wwPDB X-ray Structure Validation Summary Report (i)

Jan 26, 2026 – 04:15 PM JST

PDB ID	:	9VJJ / pdb_00009vjj
Title	:	Crystal Structure of human Latent TGF-beta1 in complex with SOF10
Authors	:	Kawauchi, H.; Kanamori, M.; Sato, I.; Fukami, T.A.; Irie, M.; Torizawa, T.; Shimada, H.
Deposited on	:	2025-06-20
Resolution	:	2.48 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:

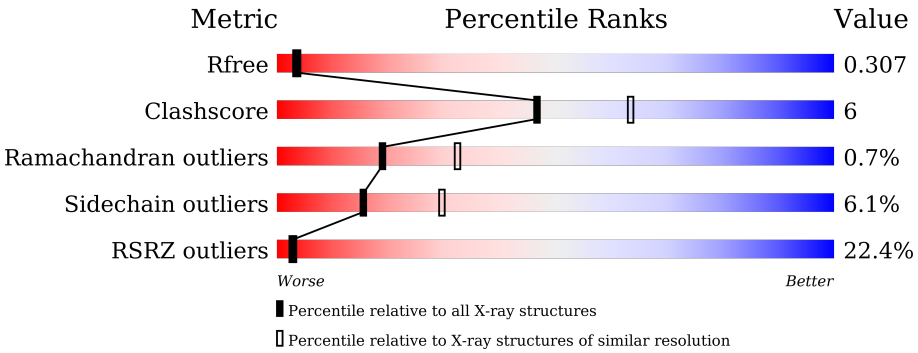
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	?.? (???), CSD ??CSD?? (????)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

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 2.48 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}	164625	7106 (2.50-2.46)
Clashscore	180529	7991 (2.50-2.46)
Ramachandran outliers	177936	7888 (2.50-2.46)
Sidechain outliers	177891	7890 (2.50-2.46)
RSRZ outliers	164620	7106 (2.50-2.46)

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	369	20% 64% 16% · 18%
1	B	369	19% 69% 16% · 14%
2	H	227	7% 46% 8% 46%
2	I	227	20% 44% 8% · 47%
3	L	215	4% 42% 7% 50%
3	M	215	13% 39% 10% · 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	C	2	100%
4	D	2	50% 50%
4	E	2	100%
4	X	2	100%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8418 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 Transforming growth factor beta-1 proprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	0	0
			2380	1514	414	436	16			
1	B	317	Total	C	N	O	S	0	0	0
			2475	1577	425	456	17			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	ASP	-	expression tag	UNP P01137
A	23	TYR	-	expression tag	UNP P01137
A	24	LYS	-	expression tag	UNP P01137
A	25	ASP	-	expression tag	UNP P01137
A	26	ASP	-	expression tag	UNP P01137
A	27	ASP	-	expression tag	UNP P01137
A	28	ASP	-	expression tag	UNP P01137
A	29	LYS	-	expression tag	UNP P01137
A	33	SER	CYS	engineered mutation	UNP P01137
B	22	ASP	-	expression tag	UNP P01137
B	23	TYR	-	expression tag	UNP P01137
B	24	LYS	-	expression tag	UNP P01137
B	25	ASP	-	expression tag	UNP P01137
B	26	ASP	-	expression tag	UNP P01137
B	27	ASP	-	expression tag	UNP P01137
B	28	ASP	-	expression tag	UNP P01137
B	29	LYS	-	expression tag	UNP P01137
B	33	SER	CYS	engineered mutation	UNP P01137

- Molecule 2 is a protein called SOF10 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	122	Total	C	N	O	S	0	0	0
			940	600	157	178	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	121	Total	C	N	O	S	0	0	0
			926	591	155	175	5			

- Molecule 3 is a protein called SOF10 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	107	Total	C	N	O	S	0	0	0
			795	497	126	169	3			
3	M	107	Total	C	N	O	S	0	0	0
			786	493	124	166	3			

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	X	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	C	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	D	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	E	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is water.

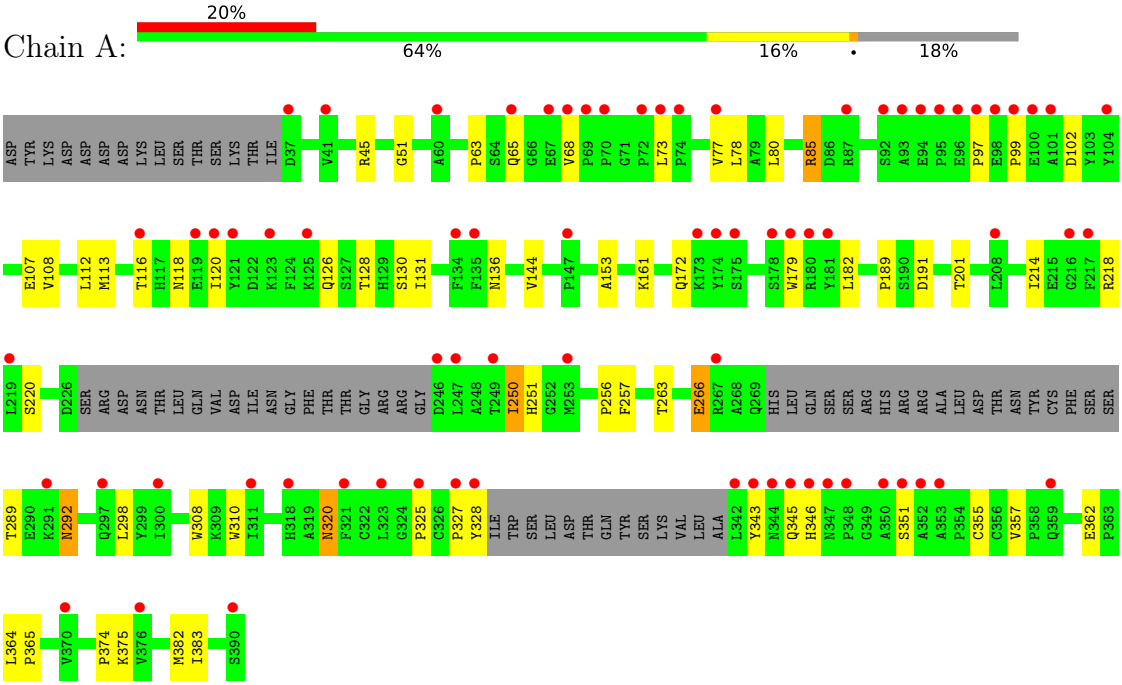
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total	O	0	0
			2	2		
5	L	1	Total	O	0	0
			1	1		
5	M	1	Total	O	0	0
			1	1		

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.

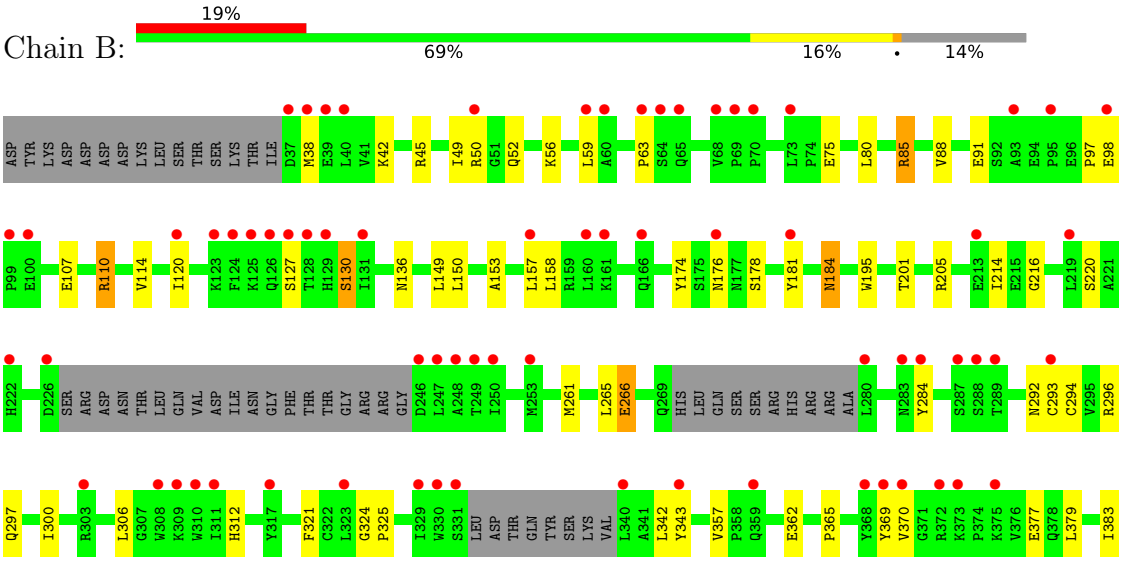
• Molecule 1: Transforming growth factor beta-1 proprotein

Chain A:



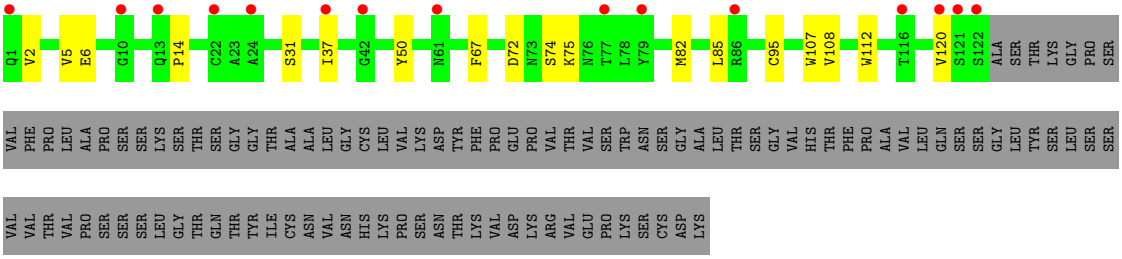
• Molecule 1: Transforming growth factor beta-1 proprotein

Chain B:



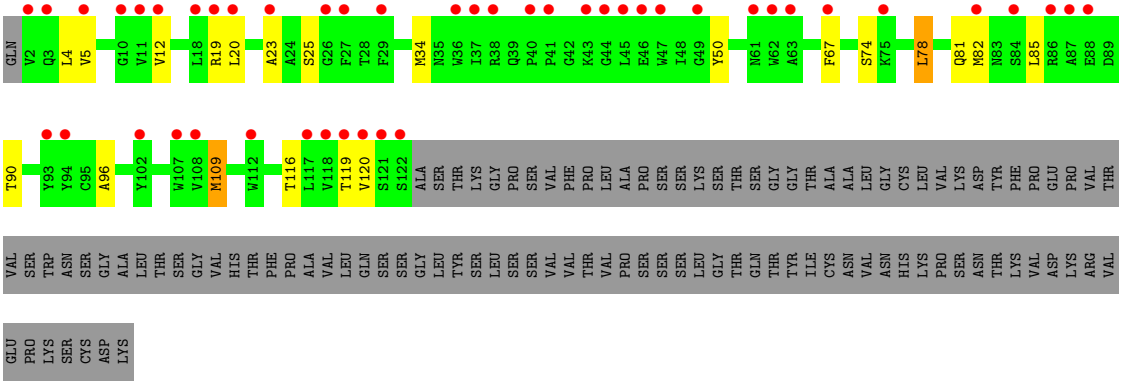
• Molecule 2: SOF10 Fab heavy chain

Chain H:



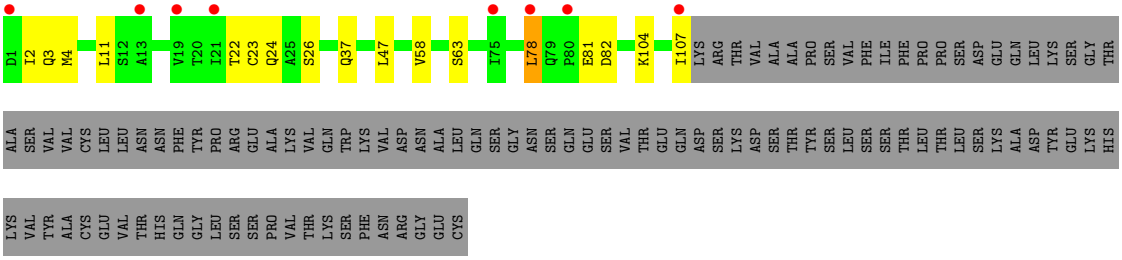
• Molecule 2: SOF10 Fab heavy chain

Chain I:



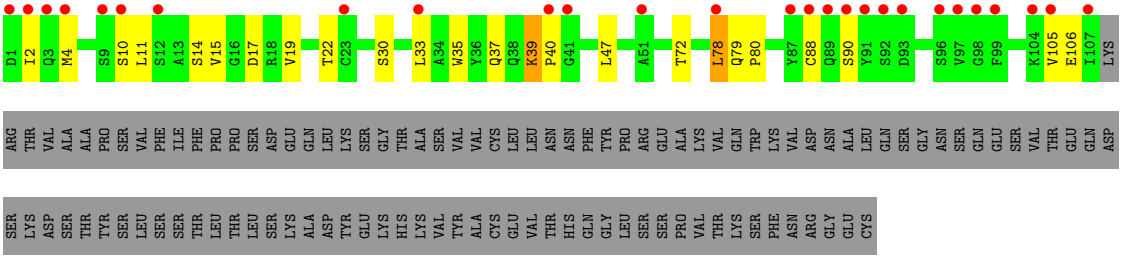
• Molecule 3: SOF10 Fab light chain

Chain L:



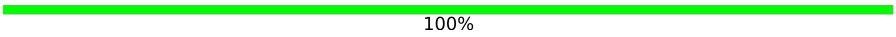
• Molecule 3: SOF10 Fab light chain

Chain M:



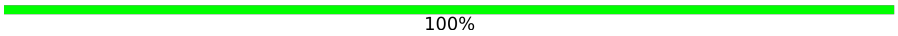
• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain X:



• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

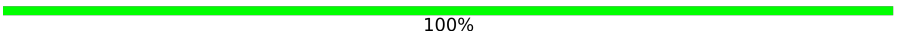


• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	225.88Å 91.71Å 126.12Å 90.00° 108.58° 90.00°	Depositor
Resolution (Å)	36.12 – 2.48 36.12 – 2.48	Depositor EDS
% Data completeness (in resolution range)	62.9 (36.12-2.48) 62.9 (36.12-2.48)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 2.48Å)	Xtriage
Refinement program	BUSTER 2.11.8	Depositor
R, R_{free}	0.281 , 0.310 0.279 , 0.307	Depositor DCC
R_{free} test set	2765 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	77.7	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 38.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	8418	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.73	0/2440	1.06	4/3319 (0.1%)
1	B	0.76	0/2539	1.10	8/3459 (0.2%)
2	H	0.73	0/966	0.97	1/1318 (0.1%)
2	I	0.65	0/952	0.96	0/1301
3	L	0.71	0/812	0.99	0/1105
3	M	0.70	1/803 (0.1%)	0.95	0/1094
All	All	0.73	1/8512 (0.0%)	1.04	13/11596 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	M	39	LYS	CA-C	5.03	1.59	1.52

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	136	ASN	CA-CB-CG	6.27	118.87	112.60
1	B	130	SER	N-CA-C	5.67	117.90	108.99
1	B	284	TYR	CA-C-N	5.59	128.04	120.44
1	B	284	TYR	C-N-CA	5.59	128.04	120.44
1	B	216	GLY	N-CA-C	5.58	120.58	111.27

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2380	0	2299	36	0
1	B	2475	0	2347	32	0
2	H	940	0	880	9	0
2	I	926	0	858	9	0
3	L	795	0	755	5	0
3	M	786	0	740	10	0
4	C	28	0	25	0	0
4	D	28	0	25	1	0
4	E	28	0	25	0	0
4	X	28	0	25	0	0
5	B	2	0	0	0	0
5	L	1	0	0	0	0
5	M	1	0	0	0	0
All	All	8418	0	7979	91	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 6.

The worst 5 of 91 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:292:ASN:HA	1:A:325:PRO:HD2	1.36	1.05
1:B:292:ASN:HA	1:B:325:PRO:HD2	1.47	0.96
1:A:116:THR:HG21	1:A:251:HIS:CE1	2.06	0.90
1:B:158:LEU:HD13	1:B:195:TRP:CE2	2.11	0.86
1:A:97:PRO:HD3	2:I:74:SER:HA	1.60	0.84

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	295/369 (80%)	273 (92%)	18 (6%)	4 (1%)	9 16
1	B	309/369 (84%)	285 (92%)	23 (7%)	1 (0%)	37 54
2	H	120/227 (53%)	116 (97%)	4 (3%)	0	100 100
2	I	119/227 (52%)	112 (94%)	7 (6%)	0	100 100
3	L	105/215 (49%)	100 (95%)	4 (4%)	1 (1%)	13 23
3	M	105/215 (49%)	91 (87%)	13 (12%)	1 (1%)	13 23
All	All	1053/1622 (65%)	977 (93%)	69 (7%)	7 (1%)	19 32

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	L	78	LEU
1	A	128	THR
1	A	320	ASN
1	B	127	SER
1	A	63	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	256/330 (78%)	243 (95%)	13 (5%)	20 38
1	B	261/330 (79%)	246 (94%)	15 (6%)	17 32
2	H	95/191 (50%)	91 (96%)	4 (4%)	25 46
2	I	93/191 (49%)	86 (92%)	7 (8%)	11 21
3	L	90/188 (48%)	82 (91%)	8 (9%)	8 15
3	M	87/188 (46%)	80 (92%)	7 (8%)	10 18
All	All	882/1418 (62%)	828 (94%)	54 (6%)	15 29

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	2	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	I	78	LEU
3	M	14	SER
2	H	5	VAL
2	I	4	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	3	GLN
3	L	79	GLN
3	M	38	GLN
3	L	24	GLN
1	B	177	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)

8 monosaccharides are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

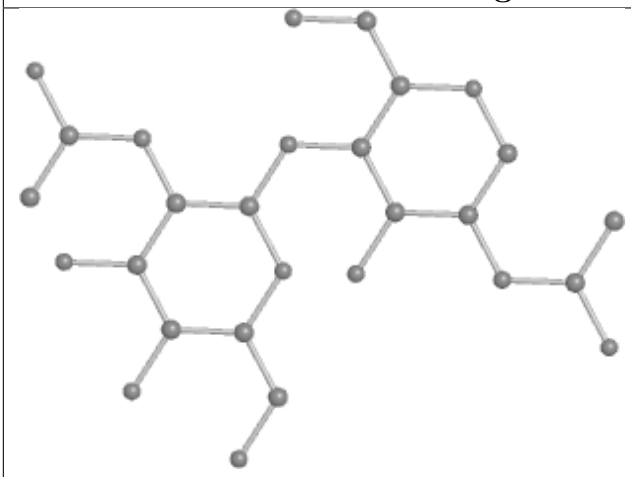
There are no torsion outliers.

There are no ring outliers.

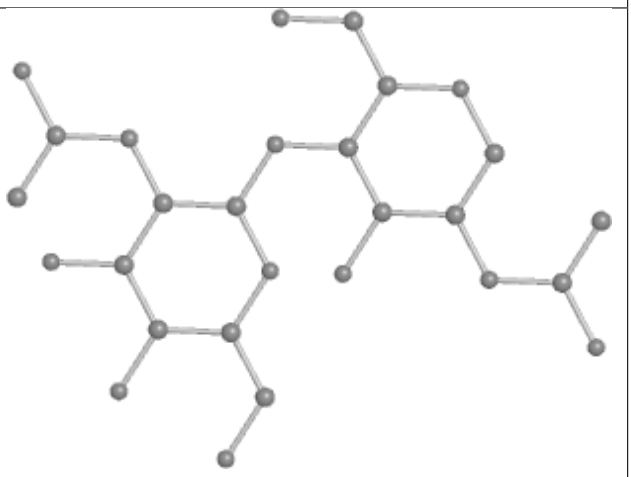
No monomer is involved in short contacts.

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

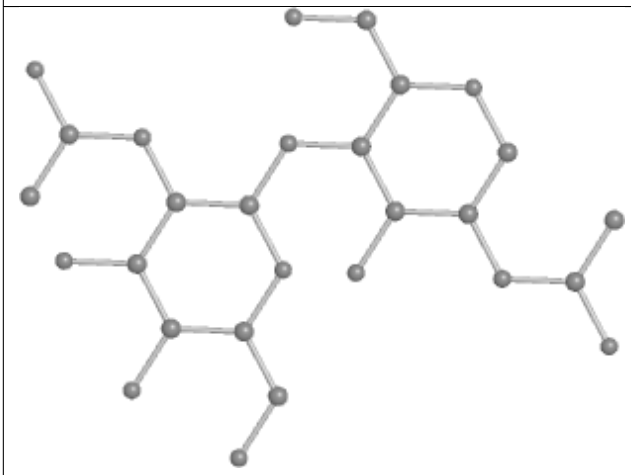
Oligosaccharide Chain X



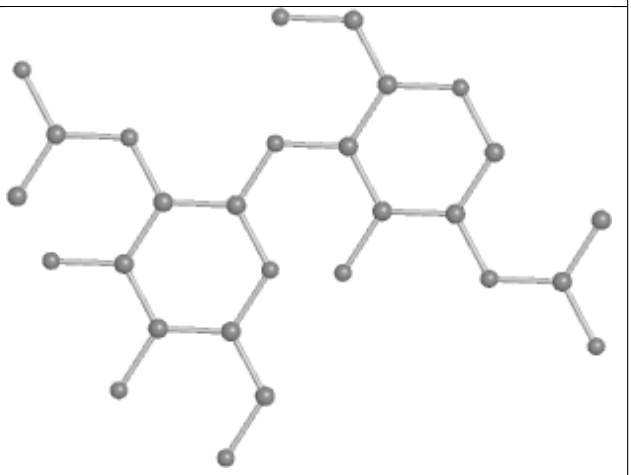
Bond lengths



Bond angles

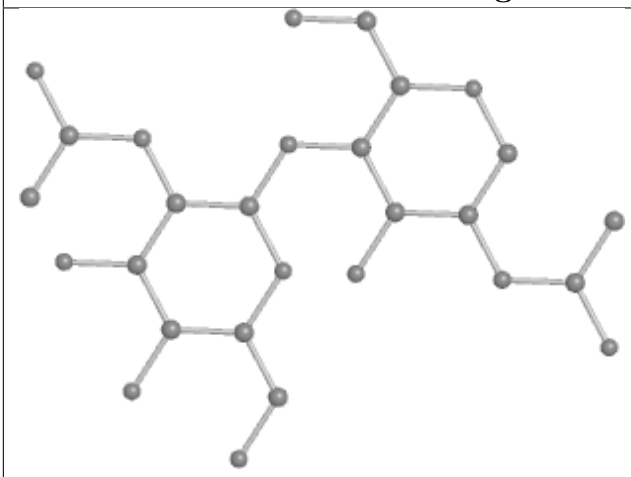


Torsions

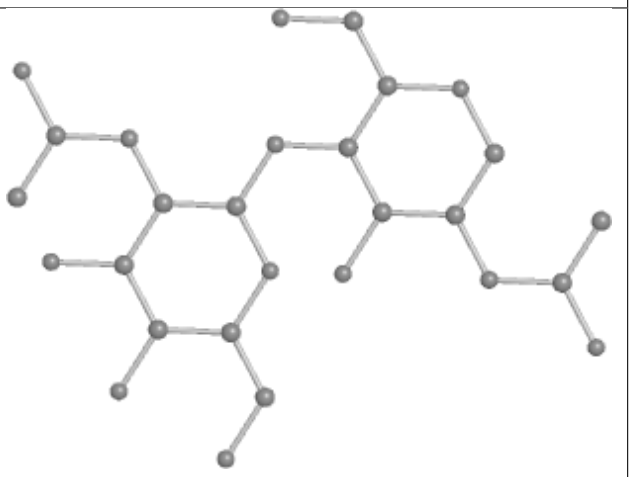


Rings

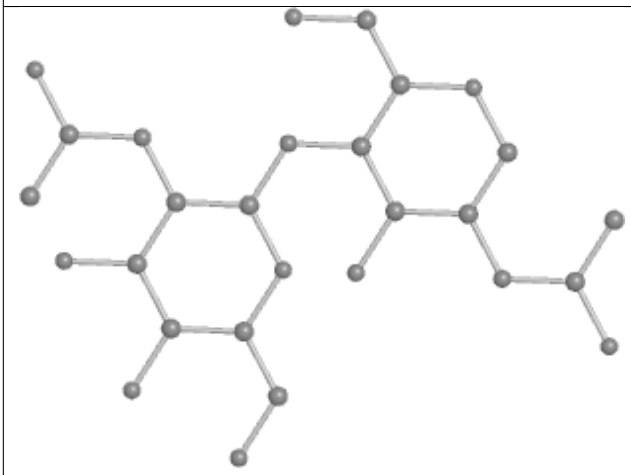
Oligosaccharide Chain C



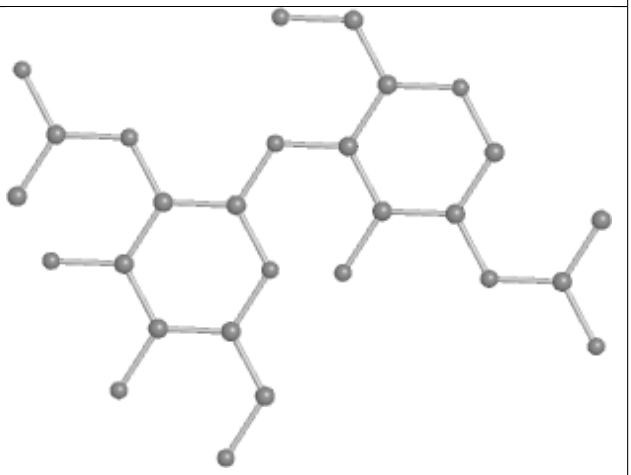
Bond lengths



Bond angles

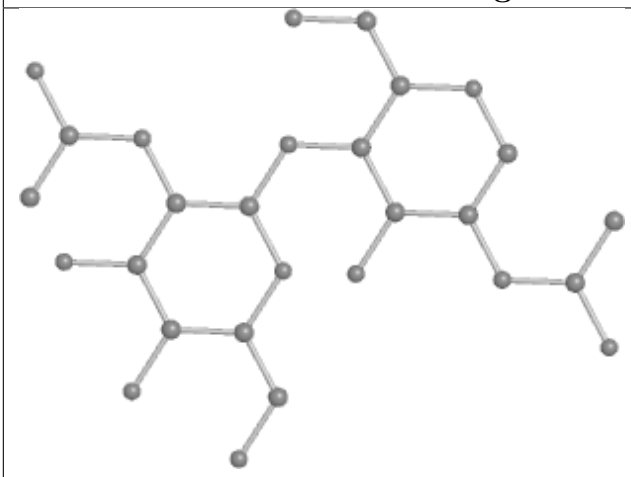


Torsions

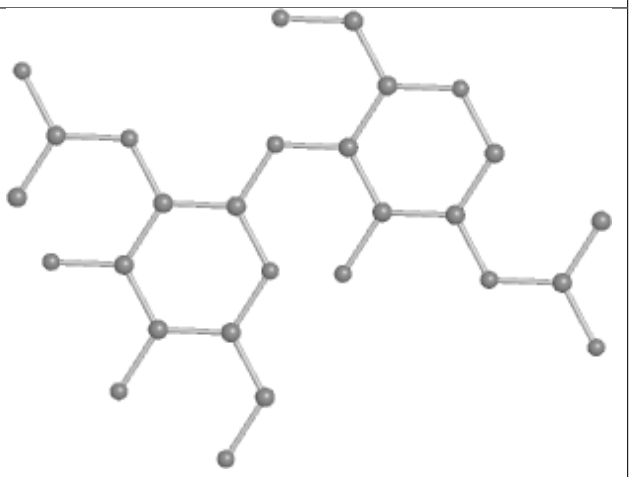


Rings

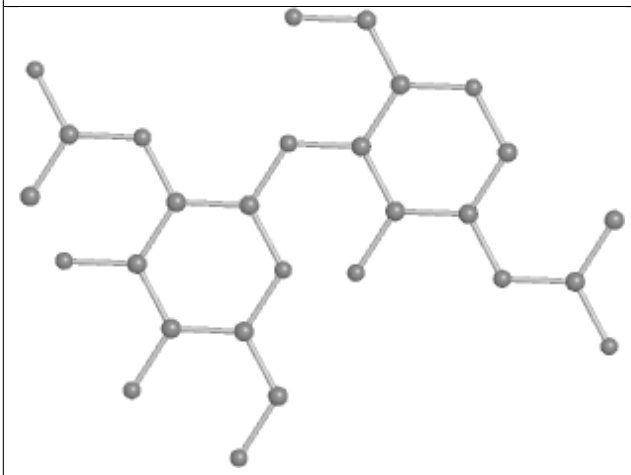
Oligosaccharide Chain D



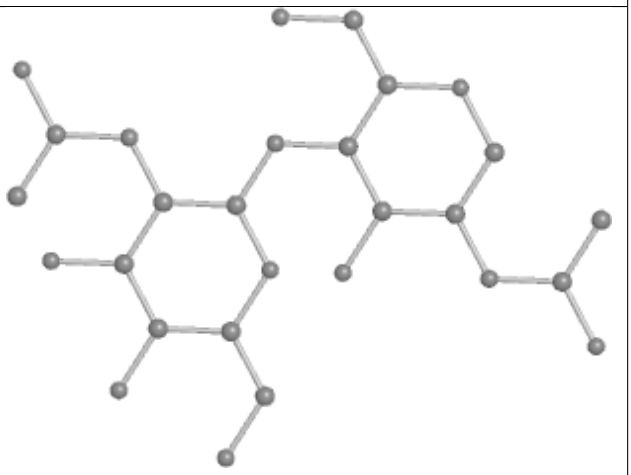
Bond lengths



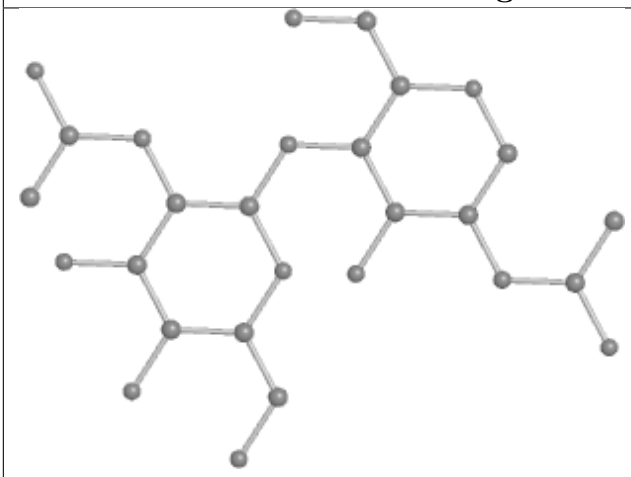
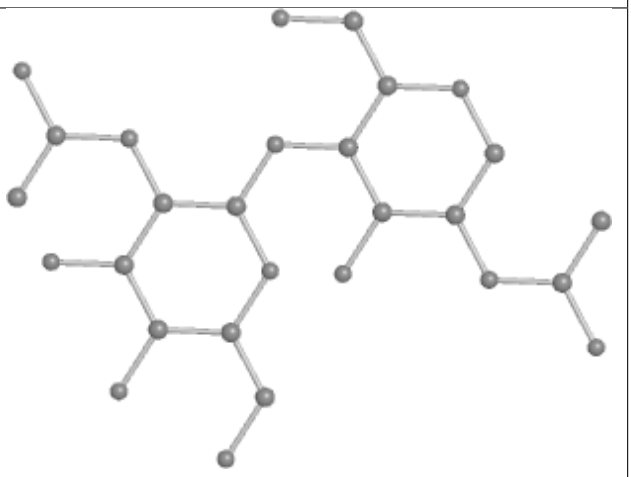
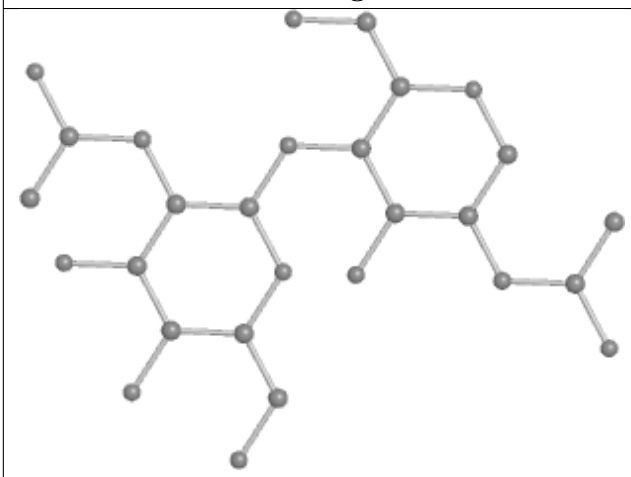
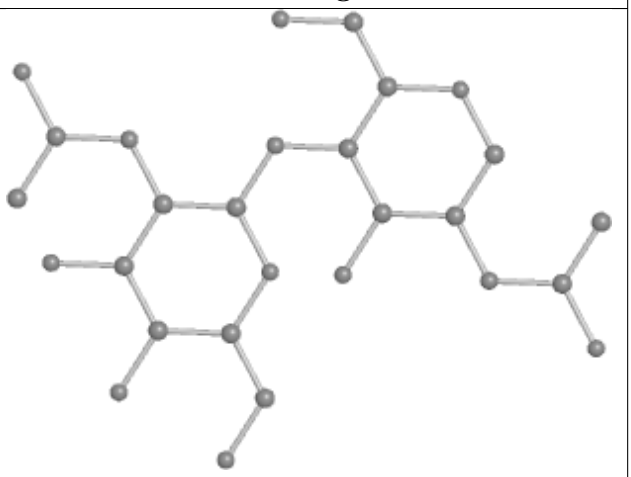
Bond angles



Torsions



Rings

Oligosaccharide Chain E	
 Bond lengths	 Bond angles
 Torsions	 Rings

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å^2)	Q<0.9
1	A	303/369 (82%)	1.48	74 (24%)	2 2	52, 77, 107, 114	0
1	B	317/369 (85%)	1.32	71 (22%)	3 3	44, 72, 105, 118	0
2	H	122/227 (53%)	1.11	15 (12%)	9 9	46, 77, 98, 106	0
2	I	121/227 (53%)	1.82	46 (38%)	1 1	62, 90, 106, 110	0
3	L	107/215 (49%)	0.73	8 (7%)	22 21	44, 65, 84, 87	0
3	M	107/215 (49%)	1.59	27 (25%)	2 2	65, 84, 104, 112	0
All	All	1077/1622 (66%)	1.37	241 (22%)	3 3	44, 77, 105, 118	0

The worst 5 of 241 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	344	ASN	10.1
1	A	95	PRO	8.2
1	A	343	TYR	7.1
1	B	280	LEU	6.7
2	I	11	VAL	6.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

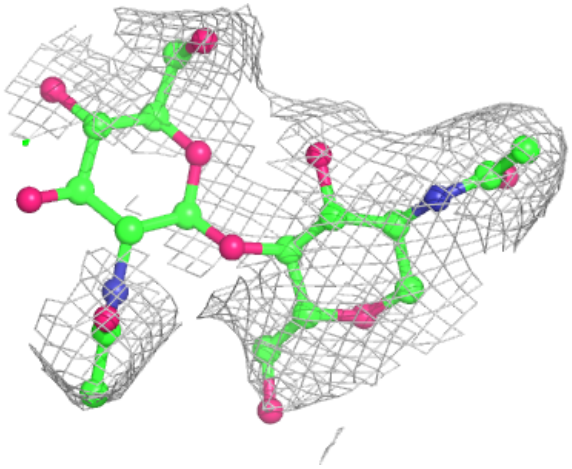
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

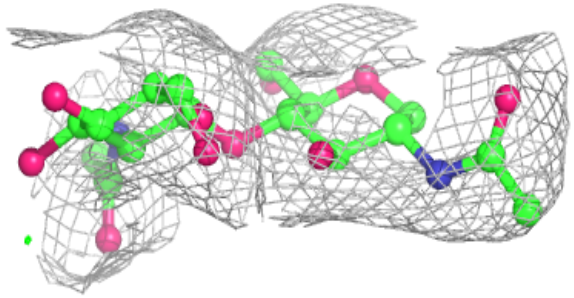
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	X	1	14/15	-	-	84,86,88,89	0
4	NAG	X	2	14/15	-	-	90,91,93,93	0
4	NAG	C	1	14/15	-	-	107,108,109,110	0
4	NAG	C	2	14/15	-	-	111,112,112,112	0
4	NAG	D	1	14/15	-	-	93,95,97,99	0
4	NAG	D	2	14/15	-	-	101,102,103,104	0
4	NAG	E	1	14/15	-	-	106,107,109,111	0
4	NAG	E	2	14/15	-	-	112,113,113,114	0

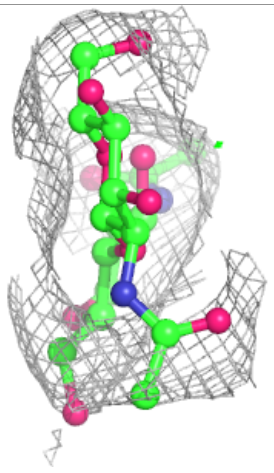
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain X:

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

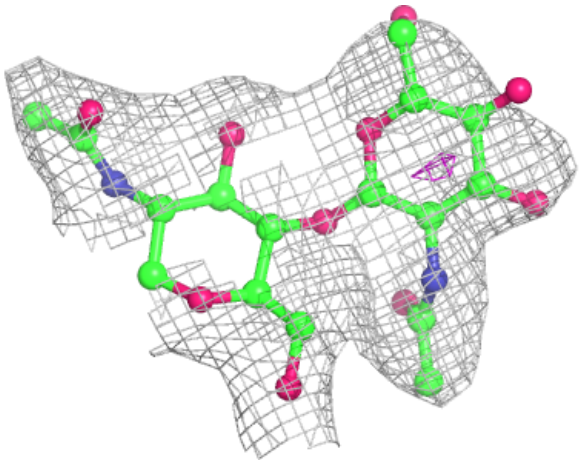


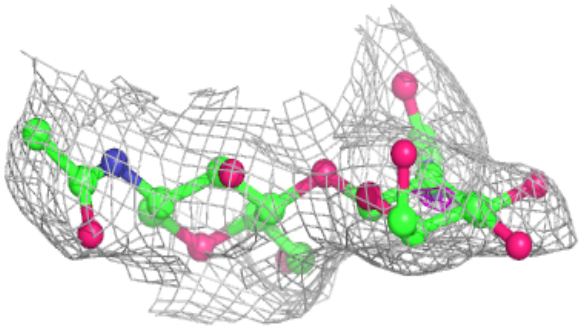


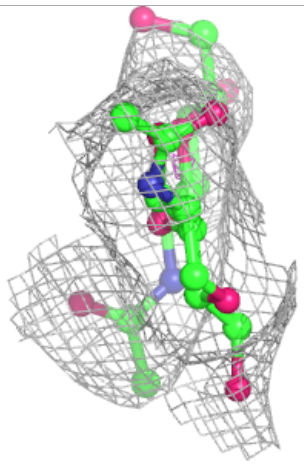


Electron density around Chain C:

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

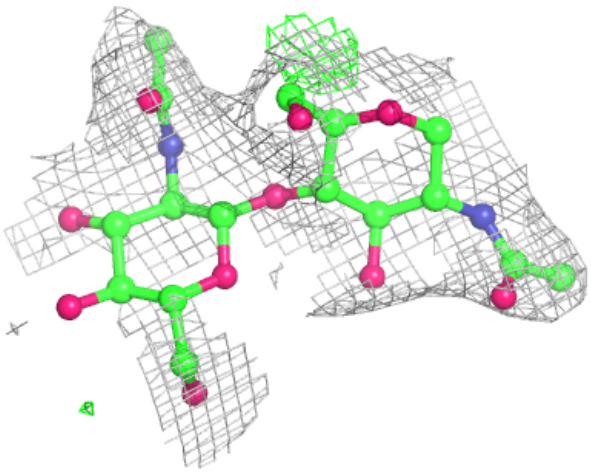


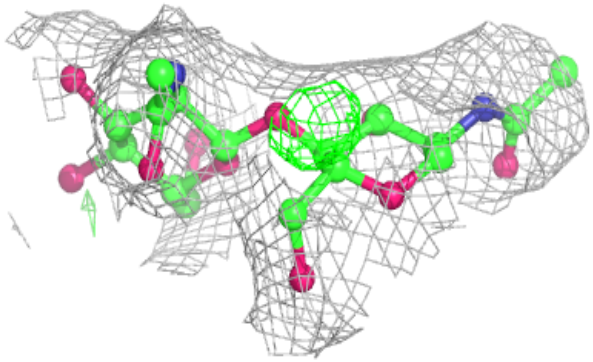


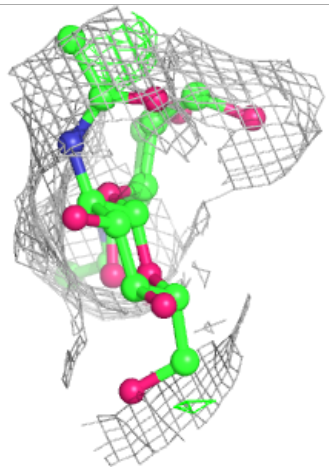


Electron density around Chain D:

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

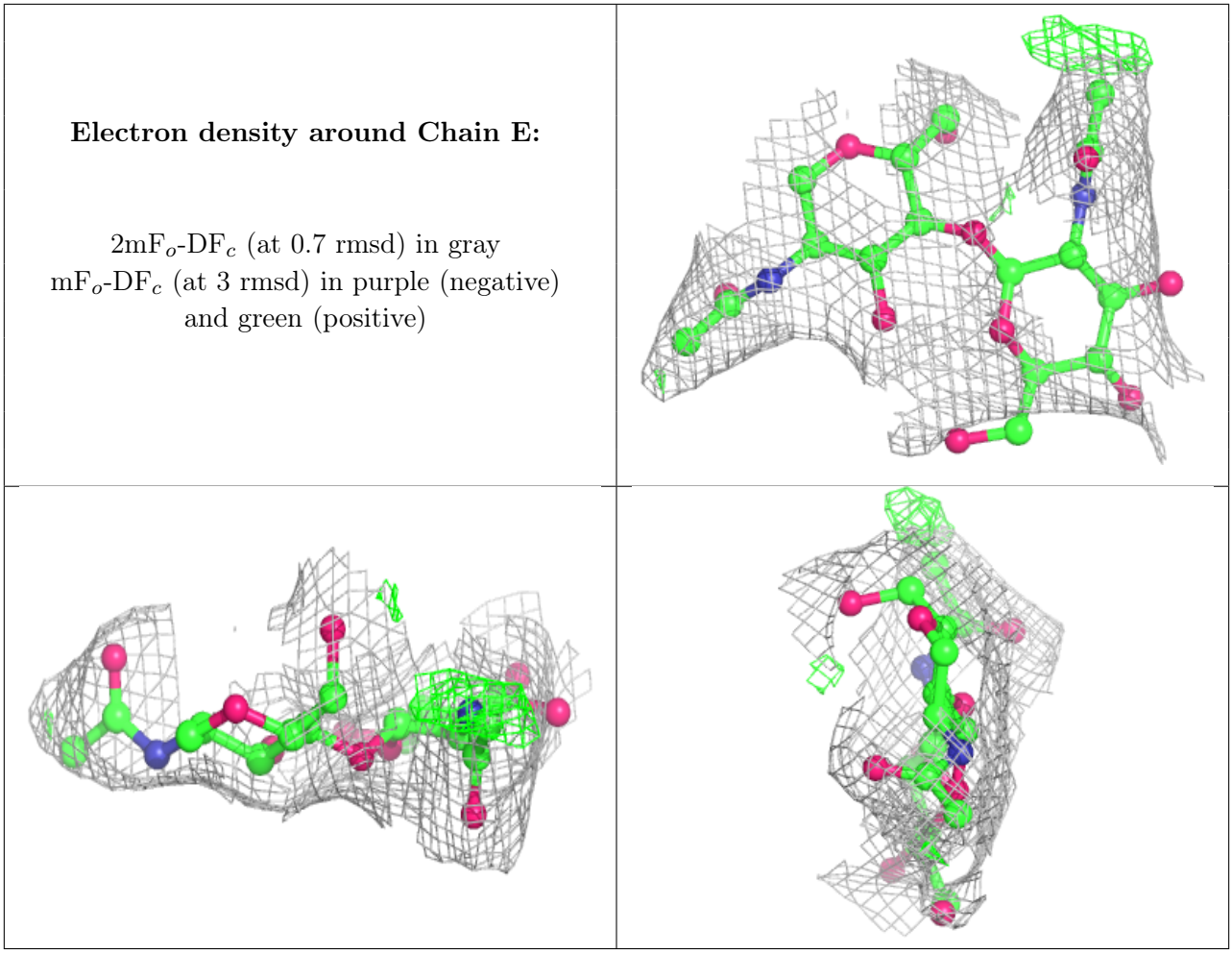






Electron density around Chain E:

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



6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.