Full wwPDB X-ray Structure Validation Report (i)

Oct 22, 2023 – 02:23 PM EDT

PDB ID	:	3AHG
Title	:	H64A mutant of Phosphoketolase from Bifidobacterium Breve complexed with a tricyclic ring form of thiamine diphosphate
Authors	:	Suzuki, R.; Katayama, T.; Kim, B.-J.; Wakagi, T.; Shoun, H.; Ashida, H.; Yamamoto, K.; Fushinobu, S.
Deposited on	:	2010-04-22
Resolution	:	1.90 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 (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:

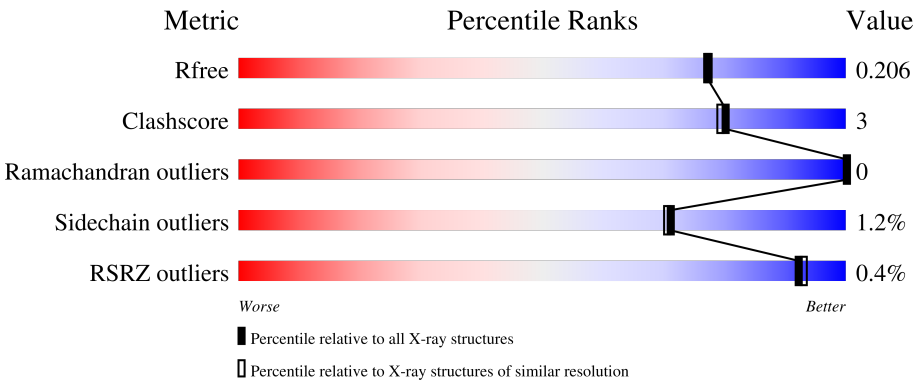
MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.90 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}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	845	84% 10% • 5%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7265 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 Xylulose 5-phosphate/fructose 6-phosphate phosphoketolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	802	6374	4050	1083	1217	24	0	0	0

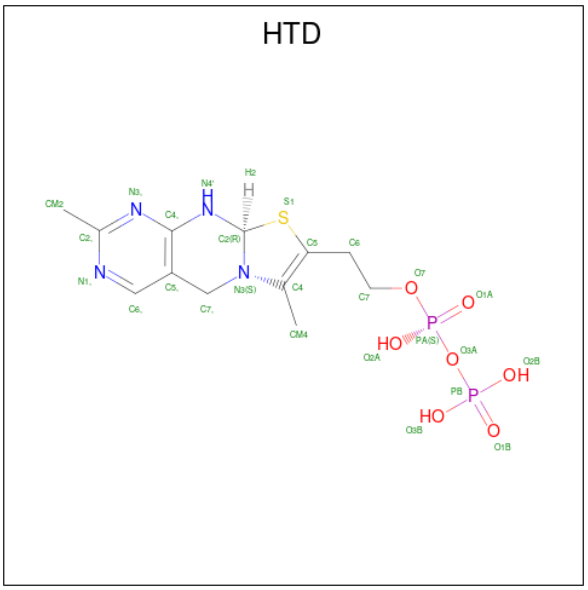
There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP D6PAH1
A	-18	GLY	-	expression tag	UNP D6PAH1
A	-17	SER	-	expression tag	UNP D6PAH1
A	-16	SER	-	expression tag	UNP D6PAH1
A	-15	HIS	-	expression tag	UNP D6PAH1
A	-14	HIS	-	expression tag	UNP D6PAH1
A	-13	HIS	-	expression tag	UNP D6PAH1
A	-12	HIS	-	expression tag	UNP D6PAH1
A	-11	HIS	-	expression tag	UNP D6PAH1
A	-10	HIS	-	expression tag	UNP D6PAH1
A	-9	SER	-	expression tag	UNP D6PAH1
A	-8	SER	-	expression tag	UNP D6PAH1
A	-7	GLY	-	expression tag	UNP D6PAH1
A	-6	LEU	-	expression tag	UNP D6PAH1
A	-5	VAL	-	expression tag	UNP D6PAH1
A	-4	PRO	-	expression tag	UNP D6PAH1
A	-3	ARG	-	expression tag	UNP D6PAH1
A	-2	GLY	-	expression tag	UNP D6PAH1
A	-1	SER	-	expression tag	UNP D6PAH1
A	0	HIS	-	expression tag	UNP D6PAH1
A	64	ALA	HIS	engineered mutation	UNP D6PAH1

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	A	1	1	1	0	0

- Molecule 3 is 2-[(9aR)-2,7-dimethyl-9a,10-dihydro-5H-pyrimido[4,5-d][1,3]thiazolo[3,2-a]pyrimidin-8-yl]ethyl trihydrogen diphosphate (three-letter code: HTD) (formula: $C_{12}H_{18}N_4O_7P_2S$).

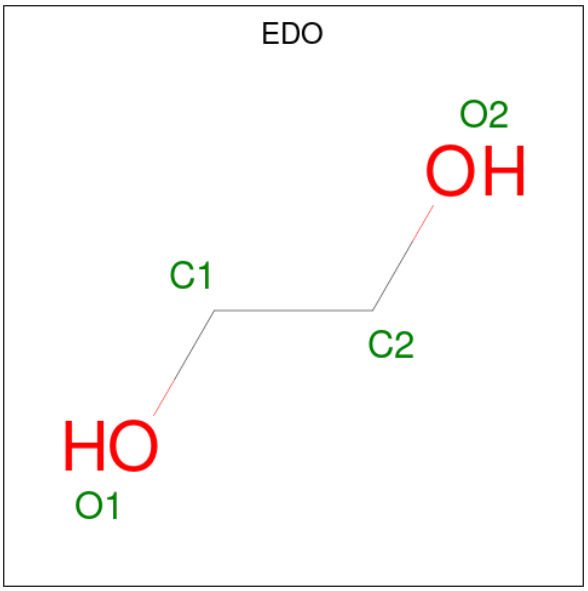


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	N	O	P	S		
3	A	1	26	12	4	7	2	1	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
4	A	1	1	1	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0

- Molecule 6 is water.

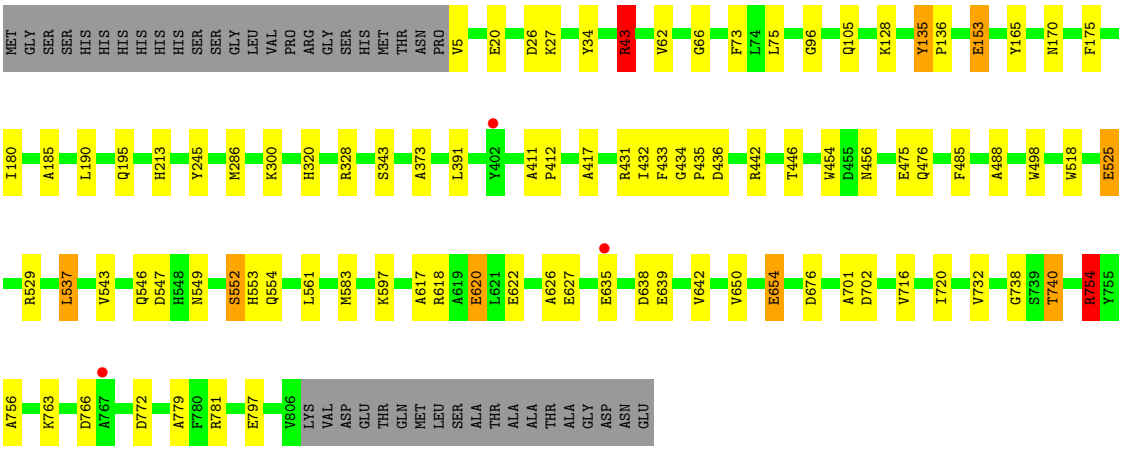
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	819	819	819	0	0

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: Xylulose 5-phosphate/fructose 6-phosphate phosphoketolase

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α, β, γ	175.03Å 175.03Å 163.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.50 – 1.90 33.50 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.1 (33.50-1.90) 98.1 (33.50-1.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.35 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.163 , 0.205 0.164 , 0.206	Depositor DCC
R_{free} test set	4849 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	24.4	Xtriage
Anisotropy	0.138	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 45.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	7265	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, EDO, MG, HTD

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	1.46	35/6547 (0.5%)	1.15	23/8901 (0.3%)

All (35) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	552	SER	CB-OG	10.34	1.55	1.42
1	A	701	ALA	CA-CB	7.87	1.69	1.52
1	A	245	TYR	CG-CD1	7.61	1.49	1.39
1	A	732	VAL	CB-CG2	7.18	1.68	1.52
1	A	373	ALA	CA-CB	6.72	1.66	1.52
1	A	642	VAL	CB-CG2	6.57	1.66	1.52
1	A	620	GLU	CB-CG	6.57	1.64	1.52
1	A	185	ALA	CA-CB	6.56	1.66	1.52
1	A	754	ARG	CD-NE	-6.54	1.35	1.46
1	A	485	PHE	CE2-CZ	6.41	1.49	1.37
1	A	165	TYR	CD1-CE1	6.33	1.48	1.39
1	A	779	ALA	CA-CB	6.21	1.65	1.52
1	A	488	ALA	CA-CB	6.18	1.65	1.52
1	A	153	GLU	CG-CD	6.00	1.60	1.51
1	A	135	TYR	CE1-CZ	5.89	1.46	1.38
1	A	195	GLN	CG-CD	5.78	1.64	1.51
1	A	20	GLU	CG-CD	5.76	1.60	1.51
1	A	417	ALA	CA-CB	5.71	1.64	1.52
1	A	654	GLU	CB-CG	5.65	1.62	1.52
1	A	525	GLU	CD-OE1	5.48	1.31	1.25
1	A	343	SER	CB-OG	5.47	1.49	1.42
1	A	433	PHE	CE2-CZ	5.41	1.47	1.37
1	A	485	PHE	CD2-CE2	5.38	1.50	1.39
1	A	135	TYR	CE2-CZ	-5.32	1.31	1.38
1	A	650	VAL	CB-CG1	5.28	1.64	1.52
1	A	175	PHE	CE2-CZ	5.24	1.47	1.37

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	43	ARG	CG-CD	5.23	1.65	1.51
1	A	62	VAL	CB-CG1	5.21	1.63	1.52
1	A	5	VAL	N-CA	5.19	1.56	1.46
1	A	180	ILE	CA-CB	-5.18	1.43	1.54
1	A	518	TRP	CE3-CZ3	5.12	1.47	1.38
1	A	34	TYR	CD1-CE1	5.06	1.47	1.39
1	A	66	GLY	N-CA	5.06	1.53	1.46
1	A	73	PHE	CG-CD1	5.06	1.46	1.38
1	A	552	SER	CA-CB	5.05	1.60	1.52

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	754	ARG	NE-CZ-NH2	-18.31	111.14	120.30
1	A	43	ARG	NE-CZ-NH2	-15.57	112.52	120.30
1	A	43	ARG	NE-CZ-NH1	13.47	127.03	120.30
1	A	754	ARG	NE-CZ-NH1	10.40	125.50	120.30
1	A	618	ARG	NE-CZ-NH1	9.68	125.14	120.30
1	A	618	ARG	NE-CZ-NH2	-8.68	115.96	120.30
1	A	529	ARG	NE-CZ-NH2	-8.01	116.29	120.30
1	A	442	ARG	NE-CZ-NH1	6.97	123.78	120.30
1	A	537	LEU	CB-CG-CD2	-6.66	99.68	111.00
1	A	43	ARG	CD-NE-CZ	6.64	132.89	123.60
1	A	26	ASP	CB-CG-OD1	6.58	124.23	118.30
1	A	26	ASP	CB-CG-OD2	-6.33	112.60	118.30
1	A	547	ASP	CB-CG-OD2	-6.25	112.68	118.30
1	A	128	LYS	CD-CE-NZ	-6.11	97.66	111.70
1	A	676	ASP	CB-CG-OD2	-5.95	112.94	118.30
1	A	328	ARG	NE-CZ-NH2	5.70	123.15	120.30
1	A	436	ASP	CB-CG-OD1	5.68	123.41	118.30
1	A	676	ASP	CB-CG-OD1	5.61	123.35	118.30
1	A	766	ASP	CB-CG-OD2	-5.47	113.38	118.30
1	A	772	ASP	CB-CG-OD2	-5.37	113.47	118.30
1	A	638	ASP	CB-CG-OD1	5.37	123.13	118.30
1	A	75	LEU	CA-CB-CG	-5.27	103.17	115.30
1	A	754	ARG	CB-CG-CD	-5.15	98.22	111.60

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	6374	0	6083	33	1
2	A	1	0	0	0	0
3	A	26	0	14	0	0
4	A	1	0	0	0	0
5	A	44	0	66	4	0
6	A	819	0	0	6	1
All	All	7265	0	6163	34	1

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 3.

All (34) 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:549:ASN:O	1:A:740:THR:HB	1.89	0.72
1:A:654:GLU:OE1	1:A:754:ARG:HD3	1.91	0.70
1:A:43:ARG:HD3	6:A:1504:HOH:O	2.01	0.59
1:A:432:ILE:CD1	1:A:446:THR:HG21	2.31	0.59
1:A:797:GLU:OE1	5:A:837:EDO:O1	2.21	0.59
1:A:654:GLU:OE1	1:A:754:ARG:CD	2.52	0.57
1:A:27:LYS:HE2	6:A:950:HOH:O	2.04	0.57
1:A:620:GLU:HG2	1:A:626:ALA:CB	2.34	0.57
1:A:597:LYS:HE3	6:A:1222:HOH:O	2.06	0.54
1:A:525:GLU:HB3	6:A:1387:HOH:O	2.06	0.54
1:A:702:ASP:OD1	5:A:829:EDO:O1	2.25	0.53
1:A:170:ASN:HD21	1:A:431:ARG:HH11	1.55	0.53
1:A:635:GLU:HG2	1:A:639:GLU:OE2	2.08	0.53
1:A:554:GLN:HE22	1:A:738:GLY:HA3	1.74	0.52
1:A:716:VAL:O	1:A:720:ILE:HG12	2.10	0.52
1:A:627:GLU:OE1	5:A:830:EDO:H11	2.12	0.49
1:A:546:GLN:OE1	1:A:553:HIS:HB3	2.12	0.49
1:A:583:MET:HA	1:A:617:ALA:HB1	1.95	0.48
1:A:756:ALA:HB2	1:A:781:ARG:CZ	2.44	0.48
1:A:300:LYS:HE2	1:A:320:HIS:HA	1.95	0.47
1:A:435:PRO:HA	1:A:476:GLN:O	2.16	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:135:TYR:CG	1:A:136:PRO:HD2	2.51	0.45
1:A:96:GLY:HA3	1:A:153:GLU:O	2.18	0.44
1:A:432:ILE:HD13	1:A:446:THR:HG21	1.99	0.44
1:A:498:TRP:CE2	1:A:537:LEU:HD13	2.53	0.43
1:A:411:ALA:HB3	1:A:412:PRO:HD3	2.02	0.42
1:A:454:TRP:CE2	1:A:456:ASN:HB3	2.54	0.42
1:A:434:GLY:O	1:A:475:GLU:HA	2.20	0.42
5:A:831:EDO:C1	6:A:1505:HOH:O	2.68	0.42
1:A:763:LYS:HB2	1:A:763:LYS:HE2	1.44	0.41
1:A:635:GLU:HA	1:A:635:GLU:OE1	2.21	0.41
1:A:543:VAL:HG23	1:A:554:GLN:HB3	2.02	0.40
1:A:190:LEU:O	1:A:190:LEU:HD13	2.21	0.40
1:A:622:GLU:HG2	6:A:1494:HOH:O	2.20	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:286:MET:CE	6:A:1074:HOH:O[6_565]	2.15	0.05

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	800/845 (95%)	769 (96%)	31 (4%)	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	676/711 (95%)	668 (99%)	8 (1%)	71 70

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	43	ARG
1	A	105	GLN
1	A	213	HIS
1	A	391	LEU
1	A	552	SER
1	A	561	LEU
1	A	740	THR
1	A	754	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	39	GLN
1	A	170	ASN
1	A	408	GLN
1	A	441	ASN
1	A	472	GLN
1	A	554	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)

Mogul failed to run properly - this section is therefore empty.

5.5 Carbohydrates (i)

Mogul failed to run properly - this section is therefore empty.

5.6 Ligand geometry (i)

Mogul failed to run properly - this section is therefore empty.

5.7 Other polymers (i)

Mogul failed to run properly - this section is therefore empty.

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	802/845 (94%)	-0.42	3 (0%) 92 93	17, 25, 37, 57	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	402	TYR	2.5
1	A	635	GLU	2.5
1	A	767	ALA	2.1

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 monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	EDO	A	832	4/4	0.80	0.33	52,59,60,60	0
5	EDO	A	831	4/4	0.84	0.18	57,60,62,62	0
5	EDO	A	830	4/4	0.88	0.15	54,54,56,57	0
5	EDO	A	836	4/4	0.88	0.18	55,56,57,62	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	EDO	A	837	4/4	0.91	0.13	37,38,41,50	0
5	EDO	A	835	4/4	0.92	0.14	36,38,41,43	0
5	EDO	A	833	4/4	0.95	0.15	35,40,42,42	0
5	EDO	A	829	4/4	0.95	0.20	38,39,42,45	0
5	EDO	A	839	4/4	0.96	0.09	37,41,43,45	0
4	NA	A	828	1/1	0.97	0.04	57,57,57,57	1
5	EDO	A	838	4/4	0.97	0.07	21,22,22,28	0
5	EDO	A	834	4/4	0.97	0.23	33,34,38,39	0
3	HTD	A	827	26/26	0.99	0.07	17,20,27,29	0
2	MG	A	826	1/1	1.00	0.08	20,20,20,20	0

6.5 Other polymers (i)

There are no such residues in this entry.