## Conservation of structural and functional domains in complement component C3 of *Xenopus* and mammals

(evolution/sequence homology/major histocompatibility complex/properdin binding site/Arg-Gly-Asp sequence)

DARIO GROSSBERGER, ANNE MARCUZ, LOUIS DU PASQUIER, AND JOHN D. LAMBRIS

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland

Communicated by Hans J. Müller-Eberhard, October 24, 1988 (received for review June 8, 1988)

The cDNA sequence and the deduced amino ABSTRACT acid sequence of the Mr 34,000 C-terminal fragment of Xenopus laevis complement component C3 are presented. The sequence of Xenopus C3 has 57% nucleotide identity to the corresponding sequence of human C3 and  $\approx$ 49% amino acid identity to C3 from human, mouse, and rabbit. The Xenopus C3 sequence shows clusters of high and of low similarity to the mammalian C3 sequences. One of these regions of high similarity represents the domain of mammalian C3b involved in the binding of properdin, a regulator of the alternative pathway of complement activation. It is not clear whether the other highly conserved regions are involved in binding to other C3 ligands. The Xenopus C3 sequence completely lacks the Arg-Gly-Asp sequence, which has been suggested to be the recognition site of the human complement receptor type 3 on the iC3b fragment of human C3. The Xenopus C3 gene is shown not to be linked to the Xenopus major histocompatibility complex, as is also the case in mammals. Since the gene of the related molecule C4 is MHC-linked in both mammals and Xenopus, the C3 and C4 genes may have separated before Xenopus and mammals speciated.

The third component of complement (C3) plays a central role in both the classical and alternative pathways of complement activation by interacting with numerous serum and surface complement proteins (1, 2). In addition to being important for complement activation, C3 has been found to play a significant role in inflammatory processes and in the immune response (1, 3).

Xenopus has both classical (4) and alternative (5, 6)pathways of complement activation. The complement proteins C3 (7) and C4 (4) isolated from Xenopus serum have subunit structures similar to mammalian C3 and C4. Human C3 fragment C3b is cleaved by Xenopus serum in a way similar to its cleavage by human serum, suggesting the presence of the regulatory proteins factors I and H in Xenopus serum (8). Receptors for Xenopus C3 fragments have been found on the surface of Xenopus macrophages (9), although it is not clear whether these are the Xenopus homologs of mammalian macrophage complement receptor type 1 (CR1) or 3 (CR3). Since the Xenopus C3 molecule shares many features with mammalian C3, including binding to mammalian CR1, properdin, and C5 (10), comparison of the Xenopus C3 sequence with mammalian C3 sequences may reveal regions that are involved in binding to its many ligands as well as the structural elements of these interactions. Such regions are expected to be relatively conserved, since their rate of evolutionary divergence would be constrained by the rate of evolution of their ligands.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

This paper describes the isolation and characterization of a cDNA clone representing the carboxyl end of the  $\alpha$  chain of the *Xenopus* C3 molecule.\* The *Xenopus* C3 sequence shows  $\approx$ 49% amino acid identity to the corresponding sequence of mammalian C3 and has clusters of conserved sequences that probably represent regions important for the structure and function of the C3 molecule. In addition, the gene for *Xenopus* C3 is shown not to be linked to the *Xenopus* major histocompatibility complex (MHC), a result with implications for understanding the evolution of complement genes and the MHC.

## MATERIALS AND METHODS

**Library Construction.** RNA was prepared from the liver of an adult female *Xenopus laevis* (11) and a cDNA library was constructed in the expression vector  $\lambda$ gt11 (12) essentially as described (13).

**Preparation of Anti-***Xenopus* **C3 Antibody.** A rabbit anti-*Xenopus* C3 antiserum was prepared by using zymosan-C3b/iC3b particles as described (7). The anti-*Xenopus* C3 antibody was affinity-purified on zymosan coated with  $\alpha$ chain fragments of *Xenopus* C3. The coated zymosan was prepared by incubation in *Xenopus* serum and washing with 0.1% sodium dodecyl sulfate (SDS) containing 2% 2-mercaptoethanol.

Screening of Library. The cDNA expression library was screened essentially as described (14). The filters were incubated successively with biotinylated goat anti-rabbit immunoglobulin (Vector Laboratories) and a complex of avidin and biotinylated horseradish peroxidase (Vector Laboratories) and were developed with the peroxidase substrate 3,3'-diaminobenzidine (Sigma).

**Immunoblots.** Immunoblots of normal *Xenopus laevis* serum were prepared as described (15). The blots were treated in the same way as the filters used in the library screening.

Sequence Determination. The *Eco*RI inserts were subcloned into the Bluescript plasmid vector (Stratagene) and several nested deletion constructs were prepared. Sequence was determined from plasmids (16) with the use of Sequenase (United States Biochemical) and was compiled with the MICROGENIE program (Beckman). Protein alignments were done with a program (written by C. Steinberg, Basel Institute for Immunology) using the Needleman–Wunsch algorithm (17).

Southern Blot. DNA was isolated from *Xenopus* erythrocytes and used as described (18). The probe was excised from the pC3-27 subclone of C3-27 by digestion with *Bam*HI,

Abbreviations: Cn, complement component n; MHC, major histocompatibility complex; CRn, complement receptor type n;  $\alpha_2$ M,  $\alpha_2$ -macroglobulin; RGD, Arg-Gly-Asp.

<sup>\*</sup>The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04493).

purified by agarose gel electrophoresis (19), and labeled to a specific activity of  $10^9$  cpm/ $\mu$ g (20).

## RESULTS

Isolation of a Xenopus C3 cDNA Clone. Activation of the alternative pathway of complement by zymosan leads to the covalent binding of C3b to zymosan through its reactive thioester site on the  $\alpha$  chain (21). This property was used to purify Xenopus C3 fragments and to produce anti-Xenopus C3 antibodies. Electrophoresis of fragments eluted by boiling a zymosan–C3b/iC3b preparation in the presence of SDS and 2-mercaptoethanol showed the expected  $\beta$  chain ( $M_r$  85,000) and the  $M_r$  34,000 C-terminal fragment of the  $\alpha$  chain of iC3b (Fig. 1A). Affinity purification was necessary to reduce the background (Fig. 1B, lane 1). The affinity-purified antibody reacts with the uncleaved  $\alpha$  chain as well as  $\alpha'$  chain and its  $M_r$  34,000 C-terminal fragment (Fig. 1B, lane 2). We do not understand why this antibody did not also recognize the  $M_r$  81,000 N-terminal  $\alpha$ -chain fragments of iC3b (7).

Screening of the  $\lambda$ gt11 expression library with the affinitypurified antibody resulted in one positive phage clone, C3-27 (Fig. 1C). To verify that this clone actually represents *Xenopus* C3, it was plated at high density on several Petri dishes and the filters were used to affinity-purify the anti-*Xenopus* C3 antibody. This antibody recognized the same protein species as the antibody purified on zymosan-bound C3  $\alpha$  chain (Fig. 1B, lane 3). This showed that C3-27 is in fact a *Xenopus* C3 clone. It also placed this clone at the C terminus of the  $\alpha$  chain, since the antibody selected on C3-27 recognized the  $M_r$  34,000 C-terminal  $\alpha$ -chain fragment of the *Xenopus* iC3b.

Sequence Analysis. To obtain longer cDNA C3 clones, a 425-base-pair *Bam*HI fragment was purified from the pC3-27 subclone of the *Eco*RI insert of C3-27. Approximately 200,000 clones were screened by hybridization, and of the resulting 16 positive clones, C3-39 was the longest (997 base pairs). The inserts in this library, as well as in four other



FIG. 1. Characterization of anti-Xenopus C3 antibody and Xenopus C3 clone. (A) Coomassie blue-stained SDS/polyacrylamide gel (22) of Xenopus C3 fragments eluted from zymosan. Markers at left indicate positions and  $M_r \times 10^{-3}$  of standard proteins. (B) Immunoblots of Xenopus serum probed with rabbit anti-Xenopus C3 antibody (lane 1), antibody affinity-purified on zymosan particles having fixed Xenopus C3  $\alpha$ -chain fragments (lane 2), or antibody affinity-purified on clone C3-27 (lane 3). (C) Reactive C3-27 (open arrow) and unreactive (solid arrow) plaques with affinity-purified anti-Xenopus C3 antibody.

libraries tested, appear to be considerably shorter than the full-length C3 mRNA, which is expected to be about 5.5 kilobases (kb).

The full pC3-27 sequence was determined, as well as additional regions of pC3-39. When the assembled nucleic acid sequence of this region of *Xenopus* C3 is aligned with the human C3 sequence (Fig. 2), it shows 57% nucleotide identity.

In Fig. 3 the cloned region of *Xenopus* C3 is aligned with the corresponding regions of the human mRNA and human protein. The human mRNA is 5.5 kb and encodes a protein that is cleaved once to give the  $\alpha$  and  $\beta$  chains of mature C3 (2). The *Xenopus* C3 sequence aligns at the 3' end of the mRNA and at the C terminus of the mature  $\alpha$  chain.

In Fig. 4 the deduced Xenopus C3 amino acid sequence is aligned with the available mammalian sequences, as well as with the sequences of the structurally related molecules C4 (26), C5 (27), and  $\alpha_2$ -macroglobulin ( $\alpha_2$ M) (28). All sequences were aligned against the Xenopus sequence. The Xenopus C3 amino acid sequence is about equally similar to the three mammalian C3 sequences and significantly less similar to the C4, C5, and  $\alpha_2$ M sequences (Table 1). Amino acid residues at positions 48 (cysteine), 63 (aspartic acid), 69 (glycine), 80 (leucine), and others (in reversed print in Fig. 4) are conserved in all complement sequences aligned, although these proteins have diverged considerably in function. These residues are mostly either cysteine or hydrophobic and may be necessary for the overall structural integrity of this family of proteins. In addition to the cysteines found in the  $M_r$  40,000 mammalian C3 fragment, Xenopus C3 has three additional cysteines at positions 25, 27, and 33. In mammalian C3 one of the cysteines is disulfidebonded to the  $M_r$  67,000 N-terminal  $\alpha$ -chain fragment. The  $M_r$ 34,000 C-terminal fragment of the  $\alpha$ -chain is smaller than the mammalian  $M_r$  40,000 fragment, so Xenopus C3 may have a different site of factor I cleavage. A potential Arg/Ser cleavage site is located at residues 37 and 38, and a fragment of  $M_r \approx$ 32,000 would result from cleavage there. If this is the case, the additional cysteines would not reside in the  $M_r$  34,000 Cterminal fragment.

Several structural features shared by these proteins are apparent from these comparisons. The properdin-binding region of human C3 (residues 84–116) (29), which has been further localized to 9 amino acids (residues 105–113; M. Daoudaki and J.L., unpublished observations), shows high degree of similarity to C3 of other species, and much less similarity to the non-properdin-binding proteins C4, C5, and  $\alpha_2$ M. Interestingly, the tripeptide RGD sequence in the human C3 sequence, reported to be the recognition site of CR3 on iC3b (30), is not completely conserved in the mouse and rabbit sequences and is absent in the *Xenopus* sequence. However, considerable conservation exists to the C-terminal side of the RGD site.

The Xenopus C3 Gene Is Not Linked to the MHC. A Southern blot was done with DNA from a family of interspecies hybrids from a cross of X. laevis and X. gilli (Fig. 5). These frogs make diploid eggs that develop upon stimulation with irradiated sperm, and are thus maintained as clones. The MHC types of these cloned frogs were determined by grafting and by mixed lymphocyte culture. In this family there are four MHC haplotypes, a, b, c, and d, which segregate independently (31). The DNAs were digested with EcoRV and Pst I, and the blot was probed with a fragment comprising 425 base pairs 5' of the BamHI site from the pC3-27 subclone (see Fig. 3). The LG6, LG15, and LG17 frogs all have the MHC haplotypes a and c. However, the LG6 DNA has a 24-kb EcoRV fragment that is not shared by LG17. Furthermore, LG15 has a 24-kb Pst I fragment that is not shared by either LG6 or LG17. LG15 also lacks a 6-kb Pst I fragment that is present in LG6 and LG17. In each of these cases, frogs with identical MHC haplotypes have different C3 restriction fragment polymorphisms. In addition, two individuals, LG3 Immunology: Grossberger et al.

Xenopus C3

G	R	v	V	T	v	Y	н.	A	L	v	I	E	ĸ	E	R	ĸ	с	L	N	F	E :	LŚ	3 0	: E	с	ĸ	G	s 1	с.	T	с	ĸ	T	s	R	s	A .	K J	A '	T I	V 5	3 )	C B	
GGA	GGG	TGG	TGA	ICC(	TG	'ACC	ATG	CTC	TTG	TCA	TAG	AGA	LAAG	GAAA	GN	AAGI	rgc(	CTTA	ACI	TTG	AAC	TGT	CTTG	TGA	TGI		GAA	GTA	:A	. AC	TGC	AAG	ACC	rcco	GAA	GTC	CAA	AGC	CAR	CAG	TGTC	CTA!	FAGAA	13
Т	П	111	111	11	111	ш			1		П		11	1	1	1	11	11	1	нт	11	I I				111	1			11		111	1	1	1	П	11	П	Т	I I	11		1 1	
TGT	CGG	TGG	TG		ATG	ACC	ATG	CTA		acca	AAG	ATC	-	CTC	ACC.	9	IGT/	AATA		TCG	ACC	TCM	LGGT	CAC	сата		CAG	CACO	GGA	AACI	AGAA	AAG	AGG	CCTO	AGG	ATG	CCA	AGA	ACA	CTA	TGA	rcc:	ITGAG	42
T.		v	v	Ŧ	м	v		2	r	T.	r	e .	x	v			c	ĸ	x	8	D	T. 1	с s	7 7	т	ĸ	P	<b>X</b> 1	> R	т	R	ĸ	R	P	0	D	2	K I	N	T	M :	<b>r</b> 1	LE	
			•	•		•		~			•		•	•	•		•		••	•					-		-			-	-		••	-	-	-								
3											_			-	÷		-					~ .			n		n					м		~	17		-	•	-		* *	•		
A	с 		ĸ	n		R		•		<b>.</b>	T		3	1	±		1	3			T			• F											•		~ ~ ~		-					
CCT	GTG	CAA	LCCC	CAT	CTCJ	U.G.	ACG	TTG	ATG	CCA	(CC)	TGI	reca	ATTA	ATT	GAT	ATT	TCCI	TG	LTGJ	CTG	GCT:	TTT	TCC:	FGAT	ACC	ATI	CCC:	Aنیانی ۲	TAG	GCTA	ATG		لالفانا 	TGG		AAT	ACA	TCT	CTA	AATA	ATG:	MG1C	21
11	11	1.1		ł.	I		11			1111	11		11 1	11	1		11				111	111			11			н			111		11	11	11 1	111	11			11	1 1	111		
TCT	GTA	CCI	ICC1	LYC	CGGG	GAG	ACC	AGG	ATC	CC3	CTA	TG	ICTJ	ATAT	TTG	CVC1	ATA	TCC	TG	LICI	CTG	GCT:	TTG	CTCC	<b>IGI</b> C	ACA	CATC	ACC	<b>FGAA</b>	CCA	GCTG	CCC	'AAT	GGT	GTTG	ACA	CAT	YCY	TCT	CCA	AGTI	ATG	AGCTG	43
I	с	T	R	¥	R	G	D	Q	D	A	T	M	s	I	L	D	I	S	M	M	T	G 1	,	N P	D	T	D	נס		C Q	L	X	N	G	v	D	R	Y	I	S I	K 7	r I	E L	
N	ĸ	G	A	N	D	ĸ	G	T	L	I	L	¥	L	D	ĸ	v	s	H	I	D	E	B (	c 1	v K	F	Y	A	H (	2 2	F	E	۷	G	F	I	Q	P	A	S	<b>v</b> .	T T	۷.	Y D	
ACA	AAG	GAG	cN	AAT	GACI	LA GO	GAA	CAC	CTT	ATTC	TCI	LACI	TTG	GACJ		GTC	rcc	CACI	TA	SACG	ANG	AAT	GTG	<b>FGAN</b>	GTTT	TAT	CTC	ATC	ATA	CTT	TGAN	GTG	GGT	TTC	ATCO	AGC	CAG	CTT	CTG	TAA	CTG	TGT	ATGAC	42
							- 1					ш	111		11						1.1	1.1		ı	н	1		11		111	1.1	п	1	ł		111			1.1	11	1	11		
			PTCT	TCC	CATI				-		TC	PACO	CTG	GACT		GTC	TCA	CACI	CTO	3366	arc	ACT	GTC	FAGC	TTTC		GTTC	acc		CTT	 TAA1	GTA		CTT	ATCO	300	CTG	GAG	CAG	TCA	ACC	TCT	ACGCC	45
	*						*		,			v	-								n	D (	- 1			T	v		~ ¥		N	v		т.	т. Т	~	P	6		v	<b>r</b> .	v	<b>Y X</b>	
	r	^		3	5	~		•	-	•	•	•	~		~	•					-		• •		-		•	-	* *	• •		•	-	-	-	*	•	•	-	•		•		
		_	_	_		_	_	_	_	_		_		-	_	~									~		-	~ .			_	-		~	_		~	~	~	-	-	~		
I	I	T	P	ט	N	ĸ	С	T	ĸ	F	I	H	v	Б	ĸ	G	s			7		6.	т (	u u			±	с .	ĸ		ĸ	ĸ	N				2	¥	8	±	<u>ь</u> (	. w	R I	
ATT	ATA:	CTC	CA	GAT	AAT	CGT	ICC A	CT	AAA:	CTTI	CACC	CATO	STG	GANG	GAA	GCCI	AGT	GCC:	TG	CTCG	CAG	CCA	TTT	GCCA	AGGG	GAT	ATA	recc	GATC	TCC	AGAJ	GAG	AAC	TGT	TTC	ILCO	CAGO	AGC	AAA	TTG	AGG	GTA	AAATC	56
111	11			11				1		11-1		111	1	111	I	I	1	1	1	111			11	111	1	11	1		1 11	111	11	111		11	1111			11		11	1	1	1 11	
ATT	ACA	ACC	CTG	GYC	GVV	AGC	CT3	CCC	CGG	LLC1	CAC(	CATO	CCG	GYY	YYC	CAC	GAT	GGN	<b>V</b> C	CTG	ACA	ACC	TCT	CCCG	TGA	<b>IGAA</b>	CTG	recc	GCTG	TCC	TGA	GAG	AAT	TGC	TTCJ	TAC		AGT	CCC	ATG	AC.	<b>A</b>	AGGTC	- 40
Y	Y	N	L	E	E	s	с	T	R	F	Y	H	₽	R	ĸ	E	D	G	ĸ	L	N	K :	L	C R	D	E	L	C	RC	: A	E	E	N	С	F	I	Q	ĸ	s	D	D	•	ĸν	
T	A	D	M	R	v	N	M	A	С	A	₽	G	v	D	F	V	Y	ĸ	A	T	L	T	<b>B</b> )	гQ	P	s	D	N	Y C	N	Y	V	M	T	I	ĸ	ĸ	۷	I	ĸ	Q (	G	T D	
CTG	СТС	ACI	TG	AGA	GTC	NAC)	ATGG	CT	rcco	GCTC	ccc	GGA	GTG	GAT	TTT	GTG	TAC	AAC	CT/	ACTO	TCA	CGG	AGT	TGCA	ecco	CAGC	GACI	ACT	ATG		CTA	IGT/	ATG	ACA	ATT/	AC	AAGC	TCA	TTA	AGC	:AAG	GCA	CAGAT	70
	1	1		I.	1	11	1 11				11		111	11	11	111	111	111	1	1	1		п	1 11		I	1	111	111		11	1	111	Т	ŧΠ.	ш	н	11	11	11	- 11	11	1.111	
ccc	TGG	-	CAR	CGG	CTC	CAC	AGG	cc	TGT	CAC	CCM	GGA	GTG	GAC	TAT	GTG	TAC	AAG	ACC	CGA	TGG	TCA	AGG	TTCA	GCT	STCC	AAT	ACT	TTGJ	CGA	GTA	CATC	CATG	CCC	ATTO	-	CAG	LCC3	TCI	ACT	CAG	GCT	CGGAT	4
Ŧ	т.	R	R	P	τ.	n	ĸ	2	c	R	P	G	v	D	¥	v	Y	ĸ	Ŧ	R	τ.	v	<b>K</b> 1	v o	ь	s	N	D	FT		Y	т	м		т	R	0	T	т	ĸ	s	G	S D	
-	-	-	-		-	-		-	-	-	-	-	•	-		•	-		-		-			•	_	-		-			-	-			-	-	-	-	-		-	-		
_			-		-	_	-		-	-			-		~		-					~			~			-			_	~	~			-	-		~	~		~	<b>v</b> <del>-</del>	
Б 		¥	Б			T	ĸ	N	F	±	5	н 	1			R			L.		R	¥				1	-	1		. v	T			- 		R	н 			1	•	3	1 1	
AGG	ATC	CT	GAG	GAC	AAG	ACA	CGCJ	AT	TTT/	ATC	IGCO	CAT	ATC	AAAT	TGC	CGA		GCT	TTA.	AATA	ITGC	AGC	TGA	ACCG	: هفا ه	ITAI	CTG	ATTT		T فأفاد	AAC	rGG:	rGAC	CTC	TGG	CGC	CAC	AT	(GA 1	CAT .	<b>r</b>	CCT	ACATC	. 8
			1	I.	11				11	1111	111	I.	111	11	111	11	11	11	1	11	11	11	11		11	11	11	111	1111	1 1	1		11	11	111		1	I		1		11	11111	
CACC	TG	AG	GTT	CCA	CVG	CYC	CGCI	ICC:	TTC	ATCJ	<b>NCC</b>	ccci	ATC	AAG:	TGC	AGA	GAA	SCC	CTG	ANG	TGG	AGG	ACA	YCYY	ACM	CTAC	CTC	ATGT	GCCC	STCT	CTC	CTC	CGAI	TTC	TGG	CCT	CYC	ULCC	:cc)	<b>UNCC</b>	TCA	GCI	ACATC	: 4
B	۷	Q	v	R	Q	Q	R	T	F	I	S	P	I	ĸ	с	R	E	X	L	ĸ	L	E	B I	ĸĸ	н	Y	L	M	W (	3 L	, s	S	D	F	W	G	E	ĸ	P	N	L	s	ΥI	
I	G	ĸ	D	T	W	M	E	W	W	P	N	B	R	R	с	Q	Q	R	E	N	Q	D	L	сD	D	F	E	T	v s	S D	N	L	R	I	v	G	с	P	N 3	FER				
ATTO	GGI	AG	GAC	ACA	TGG	ATG	GAGI	rGG:	TGG	ccci	AAT	GAG	AGG	GAG	TGT	CAA	CAA	CGT	GAG	AAC	AGG	ATC	TCT	ccca	TGA	CTTI	GAG	ACAG	TTT	CTGA	CAA	CCT	GGAG	ATT	GTT	GGT	TGT	cca	AC	rga(	GA	ACC	TAAAT	TCA
	ш	111		11	ш	11	111		н.	11	Т.	ш		11	11	111	1		(1)	111	111	1.1	1		11	11	1	1	1	1 11	1	11	11	1	11		н.			ш	1	111	1 11	11
	200		CAC	207	TGG	GTO	G	-20	TCC			620	cac	 'GBA'					C.A.C	220	202		, 724	CC2	662	CCTC	ccr	crc=				 					 TCC		20	TCP/	- -	200	· · · ·	TCC
					****	~	-				-		-	week.																														
T	~		•			**							•		~	^	n		72	N	^		<u>~</u>	<b>~</b> ~		T	0						17	17		~	~	<b>n</b>	N 4	100				

GACCANAATGTTTCCATCTTCCCAAGGTCATCTGTGAGTGTTTAAATGTTGTTAATGTGTTCAGTGTTTTATAATACAGTTCCAATAAAAGCATTTATTATGCTTAAC Poly A

FIG. 2. Alignment of the *Xenopus* and human C3 nucleotide sequences. Spaces represent gaps introduced to improve alignment. The translated amino acid sequences are in single-letter code with the nucleotide sequences.

and LG6, have identical C3 restriction fragment patterns but have different MHC haplotypes. Since the restriction fragment length polymorphisms do not segregate with the MHC haplotypes, the C3 gene is not linked to the MHC. A similar lack of cosegregation was observed with 18 other MHC-typed sibling frogs (data not shown).



FIG. 3. Comparison of cloned *Xenopus* C3 with human C3 mRNA and protein. The regions of *Xenopus* C3 represented in the clones C3-27 and C3-39 are shown aligned with the human mRNA and protein. The human protein is synthesized as a single-chain precursor, which is cleaved into  $\alpha$  and  $\beta$  chains to form the mature protein. I, factor I cleavage sites; RGD, Arg-Gly-Asp; P, properdin binding site; C3a, C3d, and C3g, proteolytic fragments of C3 ( $M_r \times 10^{-3}$  indicated) as described (2).

1087

XENC3 MUSC3 RABC3 HUMC3 HUMC4 MUSC5 HUMA2M	GRVTVYHALVIEKERKCLNFELSCECKGST-TCKTSRSAKATVSIEA A-RHLKNDDTMSII ISMMTGFSPDTDSLDR MKG LS.AKLKS.VTKK.D.RVSIRPAPE A.KPEE.N.MFL.IT-KM.GDL
XENC3 MUSC3 RABC3 HUMC3 HUMC4 MUSC5 HUMA2M	Properdin binding +9+10+12+14+15+16+17 W KYISKYEVNKGANDKGTLILYLDKVSHTDEE VK YAHQYFEVGFICHASVTVDD YTPDNR TKFHVEEGSAL AGI QGD IR A .RM.AFSN.N.I.E.I.TE.DLITKVN.LG.K.S.NLEES.R.P.KDDGW SKLHSEM .RL.AFSN.N.I.I.S.RALAKVN.LGA.K.S.NLEES.R.P.KDDGW SKLHSEM .RL.AFSN.N.I.S.SEDDLAKVN.LGA.K.AS.NLEES.R.P.KEDGW SKLHSEM S.R.V.HF.TEPHV.LYFDSVPTSRGE.V.EVP.LVA.LN.ER.SV.GAPSK.RTLSAEVCO. .QLLTD.QIKD.HVQ.NSIPSR.FL.RRIFEL.QLNM.TFEHRKQ.MITSISDTRQKVE.AA TV .KMLERSNH.SRTEVSSNHVLINQTLSLF-TVL.DVP.RDLK.I.KET.EFAIAE NAPCSKD.GNA
XENC3 MUSC3 RABC3 HUMC3 HUMC4 MUSC5	<pre>+18+19+20+21+23+24+25+25+25+25+25+25+25+25+25+25+25+25+25+25+25</pre>
XENC3 MUSC3 RABC3 HUMC3 HUMC4 MUSC5	26+27+28+30+31+32 RDF IW VTGDLWRHDGY-S-HIIGKDT WWPNERE QQRENQDLCDDEFTVSDNLEIVCPN KKWM.LSSGEKPNTVIHEAEDQKY.KQ.EEELQAFTESMVVY KHWM.LSS.F.GEKPNTVIKEDEKH.E.LGAFAESMVVF KHWM.LSS.F.GEKPNLVIKEDEKQ.JGAFTESMVVF KHWM.LSS.F.GEKPNLVIKEVIKE KHWM.LD.STYDLEH-PCILDSNS ILEM.S.L RSTRQRAA.AQLNDFLQEYGTQ.QV KQM KEVLQIK.NFSFKI-PLDSS.IIYTDTT PSCQAFVENLNNFAEDLF.NSE

RGD

## DISCUSSION

The initiation of the alternative pathway of complement activation, which occurs in the absence of antibody, relies on the slow but continuous activation of C3 (2). During complement activation the C3b fragment of C3 can be fixed on self and non-self surfaces, but in the case of self it is inactivated by widely distributed down-regulatory molecules (32). The alternative pathway may thus be a very ancient form of immunological defense, pre-dating the evolution of antibodies.

Xenopus C3 has the same subunit structure as mammalian C3 (7). This structure is also shared by the C3 from bird (33), reptile (34), and fish (35) but not by the C3 from lamprey, which has a three-subunit structure (36). C3b binds cell surfaces via its thioester site, and this property was utilized to purify Xenopus C3 bound to zymosan, which in turn was used to prepare a rabbit anti-Xenopus C3b antibody. This antibody was used to isolate an expression clone representing Xenopus C3 from a liver cDNA library. Liver is the major site of C3 synthesis in mammals (37). The anti-C3 antibody, which was affinity-purified on protein expressed by the C3-27 clone, recognized the same fragments of C3 on immunoblot strips as the antibody purified on C3 fragments fixed to zymosan (Fig. 1B, lanes 2 and 3). This verified that the clone indeed represents Xenopus C3. The  $M_r$  115,000 protein represents the  $\alpha'$  chain of C3b, and the  $M_r$  34,000 protein represents the C-terminal end of the  $\alpha$  chain of iC3b (7).

The sequence of *Xenopus* C3 has significant nucleotide and amino acid similarity with the corresponding sequence of

 Table 1. Amino acid conservation between Xenopus C3 and other related proteins

	% identity													
	Human C3	Mouse C3	Rabbit C3	Human C4	Mouse C5	Human α <sub>2</sub> M								
Xenopus C3	48	50	49	25	28	22								
Human C3		76	76	21	24	20								
Mouse C3			78	22	25	20								
Rabbit C3				22	24	18								
Human C4					17	14								
Mouse C5						13								

Values are percentages of identities for all positions that are matched in the alignment shown in Fig. 3.

FIG. 4. Comparison of the deduced Xenopus C3 (XENC3) amino acid sequence with sequences of mouse C3 (MUSC3) (23), rabbit C3 (RABC3) (24), human C3 (HUMC3) (25), human C4 (HUMC4) (26), mouse C5 (MUSC5) (27), and human  $\alpha_2 M$  (HUMA2M) (28). The numbering is from the Xenopus sequence. The first amino acid in each other case is as follows: mouse C3, 1342; rabbit C3, 405; human C3, 1342; human C4, 1403; mouse C5, 1322; human  $\alpha_2 M$ , 1313. The box labeled "RGD" encloses the Arg-Gly-Asp sequence of human C3 and the corresponding squences of the other C3 sequences. The boxes labeled "properdin binding" enclose the human C3 regions found to bind properdin (29). Identities with the Xenopus sequence are shown by dots, and gaps are shown by dashes. Residues shown in whiteon-black type are conserved in all complement sequences aligned.

human (25), mouse (23), and rabbit (24) C3, the three species for which C3 sequence is available. Alignment of the *Xenopus* C3 sequence with these sequences shows clusters of sequence conservation. Although present, these conserved clusters are not as apparent when mammalian sequences are compared among themselves, presumably because the mammalian species from which C3 sequence is available are too closely related. One of these conserved clusters correlates quite closely with the region of C3b that mediates properdin binding. It is also interesting that the human RGD sequence is absent in *Xenopus* C3. The RGD sequence has been reported to be involved in iC3b binding to CR3 (30) and functions as a general attachment site of several adhesion molecules to their receptors (38, 39). Mammalian macrophages have CR3, as well as the more widely distributed C3b



FIG. 5. MHC non-linkage of the Xenopus C3 gene. Genomic Southern blots of DNA from MHC-typed X. laevis  $\times X$ . gilli hybrid siblings. DNA was digested with Pst I (Left) or EcoRV (Right). Leftmost lane in each blot shows HindIII fragments of phage  $\lambda$  DNA used as size markers.

receptor, CR1 (1). Since Xenopus macrophages bind Xenopus C3b/iC3b-coated sheep erythrocytes (9), either this binding was via the Xenopus equivalent of CR1 or the binding of Xenopus iC3b to the CR3 receptor does not require the RGD sequence. The 20 amino acids extending to the C terminus of the human RGD sequence are highly conserved in all C3 sequences. Since these residues were present in the peptides used to show binding to CR3 (30), it may be these, rather than RGD, that are important for CR3 binding. The other extended regions of conservation with the other C3 sequences (amino acids 56-74, 170-180, 244-254, and 279-286) may be involved in binding of the other C3 ligands. C3 has many natural ligands besides properdin, some of which are factors I, H and B, C5, CR1, CR2, CR3, CR4, and membrane cofactor protein (MCP) (1, 3). Ligands must evolve to match each other, and the necessity for dual evolution is expected to make change at these sites slower. This seems to have been the case for the site of properdin binding.

The Xenopus sequence is 2–3% more similar to C4, C5, and  $\alpha_2 M$  than are the other mammalian C3 sequences. Although this difference in similarity is small, it is present in all three sequences and suggests that Xenopus C3 has evolved more slowly than mammalian C3. The nucleotide alignment shows no stretches of sequence with sufficient nucleotide similarity expected to permit the isolation of this gene by hybridization with human probes. This may explain the inability to isolate a Xenopus C4 clone by low-stringency hybridization with human or mouse probes (D.G., unpublished observations).

C3, C4, C5, and  $\alpha_2 M$  are structurally related and are thought to have arisen by gene duplication from an ancestral gene (40). The most primitive vertebrate, the lamprey, has "C3" with a subunit structure similar to mammalian C4, and it has been suggested to represent the ancestor molecule of higher vertebrate C3 and C4 (41). The C4 gene is in the MHC of Xenopus (42), humans, and mice (43), although complement proteins have not been found to interact in any way with MHC proteins. The C3 gene is not in the MHC of humans (44) or mice (45). The genes for C4 and the other complement components factor B and C2 have been said to reside in the MHC as a result of an accidental translocation (46). But these genes may have evolved within the MHC, and the C3 gene may have translocated out. To determine whether the C3 gene of *Xenopus* is linked to the *Xenopus* MHC, a Southern blot was done with DNA of a family of X. laevis  $\times X$ . gilli hybrid siblings. The polymorphisms observed did not cosegregate with the MHC alleles, and C3 is therefore not linked to the Xenopus MHC, as is the case with mammals. The translocation event that separated these duplicated genes must therefore have antedated the *Xenopus*/mammal divergence.

We thank Drs. W. Hein and K. Karjalainen for critical reading of the manuscript. The Basel Institute for Immunology was founded and is entirely supported by F. Hoffman-La Roche Ltd., Co., Basel, Switzerland.

- 1. Lambris, J. D. (1988) Immunol. Today 9, 387-393.
- Pangburn, M. K. & Müller-Eberhard, H. J. (1984) Springer Semin. Immunopathol. 7, 163–192.
- Lambris, J. D. & Müller-Eberhard, H. J. (1986) Mol. Immunol. 23, 1237–1242.
- Fujii, T., Sekizawa, A. & Katagiri, C. (1985) Immunology 56, 743– 750.
- 5. Jurd, R. J. (1978) Immunology 34, 389-396.

- Dlabac, V., Simeckova, J., Sima, P. & Hofmanova, B. (1983) Dev. Comp. Immunol. 7, 783-784.
- Sekizawa, A., Fujii, T. & Katagiri, C. (1984) J. Immunol. 133, 1436– 1443.
- 8. Kaidoh, T. & Gigli, I. (1987) J. Immunol. 139, 194-201.
- Sekizawa, A., Fujii, T. & Tochinai, S. (1984) J. Immunol. 133, 1431– 1435.
- Becherer, J. D., Daoudaki, M. E. & Lambris, J. D. (1987) Fed. Proc. Fed. Am. Soc. Exp. Biol. 46, 771 (abstr.).
- Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W. J. & Goodman, H. M. (1977) Science 196, 1313-1319.
- Young, R. A. & Davis, R. W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194–1198.
- Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in DNA Cloning. A Practical Approach, eds. Rickwood, D. & Hames, B. D. (IRL, Oxford), Vol. 1, pp. 49-78.
- 14. Young, R. A. & Davis, R. W. (1983) Science 222, 778-782.
- 15. Towbin, H. T., Staehelin, T. & Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350-4354.
- 16. Chen, E. Y. & Seeburg, P. H. (1985) DNA 4, 165-170.
- 17. Needleman, S. & Wunsch, C. (1970) J. Mol. Biol. 48, 444-453.
- Schwager, J., Grossberger, D. & DuPasquier, L. (1988) EMBO J. 7, 2409-2415.
- Vogelstein, B. & Gillespie, D. (1979) Proc. Natl. Acad. Sci. USA 76, 615-619.
- 20. Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
- 21. Tack, B. F. (1983) Springer Semin. Immunopathol. 6, 259-282.
- 22. Laemmli, U. K. (1970) Nature (London) 227, 68-685.
- Wetsel, R. A., Lundwall, A., Davidson, F., Gibson, T., Tack, B. F. & Fey, G. H. (1984) J. Biol. Chem. 259, 13857–13862.
- Kusano, M., Choi, N. H., Tomita, M., Yamamoto, K.-I., Migita, S., Sekiya, T. & Nishimura, S. (1986) *Immunol. Invest.* 15, 365-378.
   deBruijn, M. H. L. & Fey, G. H. (1985) *Proc. Natl. Acad. Sci. USA*
- 82, 708–712.
   26. Belt, K. T., Carroll, M. C. & Porter, R. R. (1984) Cell 36, 907–914.
- 27. Wetsel, R. A., Ogata, R. T. & Tack, B. F. (1987) Biochemistry 26, 737-743.
- Kan, C.-C., Solomon, E., Belt, K. T., Chain, A. C., Hiorns, L. R. & Fey, G. H. (1985) Proc. Natl. Acad. Sci. USA 82, 2282–2286.
- 29. Daoudaki, M. E., Becherer, J. D. & Lambris, J. D. (1988) J. Immunol. 140, 1577-1580.
- Wright, S. D., Reddy, P. A., Jong, M. T. C. & Erickson, B. W. (1987) Proc. Natl. Acad. Sci. USA 84, 1965–1968.
- 31. Kobel, H. R. & Du Pasquier, L. (1975) Immunogenetics 2, 87-91.
- Atkinson, J. P. & Farries, T. (1987) Immunol. Today 8, 212–215.
   Kai, C., Yoshikawa, Y., Yamanouchi, K. & Okada, H. (1983) J.
- Immunol. 130, 2814–2820.
   Eggersten, G., Lundwall, A., Hellman, U. & Sjoquist, J. (1983) J.
- Eggersten, G., Lundwan, A., Heilman, U. & Sjoquist, J. (1983) J. Immunol. 131, 1920–1923.
   Namba M. Juckin M. Nakin G. Nakin M. Mathara, T. J.
- Nonaka, M., Iwaki, M., Nakai, C., Nozaki, M., Kaidoh, T., Natsuume-Sakai, S. & Takahashi, M. (1984) J. Biol. Chem. 259, 6327-6333.
- Nonaka, M., Fujii, T., Kaidoh, T., Natsuume-Sakai, S., Yamaguchi, N. & Takahashi, M. (1984) J. Immunol. 133, 3242-3249.
- 37. Alper, C. A., Johnson, A. M., Birtch, A. G. & Moore, F. D. (1969) Science 163, 286-288.
- 8. Ruoslahti, E. & Pierschbacher, M. D. (1986) Cell 44, 517-518.
- Simmons, D., Makgoba, M. W. & Seed, B. (1988) Nature (London) 331, 624–627.
- Sottrup-Jensen, L., Stepanik, T. M., Kristensen, T., Lonblad, P. B., Jones, C. M., Wierzbicki, D. M., Magnusson, S., Domdey, H., Wetsel, R. A., Lundwall, A., Tack, B. F. & Fey, G. H. (1985) Proc. Natl. Acad. Sci. USA 82, 9-13.
- 41. Nonaka, M., Fujii, T., Kaidoh, T., Natsuume-Sakai, S., Yamaguchi, N. & Takahashi, M. (1984) J. Immunol. 133, 3242-3249.
- 42. Nakamura, T., Sekizawa, A., Fujii, T. & Katagiri, C. (1986) Immunogenetics 23, 181–186.
- 43. Alper, C. A. (1981) in *The Role of the Major Histocompatibility Complex in Immunobiology*, ed. Dorf, M. E. (Garland STPM, New York), pp. 173-220.
- Whitehead, A. S., Soloman, E., Chambers, S., Bodmer, W. F., Povey, S. & Fey, G. H. (1982) Proc. Natl. Acad. Sci. USA 79, 5021– 5025.
- 45. Ferreira, A. & Nussenzweig, V. (1975) J. Exp. Med. 141, 513-517.
- 46. Klein, J. & Figueroa, F. (1986) CRC Crit. Rev. Immunol. 6, 295-386.