Full wwPDB X-ray Structure Validation Report (i)

Jan 15, 2024 – 01:10 pm GMT

PDB ID	:	6YZY
Title	:	Crystal structure of the M295V variant of Ssl1
Authors	:	Mielenbrink, S.; Olbrich, A.; Urlacher, V.; Span, I.
Deposited on	:	2020-05-07
Resolution	:	2.28 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:

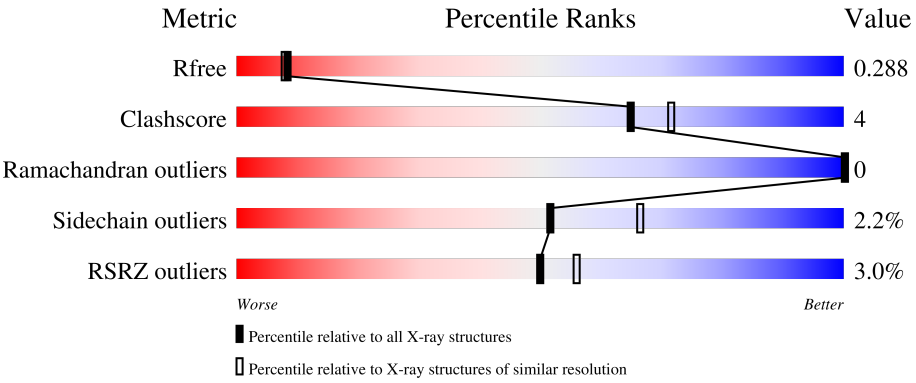
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.28 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (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	AAA	325	2% 76% 7% 17%
1	BBB	325	4% 74% 9% 17%
1	CCC	325	% 73% 10% • 17%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12346 atoms, of which 5985 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 Copper oxidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	270	Total	C	H	N	O	S	96	0	0
			4093	1309	1998	383	394	9			
1	BBB	269	Total	C	H	N	O	S	97	0	0
			4078	1304	1992	382	391	9			
1	CCC	270	Total	C	H	N	O	S	96	0	0
			4085	1306	1995	383	392	9			

There are 24 discrepancies between the modelled and reference sequences:

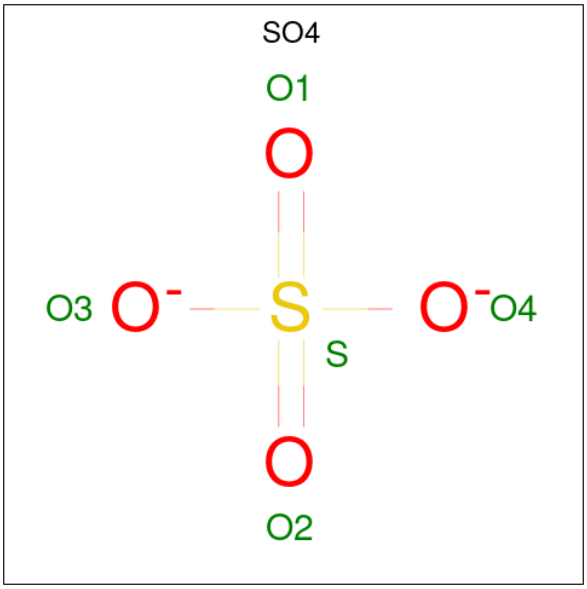
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP B5HSR1
AAA	2	HIS	-	expression tag	UNP B5HSR1
AAA	3	HIS	-	expression tag	UNP B5HSR1
AAA	4	HIS	-	expression tag	UNP B5HSR1
AAA	5	HIS	-	expression tag	UNP B5HSR1
AAA	6	HIS	-	expression tag	UNP B5HSR1
AAA	7	HIS	-	expression tag	UNP B5HSR1
AAA	295	VAL	MET	engineered mutation	UNP B5HSR1
BBB	1	MET	-	initiating methionine	UNP B5HSR1
BBB	2	HIS	-	expression tag	UNP B5HSR1
BBB	3	HIS	-	expression tag	UNP B5HSR1
BBB	4	HIS	-	expression tag	UNP B5HSR1
BBB	5	HIS	-	expression tag	UNP B5HSR1
BBB	6	HIS	-	expression tag	UNP B5HSR1
BBB	7	HIS	-	expression tag	UNP B5HSR1
BBB	295	VAL	MET	engineered mutation	UNP B5HSR1
CCC	1	MET	-	initiating methionine	UNP B5HSR1
CCC	2	HIS	-	expression tag	UNP B5HSR1
CCC	3	HIS	-	expression tag	UNP B5HSR1
CCC	4	HIS	-	expression tag	UNP B5HSR1
CCC	5	HIS	-	expression tag	UNP B5HSR1
CCC	6	HIS	-	expression tag	UNP B5HSR1
CCC	7	HIS	-	expression tag	UNP B5HSR1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	295	VAL	MET	engineered mutation	UNP B5HSR1

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total O S 5 4 1	0	0
2	BBB	1	Total O S 5 4 1	0	0
2	CCC	1	Total O S 5 4 1	0	0
2	CCC	1	Total O S 5 4 1	0	0

- Molecule 3 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	4	Total Cu 4 4	0	0
3	BBB	3	Total Cu 3 3	0	0
3	CCC	3	Total Cu 3 3	0	0

- Molecule 4 is water.

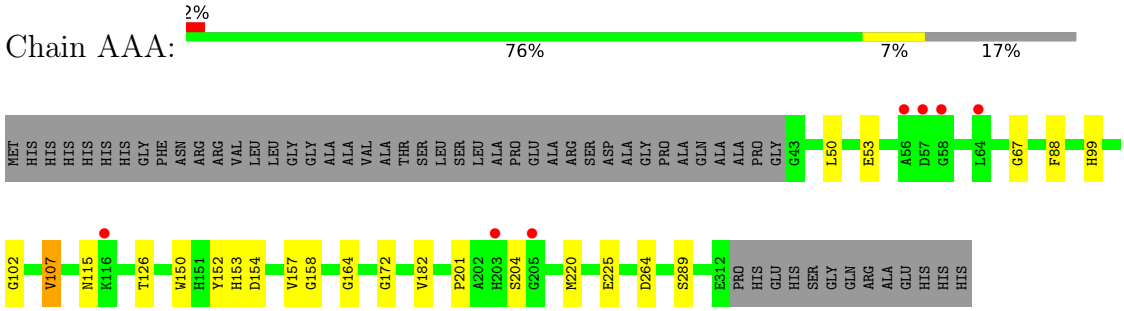
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	18	Total O 18 18	0	0
4	BBB	16	Total O 16 16	0	0
4	CCC	26	Total O 26 26	0	0

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.

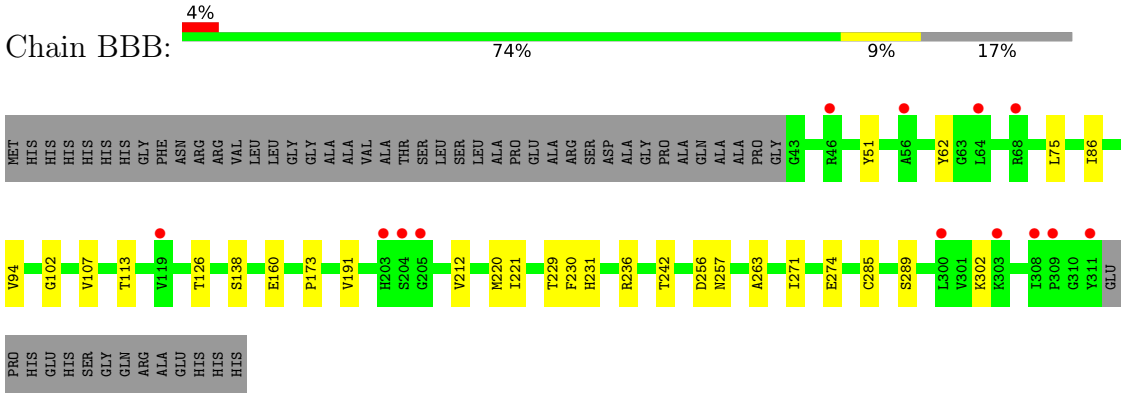
• Molecule 1: Copper oxidase

Chain AAA:



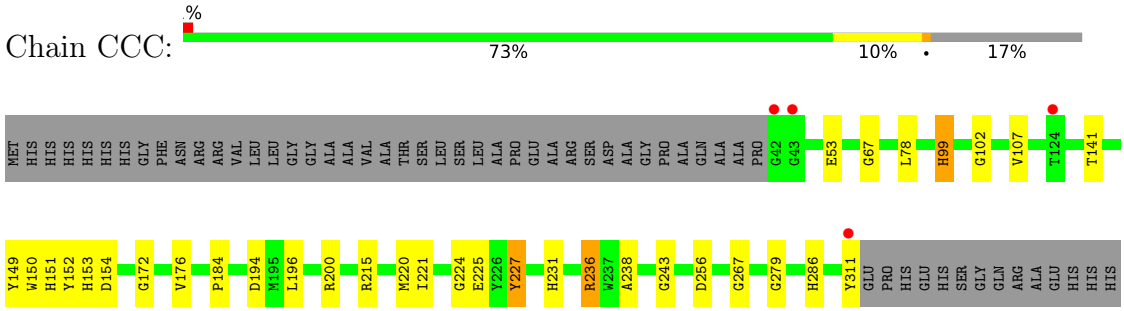
• Molecule 1: Copper oxidase

Chain BBB:



• Molecule 1: Copper oxidase

Chain CCC:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.35Å 103.96Å 161.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.99 – 2.28 48.94 – 2.28	Depositor EDS
% Data completeness (in resolution range)	98.6 (48.99-2.28) 98.7 (48.94-2.28)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.227 , 0.285 0.237 , 0.288	Depositor DCC
R_{free} test set	1994 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	32.8	Xtriage
Anisotropy	0.302	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 26.8	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	0.93	EDS
Total number of atoms	12346	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, CU

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	0.67	0/2153	0.83	0/2922
1	BBB	0.68	0/2144	0.80	0/2910
1	CCC	0.67	0/2148	0.83	0/2915
All	All	0.67	0/6445	0.82	0/8747

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	2095	1998	1978	20	0
1	BBB	2086	1992	1972	17	0
1	CCC	2090	1995	1975	23	0
2	AAA	5	0	0	0	0
2	BBB	5	0	0	0	0
2	CCC	10	0	0	0	0
3	AAA	4	0	0	0	0
3	BBB	3	0	0	0	0
3	CCC	3	0	0	0	0
4	AAA	18	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	BBB	16	0	0	0	0
4	CCC	26	0	0	1	0
All	All	6361	5985	5925	50	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 4.

All (50) 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:AAA:201:PRO:O	1:AAA:204:SER:OG	1.64	1.15
1:AAA:53:GLU:OE1	1:AAA:67:GLY:N	2.22	0.69
1:AAA:289:SER:N	1:CCC:225:GLU:OE2	2.24	0.65
1:BBB:86:ILE:O	1:BBB:126:THR:HA	2.01	0.61
1:CCC:53:GLU:OE2	1:CCC:67:GLY:N	2.35	0.56
1:AAA:102:GLY:HA3	1:AAA:150:TRP:CD2	2.40	0.56
1:AAA:158:GLY:HA3	1:AAA:164:GLY:H	1.71	0.55
1:AAA:182:VAL:HG21	1:CCC:141:THR:HG22	1.89	0.55
1:AAA:201:PRO:C	1:AAA:204:SER:OG	2.45	0.53
1:CCC:99:HIS:CD2	1:CCC:153:HIS:CE1	2.97	0.52
1:BBB:212:VAL:HG13	1:BBB:302:LYS:O	2.10	0.52
1:CCC:102:GLY:HA3	1:CCC:150:TRP:CD2	2.45	0.52
1:BBB:113:THR:HG21	1:BBB:160:GLU:OE1	2.11	0.51
1:CCC:238:ALA:O	1:CCC:243:GLY:HA2	2.11	0.50
1:AAA:99:HIS:CD2	1:AAA:153:HIS:CE1	3.00	0.49
1:AAA:225:GLU:OE1	1:BBB:289:SER:OG	2.25	0.49
1:AAA:50:LEU:O	1:AAA:88:PHE:HA	2.12	0.49
1:AAA:264:ASP:OD1	1:BBB:257:ASN:ND2	2.41	0.48
1:BBB:107:VAL:HG11	1:CCC:279:GLY:HA3	1.95	0.48
1:CCC:196:LEU:HD22	1:CCC:200:ARG:O	2.14	0.47
1:AAA:102:GLY:HA3	1:AAA:150:TRP:CE3	2.50	0.47
1:BBB:230:PHE:O	1:BBB:257:ASN:HA	2.14	0.47
1:BBB:102:GLY:O	1:CCC:236:ARG:HD3	2.14	0.47
1:BBB:263:ALA:HB2	1:CCC:286:HIS:ND1	2.30	0.47
1:AAA:107:VAL:HG13	1:AAA:115:ASN:HB3	1.97	0.46
1:CCC:78:LEU:O	1:CCC:176:VAL:HA	2.17	0.45
1:AAA:99:HIS:HD2	1:AAA:153:HIS:CE1	2.34	0.45
1:CCC:194:ASP:OD1	1:CCC:224:GLY:HA3	2.15	0.45
1:BBB:263:ALA:HB2	1:CCC:286:HIS:CE1	2.52	0.44
1:CCC:227:TYR:C	1:CCC:227:TYR:CD1	2.90	0.44
1:BBB:229:THR:O	1:BBB:285:CYS:HA	2.17	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:51:TYR:O	1:BBB:62:TYR:HA	2.17	0.44
1:CCC:99:HIS:HD2	1:CCC:153:HIS:CE1	2.35	0.43
1:BBB:271:ILE:HB	1:BBB:274:GLU:HB2	2.01	0.43
1:CCC:149:TYR:CG	1:CCC:267:GLY:HA3	2.53	0.43
1:BBB:231:HIS:HA	1:BBB:256:ASP:O	2.18	0.43
1:CCC:184:PRO:HA	1:CCC:215:ARG:HG2	2.01	0.42
1:AAA:102:GLY:O	1:BBB:236:ARG:HD3	2.19	0.42
1:AAA:154:ASP:OD2	1:AAA:157:VAL:HB	2.19	0.42
1:BBB:191:VAL:HA	1:BBB:221:ILE:O	2.18	0.42
1:CCC:151:HIS:HB3	1:CCC:221:ILE:HD13	2.01	0.42
1:AAA:152:TYR:CZ	1:AAA:172:GLY:HA3	2.54	0.42
1:AAA:53:GLU:CD	1:AAA:67:GLY:H	2.16	0.42
1:AAA:182:VAL:HG21	1:CCC:141:THR:CG2	2.49	0.42
1:CCC:236:ARG:NH1	4:CCC:504:HOH:O	2.52	0.41
1:AAA:158:GLY:HA3	1:AAA:164:GLY:N	2.34	0.41
1:CCC:152:TYR:CZ	1:CCC:172:GLY:HA3	2.56	0.41
1:CCC:149:TYR:CD2	1:CCC:267:GLY:HA3	2.55	0.41
1:BBB:75:LEU:HA	1:BBB:173:PRO:HG2	2.04	0.40
1:CCC:231:HIS:HA	1:CCC:256:ASP:O	2.21	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	268/325 (82%)	249 (93%)	19 (7%)	0	100	100
1	BBB	267/325 (82%)	254 (95%)	13 (5%)	0	100	100
1	CCC	268/325 (82%)	256 (96%)	12 (4%)	0	100	100
All	All	803/975 (82%)	759 (94%)	44 (6%)	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	AAA	217/256 (85%)	214 (99%)	3 (1%)	67 79
1	BBB	216/256 (84%)	212 (98%)	4 (2%)	57 71
1	CCC	216/256 (84%)	209 (97%)	7 (3%)	39 52
All	All	649/768 (84%)	635 (98%)	14 (2%)	52 66

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	107	VAL
1	AAA	126	THR
1	AAA	220	MET
1	BBB	94	VAL
1	BBB	138	SER
1	BBB	220	MET
1	BBB	242	THR
1	CCC	99	HIS
1	CCC	107	VAL
1	CCC	154	ASP
1	CCC	220	MET
1	CCC	227	TYR
1	CCC	236	ARG
1	CCC	311	TYR

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)

Of 14 ligands modelled in this entry, 10 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SO4	BBB	401	-	4,4,4	0.31	0	6,6,6	0.07	0
2	SO4	CCC	401	-	4,4,4	0.34	0	6,6,6	0.08	0
2	SO4	CCC	402	-	4,4,4	0.36	0	6,6,6	0.10	0
2	SO4	AAA	401	-	4,4,4	0.39	0	6,6,6	0.06	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	270/325 (83%)	0.20	7 (2%) 56 62	23, 38, 67, 112	14 (5%)
1	BBB	269/325 (82%)	0.27	13 (4%) 30 36	25, 41, 70, 107	13 (4%)
1	CCC	270/325 (83%)	-0.06	4 (1%) 73 78	20, 31, 50, 86	12 (4%)
All	All	809/975 (82%)	0.14	24 (2%) 50 56	20, 36, 66, 112	39 (4%)

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	204	SER	6.9
1	CCC	42	GLY	5.5
1	BBB	203	HIS	5.2
1	BBB	56	ALA	3.9
1	AAA	57	ASP	3.9
1	BBB	205	GLY	3.4
1	BBB	308	ILE	3.4
1	AAA	203	HIS	3.4
1	AAA	205	GLY	3.2
1	CCC	43	GLY	3.2
1	BBB	46	ARG	3.0
1	AAA	58	GLY	3.0
1	BBB	303	LYS	2.8
1	BBB	311	TYR	2.4
1	BBB	300	LEU	2.4
1	BBB	119	VAL	2.4
1	BBB	68	ARG	2.4
1	AAA	64	LEU	2.3
1	BBB	309	PRO	2.3
1	CCC	311	TYR	2.2
1	CCC	124	THR	2.1
1	AAA	116	LYS	2.1
1	BBB	64	LEU	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	AAA	56	ALA	2.0

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, 95^{th} 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
3	CU	AAA	402	1/1	0.93	0.05	101,101,101,101	0
2	SO4	CCC	402	5/5	0.94	0.13	59,60,62,64	0
2	SO4	BBB	401	5/5	0.96	0.16	50,60,61,62	0
2	SO4	AAA	401	5/5	0.97	0.13	52,53,57,64	0
2	SO4	CCC	401	5/5	0.97	0.08	47,49,50,51	0
3	CU	AAA	405	1/1	0.98	0.09	45,45,45,45	0
3	CU	BBB	402	1/1	0.98	0.13	35,35,35,35	0
3	CU	CCC	404	1/1	0.98	0.11	37,37,37,37	0
3	CU	BBB	403	1/1	0.99	0.07	44,44,44,44	0
3	CU	BBB	404	1/1	0.99	0.12	37,37,37,37	0
3	CU	CCC	403	1/1	0.99	0.10	27,27,27,27	0
3	CU	AAA	403	1/1	0.99	0.10	33,33,33,33	0
3	CU	CCC	405	1/1	0.99	0.11	40,40,40,40	0
3	CU	AAA	404	1/1	1.00	0.12	26,26,26,26	0

6.5 Other polymers (i)

There are no such residues in this entry.