wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 01:44 am BST

PDB ID	:	2XQ4
Title	:	Pentameric ligand gated ion channel GLIC in complex with tetramethylarsonium (TMAs)
Authors	:	Hilf, R.J.C.; Bertozzi, C.; Zimmermann, I.; Reiter, A.; Trauner, D.; Dutzler, R.
Deposited on	:	2010-09-01
Resolution	:	3.60 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 following versions of software and data (see references (1)) were used in the production of this report:

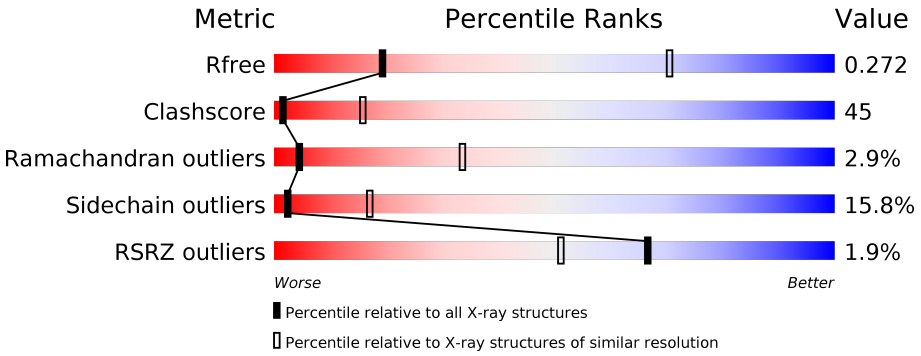
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 on 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	317	2% 34% 52% 11%
1	B	317	3% 32% 54% 11%
1	C	317	2% 35% 51% 11%
1	D	317	33% 53% 12%
1	E	317	3% 36% 51% 10%

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
2	ARS	C	1317	-	-	-	X

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12606 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 GLR4197 PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	310	Total	C	N	O	S	0	0	0
			2521	1662	403	452	4			
1	B	310	Total	C	N	O	S	0	0	0
			2521	1662	403	452	4			
1	C	310	Total	C	N	O	S	0	0	0
			2521	1662	403	452	4			
1	D	310	Total	C	N	O	S	0	0	0
			2521	1662	403	452	4			
1	E	310	Total	C	N	O	S	0	0	0
			2521	1662	403	452	4			

- Molecule 2 is ARSENIC (three-letter code: ARS) (formula: As).

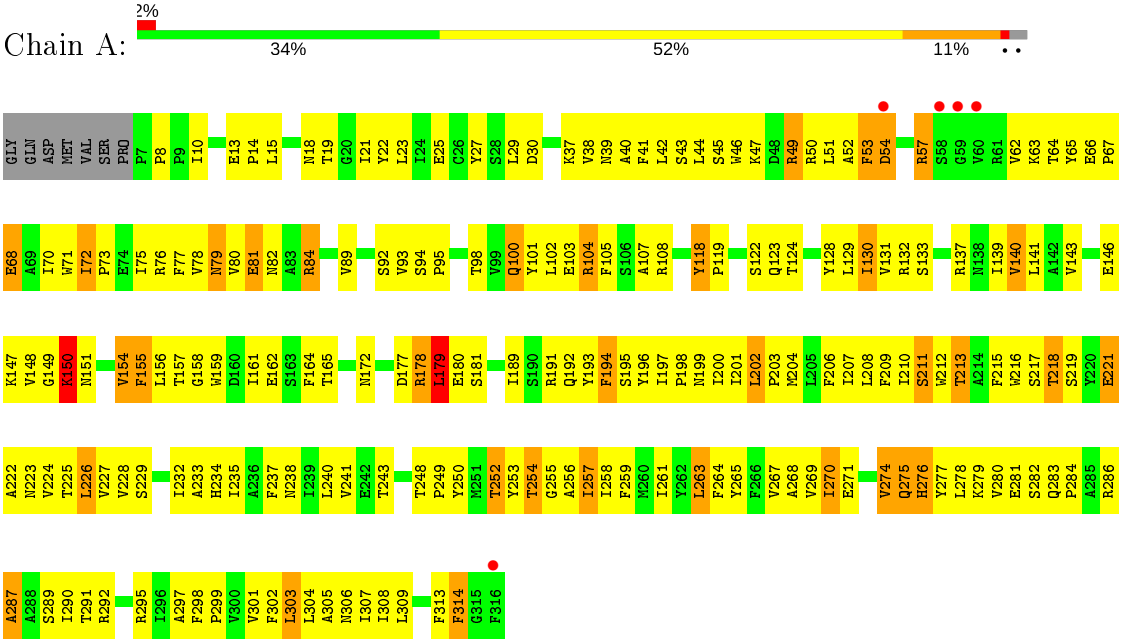
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total	As	0	0
			1	1		

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 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.

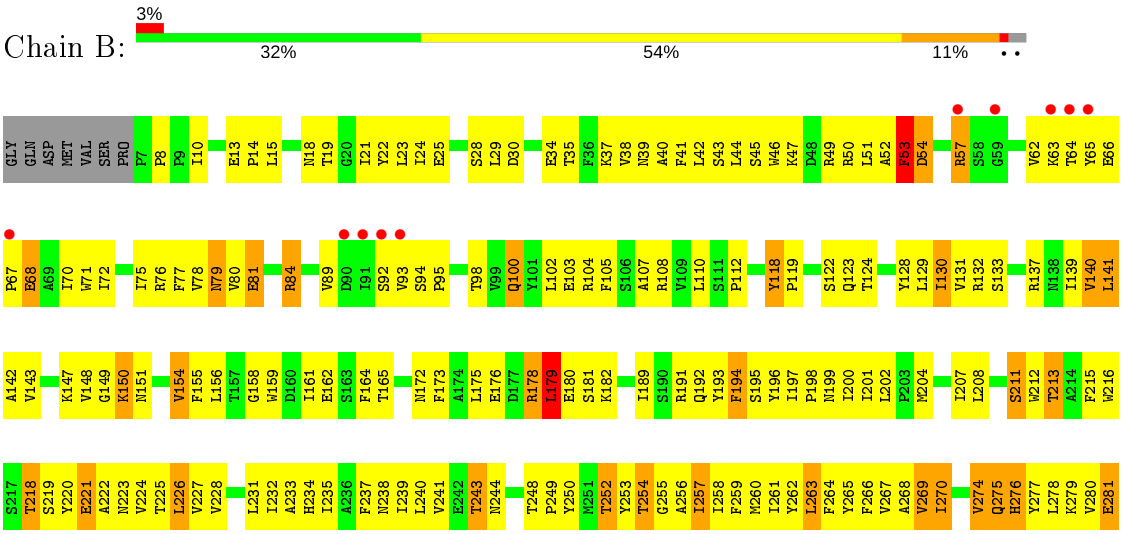
- Molecule 1: GLR4197 PROTEIN

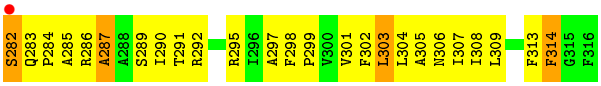
Chain A:



- Molecule 1: GLR4197 PROTEIN

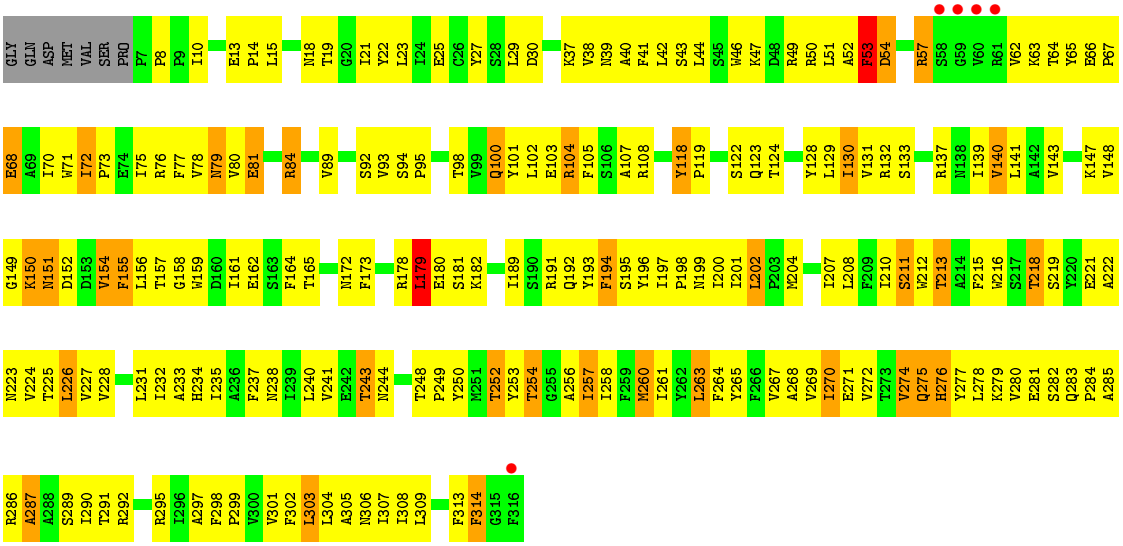
Chain B:





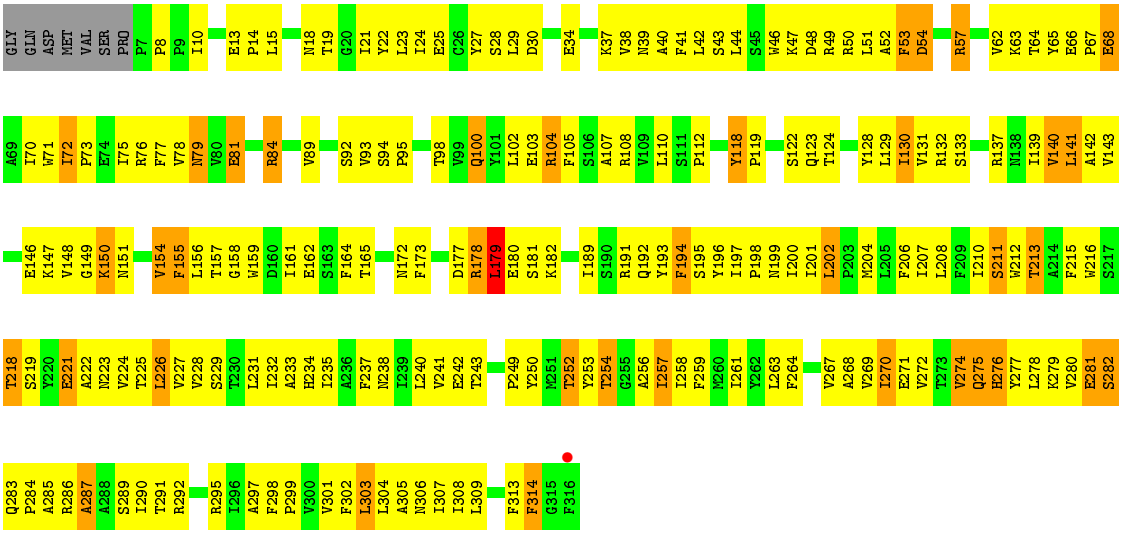
• Molecule 1: GLR4197 PROTEIN

Chain C:



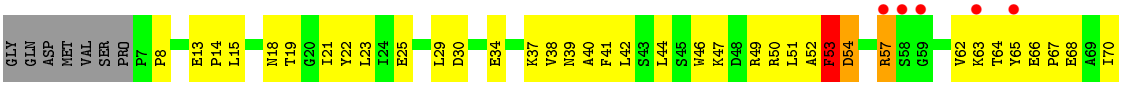
• Molecule 1: GLR4197 PROTEIN

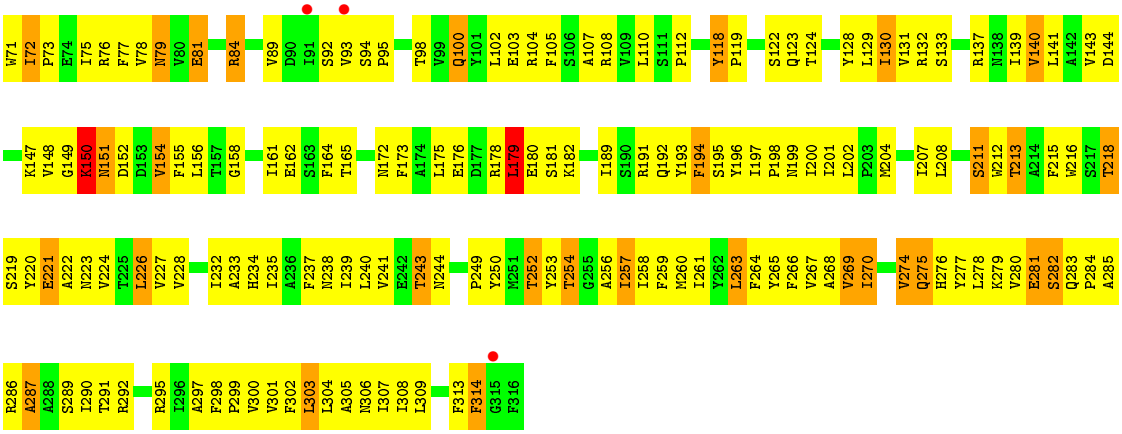
Chain D:



• Molecule 1: GLR4197 PROTEIN

Chain E:




W71 I72 P73 E74 I75 R76 F77 V78 N79 V80 E81 R84 V89 D90 I91 S92 V93 S94 P95 T98 V99 Q100 Y101 L102 E103 R104 F105 S106 A107 R108 V109 L110 S111 P112 Y118 P119 S122 Q123 T124 Y128 L129 I130 V131 R132 S133 R137 N138 I139 V140 L141 A142 V143 D144
K147 V148 G149 K150 N151 D152 D153 V154 F155 L156 T157 G158 I161 E162 S163 F164 T165 N172 F173 A174 L175 E176 D177 R178 L179 E180 S181 K182 I189 S190 R191 Q192 Y193 F194 S195 Y196 I197 P198 N199 I200 I201 L202 P203 M204 I207 L208 S211 W212 T213 A214 F215 W216 S217 T218
S219 Y220 E221 A222 N223 V224 T225 L226 V227 V228 I232 A233 H234 I235 A236 F237 N238 I239 L240 V241 E242 T243 N244 P249 Y250 M251 T252 Y253 T254 G255 A256 I257 I258 F259 M260 I261 Y262 L263 F264 Y265 F266 V267 A268 V269 I270 V274 Q275 H276 Y277 L278 K279 V280 E281 S282 Q283 P284 A285
R286 A287 A288 S289 I290 T291 R292 R295 I296 A297 F298 P299 V300 V301 F302 L303 L304 A305 N306 I307 I308 L309 F313 F314 G315 F316

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	183.31Å 128.31Å 164.37Å 90.00° 104.04° 90.00°	Depositor
Resolution (Å)	40.20 – 3.60 40.20 – 3.60	Depositor EDS
% Data completeness (in resolution range)	98.8 (40.20-3.60) 98.9 (40.20-3.60)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.05 (at 3.57Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.259 , 0.272 0.256 , 0.272	Depositor DCC
R_{free} test set	2172 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	97.7	Xtriage
Anisotropy	0.376	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 87.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	12606	wwPDB-VP
Average B, all atoms (Å^2)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% 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: ARS

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	0.49	0/2589	0.68	3/3535 (0.1%)
1	B	0.49	0/2589	0.68	4/3535 (0.1%)
1	C	0.50	0/2589	0.69	3/3535 (0.1%)
1	D	0.51	0/2589	0.69	3/3535 (0.1%)
1	E	0.51	0/2589	0.69	4/3535 (0.1%)
All	All	0.50	0/12945	0.68	17/17675 (0.1%)

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	179	LEU	CA-CB-CG	5.56	128.08	115.30
1	B	281	GLU	N-CA-C	-5.55	96.02	111.00
1	C	179	LEU	CA-CB-CG	5.53	128.01	115.30
1	E	179	LEU	CA-CB-CG	5.52	128.00	115.30
1	D	276	HIS	N-CA-CB	-5.45	100.80	110.60

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	2521	0	2537	278	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2521	0	2537	274	0
1	C	2521	0	2537	232	0
1	D	2521	0	2537	272	0
1	E	2521	0	2537	241	0
2	C	1	0	0	0	0
All	All	12606	0	12685	1150	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 45.

The worst 5 of 1150 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:155:PHE:CE1	1:B:112:PRO:HB3	1.77	1.18
1:A:27:TYR:HB3	1:B:110:LEU:HD11	1.22	1.15
1:D:155:PHE:CE1	1:E:112:PRO:HB3	1.84	1.11
1:A:222:ALA:HB2	1:B:221:GLU:HA	1.37	1.06
1:A:240:LEU:HD13	1:B:239:ILE:HG12	1.28	1.05

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	308/317 (97%)	239 (78%)	59 (19%)	10 (3%)	4	31
1	B	308/317 (97%)	238 (77%)	61 (20%)	9 (3%)	4	33
1	C	308/317 (97%)	235 (76%)	64 (21%)	9 (3%)	4	33
1	D	308/317 (97%)	238 (77%)	61 (20%)	9 (3%)	4	33
1	E	308/317 (97%)	239 (78%)	61 (20%)	8 (3%)	5	35
All	All	1540/1585 (97%)	1189 (77%)	306 (20%)	45 (3%)	4	33

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	118	TYR
1	A	148	VAL
1	A	150	LYS
1	A	194	PHE
1	B	118	TYR

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	278/284 (98%)	234 (84%)	44 (16%)	2	17
1	B	278/284 (98%)	234 (84%)	44 (16%)	2	17
1	C	278/284 (98%)	234 (84%)	44 (16%)	2	17
1	D	278/284 (98%)	234 (84%)	44 (16%)	2	17
1	E	278/284 (98%)	235 (84%)	43 (16%)	2	18
All	All	1390/1420 (98%)	1171 (84%)	219 (16%)	2	17

5 of 219 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	122	SER
1	C	260	MET
1	E	218	THR
1	C	130	ILE
1	C	211	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	151	ASN
1	C	275	GLN
1	E	151	ASN
1	C	234	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	275	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is 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, 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	310/317 (97%)	-0.11	5 (1%) 72 57	77, 116, 170, 252	0
1	B	310/317 (97%)	-0.10	11 (3%) 44 29	77, 116, 170, 251	0
1	C	310/317 (97%)	-0.16	5 (1%) 72 57	77, 115, 169, 250	0
1	D	310/317 (97%)	-0.12	1 (0%) 94 88	75, 115, 169, 252	0
1	E	310/317 (97%)	-0.14	8 (2%) 56 40	79, 115, 170, 251	0
All	All	1550/1585 (97%)	-0.12	30 (1%) 66 51	75, 116, 170, 252	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	59	GLY	5.5
1	C	59	GLY	4.4
1	A	59	GLY	4.3
1	B	91	ILE	3.9
1	C	61	ARG	3.8

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 carbohydrates in this entry.

6.4 Ligands

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, 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	B-factors(Å^2)	Q<0.9
2	ARS	C	1317	1/1	0.71	1.17	206,206,206,206	0

6.5 Other polymers (i)

There are no such residues in this entry.