wwPDB X-ray Structure Validation Summary Report (i)

Sep 13, 2023 – 09:52 PM EDT

PDB ID	:	4R5P
Title	:	Crystal structure of HIV-1 reverse transcriptase (RT) with DNA and a nucleoside triphosphate mimic alpha-carboxy nucleoside phosphonate inhibitor
Authors	:	Das, K.; Martinez, S.E.; Arnold, E.
Deposited on	:	2014-08-21
Resolution	:	2.89 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:

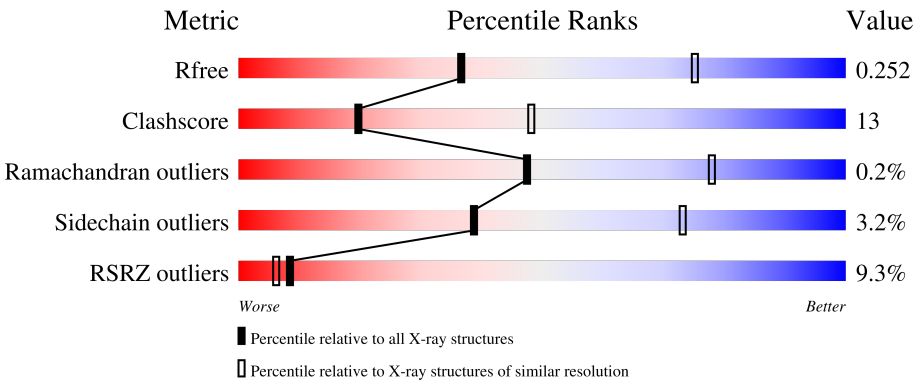
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.89 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	556	10% 66% 32% •
1	C	556	12% 65% 33% •
2	B	428	4% 70% 25% • •
2	D	428	7% 69% 27% •
3	E	27	19% 59% 30% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	T	27	11% 52% 37% 11%
4	F	21	10% 43% 48% 5% 5%
4	P	21	5% 62% 24% 10% 5%
5	G	2	50% 50%
5	H	2	50% 50%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 17714 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 HIV-1 reverse transcriptase, p66 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	555	4511	2920	751	832	8	0	0	0
1	C	554	4506	2917	750	831	8	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	expression tag	UNP P03366
A	0	VAL	-	expression tag	UNP P03366
A	258	CYS	GLN	engineered mutation	UNP P03366
A	280	SER	CYS	engineered mutation	UNP P03366
A	498	ASN	ASP	engineered mutation	UNP P03366
C	-1	MET	-	expression tag	UNP P03366
C	0	VAL	-	expression tag	UNP P03366
C	258	CYS	GLN	engineered mutation	UNP P03366
C	280	SER	CYS	engineered mutation	UNP P03366
C	498	ASN	ASP	engineered mutation	UNP P03366

- Molecule 2 is a protein called HIV-1 reverse transcriptase, p51 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	412	3391	2204	562	619	6	0	0	0
2	D	412	3393	2206	563	618	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366
D	280	SER	CYS	engineered mutation	UNP P03366

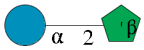
- Molecule 3 is a DNA chain called 5'-D(*TP*GP*GP*AP*CP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*CP*TP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	T	24	495	233	100	139	23	0	0	0
3	E	24	495	233	100	139	23	0	0	0

- Molecule 4 is a DNA chain called 5'-D(*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)P*CP*GP*CP*CP*G)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
4	P	20	407	195	72	120	19	1	0	0	0
4	F	20	407	195	72	120	19	1	0	0	0

- Molecule 5 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.



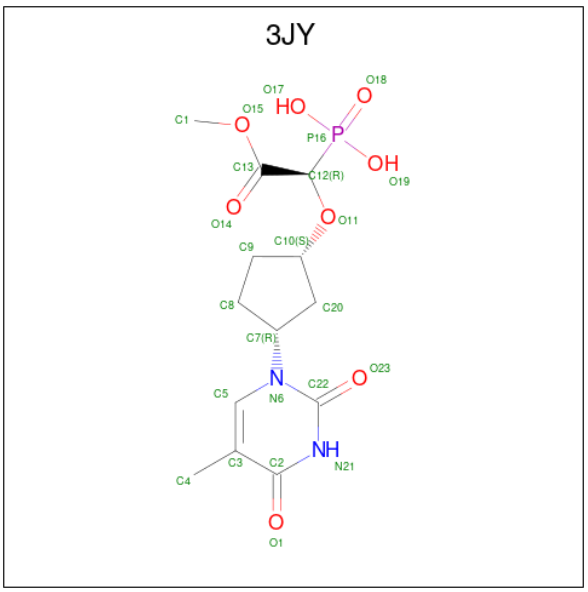
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
5	G	2	23	12	11	0	0	0
5	H	2	23	12	11	0	0	0

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
6	A	2	2	2	0	0
6	C	2	2	2	0	0
6	B	1	1	1	0	0
6	D	1	1	1	0	0

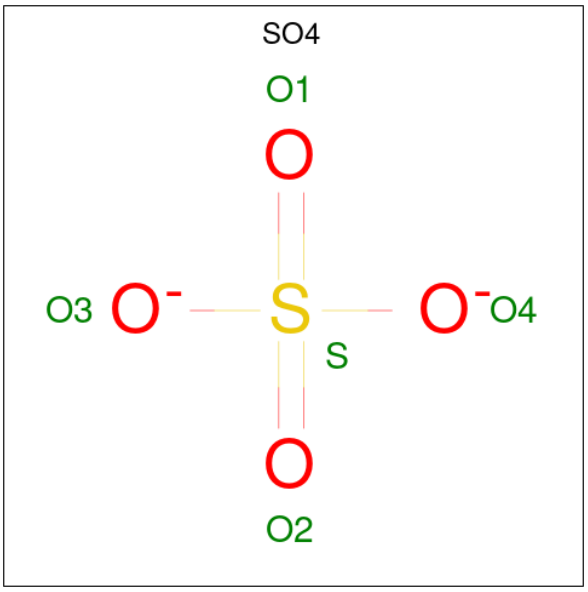
- Molecule 7 is [(1R)-2-methoxy-1-{[(1S,3R)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)cyclopentyl]oxy}-2-oxoethyl]phosphonic acid (three-letter code: 3JY) (formula:

$C_{13}H_{19}N_2O_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total 23	C 12	N 2	O 8	P 1	0	0
7	C	1	Total 23	C 12	N 2	O 8	P 1	0	0

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total 5	O 4	S 1	0	0

- Molecule 9 is water.

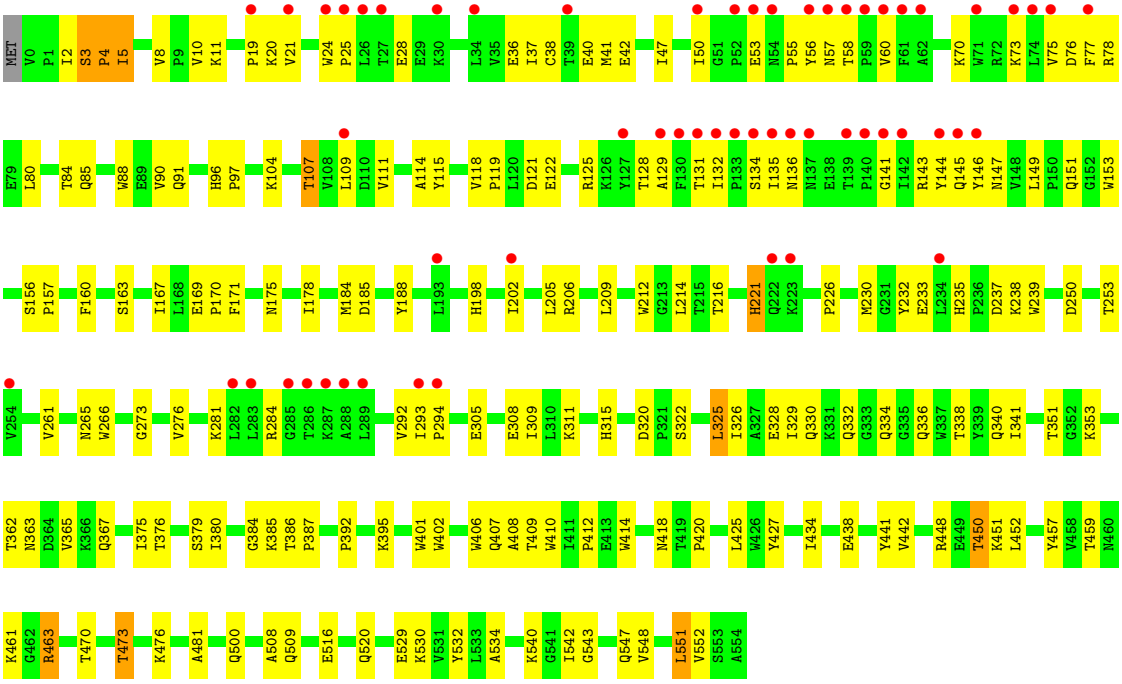
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total O 1 1	0	0
9	C	1	Total O 1 1	0	0
9	B	2	Total O 2 2	0	0
9	D	2	Total O 2 2	0	0

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.

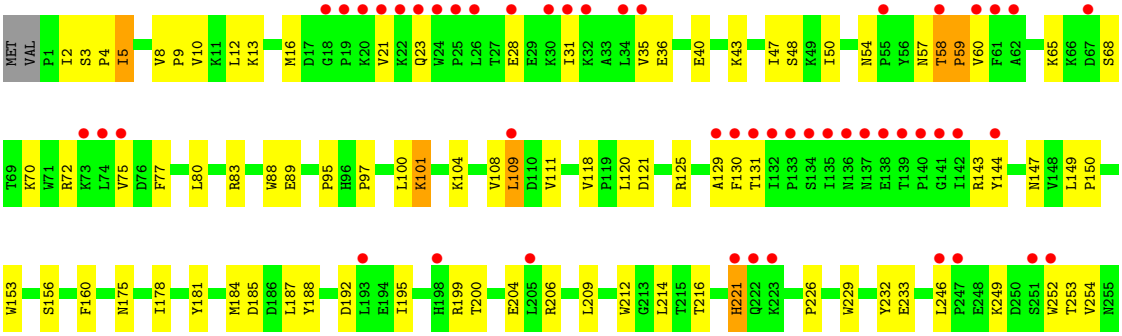
- Molecule 1: HIV-1 reverse transcriptase, p66 subunit

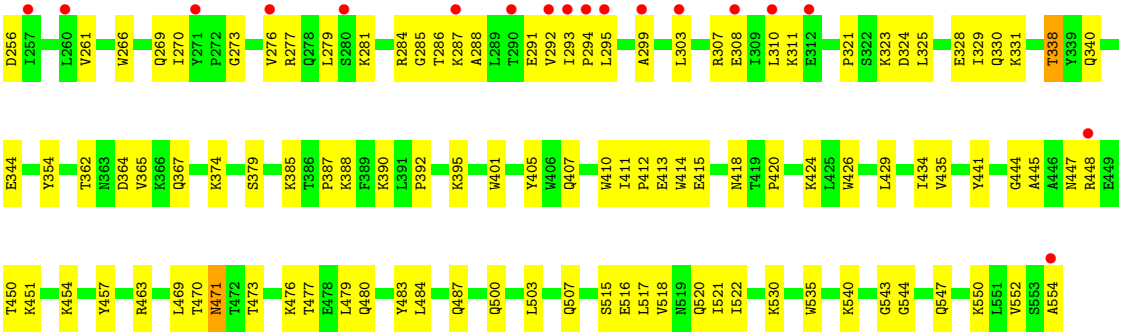
Chain A:



- Molecule 1: HIV-1 reverse transcriptase, p66 subunit

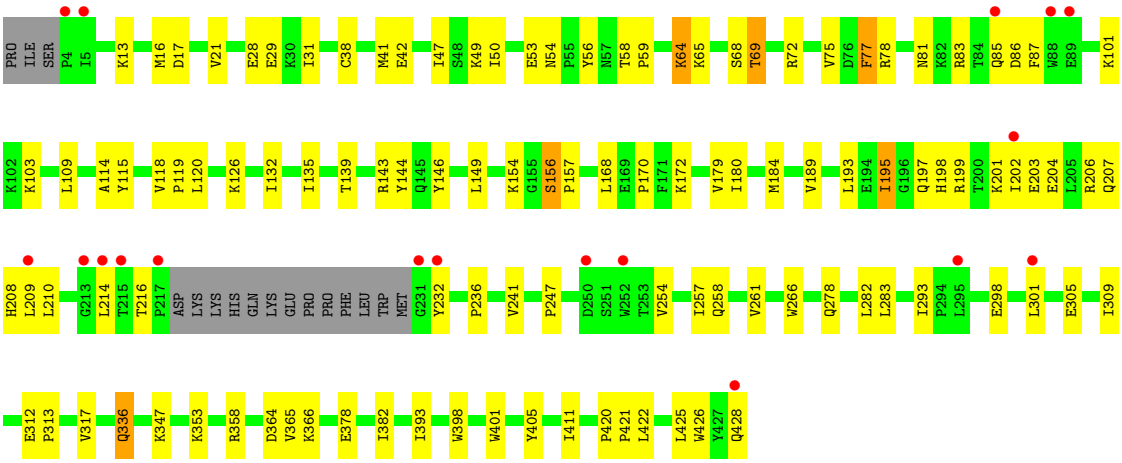
Chain C:





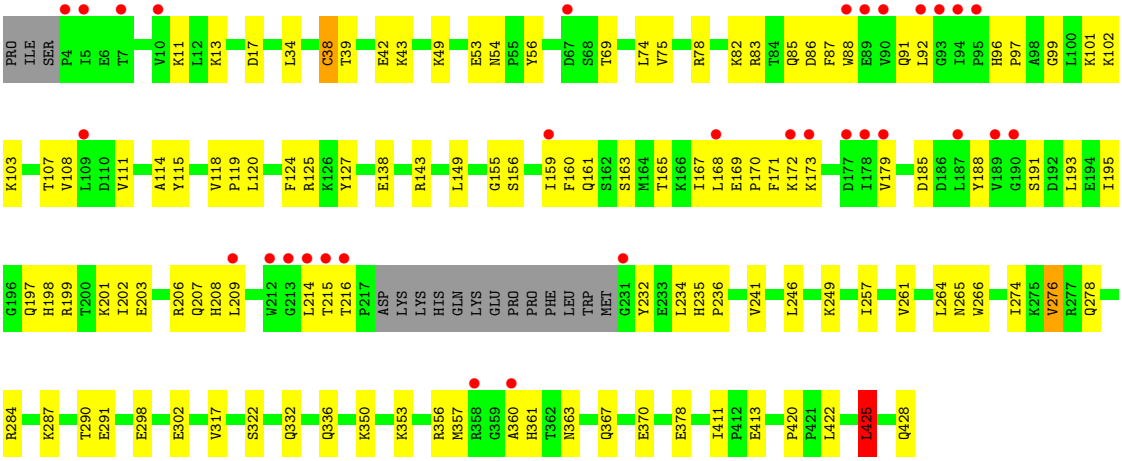
• Molecule 2: HIV-1 reverse transcriptase, p51 subunit

Chain B:



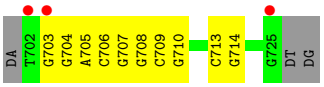
• Molecule 2: HIV-1 reverse transcriptase, p51 subunit

Chain D:



• Molecule 3: 5'-D(*TP*GP*GP*AP*CP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*CP*TP*G)-3'

Chain T:

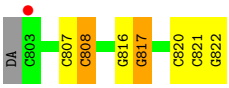


• Molecule 3: 5'-D(*TP*GP*GP*AP*CP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*CP*TP*G)-3'

Chain E:

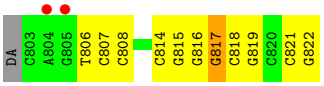
• Molecule 4: 5'-D(*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)P*CP*GP*CP*CP*G)-3'

Chain P:



• Molecule 4: 5'-D(*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)P*CP*GP*CP*CP*G)-3'

Chain F:



• Molecule 5: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain G:

• Molecule 5: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	89.91Å 133.53Å 139.22Å 90.00° 97.81° 90.00°	Depositor
Resolution (Å)	44.54 – 2.89 48.50 – 2.89	Depositor EDS
% Data completeness (in resolution range)	99.2 (44.54-2.89) 99.1 (48.50-2.89)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.195 , 0.251 0.203 , 0.252	Depositor DCC
R_{free} test set	2152 reflections (2.98%)	wwPDB-VP
Wilson B-factor (Å^2)	69.4	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 58.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	17714	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.62% 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: MG, FRU, GLC, MRG, 3JY, SO4

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	0.60	0/4629	0.75	3/6290 (0.0%)
1	C	0.54	0/4624	0.69	1/6282 (0.0%)
2	B	0.63	0/3486	0.76	1/4735 (0.0%)
2	D	0.59	1/3489 (0.0%)	0.74	1/4740 (0.0%)
3	E	0.81	0/557	0.98	2/859 (0.2%)
3	T	0.85	0/557	0.96	0/859
4	F	0.93	0/424	1.07	1/649 (0.2%)
4	P	0.93	1/424 (0.2%)	1.03	0/649
All	All	0.63	2/18190 (0.0%)	0.77	9/25063 (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	D	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	298	GLU	CG-CD	5.24	1.59	1.51
4	P	808	DC	C3'-O3'	-5.03	1.37	1.44

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	364	ASP	CB-CG-OD2	-6.29	112.63	118.30
1	A	325	LEU	CB-CG-CD1	-6.22	100.43	111.00
1	A	134	SER	CB-CA-C	5.54	120.63	110.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	703	DG	O4'-C4'-C3'	-5.25	102.40	104.50
1	A	551	LEU	CA-CB-CG	5.24	127.35	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	420	PRO	Peptide

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	4511	0	4570	126	0
1	C	4506	0	4568	124	0
2	B	3391	0	3425	89	0
2	D	3393	0	3426	83	0
3	E	495	0	268	7	0
3	T	495	0	268	11	0
4	F	407	0	229	11	0
4	P	407	0	229	9	0
5	G	23	0	21	1	0
5	H	23	0	21	4	0
6	A	2	0	0	0	0
6	B	1	0	0	0	0
6	C	2	0	0	0	0
6	D	1	0	0	0	0
7	A	23	0	14	4	0
7	C	23	0	14	2	0
8	C	5	0	0	0	0
9	A	1	0	0	0	0
9	B	2	0	0	0	0
9	C	1	0	0	0	0
9	D	2	0	0	1	0
All	All	17714	0	17053	441	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 441 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)
1:A:408:ALA:HB3	2:B:393:ILE:HG13	1.56	0.85
2:B:209:LEU:HD22	2:B:214:LEU:HD23	1.57	0.85
1:A:542:ILE:HG23	2:B:283:LEU:HD13	1.60	0.84
2:D:209:LEU:HD22	2:D:214:LEU:HD23	1.58	0.84
2:D:115:TYR:HD2	2:D:156:SER:HB3	1.46	0.81

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	553/556 (100%)	518 (94%)	34 (6%)	1 (0%)	47 78
1	C	552/556 (99%)	526 (95%)	24 (4%)	2 (0%)	34 66
2	B	408/428 (95%)	389 (95%)	19 (5%)	0	100 100
2	D	408/428 (95%)	388 (95%)	20 (5%)	0	100 100
All	All	1921/1968 (98%)	1821 (95%)	97 (5%)	3 (0%)	47 78

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	PRO
1	C	59	PRO
1	C	4	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	495/497 (100%)	478 (97%)	17 (3%)	37	71
1	C	495/497 (100%)	479 (97%)	16 (3%)	39	73
2	B	373/390 (96%)	363 (97%)	10 (3%)	44	77
2	D	373/390 (96%)	360 (96%)	13 (4%)	36	70
All	All	1736/1774 (98%)	1680 (97%)	56 (3%)	39	73

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	324	ASP
2	D	428	GLN
2	B	69	THR
2	D	425	LEU
2	D	241	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	487	GLN
1	C	507	GLN
2	D	151	GLN
2	B	278	GLN
1	C	147	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)

2 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	$\#\|Z\| > 2$
4	MRG	F	817	4,3,1	21,28,29	2.57	8 (38%)	19,39,42	1.95	7 (36%)
4	MRG	P	817	4,3,1	21,28,29	2.45	8 (38%)	19,39,42	1.63	5 (26%)

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	MRG	F	817	4,3,1	-	2/8/26/27	0/3/3/3
4	MRG	P	817	4,3,1	-	3/8/26/27	0/3/3/3

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	817	MRG	C2-N2	7.59	1.50	1.34
4	P	817	MRG	C2-N2	6.95	1.49	1.34
4	F	817	MRG	C2-N1	4.90	1.44	1.36
4	P	817	MRG	C2-N1	4.34	1.43	1.36
4	F	817	MRG	C2'-C3'	-3.49	1.43	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	817	MRG	C22-C21-N2	4.10	123.94	112.21
4	P	817	MRG	C22-C21-N2	3.96	123.54	112.21
4	F	817	MRG	O6-C6-C5	-3.58	117.38	124.37
4	F	817	MRG	C5-C6-N1	3.17	119.55	113.95
4	F	817	MRG	C21-N2-C2	2.96	128.96	123.41

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	P	817	MRG	C22-C21-N2-C2
4	P	817	MRG	N2-C21-C22-C23
4	F	817	MRG	C22-C21-N2-C2
4	F	817	MRG	N2-C21-C22-C23
4	P	817	MRG	C21-C22-C23-S24

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	817	MRG	1	0
4	P	817	MRG	1	0

5.5 Carbohydrates (i)

4 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	$\#\|Z\| > 2$
5	GLC	G	1	5	11,11,12	0.32	0	15,15,17	0.88	0
5	FRU	G	2	5	11,12,12	0.72	1 (9%)	10,18,18	1.26	1 (10%)
5	GLC	H	1	5	11,11,12	0.26	0	15,15,17	0.59	0
5	FRU	H	2	5	11,12,12	0.55	0	10,18,18	1.05	1 (10%)

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	GLC	G	1	5	-	0/2/19/22	0/1/1/1
5	FRU	G	2	5	-	3/5/24/24	0/1/1/1
5	GLC	H	1	5	-	2/2/19/22	0/1/1/1
5	FRU	H	2	5	-	4/5/24/24	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	2	FRU	O2-C2	2.01	1.44	1.40

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	G	2	FRU	C6-C5-C4	-2.38	109.34	115.09
5	H	2	FRU	C6-C5-C4	-2.04	110.18	115.09

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	G	2	FRU	O1-C1-C2-C3
5	G	2	FRU	O1-C1-C2-O2
5	G	2	FRU	O1-C1-C2-O5
5	H	2	FRU	O1-C1-C2-C3
5	H	2	FRU	O1-C1-C2-O2

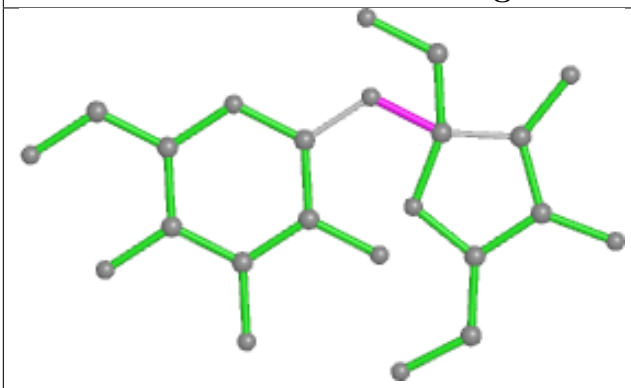
There are no ring outliers.

4 monomers are involved in 5 short contacts:

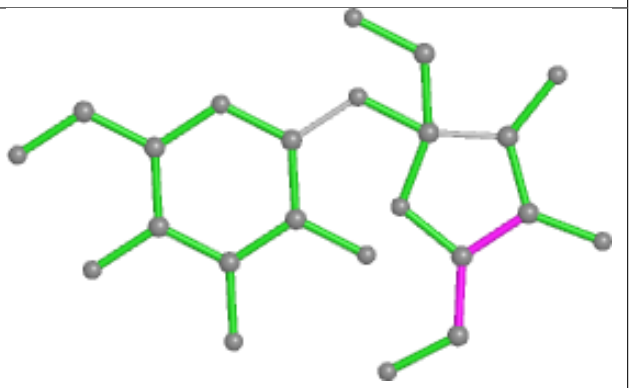
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	H	2	FRU	3	0
5	G	1	GLC	1	0
5	H	1	GLC	3	0
5	G	2	FRU	1	0

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

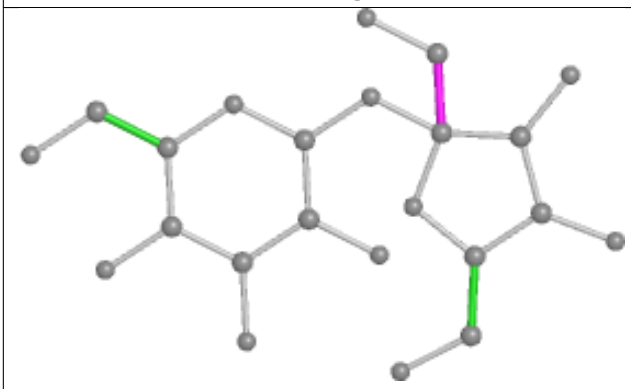
Oligosaccharide Chain G



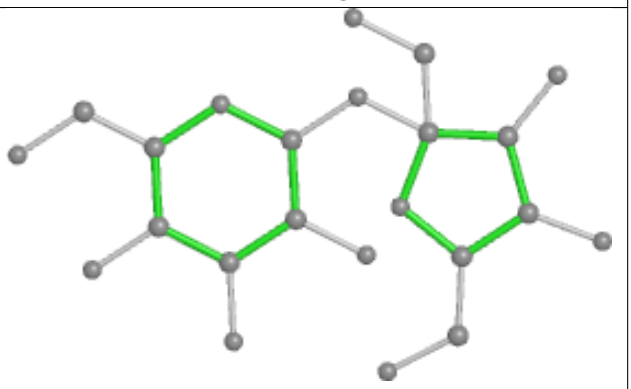
Bond lengths



Bond angles

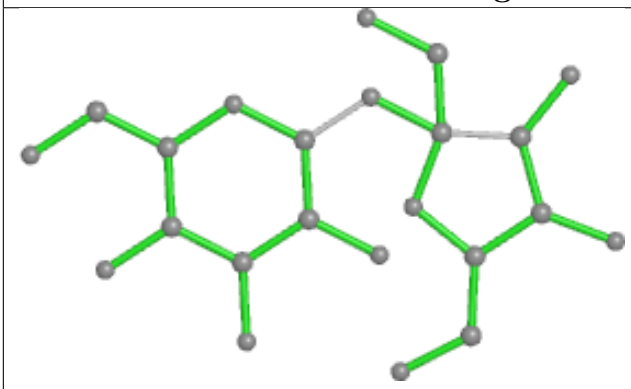


Torsions

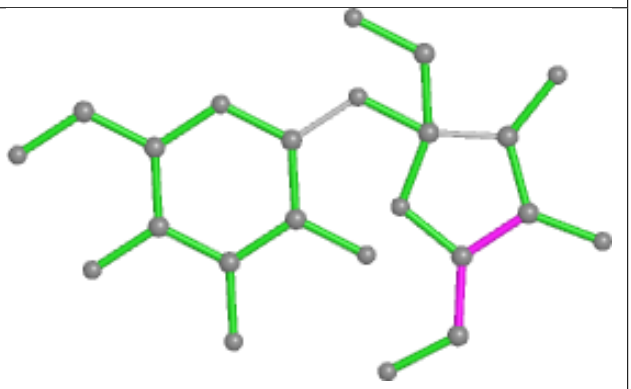


Rings

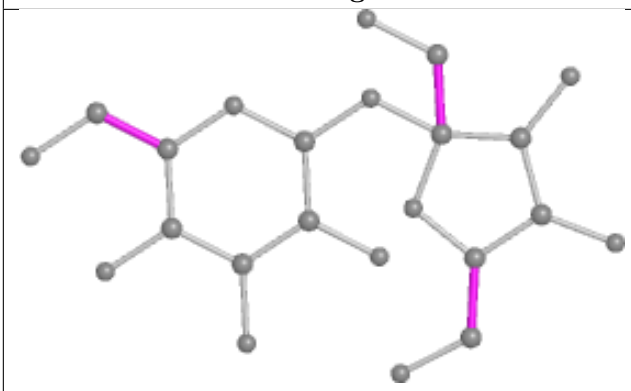
Oligosaccharide Chain H



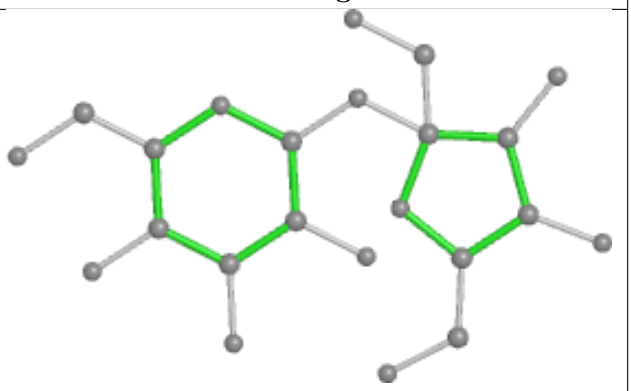
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 6 are monoatomic - leaving 3 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
7	3JY	A	602	6	21,24,25	2.61	8 (38%)	30,36,37	2.62	14 (46%)
8	SO4	C	604	-	4,4,4	0.26	0	6,6,6	0.59	0
7	3JY	C	602	6	21,24,25	2.63	7 (33%)	30,36,37	2.31	11 (36%)

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	3JY	A	602	6	-	7/12/27/29	0/2/2/2
7	3JY	C	602	6	-	6/12/27/29	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	C	602	3JY	O11-C10	-7.06	1.34	1.44
7	A	602	3JY	O11-C10	-6.90	1.35	1.44
7	C	602	3JY	O1-C2	6.12	1.35	1.23
7	A	602	3JY	O1-C2	5.98	1.35	1.23
7	C	602	3JY	C5-C3	3.31	1.40	1.34

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	602	3JY	P16-C12-C13	-5.93	97.90	110.87
7	C	602	3JY	C3-C2-N21	4.79	119.40	115.31
7	C	602	3JY	C2-N21-C22	-4.74	121.21	127.35
7	C	602	3JY	O1-C2-C3	-4.69	119.46	124.90
7	A	602	3JY	C2-N21-C22	-4.50	121.52	127.35

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	602	3JY	C20-C10-O11-C12
7	A	602	3JY	O11-C12-C13-O14
7	A	602	3JY	C13-C12-P16-O18
7	A	602	3JY	C13-C12-P16-O19
7	A	602	3JY	C13-C12-P16-O17

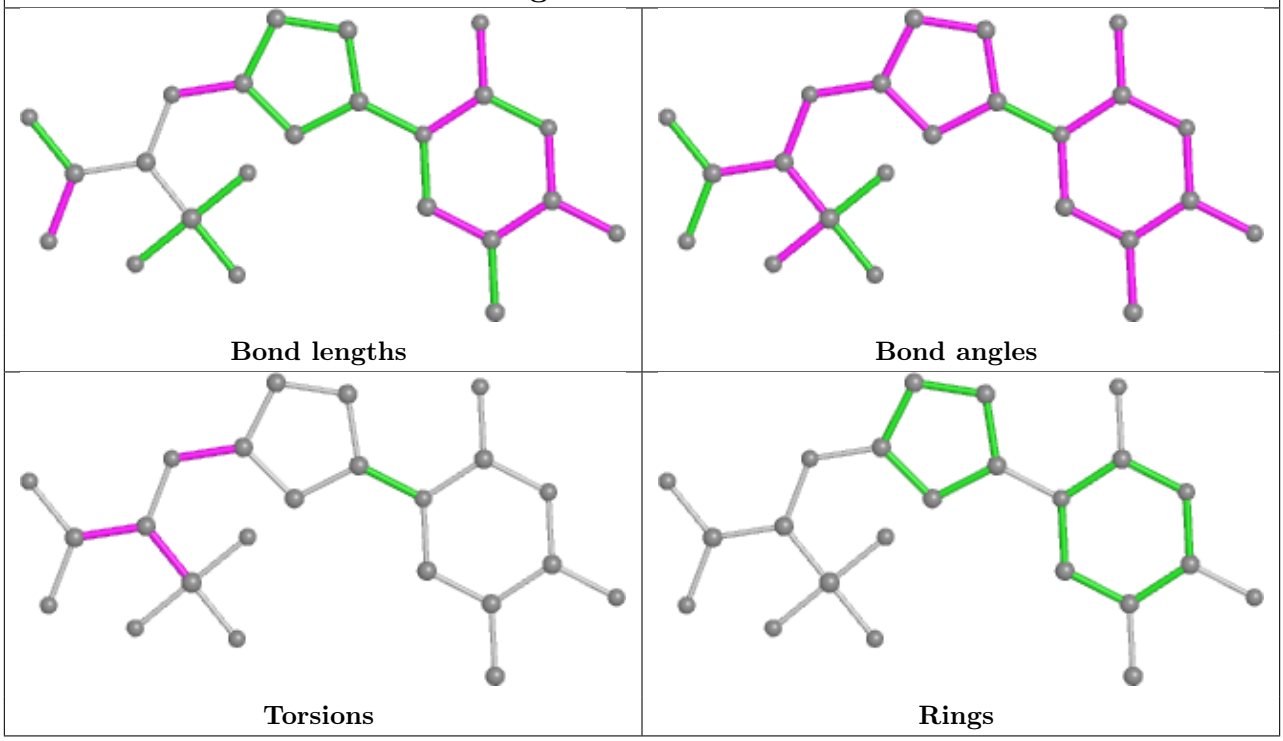
There are no ring outliers.

2 monomers are involved in 6 short contacts:

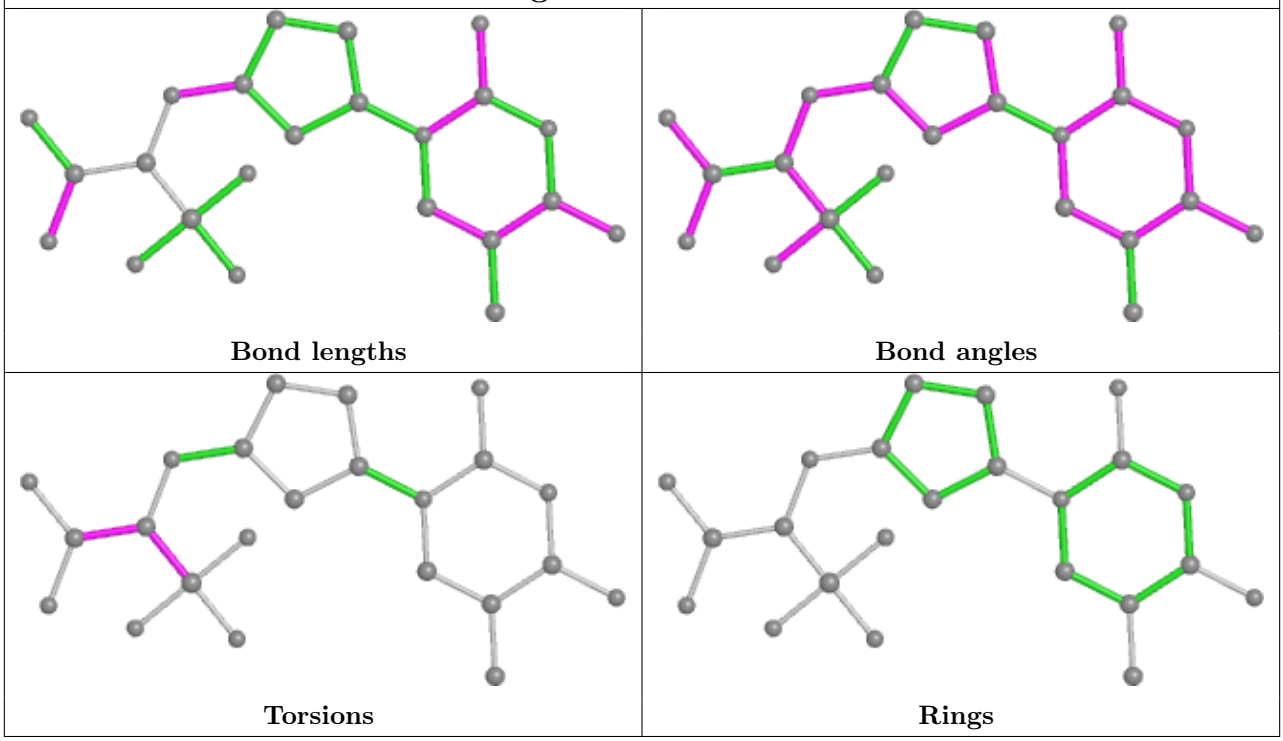
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	602	3JY	4	0
7	C	602	3JY	2	0

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 3JY A 602



Ligand 3JY C 602



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	555/556 (99%)	0.49	58 (10%)	6	5	30, 77, 151, 185	0
1	C	554/556 (99%)	0.64	68 (12%)	4	3	32, 83, 149, 180	0
2	B	412/428 (96%)	0.20	18 (4%)	34	30	28, 60, 101, 126	0
2	D	412/428 (96%)	0.36	32 (7%)	13	10	33, 74, 119, 150	0
3	E	24/27 (88%)	0.82	5 (20%)	1	0	71, 108, 176, 199	0
3	T	24/27 (88%)	0.50	3 (12%)	3	3	66, 99, 176, 190	0
4	F	19/21 (90%)	0.37	2 (10%)	6	5	59, 89, 157, 160	0
4	P	19/21 (90%)	-0.02	1 (5%)	26	22	63, 84, 142, 145	0
All	All	2019/2064 (97%)	0.44	187 (9%)	8	6	28, 74, 145, 199	0

The worst 5 of 187 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	214	LEU	13.3
2	D	4	PRO	11.3
3	T	702	DT	9.6
2	D	214	LEU	8.8
1	A	140	PRO	8.5

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. 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	MRG	P	817	26/27	0.86	0.17	83,98,119,131	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	MRG	F	817	26/27	0.88	0.19	93,111,132,135	0

6.3 Carbohydrates (i)

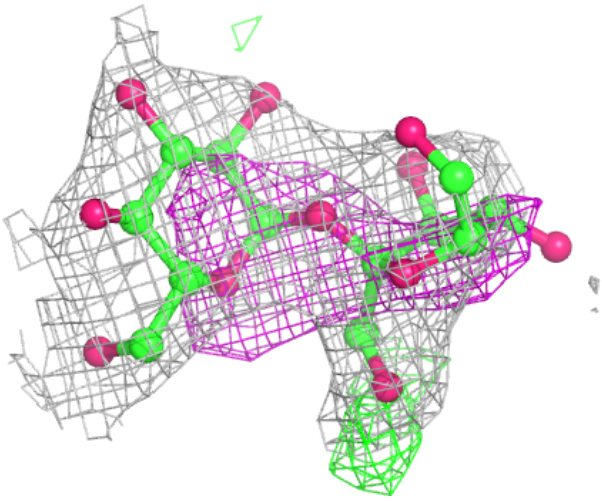
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

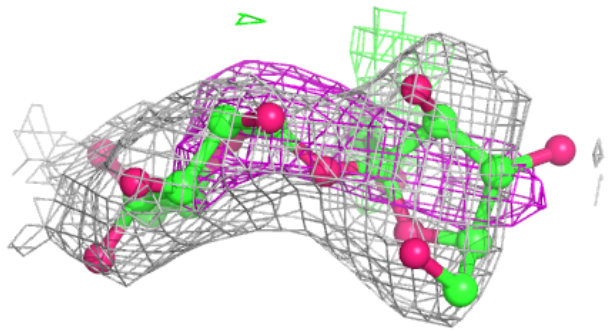
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	FRU	G	2	12/12	0.83	0.23	71,82,92,97	0
5	GLC	H	1	11/12	0.90	0.15	66,79,84,87	0
5	FRU	H	2	12/12	0.90	0.19	83,98,106,106	0
5	GLC	G	1	11/12	0.94	0.21	56,63,72,74	0

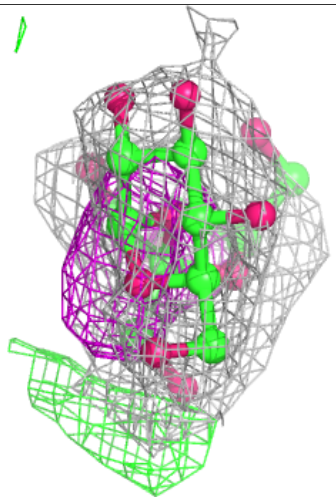
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain G:

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

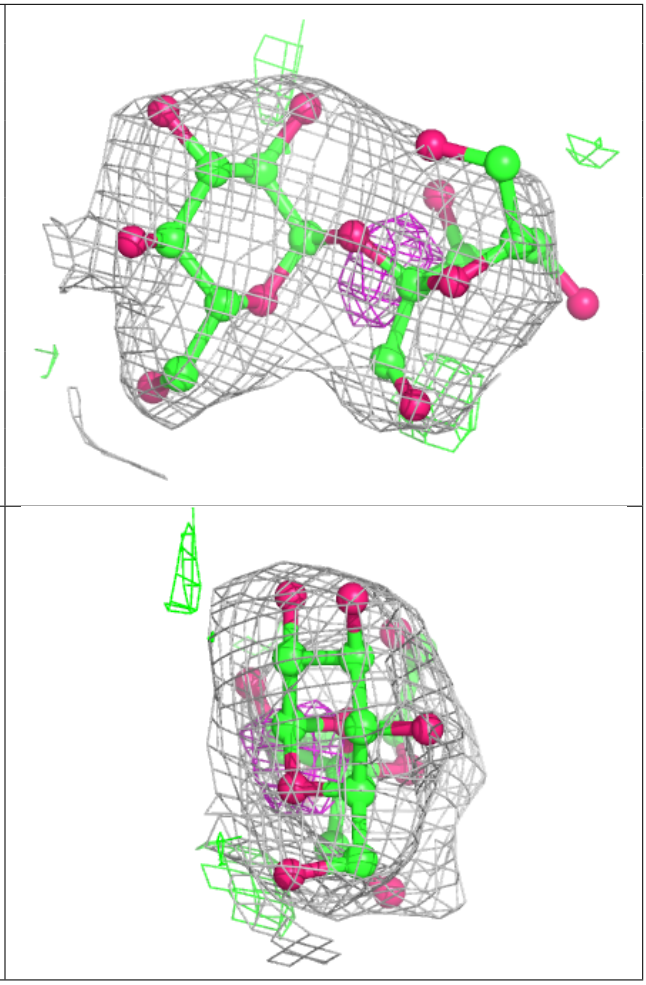






Electron density around Chain H:

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



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.

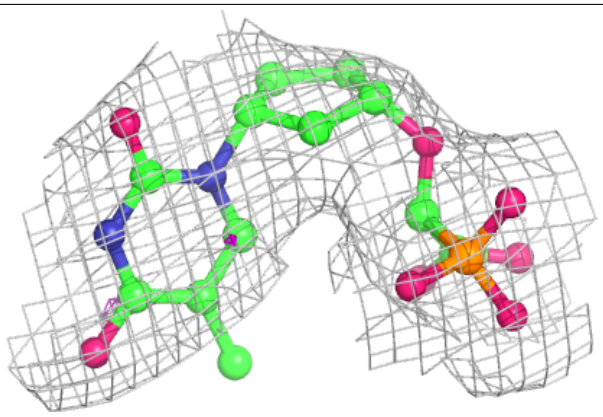
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	MG	C	601	1/1	0.79	0.11	63,63,63,63	0
6	MG	A	603	1/1	0.80	0.17	72,72,72,72	0
6	MG	B	502	1/1	0.83	0.38	51,51,51,51	0
8	SO4	C	604	5/5	0.85	0.28	74,84,89,104	0
6	MG	A	601	1/1	0.88	0.12	100,100,100,100	0
7	3JY	A	602	23/24	0.91	0.20	78,89,106,109	0
6	MG	C	603	1/1	0.92	0.07	113,113,113,113	0
7	3JY	C	602	23/24	0.93	0.19	79,94,106,114	0
6	MG	D	502	1/1	0.96	0.22	57,57,57,57	0

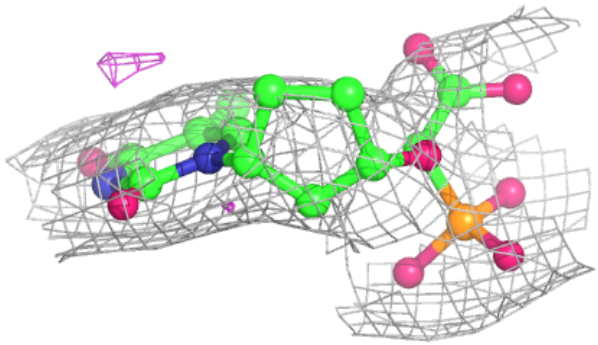
The following is a graphical depiction of the model fit to experimental electron density of all

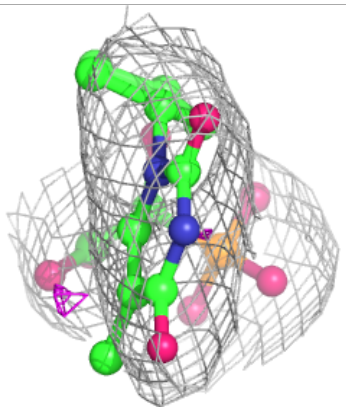
instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 3JY A 602:

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

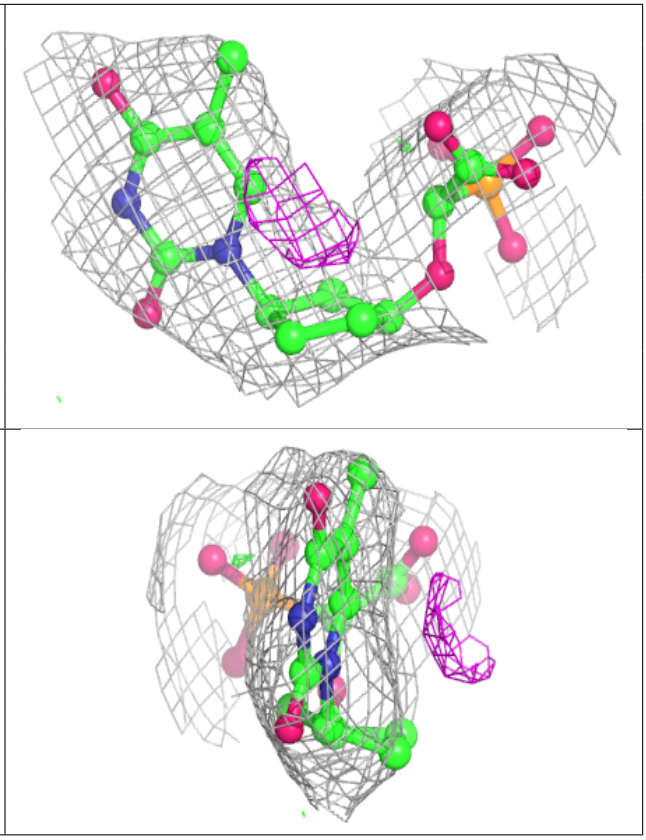






Electron density around 3JY C 602:

$2mF_o$-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.