HUMAN α_1-ANTITRYPSIN: CARBOHYDRATE ATTACHMENT AND SEQUENCE HOMOLOGY

R. W. CARRELL, J.-O. JEPPSSON*, L. VAUGHAN, S. O. BRENNAN, M. C. OWEN and D. R. BOSWELL

Molecular Pathology Laboratory, Pathology Department, Christchurch Hospital, Christchurch, New Zealand and **Department of Clinical Chemistry, University of Lund, Malmö General Hospital, S-214 01 Malmö, Sweden*

Received 27 October 1981

1. Introduction

Human α_1-antitrypsin (α_1-AT) has 3 carbohydrate sidechains [1]. The attachment point of one of these sidechains has been determined in conjunction with the amino acid sequence of the C-terminal third of the molecule [2]. We now provide further sequence data which define the attachment points of the other 2 carbohydrate sidechains. These data also show the extent of the sequence homology with antithrombin-III and ovalbumin, and provide support for a single reactive centre situated near the C-terminus.

2. Methods

Human α_1-AT was isolated from plasma by thiol exchange [3]. Cyanogen bromide fragments were prepared and isolated as in [4]. Automated sequence analysis was carried out in a Beckman 890 C sequenator using the 1 M quadrol program. PTH derivatives were identified by high-performance liquid chromatography and by back hydrolysis with HI.

Subpeptide digestion was carried out as in [5] using trypsin, chymotrypsin, thermolysin or *Staphylococcus aureus* V8 protease. Enzyme to substrate ratios of 1:30 were used with digestion at pH 8.6 in 0.1 M NH_4HCO_3 buffer. Peptide isolation and dansyl-Edman sequence determinations were performed as in [5].

Standard tryptic digestion of heat denatured α_1-AT yielded the central and C-terminal glycopeptides but not the N-terminal glycopeptide centred on Asn 46. This was obtained by denaturation of α_1-AT in 6 M guanidine hydrochloride at 70°C for 2 h, followed by tryptic digestion in 2 M guanidine hydrochloride in 0.1 M NH_4HCO_3 buffer, pH 8.0 at 37°C for 15 h with an enzyme to substrate ratio of 1:10.

The glycopeptides were separated from the bulk of the peptides by gel filtration on Sephadex G-50 in 0.05 M NH_4HCO_3. Ion-exchange chromatography on DEAE-A25 Sephadex in 0.02 M Tris–HCl (pH 8.1) buffer with a linear gradient of 0–0.5 M NaCl resolved the 3 tryptic glycopeptides. A final step of affinity chromatography on concanavalin A–Sepharose 4B was performed as described for the pronase glycopeptides in [6] to remove minor peptide contaminants. This step also gave an indication of the ratio of biantennary to triantennary oligosaccharide chains for each glycopeptide.

Sequences were aligned for maximum homology using the Needleman–Wunsch algorithm with the MDM_{78} score matrix [7]; gap penalty and matrix bias parameters were 100 and 60, respectively. The statistical significance of homology was tested by comparison with match scores for 100 randomisations of the sequences generated by selection with replacement.

3. Results and discussion

Analysis of the individual glycopeptides and of the cyanogen bromide fragments of the whole molecule confirmed the finding [1] that α_1-AT has only 3 carbohydrate sidechains. The evidence obtained for the attachment point of each of these sidechains is summarised in fig.1. The additional sequence provided by the isolation of the glycopeptides along with [2] gives a provisional sequence of some 90% of the residues of human α_1-AT. This allows placement of the carbohydrate attachment points within the molecule (fig.2).

Abbreviations: α_1-AT, α_1-antitrypsin; PTH, phenylthiohydantoin

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.75

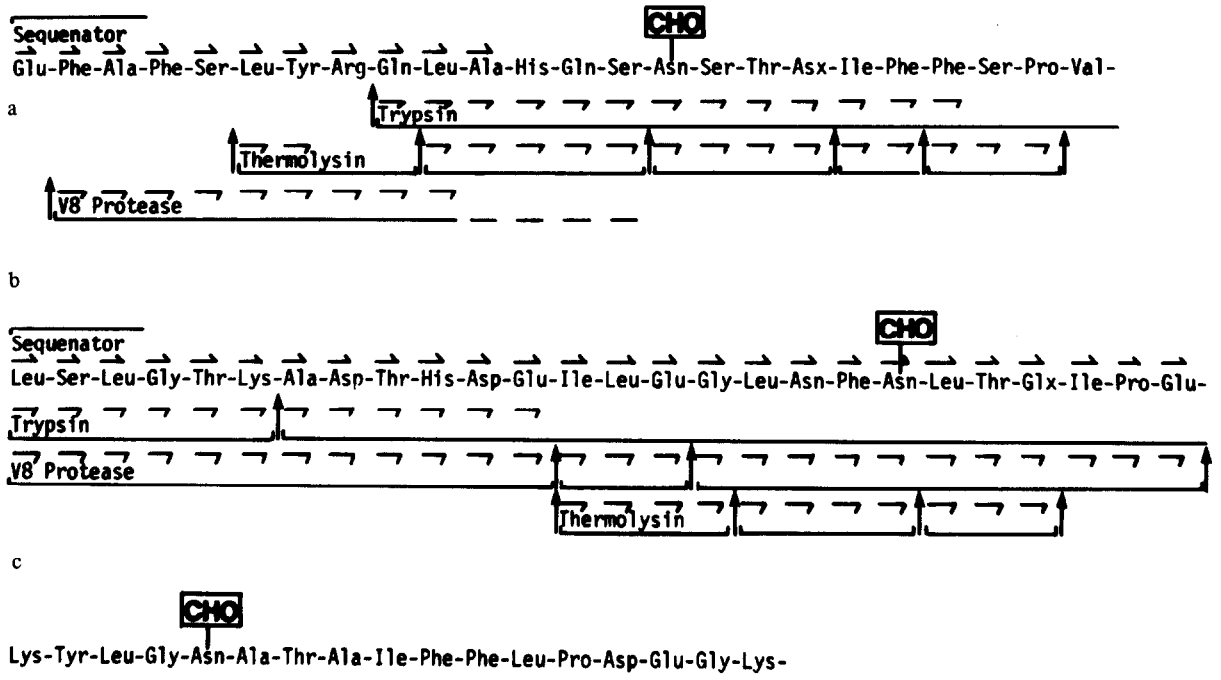

Fig.1. Sequence of the glycopeptides of α_1-antitrypsin. The peptides are shown in order of placement from the N-terminus. Data for the C-terminal glycopeptide (c) appeared in [2]. The position of the peptides in the whole molecule is shown in fig.2.

Glu-Asp-Pro-Gln-Gly-Asp-Ala-Ala-Gln-Lys-Thr-Asp-Thr-Ser-His-His-Asp-Gln-
Asp-His-Pro-Thr-Phe-Asn-Lys-Ile-Thr-Pro-Asn-Leu-Ala-Glu-Phe-Ala-Phe-Ser-
Leu-Tyr-Arg-Gln-Leu-Ala-His-Gln-Ser-[Asn]-Ser-Thr-Asx-Ile-Phe-Phe-Ser-Pro-
Val-Ser-..(4)..-Phe-Ala-Met-Leu-Ser-Leu-Gly-Thr-Lys-Ala-Asp-Thr-His-Asp-
Glu-Ile-Leu-Glu-Gly-Leu-Asn-Phe-[Asn]-Leu-Thr-Glu-Ile-Pro-Glu-Ala-Glx-Ile-
His-Glu-Gly-Phe-Glu-Glu-Leu-Leu-Arg-Thr-Leu-Asp-Gln-Pro-Asx-Ser-Gln-....
(23)-Phe-Leu-Glu-Asx-Val-His-Lys-Leu-Tyr-His-Ser-Glu-Ala-Phe-Thr-Val-Asx-
Phe-Gly-Asx-Thr-Glx-Glu-Ala-His-Lys-Glx-Ile-XXX-Asx-Tyr-Val-Glx-Lys-Gly-
Thr-Gln-Gly-Lys-..(14)..-Ala-Leu-Val-Asn-Tyr-Ile-Leu-Phe-Lys-Gly-Lys-Trp-
Glu-Arg-Pro-Phe-Glu-Val-Lys-Asp-Thr-Glu-Glu-Glu-Asp-Phe-His-Val-Asp-Gln-
Val-Thr-Thr-Val-Lys-Val-Pro-Met-Met-Lys-Arg-Leu-Gly-Met-Phe-Asn-Ile-Gln-
His-Cys-Lys-Lys-XXX XXX-Ser-Trp-Val-Leu-Leu-Met-Lys-Tyr-Leu-Gly-[Asn]-Ala-
Thr-Ala-Ile-Phe-Phe-Leu-Pro-Asp-Glu-Gly-Lys-Leu-Gln-His-Leu-Glu-Asn-Glu-
Leu-Thr-His-Asp-Ile-Ile-Thr-Lys-Phe-Leu-Glu-Asn-Glu-Asp-Arg-Arg-Ser-Ala-
Ser-Leu-His-Leu-Pro-Lys-Leu-Ser-Ile-Thr-Gly-Thr-Tyr-Asp-Leu-Lys-Ser-Val-
Leu-Gly-Gln-Leu-Gly-Ile-Thr-Lys-Val-Phe-Ser-Asn-Gly-Ala-Asp-Leu-Ser-Gly-
Val-Thr-Glu-Glu-Ala-Pro-Leu-Lys-Leu-Ser-Lys-Ala-Val-His-Lys-Ala-Val-Leu-
Thr-Ile-Asp-Glu-Lys-Gly-Thr-Glu-Ala-Ala-Gly-Ala-Met-Phe-Leu-Glu-Ala-Ile-
Pro-Met-Ser-Ile-Pro-Pro-Glu-Val-Lys-Phe-Asn-Lys-Pro-Phe-Val-Phe-Leu-Met-
Ile-Glu-Gln-Asn-Thr-Lys-Ser-Pro-Leu-Phe-Met-Gly-Lys-Val-Val-Asn-Pro-Thr-
Gln-Lys

Fig.2. Carbohydrate attachment sites: 90% of the sequence of human α_1-AT is provided by the 4 polypeptide fragments above, positioned by homology with antithrombin-III and ovalbumin. The composition but not the sequence of the intervening peptides has been determined. Their approximate lengths, as deduced from homologies are indicated by the bracketed figures. Carbohydrate attachment sites are boxed. The reactive centre is underlined. Sequence and compositional analyses exclude a similar centre near the N-terminus.

A check on the positioning of the attachment points was possible because of the close amino acid sequence homology of α_1-AT with antithrombin-III and ovalbumin [2,8,9]. This homology had been demonstrated for the C-terminal portion of the molecule but did not include the initial N-terminal fragment. However, the extension of sequence of the N-terminal fragment shows that a strong homology does commence after the 30th residue. The alignment of the first 30 residues of α_1-AT with the first 56 residues of antithrombin-III gave a match score below the mean for the randomised same-composition, i.e., there is no homology. However, alignment of residues 31–48 of α_1-AT with residues 57–75 of antithrombin gave an alignment score for the match of 3.24 giving $p < 0.001$ that the similarity was due to chance alone. Visual inspection of the aligned sequence of α_1-AT, antithrombin-III and ovalbumin confirms that the homologies extend from residue 31 of α_1-AT to the C-terminus. The homologous sequences allow the carbohydrate attachment points of α_1-AT to be precisely positioned in alignment with residues Asp 72, Asp 109 and Asp 269 of antithrombin-III.

The structure of the carbohydrate sidechains shows

some variation. These results support the findings [1] that the predominant structure is a biantennary sidechain but the central attachment site may be occupied by a triantennary carbohydrate sidechain. This variation in carbohydrate structure, mainly at the central attachment point, gives rise to isoforms that explain the observed microheterogeneity of α_1-AT [6].

A number of points could be made in relation to the overall sequence of the protein but these can await the publication of the definitive structure of human α_1-AT. However, a striking feature at this stage is the support given for the placement [5] of the reactive centre as being 36 residues from the C-terminus (fig.2). The sequence gives no support for the claimed existence of a similar site near the N-terminus [10–13].

Acknowledgements

This investigation was supported by grants from the Swedish Medical Research Council (project no. B81-13X-581-14C) and the Medical Research Council of New Zealand. Dr A. Yoshida and colleagues kindly provided a cyanogen bromide pentapeptide fragment of human α_1-antitrypsin.

References

[1] Mega, T., Lugan, E. and Yoshida, A. (1980) J. Biol. Chem. 255, 4057–4061.
[2] Carrell, R., Owen, M., Brennan, S. and Vaughan, L. (1979) Biochem. Biophys. Res. Commun. 91, 1032–1037.
[3] Laurell, C.-B. (1978) J. Chromatogr. 159, 25–31.
[4] Jeppsson, J.-O. (1976) FEBS Lett. 65, 195–197.
[5] Carrell, R. W., Boswell, D. R., Brennan, S. O. and Owen, M. C. (1980) Biochem. Biophys. Res. Commun. 93, 399–402.
[6] Vaughan, L. and Carrell, R. (1981) Biochem. Intl. 2, 461–467.
[7] Schwartz, R. M. and Dayhoff, M. O. (1978) in: Atlas of Protein Sequence and Structure. vol. 5, suppl. 3, p. 354, National Biomedical Research Foundation, Washington DC.
[8] Peterson, T. E., Dudek-Wojciechowska, Sottrup-Jensen, L. and Magnusson, S. (1979) in: The Physiological Inhibitors of Blood Coagulation and Fibrinolysis (Collen, D. et al. eds) pp. 43–54, Elsevier/North-Holland, Amsterdam, New York.
[9] Hunt, Lois T. and Dayhoff, Margaret O. (1980) Biochem. Biophys. Res. Commun. 95, 864–871.
[10] Johnson, D. A. and Travis, J. (1978) J. Biol. Chem. 253, 7142–7144.
[11] Kress, L. F., Kurecki, T., Chan, S. K. and Laskowski, M. sr (1979) J. Biol. Chem. 254, 5317–5320.
[12] Martodam, R. R. and Liener, I. E. (1981) Biochim. Biophys. Acta 667, 328–340.
[13] James, H. L. and Cohen, A. B. (1978) J. Clin. Invest. 62, 1344–1353.