Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 08:37 AM GMT

PDB ID	:	3FGB
Title	:	Crystal structure of the Q89ZH8_BACTN protein from Bacteroides thetaiotaomicron. Northeast Structural Genomics Consortium target BtR289b.
Authors	:	Vorobiev, S.M.; Lew, S.; Seetharaman, J.; Lee, D.; Foote, L.E.; Ciccosanti, C.; Janjua, H.; Xiao, R.; Acton, T.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2008-12-05
Resolution	:	2.00 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
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 (1)) were used in the production of this report:

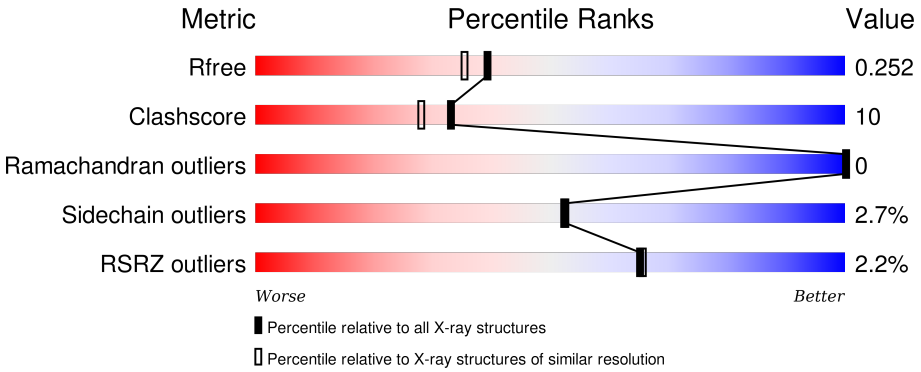
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.00 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}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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.

Mol	Chain	Length	Quality of chain
1	A	361	% 80% 15%
1	B	361	3% 80% 15%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5957 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 uncharacterized protein Q89ZH8_BACTN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	349	Total	C	N	O	S	Se	0	0	0
			2640	1670	441	523	4	2			
1	B	349	Total	C	N	O	S	Se	0	0	0
			2654	1679	443	526	4	2			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	31	MSE	-	expression tag	UNP Q89ZH8
A	384	LEU	-	expression tag	UNP Q89ZH8
A	385	GLU	-	expression tag	UNP Q89ZH8
A	386	HIS	-	expression tag	UNP Q89ZH8
A	387	HIS	-	expression tag	UNP Q89ZH8
A	388	HIS	-	expression tag	UNP Q89ZH8
A	389	HIS	-	expression tag	UNP Q89ZH8
A	390	HIS	-	expression tag	UNP Q89ZH8
A	391	HIS	-	expression tag	UNP Q89ZH8
B	31	MSE	-	expression tag	UNP Q89ZH8
B	384	LEU	-	expression tag	UNP Q89ZH8
B	385	GLU	-	expression tag	UNP Q89ZH8
B	386	HIS	-	expression tag	UNP Q89ZH8
B	387	HIS	-	expression tag	UNP Q89ZH8
B	388	HIS	-	expression tag	UNP Q89ZH8
B	389	HIS	-	expression tag	UNP Q89ZH8
B	390	HIS	-	expression tag	UNP Q89ZH8
B	391	HIS	-	expression tag	UNP Q89ZH8

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	331	Total	O	0	0
			331	331		

Continued on next page...

Continued from previous page...

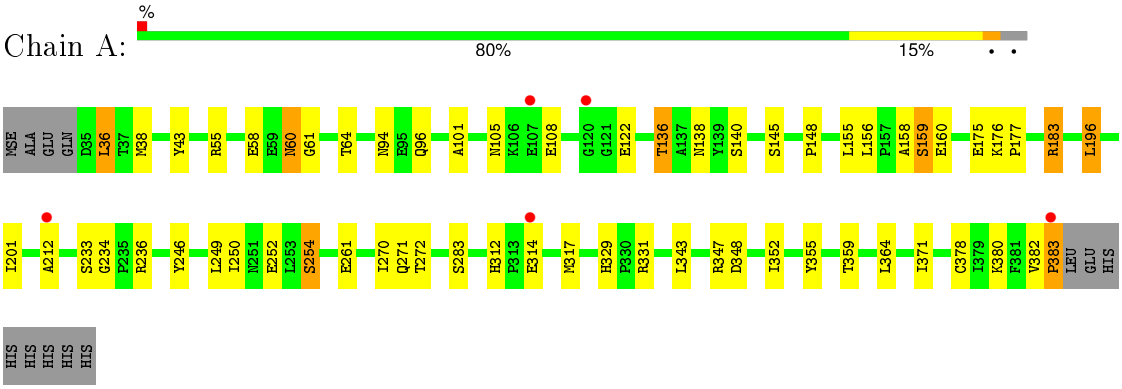
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	332	Total	O	0	0
			332	332		

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.

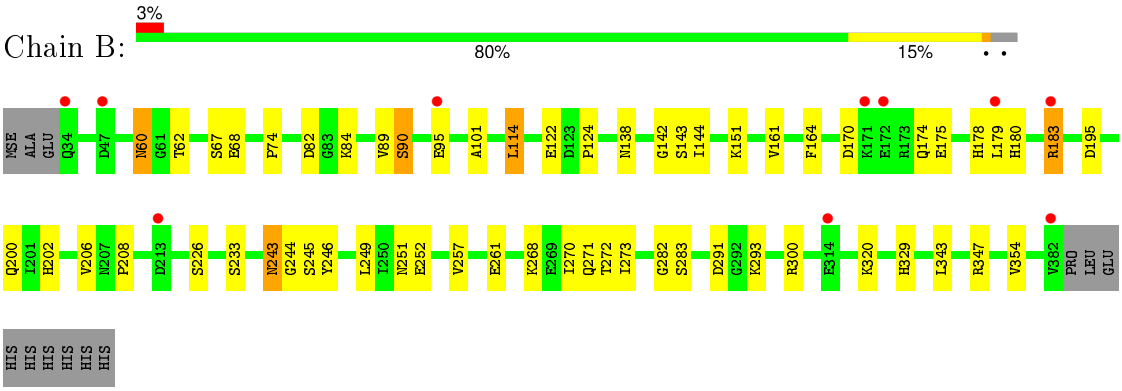
- Molecule 1: uncharacterized protein Q89ZH8_BACTN

Chain A:



- Molecule 1: uncharacterized protein Q89ZH8_BACTN

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	79.48Å 85.08Å 99.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.93 – 2.00 42.92 – 1.89	Depositor EDS
% Data completeness (in resolution range)	80.2 (42.93-2.00) 92.7 (42.92-1.89)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.79 (at 1.88Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.225 , 0.240 0.239 , 0.252	Depositor DCC
R_{free} test set	1805 reflections (4.44%)	DCC
Wilson B-factor (Å^2)	15.0	Xtriage
Anisotropy	1.068	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 42.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.31$	Xtriage
Outliers	11 of 104048 reflections (0.011%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5957	wwPDB-VP
Average B, all atoms (Å^2)	20.0	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 57.94 % 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 2.1987e-05.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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.36	0/2693	0.65	2/3653 (0.1%)
1	B	0.37	0/2707	0.66	0/3668
All	All	0.37	0/5400	0.66	2/7321 (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})
1	A	383	PRO	N-CA-CB	5.60	110.02	103.30
1	A	43	TYR	N-CA-C	-5.06	97.33	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	2640	0	2544	55	0
1	B	2654	0	2576	52	0
2	A	331	0	0	13	0
2	B	332	0	0	11	0
All	All	5957	0	5120	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 10.

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:B:144:ILE:HG23	1:B:179:LEU:HD23	1.48	0.94
1:A:36:LEU:HD11	1:A:58:GLU:HB3	1.52	0.91
1:A:196:LEU:O	1:A:196:LEU:HD12	1.78	0.82
1:B:354:VAL:HG12	2:B:688:HOH:O	1.81	0.81
1:B:268:LYS:HG3	2:B:706:HOH:O	1.81	0.80
1:B:164:PHE:CE2	1:B:179:LEU:HD21	2.17	0.79
1:A:233:SER:HA	1:A:254:SER:HB3	1.62	0.79
1:A:136:THR:CG2	1:A:138:ASN:HD21	1.97	0.78
1:B:144:ILE:HG23	1:B:179:LEU:CD2	2.13	0.78
1:B:282:GLY:O	1:B:300:ARG:HD3	1.83	0.78
1:B:243:ASN:ND2	1:B:245:SER:H	1.84	0.76
1:A:136:THR:HG23	1:A:138:ASN:HD21	1.50	0.75
1:A:261:GLU:HG3	1:A:270:ILE:HG21	1.67	0.75
1:B:243:ASN:HD22	1:B:244:GLY:N	1.85	0.74
1:A:312:HIS:HD2	1:A:314:GLU:H	1.36	0.74
1:A:136:THR:HG23	1:A:138:ASN:ND2	2.03	0.73
1:B:261:GLU:HG3	1:B:270:ILE:HG21	1.73	0.70
1:B:122:GLU:HG3	2:B:571:HOH:O	1.92	0.69
1:B:144:ILE:CG2	1:B:179:LEU:HD23	2.21	0.68
1:B:243:ASN:C	1:B:243:ASN:HD22	1.97	0.68
1:A:271:GLN:HB2	1:A:317:MSE:HE3	1.75	0.67
1:B:89:VAL:HG22	1:B:90:SER:N	2.10	0.67
1:A:55:ARG:HB2	1:A:64:THR:OG1	1.96	0.64
1:B:89:VAL:CG2	1:B:124:PRO:HG2	2.27	0.63
1:A:378:CYS:SG	2:A:852:HOH:O	2.56	0.62
1:A:60:ASN:HD22	1:A:60:ASN:C	2.03	0.62
1:A:148:PRO:HD3	1:A:159:SER:HB2	1.81	0.61
1:B:138:ASN:ND2	2:B:646:HOH:O	2.33	0.61
1:B:243:ASN:HD22	1:B:245:SER:H	1.49	0.59
1:A:175:GLU:HG2	1:A:176:LYS:HG2	1.84	0.59
1:B:82:ASP:HB3	1:B:84:LYS:HG2	1.83	0.59
1:A:246:TYR:CE1	1:A:261:GLU:HG2	2.38	0.59
1:B:60:ASN:C	1:B:60:ASN:HD22	2.06	0.58
1:A:343:LEU:N	1:A:343:LEU:HD12	2.18	0.58
1:A:312:HIS:CD2	1:A:314:GLU:H	2.20	0.58
1:A:158:ALA:HB1	2:A:651:HOH:O	2.02	0.58
1:A:122:GLU:HG3	2:A:544:HOH:O	2.01	0.58
1:B:343:LEU:N	1:B:343:LEU:HD12	2.20	0.57
1:B:291:ASP:OD1	1:B:293:LYS:HE2	2.05	0.57
1:B:183:ARG:HD2	2:B:994:HOH:O	2.04	0.56
1:A:64:THR:HG23	2:A:666:HOH:O	2.06	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:148:PRO:HG2	1:A:156:LEU:HB2	1.88	0.55
1:A:331:ARG:HG3	2:A:950:HOH:O	2.04	0.55
1:A:250:ILE:HD12	2:A:798:HOH:O	2.06	0.55
1:B:246:TYR:CE1	1:B:261:GLU:HG2	2.42	0.55
1:B:74:PRO:HA	1:B:89:VAL:O	2.07	0.54
1:A:183:ARG:NH2	2:A:402:HOH:O	2.34	0.54
1:A:359:THR:HG22	2:A:883:HOH:O	2.08	0.54
1:B:89:VAL:HG21	1:B:124:PRO:O	2.08	0.54
1:B:202:HIS:HD2	2:B:826:HOH:O	1.91	0.53
1:A:196:LEU:HD13	1:A:234:GLY:CA	2.39	0.53
1:B:329:HIS:HD2	2:B:597:HOH:O	1.92	0.52
1:A:36:LEU:CD1	1:A:58:GLU:HB3	2.33	0.50
1:B:89:VAL:HG22	1:B:90:SER:H	1.75	0.50
1:A:380:LYS:HE3	2:A:852:HOH:O	2.11	0.50
1:A:101:ALA:HB2	1:A:155:LEU:HG	1.93	0.50
1:B:89:VAL:HG21	1:B:124:PRO:HB2	1.94	0.49
1:A:136:THR:HG21	1:A:138:ASN:HD21	1.76	0.49
1:A:196:LEU:C	1:A:196:LEU:HD12	2.33	0.49
1:B:161:VAL:O	1:B:161:VAL:HG23	2.12	0.48
1:B:89:VAL:CG2	1:B:124:PRO:HB2	2.43	0.48
1:B:175:GLU:OE1	1:B:178:HIS:HE1	1.96	0.48
1:A:94:ASN:OD1	1:A:96:GLN:HB2	2.14	0.48
1:A:347:ARG:CZ	2:A:536:HOH:O	2.61	0.47
1:A:122:GLU:HB2	1:A:140:SER:HB2	1.96	0.47
1:A:261:GLU:HG3	1:A:270:ILE:CG2	2.41	0.47
1:B:89:VAL:CG2	1:B:124:PRO:CG	2.93	0.47
1:A:38:MSE:HE1	1:A:371:ILE:HD13	1.96	0.47
1:B:142:GLY:HA3	2:B:1036:HOH:O	2.14	0.47
1:B:271:GLN:HG2	1:B:272:THR:N	2.30	0.46
1:A:246:TYR:HE1	1:A:261:GLU:HG2	1.79	0.46
1:B:60:ASN:ND2	1:B:62:THR:H	2.14	0.46
1:B:89:VAL:CG2	1:B:90:SER:N	2.79	0.46
1:A:347:ARG:NE	2:A:536:HOH:O	2.48	0.46
1:A:271:GLN:HG2	1:A:272:THR:N	2.30	0.46
1:A:201:ILE:HD11	1:A:249:LEU:HD22	1.98	0.46
1:B:320:LYS:HD2	2:B:545:HOH:O	2.15	0.45
1:B:143:SER:C	1:B:179:LEU:HD22	2.37	0.45
1:A:176:LYS:HB2	1:A:177:PRO:CD	2.47	0.44
1:B:67:SER:OG	1:B:68:GLU:N	2.50	0.44
1:A:105:ASN:ND2	1:A:108:GLU:OE1	2.50	0.43
1:A:382:VAL:O	1:A:383:PRO:C	2.56	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:243:ASN:C	1:B:243:ASN:ND2	2.69	0.43
1:A:329:HIS:CD2	1:A:347:ARG:HE	2.36	0.43
1:B:101:ALA:O	1:B:114:LEU:HB2	2.18	0.43
1:B:249:LEU:O	1:B:257:VAL:HA	2.19	0.43
1:A:355:TYR:HB3	1:A:364:LEU:HB3	2.01	0.43
1:A:236:ARG:HD2	1:A:252:GLU:HB2	2.00	0.42
1:B:82:ASP:HB3	1:B:84:LYS:CG	2.48	0.42
1:A:160:GLU:N	2:A:651:HOH:O	2.52	0.42
1:A:352:ILE:HB	1:A:371:ILE:HB	2.00	0.42
1:B:200:GLN:NE2	1:B:226:SER:OG	2.52	0.42
1:A:250:ILE:CD1	2:A:798:HOH:O	2.66	0.42
1:B:195:ASP:OD1	1:B:202:HIS:HE1	2.03	0.42
1:A:252:GLU:HA	1:A:283:SER:HB2	2.00	0.42
1:A:148:PRO:HB2	1:A:212:ALA:HB2	2.02	0.42
1:A:136:THR:HG22	1:A:145:SER:HB2	2.01	0.41
1:A:60:ASN:HD22	1:A:61:GLY:N	2.17	0.41
1:B:60:ASN:C	1:B:60:ASN:ND2	2.74	0.41
1:B:233:SER:O	1:B:251:ASN:HB3	2.20	0.41
1:A:156:LEU:HD13	1:A:212:ALA:HB1	2.02	0.41
1:B:170:ASP:O	1:B:174:GLN:HB2	2.20	0.41
1:B:273:ILE:HG12	2:B:975:HOH:O	2.21	0.41
1:B:206:VAL:O	1:B:208:PRO:HD3	2.21	0.40
1:B:252:GLU:HA	1:B:283:SER:HB2	2.02	0.40
1:B:347:ARG:NH2	2:B:809:HOH:O	2.54	0.40
1:A:347:ARG:CZ	1:A:348:ASP:OD1	2.70	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	347/361 (96%)	330 (95%)	17 (5%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	347/361 (96%)	330 (95%)	17 (5%)	0	100 100
All	All	694/722 (96%)	660 (95%)	34 (5%)	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	280/297 (94%)	273 (98%)	7 (2%)	55 55
1	B	284/297 (96%)	276 (97%)	8 (3%)	51 50
All	All	564/594 (95%)	549 (97%)	15 (3%)	52 52

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	LEU
1	A	60	ASN
1	A	136	THR
1	A	159	SER
1	A	183	ARG
1	A	196	LEU
1	A	254	SER
1	B	60	ASN
1	B	90	SER
1	B	95	GLU
1	B	114	LEU
1	B	151	LYS
1	B	180	HIS
1	B	183	ARG
1	B	243	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	60	ASN
1	A	73	ASN
1	A	133	ASN
1	A	138	ASN
1	A	202	HIS
1	A	312	HIS
1	A	332	ASN
1	A	353	GLN
1	B	60	ASN
1	B	178	HIS
1	B	200	GLN
1	B	202	HIS
1	B	243	ASN
1	B	251	ASN
1	B	299	ASN
1	B	329	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	347/361 (96%)	0.38	5 (1%)	78 78	12, 19, 28, 38	0
1	B	347/361 (96%)	0.39	10 (2%)	55 56	11, 18, 27, 31	0
All	All	694/722 (96%)	0.39	15 (2%)	65 66	11, 19, 27, 38	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	120	GLY	6.7
1	B	34	GLN	5.1
1	B	171	LYS	3.6
1	B	213	ASP	2.9
1	B	47	ASP	2.9
1	B	183	ARG	2.6
1	A	383	PRO	2.5
1	B	314	GLU	2.5
1	A	314	GLU	2.4
1	B	172	GLU	2.3
1	B	179	LEU	2.3
1	A	212	ALA	2.2
1	B	382	VAL	2.2
1	B	95	GLU	2.1
1	A	107	GLU	2.0

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.