wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2020 – 10:18 PM BST

PDB ID	:	7BVH
Title	:	Crystal structure of arabinosyltransferase EmbC2-AcpM2 complex from Mycobacterium smegmatis complexed with di-arabinose
Authors	:	Zhao, Y.; Zhang, L.; Wu, L.J.; Wang, Q.; Li, J.; Besra, G.S.; Rao, Z.H.
Deposited on	:	2020-04-10
Resolution	:	3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
https://www.wwpdb.org/validation/2017/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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
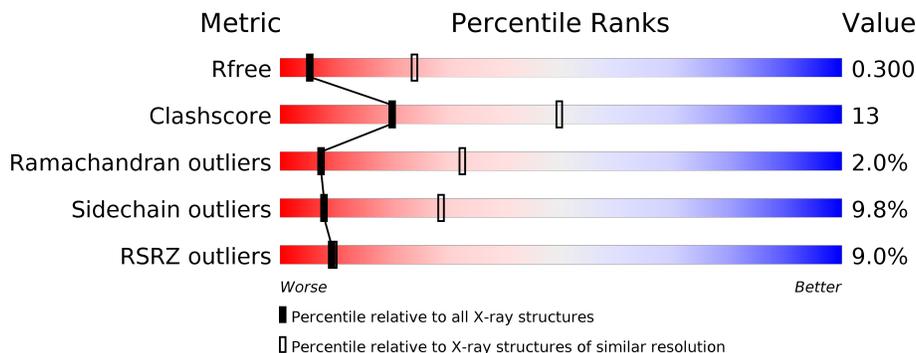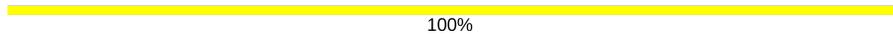

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	1113	7% / 64% / 25% / • / 7%
1	B	1113	9% / 63% / 27% / • / 7%
2	C	99	3% / 62% / 22% / • / 15%
2	D	99	12% / 68% / 14% / • / 15%
3	E	2	50% / 50%
3	J	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	2	100%
4	G	2	50% 50%
4	H	2	100%
4	I	2	100%
4	K	2	100%
4	L	2	100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GLC	G	2	-	-	X	-
4	GLC	L	2	-	-	X	-
6	PO4	B	1202	-	-	-	X
7	BXY	A	1203	-	X	-	-
7	BXY	B	1204	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 17316 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 Integral membrane indolylacetylinositol arabinosyltransferase EmbC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1035	Total	C	N	O	S	0	0	0
			7900	5101	1373	1400	26			
1	B	1035	Total	C	N	O	S	0	0	0
			7900	5101	1373	1400	26			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-12	MET	-	initiating methionine	UNP I7FMU5
A	-11	PRO	-	expression tag	UNP I7FMU5
A	-10	GLU	-	expression tag	UNP I7FMU5
A	-9	VAL	-	expression tag	UNP I7FMU5
A	-8	VAL	-	expression tag	UNP I7FMU5
A	-7	GLY	-	expression tag	UNP I7FMU5
A	-6	SER	-	expression tag	UNP I7FMU5
A	-5	TYR	-	expression tag	UNP I7FMU5
A	-4	PHE	-	expression tag	UNP I7FMU5
A	-3	GLN	-	expression tag	UNP I7FMU5
A	-2	SER	-	expression tag	UNP I7FMU5
A	-1	ASN	-	expression tag	UNP I7FMU5
A	0	ALA	-	expression tag	UNP I7FMU5
A	1075	HIS	-	expression tag	UNP I7FMU5
A	1076	LEU	-	expression tag	UNP I7FMU5
A	1077	GLY	-	expression tag	UNP I7FMU5
A	1078	GLY	-	expression tag	UNP I7FMU5
A	1079	ILE	-	expression tag	UNP I7FMU5
A	1080	LYS	-	expression tag	UNP I7FMU5
A	1081	ALA	-	expression tag	UNP I7FMU5
A	1082	PHE	-	expression tag	UNP I7FMU5
A	1083	LEU	-	expression tag	UNP I7FMU5
A	1084	GLU	-	expression tag	UNP I7FMU5
A	1085	VAL	-	expression tag	UNP I7FMU5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1086	LEU	-	expression tag	UNP I7FMU5
A	1087	PHE	-	expression tag	UNP I7FMU5
A	1088	GLN	-	expression tag	UNP I7FMU5
A	1089	GLY	-	expression tag	UNP I7FMU5
A	1090	PRO	-	expression tag	UNP I7FMU5
A	1091	HIS	-	expression tag	UNP I7FMU5
A	1092	HIS	-	expression tag	UNP I7FMU5
A	1093	HIS	-	expression tag	UNP I7FMU5
A	1094	HIS	-	expression tag	UNP I7FMU5
A	1095	HIS	-	expression tag	UNP I7FMU5
A	1096	HIS	-	expression tag	UNP I7FMU5
A	1097	HIS	-	expression tag	UNP I7FMU5
A	1098	HIS	-	expression tag	UNP I7FMU5
A	1099	HIS	-	expression tag	UNP I7FMU5
A	1100	HIS	-	expression tag	UNP I7FMU5
B	-12	MET	-	initiating methionine	UNP I7FMU5
B	-11	PRO	-	expression tag	UNP I7FMU5
B	-10	GLU	-	expression tag	UNP I7FMU5
B	-9	VAL	-	expression tag	UNP I7FMU5
B	-8	VAL	-	expression tag	UNP I7FMU5
B	-7	GLY	-	expression tag	UNP I7FMU5
B	-6	SER	-	expression tag	UNP I7FMU5
B	-5	TYR	-	expression tag	UNP I7FMU5
B	-4	PHE	-	expression tag	UNP I7FMU5
B	-3	GLN	-	expression tag	UNP I7FMU5
B	-2	SER	-	expression tag	UNP I7FMU5
B	-1	ASN	-	expression tag	UNP I7FMU5
B	0	ALA	-	expression tag	UNP I7FMU5
B	1075	HIS	-	expression tag	UNP I7FMU5
B	1076	LEU	-	expression tag	UNP I7FMU5
B	1077	GLY	-	expression tag	UNP I7FMU5
B	1078	GLY	-	expression tag	UNP I7FMU5
B	1079	ILE	-	expression tag	UNP I7FMU5
B	1080	LYS	-	expression tag	UNP I7FMU5
B	1081	ALA	-	expression tag	UNP I7FMU5
B	1082	PHE	-	expression tag	UNP I7FMU5
B	1083	LEU	-	expression tag	UNP I7FMU5
B	1084	GLU	-	expression tag	UNP I7FMU5
B	1085	VAL	-	expression tag	UNP I7FMU5
B	1086	LEU	-	expression tag	UNP I7FMU5
B	1087	PHE	-	expression tag	UNP I7FMU5
B	1088	GLN	-	expression tag	UNP I7FMU5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1089	GLY	-	expression tag	UNP I7FMU5
B	1090	PRO	-	expression tag	UNP I7FMU5
B	1091	HIS	-	expression tag	UNP I7FMU5
B	1092	HIS	-	expression tag	UNP I7FMU5
B	1093	HIS	-	expression tag	UNP I7FMU5
B	1094	HIS	-	expression tag	UNP I7FMU5
B	1095	HIS	-	expression tag	UNP I7FMU5
B	1096	HIS	-	expression tag	UNP I7FMU5
B	1097	HIS	-	expression tag	UNP I7FMU5
B	1098	HIS	-	expression tag	UNP I7FMU5
B	1099	HIS	-	expression tag	UNP I7FMU5
B	1100	HIS	-	expression tag	UNP I7FMU5

- Molecule 2 is a protein called Meromycolate extension acyl carrier protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	84	644	403	95	145	1	0	0	0
2	D	84	Total	C	N	O	S	0	0	0
			644	403	95	145	1			

- Molecule 3 is an oligosaccharide called alpha-D-arabinofuranose-(1-5)-alpha-D-arabinofuranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
3	E	2	19	10	9	0	0	0
3	J	2	Total	C	O	0	0	0
			19	10	9			

- Molecule 4 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	F	2	Total 23	C 12	O 11	0	0	0
4	G	2	Total 23	C 12	O 11	0	0	0
4	H	2	Total 23	C 12	O 11	0	0	0
4	I	2	Total 23	C 12	O 11	0	0	0
4	K	2	Total 23	C 12	O 11	0	0	0
4	L	2	Total 23	C 12	O 11	0	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total 1	Ca 1	0	0
5	A	1	Total 1	Ca 1	0	0

- Molecule 6 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P) (labeled as "Ligand of Interest" by author).

PO4

O3

O⁻

O2 O⁻ — P = O O1

P

O⁻

O4
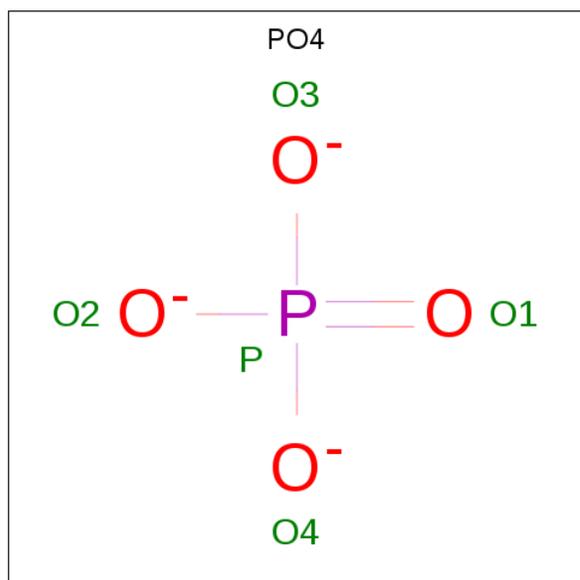
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 5	O 4	P 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total 5	O 4	P 1	0	0

• Molecule 7 is alpha-D-arabinofuranose (three-letter code: BXY) (formula: $C_5H_{10}O_5$).
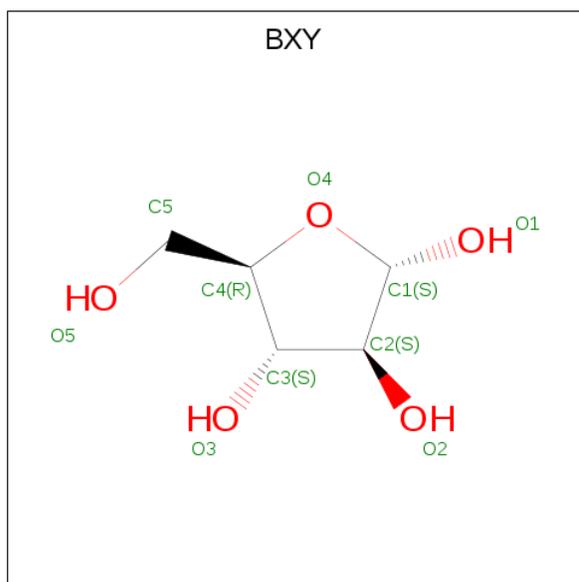
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 10	C 5	O 5	0	0
7	A	1	Total 10	C 5	O 5	0	0
7	B	1	Total 10	C 5	O 5	0	0
7	B	1	Total 10	C 5	O 5	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

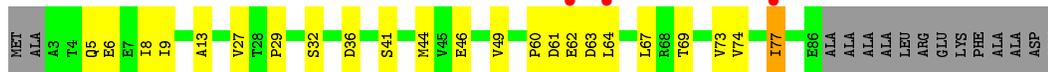- Molecule 1: Integral membrane indolylacetylinositol arabinosyltransferase EmbC
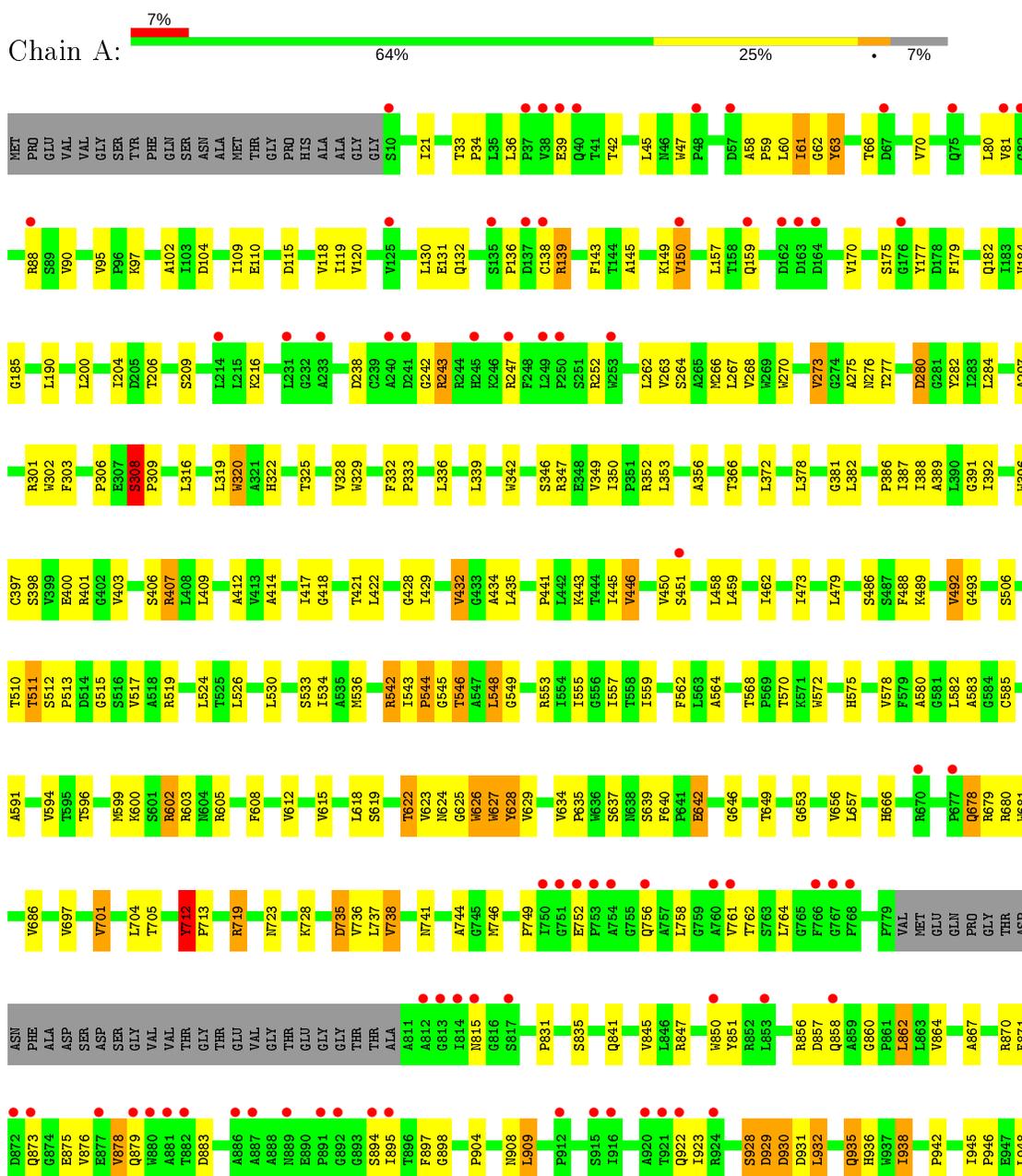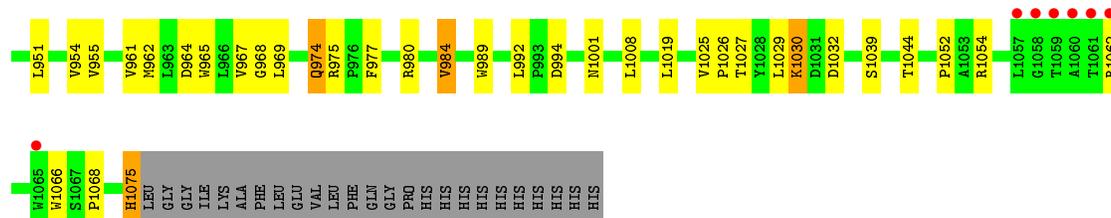
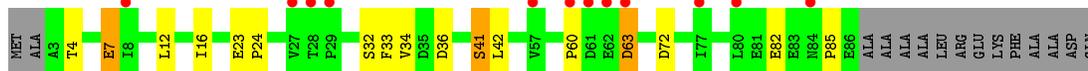• Molecule 1: Integral membrane indolylacetylinositol arabinosyltransferase EmbC
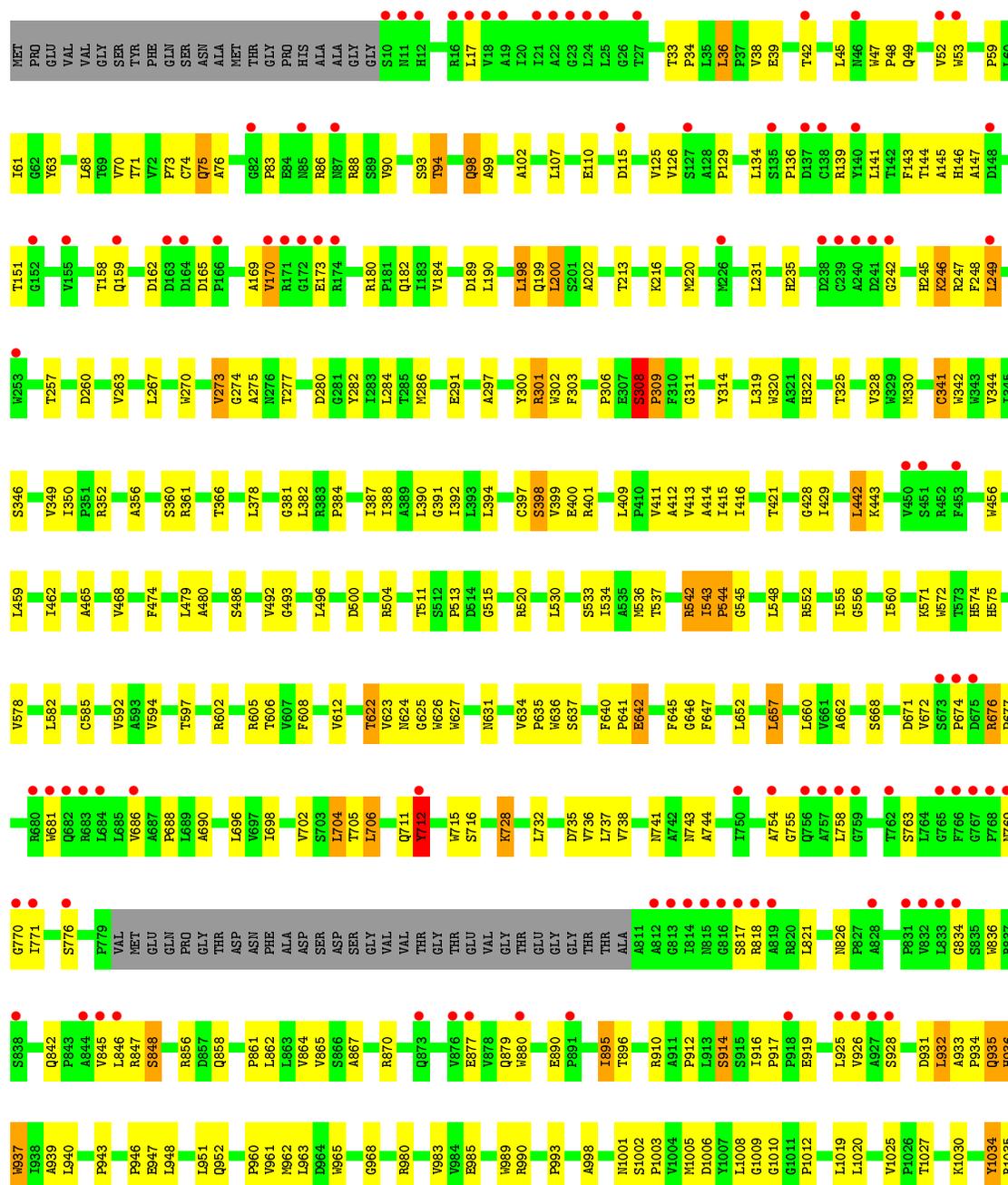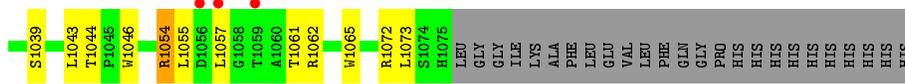
• Molecule 2: Meromycolate extension acyl carrier protein

Chain C:

• Molecule 2: Meromycolate extension acyl carrier protein

Chain D:

• Molecule 3: alpha-D-arabinofuranose-(1-5)-alpha-D-arabinofuranose

Chain E:

• Molecule 3: alpha-D-arabinofuranose-(1-5)-alpha-D-arabinofuranose

Chain J:

• Molecule 4: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain F:

• Molecule 4: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain G:

• Molecule 4: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain H:

BGC1
GLC2

● Molecule 4: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain I:

100%

BGC1
GLC2

● Molecule 4: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain K:

100%

BGC1
GLC2

● Molecule 4: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain L:

100%

BGC1
GLC2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	121.08Å 176.33Å 207.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.67 – 3.30 49.67 – 3.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.67-3.30) 100.0 (49.67-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 3.33Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.232 , 0.265 0.265 , 0.300	Depositor DCC
R_{free} test set	3362 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	33.0	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 72.9	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.78	EDS
Total number of atoms	17316	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section:
BXY, CA, GLC, BGC, PO4

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	1/8124 (0.0%)	0.75	2/11137 (0.0%)
1	B	0.50	0/8124	0.74	1/11137 (0.0%)
2	C	0.48	0/650	0.71	0/884
2	D	0.47	0/650	0.68	0/884
All	All	0.51	1/17548 (0.0%)	0.74	3/24042 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	626	TRP	CB-CG	-5.37	1.40	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	627	TRP	CA-CB-CG	5.17	123.52	113.70
1	A	627	TRP	C-N-CA	5.01	134.24	121.70
1	B	712	TYR	N-CA-C	5.01	124.52	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms
and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen
atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within
the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7900	0	7935	195	0
1	B	7900	0	7935	209	0
2	C	644	0	626	13	0
2	D	644	0	626	6	0
3	E	19	0	17	0	0
3	J	19	0	17	0	0
4	F	23	0	21	3	0
4	G	23	0	21	12	0
4	H	23	0	21	2	0
4	I	23	0	19	1	0
4	K	23	0	21	0	0
4	L	23	0	21	7	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	5	0	0	0	0
6	B	5	0	0	0	0
7	A	20	0	14	4	0
7	B	20	0	14	13	0
All	All	17316	0	17308	437	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 13.

The worst 5 of 437 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)
7:B:1204:BXY:O2	4:G:2:GLC:H3	1.31	1.24
4:I:1:BGC:O3	4:I:2:GLC:O2	1.54	1.22
1:A:841:GLN:HG3	1:A:932:LEU:O	1.36	1.21
1:B:543:ILE:HD11	1:B:592:VAL:HG12	1.31	1.09
1:A:719:ARG:HH21	1:A:719:ARG:HG2	1.12	1.07

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	1031/1113 (93%)	915 (89%)	96 (9%)	20 (2%)	8	34
1	B	1031/1113 (93%)	934 (91%)	77 (8%)	20 (2%)	8	34
2	C	82/99 (83%)	72 (88%)	7 (8%)	3 (4%)	3	20
2	D	82/99 (83%)	70 (85%)	10 (12%)	2 (2%)	6	28
All	All	2226/2424 (92%)	1991 (89%)	190 (8%)	45 (2%)	7	32

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	TYR
1	A	243	ARG
1	A	308	SER
1	A	451	SER
1	A	628	TYR

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	821/880 (93%)	741 (90%)	80 (10%)	8	29
1	B	821/880 (93%)	736 (90%)	85 (10%)	7	25
2	C	73/81 (90%)	71 (97%)	2 (3%)	44	71
2	D	73/81 (90%)	65 (89%)	8 (11%)	6	24
All	All	1788/1922 (93%)	1613 (90%)	175 (10%)	8	29

5 of 175 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1075	HIS
1	B	286	MET
1	B	1057	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	33	THR
1	B	94	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	54	GLN
1	B	235	HIS
1	B	879	GLN
1	B	113	ASN
1	B	322	HIS

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)

16 monosaccharides 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	$\#|Z| > 2$
3	BXY	E	1	3	10,10,10	0.76	0	13,14,14	1.06	1 (7%)
3	BXY	E	2	3	9,9,10	0.49	0	10,12,14	0.88	0
4	BGC	F	1	4	12,12,12	1.58	4 (33%)	17,17,17	2.34	6 (35%)
4	GLC	F	2	4	11,11,12	1.54	2 (18%)	15,15,17	2.49	7 (46%)
4	BGC	G	1	4	12,12,12	1.42	2 (16%)	17,17,17	2.38	8 (47%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z|$ > 2	Counts	RMSZ	#$|Z|$ > 2
4	GLC	G	2	4	11,11,12	1.73	3 (27%)	15,15,17	2.04	5 (33%)
4	BGC	H	1	4	12,12,12	1.29	3 (25%)	17,17,17	1.64	4 (23%)
4	GLC	H	2	4	11,11,12	2.24	3 (27%)	15,15,17	3.26	6 (40%)
4	BGC	I	1	4	12,12,12	1.33	2 (16%)	17,17,17	1.62	2 (11%)
4	GLC	I	2	4	11,11,12	2.35	3 (27%)	15,15,17	2.72	6 (40%)
3	BXY	J	1	3	10,10,10	0.70	0	13,14,14	1.39	2 (15%)
3	BXY	J	2	3	9,9,10	0.47	0	10,12,14	0.93	1 (10%)
4	BGC	K	1	4	12,12,12	1.51	3 (25%)	17,17,17	1.64	2 (11%)
4	GLC	K	2	4	11,11,12	2.24	3 (27%)	15,15,17	3.04	8 (53%)
4	BGC	L	1	4	12,12,12	1.19	1 (8%)	17,17,17	0.83	0
4	GLC	L	2	4	11,11,12	1.75	2 (18%)	15,15,17	0.90	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
3	BXY	E	1	3	-	2/2/18/18	0/1/1/1
3	BXY	E	2	3	-	0/2/15/18	0/1/1/1
4	BGC	F	1	4	-	2/2/22/22	0/1/1/1
4	GLC	F	2	4	-	0/2/19/22	0/1/1/1
4	BGC	G	1	4	-	1/2/22/22	0/1/1/1
4	GLC	G	2	4	-	0/2/19/22	0/1/1/1
4	BGC	H	1	4	-	0/2/22/22	0/1/1/1
4	GLC	H	2	4	-	1/2/19/22	0/1/1/1
4	BGC	I	1	4	-	0/2/22/22	0/1/1/1
4	GLC	I	2	4	-	2/2/19/22	0/1/1/1
3	BXY	J	1	3	-	2/2/18/18	0/1/1/1
3	BXY	J	2	3	-	2/2/15/18	0/1/1/1
4	BGC	K	1	4	-	2/2/22/22	0/1/1/1
4	GLC	K	2	4	-	0/2/19/22	0/1/1/1
4	BGC	L	1	4	-	2/2/22/22	0/1/1/1
4	GLC	L	2	4	-	1/2/19/22	0/1/1/1

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	K	2	GLC	C2-C3	-6.07	1.43	1.52
4	H	2	GLC	O5-C1	5.22	1.52	1.43
4	I	2	GLC	O5-C1	4.92	1.51	1.43
4	L	2	GLC	O5-C1	4.58	1.51	1.43
4	I	2	GLC	C2-C3	-4.40	1.46	1.52

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	K	2	GLC	C3-C4-C5	-8.51	95.06	110.24
4	I	2	GLC	C1-O5-C5	7.53	122.40	112.19
4	H	2	GLC	C2-C3-C4	-7.39	98.10	110.89
4	F	2	GLC	C3-C4-C5	-5.92	99.68	110.24
4	F	1	BGC	O4-C4-C3	-5.60	97.39	110.35

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	1	BXY	O4-C4-C5-O5
3	E	1	BXY	C3-C4-C5-O5
3	J	2	BXY	O4-C4-C5-O5
4	L	1	BGC	O5-C5-C6-O6
4	F	1	BGC	O5-C5-C6-O6

There are no ring outliers.

9 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	I	1	BGC	1	0
4	F	1	BGC	2	0
4	L	2	GLC	6	0
4	L	1	BGC	2	0
4	H	1	BGC	1	0
4	H	2	GLC	1	0
4	G	2	GLC	12	0
4	F	2	GLC	3	0
4	I	2	GLC	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.
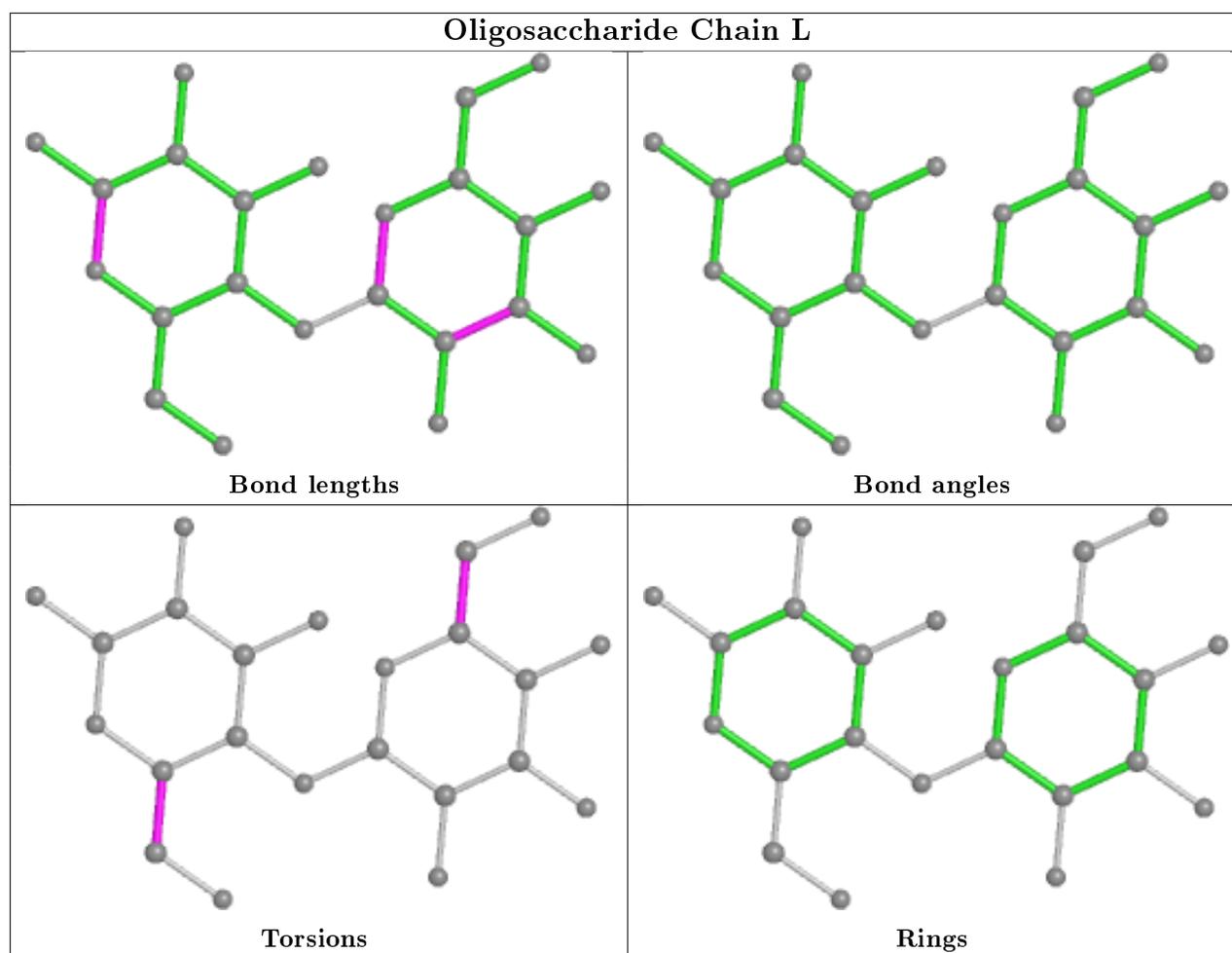
Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain J

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain G

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain H

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain I

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain K

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain L

Bond lengths

Bond angles

Torsions

Rings

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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	#$|Z| > 2$
6	PO4	A	1202	-	4,4,4	2.47	1 (25%)	6,6,6	0.56	0
7	BXY	A	1204	-	10,10,10	2.84	5 (50%)	13,14,14	1.80	3 (23%)
6	PO4	B	1202	-	4,4,4	2.50	2 (50%)	6,6,6	0.47	0
7	BXY	B	1204	-	10,10,10	3.62	7 (70%)	13,14,14	1.48	2 (15%)
7	BXY	B	1203	-	10,10,10	5.04	7 (70%)	13,14,14	2.30	5 (38%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
7	BXY	A	1203	-	10,10,10	6.39	8 (80%)	13,14,14	3.46	8 (61%)

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	BXY	A	1204	-	-	0/2/18/18	0/1/1/1
7	BXY	B	1204	-	-	2/2/18/18	0/1/1/1
7	BXY	B	1203	-	-	1/2/18/18	0/1/1/1
7	BXY	A	1203	-	-	2/2/18/18	0/1/1/1

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1203	BXY	C1-C2	-12.56	1.38	1.52
7	B	1203	BXY	C1-C2	-12.48	1.38	1.52
7	A	1203	BXY	O4-C1	-8.94	1.32	1.43
7	B	1204	BXY	C1-C2	-7.39	1.44	1.52
7	A	1203	BXY	O2-C2	-6.74	1.27	1.43

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	1203	BXY	C2-C3-C4	-6.88	89.27	102.64
7	A	1203	BXY	O1-C1-O4	-6.49	102.82	111.13
7	B	1203	BXY	C1-C2-C3	-4.68	96.44	102.30
7	A	1203	BXY	O4-C4-C5	4.59	119.13	109.21
7	A	1203	BXY	C1-C2-C3	-3.82	97.52	102.30

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	1204	BXY	O4-C4-C5-O5
7	B	1204	BXY	C3-C4-C5-O5
7	A	1203	BXY	O4-C4-C5-O5
7	A	1203	BXY	C3-C4-C5-O5
7	B	1203	BXY	O4-C4-C5-O5

There are no ring outliers.

3 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1204	BXY	2	0
7	B	1204	BXY	13	0
7	A	1203	BXY	2	0

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

Ligand PO4 A 1202

Bond lengths

Bond angles

Torsions

Rings

Ligand PO4 B 1202

Bond lengths

Bond angles

Torsions

Rings
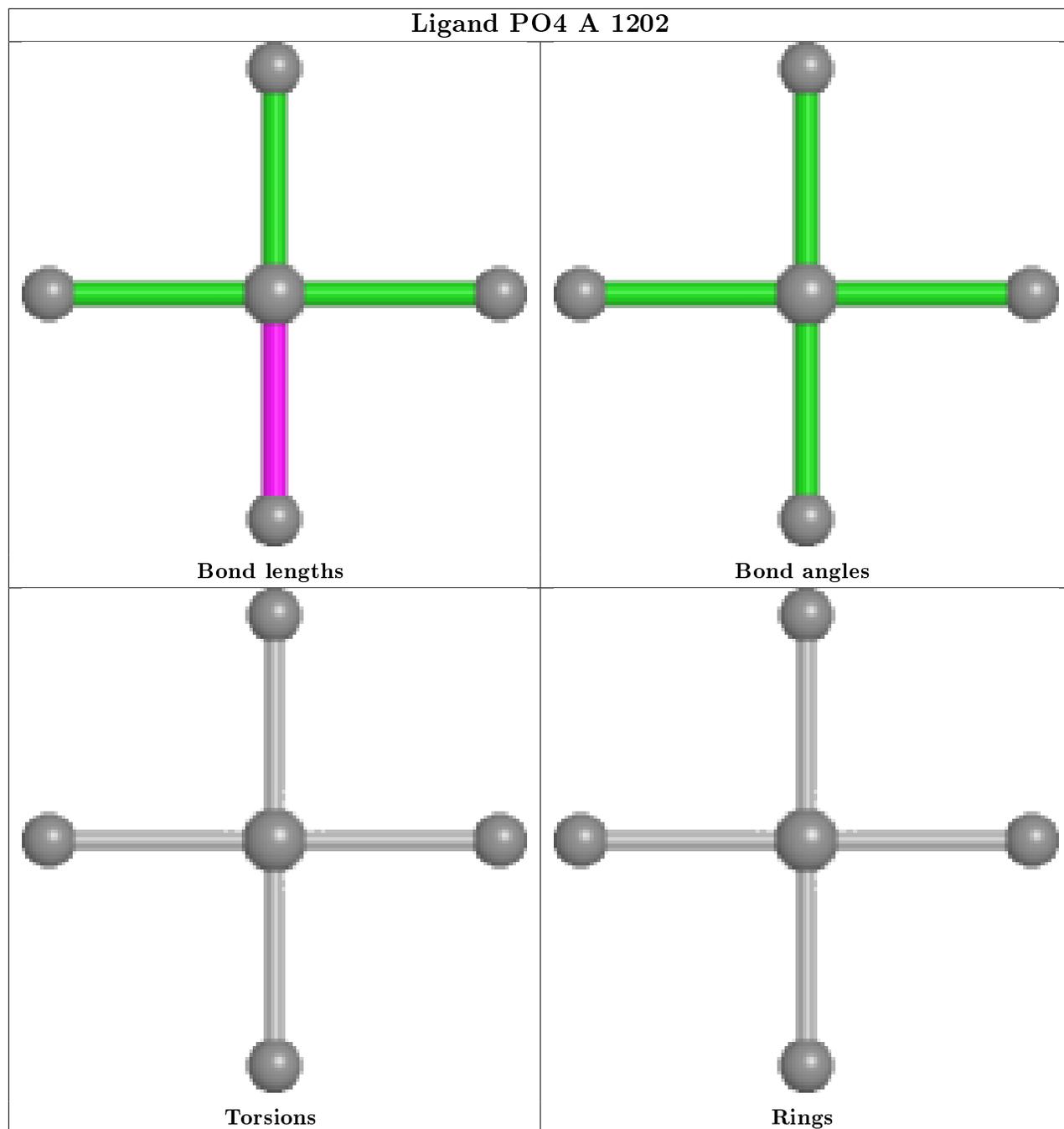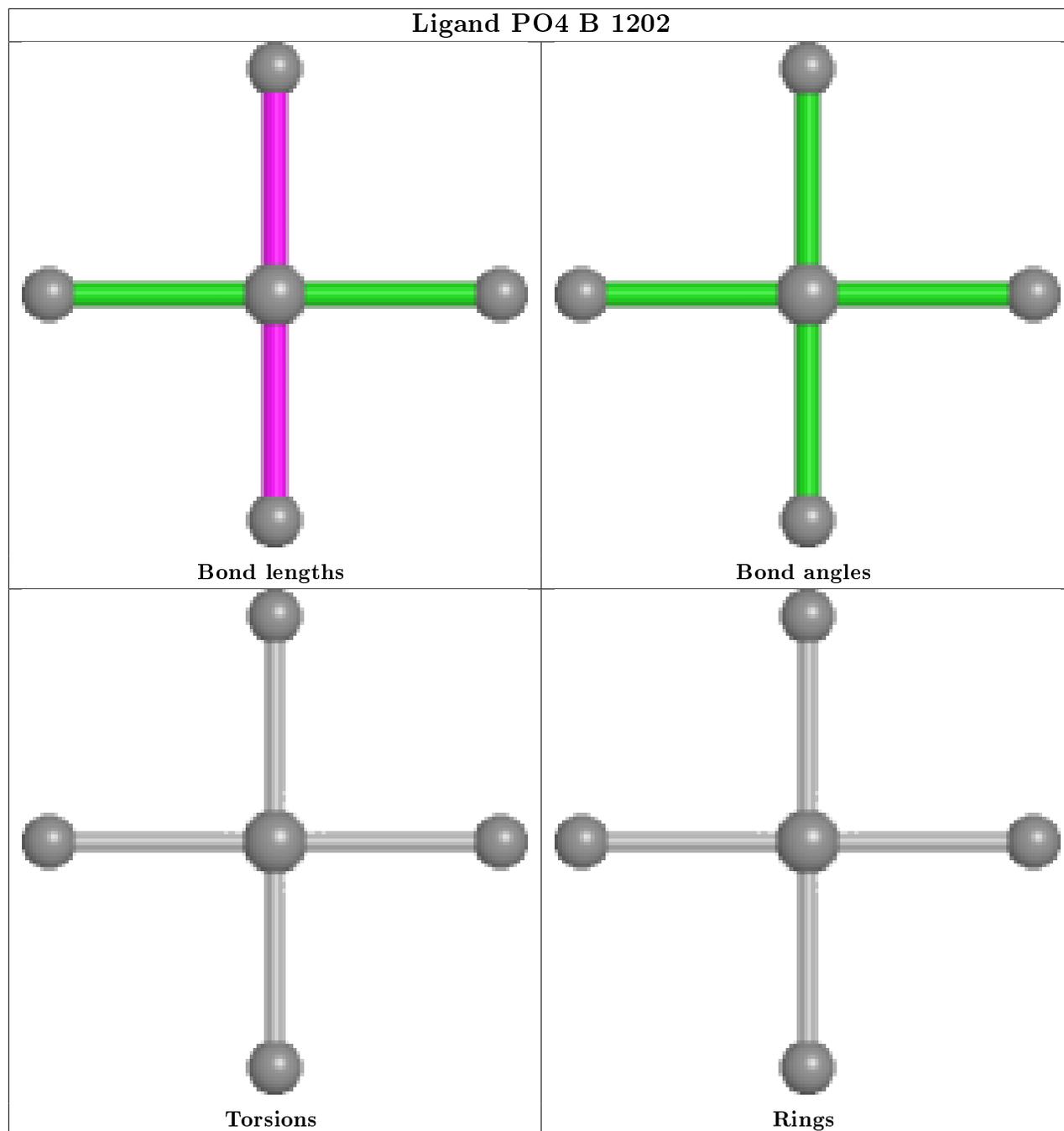
5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th 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	1035/1113 (92%)	0.20	82 (7%)	12	12	4, 68, 160, 199	0
1	B	1035/1113 (92%)	0.27	104 (10%)	7	7	5, 68, 173, 221	0
2	C	84/99 (84%)	0.41	3 (3%)	42	40	44, 82, 125, 147	0
2	D	84/99 (84%)	0.97	12 (14%)	2	2	45, 87, 126, 144	0
All	All	2238/2424 (92%)	0.27	201 (8%)	9	9	4, 71, 164, 221	0

The worst 5 of 201 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	138	CYS	10.5
1	A	922	GLN	8.4
1	A	40	GLN	6.8
1	B	10	SER	6.5
1	A	249	LEU	6.2

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. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GLC	H	2	11/12	0.73	0.24	99,102,108,109	0
4	BGC	I	1	12/12	0.74	0.39	104,106,113,116	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
4	BGC	L	1	12/12	0.82	0.57	88,95,98,99	0
3	BXY	E	1	10/10	0.82	0.28	72,77,78,79	0
3	BXY	J	1	10/10	0.82	0.51	76,82,84,86	0
4	GLC	I	2	11/12	0.84	0.35	98,101,103,103	0
4	BGC	G	1	12/12	0.85	0.21	71,79,81,81	0
4	BGC	H	1	12/12	0.86	0.20	99,107,114,116	0
3	BXY	E	2	9/10	0.87	0.25	78,79,82,85	0
3	BXY	J	2	9/10	0.88	0.33	78,80,82,84	0
4	BGC	K	1	12/12	0.88	0.19	49,59,66,66	0
4	GLC	L	2	11/12	0.89	0.37	88,90,92,94	0
4	BGC	F	1	12/12	0.91	0.18	36,40,44,45	0
4	GLC	F	2	11/12	0.93	0.19	27,29,32,33	0
4	GLC	G	2	11/12	0.93	0.14	78,80,86,87	0
4	GLC	K	2	11/12	0.95	0.18	32,36,40,41	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.
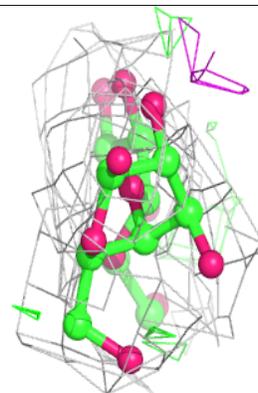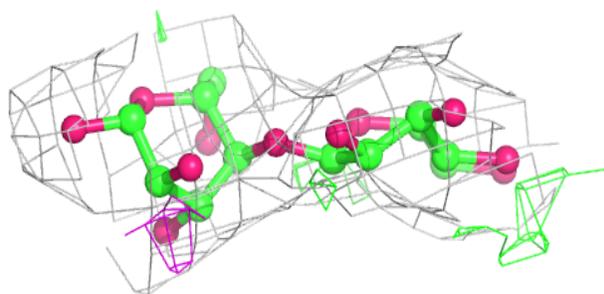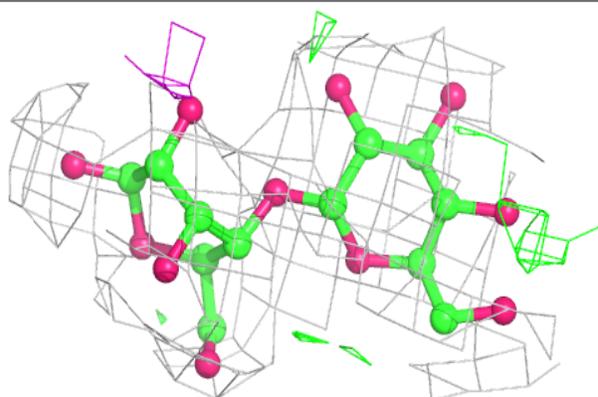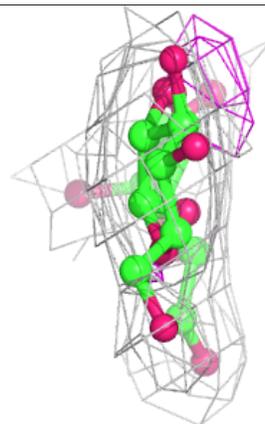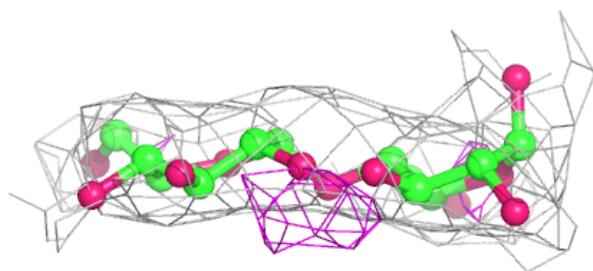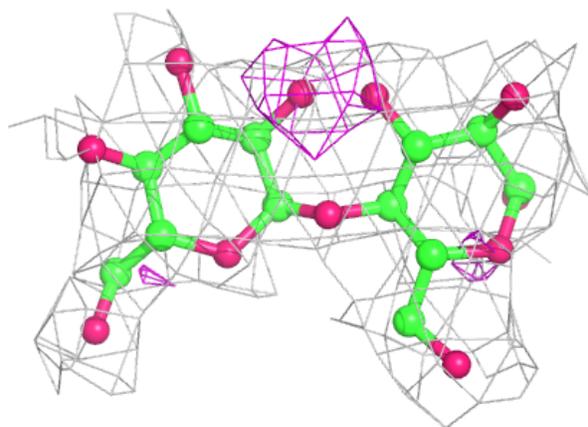

Electron density around Chain E:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
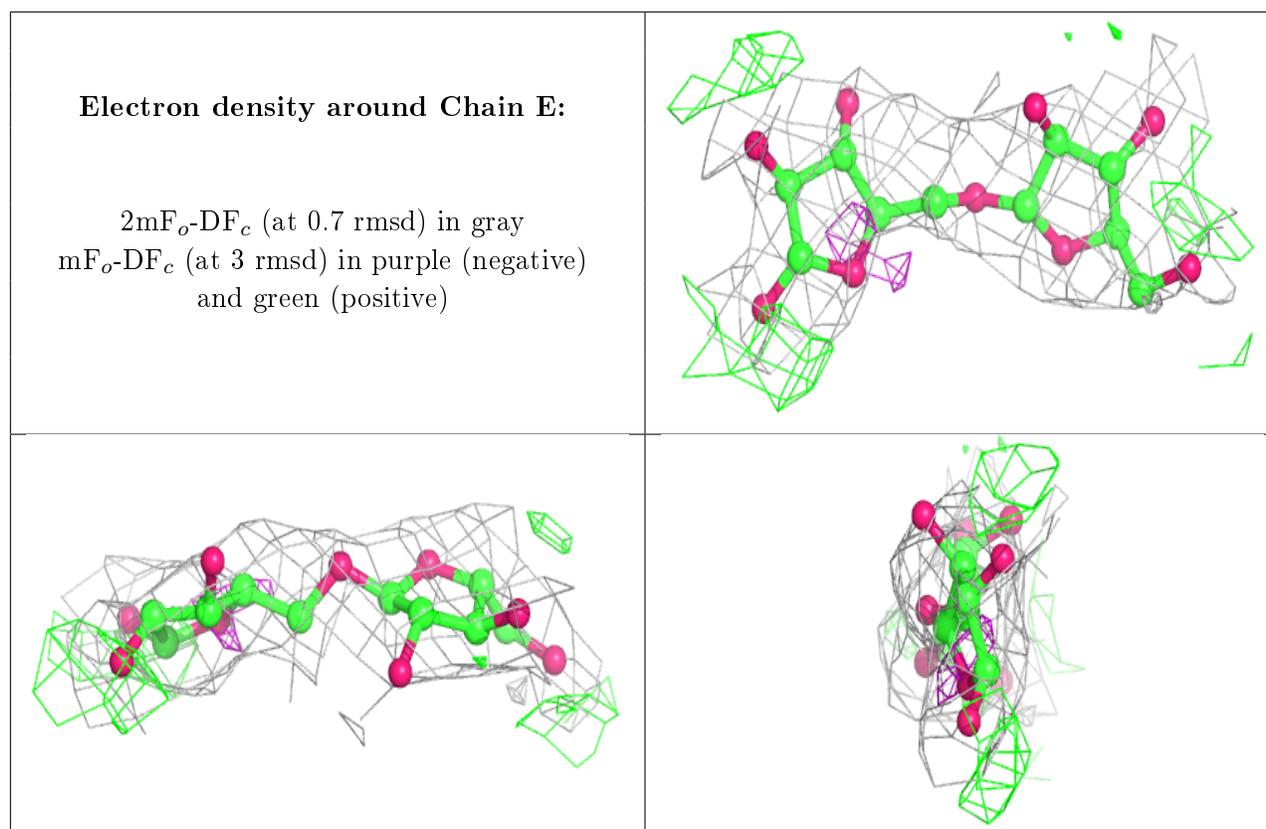
Electron density around Chain J:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
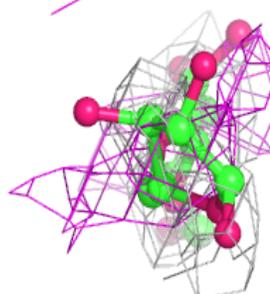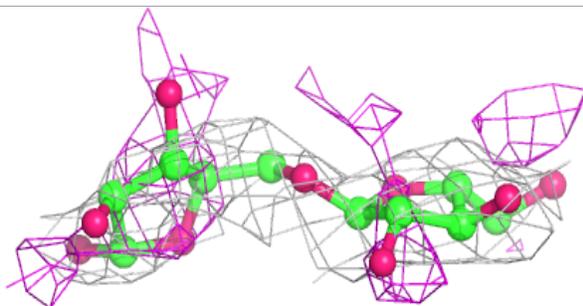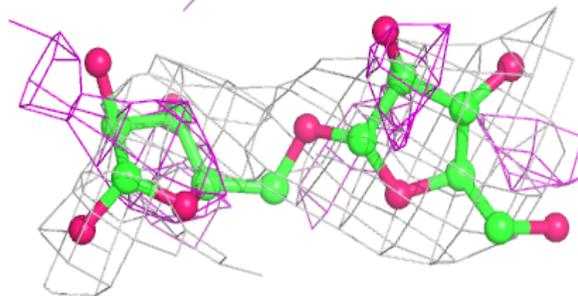
Electron density around Chain F:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
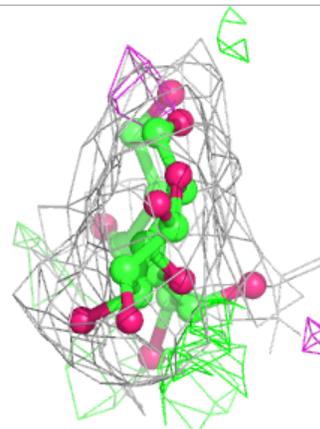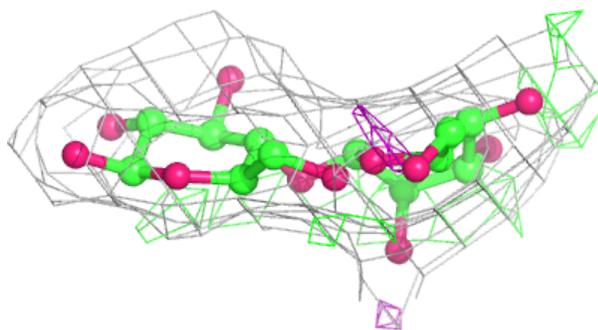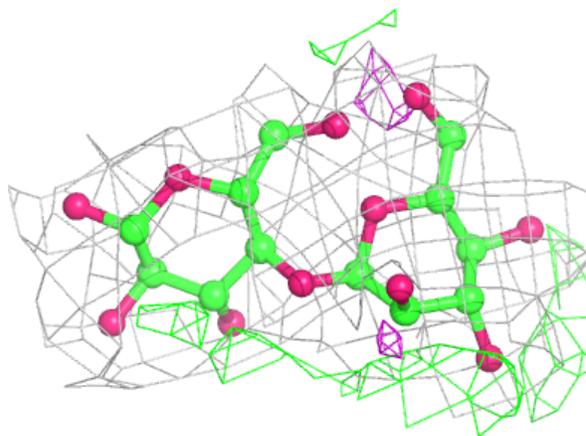
Electron density around Chain G:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
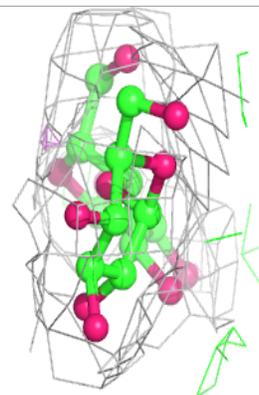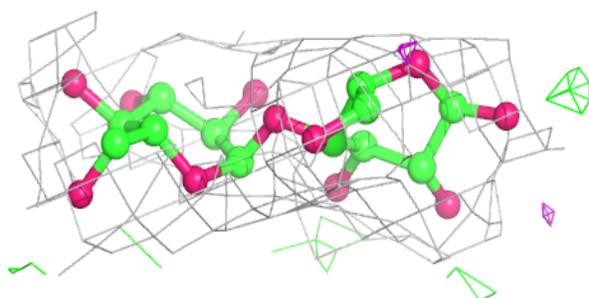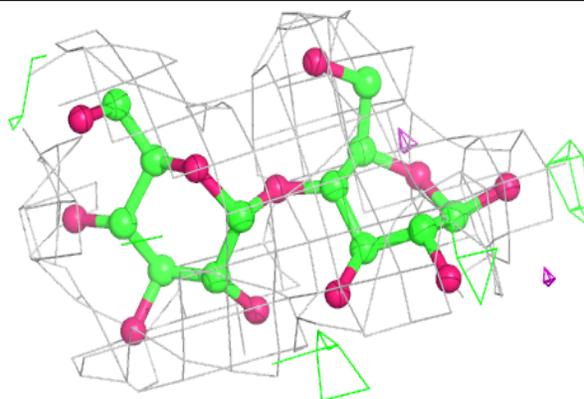
Electron density around Chain H:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around Chain I:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around Chain K:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
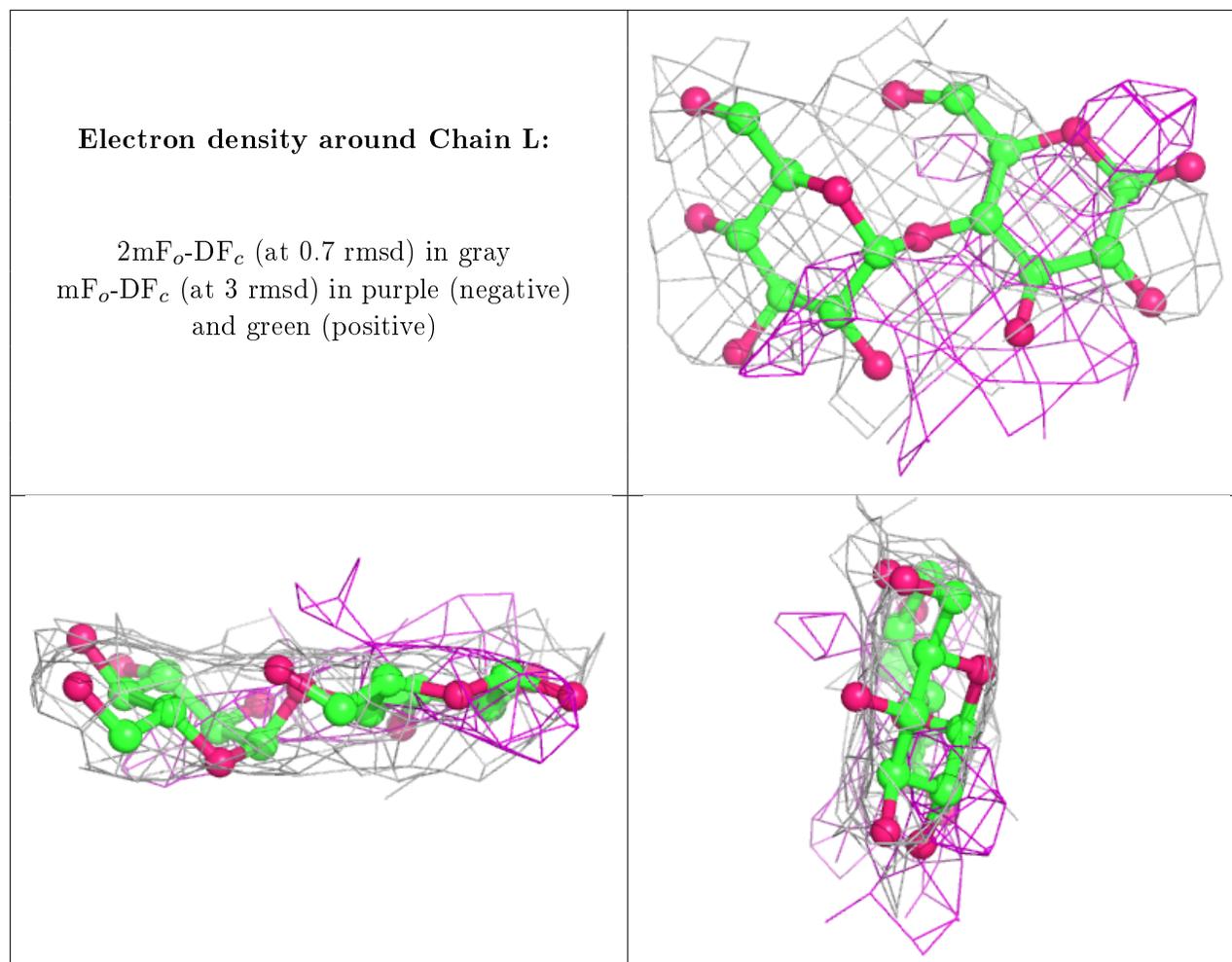and green (positive)

Electron density around Chain L:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
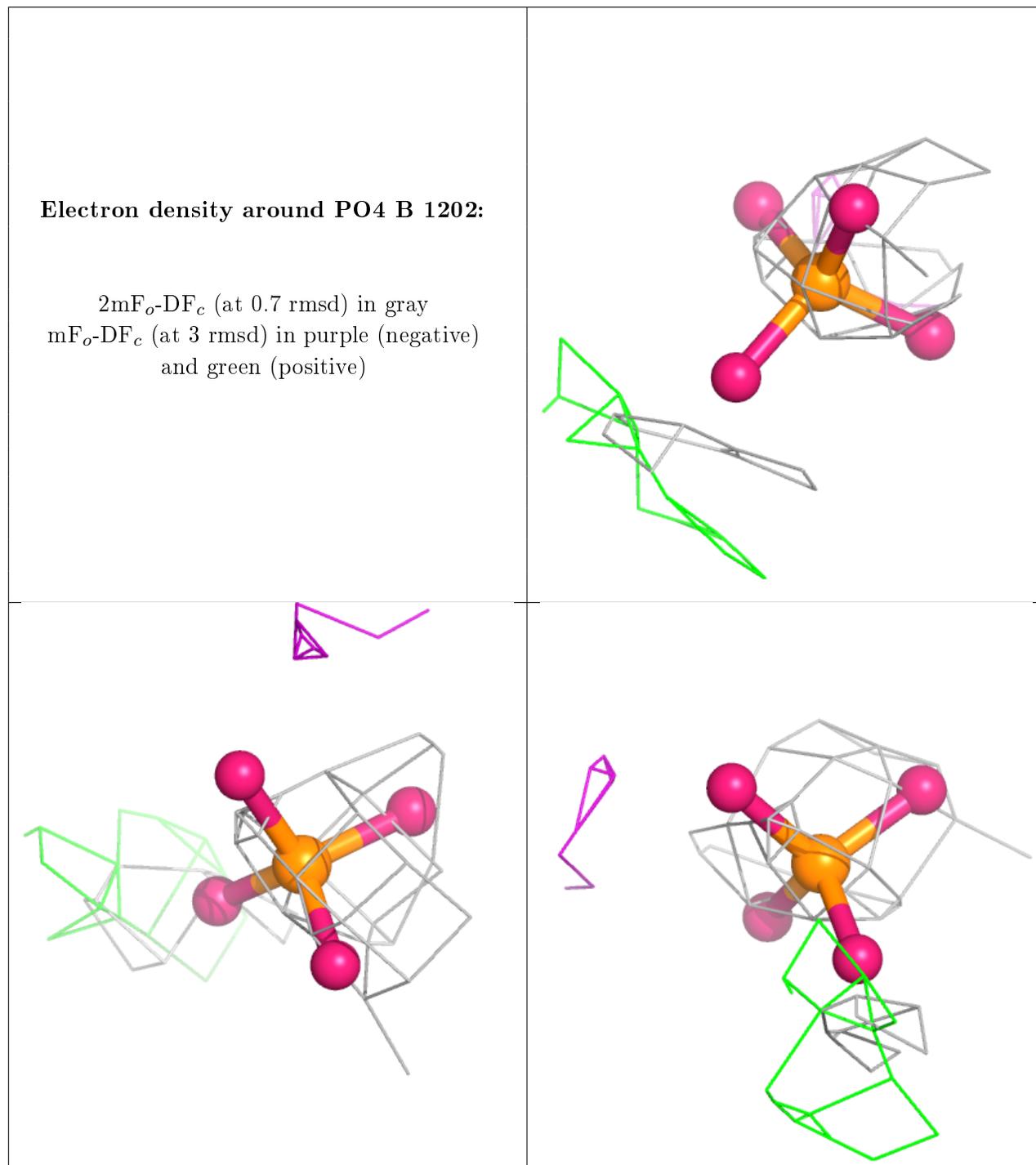mF_o-DF_c (at 3 rmsd) in purple (negative)
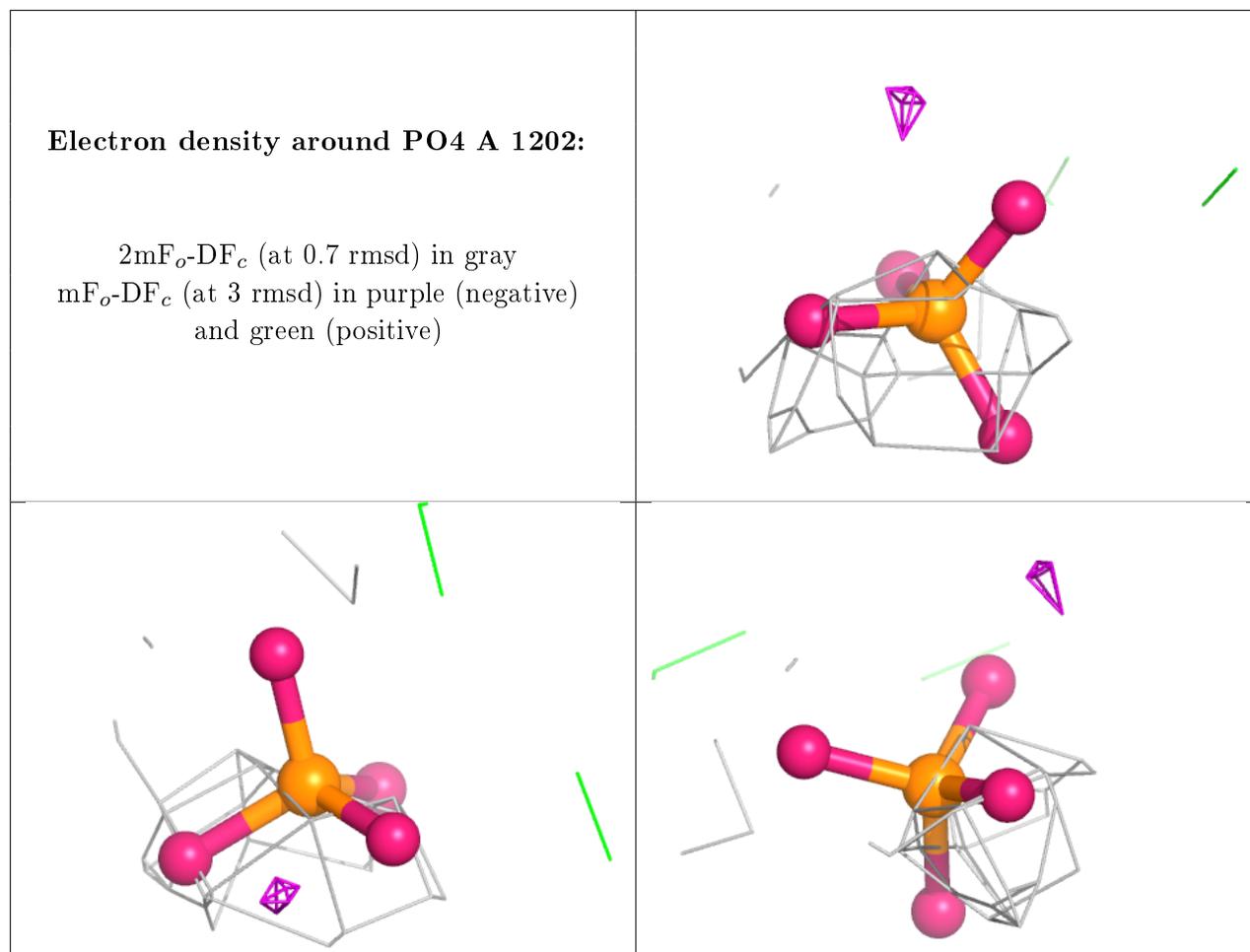and green (positive)
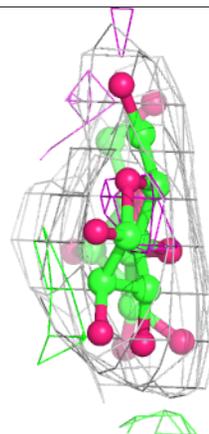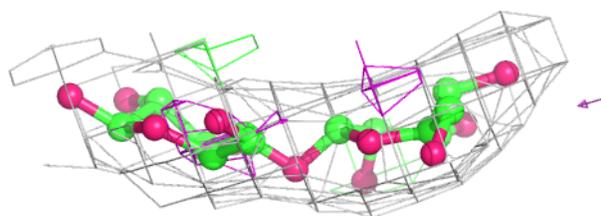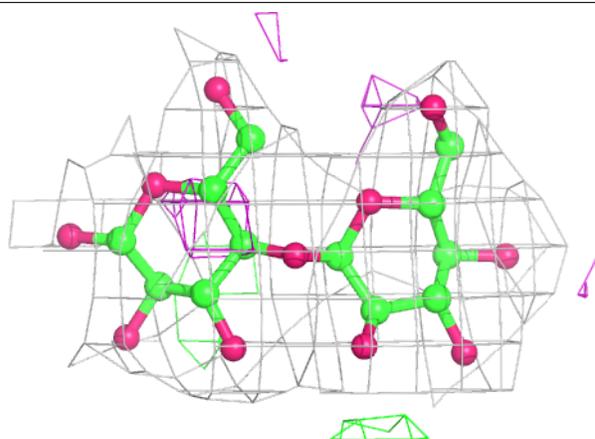
6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	CA	B	1201	1/1	0.73	0.14	109,109,109,109	0
6	PO4	B	1202	5/5	0.79	0.43	150,153,153,157	0
6	PO4	A	1202	5/5	0.84	0.33	159,159,160,162	0
7	BXY	A	1204	10/10	0.87	0.24	92,95,98,102	0
7	BXY	B	1203	10/10	0.88	0.29	82,84,85,86	0
7	BXY	A	1203	10/10	0.89	0.21	52,57,59,61	0
7	BXY	B	1204	10/10	0.93	0.16	106,111,113,114	0
5	CA	A	1201	1/1	0.94	0.15	88,88,88,88	0

The following is a graphical depiction of the model fit to experimental electron density of all

instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around PO4 B 1202:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around PO4 A 1202:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers (i)

There are no such residues in this entry.