



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:26 AM GMT

PDB ID : 3LSM
Title : Pyranose 2-oxidase H167A mutant with flavin N(5) sulfite adduct
Authors : Tan, T.C.; Spadiut, O.; Divne, C.
Deposited on : 2010-02-12
Resolution : 1.70 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
<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

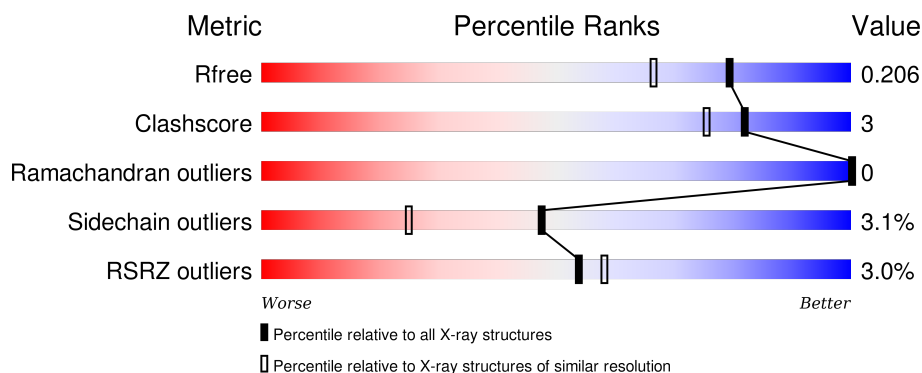
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	3190 (1.70-1.70)
Clashscore	102246	3585 (1.70-1.70)
Ramachandran outliers	100387	3527 (1.70-1.70)
Sidechain outliers	100360	3527 (1.70-1.70)
RSRZ outliers	91569	3200 (1.70-1.70)

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	A	623	<div> <div>3%</div> <div> <div></div> <div>82%</div> <div>9%</div> <div>8%</div> </div> </div>
1	B	623	<div> <div>3%</div> <div> <div></div> <div>81%</div> <div>10%</div> <div>8%</div> </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
2	SFD	A	801	X	-	-	-
2	SFD	B	801	X	-	-	-
3	SO3	A	911	-	-	-	X
4	GLC	A	912	-	-	-	X
5	12P	A	901	-	-	-	X
5	12P	B	902	-	-	-	X

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 10395 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.

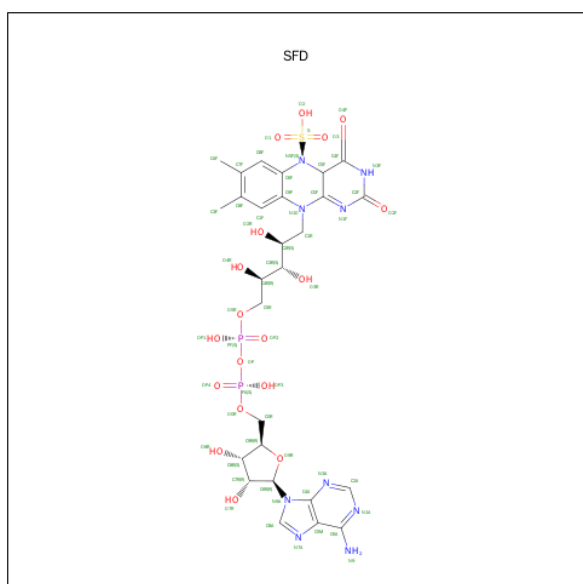
- Molecule 1 is a protein called Pyranose 2-oxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	574	Total	C	N	O	S	0	0	0
			4521	2856	773	868	24			
1	B	576	Total	C	N	O	S	0	0	0
			4537	2865	775	872	25			

There are 2 discrepancies between the modelled and reference sequences:

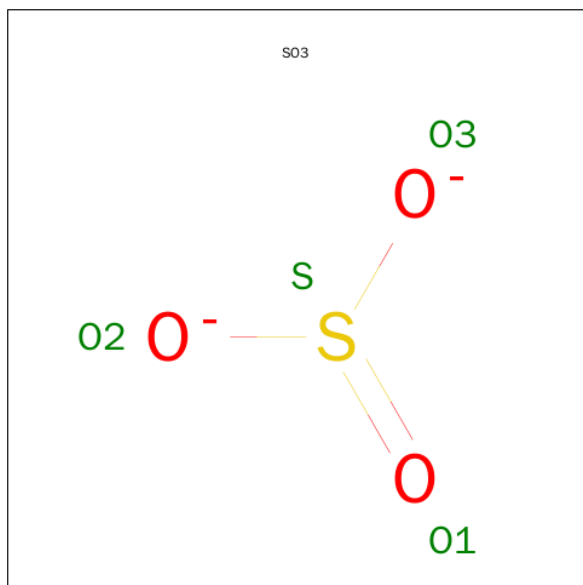
Chain	Residue	Modelled	Actual	Comment	Reference
A	167	ALA	HIS	ENGINEERED	UNP Q7ZA32
B	167	ALA	HIS	ENGINEERED	UNP Q7ZA32

- Molecule 2 is (S)-10-((2S,3S,4R)-5-((S)-((S)-(((2R,3S,4R,5R)-5-(6-AMINO-9H-PURIN-9-YL)-3,4-DIHYDROXY-TETRAHYDROFURAN-2-YL)METHOXY)(HYDROXY)PHOSPHORYLOXY)(HYDROXY)PHOSPHORYLOXY)-2,3,4-TRIHYDROXYPENTYL)-7,8-DIMETHYL-2,4-DIOXO-2,3,4,4A-TETRAHYDROBENZO[G]PTERIDINE-5(10H)-SULFONIC ACID (three-letter code: SFD) (formula: C₂₇H₃₅N₉O₁₈P₂S).



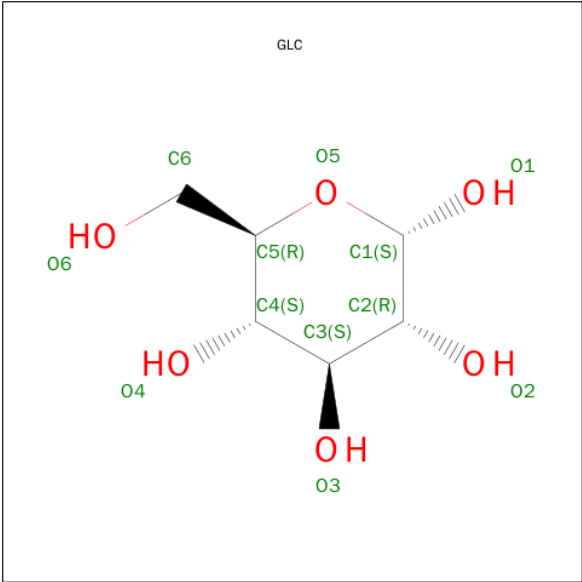
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	S	0	0
			57	27	9	18	2	1		
2	B	1	Total	C	N	O	P	S	0	0
			57	27	9	18	2	1		

- Molecule 3 is SULFITE ION (three-letter code: SO3) (formula: O₃S).



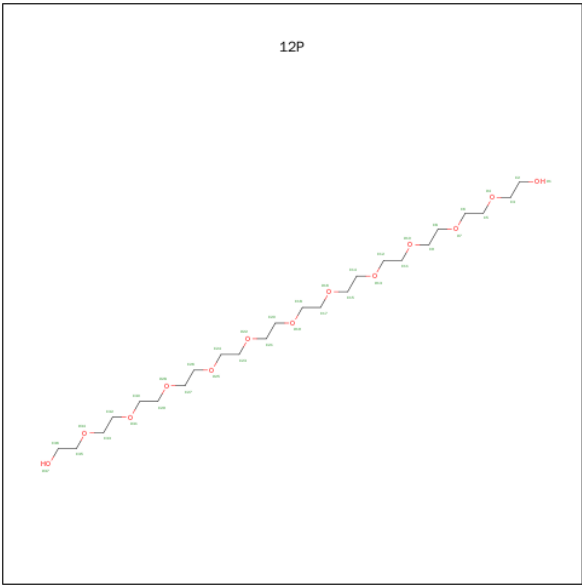
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			4	3	1		
3	A	1	Total	O	S	0	0
			4	3	1		
3	B	1	Total	O	S	0	0
			4	3	1		
3	B	1	Total	O	S	0	0
			4	3	1		

- Molecule 4 is SUGAR (ALPHA-D-GLUCOSE) (three-letter code: GLC) (formula: C₆H₁₂O₆).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			12	6	6		

- Molecule 5 is DODECAETHYLENE GLYCOL (three-letter code: 12P) (formula: C₂₄H₅₀O₁₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			12	8	4		
5	B	1	Total	C	O	0	0
			14	9	5		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	607	Total 607	O 607	0	0
6	B	562	Total 562	O 562	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	101.57Å 101.57Å 249.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.20 – 1.70 29.18 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.5 (29.20-1.70) 99.5 (29.18-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.57 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.160 , 0.195 0.174 , 0.206	Depositor DCC
R_{free} test set	2871 reflections (2.05%)	DCC
Wilson B-factor (Å ²)	17.0	Xtriage
Anisotropy	0.255	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 47.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 143007 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	10395	wwPDB-VP
Average B, all atoms (Å ²)	19.0	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 25.08 % 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 3.3927e-03.*

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: 12P, SO3, GLC, SFD

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	1.19	15/4636 (0.3%)	1.12	26/6304 (0.4%)
1	B	1.17	22/4652 (0.5%)	1.10	17/6325 (0.3%)
All	All	1.18	37/9288 (0.4%)	1.11	43/12629 (0.3%)

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	478	GLU	CD-OE1	8.10	1.34	1.25
1	A	310	GLU	CB-CG	-7.56	1.37	1.52
1	B	542	GLU	CD-OE1	7.02	1.33	1.25
1	A	482	GLU	CG-CD	6.98	1.62	1.51
1	B	432	GLU	CD-OE2	6.73	1.33	1.25

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	139	ARG	NE-CZ-NH1	24.26	132.43	120.30
1	A	139	ARG	NE-CZ-NH1	23.88	132.24	120.30
1	B	139	ARG	NE-CZ-NH2	-20.09	110.25	120.30
1	A	139	ARG	NE-CZ-NH2	-13.71	113.45	120.30
1	B	383	ARG	NE-CZ-NH1	-12.37	114.11	120.30

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	4521	0	4373	27	0
1	B	4537	0	4386	30	0
2	A	57	0	32	0	0
2	B	57	0	32	0	0
3	A	8	0	0	0	0
3	B	8	0	0	0	0
4	A	12	0	12	0	0
5	A	12	0	14	0	0
5	B	14	0	16	0	0
6	A	607	0	0	8	0
6	B	562	0	0	10	0
All	All	10395	0	8865	57	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 3.

The worst 5 of 57 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:B:231:LYS:HG2	1:B:231:LYS:O	1.64	0.95
1:A:618:PHE:HD1	1:A:618:PHE:C	1.90	0.76
1:A:618:PHE:C	1:A:618:PHE:CD1	2.63	0.72
1:B:404:HIS:HE1	6:B:1489:HOH:O	1.79	0.66
1:A:218:ARG:HD2	6:A:1144:HOH:O	1.96	0.66

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	A	572/623 (92%)	558 (98%)	14 (2%)	0	100	100
1	B	574/623 (92%)	561 (98%)	13 (2%)	0	100	100
All	All	1146/1246 (92%)	1119 (98%)	27 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	501/541 (93%)	489 (98%)	12 (2%)	57	36
1	B	503/541 (93%)	484 (96%)	19 (4%)	40	17
All	All	1004/1082 (93%)	973 (97%)	31 (3%)	47	25

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	139	ARG
1	B	185	LYS
1	B	496	ASN
1	B	168	TRP
1	B	206	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	461	GLN
1	A	563	ASN
1	B	324	HIS
1	A	404	HIS
1	B	263	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

9 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	SFD	A	801	-	50,62,62	2.92	9 (18%)	58,97,97	2.24	14 (24%)
5	12P	A	901	-	11,11,36	0.58	0	10,10,35	0.37	0
3	SO3	A	910	-	1,3,3	0.85	0	0,3,3	0.00	-
3	SO3	A	911	-	1,3,3	1.70	0	0,3,3	0.00	-
4	GLC	A	912	-	12,12,12	0.74	0	17,17,17	2.57	5 (29%)
2	SFD	B	801	-	50,62,62	3.06	13 (26%)	58,97,97	1.76	8 (13%)
5	12P	B	902	-	13,13,36	0.55	0	12,12,35	0.41	0
3	SO3	B	910	-	1,3,3	0.27	0	0,3,3	0.00	-
3	SO3	B	911	-	1,3,3	1.78	0	0,3,3	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
2	SFD	A	801	-	1/1/16/17	0/34/88/88	0/5/6/6

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	12P	A	901	-	-	0/9/9/34	0/0/0/0
3	SO3	A	910	-	-	0/0/0/0	0/0/0/0
3	SO3	A	911	-	-	0/0/0/0	0/0/0/0
4	GLC	A	912	-	-	0/2/22/22	0/1/1/1
2	SFD	B	801	-	1/1/16/17	0/34/88/88	0/5/6/6
5	12P	B	902	-	-	0/11/11/34	0/0/0/0
3	SO3	B	910	-	-	0/0/0/0	0/0/0/0
3	SO3	B	911	-	-	0/0/0/0	0/0/0/0

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	801	SFD	C6F-N5F	-11.55	1.32	1.44
2	A	801	SFD	C6F-N5F	-11.06	1.33	1.44
2	B	801	SFD	O9R-C9R	-3.18	1.37	1.45
2	B	801	SFD	C5M-C4A	-2.20	1.35	1.40
2	A	801	SFD	C4F-N3F	2.00	1.40	1.37

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	801	SFD	N3A-C2A-N1A	-7.80	122.92	128.89
2	A	801	SFD	O3-S-O1	-7.66	106.41	119.55
2	A	801	SFD	N3A-C2A-N1A	-6.87	123.63	128.89
4	A	912	GLC	C1-C2-C3	-5.87	101.69	110.43
2	B	801	SFD	O3-S-O1	-5.48	110.15	119.55

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	801	SFD	C5F
2	B	801	SFD	C5F

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	A	574/623 (92%)	-0.10	17 (2%) 54 58	11, 16, 32, 51	0
1	B	576/623 (92%)	-0.04	17 (2%) 54 58	11, 18, 34, 58	0
All	All	1150/1246 (92%)	-0.07	34 (2%) 54 58	11, 17, 34, 58	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	343	ALA	7.6
1	B	618	PHE	6.0
1	A	618	PHE	5.8
1	B	459	VAL	5.7
1	B	344	ASN	5.5

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 [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. 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(\AA^2)	Q<0.9
5	12P	A	901	12/37	0.90	0.13	12.21	33,36,40,45	0
5	12P	B	902	14/37	0.84	0.18	6.17	36,41,49,49	0
3	SO3	A	911	4/4	0.98	0.14	2.88	35,35,39,39	0
4	GLC	A	912	12/12	0.82	0.14	2.05	28,35,37,38	0
2	SFD	B	801	57/57	0.97	0.08	-0.59	10,14,30,32	0
2	SFD	A	801	57/57	0.97	0.09	-0.60	9,12,28,30	0
3	SO3	B	910	4/4	0.96	0.07	-0.61	28,38,38,40	0
3	SO3	B	911	4/4	0.97	0.07	-0.85	23,32,35,37	0
3	SO3	A	910	4/4	0.94	0.09	-	29,34,37,39	0

6.5 Other polymers [i](#)

There are no such residues in this entry.