WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2023 – 11:25 PM EDT

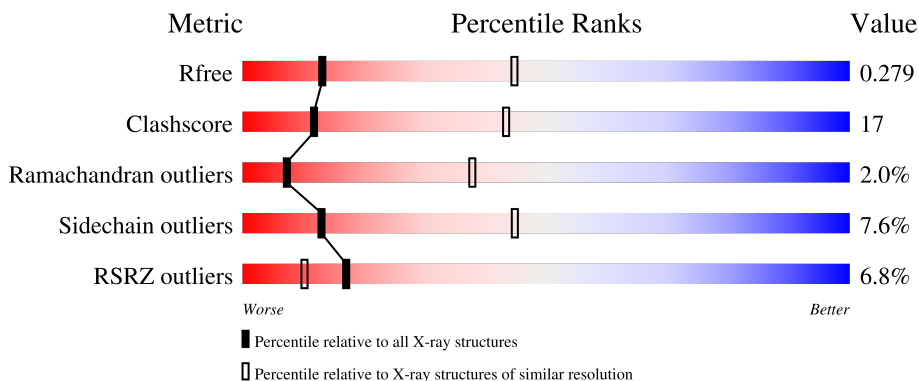|		
PDB ID	:	3EJY
Title	:	Structure of E203H mutant of E.coli Cl-/H+ antiporter, CLC-ec1
Authors	:	Lim, H.-H.; Miller, C.
Deposited on	:	2008-09-18
Resolution	:	3.20 Å(reported)

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 3.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
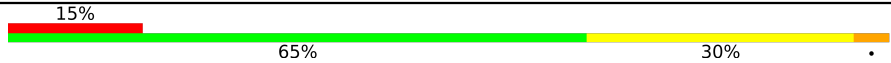
Metric	Percentile Ranks	Value
Rfree		0.279
Clashscore		17
Ramachandran outliers		2.0%
Sidechain outliers		7.6%
RSRZ outliers		6.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	473	6% — 62% / 28% / • / 6%
1	B	473	6% — 57% / 33% / • / 7%
2	C	221	2% — 62% / 30% / 6% / •
2	E	221	4% — 71% / 23% / 5%
3	D	211	8% — 55% / 43% / •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	211	15% 65% 30%

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	BR	B	503	-	-	X	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 13218 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 H(+)/Cl(-) exchange transporter clcA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	444	Total	C	N	O	S	0	0	0
			3334	2191	562	561	20			
1	B	440	Total	C	N	O	S	0	0	0
			3294	2169	551	554	20			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	HIS	GLU	engineered mutation	UNP P37019
B	203	HIS	GLU	engineered mutation	UNP P37019

- 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 BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Br	0	0
			2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Br		
4	B	2	2	2	0	0

3   Residue-property plots ⓘ

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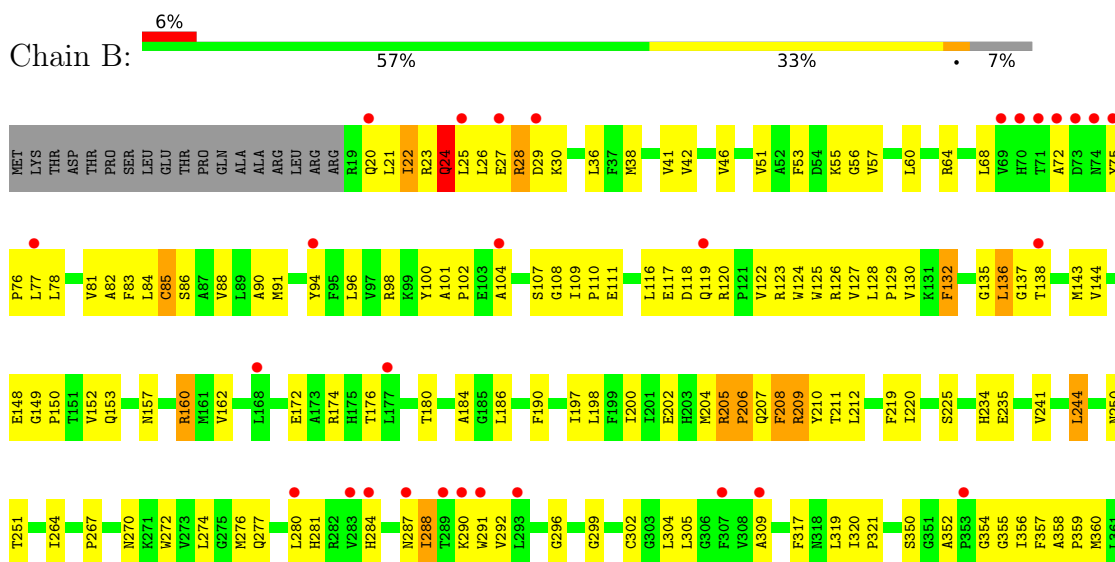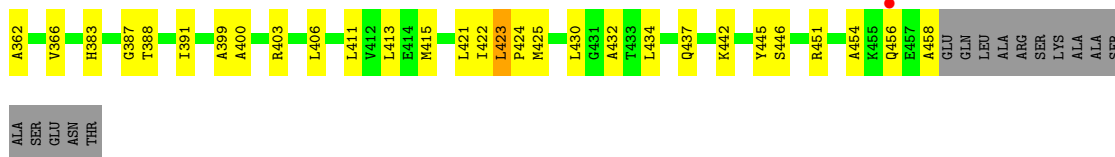 classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 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.
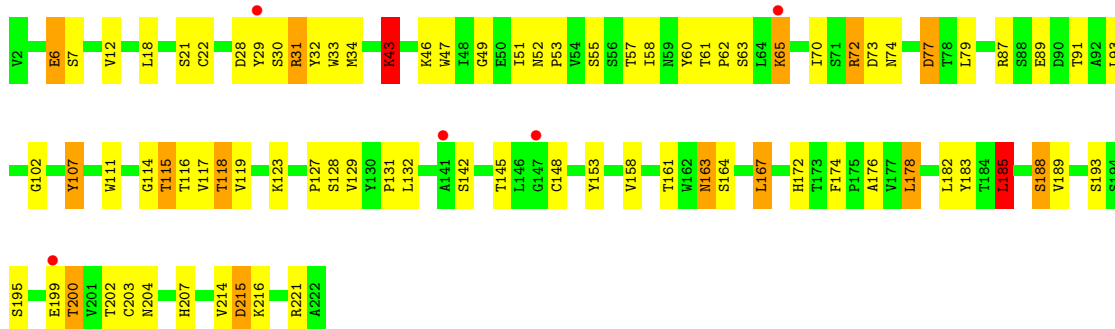
• Molecule 1: H(+)/Cl(-) exchange transporter clcA
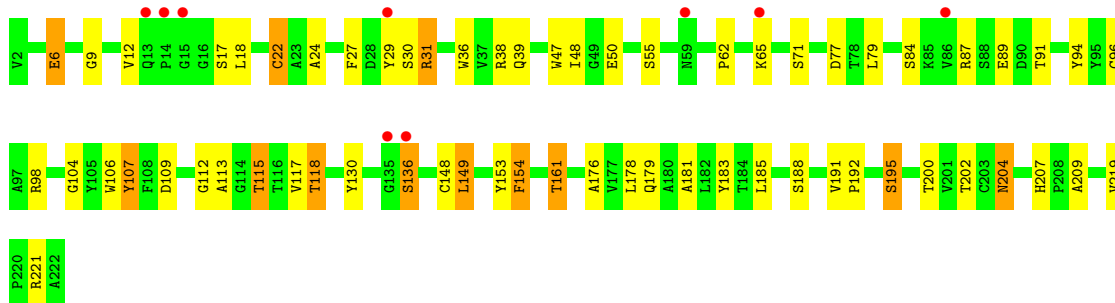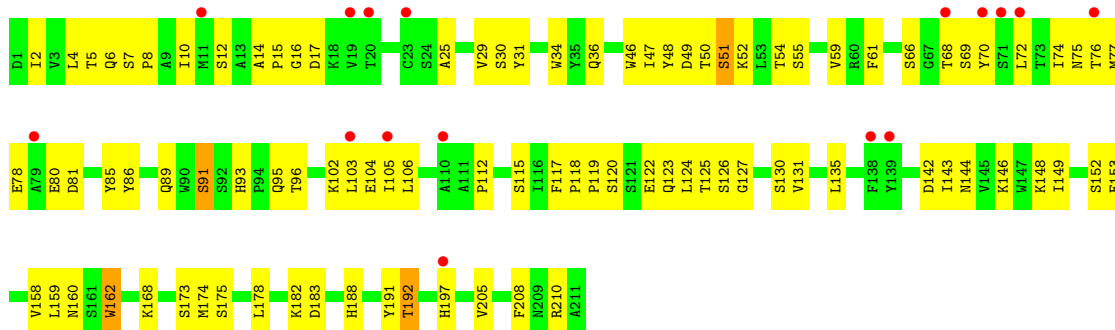

• Molecule 1: H(+)/Cl(-) exchange transporter clcA

A362 V366 H383 G387 T388 I391 A399 A400 R403 L406 L411 V412 L413 E414 M415 L421 I422 L423 P424 M425 L430 G431 A432 T433 L434 Q437 K442 Y445 S446 R451 A454 K455 Q456 E457 A458 GLU GLN LEU ALA ARG SER LYS ALA ALA SER

ALA SER GLU ASN THR

• Molecule 2: Fab fragment, Heavy chain

Chain C:

V2 E6 S7 V12 L18 S21 C22 D28 Y29 S30 R31 Y32 W33 M34 K43 K46 W47 I48 G49 E50 I51 N52 P53 V54 S55 S56 T57 I58 N59 Y60 T61 P62 S63 L64 K65 I70 S71 R72 D73 N74 D77 T78 L79 R87 S88 E89 D90 T91 A92 L93

G102 Y107 W111 G114 T115 T116 V117 T118 V119 K123 P127 S128 V129 Y130 P131 L132 A141 S142 T145 L146 G147 C148 Y153 V158 T161 W162 N163 S164 L167 H172 T173 F174 P175 A176 V177 L178 L182 Y183 T184 L185 S188 V189 S193 S194

S195 E199 T200 V201 T202 C203 N204 H207 V214 D215 K216 R221 A222

• Molecule 2: Fab fragment, Heavy chain

Chain E:

V2 E6 G9 V12 Q13 P14 G15 G16 S17 L18 C22 A23 A24 F27 D28 Y29 S30 R31 W36 V37 R38 Q39 W47 I48 G49 E50 S55 N59 P62 K65 S71 D77 T78 L79 S84 K85 V86 R87 S88 E89 D90 T91 Y94 Y95 C96

A97 R98 G104 Y105 W106 Y107 F108 D109 G112 A113 G114 T115 T116 V117 T118 Y130 G135 S136 C148 L149 Y153 F154 T161 A176 V177 L178 Q179 A180 A181 L182 Y183 T184 L185 S188 V191 P192 S195 T200 V201 T202 C203 N204 H207 P208 A209 V219

P220 R221 A222

• Molecule 3: Fab fragment, Light chain

Chain D:

D1 I2 V3 L4 T5 Q6 S7 P8 A9 I10 M11 S12 A13 A14 P15 G16 D17 K18 V19 T20 C23 S24 A25 V29 S30 Y31 W34 Y35 Q36 W46 I47 Y48 D49 T50 S51 K52 L53 T54 S55 V59 R60 F61 S66 G67 T68 S69 Y70 S71 L72 T73 I74 N75 T76 M77

E78 A79 E80 D81 Y85 Y86 Q89 W90 S91 S92 P94 Q95 T96 K102 L103 E104 I105 L106 A110 A111 P112 S115 I116 F117 P118 P119 S120 S121 Q123 L124 T125 S126 G127 S130 V131 L135 F138 Y139 D142 I143 N144 V145 K146 W147 K148 I149 S152 E153

V158 L159 N160 S161 W162 K168 S173 M174 S175 L178 K182 D183 H188 Y191 T192 H197 V205 F208 N209 R210 A211

• Molecule 3: Fab fragment, Light chain

Chain F:

15%

65%

30%

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	231.52Å 96.83Å 172.79Å 90.00° 132.81° 90.00°	Depositor
Resolution (Å)	58.52 – 3.20 57.87 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.5 (58.52-3.20) 99.6 (57.87-3.20)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.69 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.5	Depositor
R, R_{free}	0.244 , 0.279 0.237 , 0.279	Depositor DCC
R_{free} test set	2358 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	92.7	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 66.3	EDS
L-test for twinning [2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.008 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	13218	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% 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: BR

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	0/3407	0.59	1/4624 (0.0%)
1	B	0.49	0/3367	0.59	1/4572 (0.0%)
2	C	0.55	0/1721	0.67	1/2355 (0.0%)
2	E	0.51	0/1721	0.67	3/2355 (0.1%)
3	D	0.45	0/1660	0.61	0/2257
3	F	0.48	0/1660	0.63	0/2257
All	All	0.49	0/13536	0.62	6/18420 (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	B	0	1
2	E	1	0
All	All	1	1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	24	GLN	N-CA-C	6.63	128.91	111.00
2	E	154	PHE	C-N-CD	6.48	142.01	128.40
2	E	154	PHE	N-CA-C	6.41	128.32	111.00
2	E	154	PHE	N-CA-CB	5.46	120.44	110.60
1	A	18	ARG	NE-CZ-NH2	-5.30	117.65	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	E	154	PHE	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	23	ARG	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	3334	0	3483	137	0
1	B	3294	0	3444	149	0
2	C	1672	0	1654	69	0
2	E	1672	0	1654	41	0
3	D	1621	0	1546	58	0
3	F	1621	0	1546	57	0
4	A	2	0	0	2	0
4	B	2	0	0	6	0
All	All	13218	0	13327	461	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 17.

The worst 5 of 461 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:7:SER:HB3	3:F:8:PRO:HD3	1.26	1.12
1:B:107:SER:OG	4:B:503:BR:BR	2.21	1.10
3:F:7:SER:HB2	3:F:22:THR:H	1.15	1.10
3:F:7:SER:HB3	3:F:8:PRO:CD	1.86	1.05
2:E:31:ARG:HH11	2:E:31:ARG:HG2	1.22	1.03

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	442/473 (93%)	381 (86%)	55 (12%)	6 (1%)	11	46
1	B	438/473 (93%)	380 (87%)	51 (12%)	7 (2%)	9	43
2	C	219/221 (99%)	192 (88%)	18 (8%)	9 (4%)	3	21
2	E	219/221 (99%)	188 (86%)	25 (11%)	6 (3%)	5	30
3	D	209/211 (99%)	179 (86%)	27 (13%)	3 (1%)	11	46
3	F	209/211 (99%)	181 (87%)	24 (12%)	4 (2%)	8	39
All	All	1736/1810 (96%)	1501 (86%)	200 (12%)	35 (2%)	7	38

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	65	LYS
2	C	164	SER
3	D	51	SER
3	F	7	SER
3	F	199	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	A	335/358 (94%)	314 (94%)	21 (6%)	18	52
1	B	331/358 (92%)	314 (95%)	17 (5%)	24	60
2	C	181/181 (100%)	161 (89%)	20 (11%)	6	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	181/181 (100%)	164 (91%)	17 (9%)	8	33
3	D	185/185 (100%)	167 (90%)	18 (10%)	8	31
3	F	185/185 (100%)	172 (93%)	13 (7%)	15	48
All	All	1398/1448 (96%)	1292 (92%)	106 (8%)	13	45

5 of 106 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	215	ASP
3	D	168	LYS
3	F	142	ASP
3	D	29	VAL
3	D	80	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	37	GLN
3	F	136	ASN
1	B	277	GLN
1	B	270	ASN
3	F	137	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - 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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	444/473 (93%)	0.49	27 (6%)	21	12	72, 93, 122, 139	0
1	B	440/473 (93%)	0.59	30 (6%)	17	10	73, 92, 122, 133	0
2	C	221/221 (100%)	0.25	5 (2%)	60	47	66, 85, 102, 114	0
2	E	221/221 (100%)	0.22	9 (4%)	37	24	63, 85, 106, 119	0
3	D	211/211 (100%)	0.61	16 (7%)	13	7	74, 95, 109, 113	0
3	F	211/211 (100%)	0.81	32 (15%)	2	1	59, 79, 113, 115	0
All	All	1748/1810 (96%)	0.50	119 (6%)	17	10	59, 89, 114, 139	0

The worst 5 of 119 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	73	ASP	6.4
1	B	71	THR	5.8
1	A	72	ALA	5.6
1	B	20	GLN	5.5
1	B	72	ALA	5.4

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 (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. 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	B-factors(Å^2)	Q<0.9
4	BR	A	502	1/1	0.78	0.40	100,100,100,100	0
4	BR	B	503	1/1	0.94	0.33	100,100,100,100	0
4	BR	B	504	1/1	0.94	0.35	100,100,100,100	0
4	BR	A	501	1/1	0.95	0.42	100,100,100,100	0

6.5 Other polymers (i)

There are no such residues in this entry.