A rapid solid-phase protein microsequencer

John E. WALKER,* Ian M. FEARNLEY and Roger A. BLOWS

Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

A solid-phase protein microsequencer is described that has been designed to determine protein sequences with subnanomolar quantities of protein. Its utility has been demonstrated by the determination of many sequences in subunits of mitochondrial F_1 -ATPase, in a protein isolated from mouse gap junctions and in the mitochondrial phosphate-transporter protein. It has a number of advantages over liquid- and gas-phase sequencers. Firstly, the degradation cycle takes 24 min, more than twice as fast as any other sequencer. This helps to reduce exposure of proteins to inimical reagents and increases throughput of samples. Secondly, polar amino acids such as phosphoserine, and polar derivatives formed by active-site photoaffinity labelling with 8-azido-ATP, are recovered quantitatively from the reaction column and can be positively identified. In other types of sequencer these polar derivatives, being somewhat insoluble in butyl chloride, tend to remain in the reaction chamber of the instrument and so are more difficult to identify. The solid-phase protein sequencer is also more suited than the liquid-phase instrument for analysis of proteolipids from membranes. These hydrophobic proteins tend to dissolve in organic solvents during washing steps in the liquid-phase instrument and are lost. Covalent attachment as used in the solid-phase instrument solves this problem.

INTRODUCTION

Most protein sequences determined by sequence analysis of proteins (rather than DNA) depend upon the Edman degradation, whereby amino acids are removed sequentially from the *N*-terminal end of polypeptide chains [for reviews, see Edman & Henschen (1975) and Allen (1981)]. This is achieved by repetition of a cycle of reactions containing two steps: *coupling* of the *N*-terminus under basic conditions with phenyl isothiocyanate, and *cleavage* of the phenylthiocarbamyl derivative of the *N*-terminal residue by a cyclization reaction in acid. In a third reaction, *conversion*, the cleavage product, the anilinothiazolinone, is rearranged to a more stable derivative, the phenylthiohydantoin (PTH). This derivative is then identified, today usually by h.p.l.c.

Manual procedures for performing these reactions (see Allen, 1981) have been largely superseded by automated chemistry performed in protein sequencers. These are of three types; the liquid-phase (or spinning-cup) sequencer (Edman & Begg, 1967; Hunkapiller & Hood, 1978, 1980), the solid-phase sequencer (Laursen, 1971; Laursen *et al.*, 1975) and the gas-phase sequencer (Hewick *et al.*, 1981; Hunkapiller *et al.*, 1983*a,b*). The solid-phase approach to sequence analysis of proteins was introduced by Laursen (1971). The essence of the procedure is that proteins or peptides are covalently attached to an insoluble support. This is then placed in a glass column and the reagents for Edman degradation are introduced sequentially by an arrangement of pumps and valves as instructed by an electro-mechanical or electronic control device. A major advantage of this approach is that because the protein is covalently attached, by-products and excess reagents can be removed by solvent washes without concomitant risk of loss of sample. (In the liquid-phase sequencer the

protein is degraded in a liquid film and in the gas-phase sequencer it is adsorbed on to a glass-fibre disc.)

In the present paper we describe a solid-phase sequencer that has been designed with the aim of degrading subnanomolar quantities of proteins and peptides. With this instrument we have been able to successfully degrade proteins by using as little as 60 pmol of protein. Routinely we have employed 100 pmol of protein usually obtaining sequence up to at least residue 20–30. This compares favourably with other types of high-sensitivity protein sequencers. Moreover, in the course of this work we have demonstrated clear advantages of this instrument over others in terms of rapidity of degradation, and in the degradation of peptides containing very polar amino acid residues, such as phosphorylated residues or amino acids that have been derivatized by a photoactivatable derivative of ATP, 8-azido-ATP.

CONSTRUCTION OF THE SEQUENCER

(a) General design

In setting out to design an instrument capable of degrading microquantities of protein, a number of considerations were thought to be important. Firstly, in order to minimize unwanted side reactions the reagents themselves should be highly purified and maintained in an inert (oxygen-free) environment to ensure their long-term stability. For this reason the use of Teflon tubing, which is porous to oxygen, has been minimized and flow paths of reagents from reservoirs to reaction column drastically decreased. A second way in which side reactions and the effect of contaminants can be decreased is to minimize the exposure of the protein to all reagents

Abbreviations used: ATZ, anilinothiazolinone; PTH, phenylthiohydantoin; PITC, phenylisothiocyanate; TFA, trifluoroacetic acid; DITC, *p*-phenylene di-isothiocyanate.

* To whom correspondence and reprint requests should be sent.

Fig. 1. Flow paths in the solid-phase protein microsequencer

Heavy and thin lines indicate liquid and argon flow paths respectively. P1–P3 are syringe pumps for delivery of PITC, buffer and TFA respectively. V1–V14 are Altex slider valves for delivery of reagents and argon. They are pneumatically activated by activation of electromagnetic valves (not shown). For further details of reservoirs column and autoconverter, see Figs. 2 and 3 respectively.

by reduction of both the volumes of reagents that pass over the protein and the duration of exposure of the protein to chemicals. The desire to use highly purified chemicals has an additional corollary: since the purification procedures are labour-intensive, it is clearly advantageous to employ the minimum number of different chemical reagents in the degradation and conversion processes. Considerable attention has also been given to the design of the reaction column. It too has been dramatically reduced in physical size so that it accommodates only the glass beads to which protein has been covalently attached and does not require 'inert' underivatized beads to fill up dead spaces. This reduction of the size of reaction column has itself required the development of precisely controlled delivery systems for reagents.

The instrument has much in common with that described by Laursen (1971) and is shown in Fig. 1. It differs in many details dictated by the general considerations outlined above. The coupling and cleavage reactions of the Edman degradation take place in the column that contains the immobilized protein. The conversion of the released anilinothiazolinone amino acids to the corresponding phenylthiohydantoin is performed in the glass vial used to collect the cleavage

products (see also subsection g). The three reagents, phenylisothiocyanate, the coupling buffer and trifluoroacetic acid, are displaced into the reaction column by syringe pumps, P1–P3, via pneumatically activated slider valves (V1–V3). Argon can be introduced into the column via the sixth valve, V6. The solvents, methanol and benzene, which are used to wash away reaction by-products and excess reagents, are displaced through valves V4 and V5 by argon-gas pressure in the reservoirs. A second set of pneumatically operated slider valves, V9–V11, V13 and V14, deliver argon to the solvent reservoirs to ensure that they are kept under a positive inert-gas pressure and to displace the methanol or benzene when V4 or V5 is opened. The column effluent is directed by valve V7 either to waste (solvent washes), or into the autoconverter (trifluoroacetic acid). This valve also is used in the inactivated position to deliver argon to assist evaporation of the sample during autoconversion. The autoconverter heater is incorporated into the turntable where the products of the Edman degradation are collected.

The operation of valves, pumps, column heater and autoconverter are controlled by a DEC PDP-11 minicomputer. The operation of the instrument can be monitored on a visual-display unit and information to

Fig. 2. Reagent reservoirs

(a) Plan view of the cap. (b) Side view of reservoir with cap in place. The outer part of the caps (A), shown unscrewed in the lower part, are made from Delrin or, for the TFA vessel, from polypropylene. The inert plug (D) is polytetrafluoroethylene and is screwed to a Dural component (C). Through this the three inlet and outlet lines are sealed with O-ring seals screwed down with Altex fittings (B).

start or stop the operation of the instrument can be inputted from a keyboard. Commands specifying the operation of the instrument are stored on mini-tape cassettes and can be entered into memory via the cassette reader. The memory can be loaded and the instrument be operational within 5 s.

(b) Reagent reservoirs

The glass bottles (250 ml capacity for methanol, benzene and buffer, 100 ml for phenylisothiocyanate and trifluoroacetic acid) are brown Duran bottles (purchased from Schott, Postfach 2480, D-6500 Mainz 1, Germany). The rims were chamfered to fit the PTFE (polytetrafluoroethylene) insert (D in Fig. 2a) of the cap by Cambridge Glassblowing Ltd., Cottenham, Cambridge CB4 4PS, U.K. The caps (see Figs. 2a and 2b) were made in the engineering workshop of this Laboratory. The outer part (A in Fig. 2a) is either made from Delrin (du Pont), or in the case of the trifluoroacetic acid container, from polypropylene. O-Ring seals (British Standard 003 B.S.) are made of Teflon-coated Viton elastomer (British Seals and Rubber Mouldings, Romford, RM7 0AP, Essex, U.K.). Reagents were kept under an argon pressure of 20.7 kPa (3 lbf/in²). Outlet tubes can vent to an outlet manifold during purging with argon after re-filling operations. Otherwise these lines are closed by taps. The argon inlet to each bottle has its own pressure regulator. A sapphire check valve (part TCKA 6201 070A from Lee Products, P.O. Box 0424, Westbrook, CT 06498, U.S.A.) in the argon-inlet line to the trifluoroacetic acid bottle prevents back-diffusion of trifluoroacetic acid vapour into the pressure regulator.

(c) Reagent valves

The arrangement of reagent valves is similar to that described by Laursen (1971), but some changes have been made. Firstly, an additional valve, V6, has been introduced to allow argon gas to be blown through the column to displace trifluoroacetic acid after the cleavage reaction; secondly, modification of the slider in V7 permits argon to be blown into the glass vial during the conversion reaction (see subsection g).

(d) Pumps

Braun syringe pumps (designed for intravenous perfusion; Braun Melsungen AG, Melsungen, Germany) are used for the delivery of buffer, phenylisothiocyanate and trifluoroacetic acid. These have been modified so that they are driven by an Impex stepper motor (maximum torque 2.2 N·cm). The speed of the motor, and hence the rate of delivery of the reagent, is controlled by a Siemens SAA1027 driver chip using data from the LSI-11 computer (a Digital Equipment Corporation 16-bit computer). These data are latched or retained in the interface memory while the computer performs other functions. Pumping speeds in the range 50 μ l–1.5 ml/min can be selected when a 5 ml syringe is employed. In current programs, reagents are delivered at 160 μ l/min. At the end of a delivery, the delivery valve is closed and simultaneously the pump reversed, thereby refilling the syringe from its reservoir. Its backward travel is stopped by closure of a micro-switch placed toward the rear of the pump casing. It is closed by an arm attached to the syringe mounting. Then the syringe is driven forward for 15 s, the contents of the syringe being displaced towards the reservoir. This eliminates mechanical backlash from the pump and ensures that the next delivery from this pump commences immediately upon activation. The syringes (5 ml; Spectrum Laboratories, East Houston, TX 7703, U.S.A., or from Cambridge Glassblowing) have an adjustable Teflon O-ring seal on the plunger and a

Fig. 3. The reaction column and mounting

The glass column (C) is held in a Dural casing which is machined to allow the column to be seen through a Perspex (Lucite) block (J) held in place by two knurled screws (H). The column is held into the Dural block by a component (D) made of polytetrafluoroethylene, containing a porous Teflon plug. This is held in place by the outer part of the Altex fitting (K). Into this is screwed the inner part of the fitting (L) through which the flanged outlet tube, with a metal washer (E), passes.

screw-threaded outlet on the glass barrel. The outlet tubing is connected via an Altex fitting to the end of the syringe.

(e) The reaction column

The details of the reaction column (Section c) and its mount are given in Fig. 3. The mount is held by two screws, I, in a recess machined in the heating block. Either of two columns with internal volumes of $15\ \mu\text{l}$ and $7.5\ \mu\text{l}$ can be fitted. The column and holder are usually dismantled together by unfastening screws, I, before being filled with glass beads, A, as illustrated in Fig. 3.

(f) Plumbing

A major difference between the present machine and earlier instruments is that we have tried to minimize the amounts of Teflon tubing used in the instrument, and particularly to reduce the flow path of reagents and solvents. This has been achieved by placing the slider valves close to each other and to the columns, pumps and reagent reservoirs. Connections are made in Teflon tubing (0.3 mm int. diam., 1.5 mm outer diam.) with Altex fittings. The total dead volume of the flow path

from valve V4, through the column to valve V7 is $90\ \mu\text{l}$ (with the $15\ \mu\text{l}$ column in place).

(g) Autoconversion

The design of the autoconverter is shown in Fig. 2. Before a sequencing experiment, glass vials (Hewlett Packard; part no. 5080-8779; 0.3 ml capacity) are placed in the carousel and a solution of dithiothreitol in water ($120\ \mu\text{l}$; $50\ \text{pmol}/\mu\text{l}$) is introduced into each tube. At the end of the cleavage reaction the trifluoroacetic acid ($\sim 50\ \mu\text{l}$) containing the cleavage product is displaced from the reaction column by argon gas, with the waste/collect valve (V7) activated. Thereby the anilinothiazolinone is transferred from the reaction column to the conversion vial in the same reagent as used in cleavage. After collection, valve V7 is deactivated and then argon gas blown into the conversion vial by activation of valve V8. At the same time the tube heater, set at 75°C , rises to surround the tube, and remains there for 18 min when conversion is complete. Then the heater is lowered and the carousel advanced one place in a clockwise direction. (For details of the autoconversion cycle, see Table 3.) The entire autoconverter is surrounded by a Perspex (Lucite)

Fig. 4. Interface between the microprocessor and the sequencer

This illustrates the general decode routing and device information latches.

box from which trifluoroacetic acid vapour is extracted with a fan into an exhaust duct.

(h) Instrument controller

The instrument is controlled by a PDP-11 minicomputer via an interface. The interface, summarized in Fig. 4, allows the switching of 105 devices by means of a system of decoders and data latches. The eight address bits of the PDP-11 are divided into upper and lower groups of four bits which select both a binary-to-10-way master decoder, and binary-to-15-way device decoders. This potentially allows the use of seven 15-way device decoders which therefore can control 105 separate devices.

The outputs from the device decoder directly switch 'opto isolated' relays used to operate the 240 V alternating-current solenoid valves that deliver compressed air [480 kPa (70 lbf·in²)], which in turn activates the slider valves. For controlling more complex devices such as pumps, heaters, displays etc., each output from a device decoder is used to latch the eight data bits from the PDP-11.

These eight bits are used to supply the particular device with all information necessary to control pump speeds, volumes, heater temperature displays and autoconverter, through a system of look-up tables resident within the PDP-11 software.

Once the data bus information is latched, the PDP-11 is free to operate further devices in any order. By means of this interface, any 16-bit computer can operate a large number of independent devices. This arrangement enables a very wide range of control to be exercised over complex electro-mechanical devices.

MATERIALS AND METHODS

Reagents and solvents

Dithiothreitol (high purity grade) was obtained from Calbiochem. *p*-Phenylene di-isothiocyanate was purchas-

ed from Pierce and Warriner, Chester, U.K.; it was supplied as white needles and was used without further purification. Samples of this reagent obtained from other sources (Sigma, Eastman-Kodak) were recrystallized from acetone. 3-Aminopropyltriethoxysilane was purchased from Pierce and Warriner and used without purification. Dimethylformamide (AR grade; BDH Chemicals, Poole, U.K.) was redistilled *in vacuo* immediately before use. Phenyl isothiocyanate (sequencer grade; Rathburn Chemicals, Walkerburn, Peeblesshire, Scotland, U.K.) was usually used without further purification. Occasionally it was re-distilled *in vacuo* [b.p. 65 °C, 800 Pa (6 mmHg)]. Triethylamine (sequencer grade; Rathburn Chemicals) was distilled twice or thrice from ninhydrin immediately before being used to make buffer. Acetonitrile (S grade; Rathburn Chemicals) was used for liquid chromatography without further purification; for use in the sequencer it was redistilled from ninhydrin. Benzene and methanol (both AR grade, 'getrocknet'; Merck Chemicals, Darmstadt, Germany) were occasionally re-distilled under N₂ in the presence of dithiothreitol (10 mg/litre). Otherwise they were used as supplied. Trifluoroacetic acid (AR grade, 99% containing < 0.1% water; Koch-Light Chemicals, Haverhill, Suffolk, U.K., or Aldrich Chemicals) was refluxed over potassium dichromate and then redistilled. Distillates from some batches were yellow, as noted by Laursen (1967). These batches were rejected. The distillate was left overnight over dry neutral alumina (Woelm Pharma, Eschwege, Germany) then decanted, refluxed over dithiothreitol (50 mg/100 ml) and distilled (Hunkapiller *et al.*, 1983a,b). It was stored under nitrogen and used in the sequencer without addition of dithiothreitol as recommended by Hunkapiller *et al.* (1983a,b). Addition of this reagent was found to give rise to a spurious peak, eluted in the position of PTH-proline, in PTH analysis by h.p.l.c. However, dithiothreitol was added to both sequencer solvents, methanol and benzene (10 mg/l); sonication helps dissolution in benzene. The volumes of reagents consumed per cycle are summarized in Table 1.

Table 1. Quantities of reagents consumed by the solid-phase microsequencer

Reagent	Volume/cycle (μ l)
Phenyl isothiocyanate*	825
Trifluoroacetic acid	130†
Buffer‡	990
Methanol	1000
Benzene	650

* 10% in acetonitrile.
† With double cleavage.
‡ 2.5% triethylamine in aq. 65% (v/v) methanol.

Glass supports

3-Aminopropyl-glass was prepared by silylation of controlled-pore glass [200–400 mesh; mean pore size 7 nm (70Å) (Corning)] as described by Robinson *et al.* (1971), except that a 4% (w/v) solution of silylating agent in acetone was used. The coupling capacity (Schmitt & Walker, 1977) of the derivatized glass in dimethylformamide was 200–250 nmol/mg. Activation of the 3-aminopropyl-glass to the phenylene 1,4-di-isothiocyanato derivative was carried out in either dimethylformamide or benzene containing 0.2% (v/v) triethylamine, by using a 1.2-fold molar excess of *p*-phenylene di-isothiocyanate (50 mg/g of glass). Freshly distilled dimethylformamide must be used, otherwise the spontaneous breakdown product, dimethylamine, reacts with the isothiocyanate. The derivatized glass was washed with benzene rather than with methanol or dimethylformamide as described previously (Wachter *et al.*, 1973). Washing was carried out in a sintered-glass column (12 cm $\times$ 4 mm int. diam.) with gentle suction, care being taken not to draw air over the glass. Finally, the glass was dried *in vacuo* and stored under nitrogen at -15°C .

Immobilization of proteins and peptides

Cytochrome *c* from the yeast *Candida krusei* and sperm-whale myoglobin (both purchased from Sigma Biochemicals) were employed as standard proteins to assess sequencer performance. They were immobilized in bulk on DITC-aminopropyl-glass as follows. The derivatized glass support (30 mg) was added to a solution of the protein (1 nmol/ μ l in 0.1 M-NaHCO₃) to give a final loading on the glass of 0.25, 0.5 or 1.0 nmol/mg of glass. Sufficient additional buffer was added to completely cover the glass beads (total volume of solution $\sim$ 50 μ l). After 18 h reaction at room temperature, the beads were removed by filtration on a small sintered-glass column (3 cm $\times$ 0.4 cm int. diam.), suction being applied to the bottom of the column with a syringe attached to a piece of plastic tubing. Then the beads were washed with 0.02% triethylamine in methanol. During these filtration operations care was taken not to draw air through the beads. The beads were dried *in vacuo* and distributed into several capped tubes and stored under nitrogen at -20°C .

To measure the amount of protein covalently attached to the glass, a weighed amount ($\sim$ 1 mg) was washed with trifluoroacetic acid on a sintered-glass funnel with suction from an aspirator, and then dried *in vacuo*. Immobilized protein was then hydrolysed *in vacuo* at

Table 2. H.p.l.c. flow program for separation of PTH-amino acids

A du Pont Gold Series Zorbax ODS column (3 μ m particle size; 8 cm $\times$ 6.2 mm int. diam.) was employed.

Time (min)	Flow rate (ml/min)	[Acetonitrile] (%)
Initial	1.2	30
0–2.5	1.2	30–45*
2.5–5.5	1.2	45 isocratic
5.6–6.0	1.2	50 isocratic
6.0–7.0	1.2–2.0	50
7.0–7.8	2.0	50 isocratic
7.9–8.9	2.0	30 isocratic
9.0–12.8†	1.2	30

* Non-linear, convex gradient (curve 5 with Waters model 720 systems controller).

† The gradient returns to the initial conditions after 9 min. The autosampler commences the procedure for injection of the subsequent sample after 11 min and makes the injection at 12.8 min = 0 min.

110 $^{\circ}\text{C}$ for 24 h with 6 M-HCl containing 0.1% phenol. The released amino acids were measured by amino acid analysis.

Peptides with C-terminal lysine, obtained by digestion of proteins with trypsin or endoproteinase lys C, were immobilized by a similar procedure. The dried peptide was dissolved in 0.1 M-NaHCO₃ (30 μ l), added to $\sim$ 10 mg of DITC-glass, and a further amount of buffer ($\sim$ 10 μ l) added to just cover the beads. The reaction conditions and work-up procedures are as described for immobilization of standard proteins.

C-Terminal arginine residues were converted into ornithine by treatment under nitrogen with aq. 50% (v/v) hydrazine (30–40 μ l) at 70 $^{\circ}\text{C}$ for 15 min (Morris *et al.*, 1973). The peptides were then immobilized as described for C-terminal lysyl-peptides.

Peptides with C-terminal homoserine, resulting from cleavage of proteins with CNBr, were immobilized on 3-aminopropyl-glass by a scaled-down version of the procedure described by Horn & Laursen (1973).

Hydrophobic proteins and peptides were coupled to DITC-glass in buffers containing 0.1% SDS and 10–25% (v/v) ethanol.

Immobilized standard proteins (1–2 mg) were transferred to the sequencer column with a small spatula with a spoon at the end, via a funnel made by cutting down a disposable blue Eppendorf pipette tip. Remaining space in the column was filled with underivatized glass beads. Unknown proteins or peptides were immobilized on sufficient DITC-glass to completely fill the reaction column (5 or 10 mg). The glass beads were introduced into the column from the Eppendorf tube in which coupling was conducted, as illustrated in Fig. 3.

Identification of PTH amino acids

PTH amino acids are identified by reverse-phase h.p.l.c. on a Zorbax Gold Series ODS column (du Pont; 3 μ m packing; 8 cm $\times$ 6.2 mm int. diam.) maintained at 50 $^{\circ}\text{C}$ by a water jacket. Solvents are delivered by a Waters h.p.l.c. system. This consists of two model M6000

Fig. 5. Separation of PTH-amino acids by h.p.l.c.

Absorbance was measured at 254 nm or (lower line) 313 nm (for dehydroalanine, ΔS , and dehydrothreonine, ΔT); ———, [acetonitrile] gradient; - · - · - flow program (see Table 2). The capital letters signify amino acids in the one-letter code.

pumps controlled by a model M720 system controller and a model 710A WISP auto sampler. The separated PTH derivatives (see Fig. 5) are detected with a Waters model 440 u.v. detector equipped for simultaneous dual-wavelength measurement at 254 and 313 nm. The PTH amino acids are separated by gradient elution with an aqueous phase buffered with sodium acetate (0.0165 M, pH 5.1). The flow rate is programmed to change as shown in Table 2 and Fig. 5. The total time between injections is 13.0 min.

The flow diagram described in Fig. 5 separates all PTH amino acids with the exception of the derivatives of phenylalanine and lysine. These can be resolved by using a modified flow program in which the acetonitrile concentration from 5 to 11 min is maintained at 38%. PTH-glutamine and PTH-serine standards are partially resolved, but serine is recovered from the sequencer almost entirely as PTH-dehydroalanine. This product is completely resolved from other products and in addition is detected at 313 nm rather than 254 nm. The adduct between PTH-serine and dithiothreitol is eluted between PTH-glycine and PTH-alanine. The by-product, diphenylthiourea, is eluted before lysine/phenylalanine and is completely resolved from them. It should be noted that

this by-product is removed efficiently by the post-coupling washes and does not interfere seriously with PTH identification.

RESULTS

The degradation cycle

The use of a microprocessor to control the instrument has brought a number of benefits. Of these perhaps the most important is that it has given flexibility to the experimenter and has allowed parameters controlled by the program to be changed at will. This was not possible in earlier instruments. Thus we have been able to investigate systematically the degradative program. Thereby many steps have been shortened or eliminated.

The current version of the degradative cycle (see Table 3) is 24 min long. Most of the cycle (16 min) is devoted to the coupling reaction; cleavage lasts 5.0 min; collection of the ATZ derivative a further 1.5 min and washes take up 1.5 min of the cycle (see Table 3). The conversion step is carried out from the end of cleavage and continues into the subsequent coupling steps. It takes a total of 18 min. This cycle has been used successfully over a 3-year period. It is by far the shortest Edman-degradation cycle in any extant protein sequencer (see Fig. 6). Its major advantage is that it exposes the protein for the minimum time required for effective Edman degradation and thereby limits the opportunity for unwanted side reactions to occur.

Performance of the instrument

The utility of the instrument has been demonstrated in numerous successful degradations of proteins and peptides over the past 3 years, particularly in the investigation of the F_1 -ATPase complex from bovine mitochondria (see Runswick & Walker, 1983; Hollemans *et al.*, 1983; Walker *et al.*, 1985). A demonstration of the efficiency of degradation in the instrument was the determination of 54 residues (corresponding to amino acids 2–55) of a 64-residue CNBr fragment from the γ -subunit of bovine mitochondrial F_1 -ATPase (see Table 4). Most of these experiments were conducted with 0.5–2 nmol of peptides or proteins. A number of other proteins have been recovered from polyacrylamide gels after electrophoretic separation in the presence of SDS (see Table 4 and Walker *et al.*, 1982). More recently, studies of a number of relatively non-abundant proteins of biological interest have led to a greater emphasis in exploiting the ability of the instrument to operate at higher sensitivities. This was first demonstrated by the *N*-terminal degradation of 66 pmol of sperm-whale myoglobin and subsequently with a number of proteins of unknown sequence (see Table 4).

Polar amino acids

A significant difference between the solid-phase sequencer and the gas- and liquid-phase sequencers concerns the way in which the products of cleavage, the anilinothiazolinones, are removed from the reaction chamber. In the gas-phase and liquid-phase instruments this is achieved by extraction with butyl chloride. In the solid-phase sequencer the cleavage products are removed from the reaction chamber (column) dissolved in the trifluoroacetic acid that is employed in cleavage. At the end of this reaction it is displaced from the column with

Table 3. Listing of the degradation cycle

Time (min)	Command	Procedure	Function
0:00	Start	Start timer of display	Coupling
0:01	Turn on V2 and P2 (speed 60)	Deliver buffer to column	
	Turn on V14	Pressurize TFA	
0:30	Turn on V1 and P1 (speed 60)	Deliver PITS to column	
1:30	Turn off V1, P1, V2 and P2	Stop delivery of PITS and buffer	
1:31	Turn on P1 (speed -120) and P2 (-120)	Refill buffer and PITS syringes	
2:20	Turn on V1, P1 (60), V2, P2 (60)	Deliver buffer and PITS	
	Repeat cycle of 10 s deliveries of buffer and PITS and 20 s pauses 27 times		
4:00	Turn on V14	Pressurize TFA	
12:00	Turn on V9	Pressurize methanol	
16:00	Turn off V1, P1, V2, P2, V9	End of coupling	Conversion
16:05	Turn on P1, (-110), P2 (-110), V10	Refill PITS buffer and syringes	
16:30	Turn on V4; V11; Off V10	Deliver methanol to column	
16:45	Turn off V11; On V13	Pressurize buffer	
17:00	Turn off V4, V13; On V5, V14	Stop methanol delivery; deliver benzene	
17:30	Turn off V5, V14; On V6, V12	Stop benzene delivery; argon to column	
17:40	Turn off H1	Lower conversion heater	
17:50	Turn off V6, V12, V8;	Stop argon to column; deliver TFA, with waste collect valve to vial.	
	Turn on V3, P3 (60), V7, V10	Advance fraction collector	
17:53	F	Stop TFA delivery	
18:20	Turn off V3, P3	Displace TFA from column with argon	Cleavage
20:20	Turn on V6, V12; Off V10	Stop argon; deliver TFA	
20:40	Turn off V6, V12; On V3, P3 (60), V13	Stop TFA delivery	
21:00	Turn off V6, V12; On V3, P3 (60), V13	Displace TFA from column with argon	
23:00	Turn on V6, V12	Raise conversion heater	Solvent wash
23:05	Turn on H1	Stop argon delivery; deliver methanol to waste via column and argon to conversion vial	
23:30	Turn off V6, V7, V12, V13; on V4, V8	Refill TFA syringe	
23:35	Turn on P4 (-110), V11	Stop methanol delivery; increment cycle counter	
24:00	Turn off V4, V11, increment cycle counter		

argon and collected. This difference in design has an important consequence. It was first illustrated by attempts to degrade a peptide containing amino acid derivatives formed by reaction of bovine mitochondrial F_1 -ATPase with the photoactivatable active-site analogue, 8-azido- $[^3H]$ ATP. On illumination of the complex between enzyme and this substrate analogue a nitrene is formed. This then reacts with amino acid side chains in the vicinity of the active site of the enzyme. These amino acid derivatives contain phosphate groups and are very polar. Initially, in an attempt to locate the reacted amino acids, the peptide, a tryptic arginine fragment from the β -subunit of the enzyme, was degraded in a spinning-cup sequencer. The degradation proceeded successfully, except that the sequence contained gaps where apparently no amino acid derivative was released. Moreover, no radioactivity corresponding to radiolabelled side chains could be detected. Therefore the remainder of the peptide was treated with hydrazine to convert the C-terminal arginine residue to ornithine and the product immobilized by reaction with isothiocyanato-glass and degraded in the solid-phase sequencer. This time the sequence contained no gaps and radioactivity was released in the positions where gaps were noted previously (Hollemaans *et al.*, 1983). The explanation of this success in the solid-phase instrument over the failure in the liquid-phase instrument

is that, in the solid-phase sequencer, the polar ATZ derivatives of the derivatized amino acids are soluble in trifluoroacetic acid and are thus removed from the reaction column; in the liquid-phase instrument the released anilinothiazolinones, because of their polarity, are insoluble in butyl chloride and so remain in the reaction cup. Recently, it has been shown that derivatives of 8-azido ATP are not recovered from the gas-phase sequencer (Knight & McEntee, 1985).

A more commonly encountered polar amino acid is phosphoserine resulting from protein phosphorylation. Its anilinothiazolinone also cannot be extracted efficiently with butyl chloride, whereas it is readily soluble in trifluoroacetic acid and its presence can be positively identified in sequences. This was demonstrated by sequence analysis of a $[^{32}P]$ phosphorylated peptide from the branched-chain-2-oxo-acid dehydrogenase complex (Cook *et al.*, 1984).

Proteins recovered from polyacrylamide gels

A number of N-terminal sequences have been determined on proteins recovered from polyacrylamide gels (Walker *et al.*, 1982). This information has been useful in identification of genes in DNA sequences (see Table 4).

Fig. 6. Comparison of degradation cycles in different designs of protein sequencer

The liquid-phase program is based upon that described by Brauer *et al.* (1975) and the gas phase is quoted from Hunkapiller *et al.* (1983).

DISCUSSION

Advantages of the instrument

We have described above two advantages of the solid-phase microsequencer. Firstly, it is unique among the three types of extant sequencer in so far as it can be used to positively identify very polar amino acids in peptides. This is an important advantage, given that one of the major roles of direct protein sequence analysis in modern molecular biology is the identification of post-translational modifications, amongst which phosphorylation is an important example. A second role is the identification of sites of covalent modification with active-site analogues. In the case of analogues of ATP such as 8-azido-ATP (Hollemaans *et al.*, 1983; Knight & McEntee, 1985), again only the solid-phase approach can be used to positively identify modified residues.

A second advantage is the speed of degradation, which is more than twice that achieved in other instruments (see Fig. 6). Apart from increased convenience and greater efficiency in processing protein samples, this helps to minimize contact of protein with inimical reagents that might degrade the protein and give rise to side reactions.

A third advantage is that because of the way in which conversion is carried out, the instrument requires two fewer reagents than are used in the liquid- or gas-phase instruments.

A fourth advantage that has not been discussed is encountered in the sequence analysis of very hydrophobic membrane proteins. By nature they tend to be soluble in the relatively non-polar organic solvents (benzene, ethyl acetate, butyl chloride) that are used for extraction of by-products in the liquid-phase and gas-phase sequencers, and unless steps are taken to counter this tendency, they

Table 4. Some protein sequences determined in the solid-phase microsequencer

With the exception of CNBr fragment (homoserine lactone method) all proteins/peptides were coupled by the DITC method.

Protein	Quantity	Sequence	Footnote
Sperm-whale myoglobin	66 pmol	XLSEGEWQLV LHV ¹⁰	
CB Fragment of γ -subunit of bovine F_1 -ATPase	1.8 nmol	XAAAKYARAE RELKPARVYG VGSLALYEKA* ^{10 20 30} DIKTPEDKKK HLIIGVSSDR GLCGA ^{40 50}	
Bovine phosphate transporter	300 pmol	AVEEQYSXDY GSGXF/K FILXG LGGI ^{10 20}	††,‡‡
Human β T	600 pmol	XQRPPK/FIQVY SRHPPEDGK/FP NYLNXVYV ^{10 20}	‡\$,‡‡
Protein from mouse gap junction	1 nmol	XPEYSSFFGV MGASSAMVFV AMGXYG ^{10 20}	‡
Wheat chloroplast ATP synthase ϵ -subunit	400 pmol	XLNLYVLXPK RXIXDXE ^{10 15}	‡¶
Spinach chloroplast ATP synthase β -subunit	300 pmol	XINPTTSDPG VXTL ^{5 10}	‡, **
Spinach chloroplast ATP synthase subunit-I	300 pmol	XSFGFNTDXL ^{5 10}	‡, **
<i>Escherichia coli</i> TyrA	400 pmol	XAELTALRPQ XEVDK ¹⁰	‡, ††

*Walker *et al.* (1985). †In collaboration with H. Wohlrab. ‡The estimates of the quantities employed are maximal amounts. The sequencing experiments demonstrated that considerably less material was present. §Kefford *et al.* (1984). ||In collaboration with J. D. Pitts and J. Shuttleworth. ¶Howe *et al.*, 1985. **In collaboration with J. C. Gray. Samples were electroeluted from polyacrylamide gels as described by Walker *et al.* (1982). ††Hudson & Davidson, 1984. ‡‡Obliques indicate that lysine and phenylalanine were not distinguished.

are rapidly extracted from the instrument and lost. The addition of the polycationic substance, polybrene (Klapper *et al.*, 1978; Tarr *et al.*, 1978) to the spinning cup and to the glass-fibre disc of the gas-phase instrument helps, but extraction from the spinning-cup instrument is not totally prevented with very hydrophobic proteins. Few, if any, proteins in this category have been reported to have been degraded after absorptive application in the gas-phase instrument. Covalent attachment to glass beads prevents such losses, and a number of proteolipids (i.e. proteins that are soluble in mixtures of chloroform and methanol; Folch & Lees, 1951) have been sequenced without difficulty in solid-phase instruments. Covalent attachment has an additional advantage, namely that the by-products of coupling are removed very efficiently by the post-coupling washes and so they do not interfere with PTH identification. The presence of these by products (eg. diphenylthiourea) in PTHs from the liquid- and gas-phase sequencers causes difficulties in the identification process.

Disadvantages of the solid-phase approach

The covalent coupling of proteins to glass supports is more complex than application of a protein sample to a spinning-cup or gas-C phase instrument and used to be regarded by users of solid-phase sequencers as a troublesome process. This is no longer so, either in the case of attachment by the DITC or homoserine lactone coupling procedures. For success with both methods it is of paramount importance to effectively solubilize the protein or peptide in an amine-free reagent. For DITC coupling, aqueous buffers containing a wide range of chaotropes including 6 M-guanidinium chloride, 0.1–5% SDS, 2–8 M-urea, and urea/SDS mixtures have been

successfully and routinely used. In extreme cases, non-aqueous solvents have also been employed. For example, in the case of a very hydrophobic protein isolated from mitochondria, quantitative coupling to DITC-glass was performed in a chloroform/methanol (2:1, v/v) mixture containing 1% triethylamine (I. M. Fearnley & J. E. Walker, unpublished work). The key to success for homoserine lactone coupling is that the anhydrous peptide should be dissolved in a dry, high-quality (i.e. amine-free) solvent, usually dimethyl-formamide. For ready dissolution the peptide must be completely anhydrous. This is achieved by drying the peptide down from trifluoroacetic acid *in vacuo*, followed immediately by addition of the solvent and the derivatized support. With these minor precautions, quantitative coupling is achieved routinely with both methods and high yields are also obtained by the C-terminal spiro-lactone method (Wachter & Wehrhahn, 1979).

In combination with an appropriate strategy, these methods are sufficient to generate complete overlapping sequences (although with advent of DNA cloning and sequencing the need for such experiments is receding, and partial sequences are sufficient). One such strategy would be a combination of two principal digests, CNBr cleavage and enzymic digestion with endoproteinase lys C. The former generates peptides with C-terminal homoserine lactone and the latter digest generates peptides with C-terminal lysine residues. Therefore these peptides are ideally suitable for DITC or homoserine lactone coupling. The only peptides that could not be sequenced by this approach would be the C-terminal fragments of both digests (except in the somewhat rare event of the

intact protein having either C-terminal methionine or lysine. In order to determine which peptide is the C-terminal fragment in these digests, it is desirable, as in other forms of sequence analysis, to determine the amino acid compositions of peptides in advance. In order to complete the sequence analysis by the solid-phase approach alone, additional coupling methods would have to be employed. Coupling via carboxy groups with carbodi-imides has been used, but with variable success. A recent development employs activation of carboxy groups of the peptide by treatment with a mixture of 10% trifluoroacetic anhydride in trifluoroacetic acid. We have successfully used this procedure during the sequence analysis of bovine F_1 -ATPase (Walker *et al.*, 1985). In case of need, other more specialized methods have also been employed. For example, during the identification of sites of labelling of F_1 -ATPase with 8-azido-ATP, we were obliged to employ a solid-phase method for sequence analysis of a tryptic peptide with C-terminal arginine, the liquid-phase method having been found to be unsatisfactory (see above). Therefore the peptide was treated with hydrazine and the generated C-terminal ornithine residue used to couple the peptide (Holleman *et al.*, 1983). As a general method this procedure has disadvantages. The most serious is that hydrazine treatment, albeit under rather severe conditions, may lead to chain cleavage (Honegger *et al.*, 1981), although this can also be turned to advantage.

In rare cases (absence of lysine, methionine, tryptophan; presence of C-terminal amides) peptides have no suitable groups for covalent attachment; in these cases other methods for sequence analysis are more suitable.

However, covalent attachment can be advantageous as already discussed above. Moreover, it can be used to selectively scavenge proteins from contaminants that are found in protein samples recovered from polyacrylamide gels (Walker *et al.*, 1982).

Although coupling yields are usually close to quantitative, the initial yields in degradations of proteins attached by the isothiocyanate method are usually about 30%. This presumably reflects oxidation of the isothiocyanate, a problem that requires further investigation. In our experience, initial yields in liquid-phase sequencers with nanomolar samples are > 90%; in the gas-phase sequencer they are 30–90%.

An additional disadvantage of the solid-phase method is that in degradation of proteins by the di-isothiocyanate method, but not other methods, the first residue remains attached to the support and is 'silent'. However, for many purposes this is not a serious problem.

Perspectives

The gas-phase sequencer can be seen as the successor to the older liquid-phase instrument. It has the advantage over the liquid-phase instrument of having been designed specifically to accommodate analysis of subnanomolar quantities of material. This has been achieved in part by close attention to reagent purity and to delivery of minimal amounts of buffer (trimethylamine) and cleavage acid (trifluoroacetic acid) by delivering them as vapours in an argon stream. It has not, however, as demonstrated above, removed the need for solid-phase sequencing. In the area of membrane proteins, careful objective evaluation of the comparative performances of the instruments in degradations of membrane components is

required. The availability of both instruments in this laboratory now makes such a study feasible. However, it is apparent from the designs of the two instruments that the techniques are converging. This is further emphasized by attempts to degrade samples covalently coupled to glass beads in the gas-phase instrument (Strickler *et al.*, 1984).

We thank Mr. M. Bitton, Mr. C. Hellon and Mr. N. Young of the mechanical-engineering workshop of this laboratory for their help in construction of the instrument. We are grateful to Dr. J. Shuttleworth and Dr. J. D. Pitts for samples of a protein isolated from mouse gap junctions and to Dr. H. Wohlrab for samples of mitochondrial phosphate transporter.

REFERENCES

- Allen, G. (1981) in *Sequencing of Proteins and Peptides*, (Work, T. S. & Burdon, R. H., eds.), pp.174–200, North-Holland, Amsterdam, New York and Oxford
- Brauer, A. W., Margolies, M. N. & Haber, E. (1975) *Biochemistry* **14**, 3029–3035
- Cook, K. G., Bradford, A. P., Yeaman, S. J., Aitken, A., Fearnley, I. M. & Walker, J. E. (1984) *Eur. J. Biochem.* **145**, 587–591
- Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
- Edman, P. & Henschen, A. (1975) in *Protein Sequence Determination* (Needleman, S. B., ed.), pp. 232–279, Springer-Verlag, Berlin, Heidelberg and New York
- Folch, J. & Lees, M. (1951) *J. Biol. Chem.* **191**, 807–817
- Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. (1981) *J. Biol. Chem.* **256**, 7990–7997
- Holleman, M., Runswick, M. J., Fearnley, I. M. & Walker, J. E. (1983) *J. Biol. Chem.* **258**, 9307–9313
- Honegger, A., Hughes, C. J. & Wilson, K. J. (1981) *Biochem. J.* **199**, 53–59
- Horn, M. & Laursen, R. A. (1973) *FEBS Lett.* **36**, 285–289
- Howe, C. J., Fearnley, I. M., Walker, J. E., Dyer, T. A. & Gray, J. C. (1985) *Plant Mol. Biol.* **4**, 333–346
- Hudson, G. S. & Davidson, B. E. (1984) *J. Mol. Biol.* **180**, 1023–1051
- Hunkapiller, M. W. & Hood, L. E. (1978) *Biochemistry* **17**, 2124–2133
- Hunkapiller, M. W. & Hood, L. E. (1980) *Science* **207**, 523–525
- Hunkapiller, M. W., Lujan, E., Ostrander, F. & Hood, L. E. (1983a) *Methods Enzymol.* **91**, 227–236
- Hunkapiller, M. W., Hewick, R. M., Dreyer, W. J. & Hood, L. E. (1983b) *Methods Enzymol.* **91**, 399–413
- Kefford, R. F., Calabi, F., Fearnley, I. M., Burrone, O. R. & Milstein, C. (1984) *Nature (London)* **308**, 641–642
- Klapper, D. G., Wilde, C. E. & Capra, J. D. (1978) *Anal. Biochem.* **85**, 126–131
- Knight, K. L. & McEntee, K. (1985) *J. Biol. Chem.* **260**, 10185–10191
- Laursen, R. A. (1971) *Eur. J. Biochem.* **20**, 89–102
- Laursen, R. A., Bonner, A. G. & Horn, M. J. (1975) in *Instrumentation in Amino Acid Sequence Determination* (Perham, R. N., ed.), pp. 73–110, Academic Press, London, New York and San Francisco
- Morris, H. R., Dickinson, R. J. & Williams, D. H. (1973) *Biochem. Biophys. Res. Commun.* **51**, 247–255
- Robinson, P. J., Dunnhill, P. & Lilly, M. D. (1971) *Biochim. Biophys. Acta* **242**, 659–661
- Runswick, M. J. & Walker, J. E. (1983) *J. Biol. Chem.* **258**, 3081–3089
- Schmitt, H. W. & Walker, J. E. (1977) *FEBS Lett.* **81**, 403–405
- Strickler, J. E., Hunkapiller, M. W. & Wilson, K. J. (1984) *Anal. Biochem.* **140**, 553–566
- Tarr, G., Beecher, J. F., Bell, M. & McKean, D. J. (1978) *Anal. Biochem.* **84**, 622–627
- Wachter, E. & Wehrhahn, R. (1979) *Anal. Biochem.* **97**, 56–64

- Wachter, E., Machleidt, W., Hofner, H. & Otto, J. (1973) *FEBS Lett.* **35**, 97–102
- Walker, J. E., Auffret, A. D., Carne, A., Gurnett, A., Hanisch, P., Hill, D. & Saraste, M. (1982) *Eur. J. Biochem.* **123**, 253–260
- Walker, J. E., Eberle, A., Gay, N. J., Hanisch, P., Saraste, M. & Runswick, M. J. (1982) in *Methods in Peptide and Protein Sequence Analysis* (Elzinga, M., ed), pp. 337–354, Humana Press, Clifton, NJ
- Walker, J. E., Fearnley, I. M., Gay, N. J., Gibson, B. W., Northrop, F. D., Powell, S. J., Runswick, M. J., Saraste, M. & Tybulewicz, V. L. J. (1985) *J. Mol. Biol.* **184**, 677–701

Received 15 November 1985/5 February 1986; accepted 28 February 1986