Biochemistry

 Volume 22, Number 11 May 24, 1983

EDITORIAL

DOCUMENTATION OF SEQUENCE ANALYSIS OF PROTEINS

Some considerable time has gone by since the current technology of protein sequencing was introduced and applied to a variety of proteins and peptides. Many of the methods have become reliable routine procedures, and their resolving power and accuracy have been repeatedly assessed and verified. Nevertheless, most journals, including *Biochemistry*, still require detailed documentation, to a degree that renders the preparation of the data for publication a laborious and time-consuming undertaking. As a result, many authors take the shortcut of publishing in a short communication the complete protein sequence with minimum, or without any, documentary proof and defer the documentation for a more complete publication at a later time. This practice creates serious problems for the editor of the journal to which the follow-up manuscript is submitted since most journals, like *Biochemistry*, require that the same material not have been published before or be under consideration for publication elsewhere. The author is presented with a quandary: prompt publication is difficult if extensive documentation must be provided, whereas preliminary publication jeopardizes subsequent documentation in print.

In order to resolve this dilemma, we have revised our guidelines so as to lighten the burden of published proof (but not of the proof itself) in the sequence analysis of proteins and peptides. In essence, authors will be required to provide sufficient data to allow reviewers and readers to examine the rigor of the logic whereby the sequenced peptide fragments are arranged into a unique sequence corresponding to the original polypeptide chain(s). To this end, the following information should be provided as a minimum:

1. Description of, or reference to, the source of the protein (or peptide) being sequenced, the procedures used for its isolation, criteria of purity, including biological activity where applicable, and amino acid composition. Comparison of the composition determined by amino acid analysis with that deduced from the final sequence.

2. General strategy of the sequence analysis, including the major fragmentation scheme (e.g., cyanogen bromide cleavage, proteolytic digestion, etc.).

An illustration of the strategy should include the following information: (a) the peptides used to prove the sequence, grouped and identified by source of digest; (b) the position of each peptide in the final sequence; (c) the specific sequences proven within each peptide; (d) the positions of overlaps in the linear assembly of fragments. The design of such an illustration is left to the judgment and discretion of the author. An example of an illustration that includes all of the above documentation can be found in Parmelee et al. (1982) *Biochemistry 21*, 3299, Figure 2. Quantitative results of sequential degradations and identifications can be omitted. However, if the analytical results approach the sensitivity limits of the methods used, it may be necessary to provide representative documentation of the quality of identification of critical amino acid residues in the sequence.

3. Descriptions of the methods used to separate and purify a major set of primary fragments. A group separation pattern should be shown in an illustration with an indication of the peptides present in each peak fraction. A typical illustration of this sort can be found in Kimura et al. (1982) *Eur. J. Biochem. 123*, 40, Figure 2 (top), or in Sasagawa et al. (1982) *Biochemistry 21*, 2568, Figure 5. Such a diagram should enable others to determine in which peak fraction any peptide fragment of particular interest can be located. The principles of subsequent separation steps may be summarized in a table or diagram or described in the text. The separation of minor subdigests need not be detailed but special problems of peptide separation should receive comments. Subdigests of large fragments may need an illustration showing the first separation.

4. The yields and compositions of one set of major fragments, covering the whole polypeptide chain.

Authors should be prepared to submit on request, in raw form, experimental information that might be necessary to convince *reviewers* of the validity of their claims.

We recognize that these guidelines are a compromise and as such are open to criticism. The present document was drafted after extensive consultation with experienced reviewers and practitioners in the field of protein sequencing. While none of them should be held responsible for the final draft, we wish to express our special gratitude to Kenneth A. Walsh for his many valuable suggestions and thank the following colleagues for suggestions and criticism of an earlier version of this document: E. Appella, R. A. Bradshaw, J. D. Capra, B. A. Cunningham, H. F. Deutsch, R. F. Doolittle, D. Eaker, L. H. Ericsson, E. Fink, A. V. Fowler, G. Frank, J. H. Freisheim, A. Henschen, M. A. Hermodson, T. Hofmann, M. W. Hunkapiller, W. Konigsberg, C. Nolan, J. J. Pisano, H. Ponstingl, E. Schiltz, J. Spiess, L. A. Steiner, H. M. Steinman, J. J. N. Tang, K. Titani, T. C. Vanaman, B. Wittmann-Liebold, F. Wold, and I. Zabin.

Hans Neurath, Editor