wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:47 PM GMT

PDB ID	:	1QMI
Title	:	Crystal structure of RNA 3'-terminal phosphate cyclase, an ubiquitous enzyme with unusual topology
Authors	:	Palm, G.J.; Billy, E.; Filipowicz, W.; Wlodawer, A.
Deposited on	:	1999-09-28
Resolution	:	2.80 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:

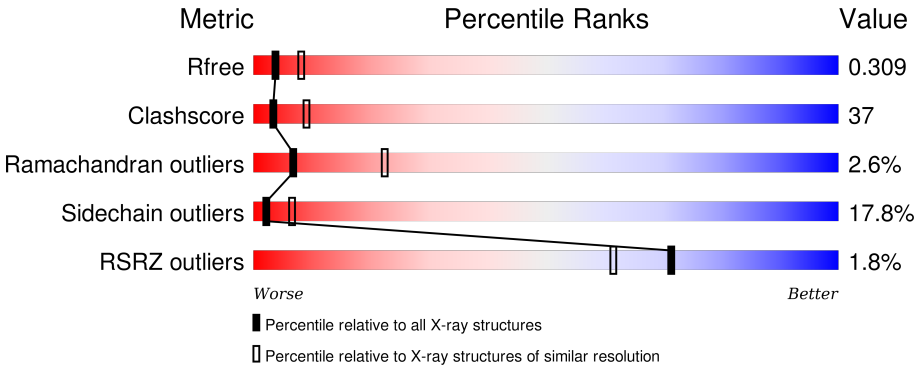
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.80 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}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	347	2% 42% 42% 11%
1	B	347	% 40% 46% 10%
1	C	347	2% 39% 45% 12%
1	D	347	% 41% 46% 8%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9972 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 RNA 3'-TERMINAL PHOSPHATE CYCLASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	335	Total	C	N	O	S	Se	0	0	0
			2493	1571	449	466	3	4			
1	B	335	Total	C	N	O	S	Se	0	0	0
			2493	1571	449	466	3	4			
1	C	335	Total	C	N	O	S	Se	0	0	0
			2493	1571	449	466	3	4			
1	D	335	Total	C	N	O	S	Se	0	0	0
			2493	1571	449	466	3	4			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	EXPRESSION TAG	UNP P46849
A	2	VAL	-	EXPRESSION TAG	UNP P46849
A	340	GLY	-	EXPRESSION TAG	UNP P46849
A	341	SER	-	EXPRESSION TAG	UNP P46849
A	342	HIS	-	EXPRESSION TAG	UNP P46849
A	343	HIS	-	EXPRESSION TAG	UNP P46849
A	344	HIS	-	EXPRESSION TAG	UNP P46849
A	345	HIS	-	EXPRESSION TAG	UNP P46849
A	346	HIS	-	EXPRESSION TAG	UNP P46849
A	347	HIS	-	EXPRESSION TAG	UNP P46849
B	1	MSE	-	EXPRESSION TAG	UNP P46849
B	2	VAL	-	EXPRESSION TAG	UNP P46849
B	340	GLY	-	EXPRESSION TAG	UNP P46849
B	341	SER	-	EXPRESSION TAG	UNP P46849
B	342	HIS	-	EXPRESSION TAG	UNP P46849
B	343	HIS	-	EXPRESSION TAG	UNP P46849
B	344	HIS	-	EXPRESSION TAG	UNP P46849
B	345	HIS	-	EXPRESSION TAG	UNP P46849
B	346	HIS	-	EXPRESSION TAG	UNP P46849
B	347	HIS	-	EXPRESSION TAG	UNP P46849
C	1	MSE	-	EXPRESSION TAG	UNP P46849

Continued on next page...

Continued from previous page...

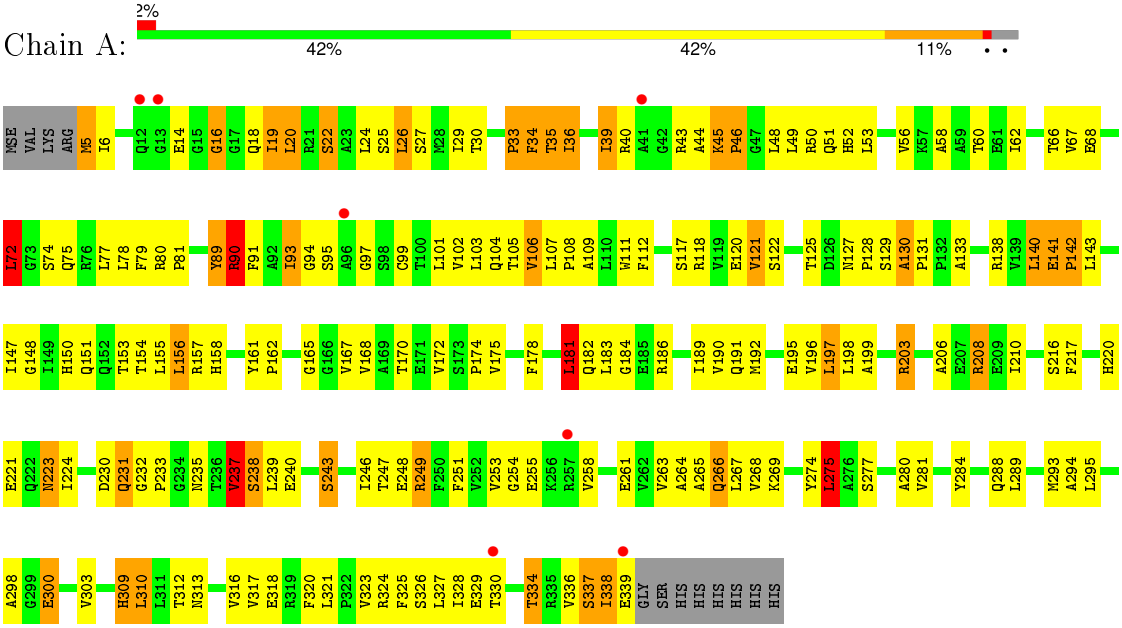
Chain	Residue	Modelled	Actual	Comment	Reference
C	2	VAL	-	EXPRESSION TAG	UNP P46849
C	340	GLY	-	EXPRESSION TAG	UNP P46849
C	341	SER	-	EXPRESSION TAG	UNP P46849
C	342	HIS	-	EXPRESSION TAG	UNP P46849
C	343	HIS	-	EXPRESSION TAG	UNP P46849
C	344	HIS	-	EXPRESSION TAG	UNP P46849
C	345	HIS	-	EXPRESSION TAG	UNP P46849
C	346	HIS	-	EXPRESSION TAG	UNP P46849
C	347	HIS	-	EXPRESSION TAG	UNP P46849
D	1	MSE	-	EXPRESSION TAG	UNP P46849
D	2	VAL	-	EXPRESSION TAG	UNP P46849
D	340	GLY	-	EXPRESSION TAG	UNP P46849
D	341	SER	-	EXPRESSION TAG	UNP P46849
D	342	HIS	-	EXPRESSION TAG	UNP P46849
D	343	HIS	-	EXPRESSION TAG	UNP P46849
D	344	HIS	-	EXPRESSION TAG	UNP P46849
D	345	HIS	-	EXPRESSION TAG	UNP P46849
D	346	HIS	-	EXPRESSION TAG	UNP P46849
D	347	HIS	-	EXPRESSION TAG	UNP P46849

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.

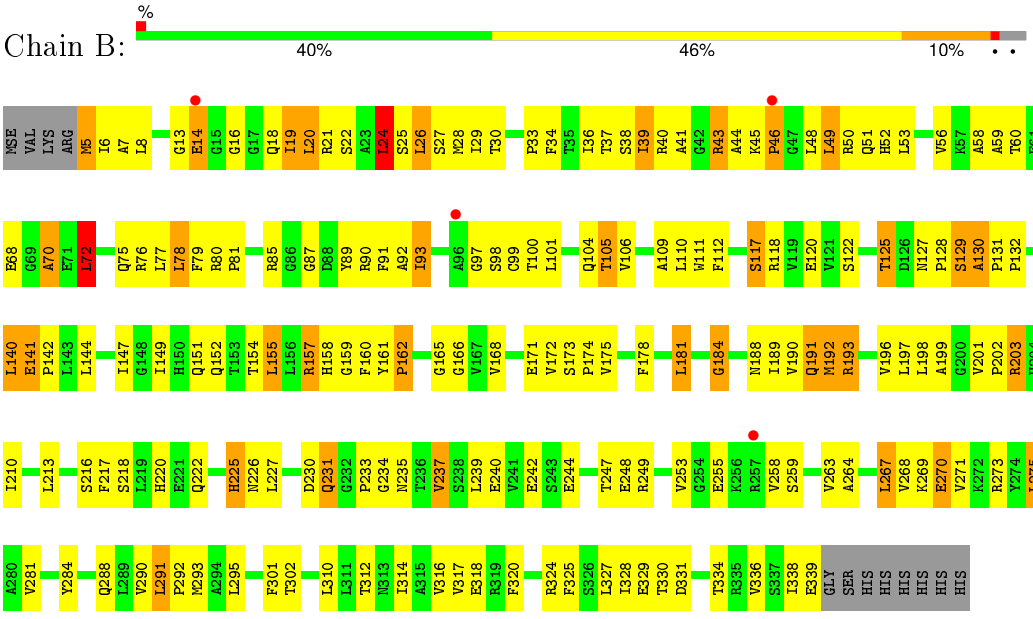
- Molecule 1: RNA 3'-TERMINAL PHOSPHATE CYCLASE

Chain A:



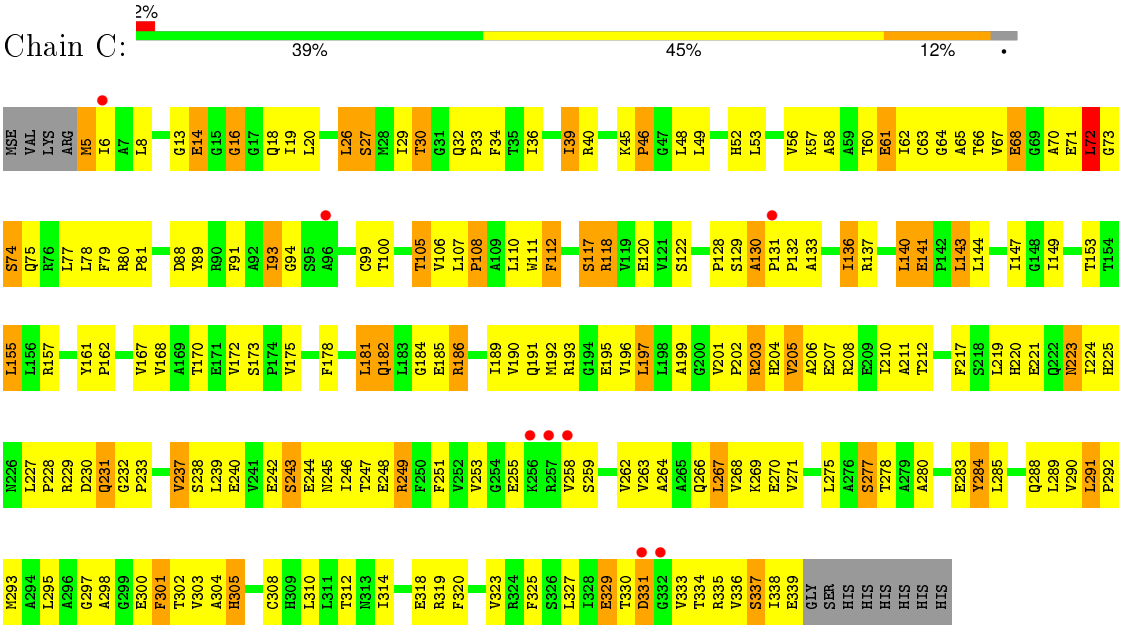
- Molecule 1: RNA 3'-TERMINAL PHOSPHATE CYCLASE

Chain B:



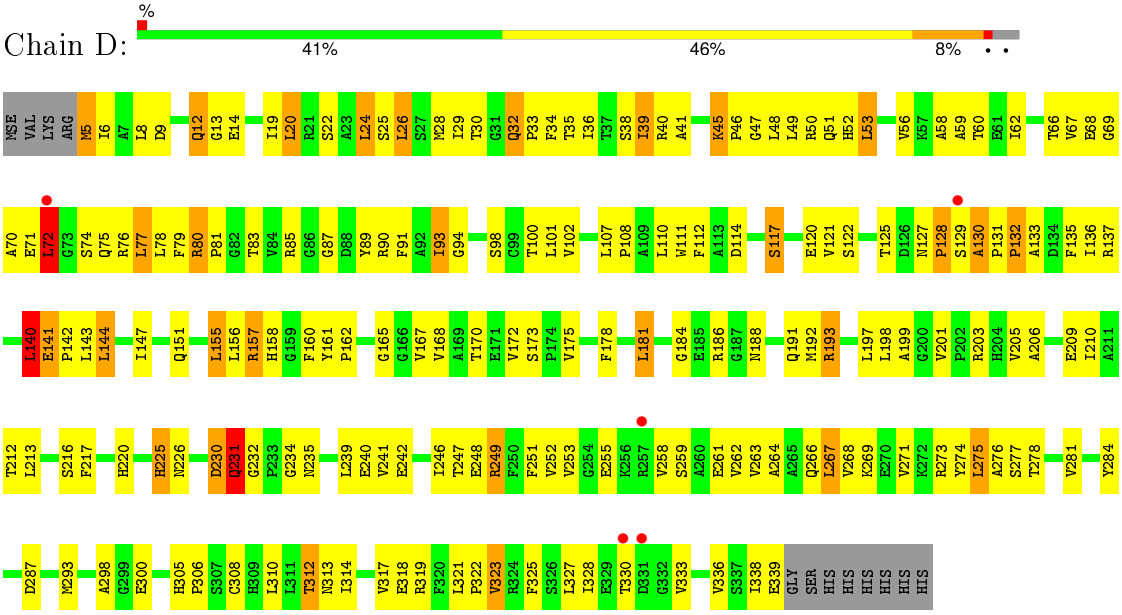
• Molecule 1: RNA 3'-TERMINAL PHOSPHATE CYCLASE

Chain C:



• Molecule 1: RNA 3'-TERMINAL PHOSPHATE CYCLASE

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	101.80Å 126.60Å 128.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.80 19.98 – 2.81	Depositor EDS
% Data completeness (in resolution range)	89.8 (10.00-2.80) 92.7 (19.98-2.81)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.76 (at 2.79Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.244 , 0.331 0.221 , 0.309	Depositor DCC
R_{free} test set	1891 reflections (5.08%)	DCC
Wilson B-factor (Å^2)	44.8	Xtriage
Anisotropy	0.440	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 92.1	EDS
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 39548 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	9972	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 40.72 % of the origin peak, indicating pseudo translational symmetry.The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.6224e-04.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

5.1 Standard geometry

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	0.93	3/2533 (0.1%)	1.07	7/3436 (0.2%)
1	B	0.87	0/2533	1.06	7/3436 (0.2%)
1	C	0.94	0/2533	1.07	4/3436 (0.1%)
1	D	0.82	0/2533	1.06	6/3436 (0.2%)
All	All	0.89	3/10132 (0.0%)	1.06	24/13744 (0.2%)

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	A	0	1
1	C	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	237	VAL	CA-CB	5.66	1.66	1.54
1	A	34	PHE	CB-CG	5.26	1.60	1.51
1	A	99	CYS	CB-SG	-5.24	1.73	1.81

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	181	LEU	CA-CB-CG	10.29	138.96	115.30
1	A	181	LEU	CA-CB-CG	8.84	135.62	115.30
1	C	181	LEU	CA-CB-CG	8.17	134.08	115.30
1	A	20	LEU	CA-CB-CG	8.10	133.92	115.30
1	C	249	ARG	NE-CZ-NH1	-6.80	116.90	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	89	TYR	Sidechain
1	C	284	TYR	Sidechain

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	2493	0	2533	186	0
1	B	2493	0	2533	197	0
1	C	2493	0	2533	201	0
1	D	2493	0	2533	163	0
All	All	9972	0	10132	741	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 37.

The worst 5 of 741 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:B:5:MSE:HE2	1:B:33:PRO:HB2	1.32	1.11
1:D:293:MSE:HB3	1:D:338:ILE:HD11	1.30	1.10
1:D:239:LEU:HD13	1:D:271:VAL:HG21	1.26	1.07
1:D:26:LEU:O	1:D:30:THR:HG22	1.55	1.05
1:B:26:LEU:O	1:B:30:THR:HG22	1.62	0.97

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	333/347 (96%)	291 (87%)	34 (10%)	8 (2%)	7 25
1	B	333/347 (96%)	295 (89%)	30 (9%)	8 (2%)	7 25
1	C	333/347 (96%)	293 (88%)	32 (10%)	8 (2%)	7 25
1	D	333/347 (96%)	294 (88%)	29 (9%)	10 (3%)	5 18
All	All	1332/1388 (96%)	1173 (88%)	125 (9%)	34 (3%)	7 22

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	ILE
1	A	130	ALA
1	B	39	ILE
1	B	130	ALA
1	C	39	ILE

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	262/269 (97%)	215 (82%)	47 (18%)	2 6
1	B	262/269 (97%)	217 (83%)	45 (17%)	2 7
1	C	262/269 (97%)	211 (80%)	51 (20%)	2 5
1	D	262/269 (97%)	218 (83%)	44 (17%)	2 8
All	All	1048/1076 (97%)	861 (82%)	187 (18%)	2 6

5 of 187 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	259	SER
1	C	108	PRO
1	D	230	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	275	LEU
1	C	26	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	288	GLN
1	C	182	GLN
1	D	52	HIS
1	B	231	GLN
1	D	179	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 carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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	331/347 (95%)	-0.26	7 (2%)	67	56	2, 26, 97, 160	0
1	B	331/347 (95%)	-0.32	4 (1%)	81	73	2, 29, 92, 186	0
1	C	331/347 (95%)	-0.38	8 (2%)	62	50	2, 27, 94, 141	0
1	D	331/347 (95%)	-0.37	5 (1%)	76	68	2, 30, 99, 162	0
All	All	1324/1388 (95%)	-0.33	24 (1%)	71	61	2, 28, 97, 186	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	96	ALA	5.0
1	C	96	ALA	3.8
1	A	257	ARG	3.7
1	D	257	ARG	3.7
1	A	330	THR	3.6

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

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.