W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 14, 2020 – 03:23 AM BST

PDB ID	:	6KYW
Title	:	S8-mSRK-S8-SP11 complex
Authors	:	Murase, K.; Hakoshima, T.; Mori, T.
Deposited on	:	2019-09-20
Resolution	:	2.60 Å(reported)

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	:	2.14.4.dev1
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)	:	2.14.4.dev1

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.60 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.
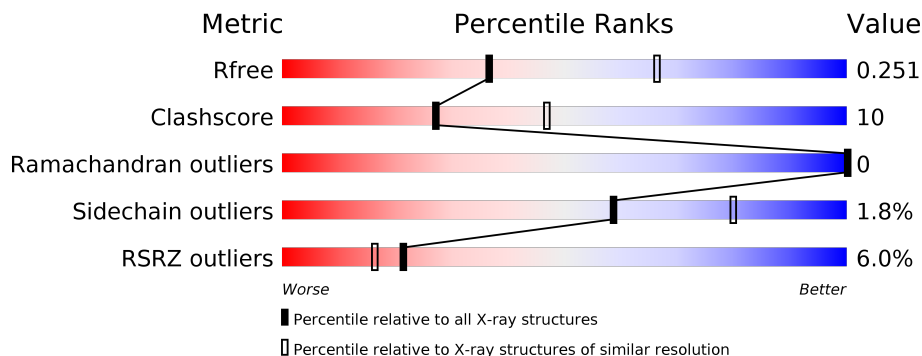
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	443	5% / 67% / 20% / 12%
1	B	443	5% / 70% / 16% / · / 13%
2	C	46	76% / 24%
2	D	46	13% / 74% / 26%

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

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	A	503	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7006 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 Receptor protein kinase SRK8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	389	Total 3034	C 1898	N 537	O 578	S 21	0	0	0
1	B	386	Total 3024	C 1893	N 532	O 578	S 21	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	79	SER	PRO	engineered mutation	UNP Q39276
A	80	GLU	TYR	engineered mutation	UNP Q39276
A	81	ARG	ILE	engineered mutation	UNP Q39276
A	108	VAL	PHE	engineered mutation	UNP Q39276
A	110	ARG	LEU	engineered mutation	UNP Q39276
A	180	ARG	LEU	engineered mutation	UNP Q39276
A	190	SER	PHE	engineered mutation	UNP Q39276
A	214	GLN	LEU	engineered mutation	UNP Q39276
A	239	SER	LEU	engineered mutation	UNP Q39276
A	248	GLU	LYS	engineered mutation	UNP Q39276
A	286	GLY	VAL	engineered mutation	UNP Q39276
A	287	ALA	VAL	engineered mutation	UNP Q39276
A	434	SER	-	expression tag	UNP Q39276
A	435	ARG	-	expression tag	UNP Q39276
A	436	LEU	-	expression tag	UNP Q39276
A	437	GLU	-	expression tag	UNP Q39276
A	438	VAL	-	expression tag	UNP Q39276
A	439	LEU	-	expression tag	UNP Q39276
A	440	PHE	-	expression tag	UNP Q39276
A	441	GLN	-	expression tag	UNP Q39276
A	442	GLY	-	expression tag	UNP Q39276
A	443	PRO	-	expression tag	UNP Q39276
B	79	SER	PRO	engineered mutation	UNP Q39276
B	80	GLU	TYR	engineered mutation	UNP Q39276
B	81	ARG	ILE	engineered mutation	UNP Q39276

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	108	VAL	PHE	engineered mutation	UNP Q39276
B	110	ARG	LEU	engineered mutation	UNP Q39276
B	180	ARG	LEU	engineered mutation	UNP Q39276
B	190	SER	PHE	engineered mutation	UNP Q39276
B	214	GLN	LEU	engineered mutation	UNP Q39276
B	239	SER	LEU	engineered mutation	UNP Q39276
B	248	GLU	LYS	engineered mutation	UNP Q39276
B	286	GLY	VAL	engineered mutation	UNP Q39276
B	287	ALA	VAL	engineered mutation	UNP Q39276
B	434	SER	-	expression tag	UNP Q39276
B	435	ARG	-	expression tag	UNP Q39276
B	436	LEU	-	expression tag	UNP Q39276
B	437	GLU	-	expression tag	UNP Q39276
B	438	VAL	-	expression tag	UNP Q39276
B	439	LEU	-	expression tag	UNP Q39276
B	440	PHE	-	expression tag	UNP Q39276
B	441	GLN	-	expression tag	UNP Q39276
B	442	GLY	-	expression tag	UNP Q39276
B	443	PRO	-	expression tag	UNP Q39276

- Molecule 2 is a protein called S locus protein 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	46	Total 356	C 210	N 68	O 69	S 9	0	0	0
2	D	46	Total 338	C 201	N 59	O 69	S 9	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by author).

NAG

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	69	Total	O	0	0
			69	69		
4	B	90	Total	O	0	0
			90	90		
4	C	13	Total	O	0	0
			13	13		
4	D	12	Total	O	0	0
			12	12		

3 Residue-property plots ⓘ

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

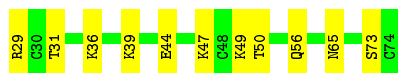- Molecule 1: Receptor protein kinase SRK8
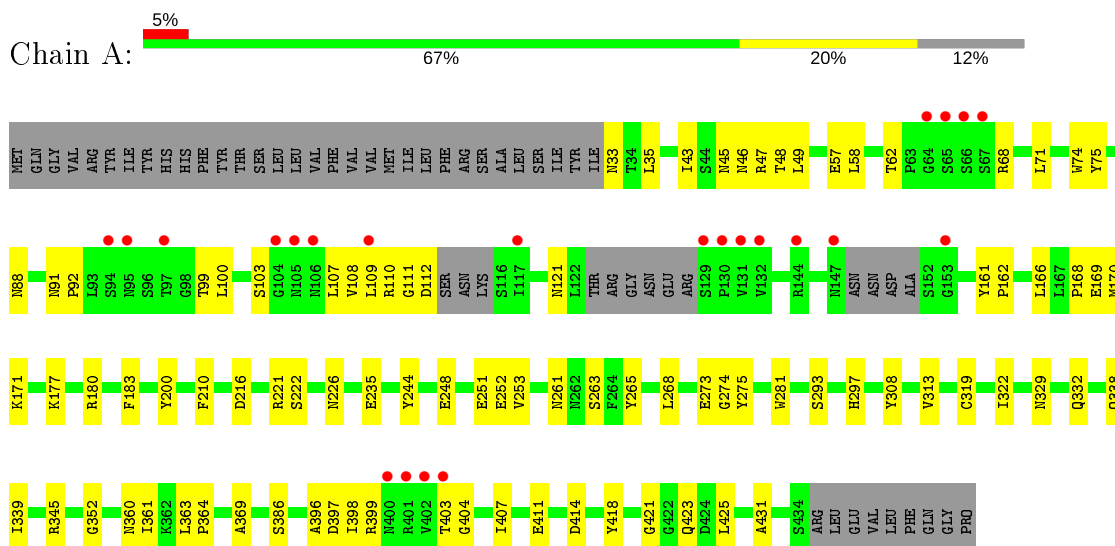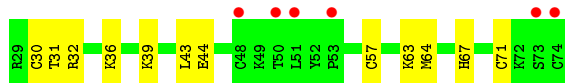

- Molecule 1: Receptor protein kinase SRK8

- Molecule 2: S locus protein 11

R29 C30 T31 K36 K39 E44 K47 C48 K49 T50 Q56 N65 S73 C74

• Molecule 2: S locus protein 11

Chain D:

13%

74%   26%

R29 C30 T31 R32 K36 K39 L43 E44 C48 K49 T50 L51 Y52 P53 C57 K63 M64 H67 C71 K72 S73 C74
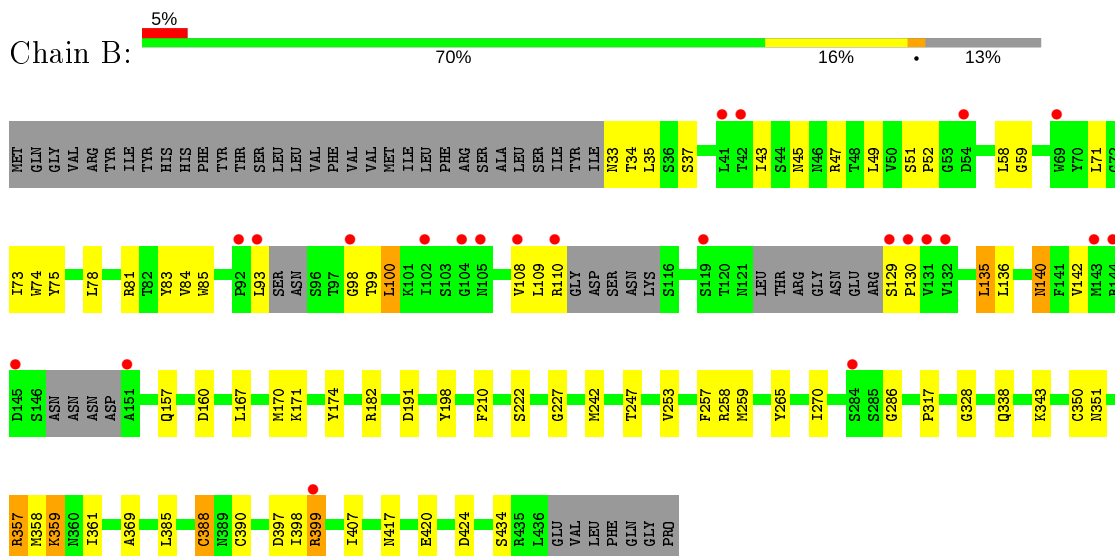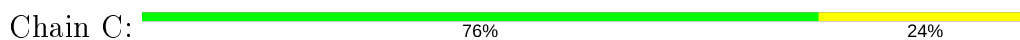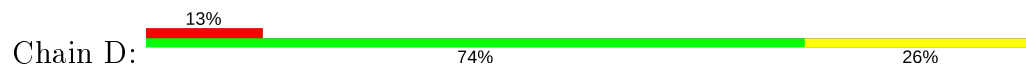
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	143.56Å 143.56Å 194.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.84 – 2.60 43.84 – 2.60	Depositor EDS
% Data completeness (in resolution range)	94.9 (43.84-2.60) 94.9 (43.84-2.60)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.22 (at 2.61Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.220 , 0.251 0.220 , 0.251	Depositor DCC
R_{free} test set	1999 reflections (3.35%)	wwPDB-VP
Wilson B-factor (Å^2)	50.8	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 55.6	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7006	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.66% 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	0.49	0/3105	0.65	0/4215
1	B	0.50	1/3094 (0.0%)	0.66	0/4198
2	C	0.46	0/359	0.71	0/475
2	D	0.43	0/341	0.60	0/454
All	All	0.49	1/6899 (0.0%)	0.65	0/9342

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	351	ASN	C-N	-6.17	1.22	1.33

There are no bond angle outliers.

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	3034	0	2801	58	0
1	B	3024	0	2811	60	0
2	C	356	0	346	9	0
2	D	338	0	313	10	0
3	A	42	0	39	1	0
3	B	28	0	26	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	69	0	0	4	0
4	B	90	0	0	2	0
4	C	13	0	0	0	0
4	D	12	0	0	2	0
All	All	7006	0	6336	131	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 10.

The worst 5 of 131 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:45:ASN:HA	1:A:46:ASN:HB2	1.42	1.00
1:B:242:MET:HG2	1:B:259:MET:HE2	1.57	0.87
1:B:242:MET:HG2	1:B:259:MET:CE	2.10	0.81
1:B:167:LEU:H	1:B:170:MET:HE3	1.47	0.79
1:A:221:ARG:NH2	1:A:363:LEU:O	2.18	0.77

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	381/443 (86%)	360 (94%)	21 (6%)	0	100	100
1	B	376/443 (85%)	350 (93%)	26 (7%)	0	100	100
2	C	44/46 (96%)	44 (100%)	0	0	100	100
2	D	44/46 (96%)	40 (91%)	4 (9%)	0	100	100
All	All	845/978 (86%)	794 (94%)	51 (6%)	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	319/394 (81%)	317 (99%)	2 (1%)	86	95
1	B	321/394 (82%)	310 (97%)	11 (3%)	37	63
2	C	43/43 (100%)	43 (100%)	0	100	100
2	D	40/43 (93%)	40 (100%)	0	100	100
All	All	723/874 (83%)	710 (98%)	13 (2%)	59	80

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	140	ASN
1	B	191	ASP
1	B	399	ARG
1	B	135	LEU
1	B	388	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	332	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)

5 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	NAG	A	503	1	14,14,15	0.40	0	17,19,21	0.58	0
3	NAG	B	502	1	14,14,15	0.26	0	17,19,21	0.61	1 (5%)
3	NAG	B	501	1	14,14,15	0.41	0	17,19,21	0.97	1 (5%)
3	NAG	A	501	1	14,14,15	0.56	0	17,19,21	0.61	0
3	NAG	A	502	1	14,14,15	0.40	0	17,19,21	0.53	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	NAG	A	503	1	-	2/6/23/26	0/1/1/1
3	NAG	B	502	1	-	2/6/23/26	0/1/1/1
3	NAG	B	501	1	-	2/6/23/26	0/1/1/1
3	NAG	A	501	1	-	0/6/23/26	0/1/1/1
3	NAG	A	502	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	501	NAG	C1-O5-C5	3.48	116.91	112.19
3	B	502	NAG	C1-O5-C5	2.05	114.97	112.19

There are no chirality outliers.

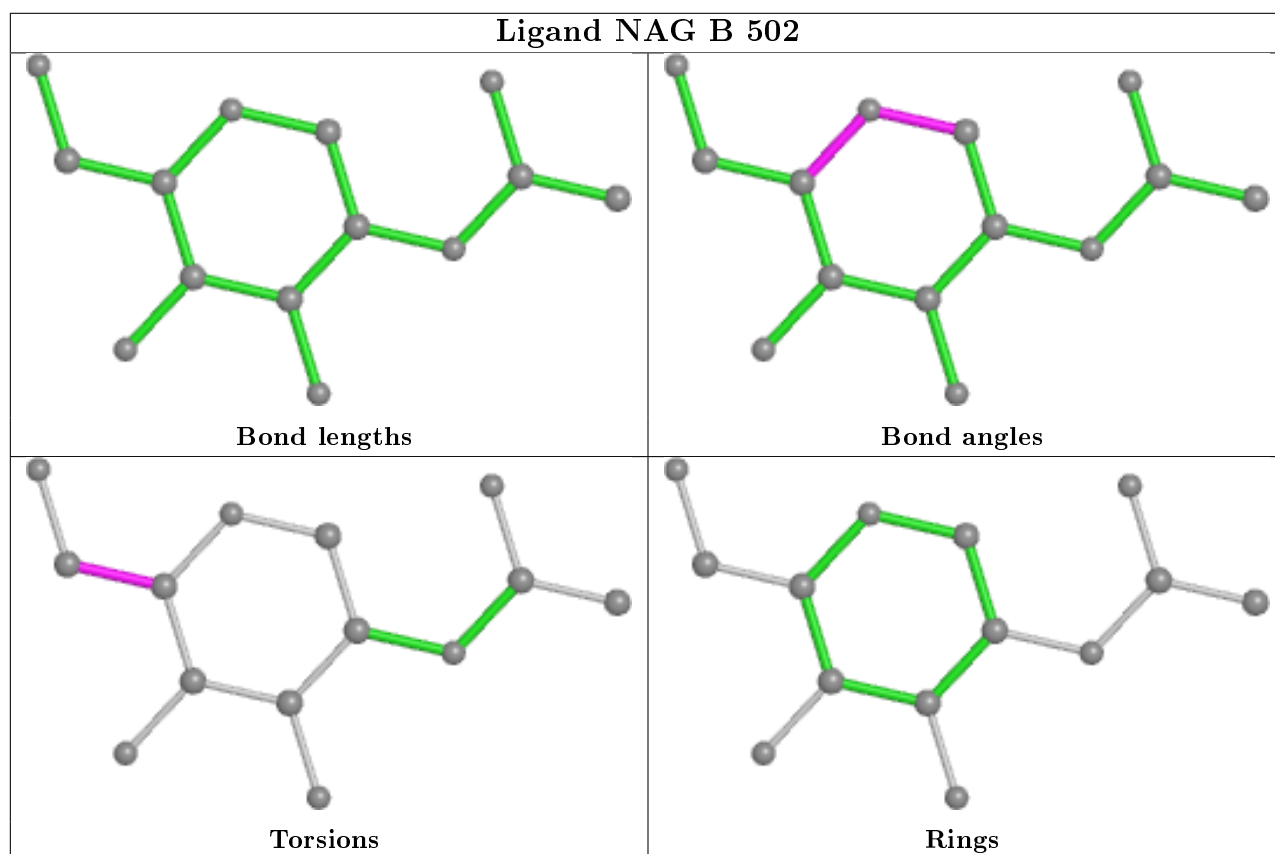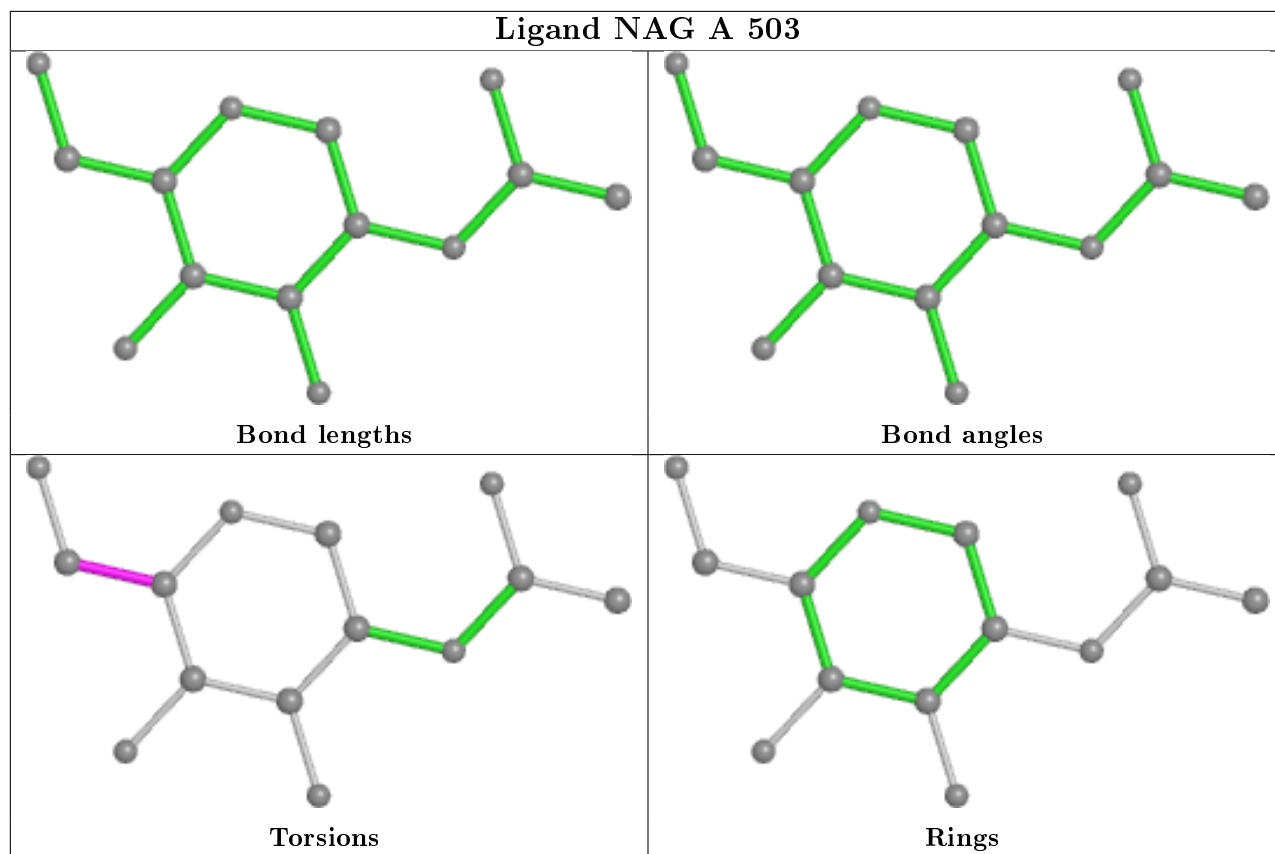5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	501	NAG	O5-C5-C6-O6
3	B	502	NAG	O5-C5-C6-O6
3	B	501	NAG	C4-C5-C6-O6
3	A	503	NAG	O5-C5-C6-O6
3	B	502	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	NAG	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 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 NAG A 503

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG B 502

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG B 501

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG A 501

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG A 502

Bond lengths

Bond angles

Torsions

Rings
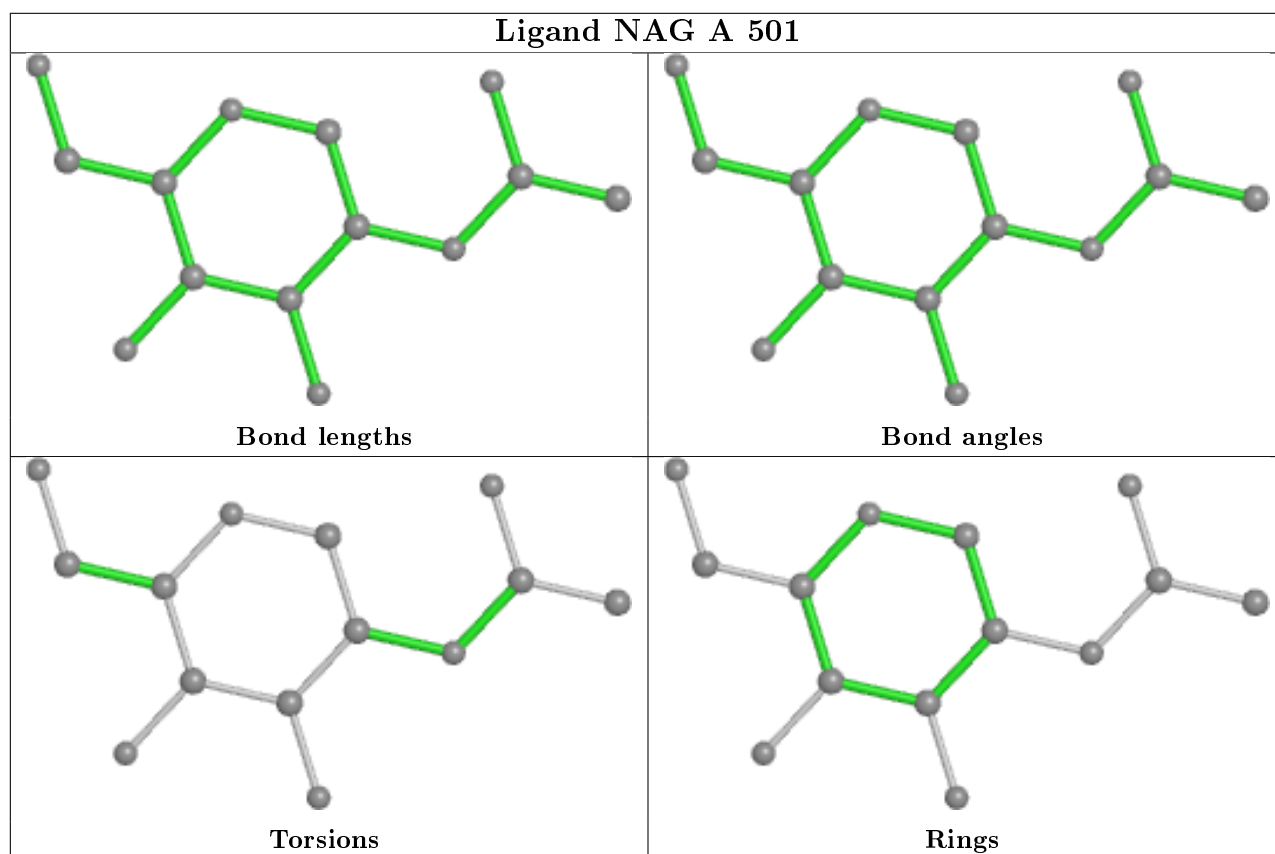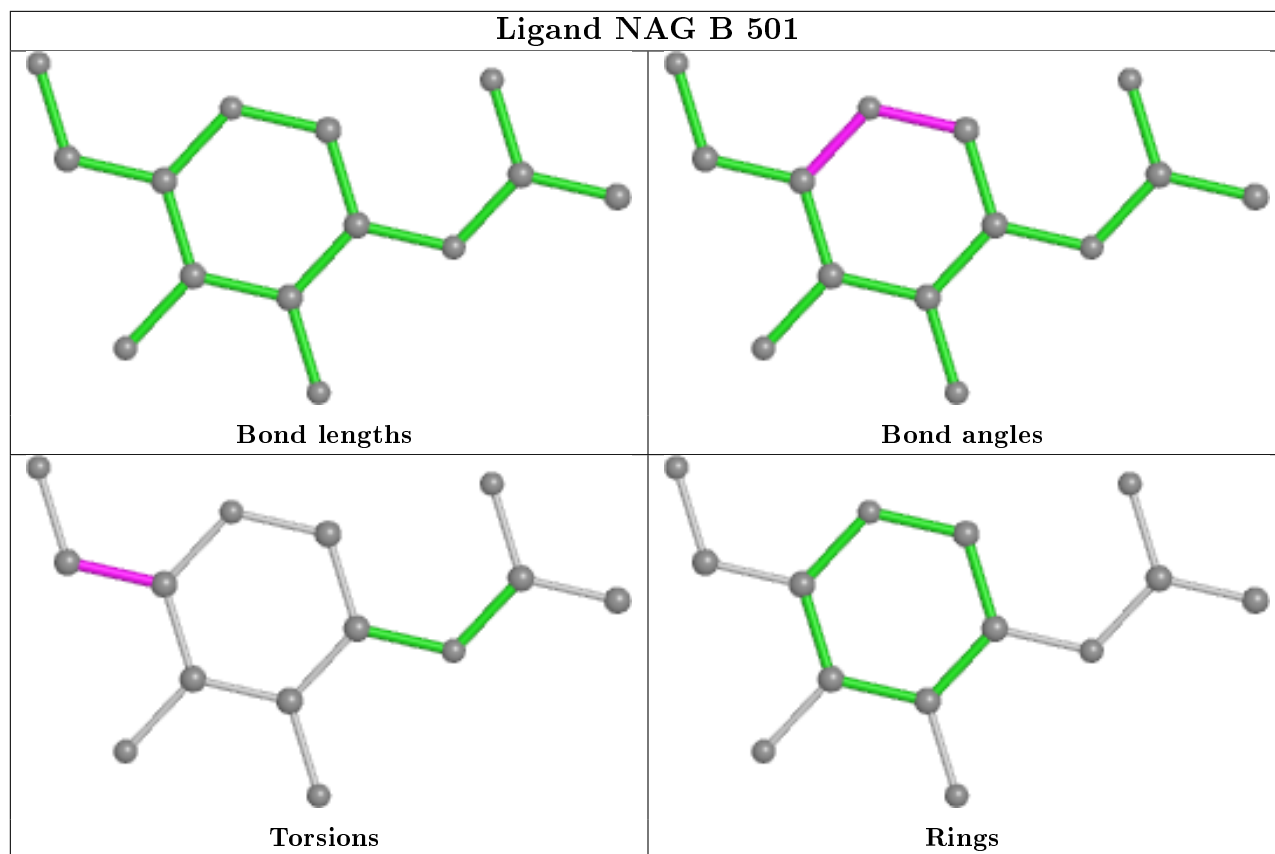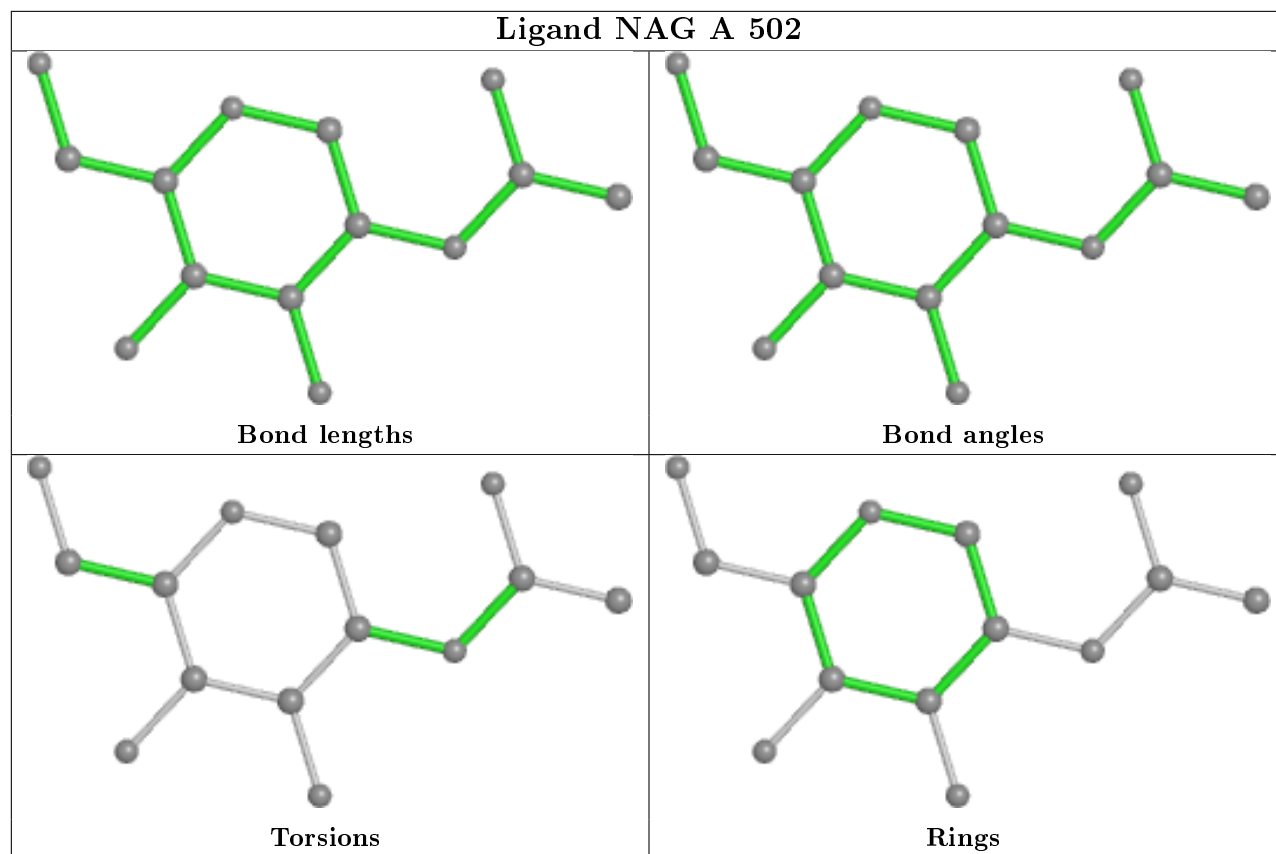
5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	389/443 (87%)	0.34	23 (5%)	22	17	33, 59, 104, 125	0
1	B	386/443 (87%)	0.23	23 (5%)	21	16	30, 53, 104, 115	0
2	C	46/46 (100%)	0.02	0	100	100	35, 57, 71, 79	0
2	D	46/46 (100%)	0.72	6 (13%)	3	2	42, 69, 98, 103	0
All	All	867/978 (88%)	0.29	52 (5%)	21	16	30, 58, 103, 125	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	51	LEU	5.2
1	B	145	ASP	4.9
1	A	401	ARG	4.8
1	B	105	ASN	4.3
1	B	132	VAL	4.3

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
3	NAG	A	503	14/15	0.57	0.43	72,85,94,112	14
3	NAG	A	502	14/15	0.66	0.26	75,82,90,91	14
3	NAG	B	501	14/15	0.72	0.24	78,85,93,94	14
3	NAG	B	502	14/15	0.76	0.40	83,91,96,103	14
3	NAG	A	501	14/15	0.78	0.31	81,87,96,97	14

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 NAG A 503:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative) and green (positive)
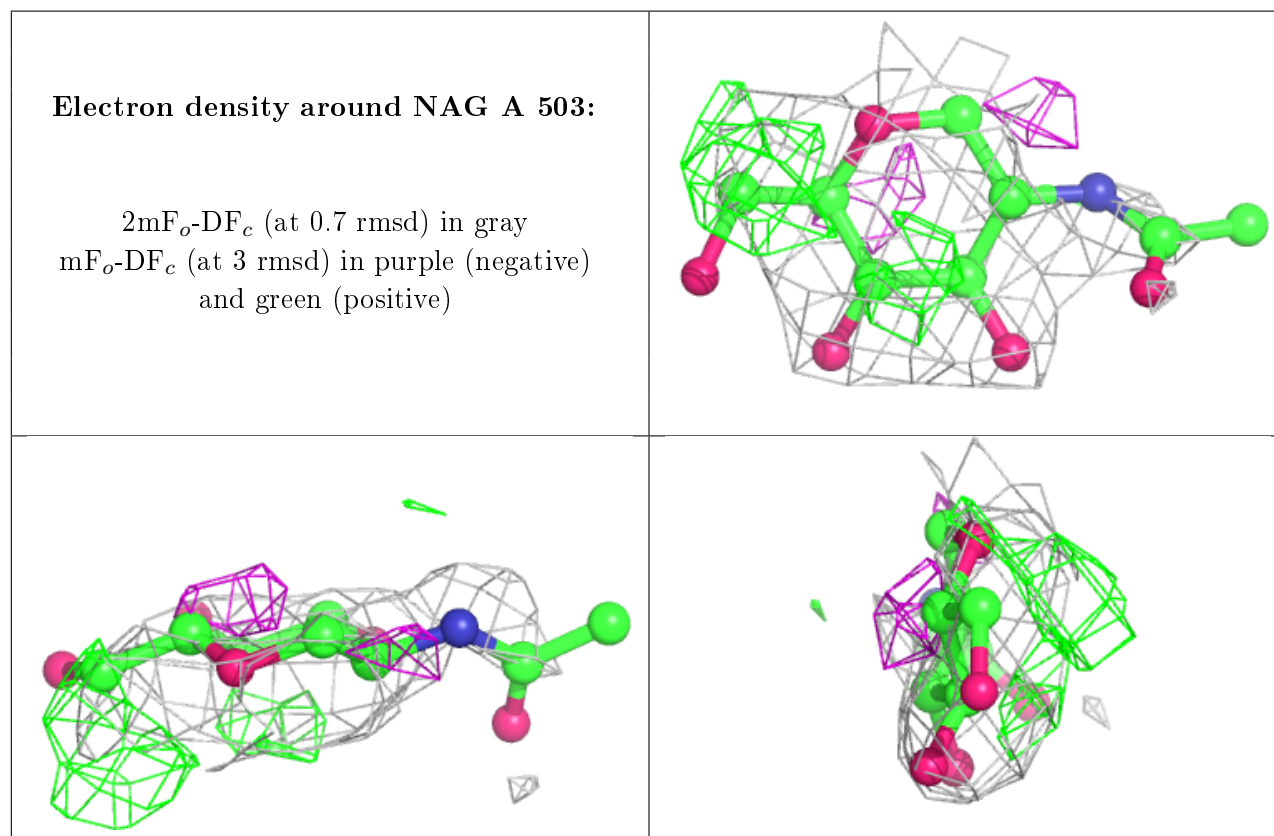
Electron density around NAG A 502:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
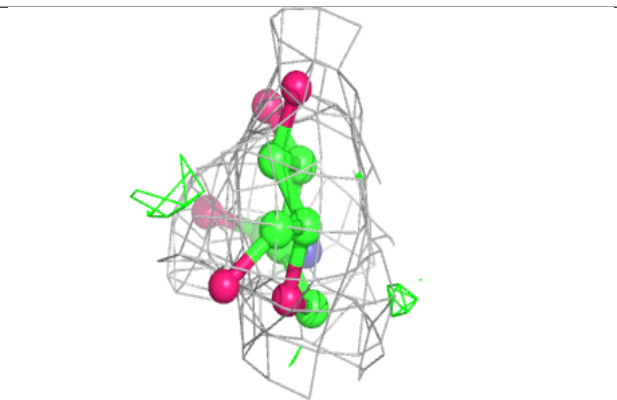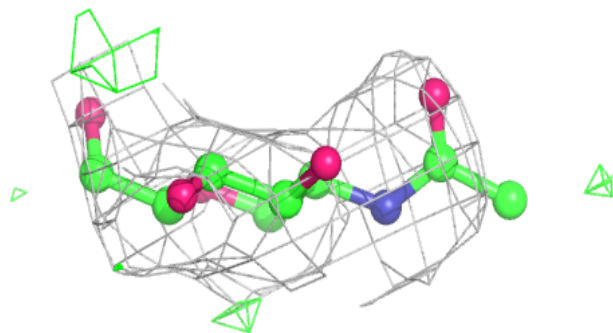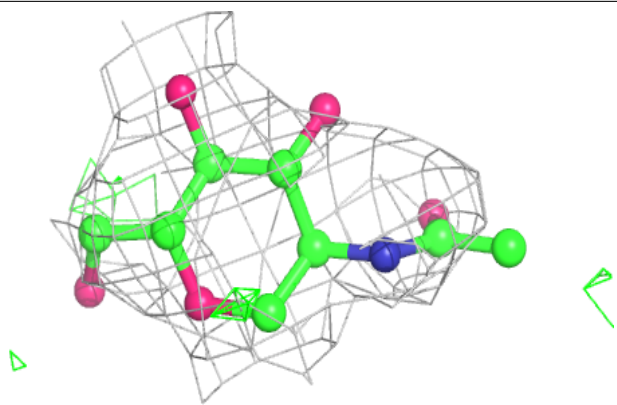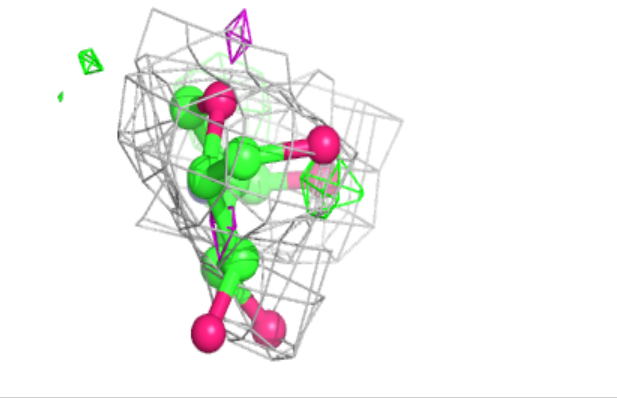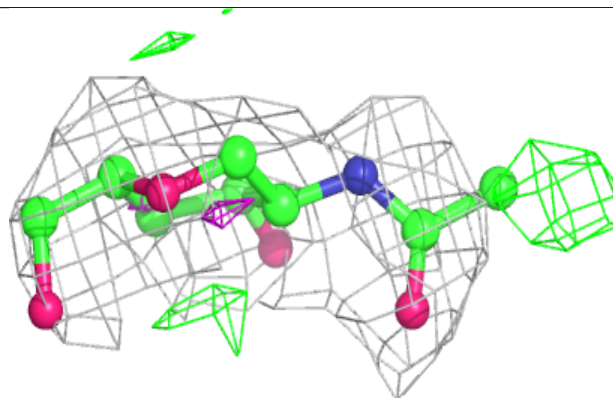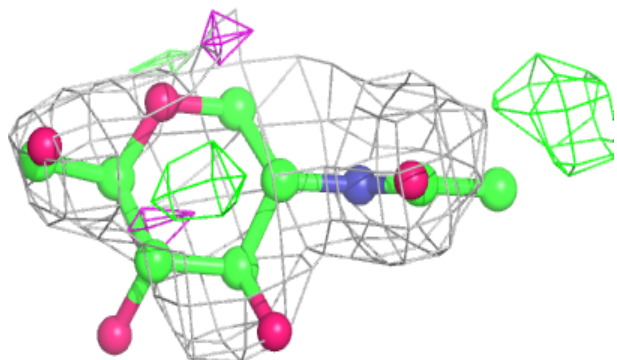mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
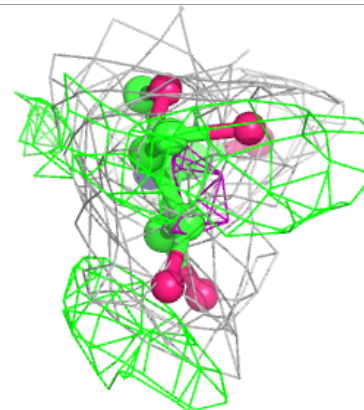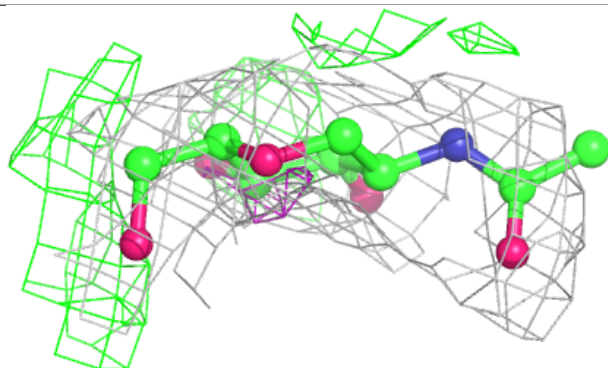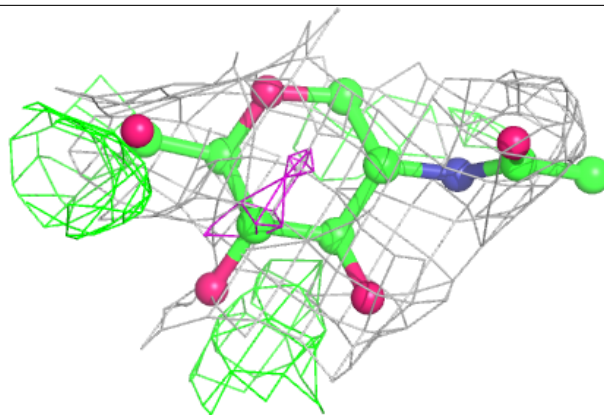
Electron density around NAG B 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
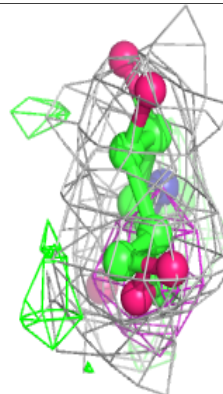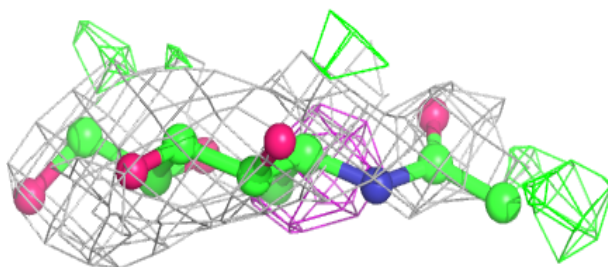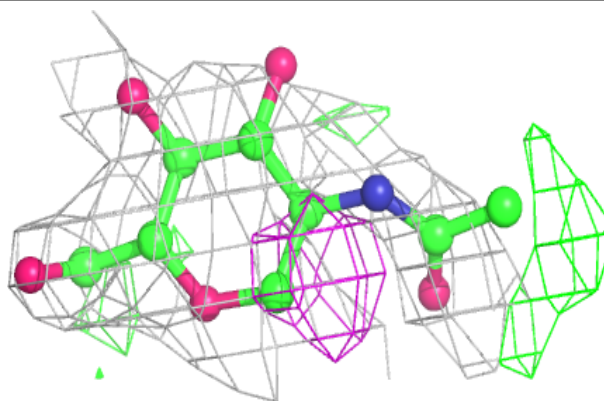
Electron density around NAG B 502:

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

Electron density around NAG A 501:

$2mF_o$-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.