



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 07:32 PM GMT

PDB ID : 4P3W
Title : Crystal structure of the human filamin A Ig-like domains 20-21 in complex with migfilin peptide
Authors : Seppala, J.; Pentikainen, U.; Ylanne, J.
Deposited on : 2014-03-10
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026982
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20026982

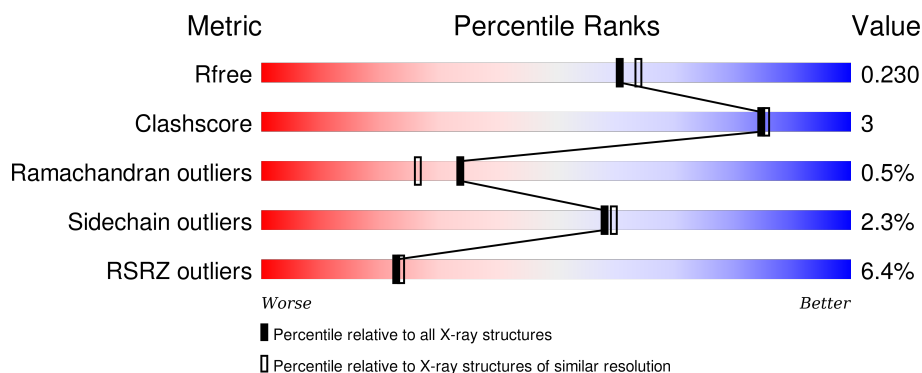
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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. 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 $\leq 5\%$. 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	A	182	<div> <div>4%</div> <div>84% 10% • 5%</div> </div>
1	B	182	<div> <div>4%</div> <div>82% 12% • 6%</div> </div>
1	C	182	<div> <div>7%</div> <div>81% 14% 5%</div> </div>
1	D	182	<div> <div>11%</div> <div>85% • 11%</div> </div>
1	E	182	<div> <div>2%</div> <div>82% 13% 5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	182	<div><div><div></div><div></div><div></div></div><div><div>7%</div><div>79%</div><div>7%</div><div>•</div><div>14%</div></div></div>
2	G	24	<div><div><div></div><div></div><div></div></div><div><div>4%</div><div>58%</div><div>•</div><div>38%</div></div></div>
2	H	24	<div><div><div></div><div></div><div></div></div><div><div>4%</div><div>58%</div><div>•</div><div>38%</div></div></div>
2	I	24	<div><div><div></div><div></div><div></div></div><div><div>4%</div><div>50%</div><div>8%</div><div>42%</div></div></div>
2	J	24	<div><div><div></div><div></div><div></div></div><div><div>54%</div><div>46%</div></div></div>
2	K	24	<div><div><div></div><div></div><div></div></div><div><div>4%</div><div>54%</div><div>•</div><div>•</div><div>38%</div></div></div>
2	L	24	<div><div><div></div><div></div><div></div></div><div><div>4%</div><div>46%</div><div>•</div><div>50%</div></div></div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8100 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 Filamin-A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	172	Total	C	N	O	S	0	0	0
			1269	796	214	254	5			
1	B	171	Total	C	N	O	S	0	0	0
			1264	794	215	250	5			
1	E	172	Total	C	N	O	S	0	0	0
			1259	792	212	250	5			
1	F	157	Total	C	N	O	S	0	0	0
			1120	703	192	220	5			
1	D	162	Total	C	N	O	S	0	0	0
			1153	724	194	230	5			
1	C	172	Total	C	N	O	S	0	0	0
			1262	794	213	250	5			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2148	GLY	-	expression tag	UNP P21333
A	2149	ALA	-	expression tag	UNP P21333
A	2150	MET	-	expression tag	UNP P21333
A	2151	GLY	-	expression tag	UNP P21333
B	2148	GLY	-	expression tag	UNP P21333
B	2149	ALA	-	expression tag	UNP P21333
B	2150	MET	-	expression tag	UNP P21333
B	2151	GLY	-	expression tag	UNP P21333
E	2148	GLY	-	expression tag	UNP P21333
E	2149	ALA	-	expression tag	UNP P21333
E	2150	MET	-	expression tag	UNP P21333
E	2151	GLY	-	expression tag	UNP P21333
F	2148	GLY	-	expression tag	UNP P21333
F	2149	ALA	-	expression tag	UNP P21333
F	2150	MET	-	expression tag	UNP P21333
F	2151	GLY	-	expression tag	UNP P21333
D	2148	GLY	-	expression tag	UNP P21333

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	2149	ALA	-	expression tag	UNP P21333
D	2150	MET	-	expression tag	UNP P21333
D	2151	GLY	-	expression tag	UNP P21333
C	2148	GLY	-	expression tag	UNP P21333
C	2149	ALA	-	expression tag	UNP P21333
C	2150	MET	-	expression tag	UNP P21333
C	2151	GLY	-	expression tag	UNP P21333

- Molecule 2 is a protein called Filamin-binding LIM protein 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	15	Total	C	N	O	0	0	0
			113	74	19	20			
2	H	15	Total	C	N	O	0	0	0
			113	74	19	20			
2	K	15	Total	C	N	O	0	0	0
			113	74	19	20			
2	L	12	Total	C	N	O	0	0	0
			90	59	16	15			
2	J	13	Total	C	N	O	0	0	0
			97	64	17	16			
2	I	14	Total	C	N	O	0	0	0
			104	69	18	17			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	E	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0

- Molecule 4 is PRASEODYMIUM ION (three-letter code: PR) (formula: Pr).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Pr 1 1	0	0
4	A	1	Total Pr 1 1	0	0
4	D	1	Total Pr 1 1	0	0
4	E	1	Total Pr 1 1	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total Na 1 1	0	0
5	E	1	Total Na 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	31	Total O 31 31	0	0
6	B	32	Total O 32 32	0	0
6	H	4	Total O 4 4	0	0
6	E	20	Total O 20 20	0	0
6	F	5	Total O 5 5	0	0

Continued on next page...

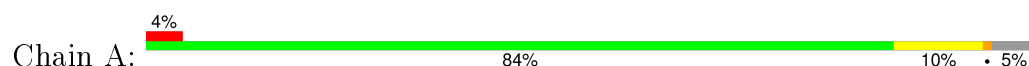
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	7	Total 7	O 7	0	0
6	C	17	Total 17	O 17	0	0
6	I	1	Total 1	O 1	0	0

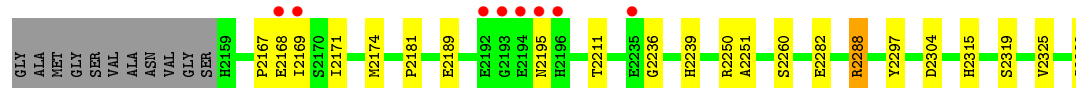
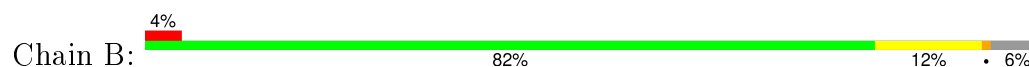
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

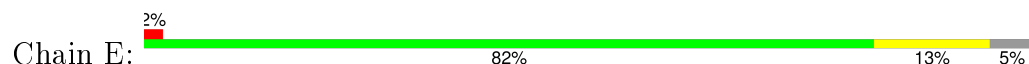
• Molecule 1: Filamin-A



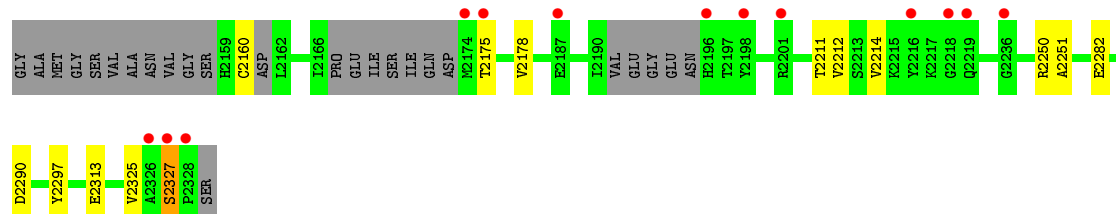
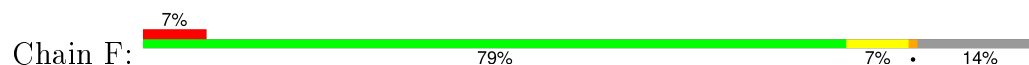
• Molecule 1: Filamin-A



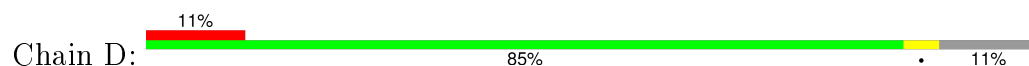
• Molecule 1: Filamin-A

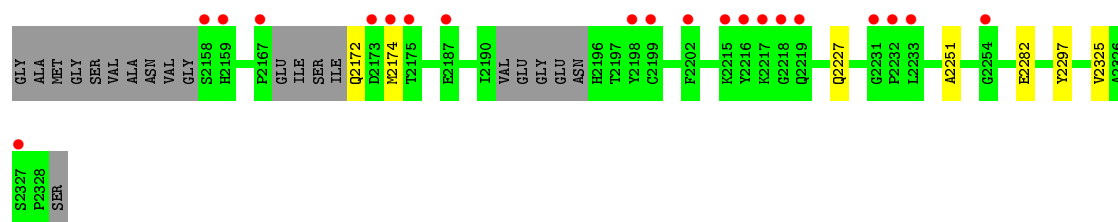


• Molecule 1: Filamin-A

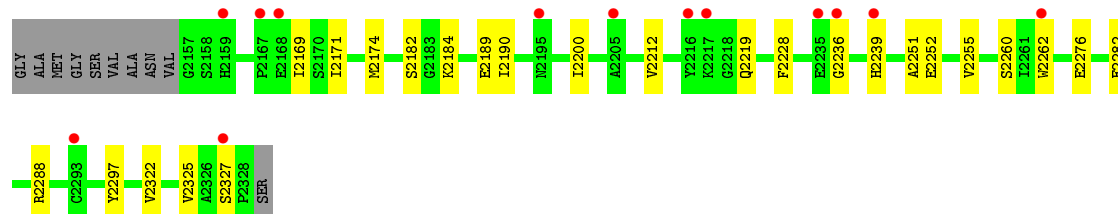
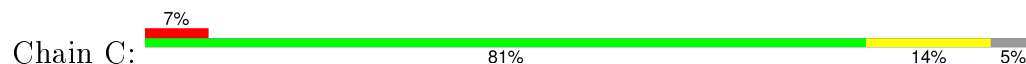


• Molecule 1: Filamin-A

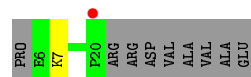




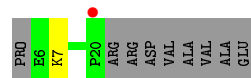
• Molecule 1: Filamin-A



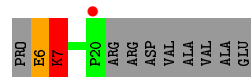
• Molecule 2: Filamin-binding LIM protein 1



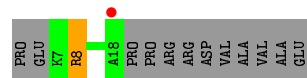
• Molecule 2: Filamin-binding LIM protein 1



• Molecule 2: Filamin-binding LIM protein 1



• Molecule 2: Filamin-binding LIM protein 1

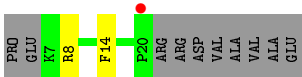


• Molecule 2: Filamin-binding LIM protein 1





● Molecule 2: Filamin-binding LIM protein 1



4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	88.15Å 88.15Å 394.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.86 – 2.00 43.86 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (43.86-2.00) 100.0 (43.86-2.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.36 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.200 , 0.226 0.211 , 0.230	Depositor DCC
R_{free} test set	3874 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	40.6	Xtriage
Anisotropy	0.360	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 45.1	EDS
Estimated twinning fraction	0.296 for -h-k,k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.37$	Xtriage
Outliers	0 of 77265 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8100	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.37% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PR, NA, SO4

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	A	1.09	1/1301 (0.1%)	1.03	3/1770 (0.2%)
1	B	1.11	1/1297 (0.1%)	1.05	6/1765 (0.3%)
1	C	0.92	2/1295 (0.2%)	0.95	2/1762 (0.1%)
1	D	0.85	0/1181	0.88	0/1609
1	E	1.10	4/1292 (0.3%)	0.99	1/1759 (0.1%)
1	F	1.00	1/1146 (0.1%)	0.99	2/1558 (0.1%)
2	G	1.03	0/115	1.05	0/156
2	H	1.13	0/115	1.18	0/156
2	I	0.91	0/106	1.02	1/144 (0.7%)
2	J	0.97	0/98	1.17	0/132
2	K	1.14	0/115	1.18	1/156 (0.6%)
2	L	0.88	0/90	1.14	0/120
All	All	1.02	9/8151 (0.1%)	0.99	16/11087 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	F	0	1
2	K	0	1
All	All	0	2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	2189	GLU	CD-OE2	-7.92	1.17	1.25
1	E	2306	GLU	CD-OE1	7.77	1.34	1.25
1	E	2189	GLU	CD-OE2	-7.46	1.17	1.25
1	B	2189	GLU	CD-OE2	-6.49	1.18	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	2276	GLU	CD-OE1	-6.00	1.19	1.25
1	E	2306	GLU	CG-CD	5.62	1.60	1.51
1	E	2260	SER	CB-OG	-5.51	1.35	1.42
1	A	2180	SER	CB-OG	-5.15	1.35	1.42
1	F	2313	GLU	CD-OE1	-5.12	1.20	1.25

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	2288	ARG	NE-CZ-NH2	-7.98	116.31	120.30
1	B	2288	ARG	NE-CZ-NH1	7.28	123.94	120.30
1	A	2288	ARG	NE-CZ-NH2	-6.43	117.08	120.30
1	B	2250	ARG	NE-CZ-NH1	6.15	123.37	120.30
1	C	2288	ARG	NE-CZ-NH1	5.74	123.17	120.30
1	A	2290	ASP	CB-CG-OD1	5.64	123.38	118.30
1	E	2288	ARG	NE-CZ-NH2	-5.62	117.49	120.30
1	A	2236	GLY	N-CA-C	-5.54	99.26	113.10
1	C	2288	ARG	NE-CZ-NH2	-5.52	117.54	120.30
1	B	2304	ASP	CB-CG-OD1	5.38	123.14	118.30
1	F	2250	ARG	NE-CZ-NH1	5.17	122.89	120.30
1	F	2290	ASP	CB-CG-OD2	-5.12	113.69	118.30
1	B	2167	PRO	N-CA-C	5.10	125.37	112.10
1	B	2304	ASP	CB-CG-OD2	-5.09	113.72	118.30
2	I	8	ARG	NE-CZ-NH1	5.04	122.82	120.30
2	K	7	LYS	CA-CB-CG	5.04	124.48	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	2327	SER	Peptide
2	K	6	GLU	Peptide

5.2 Too-close contacts

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	A	1269	0	1205	10	0
1	B	1264	0	1201	8	0
1	C	1262	0	1198	11	0
1	D	1153	0	1064	4	0
1	E	1259	0	1189	10	0
1	F	1120	0	1035	6	0
2	G	113	0	121	0	0
2	H	113	0	121	0	0
2	I	104	0	115	1	0
2	J	97	0	108	0	0
2	K	113	0	121	1	0
2	L	90	0	101	1	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
3	C	5	0	0	0	0
3	E	5	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
5	C	1	0	0	1	0
5	E	1	0	0	0	0
6	A	31	0	0	0	0
6	B	32	0	0	0	0
6	C	17	0	0	1	0
6	D	7	0	0	0	0
6	E	20	0	0	0	0
6	F	5	0	0	0	0
6	H	4	0	0	0	0
6	I	1	0	0	0	0
All	All	8100	0	7579	45	0

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 3.

All (45) 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)
1:A:2235:GLU:HG2	1:A:2235:GLU:O	1.46	1.10
5:C:2402:NA:NA	6:C:2502:HOH:O	1.54	0.79
1:F:2178:VAL:HG22	1:F:2212:VAL:HG22	1.67	0.75
1:A:2235:GLU:O	1:A:2235:GLU:CG	2.30	0.73

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:2182:SER:OG	1:C:2184:LYS:HD3	1.95	0.66
1:E:2220:HIS:HD2	1:E:2224:SER:OG	1.79	0.65
1:B:2315:HIS:HD2	1:B:2319:SER:OG	1.83	0.62
1:A:2220:HIS:HD2	1:A:2224:SER:OG	1.84	0.60
1:E:2182:SER:OG	1:E:2184:LYS:HG2	2.02	0.59
1:C:2252:GLU:HB2	1:C:2255:VAL:HG22	1.83	0.59
1:C:2212:VAL:CG2	1:C:2228:PHE:HE2	2.17	0.56
1:E:2315:HIS:HB2	2:L:8:ARG:HG2	1.88	0.55
1:A:2181:PRO:O	1:B:2211:THR:HG21	2.07	0.55
1:E:2236:GLY:O	1:E:2239:HIS:CD2	2.60	0.54
1:C:2236:GLY:O	1:C:2239:HIS:CD2	2.61	0.54
1:B:2236:GLY:O	1:B:2239:HIS:CD2	2.61	0.53
1:F:2212:VAL:HG12	1:F:2214:VAL:HG23	1.94	0.50
1:A:2211:THR:HG21	1:B:2181:PRO:O	2.12	0.50
2:K:6:GLU:HA	2:K:6:GLU:OE1	2.13	0.48
1:F:2282:GLU:O	1:F:2297:TYR:HA	2.14	0.48
1:E:2212:VAL:CG2	1:E:2228:PHE:HE1	2.28	0.46
1:E:2159:HIS:CD2	1:F:2282:GLU:HG2	2.51	0.46
1:E:2169:ILE:HD12	1:E:2174:MET:SD	2.56	0.46
1:C:2282:GLU:O	1:C:2297:TYR:HA	2.15	0.46
1:A:2251:ALA:HB3	1:A:2325:VAL:HG22	1.98	0.46
1:D:2282:GLU:O	1:D:2297:TYR:HA	2.16	0.45
1:C:2169:ILE:HD12	1:C:2174:MET:SD	2.56	0.45
1:B:2251:ALA:HB3	1:B:2325:VAL:HG22	1.97	0.45
1:C:2190:ILE:HG23	1:C:2200:ILE:HD13	1.99	0.44
1:B:2169:ILE:HD12	1:B:2174:MET:SD	2.58	0.44
1:C:2252:GLU:HB2	1:C:2255:VAL:CG2	2.48	0.44
1:E:2282:GLU:O	1:E:2297:TYR:HA	2.18	0.43
1:E:2251:ALA:HB3	1:E:2325:VAL:HG22	1.99	0.43
1:A:2290:ASP:OD2	1:C:2260:SER:OG	2.36	0.43
1:E:2181:PRO:O	1:F:2211:THR:HG21	2.18	0.42
1:C:2251:ALA:HB3	1:C:2325:VAL:HG22	2.00	0.42
1:D:2172:GLN:HB3	1:D:2174:MET:HG2	2.01	0.42
1:B:2260:SER:HB3	1:B:2288:ARG:NH2	2.35	0.42
1:A:2262:TRP:CH2	1:C:2262:TRP:HB2	2.55	0.42
1:A:2220:HIS:HD2	1:A:2224:SER:HG	1.65	0.42
1:D:2251:ALA:HB3	1:D:2325:VAL:HG22	2.01	0.41
1:F:2251:ALA:HB3	1:F:2325:VAL:HG22	2.01	0.41
1:A:2282:GLU:O	1:A:2297:TYR:HA	2.20	0.41
1:B:2282:GLU:O	1:B:2297:TYR:HA	2.21	0.40
1:D:2227:GLN:HB2	2:I:14:PHE:HB3	2.04	0.40

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	A	170/182 (93%)	164 (96%)	4 (2%)	2 (1%)	16	8
1	B	169/182 (93%)	166 (98%)	2 (1%)	1 (1%)	30	22
1	C	170/182 (93%)	168 (99%)	2 (1%)	0	100	100
1	D	156/182 (86%)	151 (97%)	5 (3%)	0	100	100
1	E	170/182 (93%)	168 (99%)	2 (1%)	0	100	100
1	F	149/182 (82%)	143 (96%)	5 (3%)	1 (1%)	26	19
2	G	13/24 (54%)	13 (100%)	0	0	100	100
2	H	13/24 (54%)	13 (100%)	0	0	100	100
2	I	12/24 (50%)	12 (100%)	0	0	100	100
2	J	11/24 (46%)	10 (91%)	1 (9%)	0	100	100
2	K	13/24 (54%)	12 (92%)	0	1 (8%)	1	0
2	L	10/24 (42%)	10 (100%)	0	0	100	100
All	All	1056/1236 (85%)	1030 (98%)	21 (2%)	5 (0%)	34	26

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	2327	SER
2	K	7	LYS
1	A	2237	GLY
1	B	2195	ASN
1	A	2195	ASN

5.3.2 Protein sidechains ⓘ

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	A	135/142 (95%)	131 (97%)	4 (3%)	48	47
1	B	134/142 (94%)	131 (98%)	3 (2%)	60	62
1	C	133/142 (94%)	129 (97%)	4 (3%)	48	47
1	D	117/142 (82%)	117 (100%)	0	100	100
1	E	132/142 (93%)	130 (98%)	2 (2%)	72	75
1	F	112/142 (79%)	110 (98%)	2 (2%)	66	69
2	G	13/20 (65%)	12 (92%)	1 (8%)	16	10
2	H	13/20 (65%)	12 (92%)	1 (8%)	16	10
2	I	12/20 (60%)	12 (100%)	0	100	100
2	J	11/20 (55%)	11 (100%)	0	100	100
2	K	13/20 (65%)	12 (92%)	1 (8%)	16	10
2	L	10/20 (50%)	9 (90%)	1 (10%)	9	5
All	All	835/972 (86%)	816 (98%)	19 (2%)	58	60

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2158	SER
1	A	2171	ILE
1	A	2201	ARG
1	A	2246	PRO
1	B	2168	GLU
1	B	2171	ILE
1	B	2329	SER
2	G	7	LYS
2	H	7	LYS
1	E	2170	SER
1	E	2171	ILE
1	F	2160	CYS
1	F	2175	THR
2	K	7	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	L	8	ARG
1	C	2171	ILE
1	C	2219	GLN
1	C	2322	VAL
1	C	2327	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2159	HIS
1	A	2220	HIS
1	B	2315	HIS
1	E	2177	GLN
1	E	2220	HIS
1	E	2239	HIS
1	C	2159	HIS
1	C	2177	GLN
1	C	2239	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	SO4	A	2401	-	4,4,4	0.67	0	6,6,6	0.55	0
3	SO4	B	2401	-	4,4,4	0.81	0	6,6,6	0.63	0
3	SO4	C	2401	-	4,4,4	0.56	0	6,6,6	0.14	0
3	SO4	E	2401	-	4,4,4	0.72	0	6,6,6	0.28	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SO4	A	2401	-	-	0/0/0/0	0/0/0/0
3	SO4	B	2401	-	-	0/0/0/0	0/0/0/0
3	SO4	C	2401	-	-	0/0/0/0	0/0/0/0
3	SO4	E	2401	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

6.1 Protein, DNA and RNA chains

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(Å ²)	Q<0.9
1	A	172/182 (94%)	0.22	7 (4%) 41 42	27, 38, 61, 87	0
1	B	171/182 (93%)	0.27	8 (4%) 35 37	26, 36, 61, 96	0
1	C	172/182 (94%)	0.45	13 (7%) 17 18	31, 47, 65, 80	0
1	D	162/182 (89%)	0.68	20 (12%) 5 6	35, 51, 77, 85	0
1	E	172/182 (94%)	0.23	4 (2%) 64 64	26, 40, 54, 65	0
1	F	157/182 (86%)	0.55	13 (8%) 14 15	28, 48, 76, 95	0
2	G	15/24 (62%)	0.40	1 (6%) 21 22	28, 34, 61, 68	0
2	H	15/24 (62%)	0.40	1 (6%) 21 22	28, 34, 54, 61	0
2	I	14/24 (58%)	0.59	1 (7%) 19 20	36, 42, 65, 65	0
2	J	13/24 (54%)	0.53	0 100 100	35, 41, 58, 67	0
2	K	15/24 (62%)	0.80	1 (6%) 21 22	33, 41, 67, 69	0
2	L	12/24 (50%)	0.89	1 (8%) 14 15	34, 36, 53, 61	0
All	All	1090/1236 (88%)	0.41	70 (6%) 23 24	26, 43, 71, 96	0

All (70) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	2167	PRO	5.5
1	D	2216	TYR	5.4
1	B	2196	HIS	5.1
1	D	2217	LYS	5.1
1	F	2216	TYR	5.0
1	B	2193	GLY	4.9
1	D	2174	MET	4.8
1	F	2327	SER	4.6
1	C	2236	GLY	4.6
1	C	2168	GLU	4.3
1	F	2328	PRO	4.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	2158	SER	4.3
1	B	2195	ASN	4.0
1	B	2235	GLU	4.0
1	A	2167	PRO	3.8
1	E	2236	GLY	3.8
1	C	2262	TRP	3.8
1	B	2169	ILE	3.8
1	D	2173	ASP	3.8
2	I	20	PRO	3.7
1	D	2175	THR	3.6
1	F	2198	TYR	3.4
1	B	2168	GLU	3.4
1	F	2174	MET	3.3
1	F	2196	HIS	3.2
1	D	2218	GLY	3.2
1	D	2202	PHE	3.1
2	G	20	PRO	3.1
1	E	2168	GLU	3.0
1	D	2327	SER	3.0
1	D	2254	GLY	3.0
1	C	2235	GLU	3.0
1	F	2236	GLY	3.0
1	C	2159	HIS	3.0
1	C	2327	SER	2.9
1	A	2235	GLU	2.9
1	D	2187	GLU	2.8
1	C	2217	LYS	2.8
1	C	2216	TYR	2.8
1	A	2195	ASN	2.8
1	D	2199	CYS	2.8
1	B	2194	GLU	2.7
1	D	2198	TYR	2.7
1	F	2187	GLU	2.6
1	F	2218	GLY	2.6
2	H	20	PRO	2.6
1	A	2169	ILE	2.6
1	D	2233	LEU	2.6
1	D	2232	PRO	2.5
1	C	2205	ALA	2.5
1	C	2239	HIS	2.5
1	E	2234	GLY	2.5
1	F	2175	THR	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	F	2219	GLN	2.3
1	B	2192	GLU	2.3
1	C	2167	PRO	2.3
1	C	2293	CYS	2.3
1	D	2231	GLY	2.3
1	A	2194	GLU	2.2
1	F	2201	ARG	2.2
1	D	2215	LYS	2.2
1	D	2219	GLN	2.2
1	E	2239	HIS	2.2
2	L	18	ALA	2.2
2	K	20	PRO	2.1
1	C	2195	ASN	2.1
1	A	2234	GLY	2.1
1	A	2168	GLU	2.1
1	F	2326	ALA	2.1
1	D	2159	HIS	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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	SO4	C	2401	5/5	0.92	0.30	1.12	57,69,71,72	0
3	SO4	E	2401	5/5	0.97	0.25	0.95	47,56,60,60	0
3	SO4	A	2401	5/5	0.95	0.17	0.07	43,50,53,56	0
3	SO4	B	2401	5/5	0.95	0.16	0.04	42,46,50,57	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	PR	D	2401	1/1	0.98	0.04	-3.76	57,57,57,57	0
5	NA	C	2402	1/1	0.99	0.35	-	66,66,66,66	1
5	NA	E	2402	1/1	0.68	0.21	-	50,50,50,50	1
4	PR	A	2402	1/1	0.98	0.08	-	53,53,53,53	0
4	PR	E	2403	1/1	0.98	0.08	-	45,45,45,45	0
4	PR	B	2402	1/1	0.97	0.08	-	53,53,53,53	0

6.5 Other polymers [i](#)

There are no such residues in this entry.