INSTITUT INTERNATIONAL DE CHIMIE SOLVAY

NEUVIÈME CONSEIL DE CHIMIE

tenu à l'Université de Bruxelles

du 6 au 14 avril 1953

LES PROTÉINES

RAPPORTS ET DISCUSSIONS

Publiés par les Secrétaires du Conseil

sous les auspices du Comité Scientifique de l'Institut

R. STOOPS

Editeur

76-78, COUDENBERG, BRUXELLES

1953

INSTITUT INTERNATIONAL DE CHIMIE SOLVAY

NEUVIÈME CONSEIL DE CHIMIE

tenu à l'Université de Bruxelles

du 6 au 14 avril 1953

LES PROTÉINES

RAPPORTS ET DISCUSSIONS

Publiés par les Secrétaires du Conseil

sous les auspices du Comité Scientifique de l'Institut

R. STOOPS

Editeur

76-78, COUDENBERG, BRUXELLES

1953

INTRODUCTION

Institut International de Chimie Solvay

EXTRAITS DES STATUTS

Article premier. — Il a été fondé, à Bruxelles, à l'initiative de M. Ernest SOLVAY et pour une période de trente années, à partir du 1er mai 1913, un *Institut International de Chimie.*

La durée avait été prorogée jusqu'en 1949. Après le décès de M. Ernest Solvay, survenu le 26 mai 1922, Mme Ernest Solvay et ses enfants ont désiré assurer l'avenir de l'Institut pour un temps plus long que celui qui avait été prévu. Dans ce but, une convention a été conclue entre les prénommés et l'Université de Bruxelles; en vertu de cette convention, l'avoir actuel de l'Institut est remis à l'Université en même temps que la somme nécessaire pour qu'à l'échéance prévue de 1949 le capital d'un million primitivement consacré par M. Ernest Solvay à l'Institut International de Chimie se trouve reconstitué.

L'Université assumera la gestion de cette somme en se conformant à toutes les dispositions des présents statuts.

Art. 2. — Le but de l'Institut est d'encourager des recherches qui soient de nature à étendre et surtout à approfondir la connaissance des phénomènes naturels à laquelle M. Ernest Solvay n'a cessé de s'intéresser.

L'Institut a principalement en vue les progrès de la Chimie, sans exclure cependant les problèmes appartenant à d'autres branches des sciences naturelles, pour autant, bien entendu, que ces problèmes se rattachent à la Chimie.

Art. 3. — L'Institut International de Chimie a son siège social à l'Université Libre de Bruxelles, qui met à la disposition de l'Institut les locaux nécessaires à la tenue des *Conseils de Chimie.*

Art. 4. — L'Institut est régi par une *Commission Administrative* comprenant *cinq* membres, belges de préférence, et par un *Comité Scientifique* international comprenant *huit* membres ordinaires

auxquels peut être ajouté un membre extraordinaire ayant les mêmes droits qu'un membre ordinaire.

. . . . . .

Art. 9. — Le Fondateur a manifesté le désir qu'avant tout, l'Institut fasse preuve dans tous ses actes d'une parfaite impartialité; qu'il encourage les recherches entreprises dans un véritable esprit scientifique, et d'autant plus que, à valeur égale, ces recherches auront un caractère plus objectif. Il lui a semblé désirable que cette tendance se reflétât dans la composition du *Comité Scientifique*. Par conséquent s'il y avait des savants qui, sans occuper une haute position officielle, pourraient être considérés, en raison de leur talent, comme de dignes représentants de la Science, ils ne devront pas être oubliés par ceux qui désigneront les candidats aux places vacantes.

COMPOSITION DE LA COMMISSION ADMINISTRATIVE

(En date du 6 avril 1953)

M. J. BORDET, Professeur honoraire et membre du Conseil d'Administration de l'Université Libre de Bruxelles, *Président*.
M. P. ERCULISSE, Professeur à l'Université Libre de Bruxelles.
M. P. HEGER-GILBERT, Professeur honoraire à l'Université Libre de Bruxelles.
M. E.-J. SOLVAY, Gérant à la Société Solvay et Cie, membre du Conseil d'Administration de l'Université Libre de Bruxelles.
M. F.-H. van den DUNGEN, Professeur à l'Université Libre de Bruxelles, *Secrétaire-administrateur*.

NEUVIEME CONSEIL DE CHIMIE

(6-14 avril 1953)

LISTE DES PARTICIPANTS

A. Le Comité Scientifique.

MM. Paul KARRER, Professeur à l'Université (Zürich), Président.
H.-J. BACKER, Professeur à la Rijksuniversiteit (Groningue).
Ch. DUFRAISSE, Professeur au Collège de France (Paris).
Sir Cyril HINSHELWOOD, Professeur à l'Université (Oxford). (Excusé.)
K. LINDERSTRÖM-LANG, Directeur du Carlsberg Laboratorium (Copenhague).
P. PASCAL, Professeur honoraire à la Sorbonne (Paris).
Sir Robert ROBINSON, Professeur à l'Université (Oxford). (Excusé.)
H. WUYTS, Professeur honoraire à l'Université Libre de Bruxelles, Secrétaire honoraire.
J. TIMMERMANS, Professeur honoraire à l'Université Libre de Bruxelles, Secrétaire.

B. Les Membres rapporteurs.

MM. M. L. ANSON, Dr. Sc., Research and Development Division, Lever Brothers Company (Edgewater, U. S. A.).
Sir Lawrence BRAGG, Professeur à l'Université (Cambridge).
A. C. CHIBNALL, Professeur à l'Université (Cambridge).
V. DESREUX, Professeur à l'Université (Liège).
K. LINDERSTRÖM-LANG, Directeur du Carlsberg Laboratorium (Copenhague).
Linus PAULING, Professeur à l'Université (Pasadena, Californie).
Kai O. PEDERSEN, Professeur à l'Université (Uppsala).
R. L. M. SYNGE, Dr. Sc., The Rowett Research Institute (Bucksburn, Grande-Bretagne).
Hugo THEORELL, Professeur à l'Institut Nobel de Médecine (Stockholm).

C. Les Membres invités.

MM. Gilbert S. ADAIR, Professeur à l'Université (Cambridge).
F. FREDERICQ, Assistant à l'Université (Liège).
Cl. FROMAGEOT, Professeur à l'Université (Paris).
E. HAVINGA, Professeur à l'Université (Leiden).
J. J. HERMANS, Professeur à l'Université (Groningue).
A. J. P. MARTIN, Dr. Sc., National Institute for Medical Research (Londres).
A. NEUBERGER, Dr. Sc., National Institute for Medical Research (Londres).
P. PUTZEYS, Professeur à l'Université (Louvain).
R. SIGNER, Professeur à l'Université (Berne).
Arne TISELIUS, Professeur à l'Université (Uppsala).

D. Les Membres Secrétaires.

MM. J. TIMMERMANS, Professeur Honoraire à l'Université Libre de Bruxelles, Secrétaire du Comité Scientifique de l'Institut.
E. J. BIGWOOD, Professeur à l'Université Libre de Bruxelles.
Mlle L. de BROUCKERE, Professeur à l'Université Libre de Bruxelles.
MM. H. CHANTRENNE, Chargé de cours.
J.-P. DUSTIN, Chef de travaux.
J. LEONIS, Chef de travaux.

E. Auditeurs invités.

MM. J. BRACHET, R. DEFAY, R. DESCAMPS, P. ERCULISSE, L. FLAMACHE, P. GOLDFINGER, J. GUILLISSEN, R. MARTIN, I. PRIGOGINE, Professeurs à l'Université Libre de Bruxelles.

Discours d'Ouverture de M. le Professeur P. Karrer, Président du Neuvième Conseil

Mesdames, Messieurs,

Invités par l'Institut de Chimie Solvay, nous sommes rassemblés ici au Neuvième Conseil de Chimie Solvay afin de traiter un problème scientifique bien déterminé. Ceux d'entre nous qui ont déjà participé à un ou plusieurs Conseils de Chimie Solvay connaissent l'atmosphère tout à la fois scientifique et amicale qui entoure nos réunions. Pendant une semaine les participants ont l'impression d'appartenir à une grande famille où ils discutent des travaux du Congrès mais où ils peuvent évoquer aussi des problèmes personnels qui les préoccupent. Je suis certain que les collègues qui sont ici aujourd'hui pour la première fois subiront très vite le charme de cette ambiance due, ne l'oublions pas, à la généreuse hospitalité de nos confrères belges. Ces derniers savent rendre agréable notre séjour à Bruxelles et nous laissent autant de liberté que nous ne sentons nulle contrainte ni dans le travail, ni dans les obligations sociales. Je me fais votre interprète en remerciant vivement nos amis belges de la peine qu'ils se sont donnée pour préparer et organiser ce Congrès et pour leur hospitalité. Mes remerciements vont aussi aux membres de la commission administrative et tout particulièrement à nos Collègues MM. les Professeurs van den Dungen et Timmermans, secrétaire dévoué du Comité Scientifique du Neuvième Conseil de Chimie Solvay.

Le sujet de ce 9^{me} Conseil de Chimie Solvay est vaste et très important. L'étude des protéines fait appel à de nombreuses disciplines scientifiques : physique, chimie physique, chimie organique, chimie analytique et chimie biologique. C'est-à-dire que les protéines sont étudiées avec des méthodes très diverses et à des points de vue fort différents; nous ne serons donc pas surpris d'apprendre que nos participants relèvent eux aussi de disciplines très diverses et

considèrent le problème des protéines à des points de vue fort différents. La discussion n'en sera peut-être pas facilitée mais y gagnera certainement en intérêt. La place occupée par les protéines dans la pensée des biologistes, des biochimistes, des chimistes et des cristallographes eux-mêmes est si grande qu'elle préoccupera encore de nombreuses générations de chercheurs. Nous n'avons donc pas la prétention de nous réunir ici pour résoudre le problème des protéines mais seulement pour tâcher d'en éclaircir certains points particuliers. Selon Goethe, gagner une clairière sur la forêt de l'inconnu est toujours exaltant et l'harmonie de la nature console du désordre humain.

Je viens de dire que notre assemblée est hétérogène au point de vue scientifique; je voudrais signaler qu'elle l'est aussi au point de vue nationalité : on y rencontre les représentants de huit pays parlant six langues différentes. Il y a lieu de s'en réjouir et notre reconnaissance va aux institutions libérales de nos pays respectifs qui permettent et facilitent de telles réunions internationales. Le caractère supranational et la liberté de la science se doivent d'être ainsi soulignés car tout le monde ne les admet pas et cependant sans ces attributs la science ne peut que végéter et s'étioler pour aboutir à un instrument de dégradation humaine.

Je vous demande donc d'aborder avec confiance les débats de notre Neuvième Conseil de Chimie Solvay et de les conduire au même succès qui a caractérisé les précédents. Je remercie au nom du Comité Scientifique et au nom des invités les Collègues qui ont eu l'amabilité de se charger d'un exposé et qui dans leurs conférences poseront les fondements des discussions ultérieures.

ACTIVITES DU NEUVIEME CONSEIL

Les travaux du Neuvième Conseil de Chimie Solvay se sont ouverts le mercredi 8 avril 1953, à 10 heures, dans les locaux de l'Université Libre de Bruxelles. Le Président du Conseil, le Professeur P. Karrer a prononcé l'allocution liminaire. Le même jour, à 17 heures, les participants ont été reçus par le Recteur de l'Université qui leur a offert un thé des plus animés.

L'exposé des rapports et leur discussion ont été suspendus le samedi 11, pour permettre aux membres du Conseil de prendre part à la réunion annuelle de la Société Belge de Biochimie, à laquelle assistait aussi bon nombre de collègues Britanniques. Le dimanche, sous l'égide de cette Société, une intéressante excursion a permis aux participants de visiter les trésors artistiques de la Ville de Bruges.

Les travaux du Conseil se sont terminés le lundi 13 avril, vers 17 heures. Le même soir, le banquet offert suivant la tradition par la famille Solvay et par la Commission Administrative de l'Institut, a réuni les membres du Conseil et les autorités universitaires. M. E.-J. Solvay, au nom de la Commission Administrative, a souligné l'intérêt tout particulier des travaux du Neuvième Conseil et le Président P. Karrer en a remercié les organisateurs et les participants.

MM^{mes} Timmermans et Erculisse ont très aimablement guidé les dames qui accompagnaient les membres du Conseil, dans leurs visites de la Ville de Bruxelles et de ses environs.

Discours de Clôture de M. Karrer

Monsieur Solvay,
Monsieur le Président,
Monsieur le Recteur,
Mesdames et Messieurs,

Voici arrivée avec la fin du Neuvième Conseil de Chimie Solvay l'heure des remerciements. Lorsque l'Institut International de Chimie Solvay fut fondé en 1912, il le fut dans un esprit libéral et universel. L'article 9 des statuts stipule en particulier : « Le Fondateur a manifesté le désir qu'avant tout, l'Institut fasse preuve dans tous ses actes d'une parfaite impartialité; qu'il encourage les recherches entreprises dans un véritable esprit scientifique et d'autant plus que, à valeur égale, ces recherches auront un caractère plus objectif. » Quelques années plus tard, la famille Solvay, puis enfin le gouvernement belge lui-même ont mis de nouveaux moyens à sa disposition pour lui permettre de poursuivre sa noble tâche pendant fort longtemps encore.

Lorsque j'exprime à la famille Solvay, au gouvernement belge et à l'Université libre de Bruxelles mes plus vifs remerciements pour le soutient accordé à l'Institut International de Chimie Solvay et tout particulièrement au Neuvième Conseil de Chimie, je le fais non seulement au nom des collègues qui m'entourent, mais au nom de tous les chimistes sans distinction de nationalité. L'organisation d'un tel congrès demande un travail considérable de la part de la Commission Administrative et du Secrétaire du Comité Scientifique, aussi nous ne ménagerons pas nos remerciements ni à M. le Professeur van den Dungen, secrétaire administrateur, et M. le Professeur Timmermans, secrétaire du Comité Scientifique, ni aux Membres secrétaires : M. le Professeur Bigwood, Mlle de Brouckère et MM. Chantrenne, Dustin, Léonis et Crockaert.

L'organisation de notre Neuvième Conseil de Chimie fut une fois de plus un modèle de précision et de perfection qu'on ne saurait

dépasser et qui pourrait servir d'exemple à nombre d'autres Congrès internationaux. L'hospitalité que nous avons reçue de la part de nos collègues et amis belges était si cordiale et si naturelle que dès le premier jour nous nous sommes sentis parfaitement à notre aise.

Ce soir se trouvent réunis avec nous de nombreuses personnes qui n'ont pas participé à nos discussions scientifiques. A celles-ci je voudrais dire que les travaux de notre Conseil furent consacrés aux protéines. La notion de protéine évoque généralement chez le profane quelque chose d'agréable : le fromage, la viande, le lait, les œufs ou d'autres friandises telles que celles qui viennent de nous être servies au dîner, ce soir. Pendant la guerre on a remarqué qu'un manque de protéine se traduisait rapidement par une perte de mémoire ce qui ne veut pas dire, naturellement, que chaque oubli doit être attribué à un manque de protéine et excusé pour cela. Toutefois, il est certain que les aliments doivent nous fournir les protéines ou plus exactement certaines protéines qui nous sont absolument indispensables. Pendant huit séances nous avons examiné ces protéines sous tous leurs aspects; nous nous sommes demandés comment étaient ordonnés leurs atomes : en hélice ? suivant de grandes hélices ? ou de petites hélices ? nous avons discuté leurs poids moléculaires, leur mode de synthèse et de dégradation chez les êtres vivants; les opinions exprimées furent naturellement contraires, mais les discussions terminées, nous nous sommes trouvés réunis une fois de plus autour d'une bonne table pour nous adonner au plaisir d'un repas admirablement préparé dans lequel les protéines ont joué le rôle principal sous forme de poisson, de viande et de fromage; il était visible que pour un certain temps les hélices et le métabolisme des protéines avaient été complètement oubliés.

Des congrès comme celui que nous tenons tous les trois ans à Bruxelles ont une importance débordant le cadre étroit des sujets qu'ils traitent. Goethe a dit une fois : « L'Histoire des sciences est une grande fugue musicale dans laquelle la voix de chaque peuple se fait entendre à tour de rôle. » Cette image est très juste. Lorsqu'aujourd'hui en Europe un scientifique élève la voix pour exprimer une idée nouvelle, un collègue américain lui répond le lendemain, le surlendemain un chercheur australien entre dans le chœur et bientôt un concert retentit, qui s'étend au monde entier. Qu'y a-t-il de plus profitable que ces discussions de pays à pays, de continents à continents, discussions qui n'ont pas pour but l'oppression des

hommes, la recherche de la force, mais instaurées pour mieux connaître et comprendre la structure de l'univers, les miracles et les secrets de la nature ? Il faut souhaiter que cet exemple de la science, partagé par les arts, ait sa réplique dans d'autres domaines de la vie humaine. La science a été comparée à un levier avec lequel on peut soulever le monde, mais n'oublions pas qu'elle n'est pas le seul levier et que des mots, des idéologies et des religions ont autant ébranlé et transformé le monde que les découvertes scientifiques. Il ne faut pas que nous, scientifiques, nous imaginions être les seuls à détenir toute la vérité; nous devons avoir la modestie de ne pas oublier que la science est seulement un des facteurs qui donne sa forme à la pensée humaine.

Je termine en renouvelant mes plus sincères remerciements au nom du Neuvième Conseil de Chimie et je lève mon verre à la prospérité du Royaume de Belgique, de l'Université libre de Bruxelles et de la famille Solvay.

RAPPORTS ET DISCUSSIONS

The Molecular Weights of the Proteins

by Kai O. Pedersen

Almost all our knowledge of the molecular weights of the proteins has been gained through investigations carried out during the last two or three decades. Before that time we knew very little about the size and shape of the soluble protein molecules. It was at that time generally assumed that proteins were colloids with greatly varying particle sizes. Some attempts had been made to get an idea of the sizes of the protein particles in solutions, but the results obtained were not convincing. The most reliable experiments were those based on osmotic pressure measurements. They showed that eggalbumin had a molecular weight, M, of about 34 000 (S.P.L. Sörensen 1916) (1) and that different haemoglobins had an average M in the neighbourhood of 66 700 (G.S. Adair 1924) (2). The results made clear that these proteins formed very large particles in solution, but these investigations could not demonstrate whether the particles were all of the same size.

Simultaneously with Adair's work on haemoglobin Svedberg was constructing his ultracentrifuge, which soon became such a powerful instrument in the study of the molecular weight of the proteins. From the first investigations on proteins in the ultracentrifuge Svedberg inferred that the soluble proteins were monodisperse, i.e. had a definite molecular weight. Some years later when more results had been gained Svedberg put forward the hypothesis that the molecular weights of the proteins are distributed on a limited number of weight classes. The molecular weights of most of the higher classes could be expressed as simple multiples of those for the lower classes. In the beginning a unit of about 35 000 was assumed, but later several proteins were found with M about 17 000. Tentatively 17 600 was taken as the unit and the other multiples became : 2, 4, 8, 16, 24, 96 (168), 192, 384 and 576 times 17 600.

A large part of the experimental basis for Svedberg's multiple hypothesis for the molecular weights of the proteins has been gained through investigations on the sedimentation behaviour of coloured proteins, especially of the respiratory proteins. These substances offered the great advantage that their sedimentation could be followed by means of the light absorption method irrespective of the presence of other noncoloured proteins. Since the specific absorption of the coloured proteins was generally quite high, it was not considered necessary, except in a few cases, to isolate and purify the respiratory proteins before the investigation but the measurement of the sedimentation was in most cases carried out directly on samples of highly diluted blood from the animal in question. This made it possible to survey a large number of respiratory proteins belonging to different classes of the animal kingdom [for details see (3)]. From this investigation it soon became evident that the respiratory proteins from species belonging to the same class of animals generally had about the same sedimentation constants, s_{20}. It sometimes happened that two different values for s_{20} were found within the same class of animals. In such cases the blood from some of the species within the same class often contained two components each representing one of the groups of s_{20} found within the class.

One of the questions was then whether all these proteins that showed the same s_{20} were also alike in other respects or showed different properties. One of the ways this could be investigated was by means of electrophoresis. By using the light absorption method for the registration in the experiments, it was also possible here to do the measurements directly on the dialyzed and diluted blood. It was hereby found that even closely related species showed distinctly different electrophoretical behaviour. Thus they had different isoelectric points and different slope on the mobility *versus* pH curve (4). Later investigations on the pH-stability of haemocyanins showed great differences between proteins from closely related species (5). Similarily some haemocyanins dissociated in the presence of certain neutral salts whereas others belonging to the same class were not at all affected (6).

The conclusion to be drawn from all these results seems to be that the respiratory proteins, which show similar values for s_{20}, must be different chemical substances. On the other hand it seems likely to assume that the similarity in s_{20} is due to alike size and

shape of these molecules. Furthermore, as they have the same function in the various species they are probably synthesized in the same manner and have very like structures, except for minor details.

It is difficult to say just how similar the molecular weights are. For the proteins with s_{20} higher than 15 S (1 $S = 1$ *Svedberg* unit $= 1 \cdot 10^{-13}$ c.g.s. units) Svedberg used a special mixture test in order to determine whether two closely situated s_{20} were different or were identical within the limit of experimental error. In this test, solutions of the two proteins to be compared were mixed in such proportion that each component was responsible for half the total light absorption of the mixture. If the sedimentation constants for the two proteins differed by more than 2 per cent from one another this could easily be detected on the sedimentation diagrams. For the most high molecular proteins even smaller differences in s could be discovered by this test. For spherical or nearly spherical molecules a difference of 2 per cent in the sedimentation constant corresponds to a difference of about 3 per cent in M. For non-spherical molecules a difference of 2 per cent in s_{20} may mean more than 3 per cent difference in M. As the shape in gross of the majority of the soluble proteins does not deviate too much from a sphere or cube, we may assume that the spread in M within a certain group showing the same s_{20} would be less than 5 per cent. It seems very likely that the actual spread in these M values is less than 5 per cent, but to prove this statement would mean extensive experimental investigations on each of several proteins within at least one of the s_{20} groups. Such a comprehensive study has never been made, but within each group a few of the more easily accessible proteins were selected for more detailed studies. Before these were carried out the proteins were generally subjected to some kind of purification. In some cases the molecular weight of the protein was determined as well from sedimentation equilibrium measurements as from studies of rate of sedimentation and diffusion. In most cases, however, the last method only was used. The M found for different members of the same group generally agreed fairly well, although sometimes differences were observed which were larger than could be ascribed to differences in the amino acid composition. I think that the large deviations sometimes found must be due either to impure protein solutions being studied or to inaccuracy in the sedimentation and diffusion measurements.

(The possible errors connected with these methods will be discussed later.) I feel quite convinced that a careful redetermination of M for respiratory proteins belonging to the same group will reduce the observed spread to an amount which may be explained mainly by differences in amino acid composition or in different prosthetic groups.

The size relationship between the different groups of respiratory proteins was first based on the values found for M [7] [8] but it was later emphazised by the result from the study of the pH-stability of the haemocyanins [5] [6]. It was thus found that these molecules could be reversibly split or associated to molecules having the same M as those belonging to other groups of respiratory proteins.

Many other types of proteins have been studied in the ultracentrifuge. Often it was found that their s_{20} coincided with one of the groups of s_{20} found earlier, for instance amongst the respiratory proteins. It was then generally considered not worth while to spend time on a further determination of their M. As soon as, however, the values found for s_{20} indicated that a hitherto unknown M could be expected, a special determination of M was made. This means that when the tables of the molecular weights of the proteins were published from Uppsala, they gave a summary of all the different values of M determined for different types of proteins with special emphasis on those values that did not fit into the hypothetical multiple system. Thus all the large groups of proteins with similar M values were only represented in the tables of M by a very small fraction of their total, whereas those which did not belong to any of the large groups were *all* put into the table. Due to this selection it is quite evident that the tables give no information whatever about the frequency with which a given M occurs amongst the proteins studied in the ultracentrifuge. A statistical analysis solely based upon the values given in such tables is bound to show that no multiples exist amongst the protein.

The multiple hypothesis was also severly critized by, for instance, Bull [9], Rothen [10], Johnston *et al.* [11] and Norris [12]. Bull made a logaritmic « spectrum » of the M values published in the aforementioned tables and he found that « there are no molecular weight classes. There is an apparent tendency for certain molecular weights to cluster around 17 000 but no one knows how many « proteins » have smaller weights than 17 000; the clustering may simply indicate that there are a large number of small molecular

weight proteins in nature ». Bull proposes a statistical analysis of the distribution of M in order to see whether or not the distribution departs significantly from a random one. Such analyses have been made as well by Johnston *et al.* as by Norris. On the basis of the values given in the tables of M they find that the multiple hypothesis is disproved by the statistical analyses.

Before this criticism was published a certain scepticism was already present in Uppsala against the general validity of the multiple hypothesis. The spread in the values for M was too great and more proteins were found that did not fit into the multiples. Furthermore, in the spring of 1940 I had found that Theorell's purest cytochrome C had a M of about 12 000, as determined from sedimentation equilibrium as well as from the rate of sedimentation and diffusion. Some years earlier a less pure sample had given $M = 15\,600$, even that a rather low value for the multiple unit. At about the same time Rothen (13) had found that ribonuclease had a M of about 13 000, a value which shortly afterwards was confirmed in Uppsala. There were thus at least two proteins with M considerably lower than the multiple unit of 17 600.

There is no doubt that the hypothesis of the multiples is not of such a general validity as first thought to be. On the other hand it seems to me that there is much evidence aginst the statement that the M values are completely randomly distributed. We have thus large numbers of proteins which serve the same purpose in different organisms but are chemically different and still have about the same value for M, or at least for their s_{20}. As examples may be mentioned several of the plasma proteins, a number of the seed proteins (14) and various groups of respiratory proteins (3) (5) (8). As mentioned earlier (page 22) a certain size relationship exists between several groups of respiratory proteins, and from the results of the pH-stability investigations (5) (6) it seems appropriate to assume that their molecular weights fit into a few multiple systems with different basic units. Whether this is true or not can hardly be answered at present. Most of the experimental work on the respiratory proteins was carried out 15-20 years ago and especially the values for M are in need of revision as seen from the more recent work of Brohult (6).

Until now we have been concerned mainly with the similarity in gross in M between various proteins of different origin. How is the situation if we go to the other side and examine individual

proteins produced in a single animal species. In this case differences have been observed from time to time between various preparations of apparently well-defined proteins. In many cases the variations have been so small that experimental inaccuracy could not be excluded. In other cases differences were observed at the same time in several properties. It was thus found that fetuin, a globulin present as well in serum from cow's fetus as in serum from newly born calves, may differ quite considerably ([15]) ([16]). The different preparations all showed about the same M, but s_{20} varied between 3.09 S and 3.29 S, D_{20} between $5.0 \cdot 10^{-7}$ and $5.5 \cdot 10^{-7}$, V_{20} between 0.692 and 0.714, the nitrogen content between 12.3 and 13.4 g per 100 g protein, and finally the phosphorus content between 188 and 101 mg per 100 g protein. In this case the variation in the protein may perhaps be due to a difference in the composition of the fetuin present in the early fetus and in the fully developed fetus or the newly born calf; just as the very pronounced difference found between fetal and adult haemoglobin.

Some times it has been disclosed that an apparently uniform protein consisted of two types of molecules with different properties. Thus Brohult ([6]) has shown that the haemocyanin from *Helix pomatia* in most respects behave like an uniform protein, but in certain salt solutions near the isoelectric point of the protein 75 per cent of the haemocyanin will dissociate to half molecules, whereas the remaining 25 per cent cannot be brought to dissociate by change in the salt concentration. The half molecules can be separated in the centrifuge and by decrease in the salt concentration the half molecules may unite to whole molecules. These new molecules will in the proper salt solution show 100 per cent dissociation to half molecules. Among the other haemocyanins Brohult has shown that some are apparently uniform proteins, they show either no dissociation to smaller molecules in for instance sodium chloride solutions or they show 100 per cent dissociation, whereas other haemocyanins under similar conditions dissociate to 30 or 75 per cent. These last proteins must consist of at least two kinds of molecules although they in many respects behave like pure proteins. We may expect that in the future more of the proteins we now consider pure will turn up to consist of more than one protein. As a further example I may mention β-lactoglobulin. For many years this was considered a prototype of a homogeneous well-defined protein. It was therefore quite

unexpected when Li [17] found that crystalline β-lactoglobulin consisted of two or three major components. Later on McMeeking *et al.* [18] by fractionated crystallization succedeed in obtaining fractions with marked difference in solubilities.

When the molecular weights of the proteins are discussed the question arises whether a given protein, say human serum albumin, is always exactly the same substance, with exactly the same chemical composition and M, or one may expect a certain minor variation in the composition and in M. The question is on one side intimately connected with the discussion of the purity of the proteins and on the other side with the reproducibility of the synthesis of the proteins. The separation of a « pure » protein from the natural mixture of proteins where it is generally found is often quite complicated and if the proteins in the mixture are very similar it may be difficult to isolate the « pure » protein free of any of the other proteins from the mixture. If different samples have been purified to a different degree one cannot expect that the values found for M should be exactly the same for different samples. Even if one has succeeded in isolating a single protein out of the mixture it may not be pure in other respects. Many proteins, for instance serum albumins, bind a number of low molecular weight substances and they are often charged with these when they are in their natural environment. During the separation process part of these low molecular weight substances are removed, but some usually remain if no special precautions are taken. Different protein samples may be different in this respect. Finally during the separation procedure the conditions may have been such that we have produced small irreversible changes in some of the protein molecules, so that they are slightly different from the genuine protein molecule present in the original mixture.

As to the reproducibility of the synthesis it might perhaps be so that some of the proteins are synthesized slightly different in some of the individuals within the same species or in the same individual at different times depending upon the availability of the various less important aminoacids. If there is such a variation in the composition of the individual proteins, we may expect that the properties of proteins, obtained from pooled material, will be statistically distributed around a certain average value. For several proteins, such a spread in the values has been observed by Alberty *et al.* [19]. If, however, the distribution is gaussian and

narrow, I think we may be content and for most purposes consider our protein to be « pure ». On the other hand we cannot expect that the average values obtained on different preparations will always be the same.

Before reporting about the actual molecular weight determinations on the proteins I would like to make a few comments on some of the methods used for these determinations.

Most of the methods applied for the determination of M are based on some kind of physico-chemical measurements on proteins in solution, preferable at a pH close to their isoelectric point. The values obtained in this way may differ significantly from those used by the organic chemists. According to their definition a molecule is the smallest particle showing *all the chemical properties* of the material. *In this report we will regard the smallest particle showing all the physical, chemical and biological properties of the material as a molecule.* In some cases this may mean that our molecule from the organic chemist point of view is a polymer. It means that the weight of a single unhydrated molecule of the protein should be M/N gram, where N is Avogadro's number.

METHODS

Osmotic pressure.

By using this method the molecular weight of a protein was determined for the first time more than 35 years ago by S.P.L. Sörensen ([1]). Since then osmotic measurements have been widely used in the study of the soluble proteins. In order to get the molecular weight, osmotic pressure measurements must be carried out on solutions over a range of protein concentration and Π/c must be plotted against c, where Π is the osmotic pressure and c the protein concentration.

From the extrapolated value for Π/c at $c = 0$ the molecular weight for the protein can be calculated. For non-electrolytes it has been shown that Π for dilute solutions will vary with the solute concentration in the following way

$$\Pi = \frac{RTc}{M} + Bc^2 \qquad (1)$$

where the coefficient B is a measure of the deviation from ideality. As, however, proteins are colloidal electrolytes, the above expression is only valid under certain conditions. It has been shown [20] that when a « swamping excess » of a simple electrolyte is used in the osmotic pressure measurements, the colloidal ions behave in the same way as if they were non-electrolytes present in the same medium. Equation (1) is therefore also valid in this case. In more dilute salt solutions valuable informations about the interaction between proteins and various ions may be gained [21]. [For reference to aqueous solutions see also [22]].

The osmotic method does not give any information about the purity of the macromolecular substance investigated. For a mixture of macromolecules it gives the number-average molecular weight. Association or dissociation will show up as a drift in the values, calculated for M, whith change in concentration.

The ultracentrifuge.

This instrument has been used to a great extent for the determination of M for proteins. Two different types of measurements may be used for this purpose :

1) sedimentation equilibrium,
2) rate of sedimentation.

The first method is a thermodynamical method like the osmotic one. For a dilute ideal solution the following expression is valid

$$M = \frac{2\,RT\,ln\,c_2/c_1}{(1 - V\rho)\,\omega^2(x_2^2 - x_1^2)} = \frac{RT\,dc/dx}{(1 - V\rho)\,\omega^2 cx} \qquad (2)$$

where R is the gas constant, T the absolute temperature, c_1 and c_2 are the concentrations at the distances x_1 and x_2 from the axis of rotation, respectively; V is the partial specific volume of the macromolecular solute, ρ the density of the solution and ω the angular velocity of the centrifuge.

A number of different average molecular weights may be calculated from sedimentation equilibrium experiments [23]. For the weight-average molecular weight we have for the ideal case

$$M_w^{ideal} = \frac{2\,RT\,(c_b - c_a)}{(1 - V\rho)\,\omega^2 C\,(b^2 - a^2)} \qquad (3)$$

where C is the original concentration of the macromolecular solute, a and b are the distances from the axis of rotation to the meniscus of the solution and to the bottom of the cell, respectively; c_a and c_b are the corresponding concentrations at equilibrium. For the non-ideal case Wales (24) has shown the following equation to be valid

$$\frac{1}{M_w^{ideal}} = \frac{1}{M_w} + BC \tag{4}$$

where M_w will be the true weight-average molecular weight and B is the same non-ideality coefficient as in equation (1). By plotting $1/M_w^{ideal}$ as calculated from equation (4) against C the slope of the line is equal to B. For $C = 0$ we have $M_w^{ideal} = M_w$.

Until recently equations (2) and (3) have been used for the calculation of M from sedimentation equilibrium measurements. Due to the non-ideality of the protein solutions such M values must therefore be too low. As, however, the protein concentration has usually been about 1 per cent or less the error committed in these experiments is generally small.

From sedimentation equilibrium measurements some knowledge may be gained about the purity of the protein solution, but the method is not very sensitive. Mixture of molecules having different M will give rise to an increase in the calculated M [equation (2)] with increase in the distance from the axis of rotation. If the mixture is due to some association-dissociation reaction the distribution of M through the cell must vary with change of the macromolecular concentration, but it should be independent of the speed of the centrifuge. If the mixture consists of independent components, the distribution of M should be the same for different concentrations (at least in dilute solutions), but it should be strongly dependent upon the speed of the centrifuge. In the case of high values for B or if V varies with the degree of association the above effects may be at least partly masked.

In sedimentation velocity experiments the independent components may often give rise to individual peaks in the sedimentation diagrams. An apparently normal and completely symmetrical peak on the diagram is not, however, a criterion of monodispersity, but may be due to a rapid association-dissociation reaction. If so, a study of the variation of s_{20} with the macromolecular concentration will often reveal such a reaction. For a purely monodisperse substance

the sedimentation coefficient decreases with increase in concentration, whereas an increase in s_{20} with increase in concentration shows the presence of an association reaction.

Most of the ultracentrifugally determined M for proteins have been calculated from the Svedberg formula

$$M = \frac{RTs}{D(1 - V\rho)} \tag{5}$$

where D is the diffusion coefficient. As a rule s^0_{20}, the sedimentation coefficient extrapolated to zero protein concentration, has been used for the calculation of M and where possible the diffusion coefficient has been extrapolated to $c = 0$. Several problems connected with the use of the ultracentrifuge and especially equation (5) for molecular weight determinations were discussed at a conference on the ultracentrifuge, sponsored by the National Academy of Sciences, Washington (25).

As equation (5) has been so widely used in calculating M it seems appropriate to discuss the different quantities in this formula. The gas constant $R = 8.3136 \cdot 10^7$ erg. degree^{-1} mole^{-1} is taken for granted. The temperature plays a very important role for these measurements. Thus for aqueous solutions at room temperature a change of 1° C alters the values for s and D by about 2 ½ per cent. Generally we know the temperature of the diffusion experiment with great accuracy and there is no problem in its determination. As regards the temperature of the sedimentation experiments the situation is different. For the moment I think that we do not know the exact temperature of the solution in the rotating cell in any of the fast running ultracentrifuges. It seems as if at present it is necessary to allow for an uncertainty in the determination of the cell temperature of at least $\pm$ 0.5° C at a speed of 60 000 r.p.m., but correspondingly less at lower speeds.

The whole question of the cell temperature is being investigated in many laboratories; it came up some years ago when it was found that sedimentation constants obtained in different laboratories for various proteins did not agree with those earlier found in Uppsala (26). When the Spinco ultracentrifuges came into general use, the values found with these centrifuges were in many cases 10 per cent lower than those found in Uppsala (27 - 30). In order to find the reason for this discrepancy a study of the temperature measurements in the oil-turbine ultracentrifuges has been going on

in Uppsala for the last couple of years. As part of this investigation a large number of melting point curves has been determined in runs at various speeds and in the different ultracentrifuges in Uppsala.

According to earlier investigations, it was assumed that a thermocouple placed very close to the rotor would show the cell temperature with an accuracy of about 1/2° C. It was soon found that this was not true, but the temperature of the cell would be about 3° higher than that of the thermocouple, when the centrifuge was running under standard conditions at 60 000 r.p.m. The temperature difference between the cell and the thermocouple was found to be approximately proportional to the square of the rotor speed. This error in the cell temperature in Uppsala accounts for the greater part of the discrepancy between the results obtained with the Spinco centrifuges and those in Uppsala, but it does not seem possible to explain all the differences.

There is, however, also some uncertainty about the cell temperature in the Spinco ultracentrifuges. Usually the cell temperature is interpolated from measurements of the rotor temperature taken immediately before the start of the centrifuge and immediately after the rotor has been brought to a stand still. It has recently been claimed, however, that « on acceleration the rotor cools, the opposite effect occurring during deceleration. Such changes are shown to be due to the adiabatic changes concomitant with the production or release of stress in the rotor. At 60 000 r.p.m. a change of — 0.9° C is observed on acceleration » (31). According to the same authors it is possible to measure the temperature of the moving rotor by making use of its infra-red radiation. The accuracy of the method is said to be 0.1° C measured with a precision of $\pm$ 0.03° C. If these findings of Waugh *et al.* are confirmed it is probable that the discrepancies between the Spinco and Uppsala values will vanish.

As the determination of s and D are usually not carried out at the same temperature, the experimental values must be reduced to the values corresponding to some standard temperature. At the same time it is convenient to refer them to some standard medium. For proteins water of 20° C is generally chosen as the standard condition. The sedimentation coefficient is thus calculated in the following way

$$s_{20} = \frac{dx/dt\ \eta_\theta}{x\omega^2\ \eta^0_{i20}} \cdot \frac{1 - V_{20}\,\rho^0_{20}}{1 - V_\theta\,\rho_\theta} \qquad (6)$$

where dx/dt is the sedimentation velocity at the distance x from the axis of rotation, $\eta_{i\theta}$ and η^0_{i20} the viscosities of the medium at θ^o and of water at 20° C respectively, and ρ_θ and ρ^0_{20} the corresponding densities.

The determination of x and dx/dt from the sedimentation diagrams is usually made sufficiently accurate by means of the Lamm scale method ([23]) or by the Fresnel diffraction method as developed by Kegeles *et al.* ([27]). The other automatic recording systems are less satisfactory for precision determination of the movement of the sedimenting boundary. As mentioned above exact knowledge of the cell temperature is necessary for the two last terms in equation (6), especially for $(\eta_{i\theta}/\eta^0_{i20})$.

For the proteins the term $(1 - V\rho)$ is of the order of 0.25 — 0.30; it means that the partial specific volume must be accurately known, and an error of 1 per cent in V causes an error of about 3 per cent in M. The introduction by MacInnes *et al.* of the magnetic float method for measuring densities of protein solutions means a great improvement in the accuracy of determining V. So far, however, it has only been applied on three proteins at 25° C ([32]). It would be highly desirable to get measurements made on more proteins and at a number of different temperatures. This would make it possible to estimate the importance of the variation of V with temperature in the last term in equation (6).

In some cases the hydration of the proteins may cause some trouble in the determination of s^0_{20}, viz. when the protein is soluble only in strong salt solutions and the density and composition of the hydration layer differ from that of the surrounding medium. Usually the salt concentration can be kept at such a level that the error committed in neglecting the hydration is small, but in extreme cases it may be appreciable ([33]) ([34]).

As to the diffusion coefficient, D, the last factor from equation (5) to be discussed, there is no problem in the temperature measurements during the experiment, but sometimes perhaps in the minor temperature fluctuations occurring during the run. For the optical observations of the diffusion process we have now a number of fine methods. Besides the old Lamm scale method several interferometric procedures have been published after the last war ([35] - [38]).

So far the new methods have mainly been tested with low molecular weight substances diffusing very fast. Very few experiments on proteins have been published, and it seems as if here some difficulties exist. The main trouble is probably connected with the construction of the diffusion cell and for the moment hardly any entirely satisfactory cell is to be found. As, however, many different laboratories are working on the problem we may hope to get satisfactory diffusion cells in a not too distant future.

The difficulties met in the individual sedimentation and diffusion experiments are either eliminated or at least much reduced in sedimentation equilibrium measurements. The main drawback with this method is that it usually takes several days, sometimes weeks before the equilibrium is established. It has been shown, however, by Archibald [39] that if just the molecular weight is wanted, it is not necessary to wait for the equilibrium to be established, but reliable values for M could be obtained much earlier. So far only a single paper has been published [40] where this method has been used. For the moment Archibald's method is being studied in details in Uppsala. Experiments with peptides having $M \sim 1\,000$ have been carried through until sedimentation equilibrium has been established. The values calculated for M from equations (2) or (3) have agreed completely with those determined by the Archibald procedure. This method is based on the following : According to the differential equation for the combined sedimentation and diffusion in a sectorial ultracentrifuge cell the flux, Φ, of material in 1 second through a surface at the distance x from the axis of rotation having unit area is given in the equation :

$$\Phi = \omega^2 xcs - D \frac{\delta c}{\delta x} \tag{7}$$

The first term on the right of the equation provides for the transport of material by sedimentation and the second by diffusion. At equilibrium this flux becomes zero for all values of x. Since, however, no solute can pass through the meniscus of the solution, $x = a$, nor through the bottom of the cell, $x = b$, the flux is also zero at these two levels at all times, i.e. even prior to the attainment of equilibrium. Thus a plot of $(\delta c/\delta x)/\omega^2 xc$ against x should extrapolate to the same value, s/D, at $x = a$ and $x = b$ at any time for a monodisperse substance. The molecular weight is then obtained from equation (5) by inserting the value for s/D. For a polydisperse

substance the M thus obtained corresponds to the weight-average M at $x = a$ and $x = b$, respectively, at the time of the measurement. As the experiment proceeds the value by the Archibald method at $x = a$ should decrease, whereas increasing values by this method should be found for $x = b$. The variation found in M with time thus gives a measure of the homogeneity of the preparation. If the extrapolated values for M corresponding to $x = a$ and $x = b$, respectively, are plotted against the time after the centrifuge has reached full speed, two curves are obtained. For $t = 0$ they will both extrapolate to the weight-average M for the substance [corresponding to the value obtained from equation (3)]. For $t \rightarrow \infty$ they will correspond to the weight-average M at equilibrium at the meniscus and at the bottom of the solution, respectively.

Light scattering.

Putzeys and Brosteaux [41] were the first to make use of Rayleigh's theory of the scattering of light for determining the size of the proteins in solution. Since then several papers dealing with light-scattering of protein solutions have been published from Putzeys' laboratory. An important contribution to the theory of light-scattering was made when Debye [42] showed that there is a close relationship between the results obtained by this method and by osmotic pressure.

In the experiment the « turbidity », τ, of the solution is determined. It is defined as $\tau = (1/l)\, ln\, (I_0/I)$, where I_0 is the initial intensity of the light, I the intensity transmitted after passing the distance l through the solution. For the ideal case, application of Rayleigh's law for independent point scatterers in solution gives

$$\tau = HcM \qquad H = \frac{32\ \pi^3\ n_0^2}{3N\ \lambda^4}\left(\frac{n - n_0}{c}\right)^2 \tag{8}$$

where c is the weight concentration in g protein per ml, λ is the wave length of the incident light, N is Avogadro's number, n_0 is the refractive index of the medium, and n is the refractive index of the solution.

For non ideal dilute solutions the following expression is valid

$$\frac{Hc}{\tau} = \frac{1}{M} + 2\,Bc \tag{9}$$

where B is a measure of the deviation from ideality.

It was shown by Zimm and Doty [44] that the M obtained from light scattering experiments is the weight-average molecular weight. It means that minute quantities of particles having quite high effective weights, such as dust and other foreign, colloidal suspended matter, will seriously affect the results obtained. This seems to be the main draw-back of this method and a number of special precautions must therefore be taken in order to remove such foreign matter. It must be emphasized, however, that the addition of light scattering measurements to the methods used in the study of proteins has been extremely valuable. It is thus quite indispensable for following the kinetics of association-dissociation reaction in protein solutions [44] [45]. For a recent review on light scattering see [46].

Besides the scattering experiments with visible light some X-ray scattering studies have been made on dilute solutions of some proteins [47] [48].

Analytical methods.

Several proteins contain metals or specific groups whose amount can be accurately determined. Under the assumption that the molecule holds only one metal atom or one specific group a minimum M may be calculated. It is, of course, necessary that the metal or specific group are integral parts of the molecule and are not just adsorbed onto the molecule. In recent years several minimum molecular weights have been calculated from the amino acid compositions of the proteins under the assumption that the number of each amino acid pro molecule must be an integer.

Recently Craig *et al.* [49] have developed a very interesting method for molecular weight determination on peptides. It is based upon the striking effect produced in the partition ratio of the peptide in countercurrent distribution (CCD) when any substitution has been made which affects either a carboxyl or an amino group in the molecule. It is therefore possible to separate derivatives representing different stages of substitution which result from incomplete reaction. Since with unchanged peptide remaining, the band occurring in the CCD pattern nearest the unchanged substance would be the monosubstituted derivative, M may be calculated from the relative amount of the substituent in the derivative. The method has also been applied on insulin, where it gave $M = 6\ 500$ [50].

In addition to the methods described reference must be made to the valuable information that may be gained from the X-ray study of protein crystals. From X-ray studies the dimension and volume of the unit cell may be calculated. From the density of the crystal the mass of the unit cell is computed. The density may be accurately measured in the gradient tube (51). By dividing the mass of the unit cell with the mass of the hydrogen atom ($1.66 \cdot 10^{-24}$ g) the « molecular weight » of the unit cell content is obtained. From the X-ray studies we do not know the number of molecules, n, in the unit cell. This number must be an integer and usually a small one. Furthermore the symmetry of the space group imposes certain limitations on n. From the « molecular weight » of the unit cell and the possible values for n a small number of possible M for the protein may be calculated. Comparison of these M values with those found by other methods will nearly always indicate which M is the right one. For further discussion about determinations of M from X-ray studies see (51 - 53).

MOLECULAR WEIGHTS OF SOME PROTEINS.

The following tables give a summary of molecular weight determinations by different methods on proteins belonging to the following groups : plasma proteins, seed proteins, enzymes, and hormones.

In the tables the following abbreviations have been made :

V_{20} = partial specific volume at 20° C. Values in parenthesis are not measured, but assumed.

s_{20} = sedimentation coefficients in Svedberg units reduced to water at 20° C; where possible extrapolated to zero protein concentration. Values marked with an asterisk (*) may be 7.5% too high (see pages 29-30).

D_{20} = diffusion coefficient in units of 10^{-7} reduced to water at 20° C.

M_{eq} = M determined from sedimentation equilibrium.

M_{sD} = M determined from sedimentation and diffusion.

M_o = M determined from osmotic pressure.

M_l = M determined from light scattering.

M_x = M determined from X-ray.

M_a = M determined from chemical analysis.

$M_{s\eta}$ = M estimated from sedimentation and viscosity.
$M_{\eta D}$ = M estimated from diffusion and viscosity.
M_e = roughly estimated value for M.

Plasma proteins.

The proteins in this group have been prepared in several different ways. Some have been isolated by modifications of the classical method of salting out with ammonium sulphate. Several have been separated by the method developed at Harvard by E.J. Cohn *et al.* (54) making use of ethanol at low temperature. For the proteins prepared according to this procedure reference has been made to the fraction number (according to the nomenclature of the Harvard group) from which it has been isolated.

The Harvard method has made possible the separation of a number of new plasma proteins besides those already known. In the case where the same protein has been prepared according to as well the last mentioned method as to the more classical ones, the different samples seem to have the same general properties. There may perhaps be somewhat greater tendency to associate in the proteins prepared by ethanol fractionation than by the other methods. Thus Cann *et. al.* (55) have found definite differences in the ultracentrifugal patterns of various γ-globulins prepared according to the ethanol fractionation and by electrophoresis-convection. The latter method yielded material where 95 per cent, or generally even more, sedimented as a single component with $s_{20} = 6\ S$ at a concentration of about 1 per cent. The ethanol separated material on the other hand showed from 5 to 20 per cent material with $s_{20} = 8 - 20\ S$. For the main component was found $s_{20} = 6.5\ S$.

From light scattering experiments on human and bovine serum albumin Halwer *et al.* (56) concluded that ethanol prepared serum albumins have a great tendency to aggregate during storage in the solid condition. They studied a number of different samples of bovine serum albumin after various durations of storage. The initial M_l-values they found for bovine albumin ranged from 73 000 to 99 000. After storage the M_l-values always increased. In one sample as high a value as $M_l = 237\,000$ was found after 18 months storage. The same sample gave $M_a = 104\,000$.

Putzeys and Brosteaux (57) fractionating serum by means of ammonium sulphate had to adopt a special procedure in order

to get consistent results for the light scattering of serum albumin. Their values for M_l were based on $M = 330\,000$ for amandin.

As seen from Table I the plasma proteins have molecular weights ranging from 44 000 to about 10^6. Under normal conditions the greater part of the protein in plasma has M-values in the range 60 000 to 200 000. In certain pathological cases more than half the amount of protein may have $M \sim 10^6$.

Seed globulins.

Besides the seed proteins mentioned in Table II a great number of proteins belonging to this group has been studied in the ultracentrifuge ([75]). They have shown s_{20} values similar to those in Table II. In the case of globulins from some of the seeds belonging to the family Leguminosae a component was found having $s_{20} \sim 19$ S. For the remainder of these proteins their molecular weights are situated in the range 25 000 to about 350 000.

The M_l-values are all based upon $M = 330\,000$ for amandin. It may be that some of these proteins show association-dissociation reactions in solution. It was thus found that s_{20} for legumin is constant within the concentration range 0 to 1.15 per cent protein. For vicilin s_{20} is constant from 0.2 to 1.5 per cent protein; below 0.2 per cent there is a small increase in s_{20}. A normal protein would show a decrease in s_{20} with increasing protein concentration.

Enzymes.

The molecular weights of a great number of these proteins have recently been determined. One of the difficulties frequently encountered with these substances is that they have often been obtained in minute quantities only. This makes the determination of M much more difficult and several of the values have only been estimated. Due to their specific activity it is probably much easier with the enzymes to ascertain whether a given sample is pure or not.

The molecular weights for the enzymes are scattered within the range of 10 000 to 500 000 or perhaps even higher. In the case of the proteases, all those studied have shown M-values below 45 000, except dihydropeptidase. They have also shown a tendency to aggregate under certain conditions, a property that may have something to do with their proteolytic activity.

Hormones.

Many of the protein hormones may be very difficult to get in quantities sufficient for molecular weight determinations. Even if they possess biological activity, it may often be difficult to determine their purity, as the biological tests are not so accurate, although they are very sensitive.

All the protein hormones have molecular weights below 100 000 and a large part of them has very low M-values. Thyroglobulin was earlier considered as a hormone, but it appears as if it is merely a carrier for the hormone. For some of the protein hormones the molecular weight is so low that they should be classified rather as peptides than as proteins. It seems likely that this question may be solved in the near future.

Several of the protein hormones show a marked tendency to aggregation. This is, for instance, the case with the pituitary growth hormone ([150]). The protein that has been studied most in this respect is, however, insulin. Very interesting investigations on the thermodynamics of this association reaction is at present going on in two groups at Harvard ([44]) ([45]) ([139]). From all the work so far done on insulin it seems established that the monomer of the insulin molecule in dilute acid solution has about $M = 12\,000$. At higher insulin concentration or on the alkaline side of the isoelectric point dimers, trimers, tetramers and under certain conditions still higher polymers exist in aqueous solutions. Attempts have been made to verify the finding of a minimum molecular weight of 6 000 by Fredericq and Neurath ([159]), but in vain ([140]). According to Craig *et al.* the « organic Chemical M » of insulin is, however, about 6 000 ([50]).

The molecular weights of many other proteins have been determined during the last twenty years, but it would carry too far to summarize them all. A study of Tables I to IV will give an indication of the range of M for the proteins belonging to the four groups : Plasma proteins, seed proteins, enzymes, and hormones.

To summarize the situation one may say : We have a very good idea of the approximate M for a large number of proteins, but we are still lacking precision determination of M. It would be very valuable to get such measurements made on a limited number of well-defined, « pure » proteins. In such an investigation all the methods applicable should be used.

TABLE I

Molecular weights of plasma proteins.

PROTEIN	V	s_{20}	D_{20}	M	ORIGIN	Literature
Mucoprotein	0.675	3.11	5.27	M_{sD} = 44 100	Human serum	(58)
Mucoprotein	—	—	—	M_a = 46 500	Human serum	(58)
Fetuin	0.692	3.1*	5.0	M_{sD} = 49 000*	Cow's foetus	(16)
Fetuin	0.714	3.3*	5.5	M_{sD} = 51 000*	Calf	(16)
Albumin, human	0.736	4.67*	5.93	M_{sD} = 72 000*	Human serum	(15)
Albumin, human	0.733	4.6	—	M_o = 69 000	Human plasma Fraction V	(59)
Albumin, human	(0.733)	4.28	6.32	M_{sD} = 61 500	Human serum	(30)
Albumin, bovine	0.734[3]	4.73*	—	—	Bovine plasma	(32) (60)
Albumin, bovine	(0.734[3])	4.31	6.15	M_{sD} = 64 500	Bovine plasma	(27)-(29) (61)
Albumin, bovine	—	—	—	M_i = 73 000 to 99 000	Bovine plasma	(62) (56)
Albumin, bovine	—	—	—	M_i = 77 000	Bovine serum	(57)
Albumin, equine	0.748	4.46*	6.1	M_{sD} = 70 000*	Horse serum	(63)
Albumin, equine	—	—	6.46	—	Horse serum	(64)
Albumin, equine	—	—	—	M_i = 74 000	Horse serum	(57)
Albumin, porcine	—	—	—	M_i = 72 000	Hog serum	(57)
Siderophilin	(0.725)	5.8*	5.82	M_{sD} = 88 000*	Hog serum	(65)
β^1-globulin	0.725	5.5	—	$M_o \sim$ 90 000	Human plasma Fraction IV-7	(59)
β^2-globulin	—	7	—	$M_e \sim$ 150 000	Human plasma Fraction III-1	(59)
Caeruloplasmin	0.73	7.5	4.7	$M_{sD} \sim$ 150 000	Human and hog serum	(66) (67)
γ-globulin	0.718	7.12*	4.0	M_{sD} = 153 000*	Human serum	(15)

TABLE I

Molecular weights of plasma proteins *(continued)*.

PROTEIN	V	s_{20}	D_{20}	M	ORIGIN	Literature
γ-globulin	(0.739)	7.3	3.7	M_e = 170 000	Human plasma Fraction II	(68)
γ-globulin	0.739	—	—	M_o = 156 000	Human plasma Fraction II	(59)
γ-globulin	—	10	—	M_e ~ 300 000	Human plasma Fraction II	(59)
Antipneumococcus serum globulin	(0.745)	6.5*	3.9	M_{sD} = 158 000*	Rabbit serum	(69)
Antibody against p-Azophenyl-arsonic acid	—	—	—	M_o = 140 000	Rabbit serum	(70)
				M_i = 158 000	Rabbit serum	(70)
Antipneumococcus serum globulin	(0.745)	7.4*	3.6	M_{sD} = 195 000*	Human serum	(69)
α^1-lipoprotein	0.841	5.0	—	$M_{s\gamma_1}$ ~ 200 000	Human plasma Fraction IV-1	(59)
α^2-globulin	0.693	9	—	M_e ~ 300 000	Human plasma Fraction IV-6	(59)
Serumglobulin	(0.745)	12*	3.2	M_{sD} = 360 000*	Serum *Lampetra fluviatilis*	(71)
Fibrinogen, human	—	9	—	$M_{s\gamma_1}$ ~ 400 000	Human plasma Fraction I-2	(59)
Fibrinogen, human	—	—	—	M_o = 580 000	Human plasma Fraction I-2	(59)
Fibrinogen, bovine	—	8.5	—	M_i = 407 000	Bovine plasma	(72) (73)
Apoferritin	0.747	17.6	3.6	M_{sD} = 465 000	Human and horse serum	(74)
Antipneumococcus serum globulin	(0.715)	18.1*	1.69	M_{sD} = 910 000*	Cow serum	(69)
Antipneumococcus serum globulin	0.715	19.3*	1.80	M_{sD} = 910 000*	Horse serum	(69)
Antipneumococcus serum globulin	(0.715)	18.0*	1.64	M_{sD} = 930 000*	Hog serum	(69)
Pathological euglobulin	0.733	19.3*	1.5	M_{sD} = 1 200 000*	Pathological human serum	(15)
β^1-lipoprotein	0.950	2.9	—	M_e = 1 300 000	Human plasma Fraction III-0	(59)

TABLE II

Molecular weights of seed globulins.

PROTEIN	V	s_{20}	D_{20}	M	ORIGIN	Literature
Gliadin	0.724	2.1	6.7	M_{SD} = 27 500	Wheat gluten	(76) (23)
Hordein	(0.729)	2.0*	6.5	M_{SD} = 27 500*	Barley	(23)
Barley Globulin α	0.72	2.5	7.4	M_{SD} = 29 000	Barley	(77) (75)
Concanavalin B	0.73	3.5*	7.4	M_{SD} = 42 000*	Jack beans	(78)
Zein	0.776	1.9	4.0	M_{SD} = 51 000	Maize	(76) (79)
Ricin	(0.75)	4.8	(6.0)	M_{SD} = 77 000	Castor beans	(80)
Concanavalin A	0.73	6.0*	5.6	M_{SD} = 96 000*	Jack beans	(78)
Canavalin	0.73	6.4*	5.1	M_{SD} = 113 000*	Jack beans	(78)
Vicilin	0.752	8.10	4.26	M_{SD} = 186 000	Seeds from pea	(75)
Barley globulin γ	0.72	8.3*	4.4	M_{SD} > 166 000*	Barley	(77) (75)
Wheat globulin γ	(0.72)	8.7*	3.6	M_{SD} = 210 000*	Wheat embryo	(75)
Globulin from *Prunus Avium*	—	—	—	M_i = 286 000 to 316 000	Seeds from sweet cherry	(81)
Globulin from *Prunus Cerasus*	—	—	—	M_i = 295 000	Seeds from cherry	(81)
Globulin from *Prunus Domestica*	—	—	—	M_i = 290 000	Seeds from plum (Reine-Claude)	(81)
Excelsin	0.743	13.3	4.26	M_{SD} = 295 000	Brazil nuts	(23)
Excelsin	—	—	—	M_i = 278 000	Brazil nuts	(57)
Edestin	0.744	12.8	3.93	M_{SD} = 310 000	Hemp seeds	(23)
Edestin	(0.744)	(12.8)	3.18	M_{SD} = 380 000	Hemp seeds	(82)
Edestin	—	—	—	M_i = 335 000	Hemp seeds	(83)
Amandin	0.746	12.5	3.62	M_{SD} = 330 000	Almonds	(23)
Amandin	(0.746)	(12.5)	3.45	M_{SD} = 346 000	Almonds	(82)
Legumin	0.735	12.64*	3.49	M_{SD} = 330 000*	Seeds from pea	(75)

TABLE III

Molecular weights of enzymes.

PROTEIN	V	s_{20}	D_{20}	M	ORIGIN	Literature
Esterases						
Pyrophosphatase	(0.75)	4.4	6.8	M_{sD} = 63 000	Yeast	(84)
Phosphatase	(0.749)	4.0	1.83	M_{sD} = 212 000	Prostate	(85)
Carbohydrases						
α-Amylase	—	—	—	M_o = 59 500	Barley malt	(86)
α and β-Amylase	0.69	4.52*	6.53	M_{sD} = 54 000*	Barley malt	(87)
α-Amylase	0.70	4.5*	8.0	M_{sD} = 45 000*	Hog pancreas	(88)
β-Amylase	(0.749)	8.9	5.77	M_{sD} = 152 000	Sweet potato	(89)
Lysozyme	—	—	—	M_x = 13 900	Chicken Eggwhite	(90)
Lysozyme	0.720	1.9	11.2	M_{sD} = 14 700	Chicken Eggwhite	(91) (92)
Lysozyme	—	—	—	M_o = 17 500	Chicken Eggwhite	(92)
Lysozyme	(0.75)	—	8.6	M_D = 12 800	Chicken Eggwhite	(93)
Lysozyme	1 P atom/mol enzyme			M_a = 11 900	Chicken Eggwhite	(93)
Lysozyme	From amino acid composition			M_a = 14 800	Chicken Eggwhite	(91) (94)
Lysozyme				M_i = 14 800	Chicken Eggwhite	(56)
Phosphorylases						
Phosphorylase b	(0.74)	13.7	3.3	M_{sD} = 385 000	Rabbit skeletal muscle	(95)
Transphosphorylases						
Hexokinase	(0.74)	6.0	5.8	M_{sD} = 97 000	*Saccharomyces cerviscia*	(96)
Nucleases						
Ribonuclease	—	—	—	M_o = 15 000	Beef pancreas	(97)
Ribonuclease	0.707	2.00	13.1	M_{sD} = 12 700	Beef pancreas	(98)
Ribonuclease	0.707	—	—	M_{eq} = 13 000	Beef pancreas	(98)
Ribonuclease	Estimated from 15 700 for the air-dried crystals			M_x = 13 700	Beef pancreas	(99)
Ribonuclease	Estimated from 15 500 for the air-dried crystals			M_x = 13 400	Beef pancreas	(100)
Amidases						
Transaminase glutamic aspartic	Estimated from sedimentation and diffusion			$M_e \sim$ 60 000	Hog heart	(101)
Transaminase alanine-glutamic	—	—	—	$M_e \sim$ 180 000	Hog heart	(101)
Urease	0.73	18.6	3.46	M_{sD} = 480 000	Jack beans	(102)
Proteases						
Carboxypeptidase	(0.75)	—	8.67	$M_{\eta D}$ = 32 000	Beef pancreas	(103)
Carboxypeptidase	(0.75)	3.07	—	M_{sD} = 34 000	Beef pancreas	(104)
Pepsinogen	—	—	—	M_o = 42 000	Hog gastric mucosa	(105) p. 81
Pepsin	—	—	—	M_o = 35 000	Hog gastric mucosa	(105) p. 75
Pepsin	—	—	—	M_o = 38 000	Hog gastric mucosa	(105) p. 81
Pepsin	(0.75)	—	—	M_{eq} = 39 000	Hog gastric mucosa	(23)
Pepsin	(0.75)	3.3	9.0	M_{sD} = 35 500	Hog gastric mucosa	(23)
Pepsin	—	—	—	M_a = 34 300	Hog gastric mucosa	(105) p. 74
Chymotrypsinogen	0.721	2.5	9.5	M_{sD} = 23 200	Beef pancreas	(106)
Chymotrypsinogen	—	—	—	$M_{s\eta}$ = 22 600	Beef pancreas	(106)
Chymotrypsinogen	—	—	—	$M_{\eta D}$ = 24 500	Beef pancreas	(106)
Chymotrypsinogen α	(0.73)	2.5	10.2	M_{sD} = 22 000	Beef pancreas	(107)
Chymotrypsin α, monomer	0.736	2.4	10.2	M_{sD} = 21 600	Beef pancreas	(108)
Chymotrypsin α, monomer	—	—	—	$M_{\eta D}$ = 20 900	Beef pancreas	(108)
Chymotrypsin α, dimer	0.731	3.52	7.4	M_{sD} = 42 900	Beef pancreas	(108)
Chymotrypsin α, dimer	—	—	—	$M_{s\eta}$ = 43 000	Beef pancreas	(108)
Chymotrypsin α, dimer	—	—	—	$M_{\eta D}$ = 42 800	Beef pancreas	(108)
Chymotrypsinogen B	(0.73)	2.5	10.4	M_{sD} = 21 600	Beef pancreas	(107)
Chymotrypsin B	(0.73)	2.6	9.9	M_{sD} = 23 600	Beef pancreas	(107)

TABLE III

Molecular weights of enzymes *(continued)*.

PROTEIN	V	s_{20}	D_{20}	M	ORIGIN	Literature
Chymotrypsin β	Combining ratio with diisopropylfluorophosphate			M_a = 24 200	Beef pancreas	(109)
Chymotrypsin γ	Combining ratio with diisopropylfluorophosphate			M_a = 23 900	Beef pancreas	(109)
Trypsin	Osmotic pressure at 5° C			M_o = 34 000	Beef pancreas	(105) p. 140
Trypsin	(0.751)	1.69	10.95	M_{SD} = 15 000	Beef pancreas	(85)
Trypsin	—	—	—	M_a = 20 700	Beef pancreas	(109)
Dihydropeptidase	—	4.9	—	M_e ~ 70 000	Beef pancreas	(110)
Iron enzymes						
Cytochrome c	0.702	1.83*	12.0	M_{SD} = 12 400*	Beef heart	(23)
Cytochrome c	—	—	—	M_a = 13 000	Beef heart	(111)
Cytochrome c	0.460 per cent Fe			M_a = 12 100	Beef heart	(112)
Cytochrome c	(0.74)	1.4	7.7	M_{SD} = 18 000	Fungus : *Ustilago sphaerogena*	(112)
Cytochrome c	0.28 per cent Fe			M_a = 20 000	Fungus : *Ustilago sphaerogena*	(112)
Peroxidase	0.69	3.85*	6.8	M_{SD} = 44 000*	Horse radish	(113)
Lactoperoxidase	0.764	5.4*	6.0	M_{SD} = 93 000*	Cow's milk	(114)
Catalase	0.73	11.3	4.1	M_{SD} = 250 000	Cow's liver	(115)
Catalase	—	11.2	—	M_e = 220 000	Human erythrocytes	(116)
Catalase	—	11.0	—	M_e = 220 000	*Micrococcus lysodeiktus*	(116)
Catalase	—	11.2	—	M_e = 225 000	Horse liver	(117)
Copper enzymes						
Tyrosinase	(0.75)	6.4	6.1	M_e ~ 100 000	*Psalliota campestris*	(118)
Ascorbic acid oxidase	(0.75)	6.9	4.6	M_{SD} = 150 000	Summer crookneck squash	(119)
Dehydrogenases						
Alcohol dehydrogenase	0.751	4.9*	6.5	M_{SD} = 73 000*	Horse liver	(120)
Alcohol dehydrogenase	—	7.6*	—	M_e ~ 150 000*	Yeast	(120)
D-glyceraldehyde-3-phosphate dehydrogenase	—	—	—	M_a = 99 000	Rabbit skeletal muscle	(121)
L-glutamic acid dehydrogenase	0.75	26.6	2.54	M_{SD} ~ 1.10^6	Rabbit skeletal muscle	(122)
Yellow enzymes						
« New » yellow enzyme (Haas enzyme)	1 Flavin mol/mol enzyme			M_a = 60 000 to 65 000	Brewer's yeast	(123)
L-Aminoacid oxidase	(0.749)	6.9	11.0	M_{SD} = 61 000	Mocassin venom	(124)
L-Aminoacid oxidase	1 Flavin mol/mol enzyme			M_a = 62 800	Mocassin venom	(124)
Triphosphopyridine nucleotide-cytochrome c reductase	1 Flavin mol/mol enzyme			M_a = 68 000	Hog liver	(125)
Cytochrome c reductase	1 Flavin mol/mol enzyme			M_a = 75 000	Top ale yeast	(126)
« Old » yellow enzyme	0.731	5.76	6.3	M_{SD} = 82 000	Brewer's yeast	(127)
« Old » yellow enzyme	—	—	—	M_{eq} = 78 000	Brewer's yeast	(127)
L-Aminoacid oxidase	(0.75)	5.0	4.0	M_e ~ 120 000	Rat kidney	(128)
Glucose oxidase	—	8	5.1	M_e ~ 150 000	*Penicillium notatum*	(129)
Desmolases						
Carbonic anhydrase	(0.749)	2.8	9	M_e ~ 30 000	Beef blood	(130)
Aldolase	—	—	—	M_a ~ 140 000	Rabbit skeletal muscle	(131)
Myogen A	0.735	7.9	4.8	M_{SD} = 150 000	Rabbit skeletal muscle	(132)
Myogen A	—	—	—	M_{eq} = 136 000	Rabbit skeletal muscle	(132)
Carboxylase α	—	29	—	M_e ~ 10^6	Wheat germs	(133)
Miscellaneous enzymes						
Enolase	0.735	5.6	8.1	M_{SD} = 64 000	Yeast	(85)
Enolase	1 Hg atom/mol enzyme			M_a = 64 000 to 68 000	Yeast	(134)
Enolase	8 S atom/mol enzyme			M_a = 67 300	Yeast	(135) (134)
Enolase	—	—	—	M_i = 66 000	Yeast	(134)
Lipoxidase	0.750	5.6*	5.6	M_{SD} = 100 000*	Soy beans	(157)
Fumarase	(0.75)	8.5	4.0	M_{SD} ~ 200 000	Hog heart	(158)

TABLE IV

Molecular weights of hormones.

PROTEIN	V	s_{20}	D_{20}	M	ORIGIN	Literature
Secretin phosphate	—	1.03	—	$M_e \sim 5\,000$	Hog intestine	(136)
Secretin	1 mol tyrosine/mol secretin			$M_a = 5\,250$	Hog intestine	(137)
Relaxin	(2.7 per cent cystine and sedimentation)			$M_e \sim 9\,000$	Pregnant sow ovary	(138)
Insulin (monomer)	—	—	—	$M_e = 12\,000$	Pancreas	(44) (45) (139)-(141)
Adrenocorticotropic (ACTH)	(0.75)	2.1	10.4	$M_{sD} = 20\,000$	Sheep and hog pituitary	(142) (143)
Adrenocorticotropic (ACTH)	(0.75)	2.16	10.8	$M_{sD} = 19\,400$	Hog pituitary	(144)
Lactogenic hormone	—	—	—	$M_o = 26\,500$	Beef and sheep pituitary	(145)
Protein from posterior lobe	(0.749)	2.7	8.5	$M_{sD} \sim 30\,000$	Beef pituitary	(146)
Growth hormone	0.76	—	7.15	$M_o = 44\,000$	Beef anterior pituitary	(147)
Growth hormone	—	—	—	$M_a = 43\,600$	Beef anterior pituitary	(147)
Growth hormone	—	—	—	$M_a = 46\,800$	Beef anterior pituitary	(148)
Growth hormone	(0.76)	3.60	7.36	$M_{sD} = 49\,200$	Beef anterior pituitary	(149)
Growth hormone	(0.76)	2.95	7.2	$M_{sD} = 41\,000$	Beef anterior pituitary	(150)
Metakentrin (ICSH)	—	—	—	$M_o = 40\,000$	Sheep pituitary	(151)
Metakentrin (ICSH)	(0.749)	5.4	5.9	$M_{sD} = 88\,000$	Hog pituitary	(152)
Follicle stimulating hormone (FSH)	0.718	4.7*	6.0	$M_{sD} = 67\,000$*	Sheep pituitary	(153)
Chorionic gonadotropin	0.76	4.3	4.4	$M_{sD} \sim 100\,000$	Human pregnancy urine	(154)
Thyroglobulin	0.72	19.2	2.65	$M_{sD} = 630\,000$	Hog thyroid	(155)
Thyroglobulin	0.723	19.4	2.60	$M_{sD} = 650\,000$	Hog thyroid	(156)

REFERENCES

(1) S.P.L. Sörensen, *Compt. rend. trav. Lab. Carlsberg*, **12,** (1917). — See also : A.V. Güntelberg, and K. Linderström-Lang, *ibid.*, sér. chim., **27,** 18 (1949).
(2) G.S. Adair, *Proc. Camb. Phil. Soc.* (Biol.), **1,** 75 (1924).
(3) T. Svedberg, and Astrid Hedenius, *Biol. Bull.*, **66,** 191 (1934).
(4) K.O. Pedersen, *Kolloid-Z*, **63,** 268 (1933) and T. Svedberg, and K.O. Pedersen, *The Ultracentrifuge* (Oxford, 1940), 355-72.
(5) Inga-Britta Eriksson-Quensel, and T. Svedberg, *Biol. Bull.*, **71,** 498 (1936).
(6) S. Brohult, *J. Phys. Colloid Chem.*, **51,** 206 (1947).
(7) T. Svedberg, *J. Biol. Chem.*, **103,** 311 (1933).
(8) T. Svedberg, and Inga-Britta Eriksson-Quensel, *J. Am. Chem. Soc.*, **56,** 1700 (1934).
(9) H.B. Bull, *Advances in Enzymol*, **1,** 1 (1941).
(10) A. Rothen, *Ann. N.Y. Acad. Sci.*, **43,** 229 (1942).
(11) J.P. Johnston, H.C. Longuet-Higgins, and A.G. Ogston, *Trans. Faraday Soc.*, **41,** 588 (1945).
(12) A.D. Norris, *Nature* (London), **157,** 408 (1946).
(13) A. Rothen, *J. Gen. Physiol.*, **24,** 203 (1940).
(14) C.E. Danielsson, *Biochem. J.*, **44,** 387 (1949).
(15) K.O. Pedersen, *Ultracentrifugal Studies on Serum and Serum Fractions* (Uppsala, 1945).
(16) K.O. Pedersen, *J. Phys. Colloid Chem.*, **51,** 164 (1947).
(17) C.H. Li, *J. Am. Chem. Soc.*, **68,** 2746 (1946).
(18) T.L. McMeeking, B.D. Polis, E.S. DellaMonica, and J.H. Custer, *ibid.*, **70,** 881 (1948).
(19) R.A. Alberty, E.A. Anderson, and J.W. Williams, *J. Phys. Colloid Chem.*, **52,** 217 (1948); E.A. Anderson, and R.A. Alberty, *ibid.*, **52,** 1345 (1948), and R.A. Alberty, *J. Am. Chem. Soc.*, **70,** 1675 (1948).
(20) G.S. Hartley, *Quarterly Reviews*, **2,** 152 (1948).
(21) G. Scatchard, A.C. Batchelder, and A. Brown, *J. Am. Chem. Soc.*, **68,** 2320 (1946): G. Scatchard, I.H. Scheinberg, and S.H. Armstrong, Jr., *ibid.*, **72,** 535, 540 (1950).
(22) A.E. Alexander, and P. Johnson, *Colloid Science* (Oxford, 1949), p. 150.
(23) T. Svedberg, and K.O. Pedersen, *The Ultracentrifuge* (Oxford, 1940).
(24) M. Wales, *J. Phys. Colloid Chem.*, **55,** 282 (1951).
(25) L.G. Longsworth, *Proc. Natl. Acad. Sci. U.S.*, **36,** 502 (1950).
(26) R. Cecil, and A.G. Ogston, *Biochem. J.*, **43,** 592 (1948); **44,** 33 (1949).
(27) G. Kegeles, and F.J. Gutter, *J. Am. Chem. Soc.*, **73,** 3770 (1951).
(28) G.L. Miller, and R.H. Golder, *Arch. Biochem. Biophys.*, **36,** 249 (1952).
(29) J.F. Taylor, *ibid.*, **36,** 357 (1952).
(30) P.A. Charlwood, *Biochem. J.*, **51,** 113 (1952).
(31) D.F. Waugh, and D.A. Yphantis, *Federation Proc.*, **11,** 168 (1952), and *Rev. Sci. Instruments*, **23,** 609 (1952).
(32) D.A. MacInnes, Margaret O. Dayhoff and B.R. Ray, *Rev. Sci. Instruments*, **22,** 642 (1951), and Margaret O. Dayhoff, Gertrude E. Perlmann and D.A. MacInnes, *J. Am. Chem. Soc.*, **74,** 2515 (1952).
(33) H.K. Schachman, and M.A. Lauffer, *ibid.*, **72,** 4266 (1950).
(34) M. Wales, and J.W. Williams, *J. Polymer Sci.*, **8,** 449 (1952).
(35) L.G. Longsworth, *J. Am. Chem. Soc.*, **69,** 2510 (1947).
(36) C.A. Coulson, J.T. Cox, A.G. Ogston, and J. St. L. Philpot, *Proc. Royal Soc.* (London), A **192,** 382 (1948).
(37) L.J. Gosting, E.M. Hanson, G. Kegeles, and Margaret S. Morris, *Rev. Sci. Instruments*, **20,** 209 (1949).
(38) H. Svensson, *Acta Chem. Scand.*, **5,** 72, 1301, 1410 (1951).
(39) W.J. Archibald, *J. Phys. Colloid Chem.*, **51,** 1204 (1947).
(40) C.H. Li, A. Tiselius, K.O. Pedersen, L. Hagdahl, and H. Carstensen, *J. Biol. Chem.*, **190,** 326 (1951).
(41) P. Putzeys, and Jeanne Brosteaux, *Trans. Faraday Soc.*, **31,** 1314 (1935).

(42) P. Debye, *J. Phys. Colloid Chem.*, **51,** 18 (1947).
(43) B.H. Zimm, and P. Doty, *J. Chem. Phys.*, **12,** 203 (1944).
(44) P. Doty, M. Gellert, and B. Rabinovitch, *J. Am. Chem. Soc.*, **74,** 2065 (1952).
(45) P. Doty, and G.E. Myers, *Trans. Faraday Soc.*, Discussions No. **13,** 51 and 94 (1953).
(46) P. Doty, and J.T. Edsall, *Adv. Protein Chem.*, **6,** 35 (1951).
(47) A. Guinier, and G. Fournet, *C. R. Acad. Sci.*, Paris, **226,** 656 (1948).
(48) O. Kratky, *J. Polymer Sci.*, **3,** 195 (1948).
(49) A.R. Battersby, and L.C. Craig, *J. Am. Chem. Soc.*, **73,** 1887 (1951); **74,** 4023 (1952).
(50) Elizabeth J. Harfenist, and L.C. Craig, *ibid.*, **74,** 3087 (1952).
(51) Barbara W. Low, and F.M. Richards, *J. Am. Chem. Soc.*, **74,** 1660 (1952).
(52) I. Fankuchen, *Ann. N. Y. Acad. Sci.*, **41,** 157 (1941).
(53) D.P. Riley in « Un Symposium sur les Protéines », *Médecine et Biologie*, n° 5 (Liège, 1947), p. 35.
(54) E.J. Cohn, L.E. Strong, L.W. Hughes, Jr., D.J. Mulford, J.N. Ashworth, M. Melin, and H.L. Taylor, *J. Am. Chem. Soc.*, **68,** 459 (1946).
(55) J.R. Cann, R.A. Brown, S.J. Singer, J.B. Shumaker, Jr., and J.G. Kirkwood, *Science*, **114,** 30 (1951).
(56) M. Halwer, G.C. Nutting, and B.A. Brice, *J. Am. Chem. Soc.*, **73,** 2786 (1951).
(57) P. Putzeys, and Jeanne Brosteaux, *Mededeel. Koninkl. Vlaam. Acad. Wetenschap.*, Belg. Klasse Wetenschap., **3,** No. 1 (1941).
(58) E.L. Smith, D.M. Brown, H.E. Weimer, and R.J. Winzler, *J. Biol. Chem.*, **185,** 569 (1950).
(59) J.L. Oncley, G. Scatchard, and A. Brown, *J. Phys. Colloid Chem.*, **51,** 184 (1947).
(60) V.L. Koenig, and K.O. Pedersen, *Arch. Biochem.*, **25,** 97 (1950).
(61) J.M. Creeth, *Biochem. J.*, **51,** 10 (1952).
(62) J.T. Edsall, H. Edelhoch, R. Lontie, and P.R. Morrison, *J. Am. Chem. Soc.*, **72,** 4641 (1950).
(63) R.A. Kekwick, *Biochem. J.*, **32,** 552 (1938).
(64) M. Champagne, *J. Chim. Phys.*, **48,** 627 (1951).
(65) C.B. Laurell, and B. Ingelman, *Acta Chem. Scand.*, **1,** 770 (1947).
(66) C.G. Holmberg, and C.B. Laurell, *ibid.*, **2,** 550 (1948).
(67) K.O. Pedersen, unpublished experiments.
(68) W.B. Bridgman, *J. Am. Chem. Soc.*, **68,** 857 (1946).
(69) E.A. Kabat, *J. Exp. Med.*, **69,** 103 (1939).
(70) D.H. Cambpell, R.H. Blaker, and A.B. Pardee, *J. Am. Chem. Soc.*, **70,** 2496 (1948).
(71) T. Svedberg, and K. Andersson, *Nature*, **142,** 147 (1938), and unpublished experiments.
(72) S. Shulman, and J.D. Ferry, *J. Phys. Colloid Chem.*, **55,** 135 (1951).
(73) C.S. Hocking, M. Laskowski, Jr., and H.A. Scheraga, *J. Am. Chem. Soc.*, **74,** 775 (1952).
(74) A. Rothen, *J. Biol. Chem.*, **152,** 679 (1944).
(75) C.E. Danielsson, *Biochem J.*, **44,** 387 (1949).
(76) J.F. Foster, and D. French, *J. Am. Chem. Soc.*, **67,** 687 (1945).
(77) O.Quensel, *Untersuchungen über die Gerstenglobuline*, Almqvist and Wiksells Boktryckeri A.-B. (Uppsala, 1942).
(78) J.B. Sumner, N. Gralén, and Inga-Britta Eriksson-Quensel, *J. Biol. Chem.*, **125,** 45 (1938).
(79) C.C. Watson, S. Arrhenius, and J.W. Williams, *Nature*, **137,** 322 (1936), and unpublished experiments.
(80) E.A. Kabat, M. Heidelberger, and Ada.E. Bezer, *J. Biol. Chem.*, **168,** 629 (1947).
(81) P. Putzeys, and M.L. Beeckmans, *Bull. soc. chim. biol.*, **28,** 503 (1946).
(82) E.M. Bevilacqua, E.B. Bevilacqua, M.M. Bender, and J.W. Williams, *Ann. N. Y. Acad. Sci.*, **46,** 309 (1945).
(83) M.L. Beeckmans, and R. Lontie, *Bull. soc. chim. biol.*, **28,** 509 (1946).

(84) M. Kunitz, *J. Gen. Physiol.*, **35,** 423 (1952), and H.K. Schachman, *ibid.*, **35,** 451 (1952).
(85) G. Bergold, *Z. Naturfoschung*, **1,** 100 (1946).
(86) S. Schwimmer, and A.K. Balls, *J. Biol. Chem.*, **179,** 1063 (1949).
(87) C.-E. Danielsson, and E. Sandegren, *Acta Chem. Scand.*, **1,** 917 (1947). and C.-E. Danielsson, *Nature*, **162,** 525 (1948).
(88) C.-E. Danielsson, *ibid.*, **160,** 899 (1947).
(89) S. Englard, and T.P. Singer, *J. Biol. Chem.*, **187,** 213 (1950).
(90) K.J. Palmer, M. Ballantyne, and J.A. Galvin, *J. Am. Chem. Soc.*, **70,** 906 (1948).
(91) J.C. Lewis, N.S. Snell, D.J. Hirschmann, and H. Fraenkel-Conrat, *J. Biol. Chem.*, **186,** 23 (1950).
(92) G. Alderton, W.H. Ward, and H.L. Fevold, *ibid.*, **157,** 43 (1945).
(93) A. Passynsky, and V. Plaskeyev, *Comt. rend. acad. sci. U.R.S.S.*, **48,** 579 (1945).
(94) C. Fromageot, and M.P. de Garilhe, *Biochim. et Biophys. Acta*, **4,** 509 (1950).
(95) J.L. Oncley, *J. Biol. Chem.*, **151,** 27 (1943), and A.A. Green, *ibid.*, **158,** 315 (1945).
(96) M. Kunitz, and Margaret R. McDonald, *J. Gen. Physiol.*, **29,** 393 (1946).
(97) M. Kunitz, *ibid.*, **24,** 15 (1940).
(98) A. Rothen, *ibid.*, **24,** 203 (1940).
(99) I. Fankuchen, *ibid.*, **24,** 315 (1940).
(100) C.H. Carlisle, and Helen Scouloudi, *Proc. Roy. Soc.* (London), A **207,** 496 (1951).
(101) D.E. Green, L.F. Leloir, and V. Nocito, *J. Biol. Chem.*, **161,** 559 (1945).
(102) J.B. Sumner, N. Gralén, and Inga-Britta Eriksson-Quensel, *ibid.*, **125,** 37 (1938).
(103) F.W. Putnam, H. Neurath, E. Elkins, and S. Segal, *ibid.*, **166,** 603 (1946).
(104) E.L. Smith, D.M. Brown, and H.T. Hanson, *ibid.*, **180,** 33 (1949).
(105) J.H. Northrop, M. Kunitz, and R.M. Herriot, *Crystalline Enzymes* 2nd Ed. (New York, 1948).
(106) G.W. Schwert, *J. Biol. Chem.*, **190,** 799 (1951).
(107) E.L. Smith, D.M. Brown, and M. Laskowski, *ibid.*, **191,** 639 (1951), and E.L. Smith, and D.M. Brown, *ibid.*, **195,** 525 (1952).
(108) G.W. Schwert, and S. Kaufman, *ibid.*, **190,** 807 (1951).
(109) E.F. Jansen, and A.K. Balls, *ibid.*, **194,** 721 (1952).
(110) J. Shack, *ibid.*, **180,** 411 (1949).
(111) H. Theorell, *J. Am. Chem. Soc.*, **63,** 1804 (1941).
(112) J.B. Neilands, *J. Biol. Chem.*, **197,** 701 (1952).
(113) H. Theorell, and A. Akesson, *Arkiv Kemi, Mineral. Geol.*, **15** B, No. 24 (1942).
(114) H. Theorell, and K.O. Pedersen, in *The Svedberg*, 1884, 30/8, 1944 (Uppsala, 1944), p. 523.
(115) J.B. Sumner, and N. Gralén, *J. Biol. Chem.*, **125,** 33 (1938).
(116) R. Cecil, and A.G. Ogston, *Biochem. J.*, **43,** 205 (1948).
(117) K. Agner, *ibid.*, **32,** 1702 (1938).
(118) M.F. Mallette, and C.R. Dawson, *Arch. Biochem.*, **23,** 29 (1949).
(119) F.J. Dunn, and C.R. Dawson, *J. Biol. Chem.*, **189,** 485 (1951).
(120) H. Theorell, and R. Bonnichsen, *Acta Chem. Scand.*, **5,** 1105 (1951).
(121) S.F. Velick, and Ethel Ronzoni, *J. Biol. Chem.*, **173,** 627 (1948).
(122) J.A. Olson, and C.B. Anfinsen, *ibid.*, **197,** 67 (1952).
(123) E. Haas, *Biochem. Z.*, **298,** 378 (1938).
(124) T.P. Singer, and Edna B. Kearney, *Arch. Biochem.*, **29,** 190 (1950).
(125) B.L. Horecker, *J. Biol. Chem.*, **183,** 593 (1950).
(126) E. Haas, B.L. Horecker, and T.R. Hogness, *ibid.*, **136,** 747 (1940).
(127) R.A. Kekwick, and K.O. Pedersen, *Biochem. J.*, **30,** 2201 (1936).
(128) D.H. Moore, *J. Biol. Chem.*, **161,** 597 (1945).
(129) R. Cecil, and A.G. Ogston, *Biochem. J.*, **42,** 229 (1948).
(130) Mary Petermann, and N.V. Hakala, *J. Biol. Chem.*, **145,** 701 (1942).
(131) S. Undenfriend, and S.F. Velick, *ibid.*, **190,** 733 (1951).

(132) N. Gralén, *Biochem. J.*, **33**, 1342 (1939).
(133) T.P. Singer, and J. Pensky, *J. Biol. Chem.*, **196**, 375 (1952).
(134) T. Bücher, *Biochim. Biophys. Acata*, **1**, 467 (1947).
(135) O. Warburg, and W. Christian, *Biochem. Z.*, **310**, 384 (1942).
(136) E. Hammarsten, G. Agren, H. Hammarsten, and O. Wilander, *ibid.*, **264**, 275 (1933).
(137) P. Edman, and G. Agren, *Arch. Biochem.*, **13**, 283 (1947).
(138) E.H. Frieden, *ibid.*, **30**, 138 (1951).
(139) J.L. Oncley, and E. Ellenbogen, *J. Phys. Chem.*, **56**, 87 (1952).
(140) F. Tietze, and H. Neurath, *J. Biol. Chem.*, **194**, 1 (1952).
(141) H. Neurath, *Trans. Faraday Soc.*, Discussions No. **13**, 93 (1953).
(142) Ellen Burtner, *J. Am. Chem. Soc.*, **65**, 1238 (1943).
(143) C.H. Li, and K.O. Pedersen, *Arch. Biochem. Biophys.*, **36**, 462 (1952).
(144) E.L. Smith, D.M. Brown, B.N. Ghosh, and G. Sayers, *J. Biol. Chem.*, **187**, 631 (1950).
(145) C.H. Li, W.R. Lyons, and H.M. Evans, *ibid.*, **140**, 43 (1940).
(146) H.B. van Dyke, B.F. Chow, R.O. Greep, and A. Rothen, *J. Pharm. Exptl. Ther*, **74**, 190 (1942).
(147) C.H. Li, *J. Phys. Colloid Chem.*, **51**, 218 (1947).
(148) A.L. Franklin, C.H. Li, and M.S. Dunn, *J. Biol. Chem.*, **169**, 515 (1947).
(149) E.L. Smith, D.M. Brown, J.B. Fishman, and A.E. Wilhelmi, *ibid.*, **177**, 305 (1949).
(150) C.H. Li, and K.O. Pedersen, *J. Biol. Chem.* (in press).
(151) C.H. Li, Miriam E. Simpson, and H.M. Evans, *J. Am. Chem. Soc.*, **64**, 367 (1942).
(152) T. Shedlovsky, A. Rothen, R.O. Greep, H.B. van Dyke, and B.F. Chow, *Science*, **92**, 178 (1940).
(153) C.H. Li, and K.O. Pedersen, *J. Gen. Physiol.*, **35**, 629 (1952).
(154) H.P. Lundgren, S. Gurin, C. Bachman, and D.W. Wilson, *J. Biol. Chem.*, **142**, 367 (1942).
(155) M. Heidelberger, and K.O. Pedersen, *J. Gen. Physiol.*, **19**, 95 (1935).
(156) Y. Derrien, R. Michel, K.O. Pedersen, and J. Roche, *Biochim. Biophys. Acta*, **3**, 436 (1949).
(157) H. Theorell, R.T. Holman, and A. Akeson, *Acta Chem. Scand.*, **1**, 571 (1947).
(158) R. Cecil, and A.G. Ogston, *Biochem. J.*, **51**, 494 (1952).
(159) E. Fredericq, and H. Neurath, *J. Am. Chem. Soc.*, **72**, 2684 (1950).

Discussion du rapport de M. K. O. Pedersen

M. Adair. — In his most interesting paper on the molecular weights of proteins, Dr. Pedersen has referred to calculations based on measurements of osmotic pressure, including the value of 34.000 for egg albumin, recorded by Sörensen in 1917 (1). It has been suggested (2) that Sörensen's data are consistent with a higher value, 43.000. Sörensen himself pointed out that the calculation of the molecular weight of egg albumin was difficult on account of the effects due to the unequal distribution of ions and the alterations in the state of aggregation of proteins caused by salts.

In studies of the osmotic pressures of solutions of haemoglobin, begun in 1921, solutions with different concentrations of protein were dialysed against large volumes of a standard salt solution or buffer mixtures. The results obtained could be represented by a modification of van der Waals equation $\pi\ (V\text{-}b) = RT$. The molecular weight was estimated as 66.700 ± 6000 (3).

Recently, it has been found that the number average molecular weight of horse haemoglobin is constant within 1 %, over a fairly wide range of concentration (4).

M. Havinga. — A particle showing all the chemical properties of the material should equally show all the biological properties. Moreover it seems difficult to imagine particles having different physical properties and identical chemical properties. Im my opinion, nowadays there will be no significant difference in the definition of a molecule used by an organic chemist, by a biologist or by a physical chemist.

The determination of M by spreading in a monolayer (Bull a. o.) has not been mentionned. Of course, this method has the disad-

(1) Sörensen : *Comptes Rendus Lab. Carlsberg*, Vol. 12 (1917).
(2) Adair . *J. Amer. Chem. Soc.* **49**, 2524 (1927).
(3) Adair : *Proc. Camb. Phil. Soc. Biol.*, **1**, 75 (1924).
(4) Adair : *Barcroft Memorial Volume Hæmoglobin.* Butterworth's Scientific Publications. London (1949).

vantage of making measurements on material in a more or less « denatured » state. However, it seems to be a valuable method in many cases. One might obtain information by this method (e. g. on dimerisation phenomena) in a less difficult way than by other methods of investigation.

I should like to stress the danger that exists in pooling material obtained from a number of individuals, if one is interested in obtaining strictly pure proteins. There is a clear example in the case of the human hemoglobin. Studies, made in Dr. Pauling's department by Dr. H. Itano a. o., have proved the existence of several essentially different « modifications » of human hemoglobin. Such hereditary differences between individuals might also exist in the synthesis of other proteins.

M. Pedersen. — It has been stated by Graig that insulin should have a minimum molecular weight around 6 000. Under normal and physiological conditions, however, such small particles cannot be isolated. Until now the physico-chemical methods have given a minimum of about 12 000 on extrapolation to zero protein concentration. The higher values for M found in solutions of insulin may be explained by the existence of dimers, trimers, etc.

M. Pauling. — Professor Havinga has suggested that I should report the work done on hemoglobin in our laboratories, and discuss its bearing on the question of the homogeneity of proteins obtained from individual animals of a species, or from several individuals.

Hemoglobin used to be considered an especially good protein, because of its ease of crystallization and its well-defined species specificity. In 1946 I had the idea that the disease sickle-cell anemia might be due to an abnormal form of the hemoglobin molecule. Work was begun on the problem by Dr. Harvey Itano, who was aided also by Drs. Singer and Wells. It was found that patients with sickle-cell anemia have in their red cells an abnormal form of hemoglobin, differing from normal hemoglobin in isoelectric point by 0.2 pH unit (Pauling, Itano, Singer, and Wells, *SCIENCE, 1949*). The carriers of the disease (parents of the patients) have roughly equal amounts of normal adult human hemoglobin and sickle-cell-anemia hemoglobin in their red cells.

Dr. Itano has recently discovered two other abnormal forms of human hemoglobin, hemoglobin c and hemoglobin d. Hemo-

globin *c* differs in isoelectric point from normal adult hemoglobin by 0.4 units. Hemoglobin *d* has the same electrophoretic properties as normal adult hemoglobin, but is much less soluble. Both hemoglobin *c* and hemoglobin *d* are associated with previously unrecognized diseases, hereditary hemolytic anemias resembling sickle-cell anemia.

Dr. Itano has found that some individuals have in their red cells three different kinds of hemoglobin; for example, sickle-cell-anemia hemoglobin, hemoglobin *c*, and fetal hemoglobin (which the anemic individual has continued to manufacture as an adult).

These observations cause us to be skeptical about the homogeneity of protein preparations from several animals of a species, and even from a single individual. It seems reasonable to assume that the situation found for hemoglobin may apply to many other proteins. The fact that different forms of adult human hemoglobin crystallize together shows that crystallization cannot be taken as a criterion of homogeneity.

M. Tiselius. — The examples given by Dr. Pauling show how necessary it is to base conclusions regarding the homogeneity of proteins on several independent methods. Thus it seems highly desirable that efforts are made in working out new and highly specific protein fractionation methods. Recent work on the determination of amino acid sequences in proteins (Sanger) must be based upon strictly homogeneous material if it is not to lead to extreme complications.

I would like to stress the importance of the application of ultracentrifugation for determination of molecular weights in the region of, say 1.000 - 10.000, where we would expect many important polypeptides to belong. Other methods meet with great difficulties in this region. It would only seem desirable to simplify and standardize the Archibald method to make it less laborious.

The fact that molecular weights of many proteins tend to group themselves into certain size classes of which each may contain chemically quite different proteins appears to me to be of great interest, even if these do not follow any simple numerical relationships, as was originally assumed. This indicates a common pattern in the structure (and biosynthesis) of many proteins. The reversible dissociation and association of some of the large protein molecules, brought about by changes in pH, seem to follow the rule of simple

multiples. Thus the dissociation occurs in 1/2, 1/4, 1/6, etc. These reversible dissociation-association reactions offer many interesting problems, particularly with regard to the specificity of the forces involved, and the degree to which they influence the biochemical properties.

It seems to me that it would be extremely important if more attention could be paid to the problem of performing *zone separation* in the ultracentrifugal cell, so that one could isolate different components completely from each other, and be able to take them out for further analyses. Such arrangements are, as you all know, common in chromatography, and recently also in electrophoresis, and have proved their value amply. In the long run, it can not be satisfactory to identify the various components only as « peaks » in diagrams obtained by optical observations. They must be made accessible to detailed chemical investigation if we are to make full use of the information obtainable from ultracentrifugation experiments.

M. Pedersen. — Archibald's method involves laborious computations, especially at the time being where we are testing the method as such. We are now running the experiments all through to equilibrium in order to compare the values obtained by Archibald's method with those from sedimentation equilibrium. The results so far obtained have been satisfactory. One may hope that when this checking period is over it may be possible to simplify and standardize the whole procedure so that this valuable method could be more generally used in the future.

M. Bragg. — Judging from the X-ray diffraction pictures, the physical differences between the different forms of haemoglobin must be restricted to very small details of the whole molecule. For instance, Professor Pauling has referred to the very different properties of normal and sickle-cell haemoglobin molecules. Perutz has shown that crystals (possibly liquid crystals) form within the blood corpuscles, owing to the much lower solubility of the sickle-cell haemoglobin when it loses oxygen. Yet normal and sickle-cell haemoglobins give identical X-ray pictures although these pictures are very sensitive to the size, shape and structure of the molecule. Correspondingly, oxy-, met-, and carboxy-haemoglobin crystallise in the same form. Perutz has followed the change from one to the other by optical methods inside the solid

crystal. To use an analogy, these molecules which have such very different physico-chemical properties would look, if we saw enlarged models of them, like automobiles of identical make but with some slightly different detail such as different number-plates.

M. Chibnall. — On the question of bulk preparation of proteins from animals, we cannot of course avoid this in structure studies on insulin. In the case of the beef variety, Craig find by his counter-current procedure that he can separate « *A* » and « *B* » peaks. With english (Boots) insulin, he finds less than 10 % of « *B* ». The overall amino-acid analysis of « *B* » differs from « *A* » only in one respect, per mol. wt. of 6.000 it contains one less amide group. « *B* » must therefore be more acidic than « *A* », irrespective of any sequence difference that future work may disclose. A small difference, but a significant one.

Craig's method of determining molecular weights depends on the separation of a peak representing a monosubstituted derivative : is it not possible that on account of some configurational feature in the molecule this is really a disubstituted derivative ?

M. Desreux. — *A*.

La méthode de détermination de la grandeur moléculaire par mesure de la pression superficielle d'un film monomoléculaire de protéine mérite plus de crédit qu'elle n'en a eu jusqu'à ce jour. Une telle détermination n'exige que des quantités extrêmement faibles de matière et, en une seule opération, on obtient les valeurs de *FA* pour différentes aires, c'est-à-dire pour différentes concentrations superficielles. Il est d'autre part aisé, à l'aide de techniques simples, d'obtenir une très bonne précision et, dans certains cas, une excellente concordance avec les poids moléculaires obtenus par sédimentation-diffusion.

Citons comme exemple celui relatif à deux papaines cristallisées (Close, Dieu) :

		P.M. séd.-diff.	pres. superf.
papaine :	*A'*	30.500	31.500
	A	24.500	23.500

Le poids moléculaire chimique de la papaine *A'*, déterminé au laboratoire du Professeur Bigwood, est de 32.500.

L'objection généralement faite à cette technique est le danger d'une dénaturation en surface; il nous semble cependant que dans

de nombreux cas ce facteur n'est pas tellement perturbateur puisque le poids moléculaire obtenu est une valeur en nombre.

Une théorie plus complète du comportement des couches monomoléculaires reste cependant à établir.

B. — Nous ne croyons pas que la mesure de l'intensité de la lumière diffusée par les solutions de protéine présente actuellement encore de grandes difficultés techniques; par contre, dans la détermination des coefficients de dissymétrie, il faut, avec certains dispositifs expérimentaux, tenir compte de facteurs correctifs dus à des réflexions multiples sur les parois de la cellule de mesure.

Ces corrections deviennent importantes dans le cas de solutions de particules très asymétriques (voir A. et J. Oth et V. Desreux par ex. : *J. Polymer. Science,* 10, 551, 1953).

C. — D'expériences récentes, il résulte que le problème de la détermination de la grandeur moléculaire de petites molécules par ultra-centrifugation peut être résolu par emploi de la nouvelle cellule imaginée par Kegeles et la firme Spinco (synthetic boundary cell).

M. Linderström-Lang. — Pointed to the possibility of measuring the temperature of the centrifuge cell colorimetrically by selecting a system consisting of a buffer and an indicator with widely different heats of ionization. Experiments carried out by Ottesen and Linderström-Lang have shown that thymol blue in glycine buffer could be applied.

Emphasized that Sörensen's value of 34 000 for the *Mw* of ovalbumin was based on few out of a large series of measurements. The majority of the measurements, although carried out at a pH value far from the isoelectric point, gave the value 45 000 if extrapolation to zero protein concentration was made.

Pointed to the possibility that the ratio $CONH_2/COOH$ for the dicarboxylic amino acids in proteins secreted e. g. by the milk glands (β-lactoglobulin) may depend upon the acid-base balance of the animal.

Pointed to the usefulness of the Edman method in testing the purity of a given protein.

M. Fredericq. — In relation to the problem of the determination of very low molecular weights, we had recently some new results on

insulin. We measured sedimentation constants in alkaline solution and obtained some evidence for a minimum molecular weight of 6.000, by extrapolation of sedimentation and diffusion data. However the determination of rate of sedimentation in patterns where the peak is not completely separated from the meniscus presents difficulties.

We made new experiments with the Pickel's « synthetic boundary » cell. Here the boundary is formed at the start of the run, so that we can measure accurately the displacement of the peak, even for small molecules. The value of $S_{20,w}$ as a function of insulin concentration is almost constant, indicating that complete dissociation has been practically reached. At low protein concentration, it is close to 1.3 Svedberg, which is indicative of a molecular weight between 6,000 and 7,000.

Since the question was raised of using surface monolayers for determining molecular weights, I should mention that we got a very satisfactory agreement between monolayers studies and sedimentation-diffusion data for the dissociation of insulin in phosphate buffers between pH 2 and 4. Under certain conditions, complete dissociation takes place.

So, we would say as a conclusion, that the submolecule of insulin probably has a molecular weight of 6,000 but that complete dissociation is reached only in a few favourable cases. Generally, the dissociation will not go further than a dimer of 12,000.

M. Hermans. — The determination of molecular weight from light scattering is completely straight-forward when we are dealing with a single solute. The problem is much more complex in multicomponent systems. It is important to realize that a charged protein with its counterions in the presence of extraneous electrolytes represents a multi-component system. I should like to ask Prof. Pedersen whether any cases have been met with where the apparent molecular weight proved to be affected by the type of extraneous electrolyte present.

In connection with the remarks made by Dr. Adair, I should like to emphasize that osmotic measurements in electrolytic systems should preferably always be made in the presence of excess electrolyte. The plots of reduced osmotic pressure π/c versus c become straight lines whose slope is approximately proportional to the reciprocal of the concentration of extraneous electrolyte (see, for

example, Pals and Hermans, *Rec. trav. chim.*, 71, 458, 1952). The lines extrapolate to the same π/c-value when c approaches zero, thus giving a more reliable value for the molecular weight.

Dr. Pedersen says in his article that s and D are always reduced to standard conditions; namely, to water at 20° C. Now, s and D separately are both affected by solvation and by shape, although the ratio s/D is not. The question arises whether the reduction of the data to standard conditions might affect the ratio s/D in those cases where s and D have not been measured under exactly similar circumstances. In this connection, I should like to ask also, whether the influence of the pressure on the partial volume in a sedimentation run is likely to be sufficiently large to affect the molecular weight calculated from s/D.

M. Pedersen. — Edsall and coworkers [my reference ([62])] have studied the effect of electrolytes on the light scattering of serum albumin.

When we reduce s and D to standard condition, we always try to do the measurements at a temperature not too far from 20° C, so the total correction is kept small. Difficulties arise when we cannot work in dilute salt solutions, but must use more concentrated ones. The correctness of the applied corrections over a larger temperature interval has not yet been verified. At present the accuracy in the determination of D is too small to justify such a test.

At Yale they have studied the compressibility of some proteins and found it extremely small. The effect of pressure in the ultracentrifuge cell cannot affect the computed molecular weights.

M. Theorell. — We have now heard a great deal about physical methods for determining the molecular weights of proteins. I would like to call your attention to the importance of chemical analyses for establishing the minimum *M. w.* Of course, a pure protein has to obey the general law for pure compounds, saying that all the constituents must be present in multiples of integers. This does not help us much in the case of some of the amino acids, present in higher numbers in the molecules, because the analytical errors may make it impossible to decide their accurate number in the molecule. But in most of the conjugated proteins, the minimum *M. w.* can be determined with a fair degree of accuracy by analyzing the content of prosthetic group. Let us consider the

hemoproteins where we have the double possibility of determining iron and hematin content. Iron can be determined with great accuracy, but one has to recall that it is never certain that all traces of non-hematin iron have been removed. However, in a pure compound, the iron values should fit to the hematin content that is in general determined spectrophotometrically, as pyridine hemochromogen. This estimation must be founded on absorption coefficients, obtained from a pure sample of hemin. Unfortunately we have recently found that such a sample cannot be prepared by recrystallisation or any method used hitherto. Our results indicate that even the highest light absorption values used by several workers, like Drabkin and by ourselves, may be too low by as much as 10 %. This would correspondingly increase the calculated values for the minimum *M.w.*, and this is a great nuisance because, at the same time, the ultracentrifugal data have in some cases to be recalculated in the opposite direction, when the sedimentation constants obtained in the oil turbine centrifuges are corrected for the temperature differences now discovered.

In order to illustrate the magnitude of the uncertainty caused by these conditions, we may consider the horse radish peroxidase. Our first crystals, ten years ago, had a hemin content, as analysed by the pyridine method and using $\varepsilon_{557m\mu} = 7.3$, of 1.48 %, and a *M.w.*, determined by sedimentation and diffusion, of 44 000. This value fitted exactly to the value calculated from the spectrophotometry. Keilin and Hartree recently stated that they had produced a purer product with 1.61 % hemin, and a *M.w.*, calculated from a sedimentation constant determined by Cecil and Ogston, of 40 000, again in good agreement with the hemin content. When they kindly sent us a sample for comparison, we got a lower spectrophotometric value, 1.49 %, in excellent agreement with our old value. But since we now have reason to believe that the value of $\varepsilon_{557m\mu}$ may be perhaps 10 % too low, the recalculated value would de *hemin = 1,37 %, M.w. = 48 000.* This shows the need for more accurate data in the chemical determination of the minimal *M.w.* But these should be comparatively easy to obtain.

M. Pedersen. — Several of the previous experimental data for s_{20} from the oil-turbine ultracentrifuge have to be corrected and the corresponding MsD values recalculated when the exast magni-

tude of the correction for the cell temperature is clear (see my report pp. 29-30). Values obtained by sedimentation equilibrium method will not, however, be affected by this correction.

M. Putzeys. — *A.* En ce qui concerne *l'élimination des poussières* des solutions destinées aux mesures de diffusion moléculaire de la lumière, celle-ci ne cause actuellement plus de difficultés, comme on l'a déjà fait remarquer. Plus difficile est l'élimination de grosses particules qui ont pris naissance aux dépens de la protéine elle-même. Leur élimination par précipitation fractionnée en cours de préparation est très efficace. La centrifugation finale, à grande vitesse, immédiatement avant l'examen de la solution, est cependant toujours indispensable. La présence de ces grosses particules se laisse fréquemment repérer par l'existence d'une dissymétrie anormale de la diffusion moléculaire de la lumière.

B. — La sensibilité de la diffusion moléculaire de la lumière à la présence de grosses particules la rend particulièrement apte à nous renseigner sur l'état de la solution examinée. Ces grosses particules sont éliminées tout au début de la centrifugation lorsqu'on fait une mesure à l'ultracentrifuge, de sorte qu'une solution qui se révèle parfaitement homogène dans cet appareil peut montrer un poids moléculaire moyen anormalement élevé lorsqu'on l'examine à la diffusion moléculaire de la lumière.

C. — A cause de cela, il est en général très difficile de trouver une protéine qui puisse servir d'étalon pour la diffusion de la lumière. L'ovalbumine surtout donne fréquemment un poids moléculaire beaucoup trop élevé et augmentant avec l'âge des solutions. La carboxyhémoglobine du cheval donne facilement un poids moléculaire correct dans le voisinage de pH 8.0., mais les remarques de Theorell nous invitent à la prudence.

Nous avons employé depuis le début l'amandine comme protéine de référence. Les préparations obtenues à plusieurs reprises, depuis une vingtaine d'années, accusent toujours la même diffusion moléculaire de la lumière endéans 5 %.

Il est indispensable aujourd'hui d'obtenir une précision plus grande. Nous examinons actuellement la possibilité de cristalliser l'amandine sous forme de son complexe mercurique. Reste à voir si le nouveau procédé de préparation donnera un résultat meilleur que le procédé employé jusqu'à présent.

D. — En ce qui concerne la question de J.-J. Hermans concernant l'influence de divers électrolytes sur le poids moléculaire obtenu par mesure de la diffusion moléculaire de la lumière, on peut dire que la nature des électrolytes (chlorures de Li, Na et K, sulfates Na, NH_4 et Mg) n'a pas d'influence sur le poids moléculaire extrapolé en fonction de la concentration en protéine, à condition que la force ionique ne soit pas telle qu'elle provoque une dissociation ou une association des molécules. La nature et la concentration des électrolytes peut influencer la pente de la courbe d'extrapolation (Edsall, Doty, Lontie, dans le *Journ. Am. Chem. Soc.).*

M. Pedersen. — It is quite clear that dust and grossly aggregated particles should be removed before any light scattering, diffusion, or sedimentation equilibrium experiment is carried out. For the determination of s this is not important, since these particles are thrown out of the solution at the beginning of the run. For runs with the Archibald's method, we have found it necessary to remove any such particles by making a short preliminary run at high speed in an ultracentrifuge cell equiped with filter paper at the bottom. Constant composition of the solution is revealed by constancy of refractive index.

M. Synge. — Dr. Synge asked Dr. Pedersen and Dr. Adair what are the prospects for determining the physical properties (sedimentation, diffusion, osmotic pressure, etc.) of proteins and peptides dissolved in multicomponent solvent mixtures such as those used by Craig and colleagues. The problem of the molecular state of insulin is topical just now, but other cases are bound to arise. It would be of great value to have physical data on solutes in systems where there are chemical grounds for suspecting dissociation effects.

M. Pedersen. — Multicomponent systems, such as used by Craig, are indeed difficult to handle from a purely physicochemical point of view. The experimental data would be very difficult to evaluate, if not completely impossible. The variation in the composition of the solvent mixture with time and position in the cell is a problem in itself on top of the protein problem.

M. Anson. — I should like to add two remarks to Pauling's discussion of genetically controlled variations in the composition of individual proteins.

My first remark deals with the possibility of detecting by spectroscopic means differences in the hemoglobin of individuals of the same species. Barcroft, in the early twenties, observed that species differences in hemoglobins could be detected by differences in the shift of the main band of oxyhemoglobin, when the oxygen was displaced by carbon monoxide. This shift, the so-called span, could be easily and accurately measured with the Hartridge reversion spectroscope. While working in Barcroft's laboratory, Mirsky and I observed further that sometimes even hemoglobin of different individual rabbits had different spans. At the time, we thought that variations in the span of individual rabbit hemoglobins might possibly have a genetic basis. It would be interesting to find out whether the span of sickle cell hemoglobin is different from the span of normal hemoglobin, and to see whether further differences in the hemoglobins of different individuals of the same species can be discovered by the easy measurements of span.

My second remark concerns the important implications for agriculture of the possibility of breeding new varieties of plants with proteins of new compositions. The common grain proteins consumed by domestic animals and man are low in lysine, and the utilization of an entire diet can be limited by its lysine content. It would be of very great importance, for instance, if new varieties of grains could be obtained with proteins even moderately richer in lysine. The Pauling work shows that breeding for improved amino acid composition is at least not out of the question. There are, indeed, reports in the literature of quite great variations in the amino acid composition of the proteins of different samples of a particular kind of plant. Unfortunately, the reported variations, although great, may be merely errors in analysis. In the one case studied carefully, the proteins of many varieties of soybean were analyzed without any important variation in amino acid composition turning up.

From the practical point of view, it would not make any difference whether an inherited change in amino acid composition were due to a change in the composition of an individual protein or to a change in the proportions of different proteins. From the theoretical point of view, however, it would be interesting to know whether both the composition of individual proteins and the proportions of different proteins can be changed by mutations.

The Configuration of Polypeptide Chains in Proteins

by Linus Pauling

The problem of the structure of proteins has been under attack by investigators during the last one hundred years. Since the work of Emil Fischer fifty years ago it has been recognized that proteins involve polypeptide chains. There are two essential parts to the problem of determining the structure of proteins: first, the determination of the configuration of polypeptide chains in proteins, and second, the determination of the sequence of amino-acid residues in the chains, and the configuration of the side-chain groups. (The determination of the structure of prosthetic groups in conjugated proteins and their relation to the polypeptide chains may be considered incidental to the problem of the determination of the structure of proteins. Nucleoproteins, a most important class of conjugated proteins, constitute a problem in themselves.)

There has recently been obtained a partial solution of the problem of the configuration of polypeptide chains in proteins. This advance in knowledge has resulted from the application of the technique of x-ray diffraction to the problem. It has resulted in large part not from the direct investigation of proteins themselves with x-rays, but rather from the precise determination of the structure of amino acids, simple peptides, and other substances related to proteins, and also from the application of the quantum mechanical theory of the chemical bond.

1. EARLY X-RAY STUDIES OF PROTEINS

The first x-ray photographs of proteins that are described in the literature were made by Herzog and Jancke [1] in 1920. These investigators studied muscle, nerve, sinew, hair, and silk. They found that these fibrous proteins produce rather diffuse x-ray fiber diagrams. The diagrams given by silk are especially good, in that they show a rather large number of moderately sharp reflections. In 1923 a detailed study of silk was reported by Brill [2]. He found that the data for silk could be accounted for by a monoclinic unit of structure, containing eight amino-acid residues. He discussed a number of alternative structures for the crystal, and concluded that the amino-acid residues are not present as small cyclic peptides, but rather are present as polypeptide chains, with two residues of each of four chains passing through the unit cell. The dimension of the unit cell in the direction of the silk fibers is 7.0 A; in 1928 Meyer and Mark [3] pointed out that this dimension corresponds to two residues in a nearly fully extended polypeptide chain, as calculated with use of the approximately known values of interatomic distances and bond angles, and they concluded that the polypeptide chains in silk fibroin are nearly completely extended, in the direction of the axis of the fiber. Other early investigations of silk fibroin were reported by Kratky [4], Kratky and Kuriyama [5], and Trogus and Hess [6].

One difficulty about the structure of silk is that the volume of the unit of structure is such that although four residues of glycine and four residues of alanine (perhaps one or two residues of serine in place of alanine) may be included within it, there seems not to be room for a tyrosine residue. Tyrosine is found by analysis to constitute approximately 6 percent of the residues. The suggestion has been made by Meyer and his collaborators that in silk the polypeptide chains have in some regions such composition as to permit them to crystallize, whereas in other regions the structure remains amorphous; this may well be the correct explanation [7].

Rapid progress in the understanding of hair, horn, and related proteins of the keratin class was made in the years following 1930, through the efforts of Astbury and his collaborators. In 1931, Astbury and Street [8] pointed out that hair, wool, and other fibers of the keratin class can exist in either a contracted form, which they named α keratin, or an extended form, β keratin. They

suggested that in β keratin the polypeptide chains are in the nearly completely extended configuration. The observed identity distance along the fiber axis in β keratin, obtained by stretching hair or horn, is 6.6 A, which is about 6 percent less than in silk; this suggests that the polypeptide chains are not quite completely extended in the β-keratin proteins. Astbury and Street also suggested that a chemical change involving the formation of rings of atoms held together by covalent bonds takes place on the contraction of β keratin to α keratin. In 1933, Astbury and Woods [9] described β keratin as involving sheets of extended polypeptide chains, the chains being held together in the sheet by the interaction of carbonyl and imino groups of the polypeptide amide groups. The interaction between the chains was recognized in 1936 as that corresponding to the formation of a CO···H — N hydrogen bond [10, 11]. Further consideration of the nature of possible cyclic structures for α keratin led to their abondonment, and it was rather generally accepted that in α keratin the polypeptide chains are folded or coiled in such a way as to give a fiberaxis length of only about 1.5 A per residue, in place of 3.3 A per residue for β keratin.

A number of possible configurations were then suggested for the coiled polypeptide chain in α keratin. A discussion of two of these configurations and of the structure of the α-keratin proteins is given below, in section 8.

Muscle and some of the protein constituents of muscle (myosin, actomyosin) give x-ray photographs similar to those of keratin [12, 13].

The x-ray diagram of collagen is quite different from that of the proteins of the keratin class. The principal spacing along the fiber axis is about 2.9 A, and this value has been assumed to be the average fiberaxis length per amino-acid residue in the polypeptide chain. Several ways of folding the polypeptide chain have been suggested for collagen [14]. These are discussed briefly in section 15. Gelatin was shown by Gerngross and Katz to have the same structure as collagen [15].

Some fibrous proteins give rise to x-ray diagrams showing a large number of diffraction maxima, corresponding to unit cells with very large dimensions, of the order of magnitude of 100 A. Porcupine quill and feather rachis are outstanding in this respect. Descriptions of these photographs and measurements of the long

spacings have been reported by Astbury and Marwick [16], Corey and Wyckoff [17], MacArthur [18], Bear [19] and others.

The first good photographs of crystalline globular proteins to be described in the literature were those of insulin made by Bernal and Crowfoot [20], by use of the technique of keeping the crystals surrounded by mother liquor during the photography. A detailed discussion of the data for this crystal, with application of the Patterson method of interpreting the data, was published by Crowfoot [21]. About a score of other globular proteins have since been investigated. The most interesting results that have been obtained are those of Bragg, Perutz, and Kendrew, and their collaborators, who have studied a large number of crystalline varieties of hemoglobin and myoglobin. The data on horse hemoglobin [22, 23] have been published as a set of sections of a three-dimensional Patterson diagram, from which it has been possible to draw the conclusion that the hemoglobin molecule is about 57 A long, and between 34 A and 57 A in other dimensions, and that there are present rods extending in the 57 A direction, and packed in a pseudohexagonal array, with the centers of the rods about 10.5 A apart. These results indicated that the rods probably have the same structure as the molecules in α keratin, which also are about 10.5 A in diameter. Kendrew [24] has interpreted his data for myoglobin as showing that similar rods are present also in this molecule. A more detailed discussion of the structure of hemoglobin and myoglobin will be given in a later section, section 10.

2. THE STRUCTURE OF CRYSTALS OF AMINO ACIDS AND PEPTIDES

Fifteen years ago, after we had made an unsuccessful effort to determine the configuration of polypeptide chains in proteins, Professor Robert B. Corey and I decided to attack the problem in an indirect way, through the determination of structures of amino acids, peptides, and other simple substances related to proteins, and the formulation of basic structural principles with the use of which acceptable configurations for polypeptide chains could be formulated.

At that time there had not yet been determined the crystal structure of any amino acid or peptide. The first x-ray investigation of crystals

of amino acids had been carried out in 1931 by Bernal [25]. He determined the unit cells and space groups of fifteen crystals of amino acids and related substances. His work, however, did not lead to the determination of the atomic arrangement for any of these crystals. Hengstenberg and Lenel [26] in the same year reported a structure determination of glycine which is incorrect.

The investigators who have taken part in our program of x-ray studies designed to lead to the determination of the configuration of polypeptide chains in proteins include Dr. Gustave Albrecht, Dr. Henri A. Lévy, Dr. E.W. Hughes, Dr. Jerry Donohue, Professor Verner Schomaker, Professor David P. Shoemaker, Dr. Walter J. Moore, Dr. Kenneth N. Trueblood, Dr. H.R. Branson, Dr. Sidney Weinbaum, Dr. R.A. Pasternak, Dr. Gene B. Carpenter, Dr. Harry L. Yakel, Jr., and Dr. Richard E. Marsh.

The first structure reported by Professor Corey was that of diketopiperazine [27]. It was found that the diketopiperazine molecule is planar; each of the two amide groups that constitute the ring is planar, with the cis configuration (C = O and N — H on the same side of the C — N axis, in the plane of the amide group). The molecules are arranged in strings, apparently held together by N — H···O hydrogen bonds, with length 2.8 A and with the oxygen atom lying very close to the N — H axis. The structure of glycine was reported by Albrecht and Corey [28] in 1939. This investigation was followed by the precise determinations of the structure of crystals of alanine [29, 30], L_s-threonine [31], hydroxy-L-proline [32], and DL-serine [33]. Crystals containing amide groups that have been subjected to complete structure determination include β-glycylglycine [34], α-glycylglycine [35], N-acetylglycine [36], and N,N′-diglycylcystine [37].

3. FORMULATION OF BASIC STRUCTURAL PRINCIPLES

In Table 1 there are given values of the interatomic distances corresponding to bonds in the amide group, as derived from the x-ray investigation of crystals of amino acids, peptides, and related compounds. The values for urea [38], the complex of urea and hydrogen peroxide [39], and urea oxalate [40] are, of course, of less direct significance to the problem of the dimensions of the

polypeptide chain than are those for the peptides. An earlier and less accurate determination of the structure of hydroxy-L-proline has been published by Zussman (41). A structure that has been reported for cysteyl-glycine sodium iodide by Dyer (42) is probably not correct. The structure of L-glutamine has been determined by Cochran and Penfold (43). The amide group is planar to within the accuracy of the structure determination, and the bond distances and bond angles are also in approximate agreement with the values reported in Tables 1 and 2.

TABLE 1

Bond Lengths of the Amide Group as Derived from Those Found in Crystals of Amino Acids, Peptides, and Related Compounds.

	N-αC Å	C′-αC Å	C′-O Å	C′-N Å	Reference
DL-Alanine*	1.50	1.54	—	—	Donohue, 1950
L-Threonine*	1.49	1.52	—	—	Shoemaker, Donohue Shoemaker & Corey, 1950
DL-Serine*	1.49	1.53	—	—	Shoemaker, Donohue Barieau & Lu
Hydroxy-L-proline* ...	1.50	1.52	—	—	Donohue & Trueblood
N-Acetylglycine*	1.45	1.50 1.51	1.24	1.32	Carpenter & Donohue, 1950
β-Glycylglycine	1.48	1.53 1.53	1.23	1.29	Hughes & Moore, 1949
α-Glycylglycine*	1.47	1.56 1.52	1.24	1.32	Hughes & Biswas
N, N′-Diglycylcystine*.	1.48	1.56 1.52	1.21	1.35	Hughes & Yakel
Urea*	—	—	1.26	1.34	Vaughan & Donohue
Urea . H_2O_2	—	—	1.24	1.34	Lu, Hughes & Giguère, 1941
Urea Oxalate*	—	—	1.26	1.34 1.35	Schuch, Merritt & Sturdivant
Selected Value	1.47	1.53	1.24	1.32	

* Three-dimensional refinement of the atomic positions.

TABLE 2

Bond Angles of the Amide Group as Derived from Those Found in Crystals of Peptides.

	Around carbonyl carbon atom			Around amide nitrogen atom		
	αC-C′-O	N-C′-O	αC-C′-N	C′-N-αC	C-N-H	αC-N-H
N-Acetylglycine*	121.0°	121.3°	117.7°	119.6°	—	100° ± 10° +
β-Glycylglycine	121	125	114	122	—	—
α-Glycylglycine*	121.1	124.2	114.4	119.3	—	—
N,N′-Diglycyl-cystine*	120.6	125.3	113.2	121.6	—	—
Selected Values	121	125	114	123	123	114

* Three-dimensional refinement.
\+ Polarized infrared study (Newman & Badger, 1951).

The selected values, $N - C = 1.47$ A, $C' - C = 1.53$ A, $C' - O = 1.24$ A, and $C' - N = 1.32$ A, can be accepted as valid to about ± 0.02 A for any peptide.

The values of bond angles for the amide group, as determined experimentally, are given in Table 2. The selected values, which range between 114° and 125°, are not reliable to more than about $\pm 2°$. Inasmuch as 2° in a length of 1.4 A corresponds to an uncertainty of 0.05 A in position, the bond angles constitute a larger source of error in predicted dimensions of polypeptide chains than do the bond distances (44).

The planarity of the amide group was indicated by theoretical considerations twenty years ago, when it was recognized that an amide can be described as a conjugating structure, and that both the $C' - O$ bond and the $C' - N$ bond have a significant amount of double-bond character. The observed interatomic distances correspond to about 60 percent and 40 percent double-bond character in these two positions, respectively. All of the amides that have been subjected to precise crystal-structure investigation in our laboratory have been found to be planar. In β-glycylglycine and acetylglycine the planarity is preserved to within about 0.03 A.

In α-glycylglycine the carbonyl-carbon atom and the three atoms bonded to it are within 0.03 A of a single plane, but the α-carbon atom bonded to the nitrogen atom is in a position corresponding to rotation by about 5° out of the plane. A rotation of about the same amount is found in N,N′-diglycylcystine. The rotations correspond to a strain energy of approximately 0.2 kcal mole^{-1}, as calculated by the equation given below. In the structure reported for the sodium and potassium benzyl penicillins ([45]) the amide group is planar to within the rather large limits of error of the determination. In the recently reported structure of cysteyl glycine sodium iodide ([42]) the amide group is nonplanar, in amount corresponding to rotation ot more than 40° around the C′ — N bond; it is likely that the atomic positions as reported are not correct.

In all of the simple substances except diketopiperazine the amide group has been found to have the trans configuration. It is not

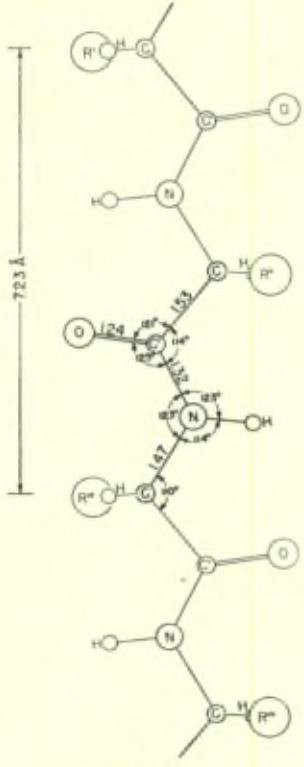


Fig. 1. — Interatomic distances and bond angles in the polypeptide chain.

unlikely, as emphasized by Mizushima and Shimanouchi [46], that the cis configuration is considerably less stable than the trans configuration for an amide group.

A discussion has been presented of the stability of polypeptides as a function of azimuthal orientation around the single bonds between the atoms N and C′ of the amide group and the α-carbon atom [47]. It is probable that some azimuthal orientations are favored over others, but at the present time the difference in energy corresponding to different orientations is not known, and in any case it is probably small.

The dimensions of the polypeptide chain, as predicted from the studies on simple substances, are shown in Figure 1.

The most important structural feature of folded polypeptide chains is the hydrogen bond between the N — H group and the carbonyl oxygen atom of the amide groups. It is found by experiment that in most simple substances the nitrogen atoms and oxygen atoms lie in such positions that it is reasonable to assume that hydrogen bonds are present, and in the case of a few crystals the electron distribution as calculated by the Fourier series method from the observed intensities of reflection has clearly shown the presence of hydrogen atoms on the line between a nitrogen atom and an oxygen atom.

Dimensions of hydrogen bonds found in crystals of amino acids and peptides are given in Table 3. Most of the hydrogen bonds have the length N — H···O equal to 2.79 ± 0.12 A; in a few cases the observed length lies outside of this range. Because of uncertainty about the position of the hydrogen atom, the value of the N — H···O angle is not listed in Table 3. Instead, values are given for the angle C — N···O. If the N — H bond lies at 110° from the C — N bond, in the plane of C — N···O, the deviation of this angle from 110° will be equal to the angle of bending of the hydrogen bond, N — H···O. It is seen from Table 3 that most of the hydrogen bonds involve a bend less than 12°, and that the exceptional hydrogen bonds, with larger bends, are in most cases (all except N,N′-diglycylcystine) the abnormally long hydrogen bonds. The largest angle of bending of the hydrogen bond in these crystals is 25.5° [48].

A general principle of importance in the formulation of configurations of polypeptide chains is the principle of closest packing of atoms. The electronic van der Waals attraction of all atoms for

TABLE 3

Dimensions of N — H...O Hydrogen Bonds Found in Crystals of Amino Acids and Peptides.

CRYSTAL	N — H ··· O	< C — N ··· O	
		Angle Found	Deviation from 110°
	Å	Degrees	Degrees
DL-Alanine*	2.80	105	5
	2.84	103	7
	2.88	116	6
L-Threonine*	2.80	98	12
	2.90	116	6
	3.10	132	22
DL-Serine*	2.79	99	11
	2.81	98	12
	2.87	121	11
Hydroxy-L-proline*	2.69	102, 113,	8.3,
	3.17	81, 133	19, 23
N-Acetylglycine*	3.03	132	22
β-Glycylglycine	2.68	100	10
	2.80	115	5
	2.81	99	11
	3.07	131	21
α-Glycylglycine*	2.67	118	8
	2.77	114	4
	2.67	88	22
	2.75	117	7
N,N′-Diglycylcystine*	2.75	84.5	25.5
	2.89	129.8	19.8
	2.75	111.8	1.8
Most probable value.	2.79 ± 0.12 Å		

* Three-dimensional analysis.

one another serves effectively to stabilize structures in which atoms are surrounded as closely as possible by other atoms. A configuration for a polypeptide chain which leaves any large volume unoccupied by atoms is consequently less likely to provide a stable structure for a protein than a configuration in which the atoms are closely packed.

It is possible to give a rough quantitative consideration to the structural principles described in the preceding paragraphs. Let us ask how great the deviation from the favored structural parameters may be without causing instability amounting to more than 0.1 kcal $mole^{-1}$.

With use of the known values of the force constants of bonds, we find that this amount of strain energy corresponds to a stretching or compression of the single bonds to the α-carbon atom by 0.02 A, and of the conjugated bonds C′ — O and C′ — N of 0.01 A.

The force constants for bending bonds are somewhat smaller (when expressed in corresponding units), and change in atomic positions by 0.02 A to 0.04 A, corresponding to 2° or 3° change in bond angle, may be produced with a strain energy of 0.1 kcal $mole^{-1}$.

The planarity of the amide group may be given a theroretical discussion. The resonance energy stabilizing the planar amide group is about 30 kcal $mole^{-1}$. If the planes of the two ends of the amide group form a dihedral angle δ, and if A is the resonance energy of the amide group for the planar configuration, the strain energy may be taken as equal to A $\sin^2 \delta$. From this expression we calculate that a distortion of the amide group by 3° from the planar configuration corresponds to strain energy 0.1 kcal $mole^{-1}$.

The average length N — H···O of the hydrogen bond is 2.79 A. There is no reliable experimental value for the energy of stretching or compressing the N — H···O bond. However, the value for the O — H···O hydrogen bond, which has length 2.76 A, can probably be applied without great error. The compressibility of ice, 12×10^{-6} cm^2 kg^{-1}, corresponds to a strain energy of 0.1 kcal $mole^{-1}$ for a stretching or compression of the hydrogen bond by 0.09 A, and hence about 0.09 A can be taken as the change in length of the N — H···O hydrogen bond that would correspond to 0.1 kcal $mole^{-1}$ strain energy. It must, of course, be recognized that changes of this order of magnitude in the length of the N — H···O hydrogen bond may also result from change in the electric charge or other characteristics of the atoms involved, through interaction with other atoms. In particular, the formation of a hydrogen bond by the oxygen atom of an amide group increases the hydrogen bond-forming power of the hydrogen atom of the same amide group.

It is difficult to estimate the strain energy of bending a hydrogen

bond. If it is assumed that a lateral displacement of an atom is associated with as much strain energy as one half the displacement of the atom along the bond, then the strain energy 0.1 kcal mole^{-1} would correspond to a bending of the hydrogen bond (change of the angle N — H···O from the value 180°) by about 6°.

A comparison of two structures with respect to close packing can be made by the calculation of the electronic van der Waals energy. From the equation

$$E = -\Sigma\, 38\, \frac{R_A R_B}{r_{AB}6} \text{ kcal mole}^{-1}$$

for the energy of the London electronic dispersion interaction, the value of the van der Waals energy can be roughly calculated for any structure in which the interatomic distances r_{AB} are known. Here the summation is to be made over all pairs of atoms AB. The quantities R_A and R_B represent the mole refraction of the atoms or groups A and B. This equation has been obtained from the approximate second-order perturbation theory with use of the average value 14 ev for excitation energy of the groups. The mole refraction R of the amide group including the α carbon atom can be taken as 13 cm^3. When this equation is applied to the γ helix (described below), for example, in which there is a cylindrical hole about 2 A in diameter, it is found that the instability corresponding to the decreased van der Waals attraction resulting from the presence of the hole amounts to about 4 kcal mole^{-1} per residue, in comparison with the α helix ([49]).

4. THE PREDICTION OF CONFIGURATIONS OF POLYPEPTIDE CHAINS

Until two years ago the configurations of polypeptide chains in proteins that had been proposed were described by their proponents only qualitatively, and no serious effort had been made to account for the intensities of x-ray diffraction by proteins in terms of the proposed structures. The first predictions about the precise configuration of polypeptide chains were made by Corey ([50]); his values were slightly revised by Corey and Donohue ([51]), and have since been subjected to further small revisions, leading to the values given in the preceding section. In 1950 atomic coordinates were

reported for a large number of configurations of polypeptide chains by Bragg, Kendrew, and Perutz ([52]), and the coordinates were used in the discussion of the x-ray data for hemoglobin and myoglobin. At the same time some precisely described configurations were formulated by Pauling and Corey, in part with the help of Dr. H.R. Branson. In this work, which has led to the discovery of a number of structures that had not previously been formulated, great emphasis was laid on rigorous adherence to the fundamental structural principles that had been found through the studies of simple substances and through application of the quantum mechanical theory of the chemical bond. In particular, the requirements that the amide group be planar and that the CO and NH groups be involved in hydrogen-bond formation were rigorously applied. In this way two helical structures, the α helix, with 3.7 residues per turn, and the γ helix, with 5.2 residues per turn, were discovered; a third helix, with 4.4 residues per turn, has since been described. The α helix seems to be the most widely occurring configuration of polypeptide chains in proteins. Several pleated sheets, involving polypeptide chains that are not quite fully extended and that form lateral hydrogen bonds with adjacent chains, have also been discovered; of these the antiparallel-chain pleated sheet seems to be present in silk fibroin, and the parallel-chain pleated sheet in the proteins of the β-keratin class. A complex structure, involving three polypeptide chains twisted about one another, has also been suggested for collagen and gelatin.

5. THE α HELIX

The most general configurations of a polypeptide chain composed of equivalent L amino acid residues are helical configurations. There are three ways in which a single polypeptide chain with planar amide groups can be coiled into a helix, and form hydrogen bonds involving little or no strain. One of these helical structures, the α helix, is much more compact than the other two, and is probably favored over them in a significant way by its greater van der Waals stabilization. The α helix has about 3.7 residues per turn; it is represented diagrammatically in Figure 2, which shows the hydrogen bonds between each amide group and the two amide groups further along the chain in either direction, and by drawings in Figures 3 and 4.

The γ helix, with about 5.2 residues per turn, has the appearance shown in the drawing of Figure 5. The γ helix seems to be essentially equivalent to the α helix with respect to planarity of the amide group and linearity of the hydrogen bond. However, it has a hole 2 A in diameter along its axis, and it seems likely that the instability

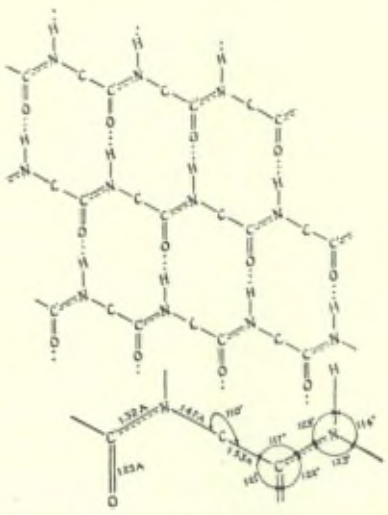


Fig. 2. — Diagrammatic representation of the α helix. The end of the polypeptide chain in each segment at the left is to be continued as the next segment, beginning at the right. Some of the bond angles shown at the bottom of the figure are incorrect.

Fig. 3. — The α helix is here correctly represented as a left-handed screw, but the small circles, representing hydrogen atoms, and large circles R, representing the side chains, are to be interchanged on the α-carbon atom to reproduce correctly the configuration of amino-acid residues in proteins.

resulting from the decrease in van der Waals energy of stabilization, mentioned in the preceding section, is great enough to prevent the γ helix from constituting a significant structural feature of proteins.

Another helix, with about 4.4 residues per turn, has been described recently ([53]). It is closely similar to the α helix, but involves the

formation of hydrogen bonds with the fourth amide group in the chain, rather than with the third. There is a hole about 1 A in diameter along the axis of this helix, and in addition some strain is produced in the formation of hydrogen bonds. The quantitative discussion of deformation energy given in an earlier section indicates

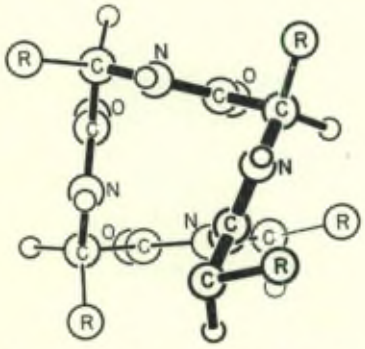


Fig. 4. — Projection of the α helix onto a plane perpendicular to the helical axis.

Fig. 5. — The γ helix, with about 5.2 residues per turn.

that the 4.4-residue helix is less stable than the α helix by about 1 kcal mole^{-1}, and it is accordingly likely that the 4.4-residue helix also is not an important structural feature of proteins.

The α helix was first predicted (54) to have 3.69 residues per turn and length per residue of 1.47 A along the fiber axis, leading to a fiber-axis length per turn of 5.44 A. Later it was pointed out (55)

that the fiber-axis length per residue 1.47 A corresponds to the hydrogen-bond distance 2.75 A, and that the length per residue increases by 0.01 A for every 0.03 A increase in the hydrogen-bond distance. The reasonable range 2.68 A to 2.92 A for the hydrogen-bond length would then correspond to the range 1.45 A to 1.53 A for the fiber-axis length per residue.

The number of residues per turn is fixed primarily by the bond angle at the α-carbon atom; it varies from 3.60 for bond angle 108.9° to 3.67 for bond angle 110.8°. These ratios correspond respectively to 18 residues in 5 turns and 11 residues in 3 turns. The minimum values and the maximum values just quoted lead to the values 5.22 A and 5.62 A, respectively, for the pitch of the helix.

TABLE 4

Atomic Coordinates for the 18-Residue
5-Turn α Helix
x, y, z, and ρ in A

Atom	x	y	z	ρ	φ
C	0.00	0.00	1.48	2.28	0.0°
N	1.19	0.00	0.62	1.57	29.1°
C′	2.42	0.00	1.09	1.61	74.6°
O	2.72	0.00	2.29	1.76	83.5°
C*	3.50	0.00	0.00	2.28	100.0°
βC_1	— 1.26	0.30	0.66	3.23	— 18.9°
βC_2	— 0.15	— 1.40	2.12	3.29	14.7°
Axis	1.75	1.47	—	0.00	—

Atomic coordinates for the α helix are given in Table 4. These coordinates differ very slightly from those previously published; they have been recalculated with use of the slightly revised values of interatomic distances and bond angles given in Tables 1 and 2.

Two alternative positions are listed for the β carbon atom. For a polypeptide chain in the form of a left-handed screw, as shown in Figure 3, and with the absolute configuration of amino acids as determined recently by Bijvoet, Peerdeman, and Van Bommel (56), position 1 for the β carbon atoms corresponds to a polypeptide

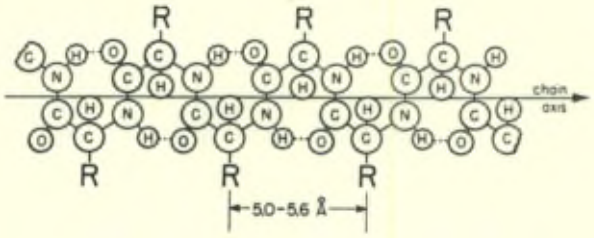
composed of D amino-acid residues, and position 2 to a polypeptide composed of L amino-acid residues. With a right-handed screw position 1 corresponds to L amino-acid residues and position 2 to D amino-acid residues. It is mentioned later (section 10) that Riley and Arndt have obtained evidence that in proteins containing the α helix the β carbon atoms are in position 2. Inasmuch as these proteins are composed of L amino-acid residues, the α helix must have the configuration of a left-handed screw, as shown in Figure 3. Figure 3 would accordingly represent the structure of these proteins if the side chains were shown in the positions represented by the small circles, rather than the large circles, adjacent to the α carbon atoms.

The average residue weight in most proteins is about 1.05, and the density is about 1.33 g cm^{-3}. Accordingly the volume per residue is about 131 A^3. If the axial length per residue is taken as 1.47 A, the cross-sectional area becomes 89 A^2. This corresponds to a diameter of 10.6 A. Accordingly we predict that the polypeptide chains of proteins with the configuration of the α helix can be described as having approximately the shape of circular cylinders, with diameter about 10.5 A.

6. SYNTHETIC POLYPEPTIDES WITH THE α STRUCTURE

The first strong evidence in support of the α helix was provided by a synthetic polypeptide, poly-γ-methyl-L-glutamate.

In 1948 it was reported by Astbury and coworkers [57] that a number of synthetic polypeptides give x-ray diagrams similar to those of the α-keratin proteins. Some oriented films and fibers of synthetic polypeptides were prepared and examined by x-ray diffraction by Bamford, Hanby, and Happey [58], and the oriented films were studied with polarized infrared radiation by Ambrose and Elliott [59]. These investigators proposed as the structure of polypeptide the α_{II} structure shown in Figure 6. This structure had originally been proposed by Huggins [60], for the α-keratin proteins, and has been discussed by Zahn [61], Simanouti and Mizushima [62], and Ambrose and Hanby [63]. The structure must be considered unsatisfactory, in that when it is constructed with

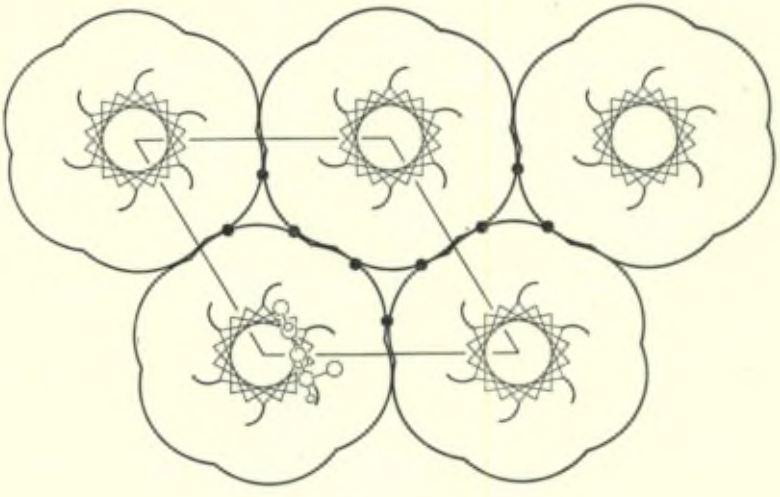


The Proposed α_{II} Structure of the Polypeptide Chain

Fig. 6. — The proposed α_{II} structure for the α-keratin proteins and α polypeptides.

planar or nearly planar amide groups the N — H · · · O bond is bent by about 60°. Atomic coordinates have not been given for this structure by any of its supporters (64).

We noticed that the x-ray data for poly-γ-methyl-L-glutamate reported by Bamford, Hanby, and Happey were compatible with a hexagonal unit of structure with the dimensions indicated for parallel packing of cylindrical molecules with the structure of the α helix. A plan of the proposed structure is shown in Figure 7.



18-Residue 5-Turn Helical Structure
of Poly-γ-methyl-L-glutamate

Fig. 7. — A plan of the proposed arrangement of α helixes in one form of poly--γ-methyl-L-glutamate.

It was found by calculation that the observed intensities could be roughly accounted for.

Almost completely convincing evidence for the correctness of a α helix was obtained by Perutz [65], who observed a strong meridional reflection at 1.50 A on x-ray photographs of poly-γ-methyl-L-glutamate and poly-γ-benzyl-L-glutamate. Perutz [66] has described the α helix in the following words : « If the α helix is compared to a spiral staircase with the residues as steps, then the height of each step is 1.5 A and the height of each turn 5.4 A, making 3.6 steps per turn. It takes eighteen steps or five turns until a step is found exactly in a vertical line above the starting point. Hence the true ' repeat ' pattern is $18 \times 1.5 = 27$ A... The 3.6-residue helix, if present in all proteins and polypeptides of the α type, should give rise to a reflection at 1.5 A spacing from planes perpendicular to the fiber axis, corresponding to the axial repeat of residues along the chain. »

A meridional reflection at 1.49 A had been reported for porcupine quill by MacArthur [18], and the presence of the reflection, as well as of reflections corresponding to multiples of this spacing, had been quoted by Pauling and Corey in support of the α helix [67], but it remained for Perutz to point out the crucial character of the 1.5 A reflection, and to show experimentally that it is present in the x-ray diffraction pattern of many polypeptides and proteins.

Strong meridional reflections at 1.47 A to 1.50 A have been reported also for several other α polypeptides, including copolymers of different amino acids [68].

Additional evidence substantiating a helical structure with about 18 residues in five turns of the helix has been obtained by Cochran, Crick, and Vand [69], who calculated the theoretical expression for the intensity of x-ray scattering by a helix. The theory leads to the prediction of relative intensities of the different layer lines in the x-ray photograph of a crystal containing helical molecules, and the theoretical expression has been found to be in good agreement with the observed patterns [68].

Bamford, Brown, et al. [70] have recently reported their x-ray data on another modification of poly-γ-L-methyl-glutamate, which has an identity distance along the fiber axis of 43.5 A, corresponding to 29 amino-acid residues in eight turns of the helix. In addition to the hexagonal or closely pseudo-hexagonal modification first described by Bamford, Hanby, and Happey [58] in 1951, and also

investigated by Dr. Harry L. Yakel, Jr., Dr. Yakel has prepared a third modification, which, like the hexagonal fibers, has 18 residues in five turns, but which is orthorhombic in symmetry, and shows rather large deviations from hexagonal packing, the intermolecular distances being about 10.5 A and 13.5 A. The 1.5 A reflection is measured at 1.49 A for this modification, and also for the hexagonal or closely pseudo-hexagonal fibers, whereas Bamford, et al., report 1.50 A. It would not be surprising if the fibers of this rather complex material could crystallize in even more than three modifications.

The synthetic polypeptides can be converted into another modification, which gives x-ray diagrams similar to those of the β-keratin proteins ([58],[59]). It is probable that the structure of the polypeptides in this form is that of the pleated sheets, described in section 11.

7. COMPOUND HELICAL CONFIGURATIONS OF POLYPEPTIDE CHAINS

The general similarity of the x-ray diagrams of the α-keratin proteins and synthetic polypeptides such as poly-γ-methyl-L-glutamate suggests strongly that the configurations of polypeptide chains in these materials are essentially the same. Evidence supporting the suggestion that the α-helix is present as the principal structural feature of the α-keratin proteins was summarized by Pauling and Corey ([67]). Strong support for the proposal was provided by the work of Perutz, quoted above, who found the 1.5 A reflection to be present, with spacing 1.50 A, on x-ray photographs of myosin, hair, and muscle; he has also reported ([66]) that Astbury has found this reflection, with spacing 1.49 A, on photographs of epidermin, tropomyosin, fibrin, and bacterial flagella.

There remained, however, some puzzling difficulties. One of these is the presence on the photographs of a strong reflection with spacing about 5.15 A, which seems to be a meridional reflection. Pauling and Corey had attempted to account for this reflection, on the basis of a theoretical calculation of the x-ray form factor for the helix, as an off-meridional reflection, resulting from the cooperation of successive turns of the helix; careful experimental studies indicate strongly, however, that the reflection is a true

meridional reflection, and such a reflection cannot be explained by polypeptide chains with the configuration of the α helix and oriented parallel to the fiber axis.

It seems likely that the solution of these difficulties is provided by the assignment to the α-keratin proteins of a somewhat more complex structure ([71], [72]). An α helix for a polypeptide chain involving repeating sequences of amino-acid residues of different kinds would not be expected to have a straight axis; instead, the axis of the helix would itself be predicted to pursue a helical course. Change in the nature of side-chain groups of amino acids adjacent to one another in the polypeptide chain or on successive turns of the helix might cause the hydrogen-bond distance to vary by 0.1 or 0.2 A about its average value, 2.79 A. This variation might be caused directly by the interaction of the side-chains with the carbonyl and imino groups of the amide groups, or indirectly by steric hindrance or by attraction between side-chains. As an example, we may consider an α helix composed of a polypeptide in which a unit of four amino-acid residues of different sorts is repeated. Two of the hydrogen bonds might be longer than the other two, by about 0.2 A. This difference in length would cause a curvature of the axis of the α helix. If the α helix has 3.67 residues per turn, the normal to the curved helical axis would be rotated by about 0.09 revolution, from one unit to the next unit of four residues along the chain. This corresponds to a complete revolution in about 11 units. The axis of the α-helix would itself describe a larger helix, with pitch approximately 66 A, the axial length of 44 residues. The radius of the larger helix would be about 1.5 A, and the sense of the larger helix would be the same as that of the α helix — it would be a left-handed screw, in the case of proteins. A complex helix of this sort is represented in Figure 8.

Another case is that of the compound helix with a repeating unit of seven amino-acid residues. An α helix with 3.60 residues per turn executes 97.2% of two turns in 7 residues, and would accordingly be expected to complete a turn of the larger helix in about 35 turns of the α helix, corresponding to about 200 A for the pitch of the helix. The radius of the large helix might easily be as great as 10 A, with the variation in hydrogen-bond length mentioned above. The sense of the large helix of the 7-residue compound α helix is opposite to that of the α helix itself: it is a righthanded screw, in the case of proteins.

A radius of 6 A for the large helix would permit three compound helixes to twist about one another, to form a three-strand rope, as shown in Figure 9. Such a rope with the sense of its twist opposite to that of the strands (and hence right-handed for proteins) would be formed by the seven-residue compound helix, or with the sense of twist the same as that of the strands (left-handed for proteins, as shown in Figure 9) by, for example, the fifteen-residue compound helix (with a repeating unit of fifteen residues, comprising nearly four turns of the α helix).

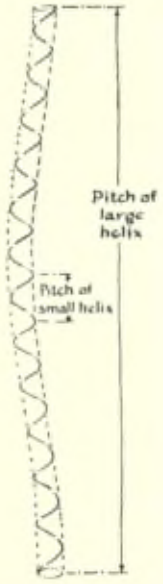


Fig. 8. — A representation of a compound helical configuration of polypeptide chains, with the axis of the α helix describing a helical course.

Fig. 9. — A compound α helix (at the left), a 7-strand α cable (in the center), and a 3-strand α rope (at the right).

Six compound helixes with radius of the large helix equal to 10 A could twist about a central α helix, to form a seven-strand cable, as shown in figure 9. The repeating unit of seven amino-acid residues, comprising nearly two turns of the α helix, is the simplest one that would give rise to compound helixes suitable to a seven-strand cable. The sense of the cable would be right-handed, in the case of proteins, as shown in the figure.

The presence of proline or hydroxyproline residues in the polypeptide chain would lead to a pronounced curvature of the α helix,

as a result of the inability of the residues to form a hydrogen bond involving the amide nitrogen atom, and also of the steric hindrance of the adjacent carbon atom of the proline ring and the atoms of the next turn of the helix.

8. THE STRUCTURE OF HAIR, HORN AND OTHER α-KERATIN PROTEINS

The evidence supporting the original assignment of the α helix to the α-keratin proteins included the indication of the x-ray diagram that the structure involves cylindrical molecules of diameter about 10.5 A, in approximately parallel orientation, the rough agreement of observed and calculated intensities of reflection in the equator, and the very rough approximation of the 5.15-A meridional reflection to the length per turn, predicted to be about 5.4 A. Additional evidence was provided by the observation of the 1.5-A meridional reflection, the interpretation of which, as an axial length per residue, is especially straightforward. In addition, Riley and Arndt have found that the powder diagram of hair, corresponding to the radial distribution function, is that of the α helix (unpublished result).

It now seems likely that the α-keratin proteins that give the 5.15-A meridional reflection have the structure indicated in Figure 10. The seven-strand α cable is about 30 A in diameter. A fiber containing

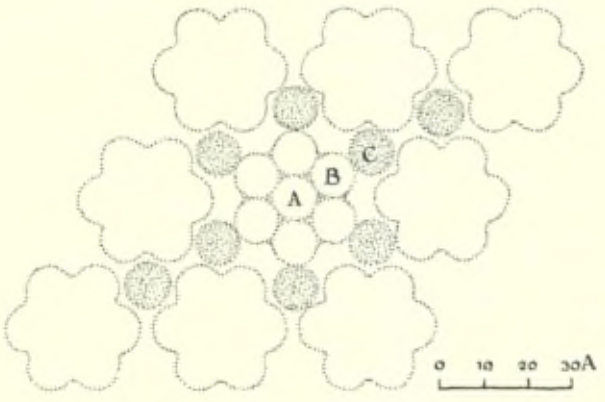


Fig. 10. — Cross section of the proposed structure for the α-keratin proteins. The structure consists of 7-strand α cables, AB_6, with single compound α helixes, C, in the interstices.

these cables in parallel orientation would have a hexagonal pseudo-unit of structure with a = approximately to 30 A. Hair, horn, and similar proteins give a moderately strong equatorial reflection with spacing 27 A; if this is interpreted as the 10.0 reflection, the value of a becomes 32.4 A. With these dimensions, there is room enough between the AB_6 cables to permit the introduction of other, single α helixes, C, as shown in figure 10. The α helixes C could fit between the strands B of the seven-strand cable; in order to fit easily, they themselves would have to have the form of a compound helix, with pitch one-sixth as great as the lead of the compound helixes B. It is probable that the lead of the helixes B (the axial length required for a complete turn) is about 400 A, as indicated by the large-spacing reflections on the x-ray diagrams of the α-keratin proteins and by electron micrography. The pitch of the helixes C would then be about 67 A. In order to fit into the interstices, these compound helixes C would be left-handed, the cables AB_6 being right-handed.

This structure, involving nine α helixes in a hexagonal pseudo-unit with a = 32.4 A and with axial length per residue 1.50 A, accounts satisfactorily for the observed density of the α-keratin proteins. Moreover, it provides an explanation of the 5.15-A meridional reflection. The B helixes of the AB_6 cable are assumed to involve a repeating unit of seven amino-acid residues. The axial length of this repeating unit is calculated to be 10.36 A. It should give rise to a very weak first-order reflection, and to a strong second-order reflection, involving reinforcement of two turns of the α helix in the repeating unit of seven residues. The predicted spacing of the strong reflection is 5.18 A, in satisfactory agreement with the observed spacing.

The structure shown in Figure 10 also accounts for the intensities of reflection in the equatorial region of the x-ray diagram.

It is probable that hair and horn can be fractionated into AB_6 cables and the interstitial C α helixes. These fractions, when spun into fibers, should give x-ray diagrams of different kinds. The AB_6 cables would give diagrams rather closely similar to those of the α-keratin proteins themselves, with, in particular, the 5.15-A meridional reflection. The individual α helixes C, however, should produce x-ray diagrams similar to those of the synthetic polypeptides, and without the 5.15-A meridional reflection. It seems not unlikely that actomyosin has the structure indicated in Figure 10, and that

its fractionation into myosin and actin is, in fact, a separation of the AB_6 cables from the interstitial α helixes. Support of this suggestion is provided by the fact that myosin, like actomyosin, gives an x-ray diagram showing the 5.15-A meridional reflection, whereas this reflection is missing on the diagrams of actin.

9. THE STRUCTURE OF FEATHER RACHIS

Feather rachis has usually been described [73] as belonging to the β-keratin class, with, however, an unusually short fiber-axis residue length, 3.07 A, in place of 3.3 A. Pauling and Corey proposed a structure involving double rows of α helixes alternating with a pleated-sheet layer [74]. Although some evidence, from x-ray data, was advanced in support of this proposal, it was far from convincing. Another structure, somewhat similar in nature, can be formulated. The x-ray evidence in support of the new structure must be given some weight, but it is not as yet great enough to prove that the structure is correct.

The x-ray diagram of feather rachis was interpreted by Pauling and Corey in terms of a pseudo-orthorhombic unit with $a_0 = 9.50$ A, $b_0 = 34.2$ A, and $c_0 = 94.6$ A. It has been observed that the diagram can be indexed in terms of a hexagonal unit, the 34.2-A reflection being the reflection 11·0. This hexagonal unit has $a_0 = 39.5$ A. Its cross-sectional area is 1.485 times that of the α-keratin protein; that is, in the ratio of 13.4 : 9. A possible structure for 13 α helixes is the hexagonal packing of right-handed AB_6 cables, with the interstices occupied by left-handed 3-strand α ropes. The identity distance along the c axis for this structure is $3 \times 94.6 = 283.8$ A. The x-ray diagram of feather rachis corresponds strikingly to that predicted for this structure with the use of the theory of Cochran, Crick, and Vand. Further investigation will be necessary, however, before the structure can be accepted as proved.

10. HEMOGLOBIN AND OTHER GLOBULAR PROTEINS

The early detailed investigation of crystals of insulin by Bernal and Crowfoot [20, 21] showed that extensive data on the diffraction patterns of globular proteins can be collected, but it was not found possible to interpret the data in such a way as to

provide significant information about the configuration of the polypeptide chains in the molecules of insulin. Bragg, Perutz, and Kendrew, and their collaborators, obtained significant information from their study of hemoglobin and myoglobin ([22], [23], [24], [52]). They were able to report that the polypeptide chains in these molecules are coiled into nearly cylindrical rods, about 10.5 A in diameter. These rods extend in the 57-A direction of the molecules, and they are packed in a pseudo-hexagonal array. The arrangement of atoms in the rods corresponds to an approximate repeat of about 5.5 A ([52]).

Identification of the cylindrical rods in hemoglobin and myoglobin with the configuration of the α helix was made by Pauling and Corey, on the basis of the radial distribution function ([75]). Perutz ([23]) had published a three-dimensional Patterson diagram for horse carbonmonoxyhemoglobin, calculated from the observed intensities of x-ray reflection. This diagram was converted into a radial distribution diagram by Pauling and Corey, and was compared with radial distribution functions calculated for the 5.2-residue helix (the γ helix) and also for the α helix. There was striking disagreement between the experimental function and the functions calculated for the γ helix, but reasonably good agreement between the experimental function and the functions calculated for the α helix. The experimental function was not sufficiently reliable to permit a distinction to be made between the left-handed α helix (β-carbon atoms in position 2) and the right-handed α helix (β-carbon atoms in position 1). In addition, a comparison was made between the calculated and observed Patterson vector components in the direction of the axis of the rods. The experimental function, calculated by Bragg, Kendrew, and Perutz for hemoglobin ([52]), shows a succession of maxima at vector lengths equal to integral multiples of 5.5 A. This agrees very well with the theoretical function calculated for the α helix ([49]).

There is evidence from chemical properties ([76]) and from the x-ray data ([77]) that the segments of the α helix in the hemoglobin molecule do not extend throughout the 57-A length of the molecule, but are instead split by the hemes into two parts, with average length about 25 A. Since there are about forty of these 25-A segments of the α helix (each composed of about five turns of the helix) in the hemoglobin molecule, and there are only five (or six) polypeptide chains, the individual segments must be linked toge-

ther by sections of the polypeptide. The molecule may accordingly be described as composed of polypeptide chains, which execute about five turns with the configuration of the α helix, and then twist about to form another section of the α helix.

An important contribution to our knowledge of the structure of globular proteins has been made by Riley and Arndt. They carried out a careful experimental determination of the radial distribution function for amorphous bovine serum albumin, prepared by the freeze-dry method ([78]). The experimental radial distribution function calculated from their x-ray measurements was found to resemble very closely the radial distribution function calculated for the α helix, with β-carbon atoms in position 2. The close approximation of the experimental function to the calculated function leaves little doubt of the correctness of the assignment of the configuration of the α helix to polypeptide chains in this globular protein. The fact that agreement is obtained with β-carbon atoms in position 2 rather than in position 1 requires that the α helix have the sense of a left-handed screw. In later work ([79]) they have verified the presence of the left-handed α helix in hemoglobin, and have shown also that the same configuration represents the principal structural feature of a number of other globular proteins, including ovalbumin, lysozyme, pepsin, insulin, and edestin. At the present time the left-handed α helix has been found to be the principal structural feature of every globular protein that has been investigated sufficiently well to permit a decision to be made as to the configuration of its polypeptide chains.

11. THE PLEATED SHEETS

It was mentioned in section 1 that during recent years silk fibroin and the proteins with β-keratin structure, such as stretched hair, had been assumed to have the configuration indicated in Figure 11, involving polypeptide chains stretched out to their maximum length, and forming lateral hydrogen bonds. It was pointed out in 1951 by Pauling and Corey ([80]) that the structural parameters are such that this planar sheet is unsatisfactory for proteins and polypeptides other than polyglycine, in that the groups CHR in adjacent chains are held so close to one another as to give rise to serious steric hindrance. A structure of the sort shown in Figure 12, involving

the bending of the polypeptide chains so that the α carbon atoms lie alternately above and below the plane of the paper, was formulated. This structure, called the polar pleated sheet, was suggested as the structure of the β-keratin proteins, but was later replaced by the two pleated sheets that we shall now describle.

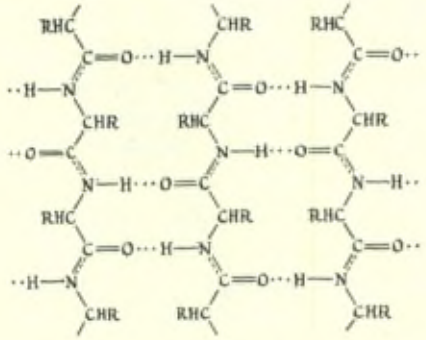


Fig. 11. — Diagrammatic representation of a hydrogen-bonded layer structure of completely extended polypeptide chains.

In the polar pleated sheet successive residues in polypeptide chains are related to one another by a reflection by a glide plane. This causes the side chains on one side of the sheet to be different in environment from those on the other side of the sheet. If the polypeptide chain involves a two-fold screw axis, instead of a glide plane, the structures shown diagrammatically in Figures 13 and 14

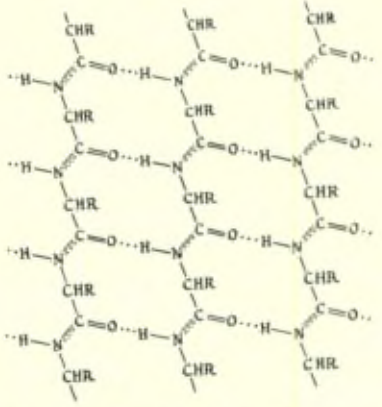


Fig. 12. — Diagrammatic representation of a hydrogen-bonded layer structure of polypeptide chains with residues similarly oriented (the polar pleated sheet).

and represented by drawings in Figures 15 and 16 are obtained (47). The structure shown in figures 13 and 15, the antiparallel-chain pleated sheet, involves polypeptide chains which alternate in direction along the fiber axis, whereas that shown in figures 14 and 16, the parallel-chain pleated sheet, involves polypeptide chains with the same orientation.

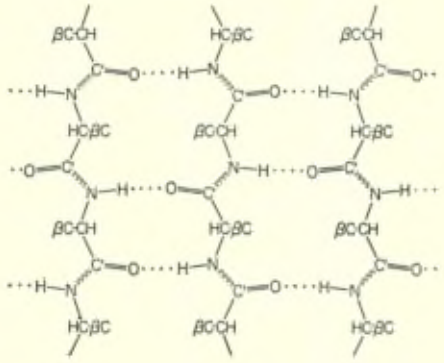


Fig. 13. — Diagrammatic representation of the antiparallel-chain pleated-sheet structure.

These two pleated-sheet structures were originally discovered in the course of an investigation of the consequences of the assumption that certain orientations about the single bonds to the α carbon atom are favored over other orientations (47). This assumption led to the conclusion that the identity distance in the fiber-axis

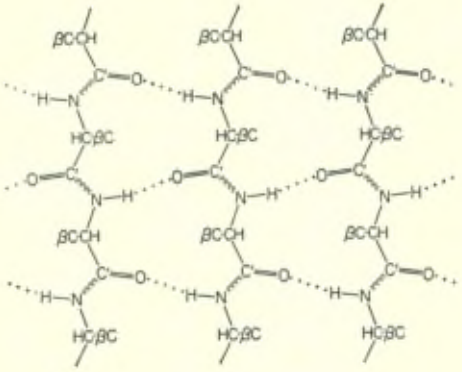


Fig. 14. — Diagrammatic representation of the parallel-chain pleated-sheet structure.

direction should be 6.68 A, corresponding to 3.34 A per residue. Since then a study has been made of the consequences of the postulate that it is more important that the hydrogen bond be linear than that favored orientations about the single bonds to the

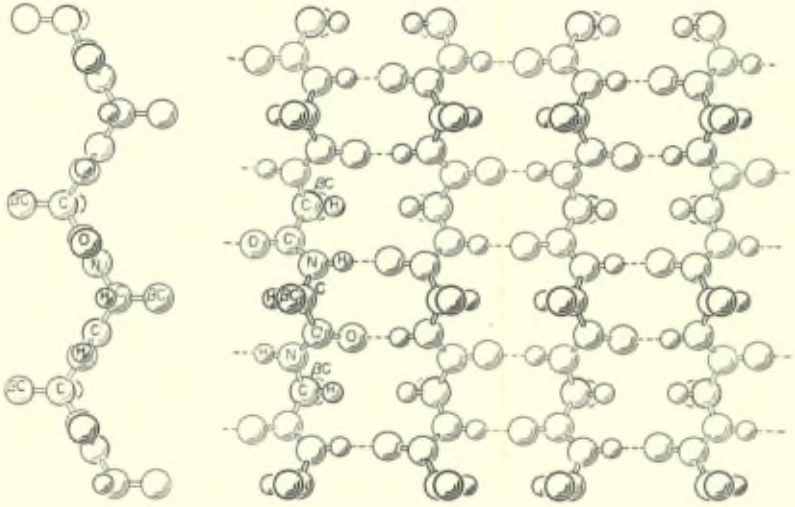


Fig. 15. — Drawing representing the antiparallel-chain pleated-sheet structure.

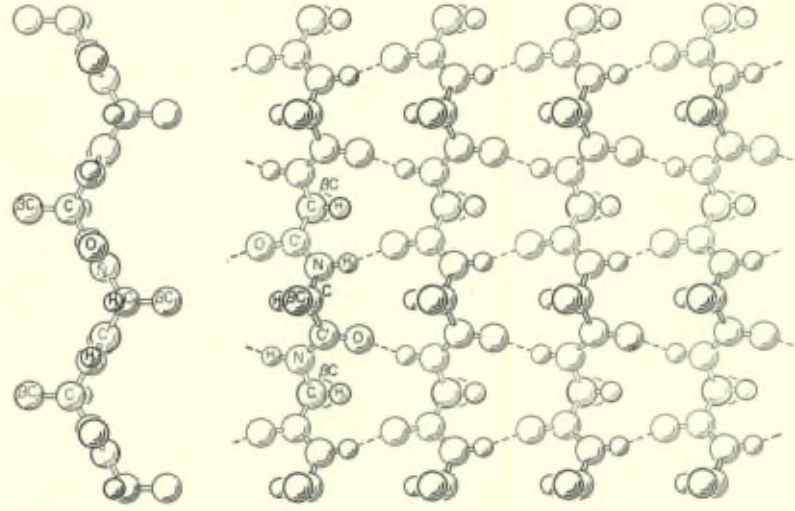


Fig. 16. — Drawing representing the parallel-chain pleated-sheet structure.

α carbon atoms be assumed. It has been found that the postulate of linearity of the hydrogen bonds leads to the predicted value 7.00 A for b_0 for the antiparallel-chain pleated sheet, and 6.50 A for the parallel-chain pleated sheet.

The drawings shown in Figures 15 and 16 correspond to the D configuration of polypeptide chains. In order to represent proteins, these drawings should be replaced by their mirror images.

Two other structures, which may be called the antiparallel-chain rippled sheet and the parallel-chain rippled sheet, may be constructed by reflecting alternate chains in figures 15 and 16 in the plane of the paper. They are represented topologically by the same diagrams. Figures 13 and 14, as the corresponding pleated sheets. The rippled-sheet structures are satisfactory for an equimolal mixture of D polypeptides and L polypeptides, or for polyglycine. The predicted values of the fiber-axis identity distance are the same as for the corresponding pleated sheets.

12. POLYGLYCINE AND OTHER POLYPEPTIDES WITH THE β CONFIGURATION

Polyglycine has not been prepared in the form of fibers. Powder photographs show a strong ring at 4.35 A, a very strong ring at 3.40 A, and a number of weaker rings. Astbury, Dalgliesh, Darmond, and Sutherland ([57]) suggested an orthorhombic cell with $a = 4.36$ A, $b =$ approximately 7 A, and $c = 3.44$ A. It was then pointed out by Astbury ([81]) that this unit leads to an improbably high density, 1.8 g cm^{-3}, and he suggested that the unit is monoclinic, with $\beta = 66°$. The value of a_0 becomes 4.77 A, in satisfactory agreement with the hydrogen-bonded distance in a sheet of the β-keratin type. Astbury suggested that the chains be completely extended, with $b_0 = 7.34$ A.

We have prepared powder photographs of polyglycine which are similar to those previously reported by Astbury, et al., and by Bamford, Hanby, and Happey ([58]). In addition, the photographs show a pronounced reflection at the angle corresponding to the spacing 1.172 A. If the structure is a pleated-sheet structure, it would be expected that the reflection 060, for which most of the atoms are in phase, would be strong. It is probable that the 1.172 A reflection is to be interpreted as 060. The value b_0 is accordingly $6 \times 1.172 = 7.03$ A.

This value is the value predicted for the antiparallel-chain pleated sheet, rather than for the parallel-chain pleated sheet. There are no observed reflections which require a_0 to be doubled — the predicted value for the antiparallel-chain pleated sheet is 9.6 A, whereas that for the parallel-chain pleated sheet is 4.8 A. It is not unlikely, however, that in the course of formation of crystalline polyglycine the segregation of the polypeptide chains into layers in which the chains alternate in direction is only partial, and that an occasional pair of adjacent chains have the same direction. Flaws of this sort involving a few percent of the pairs of chains would diminish the intensities of the reflections requiring the larger unit sufficiently to account for their non-observance.

We have carried out a detailed calculation of intensities of reflections to be expected for a pleated-sheet structure of polyglycine, resulting in reasonably good agreement with the observed intensities. The data are not extensive enough to permit an experimental decision to be made between the antiparallel-chain pleated sheet and the parallel-chain pleated sheet.

It was mentioned in section 6 that some synthetic polypeptides, in addition to polyglycine, give x-ray diagrams similar to those of the β-keratin proteins. Bamford, Hanby, and Happey ([58]) found that some copolymers containing a large fraction of glycine residues are in this class. Although it is likely that these polypeptides have the pleated-sheet structures, this conclusion has not been verified by a quantitative discussion of the x-ray diagrams.

13. THE STRUCTURE OF SILK FIBROIN

The early x-ray work on silk fibroin has been summarized in section 1. The approximation of the observed fiber-axis identity distance for silk fibroin, reported as between 6.95 A and 7.2 A, to the predicted value for the antiparallel-chain pleated sheet, 7.00 A, suggested that silk fibroin consists of pleated sheets arranged side by side in some manner. A discussion of alternative structures by comparison of calculated and observed values of intensities of x-ray reflections has been prosecuted by Dr. Richard E. Marsh. It is probable that alternate residues in the polypeptide chains of silk fibroin are glycine residues. It has been found that the intensities of reflection are not satisfactorily accounted for by

a structure in which the remaining residues (principally alanine) have their side chains directed on both sides of the pleated sheet, but that instead the data indicate that all of the side chains of a sheet are on the same side, and that pairs of sheets are oriented in such a way as to include the side chains in the region between them.

This structure applies to the silk of Bombyx mori. The structure of other forms of silk has not yet been discovered, although the values of the fiber-axis identity distance and the close similarity, except for a few reflections, of the x-ray diagrams and that of Bombyx mori indicate strongly that the structures are pleated-sheet structures.

14. THE β-KERATIN PROTEINS

When hair, horn, and other α-keratin proteins are strongly stretched they undergo a change in structure. The fiber-axis identity distance of the β-keratin proteins, produced in this way, is about 6.6 A. The approximation of this value to the predicted value for the parallel-chain pleated sheet, 6.5 A, suggests that this sheet constitutes the characteristic structural feature of the β-keratin proteins. If the polypeptide chains in a fiber of α-keratin are oriented in the same direction, the parallel-chain pleated sheet rather than the antiparallel-chain pleated sheet would, of course, be formed on stretching.

The parallel-chain pleated sheet structure for the β-keratin proteins has not as yet been verified through the comparison of calculated and observed intensities of x-ray diagrams.

15. COLLAGEN AND GELATIN

Collagen and gelatin have a structure different from that of any other class of proteins. The diffraction pattern is characterized by a strong meridional reflection at 2.86 A. The equatorial reflections are those expected for hexagonal packing of circular cylinders, with diameter about 12 A. From the density, 1.35 g cm^{-3}, and these dimensions it can be calculated that there are 3 amino-acid residues per molecule in the length 2.86 A.

A structure for collagen has been proposed ([82]) in which three nearly extended polypeptide chains, involving a sequence of two amide groups with the cis configuration and one with the trans configuration in each chain, are twisted about one another, and held to one another by lateral hydrogen bonds. Although this structure accounts reasonably well for some of the properties of collagen, including the equatorial distribution of intensity on the x-ray diagrams, it is unsatisfactory in other respects, and probably will have to be abandoned. It does not account for the off-equatorial portion of the x-ray diagram; there is a hole along the axis sufficiently large to suggest serious instability; and it is possible that the cis configuration of amide groups is sufficiently unstable to make structures of this type less probable than those involving only trans amide groups. No alternative configuration of polypeptide chains that satisfies the structural requirements described in section 3 and that seems reasonable as a collagen structure has been formulated.

16. NUCLEIC ACIDS AND NUCLEOPROTEINS

A precisely described structure has recently been proposed ([83]) for the nucleic acids. This structure involves long chains formed by phosphate di-ester links between the positions 3 and 5 of furanose residues. Three chains of this type are intertwined to form the nucleic acid molecule. The diameter of the molecule is about 20 A, and the length per nucleotide in each chain is 3.4 A. The purine and pyrimidine residues, attached in position 1 of the furanose, with the β configuration, project at right angles to the axis of the molecule, and their planes are approximately normal to the axis. There are about 3.17 nucleotides per turn of the helix formed by each of the three chains. All of the interatomic distances and other structural parameters are satisfactorily preserved in this structure. Moreover, it is compact, with no open space, unoccupied by atoms, near the axis of the three-chain helical molecule.

It seems reasonable to suggest that nucleoproteins will be found to be constructed of nucleic acids, with the configuration just described, and the polypeptide chains of proteins, with the configuration of the α helix or with some other configuration. A nucleoprotein molecule 20 A in diameter might surround itself with 6,

12, or more α helixes, each 10 A in diameter. These α helixes might pursue a slow helical course about the nucleic acid molecule.

The rapid increase in our knowledge of the structure of proteins and of nucleic acids during recent years suggests that it may well be possible to obtain, during the next few decades, a far more detailed picture of fundamental biological processes than could be formulated without the use of information about the atomic structures of these molecules.

Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California.

REFERENCES

(1) R.O. Herzog and W. Jancke, *Ber.*, **53**, 2162 (1920); « Festschrift der Kaiser Wilhelm Gesellschaft », (1921).
(2) R. Brill, *Ann.*, **434**, 204 (1923).
(3) K.H. Meyer and H. Mark, *Ber.*, **61**, 1932 (1928).
(4) O. Kratky, *Z. physik. Ch.*, B **5**, 297 (1929).
(5) O. Kratky and S. Kuriyama, *ibid.*, B **11**, 363 (1930).
(6) C. Trogus and K. Hess, *Biochem. Z.*, **260**, 376 (1933).
(7) K.H. Meyer, M. Fuld, and O. Klemm, *Helv., Chim. Acta*, **23**, 1441 (1940).
(8) W.T. Astbury and A. Street, *Phil. Trans. Roy. Soc.*, A **230**, 75 (1931).
(9) W.T. Astbury and H.J. Woods, *Phil. Trans. Roy. Soc.*, A **232**, 333 (1933).
(10) A.E. Mirsky and L. Pauling, *Proc. Nat. Acad., Sci.*, **21**, 439 (1936).
(11) M.L. Huggins, *J. Org. Chem.*, **1**, 407 (1936).
(12) G. Boehm and H.H. Weber, *Kolloid. Z.*, **61**, 269 (1932).
(13) W.T. Astbury and S. Dickinson, *Proc. Roy. Soc.*, B **129**, 307 (1940).
(14) W.T. Astbury, *J. Int. Soc. Leather Trades' Chem.*, **24**, 69 (1940).
(15) O. Gerngross and J.R. Katz, *Kolloid. Z.*, **39**, 181 (1926).
(16) W.T. Astbury and T.C. Marwick, *Nature*, **130**, 309 (1932).
(17) R.B. Corey and R.W.G. Wyckoff, *J. Biol. Chem.*, **114**, 407 (1936).
(18) I. MacArthur, *Nature*, **152**, 38 (1943).
(19) R.S. Bear, *J. Am. Chem. Soc.*, **66**, 2043 (1944).
(20) J.D. Bernal and D.M. Crowfoot, *Nature*, **133**, 794 (1934).
(21) D.M. Crowfoot, *Proc. Roy. Soc.*, A **164**, 580 (1938).
(22) J. Boyes-Watson, E. Davidson, and M.F. Perutz, *Proc. Roy. Soc.*, A **191**, 83 (1947).
(23) M.F. Perutz, *ibid.*, A **195**, 474 (1949).
(24) J.C. Kendrew, *ibid.*, A **201**, 62 (1950).
(25) J.D. Bernal, *Z. Krist.*, **78**, 363 (1931).
(26) J. Hengstenberg and F.V. Lenel, *Z. Krist.*, **77**, 424 (1931).
(27) R.B. Corey, *J. Am. Chem. Soc.*, **60**, 1598 (1938).
(28) G. Albrecht and R.B. Corey, *J. Am. Chem. Soc.*, **61**, 1087 (1939).
(29) R.B. Corey and H.A. Lévy, *J. Am. Chem. Soc.*, **63**, 2095 (1941).
(30) J. Donohue, *ibid.*, **72**, 949 (1950).
(31) R.B. Corey, J. Donohue, V. Schomaker, and D.P. Schoemaker, *J. Am. Chem. Soc.*, **72**, 2328 (1950).
(32) J. Donohue and K.N. Trueblood, *Acta Cryst.*, **5**, 414, 419 (1952).
(33) D.P. Shoemaker, R. Barieau, J. Donohue, and C.-S. Lu (in press).
(34) E.W. Hughes and W.J. Moore, *J. Am. Chem. Soc.*, **71**, 2618 (1949).
(35) E.W. Hughes and A.B. Biswas (to be published).
(36) G.B. Carpenter and J. Donohue, *J. Am. Chem. Soc.*, **72**, 2315 (1950).
(37) E.W. Hughes and H.L. Yakel, Jr. (to be published).
(38) J. Donohue and P. Vaughan, *Acta Cryst.*, **5**, 530 (1952).
(39) C.-S. Lu, E.W. Hughes, and P.A. Giguère, *J. Am. Chem. Soc.*, **63**, 1507 (1941).
(40) A.F. Schuch, L.L. Merritt, Jr., and J.H. Sturdivant (not yet published).
(41) J. Zussman, *Acta Cryst.*, **4**, 72, 493 (1951).
(42) H.B. Dyer, *Acta Cryst.*, **4**, 42 (1951).
(43) W. Cochran and B.R. Penfold, *Acta Cryst.*, **5**, 644 (1952).
(44) R.B. Corey and L. Pauling, *Proc. Roy. Soc.* (in the press).
(45) D. Crowfoot, C.W. Bunn, B.W. Rogers-Low, and A. Turner-Jones, *The Chemistry of Penicillin*, Princeton University Press (1949), p. 310.
(46) S. Mizushima and T. Shimanouchi, *J. Am. Chem. Soc.*, **74**, 5550 (1952).
(47) L. Pauling and R.B. Corey, *ibid.*, **37**, 729 (1951).
(48) See the summarizing study of the hydrogen bond by J. Donohue, *J. Phys. Chem.*, **56**, 502 (1952).
(49) L. Pauling and R.B. Corey, *Proc. Roy. Soc.* (in the press).
(50) R.B. Corey, *Chem. Revs.*, **26**, 227 (1940).

(51) R.B. Corey and J. Donohue, *J. Am. Chem. Soc.*, **72,** 2899 (1950).
(52) W.L. Bragg, J.C. Kendrew, and M.F. Perutz, *Proc. Roy. Soc.*, A **203,** 321 (1950).
(53) B.W. Low and R.B. Baybutt, *J. Am. Chem. Soc.*, **74,** 5806 (1952).
(54) L. Pauling and R.B. Corey, *J. Am. Chem. Soc.*, **71,** 5349 (1950); L. Pauling, R.B. Corey, and H.R. Branson, *Proc. Nat. Acad. Sci.*, **37,** 205 (1951).
(55) L. Pauling and R.B. Corey, *ibid.*, **37,** 235 (1951).
(56) J.N. Bijvoet, A.F. Peerdeman, and A.J. Van Bommel, *Nature*, **168,** 271 (1951).
(57) W.T. Astbury, C.E. Dalgliesh, S.E. Darmon, G.B.B.M. Sutherland, *Nature*, **162,** 596 (1948).
(58) C.H. Bamford, W.E. Hanby, and F. Happey, *Proc. Roy. Soc.*, A **205,** 30 (1951).
(59) E.J. Ambrose and A. Elliott, *ibid.*, A **205,** 47 (1951).
(60) M.L. Huggins, *Chem. Revs.*, **32,** 195 (1943).
(61) H. Zahn, *Z. Naturforsch.*, **2** B, 104 (1947).
(62) T. Simanouti and S. Mizushima, *Bull. Chem. Soc. Japan*, **21,** 1 (1948).
(63) E.J. Ambrose and W.E. Hanby, *Nature*, **163,** 483 (1949).
(64) A discussion of the structure has recently been presented by C. Robinson and E.J. Ambrose, *Trans. Faraday Soc.*, **48,** 854 (1952).
(65) M.F. Perutz, *Nature*, **167,** 1053 (1951).
(66) M.F. Perutz, in « Annual Reports on the Progress of Chemistry », *The Chemical Society of London*, **1951,** pp. 368, 370.
(67) L. Pauling and R.B. Corey, *Proc. Nat. Acad. Sci.*, **37,** 261 (1951).
(68) C.H. Bamford, L. Brown, A. Elliott, W.E. Hanby, and I.F. Trotter, *Nature*, **169,** 357 (1952).
(69) W. Cochran and F.H.C. Crick, *Nature*, **169,** 234 (1952); W. Cochran, F.H.C. Crick, and V. Vand, *Acta Cryst.*, **5,** 581 (1952).
(70) C.H. Bamford, L. Brown, A. Elliott, W.E. Hanby, and I.F. Trotter, *Proc. Roy. Soc.* (in the press).
(71) L. Pauling and R.B. Corey, *Nature*, **171,** 59 (1953).
(72) F.H.C. Crick, *Nature*, **170,** 1882 (1952).
(73) W.T. Astbury, *Cold Spring Harbor Symposia on Quantitative Biology*, **2,** 15 (1934).
(74) L. Pauling and R.B. Corey, *Proc. Nat. Acad. Sci.*, **37,** 256 (1951).
(75) L. Pauling and R.B. Corey, *Proc. Nat. Acad. Sci.*, **37,** 282 (1951).
(76) R.C.C. St. George and L. Pauling, *Science*, **114,** 629 (1951).
(77) F.H.C. Crick, *Acta Cryst.*, **5,** 381 (1952).
(78) D.P. Riley and U.W. Arndt, *Nature*, **169,** 138 (1952).
(79) D.P. Riley and U.W. Arndt, *Proc. Roy. Soc.* (tn the press).
(80) L. Pauling and R.B. Corey, *Proc. Nat. Acad. Sci.*, **37,** 251 (1951).
(81) W.T. Astbury, *Nature*, **163,** 722 (1949).
(82) L. Pauling and Robert B. Corey, *Proc. Nat. Acad. Sci.*, **37,** 272 (1951).
(83) L. Pauling and R.B. Corey, *ibid.*, **39,** February (1953).

X-Ray Analysis of Protein Stucture

by Lawrence Bragg

1. I feel it a great compliment that I have been invited to take part in a conference of chemists, although my own field of work lies so exclusively in the domain of physics. I am reminded of the fable of the stork and the fox. The fox invited the stork to dinner but his guest fared very badly because the repast was spread on flat plates from which he found it difficult to eat. When in turn the stork was the host, the viands were in jars which suited the stork's bill, but were inaccessible to the fox. I have been listening to chemical experts and trying to partake of your fare. When I now describe our investigations which are carried out by methods of physical optics, I fear you will be confronted with a profound physical jar. I propose, however, to describe to you the application of X-ray analysis to proteins, and to outline the progress which has so far been made, so that you may see where these methods are powerful and where they are limited. In the case of simpler substances such as minerals or alloys, it can fairly be claimed that the investigation into the arrangement of atoms has cast a flood of light on the chemical nature of these substances. If we could get some hints as to the general architecture of a globular protein, we cannot doubt that the findings of X-ray analysis would be equally significant. It is of course a very ambitious venture. The most difficult substances so far analysed by standard methods contain molecules with one or two hundred atoms. In a protein molecule the atoms are numbered in thousands, and in view of the trial and error procedure of X-ray analysis the problem would seem almost beyond the possibility of solution. Nevertheless certain peculiar features of protein crystals have made possible a very onsiderable advance.

2. The Cavendish Laboratory contains a research group financed by the Medical Research Council, and under the direction of

Dr. M.F. Perutz. The following problems are being studied by members of this group:

(a) The structure of haemoglobin. This was initiated by Professor Bernal and Dr. Perutz ([1]), and the latter has continued the research for over fifteen years. Haemoglobin gives highly perfect and reproducible X-ray diffraction pictures. A successful analysis of these pictures would give us a projection of the molecule on a plane in considerable detail. Such projections, it is to be hoped, would lead to an understanding of the general molecular architecture of the protein.

(b) The structure of myoglobin. Dr. Kendrew is in charge of this work. Until quite recently he has found it very hard to get good crystals of myoglobin. Recently he has found a rich source in the various species of whale. Many different crystalline types have been grown to a large size. As the molecular weight of myoglobin (17,000) is only one quarter that of haemoglobin, it may well prove to be an even more favourable type of crystal. These researches into whale myoglobin, only recently started, are being actively developed.

(c) Pleochroism. The absorption of light for different directions of polarization affords a powerful method of relating the orientations of the molecules of haemoglobin and myoglobin in different crystalline forms. Absorption is large for an electron vector in the plane of the haem group, small for one at right angles to it. The pleochroic ratio in haemoglobin is 2.6. It is far higher in some myoglobin crystals where all haem groups must be very nearly parallel.

(d) X-ray diffraction by living muscle fibres (Huxley). The spacings of the living relaxed muscle fibre under different conditions have been studied, and it is planned to extend these to contracted living muscle.

(e) The development of powerful X-ray tubes for protein analysis, so as to lessen the often very lengthy time of exposure (Broad).

(f) The study by X-rays of virus structure (Watson).

(g) The theory of diffraction by helical arrangements (Cochran and Crick) ([2]). In particular this theory, applied to artificial polypeptide chains, has strongly confirmed Pauling's proposed α helix.

(h) The study of purines, pyrimidines, and their complexes because of their importance in nucleic acids (Cochran).

(i) The structure of nucleic acid (Crick and Watson). This very recent investigation, mentioned in the previous communication by Professor Pauling, is described in the note which follows this paper.

3. I shall confine myself to the X-ray analysis of haemoglobin, treated purely as a problem of physical optics. I must not go too deeply into our physical methods of investigation, for you are no doubt far more interested in results than in techniques. I wish, however, to say something of these methods so that you may judge what contribution we can make, and where we need your help in interpreting the results.

If one is looking at a distant light source through a telescope, and a mask is placed in front of the objective which has an aperture of some particular shape, or a set of holes in a particular arrangement, a " diffraction pattern " is seen in the eyepiece. This is the optical " Fraunhofer " diffraction pattern formed by the waves which have come through the apertures, and which interfere. The pattern of spots which the X-ray crystallographer must analyse may be compared to such a Fraunhofer pattern. Instead of holes in a mask, there are atoms which scatter the X-rays, but the principle is the same. The analogy is used in an apparatus which is becoming an increasingly important weapon of X-rays analysis and which in its present form was devised by Lipson in Manchester ([3]). If one has guessed that a certain arrangement of atoms is the unit of crystal structure, a drawing is made of the projection of these atoms on one of the crystal planes. By a pantographic arrangement, a small scale version of this drawing is reproduced as a set of holes punched in an opaque film. This film is placed in an optical arrangement which yields its Fraunhofer diffraction pattern. If the crystal structure has been correctly guessed, the observed pattern of X-ray spots will closely resemble the optical interference pattern.

4. This analogy between light interference and X-ray interference is useful because it shows that X-ray diffraction is another version of a familiar effect, though it is rather more complex because the scattering patern is in three dimensions in the X-ray case. For any given crystal arrangement it is always possible to compare the predicted diffraction with that actually observed, and so to test whether the structure is correct.

If the reverse were true, and we could deduce the structure directly from the observed diffraction effects, X-ray analysis would be a

purely mechanical operation. This is not the case; although a given crystal structure can give only one kind of diffraction pattern, a given diffraction pattern could be produced by an infinite number of crystal structures. The reverse procedure is, analytically, quite impossible. This can be seen by considering the Abbe method of treating the resolving power of a microscope. It is focussed on a grating illuminated by parallel monochromatic light. The spectra formed by the grating come to a focus behind the objective in its principal focal plane, and waves from these focussed spectra continue on to produce the image by their interference. In order to produce a correct image, these waves must have both the correct amplitude and the correct phase. Their amplitudes could be measured by recording the spectra, for example, on a photographic plate, which corresponds to recording the X-ray diffraction spots. The *phases* are not recorded, however. In trying to interpret the diffraction picture we could assign arbitrarily any phase to any spot, and every change of phase would create a different type of image. A very helfpful simplification arises when the object has a centre of symmetry, for then the arbitrary phase is represented merely by an arbitrary choice of + or — sign for the amplitude of the wave owing to the symmetry, but the problem is still formidable. Consider a protein crystal like haemoglobin, which has a symmetry centre in the projection we are considering and gives perhaps 600 diffracted spots. We can measure their amplitudes; what are their phases? If we knew the answer we could do mathematically what the microscope does by interference and form the image. But the 600 observed spots could be explained by 2^{600} types of crystal, since any one may be + or —. Which of these structures is the right one?

In simple crystals, where the problem is very similar, the answer is that the right solution is one which gives a sensible result. We try various signs until a picture with the right kinds of atoms in plausible places begins to emerge. We cannot do this with a protein molecule. Not only are there far too many atoms, but we have no idea what to expect, or how to apply the criterion of what is "sensible". Some direct method of attack must be substituted for the usual indirect attack by trial and error.

5. Haemoglobin, as Perutz has shown, crystallizes in a number of different forms. A comparison of these forms has made it possible to measure rather exactly the dimensions of the haemoglobin

molecule ([4]). For instance, in one trigonal form each molecule is at the centre of a plane group of six similarly oriented molecules 56 A away. In another tetragonal form it is surrounded by four at a distance of 54 A. Clearly one section of the molecule must be nearly circular with a width of 54-56 A. Packing arguments show that it is longer in the direction perpendicular to the circular section. In a number of cases six or eight nearest neighbours are about 65 A distant. An upper limit to the greatest length of the molecule is 73 A in a crystal where this can be measured. To sum up, the

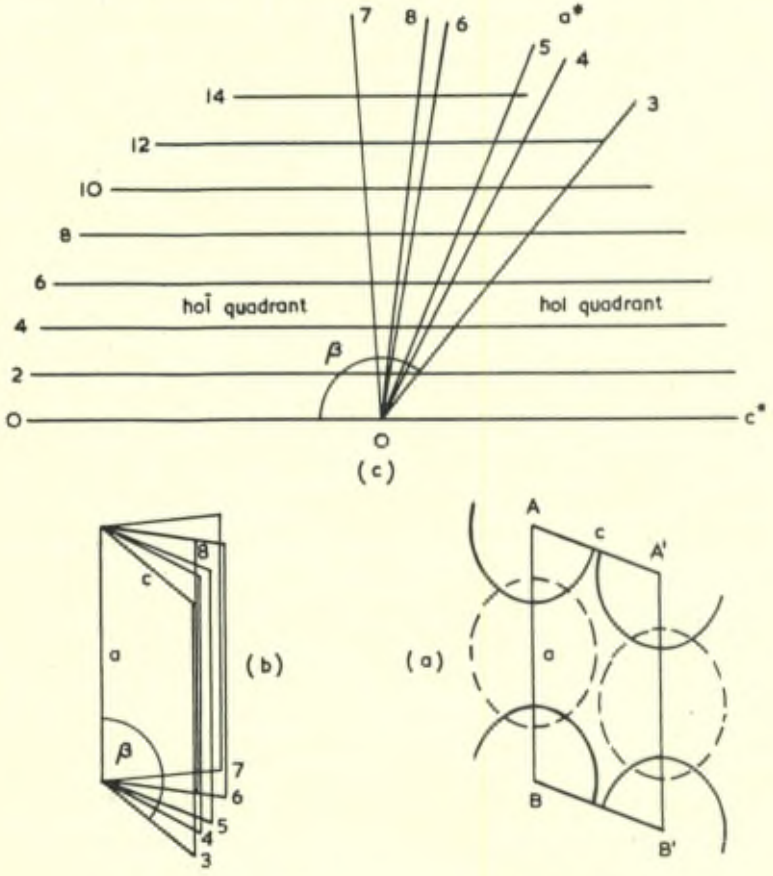


Fig. 1 (a) The unit cell of horse haemoglobin projected on the "b" face of the crystal.
(b) Different stages of shrinking or expansion of the unit cell.
(c) The layer lines. The various forms make it possible to obtain numerous samples of the amplitude of the *trans*form along these layer lines.

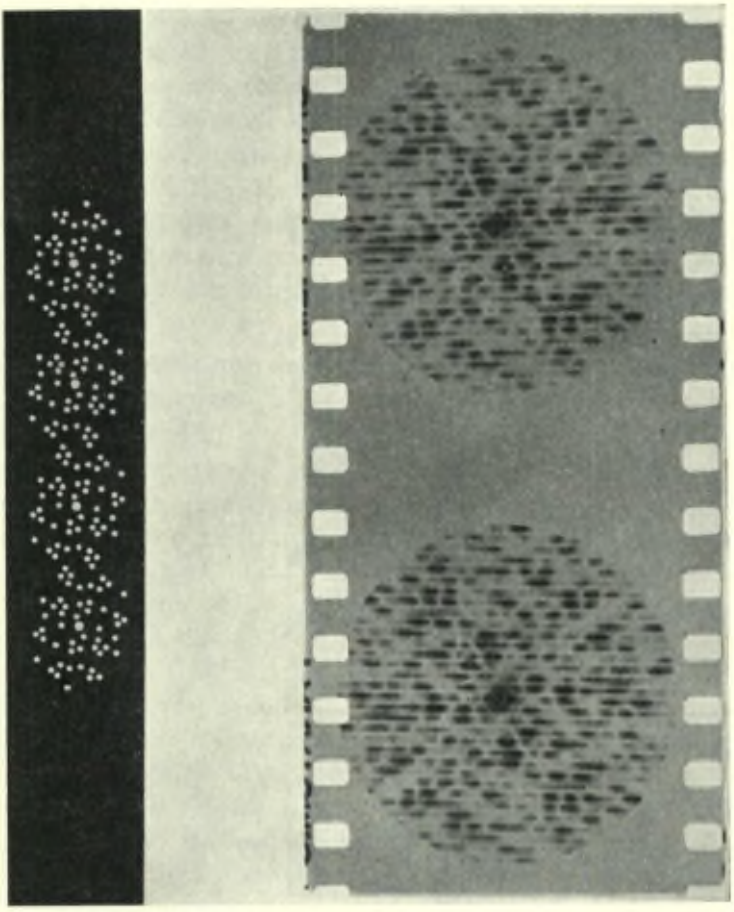


Fig. 2. — An optical Fraunhofer pattern (right) formed by a row of idealized haemoglobin molecules (left). In the lower Fraunhofer pattern, a hole has been added at the centre of each molecule, which enhances the positive loops and diminishes the negative ones, thus enabling them to be indexed with a sign.

protein molecule packs into crystals as if it were a prolate spheroid about 70 × 55 × 55 A. This is for all forms of "wet" crystal. If they are dried, the molecules approach more closely and their dimensions are about 60 × 48 × 48 A. Measurements from different crystals agree to within 2 A.

This approximate definition of the shape of the haemoglobin molecule is a necessary basis for the direct method of interpreting its X-ray diffraction.

6. There is another feature of haemoglobin crystals which may be used in analysis. Perutz has shown that horse haemoglobin passes through a series of shrinkage stages when the humidity of the surrounding air is reduced ([5]). They are illustrated in figure 1 (b), which shows the monoclinic unit cells projected on a plane perpendicular to the twofold " b " axis. The axes *a*, *b*, *c* remain constant in length, but the angle β alters. Further, Perutz has shown by various tests which need not be described here that the sheets of atoms AB, or A'B', remain constant in structure, and that the change is one by which A'B' approaches or recedes from AB.

What type of diffraction would be given by a single row of molecules AB? This can be tested by the apparatus described in Section 3 above. Figure 2 (a) shows a row of imaginary haemoglobin molecules, their structure being represented by punched holes contained within the ellipses representing the dimensions deduced from the packing, and with a centre of symmetry in each molecule. Figure 2 (b) shows the diffraction pattern produced by this row of molecules. Certain features of the diffraction pattern are characteristic. In the first place the diffracted intensity is concentrated along " layer lines " perpendicular to the row. These correspond to successive spectra of a grating; the zero layer line runs through the centre of the pattern, the next is one for which there is one wave length path difference for successive molecules in the row, the next has 2λ path difference. Next, the intensity along the layer lines is in patches separated by zero values. The amplitude must be + or — because the molecule has a centre of symmetry and if it is to change from + to — it must pass through a zero value. We thus find on a layer line a series of alternate + and — loops separated by zero nodes. Finally, it will be seen that there is a minimum " wave-number " of loops. This minimum is determined by the dimensions of the molecules perpendicular to the row; the greater the dimension, the more closely are the loops crowded together ([6]).

If haemoglobin only crystallized in one form as for instance that of figure 1 (a), we would get only one set of spots on the layer lines and it would be impossible to tell whether the sign of the amplitude had changed from + to — between any two neighbouring spots. However, since there is a number of shrinkage forms as shown in figure 1 (b), we get numerous samples of the amplitude of diffraction along the layer lines shown in figure 1 (c). Each crystal produces

spots in a different place, both because A'B' approaches AB and because the angle β alters. Hence we can sample the nodes and loops at numerous points rather like a ship taking a series of soundings along its course. The nodes and loops of the haemoglobin diffraction pattern, deduced in this way, are shown in figure 3. By noting where the amplitude runs through a zero, we can tell when it changes from + to —.

If the sign of any one loop is known, the signs of all the others, on the layer line follow because they are alternately + and —. This fortunate shrinkage behaviour of the crystal therefore reduces the problem of determining the structure from that of guessing about 100 signs in this particular case, to that of guessing 7, for the 7 layer lines in figure 3. The number of possibilities is reduced from 2^{100} to 2^7.

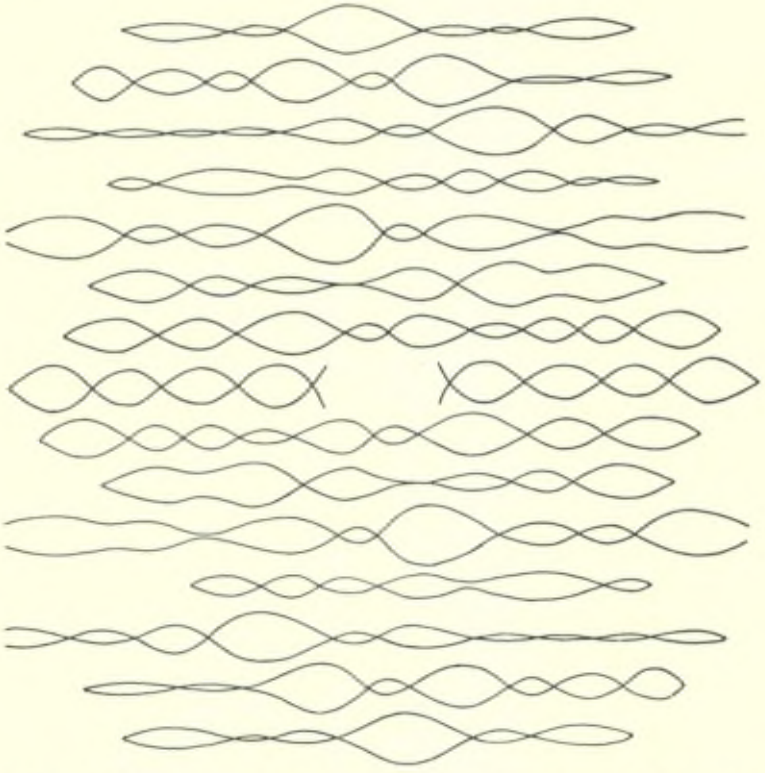


Fig. 3. — Nodes and loops in the *trans*form of the horse haemoglobin molecule, deduced from X-ray measurements of diffraction by the crystal forms shown in figure 1.

This is still quite a formidable number of alternatives, but if one were dealing with an organic molecule of moderate complexity the crystallographer would regard the problem as easily soluble from this stage onwards. A trial of alternative signs for the first two or three layer lines would produce only one reasonable solution showing the general shape of the molecule under low resolving power. Having established the right choice for these first layer lines, the picture would be sharpened up by trying alternatives for the next layer lines and so forth. The difficulty of haemoglobin analysis is due to our ignorance of what is a reasonable picture, as stressed above. It is still necessary to apply direct methods, and two lines of attack are being followed. One is to substitute salt solution for water inside the crystal, which reduces the low orders of diffraction because it lessens the contrast in density between the protein and the liquid. The change in amplitude depends upon the form of the molecule alone, not on its internal structure, and gives a hint as to the signs of the lower layer lines.

Incidentially this examination of the " salt effect " has enabled Perutz to draw interesting conclusions about the molecule. The amount of salt which enters the crystal can be measured by weighing it, and the volume which it fills can be estimated by noting where the density of the solution exactly matches that of protein, shown by the disappearance of the lowest orders of diffraction. It appears that while the protein molecule itself has a volume of 83,000 A^3, the ammonium sulphate solution is restricted to the space outside a volume of 130,000 A^3. This is attributed to the existence of a layer of bound water on the surface of the molecule, into which the salt does not penetrate.

Another method is to look for other crystalline forms which have molecules in a similar orientation, but with a different distance AB (fig. 1). Such crystals would give values of the amplitude *between* the layer lines of figure 3, making it possible to draw nodes and loops across the layer lines as well as along them, and so tying the whole diagram together. The prospect is hopeful, but success has not yet been achieved.

The signs of the zero and first layer lines are not in doubt, they follow from very general optical considerations. If we only attempt to form a picture of the protein molecule at low resolution, we can for the moment neglect the fourth and higher layer lines which merely serve to sharpen up detail. By taking either alternative in

signs for the second and third layer lines, four possible pictures of the row of haemoglobin molecules can be calculated. One of these must be right, and if only we could get some hint from our chemical colleagues as to which it is it would probably be possible to bring in the higher layer lines and show the general form of the molecule.

SUMMARY

X-ray measurements of haemoglobin crystals by Perutz have yielded rather accurate estimates of the size and shape of the molecule, in both " wet " and " dry " form. Similar measurements of myoglobin by Kendrew will probably give equally accurate data.

Much progress has been made with the seemingly hopeless problem of analysing the internal structure of the molecule by means of the X-ray data. The almost infinite number of possible interpretations of these data has been reduced to a few alternatives, and some hint as to which is the correct one, or some measurement of a further crystalline form of fortunate aspect, may well enable the last stage to be successfully completed in the near future. One will then have a projection of the haemoglobin molecule on a plane, to a resolution of some 7 A. This is insufficient to reveal atomic arrangement, but gives the main architecture of the structure. If this stage can be reached, it ought not to be difficult to extend the analysis to finer detail, for instance to determine where the haems are located. It is disappointing not to be able to announce the final success at this Conference. However, the X-ray analysts of protein at the present time feel rather like climbers who have managed to surmount a series of pitches even more forbidding in their aspect than the final one which lies between them and the summit.

REFERENCES

(1) J.D. Bernal, I. Fankuchen and M.F. Perutz, *Nature*, **141**, 523 (1938).
(2) W. Cochran and F.H.C. Crick, *Nature*, **169**, 234 (1952).
(3) H. Lipson and C.A. Taylor, *Acta Cryst.*, **4**, 458 (1951).
(4) W.L. Bragg and M.F. Perutz, *Acta Cryst.*, **5**, 277 and 323, (1952).
(5) M.F. Perutz, *Trans. Faraday Soc.*, **42 B**, 187 (1946); J. Boyes-Watson, E. Davidson and M.F. Perutz, *Proc. Roy. Soc.*, **191**, 83 (1947).
(6) W.L. Bragg and M.F. Perutz, *Proc. Roy. Soc.*, **213**, 425 (1952).

(Note complémentaire)

The Stereochemical Structure of DNA

by J. D. Watson and F. H. C. Crick

While the chemical formula of DNA suggests that the basic structure of DNA is a very long unbranched polynucleotide chain, recent X-ray diffraction data have led Wilkins and Franklin and their co-workers at King's College, London, to postulate a basic stereochemical structure containing more than one polynucleotide chain arranged together in a regular manner, possibly in the form of intertwined helices. To account for this crystallographic inference,

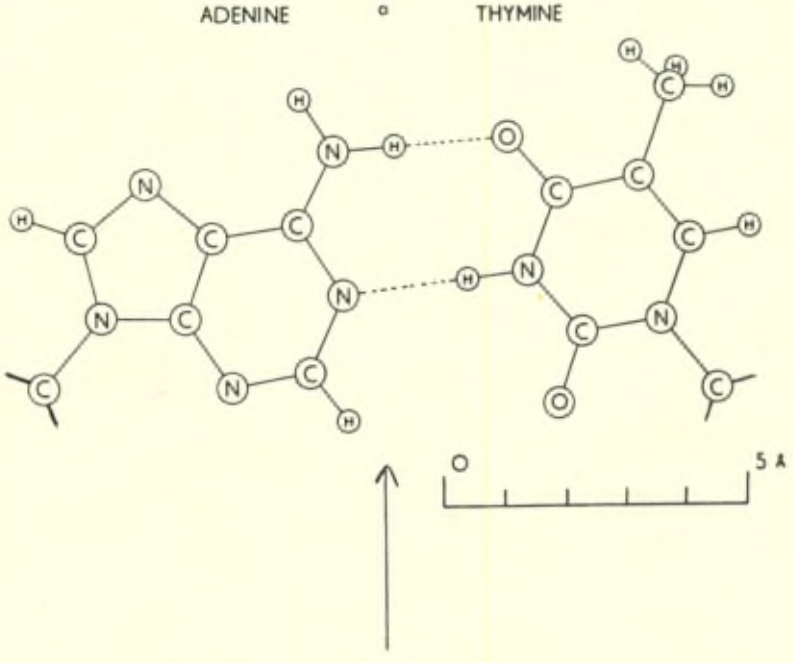


Fig. 1. — Pairing of adenine with thymine. Hydrogen bonds are shown dotted. One carbon atom of each sugar is shown.

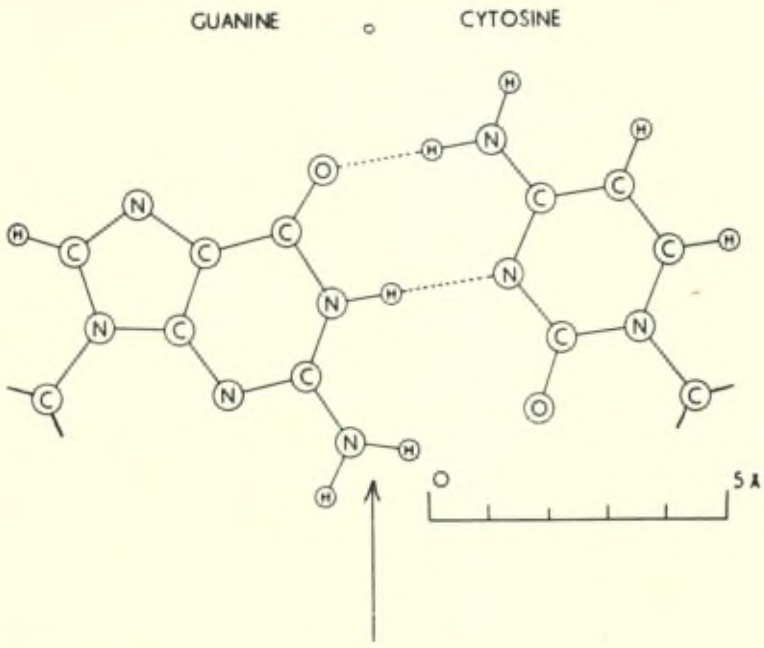


Fig. 2. — Pairing of guanine with cytosine. Hydrogen bonds shown dotted. One carbon atom of each sugar is shown.

we have proposed (Watson and Crick 1953*a*, 1953*b*) a two stranded helix in which the two chains are joined together by hydrogen bonds between the bases. The bases are joined together in pairs, a single base in one chain being hydrogen bonded to a single base from the other. The pairing is highly specific, not only must one member be a purine and the other a pyrimidine, but adenine must pair with thymine and guanine with cytosine. The way in which this is accomplished is shown in figures 1 and 2. As can be seen the links joining the bases to the sugarphosphate backbone are similarly spaced in both pairs, so that structurally the two pairs are interchangeable. A given nucleotide can occur on either chain, but when it does, its partner on the other chain is specifically determined. This results in a complementary relationship between the sequence of bases on the two chains, the sequence of which on one chain we believe to be irregular.

As far as we can tell, this structure is compatible with the X-ray evidence of the King's College group (Wilkins et al., 1953; Franklin

and Gosling, 1953), but it can in no sense be considered proved. Two lines of evidence, however, suggest to us that the complementary feature of the structure is probably right. Firstly, the analytical data (Wyatt, 1952) on the proportion of the various bases shows that for all sources of DNA analysed, the amount of adenine is close to the amount of thymine, and the amount of guanine close to the amount of cytosine, although the ratio of adenine to guanine can vary from one source to another. Secondly, the biological evidence suggesting that DNA is the carrier of at least part of the genetic specificity of the chromosomes and thus must possess the capacity for self duplication. The presence of complementary strands immediately suggests a self replicating mechanism, since if we postulate that the two chains unwind and separate, then each can serve as a template for the formation of its complement.

REFERENCES

(1) Wilkins, M.H.F., Stokes, A.R. and Wilson, H.R., *Nature*, **171**, 738 (1953).
(2) Franklin, R.E., and Gosling, R.G., *Nature*, **171**, 740 (1953).
(3) Watson, J.D., and Crick, F.H.C., *Nature*, **171**, 737 (1953).
(4) Watson, J.D., and Crick, F.H.C., *Nature*, **171**, 964 (1953).
(5) Wyatt, G.R., *J. Gen. Physiol.*, **36**, 201 (1952).

Discussion des rapports de MM. L. Pauling et L. Bragg

M. Pauling. — Although it is only two months since Professor Corey and I published our proposed structure for nucleic acid, I think that we must admit that it is probably wrong.

I have just learned, a few days ago, that Dr. Wilkins, of King's College, London, and his collaborators have shown that there are two different forms of sodium thymonucleate, giving different X-ray patterns, and that our X-ray photographs, and the older ones of Astbury and Bell, show a superposition of the two patterns. We were misled in this way; we attempted to formulate a structure accounting for a feature of one pattern (the 3.4 A meridional reflection) and a feature of the other pattern (the 27-A repeat).

We also concluded that the nitrogen bases could not be near the axis of the helical molecule. This conclusion, which Watson and Crick have shown to be unjustified, resulted from our assumption that each nucleotide residue should provide a suitable place for any nitrogen base, either purine or pyrimidine. Watson and Crick have as an extraordinary feature of their structure the occurence of purine-pyrimidine pairs (guanine-cytosine, adenine-thymosine), which lie near the axis of the helix.

Although some refinement might still be made, I feel that it is very likely that the Watson-Crick structure is essentially correct. In its feature of complementariness of the two chains it suggests a mechanism for duplication of a chain by a two-step process — a molecular mechanism that may well be the mechanism of hereditary transmission of characters.

I think that the formulation of their structure by Watson and Crick may turn out to be the greatest development in the field of molecular genetics in recent years.

M. Neuberger. — The α-helix as an important feature of the structure of many fibrous and most globular proteins has been widely accepted amongst protein chemists, but not many attempts

have so far been made to link up the crystallographic evidence or interpretation with the many, purely chemical, observations obtained during the last few years.

One of the basic assumptions in Dr. Pauling's arguments has been the postulate that the most likely structure for a protein is the one which has the lowest energy relative to those of similar structures under consideration. This approach is almost certainly correct for most fibrous proteins and for many, relatively stable, globular proteins. The majority of native globular proteins, however, are somewhat unstable and in these cases the assumption that the structure associated with the lowest energy is the most probable one, would appear to be unwarranted.

Another related problem is the following : from the known approximate dimensions of the globular proteins of molecular weights between 35,000 and 10,000, one can deduce that they must contain between 10 and 30 rods of the α-helix type. On the other hand, the end group methods reveal in some cases a cyclic structure and suggest in most cases a much smaller number of peptide chains. It is of course possible that the end group methods used at present are not really reliable; but if we accept the chemical results as approximately correct, it would follow that the α-helices bend round with an angle of about 180°. Such a bending must be associated with a distortion of either valency angles in the main chains or more likely, a distortion of the hydrogen bonds. In addition, there may be steric hindrance between side chains.

It may therefore be suggested that in the majority of globular proteins there are regions of low energy, i. e. in the straight part of the chains, and regions of higher potential energy at and near the bends. The configuration at the bends may be retained owing to electrostatic and van der Waals' interaction of the side chains. On denaturation this stabilisation is abolished and the α-helices uncoil at the bends.

In insulin the peptide chains are relatively short and no bending is required to accomodate the known facts. It is therefore understandable why this protein is unusually stable and cannot be denatured in the same way as e. g. egg albumin or haemoglobin.

M. Pauling. — I agree with Professor Neuberger that we need to obtain information about the way that the α helix bends around

corners. Many years ago (L. Pauling, *Jour. Am. Chem. Soc.*, 62, 2643, 1940) I suggested that the presence of proline residues, which prevent hydrogen-bond formation by the nitrogen atom of the amide group, would aid the polypeptide chain to get around a corner. Also side-chains might form hydrogen bonds with the amide, NH and CO groups, and thus stabilize the bent α helix.

It is of course possible, as Professor Neuberger has pointed out, that the bend is associated with an increase in energy — with strain. Presumably the reversibility of denaturation of hemoglobin and trypsin (Anson & Mirsky) shows that the bends in these molecules are not accompanied by strain; the bends in the α helix of ovalbumin, the denaturation of which has not been reversed, may well involve strain. We could thus understand the reported structure of denatured ovalbumin, from physicochemical data, as a rigid rod 600 A long (α helix, 1.5 A per residue).

M. Bragg. — Patterson diagrams of haemoglobin molecules indicate rods seen from the end; but the number of « crystallographic » chains is higher than the corresponding number of open polypeptide chains as measured by end groups methods. It is thus likely that chains are bent and turn back. If we had strictly parallel α helices, we should obtain a rather strong 1.5 A reflection, which is not the case,

M. Hermans. — Prof. Bragg has told us that a hemoglobin molecule in a salt solution is surrounded by a layer of water molecules which apparently does not contain any ions. This is based on an analysis of the salt concentration at which the *electron* density of the protein becomes equal to that of the surrounding medium. In principle this method is the same as that of ordinary density determinations (of the protein in salt solution), which has been used by physical chemists for similar purposes since many years.

However, the water molecules surrounding the protein particle have a lower electron density than either the inside or the outside, and they should therefore act like a shell of negative density for X-rays. Do the details of the X-ray picture give any direct indication of such a shell ?

M. Bragg. — Dr. Hermans is correct in principle, but the knowledge of the molecular shape is as yet too inexact for it to be

possible to draw any conclusion other than that of the total volume into which the salt does now appear to penetrate; it cannot tell us where the bound water is attached.

M. Martin. — *A*. It is to be expected that the coiling of the peptide chains should give rise to striking effects in the optical rotatory power in different directions. Can Prof. Pauling give any theoretical prediction in this matter ?

B. — It would be expected that various of the structures proposed would give rise to piezo electricity. I have observed that the cuticle of wool and hair is piezoelectric while the cortical cells are not. Have any other proteins been examined ?

C. — With the hydrogen bonds of the α-helix cell neatly satisfied and the side chains exposed, is it not surprising that the solubility of native proteins is higher than of denatured proteins, with a more disordered arrangement ?

M. Pauling. — I do not understand the theory of optical activity well enough to make any theoretical predictions about the effect of denaturation on specific rotation of proteins. The three present theories of optical activity, due to Kuhn, Eyring, and Kirkwood, are somewhat different from one another; they lead, nevertheless, to the same conclusion about absolute configurations, which agrees also with the experimental result of Bijvoet. It would, however, be difficult to extend the theory to such complex molecules as proteins.

The piezoelectric effect observed by Dr. Martin for wool cuticle can be explained by the structure of the cuticle, which is that of β-keratin. We have assigned to β-keratin the structure of the parallel-chain pleated sheet, which is a polar structure, that should give rise to piezoelectricity. We predict also that the cortical protein, with the α-keratin structure, should be piezoelectric. The observation by Dr. Martin of a negative result may be due to its having too small a piezoelectric constant to be observed; the negative result does not necessarily mean that the substance must have a non-polar structure.

The solubility of proteins may be attributed to the polar side chains; it does not involve the amide groups. The amide CO...HN hydrogen bonds are stable with respect to water, as shown for

example by the low solubility of diketopiperazine in water. The α-helix is not uncoiled on solution of a protein in water.

M. Havinga. — *A*. One might try to explain the one sidedness of nature in using almost exclusively L-aminoacids for building proteins by the assumption that D-aminoacids do not fit into the structural plan.

A superficial inspection of the arrangement of the molecules in the α-helix does not reveal any essential difficulties of building in D-aminoacids in a helix consisting for the main part of L-aminoacids. I should like to ask Prof. Pauling whether this statement remains valid upon closer examination of the models made for the helical structure.

B. — Formerly we thought of a protein monolayer on water as sheat structure with most of the hydrophilic side chains diving into the water and the hydrophobic chains of leucine and the like oriented towards the air.

Prof. Pauling reported X-ray evidence for the α-helical structure also in the case of proteins spread in a monolayer. Does this helical structure provide an explanation for the remarkable tendency of most proteins to form monomolecular layers at the interface between water and air ?

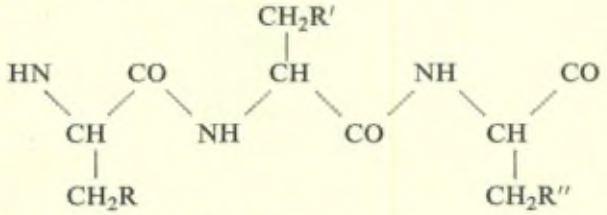
M. Pauling. — The α-helix provides space for the side chains of L-amino-acid residues in essentially arbitrary order; only rarely would steric hindrance be expected to occur. Steric hindrance would be expected to occur more often for an α-helix with both L and D residues. However, I do not feel that the steric effects would be important enough to exclude D residues. The predominance of L residues in proteins must have some other explanation.

A surface film of α-helixes would be about 10 A thick. Probably many of the side chains could twist about and bend down in such a way as to get their polar ends into contact with water molecules, thus stabilizing the film.

M. Desreux. — D'après la structure proposée par le Professeur Bragg pour l'acide desoxyribonucléique, doit-on s'attendre à ce que la molécule se comporte comme une chaine de flexibilité assez faible ?

M. Bragg. — I don't see any reason why nucleic acid molecules should not bend in solution. The chain can contract to some extent, due to small tilts of the nitrogen bases; a bending will result when the effect is acting mainly on one side.

M. Karrer. — On page 9 of the report of Professor Pauling we find for the protein-chain the following formula :



The same formula is used in other papers and even in textbooks. I should like to know what kind of formula this may be. It hardly can be a projection formula in the sense of Emil Fischer, because in this case the chain would be composed of alternating L- and D-amino-acids.

M. Pauling. — This formula is written not as a projection formula in the conventional sense, but rather as a representation of the molecule in perspective.

M. Linderström-Lang. — Could not a partial opening or « dissolution » of the α-helix take place in urea or guanidine chloride. Certain experiment on stepwise degradation of insulin by the Edman method indicates that the N-terminal peptide bond split in the thiohydantain formation is « activated » by guanidine chloride.

M. Pauling. — There is little evidence in existence as to the power of urea or guanidium ion to break NH...OC hydrogen bonds. Information could be obtained on this point by, for example, a study of the solubility of diketopiperazine in aqueous solutions containing these substances.

Chemical Constitution of the Proteins

by A. C. Chibnall

When, in 1945 [1], I reviewed the contribution of the analytical chemist to the problem of protein structure I pointed out that the tempo of research had increased considerably during the decade or so before; that new and rapid micro-methods had already supplanted the laborious gravimetric procedures upon which our earlier knowledge had been based; and that we were at last in a position to effect the complete overall amino-acid analysis of a protein, as the first necessary step in the elucidation of its structure. Since then the tempo, far from slackening, has increased still further, and researches on many different aspects of protein structure are being actively pursued. This broadening of interest has its counterpart in the many summaries and reviews which now appear at regular intervals, so that the whole field is well documented and the newer knowledge is readily available to those who seek it. In particular, the excellent review by Sanger [2] written less than two years ago, and the somewhat earlier one by Tristram [3] cover much of the ground which might come within the purview of the present Report.

As Sanger [2] has emphasized, the experimental evidence at present available does not permit us to form any general theory of protein structure, or to formulate any principles that govern the arrangement of the amino-acid residues in the peptide chains. We are, nevertheless, passing through a period of research activity in which methods that will aid in the collection of further evidence are being initiated, tested and perfected. For this reason I have deliberately restricted my discussion to topics dealing with analytical procedures, in which much progress can be reported, and have not attempted to amplify the survey on structure presented by Sanger.

In a broad sense we can envisage five successive stages of enquiry through which a protein would have to be passed if the ultimate object were the elucidation of its complete chemical structure.

1) Amino-acid analysis of the protein hydrolysate and treatment of the resulting analytical data in the conventional manner.
2) Assay of the N-terminal and C-terminal residues to indicate the minimum number of component peptide chains in the molecule.
3) Determination of the linkages between the component peptide chains.
4) Separation and purification of the component peptide chains.
5) Determination of the amino-acid sequence in the peptide chains.

In the account which follows I shall consider each of these stages in turn and try to appraise the present position.

AMINO-ACID ANALYSIS OF PROTEINS

The pioneer work of Martin and Synge and their collaborators on the chromatographic separation of amino-acids has led to many outstanding developments, and of these perhaps the most interesting in quantitative analysis is the new ion-exchange procedure of Moore and Stein (4) who use a column of the polystyrene resin Dowex-50. Elution is effected by a sequence of buffer solutions of progressively increasing pH, which yields in a single chromatogram an effluent curve in which every component emerges as a distinct peak. The curve obtained from a known mixture of amino-acids representing 3-6 mg of protein gave recoveries of $100 \pm 3\%$ on the average.

The method has been applied recently by Harfenist and Craig (5) in a comprehensive analysis of three different species of insulin. In each case the preparation used was the major component obtained from a successful fractionation by countercurrent distribution. The results obtained after hydrolysis show that six of the amino-acids are present in different amounts in the three species concerned, but that otherwise the compositions are identical. When the necessary experience has been acquired the method will probably prove to be the ideal one for the accurate routine analysis of protein hydrolysates.

In giving prominence to this work of Moore and Stein I do not wish to imply that other procedures introduced in recent years, many of which can be carried out with simple apparatus, may not be of great use. I may instance the dinitrophenylamino-acid technique of Mills (6), the indicator and carrier methods of Keston *et al* (7) and the use of filter-paper chromatograms by a number of different workers. Likewise many of the older procedures may still be useful when a partial analysis for certain types of residues is called for. The point that I wish to emphasize is that a routine procedure of relative simplicity and proved reliability meets the needs of those concerned with most aspects of protein structure, so that studies on methods of amino-acid analysis have no longer the over-riding importance that was rightly attached to them until recently. A decade or so ago the analyst was thinking exclusively in terms of the whole protein molecule and accuracy was his first aim. Today the emphasis is on the component peptide chains, and in such cases a high degree of accuracy is not always called for. It is somewhat ironical that this change of viewpoint should come about at the very time when accuracy can be achieved as a routine!

THE AMINO-ACID COMPOSITION OF PROTEINS

The results of an amino-acid analysis of a protein are generally presented in one or more formats: as g amino-acid per 100 g protein, as variants in terms of residues or N, as moles per 10^5 g protein, or as apparent number of residues per molecule. The first is conventional, while the second shows the completeness of the analysis. The third is useful for comparative purposes, especially when the molecular weight of the protein concerned is not well established. The fourth can provide an independant check of the molecular weight deduced from physico-chemical measurement as well as information which is indispensible in studies of amino-acid sequence. Whichever type of interpretation is used, due regard must be paid to the degree of accuracy with which the results were obtained (3). The data collected in Table 1 illustrates this point. The glycyl chain of insulin has a molecular weight of only 2336 and the residue numbers are so low that the question of accuracy of analysis can hardly be said to arise. As the molecular weight of myoglobin is low the residue numbers are of an order well within the accuracy

limits of most of the analytical procedures concerned. The same cannot be said for horse haemoglobin however, even though its molecular weight is not large in comparison with those of many proteins of biological interest. No analysist of experience would concede that a residue number of 40 was accurate to $\pm$ 0.5 residues. It was, of course, the failure to grasp this point, and to appreciate the inadequacies of some of the older analytical techniques, which helped to keep alive for so long the doctrine that the amino-acid residues of proteins were present in proportions which indicated a frequence of occurrence along the peptide chain.

POLYPEPTIDE CHAINS

An open polypeptide chain has a C-terminal and an N-terminal residue, so that the number of such chains in a protein can be found by estimating one or other of these residues. Cyclopeptides, assumed to be formed by peptide linkage between the terminal residues, may be present in some proteins (see below), but there is as yet no evidence for the existance of branched chains involving the ε-amino group of lysine or the ω-carboxyl groups of the dicarboxylic acids.

N-terminal Residues.

The facile dinitrophenyl (DNP) method of Sanger ([8]) is well established. The available data are not extensive enough for generalization, but twelve amino-acids have been shown to occupy an N-terminal position. Insulin (ox, pig, sheep) would appear to be unusual in that for a protein of such small molecular weight (12,000) it has four N-terminal residues (2 glycine, 2 phenylalanine). Other well characterized proteins containing more than one such residue per molecule are various haemoglobins (horse, donkey — 6 valine; human adult — 5 valine; ox, sheep, goat, 2 valine, 2 methionine), β-lactoglobulin (3 leucine) ([2]), α-chymotrypsin (alanine and isoleucine) ([9]), avidin (3 alanine) ([10]).

With certain proteins the DNP technique fails to reveal the presence of any N-terminal residue. As will be discussed later, this has been accepted as tentative evidence for cyclic structures in the proteins concerned, but it may be due, of course, to some steric effect similar to that noted by Porter ([11]) who was unable

to couple the full number of ε-amino residues of lysine in β-lactoglobulin until after the native protein had been denatured. It is to be hoped therefore that more attention will be given in the future to alternative procedures which may help to clarify the position. Several possible methods, designed primarily for stepwise degradation from the N-terminal residue, are noted in a later section, but to my knowledge the use of all but one of them with proteins has not been exploited. Another method, due to Bowman [12], would seem to be worth exploring. By reductive alkylation with formaldehyde the N-terminal residue is converted into the N, N-dimethyl derivative. After hydrolysis the dimethylamino-acids are easily separated from the unmethylated acids by their solubility in alcohol.

Evidence for cyclic structures.

Desnuelle and Casal [13] and Porter [14] when using the DNP method were unable to detect N-terminal residues in ovalbumin, suggesting among other possibilities a cyclic structure. Bailey [15] likewise could find no evidence for these residues in either tropomyosin or myosin and considered that the depolymerized units of these proteins may be cyclopeptides, a suggestion in keeping with the observation of Tsao, Bailey and Adair [16] that the axial ratio of the tropomyosin monomer is in agreement with that calculated for two α-keratin chains lying side by side. Equally impressive results have been obtained with chymotrypsinogen. Desnuelle, Rovery and Fabre [17] could find no evidence for N-terminal residues by the DNP technique, while Gladner and Neurath [18] found that carboxypeptidase failed to liberate any C-terminal amino-acid, suggesting that this particular protein is unlikely to possess any open peptide chains. All of these observations, of course, are susceptible to alternative explanations, yet it is clear that on present evidence the possible occurrence of large cyclic peptides in proteins must be given serious attention. Such structures, admittedly of much smaller molecular dimensions, are present in many antibiotics and in natural cyclic ketones such as civitone and muscone. It follows that the estimation of the number of N-terminal residues in a protein may not necessarily indicate the full number of peptide chains. Indeed Williamson and Passmann [19] have already suggested that the pepsin molecule contains one open chain

bearing an N-terminal leucine residue and one cyclic chain. The need for complimentary procedures, dealing quantitatively with the opposite end of the peptide chain, is apparent.

C-terminal residues.

In addition to the reservation mentioned above, the position with regard to C-terminal residues and the number of peptide chains may be further complicated by the presence of terminal amide residues. Such a possibility serves to emphasize how little experimental evidence has yet been adduced in favour of the so-called amide hypothesis, according to which the ammonia given in mild hydrolysis of proteins is derived exclusively from asparagine and glutamine residues in the molecule. Although direct evidence for the presence of these residues in edestin and gliadin respectively was obtained many years ago it remains true that for lack of a suitable method for their quantitative estimation the question of whether other amides, *e.g.* glycinamide, might not be present as terminal residues has really been left open. All of the chemical procedures described below can be regarded as still in the experimental stage, but even so only one of them is designed to face the wider issue just mentioned.

Nystrom and Brown ([20]) showed that lithium aluminium hydride would reduce, *inter alia*, carboxyl, amide and ester groups and it has been used by Fromageot *et al* ([21]) to reduce the terminal carboxyl groups in protein. On subsequent hydrolysis the original C-terminal residues should be present in the hydrolysate as β-amino-alcohols, amino-diols or hydroxy-amino-acids as the case may be. After appropriate separation these products were characterised by various means and estimated by reaction with periodate. Insulin, per 12 000 g gave about 2 moles of each of alaninol, γ-hydroxy-β-amino-butyric acid and ethanolamine. The first two substances are derived from C-terminal alanine and asparagine residue respectively; the origin of the ethanolamine required further investigation ([22]). Ovomucoid, per 28 000 g gave 0.8 moles of phenylalaninol.

Lithium borohydride is a milder reagent that lithium aluminium hydride and under appropriate conditions will reduce an ester group but not a carboxyl or amide group. It has accordingly been used by Chibnall and Rees ([23]) to determine the amide distribution as

well as the C-terminal residues in insulin methyl ester. On treatment the ω-carboxyl groups of glutamyl and aspartyl residues, and all C-terminal residues, are reduced to the corresponding hydroxy-compounds, but glutaminyl and asparaginyl residues remain unchanged. After hydrolysis of the reduced product it was possible to show by appropriate analysis that insulin, per molecule of molecular weight 12 000, contains 8 glutamyl, 6 glutaminyl, 4 asparaginyl and 2 C-terminal asparagine residues respectively (*c.f.* Sanger's results mentioned later). For this particular protein therefore the amide hypothesis, mentioned above, is valid. Moreover the data definitely exclude the presence of β-aspartyl and γ-glutamyl peptide linkages, though they do not exclude the possible occurrence of *iso*asparagine and *iso*glutamine residues. Neither of these two *iso*amides has however been shown to occur in natural products.

Chibnall and Rees had difficulty in interpreting the data for the C-terminal residues. Two equivalents of alaninol were found, corresponding to 2 C-terminal residues of alanine, bringing the results, thus far, into line with those of Sanger to be mentioned later. In addition, however, an equivalent of ethanolamine was found [*c.f.* Fromageot *et al* (21)] which might indicate glycine as a fifth C-terminal residue. Its origin however was traced to acyl migration in the combination -glycyl-seryl- of the phenylalanyl chain, the ester linkage thereby produced undergoing reductive cleavage to produce an ethanolamine residue at the end of one of the resulting peptides. The rearrangement appears to be associated with the disappearance of free carboxyl groups during the esterification or reduction and is presumably brought about by a change in the charge distribution in the neighbourhood of the labile linkages concerned (41). The mechanism of the change will be discussed later, but it is pertinent to remark here that neither of the two above-mentioned reduction procedures will give an unequivocal characterization of the C-terminal residues in a protein or peptide until such acyl migration can be prevented or at any rate controlled.

Carboxypeptidase will remove an amino-acid residue from the carboxyl end of a peptide chain, but the rate of hydrolysis will depend on the C-terminal residue concerned. According to Smith (24) terminal aromatic residues are most readily attacked, then, in order, come the longer, followed by the shorter aliphatic

residues. The action is slow on peptides with acidic or basic terminal residues, and almost inappreciable when proline or hydroxyproline occupy the latter position. The nature of the C-terminal residue is not however the sole factor which determines the sensitivity of the substrate, for an adjacent proline residue can greatly decrease the rate of cleavage of the terminal bond or even prevent it. In certain cases it has been found that the rate of hydrolysis can be enhanced by N-acetylation of the peptide [25] or by treatment with dinitrofluorobenzene [26]. Bearing these considerations in mind it is clear that the procedure can be no sure guide to the number of open peptide chains in a protein molecule. It can, nevertheless, be of use in revealing the presence of C-terminal residues, and examples of this are alanine and asparagine in the case of insulin [25] [27], leucine (or isoleucine) and tyrosine in that of DFP α-chymotrypsin [18] and leucine in that of lysozyme [28]. It is probable that the enzyme will be of greater use in the study of the structure of small peptides obtained from partial hydrolysates of proteins than in the study of intact proteins.

According to Akabori *et al* [29] when a protein is treated with anhydrous hydrazine under controlled conditions the C-terminal residue is liberated as the free amino-acid and the other residues are converted into hydrazides. Terminal asparagine and glutamine residues are the exception, as these also pass into hydrazides. The method has been applied to several proteins, but the results so far recorded are not encouraging. In the case of α-chymotrypsin, for instance, the C-terminal residues are stated to be glycine, alanine and aspartic acid, whereas, as mentioned above, Gladner and Neurath [18] find that leucine (or isoleucine) and tyrosine are set free by carboxypeptidase. The procedure is a simple one and therefore of potential value, but it needs further investigation.

THE COMPONENT PEPTIDE CHAINS AND THEIR SEPARATION

It is surprising how little information is available about the way in which the component peptide chains of a protein are held in the molecule. The only covalent linkages definitely known to occur in proteins are the disulphide bridge of cystine and the peptide

bond. Cystine residues are apparently present in most proteins and it might well be thought that their function was that of a binder but the evidence for this is not always conclusive. That the residues operate as such in insulin is clear from the work of Sanger ([30]), who was able to release the component peptide chains by oxidative cleavage of the disulphide bridges with performic acid.

When other proteins are considered from the same point of view the question of how many sulphydryl and disulphide residues are present in the molecule of the native and denatured product presents difficulties due to uncertainty in analysis ([31]). It may be conceded perhaps that on present evidence β-lactoglobulin contains the necessary number of disulphide residues to bind the three open chains suggested by N-terminal assay, but the same is not true for various species of haemoglobin, where, for instance, that of horse would require at least 5, and of ox at least 3 disulphide bridges per molecule. A peptide bridge, presumably formed by linking the ε-amino group of lysine with the ω-carboxyl group of one of the dicarboxylic acids is a possibility, but according to Porter and Sanger ([32]) all the ε-amino groups of the lysine residues in the various haemoglobins they examined were free. A like result was obtained by Porter ([11]) with β-lactoglobulin and several species of serum globulin, though in some cases it was necessary to denature the protein before the full number of groups concerned could be estimated. Such evidence must however be treated with reserve, as it is based on the assumption that the methods used for estimating lysine and ε-DNP-lysine are accurate to one residue in thirty or more, which is doubtful. Other types of covalent linkages (thio-esters, imide bridges, etc.) have been suggested in the past, but no evidence for them has been forthcoming and one is led perforce to consider secondary valence interactions.

It will be recalled that Sörensen ([33]) advanced the view that many proteins are systems of sub-molecular components linked by residual valencies and capable of reversible dissociation. A discussion on the reversible dissociation of proteins and peptides is outside the scope of the present Report, but one may instance the conflicting evidence for 6,000 or 12,000 as the minimum molecular weight of insulin [for references see ([5])] and the recent work of Cheesman ([34]) suggesting that the molecule of methaemoglobin dissociates into about eight, and that of bovine serum albumin

into about sixteen kinetic units when adsorbed at a benzene-water interface at pH 6.8.

The field is an important one to which scant attention has so far been given. If, indeed, the large molecular weights that have been assigned to many proteins of biological interest are realities, and the products that the chemist handles are systems of small units or polymers of particle weight about 10,000 or less then, unless our present methods of terminal group assay are misleading us, cyclic structures must be more prevalent than we have been so far prepared to concede. Further research is clearly called for, because it will be necessary to characterise the simpler units before studies on amino-acid sequences become worth while. For the same reason a re-examination of the sulphydryl and disulphide contents of a number of proteins that might be of use in structure studies is urgently needed.

In the case of insulin the two component chains studied by Sanger were so unlike in composition that he had no difficulty in separating them by simple means. With other multi-chain proteins the problem of separating the component peptide chains, be they open or cyclic, has yet to be faced and, as Dr. Synge emphasizes in his Report, it is one that invites attention.

Amino-acid Sequences in Peptide Chains.

After being in the doldrums for a decade or more the determination of the amino-acid sequence in peptide chains became a fruitful line of enquiry when Consdon *et al* ([35]) showed that the simple peptides obtained from gramicidin S on partial acid hydrolysis could be separated by paper chromatography and ionophoresis. After appropriate analysis they were able to infer that these peptides arise by the breakdown, without rearrangement, of the structure (-α-[*l*-Val]-*l*-Orn-*l*-Leu-*d*-Phe-*l*-Pro), occurring once or twice in a closed peptide chain. The simple techniques employed by these workers have been extended by many others and have recently been brought to a high degree of perfection by Sanger and his colleagues. All but the most recent work has been fairly extensively reviewed ([2]) ([36]) and as most of the procedures employed are of proved value the Reporter does not feel that they call for the extended comment which has been given to the various methods of end-group assay and stepwise degradation that are in an early stage of development.

Insulin (mol. wt. assumed to be 12,000). By the DNP technique Sanger [8] found 2 N-terminal residues of glycine and 2 of phenylalanine. It was concluded that the molecule contained four open peptide chains held together by S-S bridges and by oxidative cleavage of the latter with performic acid it was possible to separate two main fractions, an acidic fraction A and a basic one B [30].

Fraction B contained all the amino-acids of insulin except isoleucine and had phenylalanine N-terminal residues. By partial hydrolysis of DNP-insulin and also of the DNP derivative of fraction B it was possible to separate and estimate several DNP-peptides. All of the DNP-phenylalanine peptides were derived from the sequence Phe. Val. Asp. Glu. and from the yields it was deduced that all the N-terminal phenylalanine residues in insulin were combined in this sequence. Similar results were obtained with the ε-DNP-lysine peptides, which were present in the sequence Thr. Pro. Lys. Ala. From the results it was concluded that there was only one type of phenylalanine chain in insulin. Partial hydrolysis of fraction B gave a large number of peptides and in interpreting their structure the fact that six residues (aspartic acid, serine, arginine, threonine, proline and lysine) occur only once in the chain was a great advantage. Difficulty was experienced with the aromatic peptides, which were more closely linked in the chain than might have been anticipated. From the results obtained by acid and alkaline hydrolysis the presence of the following sequences could be deduced

Phe. Val. Asp. Glu. His. Leu. $CySO_3H$. Gly.
Thr. Pro. Lys. Ala.
Gly. Glu. Arg. Gly.
Tyr. Leu. Val. $CySO_3H$. Gly.
Ser. His. Leu. Val. Glu. Ala.

Only two of the five aromatic residues are accounted for, and to place them it was necessary to investigate many of the larger peptides from hydrolysates obtained with pepsin, trypsin and chymotrypsin before Sanger and Tuppy [37] were able to show that the complete sequence in the phenylalanyl chain was as follows:

Phe. Val. Asp. ($-NH_2$) Glu. ($-NH_2$) His. Leu. (CyS-).
Gly. Ser. His. Leu. Val. Glu. Ala. Leu. Tyr. Leu. Val.
(CyS-). Gly. Glu. Arg. Gly. Phe. Phe. Tyr. Thr. Pro.
Lys. Ala.

Fraction A was subject to somewhat similar treatment. All of the N-terminal glycine residues in insulin were combined in the sequence Gly. Ileu. Val. Glu. Glu. and the conclusion was drawn that insulin contains only one type of glycyl chain. The following sequences were deduced from the characterization of the peptides given on acid hydrolysis.

1) Gly. Ileu. Val. Glu. Glu. $CySO_3H$. $CySO_3H$. Ala (N-terminal sequence).
2) Ser. Leu. Tyr. Glu. Leu. Glu. Asp. Tyr. $CySO_3H$.
3) Ser. Val. $CySO_3H$.
4) $CySO_3H$. Asp.

These accounted for all the amino-acids present in the fraction, but it was not possible to deduce an unequivocal structure solely from these because the large number of cysteic acid residues made interpretation difficult and the lability of the peptide bonds involving the amino groups of serine residues meant that no peptide was present in which these bonds were intact. Hydrolysis with pepsin, chymotrypsin and papain provided a large number of longer peptides, although only one of these was needed for the final characterization. The peptide obtained in highest yield with pepsin had glycine as the N-terminal residue and also contained [$CySO_3H$, Glu. Ser. Ala., Val., Ileu., Leu]. On partial hydrolysis with acid this gave the following simple peptides, all of which can be derivated from the N-terminal sequence 1.

$CySO_3H$. $CySO_3H$. Ala.	Ileu. Val. Glu. Glu.
Glu. $CySO_3H$.	Val. Glu.
Glu. Glu.	Ileu. Val. Glu.
$CySO_3H$. Ala.	

Two other peptides, Ser. Val. and Ser. Val. $CySO_3H$, could be referred to sequence 3, and Ser. Leu. to sequence 2. As the peptide obtained on peptic hydrolysis contained no tyrosine, the leucine in the peptide Ser. Leu. must be the C-terminal residue, and the sequence of the main peptide

Gly. ILeu. Val. Glu. Glu. Glu. $CySO_3H$. $CySO_3H$. Ala. Ser. Val. $CySO_3H$. Ser. Leu.

From this, and from the fact that there are only four cysteic acid

residues, Sanger and Thompson ([25]) deduced that the glycyl chain had the following sequence:

> Gly. ILeu. Val. Glu. Glu. (-NH_2). (CyS-). (CyS-). Ala. Ser. Val. (CyS-). Ser. Leu. Tyr. Glu. (-NH_2). Leu. Glu. Asp (-NH_2). Tyr. (CyS-). Asp. (-NH_2).

The amide groups were allocated by comparing the ionophoretic rates of peptides from enzymic digests containing aspartic and glutamic aicd residues. Sanger and his colleagues have been satisfied that they have not been misled in their deductions by possible inversion of sequence in dipeptides produced during hydrolysis with dilute acid, or by the enzymic synthesis of new peptides from the products of digestion: they claim that all the peptides they have characterized will fit the sequence of the peptide chain concerned.

This pioneer work of Sanger is a brilliant achievement. Although based on a variety of analytical procedures, some of which are, in the usually accepted sense, non-quantitative, it is interesting to note that a molecule consisting of these two particular chains joined by three disulphide bridges, which has a theoretical molecular weight of 5,732, would give values for amino-acid analysis very close indeed to those found by Harfenist and Craig ([5]). The data would, of course, fit the larger, four chain, molecule of 11,464 equally well.

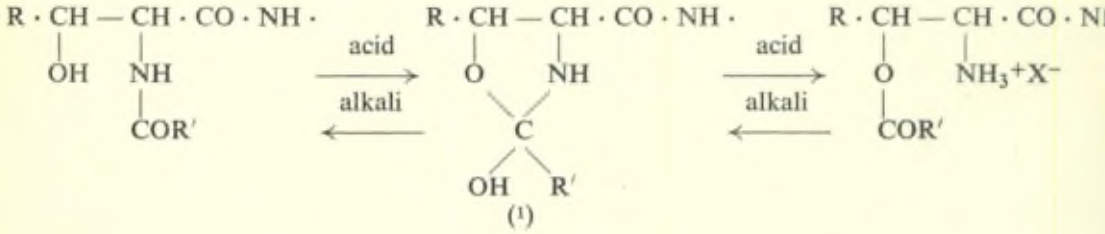
Apart from a natural desire to see such work confirmed there are, however, many reasons why alternative methods for determining amino-acid sequences are required.

Insulin would appear to be unusual among the proteins of low molecular weight in that the molecule contains more than one open peptide chain, each of which has a fairly small number of amino-acid residues. Such other proteins of this class as are suitable for study, *e.g.* lysozyme, ribonuclease and myoglobin, have, according to DNP analysis, only one open chain of 120-140 residues per molecule, a frightening array when considered from the standpoint of Sanger's procedures. What is needed in these cases is a preliminary cleavage of the chain in some specific and predictable way. It may well be that this could be effected by enzymes; nevertheless a chemical method involving a group which is not directly concerned in the peptide linkage would be more satisfactory, as it would eliminate the possibility of concomitant non-specific hydrolysis. The aliphatic hydroxyl group is a case in point.

In the penultimate stage of development of amino-acid analysis of proteins confidence in the results achieved was engendered by the agreement obtained with various procedures based on different principles, because errors inherent in one would not necessarily be present in another. At the present stage of development in the determination of amino-acid sequence in peptide chains the same dictum holds. The shortened chains mentioned above may well become amenable to Sanger's treatment, but even so it will be more satisfactory to have available alternative methods such as those which can affect a stepwise degradation by cleavage of a terminal residue.

SPECIFIC FISSION OF PEPTIDE BONDS

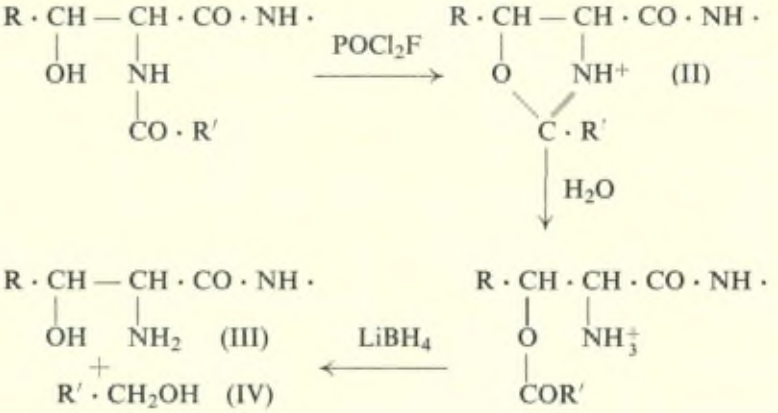
Bergmann and his collaborators showed that it was possible to bring about an N-acyl to O-acyl transformation in benzoyl derivatives of serine and postulated that the oxazoline and perhaps also the hydroxyoxazolidine structures were intermediates. The more recent work of Phillips and Baltzly [38] suggest that in acidic media the latter (I) is the more probable.



(I)

Elliott [39] treated silk fibroin with 97.5% H_2SO_4 at 21° for 3 days and found 62% of the N-peptidyl bonds involving serine were converted into the O-peptidyl form. After blocking the liberated amino group by acetylation the ester bond was readily hydrolysed by dilute alkali at room temperature. Fission by such means however was not completely selective.

Bailey [40] has used $POCl_2F$ to bring about the re-arrangement *via* the intermediary oxazoline (II)



With insulin, 50% of the N-peptidyl bonds involving serine and threonine were converted into the O-peptidyl form. As esterification had been shown ([41]) to be conducive to rearrangements of this type at certain of the labile linkages the treatment was applied to insulin methyl ester. In this case a near quantitative conversion was obtained, as measured by the increase in amino-N. Reductive cleavage was then brought about by $LiBH_4$ to give a series of peptides bearing terminal β-hydroxyalkylamide residues (III, IV). The method shows promise, especially as the carboxyl groups in the protein will have disappeared leaving the cleaved peptide fragments so much the more basic and therefore almost certainly amenable to separation by chromatography on the ion exchange resin 1RC-50 ([52]).

It must be recognised that any such specific fission is but one step in the elucidation of the structure of the intact peptide: the order in which the fragments were originally present in the chain constitutes a problem that calls for further research.

STEPWISE DEGRADATION OF PEPTIDE CHAINS

For a stepwise degradation to be successful all the reactions leading to the removal of the terminal residue must proceed quantitatively and without change in the remainder of the peptide chain. At the stage of development through which we are passing these ideal conditions will not be readily attained with proteins, if only for the reason that model synthetic substractes used in pioneer

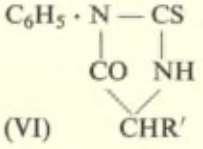
experiments may differ profoundly in solubility properties and general charge distribution from peptides obtained from natural sources.

Degradation from the N-terminal residue.

The procedure of Edman ([43]) is a logical extension of that of Bergmann *et al* ([44]). When a phenylthiocarbamyl derivative of a peptide (V) is treated with glacial acetic acid saturated with hydrogen chloride it is split into a 3-phenyl-2-thiohydantoin (VI) and a peptide with one residue less than the original. Under proper conditions the reaction is quantitative and almost instantaneous even in the cold.

$$C_6H_5 \cdot NH \cdot CS \cdot NH \cdot CHR' \cdot CO \cdot NH \cdot CHR'' \cdot CO \cdot$$

(V)

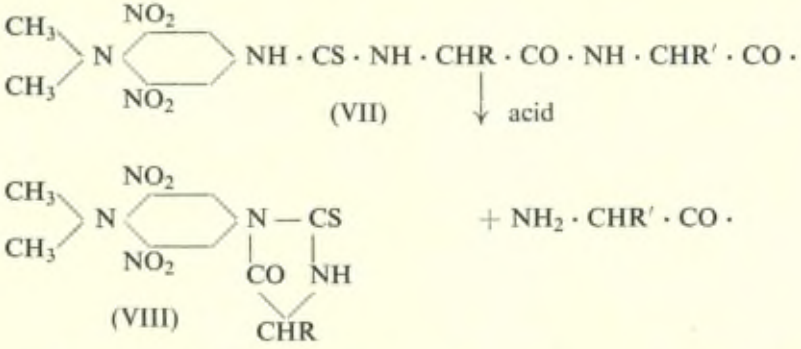


$$+ NH_2 \cdot CHR'' \cdot CO \cdot$$

(VI)

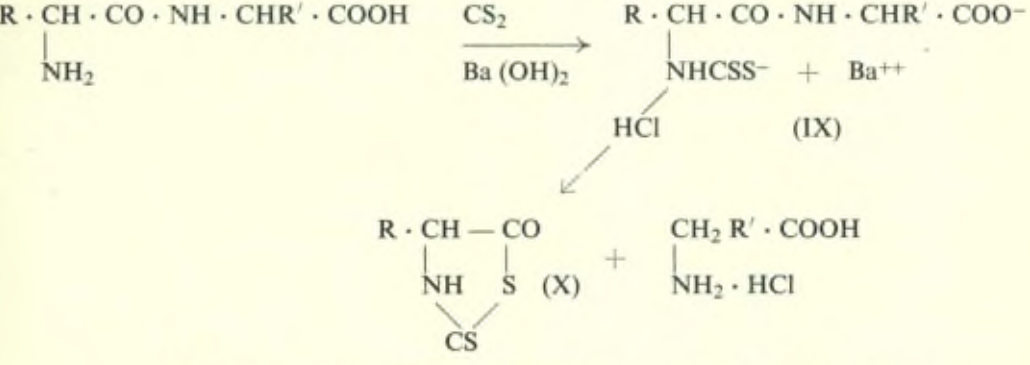
Ottesen and Wollenberger ([44a]) effect the condensation in aqueous dioxane and follow the course of the reaction, in which one hydrogen ion is liberated per ammonium group, by means of an autotitrator. Cleavage of the carbamyl peptide is then brought about at pH 1 by heating to 75° for 1 ½ — 15 hours. The modification has already been successfully applied to the peptides liberated in the transformation of ovalbumin to plakalbumin. In cases when the carbamyl peptide is insoluble the suggested splitting in aqueous 0.6 — 1.2N HCl at 36° in the presence of guanidine ([45]) may be useful.

A further extension by Reith and Waldron ([46]) has attractive possibilities. 3 : 5-dinitro-4-dimethylamino-phenyl-isothiocyanate (DDPT), which is deeply coloured, reacts in aqueous acetone at room temperature with the N-terminal group of peptides to give the corresponding thiocarbamyl (DDP) derivate (VII) in excellent yield.



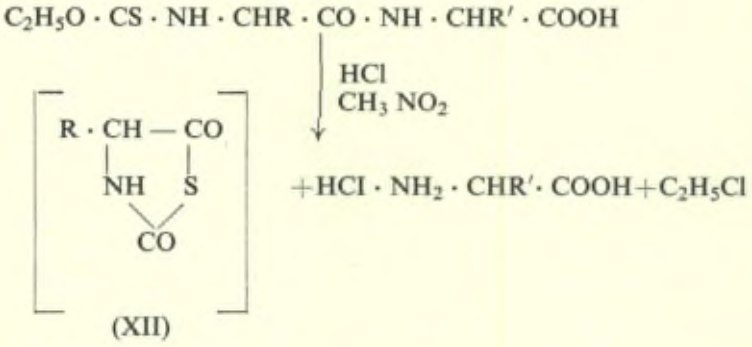
DDPT-alanyl-glycyl-glycine in aqueous acetic acid at 40° for 24 hours gave the DDPT-alanyl-hydantoin (VIII) in 82.5% yield. The degradation was repeated on the residual glycyl-glycine with success. The deeply coloured DDPT-hydantoins are amenable to chromatographic separation. The application to longer peptides with charged polar groups will be awaited with interest.

A novel procedure has been described by Levy ([47]). When peptides are titrated with cold aqueous alkali at constant pH in the presence of carbon disulphide they yield substituted dithiocarbamic acid salts. On acidification these cyclise (X) with fission of the terminal peptide link.



The employment of an aqueous medium is an advantage, but as one reaction follows the other in the same solution there is danger of contamination with unchanged peptide material. This indeed was the experience of Ottesen and Wollenberger ([48]) when they applied the procedure to the peptides which accompany the formation of plakalbumin from ovalbumin.

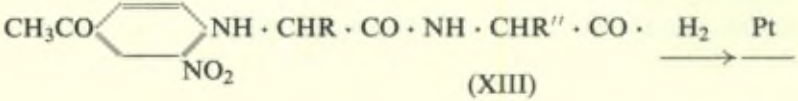
Kenner and Khorana [49] make the N-thioncarbethoxy-derivate of a peptide (XI) by condensation with alkyl methyl xanthates and treat this with nitromethane saturated with hydrogen chloride to cleave the N-terminal residue to what they think may be the substituted thiazolid-2 : 5-dione (XII).



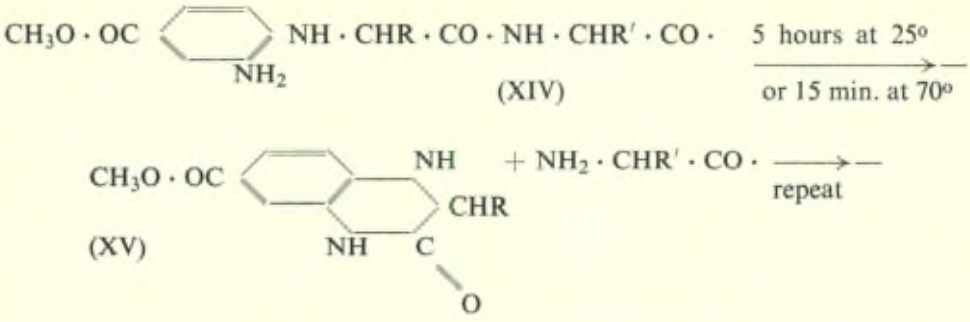
(XII)

The method has so far been applied only to simple peptides.

Holley and Holley [49a] have developed a procedure in which the sodium salt of an N-(4-carbomethoxy-2-nitrophenyl)-peptide (XIII), prepared in excellent yield with 4-carbomethoxy-2-nitrofluorobenzene, is reduced smoothly and quantitatively, using Adam's catalyst, to give the N-(2-amino-4 carbomethoxyphenyl)-peptide (XIV). Lactam formation, with removal of the terminal amino-acid as the 7-carbomethoxy-3, 4-dihydro-2 (1H)-quinoxalone (XV) takes place readily in the neutralised aqueous solution and is complete in 5 hours at room temperature or in 15 minutes at 70°. The dihydroquinoxalones are crystalline and can be identified by comparison with authentic samples prepared from amino-acids. An average yield of 84% per residue has been obtained from simple tri- and di-peptides. The authors state that modification of the procedure will be needed for residues of cystine, cysteine, methionine, tyrosine and the bases.

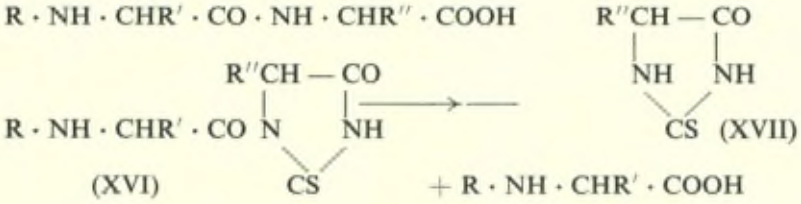


(XIII)



Degradation from the C-terminal residue.

In the original method of Schlack and Kumpf ([50]) an N-acyl peptide is converted into the acyl thiohydantoin (XVI) by the simultaneous action of acetic anhydride and ammonium cyanate under somewhat vigorous conditions. The thiohydantoin (XVII) corresponding to the C-terminal residue is then released by mild treatment with alkali.

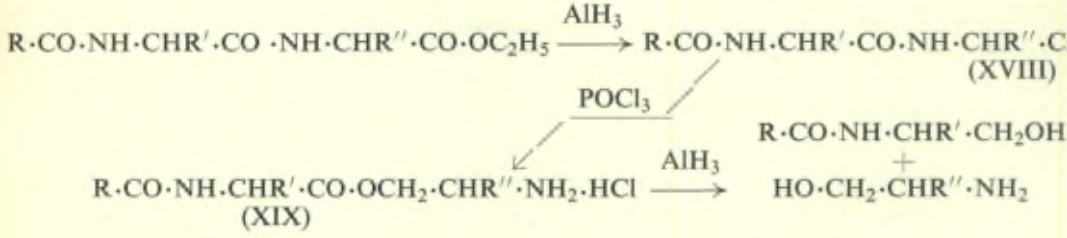


The procedure has been recently re-investigated by two groups of workers. Waley and Watson ([51]), using modern techniques, have shown that good yields can be obtained with simple peptides and that the thiohydantoin can be readily estimated after conversion to the amino-acid by HBr. When the method was applied to insulin however the yields were low, due in part to interaction between the thiohydantoin and the protein. Alanine was the only C-terminal residue identified. From studies of wool keratin and insulin Tibbs ([52]) also concluded that some modification of the procedure was necessary before it can be of use with proteins.

Khorana ([53]) has explored the reaction of carbodiimides with N-acyl peptides and subsequent alkaline degradation of the resulting acylureas as a possible method oı selective degradation from the C-terminal residue. Although mild conditions can be employed

throughout and the procedure was effective in certain cases the degraded peptide is often contaminated with the original peptide.

In the procedure under study by Bailey (54) the β-hydroxyalkylamide (XVIII) of the N-p-tosyl derivative of a peptide is the starting point. In the presence of reagents such as the phosphorus oxyhalides a rearrangement *via* the oxazoline (II) takes place to give the β-amino ester (XIX). The latter is then reduced with AlH_3 to yield the free amino alcohol and the residual peptide in a form ready for a further rearrangement.



With simple peptides yields of over 80% in the conversion and over 90% in the reductive cleavage have so far been obtained. Specific fission at serine and threonine residues (II), (III), (IV) is an extension of the procedure.

CONCLUSIONS

The foregoing survey shows a very welcome broadening of interest in analytical procedures for determining amino-acid sequences in proteins. The analyst still has a long way to go and the advance may be slow because of the exacting nature of the work ahead. Many ingenious organic reactions dealing with rearrangements and degradations have been discussed, but unless these can be made to proceed quantitatively or nearly so their use with long peptide chains derived from proteins is likely to prove disappointing. As long as progress is being made this is immaterial and it is pertinent to recall the analogous case of amino-acid analysis in which progress, slow at first, but extremely rapid when real interest was aroused, led finally to the adoption of quantitative methods which had little kinship with those used in the earlier work. Routine quantitative procedures must be the ideal in the present case and for some time to come emphasis in this field of research should be given to methods of analysis and their application rather than to the easier and

perhaps more entertaining pastime of determining sequences in nondescript peptides isolated from this or that protein. The latter will provide little more than grist to the mill of the annual reviewer, whereas the former will contribute towards a goal that no one would have thought possible of achievement a few years ago — the assignment of a chemical structure to a protein.

TABLE I

Number of Amino-acid Residues per molecule.

Amino-acid	Glycyl chain of insulin M.W. 2407	Horse myoglobin* M.W.Ca. 17,000	Horse haemoglobin* M.W. 68,000
Glycine	1	13	48
Alanine	1	15	54
Valine	2	6	50
Leucine	2	22	75
Isoleucine	1	22	0
Proline	—	5	22
Phenylalanine	—	5	30
Half-Cystine	4	0	2.5
Cysteine	—	0	3
Arginine	—	2	14
Histidine	—	9	36
Lysine	—	18	38
Aspartic acid	2	10	51
Glutamic acid	4	19	38
Amide NH_3	(4)	(8)	(36)
Serine	2	6	35
Threonine	—	7	24
Tyrosine	2	2	11
Tryptophan	—	2	5
Methionine	—	2	4.5
Total	21	143	543

* Data from Tristram (3).

Sir William Dunn Institute, University of Cambridge.

REFERENCES

(1) A.C. Chibnall, *J. Intern. Soc. Leather Trades' Chemists* 30, 1 (1946).
(2) F. Sanger, *Advances in Protein Chemistry*, 7, 1 (1952).
(3) G.R. Tristram, *Advances in Protein Chemistry*, **5**, 83 (1949).
(4) S. Moore and W.H. Stein, *J. Biol. Chem.*, **192**, 663 (1951).
(5) E.J. Harfenist and L.C. Craig, *J. Am. Chem. Soc.*, **74**, 4216 (1952); L.C. Craig, *CIBA Symposium* (1952) (in the press).
(6) G.L. Mills, *Biochem. J.*, **50**, 707 (1952).

(7) A.S. Keston, S. Udenfriend and R.K. Cannan, *J. Am. Chem. Soc.*, **71**, 249 (1949); A.S. Keston, S. Udenfriend and M. Levy, *J. Am. Chem. Soc.*, **72**, 748 (1950).
(8) F. Sanger, *Biochem. J.*, **39**, 507 (1945).
(9) P. Desnuelle, M. Rovery and C. Fabre, *CIBA Symposium* (1952) (in the press).
(10) H. Fraenkel-Courat and R.R. Porter, *Biochem. J.*, **49**, lxxvii (1951).
(11) R.R. Porter, *Biochim. et Biophys. Acta*, **2**, 575 (1948).
(12) R.E. Bowman, *J. Chem. Soc.*, 1349 (1950).
(13) P. Desnuelle and A. Casal, *Biochim. et Biophys. Acta*, **2**, 64 (1948).
(14) R.R. Porter, *Biochem. J.*, **46**, 473 (1950).
(15) K. Bailey, *Biochem. J.*, **49**, 23 (1951).
(16) T.C. Tsao, K. Bailey and G.S. Adair, *Biochem. J.*, **49**, 27 (1951).
(17) P. Desnuelle, M. Rovery and C. Fabre, *Comp. rend.*, **233**, 1496 (1951).
(18) J.A. Gladner and H. Neurath, *Biochim. et Biophys. Acta*, **9**, 335 (1952).
(19) M.B. Williamson and J.M. Passmann, *Federation Proc.*, **11**, 311 (1952).
(20) R.F. Nystrom and W.G. Brown, *J. Am. Chem. Soc.*, **69**, 1197, 2548 (1947).
(21) C. Fromageot, M. Jutisz, D. Meyer and L. Pénasse, *Biochim. et Biophys. Acta*, **6**, 283 (1950).
(22) C. Fromageot and M. Jutisz, *CIBA Symposium* (1952) (in the press).
(23) A.C. Chibnall and M.W. Rees, *Biochem. J.*, **48**, xlvii (1951); **52**, iii (1952).
(24) E.L. Smith, *Advances in Enzymol*, **12**, 191 (1951); *CIBA Symposium* (1952) (in the press).
(25) F. Sanger and E.O.P. Thompson, *Biochem. J.* (1953) (in the press).
(26) E. Waldschmidt-Leitz and K. Gauss, *Ber.*, **85**, 352 (1952).
(27) J. Lens, *Biochim. et Biophys. Acta*, **3**, 367 (1949).
(28) A.R. Thompson, *Nature*, **169**, 495 (1952).
(29) S. Akabori, K. Ohno and K. Narita, *Bull. Chem. Soc. Japan*, **25**, No. 3 (1952); S. Akabori and K. Ohno, 11e *Congrès International de Biochimie, Résumés des Communications*, *47* (1952).
(30) F. Sanger, *Biochem. J.*, **44**, 126 (1949).
(31) H. Neurath, J.P. Greenstein, F.W. Putnam and J.O. Erickson, *Chem. Rev.*, **34**, 158 (1944).
(32) R.R. Porter and F. Sanger, *Biochem. J.*, **42**, 287 (1947).
(33) S.P.L. Sörensen, *C. R. Lab. Carlsberg*, **5**, 1 (1930).
(34) D.F. Cheeseman, *Biochem. J.*, **50**, 667 (1950).
(35) R. Consdon, A.H. Gordon, A.J.P. Martin and R.L.M. Synge, *Biochem. J.*, **41**, 596 (1947).
(36) C.H. Li and J.I. Harris, *Annual Rev. Biochem.*, **21**, 603 (1952).
(37) F. Sanger and H. Tuppy, *Biochem. J.*, **49**, 463, 481 (1951).
(38) A.P. Philips and M.E. Baltzly, *J. Am. Chem. Soc.*, **69**, 200 (1947).
(39) D.F. Elliott, *Biochem. J.*, **50**, 542 (1952).
(40) J.L. Bailey, (unpublished work).
(41) A.C. Chibnall and M.W. Rees, *CIBA Symposium* (1952) (in the press).
(42) C.H. W. Hirs, S. Moore and W.H. Stein, *J. Biol. Chem.*, **195**, 669 (1952).
(43) P. Edman, *Arch. Biochem.*, **22**, 475 (1949); *Acta Chem. Scand.*, **4**, 283 (1950); *CIBA Symposium* (1952) (in the press).
(44) M. Bergmann, E. Kann and A. Miedely, *Ann.*, **458**, 56 (1927).
(44a) M. Ottesen and A. Wollenberger, *Nature*, **170**, 801 (1952).
(45) H. Fraenkel-Courat, *Federation Proc.*, **11**, 214 (1952).
(46) W.S. Reith and N.M. Waldron, *Biochem. J.* (1953) (in the press).
(47) A.L. Levy, *J. Chem. Soc.*, 404 (1950).
(48) Quoted by K.U. Linderström-Lang, Lane Memorial Lectures (1952), Stamford University Press.
(49) G.W. Kenner and H.G. Khorana, *J. Chem. Soc.*, 2076 (1952).
(49a) R.W. Holley and A.D. Holley, *J. Am. Chem. Soc.*, **74**, 5445 (1952).
(50) P. Schlack and W. Kumpf, *Z. Physiol. Chem.*, **154**, 125 (1926).
(51) S.G. Waley and J. Watson, *J. Chem. Soc.*, 2394 (1951).
(52) J. Tibbs, *Nature*, **168**, 910 (1951).
(53) H.G. Khorana, *J. Chem. Soc.*, 2081 (1952).
(54) J.L. Bailey, *Biochem. J.*, **52**, iv (1952) (unpublished work).

Discussion du rapport de M. A. C. Chibnall

M. Bigwood. — In his chapter dealing with amino acid analysis of proteins, Professor Chibnall referred to the outstanding developments following Martin and Synge's paper on chromatographic method of analysis. He pointed namely to the quatitative potentialities of Moore and Stein's ion exchange procedure. We have used this method in my laboratory with several coworkers (*) and in collaboration with Stanford Moore himself, as we started using his technique in Brussels in 1950-51. In papers going now into press we have described certain technical adjustments of his procedure to the usual conditions of our European laboratories, and the results so obtained. We have endeavoured to determine as completely as possible the quantitative composition in ninhydrine reacting nitrogenous constituants of :

A) Protein materials as pure as possible; the data being obtained with an approximation usually of the order of $\pm$ 5 %; the size of the sample available for all of the essential operations required being of the order of 50 to 200 mg.

B) Natural media hydrolysed in bulk, that is to say without extraction of the nitrogenous material prior to the hydrolysis and analysis of the latter.

A) *Purified proteins.* Examples of diagrams of chromatograms and of the corresponding tables of composition are presented, including the analysis of samples of a cristallised preparation of papain : of a Bence Jones protein; of a mixture of γ-globulins obtained from mixed human blood (a sample electrophoretically pure as far as the γ-group of globulins is concerned, supplied through Prof. E. J. Cohn's courtesy); of human chorionic gonadotrophin; of a pathological γ cryoglobulin; of a sample of hyaluronic acid containing a peptide fraction. These examples are presented

(*) J. P. Dustin, R. Crokaert, E. Schram, P. Soupart, J. Close, E. L. Adriaens and others still.

in view to underline some of the potentialities of the method; the total recovery of the nitrogenous constituants approaches closely 100 % of the total nitrogen content of the product when it is formed exclusively of ninhydrin reacting substances. A deficit in this respect is indicative of the presence in the preparation of other nitrogenous constituants. The sum of the amino acid residues as compared to the dry weight of the sample may approach or fall below 100 % according to whether the substance contains or not other constituants than the peptide fraction proper. The comparison of the recovery by weight and of the nitrogen recovery contributes to ascertain whether the non peptide fraction is or is not nitrogenous. Some strongly bound water may perhaps interfere however in the correct interpretation of such comparisons, even when the analyses refer to the dry weight of the samples as it was the case in our measurements. Some of the protocoles are presented in view to show how ninhydrin reacting nitrogenous constituants other than amino acids and ammonia, such as glucosamine for instance, can be detected and quantitatively estimated in certain protein molecules.

B) Total amino acid content of more complex natural media. Although this category of materials brings us rather apart from the subject of pure proteins, diagrams of chromatograms and the corresponding tables of composition in amino acids of various materials are presented, such as the analysis of known mixtures of free amino acids in presence of a large excess of carbohydrate, or natural media such as various foods or feeds (milk, meat, cereals, hay, seedcakes, fodder beet, cassava meal, etc.). These data are presented in view to show (a) that in suitable conditions of dilution of the sample in an excess of the hydrolysing acid solution, hydrolysis can take place with carbohydrate and other constituants present besides protein, without interference on their part with the accuracy of the determination of the amino acid content; and (b) that Moore and Stein's procedure lends itself to such operations. The sulphur containing amino acids are to be excepted however, with regard to this general statement. A method is described for determining accurately the cystin content chromatographically on ion exchange columns by converting it to cysteic acid prior to the hydrolysis.

Methionine is difficult to determine accurately. Data are presented showing that this amino acid may, although not always, be

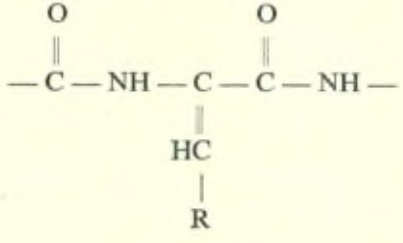
partly or completely oxidised to its sulfoxide and that this may take place not only when proteins are hydrolysed in the presence of other constituants such as carbohydrates, but also when the material submitted to hydrolysis corresponds to proteins purified as completely as possible. Both methionine and its sulfoxide can be determined quantitatively with Moore and Stein's method, although certain difficulties envolved ni these two determinations need however further investigation.

Finally, the data presented show examples of the possibility of identification and quantitative determination of unexpected constituants among the amino acids found in the products submitted to analysis. A specific example of this is given by the case of ornithine found in certain mealy preparations of cassava roots.

M. Havinga. — (1) Is it possible to separate peptide chains from each other in their intact state, when they are linked together through their side chains by peptide bonds, in stead of -S-S- linkages ?

(2) If a glycine residue were combined, as a peptide bond, with a cyclic aromatic structure, it would absorb light in the U. V. spectrum. This raises the eventual possibility of splitting peptide chains photochemically at the places where amino acid residues linked with aromatic rings function as light absorbing systems. This process has been studied in Leiden, using various synthetic substances with peptide bonds as model systems. Although these studies seem to indicate that the presence of tyrosine or phenylalamine residues is indeed favorable for photochemical splitting of nearby peptide bonds, there is a serious complication in the fact that ultraviolet light of wave length 2,000-3,000 A causes non specific splitting of peptide bonds. What is the opinion of Prof. Chibnall on the possibility of coupling the — CH_2 — groups of glycine residues to coloured aromatic residues ?

In case one could obtain products of the type :



One might cherish the hope that the conjugation of the aromatic light absorbing system (R) with the peptide bonds should enable splitting of these bonds by rather weak illumination with visible light.

M. Chibnall. — (1) To the best of my belief there is at present no experimental evidence for the suggestion that the peptide chain in any protein are linked together through their side groups by peptide bonds. Nor have any synthetic models been made, so that your query must remain unanswered for lack of information.

(2) The CH_2 group in glycine is very stable and I should think that chance of an effective substitution of the type you want is remote.

M. Synge referred to the rather dogmatic acceptance of the peptide theory. Recent observations on the structures of antibiotic peptides and related natural compounds have revealed a wealth of different types of linkage. Particularly striking are the recent data of Stoll, Hofmann and Petrzilka, *Helv. Chim. Acta*, **34**, 1544, 1951, on the ergot alkaloids.

Here are observed :

(i) Residues of α-hydroxy-α-amino acids.

(ii) A residues of *L* proline which by inversion yields *D* proline after acid hydrolysis.

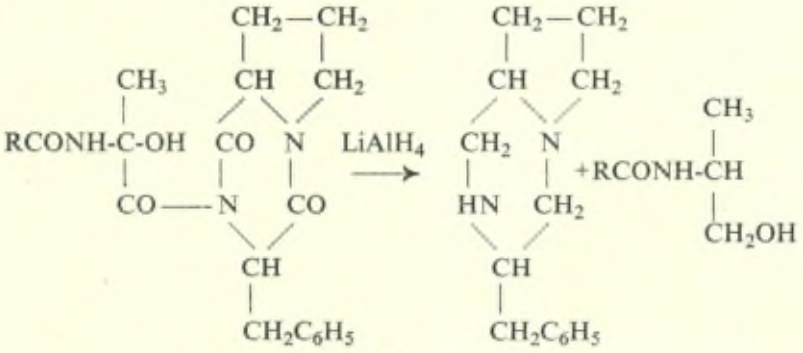
(iii) A carboxyl group in the ortho form, linked partly in ester, partly in amide linkage and presenting a new centre of asymmetry in the molecule.

(iv) The first substantiated structure of the « cyclol » type, originally postulated for proteins by Wrinch, and rejected by Pauling and Niemann (*J. Amer. Chem. Soc.*, **61**, 1860, 1939), mainly on energetic grounds.

Dr. Synge further asked Prof. Pauling whether it would be possible for closed peptide chains of, say 30 amino-acid residues, to be associated through secondary valence interactions to give molecular complexes that would resist dissociation by urea, guanidine, acid, alkali, etc. Such aggregations might well be difficult to distinguish with present techniques from molecules consisting of a single peptide chain.

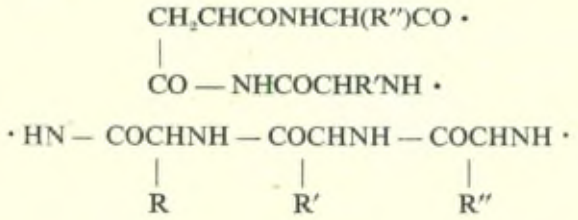
M. Karrer. — In connection with the report of Professor Chibnall and the comments of Dr. Synge, I wish to refer once again

to the paper of Stoll and coworkers, who have showed that diketopiperazine structures may occur in natural polypeptides. A tripeptide of ergotamine consists of a derivative of pyruvic acid, of phenyl-alanine and proline. Its structure has been proved by reduction with Li Al H_4 which gives a piperazine derivative.



This is the first case of a diketopiperazine derivative that has been discovered as a part of a natural polypeptide. It is possible that similar structures occur in natural proteins or that polypeptide chains are connected by carboxyl-groups.

Which encroach in peptide-groups of another chain in the following way :



M. Chibnall. — (1) It is, of course, possible that the dicarboxylic acid residues could function as connecting links in this way. The protein would then contain carboxyl groups that are neither free nor amidised so that on application of amid distribution procedure that Rees and I are exploring we should find in the hydrolysate of the reduced protein an equivalent amount of dicarboxylic acid over and above that required to satisfy the amid nitrogen. Such a finding, however, would not be specific for Prof. Karrer's suggested linkage : it would be given e. g. by an imide bridge between two dicarboxylic acid residues.

(2) With regard to Dr. Synge's question, the evidence from enzymatic work points to a large predominance of peptide structures, altough Stoll's work must be taken into consideration. « I think like most protein chemists that I rely, perhaps too heavily, on the early work on the splitting of proteins by proteolytic enzymes : Linderström-Lang, I believe, found that he could account for about 95-97 % of the assumed peptide bonds present. If other structures should be present, as suggested from Stoll's work, it may be possible to discover these when methods of step-wise degradation can be carried out easily and smoothly. »

(3) I should like to ask Prof. Pauling if he could express an opinion about the work of Cheeseman — in which dissociation was observed of methaemoglobin and of bovine serum albumin into eight and sixteen units respectively.

M. Pauling. — I feel that we have preserved an open mind about the peptide bond versus other structures; but I believe that, although other structures may be present in proteins, the peptide chain surely predominates.

In 1939, Niemann and I pointed out that the great instability of the cyclol link, 10 kcal/mole, eliminates the cyclol theory of protein structure. This argument is not invalidated by Stoll's report of a molecule containing an anthocarboxylate group. Such an unstable group may occur occasionally; but we may be confident that such a highly unstable structure would not occur for all or nearly all of the amino-acid residues in a protein.

In answer to Dr. Synge's question, I would say that I think that hydrogen bonds and other weak interactions might hold protein molecules of molecular weight about 5,000 together to form a stable larger molecule. Experiments such as Landsteiner's, with azoprotein antigens and hapten-homologons antibody, indicate that a bond strong enough to resist thermal agitation can be formed by a small haptenic group, such as p-azobenzenearsonate ion, and the antibody protein that can get in contact with it. Aggregates of such small molecules should, however, dissociate rather easily under the influence of hydrogen ion, hydroxide ion, urea, guanidine, and similar reagents.

It need not be assumed that the several polypeptide chains that are present in some protein molecules are held together by covalent bonds (S — S bonds, side-chain peptide bonds, etc.); they may

be held together by hydrogen bonds and other rather weak intermolecular forces, which cooperate to prevent their easy separation. If the chains are intertwined, their separation might be achieved only with difficulty.

Horse hemoglobin and hemocyanin are wellknown examples of molecules which consist of smaller molecules held together by weak forces. I doubt, however, that the molecules of methemoglobin dissociates into eight and that of bovine serum albumin into sixteen small molecules, as suggested by Cheeseman, and mentioned by Chibnall.

M. Anson. — (1) It is difficult to obtain large peptides from a protein by digestion, even by the use of a single specific enzyme, because even a single enzyme attacks enough linkages to produce small peptides. I should like to suggest that it might be worth while trying to decrease the number of linkages broken by a single enzyme by working not at the usual optimum pH for the activity of the enzyme, but at the extreme pH at which the enzyme is active, or by changing the conditions of digestion in some other way.

The optimum pH for the activity of a proteolytic enzyme depends on the structure of the substrate. The big protein molecule is really a mixture of substrates which are attacked most rapidly at pH's which surely differ but which differ to on unknown extent. At the extreme pH at which the enzyme is active, the relative ratio at which the different substrates are attacked must therefore differ, although to an unknown extent, from the relative rates at the optimum pH. Only experiment can decide whether or not the use of an extreme, unfavorable pH for digestion will suppress the splitting of some linkages enough to make possible the practical production of big peptides for analytical purposes.

There are other ways in which digestion conditions might be altered to change the relative rates at which different linkages in a protein are broken. One might add salts, including heavy metal salts. Or one might change the protein chemically by reactions with side chains. Again, only experiment can tell how much the specificity of the enzyme can be increased by such devices. One can only suggest that the theoretical possibilities exist, and that they have not yet been adequately explored experimentally.

There are probably large peptides formed at the beginning of

any normal digestion of proteins; large peptides which are then digested further more rapidly then the original protein is attacked. In any case, one usually obtains in practice, as Tiselius has pointed out, a mixture of unchanged proteins and small peptides. That it might be possible to isolate transitory large peptides formed at the beginning of digestion was indicated (but not proven) by my observation that a large part of the very first products of the digestion of hemoglobin could be precipitated by concentrated but not by dilute trichloracetic acid. Ways might be found of trapping the big peptides continuously as they are formed. Only when such peptides are available will it be known how homogeneous they are.

(2) Chibnall pointed out that the value for the number of end amino groups obtained by Sanger's method is sometimes uncertain. I should like to emphasize that much can be learned about reactions with protein amino and carboxyl groups from what is known about the reactions of other protein groups, especially the SH groups which have been most thoroughly studied.

A reaction with a protein group usually takes place more readiliy if the protein is denatured than if it is native. Porter has shown this to be true even of the Sanger reaction. A protein group reaction, however, often does not go smoothly to completion, even if the protein is denatured. To carry the reaction to completion, one must then find a stronger reagent, or a solvent which makes the groups more reactive, or both. And it is always desirable to show that one gets the same result despite variations (in the right direction) of both the reagent and the solvent. It would make the result of a Sanger estimation of end amino groups more convincing in any doubtful case, if the same result were obtained by the use of a stronger reagent, or by the use of guanidine hydrochloride in stronger concentration or at higher temperature. The merse use of guanidine hydrochloride does not guarantee a correct SH titration. Two reagents may give the same titration when applied to one protein in guanidine hydrochloride solution. When applied to another protein with less reactive SH groups, however, the stronger reagent may give higher titration than the weaker reagent.

M. Chibnall. — (1) Dr. Anson's suggestions are interesting. It may be that highly purified enzymes could be used under special

conditions to give the results he suggests, but unless the procedure was tested first of all on a long peptide chain of definitely known constitution the element of uncertainty would remain.

(2) The D. N. P. technique has been used under varying conditions but so far without any significance in the results obtained.

M. Neuberger. — As Dr. Anson has already pointed out, there is no certainty that all end groups react completely with the rather bulky reagents in use at present. In this connection it may be mentioned that the comparison of the number of basic groups of egg albumin as revealed by potentiometric titration (carried out by Kekwick and Cannan) with the number of basic amino-acid residues as shown by analytical techniques then available suggested the presence of several free α-amino groups. Is this deduction still valid in the light of more recent results ?

With respect to acyl migration, the question may be asked whether the oxazolidine structure is possibly performed in the insulin molecule.

M. Chibnall. — (1) I think that Dr. Neuberger is recalling an old suggestion of mine that the total cationic groups as revealed by Cannan's titration curves are in excess of those that can be accounted for in terms of the bases, so that there may be some terminal amino groups present. Since then, the analytical value for lysine has been raised, and the difference is now no longer realised.

(2) We have considered the possibility that oxazoline or hydroxyoxazolidine rings pre-exist in insulin, but we have no valid evidence. Insulin itself does not seem to undergo any change except the reduction of the disulphide linkages when passed through our reduction procedure.

M. Linderström-Lang. — (1) There is Bergmann's old hypothesis of the existence of oxazoline rings in proteins. He pointed out that thiazoline rings might also occur in some cases as well as thio ether or other bonds in general. Esters linkages involving the hydroxyl group of serine should also be considered in view of the esterase activity of proteolytic enzymes demonstrated by Neurath and his school.

(2) It is desirable to define what is meant by « salt-linkages ». Originally (Speakman) these linkages were considered to be « anhy-

drous ». Therefore in soluble proteins they have to be protected against water in the interior of the molecules. This type of bond is assumed to exist in Eyring and Stearn's theory of denaturation. A different kind of « linkage » is obtained when two hydrated ions of opposite sign are in close proximity in the protein molecule. Both types of linkage may exist in protein. In β-lactoglobulin they seem to be absent (or present in too small amounts to be measurable).

M. Neuberger. — What was suggested is not the presence of a large number of hydrogen bonds between charged groups in side chains, but the existence of electrostatic forces between ionized groups carrying charges of opposite type. It is possible that in a few cases real hydrogen bonds may also exist.

M. Putzeys. — L'importance des groupes latéraux ionisables au point de vue de la structure des protéines est mise en évidence par l'existence de zones de stabilité particulièrement nettes lorsqu'il s'agit de molécules très complexes telle que l'hémocyanine de l'escargot de Bourgogne (Helix Pomatia).

L'hémocyanine purifiée possède une zone de stabilité comprise entre pH 4,5 et 6,8 à 25°. (Déterminée par la diffusion moléculaire de la lumière.) Une légère variation du pH en dessous de 4,5 dissocie la molécule en demi-molécules et celles-ci à leur tour se dissocient en molécules plus petites par abaissement plus avancé du pH. Au delà de 6,8 la dissociation donne immédiatement des molécules de l'ordre de 1/8 ou 1/16 du poids moléculaire de la molécule entière. Ces dissociations qui sont produites aussi bien par abaissement que par élévation du pH ne peuvent être expliquées semble-t-il que par rupture de ponts hydrogène établis entre les groupes latéraux. Il est probable que l'ionisation produite par addition d'acides ou de bases intéresse tout d'abord les groupes basiques et acides non engagés dans des ponts hydrogène, mais ceux qui sont ainsi liés doivent cependant fixer ou libérer des protons lorsque l'addition d'acide ou de base a été suffisante puisque la molécule à ce moment se dissocie.

Qu'il s'agisse réellement de phénomènes où les forces électrostatiques sont les forces agissantes est encore montré par l'effet calcium décrit par J. Brosteaux [*Naturwissenschaften* (1937) **25**, 249]. En présence de 0,01 m $CaCl_2$ la zone de stabilité de l'hémo-

cyanine est étendue jusque pH 9,5. Enfin la dissociation sous l'influence des électrolytes, telle que l'a montrée Brohult [*Journ. Phys. Colloid Chem.* (1947) 51, 206] (et nous avons confirmé ces résultats par la diffusion moléculaire de la lumière), dissociation qui est pratiquement complète dès une concentration 1 molaire en chlorures alcalins, montre elle aussi l'importance de l'ionisation des groupes latéraux et des effets électrostatiques sur les liaisons entre les sous-molécules qui constituent la molécule complexe de l'hémocyanine.

M. Bragg. — It would be of great interest to know the shape of the two half molecules into which the haemoglobin molecule can split. Have any methods of investigation given an indication of the shape of the half-molecules ? Could they in any way be crystallized as half-molecules, or do they always come together again in pairs under such conditions ?

M. Theorell. — Since dissociation of hemoglobin molecules in water solution into half molecules occurs only in dilute solution, and is reversible, I am afraid the halves will come together again on crystallisation.

M. Karrer. — The report of Professor Chibnall and the discussions following it have shown that the problem of the chemical constitution of a protein molecule is not less difficult and complicated than the cristal-structure of such substances. We have certainly still much to learn of the structure and the chemical behaviour of proteins.

Electrophoresis, Chromatography and Related Physical Methods in Application to Future Requirements of Protein Chemistry

by R. L. M. Synge

The application of electrophoretic and chromatographic methods to the study of peptides and amino-acids has been one of the main factors leading to the recent remarkable advances in our understanding of protein chemistry, which are assessed in the Report of Dr Chibnall. He has dealt with the chief results of :—

(*a*) amino-acid analysis of complete hydrolysates;

(*b*) the detection and identification of free functional groups;

(*c*) partial degradation of proteins and identification of resulting peptides, etc. as a means of ascertaining sequences of amino-acid residues and, ultimately, complete chemical structures.

The electrophoretic and chromatographic methods used in such studies have in general been concerned with separations of smaller molecules (mol. wt. less than about 500). The knowledge gained as a result of the use of these methods forms a theoretical background by means of which a suitable method can be selected for almost any analytical problem requiring separation of such smaller molecules. Thus in this field, while only an infinitesimal proportion of the results has been obtained compared with those that will be required for increasing our understanding of protein structure and metabolism, the techniques already developed look as if they will be adequate, without major changes in principle, for most of the analytical problems presented by this group of compounds.

Sanger (1952), in an excellent review on the study of the arrangement of amino-acid residues in proteins, has given an outline of the main analytical methods that have so far proved useful. He has also shown how future studies will necessitate the analytical separation of much larger molecules. At the same time, very much more serious attention will have to be paid to isolating in the pure state the starting material which is to be used for structural studies. In the present paper it is intended :—

(*a*) to outline how differences in the structures of the smaller molecules are made use of in selecting an appropriate analytical technique;

(*b*) to discuss how far existing analytical procedures for smaller molecules may be extended to separations of larger molecules, and what new analytical principles show promise of being useful for such separations. The experience already gained in work on the purification of intact proteins is clearly of enormous importance for development of this field (cf. the Reports of Prof. Desreux and Dr Pedersen);

(*c*) to show that the analyst who undertakes work in these fields, whether he studies new compounds by existing procedures or undertakes the development of new analytical procedures, is not acting merely as technician to the structural organic chemist. He is also accumulating data on the molecular interactions of biologically important chemical groupings. This knowledge will sooner or later serve to bring much closer together the structural and functional sides of biochemistry, throwing new light on our understanding of the working parts of the living organism at the molecular level.

CONSIDERATIONS DETERMINING CHOICE OF ANALYTICAL METHODS WITH SMALLER MOLECULES

Electrophoresis

Molecular properties. The migration of small molecules in an electric field may be predicted with fair confidence, the mobility being the resultant of the driving force, determined by the net electric charge on the molecule, and the frictional resistance, determined by the bulk and shape of the molecule. Rules for selecting

the best pH for any given separation on the basis of the dissociation constants have been given by Consden, Gordon & Martin (1946). In general, charge differences due to differences in dissociation constant have so far been exploited as the main basis for separations, but there have been some striking separations of molecules of different bulk and similar charge [e. g. acidic peptides in hydrolysates of wool (Consden, Gordon & Martin, 1949; Consden & Gordon, 1950) and of insulin (Sanger & Thompson, 1953); oxidized dextrins (Norberg & French, 1950)]. Molecules not themselves charged may acquire charges in the presence of a complexing agent (see below).

Apparatus and procedures. Svensson (1948) gave an exhaustive review of preparative electrophoretic procedures and Tiselius (1952 *b*) has reviewed some more recent developments (see also Grassmann, 1951). In general, the grosser separations, depending on presence or absence of a charge on the molecules, are most conveniently carried out in diaphragm cells equipped with suitable membranes while separations depending on finer differences in mobility are best done in a buffered medium having such a greater concentration than the mixture being analysed that the pH and electric field (determined by conductivity) remain substantially constant throughout the medium. For preparative purposes the liquid may be immobilized by incorporation in a gel, paper or powder, etc., and changes in composition due to electrode reactions are prevented from reaching the region where the analysis is proceeding. A very interesting recent extension of such procedures is to arrange flow of the liquid perpendicular to the electric field, so that a continuous separation of the components of a mixture can be obtained (Svensson & Brattsten, 1949; Grassmann & Hannig, 1950; Durrum, 1951; Brattsten & Nilsson, 1951). This promises great advantages when electrophoretic fractionation has to be done on a large scale.

The possibilities of « isoelectric fractionation » and, generally, the migration of ions in the absence of a swamping concentration of buffer do not seem yet to have been at all thoroughly explored (see Svensson, 1948, pp. 259-264).

Molecular properties. Nearly all chromatographic procedures so far described depend on differences in the equilibrium distribution of the molecules to be separated between moving and stationary phases in the chromatographic column. Repeated countercurrent equilibration enables quite small differences in distribution coefficients to be used for complete separations. The same applies to liquid-liquid extraction trains (and to liquid-vapour fractionating columns, although few of the compounds which concern the protein chemist are volatile).

There is an enormous variety of molecular properties that can be exploited for determining equilibrium distributions in different countercurrent systems. The manifold data accumulated in chromatographic and similar work, when subjected to quantitative treatment by competent physical chemists will throw much light on molecular interactions. At present few such treatments exist, and workers in this field use rather empirical ideas developed from the foundation that each chemical grouping or residue or radical in a molecule contributes an energy term for transfer between the two phases which is characteristic for itself and for the two-phase system in which the solute is distributed. This, of course, is the theoretical justification of Traube's Rule (see Brönsted, 1931). A simple mathematical treatment has been given by A. J. P. Martin (1949). Thus predictions of the effect of substitution with a known radical of an unknown molecule or of polymerization and so forth can be made with some confidence if data on analogous substitutions are available for the system being used. Reference should be made to the systematic data of Bate-Smith & Westall (1950) on flavonoid compounds, of Isherwood & Jermyn (1951) on sugars and related compounds and of Knight (1951; cf. Pardee, 1951) on peptides. With solid adsorbents these rules are less easy to apply than with liquid-liquid systems, as purely steric features of the solute may play a bigger role in determining its distribution. Competition with adsorbed solvent or with other substances may also exert a big influence.

Ion-exchange systems are in some ways the most predictable. Depending on the coulombic interactions of electrically charged groups they give readily most of the separations that can be made

electrophoretically. However, other molecular properties than charge of the solute usually influence these equilibria : thus the aromatic groupings in the most commonly used ion-exchange resins lead to a sequence of affinities for amino-acids and peptides of otherwise similar charge (Moore & Stein, 1951) which resembles that seen with charcoal (Tiselius, 1941) and to some extent that with aromatic solvents in two-phase liquid systems. In such ways the chromatographer rather empirically comes to know what system will be most promising for separating substances differing by a known feature of molecular structure, and is conversely able to make intelligent guesses as to chemical structures in unknown substances on the basis of their behaviour in a number of different chromatographic systems.

Possible steps towards increasing the specificity of the equilibria used in chromatography and towards extending the application of chromatography to larger molecules are discussed below.

Chromatographic operating procedures. Three main types of operating procedure are usually recognized (see « *Chromatographic Analysis* » — *Discuss. Farad. Soc.* No. 7, 1949; Tiselius, 1952 *a*) : (*a*) *frontal analysis;* (*b*) *elution development;* (*c*) *displacement development.*

Frontal analysis has found relatively little application for analytical purposes, although it can be useful in dealing with unknown mixtures. It is of course used in « stripping » solutions, as in desalting with ion-exchange columns, decolorizing with charcoal columns and so forth.

Elution development has had the widest application, since it can lead to complete separations of substances. Countercurrent distributions of substances are usually done in an analogous manner. Recently, however, considerably greater interest has been shown in *displacement development.* In the form originally proposed by Tiselius (1943) this gives in the effluent from the column successive zones of the substances separated following one another closely and often with a considerable overlap. Even here there may be advantages — thus elution development in the hands of Moore & Stein (1951) has given remarkable separations of amino-acids for quantitative analysis on the mg. scale while similar columns operated by displacement development in the hands of Partridge and colleagues (see Partridge & Brimley, 1952; Partridge, 1952) have

been used to handle g. quantities preparatively. A logical extension of displacement development is « carrier displacement » in which a series of displacing agents is used which have affinities for the stationary phase intermediate between those of the substances which it is wished to separate. This yields these substances uncontaminated by one another. Successes in elution from charcoal of amino-acids and peptides with a series of alcohols (Tiselius & Hagdahl, 1950; Hall & Tiselius, 1951; Li *et al.*, 1951) and of oligosaccharides, obtained in amylolysis, with successively increasing concentrations of ethanol in water (Whistler & Durso, 1950) should be mentioned here. Many of the successes obtained by chromatography with successive changes of solvent doubtless depend on similar displacement effects. For many purposes more convenient operation results from continuously changing the composition of the developing solvent (« gradient elution ») and interesting results with such methods have been described by Williams (1952) * and by Moore & Stein (private communication).

General

The foregoing paragraphs give in outline some considerations that apply in selecting or exploring methods for the separation of amino-acids and lower peptides. As already mentioned, Sanger (1952) has given a synopsis with references of the methods currently in use or proposed for these purposes. For amino-acid analysis (which, for the protein chemist, occupies a similar place to that of elementary analysis in general organic chemistry) two-dimensional filter-paper chromatography and the quantitative chromatographic procedures of Moore & Stein may already have assumed the respective roles of Lassaigne's tests and Pregl's quantitative microanalysis.

Isotopic tracers. The use of isotopic tracers promises to be of very great value in metabolic studies on proteins. Tracers have already done much to elucidate the pathways of building up and breaking down of animo-acids in living organisms. The work of A. A. Benson, M. Calvin and colleagues on the formation of amino-acids in photosynthesis is particularly striking. However, most of

* See also Alm, R. S., Williams, R. J. P. and Tiselius, A. *Acta chem. Scand.* **6**, 826 (1952); Alm, R. S. *Acta chem. Scand.* **6**, 1186 (1952).

the data obtained so far with tracers on the incorporation of amino-acids into proteins is rather crude and difficult to interpret. It is clear, however, that as structural knowledge of proteins increases, tracer methods will allow metabolic studies to follow close on its heels. Although the work of Anfinsen & Steinberg (1951) on aspartic acid incorporation in different parts of the egg albumin molecule is rather ambiguous, owing to uncertainty whether aspartic acid or asparagine residues were measured, their more recent study (Steinberg & Anfinsen, 1952) on the alanine residues is not. Almost the only comparable data are those of Muir, Neuberger & Perrone (1952) on different types of valine residue in haemoglobin, which lead to very different conclusions. The existence of adequate analytical methods is obviously of prime importance for metabolic tracer studies.

Tracer methods may also have direct analytical application for identification or estimation in dealing with amino-acids or peptides. ^{131}I has proved of great value, after metabolic incorporation, for studying the natural I-containing amino-acids, and the isotope-dilution methods of S. Udenfriend and colleagues, using labelled *p*-iodobenzenesulphonyl groups, have had both quantitative and qualitative applications.

IMPROVEMENTS IN ANALYTICAL METHODS REQUIRED FOR THE EXTENSION OF STRUCTURAL STUDIES OF PROTEINS

In this section are discussed some recent or foreseeable developments of analytical technique which seem promising for satisfying the requirements imposed by the progress of structural studies of proteins. These requirements are simply expressed as the need to effect separations of considerably larger molecules, both for obtaining more homogeneous starting material and for dealing with larger split-products derived therefrom. As Sanger (1952) has pointed out, the first logical step in the study of partial hydrolysates of proteins, where the mixture of lower peptides obtained by prolonged random hydrolysis cannot be interpreted unequivocally, is to use more specific agents for splitting the protein into relatively few large polypeptide fragments. The two kinds of large fragments obtained

thus from insulin by Sanger and colleagues were so different that their separation was fairly easy. In other cases, however, it is likely to be much more difficult. As will be seen, some of the difficulties arise from the fact that it is possible for a structural difference between larger molecules to have less influence on their properties than the same structural difference in smaller molecules would have. Other difficulties arise from the fact that the molecules to be separated are large ones, for which the methods developed to separate small molecules are simply not suitable.

Electrophoresis

Despite a considerable current increase of interest in electrophoretic methods, these have by no means yet been fully exploited. There is a great need for technical improvement of available methods, since not nearly full advantage has been taken of known differences in mobilities. The study of reversible spreading of boundaries in the Tiselius apparatus or of failure of boundaries to reach a steady state has shown that most protein preparations giving a single boundary are still electrophoretically heterogeneous (see grouped references on Electrophoretic Heterogeneity). The electrophoresis-convection methods proposed by Kirkwood and colleagues should be useful for effecting on a preparative scale some of the more difficult separations. They have already shown very encouraging results in fractionation of antibodies from the γ-globulin fraction of sera (see grouped references on Electrophoresis-Convection).

Electrophoresis has great value for studying natural complexes in solution without disturbing their physical environment. There are also great possibilities for using extraneous complexing agents for extending the range of substances amenable to electrophoretic separations; some of these are discussed in the following section.

Chromatographic processes

Large molecules at equilibrium in two-phase systems tend to distribute themselves mostly to one phase or the other, as can be deduced from simple thermodynamic considerations (Brønsted, 1931). With proteins, this behaviour leads to useful differential solubility effects which Cohn and others have very fully exploited for protein separations, using both salts and organic solvents,

and which Prof. Desreux describes in his Report. Changes in the activity of a protein molecule in solution resulting from changes in the composition of the solvent should produce corresponding changes in its adsorption behaviour and similar properties even before any precipitation occurs. It was these ideas that led Tiselius (1948) to propose chromatography of proteins using « salting-out adsorption ». Here advantage can be taken of countercurrent repetition of the process, which is less easily effected with bulk precipitation. There are now available several reviews of the considerable work on protein chromatography which has accumulated, particularly in recent years (Zechmeister & Cholnoky, 1941; Zechmeister, 1950; A. J. P. Martin, 1950; Zechmeister & Rohdewald, 1951; Bull, 1952; Moore & Stein, 1952; Zittle, C. A., 1953, *Adv. Enzymol.* **14,** 319; Schwimner, S. & Pardee, A. B., 1953, *Adv. Enzymol.* **14**, 375). In most experiments proteins have tended to move fast or not at all, and the column arrangement has therefore had little advantage over salt fractionation or the batch adsorption processes which proved so valuable for the early work on purification of enzymes by R. Willstätter and others. However, quite a number of cases have now been found where proteins give intermediate band rates on chromatograms. The selection of solvents, adsorbents, etc. that can be used is sharply limited by the necessity of preventing protein denaturation; this may well not apply to some of the larger molecules arising from protein breakdown. Moore & Stein (1952) have done well to emphasize the importance of ascertaining whether the band positions are due to displacement or association effects with other components in the system by rechromatographing the purified material and observing its behaviour on a fresh column.

Much the same considerations apply to methods depending on distribution of these larger molecules in systems having two liquid phases. The composition of the phases must not be very different if partition coefficients not too far from unity are aimed at. Multicomponent systems not too far from the plait-point are therefore necessary. Pioneer work in this field has been done by Brooks & Badger (1950), Martin & Porter (1951), Harfenist & Craig (1951, 1952) and Porter (1953). It is interesting that Martin & Porter found that ribonuclease is adsorbed on the liquid-liquid interface of their chromatograms, whereas insulin is not (Porter, 1953).

The number of variables capable of exploitation for the elaboration of chromatographic systems to handle these larger molecules is very large, and it is already clear that more or less empirical search for suitable systems will yield results. There seem also to be no upper limits to the size of particle which can be handled by adsorption methods at solid surfaces or at liquid-liquid and gas-liquid interfaces. Thus several workers have described adsorption or chromatography of viruses or of larger intracellular particles (see grouped references on Adsorption of Larger Particles).

However, there are also possibilities of improving chromatographic systems by increasing their specificity, taking advantage of particular structural features of the molecules to be separated. The use of buffers for making use of differences in dissociation constants is already well known in chromatography. More recently, we have had « solubilizing agents » — e. g. the use of fat-soluble organic acids to form salts with hydrophilic bases (Titus & Fried, 1947, 1948; Plaut & McCormack, 1949; O'Keeffe, Dolliver & Stiller, 1949; Swart, 1949), the addition of borate to chromatographic systems to differentiate *cis*- and *trans*-hydroxyl groups (Khym & Zill, 1951; Annison, James & Morgan, 1951) and the use of complexing agents such as citrate in the chromatography of the rare earths on ion-exchange resins (cf. Spedding, 1949). With silica gel it seems that the adsorption specificity may be modified by precipitation in the presence of different compounds (Dickey, 1949; Curti & Colombo, 1952). The selective power of charcoal has also been modified, although less spectacularly, by doping it in various ways (Steenberg, 1944; Synge & Tiselius, 1949; Weiss, 1949; Dalgliesh, 1952). It is only a short step from here to the study of the specificity of proteins themselves as adsorbents. Here the work described by Klotz & Ayers (1952), Cohn *et al.* (1952) and others on interaction of proteins and small molecules has revealed interesting possibilities. Antibodies dissolved in the aqueous phase of liquid-liquid columns could be expected to have differing affinities for substances chemically related to the antigen and to affect their distribution coefficients; with haptens there would be no need for actual precipitation to occur. In this connection the experiments of Campbell, Luescher & Lerman (1951) are of great interest; here antigen protein was chemically coupled to cellulose and this was used as a specific adsorbent for antibody. Similarly enzymes could

act as modifiers of equilibria involving substances related to their substrates. Starch grains have in fact for long been used as adsorbents in the purification of amylases (Starkenstein, 1910; Holmbergh, 1933; Schwimmer & Balls, 1949). There is obvious scope for extension of this principle.

Many of these suggested means of improving the specificity of the equilibria which serve as a basis for chromatography can equally be used in electrophoretic systems when the resulting complexes are charged. As examples can be quoted the separation of sugars as complexes with boric acid (Consden & Stanier, 1952) * and of starchlike polysaccharides as complexes with iodine-iodide (Mould & Synge, 1951). Fels (1951) has given electrophoretic evidence for the occurrence of peptide-trichloroacetic acid complexes in solution. It is quite surprising, in view of the classical part played by ion-migration methods in the formation of the ionic theory and in the study of complexes, that such methods have not been much more used.

Molecular-sieve effects in chromatography

It is now generally recognized that the grains of many of the working substances most favoured for chromatography — charcoal, cellulose, starch, silica gel, ion-exchange resins — are sponge-like structures having pores which may be of molecular dimensions. Smaller molecules can penetrate while larger ones cannot (for references see Mould & Synge, 1952). Such differences could be, and in some cases have been, made the basis of useful separations according to molecular size. Molecular-sieve effects may also be playing a useful role in the selective extraction of lower molecular-weight components from heterogeneous polymers (Desreux, 1949; Williams, 1952). However, in chromatographic colums the molecular-sieve effect usually works in the opposite sense to adsorption, leading to confusing results when a polymeric series is subjected to chromatographic analysis. The alternative is to use as chromatographic column a continuous gel structure rather than a coarse powder of gel particles in liquid. Here adsorption and molecular-sieve effects would work in the same sense. This was suggested

* See also Michl, H. (1952). *Monatsh.* **83,** 37; Michel, F. et van de Kamp. F. P. (1952). *Angew. Chem.* **64,** 607.

in discussions between Dr A. J. P. Martin, Prof. A. Tiselius and the author, in which it was proposed to use electroendosmosis under an applied potential for moving liquid through the gel. (For any reasonable length of column the hydrostatic pressure required to effect movement of liquid through the gel would disrupt the gel structure). Following this proposal Mould & Synge (1952) have realized useful separations of starch derivatives using collodion ultrafiltration membranes; this method seems capable of considerable extension in the field of polymer chemistry. Where organic solvents are required, a porous cellulose structure could be regenerated from the collodion by denitration. The effects reported were observed with uncharged molecules; with charged molecules electrophoretic effects would of course be superposed on those of molecular-sieving and adsorption. A given gel structure would only handle uncharged molecules or molecules of opposite charge; its penetrability by molecules of like charge will be low, and in any case these will be migrating in the direction opposite to the electroendosmotic stream.

Diffusion methods

It is a short step from considering molecular-sieve effects in chromatography to considering separation methods based on diffusion, either in free solution or through diaphragms having an element of selective permeability. Where the membrane has selective permeability, it will often be convenient to force the solution undergoing analysis through it. However, it is of interest that C. J. Martin (1896), in his original paper on ultrafiltration, sharply distinguished the behaviour of crystalloids when allowed to diffuse or when forced through an ultrafiltration membrane. Differences of concentration arising from differences of diffusion constant, observed when material crossed the membrane by diffusion, were absent with forced flow since the membrane showed little selective permeability for these substances.

For multiplying differences in diffusion behaviour it is not possible to use a simple countercurrent arrangement in which each stage approaches equilibrium since maximum enrichment factor for the transfer of material is obtained at the earliest stage, and it is not practical to allow equilibrium to be approached. Signer *et al.*

(1946) have described a multi-stage arrangement of diaphragm diffusion cells which should be suitable for many purposes. Brooks & Badger (1948) have studied the behaviour of collodion undergoing simple diffusion in a solution immobilized in filter paper, and their arrangement might prove useful for preparative work. However, the relationship of diffusion constant to particle mass gives rather poor enrichment factors.

Kirkwood & Brown (1952) have described an interesting arrangement by which the equilibrium distribution of protein in a concentration gradient of salt or other low-molecular substance is made the basis of fractionation by convection as in the Clusius-Dickel type of column. Incidentally, diffusion serves as the means of transport for securing the necessary concentration gradients in the serological precipitation methods in gel, in which interest has been revived by Oudin, Ouchterlony and others (see grouped references on Serological Reactions in Gels). These have made possible separations both of antigens and antibodies, making apparent complexities that had previously only been suspected.

CONCLUSIONS

The foregoing survey of foreseeable developments for the separation of higher peptides and proteins shows that, leaving aside entirely the possible introduction of quite new analytical principles, there is an enormous range of possibilities for improving the technique of the analyst in this field. It is not presumptuous for the analyst to claim that, at the present stage of development of protein chemistry, he has the most important contribution to make. Inside his field, there is much work that can be done with techniques already developed, but also much requiring new techniques. All this should be platitudinous, and probably does seem so to most readers, but biochemistry has been cursed with experimenters who learn a single technique and then flog it to death attempting to answer problems for which it can only give inconclusive answers. And these have to be supported by a long chain of flimsy assumptions and special pleading. The present Report deals only with one sector of the analytical problems confronting the protein chemist,

those of effecting separations by methods that are mainly physical. The analytical use of selective chemical substitutions and degradations, etc. has been dealt with by Dr Chibnall. The two sides of protein analysis must be developed simultaneously.

The physical chemist who has become involved with analytical problems of this kind cannot, however, help noticing that while working on these he is at the same time building up a body of knowledge about the interactions of molecules, particularly « secondary valence » interactions. From what is already known of these phenomena they have been assigned a very high importance in the working of living organisms. The analyst, starting with small molecules and their interactions, is collecting information about the behaviour of the component parts of proteins. The partition coefficients, vapour pressures in solution, slopes of adsorption isotherms and displacement effects observed in chromatographic operations all supply data on energies and specificities of molecular interactions, and the chromatographic method of observation is much simpler than those first developed for physicochemical measurements of this kind. Systematization of this knowledge is already becoming a pressing need, and will lead to a detailed understanding of the interactions of larger molecules. The extension of the analyst's work towards larger molecules, sketched in the present Report, will greatly accelerate this understanding.

The behaviour of the porous substances used in chromatography — cellulose, starch, rubber, etc. — is bound moreover to throw new light on their fine structure and behaviour at the molecular level. The electrochemical and permeability properties of gels, membranes and porous structures observed in electrophoretic studies, studies with ion-exchange materials or in electrokinetic work, have already enormously clarified our ideas on cellular physiology; the extension of analytical work in the directions I have indicated above is likely to lead to further advances in this knowledge.

Of course it is really inevitable that the process of finding out what things are made of shall also throw light on how they work. The idea that the « secret of life » would be revealed as a single discovery in some laboratory is now pretty well dead. But there has been a tendency to think that knowledge of the workings of living matter will come as some great theoretical synthesis in the future. I prefer to think that there will be no sudden apocalyptic

process but that, now that we have an increasingly sound basis of physicochemical theory, the patient piecing together of information will give us an ever more detailed and accurate picture of the nature of living things.

The Rowett Research Institute, Bucksburn, Aberdeenshire, Scotland.

REFERENCES

General

Anfinsen, C.B. & Steinberg, D. (1951), *J. biol. Chem.*, **189**, 739.
Annison, E.F., James, A.T. & Morgan, W.T.J. (1951). *Biochem. J.*, **48**, 477.
Bate-Smith, E.C. & Westall, R.G. (1950). *Biochim. biophys. Acta*, **4**, 427.
Brattsten, I. & Nilsson, A. (1951). *Ark. Kemi*, **3**, 337.
Brönsted, J.N. (1931). *Z. phys. Chem.*, Bodenstein Festband, p. 257.
Brooks, M.C. & Badger, R.M. (1948). *J. phys. coll. Chem.*, **52**, 1390.
Brooks, M.C. & Badger, R.M. (1950). *J. Amer. chem. Soc.*, **72**, 1705.
Bull, H.B. (1952). *Ann. Rev. Biochem.*, **21**, 179.
Campbell, D.H., Luescher, E. & Lerman, L.S. (1951). *Proc. Nat. Acad. Sci.*, **37**, 575.
Cohn, E.J., Surgenor, D.M., Schmid, K., Batchelor, W.H. & Isliker, H.C. (1952). *Disc. Farad. Soc.* Cambridge, Aug. 1952.
Consden, R. & Gordon, A.H. (1950). *Biochem. J.*, **46**, 8.
Consden, R., Gordon, A.H. & Martin, A.J.P. (1946). *Biochem. J.*, **40**, 33.
Consden, R., Gordon, A.H. & Martin, A.J.P. (1949). *Biochem. J.*, **44**, 548.
Consden, R. & Stanier, W.M. (1952). *Nature, Lond.*, **169**, 783.
Curti, R. & Colombo, U. (1952). *J. Amer. chem. Soc.*, **74**, 3961.
Dalgliesh, C.E. (1952). *Biochem. J.*, **52**, 3.
Desreux, V. (1949). *Rec. Trav. chim. Pays-Bas.*, **68**, 789.
Dickey, F.H. (1949). *Proc. Nat. Acad. Sci.*, **35**, 227.
Durrum, E.L. (1951). *J. Amer. chem. Soc.*, **73**, 4875.
Fels, I.G. (1951). *Ark. Kemi*, **3**, 347.
Grassmann, W. (1951). *Naturwiss*, **38**, 200.
Grassmann, W. & Hannig, K. (1950). *Z. angew. Chem.*, **62**, 170.
Hall, D.A. & Tiselius, A. (1951). *Acta chem. Scand.*, **5**, 854.
Harfenist, E.J. & Craig, L.C. (1951). *J. Amer. chem. Soc.*, **73**, 877.
Harfenist, E.J. & Craig, L.C. (1952). *J. Amer. chem. Soc.*, **74**, 4216.
Holmbergh, O. (1933). *Biochem. Z.*, **258**,, 134.
Isherwood, F.A. & Jermyn, M.A. (1951). *Biochem. J.*, **48**, 515.
Khym, J.X. & Zill, L.P. (1951). *J. Amer. chem. Soc.*, **73**, 2399.
Kirkwood, J.G. & Brown, R.A. (1952). *J. Amer. chem. Soc.*, **74**, 1056.
Klotz, I.M. & Ayers, J. (1952). *Disc. Farad. Soc.* Cambridge, Aug. 1952.
Knight, C.A. (1951). *J. biol. Chem.*, **190**, 753.
Li, C.H., Tiselius, A., Pedersen, K.O., Hagdahl, L. & Carstensen, H. (1951). *J. biol. Chem.*, **190**, 317.
Martin, A.J.P. (1949). *Biochem. Soc. Symposia*, **3**, 4.
Martin, A.J.P. (1950). *Ann. Rev. Biochem.*, **19**, 517.
Martin, A.J.P. & Porter, R.R. (1951). *Biochem. J.*, **49**, 215.
Martin, C.J. (1896). *J. Physiol.*, **20**, 364.
Moore, S. & Stein, W.H. (1951). *J. biol. Chem.*, **192**, 663.
Moore, S. & Stein, W.H. (1952). *Ann. Rev. Biochem*, **21**, 521.
Mould, D.L. & Synge, R.L.M. (1951). *Biochem. J.*, **50**, xi.
Mould, D.L. & Synge, R.L.M. (1952). *Analyst.*, **77**, 964.
Muir, H.M., Neuberger, A. & Perrone, J.C. (1952). *Biochem. J.*, **52**,, 87.
Norberg, E. & French, D. (1950). *J. Amer. chem. Soc.*, **72**, 1202.
O'Keeffe, A.E., Dolliver, M.A. & Stiller, E.T. (1949). *J. Amer. chem. Soc.*, **71**, 2452.
Pardee, A.B. (1951). *J. biol. Chem.*, **190**, 757.
Partridge, S.M. (1952). *Analyst.*, **77**, 955.
Partridge, S.M. & Brimley, R.C. (1952). *Biochem. J.*, **51**, 628.
Plaut, G.W.E. & McCormack, R.B. (1949). *J. Amer. chem. Soc.*, **71**., 2264.
Porter, R.R. (1953). *Biochem. J.*, **53**, 320.
Sanger, F. (1952). *Adv. Prot. Chem.*, **7**, 1.
Sanger, F. & Thompson, E.O.P. (1953). *Biochem. J.*, **53**, 353, 366.
Schwimmer, S. & Balls, A.K. (1949). *J. biol. Chem.*, **179**, 1063; **180**, 883.

Signer, R., Hänni, H., Koestler, W., Rottenberg, W. & Tavel, P. von (1946). *Helv. chim. Acta*, **29**, 1984.
Spedding, F.H. (1949). *Discuss. Farad. Soc.*, **7**, 214.
Starkenstein, E. (1910). *Biochem. Z.*, **24**, 210.
Steenberg, B. (1944). *Adsorption & Exchange of Ions on Activated Charcoal*. Uppsala : Almquist och Wiksells Boktryckeri AB.
Steinberg, D. & Anfinsen, C.B. (1952). *J. biol. Chem.* **199**, 25.
Svensson, H. (1948). *Adv. Prot. Chem.*, **4**, 251.
Svensson, H. & Brattsten, I. (1949). *Arkiv Kemi.*, **1**, 401.
Swart, E.A. (1949). *J. Amer. chem. Soc.*, **71**, 2942.
Synge, R.L.M. & Tiselius, A. (1949). *Acta chem. Scand.*, **3**, 231.
Tiselius, A. (1941). *Ark. Kemi, Mineral. Geol.*, 15B No. 6.
Tiselius, A. (1943). *Ark. Kemi, Mineral. Geol.*, 16A No. 18.
Tiselius, A. (1948). *Ark. Kemi, Mineral. Geol.*, 26B No. 1.
Tiselius, A. (1952*a*). *Endeavour*, **11**, 5.
Tiselius, A. (1952*b*). *Discuss. Farad. Soc.*, Cambridge, Aug. 1952.
Tiselius, A. & Hagdahl, L. (1950). *Acta chem. Scand.*, **4**, 394.
Titus, E. & Fried, J. (1947). *J. biol. Chem.*, **168**, 393.
Titus, E. & Fried, J. (1948). *J. biol. Chem.*, **174**, 57.
Weiss, D.E. (1949). *Discuss. Farad. Soc.*, **7**, 142.
Whistler, R.L. & Durso, D.F. (1950). *J. Amer. chem. Soc.*, **72**, 677.
Williams, R.J.P. (1952). *Analyst.*, **77**, 905.
Zechmeister, L. & Cholnoky, L. (1941). *Principles & Practice of Chromatography*. (trans. A.L. Bacharach & F.A. Robinson). London : Chapman & Hall.
Zechmeister, L. (1950). *Progress in chromatography* (1938-1947). London : Chapman & Hall.
Zechmeister, L. & Rohdewald, M. (1951). *Fortschr. Chem. org. Naturstoffe*, **8**, 341.

Electrophoretic Heterogeneity

Alberty, R.A. (1948). *J. Amer. chem. Soc.*, **70**, 1675.
Alberty, R.A., Anderson, E.A. & Williams, J.W. (1948). *J. phys. coll. Chem.*, **52**, 217.
Anderson, E.A. & Alberty, R.A. (1948). *J. phys. coll. Chem.*, **52**, 1345.
Baldwin, R.L., Laughton, P.M. & Alberty, R.A. (1951). *J. phys. coll. Chem.*, **55**, 111.
Brown, R.A. & Cann, J.R. (1950). *J. phys. coll. Chem.*, **54**, 364.
Hess, E.L. & Deutsch, H.F. (1948). *J. Amer. chem. Soc.*, **70**, 84.
Hess, E.L. & Deutsch, H.F. (1949). *J. Amer. chem. Soc.*, **71**, 1376.
Hoch, H. (1950). *Biochem. J.*, **46**, 199.

Electrophoresis-Convection

Brown, R.A., Shumaker, J.B., Cann, J.R. & Kirkwood, J.G. (1951). *J. Amer. chem. Soc.*, **73**, 4420.
Brown, R.A., Shumaker, J.B., Timasheff, S.N. & Kirkwood, J.G. (1952). *J. Amer. chem. Soc.*, **74**, 460.
Cann, J.R., Brown, R.A., Gajdusek, D.C., Kirkwood, J.G. & Sturgeon, P. *J. Immunol.*, **66**, 137.
Cann, J.R., Brown, R.A. & Kirkwood, J.G. (1949). *J. biol. Chem.*, **181**, 161.
Cann, J.R., Brown, R.A. & Kirkwood, J.G. (1949). *J. Amer. chem. Soc.*, **71**, 1609, 2687.
Cann, J.R., Campbell, D.H., Brown, R.A. & Kirkwood, J.G. (1951). *J. Amer. chem. Soc.*, **73**, 4611.
Cann, J.R. & Kirkwood, J.G. (1950). *Cold Spr. Harb. Symp. Quant. Biol.*, **14**, 9.
Cann, J.R., Kirkwood, J.G., Brown, R.A. & Plescia, O.J. (1949). *J. Amer. chem. Soc.*, **71**, 1603.
Kirkwood, J.G., Cann, J.R. & Brown, R.A. (1950) *Biochim. biophys. Acta.*, **5**, 301.
Mathies, J.C. (1952). *Science*, **115**, 144.

Adsorption of Larger Particles

Cochran, G.W. (1947). *Phytopathology*, **37**, 850.
Davenport, F.M. & Horsfall, F.L. (1948). *J. exp. Med.*, **88**, 621.
Delbrück, M. (1940). *J. gen. Physiol.*, **23**, 631, 643.
Muller, R.H. & Rose, H.M. (1952). *Proc. Soc. exp. Biol. Med., N.Y.*, **80**, 27.
Riley, V.T. (1948). *Science*, **107**, 573.
Riley, V.T. (1950). *J. nat. Cancer Inst.*, **11**, 199, 215, 229.
Riley, V.T., Hesselbach, M. L., Fiala, S., Woods, M.W. & Burk, D. (1949). *Science*, **109**, 361.

Serological Reactions in Gels

Becker, E.L. & Munoz, J. (1949). *Proc. Soc. exp. Biol. Med. N.Y.*, **72**, 287.
Ouchterlony, Ö. (1948). *Ark. Kemi, Mineral. Geol.*, 26B, No. 14.
Ouchterlony, Ö. (1949). *Ark. Kemi*, **1**, 43, 55.
Ouchterlony, Ö. (1949). *Acta path. microbiol. scand.*, **26**, 507.
Oudin, J. (1946). *C. R. Acad. Sci.*, **222**, 115.
Oudin, J. (1947). *Bull. Soc. Chim. biol.*, **29**, 140.
Oudin, J. (1948). *Ann. Inst. Pasteur*, **75**, 30, 109.
Oudin, J. (1949). *C. R. Acad. Sci.*, **228**, 1890.

Fractionnement et purification des protéines

V. Desreux et E. Fredericq

En raison de l'étendue du sujet et du grand nombre de travaux publiés les dernières années en ce domaine, nous nous sommes limités à brosser un tableau très général du problème posé, à signaler les tendances actuelles et à ne traiter en détail que certains aspects de la question.

I. — EXTRACTION DU MILIEU BIOLOGIQUE

Le premier stade dans tout isolement d'une protéine est son extraction du milieu biologique. La solubilisation d'une protéine dans un solvant donné dépend évidemment de ses caractéristiques moléculaires mais également de la labilité de la structure cellulaire dans laquelle elle est engagée.

L'énergie avec laquelle les protéines du muscle, par exemple, sont liées à l'édifice cellulaire étant très variable, on peut, par solubilisation fractionnée, réaliser des séparations déjà très nettes dès le premier stade du fractionnement. Une connaissance plus approfondie des caractéristiques d'une protéine (domaine de stabilité, susceptibilité enzymatique, etc.) permet actuellement de mieux préciser les conditions dans lesquelles son extraction doit être effectuée afin d'être aussi sélective que possible et afin de conserver autant que faire se peut son intégrité. Rappelons l'exemple devenu classique de l'isolement du trypsinogène et du chymotrypsinogène à partir du pancréas en milieu H_2SO_4 : 0,25 N, qui inhibe l'activation de ces précurseurs et dénature la plupart des protéines inertes (1).

Le fractionnement de protéines à activité protéolytique, telles les enzymes de la papayotine par exemple, peut être précédé d'une incubation au pH d'activité maxima, incubation qui entraîne une hydrolyse des protéines étrangères et simplifie donc le fractionnement ultérieur. De même, Kunitz purifie partiellement l'extrait obtenu par plasmolyse de cellules de levure en laissant le milieu s'autolyser pendant plusieurs jours dans des conditions bien déterminées [isolement de la pyrophosphatase ([2])]. On peut également en chauffant l'extrait obtenir une dénaturation préférentielle.

Ces procédés ne sont cependant pas toujours sans danger (voir paragraphe IV).

Il s'avère d'ailleurs nécessaire de mieux définir l'objectif de chaque fractionnement, ce qui permet ainsi de préciser les conditions d'extraction et ensuite celles du fractionnement. Les exigences du physico-chimiste sont en effet différentes de celles du biologiste; le premier est particulièrement intéressé à disposer de protéines aussi homogènes que possible alors que le second est amené de plus en plus à étudier les complexes natifs tels qu'ils existent dans la cellule. Les conditions opératoires seront donc très différentes suivant le but à atteindre.

La répartition des constituants biologiques dans les différents granules des cellules étant de mieux en mieux connue, il peut être utile et même parfois nécessaire de faire précéder le fractionnement moléculaire d'un fractionnement partiel ou total de granules par centrifugation, après homogénéisation du tissu. Falconer, Jender et Taylor ([3]) préconisent l'élimination des noyaux par centrifugation d'un homogénéisat de foie par exemple en solution concentrée de sulfate ammonique (0.8 saturé), milieu dans lequel toutes les protéines sont pratiquement insolubles et où les granules sont donc « stabilisés ». Un fractionnement plus poussé (mitochondries, microsomes, surnageant) est recommandable si la protéine à isoler est concentrée dans un type déterminé de granules [ex. : la cytochrome oxydase concentrée dans la fraction mitochondries : Schneider et Hogeboom ([4]), la phosphoglucomutase dans le surnageant : Hers, Berthet et de Duve ([5])].

La libération complète d'une protéine du complexe « granule » est dans certains cas très aisée, mais dans d'autres cas elle n'est obtenue que par application de moyens très énergiques [cas de la ribonucléase : H_2SO_4 : 0,25 N, Pirotte et Desreux ([6])].

Remarquons enfin que le choix du matériel de départ est très

important et que ce facteur, responsable fréquemment de l'hétérogénéité observée d'une fraction de protéine, a été trop souvent négligé. De même, les facteurs de stabilité sont encore trop mal connus et il y aurait grand intérêt à poursuivre des études dans cette voie.

II. — FRACTIONNEMENT

1) Aspect général du problème.

Les facteurs moléculaires à considérer dans tout fractionnement sont les dimensions, la charge électrique ainsi que la structure physico-chimique de la protéine. Dans certaines techniques nous ne mettrons à profit qu'un facteur intrinsèque, par exemple la charge en électrophorèse préparative: dans d'autres techniques, nous ferons appel à des propriétés beaucoup plus complexes comme la solubilité. Dans chaque cas nous pouvons modifier dans certaines limites la valeur des facteurs intrinsèques en agissant sur les facteurs extérieurs : température, pH, force ionique, constante diélectrique du milieu.

Nous pouvons d'autre part agir sur certains facteurs de caractère chimique comme la sensibilité thermique, la susceptibilité enzymatique, la dénaturation (cf. rapport d'Anson) etc.

Dans le cas des protéines, contrairement à celui des polymères de synthèse, nous disposons donc d'un grand nombre de variables nous permettant de fixer les conditions optima du fractionnement et de la purification.

Dans le passé, on s'est contenté en général de procédés purement empiriques mais il est certain qu'à l'heure actuelle, on peut perfectionner considérablement ces procédés par une étude préalable des propriétés physico-chimiques des protéines en présence et de l'influence de différents facteurs.

Cohn et son école ont montré la voie dans le cas du plasma sanguin et grâce à leurs remarquables travaux, le fractionnement basé sur la solubilité a été considérablement affiné et est actuellement d'application générale. A cet égard l'étude des interactions ioniques est primordiale et nous lui consacrerons un peu plus d'attention.

Il est cependant essentiel dans chaque cas particulier de connaître le degré de variation que peuvent subir les paramètres

moléculaires sans entraîner de modifications irréversibles de la protéine, édifice de très grande labilité. L'étude des phénomènes d'association et de dissociation présente, à ce point de vue, beaucoup d'importance. Le développement récent de méthodes analytiques a entraîné nécessairement la recherche de procédés de purification plus efficaces et l'adaptation éventuelle en chimie des protéines de techniques basées sur des équilibrages successifs (multi-stage process).

Une revue même succincte de tous les travaux qui ont été effectués au cours de ces dernières années dans la préparation et la purification de protéines dépasse largement le cadre de ce rapport. Nous nous bornerons donc à examiner quelques aspects des méthodes basées sur la solubilité et à rappeler quelques techniques plus nouvelles susceptibles de développement.

2) Facteurs du fractionnement.

a) *Interactions ioniques.*

La formation de combinaisons stables entre protéines et ions de toute nature modifie principalement la charge sans altérer en général la masse moléculaire. Le simple fait de modifier la charge entraîne néanmoins des conséquences très importantes au point de vue solubilité et ce sont surtout les méthodes de fractionnement basées sur cette propriété qui tireront naturellement le plus de profit de l'utilisation de dérivés d'addition.

Il ne faut pourtant pas perdre de vue que des techniques comme l'électrophorèse ou la chromatographie devront tenir compte également de ces phénomènes : on sait depuis longtemps que la mobilité électrophorétique est notablement modifiée par la fixation d'ions du tampon; il y a donc là un facteur de plus sur lequel on peut agir pour obtenir une meilleure séparation.

Si l'on connaissait la structure superficielle exacte des protéines et la réactivité de leurs différents groupes, on pourrait en déduire rationnellement les composés susceptibles de réagir préférentiellement avec chaque protéine et établir un schéma de séparation comme on le fait dans la chimie des corps à bas poids moléculaire. Nous sommes encore loin d'une telle connaissance et nous devons encore trop souvent nous contenter à l'heure actuelle de données purement empiriques.

Remarquons en outre qu'en dépit du terme de spécifique dont

on qualifie nombre de ces interactions ioniques, beaucoup de substances étudiées forment des composés avec un grand nombre de protéines (colorants à structure complexe, détergents anioniques ou cationiques). Une différenciation n'est donc possible qu'en agissant sur les conditions de réaction (concentration, pH, etc.). La possibilité d'utiliser ces substances dans un fractionnement reposera alors sur la différence éventuelle de charge des protéines en présence; en particulier une variation de pH amènera la précipitation successive de protéines au fur et à mesure que leur charge nette deviendra opposée à celle de l'ion utilisé. A cet égard, l'usage des anions précipitants tels que picrate, trichloracétate, phosphotungstate, sulfosalicylate, etc., présente des possibilités encore peu explorées. Si les acides dont ils dérivent provoquent fréquemment la dénaturation, on peut par contre les utiliser sous forme de sel. Ils forment alors avec les protéines des complexes très peu solubles qui peuvent être dissociés par simple dialyse avec régénération de la protéine de départ sous son état natif.

Astrup et Birch-Andersen ([7]) notamment ont utilisé ces précipitants pour le fractionnement du plasma sanguin. L'acide lui-même convient pour la purification lorsqu'il forme un dérivé cristallin avec la protéine comme c'est le cas pour l'ovalbumine et l'acide métaphosphorique [Perlmann ([8])].

Lorsque la protéine à isoler est particulièrement résistante [rénine : Astrup ([9]); glycoprotéine du plasma : Schmid ([10]); ovomucoïde du blanc d'œuf : Lineweaver et Murray ([11]); Fredericq et Deutsch ([12])] on peut l'isoler assez aisément par dénaturation préférentielle à l'aide de ces acides.

Il faut pourtant reconnaître que la plupart des protéines ont peu d'affinité pour les anions simples et il faudra en général recourir à des molécules assez compliquées comme les détergents par exemple. Utilisés à faible concentration, ils peuvent former des dérivés d'addition avec les protéines sans les dénaturer. Mac Meekin et collaborateurs ([13]) ont suggéré l'utilisation de la combinaison cristalline : β lactoglobuline + dodécylsulfate en vue de la purification de cette protéine dont on a reconnu récemment l'hétérogénéité.

Par contre, beaucoup de protéines forment des dérivés cristallins peu solubles avec nombre d'ions métalliques simples ou engagés dans des structures complexes. La liste s'en accroît de jour en jour et permet d'espérer de vastes applications. Un bel exemple en est

fourni par le fractionnement du plasma sanguin que Cohn et son école ont poussé à un haut degré de perfectionnement en étudiant rationnellement la formation de dérivés entre protéines et cations bivalents surtout ([14]). Ainsi le zinc forme avec la plupart des protéines du plasma des complexes dont la solubilité varie énormément d'une protéine à l'autre. L'étude thermodynamique de ces combinaisons est activement poursuivie [Gurd et Goodman ([15])].

D'autres ions comme le calcium et le baryum se combinent à un plus petit nombre de protéines et agissent sur d'autres par un effet de force ionique. Enfin, dans d'autres cas, on utilisera des interactions spécifiques comme celles du thiocyanate mercurique avec le fibrinogène, du cuivre avec la céruloplasmine et la « β_1 metal-combining protein » ou encore celles du magnésium et du manganèse, etc., avec les enzymes sanguines.

La protéine est séparée de l'ion métallique par addition d'un agent complexant approprié ou par passage sur résine échangeuse d'ions.

Citons encore les études faites à Harvard sur les complexes métalliques de l'insuline [Ellenbogen ([16])], de la sérumalbumine et de la β_1-globuline [Lewin ([17])].

Enfin Morawetz et Hughes ([18]) ont récemment attiré l'attention sur les vastes possibilités offertes par l'emploi de polyélectrolytes synthétiques. On peut en effet, par le choix judicieux de copolymères, modeler plus ou moins l'ion pour lui faire jouer le rôle de précipitant d'une protéine déterminée. De plus, en faisant varier la nature et le nombre des groupes ionisables on peut espérer déplacer les protéines de leurs combinaisons complexes et effectuer ainsi de meilleures séparations. Morawetz et Hughes ont étudié la formation de différents complexes entre polyélectrolytes et sérum albumine. L'addition de petits ions au système permet encore de faire varier à l'infini les conditions de précipitation. Ils ont notamment montré qu'on pouvait séparer la sérumalbumine de l'oxyhémoglobine dans la région interisoélectrique au moyen d'acide polyméthacrylique.

b) *Phénomènes d'association et de dissociation.*

La charge et la masse moléculaire d'une protéine pourront éventuellement être influencées par une dissociation en particules plus petites ou par associations avec d'autres protéines du système. Il

faudra donc tenir compte de ces phénomènes dans toute méthode de fractionnement, puisque ces facteurs interviendront d'une façon ou d'une autre.

Dans le premier cas, la solubilité diminuera au fur et à mesure que le degré d'association augmente. C'est ainsi que le poids moléculaire de l'insuline augmente considérablement lorsque le pH varie de 2 jusque 4 ([19]). Sa solubilité suit une variation inverse et devient très faible aux environs du point isoélectrique. La fixation d'anions par l'insuline refoule la dissociation et diminue donc la solubilité [Fredericq et Neurath ([20])]; par contre, un milieu à constante diélectrique faible (eau-dioxane) exalte la dissociation et augmente la solubilité [Fredericq ([21])].

En milieu acide et à chaud, l'insuline s'associe réversiblement en fibrilles [Waugh ([22])]. On a proposé d'utiliser cette propriété spécifique pour purifier cette hormone.

Mommaerts ([23]) et Szent-Györgyi ([24]) ont réussi à purifier l'actine par ultra-centrifugation en effectuant des polymérisations et dépolymérisations successives.

Un autre exemple de l'importance de phénomènes de dissociation pour la préparation de protéines est fourni par le cas de l'hormone adrénocorticotropique : Cortis-Jones et collaborateurs ([25]) avaient trouvé que l'on pouvait purifier cette hormone par ultrafiltration, l'activité biologique se trouvant concentrée dans l'ultrafiltrat bien que le poids moléculaire de l'hormone ait été déterminé comme étant de 20.000. Or, Li et Pedersen ([26]) ont montré que cette hormone se dissociait partiellement dans l'acide acétique et que l'on pouvait séparer par dialyse à travers la cellophane deux fractions de constantes de sédimentation : 0,93 et 1,95 S.

Les myosines du muscle offrent par ailleurs l'exemple d'un système où l'état d'agrégation dépend fortement des interactions entre protéines différentes et entre protéines et ions. Il est certain que le fractionnement et la purification complète de ce système nécessiteront la connaissance exacte de ces phénomènes d'association : association de l'actine et de la myosine pour former l'actomyosine, dissociation de la tropomyosine, polymérisation de l'actine, etc. ([27]).

Les protéines du blanc d'œuf présentent des interactions considérables. Longsworth, Cannan et Mac Innes ([28]) ont indiqué par une étude électrophorétique qu'il fallait travailler dans des régions déterminées de pH et à force ionique élevée pour rompre ces

associations. Actuellement on est arrivé à isoler la plupart des composants de ce système par diverses méthodes ([29]); en se basant sur les principes développés par Cohn, on a pu établir un schéma rationnel de fractionnement à l'éthanol [Forsythe et Foster ([30])].

3) Fractionnement basé sur la solubilité.

Rappelons que la solubilité des protéines augmente en fonction de la force ionique dans le domaine des faibles concentrations salines, passe par un maximum et diminue ensuite jusqu'à devenir négligeable aux concentrations salines élevées.

La première partie de la courbe est celle du « salting in », la seconde, celle du « salting out ».

Les anciens procédés de fractionnement reposaient presque exclusivement sur le phénomène de « salting out ». Il est actuellement reconnu que les interactions : protéines-électrolytes sont de nature beaucoup plus spécifique dans la partie « salting in » de la courbe. On parvient ainsi à exalter les différences de solubilité des protéines présentes dans le mélange; l'abaissement de la constante diélectrique du milieu par addition de substances organiques comme l'acétone, les alcools, etc., entraîne alors la précipitation graduelle des différents constituants du système.

Le fractionnement classique par « salting out » réalisé à différents pH, combiné éventuellement avec l'emploi de milieux organiques, a conduit et conduit encore à de très beaux résultats surtout lorsque la protéine à isoler peut être caractérisée par une activité enzymatique. Rappelons ici l'isolement sous forme cristalline de nombreux ferments protéolytiques par Northrop, Kunitz, Herriott et Anson ([1]), celui de différentes α et β amylases par Meyer et ses collaborateurs ([21]) celui des protéines musculaires du groupe du myogène [voir la revue générale : Dubuisson ([27])] etc.

Le fractionnement poussé d'un milieu aussi complexe que le sérum ou que l'extrait du muscle et la purification ultérieure d'une protéine caractérisée par un degré d'homogénéité déjà assez élevé, ne peuvent être réalisés comme explicité plus haut, qu'en se basant sur les principes énoncés par Cohn et son école ([14], [22]) : emploi d'un système à cinq variables, concentration saline telle que les interactions protéines-sels soient élevées, emploi de sels à cations bivalents.

Tout récemment K. Schmid [10] a complété l'analyse du plasma humain réalisée à Harvard, en fractionnant la solution résiduelle qui contient environ 2 % des protéines totales, des polypeptides, des acides aminés et des carbohydrates. Un des composants principaux, une glycoprotéine, a été obtenue à l'état cristallisé sous forme de sels de Pb. Cet auteur a largement utilisé au cours de ces fractionnements, les techniques récentes de précipitation sélective et d'adsorption sur échangeur d'ions.

Il semble qu'il est préférable de combiner les fractionnements par précipitation et par extraction, les protéines restant à l'état de précipité beaucoup plus longtemps et échappant ainsi à d'éventuelles actions enzymatiques ou dénaturantes. L'équilibre entre la phase solution et la phase précipitée semble être rapide et il n'est pas nécessaire de réaliser des concentrations élevées en précipitant organique dès le début de l'opération. L'élimination initiale par extraction des produits de dégradation très solubles pouvant exercer une action solubilisante marquée sur les protéines, semble être un autre avantage de cette technique.

Un cas typique est celui de la papaïne dont la solubilité décroît considérablement à mesure que le degré de purification progresse [A. K. Balls et H. Lineweaver, J. Close [33]]. L'élimination de petits peptides éventuellement fixés sur une protéine est parfois obtenue par déplacement à l'aide d'une substance étrangère présentant une grande affinité pour la protéine; c'est ainsi que la cristallisation du complexe α-chymotrypsine-inhibiteur : diisopropylfluorophosphate est accompagnée d'une séparation de substances de nature peptidique [P. Desnuelle, M. Rovery et C. Fabre [34], E. F. Jansen, M. D. F. Nutting, R. Jang et A. K. Balls [35]]. Le même résultat peut être obtenu par lavage de la D. N. P.-protéine.

Signalons enfin qu'une extraction fractionnée peut être réalisée d'une manière continue à l'aide de dispositifs appropriés en modifiant graduellement le pouvoir solvant du milieu extracteur [36].

D'une manière générale, on peut dire qu'actuellement encore le fractionnement basé sur la solubilité reste la méthode de choix en vue de la préparation de grandes quantités de fractions protéiques. L'emploi de précipitants plus spécifiques permet déjà d'augmenter très considérablement l'efficacité de cette technique, efficacité qui sera encore accrue dans un proche avenir grâce à une connaissance plus approfondie des interactions protéines-petits ions ainsi qu'à la synthèse de nouveaux types de polyélectrolytes.

Remarque. Cristallisation.

Tout chercheur s'efforce de conduire la purification d'une protéine jusqu'à l'obtention de cristaux bien définis. De nombreuses observations montrent que la cristallisation fractionnée ne constitue pas une méthode de purification aussi efficace en chimie des protéines, qu'en chimie des petites molécules; Northrop lui-même a montré que des cristaux d'édestine pouvaient contenir jusqu'à 50 % de pepsine sans que la forme cristalline en soit profondément modifiée (1). Les cristaux de pepsine partiellement acétylée sont indiscernables de ceux de l'enzyme non traitée [Herriott (1)]. Les protéines forment aisément des solutions solides entre certaines limites de pH et il n'est donc pas étonnant que l'on ait observé à de multiples reprises que des cristallisations successives n'améliorent pas sensiblement la pureté de l'échantillon, après une première cristallisation. Il n'en subsiste pas moins que l'obtention d'une protéine à l'état cristallisé est le plus souvent l'indice d'un degré d'homogénéité assez élevé et de l'absence de toute hétérogénéité continue.

4) Techniques nouvelles.

Alors que les méthodes d'analyse des protéines se sont si prodigieusement développées au cours des dernières années, on est assez frappé de constater combien les techniques de fractionnement sont restées basées essentiellement sur les principes classiques du « salting out ». L'application de la chromatographie et de la distribution à contre-courant semble être limitée jusqu'à ce jour à quelques cas particuliers.

L'électrophorèse préparative a été beaucoup étudiée les dernières années mais, à notre connaissance, elle n'a pas encore conduit à des résultats pratiques. Par contre, l'électrophorèse-convexion est susceptible de réaliser des fractionnements beaucoup plus poussés que les méthodes classiques.

a) *Electrophorèse-convexion.*

Cette méthode a été utilisée avec succès par ses auteurs, Cann, Brown et Kirkwood dans différents cas. Ils ont notamment réalisé des fractionnements partiels de sérum sanguin correspondant à ceux qu'on peut effectuer par l'éthanol, mais avec infiniment moins

de peine et avec un meilleur rendement ([37]). De plus, ils sont arrivés à séparer la γ globuline bovine en huit fractions différentes dont les points isoélectriques moyens s'étagent de 5,74 à 7,31. Aucune de ces fractions n'est homogène mais certaines possèdent une courbe de répartition des mobilités tout à fait gaussienne ([38]). Ces expériences ont montré que l'hétérogénéité est la règle dans ce groupe très complexe des globulines. Tout ce qu'on peut espérer c'est de rétrécir le spectre des charges électriques dans une fraction donnée. Des séparations efficaces ont été réalisées par Cann, Campbell, Brown et Kirkwood ([39]) pour des anticorps de sérum de lapin et par Mathies ([40]) pour la phosphatase alcaline du rein de porc. Dans le dernier cas, la fraction obtenue, bien que n'étant pas homogène à l'électrophorèse, a une activité deux fois supérieure à celle des meilleures fractions obtenues par les procédés de précipitation. Notons encore que la méthode d'électrophorèse-convexion est d'un rendement élevé et d'une grande économie. Elle laisse les produits traités pratiquement inaltérés. Il est à penser que cette méthode, qu'on peut espérer encore voir s'améliorer du point de vue technique, est appelée à rendre de grands services.

b) *Techniques d'adsorption.*

Ces méthodes connaissent un développement lent en ce qui concerne leurs applications préparatives. Remarquons d'ailleurs que le mécanisme d'adsorption des protéines est encore très mal connu. De l'effet de la température, on est amené à supposer qu'un des stades de l'adsorption est souvent endothermique et qu'il doit donc y avoir en surface modification de la structure de la protéine.

D'autre part, l'équilibre ne semble pas être atteint pour les substances à poids moléculaires élevés; on s'est donc surtout attaché à traiter des protéines à bas poids moléculaires et à travailler dans des conditions expérimentales telles que le danger de dénaturation soit éliminé.

Tiselius ([41]) a proposé d'exalter l'adsorption d'une protéine sur un adsorbant peu actif, en diminuant sa solubilité par modification de la force ionique du milieu.

Il semble que l'emploi de nouveaux types d'adsorbants soit plus prometteur; les résines échangeuses d'ions ont été notamment utilisées par Hirs, Stein et Moore ([42]) et par Tallan et Stein ([43]) pour purifier la ribonucléase et le lysozyme sans altérer leurs propriétés biologiques.

Paleus et Neilands ([44]) ont fractionné de même le cytochrome C; Dixon, Stack-Dunne, Young et Cater ([45]) ont préparé plusieurs fractions actives de l'hormone adrénocorticotropique; une adsorption sur oxycellulose entraîne un enrichissement d'A. C. T. H. d'environ 40 fois.

Sober, Kegeles et Gutter ([46]) ont montré que des protéines telles que l'ovalbumine et la sérumalbumine bovine n'étaient pas dénaturées par passage sur ces résines; leur travail ouvre la voie dans le domaine des protéines à poids moléculaire moyennement élevé.

Polis et Shmukler ([47]) ont tout récemment isolé et cristallisé la lactoperoxidase par chromatographie de déplacement sur phosphate calcique.

Tous ces travaux n'ont cependant pas encore été réalisés à une échelle préparative macroscopique. Notons encore que Mitchell, Gordon et Haskins ([48]) combinant l'adsorption et l'extraction fractionnée, ont séparé des quantités notables d'enzymes sur une « chromatopile » de papiers filtres (fractionnement partiel de l'adénosinedésaminase, de l'amylase et de la phosphatase du complexe takadiastase) et que l'adsorption sur mousses entraînant une dénaturation préférentielle, peut conduire à des enrichissements notables [protéinase et peptidase d'actinomyces, Muggleton et Webb ([48a])]. La dénaturation sur mousses est suffisamment sélective pour permettre la séparation de la nucléoside-phosphorylase active du foie de la phosphoribomutase inactive [Abrams et Klenow ([49])].

c) *Distribution.*

On en est encore au stade des études préliminaires en ce qui concerne l'application de cette méthode à la préparation de quantités notables de protéines.

Quelques travaux récents sont cependant prometteurs dans le cas de la purification de l'insuline, par exemple, par distribution à contre-courant [Harfenist et Craig ([50])] et par chromatographie de partition [Porter ([51])]. Les premiers sont arrivés à mettre en évidence une hétérogénéité marquée dans différents échantillons d'insuline qui s'étaient révélés homogènes par d'autres méthodes. Ces résultats sont parfaitement en ligne avec les observations récentes de notre laboratoire sur l'hétérogénéité de l'insuline cristalline. En opérant un nombre suffisant de transferts, Harfenist

et Craig sont arrivés à obtenir des préparations très homogènes.

Porter a étudié toute une série de systèmes biphasés appropriés à une chromatographie de partition. Il s'agit principalement d'éthers de glycol et de solutions de phosphates. Il peut ainsi purifier rapidement l'insuline brute et la séparer du facteur hyperglycémiant.

Remarque générale : La différenciation si nette par solubilité, des sels de métaux lourds de protéines, permet de supposer qu'il serait également intéressant d'en étudier l'adsorption ainsi que le partage entre deux phases liquides.

d) *Centrifugation.*

Grâce au développement de la technique d'ultra-centrifugation, on peut prévoir que des fractionnements dans des centrifuges continues pourront être effectués même avec des solutions de fractions de poids moléculaires relativement bas.

Rappelons que la β-lipoprotéine a déjà été purifiée par cette méthode. L'emploi de milieux solvants de différentes densités doit permettre en certains cas d'excellentes séparations [lipoprotéines, par exemple : Gofman et coll. ([52])].

III. — METHODES D'ANALYSE ET CRITERES DE PURETE.

La conduite rationnelle d'un fractionnement nécessite la connaissance de la composition de l'extrait à traiter et la recherche du degré d'homogénéité de la fraction obtenue à différents stades de l'opération.

De toutes les méthodes analytiques, l'électrophorèse et l'ultracentrifugation ont été les plus utiles et les plus couramment appliquées (cf. Synge et Pedersen).

D'autre part, l'analyse complète des diagrammes d'électrophorèse, d'ultra-centrifugation, de centrifugation d'équilibre et de diffusion a été beaucoup étudiée les dernières années. Une connaissance plus parfaite des interactions moléculaires et donc du mécanisme du « boundary spreading » permettra certainement dans un proche avenir de définir avec précision un coefficient d'hétérogénéité ayant une réelle signification physique quantitative. Les techniques de mesure (optiques par exemple) ont été développées les dernières années avec très grand succès; il n'en subsiste

pas moins que de grands progrès doivent encore être réalisés afin de permettre de travailler avec des solutions de plus en plus diluées et afin d'éviter ainsi certaines complications d'interprétation.

Nous ne discuterons dans ce paragraphe que quelques points particuliers.

a) *Test d'activité biologique spécifique.*

Si la protéine est douée d'activité biologique, le test d'activité spécifique constitue un guide très précieux au cours du fractionnement et un excellent critère d'homogénéité biologique; de nombreuses recherches en témoignent [par exemple : (1) (31)].

Des facteurs perturbateurs diminuent cependant souvent la validité de ce test; un ferment peut exister partiellement sous forme inactive (complexe, précurseur) ou inhibée [ferments protéolytiques, ribonucléase (1) (6)].

D'autre part, les conditions dans lesquelles l'activité spécifique doit être mesurée, sont parfois dificiles à fixer et à reproduire au cours d'un fractionnement par suite de la présence en solution de petites molécules activatrices ou inhibitrices (papaïne par exemple).

b) *Courbes de solubilité.*

L'établissement d'une courbe de solubilité permet en principe d'avoir une image de la composition d'une fraction protéique, d'en déterminer le degré d'homogénéité et de servir donc de base au processus à suivre pour le fractionnement. Ce problème a été longuement discuté par de nombreux auteurs (Herriott, Kunitz, Steinhardt, Taylor, etc.).

Supposons provisoirement que la solubilité d'une protéine réponde aux exigences de la loi des phases.

Les diagrammes de solubilité sont le plus souvent tracés soit à température, force ionique et pH constants, mais à concentration en protéine variable (test de Kunitz-Northrop : constant solvent test) soit en modifiant le pouvoir solvant (variable solvent test) et en appliquant la relation de Cohn ($\log S = \beta - K\mu$) si la variable est la force ionique.

Deux facteurs antagonistes rendent l'établissement d'un diagramme de Kunitz-Northrop difficile dans sa partie la plus importante, c'est-à-dire à l'endroit où la courbe de solubilité présente une discontinuité : la tendance à cristalliser de tout précipité

amorphe entraînant une diminution de solubilité et d'autre part la séparation lente d'un précipité amorphe à partir d'une solution très légèrement sursaturée.

La signification de ces graphiques est plus précise lorsqu'ils sont établis par équilibrage de protéines cristallisées à condition toutefois que le taux de cristallinité soit de 100 %, ce dont il est toujours difficile de s'assurer.

Enfin, pratiquement tous les diagrammes de Kunitz-Northrop sont établis dans des conditions expérimentales telles que la concentration en protéine de la solution surnageante soit très faible ce qui peut avoir pour conséquence qu'une partie importante de la fraction protéique étudiée, échappe à l'analyse trop peu précise.

L'établissement d'une courbe de solubilité du type « salting out » échappe à certaines des critiques précédentes : d'autre part, les discontinuités d'une telle courbe sont plus marquées mais leur existence ne peut à nouveau être affirmée que sur la base de données analytiques précises et nombreuses.

Nous croyons donc que l'on ne peut admettre qu'avec circonspection des courbes à nombreuses discontinuités [Derrien, Roche : sérum, thyroglobuline ([53]) ([54])]; d'autre part, la distinction entre une courbe formée de segments de droite et une courbe continue (solution solide ou hétérogénéité continue) n'est pas aisée.

Indépendamment de ces critiques d'ordre technique, notons que la constance de solubilité en fonction de la quantité de phase solide pour une protéine dite « pure » a été mise en doute.

En principe, le précipité obtenu initialement le long du segment de droite du graphique de Kunitz-Northrop doit être pur.

De nombreuses observations montrent qu'il n'en est pas ainsi.

Dans une étude très soignée, Steinhardt ([55]) a montré que la solubilité de la pepsine diminue en fonction d'équilibrages successifs et il a attribué ce comportement à la présence de petites molécules peu solubles qui diminuent la fugacité de la pepsine dans les cristaux. Cet auteur émettait l'hypothèse que dans un tel cas, le test de solubilité serait de peu d'utilité en chimie macromoléculaire parce que trop sensible. Il semble en réalité que l'existence de molécules de pepsine légèrement différentes (solubilités différentes, activités enzymatiques qui ne sont pas strictement comparables, etc.) soit partiellement responsable de ces anomalies ([56]).

En conclusion, le tracé des courbes de solubilité a révélé à maintes reprises l'hétérogénéité d'une fraction protéique et a per-

mis même dans le cas de mélanges complexes, de préciser, par une analyse poussée, les conditions du fractionnement. C'est ainsi que Distèche [57] combinant la technique des courbes de précipitation à pH variable (force ionique constante) et à force ionique variable (pH constant) a pu étudier de manière très complète, la technique de séparation des protéines musculaires du groupe du myogène.

Cette technique peut donc être d'une très grande utilité mais, malgré sa simplicité apparente, nous croyons que les résultats ne peuvent être interprétés qu'avec beaucoup de prudence.

Le tracé de courbes de précipitation à l'aide des réactifs spécifiques mentionnés plus haut, doit permettre actuellement l'obtention de résultats plus nets et plus quantitatifs que ceux obtenus en prenant uniquement comme variable la force ionique ou le pH.

D'autre part, la technique elle-même pourrait être considérablement améliorée par application de méthodes nouvelles [turbidimétrie par exemple [58]].

c) *Analyse des groupes terminaux.*

L'analyse des groupes terminaux a également été suggérée comme critère de pureté [Desnuelle [59]]. Cette méthode basée sur celle utilisée dans le cas des peptides a déjà permis à Desnuelle, Rovery et Fabre [59] de montrer l'hétérogénéité d'un échantillon d'α chymotrypsine, constitué vraisemblablement d'une protéine et de peptides.

IV. — NOTION DE PURETE EN CHIMIE DES PROTEINES

Conjointement au développement des méthodes de fractionnement et de purification, l'analyse poussée des graphiques d'électrophorèse et d'ultra-centrifugation (Boundary spreading) a montré qu'il est dangereux de parler d'une protéine strictement homogène à la fois du point de vue de la forme moléculaire, de la charge électrique et de la structure chimique (et donc de la solubilité).

Il est nécessaire de s'entendre sur la signification à attribuer au terme « pureté » d'une protéine. Les critères de pureté devront en outre être revisés dans les cas d'association.

La fixation par une molécule de protéine, de petits peptides, acides aminés ou ions peut entraîner une hétérogénéité apparente.

L'analyse des groupes terminaux est donc de première importance afin de déceler de telles impuretés. D'autres méthodes pourraient d'ailleurs être utilisées actuellement dans le même but, après dissociation des combinaisons formées.

On pourra observer suivant les cas, soit plusieurs frontières de sédimentation [hémocyanine : Brohult ([60])] soit une frontière très étalée [insuline : Oncley et Ellenbogen ([19])].

Par contre, l'électrophorèse est souvent impuissante à distinguer les composants à degré d'association variable [conarachine : Johnson et Naismith ([61]) ([62])].

Les hétérogénéités ainsi mises en évidence dans une protéine soit disant pure, ne sont souvent dues qu'à des différences dans la nature de quelques groupes chargés ou la rupture de quelques acides aminés. Ces légères différenciations peuvent être dues à différentes causes, dont une des plus importantes est une action enzymatique superficielle au cours de la préparation.

C'est ainsi que l'on a reconnu la présence de deux et parfois trois composants dans les clichés électrophorétiques de l'ovalbumine cristalline, sans arriver à aucune séparation, par d'autres méthodes physiques ou chimiques ([63]). Perlmann ([64]) a éclairci ce problème en montrant qu'il ne s'agit que d'une différence dans la quantité de phosphore lié, le nombre de groupes phosphates variant de 0 à 2 par molécule, par suite d'une déphosphorylation enzymatique progressive. Une action semblable est exercée par des phosphatases sur l' α-caséine, la β-caséine restant inaltérée [Perlmann ([65])].

L'ovalbumine subit une modification plus profonde sous l'influence d'enzymes bactériennes et donne naissance à la plakalbumine avec libération de quelques peptides [Linderström-Lang et Ottesen ([66])]. Cette protéine ressemble à l'ovalbumine par de nombreuses propriétés, les principales différences étant la forme cristalline, la mobilité électrophorétique [Perlmann ([67])] et la solubilité très supérieure [Fredericq et Linderström-Lang ([68])].

Jacobsen ([69]) a trouvé que la solubilité et le nombre de groupes titrables d'échantillons de β-lactoglobuline préparés par diverses méthodes présentent des différences considérables qu'il attribue à des dégradations enzymatiques restreintes.

De nombreux auteurs ont également montré que l'évolution d'un organe entraînait de légères variations des propriétés chimiques et physiques des protéines.

La possibilité d'existence de structures très voisines caractérisées par la même activité biologique est démontrée par l'étude d'une protéine isolée à partir d'espèces différentes [α et β amylases : Meyer ([31]); insuline : Sanger ([70]), Harfenist et Craig ([71])]. La distinction de ces molécules très semblables ne peut être obtenue que par emploi d'une méthode très sélective et spécifique d'une propriété donnée. C'est ainsi que les insulines de bœuf, de mouton et de porc se comportent de la même manière dans l'appareil à contre-courant de Craig, mais se distinguent entre elles par de légères variations de composition chimique.

Enfin, comme signalé plus haut, le facteur stabilité des protéines mérite d'être étudié de plus près, le vieillissement entraînant dans de très nombreux cas une augmentation du degré d'hétérogénéité [voir le travail de Tallan et Stein sur le lysozyme ([72])]. Il est intéressant de signaler ici que la fixation d'ions lourds (Ca, Mn, Cd) par la trypsine, augmente la stabilité de ce ferment, l'équilibre protéine native $\rightleftharpoons$ protéine dénaturée étant déplacé vers la gauche et la vitesse d'auto-hydrolyse étant donc diminuée ([73]).

De même le plasma traité au sulfate de baryum acquiert une remarquable stabilité à la chaleur. Le caprylate de sodium a le même effet. De telles observations sont évidemment de toute première importance.

Conclusion.

Les méthodes actuelles de fractionnement permettent la résolution de milieux complexes, en groupes de protéines nettement distincts. Il serait urgent de développer des techniques plus fines permettant plus aisément la séparation des constituants d'un groupe donné comme cela a déjà été réalisé dans quelques cas particuliers; il semble bien que souvent on se heurtera en dernière analyse à une hétérogénéité continue de charge, de grandeur moléculaire, de composition chimique ou de structure superficielle; s'il en était ainsi, les propriétés que nous mesurons ne seraient que des moyennes comme pour les polymères de synthèse.

Institut de Chimie physique-Université de Liège.

BIBLIOGRAPHIE

(1) J. H. Northrop, M. Kunitz et R. M. Herriott : *Crystalline enzymes*, New-York, (1948).
(2) M. Kunitz : *J. Gen. Physiol.*, **35**, 423 (1952).
(3) J. S. Falconer, D. J. Jender et D. B. Taylor : *Discuss. Faraday Soc.*; N° **13**, 40 (1953).
(4) W. C. Schneider et J. H. Hogeboom : *J. Nat. Canc. Inst.* **10**, 969 (1950).
(5) H. G. Hers, J. H. L. Berthet et C. de Duve : *Bull. Soc. Chim. Biol.*, **33**, 21 (1951).
(6) M. Pirotte et V. Desreux : *Bull. Soc. Chim. Belg.*, **61**, 167 (1952).
(7) T. Astrup et A. Birch-Andersen : *Nature*, **160**, 637 (1947).
(8) G. E. Perlmann : *Biochem. J.*, **32**, 931 (1938).
(9) T. Astrup : *Nature*, **160**, 570 (1947).
(10) K. Schmid : *J. Am. Chem. Soc.*, **75**, 61 (1953).
(11) H. Lineweaver et C. W. Murray : *J. Biol. Chem.*, **171**, 565 (1947).
(12) E. Fredericq et H. F. Deutsch : *J. Biol. Chem.*, **181**, 499 (1949).
(13) T. M. Mac Meekin, B. D. Polis, E. S. Della Monica et J. H. Custer : *J. Am. Chem. Soc.*, **71**, 3606 (1949).
(14) E. J. Cohn, D. M. Surgenor et H. Isliker : *Discuss. Faraday Soc.*, N° **13**, 176 (1953).
(15) F. R. N. Gurd et D. S. Goodman : *J. Am. Chem. Soc.*, **74**, 670 (1952).
(16) E. Ellenbogen : *Doctoral Thesis*, Harvard (1949).
(17) J. Lewin : *J. Am. Chem. Soc.*, **73**, 3906 (1951).
(18) H. Morawetz et W. L. Hughes, Jr. : *J. Phys. Chem.*, **56**, 64 (1952).
(19) H. Gutfreund : *Biochem. J.*, **42**, 156 et 544 (1948); J. L. Oncley et E. Ellenbogen : *J. Phys. Chem.*, **56**, 85 (1952).
(20) E. Fredericq et H. Neurath : *J. Am. Chem. Soc.*, **72**, 2684 (1950).
(21) E. Fredericq : *Nature*, **171**, 570 (1953).
(22) D. F. Waugh : *J. Am. Chem. Soc.*, **70**, 1850 (1948).
(23) W. F. H. M. Mommaerts : *J. Biol. Chem.*, **188**, 559 (1951).
(24) A. Szent-Györgyi : *J. Biol. Chem.*, **192**, 361 (1951).
(25) B. Cortis-Jones, A. C. Crooke, A. A. Henley, P. Morris et C. J. O. R. Morris : *Biochem. J.*, **46**, 173 (1950).
(26) C. H. Li et K. O. Pedersen : *Arch. Biochem. Biophys.* **36**, 462, (1952).
(27) M. Dubuisson : *Annual Review of Biochemistry*, **21**, 387 (1952).
(28) L. G. Longsworth, R. K. Cannan et D. A. Mac Innes : *J. Am. Chem. Soc.*, **62**, 2580 (1940).
(29) H. L. Fevold : *Advances in Protein Chemistry*, **6**, 188 (1951).
(30) R. H. Forsythe et J. F. Foster : *J. Biol. Chem.*, **184**, 377, 385 (1950).
(31) K. H. Meyer : *Angew. Chem.*, **63**, 153 (1951).
(32) E. J. Cohn et collab. : *J. Am. Chem. Soc.*, **72**, 465 (1950).
(33) A. K. Balls et H. Lineweaver : *J. Biol. Chem.*, **130**, 669 (1939); J. Close : Thèse, Liège (1952).
(34) P. Desnuelle, M. Rovery et C. Fabre : *Biochim. Biophys. Acta*, **9**, 109 (1952).
(35) E. F. Jansen, M. D. F. Nutting, R. Jang et A. K. Balls : *J. Biol. Chem.*, **185**, 209 (1950).
(36) V. Desreux : *Rec. Trav. Chim. Pays-Bas*, **68**, 789 (1949).
(37) J. R. Cann, R. A. Brown et J. G. Kirkwood : *J. Am. Chem. Soc.*, **71**, 1609 (1949).
(38) J. R. Cann et J. G. Kirkwood : *Cold Spring Harbor Symposia on Quantitative Biology*, **14**, 9 (1950).
(39) J. R. Cann, D. H. Campbell, R. A. Brown et J. G. Kirkwood : *J. Am. Chem. Soc.*, **73**, 4611 (1951).
(40) J. C. Mathies : *Science*, **115**, 144 (1952).
(41) A. Tiselius : *Ark. Kem. Miner. Geolog.* **26** *B*, N° **1**, (1948).
(42) C. H. W. Hirs, W. H. Stein et S. Moore : *J. Am. Chem. Soc.*, **73**, 1893 (1951).

(43) H. H. Tallan et W. H. Stein : *J. Biol. Chem.*, **200**, 507 (1953).
(44) S. Paleus et J. B. Neilands : *Act. Chem. Scand.*, **4**, 1024 (1950).
(45) H. B. F. Dixon, M. P. Stack-Dunne, F. G. Young et D. D. Cater : *Nature*, **168**, 1084 (1951).
(46) H. A. Sober, G. Kegeles et F. J. Gutter : *J. Am. Chem. Soc.*, **74**, 2734 (1952).
(47) B. D. Polis et H. W. Shmukler : *J. Biol. Chem.*, **201**, 475 (1953).
(48) H. K. Mitchell, M. Gordon et F. A. Haskins : *J. Biol. Chem.*, **180**, 1071 (1949).
(48a) P. W. Muggleton et M. Webb : *Biochim. Biophys. Acta*, **8**, 526 (1952).
(49) A. Abrams et H. Klenow . *Arch. Biochem. Biophys.*, **34**, 285 (1952).
(50) E. J. Harfenist et L. C. Craig : *J. Am. Chem. Soc.*, **74**, 3083 (1952).
(51) R. R. Porter : *Biochem. J.*, **53**, 320 (1953).
(52) J. W. Gofman et coll. : *Science*, **111**, 166 (1950).
(53) Y. Derrien : *Svensk Kem. Tid.*, **59**, 139 (1947).
(54) Y. Derrien, R. Michel et G. Roche : *Biochim. Biophys. Acta*, **2**, 454 (1948).
(55) J. Steinhardt : *Cold Spring Harbor Symposia on Quantitative Biology*, **6**, 301 (1938).
(56) R. M. Herriott, V. Desreux et J. H. Northrop : *J. Gen. Physiol.*, **24**, 213 (1940).
(57) A. Distèche : *Biochim. Biophys. Acta*, **10**, 524 (1953).
(58) A. Oth : *Bull. Soc. Chim. Belg.*, **58**, 285 (1949); J. Bischoff et V. Desreux : *Bull. Soc. Chim. Belg.*, **60**, 137 (1951).
(59) P. Desnuelle : Quelques techniques nouvelles pour l'étude des protéines, *Advances in Enzymology*, **14**, 261 (1953).
P. Desnuelle, M. Rovery et C. Fabre : *Compt. Rend.*, **233**, 1496, (1951).
(60) S. Brohult : *J. Phys. Coll. Chem.*, **51**, 206 (1947).
(61) P. Johnson et W. E. F. Naismith : *Discuss. Faraday Soc.*, **13**, 98 (1953).
(62) P. Johnson, E. M. Shooter et E. K. Rideal : *Biochim. Biophys. Acta*, **5**, 376 (1950).
(63) L. G. Longsworth : *J. Am. Chem. Soc.*, **61**, 529 (1939).
(64) G. E. Perlmann : *Discuss. Faraday Soc.*, **13**, 67 (1953).
(65) G. E. Perlmann : *J. Am. Chem. Soc.*, **74**, 3191 (1952).
(66) K. Linderström-Lang et M. Ottesen : *Nature*, **159**, 807 (1947).
K. Linderström-Lang et M. Ottesen : *Compt. Rend. Trav. Lab. Carlsberg*, **26**, 403 (1949).
(67) G. E. Perlmann : *J. Am. Chem. Soc.*, **71**, 1146 (1949).
(68) E. Fredericq et K. Linderström-Lang : *Compt. Rend. Trav. Lab., Carlsberg*, **26**, 443 (1949).
(69) C. F. Jacobsen : *Compt. Rend. Trav. Lab. Carlsberg*, 26, 472 (1949).
(70) F. Sanger : *Nature*, **164**, 529 (1949).
(71) E. J. Harfenist et L. C. Craig : *J. Am. Chem. Soc.*, **74**, 4216 (1952).
(72) H. H. Tallan et W. H. Stein : *J. Am. Chem. Soc.*, **73**, 2976 (1951).
(73) M. Bier et F. F. Nord : *Arch. Biochem. Biophys.*, **31**, 335 (1951); **33**, 320 (1951); *Nature*, **171**, 1022 (1953).
L. Gorini : *Biochim. Biophys. Acta*, **7**, 318 (1951).

Discussion des rapports de MM. Synge et Desreux

M. Tiselius. — 1) In discussing the purity of proteins, particularly when we deal with preparations isolated by gentle methods, it is well to remember that a substance which from a chemical point of view should be regarded as an impurity, from a more biochemical or functional point of view may be something very essential. Reactions in a living organism are coupled reactions and therefore must, to a large extent, depend upon molecular interaction, which shows up as complex formation.

2) The zone electrophoresis methods have proved very valuable, also in the separation and characterization of peptides and polypeptides. Filter paper strips are, of course, very commonly used, and are very suitable for qualitative studies. In many cases, however, they show marked adsorption effects, which are very disturbing. Some of these effects may be eliminated by using filter paper pretreated with suitable reagents, to modify its electrical charge. For more accurate work we prefer zone electrophoresis in columns packed with starch, which as Kunkel has shown, shows much less adsorption than filter paper. The zones are « eluted » from the column (by washing with the buffer solution) into an automatic fraction collector. Thus one column can be used for many successive experiments.

3) As regards the separation of large peptides, I agree entirely with Dr. Synge that this is now a particularly important problem in the protein structure field. There seems to be quite a number of different methods available : different forms of electrophoresis, chromatography and countercurrent extraction, but it appears somewhat difficult to predict in a specific case which procedure (or combination of procedures) is to be recommended. We know too little about the rational background of some of these procedures, and their application is mostly based upon empirical results, involving many trials. It would seem desirable, however, to investigate more systematically the main principles on which such separations are based.

4) With regard to the electrokinetic ultrafiltration mentioned by Dr. Synge, I may add that it also appears possible to introduce frictional effects, depending upon molecular size, in electrophoresis, in media which are not strictly gels. Many substances, for example dextran, will influence the mobility of such ions very little but will have a retarding effect on larger molecules, depending upon their size. I believe that attempts at *zone* separations based upon molecular size are particularly desirable (cf. also my remark in the discussion of Dr. Pedersen's paper).

M. Adair. — I should like to refer to an advance in our knowledge of the chromatography of proteins, made by Boardman & Partridge. This work was published two months ago, too late for inclusion in Dr. Synge's report.

Previous investigations had indicaded that the adsorption of proteins with fairly high molecular weight is an irreversible process, not adapted for chromatographic procedures. By using suitable ion exchange resins, and modifying the experimental conditions, Boardman * succeeded in obtaining reversible adsorption and fractionation of proteins like haemoglobin and foetal sheep's haemoglobin.

Prof. Desreux and Dr. Frédéricq have referred to several criteria for the purity of proteins. One additional method is to compare the number average molecular weight (symbolized M_n) for different samples of the protein. If we have a mixture of proteins with different molecular weights, the average M_n may be altered by fractional crystallization and other procedures.

Figures calculated from unpublished measurements of the osmotic pressures of solutions of horse carbon monoxide haemoglobin are recorded in table I. It will be seen that the variations in M_n are small.

It appears in this test haemoglobin resembles a pure substance

Table I.

M_n	Sample number	Description
67,100	1	Total protein (1)
67,300	2	Fraction crystallized at pH 7.4
67,700	3	Fraction in mother liquor at pH 7.4
67,700	4	Fraction crystallized at pH 6.9

*) N. K. Boardman & S. M. Partridge. *Nature*. Vol. 171, p. 208, 1953.

(1) Total protein represents hæmoglobin separated from stroma proteins by ether and salt, but not crystallised, ref. : *Biochem. J.* **28,** 1230 (1934).

M. Martin. — The success which has attended the classical salting out procedures for the separation of proteins indicates that the difference between different proteins is really very large indeed. Such methods effect only poor separations between molecules as similar as those of the amino acids. If therefore the methods capable of separating small similar molecules can be applied to proteins, previously difficult separations should become easy. This has indeed proved to be the case with partition chromatography, where practically pure proteins may be isolated from a gross mixture in a single step, once the right two phase system has been found, the greatest part of the contaminating proteins having an Rf value of nearly zero or unity. Thus Porter was able to isolate insulin, substantially pure, in a single step from a crude pancreatic extract, and in another case, unpublished, achieved a one or two thousand fold concentration of a bacterial pencillinase in a crude extract. It is however difficult to find a suitable system, in which the desired protein shows a moderate Rf value, and is not denatured.

I would like to add my support to what Synge and Tiselius have already said of the potentialities of the methods depending upon mobility in an electric field. I feel that a great deal more effort should be put into their development, particularly the continuous flow type associated with Svensson and Brattsten; Grassmann and Durrum. By successive runs at different pH values, and in the presence of various types of ionic, and even non ionic substances, I believe that almost any protein could be purified, by this means alone. The combination of proteins with different ions, and the influence on the charge of the protein of non ionic substances can be expected to be highly specific, and while the method may not have the resolving power of partition chromatography in all cases, in some it will have a greater, and will in any event afford much preliminary purification. It should be a great deal easier to find a suitable set of conditions, and the quantity that can be handled should have wide limits. I am aware that present apparatus does not fulfill the requirements of a routine piece of apparatus, but I feel that a little more development should overcome this.

I do not think that the convection electrophoresis method of Kirkwood and his colleagues can possess the resolving power that is possible for the zone type of electrophoresis, whether the latter

is continuous or discontinuous. The principle of combining convection in one direction with a concentration change in a direction at right angles, as originally introduced by Clusius and Dickel for isotopic separations, is particularly adapted for cases where the possible concentration change is very small, and the final result can only be good where a large number of stages of effective equilibration can be obtained. This latter condition is difficult to meet in the case of proteins, where the low diffusion constant would entail working in an extremely thin layer, which at least in present apparatus is impossible. In any case, it is easy to reach extremely high values of concentration ratio in the zone type of apparatus, which cannot be made use of in the convection type of apparatus.

The convection type of apparatus is the proper one to use for the utilization of any tendency of a protein to migrate in a salt, solvent or pH gradient. It has not however yet been satisfactorily shown that the tendency to migrate is large enough to afford a basis of useful separation. Kirkwood and his colleagues have purported to have shown this, but in the experiment they describe the observed result could be explained by a small transport of water occurring in the membrane as the result of the salt gradient there. If this were in fact the explanation of the result of their experiment, as I believe, the only fractionation which would occur would be due to differences in diffusion constants of the various proteins concerned, and the result would more easily be obtained by the application of pressure, or a strong solution of some high molecular weight substance, which would cause a flow of water through the membrane.

Since we cannot have too many independent methods of separation, I think more attention should be paid to those depending on diffusion constant. Though the diffusion constant may be rather insensitive to molecular differences, diffusion against a flow can amplify the concentration ratio. I do not feel that the methods so far used have been developed enough to justify discussion here, but they should not be dismissed from consideration. Signer has had more experience of these methods than any of us. Would he care to give his views ?

M. Signer. — I. *Fractionation by dialysis.*

The multi-stage arrangement of diaphragm diffusion cells as

described in 1946 has a rather restricted field of application for three reasons :

a) The single stage is not very efficient in separating molecules of different molecular weights. Two substances with a molecular weight ratio 8 : 1, both having spherical particles, give a ratio of the diffusion constants as small as 1 : 2.

b) The arrangement of the diffusion cells is rather complicated. It needs special means to regulate the rate of flow on both sides of each diaphragm.

c) The apparatus is useful to separate amounts of material from 10 to 100 grams and is not yet adapted to micro quantities.

II. *Partition of proteins between two liquid phases.*

a) *Coacervates with phenol-water - casein.*

In the pH range 5 to 8, the casein accumulates at the phase boundary between the two liquids and a very stable emulsion is formed. By centrifugation, three layers are obtained. On top is an aqueous solution of low protein concentration. The second layer is a coacervate with drops of water dispersed in phenol, the casein being at the surface of the drops. The third layer is made of phenol with a low protein concentration.

b) *Use of three components in the critical region.*

If ethanol is added to the water-phenol-casein system in the right amount to form a critical two-phase-system, there is no emulsification and the casein is distributed between the two phases. The distribution depends on the pH value. At a pH about 8, in two successive operations electrophoretically pure α casein and a mixture of β and γ casein free from α casein can be obtained.

Details will be found in I. O. Walter's thesis, the University of Berne, 1952 : « Untersuchungen über die Fraktionierung von Casein mit zwei flüssigen, nicht mischbaren Phasen. » (Micro films available through : Stadt- und Hochschulbibliothek, Bern.)

M. Anson. — Desreux pointed out that fractionation by anion precipitants such as tungstate and trichloracetate have been little explored. Perhaps the main reason for this neglect of a very useful group of reagents is that the amount of acid needed to make the anion precipitants work is usually sufficient to cause denaturation. Thus the classical Folin tungstic acid precipitation

used for the removal of blood proteins usually results in denaturation of the proteins. There are two ways of avoiding denaturation by anion precipitants. The first is to find by careful exploration a minimum pH of precipitation at which the protein is stable. I found a very narrow range of pH within which cathepsin was stable and still precipitated by tungstic acid. A second way of obtaining anion precipitation without denaturation is by adding an amount of ammonium sulfate which is not sufficient to cause precipitation by itself but is sufficient to make the pH of anion precipitation enough high to be out of the denaturation region of pH. This salt promoted anion precipitation is analogous to salt promoted cation precipitation and, in some ways, to salt promoted adsorption.

Anion precipitation in addition to involving the risk of denaturation by acid, also is usually relatively unspecific. In some unpublished experiments on the purification of cathepsin, however, I found that anion combination together with cation combination should give a rapid and very effective sort of fractionation. When I added barium hydroxide to a tungstic acid precipitate of cathepsin and contaminating protein, in the expectation of getting insoluble barium tungstate and soluble cathepsin, I actually obtained insoluble and greatly purified cathepsin. Apparently, at a pH too acid to cause barium precipitation and too alkaline to cause tungstate precipitation there was sufficient combination of the protein with both ions to cause insolubility. This combined anion-cation fractionation offers the possibilities of new specification of fractionation by differential precipitation and also has the advantage of operation in the pH range most favorable from the point of view of protein stability.

Salt promoted anion precipitation may also be more specific than straight anion precipitation.

Even when anion precipitation is not specific, it may offer, and especially under unfavorable conditions, a very useful way of collecting protein. I was able to precipitate cathepsin with tungstic acid out of a very dilute cathepsin solution containing many protein split products which acted as peptizing agents. No other method of collection I tried was successful.

Desreux also pointed out the advantage of beginning the isolation of a protein with a selective extraction, for instance by a solvent, which denatures the bulk of the contaminating proteins but leaves

the desired protein unchanged. As a classical example of this he gave the Kunitz extraction of trypsinogen and chymotrypsinogen with 0.25 N sulfuric acid. Despite the very great individual successes which have been achieved by the use of differential denaturation, I do not know of a single case in which the isolation of a protein has been begun by a systematic study of the stability of the protein to a great variety of denaturation procedure. Such a systematic study, I feel confident, would often discover some new denaturation procedure or some unusual denaturation procedures, towards which the desired protein is peculiarly stable, and which could be the basis of purification by differential denaturation. If any stability tests at all are carried out, it is usually only tests of pH or temperature stability, not even tests of stability towards the many other denaturing agents that are known under various conditions and in the presence and absence of known promoters and inhibitors of denaturation. And no one has attempted to isolate a protein by the use of systematic differential denaturation alone. But no matter what method of systematic fractionation is used, there is a great advantage in removing the bulk of contaminating proteins at the start.

Synge stated that the separation of proteins is difficult because the rate of diffusion does not change rapidly with change in molecular mass. Another difficulty is the slowness of diffusion. Both of these difficulties can be decreased, in some measure, by the use of ion accelerated diffusion.

In the usual diffusion of protein through a porous membrane, neutral salt is added to the water on both sides of the membrane to suppress diffusion potentials, just as salt is added to suppress Donnan effects in osmotic pressure measurements. In the absence of added salt, diffusion potentials are set up if acid or base is added to isoelectric protein. The protein then moves not only by ordinary diffusion but also, and more rapidly, by electrophoresis, set up by the diffusion potential. The rate of this electrophoresis, other things being equal, depends on the titration curve of the protein. Thus proteins with similar frictional resistances but different titration curves can be separated by ion accelerated diffusion. In testing for the purity of an enzyme by diffusion experiments, it is better to show that the nitrogen and the activity move at the same rates in the presence of varying amounts of acid and base and in the absence of salt than merely to show

that the nitrogen and the activity move at the same rate when diffusion potentials are suppressed by salt.

Although ion accelerated diffusion is an interesting phenomenon, it is not clear that it has important advantages over electrophoresis set up by electrode potentials.

M. Bigwood. — With regard to procedures of extraction, separation or concentration of proteins from the natural media in which they are found, I might refer to a rather simple and convenient method, applicable in certain cases, namely when the amount of protein in solution is only very minute. We have used it for concentrating and extracting the chorionic gonadotrophic hormone from urine of pregnant women (Cl. Wodon, J. P. Dustin & E. J. Bigwood, *Archives Internationales de Physiologie*, 1951, **58**, 463; 1952, **60**, 207; J. P. Dustin, Cl. Wodon, O. Médard & E. J. Bigwood, *Annales d'Endocrinologie*, 1952, **13**, 687). The procedure consists in using porous porcelaine (Chamberland filters for bacteriological use) which acts as a selective filter towards proteins at their isoelectric point. In the case of chorionic gonadotrophin (isoelectric point close to pH 3), the pH of paper-filtered urine of pH 5 to 6 is adjusted at pH 3 and a 100 cc sample is run slowly through the porous wall after having washed the latter first with a HCl solution of pH 3. The protein is retained in the wall. A specific biological test for the hormone shows that the filtered urine is free from gonadotrophin. The latter is eluted and concentrated in 10 cc of a moderately alkaline solution. At any pH other than the one corresponding to the protein's isoelectric point, the hormone passes through the filter. With a solution of Armour's cristalline serum bovalbumin Chamberland filters will also hold the protein back in its wall unless the pH of the solution is not adjusted at pH 4,8, the isoelectric point of the protein. Similarly they will also retain haemoglobin at pH 6,8 or in acid media in which hematin precipitates, but in alkalin solutions the protein passes through the filter. It remains to be determined whether the process is still a sufficiently selective one when one deals with a mixture of proteins of different isoelectric points.

Walls of different grades of porosity have been tested. It has been noticed that the most porous ones may be the most efficient ones for the purpose.

M. Tiselius. — I agree with Dr. Martin that the continuous flow electrophoresis, based upon a cross-wise combination of flow and electrophoretic migration, appears to be the most promising approach to a large scale electrophoretic fractionation. I would like to point out, however, that work by Dr. Brattsten in my laboratory has shown that a rather complicated arrangement is necessary if one is to gain the full advantage of the method. It is not so easy, for example, to obtain an evenly distributed flow through the rectangular trough used. Thus, for preparing moderately large quantities, we rather use zone electrophoresis in starch columns.

I also believe that attention should be given to the problem of chromatography on a larger scale, particularly for preparing larger quantities of peptides from partial hydrolysates. Displacement chromatography is particularly well suited for this type of work. Partridge has been able to separate amino acids on a preparative scale in this way, and I believe such methods are capable of a very wide application.

M. Desreux. — La formation de complexes protéines-ions métalliques (Zn-Hg, etc.), entraînant une variation de solubilité très différente suivant la nature de la protéine, il serait probablement intéressant d'appliquer le principe de la chromatographie par « salting out » proposé par Tiselius à des solutions de sels de zinc de protéine, par exemple.

The Denaturation of Proteins

by M. L. Anson

INTRODUCTION

It has long been known that when a protein solution is heated to some critical temperature, the protein is suddenly changer into an insoluble form. This sudden change, which is called denaturation, is the most characteristic reaction of proteins. As we shall see, denaturation is accompanied by many changes in the properties of the protein, as well as by the change in solubility, and denaturation can be brought about in many very different ways, as well as by heating.

It has also long been known that the rate of heat denaturation, under certain conditions, is very much more sensitive to the temperature than the rate of ordinary chemical reactions. The rate of denaturation can be increased as much as 600 times for a rise of 10 degrees. The great effect of temperature on the rate of the heat destruction of enzymes was what first led to the notion that enzymes are proteins.

Any theory of denaturation, indeed any general theory of protein structure, must account for the fact that proteins can exist in two forms, native and denatured, with very different solubilities and other properties, and must account for the facts that denaturation can be brought about in many different ways and that the rate of denaturation can be extraordinarily sensitive to the temperature.

When last I reviewed the subject of denaturation in 1945 for Advances in Protein Chemistry, I dealt with the meager pre-war literature, and in particular with the then recent work on the relation of protein denaturation to the reactivity of protein groups. In the

post-war years, the literature of protein denaturation has expanded considerably, along with the very great expansion of work on proteins in general. In preparing this paper, I collected a surprisingly high pile of cards with references to the work of the last ten years. I am afraid, nevertheless, that I have no startling new discoveries to report. Some old discoveries have been worked out in more detail. The general increase in our knowledge of protein structure has made more clear the place of the study of denaturation in the study of protein structure. As a result, what is essentially the old picture can now perhaps be better presented. But not enough is yet known about the means of identifying and measuring the kinds of protein bonds which are broken in denaturation to make possible the next decisive advances.

Reviews of protein denaturation usually begin with a definition of denaturation. I believe that most such definitions are bad and useless, and that the definition in the notable review of Neurath et al. (1944), that denaturation is any non-proteolytic modification of the protein, is very bad indeed, and even goes beyond being useless to being harmful. I believe further that no rigid definition is possible at present; that the whole review must serve as a vague, complex, inadequate definition; that the only useful purpose of a definition, at present, is to separate denaturation, whatever denaturation may be, from protein changes which are definitely not denaturation, such as dissociation of the protein molecule into smaller units.

Rather than attempt the impossible task of providing a proper definition of denaturation, I should like to introduce this review, which is intended to be primarily a starting point for discussion rather than the usual compilation of literature, by giving my own vague general notions about denaturation itself, and about the place of the study of denaturation in that study of protein structure which is the main subject of this Solvay Council.

Proteins can be considered, first of all, as chains of amino acids bound together by firm peptide linkages. Others speaking before this Council will describe the triumph of this generation of protein chemists in being able to put down plain the nature and the order of amino acids in a peptide chain as much as 30 amino acids long.

The so-called straight peptide chain is not straight in the ordinary sense of the word. The amino acids, even when they are joined

end to end, and not branched, are twisted by forces between amino acid groups into a definite wriggled chain structure, which will be the subject of discussion in other sessions. In some proteins wriggled chains are joined by firm S-S linkages to form rigid complexes of chains.

Even the complex of chains is still not the typical denaturable globular protein like egg albumin. In the typical native globular protein, the wriggly chains are twisted further into some fixed, characteristic globular form by a considerable number of looser linkages between amino acids; looser linkages such as hydrogen bonds, salt bonds and van der Waals bonds; bonds which are much easier to break and to reform than the firmer peptide and S-S bonds; bonds which in proteins are not easily identified and measured by the methods available today. It is now pretty much agreed that denaturation involves the breaking of a considerable number of these looser bonds within the molecule, and that therefore the change from the globular native to the denatured form of the protein makes the protein more like the simple peptide chain. Not entirely like the simple peptide chain, because not all the looser bonds are broken.

It has perhaps not been pointed out explicitly enough that since denaturation involves the breaking of some of the looser protein bonds, the study of denaturation must be related to all those other aspects of protein chemistry which deal with similar protein bonds. It may be worth while to bring these different aspects of protein chemistry together as a group. The sorts of looser linkages involved in denaturation are also involved to varying degrees in the fixing of the geometrical structure of peptide chains; in protein-protein interactions; in interactions between proteins and many ionic and non-ionic substances present in solution with the proteins; in enzyme-substrate combinations; in antigen-antibody combinations; in protein adsorption, aggregation, precipitation and crystallization; perhaps in the reproduction of proteins. The occasion of a Solvay Council offers a special opportunity to treat the different protein reactions involving the looser linkages as relatives in a family of reactions, and to discuss the ways in which the understanding of any one of these reactions can help the understanding of the others.

Nobody would deny that a native protein can exist in a continuous series of myriad forms and still retain the solubility and other essential characteristics of native protein. The native protein

can, for instance, exist in many ionic forms. It can have many of its groups combined reversibly with many different substances. The native protein can even become somewhat uncoiled by the breaking of some of its looser linkages, and still remain native. There is no such thing as *the* native form of a protein. There is native protein always involved in many different equilibria which are influenced by the chemical environment. It ought to be, but apparently has not been, obvious that denatured protein can similarly exist in myriad forms — in myriad ionic forms, in myriad degrees of being uncoiled. Without very special evidence, one ought not to talk of different degrees or different kinds of denaturation, any more than of different degrees of nativity. One ought to talk merely of the different forms of either native or denatured protein. So far, no one has demonstrated that the properties of denatured protein in equilibrium with completely renatured protein depend on the history of the denatured protein as well as on the chemical environment.

Beyond denaturation, of course, the denatured protein can be changed more drastically. The very reagents, such as acid or alkali, which bring about denaturation, may sometimes bring about various kinds of decomposition or other irreversible alteration of the protein molecule.

My first general point, then, is that protein can exist in two quite different states, the native and the denatured, each one of which, the protein being so complex, has a very large number of forms. And that the denatured protein can be further changed by reactions which have nothing to do with denaturation. I look upon the protein molecule, considering only its two relatively stable states, as a many faceted body on a vibrating table. While the body rests on one stable face, it is bounced about into an infinity of different positions, until it is given enough of a push to be flopped over onto another relatively stable face, from which it is bounced about into another group of positions.

My second general point, presented more for discussion and as a notion for holding my review together than as something proven, is that the change from the native to the denatured state involves usually the breaking of a considerable number of the looser protein linkages and that these linkages are usually broken in an « all or none » way. By the expression « all or none » I do not mean to

imply that the linkages are not broken one at a time, but rather that the intermediate forms are relatively unstable.

To return to our mechanical model for an analogy of my second point, the many faceted body is not directly pushed from one stable face to another. Rather it rests momentarily on many unstable, say very small, faces, finally either returning to its original stable face or landing on a new stable face, representing the denatured state.

To anticipate my conclusions, the means of bringing about denaturation are the means of breaking the looser linkages; the large changes in properties are due to the great changes in configuration of the molecule as a result of the breaking of many linkages; the reversibility of denaturation, when it is reversible, is due to the fact that only looser, superficial linkages are broken, linkages which are easily reformed, while the spine of the protein remains intact; the great sensitivity of denaturation to temperature, when it is sensitive, is due to the fact that many linkages are broken in an « all or none » way.

The picture of denaturation I have given is complex; the native and denatured states both existing in many forms whose extremes may be close, and the transformation from one state to the other involving the breaking of many linkages through many steps but in an « all or none » way.

This picture, useful as it has been in suggesting experiments, may be wrong. I am sure, nevertheless, that any picture which eventually proves to be right will also be complex, that any simple view of the details of denaturation is out of the question.

I remember being present at the famous address in which General Marshall proposed the Marshall Plan. In addition to his essentially simple program for what has come to be known as Marshall Aid, the general made the powerful point that no simple approach to the complex problems of Europe could possibly be successful, that no matter how unpleasant it was to face the sad fact of complexity, there just was no alternative to understanding the complexities of Europe and to dealing with the complex problems in a complex manner. After a strange lag, Marshall's simple proposal for the United States helping Europe back on its feet captured the imagination of the world. But, as Marshall must have expected, his rather

plaintive plea for recognizing the complexity of the problems received very little attention.

The complexity of the phenomenon of denaturation then unavoidably demands complexity in the techniques of studying denaturation. First of all, since denaturation is concerned with the looser linkages of proteins, anything that is learned about similar linkages in all of protein chemistry must help in the study of protein denaturation.

Happily, Pauling and other members of this year's Council are much more competent to discuss this side of things than I am, and I shall therefore leave this side of things to them.

For the rest, one can only build up a large and complex mass of information by studying systematically and in detail the ways of bringing about and inhibiting denaturation, the changes in protein properties that result from denaturation, the reversibility and the energetics of denaturation. Unfortunately, the number of protein chemists working on denaturation has been so small and the recognition of the character of the necessary work so inadequate, that the systematic, detailed study of denaturation has not progressed very far. I shall review what has been done, restricting myself, for the most part, to those facts and arguments which bring out the few main impressions I wish to leave with you. I shall further not only leave out some subjects altogether, but for reasons of space as well as of emphasis, I shall leave out the details of experimental results. These details are readily found in the original papers. Furthermore, rather than clutter up this general presentation of the status and the problems of the study of denaturation with numerous references to individual papers, I have, for the most part, given references only to a few reviews and to a few recent or summary papers. These reviews and recent papers contain references to almost all the literature of denaturation.

One last more specific introductory remark about the restriction of my subject matter before the presentation of experimental results. Since the phenomenon and the study of denaturation are complex, it is urgent that, at the least, one devise experiments which avoid those complexities which are avoidable. For instance, for experiments on the reversibility of denaturation, it is desirable to choose a protein which is essentially a one component protein and a protein whose denaturation is completely reversible by at least one super-

ficial test. The denaturation procedure ought to be the minimum which causes the changes characteristic of denaturation and the protein ought to be exposed to the denaturation procedure for the minimum necessary time. Yet many discussions of reversibility have been based on experiments with proteins which are known to be non-homogeneous or with proteins whose denaturation was not completely reversed by any test. In some of these experiments the alteration in the protein was not the minimum required to produce the changes in properties characteristic of denaturation, or there was even the possibility of the breaking of peptide or S-S linkages and the splitting off of ammonia, hardly changes which one would expect to be reversible. Many of my own early results were not obtained with what I would now consider good systems. It takes time to discover good materials and to learn simplicity. I shall try in this review to emphasize the least complicated experiments.

THE PRODUCTION OF DENATURATION

No important new ways of bringing about denaturation by either chemical or physical means have been discovered recently. References to the earlier literature can be found in earlier reviews. (Neurath et al., 1944; Anson, 1945.)

Denaturation can be brought about either by the addition of chemical substances or by the application of energy in various forms. The chemical denaturation agents can be classified in three groups. First, there are the denaturation agents which produce ionization. These are usually, but not always, simple acids and bases. Secondly, there is a group of substances which can denature proteins at their isoelectric points. This group includes urea, guanidine hydrochloride, salicylate, thiocyanate and detergents such as sodium oleate or sodium dodecyl sulfate. Thirdly, mixtures of certain organic solvents such as alcohol and acetone with water can denature proteins. Fifty percent alcohol is a better denaturant, and hence a better disinfectant, than absolute alcohol.

In most cases both the ionizing and the neutral denaturing agents can dissolve isoelectric denatured protein, and substances which can dissolve isoelectric denatured protein are denaturing agents. In contrast, neutral denaturing mixtures of organic solvent and

water leave the denatured protein in a precipitated state. Acid mixtures of organic solvents and water, however, can be exceptionally powerful solvents.

The differences in the concentrations of denaturing agents required for denaturation are very striking. A urea solution containing one gram of urea per gram of water will usually denature a protein more slowly than a solution containing a small number of molecules of synthetic detergent per molecule of protein. (The reactions of detergents with proteins, including the denaturation réactions, have been reviewed by Putnam, 1948.)

So far as I know, there have been no systematic investigations of denaturation by organic solvents, investigations in which the nature and concentration of the organic solvent as well as the nature of the protein and the temperature and the pH of the solution, have been varied.

Just as striking as the differences in the concentrations of different denaturing agents needed for denaturation are the extreme differences in the sensitivities of different proteins to the same denaturing agent. One protein may be denatured rapidly in saturated urea solution, whereas another may be quite stable in the same solution. One protein may be quite stable even in extremely acid or alkaline solution. Other proteins may be easily denatured by small degrees of acid or alkaline ionization. Even small species differences between proteins may alter considerably their sensitivities to denaturing agents. It would be very valuable to have a compilation of the varying sensitivities of different proteins to different chemical and physical procedures and to have this compilation brought together with such other information about the denaturants and the protein properties and structures as might seem relevant. Such a compilation, which would require a very considerable amount of time to prepare and which I have been unable to make, might well stimulate a more comprehensive theoretical survey and interpretation of the facts than any which has so far been attempted. It is obvious, for instance, that the chemical denaturants must break some, at least, of the loose bonds which are broken in denaturation. And it has, indeed, often been pointed out that many of the denaturing agents break hydrogen bonds. The various suggestions which have been made, however, are far from the desired comprehensive organization and treatment of all

the known facts. The first result of such an investigation, I believe, would be the recognition of the gross inadequacy of the present scattered experimental information.

The two most common forms of denaturation by physical means are heat denaturation and surface denaturation. Denaturation can also be brought about by the application of very high pressures and by ultraviolet light. Denaturation by ultraviolet light is usually accompanied by some decomposition of the protein. The sensitivities of different proteins to denaturation by physical means, especially to denaturation by heat, vary tremendously, just as do the sensitivities of different proteins to denaturation by the addition of chemical substances.

It is interesting to note that, in general, the proteins of warm blooded animals seem to have higher temperatures of denaturation than the proteins of cold blooded animals, and that the enzymes of microorganisms which can withstand high temperatures are more heat resistant than the enzymes of heat sensitive microorganisms. It is easy to see why heat resistant organisms must have heat resistant proteins, not so easy to see why organisms that are not normally exposed to high temperatures have more heat sensitive proteins. There are, in any case, many exceptions to the rule that proteins are no better than they have to be. Papain can survive, native and active, at higher temperatures than the common animal proteinases. It is often very convenient to use hot-active enzymes, such as papain, to produce desired chemical changes at temperatures at which the ordinary microorganisms do not multiply.

The very marked differences in the sensitivies of different proteins to denaturation by chemical or physical means are of great practical, as well as theoretical, interest, for they are the basis for differential denaturation. This differential denaturation, followed by the relatively easy separation of the native and denatured proteins, has been used in countless protein fractionations since the earliest days of protein chemistry, and has been a most useful, sometimes essential, step in many of the most notable enzyme preparations. Despite the very frequent and successful use of differential denaturation, however, it has almost always been used in a hit or miss way. Protein fractionation based on systematic differential denaturation offers possibilities that have hardly been explored. Innumerable ways are known of varying the chemical

or physical means of denaturation, the conditions of denaturation, and the inhibitors of denaturation.

There is some evidence that enzymes which digest proteins can sometimes convert native proteins into forms which are still native, but more readily denatured than the forms present in the absence of the proteolytic enzymes. Since Linderstrom-Lang has been especially interested in this question, I shall leave discussion of this question to him. (See Christensen, 1952, for the most recent work in the Carlsberg Laboratory and for references to the earlier work of Lundgren and Linderstrom-Lang.)

INHIBITION OF DENATURATION

The techniques of producing denaturation were treated very briefly, with emphasis only on the theoretical and practical implications, since the ways of denaturing proteins have been known for some time and have been frequently reviewed. Less familiar, for the most part, but also of considerable theoretical and practical interest, are the ways of inhibiting denaturation. Furthermore, there has been important recent work on the inhibition of denaturation and, so far as I know, the work on inhibition has never been brought together with any completeness.

The best known and most effective way of inhibiting denaturation is by the removal of water. Dry protein is remarkably resistant to heat and other denaturing procedures. The heat resistance of dry spores is undoubtedly an example of this heat resistance of dry proteins. Recently, there has been some interesting evidence that once a protein preparation has been dried to as little as 5 percent moisture, still further drying of the protein results in increased sensitivity to heat again. (Hornibrook, 1952.)

Also well known, and long known, is inhibition by certain neutral organic water soluble substances, typified by glycerol. Although glycerol in high concentration has been used with great success for the preservation of enzymes for periods of many years, this kind of inhibition of denaturation has been studied very little, and very little is known about the quantitative effects or about the structure needed for inhibition.

One other organic, water-soluble inhibitor should be mentioned, namely common sugar in high concentration. Although sugar does not have such extreme inhibitory effects as glycerol, the effects are definite. (Ball et al., 1943.) One commercial application of sugar inhibition is in the preparation of spray dried eggs for the bakery trade. If eggs are spray dried without the addition of sugar, they lose some of the functional properties which are useful for baking purposes, presumably due to surface denaturation during the spray drying. The addition of sugar stops this loss of functional properties, presumably by inhibiting denaturation. The sugared dried eggs can be used in those baked goods which normally contain sugar anyway, although, of course, they cannot be used for scrambled eggs and the like. (See Brosteau and Eriksson-Quensel (1935) for sugar inhibition of denaturation by drying.)

The denaturation of serum albumin by heat and urea is inhibited by a variety of organic anions typified by the lower fatty acids and related compounds. These anion inhibitors which, unlike glycerol and sugar, are effective in low concentrations, have been studied extensively by Luck and his school. (Boyer, Lum et al., 1946; Boyer, Ballow, et al., 1946; Duggan and Luck, 1948.) They have shown that it is denaturation itself which is inhibited and not merely the precipitation of denatured protein. They have also studied the effects of varying the structure of the inhibitor and of varying the protein. Serum albumin is peculiarly susceptible to this new kind of inhibition. It was indeed, the practical need of stabilizing solutions of human serum albumin during the war which led to the work of Luck.

Neutral Salts : There have been many reports in the literature of neutral salts either promoting or inhibiting denaturation. In most cases the promotion or inhibition has been small. There is no doubt, however, that the inhibition can sometimes be very impressive.

The most striking case of inhibition by sodium chloride is my own observation that a saturated sodium chloride solution of edestin can actually be boiled without the edestin being denatured Edestin can, indeed, be extracted from its meal with boiling sodium chloride solution and then crystallized as usual. When edestin is heat coagulated in solutions of lower salt concentration, the exact temperature

of coagulation depends very much on the salt concentration. Heat coagulated edestin is not soluble in saturated salt solution.

Especially interesting are the recent studies on the stabilization of proteins by neutral salts with bivalent cations, such as calcium and manganese. (References in Gorini and Felix, *in the press.*) After Gorini and Fromageot showed that certain bacterial proteinases could be stabilized by calcium, similar stabilization of trypsin was observed by Gorini, and by Bier and Nord. Gorini and his colleagues have, indeed, not only accumulated many examples of cation stabilization of various proteins with and without enzymatic activity but they have provided evidence that the stabilization is due to the inhibition of denaturation.

The stabilization by bivalent cation must be distinguished from the sort of activation of dipeptidases studied by Emil Smith, in which the bivalent cations promote the combination of enzyme and substrate, not the stabilization of the enzyme protein.

Characteristic of the stabilization with bivalent salts is its specificity. In the cases studied, one or two ions were found to have very marked stabilizing effects, whereas the other ions tried were only slightly effective or not effective at all.

Specific stabilization with bivalent ions is probably related in its mechanism to the specific effects of certain ions in promoting crystallization. The best known case of this is the effect of zinc in promoting the crystallization of insulin. Just as zinc seems to be an essential in the structure of insulin crystals as they are formed under certain conditions, so we may have to consider certain ions essential to the stable native structures of some proteins.

It would obviously be desirable to have direct measurements of specific ion effects on the equilibrium between the native and denatured forms of some protein. Eisenberg and Schwert (1951) studied the effects of some salts on the equilibrium between native and denatured chymotrypsinogen, but not, however, any case of marked, specific cation effect.

Pressure : It has long been known that proteins can be denatured by the very high pressures of the sort studied by Bridgman. More recently, Johnson and his colleagues (references in Foster et al., 1949) have demonstrated that more modest pressures, pressures of a mere 5-10,000 atmospheres, can inhibit denaturation. Perhaps

the most interesting case studied by Johnson was the restoration of the luminescence of living bacteria. First, the luminescence was quenched by heating the bacteria to the minimum temperature which produced quenching. Then, while the bacteria were kept at the quenching temperature, luminescence was restored by the application of pressure. It was supposed that native luciferase, the enzyme responsible for luminescence, is in equilibrium with its denatured, inactive form and that the equilibrium is shifted to the denatured side by a rise in temperature and to the native, active side by a rise in pressure.

Effects of pressure must be related to changes in volume. The changes in volume have been studied particularly in the Carlsberg Laboratory (References in Christensen, 1952). I shall leave it to Linderstrom-Lang to discuss the significance of the volume changes and their relation to the pressure effects observed by Johnson and his colleagues.

Prosthetic Groups : If my memory of odd observations serves me right, in those cases in which simple native protein has been prepared from conjugated protein by the removal of the prosthetic group, the simple native protein has been less stable than the conjugated protein from which it was prepared. Apparently, the prosthetic groups can play a role in promoting a stable native structure similar to the role of calcium studied by Gorini.

Inhibition of Precipitation : Many cases of inhibition of heat coagulation, for instance inhibition by nucleic acids, are cases of the inhibition of the precipitation of denatured protein, not the inhibition of denaturation itself. The denatured protein is usually given a positive or negative charge by the substance which inhibits coagulation.

Stabilization of Enzymes : It is often observed in the course of the purification of enzymes that the enzyme becomes more unstable as impurities are removed, with sometimes a quite lamentable rise in instability at the last stages of purification. Everybody who has worked extensively on the isolation of enzymes can remember such experiences. The mechanism of the very real protection by impurities is usually unknown or unproven. One can see how the

removal of protective substances, such as calcium salts or inhibitors of the type studied by Luck, may be the cause of the decrease in stability as the enzyme is purified. It is also possible, of course, that there are many specific inhibitors of denaturation present in crude extracts whose nature is still unknown. Furthermore, purification may involve the elimination of two other kinds of inhibition which are familiar enough, although they have rarely, if ever, been studied carefully, inhibition by enzyme substrates and by adsorption.

Substrates : Stabilization of an enzyme by combination with its substrate may be effective enough to warrant carrying out most of the isolation of the enzyme in the presence of excess substrate. Although it is reasonable to suppose that inhibition of destruction of the enzyme is due to inhibition of denaturation, in most cases this has remained merely a supposition. The person trying to isolate the enzyme has been concerned with keeping his enzyme alive, not in studying denaturation. On the other hand, the student of denaturation can find many hints for possible inhibitors of denaturation in the vast literature of enzyme purification. Such hints ought to be followed up by direct measurements of denaturation which, in most cases, are lacking.

Adsorption : An example of stabilization by adsorption can be given from my own experience in the purification of spleen cathepsin (Anson, 1940). In the first step of the purification, a suspension of spleen in 0.3 saturated ammonium sulfate solution was acidified and heated. A great deal of the soluble protein was denatured and digested, whereas the cathepsin, which was adsorbed on to the spleen solids, remained stable. The cathepsin, along with a part of the other adsorbed protein, was then eluted by the addition of alkali and the spleen solids removed by filtration. In the absence of added ammonium sulfate, the cathepsin was not adsorbed. (An early, if not the first, case of salt-promoted adsorption.) If not adsorbed, the cathepsin was rapidly destroyed in ammonium sulfate solution at the pH and temperature chosen. Salt-promoted adsorption has since been studied under much better defined conditions, especially in the laboratory of Tiselius. Adsorption-promoted stabilization has not, to my knowledge, been studied under well-defined conditions. Nor was even the mechanism of the stabi-

lization of cathepsin properly demonstrated to be protection from denaturation.

Crystallization : Working with another enzyme, carboxypeptidase, I observed an example of still another kind of stabilization, stabilization by crystallization. Crystals of carboxypeptidase suspended in cold water, with toluol as a preservative, are stable for many years. If the crystals are dissolved by the addition of sodium chloride, the protein slowly changes into an insoluble, inactive form.

The means of denaturation or inhibition of denaturation and the differing sensitivities of proteins are important for the theory of denaturation. A complete theory of denaturation ought to explain why certain substances are denaturing agents, whereas others inhibit denaturation; why certain denaturing or inhibiting agents are more effective than others; why certain proteins are more readily denatured by a particular method than others are, or more readily protected from denaturation.

The same phenomena which are important for the theory of denaturation, and which would be more important were they better studied, are also, as we have seen, of practical importance in the every day fractionation and purification of proteins. Denaturation tests, I might add, are also of practical use in checking the purity of a biologically active protein after it has been isolated. If the protein and the enzyme are identical, susceptibility to denaturation and to inactivation ought to go together under very varying conditions, and half denaturation ought to result in half inactivation.

CHANGES IN PROPERTIES

In order to give, in a bref review, some true picture of denaturation in its varied and complex aspects, and to point out the nature of the problems and opportunities for further study, I have had to make the choice, as I said in the Introduction, of merely outlining the facts and discussing their implications, and of leaving out most details. The necessity for leaving out details was especially clear when I came to the treatment of the many changes in properties which proteins undergo when they are denatured. Indeed, apart

from composition and molecular weight, practically all the properties of a protein are changed by denaturation, which makes the subject a large one. These many changes in properties demonstrate the role of the structure produced by the looser linkages of native protein in influencing, sometimes slightly, sometimes greatly, all the familiar characteristics of native proteins.

Solubility : The most obvious change in the protein when the protein is denatured is the change from solubility at the isoelectric point in water or salt solution to practically complete insolubility in water or salt solution. This is the change so easily observed in common heat coagulation. In order to dissolve the denatured protein, it is necessary to add acid or alkali or one of the neutral denaturing agents such as urea or detergent. If the protein is denatured in the presence of a solubilizing agent, then the denatured protein must be brought to its isoelectric point and the solubilizing agent removed, in order that the loss of solubility be made apparent. Strangely enough, the exact reasons for the loss of solubility due to denaturation have never been demonstrated. One can only suggest possible reasons. The hydrophilic groups of the native protein are certainly not eliminated when the protein is denatured. To account for the loss of solubility, one must assume that either hydrophobic groups are exposed by denaturation, or that conditions are made more favourable for the hydrophilic groups of one protein molecule to react with the hydrophilic groups of another molecule than for the hydrophilic groups to react with water. Similarly, the substances which dissolve denatured protein must either add or produce hydrophilic groups, such as ionized groups, to the protein, or they must make the conditions more favourable for the hydrophilic groups to react with water than with the hydrophilic groups of other molecules.

Biological Activity : The second most obvious change in protein properties brought about by denaturation is the loss of the activity of biologically active proteins. Almost all enzyme proteins are easily denatured and all enzymes that are denatured are inactivated by the denaturation. In working out the isolation of an enzyme, it is necessary to know the stability limits of the protein at the different stages of purification. The relation of pH to stability varies greatly

from case to case. In most cases, however, the enzyme proteins are stable only in the region of neutrality. Hence most enzyme purifications are carried out in this region. Fractionation with neutral salts has long been popular, and recently heavy metal fractionation has also proven very useful, partly because such fractionations can be carried out in neutral solution and at room temperature. When alcohol or acetone is used for fractionation, the fractionation must be carried out at low temperature, as has been known since the crystallization of hemoglobin from cold alcohol-water solution eighty years ago. Some enzymes are so unstable that their fractionation is best carried out in the cold, no matter what the technique of fractionation.

As I have already pointed out, since enzymes are inactivated by denaturation, the inactivation of enzymes by denaturation can be used in tests for the purity of an isolated enzyme.

The rate of heat destruction of bacteria and viruses has the high temperature coefficient characteristic of denaturation, strong evidence that the pace setting reaction is protein denaturation. As is well known, the temperatures used for common pasteurization are the common temperatures of denaturation. Those relatively uncommon organisms which flourish at the usual pasteurization temperatures must have proteins with unusually high temperatures of denaturation. Dry spores which are not readily hydrated are destroyed only at higher temperatures, in harmony with the known heat stability of dry proteins.

In killing microorganisms in foods by heat sterilization, one usually wants to minimize the chemical changes brought about by the heat. This is best done by heating at a relatively high temperature and reducing the time during which the food material is kept hot. Since heat denaturation has a much higher temperature coefficient than the other chemical changes which it is desired to suppress, the ratio of the rate of denaturation to the rate of the other changes is higher, the higher the temperature. With sterile filling of cans now commercially feasible, it is feasible to heat very rapidly any pumpable material to a high temperature at which sterilization takes place in a minute or less (not time enough for the undesired side reactions to go far), and then to cool the material rapidly and, operating in a sterile atmosphere, to fill the sterilized and already cooled material in sterile cans.

Digestibility : Just as it has long been known, as a result of simple and crude observations, that heat can make proteins insoluble and enzymes inactive, so it was observed very early that heat can make proteins digestible. Actually, more careful modern studies have shown that whether or not the digestibility of a protein by proteinases is changed by denaturation depends on which protein and which proteinase are being used. Native edestin is digested by trypsin almost as readily as denatured edestin. In contrast, native beef hemoglobin is not digested at all by trypsin, whereas the denatured form of the protein is digested very readily. Intermediary cases are known in which both the native and denatured forms of a protein are digested, the denatured form, however, being digested more rapidly than the native.

Linderstrom-Lang will perhaps make a few remarks on his own views on the relation of denaturation to digestion by proteinases, and especially on his view that even the apparent digestion of native protein is really digestion of denatured protein.

The changes in solubility, activity and digestibility, as I have said, were observed crudely in very early days. More sophisticated observations were needed to discover the two remaining changes which will be discussed, the change in reactivity of protein groups, and the change in the shape of the protein molecule.

SH Groups : Studies of the effects of denaturation on the reactivity of protein groups began with the qualitative observation that denatured, but not native, egg albumin gives the red color with nitroprusside which is given by cysteine and other simple SH compounds. Since then, the SH groups have been studied quantitatively, the factors which influence the reactivity of SH groups in both native and denatured proteins have been studied and the results obtained with SH groups have been extended to other protein groups. References to much of the literature in this field, which is so important for the theory of denaturation and of protein structure, can be found in an earlier review of mine. (Anson, 1945)

The earliest and the most extensive work on the reactivity of protein groups has been done with SH groups, because the study of these groups is relatively easy. SH groups are the strongest reducing groups in proteins. It is easy to oxidize SH groups without oxidizing other reducing groups, such as the OH groups of tyrosine.

And there exist various specific non-oxidizing SH reagents, such as chloromercuribenzoate, as well as the convenient nitroprusside colorimetric test for SH groups.

First of all, a variety of independent quantitative methods were developed for the estimation of SH groups. It was found that the same value for the SH groups of denatured egg albumin was obtained by the application of many titration methods, despite the use of different oxidizing and non-oxidizing SH reagents for titration of the SH groups, and despite the use of different solvents and reaction conditions. Evidence was obtained that in the titrations all the SH groups of egg albumin were titrated. The SH titration is thus an estimation of the cysteine content of egg albumin, an extraordinarily easy estimation which can be carried out in a few minutes with a few milligrams of unhydrolyzed protein.

For the various procedures used to titrate the SH groups of denatured egg albumin, the reagents and the reactions were chosen to favor easy and complete reaction of the SH groups. By studying the SH groups under conditions progressively less favourable for their reaction, however, it was possible to show that the reactivity of the SH groups of denatured egg albumin is quite dependent on both the reagent and the reaction conditions, and that reagents and reaction conditions can be chosen which permit complete titration of the SH groups of free cysteine, but only partial or no titration of the SH groups in the cysteine of even denatured egg albumin.

To give a typical example of the effect of varying the reagent, although the Uric Acid Reagent, which Folin used to oxidize the SH groups of cysteine, cannot oxidize the SH groups of denatured egg albumin in a solution of the synthetic detergent, Duponol PC, yet the same SH groups in the same solution are readily oxidized by ferricyanide, a somewhat stronger oxidizing agent. To give a typical example of the effect of varying the solvent, the Uric Acid Reagent which cannot oxidize the SH groups of denatured egg albumin in Duponol solution can readily oxidize the SH groups of denatured egg albumin in urea solution. In contrast, the SH groups of *native* egg albumin are not oxidized even when the oxidizing agent is ferricyanide and the solvent urea solution. Only when in time the protein becomes denatured by the urea solution, as shown by solubility tests, do the SH groups become oxidized by the ferricyanide.

Thus, although there is clearly a gross increase in the reactivity of the SH groups of egg albumin brought about by denaturation, still there can be no question of the SH groups of the protein being reactive or unreactive in any absolute sense. The SH groups of denatured egg albumin are in any case less reactive than the SH groups of cysteine and the SH groups of denatured egg albumin can be made to react with an SH reagent either not at all or completely and quickly, depending on the choice of reagent and reaction conditions. The marked effects of varying the solvent is one more illustration of the point made early in the review, that there is no such thing as *the* denatured form of a protein, that rather denatured, like native protein, can in different chemical environments exist in many different forms with very different properties.

The reaction of the SH groups of denatured egg albumin with a given reagent in a given solvent can be decreased by aggregation of the protein and increased by even slight hydrolysis of the protein with pepsin. All the SH groups of heat denatured egg albumin in neutral water solution containing no urea, detergent, guanidine hydrochloride, or the like, are readily oxidized by ferricyanide so long as aggregation of the protein has not taken place. As aggregation of the denatured protein proceeds, however, more and more of the SH groups fail to react with ferricyanide present in slight excess. These groups which fail to reduce dilute ferricyanide are, again, not unreactive in any absolute sense, since they can be oxidized by ferricyanide in much more concentrated solution or by a dilute solution of porphyridin, a stronger oxidizing agent than ferricyanide.

We see, then, that aggregation, which is a formation of loose inter-protein linkages between different molecules of denatured protein, has the same sort of effect in decreasing the reactivity of SH groups as changing the protein from the denatured to the native state, which is the formation of loose intra-protein linkages between different groups of the same protein molecule.

The fact that the reactivity of the SH groups of egg albumin is increased by denaturation of the protein is evidence that the reactivity of protein SH is very dependent on neighboring protein structure. The same conclusion must be drawn from the fact that the reactivity of the SH groups of already denatured egg albumin can be increased still further by slight hydrolysis of the protein with

pepsin. Ferricyanide, as we have seen, can oxidize all the SH groups of denatured egg albumin in neutral Duponol solution. In pH 4.8 Duponol solution, however, only a few of the SH groups are oxidized by ferricyanide. In the same pH 4.8 Duponol solution, however, almost all of the SH groups of the slightly hydrolyzed egg albumin are oxidized by ferricyanide.

Two other instances of the effect of neighboring protein structure on the reactivity of protein SH groups may be mentioned at this point. First, many cases are known of some SH groups of a protein molecule reacting with a particular reagent under particular conditions whereas other SH groups of the same molecule, presumably with different neighboring structures, do not react. Secondly, many cases are known of the SH groups of one denatured protein reacting readily and completely with an SH reagent, whereas the SH groups of another protein, with presumably different neighboring structures, do not react at all with the same reagent under the same conditions.

SH of Native Protein: As we have seen, the SH of native egg albumin does not give a pink color with nitroprusside and is not oxidized by ferricyanide, even in urea solution. Because of such facts, it was long supposed that the SH of native egg albumin is unreactive in some absolute sense, possibly due to the SH groups being inaccessible in the interior of the native, globular molecule. It is now known, however, that some of the SH groups of native egg albumin can be made to react with various reagents, such as iodoacetamide, and that all the SH groups of native egg albumin can be oxidized by iodine. Iodine is best added in cold 1M KI solution. In cold KI solution iodine does not react with free tyrosine or with tryptophan in proteins, and thus is a specific reagent for SH. Under these conditions, furthermore, iodine does not oxidize the protein SH beyond the S-S stage. When iodine is added to native egg albumin in just the amount theoretically needed to oxidize the SH to S-S, all the iodine is used up and no S-S can be detected in the protein after denaturation. The rate of sedimentation of the native egg albumin is not changed by the oxidation of the protein's SH groups. Thus the native protein must contain an even number of SH groups, and each S-S group must be formed from two SH groups of the same molecule, not from two SH groups

from two molecules. Similarly, as shown by Frederique and Desreux (1947), the SH of native egg albumin can be abolished with chloropicrine, without any change in the molecule detectable by measurements of diffusion, sedimentation or viscosity.

Tobacco mosaic virus, like egg albumin, gives the nitroprusside test only when denatured, and the SH groups of native tobacco mosaic virus, like the SH groups of native egg albumin, can be oxidized by iodine. This iodine-oxidized virus was found to be fully infectious, the first case of a chemically altered, but still active virus. Plants infected with oxidized tobacco mosaic virus, however, produced the normal SH virus; oxidation by iodine did not result in the synthesis of a mutant.

Although the SH groups of native egg albumin do not react with nitroprusside or with some common SH reagents such as ferricyanide and mercuribenzoate, there are other native proteins whose SH groups react readily with nitroprusside and all the common SH reagents, and other native proteins some of whose SH groups react readily whereas others do not. Thus, in the case of native protein, as in the case of denatured protein, there is no such thing as absolute reactivity or absolute non-reactivity. Whether or not an SH group of a native protein reacts depends on the reaction conditions and the neighboring protein structure; in other words, on the reagent used, on the conditions under which the reagent is used, on the location of the SH groups in the protein, and on the nature of the protein.

Other Groups: There are protein groups other than SH groups which become detectable by a particular technique only when the protein becomes denatured; S-S groups that became reducible only when the protein is denatured, phenolic groups of tyrosine which react with hydroxyl ions or Folin's Phenol Reagent (Herriott, 1947) only when the protein is denatured. References to the earlier studies of S-S and tyrosine groups were given in my earlier review. Later studies have added nothing essentially new.

In my earlier review, I said that it might well be possible even to discover reagents which, under suitable conditions, would react with the amino and carboxyl groups of denatured but not of native protein. Porter (1948) has indeed obtained results with amino groups quite similar to those obtained with SH groups. All the

free lysine amino groups of denatured β-lactoglobulin and of serum globulin react with Sanger's Reagent, fluorodinitrobenzene. Over a third of the amino groups of the native forms of these proteins, however, do not react with Sanger's Reagent.

Similarly, Porter (1950) has shown that Sanger's Reagent reacts with all the four histidine iminazole groups of denatured β-lactoglobulin but with only two of the four iminazole groups of the native form of β-lactoglobulin.

Steinhardt and Zaiser (1951) have shown that many of the carboxyl groups of hemoglobin are not titratable in the native form of the protein but became titratable when the protein is denatured by acid.

It now seems fairly safe to predict that given sufficient searching, proteins, reagents and reaction conditions can be found which will permit the demonstration of an increase in the reactivity of any protein functional group when the protein is denatured, perhaps even the demonstration of no reaction in the native form with complete reaction in the denatured form.

Some of the factors which influence the reactivity of SH groups have been shown to influence the reactivities of other protein groups as well. The influence of neighboring protein structure is shown by the observation already mentioned that Sanger's Reagent combines with some but not all of the lysine amino groups of β-lactoglobulin and by the further observation of Porter and Sanger (1948) that this same reagent which does not react with all the lysine groups of native β-lactoglobulin does react with all the lysine groups of native hemoglobin. That lysine groups can react or not depending on what reagent is used is illustrated by the observation (Porter, 1948) that the lysine groups of native lactoglobulin which do not combine with Sanger's Reagent can easily be acetylated.

In general, the factors which influence the reactivity of protein SH groups have been studied, not thoroughly, but at least more thoroughly than the factors which influence the reactivities of other protein groups. One would like to know the effects on the reactivities of all protein groups of adding urea, detergent or guanidine hydrochloride to denatured protein and of aggregating or digesting the protein.

It must be remembered, in connection with the generalization of any results, that alterations of the chemical environment, such as

changes of pH or addition of urea, in addition to changing the reactivities of protein groups by altering protein structure, also change the reactivities of protein groups directly and in ways which may vary from group to group. Thus the ease with which the SH even of free cysteine and the OH even of free tyrosine are oxidized is very much influenced by the pH. And the ease with which the OH of free tyrosine is oxidized can be influenced by the addition of urea.

If it should prove true that the addition of substances such as guanidine hydrochloride or slight enzymatic hydrolysis of the protein can increase the reactivity of all protein groups, as they increase the reactivity of protein SH groups, then perhaps the experience with SH groups which has been applied to the easy estimation of cysteine in proteins can also be applied to the easy estimation in intact or slightly hydrolyzed proteins of other amino acids with characteristic groups.

There are, indeed, many practical applications in protein chemistry of any knowledge of the reactivities of protein groups. And, on the theoretical side, any theory of protein structure and of the change in protein structure which is denaturation must explain why denaturation increases the reactivity not only of SH groups, but of other, perhaps of all other protein groups as well. The theory, furthermore, must explain the known facts about how reaction conditions and neighboring protein structure influence the reactivities again not only of SH groups but of other groups as well. I have outlined the facts, of necessity very briefly. And I leave it to the structural experts attending the Council to suggest the theoretical implications.

Prosthetic Groups : The discussion of protein groups so far has dealt with the groups of amino acids bound into the protein by peptide linkages. The conjugated proteins, such as hemoglobin, contain non -amino acid groups, the so-called prosthetic groups. In hemoglobin, the prosthetic group, heme, is bound to the simple protein, globin. The properties of heme are changed greatly by its combination with globin. Free heme cannot combine reversibly with oxygen, it is rapidly oxidized to the ferric state by oxygen, and its spectrum is changed by its combination with globin. When reduced hemoglobin is denatured, that is, when it is converted into

globin hemochromogen, the properties of heme attached to the globin are changed by the denaturation of the protein. The heme, although still combined with globin, can no longer combine reversibly with oxygen and the spectrum of heme attached to denatured globin is different from the spectrum of heme attached to native globin. Furthermore, the compound of heme with native globin is a tight compound whereas the compound of heme with denatured globin is a loose, easily dissociated compound. Thus denaturation influences the properties of prosthetic groups as well as of amino acid groups. It was, indeed, the study of the effects of the denaturation of hemoglobin on the properties of heme which led to the discovery of some of the essential facts about the hemochromogens as a class.

Size : As shown by measurements of the osmotic pressures of denatured proteins in urea solutions, the molecular weights of the non dissociable proteins are not changed by denaturation. (References in Neurath et al., 1944.)

The same conclusion must be drawn from the fact, to be discussed later in this paper, that the equilibrium between the native and denatured forms of a protein, in the cases studied, is independent of the protein concentration.

Shape : There is no doubt that in a general way protein molecules are opened up by denaturation.

Astbury, Dickeman and Bailey (1935) first pointed out as a conclusion from X-ray data that the globular proteins on being denatured became more like the fibrous proteins.

That surface denaturation of protein results in very considerable opening up of protein molecules is amply proven by the work on the surface films, although there is some doubt about the detailed interpretation of the results. (References in the review of surface films by Bull, 1947.)

The observation that the viscosity of an urea solution of unaggregated protein is increased as the protein becomes denatured (Anson and Mirsky, 1932) led to the suggestion of Mirsky and Pauling (1936) that the increase in viscosity is due to an opening up of the protein molecule.

Although there have since been many studies of the viscosities of urea solutions of proteins, there is still doubt about how much of

the change in viscosity is due to change in shape and how much to change in hydration.

The opening up of the protein molecule in denaturation, however, has also been demonstrated by combined measurements of diffusion and sedimentation. (Rothen, 1942.) In the classical Svedberg calculation from such measurements, any effect of hydration is eliminated and the result, expressed as some arbitrary change of shape, is due solely to some sort of change in shape.

Protein fibers consist of elongated fibrous molecules. Astbury and his colleagues, after realizing as a result of their X-ray studies that denaturation makes corpuscular proteins more fibrous in shape, actually made fibers from denatured corpuscular protein, and ICI, to complete the story, is now making textile fibers from the denatured form of peanut protein, which is globular in its native state, elongated in its denatured state. Much earlier, von Weimarn, knowing nothing about the shapes of native and denatured proteins, had observed that fibers could be made from urea solutions of many proteins. (References in the review of the subject of protein fibers by Lundgren, 1949.)

If data about the shape of molecules of denatured protein are to become of increased use for theorizing about denaturation, information somewhat different from that, for the most part, now available in the literature must be obtained. First of all, one would want to know the change in shape due to reversible denaturation, uncomplicated by other changes in the protein. It is pretty sure that irreversible changes following reversible denaturation can result in further changes in the shape of the molecule. Secondly, it would be worth while knowing (still sticking to reversible denaturation) how the change in shape depends on the original structure of the protein and on the denaturation conditions and whether, once the protein is denatured by heat or by some denaturing agent, raising the temperature still further or increasing the concentration of the denaturing agent still further results in still further changes in the shape of the protein molecules.

Other Properties: I have not, by any means, discussed all the changes in properties which proteins undergo when they are denatured. There are the changes in immunological properties, the interesting decrease in species specificity and antigenicity, and the

changes in reactions with antibodies; the changes in reversible combinations with various ionic and non-ionic solutes; the changes in absorption of infrared and ultraviolet light, in optical rotation, flow birefringence and light scattering; the changes in rates of diffusion and sedimentation and migration in an electric field; the changes generally, in all those properties which depend on the number and distribution of the charges of the protein molecules, and on the shape and hydration of the molecules. One can, indeed, go through almost the whole of protein chemistry in the study of denaturation. But in any study of denaturation, as in any brief discussion of denaturation, one must limit oneself in the number of protein properties which are dealt with. The art of the experiment is to select for measurement those properties which are easiest to measure well and the measurement of which is most significant for the problem being studied.

REVERSIBILITY AND ENERGETICS

Reversibility : The denaturation of some proteins has been reversed completely. That is to say, all the denatured protein was converted into « renatured » protein which, by such sensitive tests as were applied, could not be distinguished from the original native protein. The denaturation of other proteins, such as egg albumin, in contrast, has never been reversed at all. In between are cases of only a part of the denatured protein being convertible back to native protein or of the renatured protein having the general properties characteristic of native protein but still being demonstrably different from the original native protein, for instance, in ease of denaturation or digestion or in solubility. (References to papers on the quantitative differences between native and renatured proteins in Roche and Chouaiech, 1940, and Neurath et al., 1944.)

Most interesting for the theory of denaturation are the cases of apparently complete reversal of denaturation, especially those cases in which mobile equilibria between the native and denatured forms of the protein have been demonstrated. It is these equilibrium cases which I want to emphasize, especially five of them, trypsin (Anson and Mirsky, 1933 a), trypsin inhibitor (Kunitz, 1948) and chymotrypsinogen (Eisenberg and Schwert, 1951) in solutions of

different temperatures, hemoglobin in solutions of different salicylate concentrations (Anson and Mirsky, 1933 b), and pepsinogen in solutions of different pH's (Herriott, 1938).

In all five cases, definite mobile equilibria were demonstrated. The percentage denaturation depended only on the final temperature, or salicylate concentration or pH. It was independent of the time, and once a steady state was reached, was the same whether the equilibrium was approached from the native or from the denatured side.

In those cases in which the effect of protein concentration was studied, the percentage denaturation was independent of the protein concentration.

Whenever the effect of adding a denaturation agent was studied, addition of a denaturation agent caused a shift of the equilibrium to the denatured side.

Whenever the percentage denaturation was measured by two independent methods, e. g. by the formation of insoluble protein and by the loss of enzymatic activity, the same value for the percentage denaturation was obtained by the two methods.

It has long been known that the rate of denaturation can be very sensitive to the temperature. Similarly, the equilibria between the native and denatured forms of trypsin, trypsin inhibitor and chymotrypsinogen were found to be very sensitive to the temperatures, the heats of reaction calculated from the Van 't Hoff constants ranging from 37,000 to 143,000 calories per mole.

In contrast, the equilibrium between the native and denatured forms of hemoglobin in salicylate solution was found to be practically independent of the temperature, but extremely sensitive to the salicylate concentration.

The enzymes and enzyme precursors, after being denatured and renatured, became soluble and crystallizable again and had their original activities, or potential activities. In the cases of trypsin inhibitor and chymotrypsinogen, the renatured proteins were found to be indistinguishable from the original native proteins by the most sensitive solubility tests.

The best of the studies of denaturation equilibria is the study of the equilibrium between native and denatured trypsin inhibitor by Kunitz. For Kunitz measured not only the mobile equilibria but also the individual rates of denaturation and of its reversal. He

accordingly was able to calculate the changes in free energy and entropy in denaturation, as well as the heat of denaturation. Similar measurements and calculations were made in the study of the denaturation of chymotrypsinogen. But here the rates of denaturation and its reversal were much faster than the rates of the denaturation of trypsin inhibitor and of its reversal, and so the measurements of rates could not be made as accurately.

In a number of cases, it was observed that denaturation was completely reversible only if the protein was not held in the denatured form too long. In time, reversible denaturation went over into irreversible denaturation.

In other cases, it was observed that completely reversible denaturation and good equilibria were obtained only if the system were kept within very definite limits of pH and temperature. Thus, in order to obtain complete reversibility and good equilibria, it is necessary not only to choose suitable proteins but also suitable conditions of denaturation and its reversal.

Given suitable proteins and suitable conditions, however, it is now no longer in doubt that a good equilibrium between the native and denatured forms of protein can be established, despite the complexity of protein structure and the complexity of the structural changes involved in denaturation. It is the denaturation equilibria which are most useful for the study of the all or none character of denaturation and of the energetics of denaturation. It is to be hoped that many more clear cut denaturation equilibria will be discovered.

The question, of course, arises why denaturation in some cases is not reversible at all. One reason is surely operative sometimes, that the denaturation procedure, in addition to producing the structural changes characteristic of denaturation, goes on to produce further irreversible changes. We have seen that in a few cases, the denatured protein is convertible completely into native protein again only if the protein is exposed to the denaturation procedure for a limited time. In other cases, actual decomposition of the molecule, of one sort or another, has been demonstrated. A second possible reason for the irreversibility of denaturation is that however gentle and quick the denaturation, the denatured protein just has, in some proteins, undergone structural changes which are too complex and too far reaching to permit practical reversal of the changes. Finally,

it is possible that the procedure chosen for the reversal of denaturation also brings about aggregation of the protein, and that this aggregation stops the reversal reaction.

That denatured protein is sometimes converted into undoubtedly native protein which, however, is not identical with the original native protein is not surprising. The result must mean that some of the bonds broken by the denaturation procedure are not reformed or are not reformed in the original way, but that such degree and kind of reformation of bonds as does take place is adequate to bring back the general native properties of the original protein.

An example of the incomplete reversal of denaturation is given by the preparation of native globin from hemoglobin (Anson and Mirsky, 1930). The hemoglobin is first denatured in acid acetone solution. Then the denatured globin is separated from the heme and is converted into native globin by neutralization. Some of the protein is precipitated as denatured globin by the neutralization. The remaining native globin, as shown by Roche, is not identical with the original globin since it does not combine with heme to form hemoglobin identical with the original hemoglobin.

It might be worth while investigating the reasons for the incompleteness of the reversal of the denaturation of globin. Perhaps some of the SH of the denatured globin is oxidized during the preparation.

All or None : By saying that denaturation, in certain cases, at least, is an all or none reaction, I mean to express the experimental observation that in an equilibrium mixture of native and denatured protein, the same value for the percentage denaturation is obtained when the percentage denaturation is measured by two different ways. This observation suggests that there are, as a first approximation, two types of protein in the equilibrium mixture, one type, the native protein as it exists in that chemical environment, the second type, the denatured protein as it exists in that chemical environment. If there are intermediate forms, they exist in relatively small quantities or, as is not out of the question, they are not detected by the tests used.

By « all denatured » it is not meant that the denatured protein may not be more uncoiled, or the like, in some other chemical environment, but that all the denatured molecules have been con-

verted into a form with all the characteristic properties of denatured protein.

The comparison of the percentages of denaturation estimated by two different methods is not necessarily significant if the protein used is a mixture of several related proteins or if the reversible denaturation is followed by or accompanied by a series of irreversible changes. Furthermore, an apparently negative test does not necessarily mean that there has been no change in a given protein property. For instance, the SH of egg albumin is not oxidized by the Uric Acid Reagent even if the protein is in Duponol solution. Yet that the protein is denatured by the detergent is shown by suitable solubility tests. And that the SH groups are changed in reactivity is shown by the fact that they are oxidized by ferricyanide after but not before the addition of the detergent. Thus the all or none notion seems to be contradicted by the tests carried out with the Uric Acid Reagent and yet is in harmony with the more suitable tests carried out with ferricyanide.

There is not much profit trying, on the basis of present data, to make more precise either the structural meaning or the proof of the all or none character of denaturation. What are needed are more measurements of many different properties of many different native and denatured proteins in many different equilibrium mixtures.

Energetics : Discussions of the energetics of denaturation have usually been based on studies of the kinetics of denaturation. (Literature in Levy and Benaglia, 1950.) A better basis, perhaps, for such discussion is provided by the still meager data on the effect of temperature on the equilibrium between the native and denatured forms of a protein. Best of all, such data are combined, as they were by Kunitz (1948), with measurements of the rates at different temperatures both of denaturation and of its reversal. The important facts are that under the conditions chosen, temperature has a very great effect on the equilibria between the native and denatured forms of trypsin, trypsin-inhibitor and chymotrypsinogen but that temperature has practically no effect on the equilibrium between the native and denatured forms of hemoglobin in salicylate solution. I should like to make some remarks on these temperature effects.

1) Native protein is in mobile equilibrium with hydrogen and hydroxyl ions, water molecules, and other constituents of the solvent solution. When the bonds which keep the protein molecule in its native structure are broken by the denaturation procedure, all the equilibria are changed, since the changed protein molecule has different affinities for the various constituents of the solution with which it is in equilibrium. The heat of denaturation calculated from the effect of temperature on the denaturation equilibrium includes not only the heat of the breaking of the loose intra-protein bonds but the heat of the many changes in equilibria between the protein and constituents of the solution. There is no observable heat of denaturation *per se*, or even any way of separating the heats of the breaking of the intra-protein linkages, the breaking or forming of linkages with constituents of the solution, and the changes in strength of the bonds that are neither made nor broken. All one can do is to measure the heats of denaturation under a great variety of conditions and try to draw some conclusions from the whole complex of facts. Unfortunately, the data now available are, as yet, hardly varied enough to permit much by way of conclusions.

2) The bonds broken by denaturation, whether they be intra-protein bonds or bonds with constituents of the solution, must be weak bonds. Strong bonds, such as peptide or S-S bonds, could not possibly be readily broken and reformed by the procedures used for denaturation and its reversal.

3) Denaturation, in some cases at least, must involve changes in many weak bonds. The heat of denaturation, when high, must be the sum of the small heats of the breaking of many weak linkages.

4) Denaturation, in some cases at least, must involve the breaking of many weak bonds in an all or none way. Otherwise, the great effect of temperature on the equilibrium would not be observable. The heat of ionization of protein per mole of *protein*, must be high, since many groups of each molecule are ionized. But one cannot observe the great effect of temperature on the equilibrium between the completely unionized and completely ionized forms because all the intermediate ionized forms are stable and the completely unionized and completely ionized forms do not have special properties by which they can be identified and measured.

5) Since the great temperature effects observed in the study of denaturation depend on the breaking of many weak bonds in an all or none way, it is not surprising that similar great temperature effects are not often observed in chemical studies. One must have a large, complex molecule to have the possibility of the breaking of many bonds, and a molecule which undergoes a very special kind of change to have the possibility of the bonds being broken in an all or none way. The only other similar cases of which I know, and those not clear cut and well studied, are the dissociation of detergent micelles by heat, which seems to involve the breaking in an all or none way of many van der Waals bonds, and the change in starch which takes place in cooking, a change still very obscure in nature.

6) Whenever a very small effect of temperature is observed, as in the case of the salicylate denaturation of hemoglobin, it must mean, so it seems to me, either (a) that under the conditions used few bonds are broken in denaturation or (b) that the energy for the breaking of the bonds comes from the change in equilibria with constituents of the solution, or from other changes in energies of bonds within the protein molecule.

7) The fact the denaturation of hemoglobin by salicylate is not sensitive to temperature does not mean that the bonds broken by heat denaturation are not also broken by salicylate denaturation, with the energy provided by changes in combination with salicylate, or the like.

8) The facts that the equilibrium between native and denatured pepsinogen is very sensitive to the pH and that the equilibrium between native and denatured hemoglobin is very sensitive to the salicylate concentration show that denaturation must involve considerable changes in the protein ionization and salicylation equilibria. And the heat of denaturation as calculated, as I have pointed out, includes the heats of all the changes in ionization or salicylation which take place.

9) In a few cases the heat of denaturation has been measured by direct calorimetry. (References in Buzzell and Sturtevant, 1952.) The results so far have been rather confusing. What is wanted is

direct calorimetry applied to the study of denaturation equilibria, if that is technically feasible.

There is not time to discuss in detail other aspects of the energetics of denaturation, particularly the change in entropy accompanying denaturation and the effect of temperature on the rate of denaturation. I hope, however, that I have made clear that the general facts about the sometimes great, sometimes small effects of temperature on denaturation are reasonable when one considers what is known about the character of the structural changes involved in denaturation. I hope, too, that I have made it clear that progress in the study of the energetics of denaturation is most likely to come from more extensive investigations utilizing suitable materials and conditions. The proteins used ought to be proteins whose denaturation is completely reversible. There ought to be measurements both of the equilibrium and of the rates of denaturation and of its reversal. A variety of proteins, of denaturation conditions and of ways of following denaturation ought to be used and, if possible, the heat of denaturation ought to be measured directly as well as calculated from the effect of temperature on the equilibrium.

I feel sure that such more extensive experimental studies under favourable conditions will be more profitable than some of the existing involved discussions, discussions sometimes based largely on data that are either not complete enough or that are concerned with irreversible reactions, often including reactions beyond denaturation.

Non-denaturable Proteins : We have seen that the denaturation of some proteins under some conditions is reversible whereas the denaturation of some other proteins has so far not been reversed under any conditions. It ought to be pointed out, for the sake of completeness, that there are proteins which do not show the denaturation phenomenon at all. At least, they do not undergo any marked changes of properties at the usual denaturation temperatures. For the most part, these non-denaturable proteins are either proteins of low molecular weights, molecular weights of 15-20,000, or somewhat elongated proteins, such as casein, or decomposed proteins, such as gelatin. They thus either have relatively simple structures or structures somewhat like that of denatured protein to begin with.

There are, however, elongated proteins, like fibrinogen, which can show the phenomenon of denaturation. They may be somewhat « denatured » in their properties. Fibrinogen and some other elongated proteins, as Haurowitz et al. (1945) have pointed out, are, even when native, readily digested by enzymes which attack some native globular proteins slowly or not at all. The SH groups of even native myosin are quite reactive.

There are protein hormones which are not readily inactivated by heat. They have not — I say this from memory — been properly studied from the denaturation point of view. Perhaps they are protected from denaturation by carbohydrate constituents they contain. Perhaps they are biologically active even after being denatured.

CONCLUSIONS

I said, in the beginning of this paper, that denaturation is a complex reaction of complex molecules, and that therefore the experimental and theoretical treatment of denaturation must be complex. Nevertheless, I have tried to leave with you as a main impression that denaturation is not quite so complicated and confusing as it is usually made out to be. I believe that there is such a thing as denaturation, as a rather restricted and characteristic sort of protein reaction which can be separated from other protein reactions, that there is such a thing as denaturation, even if we do not know what denaturation is in structural detail, even if the exact properties of denatured protein depend on the nature of the protein and on the nature of the chemical environment. I believe further that denaturation equilibria can, for many purposes, be treated as simple equilibria, and that the properties of the native and denatured forms of protein in the equilibrium mixture can be studied by a variety of simple and familiar means. The subject of denaturation, the most characteristic reaction of proteins, is, indeed, in much better shape now than it was not too long ago. It is true, of course, that at any stage, the understanding of protein denaturation must be limited by the contemporary knowledge of protein structure. But I have the feeling that we are on the verge of just that increase in knowledge of protein structure which is needed, and also that a

much greater understanding of protein denaturation can be achieved, even without any great increase in knowledge of protein structure, by a systematic expansion of the kind of experiment which has been outlined in this review.

The greater knowledge of protein structure that is needed is greater knowledge of the loose bonds which keep the native protein in place. There are at present no proper ways of identifying or locating or measuring these bonds.

The kind of systematic expansion of known types of experiments which is to be desired is not too hard to visualize. First of all, although much can be learned by studies of irreversible denaturation, there has not been nearly enough concentration on the good cases of reversible denaturation. It would be desirable to have systematic studies of the effects on denaturation equilibria of the various denaturation agents and denaturation inhibitors which have been mentioned. Many different proteins ought to be studied so that the relation of the observed effects to the protein structures can become clearer. The equilibria in the presence of various denaturation agents and inhibitors ought to be studied at various pH's and temperatures. And, where possible, the rates both of denaturation and of its reversal ought to be measured. In other words, each protein denaturation equilibrium system ought to be studied as completely as possible and the protein ought to be varied as much as possible. This as a substitute for the hit and miss observations, which, with few exceptions, are now available. I feel confident that the transition from hit and miss observations to systematic investigations planned with a broad point of view can bring about a qualitative change in our understanding of denaturation. The main purpose of this review has been to sketch out the general lines of the sort of more systematic investigation of denaturation which is possible with our present sort of knowledge.

At other sessions of this Council, there will be illustrations of what can be done by going to more extensive and systematic experimentation. The discovery of the order of amino acids in a long chain was achieved by applying many known procedures in an organized way. The new fractionations with cold alcohol and heavy metal salts based on systematic procedures and using a great deal of our contemporary knowledge of protein chemistry are a great advance on the mere crystallization with cold alcohol or even

on the many old individual cases of deliberate specific fractionation with heavy metal salts. The widely used fractionation by differential denaturation has never been applied in the same systematic way as fractionation by alcohol or heavy metal salts.

So, all told, I feel that attention ought to be concentrated on the relatively simple equilibrium systems and on the testing of relatively simple general ideas. And I believe that great progress is most likely to come from studying the simpler equilibrium systems in a much more systematic and comprehensive way than has hitherto been done. What is now needed above all, and what now can be obtained, is much more organized detail about a number of very different denaturation systems.

No matter how simple the experimental approach may be in principle, one cannot avoid dealing with the facts that there are many different kinds of proteins which can be denatured by many different procedures, that both native and denatured protein can exist in many different forms, and that denaturation involves many structural changes in a large complex molecule. In short, we must follow the policy of General Marshall and be guided both by a simplicity of general approach and a respect for the unavoidable complexities of our problem.

Lever Brothers Company Research Center,
Edgewater, New Jersey, U. S. A.

REFERENCES

Anson, M.L. (1940). *J. Gen. Physiol.*, **23**, 695.
Anson, M.L. (1945). *Advances in Protein Chemistry*, **2**, 361.
Anson, M.L., and Mirsky, A.E. (1930). *J. Gen. Physiol.*, **13**, 469.
Anson, M.L., and Mirsky, A.E. (1932). *J. Gen. Physiol.*, **15**, 341.
Anson, M.L., and Mirsky, A.E. (1933*a*). *J. Gen. Physiol.*, **17**, 393.
Anson, M.L., and Mirsky, A.E. (1933*b*). *J. Gen. Physiol.*, **17**, 399.
Astbury, W.T., Dickinson, S., and Bailey, K. (1935). *Biochem. J.*, **29**, 2351.
Ball, C.D., Hardt, C.R., and Duddles, W.J. (1943). *J. Biol. Chem.*, **151**, 163.
Boyer, P.D., Ballow, G.A., and Luck, J.M. (1946). *J. Biol. Chem.* **162**, 199.
Boyer, P.D., Lum, F.G., Ballow, G.A., Luck, J.M., and Rice, R.G. (1946). *J. Biol. Chem.*, **162**, 181.
Brosteau, J., and Eriksson-Quensel (1935). *Arch. phys. biol.* **12**, 209.
Bull, H.B. (1947). *Advances in Protein Chemistry*, **3**, 95.
Buzzell, A., and Sturtevant, J.M. (1952). *J. Am. Chem. Soc.*, **74**, 1983.
Christensen, L.K. (1952). *Compt. rend. trav. lab. Carlsberg*, **28**, 39.
Duggan, E.L., and Luck, J.M. (1948). *J. Biol. Chem.* **172**, 205.
Eisenberg, M.A., and Schwert, G.W. (1951). *J. Gen. Physiol.*, **34**, 583.
Foster, R.A., Johnson, F.H., and Miller, J.K. (1949). *J. Gen. Physiol.*, **33**, 1.
Frederique, E., and Desreux, V. (1947). *Bull. soc. chim. biol.*, **29**, 105.
Gorini, L., and Felix, F. (1953). *Biochim.* et *Biophys. Acta*, in the press.
Haurowitz, F., Tunca, M., Schwerin, P., and Goksu, V. (1945). *J. Biol. Chem.*, **157**, 621.
Herriott, R.M. (1938). *J. Gen. Physiol.*, **21**, 501.
Herriott, R.M. (1947). *Advances in Protein Chemistry*, **3**, 170.
Hornibrook, J.W. (1952). *Proc. Soc. Exp. Biol. Med.*, **79**, 534.
Kunitz, M. (1948). *J. Gen. Physiol.*, **32**, 241.
Levy, M., and Benaglia, A.E. (1950). *J. Biol. Chem.*, **186**, 829.
Lundgren, H.P. (1949). *Advances in Protein Chemistry*, **5**, 305.
Mirsky, A.E., and Pauling, L. (1936). *Proc. Nat. Acad. Sci.*, **22**, 439.
Neurath, H., Greenstein, J.P., Putnam, F.W., and Erickson (1944). *Chem. Rev.*, **34**, 157.
Porter, R.R. (1948). *Biochim. et Biophys. Acta*, **2**, 105.
Porter, R.R. (1950). *Biochem. J.*, **46**, 304.
Porter, R.R., and Sanger, F. (1948). *Biochem. J.*, **42**, 287.
Putnam, F.W. (1948). *Advances in Protein Chemistry*, **4**, 80.
Roche, J., and Chouaiech, M. (1940). *Bull. soc. chim. biol.* **22**, 263.
Rothen, A. (1942). *Ann. N.Y. Acad. Sci.*, **43**, 229.
Steinhardt, J., and Zaiser, E.M. (1951). *J. Biol. Chem.*, **190**, 197.

Discussion du rapport de M. Anson

M. Pauling. — I wish that I could make some definite statements about the structural change in proteins that accompanies denaturation. I feel that a protein molecule may be subjected to a very small structural change, which causes significant change in some sensitive properties and no noticeable change in other properties, or to a larger structural change, causing a larger change in properties, or to a very large structural change, with correspondingly great change in properties. At the present time we have little information about the magnitude of the structural change that occurs when proteins are denatured in various ways.

An example is provided by hemoglobin, which can be denatured and then have its denaturation reversed. The resulting renatured hemoglobin is closely similar to native hemoglobin in many properties. Its absorption spectrum is not appreciably different from that of native hemoglobin. It has the power of combining reversibly with oxygen; however, the equilibrium curve showing the amount of combination as a function of the partial pressure of oxygen does not have the sigmoid character shown by native hemoglobin. We conclude that on renaturation the structure of the hemoglobin molecule has returned nearly to the original structure, of native hemoglobin, but that the part of the structure involved in the interactions between the heme groups have not been completely restored.

M. Anson. — Pauling's observation that the oxygen equilibrium curve of hemoglobin prepared by the reversal of denaturation is somewhat different from the curve of hemoglobin which has never been denatured is one more illustration of the fact that one can sometimes obtain reversal of the main changes involved in denaturation, and still not get perfect reversal of all the changes. This sort of result is common in cases in which not all the protein becomes native again when denaturation is reversed. On the other hand, in the cases in which all the protein was converted back to the native form, no differences could be detected between the renatured and original native forms of the protein. It would accordingly be interesting to measure the oxygen equilibrium curve of renatured hemoglobin prepared by a reversal procedure with

a 100 % yield of native hemoglobin. For instance, one might compare the oxygen equilibrium curve of the two hemoglobins prepared from native methemoglobin and from methemoglobin prepared by the reversal of denaturation by salicylate.

M. Theorell. — I think there is good experimental reason to believe that denaturation of proteins is not an « all or none » reaction, but is initiated by small changes of configuration which are at least sometimes reversible. This is particularly easy to demonstrate in the case of flavoproteins, as I am going to mention in my lecture.

I don't agree entirely with Dr. Anson's somewhat too general statement that enzymes become less and less stable on purification. According to our experience the opposite is very often the case.

A more general phenomenon, however, seems to be the stabilization of apoenzymes by their prosthetic groups, and of enzymes by their substrates. This is understandable if we remember that both prosthetic groups and substrates are attached to the protein by more than one bond — at least for the prosthetic groups this has been experimentally proved. Obviously this can give an extra stabilization to the protein molecule.

M. Anson. — The failure of protein to combine with flavine may or may not mean that the protein is a « little bit denatured ». One would want to know whether the non-combining protein has other characteristics of denatured protein.

M. Hermans. — The report mentions organic, water soluble, substances (inhibitors), which protect the protein against denaturation. The question arises : what exactly is the protein molecule protected against ? Could it be that it is simply protected against interaction with other protein molecules ? From this point of view, it may be of interest to study the effect of protein *concentration* on the rate of denaturation.

M. Anson. — So far as I remember, the sort of experiment desired by Dr. Hermans has not been carried out.

M. Havinga. — Denaturation has been called an « all or none » process and it has been indicated that there should be only one denatured state, related to a native state. We may represent a state of the protein by a valley in the n-dimensional free-energy landscape; in each valley there are millions of slightly different

configurations. In case there is thermal equilibrium between a native state and a denatured state, there should be in the equilibrium mixture also a not neglegible number of molecules having conformations corresponding to situations in between the regime of nativity and denaturation. Moreover, it seems not improbable that generally there will be several valleys, that one might look upon as each representing a different denatured state.

M. Anson. — The present limited evidence indicates that in any single equilibrium mixture of the native and denatured forms of a protein, practically all the molecules are either in one native form (with slight thermal variations) characteristic of the chemical environment or in one denatured form characteristic of the chemical environment. A priori, one might conceive of there being different forms of denatured protein in the equilibrium mixtures, representing protein molecules in different « stages of denaturation » with different properties. But so far there is no evidence of different forms of denatured protein in a given equilibrium mixture. Of course, if the chemical environment is changed, then both the native and the denatured forms of the protein are changed. The « all or none » hypothesis does not state that denatured protein can exist in only one form, despite variations in the chemical environment, but merely that, so far as is known, the denatured protein in any given equilibrium mixture is practically all in only one form. Since the phrase « all or none » is so often misunderstood, perhaps it were best abandoned.

M. Linderström-Lang : Agreed with Dr. Anson that detailed and painstaking experiments were needed to solve the question of denaturation and regretted that he (K. L. L.), by asking questions that were not ripe for discussion, had contributed to bringing the discussion on a philosophical level.

Pointed out to the importance of distinguishing between denatured molecules in solution and precipitates of denatured molecules, and emphasized the necessity of finding methods for characterizing the former. Mentioned volume and optical rotation (Kauzmann) as possible characterization, but admitted the difficulty of handling these theoretically.

Pointed out that the « all or none » character of denaturation, so strongly emphasized by Dr. Anson, indicated that the initial phase of denaturation involved the passing of a high energy

barrier after which the subsequent steps involved occur rapidly. Thought it likely that in certain systems where the energy barrier is lowered (e. g. by addition of urea, etc.) intermediates may be found (partly denatured molecules).

Asked Prof. Pauling whether precipitated denatured proteins might not be considered having a « pleated sheet » structure.

M. Anson. — Linderström-Lang has called attention to denaturation of protein in the solid state. It is indeed unfortunate that denaturation of protein in the solid state has received so little study, especially since so much of the protein in the living cell, even of the salt-soluble protein, is in the solid state.

If the great sensitivity of denaturation to temperature is due to the fact that denaturation involves many changes in a large molecule in an all or none way, then the bigger the molecule and the more numerous the all or none changes, the closer denaturation ought to come to a sort of melting at a sharp temperature. For melting is an all or none reaction involving a very large number of changes in an indefinitely large molecule. I once measured the effect of temperature on the role of denaturation of protein in the form of an amorphous precipitate, in the hope that the unit of reaction would be the large particle of precipitate and that the denaturation point would then be as sharp as the melting point. Actually, the effect of temperature was about the same as the effect of temperature on denaturation of protein in solution; the unit of denaturation remained the usual molecule, even when the protein was in the solid state. It might be worth while repeating the experiment, measuring the effect of temperature on the rate of denaturation of protein in crystalline form.

M. Pauling. — There is evidence from X-rays, by Riley and Arndt, that bovine serum albumin denatured by heat and ovalbumin denatured by surface action retain the α-helix as their principal structural feature. Astbury and his collaborators have shown that globular proteins can be denatured to form films which, when stretched, give the X-ray pattern of β keratin.

Probably mild denaturation and even severe denaturation usually leave the α helixes intact, and involves only the unfolding of larger structures formed of segments of α helixes. Extreme denaturation may involve the uncoiling of α helixes.

An American investigator has recently reported that physicochemical study of denatured ovalbumin in solution has led to the

conclusion that the molecule is a rigid rod 600 A long, with 1,5 A per residue, which indicates that it is a single α helix.

M. Putzeys. — La structure des protéines dénaturées dépend de la façon dont la protéine, dénaturée au préalable par un agent convenable — acides, urée..., — est traitée par la suite. Un précipité de protéine dénaturée peut être remis en solution par acidification et dialyse pour éliminer les électrolytes comme les recherches de Lontie (*Bull. Soc. Chim. Biol.*, 1946, **28**, 509) l'ont montré. De petites quantités de sels sont suffisantes pour précipiter la protéine qui passe de nouveau en solution lorsque les sels sont éliminés par dialyse. Ces solutions sont poly-dispersées et présentent une très forte aggrégation. Il est bien possible que des précipités « solubles » de ce genre aient la structure du type α-kératine. Si cependant une telle protéine dénaturée est formée en film et que l'on soumet ces films à la traction, on obtient une structure β-keratine typique, comme l'ont montré les recherches d'Astbury, Dickinson et Bailey (*Biochem. J.*, 1935, **29**, 2351). A ce sujet, il ne semble pas y avoir de doute. Lorsque cependant la protéine dénaturée est mal orientée mécaniquement, comme lorsqu'on essaie d'en faire un fil, le diagramme n'est pas d'un type net, un point sur lequel Astbury et col. ont insisté.

On peut d'ailleurs répéter ces expériences en mettant la fibroïne de la soie en solution par la cuprié thylène diamine et obtenir une fibroïne soluble dans l'eau pure après acidification et dialyse. Lorsqu'on évapore la fibroïne soluble sur une surface de mercure, on obtient un film qui passe en solution au contact de l'eau, mais si on l'étire au préalable, le film devient résistant, biréfringent et insoluble dans l'eau (D. Coleman-F. O. Howitt, Society Dyers and Colourists, Fibrons Proteins, 1946).

On peut conclure de ceci qu'une protéine dénaturée complètement peut exister sous une forme soluble et une forme insoluble, et que cette dernière n'est en général obtenue que par une intervention mécanique. C'est sans doute seulement dans ces films bien orientés que la structure du type β-kératine apparaît nettement.

M. Léonis. — Le professeur Linderström-Lang vient d'intervenir en faveur d'une des thèses du Dr. Anson, en insistant sur la nécessité de réunir un matériel expérimental plus abondant pour pouvoir interpréter le phénomène de dénaturation dans son ensemble. Des expériences préliminaires sur la dénaturation de l'édestine par le chlorhydrate de guanidine nous ont indiqué, en effet, qu'un

grand nombre de phénomènes se manifestent successivement ou concouramment au cours de ce processus. Divers points singuliers apparaissent par exemple dans les courbes de viscosité, de dilatométrie, de turbidité ou d'activité optique; des groupes cationiques deviennent très rapidement accessibles, avant la libération des groupes -SH qui sont suivis à leur tour par des noyaux aromatiques (spectre U. V.). Ces expériences donnent une idée de la complexité des réactions en jeu, et montrent que les phénomènes liés à la dénaturation lente peuvent se répartir dans le temps. Il nous semble donc fructueux d'envisager l'étude de la dénaturation par une variété très grande de méthodes, mettant simultanément en relief les changements physico-chimiques de la protéine en solution et les transformations d'ordre chimique (groupes cachés) qu'elle subit.

M. Desreux. — Barbu et Joly ont récemment présenté de nombreux arguments sur la base desquel il semble que, dans certains cas de faible dénaturation, l'aggrégation de molécules de protéine entre elles, celles-ci conservant essentiellement leur caractère globulaire, est le phénomène le plus important. (Discussions of the Faraday Society : the physical chemistry of proteins, n° 13, 77, 1953.)

M. Anson. — Barbu and Joly have shown that what might be mistakenly interpreted as the formation of very elongated molecules by extreme opening up of the molecules due to denaturation is, in some cases, end to end linear aggregation of fairly globular molecules, analogous to the formation of insulin fibrils studied by Waugh. Although denaturation can sometimes involve very great opening up of protein molecules, it is not, however, claimed that denaturation *always* involves extreme or complete opening up. This was pointed out by Astbury when the paper of Barbu and Joly quoted by Desreux was presented at the Faraday Society Discussion. The experimental problem is to investigate, by the usual methods of studying the size and shape of protein molecules in solution, exactly how much the protein molecules are opened up in different cases of denaturation (preferably reversible denaturation) and exactly how much the degree of opening up depends of the nature of the solution. But precisely in such an investigation, it is necessary to avoid aggregation rather than to study aggregation. The difficulties of interpretation caused by not eliminating aggregation are illustrated by the results of Barbu and Joly.

The necessity of avoiding aggregation was also illustrated by the study of the effect of denaturation on viscosity [Anson and Mirsky : *J. Gen. Physiol.*, **15**, 341 (1931)]. The moderate increase in viscosity accompanying denaturation in urea solution was due to changes in the individual molecules. Aggregation was ruled out by the results of osmotic pressure measurements. But when denaturation was brought about by heating an aqueous solution of protein at a pH somewhat away from the isoelectric point, then great increase in viscosity, even gel formation, was observed. This great change in viscosity was due to aggregation and was sensitive to the pH, to salts and to the concentration of protein. Many similar cases of great increases of viscosity and of gel formation, due to aggregation and sensitivity to the same factors, have been described by Barbu and Machebœuf [*Annales de l'Institut Pasteur*, **75**, 429 (1948)]. It is certainly useful to have the French workers emphasize the clear distinction between the physical results of protein-protein interactions following denaturation and the physical results of the changes within individual protein molecules characteristic of denaturation itself.

M. Fredericq. — La théorie classique, selon laquelle les groupes chimiques des protéines natives ne pourraient réagir par suite d'empêchement stérique, est actuellement controuvée par de nombreux faits expérimentaux. En particulier, dans le cas des groupes -SH, il n'y a aucun rapport entre les dimensions des molécules susceptibles de réagir avec eux et la vitesse de réaction : tandis que la chloropicrine, la bromopicrine, la chloracétophénone, etc., bloquent rapidement les -SH de l'ovalbumine native, des molécules beaucoup plus petites telles que chloracétate, ions hydroxyles, cétène, oxygène, n'ont aucune action mesurable.

Suivant une autre théorie, les groupes SH seraient impliqués dans d'autres groupements chimiques dans les protéines natives. Des structures cycliques, notamment, tels les cycles thiazoline, ou thiazolidine, permettent d'expliquer certains faits expérimentaux. Je voudrais demander au Dr. Anson son avis quant à la validité de ces théories ?

Ne pensez-vous pas que le déroulement de protéines natives sous l'action de certains agents dénaturants a été parfois surestimé, à la suite de mesures de viscosité en particulier ? C'est ainsi que Rothen a trouvé pour l'ovalbumine dans l'urée un coefficient de friction de 4; ceci correspondrait à une chaîne polypeptidique

complètement déroulée avec une longueur de plusieurs centaines d'angströms. Nous n'avons pu cependant observer aucune biréfringence de flux dans ces conditions.

Il semble que l'on n'ait pas suffisamment tenu compte de la solvatation des protéines par des agents dénaturants, tels que l'urée ou la guanidine, solvatation qui est certainement très considérable.

M. Anson. — 1) So far as I can see, not enough is known about the structures of the native and denatured forms of protein to provide the basis for an adequate theory of why denaturation activates so many protein groups. The special theories which have been proposed are all open to criticism, insofar as they are at all definite. Fredericq has pointed out the difficulties of the spatial theory. The difficulty of the second theory he mentioned, that of native groups being bound in cyclic structures, is that so many different groups are activated by denaturation. One would not like to be forced to picture a native protein molecule in which all the groups activated by denaturation are bound in cyclic structures.

2) From the mere increase in viscosity as a result of denaturation in urea solution, one cannot say whether there has been a change in shape of the molecule, or increased hydration, or both. Rothen, however, calculated an increase in the Svedberg « frictional ratio » from measurements of sedimentation and diffusion. Although the exact change in shape cannot be determined from the « frictional ratio », the hydration effect is eliminated in the Svedberg calculation, since hydration affects sedimentation and diffusion equally.

M. Pauling. — a) One would conclude from the small electronegativity of sulfur that the SH group would never form strong hydrogen bonds. This conclusion is supported by many experimental facts.

M. Petersen. — As regards Dr. Anson's remarks about the use of viscosity measurements for the determination of the shape or the axial ratio of the eggalbumin after denaturation, I would like to remark that what one observes is an increase in viscosity. From the viscosity experiments alone, one cannot say anything about the change that has taken place in the eggalbumin molecule on denaturation. If one assumes, however, that the shape of the molecule is ellipsoidal, it is possible to calculate an axial ratio that may fit the experimental viscosity values. Such a calculation is, however, purely speculative.

DEGRADATION OF PROTEINS BY ENZYMES

By K. LINDERSTRÖM-LANG

1. INTRODUCTION

It is with considerable hesitation that I attempt a discussion of the enzymatic breakdown of proteins. There is a flagrant disproportion between the amount of data collected and the information obtained from them concerning the mechanism of this complex process or series of processes. Many workers will probably maintain as their opinion that the mechanism of the enzymatic hydrolysis of simpler substrates ought to be cleared up before this intricate problem is touched upon, but it seems to me that its close connection with important biological questions justifies a discussion of its chemical aspects though the treatment cannot become very deep-going.

2. GENERAL AND SPECIAL METHODS

2.1 General chemical methods.

Since hydrolysis of peptide bonds is the best established chemical reaction involved in the proteolytic degradation of proteins several methods have been developed for the determination of the carboxylic groups or the amino groups set free in this hydrolysis :

$$-R-\underset{\underset{O}{\|}}{C}-\underset{\underset{H}{|}}{N}-R_1 \xrightarrow{H_2O} RCOO^- + NH_3^+R_1$$

In order that these methods be general they should measure the groups in question irrespective of the constituents R and R_1. This is approximately true in several cases. I shall not use much space in discussing the various methods that satisfy such conditions. In

a previous paper (1) the author has given a table showing advantages and drawbacks of 4 of the chemical methods that have been used for following the proteolytic degradation: The formol titration, van Slyke's gasometric method, the Willstätter titration in alcohol and the acetone titration. In certain cases the latter method is superior to the other ones with respect to accuracy as shown by Jacobsen (2) (3). More recently developed general methods, like the titration in glacial acetic acid (4) (5) (6) or methods based on reactions with special reagents (see below) do not lend themselves so easily to determinations of the over-all yield of amino or carboxylic groups.

2.2 General physical methods.

No exhaustive discussion of the various physical properties studied in relation to proteolytic breakdown shall be attempted here. The physical change most commonly used for characterization of digests is that of solubility under fixed conditions. The methods of this type are numerous. In some cases the enzyme substrate system is heterogeneous and the degree of proteolysis is measured by the solubilization of the substrate (certain commercial and clinical methods), in others distinction is made between digested and undigested protein by precipitation of the latter with trichloracetic acid, sulfosalicylic acid and the like. Notwithstanding the fact that such a distinction is unsharp if appreciable amounts of high-molecular degradation products are present, the latter methods have been extremely useful both in routine work and, used with criticism, also in kinetic studies of protein breakdown.

Another physical change which bears a close relationship to the desintegration of the large protein molecules is the shift in molecular weight pattern of the digest during degradation.

Investigations have been carried out by Tiselius and Ericsson-Quensel (7), by Annetts (8), by the Williams school (9) (10) (11) (12) (13) and by others, and highly important results have been obtained regarding the distribution of weight classes among the degradation products. The study of Williams and Lundgren (9) (10) on thyreoglobulin has thrown considerable light on the mechanism of the initial attack of papain upon this protein. The analysis of sedimentation diagrams is however complicated and it is only quite recently that an adequate theory has been developed [Gosting (13), Williams et al. (12)].

Changes of viscosity have been widely used as a measure of changes in size and shape during degradation. Due to the simplicity of the most commonly used instruments and the high complexity of the theoretical basis the quantity of experiments carried out stands in no reasonable proportion to the understanding acquired.

The volume change accompanying proteolytic breakdown has been studied in a few cases. Owing to the relatively simple predictions which can be made regarding the volume change in plain peptide bond hydrolysis it has been possible to isolate, if not to identify, considerable volume changes connected with the collapse of the secondary structure of protein molecules in the initial stages of their enzymic degradation (14) (15) (16).

Finally mention should be made of the change in optical rotation which has been so widely used as a measure of enzymic hydrolysis. In the case of proteins recent investigations (17) (18) (19) have pointed to the value of this physical property in determining the physical state of protein molecules and especially in discriminating between native and denatured molecules.

2.3 Special methods.

Methods for the determination of the concentration of unattacked substrate protein in an enzymic digest belong to this group. They must measure some property or other of the protein which it does not share with any of its degradation products. Such criteria as power of crystallization, solubility, immunological reactions, etc., are in general unsatisfactory, in part because they are difficult to handle experimentally, in part owing to lack of sufficient specificity. The only almost certain criterion of nativity is the biological activity of enzymes and here the choice of substrate is naturally limited, especially since proteolytic enzymes must generally be excluded.

Methods for the isolation and identification of degradation products are next to be considered. In recent years a large number of highly important methods of this kind have been developed. Some of them may be classified as general chemical methods, some as general physical methods. However, they all attempt a detailed analysis of the systems they are applied to, and they often have to be remodeled to fit special cases, which is the reason why they are mentioned in the present section.

Since a detailed review of this side of the question will be given by others, the following indications may suffice.

Methods for isolation:

Electrophoresis on columns or on paper [20] [21], convection electrophoresis [22] [23].

Frontal analysis on active carbon [24], chromatography on columns (starch, ion exchange resins [25] [26] [27] or filter paper [28], counter current extraction [29] [30], or partition chromatography [31]. In addition classical or recently improved precipitation methods [32] [33].

Methods for analysis:

Endgroup determination [NH_2-groups: Dinitrofluorbenzene [34] [35], CS_2 [36] or phenylthioisocyanate [37]; COOH-groups: aluminium lithium hydride [38], arylurea [39] and others)].

Determination of sequence of amino acids in peptides [CS_2 [40] [41], phenylthioisocyanate [37] [42] [43] [44] [45]] in addition to chromatographic estimations of amino acids [25] [26].

This is a very brief and by no means complete survey of the different tools to which the enzyme chemist has access when tackling the problem of protein degradation. So far there are only few examples of detailed analysis of the course of such reactions. It must be expected, however, that in the next future a large amount of work will be carried out on this problem.

3. GENERAL REMARKS ABOUT THE ENZYMES AND SUBSTRATES

In the following we shall take it as a certainty that proteins are built up by peptide chains. We shall also assume that in native globular proteins, with which we shall deal exclusively here, the properties of the free flexible peptide chains are suppressed and that the structural unit which we call a protein molecule is stabilized by additional bonds (primarily hydrogen bonds) and by physical forces of the kind that act in crystals of amino acids and peptides. Furthermore we shall regard denaturation as a complete or partial unfolding of the molecule whereby properties of the free peptide chains reappear. The fact that denatured proteins are attacked by enzymes at a higher rate than native ones is explained

by assuming that part of the peptide bonds of the chains are masked in the native molecule, e.g. by hydrogen bonding (46) (47) (48), and are set free in the process of unfolding in the same way as SH-groups, tyrosine and lysine groups are exposed.

With regard to the proteolytic enzymes attacking proteins, the proteinases, we shall assume that they catalyze the hydrolysis of peptide bonds primarily. The existence of ester bonds in the protein molecule cannot be excluded [Neurath and coworkers (49)] but their number must be very small. The points of attack of the proteinases are determined by the amino acids directly adjacent to the peptide bonds in question [the wellknown side chain specificity (50) (51) (52)] as well as by the local and general structure of the substrate molecule.

The question whether proteinases are able to act denaturing upon their substrates without opening peptide bonds is still unsettled. It is however a comparatively well established fact that denaturation occurs at a very early stage of the process of degradation and that therefore the opening of only a few bonds is required for a complete upsetting of the native protein molecule.

As regards details of the picture outlined above the author refers to a previous article and the literature quoted here (16).

4. SOME KINETIC CONSIDERATIONS

The kinetics of the proteolytic breakdown of a protein molecule is naturally extremely complex and not open to detailed treatment. A brief discussion of a few theoretical cases may be of value for the following. The author may refer to previous treatments of similar problems especially to those by Kuhn (53), Montroll and Simha (54), Warner (55), and Sanger (56).

4.1 The protein molecule is considered as an aggregate of substrates each containing one peptide bond susceptible to the enzyme. The reaction between a given bond (substrate) and the enzyme is assumed to be independent of the fate of the rest of the aggregate.

In our consideration we shall adopt the Henri-Michaelis equation as a first approximation and put

$$-\frac{ds}{dt} = k \cdot se \tag{1}$$

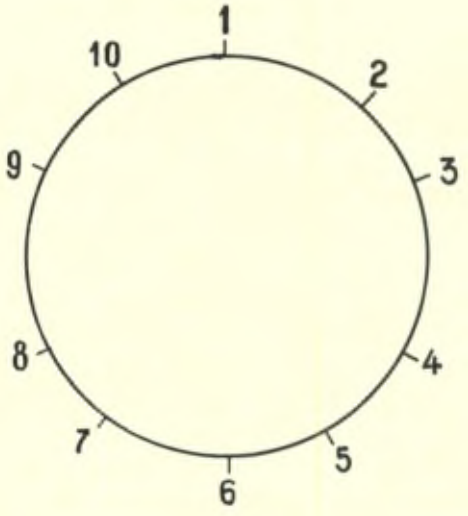
$$\frac{s \cdot e}{se} = k_m \qquad e_o = e + se \tag{2}$$

$$-\frac{ds}{dt} = \frac{k}{k_m} \cdot \frac{e_o \cdot s}{1 + \frac{s}{k_m}} = \varkappa \cdot s \cdot e \tag{3}$$

where s, se, e and e_o are the concentrations of substrate, enzyme-substrate complex, free enzyme, and total enzyme, respectively; k is the velocity constant, and k_m the Michaelis constant. The quantity $\varkappa$ is equal to k/k_m.

4.11 Proteolysis of a peptide ring.

As an example we shall consider a peptide ring (S) with n susceptible bonds. Denoting by p_i and $\varkappa_i$ the concentration of the bond i and its constant $\varkappa$ respectively we obtain



$$-\frac{ds}{dt} = s \cdot e \sum_n \varkappa_i \tag{4}$$

where e is the concentration of free enzyme.

Equation (4) cannot be solved for s and t in a simple way (since e

is generally a function of time) unless the enzyme is bound to the same degree by all components of the system, in which case

$$ln \frac{s_o}{s} = t \cdot e \sum_n \varkappa_i \tag{5}$$

However, due to our preliminary assumption we may also write

$$- \frac{dp_i}{dt} = \varkappa_i \cdot p_i \cdot e \tag{6}$$

valid for each individual bond. Consequently we obtain from (4) and (6)

$$\frac{dp_i}{ds} = \frac{\varkappa_i}{\Sigma \varkappa_i} \cdot \frac{p_i}{s} \tag{7}$$

or putting

$$\varkappa_i / \Sigma \varkappa_i = \alpha_i; \quad s/s_o = \mathrm{R} \tag{8}$$

$$\frac{p_i}{s_o} = \mathrm{R}^{\alpha_i} \tag{9}$$

(s_o initial concentration of substrate).

On the basis of this very simple equation a number of different quantities can be calculated :

Number of bonds of type i split

$$b_i = s_o (1 - \mathrm{R}^{\alpha_i}). \tag{10}$$

Total number of bonds split

$$b = s_o (n - \sum_n \mathrm{R}^{\alpha_i}). \tag{11}$$

If we define a peptide « unit » as the piece of peptide chain between two susceptible bonds — irrespective of the fact that the lengths of the units may vary from place to place in the ring, — we find :

Concentration of an open peptide with n units formed by splitting bond l

$$s_{nl} = s_o (1 - \mathrm{R}^{\alpha_l}) \, \mathrm{R}^{\sum_{n-1} \alpha_l} \tag{12}$$

(summation over all bonds except l).

Total concentration of open peptides with n units

$$s_n = s_o \sum_n (1 - R^{\alpha_l}) R^{\Sigma\alpha_i} \tag{13}$$

$$[l = 1, 2, 3 \ldots n;\ i = 1, 2 \ldots (l-1), (l+1) \ldots n]$$

Concentration of a peptide with q units formed by splitting bonds l and m and none of the $q - 1$ intermediate bonds

$$s_{qlm} = s_o (1 - R^{\alpha_l})(1 - R^{\alpha_m}) R^{\Sigma\alpha_i} \tag{14}$$

(summation over $q - 1$ intermediate bonds)

Total concentration of peptides with q units (see the figure)

$$s_q = s_o \sum_n (1 - R^{\alpha_l})(1 - R^{\alpha_m}) R^{\Sigma\alpha_i} \tag{15}$$

$$[l = 1, 2, 3 \ldots n;\ m = q + l;$$

$$i = (l+1)(l+2)\ldots(m-2)(m-1)].$$

Average chain length of degraded molecules measured in units

$$A = \frac{n \cdot s_n + \Sigma q s_q}{s_n + \Sigma s_q} = \frac{(s_o - s)n}{b} = \frac{1 - R}{1 - \frac{1}{n}\sum_n R^{\alpha_i}} \tag{16}$$

Naturally equations of this general type are of no great value unless the constants α_i are known. As a numerical example we may consider the case that all α_i are identical or according to (8)

$$\alpha_i = \frac{1}{n} \tag{17}$$

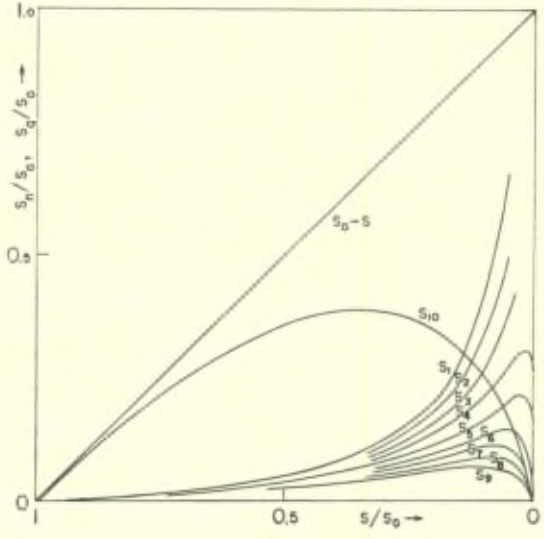


Fig. 1

where we have

$$\Sigma p_i = p = s_o \cdot n \cdot R^{\frac{1}{n}} \tag{18}$$

$$b = s_o \cdot n\,(1 - R^{\frac{1}{n}})$$

$$s_n = s_o \cdot n\,(1 - R^{\frac{1}{n}})\,R^{\frac{n-1}{n}}$$

$$s_q = s_o \cdot n\,(1 - R^{\frac{1}{n}})^2 R^{\frac{q-1}{n}}$$

Figure 1 shows the size distribution as a function of

$$R\left(=\frac{s}{s_o}\right); n = 10$$

It is apparent that a great number of intermediary products are present at all stages. Especially interesting is the distribution of s_{10}, the concentration of the open peptides with all n units. Due to the fact that they are formed by the splitting of only one bond these peptides dominate at all stages of the proteolysis except towards the very end of the process.

The enzyme therefore exhibits what has been termed a « zipper » action (16) upon the peptide ring, ripping the molecules open one after the other by splitting one bond in each practically without touching the other.

A more detailed picture may be obtained experimentally if the substrate molecule has some characteristic property which it does not share with any of the degradation products, so that $R = s/s_o$ may be determined. If furthermore any bond i may be characterized e.g. by the N-terminal amino acid of the unit whose amino group is involved in the formation of the bond we find

$$s_{term} = n_i s_o\,(1 - R^{\alpha_i})$$

where s_{term} is the number of N-terminal groups appearing of type i and n_i the number of bonds i in the substrate molecule. From this simple expression n_i and α_i may be determined provided, naturally, that our additional assumptions are valid.

4.12 Proteolysis of an open peptide chain.

We shall consider an open peptide with n bonds and $n + 1$ units. Using the same terminology as before we find

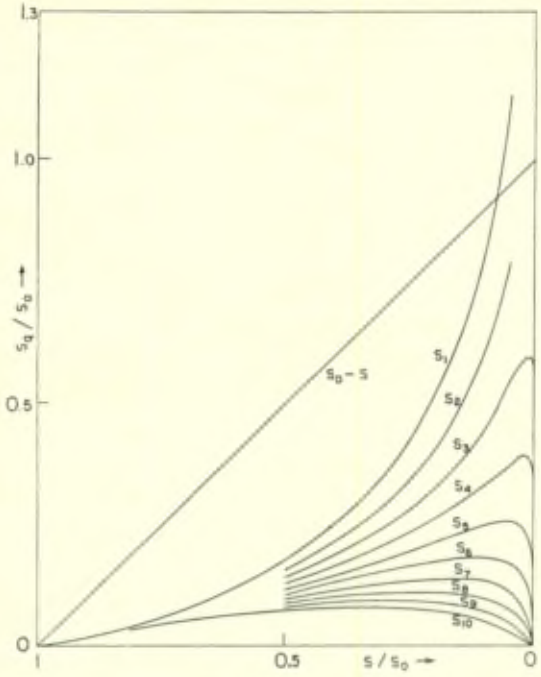


Fig. 2

$$b_i = s_o (1 - R^{\alpha_i}) \tag{10}$$

$$b = s_o (n - \Sigma R^{\alpha_i}) \tag{11}$$

where R and α_i are given by (8). Furthermore we find :

Concentration of a peptide with q units, including terminal units and formed by splitting the bonds q or $n - q + 1$

$$s_{qt} = s_o (1 - R^{\alpha_q}) R^{\Sigma\alpha_i} \text{ or } s_{qt} = s_o (1 - R^{\alpha_{n-q+1}}) R^{\Sigma\alpha_i} \tag{19}$$

$(i = 1, 2, 3 \ldots q - 1$ or $i = n, n - 1, n - 2 \ldots n - q + 2)$

Concentration of a peptide with q units not including terminal units

$$s_{qlm} = s_o (1 - R^{\alpha_l}) (1 - R^{\alpha_m}) R^{\Sigma\alpha_i} \tag{20}$$

(summation over $q - 1$ intermediate bonds).

Total concentration of peptides with q units

$$s_q = s_o (1 - R^{\alpha_q}) R^{\Sigma\alpha_i} + s_o (l - R^{\alpha_{n-q+1}}) R^{\Sigma\alpha_i} + s_o \sum_{n-q} (1 - R^{\alpha_l}) (1 - R^{\alpha_m}) R^{\Sigma\alpha_i}. \tag{21}$$

If again we put $\alpha_i = 1/n$

we get

$$b = s_o \cdot n\,(1 - R^{\frac{1}{n}})$$

$$s_q = s_o\, R^{\frac{q-1}{n}}\,(1 - R^{\frac{1}{n}})\,[(n - q)\,(1 - R^{\frac{1}{n}}) + 2] \qquad (22)$$

The results are seen in figure 2. They show a more even distribution of size classes than in the case of the peptide ring.

4.2 The peptide bonds of the protein molecule in its native form are resistent to the enzyme. However, in an irreversible reaction the enzyme is able to transform the native protein into one of the above described forms in which all susceptible bonds are accessible to the enzyme. No peptide bonds are split in the initial reaction : $N \xrightarrow{E} S$.

The kinetic equations are in this case

$$-\frac{dN}{dt} = \varkappa_N \cdot N \cdot e \qquad (23)$$

$$-\frac{ds}{dt} = s \cdot e\,\Sigma\varkappa_i - \varkappa_N \cdot N \cdot e \qquad (24)$$

$$-\frac{dp_i}{dt} = \varkappa_i \cdot p_i \cdot e - \varkappa_N \cdot N \cdot e \qquad (25)$$

from which we get

$$s = \frac{N_o}{1 - \beta_N}\,(Q^{\beta_N} - Q) \qquad (26)$$

$$p = \frac{N_o}{1 - \beta_i}\,(Q^{\beta_i} - Q) \qquad (27)$$

where

$$Q = N/N_o;\ \beta_N = \Sigma\varkappa_i/\varkappa_N;\ \beta_i = \varkappa_i/\varkappa_N \qquad (28)$$

If β_N and β_i are very small, which means that the initial reaction is fast in relation to the hydrolysis of S, we find for all values of Q except very low ones

$$s = p_i = N_o\,(1 - Q) = N_o - N$$

When practically all N has disappeared the situation is as described in section 4.1.

If β_N and β_i are very large, corresponding to a slow initial reaction, we find

$$s = \frac{QN_o}{\beta_N} \sim 0$$

$$p_i = \frac{QN_o}{\beta_i} \sim 0$$

which means that at every stage of the proteolytic process the system contains practically only N and one-unit degradation products. The process is therefore a one-by-one reaction [see (16)].

For finite β_N and β_i expressions like (10)-(16) or (19)-(21) may be derived. If S is assumed to be a peptide ring, we obtain simply

$$b_i = N_o - N - p_i = N_o\,(1 - Q - \frac{Q^{\beta_i} - Q}{1 - \beta_i}) \tag{29}$$

$$b = \sum_n b_i \tag{30}$$

The determination of the quantity and type of peptides formed is however somewhat more involved. Since at a given time the system contains an infinite number of age classes in S (including degradation products) the members of which have been under the influence of the enzyme for different lengths of time the quantity $p_i/(N_o - N)$ is not a correct expression of the probability that p_i is unhydrolyzed in any member of the total S-population. If we use the process $N \longrightarrow S$ as a « clock » the quantity

$$-\frac{dp_i}{dN_\tau} = \left(\frac{N}{N_\tau}\right)^{\beta_i} \tag{31}$$

is the probabilitiy that p_i is intact in the age class N_τ at the « time » N. From (31) we find by integrating over all age classes present (N constant)

$$p_i = -\int_{N_o}^{N} \left(\frac{N}{N_\tau}\right)^{\beta_i} dN_\tau = N_o\,\frac{Q^{\beta_i} - Q}{1 - \beta_i} \tag{32}$$

which is identical with (27). Similarly we find

$$-\frac{ds_{nl}}{dN_\tau} = \left[1 - \left(\frac{N}{N_\tau}\right)^{\beta_l}\right]\left(\frac{N}{N_\tau}\right)^{\sum_{n-1}\beta_i} \tag{33}$$

and

$$-\frac{ds_{qlm}}{dN_\tau} = \left[1 - \left(\frac{N}{N_\tau}\right)^{\beta_l}\right]\left[1 - \left(\frac{N}{N_\tau}\right)^{\beta_m}\right]\left(\frac{N}{N_\tau}\right)^{\sum_{q-1}\beta_l} \tag{34}$$

from which s_{nl}, $s_{q/m}$, s_n and s_q can easily be found. For $\beta_i = \beta_j = \cdots\cdots = 2$, $\beta_N = 20$, $n = 10$ the results shown in figure 3 are found. The expressions used are

$$s = N_o \frac{Q - Q^{20}}{19}; \quad s_n = 10\, N_o \left\{ \frac{Q - Q^{18}}{17} - \frac{Q - Q^{20}}{19} \right\}$$

$$s_q = 10\, N_o \left\{ \frac{Q^{2q-2} - Q}{1 - 2q + 2} - 2\, \frac{Q^{2q} - Q}{1 - 2q} + \frac{Q^{2q+2} - Q}{1 - 2q - 2} \right\}$$

The abundance of small degradation products at early stages of the breakdown of N is evident.

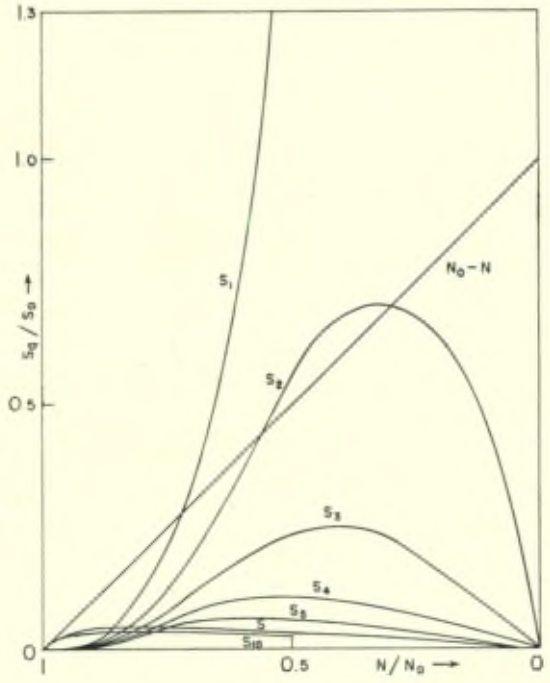


Fig. 3

4.3 The breakdown of the protein molecule is a series of consecutive reactions, the availability of a bond to the enzyme being conditioned by the hydrolysis of the preceding bond.

As an example we may consider carboxypeptidase action upon an open peptide. The units are here amino acids.

bonds $\quad n-5 \quad n-4 \quad n-3 \quad n-2 \quad n-1 \quad n$

carboxyl end

units $n-5 \quad n-4 \quad n-3 \quad n-2 \quad n-1 \quad n \quad n+1$

We have here

$$-\frac{ds}{dt} = \varkappa_n \cdot s \cdot e \qquad s \sim s_{n+1} \tag{35}$$

$$-\frac{ds_n}{dt} = \varkappa_{n-1} \cdot s_n \cdot e - \varkappa_n \cdot s \cdot e \tag{36}$$

$$-\frac{ds_{n-1}}{dt} = \varkappa_{n-2} \cdot s_{n-1} e - \varkappa_{n-1} \cdot s_n \cdot e \tag{37}$$

etc.

The substrates are indexed according to terminal unit while $\varkappa_{n-i}$ refers to bond $n - i$.

From (35) and (36) we obtain

$$\frac{ds_n}{ds} = \gamma_1 \cdot \frac{s_n}{s} - 1 \tag{38}$$

or

$$\frac{s_n}{s} = \frac{1}{1-\gamma_1}\,(R^{\gamma_1 - 1} - 1) \tag{39}$$

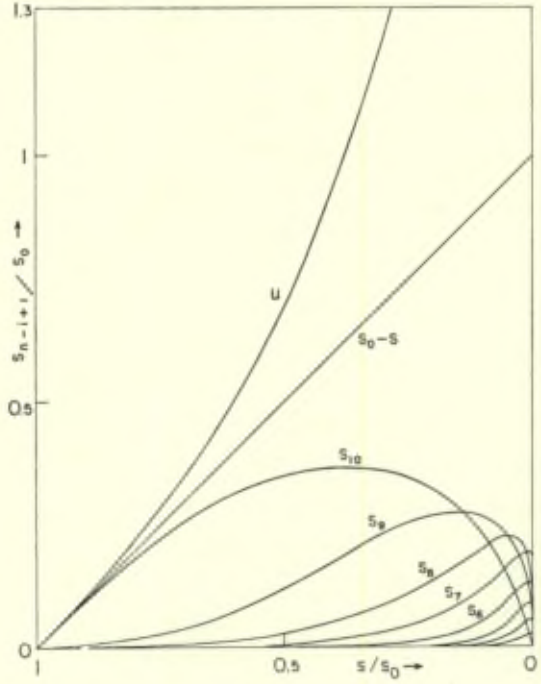


Fig. 4

where

$$\gamma_1 = \varkappa_{n-1}/\varkappa_n$$

Similarly we get

$$\frac{ds_{n-1}}{ds} = \gamma_2 \frac{s_{n-1}}{s} - \frac{\gamma_1}{1-\gamma_1}\left[\left(\frac{s}{s_o}\right)^{\gamma_1 - 1} - 1\right] \tag{40}$$

and so on. The general expression is [see (85)]

$$s_{n-i+1} = s_o P_{i-1} \sum_{=o}^{j=i} \frac{R^{\gamma_j}}{H_{ij}}, \tag{41}$$

where

$$P_{i-1} = \prod_{r=o}^{r=i-1} \gamma_r;\ H_{ij} = \prod_{r=o}^{r=i} (\gamma_r - \gamma_j), r \gtrless j; \gamma_o = 1, \gamma_r = \varkappa_{n-r}/\varkappa_n. \tag{42}$$

As in previous cases we find :

Number of peptide bonds split of the type $n - l$

$$b_{n-l} = \sum_{i=l}^{i=n-1} s_{n-i} \tag{43}$$

Total number of peptide bonds split

$$b = \sum_{l=o}^{l=n-1} \sum_{i=l}^{i=n-1} s_{n-} \tag{44}$$

Concentration of free units of type $n - l + 1$

$$u_{n-l+1} = b_{n-l} \tag{45}$$

Total number of free units $u = b$. The concentrations of residual chains are given by (41). If all γ are unity (41) is transformed into

$$s_{n-i+1} = s_o \cdot R \frac{(-\ln R)^i}{i!} \tag{46}$$

Figure 4 shows the distribution in s and u for this special case; $n = 10$.

4.4 Two adjacent bonds in the protein molecule mutually influence each other so that the rate of splitting of one bond is changed when the other is split and conversely.

We shall consider the following simple peptide

amino end —1—|—2—|—3— carboxyl end
(1 under 1, 2 under 2)

for which we have

$$-\frac{d\sigma}{dt} = (\varkappa_1 + \varkappa_2)\,\sigma \cdot e \tag{47}$$

$$-\frac{d\sigma_{23}}{dt} = \varkappa_2'\,\sigma_{23} \cdot e - \varkappa_1\,\sigma \cdot e \tag{48}$$

$$-\frac{d\sigma_{12}}{dt} = \varkappa_1'\,\sigma_{12} \cdot e - \varkappa_2\,\sigma \cdot e \tag{49}$$

where σ are the concentrations and $\varkappa$ and $\varkappa'$ the characteristic constants before and after the splitting of the adjacent bond. The solution of these equations gives

$$\sigma_{23} = \sigma_0 \frac{\alpha_1}{1-\alpha_2'}\,(\rho^{\alpha_2'} - \rho) \tag{50}$$

$$\sigma_{12} = \sigma_0 \frac{\alpha_2}{1-\alpha_1'}\,(\rho^{\alpha_1'} - \rho)$$

$$\sigma_1 = \sigma_0\,\{1 - \rho - \sigma_{12}/\sigma_0\}$$

$$\sigma_3 = \sigma_0\,\{1 - \rho - \sigma_{23}/\sigma_0\}$$

$$\sigma_2 = \sigma_0 - \sigma - \sigma_{12} - \sigma_{23}$$

where $\alpha_1 = \varkappa_1/(\varkappa_1 + \varkappa_2)$, $\alpha_1' = \varkappa_1'/(\varkappa_1 + \varkappa_2)$, etc; $\rho = \sigma/\sigma_0$
If $\alpha_1' = \alpha_2' = 0$ we naturally find $\sigma_2 = 0$, $\sigma_{12} = \sigma_3$, $\sigma_{23} = \sigma_1$, and

$$\frac{\sigma_{12}}{\sigma_{23}} = \frac{\alpha_2}{\alpha_1}$$

If the above peptide is built into a larger molecule of the type described in section 4.12 we obtain by combining (4) and (47):

$$\sigma = \mathrm{R}^{\alpha_{12}} \times \text{constant}$$

where $\alpha_{12} = (\varkappa_1 + \varkappa_2)/(\varkappa_1 + \varkappa_2 + \sum_{n-2} \varkappa_l)$

Since $\sigma_0 = s_0$ the relation

$$\rho = \mathrm{R}^{\alpha_{12}} \tag{51}$$

is valid, so that the quantities in (50) may be found in terms of R:

σ_{12} : total conc. of peptides with C-terminal unit 2
σ_1 : total conc. of peptides with C-terminal unit 1
σ_{23} : total conc. of peptides with N-terminal unit 2
σ_3 : total conc. of peptides with N-terminal unit 3
σ_2 : conc. of unit 2.

For σ_{23ql}, the concentration of a peptide with q units, formed by

splitting bond 1 and bond l, q units to the right of the former, we find

$$\sigma_{23ql} = \sigma_{23}\,(1 - R^{\alpha_l})\,R_{q-1}^{\Sigma\alpha_i} \tag{52}$$

and similarly for other peptides of this kind.

The above considerations may be extended to include cases where a higher number of mutually dependent bond-pairs are present in the protein molecule, and hence systems with rather universal interaction of bonds may be treated. Since however the possibility of testing our theoretical results experimentally is extremely slight the above suggestions may suffice.

The following case is of more general interest.

4.5 The initial reaction discussed in section 4.2 is a splitting of a peptide bond, not any particular bond, but one out of f bonds that all can be split slowly by the enzyme when the substrate is in its native state. The opening of one of the f bonds increases the rate of hydrolysis of all other bonds, including that of the remaining $f - 1$ bonds.

We have here

$$-\frac{dN}{dt} = N \cdot e \cdot \sum_f \varkappa_r \tag{53}$$

$$-\frac{ds_r}{dt} = s_r \cdot e\,(\sum_n \varkappa_j' - \varkappa_r') - N \cdot e \cdot \varkappa_r \tag{54}$$

$$-\frac{dp_i}{dt} = \varkappa_i' \cdot p_i \cdot e - N \cdot e \cdot \sum_f \varkappa_r \tag{55}$$

$$-\frac{dp_r}{dt} = \varkappa_r' p_r \cdot e + N \cdot e\,(\varkappa_r - \varkappa_r') \tag{56}$$

where the bonds p_r belong to the f-group, while p_i are members of the remaining group of $n - f$ bonds that are assumed not to be split in the native molecule.

s_r is the concentration of one of the f different « denatured » molecules formed by splitting one f-bond in N.

From the above equations we obtain

$$s_r = N_o \frac{\xi_r}{1 - \beta_{Nr}'}(Q^{\beta' Nr} - Q) \tag{57}$$

$$p_i = \frac{N_o}{1 - \xi_i'}(Q^{\xi_i'} - Q) \tag{58}$$

$$p_r = N_o \left[Q^{\xi'_r} - \frac{\xi_r - \xi'_r}{1 - \xi'_r}(Q^{\xi'_r} - Q)\right] \tag{59}$$

and

$$b = n\,N_o - (n-f)N - \sum_{n-} p_i - \Sigma p_r \tag{60}$$

where

$\beta'_{Nr} = (\sum_n \varkappa'_j - \varkappa'_r)/\sum_f \varkappa_r$; $\xi'_i = \varkappa'_i/\sum_f \varkappa_r$; $\xi'_r = \varkappa'_r/\sum_f \varkappa_r$; $\xi_r = \varkappa_r/\sum_f \varkappa_r$.

The average length of degradation products expressed in units is

$$A = \frac{N_o - N}{b}\, n \qquad \text{(for a peptide ring)} \tag{61}$$

Other quantities may be derived on this basis (s_q, etc.) but since the calculations are rather involved we shall limit the treatment to a determination of A for the case that all $\varkappa'_i$ and $\varkappa'_r$ are identical ($= \varkappa'$) and different from all $\varkappa_r$ [which also are identical ($= \varkappa$)]. In this case we find

$$A = \frac{1 - Q}{1 - Q - \dfrac{n-1}{n} \cdot \dfrac{1}{1-\xi'}(Q^{\xi'} - Q)} \tag{62}$$

where $\xi' = \varkappa'/\varkappa f$

4.6 The initial reaction in the enzymic degradation is the splitting of a peptide bond x which may re-form in a reversible process. The opening of x conditions the further hydrolysis of the $n-1$ residual peptide bonds.

The simplest case is one in which a peptide ring is broken at x and becomes susceptible to further breakdown by the enzyme. In this case we have

$$-\frac{ds}{dt} = \varkappa_x \cdot s \cdot e - \varkappa_{-x} s_x \cdot e \tag{63}$$

$$-\frac{ds_x}{dt} = s_x \cdot e \cdot \sum_{n-1} \varkappa_i + \varkappa_{-x} \cdot s_x \cdot e - \varkappa_x \cdot s \cdot e \tag{64}$$

$$-\frac{dp_i}{dt} = \varkappa_i p_i \cdot e + \varkappa_{-x} \cdot s_x \cdot e - \varkappa_x \cdot s \cdot e \tag{65}$$

where $\varkappa_{-x}$ refers to the synthetic process restoring the molecule to its original state. From (63) and (64) we obtain

$$\frac{ds_x}{ds} = \frac{\alpha_x}{\dfrac{s}{s_x} - K_x} - 1 \tag{66}$$

where $\alpha_x = \Sigma \varkappa_i / \varkappa_x$ and $K_x = \varkappa_{-x}/\varkappa_x$ is the equilibrium constant for the process $S \rightleftarrows S_x$. Equation (66) leads to

$$lns = \int \frac{(1 - u\,K_x)\,du}{K_x u^2 + (\alpha_x + K_x - 1)\,u - 1} \tag{67}$$

where $u = s_x/s$, and denoting by $-\varepsilon$ and $-\omega$ the two roots of the equation obtained by putting the denominator in (67) equal to zero we get

$$\begin{aligned} 2\,K_x\,\varepsilon &= \alpha_x + K_x - 1 - \varphi_x \\ 2\,K_x\,\omega &= \alpha_x + K_x - 1 + \varphi_x \\ \varphi_x &= \sqrt{(\alpha_x + K_x - 1)^2 + 4K_x} \end{aligned} \tag{68}$$

and find

$$\frac{s}{s_o} = Y^{\frac{\varepsilon K_x + 1}{\varphi_x}} \frac{1}{\frac{1}{\omega}\frac{s_x}{s} + 1} \quad \text{where} \quad Y = \frac{\frac{1}{\varepsilon}\cdot\frac{s_x}{s} + 1}{\frac{1}{\omega}\cdot\frac{s_x}{s} + 1} \tag{69}$$

Similarly we have

$$\begin{aligned} \frac{p_i}{s_o} = \frac{1}{\varepsilon - \omega}\Bigg[&\frac{\varepsilon\,(K_x\,\omega + 1)}{K_x\,\omega + 1 - \beta_i}\left(Y^{\frac{\beta_i}{\varphi_x}} - Y^{\frac{(K_x\,\omega + 1)}{\varphi_x}}\right) \\ &- \frac{\omega\,(K_x\,\varepsilon + 1)}{K_x\,\varepsilon + 1 - \beta_i}\left(Y^{\frac{\beta_i}{\varphi_x}} - Y^{\frac{K_x\,\varepsilon + 1}{\varphi_x}}\right)\Bigg] \end{aligned} \tag{70}$$

where β_i is given by $\beta_i = \varkappa_i/\varkappa_x$.

From (69) and (70) p_i and s_x may be found as functions of s/s_o. Since however, these expressions are not explicit in s_x/s the calculation of s_q is not possible without approximation. We shall therefore refrain from further discussion and take the above consideration as an example of the extreme complexity introduced when back-reactions are assumed to occur in the initial process.

4.7 Resynthesis occurs during the degradation.

Apart from the special and highly unlikely case that only the original bonds p_i are resynthesized and no « mixing », i.e. combining of previously separated units, occurs, this problem is hardly open to detailed treatment. However, if we assume that all bonds p are equal, we have

$$-\frac{dp}{dt} = \varkappa \cdot p \cdot e - \varkappa_- (ns_o - p)^2 \cdot e \tag{71}$$

If furthermore we assume that resynthesis of S from any of the degradation products does not occur we find as usual

$$\frac{dp}{dlns} = \frac{1}{n}\, p - \alpha_s (ns_o - p)^2 \tag{72}$$

where $\alpha_s = \varkappa_-/\Sigma\varkappa$. (72) leads to

$$\frac{p}{ns_o} = \frac{R^{\varphi/n} + \delta}{1 + \delta \cdot R^{\varphi/n}} \tag{73}$$

where $$\delta = \frac{\varphi - 1}{\varphi + 1},\quad \varphi = \sqrt{1 + 4\, ns_o K},$$

and K, the equilibrium constant, is given by

$$K = \varkappa_-/\varkappa = n\alpha_s$$

For $K = 10^{-1}$ [see e.g. ([16]) page 98] and $n = 40$, we find

$$\varphi = \sqrt{1 + 16\, s_o}$$

and assuming an average « unit » weight of 800 (Mw of protein 32 000) we calculate the figures in Table 1.

TABLE 1

s_o, grams per litre	φ	δ
1	1.00025	0.00025
10	1.0025	0.0025
100	1.025	0.024

If $p/ns_o = 0.5$ i.e. if half of the peptide bonds are split we find by comparison of (73) with the equation

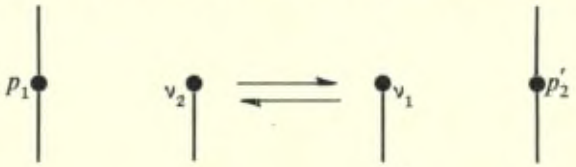
$$p = n \cdot s_o \cdot R^{\frac{1}{n}},$$

valid if no back reactions take place [see (9) and (18)], that the synthetic yield is $n \cdot s_o \cdot 0.0094$ for $s_o = 100$ g/litre. This means that 2 per cent of the peptide bonds opened have become re-closed *at random* and appear in the digest as artifacts produced by the enzyme. For smaller substrate concentration the yield is correspondingly smaller.

4.8 Exchange reactions occur during the degradation.

Due to the complexity of the problem we shall only give a relatively simple example.

The following transpeptidation reactions are assumed to occur



and

where ν indicates a free amino group. No other exchange reactions are permitted. We then have at any stage

$$\nu_1 = s_o - p_1' - p_1 \tag{74}$$
$$\nu_2 = s_o - p_2' - p_2$$

and

$$\frac{dp_1}{dlns} = \alpha_1 p_1 + \alpha_{m1} p_1 (s_o - p_2' - p_2) - \alpha_{-m1} p_2' (s_o - p_1' - p_1) \tag{75}$$

$$\frac{dp_2'}{dlns} = \alpha_2' p_2' + \alpha_{-m1} p_2' (s_o - p_1' - p_1) - \alpha_{m1} p_1 (s_o - p_2' - p_2) \tag{76}$$

$$\frac{dp_2}{dlns} = \alpha_2 p_2 + \alpha_{m2} p_2 (s_o - p_1' - p_1) - \alpha_{-m2} p_1' (s_o - p_2' - p_2) \tag{77}$$

$$\frac{dp_1'}{dlns} = \alpha_1' p_1' + \alpha_{-m2} p_1' (s_o - p_2' - p_2) - \alpha_{m2} p_2 (s_o - p_1' - p_1) \tag{78}$$

Putting $\alpha_1 = \alpha_2' = \alpha_2 = \alpha_1' = \alpha$ and $\alpha_{m1} = \alpha_{-m1} = \alpha_{m2} = \alpha_{-m2} = \alpha_m$ we obtain from these equations

$$p_1 + p_1' = p_2 + p_2' = p_1 + p_2' = p_2 + p_1' = s_o \, \mathrm{R}^{\alpha} \tag{79}$$

and

$$\frac{p_2'}{p_1 + p_2'} = \frac{p_1'}{p_2 + p_1'} = \frac{1}{2}\left\{ 1 - e^{\frac{2\alpha_m s_o}{\alpha}(1 - \mathrm{R}^{\alpha})} \cdot \mathrm{R}^{2\alpha_m s_o} \right\} \tag{80}$$

Since with our assumptions the equilibrium constants of the transpeptidation reactions are unity ($\alpha_m = \alpha_{-m}$) the ratios in (80) approach 1/2 for $R = s/s_o = 0$. The stage of degradation at which this limit is approximately reached depends strongly on α_m as will be seen from figure 5.

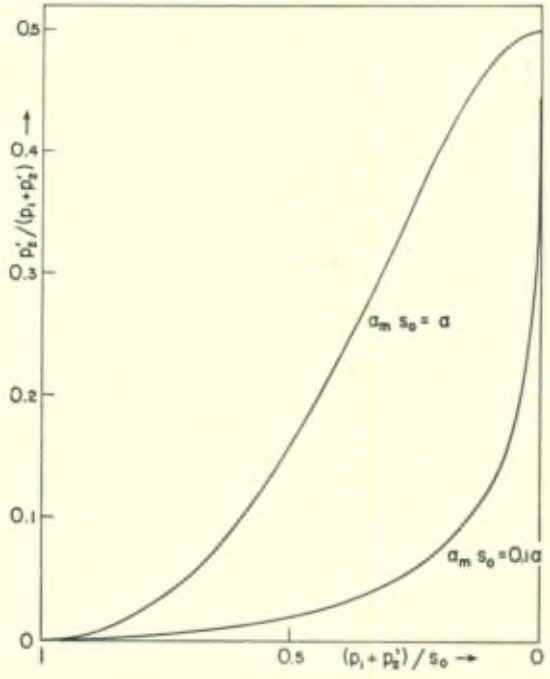


Fig. 5

According to recent results of Fruton and his group (57) (58) such cross reactions should occur in protein digests. However, at the moment it is difficult to estimate the value of α_m/α. From an experiment by Jones, Hearn, Fried, and Fruton (58) we find for the reaction :

50 μM glycyl-L-phenylalaninamide + 50 μM L-argininamide ⟶ 9.8 μM glycyl-L-phenylalanine + 12.2 μM glycyl-L-phenylalaninearginineamide + 22 μM NH_3

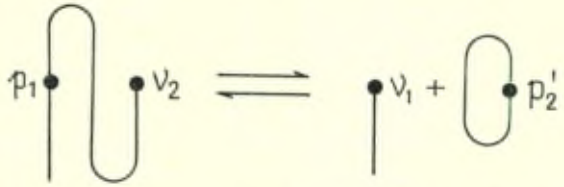
(1 *ml* reaction mixture, 1 hour reaction time, enzyme: beef spleen cathepsin at pH 7)

$$\alpha_m/\alpha = 24.8$$

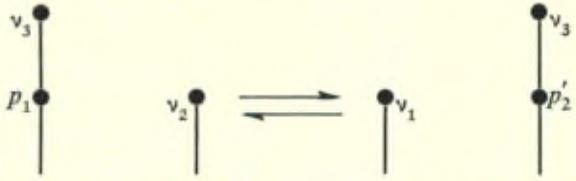
and for $s_o = 100$ g/l (see section 6)

$$2\, s_o \alpha_m / \alpha = 1.6.$$

We have here disregarded the back reaction between glycyl-L-phenylalanineargininamide and NH_3 so that the value is a minimum value. It would therefore seems as if « crossing over » of this kind might play a significant role in creating artifacts in protein degradation by enzymes (see the curve for $2\, s_o \alpha_m / \alpha = 2$ in figure 5), but it should naturally be remembered that a series of highly unlikely reactions are included in our general picture as sketched above, e.g. ring formation:



and the like. The probability for formation of « unnatural » bonds is therefore smaller than indicated. If only pieces of one unit can be exchanged



the probability of exchange will be reduced by a factor of $(1 - R^{\alpha 3})$ [see equations (10)-(15)], and if it is a further condition that this unit is an amino acid the probability may be zero (namely if no unit is as small as an amino acid). If on the other hand p_1, say, can react with $n - 1$ different amino groups ν_i to form $\Sigma p_i'$ false bonds, we may find a substantial quantity of the constituents of p_1 in artifacts even at early stages of the degradation.

Naturally none of the artifacts mentioned above will survive exhaustive degradation since we have assumed that also unnatural bonds are hydrolyzed. If therefore the protein degradation is a one-by-one process (or if degradation in a « zipper » system is complete) the digest will only contain single units, the amino acid sequences of which are untouched by the proteolytic reaction.

Permanent artifacts will only be formed if there is interaction

between the bonds. If for example an unsusceptible bond p_3 in the vicinity of p_2 becomes susceptible by the formation of p_2' while p_2' becomes resistent after the opening of p_3 we may find an unnatural sequence of amino acids in the final mixture of units. We shall leave it an open question whether situations like this are likely to exist or not.

4.9. Concluding remarks

In the above elementary considerations the disappearing of substrate has been used as a « clock » and all other reactions have been put in relation to this process. By such a procedure the unknown concentration of « free » enzyme can be eliminated and the expressions obtained are therefore free of assumptions regarding the detailed reaction of the enzyme with the products appearing during the degradation. In a few cases it has been shown that considerable difficulties are encountered when intermediates of finite concentration are formed by back reactions.

If the real time-dependence of the enzymatic reaction is sought simple kinetic equations can only be obtained in certain cases, especially if e is equal to the total enzyme concentration or if the enzyme is bound to the same degree by all degradation products (and the substrate).

Similar difficulties are encountered where two different enzymes E_1 and E_2 are assumed to act simultaneously on the same substrate. Unless the ratio e_1/e_2 is constant during the degradation no general equations can be set up. In a following chapter an interesting system of this kind will be discussed.

5. THE INTERACTION OF PROTEINS WITH ENZYMES

In the following a few examples will be given of important enzyme substrate systems studied experimentally. A complete survey of the work carried out in this field shall not be attempted.

5.1 The peptic breakdown of ovalbumin.

The total number of bonds in ovalbumin that can be split by pepsin is not exactly known, in part because there seems to be a great deal of interaction between the susceptible bonds, and the $\varkappa_i'$

(see section 4.4) are low. However, if we use the values of Calvery et al. ([59]) ([60]) and Moring-Claesson ([61]) we arrive at an approximate value of 120 bonds per molecule (Mw 45 000). If these were to be hydrolyzed according to the scheme in section 4.11 (assuming ovalbumin to be a peptide ring) we should find the values in Table 2 for the average length of the degradation products (A) measured in units (average weight $\sim$ 400).

TABLE 2

$R = \frac{s}{s_0}$	A calculation	A approx. from Moring-Claesson	$\frac{s + s_n}{s_0}$
0.9	114		0.995
0.7	101	2 — 3	0.951
0.5	87		0.848
0.2	60		0.524
0.1	47	1 (— 2)	0.232
0.01	26		0.047

This table shows the striking feature first observed by Tiselius and Eriksson-Quensel ([7]) that the breakdown of ovalbumin by pepsin does not lead to formation of average-sized degradation products as it unquestionably would do if all susceptible bonds were equally open to the enzyme in the native molecule and in the intermediary products.

The digests contain practically only native protein and degradation products consisting of 1-3 units. Even if S_n (of which there are 120 different species) is included in the fraction « unsplit substrate », because of its having the same Mw, the results are entirely incompatible with the theory (Table 2). If however we adopt the picture

in section 4.2 we obtain the results in Table 3, in the calculation of which we have used the expression

$$A = \frac{N_o - N}{b} = \frac{1 - Q}{1 - Q - \dfrac{Q^\beta - Q}{1 - \beta}} \tag{81}$$

TABLE 3

$Q = \frac{N}{N_o}$	A		
	$\beta = 1$	$\beta = 2$	$\beta = 3$
0.9	19.3	10.0	6.9
0.7	6.0	3.3	2.4
0.5	3.3	2.0	1.6
0.2	1.7	1.3	1.1
0.1	1.3	1.1	1.0

The agreement is naturally much better here. The interesting fact is that this result is obtained by assuming the existence of an initial reaction with a characteristic constant $\varkappa_N$ which is as high as one half of that for any of the peptide bonds liberated:

$$\beta = \varkappa/\varkappa_N = 2$$

If the calculation is based on the assumptions in section 4.5 we obtain from (62) essentially the same result provided

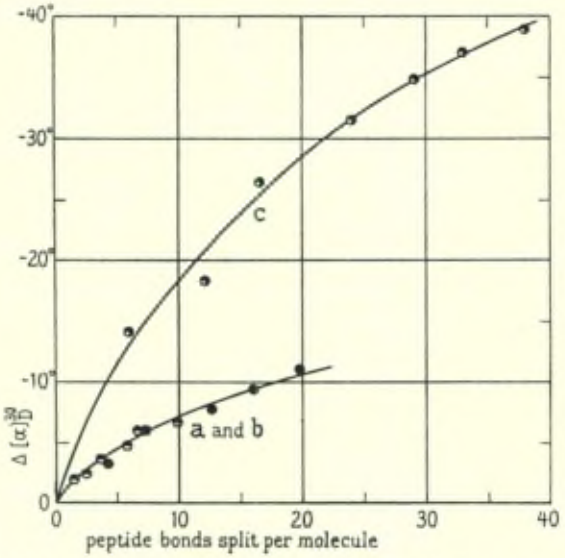
$$\xi' = \varkappa'/\varkappa f = 2$$

which means that the rate of hydrolysis of an f-bond increases $2f$ fold after the initial reaction. It should however be pointed out that in the derivation of the expressions in 4.5 it was tacitly assumed that the rearrangement of the substrate molecule, the appearance of susceptible bonds after the splitting of an f-bond, is instantaneous; or that the breaking of the bond and the rearrangement occur simultaneously. In the first case it might be imagined that the breaking of the bond *activates* the substrate molecule, in the second case that only f-bonds in *activated* molecules are susceptible. The third case: that the substrate molecules undergo spontaneous rearrangement after the opening of an f-bond and that they have a measurable lifetime in an almost native state, is theoretically very difficult

to handle. In both of the former situations however, the activation energy of the rearrangement is somehow involved and manifests itself in the $2f$ fold increase in $\varkappa$ necessary to explain the experiments on the present basis.

We must therefore consider it highly probable that a denaturation of the ovalbumin molecule takes place either prior to the opening op peptide bonds (section 4.2) or in the very first stages of the degradation (section 4.5). This view is supported by or actually based upon experiments other than the above named. We may refer to a paper by Bull and Currie ([62]) in which it was shown that the heat of activation in the initial attack of pepsin upon ovalbumin is very nearly equal to the heat of activation for the denaturation of ovalbumin in the acid region [see also Currie and Bull ([63])]. Korsgaard-Christensen ([19]) has found that addition of urea caused

Fig. 6



Hydrolysis of egg albumin by pepsin. The change in optical activity as a function of the number of peptide bonds hydrolyzed.

Curves a, b, and c correspond to curves a, b, and c of fig. 31.

Curve a: 2.5 % egg albumin, pH 1.95, 0.05 M citrate buffer, 0.092 % pepsin (A), 30° C.

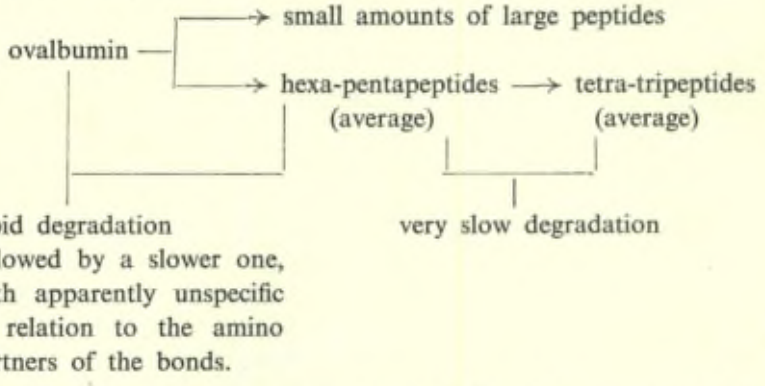
— b: 2.5 % egg albumin, pH 1.65, 0.05 M citrate buffer, 0.092 % pepsin (A), 30° C.

— c: 2.5 % egg albumin, pH 1.85, 0.05 M citrate buffer, 10 % urea, 0.092 % pepsin, 30° C.

a pronounced increase in the change of optical rotation per peptide bond split as shown in figure 6. This increase, by our formulae, corresponds to a marked decrease in β or ξ', if we ascribe an approximately constant specific optical rotation to the degradation products, different from that of the native protein. Due to the uncertainty of this assumption we shall omit quantitative calculations, but the fact that urea has this pronounced effect naturally strengthens the view that a denaturation occurs in the initial stage of the breakdown.

The detailed analysis of the ovalbumin-pepsin system does not promise very much information about the path-way of the degradation. Moring-Claesson was able to isolate some intermediary products though of very small size. By filter paper chromatography of hydrolysates of different peptide fractions obtained by frontal analysis she demonstrated the amino acids glu, asp, gly, ala, val and leu in the fractions with NH_2 - N/total - N > 0.25 (di-tetra-peptides) while all the amino acids of ovalbumin were represented in the fractions with NH_2 - N/total - N $= 0.10 - 0.20$. The actual lengths of these latter peptides naturally depend upon their content of lysin, histidine and especially arginine which contribute to the total nitrogen by more than one nitrogen per molecule.

Desnuelle, Rovery and Bonjour (64) have carried out an investigation in which they have made use of Sanger's dinitrofluorbenzene method. They arrive at the scheme below:



and their results are therefore qualitatively in good agreement with our considerations.

Bizerte and Boulanger (65) (66) have made an extensive study of the peptides formed in the reaction. They have successfully made

use of chromatography on resins for the separation of these peptides, mainly effecting a fractionation into groups of acid, neutral, and basic peptides. The amino acids in these fractions were qualitatively estimated by filter paper chromatography but no endgroup determinations were made. Due to the extreme complexity of the digests these experiments cannot give a clear picture of the detailed action of pepsin upon ovalbumin. Combined determinations of N- and C-terminal amino acids would at least give some information about the type of bonds cleaved although the matching of the original pairs is a problem beyond human efforts.

5.2 The peptic breakdown of β-lactoglobulin.

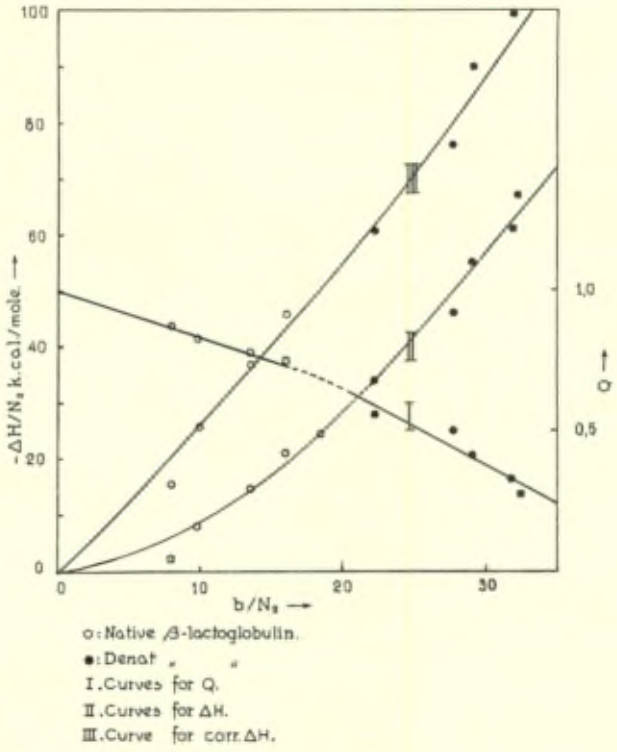
This case seems to be very much like the preceding one. Haugaard and Roberts ([67]) have studied the heat of reaction for both native and alkali-denatured lactoglobulin and found the results summarized in figure 7. They were unable to detect larger intermediary degradation products, the digest apparently being composed of unchanged protein molecules and final units of 5-7 amino acids at all stages. The values of β or ξ' must therefore be high, probably decidedly higher than in the case of ovalbumin since in dilatometric experiments these authors found a pratically linear relationship between the volume change observed in the enzymatic reaction and the release of amino groups. Their measurements were made on alkali-denatured lactoglobulin, but similar experiments carried out by Jacobsen and Linderström-Lang (unpublished) have given essentially the same results for the native protein, although a slight bend of the curve could be detected at small cleavages where $Q \sim 1$. For this reason values for β or ξ' of about 40 are probably of the correct order of magnitude (see section 5.5).

The average slopes (ΔV_{sp}) obtained, — 24 and — 29 *ml* per peptide bond split for denatured and native protein respectively, are numerically high as compared to the theoretical values of — 6 *ml* per bond calculated for free peptide chains at pH 1.5 ([14]). It is therefore evident that both in native and in *alkali denatured* lactoglobulin brought to pH 1.5, secondary structures are present which are abolished with loss of volume in the reaction with pepsin. We may therefore write:

$$\Delta V_{sp} = \Delta V_D (1 - Q)/b - 6 \qquad (82)$$

where ΔV_D is the additional volume change per mole of protein.

Fig.7



On this basis we calculate

for native protein . . . . . $\Delta V_D = -1\,400$ *m*l/mole,
and for alkali denatured protein. . . . $\Delta V_D = -\ 700$ *m*l/mole.

The certainty of these values naturally strongly depends on the determination of Q which was made by precipitation with trichloracetic acid (67). If part of the degradation products are co-precipitated with the protein in this procedure the values of $-\Delta V_D$ will be lower.

As regards the heat of reaction it shows a marked lack of proportionality to the quantity of opened bonds (figure 7), a result which seems difficult to explain in view of the fact that the reaction otherwise behaves as one-by-one process. However, even if β or ξ' are

high, a deviation from strict proportionality will be observed provided the heat of the initial reaction (ΔH_D) is markedly different from that involved in the subsequent opening of the bonds (ΔH_p). If therefore in the equation

$$\Delta H = N_o (1 - Q) \cdot \Delta H_D + b\Delta H_p \qquad (83)$$

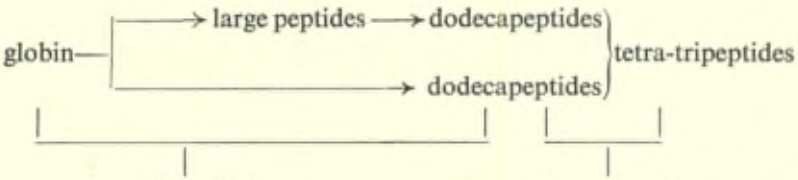
we attribute to ΔH_D a positive value we may obtain an approximately constant value for ΔH_p even if b is nearly proportional to $(1 - Q)$. Curve III in figure 7 is thus obtained by adding to $-\Delta H$ the quantity 100 $(1 - Q)$ kcal corresponding to a value for ΔH_D of + 100 kcal. In view of recent studies by Buzzell and Sturtevant ([68]) who report values of about + 50 kcal for the heat of denaturation of pepsin, this assumption seems acceptable especially since β-lactoglobulin is very stable in acid solution (in contradistinction to ovalbumin).

The corresponding value for alkali-denatured protein is still more uncertain. The upper part of curve III in figure 7 is drawn on the basis of $\Delta H_D = + 60$ kcal, but it remains a matter of taste which value to select. The value of ΔH_p obtained for both native and alkali denatured β-lactoglobulin, viz. — 3 to — 4 kcal is a little on the lower side but hardly improbable.

In view of the fact that alkali denatured lactoglobulin is hydrolyzed with an initial velocity of about three times that found for the native protein, we should expect a smaller value of β or ξ' for the former substrate. However, the experiments do not permit a decision of this question. We can only surmise that the heat of activation is very large in both cases. Korsgaard-Christensen ([19]) has studied the effect of urea upon the rate and the relation between optical rotation and liberation of carboxyl groups. The results are similar to those found for ovalbumin, but the concentration of urea required for appreciable activation is considerably higher (29 per cent). The interesting phenomenon was observed that, from zero time until about ten bonds were split per molecule, an increasing quantity of a highmolecular intermediary product was formed, distinguishable both from native protein and from lower degradation products by its insolubility in aqueous salt solution at pH 4.5. With increasing degree of degradation the substance tended to disappear again. This observation is adequately explained by assuming a substantial lowering of β or ξ' by urea.

5.3 The peptic breakdown of horse globin.

According to Desnuelle, Rovery and Bonjour (64) the breakdown of horse globin by pepsin proceeds according to the scheme



Horse globin, which is a denatured protein, has a Mw of 64 000 and contains 5-6 peptide chains in its molecule. The formation of dodecapeptides would involve the splitting of 8-10 bonds per chain while the further breakdown to tripeptides would require the opening of further 20 bonds per chain. Let us tentatively adopt the reasoning in section 4.12 and put 10 of the kappas ($\varkappa_l$) equal to each other and assume that they are g times as large as the remaining 20 ($\varkappa_m$). We then have

$$b_l/s_o = 10\,(1 - R^{\frac{1}{10}})$$

$$b_m/s_o = 20\,(1 - R^{\frac{1}{10g}})$$

valid for the breakdown of a single of the chains. Table 4 shows the results of a primitive calculation based on the figures in Table VII of Desnuelle et al.

TABLE 4

Time of hydrolysis (min.)	$\frac{b}{10\ s_o}$ per chain, found	$R = \frac{s}{s_o}$	$s_l/10\ s_o$ conc. of dodecapeptide calc.	s_l relative conc. found for non protein fraction
30	0.276	4.10^{-2}	0.124	0.194
90	0.638	4.10^{-5}	0.494	0.485

Considering the assumptions made and the difficulties of discriminating between non-protein and protein when intermediary products are abundant this agreement is satisfactory. If g is chosen sufficiently large, e.g. 10, the hydrolysis of the bonds p_m will not be measurable until most of the p_l are split.

In the absence of values for R to compare with, the argumentation is naturally somewhat airy. So much can however be said that there is nothing in these experiments that is incompatible with the idea that horse globin behaves as a free peptide in its reaction with pepsin.

A detailed calculation of the α-values for the different bonds by the formula (see section 4.11)

$$b_i = n_i(1 - R^{\alpha_i})$$

cannot be made since the n_i-values are unknown. For the susceptible bonds in which ser, phe, and ala are engaged our assumption that the values of α are identical is approximately true. The α-value for leu seems however to deviate considerably.

5.4 The peptic breakdown of ribonuclease.

Anfinsen ([69]) has studied this process and found that approximately 10 bonds are split per molecule. The resulting peptides, containing on the average 7 to 8 amino acids, do not readily disperse, but seem in part to retain the secondary structure characteristic of the ribonuclease molecule. We seem here to meet with an entirely new situation in that we have a protein molecule, which is stabilized be a very firm secondary structure and which nevertheless seems to be enzymatically degraded while in its native form, without preceding or subsequent denaturation. The 10 bonds in question must therefore be exposed in the native configuration and it should be possible to adopt the formulae of section 4.11 to the kinetic problem. If no dispersion of the peptides takes place, the concentration s_{qr} of ribonuclease molecules in which q bonds of a given kind are split is found from

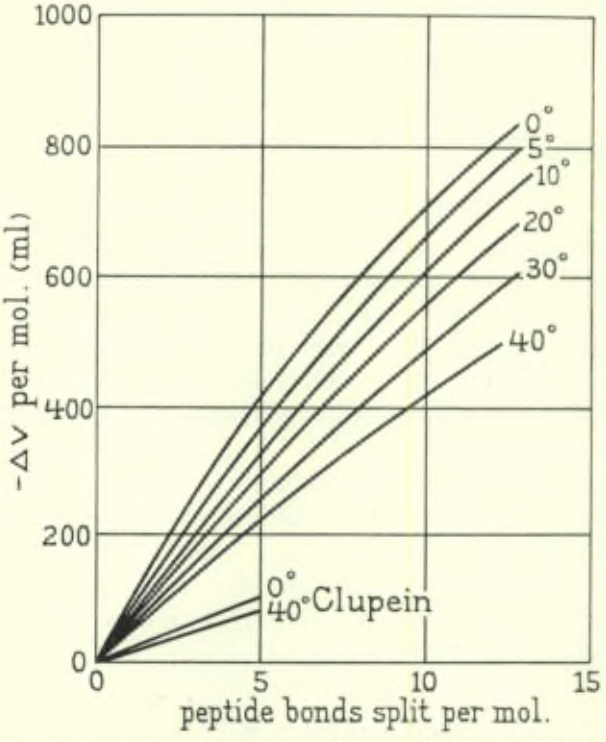
$$s_{qr} = s_o \prod_{i=1}^{i=q} (1 - R^{\alpha_i}) \prod_{i=q+1}^{i=10} R^{\alpha}$$

5.5 The tryptic breakdown of β-lactoglobulin.

This case has previously been treated by the author at some length ([15]) ([16]) and the discussion will therefore be as brief as possible.

β-lactoglobulin was the first protein used in our study of the volume change in proteolytic degradation ([14]), where it was found that the decay of the protein structure brought about by the attack of the enzyme was accompanied by a volume contraction which was considerably larger than that predicted on the basis of the electrostriction around the carboxylate and ammonium ions formed in the reaction. The latter contraction was shown to be smaller than 20 ml/mole per bond split under the given conditions (pH: 6-10, $NH_4Cl - NH_3$ buffer) and the large volume effects observed — up to 100 ml per mole bond split — was explained by attributing to the native β-lactoglobulin molecules a volume that

Fig. 8



Volume change in the tryptic breakdown of β-lactoglobulin at different temperatures. Curves for clupein given for comparison.

was about 700 ml/mole protein higher than that of its unfolded peptide chains ([15]). This view was qualitatively supported by measurements of the contraction accompanying denaturation of lactoglobulin by base or urea ([15]) ([16]) ([64]).

What interests us especially in the present connection is that the initial volume change observed was strongly dependent upon temperature showing a 2-3 fold decrease from 0° to 40° C (fig. 8). This phenomenon could not be explained by assuming a contraction of the native molecules with increasing temperature. Since however the curves relating ΔV with b are strongly bent at 0° C but approach straight lines with rising temperature and falling initial volume change, the obvious explanation of the phenomenon is that the mechanism of the breakdown is temperature dependent and that the process from being of the « zipper » type at 0° changes into a one-by-one reaction at 40° C.

If we write again

$$\Delta V = N_o (1 - Q) \Delta V_D - 20\, b \tag{82}$$

and

$$b = N_o[n (1 - Q) - (n - 1) \frac{Q^{\xi'} - Q}{1 - \xi'}] \quad \text{(section 4.5)} \tag{83}$$

where n is about 35 [see ([70])] we get at all stages of the degradation

$$\Delta V/b = -40$$

if ξ' is large, whereas for finite values of ξ' we obtain

$$\Delta V_{sp} = \frac{d\Delta V}{db} = -20 + \Delta V_D/[n + \frac{n - 1}{1 - \xi'} (\xi' Q^{\xi'-1} - 1)] \tag{84}$$

which for $Q = 1$ (the initial point) leads to

$$(\Delta V_{sp})_o = -720$$

and for $Q \longrightarrow 0$ to

$$(\Delta V_{sp})_\infty = -700 \frac{\xi' - 1}{n \xi' - 1} - 20;\ \xi' \geqq 1$$

and

$$(\Delta V_{sp})_\infty = -20;\ \xi' < 1$$

These equations indicate that if $\xi' = 0$, ΔV_{sp} will be constant and equal to -720 until all N has disappeared at which stage ΔV_{sp} will change discontinually to -20. For $\xi' \leqslant 1$, ΔV_{sp} will

change from — 720 to — 20 in the lifetime of N. The high value of $(\Delta V_{sp})_o$ as compared to that experimentally found, ~ -100, would seem to indicate that the cleavage of more than one bond is required for the rearrangement postulated in section 4.5. However, the initial slope of the ΔV-b-curves is impossible to determine with any degree of accuracy, and our experimental values for $(\Delta V_{sp})_o$ are actually more nearly equal to $\Delta V/b$ for $b/N_o \sim 1-2$, a quantity which is lower than $(\Delta V_{sp})_o$ and depends upon ξ' as shown in Table 5.

TABLE 5

ξ'	$-\left(\frac{\Delta V}{b}\right)_{b/N_o = 1}$	$Q_{b/N^o = 1}$
0	720	—
0.5	215	0.722
1	164	0.794
2	130	0.843
5	97	0.890
10	81	0.913
20	67	0.933
40	55	0.950
80	45	0.964
∞	40	—

With these assumptions values of ξ' from 5 to 80 would cover the variations in figure 8 corresponding to a twofold increase of this constant for a rise in temperature at 10° C. Since $\xi' = \varkappa'/\varkappa f$, we find

$$\frac{d\ln \varkappa f}{dT} = \frac{d\ln \varkappa'}{dT} - 0.0694$$

which, if a reasonable value for $d\ln\varkappa'/dT$ is adopted [see (15)] will lead to a very low positive value for the heat of activation in the initial reaction

$$N \longrightarrow s_r \text{ (see section 4.5).}$$

The simplest explanation of this phenomenon is that β-lactoglobulin exists in two forms which are in equilibrium with each other and of which one (S″) is resistent to trypsin, the other (S′)

not. Component S'' predominates at higher temperature so that the equilibrium constant k_s of the reaction

$$s' \rightleftharpoons k_s s''; \qquad s' + s'' = N; \qquad s' = N/(1 + {}^1/k_s)$$

increases with decreasing temperature. If the rate of equilibration is assumed to be very high we find instead of (83)

$$b = N_o \left[n(1 - Q) - (n - 1)\frac{Q^{\varepsilon'} - 1}{1 - \varepsilon'} \right] \tag{85}$$

where

$$\varepsilon' = (1 + 1/k_s)\xi' \tag{86}$$

In this way we may place the whole temperature variation on k_s, and shifting from the standard state of « (N) » to that of « (S') » the reaction

$$S' \longrightarrow S_r$$

can be given any positive heat of activation within reason. This primitive formulation may be replaced by others. For example we may take the Michaelis complex, (*es*), as standard state and place the temperature variation on k_m [see equations (1)-(3)] which must then rise with rising temperature, or we may leave the question of standard states open and assume an apparent high negative entropy of activation ([71]) in order to explain that the reaction rate is finite although ΔH^{+}_{+} is zero. Since however we have reasons to assume that denaturation plays a part in the initial reaction the latter formulation seems less adequate.

The whole question about the nature of the initial reaction shall not be discussed in detail here. The facts that β-lactoglobulin is easily denatured in weakly alkaline solution (pH 9-10) and that urea in moderate concentration is a strong activating agent in the tryptic reaction ([72]) ([19]) seem to us to support the view that denaturation is the process which changes the susceptibility of the peptide bonds in the initial phases of the degradation. As in case of the peptic hydrolysis of β-lactoglobulin Korsgaard-Christensen ([19]) has been able to isolate considerable quantities of high-molecular intermediary products from tryptic digests at pH 9.25 where the conditions are favorable for an acceleration of the initial process relative to the subsequent degradation.

For the rest we shall refer to earlier publications ([15]) ([16]).

5.6 The tryptic and chymotryptic digestion of horse globin.

Rovery, Desnuelle and Bonjour ([73]) have made a detailed analysis of these processes. The methods and general course of the investigations are the same as in section 5.3 but the results are naturally different, the main feature of the chymotryptic and tryptic degradation being the characteristic formation of a core of « protein » nature which is almost completely resistent to further attack by the enzymes.

TABLE 6

Bonds split in the chymotryptic hydrolysis of horse globin.

AMINO PARTNERS OF THE BONDS	Average number of bonds split in 1 mole of globin (63 400 g)	
	Beginning of hydrolysis 30 minutes NPN/T - N = 0.26	End of hydrolysis 4 3/4 hours NPN/T - N = 0.53
Alanine	8.4	16.6
Phenylalanine	4.5	4.2
Serine	3.0	6.8
Threonine	2.2	4.9
Histidine		8.3
Arginine		4.3
Other partners	3.0	13.0
Total (van Slyke)	23.2	68.0

TABLE 7

Distribution of amino endgroups between the « protein » and « non-protein » fractions of a chymotryptic digest of horse globin (end of reaction).

NATURE OF ENDGROUPS	NUMBER OF ENDGROUPS		
	Total hydrolysate (63 400 g)	« Protein » fraction (15 850 g)	« Non-protein » fraction (47 550 g)
Valine	4.8	2.3	2.5
Alanine + phenylalanine	20.8	0.7	20.1
Hydroxyaminoacids + glutamic acid	14.4	0.4	14.0
Histidine	8.3	0.4	7.9
Arginine	4.2	1.2	3.0
Other amino acids	9.4	0.5	8.9
Total	61.9	5.5	56.4

The results for chymotrypsin, which we may select as an example, are summarized in the Tables 6 and 7.

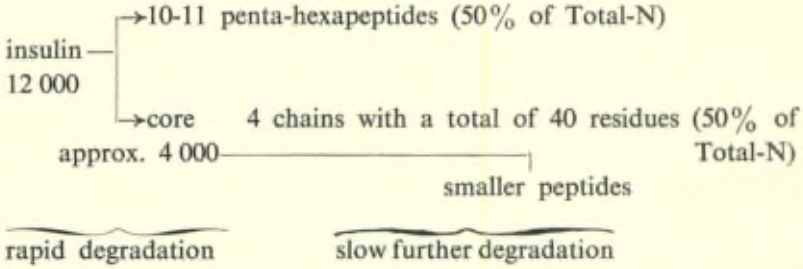
Since the specificity of chymotrypsin (and trypsin) is mainly regulated by the side chains of the carboxyl partners of the susceptible bonds, most of the N-terminal amino acids found in this analysis are presumably linked to either tyrosine, phenylalanine or methionine in horse globin ([49]). Due to a relatively high degree of bond interaction the kinetics shall not be discussed in detail.

5.7 The enzymatic breakdown of insulin.

The chemical constitution of insulin is well known through the admirable work of Sanger, Tuppy and Thompson ([56]) in which extensive use was made of different enzymes as analytical tools, and where endgroup determinations by the dinitrofluorbenzene method have played an important rôle in determining the amino acid sequence in the degradation products and thereby in the protein itself. So far however, no quantitative kinetic analysis has been made and the author shall therefore limit the discussion to a brief review of two papers by Butler, Dodds, Philips and Stephen ([74]) ([75]). One of these papers deals with the peptic degradation of insulin ([75]) and presents evidence that in a first rapid reaction involving 10 bonds per sub-molecule of 12 000, the ratio NH_2 - N/Total N in the non-protein fraction is approximately constant. It would therefore seem as if the insulin molecule had to undergo a slow initial rearrangement prior to the main attack upon the peptide bonds. It is however important to emphasize in this connection that the smaller the molecule investigated and the fewer the susceptible bonds the less pronounced is the dependence of the average size of reaction products (A) upon s/s_0 in the system treated in section 4.12. In the case of insulin, which in its sub-molecule of Mw 12 000 contains 4 peptide chains each with only 2-3 susceptible bonds, ([75]), the situation is very different from that encountered in the peptic breakdown of ovalbumin with its 120 bonds.

It is however doubtful whether the experiments are sufficiently accurate to give unequivocal information about the mechanism of the degradation.

The same is true for the interesting investigation by the same authors, ([74]), of the chymotryptic degradation of insulin. This reaction follows the scheme



where the core, which contains most of the cysteine, may be separated from the peptides by precipitation with 0.25 N TCA. The quantitative relation is approximately as given in Table 8 which shows that s_{pept}/s_{core} is constant. The ratio NH_2 - N/Total-N in the non-protein material was found constant as in the pepsin degradation.

TABLE 8

1 — R	s_{pept}/N_o	s_{core}/N_o	s_{pept}/s_{core}
0.2	0.11	0.10	1.1
0.4	0.22	0.18	1.2
0.6	0.31	0.29	1.1
0.8	0.42	0.38	1.1
0.9	0.52	0.38	1.4

It would be interesting to investigate the chemical composition of the peptide fraction for different values of R. If the number of different peptides exceeds that corresponding to the cleavage of the maximum number of bonds it would imply that interaction between bonds occurs.

5.8 The transformation of ovalbumin into plakalbumin.

The problem of protein transformation, which we finally shall consider, has its origin in the precursor-enzyme systems so admirably studied by Northrop, Kunitz, Herriott and others ([76]). When a relatively new system like the above has been chosen for the following considerations it is because it has been studied in greater detail by Ottesen and our group than most other systems of this type, a fact which in part is connected with its relative simplicity. The author regrets, that due to lack of space, other most interesting systems e.g. the transformation of chymotrypsinogen into chymotrypsin ([2]) or fibrinogen into fibrin ([77]) ([78]) ([79]) ([80]) ([81]) ([82]) can not be dealt with here. Reference is made to ([49]) ([16]).

According to Ottesen and Villee ([83]) and Ottesen and Wollenberger ([43]) there exists in ovalbumin the sequence of amino acids given below

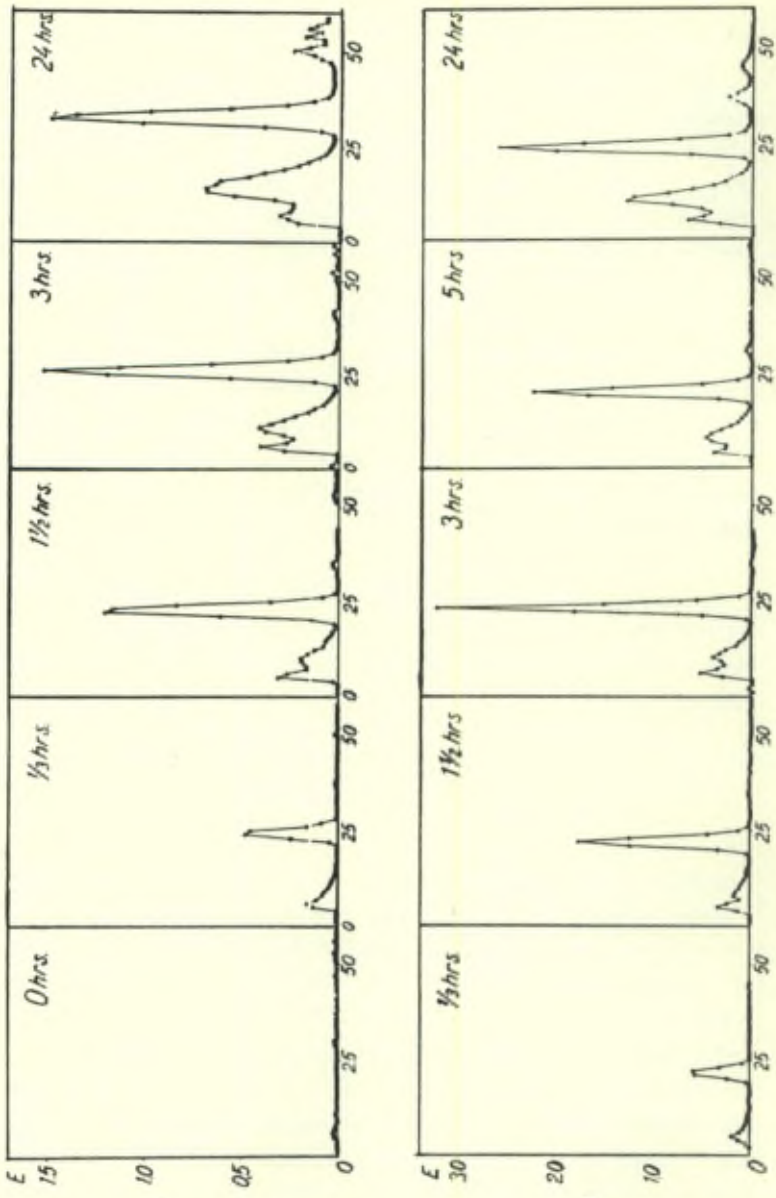
— ala — gly — val — asp — ala — ala —

By the action of a proteinase from *Bacillus subtilis* [which has now been crystallized by Güntelberg and Ottesen ([84])] 3 peptides are formed from this piece of a peptide chain namely

A ala — gly — val — asp — ala — ala
B ala — gly — val — asp
C ala — ala

These peptides have been isolated in relatively pure state by chromatography on starch columns and their amino acid sequence ([43]) ([16]) has been determined by the Edman method ([37]). In the enzymatic reaction at pH 6.4 and 20° C, A and C appear rapidly, while B is slowly formed (figs. 9 and 10).

According to Ottesen and Villee the change in physical properties of the protein, characteristic of the transformation of ovalbumin into plakalbumin [substantial increase of the solubility in ammonium sulphate at pH 6.4 besides change in crystal form ([85])] is directly correlated with the appearance of A and C so that ovalbumin which has lost its hexapeptide A *or* its dipeptide C has assumed the properties of plakalbumin. The appearance of B is considered a secondary phenomenon. In agreement with this view Linderström-Lang and Ottesen ([85]) were originally unable to separate the plakalbumin formation from the appearance of non-



A and B. Effluent chromatograms of unhydrolyzed peptides isolated from the reaction mixture at successive times.

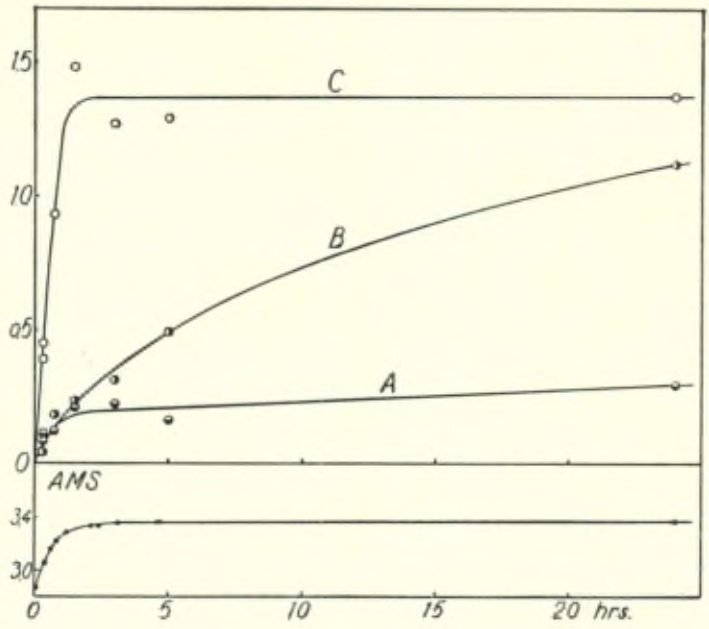
A. Experiment $^{27}/_{10}$ 1949.

B. Experiment $^{17}/_{2}$ 1950.

In each figure the ordinate is E, the extinction coefficient, and the abscissa is the fraction number.

288 Order of peaks from left to right : peptides A, B and C.

Fig. 9



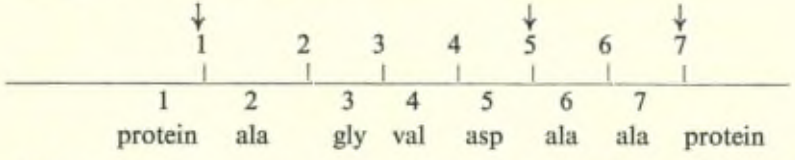
Upper figure: The increase in the amount of peaks A, B and C liberated with time. The extinction values for the fractions constituting each peak have been summed, converted to "μmols alanine color", and plotted against time. For comparison, the lower figure shows the results of concurrent titrations with 95 per cent saturated ammonium sulfate (AMS) plotted against time.

Fig. 10

protein nitrogen even though the enzyme concentration and the pH of the reaction was varied considerably.

On the basis of these facts the author has previously discussed the mechanism of the release of peptides ([16]), and the following is a more detailed presentation of this problem, supplemented by new considerations based upon recent experiments by Steinberg and by Ottesen (unpublished data).

If we consider the peptide piece:



which we assume posesses 3 susceptible bonds, we may distinguish between three different cases.

5.81 Bond 7 is open in ovalbumin.

This case corresponds to the picture given by Ottesen and Villee ([83]). Referring to section 4.4, we may write

$$\sigma_{27} = \sigma_0 \frac{\alpha_1}{1 - \alpha_5'} (\rho^{\alpha_5'} - \rho) = \sigma_A \; : \qquad \text{concentration of peptide A}$$

$$\sigma_{15} = \sigma_0 \frac{\alpha_5}{1 - \alpha_1'} (\rho^{\alpha_1'} - \rho) = \sigma_{P_I} \; : \qquad \text{conc. of plakalbumin I ([83])}$$

$$\sigma_1 = \sigma_0 \{1 - \rho - \sigma_{15}/\sigma_0\} = \sigma_{P_{II}} \; : \qquad \text{conc. of plakalbumin II}$$

$$\sigma_{67} = \sigma_0 \{1 - \rho - \sigma_{27}/\sigma_0\} = \sigma_C \; : \qquad \text{conc. of peptide C}$$

$$\sigma_{25} = \sigma_0 \{1 - \rho - \sigma_{15}/\sigma_0 - \sigma_{27}/\sigma_0\} = \sigma_B \; : \text{conc. of peptide B}$$

where nothing yet is said about the relative values of α_1, α_1', α_5, and α_5', the characteristic constants for the two susceptible bonds 1 and 5. The assumptions, now, that fit the experiments in ([85]) and ([83]) best are

$\alpha_1 = 0.20$: Relative rate const. for bond 1; release of A from ovalb.
$\alpha_5 = 0.80$: Relative rate const. for bond 5; release of C
$\alpha_1' = 0.46$: Relative rate const. for bond 1; release of B from P_I
$\alpha_5' = 0$: Bond 5 is not broken in the free hexapeptide.

We therefore get

$$\sigma_A = \sigma_0 \cdot 0.2\,(1 - \rho)$$

$$\sigma_{P_I} + \sigma_{P_{II}} = \sigma_0\,(1 - \rho)$$

$$\sigma_C = \sigma_0 \cdot 0.8\,(1 - \rho)$$

$$\sigma_B = \sigma_0 \cdot 0.8 \left\{1 - \rho - \frac{\rho^{0.46} - \rho}{0.54}\right\}$$

which give the results in Table 9 showing a fair agreement with the crude experiments by Linderström-Lang and Ottesen. Bond 5 is therefore split about 4 times faster than bond 1, initially, while *the rate constant for the latter increases more than twofold after the opening of bond 5.* The reason why nevertheless B appears so slowly is that the splitting of bond 1 here is bound to succeed the splitting of bond 5, which reaction produces the new substrate for the enzyme: plakalbumin I.

TABLE 9

ρ	Atoms of nitrogen released per mole ovalbumin	
	Found (85)	Calc.
1	0	0
0.8	0.56	0.59
0.6	1.22	1.27
0.4	2.08	2.08
0	3.49	3.15
	6.33	6.00

5.82 Bond 7 is intact in ovalbumin and is split irreversibly by the enzyme. Neither bond 1 nor bond 5 are susceptible prior to the opening of 7.

The formulae are considerably more complicated in this case and will therefore not be given *in extenso*. The calculation shows however the following main facts:

1. If the substrate, I_7, formed by the cleavage of bond 7, is indistinguishable from *ovalbumin* the situation will be as in section 5.81 provided the rate of cleavage is very high. The enzyme will then initially produce the substrate for the further degradation in the concentration σ_0. If the rate of cleavage is very small the system will again approach the conditions of section 5.81. For intermediate values of the initial rate we find that $\sigma_C/\sigma_A = \alpha_5/\alpha_1$, as before, but the relations to B and to ρ are rather complex:

$$\sigma_C = \sigma_0 \beta_5 \left[1 - \rho^{\beta_7}/\beta_7 (1 - \beta_7) - \rho/(\beta_7 - 1)\right]$$

$$\sigma_A = \sigma_C \beta_1/\beta_5$$

$$\sigma_{P_I} + \sigma_{P_{II}} = \sigma_0\left[1 - \rho - (\rho^{\beta_7} - \rho)/(1 - \beta_7)\right]$$

$$\sigma_B = \sigma_0\left\{1 - \rho - \sigma_A/\sigma_0 - \beta_5\left(\frac{\rho^{\beta_1'}}{(\beta_7 - \beta_1')(1 - \beta_1')}\right.\right.$$

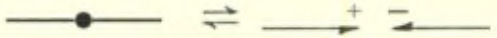
$$\left.\left. + \frac{\rho^{\beta_7}}{(\beta_1' - \beta_7)(1 - \beta_7)} + \frac{\rho}{(\beta_7 - 1)(\beta_1' - 1)}\right)\right\}$$

where $\beta_7 = \frac{\varkappa_1 + \varkappa_3}{\varkappa_7}$, $\beta_1 = \frac{\varkappa_1}{\varkappa_7}$, $\beta_1' = \frac{\varkappa_1}{\varkappa_7}$, etc.

2. If on the other hand I_7 is indistinguishable from *plakalbumin* (I or II) the whole correlation between peptide release and plakalbumin formation is disturbed unless *bond 7 is cleaved much more slowly than bond 1 and 5*, in which case the formulae of section 5.81 apply directly.

5.83 Bond 7 is intact in ovalbumin and is labilized by the enzyme without being split to any measurable extent.

We may here imagine that ΔF, the change in free energy of the reaction



is negative in the direction from right to left, e.g. due to a stress exerted by the rest of the molecule and serving to keep the bond together. If the rates of the reaction in both directions are high we may proceed as in section 5.5 and write the kinetic equation for the irreversible opening of bond 1 and 5 as follows

$$-\frac{d\sigma}{dt} = (\varkappa_1 + \varkappa_2)\,\psi \cdot \sigma \cdot e \qquad \text{(see section 4.4)}$$

where ψ is the small fraction of the molecules in which bond 7 is open. The equations in section 5.81 can therefore be directly applied substituting for $\alpha_1 = \varkappa_1/(\varkappa_1 + \varkappa_3)$ the quantity $a_1 = \varkappa_1/\psi(\varkappa_1 + \varkappa_3)$ and so on. *In this way bond 7 is « pulled apart » by the subsequent reactions*, since one can very well imagine that the reunion of the peptide ends is made impossible by the removal of shorter or larger fragments of the chain. The case is hypothetical, but not without interest.

5.84 Discussion.

The choice among these possibilities has become facilitated by recent findings of Steinberg who has made the observation that ovalbumin is not attacked by carboxypeptidase alone even in strong urea solutions where « masked » carboxyl ends should be exposed. However, a minute addition of Bacillus subtilis enzyme which by itself would produce negligible amounts of the above peptides during the experimental period is able to « activate »

carboxypeptidase so that terminal alanine is released in quantities that far exceed those of the peptides A, B and C observed here. From these interesting facts the following conclusions can be made:

1. Ovalbumin is devoid of C-terminal amino acids which can be released by carboxypeptidase.

2. The Bacillus subtilis enzyme initially attacks an internal peptide chain in ovalbumin at a bond (7) which has alanine as the carboxyl partner.

3. The rate of cleavage of this bond must be high in comparison with those of bonds 1 and 5 since otherwise the carboxypeptidase would not be « fed » with substrate sufficiently fast. This excludes that I_7 is a plakalbumin (see section 5.82). One may also consider the possibility mentioned in section 5.83 and assume that carboxypeptidase pulls bond 7 apart by removing the alanine required for the back-reaction. Preliminary experiments by Steinberg have however shown that ovalbumin, pretreated with a low concentration of Bacillus subtilis enzyme and afterwards deprived of the enzyme by addition of diisopropylfluorphosphate (Ottesen, unpublished), was able to release alanine with carboxypeptidase.

So far nothing is known with certainty about the amino partner of bond 7. According to Ottesen and collaborators the non-protein material from the Bacillus subtilis reaction always contains small quantities of other amino acids than the four species that enter into the main peptide piece. These amino acids, which are present as lower peptides, are leu or ileu, glu and ser, and they have mostly been considered as degradation products from an unspecific breakdown of plakalbumin ([83]) ([85]). It is possible however that the situation is as pictured below

$$\text{— ala — gly — val — asp — ala — ala —} \overset{7}{\underset{\uparrow}{\text{ser}}} \text{— glu — leu ——}$$

which might explain the unique properties of bond 7. It should be noted that the constellation of amino acids to the right of bond 7 is that found connected with phosphoric acid in caseinogen and also in ovalbumin [Perlmann ([86])]. An esterification of serine with phosphoric acid would unquestionably place bond 7 in a special category.

6. CONCLUDING REMARKS

It was originally intended to give in this paper a discussion of the complications introduced into the enzymatic analysis of proteins by synthetic reactions and transpeptidations. For this reason the sections 4.7 and 4.8 were written. In view of the considerable space used in discussion of other pertinent topics this original plan has been given up. The reader is referred to the above sections, and to ([87]) and ([56]) in which also some important older investigations have been reviewed.

The author is indebted to his colleagues: M. Ottesen, G. Johansen K. Max-Möller, Bodil Dahlerup-Petersen and D. Steinberg for kind help and advice.

REFERENCES

(1) Linderström-Lang, K., *Ergebn. Physiol. exper. Pharmakol.*, **35**, 415 (1933).
(2) Jacobsen, C.F., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **25**, 325 (1947).
(3) Jacobsen, C.F., *Compt. Rend. Lab. Carlsberg. sér. chim.*, **26**, 455 (1949).
(4) Riddick, J.A., *Analytical Chem.*, **24**, 41 (1952).
(5) Seaman, W., and Allen, E., *Analytical Chem.*, **23**, 592 (1951).
(6) Markunas, P.C., and Riddick, J.A., *Analytical Chem.*, **23**, 337 (1951).
(7) Tiselius, A., and Eriksson-Quensel, J.-B., *Biochem. J.*, **33**, 1752 (1939).
(8) Annetts, M., *Biochem. J.*, **30**, 1807 (1936).
(9) Williams, J.W., and Lundgren, H.P., *J. Phys. Chem.*, **43**, 989 (1939).
(10) Lundgren, H.P., *J. Biol. Chem.*, **138**, 293 (1941).
(11) Petermann, M.J., *J. Biol. Chem.*, **144**, 607 (1942).
(12) Williams, J.W., Baldwin, R.L., Saunders, W.M., and Squire, P.G., *J. Am. Chem. Soc.*, **74**, 1542 (1952).
(13) Gosting, L.J., *J. Am. Chem. Soc.*, **74**, 1548 (1952).
(14) Linderström-Lang, K., and Jacobsen, C.F., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **24**, 1 (1941).
(15) Linderstrøm-Lang, K., *Cold Spring Harbor Symp. Quant. Biol.*, **14**, 117 (1950).
(16) Linderström-Lang, K., *Lane Medical Lectures* (1952), « Proteins and Enzymes », Stanford Univ. Press, Stanford, California, U.S.A., p. 115.
(17) Aten, A.H.W., Dippel, C.J., Kenning, K.J., and van Dreven, J., *J. Colloid Sc.*, **3**, 65 (1948).
(18) Simpson, R.B., and Kauzmann, W., Personal communication.
Kauzmann, W., Simpson, R.B., Watson, M.T., Levedahl, B., Schellman, J. and Frensdorff, H.K., *Abstracts Am. Chem. Soc.*, 120th Meeting, 11 (New York, N.Y., 1951).
(19) Korsgaard Christensen, L., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **28**, 37 (1952).
(20) Brattsten, I., *Arkiv för Kemi*, **4**,, 503 (1952).
(21) Kunkel, H.G., and Tiselius, A., *J. Gen. Physiol.*, **35**, 89 (1951).
(22) Cann, J.R., and Kirkwood, J.G., *Cold Spring Harbor Symp. Quant. Biol.*, **14**, 9 (1950).
(23) Brown, R.A., Shumaker Jr., J.B., Timasheff, S.N., and Kirkwood, J.G., *J. Am. Chem. Soc.*, **74**, 460 (1952).
(24) Tiselius, A., *Adv. in Protein Chem.*, III, 67 (1947).
(25) Moore, S., and Stein, W.H., *J. Biol. Chem.*, **178**, 53 (1949).
(26) Moore, S., and Stein, W.H., *J. Biol. Chem.*, **192**, 663 (1952).
(27) Hirs, C.H.W., Stein, W.H., and Moore, S., *J. Am. Chem. Soc.*, **73**, 1893 (1951).
(28) Balston, J.N., and Talbot, B.E., *A Guide to Filter Paper and Cellulose Powder Chromatography*, H. Reeve Angel & Co. Ltd., London (1952), p. 146.
(29) Craig, L.C., Hausmann, W., Ahrens Jr., E.H., and Harfenist, E.J., *Analytical Chem.*, **23**, 1236 (1951).
(30) Gregory, J.D., and Craig, L.C., *Ann. N.Y. Acad. Sc.*, **53**, 1015 (1951).
(31) Martin, A.J.P., and Porter, R.R., *Biochem. J.*, **49**, 215 (1951).
(32) Derrien, Y., *Bull. Soc. Chim. Biol.*, **26**, 1091 (1944).
(33) Cohn, E.J., et al., *J. Am. Chem. Soc.*, **72**, 465 (1950).
(34) Sanger, F., *Biochem. J.*, **39**, 507 (1945).
(35) Green, F.C., and Kay, L.M., *Analytical Chem.*, **24**, 726 (1952).
(36) Leonis, J., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **26**, 315 (1948).
(37) Edman, P., *Acta Chem. Scand.*, **4**, 283 (1950).
(38) Fromageot, C., Jutisz, M., Meyer, D., and Penasse, L., *Biochim. et Biophys. Acta*, **6**, 283 (1950).
(39) Khorana, H.G., *J. Chem. Soc.*, 2081 (1952).
(40) Levy, A.L., *J. Chem. Soc.*, 404 (1950).
(41) Leonis, J., *Bull. Soc. Chim. Belge*, **61**, 524 (1952).
(42) Fraenkel-Conrat, H., and Fraenkel-Conrat, J., *Acta Chem. Scand.*, **5**, 1409 (1950).

(43) Ottesen, M., and Wollenberger, A., *Nature*, **170**, 801 (1952).
(44) Christensen, H.N., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **28**, 265 (1953).
(45) Dahlerup-Petersen, B., Linderström-Lang, K., and Ottesen, M., *Acta Chem. Scand.*, **6**, 1135 (1952).
(46) Mirsky, A.E., and Pauling, L., *Proc. Nat. Acad. Sc.*, Washington, **22**, 439 (1936).
(47) Pauling, L., and Corey, R.B., *Proc. Nat. Acad. Sc.*, Washington, **37**, 235, 241, 251, 256, 261, 272, 282, 729 (1951); **38**, 86 (1952).
(48) Huggins, M.L., *J. Am. Chem. Soc.*, **74**, 3963, 3963 (1952); cf. Pauling, L., and Corey, R.B., *ibid.*, 3964 (1952).
(49) Neurath, H., and Schwert, G.W., *Chem. Revs.*, **46**, 69 (1950).
(50) Bergmann, M., and Fruton, J.S., *Adv. Enzymol.*, **1**, 63 (1941).
(51) Bergmann, M., *Adv. Enzymol.*, **2**, 49 (1942).
(52) Smith, E.L., in Sumner, J.B., and Myrbäck, Ed., *The Enzymes*, **I**, part 2, 793, Academic Press Inc., Publ., New York (1951).
(53) Kuhn, W., *Ber.*, **63**, 1503 (1930).
(54) Montroll, E.W., and Simha, R., *J. Chem. Phys.*, **8**, 721 (1940).
(55) Warner, R.C., *J. Biol. Chem.*, **142**, 741 (1942).
(56) Sanger, F., *Adv. Protein Chem.*, **7**, 1 (1952).
(57) Fruton, J.S., *Symposium sur la biogénèse des protéines*, II^e Congrès International Biochim., Paris, 5-20 (1952).
(58) Jones, M.E., Hearn, W.R., Fried, M., and Fruton, J.S., *J. Biol. Chem.*, **195**, 643 (1952).
(59) Calvery, H.O., *J. Biol. Chem.*, **102**, 73 (1933).
(60) Calvery, H.O., Block, W.D., and Schock, E.D., *J. Biol. Chem.*, **113**, 21 (1936).
(61) Moring-Claesson, I., *Biochem, Biophys. Acta*, **2**, 389 (1948).
(62) Bull, H.B., and Currie, B.T., *J. Am. Chem. Soc.*, **71**, 2758 (1949).
(63) Currie, B.T., and Bull, H.B., *J. Biol. Chem.*, **193**, 29 (1951).
(64) Desnuelle, P., Rovery, M., and Bonjour, G., *Biochim. Biophys. Acta*, **5**, 116 (1950).
(65) Biserte, G., and Boulanger, P., *Compt. Rend.*, **230**, 583 (1950).
(66) Biserte, G., and Boulanger, P., *Bull. Soc. Chim. Biol.*, **32**, 601 (1950).
(67) Haugaard, G., and Roberts, R.M., *J. Am. Chem. Soc.*, **64**, 2664 (1942).
(68) Buzzell, A., and Sturtevant, J.M., *J. Am. Chem. Soc.*, **74**, 1983 (1952).
(69) Anfinsen, C.B., *J. Biol. Chem.*, **196**, 201 (1952).
(70) Miller, L., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **23**, 45 (1939).
(71) Stearn, A.E., *Adv. Enzymol.*, **9**, 25 (1949).
(72) Korsgaard Christensen, L., *Nature*, **163**, 1003 (1949).
(73) Rovery, M., Desnuelle, P., and Bonjour, G., *Biochim. Biophys. Acta*, **6**, 166 (1950).
(74) Butler, J.A.V., Dodds, E.C., Phillips, D.M.P., and Stephen, J.M.L., *Bioch. J.*, **42**, 116 (1948).
(75) Butler, J.A.V., Dodds, E.C., Phillips, D.M.P., and Stephen, J.M.L., *Bioch. J.*, **42**, 122 (1948).
(76) Northrop, J.H., Kunitz, M., and Herriott, R.M., *Crystalline Enzymes*, Columbia University Press, New York, p. 352 (1948), 2nd edit.
(77) Bailey, K., Bettelheim, F.R., Lorand, L., and Middlebrook, W.R., *Nature*, **167**, 233 (1951).
(78) Lorand, L., *Nature*, **167**, 992 (1951).
(79) Lorand, L., and Middlebrook, W.R., *Biochem. J.*, **52**, 196 (1952).
(80) Lorand, L., *Biochem. J.*, **52**, 200 (1952).
(81) Laki, K., *Science*, **114**, 435 (1951).
(82) Kowarzyk, H., *Nature*, **169**, 614 (1952).
(83) Ottesen, M., and Villee, C., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **27**, 421 (1951).
(84) Güntelberg, A.V., and Ottesen, M., *Nature*, **170**, 802 (1952).
(85) Linderström-Lang, K., and Ottesen, M., *Compt. Rend. Lab. Carlsberg, sér. chim.*, **26**, 403 (1949).
(86) Perlmann, G.E., *J. Gen. Physiol.*, **35**, 711 (1952).
(87) Linderström-Lang, K., *Erg. Physiol. & Exper. Pharmakol.*, **35**, 415 (1933).

Discussion du rapport de M. K. Linderström-Lang

M. Hermans. — In the theory of kinetics in liquid systems, one has introduced the so-called « cage effect » : once two molecules have come into contact they cannot easily get away from each other, since they are sort of caught in a cage formed by the solvent molecules surrounding them.

This effect may have some significance for the degradation of proteins, for example in the mechanism discussed by Prof. Linderström-Lang in section 4.2 of his report. The quantity β introduced there represents the ratio between reaction rates. If it is found experimentally that each protein molecule is digested practically completely once it is « denatured », this will be interpreted as meaning that β is large. However, the same experimental result would be obtained in those cases where the cage effect is operative. I should like to ask, therefore, whether the occurence of this effect would make the conclusions about the reaction rate constants dubious or whether it has already been taken into account by the introduction of the Michaelis complex.

M. Linderström-Lang. — I think that on the primitive level of my calculation the introduction of $\varkappa$, does take the « cage effect » into consideration. The conditions are that the configuration of the enzyme substrate complex in the « denaturing » process is different from that in the hydrolytic reaction, and furthermore that the energy released in the reaction is dissipated and not led specifically to subsequent centers of attack.

M. Tiselius. — 1) The early experiments of Eriksson-Quensel and myself (1939) were, of course, done at a time when we did not have the methods now available to study split products of proteins, and too much stress should not be given to our inability to find any intermediate split products, which is the basis of the hypothesis of the « all or none » reaction. Nevertheless, we are pleased that our results have been confirmed by other authors, using other methods.

It is obvious from Prof. Linderström-Lang's paper that proteins are quite individual in this respect. Thus in some cases, intermediate fragments would be possible to isolate, and several speakers

earlier during this conference have emphasized the importance of such attempts with regard to protein structure.

I would like to point out to the possibility of isolating such fragments by removing as efficiently as possible during the digestion, the primary digestion products from further contact with the enzyme. This might be done by some suitable partition system, as the intact protein and the enzyme could certainly remain in the aqueous phase, whereas with a suitable second phase, the split products could concentrate there. Of course, dialysis would also work, but dialysis may be rather slow for these comparatively large particles.

2) In order to study the « all or none » digestion, *active* proteins would be of particular interest. Thus it would be very valuable to follow, in parallel with other determinations, the successive loss of activity of, for example, ribonuclease, during digestion.

M. Synge. — Do the quantities of terminal alanine released by carboxypeptidase from egg albumin after action of *Bacillus subtilis* enzyme (p. 293), exceed those that could be literated from one locus (of the type postulated on p. 290) per egg albumin molecule ?

M. Linderström-Lang. — No, but in recent experiments Sternberg has isolated 0.7 moles of alanine per ovalbumine molecule.

M. Neuberger. — Is the term « denaturation » to describe the first step of the peptic hydrolysis of native egg albumine really justified ? In other words, denaturation as the term is ordinarily used, does not imply a splitting of peptide linkages, whilst the first step of peptic hydrolysis may be truly hydrolytic, an unfolding being only a necessary consequence of this first reaction. This would mean that « denaturation » follows the first hydrolytic step.

M. Linderström-Lang. — Inasmuch as a protein in which a few peptide bonds are opened still may possess all the properties characteristic of a native protein (e. g. plakalbumin) I do not see why the word « denaturation » cannot be used for describing the loss of nativity of such a partly degraded protein.

M. Havinga. — 1. In the case of the tryptic breakdown of β-lactoglobuline, the experiments suggest that for the initial reaction the energy of activation is very low. Therefore the entropy of activation should have a high negative value. This means that

the transition state for this reaction has a very low entropy. It seems doubtful whether a better insight is gained in explaining this interesting feature by assuming an equilibrium between two forms of β-lactoglobuline. This tends to detract attention from the really remarkable fact of the very low entropy in the transition state in this case. If one does not follow this line of thought, however (e. g. on account of certain experimental peculiarities), it might be advisable to stress not only the low energy of the « reacting form », but especially also its low entropy; implying that this form has a very low a priori probability.

2. In relation to the suggestion that the special features of bond 7 might be due to the participation of phosphorylated serine in this bond, the question arises whether dephosphorylated ovalbumin is transformed enzymatically into a corresponding plakalbumin in the same way as phosphorus-containing ovalbumin.

M. Linderström-Lang. — The picture given was formed on the basis of experiments on urea-denaturation where a strong negative temperature coefficient was observed. Here the simplest explanation is that based on the equilibrium of two states of β-lactoglobulin. The whole question needs however a more thorough theoretical and experimental investigation.

M. Putzeys. — Dans son rapport Dr. Linderström-Lang dit que la globine de l'hémoglobine de cheval (p. 278) est une protéine dénaturée, ayant un poids moléculaire de 64.000. Il est certain que le poids moléculaire moyen de cette globine dénaturée ne correspond pas à ce poids moléculaire. J. Roche, A. Roche, G. S. Adair et M. E. Adair [*Biochem. J.*, **26**, 1811 (1932)] ont trouvé par pression osmotique un poids moléculaire d'environ la moitié. Par la diffusion moléculaire de la lumière, il est facile de démontrer que les solutions de globine dénaturée sont fortement polydispersées, et que certaines particules atteignent des dimensions considérables. Cette polydispersité affecte-t-elle le résultat des calculs du degré d'hydrolyse enzymatique de la globine dénaturée ?

M. Linderström-Lang. — I do not think that my very coarse calculations will be changed at all by changing the aggregatweight in one or the other direction. The essential experimental basis is the length of the individual chains (Mw : 11 000-13 000).

Some Relations Between Prosthetic Groups and Proteins

by Hugo Theorell

HISTORICAL INTRODUCTION

Modern protein chemistry was born in the year 1926, when The Svedberg and his collaborators constructed the first ultracentrifuge and James B. Sumner crystallized the first enzyme, urease. Before that time proteins were regarded as « colloids » of variable particle sizes changing in shape and hydratation with the pH, electrolytes, temperature, and other conditions in solution. Now from 1927 on, could by the aid of gravitational fields exceeding 100,000 g, Svedberg and his collaborators (1) demonstrate that proteins in solution have molecules of uniform sizes and shapes that were, within rather wide limits, independent of pH and salt concentration. This remarkable series of investigation laid the foundation for the hope that the chemical structure of proteins could once be elucidated in all detail, since it appeared now probable that proteins existed in a chemically pure state. Now after only twentyfive years it appears that the structure of a protein — I am of course thinking of the insulin — may be known in the not too distant future.

The importance of Sumner's crystallization of urease (2) and J.H. Northrop's of pepsin (1930), and Northrop's and Kunitz's of trypsin and chymotrypsin 1930-31 (3) was not realized immediately, because it was very difficult to prove that the crystallized proteins were strictly pure. Thus it was not known whether the crystallized proteins were the enzymes themselves, or merely carriers of the enzymes, which could possibly be present as an impurity of unknown, perhaps very low concentration. The extensive work of Willstätter and his collaborators on the purification of peroxidase, saccharase, lipase and amylase had lead to a negative result; in a lecture before the *Deutsche Chemische Gesellschaft*, in 1926, Willstätter concluded that the enzymes did not belong to the groups of proteins or carbohydrates, or to any known group of

complicated organic compounds. Willstätter, in his memoirs « Aus meinen Leben » (p. 361) tells that R. Kuhn, at this time, pointed out that the « enormous » degree of purification already reached by Willstätter in his enzyme preparations did not prove that even a considerable part of one per cent of these preparations would consist of the pure enzyme. The great difficulty was that no active group was found — and still is not found — in the hydrolytic enzymes that were crystallized first. The analyses revealed nothing but the same amino acids known from inactive proteins. Therefore the discussion on the purity of Sumner's and Northrop's enzyme preparations persisted for years and many attempts were made to achieve further purification. In all probability the discussion would never have come to an end, in spite of Northrop's beautiful homogeneity tests on his crystallized preparations, if additional evidence from another group of enzymes — the oxidation enzymes — had not been presented.

Warburg's and Keilin's work on the respiratory pigment and the cytochromes, in the twenties, had led to the conclusion that these enzymes contained hematin as prosthetic group linked to a carrier of high molecular weight. Zeile and Hellström [1930, (4)] and Kuhn, Hand and Florkin [1931, (5)] presented experimental evidence for catalase and peroxidase being built up in an analogous manner. But in none of these cases could the chemical nature of the high molecular weight carrier be definitely established, because pure preparations were not obtained. Warburg's and Christian's « yellow ferment » was the first oxidation enzyme to be crystallized, by Theorell in 1934 (6). I separated the yellow, prosthetic group from the protein, found it to be a monophosphoric acid ester of riboflavin and crystallized it as calcium salt. The riboflavinphosphate, now called FMN, was the first of the coenzymes that was produced in a pure form.

The separation of FMN from protein could be carried out in such a mild manner — dialysis against dilute hydrochloric acid — that the protein component was not irreversibly denatured. The free protein was found to be inactive in our enzymatic test: the activity reappeared, however, when riboflavin phosphate was added. The yellow ferment could thus be split into two parts and resynthesized. The resynthesized ferment was identical with the original ferment in activity tests, electrophoretic mobility and sedimentation properties in the ultracentrifuge. One molecule of the prosthetic

group was united with one molecule of the protein. This simple stoichiometrical relation between protein and prosthetic group proved that in this case the « enzyme » could not possibly be present as a minor impurity; thus these experiments gave the definite demonstration of an enzyme being a protein. It must be remembered that before the isolation of the yellow enzyme we could not know that it was a protein. Warburg and Christian's (7) preparation consisted largely of yeast polysaccharides, and Kuhn et al. (8) therefore supposed the existence of « flavopolysaccharides »; and for those who were not convinced of the homogeneity of Sumner's and Northrop's enzyme preparations it appeared that no conclusive evidence had been presented to prove that enzymes were proteins. Therefore I may be justified in saying that the isolation and reversible splitting of « the yellow enzyme » became of general importance for enzyme chemistry.

We know now with a high degree of probability that all enzymes are proteins, and that a great part of the proteins, but not all of them, are enzymes.

Nomenclature. — At present time there is a great deal of confusion in the use of terms like « prosthetic group », « coenzyme », « active group » and « essential group ». It might be of value to make an attempt to establish some definitions in this field.

Essential groups. — It is more or less self-evident that all of the many atoms and groups in the large enzyme molecule cannot be of the same importance for the activity. In fact, probably most of the groups are « non essential », or inert. All the free lysine amino groups in pepsin can be acetylated, and it is possible to remove considerable parts of the chymotrypsin molecule without loss of activity in either of the cases. There may exist « partly essential » groups, too, in both protein and prosthetic groups modifying the reaction velocity without being indispensable, as for instance the vinyl groups in the protohaematin of the horse radish peroxidase. A group in an enzyme may be « essential » for different reasons: 1) it may serve to bind prosthetic groups, coenzymes or substrates to the protein, and in that case suitably be called « *conjugating group* » (in German « Haltgruppen »). Divalent metals, such as Mg, Mn or Co sometimes serve as conjugating groups between enzyme and substrate, for instance arginase, glutaminase, phosphatases and peptidases. Or 2) the

« essential group » may be an *active group*, taking part in the enzymatic reaction cycles; or 3) it may be a group closely located to conjugating or active groups and influencing their properties, for instance through their electrical charge, without being directly attached to the conjugating or the active groups. This last type of essential groups are included in what Pauling calls *« linked » groups.* This term is frequently used in hemoprotein chemistry. Wyman (9) has given some instructive examples of linked functions: « Whenever a molecule possesses two or more different functions, e.g. dissociation of protons and combination with oxygen, belonging to nearby groups in the molecule, there is the likelihood of an interdependence of the function due to interaction between the groups. When this occurs, we may conveniently speak of the functions as linked and refer to the groups as linked groups ». Wyman then mentions the linking in hemoglobin, between imidazole groups and iron, where the attachment of oxygen to iron causes a displacement of the acid dissociation constant of the imidazole, and correspondingly the addition of a proton to the acid group causes a decrease in the oxygen affinity of the iron. The imidazole may thus be regarded as an oxygen-linked acid group, the oxygen combining iron as an acid-linked, or proton-linked group. « Such groups in hemoglobin are accountable for the well-known Bohr effect. To take another example from hemoglobin, there is also a pH-dependence of the reaction with fluoride, which leads us to speak of fluoride-linked acid groups and acid-linked fluoride-binding groups. Likewise we may speak of oxidation-linked acid groups and acid-linked oxidizable groups ». The reactions of carbon monoxide or oxygen with hemoglobin iron, which are competitive and exclude one another, should, according to Wyman, be called « identically linked ».

Prosthetic groups and coenzymes. — As « prosthetic » means « predominating » the term « prosthetic » should be reserved for groups that have a dominating influence upon the whole character of the compound formed with proteins, as the heme, or hematin, in the hemoproteins, the flavinnucleotides, the pyridine nucleotides thiaminpyrophosphate, pyridoxalphosphate, and maybe, coenzyme A. It is common to call compounds like the last mentionned, « coenzymes », but heme, or hematin is never named « coenzyme ». There is a certain reason for this nomenclature since « heme » is the prosthetic group in hemoglobin and myoglobin, which are

no enzymes; and the hemoproteins do not dissociate to any appreciable degree under physiological conditions into heme, or hematin and protein, and, unlike the coenzyme-proteins, can never be separated into prosthetic group and protein simply by dialysis or heat denaturation.

However, the dissociation constant of different coenzyme-enzyme compounds vary considerably, and as will be described later, this is of decisive importance for their mode of action. Parnas ([10]) has suggested a distinction between « mobile » and « fixed » coenzymes, and gives the nicotinamide coenzymes as an example of the former type. In many, but not all, cases these coenzymes (« DPN », and « TPN ») alternate between different proteins (« apoenzymes »), being reduced when they are combined with one protein, reoxidized when combined with another. The « mobile » mechanism can thus be formulated as follows:

$$E_1 + DPN \rightleftarrows E_1 \cdot DPN \tag{1}$$

$$E_1 \cdot DPN + AH_2 \rightleftarrows E_1 \cdot DPNH + H^+ + A \tag{2}$$

$$E_1 \cdot DPNH \rightleftarrows E_1 + DPNH \tag{3}$$

$$E_2 + DPNH \rightleftarrows E_2 \cdot DPNH \tag{4}$$

$$E_2 \cdot DPNH + H^+ + B \rightleftarrows E_2 \cdot DPN + BH_2 \tag{5}$$

$$E_2 \cdot DPN \rightleftarrows E_2 + DPN \tag{6}$$

where E_1 and E_2 are different enzymes, A and B donors and acceptors of hydrogen.

The « fixed » coenzyme mechanism, as will be discussed more in detail later in this lecture, is found to operate in all flavin- and hemin enzymes, and may be schematically formulated for a flavoprotein as follows (without regard to Michaelis' one-step mechanism) :

$$E \cdot FMN + AH_2 = E \cdot FMNH_2 + A \tag{7}$$

$$E \cdot FMNH_2 + B = E \cdot FMN + BH_2 \tag{8}$$

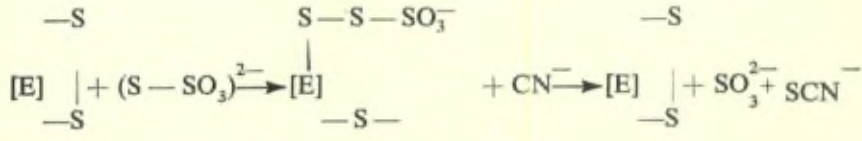
The difference between the two mechanisms is obviously profound enough to justify different names for the « mobile » and « fixed » groups, as suggested by Parnas. However, it should be pointed out that it would seem entirely logical to call the « fixed coenzymes » *prosthetic groups*, and reserve the word « *coenzyme* » for the mobile coenzymes.

Of course, it may at first sight appear complicated to call the « fixed » DPN in the phosphoglyceraldehyde dehydrogenase « prosthetic group », the « mobile » DPN in alcohol dehydrogenase, « coenzyme », but the advantage is that the different names account for the different modes of function. From a practical point of view it is rather simple to establish an experimental limit between the two classes of groups if we call those that can be separated from the apoenzyme by dialysis against water, *coenzymes*, those that cannot, *prosthetic groups*. This would be in historical agreement with Bertrand's discovery that certain enzymes require dialyzable substances for their activity, called by him « coenzymes ». The distinction between coenzymes and prosthetic groups based upon their dialyzability will on the whole turn out to coincide with the functional difference between « mobile » and « fixed » coenzymes.

The terms « *active groups* » or « *active centers* » should be distinguished from « prosthetic groups » or « coenzymes ». It may be of advantage to designate those atoms (or the atom) in a « holoenzyme » *where the cyclic enzyme reaction finds place* as « active groups », for example the pyridine ring in DPN and TPN, the oxido-reducible ring system in the flavinnucleotides, the iron atom in hemoproteins.

According to this active groups can be parts of the prosthetic groups in enzymes containing prosthetic groups; but the great many enzymes lacking prosthetic group, i.e. the hydrolytic enzymes, must nevertheless be considered to contain active groups, probably formed by ordinary amino acids in some suitable arrangement. Recent research on polypeptide hormones and antibiotics has revealed that the ordinary amino acids in special peptide combinations can form substances of surprisingly strong biological effects.

The activity of many enzymes, in some way or another, depends upon sulfhydryl or disulfide groups [for a summary, see Barron (11)]. In some of the enzymes these groups should be classified as active groups, taking part in the enzyme reaction itself. This is particularly clear in the case of rhodanese, where Sörbo (12) has found a disulfide group working as active group in the formation of rhodanide from thiosulfate and cyanide.



We have now arrived at a picture of the biological functions of proteins totally different from the one we had twenty-five years ago. There are a few proteins still that may serve only a nutritive purpose, being broken down to peptides and amino acids in the intestine, absorbed and rebuilt to other proteins. Casein may be a protein of this kind, though it must be remembered that casein is probably specially adapted to this apparently simple function by having its amino acids arranged in a way to give valuable peptide fragments in the course of digestion in the intestine. It is supposed that these fragments may promote the resorption of metal ions like Ca^{++} or Fe^{++}([12a]), and that the peptides perhaps are synthesized as such into new proteins.

However, it is now clear that of the proteins in the cell protoplasm if not all at least the greater part of them are enzymes.

Thus, since the majority of all proteins are built to serve specific activities, the problem of the chemical and functional relations between proteins and their active groups is of very great importance. Our present knowledge in this field is today very poor and scattered, but still far too large to be covered in one lecture. For that reason I shall confine myself to a discussion of some of the information available at present concerning the relations between the constituents of conjugated proteins. These are the easiest to study for manifold reasons: the prosthetic groups and coenzymes can be separated from the proteins, and the chemical constitution of these low molecular constituents has been determined in most cases. This is a great advantage, as compared with the non-conjugated enzymes, where the total structure has not yet been determined in any single case.

The characteristic absorption spectra of such prosthetic groups as hematin, flavin- or pyridinnucleotides is very favourable for spectrophotometric studies. Since they are oxido-reduction enzymes, measurements of the redox potential at different pH-values gives a good deal of information. In the case of hemoproteins magnetometric studies have proved to be very valuable.

The possibility of splitting proteins from their prosthetic groups and recombining the parts again gives further possibilities to study the nature of the conjugating groups by for instance titrations of the free protein and the recombined holoenzyme; blockade of conjugating groups by « inhibitors »; changes in the structure of the prosthetic group; and so on.

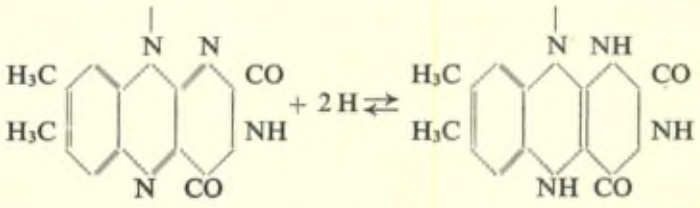
Only within three groups of conjugated proteins, flavin-, pyridine-, and hemoproteins, is enough information available concerning the chemical mode of junction between prosthetic groups and proteins and their functional relations, to justify a discussion at present time.

In the following we shall focus our attention on these classes.

THE FLAVOPROTEINS

All flavoproteins isolated so far contain as prosthetic group either flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD).

The active group is the isoalloxazine part that can be reduced to its leuco form and reoxidized again:



However, the oxidoreduction in free FMN or FAD is sluggish and does not seem to play any biological role. The combination with specific proteins: 1) accelerates the oxidoreduction reactions, and 2) adapts the holoenzyme to react more or less specifically with certain oxidants and reductants. How the proteins exert this action is the fundamental problem of enzyme chemistry. « Warum das Eiweiss so wirkt, ist heute das Problem der Fermentchemie » (Warburg). It is obvious that the specific proteins have a manifold influence on the reactivity of the prosthetic group. We know that the oxidoreduction potential is greatly influenced by the linkage to protein. This is discussed later in more detail in connection with the equilibrium between DPN and alcohol dehydrogenase. In the « old yellow ferment » the redox potential is increased from the value of — 0.185 V to — 0.06 V by the coupling to protein [Kuhn and Boulanger (13)]. Of course this great change must be of importance for the function of the enzyme, but it does not give the whole explanation of the enzymatic properties of the holoenzyme. Very probably the linkage of prosthetic group to protein

causes a decrease in activation energy, and an increase in the ability of the flavin to form monohydroradicals (Michaelis one-step mechanism). Furthermore, in some way conjugating groups may be assumed to link the oxidizing and reducing substrates to a suitable site of the enzyme surface to let them react easily and quickly with the flavin. All these complicated relations of course cannot be fully understood before the structure of the protein parts has been elucidated. At present time we can point out the active and the conjugating groups in most of the prosthetic groups, and in some few cases the conjugating groups of the proteins. In most cases, however, we are still confined to guesses on the protein side.

a) Dissociation constants.

There is a very good reason to believe that FMN and FAD are linked to proteins by their phosphoric acid groups: riboflavin seems to give a highly dissociated compound with the protein of the old yellow ferment, because Kuhn and Weygand (14) observed a weak catalytic activity in a mixture of riboflavin and the protein. The riboflavinphosphoric acid, FMN, on the contrary gives a conjugated protein with very low dissociation constant $K_{\text{Protein, FMN}} = 0.6 \times 10^{-7}$ according to Haas, Harrer and Hogness (15). This dissociation constant, « K_{ox} », has been determined for some other flavoproteins and found to vary from 2.5×10^{-7}(*) for the FAD protein « d-amino acid oxidase » (16) to 10^{-9} for the FMN protein « cytochrome *c* reductase » from yeast (17). These values were obtained for the oxidized form; the reduced form may show quite different values for the dissociation constant « K_{red} ». The reduced old yellow ferment would be expected to show a K_{red} of around 0.5×10^{-11}, based on the redox potential of the ferment being 125 mv higher than for free riboflavin, and according to the formula (17) (18)

$$E'_o \text{ complex} = E''_o \text{ free flavin} + 0.030 \log \frac{K_{ox}}{K_{red}}$$

Already from this very considerable difference between K_{ox} and

(*) It should be mentioned, however, that this figure is correct only under the assumption, not warranted by the experiments of W. and C. that the oxidizer, oxygen, was present in high enough concentration to keep the enzyme essentially in the oxidized state. Hellerman et al. (19) found around twice as high values.

K_{red} one may draw the conclusion that a direct linkage between the protein and the isoalloxazine ring system must exist in the yellow ferment, because bonds only between protein and the phosphoric acid residue could not possibly be strongly influenced by the oxidation and reduction of the prosthetic group.

Are *FMN and FAD « fixed » or « mobile »*? This question can be answered with a high degree of probability in favour of the first alternative for those flavoproteins whose dissociation constants are known hitherto. This constant, K, is a function of the association and dissociation velocities:

$$A + B \underset{k_2}{\overset{k_1}{\rightleftarrows}} AB$$

$$K = \frac{k_2}{k_1}$$

Since neither k_1 nor k_2 is known for any flavoprotein — though k_1 could easily have been determined, e.g. by aid of the disappearance of fluorescence — the dissociation constant, K, in itself does not give any information about the magnitude of k_1 or k_2. However, the calculated dissociation constant for the reduced form of the old yellow ferment, $K_{red} \sim 10^{-11}$, is so low that a « mobile » coenzyme mechanism would require improbably high values for both k_1 and k_2; of course k_1 is limited by the total number of collisions per second between prosthetic group and protein molecules. Even the highest dissociation constant so far determined for a flavoprotein, $K_{ox} = 2.5 \times 10^{-7}$ for d-amino acid oxidase, is far too low compared to the high turnover number (1 440 mol 0_2/minute $\times$ mol) to be compatible with a « mobile » coenzyme mechanism. Warburg and Christian (16) made an experiment that directly proves this: a small quantity of d-amino acid oxidase protein was mixed with an excess of FAD and d-alanine under anaerobic conditions. Practically no reduction of FAD was observed. With equimolar proportions of FAD and protein plus excess of d-alanine, on the contrary, reduction of FAD finds place. The obvious explanation was that in the first case the reaction stops as soon as the small amount of FAD-protein was reduced to $FADH_2$-protein, in which the dissociation rate (k_2) according to the experiment must be low. We may draw the conclusion that *FMN and FAD remain attached to their proteins all through the oxidoreduction cycles* and thus are

to be regarded as prosthetic groups rather than coenzymes. As will be shown later, the pyridine nucleotides in most cases are mobile. This means that when hydrogen has to be transferred from reduced pyridine nucleotide to isoalloxazine the former has to dissociate form its specific protein and combine with the flavoprotein. Early experiments by Haas in Warburg's institute seemed to demonstrate such a mechanism. However, since that time no more work seems to have been done along this line.

b) Absorption bands.

The coupling between protein and FMN or FAD causes a slight displacement of the absorption bands in the blue and the near ultraviolet, as seen from table 1, where some examples are collected.

Table 1

Flavin Compound	Prosthetic group	Absorption maxima mμ
Riboflavin . . . . . (20)		365,445
FMN . . . . . (21)		365,445
FAD . . . . . (16)		375,450
Old yellow ferment . . . . . (22)	FMN	380,465
d-amino acid oxidase . . . . . (16)	FAD	380,450
Haas yeast enzyme . . . . . (23)	FAD	377,455
Straubs diaphorase (*) . . . . . (24)	FAD	359,451
Cytochrome c-reductase, yeast . . . . (25)	FMN	380,455
Notatin . . . . . (26)	FAD	377,455
Vernons DPN-cytochrone c-reductase . . (27)		
Ophio-1-aminoacid oxidase . . . . . (28)	FAD	389,465

(*) Fluorescent.

The coupling to protein in most cases causes the bands to shift towards higher wavelengths. Warburg (29) took this as an indication that the activation energy of the alloxazin would be lowered, the reactive ability increased by the coupling to protein. A recent example, the DPN-cytochrome c-reductase of Vernon (27), however, shows that the bands may move the opposite way. An analogous case had been found earlier by us : the band of DPNH moves to shorter wavelengths when it is coupled to the enzyme ADH, see below page 318. The reactivity can thus be enhanced by proteins

even if the absorption maxima are displaced to shorter wavelengths.

Nevertheless, these band shifts give another strong indication of a coupling between protein and the isoalloxazine ring.

c) Fluorescense.

The strongest evidence for this type of coupling, however, comes from the fact that the greenish fiuorescence of free FMN and FAD are in most cases absent in the holoenzymes. The only known exception is Straub's diaphorase. Kuhn and Moruzzi (30) found the fluorescence of lumiflavin and riboflavin to be abolished in acid and alkaline solution; the disappearance followed dissociation curves with pH = 1.7 and 10.2. Bessey et al. ([31]) studied the fluorescence of riboflavin, FMN and FAD at different pH-values and found that FAD gives only 9% as strong fluorescence as riboflavin and FMN around neutral reaction. Around pH 3 FAD, however, gave a much stronger fluorescence, approaching the values for FMN. Weber ([32]) found that the quenching of fluorescence in FAD was caused by the formation of an internal complex between the isoalloxazine part and the amino group of the adenine; the complex is non-fluorescent and exists in equilibrium with it fluorescent form; the equilibrium is dependent on pH. Weber showed that the value of the acid pK ($\sim$ 2) obtained by Kuhn and Moruzzi by measuring the fluorescence is not the true dissociation constant, because the mean life time of the excited molecules is shortened at this pH. The value pK = — 0.2 determined spectrophotometrically by Michaelis et al. ([33]) is very probably the more correct one. Kuhn's pK = 10.2 in the alkaline range was, however, confirmed.

This is of considerable importance, since Kuhn and Boulanger ([13]) based their well-known assumption of a bond between the protein and the 3-imino group of the isoalloxazine ring on a pK value of 10.2 being likely for this imino group. Such a bond would be expected to quench the fluorescence just in the same way as this disappears on the addition of alkali. Weber suggests that « aromatic groups in the protein (tyrosine) » are the most likely to act as quenchers. However, another explanation is possible when FAD is the prosthetic group; union with the protein might simply prevent the complex from reverting to the fluorescent form in notatin, while

for diaphorase (the fluorescent enzyme of Straub) « the ease of this transition would not be affected ».

The disappearance of fluorescence upon coupling of prosthetic group with protein offers a very spectacular method of following the rate of this reaction as shown by Theorell in 1937 (34). The amount of active protein from « old yellow ferment » could even be titrated with FMN, since the fluorescence of FMN added to an excess of the protein disappears instantaneously. Today it should be possible to gain much valuable information on the velocity constant of the holoenzyme formation, using the rapid flow methods of B. Chance. However, such work has not yet been reported.

In the work of 1937 we observed that the free protein of the old yellow ferment was remarkably sensitive to even slightly elevated temperatures, as 38° C. At pH-values around 5 after a few minutes at 38° C, with subsequent cooling to 20°, a solution of the protein had entirely lost its coupling capacity, but this returned gradually to a great extent, around 60%, in 30-60 minutes. The explanation of this peculiar behaviour is still obscure. Sulfhydryl groups, reacting with nitroprussid, do not appear in this inactivated state; one might imagine that the warming causes a temporary deformation of the protein molecule, preventing the conjugating groups in prosthetic group and protein from fitting to one another.

d) Inhibitors.

Valuable information on the nature of the different groups responsible for enzyme activity may be obtained from experiments with inhibitors. This is of course particularly true for hemoproteins and other metalloproteins.

Greater difficulties are involved in the interpretation of the effect of inhibitors on metal-free, conjugated enzymes, because the inhibitor may interfere either with active or conjugating groups in the prosthetic group or the protein, or with the groups serving to attach the substrate on the enzyme surface.

Systematic research on inhibitors of flavin enzymes has been carried out by Hellerman and his collaborators (35). Their work on the inhibitors of d-amino acid oxidase has shed some light upon

different types of inhibition that may find place in any of the partial reactions :

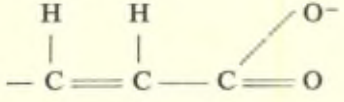
$$\text{FAD} : \text{P} \rightleftharpoons \text{FAD} + \text{Protein (P)} \quad \ldots\ldots\ldots\ldots\ldots \quad (1)$$

$$\text{FAD} : \text{P} : \text{CH}_3\text{—}\underset{\displaystyle \text{NH}_2}{\overset{\displaystyle \text{H}}{\underset{|}{\overset{|}{\text{C}}}}}\text{—COO}^- \rightleftharpoons \text{FADH}_2 : \text{P} + \text{CH}_3\text{—}\underset{\displaystyle \text{NH}}{\overset{}{\underset{\|}{\text{C}}}}\text{—COO}^- \quad (2)$$

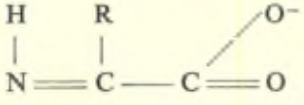
$$\text{FAD} : \text{P} : \text{substrate} \rightleftharpoons \text{FAD} : \text{P} + \text{substrate} \quad \ldots\ldots\ldots \quad (3)$$

After Haas ([36]) had found that the aromatic base quinacrine strongly inhibited the isolated cytochrome C— reductase from yeast, Hellerman et al. ([19]) found that the members of several different groups of basic aromatic compounds like quinolines, acridines, anilines, etc., inhibit d-amino acid oxidase by competition with FAD for the protein in reaction (1). With some of the compounds the inhibition was completely, with others only partially reversible.

No inhibitors are known for reaction (2). The reaction (3) is competitively inhibited by many organic acids possessing the grouping



This is not surprising, when we consider the similar constitution of the reaction products of oxidation, the iminoketonic acids



Sulfhydryl groups are essential in some of the flavoproteins like d-amino acid oxidase and xanthine oxidase, because strong inhibition occurs through the addition of the mercaptide reagent p-chloromercuribenzoate. Whether the sulhydryl group is concerned with the combination of flavonucleotide or substrate with the protein or has some other function remains to be determined.

THE PYRIDINE NUCLEOTIDE PROTEINS

At the time, in 1934, when TPN was first isolated in Warburg's laboratory, the combination of flavin phosphate with protein to form the old yellow ferment had already been demonstrated. It then appeared probable by analogy that the « Zwischenferment », an impure protein fraction from yeast, necessary for the catalytic function of DPN, could contain the specific protein component for forming an active, conjugated enzyme with the TPN. « Es lag von vornherein gewissermassen auf der Hand anzunehmen, dass das Zwischenferment, das ja ein Eiweisskörper ist, sich mit dem Co-Ferment verbinden könnte und so als ein kolloider Träger fuer das Co-Ferment wirke » [Theorell ([37])]. My attempts to demonstrate a combination between TPN and protein by determining the electrophoretic mobility in mixtures of DPN and Zwischenferment compound failed, probably because of too high dissociation of the cmpounds, but the existence of DPN– or TPN– proteins was soon afterwards proved in different ways.

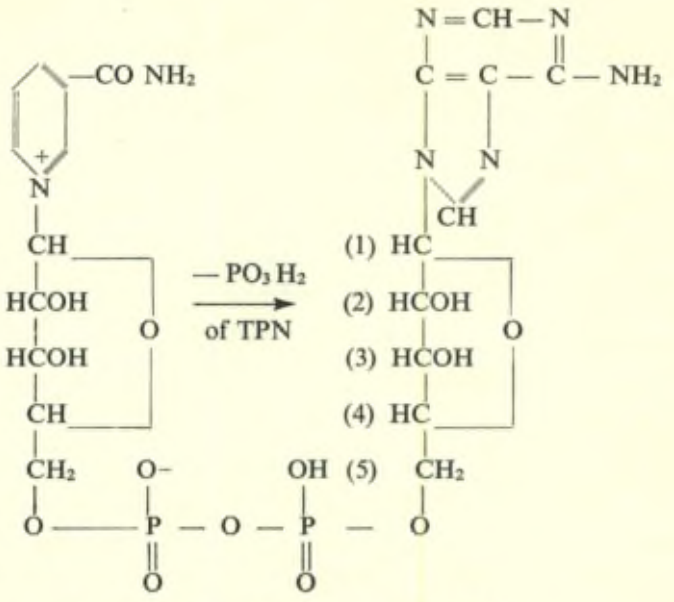
It might be of interest to recall that our early electrophoresis experiments on TPN ([38]) before either TPN or DPN were prepared pure demonstrated with a high degree of probability that TPN was a phosphoric acid ester compound, since the migration velocity of the activity — which is independent of the purity — followed a curve for different pH-values that gave two acid dissociation constants of pK = 1.8 and 6.2, typical for phosphoric acid esters.

The further development after Warburg's discovery of the nicotinic acid amide as being the active group is too well known to be repeated here.

The pyridine nucleotides DPN and TPN have, as is evident from their constitution, manifold possibilities to conjugate with proteins.

The position of the third phosphoric acid group in TPN is still somewhat doubtful. According to recent results by Kornberg and Pricer ([39]) and Kaplan, Colowick and Ciotti ([40]) it may be esterified to carbon atom 2 of the ribose next to the adenine.

Let us first consider whether the adenosine part serves as conjugating group to the protein. Williams ([41]) started from the observation that feeding adenine to dogs could provoke « black tongue », a vitamin deficiency associated, among others, with lack of niacin amide. He therefore investigated the influence of additions of adenine, adenosine, or ATP on the reaction velocity in the



Disphosphopyridine nucleotide, DPN

isolated malic dehydrogenase — DPN system. It was found that all three of the adenine compounds had a marked inhibitory effect of the competitive type. Since the results with adenine and adenosine were rather similar ($Km = 0.032$ M resp. 0.046 M) it was concluded that conjugation with the protein occurred through the adenine part and not through the ribose that rather lowered the « affinity ». Some stabilization was apparently caused by the phosphoric acid group in ATP ($Km = 0.016$) but the difference was not considered great enough for drawing the conclusion that phosphoric acid residues are conjugated to the protein in the case of DPN.

According to Pullman, Colowick and Kaplan (42) the amino group of the adenine seems to play a role that we would, according to our therminology, call « partly essential ». Kaplan, Colowick and Ciotti (40) found a deaminase in takadiastase that was active on 5-adenylic acid, adenylic acid b (3-adenylic acid), DPN, adeninediphosphate ribose, ADP and adenosine, but not on adenine, TPN or adenylic acid a (= 2-adenylic acid?). TPN can be dephosphorylated by prosthata phosphatase to give DPN. Since enzymes exist that can transform both DPN and TPN to the corresponding

deamino (inosine-) derivatives, it seemed possible that deamino-DPN or —TPN are of physiological importance. Therefore the activity of DPN and deamino-DPN were compared in a number of enzyme systems. With alcohol dehydrogenase from liver, cytochrome C— reductase from rabbit heart and pyridine nucletide transdehydrogenase from Ps. fluorescens DPN and deamino-DPN worked equally well. In the following enzyme systems the deamino-DPN showed a diminishing activity compared with DPN: formic dehydrogenase from peas, alcohol dehydrogenase from yeast, acetaldehyde dehydrogenase from liver, triosephosphate dehydrogenase from rabbit muscle, and β-oxybutyric acid dehydrogenase from rat liver.

The lactic dehydrogenase from pig heart worked better with DPN in the oxidation of lactic to pyruvic acid, but better with deamino-DPN in the reverse reaction. The authors assume the cases of lower activity for deamino-DPN to depend upon lower affinity of deamino-DPN for the protein, or lower activity (turnover number) of the holoenzyme. However, as we will see later (p. 324), the *higher* activity of deamino-DPN in the reaction.

$$\text{Pyruvic acid} + 2\,H \rightleftharpoons \text{lactic acid}$$

can just as well depend upon a *lower* affinity of deamino-DPN for the protein.

To summarize, the adenine part of the phosphopyridine nucleotides at least sometimes takes part in the conjugation to protein, while in other cases its primary amino group is involved.

By analogy with the flavin nucleotides we would expect the phosphoric acid residues to be linked to the proteins, too. This seems to be the case, since phosphate causes a reversible inhibition in the TPN— « Zwischen ferment » system (42a).

There are different reasons to believe that the nicotinic acid amide besides being the active group is conjugated directly to the protein.

Feigelson, Williams and Elvehjem (43) found free niacin-amide to act as a competitive inhibitor for DPN both in liver homogenates and in the isolated malic dehydrogenase system. Rather high concentrations had to be used, because the Michaelis constant for the niacin amide was 3,100 times as high as for DPN in the malic dehydrogenase system.

Alivisatos and Denstedt (44) have extended these observations to lactic dehydrogenase + DPN and glucose-6-phosphate dehydrogenase + TPN with similar results.

Our studies on the liver alcohol dehydrogenase (ADH) + DPN (17) (18) gave results that cannot be explained without assuming direct bonds between the niacin amide part of DPN and the protein.

The light absorption maximum at 340 mμ for free DPNH was shifted to 320-325 mμ in the compound with ADH (see fig. 1).

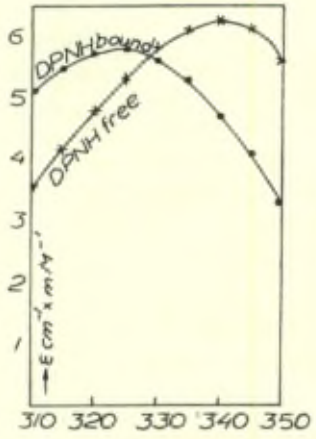
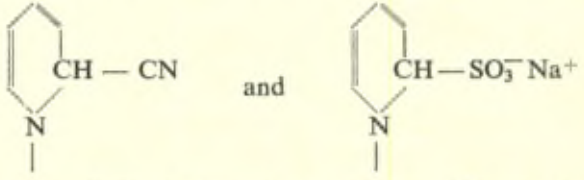


Fig. 1
Millimolar extinction of free and ADH-bound DPNH.
(From Theorell and Bonnichsen, *Acta Chem. Scand.* 5, 1105, 1951.)

Since this band derives from the reduced niacin amide it is highly likely that a bond formation between the pyridine ring and some group in the protein is needed to give such a remarkable displacement of the band; and Meyerhof, Ohlmeyer and Möhle (45) already in 1938 had shown that cyanide or bisulfite gives addition compounds with DPN, supposed to have the formulae:



These compounds had absorption bands at 320 mμ, just as the DPNH-ADH. Therefore it seems likely that some group in the ADH-protein is bound to the carbon atom next to the pyridine-N of DPN. This group could be a sulfhydryl, because it was found (18) that the addition of p-chloromercuribenzoate not only inhibits the

enzymatic reaction, but also causes the band maximum of ADH-DPNH to move back from 325 to 340 mμ. However, other possibilities must be considered, too. If sulfhydryl groups are conjugated to some other site of DPN, the addition of PCMB could release the DPNH and cause the band to go back to 340 mμ. Studies on the stoichiometry and kinetics of the reaction of PCMB with ADH · DPNH are in progress in collaboration with B. Chance. Neilands, while working in my institute, started an investigation on the system lactico dehydrogenase (LDH) + DPN in order to find out whether the band 340 mμ was shifted to lower wavelengths in the compound LDH · DPNH. The work was finished in the U. S. A., partly in collaboration with B. Chance ([46]). In this system some difficulties were met with because LDH · DPNH is much more dissociated than ADH · DPNH. Nevertheless, they could

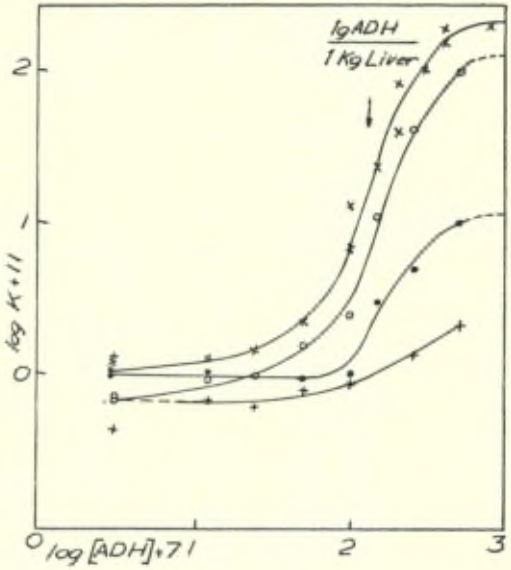


Fig. 2

Equilibrium constant K, at varied [ADH] (M = 73,000) and pH. Log K + 11 plotted against [ADH] + 7.

× pH = 7
O pH = 8
● pH = 9
\+ pH = 10

(From Theorell and Bonnichsen, *Acta Chem. Scand.* 5, 1105, 1951).

under suitable conditions demonstrate a band shift to 330 mμ for the compound. Recently Chance (47) reported a shift of the DPNH band to shorter wavelengths in living yeast and bacteria.

One molecule of ADH (Mol. weight = 73,000) binds two molecules of DPNH (18) around neutral reaction, at higher pH less, so that at pH only one DPNH is bound per one ADH.

The dissociation constants of the DPN- and DPNH-compounds with ADH are of great interest. In the equilibrium

$$K = \frac{[DPNH]\ [Aldehyde]\ [H^+]}{[DPN]\ \ [Alcohol]}$$

we found the value of K to be strongly dependant on the concentration of ADH. If [ADH] was kept negligibly low compared with [DPN] and [DPNH], K was $= 0.86 \times 10^{-11}$ and independent of pH. With an excess of ADH K was found to be 200 times greater at pH 7 (see fig. 2).

This effect decreased with increasing pH, so that at pH 10 K was nearly independent of [ADH].

The reason for the influence of [ADH] is simply that the dissociation constants K_{ox} and K_{red} of ADH · DPN and of ADH · DPNH are different. At pH 7 the ratio $K_{ox} : K_{red}$ is thus = 200, at pH 8 = 135, at pH 9 = 10 and pH 10 sligthly above 1.

By the aid of the formula for the relation between dissociation constants and redox potential

$$E'_o \text{ (complex)} = E''_o \text{ (free DPN-DPNH)} + 0{,}030 \log \frac{K_{ox}}{K_{red}}$$

the redox potential of the « holoenzyme » at pH-values from 7 to 10 was calculated as shown in figure 3.

It is seen that the coupling of the coenzyme to the protein increases the potential so considerably that it comes rather close to the potential of alcohol-acetaldehyde, facilitating the reversible reaction between enzyme and substrate. The increased potential of the holoenzyme makes it more understandable how the enzyme can oxidize alcohol physiologically, because aldehyde is present to a

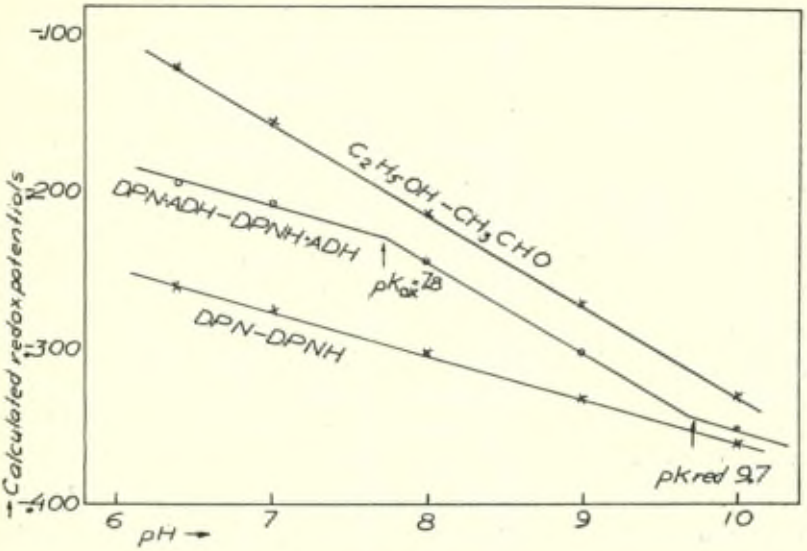


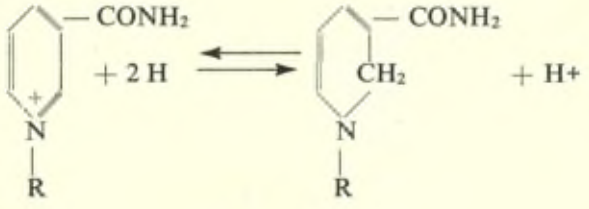
Fig. 3

Redox potentials in the ADH-system.

The straight lines for C_2H_5OH-CH_3CHO and DPN-DPNH are calculated according to Borsook (7) and Kalckar (8). The points (o) on the DPN · ADH - DPNH · ADH line are calculated from the equilibrium data at high [ADH]. (From Theorell and Bonnichsen, *Acta Chem. Scand.* 5, 1105, 1951.)

reasonable extent in the equilibrium state, and is continuously removed by aldehyde oxidases in the liver.

The great difference between K_{ox} and K_{red} at pH 7 gives another strong reason to assume a coupling between protein and the niacin amide part of the coenzyme, because this is the only place where the oxidation-reduction results in a structural change:



DPN DPNH

In addition to its favourable effect on the redox potential, the direct coupling of a group in the protein — maybe a sulfhydryl group — to the pyridine ring may serve to facilitate the oxidoreduction reaction between pyridine and substrate. The recent results of Racker and Krimsky on triosephosphate dehydrogenase (see below p. 325) are in agreement with this idea.

As seen in figure 3 the calculated curve for the redox potential of the holoenzyme shows two breaks at pH $\sim$ 7.8 and $\sim$ 9.7, indicating that an acid group with pK′ $\sim$ 7.8 is present in ADH · DPN, but not in ADH · DPNH, and that another acid group with pK′ $\sim$ 9.7 exists in ADH · DPNH, but not in ADH · DPN. This is most easily explained if we are dealing with the same group, the pK′ of which is influenced by the oxidation state of the pyridine.

The question whether DPN is « mobile » or « fixed » to ADH was answered through the work of Theorell and Chance ([8]). The mobile mechanism in this case should work as follows:

$$\mathrm{ADH} + \mathrm{DPNH} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{ADH \cdot DPNH} \quad \ldots\ldots\ldots\ldots \quad (1)$$

$$\mathrm{H^+} + \mathrm{ADH \cdot DPNH} + \mathrm{aldehyde} \overset{k_4}{\rightleftharpoons} \mathrm{ADH \cdot DPN^+} + \mathrm{alcohol} \quad (2)$$

$$\mathrm{ADH \cdot DPN^+} \underset{k_5}{\overset{k_3}{\rightleftharpoons}} \mathrm{ADH} + \mathrm{DPN^+} \quad \ldots\ldots\ldots\ldots \quad (3)$$

Experimental data for testing the validity of these formula were obtained by the aid of rapid spectrophotometry. The band shift from 340 to 325 mμ upon combination of DPNH with ADH was used for determing k_1 in a set of apparatus with two Beckman spectrophotometers operating simultaneously at 328 and 353 mμ. At these wavelengths the extinction is equal for free DPNH, but different for ADH-bound DPNH. The apparatus was arranged to record only differences between the light absorption at the two wavelengths. k_1 was found to be $= 0.4 \times 10^6\ \mathrm{M^{-1}} \times \mathrm{Sec^{-1}}$.

The dissociation constant

$$\mathrm{K_{red}} = \frac{k_2}{k_1}$$

was determined spectrophotometrically at very high dilution and found to be 10^{-7} M. Hence $k_2 = 0.4\ \mathrm{Sec^{-1}}$. k_4 is the overall reaction

velocity constant that differs for different substrates. Acetaldehyde gave inconveniently high reaction velocities, but good data were obtained with formaldehyde that reacts about 100 times more slowly.

In order to find k_3 we had to compute the kinetics formula valid under the assumptions given in equations 1-3. This lead up to some relations given in table 2.

TABLE 2

	1	2	3	4	5	6
[DPNH] $= x_1$	∞	$\frac{k_3}{k_1}$	∞	$\frac{k_2+k_4 a_1}{k_1}$	∞	∞
[Aldehyde] $= a_1$ . . .	∞	∞	small	small	variable	$\frac{k_3}{k_4}$
Turnover number . . . $\frac{1}{e} \cdot \frac{da_1}{dt}$	k_3	$\frac{k_3}{2}$	$k_4 a_1$	$\frac{k_4 a_1}{2}$	$\frac{1}{\frac{1}{k_4 a_1}+\frac{1}{k_3}}$	$\frac{k_3}{2}$

These relations are valid for the initial reaction velocities, when [DPN] and [alcohol] are still small; but it should be noticed that they can be used equally well for describing the reaction at the start from the opposite side, with DPN + alcohol + ADH.

From col. 1 we find that the maximal reaction velocity with both DPNH and aldehyde in excess is $= k_3$, that is the dissociation velocity constant of ADH · DPN. Under these conditions practically all of the ADH is present in the form of ADH · DPN, and the overall reaction velocity is limited by k_3. At pH 7 k_3 is $= 42$. Since the dissociation constant of ADH · DPN at pH 7 already had been determined to be 200×10^{-7}, k_5, the association velocity constant for ADH + DPN, is calculated to 2×10^6 $M^{-1} \times Sec^{-1}$. This values is within the limits of error equal to $k_1 = 4 \times 10^6$ $M^{-1} \times Sec^{-1}$. The association velocities of DPN and DPNH to ADH are thus roughly the same, which would be expected from the fact that they are both bound to the same site on the protein.

Col. 2 shows the interesting fact that the concentration of DPNH giving half of the maximal velocity, that is the Michaelis constant

for DPNH with excess of aldehyde is not at all equal to the dissociation constant of ADH · DPNH, $K_{red} = \frac{k_2}{k_1}$; instead it is $\frac{k_3}{k_1}$, the dissociation velocity constant of ADH · DPN divided by the association velocity constant of ADH + DPNH. Since, as just mentioned, k_1 is rougly equal to k_5, we find

$$\frac{k_3}{k_1} \sim \frac{k_3}{k_5} = K_{ox}.$$

The Michaelis constant for *DPNH* is thus $\approx$ the dissociation constant of ADH · *DPN*. With small concentrations of acetaldehyde the reaction velocity is proportional to [aldehyde], $= k_4 a_1$ (col. 3), and the Michaelis constant under these these conditions

$$= \frac{k_2 + k_4 a}{k_1}.$$

Excess of DPNH (col. 5) and varied a_1 gives

$$\frac{1}{e} \cdot \frac{da_1}{dt} = \frac{1}{\frac{1}{k_4 a_1} + \frac{1}{k_3}}.$$

The Michaelis constant for aldehyde, with excess of DPNH is $= \frac{k_3}{k_4}$ (col. 6).

ADH from liver is a surprisingly sluggish enzyme with a maximal turnover number of only around 140 moles of ethanol/mol ADH and minute at physiological pH. The reason for this is easy to see after our kinetics investigation; the reaction velocity in the oxidation of alcohol is limited by the low dissociation velocity of the reduced holoenzyme. The ADH from yeast works much faster; this brought Schlenk to be « tempted to suspect an evolutionary degeneration of the liver enzyme due to the more limited role of alcohol in the mammalian organism as compared with yeast » ([48]).

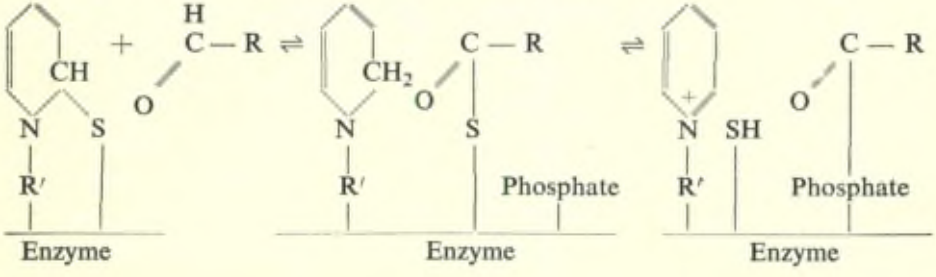
The sluggishness of the liver ADH depends upon the DPN functioning as « mobile » coenzyme, and not as « fixed » prosthetic group. Could the reduction and reoxidation of DPN occur without separation from the ADH, the reaction velocity would of course not be limited by the dissociation velocity constants of the complex. One could very well imagine a fixed mechanism to operate under physiological conditions in the liver; however, this seems not to

be the case, since our calculations of the amount of ADH needed to be present in the liver in order to account for the wellknown combustion velocity of alcohol in man gave 1.5 g per kg wet weight. This figure roughly agrees with direct determinations of ADH in crude extracts (5 g per kg in horse liver). Moreover, the combustion velocity of alcohol in the body proceeds at a rate that is independent of the alcohol concentration. This was to be expected from the low value for the Michaelis constant for alcohol at physiological pH (500 — 1,000 μM, corresponding to 0.002-0.004% ethanol). It is thus highly probable that the physiological combustion of alcohol is carried out mainly by the ADH, with DPN working as mobile coenzyme.

It appears probable that the pyridine nucleotides in the great majority of cases function in the mobile way, since the reaction velocities and dissociation constants of the coenzyme-enzyme compounds are in most cases greater than for ADH.

The glyceraldehyde-3-phosphate dehydrogenase (GA-3-PD) forms a remarkable exception to this rule. Cori et al. [49] observed that DPN and DPNH were very firmly attached to the protein, probably at double sites, one of them with a dissociation constant of 10^{-10}, the other one at least 100 times greater. Sulfhydryl inhibitors, like iodoacetate, completely stopped the enzymatic reaction. The authors point out that reoxidation of the firmly attached DPNH may occur through another enzyme system, i.e. lacticodehydrogenase + pyruvate.

Krimsky and Racker [50] found that GA-3-PD contains a firmly bound glutathione group that is essential for the activity, and Racker and Krimsky [51] presented evidence that the oxidation of aldehydes by this enzyme occurs in two steps: the aldehyde first reacts with the — SH group of the enzyme and is oxidized to a thiol ester. In the second step the acyl group of the thiol ester is transferred to phosphate to form an acyl phosphate:



HEMOPROTEINS

In none of the classes of conjugated proteins is the modifying and differentiating action of the protein part on the mode of action of the prosthetic group so strikingly manifold as in the hemoproteins. In the majority of the cases — hemoglobin, myoglobin, erythrocruorins, cytochromes *b* and *c*, catalases, and peroxidases from plants — the prosthetic group is iron protoporphyrin. Cytochrome *a*, cytochrome oxidase, chlorocruorins and animal peroxidases (myelo- and lacto-) contain other hemes, the structure of which is not yet definitely established. We do not yet know whether the deviations from protoheme are of any functional importance, or whether they are to be regarded as remnants from experiments, made by Nature in the course of phylogenetic evolution. Perhaps the second alternative is the more likely one, because none of these hemoproteins seem to have any special functions that could not as well be fulfilled by protoheme in combination with a suitable protein.

Three different types of functions may be distinguished in the hemoproteins:

1. Gases like oxygen, carbon monoxide and others may be attached reversibly to divalent heme iron without any valency change. This occurs in the natural oxygen carriers hemoglobins, myoglobins and erythrocruorins

$$Fe^{++} + O_2 \rightleftarrows Fe^{++}O_2 \qquad (1)$$

2. In the cytochromes the iron oscillates between the di- and trivalent states. They thus work as intermediate electron carriers in the series of enzymatic oxidation-reduction reactions that lead electrons from the hydrogen of substrates to molecular oxygen

$$Fe^{+++} + e \rightleftarrows Fe^{++} \qquad (2)$$

It should be mentioned that cytochrome oxidase (Warburg's respiratory pigment) may be supposed to react according to both (1) and (2), so that the oxygen attached according to (1) takes up an electron from Fe^{++}; the Fe^{+++} thus formed is reduced again by taking up an electron from the next catalyst in the chain, presumably cytochrome *a* or *c*.

3. In the hydroperoxidases (catalases and peroxidases) conditions are more complicated. The hematin iron in these enzymes is ferric, but there has been much dispute about the question whether they remain so in their reaction cycles with hydrogen (or alkyl-) peroxide

[Sumner ([52]), Theorell ([53]), and others], or whether they oscillate between the ferric and ferrous states [Keilin and Hartree ([54])].

Very recent results have to some extent clarified these complicated reactions. It has been known for several years that if a small amount of H_2O_2 is added to horseradish peroxidase (HRP) the colour turns greenish [« Compound I », Theorell ([55])]. This colour is in a more or less short time depending on the amount of hydrogen donor present [Chance ([56])] substituted by a reddish one [« Compound II », first observed by Keilin and Mann ([57])]. Compound II then disappears, and free HRP appears again. George and Irvine ([58]) demonstrated that the hydrogen peroxide compound of ferrimyoglobin had only one oxidation equivalent left instead of the two present in one molecule of hydrogen peroxide, and the same was later found for HRP-II [George ([59])]. The picture is rendered still more intricate by the appearance of another type of red compounds of H_2O_2 and catalases, peroxidases, ferrihemoglobin and ferrimyoglobin. They are numbered « III », are bright red in colour and are spectroscopically different from type II, the absorption bands in the visible being shifted towards the red. The compound « III » of HRP appears with excess of H_2O_2 [Keilin and Mann ([57])], an in ferrihemo- and myoglobin the compounds III are the only ones that are known. The compounds III are sluggish and do not seem to take part in the enzyme reactions under physiological conditions.

Chance ([60]) recently succeeded in titrating the number of oxidation equivalents in the primary, green HRP-I, and as would have been expected found the whole oxidation capacity of H_2O_2, two equivalents per mol, to be present in this compound.

Our magnetic measurements ([61]) on the peroxide compounds of HRP, catalase and ferrimyoglobin revealed that the compounds « I » have ferric iron with three odd electrons, a rare bond type intermediate between covalent and ionic. The compounds II have two odd electrons. These cannot both be present in ferric iron, but since only one oxidation equivalent is present, one of the odd electrons may reasonably be assumed to be present in a radical, the other in ferric iron held by covalent bonds. George and Irvine ([58]) prefer to assume that the iron in types II (and III) is tetravalent. It is at present impossible to decide between the two alternatives, since both are compatible with the magnetic data. We prefer the first because of the spectroscopical similarity of the

compounds II and III with other ferric, covalent hemoprotein compounds.

Keilin's and Hartree's (54) arguments for an oxidation- reduction of the catalase iron occuring in the decomposition of H_2O_2 have recently been disproved [Theorell and Ehrenberg (62)]. They were based upon the formation of a ferrous catalase compound with H_2O_2 + azide. We found that azide acts not as a common inhibitor, but as a cosubstrate in this reaction, and is oxidized itself to nitrous oxide + nitrogen. In other words, there is no indication of a reduction of catalase iron by H_2O_2, unless a suitable substrate, like azide, is present at the same time.

After all these recent results we may venture to formulate the reaction cycles of the hydroperoxidases in the following schematic way :

$$Fe^{+++} + H_2O_2 \longrightarrow \underset{\text{(Compound I)}}{Fe^{+++} \cdot H_2O_2} \quad \ldots\ldots\ldots\ldots\ldots \quad (3\ a)$$

$$Fe^{+++} \cdot H_2O_2 + AH_2 \longrightarrow \underset{\text{(Compound II)}}{Fe^{+++} \cdot \dot{O}H} + \dot{A}H + H_2O \quad \ldots \quad (3\ b)$$

$$Fe^{+++} \cdot \dot{O}H + \dot{A}H \longrightarrow Fe^{+++} + H_2O + A \quad \ldots\ldots\ldots \quad (3\ c)$$

In the case of catalases AH_2 may be either H_2O_2, or some organic compound, e.g. alcohol. Reaction 3 *c* is then so rapid that no compound II has yet been observed. The peroxidases do not decompose H_2O_2, so AH_2 cannot be = H_2O_2. They give more stable compounds II.

***

All three of these types of action are inherent in the free ferro- and ferriheme molecule. Ferroheme in solution can add carbon monoxide — but not oxygen, that instead of being attached as in hemoglobin will oxidize ferro- to ferriheme. Heme can serve as an oxidation catalyst like the cytochromes, for instance in the aerobic oxidation of cysteine. And, finally, ferriheme has some catalatic as well as peroxidatic activity.

The function of the protein part of hemoproteins is to specialize the heme for one of the three types of reaction, and for certain substrates, and to accelerate the reaction velocity, sometimes thousands of times. How does the protein promote these effects?

This is a formidable problem to solve. A short survey of our present knowledge on hemoglobin, myoglobin, cytochrome *c* and HRP will serve to illustrate our present knowledge in this field.

Hemoglobin.

Wymans (63) well-known differential acid-base titrations on oxyhemoglobin and hemoglobin gave very strong evidence in favour of Conants earlier idea, that imidazole rings of histidine residues were linked to heme iron in hemoglobin. Coryell and Pauling (64) gave an elegant interpretation of how the Bohr effect could derive from this linkage, based on the fact [Pauling and Coryell (65)] that the iron is held by essentially ionic bonds in ferro- and ferrihemoglobin, but by covalent bonds in oxyhemoglobin. The consequence of Coryell's and Pauling's formulae seems to be that a Bohr effect must inevitably occur in all hemoglobins containing this iron-imidazole linkage. However, Wyman (66) has recently emphasized that the Bohr effect shows very great variations in different species of hemoglobin, and Riggs (67) found an extreme example: the hemoglobin of the adult bull frog has a very large Bohr effect, whereas the tadpole hemoglobin shows no Bohr effect at all. The conclusion would seem to be that iron is bound to imidazole in the hemoglobins that show a Bohr effect, but not in those where this effect is lacking. Wyman offers an alternative explanation: The whole shape of the hemoglobin molecule is profoundly changed when it is oxygenated or oxidized. The change in shape results in differences in crystal structure and solubility. According to Wyman the Bohr effect may be due not to changes of bond type in the heme-globin linkage, as presumed by Coryell and Pauling, but to changes in the position of acid groups near the heme, resulting in changes in their dissociation constants. Wyman's theory ascribes not only the Bohr effect but also the heme-heme interactions to entropy effects associated with these configurational alterations. Obviously this explanation can easily account for any species differences in Bohr effect between hemoglobins. On the other hand the question of whether imidazole rings are conjugated to heme — that for a decade of years was regarded as a corner stone of hemoglobin chemistry — must again be discussed.

In a recent paper from Wyman's institute Riggs (68) has proved

that the heme-heme interaction is strongly dependant on free sulfhydryl groups. Why this is so still remains to be explained.

Myoglobin.

As compared with hemoglobin, myoglobin offers some obvious advantages for the study of the relations between heme and protein. It contains only one heme per molecule, so that no heme-heme interaction can occur; it can easily be dissociated reversibly into hematin and protein ([69]). The free protein is homogeneous in the ultracentrifuge with a sedimentation constant nearly equal to that of myoglobin. The recombined myoglobin is indistinguishable from ordinary myoglobin in sedimentation constant, electrophoretic mobility and light absorption. It is hoped that further experiments on the free protein component and the reconstituted myoglobin may give valuable information about the chemical nature of the heme-linked groups.

Our magnetic and spectrophotometric investigations on myoglobin ([70]) revealed some obvious differences between the heme-linked groups in myoglobin and hemoglobin. Some of our data on myoglobin are compared with data on ferrihemoglobin from the literature in table 3.

TABLE 3

pK-values of heme linked groups in ferri-myoglobin, ferrihemoglobin and their fluorides.

	Hb+	Mb+	HbF	MbF
pK_1	5.3-5.45 Mo Pi Si To	5.3 Mi Pi? Si	5.5-5.8 Mi? So	6.0 Mo So
pK_2	6.65 Mi Po Si	7.4 Mi P? Si	6.9 Mi? Si	8.1 Mo Si
pK_3	7.88 Mo So To	8.84 Mo So To	—	—

M = magnetically
P = potentiometrically
S = spectrophotometrically
T = titrimetrically
o = operable
i = inoperable

The $pK_1 = 5.3$ for Hb^+ was determined magnetometrically by Coryell, Stitt and Pauling (71); the value 5.45 was found titrimetrically by Theorell (72). Mb^+ contains a heme-linked group of about the same pK. Since this group is magnetically inoperable in Mb^+ it had to be derived from spectrophotometric and magnetic data of the fluoride compound. It is thus probably not identical with the corresponding group in ferrihemoglobin.

The second acid heme-linked group has a higher pK value, 7.4, for Mb^+ than for Hb^+, and a similar difference occurs for the third acid group, that is the iron atom itself in the equilibrium

$$Fe^{+++} \cdot H_2O \rightleftarrows Fe^{+++}OH^- + H^+$$

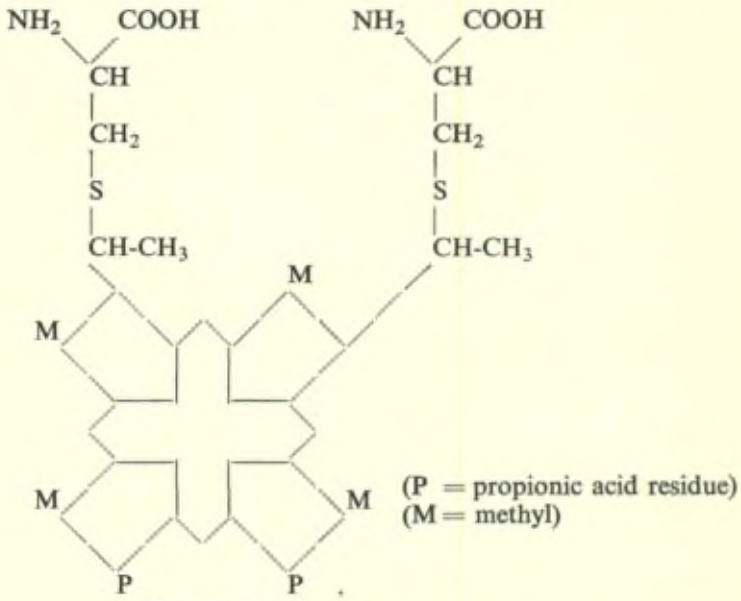
The general conclusion could be drawn that the groups in the vicinity of the iron in myoglobin must have a more negative character than in the hemoglobin, because not only the pK_3 is about one unit higher in Mb^+ than in Hb^+, but also the dissociation constants for cyanide and fluoride are ten times higher in ferrimyoglobin than in ferrihemoglobin.

The paramagnetic susceptibilities for Mb^+ and Hb^+ are different, too; the alkaline forms give, for example $\chi_m = 11{,}000 \times 10^{-6}$ and $8{,}000 \times 10^{-6}$ c.g.s. e.m.u. respectively. Finally, as would be expected from the nearly total absence of Bohr effect in myoglobin, ferromyoglobin and carbon monoxide myoglobin gave identical titration curves. There is thus no experimental evidence whatsoever in favour of imidazole being heme-linked in myoglobin. On the contrary, the effects predicted by Coryell and Pauling for this type of linkage is absent.

Cytochrome c.

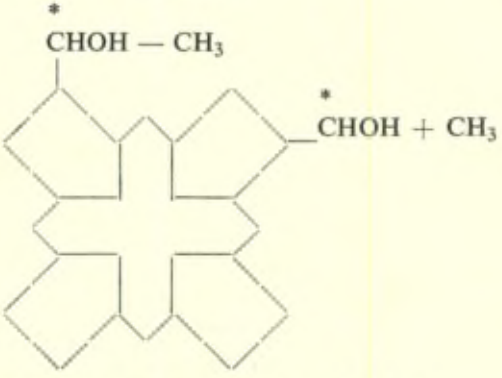
A. *Porphyrin-protein linkages.* — The conjugating groups in this interesting hemoprotein have been subjected to a great deal of studies. From acid hydrolysates Theorell (73) isolated a « porphyrin *c* » that contained N, amino-N, S, COOH and Fe (subsequently introduced) in the ratios 6 : 2 : 2 : 4 : 1, and from this preparation l-cystine was isolated (74). Since the sulfur in the porphyrin *c* was present in thioether bonds, and for other reasons,

we could draw the conclusion that « porphyrin *c* » was a dicysteine adduct to protoporphyrin:



Porphyrin c

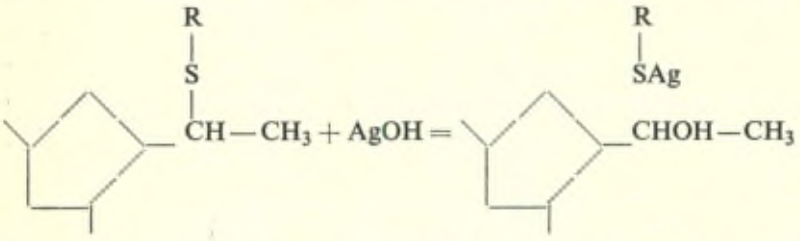
This structure was recently definitely confirmed by Paul ([75]), who had found that silver salts split cytochrome into a colourless protein and a hematin that was identified as one of the four possible stereoisomers of hematohemin:



Hematohemin (* C = asymmetrical C-atom)

The optical activity was — 97° in dioxane, — 31° in glacial acetic

acid, showing that at least one of the hydroxyl groups must be in the α-position of the hydroxyethyl group. That this was the case for both hydroxyl groups was evident from the facts that the tetramethyl hematoporphyrin yielded protoporphyrin dimethyl ester when heated in vacuo to 135-150° for 5-10 minutes, and that neither of the hydroxyl groups in the porphyrin could be tritylated, as expected for α-hydroxyls, but not for β-hydroxyls. The splitting of cytochrome *c* with silver salts thus occurs in the following way:



On the other hand it follows that the prosthetic group in cytochrome *c* is protohemin; the vinyl groups conjugate with cysteine-SH in the protein through an addition reaction, leading to a particularly firm combination between prosthetic group and protein.

It may be remarked that this seems to be *the only case hitherto, in which the conjugating groups in both protein and prosthetic group* have been finally established by the aid of the methods of ordinary organic chemistry; and besides, the hematoporphyrin *c* was the first preparation of an optically active porphyrin.

B. *Iron-protein linkages.* — On the basis of spectrophotometric studies titration curves and magnetic experiments we, in 1941 ([76]), advanced the hypothesis that the two nitrogenous hemochrome-forming groups in cytochrome *c* were likely to be imidazoles belonging to histidine residues in the protein component. This gave the picture of the hematin disc being imbedded in a crevice of the protein moiety, making the iron inaccessible to molecular oxygen; this explains why cytochrome *c* is not autoxidizable. On the other hand the iron would be readily accessible for electrons through the covalent bonds e.g. imidazole that could be imagined to form radicals and exchange electrons with the iron ([72]).

In the meantime some more evidence has appeared in favour of the imidazole theory [Paul ([77])]. However, quite a deal of conflicting data are still to be cleared up, so we shall at this occasion confine ourselves to the description of some as yet unpublished

work ([78]) that has shed some new light upon the nature of the hemochrome-forming groups.

The spectrophotometric determinations in 1941 indicated two dissociation constants to occur in the acid region with pK = 2.5 and 0.5. We considered the first to depend upon the rupture of one of the nitrogen-Fe bonds, the second one to the other bond being broken. Since Paul has found that the lowest pH tolerated by cytochrome *c* without loss of enzymatic activity was 1.6, and Maehly ([79]) had found interesting effects of the chloride ion on the acid splitting of horse radish peroxidase, we found it advisable to reinvestigate the behaviour of cytochrome *c* in acid solution at varied chloride concentrations, by the aid of both spectrophotometry and magnetometry. Some of the results are given in the figures 4-7.

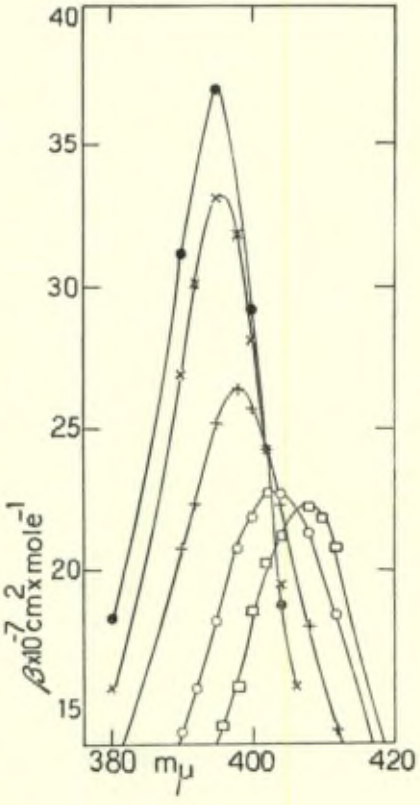


Fig. 4

pH.
● 1.70 × 1.87 + 2.20 o 2.53 □ 3.36

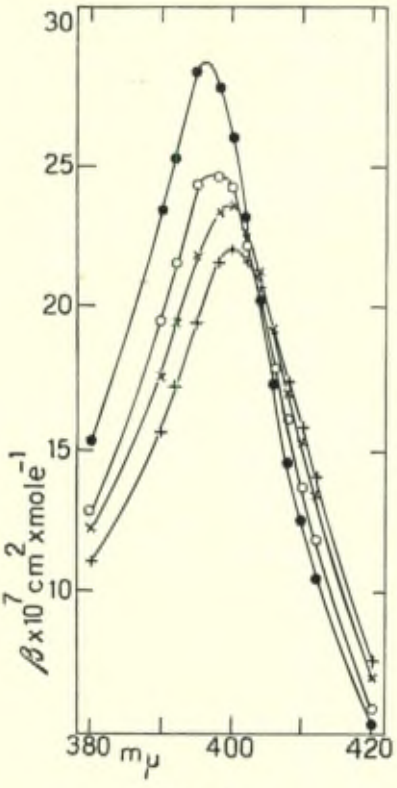


Fig. 5

pH. 1.0

Molarity of added Na Cl

● 0.0 o 0.066 × 0.132 + 0.528

Figure 4 shows the Soret band of ferricytochrome *c*, when pH is varied at low [Cl⁻]. Not much more is seen than a transition of one form, neutral cytochrome *c* (form « III » of 1941 with its maximum at 408 mμ) to an acid form (« II » of 1941) with its maximum at 395. However, since there is no isosbestic point the suspicion arises that some intermediate form may exist. This is clearly found to be the case in figure 5, where pH is kept constant, but [Cl⁻] varied: at increasing [Cl⁻] the peak of the Soret band is lowered and approaches a new position 402 mμ. And this compound can be asymptotically approached by increasing [Cl⁻] at both pH 1, 2 and 3 (fig. 6). Figure 7 gives some data showing that the light absorption (at 395 mμ) is dependant upon both pH and [Cl⁻]. These and many other data proved that:

1. At high enough acidity the bonds between the iron atom and the two hemichrome-forming nitrogen atoms are broken by two protons being added, one to each of them. The entrance of one proton facilitates that of the second one so strongly that the two steps cannot be separated ($n = 2.0$, maximal interaction).

2. The cyctochrome molecule as such is not split, at least above pH = 1; the heme part remains attached to the protein by the two thioether bonds. Thus we may assume that after breaking the bonds to iron by the addition of protons the $\equiv \overset{+}{\mathrm{N}}$ groups remain essentially in the same positions relative to the positively charged iron atom.

3. Anions, like Cl^-, must be stronly attracted into the two gaps in the sequence $\equiv \overset{+}{\mathrm{N}}\mathrm{H}\ \mathrm{Fe}^+\ \overset{+}{\mathrm{H}}\mathrm{N}$ to form a dipole chain : $\equiv \overset{+}{\mathrm{N}}\mathrm{H}\ \mathrm{Cl}^-\ \mathrm{Fe}^+\ \mathrm{Cl}^-\ \overset{+}{\mathrm{H}}\mathrm{N} \equiv$. This is the dichloride compound with its Soret band maximum at 402 mμ.

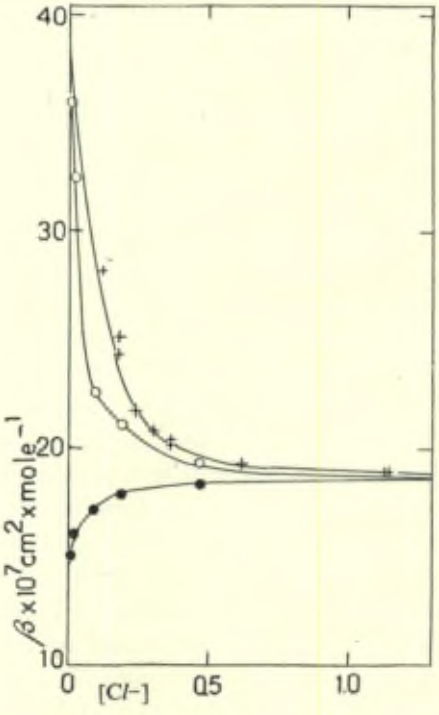


Fig. 6

pH : 1.0 (+); 2.0 (o); 3.0 (●); λ = 395 mμ

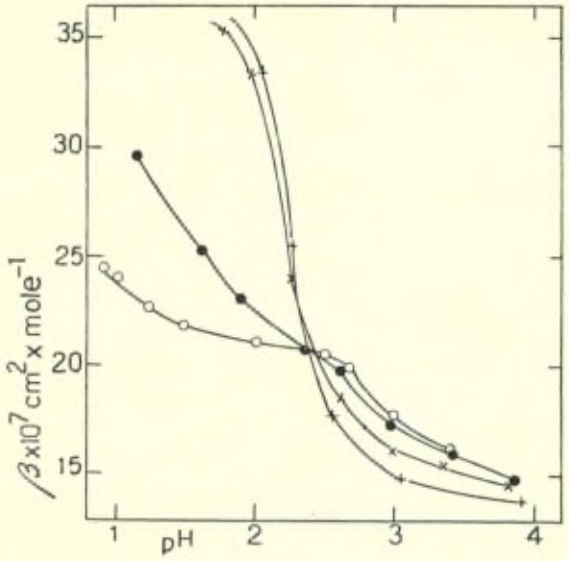


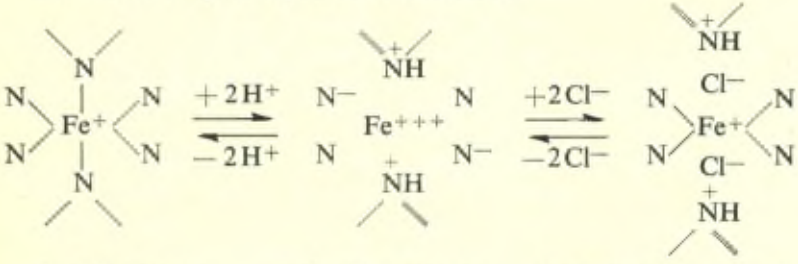
Fig. 7

+ 0.009 *g.eq.* Cl^-/l × 0.019 *g.eq.* Cl^-/l
● 0.094 *g.eq.* Cl^-/l o 0.189 *g.eq.* Cl^-/l

4. It should be observed that protons and chloride ions synergistically cooperate in splitting off the hemichromeforming groups from the iron atom; the degree of splitting at any given pH is greatly enhanced with increasing chloride ion concentration.

The magnetic properties of the compounds are interesting: in the neutral ferricytochrome *c* the iron is held by covalent bonds (one odd electron); in the acid form, with disrupted N — Fe bonds, ionic bonds (5 odd electrons); in the dichloride intermediate bonds with 3 odd electrons.

The results may be represented as follows:



1 odd electron, covalent d^2sp^3 bonds

5 odd electrons, essentialy ionic bonds

3 odd electrons, square dsp^2 bonds to the pyrrol nitrogens

Peroxidases.

In myeloperoxidase and lactoperoxidase the greenish hematin (of unknown constitution) is firmly bound to the protein. Since it is split off neither by acetone + hydrochloric acid, nor by silver salts. In horse radish peroxidase « HRP », which we crystallized in 1942, the linkages between protohematin and protein are such as to allow a reversible separation of hematin and protein without any great loss of active protein. It was therefore possible to compare the titration curves of free protein and recombined enzyme ([72]). The result was compatible with the assumption of a carboxyl group being bound to the iron atom, and that at least one of the propionic acid residues is bound to some weakly acid groups (— NH_2?) in the protein.

Maehly ([79]) has approached the same problem by following the spectrophotometrical changes occurring when HRP is split by mineral acids. When working with HCl he observed by using rapid spectrophotometry, first the formation of a chloride compound (« A »), then a compound « B » that was supposed to represent a product with the (hypothetic) carboxyl group split from the iron, followed by free protohemin (« C »), and polymerized protohemin. These studies are continued.

An interesting difference between HRP on one side, ferrihemoglobin and myoglobin on the other is that the pK of the transition from the brown, neutral form to the red, alkaline one occurs at much higher alkalinity in HRP (pK = 11 for HRP, 8 for Hb, 9 for Mb). This difference could very well be caused by the influence of a (negatively charged) carboxyl group in the vicinity of the iron atom in HRP. The magnetic data for the alkaline forms differ greatly, too ([70]): 1 odd electron in HRP, 3 in HbOH, 5 in MbOH.

Having the free HRP protein at our disposal we studied some « synthetic » peroxidases with hemins other than protohemin ([80]) and ([81]). Gjessing and Summer ([82]) independently used the protein, described by us, for similar investigations. It was found that meso- and deutero-hemin gave active peroxidases, though the activity was somewhat lower than for the natural enzyme. Hematohemin was found to be active by the American authors, whereas we found it to be inactive. Gjessing and Sumner were right on this point, since pure specimens of hematohemin, prepared by Paul in Stockholm, now were found to be active ([83]). The vinyl groups are

thus of no great importance for the enzyme activity, since they could be removed (deutero-) or saturated with hydrogen (meso) or with water (hemato-) without any drastic loss of activity. Both of the propionic acid residues, on the contrary, were found to be essential for the activity, as would be expected if they are used as conjugating groups, see above.

***

In this lecture I have confined myself to a brief summary of *experimental results* on the subject of relations between prosthetic groups and proteins, and therefore avoided to discuss the large literature on more general explanations of the enzymes working mechanism. Stimulating as these theories may be — confer for example Geissman [84] — they can serve only as working hypotheses, but can never substitute for the patient collecting of chemical and physical data on each protein. This implies an amount of work, the magnitude of which cannot even be estimated. The rapid improvement of our methods, and the rapid general progress in biochemistry, however, allows us to hope for a much deeper insight in these problems in a not too distant future.

LITERATURE

(1) Svedberg, The and Pedersen, K.O., *The Ultracentrifuge*, Oxford (1940)
(2) Sumner, J.B., *J. Biol. Chem.*, **69**, 435 (1926).
(3) Northrop, J.H., Kunitz, M. and Herriott, R.M., *Crystalline Enzymes*, Columbia University Press, New York (1948).
(4) Zeile, K., and Hellström, H., *Zeitschr. physiol. Chem.*, **192**, 171 (1930).
(5) Kuhn, R., Hand, D. and Florkin, M., *Zeirschr. physiol. Chem.*, **201**, 255 (1931).
(6) Theorell, H., *Biochem. Z.*, **272**, 155 (1934).
(7) Warburg, O. and Christian, W., *Biochem. Z.*, **266**, 377 (1933).
(8) György, P., Kuhn, R. and Wagner-Jauregg, Th., *Naturwiss.* **21**, 560 (1933); Wagner-Jauregg, Th. and Ruska, H., *Ber.* **66**, 1298 (1933).
(9) Wyman, J., *Advances in protein Chem.*, IV, 407 (1948).
(10) Parnas, J., *Nature*, **151**, 577 (1943).
(11) Barron, E.S.G., *Advances in Enzymology*, XI, 201 (1951).
(12) Sörbo, B., *Acta Chem. Scand.*, **5**, 1218 (1951).
(12a) Mellander, O., *Uppsala Läkarfören, Förh.* **52**. 107, (1947); **55**, 239, (1950); **55**, 247, (250).
(13) Kuhn, R. and Boulanger, P. *Ber.* **69**, 1557 (1936).
(14) Kuhn, R. and Weygand, F., *Ber.* **68**, 166 (1935).
(15) Haas, E., Harrer, C.J. and Hogness, T.R., *J. Biol. Chem.*, **143**, 341 (1942).
(16) Warburg, O. and Christian, W., *Biochem. Z.*, **298**, 150 (1938).
(17) Theorell, H. and Bonnichsen, R., *Acta Chem. Scand.*, **5**, 1105 (1951).
(18) Theorell, H. and Chance, B., *ibid.*, **5**, 1127 (1951).
(19) Hellerman, L., Lindsay, A., and Bovarnick, M., *J. Biol. Chem.*, **163**, 553 (1946).
(20) Kuhn, R., György, P. and Wagner-Jauregg, Th., *Ber.*, **66**, 1034 (1933).
(21) Theorell, H., *Biochem. Z.* **275**, 37 (1934).
(22) Theorell, H., *Biochem. Z.* **278**, 263 (1935).
(23) Haas, E., *Biochem. Z.*, **298**, 378 (1938).
(24) Straub, F.B., *Biochem. J.* **33**, 787 (1939).
(25) Altschul, A.M., Persky, H. and Hogness, T.R., *Science*, **94**, 349 (1941).
(26) Keilin, D., and Hartree, E.F., *Biochem. J.*, **42**, 221 and 230 (1948).
(27) Mahler, H.R., Sarkar, N.K. and Vernon, L.P., *J. Biol. Chem.*, **199**, 585 (1952); Vernon, L.P., Mahler, H.R. and Sarkar, *ibid.*, **199**, 599 (1952).
(28) Singer, T.P. and Kearney, E.B., *Arch. Biochem.*, **27**, 348 (1950); *ibid.*, **29**, 190 (1950).
(29) Warburg, O., *Erg. Enz. Forsch.*, VII, 224 (1938).
(30) Kuhn, R. and Moruzzi, G., *Ber.*, **67**, 888 (1934).
(31) Bessey, O.A., Lowry, O.H. and Love, R., *J. Biol. Chem.*, **180**, 755 (1949).
(32) Weber, G., *Biochem. J.*, **47**, 114 (1950).
(33) Michaelis, L., Schuberth, M.P. and Smythe, C.V., *J. Biol. Chem.*, **116**, 587 (1936).
(34) Theorell, H., *Biochem. Z.*, **290**, 293 (1937).
(35) Hellerman, L., *Biological Antioxidants, Trans.* II. *Conf.*, Josiah Macy Foundation, p.78 (1947).
(36) Haas, E., *J. Biol. Chem.*, **155**, 321 (1944).
(37) Theorell, H., *Bioch. Z.*, **275**, 30 (1934).
(38) Theorell, H., *Naturwiss.*, **22**, 290 (1934); *Biochem. Z.* 275, 11 (1934).
(39) Kornberg, A., and Pricer, W.E., *J. Biol. Chem.*, **186**, 557 (1950).
(40) Kaplan, N.O., Colowick, S.P. and Ciotti, M., *J. Biol. Chem.*, **194**, 579 (1952).
(41) Williams, J.N. jr., *J. Biol. Chem.*, **195**, 629 (1952).
(42) Pullman, M.E., Colowick, S.P. and Kaplan, N.O., *J. Biol. Chem.*, **194**, 593 (1952).
(42a) Theorell, H., *Biochem. Z.* 275, 416, (1935).
(43) Feigelson, P., Williams, J.N. jr., and Elvehjem, C.A., *J. Biol. Chem.*, **189**, 361 (1951).
(44) Alivisatos, S.G.A., and Denstedt, O.F., *J. Biol. Chem.*, **199**, 493 (1952).
(45) Meyerhof, O., Ohlmeyer, P., and Möhle, W., *Biochem. Z.*, **297**, 90 (1938).

(46) Neilands, J., *J. Biol. Chem.*, **199**, 373 (1952); Neilands, J. and Chance, B., *ibid.*, **199**, 383 (1952).
(47) Chance, B., *Nature*, **169**, 215 (1952).
(48) Schlenk, F., Sumner-Myrbäck, « *The Enzymes* », Academic Press, New York (1951), II, Part I, p. 291.
(49) Cori, C.F., Velick, S.F. and Cori, G.T., *Biochim. et Biophys. Acta*, **4**, 160 (1950).
(50) Krimsky, I. and Racker, E., *J. Biol. Chem.*, **198**, 721 (1952).
(51) Racker, E., and Krimsky, I., *J. Biol. Chem.*, **198**, 731 (1952).
(52) Sumner, J.B., *Adv. in Enzymology*, I, 168 (1941).
(53) Theorell, H. and Agner, K., *Ark. Kemi, Min., Geol.*, **16***A*, N°. 7 (1952).
(54) Keilin, D. and Hartree, E.F., *Proc. Roy. Soc.* (London), ***B*124**, 397 (1938).
(55) Theorell, H., *Enzymologia*, **10**, 250 (1941).
(56) Chance, B., *Arch. Biochem.*, **22**, 224 (1949).
(57) Keilin, D. and Mann, T. *Proc. Roy. Soc.* (London), ***B*122**, 119 (1937).
(58) George, P. and Irvine, D.H., *Nature*, **168**, 184 (1951).
(59) George, P., *Advances in Catalysis*, **4**, 367 (1952).
(60) Chance, B. *Arch. Biochem. Biophys.*, **41**, 416 (1952).
(61) Theorell, H., and Ehrenberg, A. *Arch. Biochem. Biophys*, **41**, 442 (1952).
(62) Theorell, A., and Ehrenberg, A. *ibid.*, **41**, 462 (1952).
(63) Wyman, J., *J. Biol. Chem.*, **127**, 581 (1939).
(64) Coryell, C.D. and Pauling, L. *ibid.*, **132**, 769 (1940).
(65) Pauling, L. and Coryell, C.D., *Proc. Nat. Acad. Sci. U. S.*, **22**, 159 and 210 (1936).
(66) Wyman, J. and Allen, D.W., *J. Polymer Sci.*, **7**, 499 (1951).
(67) Riggs, A. *J. Gen. Physiol.*, **35**, 23 (1951).
(68) Riggs, A. *J. Gen. Physiol.*, **36**, 1 (1952).
(69) Drabkin, D. L., *J. Biol. Chem.r* **158**, 721 (1945).
(70) Theorell, H. and Ehrenberg, A., *Acta Chem. Scand.*, **5**, 823 (1951).
(71) Coryell, E.D., Stitt, F. and Pauling, L., *J. Amer. Chem. Soc.*, **59***A*, 633 (1937).
(72) Theorell, H., Ark. Kemi, *Min. Geol.*, **16***A*, N°. 14 (1943).
(73) Theorell, H., *Biochem. Z.*, **298**, 242 (1938).
(74) Theorell, H., *Enzymologia*, **6**, 88 (1939).
(75) Paul, K.G., *Acta Chem. Scand.*, **5**, 389 (1951).
(76) Theorell, H., and Åkeson, A., *J. Amer. Chem. Soc.*, **63**, 1804 (1941).
(77) Paul, K.G., *Acta Chem. Scand.*, **5**, 379 (1951).
(78) Boeri, E., Ehrenberg, A., Paul, K.G. and Theorell, H., to be published.
(79) Maehly, A., *Biochim. et Biophys. Acta*, **8**, 1 (1952).
(80) Theorell, H., Bergström, S. and Akeson, Å., *Ark. Kemi, Geol.*, **16***A*, N°. 13 (1943).
(81) Theorell, H., and Maehly, A., *Acta Chem. Scand.*, **4**, 422 (1950).
(82) Gjessing, E.C. and Sumner, J.B. *Arch. Biochem.*, **1**, 1 (1943).
(83) Unpublished.
(84) Geissman, T.A. *Quat. Rev. of Biol.*, **24**, 309 (1949).

Discussion du rapport de M. H. Theorell

M. Theorell. — I would be most grateful to have Dr. Pauling's opinion on whether he feels that Wyman's new picture of hemoglobin structure implies that the Bohr effect can now appear, or be absent, independently of whether imidazole groups are bound to the iron or not.

M. Pauling. — 1) The change in acid constant of an imidazole ring with change in nature of the bond to the iron atom, as discussed by Coryell and myself, no doubt takes place, but its magnitude may be affected somewhat by the environment, such as an arrangement of the protein that holds the ring away from the iron atom. There might occur a range of values of the Bohr interaction energy for a haem-linked imidazole ring, and another range (nearer to zero) for a haem not linked to a histidine residue. These ranges might overlap, making it difficult to decide whether or not the haem is linked to a histidine residue in a particular protein.

2) The structure of acid cytochrome-*c* dichloride is a striking one. It is another example of the rare intermediate electronic structure of a compound of tripositive iron, in which it has three odd electrons, instead of one or five. The first example of this rare structure was ferrihaemoglobin hydroxide (Pauling, Coryell, and Stitt, 1947).

It seems likely that the iron atom forms five covalent bonds, which resonate among the positions to the four nitrogen atoms and the two chlorine atoms.

M. Theorell. — May I take the opportunity to ask for Dr. Pauling's opinion on the nature of the iron-nitrogen bonds in cytochrome-*c*-dichloride ? The 3 odd electrons would fit to a picture of four, square covalent DSP^2 bonds to the four pyrrol nitrogens, and ionic bonds to the chlorine ions.

M. Pauling. — The groups in hemoglobin and related proteins

that are known to have acid constants in the range near neutrality are the imidazole groups. It is probable accordingly that change in acidity with change in structure or composition is to be attributed to these groups. The new picture of the hemoglobin molecule provides, however, a more general explanation of the way in which structural change can lead to a change in the acid constant of the imidazole group than has been given before. In our early discussion of this point, Coryell and I mentioned that a change in the nature of the bond from the iron atom to the nitrogen atom of an imidazole group would change its acid constant, and also that the displacement of the imidazole group by the oxygen molecule, attaching itself to the iron atom, could change the acid constant of this group. More generally, we may now say that the change in the structure of the protein, in the neighborhood of the heme, when an oxygen molecule is attached to the iron atom of the heme may cause a change in the acid constant of any imidazole group in the region subject to steric effects. I feel that the major Bohr effect is still to be attributed to an imidazole ring directly attached to the iron atom.

As to the nature of the bonds in cytochrome-*c*-dichloride, in which the iron atom has three odd electrons, I would agree that the iron atom may be forming four bonds with the four nitrogen atoms; in addition, however, it may form two half-bonds with the two chlorine atoms. This would require a total of five orbitals (with resonance of one bond between the two chlorine atoms), leaving enough orbitals on the iron atom to permit three of the electrons to be unpaired.

M. Theorell. — Thank you; I understand you mean that we have still to assume that iron is bound to imidazole in hemoglobin, but that almost any changes in the magnitude of the Bohr effect could be brought about by changes in the molecules shape, as discussed by Wyman.

M. Bragg. — Can the whole shape of the haemoglobin molecule be profoundly changed ? (Prof. Theorell's report, p. 329). One of the rather definite and striking results of X-ray measurements of protein crystals, is that the overall shape of the molecule is very much the same in different crystalline forms. It appears that the dimensions of the molecule in any direction do not differ by more than two or three per cent at most.

The pleochroic ratios of haemoglobin and myoglobin crystals, which Perutz and Kendrew have been investigating, may cast some light on what happens when chemical changes take place at the haem position. Oxyhaemoglobin of the horse has a pleochroic ratio of 2.6, the large absorption being for an electric vector perpendicular to the axis of the crystal. The long axis of the molecule makes an angle of about 20° with the *a* axis. The reduced form shows a much higher pleochroic ratio, about 4.5 : 1. Further, certain myoglobin crystals have an extremely high pleochroic ratio. If the haem group is responsible for the pleochroism, we suppose it has a large absorption in its plane and a small absorption for a light vector perpendicular to its plane. The highly pleochroic myoglobins must have very nearly parallel haem groups. As haemoglobin has a ratio of only 2.6, the haem groups must be inclined at an angle of some 20° to 30° with the *a* axis. Does this increase when reduced imply a change in orientation of the haem, or is it due to a change in its absorption coefficent ?

M. Theorell. — With regard to Sir Lawrence Bragg's comments, I might add that since myoglobin contains only one hematin per molecule I understand that the high pleiochroism must indicate that the flat hematin disks are normally strictly parallel in good crystals. In hemoglobin with its 4 hematins per molecule the possibility of tilting the hematin disks in relation to one another will probably, among other things, depend upon whether they are bound to the protein only at the central iron atom or whether additional bonds are present at the periphery of the hematin molecules. Some titration curves we made some ten years ago seem to indicate that one or two of the propionic acid residues would probably be bound to the protein component, but I would not dare to say that our evidence is absolutely convincing.

M. Synge. — 1) I have objected to the usage of the term « link », « linkage » (p. 304) as being liable to confusion with chemical linkages in the sense of direct bonds between atomic groupings, and proposed « interact », « interaction », as possibly a less confusing terminology.

2) The detailed study of specific interactions of proteins and the low molecular compounds reviewed by Prof. Theorell seems to me

to have made, during the early 1930's, a revolution in the intellectual climate of biochemistry. At that time it was fashionable to be exclusively interested in low molecular vitamins, coenzymes, etc., and to regard proteins as « inert colloidal carriers ».

The work of Warburg and others, showing the specific roles of various proteins in the action of these selfsame coenzymes was of decisive importance for diverting this trend of thought, and it is questionable whether the date 1926 given by Prof. Theorell as the birth of modern protein chemistry should not be postponed seven years or so.

M. Pauling. — In the United States the term « chemical bond » (rather than « link ») is used for the force holding two adjacent atoms together in a molecule. Coryell and I, fifteen years ago, used the expression « Haem-linked groups » to refer to groups which are interacting with the haems. By this expression we intended to indicate that an interaction may operate, but not that it necessarily involves covalent chemical bonds. The British usage of the words is somewhat different from the American.

M. Havinga. — 1) It is known, especially by the work of Green and his collaborators, that many of our classical enzymes are linked together in the living cell forming complexes like the cyclophorase system (mitochondria etc.). Whereas in the case of many isolated dehydrogenases the coenzyme (D. N. P. or T. N. P.) dissociates off easily, this D. N. P. cannot be removed from the intact cyclophorase system. Would it still be preferable in these cases to speak of coenzymes in the restricted sense that has been proposed ?

2) It seems doubtful whether the shift in normal redoxpotential brought about by the adsorption of the coenzyme (D. P. N.-D. P. N. 2 H) at the surface of the protein is essential in explaining its function as an intermediate in the oxydation of the substrate (alcohol). The coupled coenzyme is supposed to be in equilibrium with free D. N. P.-D. N. P 2 H. From a kinetic point of view one may gain on the one hand by shifting the apparent normal potential but one looses just as much by the fact that the strong adsorption of reduced D. P. N. makes the transport of the coenzyme difficult (compare the results obtained with A. D. H. from liver !).

M. Theorell. — With regard to Professor Synge's remark, I might mention :

1) That the experiments on the promotion of the resorption of calcium and iron by peptides mentioned on page 307 were made by my colleague Professor O. Mellander in Gothenburg. The references will be included in the literature.

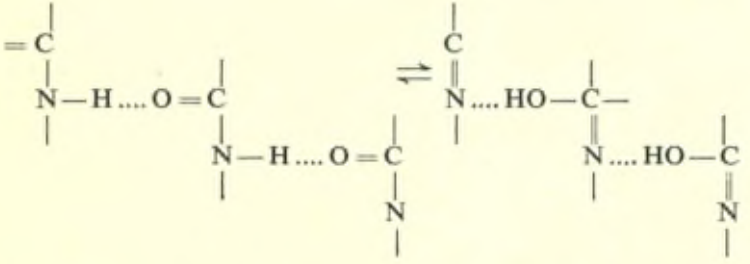
2) That the year 1926 would be too early to be designated as starting-time for modern protein chemistry I would like to say that it may be still too early to estimate the importance of the different discoveries around that time. Dr. Anson discussed it with me the other day whether 1926 is not too late since important discoveries in this field, like the crystallisation of many proteins and the measurement of the molecular weight of hemoglobin by Professor Adair, had been made even before 1926. Personally I think, however, it will be agreed in the future that the important discoveries made around 1926 may justify my suggestion.

As for Professor Havinga's comment, it is quite in accordance with my suggestions concerning the nomenclature that the non-dissociable D. P. N. or T. P. N. in mitochondria or cyclophorase should be called prosthetic groups. Furthermore, I agree with his remark that the redox potentials are not always a limiting factor for the reaction velocity. But in the case of the alcohol dehydrogenase system it is rather obvious that the oxidation of alcohol to acetaldehyde could not possibly proceed rapidly and smoothly at the low potential of the free D. P. N.-D. P. N. H. At physiological pH equal amounts of D. P. N. and D. P. N. H would be in equilibrium with 0.05 % acetaldehyde and 99.95 % alcohol. The coupling of the A. D. H.-protein from liver changes this figures to 10 respectively 90 %, which is good enough because the aldehyde formed is continuously removed through the action of aldehyde oxidase. The so called alcohol dehydrogenase from yeast does not change the redox potential of the D. P. N. H. to any appreciable extent; and in accordance with this it does not work as an alcohol dehydrogenase but as an aldehyde reductase.

M. Anson. — Theorell stated, in his report and in his comments on my report, that the stabilizations of proteins by prosthetic groups and by substrates must have similar mechanisms. I should like

to reverse this statement of Theorell and say that the effects of proteins on prosthetic groups and on substrates must have similar mechanisms. Enzyme-substrate compounds may be considered as typical conjugated proteins, with the substrates as typical prosthetic groups. Thus the conjugated proteins are much more common than they are usually considered to be. And the sort of chemistry discussed by Theorell can be applied to the enzyme-substrate compounds, both the chemistry dealing with the mechanism of combination of protein and prosthetic group and the chemistry dealing with the changes in the properties of the prosthetic group as a result of its combination with protein.

M. Linderström-Lang. — To what an extent does the protein molecule of an enzyme as a whole contribute to the catalytic activity ascribed to the prosthetic group? To what an extent has one to consider theories put forward by Szent Gyorgyi, Wirtz, Evans and Gergely or Geissmann and according to which the catalytic power of enzyme proteins is ascribed to a resonance system of the type



where a « flow » of π electrons may be directed to the active center (prosthetic group).

M. Pauling. — I do not think that the system of amide groups forming hydrogen bonds, as in the α-helix, can serve in an especially effective way to conduct electrons or energy over long distances. There would, of course, be a transmission of effects along the system, but probably of less significance than along a conjugated chain of alternating single and double bonds, and comparable to that along a saturated chain, which dies out rapidly.

TABLE DES MATIÈRES

Notice sur l'Institut International de Chimie Solvay.

Rapports et Discussions.

LES PREMIERS CONSEILS DE CHIMIE SOLVAY

Premier Conseil, 1922 : *Cinq Questions d'Actualité.*

Rapports de MM. F. Soddy, F.-W. Aston, J. Perrin, W.-H. Bragg, W.-J. Pope, T.-M. Lowry, Ch. Mauguin et A. Job.

Deuxième Conseil, 1925 : *Structure et Activité chimique.*

Rapports de MM. W.-B. Hardy, W.-L. Bragg, J. Duclaux, T.-M. Lowry, F. Swarts, M. Tiffeneau et Okékhoff, J. Perrin, A. Job, E.-K. Rideal, E.-F. Armstrong et P. Hilditch, Ch. Moureu et Ch. Dufraisse, H.-E. Armstrong et H. von Euler.

Troisième Conseil, 1928 : *Question d'Actualité.*

Rapports de MM. E.-K. Rideal, W. Mund, A. Berthoud, S. Price, F.-G. Donnan, P. Girard, G. Urbain, N.-V. Sidgwick, P. Walden et P. Karrer.

Quatrième Conseil, 1931 : *Constitution et Configuration des molécules organiques.*

Rapports de MM. W.-H. Mills, J. Boeseken, H. Staudinger, J. Timmermans, Mme Ramart-Lucas, MM. S. Sugden, R. Kuhn, B. Holmberg, R. Robinson et W. Schlenk.

Cinquième Conseil, 1934 : *L'Oxygène, ses réactions chimiques et biologiques.*

Rapports de MM. M. Bodenstein, J.-A. Christiansen, H. Wieland, W.-P. Jorissen, W.-H. Bone, Ch. Dufraisse, R. Wurmser, O. Warburg et O. Meyerhoff.

Sixième Conseil, 1937 : *Les Vitamines et les Hormones.*

Rapports de MM. G. Bertrand, P. Karrer, A. Szent-Györgyi, A. Windaus, H. von Euler, W.-N. Haworth, L. Ruzicka, E. Laqueur, F. Kögl et E.-C. Dodds.

Septième Conseil, 1947 : *Les Isotopes.*

Rapports de MM. F. Joliot, K.-T. Bainbridge, C.-K. Ingold, M. de Hemptinne, F.-A. Paneth, A. Langseth, G. de Hevesy, M. Calvin et D. Rittenberg.

Huitième Conseil, 1950 : *Le Mécanisme de l'Oxydation.*

Rapports de MM. M.-G. Evans, W.-A. Waters, P. Laffite et E. Freling, C. H. N. Chamberlain et A.-D. Walsh, M.-S. Kharasch, A. Schönberg, P. Fleury et J. Courtois, A.-G. Ogston, G. Schwarzenbach et M. Pourbaix.