Journal of Chromatography, 266 (1983) 545–554
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 035

PURIFICATION OF CYANOGEN BROMIDE FRAGMENTS FROM β-2-GLYCOPROTEIN I BY HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

JAY LOZIER*, NOBUHIRO TAKAHASHI and FRANK W. PUTNAM
Department of Biology, Indiana University, Bloomington, IN 47405 (U.S.A.)

SUMMARY

Reversed-phase high-performance liquid chromatography (HPLC) with a linear elution gradient of 1-propanol in 0.1 % aqueous trifluoroacetic acid was used to purify cyanogen bromide fragments from the human plasma protein β-2-glycoprotein I. The fragments, ranging from 33 to 119 amino acids in length, were obtained in sufficient purity and yield for automated sequence analysis and enzymatic digestion. One fragment, which contained most of the carbohydrate, could be isolated as an aggregate or as two heterogeneous monomers. The elution of this fragment was much earlier than expected on the basis of its retention coefficient calculated from amino acid composition. Our results demonstrate the applicability of recently developed HPLC techniques to the separation of cyanogen bromide fragments from a carbohydrate-rich glycoprotein whose structure is not completely known.

INTRODUCTION

High-performance liquid chromatography (HPLC) is rapidly displacing traditional procedures for the purification of peptides for structural study of proteins[1]. With the increasing use of automated sequencing techniques the strategy for structural studies of proteins has shifted from sequencing numerous small peptides derived from complete enzymatic digests to sequencing fewer but larger fragments derived from partial enzymatic digests or chemical cleavage by reagents such as cyanogen bromide (CNBr) or N-bromosuccinimide. In many cases gel filtration alone is not sufficient for the complete purification of such fragments, and other methods such as ion-exchange chromatography must be used. These methods are often limited by the intrinsic hydrophobicity of large fragments which may require the use of chaotropic salts, extremes of pH or organic solvents to maintain solubility[2]; such measures may interfere with the detection of the fragments, or may necessitate extra steps to remove the added materials.

In our structural studies of β-2-glycoprotein I, a human plasma protein, we have utilized CNBr cleavage to obtain many small overlapping peptides whose sequence was determined in preliminary studies[3]. Previously we were hindered by the inability to purify large carbohydrate-containing peptides and fragments which are

0021-9673/83/$03.00 © 1983 Elsevier Science Publishers B.V.

clustered in the middle of the protein. Such peptides tend to exhibit broad, irregular elution profiles due to carbohydrate heterogeneity and are recovered in poor yield due to strong interactions with commonly used ion-exchange resins[3]. In work described here we have used gradient elution with 1-propanol in trifluoroacetic acid (TFA) as a counter-ion to purify all of the CNBr fragments from β-2-glycoprotein I, including two which contain carbohydrates. For this purpose a large-pore C_{18} reversed-phase HPLC column was used after preliminary fractionation by gel chromatography. The fragments, ranging in size from 33 to 119 amino acids in length, were obtained in sufficient yield for automated amino acid sequencing and subdigestion. Our results represent the extension of recently developed techniques[2,4] to the solution of a structural problem in a protein whose structure has not been reported.

EXPERIMENTAL

Chemical modification, CNBr cleavage and enzymatic digestion

Crystalline human β-2-glycoprotein I, obtained from Behringwerke (Marburg/Lahn, G.F.R.) was subjected to several criteria of purity; it appeared to be homogeneous except for a trace of possible aggregate detected by electrophoresis in polyacrylamide gel containing sodium dodecyl sulfate (SDS). The protein (200 mg) was reduced and alkylated with iodoacetic acid by the method of Crestfield *et al.*[5] except that dithiothreitol (Sigma, St. Louis, MO, U.S.A.) was used as the reducing reagent. The complete carboxymethylation of cysteines was verified by amino acid analysis. The carboxymethylated protein was subjected to cleavage with cyanogen bromide (Eastman, Rochester, NY, U.S.A.) in 70% formic acid for 48 h (CNBr/methionine ratio of 200:1)[6]. The CNBr fragments were purified as described below. The amino acid sequence of most of the fragments could be determined directly by automated Edman degradation, but the largest fragment (CB2) had to be further digested with trypsin (Worthington, Freehold, NJ, U.S.A.) after citraconylation of lysine residues[7]. Thermolysin (Calbiochem, San Diego, CA, U.S.A.) was used to digest the monomeric forms of CB3.

Purification of CNBr fragments

The mixture of CNBr fragments was first purified by gel chromatography on Sephadex G-75 (Pharmacia, Uppsala, Sweden) and then on a large-pore (300 Å) C_{18} reversed-phase HPLC column (Synchropak RP-P, 250 × 4.1 mm I.D.; SynChrom Inc., Linden, IN, U.S.A.). Tryptic subdigest peptides from the largest CNBr fragment were purified using a C_{18} reversed-phase HPLC column (Ultrasphere ODS, 250 × 4.6 mm I.D.; Altex, Berkeley, CA, U.S.A.). A Beckman 421 controller was used to programm a gradient of 1-propanol in water delivered by two Beckman 110 A solvent delivery pumps. An Altex 210 injection valve with a 250-μl sample loop was used to inject samples. Absorbance was monitored with a Gilson Holochrome monitor. 1-Propanol (glass-distilled) was obtained from Burdick & Jackson Labs. (Muskegon, MI, U.S.A.). Sequenal grade TFA was obtained from Pierce (Rockford, IL, U.S.A.) and added directly to both water and 1-propanol at a level of 0.1%.

Analysis of CNBr fragments

Molecular weights of the CNBr fragments were estimated by electrophoresis

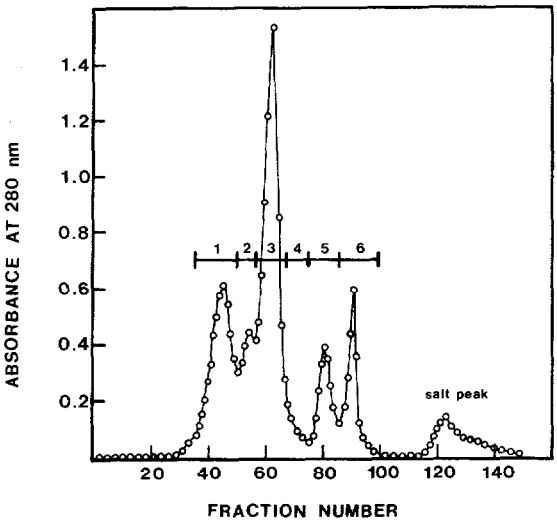

Fig. 1. Gel chromatography of CNBr cleavage fragments of β-2-glycoprotein I. The mixture of CNBr fragments was applied to a SF-Sephadex G-75 column (110 × 2.6 cm) and eluted with 10% formic acid at a flow-rate of 10 ml/h. Fractions of 5 ml were collected and absorbance at 280 nm was measured with a Beckman DU-7 spectrophotometer. Fractions were pooled as indicated by the bars numbered 1–6. The salt peak contains CNBr and 70% formic acid in which the protein was cleaved.

on a 12–30% polyacrylamide gradient gel system using 0.1% SDS and a Tris–glycine running buffer modified from Laemmli[8]. Estimation of relative amounts of material in gel bands was done with the use of a Helena Quick Scan R&D densitometer. The amino acid content of the CNBr fragments and subdigest peptides was determined by analysis of 24-h hydrochloric acid hydrolysates (110°C) on a Beckman 121 MB single-column amino acid analyzer. Glucosamine could be detected as a broad peak appearing between phenylalanine and histidine on the chromatogram. The amino acid sequence of CNBr fragments and subdigest peptides was determined by a Beckman 890C automated sequencer, using an 0.1 *M* Quadrol program with S1, S2 washes (Beckman program 121078), and Polybrene (Sigma) as a carrier[9]. Phenylthiohydantoin (PTH) derivatives were identified by C_{18} reversed-phase chromatography on a Hewlett-Packard 1084 A automated chromatograph, using gradients of sodium acetate–acetonitrile.

RESULTS AND DISCUSSION

The elution pattern of the initial gel chromatography of the CNBr fragments is shown in Fig. 1. From the gel electrophoresis pattern (Fig. 2) it is clear that gel filtration alone is inadequate for complete purification of the fragments. As an example, the main component of pool 3 is present in pools 1–4, indicating that it was eluted in a broad zone. Pools 1 and 2 were not purified further by HPLC due to the presence of high-molecular-weight aggregates and/or incomplete cleavage fragments.

The purification by HPLC of the CNBr fragments in pool 3 of Fig. 1, shown in Fig. 3, yielded two peaks, the first (I) of which was eluted at 22% 1-propanol. Gel electrophoresis of this peak (Fig. 3, inset) revealed a major component of approxi-

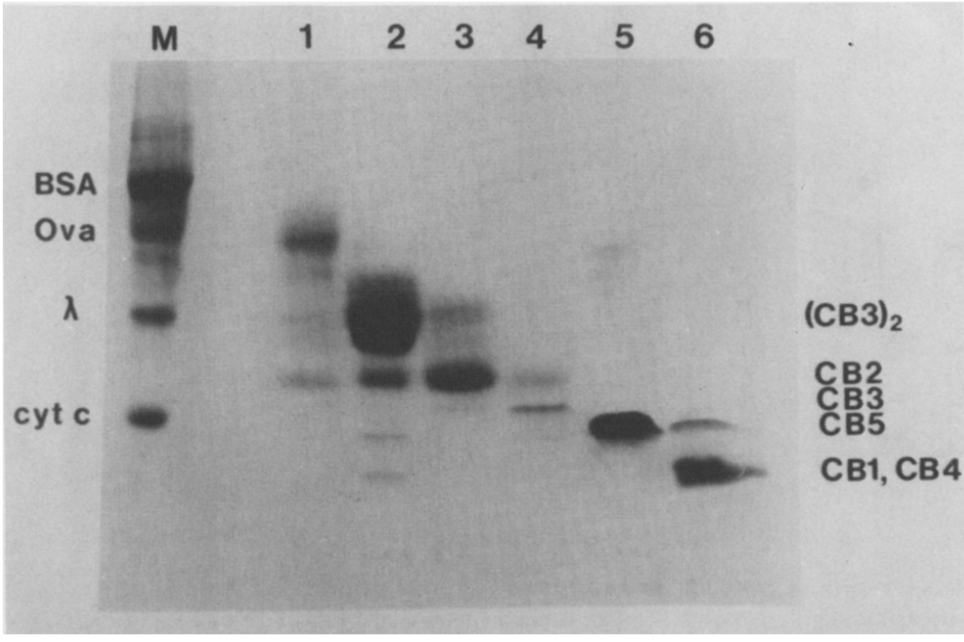

Fig. 2. Gel electrophoresis of the gel filtration pools from Fig. 1. M = Markers: BSA (bovine serum albumin, 67,000 daltons); Ova (ovalbumin, 43,000 daltons); λ (lambda light chain, 23,000 daltons) and cyt c (cytochrome *c*, 13,000 daltons). Lanes 1–6 = pools 1–6 from Fig. 1. The positions of the CNBr fragments are indicated in the right margin (see Fig. 4). Electrophoresis was performed in a 12–30% gradient of polyacrylamide with a 6% polyacrylamide stacking gel by the method of Laemmli[8]. Samples dissolved in the pH 6.8 sample buffer were electrophoresed in pH 8.3 Tris–glycine SDS buffer at 25 mA through the stacking gel and then at 50 mA.

mately 23,600 daltons and also a minor component of about 12,900 daltons; the latter accounted for only 6% of the material, as estimated by densitometric scanning of the gel. Amino acid analysis of the material in peak I gave a composition corresponding to CB3, the third CNBr fragment in the protein. The alignment of the CNBr fragments, which was determined by other procedures, is shown in Fig. 4. The amino acid sequence determined for this fragment is the same as that we found earlier for CB3 isolated by a similar procedure. Since the major fragment in peak I has a molecular weight about twice that expected for CB3, it probably represents a dimer (or aggregate) of CB3, and the minor component is the monomer form. This accords with the failure to detect any contaminant during sequence analysis of peak I, although the possibility that the minor component has a blocked amino terminus cannot be excluded. The basis of such dimer formation is unknown; it is unlikely that it results from disulfide linkage, since the protein was completely reduced and alkylated before cleavage. Furthermore, gel electrophoresis was done in the presence of the reducing reagent 2-mercaptoethanol which should break any disulfide bonds.

Peak II in Fig. 3, which was eluted by 30% 1-propanol, was shown by SDS electrophoresis to have a molecular weight of about 16,000 daltons. Amino acid analysis and sequence data indicated that peak II was the second CNBr fragment in the protein (CB2 in Fig. 4). In order to determine the amino acid sequence of CB2, the

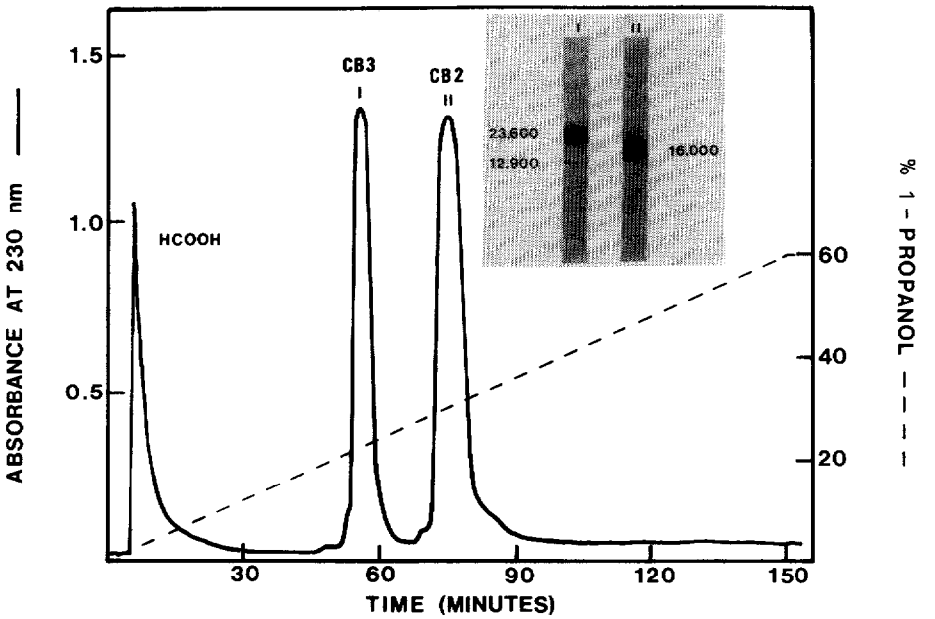

Fig. 3. C_{18} Synchropak RP-P HPLC of CNBr fragments from pool 3 of Fig. 1. The contents of pool 3 were dissolved in 10% formic acid–0.1% TFA and injected at time $T = 0$ into the HPLC system described in the Experimental section. A gradient of 0–60% 1-propanol was used to elute the fragments. Flow-rate was 0.7 ml/min. Inset: gel electrophoresis of the material from peaks I and II.

largest CNBr fragment, it was necessary to subdigest it with trypsin and separate the subdigest peptides by C_{18} reversed-phase HPLC, as shown in Fig. 5. The peptides recovered in the two overlapping peaks d and e both contained carbohydrate. The two peptides both had the amino acid sequence Val-Tyr-Lys-Pro-Ser-Ala-Gly(Asn)-Asn-Ser-Leu-Tyr-Arg, in which the glucosamine oligosaccharide is presumably attached to the first asparagine residue. Placement of the carbohydrate rests on the successful determination of sequence for all other residues in the peptide with only an Asx left in the composition which presumably is the asparagine of the Asn-X-Ser/Thr

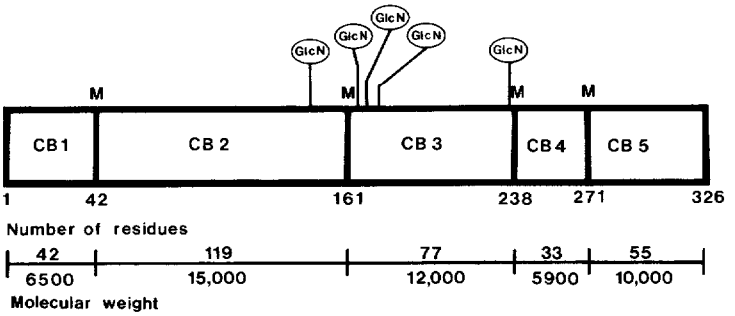

Fig. 4. Structural model of β-2-glycoprotein I. The alignment of the five CNBr fragments of the protein (CB1–5) is based on sequence analysis of peptides not described here. The position of methionine residues (M) and of glucosamine-containing oligosaccharides (GlcN) is shown at the top. The molecular weights of the fragments are those determined by gel electrophoresis (Fig. 2). The numbering of residues is tentative.

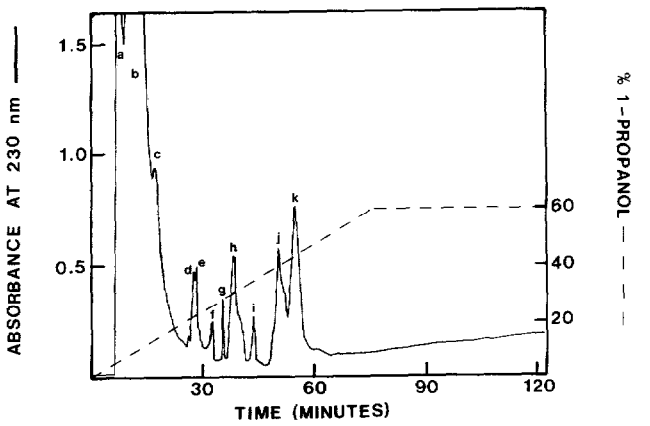

Fig. 5. C_{18} Ultrasphere ODS reversed-phase HPLC resolution of peptides from the tryptic subdigest of fragment CB2. CB2 (300 nmol) were citraconylated[7] and digested with trypsin (1%, w/w) for 6 h at 37°C in 0.1 *M* N-ethylmorpholine (pH 8.0). The lyophilized digest was redissolved in 10% formic acid–0.1% TFA and kept at 4°C overnight to allow hydrolysis of the citraconyl groups from the lysine residues to take place. Elution conditions as in Fig. 3.

acceptor sequence for glucosamine-containing oligosaccharides[10]. This resolution by HPLC of glycopeptides of identical amino acid composition may either indicate intrinsic carbohydrate heterogeneity or else extrinsic heterogeneity resulting as an artifact of the CNBr cleavage of the protein in strong acid (70% formic acid).

Fig. 6 shows the elution pattern of CNBr fragments from pool 4 of Fig. 1.

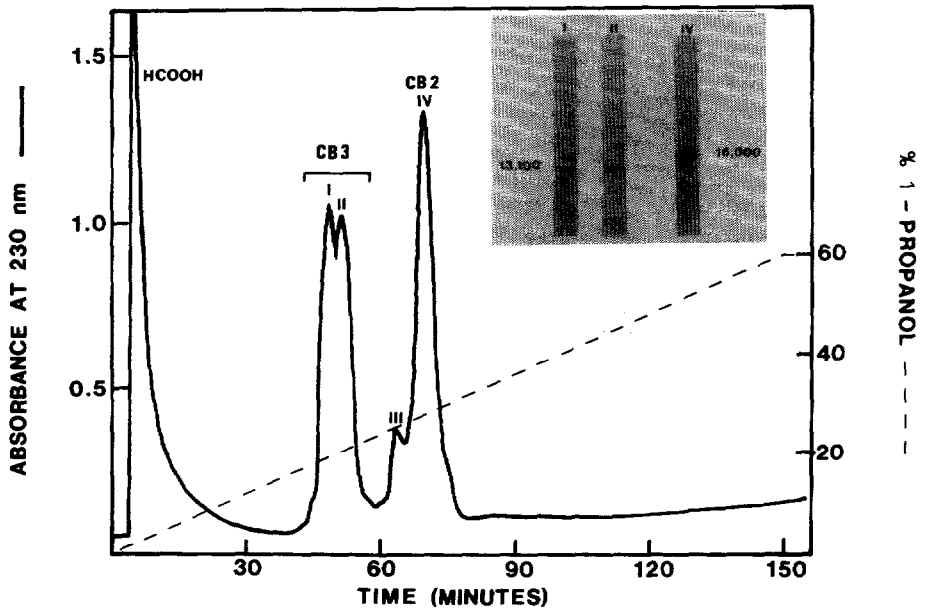

Fig. 6. C_{18} Synchropak RP-P HPLC purification of CNBr fragments from pool 4 of Fig. 1. Material from pool 4 of Fig. 1, dissolved in 10% formic acid–0.1% TFA, was chromatographed as described in Fig. 3. Inset: gel electrophoresis of material from peaks I, II and IV.

TABLE I

AMINO ACID CONTENT OF CNBr FRAGMENTS FROM β-2-GLYCOPROTEIN I

Values in parentheses are from sequence data. Values for glucosamine are given as moles glucosamine per mole of peptide. Serine values are corrected for 10% destruction during hydrolysis, and threonine values are corrected for 5% destruction during hydrolysis. Tryptophan values are uncorrected for destruction during hydrolysis. All values are from a 24-h hydrolysis in 6 *M* hydrochloric acid with two drops of 5% phenol added to prevent destruction of tyrosine. CMC = Carboxymethylcysteine; HSL = homoserine lactone.

Amino acid	*CB1 (peak III, Fig. 9)*	*CB2 (peak II, Fig. 3)*	*CB3 (peak I, Fig. 3)*	*CB4 (peak I, Fig. 9)*	*CB5 (peak III, Fig. 8)*
Aspartic acid	2.13 (2)	9.29 (9)	9.79 (11)	1.15 (1)	5.57 (6)
Threonine	3.52 (4)	11.22 (11)	7.85 (8)	1.02 (1)	2.70 (3)
Serine	2.78 (3)	6.80 (6)	3.06 (3)	1.92 (2)	4.65 (6)
Glutamic acid	3.77 (3)	7.91 (7)	5.73 (5)	4.47 (4)	5.10 (5)
Proline	5.88 (6)	12.67 (14)	6.29 (7)	1.86 (2)	2.08 (2)
Glycine	4.55 (5)	7.23 (7)	6.67 (7)	2.14 (2)	2.36 (2)
Alanine	0.22 (0)	8.36 (9)	3.03 (4)	2.00 (2)	2.88 (3)
CMC	3.58 (2)	8.24 (9)	5.36 (3)	2.12 (2)	6.97 (4)
Valine	2.45 (3)	5.51 (5)	2.30 (2)	4.29 (5)	2.57 (3)
HSL	+ (1)	+ (1)	+ (1)	+ (1)	− (0)
Isoleucine	1.06 (1)	6.28 (7)	1.72 (2)	0.95 (1)	1.86 (2)
Leucine	2.09 (2)	9.37 (9)	4.29 (4)	0.37 (0)	1.99 (2)
Tyrosine	2.84 (3)	5.40 (5)	4.20 (4)	1.12 (1)	1.00 (1)
Phenylalanine	2.09 (2)	7.04 (7)	4.13 (4)	1.08 (1)	3.72 (4)
Lysine	3.23 (3)	5.52 (5)	6.64 (6)	6.42 (7)	8.05 (9)
Histidine	0.14 (0)	1.27 (1)	2.16 (2)	0.10 (0)	1.79 (2)
Arginine	1.66 (2)	4.98 (5)	2.20 (2)	1.01 (1)	0.00 (0)
Tryptophan	0.00 (0)	0.65 (2)	0.63 (2)	0.00 (0)	0.30 (1)
Glucosamine	0.00	2.28	5.76	0.00	0.00
Total	42	119	77	33	55
Yield (%)	51.8	22.8	21.7	31.4	27.7

Peaks I and II of Fig. 6, which are eluted close together by 19.4 and 20.4% 1-propanol, respectively, both show a predominant 13,100-dalton component on electrophoresis (inset Fig. 6). In both cases the amino acid composition (Table I) was that expected for CB3 of Fig. 4, the CNBr fragment that was earlier eluted as a dimer (peak I of Fig. 3). The two 13,000-dalton fragments (peaks I and II of Fig. 6) were subdigested with thermolysin, and the peptides were separated by C_{18} reversed-phase HPLC. Thermolysin was chosen to subdigest the fragments since other enzymes such as trypsin would not cleave the peptide in the region of the three clustered oligosaccharides. As can be seen in Fig. 7, the elution profiles of the digests were similar though not identical, suggesting heterogeneity of the attached carbohydrates or possibly of the protein.

The electrophoresis and amino acid composition of the peptide in peak IV of Fig. 6 indicate that it was CB2, the largest CNBr fragment, which previously had been purified as peak II of Fig. 3 and which appears in several pools in Fig. 2.

Purification by HPLC of pool 5 of Fig. 1 gave a principal component as peak

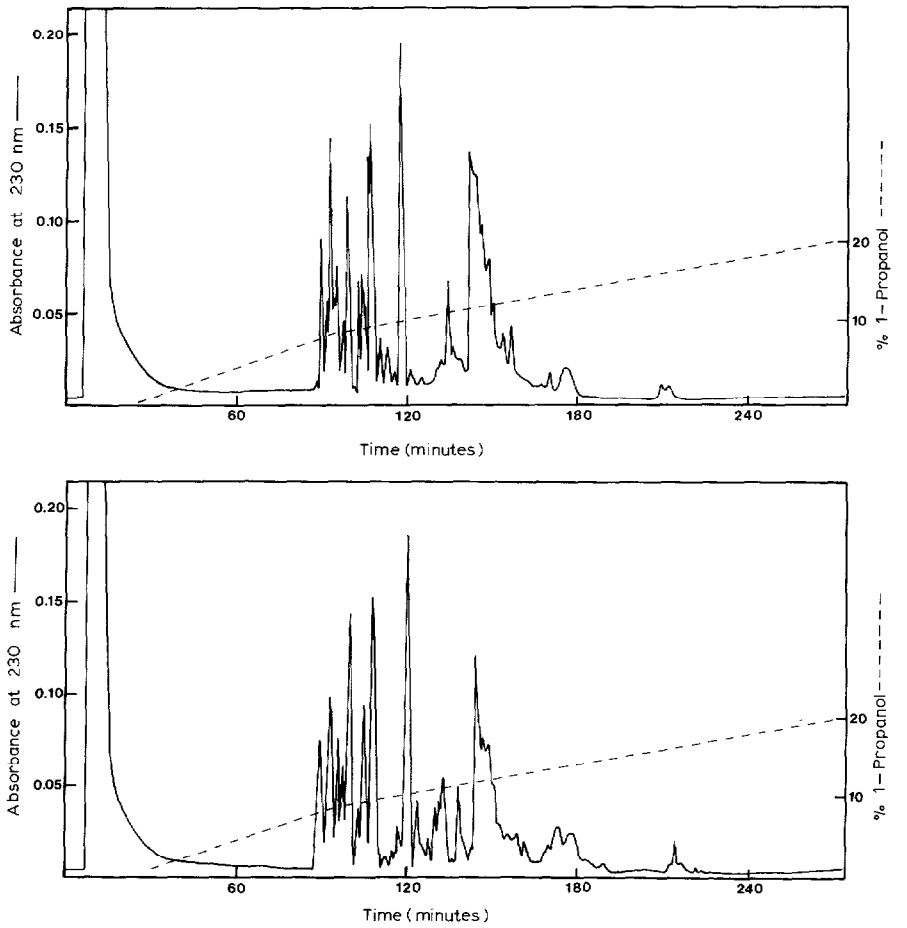

Fig. 7. C_{18} Ultrasphere ODS reversed-phase HPLC resolution of thermolysin subdigest peptides of peaks I and II of Fig. 6. The contents of peaks I and II of Fig. 6 were subdigested with thermolysin (2% w/w) for 12 h at 37°C in 0.1 *M* N-ethylmorpholine (pH 8.0)–1 m*M* calcium chloride. Each digest was chromatographed on the HPLC system described in the Experimental section. The peptides were eluted by 20 min of isocratic flow (0% 1-propanol), a 60-min 0–10% gradient of 1-propanol and a 150-min 10–20% gradient of 1-propanol. Flow-rate was 0.7 ml/min. Above: elution profile of peptides from digest of peak I. Below: elution profile of peptides from digest of peak II.

III of Fig. 8. The molecular weight of 11,500 for peak III as given by electrophoresis (inset, Fig. 8) and the lack of homoserine lactone in the amino acid composition suggested that this is the carboxyl-terminal fragment, CB5 (Fig. 4); this was confirmed by amino acid sequence analysis.

The last pool to be purified (pool 6 of Fig. 1) showed two major components and a minor one in its elution profile (Fig. 9). The first major component, which was eluted by 24% 1-propanol (peak I in Fig. 9) was apparently pure as judged by electrophoresis (inset, Fig. 9). The amino acid composition (Table I) and amino acid sequence indicate its identity as CB4, the penultimate CNBr fragment in the protein. The second major component, which was eluted by 28.8% 1-propanol (peak III, Fig. 9), is the amino-terminal CNBr fragment (CB1) and was contaminated by a trace of the minor component represented by peak II. The minor component appeared to be the carboxyl-terminal fragment (CB5) as judged by electrophoresis, and probably represents spillover of this fragment from pool 5 into pool 6.

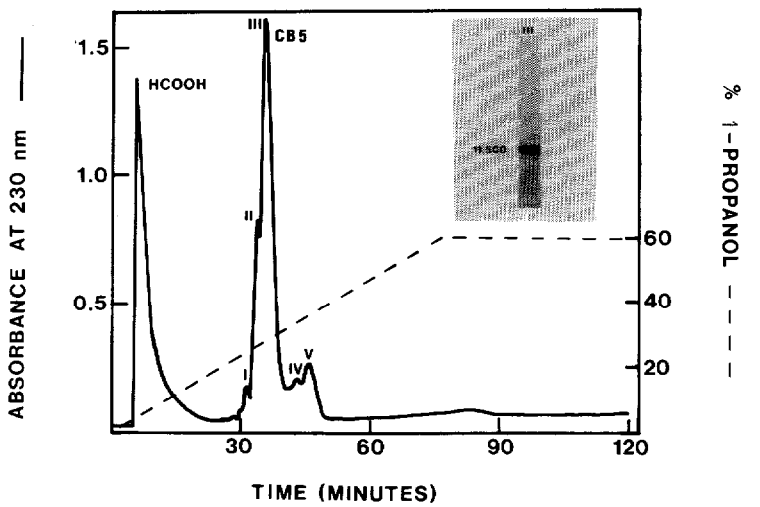

Fig. 8. C_{18} Synchropak RP-P HPLC of the main CNBr fragment from pool 5 of Fig. 1. The material in pool 5 of Fig. 1 was redissolved in 10% formic acid–0.1% TFA and chromatographed on the HPLC system described in the Experimental section. Elution conditions as in Fig. 3. Inset: gel electrophoresis of the material in peak III.

Our results indicate that HPLC systems used by other investigators[2,4] for model proteins of known structure can be successfully applied to the separation of large polypeptide fragments from a protein whose structure is not known. The protein selected, β-2-glycoprotein I also illustrates the problem of separating large, carbohydrate-rich fragments which tend to form aggregates and have therefore been difficult to purify[3]. The basis for such aggregate formation is unknown, but the phenomenon has also been observed for another carbohydrate-rich glycoprotein, the pregnancy-specific β-1-glycoprotein[11]. The latter protein contains 33% carbohydrate and in its native state forms dimers and trimers that dissociate only slowly even in the

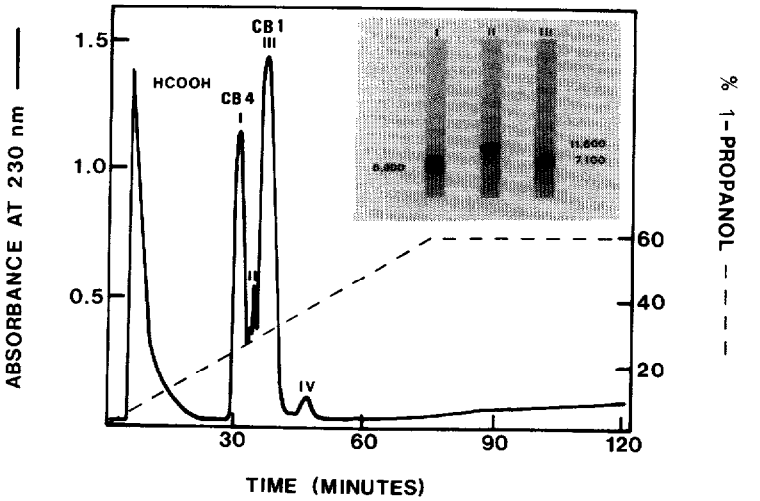

Fig. 9. C_{18} Synchropak RP-P HPLC purification of CNBr fragments from pool 6 of Fig. 1. Material from pool 6 of Fig. 1 was redissolved in 10% formic acid–0.1% TFA and chromatographed as described in Fig. 8. Inset: gel electrophoresis of material from peaks I, II and III.

TABLE II

ELUTION ORDER OF CNBr FRAGMENTS OF β-2-GLYCOPROTEIN I FROM A C_{18} SYNCHROPAK RP-P HPLC COLUMN

Elution order	*CNBr fragment*	*% 1-Propanol*	*Molecular weight*	*Retention coefficient*[12]
1	CB3 monomer	19.4, 20.4	13,000	40.61
	CB3 dimer	22.0	23,600	40.61
2	CB4	24.0	6000	16.10
3	CB5	27.9	11,500	26.88
4	CB1	28.8	7100	26.18
5	CB2	30.1	16,000	79.78

presence of strong denaturing agents. It is possible that pools 1 and 2 from our initial gel chromatography (Figs. 1 and 2) may contain similar aggregates of the carbohydrate-rich CNBr fragments, CB2 and CB3. This may explain the apparent formation of dimers of CB3.

The order of elution of CNBr fragments of β-2-glycoprotein I from the Synchropak RP-P reversed-phase HPLC column is given in Table II. It is much as expected on the basis of calculated retention coefficients[12] with the exception of CB3, which was eluted much earlier than expected. This anomalous elution may be explained by the presence of its four carbohydrate groups, shown in Fig. 4, which greatly increase its hydrophilicity and would thus decrease its elution time in reversed-phase HPLC.

Our results demonstrate that methods developed for the separation of peptides from model proteins can be applied successfully to the separation of CNBr fragments from a protein whose complete structure was not known. In addition, we have shown that these HPLC techniques are useful for the purification of carbohydrate-rich fragments that we previously could not purify by other methods such as gel filtration and ion-exchange chromatography[3].

ACKNOWLEDGEMENTS

We appreciate the technical assistance of P. Davidson, S. Dorwin, J. Dwulet, K. Huss, J. Madison and Y. Takahashi.

This work was supported by National Institutes of Health grant AM 19221.

REFERENCES

1 M. Elzinga (Editor), *Methods in Protein Sequence Analysis*, Humana Press, Clifton, NJ, 1982, p. 383.
2 W. C. Mahoney and M. A. Hermodsen, *J. Biol. Chem.*, 255 (1982) 11,199.
3 V. Liu, unpublished results.
4 J. D. Pearson, W. C. Mahoney, M. A. Hermodsen and F. E. Regnier, *J. Chromatogr.*, 207 (1981) 325.
5 A. M. Crestfield, S. Moore and W. H. Stein, *J. Biol. Chem.*, 238 (1963) 622.
6 E. Steers, Jr., G. R. Craven, C. B. Anfinsen and J. L. Bethune, *J. Biol. Chem.*, 240 (1965) 2478.
7 A. F. S. A. Habeeb and M. Z. Atassi, *Biochemistry*, 9 (1970) 4939.
8 U. K. Laemmli, *Nature (London)*, 227 (1970) 681.
9 D. G. Klapper, C. E. Wilde and J. D. Capra, *Anal. Biochem.*, 85 (1978) 126.
10 W. Lennarz (Editor), *The Biochemistry of Glycoproteins and Proteoglycans*, Plenum Press, New York, 1980.
11 J. C. Osborne, Jr., S. W. Rosen, B. Nilsson, I. Calvert and H. Bohn, *Biochemistry*, 21 (1982) 5523.
12 R. F. Rekker, *The Hydrophobic Fragmental Constant*, Elsevier, Amsterdam, 1977, p. 301.