Reprinted from the 27th Symposium of the Society for Developmental Biology
DEVELOPMENTAL BIOLOGY SUPPLEMENT 2, 1968
Copyright © 1968 by Academic Press Inc. *Printed in U. S. A.*

DEVELOPMENTAL BIOLOGY SUPPLEMENT **2**, 1–20 (1968)

I. SELF-ASSEMBLY OF MACROMOLECULAR STRUCTURES

Spontaneous Formation of the Three-Dimensional Structure of Proteins

CHRISTIAN B. ANFINSEN

Laboratory of Chemical Biology, National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Bethesda, Maryland

INTRODUCTION

Our major consideration in this symposium will be the emergence of order during cellular differentiation and growth. The concept "emerging order" implies an organized, genetically complex process taking place over a reasonably extended stretch of time. In contrast, the restatement of linear genetic information in the form of three-dimensional protein structure results from a rapid and spontaneous interaction of amino acid side chains with each other, with the completed polypeptide backbone, and with the environment, without the necessity for additional genetic information (Anfinsen, 1967; Epstein *et al.*, 1963). The achievement of this unique geometry might be visualized as a rather helter-skelter process. An almost infinite number of sets of interactions are possible as an extended polypeptide chain coils upon itself (Fig. 1). If the process of folding involved even a small fraction of this number of conformational states, the specific folding of the chain could clearly require considerable time. It is probable that the rapidity of folding is made possible through the formation of one or more "nucleation sites" by side chain interactions that would predispose, during subsequent interactions, to the tertiary structural characteristics of the native structure. The only obvious driving force during this approach to native conformation is the selection of progressively more stable conformations with ultimate fixation of geometry in the form possessing the most favorable free energy of conformation, the native protein. Thus, unlike the complex predetermined pattern of successive changes occurring during differentiation, the cell must rely, in its first steps of development, on a relatively *random*

© 1968 by Academic Press Inc.

process but involving *explicit* information—the amino acid sequence of a polypeptide chain.

It has been suggested (Phillips, 1967) as an alternative mechanism that a polypeptide chain may progressively assume a three-dimensional conformation similar or identical to that which it occupies in the completed protein molecule, as synthesis proceeds from the NH_2-terminus toward the COOH-terminal end of the chain. However, the weight of evidence available at the present time, some of which I shall mention

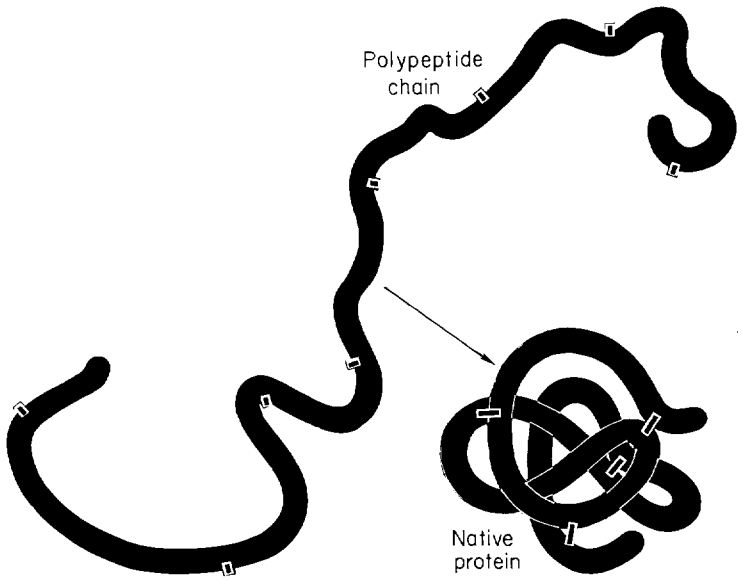

FIG. 1. Schematic drawing showing the conversion of an extended polypeptide chain to a native protein. During this oxidative process, sulfhydryl groups are paired to form disulfide bonds, and amino acid residues, widely separated in a linear sense, are brought into spatial proximity to form an active center.

below, appears to be consistent with a process in which tertiary structure appears only upon *completion* of translation of the genetic quantum of information.

With the exception of the synthesis of certain RNA molecules, the information in a chain is expressed in a form useful to a cell as linear "bursts" of polypeptide chains. Each chain represents the raw material for a function that is performed by the corresponding protein molecule. Evolution in its simplest form has consisted of the continuous selection

of organisms on the basis of the adequacy of the summation of their proteins to constitute a cell system favorable to self-reproduction under the current ecological situation. The sequences of the polypeptide chains that are synthesized are so constituted that they assume, in a spontaneous manner, unique geometric shapes that are endowed with the function in question.

Most of our information has come from a study of proteins that contain disulfide bonds as cross-links and the reversibility of refolding has been tested by examining the reformation of correct pairs of half-cystine residues, together with the restoration of biological activity and various physicochemical properties. The statistics of the situation are shown in Table 1, which lists the number of possible ways in which a given number of half-cystine residues can combine

TABLE 1

THE NUMBER OF WAYS IN WHICH $2n$ SULFHYDRYL GROUPS CAN COMBINE TO FORM j DISULFIDE BONDS

Number of bonds	Number of combinations
1	1
2	3
3	15
4	105
5	945
6	10395
7	135135
8	2027025
9	34459425
10	654729075
11	13749310575
12	316234143225
13	7905853580625
14	213458046676875
15	6190283353629375
16	191898783962510625
17	6332859870762850625
18	221643095476699771875
19	8220079453263789155937 5
20	319830986772877770815625
21	13111307045768798860344 0625
22	563862029680583509947946875
23	25373791335626257947657609375
24	1192568192774434123539907640625
25	58435841445947272053455474390625

$$N_{2n}{}^{j} = \frac{(2n)!}{2^{j}(2n-2j)!j!}$$

to form SS bonds upon oxidation. These numbers show, for example, that in the case of the γ-globulin molecules, the random chance of forming the correct 23 SS bonds from the available 46 half-cystine residues is 1 in 2×10^{28}. In the case of pancreatic ribonuclease, which contains 8 half-cystine residues, 105 possible sets of 4 SS bonds can be made, only one of which is the native structure. Since much of the evidence for the spontaneity and uniqueness of polypeptide folding has been summarized earlier, I shall present here only a schematic picture. Figure 2 depicts the renaturation of what we have called a

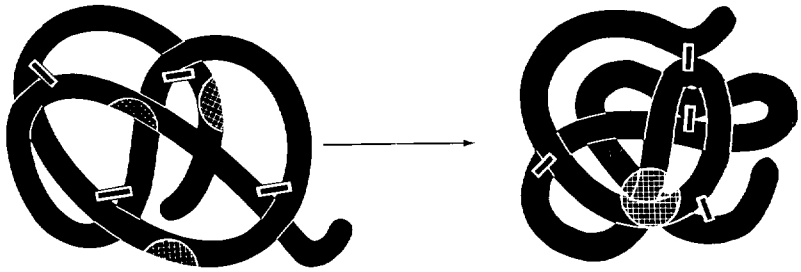

FIG. 2. The spontaneous conversion of a randomly crosslinked protein derivative to the native form under conditions favoring disulfide interchange. Structural regions of the molecule that are involved in the active center are indicated by crosshatching.

"scrambled" ribonuclease molecule. After complete reduction of the 4 disulfide bonds in the native protein, the reduced random chain was allowed to reoxidize under conditions leading to a random mixture of disulfide bonds (Haber and Anfinsen, 1962), shown diagrammatically in the upper portion of the figure. The thermodynamic instability of this scrambled mixture is demonstrated by the observation that exposure to conditions favoring disulfide interchange induced rapid rearrangement of the disulfide bonds with the formation in almost quantitative yields of the native enzyme with its correct SS pairs. By using as a catalyst for the interchange process an enzyme from microsomal membranes that we have recently isolated, the renaturation process can be made to occur *in vitro* (Fuchs *et al.*, 1967) at a rate which is quite consistent with the estimated length of time required for the synthesis of a ribonuclease molecule *in vivo*, namely about 2 minutes (Dintzis, 1961; Canfield and Anfinsen, 1963). This experimental result militates against the concept of obligatory progressive folding during

the NH_2-terminal to COOH-terminal synthesis of the chain since the scrambled collection of isomers is devoid of the features of tertiary structure that one finds in the native enzyme.

We have recently carried out some pertinent experiments on the thermodynamic stability of the RNase derivative, RNase-S (Kato and Anfinsen, unpublished results). This material, prepared by the controlled cleavage of a single bond between residues 20 and 21 in bovine pancreatic ribonuclease by the enzyme, subtilisin, may be separated into its two noncovalently bonded components, RNase-S-protein and RNase-S-peptide (Richards and Vithayathil, 1959). The former, containing all the four disulfide bonds of the native protein, is inactive without the addition of the peptide moiety. To test whether the S-protein portion contained sufficient information to determine the specific folding that would lead to proper pairing of the eight half-cystine residues, samples were subjected to conditions of disulfide interchange under catalysis by the rearranging enzyme from microsomes mentioned above. This enzyme, after prereduction of its single essential SH group, will catalyze disulfide rearrangement without need for added mercaptoethanol or other SH reagent. As summarized in Fig. 3, addition of the enzyme to S-protein solutions caused rapid loss of the capacity of the S-protein to be activated by addition of 1.3 equivalents of S-peptide. Peptide maps of pepsin digests indicated the presence of random SS pairing. [The residual activity may represent material which does not contain all the normal four SS bonds of ribonuclease. The recent observations of Neumann *et al.* (1967) on the preparation of a fully active derivative of RNase containing only two intact disulfide bonds indicate that two of the native disulfide linkages in this protein are superfluous from the standpoint of *in vitro* activity. Consistent with this view is the observation that fully reduced S-protein, when allowed to oxidize in the absence of S-peptide, with complete conversion of its 8 SH groups to 4 SS bonds, yields low levels of active material (Kato, unpublished; Haber and Anfinsen, 1961).] Upon addition of S-peptide to the largely inactivated S-protein solution, the bulk of the activity was regenerated.

Similar conclusions may be drawn from parallel experiments in which the formation of intermolecular, disulfide bonded aggregates of S-protein was studied in the presence and absence of S-peptide (Fig. 4) by turbidity measurements. Once again, the information contained in the S-peptide portion of RNase-S was required to deter-

mine the native structure which, by inference, must represent the most thermodynamically stable form.

Experiments similar to those I have just described for ribonuclease and S-protein have been carried out on a wide variety of protein molecules, both large and small, and the phenomenon appears to be a general one (Anfinsen, 1967). Perhaps the most dramatic example is

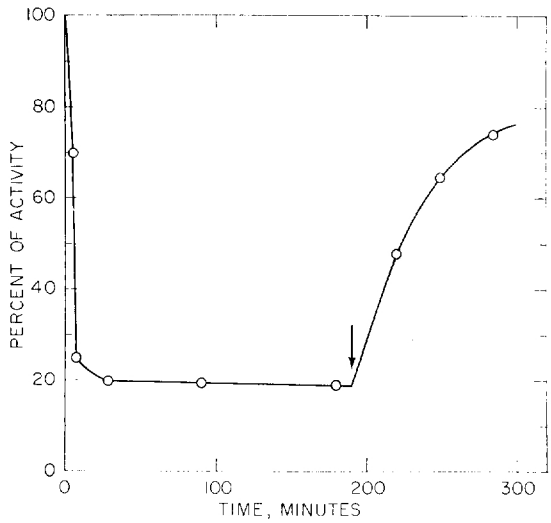

FIG. 3. Inactivation and disulfide interchange of native RNase-S-protein catalyzed by prereduced interchange enzyme (I. Kato and C. B. Anfinsen, unpublished results; Fuchs *et al.*, 1967). The arrow indicates the time of addition of RNase-S-peptide (1.3 equivalents relative to S-protein) to the reaction mixture. RNase-S-peptide (1.3 equivalents) was added to aliquots taken prior to the time marked by the arrow, and the mixtures were assayed for RNase activity.

given by recent studies by Freedman and Sela (1966) on γ-globulins. Both Haber (1964) and Whitney and Tanford (1965) showed that the $(Fab)_2$ fragment of 7 S γ_G antibodies, produced by papain digestion, could be subjected to full reduction of SS bonds with subsequent restoration of significant levels of specific antibody activity upon reoxidation. Freedman and Sela were able to repeat such experiments using *undegraded,* native antibody molecules by the trick of massive polyalanylation of the ϵ-amino groups of the purified rabbit-antibovine serum albumin. The addition of DL-polyalanyl side chains on proteins and polypeptides has been shown, in several instances, to confer much

greater solubility on the products than that shown by the unpeptidylated material. The 23 disulfide bonds of the protein (whose immunological activity was unchanged by the peptidylation) could then be reduced without formation of the otherwise insoluble, reduced heavy chain, a product of reduction that had been avoided by use of papain fragments in the earlier experiments. The reduced forms of the soluble, polyalanylated light and heavy chains were reoxidized separately and finally recombined through oxidative formation of the

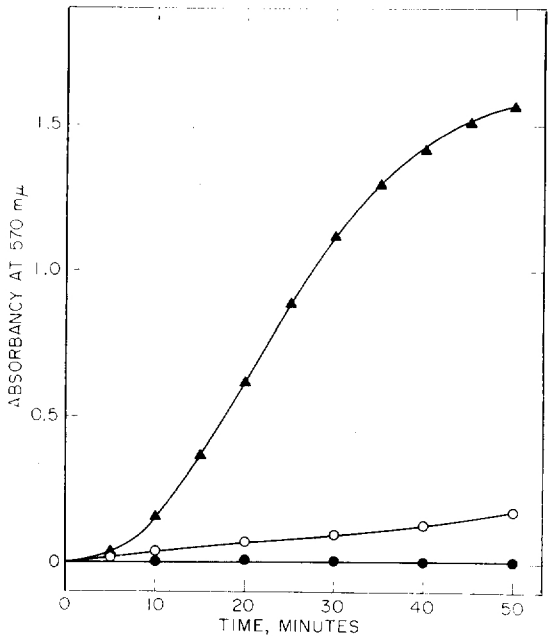

FIG. 4. Disulfide interchange in S-protein as evidenced by turbidity formation. S-protein (1 mg/ml) was incubated in 10^{-3} *M* β-mercaptoethanol, 0.1 *M* Tris buffer, pH 7.4. △——▲; with interchange enzyme (7 μg/ml); ○——○, without enzyme; ●——●, with enzyme (7 μg/ml) and 1.3 equivalents of S-peptide relative to S-protein.

interchain SS bonds to yield regenerated γ-globulin with over 50% of the initial antibody activity. The unlikelihood of this process, unless completely determined by amino acid sequence, is certainly emphasized by the figures listed in Table 1.

For completeness I should mention that certain polypeptide systems

can form native tertiary structures only in the presence of ligands, such as metal ions and prosthetic groups. In the case of Taka-amylase, for example, which contains 9 half-cystine residues, the final formation of the fourth SS bond and the preservation of the remaining single SH group is dependent upon the addition of calcium ions (Friedmann and Epstein, 1967). Similarly, the final native structure of myoglobin is achieved only when heme is added to the slightly "relaxed" apomyoglobin structure (Schechter and Epstein, 1968; Harrison and Blout, 1965).

FUNCTION AND GEOMETRY

The increasing library of sequence data on functionally related proteins has made it extremely likely, simply on the basis of sequence homology, that many groups of these macromolecules have been derived from the same primordial ancestral protein molecule. Furthermore, the crystallographic information available on the heme proteins, myoglobin, and the hemoglobins, indicates that three-dimensional structure has been preserved in the face of very large changes in the details of amino acid sequence. Thus, a particular spatial arrangement of the polypeptide chain has been "imprinted" and a variety of solutions to the geometric problem have been evolved. Although natural selection obviously operates at the level of the organism, this principle of "conservation of geometry" at the protein level seems likely to be a central molecular mechanism in evolution. A stereochemical arrangement consistent with a particular kind of function, once established through chance mutation of a primordial gene, would become established in a line of organisms because of its selective advantage.

Because of such considerations, the problem of determining the nature of the forces that determine and stabilize three-dimensional structure is now a major concern of protein chemists. The role of *hydrophobic side chains* in the internal stabilization of protein structure in solution was examined theoretically by Walter Kauzmann in 1959 (Kauzmann, 1959). Recent crystallographic work has clearly confirmed the predominant location of such side chains within the interior of proteins, secluded from the aqueous environment. The great importance of hydrophobic interaction in the determination of tertiary structure has become even more apparent from considerations by Perutz (1965) and his colleagues (Perutz *et al.*, 1965), Epstein (1964), and others of the amino acid replacements that have occurred

in certain groups of proteins during evolution and as the result of point mutations (for example, in the abnormal hemoglobins). Perutz and his associates point out that, in contrast to the extensive substitution of the less hydrophobic externally situated amino acids in the large series of heme proteins that have been sequenced, a central "core" of nonpolar residues have either remained unchanged or have undergone extremely conservative replacement with residues of closely similar volume and polarity. One must infer that these invariant residues in the sequences are a most important part of the "program" for tertiary structure. Epstein has presented statistics on the heme proteins together with a number of examples of species variants that indicate that replacements generally involve substitution of one amino acid with another of similar polarity. A recent comparison of the *sequence* of rat pancreatic ribonuclease with the three-dimensional structure of bovine pancreatic ribonuclease-S, which I shall discuss in more detail below, offers a particularly compelling set of results in this connection.

We have obtained data in accord with these observations from a study of the influence of changes in the surface stereochemistry and net charge of the ribonuclease molecule on the ability of this protein to regain its native conformation after SS bond reduction and complete denaturation.

As referred to above in regard to γ-globulins, proteins may be reacted with *N*-carboxyamino acid anhydrides at neutral pH to yield derivatives containing polypeptidyl chains on the ϵ-amino groups of the majority of the lysine side chains. Using the *N*-carboxyamino acid anhydride of DL-alanine, eight polyalanyl chains, each containing 5–7 residues of alanine, may be attached to pancreatic ribonuclease (Fig. 5) without loss of enzymatic activity. After reduction of the SS bonds of this derivative in 8 *M* urea and mercaptoethanol, removal of reagents, and exposure of the reduced, random chain to air, oxidation causes essentially complete regeneration of enzymatic activity and of the physical properties characteristic of the starting material. These experiments (Anfinsen *et al.*, 1962; Cooke *et al.*, 1963) indicate that, in spite of a large number of bulky polyalanine chains, the folding of the molecule and the formation of the native pairs of half-cystine residues can proceed normally. The interaction of hydrophobic residues to form the internal structure of the protein can thus proceed effectively in spite of the large change in external stereochemistry. Similar studies have been performed in which amino groups have

been acylated or succinylated with the replacement of positively charged side chains by uncharged acylamino- or negatively charged succinylamino- groups, once again without destroying the capacity of the reduced derivatives to refold correctly (Epstein and Goldberger, 1963).

It is hopeful that the complexity of computer programs now being employed in attempts to calculate tertiary structure of proteins from

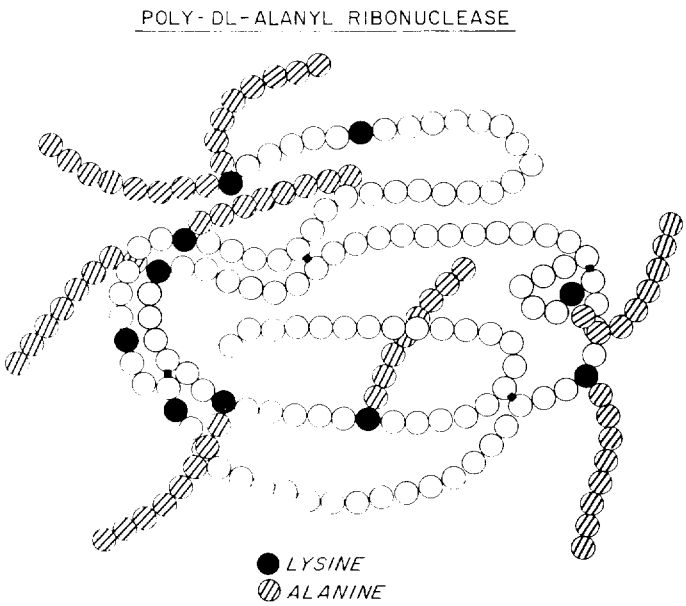

FIG. 5. Schematic representation of a fully active polyalanyl-ribonuclease molecule. The crosshatched circles indicate alanyl residues, attached in chains to ϵ-amino groups.

the information encoded in amino acid sequences, may eventually be simplified when we learn to detect and employ only those portions of the total information that are essential and sufficient. Results such as those on polyalanyl-RNase would certainly suggest that much of the polypeptide structure destined to become external in the native protein may contribute very little to the thermodynamic forces involved in chain folding and stabilization.

Although our catalog of three-dimensional solutions is still quite limited, it would be surprising to find that the structures of the closely

chemically related proteases, chymotrypsin and trypsin, or of the large number of well studied cytochromes *c*, are not extremely similar. The same situation might be expected for egg white lysozyme and the α-lactalbumin of milk whose sequences are remarkably homologous. An interesting analysis of the *sequence* of rat pancreatic RNase (Beintema and Gruber, 1967) has recently been made by Wyckoff, Richards, and their colleagues (see Wyckoff, 1968) in relation to the *three-dimensional structure* of bovine RNase-S. The sequences of these two homologous proteins are shown in Fig. 6. When considered in the

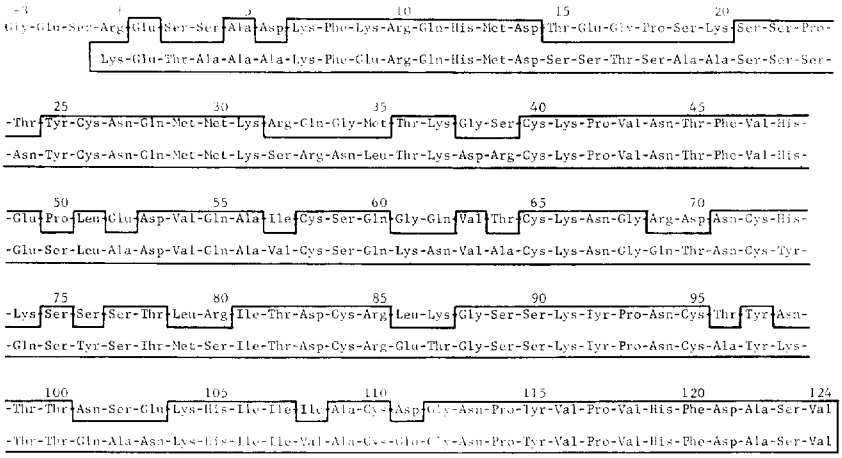

FIG. 6. A comparison of the amino acid sequences of rat (above) and bovine (below) pancreatic ribonucleases. The enclosed area contains the regions of identical sequence (Beintema and Gruber, 1967; Wyckoff, 1968).

context of the bovine geometry, differences in sequences in the rat protein, often occurring in pairs and frequently far separated on the chain, make good sense in terms of structural stabilization. Many of these double replacements appear to permit the retention of interaction between neighboring lengths of the polypeptide chain that form stabilized, structural features of the three-dimensional model. For example, the substitutions of arginine and glutamic acid at positions 80 and 103, replacing the neutral serine-asparagine interaction in the bovine enzyme, may help maintain the stability of a loop in the structure, but now by an electrostatic interaction. Other replacements lead to a conservation polarity or specific net charge in certain areas of the surface. Thus, replacement of the hydrophobically interacting methio-

nine residue 79 in the bovine enzyme with leucine in the rat, involves little change in volume but a definite change in shape. Since the former residue is partly exposed in a pit in the bottom of the three-dimensional model, the change in shape can be accommodated and actually makes room for the extra volume of isoleucine 57 which replaces valine 57 in the bovine protein.

Some of the double changes are less understandable when considered in the context of other experimental data. The pair of conformationally neighboring residues, Lys-61 and Gln-74, in the bovine enzyme became Gly and Lys, respectively, in the rat protein. Local charge is preserved by this set of replacements, but an examination of the three-dimensional model does not suggest any more subtle reason for "conservatism," such as preservation of a stabilizing interaction or the avoidance of a "hole" in the structure. Nevertheless, our studies on polyalanylated RNase, referred to above, show clearly that the ϵ-amino group of lysine-61 may be modified by the addition of a chain of 5–8 alanyl residues without interference with either activity or the capacity of the fully reduced polyalanyl-RNase to refold correctly after complete reduction and denaturation. Such a modification, although preserving net charge, moves the ionized amino group about 20 Å from the position of the original ϵ-amino group. Intracellular requirements of a more complex nature must underlie the genetic changes that lead to double replacements of this sort; it is clear that we have much to learn about the "design" of proteins in relation to function.

EFFECTS OF INTERRUPTION OR MODIFICATION OF GENETIC INFORMATION

Since function is a consequence of precise geometry, spontaneous and correct folding of a polypeptide chain might not occur after tampering with the integrity of the translated genetic information. It is of interest, therefore, to examine the adequacy of the information for folding in multichained proteins after various limited cleavages.

Multichained proteins may be classified as follows:

1. Naturally occurring proteins containing more than one chain resulting from specific *in vivo* cleavage; this group includes, to my knowledge, only two examples—chymotrypsin and insulin.

2. Biologically active multichained molecules derived from single-chained proteins, produced by deliberate experimental cleavage of peptide bonds by protease treatment. This group of man-made deriva-

tives is very small; RNase-S (Richards and Vithayathil, 1959), RNase-E (Klee, 1965), RNase-T (Ooi *et al.*, 1963) (Fig. 7), and nuclease-T, -S, and -C (see Figs. 8 and 9).

3. Naturally occurring multichained proteins formed by disulfide bonding of two or more separately synthesized chains—the immunologically active globulins.

4. Oligomeric proteins, made up of noncovalently aggregated single chains. This very large group includes a variety of intracellular proteins whose multimeric structures permit "allosteric" modifications due to ligand interaction.

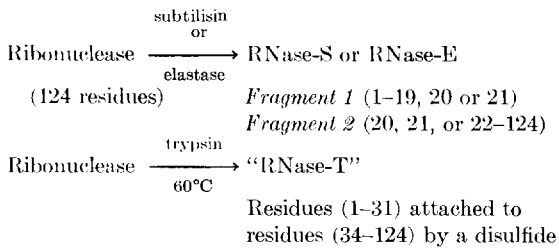

FIG. 7. The limited cleavage of bovine pancreatic ribonuclease with subtilisin, elastase, and trypsin to yield active derivatives. The products produced by elastase and subtilisin may be separated into two chains which may be recombined through noncovalent interactions to yield full activity. In the trypsin product, the two stretches of sequence are held together through an SS bond and, after separation by reduction of this and the other 3 SS bonds, do not recombine correctly upon SH oxidation.

Both examples in the first group have been examined with respect to the stability of their conformations to conditions favoring disulfide interchange (Givol *et al.*, 1965). Whereas the precursor zymogen chymotrypsinogen, a single-chained protein, is quite stable to sulfhydryl reagents and to the action of the disulfide rearranging enzyme mentioned earlier, its product of activation, chymotrypsin, is rapidly inactivated under such conditions, through "scrambling" of its disulfide bonds. One may conclude, therefore, that the information in the three polypeptide chains of the active protease is not sufficient to determine the correct structure and half-cystine pairing that one finds in this "derived" protein.

A similar inactivation and structural disorganization occurs with insulin. This phenomenon led us (Givol *et al.*, 1965) to suggest that insulin, like chymotrypsin, might be synthesized as a "proinsulin" in

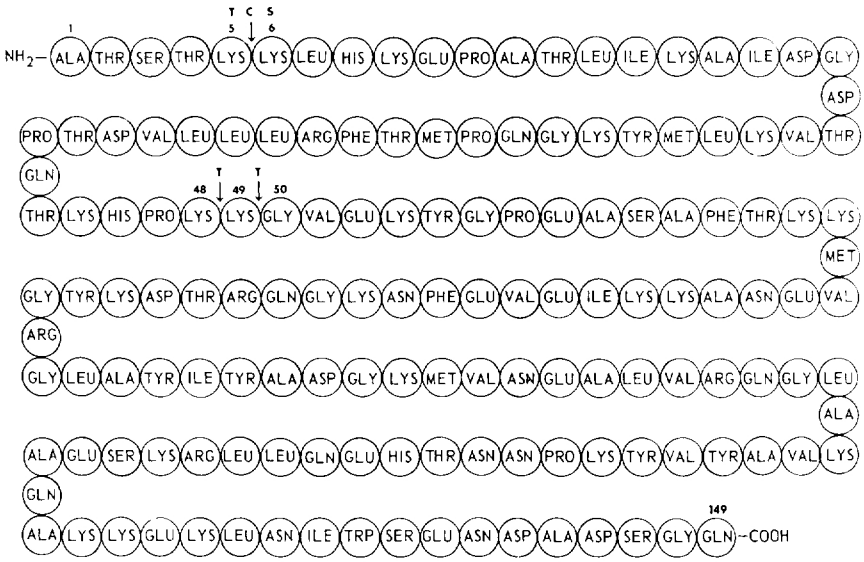

FIG. 8. The amino acid sequence of an extracellular nuclease of *Staphylococcus aureus*. Specific points of cleavage, during digestion in the presence of deoxythymidine-3',5'-diphosphate and calcium ions by trypsin (T), chymotrypsin (C), and subtilisin (S) are indicated by the arrows.

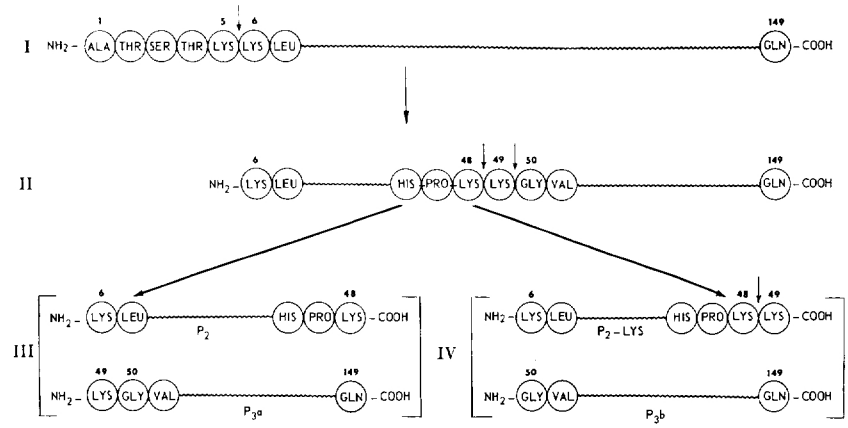

FIG. 9. The formation of "nuclease-T" during the limited trypsin cleavage of staphylococcal nuclease (see also Fig. 8). As discussed in the text, fragments P2 and P3 associate, noncovalently, in solution to form an enzymatically active complex.

which the normal chains are connected through a linking peptide joining the COOH-terminal residue of one with the NH_2-terminus of the other. Following the properly directed pairing of half-cystine residues, the linking peptide might then be removed by a proteolytic process to yield the interchange-prone hormone. The recent discovery by Steiner and his colleagues of such a "proinsulin" molecule (Steiner, 1967), lends strong support to the general idea that thermodynamic instability of the structure of a protein indicates a precursor-product

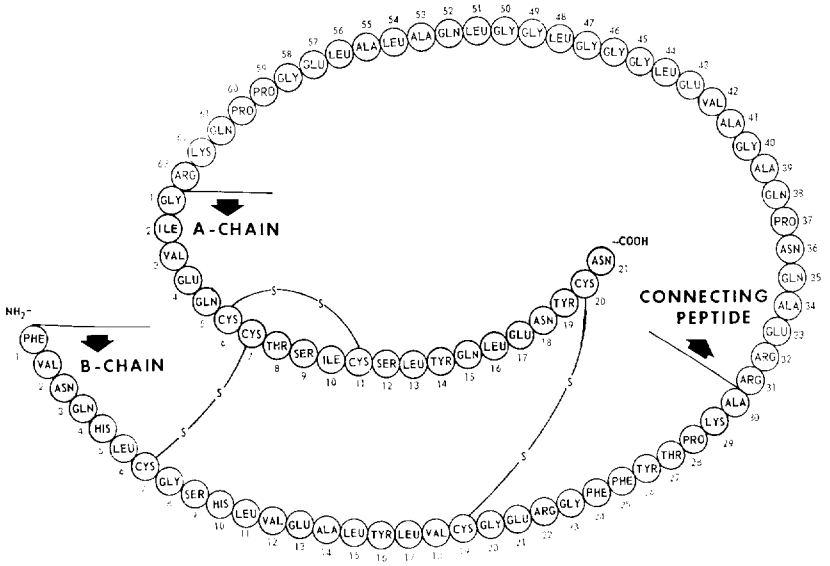

FIG. 10. The structure of porcine proinsulin, including the amino acid sequence of the connecting peptide that joins the B chain to the A chain. Courtesy of Drs. Chance, Ellis, and Bromer, Eli Lilly Co., Indianapolis, Indiana (Chance *et al.*, 1968).

relationship involving deletion of essential information. The structure of the porcine proinsulin is given in Fig. 10 (Chance *et al.*, 1968). It is striking that recombination of the two reduced chains of insulin itself, through disulfide bond formation, can only be made to take place in high yield when certain ingenious chemical manipulations are employed in the process that favor the formation of the desired SS bonds.

We have already discussed the case of RNase-S. This disulfide

bonded protein could be studied by the estimation of the degree of "scrambling" of SS bonds under interchange conditions. A closely related phenomenon has recently been observed with another protein which lacks SS bonds. Staphylococcal nuclease, whose structure (Taniuchi *et al.*, 1967a; Cusumano *et al.*, 1968) is shown in Fig. 8, may be subjected to a limited proteolytic cleavage with trypsin, chymotrypsin, or subtilisin, when the digestion is carried out in the presence of calcium ions and a tightly bound substrate analog, deoxythymidine-3′,5′-diphosphate (Taniuchi *et al.*, 1967b). These ligands stabilize the structure in a manner that restricts peptide bond cleavage to those specific bonds indicated in Fig. 9. The two large fragments resulting from trypsin attack may be separated from one another and, upon mixing in solution, regenerate the full activity of the original nuclease-T. The dissociation constant of the P2-P3 complex is approximately 10^{-7}, indicating a very precise and strong set of noncovalent interactions between the two peptide fragments (Taniuchi and Anfinsen, 1968).

The γ-globulins constitute a class of multichained proteins which are stable to disulfide interchange, in contrast to insulin and chymotrypsin. We have already described the experiments of Haber (1964), Whitney and Tanford (1965), and Freedman and Sela (1966), which clearly show that a precise, antigen-specific structure is determined by the amino acid sequences of the two kinds of component chains. The stability to SS interchange, and the "informational sufficiency" may be explained by assuming that the sequences of the light and heavy chains are coded for by closely related genes and that the complete γ-globulin molecule is a disulfide-linked oligomer rather than a combination of basically different individual chains. Light chains and heavy chains recombine to form active antibody, even after reduction and carboxymethylation of the half-cystine residues involved in interchain bonding (Edelman *et al.*, 1963). The introduction of such disulfide bonds may have been an event in the natural selection of divalent, precipitating antibodies.

Whereas "derived" multichained proteins such as chymotrypsin and insulin are thermodynamically unstable, proteins such as β-galactosidase (Zipser, 1963; Steers *et al.*, 1965; Shifrin and Steers, 1967) (containing four *identical* subunits) and aldolase (Penhoet *et al.*, 1967) or hemoglobin (Kawahara *et al.*, 1965) (with four homologous subunits) are conformationally stable and exhibit revers-

ible denaturation. The latter proteins presumably represent examples of oligomers of closely related chains whose sequences are determined by duplicated homologous genes. Their oligomeric states appear to be involved with mechanisms of metabolic control (Monod *et al.*, 1965).

SUMMARY

Let me summarize the points I have made about the way in which conformational order is achieved at the point of transition from the linear information of the genotype to phenotypic function. First, the amino acid sequence coded for by a genetic cistron in turn codes for a specific three-dimensional structure. This conversion from linearity to spatial organization appears to be a spontaneous process. The native proteins that we find in cells are the polypeptide translations of genetic information, arranged in a form possessing maximum thermodynamic stability under physicological conditions. A particularly important factor in the determination of tertiary structures seems to be the internal and external positioning of hydrophobic and hydrophilic side chains, respectively.

Second, the solution of a functional problem in terms of the three-dimensional arrangement of a polypeptide chain permits subsequent evolutionary changes in sequence only through mutations that are consistent with maintenance of the geometry of the prototypic protein. Although insufficient data now exist, we may expect to find that a particular protein, or class of related proteins, isolated from a variety of species may have very similar three-dimensional structures.

Finally, an examination of the extents to which various natural and "derived" multichained proteins undergo reversible denaturation suggests that interruption or deletion of information in the polypeptide chain of single-chained proteins is generally not permissible, and that only those multichained proteins that are made up of identical or genetically related subunits may be reversible denatured. Studies on the thermodynamic stability of proteins thus reinforces the finding of genetics that a cistron, or gene, determines the more-or-less irreducible unit of function, the correctly folded polypeptide chain.

REFERENCES

ANFINSEN, C. B. (1967). The formation of the tertiary structure of proteins. *Harvey Lectures Ser.* 61, 95–116.

ANFINSEN, C. B., SELA, M., and COOKE, J. P. (1962). The reversible reduction of disulfide bonds in polyalanyl ribonuclease. *J. Biol. Chem.* **237**, 1825–1831.

BEINTEMA, J. J., and GRUBER, M. (1967). Amino acid sequence in rat pancreatic ribonuclease. *Biochim. Biophys. Acta* **147**, 612–614.

CANFIELD, R. E., and ANFINSEN, C. B. (1963). Nonuniform labeling of egg white lysozyme. *Biochemistry* **2**, 1073–1087.

CHANCE, R. E., ELLIS, R. M., and BROMER, W. W. (1968). Porcine proinsulin: characterization and amino acid sequence. *Science* **161**, 165–167.

COOKE, J. P., ANFINSEN, C. B., and SELA, M. (1963). The identification of unreactive amino groups in ribonuclease and their significance to enzymatic activity. *J. Biol. Chem.* **238**, 2034–2039.

CUSUMANO, C., TANIUCHI, H., and ANFINSEN, C. B. (1968). Staphylococcal nuclease (Foggi strain). I. Order of cyanogen bromide fragments and a "fourth" histidine. *J. Biol. Chem.* in press.

DINTZIS, H. M. (1961). Assembly of the peptide chains of hemoglobin. *Proc. Natl. Acad. Sci. U. S.* **47**, 247–261.

EDELMAN, G. M., OLINS, D. E., GALLY, J. A., and ZINDER, N. D. (1963). Reconstruction of immunologic activity by interaction of polypeptide chains of antibodies. *Proc. Natl. Acad. Sci. U. S.* **50**, 753–761.

EPSTEIN, C. J. (1964). Relation of protein evolution to tertiary structure. *Nature* **203**, 1350–1352.

EPSTEIN, C. J., and GOLDBERGER, R. F. (1963). A study of factors influencing the reactivation of reduced egg white lysozyme. *J. Biol. Chem.* **238**, 1380–1383.

EPSTEIN, C. J., GOLDBERGER, R. F., and ANFINSEN, C. B. (1963). The genetic control of tertiary protein structure: studies with model systems. *Cold Spring Harbor Symp. Quant. Biol.* **28**, 439–449.

FRIEDMANN, T., and EPSTEIN, C. J. (1967). The incorporation of ^{3}H-leucine into protein by tetraphenylboron- and citrate-dispersed rat liver parenchymal cells. *Biochim. Biophys. Acta* **138**, 622–624.

FUCHS, S., DELORENZO, F., and ANFINSEN, C. B. (1967). Studies on the mechanism of the enzymic catalysis of disulfide interchange in proteins. *J. Biol. Chem.* **242**, 398–402.

GIVOL, D., DELORENZO, F., GOLDBERGER, R. F., and ANFINSEN, C. B. (1965). Disulfide interchange and the three-dimensional structure of proteins. *Proc. Natl. Acad. Sci. U. S.* **53**, 676–684.

HABER, E. (1964). Recovery of antigenic specificity after denaturation and complete reduction of disulfides in a papain fragment of antibody. *Proc. Natl. Acad. Sci. U. S.* **52**, 1099–1106.

HABER, E., and ANFINSEN, C. B. (1961). Regeneration of enzyme activity by air oxidation of reduced, subtilisin-modified ribonuclease. *J. Biol. Chem.* **236**, 422–424.

HABER, E., and ANFINSEN, C. B. (1962). Side chain interactions governing the pairing of half-cystine residues in ribonuclease. *J. Biol. Chem.* **237**, 1839–1844.

HARRISON, S. C., and BLOUT, E. R. (1965). Reversible conformational changes of myoglobin and apomyoglobin. *J. Biol. Chem.* **240**, 299–303.

KAUZMANN, W. (1959). Some factors in the interpretation of protein denaturation. *Advan. Protein Chem.* **14**, 1–63.

KAWAHARA, K., KIRSHNER, A. G., and TANFORD, C. (1965). Dissociation of human CO-hemoglobin by urea, guanidine hydrochloride, and other reagents. *Biochemistry* **4**, 1203–1213.

KLEE, W. (1965). Ribonuclease-E: an intermediate in the degradation of ribonuclease by porcine elastase. *J. Biol. Chem.* **240**, 2900–2906.

MONOD, J., WYMAN, J., and CHANGEUX, J. P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88–118.

NEUMANN, H., STEINBERG, I. Z., BROWN, J. R., GOLDBERGER, R. F., and SELA, M. (1967). On the non-essentiality of two specific disulfide bonds in ribonuclease for its biological activity. *European J. Biochem.* **3**, 171–182.

OOI, T., RUPLEY, J. A., and SCHERAGA, H. A. (1963). Structural studies of ribonuclease. VIII. Tryptic hydrolysis of ribonuclease A at elevated temperatures. *Biochemistry* **2**, 432–437.

PENHOET, E., KOCHMAN, M., VALENTINE, R., and RUTTER, W. J. (1967). The subunit structure of mammalian fructose diphosphate aldolase. *Biochemistry* **6**, 2940–2949.

PERUTZ, M. F. (1965). Structure and function of haemoglobin. I. A tentative atomic model of horse oxyhaemoglobin. *J. Mol. Biol.* **13**, 646–668.

PERUTZ, M. F., KENDREW, J. C., and WATSON H. C. (1965). Structure and function of haemoglobin. II. Some relations between polypeptide chain configuration and amino acid sequence. *J. Mol. Biol.* **13**, 669–678.

PHILLIPS, D. C. (1967). The hen egg-white lysozyme molecule. *Proc. Natl. Acad. Sci. U. S.* **57**, 484–495.

RICHARDS, F. M., and VITHAYATHIL, P. J. (1959). The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components. *J. Biol. Chem.* **234**, 1459–1465.

SCHECHTER, A. N., and EPSTEIN, C. J. (1968). Spectral studies on the denaturation of myoglobin. *J. Mol. Biol.* in press.

SHIFRIN, S., and STEERS, E., JR. (1967). The effect of urea on subunit interactions of β-galactosidase from *Escherichia coli* K12. *Biochim. Biophys. Acta* **133**, 463–471.

STEERS, E., JR., CRAVEN, G. R., and ANFINSEN, C. B. (1965). Comparison of β-galactosidases from normal ($i^-o^+z^+$) and operator constitutive ($i^-o^cz^+$) strains of *E. coli Proc. Natl. Acad. Sci. U. S.* **54**, 1174–1181.

STEINER, D. F. (1967). Evidence for a precursor in the biosynthesis of insulin. *Trans. N. Y. Acad. Sci. Ser. II* **30**, 60–68.

TANIUCHI, H., and ANFINSEN, C. B. (1968). Steps in the formation of active derivatives of staphylococcal nuclease during trypsin digestion. *J. Biol. Chem.* in press.

TANIUCHI, H., ANFINSEN, C. B., and SODJA, A. (1967a). The amino acid sequence of an extracellular nuclease of *Staphylococcus aureus*. III. Complete amino acid sequence. *J. Biol. Chem.* **242**, 4752–4758.

TANIUCHI, H., ANFINSEN, C. B., and SODJA, A. (1967b). Nuclease-T: an active

derivative of staphylococcal nuclease composed of two noncovalently bonded peptide fragments. *Proc. Natl. Acad. Sci. U. S.* **58**, 1235–1242.

WHITNEY, P. L., and TANFORD, C. (1965). Recovery of specific activity after complete unfolding and reduction of an antibody fragment. *Proc. Natl. Acad. Sci. U. S.* **53**, 524–532.

WYCKOFF, H. (1968). Discussion of comparative structures of rat and beef pancreatic ribonucleases. *Brookhaven Symp. Biol.* in press.

ZIPSER, D. (1963). A study of the urea-produced subunits of β-galactosidase. *J. Mol. Biol.* **7**, 113–121.