wwPDB X-ray Structure Validation Summary Report (i)

Jun 24, 2024 – 04:30 PM EDT

PDB ID	:	6TV2
Title	:	Heme d1 biosynthesis associated Protein NirF
Authors	:	Kluenemann, T.; Layer, G.; Blankenfeldt, W.
Deposited on	:	2020-01-08
Resolution	:	1.56 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 (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 (1)) were used in the production of this report:

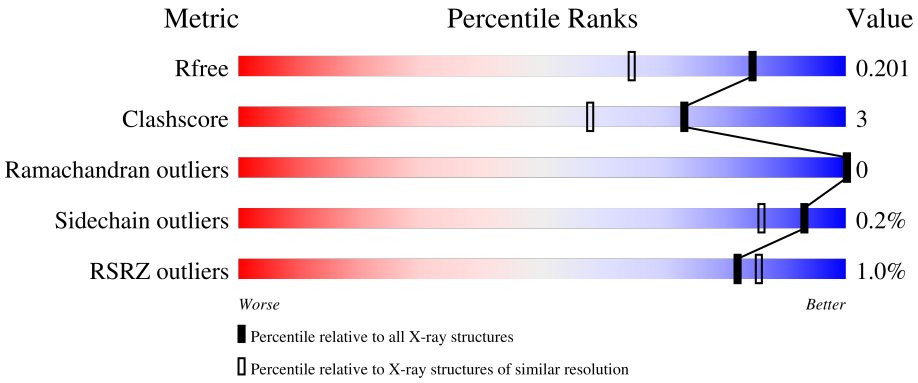
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 1.56 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.



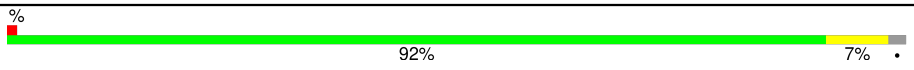
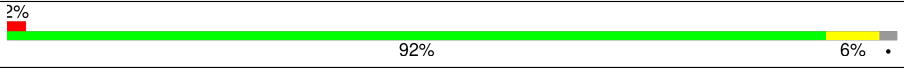
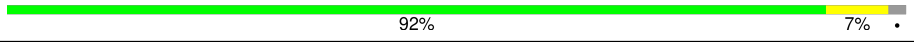
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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 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	A	378	93% 5%
1	B	378	93% 6%
1	C	378	92% 6%
1	D	378	2% 90% 8%
1	E	378	92% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	378	
1	G	378	
1	H	378	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 50955 atoms, of which 23539 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 Protein NirF.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	372	Total	C	H	N	O	S	0	10	0
			5908	1880	2933	544	545	6			
1	B	372	Total	C	H	N	O	S	0	5	0
			5837	1856	2897	536	542	6			
1	C	372	Total	C	H	N	O	S	0	8	0
			5880	1869	2920	540	544	7			
1	D	372	Total	C	H	N	O	S	0	18	0
			5992	1906	2978	548	554	6			
1	E	372	Total	C	H	N	O	S	0	12	0
			5933	1886	2951	543	547	6			
1	F	372	Total	C	H	N	O	S	0	10	0
			5900	1874	2932	540	547	7			
1	G	372	Total	C	H	N	O	S	0	9	0
			5889	1871	2925	541	546	6			
1	H	372	Total	C	H	N	O	S	0	12	0
			5943	1889	2957	545	545	7			

There are 24 discrepancies between the modelled and reference sequences:

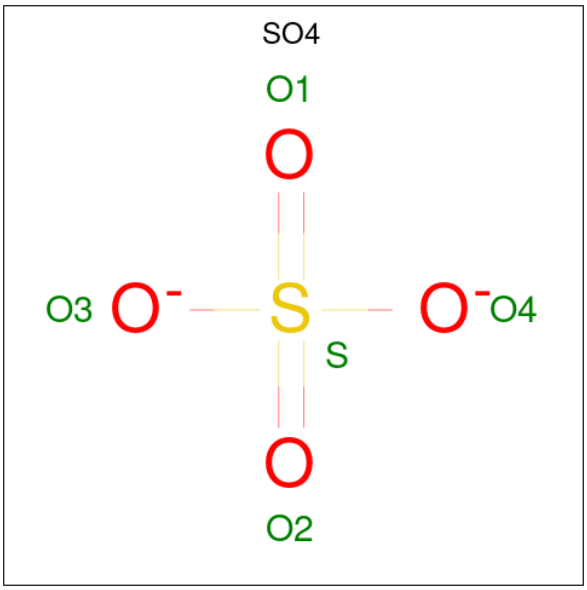
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	HIS	-	expression tag	UNP Q51480
A	2	MET	-	expression tag	UNP Q51480
A	3	MET	-	expression tag	UNP Q51480
B	1	HIS	-	expression tag	UNP Q51480
B	2	MET	-	expression tag	UNP Q51480
B	3	MET	-	expression tag	UNP Q51480
C	1	HIS	-	expression tag	UNP Q51480
C	2	MET	-	expression tag	UNP Q51480
C	3	MET	-	expression tag	UNP Q51480
D	1	HIS	-	expression tag	UNP Q51480
D	2	MET	-	expression tag	UNP Q51480
D	3	MET	-	expression tag	UNP Q51480
E	1	HIS	-	expression tag	UNP Q51480

Continued on next page...

Continued from previous page...

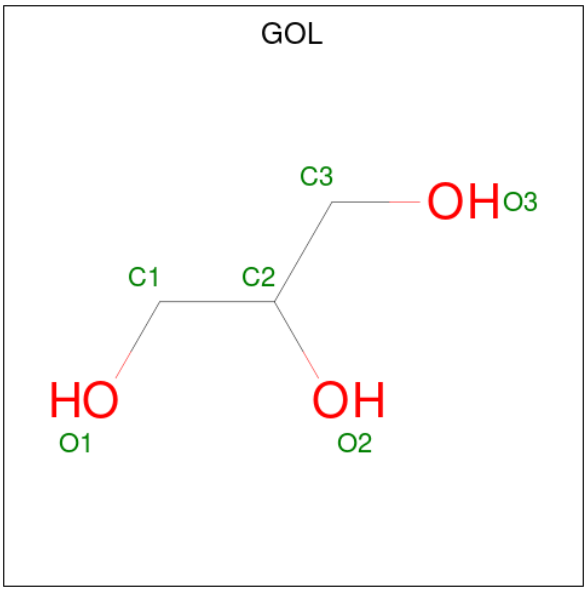
Chain	Residue	Modelled	Actual	Comment	Reference
E	2	MET	-	expression tag	UNP Q51480
E	3	MET	-	expression tag	UNP Q51480
F	1	HIS	-	expression tag	UNP Q51480
F	2	MET	-	expression tag	UNP Q51480
F	3	MET	-	expression tag	UNP Q51480
G	1	HIS	-	expression tag	UNP Q51480
G	2	MET	-	expression tag	UNP Q51480
G	3	MET	-	expression tag	UNP Q51480
H	1	HIS	-	expression tag	UNP Q51480
H	2	MET	-	expression tag	UNP Q51480
H	3	MET	-	expression tag	UNP Q51480

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



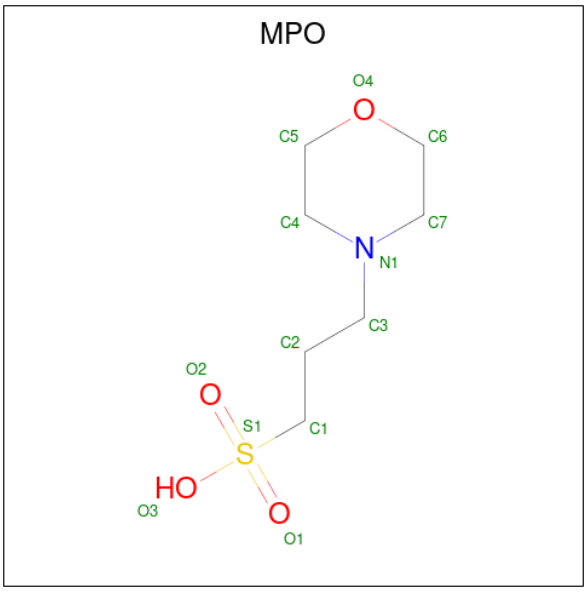
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	G	1	Total O S 5 4 1	0	0
2	H	1	Total O S 5 4 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	E	1	Total	C	H	O	0	0
			14	3	8	3		
3	H	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is 3[N-MORPHOLINO]PROPANE SULFONIC ACID (three-letter code: MPO) (formula: $C_7H_{15}NO_4S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	E	1	Total	C	H	N	O	S	0	0
			28	7	15	1	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	F	1	Total	C	H	N	O	S	0	0
			28	7	15	1	4	1		

- Molecule 5 is water.

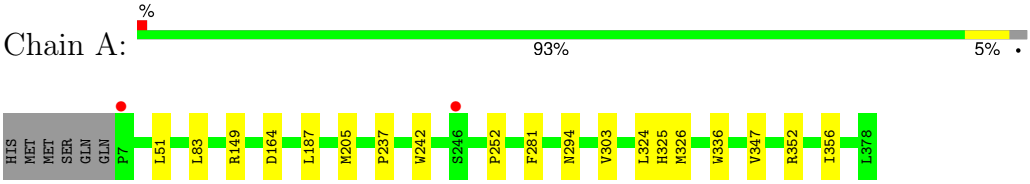
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	471	Total	O	0	0
			471	471		
5	B	463	Total	O	0	0
			463	463		
5	C	407	Total	O	0	0
			407	407		
5	D	453	Total	O	0	0
			453	453		
5	E	426	Total	O	0	0
			426	426		
5	F	456	Total	O	0	0
			456	456		
5	G	434	Total	O	0	0
			434	434		
5	H	459	Total	O	0	0
			459	459		

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.

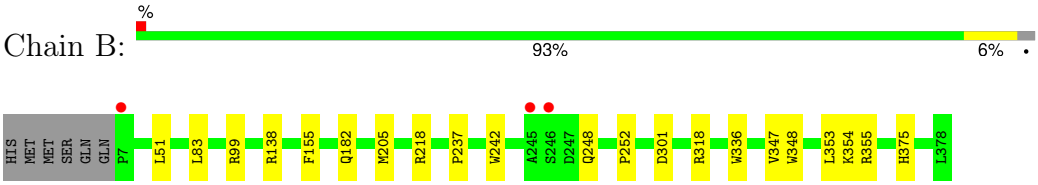
• Molecule 1: Protein NirF

Chain A:



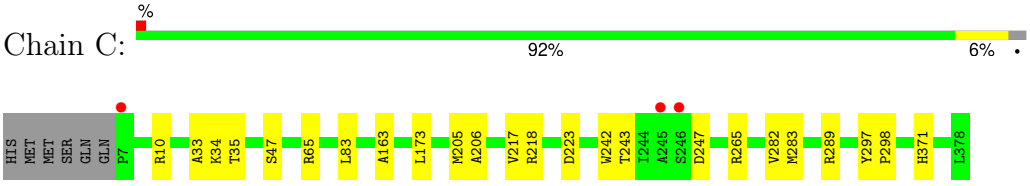
• Molecule 1: Protein NirF

Chain B:



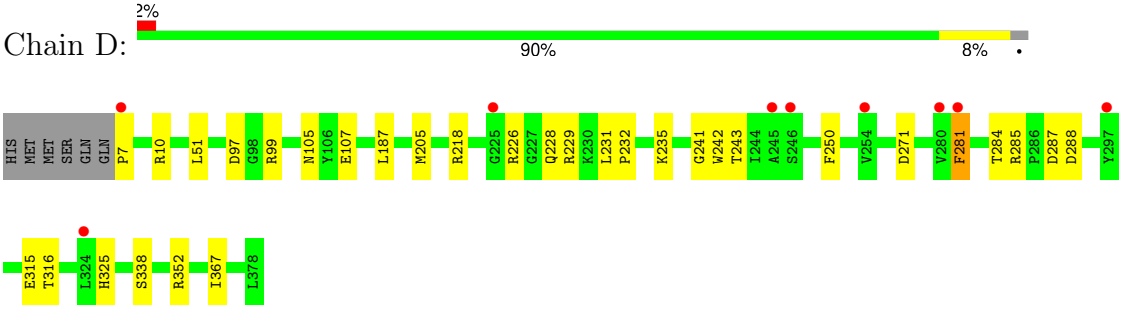
• Molecule 1: Protein NirF

Chain C:



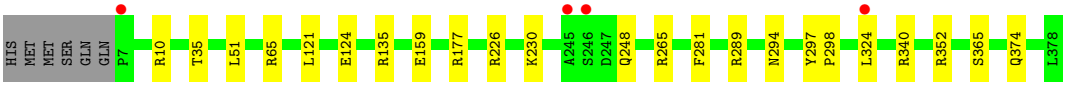
• Molecule 1: Protein NirF

Chain D:

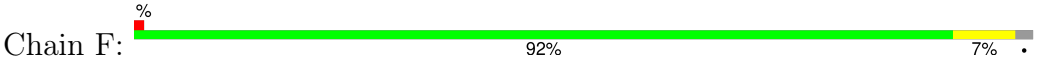


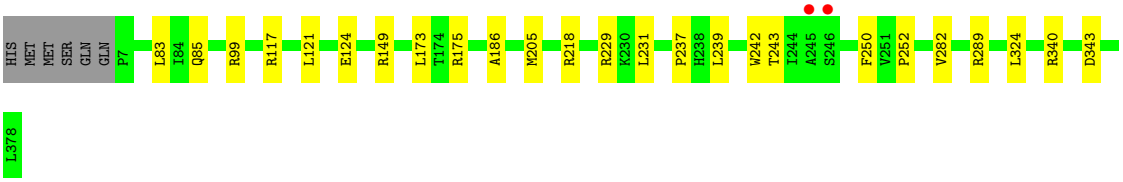
• Molecule 1: Protein NirF

Chain E:

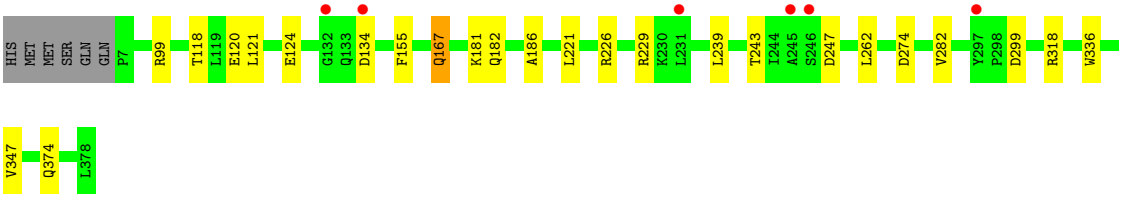


• Molecule 1: Protein NirF

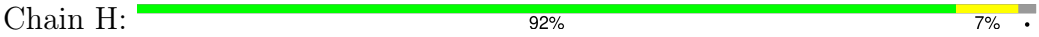


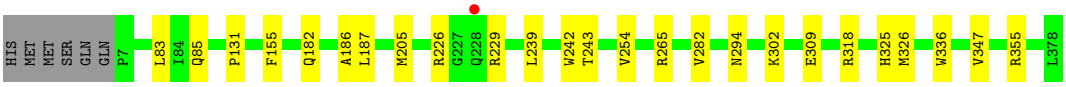


• Molecule 1: Protein NirF



• Molecule 1: Protein NirF





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	101.13Å 147.83Å 108.72Å 90.00° 98.21° 90.00°	Depositor
Resolution (Å)	59.46 – 1.56 107.60 – 1.56	Depositor EDS
% Data completeness (in resolution range)	64.9 (59.46-1.56) 64.9 (107.60-1.56)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 1.56Å)	Xtriage
Refinement program	PHENIX dev-3742	Depositor
R, R_{free}	0.163 , 0.202 0.163 , 0.201	Depositor DCC
R_{free} test set	14295 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	13.0	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 48.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	50955	wwPDB-VP
Average B, all atoms (Å^2)	22.0	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 38.53 % 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 3.6365e-04. 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, 0.333 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: MPO, SO4, GOL

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	0.37	0/3073	0.62	0/4162
1	B	0.36	0/3025	0.60	0/4099
1	C	0.34	0/3051	0.60	0/4133
1	D	0.37	0/3133	0.61	0/4244
1	E	0.35	0/3085	0.60	0/4179
1	F	0.37	0/3065	0.61	0/4152
1	G	0.34	0/3061	0.59	0/4148
1	H	0.37	0/3093	0.63	0/4188
All	All	0.36	0/24586	0.61	0/33305

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	A	2975	2933	2943	15	0
1	B	2940	2897	2898	16	0
1	C	2960	2920	2926	22	0
1	D	3014	2978	2994	23	1
1	E	2982	2951	2960	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	2968	2932	2935	20	0
1	G	2964	2925	2929	23	0
1	H	2986	2957	2964	17	1
2	A	5	0	0	0	0
2	C	5	0	0	0	0
2	G	5	0	0	0	0
2	H	5	0	0	1	0
3	E	6	8	7	1	0
3	H	6	8	7	1	0
4	E	13	15	15	0	0
4	F	13	15	14	1	0
5	A	471	0	0	1	4
5	B	463	0	0	6	2
5	C	407	0	0	7	1
5	D	453	0	0	5	2
5	E	426	0	0	11	1
5	F	456	0	0	7	2
5	G	434	0	0	11	1
5	H	459	0	0	6	5
All	All	27416	23539	23592	154	10

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 3.

The worst 5 of 154 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:F:243[A]:THR:HG23	1:F:282:VAL:O	1.76	0.84
1:D:231:LEU:CD2	1:D:235:LYS:HB2	2.11	0.81
1:D:231:LEU:HD21	1:D:235:LYS:HB2	1.66	0.76
1:F:117:ARG:NH1	5:F:501:HOH:O	2.20	0.74
3:E:401:GOL:O1	5:E:501:HOH:O	2.04	0.72

The worst 5 of 10 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:A:930:HOH:O	5:F:940:HOH:O[1_655]	2.00	0.20
5:B:757:HOH:O	5:H:639:HOH:O[1_554]	2.06	0.14
5:D:770:HOH:O	5:H:690:HOH:O[1_455]	2.06	0.14

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:C:878:HOH:O	5:G:584:HOH:O[2_554]	2.08	0.12
5:A:829:HOH:O	5:F:695:HOH:O[1_655]	2.09	0.11

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	380/378 (100%)	374 (98%)	6 (2%)	0	100	100
1	B	375/378 (99%)	369 (98%)	6 (2%)	0	100	100
1	C	378/378 (100%)	370 (98%)	8 (2%)	0	100	100
1	D	388/378 (103%)	384 (99%)	4 (1%)	0	100	100
1	E	382/378 (101%)	375 (98%)	7 (2%)	0	100	100
1	F	380/378 (100%)	376 (99%)	4 (1%)	0	100	100
1	G	379/378 (100%)	371 (98%)	8 (2%)	0	100	100
1	H	382/378 (101%)	377 (99%)	5 (1%)	0	100	100
All	All	3044/3024 (101%)	2996 (98%)	48 (2%)	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	316/312 (101%)	316 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	311/312 (100%)	311 (100%)	0	100 100
1	C	314/312 (101%)	313 (100%)	1 (0%)	92 85
1	D	324/312 (104%)	322 (99%)	2 (1%)	86 73
1	E	318/312 (102%)	317 (100%)	1 (0%)	92 85
1	F	316/312 (101%)	316 (100%)	0	100 100
1	G	315/312 (101%)	314 (100%)	1 (0%)	92 85
1	H	318/312 (102%)	318 (100%)	0	100 100
All	All	2532/2496 (101%)	2527 (100%)	5 (0%)	93 86

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	247	ASP
1	D	281[A]	PHE
1	D	281[B]	PHE
1	E	265	ARG
1	G	167	GLN

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)

8 ligands 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	$\#\|Z\| > 2$
2	SO4	G	401	-	4,4,4	0.21	0	6,6,6	0.15	0
4	MPO	E	402	-	13,13,13	0.27	0	17,17,17	0.49	0
3	GOL	H	401	-	5,5,5	1.36	0	5,5,5	0.67	0
2	SO4	C	401	-	4,4,4	0.24	0	6,6,6	0.09	0
2	SO4	A	401	-	4,4,4	0.32	0	6,6,6	0.21	0
2	SO4	H	402	-	4,4,4	0.23	0	6,6,6	0.12	0
4	MPO	F	401	-	13,13,13	0.21	0	17,17,17	0.52	0
3	GOL	E	401	-	5,5,5	0.90	0	5,5,5	1.10	1 (20%)

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
4	MPO	E	402	-	-	5/7/15/15	0/1/1/1
3	GOL	E	401	-	-	0/4/4/4	-
4	MPO	F	401	-	-	0/7/15/15	0/1/1/1
3	GOL	H	401	-	-	0/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	401	GOL	C3-C2-C1	-2.07	104.20	111.80

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	402	MPO	C2-C1-S1-O1
4	E	402	MPO	C2-C1-S1-O2
4	E	402	MPO	C2-C1-S1-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	E	402	MPO	C2-C3-N1-C7
4	E	402	MPO	C2-C3-N1-C4

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	H	401	GOL	1	0
2	H	402	SO4	1	0
4	F	401	MPO	1	0
3	E	401	GOL	1	0

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	A	372/378 (98%)	-0.52	2 (0%)	91	93	8, 15, 33, 65	0
1	B	372/378 (98%)	-0.53	3 (0%)	86	89	8, 15, 36, 64	0
1	C	372/378 (98%)	-0.46	3 (0%)	86	89	10, 19, 39, 70	0
1	D	372/378 (98%)	-0.40	9 (2%)	59	65	6, 15, 39, 73	0
1	E	372/378 (98%)	-0.50	4 (1%)	80	84	8, 16, 38, 61	0
1	F	372/378 (98%)	-0.56	2 (0%)	91	93	8, 15, 33, 52	0
1	G	372/378 (98%)	-0.45	6 (1%)	72	77	10, 19, 39, 64	0
1	H	372/378 (98%)	-0.56	1 (0%)	94	95	6, 15, 33, 55	0
All	All	2976/3024 (98%)	-0.50	30 (1%)	82	86	6, 16, 37, 73	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	245	ALA	5.2
1	B	246	SER	4.3
1	E	245	ALA	4.3
1	B	7	PRO	4.1
1	D	246	SER	3.7

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 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	GOL	E	401	6/6	0.87	0.14	18,31,39,45	0
4	MPO	E	402	13/13	0.91	0.17	41,50,59,60	0
4	MPO	F	401	13/13	0.92	0.14	35,48,57,60	0
3	GOL	H	401	6/6	0.93	0.13	14,29,37,44	0
2	SO4	C	401	5/5	0.95	0.11	45,48,52,55	0
2	SO4	G	401	5/5	0.96	0.15	47,47,51,60	0
2	SO4	A	401	5/5	0.97	0.10	30,30,38,39	0
2	SO4	H	402	5/5	0.98	0.12	35,37,51,53	0

6.5 Other polymers (i)

There are no such residues in this entry.