wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 08:24 AM GMT

PDB ID	:	3EI0
Title	:	Structure of the E221A mutant of the Gloebacter violaceus pentameric ligand gated ion channnel (GLIC)
Authors	:	Hilf, R.J.C.; Dutzler, R.
Deposited on	:	2008-09-15
Resolution	:	3.50 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:

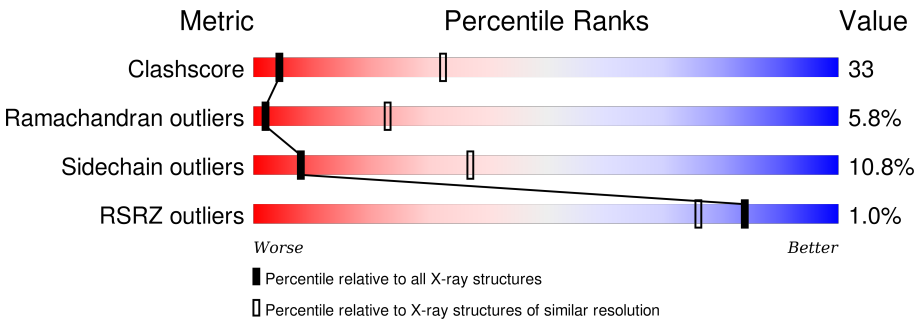
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.50 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))
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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	317	% 41% 46% 10% • •
1	B	317	% 39% 47% 11% • •
1	C	317	39% 49% 9% • •
1	D	317	% 40% 47% 9% • •
1	E	317	% 41% 47% 10% • •

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 12585 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
			Total	C	N	O	S			
1	A	310	2517	1660	403	450	4	0	0	0
1	B	310	2517	1660	403	450	4	0	0	0
1	C	310	2517	1660	403	450	4	0	0	0
1	D	310	2517	1660	403	450	4	0	0	0
1	E	310	2517	1660	403	450	4	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	EXPRESSION TAG	UNP Q7NDN8
A	1	GLN	-	EXPRESSION TAG	UNP Q7NDN8
A	2	ASP	-	EXPRESSION TAG	UNP Q7NDN8
A	3	MET	-	EXPRESSION TAG	UNP Q7NDN8
A	4	VAL	-	EXPRESSION TAG	UNP Q7NDN8
A	5	SER	-	EXPRESSION TAG	UNP Q7NDN8
A	6	PRO	-	EXPRESSION TAG	UNP Q7NDN8
A	221	ALA	GLU	ENGINEERED	UNP Q7NDN8
B	0	GLY	-	EXPRESSION TAG	UNP Q7NDN8
B	1	GLN	-	EXPRESSION TAG	UNP Q7NDN8
B	2	ASP	-	EXPRESSION TAG	UNP Q7NDN8
B	3	MET	-	EXPRESSION TAG	UNP Q7NDN8
B	4	VAL	-	EXPRESSION TAG	UNP Q7NDN8
B	5	SER	-	EXPRESSION TAG	UNP Q7NDN8
B	6	PRO	-	EXPRESSION TAG	UNP Q7NDN8
B	221	ALA	GLU	ENGINEERED	UNP Q7NDN8
C	0	GLY	-	EXPRESSION TAG	UNP Q7NDN8
C	1	GLN	-	EXPRESSION TAG	UNP Q7NDN8
C	2	ASP	-	EXPRESSION TAG	UNP Q7NDN8

Continued on next page...

Continued from previous page...

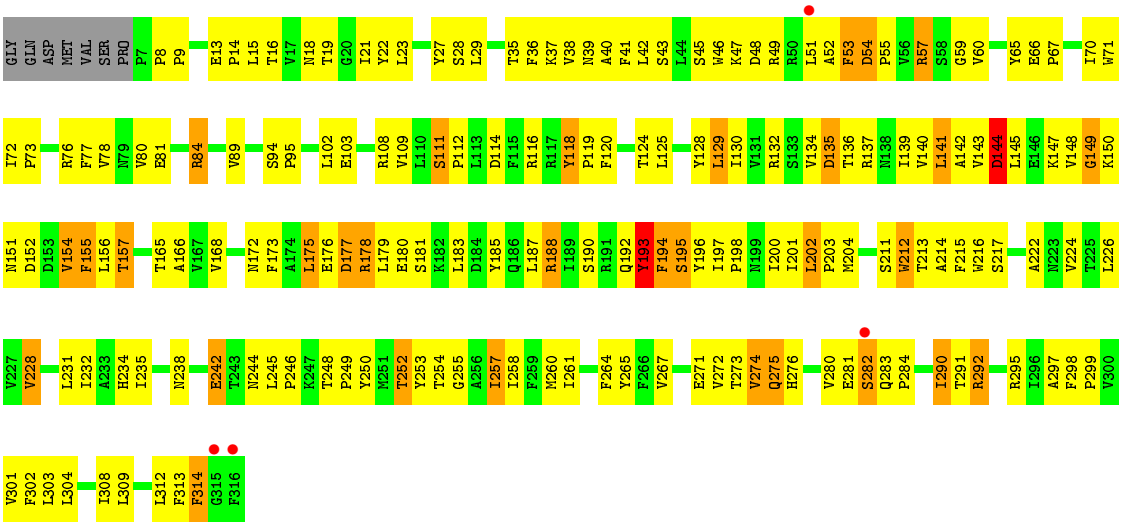
Chain	Residue	Modelled	Actual	Comment	Reference
C	3	MET	-	EXPRESSION TAG	UNP Q7NDN8
C	4	VAL	-	EXPRESSION TAG	UNP Q7NDN8
C	5	SER	-	EXPRESSION TAG	UNP Q7NDN8
C	6	PRO	-	EXPRESSION TAG	UNP Q7NDN8
C	221	ALA	GLU	ENGINEERED	UNP Q7NDN8
D	0	GLY	-	EXPRESSION TAG	UNP Q7NDN8
D	1	GLN	-	EXPRESSION TAG	UNP Q7NDN8
D	2	ASP	-	EXPRESSION TAG	UNP Q7NDN8
D	3	MET	-	EXPRESSION TAG	UNP Q7NDN8
D	4	VAL	-	EXPRESSION TAG	UNP Q7NDN8
D	5	SER	-	EXPRESSION TAG	UNP Q7NDN8
D	6	PRO	-	EXPRESSION TAG	UNP Q7NDN8
D	221	ALA	GLU	ENGINEERED	UNP Q7NDN8
E	0	GLY	-	EXPRESSION TAG	UNP Q7NDN8
E	1	GLN	-	EXPRESSION TAG	UNP Q7NDN8
E	2	ASP	-	EXPRESSION TAG	UNP Q7NDN8
E	3	MET	-	EXPRESSION TAG	UNP Q7NDN8
E	4	VAL	-	EXPRESSION TAG	UNP Q7NDN8
E	5	SER	-	EXPRESSION TAG	UNP Q7NDN8
E	6	PRO	-	EXPRESSION TAG	UNP Q7NDN8
E	221	ALA	GLU	ENGINEERED	UNP Q7NDN8

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.

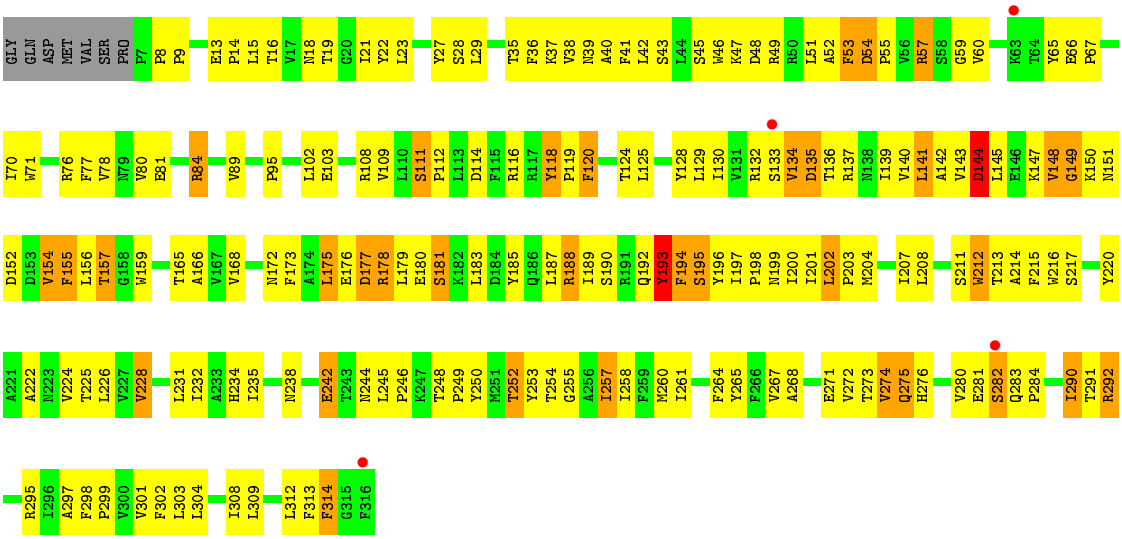
- Molecule 1: Glr4197 protein

Chain A:



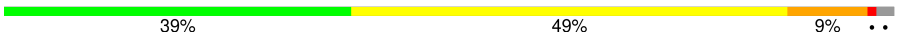
- Molecule 1: Glr4197 protein

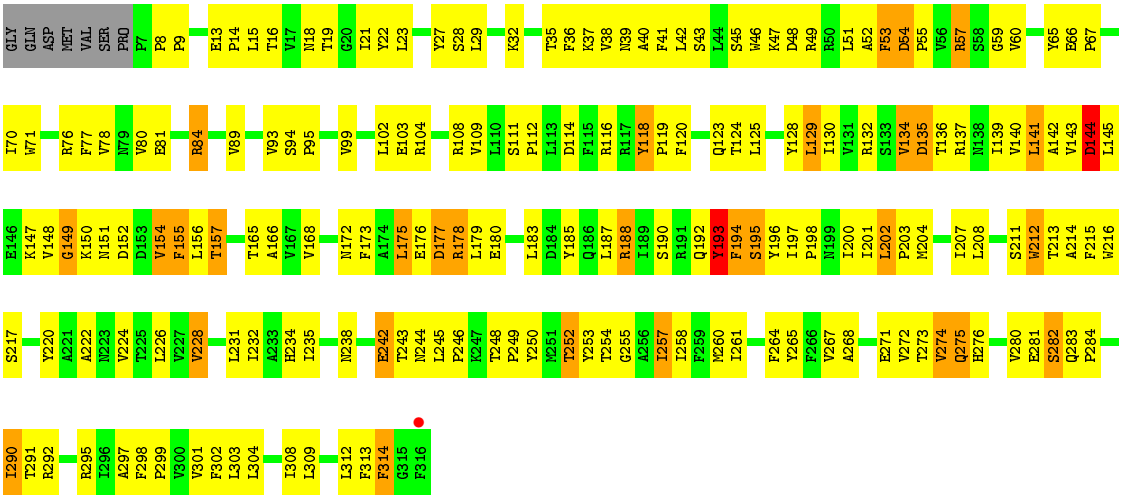
Chain B:



- Molecule 1: Glr4197 protein

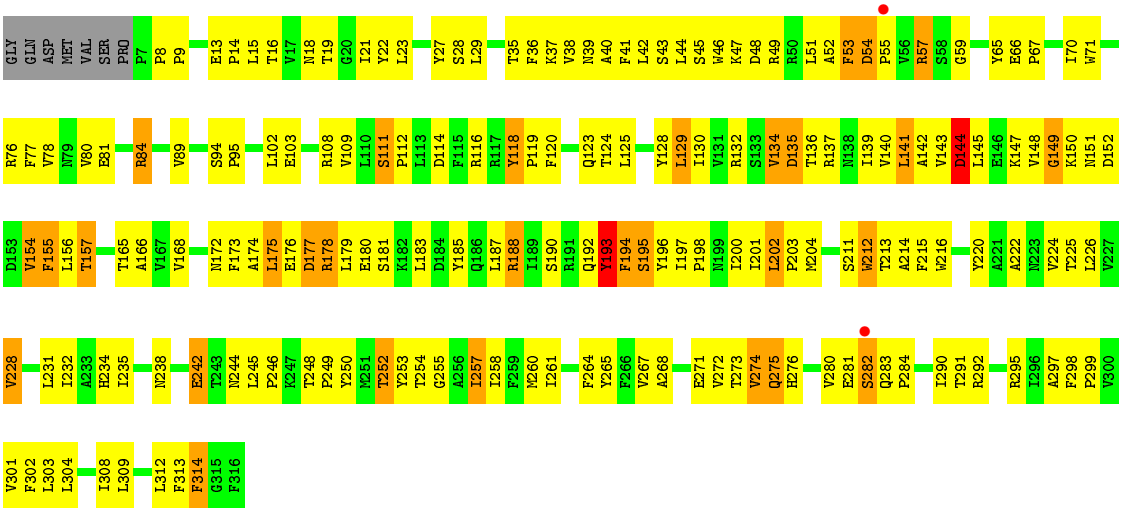
Chain C:





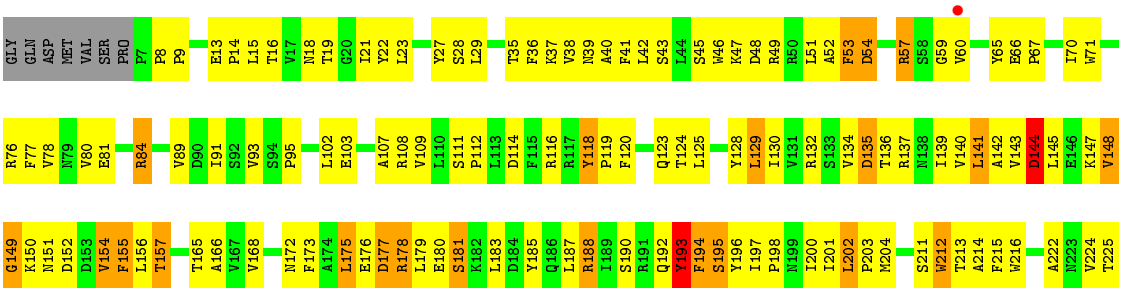
- Molecule 1: Glr4197 protein

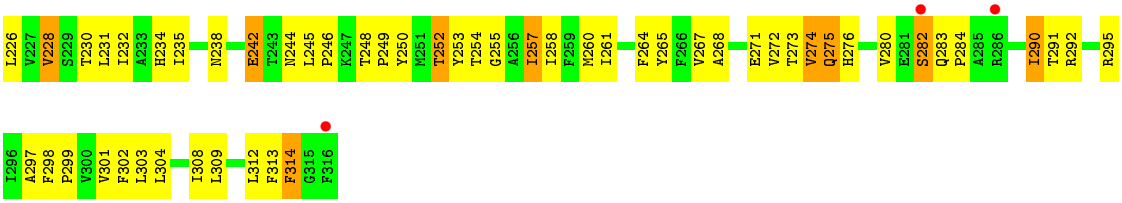
Chain D:



- Molecule 1: Glr4197 protein

Chain E:




L226 V227 V228 S229 T230 L231 I232 A233 H234 I235
N238
E242 T243 N244 L245 P246 K247 T248 P249 Y250 M251 T252 Y253 T254 G255 A256 I257 I258 F259 M260 I261
F264 Y265 F266 V267 A268
E271 V272 T273 V274 Q275 H276
V280 E281 S282 Q283 P284 A285 R286
I290 T291 R292
R295
I296 A297 F298 P299 V300 V301 F302 L303 L304
I308 L309
L312 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	179.64Å 133.76Å 161.65Å 90.00° 101.67° 90.00°	Depositor
Resolution (Å)	20.00 – 3.50 39.58 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.2 (20.00-3.50) 99.1 (39.58-3.50)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.23 (at 3.48Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.255 , 0.276 0.251 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	85.5	Xtriage
Anisotropy	0.311	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 69.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Outliers	0 of 47328 reflections	Xtriage
F_o,F_c correlation	0.84	EDS
Total number of atoms	12585	wwPDB-VP
Average B, all atoms (Å^2)	107.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.56% 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 (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	A	0.38	0/2585	0.54	0/3530
1	B	0.38	0/2585	0.54	0/3530
1	C	0.38	0/2585	0.54	0/3530
1	D	0.40	0/2585	0.55	0/3530
1	E	0.39	0/2585	0.54	0/3530
All	All	0.39	0/12925	0.54	0/17650

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2517	0	2536	168	0
1	B	2517	0	2536	172	0
1	C	2517	0	2536	170	0
1	D	2517	0	2536	170	0
1	E	2517	0	2536	170	0
All	All	12585	0	12680	825	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 33.

The worst 5 of 825 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:D:22:TYR:HA	1:D:149:GLY:HA2	1.43	0.99
1:C:22:TYR:HA	1:C:149:GLY:HA2	1.45	0.98
1:B:22:TYR:HA	1:B:149:GLY:HA2	1.45	0.96
1:A:147:LYS:HE2	1:A:165:THR:HA	1.48	0.96
1:E:22:TYR:HA	1:E:149:GLY:HA2	1.44	0.96

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%)	246 (80%)	44 (14%)	18 (6%)	2 23
1	B	308/317 (97%)	250 (81%)	40 (13%)	18 (6%)	2 23
1	C	308/317 (97%)	247 (80%)	43 (14%)	18 (6%)	2 23
1	D	308/317 (97%)	249 (81%)	42 (14%)	17 (6%)	2 24
1	E	308/317 (97%)	247 (80%)	43 (14%)	18 (6%)	2 23
All	All	1540/1585 (97%)	1239 (80%)	212 (14%)	89 (6%)	2 23

5 of 89 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	149	GLY
1	A	178	ARG
1	A	193	TYR
1	B	149	GLY
1	B	178	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	277/283 (98%)	248 (90%)	29 (10%)	8	38
1	B	277/283 (98%)	246 (89%)	31 (11%)	7	35
1	C	277/283 (98%)	248 (90%)	29 (10%)	8	38
1	D	277/283 (98%)	248 (90%)	29 (10%)	8	38
1	E	277/283 (98%)	246 (89%)	31 (11%)	7	35
All	All	1385/1415 (98%)	1236 (89%)	149 (11%)	8	37

5 of 149 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	141	LEU
1	C	257	ILE
1	E	202	LEU
1	C	154	VAL
1	C	190	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	123	GLN
1	C	199	ASN
1	E	192	GLN
1	C	151	ASN
1	C	192	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 carbohydrates 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	A	310/317 (97%)	-0.21	4 (1%)	79	70	76, 103, 142, 169	0
1	B	310/317 (97%)	-0.24	4 (1%)	79	70	76, 103, 142, 169	0
1	C	310/317 (97%)	-0.26	1 (0%)	94	91	76, 103, 142, 169	0
1	D	310/317 (97%)	-0.22	2 (0%)	90	85	76, 103, 142, 169	0
1	E	310/317 (97%)	-0.25	4 (1%)	79	70	76, 103, 142, 169	0
All	All	1550/1585 (97%)	-0.24	15 (0%)	84	76	76, 103, 143, 169	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	316	PHE	3.7
1	A	316	PHE	3.1
1	E	282	SER	2.9
1	D	282	SER	2.6
1	E	316	PHE	2.6

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.