



# wwPDB X-ray Structure Validation Summary Report

Aug 8, 2014 – 01:55 PM EDT

PDB ID : 4D0Z  
Title : GalNAc-T2 crystal soaked with UDP-5SGalNAc, mEA2 and manganese (Higher resolution dataset)  
Authors : Lira-Navarrete, E.; Iglesias-Fernandez, J.; Zandberg, W.F.; Companon, I.; Kong, Y.; Corzana, F.; Pinto, B.M.; Clausen, H.; Peregrina, J.M.; Vocadlo, D.; Rovira, C.; Hurtado-Guerrero, R.  
Deposited on : 2014-04-30  
Resolution : 2.20 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

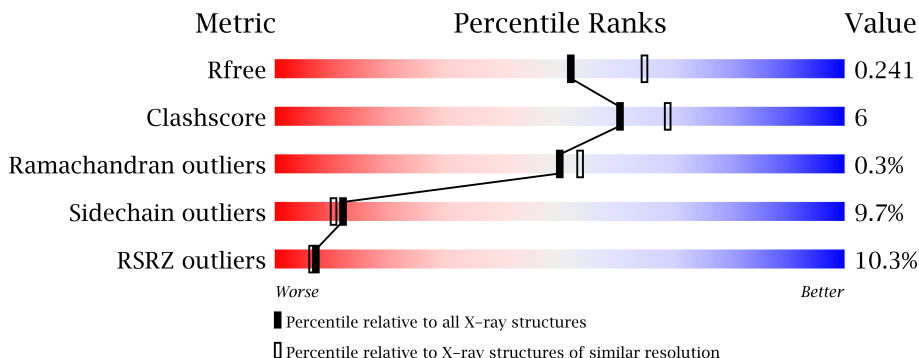
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23489  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23489

# 1 Overall quality at a glance

The reported resolution of this entry is 2.20 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))
$R_{free}$	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (2.20-2.20)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	571	
1	B	571	
1	C	571	
1	D	571	
1	E	571	
1	F	571	
2	X	6	
2	Y	6	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	EDO	A	1571	-	X
4	EDO	A	1572	-	X
4	EDO	A	1573	-	X
4	EDO	A	1576	-	X
4	EDO	A	1577	-	X
4	EDO	A	1579	-	X
4	EDO	A	1580	-	X
4	EDO	B	1573	-	X
4	EDO	B	1575	-	X
4	EDO	B	1576	-	X
4	EDO	B	1578	-	X
4	EDO	D	1576	-	X
4	EDO	D	1577	-	X
4	EDO	E	1570	-	X
4	EDO	E	1573	-	X
4	EDO	E	1576	-	X
4	EDO	E	1577	-	X
4	EDO	E	1578	-	X
4	EDO	E	1580	-	X
4	EDO	E	1581	-	X
5	BBK	B	1580	-	X
5	BBK	D	1572	-	X
5	BBK	E	1584	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 24773 atoms, of which 0 are hydrogen and 0 are deuterium.

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 POLYPEPTIDE N-ACETYL GALACTOSAMINYLTRANSFERASE2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	495	Total	C	N	O	S	20	1	0
			3975	2502	722	727	24			
1	B	495	Total	C	N	O	S	20	0	0
			3967	2497	719	727	24			
1	C	476	Total	C	N	O	S	20	0	0
			3826	2407	693	702	24			
1	D	495	Total	C	N	O	S	20	0	0
			3967	2497	719	727	24			
1	E	495	Total	C	N	O	S	20	0	0
			3967	2497	719	727	24			
1	F	491	Total	C	N	O	S	20	0	0
			3926	2470	710	722	24			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	516	ASP	ASN	ENGINEERED MUTATION	UNP Q10471
B	516	ASP	ASN	ENGINEERED MUTATION	UNP Q10471
C	516	ASP	ASN	ENGINEERED MUTATION	UNP Q10471
D	516	ASP	ASN	ENGINEERED MUTATION	UNP Q10471
E	516	ASP	ASN	ENGINEERED MUTATION	UNP Q10471
F	516	ASP	ASN	ENGINEERED MUTATION	UNP Q10471

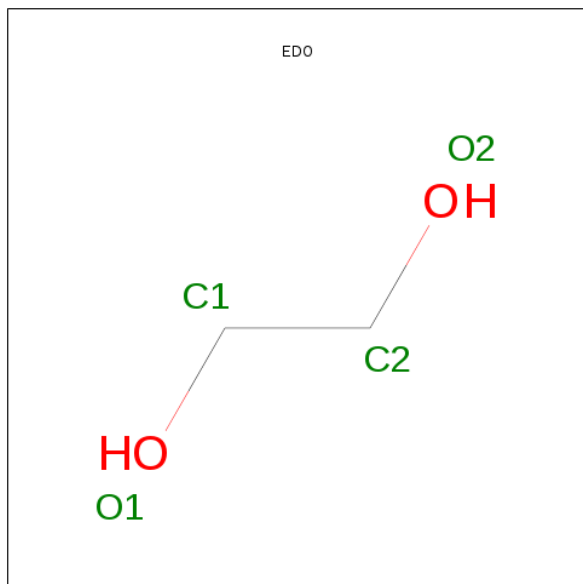
- Molecule 2 is a protein called PEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	6	Total	C	N	O	S	0	0	1
			32	18	6	7	1			
2	Y	6	Total	C	N	O	S	0	0	1
			32	18	6	7	1			

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total	Mn	0	0
			1	1		
3	E	1	Total	Mn	0	0
			1	1		
3	B	1	Total	Mn	0	0
			1	1		
3	C	1	Total	Mn	0	0
			1	1		
3	A	1	Total	Mn	0	0
			1	1		
3	F	1	Total	Mn	0	0
			1	1		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



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

*Continued on next page...*

*Continued from previous page...*

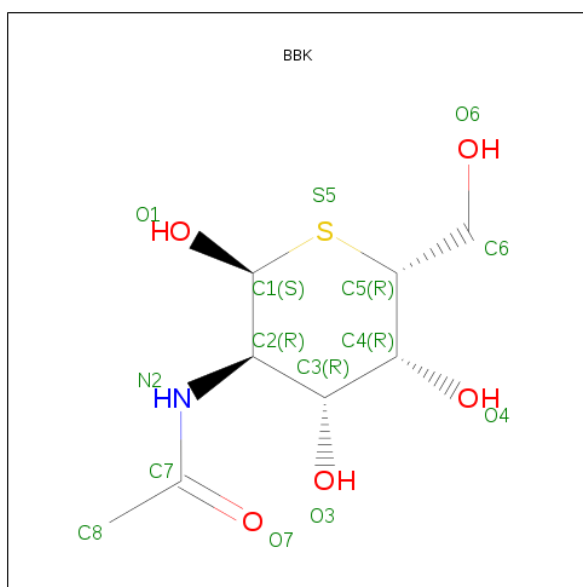
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0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

*Continued on next page...*

Continued from previous page...

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

- Molecule 5 is SUGAR (2-(ACETYLAMINO)-2-DEOXY-5-THIO-ALPHA-D-GALACTOPYRANOSE) (three-letter code: BBK) (formula:  $C_8H_{15}NO_5S$ ).



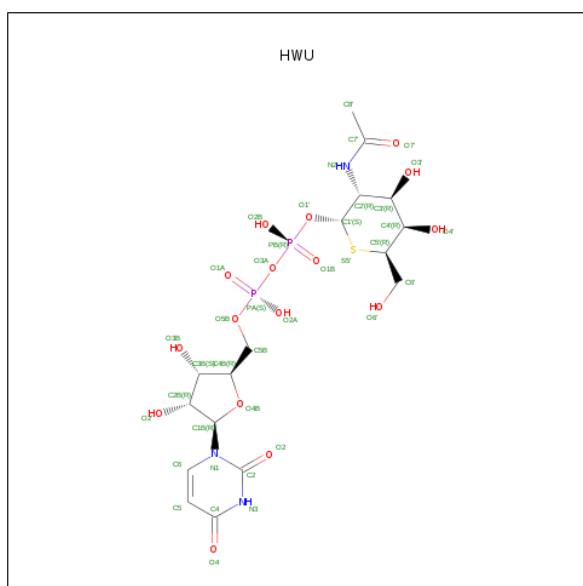
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			15	8	1	5	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	S	0	0
			15	8	1	5	1		
5	D	1	Total	C	N	O	S	0	0
			15	8	1	5	1		
5	E	1	Total	C	N	O	S	0	0
			15	8	1	5	1		

- Molecule 6 is (2R,3R,4R,5R,6R)-3-(ACETYLAMINO)-4,5-DIHYDROXY-6-(HYDROXYMETHYL)TETRAHYDRO-2H-THIOPYRAN-2-YL[(2R,3S,4R,5R)-5-(2,4-DIOXO-3,4-DIHYDROPYRIMIDIN-1(2H)-YL)-3,4-DIHYDROXYTETRAHYDROFURAN-2-YL]METHYLDIHYDROGEN DIPHOSPHATE (three-letter code: HWU) (formula: C<sub>17</sub>H<sub>27</sub>N<sub>3</sub>O<sub>16</sub>P<sub>2</sub>S).



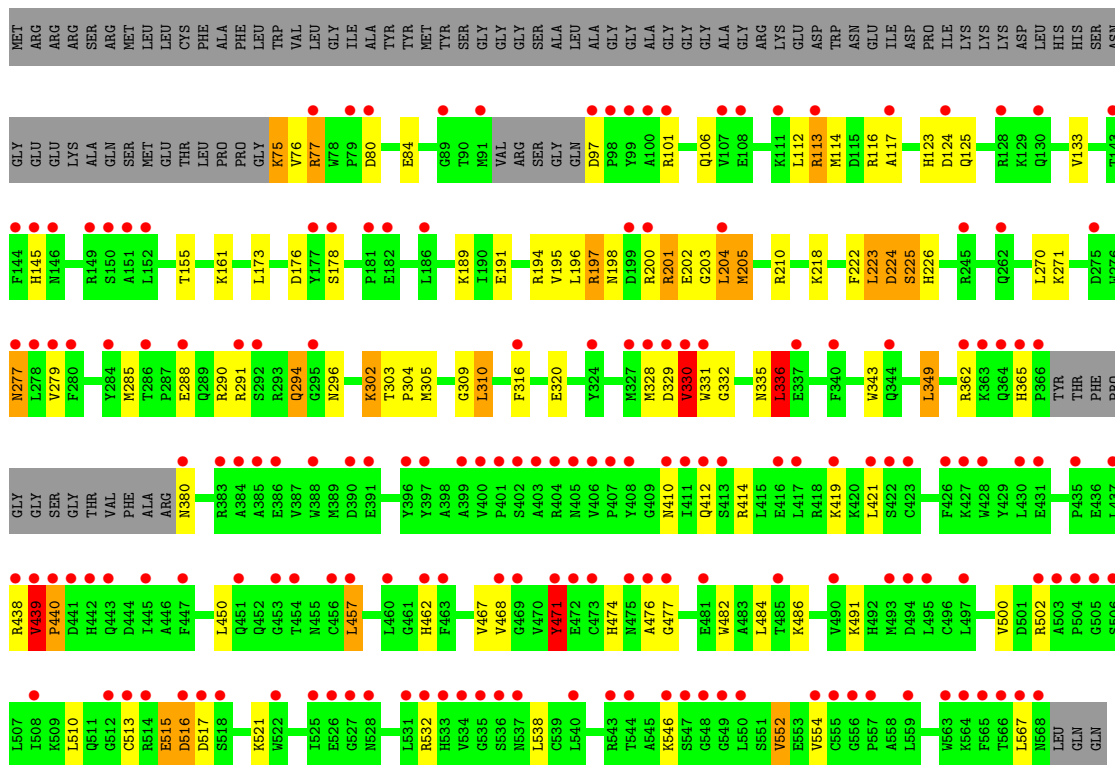
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	A	1	Total 39	C 17	N 3	O 16	P 2	S 1	0	0
6	B	1	Total 39	C 17	N 3	O 16	P 2	S 1	0	0
6	C	1	Total 39	C 17	N 3	O 16	P 2	S 1	0	0
6	D	1	Total 39	C 17	N 3	O 16	P 2	S 1	0	0
6	E	1	Total 39	C 17	N 3	O 16	P 2	S 1	0	0
6	F	1	Total 39	C 17	N 3	O 16	P 2	S 1	0	0

- Molecule 7 is water.



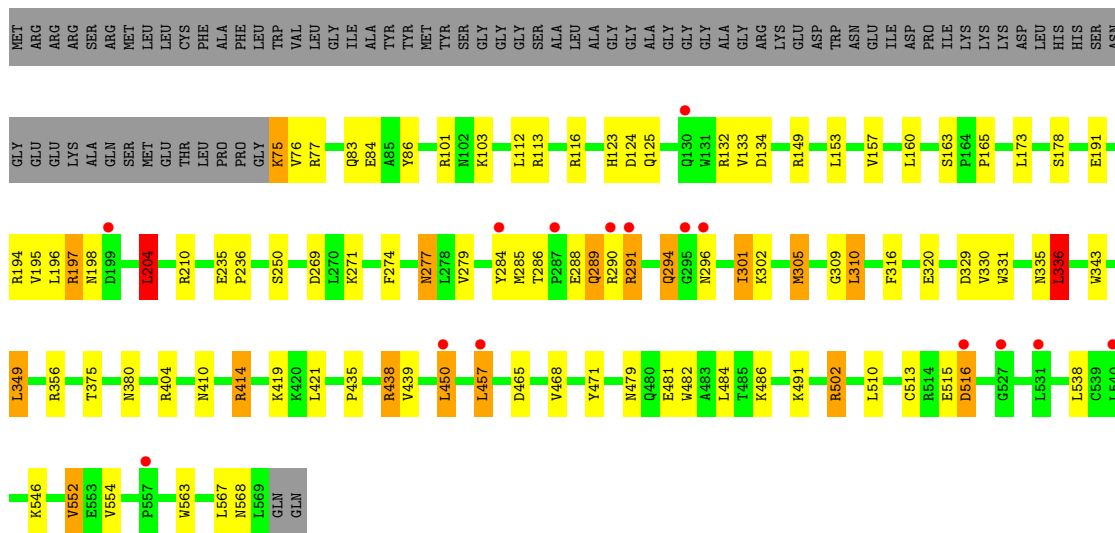
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	144	Total 144	O 144	0	0
7	B	125	Total 125	O 125	0	0
7	C	13	Total 13	O 13	0	0
7	D	122	Total 122	O 122	0	0
7	E	168	Total 168	O 168	0	0
7	F	61	Total 61	O 61	0	0





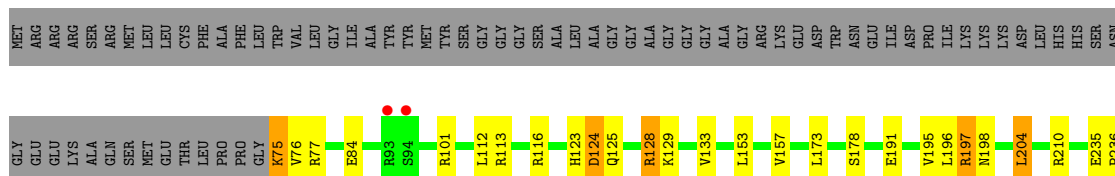
- Molecule 1: POLYPEPTIDE N-ACETYLGALACTOSAMINYLTRANSFERASE2

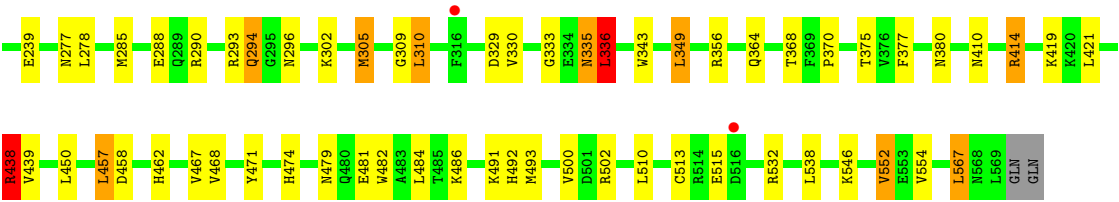
Chain D:



- Molecule 1: POLYPEPTIDE N-ACETYLGALACTOSAMINYLTRANSFERASE2

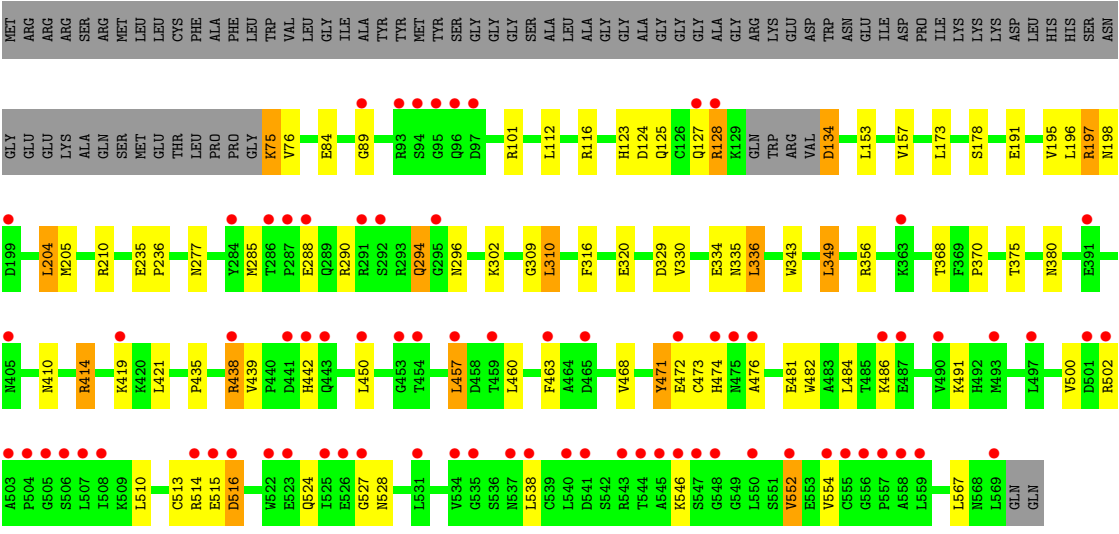
Chain E:





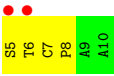
• Molecule 1: POLYPEPTIDE N-ACETYLGALACTOSAMINYLTRANSFERASE2

Chain F:



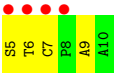
• Molecule 2: PEPTIDE

Chain X:



• Molecule 2: PEPTIDE

Chain Y:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.48Å 121.14Å 249.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	249.39 – 2.20 19.95 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (249.39-2.20) 99.8 (19.95-2.20)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.51 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.196 , 0.234 0.206 , 0.241	Depositor DCC
$R_{free}$ test set	4936 reflections (2.85%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.8	Xtriage
Anisotropy	0.540	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 38.2	EDS
Estimated twinning fraction	0.016 for k,h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 178449 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	24773	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.55% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: HWU, BBK, EDO, MN

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	0.85	4/4070 (0.1%)	0.98	25/5503 (0.5%)
1	B	0.79	3/4059 (0.1%)	0.91	9/5489 (0.2%)
1	C	0.88	7/3912 (0.2%)	0.98	26/5287 (0.5%)
1	D	1.01	3/4059 (0.1%)	0.98	19/5489 (0.3%)
1	E	0.92	3/4059 (0.1%)	0.98	19/5489 (0.3%)
1	F	0.89	3/4015 (0.1%)	0.90	12/5427 (0.2%)
2	X	1.42	0/32	1.09	0/44
2	Y	1.82	1/32 (3.1%)	1.53	0/44
All	All	0.89	24/24238 (0.1%)	0.95	110/32772 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	C	0	1
1	F	0	1
All	All	0	4

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	75	LYS	CB-CG	-33.78	0.61	1.52
1	C	84	GLU	CB-CG	-26.41	1.01	1.52
1	D	486	LYS	CB-CG	-25.36	0.84	1.52
1	F	75	LYS	CB-CG	-25.00	0.85	1.52
1	E	84	GLU	CB-CG	-23.77	1.06	1.52

The worst 5 of 110 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	75	LYS	CA-CB-CG	21.27	160.19	113.40
1	D	194	ARG	NE-CZ-NH2	-17.68	111.46	120.30
1	E	75	LYS	CA-CB-CG	16.43	149.55	113.40
1	F	75	LYS	CA-CB-CG	15.46	147.41	113.40
1	D	75	LYS	CB-CG-CD	13.48	146.66	111.60

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	89	GLY	Peptide
1	B	89	GLY	Peptide
1	C	223	LEU	Mainchain
1	F	89	GLY	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3975	0	3903	53	0
1	B	3967	0	3890	28	0
1	C	3826	0	3748	78	0
1	D	3967	0	3890	53	0
1	E	3967	0	3890	47	0
1	F	3926	0	3849	32	0
2	X	32	0	28	2	0
2	Y	32	0	28	1	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
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	40	0	60	9	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	32	0	48	6	0
4	D	20	0	30	5	0
4	E	56	0	84	15	0
5	A	15	0	0	3	0
5	B	15	0	0	2	0
5	D	15	0	0	2	0
5	E	15	0	0	3	0
6	A	39	0	0	0	0
6	B	39	0	0	1	0
6	C	39	0	0	6	0
6	D	39	0	0	0	0
6	E	39	0	0	1	0
6	F	39	0	0	2	0
7	A	144	0	0	0	0
7	B	125	0	0	3	0
7	C	13	0	0	6	0
7	D	122	0	0	3	0
7	E	168	0	0	3	0
7	F	61	0	0	0	0
All	All	24773	0	23448	281	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

The worst 5 of 281 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:203:GLY:HA3	1:C:330:VAL:CG2	1.71	1.19
1:C:203:GLY:CA	1:C:330:VAL:CG2	2.26	1.13
1:A:291:ARG:CZ	1:D:435:PRO:HB2	1.79	1.11
1:C:203:GLY:HA3	1:C:330:VAL:HG22	1.20	1.10
1:C:330:VAL:HG12	1:C:331:TRP:HB3	1.32	1.08

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	494/571 (86%)	479 (97%)	14 (3%)	1 (0%)	56	62
1	B	493/571 (86%)	476 (97%)	16 (3%)	1 (0%)	56	62
1	C	470/571 (82%)	451 (96%)	16 (3%)	3 (1%)	33	32
1	D	493/571 (86%)	478 (97%)	14 (3%)	1 (0%)	56	62
1	E	493/571 (86%)	479 (97%)	13 (3%)	1 (0%)	56	62
1	F	487/571 (85%)	473 (97%)	13 (3%)	1 (0%)	56	62
2	X	4/6 (67%)	3 (75%)	1 (25%)	0	100	100
2	Y	4/6 (67%)	3 (75%)	1 (25%)	0	100	100
All	All	2938/3438 (86%)	2842 (97%)	88 (3%)	8 (0%)	50	53

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	330	VAL
1	C	440	PRO
1	B	330	VAL
1	C	332	GLY
1	C	516	ASP

### 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	429/485 (88%)	388 (90%)	41 (10%)	12	11
1	B	428/485 (88%)	390 (91%)	38 (9%)	14	13
1	C	414/485 (85%)	371 (90%)	43 (10%)	10	9
1	D	428/485 (88%)	385 (90%)	43 (10%)	11	10
1	E	428/485 (88%)	389 (91%)	39 (9%)	14	13
1	F	424/485 (87%)	382 (90%)	42 (10%)	11	10
2	X	4/4 (100%)	3 (75%)	1 (25%)	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	Y	4/4 (100%)	2 (50%)	2 (50%)	0	0
All	All	2559/2918 (88%)	2310 (90%)	249 (10%)	12	10

5 of 249 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	471	TYR
1	D	296	ASN
1	F	438	ARG
1	C	502	ARG
1	D	116	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 98 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	452	GLN
1	D	335	ASN
1	F	380	ASN
1	C	462	HIS
1	D	123	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 53 ligands modelled in this entry, 6 are monoatomic - leaving 47 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	EDO	A	1571	-	3,3,3	0.89	0	2,2,2	0.27	0
4	EDO	A	1572	-	3,3,3	0.54	0	2,2,2	0.58	0
4	EDO	A	1573	-	3,3,3	0.85	0	2,2,2	0.04	0
4	EDO	A	1574	-	3,3,3	0.57	0	2,2,2	0.14	0
4	EDO	A	1575	-	3,3,3	0.40	0	2,2,2	0.70	0
4	EDO	A	1576	-	3,3,3	0.34	0	2,2,2	0.80	0
4	EDO	A	1577	-	3,3,3	0.59	0	2,2,2	0.18	0
4	EDO	A	1578	-	3,3,3	0.73	0	2,2,2	0.64	0
4	EDO	A	1579	-	3,3,3	0.73	0	2,2,2	0.43	0
4	EDO	A	1580	-	3,3,3	0.98	0	2,2,2	1.18	0
5	BBK	A	1581	-	15,15,15	8.44	2 (13%)	21,21,21	2.70	7 (33%)
6	HWU	A	1582	3	41,41,41	5.96	6 (14%)	58,62,62	1.85	14 (24%)
6	HWU	B	1571	3	41,41,41	6.21	12 (29%)	58,62,62	2.62	19 (32%)
4	EDO	B	1572	-	3,3,3	0.54	0	2,2,2	0.61	0
4	EDO	B	1573	-	3,3,3	0.30	0	2,2,2	0.36	0
4	EDO	B	1574	-	3,3,3	0.64	0	2,2,2	0.36	0
4	EDO	B	1575	-	3,3,3	0.49	0	2,2,2	0.39	0
4	EDO	B	1576	-	3,3,3	0.68	0	2,2,2	0.15	0
4	EDO	B	1577	-	3,3,3	0.30	0	2,2,2	1.17	0
4	EDO	B	1578	-	3,3,3	0.58	0	2,2,2	0.42	0
4	EDO	B	1579	-	3,3,3	0.74	0	2,2,2	0.05	0
5	BBK	B	1580	-	15,15,15	8.32	2 (13%)	21,21,21	2.48	5 (23%)
6	HWU	C	1569	3	41,41,41	7.45	8 (19%)	58,62,62	2.36	20 (34%)
6	HWU	D	1571	3	41,41,41	6.33	12 (29%)	58,62,62	2.41	17 (29%)
5	BBK	D	1572	-	15,15,15	8.78	3 (20%)	21,21,21	3.31	8 (38%)
4	EDO	D	1573	-	3,3,3	0.51	0	2,2,2	0.79	0
4	EDO	D	1574	-	3,3,3	0.84	0	2,2,2	0.50	0
4	EDO	D	1575	-	3,3,3	0.93	0	2,2,2	0.37	0
4	EDO	D	1576	-	3,3,3	0.63	0	2,2,2	0.60	0
4	EDO	D	1577	-	3,3,3	0.50	0	2,2,2	0.73	0
4	EDO	E	1570	-	3,3,3	0.78	0	2,2,2	0.37	0
4	EDO	E	1571	-	3,3,3	0.72	0	2,2,2	0.26	0
4	EDO	E	1572	-	3,3,3	0.45	0	2,2,2	0.43	0
4	EDO	E	1573	-	3,3,3	0.34	0	2,2,2	1.02	0
4	EDO	E	1574	-	3,3,3	0.40	0	2,2,2	0.88	0
4	EDO	E	1575	-	3,3,3	0.50	0	2,2,2	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	E	1576	-	3,3,3	0.58	0	2,2,2	0.57	0
4	EDO	E	1577	-	3,3,3	0.39	0	2,2,2	0.80	0
4	EDO	E	1578	-	3,3,3	0.44	0	2,2,2	0.72	0
4	EDO	E	1579	-	3,3,3	0.50	0	2,2,2	0.65	0
4	EDO	E	1580	-	3,3,3	0.67	0	2,2,2	0.27	0
4	EDO	E	1581	-	3,3,3	0.39	0	2,2,2	0.92	0
4	EDO	E	1582	-	3,3,3	0.40	0	2,2,2	0.34	0
4	EDO	E	1583	-	3,3,3	0.63	0	2,2,2	0.16	0
5	BBK	E	1584	-	15,15,15	8.50	2 (13%)	21,21,21	2.61	7 (33%)
6	HWU	E	1585	3	41,41,41	6.23	7 (17%)	58,62,62	2.15	15 (25%)
6	HWU	F	1571	3	41,41,41	5.86	7 (17%)	58,62,62	1.92	12 (20%)

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	EDO	A	1571	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1572	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1573	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1574	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1575	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1576	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1577	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1578	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1579	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1580	-	-	0/1/1/1	0/0/0/0
5	BBK	A	1581	-	-	2/6/26/26	0/1/1/1
6	HWU	A	1582	3	-	0/24/63/63	0/3/3/3
6	HWU	B	1571	3	-	0/24/63/63	0/3/3/3
4	EDO	B	1572	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1573	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1574	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1575	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1576	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1577	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1578	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1579	-	-	0/1/1/1	0/0/0/0
5	BBK	B	1580	-	-	0/6/26/26	0/1/1/1
6	HWU	C	1569	3	-	1/24/63/63	0/3/3/3
6	HWU	D	1571	3	-	0/24/63/63	0/3/3/3

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	BBK	D	1572	-	-	0/6/26/26	0/1/1/1
4	EDO	D	1573	-	-	0/1/1/1	0/0/0/0
4	EDO	D	1574	-	-	0/1/1/1	0/0/0/0
4	EDO	D	1575	-	-	0/1/1/1	0/0/0/0
4	EDO	D	1576	-	-	0/1/1/1	0/0/0/0
4	EDO	D	1577	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1570	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1571	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1572	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1573	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1574	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1575	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1576	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1577	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1578	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1579	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1580	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1581	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1582	-	-	0/1/1/1	0/0/0/0
4	EDO	E	1583	-	-	0/1/1/1	0/0/0/0
5	BBK	E	1584	-	-	0/6/26/26	0/1/1/1
6	HWU	E	1585	3	-	0/24/63/63	0/3/3/3
6	HWU	F	1571	3	-	0/24/63/63	0/3/3/3

The worst 5 of 61 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	1569	HWU	C1'-S5'	-38.24	1.42	1.81
6	E	1585	HWU	C1'-S5'	-32.71	1.47	1.81
6	B	1571	HWU	C1'-S5'	-32.55	1.48	1.81
6	D	1571	HWU	C1'-S5'	-31.19	1.49	1.81
6	A	1582	HWU	C1'-S5'	-31.18	1.49	1.81

The worst 5 of 124 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	1571	HWU	N3-C2-N1	10.62	124.83	115.97
6	D	1571	HWU	C2-N1-C1B	9.24	124.00	118.21
5	D	1572	BBK	C1-C2-N2	-9.24	93.88	111.53
6	E	1585	HWU	N3-C2-N1	9.23	123.68	115.97
5	A	1581	BBK	C1-S5-C5	8.81	116.75	96.41

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1581	BBK	O7-C7-N2-C2
5	A	1581	BBK	C8-C7-N2-C2
6	C	1569	HWU	PB-O1'-C1'-C2'

There are no ring outliers.

## 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 ⓘ

### 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	495/571 (86%)	-0.11	15 (3%) 48 48	26, 42, 70, 130	5 (1%)
1	B	495/571 (86%)	-0.17	14 (2%) 50 51	26, 40, 66, 106	5 (1%)
1	C	476/571 (83%)	1.82	173 (36%) 1 0	42, 92, 135, 168	5 (1%)
1	D	495/571 (86%)	-0.11	15 (3%) 48 48	26, 41, 66, 142	5 (1%)
1	E	495/571 (86%)	-0.27	4 (0%) 83 85	23, 36, 58, 90	5 (1%)
1	F	491/571 (85%)	0.54	78 (15%) 3 2	34, 57, 112, 140	5 (1%)
2	X	6/6 (100%)	1.17	2 (33%) 1 1	40, 73, 84, 87	0
2	Y	6/6 (100%)	2.58	4 (66%) 0 0	42, 71, 77, 83	0
All	All	2959/3438 (86%)	0.28	305 (10%) 7 6	23, 45, 109, 168	30 (1%)

The worst 5 of 305 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	331	TRP	10.8
1	F	547	SER	9.7
1	C	330	VAL	9.5
1	F	556	GLY	9.0
1	C	476	ALA	8.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	EDO	E	1578	4/4	0.22	12.47	64,64,66,82	0
4	EDO	A	1577	4/4	0.47	11.36	62,68,71,76	0
5	BBK	E	1584	15/15	0.31	8.19	96,105,117,124	0
5	BBK	D	1572	15/15	0.30	7.51	96,105,117,124	0
4	EDO	E	1580	4/4	0.39	7.44	59,70,71,76	0
4	EDO	A	1576	4/4	0.28	6.95	57,62,66,75	0
5	BBK	B	1580	15/15	0.28	6.84	69,86,120,142	0
4	EDO	E	1570	4/4	0.37	6.60	44,48,59,61	0
4	EDO	B	1576	4/4	0.20	5.71	65,69,73,76	0
4	EDO	D	1577	4/4	0.18	5.43	67,68,69,69	0
4	EDO	A	1580	4/4	0.20	4.45	45,52,57,59	0
4	EDO	A	1573	4/4	0.20	4.43	56,68,70,70	0
4	EDO	E	1576	4/4	0.17	4.24	68,68,69,69	0
4	EDO	E	1577	4/4	0.25	4.00	61,61,64,68	0
4	EDO	B	1573	4/4	0.23	3.87	40,41,42,48	0
4	EDO	A	1579	4/4	0.19	3.84	51,52,54,64	0
4	EDO	E	1573	4/4	0.24	2.93	49,58,60,71	0
4	EDO	B	1575	4/4	0.28	2.92	66,68,69,69	0
4	EDO	A	1572	4/4	0.23	2.72	57,66,67,69	0
4	EDO	A	1571	4/4	0.23	2.48	43,46,46,52	0
4	EDO	D	1576	4/4	0.17	2.36	39,43,45,49	0
4	EDO	E	1581	4/4	0.17	2.22	53,56,60,68	0
4	EDO	B	1578	4/4	0.28	2.08	55,58,62,69	0
4	EDO	A	1578	4/4	0.20	1.95	38,45,48,63	0
4	EDO	E	1583	4/4	0.14	1.91	58,61,62,68	0
4	EDO	D	1573	4/4	0.13	1.71	38,40,40,45	0
4	EDO	A	1574	4/4	0.14	1.61	51,52,54,57	0
4	EDO	B	1577	4/4	0.19	1.48	42,43,48,57	0
4	EDO	E	1572	4/4	0.21	1.39	38,45,48,51	0
5	BBK	A	1581	15/15	0.23	1.36	69,86,120,142	0
4	EDO	D	1574	4/4	0.15	1.15	49,50,52,58	0
4	EDO	B	1579	4/4	0.12	0.47	41,48,56,56	0
4	EDO	D	1575	4/4	0.18	0.37	60,63,68,70	0
4	EDO	E	1574	4/4	0.19	0.12	62,68,68,75	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	HWU	F	1571	39/39	0.11	0.09	38,44,55,60	0
4	EDO	E	1575	4/4	0.15	-0.02	64,69,69,76	0
6	HWU	D	1571	39/39	0.11	-0.03	24,35,45,46	0
6	HWU	B	1571	39/39	0.12	-0.06	30,36,44,49	0
4	EDO	A	1575	4/4	0.10	-0.09	53,55,57,59	0
6	HWU	A	1582	39/39	0.11	-0.13	26,34,45,48	0
4	EDO	E	1582	4/4	0.11	-0.18	47,48,51,56	0
4	EDO	B	1574	4/4	0.10	-0.29	40,40,42,45	0
6	HWU	E	1585	39/39	0.11	-0.34	26,36,54,58	0
4	EDO	E	1571	4/4	0.12	-0.44	38,46,47,50	0
4	EDO	E	1579	4/4	0.10	-0.57	43,44,52,53	0
6	HWU	C	1569	39/39	0.17	-0.61	51,76,89,92	0
4	EDO	B	1572	4/4	0.08	-1.14	34,35,35,42	0
3	MN	A	1570	1/1	0.07	-1.34	27,27,27,27	0
3	MN	B	1570	1/1	0.08	-1.56	33,33,33,33	0
3	MN	D	1570	1/1	0.07	-1.84	29,29,29,29	0
3	MN	E	1586	1/1	0.07	-2.37	26,26,26,26	0
3	MN	C	1570	1/1	0.03	-2.52	59,59,59,59	0
3	MN	F	1570	1/1	0.03	-9.30	42,42,42,42	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.