Primary sequence differences between Chido and Rodgers variants of tryptic C4d of the human complement system

Ulf Hellman, Gösta Eggertsen, Åke Lundwall*, Åke Engström and John Sjöquist

Department of Medical and Physiological Chemistry, PO Box 575, S-751 23 Uppsala, Sweden

Received 13 March 1984

Human tryptic C4d of the Chido and Rodgers variant was fragmented by cyanogen bromide and trypsin. The fragments were characterized by amino acid analysis and sequence determination. Polymorphism between the two genetic variants was detected in 5 positions. Four were closely located (residues 141, 142, 145, 146), where Leu, Ser, Ile, His occurred in the Chido variant and Pro, Cys, Leu, Asp in the Rodgers variant, respectively. In position 94 Gly was found in Chido and Asp in Rodgers. Alignment of the fragments was performed and it is concluded that tryptic C4d of both variants contains 346 residues.

Complement *Human C4d* *Sequence* *Polymorphism* *Blood group*

1. INTRODUCTION

The complement factor C4 is a serum protein consisting of three disulfide linked polypeptide chains with M_r 90000 (α-chain), 72000 (β-chain) and 32000 (γ-chain) [1,2]. As a constitutive part of the classical pathway of the complement system, C4, when activated, forms the classical C3-convertase in combination with C2 (review [3]). In serum C4 is degraded by factor I and C4-binding protein (C4bp) to C4c (M_r 140000) and C4d (M_r 50000), of which the latter contains the labile binding site [4,5]. C4 is coded for by two separate loci, C4A and C4B, located within the major histocompatibility complex (HLA in man) on chromosome 6 [6]. Based on electrophoretic mobility in agarose gels multiple alleles, including null alleles have been assigned to each locus. These differences appear to be localized to the α-chain, especially the C4d region [7].

Authors in [8] have shown that the blood groups Chido (i.e., C4B) and Rodgers (i.e., C4A) were associated with C4. The antigenic determinants for these blood groups have been located in the C4d part of the molecule in [9].

The C4d fragment has been isolated from pooled serum and the amino acid sequence determined for 196 residues. While a homogenous sequence was observed for the N-terminal 90 amino acids [10] heterogeneity was detected at 3 out of 106 positions in a second region of C4d [11]. Comparison with the corresponding sequence predicted from the nucleotide sequence of a 304 base pair C4 cDNA clone suggested heterogeneity even at a fourth position [12].

We have isolated and separated C4d components of the Chido and Rodgers types from human C4 after digestion by trypsin [13]. Tryptic C4d (try-C4d) was found to consist of a fragment with an M_r of 28000–30000 (designated A) carrying either the Chido or Rodgers determinants, and in addition a fragment with an M_r of 7500 (designated B) non-covalently associated with the larger polypeptide. This report describes detection of primary sequence differences at 5 positions between the Chido and Rodgers try-C4d fragments.

2. MATERIALS AND METHODS

2.1. *Materials*

Sephadex G-75 (Pharmacia Fine Chemicals, Uppsala); TPCK-trypsin (Worthington Biochem.

* Present address: Scripps Clinic and Research Foundation, La Jolla, 10666 North Torrey Pines Rd, CA 92037, USA

Published by Elsevier Science Publishers B.V.

[^{14}C]methylamine and iodo-[^{3}H]acetic amide (New England Nuclear, Boston, MA).
(DITC) (LKB, Bromma, Sweden); controlled pore glass (CPG 10) (Sera Feinbiochemica, Heidelberg); [^{14}C]methylamine and iodo-[^{3}H]acetic amide (New England Nuclear, Boston, MA).

2.2. *Isolation of tryptic C4d*

Tryptic C4d of either the Chido or Rodgers variant and B-fragments from combined Chido and Rodgers variants was prepared from purified human C4 as in [13], with the glutamic acid residue and cysteine residue in the thiolester labelled with [^{14}C]methylamine and the iodo-[^{3}H]acetamide, respectively. The larger fragment (A) was separated from the 7500-M_r fragment (B) by gel filtration in 1.0 M acetic acid as in [13].

2.3. *Preparation of cyanogen bromide fragments*

A-fragments of either the Chido or Rodgers variant of try-C4d were digested with cyanogen bromide as in [14]. The material was then chromatographed on Sephadex G-75 superfine (1.0 × 135 cm) in 1.0 M acetic acid with a flow rate of 3 ml/h.

2.4. *Preparation of tryptic fragments*

Peptide fragments were digested with TPCK trypsin (10 units trypsin/80 nmol peptide) at 37°C for 3 h in 0.1 M NH_4HCO_3. The digestion was stopped by the addition of a double weight amount of SBTI over trypsin. Fragments were isolated by HPLC using a C 18 column (Waters radial compression separation system) in 0.16% trifluoroacetic acid in H_2O (v/v) with a linear gradient of acetonitrile containing 0.10% trifluoroacetic acid (v/v) from 0–60%. The column was operated with a flow rate of 1.0 ml/min at room temperature, and the elution was monitored at 220 nm.

2.5. *Amino acid analysis and sequence determination*

Amino acid analysis was performed as in [15]. Manual Edman degradation was done by the direct phenylthiohydantoin method [16]. Automatic Edman degradation on solid phase was performed on an LKB 4020 solid phase peptide sequencer. CNBr peptides with a C-terminal homoserine were immobilized onto aminopropyl glass as in [17]. Peptides without C-terminal homoserine were coupled to DITC-glass [18]. Automatic Edman degradation in liquid phase with polybrene [19] was performed in a Beckman 890C liquid phase sequencer equipped with a cold trap using a program described in [20]. The PTH amino acids were identified by HPLC on a Nucleosil ODS column (4 × 300 mm) in 6 mM sodium acetate buffer (pH 4.9), using a gradient of acetonitrile.

3. RESULTS AND DISCUSSION

Separation of CNBr peptides from either Chido or Rodgers variants of the A-fragment on G-75 yielded 5 defined peaks (fig.1). The elution profile was identical for material from both genetic variants. The material in each peak was subjected to amino acid analysis and to manual and automatic Edman degradation, and 6 different fragments were identified, designated CN1–6. CN3 and CN4 eluted together, and this fraction contained the radioactive label. Manual Edman degradation of the CN3/CN4 fraction revealed a

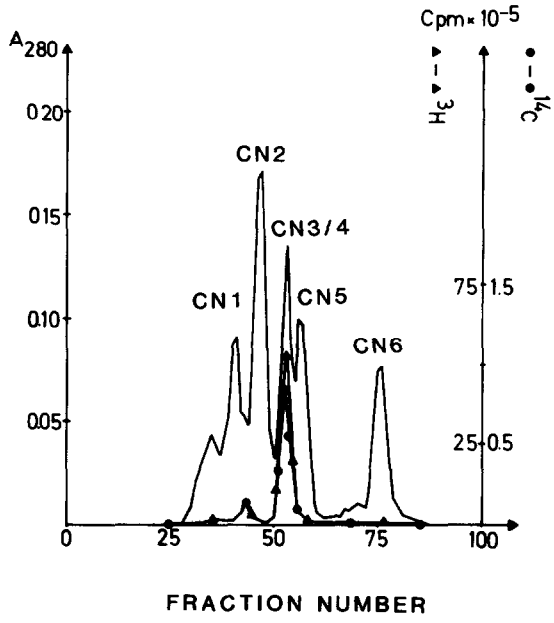

Fig.1. Chromatography of cyanogen bromide fragments from Chido or Rodgers try-C4d on Sephadex G-75 superfine in 1.0 M acetic acid (column 1 × 135 cm, flow rate 3.0 ml, fraction volume 1.0 ml). Approximately 90 nmol of digest were dissolved in 50% (v/v) acetic acid and applied. Appropriate pools were taken. (——) Absorbance at 280 nm, (▲—▲, ●—●) radioactivity.

double sequence. One of these (i.e., CN4) was identical to the N-terminal sequence of try-C4d reported by us in [12]. The same sequence is also recognized in human C4d [10], located 23 amino acids from its N-terminus (position 1 in fig.3), and the second sequence in the CN3/CN4 fraction (i.e., CN3) was found to be identical to the sequence following the methionine in position 36 in fig.3. To obtain the C-terminal part of CN3 (which is not covered by the data in [10]) the CN3/CN4 pool was digested with trypsin and fragments isolated by HPLC on a C18 column. A tryptic nonapeptide (residues 64–72, fig.3) was found to overlap with the published sequence [10]. The tripeptide Gly-Tyr-Hse was considered to constitute the C-terminal of CN3, and its position immediately C-terminal to the nonapeptide was determined by comparison with an overlapping peptide (residues 48–76, fig.3) (J. Gagnon and R.D. Campbell, personal communication).

CN1 was found to be contaminated with a peptide which had an N-terminal sequence identical to CN3, probably representing incompletely digested material. CN1/CN3 were co-sequenced for 15 cycles whereafter the N-terminal sequence for CN1 was deduced by subtracting the sequence of CN3. CN6 was found to consist of a dipeptide, Ala-Hse. CN5 lacked homoserine, and was considered to represent the C-terminal of the A fragment. Since it also lacked lysine, the peptide was coupled to DITC-glass after conversion of the C-terminal arginine to ornithine [18] before 21 cycles of solid phase sequencing was performed.

CN2 was of particular interest, since several amino acids in the composition differed between fragments derived from the Chido and Rodgers variants (see table 1). Solid phase sequencing of the 35 N-terminal amino acids revealed only polymorphism in residue 94, where the Chido variant contained glycine and the Rodgers variant an aspartic acid residue. The former statement was confirmed

Table 1

Amino acid composition of try-C4d CN2T5 from the Chido and Rodgers variants

Amino acid	Chido	Rodgers
Trp	ND (3)	ND (3)
Lys	3.07 (3)	3.06 (3)
His	0.97 (1)	0.20 (0)
Arg	4.05 (4)	3.92 (4)
Asp	4.19 (4)	5.45 (6)
Thr	3.29 (3)	3.31 (3)
Ser	10.69 (11)	9.66 (10)
Hse	0.30 (1)	0.30 (1)
Glu	12.37 (12)	12.16 (12)
Pro	2.22 (2)	3.19 (3)
Gly	5.29 (5)	4.53 (4)
Ala	5.87 (6)	5.75 (6)
Cys	– (0)	1.20 (1)
Val	3.94 (4)	3.89 (4)
Ile	1.61 (2)	1.21 (1)
Leu	9.00 (9)	9.00 (9)
Tyr	1.12 (1)	1.17 (1)
Phe	2.99 (3)	3.07 (3)

Values are given as mol of residue/mol of peptide. ND, not determined. Thr and Ser were corrected using the recovery factors 0.96 and 0.89, respectively. Cys was determined as cysteic acid. Numbers in parentheses represent residues found in the sequence determination

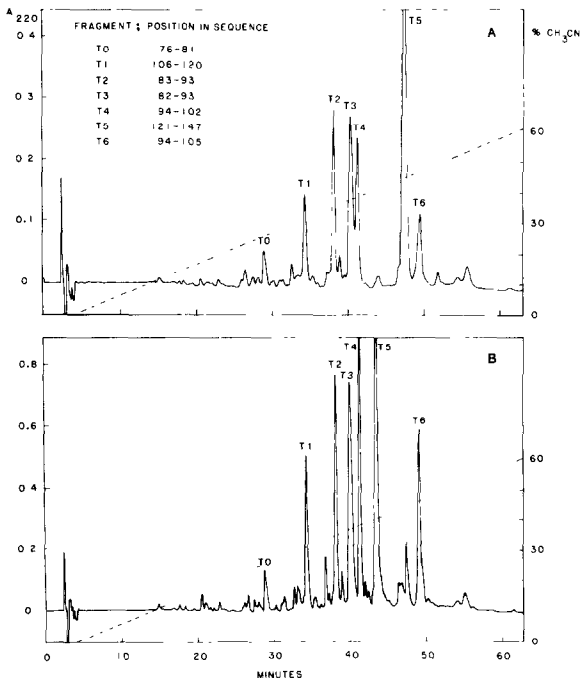

Fig.2. Separation of tryptic fragments from Chido CN2 (A) and Rodgers CN2 (B) by HPLC in 0.16% (v/v) trifluoroacetic acid. The digestion mixture (220 μl, containing 40–80 nmol of peptide) was directly injected to the C18 column (8 × 100 mm). Flow rate 1.0 ml/min, fraction volume 0.4 ml. (——) Absorbance at 220 nm, (– – –) acetonitrile gradient.

by an extended automatic sequence determination of a Chido CN2 which covered 61 of the amino acids in the peptide. In order to complete the sequence, CN2 from the Chido or Rodgers variants was digested with trypsin, and the fragments were separated by HPLC on a C18 column (fig.2). Of the 7 fragments obtained, designated T0–T6, T5 from Chido eluted later than the corresponding Rodgers fragment. Automatic N-terminal sequence analysis of the whole T5 fragments revealed that the Chido and Rodgers variants differed in positions 141 (Leu/Pro), 142 (Ser/Cys), 145 (Ile/Leu) and 146 (His/Asp). No other differences were found between the tryptic fragments from the Chido and Rodgers variants of CN2 except in T4, where the polymorphism in position 94 (Gly/Asp) was confirmed. The localisation of serine in position 148 was deduced from the amino acid composition and confirmed by J. Gagnon and R.D. Campbell (personal communication).

We were unable to detect any differences in the amino acid composition between Chido and Rodgers variants of the B-fragment, why pooled material from both variants were processed together. The N-terminal sequence was determined for 20 residues and B was found to be a part of the C4d segment sequenced in [11], although their data do not cover the C-terminal end. The cyanogen bromide fragments from fragment B (designated CN7–9) were sequenced to the extent shown in fig.3. Trypsin digestion was performed on intact B, whereafter the digest was fractionated on C18 (not shown). One of the isolated fragments (CN7) which overlapped with the C-terminal sequence described in [11] was completely sequenced (25 residues), and considering the molecular mass of fragment B it was concluded to represent its C-terminal end.

The alignment of the fragments is outlined in fig.3. CN3 is interposed between CN4 and CN2. The positions of CN2 and CN1 are based on two overlaps (J. Gagnon and R.D. Campbell, personal communication), first the sequence Gly-Tyr-Met-Arg, which corresponds to the linkage between CN3 and CN2, and then Ser-Met-Gln-Gly-Gly-, which is found between CN2 and CN1. The pro-

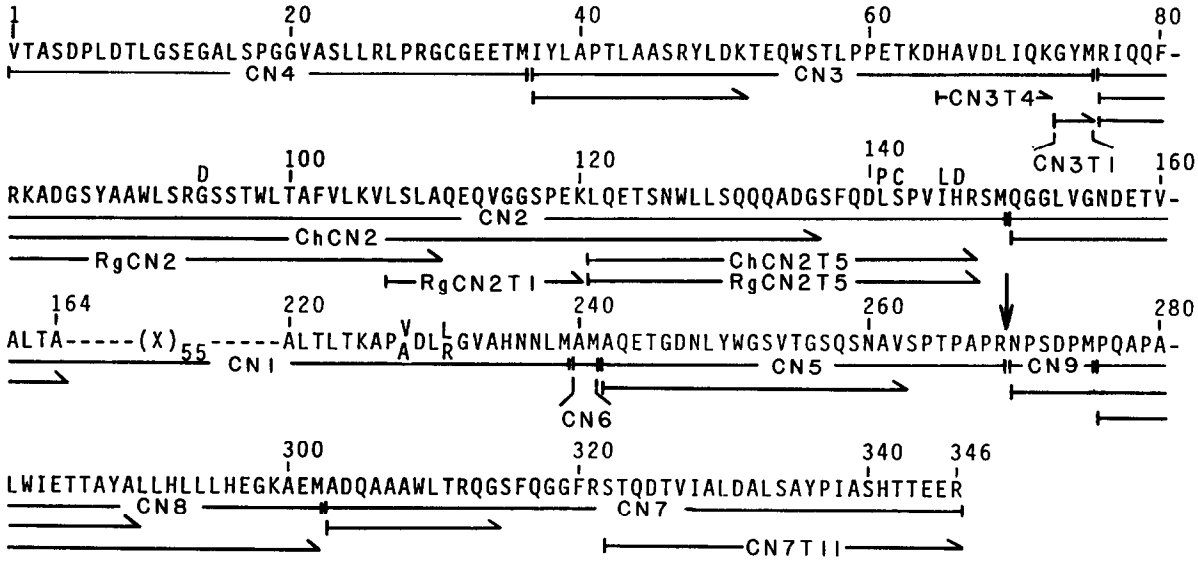

Fig.3. Amino acid sequence and alignment of cyanogen bromide (CN) and tryptic (T) fragments of try-C4d. The sequence of residues 1–67 is taken from [10] and of residues 220–325 from [11]. The Chido sequence is displayed, and the variant amino acids found in Rodgers CN2 (residues 94, 141, 142, 145, 146) are shown above the Chido sequence. The polymorphism reported in [11] at positions 228 and 231 are indicated. Arrows denote positive identification of PTH-amino acids from a particular fragment. The vertical arrow indicates the tryptic cleavage site between fragments A and B; 30–60 nmol of peptide were subjected to automatic sequence determination, and repetitive yields were 96–98% and 90–92% for the liquid and solid phase sequencer, respectively.

posed length of CN1 is based on the corresponding cDNA sequence (M. Carroll, personal communication). The N-terminal sequence of the segment reported in [11] will then constitute the C-terminal part of CN1, followed by CN6 and then CN5, which locates the B fragment C-terminally to the A-fragment. We have not found any evidence for the two polymorphic sites (positions 228 and 231) reported in the C-terminal part of CN1 [11], but the contamination of CN1 with undegraded CN3-CN2 limited the possibilities to discover such differences. Our sequence data of fragment B confirmed the results in [11], except that only arginine (and no valine) was detected in position 321, and histidine was found in position 292 instead of valine. In the latter case the cDNA was reported to code for histidine [12]. On the other hand, the cDNA codes for serine in position 307, where we and authors in [11] found alanine.

Compared to tryptic C4d (346 residues) C4d generated with factor I and C4 binding protein [4] contains an additional 23 amino acids in the N-terminal end [10] and an extra 10–15 amino acids in the C-terminus, the latter of which carries carbohydrate (J. Gagnon and R.D. Campbell, personal communication). Although the Chido and Rodgers variants of fragment A seem to have the same polypeptide chain length they migrate with different velocities on SDS–PAGE [13], suggesting that conformational differences exist between them, since carbohydrate has not been detected in try-C4d during this investigation. The same migration patterns have also been observed with the intact α-chains of the two variants [21]. Marked differences in hemolytic activity have been found for C4 from the two gene loci [6]. Even if it cannot be stated that the primary sequence differences reported here are responsible, it still remains a thrilling explanation.

ACKNOWLEDGEMENTS

The skilful technical assistance of Mrs Lena Möller and Mrs Inga Hägglöv is gratefully appreciated. We are also in great debt to Dr Michael Carroll, Department of Biochemistry, Oxford University, England, for valuable discussions. This project was supported by a grant from the Swedish Medical Research Council (project 13x-02518).

REFERENCES

[1] Schreiber, R.D. and Müller-Eberhard, H.J. (1974) J. Exp. Med. 140, 1324–1335.
[2] Lundwall, Å., Malmheden, I., Stålenheim, G. and Sjöquist, J. (1981) Eur. J. Biochem. 117, 141–146.
[3] Porter, R.R. and Reid, K.B.M. (1979) Adv. Prot. Chem. 33, 1–71.
[4] Shiraishi, S. and Stroud, R.M. (1975) Immunochemistry 12, 935–939.
[5] Scharfstein, J., Ferreira, A., Gigli, I. and Nussenzweig, V. (1978) J. Exp. Med. 148, 207–222.
[6] Awdeh, Z.L. and Alper, C.A. (1980) Proc. Natl. Acad. Sci. USA 77, 3576–3580.
[7] Mevåg, B., Olaisen, B. and Teisberg, P. (1981) Scand. J. Immunol. 14, 303–307.
[8] O'Neill, G.J., Yang, S.Y., Tegoli, J., Berger, R. and Dupont, B. (1978) Nature 273, 668–670.
[9] Tilley, C.A., Romans, D.G. and Crookston, M.C. (1978) Nature 276, 713–715.
[10] Campbell, R.D., Gagnon, J. and Porter, R.R. (1981) Biochem. J. 199, 359–370.
[11] Chakravarti, D.N., Campbell, R.D. and Gagnon, J. (1983) FEBS Lett. 154, 387–390.
[12] Carroll, M.C. and Porter, R.R. (1983) Proc. Natl. Acad. Sci. USA 80, 264–267.
[13] Lundwall, Å., Hellman, U., Eggertsen, G. and Sjöquist, J. (1982) Molec. Immunol. 19, 1655–1665.
[14] Gross, E. and Witkop, B. (1962) J. Biol. Chem. 237, 1856–1860.
[15] Spackman, D.H., Stein, W.H. and Moore, S. (1958) Anal. Chem. 30, 1190–1206.
[16] Iwanaga, S., Wallén, P., Gröndahl, N.J., Henschen, A. and Blombäck, B. (1969) Eur. J. Biochem. 8, 189–199.
[17] Horn, M.J. and Laursen, R.A. (1973) FEBS Lett. 36, 285–288.
[18] Laursen, R.A., Horn, M.J. and Bonner, A.G. (1972) FEBS Lett. 21, 67–70.
[19] Tarr, G.E., Beecher, J.F., Bell, M. and McKean, D.J. (1978) Anal. Biochem. 84, 622–627.
[20] Thomsen, J., Bucher, D., Brunfeldt, K., Nexo, E. and Olesen, H. (1976) Eur. J. Biochem. 69, 87–96.
[21] Roos, M.H., Mollenhauer, E., Démant, P. and Rittner, C. (1982) Nature 298, 854–856.