Vol. 88, No. 1, 1979 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 14, 1979 Pages 258-265

# MICROHETEROGENEITY OF SERUM TRANSFERRIN, HAPTOGLOBIN AND $\alpha_2$HS GLYCOPROTEIN EXAMINED BY HIGH RESOLUTION TWO-DIMENSIONAL ELECTROPHORESIS

N. L. Anderson and N. G. Anderson

Molecular Anatomy Program
Division of Biological and Medical Research
Argonne National Laboratory
Argonne, Illinois 60439

Received April 5, 1979

SUMMARY: Most human serum proteins consistently appear as more than one spot when analyzed by high resolution two-dimensional electrophoresis. We have characterized the "micro-patterns" corresponding to several of the principal proteins. Three minor forms of transferrin are observed (one larger than the predominant form and probably a precursor) and each shows carbohydrate heterogeneity similar to the main form. The haptoglobin β-chain exhibits stepwise heterogeneity in both pI (sialic acid variation) and SDS molecular weight (neutral sugar structure attachment), and is shown to give rise to a smaller, likewise heterogeneous product. The two frequent alleles of $\alpha_2$HS glycoprotein are shown to differ predominantly in sialic acid content rather than polypeptide charge, suggesting a sequence difference at a carbohydrate attachment site. All of the expected sources of microheterogeneity (sialic acid, proteolysis, allelism) are observed, plus an SDS molecular weight microheterogeneity probably due to variation in attachment of large neutral sugar structures.

While the human plasma proteins have been the subject of numerous biochemical investigations, from studies of biosynthesis to complete sequencing in many cases, little is known about the origins of the so-called "microheterogeneity" revealed in them by high-resolution analytical techniques. Three causes are generally considered responsible for most microheterogeneity: (a) variation in sialic acid content, leading to charge heterogeneity in glycoproteins, (b) proteolysis leading to size heterogeneity in the affected polypeptide, and (c) genetic polymorphism, in which the expression of more than one gene results in the appearance of multiple forms. In some cases, careful work has revealed the nature of observed microheterogeneity; plasminogen, for instance, has been found to undergo proteolysis *in vivo* to yield two forms having different n-terminal amino acids, each of which is also heterogeneous with respect to carbohydrate



0006-291X/79/090258-08$01.00/0

content (1). Such investigations have been difficult, however, due to the fact that most analytical techniques reveal perhaps one or two of the three types of heterogeneity mentioned above, but not all three. This paper describes the use in this context of two-dimensional electrophoresis, in which the first (isoelectric focusing) dimension is sensitive to charge (pI) variation, while the second (SDS electrophoresis) is sensitive to size variation (2-5). In conjunction with genetic analysis and neuraminidase digestion, this technique has enabled us to characterize partially the microheterogeneity in a number of proteins, three of which will be discussed here.

## MATERIALS AND METHODS

Serum proteins were isolated using specific antibodies (Behring Diagnostics) bound to protein A Sepharose (Pharmacia). Typically one volume of swollen protein A Sepharose (20 µl) was used in a small Eppendorff pipette tip column to bind the antibody in 5-10 volumes of antiserum (100-200 µl) and this, after washing with phosphate buffered saline, was used to bind the antigen (desired protein) in 5 volumes (100 µl) of human serum. Neuraminidase digestion was performed, when required, by incubating the Sepharose in 0.5 mg/ml Cl. perfringens neuraminidase (Boehringer Mannheim) at pH 7, 37°C for one hour. Twenty µl of 1% SDS, 5% mercaptoethanol, 10% glycerol, 0.1 M cyclohexylaminoethane sulfonic acid (CHES) buffer, pH 9.3 at 90°C was then used to elute all non-covalently bound proteins; the eluting solution was applied after the column had been spun dry (5 minutes at 3,000 rpm in the Beckman TJ-6 centrifuge) and was collected by a similar centrifugation into a clean tube.

Samples were analyzed directly using the procedures described in detail in refs. 3-5. Briefly, these consist of SDS, mercaptoethanol denaturation followed by tube gel isoelectric focusing in 8 M urea, 2% NP-40 for 20 hr at 400 V, followed by SDS electrophoresis in a 10-20% gradient acrylamide slab gel. All gels shown were stained with Coomassie Brilliant Blue.

Gels are shown with the acid end to the left and SDS electrophoresis proceeding downwards, so that removal of sialic acid residues (negative charges) causes incremental shifts to the right. The isoelectric point for one spot in each figure is given relative to a set of 37 carbamylation product spots (charge shift analogues) of rabbit muscle creatine phosphokinase as described in ref. 6; these are internal pI markers. The term "SDS-molecular weight" is used here to emphasize the fact that mobility in SDS electrophoresis is not a direct measure of true molecular weight (particularly for glycoproteins). Values given for SDS-molecular weight in figures are determined from high resolution rat heart standards (7) and are for comparative purposes within the figures.

## RESULTS AND INTERPRETATION

### Transferrin

The pattern of spots comprising human serum transferrin is shown in Fig. 1a. On gels run at relatively low protein loading, a basic complex of four major and

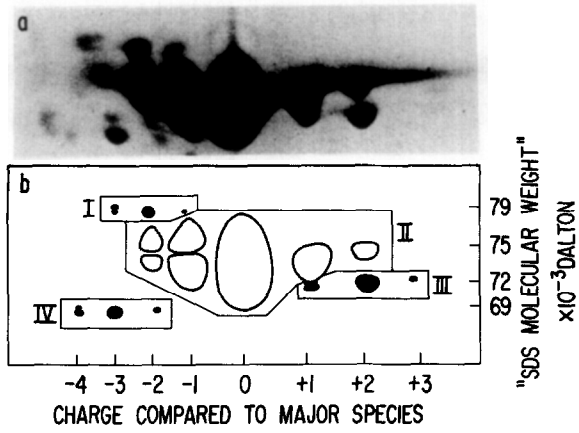


Fig. 1. Small region of a two-dimensional gel showing spots corresponding to human serum transferrin isolated by binding to specific antibody on Protein A Sepharose. Isoelectric focusing dimension is horizontal with acid end to the left; SDS electrophoresis proceeds downwards. (a) Photograph of gel loaded with approximately 20 µg transferrin. (b) Diagram illustrating pattern as observed over a range of loadings. The minor complexes (I, III, IV) reproduce the structure of the main complex (II) but are approximately 40-fold less abundant. Neuraminidase digestion (not shown) indicates that the main spot, at position 0 (= -4.0 RCPK), bears four sialic acid residues removable by neuraminidase.

three minor spots is visible (Fig. 1b; open ellipses). As the loading is increased, three additional complexes become visible; one above and two below the main complex. By running gels over a wide range of loading, these complexes are observed to have the same structure as the main complex. The small complexes are present at 2-5% of the abundance of the major complex, and constitute authentic transferrin since they are co-isolated with the major complex by immunoabsorption.

Neuraminidase digestion (data not shown) indicates that the charge heterogeneity within the basic complex is due to variation in sialic acid content. It seems likely that the small SDS molecular weight heterogeneity in the basic complex is due to addition of neutral carbohydrate, but this would be difficult to prove. The main point, however, is that the process that generates the three minor complexes is of a different kind, since it results in an overall movement of the basic complex pattern without perturbing its internal structure. The later effect is the expected result when different types of heterogeneity are superimposed. The most likely explanation for the inter-complex charge and

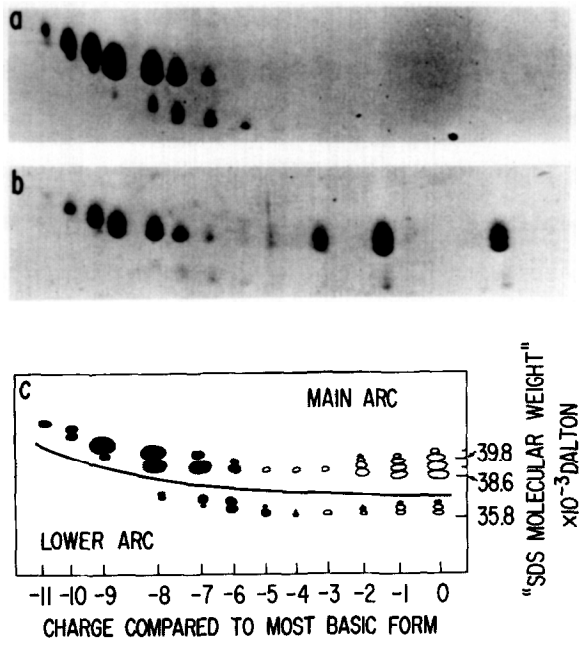


Fig. 2. Region of a two-dimensional gel showing (a) human haptoglobin β-chain (bound by specific antibody), (b) native β-chain plus products of extensive neuraminidase digestion, (c) diagram illustrating relationship of major and minor arc (filled ellipses) and their neuraminidase digested forms (open ellipses). Minor arc reproduces structure of major arc, but is shifted down (∿ 2800 SDS-daltons) and two charges to the right (+2); neuraminidase digested forms, however, are matched directly above one another as far as desialation proceeds. Micro-SDS-molecular weight heterogeneity in the form at position 0 exhibits increments of ∿ 600 SDS-daltons. Position -8 = -17.8 RCPK.

molecular weight shifts is proteolytic removal of short (1-3,000 dalton) peptides from one or both ends of the transferrin molecule starting from the top (highest molecular weight) complex. This scheme requires that minor complex I actually be the larger precursor of the predominant circulating form.

Haptoglobin β-chain

The normal pattern for the haptoglobin β-chain (Fig. 2a) consists of a main curved arc of spots separated horizontally by single charge shifts and a minor arc of precisely similar structure displaced downwards by ∿ 3,000 SDS-daltons and horizontally by precisely two positive charges.

Within the main arc (whose structure is typical of a number of plasma and cell membrane glycoproteins), the charge heterogeneity is due to the presence of sialic acid; neuraminidase digestion shifts all species towards a single

isoelectric point at the right (Fig. 2b). The minor arc is also shifted by neuraminidase digestion, but it converges toward the same pI as the main arc instead of towards a point two charge shifts to the right of it. It is clear therefore that the isoelectric point difference between major and minor arcs is due to different sialic acid content instead of a difference in polypeptide charge. Since the SDS molecular weight difference between arcs is almost certainly due to proteolytic removal of a 3-5,000 dalton peptide, one can also deduce that this peptide is neutral near the protein pI and that it contains a carbohydrate attachment site normally bearing two sialic acid residues (plus unknown neutral sugars). This is consistent with sequence data showing a carbohydrate attachment site at ASN 23 (i.e., within 3,000 daltons of the $NH_2$ terminus, ref. 8).

The most intriguing feature is, however, the SDS molecular weight microheterogeneity remaining in the main arc protein after neuraminidase treatment (2 b,c). The four spots differ by apparently equal increments of about 600 SDS-daltons. A number of reasons lead us to believe that these species differ only in the number of neutral sugar structures attached to the protein. First, the number of forms (four) agrees fairly well with the reported presence of 3-4 carbohydrate attachment sites (9). Second, the neuraminidase digested minor arc shows only three spots, consistent with the removal of one carbohydrate attachment site by the aforementioned proteolytic cleavage. Third, the number of spots visible in the neuraminidase-digested rat haptoglobin β-chain is only two, consistent with the absence of at least one of the carbohydrate binding sites in the human sequence (8). Lastly, the only other plausible explanation (proteolytic removal of several short, equal peptides from one end of the protein) is untenable, since this type of N- or C-terminal heterogeneity would certainly have been seen in sequencing studies but was not (10).

$\alpha_2$HS glycoprotein

Two allelic forms of $\alpha_2$HS glycoprotein exist (N and L; ref. 2) with approximate gene frequencies of 0.36 and 0.64 in the North American population (N. L.

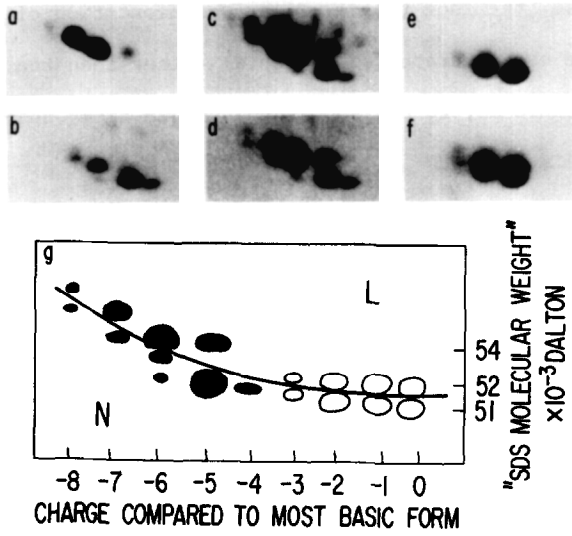


Fig. 3. Sections of gels showing forms of human serum $\alpha_2$HS glycoprotein (bound by specific antibody): (a) type L, (b) type N, (c) mixture of N and L, (d) heterozygote NL, (e) neuraminidase digested L, (f) neuraminidase digested mixture NL, (g) diagram illustrating relationship of various forms (filled ellipses are native spots, empty ellipses are spots resulting from neuraminidase digestion). The charge difference between native N and L (1 to 2 charges) disappears upon neuraminidase digestion (both end up at position 0). A molecular weight difference of approximately 1,000 SDS-daltons separates desialated N and L; this could be due to addition of a neutral sugar structure. Position -5 = -31.5 RCPK.

Anderson and J. V. Neel, unpublished observations). These forms are each heterogeneous (Figs. 3 a,b), and differ from each other in both SDS molecular weight and pI (1-2 charges). The charge difference is, however, abolished when sialic acid is removed with neuraminidase (3 e,f). Since, therefore, the alleles differ in charged carbohydrate (sialic acid), it is reasonable to postulate that the small SDS molecular weight difference is due to neutral carbohydrate bound to one and not the other; the two sorts of electrophoretic difference between alleles could then be attributed to a single protein sequence difference eliminating a sugar attachment site in one of them. Alternative explanations are that the L (smaller) form has suffered a small deletion near one end of the polypeptide (the deleted region bearing a carbohydrate attachment site), that only the L form is susceptible to proteolytic cleavage, or that the genetic difference observed occurs in an $\alpha_2$HS-specific glycosyl trans-

ferase. In any case, the fact that a primary manifestation of the allelic difference lies in the carbohydrate (sialic acid) rather than the polypeptide portion of the molecule is, to our knowledge, unique among non-blood group polymorphic systems.

## DISCUSSION

Are the micro-patterns described here truly a reflection of protein-modifying processes occurring in vivo? We believe so, for a number of reasons. Each pattern is extremely reproducible, and can be observed in gels of whole serum, plasma, immunoprecipitates from serum, and in classically purified proteins. So far, there is no reason to suspect that the analytical technique itself generates any heterogeneity. Proteins that have been thoroughly characterized as having only one form (prealbumin or haptoglobin $\alpha^{1F}$) show only a single spot, and as would be expected, any glycoprotein known to contain sialic acid shows charge heterogeneity. The primary reason why two-dimensional electrophoresis reveals so much microheterogeneity is sheer resolution; this in turn allows detection of minor components (1-5% of the total) which would not ordinarily be seen.

The three generally recognized types of protein microheterogeneity (sialic acid, proteolytic, allelic) are observed in the molecules described here. It appears that a fourth type must be added, however; the stepwise SDS-molecular weight heterogeneity observed in the haptoglobin β-chain. We believe this to be due to the hit-or-miss attachment of large, uniform, probably prefabricated neutral sugar structures (11) to the protein resulting in shifts of 500-1000 SDS-daltons per structure. Further work with specific endoglycosidases should allow testing of this hypothesis.

## ACKNOWLEDGMENT

We wish to thank Barbara J. Hickman for excellent technical assistance. This work was supported by U.S. Department of Energy.

## REFERENCES

1. Robbins, K. C., Boreisha, I. G., Arzadon, L., Summaria, L., and Barlow, G. H. (1975) J. Biol. Chem. 250, 4044-4047.

2. O'Farrell, P. H. (1975) J. Biol. Chem. 250, 4007-4021.
3. Anderson, N. L. and Anderson, N. G. (1977) Proc. Natl. Acad. Sci. USA 74, 5421-5425.
4. Anderson, N. G. and Anderson, N. L. (1978) Anal. Biochem. 85, 331-340.
5. Anderson, N. L. and Anderson, N. G. (1978) Anal. Biochem. 85, 341-354.
6. Anderson, N. L. and Hickman, B. J. (1979) Anal. Biochem., in press.
7. Anderson, N. G. and Anderson, N. L. (1979) Proc. Behring Inst., in press.
8. Barnett, D. R. and Kurosky, A., personal communication quoted by F. W. Putnam, "The Plasma Protein," Vol. II, F. W. Putnam, ed., Academic Press, p. 29.
9. Kurosky, A., Barnett, D. R., Rasco, M. A., Lee, T.-H., Hay, R. E., Kim, H.-H., Touchstone, B., and Bowman, B. H. in "Atlas of Protein Sequence and Structure," Vol. 5, Suppl. 2, M. O. Daykoff, Ed., National Biomedical Research Foundation, Washington, D.C. USA (1976).
10. Barrett, D. R., Lee, T.-H., and Bowman, B. H. (1972) Biochemistry 11, 1189-1200.
11. Lennarz, W. J. (1975) Science 188, 986-991.