Application of Sequenator Analyses to the Study of Proteins†

Mark A. Hermodson, Lowell H. Ericsson, Koiti Titani, Hans Neurath,*
and Kenneth A. Walsh*

ABSTRACT: The method of sequenator analysis described by Edman and Begg (Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1*, 80) has been modified and applied to proteins and protein fragments. Significant modifications include the replacement of Quadrol by a volatile buffer (dimethylbenzylamine), the introduction of thiols to stabilize the reaction products, and the identification of the reaction products as silylated phenylthiohydantoins by automated gas–liquid chromatography. With these and other modifications, 30–50 amino acid residues can be identified and recovered with a repetitive yield of approximately 96%. This modified method has been tested on thermolysin and its cyanogen bromide fragments and found to be reliable in determining amino acid sequences. It has also been applied to porcine trypsin and found to be of use in determining purity, allotypic variants, and internal peptide-bond cleavage. In addition, the chemical nature of protein subunits can be identified by this method.

The sequential degradation of peptides by the method of Edman (1956) is an important procedure for the determination of amino acid sequences of proteins. The method combines the specificity of end-group analysis with the advantages of a cyclic stepwise process and normally yields 7–15 unambiguous degradations. In 1967, Edman and Begg automated the process by designing an instrument called the "sequenator" and demonstrated its successful application to the identification of 60 amino-terminal residues of apomyoglobin. Since then, other sequenators have been constructed, built on the principles of Edman and Begg. According to published accounts, these instruments are capable of 20–50 consecutive degradations (*e.g.*, Niall and Edman, 1967; Morgan and Henschen, 1969; Rochat *et al.*, 1970; Hood *et al.*, 1970; Niall *et al.*, 1970; Brewer and Ronan, 1970; Niall *et al.*, 1971; Reeck *et al.*, 1971; Hermodson *et al.*, 1971; Glenner *et al.*, 1971; Jaton *et al.*, 1971; Smithies *et al.*, 1971; Titani *et al.*, 1972a).

The capability of the sequenator to determine long amino acid sequences has altered the general strategy of sequence analysis. Instead of fragmenting the protein into a large number of short peptides whose sequences can be determined by manual Edman degradations and by digestion with carboxypeptidases, the protein is cleaved into a small number of large fragments, usually by chemical procedures (*e.g.*, cyanogen bromide, hydroxylamine), and the separated fragments are directly subjected to automated sequence analysis. Only those segments which cannot be reached by the sequenator are subsequently analyzed by the classical procedures.

Sequenator analysis has also been effective for screening proteins for homology, simply by applying the sequential analysis to the amino-terminal region of the protein or to other regions adjacent to existing or newly created α-amino groups. Such initiation points for consecutive degradations can be established by chemical reactions or by limited enzymatic proteolysis.

Because of its sensitivity and the small amount of protein required for but a few turns, sequenator analysis is a rapid and accurate test for protein purity and, *inter alia*, for determining the number of polypeptide chains in a pure oligomeric protein. The method also has proved useful in following the changes in covalent structure attending the activation of various pancreatic zymogens (Hermodson *et al.*, 1971; Pétra *et al.*, 1971).

We have applied the automated technique to proteins and protein fragments. Several changes in methodology developed in the course of this work have improved the reliability of the analysis and in most cases have enabled 30–50 sequential degradations. The direct identification of the reaction products as phenylthiohydantoylamino acids[1] by automated gas–liquid chromatography has proved particularly useful. Other methodological refinements include the replacement of Quadrol by dimethylbenzylamine in the coupling buffer and the use of thiols to stabilize the reaction products.

Because of the rapidly developing interest in the use of sequenators for sequence analysis of proteins, we wish to describe in this communication our present methodology and its application to several kinds of problems in protein chemistry.

Experimental Procedure

Reagents and Solvents. Phenyl isothiocyanate and heptafluorobutyric acid were prepared according to the method of Edman and Begg (1967).

Heptane, benzene, chlorobutane, and ethyl acetate were the "glass-distilled" grade of Burdick and Jackson Laboratories, Inc. 1-Propanol was purchased as the "Sequenal" grade from Pierce Chemical Co. These solvents were used without further purification.

N,N-Dimethylbenzylamine (Baker) was purified by treatment for 30 min with sodium borohydride (1 g/100 ml of DMBA) followed by four washes with distilled water. The reagent was then dried over $CaSO_4$ and vacuum distilled through a 40-cm Vigreaux column from phthalic anhydride (5 g/100 ml of DMBA) under aspirator vacuum. The constant

† From the Department of Biochemistry, University of Washington, Seattle, Washington 98195. *Received June 8, 1972.* This work has been supported by research grants from the National Institutes of Health (GM 15731) and the American Cancer Society (NP-18N).

[1] The following abbreviations are used: DMBA, *N,N*-dimethylbenzylamine; Quadrol *N,N,N',N'*-tetrakis(2-hydroxypropyl)ethylenediamine; PTH, phenylthiohydantoin; DTE, 1,4-dithioerythritol.

TABLE I: Programs for Use with DMBA Buffer.[a]

Step	Program Time (sec)		Cup Speed (rpm)
	General	Abbreviated	
Heptafluorobutyric acid delivery (reagent 3)	30	30	1200
Cleavage	200	200	1200
Restricted vacuum	30	30	1200
Rough vacuum	30	20	1200
Fine vacuum, N_2 flushing	20	20	1200
Chlorobutane extraction (solvent 3)	150 (5 ml)	100 (4 ml)	1800
Restricted vacuum	100	80	1800
Rough vacuum	20	10	1800
Fine vacuum, N_2 flushing	800	300	1800
Phenyl isothiocyanate delivery (reagent 1)	7	7	1800
Restricted vacuum	30	30	1800
N_2 blow to atmosphere	30	30	1800
Buffer delivery (reagent 2)	20	20	1200
Coupling reaction	1800	1800	1200
Restricted vacuum	400	300	1200
Rough vacuum	40	40	1200
Fine vacuum, N_2 flushing	50	100	1200
Benzene extraction (solvent 1)	300	300	1800
Restricted vacuum	100	80	1800
Rough vacuum	20	10	1800
Fine vacuum, N_2 flushing	1200	600	1800

[a] Bottle venting and pressuring and some N_2 blowing steps are not shown. Total time per cycle (including all steps) is approximately 95 min for the general program and 56 min for the abbreviated program. The complete program (39 steps) can be obtained from this laboratory on request.

boiling fraction was then redistilled through a 40-cm Vigreaux column under vacuum (approximately 1 mm).[2]

The following reagents and solvents were used in the operation of the sequenator (Table I).

REAGENT 1 is 5% (v/v) phenyl isothiocyanate in heptane.

REAGENT 2, the buffer, is a mixture of 12 ml of DMBA, 41 ml of propanol, and 48 ml of distilled water, titrated to pH 9.4 with reagent grade glacial acetic acid. This buffer is approximately 0.8 M in DMBA. Since this concentration is very close to the limit of solubility of DMBA in the mixture, the solution should be clarified, if necessary, by dropwise addition of propanol.

REAGENT 3 AND SOLVENT 1 are heptafluorobutyric acid and benzene, respectively.

SOLVENT 3 is 0.1% (v/v) fresh reagent grade ethanethiol in chlorobutane. When placing this reagent in the instrument the reagent is flushed with nitrogen for only 2 min. Longer flushing removes ethanethiol.

Sequenator. The sequenator used in these experiments was a Beckman Sequencer (Model 890A) designed to adapt the principles described by Edman and Begg (1967). A film of protein on the wall of a spinning cup is coupled in a nitrogen atmosphere with phenyl isothiocyanate in an appropriate buffer. Non-protein components are removed by a "rough" vacuum followed by a "fine" vacuum and finally by extraction with organic solvents. The dried protein film is then exposed to anhydrous heptafluorobutyric acid and the amino-terminal residue extracted in chlorobutane as a mixture of phenylthiohydantoin and thiazolinone. At this stage, one cycle of the Edman degradation is complete and a new cycle of the coupling and acid cleavage is begun. The chlorobutane extracts are separately treated to convert thiazolinone to thiohydantoin which is then identified as outlined below.

The particular instrument used was modified by insertion of a solenoid valve and a needle valve into the nitrogen line leading to the reaction chamber to introduce a regulated flow of nitrogen through the reaction chamber during the "fine" vacuum steps (see Table I). Thus the reaction chamber is purged with nitrogen during the evacuation, and the amount of condensation in the reaction chamber during operation with volatile buffers is greatly reduced. The N_2 flushing system is an optional addition to the Sequencer available from Beckman Instruments, Inc.

Programs. The standard programs of operation are shown in Table I. The general program is used when more than ten cycles are desired. The abbreviated program is faster and is satisfactory for ten cycles or less. The abbreviated program is sufficient for estimation of protein purity, for preliminary examination for internal breaks in a protein chain, or for amino-terminal analyses. Its usefulness is limited by the shortened drying periods which cause accumulation of condensates. Both programs start with a cleavage step so that the protein is always in the "coupled" state when the instrument is shut down automatically.

A cup speed of 1200 rpm was employed throughout the coupling and cleavage reactions. At the end of the delivery to the cup, the level of the heptafluorobutyric acid was 2–3 mm below the initial level of the buffer. This adjustment allowed for the partial evaporation of the buffer film during the coupling period and for the tendency of the acid to saturate the protein by capillary action. As a result the protein was completely covered during both the cleavage and coupling steps. Single cleavage was found to be adequate.

The temperature in the heated housing around the cup was 57°, measured with a calibrated thermometer lying on the metal base plate of the cup assembly.

Sequenator Maintenance. With proper programming and nitrogen flushing, there was little condensation on either the outer wall of the cup or the surrounding glass cylinder. However, some condensation inevitably occurred in the metal well housing the drive bearings below the cup. Hence the cup bearing assembly was removed after every extended run or after every 2 days and cleaned thoroughly with ethanol.

The use of a volatile buffer increased the accumulation of foreign material in the vacuum pump oil. With continuous operation of the sequenator, the oil in both pumps was replaced at least three times a week.

All glass test tubes in contact with sequenator products were cleaned in a hot solution of nitric and sulfuric acids (2:1, v/v), then rinsed exhaustively with distilled water, and dried.

Analysis of Sequenator Products. In order to avoid removal of the volatile thiol, the chlorobutane solutions of sequenator products were not automatically dried in the sequenator. Instead, the samples from 1 day's operation were dried quickly at approximately 60° under a stream of purified nitrogen and

[2] DMBA marketed as Sequenal grade by Pierce Chemical Co. is suitable for use without further purification.

immediately dissolved in 0.2 ml of 1 N HCl containing ethanethiol (1%, v/v). After 10-min standing at 80° to complete the cyclization of the phenylthiohydantoins (Edman and Begg, 1967), each solution was extracted with ethyl acetate. The ethyl acetate phase and the aqueous phase were separately examined for phenylthiohydantoins.

The ethyl acetate extracts were dried under nitrogen, silylated by adding 25–50 μl of *N,O*-bis(trimethylsilyl)acetamide (1-ml ampoules from Pierce Chemical Co., stored at 4°), agitated in a Vortex mixer, and heated to 60° for 10 min.

The silylated products were identified by gas chromatographic analysis using an adaptation of the procedure of Pisano *et al.* (1972). A silylated glass column (2 mm i.d. × 4 ft) was used. It contained "10%" SP-400 (Supelco Inc.) on deactivated acid-washed Chromosorb W. The sample (2–5 μl) contained 5–40 nmoles of the PTH-amino acid. The gas flow was 140 cm^3/min of helium and the instrument was programmed for temperatures ranging from approximately 190 to 290°. Programmatic adjustment of temperature was either linear or stepwise, depending on the capability of the instrument used. For optimal separation of silylated standard PTH-amino acids, minor adjustments in the temperature program were made for each newly packed column. Analyses were performed with either a Beckman GC-5 or a Hewlett-Packard 7620A instrument[3] equipped with flame ionization detectors. The yield of a PTH-amino acid was estimated by comparing the area under its peak on a gas chromatogram to that of a standard. Figure 1 shows the relative retention times of the silylated PTH-amino acids for a typical SP-400 column.

The large variations in polarity, volatility, and chemical stability among the PTH-amino acids make it necessary to sacrifice some desirable features of analysis in order to optimize others. A relatively polar stationary phase (*e.g.*, SP-400) is necessary in order to separate the phenylthiohydantoins of alanine, δ-serine, valine, glycine, leucine, and isoleucine, but this impairs the quantitative chromatography of the polar and high-boiling PTH derivatives. A less polar stationary phase such as XE-60 gives better quantitation of the polar derivatives but for general sequencing work the added precision does not justify the additional time and effort of operating and servicing the columns required for this purpose.

On SP-400 the silylated PTH derivatives of alanine, glycine, valine, leucine, isoleucine, methionine, phenylalanine, proline, and tyrosine chromatograph very well and are easily quantitated. PTH-tryptophan is slightly more susceptible to destruction, especially by oxidation, and therefore is less easily quantitated. PTH-serine and PTH-threonine are exceedingly labile and any attempt to quantitate these residues appears futile. PTH-lysine, -asparagine, and -glutamine are polar and chromatograph poorly on SP-400; the peaks observed are symmetrical and discrete but the detector response does not correspond to their molecular weights. Since areas under the peaks of these compounds are not proportional to the amount of compound injected, quantitation is difficult. The sequenator products of aspartic acid and glutamic acid do not extract well from the cup (Edman and Begg, 1967) and hence their quantitation is not meaningful. PTH-*S*-carboxymethylcysteine is degraded in the inlet port of the gas chromatograph and only PTH-δ-serine is observed. *S*-Methylcysteine appears to be the only cysteine derivative which can be quantitated by gas chromatography.

During repeated degradations, cumulative nonspecific cleavage of the protein unavoidably occurs. Consequently the background of PTH-amino acids gradually increases in amount with succeeding cycles. Thus in the later cycles, residues are identified by observing which peak increases relative to the background. Since the degradation is inevitably incomplete (approximately 96% recovery per cycle), the same residue will also be slightly above background in the subsequent cycle (termed "overlap").

Identification of residues that are labile or chromatograph poorly becomes difficult when the background and overlap are high and the overall yield of the residue is low. Serine, for instance, may be impossible to identify in a late cycle whereas valine, leucine, and alanine may be clearly identifiable ten or more cycles later. Hence sequenator experiments may end with blanks in the sequence rather than with an abrupt termination of useful data.

Thin-layer chromatography has essentially the same problems of residue identification. It is, however, easier to differentiate the residue at a given cycle from background on a gas chromatogram since gas chromatography is semiquantitative and not dependent on a subjective evaluation of spot intensity.

Identification of Histidine, Arginine, and Pyridylethylcysteine. Three PTH-amino acids remain in the aqueous phase after acid cyclization. Of these, PTH-histidine and PTH-arginine can be detected by diazotized *p*-anisidine and phenanthrenequinone spot tests, respectively,[4] whereas PTH-pyridylethylcysteine is identified by thin-layer chromatography. The aqueous solutions are dried at 60° in a stream of nitrogen, then dissolved in 0.02 ml of methanol. A series of 0.005-ml aliquots of each methanolic solution is dried as spots on a strip of chromatography paper for detection of PTH-histidine and separately for detection of PTH-arginine. It is particularly important that all the aqueous layers be spotted in sequence on the paper to clearly differentiate histidine and arginine residues from the rising background and from overlap of the preceding cycle.

The histidine test will detect 15 nmoles and is performed according to the method of Sanger and Tuppy (1951). It is important that diazotization of *p*-anisidine proceeds for only 3–5 min before spraying the paper.

The arginine test will detect 10 nmoles and is performed essentially according to the procedure of Yamada and Itano (1966). Phenanthrenequinone (5 mg) is dissolved in 25 ml of absolute ethanol. A 4-ml aliquot of this solution is swirled on a Vortex mixer while adding 1 ml of 25% NaOH. The paper is dipped in this solution and air-dried for 20 min. Fluorescence under uv light is taken as a positive indication of PTH-arginine.

PTH-pyridylethylcysteine is identified by thin-layer chromatography with ethyl acetate on fluorescent silica gel (F-254, Brinkmann). The R_F of PTH-pyridylethylcysteine is approximately 0.4 while PTH-histidine and PTH-arginine remain at the origin.

[3] Aliquots (6 μl) can be removed from 15-μl samples in the automated Hewlett-Packard instrument by the use of disposable polyethylene 0.4-ml microcentrifuge tube inserts (Arthur H. Thomas 2591-D15). These tube inserts are forced into moistened glass ampoules (Hewlett-Packard sample vials) and the protruding tops are removed with a razor blade. The silylated sample is added and the ampoule capped. A special circuit board available from Hewlett-Packard (Option 504) controls a rinse of the syringe after each sample injection.

[4] If the water layers are dried and silylated, both PTH-histidine and PTH-arginine can be recognized by gas chromatography. The histidine derivative appears as a discrete peak just before silylated PTH-tyrosine; the arginine derivative appears as a series of broad peaks near PTH-glutamic acid, probably corresponding to pyrolysis products.

TABLE II: Sequenator Analysis of Thermolysin and Fragments.[a]

Fragment	Method of Preparation	Residue No.[b]	Residues Placed by Sequenator: No. of Degradations	Residues Placed by Sequenator: Residues Identified
Thermolysin		1–316	15	1–15
F_{III}	Cleavage with CNBr[a]	1–120	33	1–33
S-F–$_{III}T_1$	Tryptic cleavage of succinylated F_{III}[b]	48–90	34	48–81
F_I	Cleavage with CNBr[a]	121–205	48	121–168
F_I–T_2	Tryptic cleavage of F_I[b]	183–205	8	183–190
F_{II}	Cleavage with CNBr[a]	206–316	32	206–237
F_{II}–HA_2	Cleavage of F_{II} with NH_2OH[c]	228–316	44	228–271

[a] The preparation of thermolysin and its cyanogen bromide fragments has been described by Titani *et al.* (1972b). [b] See Titani *et al.* (1972a). [c] Titani *et al.* (1972a) using the method of Bornstein (1969).

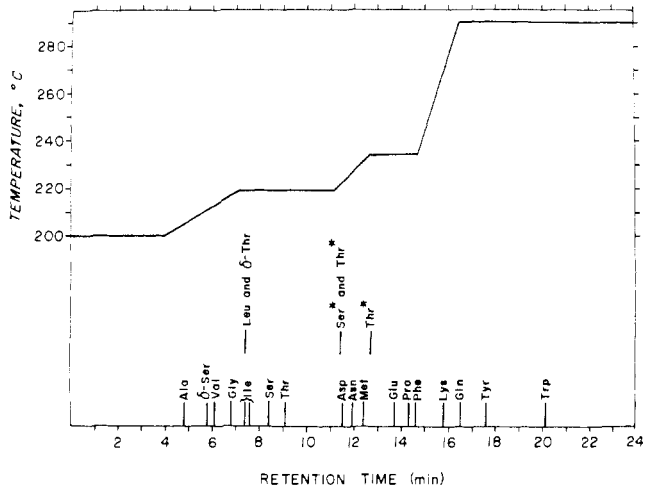

FIGURE 1: Retention times of silylated sequenator products during gas–liquid chromatography. The solid line traces the temperature program. Isoleucine residues yield a doublet at 7.4 and 7.6 min whereas leucine residues yield a single peak at 7.4 min (probably due to the formation of a diasteriomeric pair during silylation of PTH-isoleucine). Thiols are included in extraction and cyclization procedures and influence the nature of the silylation products of PTH-proline, PTH-glycine, PTH-serine, and PTH-threonine (see text). With ethanethiol, serine residues yield both silylated PTH-δ-serine (the basis of their identification) and a presumed reaction product of silylated PTH-δ-serine and ethanethiol which cochromatographs with the aspartyl derivative (*). Threonine residues also yield multiple peaks: silylated PTH-δ-threonine which cochromatographs with the leucyl derivative (in low yield), silylated PTH-threonine, and a pair of thiol-dependent peaks (*) near the aspartyl and methionyl derivatives.

Preparation of Proteins for Sequenator Analysis. In general polypeptides are prepared for sequence analysis under conditions which minimize contamination or cleavage by proteases. Prior to analysis the polypeptides are tested for homogeneity by polyacrylamide gel electrophoresis in the presence and absence of sodium dodecyl sulfate.

Cysteine and cystine are alkylated by an agent of choice.[5] *S*-β-Pyridylethylation with 4-vinylpyridine (Friedman *et al.*, 1970) yields a suitable product. The derivative, pyridylethylcysteine, carries a positive charge in acid and, since this charge is not lost during treatment with phenyl isothiocyanate, extraction of the peptide in the chlorobutane is minimized.

The PTH derivative of *S*-carboxymethylcysteine appears to break down in the inlet port of the gas chromatograph since only PTH-δ-serine is observed after injection of PTH-*S*-carboxymethylcysteine. This renders the identification of both serine and cysteine ambiguous since PTH-serine is recognized primarily on the basis of the accompanying PTH-δ-serine peak. A simple way to resolve this ambiguity is to carboxymethylate the protein with [^{14}C]iodoacetate (giving about 10,000 cpm/μmole of cysteine) and to scan aliquots of the chlorobutane extracts for ^{14}C.

Aminoethylation with ethylenimine (Raftery and Cole, 1963) is not desirable for sequenator analysis since the new amino groups react with phenyl isothiocyanate adding to the hydrophobicity and lowering the solubility of the protein in the coupling buffer. In addition, the PTH derivative of *S*-aminoethylcysteine does not chromatograph on SP-400 and has the mobility of PTH-lysine on thin-layer chromatography.

S-Methylation with methyl-*p*-nitrobenzensulfonate (Heinrikson, 1971) yields a derivative which chromatographs well in a unique position on SP-400 (between silylated PTH-threonine and PTH-aspartic acid). However this does not appear to be a suitable derivative because S-methylated polypeptides are quite insoluble and difficult to purify.

The *S*-alkyl proteins are exhaustively dialyzed or passed over Sephadex to remove all salts and reagents. It is especially important to remove urea or guanidine salts since both are deleterious to subsequent sequenator analysis. The salt-free protein (5–10 mg) is dissolved in 0.4 ml of a volatile solvent (*e.g.*, 10–50% aqueous acetic acid) and dried to a film on the lower one-third of the wall of the sequenator cup.

Proteins. PORCINE TRYPSIN (Novo Industri, Copenhagen) was purified by affinity chromatography on insoluble chicken ovomucoid according to Robinson *et al.* (1971).

Three fragments of THERMOLYSIN were derived by cleavage with cyanogen bromide and gel filtration (Titani *et al.*, 1972b). These fragments and certain subfragments derived from them are identified in Table II.

Results

Choice of Buffer for the Coupling Reaction. Quadrol buffer (as described by Edman and Begg, 1967) has a number of desirable characteristics for automated sequencing of poly-

[5] It was possible to avoid this procedure in some cases, *e.g.*, bovine trypsin. Unmodified Cys_7 in trypsin yielded no extractable product but residues 8, 9, etc., were normally degraded.

peptides. It is nonvolatile and hence does not cause condensation in the reaction chamber; it is an excellent protein solvent having a pK_a (9) in the desired pH range. However, Quadrol has a number of undesirable characteristics. It is difficult to purify; it is poorly soluble in benzene; and since the buffer salt is practically insoluble in benzene, extraction with ethyl acetate is required to remove the buffer from the protein film. Besides, hydrophobic peptides and short peptides (less than 40 residues) are also extracted by ethyl acetate and their loss represents the most serious practical problem encountered in the use of Quadrol.

Buffers of the same general formula ($RR'NCH_2CH_2N$-$R''R'''$) have pK_a values in the desired pH range. However, salts of these compounds are generally very insoluble in organic solvents including ethyl acetate. Tertiary amines (RR'-$R''N$) are too basic for use if all three substituents are simple alkyl groups. Substituted anilines are not basic enough. Substituted allylamines and benzylamines dissociate in the required range. However, dimethylallylamine buffer, commonly used in manual Edman degradations (Edman, 1970) and in the sequenator (Niall *et al.*, 1969), is quite volatile and must be added repeatedly during the coupling period to maintain the pH in the desired range.

Dimethylbenzylamine (DMBA) appears to be a satisfactory alternative. It has a pK_a of 8.9 (Perrin, 1965), it boils at 180° under atmospheric pressure making it quite nonvolatile in the sequenator, and its acetate, trifluoroacetate, and heptafluorobutyrate salts are quite soluble in benzene, eliminating the need for extraction with ethyl acetate.

Satisfactory coupling requires simultaneous solution of buffer, phenyl isothiocyanate, and polypeptide. To achieve this condition, 1-propanol–water or pyridine–water mixtures were used with varying success. With sperm-whale myoglobin at pH 9.2 in DMBA, consecutive experiments gave repetitive yields of 94% in 40% aqueous pyridine buffer and just over 96% in the aqueous propanol buffer described in Methods. (These yields were based on the recoveries of Val_1, Val_{10}, Leu_2, and Leu_{11} in the myoglobin sequence.) Since an increase in the repetitive yield from 94 to 96% permits the identification of 15–20 additional residues, other factors being equal, the propanol buffer was judged to be superior provided the protein remained soluble.

DMBA was found to be suitable for protein fragments containing 50–150 residues. Since DMBA is semivolatile at 50–60°, it was difficult to avoid condensation of DMBA-heptafluorobutyrate on the walls of the reaction chamber. Static vacuums, no matter how low the pressure, left traces of buffer and acid after each step of the cycle, forming a fog of condensate in the succeeding step. By modifying the sequenator to allow a slow bleed of nitrogen through the cup during the fine vacuum steps, the atmosphere in the reaction chamber was purged and very little condensation occurred.

Preliminary testing indicated that *N*-allylpiperidine has similar properties to DMBA and may also be a suitable buffer.

Thiols. Addition of dithioerythritol (DTE) to the chlorobutane preserved recognizable products corresponding to serine and threonine (Hermodson *et al.*, 1970) and at the same time increased the yields of the other PTH-amino acids to 90–100%.

Since even the most highly purified DTE available interfered to some extent with gas chromatography of the sequenator products, alternate thiols were examined. Mercaptoethyl ether and butanedithiol proved acceptable with regard to stability of PTH derivatives, particularly those of serine and threonine, but these thiols tended to produce increasing amounts of spurious gas chromatographic peaks during storage, probably by oxidative polymerization. Initial attempts to use highly volatile thiols in the sequenator failed because they evaporated during the automatic sample-drying procedure. This problem was not alleviated by storing the products in the chlorobutane solution containing ethanethiol, since decomposition of the serine and threonine derivatives took place during the aqueous acid treatment. Optimal yields of PTH-serine were obtained by including ethanethiol in both the extracting and cyclizing solutions (see Methods), and this thiol was adopted. In the presence of ethanethiol, serine and threonine yielded additional gas chromatographic peaks. The additional peak for serine appeared near silylated PTH-aspartic acid (illustrated later in Figure 3, cycle 14) and the two additional peaks for threonine appeared near PTH-aspartic acid and after PTH-methionine (Figure 1). Thus the mercaptan may be reacting with degradation products of PTH-serine and PTH-threonine (*e.g.*, PTH-δ-serine and PTH-δ-threonine) producing new stable compounds.[6]

PTH-proline and PTH-hydroxyproline also react with the silylating agent when any thiol is used. The product of PTH-proline appears between silylated PTH-glutamic acid and PTH-phenylalanine, whereas the product of PTH-hydroxyproline appears between PTH-glutamine and PTH-tyrosine. Small amounts (5–10%) of unsilylated PTH-proline and PTH-hydroxyproline are observed in the sequenator samples. Unsilylated PTH-proline chromatographs near silylated PTH-glycine (illustrated later in Figure 2, cycle 3). The mobilities of other PTH-amino acids are not affected by the inclusion of thiols.

Incomplete Degradation of Proline. When the sequenator was operated at approximately 50°, the degradation of proline residues was incomplete and an overlap of approximately 20% was observed in the succeeding cycles. The yield was not improved by cleaving twice, by changing buffers (DMBA in propanol or pyridine, Quadrol, and *n*-allylpiperidine), by lengthening or shortening the coupling time, or by coupling twice. The problem was minimized by raising the temperature and was aggravated by lowering the temperature in the reaction chamber. At 56–57° the overlap after proline residues was about 5–8%.

Limitation of Size of Polypeptides. The size of the peptide influences the number of definitive degradations. Since polypeptides containing more than 250–300 amino acid residues are poorly soluble in the coupling buffer, significant overlap occurs due to incomplete reaction. Large proteins also generate background quickly since nonspecific cleavage can occur at more sites. Also the larger the protein, the smaller the molar quantity which can be placed into the cup (10 mg appears to be a practical upper limit in most cases). Small peptides (less than 30 residues) tend to be extracted from the cup into the chlorobutane phase, particularly if they are hydrophobic or if the film of peptide is not continuous on the wall of the cup. To prevent this loss, the sample size should be 7–10 mg.

Peptides derived by *enzymatic* fragmentation of proteins are usually too small for sequenator analysis with the excep-

[6] One anomaly has been observed with the use of ethanethiol. Occasionally the conversion of PTH-valine and PTH-isoleucine does not go to completion, resulting in a peak appearing prior to any PTH-amino acid in the gas chromatogram and in diminished yield of PTH-valine or PTH-isoleucine (*e.g.*, see Figure 3, cycle 1). Longer conversion times or careful mixing of the HCl solutions usually prevents this anomaly. The occurrence of this phenomenon is erratic but has not interfered with the positive identification of the residue.

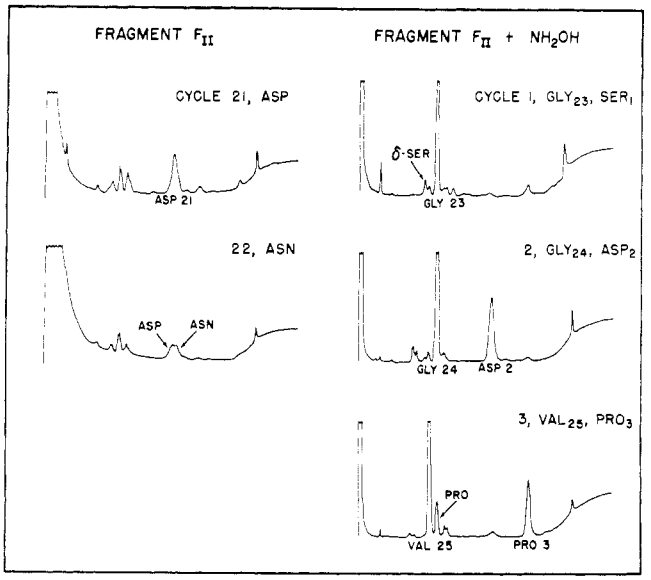

FIGURE 2: Gas–liquid chromatographic analyses of silylated sequenator products derived from the sequence Asp_{21}-Asn-Gly-Gly-Val_{25} in fragment F_{II} of thermolysin (residues 226–230, Table II). The number above each chromatogram refers to the degradation cycle under examination. The number beside each amino acid refers to its position in the sequence of F_{II}. (Since F_{II} represents residues 206–316 of thermolysin, Asp_{21} corresponds to Asp_{226} of thermolysin.) The vertical scale records the response of the detector in arbitrary units. The tops of large peaks are not always illustrated. The horizontal scale is approximately 20 min/analysis. Left: a sharp drop in product yields is seen between cycles 21 and 22, corresponding to Asp_{21} and Asn_{22} in the sequence Asp_{21}-Asn-Gly-. Right: after cleavage of F_{II} with hydroxylamine (see text), the first three cycles yield both the amino-terminal sequence of F_{II} (Ser-Asp-Pro) and the internal sequence Gly_{23}-Gly-Val-. (The yield of Ser-Asp-Pro is low because the peptide representing residues 1–22 was partially lost during dialysis.)

tion of fragments obtained by tryptic cleavage of N-acylated proteins. The carboxyl-terminal arginine residues reduce the solubility of such fragments in the extracting solvents.

Three methods of *chemical* cleavage of proteins yield fragments suitable for sequenator analysis. Fragmentation by cyanogen bromide at methionine residues is both specific and quantitative (Gross, 1967). Cleavage at tryptophan residues has been successfully employed by Niall *et al.* (1971). Cleavage with hydroxylamine at asparaginyl–glycyl bonds (Bornstein, 1969) has been applied successfully on several occasions in our laboratory (*e.g.*, Titani *et al.*, 1972a). Since asparaginylglycine sequences are relatively infrequent, the resulting fragments are quite large.

Low Repetitive Yields. The yield of product at each cycle of the degradation is 95–96% of the product yield in the previous cycle (this has been termed the "stepwise" or "repetitive" yield). Lower yields occur occasionally with particular sequences. For example, during acid cleavage of the preceding residue, certain glutamine residues are partially converted to pyrrolidone derivatives and hence their yield drops 5–10%. The product of this side reaction is stable and does not interfere with subsequent degradation of the unblocked fraction. The low yield of prolyl residues can be largely overcome by raising the temperature to 57° (see above).

Asparaginylglycine sequences also undergo cyclization in acid (Bornstein, 1969) and since the resulting cyclic imide is not susceptible to Edman degradation, the stepwise yield will be low (approximately 50%). However, hydroxylamine cleaves the imide (Bornstein, 1969) and generates an amino-terminal glycyl residue which can serve as a new starting point for sequenator analysis. These results are illustrated in Figure 2 for cyanogen bromide fragment F_{II} of thermolysin (see also Table II). It is evident that the repetitive yield falls off sharply between Asp_{21} and Asn_{22}. When F_{II} was first exposed to 1 M hydroxylamine, dialyzed overnight against 5% acetic acid, and then subjected to sequenator analysis, the original amino-terminal sequence and the newly formed sequence beginning with Gly_{23} (Gly_{23}-Gly_{24}-Val_{25}-) were degraded simultaneously (Figure 2). Since the amino-terminal sequence of the intact fragment F_{II} had already been established, there was no need to separate the products of hydroxylamine treatment prior to analysis.

Degradation of Specific Proteins. THERMOLYSIN. The reliability and efficiency of the sequenator technique were demonstrated by sequence analysis of the cyanogen bromide fragments and subfragments of thermolysin (Titani *et al.*, 1972a,b). The identification of the fragments and the portions placed by the sequenator are given in Table II. The sequences derived by sequenator analyses were confirmed by conventional methods and in no instance did a positive identification by the sequenator fail to be corroborated by conventional procedures. Many of the tryptic peptides from thermolysin could be aligned solely from the sequenator data. These alignments were confirmed both by isolating overlapping peptides (Titani *et al.*, 1972a) and by agreement with crystallographic analysis (Matthews *et al.*, 1972).

The analysis of thermolysin demonstrated several other advantages of the sequenator technique. First, a number of sequences were readily determined which are difficult by conventional methods, *e.g.*, Tyr-Tyr-Tyr-Leu (residues 27–30), Gln-Asp-Asn (residues 31–33), Asp-Asn-Gln (residues 59–61), Gln-Asn-Glu (residues 158–160), and Gln-Asp-Asn (residues 225–227). Second, three residues (140, 237, and 256) were placed which could have been overlooked in conventional analyses because they involve peptide bonds that are relatively resistant to acid hydrolysis (Val-Val and Ile-Ile). Third, chymotryptic and thermolytic peptides of a tryptic peptide of cyanogen bromide fragment F_I (Val_{121}–Lys_{182}), which could not be ordered by conventional means, were aligned by sequenator analysis of the first 48 residues of fragment F_I.

Analyses of fragments of thermolysin were greatly facilitated by the introduction of the DMBA buffer. The three cyanogen bromide fragments, F_{III}, F_I, and F_{II} (Table II), contain 120, 85, and 111 residues, respectively. Analysis in Quadrol buffer involved troublesome losses of peptides from the sequenator cup and fewer than 15 degradations of each of these fragments were definitive. In DMBA buffer as many as 48 residues of fragment F_I and 44 residues of fragment $F_{II}HA_2$ could be identified (Table II). In addition, 34 residues were placed in a 43-residue tryptic peptide from fragment F_{III} (S–F_{III}–T_1 in Table II).

To illustrate the data obtained in an extended analysis, fragment F_I (see Table II) was subjected to 52 consecutive degradations. Gas chromatograms of 20 of these products are traced in Figure 3 where cycles 1–52 of fragment F_I correspond to residues 121–172 in the thermolysin sequence (Titani *et al.*, 1972a). The first six analyses demonstrate the purity of the fragment (see footnote 6 regarding the two valine derivatives in cycle 1). In each cycle the peak of a single major PTH-amino acid increases above background levels and decreases in the following cycle. As degradations proceed, a progressive increase in "overlap" is seen (*e.g.*, by contrasting

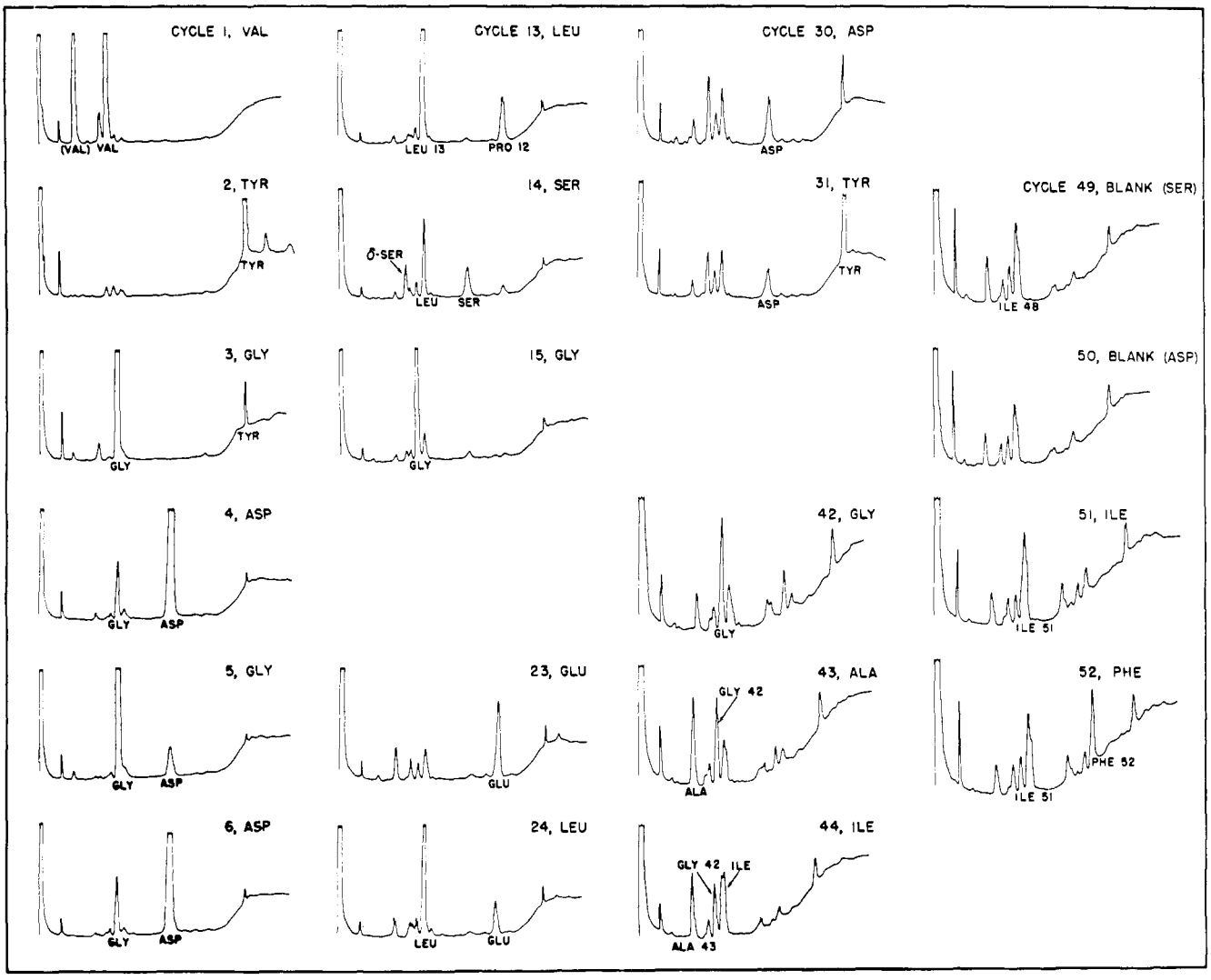

FIGURE 3: Selected gas chromatographic analyses of silylated sequenator products of fragment F_1 of thermolysin (residues 121-205, Table II). The number above each analysis refers to the degradation cycle under examination. (Since F_1 represents residues 121-205 of thermolysin, valine in cycle 1 corresponds to Val_{121} in thermolysin.) The horizontal scale represents approximately 20 min/analysis; the vertical scale (arbitrary units) is the detector response which is electronically attenuated to a greater degree in the earlier cycles than in the later cycles. The tops of large peaks are not always illustrated. The injected aliquot is progressively increased from 3 μl of a 50-μl sample in cycle 1 to 6 μl of a 20-μl sample in cycle 52. Cycles 41 and 48 (not shown) yielded serine and isoleucine; residues 49 and 50 were identified by other means as serine and aspartic acid.

the very small overlap of Gly_3 in cycle 4 with the much larger overlap of Gly_{42} in cycle 43). In addition, the background level of PTH-amino acids rises relative to the size of the principal product. This is the result of a gradual formation of new amino-terminal residues by nonspecific acid cleavage and a gradual decrease in the principal product arising from incomplete stepwise yield. This aspect of the analysis is somewhat distorted in Figure 3 by the gradual expansion of the vertical scale with increasing cycle numbers. In spite of these problems of interpretation, definitive identifications were possible up to Ile_{48} (corresponding to Ile_{168} in thermolysin). Neither cycle 49 nor cycle 50 yielded an identifiable product other than the overlap of Ile_{48}. Increases in isoleucine in cycle 51 and in phenylalanine in cycle 52 were the last identifiable products in this analysis.

PORCINE TRYPSIN. Sequenator analysis of porcine trypsin has provided evidence of heterogeneity due to (1) amino acid substitution and (2) an internal split in the polypeptide chain. Both of these phenomena have been previously observed in proteolytic enzymes: amino acid substitutions in dogfish trypsin (Bradshaw *et al.*, 1970) and in bovine carboxypeptidase A (Pétra *et al.*, 1969) and internal splits in bovine trypsin (Schroeder and Shaw, 1968; Smith and Shaw, 1969; Maroux and Desnuelle, 1969) and in bovine carboxypeptidase B (Reeck *et al.*, 1971). An internal split between residues 131 and 132 of porcine trypsin[7] was manifest by the simultaneous appearance of a major and minor (5%) peptide sequence when the reduced and carboxymethylated enzyme was subjected to sequenator analysis. The major sequence (β-trypsin) was homologous to the amino terminus of bovine trypsin (11 degradations) and the minor sequence (of α-trypsin) to the sequence of the bovine enzyme starting with residue 132. Gas chromatograms of cycles 10–13 are shown in Figure 4. The major sequence yields in cycle 10 asparagine and the

[7] Autolytic cleavage at positions 131–132 of bovine trypsin has been characterized by Schroeder and Shaw (1968).

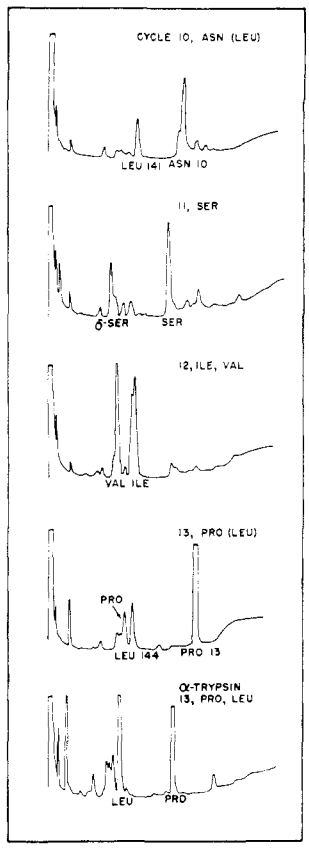

FIGURE 4: Gas–liquid chromatographic analyses of silylated sequenator products from cycles 10–13 of 20:1 mixtures of β- and α-porcine trypsins. The tops of large peaks are not always illustrated. The leucine residues in parentheses are derived from α-trypsin at positions 141 and 144. Cycle 12 reveals two major products, apparently resulting from two allotypic species of trypsin. In the bottom two chromatograms cycle 13 of α-trypsin is compared with that of the 20:1 mixture of β:α (second from bottom). The sequence of porcine trypsin (L. H. Ericsson and M. A. Hermodson, unpublished data) through these areas is:

$$Ile_1\text{-}Val_2\text{-}Gly_3\text{-}Gly_4\text{-}Tyr_5\text{-}Thr_6\text{-}Cys_7\text{-}Ala_8\text{-}Ala_9\text{-}Asn_{10}\text{-}Ser_{11}\text{-}\frac{Ile}{Val_{12}}\text{-}$$

$$Pro_{13}\text{-}\cdots\text{-}Ser_{132}\text{-}Ser_{133}\text{-}Gly_{134}\text{-}Ser_{135}\text{-}Ser_{136}\text{-}Tyr_{137}\text{-}Pro_{138}\text{-}Ser_{139}\text{-}$$

$$Leu_{140}\text{-}Leu_{141}\text{-}Gln_{142}\text{-}Cys_{143}\text{-}Leu_{144}$$

minor sequence leucine (residue 141). In cycle 11, serine is the major product whereas the corresponding residue of the minor sequence (glutamine) is not detectable. In cycle 12, valine and isoleucine (in a 2:1 ratio) together account for the major sequence in this position, whereas carboxymethylcysteine of the minor sequence was not detectable. In cycle 13, proline is clearly the major product and leucine the minor one.

Recently the porcine enzyme containing the internal split at residues 131–132 (α-trypsin) was separated from the single-chain β form (L. H. Ericsson and M. A. Hermodson, unpublished data). Sequenator analysis of α-trypsin yielded two residues per cycle through 28 degradations. Since the amino-terminal sequence of the intact polypeptide chain had already been determined, the sequence of the 28 residues following residue 131 could be established from these data. Cycles 11 and 12 placed glutamine and carboxymethylcysteine at positions 142 and 143 and confirmed the isoleucine–valine replacement at position 12. Cycle 13 of α-trypsin reveals comparable quantities of Pro_{13} and Leu_{144} (Figure 4).

Discussion

The primary purpose of the automated Edman degradation method is the determination of amino acid sequences. At the present stage of development, 30–50 consecutive residues can be identified and recovered with a stepwise yield of approximately 96%. When the sequenator is in full operation, approximately 90 amino acid residues can be determined per week. Reliability of the method of analysis has been proven by the complete agreement between amino acid sequences of thermolysin derived by conventional methods, on the one hand, and by the sequenator on the other. Sequenator analysis has the additional advantages of rapidity, of direct identification of the amide-containing residues, and of determination of repetitive residues which can present difficulties in conventional sequence determinations (*e.g.*, Val-Val and Tyr-Tyr-Tyr).

The major advantage of sequenator analysis is its ability to determine extended amino acid sequences (30–50 residues) in a single operation. This procedure eliminates the necessity for extensive fragmentation of the protein prior to sequence analysis and consequently reduces the number of overlaps necessary for aligning the fragments. In the specific case of thermolysin (Titani *et al.*, 1972a) the conventional approach involved the isolation and structural analysis of 173 small peptides from enzymatic digests. In contrast, sequenator analysis of six large fragments provided 60% of the total structure. Although in this case sequenator analysis served to confirm the structure previously established by conventional methods, sequential degradation of fewer and larger fragments would have provided a better starting point for the complete structure analysis. In the case of amyloid protein A containing 76 amino acid residues (Hermodson *et al.*, 1972), sequenator analysis of the whole protein and of a large cyanogen bromide fragment provided the sequence of all but the six carboxyl-terminal amino acid residues.

In addition to the primary purpose of accelerating the determination of amino acid sequences, the sequenator has proven useful in solving other problems usually encountered in the structural analysis of proteins. For instance, in the past a crucial test for purity of a protein has been the demonstration of a single amino-terminal residue by the quantitative cyanate method (Stark, 1967). A few degradations in the sequenator can provide the same information with a smaller quantity of protein and in addition can identify succeeding residues. This additional information eliminates the possibility that the protein in question may contain a mixture of other proteins having the same amino-terminal residue.

However, single chain proteins may yield multiple sequences in the sequenator if internal peptide bonds have been cleaved prior to sequence analysis. The two most prevalent causes of peptide-bond cleavage are limited proteolysis and instability of certain peptide bonds toward acids. The folllowing examples show the usefulness of sequenator analysis in locating the internal split and in relating the simultaneous products of sequenator analysis to the amino acid sequence of the original, single polypeptide chain. (1) Sequenator analysis of human amyloid protein A (Benditt *et al.*, 1971) revealed two residues per cycle in a yield ratio of approximately 7:1. The major sequence was H-Arg-Ser-Phe-Phe-Ser-Phe-Leu-Gly-, the minor sequence H-Phe-Ser-Phe-Leu-Gly-. The two sequences were related in that the minor sequence lacked the amino-terminal tripeptide Arg-Ser-Phe. (2) Analysis of nerve growth factor from mouse submaxillary glands (Angeletti and Bradshaw, 1971) gave two residues per

cycle in approximately equal yield (M. A. Hermodson, R. H. Angeletti, and R. A. Bradshaw, unpublished data). No unique sequences could be deduced from the residue yields in each cycle until a repetitive pattern was observed beginning with the ninth cycle by pairing the residues in cycles 1 and 9, 2 and 10, 3 and 11, etc. These data showed that approximately one-half of the preparation of nerve growth factor lacked the first eight amino acid residues. (3) Heterogeneity due to the acid lability of aspartyl–proline bonds (Piszkiewicz *et al.*, 1970) was clearly evident in the carboxyl-terminal cyanogen bromide fragment of thermolysin (residues 206–316). Sequenator analysis revealed two minor sequences comprising approximately 5–10% of the major sequence. Both minor sequences started with proline, one appearing at cycle 3 of the major sequence and the other at cycle 9 (corresponding to residues 208 and 214, respectively).

In several instances limited proteolysis of internal peptide bonds has resulted in a protein which is apparently pure by the usual electrophoretic and hydrodynamic criteria but which yields multiple sequences in sequenator analysis. The newly established amino-terminal residue provided an opportunity for the determination of extended, internal amino acid sequences. Bovine carboxypeptidase B is a case in point: it yielded an additional 26 amino acid residue sequence in the interior of the molecule and served to strengthen the hypothesis of homology with carboxypeptidase A (Bradshaw *et al.*, 1969; Reeck *et al.*, 1971).

Sequenator analysis of whole proteins has facilitated the proof of homology. For instance, comparison of the first 20 amino-terminal residues of bovine, dogfish, and lungfish trypsinogen has established their homologous relationships (Hermodson *et al.*, 1971). Smithies *et al.* (1971) have made similar use of sequenator analysis to compare light chains of Bence-Jones proteins; Rochat *et al.* (1970) have compared various toxins; and Niall *et al.* (1971) have examined related hormone structures. In our laboratory the homology between amyloid proteins of man, monkey, and Pekin duck (Benditt *et al.*, 1971; L. H. Ericsson and E. P. Benditt, unpublished observations) has been established by analysis of the whole proteins in the sequenator. The strength of the argument for homology based on sequenator data lies in the direct comparison of common regions in the respective polypeptide chains.

The direct determination of amino acid sequences in the sequenator has also given ready evidence for amino acid substitutions in proteins. A protein preparation obtained from pooled animal tissue may contain a mixture of allotypes which would complicate the assembly of sequence information. For instance, bovine β-lactoglobulin and bovine pancreatic carboxypeptidase A each occur in the form of two allotypes of comparable gene frequencies (Ashaffenburg and Drewry, 1957; Pétra *et al.*, 1969). If the allotypic change occurs in a portion of the sequence accessible to the sequenator, such a mutation is readily and unequivocally identified. In the case of porcine trypsin, such a mutation apparently has occurred in position 12 (Figure 4).

Sequenator analysis can resolve the number and identity of protomers in an oligomeric protein. In one specific case examined in this laboratory (aspartokinase–homoserine dehydrogenase I)[8] sequence analysis of the whole protein clearly established that the four subunits are identical throughout the first eight amino-terminal residues, suggesting that the subunits are either identical or homologous.

The examples given in this report are illustrative of the variable applications of sequenator analysis in determining the covalent structure of proteins. The versatility of the method is evident from the examples given, but further improvements in the operation of the method, particularly in the number of consecutive degradations, will further enhance the usefulness of the method in the direct determination of the amino acid sequence of yet larger protein fragments.

[8] This enzyme was provided by Professor Georges Cohen, Institut Pasteur, Paris, and the detailed results will be communicated separately by his laboratory.

References

Angeletti, R. H., and Bradshaw, R. A. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2417.

Ashaffenburg, R., and Drewry, J. (1957), *Nature (London) 180*, 376.

Benditt, E. P., Eriksen, N., Hermodson, M. A., and Ericsson, L. H. (1971), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 19*, 169.

Bornstein, P. (1969), *Biochem. Biophys. Res. Commun. 36*, 957.

Bradshaw, R. A., Neurath, H., Tye, R. W., Walsh, K. A., and Winter, W. P. (1970), *Nature (London) 226*, 237.

Bradshaw, R. A., Neurath, H., and Walsh, K. A. (1969), *Proc. Nat. Acad. Sci. U. S. 63*, 406.

Brewer, H. B., Jr., and Ronan, R. (1970), *Proc. Nat. Acad. Sci. U. S. 67*, 1862.

Edman, P. (1956), *Acta Chem. Scand. 10*, 761.

Edman, P. (1970), *in* Protein Sequence Determination, Needleman, S. B., Ed., Berlin, Springer-Verlag, p 211.

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1*, 80.

Friedman, M., Krull, L. H., and Cavins, J. F. (1970), *J. Biol. Chem. 245*, 3868.

Glenner, G. G., Terry, W., Harada, M., Isersky, C., and Page, D. (1971), *Science 172*, 1150.

Gross, E. (1967), *Methods Enzymol. 11*, 238.

Heinrikson, R. L. (1971), *J. Biol. Chem. 246*, 4090.

Hermodson, M. A., Ericsson, L. H., and Walsh, K. A. (1970), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 29*, 728.

Hermodson, M. A., Kuhn, R. W., Walsh, K. A., Neurath, H., Eriksen, N., and Benditt, E. P. (1972), *Biochemistry 11*, 2934.

Hermodson, M. A., Tye, R. W., Reeck, G. R., Neurath, H., and Walsh, K. A. (1971), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 14*, 222.

Hood, L. E., Potter, M., and McKean, D. J. (1970), *Science 170*, 1207.

Jaton, J. C., Waterfield, M. D., Margolies, M. N., Bloch, K. J., and Haber, E. (1971), *Biochemistry 10*, 1583.

Maroux, S., and Desnuelle, P. (1969), *Biochim. Biophys. Acta 181*, 59.

Matthews, B. W., Jansonius, J. N., Colman, P. M., Schoenborn, B. P., and Dupourque, D. (1972), *Nature (London) New Biol. 238*, 37.

Morgan, F. J., and Henschen, A. (1969), *Biochim. Biophys. Acta 181*, 93.

Niall, H. D., and Edman, P. (1967), *Nature (London) 216*, 262.

Niall, H. D., Keutmann, H. T., Copp, D. H., and Potts, J. T., Jr. (1969), *Proc. Nat. Acad. Sci. U. S. 64*, 771.

Niall, H. D., Keutmann, H. T., Sauer, R., Hogan, M., Dawson, B., Aurbach, G., and Potts, J. T., Jr. (1970), *Hoppe-Seyler's Z. Physiol. Chem. 351*, 1586.

Perrin, D. D. (1965), *in* Dissociation Constants of Organic Bases in Aqueous Solutions, London, Butterworth and Co., Ltd., p 118.

Petra, P. H., Bradshaw, R. A., Walsh, K. A., and Neurath, H. (1969), *Biochemistry 8*, 2762.
Petra, P. H., Hermodson, M. A., Walsh, K. A., and Neurath, H. (1971), *Biochemistry 10*, 4023.
Pisano, J. J., Bronzert, T. J., and Brewer, H. B., Jr. (1972), *Anal. Biochem. 45*, 43.
Piszkiewicz, D., Landon, M., and Smith, E. L. (1970), *Biochem. Biophys. Res. Commun. 40*, 1173.
Raftery, M. A., and Cole, R. D. (1963), *Biochem. Biophys. Res. Commun. 10*, 467.
Reeck, G. R., Walsh, K. A., Hermodson, M. A., and Neurath, H. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 1226.
Robinson, N. C., Tye, R. W., Neurath, H., and Walsh, K. A. (1971), *Biochemistry 10*, 2743.
Rochat, H., Rochat, C., Kupeyan, C., Miranda, F., Lissitzky, S., and Edman, P. (1970), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 10*, 349.
Sanger, F., and Tuppy, H. (1951), *Biochem. J. 49*, 463.
Schroeder, D. D., and Shaw, E. (1968), *J. Biol. Chem. 243*, 2943.
Smith, R. L., and Shaw, E. (1969), *J. Biol. Chem. 244*, 4704.
Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971), *Biochemistry 10*, 4912.
Stark, G. R. (1967), *Methods Enzymol. 11*, 125.
Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1972a), *Nature (London) New Biol. 238*, 35.
Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1972b), *Biochemistry 11*, 2427.
Yamada, S., and Itano, H. A. (1966), *Biochim. Biophys. Acta 130*, 538.

Mechanism of Bovine Prothrombin Activation by an Insoluble Preparation of Bovine Factor X_a (Thrombokinase)†

K. S. Stenn* and E. R. Blout

ABSTRACT: Sepharose-factor X_a (TK-resin) has been prepared and used for the investigation of the generation of thrombin from a purified bovine prothrombin. Thrombin generation was followed by assay for clotting activity and by disc gel electrophoresis for reaction products. Under the conditions employed (25% w/v sodium citrate, pH 6.5, 24°) prothrombin activation is essentially complete (80–100%) when compared to a two-stage activation assay. Autocatalysis could not be demonstrated either by withdrawing TK-resin before complete activation or by adding activation product to fresh prothrombin. Prothrombin in the presence of 25% sodium citrate alone, in the presence of activation products, or in the presence of thrombin does not generate clotting activity; however, prothrombin in the presence of TK-resin, with or without sodium citrate, will produce clotting activity. The complex activation pattern as observed by disc gel electrophoresis is best explained by two activation pathways: one initiated by factor X_a, and the other by thrombin. The factor X_a pathway leads from prothrombin to a single-chain molecule, P_3, plus an intermediate, F_X, which (in the presence of thrombin) rapidly decomposes. The thrombin-initiated pathway leads from prothrombin to a fragment, F_A, and an intermediate, P_2. The latter is further broken down to two single-chain molecules: a fragment F_B and an intermediate P_3. Both pathways appear to converge to the same thrombin precursor, P_3. Thrombin forms from P_3 only in the presence of factor X_a. The double-pathway mechanism was tested by reacting prothrombin (1) with TK-resin in the absence of thrombin (in the presence of $0.0079 M$ DFP) and (2) with thrombin in the absence of TK-resin. Results from these experiments and the estimated molecular weights of the intermediates are consistent with the proposed mechanism. The implications of a double pathway in physiological prothrombin activation are discussed.

Central to the phenomenon of hemostasis is the activation of prothrombin to thrombin, which catalyzes the polymerization of fibrinogen and thus forms the definitive blood clot. Despite its importance and an intensive investigative effort over many years the activation mechanism remains incompletely understood. This is due in part to the difficulty in preparation and stabilization of the zymogen and its intermediates and to the apparent complexity of the pathway. The chemistry of prothrombin and thrombin and a discussion of proposed activation mechanisms have been reviewed recently (Magnusson, 1971). As discussed by Magnusson it is generally accepted that bovine prothrombin, a single-chain plasma glycoprotein without clotting activity, of apparent molecular weight 68,000–74,000, liberates a double-chain serine protease with clotting activity, having a molecular weight of 33,700–40,000. From these facts one may assume that the activation mechanism entails (1) at least two proteolytic cleavages of the zymogen to yield the double-chain enzyme, and (2) the liberation of a large piece(s) of the prothrombin molecule.

Although it is not certain how prothrombin is activated *in vivo*, in the laboratory the conversion of prothrombin to

† From the Department of Pathology, Yale School of Medicine, New Haven, Connecticut 06510 (K. S. S.), and from the Department of Biological Chemistry, Harvard Medical School (E. R. B.), Pathology Department, Peter Bent Brigham Hospital (K. S. S.), Boston, Massachusetts 02115. *Received July 18, 1972.* Supported in part by Public Health Service RR 05358-10, National Institutes of Health Grant AM 07300, Public Health Service Grant 5T0.-HE 05274-12, by a grant-in-aid from the American Heart Association (Grant 72670) and in part during the term of U. S. Public Health Special Postdoctoral Fellowship 1 F03 HE 42488-01 (K. S. S.).