Full wwPDB X-ray Structure Validation Report ⓘ

Oct 6, 2025 – 10:12 AM JST

PDB ID	:	9JVE / pdb_00009jve
Title	:	Crysral structure of 2-keto-3-deoxypentonate 4-dehydrogenase from Herbaspirillum huttiense (apo form)
Authors	:	Watanabe, S.; Akagashi, M.
Deposited on	:	2024-10-09
Resolution	:	2.27 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

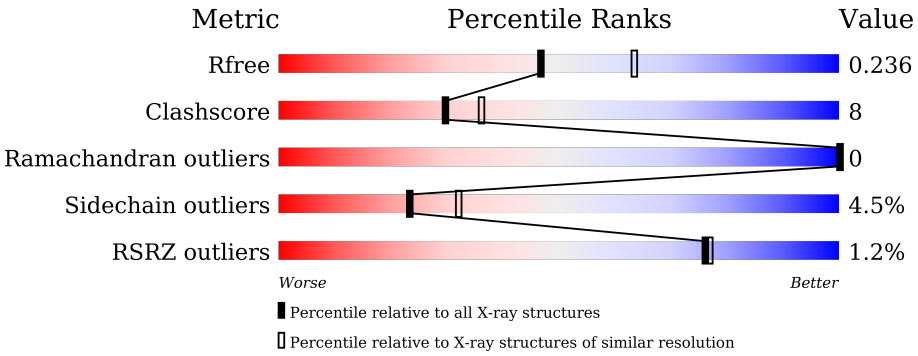
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

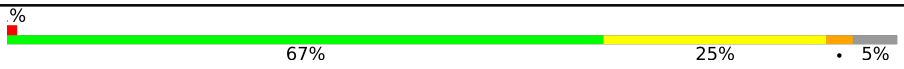
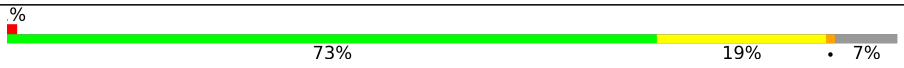
The reported resolution of this entry is 2.27 Å.

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}	164625	8487 (2.30-2.26)
Clashscore	180529	9437 (2.30-2.26)
Ramachandran outliers	177936	9341 (2.30-2.26)
Sidechain outliers	177891	9342 (2.30-2.26)
RSRZ outliers	164620	8487 (2.30-2.26)

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% 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	261	
1	B	261	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3802 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 SDR family NAD(P)-dependent oxidoreductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	248	Total	C	N	O	S	0	1	0
			1846	1159	325	356	6			
1	B	242	Total	C	N	O	S	0	1	0
			1782	1120	312	344	6			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	MET	-	initiating methionine	UNP A0AAJ2H7U7
A	-9	ARG	-	expression tag	UNP A0AAJ2H7U7
A	-8	GLY	-	expression tag	UNP A0AAJ2H7U7
A	-7	SER	-	expression tag	UNP A0AAJ2H7U7
A	-6	HIS	-	expression tag	UNP A0AAJ2H7U7
A	-5	HIS	-	expression tag	UNP A0AAJ2H7U7
A	-4	HIS	-	expression tag	UNP A0AAJ2H7U7
A	-3	HIS	-	expression tag	UNP A0AAJ2H7U7
A	-2	HIS	-	expression tag	UNP A0AAJ2H7U7
A	-1	HIS	-	expression tag	UNP A0AAJ2H7U7
A	0	GLY	-	expression tag	UNP A0AAJ2H7U7
A	1	SER	-	expression tag	UNP A0AAJ2H7U7
A	42	ASP	GLU	conflict	UNP A0AAJ2H7U7
A	72	GLN	ARG	conflict	UNP A0AAJ2H7U7
B	-10	MET	-	initiating methionine	UNP A0AAJ2H7U7
B	-9	ARG	-	expression tag	UNP A0AAJ2H7U7
B	-8	GLY	-	expression tag	UNP A0AAJ2H7U7
B	-7	SER	-	expression tag	UNP A0AAJ2H7U7
B	-6	HIS	-	expression tag	UNP A0AAJ2H7U7
B	-5	HIS	-	expression tag	UNP A0AAJ2H7U7
B	-4	HIS	-	expression tag	UNP A0AAJ2H7U7
B	-3	HIS	-	expression tag	UNP A0AAJ2H7U7
B	-2	HIS	-	expression tag	UNP A0AAJ2H7U7
B	-1	HIS	-	expression tag	UNP A0AAJ2H7U7
B	0	GLY	-	expression tag	UNP A0AAJ2H7U7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	SER	-	expression tag	UNP A0AAJ2H7U7
B	42	ASP	GLU	conflict	UNP A0AAJ2H7U7
B	72	GLN	ARG	conflict	UNP A0AAJ2H7U7

- Molecule 2 is water.

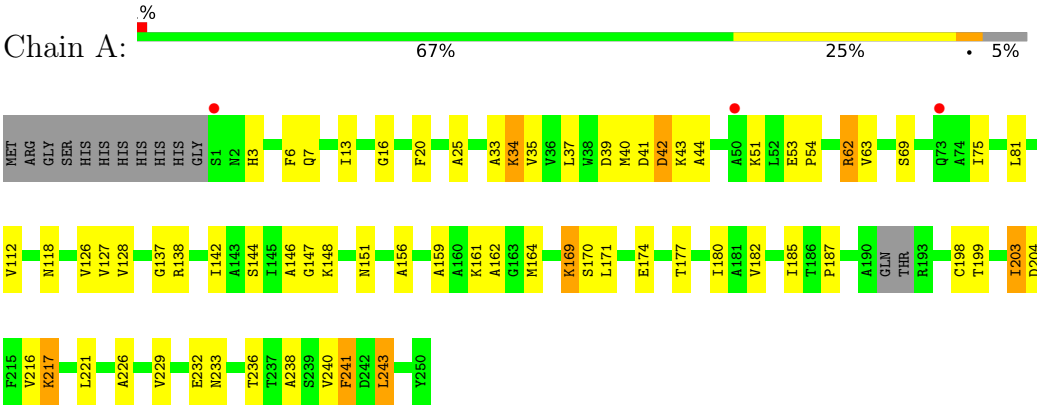
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	87	Total O 87 87	0	0
2	B	87	Total O 87 87	0	0

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 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.

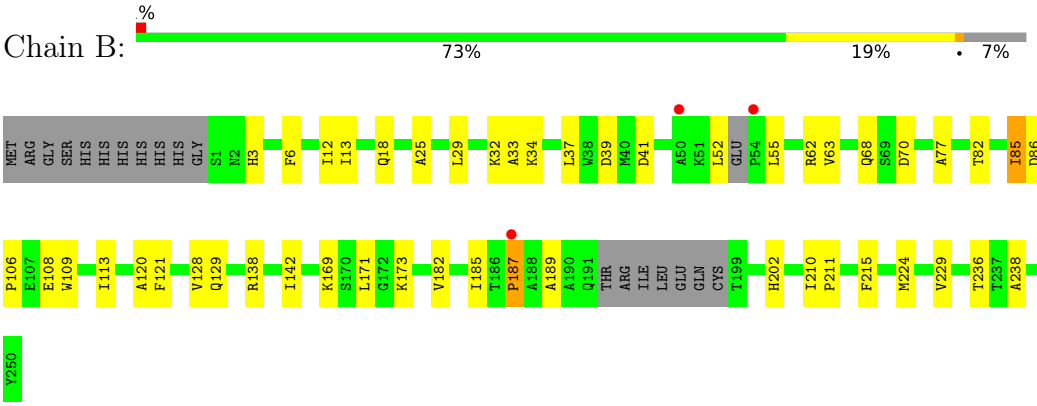
• Molecule 1: SDR family NAD(P)-dependent oxidoreductase

Chain A:



• Molecule 1: SDR family NAD(P)-dependent oxidoreductase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	76.94Å 84.95Å 69.72Å 90.00° 103.36° 90.00°	Depositor
Resolution (Å)	47.49 – 2.27 47.49 – 2.27	Depositor EDS
% Data completeness (in resolution range)	98.4 (47.49-2.27) 98.4 (47.49-2.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.44 (at 2.27Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.219 , 0.234 0.221 , 0.236	Depositor DCC
R_{free} test set	952 reflections (4.70%)	wwPDB-VP
Wilson B-factor (Å^2)	13.2	Xtriage
Anisotropy	0.821	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 33.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	3802	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.12% 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 (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	1.60	21/1873 (1.1%)	1.63	10/2546 (0.4%)
1	B	1.45	14/1808 (0.8%)	1.58	9/2458 (0.4%)
All	All	1.53	35/3681 (1.0%)	1.61	19/5004 (0.4%)

All (35) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	151	ASN	C-N	12.09	1.48	1.33
1	A	207	LEU	C-O	-6.80	1.15	1.24
1	A	187	PRO	C-O	-6.54	1.15	1.23
1	A	112	VAL	C-O	-6.40	1.16	1.24
1	A	146	ALA	C-O	-6.36	1.16	1.24
1	A	3	HIS	C-O	-6.33	1.16	1.24
1	A	35	VAL	C-O	-6.16	1.17	1.24
1	B	106	PRO	C-O	-5.96	1.16	1.24
1	A	147	GLY	C-O	-5.94	1.16	1.23
1	B	3	HIS	C-O	-5.94	1.16	1.24
1	A	156	ALA	C-O	-5.88	1.17	1.24
1	A	164	MET	C-O	-5.76	1.17	1.24
1	B	187	PRO	C-O	-5.67	1.16	1.23
1	A	241	PHE	C-O	-5.66	1.17	1.24
1	A	174	GLU	CD-OE1	-5.65	1.14	1.25
1	B	129	GLN	C-O	-5.64	1.17	1.24
1	A	144	SER	C-O	-5.63	1.17	1.23
1	A	161	LYS	C-O	-5.63	1.17	1.24
1	B	243	LEU	C-O	-5.62	1.18	1.23
1	A	177	THR	C-O	-5.60	1.16	1.24
1	A	226	ALA	C-O	-5.51	1.17	1.24
1	B	86	ASP	C-O	-5.51	1.16	1.24
1	B	120	ALA	C-O	-5.50	1.17	1.24
1	B	108	GLU	C-O	-5.50	1.17	1.24
1	A	204	ASP	C-O	-5.37	1.17	1.24
1	A	69	SER	C-O	-5.28	1.17	1.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	37	LEU	C-O	-5.27	1.17	1.23
1	B	90	HIS	C-O	-5.25	1.17	1.24
1	B	210	ILE	N-CA	5.22	1.50	1.46
1	A	233	ASN	C-O	-5.19	1.18	1.23
1	B	55	LEU	C-O	-5.18	1.18	1.24
1	B	91	SER	C-O	-5.17	1.17	1.24
1	A	159	ALA	C-O	-5.16	1.18	1.24
1	B	211	PRO	C-O	-5.15	1.17	1.24
1	A	232	GLU	C-O	-5.07	1.17	1.24

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	151	ASN	CA-C-N	10.18	130.83	119.83
1	A	151	ASN	C-N-CA	10.18	130.83	119.83
1	A	42	ASP	CA-CB-CG	6.89	119.49	112.60
1	A	151	ASN	O-C-N	-6.55	115.61	121.31
1	A	118	ASN	CA-CB-CG	5.85	118.45	112.60
1	B	41	ASP	CA-CB-CG	5.79	118.39	112.60
1	B	70	ASP	CA-CB-CG	5.72	118.32	112.60
1	B	34	LYS	CB-CA-C	5.71	119.14	109.84
1	A	62	ARG	CB-CA-C	5.50	118.83	109.80
1	A	170	SER	CA-C-N	5.43	127.50	120.44
1	A	170	SER	C-N-CA	5.43	127.50	120.44
1	B	77	ALA	CA-C-N	5.34	127.44	120.28
1	B	77	ALA	C-N-CA	5.34	127.44	120.28
1	B	202	HIS	CA-CB-CG	-5.24	108.56	113.80
1	A	162	ALA	CA-C-N	5.20	125.76	119.98
1	A	162	ALA	C-N-CA	5.20	125.76	119.98
1	B	121	PHE	CA-C-O	-5.16	115.08	120.55
1	B	82	THR	CA-C-N	5.01	127.25	120.44
1	B	82	THR	C-N-CA	5.01	127.25	120.44

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	1846	0	1806	33	0
1	B	1782	0	1733	27	0
2	A	87	0	0	1	0
2	B	87	0	0	0	0
All	All	3802	0	3539	60	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 8.

All (60) 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:142:ILE:HG12	1:A:185:ILE:HD11	1.32	1.09
1:B:142:ILE:HG12	1:B:185:ILE:HD11	1.35	1.03
1:B:224[A]:MET:CE	1:B:241:PHE:HB3	1.97	0.93
1:B:224[A]:MET:HE3	1:B:241:PHE:HB3	1.55	0.87
1:A:216:VAL:HG21	1:A:243:LEU:CD2	2.03	0.86
1:A:7:GLN:HE21	1:A:7:GLN:HA	1.43	0.82
1:A:216:VAL:HG21	1:A:243:LEU:HD21	1.63	0.80
1:B:224[A]:MET:CE	1:B:241:PHE:CB	2.64	0.76
1:A:41:ASP:OD2	1:A:44:ALA:N	2.20	0.73
1:B:6:PHE:HB3	1:B:33:ALA:HB2	1.71	0.72
1:A:142:ILE:HG12	1:A:185:ILE:CD1	2.16	0.72
1:A:216:VAL:HG21	1:A:243:LEU:HD23	1.73	0.69
1:A:7:GLN:HA	1:A:7:GLN:NE2	2.07	0.69
1:A:53:GLU:N	1:A:54:PRO:HD2	2.10	0.66
1:B:224[A]:MET:HE1	1:B:241:PHE:HB3	1.78	0.66
1:A:34:LYS:HE2	1:A:81:LEU:O	1.97	0.64
1:B:13:ILE:HD12	1:B:25:ALA:HB2	1.80	0.64
1:B:185:ILE:HG22	1:B:224[A]:MET:HE2	1.81	0.62
1:A:42:ASP:OD1	1:A:62:ARG:HD3	2.00	0.62
1:A:20:PHE:CD1	1:A:221:LEU:HD13	2.36	0.61
1:B:224[A]:MET:HE3	1:B:241:PHE:CB	2.27	0.60
1:B:12:ILE:HD11	1:B:85:ILE:HD13	1.84	0.59
1:B:185:ILE:HA	1:B:241:PHE:O	2.03	0.59
1:A:138:ARG:HB3	1:A:229:VAL:HA	1.84	0.58
1:A:137:GLY:O	1:A:180:ILE:HA	2.07	0.55
1:A:148:LYS:HE3	1:A:240:VAL:HG11	1.89	0.54
1:A:6:PHE:CB	1:A:33:ALA:HB2	2.38	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:198:CYS:HB2	1:A:203[A]:ILE:HD11	1.90	0.53
1:A:6:PHE:HB2	1:A:33:ALA:HB2	1.90	0.52
2:A:302:HOH:O	1:B:173:LYS:HE2	2.09	0.52
1:A:7:GLN:NE2	1:A:7:GLN:CA	2.72	0.51
1:B:236:THR:HG21	1:B:241:PHE:CZ	2.46	0.51
1:A:16:GLY:HA3	1:A:37:LEU:HD22	1.94	0.49
1:B:187:PRO:HA	1:B:243:LEU:HB3	1.94	0.48
1:A:13:ILE:HD12	1:A:25:ALA:HB2	1.96	0.48
1:B:6:PHE:CB	1:B:33:ALA:HB2	2.43	0.48
1:A:42:ASP:OD1	1:A:62:ARG:CD	2.63	0.47
1:B:171:LEU:HD23	1:B:182:VAL:HG11	1.97	0.46
1:A:199:THR:O	1:A:203[A]:ILE:HD12	2.16	0.46
1:B:138:ARG:HB3	1:B:229:VAL:HA	1.97	0.46
1:B:169:LYS:HA	1:B:238:ALA:HB1	1.98	0.45
1:A:214:ARG:NH1	1:A:217:LYS:HE3	2.32	0.45
1:A:142:ILE:CG1	1:A:185:ILE:HD11	2.24	0.45
1:B:29:LEU:CD1	1:B:52:LEU:HD22	2.47	0.44
1:B:39:ASP:O	1:B:62:ARG:HA	2.16	0.44
1:A:169:LYS:HA	1:A:238:ALA:HB1	1.99	0.44
1:B:236:THR:HG21	1:B:241:PHE:HZ	1.83	0.44
1:B:109:TRP:C	1:B:109:TRP:CD1	2.95	0.43
1:A:88:LEU:HB2	1:A:127:VAL:HG11	2.01	0.43
1:A:109:TRP:CD1	1:A:109:TRP:C	2.96	0.43
1:B:109:TRP:O	1:B:113:ILE:HG12	2.19	0.43
1:B:224[A]:MET:CE	1:B:241:PHE:HB2	2.46	0.42
1:A:75:ILE:HG21	1:A:126:VAL:HB	2.02	0.41
1:A:236:THR:HG21	1:A:241:PHE:CZ	2.55	0.41
1:B:187:PRO:HA	1:B:243:LEU:O	2.21	0.41
1:B:189:ALA:HB3	1:B:215:PHE:CD1	2.55	0.41
1:A:128:VAL:HG21	1:A:171:LEU:HD11	2.02	0.41
1:A:198:CYS:HB2	1:A:203[A]:ILE:CD1	2.50	0.41
1:A:236:THR:HG21	1:A:241:PHE:HZ	1.86	0.41
1:B:128:VAL:HG21	1:B:171:LEU:HD11	2.03	0.40

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	245/261 (94%)	238 (97%)	7 (3%)	0	100	100
1	B	237/261 (91%)	232 (98%)	5 (2%)	0	100	100
All	All	482/522 (92%)	470 (98%)	12 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	182/205 (89%)	170 (93%)	12 (7%)	14	17
1	B	174/205 (85%)	169 (97%)	5 (3%)	37	51
All	All	356/410 (87%)	339 (95%)	17 (5%)	23	29

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	LYS
1	A	39	ASP
1	A	40	MET
1	A	43	LYS
1	A	51	LYS
1	A	63	VAL
1	A	169	LYS
1	A	182	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	203[A]	ILE
1	A	203[B]	ILE
1	A	217	LYS
1	A	243	LEU
1	B	18	GLN
1	B	32	LYS
1	B	63	VAL
1	B	68	GLN
1	B	85	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	7	GLN
1	A	30	GLN
1	A	72	GLN
1	A	98	ASN
1	B	3	HIS
1	B	18	GLN
1	B	200	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 oligosaccharides 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

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	248/261 (95%)	0.25	3 (1%)	76	77	8, 16, 30, 34	1 (0%)
1	B	242/261 (92%)	0.21	3 (1%)	76	77	6, 15, 29, 33	1 (0%)
All	All	490/522 (93%)	0.23	6 (1%)	76	77	6, 16, 30, 34	2 (0%)

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	50	ALA	2.7
1	A	50	ALA	2.5
1	A	1	SER	2.4
1	B	54	PRO	2.2
1	B	187	PRO	2.1
1	A	73	GLN	2.1

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 oligosaccharides 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.