



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 09:28 pm GMT

PDB ID : 4CA6  
Title : Human Angiotensin converting enzyme N-domain in complex with a phosphinic tripeptide FI  
Authors : Masuyer, G.; Akif, M.; Czarny, B.; Beau, F.; Schwager, S.L.U.; Sturrock, E.D.; Isaac, R.E.; Dive, V.; Acharya, K.R.  
Deposited on : 2013-10-07  
Resolution : 1.91 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

<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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

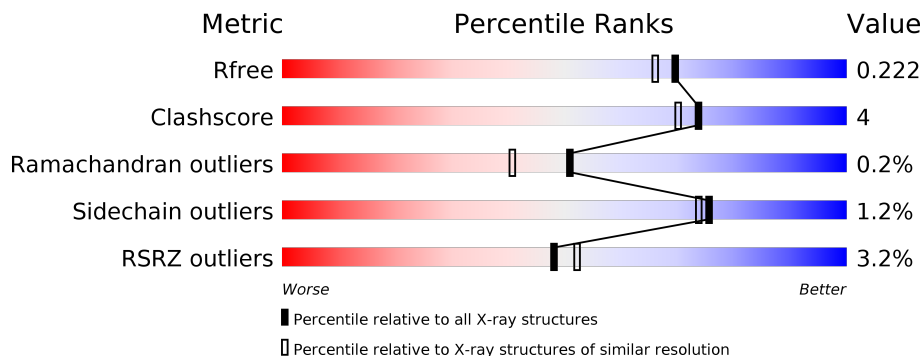
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.91 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))
$R_{free}$	100719	6276 (1.94-1.90)
Clashscore	112137	7025 (1.94-1.90)
Ramachandran outliers	110173	6947 (1.94-1.90)
Sidechain outliers	110143	6948 (1.94-1.90)
RSRZ outliers	101464	6332 (1.94-1.90)

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	610	<div> <div>3%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div>.</div> </div> </div>
1	B	610	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>8%</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
10	NAG	B	1616	-	-	-	X
11	P6G	B	1622	-	-	-	X
6	NAG	A	1616	X	-	-	-
7	PEG	A	1626	-	-	-	X
7	PEG	B	1623	-	-	-	X
7	PEG	B	1624	-	-	-	X

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 10900 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 ANGIOTENSIN-CONVERTING ENZYME N-DOMAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	608	Total	C	N	O	S	0	1	0
			4967	3190	853	905	19			
1	B	608	Total	C	N	O	S	0	2	0
			4970	3193	853	905	19			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	GLN	ASN	CONFLICT	UNP P12821
A	25	GLN	ASN	CONFLICT	UNP P12821
A	82	GLN	ASN	CONFLICT	UNP P12821
A	117	GLN	ASN	CONFLICT	UNP P12821
A	289	GLN	ASN	CONFLICT	UNP P12821
A	545	ARG	GLN	ENGINEERED MUTATION	UNP P12821
A	576	LEU	PRO	ENGINEERED MUTATION	UNP P12821
B	9	GLN	ASN	CONFLICT	UNP P12821
B	25	GLN	ASN	CONFLICT	UNP P12821
B	82	GLN	ASN	CONFLICT	UNP P12821
B	117	GLN	ASN	CONFLICT	UNP P12821
B	289	GLN	ASN	CONFLICT	UNP P12821
B	545	ARG	GLN	ENGINEERED MUTATION	UNP P12821
B	576	LEU	PRO	ENGINEERED MUTATION	UNP P12821

- 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		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 1 1	0	0
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

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

- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total C N O 28 16 2 10	0	0
5	B	2	Total C N O 28 16 2 10	0	0

- Molecule 6 is a polymer of unknown type called SUGAR (4-MER).

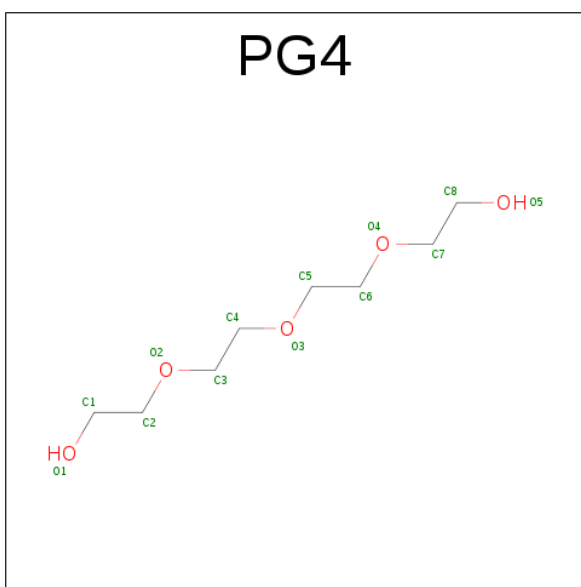
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	4	Total C N O 49 28 2 19	0	0

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



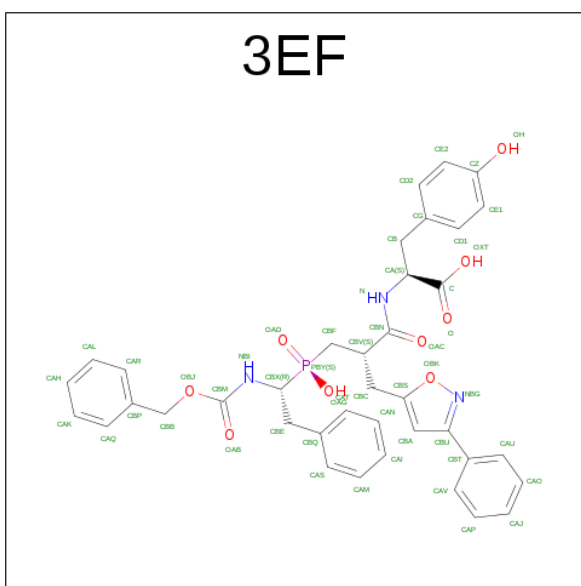
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	4	3		
7	A	1	Total	C	O	0	0
			7	4	3		
7	A	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 9 is N-{(2S)-3-[(S)-[(1R)-1-[(BENZYLOXY)CARBONYL]AMINO}-2-PHENYLETHYL](HYDROXY)PHOSPHORYL]-2-[(3-PHENYL-1,2-OXAZOL-5-YL)METHYL]PROPANOYL}-L-TYROSINE (three-letter code: 3EF) (formula: C<sub>38</sub>H<sub>38</sub>N<sub>3</sub>O<sub>9</sub>P).

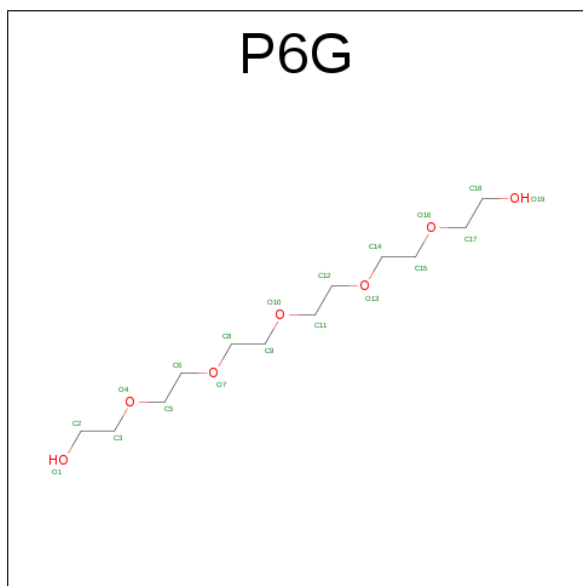


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	A	1	Total 51	C 38	N 3	O 9	P 1	0	0
9	B	1	Total 51	C 38	N 3	O 9	P 1	0	0

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	B	3	Total	C	N	O	0	0
			39	22	2	15		

- Molecule 11 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula:  $C_{12}H_{26}O_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	B	1	Total	C	O	0	0
			19	12	7		

- Molecule 12 is water.

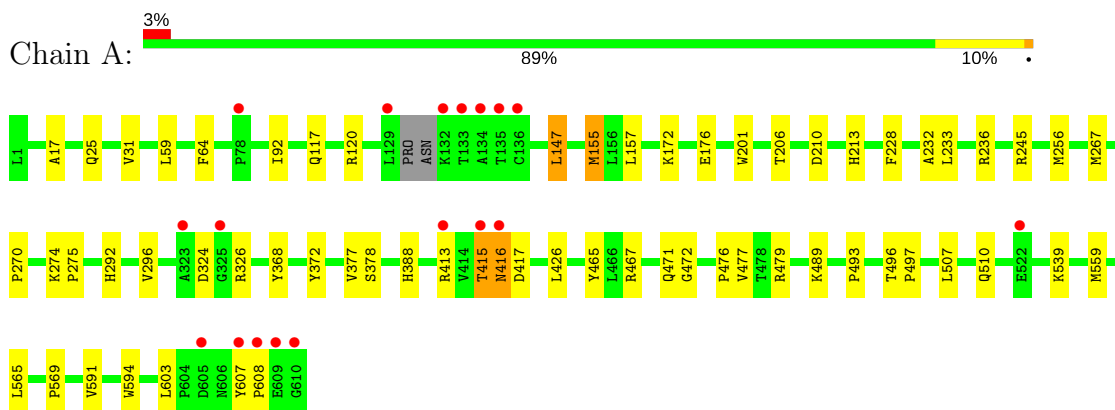
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	344	Total	O	0	0
			344	344		
12	B	250	Total	O	0	0
			250	250		



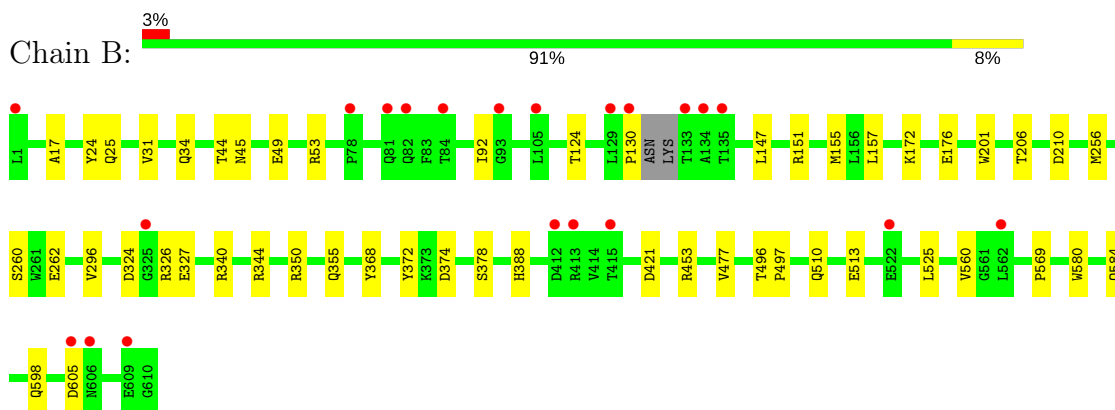
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: ANGIOTENSIN-CONVERTING ENZYME N-DOMAIN



#### • Molecule 1: ANGIOTENSIN-CONVERTING ENZYME N-DOMAIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.92Å 76.64Å 82.55Å 88.62° 64.22° 75.58°	Depositor
Resolution (Å)	29.74 – 1.91 29.73 – 1.91	Depositor EDS
% Data completeness (in resolution range)	96.4 (29.74-1.91) 93.8 (29.73-1.91)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.83 (at 1.92Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.187 , 0.223 0.186 , 0.222	Depositor DCC
$R_{free}$ test set	5779 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.0	Xtriage
Anisotropy	0.158	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 40.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10900	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.37% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, BMA, NAG, CL, 3EF, PG4, FUC, P6G, PEG

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	0.43	0/5126	0.59	1/6982 (0.0%)
1	B	0.41	0/5132	0.56	1/6990 (0.0%)
All	All	0.42	0/10258	0.58	2/13972 (0.0%)

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
6	A	1	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	130	PRO	N-CA-CB	6.54	111.14	103.30
1	A	147	LEU	CA-CB-CG	-5.27	103.19	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	1616	NAG	C1

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	4967	0	4742	45	0
1	B	4970	0	4745	30	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	24	0	22	0	0
4	B	24	0	22	1	0
5	A	28	0	25	0	0
5	B	28	0	25	0	0
6	A	49	0	43	2	0
7	A	21	0	30	3	0
7	B	21	0	30	1	0
8	A	10	0	13	3	0
9	A	51	0	36	1	0
9	B	51	0	36	0	0
10	B	39	0	34	0	0
11	B	19	0	26	1	0
12	A	344	0	0	3	0
12	B	250	0	0	3	0
All	All	10900	0	9829	77	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 4.

All (77) 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)
9:A:1630:3EF:NBG	9:A:1630:3EF:OBK	1.62	1.03
1:B:147:LEU:HD22	1:B:256:MET:HA	1.56	0.88
8:A:1625:PG4:H32	12:A:2341:HOH:O	1.73	0.86
1:A:467:ARG:HH11	1:A:471:GLN:HE22	1.27	0.83
1:A:147:LEU:HD22	1:A:256:MET:HA	1.62	0.80
1:B:350:ARG:H	1:B:355:GLN:HE21	1.30	0.79
1:B:147:LEU:CD2	1:B:256:MET:HA	2.12	0.79

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:245:ARG:HG2	1:A:591:VAL:HG11	1.65	0.78
1:A:17:ALA:HB1	1:A:92:ILE:HD11	1.66	0.77
1:A:17:ALA:HB1	1:A:92:ILE:CD1	2.18	0.73
1:A:157:LEU:HD11	1:A:477:VAL:HG13	1.72	0.71
1:B:262:GLU:OE2	12:B:2129:HOH:O	2.11	0.69
1:A:206:THR:HG23	1:A:210:ASP:OD2	1.94	0.67
1:B:124:THR:HG22	1:B:327:GLU:HG2	1.78	0.65
1:A:236:ARG:HD3	1:A:267:MET:HE3	1.79	0.65
1:B:147:LEU:HD22	1:B:256:MET:CA	2.27	0.64
1:B:157:LEU:HD11	1:B:477:VAL:HG13	1.79	0.63
1:A:270:PRO:HD3	1:A:426:LEU:HD23	1.83	0.61
1:B:324:ASP:OD1	1:B:326:ARG:HB2	2.01	0.60
1:A:117:GLN:HG2	1:A:120:ARG:NH2	2.17	0.59
1:B:49:GLU:HG3	1:B:53:ARG:NH1	2.17	0.59
1:A:213[B]:HIS:HD2	12:A:2145:HOH:O	1.84	0.59
8:A:1625:PG4:O4	1:B:453:ARG:NH1	2.36	0.58
1:A:155:MET:HA	1:A:155:MET:HE3	1.85	0.58
1:A:507:LEU:HD13	1:A:565:LEU:CD2	2.35	0.57
1:A:324:ASP:OD1	1:A:326:ARG:HB2	2.06	0.56
1:B:344:ARG:HH22	7:B:1624:PEG:H42	1.70	0.56
1:B:17:ALA:HB1	1:B:92:ILE:HD11	1.88	0.56
1:A:25:GLN:HE21	1:A:378:SER:H	1.52	0.56
1:A:245:ARG:HG2	1:A:591:VAL:CG1	2.36	0.56
1:B:260:SER:OG	1:B:262:GLU:OE1	2.15	0.55
1:A:155:MET:HA	1:A:155:MET:CE	2.38	0.54
1:A:274:LYS:HB3	1:A:275:PRO:HD2	1.88	0.54
1:A:467:ARG:NH1	1:A:471:GLN:HE22	2.01	0.53
1:A:233:LEU:HD23	1:A:267:MET:CE	2.38	0.53
1:A:232:ALA:HB2	7:A:1626:PEG:H31	1.91	0.53
1:B:340:ARG:HH11	1:B:374:ASP:HA	1.74	0.52
1:A:270:PRO:HD3	1:A:426:LEU:CD2	2.39	0.52
1:A:507:LEU:HD13	1:A:565:LEU:HD23	1.92	0.52
1:A:233:LEU:HD23	1:A:267:MET:HE2	1.93	0.51
1:B:31:VAL:O	1:B:34:GLN:HG3	2.12	0.50
1:B:44:THR:O	1:B:326:ARG:HD2	2.12	0.49
1:A:479:ARG:H	1:B:598:GLN:NE2	2.10	0.49
1:B:151:ARG:HD2	12:B:2067:HOH:O	2.12	0.49
1:A:465:TYR:CE2	8:A:1625:PG4:H31	2.48	0.48
1:B:580:TRP:O	1:B:584:GLN:HG2	2.14	0.47
1:B:206:THR:HG23	1:B:210:ASP:OD2	2.16	0.46
1:B:25:GLN:HE21	1:B:378:SER:H	1.63	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:228:PHE:CD1	7:A:1626:PEG:H41	2.52	0.45
1:B:172:LYS:O	1:B:176:GLU:HG3	2.16	0.45
1:A:415:THR:HB	1:A:417:ASP:OD2	2.16	0.45
1:B:296:VAL:HG22	11:B:1622:P6G:H171	1.98	0.45
1:A:416:ASN:OD1	6:A:1616:NAG:C7	2.65	0.45
1:B:201:TRP:HZ3	1:B:497:PRO:HG2	1.81	0.45
1:B:24:TYR:HD2	1:B:25:GLN:HG3	1.82	0.45
1:A:489:LYS:O	1:A:493:PRO:HD2	2.18	0.44
4:B:1611:FUC:H3	12:B:2213:HOH:O	2.17	0.44
1:B:510:GLN:HG2	1:B:569:PRO:HG2	1.99	0.44
1:A:228:PHE:CE1	7:A:1626:PEG:H41	2.54	0.43
1:A:416:ASN:OD1	6:A:1616:NAG:N2	2.46	0.42
1:B:155:MET:HE3	1:B:155:MET:HA	2.02	0.42
1:A:117:GLN:HG2	1:A:120:ARG:HH22	1.83	0.42
1:A:31:VAL:HG21	1:A:64:PHE:CD1	2.55	0.42
1:A:59:LEU:HD22	12:A:2031:HOH:O	2.19	0.42
1:B:124:THR:HG22	1:B:327:GLU:CG	2.46	0.42
1:A:157:LEU:HD13	1:A:476:PRO:HB2	2.00	0.42
1:A:539:LYS:HE3	1:A:559:MET:O	2.19	0.42
1:B:513:GLU:HA	1:B:525:LEU:HD11	2.01	0.42
1:A:510:GLN:HG2	1:A:569:PRO:HG2	2.02	0.42
1:A:201:TRP:CZ3	1:A:497:PRO:HG2	2.56	0.41
1:A:496:THR:HA	1:A:497:PRO:HD3	1.81	0.41
1:A:172:LYS:O	1:A:176:GLU:HG3	2.21	0.41
1:A:472:GLY:HA2	1:A:594:TRP:CE2	2.56	0.40
1:B:496:THR:HA	1:B:497:PRO:HD3	1.91	0.40
1:A:607:TYR:HA	1:A:608:PRO:HA	1.91	0.40
1:A:292:HIS:O	1:A:296:VAL:HG23	2.22	0.40
1:A:477:VAL:HG12	1:A:603:LEU:HD21	2.04	0.40

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	605/610 (99%)	593 (98%)	12 (2%)	0	100	100
1	B	606/610 (99%)	590 (97%)	14 (2%)	2 (0%)	44	33
All	All	1211/1220 (99%)	1183 (98%)	26 (2%)	2 (0%)	51	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	45	ASN
1	B	605	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	522/524 (100%)	514 (98%)	8 (2%)	70	66
1	B	522/524 (100%)	517 (99%)	5 (1%)	80	79
All	All	1044/1048 (100%)	1031 (99%)	13 (1%)	75	74

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	155	MET
1	A	368	TYR
1	A	372	TYR
1	A	377	VAL
1	A	388	HIS
1	A	413	ARG
1	A	415	THR
1	A	416	ASN
1	B	368	TYR
1	B	372	TYR
1	B	388	HIS
1	B	421	ASP
1	B	560	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	9	GLN
1	A	25	GLN
1	A	217	GLN
1	A	371	GLN
1	A	471	GLN
1	A	491	HIS
1	B	25	GLN
1	B	355	GLN
1	B	371	GLN
1	B	491	HIS
1	B	598	GLN

### 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 [i](#)

15 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	$\# Z  > 2$
4	FUC	A	1611	4	9,10,11	0.62	0	13,14,16	1.46	2 (15%)
4	NAG	A	1612	1,4	14,14,15	0.53	0	15,19,21	1.02	1 (6%)
5	NAG	A	1614	1,5	14,14,15	0.53	0	15,19,21	1.13	3 (20%)
5	NAG	A	1615	5	14,14,15	0.44	0	15,19,21	1.12	1 (6%)
6	NAG	A	1616	1,6	14,14,15	1.25	2 (14%)	15,19,21	1.54	3 (20%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	A	1617	6	14,14,15	0.53	0	15,19,21	1.06	1 (6%)
6	BMA	A	1618	6	11,11,12	0.44	0	13,15,17	1.22	2 (15%)
6	FUC	A	1619	6	9,10,11	0.64	0	13,14,16	0.71	0
4	FUC	B	1611	4	9,10,11	0.76	0	13,14,16	2.06	4 (30%)
4	NAG	B	1612	1,4	14,14,15	0.55	0	15,19,21	1.15	1 (6%)
5	NAG	B	1614	1,5	14,14,15	0.62	0	15,19,21	1.05	1 (6%)
5	NAG	B	1615	5	14,14,15	0.45	0	15,19,21	1.77	2 (13%)
10	NAG	B	1616	1,10	14,14,15	0.54	0	15,19,21	0.76	0
10	NAG	B	1617	10	14,14,15	0.49	0	15,19,21	1.68	2 (13%)
10	BMA	B	1618	10	11,11,12	0.48	0	13,15,17	1.45	2 (15%)

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	FUC	A	1611	4	-	0/0/17/20	0/1/1/1
4	NAG	A	1612	1,4	-	0/6/23/26	0/1/1/1
5	NAG	A	1614	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1615	5	-	0/6/23/26	0/1/1/1
6	NAG	A	1616	1,6	1/1/5/7	0/6/23/26	0/1/1/1
6	NAG	A	1617	6	-	0/6/23/26	0/1/1/1
6	BMA	A	1618	6	-	0/2/19/22	0/1/1/1
6	FUC	A	1619	6	-	0/0/17/20	0/1/1/1
4	FUC	B	1611	4	-	0/0/17/20	0/1/1/1
4	NAG	B	1612	1,4	-	0/6/23/26	0/1/1/1
5	NAG	B	1614	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	1615	5	-	0/6/23/26	0/1/1/1
10	NAG	B	1616	1,10	-	0/6/23/26	0/1/1/1
10	NAG	B	1617	10	-	0/6/23/26	0/1/1/1
10	BMA	B	1618	10	-	0/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1616	NAG	O5-C1	-3.62	1.37	1.43
6	A	1616	NAG	C1-C2	2.56	1.56	1.52

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1617	NAG	O5-C1-C2	-3.10	107.17	111.47
6	A	1616	NAG	O5-C1-C2	-2.63	107.81	111.47
5	A	1614	NAG	O5-C1-C2	-2.29	108.29	111.47
4	A	1612	NAG	C2-N2-C7	-2.19	119.75	122.94
6	A	1616	NAG	C3-C4-C5	-2.02	106.65	110.22
5	A	1614	NAG	C4-C3-C2	2.01	113.96	111.02
6	A	1618	BMA	C1-O5-C5	2.03	114.96	112.17
5	B	1614	NAG	C4-C3-C2	2.17	114.19	111.02
5	B	1615	NAG	C3-C4-C5	2.23	114.15	110.22
6	A	1618	BMA	C1-C2-C3	2.28	112.54	109.65
5	A	1614	NAG	C1-O5-C5	2.32	115.36	112.17
4	B	1611	FUC	C3-C4-C5	2.47	113.56	109.68
4	B	1612	NAG	C4-C3-C2	2.84	115.18	111.02
10	B	1618	BMA	C1-C2-C3	2.89	113.32	109.65
4	A	1611	FUC	O5-C1-C2	2.99	115.47	110.79
4	B	1611	FUC	C1-O5-C5	3.07	119.17	112.39
10	B	1617	NAG	C4-C3-C2	3.27	115.81	111.02
10	B	1618	BMA	C1-O5-C5	3.59	117.11	112.17
4	A	1611	FUC	C1-C2-C3	3.63	114.26	109.65
5	A	1615	NAG	C1-O5-C5	3.73	117.30	112.17
4	B	1611	FUC	O5-C1-C2	3.82	116.77	110.79
6	A	1616	NAG	C1-O5-C5	4.07	117.78	112.17
4	B	1611	FUC	C1-C2-C3	4.15	114.91	109.65
10	B	1617	NAG	C1-O5-C5	4.29	118.08	112.17
5	B	1615	NAG	C1-O5-C5	5.80	120.16	112.17

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	1616	NAG	C1

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1616	NAG	2	0
4	B	1611	FUC	1	0

## 5.6 Ligand geometry

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	$\# Z  > 2$
7	PEG	A	1622	-	6,6,6	0.37	0	5,5,5	0.53	0
7	PEG	A	1624	-	6,6,6	0.45	0	5,5,5	0.44	0
8	PG4	A	1625	-	9,9,12	0.42	0	8,8,11	0.45	0
7	PEG	A	1626	-	6,6,6	0.53	0	5,5,5	0.24	0
9	3EF	A	1630	2	46,55,55	2.07	8 (17%)	52,75,75	1.24	8 (15%)
7	PEG	B	1621	-	6,6,6	0.49	0	5,5,5	0.26	0
11	P6G	B	1622	-	18,18,18	0.63	0	17,17,17	0.32	0
7	PEG	B	1623	-	6,6,6	0.47	0	5,5,5	0.25	0
7	PEG	B	1624	-	6,6,6	0.52	0	5,5,5	0.27	0
9	3EF	B	1630	2	46,55,55	2.02	7 (15%)	52,75,75	1.09	4 (7%)

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
7	PEG	A	1622	-	-	0/4/4/4	0/0/0/0
7	PEG	A	1624	-	-	0/4/4/4	0/0/0/0
8	PG4	A	1625	-	-	0/7/7/10	0/0/0/0
7	PEG	A	1626	-	-	0/4/4/4	0/0/0/0
9	3EF	A	1630	2	-	0/38/48/48	0/4/5/5
7	PEG	B	1621	-	-	0/4/4/4	0/0/0/0
11	P6G	B	1622	-	-	0/16/16/16	0/0/0/0
7	PEG	B	1623	-	-	0/4/4/4	0/0/0/0
7	PEG	B	1624	-	-	0/4/4/4	0/0/0/0
9	3EF	B	1630	2	-	0/38/48/48	0/4/5/5

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	B	1630	3EF	CBT-CBU	-10.00	1.33	1.48
9	A	1630	3EF	CBT-CBU	-9.66	1.34	1.48
9	A	1630	3EF	CB-CG	-4.81	1.39	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	B	1630	3EF	CB-CG	-4.56	1.40	1.51
9	A	1630	3EF	CBB-CBP	-3.92	1.41	1.50
9	B	1630	3EF	CBE-CBQ	-3.83	1.42	1.51
9	A	1630	3EF	CBE-CBQ	-3.48	1.42	1.51
9	A	1630	3EF	PBY-CBF	-3.47	1.76	1.79
9	B	1630	3EF	CBB-CBP	-3.35	1.42	1.50
9	B	1630	3EF	CBA-CBU	-2.78	1.35	1.40
9	A	1630	3EF	CBA-CBS	-2.68	1.34	1.39
9	B	1630	3EF	CBA-CBS	-2.46	1.35	1.39
9	A	1630	3EF	CBA-CBU	-2.33	1.36	1.40
9	B	1630	3EF	CBC-CBS	2.07	1.53	1.51
9	A	1630	3EF	CBC-CBS	2.50	1.54	1.51

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	1630	3EF	CBA-CBU-CBT	-2.94	125.35	129.44
9	B	1630	3EF	CBA-CBU-CBT	-2.17	126.43	129.44
9	A	1630	3EF	OAB-CBM-NBI	-2.04	121.36	124.87
9	A	1630	3EF	PBY-CBX-NBI	-2.03	104.55	109.13
9	A	1630	3EF	CBT-CBU-NBG	2.07	124.21	120.98
9	A	1630	3EF	CB-CA-N	2.46	113.84	108.94
9	A	1630	3EF	CBB-OBJ-CBM	2.47	121.73	115.91
9	B	1630	3EF	CB-CA-N	2.51	113.93	108.94
9	B	1630	3EF	OBJ-CBM-NBI	2.56	116.05	110.54
9	A	1630	3EF	OBJ-CBM-NBI	2.84	116.64	110.54
9	B	1630	3EF	CBU-CBA-CBS	3.50	110.20	105.43
9	A	1630	3EF	CBU-CBA-CBS	4.64	111.75	105.43

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	1625	PG4	3	0
7	A	1626	PEG	3	0
9	A	1630	3EF	1	0
11	B	1622	P6G	1	0
7	B	1624	PEG	1	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

### 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	608/610 (99%)	-0.14	18 (2%) 51 55	15, 25, 41, 94	0
1	B	608/610 (99%)	-0.00	21 (3%) 44 49	15, 29, 52, 82	0
All	All	1216/1220 (99%)	-0.07	39 (3%) 48 52	15, 27, 49, 94	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	325	GLY	10.5
1	B	130	PRO	7.7
1	B	135	THR	6.5
1	A	133	THR	5.2
1	B	134	ALA	4.9
1	A	134	ALA	4.7
1	A	132	LYS	4.5
1	B	133	THR	4.4
1	B	325	GLY	4.2
1	A	413	ARG	4.0
1	B	82	GLN	3.8
1	B	609	GLU	3.6
1	B	605	ASP	3.6
1	A	135	THR	3.5
1	A	323	ALA	3.5
1	B	129	LEU	3.5
1	B	415	THR	3.4
1	B	78	PRO	3.4
1	A	78	PRO	3.3
1	B	81	GLN	3.2
1	B	606	ASN	3.2
1	B	413	ARG	3.1
1	A	609	GLU	2.9
1	A	610	GLY	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	562	LEU	2.8
1	A	607	TYR	2.8
1	B	84	THR	2.5
1	A	129	LEU	2.4
1	A	416	ASN	2.4
1	B	93	GLY	2.2
1	A	136	CYS	2.2
1	A	522	GLU	2.2
1	B	105	LEU	2.2
1	A	415	THR	2.1
1	B	522	GLU	2.1
1	A	605	ASP	2.1
1	A	608	PRO	2.1
1	B	412	ASP	2.1
1	B	1	LEU	2.1

## 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(Å <sup>2</sup> )	Q<0.9
10	NAG	B	1616	14/15	0.84	0.23	4.39	52,55,58,60	0
5	NAG	B	1614	14/15	0.85	0.25	1.64	42,45,51,59	0
5	NAG	A	1614	14/15	0.85	0.20	1.22	42,44,47,51	0
4	NAG	A	1612	14/15	0.88	0.13	1.09	38,44,48,54	0
4	NAG	B	1612	14/15	0.86	0.13	1.07	40,45,53,54	0
6	NAG	A	1616	14/15	0.90	0.14	0.11	41,43,46,47	0
6	FUC	A	1619	10/11	0.82	0.17	-	46,47,49,50	0
5	NAG	B	1615	14/15	0.79	0.47	-	66,73,75,75	0
6	BMA	A	1618	11/12	0.82	0.23	-	57,59,60,60	0
10	BMA	B	1618	11/12	0.73	0.35	-	77,80,82,84	0
4	FUC	B	1611	10/11	0.61	0.26	-	59,63,65,65	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	NAG	A	1617	14/15	0.90	0.19	-	47,50,54,56	0
10	NAG	B	1617	14/15	0.85	0.37	-	65,69,71,74	0
4	FUC	A	1611	10/11	0.60	0.37	-	62,68,72,73	0
5	NAG	A	1615	14/15	0.77	0.30	-	50,58,64,65	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
7	PEG	B	1624	7/7	0.81	0.20	13.25	51,54,56,56	0
7	PEG	B	1623	7/7	0.88	0.20	9.12	42,43,46,48	0
11	P6G	B	1622	19/19	0.70	0.14	5.43	38,45,50,51	0
7	PEG	A	1626	7/7	0.84	0.18	2.78	34,35,37,38	0
8	PG4	A	1625	10/13	0.91	0.12	1.08	47,47,49,49	0
9	3EF	B	1630	51/51	0.95	0.11	0.70	18,23,28,30	0
9	3EF	A	1630	51/51	0.95	0.11	0.18	19,25,30,30	0
7	PEG	A	1624	7/7	0.87	0.10	-0.69	44,45,45,45	0
3	CL	A	1002	1/1	1.00	0.08	-1.18	19,19,19,19	0
3	CL	B	1003	1/1	0.99	0.07	-1.21	24,24,24,24	0
7	PEG	B	1621	7/7	0.89	0.13	-	42,43,46,46	0
2	ZN	A	1001	1/1	1.00	0.09	-	18,18,18,18	0
7	PEG	A	1622	7/7	0.92	0.17	-	41,41,43,43	0
2	ZN	B	1001	1/1	1.00	0.06	-	16,16,16,16	0

## 6.5 Other polymers

There are no such residues in this entry.