THE AMINO ACID SEQUENCE OF "HEAVY CHAIN DISEASE" PROTEIN ZUC. STRUCTURE OF THE Fc FRAGMENT OF IMMUNOGLOBULIN G3

C. Wolfenstein-Todel, B. Frangione+, F. Prelli, and E.C. Franklin

Irvington House Institute
Department of Medicine and Pathology+
New York University Medical Center
550 First Avenue
New York, New York 10016

Received May 25, 1976

SUMMARY

The sequence of the Fc fragment of human IgG3 was studied, using a naturally-occurring γ3 heavy chain variant (ZUC). Though the molecule is internally deleted, it contains 248 residues, including the entire Fc fragment. The almost complete sequence of the C_H2 and C_H3 domains (position 234 to 446) indicates an extremely close evolutionary relationship with γ1 and γ4 chains. There is a 95% homology between IgG3 and IgG1 and 92% between IgG3 and IgG4 in the C_H2 and C_H3 domains.

INTRODUCTION

Of the four subclasses of γ chains, the complete amino acid sequence is known only for the γ1 chain (1) and a partial sequence for the γ4 chain (2). Since the subclasses differ in some of the biologic properties located in the Fc fragment (see review, 3), it seemed of value to attempt to elucidate the amino acid sequences of the other subclasses of γ chains in the hope of correlating structural features with biological functions. On the basis of peptide map analyses of a number of γ3 heavy chains and Fc fragments, the number of differences expected is small (4). Since the Fc fragments, especially of the γ3 subclass are sometimes difficult to obtain due to amino terminal heterogeneity and low recovery (5), we chose to study a naturally occurring γ3 chain variant (ZUC) (6). This report presents a summary of these studies; details will be presented elsewhere.

Copyright © 1976 by Academic Press, Inc.
All rights of reproduction in any form reserved.

MATERIALS AND METHODS

Protein ZUC, γ3, carrying Gm(b) allotype (3), was isolated from the urine by starch zone electrophoresis at pH 8.6, and tested for purity by immunoelectrophoretic analysis. Complete reduction and alkylation, cyanogen bromide digestion and separation of the fragments were carried out as described (6).

Enzyme digestions and separation of peptides: The labelled protein or the cyanogen bromide fragments were digested with L (1-tosylamido-2-phenyl) ethyl chloromethyl ketone trypsin (Worthington) in 0.2M ammonium bicarbonate (pH 8.3) for 16 hours at 37°, enzyme substrate ratio 1:50 w/w. The digest was freeze-dried and the peptides were fractionated by gel filtration. Purification of the peptides by high voltage paper electrophoresis, amino acid analyses and determination of amino acid sequences were done as described previously (7) either by manual or automatic Edman degradation using a Beckman 890 Sequencer (8). Molecular weight determinations were performed in 7% polyacrylamide gels containing 0.1% sodium dodecyl sulfate, using the Weber-Osborn techniques (9).

RESULTS

The molecular weight of protein ZUC, as determined by SDS-acrylamide gel electrophoresis was 31,000. The cyanogen bromide fragments of this protein were separated on a column of Sephadex G-75 equilibrated with 5% formic acid as previously reported (6). There were four methionine residues in the molecule, so that five fragments were expected. Peak I (6) contained two fragments corresponding to C_H2 and part of the C_H3 domain since the Met residue between the two domains (position 358) (1), is extremely resistant to cleavage. Peak II contained the 55 residues from the amino terminal of the molecule, previously sequenced (6). Peak III contained two peptides coming from the carboxyl end of the molecule (CBIII and CBIV), one of which is the carboxy terminal octadecapeptide (10).

The tryptic digest of Peak I was fractionated by chromatography on a column of Sephadex G-50, and the peptides obtained, as well as those derived from the intact molecule, were purified by high voltage paper electrophoresis. Although the majority of the peptides were sequenced, the amount of

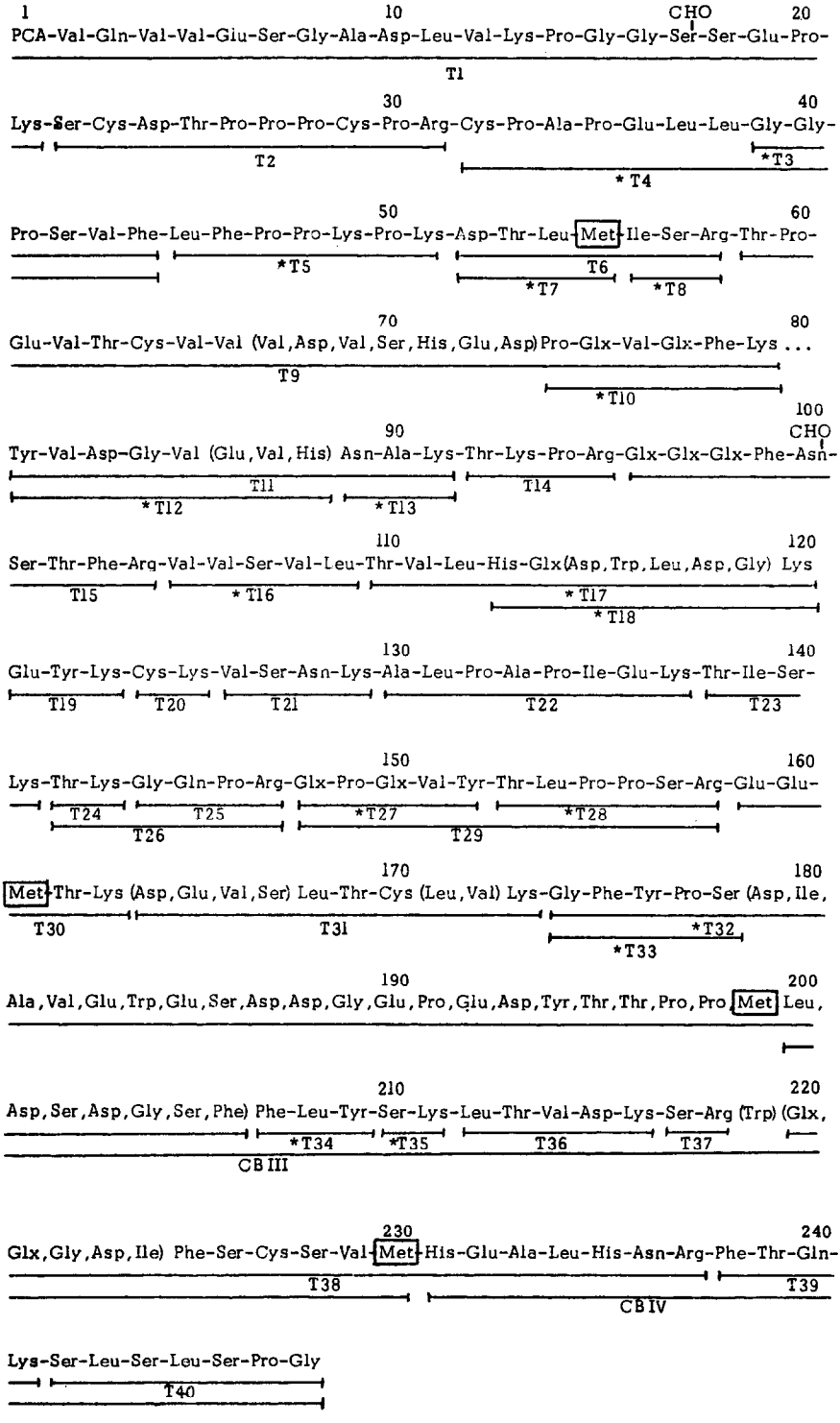
TABLE I. Amino Acid Compositions of Peptides (T1 to T40) and Cyanogen Bromide Fragments (CBIII and CBIV) from protein ZUC[a]

Peptides[b]	T1	T2	*T3	*T4	*T5	T6	*T7	*T8	T9	*T10	T11	*T12	*T13	T14	T15
Lysine	2.15				2.08				1.21	1.00	1.13		1.22	1.07	
Histidine									0.76		0.82	0.78			
Arginine		1.00				0.84		0.94						1.00	0.92
CMCys		1.62		0.48					0.54						
Aspartic acid	1.10	1.10				1.00	1.00		2.10		1.74	1.22	1.00		1.00
Threonine		0.95				0.96	0.82		1.57					0.95	0.96
Serine	2.52	1.23	1.00	0.87		1.00		1.00	1.23						1.22
Glutamic acid	4.15			1.03					3.96	2.05	1.00	1.00			2.99
Proline	1.75	3.63	0.79	2.72	2.82				1.58	0.95				0.87	
Glycine	3.20		2.19	1.93							1.07	1.09			
Alanine	1.18			0.85							0.93		0.92		
Valine	3.12		0.83	0.75					4.10	0.78	2.03	2.32			
Methionine						0.70	1.00								
Isoleucine						0.88		1.02							
Leucine	0.75			2.00	1.00	0.96	1.22								
Tyrosine											0.61	0.82			
Phenylalanine			1.06	0.68	1.08				0.82	0.79					1.95
Tryptophan[c]															
Homoserine															
CHO[d]	+														+
Mobilities at 6.5[e]	0.30	0.5	0	0.43	-0.6	0	0.46	-0.55	0.49	-0.04	0.04	0.35	-0.5	-0.8	0.15
N-terminal	Neg	Ser	Gly	ND[f]	Leu	Asp	Asp	Ile	Thr	Pro	Tyr	Tyr	Asp	Thr	Glu

Peptides[b]	*T16	*T17	*T18	T19	T20	T21	T22	T23	T24	T25	T26	*T27	*T28	T29	T30
Lysine		1.00	1.00	1.22	1.00	0.76	0.91	1.00	1.00		1.21				1.00
Histidine		0.78	0.95												
Arginine										1.00	0.85		0.67	0.95	
CMCys					0.80										
Aspartic acid		1.80	1.90			1.06									
Threonine		1.00						0.86	0.86		1.00		1.00	0.98	0.86
Serine	1.12					1.00		1.05					1.26	1.22	
Glutamic acid		1.12	1.13	1.00			1.00			1.29	1.00	2.07		2.21	1.97
Proline							1.92			1.08	0.91	0.91	1.82	2.66	
Glycine		1.12	1.07							1.28	1.00				
Alanine							2.04								
Valine	2.05	0.82				0.98						1.00		1.23	
Methionine															0.75
Isoleucine							1.00	0.90							
Leucine	1.00	2.01	1.03				1.07						1.00	0.90	
Tyrosine				0.62								0.94		1.28	
Phenylalanine															
Tryptophan[c]		+	+												
Homoserine															
CHO[d]															
Mobilities at 6.5[e]	0	-0.23	-0.10	0	0	-0.48	-0.06	-0.4	-0.7	-0.45	-0.68	0.35	-0.35	0	0.35
N-terminal	Val	Thr	ND[f]	Glu	ND[f]	Val	Ala	Thr	Thr	Gly	Thr	Glu	Thr	Glu	Glu

Peptides[b]	T31	*T32	*T33	*T34	*T35	T36	T37	T38	T39	T40	CBIII	CBIV
Lysine	1.00				1.00	1.13			1.00		2.26	1.21
Histidine								1.55				1.85
Arginine							1.00	0.73			1.21	1.20
CMCys	0.82							0.49			0.95	
Aspartic acid	1.12	5.51				1.00		2.00			3.76	1.02
Threonine	0.97	1.57				0.91			1.26		1.25	0.85
Serine	1.17	4.42	1.00		0.95		0.72	1.80		2.72	5.07	2.99
Glutamic acid	1.15	4.12						2.92	1.00		2.16	2.04
Proline		3.60	0.78							0.78		0.93
Glycine		2.98	1.05					1.13		1.00	3.98	1.27
Alanine		1.10						1.08				1.12
Valine		1.21				1.17		1.02			2.21	
Methionine		0.70						0.85				
Isoleucine		0.61						0.64			0.90	
Leucine	2.38	1.02		1.00		1.05		1.14		2.05	3.19	2.80
Tyrosine		1.98	0.95	1.03							0.36	
Phenylalanine		2.39	1.02	0.93				0.66	0.67		2.20	1.00
Tryptophan[c]		+									+	
Homoserine											+	
CHO[d]												
Mobilities at 6.5[e]	0	0.49	0	0	-0.7	0	-0.7	-0.07	-0.45	0	0.2	-0.35
N-terminal	ND[f]	Gly	Gly	Phe	Ser	Leu	Ser	ND[f]	Phe	Ser	ND[f]	ND[f]

T = tryptic; *T = tryptic-chymotryptic
a = Compositions are reported as moles of amino acid per mole of peptide.
b = Hydrolysis for 20 hr.
c = Detected by staining on paper with Ehrlich's reagent.
d = Detected on amino acid analyzer.
e = Mobilities are given relative to the distance between ϵ-N_2ph-lysine and aspartic acid.
f = ND, not done.


1
10
CHO
20
PCA-Val-Gln-Val-Val-Glu-Ser-Gly-Ala-Asp-Leu-Val-Lys-Pro-Gly-Gly-Ser-Ser-Glu-Pro-
T1
30
40
Lys-Ser-Cys-Asp-Thr-Pro-Pro-Pro-Cys-Pro-Arg-Cys-Pro-Ala-Pro-Glu-Leu-Leu-Gly-Gly-
T2
*T3
*T4
50
60
Pro-Ser-Val-Phe-Leu-Phe-Pro-Pro-Lys-Pro-Lys-Asp-Thr-Leu-Met-Ile-Ser-Arg-Thr-Pro-
*T5
T6
*T7
*T8
70
80
Glu-Val-Thr-Cys-Val-Val (Val,Asp,Val,Ser,His,Glu,Asp) Pro-Glx-Val-Glx-Phe-Lys ...
T9
*T10
100
90
CHO
Tyr-Val-Asp-Gly-Val (Glu,Val,His) Asn-Ala-Lys-Thr-Lys-Pro-Arg-Glx-Glx-Glx-Phe-Asn-
T11
T14
*T12
*T13
110
120
Ser-Thr-Phe-Arg-Val-Val-Ser-Val-Leu-Thr-Val-Leu-His-Glx(Asp,Trp,Leu,Asp,Gly) Lys
T15
*T16
*T17
*T18
130
140
Glu-Tyr-Lys-Cys-Lys-Val-Ser-Asn-Lys-Ala-Leu-Pro-Ala-Pro-Ile-Glu-Lys-Thr-Ile-Ser-
T19
T20
T21
T22
T23
150
160
Lys-Thr-Lys-Gly-Gln-Pro-Arg-Glx-Pro-Glx-Val-Tyr-Thr-Leu-Pro-Pro-Ser-Arg-Glu-Glu-
T24
T25
*T27
*T28
T26
T29
170
180
Met-Thr-Lys (Asp,Glu,Val,Ser) Leu-Thr-Cys (Leu,Val) Lys-Gly-Phe-Tyr-Pro-Ser (Asp,Ile,
T30
T31
*T32
*T33
190
200
Ala,Val,Glu,Trp,Glu,Ser,Asp,Asp,Gly,Glu,Pro,Glu,Asp,Tyr,Thr,Thr,Pro,Pro Met Leu,
210
220
Asp,Ser,Asp,Gly,Ser,Phe) Phe-Leu-Tyr-Ser-Lys-Leu-Thr-Val-Asp-Lys-Ser-Arg (Trp) (Glx,
*T34
*T35
T36
T37
CB III
230
240
Glx,Gly,Asp,Ile) Phe-Ser-Cys-Ser-Val-Met-His-Glu-Ala-Leu-His-Asn-Arg-Phe-Thr-Gln-
T38
T39
CB IV
Lys-Ser-Leu-Ser-Leu-Ser-Pro-Gly
T40

protein available was not sufficient to isolate the peptides containing the overlaps. Therefore, they were placed by homology with the Fc fragment of γ1 heavy chain protein Eu (1). Table I gives the amino acid composition and amino terminal residues of these peptides. Figure 1 shows the sequence of protein ZUC. Several points deserve comment. The molecule contains 248 residues and two carbohydrate moieties: one at position 17 attached to a serine residue, and the other at position 100 attached to an asparagine residue.
Protein ZUC is an internally deleted glycoprotein, carrying Gm(b) allotype, which has a gap of about 200 residues, but contains the intact Fc fragment (6). Since the deletion starts 18 residues after the amino terminus, position 19 in ZUC corresponds to position 216 in a normal γ1 chain (1).

A total of 40 peptides was isolated. Of the 28 tryptic peptides expected, only 22 were recovered because of two Lys-Pro bonds and some chymotryptic cleavage. The sequence of 30 residues starting with residue 56, as well as some ambiguities between residue 179-186 and 219-224 (Fig. 1) were confirmed by automatic sequencing of a cyanogen bromide fragment and the homologous peptides from another γ3 heavy chain of the same Gm type as shown in Figure 2. Figure 2 also compares the C_H2 and C_H3 domains of γ3 with those of γ1 (1) and γ4 heavy chains (2). It should be emphasized that the γ3 chain is larger than the other γ chains (11) and that the γ1 numbering (1) will be used until the complete amino acid sequence of γ3 heavy chain is known.

In the CH2 domain (234 to 341), there are at least five substitutions: Lys (274) by Glx, Asn (276) by Lys, Ala (339) by Thr, as well as the previously

Figure 1 - Amino acid sequence of a γ3 "Heavy Chain Disease" protein (ZUC) (internally deleted immunoglobulin heavy chain). Met residues are in boxes. Lines indicate tryptic (T) and tryptic-chymotryptic (*T) peptides and two of the cyanogen bromide fragments (200-230 and 231-248) (see text). Dashed line (position 80) indicates an unidentified residue. Residue 19 corresponds to position 216 in γ1 heavy chain (1). All peptides were sequenced, except for the residues in brackets ( ).

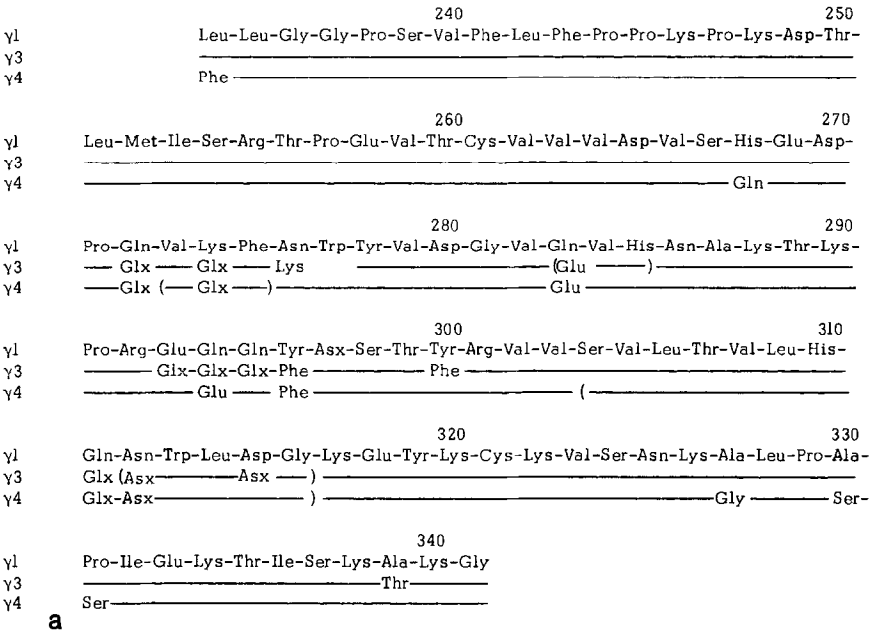


Figure 2a - Comparison of the C_H2 region of IgG1 (1), IgG3, and IgG4 (2). Residue numbering from protein Eu (1). Solid lines indicate identical residues. Residue 277 was not identified.

Figure 2b - Comparison of the C_H3 region of IgG1 (1), IgG3, and IgG4 (2). Residue numbering from protein Eu (1). Solid lines indicate identical residues. Residue 392 was not identified.

reported substitution of Tyr by Phe at residues 296 and 300, which appear to be related to the Gm(b) allotypic specificity (12) (Fig. 2a). Trp (277) was not identified from protein ZUC, but its presence is suggested by the existence of four residues of tryptophan in the molecule.

In the C_H3 domain (342 to 446) there are at least five differences including Ile for Val at position 422, and the previously reported Gm(b) related substitution of His-Tyr at position 435-436 by Arg-Phe. The third cyanogen bromide fragment starts at position 398 since Val (397) has been replaced by Met. The sequence between 384 and 404 could not be established due to the insolubility of the peptide, therefore it is not possible to exclude the exist- of other differences in addition to the absence of Lys (392). This is a region where a high degree of variability has been detected in different species (13). Taking the 212 residues from 234-446 of the γ3 heavy chain, 83.5% of the Fc fragment has been sequenced. Excluding the unknown amides, this region shows a 95% homology between IgG3 and IgG1 and 92% between IgG3 and IgG4 (see Fig. 2). All the amino acid interchanges can be accounted for by single base substitutions. In contrast, the more striking differences in the hinge regions appear to be due to deletions and duplications (14).

Future studies, perhaps using synthetic peptides, will be needed to relate these differences to functional properties such as complement fixation which appear to reside in the C_H2 domain of γ chains (15, 16).

ACKNOWLEDGEMENTS

This investigation was supported in part by USPHS Grants Nos. AM 01431 and AM 02594; The Irvington House Institute, and the Helen and Michael Schaffer Fund.

We thank Dr. Elliot Osserman for making protein ZUC available to us, Mrs. Ellen Rosenwasser and Mr. Rostyslaw Kutny for expert technical assistance, and Ms. Marlene Chavis for secretarial help.

REFERENCES

1. Edelman, G.M., Cunningham, B.A., Gall, W.E., Gottlieb, P.D., Rutishauser, U. and Waxdal, M.S. (1969), Proc.Nat.Acad.Sci. USA, 63, 78-85
2. Pink, J.R.L., Buttery, S.H., DeVries, G.M. and Milstein, C. (1970), Biochem.J., 117, 33-47.
3. Natvig, J.B. and Kunkel, H.G. (1973), Adv.Immunol., 16, 1-59.
4. Frangione, B., Franklin, E.C., Fudenberg, H.H. and Koshland, M.E. (1966), J.Exp.Med., 124, 715-732.
5. Michaelsen, T.E. and Natvig, J.B. (1974), Scand.J.Immunol., 3, 865-870.
6. Frangione, B. and Milstein, C. (1969), Nature, 224, 597-599.
7. Wolfenstein, C., Frangione, B., Mihaesco, E, and Franklin, E.C. (1971), Biochemistry, 10, 4140-4145.
8. Edman, P. and Begg, G. (1967), Eur.J.Biochem., 1, 80-91.
9. Weber, K. and Osborn, M. (1969), J.Biol.Chem., 244, 4406-4412.
10. Prahl, J.W. (1967), Biochem.J., 105, 1019-1028.
11. Michaelsen, T.E. and Natvig, J.B. (1974), J.Biol.Chem., 249, 2778-2785.
12. Grey, H.M. and Abel, C.A. (1970), Protides Biol.Fluids, Proc.Colloq., 17, 229-231.
13. Melamed, M.D. (1976), Immunochemistry, 13, 271-279.
14. Adlersberg, J.B., Franklin, E.C. and Frangione, B. (1975), Proc.Nat. Acad.Sci. USA, 72, 723-727.
15. Kehoe, J.M. and Fougereau, M. (1969), Nature, 224, 1212-1213.
16. Yasmeen, G.D., Ellerson, J.R., Dorrington, K.J. and Painter, R.H., (1976), J.Immunol., 116, 518-526