Vol. 166, No. 1, 1990
January 15, 1990

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 139-145

INTERNAL AMINO ACID SEQUENCING OF PROTEINS BY *IN SITU* CYANOGEN BROMIDE CLEAVAGE IN POLYACRYLAMIDE GELS

Willi Jahnen[1,2], Larry D. Ward[1], Gavin E. Reid[1], Robert L. Moritz[1] and Richard J. Simpson[1,3]

[1]Joint Protein Structure Laboratory
Ludwig Institute for Cancer Research (Melbourne Branch), and
The Walter and Eliza Hall Institute of Medical Research
Parkville, Victoria 3050, Australia

[2]Plant Cell Biology Research Centre
School of Botany, University of Melbourne, Parkville 3052, Australia

Received October 31, 1989

A new method was developed for generating peptide fragments for amino acid sequence analysis from polyacrylamide-gel separated proteins. This method involves *in situ* CNBr treatment of proteins in the polyacrylamide gel after their separation by electrophoresis. Pure CNBr peptides were recovered either by solvent extraction followed by microbore column reversed-phase HPLC or, alternatively, by a second electrophoretic separation step (SDS-PAGE) followed by electrotransfer of the peptides onto polyvinylidene difluoride (PVDF) membranes. These approaches yielded sequence data at subnanomole levels for a wide range of CNBr fragments recovered from gel-separated proteins.

Although significant advances have been made in the area of protein micropurification and amino acid sequencing technologies (for recent reviews see Simpson *et al* (1) and Shively *et al* (2)), the high cost and substantial skills involved with the use of specialized equipment (e.g. HPLC) still pose major obstacles to this technology having a more general application. While protein HPLC purification technology often remains in the domain of specialized laboratories, electrophoretic separation techniques are now standard analytical tools in many biological laboratories.

With the recent advent of immobilizing matrices that are chemically inert to the Edman degradation procedure, the power of electrophoretic separation techniques has been further enhanced since it is now possible to directly sequence proteins from one- (1D) and two-dimensional (2D) polyacrylamide gels. Using this simplified approach, proteins are first resolved by gel electrophoresis and then electrotransferred from the gel to an immobilizing matrix (e.g. derivatized glass fiber paper (3-5), PVDF (6)). Proteins

[3]To whom correspondence should be addressed.

0006-291X/90 $1.50

are then visualized (e.g. by staining with Coomassie blue R-250), excised, and, provided their N-terminus is unblocked, sequenced directly.

For proteins with blocked N-termini, internal amino acid sequence data provides invaluable information for protein identification. To obtain such information from gel-separated proteins, as a first step, proteins must be fragmented either proteolytically or chemically and the generated peptides isolated in a pure form. Various procedures have been reported for achieving the above. For instance, electroblotted proteins can be enzymically digested *in situ* on the immobilizing matrix (7-9) or following elution of the protein from the membrane (10); in all cases peptides were purified by microbore reversed-phase HPLC. Alternatively, samples can be proteolytically digested *in situ* in the gel matrix and the generated peptides recovered either by tandem SDS-PAGE/electroblotting (11-14) or by tandem solvent extraction from the gel/microbore reversed-phase HPLC (14-16).

We report here a new method for the *in situ* CNBr cleavage of proteins in polyacrylamide gel matrix and subsequent isolation of generated peptides for amino acid sequence analysis.

MATERIALS AND METHODS

Materials: Self-incompatible clones of *Nicotiana alata* were produced and maintained using procedures described by Anderson *et al* (17). Stylar S-glycoproteins S_1 and S_7 were purified according to the procedure of Jahnen *et al* (10,18). Lysozyme (g-type) was isolated from the egg white of the Australian black swan (19). Myoglobin (horse-heart), thioglycollic acid, tris(hydroxymethyl)-aminomethane (Tris) base were purchased from Sigma (St. Louis, MO, USA). SDS and CNBr were obtained from British Drug Houses (Poole, UK). Immobilon (PVDF) was a Millipore product. Recombinant mouse interleukin-6 (IL-6) was prepared as described elsewhere (20).

Electrophoretic techniques: Myoglobin and lysozyme (g-type) were electrophoresed (SDS-PAGE) using 12.5% polyacrylamide resolving gels (dimensions: 8 x 7 x 0.15 cm) using the Laemmli buffer system (21). After electrophoresis, proteins were visualized by Coomassie blue R-250 staining (0.2% w/v Coomassie blue in 2-propanol/acetic acid/water, 25:10:65) for 1 h and destaining in several changes of 2-propanol/acetic acid/water (10:10:80) until the background was clear. Protein-containing gel pieces were excised and stored in destaining solution at 4°C (up to one week).

2D gel electrophoresis of style extract proteins was performed using the non-equilibrium pH gradient (NEPHGE) procedure (22) in the first dimension as described elsewhere (10).

CNBr fragments were resolved by Tricine-SDS-PAGE according to the procedure of Schagger and von Jagow (23). Briefly, dried gel slices containing peptides generated by *in situ* CNBr treatment were equilibrated for 5 min with 1 ml of 1 M Tris-HCl buffer, pH 8.0. This procedure was repeated three times. The gel slices were then positioned in sample wells of a 1.5 mm thick Tricine-SDS-polyacrylamide gel (stacking gel - 4% T, 3% C, 6 cm long; spacer gel - 10% T, 3% C, 3 cm long; resolving gel - 16.6% T, 6% C, 11 cm long) and overlayed with Laemmli SDS-running buffer (21). Gels were electrophoresed at 20 mA constant current for 4 h and then at 50 mA for 16 h.

Electroblotting and visualization of proteins on PVDF membranes: Following Tricine-SDS-PAGE, gels were equilibrated with transfer buffer (10 mM 3-(cyclohexylamino)-1-propane sulfonic acid, 10% methanol, pH 11.0) prior to electroblotting (10 V constant voltage for 30 min using a Milli Blot-SDE "semi-dry" blotting apparatus from Millipore) onto PVDF membranes and visualized (by staining with 0.2% (w/v) Coomassie blue R-250 in aqueous 50% methanol / 10% acetic acid for 5 min) as previously described (10).

In situ CNBr cleavage: Protein-containing gel slices were dried by lyophilization and then treated with CNBr in 0.5 ml of aqueous 70% (v/v) formic acid for 16 h at 25°C at a protein CNBr mass ratio of 1:20-100. At the completion of the reaction, CNBr and

formic acid were removed by centrifugal-lyophilization in a speed-vac centrifuge (Savant, Hicksville, NY, USA).

HPLC purification of CNBr peptides extracted from gel slices: CNBr peptides were extracted at 4°C from gel slices with three 0.5 ml aliquots of 1% (v/v) aqueous trifluoroacetic acid (for 1 h, 4 h and 16 h, respectively). The extracts were combined (1.5 ml), centrifuged (5 min, bench top Eppendorf centrifuge) and the peptides fractionated by reversed-phase HPLC on a Hewlett-Packard model 1090A liquid chromatograph as described elsewhere (24). Fractions were collected manually in 1.5 ml polypropylene tubes (Eppendorf) and stored at -4°C.

Amino acid sequence analysis: Automated Edman degradation of samples was performed using Applied Biosystems sequencers (models 470A and 477A), equipped with model 120A on-line Pth-amino acid analyzers modified for total sample injection (25). Electroblotted samples, typically, were partially sliced and positioned upon a polybrene-treated glass filter in such a way to achieve optimal reagent/solvent flow (1).

RESULTS AND DISCUSSION

Fig. 1 shows a 2D gel of proteins extracted from mature styles of the ornamental tobacco, *Nicotiana alata*; self-incompatibility genotype S_7S_7. Since the products of the *S*-gene, *S*-glycoproteins, are highly basic (18), NEPHGE was employed in the first electrophoretic dimension to achieve optimal resolution, followed by size fractionation (SDS-PAGE) in the second dimension. In our previous studies (10), we identified a number of *N.alata* style proteins by direct sequencing from 2D gels. Protein identification using this approach, however, is limited to proteins with an unblocked N-terminus and to proteins which can be efficiently electrotransferred from the gel to the immobilizing membrane.

To obtain internal sequence data from gel-separated proteins, we explored the possibility of chemically fragmenting proteins with CNBr *in situ* in gel slices. Previously,

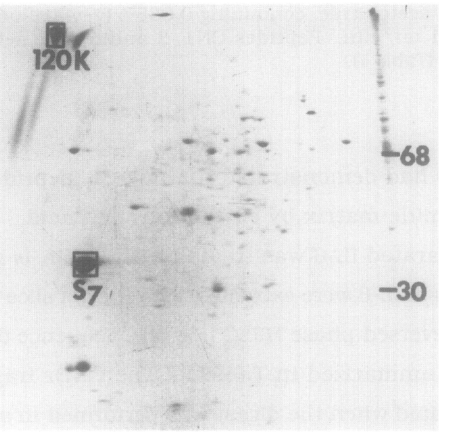

Figure 1 Pattern of *N.alata* (S_7S_7) style proteins obtained by 2D-gel electrophoresis and electroblotting onto PVDF membrane. 50 µg of total protein extract was separated by NEPHGE (horizontal axis) and SDS-PAGE (vertical axis). Proteins were stained with Coomassie blue R-250. Squares around marked proteins indicate those selected for *in situ* CNBr cleavage (Fig. 3) and internal amino acid sequence analysis (Table 1). M_r (kDa), determined from standard M_r markers (bovine serum albumin and carbonic anhydrase), are indicated on the right-hand axis. An extract from S_1S_1 plants was treated similarly (not shown).

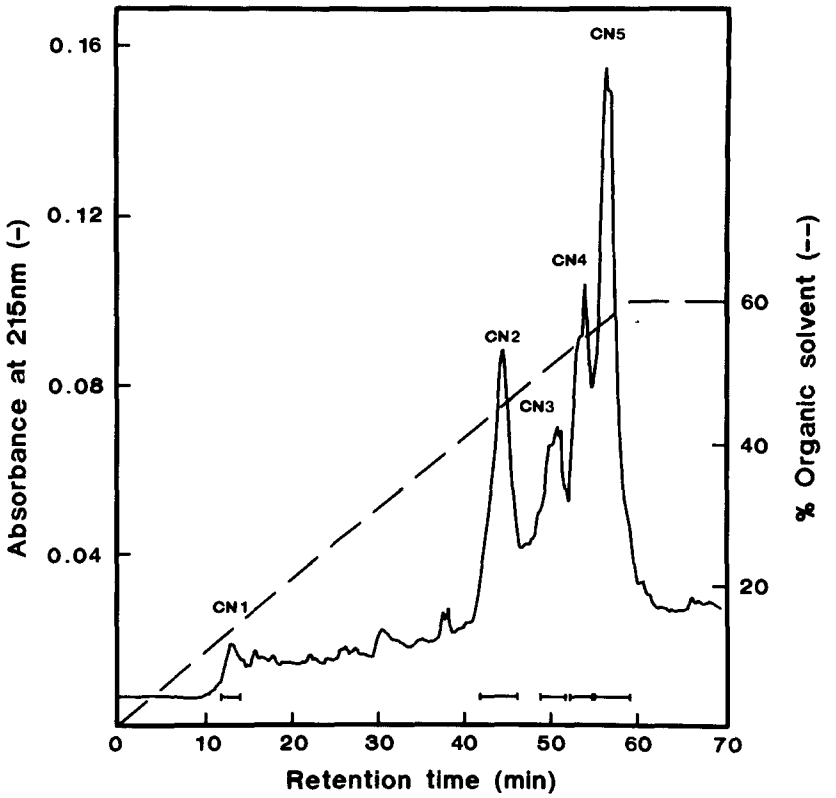

Figure 2 Reversed-phase HPLC peptide mapping of CNBr fragments of interleukin-6. Acrylamide gel containing IL-6 was excised and then treated with CNBr in 70% aq. formic acid. Generated peptides were extracted with 1% (v/v) aq. trifluoroacetic acid (total volume 1.5 ml), and applied to a reversed-phase column (Brownlee RP-300, 30 x 2.1 mm I.D.). Column was developed with a linear 60-min gradient from 0-100% B, where solvent A was 0.1% (v/v) aq. trifluoroacetic acid and eluent B was 60% aq. acetonitrile, containing 0.085% (v/v) trifluoroacetic acid. Flow rate, 0.1 ml/min. Peptides CN1, 2 and 5 were selected for sequence analysis (Table 1).

we (14) and others (15) had demonstrated that low M_r peptides could be efficiently recovered from acrylamide matrix by extraction with acid. Using this approach, electrophoretically-separated IL-6 was treated with CNBr, *in situ*, in the acrylamide matrix. CNBr fragments of IL-6 were extracted from the gel slice with 1% trifluoroacetic acid and separated by reversed-phase HPLC (Fig. 2). Sequence data, obtained from the indicated peptides is summarized in Table 1. The CNBr fragments identified are identical to those generated when the digest was performed in solution (24) indicating that the digest was complete and all fragments could be successfully extracted from the gel. This approach is particularly useful for proteins that are refractory to electrotransfer or electroelution (e.g. the highly basic 120 KDa stylar protein shown in Fig. 1, manuscript in preparation).

An alternative, and much simpler approach to isolating CNBr fragments from the acrylamide matrix relies on the tandem use of a second electrophoretic step and

TABLE 1: Sequence analysis of peptides generated by *in situ* CNBr treatment of proteins in polyacrylamide matrix[a]

Protein	Initial yield in 1st cycle (pmol)	Sequence	Residues sequenced
Myoglobin [b,e]			
peptide - 1	20	G L S D G E X Q Q V L N V X G	1 - 15
peptide - 2	19	K A S E D L K K H G T V V L T	56 - 70
Lysozyme (g-type) [b,e]			
peptide - 1	10	X R Y K T I I K K V G E K L X V E P A V I A G	46 - 68
peptide - 2	5	Q V D K R X H K P Q G T X N X E V	95 - 111
S_1-Glycoprotein [c,e]			
peptide - 1	5	Q L K F E Q D Y G R X E Q P	Unknown
peptide - 2	15	D F E X L Q L V L T W P A S F X Y A N H X E R I	1 - 24
S_7-Glycoprotein [c,e]			
peptide - 1	6	Q L V L Q X P X A F	6 - 10
peptide - 2	25	Q L V L Q X P T A F X H T T P X K R I P N K F	6 - 23
Interleukin-6 [b,f]			
CN 1	40	R K E L X N G N S D	42 - 51
CN 2	150	N N D D A L A E N N L K L P E I Q	54 - 70
CN 5	160	K N N L K D N K K D K A R V L Q R D T E T	104 - 124

[a]Amount of Coomassie blue-stained protein in polyacrylamide matrix (estimated by comparison with staining intensity of myoglobin): myoglobin, 60 µg (3.5 nmol); lysozyme, 60 µg (3 nmol); S_1-glycoprotein, 40 µg (1.3 nmol); S_7-glycoprotein, 40 µg (1.3 nmol); IL-6, 20 µg (1.0 nmol). Amino acid sequence is given in the one-letter notation. X, positions where an unambiguous Pth-amino acid assignment could not be made. [b]Protein isolated from 1D gel. [c]Protein isolated from 2D gel. CNBr peptides isolated by:[e]tandem Tricine-SDS-PAGE/electroblotting onto PVDF membranes; [f]tandem solvent extraction/microbore reversed-phase HPLC.

electroblotting. In the examples shown in Fig. 3, a Tricine-SDS-PAGE (23) step was used for fractionating CNBr fragments followed by electrotransfer of the separated fragments onto PVDF membranes. Selected fragments were subjected to sequence analysis and the data are summarized in Table 1. The low initial yields of phenylthiohydantoin-amino acid derivatives obtained for these peptides are most likely due to a combination of - (i) incomplete cleavage since the observed number of CNBr fragments (myoglobin-4, lysozyme-7 - see Fig. 3) exceeds the expected number of fragments based on possible methionine cleavage sites in these proteins, 3 and 4, respectively, - (ii) losses incurred during the additional electrophoresis step (26) and - (iii) poor recoveries during electrotransfer onto PVDF membranes. Such losses are emphasized by the much higher initial sequencing yields obtained for the acid extracted CNBr fragments of IL-6 (Table 1). In our experience (10,14), as well as others (27), such losses can be as high as 60-70%. Higher efficiency of electrotransfer onto PVDF membranes can be achieved by using polybrene-coated PVDF membranes (28) and optimizing the electrophoretic field, buffer concentration and methanol concentration in the transfer buffer (29,30). A further limitation of the electrotransfer step is the potential chemical damage to proteins, particularly tryptophan (note poor recovery of tryptophan in lysozyme and myoglobin fragments in Table 1). In our experience, destruction of tryptophan does not occur during electrophoresis *per se* (14) provided that

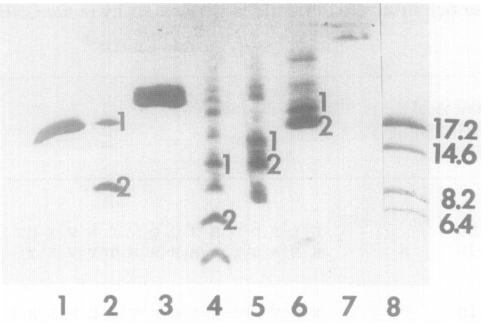

Figure 3 Tricine-SDS-PAGE peptide mapping of proteins isolated by 1D or 2D gel electrophoresis. Protein-containing gel slices were subjected to *in situ* CNBr treatment, neutralized, placed in the sample wells of a Tricine-SDS-PAGE gel and electrophoresed as described in Methods. Separated peptides were electroblotted onto PVDF membranes and stained with Coomassie blue R-250. The numbered peptides indicate those selected for sequence analysis (Table 1). Lane 1, myoglobin (20 µg), untreated; Lane 2, myoglobin (20 µg), CNBr-treated; Lane 3, lysozyme (20 µg), untreated; Lane 4, lysozyme (20 µg), CNBr-treated; Lane 5, S_1-glycoprotein, CNBr-treated; Lane 6, S_7-glycoprotein, CNBr-treated; Lane 7, 120 KD protein from Fig. 1; Lane 8, M_r markers derived by CNBr cleavage of myoglobin (Pharmacia).

appropriate precautions are taken (e.g. pre-electrophoresis with 1 mM thioglycollic acid), but during the electrotransfer step (31).

In this study it was interesting to note the dissimilarity in the CNBr fingerprints of the S_1- and S_7-glycoproteins (see Fig. 3). Both proteins are allelic forms of one gene, have similar M_r, isoelectric points, glycosylation patterns, and exert essentially the same biological function, i.e. rejection of self-pollen within the style (17). Although the dissimilarity in CNBr fragment pattern can be explained by incomplete fragmentation (as evidenced by two N-terminal fragments for S_7, Table 1), it would appear that the methionine pattern in S_1- and S_7-glycoprotein differs.

B The method presented here extends the usefulness of protein fingerprinting following *in situ* CNBr cleavage of proteins in SDS-polyacrylamide gels (32-34), since the fragments can be readily identified by direct sequence analysis. This approach should prove extremely useful in protein sequencing strategies since, historically, large CNBr fragments have proved difficult to separate by ion-exchange and reversed-phase HPLC.

ACKNOWLEDGMENTS

The work reported on the plant self-incompatibility glycoproteins is part of an on-going collaboration between the Plant Cell Biology Research Centre and the Joint Protein Structure Laboratory. The authors would like to thank Dr. M. Anderson and Professor A.E. Clarke for providing material and for valuable discussions.

REFERENCES

1. Simpson, R.J., Moritz, R.L., Begg, G.S., Rubira, M.R. and Nice, E.C. (1989) Anal. Biochem. 177, 221-236.
2. Shively, J.E., Paxton, R.J. and Lee, T.D. (1989) Trends Biochem. Sci. 14, 246-255.

3. Vandekerckhove, J., Bauw, G., Puype, M. Van Damme, J. and Montagu, M. (1985) Eur. J. Biochem. 152, 9-19.
4. Aebersold, R.H., Teplow, D.B., Hood, L.E. and Kent, S.B.H. (1986) J. Biol. Chem. 261, 4229-4238.
5. Eckerskorn, C., Mewes, W., Goretzki, H. and Lottspeich, F. (1988) Eur. J. Biochem. 176, 509-519.
6. Matsudaira, P. (1987) J. Biol. Chem. 262, 10035-10038.
7. Aebersold, R., Leavitt, J., Saevedra, R., Hood, L. and Kent, S.B.H. (1987) Proc. Natl. Acad. Sci. USA 84, 6970-6974.
8. Bauw, G., Van den Bulcke, M., Van Damme, J., Puype, M., Van Montagu, R.L.M. and Vandekerckhove, J. (1988) J. Prot. Chem. 713, 194-196.
9. Choli, T., Kapp, U. and Wittmann-Liebold, B. (1989) J. Chromatogr. 476, 59-72.
10. Simpson, R.J., Ward, L.D., Reid, G.E., Batterham, M.P. and Moritz, R.L. (1989) J. Chromatogr. 476, 345-361.
11. Kennedy, T.E., Gawinowicz, M.A., Barzilai, A., Kandel, E.R. and Sweatt, J.D. (1988) Proc. Natl. Acad. Sci. USA 85, 7008-7012.
12. Brown, T.A., Prahl, J.M. and De Luca, H.F. (1988) Proc. Natl. Acad. Sci. USA 85, 2454-2458.
13. Walsh, M.J., McDougall, J. and Wittmann-Liebold, B. (1988) Biochemistry 27, 6867-6876.
14. Ward, L.D., Reid, G.E., Moritz, R.L. and Simpson, R.J. In Current Research in Protein Chemistry (J.J. Villafranca, Ed.), Academic Press, (in press).
15. Eckerskorn, C. and Lottspeich, F. (1989) Chromatographia (in press).
16. Karwan, R. and Kamp, R-M. (1988) J. Prot. Chem. 7, 244-246.
17. Anderson, M.A., Cornish, E.C., Mau, S.L., Williams, E.G., Hoggart, R., Atkinson, A., Bonig, I., Grego, B., Simpson, R., Roche, P.J., Haley, J.D., Penshow, J.D., Niall, H.D., Tregear, G.W., Coghlan, J.P., Crawford, R.J. and Clarke, A.E. (1986) Nature 321, 38-44.
18. Jahnen, W., Batterham, M.P., Clarke, A.E., Moritz, R.L. and Simpson, R.J. (1989) The Plant Cell 1, 493-499.
19. Simpson, R.J., Begg, G.S., Dorow, D.S. and Morgan, F.J. (1980) Biochemistry 19, 1814-1819.
20. Simpson, R.J., Moritz, R.L., Van Roost, E. and Van Snick, J. (1988) Biochem. Biophys. Res. Commun. 157, 364-372.
21. Laemmli, U.K. (1970) Nature 227, 680-685.
22. O'Farrell, P.Z., Goodman, H.M. and O'Farrell, P.H. (1977) Cell 12, 1133-1142.
23. Schägger, H. and von Jagow, G. (1987) Anal. Biochem. 166, 368-379.
24. Simpson, R.J., Moritz, R.L., Rubira, M.R. and Van Snick, J. (1988) Eur. J. Biochem. 176, 187-197.
25. Begg, G.S. and Simpson, R.J. (1989) In Techniques in Protein Chemistry (T.E. Hugli, Ed.) pp. 79-88, Academic Press, New Nork.
26. Kapadia, G. and Chrambach, A. (1972) Anal. Biochem. 48, 90-102.
27. Tous, G.I., Fausnaugh, J.L., Akinyosoye, O., Lackland, H., Winter-Cash, P., Vitorica, F.J. and Stein, S. (1989) Anal. Biochem. 179, 50-55.
28. Xu, Q-Y. and Shively, J.E. (1988) Anal. Biochem. 170, 19-30.
29. Plough, M., Jensen, A.L. and Barkholt, V. (1989) Anal. Biochem. 181, 33-39.
30. Speicher, D.W. (1989) In Techniques in Protein Chemistry (T.E. Hugli, Ed.), pp. 24-35, Academic Press, New York.
31. Moos, M., Nguyen, N.Y. and Liu, T-Y. (1988) J. Biol. Chem. 263, 6005-6008.
32. Nikodem, V. and Fresco, J.R. (1979) Anal. Biochem. 97, 382-386.
33. Lonsdale-Eccles, J.D., Lynley, A.M. and Dale, B.A. (1981) Biochem. J. 197, 591-597.
34. Skopp, R.N. and Lane, L.C. (1988) Appl. Theoret. Electrophoresis 1, 61-64.