THE AMINO-TERMINAL SEQUENCE OF HUMAN PROTEIN HC

Blas FRANGIONE*, Edward C. FRANKLIN**, Anders GRUBB+, and L. TEJLER+

*Departments of Pathology **and Medicine, Irvington House Institute, New York University Medical Center, 550 First Avenue, New York, New York 10016, USA and +Department of Clinical Chemistry, University of Lund, Malmo General Hospital, S-214 01, Malmo, Sweden*

Received 17 September 1976

1. Introduction

Protein HC (human complex-forming glycoprotein, heterogeneous in charge) is a recently discovered glycoprotein present in normal plasma, urine and cerebrospinal fluid [1]. It forms complexes with other plasma proteins; the IgA-complex being the quantitatively predominant one. The protein has a tendency to polymerize and has been shown to be present on the surface of a large percentage of normal human peripheral blood lymphocytes [2]. On SDS–polyacrylamide electrophoresis, the carbohydrate-rich protein (22% carbohydrate) forms a single band and has an apparent molecular weight of about 30 000. Agarose gel electrophoresis and isoelectric focusing have demonstrated a marked charge heterogeneity of the glycoprotein despite the fact that it was purified from a single individual. Complete desialylation of the glycoprotein did not diminish its charge heterogeneity [1]. The present study demonstrates that protein HC has a unique NH_2-terminal amino acid sequence up to residue 25.

2. Materials and methods

Protein HC was purified from the urine of a single individual (JL) suffering from tubular proteinuria by high-pressue ultrafiltration and ion-exchange chromatography followed by gel filtration and immunoabsorption as earlier described [1]. It was pure as judged by SDS–polyacrylamide electrophoresis and by immunoelectrophoresis. An immunochemically closely related protein has recently been isolated from pooled urines by a different procedure [1,3]. For amino acid analyses, the protein was hydrolyzed at 110°C for 20 h with six normal HCl containing 0.5% phenol in sealed tubes. Quantitative amino acid analyses were performed on the Durrum Amino Acid Analyzer Model D-500 (Durrum Instruments Corp., Palo Alto, Calif.) N-terminal residue was determined by the dansyl-chloride method and identification was done on 15 × 15 cm polyamide thin layer plates with standards plus samples on one side and only samples on the other [4,5]. The plates were run in two dimensions in three different solvent mixtures [5]. Automated amino acid sequence analysis was carried out on a Beckman 890C Sequencer using the Beckman DMAA peptide program. Identification of amino acids was generally done by three methods. Gas–liquid chromatography (GLC) was performed with a 7620A Hewlett-Packard gas chromatograph packed with Beckman coated support No. 56796 [6]. Thin layer chromatography (TLC) of PTH-amino acids was done on 5 × 5 cm polyamide plates in two dimensions [7]. Standards were placed on one side and samples on the other. The amines of Asp or Glu were generally discriminated by TLC. The remainder of the aliquot was hydrolyzed under vacuum with hydroiodic acid (HI) and identified by amino acid analysis [8].

3. Results and discussion

Table 1 shows the amino acid composition of purified protein HC from patient JL. Its amino terminal was glycine. Figure 1 shows the amino acid

Table 1
Amino acid composition[a] of a urinary protein HC[b] from patient JL

Lys	10.51
His	4.27
Arg	10.03
Asp	15.79
Thr	16.31
Ser	9.90
Glu	23.53
Pro	13.00
Gly	15.81
Ala	10.51
Half Cysteine	2.98
Val	11.60
Met	4.96
Ile	11.97
Leu	11.93
Tyr	8.98
Phe	6.77
Trp[c]	N.D.

[a] Compositions are reported as mol of amino acid per mol of protein, based on a molecular weight of 30 000 (minus 22% for carbohydrate)
[b] Mean of two hydrolyses – 20 h (one oxidized and one unoxidized)
[c] N.D. = Not done.

sequence of the first 25 residues from the NH_2-terminal of protein HC purified from the individual JL. Residues at positions 5 and 16 were not identified. As can be seen the amino acid sequence of the NH_2-terminal of protein HC does not show heterogeneity. Therefore the charge heterogeneity of protein HC from single individuals cannot be due to polymorphism of this part of the polypeptide chain but must be due to variation in other parts of the molecule. Computer search by Dr M. Dayhoff failed to show any significant homology of the NH_2-terminal

```
1                                   10
Gly-Pro-Val-Pro-(   )-Pro-Pro-Asp-Asn-Ile

                                    20
Gln-Val-Gln-Glu-Asx-( )-Phe-Ile-Ser-Arg

                25
Ile-Tyr-Gly-Arg-Trp
```

Fig.1. Amino acid sequence of the amino-terminal end of protein HC.

sequence of protein HC to any known plasma or cell membrane glycoprotein [9]. Work is in progress to determine the complete amino acid sequence of the protein and to localize its carbohydrate groups in order to determine the reasons for its intra-individual charge heterogeneity and to detect possible homologies with other proteins further on in the molecule. We have recently shown [2, and unpublished results] that protein HC is present on the surface of the majority of normal human B- and T-lymphocytes as well as on the surface of a varying percentage of several B- and T-cell lines. Moreover, the glycoprotein shows capping on the surfaces of these cells. Hence it would be of interest if this molecule were to have homologies with other surface proteins of lymphoid cells.

Acknowledgements

This work was supported by USP HS grants No. AM 01431, AM 02594, and the Irvington House Institute; and from the Medical Faculty, University of Lund, the Swedish Medical Research Council (project no. B76-13X-00581-12B) and the Alfred Osterlunds Foundation. We would like to thank Mr Rostyslaw Kutny for expert technical assistance.

References

[1] Tejler, L. and Grubb, A. O. (1976) Biochim. Biophys. Acta 439, 82–94.
[2] Tejler, L., Grubb, A. O. and Turesson, I. (1976) Acta Med. Scand. 199, 425–427.
[3] Ekstrom, B., Peterson, P. A. and Berggard, I. (1975) Biochem. Biophys. Res. Commun. 65, 1427–1433.
[4] Gray, W. R. (1967) Method. Enzymol. 11, 469–474.
[5] Hartley, B. S. (1970) Biochem. J. 119, 805–822.
[6] Pisano, F. F. and Bronzert, T. J. (1969) J. Biol. Chem. 244, 5597–5607.
[7] Summer, M. R., Smythers, G. W. and Oroszlan, S. (1973) Anal. Biochem. 53, 624–628.
[8] Smithies, O., Gibson, D. M., Fanning, E. M., Goodfliesh, R. M., Gilman, F. G. and Ballantyne, D. L. (1971) Biochemistry 10, 4912–4921.
[9] Dayhoff, M. O. (1973) in: Atlas of Protein Sequence and Structure, Volume 5.