wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 05:12 AM EDT

PDB ID	:	2GID
Title	:	Crystal structures of trypanosoma bruciei MRP1/MRP2
Authors	:	Schumacher, M.A.; Karamooz, E.; Zikova, A.; Trantirek, L.; Lukes, J.
Deposited on	:	2006-03-28
Resolution	:	3.35 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

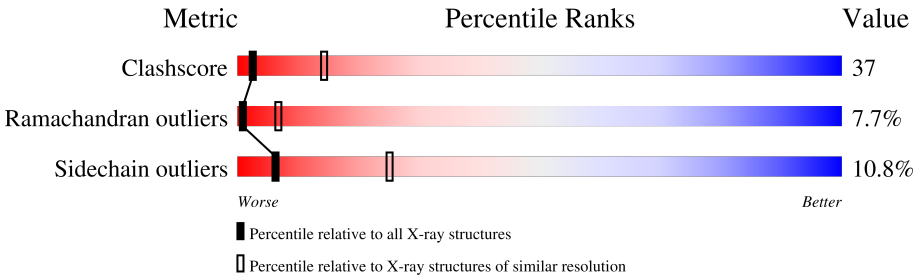
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.35 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	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	195	22% 46% 9% • 23%
1	G	195	32% 34% 11% • 23%
1	H	195	28% 42% 8% 23%
1	P	195	23% 48% 7% • 23%
2	B	187	26% 42% 10% • 22%
2	D	187	39% 35% 5% 22%
2	J	187	37% 37% • 22%
2	K	187	30% 40% 7% • 22%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9747 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 mitochondrial RNA-binding protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	150	Total	C	N	O	S	Se	0	0	0
			1224	774	223	223	1	3			
1	G	151	Total	C	N	O	S	Se	0	0	0
			1228	778	224	222	1	3			
1	H	151	Total	C	N	O	S	Se	0	0	0
			1231	779	224	224	1	3			
1	P	151	Total	C	N	O	S	Se	0	0	0
			1228	778	224	222	1	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
A	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
A	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
G	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
G	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
G	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
H	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
H	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
H	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
P	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
P	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
P	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2

- Molecule 2 is a protein called mitochondrial RNA-binding protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	146	Total	C	N	O	S	Se	0	0	0
			1195	754	220	217	1	3			
2	D	146	Total	C	N	O	S	Se	0	0	0
			1195	754	220	217	1	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	K	146	Total	C	N	O	S	Se	0	0	0
			1195	754	220	217	1	3			
2	J	146	Total	C	N	O	S	Se	0	0	0
			1195	754	220	217	1	3			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	43	GLU	LEU	CONFLICT	UNP P90629
B	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
B	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
B	139	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	43	GLU	LEU	CONFLICT	UNP P90629
D	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	139	MSE	MET	MODIFIED RESIDUE	UNP P90629
K	43	GLU	LEU	CONFLICT	UNP P90629
K	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
K	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
K	139	MSE	MET	MODIFIED RESIDUE	UNP P90629
J	43	GLU	LEU	CONFLICT	UNP P90629
J	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
J	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
J	139	MSE	MET	MODIFIED RESIDUE	UNP P90629

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	O	0	0
			4	4		
3	B	11	Total	O	0	0
			11	11		
3	D	6	Total	O	0	0
			6	6		
3	G	1	Total	O	0	0
			1	1		
3	H	4	Total	O	0	0
			4	4		
3	K	8	Total	O	0	0
			8	8		
3	J	7	Total	O	0	0
			7	7		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	15	Total	O	0	0
			15	15		

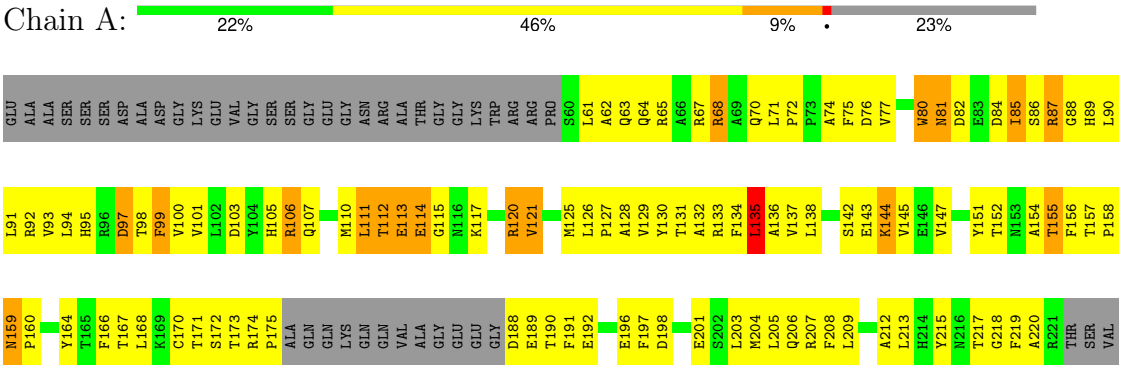
3 Residue-property plots (i)

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. 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. 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.

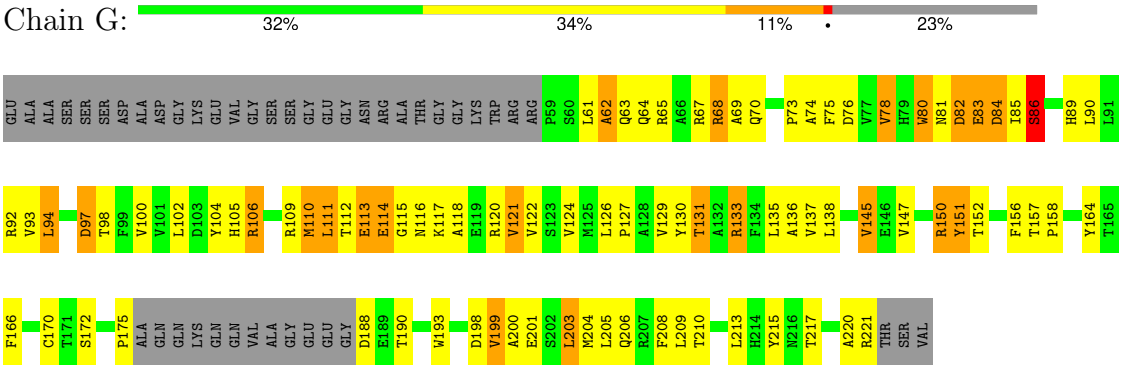
• Molecule 1: mitochondrial RNA-binding protein 2

Chain A:



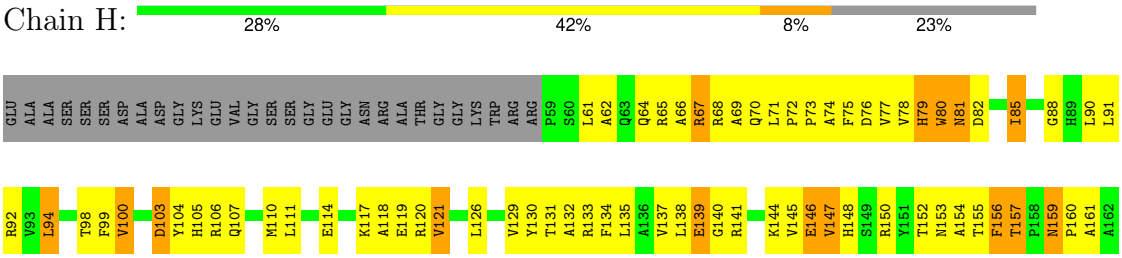
• Molecule 1: mitochondrial RNA-binding protein 2

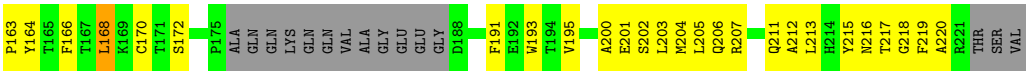
Chain G:



• Molecule 1: mitochondrial RNA-binding protein 2

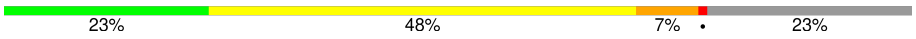
Chain H:

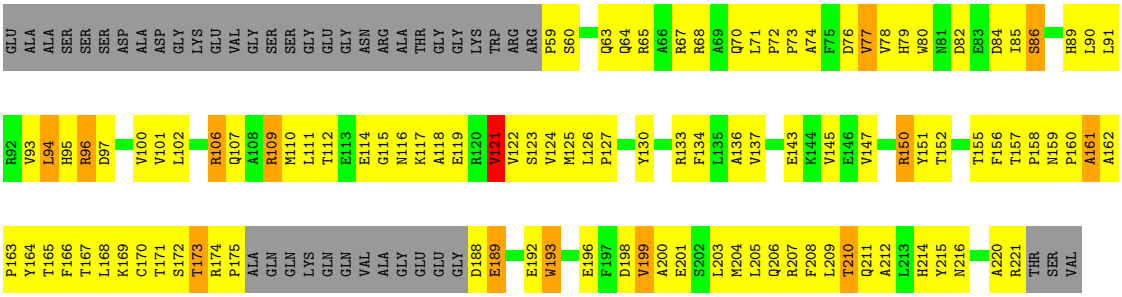




• Molecule 1: mitochondrial RNA-binding protein 2

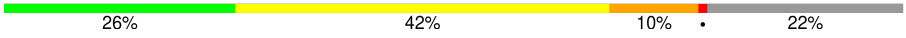
Chain P:

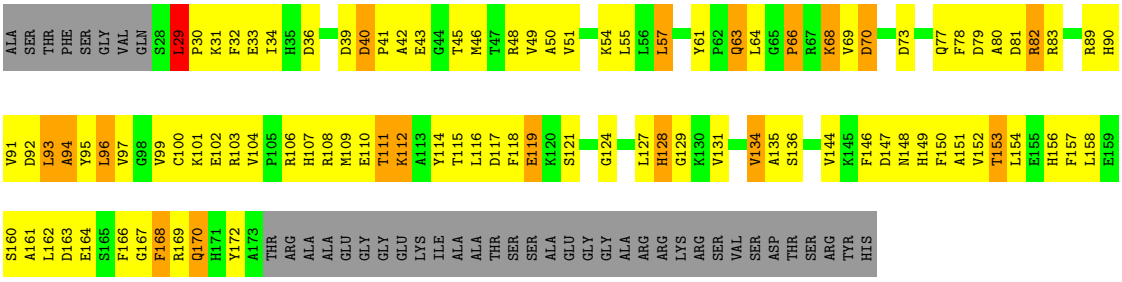




• Molecule 2: mitochondrial RNA-binding protein 1

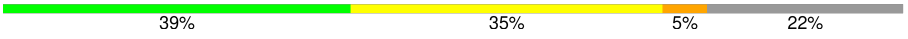
Chain B:

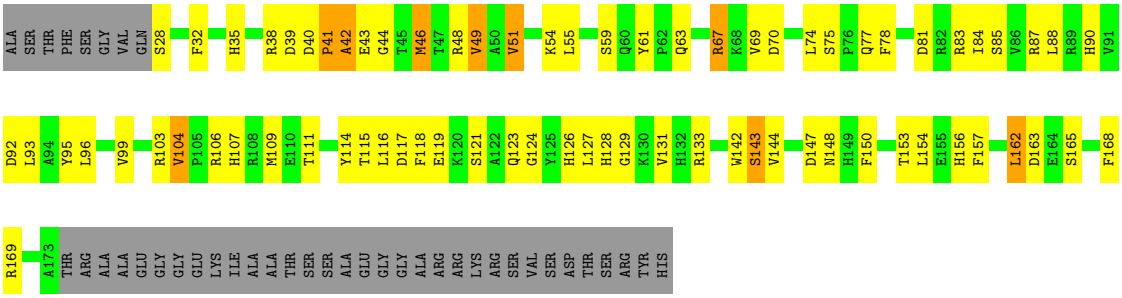




• Molecule 2: mitochondrial RNA-binding protein 1

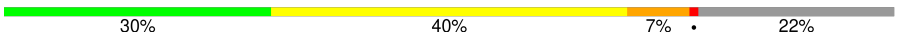
Chain D:

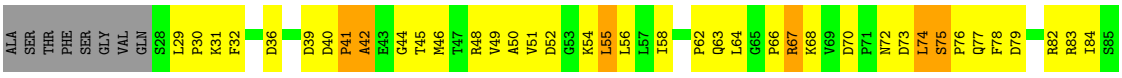


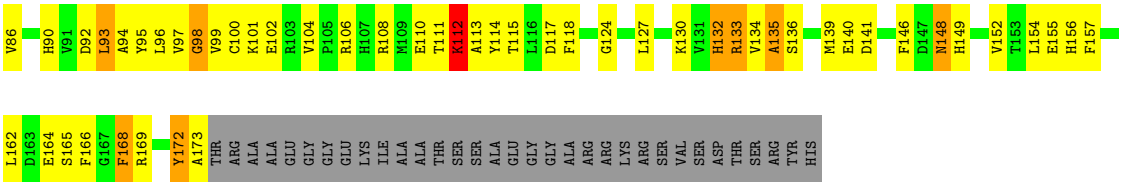


• Molecule 2: mitochondrial RNA-binding protein 1

Chain K:

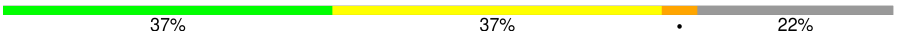


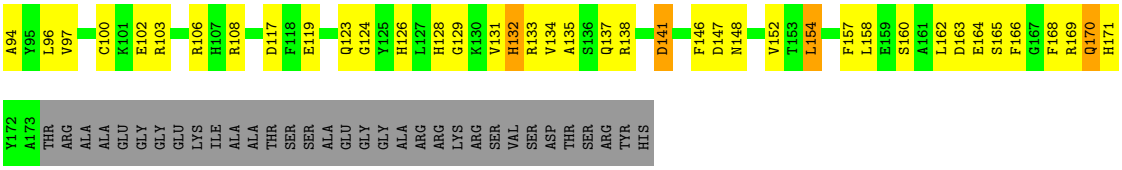




• Molecule 2: mitochondrial RNA-binding protein 1

Chain J:





4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	235.99Å 235.99Å 85.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	78.84 – 3.35	Depositor
% Data completeness (in resolution range)	99.2 (78.84-3.35)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.233 , 0.299	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9747	wwPDB-VP
Average B, all atoms (Å^2)	57.0	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	0.44	0/1250	0.73	0/1693
1	G	0.44	0/1255	0.70	1/1700 (0.1%)
1	H	0.45	0/1258	0.72	0/1704
1	P	0.47	0/1255	0.72	0/1700
2	B	0.47	0/1222	0.79	1/1645 (0.1%)
2	D	0.49	0/1222	0.74	0/1645
2	J	0.48	0/1222	0.76	0/1645
2	K	0.49	0/1222	0.77	0/1645
All	All	0.47	0/9906	0.74	2/13377 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	29	LEU	CA-CB-CG	9.37	136.84	115.30
1	G	86	SER	N-CA-C	-5.39	96.45	111.00

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	1224	0	1195	120	0
1	G	1228	0	1201	89	0
1	H	1231	0	1203	103	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	P	1228	0	1201	120	0
2	B	1195	0	1164	99	0
2	D	1195	0	1164	65	0
2	J	1195	0	1164	72	0
2	K	1195	0	1164	86	0
3	A	4	0	0	0	0
3	B	11	0	0	0	0
3	D	6	0	0	0	0
3	G	1	0	0	0	0
3	H	4	0	0	0	0
3	J	7	0	0	0	0
3	K	8	0	0	0	0
3	P	15	0	0	0	0
All	All	9747	0	9456	711	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 711 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:H:90:LEU:HD13	1:H:91:LEU:N	1.71	1.05
1:G:100:VAL:HB	1:G:126:LEU:HB2	1.44	0.99
1:G:113:GLU:HG3	1:G:114:GLU:H	1.27	0.99
1:A:82:ASP:HA	1:A:111:LEU:HD22	1.45	0.99
1:H:90:LEU:HD13	1:H:91:LEU:H	1.19	0.98

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	146/195 (75%)	104 (71%)	32 (22%)	10 (7%)	1 9
1	G	147/195 (75%)	111 (76%)	19 (13%)	17 (12%)	0 2
1	H	147/195 (75%)	98 (67%)	32 (22%)	17 (12%)	0 2
1	P	147/195 (75%)	105 (71%)	33 (22%)	9 (6%)	1 11
2	B	144/187 (77%)	111 (77%)	24 (17%)	9 (6%)	1 10
2	D	144/187 (77%)	121 (84%)	18 (12%)	5 (4%)	3 23
2	J	144/187 (77%)	116 (81%)	20 (14%)	8 (6%)	2 12
2	K	144/187 (77%)	102 (71%)	28 (19%)	14 (10%)	0 4
All	All	1163/1528 (76%)	868 (75%)	206 (18%)	89 (8%)	1 7

5 of 89 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	87	ARG
1	A	113	GLU
1	A	128	ALA
1	A	219	PHE
2	B	70	ASP

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	131/159 (82%)	115 (88%)	16 (12%)	5 19
1	G	131/159 (82%)	116 (88%)	15 (12%)	5 22
1	H	132/159 (83%)	118 (89%)	14 (11%)	6 26
1	P	131/159 (82%)	119 (91%)	12 (9%)	9 32
2	B	129/155 (83%)	112 (87%)	17 (13%)	4 17
2	D	129/155 (83%)	114 (88%)	15 (12%)	5 21
2	J	129/155 (83%)	121 (94%)	8 (6%)	18 49
2	K	129/155 (83%)	114 (88%)	15 (12%)	5 21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1041/1256 (83%)	929 (89%)	112 (11%)	6 25

5 of 112 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	145	VAL
1	P	157	THR
1	H	150	ARG
1	P	150	ARG
2	J	170	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	J	137	GLN
2	J	170	GLN
1	P	116	ASN
2	D	90	HIS
2	D	77	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 monosaccharides 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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.