Partial Amino Acid Sequence in the *N*-Terminal Region of an Anti-Pneumococcal Immunoglobulin Heavy Chain of Allotype a_2

By JEAN-CLAUDE JATON and JOSEPH HAIMOVICH*
Basel Institute for Immunology, Basel, Switzerland

(Received 19 December 1973)

The amino acid sequence of the *N*-terminal 48 residues of the heavy chain derived from a homogeneous rabbit antibody to type III pneumococci is described. This chain of allotype a_2 is compared with other rabbit heavy chains of allotypes a_1 , a_2 and a_3 . Within the *N*-terminal 25 positions, two chains which carry the same allotype a_2 possess identical amino acid sequences, but differ markedly from heavy chains of allotypes a_1 and a_3 . Sequence variability is observed in residues 26–27 and 30–34, but not in residues 35–48.

Unlike immunoglobulins of all other species so far investigated, rabbit immunoglobulins G, A, M and E carry allotypic specificities which are controlled by genes at a single locus with three alleles a_1 , a_2 and a_3 (Prah *et al.*, 1970). These three allelic forms correlate with extensive sequence variations within the *N*-terminal 33 positions of the H chains† (Wilkinson, 1969) and possibly within residues 80–85 (Mole *et al.*, 1971); recent findings obtained with homogeneous antibodies suggest, however, that no simple correlation exists between the identity of a single amino acid residue and the allotype in residues 80–94 (Jaton *et al.*, 1973). Sequence information on rabbit antibody heavy chains is limited mainly because of technical difficulties inherent in the sequencing of heterogeneous pools. The availability of homogeneous antibodies overcomes this inconvenience.

The present report describes the partial amino acid sequence of a heavy chain of allotype a_2 derived from a homogeneous antibody (designated K-25) directed against type III pneumococcal polysaccharide.

Materials and methods

Antibody K-25 was isolated as a single component by preparative electrophoresis in agarose from the serum of a rabbit hyperimmunized with type III pneumococcal vaccine (Braun & Jaton, 1973). It carries a_2b_4 allotypic markers. The separation of H and L chains, as well as the isolation of the *N*-terminal half of the heavy chain (fragment C1),

from a CNBr digest of the H chain, and of the *N*-terminal 33-residue peptide from fragment C1 after complete reduction and alkylation with iodo-[2- 14 C]acetic acid, were performed as described by Jaton & Braun (1972).

Enzymic digestions, isolation of peptides by preparative high-voltage paper electrophoresis, amino acid analyses and the determination of the sequences of small peptides by the dansyl-Edman procedure were performed as described by Jaton (1974). The analysis of the sequence of large fragments was done with a Beckman 890B protein sequencer, by using a volatile buffer instead of Quadrol and 4-sulphophenyl isothiocyanate (Braunitzer *et al.*, 1970). Details of the technique and of the identification of the phenylthiohydantoin derivatives are given elsewhere (Jaton, 1974).

Results

The *N*-terminal half of the H chain K-25 (CNBr fragment C1) was subjected to ten automated Edman degradations. Two different residues in comparable yield were detected in most of the Edman-degradation steps, suggesting the presence of two distinct sequences, as reported for a few other homogeneous anti-(polysaccharide H) chains (Jaton *et al.*, 1973). Extended homology with known sequences from rabbit H chains (Mole *et al.*, 1971; Jaton *et al.*, 1973) suggests that fragment C1 of H chain K-25 was cleaved between residues 33–34 and 79–80. Thus fragment C1 is likely to yield three fragments on cleavage with CNBr, accounting for residues 1–33, 34–79 and 80–253. The first and third fragments are assumed to be held together by a disulphide bridge between residues 22 and 92 (O'Donnell *et al.*, 1970). The second peptide (residues 34–79), which is free of half-cystine residues, was isolated from fragment C1 by gel filtration of Sephadex G-100 in 5M-guanidine

* Present address: Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot, Israel.

† Abbreviations: H and L chains, heavy and light chains; CmCys, [Glu (in amino acid sequences), *S*-carboxymethylcysteine and pyrrolidonecarboxylic acid respectively.

	1	5	10	15	20	24
a ₁ BS-5	[Glu.....Ser (Val, Glu)Glu-Ser-Gly-Gly-Arg-Leu-Val-Thr-Pro-Thr-Pro-Gly-Leu-Thr-Leu-Thr-Cys-Thr-Val-Ser-Ala					
a ₁ Pool	Leu					
a ₂ R-2690	Lys—Glu—Gly—Phe—Lys—Asn—Thr					
a ₂ K-25	Lys—Glu—Gly—Phe—Lys—Asp—Thr					
a ₃ Pool	Gln—Leu—Asp—Val—Lys—Gly—Ala—Ser—Ala					
				(?)		
	25	30	35	40	45	48
a ₁ BS-5	Gly-Phe-Ser-Leu-Ser-Tyr-Asp-Met-Gly-Trip-Val-Arg-Gln-Ala-Pro-Gly-Lys-Gly-Leu-Glu-Trip-Ile-Gly-Tyr					
a ₁ Pool	Ala-(?)—Ser					
a ₂ R-2690	Ile (Asp—) Gly-Val-Ser—Asp					
a ₂ K-25	Gly—Ser—					
a ₃ Pool	Asn-Gly—Tyr—Phe (?)					

Fig. 1. Comparison of the N-terminal 48 residues of rabbit heavy chains of allotypes a₁, a₂ and a₃

The structure of the N-terminal region of H chain K-25 (allotype a₂) comprises the sequence of the N-terminal 33-residue peptide and the partial sequence of the peptide containing residues 34-79 determined by automated Edman degradation (100 nmol). Sequences of a₁ and a₃ normal pooled H chains are from Wilkinson (1969) and Mole *et al.* (1971), that of anti-(type III pneumococci) H chain BS-5 (allotype a₁) is from Jaton & Braun (1972) and that of anti-(group C streptococcal carbohydrate) H chain R-2690 (allotype a₂) is from Fleischman (1971). The solid line indicates identical sequence to that shown immediately above. In a₁ and a₃ pools, amino acids placed above or below the line at some positions in the sequence are alternative residues.

Fig. 2. Amino acid sequence of the N-terminal fragment of H chain K-25 (allotype a₂)

T-Ch stands for tryptic-chymotryptic peptides. The subscript N indicates that the peptide is neutral at pH 6.5. The mobility of the peptide at pH 6.5 (relative to Asp = -1.0) is given in parentheses after the peptide designation. —, Residues identified by the dansyl-Edman procedure. =, Identified as dansyl residue without hydrolysis. —, By carboxypeptidase A or B. *, Cysteine-containing peptides detected by radioautography. The position of CmCys was determined by counting the radioactivity of the dansyl derivative.

hydrochloride (elution volume relative to exclusion volume = 2.0). It was recovered in low yield (25%), as it interacts strongly with the other two fragments by non-covalent forces (Jaton *et al.*, 1973). Its *N*-terminal sequence was determined in the protein sequencer; only the first 15 residues are shown in Fig. 1 (residues 34–48 of H chain K-25).

The *N*-terminal 33-residue peptide of the H chain was isolated in 55% yield and subjected to tryptic-chymotryptic digestion (250nmol). Ten peptides were separated by electrophoresis at pH 6.5 and/or 3.5. Their sequence is presented in Fig. 2. Only two lysine residues were accounted for in the 33-residue peptide: as one of the two is adjacent to a proline residue in peptide T- Ch_{N_4} (Fig. 2), it was expected that tryptic cleavage would be restricted to the other lysine residue located at position 4 from the *N*-terminus. The *N*-terminal fragment (250nmol) was thus split with trypsin and the whole digest was subjected directly to automated sequence analysis. A single sequence of 20 residues was determined starting with the residue located in position 5 of rabbit H chains, since the *N*-terminal tetrapeptide T- Ch_{N_2} is blocked (Fig. 2). This sequence overlaps the large tryptic-chymotryptic peptides T- Ch_5 , T- Ch_1^* and T- Ch_2^* and establishes their order. The determination of the sequence of the *N*-terminal 33-residue peptide is summarized in Fig. 2.

Discussion

On the basis of the poor recovery of the peptide containing residues 34–79, it cannot be excluded that H chain K-25 is heterogeneous; however, in view of the relatively high yield of the *N*-terminal 33-residue peptide (55%) and of the unique amino acid sequence of the variable region of the corresponding L chain (J.-C. Jaton, unpublished work), it is likely that the partial sequence reported here is representative of all molecules present in the preparation of H chain K-25.

The comparison of the amino acid sequence of the *N*-terminal 48 residues of H chains of allotypes a_1 , a_2 and a_3 is shown in Fig. 1. Although of different allotype, the two anti-(type III pneumococci) H chains BS-5 and K-25 differ beyond residue 16 by only two amino acid residues, in positions 30 and 34, a section which encompasses the first hypervariable region of the H chain possibly involved in antigen-binding specificity. R-2690 H chain, derived from an anti-(group C streptococci) antibody (Fleischman, 1971), exhibits three interchanges in residues 30–34 compared with K-25 and BS-5 chains.

Comparison of the two a_2 chains R-2690 and K-25

with a_1 and a_3 pooled chains reveals linked amino acid substitutions in positions 4, 7, 9, 11–12 and 15–16. Chains R-2690 and K-25 differ from each other by three interchanges within the first 29 residues, in positions 15, 26 and 27. The interchange Asn/Asp in position 15 might be attributed to technical difficulties. As residues 26–29 of H chain K-25 (allotype a_2) are identical with those present in the corresponding positions of a_1 chains, it seems unlikely that these residues are allotype-related. On the basis of the present data, allotypically related differences in the *N*-terminal 33-residue fragment would appear to occur within the first 16 residues. Alternatively, the identification of slightly differing sequences in the *N*-terminal fragment of two H chains expressing the same allotype might be explained by the results of Kindt *et al.* (1973), who demonstrated that homogeneous antibodies appear to possess only a portion of the spectrum of the group *a* allotypic specificities present in total immunoglobulins G. This observation may thus reflect amino acid sequence variations in the variable region of H chains of a single allotype. If this is true, the identification of the structural correlates of the group *a* specificities may be complicated by the presence of several antigenic variants.

The section residues 35–48 is rather constant, since only one amino acid substitution is found in R-2690 H chain (Asp-42). Extensive homologies with H chains from man, mouse and guinea pig have been noted (Jaton & Braun, 1972).

The authors thank Dr. D. G. Braun for helpful discussions and Mr. W. Hunkeler for valuable technical assistance.

- Braun, D. G. & Jaton, J. C. (1973) *Immunochemistry* **10**, 387–395
- Braunitzer, G., Schrank, B. & Ruhfus, A. (1970) *Hoppe-Seyler's Z. Physiol. Chem.* **351**, 1589–1590
- Fleischman, J. B. (1971) *Biochemistry* **10**, 2753–2761
- Jaton, J. C. (1974) *Biochem. J.* in the press
- Jaton, J. C. & Braun, D. G. (1972) *Biochem. J.* **130**, 539–546
- Jaton, J. C., Braun, D. G., Strosberg, A. D., Haber, E. & Morris, J. E. (1973) *J. Immunol.* **111**, 1838–1843
- Kindt, T. J., Seide, R. K., Tack, B. F. & Todd, C. W. (1973) *J. Exp. Med.* **138**, 33–43
- Mole, L. E., Jackson, S. A., Porter, R. R. & Wilkinson, J. M. (1971) *Biochem. J.* **124**, 301–318
- O'Donnell, I. J., Frangione, B. & Porter, R. R. (1970) *Biochem. J.* **116**, 261–268
- Prahl, J. W., Mandy, W. J., David, G. S., Steward, M. W. & Todd, C. W. (1970) *Protides Biol. Fluids Proc. Colloq.* **17**, 125–130
- Wilkinson, J. M. (1969) *Biochem. J.* **112**, 173–185