



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 03:34 AM BST

PDB ID : 2YHW  
Title : High-resolution crystal structures of N-Acetylmannosamine kinase: Insights about substrate specificity, activity and inhibitor modelling.  
Authors : Martinez, J.; Nguyen, L.D.; Tauberger, E.; Hinderlich, S.; Zimmer, R.; Tauberger, E.; Reutter, W.; Saenger, W.; Fan, H.; Moniot, S.  
Deposited on : 2011-05-08  
Resolution : 1.64 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 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) : 1.13  
EDS : 2.13.1  
buster-report : 1.1.7 (2018)  
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)  
Validation Pipeline (wwPDB-VP) : 2.13.1

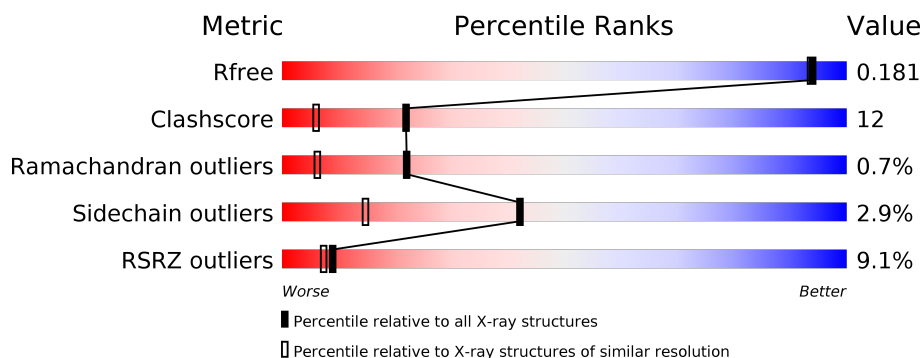
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.64 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	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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 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	343	<div> <div>8%</div> <div>78%</div> <div>10%</div> <div>10%</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
5	EDO	A	1721	-	-	X	-
5	EDO	A	1726[A]	-	-	X	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	EDO	A	1726[B]	-	-	X	-

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 2802 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 BIFUNCTIONAL UDP-N-ACETYLGLUCOSAMINE 2-EPI MERASE/N-ACETYLMANNOSAMINE KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	309	Total	C	N	O	S	0	24	0
			2471	1545	439	471	16			

There are 28 discrepancies between the modelled and reference sequences:

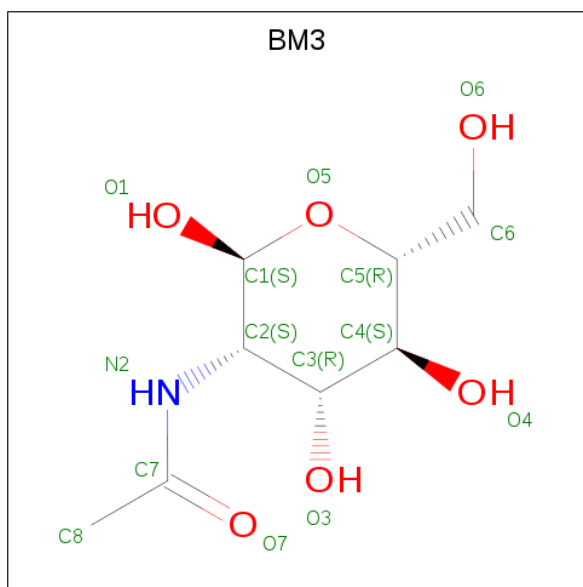
Chain	Residue	Modelled	Actual	Comment	Reference
A	378	MET	-	expression tag	UNP Q9Y223
A	379	GLY	-	expression tag	UNP Q9Y223
A	380	SER	-	expression tag	UNP Q9Y223
A	381	SER	-	expression tag	UNP Q9Y223
A	382	HIS	-	expression tag	UNP Q9Y223
A	383	HIS	-	expression tag	UNP Q9Y223
A	384	HIS	-	expression tag	UNP Q9Y223
A	385	HIS	-	expression tag	UNP Q9Y223
A	386	HIS	-	expression tag	UNP Q9Y223
A	387	HIS	-	expression tag	UNP Q9Y223
A	388	SER	-	expression tag	UNP Q9Y223
A	389	SER	-	expression tag	UNP Q9Y223
A	390	GLY	-	expression tag	UNP Q9Y223
A	391	LEU	-	expression tag	UNP Q9Y223
A	392	VAL	-	expression tag	UNP Q9Y223
A	393	PRO	-	expression tag	UNP Q9Y223
A	394	ARG	-	expression tag	UNP Q9Y223
A	395	GLY	-	expression tag	UNP Q9Y223
A	396	SER	-	expression tag	UNP Q9Y223
A	397	HIS	-	expression tag	UNP Q9Y223
A	398	MET	-	expression tag	UNP Q9Y223
A	399	GLU	-	expression tag	UNP Q9Y223
A	400	ASN	-	expression tag	UNP Q9Y223
A	401	LEU	-	expression tag	UNP Q9Y223
A	402	TYR	-	expression tag	UNP Q9Y223
A	403	PHE	-	expression tag	UNP Q9Y223

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	404	GLN	-	expression tag	UNP Q9Y223
A	405	GLY	-	expression tag	UNP Q9Y223

- Molecule 2 is 2-acetamido-2-deoxy- $\alpha$ -D-mannopyranose (three-letter code: BM3) (formula:  $C_8H_{15}NO_6$ ).

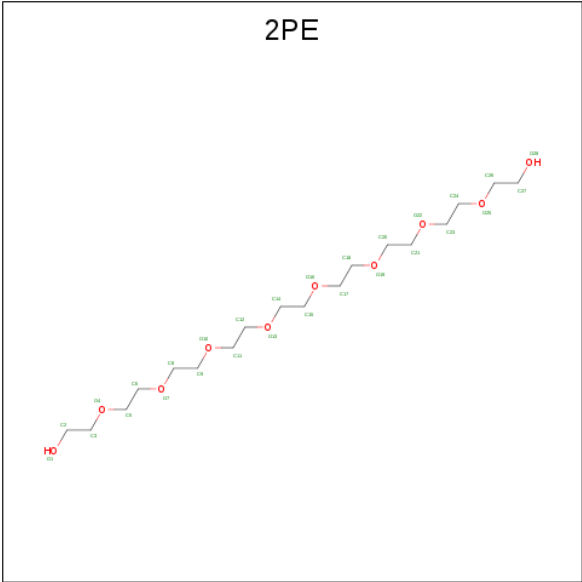


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula:  $Zn$ ).

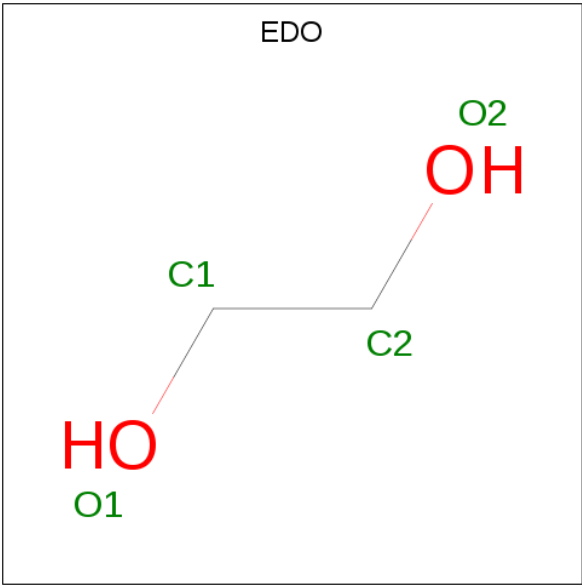
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is NONAETHYLENE GLYCOL (three-letter code: 2PE) (formula:  $C_{18}H_{38}O_{10}$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			19	12	7		

- Molecule 5 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
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

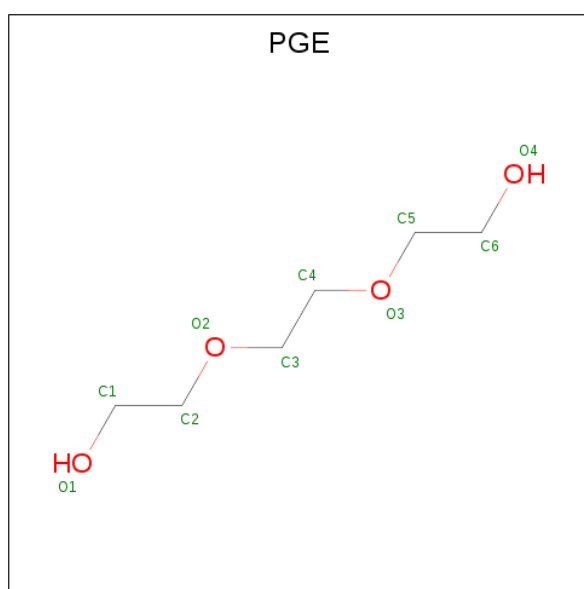
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	1
			8	4	4		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Ca	0	1
			2	2		

- Molecule 7 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	4	3		
7	A	1	Total	C	O	0	0
			7	4	3		

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



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

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Cl	0	0
			1	1		

- Molecule 10 is water.

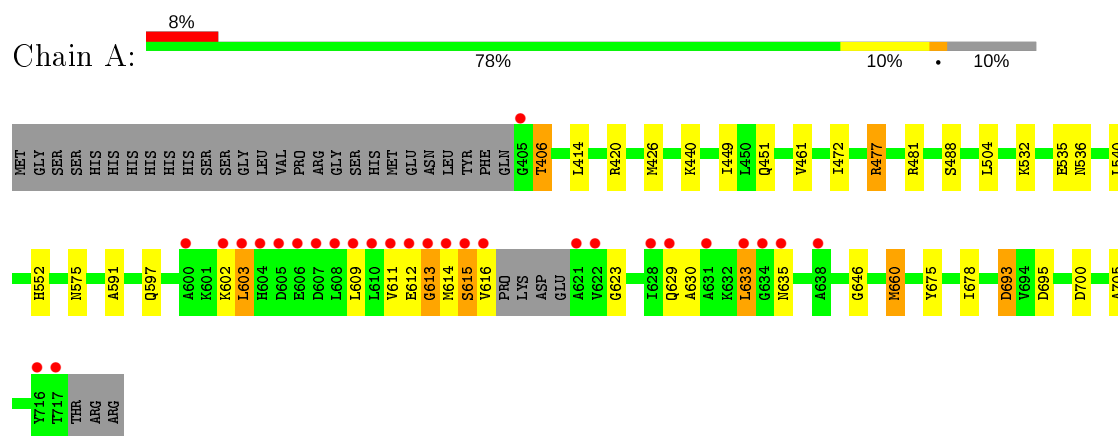
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	255	Total	O	0	0
			255	255		



### 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: BIFUNCTIONAL UDP-N-ACETYLGLUCOSAMINE 2-EPIMERASE/N-ACETYLMANNOSAMINE KINASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.73 Å 90.73 Å 100.78 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.37 – 1.64 45.37 – 1.64	Depositor EDS
% Data completeness (in resolution range)	100.0 (45.37-1.64) 100.0 (45.37-1.64)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.77 (at 1.64 Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.148 , 0.170 0.164 , 0.181	Depositor DCC
$R_{free}$ test set	2605 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.1	Xtriage
Anisotropy	0.267	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2802	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 5.98% 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: ZN, PGE, CL, SNN, CA, EDO, BM3, 2PE, ACT

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	1.15	5/2493 (0.2%)	1.02	7/3377 (0.2%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	675	TYR	CE2-CZ	-5.80	1.31	1.38
1	A	705	ALA	CA-CB	5.63	1.64	1.52
1	A	591	ALA	CA-CB	5.30	1.63	1.52
1	A	660[A]	MET	CG-SD	-5.02	1.68	1.81
1	A	660[B]	MET	CG-SD	-5.02	1.68	1.81

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	700	ASP	CB-CG-OD1	5.71	123.44	118.30
1	A	477	ARG	NE-CZ-NH1	5.67	123.14	120.30
1	A	660[A]	MET	CG-SD-CE	5.49	108.98	100.20
1	A	660[B]	MET	CG-SD-CE	5.49	108.98	100.20
1	A	693[A]	ASP	CB-CG-OD1	5.48	123.23	118.30
1	A	693[B]	ASP	CB-CG-OD1	5.48	123.23	118.30
1	A	540	LEU	CB-CG-CD1	-5.21	102.15	111.00

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	2471	0	2476	53	0
2	A	15	0	15	1	0
3	A	1	0	0	0	0
4	A	19	0	25	0	0
5	A	20	0	29	26	0
6	A	2	0	0	0	0
7	A	14	0	18	1	0
8	A	4	0	3	0	0
9	A	1	0	0	0	0
10	A	255	0	0	23	0
All	All	2802	0	2566	61	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 12.

All (61) 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)
5:A:1726[B]:EDO:H12	10:A:2234:HOH:O	1.16	1.29
5:A:1726[A]:EDO:H12	10:A:2236:HOH:O	1.09	1.27
1:A:532[B]:LYS:HE2	10:A:2127:HOH:O	1.10	1.26
1:A:426[A]:MET:SD	10:A:2016:HOH:O	2.00	1.19
1:A:536[B]:ASN:OD1	10:A:2137:HOH:O	1.63	1.15
5:A:1726[A]:EDO:C1	10:A:2236:HOH:O	1.65	1.13
1:A:693[A]:ASP:O	5:A:1726[A]:EDO:H22	1.56	1.06
1:A:406[B]:THR:HG22	10:A:2003:HOH:O	1.68	0.93
1:A:426[B]:MET:SD	10:A:2018:HOH:O	2.31	0.89
1:A:693[B]:ASP:OD1	5:A:1726[B]:EDO:H21	1.76	0.86
1:A:695:ASP:OD1	5:A:1726[A]:EDO:H11	1.74	0.85
1:A:693[B]:ASP:OD1	5:A:1726[B]:EDO:C2	2.28	0.82
1:A:678[B]:ILE:CD1	10:A:2197:HOH:O	2.31	0.78
5:A:1726[B]:EDO:C1	10:A:2234:HOH:O	1.90	0.78
1:A:695:ASP:CG	5:A:1726[A]:EDO:H11	2.03	0.78
1:A:611:VAL:HG22	1:A:614:MET:CE	2.14	0.76
1:A:451[B]:GLN:CD	5:A:1721:EDO:H11	2.06	0.75
1:A:660[B]:MET:O	1:A:660[B]:MET:CG	2.36	0.72
1:A:660[B]:MET:HG3	1:A:660[B]:MET:O	1.91	0.70
1:A:451[B]:GLN:OE1	5:A:1721:EDO:H11	1.91	0.70
1:A:532[B]:LYS:CE	10:A:2127:HOH:O	1.91	0.68

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:614:MET:HE2	1:A:630:ALA:HA	1.79	0.65
1:A:536[C]:ASN:OD1	10:A:2135:HOH:O	2.14	0.65
1:A:472:ILE:HD11	1:A:504:LEU:HD22	1.82	0.61
1:A:693[A]:ASP:O	5:A:1726[A]:EDO:C2	2.41	0.61
1:A:611:VAL:O	1:A:613:GLY:N	2.35	0.59
1:A:695:ASP:OD1	5:A:1726[A]:EDO:O2	2.22	0.58
1:A:678[B]:ILE:HD12	10:A:2197:HOH:O	1.98	0.57
1:A:693[B]:ASP:OD1	5:A:1726[B]:EDO:O2	2.22	0.57
1:A:451[B]:GLN:CD	5:A:1721:EDO:C1	2.74	0.56
1:A:660[B]:MET:HG2	10:A:2136:HOH:O	2.07	0.55
1:A:695:ASP:OD1	5:A:1726[A]:EDO:C1	2.52	0.54
1:A:451[B]:GLN:CG	5:A:1721:EDO:H11	2.38	0.53
1:A:611:VAL:H	1:A:614:MET:HE3	1.72	0.53
1:A:629:GLN:O	1:A:633:LEU:HD13	2.10	0.52
1:A:611:VAL:HG22	1:A:614:MET:HE1	1.92	0.51
1:A:597[B]:GLN:HG2	1:A:623:GLY:HA2	1.93	0.51
1:A:440:LYS:HE2	7:A:1727:PGE:O1	2.10	0.51
1:A:646:GLY:HA3	1:A:678[A]:ILE:HG21	1.94	0.50
1:A:660[B]:MET:CG	10:A:2136:HOH:O	2.61	0.49
1:A:451[B]:GLN:HG2	5:A:1721:EDO:H11	1.95	0.49
1:A:695:ASP:OD1	5:A:1726[B]:EDO:O2	2.32	0.48
1:A:611:VAL:CG2	1:A:614:MET:CE	2.90	0.48
1:A:488:SER:HB3	2:A:1718:BM3:H8C1	1.95	0.47
1:A:532[B]:LYS:CD	10:A:2127:HOH:O	2.46	0.47
1:A:609:LEU:O	1:A:635:ASN:ND2	2.47	0.47
5:A:1726[A]:EDO:H12	10:A:2237:HOH:O	2.15	0.47
5:A:1726[A]:EDO:C2	10:A:2237:HOH:O	2.63	0.46
5:A:1726[A]:EDO:C1	10:A:2237:HOH:O	2.63	0.46
1:A:461:VAL:HG21	5:A:1722:EDO:H21	1.99	0.45
1:A:535[A]:GLU:HG2	10:A:2130:HOH:O	2.17	0.45
1:A:451[B]:GLN:HG2	5:A:1721:EDO:C1	2.49	0.43
1:A:536[C]:ASN:HD22	1:A:552:HIS:HA	1.83	0.43
1:A:615:SER:O	1:A:616:VAL:HB	2.19	0.42
1:A:481:ARG:NH2	10:A:2075:HOH:O	2.53	0.42
5:A:1726[A]:EDO:H11	10:A:2236:HOH:O	1.68	0.42
1:A:414:LEU:HD21	1:A:449:ILE:HD11	2.02	0.42
1:A:678[B]:ILE:HG13	10:A:2210:HOH:O	2.19	0.41
1:A:603:LEU:HD12	1:A:603:LEU:HA	1.93	0.41
1:A:695:ASP:OD2	5:A:1726[A]:EDO:H11	2.19	0.41
1:A:616:VAL:O	1:A:616:VAL:HG12	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	326/343 (95%)	316 (97%)	8 (2%)	2 (1%)	25 8

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	612	GLU
1	A	613	GLY

### 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	267/278 (96%)	258 (97%)	9 (3%)	37 11

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	406[A]	THR
1	A	406[B]	THR
1	A	420[A]	ARG
1	A	420[B]	ARG
1	A	477	ARG
1	A	602	LYS
1	A	603	LEU
1	A	615	SER
1	A	633	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	677	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
1	SNN	A	575	1	7,8,8	2.88	3 (42%)	7,11,11	3.44	5 (71%)

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
1	SNN	A	575	1	-	-	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	575	SNN	O5-C5	5.17	1.33	1.23
1	A	575	SNN	C4-C5	4.13	1.56	1.51
1	A	575	SNN	C2-N1	-3.18	1.33	1.37

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	575	SNN	O2-C2-C3	-5.91	121.91	126.18
1	A	575	SNN	C3-C2-N1	5.39	111.46	107.30
1	A	575	SNN	C3-C4-C5	2.34	106.21	103.50
1	A	575	SNN	O5-C5-C4	2.24	129.35	126.39
1	A	575	SNN	C4-C3-C2	-2.11	102.01	104.46

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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 4 are monoatomic - leaving 10 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	EDO	A	1721	-	3,3,3	0.32	0	2,2,2	0.34	0
8	ACT	A	1728	-	1,3,3	4.61	1 (100%)	0,3,3	0.00	-
7	PGE	A	1727	-	6,6,9	0.56	0	5,5,8	0.40	0
5	EDO	A	1726[B]	-	3,3,3	0.69	0	2,2,2	0.46	0
5	EDO	A	1726[A]	-	3,3,3	1.09	0	2,2,2	0.08	0
5	EDO	A	1723	-	3,3,3	0.50	0	2,2,2	0.25	0
4	2PE	A	1720	-	18,18,27	0.55	0	17,17,26	0.85	0
7	PGE	A	1725	-	6,6,9	0.46	0	5,5,8	0.33	0
2	BM3	A	1718	-	15,15,15	1.13	2 (13%)	21,21,21	0.90	0
5	EDO	A	1722	-	3,3,3	0.47	0	2,2,2	0.58	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	EDO	A	1721	-	-	0/1/1/1	-
7	PGE	A	1727	-	-	3/4/4/7	-
5	EDO	A	1726[B]	-	-	1/1/1/1	-
5	EDO	A	1726[A]	-	-	1/1/1/1	-
5	EDO	A	1723	-	-	1/1/1/1	-
4	2PE	A	1720	-	-	1/16/16/25	-
7	PGE	A	1725	-	-	0/4/4/7	-
2	BM3	A	1718	-	-	0/6/26/26	0/1/1/1
5	EDO	A	1722	-	-	1/1/1/1	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	A	1728	ACT	CH3-C	4.61	1.54	1.48
2	A	1718	BM3	C2-N2	2.64	1.50	1.45
2	A	1718	BM3	C8-C7	2.00	1.54	1.50

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	1727	PGE	O1-C1-C2-O2
5	A	1726[B]	EDO	O1-C1-C2-O2
5	A	1723	EDO	O1-C1-C2-O2
7	A	1727	PGE	C4-C3-O2-C2
5	A	1726[A]	EDO	O1-C1-C2-O2
7	A	1727	PGE	O2-C3-C4-O3
4	A	1720	2PE	C2-C3-O4-C5
5	A	1722	EDO	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 28 short contacts:

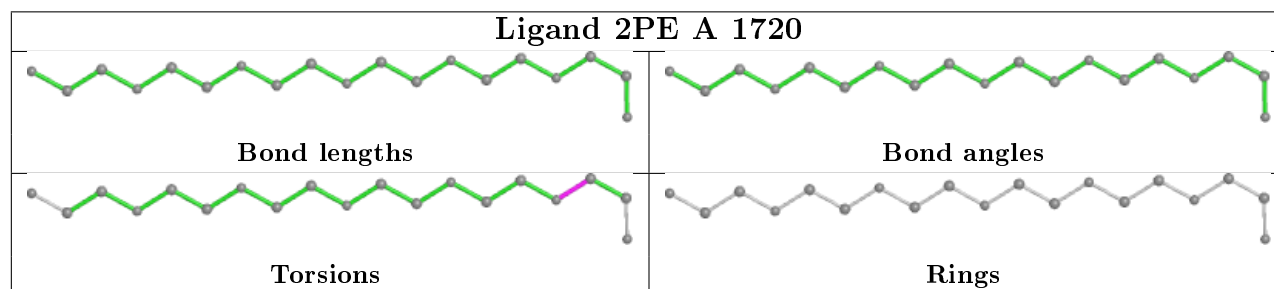
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1721	EDO	6	0
7	A	1727	PGE	1	0
5	A	1726[B]	EDO	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1726[A]	EDO	13	0
2	A	1718	BM3	1	0
5	A	1722	EDO	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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<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	308/343 (89%)	0.19	28 (9%) 9 7	12, 19, 66, 93	3 (0%)

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	611	VAL	7.9
1	A	615	SER	7.1
1	A	616	VAL	6.7
1	A	610	LEU	6.4
1	A	604	HIS	6.1
1	A	608	LEU	6.0
1	A	621	ALA	5.5
1	A	612	GLU	5.2
1	A	633	LEU	4.9
1	A	609	LEU	4.8
1	A	717	THR	4.8
1	A	607	ASP	4.7
1	A	605	ASP	4.5
1	A	631	ALA	4.2
1	A	634	GLY	4.0
1	A	603	LEU	3.4
1	A	622	VAL	3.4
1	A	716	TYR	3.0
1	A	638	ALA	3.0
1	A	635	ASN	2.8
1	A	614	MET	2.7
1	A	405[A]	GLY	2.6
1	A	606	GLU	2.6
1	A	629	GLN	2.5
1	A	613	GLY	2.5
1	A	602	LYS	2.3
1	A	600	ALA	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	628	ILE	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	SNN	A	575	8/8	0.94	0.09	19,22,31,31	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

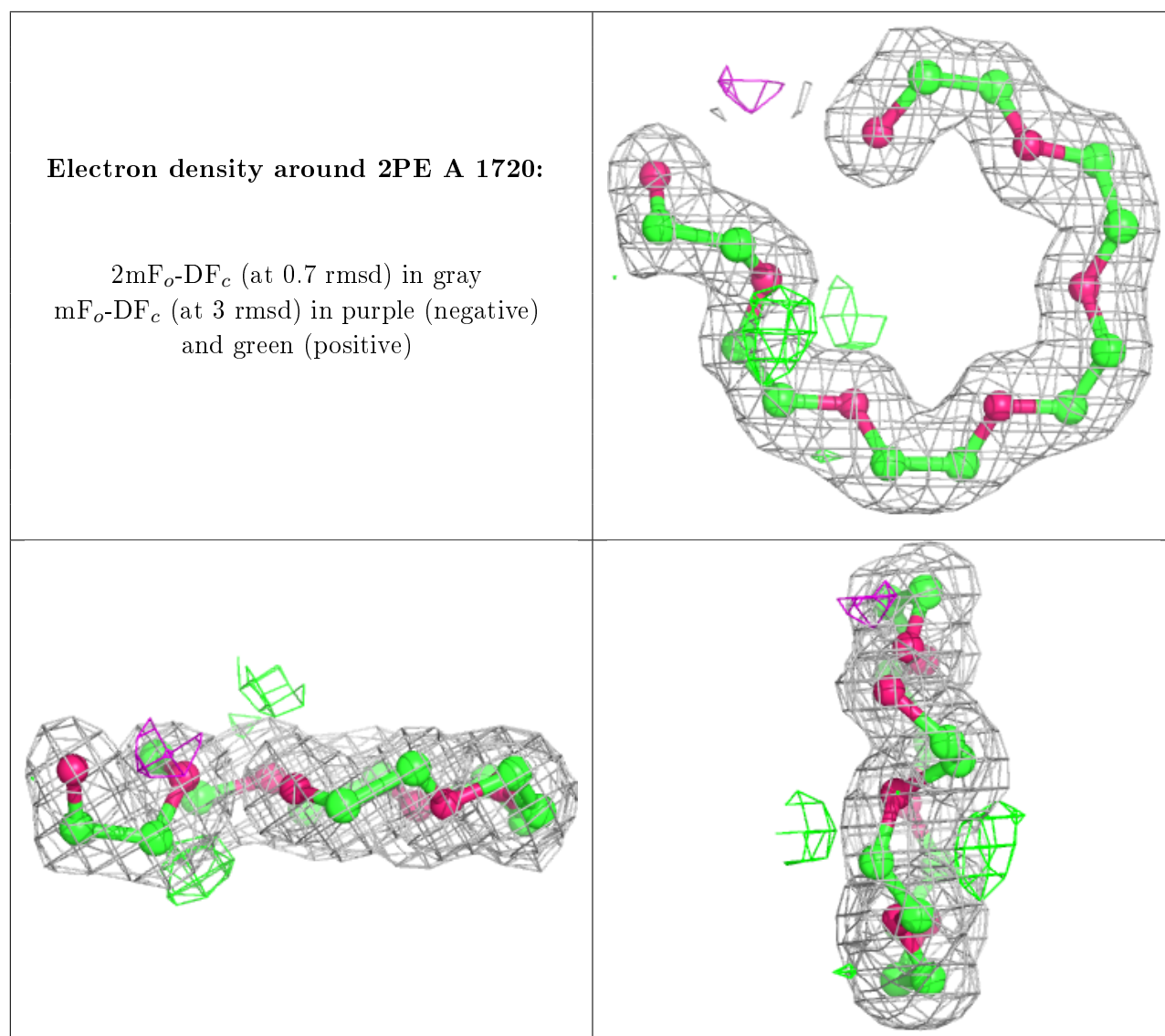
## 6.4 Ligands [i](#)

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
7	PGE	A	1727	7/10	0.67	0.20	71,73,74,74	0
5	EDO	A	1723	4/4	0.77	0.18	54,54,54,57	0
7	PGE	A	1725	7/10	0.88	0.17	48,51,56,59	0
5	EDO	A	1726[B]	4/4	0.89	0.15	7,21,24,31	4
5	EDO	A	1726[A]	4/4	0.89	0.15	14,23,24,25	4
5	EDO	A	1721	4/4	0.90	0.30	38,38,38,38	4
5	EDO	A	1722	4/4	0.93	0.44	30,34,35,36	4
4	2PE	A	1720	19/28	0.95	0.07	20,25,34,34	0
8	ACT	A	1728	4/4	0.97	0.10	24,25,25,29	0
6	CA	A	1724[A]	1/1	0.99	0.10	12,12,12,12	1
2	BM3	A	1718	15/15	0.99	0.10	11,12,13,14	0
6	CA	A	1724[B]	1/1	0.99	0.10	23,23,23,23	1
9	CL	A	1729	1/1	1.00	0.10	13,13,13,13	1
3	ZN	A	1719	1/1	1.00	0.08	14,14,14,14	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers ⓘ

There are no such residues in this entry.