



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 03:48 PM GMT

PDB ID : 4DHL  
Title : Crystal structure of red kidney bean purple acid phosphatase in complex with Maybridge fragment MO07123  
Authors : Feder, D.; Clayton, D.J.; Hussein, W.M.; Schenk, G.; McGeary, R.; Guddat, L.W.  
Deposited on : 2012-01-29  
Resolution : 2.30 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](mailto: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

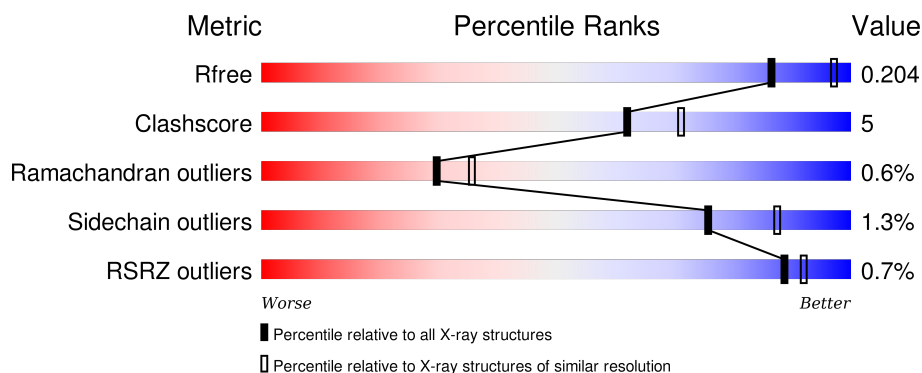
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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  $\geq 3$ , 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	426	<div> <div>87%</div> <div>12%</div> </div>
1	B	426	<div> <div>%</div> <div>85%</div> <div>14%</div> <div>.</div> </div>
1	C	426	<div> <div>%</div> <div>88%</div> <div>11%</div> <div>.</div> </div>
1	D	426	<div> <div>%</div> <div>86%</div> <div>12%</div> <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
4	OK7	A	503	-	-	-	X
4	OK7	B	503	-	-	-	X
4	OK7	C	503	-	-	-	X
4	OK7	D	503	-	-	-	X
6	SO4	A	506	-	-	-	X
6	SO4	B	506	-	-	-	X
6	SO4	D	505	-	-	-	X
7	EDO	D	506	-	-	-	X
7	EDO	D	507	-	-	-	X
7	EDO	D	508	-	-	-	X
7	EDO	D	510	-	-	-	X
7	EDO	D	511	-	-	-	X
8	NAG	A	508	-	-	-	X
8	NAG	B	507	-	-	-	X
8	NAG	B	508	-	-	-	X
8	NAG	C	508	-	-	-	X
8	NAG	D	517	-	-	-	X
9	NAG	A	509	-	-	-	X

## 2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 16091 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 Purple acid phosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	424	Total	C	N	O	S	0	1	0
			3502	2248	610	634	10			
1	B	425	Total	C	N	O	S	0	3	0
			3524	2260	612	641	11			
1	D	423	Total	C	N	O	S	7	1	0
			3495	2245	607	633	10			
1	C	424	Total	C	N	O	S	0	1	0
			3499	2248	606	634	11			

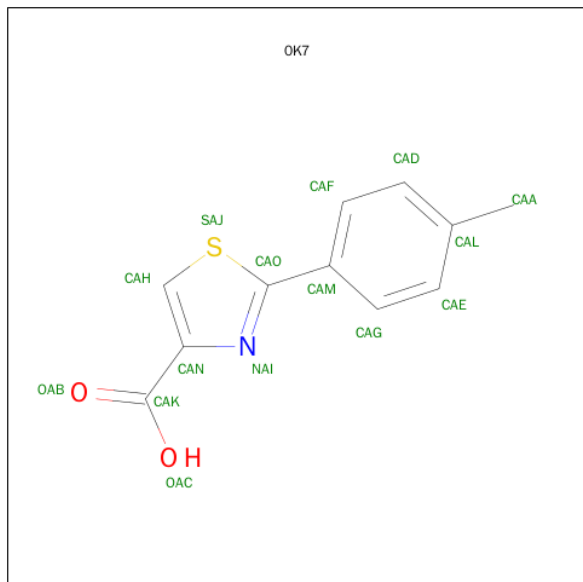
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Fe	0	0
			1	1		
3	A	1	Total	Fe	0	0
			1	1		
3	D	1	Total	Fe	0	0
			1	1		
3	C	1	Total	Fe	0	0
			1	1		

- Molecule 4 is 2-(4-METHYLPHENYL)-1,3-THIAZOLE-4-CARBOXYLIC ACID (three-letter code: OK7) (formula:  $C_{11}H_9NO_2S$ ).



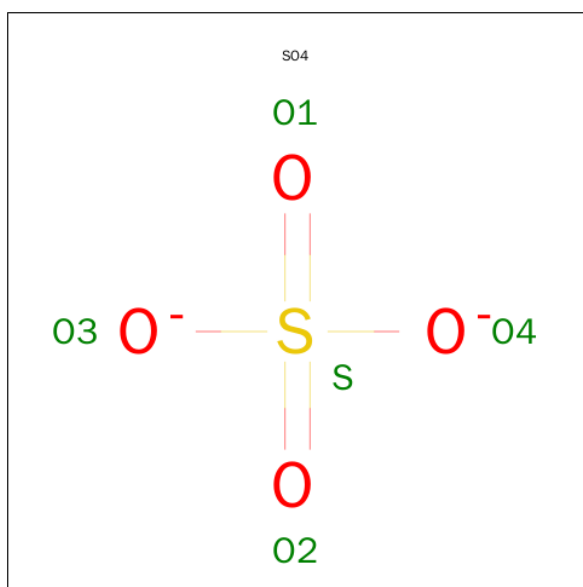
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			15	11	1	2	1		
4	B	1	Total	C	N	O	S	0	0
			15	11	1	2	1		
4	D	1	Total	C	N	O	S	0	0
			8	4	1	2	1		
4	C	1	Total	C	N	O	S	0	0
			15	11	1	2	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



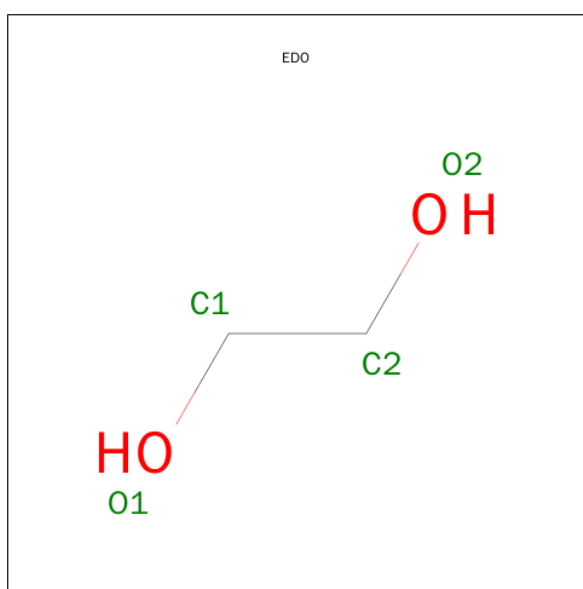
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

Continued from previous page...

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

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



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

- Molecule 8 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	N	O	0	0
			14	8	1	5		
8	B	1	Total	C	N	O	0	0
			14	8	1	5		
8	B	1	Total	C	N	O	0	0
			14	8	1	5		
8	D	1	Total	C	N	O	0	0
			14	8	1	5		
8	D	1	Total	C	N	O	0	0
			14	8	1	5		
8	C	1	Total	C	N	O	0	0
			14	8	1	5		
8	C	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 9 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	A	3	Total	C	N	O	0	0
			38	22	2	14		
9	A	3	Total	C	N	O	0	0
			38	22	2	14		
9	B	3	Total	C	N	O	0	0
			38	22	2	14		
9	D	3	Total	C	N	O	0	0
			38	22	2	14		
9	C	3	Total	C	N	O	0	0
			38	22	2	14		



- Molecule 10 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	A	3	Total	C	N	O	0	0
			38	22	2	14		

- Molecule 11 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	B	2	Total	C	N	O	0	0
			24	14	1	9		
11	D	2	Total	C	N	O	0	0
			24	14	1	9		
11	C	2	Total	C	N	O	0	0
			24	14	1	9		

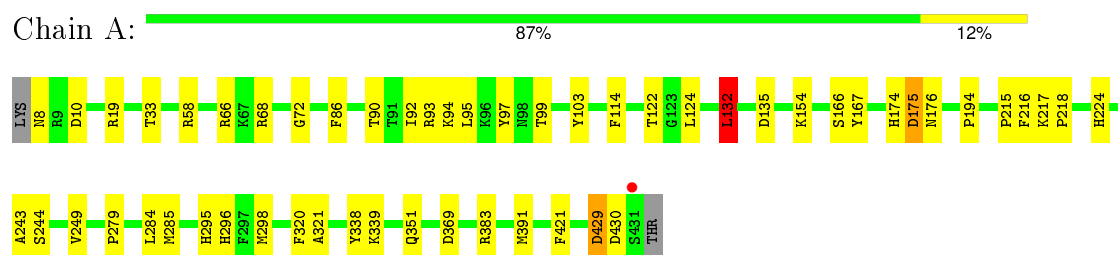
- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	403	Total	O	0	0
			403	403		
12	B	378	Total	O	0	0
			378	378		
12	D	380	Total	O	0	0
			380	380		
12	C	384	Total	O	0	0
			384	384		

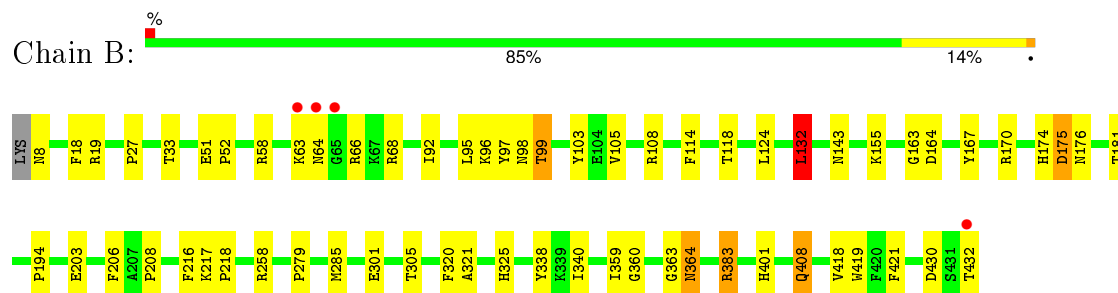
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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.

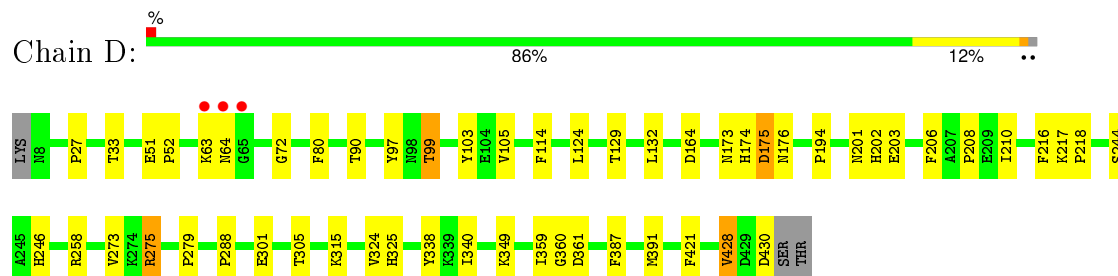
#### • Molecule 1: Purple acid phosphatase



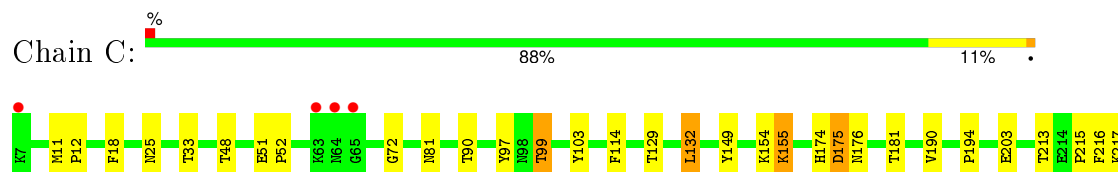
#### • Molecule 1: Purple acid phosphatase



#### • Molecule 1: Purple acid phosphatase



#### • Molecule 1: Purple acid phosphatase



P218	R258	M285	E301	T305	F320	A321	H325	Y338	K339	I340	M371	M372	R383	M391	S403	F421	Y426	P427	D430	SER	THR
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	126.10Å 126.10Å 297.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.82 – 2.30 19.82 – 2.30	Depositor EDS
% Data completeness (in resolution range)	94.1 (19.82-2.30) 94.1 (19.82-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 2.30Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, $R_{free}$	0.158 , 0.206 0.157 , 0.204	Depositor DCC
$R_{free}$ test set	5774 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.5	Xtriage
Anisotropy	0.009	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 56.3	EDS
Estimated twinning fraction	0.038 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 115055 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	16091	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% 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, 0.375 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: GOL, ZN, NAG, EDO, FUC, FE, 0K7, SO4

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	0.38	0/3624	0.54	1/4927 (0.0%)
1	B	0.37	0/3643	0.52	2/4953 (0.0%)
1	C	0.38	0/3618	0.52	0/4919
1	D	0.38	0/3614	0.53	1/4915 (0.0%)
All	All	0.38	0/14499	0.53	4/19714 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	132	LEU	CA-CB-CG	-6.12	101.22	115.30
1	B	132	LEU	CA-CB-CG	-5.68	102.24	115.30
1	B	383	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	D	275	ARG	NE-CZ-NH2	-5.26	117.67	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	3502	0	3321	38	0
1	B	3524	0	3334	42	0
1	C	3499	0	3312	35	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	3495	0	3310	36	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	15	0	8	3	0
4	B	15	0	8	5	0
4	C	15	0	8	2	0
4	D	8	0	1	1	0
5	A	6	0	8	0	0
5	C	6	0	8	0	0
5	D	6	0	8	1	0
6	A	10	0	0	1	0
6	B	5	0	0	0	0
6	C	5	0	0	0	0
6	D	5	0	0	0	0
7	A	4	0	6	0	0
7	D	20	0	30	4	0
8	A	14	0	13	0	0
8	B	28	0	26	1	0
8	C	28	0	26	0	0
8	D	28	0	26	0	0
9	A	76	0	68	0	0
9	B	38	0	34	1	0
9	C	38	0	34	0	0
9	D	38	0	34	0	0
10	A	38	0	34	2	0
11	B	24	0	22	0	0
11	C	24	0	22	0	0
11	D	24	0	22	0	0
12	A	403	0	0	3	0
12	B	378	0	0	3	0
12	C	384	0	0	2	0
12	D	380	0	0	4	0
All	All	16091	0	13723	153	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 153 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:D:173:ASN:H	7:D:510:EDO:H22	1.43	0.83
1:D:124:LEU:HD12	1:D:279:PRO:HG3	1.60	0.81
1:B:408[A]:GLN:HE21	1:B:408[A]:GLN:H	1.27	0.81
1:B:325:HIS:HE1	4:B:503:OK7:H8	1.50	0.77
1:A:217:LYS:HB3	1:A:218:PRO:HD3	1.67	0.75

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	423/426 (99%)	402 (95%)	19 (4%)	2 (0%)	34	41
1	B	426/426 (100%)	402 (94%)	19 (4%)	5 (1%)	16	16
1	C	423/426 (99%)	395 (93%)	26 (6%)	2 (0%)	34	41
1	D	422/426 (99%)	401 (95%)	20 (5%)	1 (0%)	52	64
All	All	1694/1704 (99%)	1600 (94%)	84 (5%)	10 (1%)	30	36

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	175	ASP
1	B	64	ASN
1	B	175	ASP
1	D	175	ASP
1	C	175	ASP

### 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	374/375 (100%)	370 (99%)	4 (1%)	80	90
1	B	377/375 (100%)	370 (98%)	7 (2%)	65	81
1	C	373/375 (100%)	369 (99%)	4 (1%)	80	90
1	D	373/375 (100%)	366 (98%)	7 (2%)	65	81
All	All	1497/1500 (100%)	1475 (98%)	22 (2%)	76	85

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	408[A]	GLN
1	D	99	THR
1	C	176	ASN
1	B	408[B]	GLN
1	D	64	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	295	HIS
1	B	371	ASN
1	C	295	HIS
1	A	294	ASN
1	D	64	ASN

### 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 ⓘ

24 carbohydrates 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$
9	NAG	A	509	9,1	14,14,15	0.46	0	15,19,21	1.25	2 (13%)
9	NAG	A	510	9	14,14,15	0.47	0	15,19,21	0.86	1 (6%)
9	FUC	A	511	9	10,10,11	0.58	0	14,14,16	0.79	1 (7%)
10	NAG	A	512	1,10	14,14,15	0.60	0	15,19,21	0.83	0
10	FUC	A	513	10	10,10,11	0.56	0	14,14,16	1.17	1 (7%)
10	NAG	A	514	10	14,14,15	0.50	0	15,19,21	0.77	0
9	NAG	A	515	9,1	14,14,15	0.56	0	15,19,21	1.14	2 (13%)
9	NAG	A	516	9	14,14,15	0.55	0	15,19,21	0.80	0
9	FUC	A	517	9	10,10,11	0.67	0	14,14,16	1.20	1 (7%)
11	NAG	B	504	11,1	14,14,15	0.54	0	15,19,21	1.04	1 (6%)
11	FUC	B	505	11	10,10,11	0.52	0	14,14,16	0.63	0
9	NAG	B	509	9,1	14,14,15	0.54	0	15,19,21	0.96	1 (6%)
9	NAG	B	510	9	14,14,15	0.50	0	15,19,21	0.67	0
9	FUC	B	511	9	10,10,11	0.70	0	14,14,16	0.98	1 (7%)
11	NAG	C	506	11,1	14,14,15	0.48	0	15,19,21	1.22	1 (6%)
11	FUC	C	507	11	10,10,11	0.59	0	14,14,16	0.75	0
9	NAG	C	509	9,1	14,14,15	0.58	0	15,19,21	1.31	2 (13%)
9	NAG	C	510	9	14,14,15	0.55	0	15,19,21	1.02	1 (6%)
9	FUC	C	511	9	10,10,11	0.61	0	14,14,16	0.70	0
9	NAG	D	512	9,1	14,14,15	0.62	0	15,19,21	0.70	0
9	NAG	D	513	9	14,14,15	0.59	0	15,19,21	1.07	1 (6%)
9	FUC	D	514	9	10,10,11	0.59	0	14,14,16	0.70	0
11	NAG	D	515	11,1	14,14,15	0.49	0	15,19,21	0.89	0
11	FUC	D	516	11	10,10,11	0.62	0	14,14,16	0.95	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
9	NAG	A	509	9,1	-	0/6/23/26	0/1/1/1
9	NAG	A	510	9	-	0/6/23/26	0/1/1/1
9	FUC	A	511	9	-	0/0/17/20	0/1/1/1
10	NAG	A	512	1,10	-	0/6/23/26	0/1/1/1
10	FUC	A	513	10	-	0/0/17/20	0/1/1/1
10	NAG	A	514	10	-	0/6/23/26	0/1/1/1
9	NAG	A	515	9,1	-	0/6/23/26	0/1/1/1
9	NAG	A	516	9	-	0/6/23/26	0/1/1/1
9	FUC	A	517	9	-	0/0/17/20	0/1/1/1
11	NAG	B	504	11,1	-	0/6/23/26	0/1/1/1
11	FUC	B	505	11	-	0/0/17/20	0/1/1/1
9	NAG	B	509	9,1	-	0/6/23/26	0/1/1/1
9	NAG	B	510	9	-	0/6/23/26	0/1/1/1
9	FUC	B	511	9	-	0/0/17/20	0/1/1/1
11	NAG	C	506	11,1	-	0/6/23/26	0/1/1/1
11	FUC	C	507	11	-	0/0/17/20	0/1/1/1
9	NAG	C	509	9,1	-	0/6/23/26	0/1/1/1
9	NAG	C	510	9	-	0/6/23/26	0/1/1/1
9	FUC	C	511	9	-	0/0/17/20	0/1/1/1
9	NAG	D	512	9,1	-	0/6/23/26	0/1/1/1
9	NAG	D	513	9	-	0/6/23/26	0/1/1/1
9	FUC	D	514	9	-	0/0/17/20	0/1/1/1
11	NAG	D	515	11,1	-	0/6/23/26	0/1/1/1
11	FUC	D	516	11	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	509	NAG	C2-N2-C7	-3.37	118.71	123.04
9	D	513	NAG	C2-N2-C7	-3.27	118.84	123.04
9	C	509	NAG	C2-N2-C7	-3.19	118.94	123.04
9	C	510	NAG	C2-N2-C7	-3.07	119.10	123.04
10	A	513	FUC	C2-C3-C4	-2.91	106.09	111.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	A	513	FUC	2	0
10	A	514	NAG	2	0
9	B	509	NAG	1	0
9	B	511	FUC	1	0

## 5.6 Ligand geometry [i](#)

Of 33 ligands modelled in this entry, 8 are monoatomic - leaving 25 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$
4	0K7	A	503	3,2	11,16,16	1.30	2 (18%)	9,22,22	2.05	1 (11%)
5	GOL	A	504	-	5,5,5	0.33	0	5,5,5	0.21	0
6	SO4	A	505	-	4,4,4	0.30	0	6,6,6	0.20	0
6	SO4	A	506	-	4,4,4	0.10	0	6,6,6	0.16	0
7	EDO	A	507	-	3,3,3	0.53	0	2,2,2	0.49	0
8	NAG	A	508	1	14,14,15	0.49	0	15,19,21	0.85	1 (6%)
4	0K7	B	503	3,2	11,16,16	1.00	1 (9%)	9,22,22	2.06	1 (11%)
6	SO4	B	506	-	4,4,4	0.08	0	6,6,6	0.32	0
8	NAG	B	507	1	14,14,15	0.45	0	15,19,21	1.53	2 (13%)
8	NAG	B	508	1	14,14,15	0.52	0	15,19,21	0.68	0
4	0K7	C	503	3,2	11,16,16	1.04	2 (18%)	9,22,22	2.09	1 (11%)
5	GOL	C	504	-	5,5,5	0.34	0	5,5,5	0.27	0
6	SO4	C	505	-	4,4,4	0.12	0	6,6,6	0.08	0
8	NAG	C	508	1	14,14,15	0.51	0	15,19,21	1.59	3 (20%)
8	NAG	C	512	1	14,14,15	0.63	0	15,19,21	1.09	1 (6%)
4	0K7	D	503	3	2,8,16	0.55	0	2,10,22	4.62	2 (100%)
5	GOL	D	504	-	5,5,5	0.24	0	5,5,5	0.43	0
6	SO4	D	505	-	4,4,4	0.19	0	6,6,6	0.17	0
7	EDO	D	506	-	3,3,3	0.55	0	2,2,2	0.48	0
7	EDO	D	507	-	3,3,3	0.51	0	2,2,2	0.41	0
7	EDO	D	508	-	3,3,3	0.60	0	2,2,2	0.17	0
8	NAG	D	509	1	14,14,15	0.48	0	15,19,21	0.68	0
7	EDO	D	510	-	3,3,3	0.55	0	2,2,2	0.32	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	EDO	D	511	-	3,3,3	0.51	0	2,2,2	0.35	0
8	NAG	D	517	1	14,14,15	0.41	0	15,19,21	1.38	2 (13%)

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
4	0K7	A	503	3,2	-	0/4/8/8	0/2/2/2
5	GOL	A	504	-	-	0/4/4/4	0/0/0/0
6	SO4	A	505	-	-	0/0/0/0	0/0/0/0
6	SO4	A	506	-	-	0/0/0/0	0/0/0/0
7	EDO	A	507	-	-	0/1/1/1	0/0/0/0
8	NAG	A	508	1	-	0/6/23/26	0/1/1/1
4	0K7	B	503	3,2	-	0/4/8/8	0/2/2/2
6	SO4	B	506	-	-	0/0/0/0	0/0/0/0
8	NAG	B	507	1	-	0/6/23/26	0/1/1/1
8	NAG	B	508	1	-	0/6/23/26	0/1/1/1
4	0K7	C	503	3,2	-	0/4/8/8	0/2/2/2
5	GOL	C	504	-	-	0/4/4/4	0/0/0/0
6	SO4	C	505	-	-	0/0/0/0	0/0/0/0
8	NAG	C	508	1	-	0/6/23/26	0/1/1/1
8	NAG	C	512	1	-	0/6/23/26	0/1/1/1
4	0K7	D	503	3	-	0/0/4/8	0/1/1/2
5	GOL	D	504	-	-	0/4/4/4	0/0/0/0
6	SO4	D	505	-	-	0/0/0/0	0/0/0/0
7	EDO	D	506	-	-	0/1/1/1	0/0/0/0
7	EDO	D	507	-	-	0/1/1/1	0/0/0/0
7	EDO	D	508	-	-	0/1/1/1	0/0/0/0
8	NAG	D	509	1	-	0/6/23/26	0/1/1/1
7	EDO	D	510	-	-	0/1/1/1	0/0/0/0
7	EDO	D	511	-	-	0/1/1/1	0/0/0/0
8	NAG	D	517	1	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	503	0K7	CAO-SAJ	-2.70	1.70	1.73
4	C	503	0K7	CAO-SAJ	-2.33	1.70	1.73
4	C	503	0K7	CAO-NAI	2.41	1.34	1.31
4	B	503	0K7	CAO-NAI	2.50	1.35	1.31

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	503	0K7	CAO-NAI	3.22	1.36	1.31

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	503	0K7	CAN-CAH-SAJ	-6.01	104.48	111.83
4	A	503	0K7	CAN-CAH-SAJ	-5.79	104.68	111.79
4	B	503	0K7	CAN-CAH-SAJ	-5.74	104.74	111.79
4	C	503	0K7	CAN-CAH-SAJ	-5.37	105.20	111.79
8	C	512	NAG	C2-N2-C7	-3.30	118.80	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	503	0K7	3	0
6	A	505	SO4	1	0
4	B	503	0K7	5	0
8	B	508	NAG	1	0
4	C	503	0K7	2	0
4	D	503	0K7	1	0
5	D	504	GOL	1	0
7	D	507	EDO	1	0
7	D	508	EDO	1	0
7	D	510	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](#)

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	424/426 (99%)	-0.88	1 (0%) 95 97	14, 21, 35, 52	0
1	B	425/426 (99%)	-0.79	4 (0%) 85 89	16, 24, 38, 69	0
1	C	424/426 (99%)	-0.78	4 (0%) 85 89	15, 23, 38, 75	1 (0%)
1	D	423/426 (99%)	-0.83	3 (0%) 89 92	16, 22, 36, 72	0
All	All	1696/1704 (99%)	-0.82	12 (0%) 89 92	14, 23, 37, 75	1 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	64	ASN	5.7
1	A	431	SER	3.9
1	B	64	ASN	3.5
1	B	432	THR	3.3
1	D	64	ASN	3.3

### 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](#)

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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
9	NAG	A	509	14/15	0.95	0.12	2.80	35,41,50,52	0
9	NAG	D	512	14/15	0.96	0.10	1.01	23,31,41,50	0
9	NAG	C	509	14/15	0.96	0.10	-0.29	24,34,40,47	0
9	NAG	B	509	14/15	0.96	0.10	-	19,24,33,33	0
9	FUC	C	511	10/11	0.90	0.19	-	46,53,57,59	0
9	FUC	A	511	10/11	0.86	0.40	-	55,63,69,69	0
9	NAG	A	510	14/15	0.90	0.27	-	45,55,59,60	0
10	FUC	A	513	10/11	0.93	0.23	-	40,50,56,63	0
9	NAG	D	513	14/15	0.93	0.22	-	29,42,55,69	0
9	NAG	A	515	14/15	0.98	0.09	-	20,25,30,34	0
10	NAG	A	512	14/15	0.96	0.14	-	29,35,46,49	0
11	NAG	D	515	14/15	0.84	0.23	-	42,68,74,84	0
11	FUC	B	505	10/11	0.87	0.29	-	48,61,65,66	0
9	NAG	C	510	14/15	0.92	0.27	-	32,49,64,74	0
11	NAG	C	506	14/15	0.85	0.18	-	37,46,62,64	0
9	FUC	B	511	10/11	0.96	0.12	-	28,29,35,41	0
10	NAG	A	514	14/15	0.86	0.28	-	45,55,61,66	0
9	NAG	B	510	14/15	0.95	0.15	-	26,31,38,42	0
11	FUC	D	516	10/11	0.83	0.38	-	68,79,84,85	0
11	NAG	B	504	14/15	0.91	0.22	-	37,48,55,57	0
9	NAG	A	516	14/15	0.94	0.17	-	23,30,35,36	0
9	FUC	A	517	10/11	0.97	0.17	-	29,36,39,42	0
11	FUC	C	507	10/11	0.90	0.28	-	62,70,76,77	0
9	FUC	D	514	10/11	0.91	0.18	-	46,52,58,61	0

## 6.4 Ligands ⓘ

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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	OK7	A	503	15/15	0.87	0.30	13.70	28,57,70,77	0
4	OK7	B	503	15/15	0.87	0.30	12.02	28,55,70,74	0
8	NAG	C	508	14/15	0.86	0.19	11.01	38,42,51,52	0
4	OK7	C	503	15/15	0.88	0.31	10.02	33,61,78,78	0
7	EDO	D	508	4/4	0.88	0.14	8.35	30,33,34,38	0
8	NAG	D	517	14/15	0.95	0.18	7.46	39,44,48,54	0
8	NAG	B	507	14/15	0.93	0.17	5.40	33,41,45,47	0
4	OK7	D	503	8/15	0.91	0.16	5.16	26,40,50,83	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	SO4	A	506	5/5	0.95	0.24	3.88	38,40,55,62	0
7	EDO	D	507	4/4	0.94	0.16	3.58	33,33,35,35	0
6	SO4	B	506	5/5	0.94	0.24	3.20	40,47,60,79	0
7	EDO	D	506	4/4	0.95	0.10	3.20	31,32,33,37	0
7	EDO	D	511	4/4	0.95	0.20	2.83	28,33,37,39	0
6	SO4	D	505	5/5	0.97	0.25	2.77	30,51,55,57	0
7	EDO	D	510	4/4	0.87	0.17	2.45	36,40,44,52	0
8	NAG	A	508	14/15	0.96	0.12	2.36	16,29,33,42	0
8	NAG	B	508	14/15	0.96	0.13	2.15	23,34,40,47	0
8	NAG	D	509	14/15	0.96	0.12	1.42	23,29,37,39	0
7	EDO	A	507	4/4	0.94	0.10	1.16	30,33,36,37	0
8	NAG	C	512	14/15	0.97	0.10	1.00	23,29,37,40	0
6	SO4	A	505	5/5	0.99	0.07	-1.11	29,30,36,43	0
3	FE	B	502	1/1	0.99	0.06	-1.26	21,21,21,21	1
5	GOL	A	504	6/6	0.99	0.05	-1.50	17,20,23,23	0
5	GOL	C	504	6/6	0.98	0.05	-1.93	21,23,25,26	0
2	ZN	A	501	1/1	1.00	0.02	-3.08	28,28,28,28	0
3	FE	C	502	1/1	0.99	0.02	-3.77	28,28,28,28	1
3	FE	A	502	1/1	0.99	0.02	-3.83	27,27,27,27	0
2	ZN	C	501	1/1	1.00	0.02	-3.85	30,30,30,30	0
3	FE	D	502	1/1	1.00	0.04	-3.95	29,29,29,29	1
2	ZN	D	501	1/1	1.00	0.03	-4.35	23,23,23,23	1
2	ZN	B	501	1/1	1.00	0.01	-4.81	27,27,27,27	0
6	SO4	C	505	5/5	0.90	0.30	-	58,63,88,95	0
5	GOL	D	504	6/6	0.89	0.13	-	29,39,43,50	0

## 6.5 Other polymers

There are no such residues in this entry.