Full wwPDB X-ray Structure Validation Report (i)

May 22, 2020 – 10:47 am BST

PDB ID	:	4KI5
Title	:	Cystal structure of human factor VIII C2 domain in a ternary complex with murine inhbitory antibodies 3E6 and G99
Authors	:	Walter, J.D.; Meeks, S.L.; Healey, J.F.; Lollar, P.; Spiegel, P.C.
Deposited on	:	2013-05-01
Resolution	:	2.47 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

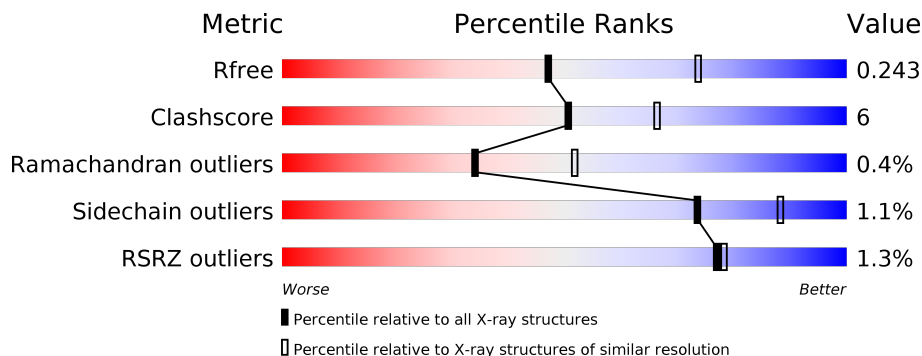|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.47 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.243
Clashscore		6
Ramachandran outliers		0.4%
Sidechain outliers		1.1%
RSRZ outliers		1.3%

Worse   Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	5857 (2.50-2.46)
Clashscore	141614	6594 (2.50-2.46)
Ramachandran outliers	138981	6469 (2.50-2.46)
Sidechain outliers	138945	6471 (2.50-2.46)
RSRZ outliers	127900	5738 (2.50-2.46)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	C	219	90% 9%
2	D	213	% 85% 14% •
3	E	224	88% 12% •
4	F	214	3% 78% 18% ••
5	M	183	% 75% 9% 16%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8074 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called MURINE MONOCLONAL 3E6 FAB HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	218	Total	C	N	O	S	0	0	0
			1647	1045	270	326	6			

- Molecule 2 is a protein called MURINE MONOCLONAL 3E6 FAB LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	211	Total	C	N	O	S	0	0	0
			1620	1009	269	333	9			

- Molecule 3 is a protein called MURINE MONOCLONAL G99 FAB HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	224	Total	C	N	O	S	0	0	0
			1685	1065	272	341	7			

- Molecule 4 is a protein called MURINE MONOCLONAL G99 FAB LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	209	Total	C	N	O	S	0	0	0
			1624	1011	271	335	7			

- Molecule 5 is a protein called Coagulation factor VIII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	M	154	Total	C	N	O	S	0	0	0
			1231	785	214	224	8			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	2150	MET	-	EXPRESSION TAG	UNP P00451
M	2151	GLY	-	EXPRESSION TAG	UNP P00451

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
M	2152	SER	-	EXPRESSION TAG	UNP P00451
M	2153	SER	-	EXPRESSION TAG	UNP P00451
M	2154	HIS	-	EXPRESSION TAG	UNP P00451
M	2155	HIS	-	EXPRESSION TAG	UNP P00451
M	2156	HIS	-	EXPRESSION TAG	UNP P00451
M	2157	HIS	-	EXPRESSION TAG	UNP P00451
M	2158	HIS	-	EXPRESSION TAG	UNP P00451
M	2159	HIS	-	EXPRESSION TAG	UNP P00451
M	2160	SER	-	EXPRESSION TAG	UNP P00451
M	2161	SER	-	EXPRESSION TAG	UNP P00451
M	2162	GLY	-	EXPRESSION TAG	UNP P00451
M	2163	LEU	-	EXPRESSION TAG	UNP P00451
M	2164	VAL	-	EXPRESSION TAG	UNP P00451
M	2165	PRO	-	EXPRESSION TAG	UNP P00451
M	2166	ARG	-	EXPRESSION TAG	UNP P00451
M	2167	GLY	-	EXPRESSION TAG	UNP P00451
M	2168	SER	-	EXPRESSION TAG	UNP P00451
M	2169	HIS	-	EXPRESSION TAG	UNP P00451
M	2170	MET	-	EXPRESSION TAG	UNP P00451

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	76	Total 76 O 76	0	0
6	D	54	Total 54 O 54	0	0
6	E	46	Total 46 O 46	0	0
6	F	40	Total 40 O 40	0	0
6	M	51	Total 51 O 51	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
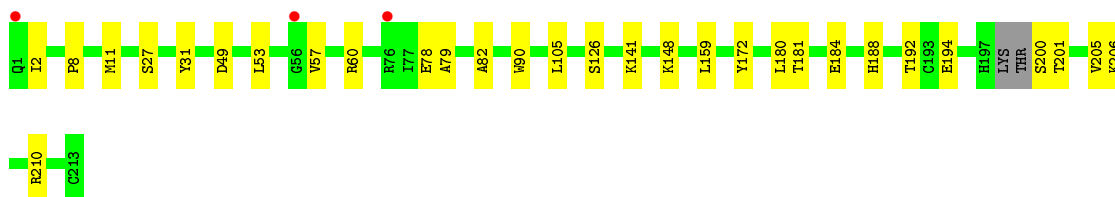
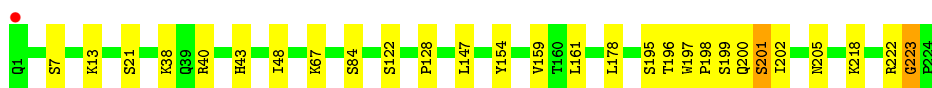- Molecule 1: MURINE MONOCLONAL 3E6 FAB HEAVY CHAIN

Chain C:
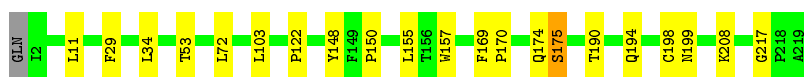
- Molecule 2: MURINE MONOCLONAL 3E6 FAB LIGHT CHAIN

Chain D:

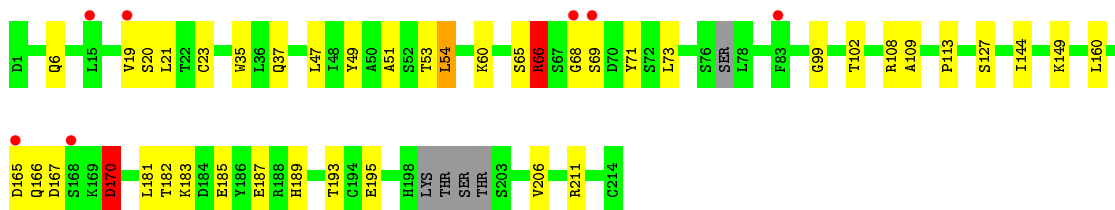- Molecule 3: MURINE MONOCLONAL G99 FAB HEAVY CHAIN

Chain E:

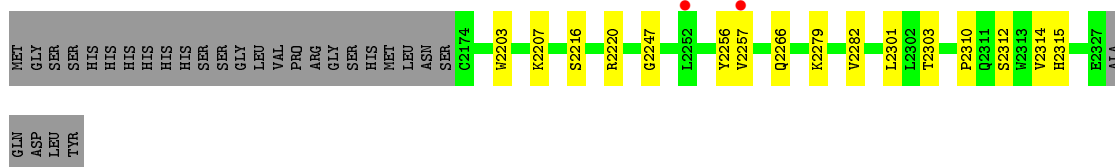- Molecule 4: MURINE MONOCLONAL G99 FAB LIGHT CHAIN

Chain F:

- Molecule 5: Coagulation factor VIII

Chain M:

MET GLY SER HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET LEU ASN SER C2174 W2203 K2207 S2216 R2220 G2247 L2252 Y2256 V2257 Q2266 K2279 V2282 L2301 L2302 T2303 P2310 Q2311 S2312 W2313 V2314 H2315 E2327 ALA

GLN ASP LEU TYR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	54.25Å 71.16Å 277.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	56.43 – 2.47 56.43 – 2.47	Depositor EDS
% Data completeness (in resolution range)	99.8 (56.43-2.47) 99.8 (56.43-2.47)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.80 (at 2.48Å)	Xtriage
Refinement program	REFMAC, PHENIX 1.8_1069	Depositor
R, R_{free}	0.188 , 0.238 0.199 , 0.243	Depositor DCC
R_{free} test set	2000 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.663	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 49.7	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8074	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.36% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	C	0.32	0/1690	0.47	0/2309
2	D	0.30	0/1658	0.47	0/2253
3	E	0.34	0/1731	0.47	0/2365
4	F	0.38	0/1657	0.50	0/2244
5	M	0.29	0/1263	0.45	0/1713
All	All	0.33	0/7999	0.47	0/10884

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1647	0	1611	13	0
2	D	1620	0	1541	19	1
3	E	1685	0	1619	23	0
4	F	1624	0	1551	31	1
5	M	1231	0	1218	9	0
6	C	76	0	0	1	0
6	D	54	0	0	0	0
6	E	46	0	0	0	0
6	F	40	0	0	1	0
6	M	51	0	0	2	0
All	All	8074	0	7540	88	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (88) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:178:LEU:CD2	4:F:160:LEU:HB3	1.97	0.95
3:E:178:LEU:HD21	4:F:160:LEU:HB3	1.49	0.94
1:C:174:GLN:HG3	2:D:159:LEU:CD1	1.98	0.94
1:C:174:GLN:HG3	2:D:159:LEU:HD12	1.49	0.93
4:F:66:ARG:HD2	4:F:71:TYR:CE1	2.13	0.82
1:C:194:GLN:NE2	1:C:194:GLN:HA	1.95	0.79
4:F:108:ARG:NH1	4:F:109:ALA:O	2.18	0.76
3:E:128:PRO:HB3	3:E:154:TYR:HB3	1.71	0.73
3:E:154:TYR:CE1	3:E:159:VAL:HG23	2.25	0.72
4:F:65:SER:O	4:F:66:ARG:HB3	1.92	0.69
4:F:195:GLU:HG2	4:F:206:VAL:HG12	1.74	0.69
2:D:194:GLU:HG2	2:D:205:VAL:HG12	1.75	0.69
2:D:2:ILE:HG12	2:D:27:SER:HB3	1.73	0.69
1:C:174:GLN:HG3	2:D:159:LEU:HD11	1.76	0.67
3:E:13:LYS:HG3	3:E:122:SER:HA	1.77	0.65
3:E:201:SER:OG	3:E:218:LYS:HE3	1.98	0.64
4:F:37:GLN:HB2	4:F:47:LEU:HD11	1.81	0.62
4:F:19:VAL:HG12	4:F:20:SER:N	2.15	0.61
5:M:2257:VAL:HG13	5:M:2312:SER:O	2.01	0.61
4:F:21:LEU:HD22	4:F:102:THR:HG21	1.83	0.60
1:C:208:LYS:NZ	6:C:361:HOH:O	2.36	0.59
2:D:8:PRO:HD2	2:D:11:MET:HE2	1.85	0.58
1:C:53:THR:HA	1:C:72:LEU:HD21	1.85	0.57
2:D:31:TYR:HB3	2:D:49:ASP:HA	1.86	0.56
4:F:165:ASP:OD1	4:F:166:GLN:N	2.36	0.56
2:D:180:LEU:HB3	2:D:184:GLU:HG3	1.87	0.56
3:E:178:LEU:HD23	4:F:160:LEU:HB3	1.87	0.55
3:E:195:SER:O	3:E:199:SER:OG	2.25	0.55
3:E:38:LYS:HB2	3:E:48:ILE:HD11	1.90	0.54
3:E:154:TYR:CE1	3:E:159:VAL:CG2	2.91	0.53
4:F:66:ARG:HD2	4:F:71:TYR:CZ	2.44	0.53
2:D:200:SER:OG	2:D:201:THR:N	2.43	0.51
3:E:147:LEU:HD12	3:E:202:ILE:HD13	1.92	0.51
4:F:19:VAL:CG1	4:F:20:SER:N	2.74	0.51
4:F:66:ARG:O	4:F:66:ARG:HG3	2.09	0.50
2:D:82:ALA:HB2	2:D:105:LEU:HD13	1.93	0.50
3:E:222:ARG:O	3:E:223:GLY:O	2.30	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:M:2282:VAL:O	6:M:2448:HOH:O	2.20	0.49
2:D:60:ARG:CZ	2:D:78:GLU:HG3	2.44	0.47
3:E:178:LEU:HD21	4:F:160:LEU:CB	2.35	0.47
4:F:51:ALA:O	6:F:307:HOH:O	2.20	0.47
4:F:113:PRO:HG3	4:F:144:ILE:HD11	1.96	0.47
3:E:197:TRP:CG	3:E:198:PRO:HA	2.49	0.47
2:D:141:LYS:HG3	2:D:172:TYR:CG	2.49	0.47
1:C:190:THR:O	1:C:194:GLN:HB2	2.15	0.46
4:F:35:TRP:CD2	4:F:73:LEU:HB2	2.51	0.46
4:F:35:TRP:CE2	4:F:73:LEU:HB2	2.51	0.46
3:E:222:ARG:C	3:E:223:GLY:O	2.51	0.46
5:M:2257:VAL:HG21	5:M:2310:PRO:HB3	1.98	0.46
5:M:2314:VAL:O	5:M:2315:HIS:HB2	2.15	0.46
4:F:181:LEU:HB3	4:F:185:GLU:HG3	1.97	0.45
4:F:6:GLN:HG3	4:F:99:GLY:HA3	1.98	0.45
4:F:182:THR:OG1	4:F:185:GLU:HG2	2.17	0.45
3:E:197:TRP:CD1	3:E:198:PRO:HA	2.52	0.45
4:F:54:LEU:HD11	4:F:60:LYS:HA	1.98	0.45
5:M:2203:TRP:CD2	5:M:2220:ARG:HD3	2.52	0.45
1:C:122:PRO:HB3	1:C:148:TYR:HB3	1.99	0.44
5:M:2207:LYS:HD3	5:M:2216:SER:O	2.17	0.44
3:E:67:LYS:HE2	3:E:84:SER:O	2.18	0.44
4:F:183:LYS:HB3	4:F:183:LYS:HE3	1.81	0.44
2:D:181:THR:OG1	2:D:184:GLU:HG2	2.18	0.43
2:D:206:LYS:HA	2:D:206:LYS:HD3	1.83	0.43
3:E:7:SER:HG	3:E:21:SER:HG	1.63	0.43
4:F:6:GLN:HG2	4:F:23:CYS:SG	2.59	0.43
1:C:11:LEU:HD22	1:C:150:PRO:HD3	2.01	0.42
2:D:79:ALA:HA	2:D:105:LEU:HD22	2.00	0.42
4:F:167:ASP:OD2	4:F:170:ASP:HB3	2.19	0.42
3:E:7:SER:OG	3:E:21:SER:OG	2.29	0.42
2:D:148:LYS:HB2	2:D:192:THR:HB	2.02	0.42
5:M:2266:GLN:HE22	5:M:2303:THR:HA	1.84	0.42
4:F:68:GLY:HA2	4:F:69:SER:HA	1.52	0.41
3:E:40:ARG:HB2	3:E:43:HIS:HB2	2.01	0.41
1:C:174:GLN:N	1:C:175:SER:HA	2.34	0.41
3:E:196:THR:O	3:E:200:GLN:HB2	2.21	0.41
3:E:197:TRP:CD1	3:E:202:ILE:HD12	2.55	0.41
3:E:161:LEU:HA	3:E:205:ASN:O	2.21	0.41
5:M:2247:GLY:HA3	5:M:2256:TYR:HB3	2.03	0.41
4:F:183:LYS:O	4:F:187:GLU:HG3	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:159:LEU:HA	2:D:159:LEU:HD23	1.91	0.41
2:D:53:LEU:HG	2:D:57:VAL:HB	2.02	0.41
4:F:189:HIS:O	4:F:211:ARG:HD3	2.21	0.41
4:F:49:TYR:CZ	4:F:53:THR:HG21	2.56	0.41
4:F:149:LYS:HB2	4:F:193:THR:HB	2.04	0.40
2:D:188:HIS:O	2:D:210:ARG:HD3	2.21	0.40
5:M:2279:LYS:NZ	6:M:2445:HOH:O	2.54	0.40
1:C:155:LEU:HA	1:C:199:ASN:O	2.21	0.40
1:C:169:PHE:HA	1:C:170:PRO:HD3	1.93	0.40
1:C:157:TRP:CZ3	1:C:198:CYS:HB3	2.56	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:126:SER:OG	4:F:127:SER:O[4_535]	2.01	0.19

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	216/219 (99%)	207 (96%)	8 (4%)	1 (0%)	29	46
2	D	207/213 (97%)	201 (97%)	6 (3%)	0	100	100
3	E	222/224 (99%)	216 (97%)	5 (2%)	1 (0%)	29	46
4	F	203/214 (95%)	194 (96%)	7 (3%)	2 (1%)	15	26
5	M	152/183 (83%)	145 (95%)	7 (5%)	0	100	100
All	All	1000/1053 (95%)	963 (96%)	33 (3%)	4 (0%)	34	52

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	F	170	ASP
4	F	66	ARG
1	C	217	GLY
3	E	223	GLY

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	185/186 (100%)	181 (98%)	4 (2%)	52	75
2	D	186/188 (99%)	185 (100%)	1 (0%)	88	95
3	E	191/191 (100%)	190 (100%)	1 (0%)	88	95
4	F	186/191 (97%)	183 (98%)	3 (2%)	62	82
5	M	139/164 (85%)	138 (99%)	1 (1%)	84	93
All	All	887/920 (96%)	877 (99%)	10 (1%)	73	88

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	29	PHE
1	C	34	LEU
1	C	103	LEU
1	C	175	SER
2	D	90	TRP
3	E	201	SER
4	F	54	LEU
4	F	66	ARG
4	F	170	ASP
5	M	2301	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	194	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	C	218/219 (99%)	-0.08	0	100	100	36, 48, 63, 72	0
2	D	211/213 (99%)	-0.09	3 (1%)	75	77	34, 47, 62, 74	0
3	E	224/224 (100%)	-0.02	1 (0%)	92	93	39, 52, 66, 77	0
4	F	209/214 (97%)	0.09	7 (3%)	46	49	35, 49, 68, 77	0
5	M	154/183 (84%)	0.12	2 (1%)	77	78	42, 54, 72, 81	0
All	All	1016/1053 (96%)	-0.00	13 (1%)	77	78	34, 50, 66, 81	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	M	2252	LEU	5.4
4	F	19	VAL	3.4
2	D	1	GLN	3.2
4	F	68	GLY	2.9
3	E	1	GLN	2.8
4	F	168	SER	2.6
2	D	76	ARG	2.5
4	F	69	SER	2.4
2	D	56	GLY	2.3
4	F	165	ASP	2.3
5	M	2257	VAL	2.1
4	F	83	PHE	2.0
4	F	15	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.