cDNA heterogeneity suggests structural variants related to the high-affinity IgE receptor

(immediate hypersensitivity/mast cell/rat basophilic leukemia cell/mRNA splicing)

FU-TONG LIU, KEITH ALBRANDT, AND MICHAEL W. ROBERTSON

Medical Biology Institute, Division of Molecular Biology, 11077 North Torrey Pines Road, La Jolla, CA 92037

Communicated by K. Frank Austin, April 4, 1988

ABSTRACT The high-affinity IgE receptor present on mast cells and basophils is responsible for the IgE-mediated activation of these cells. The current model for this receptor depicts a four-subunit structure, $\alpha\beta\gamma_2$. A cDNA for the α subunit was recently cloned and predicts a structure consisting of two homologous extracellular domains, a transmembrane segment, and a cytoplasmic tail. Using a synthetic oligonucleotide corresponding to the amino-terminal sequence of the α subunit, we identified a number of cDNA clones from a rat basophilic leukemia cell cDNA library. Nucleotide sequencing established four different forms of cDNA: one is nearly identical to the published cDNA; the second differs from the first in the 5' untranslated sequence: the other two forms use either one or the other of the 5'-end sequences as above and lack 163 base pairs in the region coding for the second extracellular domain. RNase protection analysis with radioactive RNA probes established the heterogeneity of rat basophilic leukemia cell mRNA with regard to both the 5' and the internal sequences. Our results suggest the existence of at least four different protein forms related to the α subunit of the highaffinity IgE receptor.

The high-affinity IgE receptor ($Fc \in R$ -I, which recognizes the Fc region of IgE) present on mast cells and basophils is a key component involved in IgE-mediated hypersensitivity reactions (1). Crosslinking of receptors with IgE-antigen complexes or anti-receptor antibodies leads to a series of cellular biochemical events culminating in the release of a variety of chemical substances responsible for the clinical manifestation of hypersensitivity. The study of IgE-mediated activation of mast cells and basophils has contributed to our understanding of the mechanisms of immediate hypersensitivity reactions and has provided a model system for the study of transmembrane signaling by receptor aggregation and of the cascade of events involved in cellular activation (2).

Fc ε R-I has been extensively characterized biochemically (3) and a tetrameric $\alpha\beta\gamma_2$ structure has been proposed (4): the α subunit is a glycoprotein of M_r 45,000 that is exposed on the extracellular surface; the β subunit (M_r 33,000) and the γ subunit (M_r 6000) are hydrophobic and are located on the inner portion of the plasma membrane. The α subunit of the receptor by itself accounts for IgE-binding activity, but the roles of the β and γ subunits in the assembly and function of the receptor are unknown. Both the β and the γ subunits are necessary for reconstitution of Fc ε R-I in model membranes (5), suggesting a possible supportive role of these subunits in the expression of the receptor on the cell surface.

The amino acid sequence of the α subunit of Fc ϵ R-I (Fc ϵ R-I α) has been deduced through cDNA cloning (6) and predicts a 180-amino acid extracellular portion with two homologous domains, a 20-amino acid transmembrane seg-

ment, and a 27-amino acid cytoplasmic carboxyl terminus. The identity of the cloned cDNA was confirmed by correlating the predicted protein sequence with the sequences of a number of peptic fragments of $Fc\epsilon R$ -I α (6). In addition, antibodies to synthetic peptides corresponding to portions of the predicted sequence were found to recognize native $Fc\epsilon R$ -I α (7). $Fc\epsilon R$ -I α shares significant sequence homology with an IgG receptor for which cDNA has been cloned and sequenced (8–10).

We prepared a synthetic oligonucleotide corresponding to a portion of the published sequence of $Fc \varepsilon R$ -I α and used it to screen a rat basophilic leukemia (RBL) cell cDNA library. By sequence analysis we found structural heterogeneity of cloned cDNA at both the 5' end and an internal region.* The existence of specific RBL cell mRNA corresponding to these various forms was confirmed by RNase protection analysis. The results predict the existence of at least four different protein products related to $Fc\varepsilon R$ -I α .

MATERIALS AND METHODS

cDNA Cloning, Screening, and Sequencing. The RBL cell line (11) was provided by H. Metzger (National Institutes of Health). Extraction of total cytoplasmic RNA from RBL cells by the phenol/chloroform method and isolation of $poly(A)^+$ RNA by oligo(dT)-cellulose affinity chromatography have been described (12). A cDNA library was prepared (13) from RBL cell mRNA. Escherichia coli MC1061 transformants were screened by colony hybridization (14) with a ³²P-labeled 30-mer oligodeoxyribonucleotide (3'-AGACATCACAGGA-ACCTGGGTGGCACCTAA-5') corresponding to nucleotides 108–137 of the Fc ε R-I α cDNA sequence previously reported (6). Hybridization was at 45°C in 0.75 M NaCl/75 mM Tris HCl, pH 8.0/5 mM EDTA/0.04% bovine serum albumin/0.04% Ficoll/0.04% polyvinylpyrrolidone/0.1% NaDodSO₄ containing 100 μ g of heat-denatured, sonicated salmon sperm DNA per ml and was followed by washing at 65°C in 0.03 M NaCl/3 mM sodium citrate, pH 7.0/0.1% NaDodSO₄ and autoradiography. DNA was sequenced by dideoxy chain-termination (15) with modified phage T7 DNA polymerase (Sequenase; United States Biochemical, Cleveland) and by chemical cleavage (16).

RNase Analysis with RNA Probes. Restriction fragments of one of the cDNA clones (R8-2b) were cloned into plasmid pSP64 or pSP65, and ³²P-labeled RNA was synthesized as described (ref. 17; Ribosystem II, Promega Biotec, Madison, WI) and purified by electrophoresis in a 5% acrylamide gel. Hybridizations were performed by mixing 5×10^5 cpm of ³²P-labeled RNA probe with 1 µg of RBL cell poly(A)⁺ RNA

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.

Abbreviations: RBL, rat basophilic leukemia; Fc ϵ R-I, high-affinity IgE receptor [which recognizes the Fc ("crystallizable fragment") region of IgE]; Fc ϵ R-I α , α subunit of Fc ϵ R-I.

^{*}The sequences reported in this paper are being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03811).

in 30 μ l of hybridization solution [80% (vol/vol) formamide/ 400 mM NaCl/40 mM Pipes, pH 6.4/1 mM EDTA). The solution was heated at 85°C for 10 min and then immediately transferred to 50°C and incubated for 10 hr. After hybridization, the mixture was treated with RNase A (40 μ g/ml) and RNase T1 (2 μ g/ml) in 350 μ l of 300 mM NaCl/10 mM Tris·HCl, pH 7.5/5 mM EDTA at 15°C or 30°C for 0.5–2 hr (the specific temperature and time are indicated in the legend of Fig. 2). Digestion was terminated by adding 10 μ l of 20% (wt/vol) NaDodSO₄ and 2 μ l of a 25-mg/ml solution of proteinase K, and the mixture was further incubated for 15 min at 37°C. After phenol/chloroform (1:1, vol/vol) extraction, RNA was precipitated together with 5 μ g of carrier tRNA, washed with 70% (vol/vol) ethanol, and electrophoresed in a 6% acrylamide/8 M urea gel.

RESULTS

Structural Heterogeneity of cDNA at the 5' End. A 30-mer corresponding to the nucleotide sequence (6) coding for the first 10 amino acids of the mature Fc ε R-I α was used to screen a cDNA library derived from RBL cells. From $\approx 7 \times 10^5$ colonies, 73 positive clones were identified, from which 18 were randomly selected for further analysis. The cDNA-insert length for all 18 clones ranged from ≈ 1000 to ≈ 1200 base pairs (bp). Two of the longer ones (R8-2b and R3-3) were sequenced for comparison with the published (6) Fc ε R-I α cDNA sequence.

The sequence of clone R8-2b was found to be nearly identical with the $Fc\epsilon R-I\alpha$ cDNA sequence (Fig. 1); however, the two sequences differ at the 5' end, resulting in two distinct 5' untranslated sequences and different translation initiation sites. The initiation codon used in the published cDNA and an additional 46 bp coding for a major portion of the predicted signal-peptide sequence are not present in clone R8-2b. Instead, the first potential initiation codon in clone R8-2b corresponds to that of the third-to-the-last amino acid of the signal peptide of $Fc \epsilon R - I \alpha$. Clone R8-2b also contains an additional 98 bp of 3' untranslated sequence that contributes a third AATAAA poly(A)-addition signal (Fig. 1).

In addition to the above sequence divergence, there are a few minor sequence differences between clone R8-2b and the published Fc ε R-I α clone (see legend to Fig. 1). The most significant of these is the omission of an adenine residue after nucleotide position 848, resulting in a glycine residue at the carboxyl terminus (GGT codon followed immediately by a termination codon) instead of the -Arg-Leu-Lys-Pro-Asn-Ser sequence reported previously (6).

Identification of a cDNA Clone with Deletion of an Internal Sequence. Another clone, R3-3, differed slightly from R8-2b by restriction endonuclease mapping. Complete nucleotide sequence analysis of this clone revealed the deletion of an internal sequence (163 bp) coding for a major portion of the second extracellular domain of $Fc \in R-I\alpha$. The deleted sequence contains GT and AG dinucleotides at its 5' and 3' ends, respectively, and thus appears to be a possible intronlike sequence. Significantly, the deletion results in alteration of the reading frame with a new termination codon 101 bp 5' to that found in R8-2b (Fig. 1). This results in a new sequence of 17 amino acid residues at the carboxyl terminus of the predicted R3-3 protein and a loss of the transmembrane sequence of Fc ϵ R-I α . Thus, the R3-3 protein contains 137 amino acid residues instead of the 225 found for R8-2b. In addition, the predicted sequence lacks a long stretch of hydrophobic sequence. The 5'-end sequence of clone R3-3 is identical to that of R8-2b and is different from that published

a)	ATGO	GAGG	AGGC	TAGA	GGAA	AGAG	AGCT	CCTG	стсса	ACAT	GATC	ATTA	AAGC	TCCA	GGGA	IGTG	A GGTC	TTAT. CCAT	AGAC. FCCC.	AGATGCCA ATTCCGAT	17 100
b) c)	TTTT	TCC:	TCT	CCTT	CGTG	GTTT	CAAG	CCTA	CTTT	rgaa	GCCG	FAGC	FCAC	T TGG:	GCAG ::::	FTAG	CACC	-GAA T:::	GGCA	CGGGCA :AG:::::	
a)	CCCZ	AGAG	CTCT	GCA	TCGA	GGCA	CACT	GGGG	ACAG	ACAC	TAAC	FTAA	CCTT	TATT	GATT	TTGC	AGCT	CTG	GGT	GTC	176
b)	ATG	GGT	ACC	GGA	GGA	TCT	GCC	CGG	CTG	TGC	CTA	GCA	TTA	GTG	CTC	ATA	т::	:::	:::	:::	
c)	:::	:A:	::T	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	:::	
								_	_	-	_		_		_		-	_			
b)	MET	Gly	Thr	Gly	Gly	Ser	Ala	Arg	Leu	Cys	Leu		Leu		Leu	TTe	Ser	Leu	GLY		
C)	*	Asp	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	
a١	ATC	ста	202	acc	аст	CAG	***	TCT	ста	GTG	TCC	ττG	GAC	CCA	CCG	TGG	АТТ	AGA	ата	CTT	236
α,	Met	Len	Thr	Ala	Thr	Gln	Lvs	Ser	Val	Val	Ser	Leu	Asp	Pro	Pro	Tro	Ile	Arg	Ile	Leu	200
		200				•=	-10														
	ACA	GGA	GAT	ААА	GTG	ACT	CTT	ATA	TGC	AAT	GGG	AAC	AAT	TCC	TCT	CAA	ATG	AAC	TCT	ACT	296
	Thr	Gly	Asp	Lys	Val	Thr	Leu	Ile	Cys	Asn	Gly	Asn	Asn	Ser	Ser	Gln	Met	Asn	Ser	Thr	
	AAA	TGG	ATC	CAC	AAT	GAT	AGC	ATC	TCT	AAT	GTG	AAA	TCG	TCA	CAT	TGG	GTC	ATT	GTG	AGT	356
	Lys	Trp	11e	HIS	ASN	Asp	Ser	IIe	Ser	ASN	vai	Lys	Ser	Ser	HIS	тр	vai	11e	var	Ser	
	000	ACC	атт	CAA	GAC	AGT	GGA	ААА	TAC	АТА	TGT	CAG	AAG	CAA	GGA	ттт	ТАТ	AAG	AGC	ААА	416
	Ala	Thr	Ile	Gln	Asp	Ser	Gly	Lys	Tyr	Ile	Cys	Gln	Lys	Gln	Gly	Phe	Tyr	Lys	Ser	Lys	
					•		-	-	-		-		-		-		-	-		-	
	CCT	GTG	TAC	TTG	AAC	GTG	ATG	CAA	GAG	TGG	CTG	CTG	CTC	CAA	TCT	TCT	GCT	GAC	GTG	GTC	476
	Pro	Val	Tyr	Leu	Asn	Val	Met	Gln	Glu	Trp	Leu	Leu	Leu	Gln	Ser	Ser	Ala	Asp	Val	Val	
	ጥጥእ	CNC	***	CCN	TCC	THE	CAC	ልጥሮ	ACA	TCC	COT	200	TCC	MAG	***	TCC	***	GTC	CAC	AAC	536
	Len	Asp	Asn	Glv	Ser	Phe	Asp	Tle	Ara	Cvs	Ara	Ser	Tro	LVS	Lvs	Tro	LVS	Val	His	LVS	550
,														-1-	-1-		-10				
	GTG	ATC	TAC	TAC	AAG	GAC	GAC	ATT	GCT	TTC	AAG	TAC	TCT	TAT	GAC	AGC	AAC	AAC	ATC	TCC	596
	Val	Ile	Tyr	Tyr	Lys	Asp	Asp	Ile	Ala	Phe	Lys	Tyr	Ser	Tyr	Asp	Ser	Asn	Asn	Ile	Ser	
	አጥጥ	202	220		202	(THE TOP IN THE TOP IN	እእጥ	CNC	ACT	ccc	ACC	TAC	CNC	TCC	ACA	ccc	ጥአጥ	TTC	330	220	656
	Tle	Ara	LVS	Ala	Thr	Phe	Asn	Asp	Ser	Glv	Ser	Tvr	His	Cvs	Thr	Glv	Tvr	Leu	Asn	LVS	0.00
			-10							1		-1-		-1-	_		-1-			-10	
	GTT	GAA	TGT	AAA	TCT	GAT	AAA	TTC	AGT	ATT	GCT	GTA	GTA	AAA	GAT	TAC	ACA	ATT	GAG	TAT	716
	Val	Glu	Cys	Lys	Ser	Asp	Lys	Phe	Ser	Ile	Ala	Val	Val	Lys	Asp	Tyr	Thr	Ile	Glu	Tyr	
															-110	e Th	r GL	n Le	a Se	r Ile	
	CGT	TGG	СТА	CAA	CTC	АТТ	TTC	CCA	TCA	TTG	606	GTG	ልተተ	CTG	ттт	GCT	GTG	GAC	аст	666	776
	Arg	Trp	Leu	Gln	Leu	Ile	Phe	Pro	Ser	Leu	Ala	Val	Ile	Leu	Phe	Ala	Val	Asp	Thr	Glv	
Val Gly Tyr Asn Ser Phe Ser His His Trp Arg																					
	TTA	TGG	TTC	TCA	ACC	CAC	AAA	CAG	TTC	GAA	TCC	ATC	TTG	AAG	ATT	CAG	AAG	ACT	GGA	AAA	836
	Leu	Trp	Phe	ser	Thr	HIS	LYS	GIN	Phe	GIU	Ser	11e	Leu	rys	Ile	Gln	ŗða	Thr	Gly	Lys	
	GGC	ААА	ааа	ААА	GGT	TGA	AACC	TAAC	CTT2	AACC	AAGG	TATA'	[AAA	GGAA	CTAA	IGTC/	ATCG	CTTA	AGAG	ACAA	910

GGC AAA AAA GGT TGAAACCTAACTCTTAACCAAGGTATATAAAGGAACTAATGTCATCGCTTAAGAGACAA Gly Lys Lys Gly

TTCTTAACAATTATTTCCCACAGTATCTTCAATAGCCTTTCAACTGTCAAAGGACACTCATGTTATCCATAGAAATGTC TGTACCCCAGGAATTGCATAAATGCTTCATTAAACCAACAGCAGCTGGTTAAGTGACATGCAATAAAAAAATACCA AATAAACACTGGTTTAAGAAGT<u>TAAAGAATGACAGGAGGATGGATGCTATATCTGTAAAAGAGAAGTACAATTTGAATA</u> AAGTATGGAATCATGTAGAAT<u>GTACTTTTAC</u>

FIG. 1. Nucleotide sequence and predicted amino acid sequence of various forms of FceR-I α cDNA. The top seven lines include 5' un-76 translated sequence and leader-coding sequence of the various $Fc \in R$ -I α -related cDNAs. Lines a, clones R8-2b and R3-3; lines b, the published cDNA (6); lines c, clones R3-4 and R6-1 (the first nucleotide for R3-4 is 31 nucleotides downstream of R6-1). The eighth and ninth lines represent deduced amino acid sequences from the nucleotide sequences in the sixth and seventh lines, respectively. From the tenth line down are the nucleotide and amino acid sequences corresponding to clone R8-2b. Colons indicate nucleotide sequence identity, and stars amino acid identity, with the sequence immediately above; dashes represent gaps introduced for alignment. The initiation codons of the various forms are underlined and the putative initiation methionine for R8-2b is italicized. The putative transmembrane domain (amino acids 184-203) is underlined. Nucleotide sequence at the 3' end not previously found in $Fc \in R-I\alpha$ (6) is underlined. Other differences between clone R8-2b and the published $Fc \in R-I\alpha$ cDNA sequence: base 770 is cytosine instead of thymine (silent mutation); base 827 is guanine instead of cytosine, resulting in prediction of lysine instead of asparagine; an adenine residue is omitted after nucleotide 848, resulting in a change in the carboxyl-terminal sequence of the protein (see text); thymine is omitted after nucleotide 1085; CATTT is omitted after nucleotide 1089. The sequence deleted from clones R3-3 and R6-1 cDNA is boxed. The change in amino acid sequence downstream of this deletion due to the change in reading frame is indicated in italics.

Numbers at right refer to nucleotides.

989

1068

1147

for Fc ε R-I α cDNA (6). Thus, both the R3-3- and the R8-2bpredicted protein lack a signal peptide.

Presence of at Least Four Different Forms of $Fc \in R-I\alpha$ cDNA. The isolation of two cDNA clones containing structural features distinct from the published cDNA sequence prompted us to survey other cDNA clones for the existence of additional forms. Oligonucleotide probes corresponding to $Fc \in R-I\alpha$ sequences (6) in the predicted signal peptide, extracellular domain 2, and cytoplasmic tail were synthesized and used to analyze the 18 cDNA clones mentioned above by dot blot hybridization. The cDNA clones fall into four groups (Table 1). Four clones represented by R8-2b lack the sequence coding for the signal peptide; clone R3-3 is unique in that it also lacks the sequence recognized by the extracellular-domain-2 probe. Twelve of the 18 clones, represented by R3-4, were found to hybridize with all probes and were considered to correspond to the published sequence. One clone (R6-1) hybridized with the signal-peptide probe but not with the extracellular-domain-2 probe. This result suggests that clone R6-1 contains the 5'-end signal sequence of $Fc \in R-I\alpha$ but has an internal deletion analogous to that of R3-3.

The 5' portions of clones R3-4 and R6-1 were subsequently sequenced (Fig. 1). Both clones have a 5' untranslated sequence and an amino-terminal signal-peptide sequence nearly identical to the published sequence. Internal regions of clone R6-1 were also sequenced, and the results confirm the deletion of an internal sequence, identical to that deleted in R3-3. Sequencing of internal regions of clone R3-4 confirmed its identity with the published sequence of $Fc\epsilon R-I\alpha$ cDNA (6).

The above analysis establishes the existence of two types of variations in cDNA coding for $Fc_{\varepsilon}R$ -I α -related sequences: one near the 5' end and one in the second extracellular domain coding region. All four possible combinations of these two variations have been identified in our cDNA library.

Structural Heterogeneity of $Fc\epsilon R$ -I α -Related mRNA. The finding of different forms of $Fc\epsilon R$ -I α cDNA suggested structural heterogeneity of mRNA. However, the possibility of cloning artifacts needed to be excluded. To address this question, we tested the ability of RBL cell mRNA to protect radioactive RNA probes from RNase digestion. For the detection of 5'-end sequence heterogeneity, we used an RNA probe derived from a *Pst I–Bam*HI fragment that spans the 5'-most sequence of the R8-2b cDNA clone (Fig. 2C). This fragment contains 140 bp unique to R8-2b and 165 bp common to R8-2b and R3-4. The probe [which contains some vector sequence and a poly(C) stretch, introduced during the cloning procedure] was hybridized to poly(A)⁺ RNA from RBL cells and then treated with RNases A and T1. Two major bands

Table 1. Grouping of cDNA clones by their pattern of hybridization with oligonucleotide probes representing various domains of $Fc\epsilon R$ -I α

Clor	nes	Hybridization with probe								
Туре	No.	SIG	EC-1	EC-2	CYT					
R3-4	12	+	+	+	+					
R6-1	1	+	+	_	+					
R8-2b	4	_	+	+	+					
R3-3	1	_	+	-	+					

The sequence (3' to 5') and nucleotide positions of each oligonucleotide probe are as follows: SIG probe (signal peptide), TACC-CATGGCCTCCTAGACGGG, nucleotides 27–48; EC-1 probe (extracellular domain 1), AGACATCACAGGAACCTGGGTGG-CACCTAA, nucleotides 198–227; EC-2 probe (extracellular domain 2), TTACTGTCACCGTCGAT, nucleotides 524–541; CYT probe (cytoplasmic domain), GGTAGAACTTCTAAGTC, nucleotides 718–734.



FIG. 2. RNase protection analysis with RNA probes. Autoradiographs of 6% acrylamide/8 M urea gels show RNA probes protected from RNase digestion by RBL cell poly(A)⁺ RNA. (A) ³²P-labeled RNA probe was the transcription product of pSP65 containing a 321-bp Pst I-BamHI fragment of R8-2b cDNA [nucleotides 1-305 in Fig. 1 plus (dG)₁₆] linearized at the Pvu II site in the vector. Lane 1, RNA probe alone. Note that the size of the probe is shorter than predicted (i.e., cDNA plus vector sequence) probably because of premature termination of transcription at the poly(dG) tail introduced at the 5' end of the R8-2b cDNA during the cloning procedure and represented in this fragment. Lane 2, RNA probe protected by hybridization with RBL cell poly(A)⁺ RNA and subsequently treated with RNases for 1 hr at 30°C. Lane 3, unprotected probe treated with RNases as in lane 2. Lane 4, mixture of RNase-treated protected RNA probe (lane 2 products) and undigested probe (lane 1 products). Autoradiography was at -20° C for 72 hr. Size markers were ³²P-end-labeled Hpa II fragments of pBR322 (lengths in bp at left; not all bands are indicated). (B) RNA probe was the transcription product of pSP64 containing an 83-bp BamHI-Rsa I fragment of R8-2b cDNA (nucleotides 486-569, Fig. 1) linearized at the Pvu II site in the vector. Lane 1, RNA probe alone (83 bp of cDNA plus 220 bp of vector sequence). Lanes 2, 4, and 6, RNA probe protected by hybridization with RBL cell poly(A)⁺ RNA and subsequently treated with RNases for 0.5, 1, and 2 hr, respectively, at 15°C. Lanes 3, 5, and 7, unprotected probe treated with RNases for 0.5, 1, and 2 hr, respectively, at 15°C. Autoradiography was at 70°C for 48 hr (with intensifier). Size markers were as described for

A. (C) Corresponding locations of the RNA probes employed above. Probe a, *Pst* I-*Bam*HI fragment; probe b, *Bam*HI-*Rsa* I fragment. Selected restriction endonuclease sites in the four cDNAs are indicated: P, *Pst* I; B, *Bam*HI; R, *Rsa* I; Q, *Taq* I. The filled and open boxes indicate the two different 5'-end sequences of these cDNAs. The hatched boxes represent the internal sequence deleted in R3-3 and R6-1 cDNAs. were detected—one of 305 bases and the other of 140 bases (Fig. 2A, lane 2). Neither band was detected when the probe alone was treated with RNases (lane 3). The 305-base band was distinct from the probe itself, as it could be separated from the RNA probe (lane 4).

Protection of a 140-base fragment is exactly what would be predicted to result from the presence of mRNA corresponding to R3-4 cDNA (i.e., corresponding to the sequence previously identified for Fc ϵ R-I α mRNA). In addition, the presence of a fully protected species (305 bases) suggests the presence of mRNA corresponding to R8-2b cDNA. Densitometric analysis of these protected bands normalized to fragment length gave a ratio of 3.5:1, close to the ratio of the numbers of cDNA clones isolated from our library containing these respective 5'-end sequences (i.e., no. of R3-4 + R6-1 clones/no. of R8-2b + R3-3 clones).

The presence of mRNA corresponding to cDNA with the internal deletion was established by a similar approach using an RNA probe derived from a BamHI-Rsa I fragment of clone R8-2b. This fragment contains 50 bp of sequence common to all the cDNA clones and 33 bp of the sequence that is deleted in clones R3-3 and R6-1 (Fig. 2C). As expected, a major band of 83 bases was detected (Fig. 2B, lanes 2, 4, and 6), which represents RNA probe fully protected by full-length mRNA. However, a minor band of ≈ 50 bases was also detected, consistent with the presence of mRNA with an internal deletion. The absence of this band in the samples from RNase treatment of the RNA probe itself (Fig. 2B. lanes 3, 5, and 7) indicates that this band does not result from the partial digestion of unprotected RNA. Again, the ratio of the intensities of the two bands normalized to fragment length (8.4:1) was consistent with the number of cDNA clones isolated containing the respective sequences. (The additional band just above the major 83-base band in Fig. 2B most likely resulted from the protection of a few extra bases of the RNA probe corresponding to plasmid vector that, by chance, has some sequence similarity to the mRNA.)

DISCUSSION

We have identified four different forms of cDNA coding for sequences related to Fc ε R-I α from a RBL cell cDNA library. The major form (12 clones) is nearly identical to the published cDNA (6), and the second most abundant form (4 clones) differs only at the 5' end, including the entire 5' untranslated sequence. Two additional forms are shorter versions of the above two types of cDNA, resulting from the deletion of 163 bp of internal sequence. Analyses of RNase protection of RNA probes strongly support the structural heterogeneity of Fc ε R-I α mRNA at the 5' end and provide evidence for the presence of mRNA with an internal deletion as predicted from the cDNA clones.

Our results suggest the presence of four different forms of protein products related to $Fc\epsilon R-I\alpha$ (Fig. 3). The first form (represented by clone R3-4) corresponds to the published (6) protein structure deduced from the cDNA sequence and

contains a signal peptide, two extracellular domains, a transmembrane sequence, and a cytoplasmic tail. The second form (R6-1) differs from the first one by deletion of 54 amino acids in the second extracellular domain. In addition, the deletion alters the reading frame, resulting in a truncation and alteration of the amino acid sequence that follows the deleted region. The original transmembrane and cytoplasmic sequences are not present, and the new sequence does not appear to contain a transmembrane sequence. The third form (R3-3) and the fourth form (R8-2b) differ from R6-1 and R3-4, respectively, essentially by the lack of the signal peptide.

The presence of mRNAs differing at the 5' end (e.g., the two forms corresponding to R8-2b and R3-4 cDNA) may result from transcription of a single gene at two initiation sites, followed by differential splicing. The complete identity of the coding sequence following the 5'-sequence divergence makes the possibility of two separate genes less likely. Genes using multiple transcription initiation sites and differential splicing have been well documented (reviewed in refs. 18 and 19). One remarkable aspect of these forms of $Fc \in \mathbb{R}$ -I α -related mRNAs, however, is that one of the predicted translation products differs from the other only in lacking the signal sequence. This is analogous to the case of veast invertase (20, 21), for which two mRNAs with different 5' ends have been found. Specifically, one codes for the precursor of secreted invertase and contains a signal sequence, whereas the other is a shorter form that starts at a downstream initiation codon and codes for an intracellular protein lacking a signal peptide. The situation for $Fc \in R$ -I α is somewhat different in that an alternative 5' untranslated sequence is used instead of a truncation of the same sequence. The lack of a signal sequence would preclude protein transport and expression on the cell surface, although we cannot exclude the possibility that internal sequences may be recognized as a signal, especially because the specificity of signal-sequence recognition may be low (22)

The internal deletion seen in two cDNA clones (R3-3 and R6-1) is also likely a result of differential splicing. Whether the deleted sequence represents a separate exon that is alternatively spliced remains to be determined. The presence of GT and AG dinucleotides at the 5' and 3' ends, respectively, of the deleted sequence suggests that it has the potential of being removed by a mechanism similar to that used for introns. The presence of different forms of mRNA as a result of differential splicing of internal sequences has been found for numerous proteins, including a number of other important components of the immune system (19). One example is the human interleukin-2 receptor (23), for which two cDNAs have been found, one with the deletion of a 216-bp segment containing GT and AG dinucleotides at its 5' and 3' ends, respectively. Another example is the finding of two different forms of mRNA coding for rabbit poly(immunoglobulin) receptor. A smaller mRNA results from alternative splicing and codes for a protein that lacks two of the five receptor domains (24). In this case, a protein product with the internal deletion has been isolated and characterized.



FIG. 3. Four different protein forms related to $Fc_{\mathcal{E}}R$ -I α predicted from cDNA. SIG, signal sequence; EC-1, extracellular domain 1; EC-2, extracellular domain 2; TM, transmembrane segment; C, cytoplasmic tail. Wide bars indicate coding sequences and narrow bars represent noncoding sequences. Identical shading indicates identical sequence. Unlike the above two examples, an unusual feature presented by clones R3-3 and R6-1 is that the reading frame is not preserved by the alternative splicing. The consequence is that the transmembrane sequence in Fc ϵ R-I α is replaced by a completely distinct sequence lacking a long stretch of hydrophobic residues, and thus the translated products would likely be soluble proteins. The mRNA with the internal deletion also exists in two forms that differ at the 5' end, coding for proteins with or without the signal sequence.

Polyadenylylation at different sites may be a third mechanism to generate diverse RNA species coding for $Fc\epsilon R$ -I α . Whereas the clones we have sequenced represent mRNAs that use the third AATAAA polyadenylylation signal in the 3' untranslated region, the $Fc\epsilon R$ -I α cDNA reported previously (6) represented mRNA that was polyadenylylated after the second AATAAA sequence. Differential polyadenylylation, together with 5'-end sequence variation and alternative RNA splicing, may contribute to the regulation of $Fc\epsilon R$ -I α RNA processing and overall expression of the IgE receptor.

Our study suggests possible structural heterogeneity of $Fc \in R-I\alpha$. Structural diversity has been demonstrated for the IgG receptor, $Fc\gamma R$: two distinct receptors share a highly conserved extracellular domain but contain different transmembrane and cytoplasmic segments (10). Furthermore, one of these receptors has two forms differing only by the insertion of an additional 46 amino acids in the cytoplasmic domain, likely resulting from alternative RNA splicing. Recently, structural variation of the human low-affinity IgE receptor, FceR-II, has also been found. A cDNA derived from a T-cell lymphoma was found to differ mainly at the 5' end (M. Nonaka, D. Hsu, C. Hanson, and D. H. Katz, personal communication) from that derived from B cells. Here too, two corresponding mRNAs appear to differ in the 5' untranslated sequence and in a small portion of 5' coding sequence, resulting in a difference of six amino-terminal residues. All these structural heterogeneities may be related to functional complexities of various Fc receptors.

The various possible forms of $Fc \varepsilon R - I\alpha$ -related proteins reported here have not yet been detected in the supernatants or lysates of RBL cells. Some structural heterogeneity of $Fc \varepsilon R - I\alpha$ has been demonstrated at the protein level by IgE-Sepharose 4B affinity purification of M_r 45,000 and M_r 55,000 protein components from RBL cells (25), although it is not known whether these two forms correspond to any of the variants suggested from the cDNA cloning. Truncated variants of $Fc \varepsilon R - I\alpha$ may not recognize IgE because they lack a major portion of the second extracellular domain; alternatively, they may have escaped detection because of their low abundance, suggested by the frequency of their corresponding cDNAs.

If the various predicted FceR-I α -related proteins can be detected in RBL cells, it will be important to determine whether they are unique to the transformed RBL cells or are also found in normal mast cells and basophils. If they are present only in RBL cells, these FceR-I α variants may contribute to differences in cellular properties between these cells and normal mast cells and basophils. This is important to establish because RBL cells have been widely employed as a model system for studying the cellular and molecular basis of immediate hypersensitivity. It would also be significant to establish the functional roles of the α -subunit variants once they are detected. An exciting possibility is that the expression of various mRNAs is differentially controlled and that IgE-mediated activation of mast cells and basophils is affected by varying expression of different forms of Fc ϵ R-I α .

Note. While this manuscript was being reviewed, a report on the cloning of human and rat mast cell $Fc\epsilon R-I\alpha$ was published (26), including one of the RBL cell cDNA variants (R6-1) reported here.

We thank Dr. D. Noonan for his generous gift of the pCDV1 vector and pL1 linker and his advice in cDNA cloning procedures; C. Kennedy and A. Reid for their excellent assistance in the preparation of the manuscript; and Drs. R. Ogata, C. Gritzmacher, M. Wilson, and L. Frigeri for helpful discussions and critically reviewing the manuscript. This work was supported by National Institutes of Health Grants Al19747 and Al20958 to F.-T.L.

- 1. Ishizaka, T. & Ishizaka, K. (1984) Prog. Allergy 34, 188-235.
- 2. Schwartz, L. B. & Austen, K. R. (1984) Prog. Allergy 34, 271-321.
- 3. Froese, A. (1980) CRC Crit. Rev. Immunol. 1, 79-130.
- Metzger, H., Alcaraz, G., Hohman, R., Kinet, J.-P., Pribluda, V. & Quarto, R. (1986) Annu. Rev. Immunol. 4, 419–470.
- Rivnay, B., Rossi, G., Henkart, M. & Metzger, H. (1984) J. Biol. Chem. 259, 1212-1217.
- Kinet, J.-P., Metzger, H., Hakimi, J. & Kochan, J. (1987) Biochemistry 26, 4605-4610.
- 7. Ra, C., Kochan, J., Hakimi, J., Danho, W., Metzger, H. & Kinet, J.-P. (1988) *FASEB J.* 2, A1248 (abstr.).
- Hibbs, M. L., Walker, I. D., Kirszbaum, L., Pietersz, G. A., Deacon, N. J., Chambers, G. W., McKenzie, I. F. C. & Hogarth, P. M. (1986) Proc. Natl. Acad. Sci. USA 83, 6980–6984.
- Lewis, V. A., Kock, T., Plutner, H. & Mellman, I. (1986) Nature (London) 324, 372–375.
- Ravetch, J. V., Luster, A. D., Weinshank, R., Kochan, J., Pavlovec, A., Portnoy, D. A., Hulmes, J., Pan, Y.-C. E. & Unkeless, J. C. (1986) Science 234, 718-725.
- 11. Kulczycki, A., Jr., Isersky, C. & Metzger, H. (1974) J. Exp. Med. 139, 600-616.
- 12. Liu, F.-T. & Orida, N. (1984) J. Biol. Chem. 259, 10649-10652.
- 13. Okayama, H. & Berg, P. (1983) Mol. Cell. Biol. 3, 280-289.
- 14. Hanahan, D. & Meselson, M. (1980) Gene 10, 63-67.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- Maxam, A. M. & Gilbert, W. (1980) Methods Enzymol. 64, 499-560.
- Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) Nucleic Acids Res. 12, 7035– 7056.
- Leff, S. E., Rosenfeld, M. G. & Evans, R. M. (1986) Annu. Rev. Biochem. 55, 1091-1117.
- 19. Breitbrat, R. E., Andreadis, A. & Nadal-Ginard, B. (1987) Annu. Rev. Biochem. 56, 467-495.
- 20. Carlson, M. & Botstein, D. (1982) Cell 28, 145-154.
- Perlman, D., Raney, P. & Halvorron, H. O. (1984) Mol. Cell. Biol. 4, 1682–1688.
- Kaiser, C. A., Preuss, D., Grisafi, P. & Botstein, D. (1987) Science 235, 312–317.
- Leonard, W. J., Depper, J. M., Crabtree, G. R., Rudikoff, S., Pumphrey, J., Robb, R. J., Kronke, M., Svetlik, P. B., Peffer, N. J., Waldmann, T. A. & Greene, W. C. (1984) Nature (London) 311, 626-631.
- 24. Deitcher, D. L. & Mostov, K. E. (1986) Mol. Cell. Biol. 6, 2712-2715.
- 25. Pecoud, A. R., Ruddy, S. & Conrad, D. H. (1981) J. Immunol. 126, 1624–1629.
- 26. Shimizu, A., Tepler, I., Benfey, P. N., Berenstein, E. H., Siraganian, R. P. & Leder, P. (1988) *Proc. Natl. Acad. Sci.* USA 85, 1907-1911.