Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 03:42 PM GMT

PDB ID	:	4D52
Title	:	CRYSTAL STRUCTURE OF FUCOSE BINDING LECTIN FROM ASPERGILLUS FUMIGATUS (AFL) IN COMPLEX WITH L-GALACTOPYRANOSE.
Authors	:	Houser, J.; Cioci, G.; Komarek, J.; Wimmerowa, M.; Kostlanova, N.; Lahmann, M.; Varrot, A.; Imberty, A.
Deposited on	:	2014-11-01
Resolution	:	1.76 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

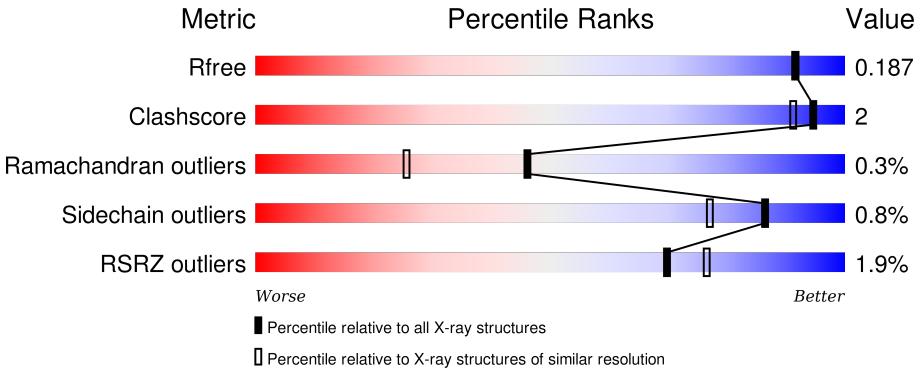
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.76 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	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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 >=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 <=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	315	2% 95% 5%
2	B	315	% 93% 6%
2	D	315	4% 97% .
3	C	315	% 95% . .

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
11	PGE	B	1324	-	-	X	-
5	PEG	A	1321	-	-	-	X
5	PEG	C	1320	-	-	-	X
5	PEG	D	1323	-	-	-	X
6	GXL	D	1320[A]	-	-	-	X
8	GIV	B	2326	-	-	-	X
8	GIV	D	2320[B]	-	-	-	X

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 11303 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 FUCOSE-SPECIFIC LECTIN FLEA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	314	Total	C	N	O	S	0	1	0
			2449	1560	423	461	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	SER	LEU	CONFLICT	UNP Q4WW81
A	111	CYS	ARG	CONFLICT	UNP Q4WW81

- Molecule 2 is a protein called FUCOSE-SPECIFIC LECTIN FLEA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	314	Total	C	N	O	S	0	3	0
			2455	1564	421	465	5			
2	D	314	Total	C	N	O	S	0	1	0
			2446	1558	420	464	4			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	SER	LEU	CONFLICT	UNP Q4WW81
B	111	CSD	ARG	CONFLICT	UNP Q4WW81
D	20	SER	LEU	CONFLICT	UNP Q4WW81
D	111	CSD	ARG	CONFLICT	UNP Q4WW81

- Molecule 3 is a protein called FUCOSE-SPECIFIC LECTIN FLEA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	312	Total	C	N	O	S	0	2	0
			2439	1552	419	464	4			

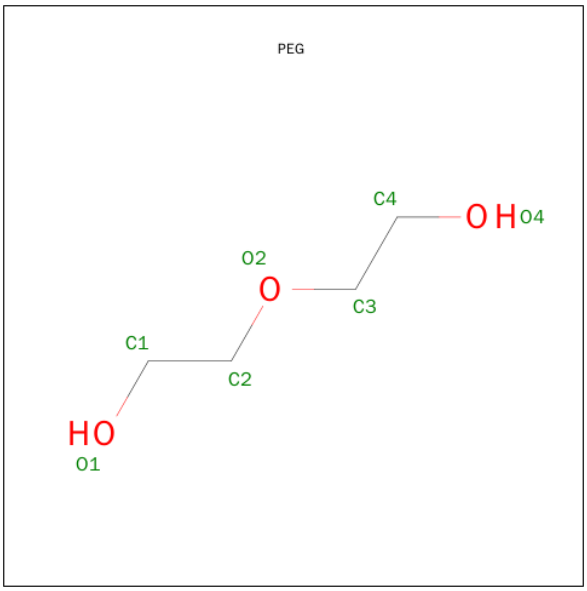
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	20	SER	LEU	CONFLICT	UNP Q4WW81
C	111	CSD	ARG	CONFLICT	UNP Q4WW81

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

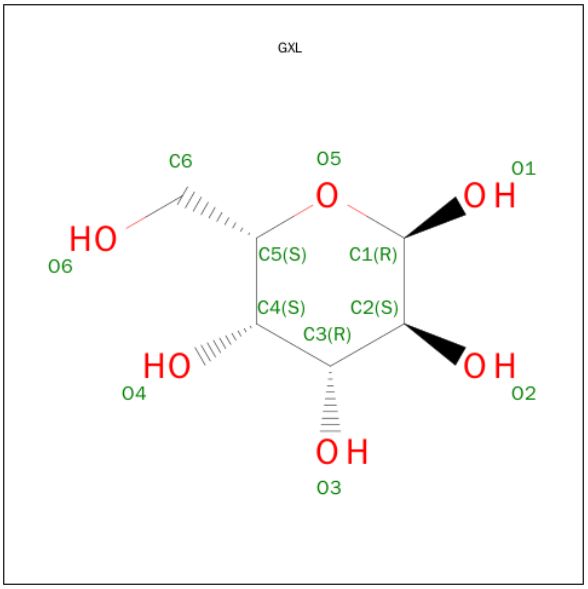
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	2	Total	C	O	0	2
			24	12	12		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			7	4	3		
5	C	1	Total	C	O	0	0
			7	4	3		
5	C	1	Total	C	O	0	0
			7	4	3		
5	D	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is SUGAR (ALPHA-L-GALACTOPYRANOSE) (three-letter code: GXL) (formula: $C_6H_{12}O_6$).



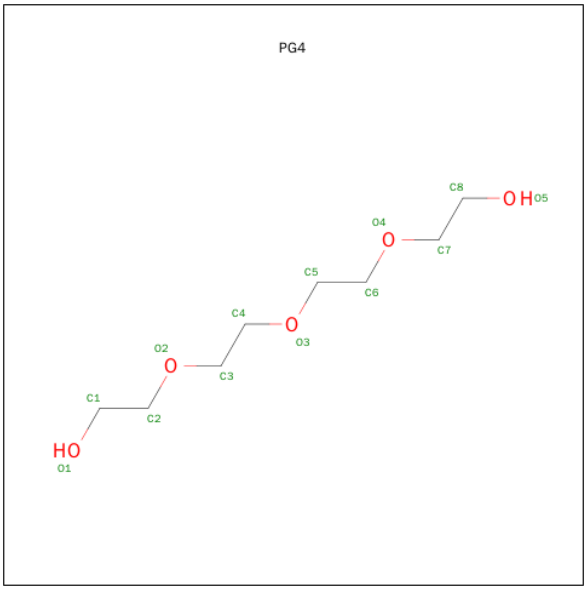
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 12 6 6	0	0
6	A	1	Total C O 12 6 6	0	1
6	A	1	Total C O 12 6 6	0	1
6	B	1	Total C O 12 6 6	0	1
6	B	1	Total C O 12 6 6	0	0
6	B	1	Total C O 12 6 6	0	1
6	C	1	Total C O 12 6 6	0	1
6	C	1	Total C O 12 6 6	0	0
6	C	1	Total C O 12 6 6	0	1
6	C	1	Total C O 12 6 6	0	1
6	D	1	Total C O 12 6 6	0	0
6	D	1	Total C O 12 6 6	0	1

Continued on next page...

Continued from previous page...

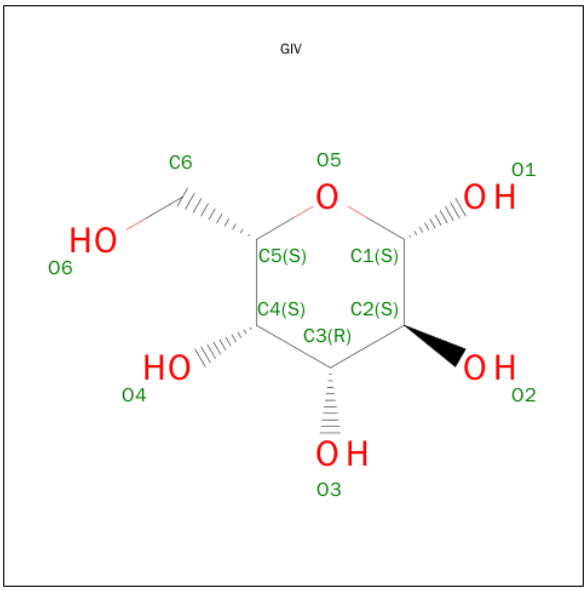
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	1	Total C O 12 6 6	0	1

- Molecule 7 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 13 8 5	0	0
7	B	1	Total C O 13 8 5	0	0
7	C	1	Total C O 13 8 5	0	0
7	C	1	Total C O 13 8 5	0	0
7	D	1	Total C O 13 8 5	0	0

- Molecule 8 is SUGAR (BETA-L-GALACTOPYRANOSE) (three-letter code: GIV) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 12 6 6	0	1
8	A	1	Total C O 12 6 6	0	1
8	B	1	Total C O 12 6 6	0	1
8	B	1	Total C O 12 6 6	0	1
8	B	1	Total C O 12 6 6	0	0
8	C	1	Total C O 12 6 6	0	1
8	C	1	Total C O 12 6 6	0	1
8	C	1	Total C O 12 6 6	0	1
8	D	1	Total C O 12 6 6	0	0
8	D	1	Total C O 12 6 6	0	1
8	D	1	Total C O 12 6 6	0	1

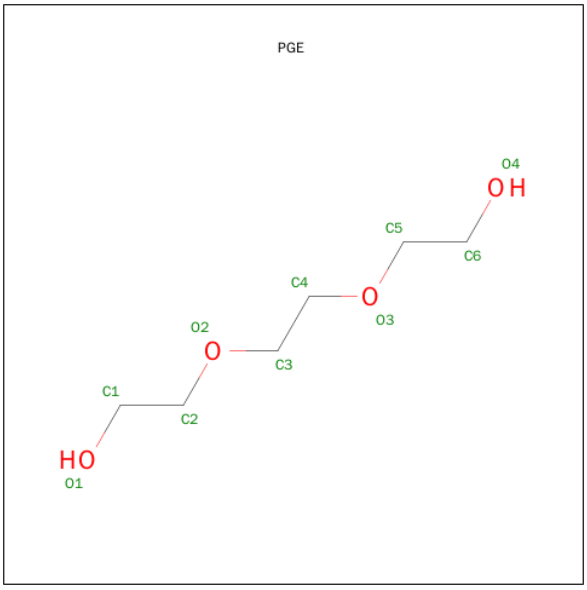
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	1	Total Zn 1 1	0	0

- Molecule 10 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	B	1	Total Na 1 1	0	0
10	D	1	Total Na 1 1	0	0
10	C	1	Total Na 1 1	0	0

- Molecule 11 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	B	1	Total C O 10 6 4	0	0
11	C	1	Total C O 10 6 4	0	0
11	D	1	Total C O 10 6 4	0	0

- Molecule 12 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	345	Total O 345 345	0	0
12	B	242	Total O 242 242	0	0
12	C	253	Total O 253 253	0	0

Continued on next page...

Continued from previous page...

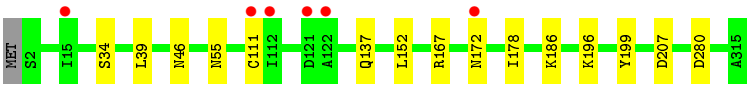
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	D	220	Total	O	0	0
			220	220		

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 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 (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: FUCOSE-SPECIFIC LECTIN FLEA

Chain A:

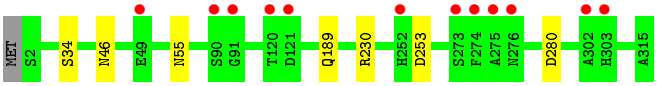


- Molecule 2: FUCOSE-SPECIFIC LECTIN FLEA

Chain B:

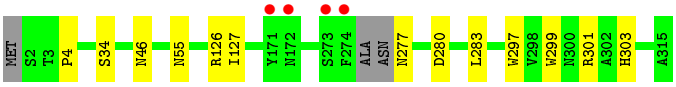
- Molecule 2: FUCOSE-SPECIFIC LECTIN FLEA

Chain D:



- Molecule 3: FUCOSE-SPECIFIC LECTIN FLEA

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	80.03Å 70.44Å 117.80Å 90.00° 108.34° 90.00°	Depositor
Resolution (Å)	111.82 – 1.76 33.59 – 1.76	Depositor EDS
% Data completeness (in resolution range)	97.3 (111.82-1.76) 97.3 (33.59-1.76)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.66 (at 1.76Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.154 , 0.178 0.166 , 0.187	Depositor DCC
R_{free} test set	6027 reflections (5.28%)	DCC
Wilson B-factor (Å^2)	16.0	Xtriage
Anisotropy	0.121	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 44.7	EDS
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 121067 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11303	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.86% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GIV, ZN, PGE, GXL, CSD, PG4, NA, 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	0.52	0/2520	0.70	0/3435
2	B	0.52	0/2527	0.70	2/3445 (0.1%)
2	D	0.49	0/2512	0.70	0/3426
3	C	0.50	0/2496	0.70	0/3403
All	All	0.51	0/10055	0.70	2/13709 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	30	ASP	CB-CG-OD1	7.58	125.12	118.30
2	B	12	ARG	NE-CZ-NH2	-5.35	117.62	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	2449	0	2333	7	0
2	B	2455	0	2336	15	0
2	D	2446	0	2318	3	0
3	C	2439	0	2305	8	0
4	A	24	0	12	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	11	0	15	1	0
5	B	11	0	15	0	0
5	C	14	0	20	0	0
5	D	7	0	10	0	0
6	A	36	0	36	0	0
6	B	36	0	36	0	0
6	C	48	0	48	0	0
6	D	36	0	36	0	0
7	A	13	0	18	0	0
7	B	13	0	18	1	0
7	C	26	0	36	0	0
7	D	13	0	18	1	0
8	A	24	0	24	0	0
8	B	36	0	36	0	0
8	C	36	0	36	0	0
8	D	36	0	36	0	0
9	B	1	0	0	0	0
10	B	1	0	0	0	0
10	C	1	0	0	0	0
10	D	1	0	0	0	0
11	B	10	0	14	6	0
11	C	10	0	14	0	0
11	D	10	0	14	1	0
12	A	345	0	0	1	0
12	B	242	0	0	0	0
12	C	253	0	0	2	0
12	D	220	0	0	0	0
All	All	11303	0	9784	33	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (33) 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)
2:B:281:LEU:HD11	11:B:1324:PGE:H52	1.63	0.80
2:B:242:GLN:HE22	11:B:1324:PGE:H3	1.54	0.73
2:B:230:ARG:HE	11:B:1324:PGE:H5	1.57	0.70
3:C:277:ASN:OD1	3:C:301:ARG:HD2	1.92	0.69
2:B:230:ARG:CZ	2:B:244[A]:CYS:SG	2.88	0.62
2:B:189:GLN:NE2	2:B:251:TYR:OH	2.39	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:140:MET:HE2	2:B:147:LYS:HE2	1.89	0.54
1:A:196:LYS:HD2	1:A:199:TYR:CE2	2.43	0.54
2:B:220:PHE:CZ	2:B:279:PRO:HG3	2.46	0.51
2:D:230:ARG:HE	11:D:1322:PGE:H62	1.77	0.50
2:B:230:ARG:NH1	2:B:244[A]:CYS:SG	2.86	0.49
1:A:167:ARG:HG2	1:A:167:ARG:HH11	1.79	0.47
2:B:281:LEU:CD1	11:B:1324:PGE:H52	2.40	0.47
3:C:301:ARG:HG2	12:C:3222:HOH:O	2.14	0.47
3:C:34:SER:OG	3:C:55:ASN:ND2	2.48	0.47
2:B:189:GLN:HE22	7:B:1323:PG4:H42	1.80	0.47
3:C:303:HIS:HB2	12:C:3234:HOH:O	2.15	0.46
2:D:189:GLN:HE22	7:D:1321:PG4:H42	1.81	0.45
3:C:297:TRP:CH2	3:C:299:TRP:HB2	2.51	0.45
3:C:283:LEU:HD12	3:C:297:TRP:CE3	2.53	0.44
1:A:186:LYS:HG2	1:A:207:ASP:O	2.17	0.44
2:B:230:ARG:HE	11:B:1324:PGE:C5	2.27	0.44
2:D:34:SER:OG	2:D:55:ASN:ND2	2.51	0.44
2:B:192:TYR:CE2	2:B:197:GLY:HA2	2.54	0.43
5:A:1317:PEG:H12	12:A:3339:HOH:O	2.18	0.42
1:A:167:ARG:NH2	1:A:172:ASN:HA	2.35	0.42
1:A:137:GLN:OE1	3:C:4:PRO:HG2	2.19	0.42
3:C:126:ARG:C	3:C:127:ILE:HD13	2.40	0.42
1:A:152:LEU:HD13	1:A:178:ILE:HD11	2.02	0.41
2:B:230:ARG:NE	11:B:1324:PGE:H5	2.30	0.41
2:B:171:TYR:HE2	2:B:195:HIS:CE1	2.39	0.40
1:A:34:SER:OG	1:A:55:ASN:ND2	2.55	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	313/315 (99%)	305 (97%)	7 (2%)	1 (0%)	46 25
2	B	314/315 (100%)	307 (98%)	6 (2%)	1 (0%)	46 25
2	D	312/315 (99%)	304 (97%)	7 (2%)	1 (0%)	46 25
3	C	308/315 (98%)	300 (97%)	7 (2%)	1 (0%)	46 25
All	All	1247/1260 (99%)	1216 (98%)	27 (2%)	4 (0%)	46 25

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	46	ASN
2	B	46	ASN
3	C	46	ASN
2	D	46	ASN

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	250/250 (100%)	246 (98%)	4 (2%)	70 52
2	B	250/249 (100%)	248 (99%)	2 (1%)	86 77
2	D	247/249 (99%)	245 (99%)	2 (1%)	86 77
3	C	246/248 (99%)	245 (100%)	1 (0%)	93 90
All	All	993/996 (100%)	984 (99%)	9 (1%)	86 72

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

Mol	Chain	Res	Type
1	A	39	LEU
1	A	111[A]	CYS
1	A	111[B]	CYS
1	A	280	ASP
2	B	127	ILE
2	B	280	ASP
3	C	280	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	253	ASP
2	D	280	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	46	ASN
1	A	55	ASN
2	B	46	ASN
2	B	55	ASN
2	B	105	GLN
2	B	189	GLN
2	B	242	GLN
3	C	46	ASN
3	C	55	ASN
2	D	46	ASN
2	D	55	ASN
2	D	189	GLN
2	D	303	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
2	CSD	B	111	2	3,7,8	1.45	1 (33%)	3,8,10	2.68	1 (33%)
3	CSD	C	111	3	3,7,8	1.49	1 (33%)	3,8,10	1.26	0
3	CSD	C	244	3	3,7,8	1.31	1 (33%)	3,8,10	1.93	2 (66%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	CSD	D	111	2	3,7,8	1.20	0	3,8,10	1.69	1 (33%)

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
2	CSD	B	111	2	-	0/2/6/8	0/0/0/0
3	CSD	C	111	3	-	0/2/6/8	0/0/0/0
3	CSD	C	244	3	-	0/2/6/8	0/0/0/0
2	CSD	D	111	2	-	0/2/6/8	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	111	CSD	CB-SG	-2.35	1.65	1.79
2	B	111	CSD	CB-SG	-2.21	1.66	1.79
3	C	244	CSD	CB-SG	-2.10	1.67	1.79

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	111	CSD	OD1-SG-CB	-4.24	98.33	105.40
3	C	244	CSD	CB-CA-C	-2.52	104.56	111.46
3	C	244	CSD	O-C-CA	-2.20	119.77	125.49
2	D	111	CSD	OD1-SG-CB	2.45	109.48	105.40

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)

2 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	GXL	A	1316[A]	-	12,12,12	0.64	0	17,17,17	0.97	1 (5%)
4	GIV	A	2316[A]	-	12,12,12	0.61	0	17,17,17	0.97	1 (5%)

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	GXL	A	1316[A]	-	-	0/2/22/22	0/1/1/1
4	GIV	A	2316[A]	-	-	0/2/22/22	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1316[A]	GXL	O5-C5-C6	2.44	112.52	106.36
4	A	2316[A]	GIV	O5-C5-C6	2.44	112.52	106.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Ligand geometry (i)

Of 43 ligands modelled in this entry, 4 are monoatomic - leaving 39 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	PEG	A	1317	-	3,3,6	0.35	0	2,2,5	0.47	0
6	GXL	A	1318	-	12,12,12	0.65	0	17,17,17	0.61	0
6	GXL	A	1319[A]	-	12,12,12	0.53	0	17,17,17	0.80	0
7	PG4	A	1320	-	12,12,12	0.54	0	11,11,11	1.01	1 (9%)
5	PEG	A	1321	-	6,6,6	0.46	0	5,5,5	0.22	0
6	GXL	A	1322[A]	-	12,12,12	0.47	0	17,17,17	0.82	0
8	GIV	A	2319[B]	-	12,12,12	0.50	0	17,17,17	0.76	0
8	GIV	A	2322[B]	-	12,12,12	0.44	0	17,17,17	0.82	0
6	GXL	B	1319[A]	-	12,12,12	0.53	0	17,17,17	0.73	0
5	PEG	B	1320	-	3,3,6	0.41	0	2,2,5	0.01	0
6	GXL	B	1321	-	12,12,12	0.44	0	17,17,17	0.66	0
6	GXL	B	1322[A]	-	12,12,12	0.52	0	17,17,17	0.68	0
7	PG4	B	1323	-	12,12,12	0.52	0	11,11,11	0.38	0
11	PGE	B	1324	-	9,9,9	0.53	0	8,8,8	0.51	0
5	PEG	B	1325	-	6,6,6	0.41	0	5,5,5	0.32	0
8	GIV	B	2319[B]	-	12,12,12	0.52	0	17,17,17	0.73	0
8	GIV	B	2322[B]	-	12,12,12	0.49	0	17,17,17	0.72	0
8	GIV	B	2326	-	12,12,12	0.51	0	17,17,17	1.04	0
6	GXL	C	1319[A]	-	12,12,12	0.44	0	17,17,17	0.71	0
5	PEG	C	1320	-	6,6,6	0.55	0	5,5,5	0.51	0
6	GXL	C	1321	-	12,12,12	0.68	0	17,17,17	0.59	0
6	GXL	C	1322[A]	-	12,12,12	0.58	0	17,17,17	0.72	0
7	PG4	C	1323	-	12,12,12	0.50	0	11,11,11	0.37	0
7	PG4	C	1324	-	12,12,12	0.44	0	11,11,11	0.46	0
11	PGE	C	1325	-	9,9,9	0.49	0	8,8,8	0.57	0
5	PEG	C	1326	-	6,6,6	0.39	0	5,5,5	0.30	0
6	GXL	C	1327[A]	-	12,12,12	0.45	0	17,17,17	0.75	0
8	GIV	C	2319[B]	-	12,12,12	0.41	0	17,17,17	0.72	0
8	GIV	C	2322[B]	-	12,12,12	0.59	0	17,17,17	0.78	0
8	GIV	C	2327[B]	-	12,12,12	0.50	0	17,17,17	0.73	0
8	GIV	D	1318	-	12,12,12	0.50	0	17,17,17	1.20	3 (17%)
6	GXL	D	1319	-	12,12,12	0.67	0	17,17,17	0.90	1 (5%)
6	GXL	D	1320[A]	-	12,12,12	0.46	0	17,17,17	0.96	2 (11%)
7	PG4	D	1321	-	12,12,12	0.58	0	11,11,11	0.37	0
11	PGE	D	1322	-	9,9,9	0.53	0	8,8,8	0.30	0
5	PEG	D	1323	-	6,6,6	0.37	0	5,5,5	0.36	0
6	GXL	D	1324[A]	-	12,12,12	0.54	0	17,17,17	0.99	0
8	GIV	D	2320[B]	-	12,12,12	0.42	0	17,17,17	1.04	2 (11%)
8	GIV	D	2324[B]	-	12,12,12	0.52	0	17,17,17	0.96	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	PEG	A	1317	-	-	0/1/1/4	0/0/0/0
6	GXL	A	1318	-	-	0/2/22/22	0/1/1/1
6	GXL	A	1319[A]	-	-	0/2/22/22	0/1/1/1
7	PG4	A	1320	-	-	0/10/10/10	0/0/0/0
5	PEG	A	1321	-	-	0/4/4/4	0/0/0/0
6	GXL	A	1322[A]	-	-	0/2/22/22	0/1/1/1
8	GIV	A	2319[B]	-	-	0/2/22/22	0/1/1/1
8	GIV	A	2322[B]	-	-	0/2/22/22	0/1/1/1
6	GXL	B	1319[A]	-	-	0/2/22/22	0/1/1/1
5	PEG	B	1320	-	-	0/1/1/4	0/0/0/0
6	GXL	B	1321	-	-	0/2/22/22	0/1/1/1
6	GXL	B	1322[A]	-	-	0/2/22/22	0/1/1/1
7	PG4	B	1323	-	-	0/10/10/10	0/0/0/0
11	PGE	B	1324	-	-	0/7/7/7	0/0/0/0
5	PEG	B	1325	-	-	0/4/4/4	0/0/0/0
8	GIV	B	2319[B]	-	-	0/2/22/22	0/1/1/1
8	GIV	B	2322[B]	-	-	0/2/22/22	0/1/1/1
8	GIV	B	2326	-	-	0/2/22/22	0/1/1/1
6	GXL	C	1319[A]	-	-	0/2/22/22	0/1/1/1
5	PEG	C	1320	-	-	0/4/4/4	0/0/0/0
6	GXL	C	1321	-	-	0/2/22/22	0/1/1/1
6	GXL	C	1322[A]	-	-	0/2/22/22	0/1/1/1
7	PG4	C	1323	-	-	0/10/10/10	0/0/0/0
7	PG4	C	1324	-	-	0/10/10/10	0/0/0/0
11	PGE	C	1325	-	-	0/7/7/7	0/0/0/0
5	PEG	C	1326	-	-	0/4/4/4	0/0/0/0
6	GXL	C	1327[A]	-	-	0/2/22/22	0/1/1/1
8	GIV	C	2319[B]	-	-	0/2/22/22	0/1/1/1
8	GIV	C	2322[B]	-	-	0/2/22/22	0/1/1/1
8	GIV	C	2327[B]	-	-	0/2/22/22	0/1/1/1
8	GIV	D	1318	-	-	0/2/22/22	0/1/1/1
6	GXL	D	1319	-	-	0/2/22/22	0/1/1/1
6	GXL	D	1320[A]	-	-	0/2/22/22	0/1/1/1
7	PG4	D	1321	-	-	0/10/10/10	0/0/0/0
11	PGE	D	1322	-	-	0/7/7/7	0/0/0/0
5	PEG	D	1323	-	-	0/4/4/4	0/0/0/0
6	GXL	D	1324[A]	-	-	0/2/22/22	0/1/1/1
8	GIV	D	2320[B]	-	-	0/2/22/22	0/1/1/1
8	GIV	D	2324[B]	-	-	0/2/22/22	0/1/1/1

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	D	1319	GXL	C4-C3-C2	-2.27	106.55	110.79
6	D	1320[A]	GXL	O2-C2-C3	-2.12	105.57	110.34
8	D	2320[B]	GIV	O2-C2-C3	-2.12	105.57	110.34
8	D	1318	GIV	O1-C1-C2	-2.11	103.55	109.21
8	D	1318	GIV	C1-O5-C5	-2.10	109.58	113.47
8	D	1318	GIV	O3-C3-C4	2.04	114.93	110.34
6	D	1320[A]	GXL	O5-C1-C2	2.35	113.54	109.80
8	D	2320[B]	GIV	O5-C1-C2	2.35	113.54	109.80
7	A	1320	PG4	C3-O2-C2	2.67	124.79	113.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1317	PEG	1	0
7	B	1323	PG4	1	0
11	B	1324	PGE	6	0
7	D	1321	PG4	1	0
11	D	1322	PGE	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^{th} 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^2)	Q<0.9
1	A	314/315 (99%)	-0.18	6 (1%) 70 76	10, 15, 25, 40	0
2	B	313/315 (99%)	-0.15	2 (0%) 90 92	11, 15, 26, 35	0
2	D	313/315 (99%)	0.07	12 (3%) 44 50	11, 17, 31, 51	0
3	C	310/315 (98%)	-0.24	4 (1%) 79 85	10, 15, 27, 57	0
All	All	1250/1260 (99%)	-0.13	24 (1%) 70 76	10, 15, 27, 57	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	275	ALA	5.1
3	C	274	PHE	4.5
2	D	121	ASP	4.4
3	C	172	ASN	4.3
2	D	274	PHE	3.5
3	C	273	SER	3.4
2	D	273	SER	3.3
1	A	121	ASP	2.9
2	D	91	GLY	2.6
2	D	276	ASN	2.6
2	B	171	TYR	2.6
2	D	90	SER	2.5
2	D	252	HIS	2.4
2	D	120	THR	2.4
3	C	171	TYR	2.4
1	A	15	ILE	2.3
2	D	303	HIS	2.3
1	A	172	ASN	2.3
2	D	302	ALA	2.2
2	D	49	GLU	2.2
1	A	122	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	111[A]	CYS	2.1
2	B	172	ASN	2.0
1	A	112	ILE	2.0

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. 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^{th} 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^2)	Q<0.9
3	CSD	C	111	8/9	0.88	0.12	-	14,17,34,38	0
2	CSD	B	111	8/9	0.92	0.11	-	12,15,30,33	0
2	CSD	D	111	8/9	0.90	0.12	-	13,15,27,28	0
3	CSD	C	244	8/9	0.96	0.08	-	16,17,23,23	2

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^{th} 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^2)	Q<0.9
4	GXL	A	1316[A]	12/12	0.98	0.06	-1.49	11,11,11,11	12
4	GIV	A	2316[A]	12/12	0.97	0.06	-1.64	11,11,11,11	12

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. 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^{th} 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^2)	Q<0.9
8	GIV	B	2326	12/12	0.90	0.50	15.03	16,17,18,19	12
6	GXL	D	1320[A]	12/12	0.89	0.13	5.52	18,20,23,23	12
5	PEG	C	1320	7/7	0.82	0.15	5.22	26,27,33,34	0
5	PEG	A	1321	7/7	0.91	0.10	3.62	27,28,30,30	0
5	PEG	D	1323	7/7	0.91	0.16	3.34	39,39,42,46	0
8	GIV	D	2320[B]	12/12	0.90	0.13	2.24	18,20,22,23	12
7	PG4	C	1324	13/13	0.83	0.15	1.86	36,39,42,45	0
6	GXL	D	1319	12/12	0.93	0.10	1.10	18,19,22,24	0
7	PG4	C	1323	13/13	0.92	0.09	1.02	18,20,31,36	0
5	PEG	B	1325	7/7	0.87	0.12	0.78	35,37,39,40	0
7	PG4	B	1323	13/13	0.92	0.10	0.70	21,23,31,32	0
11	PGE	D	1322	10/10	0.74	0.14	0.69	30,35,41,42	0
11	PGE	B	1324	10/10	0.87	0.12	0.65	30,32,37,40	0
11	PGE	C	1325	10/10	0.84	0.12	0.64	25,29,36,39	0
5	PEG	A	1317	4/7	0.89	0.11	0.48	32,32,35,36	0
7	PG4	A	1320	13/13	0.93	0.09	0.20	17,18,29,31	0
7	PG4	D	1321	13/13	0.96	0.08	0.19	20,21,37,43	0
6	GXL	B	1322[A]	12/12	0.96	0.07	0.17	14,15,16,16	12
8	GIV	A	2319[B]	12/12	0.96	0.07	-0.30	13,15,17,18	12
6	GXL	A	1318	12/12	0.95	0.07	-0.32	13,15,17,19	0
6	GXL	A	1322[A]	12/12	0.97	0.07	-0.39	11,13,15,16	12
6	GXL	A	1319[A]	12/12	0.97	0.07	-0.45	13,15,17,18	12
6	GXL	B	1319[A]	12/12	0.96	0.08	-0.51	14,15,16,16	12
5	PEG	B	1320	4/7	0.96	0.07	-0.54	26,27,27,28	0
8	GIV	A	2322[B]	12/12	0.96	0.06	-0.56	11,13,14,16	12
6	GXL	C	1319[A]	12/12	0.97	0.06	-0.59	13,14,14,15	12
8	GIV	C	2327[B]	12/12	0.97	0.06	-0.61	13,15,17,17	12
6	GXL	D	1324[A]	12/12	0.96	0.07	-0.61	19,20,23,23	12
8	GIV	D	1318	12/12	0.95	0.08	-0.61	17,19,21,21	0
8	GIV	B	2322[B]	12/12	0.96	0.06	-0.72	14,15,16,16	12
8	GIV	C	2322[B]	12/12	0.97	0.06	-0.73	14,15,17,17	12
6	GXL	C	1322[A]	12/12	0.97	0.06	-0.74	14,16,17,17	12
8	GIV	C	2319[B]	12/12	0.97	0.06	-0.74	13,14,14,15	12
6	GXL	C	1327[A]	12/12	0.97	0.06	-0.77	13,15,17,17	12
6	GXL	B	1321	12/12	0.95	0.06	-0.77	14,15,16,17	0
6	GXL	C	1321	12/12	0.95	0.06	-0.80	15,15,17,18	1
8	GIV	B	2319[B]	12/12	0.95	0.07	-1.03	14,15,15,16	12
8	GIV	D	2324[B]	12/12	0.97	0.06	-1.47	19,20,22,23	12
9	ZN	B	1317	1/1	1.00	0.03	-2.34	16,16,16,16	1
10	NA	C	1318	1/1	0.99	0.07	-	20,20,20,20	0
10	NA	D	1317	1/1	0.98	0.14	-	15,15,15,15	0
5	PEG	C	1326	7/7	0.88	0.12	-	38,39,42,43	0
10	NA	B	1318	1/1	0.99	0.17	-	13,13,13,13	0

6.5 Other polymers (i)

There are no such residues in this entry.