# Hemophilic factor VIII C1- and C2-domain missense mutations and their modeling to the 1.5-angstrom human C2-domain crystal structure

Miao-Liang Liu, Betty W. Shen, Shelley Nakaya, Kathleen P. Pratt, Kazuo Fujikawa, Earl W. Davie, Barry L. Stoddard, and Arthur R. Thompson

Factor VIII C domains contain key binding sites for von Willebrand factor (vWF) and phospholipid membranes. Hemophilic patients were screened for factor VIII Cdomain mutations to provide a well-characterized series. Mutated residues were localized to the high-resolution C2 structure and to a homology model of C1. Of 30 families found with mutations in the C domains, there were 14 missense changes, and 9 of these were novel. Of the missense mutations, 10 were associated with reduced vWF binding and 8 were at residues with surface-exposed side chains. Six of the 10 mutants had nearly equivalent factor VIII clotting activity and antigen level, suggesting that reduced vWF binding could cause hemophilia by reducing factor VIII stability in circulation. When the present series was combined with previously described mutations from an online international database, 11 C1 and C2 mutations in patients with mild or moderately severe hemophilia A were associated with antibody-inhibitor development in at least one affected individual. Of these substitutions, 6 occurred at surface-exposed residues. As further details of the C1 structure and its interface with C2 become available, and as binding studies are performed on the plasma of more patients with hemophilic C-domain mutations, prediction of surface binding sites should improve, allowing confirmation by site-specific mutagenesis of surface-exposed residues. (Blood. 2000;96: 979-987)

© 2000 by The American Society of Hematology

#### Introduction \_

Factor VIII circulates as a precursor to factor VIIIa, an essential cofactor for intrinsic factor X activation that is a critical early step in coagulation. The factor VIII gene is 186 kilobases (kb) and contains 26 exons.<sup>1,2</sup> The transcribed protein contains a signal peptide and a mature sequence of 2332 amino acid residues. Domain structures from the amino terminus are A1-A2-B-A3-C1-C2. The carboxy-terminal 313 amino acids form 2 highly homologous C domains. The C2 domain, and possibly C1, contribute to von Willebrand factor (vWF) binding,<sup>3,4</sup> which is essential for the stabilization of factor VIII in circulation. When thrombin activates factor VIII, factor VIIIa dissociates from vWF and is concentrated by binding to a phospholipid surface, where it interacts with factor IXa. Factor VIIIa enhances the rate of activation of factor X by factor IXa by more than 10<sup>5</sup>-fold.<sup>5</sup> Lipid binding involves the C2 domain<sup>6-8</sup> and possibly other sites in the A3-C1-C2 light chain of factor VIIIa.<sup>9</sup> The factor VIII C domains contain surface epitopes both for clinically significant alloimmune and autoimmune inhibitors of factor VIII and for monoclonal antibodies.<sup>10-13</sup> C2 may also bind to factor Xa.14

A deficient factor VIII clotting activity leads to hemophilia A, a congenital bleeding tendency of variable severity that is due to distinct factor VIII gene mutations. Some hemophilic mutations lead to circulating dysfunctional proteins whereas others affect expression, secretion, or stability in circulation. Comparison of the baseline clotting activity with the antigen level in the plasma of hemophilic patients provides an estimate of the specific activity of

their factor VIII and helps identify those mutations associated with a dysfunctional protein. An international database of hemophilic point mutations lists 48 missense mutations in the 2 C domains.<sup>15</sup> Unfortunately, only a quarter of these C-domain mutations have had plasma factor VIII antigen levels reported; of these, only 2 circulate dysfunctional antigen.

Each of the factor VIII C domains has a single disulfide bond connecting the cysteine residues that are near the amino- and carboxy-terminal ends.16 Comparisons of sequences of human factors VIII and V show that the C domains share about a 45% sequence identity<sup>17</sup>; human factor VIII C1 is 7 amino acids shorter than C2 but still shares 42% identity.<sup>1</sup> A comparison of factor VIII complementary DNA sequences from different species shows that predicted amino acid residues in the C1 domains are more than 90% identical and the C2 domains are around 80% identical.<sup>18-20</sup> The C2 domain of factor VIII also shares 20% sequence identity with galactose oxidase, and a molecular model based on the crystal structure of galactose oxidase correctly predicted an 8-stranded β-sandwich core.<sup>21</sup> Pratt et al<sup>22</sup> reported the 1.5 Å crystal structure of a recombinant human C2 domain. This structure provided a basis for predictions of surface binding sites and generation of models to examine homologous portions of the C1 domain. Furthermore, the crystal structure allowed modeling of hemophilic missense mutations in both the C1 and the C2 domains, which can indicate residues that are important for function or stability of the protein.

In this study, unrelated hemophilia A families were screened for

From the Department of Medicine, University of Washington, Puget Sound Blood Center; Fred Hutchinson Cancer Research Center; and Department of Biochemistry, University of Washington, Seattle, WA.

Submitted February 23, 2000; accepted April 4, 2000.

M.-L.L. and B.W.S. contributed equally to this study.

Reprints: Arthur R. Thompson, Puget Sound Blood Center, 921 Terry Ave, Seattle WA 98104-1256; e-mail: arthomps@u.washington.edu.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 U.S.C. section 1734.

© 2000 by The American Society of Hematology

Supported in part by the American Heart Association Grants-in-Aid 9750412N (A.R.T.) and 0050336N (B.L.S.); by Public Health Service grants GM49857 and HL62470 (B.L.S.) and HL16919 (E.W.D.) from the National Institutes of Health; and by Baxter, Novo-Nordisk and Speywood for mutation screening (A.R.T.).

heteroduplex formation in the 7 exonic sequences coding for the 2 C domains of the factor VIII gene: exons 20 through 26. Thirty families had a heteroduplex band in 1 exon, and DNA sequencing identified a mutation in each. Twenty-two families had 14 distinct missense mutations, including 9 that have not been reported previously.<sup>15</sup> Missense mutations were localized within the crystal-lographic structure of C2 and a homology model of the C1 domain; preliminary results were presented.<sup>23</sup>

#### Patients and methods

#### Patients and family members

DNA samples were from 76 unrelated families including patients or obligate carriers with hemophilia A not previously described. Those with severe hemophilia A were negative for gene inversions, recurrent arginine to stop codons, and gross deletions. This series did not include 82 families previously diagnosed with factor VIII gene inversions, 2 families with partial gene deletions, or 96 small mutations within the coding sequences for the A domains in 121 families, the majority of which have been published.<sup>24-26</sup> Samples were submitted for clinical carrier testing or obtained following informed consent under a protocol approved by the University of Washington Human Subjects Review Committee.

#### Factor VIII determinations

Factor VIII clotting activities were performed by one-stage assays.<sup>27</sup> Antigen levels were measured by modification of an enzyme-linked immunosorbent assay (ELISA)<sup>26,28</sup> with the use of a commercial factor VIII antigen kit with monoclonal antibodies (Immuno AG, Vienna, Austria).<sup>29</sup> The commercial ELISA kit is sensitive to 0.2% of the pooled normal plasma level, and antigen was undetectable in plasmas from the 2 patients with frameshift mutations when they had no residual transfused factor VIII.

#### vWF binding ELISA

A factor VIII–vWF binding assay using a polyclonal antibody against vWF and monoclonal antibody against the factor VIII light chain was carried out to detect the vWF binding abilities of hemophilic plasmas.<sup>30-32</sup> Briefly, microtiter wells were coated with rabbit anti-vWF (Sigma, St Louis, MO, diluted 1 to 8000 in 0.1 mol/L NaHCO<sub>3</sub>, pH 9.5, coating buffer) and incubated in 2% bovine serum albumin in phosphate-buffered saline (PBS). After washing with PBS and incubation with a 1:100-fold dilution of normal plasma as a source of vWF, wells were washed with 0.4 mol/L CaCl<sub>2</sub> to remove bound plasma factor VIII and were reequilibrated in PBS. Dilutions of patients' plasmas, normal plasma, or a purified recombinant factor VIII standard were then added, and the bound factor VIII was detected with the same anti–factor VIII light chain monoclonal antibody (conjugated with horseradish peroxidase) that was used in the standard factor VIII ELISA, described above.

#### Polymerase chain reaction amplification conditions

Fifty to 100 ng of genomic DNA from different patients' leukocytes were used for polymerase chain reaction (PCR) amplification. DNA amplification was performed in 50 mmol/L KCl, 10 mmol/L Tris-HCl, pH 8.3, 5 mmol/L MgCl<sub>2</sub>, 200 µmol/L dNTPs (ATP, GTP, CTP, and TTP), and 0.5 µmol/L for each primer set (Table 1). AmpliTaq DNA polymerase (1 to 2 units, Perkin Elmer, Foster City, CA) was added for each 100 µL reaction. PCRs were performed in a TwinBlock (Ericomp, San Diego, CA) or PTC-100 (MJ Research, Incline Village, NV) programmable thermal controller. Cycling parameters for the reaction were optimized for each exon. Amplified PCR products included both 5' and 3' intronic splice junctions and from 19 to 114 base pairs of intronic sequence (excluding the primer sequences), depending on the amount of published sequence and size of the exon. Intragenic polymorphic alleles were determined as before.<sup>25,26</sup>

#### Heteroduplex analysis

For heteroduplex formation, 5 to 7 µmol/L of amplified patient and normal fragments were mixed and then cycled through a mutation-detection enhancer (MDE) denature-annealing program:  $94^{\circ}C$  for 3 minutes,  $80^{\circ}C$  for 2 minutes,  $70^{\circ}C$  for 2 minutes,  $65^{\circ}C$  for 15 minutes,  $50^{\circ}C$  for 5 minutes,  $40^{\circ}C$  2 minutes, and  $37^{\circ}C$  for 15 minutes. Standard loading dye (3 to 4 µL) was added to each mixture before loading. Electrophoresis was performed on  $43 \times 26$ -cm gels consisting of  $1 \times$  MDE gel solution (FMC Bioproducts, Rockland, ME) with 15% (wt/vol) urea in  $0.6 \times$  TBE buffer ( $1 \times$  is 8.9 mmol/L Tris base, 8.9 mmol/L boric acid, and 0.2 mmol/L EDTA, pH 8.0) for 16 to 24 hours at 800 V at room temperature. The gels were stained with ethidium bromide to visualize heteroduplex bands.<sup>26,33</sup>

#### **DNA sequencing**

Amplified fragments were sequenced in both directions in an Applied Biosystems (Foster City, CA) automated sequencer, model 373, with the use of amplification primers. Separate amplified fragments were used to verify positive findings with either predicted restriction recognition site changes (in 16) or repeat sequencing (in 14) of the distinct mutations. For 10 of the 30 families in which a mutation was determined, verification was also obtained on a sample from an obligate carrier or another affected family member.

#### Structural analyses of the C1 domain and missense mutation sites

A model of factor VIII C1 domain was built and energy minimized with the use of the program Modeller (Molecular Simulations, San Diego, CA) based on a 1.5 Å resolution x-ray structure of the highly homologous C2 domain of the same molecule. The sequences of the C1 and C2 domains share 42% identity (85% homology) as shown in Figure 1. Sequence alignment by Bestfit program of Genebank and Align of Modeller showed that most hydrophobic residues in the  $\beta$ -sheet sandwich core of the C2 domain were conserved (Figure 1). Additional gaps in the sequences of both

Table 1.	Primers	and PCR	amplification	conditions f	for factor	VIII mutation	screening
----------	---------	---------	---------------	--------------	------------	---------------	-----------

	5' Primer	3' Primer	Size	Annealing	Extension,* sec at 72°
Exon	(primers listed from	m 5' to 3' sequence)	(base pairs)	time (sec), °C	
20	ACG TTG AGT ACA GTT CTT GG	ACT AAT AGA AGC ATG GAG ATG	320	30, 52°	50
21	GAA TTT AAT CTC TGA TTT CT	GAG TGA ATG TGA TAC ATT TC	279	30, 52°	50
22	AAA TAG GTT AAA ATA AAG TGT TAT	TTA ATG GTA TGT AAT TAG TCA TTT A	218	35, 43°	50
23	ACT CTG TAT TCA CTT TCC AT	AAC TAG AAC AGT TAG TCA CC	258	40, 55°	50
24	ATA ACT GAG GCT GAA GCA TG	CTC TGA GTC AGT TAA ACA GT	270	30, 54°	60
25	GAG TGA GAA GTG CTG TGG	TTG CTC TGA AAA TTT GGT CAT A	381	30, 58°	50
26	GCT TTG CAG TGA CCA TTG TC	AGC TGA GGA GGG AGA GGT GA	265	40, 60°	50

\*The first step is 93°C for 2 min for each exonic fragment followed by 35 cycles each of 1 min at 93°, annealing, and then extension.

Figure 1. Alignment of C domain sequences. Amino acids are aligned by CLUSTAL W 1.7, GenomeNet. Factor VIII (VIII) sequences are as indicated for different species: human (h),<sup>1</sup> porcine (p),<sup>19</sup> murine (m),<sup>18</sup> canine (c)<sup>20</sup>; factor V(V) sequences are: human (h), $^{17,34}$  murine (m),35 bovine (b).36 Bold residues indicate differences from human factor VIII, underlined are C2 residues that may bind to phospholipids, and boxes indicate a potential Asn-linked carbohydrate. HphA (above each human factor VIII sequence row) indicates hemophilic mutations: an asterisk is for those associated with an inhibitor (see Table 3). Shading is for β-strands in the human C2 crystal structure.22 In panel A. C1 homologous sequences are shown and  $\beta$ -strands are distinguished from those in C2 with a prime, with no  $\beta 3'$  nor  $\beta 4'$  (see Figure 2). In panel B, C2 homologous sequences are shown including a C2 inhibitor epitope as defined by Healey et al.13 Note that exact residues at the ends of B-strands are subject to interpretation as they often end within a residue (coordinates are available: see Acknowledgments).

A. CI domains	β1'Ε	β <b>2</b> ′	β	5'	β6'	
HphA	. v .	. P. R	• •	F	. G	
hVIIIC1 2020	KCQTPLGMASGHIRD	FQITASGQYGQWAI	KLARLHYSGSINA	STKEPFS-	WIKVDLLAP	2078
pVIIIC1	ECQAPLGMASGRIRD	FQITASGQYGQWAI	KLARLHYSGSINA	WSTKDPHS-	WIKVDLLAP	
mVIIIC1	OCOIPLGMASGSIRDI	FÖTTASGHYGOWAI	NLARLHYSGSINA	WSTKEPFS-	WIKVDLLAP	
CVIIICI	KCOTPLOMASCHIRD	FOTTASGOYGOWAI	KLARLHYSGSTNA	STKDPES-	WIKVDULAP	
bryc1	DCDMDMCI STCT I SD	OTVACEPTOVWE	DIADINNOCCVNA	CUERTAAF	TACKDWITOUDMOKE	
IFVCI	DCRMPMGLSIGIISD.	SQIKASEF LGIWEI	REAR LINNOGS INA	NO VEREMAN	TRAKEWIQVDMQKB	
mFVCI	ECKMPMGLSTGVISD	SQIKASEYLTYWE	RLARLINNAGSYNA	VSTERTALL	IL DIKLAIÖADWÖKE	
bFVC1	ECKMPMGLSTGLIAD	SQIQASEFWGYWE	PKLARLNNGGSYNA	VIAEKLSTE	FNPEPWIQVDMQKE	
	67' 68' E 69'	β10' R β11	B12+ B1	3 614	B15 <sup>7</sup> B16 <sup>7</sup>	
HphA	*NRS C	.L	C* . P	¥. ''''	S	
VIIIC1 2079	MII HGIKTOG AROL	KFSSLYISOFIIM	SLDGKKWOTY RG	NSTGTLMVF	FGNVDSSGIKHNIF	2140
pVIIIC1	MIT HOIMTOG AROL	KESSLYISOFIIM	SLDGRNWOSY RG	STGTLMVF	FGNVDASGIKHNIF	1
mVTTTC1	MTW HOTKTOG AROL	RESELVISOFITM	STUCKKWESV OC	NSTOTIMVE	FONUDSSCIKHNS	i
WIIICI	NTT UCTNIES ADO	ADOLITOQT TIM	OLDONULUOV DO		TONVDECCTUNT	
CVITICI	MIT HEIMIGE ARO	APSSLIVSQFIIM.	SEDGNAWRSI KG	NSIGIUMVE	FGNVDSSGIKHNIF	_
hFVC1	VIV TGIQTQG AKH	<b>LKSCYTTEFYVA</b>	ssnginwqif kq	NSTRNVMYF	NGNSDASTIKENOF	•
mFVC1	VVV TGIQTQG AKH	LKSCFTTEFQVA	SSDQTNWQIF RG	KSGKSVMYF	TGNSDGSTIKENRL	
bFVC1	VLL TGIQTQG AKHY	LKPYYTTEFCVA	SLDRKNWRIF KG	NSTRNVMYF	GGNSDASTIKENQI	
	B17/ B19/	C 8197		_		
HphA		H CR	% ide	entity with	number of	
	H* QI	L. *HV S	. h)	VIII CI	residues	
hVIIIC1 2141	NPPIIA RYIRLHPTH	HYSIRSTLRMELM	CDLN 2172	(100)	153	
pVIIIC1	NPPIVA RYIRLHPTI	TYSIRSTLRMELM	CDLN	92	153	
mVIIIC1	NPPIIA RYIRLHPTH	ISIRSTLRMELMO	CDLN	92	153	
cVIIIC1	NPPIIA QYIRLHPTH	HYSIRSTLRMELLC	CDFN	94	153	
hFVC1	DPPTVA RYTRTSPTI	RAYNEPTLELO	CEVN	48	159	
mEVCl	DUTUA PVTPTUDT	COVNEDTIRI.FL.O	CEVN	45	159	
hevel	DEDUNA BYIDTODY	OVNEDAL DI EL O	CENN	15	150	
Drvci	DPPVVA KIIKISPI	30 INKEADADDQ	C LIVIN	40		
B. C2 domains			() <b>*</b>			
B. C2 domains	ßı	NB2 B3	Q* β4 ī. <i>in</i> .	hihitor B5	enitone B6	
B. C2 domains HphA	β1 Ι .D* <u>Τ</u>	<u>Nβ2 β3</u>	β4L in. - G.	hibitor β5	<i>epitope</i> β6 M *C. A*	
B. C2 domains HphA hVIIIC2 2173	β1 I .D* T SCSMPLGMESKAISDA	<u>Nβ2 β3</u> . P AQITASSYFTN <u>MF</u> 2	$\beta 4$ <u>L</u> in. G. WSPSKARLHI	<i>hibitor</i> β5 LQG <b>R</b> SNAWR	<u>epitope</u> β6 M*C.A* PQVNNPKEWLQVDF	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2	β1 I.D* T SCSMPLGMESKAISD SCSMPLGMQNKAISD	Nβ2 β3 P . QITASSYFTN <u>MF</u> SQITASSHLSNIF7	β4 <u>L</u> in. G. MIWSPSKARLHI MIWSPSQARLHI	<i>hibitor</i> β5 LQG <u>R</u> SNAWR LQGR <b>T</b> NAWR	<u>epitope</u> β6 M*C.A* PQVNNPKEWLQVDF PRVSSAEEWLQVDL	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 mVIIIC2	β1 I .D* T SCSMPLGMESKAISDJ SCSMPLGMESKVISD SCSTPLGMESKVISD	<u>Nβ2</u> β3 P.J AQITASSYFTN <u>MF</u> 2 SQITASSHLSNIF2 FOITASSYFTNMF2	β4 <u>L</u> in G. MTWSPSKARLHI MTWSPSQARLHI	<i>hibitor</i> β5 LQG <u>R</u> SNAWR LQGRTNAWR LOGRTNAWR	epitope β6 M *C. A*- PQYNNPKEWLQVDF PRVSSAEEWLQVDL POVNDPKOWLOVDL	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 mVIIIC2 cVIIIC2	β1 SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ SCSIPLGMESKVISD SCSMPLGMESKAISDJ	<u>Nβ2 β3</u> - P . 1 - AQITASSYFTN <u>MF</u> - OITASSYFTNMF - OITASSYLSSML	β4 L im. G. MIWSPSKARLHI MIWSPSQARLHI MIWSPSQARLHI	hibitor β5 LQG <u>R</u> SNAWR LQGRTNAWR LQGRTNAWR LOGRTNAWR	epitope β6 M *C.A*- PQVNNPKEWLQVDF PRVSSAEEWLQVDL PQVNDPKQWLQVDL PQANNPKEWLQVDF	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 mVIIIC2 cVIIIC2 bFVC2	β1 I .D* T SCSMPLGMESKAISD SCSMPLGMESKAISD SCSMPLGMESKVISD SCSMPLGMESKAISD	<u>Nβ2</u> β3 <u>P</u> 1 AQITASSYFTN <u>MP</u> 5QITASSHLSNIF CQITASSYFTNMF/ AQITASSYLSSML/ COITASSFKSWM/	Q* <u>G</u> <u>G</u> <u>G</u> <u>G</u> <u>G</u> <u>G</u> <u>G</u> <u>G</u>	hibitor <u>65</u> LQG <u>R</u> SNAWR LQGRTNAWR LQGRTNAWR LQGRTNAWR	epitope β6 M *C. A* PQVNNPKEWLQVDF PRVSSAEEWLQVDL PQVNDPKQWLQVDL PQANNPKEWLQVDF BKANNNKOWL BTDL	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 mVIIIC2 eVIIIC2 hFVC2 mFVC2	β1 J D* T SCSMPLGMESKAISDJ SCSMPLGMONKAISDJ SCSMPLGMESKAISDJ GCSTPLGMESKAISDJ GCSTPLGMESKAISDJ	Nβ2 β3 . P . 1 AQITASSYFTN <u>MP</u> 2 SQITASSHISNIF2 IQITASSYISSML2 KQITASSFKKSWWC COLTASSFKSFKSWWC	G A G M M M M M M M M M M M M M	hibitor β5 LQGRSNAWR LQGRTNAWR LQGRTNAWR LQGRTNAWR AQGRVNAWQ	epitope <u>β6</u> <u>M</u> *C. <u>A</u> * PQVNNPKEWLQVDF PRVSSAEEWLQVDL PQVNDPKQWLQVDL PQANNPKEWLQVDF AKANNNKQWLBIDL AKANNNKQWLBIDL	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 cVIIIC2 hFVC2 mFVC2 hFVC2	β1 SCSMPLGMESKAISD SCSMPLGMONKAISD SCSTPLGMESKVISD SCSMPLGMESKAISD GCSTPLGMENGKIEN GCSTPLGLEDGRIQDI GCSTPLGLEDGRIQDI	Nβ2 β3 P 1 AQITASSYFTNMF SQITASSHLSNIF QITASSYLSSML QITASSFKSWW QITASSFKSWW QITASSFKSWW	G4 G. MTWSPSKARLHI MTWSPSQARLHI MTWSPSQARLHI MTWSPSQARLHI MTWSPSQARLHI MTWEPFRARLN2 MDYWEPSLARLN2	hibitor β5 LQGRSNAWR LQGRTNAWR LQGRTNAWR LQGRTNAWR AQGRVNAWQ AQGRVNAWQ	epitope <u>B6</u> M *C. A*- PQVINPKEWLQVDF PQVNDPKQWLQVDL PQANNPKEWLQVDL AKANNNKQWLBIDL AKANNNKQWLBIDL	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 mVIIIC2 cVIIIC2 hFVC2 mFVC2 bFVC2	β1 .D* T SCSMPLGMESKALSDJ SCSMPLGMESKALSDJ SCSTPLGMESKVISD SCSTPLGMESKVISD GCSTPLGMESGKIEN GCSTPLGMESGKIEN GCSTPLGMESGKIEN	Nβ2 β3 . P . 1 AQITASSYFTNMP QITASSHLSNIF AQITASSYLSSML QITASSFKSWWC QITASSFKKSWC QITASSFKKSWC QITASSFKKSWC QITASSFKKSWC QITASSFK QITASFK QIT	<u>B4</u> L im T WSPSKARLHI T WSPSQARLHI T WSPSQARLHI T WSPSQARLHI BDY WEPFRARLN2 BDY WEPFLARLN2 DY WEPFLARLN2	hibitor β5 LQGRSNAWR LQGRTNAWR LQGRTNAWR AQGRVNAWQ AQGRVNAWQ AQGRVNAWQ	<u>epitope</u> <u>B6</u> *C.A*- PQVNNPKEWLQVDF PRVSRAEEWLQVDF PQVNDPKQWLQVDL PQANNPKEWLQVDF AKANNNKQWLBIDL AKANNNKQWLQTDL	2234
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 2173 pVIIIC2 cVIIIC2 cVIIIC2 hFVC2 hFVC2 hFVC2 hVIIIC1 2020	β1 .D* T SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ GCSTPLGMESKAISDJ GCSTPLGMESKAIEND GCSTPLGMESGKIEND KCQTPLGMESGHIRDJ	NB2 B3 . P . . P .	<u>B4</u> <u>L</u> <u>M</u> <u>H</u> <u>H</u> <u>H</u> <u>H</u> <u>H</u> <u>H</u> <u>H</u> <u>H</u> <u>H</u> <u>H</u>	hibitor <u>β5</u> LQGRSNAWR LQGRTNAWR LQGRTNAWR LQGRVNAWQ AQGRVNAWQ AQGRVNAWQ XSGSINAWS	epitope <u>B6</u> M *C A* PQUNNPKEWLQUDF PRVSSABEWLQUDL PQUNDPKQWLQUDE AKANNNKQWLQUDE AKANNNKQWLQUDL TKEPFSWIKVDL	2234 2075
B. C2 domains HphA hvIIIC2 2173 pVIIIC2 2173 vVIIIC2 cVIIIC2 cVIIIC2 hFVC2 bFVC2 bFVC2 bFVC2 hVIIIC1 2020 HobA	β1T SCSMPLGMESKAISD SCSMPLGMONKAISD SCSIPLGMESKVISD SCSMPLGMESKVISD GCSTPLGMENGKIENI GCSTPLGMENGKIENI KCQTPLGMASGHIENI β7 β8	Nβ2 β3 P 1 AQITASSYLTNMFJ SQITASSHLSNIF/ TQITASSYLSSML/ AQITASSFKSWWC QUITASSFKSWWC QUITASSFKSWWC QUITASSFKSWWC CQITASSFKSWWC CQITAS	<u>β4</u> L im <u>β4</u> L im <u>β</u> wspskarlhi TTwspsQarlhi TTwspsQarlhi DY - wspfQarlhi DY - wspfLarlni DY - wspfLarlni DY - wspfLarlni BY - wspfLarlni B11 β12	hibitor <u>b</u> Logrennawr Logrennawr Logrennawr Logrennawr Aogrvnawg Aogrvnawg Aggrvnawg Xsgsinaws B13 B1	epitope	2234
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 cviliC2 hFVC2 mFVC2 bFVC2 hviliC1 2020 HphA	β1 D* T SCSMPLGMESKALSDJ SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKVISD GCSTPLGMESKALSDJ GCSTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMASGHIRDJ A VA R	NB2 B3 AQITASSYFTNMF7 QITASSYFTNMF7 QITASSYFTNMF7 QITASSFKSWWG QITASSFKKSWWG QITASSFKKSWWG QITASSFKKSWWG A A A A A A A A A A A A A	$\begin{array}{c c} & & & & & & \\ & & & & & & & \\ \hline & & & &$	hibitor <u>b5</u> LOGRINAWR LOGRINAWR LOGRINAWR LOGRINAWR AOGRVNAWO AOGRVNAWO XSGSINAWS B13 B1	<u>epitope</u> <u>β6</u> *C. <u>λ</u> * PQVNNPKEWLQVDF PQVNDPKQWLQVDL PQVNDPKQWLQVDL PQANNPKEWLQVDF AKANNNKQWLBIDL AKANNNKQWLQUDL TKEPFSWIKVDL 4 β15 V	2234
B. C2 domains HphA hvIIIC2 2173 pVIIIC2 cVIIIC2 hVIIC2 hFVC2 hVIIC2 hVIIIC1 2020 HphA hVIIIC2 2235	β1 D* T SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ GCSTPLGMESKAISDJ GCSTPLGMESGKIENJ KCQTPLGMESGKIENJ KCQTPLGMASGHIRDJ <u>β7</u> β8 V· βR QKTMKV TGVTTQG J	NB2 B3 AQITASSYFTNME/ SQITASSYLSML AQITASSYLSML AQITASSFKSWMC (QITASSFKSWMC (QITASSFKSWMC FQITASC β βl0 c ×I /KSLLTSMYVKEPI	β4         L           G         G           G         G           MSDSQARLHI         G           MSDSQARLHI         G           MSDSQARLHI         G           MSDSQARLHI         G           MSDSQARLHI         G           MSDSQARLHI         G           MSDY - WEPSLARLNI         G           MSY - WEPSLARLNI         G           MSSSQGQAPKLARLHI         G           MS         G	hibitor β5 LOGRSNAWB LOGRTNAWR LOGRTNAWR LOGRTNAWR AQGRVNAWQ AQGRVNAWQ AGGRVNAWQ XSGSINAWS β13 β1 FONGKV	epitope <u>66</u> M *C A*- POVINPKEWLQVDF PRVSSAESWLQVDL PQVNDPKQWLQVDL PQUNDPKEWLQVDF AKANNNKQWLQVDE AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNNQWLQIDL T KEPFSWIKVDL 4 615 V KVFQGNQDSFT 22	2234 2075 91
B. C2 domains HphA hvIIIC2 2173 pVIIIC2 2173 invIIIC2 cVIIIC2 hFVC2 hFVC2 hFVC2 hVIIIC1 2020 HphA hVIIIC2 2235 pVIIIC2	β1 CSMPLGMESKAISD SCSMPLGMESKAISD SCSMPLGMESKVISD SCSMPLGMESKVISD GCSTPLGMENGKIEN GCSTPLGMENGKIEN KCQTPLGMASGHIRD β7 β8 V AR SKTWKV TGITTOG M	NB2 B3 P 1 AQITASSHISNIF/ CQITASSHISNIF/ AQITASSHISML/ AQITASSFKKSWWC (QITASSFKKSWWC (QITASSFKKSWWC CQITASSFKKSWWC POITAS β9 β10 c KSLLISMYVKEPI /KSLLSSMYVKEPI	$\begin{array}{c c} & & & & & \\ & & & & & \\ \hline & & & & \\ &$	hibitor β5 LQGRSNAWR LQGRTNAWR LQGRTNAWR LQGRTNAWR AQGRVNAWQ AQGRVNAWQ YSGSINAWS β13 β1 FQNGKV LQDGHT	epitope <u>B6</u> *C. A* PQVNNPKEWLQVDF PRVSSAEEWLQVDI PQANNPKEWLQVDF AKANNNKQWLEDL IAKANNNKQWLEUDL IAKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNKQWLQVD AKA	2234 2075 91
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 cviliC2 hFVC2 bFVC2 hviliC1 2020 HphA hviliC2 2235 pviliC2 mviliC2	β1 I D* T SCSMPLGMESKAISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKVISD GCSTPLGMESKAISD GCSTPLGMESGKIENT KCQTPLGMESGKIENT KCQTPLGMESGKIENT KCQTPLGMASGHIRD GCSTPLGMESGKIENT KCQTPLGMASGHIRD GCSTPLGMESGKIENT KCQTPLGMASGHIRD GCSTPLGMESGKIENT KCQTPLGMASGHIRD GCSTPLGMESGKIENT GCSTPLGMESGKIENT GCSTPLGMESGKIENT GCSTPLGMESGKIENT GCSTPLGMESGKIENT GCSTPLGMESGKIENT GCSTPLGMESKVISD GCSTPLGMESGKIENT	NB2 B3 AQITASSYFTNMF/ AQITASSYFTNMF/ AQITASSYFKNMF/ AQITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKSWWG (QITASSFKKSWWG (QITASSF (QITASSF (Q	$\begin{array}{c c} & & & & & \\ & & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & &$	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AQGRVNAWQ AQGRVNAWQ AGGRVNAWG β13 β1 FONGKV LODGKV LODGKV	<u>epitope</u> <u>β6</u> PQVNPKEWLQVDF PRVSSAEEWLQVDL PQVNDPKQWLQVDL PQANNPKEWLQVDL PQANNPKEWLQVDL AKANNNKQWLQVDL AKANNNKQWLQUDL AKANNNKQWLQUDL TKEPFSWIKVDL 4 β15 V KVFQGNQDSST KVFQGNQDSST	2234 2075 91
B. C2 domains HphA hvIIIC2 2173 pVIIIC2 cVIIIC2 hVIIC2 hFVC2 hVIIIC1 2020 HphA hVIIIC2 2235 pVIIIC2 mVIIIC2 cVIIIC2	β1 D* T SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ SCSMPLGMESKAISDJ GCSTPLGMESKAISDJ GCSTPLGMESGKIENJ KCQTPLGMESGKIENJ KCQTPLGMASGHIRDJ β7 β8 OKTMKV TGVTTQG V QKTMKV TGVTTQG V QKTMKV TGITTQG V	NB2 B3 AQITASSYFTNMF/ AQITASSYFTNMF/ AQITASSYFSML/ AQITASSFKSSWL/ QITASSFKSSWL/ QITASSFKSSWL/ QITASSFKSWL/ PQITAS( β9 β10 c *I /KSLLISMYVKEFI /KSLLISMYVKEFI /KSLLISMYVKEFI	β4         L           G         G           G         G           MSDSQARLHI         H           MSDSQUAPKLARLNI         H           MSINY - WEPFLARLNI         H           MSSSQCGWAPKLARLHI         B12           MSSSQDGHQWTLF         MSSSQDGHQWTLF           MSSSQDGHWTLF         ISSSQDGHWTQI           ISSSQDGHWTQI         ISSSQDGHWTQI	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ XSGSINAWS β13 β1 FONG - KV LOBG - HT LYNG - KV	<u>epitope</u> <u>β6</u> <u>M</u> *C <u>A</u> *- PQVNPKEWLQVDF PQVNDPKQWLQVDL PQVNDPKQWLQVDL PQNNPKEWLQVDF AKANNNKQWLBIDL AKANNNKQWLUDL TKEPFSWIKVDL 4 β15 V KVFQGNQDSST KVFQGNQDSST KVFQGNQDSST KVFQGNRDSST	2234 2075 91
B. C2 domains HphA hvilic2 2173 pvilic2 mvilic2 mvilic2 dvilic2 hvilic2 mFvc2 hvilic1 2020 HphA hvilic2 2235 pvilic2 mvilic2 evilic2 hvilic2	$\begin{array}{c} \beta 1 \\ \textbf{I} & \textbf{D} \star \overline{\textbf{T}} \\ \text{SCSMPLGMESKAISD} \\ \text{SCSMPLGMESKAISD} \\ \text{SCSMPLGMESKVISD} \\ \text{SCSMPLGMESKVISD} \\ \text{SCSMPLGMESKVISD} \\ \text{GCSTPLGMESKVISD} \\ \text{GCSTPLGMESGKIENI } \\ \text{KCQTPLGMASGHIRDI } \\ \hline \begin{array}{c} \beta 7 \\ \textbf{V} \\ \textbf{A} \\ \text{KCQTPLGMASGHIRDI } \\ \text{OKTMKV TGUTTQG } \\ \text{OKTMKV TGUTTQG } \\ \text{KKMKV TGUTTQG } \\ \end{array} \right.$	Nβ2         β3           QITASSYFTMP/         AQITASSYFTMF/           QITASSYLSNIF/         CONSTRUCT           QITASSFKSWW         QITASSFKSWW           QITASSFKSWW         QITASSFKSWW           QITASSFKSWW         QITASSFKSWW           QITASSFKSWW         QITASSFKSWW           QITASSFKSWW         QITASSFKSWW           QITASSFKSWW         QITASSFKSWW           VGITASSFKSWW         YKSLLSSWYKEN           JKSLLTSMYVKEN         YKSLSFTSMFVKEN           JKSLLISMYVKEN         YKSLLSSWYKEN           JKSLLISMYVKEN         YKSLSFTSMFVKEN	β4 L m G G G G G G G G G C M S S S S S S S S S S S S S	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ ASGSINAWS β13 β1 FONG - KV LÖDG - KV LONG - KV LONG - KV LONG - KV	<u>epitope</u> <u>β6</u> M *C. A*- PQVNNPKEWLQVDF PRVSSAEEWLQVDI PQVNDFKQWLQVDI PQANNPKEWLQVDF AKANNNKQWLEDLI AKANNNKQWLQUDL AKANNNKQWLQUDL AKANNNKQWLQUDL KVFQGNQDSFT 22: KVFQGNQDSST KVFQGNQDSST KVFQGNQDSST KVFQGNTSST KVFQGNTSST	2234 2075 91
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 mviliC2 hFVC2 hVIIC2 2020 HphA hviliC2 2235 pviliC2 mviliC2 cviliC2 hFVC2 hVIC2 hVIC2 hVIC2 hVIC2 hFVC2	β1 J D* T SCSMPLGMESKAISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKAISD GCSTPLGMESGKIENI GCSTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI CSTPLGMESGKIENI CSTPLGMESGKIENI CSTPLGMESGKIENI CSTPLGMESGKIENI CSTPLGMESGKIENI CSTPLGMESGKIENI SCSTPLGMESKVISD SCS	Nβ2         β3           AQITASSYFTNMF/         AQITASSYFTNMF/           AQITASSYFTNMF/         AQITASSYFKSWWG           AQITASSFKSWWG         AQITASSFKSWWG           AGITASSFKSWWG         AQITASSFKSWWG           AGITASSFKSWWG         AQITASSFKSWWG           AGITASSFKSWWG         AQITASSFKSWWG           AGITASSFKSWWG         AQITASSFKSWWG           AGITASSFKSWWG         AQITASSFKSWWG           AGITASSFKSWWG         AQITASSFKSWWG           AGITASSWWG         AQITASSFKSWWG           AGITASSFKSWWG <td><math display="block">\begin{array}{c c} &amp; &amp; &amp; &amp; &amp; &amp; \\ &amp; &amp; &amp; &amp; &amp; &amp; \\ &amp; &amp; &amp; &amp; &amp; </math></td> <td>hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ ACGRVNAWQ ACGRVNAWQ AGGRVNAWG B13 β1 FONG - KV LODG - HT LODG - KV LONG - KV RLKSSMVD</td> <td><u>epitope</u><u>β6</u> *C. A*- PQVNNPKEWLQVDF PQVNDPKQWLQVDL PQVNDPKQWLQVDL PQANNPKEWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL KUPQGNQDST KVFQGNQDSST KVFQGNQDSST KVFQGNQDSST KVFQGNQDSST KVFQGNRDSST KIFEGNTNTKG KIFEGNTNTKG</td> <td>2234 2075 91</td>	$\begin{array}{c c} & & & & & & \\ & & & & & & \\ & & & & & $	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ ACGRVNAWQ ACGRVNAWQ AGGRVNAWG B13 β1 FONG - KV LODG - HT LODG - KV LONG - KV RLKSSMVD	<u>epitope</u> <u>β6</u> *C. A*- PQVNNPKEWLQVDF PQVNDPKQWLQVDL PQVNDPKQWLQVDL PQANNPKEWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL AKANNNKQWLQVDL KUPQGNQDST KVFQGNQDSST KVFQGNQDSST KVFQGNQDSST KVFQGNQDSST KVFQGNRDSST KIFEGNTNTKG KIFEGNTNTKG	2234 2075 91
B. C2 domains HphA hviliC2 2173 pVIIIC2 mVIIIC2 hVIIIC2 hFVC2 hVIIIC1 2020 HphA hVIIIC2 2235 pVIIIC2 mVIIIC2 mVIIIC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2	β1 SCSMPLGMESKAISD SCSMPLGMESKAISD SCSMPLGMESKVISD SCSMPLGMESKAISD GCSTPLGMESKAISD GCSTPLGMESGKIEN KCQTPLGMESGKIEN KCQTPLGMASGHIRD β7 β8 QKTMKV TGVTTQG V QKTMKV TGVTTQG V QKTMKV TGITTQG V LKIKKI TAIITQG Q LKIKKI TAIITQG Q LKIKKI TAIVTQG Q	Nβ2 β3 AQITASSYFTNMF/ AQITASSYFTNMF/ AQITASSYFSML/ AQITASSYFSML/ AQITASSFKKSWA (QITASSFKKSWA (QITASSFKKSWA (QITASSFKKSWA (QITASSFKKSWA (QITASSFX (QITASSFX (	$\begin{array}{c c} & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & & \\ & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & &$	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ YSGSINAWS β13 β1 FONG - KV LOBG - HV LOBG - KV LOBG - KV LONG - KV RLKSSMVD RCKSSMVD RCKSSMVD	epitope         β6           M         *C         A*           M         *C         A*           PQVNNPKEWLQVDE         PQVNDPKQWLQVDE           PQVNDPKQWLQVDE         PQVNDPKQWLQVDE           PQANNPKEWLQVDE         AKANNNKQWLQUDE           AKANNNKQWLQIDL         T KEPFSWIKVDL           AKANNNKQWLQUDE         X           T KEPFSWIKVDL         4           4         B15           V         KVFQGNQDSST           KVFQGNQDSST         KVFQGNRDSST           KVFQGNRDSST         KVFQGNRDSST           KIFEGNTNTKG         KIFEGNSTKG           KIFEGNSTKG         KIFEGNNWYBG	2234 2075 91
B. C2 domains HphA hvilic2 2173 pvilic2 mvilic2 mvilic2 hvilic2 hvilic2 bFvc2 hvilic1 2020 HphA hvilic2 2235 pvilic2 mvilic2 evilic2 hvilic2 hvilic2 bFvc2 hvilic2 h	β1 J D* T SCSMPLGMESKAISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKVISD GCSTPLGMESGKIEN KCQTPLGMESGKIEN KCQTPLGMESGKIEN COTTLO	Nβ2         β3           P         P           AQITASSYFTMF/         QITASSYFTMF/           QITASSYKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           VGITASSFKSWG         QITASSFKSWG           VSLSSMYVKET         YKSLFSMYVKET           VKSLFSMYVKET         YKSLSSEMYVKSY           VKSLSSEMYVKSY         YKSLSSEMYVKSY           YKSLSSEMYVKSY         YKSLSSEMYVKSY           YKSLSSEMYVKSY         YKSLSSEMYVKSY           YKSLSSEMYVKSY         YKSLSSEMYVKSY           YKSLSSEMYVKSY         YKSLSSEMYVKSY	β4 L im β4 L im T WSPSQARLHI T WSPSQARLHI T WSPSQARLHI T WSPSQARLHI BDY WEPFRARLNI BDY WEPFRARLNI BDY WEPFLARLNI BY WEPFLARLNI BY WEPFLARLNI BY WEPFLARLNI BY WEPFLARLNI BY WEPFLARLNI BY WEPFLARLNI DY WEPFLA	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ ASGSINAWS β13 β1 FONG - KV LÓDG - KV LÓNG - KV LONG - KV	epitope         β6           M         *C. Å*           PQVNNPKEWLQUDE           PQVNDFKQWLQUDE           PQVNDFKQWLQUDE           PQANNPKEWLQUDE           AKANNKQWLEDLE           AKANNKQWLEDLE           AKANNKQWLEDLE           KVFQGNQDST           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNNTKG           KIFEGNNTKG           KVFEGNNTSSG	2234 2075 91
B. C2 domains HphA hVIIIC2 2173 pVIIIC2 mVIIIC2 mVIIIC2 hFVC2 hVIIIC1 2020 HphA hVIIIC2 2235 pVIIIC2 eVIIIC2 eVIIIC2 hFVC2 hVIIIC2 hFV	$\begin{array}{c c} & \beta 1 & \\ \mathbf{I} & \mathbf{D}^{\star} & \mathbf{T} \\ & \text{SCSMPLGMESKALSD} \\ & \text{SCSMPLGMESKVISD} \\ & \text{SCSTPLGMESKVISD} \\ & \text{SCSTPLGMESKALSD} \\ & \text{GCSTPLGMESKALSD} \\ & \text{GCSTPLGMESGKIENI } \\ & \text{GCSTPLGMESGKIENI \\ & \text{GCSTPLGMESGKIENI } \\ & \text{GCSTPLGMESGKIENI } \\ & GCSTPLGMESGKIENI \\ & \text{GCSTPLGMESGKIENI \\ & \text{GCSTPLGME$	Nβ2         β3           AQITASSYFTME/         P           AQITASSYFTME/         QUITASSYFTME/           YQITASSYFTME/         QUITASSYFKSWWC           YQITASSFKKSWWC         QUITASSFKKSWWC           YQITASSFKKSWWC         QUITASSFKKSWWC           YQITASSFKKSWWC         QUITASSFKKSWWC           YQITASSFKKSWWC         QUITASSFKKSWWC           YQITASSFKKSWWC         Y           YSLLISMYVKEPI         YKSLLSSWYVKEPI           YKSLLSSWYVKEYI         YKSLSSEMYVKSYI           YKSLSSEMYVKSYI         YKSLSSEMYVKSYI           YKSLSSEMYVKSYI         YKSLSSEMYVKSYI           YKSLSSEMYVKSYI         YKSLSSEMYVKSYI           YKSLSSEMYVKSYI         YKSLSSEMYVKSYI	β4 L im β4 L im T - WSPSKARLHI T - WSPSQARLHI T - WSPSQARLHI T - WSPSQARLHI DY - WEPSLARLNI BY - WEPSLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BJ1 β12 T SSQDGHQWTLF USSSQDGRWTLF USSSQDGRWTLF USSSQDGHWTQI USSSQDGHWTQI USSSQDGHWTQI USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGHWTLF USSSQDGWWFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWKFY USSSQDGVWFY USSSQDF USSSS USSSS USSS USSS USSSS USSS USSS USSSS USSS USSS	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ ACGRVNAWQ ACGRVNAWQ ASGSINAWS β13 β1 FONG - KV LODG - KV LODG - KV LODG - KV LODG - KV LONG - KV RLKSSMVD RCKSSMVD RCKSSMVD RCKSSMVD	epitope         β6           YC         A*           PQVNNPKEWLQVDF           PRVSSAEEWLQVDI           PQVNDPKQWLQVDI           PQANNPKEWLQVDI           PQANNPKEWLQVDI           PQANNPKEWLQVDI           AKANNKQWLQVDI           AKANNKQWLQVDI           AKANNKQWLQVDI           AKANNKQWLQVDI           KVFQGNQDST           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNSTST           KVFQGNNDSST           KIFEGNNTKG           KIFEGNNTKG           KIFEGNNTKG           KIFEGNNTKG           KIFEGNNTKG	2234 2075 91
B. C2 domains HphA hviliC2 2173 pVIIIC2 mVIIIC2 mVIIIC2 hFVC2 hVIIIC1 2020 HphA hVIIIC2 2235 pVIIIC2 mVIIIC2 cVIIIC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hVIIIC2 hFV	β1	Nβ2         β3           P         P           AQITASSYFTNMP/ AQITASSYLSSML/ AQITASSYLSSML/ CQITASSFKKSWC           YESSEKSWC           YESLISSFKKSWC           PQITAS           PQITAS           PGITASSYFKSWC           YESLISSFKKSWC           YESLISSFKYCEFI           YESLISSFYVEFI           YESLSSEMYVEST           YESLSSEMYVEST           QUESSEYYSEY           QUESSEYYSEY	β4         L         L           G         G         G           T         -WSPSQARLHI         WSPSQARLHI           T         -WSPSQARLHI         WSPSQARLHI           T         -WSPSQARLHI         WSPSQARLHI           T         -WSPSQARLHI         MS           BDY - WEPFLARLNN         WY-WEPFLARLNN           WY-WEPFLARLNN         MY-WEPFLARLNN           WYSSQCHAWKLARLHY         B12           T         -           I.SSSQDGHOWTLF         I.SSSQDGHWTLF           I.SSSQDGHNWTLF         I.SSSQDGHNWTLF           I.SSSQDGHNWTLF         YSSQUYSDQCYAWKPY           THYSLOGKKWQTY         TMYSLDGKKWQTY	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ YSGSINAWS β13 β1 FONG - KV LODG - HV LODG - HV LODG - KV LODG - KV RLKSSMVD ROKSSMVD RGSSTGTI	epitope         β6           M         *C         A*           M         *C         A*           PQVNNPKEWLQVDF         PRVSSAEsWLQVDF           PQVNDPKQWLQVDL         PQUNDPKQWLQVDL           PQANNPKEWLQVDF         AKANNNKQWLQUDL           AKANNNKQWLQIDL         T KEPFSWIKVDL           4         β15         V           KVFQGNQDSST         KVFQGNQDSST           KVFQGNQDSST         KVFQGNQDSST           KVFQGNRDSST         KIFEGNTNTKG           KIFEGNNTWRG         KIFEGNNTWRG           MVFFGNVDSSG         21	2234 2075 91
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 mviliC2 hviliC2 hviliC2 hviliC2 hviliC2 2020 HphA hviliC2 eviliC2 mviliC2 eviliC2 hv	$\begin{array}{c} \beta 1 \\ \textbf{I} & \textbf{D} \star \ensuremath{\overline{\mathbf{T}}} \\ \text{SCSMPLGMESKALSD} \\ \text{SCSMPLGMESKVISD} \\ \text{SCSTPLGMESKVISD} \\ \text{SCSTPLGMESKKISD} \\ \text{GCSTPLGMESKKISD} \\ \text{GCSTPLGMESKKIEN} \\ \text{KCQTPLGMESKIEN} \\ \text{KCQTPLGMESKIEN \\ \text{KCQTPLGMESKIEN} \\ \text{KCQTPLGMESKIEN \\ \text{KCQTPLGMESKIEN} \\ KCQTPLGMESKIEN \\ \text{KCQTPLGMESKIEN \\ \text{KCQTPL$	NB2         B3           P         P           AQITASSYFTMF/         QITASSYFTMF/           QITASSYKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           QITASSFKSWG         QITASSFKSWG           VGITASSFKSWG         QITASSFKSWG           VSLSSFWYKET         YKSLFSMYVKET           VKSLFSMYVKET         YKSLSSEMYVKSY           YKSLSSEMYVKSY         YKSLSSEMYVKSY           CSLSSEMYVKSY         QISLS           Q         J318	β4 L m G G G G G G G G G G G G G	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ ASGSINAWS β13 β1 FONG - KV LÓDG - KV LÓDG - KV LÓNG - KV LONG - KV RLKSSMVD REKSSMVD REKSSMVD REKSSMVD	epitope         β6           M         C. A*           PQVNNPKEWLQUDE           PQVNDFKQWLQUDE           PQVNDFKQWLQUDE           PQVNDFKQWLQUDE           AKANNKQWLEDLE           AKANNKQWLEDLE           AKANNKQWLEDLE           KVFQGNQDST           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSST           KVFFGSNDSS           KVFFGSNDSS           KVFGSNDSS           KVFGSNDSS           KVFGSNDSS           KVFGSNDSS           KVFGSNDSS           KVFSS           KVFSS           KVFSSS           KVFSSS           KVFSSS           KVFSSSS           KVFSSSS	2234 2075 91
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 mviliC2 hFVC2 hVIIC1 2020 HphA hviliC2 cviliC2 hviliC2 cviliC2 hFVC2 hVIIC2 cviliC2 hFVC2 hFVC2 hVIIC2 hFVC2	β1 I D* T SCSMPLGMESKAISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKAISD GCSTPLGMESGKIEMI GCSTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI KCQTPLGMESGKIEMI CSTPLGMESGKIEMI SCSTP	NB2         B3           AQITASSYFTMMF/ QUITASSYFTNMF/ AQITASSYFKSWMC         AQITASSYFTNMF/ AQITASSFKSWMC           AQITASSFKSWMC         AQITASSFKSWMC           VKSLLISMYVKEFI         XKSLFSMFVKSYT           XKSLSSEMYVKSYT         XKSLSSEMYVKSYT           AQKFSSLYISQFT         Q           318         L           LETHPOSWVHOTAT         A	β4 L im β4 L im T - WSPSKARLHI T - WSPSQARLHI T - WSPSQARLHI T - WSPSQARLHI DY - WEPSLARLNI BY - WEPFLARLNI BY - WEPFLARLNI DY - WEP	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ ACGRVNAWQ ACGRVNAWG SGSINAWS β13 β1 FONGKU LODGHT LYNGKU LODGHT LYNGKU LODGKU RCKSSMVD RCKSSMVD RCKSSMVD RCKSSMVD RCKSSMVD RCKSSMVD RCKSSMVD RCKSSMVD	epitope         β6           *C. A*           PQVNNPKEWLQUDF           PRVSSAEEWLQUDE           PQVNDPKQWLQUDE           PQANNPKEWLQUDE           PQANNPKEWLQUDE           AKANNKQWLEUDE           AKANNKQWLQUDE           AKANNKQWLQUDE           AKANNKQWLQUDE           KVFQGNQDSST           MVFFGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           Gidentity           NUMP           IAO           IAO	2234 2075 91 34
B. C2 domains HphA hviliC2 2173 pVIIIC2 mVIIIC2 eVIIIC2 hFVC2 hVIIIC1 2020 HphA hviliC2 eVIIIC2 eVIIIC2 hVIIC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hVIIC2 hFVC2 hVIIC2 hFVC2 hFVC2 hVIIC2 hFVC2 hFVC2 hVIIC2 hFVC2	$\begin{array}{c c} \beta 1 & & \\ \mathbf{I} & \mathbf{D} \star \mathbf{\overline{T}} \\ \text{SCSMPLGMESKAISD} \\ \text{SCSMPLGMESKVISD} \\ \text{SCSMPLGMESKVISD} \\ \text{SCSMPLGMESKVISD} \\ \text{SCSMPLGMESKAISD} \\ GCSTPLGEDGRIQDI \\ GCSTPLGEDGRIQDI \\ GCSTPLGMESGKIENI \\ KCQTPLGMASGHIRDI \\ GCSTPLGMESGKIENI \\ KCQTPLGMESGKIENI \\ KCQTPLGMESGKIENI \\ KCQTPLGMESGKIENI \\ KCQTPLGMESGKIENI \\ KCQTPLGMESGKIENI \\ KCQTPLGMESGKIENI \\ CSTPLGMESGKIENI \\ CSTPLGMESGKIENI \\ CSTPLGMESGKIENI \\ CSTPLGMESGKIENI \\ CSTPLGMESGKIENI \\ KCQTPLGMESGKIENI \\ CSTPLGMESGKIENI \\ CSTPLG$	Nβ2         β3           AQITASSYFTNMP/ AQITASSYLSNIF/ TOITASSYLSNIF/ TOITASSYLSNIF/ AQITASSFKSWMC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSFKSWC (QITASSF (QITASSFKSWC (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITASSF (QITAS) (QITASSF (QITASSF (QITAS) (QITASSF (QITAS) (QITASSF (QITAS) (QITASSF (QITAS) (QITAS) (QITASSF (QITAS) (QITASSF (QITAS) (QITA	β4         L         β4         L           T         -WSPSKARLHI         -WSPSQARLHI         -WSPSQARLHI           T         -WSPSQARLHI         -WSPSQARLHI         -WSPSQARLHI           T         -WSPSQARLHI         -WSPSQARLHI         -WSPSQARLHI           SDY - WEPFRARLN2         -WSPSQARLHI         -WSPSQARLHI           NY - WEPFLARLN2         -WSSQDGRWLARLHY         -WSSSQDGHQWTLF           JISSSQDGHQWTLF         -ISSSQDGHWTLF         -ISSSQDGHWTLF           ISSSQDGHWTLF         -WSSQDGRWTLF         -WSSQDGRWTLF           ISSSQDGHWTLF         -WSSQDGRWTLF         -WSSQDGRWTLF           ISSSQDGHWTLF         -WSSQDGRWTLF         -WSSQDGRWTLF           ISSSQDGHWTLF         -WSSQDGRWTLF         -WSSQDGRWTLF           ISSSQDGHWTLF         -WSSQDGTDWKPY         -WSSQDGTDWKPY           THYSDQCTDWKPY         -WSSLDGKKWQYY         T           T	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ AGCRVNAWQ AGCRVNAWG AGCRVNAWG FONG - KV LODG - HV LODG - HV LODG - KV LODG - KV LODG - KV LONG - KV LONG - KV SKSSWVD ROKSSMVD RGNSTGTI %	epitope         β6           MNNPKEWLQUDF           PQVNNPKEWLQUDF           PRVSSAEEWLQUDE           PQVNDPKQWLQUDL           PQNNPKEWLQUDE           AKANNNKQWLGUDL           AKANNNKQWLQUDL           AKANNNKQWLQUDL           TKEPFSWIKUDL           4 β15           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNNDSST           KVFQGNNVRG           KIFEGNNTKG           KIFEGNNTKG           MVFFGNVDSSG 21           MUFFGNVDSSG 21           KIEND           MUFFGNVDSSG 21           MUFFGNVDSSG 21	2234 2075 91 34 of
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 mviliC2 hviliC2 hviliC2 bFVC2 hviliC1 2020 HphA hviliC2 cviliC2 mviliC2 mviliC2 hvil	$\begin{array}{c c} \beta 1 & & \\ I & & \\ SCSMPLGMESKALSDSCSMPLGMESKALSDSCSTPLGMESKVISDSCSTPLGMESKVISDSCSTPLGMESKKISDGCSTPLGMESGKIENGCSTPLGMESGKIENCCTPLGMESGKIENKCQTPLGMASGHIRDGCSTPLGMESGKIENCCTPLGMESGKIEN$	NB2         B3           P         P           AQITASSYFTNMFA         QUITASSYFTNMFA           QUITASSYFTNMFA         QUITASSYFKSWWC           QUITASSFKSWWC         QUITASSFKSWWC           QUITASSFKSWWC         QUITASSFKSWWC           QUITASSFKSWWC         QUITASSFKSWWC           QUITASSFKSWWC         QUITASSFKSWWC           QUITASSFKSWWC         QUITASSFKSWWC           VGITASSFKSWWC         QUITASSFKSWWC           VSLLSSMYVKET         YSLLSSMYVKET           VKSLLSSMYVKSYT         YSLSSEMYVKSYT           CSLSSEMYVKSYT         QUITASSFKSWYC           QUITASSFKSWYC         QUITASSFKSWYC           QUITASSFKSWYC         QUITASSFKSWYC           QUITASSFKSWYC         QUITASSFKSWYC           QUITASSFKSWYC         QUITASSFKSWYC           QUITASSFKSWYC         QUITASSFKSWYC           QUITASSFKSWYC         QUITASSFKSWYC           QUITASSFKWYC         QUITASSFKWYC	β4         L         β4         C           G         G         G         G           Marcel Sepson         G         G         G           Marcel Sepson         Marcel Sepson         G         G           Marcel Sepson         Marcel Sepson         Marcel Sepson         G           Marcel Sepson         S         RMEVLGCEAQDLY         RLEVLGCEAQDLY	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ AOGRVNAWQ ASGSINAWS FONG - KV LODG - HT LYNG - KV LODG - HT LYNG - KV LONG - KV RLKSSMVD REKSSMVD REKSSMVD RGSSTGTI % 2332	epitope         β6           M         *C. Å*           PQVNPKEWLQUDL           PQVNPKEWLQUDL           PQVNPKEWLQUDL           PQVNPKEWLQUDL           AKANNKQWLEDL           AKANNKQWLEDL           AKANNKQWLEDL           AKANNKQWLEDL           KANNKQWLEDL           KANNKQWLEDL           KANNKQWLEDL           KOFOGNQDST           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           MVFFGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           160           79           160           83	2234 2075 91 34 of
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 mviliC2 hFVC2 hFVC2 hviliC2 2020 HphA hviliC2 2235 pviliC2 mviliC2 hFVC2	β1 I D* T SCSMPLGMESKAISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKVISD SCSTPLGMESKISD GCSTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI KCQTPLGMESGKIENI CSTPLGMESGKIENI SCSTPLG	NB2         B3           AQITASSYFTMMF/ AQITASSYFTNMF/ AQITASSYFTNMF/ AQITASSYFKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (QITASSFKKSWWG (XILTSMFVKEFI) /KSLLISMYVKEFI /KSLLISMYVKEFI /KSLLISMYVKEFI /KSLLISMYVKEFI /KSLLISMYVKEFI /KSLLISMYVKEFI /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEMYVKSY: /KSLSSEM/VKSY: /KSLSSSM/VKSY: /KSLSSSM/VKSY: /KSLSSSM/VKSY: /KSLSSM/VKSY: /KSLSSSM/VKSY	β4 L im β4 L im β4 L im T - WSPSQARLHI T - WSPSQARLHI T - WSPSQARLHI T - WSPSQARLHI DY - WEPSLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BY - WEPFLARLNI BY - WEPFLARLNI DY - WEPFLARLNI	hibitor β5 Logranawr Logranawr Logranawr Logranawr Logranawr Aggrwnawg Aggrwnawg Sgallawr Fong - Kv Long - Kv Sswyd Reksswyd Romstati 2332	epitope         β6           *C. A*           PQVNNPKEWLQUDF           PRVSSAEEWLQUDE           PQVNDPKQWLQUDE           PQVNDPKQWLQUDE           PQVNDPKQWLQUDE           AKANNNKQWLQUDE           AKANNNKQWLQUDE           AKANNNKQWLQUDE           AKANNNKQWLQUDE           KVFQGNQDSST           KVFGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           6           1600           83           160           83           160	2234 2075 91 34 of
B. C2 domains HphA hviliC2 2173 pVIIIC2 mVIIIC2 mVIIIC2 hFVC2 hFVC2 hVIIIC1 2020 HphA hviliC2 cVIIIC2 hVIIIC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hVIIIC2 2292 pVIIIC2 mVIIIC2 cVIIIC2 hVIIC2 hVIC2 hVIC2	β1	NB2         B3           AQITASSYFTMMP/ AQITASSYLSNIF/ TOITASSYLSNIF/ TOITASSYLSNIF/ AQITASSFKSWMC (XSUTASSFKSWMC (QITASSFKSWMC (XSUTASSFKSWMC (QITASSFKSWMC (XSUTASSFKSWMC (QITASSFKSWMC (XSUTASSFKSWMC (QITASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (QITASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (QITASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSFKSWMC (XSUTASSF (XSUTASSF (XSUTAS) (XSUTASSF (XSUTASSF (XSUTAS) (XSUTASSF (XSUTAS) (XSUTASSF (XSUTAS) (XSUTASSF (XSUTAS) (XSUTASSF (XSUTAS) (XSUTAS) (XSUTASSF (XSUTAS) (XSUTAS) (XSUTAS) (XSUTASSF (XSUTAS) (XS	β4         L         magnetic           β4         L         magnetic           T         -WSPSQARLHI         T           T         -WSPSQARLHI         T           T         -WSPSQARLHI         T           T         -WSPSQARLHI         T           SDY         -WEPFRARLN2           DDY         -WEPFLARLN2           QYGQWAPKLARLH2         B12           β11         B12           T         JSSSQDGHQWTLF           JSSSQDGHWTLF         ISSSQDGHWTLF           ISSSQDCHWWTLF         SIQYSQQVAWKPY           THYSDQCTDWKPY         THYSLDCKKWQTY           T         S           RMEVLGCEAQDLY         RLEVIGCEAQLY           RLEVIGCEAQUY         RLEVIGCEAQUY           RLEVIGCEAQUY         RLEVIGCEAQUY	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AQGRVNAWQ XGGSINAWS β13 β1 FONGKV LODGHV LODGKV LODGKV LODGKV LONGKV LONGKV LONGKV RKSSMVD RGSSMVD RGSSMVD RGSSMVD	epitope         β6           MNNFKEWLQVDF           PQVNNPKEWLQVDF           PRVSSAEEWLQVDL           PQVNDPKQWLQVDL           PQNNPKEWLQVDL           PQNNPKEWLQVDL           PQNNPKEWLQVDL           AKANNNKQWLGUDL           TKEFFSWIKVDL           4 β15           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNNDSST           KVFQGNNVRG           KIFEGNNNVRG           MVFFGNVDSSG 21           6 identity           number           identity           160           83           60           81           60	2234 2075 91 34 of
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 mviliC2 hviliC2 hviliC2 bFVC2 hviliC1 2020 HphA hviliC2 cviliC2 hvil	β1	NB2         B3           P         AQITASSYFTMP/ AQITASSYFTMF/ AQITASSYLSSML           TQITASSYLSSML           QUITASSFKSWW           VGITASSFKSWW           VGITASSFKSWW           VGITASSFKSWW           VGITASSFKSWW           VSLLSSWYKEFI           VKSLFTSMFVKEFI           VKSLSSEMYVKSYI           VKSLSSEMYVKSYI           SKSLSSEMYVKSYI           CSLSSEMYVKSYI           Q           318           L           LRIHPOSWVHQIAI           RIHPOIWEHQIAI           RLHPOSWAHIAI           RUHPOSWAHIAI           RUHPOSWAHIAI           RUHPOSWAHIAI           RUHPOSWAHIAI           RUHPOSWAHIAI           RUHPOSWAHIAI           RUHPOSWAHIAI           RUHPOSWAHIAI	β4 L im β4 L im β4 L im G G G G G G G G G G G G G	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR LOGRTNAWR AOGRVNAWQ AOGRVNAWQ ASGSINAWS β13 β1 FONG - KV LODG - HT LYNG - KV LODG - HT LYNG - KV RLKSSMVD RCKSSMVD REKSSMVD REKSSMVD RGSSTGTI	epitope         β6           M         *C. Å*           PQVNNPKEWLQUDE           PRVSSAEEWLQUDI           PQVNDFKQWLQUDI           PQVNDFKQWLQUDI           AKANNKQWLEDLI           AKANNKQWLEDLI           AKANNKQWLEDLI           AKANNKQWLEDLI           KANNKQWLEDLI           KANNKQWLEDLI           KOPGRODST           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           KVFGGNQDSST           MVFFGNVDSSG           MVFFGNVDSSG           Sidentity           number           itch VIII           residue           (100)         160           81         160           81         160           40         160	2234 2075 91 34 of
B. C2 domains HphA hviliC2 2173 pviliC2 mviliC2 mviliC2 hFVC2 hFVC2 hviliC2 2020 HphA hviliC2 2235 pviliC2 hVIIC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hFVC2 hVIIC2 2076 HphA hviliC2 2292 pviliC2 mVIIC2 hFVC2 hFVC2 hVIIC2 2076 HphA	β1         D*           I         D*         T           SCSMPLGMESKNISD         SCSMPLGMESKNISD           SCSTPLGMESKVISD         SCSMPLGMESKVISD           SCSMPLGMESKVISD         SCSMPLGMESKVISD           SCSMPLGMESKVISD         SCSMPLGMESKVISD           SCSMPLGMESKVISD         SCSMPLGMESKVISD           GCSTPLGMESKVISD         SCSMPLGMESKVISD           GCSTPLGMESKNISD         GCSTPLGMESKVISD           GCSTPLGMESKNISD         GCSTPLGMESKVISD           GCSTPLGMESKNISD         GCSTPLGMESKVISD           GCSTPLGMESKNISD         GCSTPLGMESKVISD           GCSTPLGMESKNISD         GCSTPLGMESKVISD           GCSTPLGMESKVISD         GCSTPLGMESKVISD           GCSTPLGMESKVITGITOG         N           QKTMKV TGVITOG         Y           RKTMKV TGITTOG         Y           RKTMKV TGITTOG         Y           RKTMKV TGITTOG         Y           LKIKKI TAIVTOG         L           LKIKKI TAIVTOG         L*           B16         B17         G           L*         C*           PVWNSLDPPLIT <ry< td="">         PVWNALDPPLYA           PVKNFNPPIS         RT           HVKNFFNPPIS         RT  </ry<>	NB2         B3           P         I           AQITASSYFTMF/         I           AQITASSYFKSWW         I           AQITASSFKSWW         I           VSLISSWW         I           VKSLISMFVKEPI         I           VKSLSSEMYVKSY:         I           SKSLSSEMYVKSY:         I           C         Q           318         L           LRIHPTSWAQHIAI         I           JRUHPTSWAQHAHAI         I           JRUHPQSWHOIAI         I           THIPTSWAQHAHAI         I           JRUHPANNON         I           JRUHPANNON         I           LINIPTWNOSITI         I	β4 L im β4 L im β4 L im T - WSPSQARLHI T - WSPSQARLHI T - WSPSQARLHI T - WSPSQARLHI DY - WEPSLARLNI BY - WEPSLARLNI DY - WEPSLARLNI	hibitor β5 Logrinawr Logrinawr Logrinawr Logrinawr Logrinawr Aggrvnawg Aggrvnawg Sissinaws β13 β1 Fong - Ku Lodg - Hi Lyng - Ku Long - Ku Rom - K	epitope         β6           *C. A*           PQVNNPKEWLQVDF           PRVSSAEEWLQVDL           PQVNDPKQWLQVDL           PQVNDPKQWLQVDL           PQANNPKEWLQVDL           PQANNPKEWLQVDL           PQANNPKEWLQVDL           AKANNKQWLQUDL           AKANNKQWLQUDL           AKANNKQWLQUDL           KVFQGNQDSST           KVFGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           83           160           83           160           81           60           81           60           40           160           160	2234 2075 91 34
B. C2 domains HphA hviliC2 2173 pVIIIC2 mVIIIC2 mVIIIC2 hFVC2 hVIIIC1 2020 HphA hvIIIC2 2235 pVIIIC2 hVIIIC2 cVIIIC2 hFVC2 hFVC2 hFVC2 hVIIIC2 2292 pVIIIC2 mVIIIC2 cVIIIC2 hVIIIC2 hVIIIC2 hFVC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hFVC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIIIC2 hVIC2 hV	$\begin{array}{c c} \beta 1 & & \\ I & D \star & T \\ SCSMPLGMESKAISDSCSMPLGMESKAISDSCSMPLGMESKVISDSCSMPLGMESKVISDSCSMPLGMESKVISDSCSMPLGMESKAISDGCSTPLGMESGKIENIKCQTPLGMASGHIRDIGCSTPLGMESGKIENIKCQTPLGMASGHIRDIGCSTPLGMESGKIENIKCQTPLGMASGHIRDIGCSTPLGMESGKIENIKCQTPLGMASGHIRDIGCSTPLGMESGKIENIKCQTPLGMASGHIRDIGCSTPLGMESGKIENIKCQTPLGMASGHIRDIGCSTPLGMESGKIENI$	NB2         B3           AQITASSYFTMMP/ AQITASSYFTMMP/ AQITASSYLSNIF/ TOITASSYLSNIF/ TOITASSYLSNIF/ AQITASSFKSWMC (XSI) (XS	$\begin{array}{c} \mathbf{\beta}^{4} \\ \mathbf{\beta}^{5} \\ \mathbf{\beta}^{5$	hibitor β5 LOGRSNAWR LOGRTNAWR LOGRTNAWR AQGRVNAWQ AQGRVNAWQ XSGSINAWS β13 β1 FONGKV LODGHV LODGKV LODGKV LODGKV LONGKV LONGKV LONGKV SSSWVD RGSSMVD RGSSMVD RGSSTGTI	epitope         β6           Yey         K           PQVNNPKEWLQVDF           PRVSSAEWLQVDL           PQVNDPKQWLQVDL           PQVNDPKQWLQVDL           PQNNPKEWLQVDL           PQNNPKEWLQVDL           PQNNPKEWLQVDL           AKANNNKQWLBIDL           AKANNNKQWLQIDL           TKEPFSWIKVDL           4 β15           V           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFQGNQDSST           KVFGGNVDSSG           MVFFGNVDSSG           MVFFGNVDSSG           S1           60           83           160           81           160           43           160           43	2234 2075 91 34 of

C1 and C2 were introduced in the above alignment to eliminate large distances (greater than 8 Å) between the C $\alpha$  atoms of the C1 model and C2 template. The structures of loops corresponding to the gaps in Figure 1 were refined separately starting from the initial model generated by Modeller. Models with the lowest energy and best stereogeometry were selected for further refinement. The final model displayed no outliers in the Ramachandran plot produced by Procheck.<sup>37</sup> Solvent accessibility of the side chains in the molecular models was analyzed with the use of the program Naccess,<sup>38</sup> where values range from 0% (completely buried) to 100% (completely exposed) as compared with the same residue in a fully extended tripeptide sequence, "A-X-A" (A is single-letter amino acid code and X, the residue being considered).

#### **Results and discussion**

#### **Mutation identification**

Exon fragments coding for the C domains, exons 20 through 26 and flanking regions of the factor VIII gene, were examined in 76 families with hemophilia A. A single fragment with a heteroduplex band was found in 30 families. Direct sequencing revealed a single

point mutation in each of these positive fragments. Among the 20 distinct mutations, 13 were novel, having not been reported to an international database of hemophilic factor VIII mutations.<sup>15</sup> We found 6 non-missense mutations in 8 families with severe hemophilia A. Of these families, 2 had a single base microdeletion (A in codons 2184 to 2185 and T in codons 2139 to 2140) predicting frameshifts and premature termination. We found 4 nonsense mutations in the other 6 families (in codons W2111 [single-letter amino acid code], R2116 in 3 families each with a different haplotype, W2203; and R2209). Of these nonsense mutations, R2209 and one of the families with R2116 were previously reported.<sup>39</sup> The remaining 22 families had 14 distinct missense mutations, and 9 of these were novel (Table 2). Two of these mutations in C1 were associated with normal levels of dysfunctional factor VIII antigen, and an additional 3 in C2 had significant excess antigen over clotting activity, indicating at least partial dysfunction. vWF-bound antigen level was less than the total factor VIII antigen in all mutations except for 3: R2164, C2304, and T2320. In T2320, it was actually enhanced.

Table 2. Newly reported hemophilic C-domain missense mutations

	Mutations	Fa	Factor VIII (% normal)				
Exon	(no. families)	VIII:C	VIII:Ag	VIII-vWF	(severity)		
C1 domain							
22	Q2087E†	11	11	1	Mild		
22	R2090C†	14	107	2	Mild		
23	R2150C† (3)	33	30	<1	Mild		
23	R2159C (2)	20	104	<1	Mild		
23	R2163C	1	1	<1	Moderate		
23	M2164R†	2	2	3	Moderate		
C2 domain							
24	S2173l† (3)	16	41	14	Mild		
24	A2201P†	7	36	1	Mild		
25	V2232A†	9	8	2	Mild + Inh		
26	P2300L (2)	18	18	5	Mild + Inh		
26	R2304C (3)	2	3	2	Mod + Inh		
26	R2304G†	8	3	<1	Moderate		
26	R2307Q	10	10	<1	Mild		
26	R2320T†	5	20	100	Moderate		

Bold Ag levels indicate dysfunctional protein; bold vWF levels, bound Ag, where significantly different from total Ag; *Inh*, high titer inhibitor in at least one of the reported patients with this mutation.

\*Factor VIII determined in patient plasma as clotting activity (VIII:C), total factor VIII antigen (VIII:Ag), or vWF-bound factor VIII antigen (VIII-vWF) and reported as percentage of clotting reactivity in a pool of normal donor plasmas.

†Novel mutation, not previously reported in database.15

#### Homology model of the C1 domain

The final C1 model is similar to the C2 template with a root mean squared deviation (rmsd) of 0.9 Å between equivalent C $\alpha$  atoms in the superposition of the 2 domains (Figure 2). Loop structures connecting the  $\beta$ -strands are highly similar except for 3 fewer carboxy-terminal residues in C1 than in C2, and 3 loops with length differences relative to C2, where the model may be less predictive of C1 structure. The loops connecting  $\beta5$  to  $\beta6$  and  $\beta13$  to  $\beta14$  are 2 residues shorter and longer in C1 than C2, respectively. A third loop where the C1 and C2 domain structures differ is at the first of 2  $\beta$ -strand hairpins ( $\beta3$ -turn- $\beta4$ ) in C2 that has 4 residues fewer and only a  $\beta$ -turn in C1 (Figure 2).



Figure 2. Homology model of the factor VIII C1 domain. The homology model of the C1 domain (red) was created as described in the text and is superimposed onto the crystal structure of the factor VIII C2 domain (blue) and is shown in stereo views. The locations of the 19  $\beta$ -strand structures in C2 are indicated,<sup>22</sup> referred to as  $\beta'$  strands in C1. The major differences are in the loop lengths (Figure 1), which are indicated by gray arrows. The loop corresponding to the  $\beta3$ - $\beta4$  hairpin structure within the C2 domain, a putative membrane-binding surface, is shortened to a turn in C1 (thus, there is no  $\beta3'$  or  $\beta4'$  strand in C1; homologous numbering of strands 5' through 19' are preserved). The  $\beta5$ -to- $\beta6$  loop is 2 residues shorter in C1. The loop between  $\beta$ -strands 13 and 14 is 2 residues longer in C1.

For the C2 domain, a membrane-binding surface has been proposed that consists of the 2 tight  $\beta$ -hairpin turns that present M2199, F2200, L2251, and L2252 to solvent and a loop that presents V2223 on the same surface.<sup>22</sup> These protruding hydrophobic residues are surrounded by an underlying cleft lined with basic side chains, consistent with anionic phospholipid membrane binding. In the C1-domain model, the loops corresponding to this surface are truncated and are less hydrophobic overall. F2093 is solvent-exposed at the homologous surface and is surrounded by basic side chains (R2090, R2159, R2163, and K2092). It remains to be established whether the C1 domain contacts the membrane surface when native factor VIIIa is bound. Modeling studies based on electron microscopy and 2-D electron diffraction of factors Va and VIIIa have been interpreted in favor of a bound orientation in which only the C2 domain contacts the membrane surface.<sup>40,41</sup> However, a sterically reasonable model of a C1-C2 construct can be generated that would form a surface with the  $\beta$ -loops of C1 described above in the same plane as their counterpart hairpins in C2; therefore, C1 might also contact the membrane (B.L.S., unpublished observations, November 1999). Resolution of this issue awaits structural analyses of larger factor VIIIa constructs in the presence and absence of an associated membrane bilayer and structural data on the C1-C2 interface.

The C1-domain protein core is predominantly hydrophobic: out of 44 residues with less than 10% surface exposure, 34 (77%) are nonpolar, aromatic, or glycine residues. The remaining 10 residues are more polar, such as T2086, which is conserved in the structure of C2 (T2245). Other buried polar residues in C1 include 3 sites of hemophilic missense mutations: S2069, Q2087, and N2129. In comparison, the C2 domain core contains 47 residues with less than 10% surface exposure; of these, 36 (77%) are hydrophobic, aromatic, or glycine residues. Three of the remaining 11 buried polar residues (R2209, Q2246, and R2320) are sites of hemophilic missense mutations.

#### Hemophilic mutation localization

Clinical data and the factor VIII antigen levels (where available) for hemophilic missense mutations are listed in Table 3, combining the current series (Table 2) with previously reported C-domain missense mutations.<sup>15</sup> As shown, 57 missense mutations have been reported in the factor VIII C domains, corresponding to substitutions at 43 different residues. These sites were localized to the factor VIII C2 domain crystal structure and the C1 homology model (Figure 3). Missense mutation sites are evenly divided between the 2 C domains; together they occur at 14% of the C-domain residues. Figure 3 also distinguishes the severity of the clinical bleeding tendencies; only 20% are associated with clinically severe hemophilia A.

Of the 12 missense mutations in which at least one reported individual was clinically severe, the native side chain has over 20% surface exposure in only one residue, R2116. Therefore, C-domain missense mutations associated with severe hemophilia A most often appear to destabilize protein folding and/or alter the trafficking and secretion of factor VIII by disrupting the protein core. One would predict that these mutations should be associated with very low factor VIII antigen levels; unfortunately, most of these cases have not had antigen levels reported. The possible effects of hemophilia A missense mutations are discussed below.

#### Table 3. Molecular consequences of hemophilic C-domain missense mutations

C1 Semain           G2026E         8 (?)         Mid         UUUU         UVU         <	Mutations	Factor VIII:C(Ag) % normal	Phenotype (severity)	Factor VIII species hmpc	Factor V species hmb	Surface exposure	Localization
Gat22et         6 (?)         Mid         Boodd         0000         0000         0%         Core, Loop 1, G2179 in C2           A0339         4 (?)         Moderatin         Abaa         Abaa         Bob         Own, B/C 3129 in C2           A0339         4 (?)         Moderatin         Name         Moderatin         Name         PK           S006F         -1 (?)         Sovere         S550         ALT         Moderatin         Core, Loop 1, G2278 in C2           S006F         -1 (?)         Moderatin         TTTT<				C1 dom	ain		
Ca325V         -1 (r)         Sevele         0.093         0.93         0.94         Corr., (p) * (247) Pn C2           W2048         ? (?)         Moderatie         mmm         mm         9%         Carr., (p) * (247) Pn C2           W2048         ? (?)         Moderatie         mm<	G2026E	8 (?)	Mild				
A2035P         4 (17)         Moderate         ABAA         ABAA         ABAA         ABAA         ABAA         ABAA         ABAA         ABAA         ABAA         Core, hop 2, Var23 in C2           S2065F         -1 (1)         Severe         BBBA         ALT         3%         Core, hop 2, Var23 in C2           S2065F         -1 (1)         Mild         DTTT< TT	G2026V	<1 (?)	Severe	GGGG	GGG	0%	Core, Loop 1', G2179 in C2
Worker         ? (?)         Moderate         Work         Work         Work         PH         Core, hop 7, W2203 in C2           250766         7 (13)         Mid         pmm         pm         15%         Subme, pt, D223 in C2           263767         11 (11)         Mid         pmm         pm         15%         Subme, pt, D2246 in C2           263877         21 (11)         Mid         pmm	A2039P	4 (?)	Moderate	AAAA	AAA	6%	Core, β2', A2192 in C2
S2066F	W2046R	? (?)	Moderate	WWWW	WWW	9%	Core, loop 2', W2203 in C2
D277-60         7         Mid         Mid         non         non         non         15%         Surface, β67         D223 in C2           C2057F         11 (11)         Mid         000         0%         Core, β87, 02246 in C2           C2057F         11 (11)         Mid         0000         0%         Core, β87, 02246 in C2           C2057F         22 (7)         Mid         0000         0%         Core, β87, 02246 in C2           C2050F         11 (7)         Mid         0000         ERE         19%         Surface, β17, 5226 in C2           C2100F         9 (12)         Moderate         0000         ERE         19%         Surface, β17, 5226 in C2           V2105C         11 (7)         Mid         Inth         YYYY         YYY         36%         Surface, β17, 5226 in C2           S1129         6 (8)         Mide Inh         SSS         855         84%         Surface, β17, 6227 in C2           S21197         6 (8)         Mide Inh         SSS         855         84%         Surface, β17, 6227 in C2           S21105         6 (16)         Mide Inh         SSS         855         84%         Surface, β17, 6226 in C2           S21105         Core, β17, Mide         TTT <td>S2069F</td> <td>&lt;1 (?)</td> <td>Severe</td> <td>SSSS</td> <td>ALT</td> <td>3%</td> <td>Core, loop 5', E2228 in C2</td>	S2069F	<1 (?)	Severe	SSSS	ALT	3%	Core, loop 5', E2228 in C2
T2080N         3 (?)         Mod + hh         TTT         TT         TT         Off         Core, 89. T224 (in C2           0200FF         22 (?)         Mid         0000         000         0%         Core, 89. C224 (in C2           0200FF         11 (11)         Mid         Make         Make         0000         0%         Core, 89. C224 (in C2           0200FF         14 (107)         Mid         Make         Make         Make         0000         E8K         Sufface, 89', K224 (in C2           02100F         9 (5)         Mid         PEFF         PFF         0%         Core, 89. C224 (in C2           72016C         11 (?)         Mid         PFFF         PFF         0%         Sufface, 89', K226 (in C2           72016C         11 (?)         Mid         PFFF         PFF         0%         Sufface, 89', K226 (in C2           7216SC         11 (?)         Mid         Mid         PFFF         PFF         0%         Sufface, 81', F226 in C2           7216SC         3 (30)         Mid         BETF         PFF         0%         Sufface, 81', F226 in C2           7215SC         3 (30)         Mid         PTTT<	D2074G	7 (13)	Mild	DDDD	DDD	15%	Surface, $\beta 6'$ , D2233 in C2
C20057E*         11 (11)         Mild         COCO         COCO         OP         Core, BF: C224 in C2           C200507         22 (?)         Mild         BERE         DCC	T2086N	3 (?)	Mod + Inh	TTTT	TTT	0%	Core, $\beta 8'$ , T2245 in C2
Carbor         22 (?)         Mid         j         Could field         Description           Carbor         1 (?0)         Severe         Good         Goo	Q2087E*	11 (11)	Mild }	0000	000	0%	Core 68' 02246 in C2
G2088         1 (70)         Severe         0000         0000         0000         0000         0000         0000         00000         00000         00000         00000         00000         00000         000000         000000         000000         000000         000000         000000         000000         0000000         00000000         0000000000         000000000000000000000000000000000000	Q2087R	22 (?)	Mild	2000	222	070	0010, p0 , 02240 11 02
R2090C*         14 (07)         Mid         BERR         EXCC         66%         Surface, 107, 12249 in C2           P2101L         9 (6)         Mid         PPPP         PPP         0%         Surface, 117, 12249 in C2           P2101L         9 (6)         Mid         PPPP         PPP         0%         Surface, 117, 12249 in C2           R2119P         <1 (?)	G2088S	1 ( <b>70</b> )	Severe	GGGG	GGG	0%	Core, $\beta 8'$ , G2247 in C2
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	R2090C*	14 ( <b>107</b> )	Mild	RRRR	KKK	66%	Surface, $\beta 9'$ , K2249 in C2
F2101L         9 (b)         Mid         FFFP         FFFP         0%         Core, β11, F220b in C2           R2116P         <1 (?)	Q2100R	9 (12)	Moderate	QQQQ	EEE	19%	Surface, $\beta 11'$ , E2259 in C2
Y2105C         11 (?)         Mid + <i>hh</i> YYY         YYY         30%         Surface, β11, F2274 in C2           S2115P         6 (θ)         Mid-Mod         SS85         SS8         84%         Surface, log 13, G2278 in C2           S2115P         6 (θ)         Mid-Mod         SS85         SS8         84%         Surface, log 15, G228 in C2           R2150C*         33 (30)         Mid         BERR         BERR         2%         Core, log 15', R2280 in C2           R2150C*         33 (30)         Mid         PEPP         PEP         PFP         9%         Core, log 16', R2307 in C2           R2150C*         5.20 (6, 104)         Mid         PEPP         PEP         PFP         9%         Surface, log 16', R2307 in C2           R2159L         2 (12)         Mid         BERR         RER         5%         Surface, log 16', R230 in C2           R2159L         2 (2)         Moderate         BERR         RER         5%         Surface, log 16', R230 in C2           R2158L         2 (7)         Mid         BERR         RER         7%         Core, β19', R2320 in C2           R2158L         2 (7)         Mid         BERR         RER         7%         Core, β19', M321 in C2           <	F2101L	9 (5)	Mild	FFFF	FFF	0%	Core, β11', F2260 in C2
R2116P	Y2105C	11 (?)	Mild + Inh	YYYY	YYY	36%	Surface, $\beta 11'$ , S2264 in C2
SZ119Y         6 (8)         Mid-Mod         SSSS         SSS         SSS         SSS         SSS         SSS         SUTAGE, foop 15', f22278 in f22           RZ150C*         33 (30)         Mid         RZ150C*         33 (30)         Mid         RZ150C*         SUTAGE, foop 15', f2207 in C2           RZ150C*         33 (6)         Mid         PPPP         PPP         9%         Core, loop 15', f2203 in C2           RZ150C*         5, 20 (6, 104)         Mid         TTTT         TTT         3%         Surface, loop 18', f2201 in C2           RZ150C*         5, 20 (6, 104)         Mid         RER         ERR         5%         Surface, loop 18', f2201 in C2           RZ151         6 (7)         Mid         RER         ERR         5%         Surface, loop 18', f2201 in C2           RZ151         18 (15)         Mid         RER         ERR         7%         Core, b19', R2320 in C2           RZ1544*         2 (2)         Moderate         LLL         14'         16'         Core, b19', R2320 in C2           RZ1547*         16 (41)         Mid         SSSS         000         100%         N-terminus C2'; connects with C1           RZ1657         23 (7)         Mid         IIIII         IIII         5%	R2116P	<1 (?)	Severe	R <b>Q</b> RR	KRK	76%	Surface, β13', F2275 in C2
N2128       4 (7)       Moderate       NRN       PR       2%       Core, loop 15, N2228 in L22         R21500*       53 (30)       Mid       Mid       REFE       RER       22%       Surface, B18', R2307 in C2         R21501*       6 (6)       Mid       TTT       TTT       35%       Surface, B18', R2307 in C2         R21504*       6 (7)       Mid       TTT       TTT       35%       Surface, B19', R2301 in C2         R21504*       6 (7)       Mid       TTTT       TTT       35%       Surface, B19', R2320 in C2         R21531*       18 (15)       Mid       RERR       RER       RER       58%       Surface, B19', R2320 in C2         R21634*       <16 (7)	S2119Y	6 (9)	Mild-Mod	SSSS	SSS	84%	Surface, loop 13', G2278 in C2
K2150C <sup>-</sup> G (6)         Mild + Inh         REFR         REFR         22%         Surface, β18 <sup>-</sup> , R2307 in C2           P2153Q         3 (6)         Mild         PPPP         PPP         9%         Core, loop 18 <sup>-</sup> , P2310 in C2           R2159H         22 (12)         Mild         PTTT         TTT         35%         Surface, loop 18 <sup>-</sup> , 02316 in C2           R2159H         22 (12)         Mild         RERR         RER         58%         Surface, loop 18 <sup>-</sup> , 02316 in C2           R2159H         22 (12)         Mild         RERR         RER         7%         Core, β19 <sup>-</sup> , R2320 in C2           R2153C <sup>-</sup> 1 (1)         Moderate         RERR         RER         7%         Core, β19 <sup>-</sup> , R2320 in C2           R2164H <sup>+</sup> 2 (2)         Moderate         LLL         LLL         1%         Core, β19 <sup>-</sup> , R2320 in C2           R2164H <sup>+</sup> 2 (2)         Moderate         LLL         LLL         1%         Core, β19 <sup>-</sup> , R2320 in C2           R2164H <sup>+</sup> 2 (2)         Moderate         LLL         LLL         1%         Core, β19 <sup>-</sup> , R2320 in C2           R2164H <sup>+</sup> 2 (2)         Mild         Surface, loop 1         Core, β2         Core, β19 <sup>-</sup> , R2320 in C2           R2173H <sup>+</sup>	N2129S	4 (?)	Moderate	NNNN	NNN	2%	Core, loop 15', N2286 in C2
Krishim         6 (b)         Mild         PPPP         PPP         PPP         9%         Core, loop 18', P2310 in C2           721541         6 (r)         Mild         TTTT         TTT         35%         Surface, loop 18', P2310 in C2           R21590+         22 (12)         Mild         TTTT         TTT         35%         Surface, loop 18', P2310 in C2           R21591         11 (1)         Moderate         RERR         RER         Set         Surface, loop 18', P2310 in C2           R21591         11 (1)         Moderate         RERR         RER         Set         Surface, loop 18', P2310 in C2           R21591         <16 (r)	R2150C*	33 (30)	Mild	RRRR	RRR	22%	Surface, $\beta 18'$ , R2307 in C2
Priskul         3 (b)         Mild         PPPP         PPP         PPP <t< td=""><td>R2150H</td><td>6 (6)</td><td>Mild + Inn )</td><td></td><td></td><td>00/</td><td>0</td></t<>	R2150H	6 (6)	Mild + Inn )			00/	0
$ \begin{array}{c c c c c c } \mbox{Int} 0 (7) & \mbox{Mid} & \mbox{Int} 1 & \mbox{Int} 1 & \mbox{Int} 2 & \mbox{Int} 0 & \mbox{Int} 0 & \mbox{Int} 1 & \mbox{Int} 0 & $	P2153Q	3 (6)	Mild	PPPP	PPP	9%	Core, 100p 18', P2310 In C2
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	121341	б(?) Г. 20 (С. <b>404</b> )		1.1.1.1.1.	1.1.1.	35%	Surface, loop 18, Q2311 In C2
R2159122 (12)MildRackRackSurface, 000 Fe, 023 is In 0.2R2159116 (15)MildModerateRERRRRR7%Core, $\beta 19', R2320$ in 0.2R21534<1-6 (?)	R2159C	5, 20 (6, 104)	Mild			500/	
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	R2159H	ZZ (1Z)	Mild	RKKK	RRK	58%	Surface, 100p 18", Q2316 In C2
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	R2139L	10 (15)	Madarata				
$ \begin{array}{c c} (1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,$	R2103C	I (I) <1-6 (2)	Mod/Sev + Inh	RRRR	RRR	7%	Core, β19', R2320 in C2
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	M216/P*	2 (2)	Moderate )				
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	M2164V	28 (2)	Mild	MMMM	LLL	2%	Core, β19', M2321 in C2
First and the second	L2166S	<1 (?)	Severe	L.L.L.L	T.T.T.	1%	Core 619' V2323 in C2
S21731'       16 (41)       Mild       SSSS       GGG       100%       N-terminus C2; connects with C1         S21731'       16 (41)       Mild       SSSS       GGG       100%       N-terminus C2; connects with C1         E2181D       36 (7)       Mild       IIII       III       SSSS       Core, loop 1         12185T       23 (?)       Mild       IIII       III       0%       Core, lo2         A2192P       1 (?)       Moderate       AAAA       AAA       0%       Core, lo2         A2201P'       7 (36)       Mild       AAAA       GGG       36%       Surface, lo4; ? membrane binding         R2209G       <1 (?)	221000	~ (.)	000010	02.4		170	0010, p10 , v2020 11 02
S21731*f6 (41)MildSSSSGGG100%N-terminus C2; connects with C1E2181D36 (?)Mild + InhEQEEEEE53%Surface, loop 1I2185T23 (?)MiddIIIIIIIIIIII0%Core, b212190N2 (?)ModerateAAAAAAA0%Core, b2A2192P1 (?)ModerateAAAAAAA0%Core, b2A2201P*7 (36)MildAAAAAAA0%Core, b2R2209Q<1 (?)				C2 doma	an		
E2181D36 (?)Mild + InhEQEEEEE53%Surface, loop 1I2185T23 (?)MildIIIIIIIIS%Core, loop 1I2190N2 (?)ModerateIIIIIIII0%Core, $\beta 2$ A2192P1 (?)ModerateAAAAAAAGGG36%Surface, $\beta 4$ ; ? membrane bindingR2209G<1 (?)	S2173I*	16 ( <b>41</b> )	Mild	SSSS	GGG	100%	N-terminus C2; connects with C1
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	E2181D	36 (?)	Mild + Inh	EQEE	EEE	53%	Surface, loop 1
$ \begin{array}{ c c c c c c c } 111 & 111 & 0\% & Core, \beta2 \\ A2192P & 1 (?) & Moderate & AAAA & AAA & AAA & 0\% & Core, \beta2 \\ A2201P' & 7 (36) & Mild & AAAA & GG & 36\% & Surface, \beta4; ? membrane binding \\ R2209G & <1 (?) & Severe \\ R2209Q & <1-7 (4-130) & Mild + lnh & RRRR & RRR & 1% & Core, loop 4 \\ \hline \\ V2223M & ? (?) & - & VVV & AAA & 67\% & Surface, loop 5; ? membrane binding \\ W2229C & 3-10 (?) & Mild/Mod + lnh & WWW & WWW & 25\% & Surface, loop 5; ? membrane binding \\ W223Q & $ (?) & Mild/Mod + lnh & WWW & WWW & 25\% & Surface, loop 5; ? membrane binding \\ W223Q & $ (?) & Mild/Mod + lnh & WWW & WWW & 25\% & Surface, loop 5; ? membrane binding \\ W223Q & $ (?) & Mild/Mod + lnh & WWW & WWW & 25\% & Surface, loop 5; ? membrane binding \\ W223Q & $ (?) & Mild/Mod + lnh & WWW & WWW & 25\% & Surface, loop 5; ? membrane binding \\ W223Q & $ (?) & Mod/Sev & MWVM & KKK & 11\% & Surface, loop 6; \\ W223Q & $ (?) & Mod/Sev & MWVM & KKK & 11\% & Surface, loop 6; \\ W223Q & $ (10) & Mild & TTT & TT & 1\% & Core, loop 6; \\ F2260I & 3 (?) & Mod + lnh & FFFF & YYY & 0\% & Core, loop 15 \\ F2260C & $ (?) & Severe & IIVI & III & 0\% & Core, loop 15 \\ F2260C & $ (?) & Severe & IIVI & IIII & 0\% & Core, loop 15 \\ F2260C & $ (?) & Severe & IIVI & III & 0\% & Core, loop 15 \\ F2260C & $ (?) & Severe & IIVI & III & 0\% & Core, loop 15 \\ F2260C & $ (?) & Severe & IIVI & III & 0\% & Core, loop 15 \\ F2260C & $ (?) & Severe & IIVI & III & 0\% & Core, loop 15 \\ F2260C & $ (?) & Severe & IIVI & III & 0\% & Core, loop 16 \\ F2260C & $ (?) & Mild & PPPP & PPP & 37\% & Surface, loop 16 \\ F2200C & $ (?) & Mild & Inh & PPPP & PPP & 37\% & Surface, loop 16 \\ F2200C & $ (3) & Moderate & RRRR & RRR & RRR & 19\% & Surface, loop 16 \\ F2200C & $ (3) & Moderate & RRRR & RRR & RRR & 19\% & Surface, loop 16 \\ F2200C & $ (20) & Moderate & RRRR & RRR & RRR & 19\% & Core, l18 \\ F220C & $ (20) & Moderate & RRRR & RRR & RRR & 19\% & Core, l19; salt bridge to D2233 \\ F2200C & $ (20) & Moderate & RRRR & RRR & 3\% & Core, l19; mear Q2246 \\ F2222S & $ (?) & Severe & GGGG & GGG & 0\% & Core$	I2185T	23 (?)	Mild	IIII	III	5%	Core, loop 1
A2192P1 (?)ModerateAAAAAAAAAAO%Core, $\beta 2$ A2201P*7 (36)MildAAAAGGG36%Suface, $\beta 4$ ; ? membrane bindingR2209G<1 (?)	I2190N	2 (?)	Moderate	IIII	III	0%	Core, β2
A2201P*7 (36)MildAAAAGGG36%Surface, $\beta4$ ; ? membrane bindingR2209G<1 (?)	A2192P	1 (?)	Moderate	AAAA	AAA	0%	Core, β2
R2209G<1 (?)SevereR2209L3 (3)ModerateRRRRRRRRRCore, loop 4R2209Q<1-7 (4-130)	A2201P*	7 (36)	Mild	AAAA	GGG	36%	Surface, β4; ? membrane binding
R2209L3 (3)ModerateRRRRRRRRR1%Core, loop 4R2209Q<1-7 (4-130)	R2209G	<1 (?)	Severe				
R2209Q<1-7 (4-130)Mild + InhMild + InhVUVAAAA67%Surface, loop 5; ? membrane bindingV2229C3-10 (?)Mild/Mod + InhWWWWWW25%Surface, loop 5; ? membrane bindingV2232A*9 (8)Mild + InhWVVVIVI1%Core, $\beta 6$ W2238V<2 (?)	R2209L	3 (3)	Moderate	RRRR	RRR	1%	Core, loop 4
V2223M? (?)-VVVAAAA67%Surface, loop 5; ? membrane bindingW229C3-10 (?)Mild/Mol + InhWWWWWW25%Surface, g6V232A*9 (8)Mild + InhVVVVIVI1%Core, g6M2238V<2 (?)	R2209Q	<1-7 (4- <b>130</b> )	Mild + Inh				
W2229C3-10 (?)Mild/Mod + InhWWWWWW25%Surface, $\beta \beta$ V2232A*9 (8)Mild + InhVVVVIVI1%Core, $\beta \beta$ M2238V<2 (?)	V2223M	? (?)	—	AVVV	AAA	67%	Surface, loop 5; ? membrane binding
V2232A*9 (8)Mild + InhVVVVIVI1%Core, $\beta 6$ M2238V<2 (?)	W2229C	3-10 (?)	Mild/Mod + Inh	WWWW	WWW	25%	Surface, β6
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	V2232A*	9 (8)	Mild + Inh	VVVV	IVI	1%	Core, β6
12245A/ (10)MildTTTTTT1%Core, $\beta 8$ Q2246R4 (1)ModerateQQQQQQQ3%Core, $\beta 8$ F226013 (?)Mod + InhFFFFYYY0%Core, $\beta 11$ F22602<2 (?)	M2238V	<2 (?)	Mod/Sev	MM <b>V</b> M	KKK	11%	Surface, β7
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	T2245A	7 (10)	Mild	TTTT	TTT	1%	Core, β8
F226013 (?)Mod + Inn Mod + Inn F260CFFFFYYY0%Core, $\beta 11$ F2260C<2 (?)	Q2246R	4 (1)	Moderate	QQQQ	QQQ	3%	Core, β8
P2260C $< < 2 (r)$ SevereIIVIIII0%Core, $\beta 11$ 12262T? (?)SevereIIVIIII0%Core, $\beta 11$ G2285V14 (16)MildGGGGGGG0%Core, loop 15P2300S16 (?)MildPPPPPPP37%Surface, loop 16P2300L*18 (18)Mild + InhPPPPPPP37%Surface, $\beta 18$ R2304C*2 (3)Mod + InhRRRRRRRRRR41%Surface, $\beta 18$ R2304G*8 (3)ModerateRRRRRRR19%Surface, $\beta 18$ ; salt bridge to D2233R2307Q*10 (10)MildRRRRRRRRRR19%Surface, $\beta 18$ ; salt bridge to D2233R2307L<1-2 (4)	F22601	3 (?)		FFFF	YYY	0%	Core, β11
$ \begin{array}{c ccccc} 1 & 1 & 1 & 1 & 0\% & Cole, \beta \\ \hline \end{tabular} 1 & 1 & 1 & 0\% & Cole, \beta \\ \hline \end{tabular} 2 & 14 & (16) & Mild & Gegg & Geg & 0\% & Core, loop 15 \\ \hline \end{tabular} 2 & 16 & (?) & Mild & PPPP & PPP & 37\% & Surface, loop 16 \\ \hline \end{tabular} 2 & 2 & (3) & Mod + Inh & RRR & RRR & RRR & 41\% & Surface, \beta 18 \\ \hline \end{tabular} 2 & 2 & (3) & Mod erate & RRR & RRR & RRR & 41\% & Surface, \beta 18 \\ \hline \end{tabular} 2 & 3Moderate & RRR & RRR & RRR & 19\% & Surface, \beta 18 \\ \hline \end{tabular} 2 & 3Moderate & RRR & RRR & RRR & 19\% & Surface, \beta 18; salt bridge to D2233 \\ \hline \end{tabular} 2 & 3Moderate & RRR & RRR & RRR & 19\% & Surface, \beta 18; salt bridge to D2233 \\ \hline \end{tabular} 2 & 3Moderate & RRR & RRR & RRR & 3\% & Core, \beta 19; near Q2246 \\ \hline \end{tabular} 2 & 3Severe & Gggg & Ggg & 0\% & Core, \beta 19 \\ \hline \end{tabular}$	F2260C	<2 (?)	Severe j	T T		09/	Coro 011
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	122021	? (?) 14 (16)	Severe	1101	111	0%	Core, BTT
P23001*     16 (1)     Mild     PPP     PPP     37%     Surface, loop 16       P2300L*     18 (18)     Mild + Inh     PPPP     PPP     37%     Surface, loop 16       R2304C*     2 (3)     Mod + Inh     RRR     RRR     A1%     Surface, β18       R2304G*     8 (3)     Moderate     RRRR     RRR     10 (?)     Surface, β18       R2307L     10 (10)     Mild     RRRR     RRR     19%     Surface, β18; salt bridge to D2233       R2307L     <1-2 (4)	G2203V	14 (10)	Mild )	GGGG	GGG	0%	Core, 100p 15
R2304C*     2 (3)     Mod + Inh       R2304G*     8 (3)     Moderate       R2304H     10 (?)     Mild       R2307Q*     10 (10)     Mild       R2307L     <1-2 (4)	P23003	10 ( ? )	Mild + Inh	PPPP	PPP	37%	Surface, loop 16
R2304G*     8 (3)     Moderate     RRR     RRR     41%     Surface, β18       R2304H     10 (?)     Mild           R2307Q*     10 (10)     Mild     RRRR     RRR     19%     Surface, β18; salt bridge to D2233       R2307L     <1-2 (4)	R2304C*	2 (2)	Mod + Inh				
R2304H     10 (?)     Mild       R2307Q*     10 (10)     Mild       R2307L     <1-2 (4)	R2304G*	2 (3) 8 (3)	Moderate	ססקק	999	41%	Surface 618
R2307Q*     10 (10)     Mild       R2307L     <1-2 (4)	R2304H	10 (2)	Mild	MILL	IVIVIV	71/0	Canado, pro
R2307L         <1-2 (4)         Mild-Sev         RRR         RRR         19%         Surface, β18; salt bridge to D2233           R2320T*         5 (20)         Moderate         RRR         RRR         3%         Core, β19; near Q2246           G2325S         ? (?)         Severe         GGGG         GGG         0%         Core, β19	R23070*	10 (10)	Mild )				
R2320T*         5 (20)         Moderate         RRR         RRR         3%         Core, β19; near Q2246           G2325S         ? (?)         Severe         GGGG         GGG         0%         Core, β19	R2307L	<1-2 (4)	Mild-Sev	RRRR	RRR	19%	Surface, $\beta$ 18; salt bridge to D2233
G2325S         ? (?)         Severe         GGGG         GGG         O%         Core, β19	R2320T*	5 (20)	Moderate	RRRR	RRR	3%	Core. β19: near Q2246
······································	G2325S	? (?)	Severe	GGGG	GGG	0%	Core, β19

Mutations are from database<sup>15</sup> and if marked by an asterisk, from Table 2. Factor VIII was determined in patient plasma as clotting activity (VIII:C) or total antigen (Ag) with bold for dysfunctional (excess) antigen; "?" indicates value not reported.<sup>15</sup> **Inh**, high titer inhibitor in at least one patient; β, beta strand (numbered) with loops numbered following strands. Factor VIII, homologous residues in h, human; m, murine; p, porcine; c, canine. Factor V, homologous residues in h, human; m, murine; b, bovine (see Figure 1 for sequence references).

The mutations were grouped according to the positions of the native residues in the protein core, in a putative membraneassociated surface, or at other positions on the protein surface.

#### Missense mutations in the protein core

At 25 of the 43 residues in C1 and C2 that are sites of missense mutations (Table 3), surface exposure is less than 10%, and these are classified as core positions. Of these 25 residues, 12 reported mutations at 11 sites replace the native side chains with bulkier groups, and only 3 of these are associated with mild hemophilic phenotypes. Factor VIII antigen levels were reported for 5 of these 12 mutations, and significant dysfunctional protein was found in 1: G2088S. Nine substitutions at 8 residues incorporate smaller side chains and may disrupt the core structure by altering internal hydrogen bonds or by causing cavities that decrease van der Waals contact stabilization of the protein core. Curiously, substitution of G2026 to E in C1 has a mild effect on protein function, while substitution of the same residue to V is associated with a severe bleeding tendency. Clearly, the effect of mutations in the protein core are generally more disruptive than surface mutations. Moreover, the effect of any one particular substitution is a function of the inherent flexibility and chemistry of the side chain's environment, as well as the backbone and side-chain torsion angles supported by the native and mutant residues.

## Missense mutations at a possible membrane association surface

It is interesting to note that a relatively small percentage of the residues thought to be involved in membrane binding have been identified as mutation sites in hemophilia A patients. In the C2 domain, which is likely to be the primary anchor for membrane binding, only one mutation of a hydrophobic residue on this surface (V2223M) has been found in a hemophilic patient. Details of the phenotype for this mutation were not reported<sup>15</sup>; this same residue is alanine in canine factor VIII and in factor V (Table 3). In the C1 model, a hemophilic mutation site (R2159) is found on the same relative surface of the protein as V2223 and is significantly surface-exposed. Mutation of this residue to C, H, or L is associated with a clinically mild bleeding tendency. If this surface of the C1 domain within factor VIII a is located near the membrane, this site may represent a basic residue that also makes electrostatic contact with phospholipid head groups.

Another mutation predicted to affect the membrane association surface of C2 is A2201P, at the beginning of the  $\beta$ 4 strand. This mutation may alter the orientation of the preceding M-F2200  $\beta$ -hairpin, a central feature of the proposed binding site.<sup>22</sup> The fact that this patient also displayed a mild phenotype, and the general underrepresentation of hemophilic mutations in residues thought to interact with a phospholipid membrane, suggest that factor VIIIa generates binding energy to anionic lipid bilayers through a large number of relatively nonspecific interactions. In addition, unlike packing interactions in protein cores or at protein-protein interfaces, a change in the size or dimension of a single side chain is unlikely to cause an energetically costly steric clash against the lipid bilayer. Supporting evidence for this suggestion comes from the observations that in porcine factor VIII M2199 is I and in murine factor VIII L2252 is F. Furthermore, the residues homologous to M2199-F2200 in human factor VIII are W-W in the 3 factor V sequences reported (Figure 1). The structure of a human factor V C2 fragment also places these residues at a potential membrane binding site.42

#### Missense mutations across the protein surface

The remaining hemophilic missense mutation sites correspond to residues with side chains that are partially to extensively surfaceexposed across regions separate from the putative membranebinding surface. Amino-acid side chains having lower solvent accessibility will usually have significant core interactions. Thus, their substitutions could affect either core stability or a surface binding site. The interface between C1 and C2 remains to be defined and will undoubtedly involve burial of some sites listed here as surface-exposed. In light of these limitations, residues with greater than 10% surface exposure that are sites of mutations are identified in yellow in Figure 4, a van der Waals surface representation of each C domain. In the C1 model, 9 hemophilic mutation sites have 15% to 84% solvent accessibility (Table 3). Similarly, 9 exposed residues in C2 that are also sites of hemophilic mutations exhibit 11% to 100% accessibility (Table 3). The latter include A2201 and V2223 discussed under possible membrane association, above.

Regions of the factor VIII C-domain surfaces that are sites of hemophilic missense mutations, other than the proposed membraneassociated region, are likely to include surfaces that interact with other proteins such as vWF. Interaction surfaces between domains of native factor VIIIa, including A3-C1 and C1-C2 interfaces, might also be represented by clusters of such sites. Mutations of at least some of the residues that constitute protein-protein interaction surfaces should produce defects in protein-protein binding or affect factor VIII stability and function. In particular, disruption of vWF binding should lead to increased degradation and clearance of factor VIII from the circulation. Clustering of surface-exposed hemophilic mutations might be more likely to represent potential areas for these interactions.

In the surface representations of the C1 and C2 domains, 3 regions are particularly apparent as clusters of hemophilic mutation sites (Figure 4, left panels). The first is localized to the C2 domain and is composed of R2307 and W2229. These residues are in immediate contact with one another, where the guanidino group of the R side chain makes a polar contact with the imidazole ring face of the W residue. These residues display approximately 10 nm<sup>2</sup> of exposed surface area. Mutations of R2307 are associated with reduced vWF binding (Table 2). R2304 is found within 1 nm of this cluster. A second cluster of hemophilic mutation sites occurs on C1 at surface-exposed residues: T2154, Q2100, and R2150. The mutation R2150H was associated with reduced vWF binding,<sup>43</sup> as was R2150C (Table 2). Although this suggests that R2150 is involved in vWF binding, it is also possible that the native R side chain is buried in the C1-C2 interface. In the latter scenario, the substitutions at R2150 could have profound effects on the protein structure.

As striking as the surfaces described above are with respect to their exposed hemophilic mutation sites, the opposite sides of the C1 and C2 domains are notable for the relative paucity of such sites (Figure 4, right panels). Across these opposite surfaces, only the mutation E2181D in C2 occurs at a nonstructural, exposed residue that might be expected to interact with other macromolecules. This residue is a glutamine in murine factor VIII. Thus, the sides of C1 and C2 shown in the left panels of Figure 4 are primary candidates for sites of other interactions. These surfaces consist of the exterior of  $\beta$ -strands 6 and 18 in C2 and the corresponding region in C1.

A third cluster of polar surface-exposed residues is found near the N-terminus of the C1 domain. This cluster is composed of S2119, R2116, and Y2105. On the basis of the proximity of these



Figure 3. Ribbon diagrams of the C domains: localization of hemophilic mutations. Sites of hemophilic mutations (Table 3) are shown as spheres at the positions of their  $C_{\alpha}$  atoms in the protein structure in stereo views. Spheres are color-coded according to the reported severity; lighter lavender represents mild to moderately severe, and darker violet represents more severe phenotypes (as reported for at least one affected individual). Core  $\beta$ -strands are green (foreground) and red (opposite side) ribbons with additional strands blue; each pair is oriented as in Figure 2. (A) The C1 model is based on homology modeling against the C2-domain crystal structure as described in Figure 2 and the text. (B) C2 is from the previously described crystal structure.<sup>22</sup>

residues to the amino terminus of the C1 domain, it seems plausible that this surface associates with the A3 domain of factor VIII. An A3-domain model, based on the structure of ceruloplasmin (see website<sup>15</sup>), predicts only 5 amino acid residues between the last  $\beta$ -strand of A3 and C2121 in C1. Further analysis and verification of this and other proposed binding surfaces awaits continued structural and biochemical studies.

#### Inhibitory antibodies

The development of inhibitory antibodies is a serious complication of factor VIII concentrate therapy that occurs in 10% to 30% of patients with severe hemophilia A and, as autoantibodies, in acquired hemophilia. Although uncommon, several cases of high titer inhibitors in mild or moderately severe hemophilia A patients have also occurred.<sup>44</sup> The inhibitor antibodies bind predominantly to epitopes in the A2 and/or C2 domains of factor VIII<sup>11,13,28,44,45</sup> and often interfere with the binding of factor VIII to vWF or to phospholipid.<sup>3,10,45-47</sup> Phage display experiments indicated that significant inhibitor binding required 157 amino acids of C2, including its disulfide bond.<sup>48</sup>

In families with 11 hemophilic C-domain mutations, at least one affected individual developed an inhibitor. Of these mutations, 6 were at surface-exposed residues (Table 3, Figure 4). Thus, a bulky core residue substitution is not a prerequisite to alloimmune inhibitor formation. The substitution of W2229 to C was associated with development of an inhibitor in 7 of 18 affected individuals.<sup>44</sup> W2229 is partially exposed on the surface of the  $\beta$ 6 strand in C2 (Table 3) and, as discussed above, may be part of a binding surface.

The antibodies of patients with less severe phenotypes usually neutralize the activity of native factor VIII as well as the mutant factor VIII that previously circulated at low but readily detectable levels. This converts the patient's hemophilia to a severe bleeding tendency. Over time, the inhibitors may predominantly neutralize native as opposed to the mutant factor VIII. One exception to this pattern of response was an inhibitor in a patient with a R2150H mutation. This inhibitor was first associated with allogeneic but not autologous inhibitory factor VIII antibodies that interfered with factor VIII–vWF binding, and was predicted to reduce factor VIII stability in circulation.<sup>42</sup> In support of this prediction, the patient had comparable levels of factor VIII clotting activity and antigen. Similarly, a patient with a mutation to C2150 (Table 2) had an antigen level that was comparably reduced to his clotting activity.

An analysis of the effect of hemophilia A missense mutations in the factor VIII C domains has been recently appeared<sup>49</sup> and is based on previously published homology models against discoidin domains.<sup>21</sup> These models, while correctly predicting the general fold of the  $\beta$ -sandwich cores of the C domains, exhibit substantial differences from the recently determined crystal structures of the factors V<sup>43</sup> and VIII<sup>22</sup> C2 domains (rmsd approximately 2 Å for backbone atoms in the protein core; rmsd greater than 5 Å in the peripheral loop regions). The published assignment of hemophilic missense sites as buried or exposed, based on the discoidin-based homology models,<sup>49</sup> was compared with the same assignments based on the factor VIII C2-crystal structure (Table 3). There was significant disagreement in the calculated surface area for approximately one third of the residues that constitute hemophilic mutation sites.

Structural definition of the factor VIII C2 domain<sup>22</sup> has allowed identification of residues that contribute to exposed surfaces. By



Figure 4. Van der Waals surface space–filled model of factor VIII C domains. C1 (upper panels) and C2 (lower panels) are viewed in similar orientations to Figure 3A-B, with the left series being the "front" and the right, the "back" surface and where the "back" surface corresponds to a 180° rotation about the vertical axis. The domains are color-coded according to element type—carbon atoms are light green; oxygen, red; and nitrogen, blue—except that side chains of residues associated with hemophilia A point mutations have been colored yellow and labeled where visible. Residues that are sites of other hemophilic mutations that are either not shown or not labeled exhibit less than 10% relative solvent accessibility and are defined as buried in the hydrophobic core (Table 3).

Dr John J. Peutz, St. Louis University, St. Louis, MO; Drs

Frederick R. Rickles and Sidney Stein, Emory University, Atlanta,

GA, for providing samples on a patient with an inhibitor; and

members of the Fred Hutchinson Cancer Research Center's

Structural Biology Program for their support. The coordinates for

the homology model of the factor VIII C1 domain are available for

download at http://www.fhcrc.org/science/basic/labs/stoddard/

coords.html. Requests for structural information should be directed

modeling the homologous C1 domain, one can tentatively localize similar areas, pending future structural studies on C1 and C1-C2. Surface areas suggest potential sites of interactions with phospholipid membranes and vWF. Furthermore, they should include epitopes for common immune responses to factor VIII. By localizing hemophilic mutations and comparing the phenotypes and, where available, properties of the hemophilic proteins, one can identify specific surface areas as candidates for future studies using site-specific mutagenesis. These may indicate potential therapeutic strategies, such as creation of a recombinant factor VIII that is relatively resistant to antibody inhibitors,<sup>50</sup> and new approaches to antithrombotic therapy that target intrinsic factor X activation.

#### References

- Gitschier J, Wood WI, Goralka TM, et al. Characterization of the human factor VIII gene. Nature. 1984;312:326-330.
- Wood WI, Capon DJ, Simonsen CC, et al. Expression of active human factor VIII from recombinant DNA clones. Nature. 1984;312:330-337.
- Saenko EL, Scandella D. The acidic region of the factor VIII light chain and the C2 domain together form the high affinity binding site for von Willebrand factor. J Biol Chem. 1997;272:18007-18014.
- Jacquemin M, d'Oiron R, Lavend'homme R, et al. Point mutations scattered in the factor VIII C1 and C2 domains reduce factor VIII binding to vWF and are associated with mild/moderate hemophilia A [abstract]. Thromb Haemost. 1999; 82(suppl):229.
- van Dieijen G, Tans G, Rosing J, Hemker HC. The role of phospholipid and factor VIIIa in the activation of bovine factor X. J Biol Chem. 1981; 256:3433-3442.
- Foster PA, Fulcher CA, Houghten RA, Zimmerman TS. An immunogenic region within residues Val1670-Glu1684 of the factor VIII light chain induces antibodies which inhibit binding of factor VIII to von Willebrand factor. J Biol Chem. 1988; 263:5230-5234.
- Gilbert GE, Baleja JD. Membrane-binding peptide from the C2 domain of factor VIII forms an amphipathic structure as determined by NMR spectroscopy. Biochemistry. 1995;34:3022-3031.
- Saenko E, Sarafanov A, Greco N, et al. Use of surface plasmon resonance for studies of proteinprotein and protein-phospholipid membrane interactions: application to the binding of factor VIII to von Willebrand factor and to phosphatidylserinecontaining membranes. J Chromatogr A. 1999; 852:59-71.
- Takeshima K, Fujikawa K. The phospholipid binding property of the C2 domain of human factor VIII [abstract]. Thromb Haemost. 1999;82(suppl): 234.
- Shima M, Scandella D, Yoshioka A, et al. A factor VIII neutralizing monoclonal antibody and a human inhibitor alloantibody recognizing epitopes in the C2 domain inhibit factor VIII binding to von Willebrand factor and to phosphatidylserine. Thromb Haemost. 1993;69:240-246.
- Scandella D, Gilbert GE, Shima M, et al. Some factor VIII inhibitor antibodies recognize a common epitope corresponding to C2 domain amino acids 2248 through 2312, which overlap a phospholipid binding site. Blood. 1995;86:1811-1819.
- Sawamoto Y, Prescott R, Zhong D, et al. Dominant C2 domain epitope specificity of inhibitor antibodies elicited by a heat pasteurized product, factor VIII CPS-P, in previously treated hemophilia A patients without inhibitors. Thromb Haemost. 1998;79:62-68.
- 13. Healey JF, Barrow RT, Tamim HM, et al. Residues Glu2181-Val2243 contain a major determinant of

the inhibitory epitope in the C2 domain of human factor VIII. Blood. 1998;92:3701-3709.

Acknowledgments

to B.L.S. (bstoddar@fhcrc.org).

- Nogami K, Shima M, Hosokawa K, et al. Role of factor VIII C2 domain in factor VIII binding to factor Xa. J Biol Chem. 1999:274:31000-31007.
- Kemball-Cook G, Tuddenham EGD, Wacey AI. The factor VIII structure and mutation resource site: HAMSTeRS version 4. Nucleic Acids Res. 1998;26:216-219. Available at http://europium. csc.mrc.ac.uk/. Accessed May 8, 2000.
- McMullen BA, Fujikawa K, Davie EW, Hedner U, Ezban, M. Locations of disulfide bonds and free cysteines in the heavy and light chains of recombinant human factor VIII (antihemophilic factor A). Protein Sci. 1995;4:740-746.
- Kane WH, Davie EW. Blood coagulation factors V and VIII: structural and functional similarities and their relationship to hemorrhagic and thrombotic disorders. Blood. 1988;71:539-555.
- Elder B, Lakich D, Gitschier J. Sequence of the murine factor VIII cDNA. Genomics. 1993;16:374-379.
- Healey JF, Lubin IM, Lollar P. The cDNA and derived amino acid sequence of porcine factor VIII. Blood. 1996;88:4209-4214.
- Cameron C, Notley C, Hoyle S, et al. The canine factor VIII cDNA and 5' flanking sequence. Thromb Haemost. 1998;79:317-322.
- Pellequer J-L, Gale AJ, Griffin JH, Getzoff ED. Homology models of the C domains of blood coagulation factors V and VIII: a proposed membrane binding mode for FV and FVIII C2 domains. Blood Cells Mol Dis. 1998;24:448-461.
- Pratt KP, Shen BW, Takeshima K, Davie EW, Fujikawa K, Stoddard BL. Structure of the factor VIII C2 domain at 1.5 Å resolution. Nature. 1999;402: 439-442.
- Liu ML, Pratt KP, Shen BW, Nakaya S, Stoddard BL, Thompson AR. Modeling hemophilic factor VIII (F.VIII) C1 and C2 domain mutations into the 1.5 Å human C2 domain crystal structure [abstract]. Blood. 1999;94:454a.
- Weinmann AF, Schoof JM, Thompson AR. Clinical correlates among 49 families with hemophilia A and factor VIII gene inversions. Am J Hematol. 1996;51:192-199.
- Liu ML, Thompson AR. Factor VIII gene inversions and a Xba I polymorphism: nonradioactive detection and clinical usage. Haemophilia. 1999; 5:26-31.
- Liu M-L, Murphy MEP, Thompson AR. A domain mutations in 65 hemophilia A families and molecular modeling of dysfunctional factor VIII proteins. Brit J Haematol. 1998;103:1051-1060.
- Thompson AR, Chen SH. Characterization of factor IX defects in hemophilia B patients. Methods Enzymol. 1993;222:143-169.
- 28. Thompson AR, Murphy MEP, Liu ML, et al. Loss of tolerance to exogenous and endogenous factor

VIII in a mild hemophilia A patient with an Arg593 to Cys mutation. Blood. 1997;90:1902-1910.

- Moritz B, Broz C, Turecek PL, Schwarz HP, Lang, H. IMMUNOZYM FVIII:Ag, a novel assay for the determination of FVIII antigen [abstract]. Thromb Haemost. 1997;77:31.
- Vlot AJ, Koppelman SJ, van den Berg MH, Bouma BN, Sixma JJ. The affinity and stoichiometry of binding human factor VIII to von Willebrand factor. Blood. 1995;85:3150-3157.
- Casonato A, Pontara E, Zerbinati P, Zucchetto A, Girolami A. The evaluation of factor VIII binding activity of von Willebrand factor by means of an ELISA method: significance and practical implications. Am J Clin Pathol. 1997;109:347-352.
- Liu M, Nakaya S, Thompson AR. Hemophilic factor VIII (FVIII) C domain mutations and von Willebrand factor (vWF) binding [abstract]. Thromb Haemost. 1999;82(suppl):2.
- Chen SH, Schoof JM, Weinmann AF, Thompson AR. Heteroduplex screening for molecular defects in factor IX genes from haemophilia B families. Br J Haematol. 1995;89:409-412.
- Jenny RJ, Pittman DD, Toole JJ, et al. Complete cDNA and derived amino acid sequence of human factor V. Proc Natl Acad Sci U S A. 1987;84: 4846-4850.
- Yang TL, Cui JS, Rehumtulla A, et al. The structure and function of murine factor V and its inactivation by protein C. Blood. 1998;91:4593-4599.
- Guinto ER, Esmon CT, Mann KG, MacGillivray RT. The complete cDNA sequence of bovine coagulation factor V. J Biol Chem. 1992;267:2971-2978.
- Laskowski RJ, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystaligr. 1993;26:283-290.
- NACCESS [computer program]. London: SJ Hubbard and JM Thornton, Department of Biochemistry and Molecular Biology, University College; 1993.
- Reiner AP, Thompson AR. Screening for nonsense mutations in patients with severe hemophilia A can provide rapid, direct carrier detection. Hum Genet. 1992;89:88-94.
- Stoylova S, Mann KG, Brisson A. Structure of membrane-bound human factor Va. FEBS Lett. 1994;351:330-334.
- Stoylova SS, Lenting PJ, Kemball-Cook G, Holzenburg A. Electron crystallography of human blood coagulation factor VIII bound to phospholipid monolayers. J Biol Chem. 1999;274:36573-36578.
- Macedo-Ribeiro S, Bode W, Huber R, et al. Crystal structures of the membrane-binding C2 domain of human coagulation V. Nature. 1999;402: 434-439.
- 43. Peerlinck K, Jacquemin MG, Arnout J, et al. Antifactor VIII antibody inhibiting allogeneic but not

autologous factor VIII in patients with mild hemophila A. Blood. 1999;93:2267-2273.

- Hay C, Ludlam CA, Colvin BT, et al. Factor VIII inhibitors in mild and moderate-severity haemophilia A: UK Haemophilia Centre Directors Organisation. Thromb Haemost. 1998;79:762-766.
- Shima M, Nakai H, Scandella D, et al. Common inhibitory effects of human anti-C2 domain inhibitor alloantibodies on factor VIII binding to von Willebrand factor. Br J Haematol. 1995;91:714-721.
- 46. Prescott R, Nakai H, Saenko EL, et al. The inhibitor antibody response is more complex in hemo-

philia A patients than in most nonhemophiliacs with factor VIII autoantibodies: Recombinate and Kogenate Study Groups. Blood. 1997;89:3663-3671.

- Fijnvandraatet K, Celie PH, Turenhout EA, et al. A human alloantibody interferes with binding of factor IXa to the factor VIII light chains. Blood. 1998; 91:2347-2352.
- Kuwabara I, Maruyama H, Kamisue S, Shima M, Yoshioka A, Maruyama IN. Mapping of the minimal domain encoding a conformational epitope by lambda phage surface display: factor VIII in-

hibitor antibodies from haemophilia A patients. J Immunol Methods. 1999;224:89-99.

- Gale AJ, Pellequer J-L, Getzoff ED, Griffin JH. Structural basis for hemophilia A caused by mutations in the C domains of blood coagulation factor VIII. Thromb Haemost. 2000;83:78-85.
- Barrow RT, Healey JF, Gailani D, Scandella D, Lollar P. Reduction of the antigenicity of factor VIII toward complex inhibitory antibody plasmas using multiply-substituted hybrid human/porcine factor VIII molecules. Blood. 2000;95:564-568.

### Erratum

In the article by Alcobia et al entitled "Spatial associations of centromeres in the nuclei of hematopoietic cells: evidence for cell-type-specific organizational patterns," which appeared in the March 1, 2000, issue of *Blood* (95:1608-1615), some headings in Table 1 were incorrect. Following is a corrected version of the table.

Table 1.	Frequencies of	f association of	specific cent	romeres in qu	uiescent human	peripheral I	blood cells
----------	----------------	------------------	---------------	---------------	----------------	--------------	-------------

Centromeres	A/B	B/C	A/C	A/B/C	n-a	Centromeres	PBM a	PBG a
1/2/3	0.18	0.16 ± 0.01	0.23	0.01	0.56	1/4	0.38*	0.47*
2/3/4	$0.16\pm0.01$	0.18	$0.20\pm0.04$	0.01	0.61	3/4	0.24	0.24
3/4/6	0.15	0.20	$0.15\pm0.02$	0	0.55	3/6	0.23	0.30*
6/7/8	0.26	$0.29\pm0.06$	0.22	0.05	0.49	4/6	0.21	0.16
7/8/9	$0.29\pm0.06$	0.38	$0.43\pm0.03$	0.08	0.23	7/9	0.36	ND
9/10/11	0.24	0.42	0.35	0.05	0.34	10/11	0.26*	0.24*
10/11/12	0.42	0.37	0.26	0.06	0.29	11/12	0.22*†	0.52*†
12/15/16	0.24	$0.46\pm0.02$	0.20	0.05	0.21	13/21	0.82*	0.70
15/16/17	$0.46\pm0.02$	$0.25\pm0.02$	$0.35\pm0.03$	0.07	0.24	14/22	0.79	ND
17/18/20	0.34	0.41	0.21	0.07	0.22	15/16	0.44	0.34
20/X/Y	0.22	0.13	0.09	0.01	0.60	18/20	0.56*	ND
Y/1/2	0.16	0.19	0.10	0	0.57			
5/19	0.23				0.77			
13/21	0.64				0.36			
14/22	0.69				0.31			
1/4	0.19				0.81			
2/6	0.18				0.82			
7/11	0.30				0.70			

PBL, peripheral blood lymphocytes; PBM, peripheral blood monocytes; PBG, peripheral blood granulocytes; A-B-C, triad of centromeres analyzed. A/B, B/C, A/C, A/B/C, respectively, associations between the first and the second centromere of the triad, the second and the third, the first and the third, and all three. Association between homologous centromeres is not indicated (see text). a, associated; n-a, nonassociated; ND, not done.

\*Comparison with lymphocytes, P < .05.

†Comparison between monocytes and granulocytes, P < .05.