



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 08:57 PM GMT

PDB ID : 4UE3
Title : The Mechanism of Hydrogen Activation by NiFe-hydrogenases and the Importance of the active site Arginine
Authors : Evans, R.M.; Wehlin, S.A.M.; Nomerotskaia, E.; Sargent, F.; Carr, S.B.; Phillips, S.E.V.; Armstrong, F.A.
Deposited on : 2014-12-15
Resolution : 1.40 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

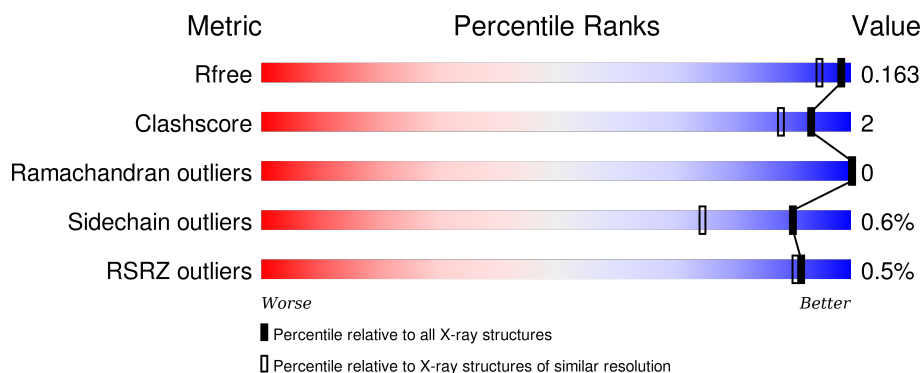
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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	1199 (1.40-1.40)
Clashscore	102246	1295 (1.40-1.40)
Ramachandran outliers	100387	1259 (1.40-1.40)
Sidechain outliers	100360	1258 (1.40-1.40)
RSRZ outliers	91569	1198 (1.40-1.40)

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 $\leq 5\%$. 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	582	<div> <div>91%</div> <div>8%</div> <div>.</div> </div>
1	M	582	<div> <div>92%</div> <div>7%</div> <div>.</div> </div>
2	S	335	<div> <div>%</div> <div>71%</div> <div>7%</div> <div>.</div> <div>21%</div> </div>
2	T	335	<div> <div>%</div> <div>71%</div> <div>8%</div> <div>.</div> <div>21%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	SO4	S	407	-	-	-	X
5	SO4	T	408	-	-	-	X

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 14776 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 HYDROGENASE-1 SMALL CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	581	Total	C	N	O	S	0	29	0
			4715	2998	824	864	29			
1	M	581	Total	C	N	O	S	0	19	0
			4648	2958	806	856	28			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	509	LYS	ARG	ENGINEERED MUTATION	UNP P0ACD8
M	509	LYS	ARG	ENGINEERED MUTATION	UNP P0ACD8

- Molecule 2 is a protein called HYDROGENASE-1 LARGE CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	S	265	Total	C	N	O	S	0	9	0
			2099	1333	360	385	21			
2	T	265	Total	C	N	O	S	0	10	0
			2107	1339	363	384	21			

There are 16 discrepancies between the modelled and reference sequences:

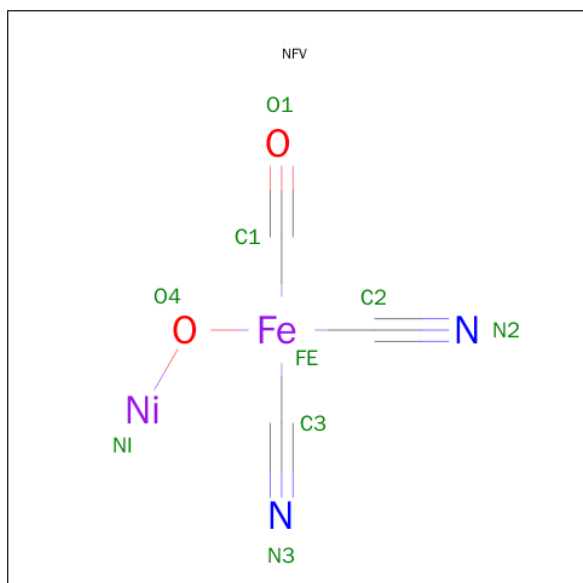
Chain	Residue	Modelled	Actual	Comment	Reference
S	328	ARG	-	EXPRESSION TAG	UNP P69739
S	329	SER	-	EXPRESSION TAG	UNP P69739
S	330	HIS	-	EXPRESSION TAG	UNP P69739
S	331	HIS	-	EXPRESSION TAG	UNP P69739
S	332	HIS	-	EXPRESSION TAG	UNP P69739
S	333	HIS	-	EXPRESSION TAG	UNP P69739
S	334	HIS	-	EXPRESSION TAG	UNP P69739
S	335	HIS	-	EXPRESSION TAG	UNP P69739
T	328	ARG	-	EXPRESSION TAG	UNP P69739
T	329	SER	-	EXPRESSION TAG	UNP P69739

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
T	330	HIS	-	EXPRESSION TAG	UNP P69739
T	331	HIS	-	EXPRESSION TAG	UNP P69739
T	332	HIS	-	EXPRESSION TAG	UNP P69739
T	333	HIS	-	EXPRESSION TAG	UNP P69739
T	334	HIS	-	EXPRESSION TAG	UNP P69739
T	335	HIS	-	EXPRESSION TAG	UNP P69739

- Molecule 3 is NI-FE OXIDIZED ACTIVE CENTER (three-letter code: NFV) (formula: $C_3FeN_2NiO_2$).



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

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

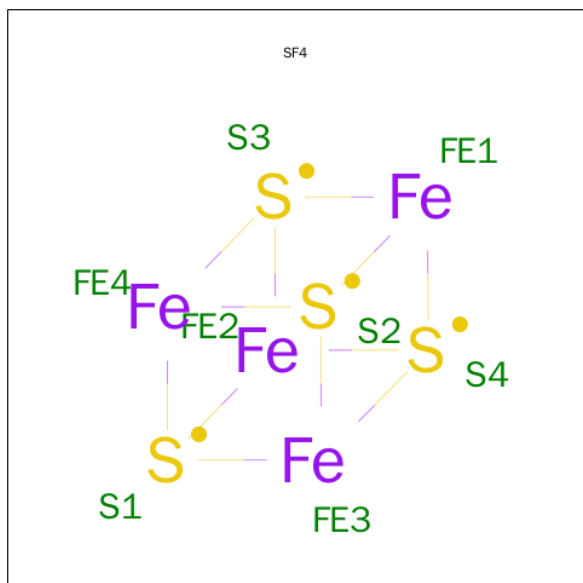
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	L	1	Total	Mg	0	0
			1	1		
4	M	1	Total	Mg	0	0
			1	1		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



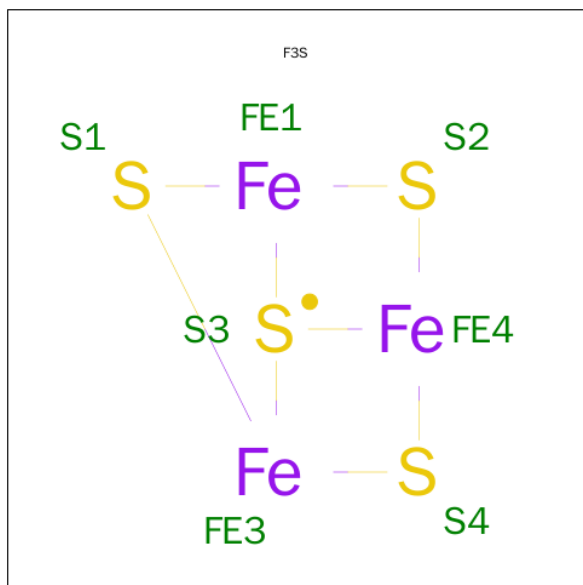
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	M	1	Total	O	S	0	0
			5	4	1		
5	S	1	Total	O	S	0	0
			5	4	1		
5	T	1	Total	O	S	0	0
			5	4	1		
5	T	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



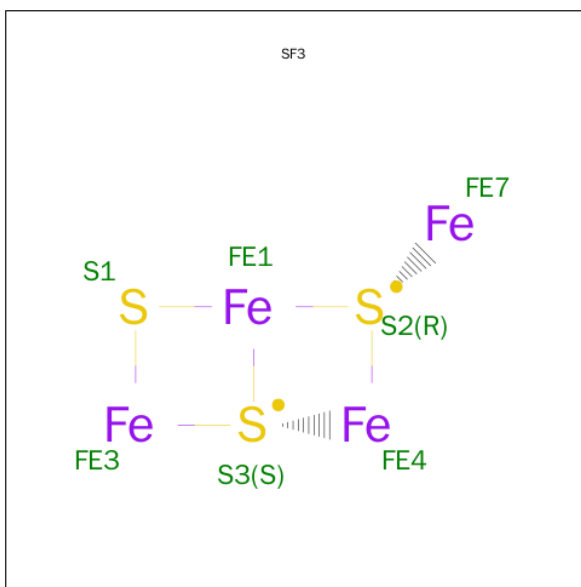
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	S	1	Total	Fe	S	0	0
			8	4	4		
6	T	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 7 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe_3S_4).



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

- Molecule 8 is FE4-S3 CLUSTER (three-letter code: SF3) (formula: Fe_4S_3).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	S	1	Total	Fe	S	0	1
			8	5	3		
8	T	1	Total	Fe	S	0	1
			8	5	3		

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	T	2	Total	Cl	0	0
			2	2		
9	S	2	Total	Cl	0	0
			2	2		

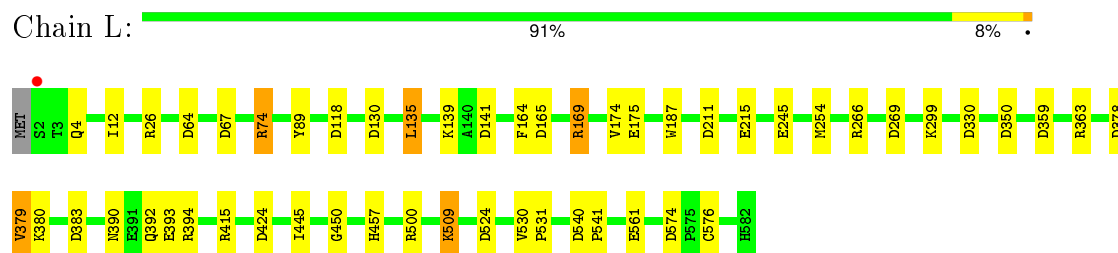
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	L	415	Total	O	0	0
			415	415		
10	M	446	Total	O	0	0
			446	446		
10	S	135	Total	O	0	0
			135	135		
10	T	121	Total	O	0	0
			121	121		

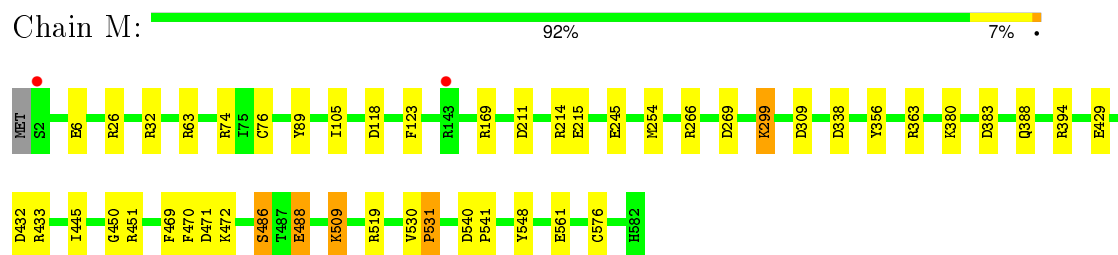
3 Residue-property plots

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.

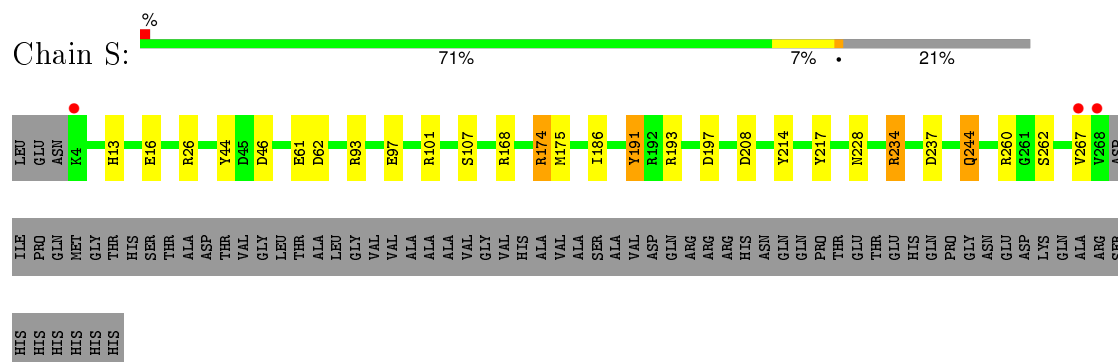
- Molecule 1: HYDROGENASE-1 SMALL CHAIN



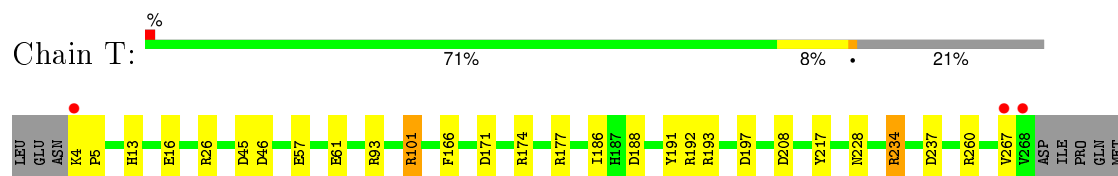
- Molecule 1: HYDROGENASE-1 SMALL CHAIN



- Molecule 2: HYDROGENASE-1 LARGE CHAIN



- Molecule 2: HYDROGENASE-1 LARGE CHAIN



GLY	THR	HIS	SER	THR	ALA	ASP	THR	VAL	GLY	LEU	THR	ALA	ALA	GLY	VAL	VAL	ALA	ALA	ALA	ALA	VAL	VAL	GLY	HIS	ALA	VAL	SER	ALA	ALA	ALA	VAL	ASP	GLN	ARG	ARG	ARG	HIS	ASN	GLN	GLN	PRO	THR	THR	GLU	THR	GLU	HIS	HIS	GLN	PRO	GLY	ASN	GLY	ASP	LYS	GLN	ALA	ALA	ARG	SER	HIS	HIS	HIS	HIS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

HIS
HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	93.52Å 97.41Å 183.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.61 – 1.40 48.70 – 1.40	Depositor EDS
% Data completeness (in resolution range)	96.8 (91.61-1.40) 96.8 (48.70-1.40)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.92 (at 1.40Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.136 , 0.155 0.146 , 0.163	Depositor DCC
R_{free} test set	16179 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	11.2	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 41.1	EDS
Estimated twinning fraction	0.014 for k,h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 316925 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	14776	wwPDB-VP
Average B, all atoms (Å ²)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.43% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, CL, SF4, SF3, NFV, F3S, SO4

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	1.04	6/4912 (0.1%)	1.14	44/6675 (0.7%)
1	M	1.04	10/4824 (0.2%)	1.14	36/6558 (0.5%)
2	S	1.16	6/2178 (0.3%)	1.37	26/2955 (0.9%)
2	T	1.16	6/2189 (0.3%)	1.34	22/2969 (0.7%)
All	All	1.08	28/14103 (0.2%)	1.21	128/19157 (0.7%)

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	L	0	1
1	M	0	1
2	S	0	1
2	T	0	1
All	All	0	4

All (28) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	260	ARG	CZ-NH1	10.55	1.46	1.33
2	S	260	ARG	CZ-NH1	10.40	1.46	1.33
1	M	245	GLU	CD-OE2	9.40	1.35	1.25
1	L	245	GLU	CD-OE2	8.94	1.35	1.25
1	M	488[A]	GLU	CD-OE2	-8.71	1.16	1.25
1	M	488[B]	GLU	CD-OE2	-8.71	1.16	1.25
2	S	262	SER	CB-OG	7.16	1.51	1.42
2	T	188	ASP	CG-OD2	-6.72	1.09	1.25
1	L	392	GLN	CD-OE1	6.71	1.38	1.24
1	L	561	GLU	CD-OE1	6.42	1.32	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	561	GLU	CG-CD	6.06	1.61	1.51
2	S	97	GLU	CD-OE1	-5.77	1.19	1.25
2	S	107	SER	CB-OG	-5.65	1.34	1.42
2	T	267	VAL	CA-CB	5.48	1.66	1.54
1	M	561	GLU	CD-OE1	5.46	1.31	1.25
2	T	193	ARG	CZ-NH1	-5.43	1.25	1.33
2	S	44	TYR	CE2-CZ	-5.33	1.31	1.38
1	L	379[A]	VAL	N-CA	-5.33	1.35	1.46
1	L	379[B]	VAL	N-CA	-5.33	1.35	1.46
1	M	486[A]	SER	CB-OG	5.32	1.49	1.42
1	M	486[B]	SER	CB-OG	5.32	1.49	1.42
1	M	488[A]	GLU	CG-CD	5.28	1.59	1.51
1	M	488[B]	GLU	CG-CD	5.28	1.59	1.51
1	M	76	CYS	CA-CB	-5.14	1.42	1.53
2	T	57	GLU	CD-OE2	-5.14	1.20	1.25
1	M	6	GLU	CG-CD	-5.13	1.44	1.51
2	S	244	GLN	CD-NE2	-5.04	1.20	1.32
2	T	217	TYR	CZ-OH	5.02	1.46	1.37

All (128) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	S	260	ARG	NE-CZ-NH2	-21.42	109.59	120.30
2	T	260	ARG	NE-CZ-NH2	-20.00	110.30	120.30
2	S	237	ASP	CB-CG-OD2	-17.80	102.28	118.30
2	S	197	ASP	CB-CG-OD1	16.29	132.97	118.30
1	M	169	ARG	NE-CZ-NH2	-15.41	112.59	120.30
2	T	260	ARG	NE-CZ-NH1	14.95	127.77	120.30
2	S	260	ARG	NE-CZ-NH1	14.87	127.74	120.30
2	T	237	ASP	CB-CG-OD2	-14.54	105.22	118.30
2	T	193	ARG	NE-CZ-NH2	-13.10	113.75	120.30
1	L	169	ARG	NE-CZ-NH2	-12.43	114.08	120.30
2	T	197	ASP	CB-CG-OD1	12.11	129.19	118.30
1	M	269[A]	ASP	CB-CG-OD1	11.71	128.84	118.30
1	M	269[B]	ASP	CB-CG-OD1	11.71	128.84	118.30
1	M	363	ARG	NE-CZ-NH1	11.16	125.88	120.30
1	L	26	ARG	NE-CZ-NH1	10.70	125.65	120.30
2	S	237	ASP	OD1-CG-OD2	10.57	143.38	123.30
1	L	269[A]	ASP	CB-CG-OD1	10.31	127.58	118.30
1	L	269[B]	ASP	CB-CG-OD1	10.31	127.58	118.30
2	S	168	ARG	NE-CZ-NH2	-9.71	115.45	120.30
2	T	46	ASP	CB-CG-OD1	9.62	126.96	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	M	215	GLU	OE1-CD-OE2	-9.40	112.02	123.30
1	L	266	ARG	NE-CZ-NH2	-9.36	115.62	120.30
2	S	168	ARG	NE-CZ-NH1	9.25	124.92	120.30
2	T	192	ARG	NE-CZ-NH2	-9.22	115.69	120.30
1	M	269[A]	ASP	CB-CG-OD2	-9.21	110.01	118.30
1	M	269[B]	ASP	CB-CG-OD2	-9.21	110.01	118.30
1	L	215	GLU	OE1-CD-OE2	-9.14	112.33	123.30
2	S	93	ARG	NE-CZ-NH1	9.05	124.82	120.30
2	S	193	ARG	NE-CZ-NH2	-8.78	115.91	120.30
2	T	237	ASP	OD1-CG-OD2	8.75	139.92	123.30
2	T	193	ARG	NE-CZ-NH1	8.33	124.47	120.30
1	M	266	ARG	NE-CZ-NH2	-8.27	116.17	120.30
1	L	379[A]	VAL	CB-CA-C	8.00	126.59	111.40
1	L	379[B]	VAL	CB-CA-C	8.00	126.59	111.40
1	M	169	ARG	NE-CZ-NH1	7.88	124.24	120.30
1	L	363[A]	ARG	NE-CZ-NH2	-7.84	116.38	120.30
1	L	363[B]	ARG	NE-CZ-NH2	-7.84	116.38	120.30
1	L	89	TYR	CB-CG-CD1	7.79	125.67	121.00
1	M	118	ASP	CB-CG-OD1	-7.73	111.34	118.30
1	M	254	MET	CG-SD-CE	7.62	112.40	100.20
1	L	266	ARG	NE-CZ-NH1	7.51	124.06	120.30
1	L	169	ARG	NE-CZ-NH1	7.42	124.01	120.30
1	M	214	ARG	NE-CZ-NH2	-7.36	116.62	120.30
1	M	214	ARG	NE-CZ-NH1	7.25	123.92	120.30
1	M	488[A]	GLU	OE1-CD-OE2	-7.23	114.63	123.30
1	M	488[B]	GLU	OE1-CD-OE2	-7.23	114.63	123.30
2	S	217	TYR	CB-CG-CD2	-7.18	116.69	121.00
1	M	245	GLU	CG-CD-OE1	-7.09	104.12	118.30
1	M	394	ARG	NE-CZ-NH2	-7.09	116.76	120.30
2	T	93	ARG	NE-CZ-NH1	7.05	123.82	120.30
1	L	89	TYR	CB-CG-CD2	-7.00	116.80	121.00
1	L	394[A]	ARG	NE-CZ-NH2	-6.96	116.82	120.30
1	L	394[B]	ARG	NE-CZ-NH2	-6.96	116.82	120.30
1	M	338	ASP	CB-CG-OD1	6.90	124.51	118.30
1	L	118	ASP	CB-CG-OD1	-6.74	112.23	118.30
1	L	330	ASP	CB-CG-OD2	-6.71	112.26	118.30
1	L	74	ARG	NE-CZ-NH2	-6.57	117.01	120.30
2	S	62	ASP	CB-CG-OD2	-6.57	112.39	118.30
1	L	175	GLU	OE1-CD-OE2	6.55	131.16	123.30
1	M	26	ARG	NE-CZ-NH1	6.51	123.56	120.30
1	L	380[A]	LYS	N-CA-C	-6.50	93.45	111.00
1	L	380[B]	LYS	N-CA-C	-6.50	93.45	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	245	GLU	CG-CD-OE1	-6.44	105.42	118.30
1	L	379[A]	VAL	N-CA-C	-6.40	93.73	111.00
1	L	379[B]	VAL	N-CA-C	-6.40	93.73	111.00
2	S	267	VAL	CA-CB-CG2	6.35	120.42	110.90
2	S	197	ASP	OD1-CG-OD2	-6.27	111.38	123.30
2	T	267	VAL	CG1-CB-CG2	-6.24	100.92	110.90
2	S	46	ASP	CB-CG-OD1	6.23	123.91	118.30
1	M	309	ASP	CB-CG-OD2	-6.22	112.70	118.30
2	T	46	ASP	OD1-CG-OD2	-6.19	111.55	123.30
1	M	245	GLU	OE1-CD-OE2	6.11	130.63	123.30
1	M	32	ARG	NE-CZ-NH1	6.08	123.34	120.30
1	L	350	ASP	CB-CG-OD2	-6.08	112.83	118.30
1	L	359	ASP	CB-CG-OD1	6.05	123.75	118.30
2	S	234[A]	ARG	NE-CZ-NH1	6.05	123.32	120.30
2	S	234[B]	ARG	NE-CZ-NH1	6.05	123.32	120.30
2	T	267	VAL	CA-CB-CG2	5.99	119.89	110.90
1	L	500	ARG	NE-CZ-NH2	-5.98	117.31	120.30
1	M	26	ARG	NE-CZ-NH2	-5.94	117.33	120.30
1	L	424	ASP	CB-CG-OD2	5.94	123.64	118.30
2	T	45	ASP	CB-CG-OD2	-5.86	113.03	118.30
1	M	63	ARG	NE-CZ-NH2	5.84	123.22	120.30
1	L	415	ARG	NE-CZ-NH1	5.84	123.22	120.30
1	L	141	ASP	CB-CG-OD1	5.83	123.55	118.30
1	L	574	ASP	CB-CG-OD2	-5.81	113.07	118.30
1	L	211	ASP	CB-CG-OD1	5.77	123.49	118.30
2	S	214	TYR	CB-CG-CD2	-5.74	117.55	121.00
1	M	470	PHE	CB-CG-CD2	5.69	124.78	120.80
1	L	299	LYS	CD-CE-NZ	5.68	124.77	111.70
2	S	191	TYR	CB-CG-CD1	5.63	124.38	121.00
2	S	62	ASP	CB-CG-OD1	5.62	123.36	118.30
2	T	174[A]	ARG	NE-CZ-NH1	5.62	123.11	120.30
2	T	174[B]	ARG	NE-CZ-NH1	5.62	123.11	120.30
2	S	174	ARG	NE-CZ-NH1	5.60	123.10	120.30
2	T	177	ARG	NE-CZ-NH1	5.58	123.09	120.30
1	M	299	LYS	CD-CE-NZ	5.57	124.51	111.70
1	M	356	TYR	CB-CG-CD1	5.53	124.32	121.00
1	L	383	ASP	CB-CG-OD1	5.52	123.27	118.30
1	L	26	ARG	NE-CZ-NH2	-5.50	117.55	120.30
2	S	193	ARG	NE-CZ-NH1	5.48	123.04	120.30
1	M	471	ASP	CB-CG-OD2	-5.47	113.38	118.30
1	L	211	ASP	CB-CG-OD2	-5.40	113.44	118.30
1	L	130	ASP	CB-CG-OD2	-5.33	113.50	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	208	ASP	CB-CG-OD2	-5.30	113.53	118.30
2	S	208	ASP	CB-CG-OD2	-5.27	113.55	118.30
1	M	211	ASP	CB-CG-OD1	5.26	123.04	118.30
1	L	67	ASP	CB-CG-OD2	5.25	123.03	118.30
1	L	64[A]	ASP	CB-CG-OD2	5.25	123.03	118.30
1	L	64[B]	ASP	CB-CG-OD2	5.25	123.03	118.30
2	S	267	VAL	CG1-CB-CG2	-5.23	102.53	110.90
1	M	451	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	M	519	ARG	CG-CD-NE	5.20	122.72	111.80
2	T	193	ARG	CG-CD-NE	-5.20	100.89	111.80
1	M	548	TYR	CB-CG-CD1	5.18	124.11	121.00
2	S	168	ARG	CD-NE-CZ	5.17	130.83	123.60
1	L	135	LEU	CB-CG-CD1	5.11	119.69	111.00
2	T	234[A]	ARG	NE-CZ-NH1	5.10	122.85	120.30
2	T	234[B]	ARG	NE-CZ-NH1	5.10	122.85	120.30
1	L	174	VAL	CA-CB-CG1	5.09	118.53	110.90
1	M	89	TYR	CA-CB-CG	5.07	123.03	113.40
1	M	356	TYR	CB-CG-CD2	-5.07	117.96	121.00
2	S	214	TYR	CD1-CE1-CZ	-5.05	115.25	119.80
2	T	171	ASP	CB-CG-OD1	5.05	122.85	118.30
2	S	217	TYR	CG-CD1-CE1	-5.05	117.26	121.30
1	M	432	ASP	CB-CG-OD1	5.04	122.84	118.30
1	L	165	ASP	CB-CG-OD2	-5.03	113.77	118.30
1	M	383	ASP	CB-CG-OD2	-5.02	113.78	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	L	74	ARG	Sidechain
1	M	74	ARG	Sidechain
2	S	26	ARG	Sidechain
2	T	26	ARG	Sidechain

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	L	4715	0	4663	16	0
1	M	4648	0	4578	16	0
2	S	2099	0	2049	9	0
2	T	2107	0	2067	12	0
3	L	9	0	0	0	0
3	M	9	0	0	0	0
4	L	1	0	0	0	0
4	M	1	0	0	0	0
5	M	5	0	0	0	0
5	S	5	0	0	0	0
5	T	10	0	0	0	1
6	S	8	0	0	0	0
6	T	8	0	0	0	0
7	S	7	0	0	0	0
7	T	7	0	0	0	0
8	S	8	0	0	0	0
8	T	8	0	0	0	0
9	S	2	0	0	0	0
9	T	2	0	0	0	0
10	L	415	0	0	7	1
10	M	446	0	0	9	0
10	S	135	0	0	1	0
10	T	121	0	0	2	0
All	All	14776	0	13357	50	1

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 2.

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)
2:S:61[B]:GLU:OE2	2:S:101[B]:ARG:NH1	1.84	1.10
1:M:486[A]:SER:OG	1:M:488[A]:GLU:OE2	1.77	1.01
2:S:234[A]:ARG:HD3	2:T:234[A]:ARG:NH1	1.82	0.94
1:L:139:LYS:HE3	10:L:2118:HOH:O	1.77	0.84
1:L:378:ASP:OD1	1:L:379[A]:VAL:O	1.95	0.84
2:T:61[B]:GLU:HG2	10:T:2022:HOH:O	1.79	0.83
2:T:101:ARG:HH11	2:T:101:ARG:HG3	1.47	0.79
1:M:123:PHE:CZ	10:M:2256:HOH:O	2.39	0.75
1:M:105:ILE:HG23	10:M:2116:HOH:O	1.86	0.74
1:M:472[B]:LYS:HD2	10:M:2396:HOH:O	1.89	0.72
1:M:488[A]:GLU:CD	1:M:488[A]:GLU:H	1.97	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:M:469:PHE:CD1	10:M:2116:HOH:O	2.47	0.67
1:M:429[B]:GLU:HG3	10:M:2375:HOH:O	1.94	0.67
1:L:164:PHE:CD1	10:L:2141:HOH:O	2.50	0.64
1:L:139:LYS:CE	10:L:2118:HOH:O	2.38	0.63
1:M:123:PHE:CE2	10:M:2256:HOH:O	2.52	0.63
2:S:61[B]:GLU:HG2	10:S:2028:HOH:O	2.00	0.62
1:L:254[B]:MET:HA	1:L:254[B]:MET:HE2	1.82	0.61
10:M:2035:HOH:O	2:T:16[B]:GLU:HG2	2.05	0.55
1:L:390:ASN:ND2	1:L:393[B]:GLU:OE2	2.40	0.54
2:T:61[B]:GLU:OE1	2:T:101:ARG:NH2	2.39	0.54
2:T:101:ARG:HG3	2:T:101:ARG:NH1	2.17	0.54
1:M:429[A]:GLU:OE2	1:M:433[A]:ARG:NH2	2.40	0.53
10:L:2027:HOH:O	2:S:16[B]:GLU:HG2	2.08	0.52
1:L:164:PHE:CE1	10:L:2141:HOH:O	2.63	0.52
2:S:234[A]:ARG:NH1	2:T:234[A]:ARG:HD3	2.25	0.52
1:M:123:PHE:CE1	10:M:2256:HOH:O	2.63	0.51
1:M:380[A]:LYS:HB2	1:M:388:GLN:HB2	1.93	0.51
1:L:135:LEU:HD22	1:L:187:TRP:CD1	2.47	0.50
1:L:530:VAL:CG1	1:L:531:PRO:HD2	2.43	0.49
2:T:4:LYS:HG3	2:T:5:PRO:HD2	1.95	0.48
1:L:457:HIS:HB2	10:L:2350:HOH:O	2.12	0.48
1:M:530:VAL:CG1	1:M:531:PRO:HD2	2.45	0.47
1:L:139:LYS:NZ	10:L:2118:HOH:O	2.48	0.47
1:M:509:LYS:HE2	1:M:576:CYS:HB2	1.97	0.46
2:S:186:ILE:HD11	2:S:228:ASN:HB3	1.98	0.46
2:T:186:ILE:HD11	2:T:228:ASN:HB3	1.99	0.45
1:M:540:ASP:HB2	1:M:541:PRO:CD	2.46	0.45
1:L:254[B]:MET:HE1	2:T:166:PHE:HE1	1.82	0.45
2:S:174:ARG:NH2	2:S:175[A]:MET:CE	2.80	0.45
1:L:445:ILE:O	1:L:450:GLY:HA3	2.17	0.44
1:L:509:LYS:HE2	1:L:576:CYS:HB2	1.99	0.44
1:L:4:GLN:HA	1:L:12:ILE:O	2.18	0.43
2:T:101:ARG:NH1	2:T:101:ARG:CG	2.81	0.43
1:L:540:ASP:HB2	1:L:541:PRO:CD	2.47	0.43
1:M:105:ILE:HA	10:M:2116:HOH:O	2.18	0.43
1:M:445:ILE:O	1:M:450:GLY:HA3	2.18	0.43
2:S:13:HIS:HD1	2:S:16[B]:GLU:CD	2.23	0.42
2:T:13:HIS:HD2	10:T:2003:HOH:O	2.02	0.42
2:S:234[A]:ARG:NH2	2:S:244:GLN:HE22	2.19	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:T:408:SO4:O4	10:L:2151:HOH:O[4_445]	1.74	0.46

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	608/582 (104%)	590 (97%)	18 (3%)	0	100	100
1	M	598/582 (103%)	581 (97%)	17 (3%)	0	100	100
2	S	272/335 (81%)	259 (95%)	13 (5%)	0	100	100
2	T	273/335 (82%)	259 (95%)	14 (5%)	0	100	100
All	All	1751/1834 (96%)	1689 (96%)	62 (4%)	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	L	508/481 (106%)	505 (99%)	3 (1%)	90	75
1	M	499/481 (104%)	496 (99%)	3 (1%)	90	75
2	S	227/274 (83%)	226 (100%)	1 (0%)	93	81
2	T	228/274 (83%)	226 (99%)	2 (1%)	84	63
All	All	1462/1510 (97%)	1453 (99%)	9 (1%)	90	75

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	169	ARG
1	L	509	LYS
1	L	524	ASP
1	M	299	LYS
1	M	509	LYS
1	M	531	PRO
2	S	191	TYR
2	T	101	ARG
2	T	191	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](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 6 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	$\# Z > 2$
3	NFV	L	601	1	3,8,8	0.99	0	0,10,10	0.00	-
3	NFV	M	601	1	3,8,8	1.13	0	0,10,10	0.00	-
5	SO4	M	604	-	4,4,4	1.04	0	6,6,6	0.58	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
6	SF4	S	401	2	0,12,12	0.00	-	0,24,24	0.00	-
7	F3S	S	402	2	0,9,9	0.00	-	0,15,15	0.00	-
8	SF3	S	403[B]	2	0,8,8	0.00	-	0,12,12	0.00	-
8	SF3	S	403[C]	2	0,8,8	0.00	-	0,12,12	0.00	-
5	SO4	S	407	-	4,4,4	1.73	1 (25%)	6,6,6	1.00	0
6	SF4	T	401	2	0,12,12	0.00	-	0,24,24	0.00	-
7	F3S	T	402	2	0,9,9	0.00	-	0,15,15	0.00	-
8	SF3	T	403[B]	2	0,8,8	0.00	-	0,12,12	0.00	-
8	SF3	T	403[C]	2	0,8,8	0.00	-	0,12,12	0.00	-
5	SO4	T	408	-	4,4,4	0.93	0	6,6,6	0.55	0
5	SO4	T	409	-	4,4,4	0.48	0	6,6,6	0.59	0

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
3	NFV	L	601	1	-	0/0/12/12	0/0/0/0
3	NFV	M	601	1	-	0/0/12/12	0/0/0/0
5	SO4	M	604	-	-	0/0/0/0	0/0/0/0
6	SF4	S	401	2	-	0/0/48/48	0/6/5/5
7	F3S	S	402	2	-	0/0/24/24	0/0/3/3
8	SF3	S	403[B]	2	-	0/0/17/17	0/2/2/2
8	SF3	S	403[C]	2	-	0/0/17/17	0/2/2/2
5	SO4	S	407	-	-	0/0/0/0	0/0/0/0
6	SF4	T	401	2	-	0/0/48/48	0/6/5/5
7	F3S	T	402	2	-	0/0/24/24	0/0/3/3
8	SF3	T	403[B]	2	-	0/0/17/17	0/2/2/2
8	SF3	T	403[C]	2	-	0/0/17/17	0/2/2/2
5	SO4	T	408	-	-	0/0/0/0	0/0/0/0
5	SO4	T	409	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	S	407	SO4	O2-S	2.47	1.55	1.47

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
5	T	408	SO4	0	1

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	581/582 (99%)	-0.40	1 (0%) 95 94	7, 12, 23, 51	0
1	M	581/582 (99%)	-0.47	2 (0%) 94 93	7, 12, 21, 35	0
2	S	265/335 (79%)	-0.27	3 (1%) 82 81	7, 11, 20, 56	0
2	T	265/335 (79%)	-0.26	3 (1%) 82 81	7, 11, 22, 54	0
All	All	1692/1834 (92%)	-0.38	9 (0%) 91 90	7, 11, 22, 56	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	S	268	VAL	10.0
2	T	268	VAL	9.7
2	T	267	VAL	6.0
2	S	267	VAL	4.5
2	T	4	LYS	3.3
2	S	4	LYS	3.1
1	L	2	SER	2.3
1	M	2	SER	2.3
1	M	143	ARG	2.2

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 carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
5	SO4	T	408	5/5	0.80	0.39	25.94	42,55,78,79	0
5	SO4	S	407	5/5	0.94	0.21	9.19	21,23,24,27	0
4	MG	M	603	1/1	1.00	0.07	0.09	7,7,7,7	0
3	NFV	M	601	9/9	0.98	0.07	-0.11	9,10,15,19	0
3	NFV	L	601	9/9	0.99	0.07	-0.17	9,10,17,19	0
4	MG	L	603	1/1	1.00	0.06	-0.82	7,7,7,7	0
9	CL	S	405	1/1	0.99	0.05	-1.08	13,13,13,13	0
8	SF3	T	403[B]	7/7	0.99	0.05	-1.12	9,11,13,13	1
8	SF3	T	403[C]	7/7	0.99	0.05	-1.12	10,11,13,17	1
7	F3S	S	402	7/7	1.00	0.05	-1.20	8,8,9,9	0
9	CL	S	406	1/1	0.99	0.05	-1.36	20,20,20,20	0
8	SF3	S	403[C]	7/7	0.99	0.05	-1.39	9,11,13,13	1
8	SF3	S	403[B]	7/7	0.99	0.05	-1.39	9,11,12,13	1
7	F3S	T	402	7/7	1.00	0.05	-1.43	8,9,9,10	0
6	SF4	S	401	8/8	0.99	0.04	-1.45	8,9,9,10	0
6	SF4	T	401	8/8	0.99	0.04	-1.60	8,9,9,10	0
9	CL	T	405	1/1	0.99	0.04	-1.71	14,14,14,14	0
9	CL	T	406	1/1	0.98	0.05	-2.43	21,21,21,21	0
5	SO4	T	409	5/5	0.86	0.29	-	67,73,82,90	0
5	SO4	M	604	5/5	0.89	0.22	-	48,49,58,63	0

6.5 Other polymers ⓘ

There are no such residues in this entry.