Full wwPDB X-ray Structure Validation Report ⓘ

Jul 28, 2024 – 02:18 am BST

PDB ID	:	8CA6
Title	:	Crystal structure of SARS-CoV-2 Mpro-Q189K mutant, free enzyme
Authors	:	El Kilani, H.; Hilgenfeld, R.
Deposited on	:	2023-01-24
Resolution	:	1.92 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:

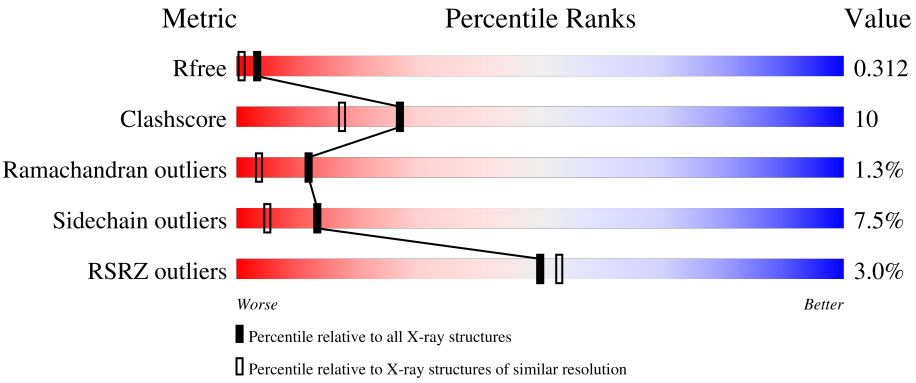
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

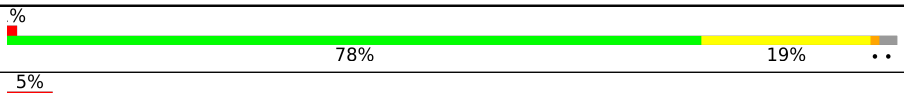
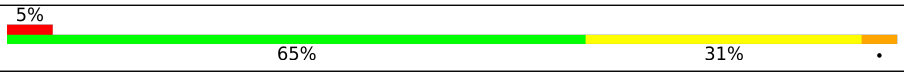
The reported resolution of this entry is 1.92 Å.

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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	AAA	306	
1	BBB	306	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4769 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 Non-structural protein 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	301	Total	C	N	O	S	0	0	0
			2329	1475	396	436	22			
1	BBB	306	Total	C	N	O	S	0	0	0
			2367	1500	402	443	22			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	189	LYS	GLN	engineered mutation	UNP P0DTC1
BBB	189	LYS	GLN	engineered mutation	UNP P0DTC1

- Molecule 2 is water.

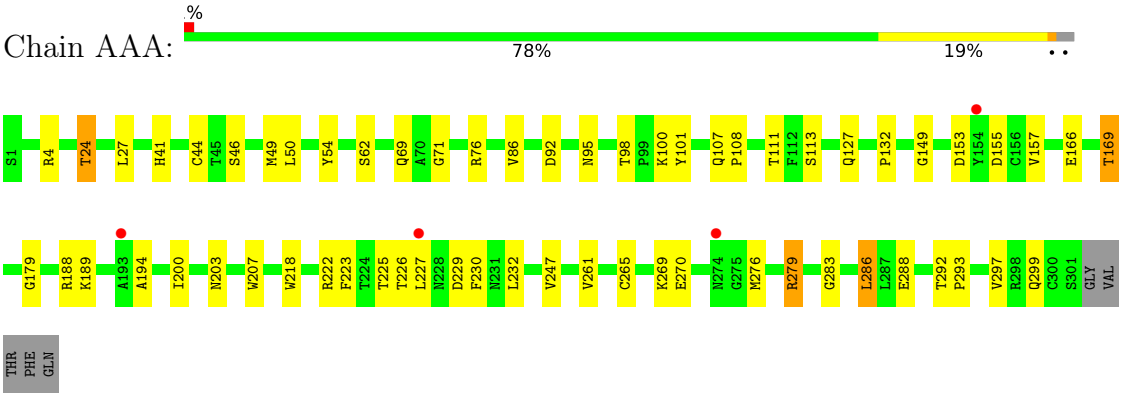
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	43	Total	O	0	0
			43	43		
2	BBB	30	Total	O	0	0
			30	30		

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.

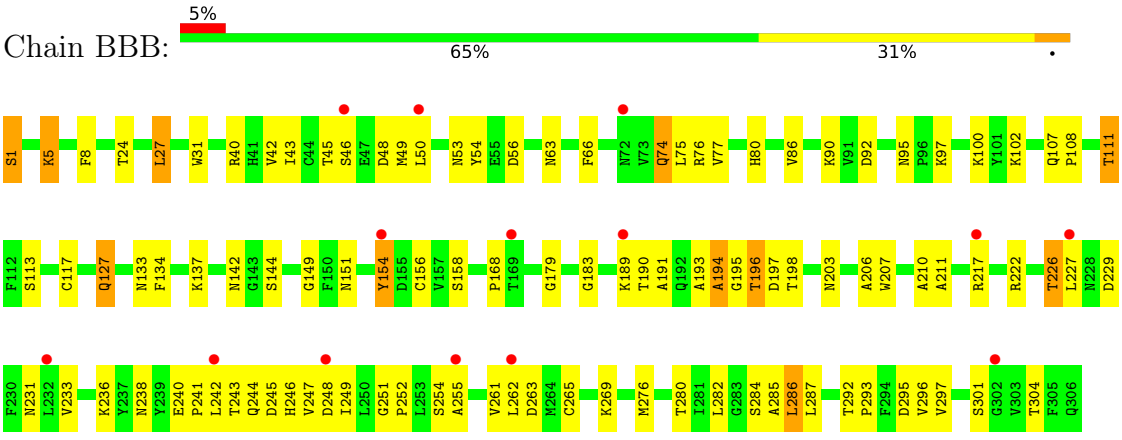
• Molecule 1: Non-structural protein 11

Chain AAA:



• Molecule 1: Non-structural protein 11

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.75Å 106.63Å 53.59Å 90.00° 104.72° 90.00°	Depositor
Resolution (Å)	46.61 – 1.92 46.61 – 1.92	Depositor EDS
% Data completeness (in resolution range)	98.3 (46.61-1.92) 98.3 (46.61-1.92)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.68 (at 1.92Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.236 , 0.294 0.252 , 0.312	Depositor DCC
R_{free} test set	1966 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	34.1	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 41.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4769	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.65% 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

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	AAA	0.72	0/2381	0.86	0/3235
1	BBB	0.74	0/2420	0.88	0/3288
All	All	0.73	0/4801	0.87	0/6523

There are no bond length outliers.

There are no bond angle outliers.

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	AAA	2329	0	2283	37	0
1	BBB	2367	0	2319	63	0
2	AAA	43	0	0	4	0
2	BBB	30	0	0	1	0
All	All	4769	0	4602	95	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 (95) 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:BBB:76:ARG:HB3	1:BBB:92:ASP:OD2	1.84	0.77

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:49:MET:HB3	1:AAA:189:LYS:HG2	1.65	0.77
1:AAA:86:VAL:HG13	1:AAA:179:GLY:HA2	1.71	0.71
1:BBB:53:ASN:ND2	1:BBB:56:ASP:OD2	2.26	0.68
1:AAA:41:HIS:HD2	2:AAA:430:HOH:O	1.79	0.65
1:BBB:252:PRO:O	1:BBB:255:ALA:HB3	1.96	0.65
1:AAA:276:MET:O	1:AAA:279:ARG:HB2	1.94	0.65
1:BBB:127:GLN:HA	1:BBB:127:GLN:HE21	1.64	0.61
1:AAA:293:PRO:O	1:AAA:297:VAL:HG23	2.02	0.60
1:BBB:203:ASN:ND2	1:BBB:292:THR:HA	2.17	0.60
1:AAA:166:GLU:OE2	1:BBB:1:SER:N	2.33	0.59
1:AAA:86:VAL:HG13	1:AAA:179:GLY:CA	2.33	0.58
1:AAA:229:ASP:HA	2:AAA:418:HOH:O	2.04	0.57
1:BBB:86:VAL:HG13	1:BBB:179:GLY:CA	2.35	0.56
1:AAA:95:ASN:HB3	1:AAA:98:THR:OG1	2.06	0.56
1:BBB:113:SER:O	1:BBB:149:GLY:HA2	2.06	0.55
1:AAA:203:ASN:ND2	1:AAA:292:THR:HA	2.21	0.55
1:BBB:286:LEU:HD13	1:BBB:287:LEU:H	1.70	0.55
1:BBB:86:VAL:HG13	1:BBB:179:GLY:HA2	1.88	0.55
1:BBB:293:PRO:O	1:BBB:297:VAL:HG23	2.08	0.54
1:BBB:243:THR:O	1:BBB:247:VAL:HG23	2.08	0.53
1:BBB:74:GLN:NE2	1:BBB:75:LEU:O	2.41	0.53
1:BBB:100:LYS:HE2	1:BBB:156:CYS:HB2	1.89	0.53
1:BBB:233:VAL:HG21	1:BBB:269:LYS:HG3	1.90	0.52
1:BBB:53:ASN:CG	1:BBB:56:ASP:OD2	2.48	0.52
1:BBB:27:LEU:HG	1:BBB:42:VAL:HG23	1.91	0.51
1:AAA:113:SER:O	1:AAA:149:GLY:HA2	2.10	0.51
1:BBB:245:ASP:O	1:BBB:249:ILE:HG13	2.10	0.50
1:BBB:198:THR:OG1	1:BBB:240:GLU:OE2	2.25	0.50
1:AAA:286:LEU:HB3	1:BBB:285:ALA:CB	2.41	0.49
1:AAA:230:PHE:CD1	1:AAA:265:CYS:HB3	2.47	0.49
1:AAA:111:THR:HG23	1:AAA:292:THR:HG23	1.94	0.49
1:AAA:169:THR:HG21	1:AAA:194:ALA:HB2	1.94	0.49
1:AAA:229:ASP:OD1	1:AAA:269:LYS:NZ	2.46	0.49
1:BBB:203:ASN:O	1:BBB:206:ALA:HB3	2.13	0.49
1:BBB:211:ALA:HA	1:BBB:282:LEU:HD21	1.95	0.48
1:AAA:225:THR:HA	2:AAA:427:HOH:O	2.13	0.48
1:BBB:80:HIS:CG	1:BBB:80:HIS:O	2.66	0.47
1:AAA:207:TRP:CG	1:AAA:288:GLU:HB3	2.50	0.47
1:AAA:4:ARG:O	1:AAA:299:GLN:NE2	2.41	0.47
1:AAA:24:THR:CG2	1:AAA:24:THR:O	2.62	0.47
1:BBB:247:VAL:HG13	1:BBB:261:VAL:HG21	1.96	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:247:VAL:HG13	1:AAA:261:VAL:HG11	1.96	0.47
1:BBB:31:TRP:CE2	1:BBB:95:ASN:HB2	2.49	0.47
1:BBB:66:PHE:HB2	1:BBB:77:VAL:HG21	1.97	0.47
1:BBB:46:SER:HB2	2:BBB:427:HOH:O	2.14	0.47
1:BBB:108:PRO:HG3	1:BBB:134:PHE:CE1	2.50	0.47
1:AAA:108:PRO:HB3	1:AAA:132:PRO:HA	1.95	0.46
1:BBB:134:PHE:HB3	1:BBB:183:GLY:N	2.31	0.46
1:BBB:243:THR:HG23	1:BBB:246:HIS:CD2	2.50	0.46
1:AAA:218:TRP:CE2	1:AAA:279:ARG:HG2	2.51	0.46
1:AAA:69:GLN:HG2	1:AAA:71:GLY:O	2.16	0.46
1:BBB:244:GLN:HA	1:BBB:247:VAL:HB	1.98	0.46
1:BBB:244:GLN:O	1:BBB:248:ASP:N	2.35	0.46
1:BBB:107:GLN:HA	1:BBB:107:GLN:HE21	1.80	0.45
1:AAA:188:ARG:HD2	2:AAA:425:HOH:O	2.17	0.45
1:BBB:195:GLY:O	1:BBB:196:THR:C	2.55	0.45
1:BBB:226:THR:O	1:BBB:229:ASP:HB3	2.16	0.45
1:BBB:8:PHE:CE2	1:BBB:151:ASN:HA	2.53	0.44
1:AAA:286:LEU:HB3	1:BBB:285:ALA:HB2	1.99	0.44
1:BBB:133:ASN:ND2	1:BBB:197:ASP:HB2	2.33	0.44
1:BBB:40:ARG:O	1:BBB:43:ILE:HG12	2.18	0.44
1:BBB:193:ALA:O	1:BBB:194:ALA:O	2.36	0.44
1:AAA:4:ARG:NE	1:BBB:137:LYS:O	2.51	0.43
1:BBB:63:ASN:OD1	1:BBB:80:HIS:ND1	2.46	0.43
1:AAA:95:ASN:HB3	1:AAA:98:THR:HG1	1.83	0.43
1:BBB:5:LYS:HD3	1:BBB:127:GLN:O	2.18	0.43
1:BBB:48:ASP:O	1:BBB:50:LEU:N	2.52	0.43
1:BBB:207:TRP:O	1:BBB:210:ALA:HB3	2.19	0.43
1:AAA:227:LEU:HD12	1:AAA:227:LEU:HA	1.88	0.42
1:AAA:218:TRP:CZ2	1:AAA:279:ARG:HB3	2.55	0.42
1:BBB:45:THR:O	1:BBB:48:ASP:N	2.53	0.42
1:AAA:44:CYS:SG	1:AAA:54:TYR:CE1	3.13	0.42
1:BBB:280:THR:HA	1:BBB:284:SER:O	2.19	0.42
1:BBB:27:LEU:HD21	1:BBB:42:VAL:HB	2.01	0.42
1:BBB:292:THR:O	1:BBB:296:VAL:HG23	2.19	0.42
1:BBB:76:ARG:CB	1:BBB:92:ASP:OD2	2.63	0.42
1:AAA:283:GLY:O	1:BBB:286:LEU:HG	2.20	0.41
1:BBB:134:PHE:HB2	1:BBB:183:GLY:CA	2.50	0.41
1:BBB:134:PHE:HB2	1:BBB:183:GLY:HA3	2.01	0.41
1:BBB:134:PHE:CB	1:BBB:183:GLY:CA	2.99	0.41
1:BBB:117:CYS:O	1:BBB:144:SER:HA	2.20	0.41
1:BBB:231:ASN:OD1	1:BBB:241:PRO:HA	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:242:LEU:HD12	1:BBB:242:LEU:HA	1.97	0.41
1:AAA:100:LYS:HD2	1:AAA:155:ASP:OD2	2.20	0.40
1:AAA:101:TYR:HA	1:AAA:157:VAL:O	2.21	0.40
1:BBB:40:ARG:CZ	1:BBB:54:TYR:CD2	3.05	0.40
1:AAA:132:PRO:HD3	1:AAA:200:ILE:HD11	2.03	0.40
1:BBB:262:LEU:HD23	1:BBB:265:CYS:SG	2.60	0.40
1:AAA:76:ARG:HB3	1:AAA:92:ASP:OD2	2.21	0.40
1:AAA:207:TRP:CD1	1:AAA:288:GLU:HB3	2.56	0.40
1:BBB:102:LYS:HG3	1:BBB:156:CYS:SG	2.62	0.40
1:BBB:111:THR:HG21	1:BBB:295:ASP:OD2	2.21	0.40
1:BBB:227:LEU:CD1	1:BBB:262:LEU:HD11	2.52	0.40
1:BBB:233:VAL:O	1:BBB:236:LYS:N	2.54	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	AAA	299/306 (98%)	278 (93%)	20 (7%)	1 (0%)	41 31
1	BBB	304/306 (99%)	275 (90%)	22 (7%)	7 (2%)	6 1
All	All	603/612 (98%)	553 (92%)	42 (7%)	8 (1%)	12 4

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	194	ALA
1	AAA	223	PHE
1	BBB	49	MET
1	BBB	190	THR
1	BBB	191	ALA
1	BBB	196	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	BBB	154	TYR
1	BBB	251	GLY

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	AAA	259/263 (98%)	244 (94%)	15 (6%)	20 10
1	BBB	263/263 (100%)	239 (91%)	24 (9%)	9 3
All	All	522/526 (99%)	483 (92%)	39 (8%)	13 5

All (39) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	24	THR
1	AAA	27	LEU
1	AAA	46	SER
1	AAA	50	LEU
1	AAA	62	SER
1	AAA	107	GLN
1	AAA	127	GLN
1	AAA	153	ASP
1	AAA	169	THR
1	AAA	222	ARG
1	AAA	226	THR
1	AAA	232	LEU
1	AAA	270	GLU
1	AAA	279	ARG
1	AAA	286	LEU
1	BBB	1	SER
1	BBB	5	LYS
1	BBB	24	THR
1	BBB	27	LEU
1	BBB	74	GLN
1	BBB	90	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	BBB	97	LYS
1	BBB	111	THR
1	BBB	127	GLN
1	BBB	142	ASN
1	BBB	154	TYR
1	BBB	158	SER
1	BBB	168	PRO
1	BBB	189	LYS
1	BBB	217	ARG
1	BBB	222	ARG
1	BBB	226	THR
1	BBB	238	ASN
1	BBB	254	SER
1	BBB	263	ASP
1	BBB	276	MET
1	BBB	286	LEU
1	BBB	301	SER
1	BBB	304	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

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	AAA	301/306 (98%)	-0.06	4 (1%)	77	79	27, 46, 69, 95	0
1	BBB	306/306 (100%)	0.23	14 (4%)	32	35	29, 50, 81, 97	0
All	All	607/612 (99%)	0.09	18 (2%)	50	53	27, 48, 76, 97	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	154	TYR	4.5
1	BBB	262	LEU	4.1
1	BBB	154	TYR	3.7
1	BBB	232	LEU	3.5
1	BBB	50	LEU	3.2
1	BBB	72	ASN	2.8
1	BBB	189	LYS	2.7
1	BBB	242	LEU	2.7
1	BBB	255	ALA	2.6
1	BBB	217	ARG	2.2
1	AAA	193	ALA	2.2
1	BBB	169	THR	2.1
1	AAA	274	ASN	2.1
1	AAA	227	LEU	2.0
1	BBB	46	SER	2.0
1	BBB	302	GLY	2.0
1	BBB	248	ASP	2.0
1	BBB	227	LEU	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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.