Full wwPDB X-ray Structure Validation Report ⓘ

Apr 30, 2024 – 08:17 PM JST

PDB ID	:	8WG0
Title	:	Crystal structure of GH97 glucodextranase from Flavobacterium johnsoniae in complex with glucose
Authors	:	Kurata, R.; Nakamura, S.; Miyazaki, T.
Deposited on	:	2023-09-20
Resolution	:	1.95 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ 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 ⓘ) were used in the production of this report:

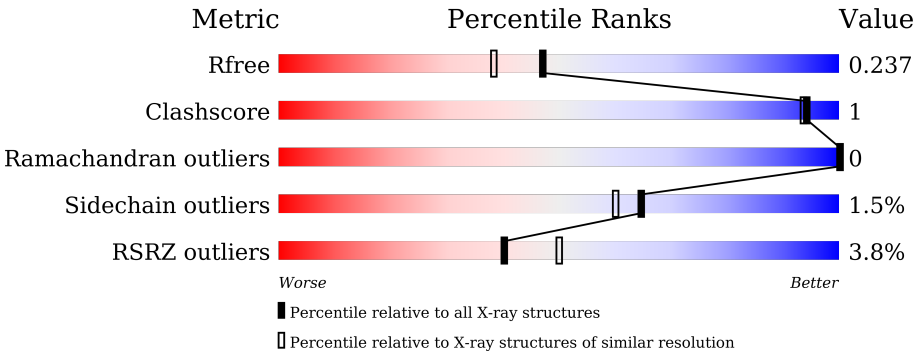
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	710	3% 90% 6% •
1	B	710	6% 89% 6% • 5%
1	C	710	% 90% 5% • •
1	D	710	5% 90% 5% • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 43981 atoms, of which 21162 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 Candidate alpha-glucosidase Glycoside hydrolase family 97.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	682	Total	C	H	N	O	S	165	1	0
			10749	3490	5273	918	1049	19			
1	B	678	Total	C	H	N	O	S	168	1	0
			10700	3477	5248	913	1043	19			
1	C	682	Total	C	H	N	O	S	166	1	0
			10759	3492	5280	920	1048	19			
1	D	684	Total	C	H	N	O	S	166	0	0
			10761	3494	5277	919	1052	19			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	initiating methionine	UNP A5FBI0
A	-1	GLY	-	expression tag	UNP A5FBI0
A	0	SER	-	expression tag	UNP A5FBI0
A	1	SER	-	expression tag	UNP A5FBI0
A	2	HIS	-	expression tag	UNP A5FBI0
A	3	HIS	-	expression tag	UNP A5FBI0
A	4	HIS	-	expression tag	UNP A5FBI0
A	5	HIS	-	expression tag	UNP A5FBI0
A	6	HIS	-	expression tag	UNP A5FBI0
A	7	HIS	-	expression tag	UNP A5FBI0
A	8	SER	-	expression tag	UNP A5FBI0
A	9	SER	-	expression tag	UNP A5FBI0
A	10	GLY	-	expression tag	UNP A5FBI0
A	11	LEU	-	expression tag	UNP A5FBI0
A	12	VAL	-	expression tag	UNP A5FBI0
A	13	PRO	-	expression tag	UNP A5FBI0
A	14	ARG	-	expression tag	UNP A5FBI0
A	15	GLY	-	expression tag	UNP A5FBI0
A	16	SER	-	expression tag	UNP A5FBI0
A	17	HIS	-	expression tag	UNP A5FBI0
A	18	MET	-	expression tag	UNP A5FBI0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	ALA	-	expression tag	UNP A5FBI0
A	20	SER	-	expression tag	UNP A5FBI0
B	-2	MET	-	initiating methionine	UNP A5FBI0
B	-1	GLY	-	expression tag	UNP A5FBI0
B	0	SER	-	expression tag	UNP A5FBI0
B	1	SER	-	expression tag	UNP A5FBI0
B	2	HIS	-	expression tag	UNP A5FBI0
B	3	HIS	-	expression tag	UNP A5FBI0
B	4	HIS	-	expression tag	UNP A5FBI0
B	5	HIS	-	expression tag	UNP A5FBI0
B	6	HIS	-	expression tag	UNP A5FBI0
B	7	HIS	-	expression tag	UNP A5FBI0
B	8	SER	-	expression tag	UNP A5FBI0
B	9	SER	-	expression tag	UNP A5FBI0
B	10	GLY	-	expression tag	UNP A5FBI0
B	11	LEU	-	expression tag	UNP A5FBI0
B	12	VAL	-	expression tag	UNP A5FBI0
B	13	PRO	-	expression tag	UNP A5FBI0
B	14	ARG	-	expression tag	UNP A5FBI0
B	15	GLY	-	expression tag	UNP A5FBI0
B	16	SER	-	expression tag	UNP A5FBI0
B	17	HIS	-	expression tag	UNP A5FBI0
B	18	MET	-	expression tag	UNP A5FBI0
B	19	ALA	-	expression tag	UNP A5FBI0
B	20	SER	-	expression tag	UNP A5FBI0
C	-2	MET	-	initiating methionine	UNP A5FBI0
C	-1	GLY	-	expression tag	UNP A5FBI0
C	0	SER	-	expression tag	UNP A5FBI0
C	1	SER	-	expression tag	UNP A5FBI0
C	2	HIS	-	expression tag	UNP A5FBI0
C	3	HIS	-	expression tag	UNP A5FBI0
C	4	HIS	-	expression tag	UNP A5FBI0
C	5	HIS	-	expression tag	UNP A5FBI0
C	6	HIS	-	expression tag	UNP A5FBI0
C	7	HIS	-	expression tag	UNP A5FBI0
C	8	SER	-	expression tag	UNP A5FBI0
C	9	SER	-	expression tag	UNP A5FBI0
C	10	GLY	-	expression tag	UNP A5FBI0
C	11	LEU	-	expression tag	UNP A5FBI0
C	12	VAL	-	expression tag	UNP A5FBI0
C	13	PRO	-	expression tag	UNP A5FBI0
C	14	ARG	-	expression tag	UNP A5FBI0

Continued on next page...

Continued from previous page...

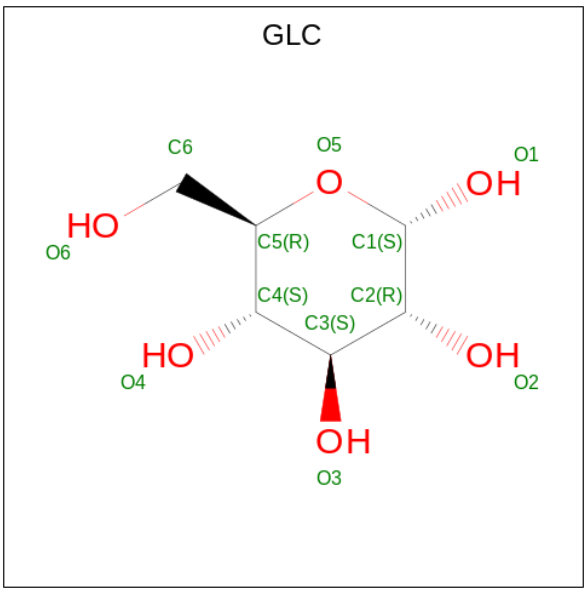
Chain	Residue	Modelled	Actual	Comment	Reference
C	15	GLY	-	expression tag	UNP A5FBI0
C	16	SER	-	expression tag	UNP A5FBI0
C	17	HIS	-	expression tag	UNP A5FBI0
C	18	MET	-	expression tag	UNP A5FBI0
C	19	ALA	-	expression tag	UNP A5FBI0
C	20	SER	-	expression tag	UNP A5FBI0
D	-2	MET	-	initiating methionine	UNP A5FBI0
D	-1	GLY	-	expression tag	UNP A5FBI0
D	0	SER	-	expression tag	UNP A5FBI0
D	1	SER	-	expression tag	UNP A5FBI0
D	2	HIS	-	expression tag	UNP A5FBI0
D	3	HIS	-	expression tag	UNP A5FBI0
D	4	HIS	-	expression tag	UNP A5FBI0
D	5	HIS	-	expression tag	UNP A5FBI0
D	6	HIS	-	expression tag	UNP A5FBI0
D	7	HIS	-	expression tag	UNP A5FBI0
D	8	SER	-	expression tag	UNP A5FBI0
D	9	SER	-	expression tag	UNP A5FBI0
D	10	GLY	-	expression tag	UNP A5FBI0
D	11	LEU	-	expression tag	UNP A5FBI0
D	12	VAL	-	expression tag	UNP A5FBI0
D	13	PRO	-	expression tag	UNP A5FBI0
D	14	ARG	-	expression tag	UNP A5FBI0
D	15	GLY	-	expression tag	UNP A5FBI0
D	16	SER	-	expression tag	UNP A5FBI0
D	17	HIS	-	expression tag	UNP A5FBI0
D	18	MET	-	expression tag	UNP A5FBI0
D	19	ALA	-	expression tag	UNP A5FBI0
D	20	SER	-	expression tag	UNP A5FBI0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

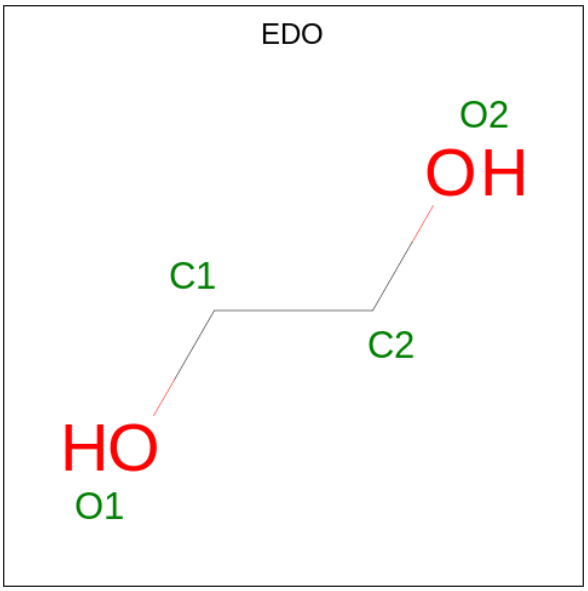
- Molecule 3 is alpha-D-glucopyranose (three-letter code: GLC) (formula: $C_6H_{12}O_6$) (labeled

as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 24 6 12 6	3	0
3	B	1	Total C H O 24 6 12 6	3	0
3	C	1	Total C H O 24 6 12 6	3	0
3	C	1	Total C H O 24 6 12 6	4	0
3	D	1	Total C H O 24 6 12 6	3	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 10 2 6 2	1	0
4	A	1	Total C H O 10 2 6 2	1	0
4	C	1	Total C H O 10 2 6 2	1	0
4	D	1	Total C H O 10 2 6 2	1	0

- Molecule 5 is water.

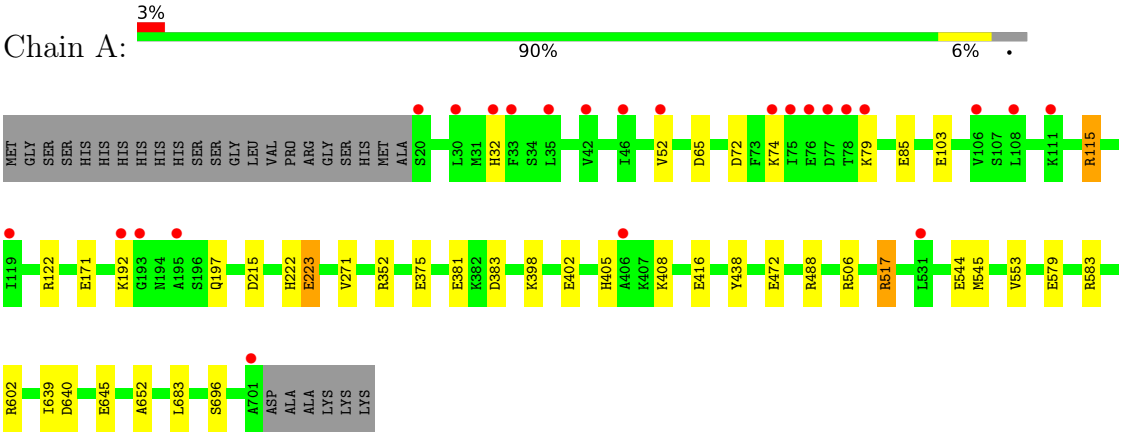
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	250	Total O 250 250	0	0
5	B	165	Total O 165 165	0	0
5	C	277	Total O 277 277	0	0
5	D	156	Total O 156 156	0	0

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.

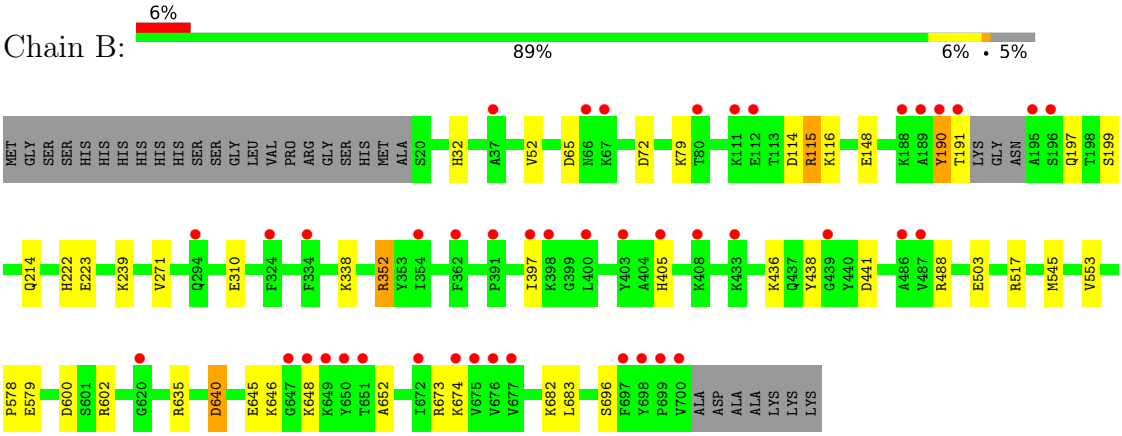
• Molecule 1: Candidate alpha-glucosidase Glycoside hydrolase family 97

Chain A:



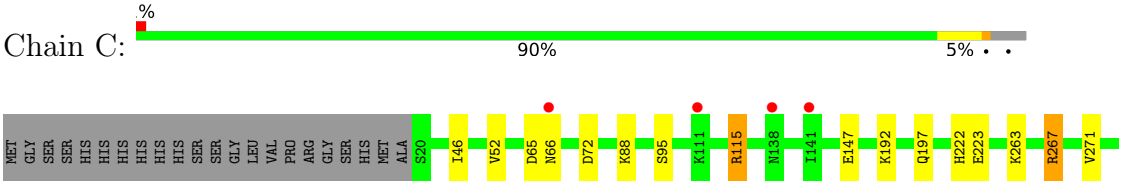
• Molecule 1: Candidate alpha-glucosidase Glycoside hydrolase family 97

Chain B:



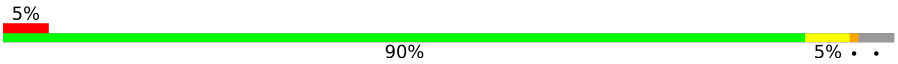
• Molecule 1: Candidate alpha-glucosidase Glycoside hydrolase family 97

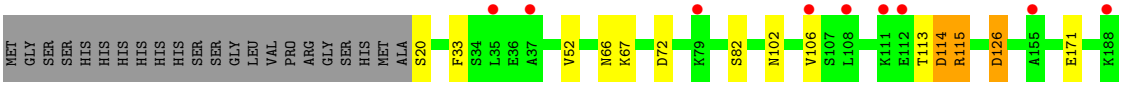
Chain C:

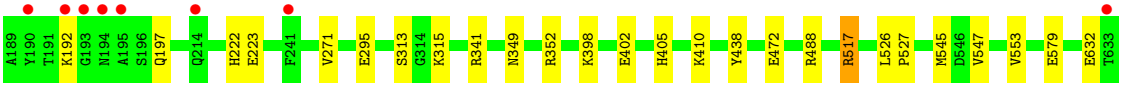


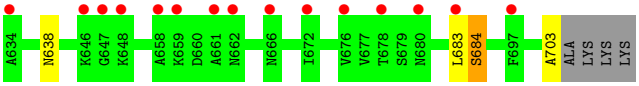
• Molecule 1: Candidate alpha-glucosidase Glycoside hydrolase family 97

Chain D:









4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	307.93Å 103.69Å 95.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.13 – 1.95 49.13 – 1.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.13-1.95) 100.0 (49.13-1.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.64 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0405	Depositor
R, R_{free}	0.189 , 0.229 0.201 , 0.237	Depositor DCC
R_{free} test set	10922 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	31.5	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 35.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	43981	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.88% 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: EDO, GLC, CA

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	0.79	12/5623 (0.2%)	0.98	10/7626 (0.1%)
1	B	0.68	2/5599 (0.0%)	0.98	9/7593 (0.1%)
1	C	0.74	5/5626 (0.1%)	0.97	8/7629 (0.1%)
1	D	0.70	6/5628 (0.1%)	0.94	4/7633 (0.1%)
All	All	0.73	25/22476 (0.1%)	0.97	31/30481 (0.1%)

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
1	A	0	3
1	B	0	2
1	C	0	4
1	D	0	2
All	All	0	11

All (25) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	645	GLU	CD-OE1	7.79	1.34	1.25
1	D	472	GLU	CD-OE2	-7.64	1.17	1.25
1	A	85	GLU	CD-OE2	7.64	1.34	1.25
1	A	223	GLU	CD-OE1	7.53	1.33	1.25
1	A	579	GLU	CD-OE1	7.43	1.33	1.25
1	C	605	GLU	CD-OE1	-7.20	1.17	1.25
1	D	632	GLU	CD-OE2	-7.05	1.17	1.25
1	D	171	GLU	CD-OE2	6.56	1.32	1.25
1	A	645	GLU	CD-OE2	6.53	1.32	1.25
1	A	381	GLU	CD-OE2	-6.50	1.18	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	503	GLU	CD-OE2	6.48	1.32	1.25
1	A	103	GLU	CD-OE1	6.41	1.32	1.25
1	A	375	GLU	CD-OE2	6.06	1.32	1.25
1	D	295	GLU	CD-OE1	-6.00	1.19	1.25
1	C	381	GLU	CD-OE2	-6.00	1.19	1.25
1	C	416	GLU	CD-OE1	-5.98	1.19	1.25
1	A	544	GLU	CD-OE2	5.97	1.32	1.25
1	D	632	GLU	CD-OE1	-5.88	1.19	1.25
1	A	171	GLU	CD-OE1	5.34	1.31	1.25
1	A	472	GLU	CD-OE2	-5.29	1.19	1.25
1	A	416	GLU	CD-OE1	-5.22	1.20	1.25
1	C	383	ASP	CG-OD1	5.16	1.37	1.25
1	D	472	GLU	CD-OE1	-5.12	1.20	1.25
1	B	148	GLU	CD-OE1	5.06	1.31	1.25
1	C	147	GLU	CD-OE2	5.00	1.31	1.25

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	579	GLU	CB-CA-C	-9.71	90.97	110.40
1	C	341	ARG	CG-CD-NE	-8.54	93.86	111.80
1	D	579	GLU	CB-CA-C	-8.43	93.54	110.40
1	B	115	ARG	NE-CZ-NH2	-8.16	116.22	120.30
1	C	579	GLU	CB-CA-C	-7.79	94.82	110.40
1	C	383	ASP	CB-CG-OD2	-7.65	111.42	118.30
1	B	32[A]	HIS	CA-CB-CG	7.10	125.67	113.60
1	B	32[B]	HIS	CA-CB-CG	7.10	125.67	113.60
1	A	215	ASP	CB-CG-OD1	7.06	124.65	118.30
1	A	517	ARG	NE-CZ-NH2	6.98	123.79	120.30
1	D	703	ALA	CA-C-O	-6.87	105.68	120.10
1	C	115	ARG	NE-CZ-NH2	-6.85	116.87	120.30
1	A	215	ASP	CB-CG-OD2	-6.85	112.14	118.30
1	C	341	ARG	NE-CZ-NH1	-6.50	117.05	120.30
1	D	341	ARG	CG-CD-NE	-6.41	98.35	111.80
1	B	640	ASP	CB-CA-C	6.25	122.89	110.40
1	A	115	ARG	NE-CZ-NH2	-6.16	117.22	120.30
1	C	583	ARG	NE-CZ-NH2	6.08	123.34	120.30
1	B	115	ARG	NE-CZ-NH1	6.06	123.33	120.30
1	D	115	ARG	CG-CD-NE	-6.01	99.18	111.80
1	A	32	HIS	CA-CB-CG	5.99	123.77	113.60
1	A	383	ASP	CB-CG-OD2	-5.95	112.94	118.30
1	A	122	ARG	NE-CZ-NH1	5.95	123.28	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	506	ARG	NE-CZ-NH2	5.90	123.25	120.30
1	C	635	ARG	NE-CZ-NH1	5.51	123.06	120.30
1	B	441	ASP	CB-CG-OD2	-5.38	113.46	118.30
1	A	640	ASP	N-CA-CB	-5.38	100.92	110.60
1	A	115	ARG	CG-CD-NE	5.27	122.88	111.80
1	B	352	ARG	NE-CZ-NH1	5.11	122.85	120.30
1	C	267	ARG	NE-CZ-NH1	5.10	122.85	120.30
1	B	635	ARG	NE-CZ-NH2	-5.00	117.80	120.30

There are no chirality outliers.

All (11) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	352	ARG	Sidechain
1	A	583	ARG	Sidechain
1	A	602	ARG	Sidechain
1	B	190	TYR	Peptide
1	B	352	ARG	Sidechain
1	C	341	ARG	Sidechain
1	C	352[A]	ARG	Sidechain
1	C	352[B]	ARG	Sidechain
1	C	583	ARG	Sidechain
1	D	115	ARG	Sidechain
1	D	517	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	5476	5273	5257	10	0
1	B	5452	5248	5230	15	0
1	C	5479	5280	5264	16	0
1	D	5484	5277	5260	16	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	12	12	11	0	0
3	B	12	12	10	0	0
3	C	24	24	23	3	0
3	D	12	12	10	0	0
4	A	8	12	12	0	0
4	C	4	6	6	0	0
4	D	4	6	6	0	0
5	A	250	0	0	1	0
5	B	165	0	0	2	0
5	C	277	0	0	2	0
5	D	156	0	0	1	0
All	All	22819	21162	21089	57	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 1.

All (57) 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:639:ILE:HD12	1:A:683:LEU:HB3	1.51	0.90
1:B:190:TYR:CE2	1:B:199:SER:HB3	2.21	0.75
1:B:52:VAL:HG11	1:B:271:VAL:HG21	1.75	0.68
1:B:600:ASP:OD2	1:B:602:ARG:NH1	2.29	0.64
1:B:65:ASP:OD2	1:B:115:ARG:NH2	2.34	0.60
1:D:192:LYS:HA	1:D:197:GLN:HE22	1.68	0.59
1:A:65:ASP:OD2	1:A:115:ARG:NH2	2.34	0.59
1:B:114:ASP:O	1:B:116:LYS:HE2	2.03	0.59
1:C:65:ASP:OD2	1:C:115:ARG:NH2	2.35	0.58
1:A:74:LYS:HE3	5:A:1102:HOH:O	2.07	0.55
1:C:192:LYS:HA	1:C:197:GLN:HE22	1.72	0.55
1:B:397:ILE:HG22	5:B:1037:HOH:O	2.04	0.55
1:C:597:ASP:OD2	3:C:804:GLC:H62	2.06	0.55
1:C:52:VAL:HG11	1:C:271:VAL:HG21	1.88	0.54
1:B:190:TYR:CE1	1:B:197:GLN:OE1	2.61	0.54
1:C:545:MET:HG3	1:C:553:VAL:HB	1.90	0.54
1:D:82:SER:OG	1:D:102:ASN:OD1	2.26	0.53
1:D:52:VAL:HG11	1:D:271:VAL:HG21	1.90	0.53
1:C:544:GLU:OE2	1:C:583:ARG:NH2	2.42	0.52
1:B:640:ASP:OD1	1:B:682:LYS:HE2	2.09	0.52
1:A:192:LYS:HA	1:A:197:GLN:HE22	1.75	0.51
1:A:398:LYS:HE3	1:A:402:GLU:OE1	2.11	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:545:MET:HG3	1:A:553:VAL:HB	1.95	0.49
1:B:190:TYR:CZ	1:B:197:GLN:OE1	2.67	0.48
1:D:405:HIS:HE1	1:D:438:TYR:O	1.95	0.48
1:C:299:ILE:O	3:C:804:GLC:H3	2.14	0.48
1:A:52:VAL:HG11	1:A:271:VAL:HG21	1.96	0.47
1:B:645:GLU:OE2	1:B:648:LYS:HD2	2.16	0.46
1:C:46:ILE:HG22	5:C:1100:HOH:O	2.16	0.46
1:C:301:PRO:HD2	3:C:804:GLC:H5	1.97	0.46
1:D:126:ASP:N	1:D:126:ASP:OD1	2.49	0.46
1:D:349:ASN:O	1:D:352:ARG:HB2	2.16	0.46
1:C:683:LEU:HD23	1:C:683:LEU:HA	1.84	0.44
1:A:222:HIS:CG	1:A:223:GLU:H	2.36	0.44
1:D:545:MET:HG3	1:D:553:VAL:HB	2.00	0.44
1:D:398:LYS:HE3	1:D:402:GLU:OE1	2.18	0.43
1:B:310:GLU:HG3	5:B:1025:HOH:O	2.18	0.43
1:C:222:HIS:CG	1:C:223:GLU:H	2.36	0.43
1:C:422:ARG:O	1:C:426:ARG:HG3	2.19	0.43
1:A:652:ALA:HA	1:A:696:SER:O	2.20	0.42
1:B:405:HIS:HE1	1:B:438:TYR:O	2.03	0.42
1:C:444:LYS:HE3	1:C:484:HIS:CE1	2.55	0.42
1:D:33:PHE:CE2	1:D:106:VAL:HG11	2.55	0.42
1:D:66:ASN:OD1	1:D:67:LYS:N	2.53	0.41
1:B:545:MET:HG3	1:B:553:VAL:HB	2.01	0.41
1:D:113:THR:O	1:D:114:ASP:HB2	2.20	0.41
1:B:222:HIS:CG	1:B:223:GLU:H	2.37	0.41
1:B:652:ALA:HA	1:B:696:SER:O	2.20	0.41
1:C:88:LYS:HE3	1:C:95:SER:O	2.20	0.41
1:A:405:HIS:HE1	1:A:438:TYR:O	2.04	0.41
1:D:638:ASN:OD1	1:D:684:SER:HB2	2.21	0.41
1:D:526:LEU:N	1:D:527:PRO:CD	2.84	0.40
1:D:313:SER:HB2	1:D:547:VAL:HB	2.04	0.40
1:D:410:LYS:HE3	5:D:918:HOH:O	2.21	0.40
1:C:652:ALA:HA	1:C:696:SER:O	2.22	0.40
1:C:310:GLU:HG3	5:C:1044:HOH:O	2.20	0.40
1:D:222:HIS:CG	1:D:223:GLU:H	2.39	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	681/710 (96%)	652 (96%)	29 (4%)	0	100	100
1	B	675/710 (95%)	647 (96%)	28 (4%)	0	100	100
1	C	681/710 (96%)	646 (95%)	35 (5%)	0	100	100
1	D	682/710 (96%)	648 (95%)	34 (5%)	0	100	100
All	All	2719/2840 (96%)	2593 (95%)	126 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	582/603 (96%)	577 (99%)	5 (1%)	78	77
1	B	580/603 (96%)	566 (98%)	14 (2%)	49	40
1	C	582/603 (96%)	575 (99%)	7 (1%)	71	68
1	D	582/603 (96%)	573 (98%)	9 (2%)	65	60
All	All	2326/2412 (96%)	2291 (98%)	35 (2%)	65	60

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	72	ASP
1	A	79	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	408	LYS
1	A	488	ARG
1	A	517	ARG
1	B	72	ASP
1	B	79	LYS
1	B	191	THR
1	B	214	GLN
1	B	239	LYS
1	B	338	LYS
1	B	436	LYS
1	B	488	ARG
1	B	517	ARG
1	B	578	PRO
1	B	646	LYS
1	B	673	ARG
1	B	674	LYS
1	B	683	LEU
1	C	66	ASN
1	C	72	ASP
1	C	263	LYS
1	C	267	ARG
1	C	488	ARG
1	C	517	ARG
1	C	579	GLU
1	D	20	SER
1	D	72	ASP
1	D	114	ASP
1	D	126	ASP
1	D	315	LYS
1	D	488	ARG
1	D	517	ARG
1	D	683	LEU
1	D	684	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	197	GLN
1	A	405	HIS
1	B	66	ASN
1	B	86	ASN
1	B	197	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	405	HIS
1	B	468	GLN
1	C	197	GLN
1	D	197	GLN

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)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 13 ligands modelled in this entry, 4 are monoatomic - leaving 9 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	$\#\|Z\| > 2$
3	GLC	A	802	2	12,12,12	1.82	2 (16%)	17,17,17	2.18	6 (35%)
3	GLC	C	802	2	12,12,12	1.96	2 (16%)	17,17,17	2.19	4 (23%)
4	EDO	D	803	-	3,3,3	0.46	0	2,2,2	0.75	0
3	GLC	B	802	2	12,12,12	1.29	1 (8%)	17,17,17	1.84	5 (29%)
4	EDO	C	803	-	3,3,3	0.25	0	2,2,2	0.43	0
3	GLC	C	804	-	12,12,12	2.24	6 (50%)	17,17,17	1.70	4 (23%)
3	GLC	D	802	2	12,12,12	1.65	3 (25%)	17,17,17	1.81	5 (29%)
4	EDO	A	804	-	3,3,3	0.61	0	2,2,2	0.56	0
4	EDO	A	803	-	3,3,3	0.73	0	2,2,2	0.36	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GLC	A	802	2	-	0/2/22/22	0/1/1/1
3	GLC	C	802	2	-	0/2/22/22	0/1/1/1
4	EDO	D	803	-	-	0/1/1/1	-
3	GLC	B	802	2	-	0/2/22/22	0/1/1/1
4	EDO	C	803	-	-	0/1/1/1	-
3	GLC	C	804	-	-	0/2/22/22	0/1/1/1
3	GLC	D	802	2	-	0/2/22/22	0/1/1/1
4	EDO	A	804	-	-	0/1/1/1	-
4	EDO	A	803	-	-	0/1/1/1	-

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	802	GLC	O1-C1	6.09	1.59	1.39
3	A	802	GLC	O1-C1	5.07	1.55	1.39
3	C	804	GLC	C3-C2	3.74	1.61	1.52
3	D	802	GLC	O1-C1	3.46	1.50	1.39
3	C	804	GLC	O5-C5	3.30	1.52	1.44
3	C	804	GLC	C1-C2	3.24	1.60	1.52
3	A	802	GLC	O5-C1	2.66	1.49	1.42
3	C	804	GLC	C4-C5	2.37	1.58	1.53
3	B	802	GLC	O5-C1	2.37	1.48	1.42
3	C	804	GLC	C4-C3	2.33	1.58	1.52
3	C	804	GLC	O5-C1	2.32	1.48	1.42
3	C	802	GLC	O5-C1	2.24	1.48	1.42
3	D	802	GLC	O4-C4	-2.09	1.38	1.43
3	D	802	GLC	O5-C1	2.01	1.47	1.42

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	802	GLC	O1-C1-O5	6.56	130.08	110.38
3	A	802	GLC	O1-C1-O5	5.85	127.94	110.38
3	B	802	GLC	O1-C1-O5	4.87	125.00	110.38
3	C	802	GLC	O5-C1-C2	-4.34	102.55	110.28
3	D	802	GLC	O1-C1-O5	4.19	122.94	110.38
3	A	802	GLC	O5-C1-C2	-4.06	103.03	110.28
3	C	804	GLC	O2-C2-C3	3.51	118.45	110.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	802	GLC	C4-C3-C2	-3.36	104.95	110.82
3	A	802	GLC	C1-O5-C5	3.25	119.80	113.66
3	D	802	GLC	O2-C2-C3	3.22	117.80	110.35
3	B	802	GLC	O5-C1-C2	-2.96	105.00	110.28
3	C	804	GLC	O4-C4-C3	2.93	117.12	110.35
3	B	802	GLC	O2-C2-C3	2.83	116.89	110.35
3	C	804	GLC	O3-C3-C2	2.82	116.86	110.35
3	C	804	GLC	O1-C1-C2	2.65	116.49	109.03
3	C	802	GLC	O2-C2-C3	2.52	116.18	110.35
3	B	802	GLC	C1-O5-C5	2.47	118.31	113.66
3	D	802	GLC	C1-O5-C5	2.45	118.29	113.66
3	A	802	GLC	C4-C3-C2	-2.18	107.02	110.82
3	A	802	GLC	O4-C4-C5	-2.15	103.97	109.30
3	C	802	GLC	C1-O5-C5	2.10	117.62	113.66
3	A	802	GLC	O3-C3-C4	2.09	115.18	110.35
3	B	802	GLC	O2-C2-C1	2.05	113.90	109.16
3	D	802	GLC	O4-C4-C5	-2.00	104.32	109.30

There are no chirality outliers.

There are no torsion outliers.

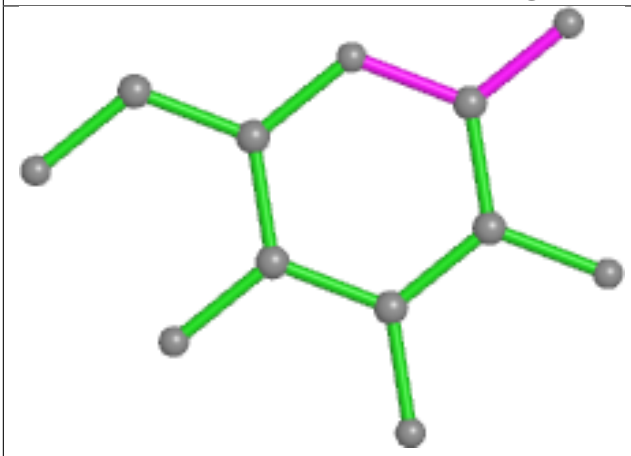
There are no ring outliers.

1 monomer is involved in 3 short contacts:

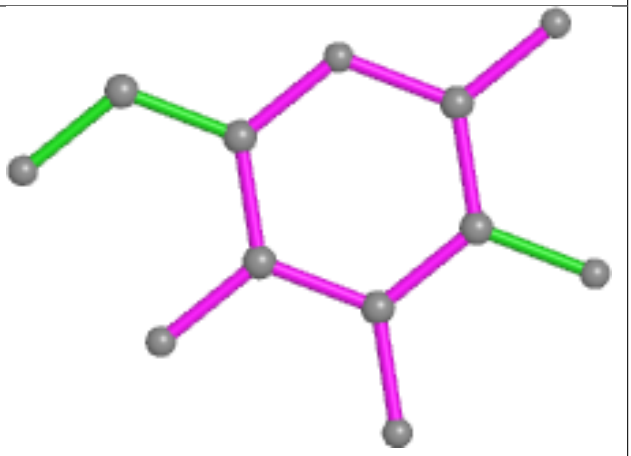
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	804	GLC	3	0

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.

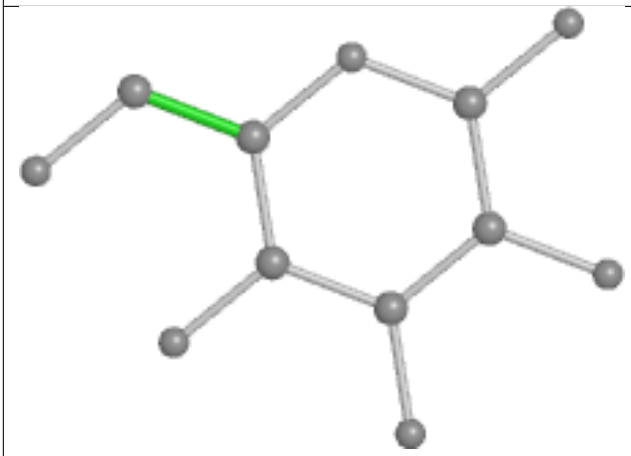
Ligand GLC A 802



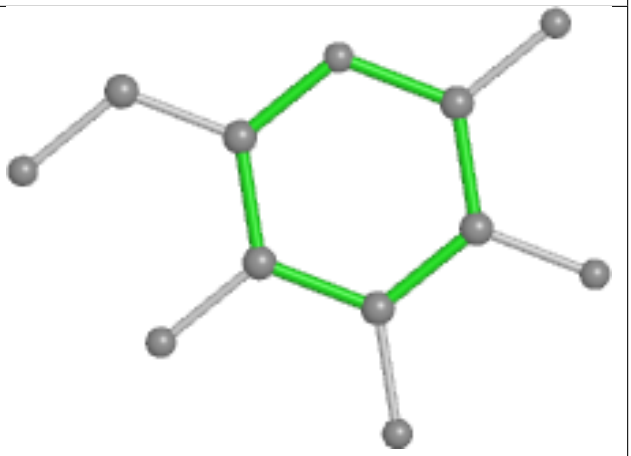
Bond lengths



Bond angles

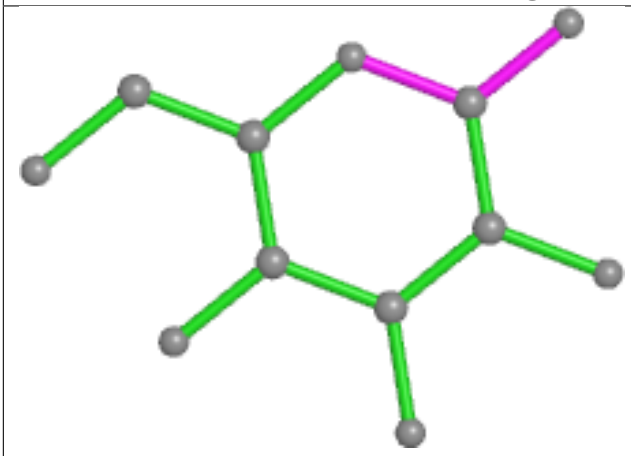


Torsions

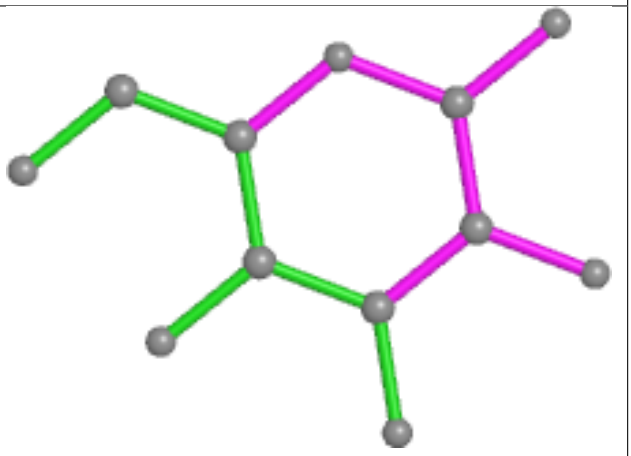


Rings

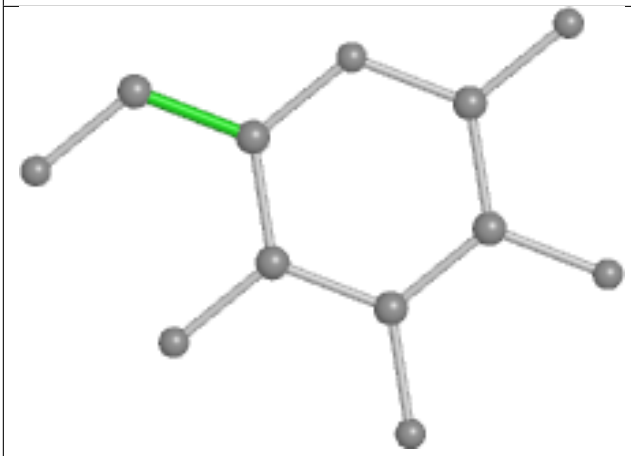
Ligand GLC C 802



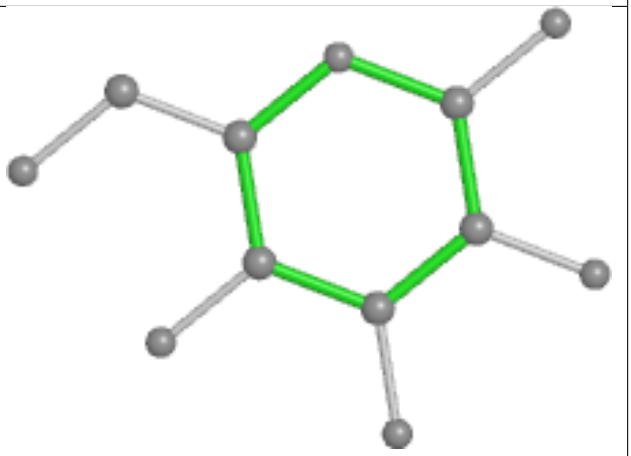
Bond lengths



Bond angles

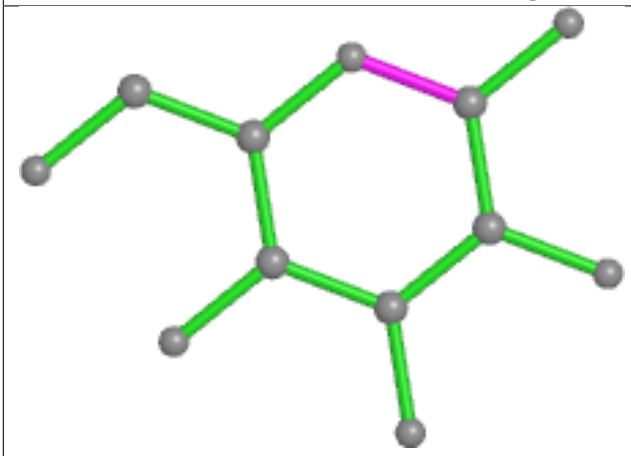


Torsions

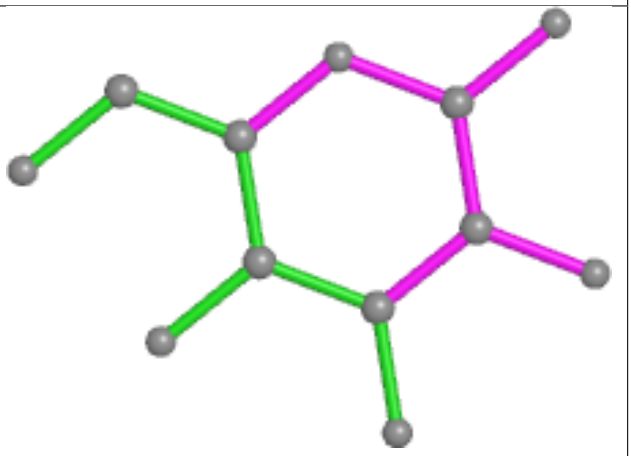


Rings

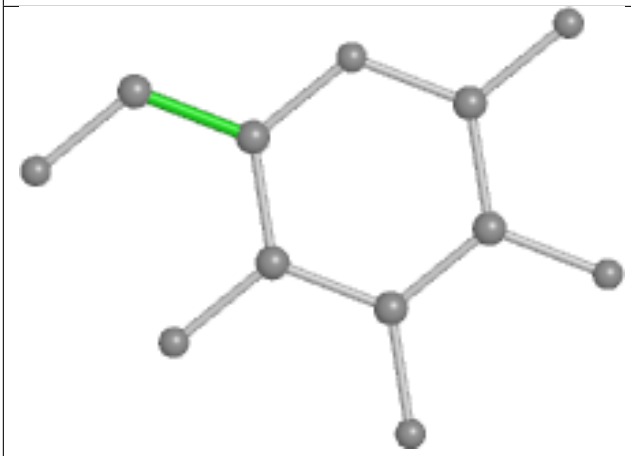
Ligand GLC B 802



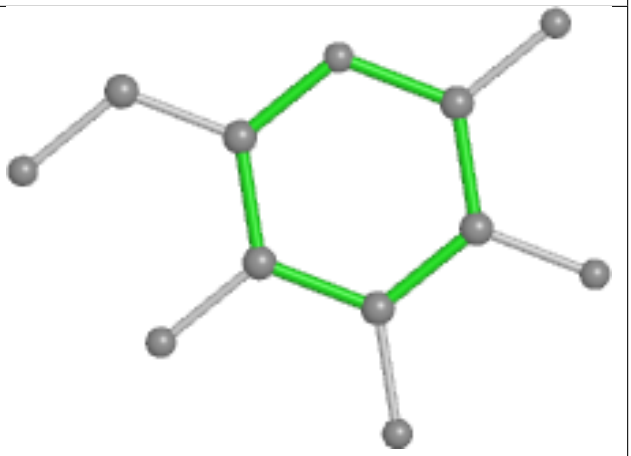
Bond lengths



Bond angles

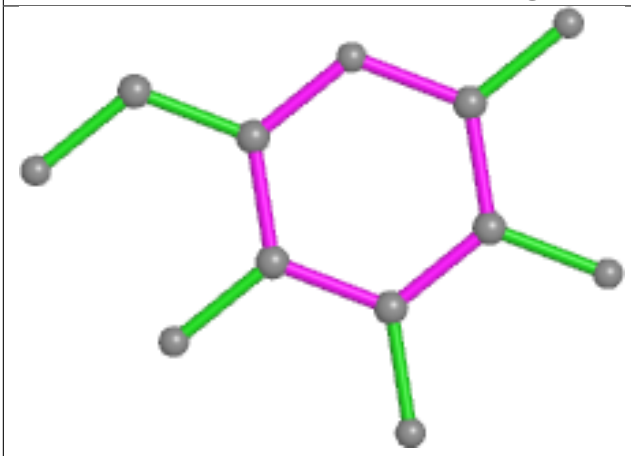


Torsions

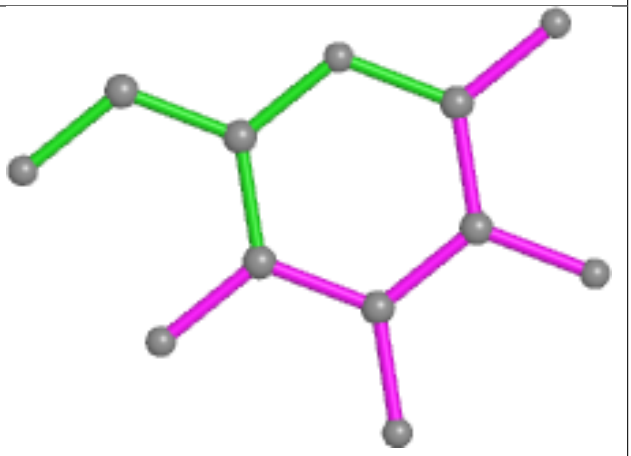


Rings

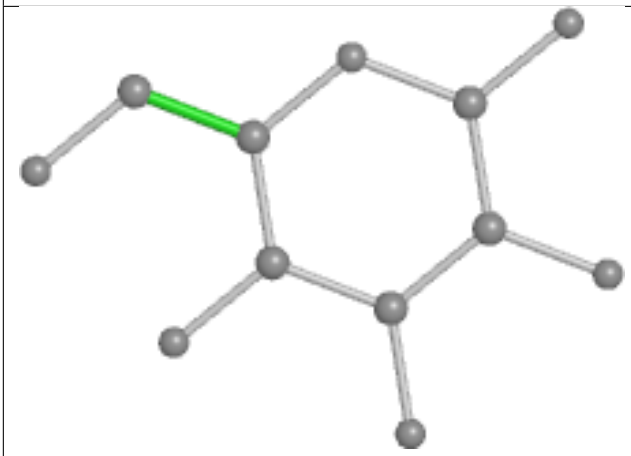
Ligand GLC C 804



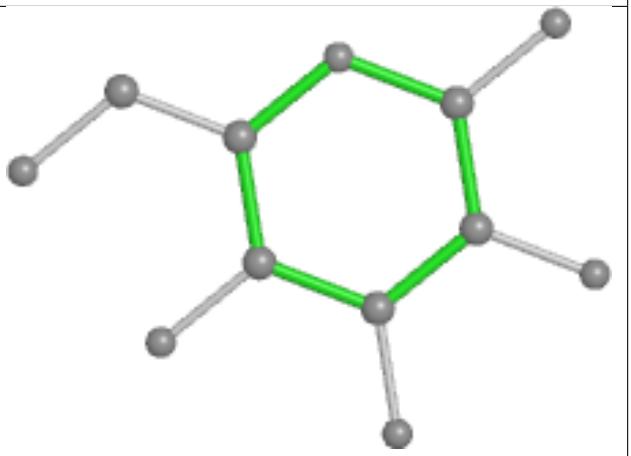
Bond lengths



Bond angles

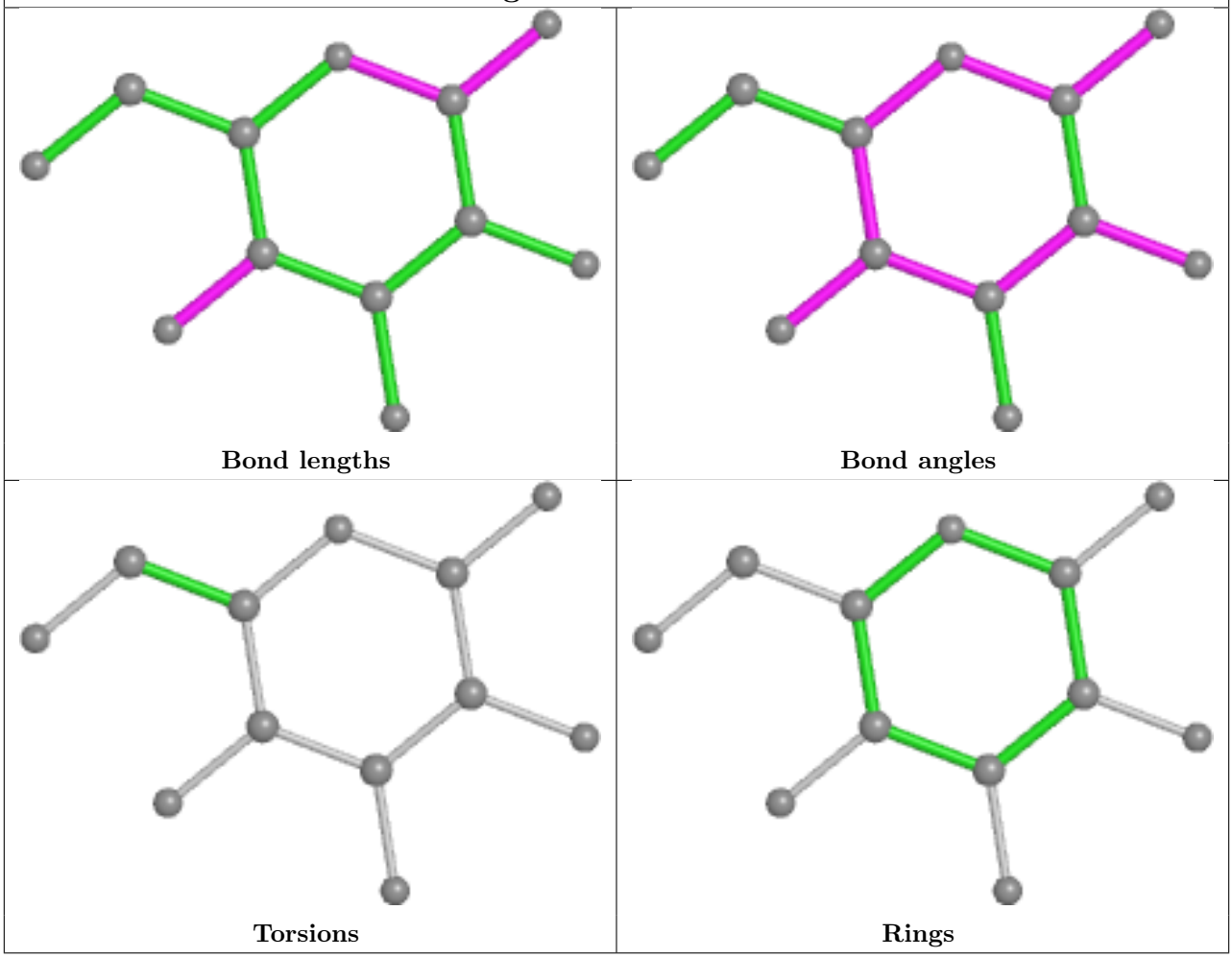


Torsions



Rings

Ligand GLC D 802



5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	682/710 (96%)	0.19	24 (3%) 44 53	22, 35, 61, 81	0
1	B	678/710 (95%)	0.52	43 (6%) 20 28	27, 47, 72, 98	0
1	C	682/710 (96%)	0.07	5 (0%) 87 92	24, 34, 58, 86	0
1	D	684/710 (96%)	0.30	32 (4%) 31 41	26, 47, 72, 106	0
All	All	2726/2840 (95%)	0.27	104 (3%) 40 50	22, 40, 69, 106	0

All (104) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	193	GLY	7.9
1	D	192	LYS	6.0
1	B	700	VAL	4.8
1	D	646	LYS	4.8
1	B	188	LYS	4.6
1	D	659	LYS	4.1
1	A	701	ALA	4.0
1	A	33	PHE	4.0
1	D	647	GLY	4.0
1	C	66	ASN	3.9
1	B	189	ALA	3.9
1	B	37	ALA	3.8
1	C	701	ALA	3.8
1	D	37	ALA	3.7
1	D	155	ALA	3.6
1	A	195	ALA	3.5
1	D	190	TYR	3.5
1	B	697	PHE	3.4
1	B	405	HIS	3.4
1	B	354	ILE	3.2
1	B	191	THR	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	188	LYS	3.2
1	A	32	HIS	3.2
1	A	77	ASP	3.2
1	D	661	ALA	3.2
1	A	106	VAL	3.2
1	B	648	LYS	3.1
1	B	112	GLU	3.1
1	B	699	PRO	3.1
1	A	108	LEU	3.1
1	B	647	GLY	3.1
1	D	680	ASN	3.1
1	B	439	GLY	2.9
1	B	620	GLY	2.9
1	B	649	LYS	2.9
1	B	676	VAL	2.9
1	B	650	TYR	2.8
1	B	408	LYS	2.8
1	A	46	ILE	2.8
1	B	675	VAL	2.8
1	D	111	LYS	2.8
1	A	406	ALA	2.8
1	D	676	VAL	2.7
1	B	487	VAL	2.7
1	B	196	SER	2.7
1	A	79	LYS	2.7
1	A	52	VAL	2.7
1	D	658	ALA	2.6
1	B	403	TYR	2.6
1	B	698	TYR	2.6
1	B	111	LYS	2.6
1	A	192	LYS	2.6
1	B	677	VAL	2.6
1	C	141	ILE	2.6
1	B	397	ILE	2.6
1	D	194	ASN	2.5
1	B	362	PHE	2.5
1	B	334	PHE	2.5
1	B	66	ASN	2.5
1	B	651	THR	2.5
1	B	195	ALA	2.4
1	D	195	ALA	2.4
1	D	634	ALA	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	138	ASN	2.4
1	B	672	ILE	2.4
1	A	74	LYS	2.4
1	D	108	LEU	2.4
1	D	678	THR	2.4
1	B	400	LEU	2.4
1	B	324	PHE	2.4
1	B	391	PRO	2.4
1	A	119	ILE	2.4
1	A	30	LEU	2.3
1	D	683	LEU	2.3
1	B	190	TYR	2.3
1	D	672	ILE	2.3
1	D	648	LYS	2.2
1	B	674	LYS	2.2
1	A	78	THR	2.2
1	D	214	GLN	2.2
1	B	433	LYS	2.2
1	D	697	PHE	2.2
1	A	42	VAL	2.2
1	A	76	GLU	2.2
1	A	20	SER	2.2
1	A	193	GLY	2.2
1	D	35	LEU	2.1
1	D	79	LYS	2.1
1	B	398	LYS	2.1
1	C	111	LYS	2.1
1	D	633	THR	2.1
1	A	111	LYS	2.1
1	B	294	GLN	2.1
1	B	80	THR	2.1
1	D	106	VAL	2.1
1	A	75	ILE	2.1
1	B	486	ALA	2.1
1	D	662	ASN	2.1
1	D	666	ASN	2.1
1	A	531	LEU	2.1
1	D	241	PHE	2.0
1	A	35	LEU	2.0
1	D	112	GLU	2.0
1	B	67	LYS	2.0

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)

There are no monosaccharides in this entry.

6.4 Ligands (i)

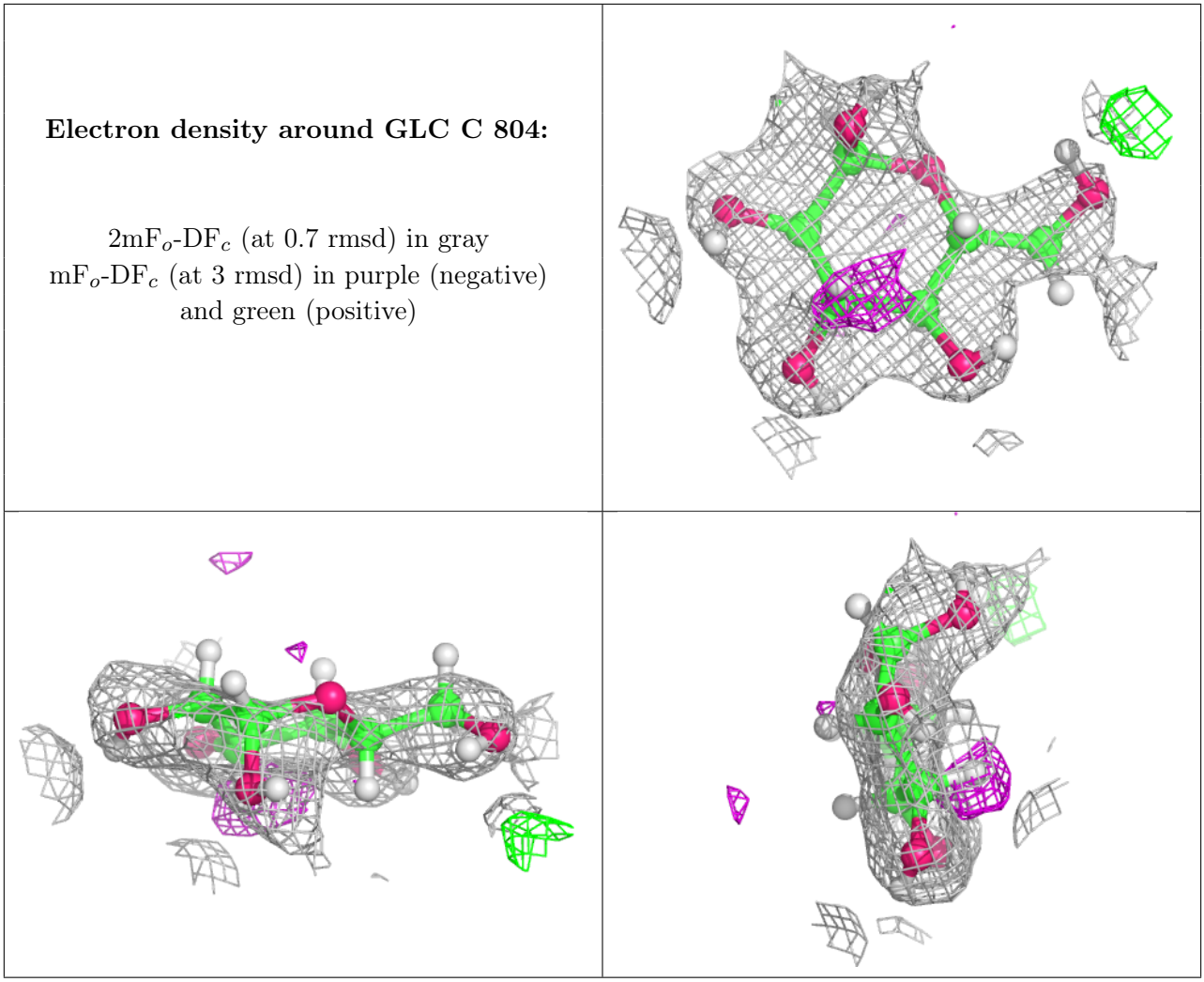
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	GLC	C	804	12/12	0.86	0.28	30,60,75,91	4
4	EDO	D	803	4/4	0.92	0.13	30,38,42,43	1
4	EDO	A	803	4/4	0.95	0.23	29,35,39,40	1
3	GLC	B	802	12/12	0.96	0.16	30,39,41,43	3
3	GLC	D	802	12/12	0.96	0.09	30,39,47,48	3
2	CA	D	801	1/1	0.97	0.11	41,41,41,41	0
3	GLC	C	802	12/12	0.97	0.20	27,30,35,41	3
4	EDO	C	803	4/4	0.98	0.13	28,30,32,34	1
3	GLC	A	802	12/12	0.98	0.14	25,30,33,37	3
2	CA	B	801	1/1	0.99	0.15	39,39,39,39	0
4	EDO	A	804	4/4	0.99	0.15	30,39,44,45	1
2	CA	A	801	1/1	1.00	0.11	26,26,26,26	0
2	CA	C	801	1/1	1.00	0.15	30,30,30,30	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.

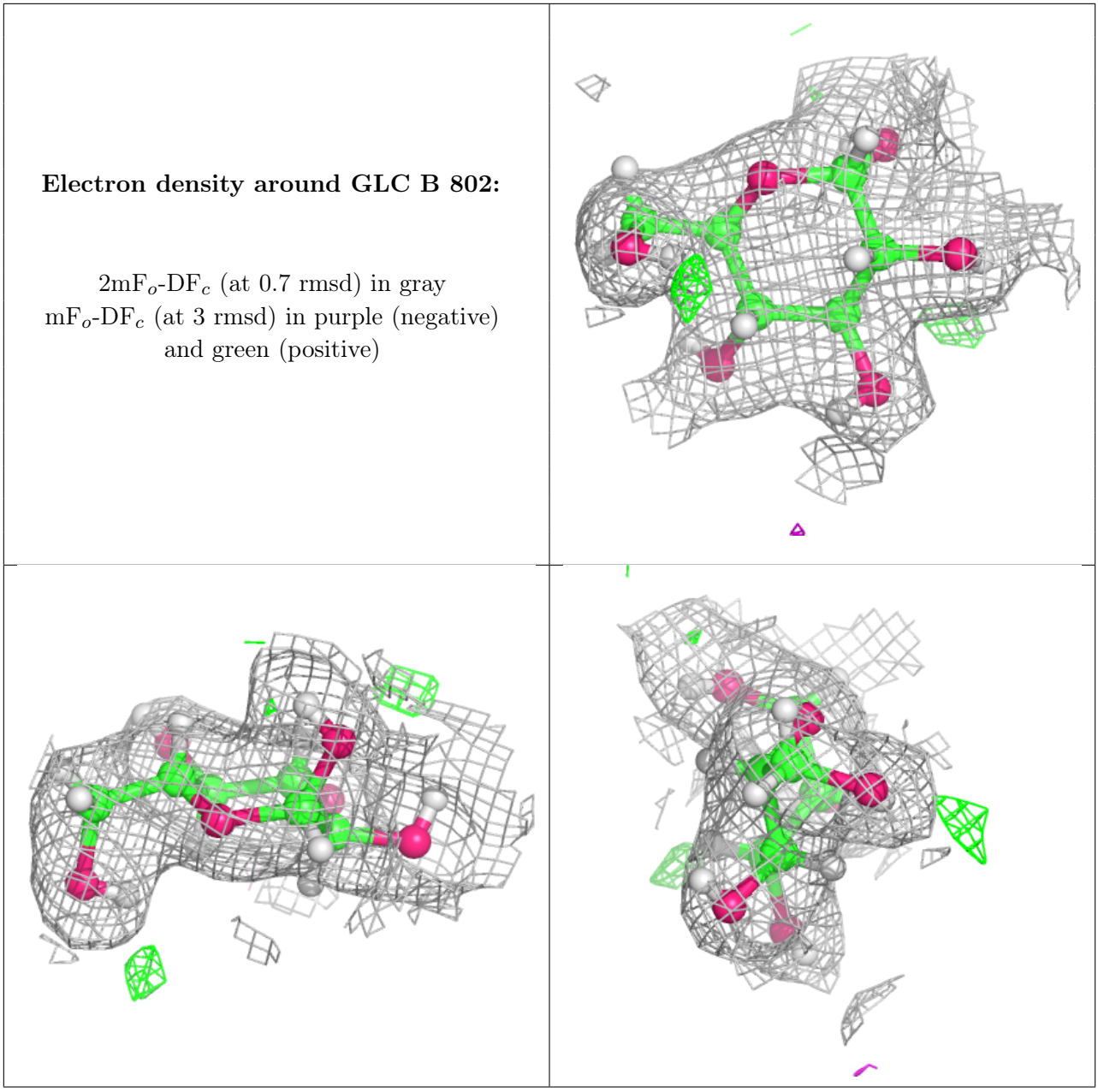
Electron density around GLC C 804:

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



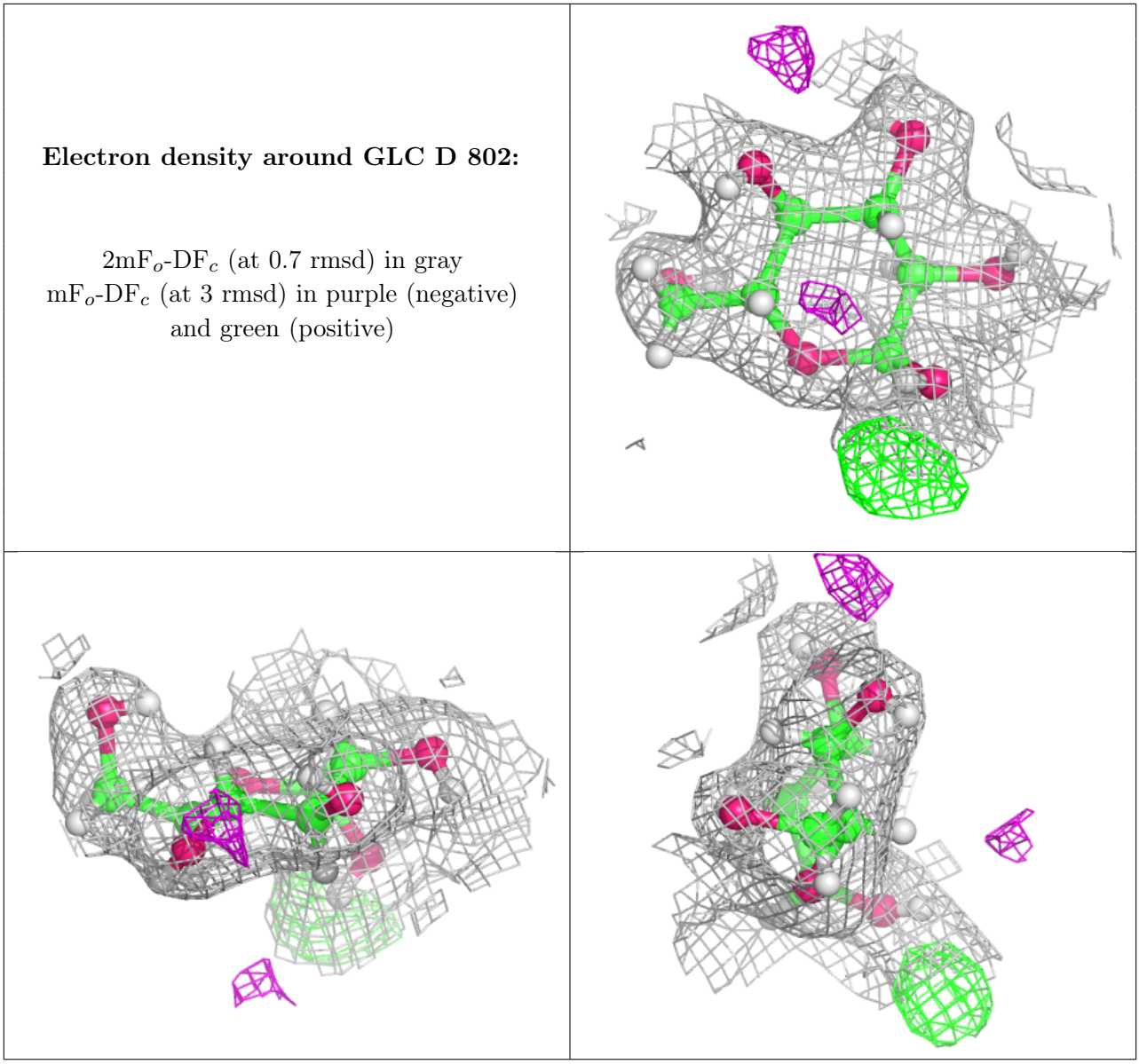
Electron density around GLC B 802:

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



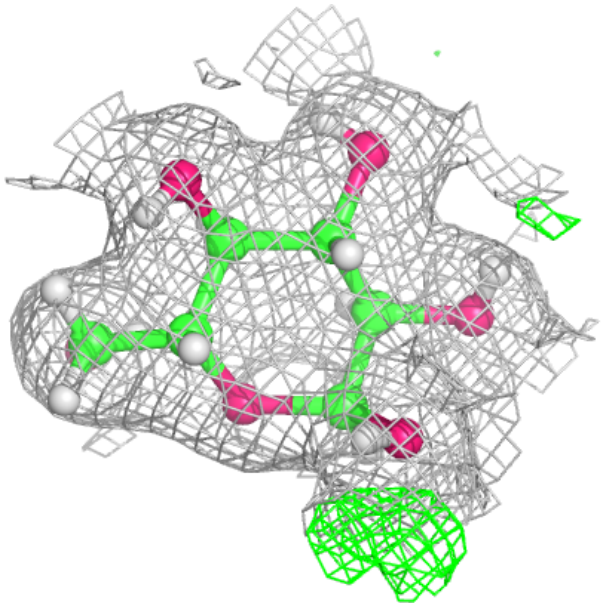
Electron density around GLC D 802:

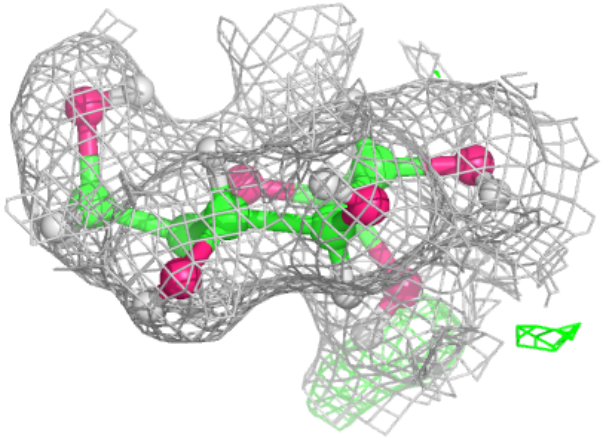
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

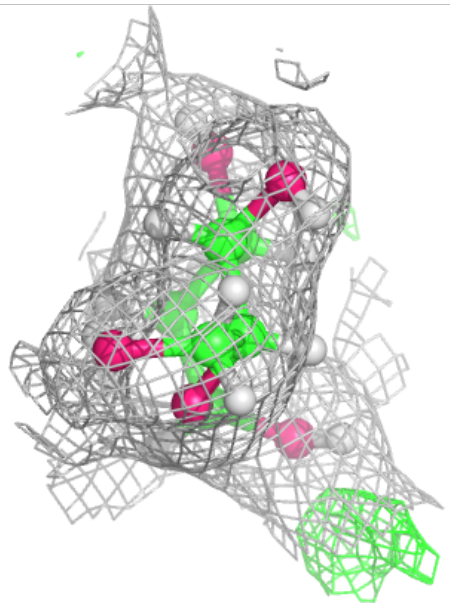


Electron density around GLC C 802:

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

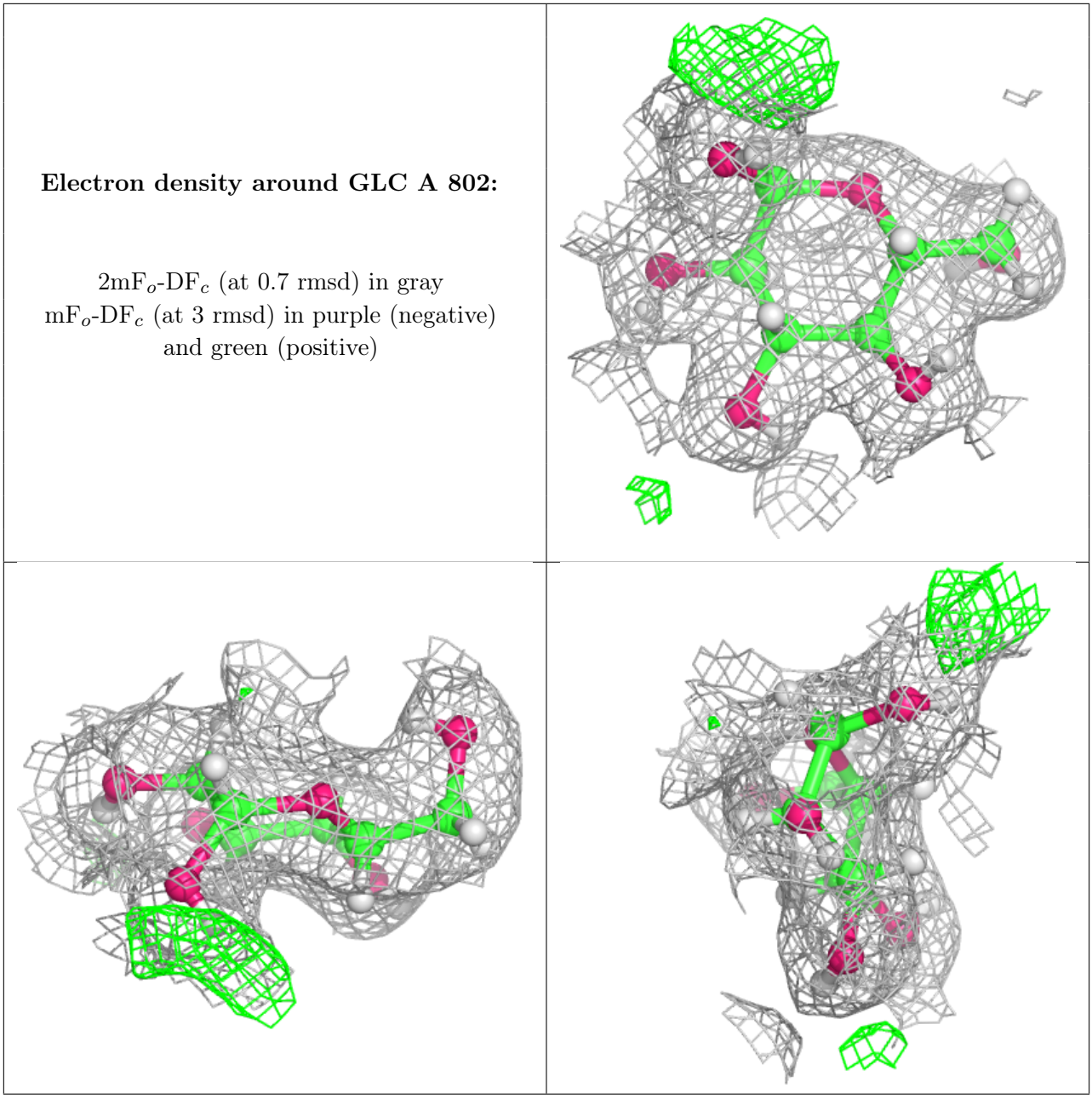






Electron density around GLC A 802:

$2mF_o$-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.