



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:46 PM GMT

PDB ID : 4MOS
Title : Pyranose 2-oxidase H450G/V546C double mutant with 2-fluorinated galactose
Authors : Tan, T.C.; Spadiut, O.; Gandini, R.; Haltrich, D.; Divne, C.
Deposited on : 2013-09-12
Resolution : 1.80 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

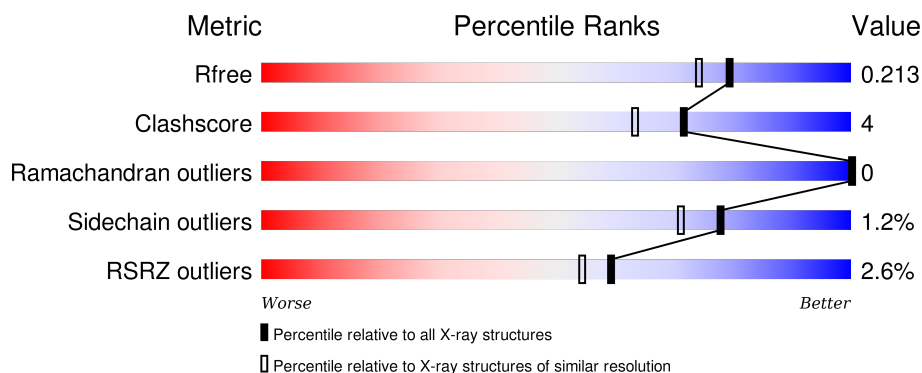
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

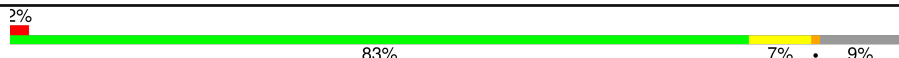
The reported resolution of this entry is 1.80 Å.

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	4533 (1.80-1.80)
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)
RSRZ outliers	91569	4547 (1.80-1.80)

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	633	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 5035 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.

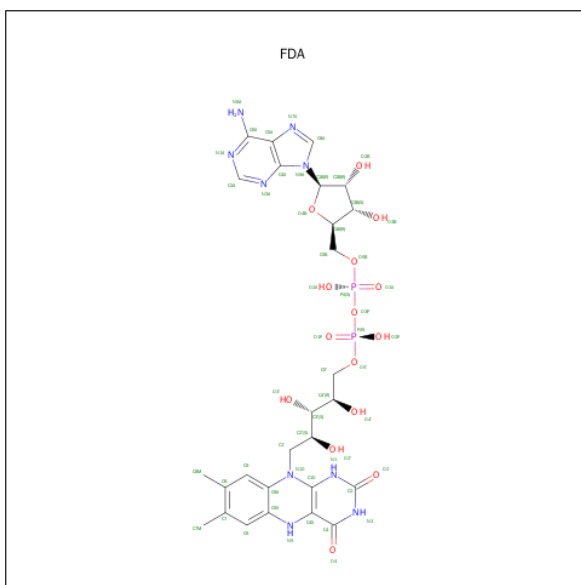
- Molecule 1 is a protein called Pyranose 2-oxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	576	Total	C	N	O	S	0	1	0
			4541	2865	776	874	26			

There are 14 discrepancies between the modelled and reference sequences:

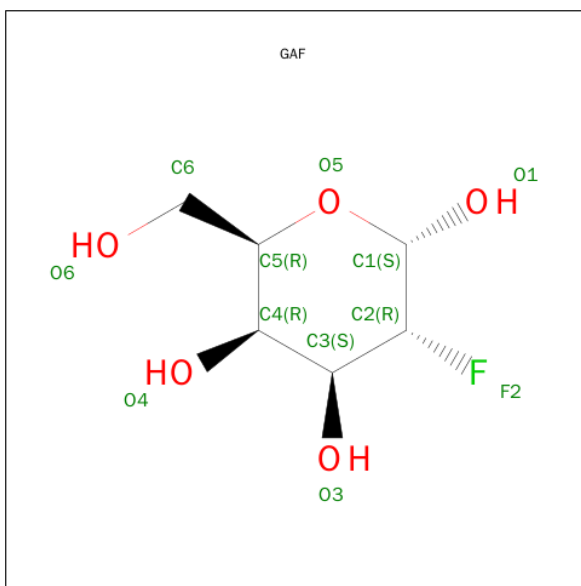
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	SER	SEE REMARK 999	UNP Q7ZA32
A	450	GLY	HIS	ENGINEERED MUTATION	UNP Q7ZA32
A	546	CYS	VAL	ENGINEERED MUTATION	UNP Q7ZA32
A	623	ALA	-	EXPRESSION TAG	UNP Q7ZA32
A	624	ALA	-	EXPRESSION TAG	UNP Q7ZA32
A	625	ALA	-	EXPRESSION TAG	UNP Q7ZA32
A	626	LEU	-	EXPRESSION TAG	UNP Q7ZA32
A	627	GLU	-	EXPRESSION TAG	UNP Q7ZA32
A	628	HIS	-	EXPRESSION TAG	UNP Q7ZA32
A	629	HIS	-	EXPRESSION TAG	UNP Q7ZA32
A	630	HIS	-	EXPRESSION TAG	UNP Q7ZA32
A	631	HIS	-	EXPRESSION TAG	UNP Q7ZA32
A	632	HIS	-	EXPRESSION TAG	UNP Q7ZA32
A	633	HIS	-	EXPRESSION TAG	UNP Q7ZA32

- Molecule 2 is DIHYDROFLAVINE-ADENINE DINUCLEOTIDE (three-letter code: FDA) (formula: C₂₇H₃₅N₉O₁₅P₂).



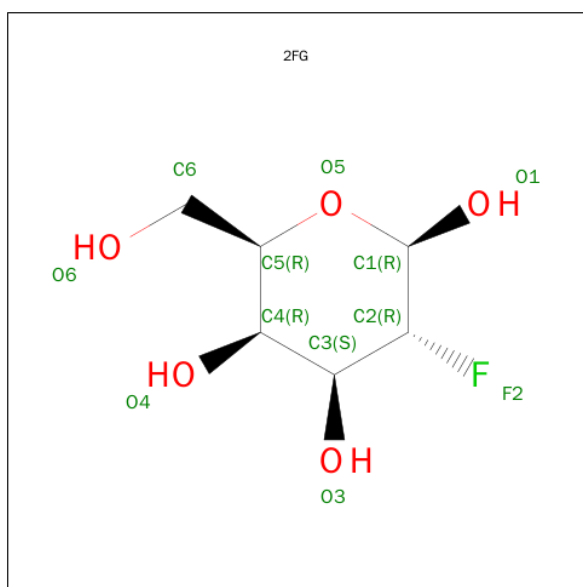
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 3 is 2-DEOXY-2-FLUORO-ALPHA-D-GALACTOPYRANOSE (three-letter code: GAF) (formula: $C_6H_{11}FO_5$).



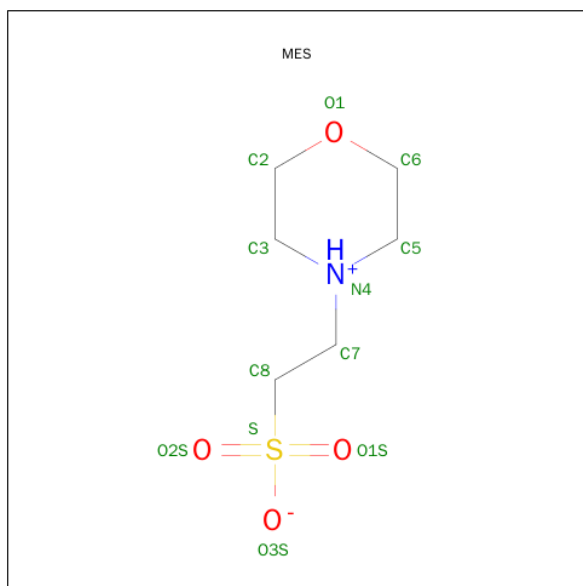
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	F	O	0	0
			12	6	1	5		

- Molecule 4 is SUGAR (2-FLUORO-2-DEOXY-BETA-D-GALACTOPYRANOSE) (three-letter code: 2FG) (formula: $C_6H_{11}FO_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	F	O	0	0
			12	6	1	5		

- Molecule 5 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).

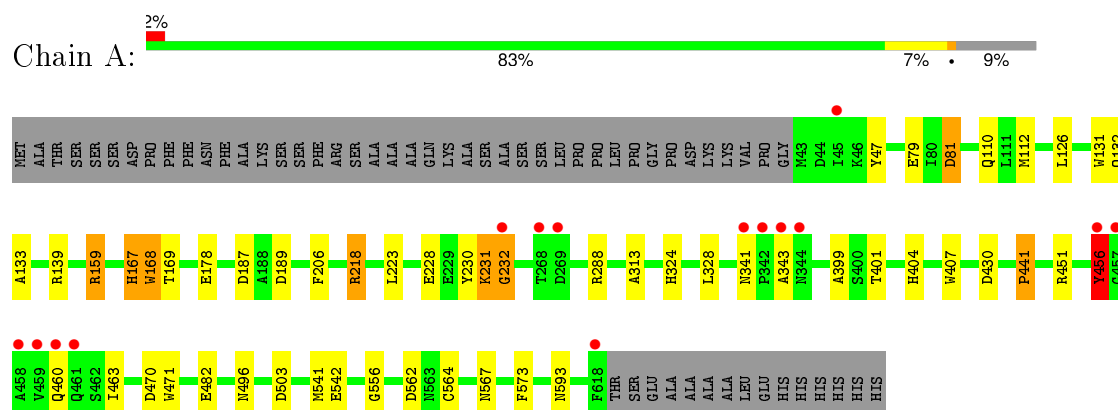


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	405	Total 405	O 405	0	0

- Molecule 1: Pyranose 2-oxidase



4 Data and refinement statistics

Property	Value	Source
Space group	P 4 ₂ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	102.42Å 102.42Å 119.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.04 – 1.80 47.04 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.04-1.80) 100.0 (47.04-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.01 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.162 , 0.199 0.173 , 0.213	Depositor DCC
R_{free} test set	1022 reflections (1.76%)	DCC
Wilson B-factor (Å ²)	18.9	Xtriage
Anisotropy	0.180	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 45.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	0 of 59195 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5035	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: FDA, GAF, 2FG, MES

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	1.05	9/4656 (0.2%)	1.08	19/6329 (0.3%)

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	A	0	3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	168	TRP	CD2-CE2	6.32	1.49	1.41
1	A	167	HIS	C-N	5.97	1.47	1.34
1	A	232	GLY	N-CA	-5.96	1.37	1.46
1	A	471	TRP	CD2-CE2	5.82	1.48	1.41
1	A	482	GLU	CD-OE1	5.60	1.31	1.25

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	139	ARG	NE-CZ-NH2	-23.37	108.62	120.30
1	A	139	ARG	NE-CZ-NH1	16.87	128.74	120.30
1	A	81	ASP	CB-CG-OD1	-9.00	110.20	118.30
1	A	139	ARG	CD-NE-CZ	8.42	135.38	123.60
1	A	232	GLY	N-CA-C	-8.21	92.58	113.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	230	TYR	Peptide
1	A	231	LYS	Peptide
1	A	456	TYR	Mainchain

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	4541	0	4384	36	0
2	A	53	0	30	5	0
3	A	12	0	11	3	0
4	A	12	0	11	0	0
5	A	12	0	13	0	0
6	A	405	0	0	8	0
All	All	5035	0	4449	39	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.

The worst 5 of 39 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:A:399:ALA:HB1	1:A:401:THR:HG22	1.48	0.95
1:A:231:LYS:HA	6:A:1277:HOH:O	1.73	0.88
1:A:110:GLN:HE21	1:A:167:HIS:HD1	1.25	0.85
1:A:404:HIS:HE1	6:A:1018:HOH:O	1.74	0.70
1:A:228:GLU:O	1:A:231:LYS:HD3	1.94	0.68

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	575/633 (91%)	559 (97%)	16 (3%)	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	504/547 (92%)	498 (99%)	6 (1%)	78	71

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	206	PHE
1	A	593	ASN
1	A	441	PRO
1	A	168	TRP
1	A	496	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	324	HIS
1	A	341	ASN
1	A	404	HIS
1	A	460	GLN

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](#)

4 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	FDA	A	801	1	48,58,58	1.77	11 (22%)	54,89,89	3.71	13 (24%)
3	GAF	A	802	-	11,12,12	1.34	0	15,17,17	1.31	2 (13%)
4	2FG	A	803	-	11,12,12	0.98	0	15,17,17	2.37	3 (20%)
5	MES	A	804	-	11,12,12	1.15	1 (9%)	14,16,16	2.57	4 (28%)

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	FDA	A	801	1	-	0/30/50/50	0/6/6/6
3	GAF	A	802	-	-	0/2/22/22	0/1/1/1
4	2FG	A	803	-	-	0/2/22/22	0/1/1/1
5	MES	A	804	-	-	0/6/14/14	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	FDA	C10-N10	-4.27	1.34	1.39
2	A	801	FDA	C2B-C3B	-2.50	1.46	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	FDA	C2A-N1A	2.04	1.37	1.33
2	A	801	FDA	C5'-C4'	2.10	1.54	1.51
2	A	801	FDA	C7M-C7	2.13	1.55	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	801	FDA	N3A-C2A-N1A	-13.75	118.37	128.89
2	A	801	FDA	C4-C4X-C10	-9.94	113.58	119.94
5	A	804	MES	O2S-S-C8	-7.55	100.47	106.91
2	A	801	FDA	C4X-C4-N3	-7.54	113.28	123.59
2	A	801	FDA	C6-C5X-N5	-4.64	113.00	118.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	FDA	5	0
3	A	802	GAF	3	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 [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	576/633 (90%)	-0.23	15 (2%) 59 54	12, 20, 40, 100	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	457	GLY	8.6
1	A	459	VAL	7.8
1	A	343	ALA	5.7
1	A	458	ALA	4.7
1	A	456	TYR	4.5

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(Å ²)	Q<0.9
4	2FG	A	803	12/12	0.93	0.12	1.67	27,31,33,33	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	MES	A	804	12/12	0.96	0.11	0.98	23,26,27,28	0
3	GAF	A	802	12/12	0.91	0.13	0.50	27,30,34,44	0
2	FDA	A	801	53/53	0.98	0.07	-0.60	11,13,16,17	0

6.5 Other polymers [i](#)

There are no such residues in this entry.