



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 4, 2024 – 03:31 AM EST

PDB ID : 2HP0
Title : Crystal structure of iminodisuccinate epimerase
Authors : Lohkamp, B.; Bauerle, B.; Rieger, P.G.; Schneider, G.
Deposited on : 2006-07-17
Resolution : 1.50 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 ⓘ 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 ⓘ](#)) 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	:	FAILED
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

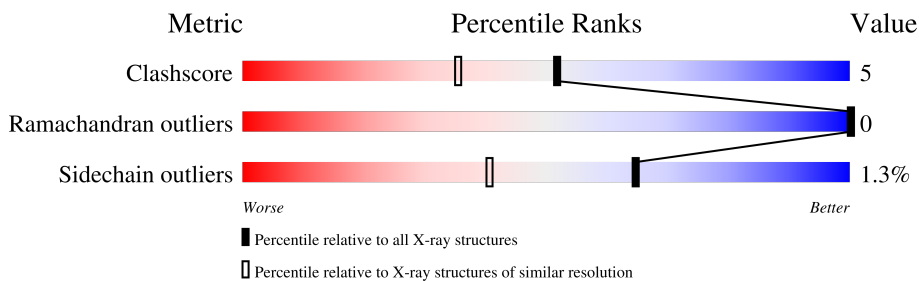
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 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))
Clashscore	180529	4048 (1.50-1.50)
Ramachandran outliers	177936	3970 (1.50-1.50)
Sidechain outliers	177891	3967 (1.50-1.50)

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 $\leq 5\%$. 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.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	466	
2	B	466	

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
3	EDO	A	1003	-	-	X	-
4	UNX	A	1102	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	UNX	A	1103	-	-	X	-
4	UNX	A	1107	-	-	X	-
4	UNX	B	1117	-	-	X	-
5	DTU	B	1020	-	X	-	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7982 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 IDS-epimerase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	447	Total	C	N	O	S	Se	31	20	0
			3452	2207	580	640	6	19			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MSE	-	SEE REMARK 999	UNP Q1L4E3
A	-18	GLY	-	cloning artifact	UNP Q1L4E3
A	-17	SER	-	cloning artifact	UNP Q1L4E3
A	-16	SER	-	cloning artifact	UNP Q1L4E3
A	-15	HIS	-	expression tag	UNP Q1L4E3
A	-14	HIS	-	expression tag	UNP Q1L4E3
A	-13	HIS	-	expression tag	UNP Q1L4E3
A	-12	HIS	-	expression tag	UNP Q1L4E3
A	-11	HIS	-	expression tag	UNP Q1L4E3
A	-10	HIS	-	expression tag	UNP Q1L4E3
A	-9	SER	-	cloning artifact	UNP Q1L4E3
A	-8	SER	-	cloning artifact	UNP Q1L4E3
A	-7	GLY	-	cloning artifact	UNP Q1L4E3
A	-6	LEU	-	cloning artifact	UNP Q1L4E3
A	-5	VAL	-	cloning artifact	UNP Q1L4E3
A	-4	PRO	-	cloning artifact	UNP Q1L4E3
A	-3	ARG	-	cloning artifact	UNP Q1L4E3
A	-2	GLY	-	cloning artifact	UNP Q1L4E3
A	-1	SER	-	cloning artifact	UNP Q1L4E3
A	0	HIS	-	cloning artifact	UNP Q1L4E3
A	1	MSE	MET	modified residue	UNP Q1L4E3
A	73	MSE	MET	modified residue	UNP Q1L4E3
A	82	MSE	MET	modified residue	UNP Q1L4E3
A	96	MSE	MET	modified residue	UNP Q1L4E3
A	196	MSE	MET	modified residue	UNP Q1L4E3
A	200	MSE	MET	modified residue	UNP Q1L4E3
A	217	MSE	MET	modified residue	UNP Q1L4E3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	236	MSE	MET	modified residue	UNP Q1L4E3
A	249	MSE	MET	modified residue	UNP Q1L4E3
A	285	MSE	MET	modified residue	UNP Q1L4E3
A	313	MSE	MET	modified residue	UNP Q1L4E3
A	353	MSE	MET	modified residue	UNP Q1L4E3
A	357	MSE	MET	modified residue	UNP Q1L4E3
A	365	MSE	MET	modified residue	UNP Q1L4E3
A	404	MSE	MET	modified residue	UNP Q1L4E3
A	420	MSE	MET	modified residue	UNP Q1L4E3

- Molecule 2 is a protein called IDS-epimerase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	447	Total	C	N	O	S	Se	38	31	0
			3520	2249	595	650	7	19			

There are 37 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MSE	-	SEE REMARK 999	UNP Q1L4E3
B	-18	GLY	-	cloning artifact	UNP Q1L4E3
B	-17	SER	-	cloning artifact	UNP Q1L4E3
B	-16	SER	-	cloning artifact	UNP Q1L4E3
B	-15	HIS	-	expression tag	UNP Q1L4E3
B	-14	HIS	-	expression tag	UNP Q1L4E3
B	-13	HIS	-	expression tag	UNP Q1L4E3
B	-12	HIS	-	expression tag	UNP Q1L4E3
B	-11	HIS	-	expression tag	UNP Q1L4E3
B	-10	HIS	-	expression tag	UNP Q1L4E3
B	-9	SER	-	cloning artifact	UNP Q1L4E3
B	-8	SER	-	cloning artifact	UNP Q1L4E3
B	-7	GLY	-	cloning artifact	UNP Q1L4E3
B	-6	LEU	-	cloning artifact	UNP Q1L4E3
B	-5	VAL	-	cloning artifact	UNP Q1L4E3
B	-4	PRO	-	cloning artifact	UNP Q1L4E3
B	-3	ARG	-	cloning artifact	UNP Q1L4E3
B	-2	GLY	-	cloning artifact	UNP Q1L4E3
B	-1	SER	-	cloning artifact	UNP Q1L4E3
B	0	HIS	-	cloning artifact	UNP Q1L4E3
B	1	MSE	MET	modified residue	UNP Q1L4E3
B	73	MSE	MET	modified residue	UNP Q1L4E3
B	82	MSE	MET	modified residue	UNP Q1L4E3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	96	MSE	MET	modified residue	UNP Q1L4E3
B	196	MSE	MET	modified residue	UNP Q1L4E3
B	200	MSE	MET	modified residue	UNP Q1L4E3
B	217	MSE	MET	modified residue	UNP Q1L4E3
B	236	MSE	MET	modified residue	UNP Q1L4E3
B	249	MSE	MET	modified residue	UNP Q1L4E3
B	271	CSO	CYS	modified residue	UNP Q1L4E3
B	285	MSE	MET	modified residue	UNP Q1L4E3
B	313	MSE	MET	modified residue	UNP Q1L4E3
B	353	MSE	MET	modified residue	UNP Q1L4E3
B	357	MSE	MET	modified residue	UNP Q1L4E3
B	365	MSE	MET	modified residue	UNP Q1L4E3
B	404	MSE	MET	modified residue	UNP Q1L4E3
B	420	MSE	MET	modified residue	UNP Q1L4E3

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

Continued on next page...

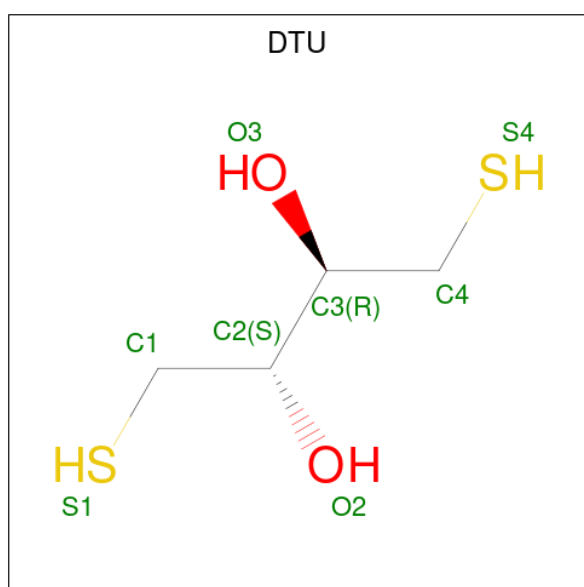
Continued from previous page...

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

- Molecule 4 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	7	Total	X	0	0
			7	7		
4	B	7	Total	X	0	0
			7	7		

- Molecule 5 is (2R,3S)-1,4-DIMERCAPTOBUTANE-2,3-DIOL (three-letter code: DTU) (formula: $C_4H_{10}O_2S_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	O	S	7	0
			8	4	2	2		
5	B	1	Total	C	O	S	7	0
			8	4	2	2		

- Molecule 6 is water.

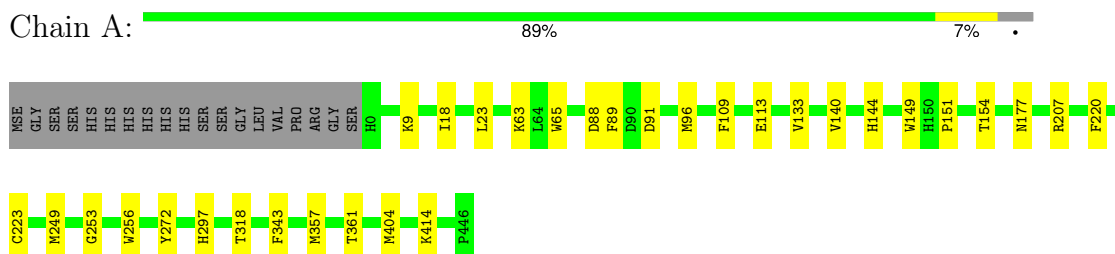
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	503	Total 509	O 509	0	5
6	B	445	Total 447	O 447	0	2

3 Residue-property plots

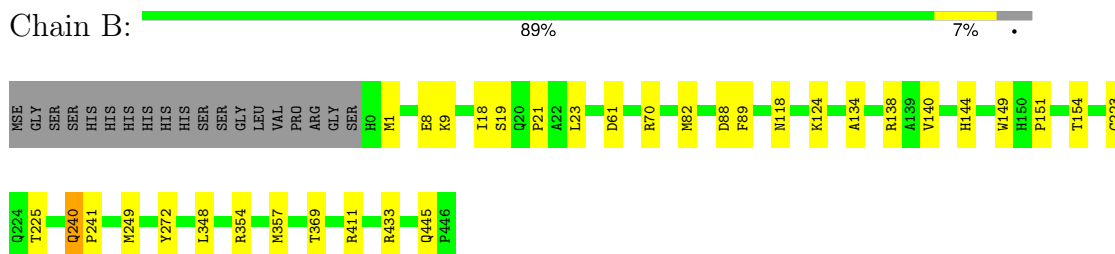
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

Note EDS failed to run properly.

- Molecule 1: IDS-epimerase



- Molecule 2: IDS-epimerase



4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.60Å 104.97Å 78.64Å 90.00° 102.51° 90.00°	Depositor
Resolution (Å)	62.02 – 1.50	Depositor
% Data completeness (in resolution range)	99.7 (62.02-1.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.90 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.151 , 0.174	Depositor
Wilson B-factor (Å ²)	16.4	Xtriage
Anisotropy	0.043	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7982	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.48% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: UNX, CSO, DTU, 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	0.60	1/3580 (0.0%)	0.68	2/4825 (0.0%)
2	B	0.57	1/3653 (0.0%)	0.66	0/4920
All	All	0.58	2/7233 (0.0%)	0.67	2/9745 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	357	MSE	SE-CE	-5.98	1.60	1.95
1	A	357	MSE	SE-CE	-5.93	1.60	1.95

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	207	ARG	NE-CZ-NH1	5.47	123.03	120.30
1	A	91	ASP	CB-CG-OD1	5.19	122.97	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	3452	0	3459	35	0
2	B	3520	0	3551	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	16	0	24	10	0
3	B	8	0	12	0	0
4	A	7	0	0	6	0
4	B	7	0	0	4	0
5	B	16	0	20	2	0
6	A	509	0	0	6	0
6	B	447	0	0	10	0
All	All	7982	0	7066	69	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 5.

The worst 5 of 69 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:96[A]:MSE:SE	1:A:133[A]:VAL:CG1	2.36	1.24
1:A:9:LYS:HE3	6:A:1332:HOH:O	1.40	1.20
1:A:223:CYS:SG	5:B:1020:DTU:S1	2.38	1.18
6:A:1380:HOH:O	2:B:149[B]:TRP:CH2	1.95	1.17
6:A:1380:HOH:O	2:B:149[B]:TRP:CZ3	1.99	1.16

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	466/466 (100%)	463 (99%)	3 (1%)	0	100	100
2	B	474/466 (102%)	471 (99%)	3 (1%)	0	100	100
All	All	940/932 (101%)	934 (99%)	6 (1%)	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	367/345 (106%)	364 (99%)	3 (1%)	79	62
2	B	373/344 (108%)	367 (98%)	6 (2%)	58	32
All	All	740/689 (107%)	731 (99%)	9 (1%)	65	44

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	272	TYR
2	B	433	ARG
2	B	61	ASP
2	B	88	ASP
2	B	89	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	CSO	B	271[B]	-	3,6,7	0.79	0	1,6,8	0.24	0
2	CSO	B	271[A]	-	3,6,7	0.78	0	1,6,8	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	CSO	B	271[B]	-	-	0/1/5/7	-
2	CSO	B	271[A]	-	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 14 are unknown - leaving 8 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
5	DTU	B	1021	-	7,7,7	0.76	0	4,8,8	2.17	1 (25%)
3	EDO	B	1002	-	3,3,3	0.57	0	2,2,2	0.25	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	EDO	A	1006	-	3,3,3	0.38	0	2,2,2	0.42	0
5	DTU	B	1020	-	7,7,7	3.04	1 (14%)	4,8,8	1.29	1 (25%)
3	EDO	A	1004	-	3,3,3	0.48	0	2,2,2	0.26	0
3	EDO	A	1003	-	3,3,3	0.27	0	2,2,2	0.19	0
3	EDO	A	1005	-	3,3,3	0.29	0	2,2,2	0.41	0
3	EDO	B	1001	-	3,3,3	0.28	0	2,2,2	0.33	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
5	DTU	B	1021	-	-	7/8/8/8	-
3	EDO	B	1002	-	-	1/1/1/1	-
3	EDO	A	1006	-	-	1/1/1/1	-
5	DTU	B	1020	-	-	8/8/8/8	-
3	EDO	A	1004	-	-	1/1/1/1	-
3	EDO	A	1003	-	-	1/1/1/1	-
3	EDO	A	1005	-	-	1/1/1/1	-
3	EDO	B	1001	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1020	DTU	C1-S1	7.67	1.97	1.81

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1021	DTU	C2-C1-S1	3.89	125.31	114.43
5	B	1020	DTU	O2-C2-C3	2.02	113.81	109.57

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	1020	DTU	S1-C1-C2-O2
5	B	1020	DTU	S1-C1-C2-C3
5	B	1020	DTU	C1-C2-C3-O3
5	B	1020	DTU	C1-C2-C3-C4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	B	1020	DTU	O2-C2-C3-O3

There are no ring outliers.

5 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1021	DTU	1	0
3	A	1006	EDO	1	0
5	B	1020	DTU	1	0
3	A	1003	EDO	7	0
3	A	1005	EDO	2	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](#)

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

6.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.