



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 04:09 pm BST

PDB ID : 5XVD
Title : [NiFe]-hydrogenase (Hyb-type) from *Citrobacter* sp. S-77 in an air-oxidized condition
Authors : Nishikawa, K.; Matsuura, H.; Muhd Noor, N.D.; Tai, H.; Hirota, S.; Kim, J.; Kang, J.; Tateno, M.; Yoon, K.S.; Ogo, S.; Shomura, Y.; Higuchi, Y.
Deposited on : 2017-06-27
Resolution : 1.57 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

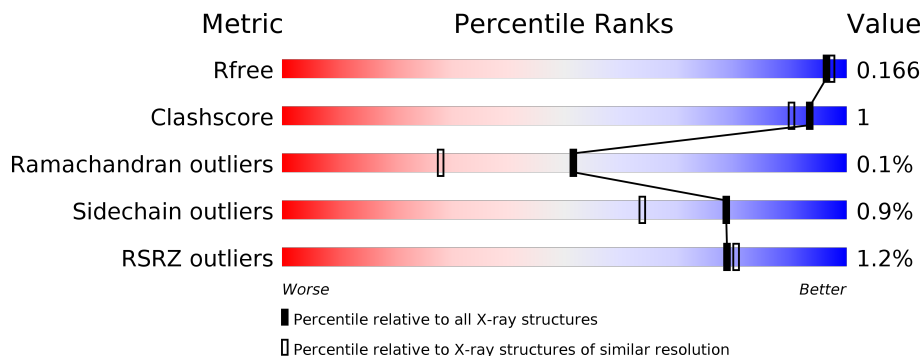
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.57 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	5534 (1.60-1.56)
Clashscore	141614	5861 (1.60-1.56)
Ramachandran outliers	138981	5708 (1.60-1.56)
Sidechain outliers	138945	5703 (1.60-1.56)
RSRZ outliers	127900	5431 (1.60-1.56)

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 $\leq 5\%$. 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	L	552	
1	M	552	
2	S	335	
2	T	335	

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 14075 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 [NiFe]-hydrogenase 2 large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	551	4361	2762	757	822	20	0	11	0
1	M	551	4325	2744	748	812	21	0	6	0

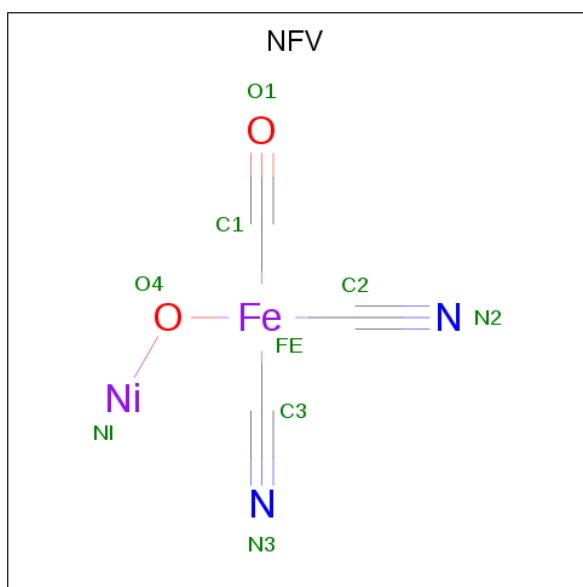
- Molecule 2 is a protein called Hydrogenase-2 small chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	S	269	2086	1317	371	381	17	0	6	0
2	T	269	2077	1312	370	378	17	0	5	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	L	1	Total	Mg	0	0
			1	1		
3	M	1	Total	Mg	0	0
			1	1		

- Molecule 4 is NI-FE OXIDIZED ACTIVE CENTER (three-letter code: NFV) (formula: C₃FeN₂NiO₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	Fe	N	Ni			O
4	L	1	9	3	1	2	1	2	0	0
4	M	1	9	3	1	2	1	2	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	L	1	6	3	3	0	0
5	L	1	6	3	3	0	0

Continued on next page...

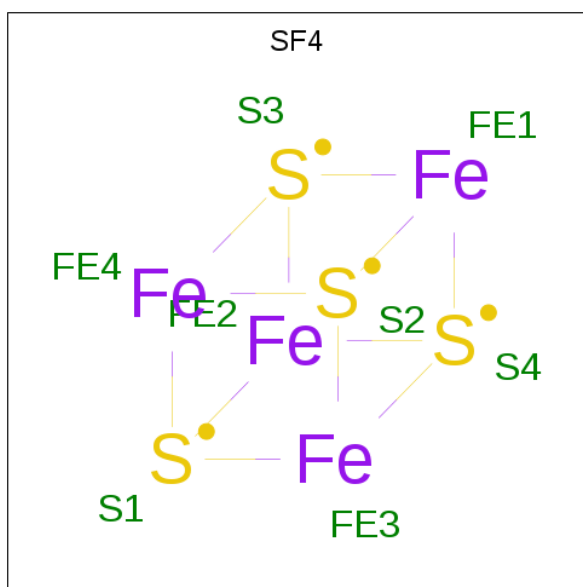
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	S	1	Total	C	O	0	0
			6	3	3		
5	T	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

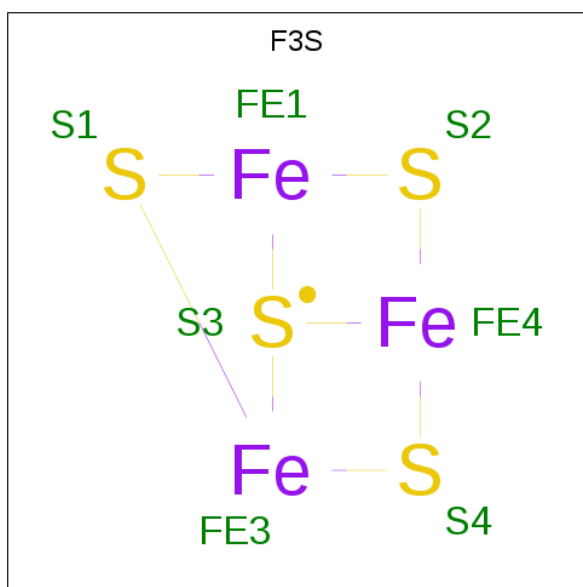
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	L	1	Total	Na	0	0
			1	1		
6	M	1	Total	Na	0	0
			1	1		

- Molecule 7 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



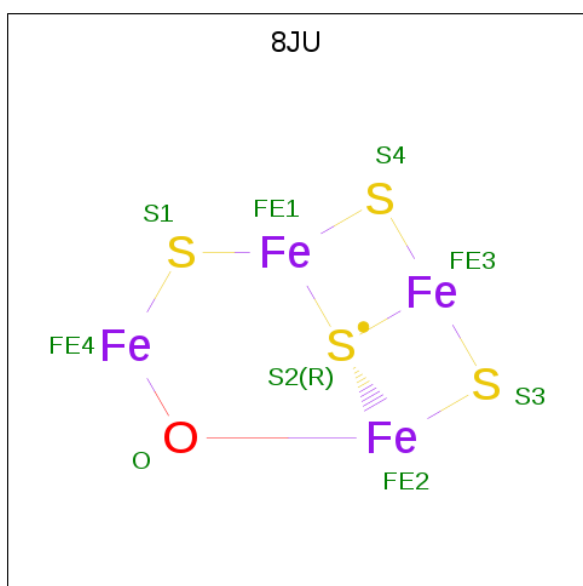
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	S	1	Total	Fe	S	0	0
			8	4	4		
7	S	1	Total	Fe	S	0	1
			8	4	4		
7	T	1	Total	Fe	S	0	0
			8	4	4		
7	T	1	Total	Fe	S	0	1
			8	4	4		

- Molecule 8 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe₃S₄).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	S	1	Total Fe S 7 3 4	0	0
8	T	1	Total Fe S 7 3 4	0	0

- Molecule 9 is FE4-S4-O CLUSTER (three-letter code: 8JU) (formula: Fe_4OS_4).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	S	1	Total Fe O S 9 4 1 4	0	1
9	T	1	Total Fe O S 9 4 1 4	0	1

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	L	424	Total 424	O 424	0	0
10	S	203	Total 203	O 203	0	0
10	M	319	Total 319	O 319	0	0
10	T	170	Total 170	O 170	0	0

VAL
ALA
GLY
VAL
SER
VAL
MET
ALA
VAL
ARG
GLU
LEU
GLY
ARG
GLN
GLN
LYS
LYS
ASP
ASN
ALA
ALA
ASP
SER
SER
ARG
GLY
GLU

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	63.94Å 118.98Å 96.81Å 90.00° 100.57° 90.00°	Depositor
Resolution (Å)	31.72 – 1.57 44.92 – 1.57	Depositor EDS
% Data completeness (in resolution range)	99.3 (31.72-1.57) 99.3 (44.92-1.57)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.43 (at 1.57Å)	Xtrriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.121 , 0.166 0.121 , 0.166	Depositor DCC
R_{free} test set	9831 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	14.4	Xtrriage
Anisotropy	0.393	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 58.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	14075	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: GOL, CSO, NA, SF4, MG, 8JU, NFV, F3S

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	L	0.32	0/4459	0.53	0/6076
1	M	0.30	0/4423	0.51	0/6025
2	S	0.30	0/2145	0.53	0/2921
2	T	0.30	0/2136	0.51	0/2909
All	All	0.31	0/13163	0.52	0/17931

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 [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	L	4361	0	4277	14	0
1	M	4325	0	4248	14	0
2	S	2086	0	1991	2	0
2	T	2077	0	1986	6	0
3	L	1	0	0	0	0
3	M	1	0	0	0	0
4	L	9	0	0	0	0
4	M	9	0	0	0	0
5	L	12	0	16	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	S	6	0	8	0	0
5	T	6	0	8	0	0
6	L	1	0	0	0	0
6	M	1	0	0	0	0
7	S	16	0	0	0	0
7	T	16	0	0	0	0
8	S	7	0	0	0	0
8	T	7	0	0	0	0
9	S	9	0	0	0	0
9	T	9	0	0	0	0
10	L	424	0	0	5	0
10	M	319	0	0	3	0
10	S	203	0	0	1	0
10	T	170	0	0	2	0
All	All	14075	0	12534	35	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 1.

The worst 5 of 35 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:M:4:ARG:HE	1:M:22:GLU:HG3	1.49	0.76
2:T:277:ASN:O	2:T:277:ASN:ND2	2.22	0.72
1:M:46:LYS:NZ	10:M:703:HOH:O	2.24	0.69
1:L:414:GLU:OE2	10:L:701:HOH:O	2.12	0.66
2:T:173:LYS:NZ	10:T:503:HOH:O	2.32	0.62

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	L	559/552 (101%)	537 (96%)	21 (4%)	1 (0%)	47	25
1	M	554/552 (100%)	535 (97%)	18 (3%)	1 (0%)	47	25
2	S	273/335 (82%)	262 (96%)	11 (4%)	0	100	100
2	T	272/335 (81%)	264 (97%)	8 (3%)	0	100	100
All	All	1658/1774 (94%)	1598 (96%)	58 (4%)	2 (0%)	51	28

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	211	LYS
1	M	211	LYS

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	L	470/460 (102%)	465 (99%)	5 (1%)	73	55
1	M	465/460 (101%)	462 (99%)	3 (1%)	86	76
2	S	217/261 (83%)	213 (98%)	4 (2%)	59	34
2	T	216/261 (83%)	215 (100%)	1 (0%)	88	80
All	All	1368/1442 (95%)	1355 (99%)	13 (1%)	78	59

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	S	22[A]	CYS
2	S	22[B]	CYS
1	M	382	ASN
1	L	463	ILE
1	M	242	TYR

Some 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](#)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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
1	CSO	L	546	1,4	3,6,7	1.10	0	0,6,8	0.00	-
1	CSO	M	546	1,4	3,6,7	0.95	0	0,6,8	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CSO	L	546	1,4	-	0/1/5/7	-
1	CSO	M	546	1,4	-	0/1/5/7	-

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	L	546	CSO	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 4 are monoatomic - leaving 14 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
7	SF4	T	401	2	0,12,12	0.00	-	-		
9	8JU	S	404[B]	2	0,11,11	0.00	-	-		
8	F3S	T	402	2	0,9,9	0.00	-	-		
5	GOL	L	604	6	5,5,5	0.41	0	5,5,5	0.58	0
5	GOL	L	603	-	5,5,5	0.42	0	5,5,5	0.22	0
7	SF4	S	401	2	0,12,12	0.00	-	-		
4	NFV	L	602	1	3,8,8	2.39	1 (33%)	-		
7	SF4	T	403[A]	2	0,12,12	0.00	-	-		
7	SF4	S	403[A]	2	0,12,12	0.00	-	-		
8	F3S	S	402	2	0,9,9	0.00	-	-		
5	GOL	S	405	-	5,5,5	0.42	0	5,5,5	0.27	0
5	GOL	T	405	-	5,5,5	0.35	0	5,5,5	0.37	0
9	8JU	T	404[B]	2	0,11,11	0.00	-	-		
4	NFV	M	602	1	3,8,8	2.72	1 (33%)	-		

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	SF4	T	401	2	-	-	0/6/5/5
9	8JU	S	404[B]	2	-	-	0/2/3/3
8	F3S	T	402	2	-	-	0/3/3/3
5	GOL	L	604	6	-	0/4/4/4	-
5	GOL	L	603	-	-	0/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	SF4	S	401	2	-	-	0/6/5/5
7	SF4	T	403[A]	2	-	-	0/6/5/5
7	SF4	S	403[A]	2	-	-	0/6/5/5
8	F3S	S	402	2	-	-	0/3/3/3
5	GOL	S	405	-	-	0/4/4/4	-
5	GOL	T	405	-	-	0/4/4/4	-
9	8JU	T	404[B]	2	-	-	0/2/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	M	602	NFV	O1-C1	-4.23	1.09	1.16
4	L	602	NFV	O1-C1	-3.32	1.11	1.16

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 [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	L	550/552 (99%)	-0.47	3 (0%) 91 91	9, 15, 25, 43	0
1	M	550/552 (99%)	-0.12	10 (1%) 68 70	12, 19, 33, 55	0
2	S	269/335 (80%)	-0.44	2 (0%) 87 88	11, 16, 28, 45	0
2	T	269/335 (80%)	-0.27	4 (1%) 73 75	11, 19, 37, 46	0
All	All	1638/1774 (92%)	-0.31	19 (1%) 79 80	9, 17, 31, 55	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	9	PRO	3.9
1	M	392	GLU	3.6
1	M	415	VAL	3.6
1	M	309	LEU	3.5
1	L	2	SER	3.5

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	CSO	M	546	7/8	0.97	0.12	13,15,20,20	1
1	CSO	L	546	7/8	0.98	0.09	10,11,16,16	1

6.3 Carbohydrates [i](#)

There are no carbohydrates 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	GOL	L	604	6/6	0.92	0.14	18,23,30,33	0
5	GOL	S	405	6/6	0.95	0.09	22,23,25,25	0
5	GOL	L	603	6/6	0.96	0.10	21,25,26,26	0
6	NA	M	603	1/1	0.97	0.06	29,29,29,29	0
5	GOL	T	405	6/6	0.98	0.07	24,27,29,32	0
7	SF4	T	401	8/8	0.99	0.05	11,12,12,13	0
7	SF4	T	403[A]	8/8	0.99	0.05	19,23,24,25	8
7	SF4	S	403[A]	8/8	0.99	0.04	14,18,19,19	8
7	SF4	S	401	8/8	0.99	0.05	11,12,12,12	0
8	F3S	T	402	7/7	0.99	0.05	12,12,13,13	0
3	MG	M	601	1/1	0.99	0.07	12,12,12,12	0
9	8JU	T	404[B]	9/9	1.00	0.03	13,16,22,25	9
8	F3S	S	402	7/7	1.00	0.04	11,11,12,12	0
6	NA	L	605	1/1	1.00	0.12	11,11,11,11	0
3	MG	L	601	1/1	1.00	0.03	10,10,10,10	0
4	NFV	L	602	9/9	1.00	0.07	8,9,12,13	0
9	8JU	S	404[B]	9/9	1.00	0.03	12,14,20,22	9
4	NFV	M	602	9/9	1.00	0.07	12,13,14,16	0

6.5 Other polymers [i](#)

There are no such residues in this entry.