Amino-terminal Sequence Analysis of Proteins Purified on a Nanomole Scale by Gel Electrophoresis*

(Received for publication, February 11, 1972)

ALAN M. WEINER,† TERRY PLATT,‡ AND KLAUS WEBER

From the Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138

SUMMARY

A simple, rapid, manual technique is described for determining the amino-terminal amino acid sequence of proteins on a nanomole scale. In this modification of the 5-dimethylaminonaphthalene-1-sulfonyl-Edman degradation, inorganic carriers permit convenient manipulation of small amounts of protein, and use of the detergent sodium dodecyl sulfate throughout the procedure maintains protein solubility. Nanomole quantities of pure protein for such sequence analysis are readily isolated from multicomponent systems by analytical scale polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. Proteins are recovered quantitatively from the gel by elution. The method is therefore suitable for characterization of the proteins derived from multichain enzymes and viruses.

The decision to determine a protein's amino-terminal sequence depends on availability of the purified species in sufficient quantity. Techniques such as the recently automated Edman degradation (1) have been extensively developed for proteins which can be purified in milligram amounts. Since many proteins can be purified only in microgram quantities, due to limitations in starting material for the purification or lack of convenient milligram scale preparative techniques, amino-terminal sequence analysis often requires a nanomole scale procedure.

The dansyl technique, introduced by Gray (2), is sufficiently sensitive for end group determination at the 0.1 nmole level (3). A combination of this technique with stepwise degradation by the Edman procedure has been successfully used for peptide sequencing by Gray and Smith (4) and by Bruton and Hartley (5). However, microtechniques have not been developed for purifying the amino-terminal peptide derived from nanomole amounts of protein by chemical or enzymatic cleavage.

The dansyl-Edman combination has rarely been applied to intact proteins. Although dansylation has been used to identify the amino-terminal residue of proteins directly, the traditional manual Edman degradation generally has not worked well. One of the main difficulties is that, in contrast to peptides, many proteins become insoluble after one or two cycles of degradation. Since insoluble material does not react completely in subsequent cycles, unambiguous sequence determination is impossible.

We have developed a reliable manual dansyl-Edman procedure which can be used to determine an amino terminal sequence of 5 to 10 residues for any protein with an unblocked amino terminus which can be purified on a nanomole scale. We have overcome protein solubility problems by replacing the aqueous organic solvents of the traditional Edman procedure with an aqueous inorganic buffer containing the detergent sodium dodecyl sulfate. SDS, which is known to denature and to solubilize proteins in aqueous solution (6-8), has also proved useful in the dansylation reaction.

Electrophoresis on analytical polyacrylamide gels in buffers containing SDS or urea is a widely used tool to characterize the individual polypeptide chains of multichain enzymes, viruses, and membrane preparations by charge or molecular weight. The high resolving power of these gels and their simple experimental design make them very attractive as a preparative technique for the study of multicomponent systems on a nanomole scale. We have successfully applied our SDS-dansyl-Edman procedure to nanomole quantities of protein separated on and quantitatively eluted from SDS polyacrylamide gels.

Proteins of both known and unknown amino-terminal sequence have been subjected to the SDS-dansyl-Edman degradation. For proteins of known sequence, our results agree with those found by conventional techniques. For proteins of unknown sequence, such as minor capsid proteins from the coliphages Qβ and R17, our amino-terminal sequence data are compatible with known RNA sequences from the phage genome.

MATERIALS AND METHODS

Chemicals—Reagents for polyacrylamide gel electrophoresis have been described previously (9, 10). Dowex 50WX8 was obtained from Bio-Rad.

Phenylation othiocyanate (sequenation grade) was an Eastman product. Sodium dodecyl sulfate, trifluoroacetic acid (both sequential grade), and dansyl chloride (10%, w/v, in acetone) were purchased from Pierce Chemical Company (Rockford, Illinois). Guanidine hydrochloride was obtained from Schwarz/Mann. A Beckman amino acid calibration mixture (Type I) was used. Cheng-Chin polyamide layer sheets were obtained from Gallard-Schlesinger (Carle Place, New York). Paraform
is an American Can Company product (Marathon Products Division). All chemicals used were reagent grade if not otherwise specified.

Proteins—Aspartate transcarbamylase from *Escherichia coli* was purified according to the method of Gerhart and Holoubek (11). β-Galactosidase from *E. coli* was purchased from Boehringer as a crystalline suspension. Egg white lysozyme (two times crystallized) was a Schwarz/Mann product. *Lac* repressor from *E. coli* was purified by the procedure of Gilbert and Mueller-Hill (12). Histone H11 was obtained from Worthington. Subunit II of *Q*̇*B* replicase (gift of Dr. T. Blumenthal) was repurified by gel electrophoresis. Tryptophan synthetase A protein was a gift from Dr. C. Yanofsky. Bovine serum albumin (Pentex recrystallized) was obtained from Miles Laboratories. *Q*̇*B* and R17 bacteriophage were purified as described previously (13).

Equipment—A heated vacuum desiccator (Fisher Scientific) is maintained at 60° and connected to an oil pump through a Dry Ice-acetone cold trap and a trap containing potassium hydroxide pellets to keep water and acids out of the pump. A good water aspirator can be substituted for the oil pump, if drying times are increased by about 50%.

A desk top clinical centrifuge (International) is used throughout the procedure, since centrifugation at higher speed is unnecessary. The centrifuge must have swinging buckets, so that precipitates will be packed on the bottom of the tube and supernatants can be cleanly withdrawn with a Pasteur pipette. The reactions are done in either disposable culture tubes (10 mm × 75 mm) or in 15-ml conical centrifuge tubes. The disposable culture tubes (6 mm × 50 mm) which are used for danylation can be centrifuged inside a larger tube by using a paper towel insert.

Preparative Gel Electrophoresis in Sodium Dodecyl Sulfate

The 10% polyacrylamide SDS gels were prepared according to the method of Weber and Osborn (10). For 5% gels, both acrylamide and methylene bisacrylamide were halved. After polymerization, the gels were aged at room temperature for at least 4 days. They were then removed from the casting tubes by rimming and pre-eluted with stirring at room temperature into three changes of running buffer (0.1 M sodium phosphate, 0.1% SDS, pH 7.2) at 12-hour intervals.2 Gels were sucked back into the casting tubes by drawing on the opposite end of the tube with a 50-ml plastic syringe connected by a short length of a 26 gauge hypodermic needle which has been dipped in India ink (16). We have found that this procedure permits more accurate determination of mobilities than slicing through the dye band with a razor blade. The analytical gel is stained in Coomassie brilliant blue (10) for 2 to 12 hours and destained by diffusion at 37° or on a roller in a culture tube which contains destaining solution and sufficient Dowex 501-X8 to absorb unbound stain irreversibly (17). A 4-ml gel requires approximately 2 g of Dowex. Gels destained in this fashion are clean and free of the artifacts sometimes produced by electrophoretic destaining. Since both mobilities and band width increase slightly with heavy loading, analytical gels should be used only as a guide to determine an approximate mobility range for each protein band.

Determining Maximal Loading—On SDS gels, an overloaded protein band "trails" severely, although the leading edge of the protein band remains sharp with a slightly increased mobility. A single gel which is overloaded with respect to one or more proteins may nonetheless give narrow symmetrical bands for other proteins present in lower concentration. This situation often arises when intact virions are subjected to preparative electrophoresis. Minor viral components can be clearly resolved despite smearing of the major capsid proteins, as can be seen in Fig. 2. Maximal loading depends on how well the protein of interest is separated from adjacent protein bands. In the simplest case, a protein which cannot be resolved from adjacent bands on 6-cm gels may be cleanly separated on gels of 12- or 18-cm length. This is possible because electrophoresis is much faster than diffusion of the protein bands. If proteins with similar mobilities are also present in much greater quantity than the protein of interest, trailing will be more of a problem. In this case, even longer gels may not give adequate separation and loading must be reduced. In the worst possible case, a protein which cannot be purified in a single step must be subjected to a second preparative run. After preparative electrophoresis and elution of the protein bands, a small aliquot of each protein is always checked for purity by analytical polyacrylamide gel electrophoresis in SDS.

Removing Nucleic Acid—Intact viral DNA's and RNA's are excluded from 5% and 10% polyacrylamide gels. At current densities exceeding 6 ma per cm² of gel cross-section, such highly charged polyanions may cause the top of the gel to collapse. For this reason, when intact virions are disrupted for preparative gel electrophoresis by heating with SDS and 2-mercaptoethanol, the gels are run at a current density not exceeding 6 ma per cm² until the tracking dye has moved at least 1.5 cm into the gel. The nucleic acid is then flushed from the top of the gel and removed with a Pasteur pipette, taking care not to touch the gel surface. After this, the current is raised to its normal level of 25 ma per cm².

Localizing Protein Bands in Preparative Gels—We have used the following three techniques.

a. A lightly loaded analytical gel is run in parallel with each batch of preparative gels. Since protein mobilities vary slightly, depending on the precise conditions of polymerization (for a review see Reference 15), the analytical gel must be poured at the same time and in the same diameter tube as the preparative gels. After electrophoresis, the middle of the tracking dye band, as judged by eye, is marked by inserting across the gel diameter a 26 gauge hypodermic needle which has been dipped in India ink (16). When there is reason to suspect that the sample is contaminated with proteases, special precautions may be taken to avoid proteinases during sample preparation as discussed by Pringle (14).

b. After electrophoresis, a single longitudinal guidestrip is cut from the edge of the preparative gel with a modified version of the gel slicer described by Fairbanks, Levinthal, and Reeder (18). The bulk of the gel is kept at 4° while the guidestrip is stained for 15 min and destained as in a. Because guidestrips cut from 10% polyacrylamide gels may stretch as much as 5% in length, this method is less accurate than the first, but prefer-
able for preparative gels with diameter of 1.0 cm or larger. The method should not be used for 5% polyacrylamide gels, since 5% polyacrylamide gels do not stretch uniformly.

c. When adjacent protein bands are too poorly resolved for mobilities to be a reliable guide, the entire preparative gel should be frozen and sliced transversely into 1-mm sections. These sections are then eluted individually into sufficient 5 mm NaHCO₃ containing 0.05% SDS to cover the slice, and an aliquot from each elute is rerun on small analytical gels which are stained as usual to localize the protein bands. Eluates containing a single protein species are pooled, and those containing a mixture are rerun on preparative gels.

**Eluting Protein Bands**—Gel sections no thinner than 2 mm are eluted with agitation at 37° into 10 volumes (up to 5 ml) of 5 mm NaHCO₃ containing 0.05% SDS. After 12 hours, acrylamide fragments are spun out in a clinical centrifuge and the supernatant is lyophilized. The lyophilized residue can then be dissolved directly in coupling buffer as described below.

**Stepwise Sodium Dodecyl Sulfate-Dansyl-Edman Procedure**

Each cycle of SDS-dansyl-Edman degradation removes the NH₂-terminal amino acid and exposes a new amino-terminal residue, which is identified by dansylating an aliquot of the protein. A complete cycle consists of four steps, as shown in Fig. 1: coupling, precipitation, cyclization and cleavage, and finally removal of an aliquot of protein for dansylation before the cycle is repeated. Allow enough starting material so that 0.25 to 0.5 nmole in 10 μl of 10% SDS solution (w/v, in acetone) is diluted 1:20 into acetone to give a few milliliters of working solution at 5 mg per ml. Small aliquots of stock solution are withdrawn from the “Hypovial” using a hypodermic syringe. At 4° in the dark, the stock solution may be used for at least 1 year, but the working solution is replaced weekly.

**Dansylation**—The stock solution of dansyl chloride (10%, w/v, in acetone) is diluted 1:20 into acetone to give a few milliliters of working solution at 5 mg per ml. Small aliquots of stock solution are withdrawn from the “Hypovial” using a hypodermic syringe. At 4° in the dark, the stock solution may be used for at least 1 year, but the working solution is replaced weekly.

**Coupling**—The sample is dissolved in 200 μl of coupling buffer (0.5 x NaHCO₃ adjusted to pH 9.8 with NaOH) and no less than 20 μl of 10% SDS solution (w/v). For efficient detergent action, SDS must be present in at least 1.4-fold weight excess over protein (6-8). Then 10 μl of phenylisothiocyanate are added to the tube, which is immediately warmed to 50° in a water bath, flushed with nitrogen, and sealed with several layers of Parafilm. Most of the phenylisothiocyanate will remain as a separate organic phase at the bottom of the tube. The reaction proceeds for 30 min at 50°, with occasional vortexing every 5 or 10 min to keep the coupling buffer saturated with phenylisothiocyanate.

**Precipitation**—Nine volumes of acetone are added while vortexing the reaction tube, and the flocculent white precipitate is spun down for 2 min in the clinical centrifuge. The precipitate is washed with 2 ml of 100% acetone, spun down, and the colorless supernatant again discarded. The precipitate is spread around the walls of the tube by vortexing, and the tube is set nearly horizontal in the heated vacuum desiccator where the pump slowly draws acetone through the tube and the acetone is flushed from the tube. The slurry must remain on the walls of the tube when it is set in the desiccator. Rapid evaporation of acetone under reduced pressure will cause lumps to form on the bottom to explode and be blown out the mouth of the tube.

**Cyclization and Cleavage**—Tubes are allowed to cool to room temperature, and 0.2 ml of trifluoroacetic acid is added to each. If the resulting solution is not clear, the precipitate from the coupling step was not completely dry. The degradation is continued in any case, but drying times should be increased by 5 to 10 min in subsequent cycles. The tubes are warmed to 50°, flushed with nitrogen, sealed with Parafilm, and incubated for 30 min at 50°. The Parafilm is removed, a small crystal of dry SDS (about 1 mg) is added, and while still warm the tube is vortexed to dissolve the SDS. The trifluoroacetic acid is blown away in the hood under a stream of nitrogen, so that a residue forms on the walls of the tube. The tube is then set nearly horizontal in the heated vacuum desiccator and dried in vacuo for 20 min.

The residue from trifluoroacetic acid treatment is taken up immediately in 0.2 ml of coupling buffer and at least 20 μl of 10% SDS solution. The sample foams as bicarbonate neutralizes any remaining trifluoroacetic acid. Heating to 50° may be required to bring all the SDS into solution, if too large a crystal was added after cyclization and cleavage. Clumps of partially insolubilized protein are divided as finely as possible by trituration with the tip of a Pasteur pipette. The pH is checked with pH paper which has been standardized in a buffer of known pH containing SDS. Indicator dyes have higher pH values when dissolved in SDS micelles. An aliquot (0.25 to 0.5 nmole in 10 to 100 μl, including a small portion of any insolubilized material) is withdrawn for dansylation into a disposable culture tube (6 mm × 50 mm). The degradation may be continued by adding 10 μl of phenylisothiocyanate to the remaining solution, and proceeding as described under “Coupling.”

**Densylation**—The stock solution of dansyl chloride (10%, w/v, in acetone) is diluted 1:20 into acetone to give a few milliliters of working solution at 5 mg per ml. Small aliquots of stock solution are withdrawn from the “Hypovial” using a hypodermic syringe. At 4° in the dark, the stock solution may be used for at least 1 year, but the working solution is replaced weekly.

One-half volume of the working dansyl chloride solution is added to the protein aliquot withdrawn for dansylation, making it 33% in acetone. A precipitate may form, which will not interfere with dansylation. The tube is covered with Parafilm and incubated at 37°. Prewarming the tube is unnecessary and may even cause evaporation of the acetone, which is needed to solubilize the dansyl chloride. The reaction is terminated after 20 min by acid precipitation, following the suggestion of Gros and Labouesse (19). At protein concentrations greater than 50...
µg per ml (for example, 5 µg/100 µl), the dansylated protein can be precipitated directly by cautious addition of ice-cold 20\% trichloroacetic acid. The trichloroacetic acid solution must be added dropwise and with constant vortexing, or evolving carbon dioxide will cause uncontrolled foaming in the detergent solution. Below 50 µg per ml, 5 µg of succinylated carrier protein\(^1\) are added before trichloroacetic acid precipitation to assure quantitative recovery.

In either case, the precipitate is spun down and the supernatant withdrawn with the finely drawn tip of a Pasteur pipette\(^4\) and discarded. The precipitate is washed once with 0.2 ml of 1 N HCl to remove dansylic acid. After centrifugation the supernatant is withdrawn, and 50 µl of 6 N HCl (reagent grade concentrated HCl diluted 1:1 with distilled water) are added. A neck is drawn on the tube using a fine oxygen torch, and the tube is sealed under water aspirator vacuum (oil pump unnecessary). After hydrolysis for 4 to 6 hours at 105°, as recommended by Gros and Labouesse (19), the tubes are opened and the hydrolysates dried down under vacuum in the heated vacuum desicator. \(\text{NH}_2\)-terminal isoleucine, leucine, and valine occasionally require longer hydrolysis, as discussed under "Chromatography."

Chromatography—Chromatography is performed essentially according to the method of Hartley (3) as modified by Neuhoff (cited in Reference 3). Polyamide plates coated on both sides of the plate, registry of the spots should be sufficiently accurate to distinguish dansyl-isoleucine from dansyl-leucine after hydrolysis residue present on the white polyamide sheet.

\(^1\) Succinylated carrier protein was prepared according to the method of Habeeb et al. (20). Protein was dissolved at 10 mg per ml in guanidine hydrochloride buffer with 1 M NaClO\(_4\) and a 36-fold molar excess of solid succinic anhydride over amino groups was added in three equal aliquots at 1-hour intervals.

\(^4\) Capillary Micropipettes—Kimax melting point capillaries and a microburner without flame-spreader are used. The center of the capillary is rotated just above the flame, taking care to heat as short a length of tubing as possible. When the glass has softened, the capillary is removed from the flame and immediately drawn out. If the capillary is drawn in the flame, it will melt in two. With the capillary resting on a clean towel, a single scratch is made at the midpoint with a diamond pencil. When the capillary is bent slowly into a "U" with the scratch on the convex side, the resulting break is even. Finely drawn Pasteur pipettes are made similarly.

by judged at the time of application by the amount of brown hydrolysis residue present on the white polyamide sheet.

In general, the smallest sample is applied which permits unambiguous identification of the unknown dansyl-amino acid. A larger sample must be applied when chromatography reveals no dansyl derivatives other than \(\epsilon\)-dansyl-lysine. A smaller sample may be advantageous by reducing background when multiple spots are observed. Incomplete hydrolysis of an aliphatic \(\text{NH}_2\)-terminal dansyl derivative may result in two bright spots without other significant background (2, 19). In this case the sample should be dried down, taken up in 6 N HCl, and rehydrolyzed for an additional 8 hours.

A standard mixture containing about 0.02 nmole of each dansyl-amino acid (see below for preparation of this mixture) is spotted on the back of the plate opposite the unknown. The plate is held in a 250-ml beaker under a hot air stream for at least 5 min. Incomplete evaporation of the pyridine will produce trailing in the first dimension, due to local neutralization of the 1.5% formic acid.

Ascending chromatography in four successive solvents is performed according to the method of Woods and Wang (21) as modified by Hartley (6). Solvent II, III, and IV are run in the same direction perpendicular to Solvent I.

Solvent I 1.5\% formic acid in water
Solvent II benzene-acetic acid, 9:1
Solvent III ethyl acetate-acetic acid-methanol, 20:1:1
Solvent IV 0.05 M NaPO\(_4\), in 25% aqueous ethanol

The four solvents can be made in advance and stored in tightly capped bottles at room temperature. Solvents should be kept covered at all times and never used for longer than 3 hours, since differential evaporation affects their chromagographic properties. Solvents should be discarded whenever they cease to resolve the standard dansyl-amino acid mixture adequately.

Altogether, chromatography requires less than an hour: 5 min each for Solvents I, II, and III; 10 min in Solvent IV; at least 10 min of drying after Dimension I and 5 min each after Dimensions II and III. Dimension IV requires only brief drying before examination under ultraviolet light. Solvents II, III, and IV need a solvent-saturated atmosphere, and can be run in 150-ml beakers covered with a stretched sheet of Parafilm. The solvent should just cover the bottom of the beaker evenly. Plates are placed in beakers and withdrawn using tongs. The resulting abrasion marks are not harmful, however, since the solvent front is never allowed to run to the edge. Plates are kept as vertical as possible during chromatography. The lower edges of the plate must not touch the walls of the beaker or capillarity may distort the solvent front.

Chromatography is not difficult but the manipulations require practice. Before several trial cycles of SDS-dansyl-Edman are attempted on a protein of known sequence, reliable chromatographic identification of the dansyl derivative must be possible.

When the standard dansyl-amino acid mixture is applied to both sides of the plate, registry of the spots should be sufficiently accurate to distinguish dansyl-isoleucine from dansyl-leucine after Dimension II; dansyl-alanine from dansyl-amine, dansyl-threonine from dansyl-serine, dansyl-glutamic acid from dansyl aspartic acid, \(\alpha\)-dansyl-histidine\(^5\) from \(\epsilon\)-dansyl-lysine after

\(^5\) We have found that \(\alpha\)-dansyl-histidine cannot be resolved from \(\epsilon\)-dansyl-lysine in Dimension III unless (a) the polyamide thin layers are presoaked for at least 10 min in Solvent I and (b)
This problem does not arise when polyamide thin layers are reused after regeneration according to the method of Wang and Grey (22).

The dansylation reaction is terminated by adding excess formic acid. In this way the bicarbonate buffer is not applied to the polyamide sheet, where it may cause smearing in the first dimension by neutralizing the 1.5% formic acid solvent.

A 100-µl aliquot of an amino acid standard mix (Beckman Instruments amino acid calibration mixture, 2.5 mg in each amino acid in 0.01 N HCl) is dried down in the heated vacuum desiccator, resuspended in 0.5 ml of 0.1 N HCl, and dansylated by adding 0.25 ml of fresh working dansyl chloride solution at 5 mg per ml in acetone. The tube is covered with Parafilm and incubated at 37°C for 30 min. Reaction is terminated by adding 25 µl of 88% formic acid. The dansylated standard mixture is used directly for spotting (0.1 µl contains 0.025 nmole of each amino acid) and stored in the cold. This dansyl standard mixture contains didansyl-histidine, which is converted to α-dansyl-histidine by acid hydrolysis. To obtain α-dansyl-histidine, an aliquot of the standard mix may be dried down and hydrolyzed, or alternatively a stock solution of histidine can be dansylated with a 2-fold excess of amino acid over dansyl chloride as described by Gray (2).

RESULTS

Sequence Results—Amino-terminal sequence results obtained with the SDS-dansyl-Edman microtechnique are presented in Table I. Proteins ranging in molecular weight from 14,000 (Qβ coat protein) to 135,000 (β-galactosidase) have been studied. Those purified by gel electrophoresis in the presence of SDS are indicated in the table with a superscript c. No more than 0.25 nmole of protein per step was needed for amino-termin-
nal sequence analysis, although more was used in several instances when sufficient starting material was available.

As an illustration of the SDS-dansyl-Edman procedure as applied to proteins purified on a nanomole scale by gel electrophoresis, we describe here a model experiment which yielded the amino-terminal sequences of the two minor capsid proteins of bacteriophage Qβ. The major coat protein species accounts for 95% of the capsid protein, and the two minor protein species IIb and IIa for 4% and less than 1%, respectively (31). In the model experiment, 6 mg of purified Qβ virions were disrupted for 2 min at 95°C with a 2-fold weight excess of SDS and 2% 2-mercaptoethanol. After reduction for 4 hours at 37°C, equal aliquots of the denatured virus were layered onto six "pre-eluted" 10% polyacrylamide SDS gels (0.8 cm × 12 cm). About 5% of the sample was run in parallel on a fifth gel for use as a guide during elution. After electrophoresis at 3 mA each for 1 hour, viral RNA which had been excluded from the gel was removed with a Pasteur pipette. The current could then be raised to 8 mA each, and the gels run for 16 hours. The preparative gels were frozen at -20°C while the guide gel was stained and destained as usual. Using mobilities determined from the guide gel, the IIa and IIb bands were cut out separately from each preparative gel and eluted into 4 ml of 0.05% SDS buffered with 5 mM NH₂HCO₃ for 12 hours at 37°C with agitation. Elution under these conditions permits quantitative recovery of the minor plaque proteins after gel electrophoresis (see Table II). To determine purity, 5% of each elute was rerun on analytical SDS polyacrylamide gels. Fig. 2 demonstrates that after elution, each minor plaque protein is completely free of the other and of the major coat protein species. The eluates were then lyophilized, redissolved in coupling buffer, and subjected to three cycles of SDS-dansyl-Edman degradation. Using 40 µg of IIa protein we obtained the amino-terminal sequence Pro-Lys-Leu-Pro, and using 100 µg of IIb protein the sequence Ala-Lys-Leu-Glu-Thr-Val-Thr-Leu (Table I). The biological implications of these sequence data are considered in the "Discussion."

Preparative Gel Electrophoresis—Proteins for SDS-dansyl-Edman sequence analysis can be quantitatively recovered from "pre-eluted" gels. In the model experiment described in Table II, greater than 99% recovery was obtained for Qβ coat protein (molecular weight 14,000) run on 10% polyacrylamide SDS gels and for β-galactosidase (molecular weight 135,000) run on 5% gels. These gels combine experimental convenience, extraordinary sensitivity (less than 0.5 µg of protein is easily visible after staining), and high resolution, although they separate polypeptide chains almost entirely on the basis of molecular weight (10, 33, 34).

When we attempted to apply our SDS-dansyl-Edman technique to microgram quantities of protein eluted from such gels, the first results were disappointing. Gel electrophoresis and subsequent elution of the protein reduced the sensitivity of the sequencing technique by 5-fold, so that 2.5 nmoles of protein instead of 0.25 n mole were required per step. Since we had expected quantitative recovery of protein, we attributed the decrease in sensitivity to either (a) chemical modification of the NH₂-terminus of the protein during electrophoresis, or (b) direct interference with the SDS-dansyl-Edman degradation by material eluted together with the protein from the polyacrylamide. Pre-electrophoresis with running buffer or with methyl green (34) to eliminate persulfate artifacts (for review see Reference 15) did not increase sensitivity, nor did pre-running the gels...
with a low molecular weight protein (9). Because the interfering material seemed to be electrophoretically neutral, we resorted to diffusion as suggested by Weber and Kuter (9). In a model experiment, 100 µg of lac repressor were recovered from such a pre-eluted gel (see "Preparative Gel Electrophoresis" under "Materials and Methods"), and four steps of SDS-dansyl-Edman degradation were successfully completed by withdrawing 10-µg (0.25 nmole) aliquots at each step for dansylation. We do not know whether the interfering substance is residual acrylamide monomer, polyacrylamide, or polyacrylic acid, a contaminant in one of the reagents, or a by-product of the polymerization reaction (for a review see Reference 15); however, it appears to be removed by diffusion.

### Stepwise Sodium Dodecyl Sulfate-Dansyl-Edman Degradation—

A complete cycle of SDS-dansyl-Edman degradation consists of four steps as shown in Fig. 1.

**Coupling Reaction**—In this step the protein is coupled with the water-insoluble Edman reagent, phenylisothiocyanate, to form a phenylthiocarbamyl derivative. The manual Edman degradation has usually been restricted to peptides, since many proteins are insoluble in aqueous organic solvents such as 50% pyridine which have been used to dissolve the Edman reagent. Other proteins become insoluble after one or two cycles of degradation.

The detergent SDS is known to solubilize and denature proteins in aqueous solution (6-8). Preliminary attempts to avoid solubility problems by adding SDS directly to the aqueous organic solvents of the traditional manual Edman procedure were unsuccessful. Model experiments revealed that the detergent action of SDS was greatly reduced in 50% pyridine. However, above its critical micellar concentration (7) SDS has been shown to bring not only proteins but also water-insoluble organic dyes into aqueous solution (for a review see Reference 6). We therefore used SDS in an aqueous inorganic coupling buffer to solubilize both the protein and the water-insoluble Edman reagent. When excess phenylisothiocyanate is present and forms a two-phase system, SDS mixtures presumably transport molecules of phenylisothiocyanate from the denser organic phase to SDS-protein complexes in the aqueous phase.

**Precipitation**—It is necessary to remove excess phenylisothiocyanate and nonvolatile by-products of the coupling reaction (4) before cyclization of the PTC-protein in anhydrous acid (1). Incomplete removal can lead to elevated backgrounds during identification of the dansylated NH₄-terminal amino acid SDS in the aqueous coupling buffer precludes repeated extractions with an immiscible organic solvent to remove phenylisothiocyanate and its by-products, since (a) the detergent action of amphiphilic SDS prevents formation of a sharp interface, and (b) PTC-protein might partition into the organic phase in the presence of high concentrations of detergent. We therefore chose precipitation with nine parts acetone since SDS, phenylisothiocyanate, and the by-products of the coupling reaction are all soluble in aqueous acetone.

When a solution of protein in aqueous SDS is buffered with an organic compound, the protein precipitates from 90% acetone as a gummy residue while the buffer remains soluble. In model experiments with lysozyme and R17 coat protein, the PTC-protein was precipitated with nine parts acetone from aqueous SDS buffered with dimethylallylamine-trifluoroacetic acid (35). The precipitates were thoroughly washed with acetone and dried under vacuum at 60°C. In neither case would the dry residue dissolve completely in anhydrous trifluoroacetic acid, even after prolonged vortexing at 50°C. Preliminary performic acid oxidation of the proteins (36) did not affect these results.

Use of a strong 0.5 M sodium bicarbonate-carbonate buffer in the coupling reaction solves this problem, since the inorganic salt NaHCO₃ gives a flocculent precipitate in aqueous acetone. When PTC-protein and NaHCO₃ co-precipitate from 90% acetone, the flocculent inorganic precipitate acts as a carrier to keep the insoluble PTC-residue finely dispersed. This coprecipitate, when thoroughly washed with acetone and dried briefly in vacuo at 60°C, dissolves immediately in anhydrous trifluoroacetic acid. Moreover, the acetone precipitation extracts sodium trifluoroacetate formed in the previous cycle of degradation by neutralization of NaHCO₃ and Na₂CO₃. Since sodium trifluoroacetate is readily soluble to at least 1 M in trifluoroacetic acid and to at least 4 M in 90% acetone, there is no accumulation of salt during the degradation.

Co-precipitation from sodium bicarbonate-carbonate buffer has a second, unexpected advantage. In model experiments with R17 coat protein and histone H1, we found that 90% acetone quantitatively precipitates the PTC-derivatives of these two proteins from coupling buffer at final concentrations as low as 10 µg of protein per ml. Histidine H1 contains 27% lysine on a molar basis (37), in contrast to 5% in R17 coat protein (27). Since the e-amino group of lysine reacts with phenylisothiocyanate to form a PTC-derivative, thereby converting a positively charged polar residue to an uncharged hydrophobic derivative, the PTC-histone would be expected to have greater solubility than most PTC-proteins in 90% acetone. Quantitative recovery of both histone H1 and R17 coat protein therefore suggests that the precipitation procedure is perfectly general.

**Cyclization and Cleavage—**Incubation with trifluoroacetic acid catalyzes cyclization and cleavage of the PTC-protein to produce the 2-anilino-5-thiazolinone derivative of the NH₄-terminal amino acid and to expose a new NH₄ terminus. Edman (for re-

* The presence of sodium carbonate in the coupling buffer slightly complicates this picture. Model experiments show that solutions of Na₂CO₃ at concentrations of 0.1 to 1.0 M separate into two phases when made 90% in acetone. The viscous lower phase, with less than 5% of the total volume, contains greater than 95% of the original sodium carbonate, as judged by constant weight after extensive drying under vacuum at 60°C. When 0.5 M coupling buffer at pH 9.8 is made 90% in acetone, the two salts seem to be independent; a flocculent precipitate characteristic of sodium bicarbonate can be spun out, while a second phase characteristic of sodium carbonate coats the test tube wall with small oily beads. The sodium carbonate does not interfere with the ability of sodium bicarbonate to act as an effective protein carrier.

* R17 coat protein and histone H1 were dissolved at 1.0 mg per ml in coupling buffer containing 1% SDS. A 20 µl aliquot of each protein solution was diluted into 0.2 ml of the same buffer, coupled with phenylisothiocyanate, precipitated with acetone, and dried as described under "Materials and Methods." The dry acetone precipitates were taken up in 4 ml of 6 N HCl with a crystal of phenol, sealed under oil pump vacuum, and hydrolyzed for 24 hours at 105°C. Since the presence of inorganic salts during acid hydrolysis can lead to reduced recovery of certain amino acids, the following controls were included. Tubes containing coupling buffer and SDS but not protein were carried in parallel through the coupling, precipitation, and drying steps. The protein-free precipitates of NaHCO₃ and Na₂CO₃ were taken up in 6 N HCl with phenol, and 20 µl of the appropriate protein solution were added before hydrolysis. The efficiency of acetone precipitation, as judged by recovery of alanine and aspartic acid, was normalized to these controls.
after cyclization and cleavage in trifluoroacetic acid is identified for tryptophan synthetase A protein in agreement with the published sequence Met-Gln-Arg-Tyr of Guest et al. (23).

A model experiment, \(^{19}\) incubation of Q3 coat protein for 15 min at 50° in trifluoroacetic acid containing 2.5% water revealed no significant hydrolysis of internal peptide bonds.

As in the traditional Edman degradation, after removal of the trifluoroacetic acid either by evaporation or lyophilization, the protein residue will not redissolve completely in coupling buffer even in the presence of SDS. We solved this problem, however, by adding a small crystal of dry SDS to the trifluoroacetic acid solution just before evaporation. The resulting residue dissolves easily when heated to 50° in coupling buffer. Presumably, a detergent-protein complex has already formed in the lyophilized material and need only be rehydrated.

The residue from trifluoroacetic acid treatment should always be redissolved as quickly as possible in coupling buffer. Reducing the cumulative exposure to strong acid in this way will (a) reduce chromatographic backgrounds due to nonspecific acid hydrolysis of the polypeptide chain, and (b) minimize the danger of blocking the SDS-dansyl-Edman degradation by acid-activated cyclization of a newly exposed NH\textsubscript{2}-terminal glutamine to pyrrolidone carboxylic acid (41). Internal glutamine residues which have become NH\textsubscript{2}-terminal in the course of degradation do not cyclize to an appreciable extent under our conditions. In a model experiment using 0.3 nmole of protein per step, we obtained an amino-terminal sequence Met-Glx-Arg-Tyr (Table I) for trichloroacetic acid residues from the Edman degradation. Using 0.2 ml volumes of coupling buffer throughout, 1-μg aliquots (0.07 nmole) of protein were withdrawn for dansylation at each step in volumes of 40 to 100 μl. The first three

19 A 5-μl aliquot of Q3 bacteriophage at 17 mg per ml in buffer (0.1 M Tris-HCl-10 mM MgCl\textsubscript{2}-0.1 M NaCl, pH 7.2) was added to 0.2 ml of anhydrous trifluoroacetic acid, flushed with nitrogen, and incubated at 50°. After 15 min a small crystal (about 1 mg) of dry SDS was added to the solution and the sample was dried down, taken up in coupling buffer, dansylated, hydrolyzed, and chromatographed as usual. In the control experiment, 5 μl of the Q3 solution were extensively dried in the heated vacuum desicicator before addition of the trifluoroacetic acid. In both the model experiment containing 2.5% water and in the anhydrous control, incubation for 15 min at 50° revealed less than 1% NH\textsubscript{2}-terminal proline relative to NH\textsubscript{2}-terminal alanine as judged by visual densitometric analysis. A 10-fold dilution of the sample into 15% pyridine yields a dansyl-alanine spot of intensity comparable to the dansyl-proline contaminant in the undiluted sample. The NH\textsubscript{2}-terminal proline can be attributed entirely to the Q3 maturation protein (see "Sequence Results") which constitutes less than 1% of the total capsid protein. This confirms earlier work by Smyth, Stark, and Koningsberg (cited in Reference 30) in which new end groups could be detected by the cyanate technique (40) after incubation for 24 hours at 25° in moist trifluoroacetic acid.

11 In a model experiment, 5 μg (0.35 nmole) of hen egg white lysozyme were subjected to four cycles of SDS-dansyl-Edman degradation. Using 0.2 ml volumes of coupling buffer throughout, 1-μg aliquots (0.07 nmole) of protein were withdrawn for dansylation at each step in volumes of 40 to 100 μl. The first three
proteins are recovered quantitatively from the gel by elution is on polyacrylamide gels in the presence of SDS. Under favorable circumstances, proteins from membrane preparations also remain very near the origin in Dimensions I, II, and III (Fig. 3).

The amino-terminal sequence of 5 to 10 residues can be obtained for any protein with an unblocked amino terminus which is available on a nanomole scale. Nanomole quantities of protein for sequence data for the corresponding proteins can be used in conjunction with ribonucleotide sequence data to identify which of many potential initiation sites for protein synthesis actually function in vivo. The SDS-dansyl-Edman microtechnique must be used when the protein of interest is available only in nanomole quantities. For instance, in the sequence of Billeter et al. (25) there are three AUG and two GUG codons within the first 175 nucleotides from the 5'-end of bacteriophage Q8 RNA, any one of which could potentially function as an initiation site (for review see Reference 46). The amino-terminal sequence of the Q8 maturation protein obtained in the model experiment described under "Sequence Results" and presented in Table 1, enabled us to identify one of these codons (the AUG codon beginning at position 62) as the true initiation site and incidentally to place the maturation protein cistron at the 5'-end of the RNA genome (31).

RNA sequences coding for the initiation of protein biosynthesis can also be identified as ribosome binding sites, which are oligonucleotide tracts of messenger RNA protected from mild RNAase digestion by the initiation complex for protein synthesis. In this complex the ribosome is specifically bound to an mRNA molecule in the presence of initiation factors, fMet-tRNA, and GTP (46). Currently, 6 ribosome binding sites have been sequenced, all from the family of single-stranded RNA coliphages such as Qβ (29) and R17 (24). In general, ribosome binding sites can be assigned unambiguously to specific phage cistrons only by using the genetic code to correlate the RNA sequence with the corresponding amino-terminal sequence of the cistron's protein product. For instance, the amino-terminal sequence of the phage-specific subunit II of Qβ replicase, which is presented in Table I, enabled Hindley and Staples (29) to assign their ribosome binding sequence to the Qβ replicase cistron. In a similar case, the amino-terminal sequence of the R11 maturation protein (Table 1) not only confirmed the identity of the maturation cistron initiation site (24) but also extended the RNA sequence at the initiation site.

b. Amino-terminal sequence analysis can demonstrate relationships between viral proteins. For instance, sequence data obtained in the same model experiment described under "Sequence Results" and presented in Table 1 reveal that the major coat protein (molecular weight 14,000) of bacteriophage Qβ and the minor capsid component Ib (molecular weight 30,000) have identical amino-terminal sequences for at least 8 residues. This extraordinary fact led us to postulate that the Ib protein is a read-through product resulting from polypeptide chain elongation past an inefficient UGA stop signal at the natural termination site of the Qβ coat protein cistron. Growth of wild type Q8 on a high level UGA suppressor confirmed this hypothesis by
increasing the molar fraction of 1Hb protein in purified phase from 2 to 7% (31).

Precursor-product relationships between distinct capsid proteins have already been established by other techniques for bacteriophage T4 (47) and for poivivirus (48). In the future we can expect that characterization of viral proteins by amino-terminal sequence will yield similar results.

c. SDS-dansyl-Edman sequence analysis can also demonstrate relationships between mutant and wild type proteins. Work in progress on E. coli lac repressor has shown that protein biosynthesis can reinitiate at an internal methionine codon after polypeptide chain termination at an early amber block (43). The reinitiated fragment of lac repressor was first isolated from a crude cellular extract by precipitation with antibody prepared against wild type repressor (49),13 and then separated from the antibody polypeptide chains by electrophoresis on 10% polyacrylamide SDS gels. Because the cross-reacting repressor fragment constitutes only 0.02% of a crude cellular extract, antibody precipitation yielded very small amounts of material which could be sequenced only by the SDS-dansyl-Edman microtechnique. The results are presented in Table 1. When this sequence was compared with that of wild type lac repressor which had been purified and partially sequenced by conventional means,6 it was found that residues 49 through 46 of the wild type sequence correspond to the amino-terminal sequence of the reinitiated fragment. This indicates that an AUG triplet which codes for the internal residue methionine 42 in wild type lac repressor can function as an initiation site for protein synthesis when activated by an early amber mutation in the gene. Stewart et al. (50) have drawn similar conclusions in their work with a eukaryotic yeast, using conventional techniques for protein purification, peptide mapping, and peptide sequencing.

Nanomole quantities of protein eluted from SDS polyacrylamide gels can be hydrolyzed for amino acid analysis (51). It may also be possible to subject these proteins to carboxy-terminal analysis, since both carboxypeptidase A and B are active in SDS (52). The free amino acids released by carboxypeptidase digestion can be identified by dansylation.14 Moreover, these same proteins can also be characterized by enzymatic (9, 15) and immunological activity (53),14 since Weber and Kuter (9) have recently shown that many proteins can be renatured from SDS with at least partial recovery of activity after removing the detergent.

Acknowledgments—We thank Dr. H. G. Khornann for use of his Packard sample oxidizer, and Miss C. Hering for cheerful, expert technical assistance.

REFERENCES

42. Edman, P. (1970) in S. N. Deedeeble (Editor), Protein sequence determination, Springer-Verlag, New York

14 A. M. Weiner, unpublished experiments.
Amino-terminal Sequence Analysis of Proteins Purified on a Nanomole Scale by Gel Electrophoresis

Alan M. Weiner, Terry Platt and Klaus Weber


Access the most updated version of this article at [http://www.jbc.org/content/247/10/3242](http://www.jbc.org/content/247/10/3242)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/247/10/3242.full.html#ref-list-1](http://www.jbc.org/content/247/10/3242.full.html#ref-list-1)