wwPDB X-ray Structure Validation Summary Report (i)

Oct 6, 2024 – 07:00 AM EDT

PDB ID	:	2H2P
Title	:	Crystal structure of CLC-ec1 in complex with Fab fragment in SeCN-
Authors	:	Nguitragool, W.; Miller, C.
Deposited on	:	2006-05-19
Resolution	:	3.10 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 (i)) were used in the production of this report:

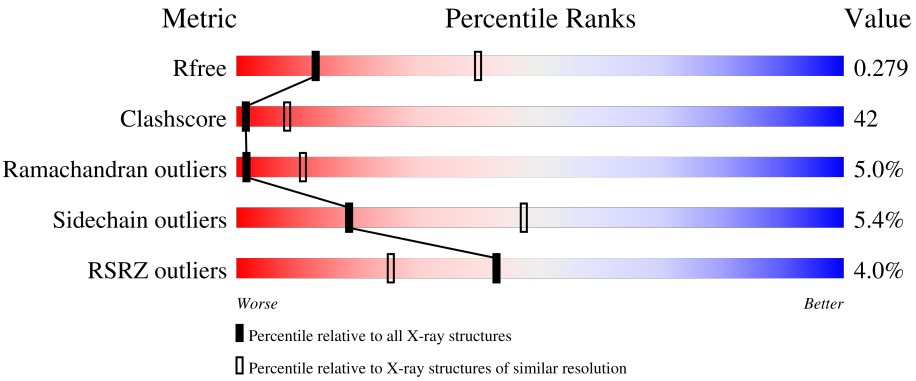
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.10 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.



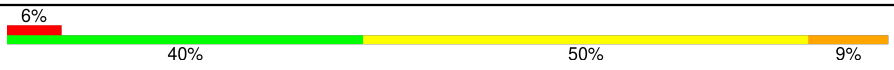
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1351 (3.10-3.10)
Clashscore	180529	1454 (3.10-3.10)
Ramachandran outliers	177936	1391 (3.10-3.10)
Sidechain outliers	177891	1391 (3.10-3.10)
RSRZ outliers	164620	1351 (3.10-3.10)

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	A	465	3% 46% 47% • 5%
1	B	465	4% 42% 49% • 5%
2	C	221	6% 51% 44% 5%
2	E	221	3% 52% 39% 8%
3	D	211	2% 29% 61% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	211	

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	SEK	A	466	-	-	X	-
4	SEK	B	466	-	-	X	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 13231 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 CLC Cl transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	444	Total	C	N	O	S	0	0	0
			3333	2190	560	563	20			
1	B	441	Total	C	N	O	S	0	0	0
			3304	2174	553	557	20			

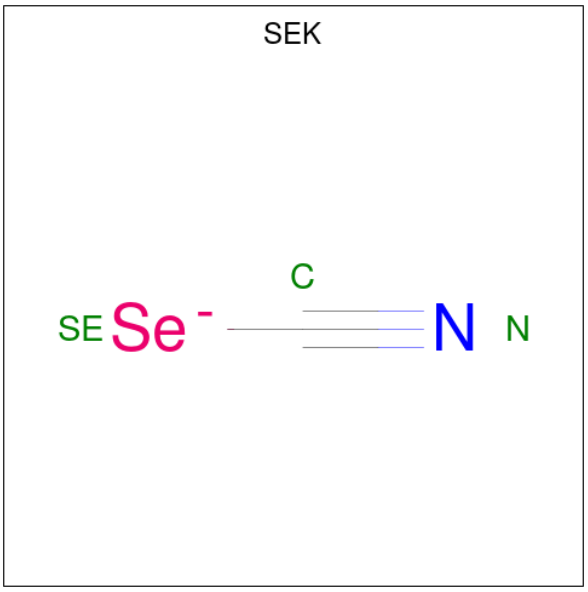
- Molecule 2 is a protein called FAB fragment, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	221	Total	C	N	O	S	0	0	0
			1672	1077	274	315	6			
2	E	221	Total	C	N	O	S	0	0	0
			1672	1077	274	315	6			

- Molecule 3 is a protein called FAB fragment, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			
3	F	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			

- Molecule 4 is SELENOCYANATE ION (three-letter code: SEK) (formula: CNSe).



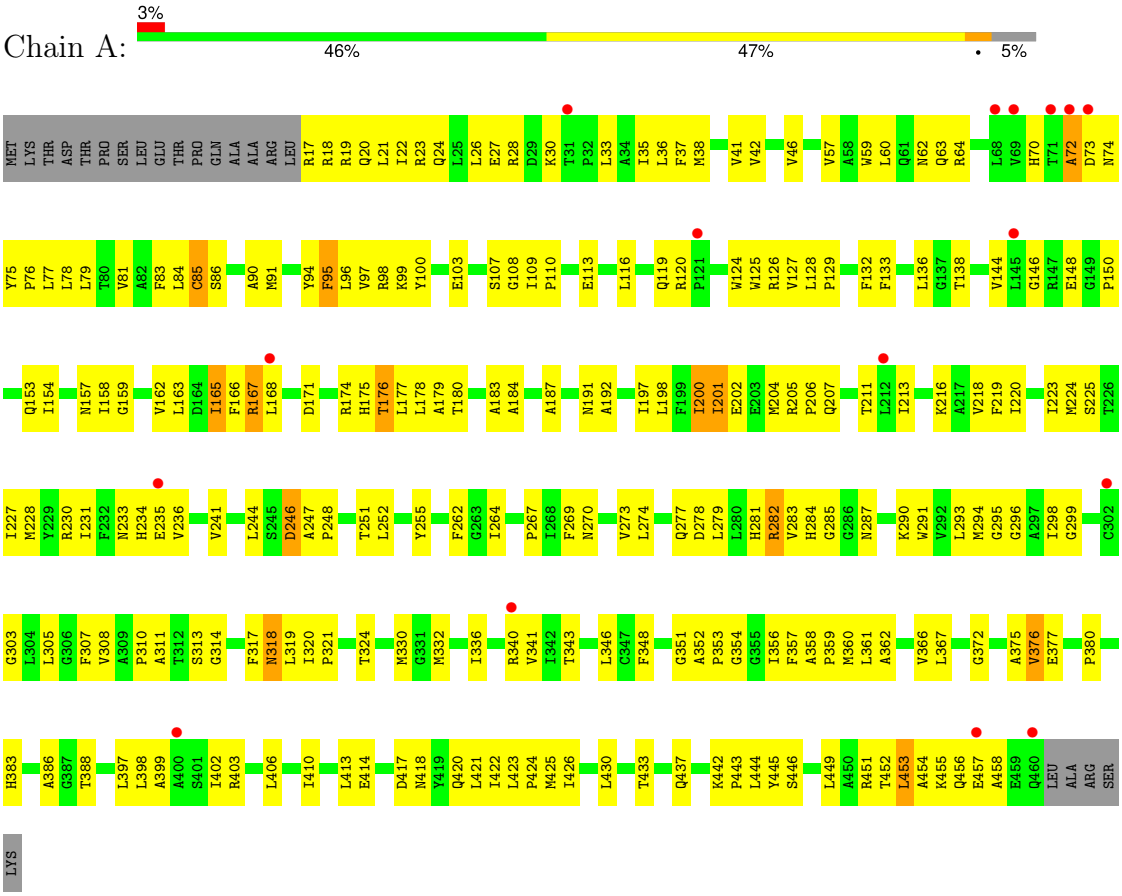
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Se 1 1	0	0
4	A	1	Total Se 1 1	1	0
4	A	1	Total Se 1 1	1	0
4	A	1	Total Se 1 1	1	0
4	B	1	Total Se 1 1	0	0
4	B	1	Total Se 1 1	1	0
4	B	1	Total Se 1 1	1	0
4	B	1	Total Se 1 1	1	0

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.

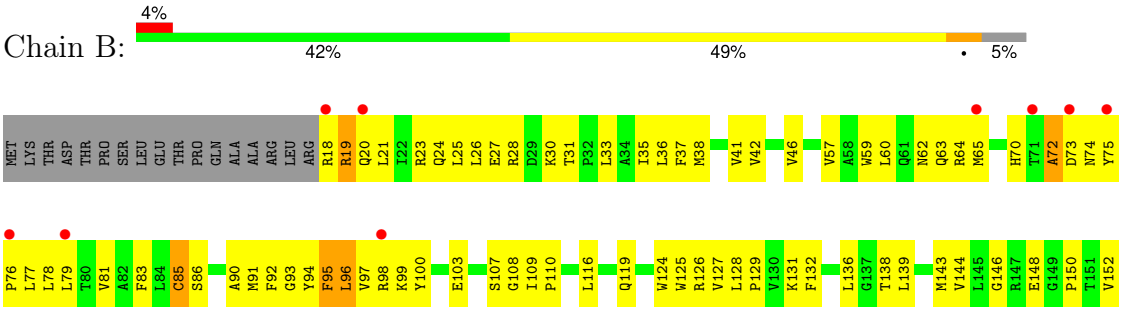
• Molecule 1: CLC Cl transporter

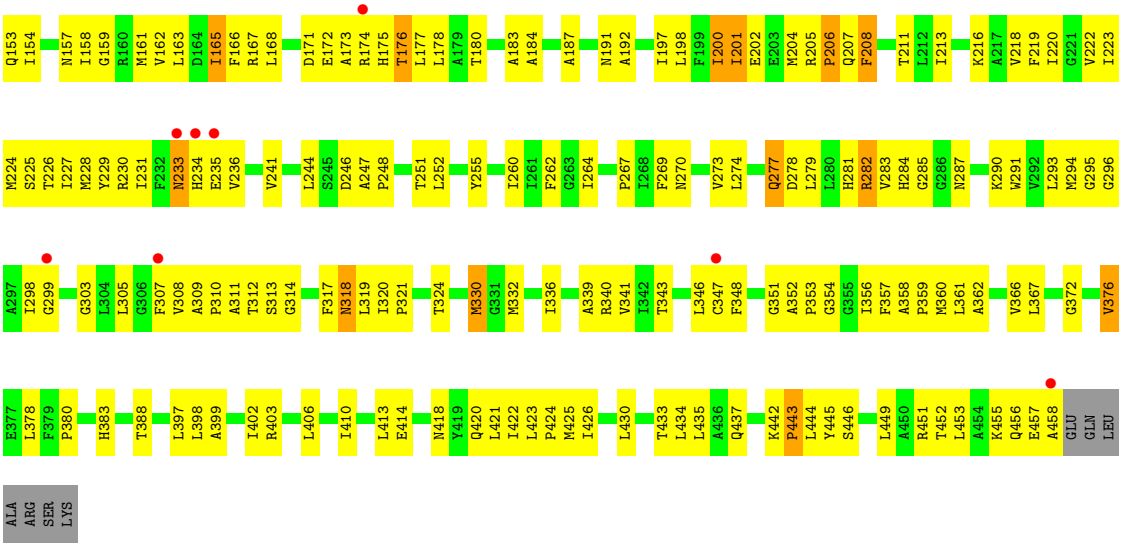
Chain A:



• Molecule 1: CLC Cl transporter

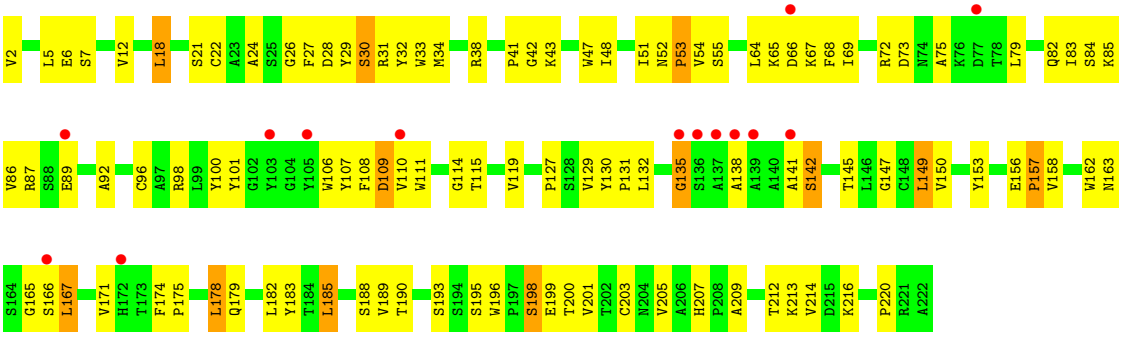
Chain B:





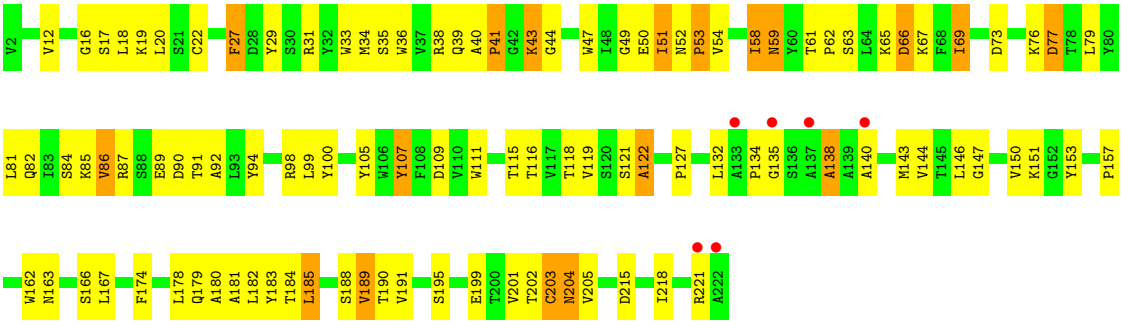
• Molecule 2: FAB fragment, heavy chain

Chain C:



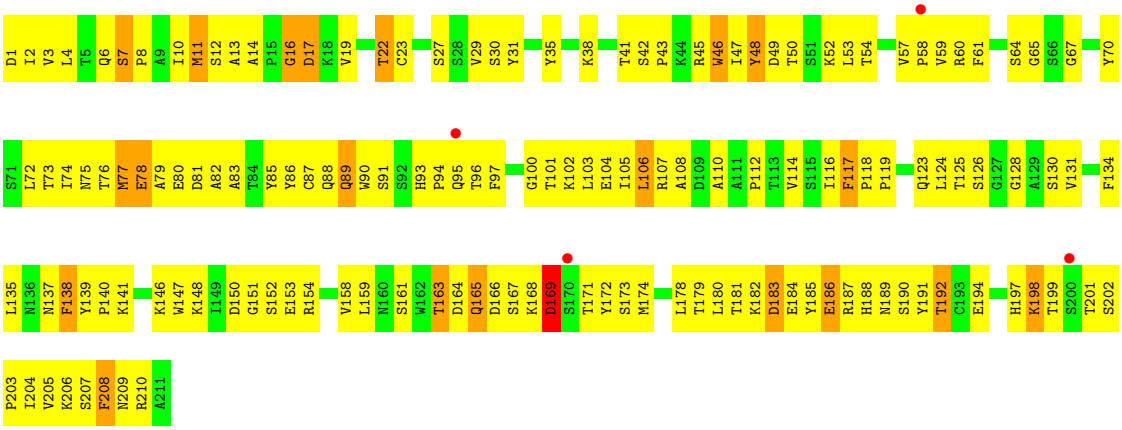
• Molecule 2: FAB fragment, heavy chain

Chain E:



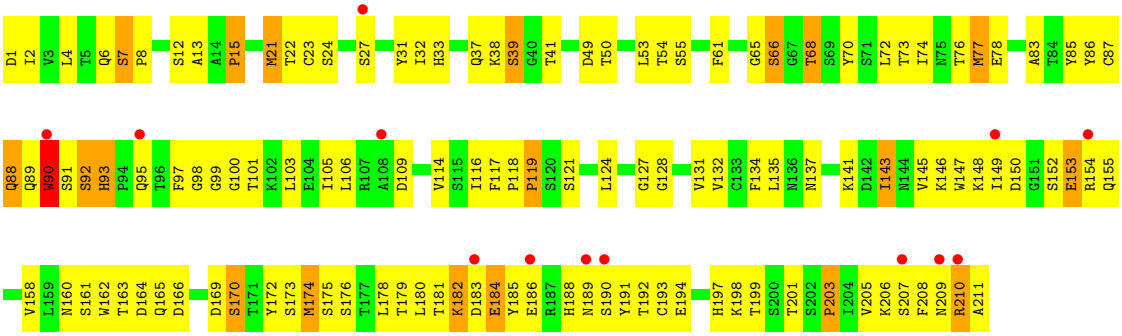
• Molecule 3: FAB fragment, light chain

Chain D:



• Molecule 3: FAB fragment, light chain

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	220.58Å 121.58Å 151.24Å 90.00° 128.25° 90.00°	Depositor
Resolution (Å)	50.00 – 3.10 50.00 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.1 (50.00-3.10) 99.0 (50.00-3.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.08 (at 3.01Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.278 , 0.280 0.274 , 0.279	Depositor DCC
R_{free} test set	2835 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	90.1	Xtriage
Anisotropy	0.676	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 31.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	13231	wwPDB-VP
Average B, all atoms (Å^2)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.25% 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

5.1 Standard geometry

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

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	0.48	0/3405	0.64	0/4621
1	B	0.48	0/3376	0.63	0/4583
2	C	0.52	0/1721	0.76	0/2355
2	E	0.52	0/1721	0.78	0/2355
3	D	0.48	0/1660	0.73	1/2257 (0.0%)
3	F	0.52	0/1660	0.77	0/2257
All	All	0.49	0/13543	0.70	1/18428 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	16	GLY	N-CA-C	-6.80	96.11	113.10

There are no chirality outliers.

There are no planarity outliers.

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	3333	0	3482	267	0
1	B	3304	0	3456	274	0
2	C	1672	0	1654	122	0
2	E	1672	0	1654	125	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	1621	0	1544	230	0
3	F	1621	0	1546	172	0
4	A	4	0	0	2	0
4	B	4	0	0	2	0
All	All	13231	0	13336	1109	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 42.

The worst 5 of 1109 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)
2:C:107:TYR:CE1	3:D:45:ARG:NH1	1.70	1.56
1:A:274:LEU:CD2	1:A:451:ARG:NH1	1.71	1.53
3:D:186:GLU:HA	3:D:210:ARG:NH2	1.27	1.45
2:E:163:ASN:HD22	2:E:167:LEU:CD1	1.35	1.39
1:A:18:ARG:NH1	1:B:456:GLN:OE1	1.62	1.30

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	442/465 (95%)	344 (78%)	80 (18%)	18 (4%)	2 13
1	B	439/465 (94%)	336 (76%)	85 (19%)	18 (4%)	2 13
2	C	219/221 (99%)	187 (85%)	22 (10%)	10 (5%)	2 12
2	E	219/221 (99%)	184 (84%)	24 (11%)	11 (5%)	1 11
3	D	209/211 (99%)	153 (73%)	40 (19%)	16 (8%)	1 4
3	F	209/211 (99%)	168 (80%)	28 (13%)	13 (6%)	1 7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1737/1794 (97%)	1372 (79%)	279 (16%)	86 (5%)	1 11

5 of 86 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	201	ILE
1	A	283	VAL
1	B	201	ILE
1	B	283	VAL
2	C	53	PRO

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	335/353 (95%)	328 (98%)	7 (2%)	48 72
1	B	332/353 (94%)	323 (97%)	9 (3%)	40 67
2	C	181/181 (100%)	172 (95%)	9 (5%)	20 50
2	E	181/181 (100%)	165 (91%)	16 (9%)	8 30
3	D	185/185 (100%)	168 (91%)	17 (9%)	7 28
3	F	185/185 (100%)	168 (91%)	17 (9%)	7 28
All	All	1399/1438 (97%)	1324 (95%)	75 (5%)	18 47

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	204	ASN
3	F	173	SER
3	F	15	PRO
3	F	119	PRO
2	C	203	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	157	ASN
2	E	39	GLN
1	B	284	HIS
3	F	6	GLN
3	D	6	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry

Of 8 ligands modelled in this entry, 8 are modelled with single atom - leaving 0 for Mogul analysis.

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

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	444/465 (95%)	0.44	16 (3%)	46	27	77, 77, 77, 77	0
1	B	441/465 (94%)	0.42	17 (3%)	44	26	77, 77, 77, 77	0
2	C	221/221 (100%)	0.61	14 (6%)	27	16	77, 77, 77, 77	0
2	E	221/221 (100%)	0.58	6 (2%)	56	36	77, 77, 77, 77	0
3	D	211/211 (100%)	0.47	4 (1%)	66	47	77, 77, 77, 77	0
3	F	211/211 (100%)	0.79	13 (6%)	28	17	77, 77, 77, 77	0
All	All	1749/1794 (97%)	0.52	70 (4%)	43	25	77, 77, 77, 77	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	72	ALA	5.4
1	A	235	GLU	5.2
3	D	95	GLN	5.1
3	F	190	SER	4.3
2	C	139	ALA	4.3

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

LIGAND-RSR INFOmissingINFO

6.5 Other polymers (i)

There are no such residues in this entry.