wwPDB X-ray Structure Validation Summary Report (i)

Jun 19, 2024 – 06:55 AM EDT

PDB ID	:	4G80
Title	:	Crystal structure of voltage sensing domain of Ci-VSP with fragment antibody (WT, 3.8 A)
Authors	:	Li, Q.
Deposited on	:	2012-07-20
Resolution	:	3.58 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

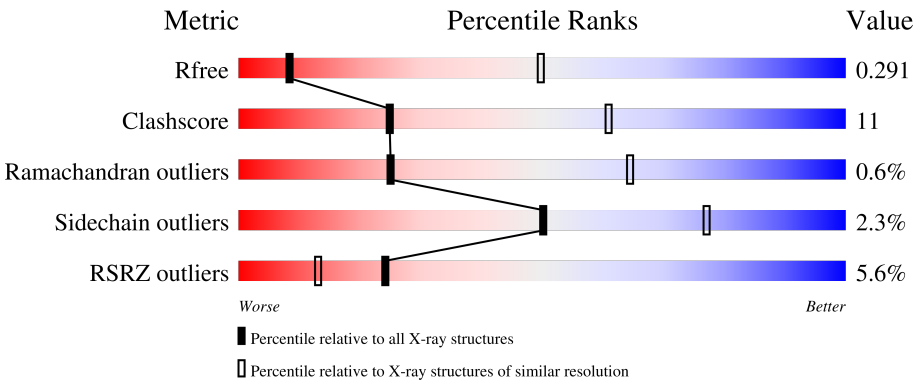
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.58 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}	130704	1094 (3.66-3.50)
Clashscore	141614	1181 (3.66-3.50)
Ramachandran outliers	138981	1143 (3.66-3.50)
Sidechain outliers	138945	1143 (3.66-3.50)
RSRZ outliers	127900	1012 (3.66-3.50)

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 of 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	I	155	12% 68% 21% • 10%
1	J	155	10% 72% 16% • 10%
1	S	155	6% 65% 17% • 17%
1	T	155	4% 70% 16% • • 10%
2	A	224	70% 28% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	224	6% 72% 25%
2	E	224	2% 73% 25%
2	G	224	14% 72% 26%
3	B	211	77% 22%
3	D	211	4% 78% 21%
3	F	211	2% 76% 24%
3	H	211	7% 80% 20%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 17446 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 Voltage-sensor containing phosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	S	128	1020	677	165	173	5	0	0	0
			Total	C	N	O	S			
1	T	139	1111	734	181	190	6	0	0	0
			Total	C	N	O	S			
1	I	139	1111	734	181	190	6	0	0	0
			Total	C	N	O	S			
1	J	139	1111	734	181	190	6	0	0	0

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	90	MET	-	EXPRESSION TAG	UNP Q4W8A1
S	91	ARG	-	EXPRESSION TAG	UNP Q4W8A1
S	92	GLY	-	EXPRESSION TAG	UNP Q4W8A1
S	93	SER	-	EXPRESSION TAG	UNP Q4W8A1
S	94	HIS	-	EXPRESSION TAG	UNP Q4W8A1
S	95	HIS	-	EXPRESSION TAG	UNP Q4W8A1
S	96	HIS	-	EXPRESSION TAG	UNP Q4W8A1
S	97	HIS	-	EXPRESSION TAG	UNP Q4W8A1
S	98	HIS	-	EXPRESSION TAG	UNP Q4W8A1
S	99	HIS	-	EXPRESSION TAG	UNP Q4W8A1
S	100	GLY	-	EXPRESSION TAG	UNP Q4W8A1
S	101	GLU	-	EXPRESSION TAG	UNP Q4W8A1
S	102	ASN	-	EXPRESSION TAG	UNP Q4W8A1
S	103	LEU	-	EXPRESSION TAG	UNP Q4W8A1
S	104	TYR	-	EXPRESSION TAG	UNP Q4W8A1
S	105	PHE	-	EXPRESSION TAG	UNP Q4W8A1
T	90	MET	-	EXPRESSION TAG	UNP Q4W8A1
T	91	ARG	-	EXPRESSION TAG	UNP Q4W8A1
T	92	GLY	-	EXPRESSION TAG	UNP Q4W8A1
T	93	SER	-	EXPRESSION TAG	UNP Q4W8A1
T	94	HIS	-	EXPRESSION TAG	UNP Q4W8A1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
T	95	HIS	-	EXPRESSION TAG	UNP Q4W8A1
T	96	HIS	-	EXPRESSION TAG	UNP Q4W8A1
T	97	HIS	-	EXPRESSION TAG	UNP Q4W8A1
T	98	HIS	-	EXPRESSION TAG	UNP Q4W8A1
T	99	HIS	-	EXPRESSION TAG	UNP Q4W8A1
T	100	GLY	-	EXPRESSION TAG	UNP Q4W8A1
T	101	GLU	-	EXPRESSION TAG	UNP Q4W8A1
T	102	ASN	-	EXPRESSION TAG	UNP Q4W8A1
T	103	LEU	-	EXPRESSION TAG	UNP Q4W8A1
T	104	TYR	-	EXPRESSION TAG	UNP Q4W8A1
T	105	PHE	-	EXPRESSION TAG	UNP Q4W8A1
I	90	MET	-	EXPRESSION TAG	UNP Q4W8A1
I	91	ARG	-	EXPRESSION TAG	UNP Q4W8A1
I	92	GLY	-	EXPRESSION TAG	UNP Q4W8A1
I	93	SER	-	EXPRESSION TAG	UNP Q4W8A1
I	94	HIS	-	EXPRESSION TAG	UNP Q4W8A1
I	95	HIS	-	EXPRESSION TAG	UNP Q4W8A1
I	96	HIS	-	EXPRESSION TAG	UNP Q4W8A1
I	97	HIS	-	EXPRESSION TAG	UNP Q4W8A1
I	98	HIS	-	EXPRESSION TAG	UNP Q4W8A1
I	99	HIS	-	EXPRESSION TAG	UNP Q4W8A1
I	100	GLY	-	EXPRESSION TAG	UNP Q4W8A1
I	101	GLU	-	EXPRESSION TAG	UNP Q4W8A1
I	102	ASN	-	EXPRESSION TAG	UNP Q4W8A1
I	103	LEU	-	EXPRESSION TAG	UNP Q4W8A1
I	104	TYR	-	EXPRESSION TAG	UNP Q4W8A1
I	105	PHE	-	EXPRESSION TAG	UNP Q4W8A1
J	90	MET	-	EXPRESSION TAG	UNP Q4W8A1
J	91	ARG	-	EXPRESSION TAG	UNP Q4W8A1
J	92	GLY	-	EXPRESSION TAG	UNP Q4W8A1
J	93	SER	-	EXPRESSION TAG	UNP Q4W8A1
J	94	HIS	-	EXPRESSION TAG	UNP Q4W8A1
J	95	HIS	-	EXPRESSION TAG	UNP Q4W8A1
J	96	HIS	-	EXPRESSION TAG	UNP Q4W8A1
J	97	HIS	-	EXPRESSION TAG	UNP Q4W8A1
J	98	HIS	-	EXPRESSION TAG	UNP Q4W8A1
J	99	HIS	-	EXPRESSION TAG	UNP Q4W8A1
J	100	GLY	-	EXPRESSION TAG	UNP Q4W8A1
J	101	GLU	-	EXPRESSION TAG	UNP Q4W8A1
J	102	ASN	-	EXPRESSION TAG	UNP Q4W8A1
J	103	LEU	-	EXPRESSION TAG	UNP Q4W8A1
J	104	TYR	-	EXPRESSION TAG	UNP Q4W8A1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	105	PHE	-	EXPRESSION TAG	UNP Q4W8A1

- Molecule 2 is a protein called fragment antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	221	1661	1052	276	327	6	0	0	0
2	C	219	1646	1044	274	322	6	0	0	0
2	E	221	1661	1052	276	327	6	0	0	0
2	G	221	1661	1052	276	327	6	0	0	0

- Molecule 3 is a protein called fragment antibody light chain.

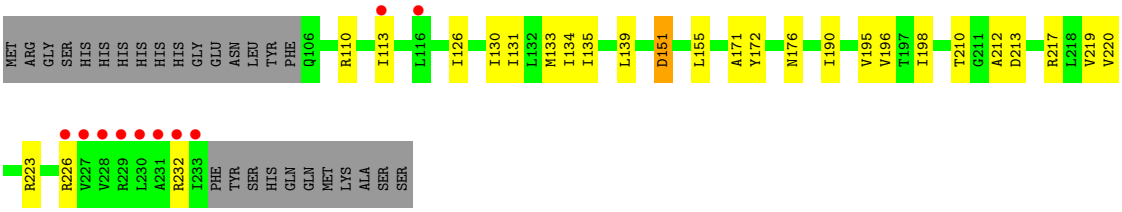
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	211	1616	1011	272	328	5	0	0	0
3	D	211	1616	1011	272	328	5	0	0	0
3	F	211	1616	1011	272	328	5	0	0	0
3	H	211	1616	1011	272	328	5	0	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

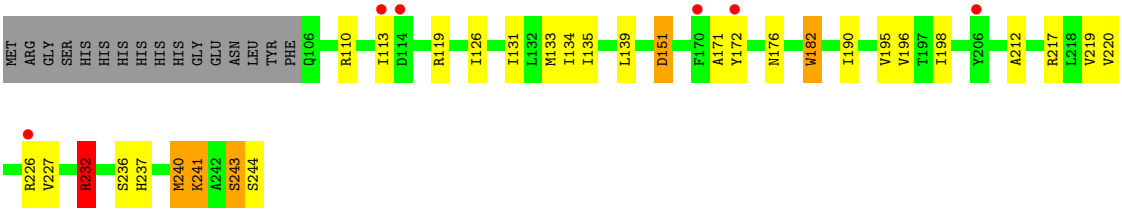
• Molecule 1: Voltage-sensor containing phosphatase

Chain S:



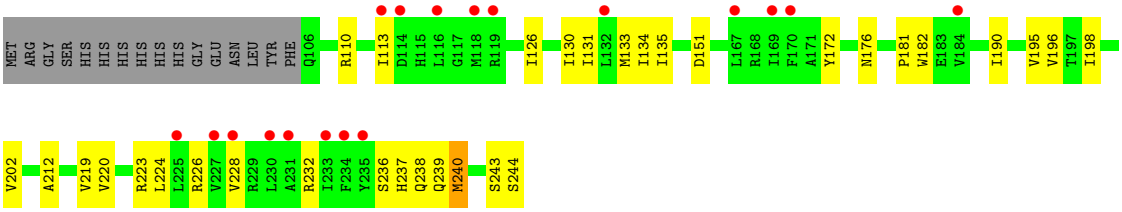
• Molecule 1: Voltage-sensor containing phosphatase

Chain T:



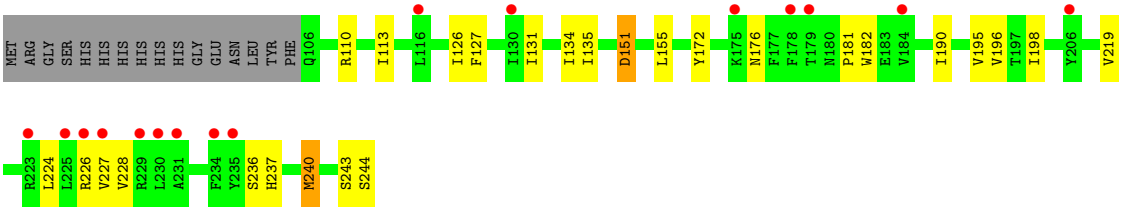
• Molecule 1: Voltage-sensor containing phosphatase

Chain I:



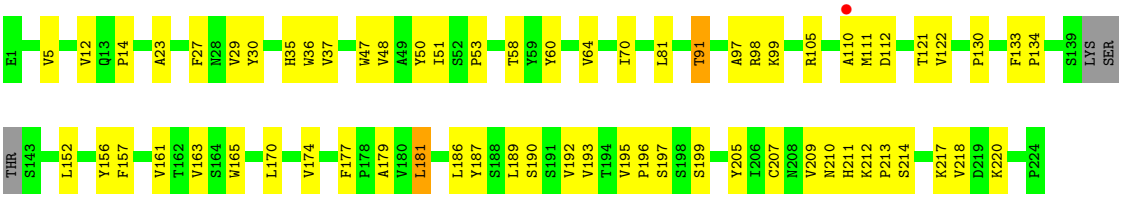
• Molecule 1: Voltage-sensor containing phosphatase

Chain J:



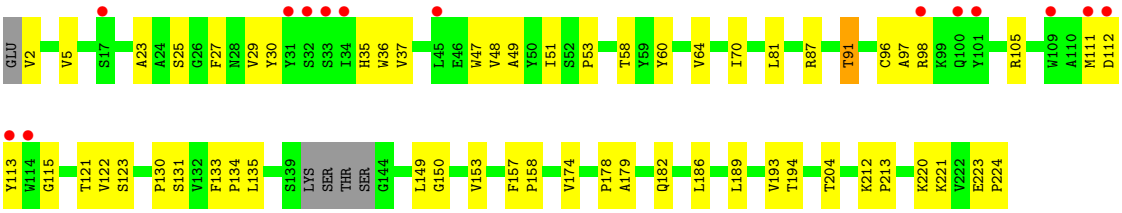
• Molecule 2: fragment antibody heavy chain

Chain A:



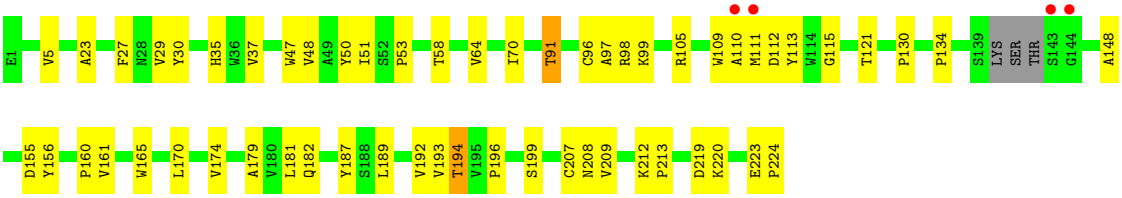
• Molecule 2: fragment antibody heavy chain

Chain C:



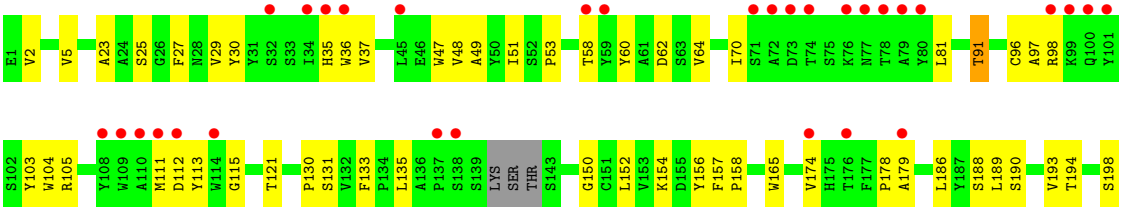
• Molecule 2: fragment antibody heavy chain

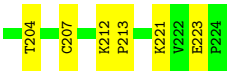
Chain E:



• Molecule 2: fragment antibody heavy chain

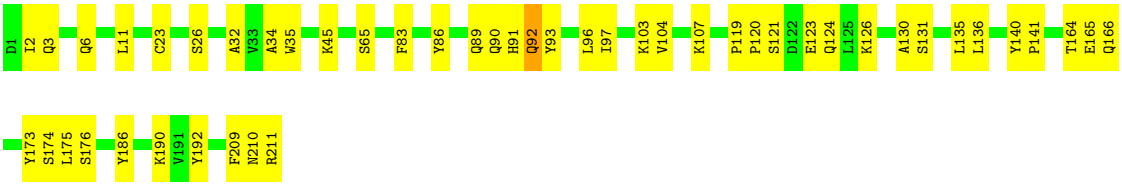
Chain G:





• Molecule 3: fragment antibody light chain

Chain B:



• Molecule 3: fragment antibody light chain

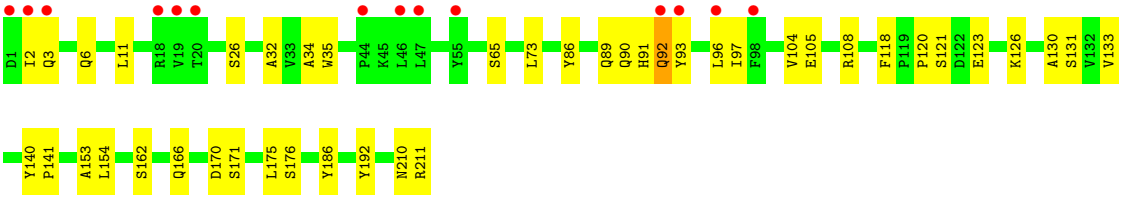
Chain D:

• Molecule 3: fragment antibody light chain

Chain F:

• Molecule 3: fragment antibody light chain

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	77.29Å 94.24Å 193.95Å 102.63° 93.45° 105.25°	Depositor
Resolution (Å)	50.00 – 3.58 45.93 – 3.58	Depositor EDS
% Data completeness (in resolution range)	92.6 (50.00-3.58) 92.6 (45.93-3.58)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	1.52 (at 3.57Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.248 , 0.292 0.246 , 0.291	Depositor DCC
R_{free} test set	2803 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	130.0	Xtriage
Anisotropy	0.142	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 70.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	17446	wwPDB-VP
Average B, all atoms (Å^2)	149.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.27% 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	I	0.34	0/1135	0.34	0/1536
1	J	0.35	0/1135	0.36	0/1536
1	S	0.40	1/1041 (0.1%)	0.37	0/1411
1	T	0.89	3/1135 (0.3%)	0.55	3/1536 (0.2%)
2	A	0.45	0/1705	0.39	0/2326
2	C	0.45	0/1690	0.39	0/2306
2	E	0.45	0/1705	0.38	0/2326
2	G	0.44	0/1705	0.37	0/2326
3	B	0.35	0/1650	0.37	0/2240
3	D	0.34	0/1650	0.37	0/2240
3	F	0.34	0/1650	0.36	0/2240
3	H	0.34	0/1650	0.36	0/2240
All	All	0.44	4/17851 (0.0%)	0.38	3/24263 (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	T	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	T	232	ARG	CZ-NH2	18.55	1.57	1.33
1	T	232	ARG	NE-CZ	17.79	1.56	1.33
1	T	232	ARG	CD-NE	11.31	1.65	1.46
1	S	232	ARG	CZ-NH2	7.54	1.42	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	T	232	ARG	NE-CZ-NH2	14.69	127.64	120.30
1	T	232	ARG	CD-NE-CZ	5.25	130.95	123.60
1	T	232	ARG	NH1-CZ-NH2	-5.06	113.84	119.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	T	232	ARG	Sidechain

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	I	1111	0	1140	25	0
1	J	1111	0	1140	16	0
1	S	1020	0	1057	16	0
1	T	1111	0	1140	38	0
2	A	1661	0	1598	55	0
2	C	1646	0	1584	41	0
2	E	1661	0	1598	38	0
2	G	1661	0	1598	39	0
3	B	1616	0	1576	37	0
3	D	1616	0	1576	37	0
3	F	1616	0	1576	37	0
3	H	1616	0	1576	33	0
All	All	17446	0	17159	375	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 11.

The worst 5 of 375 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:T:182:TRP:CH2	1:T:240:MET:HA	1.82	1.12
1:T:182:TRP:CZ3	1:T:240:MET:HA	1.89	1.07

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:182:TRP:HZ3	1:I:243:SER:HG	1.01	0.99
1:J:237:HIS:O	1:J:240:MET:HG3	1.65	0.95
1:T:182:TRP:CE3	1:T:240:MET:HB2	2.11	0.85

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	I	137/155 (88%)	126 (92%)	10 (7%)	1 (1%)	22	62
1	J	137/155 (88%)	126 (92%)	11 (8%)	0	100	100
1	S	126/155 (81%)	115 (91%)	10 (8%)	1 (1%)	19	59
1	T	137/155 (88%)	126 (92%)	10 (7%)	1 (1%)	22	62
2	A	217/224 (97%)	205 (94%)	11 (5%)	1 (0%)	29	67
2	C	215/224 (96%)	200 (93%)	14 (6%)	1 (0%)	29	67
2	E	217/224 (97%)	209 (96%)	5 (2%)	3 (1%)	11	48
2	G	217/224 (97%)	208 (96%)	8 (4%)	1 (0%)	29	67
3	B	209/211 (99%)	194 (93%)	14 (7%)	1 (0%)	29	67
3	D	209/211 (99%)	195 (93%)	13 (6%)	1 (0%)	29	67
3	F	209/211 (99%)	195 (93%)	12 (6%)	2 (1%)	15	55
3	H	209/211 (99%)	195 (93%)	13 (6%)	1 (0%)	29	67
All	All	2239/2360 (95%)	2094 (94%)	131 (6%)	14 (1%)	25	64

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	95	SER
1	S	212	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	155	ASP
3	F	92	GLN
2	A	91	THR

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	I	120/134 (90%)	115 (96%)	5 (4%)	30	63
1	J	120/134 (90%)	114 (95%)	6 (5%)	24	59
1	S	110/134 (82%)	107 (97%)	3 (3%)	44	74
1	T	120/134 (90%)	110 (92%)	10 (8%)	11	42
2	A	182/187 (97%)	179 (98%)	3 (2%)	62	83
2	C	180/187 (96%)	177 (98%)	3 (2%)	60	83
2	E	182/187 (97%)	180 (99%)	2 (1%)	73	88
2	G	182/187 (97%)	178 (98%)	4 (2%)	52	78
3	B	185/185 (100%)	183 (99%)	2 (1%)	73	88
3	D	185/185 (100%)	182 (98%)	3 (2%)	62	83
3	F	185/185 (100%)	183 (99%)	2 (1%)	73	88
3	H	185/185 (100%)	184 (100%)	1 (0%)	88	95
All	All	1936/2024 (96%)	1892 (98%)	44 (2%)	50	77

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	62	ASP
1	I	151	ASP
2	G	131	SER
3	H	65	SER
1	I	240	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such

sidechains are listed below:

Mol	Chain	Res	Type
3	F	94	ASN
3	F	160	GLN
1	J	237	HIS
2	G	182	GLN
2	E	182	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 monosaccharides 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(Å^2)	Q<0.9
1	I	139/155 (89%)	0.45	18 (12%)	3	2	160, 217, 278, 289	0
1	J	139/155 (89%)	0.15	16 (11%)	4	3	146, 202, 271, 292	0
1	S	128/155 (82%)	0.23	10 (7%)	13	7	95, 150, 248, 278	0
1	T	139/155 (89%)	0.05	6 (4%)	35	21	106, 157, 255, 293	0
2	A	221/224 (98%)	-0.09	1 (0%)	91	82	70, 104, 159, 214	0
2	C	219/224 (97%)	0.10	14 (6%)	19	10	66, 135, 208, 261	0
2	E	221/224 (98%)	-0.04	4 (1%)	68	51	68, 113, 171, 223	0
2	G	221/224 (98%)	0.48	31 (14%)	2	2	104, 170, 228, 263	0
3	B	211/211 (100%)	-0.18	0	100	100	67, 112, 146, 185	0
3	D	211/211 (100%)	-0.05	9 (4%)	35	21	70, 125, 186, 237	0
3	F	211/211 (100%)	-0.03	4 (1%)	66	49	90, 135, 178, 201	0
3	H	211/211 (100%)	0.12	14 (6%)	18	9	91, 157, 212, 269	0
All	All	2271/2360 (96%)	0.08	127 (5%)	24	13	66, 140, 239, 293	0

The worst 5 of 127 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	110	ALA	9.1
1	T	170	PHE	8.1
3	H	2	ILE	7.4
2	C	32	SER	7.1
1	I	234	PHE	7.0

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 monosaccharides 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.