Reports

New Protein Sequenator with Increased Sensitivity

Abstract. *New developments in automated amino acid sequence analysis permit a 10^4-fold increase in sensitivity (to ~ 10-picomole level) over that initially described by Per Edman in 1967.*

Interferon was discovered in 1957 (*1*) and, despite its potential as both an antiviral and antitumor agent (*2*), has remained uncharacterized for 22 years with regard to its primary structure. This lack of progress arose from the difficulty in purifying and analyzing the very small amounts of interferon available. In this report, we describe advances in automated amino acid sequence analysis that permit us to analyze microgram quantities of scarce polypeptides. In the three accompanying reports (*3*), we use this methodology to provide sequence analysis of several interferon species.

Sequential removal and identification of amino acids from the amino termini of polypeptides by automated Edman degradation is the most widely used method for the primary structural characterization of proteins and peptides (*4*). The original Edman-Begg sequenator was designed for analysis of a few hundred nanomoles (several milligrams) of protein. However, since many interesting proteins are available only in microgram rather than milligram quantities, much effort has been devoted to improving the chemistry and instrumentation for microsequence analysis. One approach that permits the analysis of microgram quantities of polypeptides is the biosynthetic incorporation of radiolabeled amino acids into polypeptide chains (*5*). However, when incorporation of radiolabel is prevented by expense or technical difficulties, as is frequently the case, one generally must have 5 to 20 nmole of protein for analysis with commercially available sequenators (*5*).

In order to increase the sensitivity of sequence analysis to the subnanomole level, we previously combined a series of modifications on a commercial spinning cup sequenator with changes in the Edman chemical procedure, the use of a nonprotein carrier (Polybrene), and analysis by high-performance liquid chromatography (HPLC) of amino acid phenylthiohydantoins (Pth) released by the sequenator (*5*). These modifications allowed extended amino terminal analysis of proteins with less than 1 nmole of protein and as little as 1 nmole of short peptides. Since this approach did not depend on the introduction of radiolabel, it was a general microsequencing technique.

In order to apply automated Edman degradation to even smaller amounts of polypeptides that are not internally radiolabeled, it is necessary to adapt all aspects of the technique—peptide preparation, reagent purification, sequenator performance, and Pth amino acid analysis—to meet the problems posed by microsequencing. These problems include sample loss by washout of the polypeptide from the spinning cup during solvent extractions, side reactions that block the Edman degradation, incomplete removal of reagents and byproducts that contaminate the polypeptide film, low and nonreproducible yields of the amino acid derivatives generated in the sequenator, and insufficient sensitivity in the analytic techniques used to analyze these derivatives.

By combining improvements in each of the elements of the sequencing process mentioned above, we have achieved another tenfold increase in sensitivity over that which we reported previously (*5*). This was made possible by construction at the California Institute of Technology of a new spinning cup sequenator with several novel design features. With this instrument, we can perform extended amino terminal sequence analysis on 20 pmole of proteins and large polypeptides and 200 pmole of shorter peptides.

A variety of peptides and proteins analyzed on the Caltech sequenator illustrate its versatility and sensitivity (Table 1). Thus, we have been able to analyze submicrogram quantities of both proteins and peptides. Figure 1 shows the yields of selected Pth amino acids from a sequenator run on 20 pmole of sperm whale apomyoglobin. These results represent approximately a 10^4-fold increase in the sensitivity of automated Edman degradation analysis of proteins over Edman's original sequenator (*4*). The overall increase in sensitivity for peptide analysis is similar.

We have incorporated in the new spinning cup sequenator many of the design features described previously as modifications of a commercial sequenator (*5*, *6*). To these, we added additional refinements in design and materials that improve the microsequencing capability of the instrument. The most important features incorporated in the new sequenator are summarized in Table 2. Use of the original straightedge-type spinning cup manufactured by Beckman Instruments is noteworthy. It supports a thinner solvent film during solvent extractions

Table 1. Polypeptides analyzed with the Caltech sequenator.

Polypeptide	Residues identified	Quantity	
		Picomoles	Micrograms
Somatostatin from *Escherichia coli*	14/14	1000	1.5
Human carbonic anhydrase fragment	23/23	200	0.5
Aplysia egg-laying hormone	36/36	500	2.0
Dynorphin (long leucine-enkephalin)	17/17	500	1.0
Sperm whale apomyoglobin	36/153	20	0.3
Human C'4* precursor	25/1800	100	20
Mouse gap junction protein	25/90	100	1.0
Human fibroblast interferon	13/150	18	0.4
Human lymphoblastoid interferon	20/150	75	1.4
Mouse fibroblast interferon	24/150	30	0.6

*Fourth component of complement.

0036-8075/80/0201-0523$00.50/0

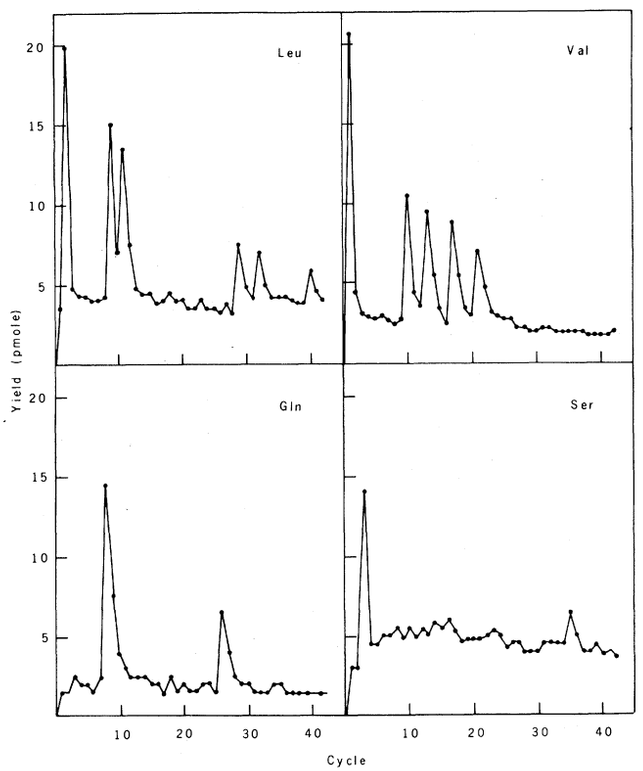
than does the current Beckman undercut cup, and, as a result, efficient extraction of the protein film can be effected with less than half the solvent required with the undercut version. The decreased solvent wash significantly reduces the washout of subnanomole samples of polypeptide.

Analysis of the Pth amino acids produced by the sequenator is performed by HPLC on a Du Pont Zorbax CN column with the use of a sodium acetate buffer, methanol, acetonitrile gradient system similar to that described (*7*). By using highly purified solvents and computer subtraction of the ultraviolet baseline absorption of the HPLC solvent, we can detect less than 1 pmole of the Pth amino acids (Fig. 2).

Preparation of subnanomole quantities of peptides and proteins for sequence analysis requires considerable effort to eliminate contamination by compounds that might interfere either with the Edman degradation or with the HPLC analysis. Removal of nonvolatile buffer salts is particularly important. With small peptides, we generally desalt the sample by gel filtration on a small column of Bio-Gel P-2 in 1*N* acetic acid or Sephadex G-10 in 1*N* NH_4OH and subsequent lyophilization. Larger peptides and proteins can be desalted by dialysis against three changes of 0.15*M* NaCl containing 0.1 percent sodium dodecyl sulfate (SDS; Bio-Rad electrophoresis grade) and then against three changes of 0.02 percent SDS (12 hours between each change). If the proteins have been eluted from gels

Table 2. Novel design features of Caltech sequenator.

Component	Benefit
Reagent and solvent delivery valves: pneumatic-actuated, Teflon and Kel-F manifold, leakproof, zero-dead-volume	Prevents cross contamination of reagents and solvents
Spinning cup: straight rather than undercut wall	Improved extraction of polypeptide film with less solvent
Vacuum system: high vacuum solenoid valves, liquid nitrogen trap	More efficient removal of volatile contaminants from polypeptide film
Vacuum and pressure seals: Teflon and Kel-F rather than rubber O rings	Minimizes seal failure rate and eliminates contamination from rubber elastomers
Inert atmosphere: argon purified by inline Gettering furnace to < 0.1 ppm of O_2 and H_2O	Minimizes oxidative and hydrolytic side reactions
Pth conversion system: automated with the use of 25 percent CF_3COOH	Provides high and reproducible phenylthiohydantoin yields
Programmer: preset separate cleavage times for glutamine and proline residues	Minimizes blocking cyclization of amino terminal glutamine residues and incomplete cleavage of amino terminal proline residues



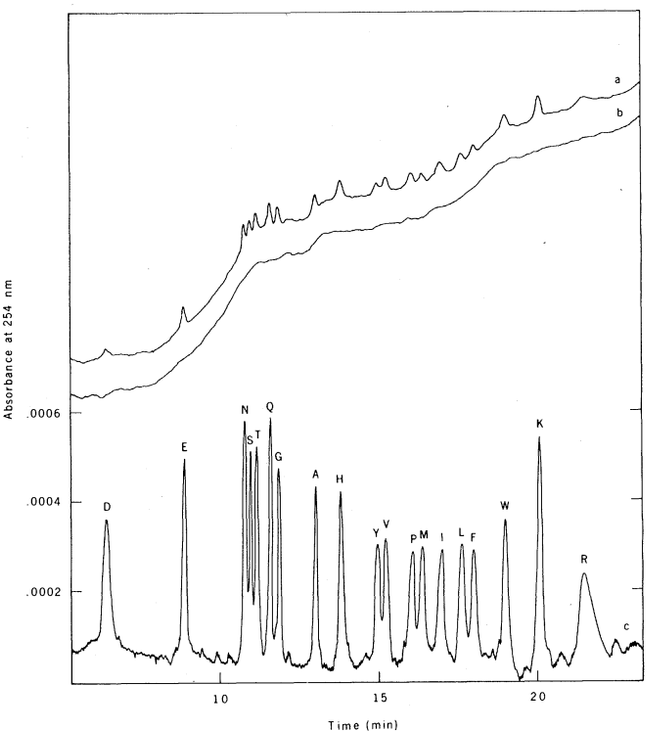


Fig. 1 (left). Yields of Pth-Leu, -Val, -Gln, and -Ser from an automated NH_2-terminal analysis of 20 pmole of sperm whale apomyoglobin. A sample (70 percent) from each cycle was analyzed by HPLC, peaks were quantified by comparison of peak heights with values for a standard Pth amino acid mixture, and yields were normalized to 100 percent injection. Leu is identified at cycles 2, 9, 11, 29, 32, and 40; Val at cycles 1, 10, 13, 17, and 21; Gln at cycles 8 and 26; and Ser at cycles 3 and 35. The repetitive cycle yield averaged 93 ± 2 percent. Abbreviations: Leu, leucine; Val, valine; Gln, glutamine; Ser, serine. Fig. 2 (right). High-sensitivity HPLC analysis of Pth amino acids. Ten microliters of a CH_3CN solution containing 5 pmole of 19 Pth amino acids was analyzed on a Du Pont Zorbax CN column with the gradient system described in (*7*). A background chromatogram (b) was subtracted from the sample chromatogram (a) to allow peak height expansion (c) for improved visualization and measurement. Subtraction was performed with a laboratory automation computer (Hewlett-Packard model 3354).

after SDS-polyacrylamide gel electrophoresis, the eluate should be centrifuged at 20,000 rev/min (Sorvall SS-34 rotor) for 20 minutes to remove particulate matter before dialysis. Coomassie blue staining of the gels to locate protein bands does not interfere with subsequent sequenator analysis.

New technologies such as the improved amino acid sequencing method described above lead to new research opportunities. With the greater sensitivity provided by this technique, we now can obtain amino acid sequence information on both proteins and peptides with submicrogram (picomole) quantities. This sensitivity should permit analysis of biomedically relevant molecules—such as the interferons—that can only be obtained in microgram quantities, and this ability opens possibilities for further study of these molecules. For example, knowledge of the amino acid sequence permits the synthesis of corresponding DNA probes and opens the possibility of new strategies for isolating genes, such as those for interferons, that express low levels of messenger RNA's (*8*).

MICHAEL W. HUNKAPILLER
LEROY E. HOOD
Division of Biology, California Institute of Technology, Pasadena 91125

References and Notes

1. A. Isaacs and J. Lindenmann, *Proc. R. Soc. London Ser. B* **147**, 258 (1957).
2. T. Merigan, *N. Engl. J. Med.* **300**, 42 (1979).
3. E. Knight, Jr., M. W. Hunkapiller, B. D. Korant, R. W. F. Hardy, L. E. Hood, *Science* **207**, 525 (1980); K. C. Zoon, M. E. Smith, P. J. Bridgen, C. B. Anfinsen, M. W. Hunkapiller, L. E. Hood, *ibid.*, p. 527; H. Taira, R. J. Broeze, B. M. Jayaram, P. Lengyel, M. W. Hunkapiller, L. E. Hood, *ibid.*, p. 528.
4. P. Edman and G. Begg, *Eur. J. Biochem.* **1**, 80 (1967).
5. For a review of these methods, see M. Hunkapiller and L. Hood, *Biochemistry* **17**, 2124 (1978).
6. B. Wittmann-Liebold, *Hoppe-Seyler's Z. Physiol. Chem.* **354**, 1415 (1973).
7. N. Johnson, M. Hunkapiller, L. Hood, *Anal. Biochem.* **100**, 335 (1979).
8. B. Noyes, M. Mevarech, R. Stein, K. Agarwal, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1770 (1979).
9. Supported by a gift from the Ben Weingart Foundation.

29 November 1979

Human Fibroblast Interferon: Amino Acid Analysis and Amino Terminal Amino Acid Sequence

Abstract. *The purification of human fibroblast interferon has been simplified to a two-step procedure consisting of affinity chromatography on Blue Sepharose and sodium dodecyl sulfate polyacrylamide gel electrophoresis. A preliminary amino acid composition and the sequence of the 13 amino-terminal residues of homogeneous interferon prepared by this method is reported.*

Since the discovery of interferon, its purification and chemical characterization have been primary goals of interferon research. Although their attainment has been slow because of the small quantities of interferon proteins available, purification to homogeneity has now been achieved with some interferons. However, only microgram quantities have been available for characterization—human fibroblast interferon (*1*, *2*), human lymphoblastoid interferon (*3*), human leukocyte interferon (*4*), mouse interferon (*5*)—and only limited structural information has been acquired (*4*, *6*).

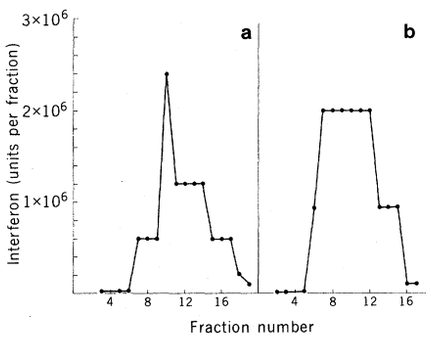


Fig. 1. (a) Fractionation of crude interferon on a large column of Blue Sepharose. Elution of interferon with 50 percent ethylene glycol in column buffer begins at fraction 1. (b) Small Blue Sepharose column. Fractions 7 to 17 in (a) were pooled, passed through the small column, and eluted with 50 percent ethylene glycol in column buffer (fractions 1 to 20).

A thorough understanding at the molecular level of the numerous phenomena that are caused by interferon in cells in culture and in animals will not be possible until the elucidation of primary and secondary structures of the interferon proteins is achieved. This structural information will permit (i) comparison of amino acid sequences of interferons from various cell types and animal species, (ii) identification of the polypeptide segments involved in binding to interferon-specific cell-surface receptors, and (iii) chemical synthesis of interferons.

We now report an improved procedure for the purification of human fibroblast interferon that can be used to provide enough protein for structural studies. Using the automated protein microsequencing technique described in (*7*), we have determined the sequence of the 13 amino acid residues at the amino terminus of the interferon prepared by this method. We also report a preliminary amino acid composition of the protein.

Human diploid fibroblast cells (FS-4) were cultured and interferon was produced (*1*). Interferon was assayed by a microtechnique (*8*) with vesicular stomatitis virus as the challenge virus. Interferon units are given in National Institutes of Health human fibroblast interferon units.

The crude interferon, 10 to 15 liters produced in the absence of serum, was made 1*M* in NaCl and passed at room temperature through a column (4 by 10 cm) of Blue Sepharose (Pharmacia, Inc.) equilibrated with 0.02*M* sodium phosphate buffer, *p*H 7.2, containing 1*M* NaCl. The interferon was retained whereas more than 95 percent of the total protein passed through the column. The interferon was eluted with a mixture of the column buffer and ethylene glycol (1:1), and each fraction was diluted immediately with 0.5 volume of the buffer (Fig. 1a). Fractions containing interferon activity were pooled, diluted with two volumes of the column buffer, and passed through a small (1 by 6 cm) Blue Sepharose column for concentration. The interferon was eluted as described above (Fig. 1b).

Fractions containing interferon were pooled, dialyzed against 1 m*M* tris-HCl,

Table 1. Amino acid composition of human fibroblast interferon.

Amino acid	Composition	
	Mole percent	Residues per 20,000 daltons
Asp	11.1	18.9
Thr	4.0	6.8
Ser	6.2	10.5
Glu	15.9	27.0
Pro	1.6	2.7
Gly*	4.6	7.8
Ala	5.9	10.0
Cys†	1.0	1.7
Val	3.5	6.0
Met	1.7	2.9
Ile	5.3	9.0
Leu	12.0	20.4
Tyr	4.4	7.5
Phe	5.5	9.4
His	2.9	4.9
Lys	6.8	11.6
Arg	6.4	10.9
Trp‡	0.6	1.0

*Includes correction for free glycine present in unhydrolyzed protein. †Determined after performic acid oxidation. ‡Determined after hydrolysis with mercaptoethanesulfonic acid.

0036-8075/80/0201-0525$00.50/0