Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 02:52 am BST

PDB ID	:	4YXT
Title	:	PksG, a HMG-CoA Synthase from Bacillus subtilus
Authors	:	Till, M.; Nair, A.V.; Robson, A.; Race, P.R.
Deposited on	:	2015-03-23
Resolution	:	2.10 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 following versions of software and data (see references ⓘ) were used in the production of this report:

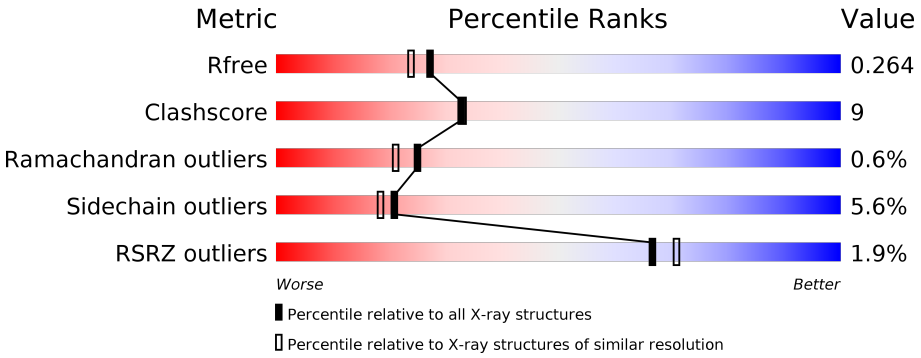
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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	420	3% 76% 18%
1	B	420	% 75% 17% 5%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6388 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 Polyketide biosynthesis 3-hydroxy-3-methylglutaryl-ACP synthase PksG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	405	Total	C	N	O	S	0	0	0
			3136	1987	528	597	24			
1	B	401	Total	C	N	O	S	0	0	0
			3107	1973	520	589	25			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	ALA	GLU	engineered mutation	UNP P40830
B	82	ALA	GLU	engineered mutation	UNP P40830

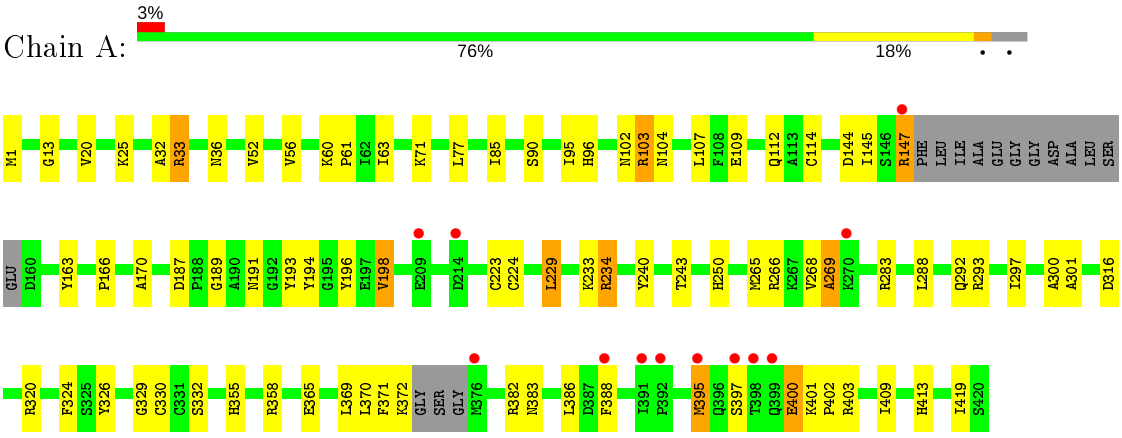
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	64	Total	O	0	0
			64	64		
2	B	81	Total	O	0	0
			81	81		

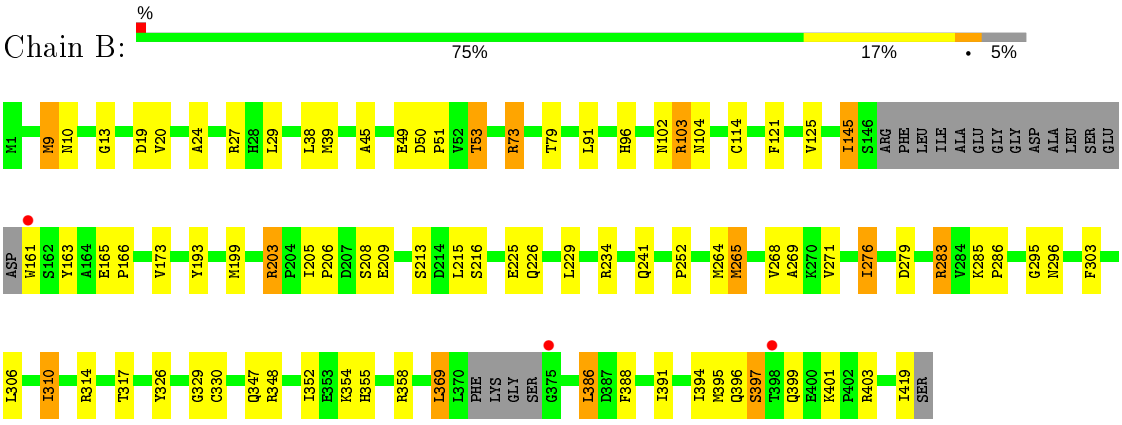
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Polyketide biosynthesis 3-hydroxy-3-methylglutaryl-ACP synthase PksG



• Molecule 1: Polyketide biosynthesis 3-hydroxy-3-methylglutaryl-ACP synthase PksG



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	81.16Å 86.45Å 107.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	64.70 – 2.10 64.70 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (64.70-2.10) 99.9 (64.70-2.10)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.89 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.198 , 0.254 0.208 , 0.264	Depositor DCC
R_{free} test set	2247 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	26.8	Xtriage
Anisotropy	0.065	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 50.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6388	wwPDB-VP
Average B, all atoms (Å^2)	31.0	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 31.62 % 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 1.0750e-03. 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.78	0/3199	0.89	5/4318 (0.1%)
1	B	0.83	0/3171	0.97	8/4280 (0.2%)
All	All	0.81	0/6370	0.93	13/8598 (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

There are no bond length outliers.

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	283	ARG	NE-CZ-NH2	-10.78	114.91	120.30
1	B	283	ARG	NE-CZ-NH1	9.34	124.97	120.30
1	B	103	ARG	NE-CZ-NH1	8.52	124.56	120.30
1	B	103	ARG	NE-CZ-NH2	-7.85	116.37	120.30
1	A	103	ARG	NE-CZ-NH2	-7.78	116.41	120.30
1	B	265	MET	CG-SD-CE	7.70	112.53	100.20
1	A	234	ARG	NE-CZ-NH1	7.42	124.01	120.30
1	A	234	ARG	NE-CZ-NH2	-6.22	117.19	120.30
1	B	314	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	A	293	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	A	103	ARG	NE-CZ-NH1	5.52	123.06	120.30
1	B	386	LEU	CA-CB-CG	-5.29	103.12	115.30
1	B	314	ARG	NE-CZ-NH2	-5.29	117.66	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	269	ALA	Peptide

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	3136	0	3049	55	0
1	B	3107	0	3032	56	0
2	A	64	0	0	3	0
2	B	81	0	0	1	0
All	All	6388	0	6081	107	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 9.

All (107) 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:268:VAL:HG11	2:A:549:HOH:O	1.78	0.81
1:A:52:VAL:HG13	1:A:95:ILE:HD13	1.66	0.78
1:B:13:GLY:H	1:B:355:HIS:HE1	1.31	0.76
1:A:187:ASP:OD2	1:A:320:ARG:NH1	2.20	0.75
1:B:53:THR:HG23	2:B:503:HOH:O	1.88	0.73
1:B:10:ASN:OD1	1:B:352:ILE:HD11	1.89	0.73
1:A:90:SER:HB2	1:A:109:GLU:HB2	1.72	0.70
1:A:20:VAL:HG13	1:A:166:PRO:HB3	1.75	0.69
1:B:49:GLU:HA	1:B:53:THR:HG21	1.75	0.67
1:A:297:ILE:HG22	1:A:300:ALA:HB3	1.78	0.66
1:A:13:GLY:H	1:A:355:HIS:HE1	1.44	0.65
1:A:52:VAL:HA	1:A:95:ILE:HD11	1.78	0.65
1:B:29:LEU:HD12	1:B:163:TYR:CE1	2.31	0.65
1:B:369:LEU:HD12	1:B:391:ILE:HD11	1.79	0.64
1:B:10:ASN:OD1	1:B:352:ILE:CD1	2.46	0.64
1:A:401:LYS:O	1:A:403:ARG:NH1	2.31	0.64
1:B:279:ASP:OD1	1:B:283:ARG:HD2	1.98	0.64
1:A:147:ARG:NE	1:A:147:ARG:H	1.95	0.63
1:B:265:MET:HE1	1:B:269:ALA:HB2	1.80	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:388:PHE:CE2	1:B:419:ILE:HD11	2.34	0.63
1:A:33:ARG:HG2	1:A:163:TYR:CD1	2.34	0.62
1:B:24:ALA:HA	1:B:29:LEU:HD23	1.81	0.62
1:B:354:LYS:HD3	1:B:358:ARG:CZ	2.30	0.62
1:B:271:VAL:HG21	1:B:276:ILE:CD1	2.30	0.61
1:A:104:ASN:HD21	1:B:234:ARG:HH21	1.49	0.61
1:A:401:LYS:HB3	1:A:402:PRO:CD	2.32	0.60
1:B:96:HIS:HE1	1:B:102:ASN:O	1.84	0.60
1:A:32:ALA:O	1:A:36:ASN:ND2	2.35	0.59
1:B:79:THR:HG21	1:B:91:LEU:HB2	1.86	0.58
1:A:297:ILE:O	1:A:297:ILE:HG22	2.04	0.57
1:B:271:VAL:HG11	1:B:276:ILE:CD1	2.35	0.56
1:A:63:ILE:CG2	1:A:71:LYS:HE3	2.34	0.56
1:B:9:MET:H	1:B:347:GLN:HE22	1.53	0.55
1:A:77:LEU:HB3	1:A:107:LEU:HD23	1.88	0.55
1:A:243:THR:HG22	1:A:320:ARG:HD2	1.90	0.54
1:B:396:GLN:O	1:B:397:SER:CB	2.54	0.54
1:B:45:ALA:HB3	1:B:145:ILE:HD13	1.89	0.54
1:B:306:LEU:O	1:B:310:ILE:HD13	2.08	0.54
1:A:103:ARG:HD3	2:A:540:HOH:O	2.09	0.53
1:B:279:ASP:OD1	1:B:283:ARG:CD	2.58	0.52
1:B:226:GLN:HA	1:B:226:GLN:HE21	1.75	0.52
1:A:229:LEU:O	1:A:233:LYS:HG3	2.09	0.52
1:A:401:LYS:HE3	1:A:402:PRO:HD3	1.92	0.51
1:B:271:VAL:HG21	1:B:276:ILE:HD11	1.91	0.51
1:A:102:ASN:ND2	1:A:104:ASN:HB3	2.26	0.51
1:B:271:VAL:HG21	1:B:276:ILE:HD13	1.91	0.51
1:B:396:GLN:O	1:B:397:SER:OG	2.28	0.51
1:A:144:ASP:C	1:A:145:ILE:HD12	2.33	0.50
1:A:234:ARG:HH21	1:B:104:ASN:HD21	1.60	0.50
1:B:203:ARG:HD3	1:B:209:GLU:OE2	2.12	0.49
1:B:96:HIS:CE1	1:B:102:ASN:O	2.66	0.49
1:B:45:ALA:CB	1:B:145:ILE:HD13	2.42	0.48
1:B:399:GLN:CB	1:B:403:ARG:NH1	2.77	0.48
1:A:388:PHE:O	1:A:395:MET:HG3	2.14	0.47
1:A:60:LYS:HB3	1:A:61:PRO:HD3	1.96	0.47
1:B:27:ARG:HB3	1:B:29:LEU:CD2	2.44	0.47
1:A:96:HIS:HE1	1:A:102:ASN:O	1.97	0.47
1:A:56:VAL:HG23	1:A:95:ILE:HD12	1.96	0.47
1:A:187:ASP:CG	1:A:320:ARG:NH1	2.68	0.47
1:A:316:ASP:HB2	2:A:550:HOH:O	2.15	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:401:LYS:HB3	1:A:402:PRO:HD2	1.96	0.47
1:B:39:MET:HE1	1:B:295:GLY:HA2	1.96	0.46
1:B:13:GLY:H	1:B:355:HIS:CE1	2.21	0.46
1:B:173:VAL:HG11	1:B:303:PHE:HB3	1.97	0.46
1:A:223:CYS:HB3	1:A:332:SER:OG	2.16	0.46
1:B:310:ILE:CD1	1:B:310:ILE:N	2.79	0.46
1:B:397:SER:OG	1:B:397:SER:O	2.33	0.46
1:A:114:CYS:HB2	1:A:326:TYR:O	2.16	0.46
1:A:240:TYR:HE2	1:A:265:MET:HE2	1.80	0.46
1:A:297:ILE:CG2	1:A:300:ALA:HB3	2.46	0.46
1:A:224:CYS:HA	1:A:324:PHE:CE1	2.50	0.46
1:A:388:PHE:CE2	1:A:419:ILE:HD11	2.51	0.45
1:A:240:TYR:CD2	1:A:265:MET:HE1	2.52	0.45
1:A:194:TYR:HA	1:B:103:ARG:O	2.17	0.45
1:A:170:ALA:C	1:A:297:ILE:HD12	2.37	0.44
1:A:250:HIS:ND1	1:A:301:ALA:HB3	2.33	0.44
1:B:121:PHE:O	1:B:125:VAL:HG23	2.17	0.44
1:A:189:GLY:C	1:A:191:ASN:HD22	2.21	0.44
1:B:252:PRO:O	1:B:296:ASN:HB2	2.17	0.44
1:B:79:THR:HG21	1:B:91:LEU:CB	2.47	0.43
1:B:20:VAL:HG13	1:B:166:PRO:HB3	2.01	0.43
1:B:394:ILE:O	1:B:396:GLN:O	2.35	0.43
1:A:329:GLY:N	1:A:330:CYS:HA	2.34	0.43
1:B:39:MET:CE	1:B:295:GLY:HA2	2.49	0.43
1:B:19:ASP:C	1:B:19:ASP:OD1	2.58	0.42
1:A:383:ASN:HA	1:A:409:ILE:O	2.19	0.42
1:B:165:GLU:N	1:B:166:PRO:CD	2.83	0.42
1:B:49:GLU:CA	1:B:53:THR:HG21	2.46	0.41
1:B:114:CYS:HB2	1:B:326:TYR:O	2.20	0.41
1:B:50:ASP:HB2	1:B:51:PRO:CD	2.50	0.41
1:B:329:GLY:N	1:B:330:CYS:HA	2.35	0.41
1:B:225:GLU:HG3	1:B:264:MET:CE	2.51	0.41
1:A:33:ARG:HG2	1:A:163:TYR:CE1	2.56	0.41
1:B:73:ARG:HD2	1:B:73:ARG:HA	1.82	0.41
1:A:196:TYR:CE2	1:A:198:VAL:HG13	2.56	0.41
1:A:240:TYR:CE2	1:A:265:MET:CE	3.04	0.41
1:B:241:GLN:NE2	1:B:269:ALA:HA	2.36	0.41
1:A:413:HIS:HB2	1:B:199:MET:HG3	2.03	0.41
1:A:370:LEU:O	1:A:372:LYS:N	2.54	0.40
1:B:205:ILE:HG22	1:B:206:PRO:HD2	2.03	0.40
1:A:266:ARG:O	1:A:269:ALA:HB2	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:370:LEU:C	1:A:372:LYS:N	2.75	0.40
1:A:52:VAL:HG13	1:A:95:ILE:CD1	2.44	0.40
1:A:265:MET:CE	1:A:283:ARG:NH2	2.85	0.40
1:A:288:LEU:O	1:A:292:GLN:HG3	2.21	0.40
1:A:196:TYR:HE2	1:A:198:VAL:CG1	2.35	0.40
1:B:285:LYS:N	1:B:286:PRO:CD	2.84	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	399/420 (95%)	373 (94%)	23 (6%)	3 (1%)	19	15
1	B	395/420 (94%)	379 (96%)	14 (4%)	2 (0%)	29	26
All	All	794/840 (94%)	752 (95%)	37 (5%)	5 (1%)	25	21

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	400	GLU
1	B	369	LEU
1	A	112	GLN
1	A	371	PHE
1	B	397	SER

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	329/347 (95%)	313 (95%)	16 (5%)	25	23
1	B	327/347 (94%)	306 (94%)	21 (6%)	17	14
All	All	656/694 (94%)	619 (94%)	37 (6%)	21	18

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	25	LYS
1	A	33	ARG
1	A	85	ILE
1	A	147	ARG
1	A	193	TYR
1	A	198	VAL
1	A	229	LEU
1	A	358	ARG
1	A	365	GLU
1	A	369	LEU
1	A	382	ARG
1	A	386	LEU
1	A	395	MET
1	A	397	SER
1	A	400	GLU
1	B	9	MET
1	B	38	LEU
1	B	53	THR
1	B	73	ARG
1	B	145	ILE
1	B	161	TRP
1	B	193	TYR
1	B	203	ARG
1	B	208	SER
1	B	213	SER
1	B	215	LEU
1	B	216	SER
1	B	229	LEU
1	B	268	VAL
1	B	276	ILE
1	B	310	ILE
1	B	317	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	348	ARG
1	B	386	LEU
1	B	395	MET
1	B	401	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	96	HIS
1	A	104	ASN
1	A	112	GLN
1	A	191	ASN
1	A	226	GLN
1	A	312	GLN
1	A	355	HIS
1	B	96	HIS
1	B	104	ASN
1	B	191	ASN
1	B	226	GLN
1	B	347	GLN
1	B	355	HIS

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	405/420 (96%)	-0.13	12 (2%) 50 56	16, 30, 59, 95	0
1	B	401/420 (95%)	-0.26	3 (0%) 87 89	14, 25, 53, 76	0
All	All	806/840 (95%)	-0.20	15 (1%) 66 71	14, 27, 56, 95	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	398	THR	5.5
1	A	376	MET	4.7
1	A	399	GLN	3.3
1	B	398	THR	3.2
1	B	161	TRP	3.0
1	B	375	GLY	2.8
1	A	395	MET	2.7
1	A	392	PRO	2.5
1	A	388	PHE	2.4
1	A	270	LYS	2.4
1	A	391	ILE	2.2
1	A	209	GLU	2.2
1	A	147	ARG	2.2
1	A	397	SER	2.1
1	A	214	ASP	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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.