



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 31, 2024 – 12:31 PM JST

PDB ID : 8HSN  
Title : Crystal structure of DFA I-forming Inulin Lyase from *Streptomyces peucetius* subsp. *caesius* ATCC 27952  
Authors : Cheng, M.; Rao, Y.J.; Mu, W.M.  
Deposited on : 2022-12-20  
Resolution : 1.69 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](mailto: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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

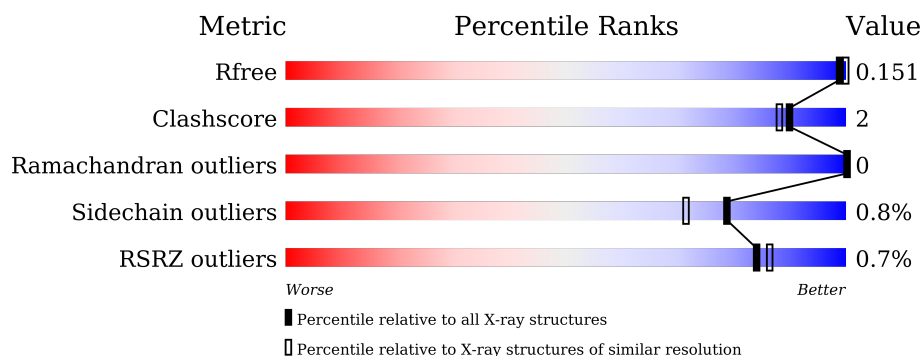
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.69 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))
$R_{free}$	164625	5161 (1.70-1.70)
Clashscore	180529	5671 (1.70-1.70)
Ramachandran outliers	177936	5594 (1.70-1.70)
Sidechain outliers	177891	5594 (1.70-1.70)
RSRZ outliers	164620	5159 (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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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.

Mol	Chain	Length	Quality of chain
1	A	400	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>5%</div> <div>.</div> </div> </div>
1	B	400	<div> <div></div> <div>94%</div> <div>.</div> <div>.</div> </div>
1	C	400	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>.</div> <div>..</div> </div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 18817 atoms, of which 9067 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 Fructotransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	395	Total	C	H	N	O	S	0	7	0
			5862	1857	2892	523	584	6			
1	B	394	Total	C	H	N	O	S	0	5	0
			5866	1856	2899	521	584	6			
1	C	395	Total	C	H	N	O	S	0	6	0
			5888	1870	2903	525	584	6			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	395	HIS	-	expression tag	UNP A0A2D3U3Z1
A	396	HIS	-	expression tag	UNP A0A2D3U3Z1
A	397	HIS	-	expression tag	UNP A0A2D3U3Z1
A	398	HIS	-	expression tag	UNP A0A2D3U3Z1
A	399	HIS	-	expression tag	UNP A0A2D3U3Z1
A	400	HIS	-	expression tag	UNP A0A2D3U3Z1
B	395	HIS	-	expression tag	UNP A0A2D3U3Z1
B	396	HIS	-	expression tag	UNP A0A2D3U3Z1
B	397	HIS	-	expression tag	UNP A0A2D3U3Z1
B	398	HIS	-	expression tag	UNP A0A2D3U3Z1
B	399	HIS	-	expression tag	UNP A0A2D3U3Z1
B	400	HIS	-	expression tag	UNP A0A2D3U3Z1
C	395	HIS	-	expression tag	UNP A0A2D3U3Z1
C	396	HIS	-	expression tag	UNP A0A2D3U3Z1
C	397	HIS	-	expression tag	UNP A0A2D3U3Z1
C	398	HIS	-	expression tag	UNP A0A2D3U3Z1
C	399	HIS	-	expression tag	UNP A0A2D3U3Z1
C	400	HIS	-	expression tag	UNP A0A2D3U3Z1

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



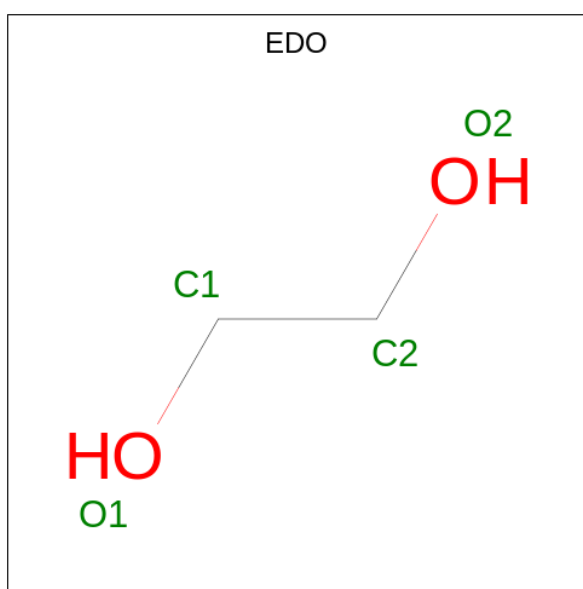
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	B	1	Total	C	H	O	0	0
			14	3	8	3		
2	B	1	Total	C	H	O	0	0
			14	3	8	3		
2	B	1	Total	C	H	O	0	0
			14	3	8	3		
2	B	1	Total	C	H	O	0	0
			14	3	8	3		
2	C	1	Total	C	H	O	0	0
			14	3	8	3		
2	C	1	Total	C	H	O	0	0
			14	3	8	3		
2	C	1	Total	C	H	O	0	0
			14	3	8	3		
2	C	1	Total	C	H	O	0	0
			14	3	8	3		

*Continued on next page...*

Continued from previous page...

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

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



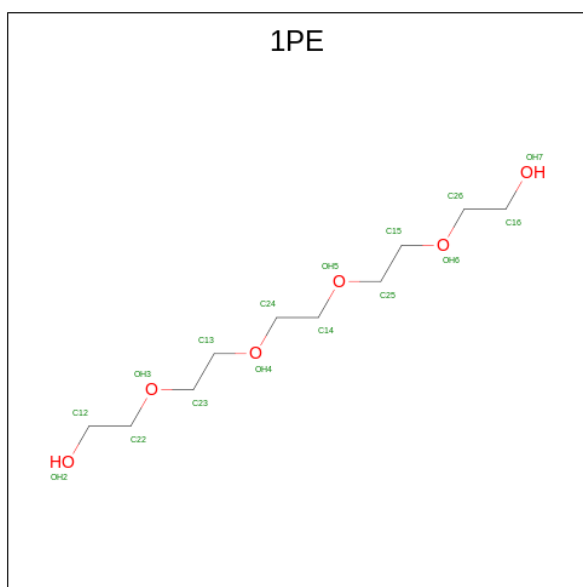
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			10	2	6	2		
3	A	1	Total	C	H	O	0	0
			10	2	6	2		
3	A	1	Total	C	H	O	0	0
			10	2	6	2		
3	A	1	Total	C	H	O	0	0
			10	2	6	2		
3	A	1	Total	C	H	O	0	0
			10	2	6	2		
3	B	1	Total	C	H	O	0	0
			10	2	6	2		

Continued on next page...

*Continued from previous page...*

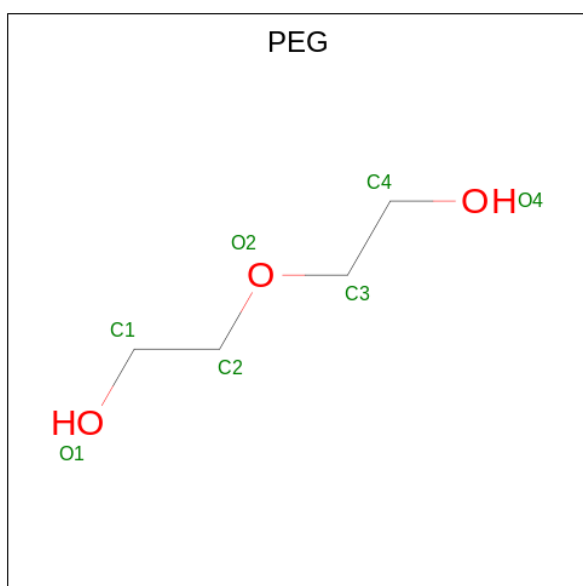
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	H	O	0	0
			10	2	6	2		
3	B	1	Total	C	H	O	0	0
			10	2	6	2		
3	B	1	Total	C	H	O	0	0
			10	2	6	2		
3	B	1	Total	C	H	O	0	0
			10	2	6	2		
3	B	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		
3	C	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 4 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			38	10	22	6		
4	B	1	Total	C	H	O	0	0
			38	10	22	6		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



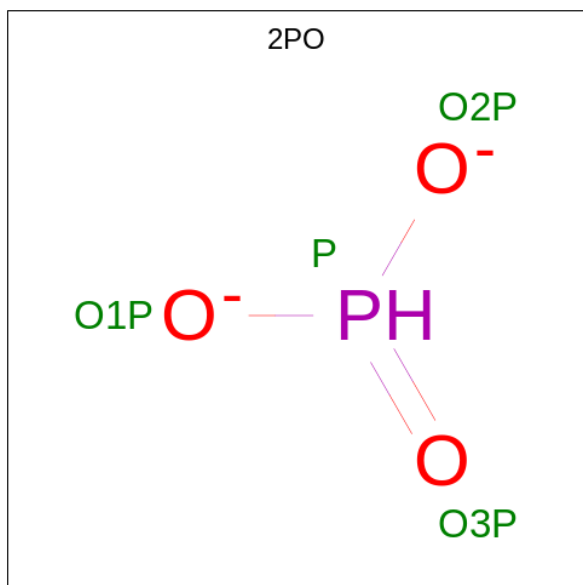
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			17	4	10	3		
5	B	1	Total	C	H	O	0	0
			17	4	10	3		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	0
			17	4	10	3		
5	B	1	Total	C	H	O	0	0
			17	4	10	3		
5	C	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 6 is PHOSPHONATE (three-letter code: 2PO) (formula:  $\text{HO}_3\text{P}$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	H	O	P	0	0
			5	1	3	1		
6	B	1	Total	H	O	P	0	0
			5	1	3	1		
6	C	1	Total	H	O	P	0	0
			5	1	3	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	171	Total	O	0	0
			171	171		
7	B	189	Total	O	0	0
			189	189		
7	C	193	Total	O	0	0
			193	193		



### 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: Fructotransferase

Chain A: 



- Molecule 1: Fructotransferase

Chain B: 



- Molecule 1: Fructotransferase

Chain C: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.68Å 102.51Å 132.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.81 – 1.69 47.81 – 1.69	Depositor EDS
% Data completeness (in resolution range)	99.0 (47.81-1.69) 99.0 (47.81-1.69)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.65 (at 1.69Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.137 , 0.151 0.137 , 0.151	Depositor DCC
$R_{free}$ test set	147038 reflections (1.35%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	11.9	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.45 , 35.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.005 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	18817	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.83% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: PEG, 2PO, GOL, EDO, 1PE

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	0.41	0/3045	0.67	0/4154
1	B	0.42	0/3033	0.68	0/4138
1	C	0.40	0/3071	0.66	0/4193
All	All	0.41	0/9149	0.67	0/12485

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 [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	2970	2892	2878	10	0
1	B	2967	2899	2894	10	0
1	C	2985	2903	2895	13	0
2	A	36	48	48	0	0
2	B	24	32	32	0	0
2	C	48	64	64	0	0
3	A	24	36	36	0	0
3	B	28	42	42	0	0
3	C	36	54	54	0	0
4	A	16	22	22	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	16	22	22	0	0
5	A	7	10	10	0	0
5	B	21	30	30	0	0
5	C	7	10	10	0	0
6	A	4	1	1	0	0
6	B	4	1	1	0	0
6	C	4	1	1	0	0
7	A	171	0	0	0	0
7	B	189	0	0	0	0
7	C	193	0	0	0	0
All	All	9750	9067	9040	31	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 2.

All (31) 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:C:344:ASN:ND2	1:C:370:ARG:HH12	2.09	0.49
1:C:303:MET:HB2	1:C:336:ILE:O	2.14	0.47
1:B:281:SER:HA	1:B:314:THR:O	2.15	0.46
1:A:167:GLU:HA	1:A:189:THR:O	2.16	0.44
1:A:260:LYS:HB2	1:C:234:GLN:HG3	1.98	0.44
1:B:234:GLN:HE22	1:C:237:GLY:HA2	1.83	0.44
1:B:167:GLU:HA	1:B:189:THR:O	2.17	0.44
1:C:167:GLU:HA	1:C:189:THR:O	2.18	0.44
1:B:314:THR:HA	1:B:348:ALA:O	2.18	0.43
1:B:212:THR:HA	1:B:236:SER:O	2.19	0.43
1:A:153:THR:HA	1:A:176:HIS:O	2.18	0.43
1:A:281:SER:HA	1:A:314:THR:O	2.18	0.43
1:C:212:THR:HA	1:C:236:SER:O	2.19	0.43
1:B:106:ASN:OD1	1:B:106:ASN:C	2.58	0.42
1:C:5:VAL:HG22	1:C:47:VAL:CG1	2.49	0.42
1:C:257:THR:HA	1:C:281:SER:O	2.20	0.42
1:C:281:SER:HA	1:C:314:THR:O	2.20	0.42
1:B:133:GLN:HA	1:B:167:GLU:O	2.19	0.42
1:C:133:GLN:HA	1:C:167:GLU:O	2.20	0.42
1:B:257:THR:HA	1:B:281:SER:O	2.21	0.41
1:B:139:GLY:HA3	1:B:173:TYR:O	2.20	0.41
1:B:180:ILE:HG22	1:B:183:ALA:HB2	2.02	0.41
1:A:257:THR:HA	1:A:281:SER:O	2.20	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:335:THR:HG21	1:C:338:LEU:HB2	2.03	0.41
1:C:189:THR:HA	1:C:212:THR:O	2.20	0.41
1:A:133:GLN:HA	1:A:167:GLU:O	2.20	0.41
1:A:190:ASN:HA	1:A:213:ASN:O	2.21	0.41
1:A:300:LYS:HE2	1:A:332:GLN:HB3	2.02	0.41
1:A:328:THR:HA	1:A:329:PRO:C	2.41	0.41
1:A:189:THR:HA	1:A:212:THR:O	2.21	0.40
1:C:139:GLY:HA3	1:C:173:TYR:O	2.21	0.40

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	396/400 (99%)	380 (96%)	16 (4%)	0	100	100
1	B	395/400 (99%)	379 (96%)	16 (4%)	0	100	100
1	C	399/400 (100%)	385 (96%)	14 (4%)	0	100	100
All	All	1190/1200 (99%)	1144 (96%)	46 (4%)	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	326/328 (99%)	323 (99%)	3 (1%)	75	67
1	B	325/328 (99%)	323 (99%)	2 (1%)	84	78
1	C	327/328 (100%)	324 (99%)	3 (1%)	75	67
All	All	978/984 (99%)	970 (99%)	8 (1%)	79	71

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	173	TYR
1	A	234	GLN
1	A	253	ARG
1	B	233	LEU
1	B	253	ARG
1	C	173	TYR
1	C	234	GLN
1	C	370	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	87	GLN
1	B	234	GLN
1	C	87	GLN
1	C	234	GLN
1	C	344	ASN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry

50 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	EDO	B	514	-	3,3,3	0.43	0	2,2,2	0.52	0
6	2PO	C	501	-	0,3,3	-	-	0,3,3	-	-
3	EDO	B	501	-	3,3,3	0.55	0	2,2,2	0.13	0
6	2PO	A	515	-	0,3,3	-	-	0,3,3	-	-
4	1PE	B	507	-	15,15,15	0.33	0	14,14,14	0.36	0
3	EDO	C	507	-	3,3,3	0.49	0	2,2,2	0.45	0
3	EDO	A	504	-	3,3,3	0.71	0	2,2,2	0.12	0
3	EDO	A	505	-	3,3,3	0.70	0	2,2,2	0.36	0
5	PEG	B	512	-	6,6,6	0.19	0	5,5,5	0.14	0
2	GOL	A	501	-	5,5,5	0.84	0	5,5,5	1.04	0
2	GOL	A	510	-	5,5,5	0.99	0	5,5,5	1.09	0
3	EDO	C	517	-	3,3,3	0.44	0	2,2,2	0.62	0
2	GOL	A	509	-	5,5,5	1.06	0	5,5,5	1.86	2 (40%)
3	EDO	C	516	-	3,3,3	0.44	0	2,2,2	1.17	0
2	GOL	C	511	-	5,5,5	1.42	1 (20%)	5,5,5	1.04	0
2	GOL	C	513	-	5,5,5	0.84	0	5,5,5	1.06	0
3	EDO	A	506	-	3,3,3	0.40	0	2,2,2	0.49	0
5	PEG	B	508	-	6,6,6	0.13	0	5,5,5	0.11	0
6	2PO	B	502	-	0,3,3	-	-	0,3,3	-	-
2	GOL	C	510	-	5,5,5	1.16	0	5,5,5	1.19	1 (20%)
2	GOL	A	511	-	5,5,5	1.14	0	5,5,5	0.83	0
5	PEG	B	505	-	6,6,6	0.12	0	5,5,5	0.14	0
2	GOL	B	509	-	5,5,5	0.67	0	5,5,5	1.22	1 (20%)
2	GOL	B	511	-	5,5,5	1.20	0	5,5,5	0.99	0
2	GOL	C	509	-	5,5,5	0.76	0	5,5,5	0.89	0
3	EDO	B	503	-	3,3,3	0.46	0	2,2,2	0.61	0
3	EDO	C	508	-	3,3,3	0.35	0	2,2,2	0.89	0
3	EDO	C	505	-	3,3,3	0.44	0	2,2,2	0.32	0
3	EDO	A	514	-	3,3,3	0.49	0	2,2,2	0.42	0
2	GOL	A	512	-	5,5,5	0.68	0	5,5,5	0.87	0
3	EDO	C	518	-	3,3,3	0.45	0	2,2,2	0.39	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	EDO	B	515	-	3,3,3	0.39	0	2,2,2	0.61	0
3	EDO	A	508	-	3,3,3	0.61	0	2,2,2	0.79	0
2	GOL	C	514	-	5,5,5	1.05	0	5,5,5	0.94	0
2	GOL	B	510	-	5,5,5	0.86	0	5,5,5	0.68	0
2	GOL	B	504	-	5,5,5	0.49	0	5,5,5	1.04	0
3	EDO	B	516	-	3,3,3	0.50	0	2,2,2	0.54	0
2	GOL	C	502	-	5,5,5	0.99	0	5,5,5	1.07	0
3	EDO	B	506	-	3,3,3	0.48	0	2,2,2	0.34	0
2	GOL	C	512	-	5,5,5	0.98	0	5,5,5	1.12	1 (20%)
4	1PE	A	507	-	15,15,15	0.36	0	14,14,14	0.48	0
3	EDO	C	504	-	3,3,3	0.43	0	2,2,2	0.85	0
3	EDO	B	513	-	3,3,3	0.58	0	2,2,2	0.28	0
3	EDO	C	519	-	3,3,3	0.43	0	2,2,2	0.54	0
3	EDO	A	503	-	3,3,3	0.58	0	2,2,2	0.60	0
5	PEG	C	515	-	6,6,6	0.28	0	5,5,5	0.10	0
5	PEG	A	513	-	6,6,6	0.10	0	5,5,5	0.27	0
3	EDO	C	506	-	3,3,3	0.52	0	2,2,2	0.49	0
2	GOL	A	502	-	5,5,5	0.81	0	5,5,5	0.73	0
2	GOL	C	503	-	5,5,5	0.95	0	5,5,5	0.69	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
3	EDO	B	514	-	-	1/1/1/1	-
3	EDO	B	501	-	-	1/1/1/1	-
4	1PE	B	507	-	-	9/13/13/13	-
3	EDO	C	507	-	-	1/1/1/1	-
3	EDO	A	504	-	-	1/1/1/1	-
3	EDO	A	505	-	-	0/1/1/1	-
5	PEG	B	512	-	-	4/4/4/4	-
2	GOL	A	501	-	-	0/4/4/4	-
2	GOL	A	510	-	-	0/4/4/4	-
3	EDO	C	517	-	-	1/1/1/1	-
2	GOL	A	509	-	-	2/4/4/4	-
3	EDO	C	516	-	-	1/1/1/1	-
2	GOL	C	511	-	-	2/4/4/4	-
2	GOL	C	513	-	-	2/4/4/4	-
3	EDO	A	506	-	-	1/1/1/1	-

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PEG	B	508	-	-	2/4/4/4	-
2	GOL	C	510	-	-	0/4/4/4	-
2	GOL	A	511	-	-	2/4/4/4	-
5	PEG	B	505	-	-	1/4/4/4	-
2	GOL	B	509	-	-	2/4/4/4	-
2	GOL	B	511	-	-	2/4/4/4	-
2	GOL	C	509	-	-	2/4/4/4	-
3	EDO	B	503	-	-	1/1/1/1	-
3	EDO	C	508	-	-	1/1/1/1	-
3	EDO	C	505	-	-	1/1/1/1	-
3	EDO	A	514	-	-	1/1/1/1	-
2	GOL	A	512	-	-	0/4/4/4	-
3	EDO	C	518	-	-	1/1/1/1	-
3	EDO	B	515	-	-	1/1/1/1	-
3	EDO	A	508	-	-	1/1/1/1	-
2	GOL	C	514	-	-	2/4/4/4	-
2	GOL	B	510	-	-	2/4/4/4	-
2	GOL	B	504	-	-	0/4/4/4	-
3	EDO	B	516	-	-	1/1/1/1	-
2	GOL	C	502	-	-	0/4/4/4	-
3	EDO	B	506	-	-	1/1/1/1	-
2	GOL	C	512	-	-	1/4/4/4	-
4	1PE	A	507	-	-	7/13/13/13	-
3	EDO	C	504	-	-	0/1/1/1	-
3	EDO	B	513	-	-	0/1/1/1	-
3	EDO	C	519	-	-	1/1/1/1	-
3	EDO	A	503	-	-	1/1/1/1	-
5	PEG	C	515	-	-	3/4/4/4	-
5	PEG	A	513	-	-	2/4/4/4	-
3	EDO	C	506	-	-	0/1/1/1	-
2	GOL	A	502	-	-	0/4/4/4	-
2	GOL	C	503	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	511	GOL	C3-C2	2.03	1.60	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	509	GOL	C3-C2-C1	-3.24	99.09	111.70
2	B	509	GOL	C3-C2-C1	-2.36	102.55	111.70
2	C	510	GOL	C3-C2-C1	-2.26	102.92	111.70
2	A	509	GOL	O2-C2-C1	2.19	118.75	109.12
2	C	512	GOL	C3-C2-C1	-2.11	103.50	111.70

There are no chirality outliers.

All (65) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	509	GOL	C1-C2-C3-O3
2	A	511	GOL	O1-C1-C2-C3
2	B	509	GOL	O1-C1-C2-C3
2	B	510	GOL	O1-C1-C2-C3
2	C	511	GOL	C1-C2-C3-O3
2	C	514	GOL	O1-C1-C2-C3
4	B	507	1PE	OH6-C15-C25-OH5
4	A	507	1PE	OH5-C14-C24-OH4
2	B	510	GOL	O1-C1-C2-O2
2	C	513	GOL	O2-C2-C3-O3
5	B	508	PEG	O1-C1-C2-O2
5	B	512	PEG	O1-C1-C2-O2
5	B	512	PEG	O2-C3-C4-O4
5	C	515	PEG	O2-C3-C4-O4
2	B	511	GOL	C1-C2-C3-O3
2	C	509	GOL	C1-C2-C3-O3
2	C	513	GOL	C1-C2-C3-O3
2	A	509	GOL	O2-C2-C3-O3
2	B	509	GOL	O1-C1-C2-O2
2	B	511	GOL	O2-C2-C3-O3
2	C	511	GOL	O2-C2-C3-O3
5	C	515	PEG	O1-C1-C2-O2
3	A	504	EDO	O1-C1-C2-O2
3	A	514	EDO	O1-C1-C2-O2
3	B	503	EDO	O1-C1-C2-O2
3	B	514	EDO	O1-C1-C2-O2
3	B	515	EDO	O1-C1-C2-O2
3	C	507	EDO	O1-C1-C2-O2
4	A	507	1PE	OH6-C15-C25-OH5
4	B	507	1PE	OH2-C12-C22-OH3
2	A	511	GOL	O1-C1-C2-O2
5	A	513	PEG	O2-C3-C4-O4
3	C	508	EDO	O1-C1-C2-O2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	C	512	GOL	O2-C2-C3-O3
2	C	514	GOL	O1-C1-C2-O2
3	A	506	EDO	O1-C1-C2-O2
3	B	516	EDO	O1-C1-C2-O2
4	B	507	1PE	C13-C23-OH3-C22
4	B	507	1PE	C24-C14-OH5-C25
4	B	507	1PE	C25-C15-OH6-C26
3	C	505	EDO	O1-C1-C2-O2
5	C	515	PEG	C1-C2-O2-C3
4	B	507	1PE	OH7-C16-C26-OH6
4	B	507	1PE	OH5-C14-C24-OH4
5	B	512	PEG	C4-C3-O2-C2
5	B	508	PEG	C1-C2-O2-C3
4	B	507	1PE	C16-C26-OH6-C15
4	A	507	1PE	OH4-C13-C23-OH3
4	A	507	1PE	OH2-C12-C22-OH3
5	A	513	PEG	O1-C1-C2-O2
5	B	505	PEG	C4-C3-O2-C2
2	C	509	GOL	O2-C2-C3-O3
5	B	512	PEG	C1-C2-O2-C3
3	C	516	EDO	O1-C1-C2-O2
4	A	507	1PE	C14-C24-OH4-C13
3	A	508	EDO	O1-C1-C2-O2
3	C	517	EDO	O1-C1-C2-O2
3	C	518	EDO	O1-C1-C2-O2
4	A	507	1PE	C13-C23-OH3-C22
3	A	503	EDO	O1-C1-C2-O2
3	B	501	EDO	O1-C1-C2-O2
3	B	506	EDO	O1-C1-C2-O2
4	A	507	1PE	C12-C22-OH3-C23
3	C	519	EDO	O1-C1-C2-O2
4	B	507	1PE	C15-C25-OH5-C14

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	395/400 (98%)	-0.93	4 (1%)	79 81	5, 12, 24, 46	1 (0%)
1	B	394/400 (98%)	-0.91	1 (0%)	90 91	5, 13, 22, 35	2 (0%)
1	C	395/400 (98%)	-0.93	3 (0%)	82 85	6, 12, 24, 45	2 (0%)
All	All	1184/1200 (98%)	-0.93	8 (0%)	84 86	5, 12, 23, 46	5 (0%)

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	396	HIS	4.0
1	C	395	HIS	3.7
1	A	395	HIS	3.6
1	C	396	HIS	3.5
1	C	331	GLY	2.9
1	A	234	GLN	2.8
1	B	395	HIS	2.5
1	A	331	GLY	2.2

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	EDO	B	503	4/4	0.71	0.20	25,32,43,51	0
3	EDO	B	515	4/4	0.73	0.22	35,42,49,53	0
6	2PO	C	501	4/4	0.73	0.18	26,28,72,86	0
4	1PE	B	507	16/16	0.75	0.27	21,45,57,59	0
3	EDO	A	506	4/4	0.77	0.19	27,39,47,51	0
5	PEG	A	513	7/7	0.79	0.18	29,36,46,53	0
4	1PE	A	507	16/16	0.80	0.21	13,38,53,53	0
3	EDO	C	506	4/4	0.80	0.15	31,38,43,50	0
5	PEG	C	515	7/7	0.81	0.21	30,49,63,63	0
5	PEG	B	512	7/7	0.81	0.19	29,37,61,61	0
3	EDO	B	516	4/4	0.82	0.17	30,37,44,50	0
2	GOL	C	503	6/6	0.82	0.15	26,34,44,49	0
3	EDO	C	516	4/4	0.82	0.17	30,36,40,49	0
3	EDO	A	504	4/4	0.83	0.17	25,35,51,51	0
3	EDO	C	519	4/4	0.84	0.15	23,37,44,48	0
3	EDO	C	504	4/4	0.85	0.14	28,34,44,44	0
3	EDO	A	514	4/4	0.85	0.19	24,41,56,56	0
2	GOL	A	511	6/6	0.86	0.17	28,40,48,57	0
2	GOL	C	514	6/6	0.86	0.17	18,42,55,66	0
3	EDO	B	514	4/4	0.86	0.16	24,40,46,56	0
2	GOL	A	512	6/6	0.87	0.18	16,47,57,66	0
2	GOL	C	513	6/6	0.87	0.13	27,44,55,55	0
2	GOL	C	512	6/6	0.88	0.14	23,37,45,54	0
3	EDO	C	517	4/4	0.88	0.14	34,41,48,48	0
2	GOL	B	509	6/6	0.88	0.12	25,36,48,52	0
2	GOL	B	511	6/6	0.89	0.12	20,33,42,47	0
2	GOL	B	510	6/6	0.90	0.13	21,26,32,35	0
3	EDO	A	505	4/4	0.90	0.12	17,32,36,39	0
6	2PO	B	502	4/4	0.90	0.20	26,26,66,80	0
5	PEG	B	508	7/7	0.90	0.16	19,34,59,59	0
3	EDO	B	513	4/4	0.91	0.15	14,28,49,49	0
2	GOL	A	509	6/6	0.91	0.14	15,31,42,45	0
3	EDO	B	501	4/4	0.91	0.15	16,36,51,54	0
6	2PO	A	515	4/4	0.91	0.12	19,43,62,75	0
2	GOL	C	511	6/6	0.91	0.13	17,35,49,50	0
5	PEG	B	505	7/7	0.91	0.14	13,35,58,58	0
3	EDO	A	508	4/4	0.92	0.11	16,26,38,38	0
3	EDO	C	507	4/4	0.92	0.14	16,43,51,61	0
3	EDO	C	505	4/4	0.92	0.15	26,35,48,58	0
3	EDO	C	518	4/4	0.93	0.14	24,39,50,50	0
2	GOL	C	509	6/6	0.93	0.10	20,28,40,48	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	C	508	4/4	0.93	0.12	20,28,47,47	0
3	EDO	A	503	4/4	0.94	0.10	17,22,41,49	0
2	GOL	C	510	6/6	0.95	0.10	15,28,34,34	0
2	GOL	A	502	6/6	0.95	0.09	18,23,30,31	0
3	EDO	B	506	4/4	0.96	0.11	17,29,35,35	0
2	GOL	A	510	6/6	0.96	0.09	15,22,42,51	0
2	GOL	C	502	6/6	0.97	0.06	12,17,20,23	0
2	GOL	B	504	6/6	0.97	0.06	12,16,21,21	0
2	GOL	A	501	6/6	0.98	0.05	11,14,18,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.