Full wwPDB X-ray Structure Validation Report (i)

Jun 24, 2024 – 04:48 PM EDT

PDB ID	:	5O1P
Title	:	Crystal structure of human aminoadipate semialdehyde synthase, saccharopine dehydrogenase.
Authors	:	Kopec, J.; Rembeza, E.; Pena, I.A.; Williams, E.; Velupillai, S.; Kupinska, K.; Strain-Damerell, C.; Goubin, S.; Talon, R.; Collins, P.; Krojer, T.; Burgess-Brown, N.; Arrowsmith, C.; Edwards, A.; Bountra, C.; von Delft, F.; Arruda, P.; Yue, W.W.
Deposited on	:	2017-05-18
Resolution	:	1.90 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
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 (i)) were used in the production of this report:

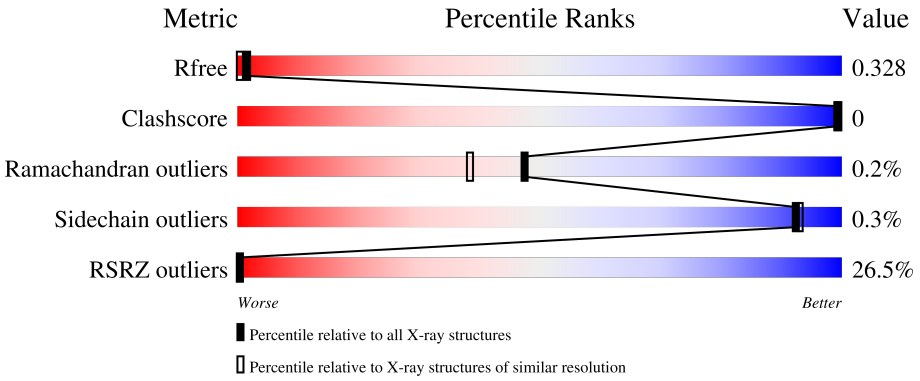
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)

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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	497	24% 88% · 11%

Validation Pipeline (wwPDB-VP) : 2.37.1

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3371 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.

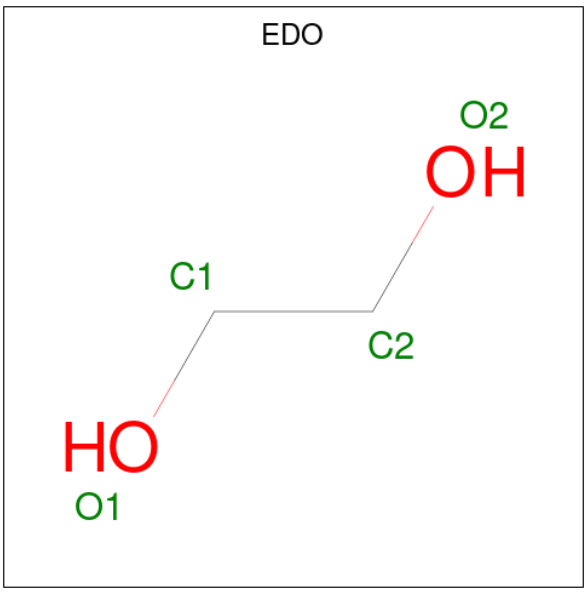
- Molecule 1 is a protein called Alpha-aminoadipic semialdehyde synthase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	442	3325	2124	543	638	20	0	1	0

There are 26 discrepancies between the modelled and reference sequences:

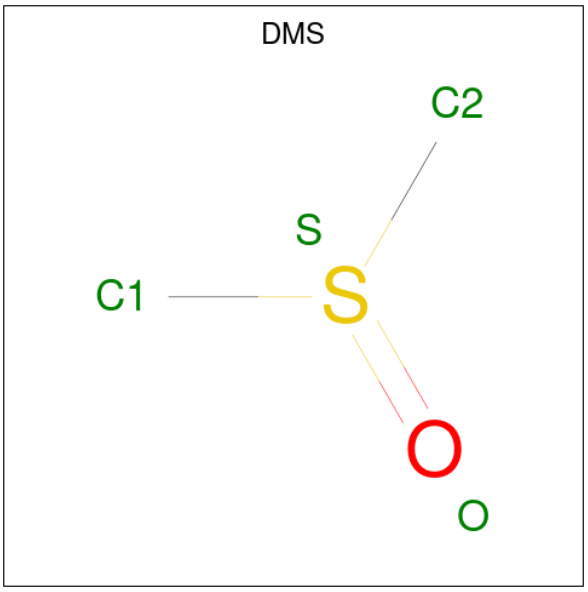
Chain	Residue	Modelled	Actual	Comment	Reference
A	430	MET	-	initiating methionine	UNP Q9UDR5
A	431	GLY	-	expression tag	UNP Q9UDR5
A	432	HIS	-	expression tag	UNP Q9UDR5
A	433	HIS	-	expression tag	UNP Q9UDR5
A	434	HIS	-	expression tag	UNP Q9UDR5
A	435	HIS	-	expression tag	UNP Q9UDR5
A	436	HIS	-	expression tag	UNP Q9UDR5
A	437	HIS	-	expression tag	UNP Q9UDR5
A	438	SER	-	expression tag	UNP Q9UDR5
A	439	SER	-	expression tag	UNP Q9UDR5
A	440	GLY	-	expression tag	UNP Q9UDR5
A	441	VAL	-	expression tag	UNP Q9UDR5
A	442	ASP	-	expression tag	UNP Q9UDR5
A	443	LEU	-	expression tag	UNP Q9UDR5
A	444	GLY	-	expression tag	UNP Q9UDR5
A	445	THR	-	expression tag	UNP Q9UDR5
A	446	GLU	-	expression tag	UNP Q9UDR5
A	447	ASN	-	expression tag	UNP Q9UDR5
A	448	LEU	-	expression tag	UNP Q9UDR5
A	449	TYR	-	expression tag	UNP Q9UDR5
A	450	PHE	-	expression tag	UNP Q9UDR5
A	451	GLN	-	expression tag	UNP Q9UDR5
A	452	SER	-	expression tag	UNP Q9UDR5
A	453	MET	-	expression tag	UNP Q9UDR5
A	454	ALA	-	expression tag	UNP Q9UDR5
A	615	SER	THR	conflict	UNP Q9UDR5

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0

- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O S 4 2 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O S 4 2 1 1	0	0
3	A	1	Total C O S 4 2 1 1	0	0

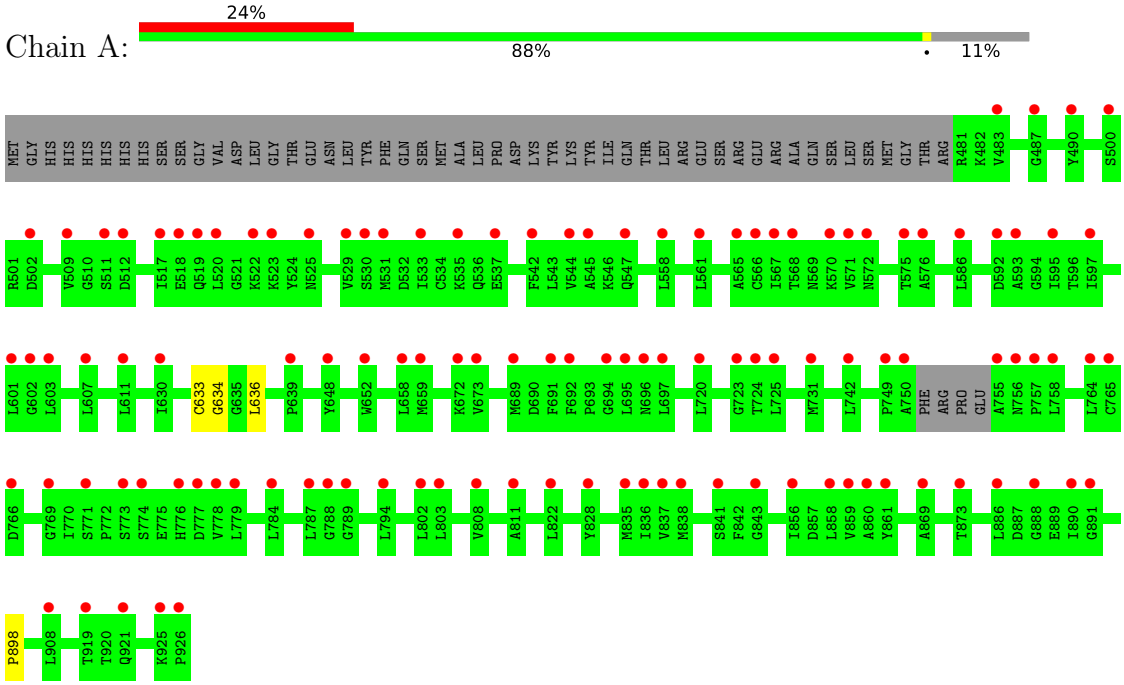
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	26	Total O 26 26	0	0

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: Alpha-aminoadipic semialdehyde synthase, mitochondrial



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	91.46Å 91.46Å 150.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.03 – 1.90 49.03 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.03-1.90) 100.0 (49.03-1.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.04 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.311 , 0.323 0.318 , 0.328	Depositor DCC
R_{free} test set	2554 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	44.0	Xtriage
Anisotropy	0.579	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 57.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3371	wwPDB-VP
Average B, all atoms (Å^2)	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% of the height of the origin peak. No significant pseudotranslation is detected.*

[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: DMS, EDO

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	0.36	0/3387	0.51	0/4598

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	3325	0	3315	1	0
2	A	8	0	12	0	0
3	A	12	0	18	0	0
4	A	26	0	0	0	0
All	All	3371	0	3345	1	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 0.

All (1) 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:633:CYS:SG	1:A:634:GLY:N	2.70	0.63

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	439/497 (88%)	419 (95%)	19 (4%)	1 (0%)	47 38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	898	PRO

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	355/418 (85%)	354 (100%)	1 (0%)	92 93

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	636	LEU

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)

5 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	$\#\|Z\| > 2$
3	DMS	A	1005	-	3,3,3	0.54	0	3,3,3	0.49	0
3	DMS	A	1003	-	3,3,3	0.54	0	3,3,3	0.51	0
3	DMS	A	1004	-	3,3,3	0.54	0	3,3,3	0.51	0
2	EDO	A	1001	-	3,3,3	0.45	0	2,2,2	0.32	0
2	EDO	A	1002	-	3,3,3	0.46	0	2,2,2	0.28	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
2	EDO	A	1001	-	-	1/1/1/1	-
2	EDO	A	1002	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1002	EDO	O1-C1-C2-O2
2	A	1001	EDO	O1-C1-C2-O2

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

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	442/497 (88%)	1.53	117 (26%)	0	0	13, 81, 113, 135	0

All (117) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	787	LEU	12.2
1	A	776	HIS	7.9
1	A	774	SER	7.9
1	A	755	ALA	7.7
1	A	750	ALA	7.5
1	A	567	ILE	7.2
1	A	777	ASP	6.4
1	A	926	PRO	6.3
1	A	789	GLY	5.7
1	A	568	THR	5.2
1	A	672	LYS	4.6
1	A	808	VAL	4.4
1	A	861	TYR	4.3
1	A	837	VAL	4.1
1	A	779	LEU	4.1
1	A	860	ALA	4.1
1	A	891	GLY	4.0
1	A	483	VAL	4.0
1	A	756	ASN	4.0
1	A	773	SER	4.0
1	A	890	ILE	3.9
1	A	531	MET	3.9
1	A	697	LEU	3.8
1	A	533	ILE	3.8
1	A	571	VAL	3.8
1	A	757	PRO	3.7
1	A	519	GLN	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	803	LEU	3.7
1	A	509	VAL	3.6
1	A	836	ILE	3.5
1	A	639	PRO	3.5
1	A	517	ILE	3.5
1	A	542	PHE	3.4
1	A	490	TYR	3.3
1	A	502	ASP	3.3
1	A	758	LEU	3.3
1	A	630	ILE	3.3
1	A	658	LEU	3.2
1	A	692	PHE	3.2
1	A	595	ILE	3.2
1	A	603	LEU	3.2
1	A	512	ASP	3.2
1	A	529	VAL	3.1
1	A	652	TRP	3.1
1	A	788	GLY	3.1
1	A	764	LEU	3.1
1	A	695	LEU	3.0
1	A	607	LEU	3.0
1	A	784	LEU	2.9
1	A	749	PRO	2.9
1	A	518	GLU	2.8
1	A	838	MET	2.8
1	A	547	GLN	2.8
1	A	602	GLY	2.8
1	A	525	ASN	2.8
1	A	561	LEU	2.8
1	A	530	SER	2.8
1	A	648	TYR	2.7
1	A	921	GLN	2.6
1	A	925	LYS	2.6
1	A	869	ALA	2.5
1	A	593	ALA	2.5
1	A	520	LEU	2.5
1	A	720	LEU	2.5
1	A	725	LEU	2.5
1	A	575	THR	2.5
1	A	778	VAL	2.5
1	A	835	MET	2.5
1	A	566	CYS	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	794	LEU	2.5
1	A	659	MET	2.5
1	A	558	LEU	2.5
1	A	858	LEU	2.5
1	A	586	LEU	2.4
1	A	822	LEU	2.4
1	A	673	VAL	2.4
1	A	689	MET	2.4
1	A	769	GLY	2.3
1	A	859	VAL	2.3
1	A	888	GLY	2.3
1	A	544	VAL	2.3
1	A	572	ASN	2.3
1	A	696	ASN	2.3
1	A	765	CYS	2.3
1	A	487	GLY	2.3
1	A	694	GLY	2.3
1	A	742	LEU	2.3
1	A	523	LYS	2.2
1	A	811	ALA	2.2
1	A	601	LEU	2.2
1	A	828	TYR	2.2
1	A	843	GLY	2.2
1	A	802	LEU	2.2
1	A	908	LEU	2.2
1	A	592	ASP	2.2
1	A	537	GLU	2.2
1	A	535	LYS	2.2
1	A	570	LYS	2.2
1	A	597	ILE	2.1
1	A	724	THR	2.1
1	A	500	SER	2.1
1	A	691	PHE	2.1
1	A	919	THR	2.1
1	A	856	ILE	2.1
1	A	886	LEU	2.1
1	A	731	MET	2.1
1	A	841	SER	2.1
1	A	576	ALA	2.1
1	A	611	LEU	2.1
1	A	511	SER	2.0
1	A	522	LYS	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	766	ASP	2.0
1	A	873	THR	2.0
1	A	771	SER	2.0
1	A	545	ALA	2.0
1	A	565	ALA	2.0
1	A	723	GLY	2.0

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 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	DMS	A	1003	4/4	0.80	0.33	136,136,137,137	0
2	EDO	A	1001	4/4	0.82	0.44	86,87,87,88	0
2	EDO	A	1002	4/4	0.85	0.13	92,92,93,93	0
3	DMS	A	1004	4/4	0.86	0.14	119,119,120,120	0
3	DMS	A	1005	4/4	0.91	0.25	115,115,116,116	0

6.5 Other polymers (i)

There are no such residues in this entry.