wwPDB X-ray Structure Validation Summary Report (i)

Feb 19, 2016 – 08:14 PM GMT

PDB ID	:	4XGZ
Title	:	Crystal structure of human paxillin LD2 motif in complex with Fab fragment
Authors	:	Nocula-Lugowska, M.; Lugowski, M.; Salgia, R.; Kossiakoff, A.A.
Deposited on	:	2015-01-04
Resolution	:	2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

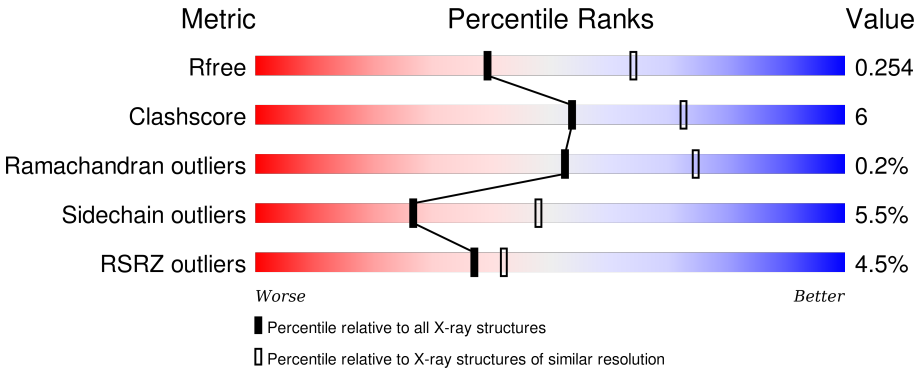
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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.



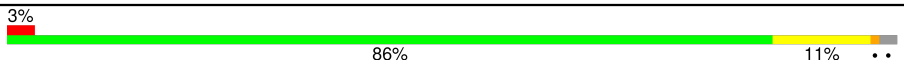
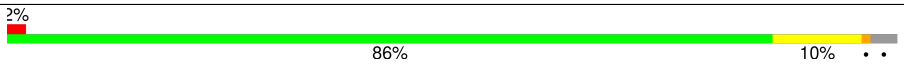
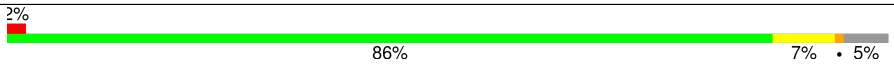
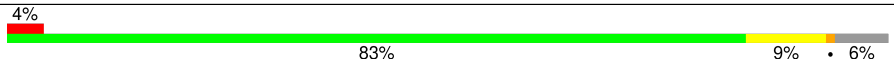
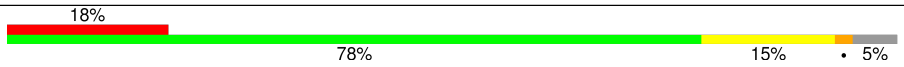
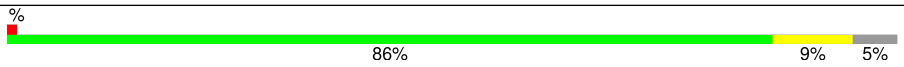
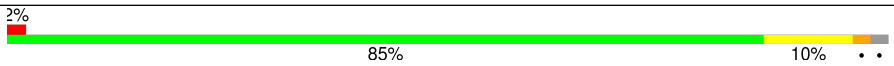
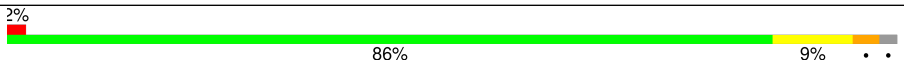
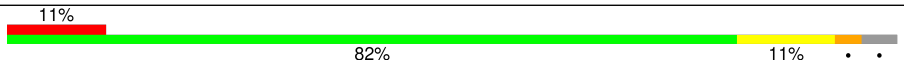
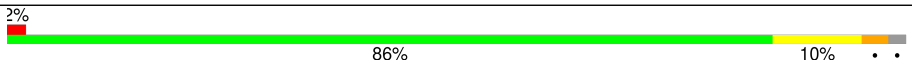
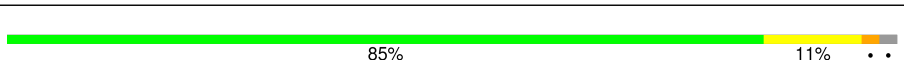
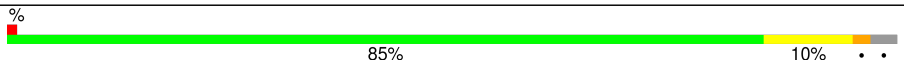
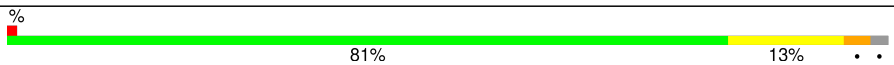
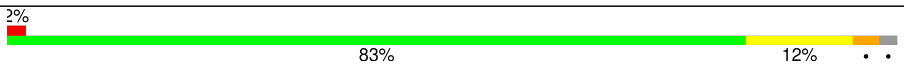
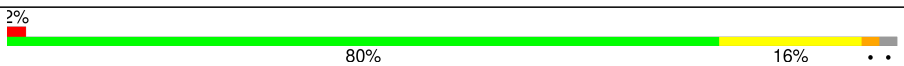
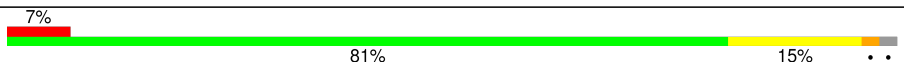
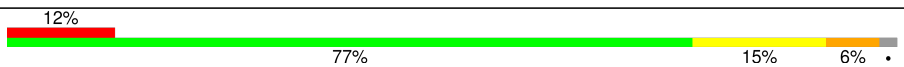
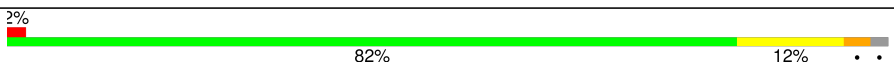
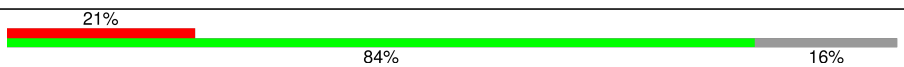
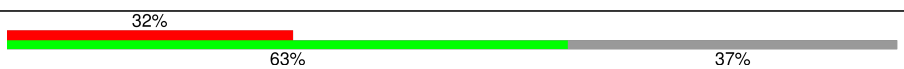
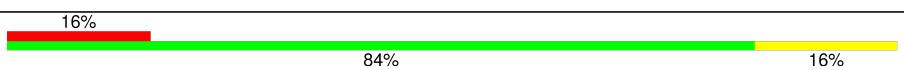
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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 <=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	229	86% 9% 5%
1	C	229	% 84% 10% • •
1	E	229	13% 83% 9% 7%
1	G	229	2% 86% 8% • 5%
1	H	229	85% 9% • •

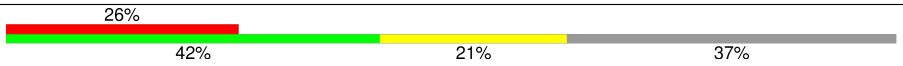
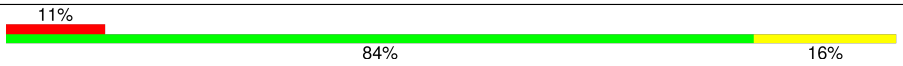
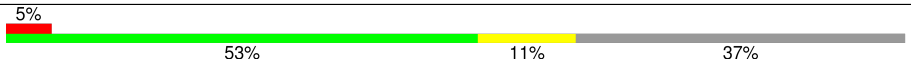
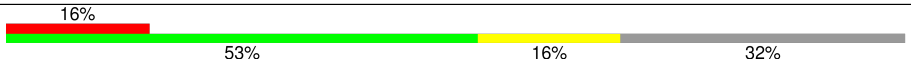
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	229	
1	M	229	
1	O	229	
1	Q	229	
1	S	229	
1	U	229	
1	W	229	
2	B	215	
2	D	215	
2	F	215	
2	I	215	
2	K	215	
2	L	215	
2	N	215	
2	P	215	
2	R	215	
2	T	215	
2	V	215	
2	X	215	
3	a	19	
3	c	19	
3	e	19	
3	g	19	
3	h	19	
3	j	19	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	m	19	
3	o	19	
3	q	19	
3	s	19	
3	u	19	
3	w	19	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NO3	B	301	-	-	-	X
4	NO3	D	301	-	-	-	X
4	NO3	F	301	-	-	-	X
4	NO3	K	301	-	-	-	X
4	NO3	L	301	-	-	-	X
4	NO3	P	301	-	-	-	X
4	NO3	T	301	-	-	-	X
4	NO3	V	301	-	-	-	X
5	EDO	Q	301	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 40426 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 FAB HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	217	Total	C	N	O	S	0	0	0
			1620	1027	266	320	7			
1	C	219	Total	C	N	O	S	0	0	0
			1633	1034	268	324	7			
1	E	212	Total	C	N	O	S	0	0	0
			1586	1009	260	311	6			
1	G	218	Total	C	N	O	S	0	0	0
			1624	1029	267	321	7			
1	H	219	Total	C	N	O	S	0	0	0
			1634	1036	269	323	6			
1	J	224	Total	C	N	O	S	0	0	0
			1662	1051	273	331	7			
1	M	220	Total	C	N	O	S	0	0	0
			1643	1040	270	326	7			
1	O	223	Total	C	N	O	S	0	0	0
			1660	1049	273	331	7			
1	Q	217	Total	C	N	O	S	0	0	0
			1618	1025	265	321	7			
1	S	215	Total	C	N	O	S	0	0	0
			1612	1025	264	317	6			
1	U	217	Total	C	N	O	S	0	0	0
			1618	1026	266	319	7			
1	W	218	Total	C	N	O	S	0	0	0
			1624	1029	267	321	7			

- Molecule 2 is a protein called FAB LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			
2	D	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	207	Total	C	N	O	S	0	0	0
			1568	979	263	322	4			
2	I	211	Total	C	N	O	S	0	0	0
			1598	995	270	328	5			
2	K	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			
2	L	209	Total	C	N	O	S	0	0	0
			1587	989	268	326	4			
2	N	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			
2	P	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			
2	R	211	Total	C	N	O	S	0	0	0
			1601	997	270	328	6			
2	T	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			
2	V	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			
2	X	210	Total	C	N	O	S	0	0	0
			1593	992	269	327	5			

- Molecule 3 is a protein called PAXILLIN LD2.

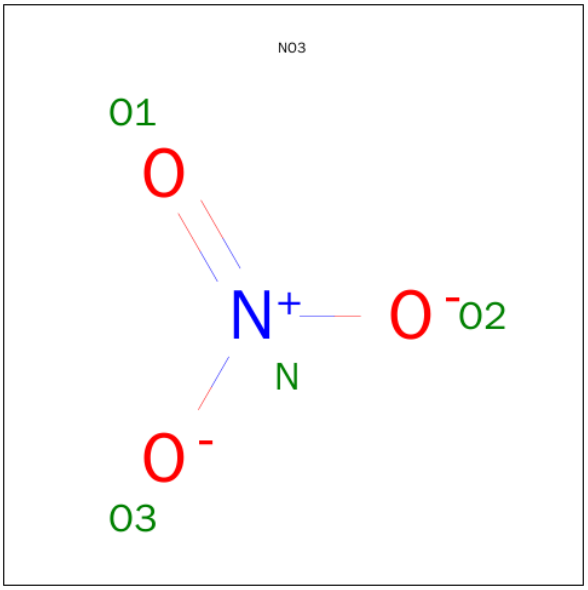
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	a	12	Total	C	N	O	0	0	0
			99	63	16	20			
3	c	16	Total	C	N	O	0	0	0
			128	80	22	26			
3	e	13	Total	C	N	O	0	0	0
			107	67	18	22			
3	g	13	Total	C	N	O	0	0	0
			107	67	18	22			
3	h	12	Total	C	N	O	0	0	0
			99	63	16	20			
3	j	19	Total	C	N	O	0	0	0
			153	95	28	30			
3	m	13	Total	C	N	O	0	0	0
			107	67	18	22			
3	o	12	Total	C	N	O	0	0	0
			99	63	16	20			
3	q	19	Total	C	N	O	0	0	0
			153	95	28	30			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	s	12	Total C N O 99 63 16 20	0	0	0
3	u	13	Total C N O 107 67 18 22	0	0	0
3	w	12	Total C N O 99 63 16 20	0	0	0

- Molecule 4 is NITRATE ION (three-letter code: NO3) (formula: NO_3).



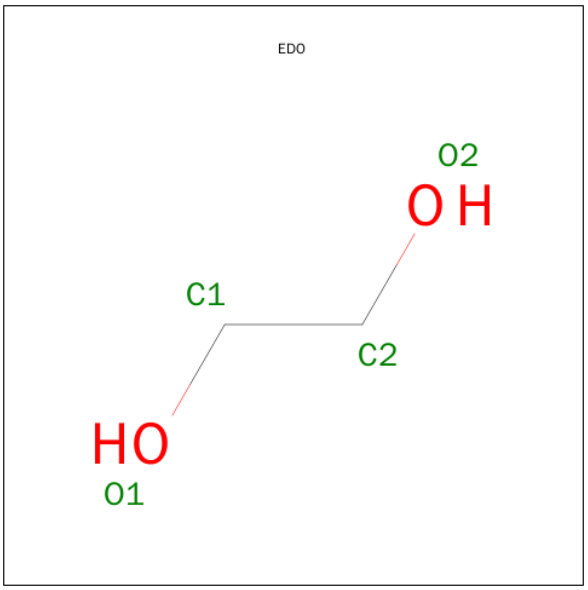
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total N O 4 1 3	0	0
4	D	1	Total N O 4 1 3	0	0
4	F	1	Total N O 4 1 3	0	0
4	I	1	Total N O 4 1 3	0	0
4	K	1	Total N O 4 1 3	0	0
4	L	1	Total N O 4 1 3	0	0
4	N	1	Total N O 4 1 3	0	0
4	P	1	Total N O 4 1 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	R	1	Total N O 4 1 3	0	0
4	T	1	Total N O 4 1 3	0	0
4	V	1	Total N O 4 1 3	0	0
4	X	1	Total N O 4 1 3	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	N	1	Total C O 4 2 2	0	0
5	Q	1	Total C O 4 2 2	0	0
5	W	1	Total C O 4 2 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	11	Total O 11 11	0	0
6	B	5	Total O 5 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	19	Total O 19 19	0	0
6	D	16	Total O 16 16	0	0
6	E	10	Total O 10 10	0	0
6	F	2	Total O 2 2	0	0
6	G	19	Total O 19 19	0	0
6	H	29	Total O 29 29	0	0
6	I	14	Total O 14 14	0	0
6	J	17	Total O 17 17	0	0
6	K	13	Total O 13 13	0	0
6	L	17	Total O 17 17	0	0
6	M	25	Total O 25 25	0	0
6	N	31	Total O 31 31	0	0
6	O	13	Total O 13 13	0	0
6	P	11	Total O 11 11	0	0
6	Q	21	Total O 21 21	0	0
6	R	20	Total O 20 20	0	0
6	S	19	Total O 19 19	0	0
6	T	10	Total O 10 10	0	0
6	U	10	Total O 10 10	0	0
6	V	4	Total O 4 4	0	0
6	W	12	Total O 12 12	0	0

Continued on next page...

Continued from previous page...

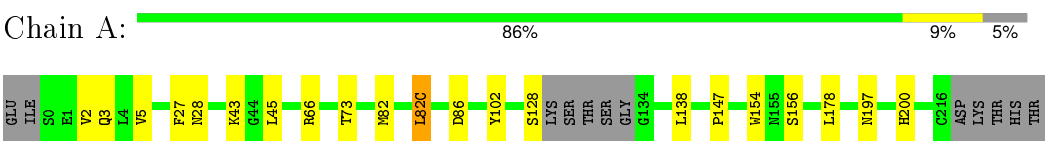
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	X	12	Total O 12 12	0	0
6	a	1	Total O 1 1	0	0
6	c	1	Total O 1 1	0	0
6	e	1	Total O 1 1	0	0
6	g	2	Total O 2 2	0	0
6	j	4	Total O 4 4	0	0
6	m	1	Total O 1 1	0	0
6	o	1	Total O 1 1	0	0
6	q	3	Total O 3 3	0	0
6	s	1	Total O 1 1	0	0
6	u	1	Total O 1 1	0	0
6	w	1	Total O 1 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.

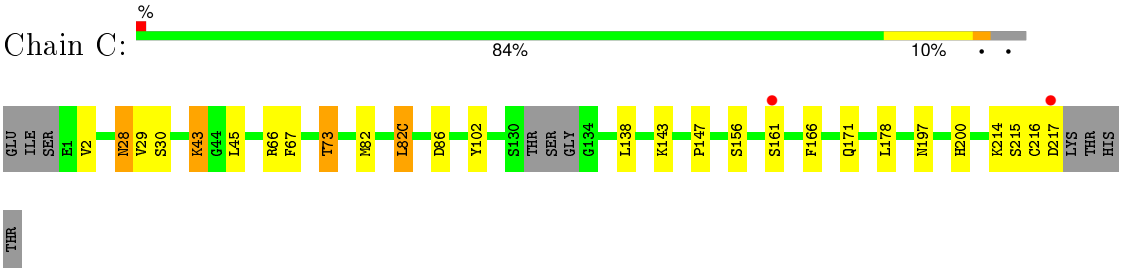
- Molecule 1: FAB HEAVY CHAIN

Chain A:



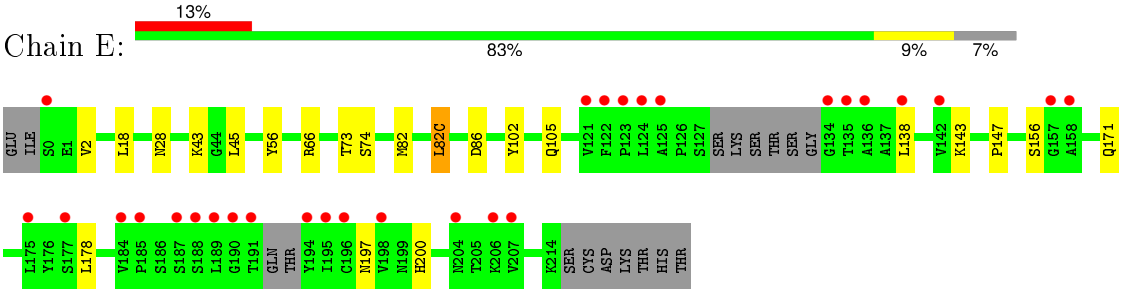
- Molecule 1: FAB HEAVY CHAIN

Chain C:



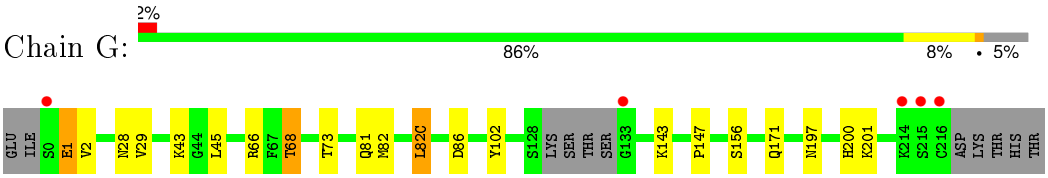
- Molecule 1: FAB HEAVY CHAIN

Chain E:



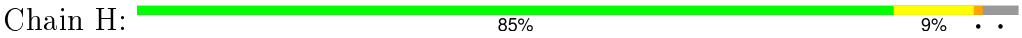
- Molecule 1: FAB HEAVY CHAIN

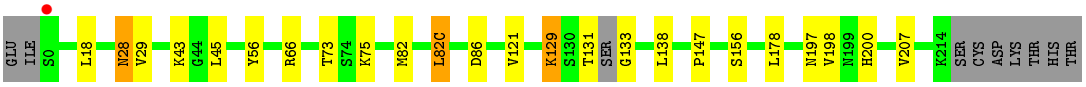
Chain G:



- Molecule 1: FAB HEAVY CHAIN

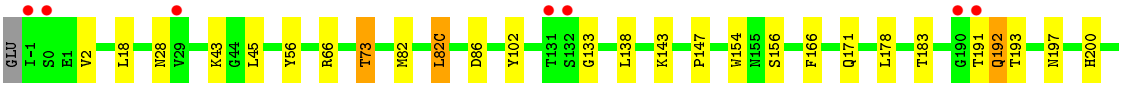
Chain H:

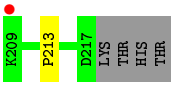




• Molecule 1: FAB HEAVY CHAIN

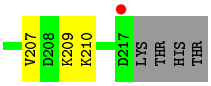
Chain J:





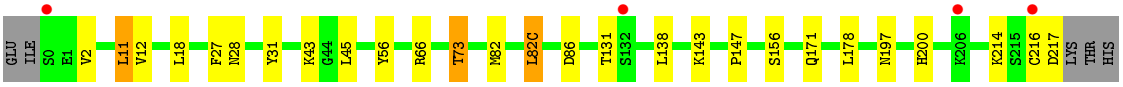
• Molecule 1: FAB HEAVY CHAIN

Chain M:



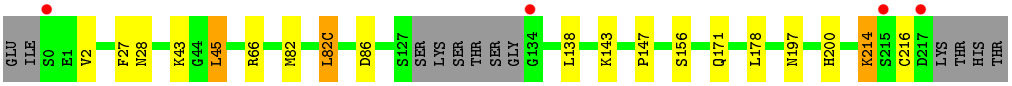
• Molecule 1: FAB HEAVY CHAIN

Chain O:



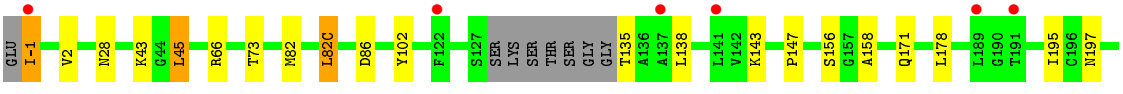
• Molecule 1: FAB HEAVY CHAIN

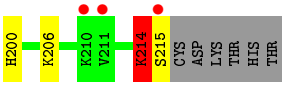
Chain Q:



• Molecule 1: FAB HEAVY CHAIN

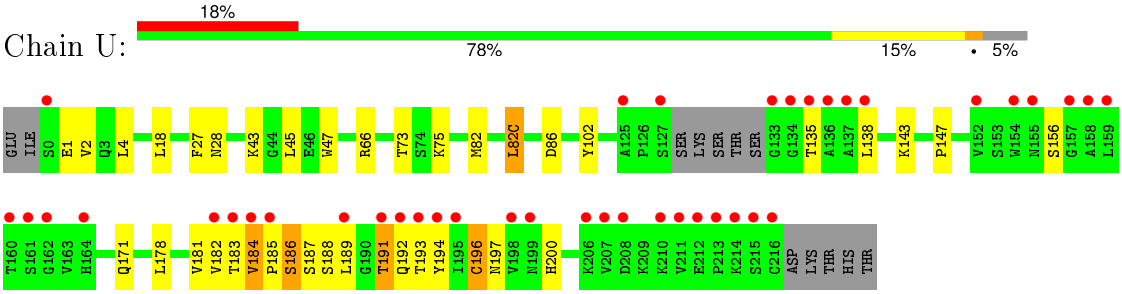
Chain S:





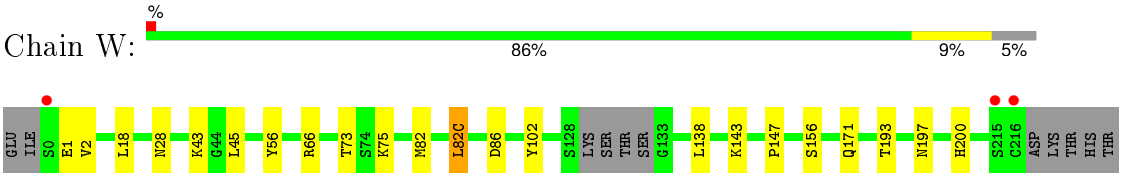
• Molecule 1: FAB HEAVY CHAIN

Chain U:



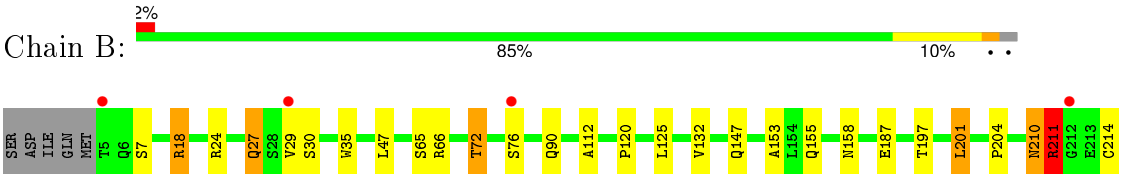
• Molecule 1: FAB HEAVY CHAIN

Chain W:



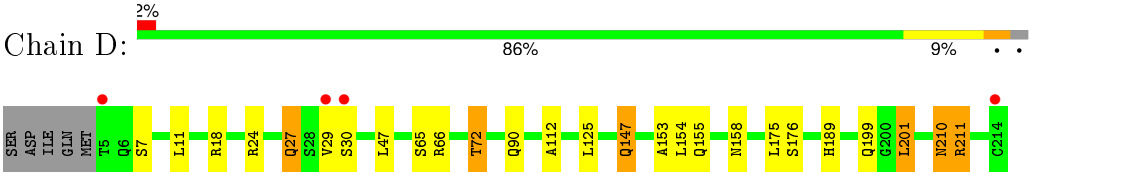
• Molecule 2: FAB LIGHT CHAIN

Chain B:



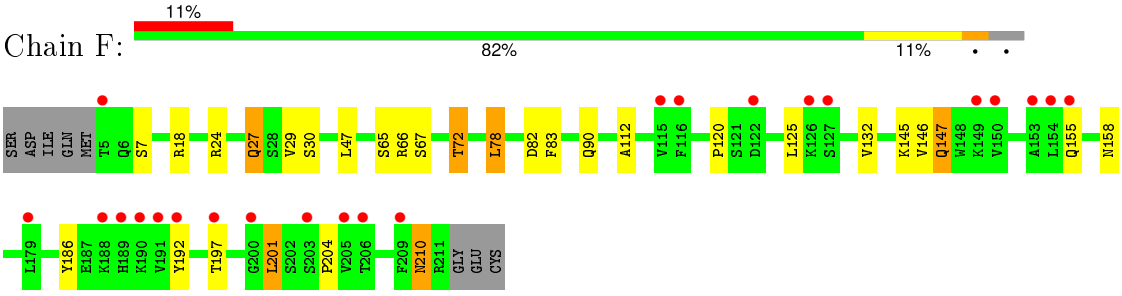
• Molecule 2: FAB LIGHT CHAIN

Chain D:



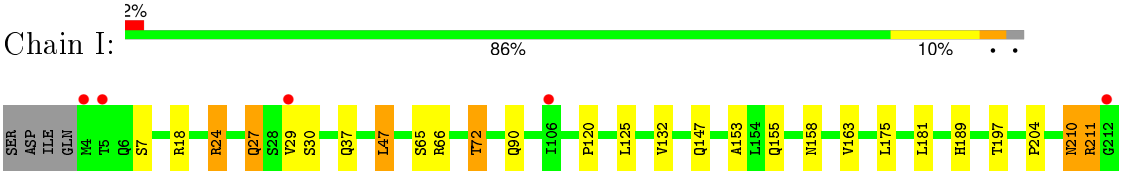
• Molecule 2: FAB LIGHT CHAIN

Chain F:



• Molecule 2: FAB LIGHT CHAIN

Chain I:

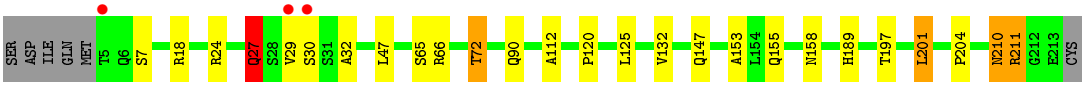


• Molecule 2: FAB LIGHT CHAIN

Chain K:

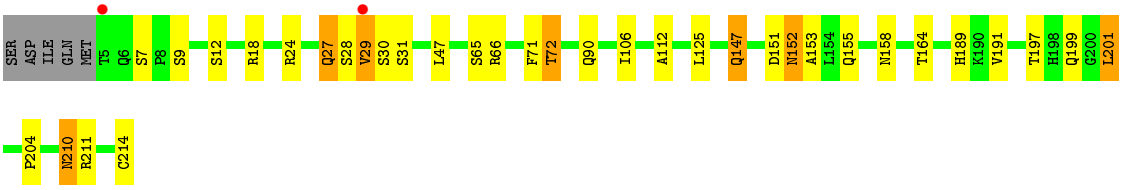
• Molecule 2: FAB LIGHT CHAIN

Chain L:



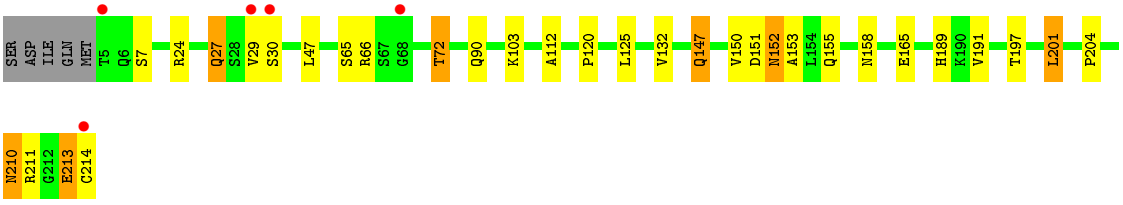
• Molecule 2: FAB LIGHT CHAIN

Chain N:



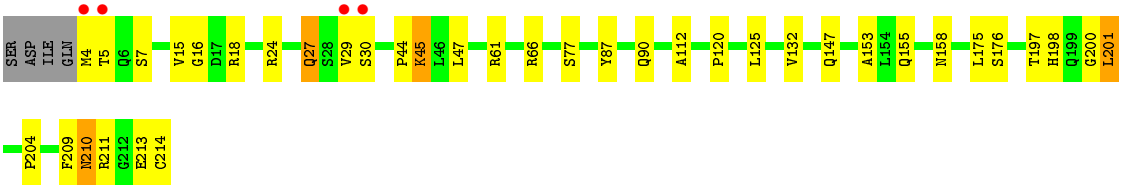
• Molecule 2: FAB LIGHT CHAIN

Chain P:



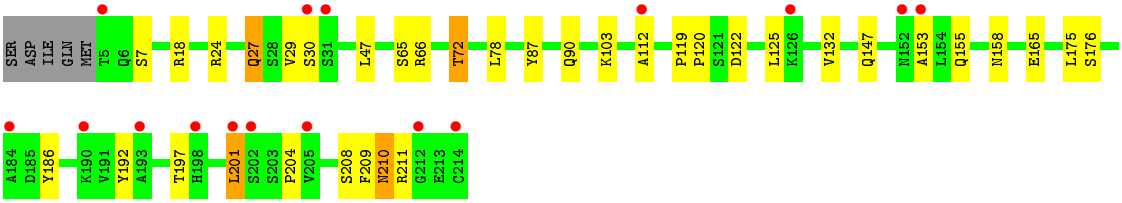
• Molecule 2: FAB LIGHT CHAIN

Chain R:

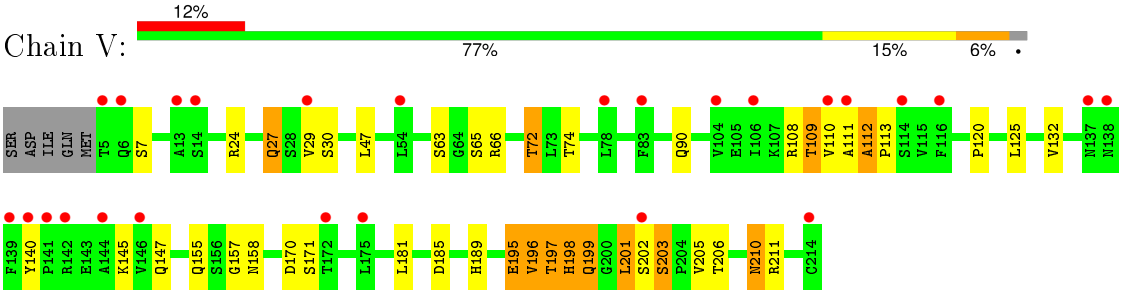


• Molecule 2: FAB LIGHT CHAIN

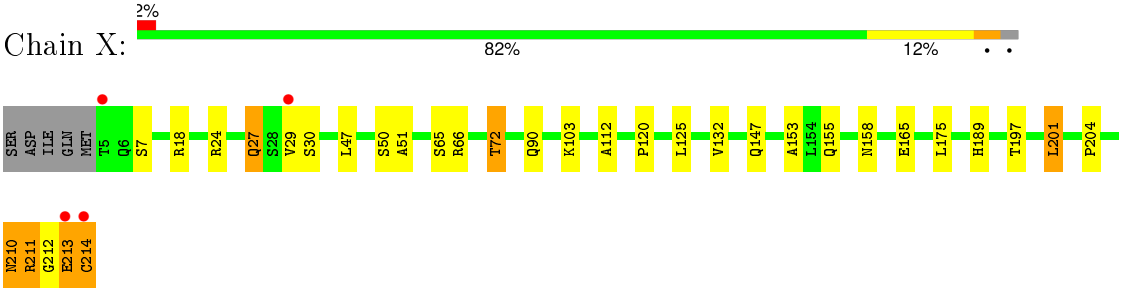
Chain T:



• Molecule 2: FAB LIGHT CHAIN

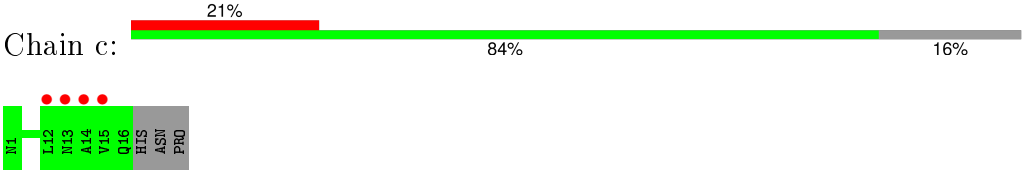


• Molecule 2: FAB LIGHT CHAIN



• Molecule 3: PAXILLIN LD2

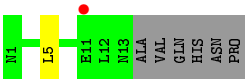
• Molecule 3: PAXILLIN LD2



• Molecule 3: PAXILLIN LD2

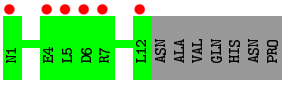
• Molecule 3: PAXILLIN LD2

Chain g:



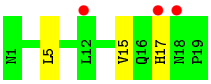
- Molecule 3: PAXILLIN LD2

Chain h:



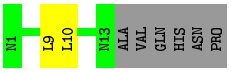
- Molecule 3: PAXILLIN LD2

Chain j:



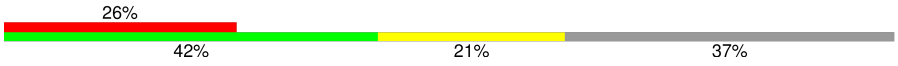
- Molecule 3: PAXILLIN LD2

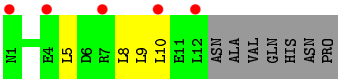
Chain m:



- Molecule 3: PAXILLIN LD2

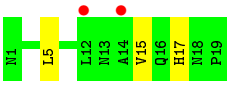
Chain o:





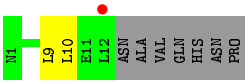
- Molecule 3: PAXILLIN LD2

Chain q:



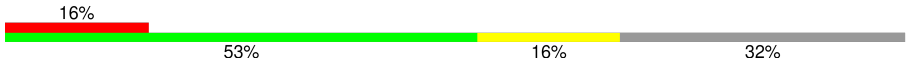
- Molecule 3: PAXILLIN LD2

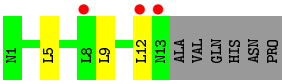
Chain s:



- Molecule 3: PAXILLIN LD2

Chain u:





- Molecule 3: PAXILLIN LD2

Chain w:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	108.19Å 111.00Å 143.87Å 94.65° 95.28° 114.53°	Depositor
Resolution (Å)	142.00 – 2.50 100.11 – 2.50	Depositor EDS
% Data completeness (in resolution range)	94.3 (142.00-2.50) 89.9 (100.11-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.228 , 0.245 0.237 , 0.254	Depositor DCC
R_{free} test set	9884 reflections (5.32%)	DCC
Wilson B-factor (Å^2)	46.4	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 30.0	EDS
Estimated twinning fraction	0.013 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 195832 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	40426	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, NO3

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	0.47	1/1661 (0.1%)	0.54	0/2265
1	C	0.48	0/1674	0.56	0/2283
1	E	0.48	1/1626 (0.1%)	0.55	1/2216 (0.0%)
1	G	0.46	0/1665	0.55	0/2270
1	H	0.47	0/1675	0.56	0/2283
1	J	0.48	1/1704 (0.1%)	0.57	0/2325
1	M	0.46	0/1684	0.56	0/2295
1	O	0.44	0/1702	0.54	0/2321
1	Q	0.44	0/1659	0.54	0/2264
1	S	0.45	0/1653	0.53	1/2255 (0.0%)
1	U	0.51	1/1659 (0.1%)	0.67	1/2262 (0.0%)
1	W	0.47	0/1665	0.55	0/2270
2	B	0.40	1/1625 (0.1%)	0.59	3/2205 (0.1%)
2	D	0.41	0/1625	0.58	1/2205 (0.0%)
2	F	0.41	0/1600	0.52	0/2173
2	I	0.41	0/1630	0.60	2/2212 (0.1%)
2	K	0.41	0/1625	0.60	2/2205 (0.1%)
2	L	0.39	0/1619	0.57	2/2197 (0.1%)
2	N	0.43	0/1625	0.57	0/2205
2	P	0.41	0/1625	0.56	0/2205
2	R	0.41	0/1633	0.62	2/2215 (0.1%)
2	T	0.43	1/1625 (0.1%)	0.59	2/2205 (0.1%)
2	V	0.48	0/1625	0.65	1/2205 (0.0%)
2	X	0.45	1/1625 (0.1%)	0.57	1/2205 (0.0%)
3	a	1.17	2/98 (2.0%)	1.24	2/131 (1.5%)
3	c	0.40	0/127	0.63	0/171
3	e	0.41	0/106	0.61	0/142
3	g	0.40	0/106	0.61	0/142
3	h	0.48	0/98	0.62	0/131
3	j	0.48	0/154	0.73	0/209
3	m	0.39	0/106	0.66	0/142
3	o	0.41	0/98	0.79	0/131

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
3	q	0.48	0/154	0.70	0/209
3	s	0.50	0/98	0.60	0/131
3	u	0.46	0/106	0.86	0/142
3	w	0.38	0/98	0.61	0/131
All	All	0.45	9/40858 (0.0%)	0.58	21/55558 (0.0%)

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	J	0	1
1	Q	0	1
All	All	0	2

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	a	11	GLU	CD-OE2	-7.46	1.17	1.25
3	a	11	GLU	CD-OE1	6.15	1.32	1.25
2	X	213	GLU	CG-CD	5.72	1.60	1.51
1	E	105	GLN	C-O	5.60	1.33	1.23
2	B	35	TRP	CD2-CE2	5.11	1.47	1.41

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	U	196	CYS	CA-CB-SG	12.72	136.90	114.00
3	a	11	GLU	CG-CD-OE2	-8.51	101.28	118.30
2	K	211	ARG	NE-CZ-NH1	-8.36	116.12	120.30
3	a	11	GLU	CG-CD-OE1	8.32	134.93	118.30
1	E	105	GLN	N-CA-C	-8.07	89.21	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	J	213	PRO	Peptide
1	Q	214	LYS	Peptide

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	A	1620	0	1567	10	1
1	C	1633	0	1573	20	0
1	E	1586	0	1537	15	0
1	G	1624	0	1570	10	0
1	H	1634	0	1586	23	0
1	J	1662	0	1600	20	1
1	M	1643	0	1589	33	0
1	O	1660	0	1605	18	0
1	Q	1618	0	1555	9	0
1	S	1612	0	1566	26	0
1	U	1618	0	1565	57	0
1	W	1624	0	1570	16	0
2	B	1593	0	1555	12	0
2	D	1593	0	1555	21	0
2	F	1568	0	1531	16	1
2	I	1598	0	1557	13	0
2	K	1593	0	1555	13	0
2	L	1587	0	1551	15	0
2	N	1593	0	1555	27	0
2	P	1593	0	1555	22	0
2	R	1601	0	1564	34	1
2	T	1593	0	1556	16	0
2	V	1593	0	1555	48	0
2	X	1593	0	1555	18	0
3	a	99	0	108	0	0
3	c	128	0	136	0	0
3	e	107	0	114	0	0
3	g	107	0	114	0	0
3	h	99	0	108	0	0
3	j	153	0	156	0	0
3	m	107	0	114	0	0
3	o	99	0	108	0	0
3	q	153	0	156	0	0
3	s	99	0	108	0	0
3	u	107	0	114	0	0
3	w	99	0	108	0	0
4	B	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	4	0	0	0	0
4	F	4	0	0	0	0
4	I	4	0	0	1	0
4	K	4	0	0	0	0
4	L	4	0	0	0	0
4	N	4	0	0	0	0
4	P	4	0	0	0	0
4	R	4	0	0	0	0
4	T	4	0	0	0	0
4	V	4	0	0	0	0
4	X	4	0	0	1	0
5	N	4	0	6	0	0
5	Q	4	0	6	0	0
5	W	4	0	6	0	0
6	A	11	0	0	0	0
6	B	5	0	0	1	0
6	C	19	0	0	0	0
6	D	16	0	0	0	0
6	E	10	0	0	0	0
6	F	2	0	0	0	0
6	G	19	0	0	0	0
6	H	29	0	0	3	0
6	I	14	0	0	0	0
6	J	17	0	0	2	0
6	K	13	0	0	0	0
6	L	17	0	0	1	0
6	M	25	0	0	0	0
6	N	31	0	0	2	0
6	O	13	0	0	0	0
6	P	11	0	0	0	0
6	Q	21	0	0	0	0
6	R	20	0	0	1	0
6	S	19	0	0	2	0
6	T	10	0	0	0	0
6	U	10	0	0	0	0
6	V	4	0	0	0	0
6	W	12	0	0	2	0
6	X	12	0	0	0	0
6	a	1	0	0	0	0
6	c	1	0	0	0	0
6	e	1	0	0	0	0
6	g	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	j	4	0	0	0	0
6	m	1	0	0	0	0
6	o	1	0	0	0	0
6	q	3	0	0	0	0
6	s	1	0	0	0	0
6	u	1	0	0	0	0
6	w	1	0	0	0	0
All	All	40426	0	38989	451	2

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.

The worst 5 of 451 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:U:184:VAL:HG21	1:U:194:TYR:CZ	1.64	1.31
2:V:108:ARG:NH1	2:V:109:THR:O	1.68	1.24
1:M:193:THR:HG22	2:R:77:SER:OG	1.08	1.24
2:V:198:HIS:HB3	2:V:201:LEU:HB3	1.27	1.16
2:V:201:LEU:HD21	2:V:203:SER:O	1.50	1.12

All (2) 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:F:18:ARG:NH2	2:R:213:GLU:OE2[1_454]	2.07	0.13
1:A:5:VAL:CG1	1:J:191:THR:CG2[1_554]	2.15	0.05

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	213/229 (93%)	213 (100%)	0	0	100	100
1	C	215/229 (94%)	215 (100%)	0	0	100	100
1	E	206/229 (90%)	206 (100%)	0	0	100	100
1	G	214/229 (93%)	211 (99%)	2 (1%)	1 (0%)	34	55
1	H	215/229 (94%)	213 (99%)	2 (1%)	0	100	100
1	J	222/229 (97%)	220 (99%)	1 (0%)	1 (0%)	34	55
1	M	216/229 (94%)	215 (100%)	0	1 (0%)	34	55
1	O	221/229 (96%)	218 (99%)	2 (1%)	1 (0%)	34	55
1	Q	213/229 (93%)	212 (100%)	1 (0%)	0	100	100
1	S	211/229 (92%)	210 (100%)	1 (0%)	0	100	100
1	U	213/229 (93%)	211 (99%)	1 (0%)	1 (0%)	34	55
1	W	214/229 (93%)	212 (99%)	1 (0%)	1 (0%)	34	55
2	B	208/215 (97%)	202 (97%)	6 (3%)	0	100	100
2	D	208/215 (97%)	202 (97%)	6 (3%)	0	100	100
2	F	205/215 (95%)	199 (97%)	6 (3%)	0	100	100
2	I	209/215 (97%)	203 (97%)	6 (3%)	0	100	100
2	K	208/215 (97%)	202 (97%)	6 (3%)	0	100	100
2	L	207/215 (96%)	200 (97%)	7 (3%)	0	100	100
2	N	208/215 (97%)	201 (97%)	6 (3%)	1 (0%)	34	55
2	P	208/215 (97%)	202 (97%)	6 (3%)	0	100	100
2	R	209/215 (97%)	203 (97%)	6 (3%)	0	100	100
2	T	208/215 (97%)	201 (97%)	7 (3%)	0	100	100
2	V	208/215 (97%)	199 (96%)	7 (3%)	2 (1%)	19	34
2	X	208/215 (97%)	200 (96%)	7 (3%)	1 (0%)	34	55
3	a	10/19 (53%)	10 (100%)	0	0	100	100
3	c	14/19 (74%)	14 (100%)	0	0	100	100
3	e	11/19 (58%)	11 (100%)	0	0	100	100
3	g	11/19 (58%)	11 (100%)	0	0	100	100
3	h	10/19 (53%)	10 (100%)	0	0	100	100
3	j	17/19 (90%)	15 (88%)	2 (12%)	0	100	100
3	m	11/19 (58%)	11 (100%)	0	0	100	100
3	o	10/19 (53%)	10 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	q	17/19 (90%)	15 (88%)	2 (12%)	0	100	100
3	s	10/19 (53%)	10 (100%)	0	0	100	100
3	u	11/19 (58%)	11 (100%)	0	0	100	100
3	w	10/19 (53%)	10 (100%)	0	0	100	100
All	All	5209/5556 (94%)	5108 (98%)	91 (2%)	10 (0%)	52	75

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	128	SER
2	N	29	VAL
1	G	1	GLU
2	V	198	HIS
2	X	211	ARG

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	A	178/189 (94%)	170 (96%)	8 (4%)	34	59
1	C	179/189 (95%)	170 (95%)	9 (5%)	30	53
1	E	173/189 (92%)	168 (97%)	5 (3%)	50	77
1	G	178/189 (94%)	169 (95%)	9 (5%)	29	52
1	H	179/189 (95%)	173 (97%)	6 (3%)	44	72
1	J	182/189 (96%)	175 (96%)	7 (4%)	40	67
1	M	181/189 (96%)	175 (97%)	6 (3%)	45	73
1	O	183/189 (97%)	175 (96%)	8 (4%)	35	60
1	Q	177/189 (94%)	172 (97%)	5 (3%)	51	78
1	S	177/189 (94%)	168 (95%)	9 (5%)	29	52
1	U	177/189 (94%)	165 (93%)	12 (7%)	20	36
1	W	178/189 (94%)	172 (97%)	6 (3%)	44	72

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	185/190 (97%)	171 (92%)	14 (8%)	16 30
2	D	185/190 (97%)	176 (95%)	9 (5%)	31 55
2	F	182/190 (96%)	171 (94%)	11 (6%)	24 43
2	I	185/190 (97%)	175 (95%)	10 (5%)	27 49
2	K	185/190 (97%)	174 (94%)	11 (6%)	24 44
2	L	184/190 (97%)	174 (95%)	10 (5%)	27 49
2	N	185/190 (97%)	174 (94%)	11 (6%)	24 44
2	P	185/190 (97%)	173 (94%)	12 (6%)	21 39
2	R	186/190 (98%)	175 (94%)	11 (6%)	24 44
2	T	185/190 (97%)	174 (94%)	11 (6%)	24 44
2	V	185/190 (97%)	168 (91%)	17 (9%)	11 21
2	X	185/190 (97%)	174 (94%)	11 (6%)	24 44
3	a	12/18 (67%)	11 (92%)	1 (8%)	14 26
3	c	15/18 (83%)	15 (100%)	0	100 100
3	e	13/18 (72%)	13 (100%)	0	100 100
3	g	13/18 (72%)	12 (92%)	1 (8%)	16 30
3	h	12/18 (67%)	12 (100%)	0	100 100
3	j	18/18 (100%)	15 (83%)	3 (17%)	3 5
3	m	13/18 (72%)	11 (85%)	2 (15%)	3 6
3	o	12/18 (67%)	8 (67%)	4 (33%)	0 0
3	q	18/18 (100%)	15 (83%)	3 (17%)	3 5
3	s	12/18 (67%)	10 (83%)	2 (17%)	3 5
3	u	13/18 (72%)	10 (77%)	3 (23%)	1 1
3	w	12/18 (67%)	11 (92%)	1 (8%)	14 26
All	All	4522/4764 (95%)	4274 (94%)	248 (6%)	27 48

5 of 248 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	82(C)	LEU
2	P	47	LEU
3	g	5	LEU
2	N	7	SER
2	N	214	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 99 such sidechains are listed below:

Mol	Chain	Res	Type
2	K	155	GLN
2	N	155	GLN
1	W	200	HIS
2	L	27	GLN
1	M	28	ASN

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)

15 ligands are modelled in this entry.

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	$\#\|Z\| > 2$
4	NO3	B	301	-	1,3,3	0.30	0	0,3,3	0.00	-
4	NO3	D	301	-	1,3,3	0.26	0	0,3,3	0.00	-
4	NO3	F	301	-	1,3,3	0.39	0	0,3,3	0.00	-
4	NO3	I	301	-	1,3,3	0.35	0	0,3,3	0.00	-
4	NO3	K	301	-	1,3,3	0.55	0	0,3,3	0.00	-
4	NO3	L	301	-	1,3,3	0.19	0	0,3,3	0.00	-
5	EDO	N	301	-	3,3,3	0.52	0	2,2,2	0.12	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NO3	N	302	-	1,3,3	0.52	0	0,3,3	0.00	-
4	NO3	P	301	-	1,3,3	0.25	0	0,3,3	0.00	-
5	EDO	Q	301	-	3,3,3	0.44	0	2,2,2	0.21	0
4	NO3	R	301	-	1,3,3	0.39	0	0,3,3	0.00	-
4	NO3	T	301	-	1,3,3	0.53	0	0,3,3	0.00	-
4	NO3	V	301	-	1,3,3	0.31	0	0,3,3	0.00	-
5	EDO	W	301	-	3,3,3	0.31	0	2,2,2	0.55	0
4	NO3	X	301	-	1,3,3	0.36	0	0,3,3	0.00	-

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
4	NO3	B	301	-	-	0/0/0/0	0/0/0/0
4	NO3	D	301	-	-	0/0/0/0	0/0/0/0
4	NO3	F	301	-	-	0/0/0/0	0/0/0/0
4	NO3	I	301	-	-	0/0/0/0	0/0/0/0
4	NO3	K	301	-	-	0/0/0/0	0/0/0/0
4	NO3	L	301	-	-	0/0/0/0	0/0/0/0
5	EDO	N	301	-	-	0/1/1/1	0/0/0/0
4	NO3	N	302	-	-	0/0/0/0	0/0/0/0
4	NO3	P	301	-	-	0/0/0/0	0/0/0/0
5	EDO	Q	301	-	-	0/1/1/1	0/0/0/0
4	NO3	R	301	-	-	0/0/0/0	0/0/0/0
4	NO3	T	301	-	-	0/0/0/0	0/0/0/0
4	NO3	V	301	-	-	0/0/0/0	0/0/0/0
5	EDO	W	301	-	-	0/1/1/1	0/0/0/0
4	NO3	X	301	-	-	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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	I	301	NO3	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	X	301	NO3	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95^{th} 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	217/229 (94%)	0.08	0 100 100	30, 45, 74, 109	0
1	C	219/229 (95%)	0.10	2 (0%) 85 88	24, 36, 67, 108	0
1	E	212/229 (92%)	0.76	29 (13%) 4 4	38, 61, 111, 133	0
1	G	218/229 (95%)	0.11	5 (2%) 64 67	29, 39, 59, 122	0
1	H	219/229 (95%)	0.08	1 (0%) 91 92	25, 36, 62, 81	0
1	J	224/229 (97%)	0.31	8 (3%) 46 51	27, 46, 90, 110	0
1	M	220/229 (96%)	0.08	4 (1%) 71 75	24, 36, 59, 109	0
1	O	223/229 (97%)	0.22	4 (1%) 71 75	36, 47, 76, 108	0
1	Q	217/229 (94%)	0.12	4 (1%) 71 75	24, 40, 75, 101	0
1	S	215/229 (93%)	0.26	9 (4%) 40 45	30, 49, 91, 113	0
1	U	217/229 (94%)	0.87	41 (18%) 2 1	34, 52, 125, 144	0
1	W	218/229 (95%)	0.15	3 (1%) 78 80	31, 44, 63, 118	0
2	B	210/215 (97%)	0.25	4 (1%) 70 73	42, 56, 73, 101	0
2	D	210/215 (97%)	0.15	4 (1%) 70 73	30, 42, 61, 101	0
2	F	207/215 (96%)	0.78	23 (11%) 7 7	45, 71, 97, 107	0
2	I	211/215 (98%)	0.15	5 (2%) 62 66	31, 42, 63, 104	0
2	K	210/215 (97%)	0.27	1 (0%) 91 92	29, 48, 72, 97	0
2	L	209/215 (97%)	0.16	3 (1%) 78 80	31, 43, 61, 93	0
2	N	210/215 (97%)	0.18	2 (0%) 84 86	30, 41, 60, 90	0
2	P	210/215 (97%)	0.29	5 (2%) 62 66	32, 50, 74, 90	0
2	R	211/215 (98%)	0.11	4 (1%) 70 73	25, 39, 61, 107	0
2	T	210/215 (97%)	0.52	16 (7%) 17 18	38, 59, 82, 111	0
2	V	210/215 (97%)	0.91	26 (12%) 5 5	50, 69, 99, 114	0
2	X	210/215 (97%)	0.17	4 (1%) 70 73	33, 45, 68, 109	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
3	a	12/19 (63%)	1.53	5 (41%) 0 0	64, 75, 88, 94	0
3	c	16/19 (84%)	1.35	4 (25%) 1 1	57, 75, 101, 107	0
3	e	13/19 (68%)	0.77	0 100 100	57, 68, 85, 90	0
3	g	13/19 (68%)	1.05	1 (7%) 16 18	60, 71, 88, 93	0
3	h	12/19 (63%)	2.02	6 (50%) 0 0	77, 87, 97, 99	0
3	j	19/19 (100%)	0.98	3 (15%) 3 2	43, 55, 83, 89	0
3	m	13/19 (68%)	0.39	0 100 100	51, 59, 72, 80	0
3	o	12/19 (63%)	1.62	5 (41%) 0 0	88, 96, 107, 113	0
3	q	19/19 (100%)	0.86	2 (10%) 8 8	42, 53, 79, 80	0
3	s	12/19 (63%)	0.63	1 (8%) 14 15	51, 65, 86, 86	0
3	u	13/19 (68%)	1.37	3 (23%) 1 1	76, 84, 99, 124	0
3	w	12/19 (63%)	0.88	0 100 100	72, 83, 95, 105	0
All	All	5303/5556 (95%)	0.32	237 (4%) 37 42	24, 46, 88, 144	0

The worst 5 of 237 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	U	136	ALA	11.0
1	J	-1	ILE	7.9
1	S	215	SER	7.8
1	U	161	SER	7.5
2	V	214	CYS	6.6

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, 95^{th} 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(Å^2)	Q<0.9
4	NO3	V	301	4/4	0.80	0.61	13.87	37,39,39,39	4
4	NO3	K	301	4/4	0.95	0.30	7.08	41,42,42,47	0
4	NO3	P	301	4/4	0.92	0.30	6.77	30,31,31,32	4
4	NO3	D	301	4/4	0.98	0.27	6.36	41,45,45,47	0
4	NO3	B	301	4/4	0.84	0.30	6.03	59,59,63,66	0
5	EDO	Q	301	4/4	0.95	0.26	5.91	32,33,33,33	0
4	NO3	T	301	4/4	0.97	0.30	5.63	22,23,23,23	4
4	NO3	F	301	4/4	0.93	0.32	5.41	26,26,27,28	4
4	NO3	L	301	4/4	0.87	0.30	5.30	26,26,27,27	4
4	NO3	R	301	4/4	0.98	0.19	1.92	38,38,38,39	0
4	NO3	I	301	4/4	0.93	0.20	1.67	49,49,52,55	0
4	NO3	X	301	4/4	0.96	0.17	1.44	47,49,49,50	0
5	EDO	N	301	4/4	0.91	0.16	-0.28	38,39,39,40	0
4	NO3	N	302	4/4	0.98	0.14	-0.45	41,42,43,45	0
5	EDO	W	301	4/4	0.74	0.24	-	47,47,48,48	0

6.5 Other polymers

There are no such residues in this entry.