

Direct Microsequence Analysis of Polypeptides Using an Improved Sequenator, a Nonprotein Carrier (Polybrene), and High Pressure Liquid Chromatography[†]

Michael W. Hunkapiller[†] and Leroy E. Hood*

ABSTRACT: We have combined the use of a nonprotein carrier (Polybrene), high pressure liquid chromatography, and modifications in Edman chemistry with the improvements of a commercial spinning cup sequenator suggested by Wittmann-Liebold [Wittmann-Liebold, B. (1973) *Hoppe-Seyler's Z. Physiol. Chem.* 354, 1415] to analyze amino acid phenylthiohydantoins obtained from automated Edman degradation of microquantities of polypeptide directly without the use of radiolabel. This approach has allowed us to determine the sequence of the N-terminal 47 residues of sperm whale myo-

globin starting with 200 pmol of protein, 77 residues of an antibody light chain with 5 nmol of protein, and 54 residues of an antibody heavy chain with 8 nmol of protein. In addition, we completely sequenced a hydrophobic 14-residue peptide at the 1.5-nmol level. Our technique of direct analysis for microsamples is capable of providing routine, extended N-terminal sequence analysis for nanomole and subnanomole levels of polypeptides and proteins, and it also is applicable to analysis of more classical sample quantities.

An important new phase of protein chemistry is underway with the attempt to determine the amino acid sequence of very small amounts of protein. The microsequence analysis of membrane, cytoplasmic, and nuclear protein components will be an essential element in extending our understanding of a variety of important problems including cell-surface recognition phenomena, the structure-function relationships of cytoskeletal elements such as actin, desmin, and tubulin, and the structure-function relationships of various chromosomal proteins. In order to provide a routine microsequencing approach for such applications, several research groups have devoted considerable effort to improving the chemistry and instrumentation of automated Edman degradation (Edman & Begg, 1967). These efforts have centered on improvements employing classical analytic techniques used in macrosequencing experiments and on the introduction of radioactive labels into the proteins themselves or into the amino acid derivatives derived in the sequenator process.

Edman & Begg (1967) relied upon thin-layer chromatography on silica gel plates for the separation and identification of amino acid phenylthiohydantoins. However, most quantitative data are now obtained using gas chromatographic analysis of these derivatives (Pisano et al., 1972) or amino acid analysis after back hydrolysis to the parent amino acids (Smithies et al., 1971; Mendez & Lai, 1975). Usually, some combination of the three techniques is necessary for the identification of all amino acids, and even then difficulties arise in the identification of tryptophan, cysteine, serine, threonine, asparagine, and glutamine, which are obtained in low yields or are totally destroyed either by conversion to the amino acid phenylthiohydantoins or by the back hydrolysis. As a result, these analytic techniques generally require protein samples of 50 to 100 nmol or more for positive identification of all amino acids. However, since some of the more stable amino acid de-

rivatives can be analyzed for smaller quantities of protein, the original Edman and Begg sequenator program designed for a few hundred nanomoles of protein has been modified to prevent losses of smaller quantities of protein from the sequenator spinning cup by introduction of polymeric carriers (Niall et al., 1974; Tarr et al., 1978; Klapper et al., 1978) or reduced coupling buffer concentrations that minimize the need for solvent extractions (Brauer et al., 1975; Crewther & Inglis, 1975). This has permitted limited, partial sequencing of proteins at the 5–25 nmol level. Further increases in the sensitivity of these techniques have been limited, however, by contaminants of the amino acid derivatives introduced by the sequenator process and by remaining problems with sample washout from the sequenator.

Attempts to exploit the sensitivity of radioactive analytic techniques have, in some cases, overcome some of these problems with the classical, direct analysis methods. *Intrinsic* (internal) labeling of the protein itself has been achieved by use of short term tissue culture procedures that readily incorporate several radiolabeled amino acids into newly synthesized polypeptides (Silver & Hood, 1976; Silver, 1977), by use of longer term tissue culture techniques employing a mixture of radiolabeled amino acids and Krebs cycle intermediates (Ballou et al., 1976), or by cell free translation of mRNA with radiolabeled amino acids (Burstin & Schecter, 1977). In the former, sequenator analysis of the partially labeled proteins results in partial amino acid sequence data with the unlabeled residues being registered as blanks in the sequence, while in the latter cases complete sequence information is theoretically possible. These schemes, although very useful in specific cases, have several serious drawbacks as routine microsequencing procedures. They are expensive. Obtaining biosynthetically labeled proteins is not always possible. The differences in relative specific activities of the different amino acids present problems in quantitative analysis of these proteins.

Several attempts have been made to introduce a radioactive tag onto amino acids by using [³H]- and [¹⁴C]phenyl isothiocyanate or phenyl [³⁵S]isothiocyanate as the coupling reagent in spinning cup sequencing (Oroszlan et al., 1975; Jacobs & Niall, 1975) or solid phase sequencing (Laursen, 1971;

[†] From the Division of Biology, California Institute of Technology, Pasadena, California 91125. Received November 18, 1977. This work was supported by Grant GM 06965 from the National Institutes of Health.

* Supported by a National Institutes of Health Postdoctoral Fellowship.

Bridgen, 1975). With these *extrinsic labeling* procedures, however, the large excess of reagent required for coupling and the instability of the radiolabeled phenyl isothiocyanate make this an expensive and troublesome process. Moreover, radioactive reagent or by-products extracted with the anilinothiazolinones can seriously interfere with the analytic schemes.

We describe in this paper an improved method for microsequencing based upon the *direct analysis* of amino acid phenylthiohydantoin from automated Edman degradation. This technique does not require radiolabel. In order to develop this routine and generally applicable procedure for microsequencing analysis of both proteins and peptides, we have combined and developed innovations in the Edman degradation technique, the automatic sequenator (originally suggested by Wittmann-Liebold, 1973), and the analytic systems (patterned after the high pressure liquid chromatographic identification of phenylthiohydantoin developed by Zimmerman et al., 1977). These techniques have largely overcome several of the special problems posed by microsequencing applications: washout of the polypeptide from the spinning cup, side reactions that block or otherwise interfere with the degradation, incomplete reagent and by-product removal that contaminates the polypeptide film or extracted anilinothiazolinones, low and nonreproducible yields of the derivatives used for analysis of the amino acids, and low sensitivity of the standard analytic techniques. These improvements are described below.

Materials and Methods

Proteins and Peptides. Spermin whale myoglobin (Sigma) was purified and converted to the apoprotein according to the method of Edman & Begg (1967). Sindbis virus E1 and E2 glycoproteins were prepared by J. R. Bell and J. H. Strauss (California Institute of Technology). Rauscher leukemia virus gp70 was prepared by S. Oroszlan (Flow Laboratories) and was a gift of R. Gilden (Frederick Cancer Research Institute) and W. J. Dreyer (California Institute of Technology). [Ala^{3,14}]-Somatostatin was synthesized by J. Rivier and R. Guillemin and was obtained from W. J. Dreyer. The heavy chain (α) from Balb/c myeloma protein J558 was prepared by J. Schilling (California Institute of Technology). The light chain (κ) from NZB myeloma protein 4050 was a gift of M. Weigert (Institute for Cancer Research, Philadelphia, Pa.).

Reagents and Solvents. Heptafluorobutyric, trifluoroacetic, and acetic acids (Pierce Sequanal Grade) were purified by distillation under argon on a 30-cm Widmer column first from chromium trioxide and then from alumina (Woelm activity 1). Phenyl isothiocyanate (Pierce Sequanal Grade) was purified by vacuum distillation (0.2 mmHg). 1-Propanol (Burdick and Jackson Distilled-in-Glass) was treated with charcoal at 50 °C for 1 h, filtered, stirred overnight with 2,4-dinitrophenylhydrazine (2 g/L) and a trace of H₂SO₄ at room temperature, filtered, and distilled under vacuum (20 mmHg). *n*-Heptane (Pierce Sequanal Grade) was distilled under argon in an 80-cm Widmer column. Water was double deionized-distilled.

Quadrol-trifluoroacetic acid buffer (1.0 M), pH 9.0, was Pierce Sequanal Grade. Dithiothreitol was obtained from Calbiochem, sodium acetate was Mallinckrodt Reagent Grade, and Polybrene was purchased from Aldrich. Benzene, ethyl acetate, acetonitrile, and 1-chlorobutane were Burdick and Jackson Distilled-in-Glass Solvents.

The composition of the reagents and solvents used in the sequenator is listed in Table I.

Sequenator. A Beckman Instruments Model 890B (updated) sequenator equipped with an undercut cup was ex-

TABLE I: Reagents and Solvents for the Sequenator.

R1	5% phenyl isothiocyanate in <i>n</i> -heptane
R2	0.33 M Quadrol-trifluoroacetic acid buffer, pH 9.0, in 1-propanol/water (3/4, v/v)
R3	Heptafluorobutyric acid
R4	Trifluoroacetic acid/water (1/3, v/v)
S1	Benzene
S2	Ethyl acetate/0.1% acetic acid
S3	1-Chlorobutane/0.001% dithiothreitol
S4	Acetonitrile/0.001% dithiothreitol

tensively modified essentially according to the method of Wittmann-Liebold (1973) and Wittmann-Liebold et al. (1976). Details of the modifications can be obtained from Dr. Wittmann-Liebold at the Max Planck Institute for Molecular Genetics, 1 Berlin 33 (Dahlem). These changes are briefly summarized as follows. (1) The vacuum line includes a removable glass trap that is immersed in a stainless steel liquid nitrogen Dewar. The trap is removed and emptied every 8 to 10 days. (2) The restricted, rough, and fine vacuum functions on the Beckman system are replaced by a two-stage vacuum drawn by a single large capacity pump (Welch Duo-Seal Model 1397B). Two solenoid valves (Leybold-Heraeus) were placed in parallel in the vacuum line, and the one controlling the restricted vacuum connects the cup to the vacuum through a 0.5 mm × 1 cm restriction. The vacuum pump oil is changed every 4 to 5 days, and the vacuum valves and their housing are disassembled and cleaned every month. (3) Vacuum lines and connections are Leybold-Heraeus seamless stainless steel hose and Quick-Flange fittings with Viton O-rings. (4) The Beckman reagent, solvent, nitrogen, waste, and collect valves are replaced by a series of pneumatic-actuated, Teflon diaphragm valves designed by H. Graffunder and B. Wittmann-Liebold (Max Planck Institute for Molecular Genetics, Berlin). Three-way universal solenoid valves (ITT Controls) are used to supply compressed air (80 to 90 psi) to keep the diaphragm valves closed or vacuum from a Welch Duo-Seal Model 1402 pump to open them. The diaphragm valves are arranged in a manifold in a zero dead volume configuration with a single exit line for all reagents and solvents delivered to the cup. This line is emptied into the cup by flushing with argon following a reagent or solvent delivery. (5) A secondary reaction vessel similar to the one described by Wittmann-Liebold et al. (1976) is inserted between the cup and fraction collector to provide automated conversion of the anilinothiazolinones extracted from the cup to amino acid phenylthiohydantoin. The vessel is thermostated at 55 °C by a water jacket connected to a Lauda Model K-2 circulating water heater. (6) Positions 36 to 42 on the Beckman programmer are wired to relays to control the functions needed for the conversion flask operation. The R4 and S4 delivery functions are interlocked to the flask pressurize function in a manner similar to that by which the other reagent and solvent delivery functions are interlocked to the cup pressurize function in the Beckman system.

Program. The sequenator program is listed in Table II. The main program, governing operations in the spinning cup, is similar to that described by Edman & Begg (1967). Changes include the following. (1) The coupling buffer, R2, is diluted to 0.33 M Quadrol and is added in two aliquots at the beginning and in the middle of the coupling stage. (2) The cup spin rate is cycled between low (1200 rpm) and high (1800 rpm) several times during the coupling stage. (3) Argon, rather than high purity nitrogen, is used to provide an inert atmosphere. (4) The reagent/solvent delivery line is emptied into the cup by a stream of argon (line flush) following actuation of the

TABLE II: Sequenator Program.

Step	Cup functions	Flask functions	Step time (s)	Vol (mL)	Cup speed (rpm)
1	Stop slew	Conversion	2		1200
2	Pressurize	Conversion	10		1200
3	Cup dry	Conversion	4		1200
4	R1 vent and pressurize	Conversion	8		1200
5	R1 deliver, cup vent	Conversion	3	0.35	1200
6	R2 vent and pressurize	Conversion	8		1200
7	R2 deliver, cup vent	Conversion	14	0.38	1200
8	Line flush, cup dry	Conversion	50		1200
9	Coupling	Conversion	10		1200
10	Coupling	Conversion	10		1800
11	Coupling	Conversion	10		1200
12	Coupling	Conversion	10		1800
13	Coupling	Conversion	400		1200
14	Coupling	Conversion	400		1800
15	Cup dry	Conversion	10		1200
16	R2 vent and pressurize	Conversion	8		1200
17	R2 deliver, cup vent	Conversion	12	0.35	1200
18	Line flush, cup dry	Conversion	50		1200
19	Coupling	Conversion	10		1200
20	Coupling	Conversion	10		1800
21	Coupling	Conversion	10		1200
22	Coupling	Conversion	10		1800
23	Coupling	Conversion	300		1200
24	Coupling	Line flush, flask dry	200		1800
25	Coupling	Vacuum, line flush, pressurize	300		1200
26	Coupling	Vacuum	100		1800
27	Coupling	Pressurize	10		1800
28	Coupling	Flask dry	4		1800
29	Coupling	S4 vent and pressurize	4		1800
30	Coupling	S4 deliver, flask vent	30	2.5	1800
31	Coupling	Line flush, transfer to collector	40		1800
32	Coupling	S4 deliver, flask vent	20	1.8	1800
33	Coupling	Line flush, transfer to collector	40		1800
34	Rough vacuum, line flush	Line flush, flask dry, collector step	300		1800
35	Fine vacuum, argon bleed		200		1800
36	Fine vacuum		200		1800
37	Pressurize		10		1800
38	Cup dry		4		1800
39	S1 & S2 vent and pressurize		4		1800
40	S1 deliver, cup vent		10	0.6	1800
41	Precipitation		80		1800
42	Cup dry		4		1800
43	S1 deliver, cup vent		200	9.0	1800
44	Cup dry		4		1800
45	S2 deliver, cup vent	Vacuum	400	28.0	1800
46	Line flush, cup dry		40		1800
47	Rough vacuum, line flush		140		1800
48	Fine vacuum, argon bleed		100		1200
49	Fine vacuum		200		1200
50	Pressurize		10		1200
51	Cup dry		4		1200
52	R3 vent and pressurize		8		1200
53	R3 deliver, cup vent		3	0.3	1800
54	Line flush, cup dry		60		1800
55	Cleavage		160		1800
56	Rough vacuum, line flush		80		1800
57	Fine vacuum		40		1800
58	Pressurize		10		1800
59	Cup dry		4		1800
60	S3 vent and pressurize		4		1800
61	S3 deliver, cup vent		12	0.6	1800
62	Precipitation		80		1800
63	Cup dry	Pressurize	10		1800
64	S3 deliver, transfer to flask	Vent	120	5.0	1800
65	Line flush, transfer to flask	Vent	40		1800
66	Rough vacuum, line flush	Line flush, flask dry	140		1800
67	Fine vacuum	Line flush, flask dry	80		1200
68	Pressurize	Line flush, flask dry	10		1200
69	Cup dry	Line flush, flask dry	4		1200
70	R3 vent and pressurize	Line flush, flask dry	8		1200
71	R3 deliver, cup vent	Line flush, flask dry	3	0.3	1800

TABLE II (continued)

Step	Cup functions	Flask functions	Step time (s)	Vol (mL)	Cup speed (rpm)
72	Line flush, cup dry	Line flush, flask dry	60		1800
73	Cleavage	Vacuum	120		1800
74	Rough vacuum, line flush	Pressurize	10		1800
75	Rough vacuum, line flush	Flask dry	4		1800
76	Rough vacuum, line flush	R4 vent and pressurize	4		1800
77	Rough vacuum, line flush	R4 deliver, flask vent	4		1800
78	Rough vacuum, line flush	Line flush, flask dry	20		1800
79	Rough vacuum, line flush	Conversion	40		1800
80	Fine vacuum	Conversion	40		1800
81	Pressurize	Conversion	10		1800
82	Cup dry	Conversion	4		1800
83	S3 vent and pressurize	Conversion	4		1800
84	S3 deliver, cup vent	Conversion	12	0.6	1800
85	Precipitation	Conversion	80		1800
86	Cup dry	Conversion	8		1800
87	S3 deliver, cup vent	Conversion	140	6.0	1800
88	Line flush, cup dry	Conversion	40		1800
89	Rough vacuum, line flush	Conversion	200		1800
90	Fine vacuum, argon bleed	Conversion	100		1200
91	Fine vacuum	Conversion	200		1200
92	Start slew	Conversion			1200
93	Conditional stop, fine vacuum	Conversion			1200

delivery valves. (5) A fine stream of argon is metered into the reaction chamber during some fine vacuum steps to purge semivolatile material (Hermodson et al., 1972). This argon bleed is adjusted with the fine vacuum actuated to give a vacuum of 250 μ m of Hg.

The secondary program, controlling the automated conversion of anilinothiazolinones to phenylthiohydantoins, is similar to that described by Wittmann-Liebold et al. (1976). The chlorobutane extract from the first cleavage stage is directed into a water-jacketed (55 °C) glass vessel where the solvent is removed first by bubbling argon through it and then by vacuum from the main Welch Model 1397B pump. The conversion is effected by addition of an aqueous trifluoroacetic acid solution to the residue. After conversion is complete and the acid is removed by vacuum, the phenylthiohydantoins are transferred to the fraction collector by extraction with acetonitrile.

Carrier. Polybrene, a polymeric quaternary ammonium salt, is used as a carrier to retain the protein or peptide in the spinning cup. The Polybrene (3 to 6 mg) and glycylglycine (50–100 nmol) dissolved in 0.5 mL of water is applied to the cup, dried under vacuum, and subjected to three complete degradation cycles beginning with step 1 of the program (Table II). A solution of the sample is then applied to the cup and dried under vacuum, and if the sample is a small peptide the program is started at step 1. If the sample is a large peptide or protein, the program is started at step 37, carried through step 91, run from step 1 to step 49 on the next cycle, and then restarted at step 1 to effect double coupling for the first degradative cycle.

Identification of Amino Acid Phenylthiohydantoins. The samples are removed from the fraction collector of the sequenator, and the solvent is evaporated with a stream of nitrogen. The residue is redissolved in acetonitrile, and an aliquot is analyzed by high pressure liquid chromatography on a Du Pont Zorbax ODS column (4.6 mm \times 25 cm). The column is housed in an oven thermostated at 55 °C. It is initially equilibrated with 0.02 M sodium acetate buffer, pH 5.0, and the amino acid phenylthiohydantoins are eluted by a gradient of increasing acetonitrile concentration.

The chromatography is performed on either a Du Pont or a Waters Associates chromatographic system. The Du Pont

system consists of a Model 830 liquid chromatograph with a 254-nm absorbance detector and a septumless injector, a Model 838 gradient programmer, a Model 833 flow controller, and a Beckman Instruments strip chart recorder. The Waters system consists of two Model 6000A pumps, a Model UK6 injector, a Model 440 dual channel (254 and 313 nm) absorbance detector, a Houston Instruments Omniscrite dual channel strip chart recorder, and a multilinear gradient programmer of our own design that controls the output of the two Waters pumps.

Results

Sequenator Operation. The design modifications of Wittmann-Liebold (1973) and Wittmann-Liebold et al. (1976) have improved several aspects of sequenator performance important to direct microsequencing experiments. (1) Changes in the vacuum system have improved both the initial quality and long term stability of the vacuum. For example, our vacuum is as low as 3–5 μ mHg at the start of a sequenator run, and it remains at less than 10 μ mHg after 60 degradation cycles without a change in pump oil. This improved vacuum has reduced oxidative losses in the sequenator, the buildup of semivolatile substances in the spinning cup, and contamination of the amino acid phenylthiohydantoins with reagents and by-products. (2) The addition of leak-proof, zero dead volume delivery valves has eliminated cross-contamination of reagents and solvents and their leakage into the reaction chamber, improved the solvent extraction efficiency, reduced polypeptide washout from the cup, and decreased contamination of the amino acid phenylthiohydantoins. (3) Addition of the device for automatic conversion of the anilinothiazolinones extracted from the cup to phenylthiohydantoins under an inert atmosphere has improved the yields of the more unstable amino acid derivatives (Table III) and provided conversion efficiencies that are reproducible from cycle to cycle. Operation of the modified sequenator without the automatic conversion device (i.e., with manual conversion) eliminates much of the improvement in recoveries shown in Table III. These improvements are critical for the analysis of microsamples (less than 25 nmol) and for the analysis of extended sequenator runs (70–80 cycles) because small signal changes can readily be

TABLE III: Yields of Amino Acid Phenylthiohydantoin from a Standard and Modified Beckman Spinning Cup Sequenator.^a

Amino acid phenylthiohydantoin	% recovery	
	Standard sequenator ^b	Modified sequenator ^c
Ala	70 ± 10	97 ± 3
Arg	40 ± 10	70 ± 10
Asn	35 ± 10	90 ± 5
Asp	(plus 30% as Asp)	(plus 5% as Asp)
Cys/cystine	60 ± 10	95 ± 5
	0	25 ± 5
		(as unknown derivative)
Glu	60 ± 10	95 ± 5
Gln	35 ± 10	85 ± 5
	(plus 30% as Glu)	(plus 10% as Glu)
Gly	60 ± 10	80 ± 10
His	35 ± 15	80 ± 10
Ile	70 ± 10	97 ± 3
Leu	70 ± 10	97 ± 3
Lys	60 ± 10	95 ± 5
Met	60 ± 10	90 ± 5
Phe	70 ± 10	97 ± 3
Pro	60 ± 10	75 ± 10
Ser	0	10 ± 5
	(5% as dehydro form)	(plus 50% as unknown derivative)
Thr	0	12 ± 5
	(15% as dehydro form)	(plus 50% as dehydro form)
Trp	35 ± 15	70 ± 15
Tyr	70 ± 10	97 ± 3
Val	70 ± 10	97 ± 3

^a Cumulative results from a variety of protein N-terminal sequenator runs analyzed by high pressure liquid chromatography. They represent recoveries of the amino acid derivatives at random cycles in the sequenator runs compared with the theoretical yields at those cycles. ^b Beckman 890B (updated) using dimethylbenzylamine program (Hermodson et al., 1972) (or a standard Beckman Quadrol program) and manual conversion of anilinothiazolinones with 1 N HCl/0.1% EtSH at 80 °C for 10 min. ^c Modified according to Wittman-Liebold (1973) using Quadrol program and automatic conversion with 25% trifluoroacetic acid at 55 °C for 30 min (similar recoveries are obtained with dimethylbenzylamine buffer instead of Quadrol, but contamination of the phenylthiohydantoin with ultraviolet-absorbing impurities interferes with microsequence analysis on the high pressure liquid chromatograph).

detected upon a stable background of amino acid phenylthiohydantoin.

We find the original Edman Quadrol program, with slight modification, superior to the newer programs employing more volatile buffers such as dimethylallylamine (Niall et al., 1974) or dimethylbenzylamine (Hermodson et al., 1972) or to those employing very dilute Quadrol with different solvent extractions (Crewther & Inglis, 1975; Brauer et al., 1975). The newer buffer systems are generally used to reduce washout of small peptides because they require much reduced solvent extractions following the coupling stage, but they leave unacceptable levels of contaminants that interfere with the high pressure liquid chromatographic analysis of the phenylthiohydantoin for microsequence analysis. The use of Polybrene as a carrier for both proteins and peptides, as shown below, is a much more effective way of preventing washout and at the same time allowing use of the older Quadrol program which is cleaner because of its extensive solvent extractions.

In order to achieve efficient coupling, we have found it necessary to add the coupling buffer (at one third the concentration recommended by Edman) in two aliquots at the beginning and end of the coupling stage. This double addition and cycling the cup speed from high to low several times help

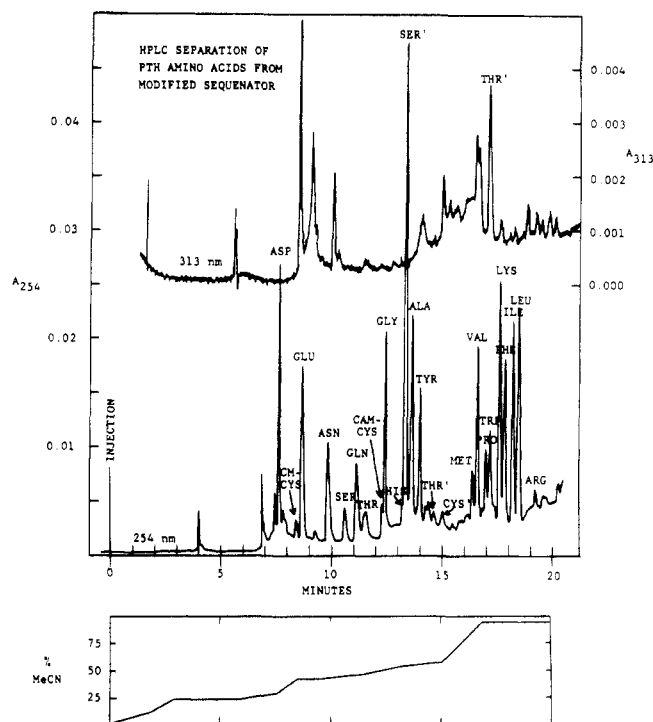


FIGURE 1: High pressure liquid chromatographic separation of amino acid phenylthiohydantoin from a sequenator cycle. The chromatography was performed on a Du Pont Zorbax ODS reverse phase column using the aqueous 0.02 M sodium acetate buffer (pH 5.0)-acetonitrile gradient for development as shown below the chromatogram. A 10% aliquot of the sequenator sample (a serine residue at cycle 30 of a murine myeloma protein, XRPC-25, light chain) was injected with the absorbance detector set at 0.05 absorbance full scale for the 254-nm channel and 0.01 absorbance full scale for the 313-nm channel. Flow rate was 1.0 mL/min, column temperature was 54 °C, and the recycle time was 26 min. Elution times for the serine and threonine derivatives with unaltered side chains are 10.6 and 11.6 min, for dehydrothreonine the elution time is 17.2 min (313-nm channel), for the uncharacterized forms of serine, threonine, and cysteine times are 13.4, 14.4, and 14.6 min, for carboxamidomethylcysteine the time is 12.3 min, and for carboxymethylcysteine it is 8.4 min.

keep the protein dissolved. Another feature of note in the coupling stage of our program, and in Edman's original program, is that the heptane in which phenyl isothiocyanate is dissolved is not removed from the cup by drying before addition of the coupling buffer. It evaporates during the coupling stage leaving a very evenly dispersed suspension of the coupling reagent in the buffer.

High Pressure Liquid Chromatography Performance. We have found chromatography on Du Pont Zorbax ODS columns to be an excellent method for analysis of amino acid phenylthiohydantoin. The absorbance detectors on commercial liquid chromatographs possess high sensitivity, with full scale recorder deflection on as little as 50 pmol, and phenylthiohydantoin of all common amino acids can be separated on a single column run. Moreover, the cycle time for this procedure is as low as 25 min.

A typical column separation for the mixture of amino acid phenylthiohydantoin obtained from a single cycle on our sequenator is shown in Figure 1. This separation is similar to that obtained for standard mixtures of phenylthiohydantoin obtained by Zimmerman et al. (1977). Several features are noteworthy: (1) Serine, threonine, and cysteine are detected as multiple peaks. The yields of authentic phenylthiohydantoin (Pth)¹-serine and Pth-threonine, which chromatograph near

¹ Abbreviations used: Pth, phenylthiohydantoin; TLC, thin layer chromatography; GC, gas chromatography; HPLC, high pressure liquid chromatography.

Pth-glutamine, are 5–15% of theoretical, while Pth-dehydrothreonine, detected in the 313-nm absorbance channel, and an uncharacterized derivative of serine, which chromatographs near Pth-alanine, are obtained in 40–60% yield. (2) Interference from ultraviolet-absorbing contaminants is minimal. Typically, injection of 50% of the sample from one sequenator cycle gives no peaks other than those for amino acid phenylthiohydantoin or dithiothreitol in the 254-nm absorbance channel larger than 10% of full scale deflection on the most sensitive absorbance setting for the high pressure liquid chromatograph detectors. Most extraneous peaks in the chromatogram arise from artifacts in the chromatography system such as impurities in the chromatography solvents. We are now in the process of removing these artifacts by thorough purification of these solvents. (3) Pth-arginine and Pth-histidine can be analyzed along with the acidic and neutral amino acid phenylthiohydantoin. However, on some columns they chromatograph as broad peaks unless a steeper gradient or higher salt concentration, e.g., 0.07 M sodium acetate buffer, is used. Furthermore, their elution times may change significantly as the column ages. Therefore, quantitative analysis of these derivatives is often made by high pressure liquid chromatography of the aqueous phase from a 0.1 N HCl/ethyl acetate fractionation of the sequenator cycle if the intact fraction indicates Pth-arginine or Pth-histidine as a broad peak obscured by other amino acid phenylthiohydantoin. (We are currently investigating the use of Du Pont Zorbax CN high performance columns as a means of eliminating this inconvenience. Preliminary experiments suggest they possess high resolution capabilities for Pth-arginine and histidine as well as the other phenylthiohydantoin.) (4) Usable column lifetimes in our chromatography systems are 3–6 weeks or 500–1000 injections. A typical gradient is shown in Figure 1, but the gradient must sometimes be adjusted slightly for older columns in order to provide optimum resolution.

Polybrene as a Protein or Polypeptide Carrier. Polybrene, originally suggested as a carrier for larger quantities of small peptides by Tarr et al. (1978) and Klapper et al. (1978), prevents small quantities of either proteins or polypeptides from washing out of our improved spinning cup sequenator. Moreover, Polybrene is not a polypeptide itself, so direct analysis of the microsamples of amino acid phenylthiohydantoin from the sequenator is possible. The Polybrene itself does not contribute amino acid background as would an N-terminal blocked protein carrier. It does increase retention of smaller fragments of the protein formed in the sequenator by random peptide cleavage, and hence the background observed from this random cleavage, but this does not seriously hinder its effectiveness.

The effectiveness of the Polybrene is illustrated by the degree of washout observed with and without Polybrene for small quantities of a protein, myoglobin (151 residues), and a peptide, [Ala^{3,14}]-somatostatin (14 residues). Polybrene allowed the analysis of 47 residues (96% repetitive yield) of myoglobin at the 200-pmol level and all 14 residues (88% repetitive yield) of the synthetic hormone at the 1.5-nmol level. Without Polybrene, these samples were washed out of the cup at cycles 16 (84% repetitive yield) and 3 (20% repetitive yield), respectively.

The effectiveness of Polybrene as an agent for retaining small quantities of proteins in our sequenator contrasts with its failure in the system of Tarr et al. (1978). They observed a lower repetitive yield with Polybrene as a protein carrier than without it. We also observe very low repetitive yields if the protein sequencing is begun without precycling the Polybrene in the sequenator with a small peptide. The peptide (glycyl-

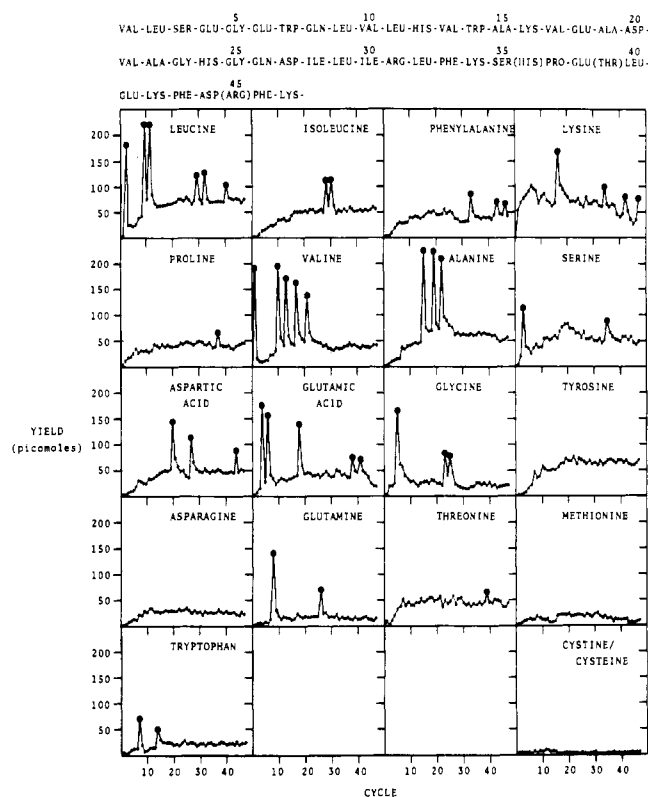


FIGURE 2: Yields of amino acid phenylthiohydantoin from an N-terminal sequenator analysis of sperm whale apomyoglobin. Aliquots (10–50%) of each cycle were analyzed by high pressure liquid chromatography, peak heights were converted to picomoles for each derivative using values for standard mixtures of amino acid phenylthiohydantoin, and the yields were normalized to a 100% injection. Parentheses indicate some uncertainty in assignment. Data for arginine and histidine are not shown. These derivatives were not well resolved on the column used for this run with the standard gradient system, and aqueous extraction was not performed.

glycine) apparently reacts with material in the Polybrene that otherwise would react with the protein and cause blockage of its N terminus.

Microsequencing Capability for Proteins. The use of the modified spinning cup sequenator, Polybrene as a carrier for the protein, and high pressure liquid chromatography as an analytic technique allows direct (i.e., without radioactive label) sequence analysis of extended portions of the N terminus of proteins using as little as 100 pmol of protein. Figures 2, 3, and 4 show the N-terminal sequence determined for sperm whale apomyoglobin (47 residues starting with 200 pmol), 4050 light chain (77 residues starting with 5 nmol), and J558 heavy chain (54 residues starting with 8 nmol) using this system. Also included in the figures are liquid chromatographic data giving yields for the various amino acids at each cycle for these sequenator runs. The excellent conversion reproducibility effected by the automated system of Wittmann-Liebold et al. (1976) is illustrated by the stable backgrounds shown in these figures. Cycles in which the yield from the primary sequence represents a relatively small absolute increase compared with the background are clearly assignable.

Several features of the data in Figures 2–4 are important. (1) The level of and variation in the background due to random cleavage of the polypeptide in the sequenator is different for each amino acid and from one protein to another. These differences arise from a variety of factors, including the size of the protein, its amino acid composition, and the degree to which certain amino acid sequences are susceptible to the acid catalyzed cleavage. (2) Some amino acid residues, most notably

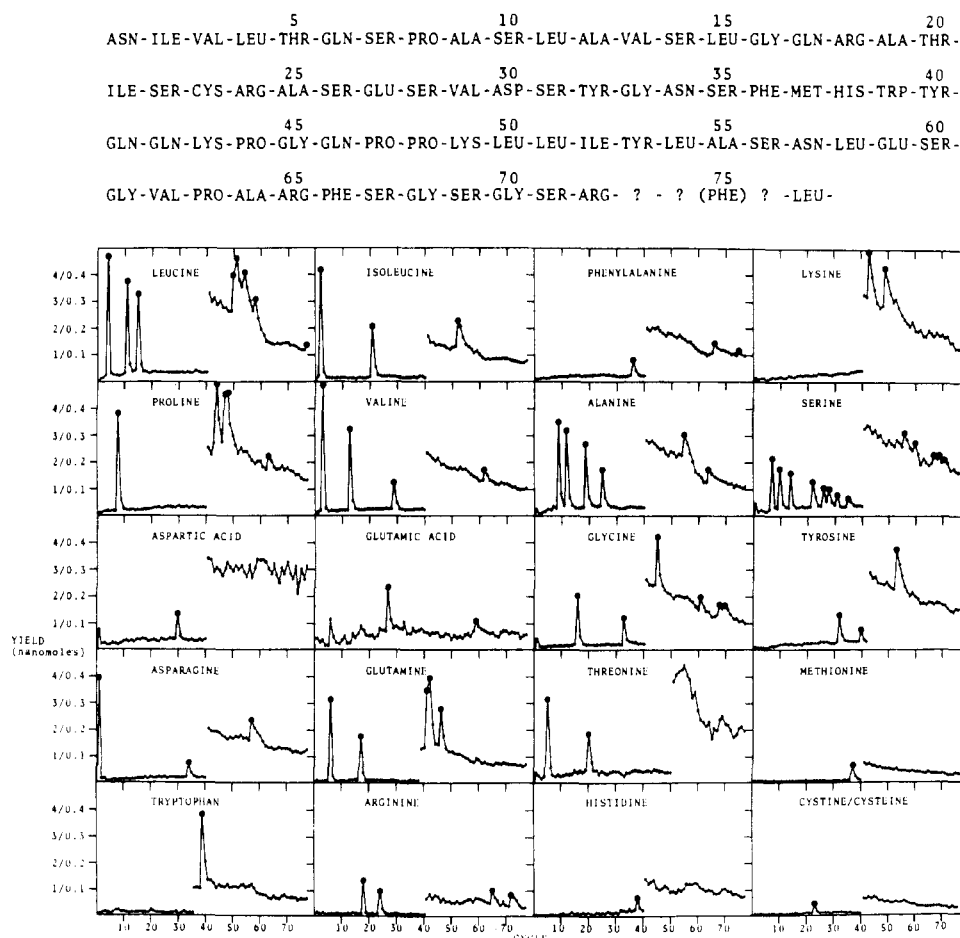


FIGURE 3: Yields of amino acid phenylthiohydantoin from an N-terminal sequenator analysis of 5.0 nmol of NZB myeloma protein 4050 light chain (peptide mol wt ~23 000). Yields were calculated as described in the legend to Figure 2. Question marks indicate no assignment, and parentheses indicate some uncertainty in assignment.

tryptophan, are slowly destroyed during sequenator operation. This results in a more rapid decline in yield of these derivatives for successive cycles than is observed for other amino acids. (3) Deamination of the side chains of asparagine and glutamine is minimal both in the spinning cup reactions and in the automated conversion from anilinothiazolinones to phenylthiohydantoin. (4) Since residue assignment depends on the signal to noise (background) for the phenylthiohydantoin at each cycle, some amino acids with low background and high yield can be identified reliably later in the sequence analysis than others. A few of these amino acids probably could have been identified past the termination points in the sequenator analyses shown. This might be useful, for instance, in locating methionine residues which provide sites for specific chemical cleavage of large peptides or proteins.

We have obtained, in addition to that for the soluble globular proteins listed above, extended N-terminal sequence data on several hydrophobic, membrane glycoproteins using the direct microsequencing techniques. These include Rauscher leukemia virus gp70 (46 residues starting with 3.5 nmol), Sindbis virus E1 (53 residues starting with 24 nmol), and Sindbis virus E2 (52 residues starting with 15 nmol). The sequenator data from these experiments will be communicated in later manuscripts.

Microsequencing Capability for Peptides. Figure 5 shows the liquid chromatographic data for a sequenator run with 1.5 nmol of a 14-residue peptide, [Ala^{3,14}]-somatostatin. All residues, including the C-terminal alanine, are clearly identifiable. As we noted above, without adding Polybrene to the cup

as a carrier, we could identify only the first three residues.

Klapper et al. (1978) were also able to sequence fully short peptides using Polybrene as carrier, but they used considerably more sample (approximately 30 nmol) because of the previously mentioned contaminants associated with their peptide program employing a volatile buffer (dimethylallylamine), minimal solvent extractions, and single cleavage program. This procedure resulted in high pressure liquid chromatograms containing numerous peaks from impurities. Several contaminant peaks were much more intense than peaks from the amino acid phenylthiohydantoin from the peptides. These impurities would seriously interfere with analysis of peptides at the nanomole or subnanomole level. The performance of the Wittmann-Liebold modified sequenators permits the sequencing of short peptides with the much cleaner Quadrol-protein program even at this low sample level.

Discussion

Performance of the Modified Spinning Cup Sequenator Reasonably simple and inexpensive (\$8000 to \$10 000) modifications of a commercial sequenator have significantly improved its operation, particularly for microsequencing operations. These modifications, similar to those described by Wittmann-Liebold (1973) and Wittmann-Liebold et al. (1976), have improved the vacuum and reagent/solvent delivery systems so as to provide cleaved amino acid derivatives more nearly free of contaminants than can be provided with standard commercial instruments. Addition of the device for automatic conversion of the cleaved anilinothiazolinones to

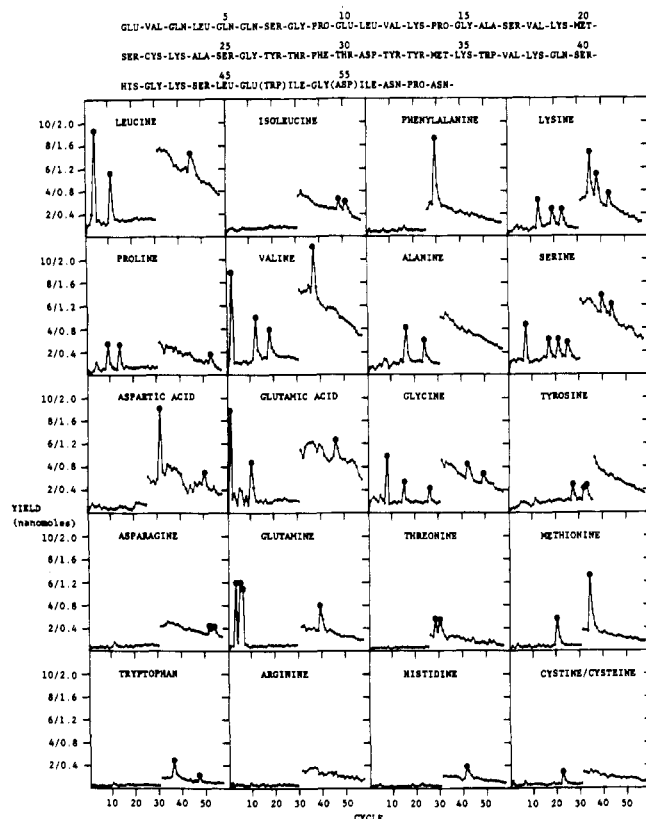


FIGURE 4: Yields of amino acid phenylthiohydantoin from an N-terminal sequenator analysis of 8.0 nmol of Balb/c myeloma protein J558 heavy chain (peptide mol wt ~55 000). Yields were calculated as described in the legend to Figure 2. Parentheses indicate some uncertainty in assignment.

amino acid phenylthiohydantoin has significantly improved yields of several of the phenylthiohydantoin and has remarkably improved the reproducibility in conversion yields from cycle to cycle.

The improvements in sequenator repetitive yield and reliability have allowed routine analysis of extended portions (beyond 70 residues) of the N terminus of many proteins and large polypeptides at the 50 nmol or larger sample size (Hunkapiller & Hood, unpublished). The use of Polybrene as a noncovalent carrier to retain small quantities of polypeptides in the spinning cup of the modified sequenator has also allowed extended sequenator analysis (50 residues) of nanomole and subnanomole quantities of polypeptides. This technique is applicable both to N-terminal microanalysis of proteins and large peptides and to complete microanalysis of smaller peptides.

An assessment of the relative contributions of the various refinements of the sequenator techniques to the ability to perform extended microsequence N-terminal analysis is somewhat subjective at present. We believe, however, that all of them make significant contributions. The Polybrene is primarily responsible for retaining the small quantity of polypeptide in the spinning cup, but the improved vacuum and reagent delivery systems are required to provide samples sufficiently free of contaminants for direct analysis. The automated conversion system is necessary to provide high yields of phenylthiohydantoin from the more unstable anilinothiazolinones. Further, it provides conversion reproducibility that permits residue assignments farther into the sequence than is possible with manual conversion. Finally, high pressure chromatographic analysis of the phenylthiohydantoin provides a rapid, high sensitivity, quantitative determination of all

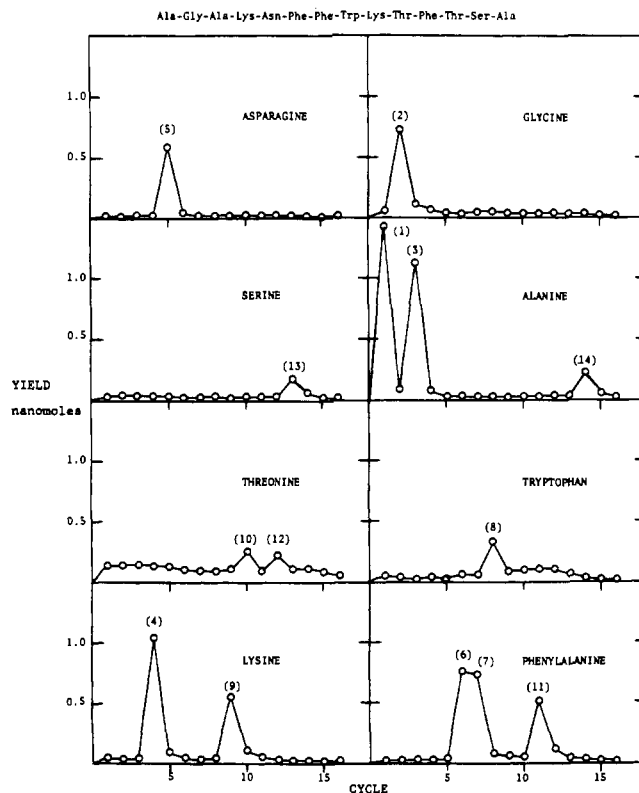


FIGURE 5: Yields of amino acid phenylthiohydantoin from a sequenator analysis of 1.5 nmol of [Ala^{3,14}]-somatostatin. Yields were calculated as described in the legend to Figure 2.

amino acids released in each sequenator cycle so that the residue can be identified unambiguously.

Performance of High Pressure Liquid Chromatography as an Analytic System. Liquid chromatographic analysis on a high performance reverse phase support (Zorbax ODS) provides a rapid, sensitive, and quantitative method for analysis of the mixtures of amino acid phenylthiohydantoin obtained from the cycles of automated Edman degradation by the sequenator. When the phenylthiohydantoin are essentially free of ultraviolet-absorbing contaminants, as they are with the modified sequenator, the liquid chromatographic system can be used for direct microsequence analysis. No radioactive labels are required to provide the needed analytic sensitivity down to the 100-pmol protein level. Moreover, if internally radiolabeled proteins are available, the high pressure liquid chromatographic technique can be used to fractionate the labeled amino acid phenylthiohydantoin from the sequenator by mixing them with a carrier solution of unlabeled phenylthiohydantoin. This permits sequencing of proteins in the 1- to 100-pmol range, depending on the specific radioactivity of the amino acids incorporated in the protein (McMillan et al., 1977). Further improvements in solvent purity may, in fact, allow direct microsequencing in this range.

Comparison with Other Microsequencing Methods: Sensitivity. Table IV lists the results from a variety of microsequencing experiments on sperm whale myoglobin by several investigators. The direct microsequencing technique described in this work is clearly superior in terms of the length of sequence obtainable in a single run and the quantity of protein required. Whereas 47 cycles could be identified with 200 pmol of myoglobin by our technique, schemes using direct analytic techniques or sequenator added radiolabels were successful for only 10 to 30 cycles using 1 to 10 nmol of protein.

Table V lists the results from microsequencing experiments

TABLE IV: Comparison of Sequenator Microanalyses for Myoglobin.^a

Quantity (nmol)	Cycles identified	Analytic technique	Remarks	Reference
0.2	47	HPLC	Modified sequenator with Polybrene carrier	This work
6	21	GC & TLC	Poly(Nle-Arg) carrier	Niall et al. (1974)
7	30	GC & TLC	0.1 M Quadrol	Brauer et al. (1975)
3	20	GC & TLC	0.1 M Quadrol	Brauer et al. (1975)
8	23	TLC	Phenyl [³⁵ S]isothiocyanate	Jacobs & Niall (1975)
1	10	TLC	Phenyl [³⁵ S]isothiocyanate	Jacobs & Niall (1975)
10	21	TLC	[¹⁴ C]Phenyl isothiocyanate with Ala methylthiohydantoin and succinylated polyornithine carriers	Oroszlan et al. (1975)

^a N-terminal sequence analysis for sperm whale apomyoglobin performed on spinning cup sequenators.

TABLE V: Comparison of Sequenator Microanalyses for Several Proteins.

Protein	Quantity (nmol)	Cycles identified	Sequenator	Analytic technique	Remarks	Reference
Sindbis virus E1 glycoprotein	24	53	Modified spinning cup	HPLC	Polybrene carrier	This work
Sindbis virus E2 glycoprotein	15	52	Modified spinning cup	HPLC	Polybrene carrier	This work
IgA heavy chain J558	8	54	Modified spinning cup	HPLC	Polybrene carrier	This work
κ light chain 4050	5	77	Modified spinning cup	HPLC	Polybrene carrier	This work
Rauscher leukemia virus gp70	3.5	46	Modified spinning cup	HPLC	Polybrene carrier	This work
Glucose-6-P dehydrogenase	35	25	Spinning cup	GC & TLC	0.1 M Quadrol	Brauer et al. (1975)
Staphylococcal protease	15	15	Spinning cup	TLC	Phenyl [³⁵ S]isothiocyanate	Jacobs & Niall (1975)
Several gp30 molecules	10	20-25	Spinning cup	TLC	[¹⁴ C]Phenyl isothiocyanate	Oroszlan et al. (1975)
HLA ^a	0.1	25	Spinning cup	Amino acid analyzer	Internally ¹⁴ C-labeled protein	Ballou et al. (1976)
HLA ^a	1	16	Solid phase	TLC	Phenyl [³⁵ S]isothiocyanate	Bridgen et al. (1976)
Murine λ light chain precursor	0.0001	24	Spinning cup	TLC	In vitro translation of mRNA with ³ H-labeled amino acids	Burstein & Schecter (1977)

^a HLA, human lymphocyte antigen.

TABLE VI: Comparison of Sequenator Microanalyses for Several Peptides.

Peptide	Quantity (nmol)	Cycles identified	Sequenator	Analytic technique	Remarks	Reference
[Ala ^{3,14}]-somatostatin	1.5	14/14	Modified spinning cup	HPLC	Polybrene carrier	This work
Many hydrophobic peptides	30	2/2 to 20/20	Spinning cup	HPLC	Dimethylallylamine program with Polybrene carrier	Klapper et al. (1978)
Several hydrophobic peptides	200	3/3 to 15/15	Spinning cup	HPLC	Dimethylallylamine program with Polybrene carrier	Tarr et al. (1978)
Several polar & hydrophobic peptides	300-800	3/3 to 28/28	Spinning cup	TLC	0.07 M Quadrol	Crewther & Inglis (1975)
Porcine insulin A chain	300	21/21	Spinning cup	GC & TLC	0.1 M Quadrol	Brauer et al. (1975)
Porcine insulin B chain	280	28/30	Spinning cup	GC & TLC	0.1 M Quadrol & Braunitzer's reagent	Brauer et al. (1975)
Bovine insulin B chain	5	7/30	Solid phase	TLC	Phenyl [³⁵ S]isothiocyanate	Bridgen (1975)

on a variety of proteins by several investigators. The data show that, for comparable protein loads, the direct sequencing approach described in this work is capable of determining sequences for two to three times as many cycles as can be defined by single runs with the other techniques (50 to 80 cycles compared with 15 to 25 cycles). The only generally applicable technique of comparable (and occasionally higher) sensitivity

to that described in the present paper is that developed by Ballou et al. (1976) using protein biosynthetically labeled by cell cultures incubated with ¹⁴C-enriched amino acids. The technique of Burstein & Schecter, although considerably more sensitive, is limited by the availability of mRNA for the in vitro translation with radiolabeled amino acids.

Table VI lists the results from microsequencing experiments

on a variety of small peptides by several investigators. The direct sequencing approach described in this work is capable of complete sequencing of even hydrophobic peptides at the 1-nmol level. Although other investigators using spinning cup sequenators with special peptide programs have succeeded in sequencing peptides, they usually require more than 100 times as much sample. Klapper et al. (1978), using Polybrene carrier and a volatile buffer program to reduce washout employed 20 times as much peptide as is required in our technique employing standard Quadrol-protein programs. Further refinements in our program (reduced Quadrol concentration, reduced solvent extractions, and single rather than double cleavage) should increase the sensitivity for peptide analysis to the 100-pmol level. Even highly sensitive (phenyl [³⁵S]isothiocyanate) techniques of Bridgen (1975) incorporated in solid phase sequenators developed specifically for peptide analysis are much less effective than the present spinning cup approach.

Applicability. The present technique is applicable to sequence analysis of any protein or peptide that can be obtained in a reasonably homogeneous state in the quantities described above. Although combining biosynthetic labeling of the protein with specific antibody purification schemes can provide comparable and occasionally higher sensitivity, many proteins of interest cannot be labeled readily with sufficient specific activity or uniformly to permit total sequence analysis. The intrinsic labeling approach offers distinct advantages, however, when separation of the radiolabeled proteins from unlabeled proteins such as antibodies used in purification steps is difficult. When proteins can be isolated free of other polypeptides, however, the direct sequencing approach described in this work is considerably simpler and more reliable.

Currently, the sensitivity of our direct microsequencing technique is about 100 pmol for extended (>40 cycles) N-terminal polypeptide analyses and somewhat less for shorter analyses. The primary sources of limitations to the sensitivity are impurities in the sequenator and chromatograph reagents and solvents. We have spent considerable time and effort to purify several of the sequenator reagents (R1, R3, and R4), and the sensitivity obtainable with our technique is due in part to the lack of contaminants in them. We are presently purifying the Quadrol buffer (R2) since the commercial product contains traces of amino acids as well as considerable amounts of aldehydic impurities. Both impurities interfere with microsequence analysis, the former by providing background amino acid phenylthiohydantoin visible in the chromatograms and the latter by reducing the repetitive cycle yield in the sequenator. Purification of the Quadrol and the chromatography solvents should increase the sensitivity of the microsequence analysis to the 10-pmol level.

Acknowledgments

We gratefully acknowledge the assistance of Dr. B. Wittmann-Liebold, Mr. H. Graffunder, and Dr. D. Braun for their invaluable and kind assistance in modification of our sequenators. We are also grateful to Dr. J. H. Strauss and J. R. Bell (Sindbis E1 and E2), to Drs. R. Gilden, S. Oroszlan, and W. J. Dreyer (Rauscher gp70), to Drs. J. Rivier, R. Guillemin, and

W. J. Dreyer (synthetic somatostatin), Dr. M. Weigert (4050 light chain), and to Drs. D. A. Klapper and D. J. McKean for preprints describing their peptide sequencing techniques using Polybrene. Ms. C. K. Kim provided excellent technical assistance in the reagent purification.

References

- Ballou, B., McKean, D. J., Freedlender, E. F., & Smithies, O. (1976) *Proc. Natl. Acad. Sci. U.S.A.* 73, 4487.
- Brauer, A. W., Margolies, M. N., & Haber, E. (1975) *Biochemistry* 14, 3029.
- Bridgen, J. (1975) *FEBS Lett.* 50, 159.
- Bridgen, J., Snary, D., Crumpton, M. J., Barnstable, C., Goodfellow, P., & Bodmer, W. F. (1976) *Nature (London)* 261, 200.
- Burstein, Y., & Schechter, I. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 716.
- Crewther, W. G., & Inglis, A. S. (1975) *Anal. Biochem.* 68, 572.
- Edman, P. (1956) *Acta Chem. Scand.* 10, 761.
- Edman, P., & Begg, G. (1967) *Eur. J. Biochem.* 1, 80.
- Hermanson, M. A., Ericsson, L. H., Titani, K., Neurath, H., & Walsh, K. A. (1972) *Biochemistry* 11, 4493.
- Jacobs, J. W., & Niall, H. D. (1975) *J. Biol. Chem.* 250, 3629.
- Klapper, D. A., Wilde, C. E., & Capra, J. D. (1978) *Anal. Biochem.* 85, 126.
- Laursen, R. A. (1971) *Eur. J. Biochem.* 20, 89.
- McKean, D. J., Peters, E. H., Waldby, J. I., & Smithies, O. (1974) *Biochemistry* 13, 3048.
- McMillan, M., Cecka, J. M., Murphy, D. B., McDevitt, H. O., & Hood, L. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5135.
- Mendez, E., & Lai, C. Y. (1975) *Anal. Biochem.* 68, 47.
- Niall, H. D., Jacobs, J. W., van Rietschoten, J., & Tregear, G. W. (1974) *FEBS Lett.* 41, 62.
- Oroszlan, S., Copeland, T., Summers, M., & Smythers, G. (1975) in *Solid Phase Methods in Protein Sequence Analysis* (Laursen, A., Ed.) p 179, Pierce Chemical Co. Press, Rockford, Ill.
- Pisano, J. J., Bronzert, T. J., & Brewer, H. B., Jr. (1972) *Anal. Biochem.* 45, 43.
- Silver, J. (1978) *Methods Enzymol.* 47, 247.
- Silver, J., & Hood, L. (1976) *Contemp. Top. Mol. Immunol.* 5, 35.
- Smithies, O., Gibson, D. M., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., & Ballantyne, D. L. (1971) *Biochemistry* 10, 4912.
- Tarr, G. E., Beecher, J. F., Bell, M., & McKean, D. J. (1978) *Anal. Biochem.* 84, 622.
- Theorell, H., & Akeson, A. (1954) *Am. Acad. Sci. Fenn. Ser. AII* 60, 303.
- Wittmann-Liebold, B. (1973) *Hoppe-Seyler's J. Physiol. Chem.* 354, 1415.
- Wittmann-Liebold, B., Graffunder, H., & Kohls, H. (1976) *Anal. Biochem.* 75, 621.
- Zimmerman, C. L., Appella, E., & Pisano, J. J. (1977) *Anal. Biochem.* 77, 569.