wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 03:10 PM EDT

PDB ID	:	4HG6
Title	:	Structure of a cellulose synthase - cellulose translocation intermediate
Authors	:	Zimmer, J.
Deposited on	:	2012-10-07
Resolution	:	3.25 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

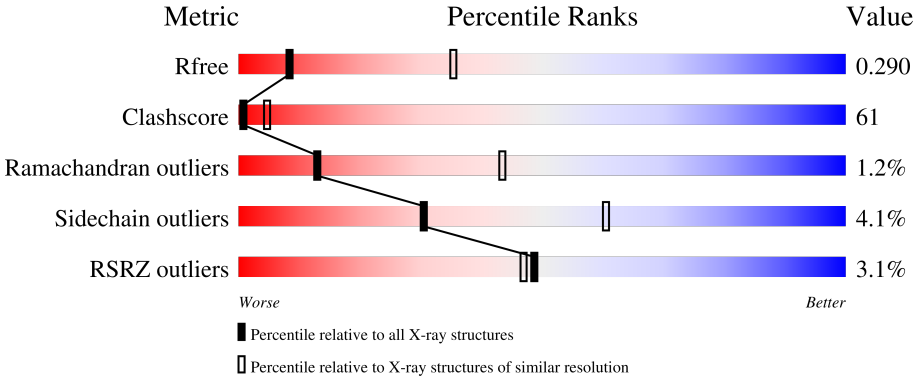
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.25 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	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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 of 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	802	3%; 40%; 51%; • ; 7%
2	B	707	3%; 42%; 45%; 5%; 7%
3	C	18	11%; 28%; 61%

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
3	BGC	C	5	-	-	X	-
4	UDP	A	919	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 11029 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 Cellulose Synthase Subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	747	5893	3822	1029	1011	31	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	789	HIS	-	expression tag	UNP Q3J125
A	790	HIS	-	expression tag	UNP Q3J125
A	791	HIS	-	expression tag	UNP Q3J125
A	792	HIS	-	expression tag	UNP Q3J125
A	793	HIS	-	expression tag	UNP Q3J125
A	794	HIS	-	expression tag	UNP Q3J125
A	795	LYS	-	expression tag	UNP Q3J125
A	796	LEU	-	expression tag	UNP Q3J125
A	797	HIS	-	expression tag	UNP Q3J125
A	798	HIS	-	expression tag	UNP Q3J125
A	799	HIS	-	expression tag	UNP Q3J125
A	800	HIS	-	expression tag	UNP Q3J125
A	801	HIS	-	expression tag	UNP Q3J125
A	802	HIS	-	expression tag	UNP Q3J125

- Molecule 2 is a protein called Cellulose Synthase Subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	655	4887	3100	864	907	16	0	0	0

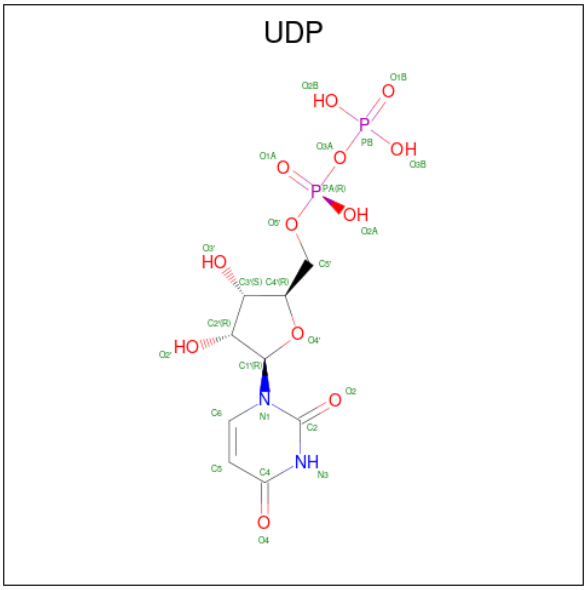
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	19	MET	-	expression tag	UNP Q3J126
B	20	GLY	-	expression tag	UNP Q3J126

- Molecule 3 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose.

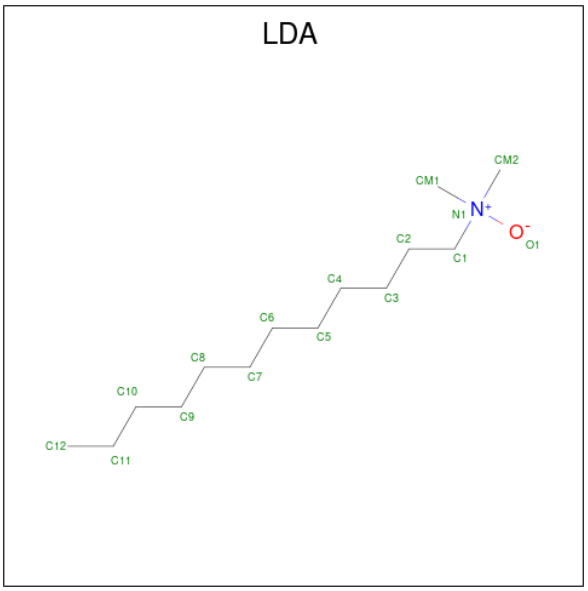
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
3	C	18	199	108	91	0	0	0

- Molecule 4 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: $C_9H_{14}N_2O_{12}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	25	9	2	12	2	0	0

- Molecule 5 is LAURYL DIMETHYLAMINE-N-OXIDE (three-letter code: LDA) (formula: $C_{14}H_{31}NO$).



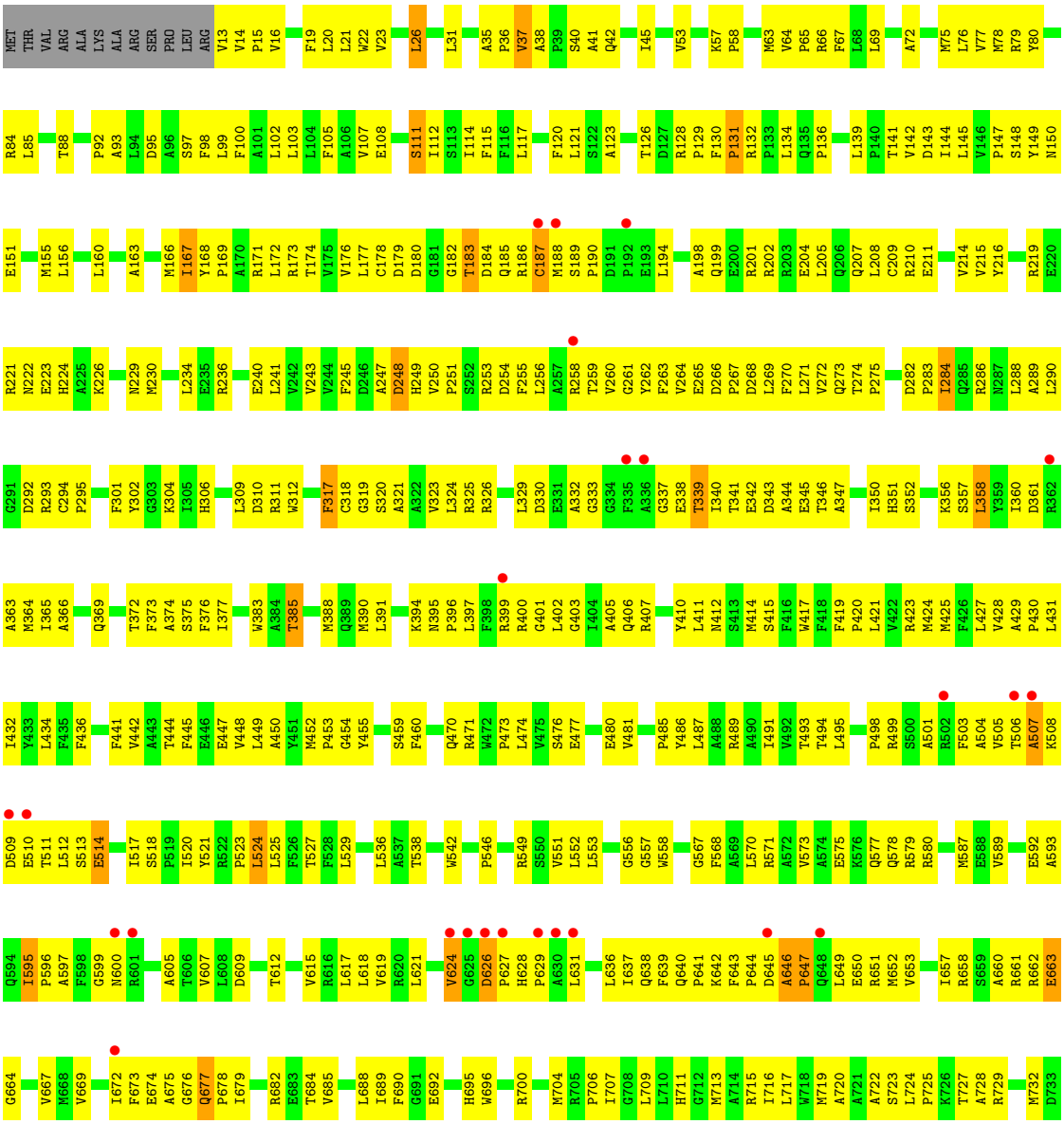
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	9	7	1	1	0	0
			Total	C	N	O		
5	A	1	16	14	1	1	0	0

3 Residue-property plots (i)

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

- Molecule 1: Cellulose Synthase Subunit A

Chain A: 3% 40% 51% • 7%



E734 P735 E743 E744 P745 H751 L752 L753 A754 F755 G756 T757 D758 F759 SER THR GLU PRO ASP TRP ALA GLY GLU LEU LEU ASP PRO THR ALA GLN VAL SER ALA ARG PRO ASN THR VAL ALA TRP GLY SER ASN HIS HIS HIS HIS HIS HIS LYS LEU HIS HIS HIS HIS HIS HIS

• Molecule 2: Cellulose Synthase Subunit B

Chain B: 3% 42% 45% 5% 7%

MET GLY GLN ASP ALA PRO MET ILE VAL ILE GLU GLY LEU THR SER GLU GLU PRO GLN ALA SER PRO ASP ALA VAL ALA GLU ALA VAL PRO ALA ALA GLU VAL ALA P54 W55 L59 A63 E64 T65 A66 Q67 V68 G69 P70 L71 F72 R73 L74 Q75 G76 Q77 R80 F83

L87 P88 T95 L96 T97 L98 A99 Q100 R101 S102 S103 I104 D105 I106 L107 P108 E109 S110 S111 Q112 I113 I114 V115 R116 M117 N118 D119 Q120 E121 I122 G123 R124 F125 T126 P127 R128 Q129 A132 L133 G134 A135 V136 T137 M138 P139 L140 G141 E142 A143 V144 R145 A146 G147 D148 N149 L150 V151 T152

I153 E154 H157 R158 H159 R160 I161 Y162 C163 G164 A165 E168 F169 D170 L171 W172 T173 D176 Q179 A183 L184 A188 I189 G190 T191 E192 P193 F196 I197 A198 T201 A202 Q203 S206 G207 R208 P209 V210 E211 I212 R213 T216 P217 P218 L223 L230 G231 R232

P233 L234 L241 S244 K245 P246 W247 S248 A249 E250 T254 I258 T259 L260 L261 P262 S263 D264 A265 D266 R267 V268 S269 I270 G273 A277 V278 V281 L282 E283 H284 P285 P286 E287 G288 S289 A292 S293 L294 V295 L298 L299 P303 T309 L310 P311 Q312 I313 P314

P315 G316 R317 V318 V319 T320 L321 A322 D323 M324 G325 V326 D327 T328 I329 L330 T331 D332 N333 R334 Y335 F336 N337 R338 D339 I340 D341 F342 Q343 L344 P345 D346 D347 W348 L349 L350 L351 A352 S353 Q354 K355 I358 G359 I360 D361 Y362 G363 F364 A365 G366 L373 L374 L375 V376 K377 V378 N379 G380

T381 T382 V383 R384 L388 D389 R390 D391 A392 A393 K396 P397 R398 I401 R402 F403 P404 A405 R406 L407 L408 H409 P410 G411 P412 N413 R414 L415 S416 S419 G423 N424 P425 P429 C430 G435 D436 L437 M438 Q439 V440 L441 S442 S443 T444 D445 P450 S451 P452 M456 A457

D458 M459 A460 R461 D462 L463 A464 Q465 V466 S470 A474 G478 L479 A480 R481 T482 L483 P484 F485 R490 D494 A495 A496 P497 V498 D499 L500 T501 V502 A503 G504 L505 H506 D507 I508 A509 T510 V511 P512 L513 N514 E515 E516 G517 L518 T519 P520 R521 L522 L523 A524 L525 T526 L527 L528

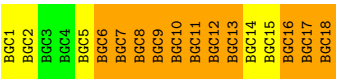
P529 S530 T531 VAL SER ARG LEU VAL GLU ARG PRO ALA THR PRO ALA G544 P545 L550 L553 G554 A555 V561 M562 P563 V566 E567 S568 N569 D572 R573 A574 Q575 T576 F577 V578 Q583 P584 V585 I586 Q587 T588 L593 R594 P595 G596 D597 L604 K608 G609 M612

L613 L614 A615 P616 E617 P618 G619 K620 L621 W622 V623 I624 L625 G626 P627 E628 A629 P642 R643 S644 P645 P648 R649 G650 A653 S663 K666 P667 G668 L669 L670 P671 E672 L673 R674 E675 S678 L679 D680 N681 V682 R683 V686 A690 S691 A692 R693 P694 P695 L696 L697

M701 L702 W706 I707 S708 A709 A710 I711 A712 V713 G714 F715 V716 L717 R718 T719 R720 ARG LYS GLY LEU LYS

• Molecule 3: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4) -beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-gluc opyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1 -4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-gl ucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain C: 11% 28% 61%


BGC1
BGC2
BGC3
BGC4
BGC5
BGC6
BGC7
BGC8
BGC9
BGC10
BGC11
BGC12
BGC13
BGC14
BGC15
BGC16
BGC17
BGC18

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	103.12Å 103.12Å 468.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.81 – 3.25 34.81 – 3.25	Depositor EDS
% Data completeness (in resolution range)	99.9 (34.81-3.25) 99.9 (34.81-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.91 (at 3.25Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.213 , 0.282 0.229 , 0.290	Depositor DCC
R_{free} test set	2055 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	117.9	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 95.4	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	0.93	EDS
Total number of atoms	11029	wwPDB-VP
Average B, all atoms (Å^2)	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: LDA, BGC, UDP

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	0.41	0/6044	0.63	2/8217 (0.0%)
2	B	0.39	1/5006 (0.0%)	0.63	2/6865 (0.0%)
All	All	0.40	1/11050 (0.0%)	0.63	4/15082 (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
2	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	345	PRO	N-CD	-10.24	1.33	1.47

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	346	ASP	CB-CA-C	-7.15	96.10	110.40
2	B	345	PRO	CA-N-CD	6.14	120.30	111.70
1	A	646	ALA	C-N-CD	-6.03	107.33	120.60
1	A	663	GLU	N-CA-C	-5.53	96.06	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	390	ARG	Sidechain

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	5893	0	5996	742	0
2	B	4887	0	4966	603	0
3	C	199	0	165	29	0
4	A	25	0	11	2	0
5	A	25	0	45	7	0
All	All	11029	0	11183	1350	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 61.

The worst 5 of 1350 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:330:LEU:CD2	2:B:439:GLN:HG3	1.33	1.58
2:B:377:LYS:CG	2:B:382:THR:HG22	1.32	1.56
1:A:643:PHE:HB3	1:A:646:ALA:CB	1.42	1.49
2:B:107:LEU:CD1	2:B:160:ARG:HA	1.47	1.41
1:A:317:PHE:CE2	1:A:347:ALA:HB2	1.58	1.38

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	745/802 (93%)	674 (90%)	64 (9%)	7 (1%)	17	50
2	B	651/707 (92%)	574 (88%)	67 (10%)	10 (2%)	10	39
All	All	1396/1509 (92%)	1248 (89%)	131 (9%)	17 (1%)	13	43

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	131	PRO
1	A	514	GLU
1	A	647	PRO
2	B	146	ALA
2	B	529	PRO

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	614/661 (93%)	595 (97%)	19 (3%)	40	67
2	B	520/559 (93%)	492 (95%)	28 (5%)	22	53
All	All	1134/1220 (93%)	1087 (96%)	47 (4%)	30	60

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	310	LEU
2	B	451	SER
2	B	312	GLN
2	B	333	ASN
2	B	500	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	67	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	337	ASN
2	B	413	ASN
2	B	343	GLN
1	A	412	ASN

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)

18 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	BGC	C	1	3	12,12,12	0.49	0	17,17,17	1.86	3 (17%)
3	BGC	C	10	3	11,11,12	0.64	0	15,15,17	1.06	1 (6%)
3	BGC	C	11	3	11,11,12	0.85	0	15,15,17	1.34	2 (13%)
3	BGC	C	12	3	11,11,12	0.99	1 (9%)	15,15,17	1.85	3 (20%)
3	BGC	C	13	3	11,11,12	0.75	0	15,15,17	1.65	4 (26%)
3	BGC	C	14	3	11,11,12	0.62	0	15,15,17	2.77	6 (40%)
3	BGC	C	15	3	11,11,12	0.58	0	15,15,17	1.37	2 (13%)
3	BGC	C	16	3	11,11,12	0.62	0	15,15,17	1.38	3 (20%)
3	BGC	C	17	3	11,11,12	0.71	0	15,15,17	2.62	5 (33%)
3	BGC	C	18	3	11,11,12	0.70	0	15,15,17	1.94	3 (20%)
3	BGC	C	2	3	11,11,12	0.66	0	15,15,17	2.16	3 (20%)
3	BGC	C	3	3	11,11,12	0.59	0	15,15,17	1.07	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	BGC	C	4	3	11,11,12	0.62	0	15,15,17	0.80	0
3	BGC	C	5	3	11,11,12	0.64	0	15,15,17	0.81	0
3	BGC	C	6	3	11,11,12	0.81	0	15,15,17	1.92	5 (33%)
3	BGC	C	7	3	11,11,12	0.74	0	15,15,17	1.81	3 (20%)
3	BGC	C	8	3	11,11,12	0.61	0	15,15,17	1.15	2 (13%)
3	BGC	C	9	3	11,11,12	0.55	0	15,15,17	1.55	3 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	BGC	C	1	3	-	2/2/22/22	0/1/1/1
3	BGC	C	10	3	-	2/2/19/22	0/1/1/1
3	BGC	C	11	3	-	0/2/19/22	0/1/1/1
3	BGC	C	12	3	-	1/2/19/22	0/1/1/1
3	BGC	C	13	3	-	1/2/19/22	0/1/1/1
3	BGC	C	14	3	-	0/2/19/22	0/1/1/1
3	BGC	C	15	3	-	0/2/19/22	0/1/1/1
3	BGC	C	16	3	-	1/2/19/22	0/1/1/1
3	BGC	C	17	3	-	1/2/19/22	0/1/1/1
3	BGC	C	18	3	-	0/2/19/22	0/1/1/1
3	BGC	C	2	3	-	2/2/19/22	0/1/1/1
3	BGC	C	3	3	-	1/2/19/22	0/1/1/1
3	BGC	C	4	3	-	2/2/19/22	0/1/1/1
3	BGC	C	5	3	-	0/2/19/22	0/1/1/1
3	BGC	C	6	3	-	0/2/19/22	0/1/1/1
3	BGC	C	7	3	-	2/2/19/22	0/1/1/1
3	BGC	C	8	3	-	2/2/19/22	0/1/1/1
3	BGC	C	9	3	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	12	BGC	O5-C5	-2.02	1.39	1.43

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	17	BGC	C1-O5-C5	8.20	123.18	112.19
3	C	14	BGC	C1-O5-C5	6.97	121.52	112.19
3	C	2	BGC	C1-O5-C5	6.72	121.19	112.19
3	C	18	BGC	C1-C2-C3	5.67	117.90	109.64
3	C	1	BGC	O5-C5-C4	5.40	119.42	109.70

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

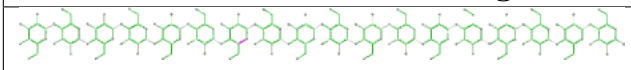
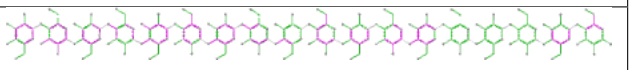
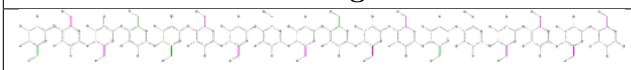
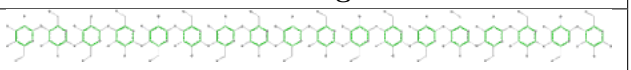
Mol	Chain	Res	Type	Atoms
3	C	7	BGC	O5-C5-C6-O6
3	C	10	BGC	O5-C5-C6-O6
3	C	7	BGC	C4-C5-C6-O6
3	C	10	BGC	C4-C5-C6-O6
3	C	2	BGC	C4-C5-C6-O6

There are no ring outliers.

12 monomers are involved in 29 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	11	BGC	5	0
3	C	13	BGC	3	0
3	C	5	BGC	6	0
3	C	8	BGC	2	0
3	C	9	BGC	3	0
3	C	10	BGC	4	0
3	C	12	BGC	4	0
3	C	18	BGC	3	0
3	C	6	BGC	1	0
3	C	7	BGC	1	0
3	C	17	BGC	4	0
3	C	16	BGC	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 C	
 Bond lengths	 Bond angles
 Torsions	 Rings

5.6 Ligand geometry (i)

3 ligands 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	$\#\|Z\| > 2$
5	LDA	A	921	-	13,15,15	2.14	2 (15%)	14,17,17	0.46	0
5	LDA	A	920	-	6,8,15	3.49	2 (33%)	7,10,17	0.43	0
4	UDP	A	919	-	25,26,26	0.97	1 (4%)	38,40,40	1.58	5 (13%)

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	LDA	A	921	-	-	4/13/13/13	-
5	LDA	A	920	-	-	3/6/6/13	-
4	UDP	A	919	-	-	10/16/32/32	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	920	LDA	O1-N1	-6.79	1.25	1.42
5	A	921	LDA	O1-N1	-6.30	1.26	1.42
5	A	920	LDA	C1-N1	-5.15	1.46	1.51
5	A	921	LDA	C1-N1	-4.30	1.47	1.51
4	A	919	UDP	C5-C4	-2.09	1.39	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	919	UDP	C4-N3-C2	-5.58	119.69	126.61
4	A	919	UDP	N3-C2-N1	4.19	120.35	114.89
4	A	919	UDP	C5-C4-N3	3.37	119.52	114.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	919	UDP	O4-C4-C5	-2.68	120.54	125.16
4	A	919	UDP	O2-C2-N1	-2.67	119.33	122.80

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

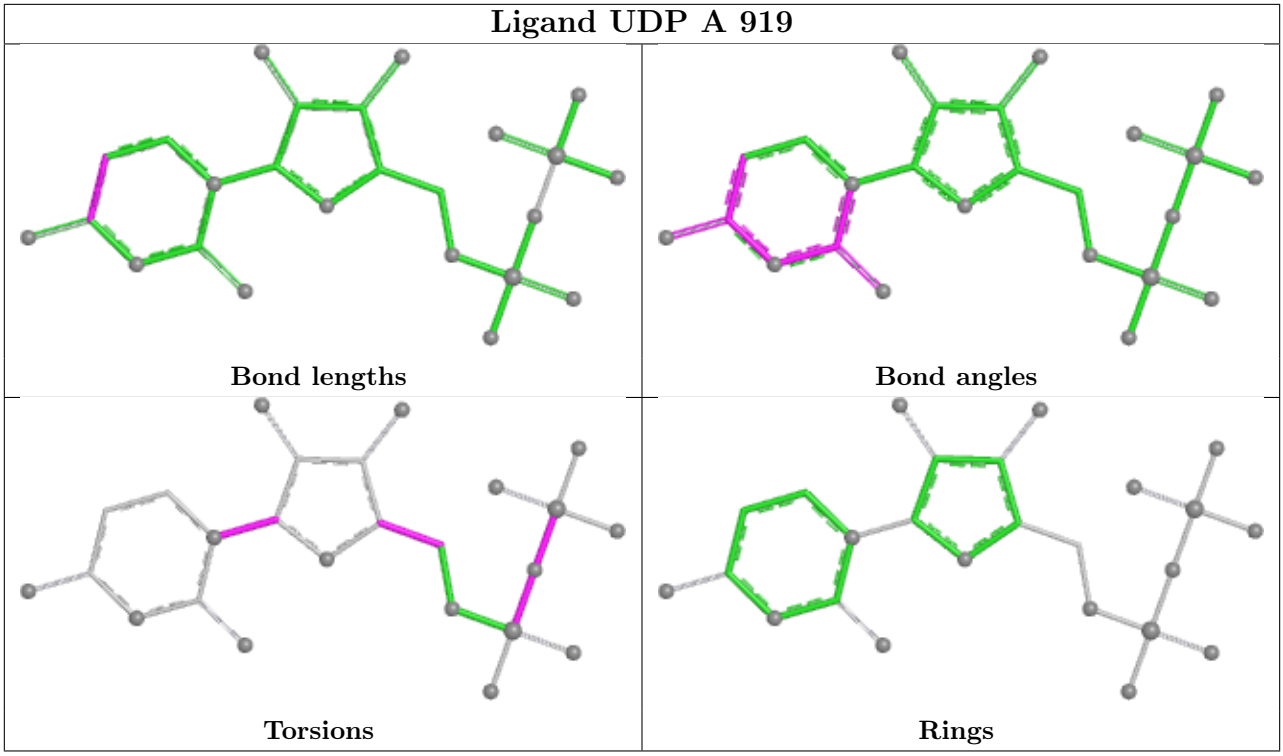
Mol	Chain	Res	Type	Atoms
5	A	920	LDA	C2-C1-N1-O1
5	A	920	LDA	C2-C1-N1-CM1
5	A	920	LDA	C2-C1-N1-CM2
4	A	919	UDP	O4'-C4'-C5'-O5'
5	A	921	LDA	C11-C10-C9-C8

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	921	LDA	2	0
5	A	920	LDA	5	0
4	A	919	UDP	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.



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	747/802 (93%)	-0.11	26 (3%)	44	40	77, 127, 205, 236	0
2	B	655/707 (92%)	-0.10	18 (2%)	54	51	82, 143, 210, 235	0
All	All	1402/1509 (92%)	-0.10	44 (3%)	49	47	77, 135, 208, 236	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	544	GLY	7.2
2	B	183	ALA	5.0
2	B	545	PRO	4.6
1	A	627	PRO	3.8
2	B	596	GLY	3.7

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
3	BGC	C	1	12/12	0.90	0.39	176,192,196,198	0
3	BGC	C	17	11/12	0.92	0.32	97,106,119,140	0
3	BGC	C	8	11/12	0.95	0.15	78,92,104,118	0
3	BGC	C	12	11/12	0.95	0.34	86,103,140,157	0
3	BGC	C	13	11/12	0.95	0.39	88,107,120,127	0

Continued on next page...

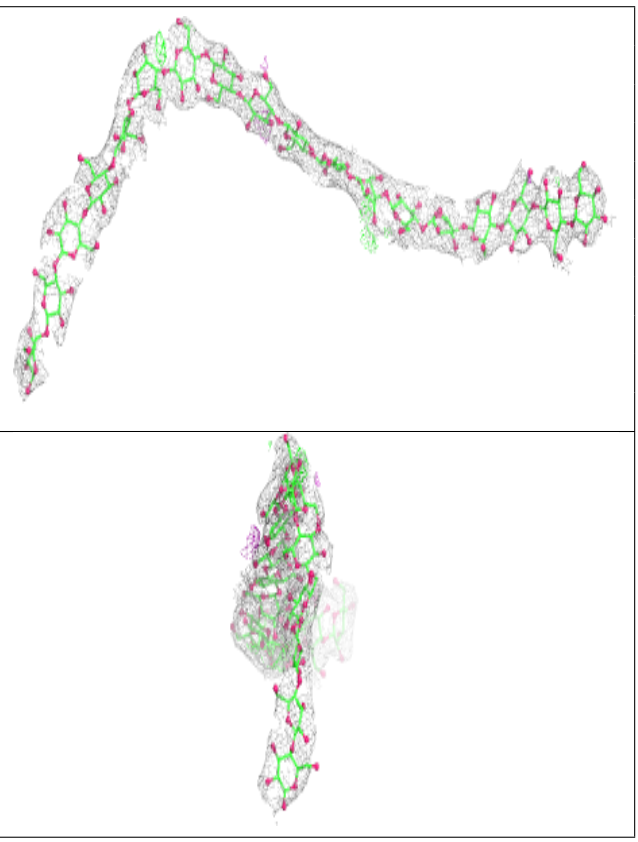
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	BGC	C	5	11/12	0.95	0.15	95,114,142,143	0
3	BGC	C	2	11/12	0.96	0.20	149,170,193,198	0
3	BGC	C	6	11/12	0.96	0.15	77,96,110,146	0
3	BGC	C	18	11/12	0.96	0.28	116,121,136,142	0
3	BGC	C	4	11/12	0.97	0.11	97,115,128,142	0
3	BGC	C	15	11/12	0.97	0.33	108,116,170,173	0
3	BGC	C	16	11/12	0.97	0.26	90,104,129,158	0
3	BGC	C	10	11/12	0.97	0.17	82,97,109,114	0
3	BGC	C	7	11/12	0.97	0.17	79,109,137,143	0
3	BGC	C	3	11/12	0.98	0.09	100,112,144,154	0
3	BGC	C	14	11/12	0.98	0.32	97,114,126,129	0
3	BGC	C	11	11/12	0.99	0.22	83,93,108,111	0
3	BGC	C	9	11/12	0.99	0.18	70,96,103,127	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 C:

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,

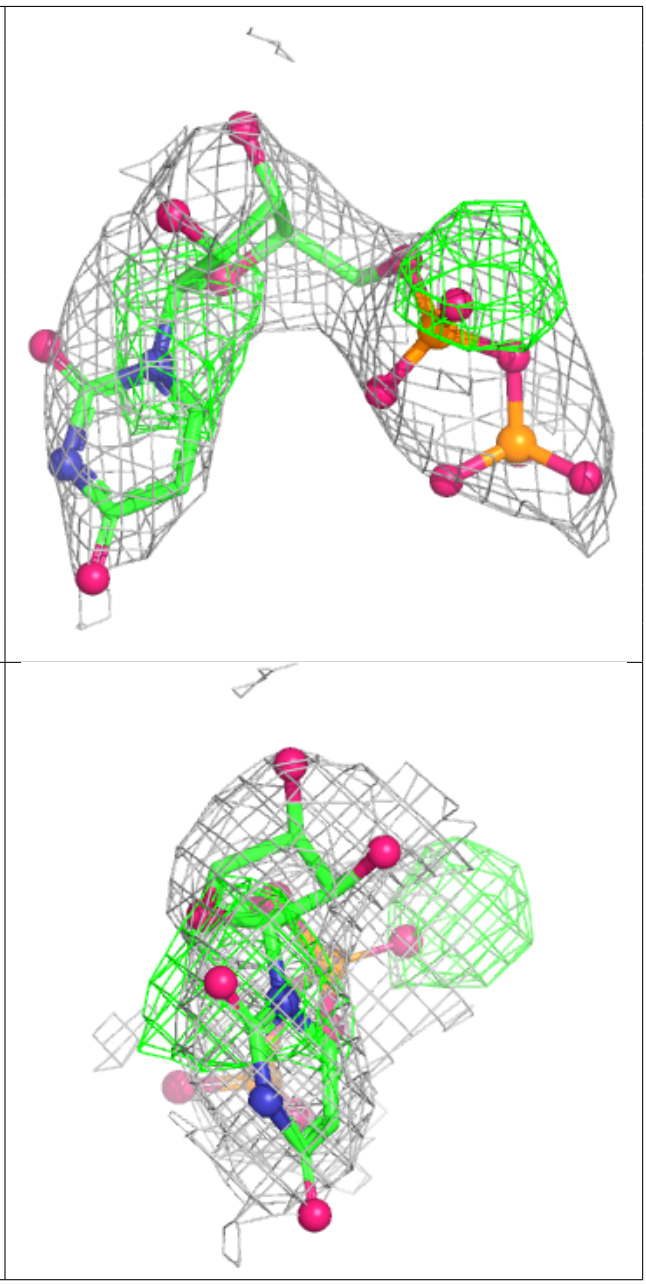
median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	UDP	A	919	25/25	0.79	0.70	57,85,113,131	25
5	LDA	A	921	16/16	0.79	0.36	105,128,177,182	0
5	LDA	A	920	9/16	0.91	0.31	92,109,143,144	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 UDP A 919:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.