Terminology of Protein Structure

I WISH to suggest a revision in the practice of discussing proteins in terms of primary, secondary, and tertiary structures. This terminology, introduced by Linderstrøm-Lang in his Lane Memorial Lectures[1], served well in the few years since it was invented, and has thoroughly permeated the literature. Developments since that time have lessened its usefulness. Arguments to demonstrate this, and suggestions for a more useful general terminology, are the substance of this communication.

One difficulty connected with the terminology is that there is major and widespread disagreement on definitions. For example, in refs 2–4 disulphide bridges are called primary structures, while in refs. 5–7 they are known as tertiary. Disagreement can also be anticipated with other covalent chain-branches and cross-links. Further, there is confusion about secondary structures: some writers accept only helices as secondary structures[6,7,9], others go a step further and include pleated-sheet models[2,4] and some few[5] maintain that a peptide chain segment or loop may have a meaningful non-random structure in the absence of peptide hydrogen bonding; for example, for steric reasons. Furthermore, it is not generally agreed whether tertiary structure is to mean "the way in which primary and secondary structures are packed together in the molecule"[2] and the resultant molecular shape[2,5,7,8]; or whether tertiary structure should direct attention to the multiple forces maintaining the particular folding of α-helices or other chain conformations in a protein[3,9].

The designations, primary, secondary, and tertiary, tend to connote relative importance which can be misleading. I do not disagree with the primacy of covalent bonds in proteins, but there is no compelling evidence that peptide hydrogen bonds contribute more strongly than side-chain interactions to conformational stability; in fact, there are numerous suggestions to the contrary[2,5,8,10].

Finally, there is a sequential connotation in the primary, secondary, tertiary terminology. While present hypotheses of protein biosynthesis support the notion that formation of peptide bonds in an ordered sequence precedes the change to the three-dimensional structure of the protein, there is little knowledge of processes involved in this transformation. The same issue arises in the reversible denaturation of proteins. It is uncertain whether side-chain interactions lead or follow formation of peptide hydrogen-bonds, or whether, as is my view, there may be an advantage in considering the transformation to be a concerted, co-operative process.

To summarize, considerable ambiguity exists at present in the application of the primary, secondary, tertiary terminology. Further objections to this terminology are that it tends to connote relative importance and a sequence of formation of structural elements. Although it is conceivable that general agreement on definitions might be achieved, the other objections remain. I therefore recommend that use of this terminology be discontinued entirely.

In fact, a general move in this direction has already been noted: a recent symposium on protein structure and function[11] ran its full three-day course with scarcely a reference to secondary and tertiary structure. Apparently several other investigators concerned with protein structure have recognized shortcomings in this terminology. If my present arguments do not succeed in persuading writers to abandon the primary, secondary, and tertiary terminology, I hope that it will at least lead some to a greater awareness of its limitations.

I suggest that protein structure can be discussed generally in terms of amino-acid sequences of peptide chains, called chain sequence, and the ways in which peptide chains and side-chains are arranged in space, called chain conformations. (It has already been noted[12] that configuration has often been used synonymously with conformation. This is undesirable, since configuration traditionally has applied to optically asymmetric centres, and should be reserved for this use We should speak of the configuration of an amino-acid residue, and of the conformation of a peptide chain.) It appears desirable to subdivide the latter term into side-chain conformations and peptide chain or backbone conformations, without implying that these two subdivisions are mutually exclusive, and certainly without implying rank or sequence of development. Taken together, chain sequences and chain conformations are sufficient to define a structure, which is the three-dimensional location of all the atoms and covalent bonds of a molecule. The factors determining a structure, such as covalent bond lengths and angles, rotational barriers, steric repulsions, hydrogen bonds, electrostatic forces, 'hydrophobic bonds', etc., should not be confused with the structure itself. No general designation seems necessary for the arrangements of protein sub-units into aggregate molecules, as in insulin, hæmoglobin, collagen, the caseins, etc. However, if such a term must be employed, 'sub-unit array' seems preferable to 'quaternary structure'[13].

Finally, I wish to emphasize that I have intended criticism only for general ways of discussing protein structure; no criticism is meant for terms describing specific structural elements, such as α-helices, pleated-sheets, polyglycine II type structures and constrained non-hydrogen-bonded backbone loops.

Since there is nothing intrinsically proteinaceous in the designations 'chain sequence' and 'chain conformation' (main and side-chain), these terms may be applicable in structural discussions of chain polymers in general.

I am grateful for discussions of this matter with Dr. E. R. Blout, Prof. L. C. Craig, Prof. J. T. Edsall and Prof. D. Waugh.

This investigation was supported in part by a grant from the National Institutes of Health, *USPHS* 1*A*–2558.

D. B. WETLAUFER*

Children's Cancer Research Foundation,
Boston, Massachusetts.

* Present address: Indiana University School of Medicine, Department of Biochemistry, Indianapolis 7, Ind.

[1] Linderstrøm-Lang, K., *Proteins and Enzymes, Lane Memorial Lectures*, 54 (Stanford University Press, 1952).
[2] Kauzmann, W., *Advances in Protein Chemistry*, **14**, 5 (Academic Press, New York, 1959).
[3] Bailey, K., *Proc. Chem. Soc.*, 92 (London, 1960).
[4] Lumry, R., and Eyring, H., *J. Phys. Chem.*, **58**, 110 (1954).
[5] Linderstrøm-Lang, K., and Schellman, J., in *The Enzymes*, edit. by Lardy, Boyer and Myrback, **1**, 443 (Academic Press, New York, 1959).
[6] Leach, S., *Rev. Pure App. Chem.*, **9**, 63 (1959).
[7] Perlmann, G., and Diringer, R., *Ann. Rev. Biochem.*, **29**, 151 (1960).
[8] Fraenkel-Conrat, H., and Ramachandran, L. K., *Advances in Protein Chemistry*, **14**, 203 (Academic Press, New York, 1959).
[9] Holter, H., *C.R. Trav. Lab. Carlsberg, Ser. Chim.*, **32**, xvii (1960).
[10] Blout, E. R., deLoze, Ch., Bloom, S. M., and Fasman, G. D., *J. Amer. Chem. Soc.*, **82**, 3787 (1960).
[11] Symposium on Protein Structure and Function, Brookhaven National Laboratory, June 1960 (proceedings to be published).
[12] Blout, E. R., *Optical Rotatory Dispersion*, edit. by Djerassi, C., 238 (McGraw-Hill, New York, 1960).
[13] Bernal, J. D., *Disc. Farad. Soc.*, **25**, 7 (1958).