wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 12:41 AM GMT

PDB ID	:	2BA0
Title	:	Archaeal exosome core
Authors	:	Buttner, K.; Wenig, K.; Hopfner, K.P.
Deposited on	:	2005-10-13
Resolution	:	2.70 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 (i)) were used in the production of this report:

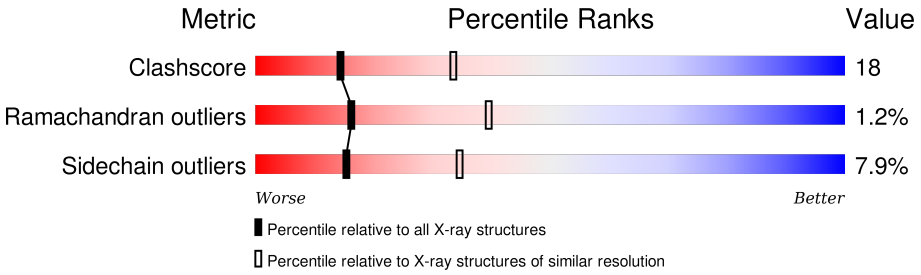
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 2.70 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	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	229	66% 24% • 5%
1	B	229	64% 28% 5% •
1	C	229	62% 28% 5% 5%
2	D	258	62% 28% • • 6%
2	E	258	62% 27% 5% 6%
2	F	258	64% 26% • 5%
3	G	259	71% 25% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	259	
3	I	259	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 16910 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 Archaeal exosome RNA binding protein RRP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	68	0	0
			1703	1089	304	305	5			
1	B	222	Total	C	N	O	S	122	0	0
			1731	1107	309	310	5			
1	C	218	Total	C	N	O	S	101	0	0
			1703	1089	304	305	5			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	224	HIS	-	EXPRESSION TAG	UNP O29758
A	225	HIS	-	EXPRESSION TAG	UNP O29758
A	226	HIS	-	EXPRESSION TAG	UNP O29758
A	227	HIS	-	EXPRESSION TAG	UNP O29758
A	228	HIS	-	EXPRESSION TAG	UNP O29758
A	229	HIS	-	EXPRESSION TAG	UNP O29758
B	224	HIS	-	EXPRESSION TAG	UNP O29758
B	225	HIS	-	EXPRESSION TAG	UNP O29758
B	226	HIS	-	EXPRESSION TAG	UNP O29758
B	227	HIS	-	EXPRESSION TAG	UNP O29758
B	228	HIS	-	EXPRESSION TAG	UNP O29758
B	229	HIS	-	EXPRESSION TAG	UNP O29758
C	224	HIS	-	EXPRESSION TAG	UNP O29758
C	225	HIS	-	EXPRESSION TAG	UNP O29758
C	226	HIS	-	EXPRESSION TAG	UNP O29758
C	227	HIS	-	EXPRESSION TAG	UNP O29758
C	228	HIS	-	EXPRESSION TAG	UNP O29758
C	229	HIS	-	EXPRESSION TAG	UNP O29758

- Molecule 2 is a protein called Archaeal exosome RNA binding protein RRP41.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	246	Total	C	N	O	S	55	0	0
			1928	1215	341	359	13			
2	E	243	Total	C	N	O	S	46	0	0
			1903	1199	337	354	13			
2	D	243	Total	C	N	O	S	57	0	0
			1903	1199	337	354	13			

- Molecule 3 is a protein called Archaeal exosome RNA binding protein RRP42.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	253	Total	C	N	O	S	63	0	0
			1969	1243	326	395	5			
3	H	256	Total	C	N	O	S	84	0	0
			1989	1255	329	400	5			
3	G	255	Total	C	N	O	S	60	0	0
			1980	1250	328	397	5			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	12	Total	O	0	0
			12	12		
4	B	10	Total	O	0	0
			10	10		
4	C	6	Total	O	0	0
			6	6		
4	D	13	Total	O	0	0
			13	13		
4	E	8	Total	O	0	0
			8	8		
4	F	9	Total	O	0	0
			9	9		
4	G	16	Total	O	0	0
			16	16		
4	H	12	Total	O	0	0
			12	12		
4	I	15	Total	O	0	0
			15	15		

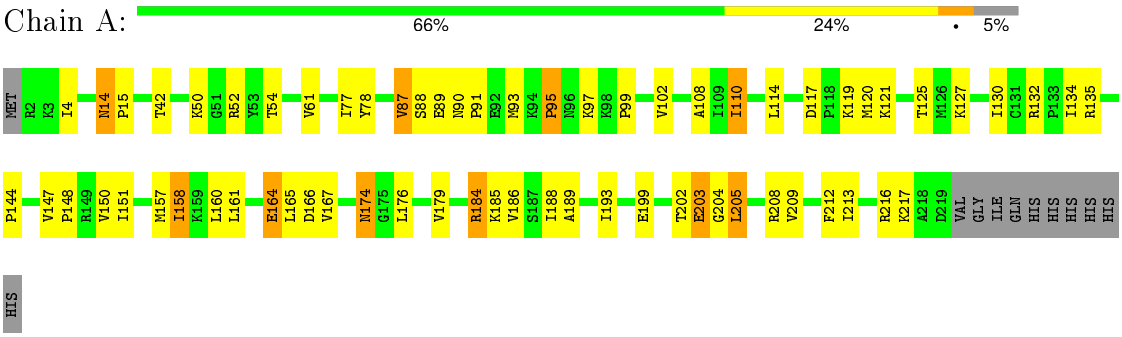
3 Residue-property plots (i)

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.

Note EDS was not executed.

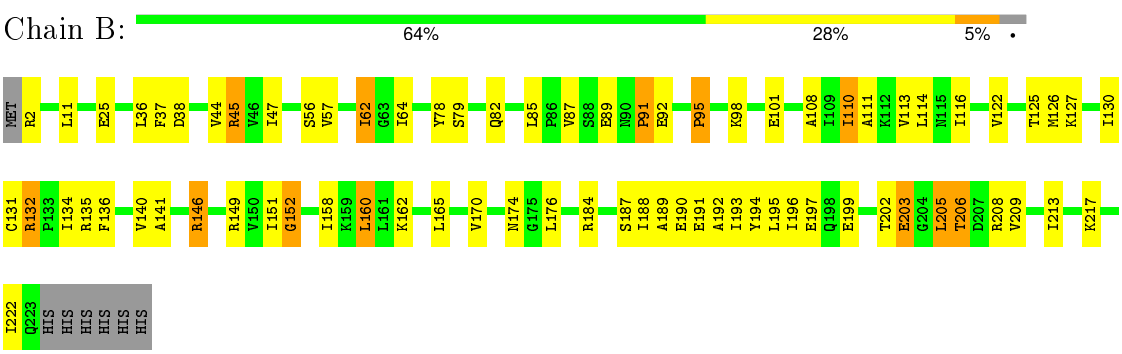
- Molecule 1: Archaeal exosome RNA binding protein RRP4

Chain A:



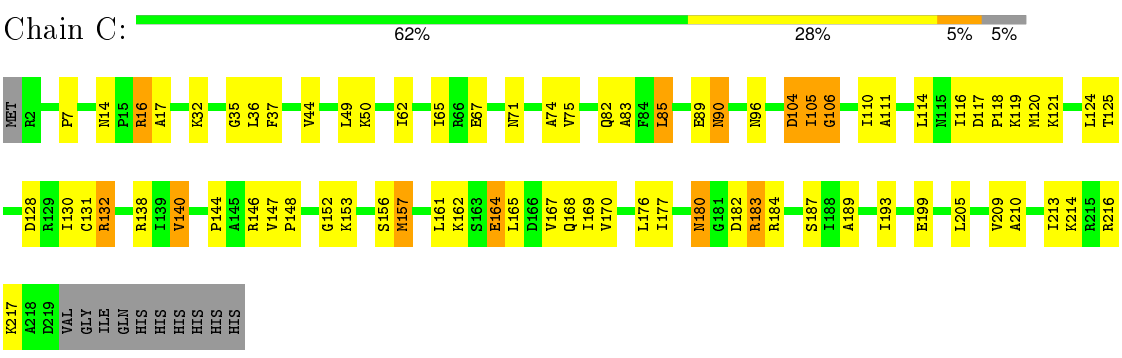
- Molecule 1: Archaeal exosome RNA binding protein RRP4

Chain B:



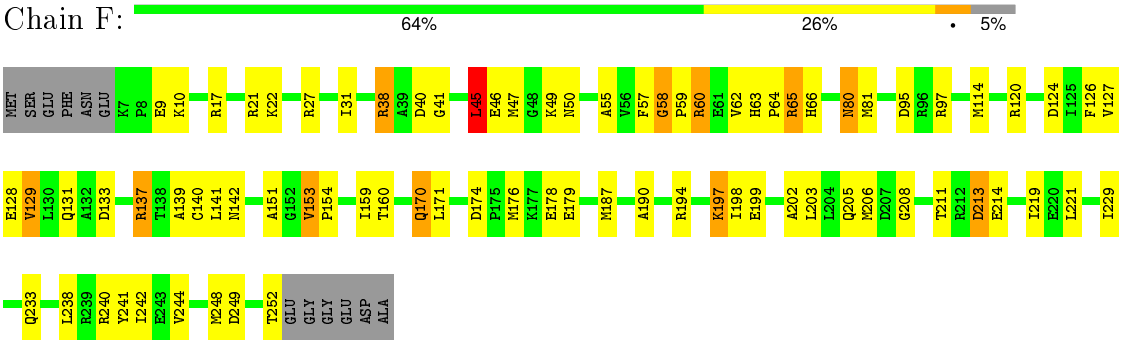
- Molecule 1: Archaeal exosome RNA binding protein RRP4

Chain C:



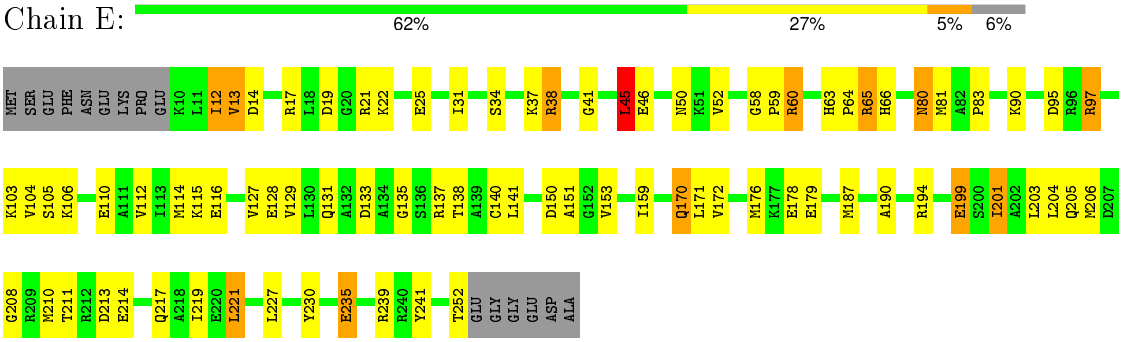
- Molecule 2: Archaeal exosome RNA binding protein RRP41

Chain F:



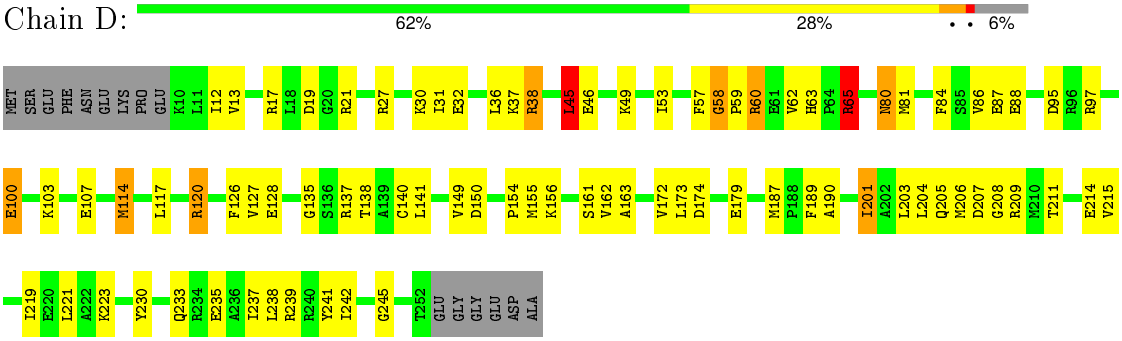
- Molecule 2: Archaeal exosome RNA binding protein RRP41

Chain E:



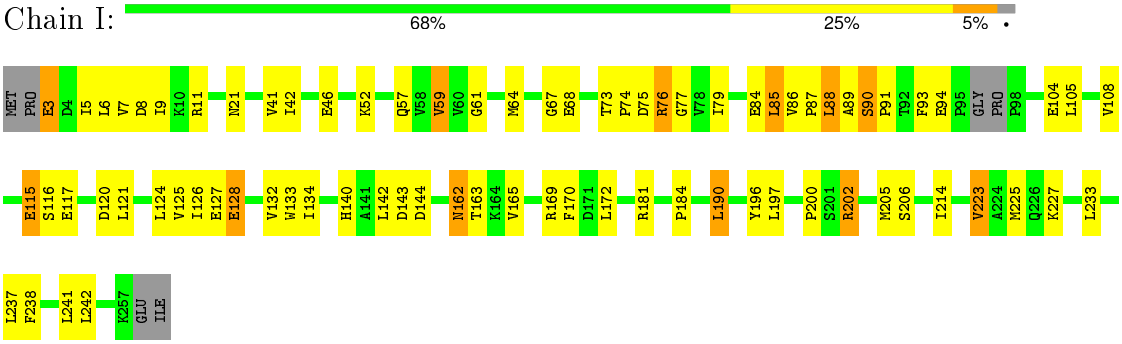
- Molecule 2: Archaeal exosome RNA binding protein RRP41

Chain D:



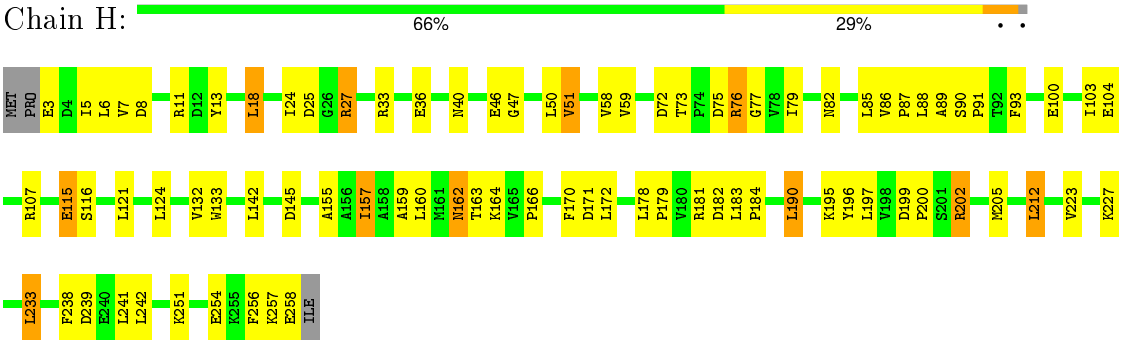
- Molecule 3: Archaeal exosome RNA binding protein RRP42

Chain I:



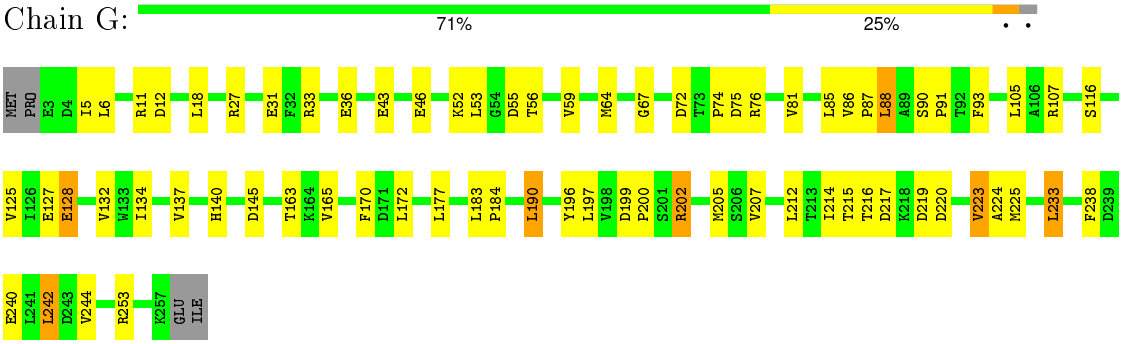
• Molecule 3: Archaeal exosome RNA binding protein RRP42

Chain H:



• Molecule 3: Archaeal exosome RNA binding protein RRP42

Chain G:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	101.57Å 129.59Å 102.33Å 90.00° 101.56° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-2.70)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.216 , 0.264	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16910	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

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	0.35	0/1732	0.63	0/2343
1	B	0.33	0/1760	0.61	0/2381
1	C	0.32	0/1732	0.59	0/2343
2	D	0.41	0/1929	0.68	1/2588 (0.0%)
2	E	0.40	0/1929	0.66	1/2588 (0.0%)
2	F	0.41	0/1955	0.66	1/2623 (0.0%)
3	G	0.39	0/2008	0.63	0/2725
3	H	0.40	0/2017	0.66	0/2737
3	I	0.38	0/1995	0.65	1/2704 (0.0%)
All	All	0.38	0/17057	0.64	4/23032 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	45	LEU	CA-CB-CG	5.78	128.60	115.30
2	E	45	LEU	CA-CB-CG	5.67	128.34	115.30
2	F	45	LEU	CA-CB-CG	5.35	127.61	115.30
3	I	88	LEU	CA-CB-CG	5.18	127.21	115.30

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	1703	0	1794	55	0
1	B	1731	0	1825	44	0
1	C	1703	0	1794	61	0
2	D	1903	0	1959	85	0
2	E	1903	0	1959	79	0
2	F	1928	0	1985	82	0
3	G	1980	0	2000	67	0
3	H	1989	0	2006	96	0
3	I	1969	0	1990	88	0
4	A	12	0	0	1	0
4	B	10	0	0	0	0
4	C	6	0	0	0	0
4	D	13	0	0	1	0
4	E	8	0	0	1	0
4	F	9	0	0	1	0
4	G	16	0	0	1	0
4	H	12	0	0	0	0
4	I	15	0	0	2	0
All	All	16910	0	17312	590	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 18.

The worst 5 of 590 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:E:38:ARG:HG2	3:G:202:ARG:NH2	1.73	1.03
1:C:110:ILE:HD11	1:C:170:VAL:HG11	1.44	0.99
3:I:6:LEU:H	3:I:6:LEU:HD22	1.24	0.97
3:I:89:ALA:CB	3:I:143:ASP:HA	1.94	0.97
3:I:115:GLU:HB3	3:I:223:VAL:HG22	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	216/229 (94%)	199 (92%)	15 (7%)	2 (1%)	21 49
1	B	220/229 (96%)	199 (90%)	15 (7%)	6 (3%)	6 16
1	C	216/229 (94%)	198 (92%)	14 (6%)	4 (2%)	10 25
2	D	241/258 (93%)	233 (97%)	5 (2%)	3 (1%)	16 39
2	E	241/258 (93%)	225 (93%)	12 (5%)	4 (2%)	11 29
2	F	244/258 (95%)	230 (94%)	11 (4%)	3 (1%)	16 39
3	G	253/259 (98%)	235 (93%)	17 (7%)	1 (0%)	39 69
3	H	254/259 (98%)	241 (95%)	13 (5%)	0	100 100
3	I	249/259 (96%)	238 (96%)	9 (4%)	2 (1%)	24 51
All	All	2134/2238 (95%)	1998 (94%)	111 (5%)	25 (1%)	16 39

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	95	PRO
1	B	203	GLU
1	C	105	ILE
2	F	58	GLY
2	F	65	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	184/194 (95%)	167 (91%)	17 (9%)	11 25
1	B	187/194 (96%)	167 (89%)	20 (11%)	8 19
1	C	184/194 (95%)	167 (91%)	17 (9%)	11 25
2	D	202/214 (94%)	191 (95%)	11 (5%)	27 56
2	E	202/214 (94%)	185 (92%)	17 (8%)	14 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	F	205/214 (96%)	189 (92%)	16 (8%)	16 35
3	G	223/227 (98%)	212 (95%)	11 (5%)	31 61
3	H	224/227 (99%)	204 (91%)	20 (9%)	12 27
3	I	222/227 (98%)	207 (93%)	15 (7%)	20 43
All	All	1833/1905 (96%)	1689 (92%)	144 (8%)	15 34

5 of 144 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	178	GLU
2	E	153	VAL
3	G	12	ASP
2	F	213	ASP
2	E	38	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	233	GLN
2	E	170	GLN
3	G	221	ASN
2	E	80	ASN
2	E	217	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

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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.