wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 11:43 am BST

PDB ID	:	5OE6
Title	:	Crystal structure of the N-terminal domain of PqsA in complex with 6-fluoro anthraniloyl-AMP (crystal form 1)
Authors	:	Witzgall, F.; Ewert, W.; Blankenfeldt, W.
Deposited on	:	2017-07-07
Resolution	:	1.67 Å(reported)

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

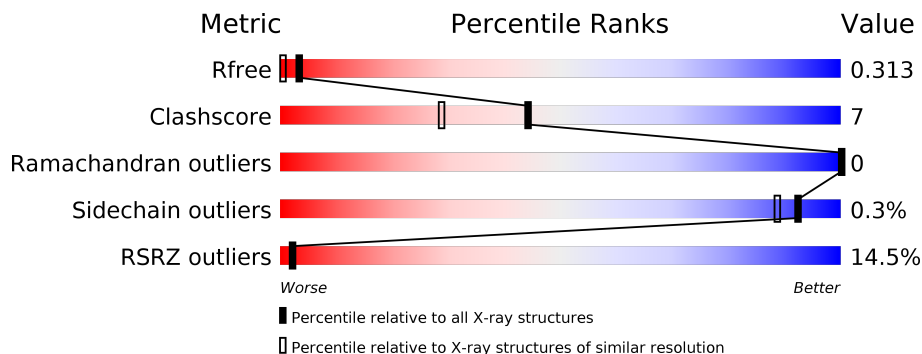|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.67 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	Percentile Ranks	Value
Rfree		0.313
Clashscore		7
Ramachandran outliers		0
Sidechain outliers		0.3%
RSRZ outliers		14.5%

Worse　　　　　　　　　　　　　　　　　　*Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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	407	23% / 84% 13% •
1	B	407	2% / 91% 5% •
1	C	407	28% / 82% 14% •
1	D	407	3% / 89% 7% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 25754 atoms, of which 12038 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 Anthranilate–CoA ligase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	394	Total 6073	C 1947	H 3017	N 539	O 557	S 13	0	8	0
1	B	392	Total 6054	C 1937	H 3012	N 543	O 549	S 13	0	6	0
1	C	393	Total 5985	C 1923	H 2964	N 535	O 550	S 13	0	5	0
1	D	391	Total 5970	C 1919	H 2959	N 532	O 548	S 12	0	3	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	400	LEU	-	expression tag	UNP Q9I4X3
A	401	GLU	-	expression tag	UNP Q9I4X3
A	402	HIS	-	expression tag	UNP Q9I4X3
A	403	HIS	-	expression tag	UNP Q9I4X3
A	404	HIS	-	expression tag	UNP Q9I4X3
A	405	HIS	-	expression tag	UNP Q9I4X3
A	406	HIS	-	expression tag	UNP Q9I4X3
A	407	HIS	-	expression tag	UNP Q9I4X3
B	400	LEU	-	expression tag	UNP Q9I4X3
B	401	GLU	-	expression tag	UNP Q9I4X3
B	402	HIS	-	expression tag	UNP Q9I4X3
B	403	HIS	-	expression tag	UNP Q9I4X3
B	404	HIS	-	expression tag	UNP Q9I4X3
B	405	HIS	-	expression tag	UNP Q9I4X3
B	406	HIS	-	expression tag	UNP Q9I4X3
B	407	HIS	-	expression tag	UNP Q9I4X3
C	400	LEU	-	expression tag	UNP Q9I4X3
C	401	GLU	-	expression tag	UNP Q9I4X3
C	402	HIS	-	expression tag	UNP Q9I4X3
C	403	HIS	-	expression tag	UNP Q9I4X3
C	404	HIS	-	expression tag	UNP Q9I4X3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	405	HIS	-	expression tag	UNP Q9I4X3
C	406	HIS	-	expression tag	UNP Q9I4X3
C	407	HIS	-	expression tag	UNP Q9I4X3
D	400	LEU	-	expression tag	UNP Q9I4X3
D	401	GLU	-	expression tag	UNP Q9I4X3
D	402	HIS	-	expression tag	UNP Q9I4X3
D	403	HIS	-	expression tag	UNP Q9I4X3
D	404	HIS	-	expression tag	UNP Q9I4X3
D	405	HIS	-	expression tag	UNP Q9I4X3
D	406	HIS	-	expression tag	UNP Q9I4X3
D	407	HIS	-	expression tag	UNP Q9I4X3

- Molecule 2 is 6-fluoroanthraniloyl-AMP (three-letter code: 9SN) (formula: $C_{17}H_{18}FN_6O_8P$).
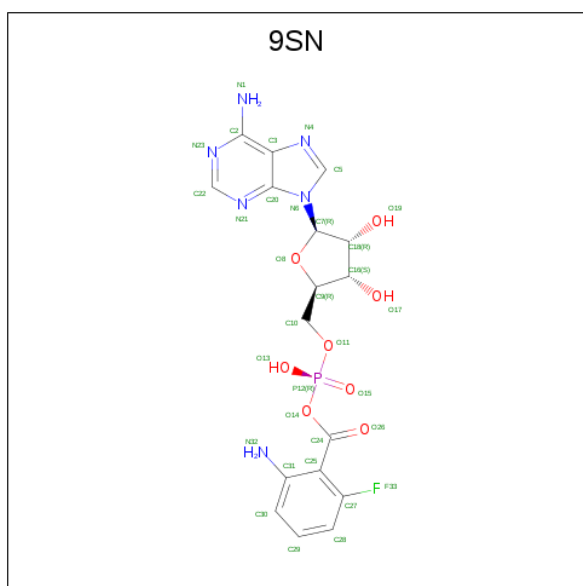
Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
2	A	1	Total	C	F	H	N	O	P	0	0
			51	17	1	18	6	8	1		
2	B	1	Total	C	F	H	N	O	P	0	0
			51	17	1	18	6	8	1		
2	C	1	Total	C	F	H	N	O	P	0	0
			51	17	1	18	6	8	1		
2	D	1	Total	C	F	H	N	O	P	0	0
			51	17	1	18	6	8	1		

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total	C	H	O	0	0
			24	6	14	4		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	360	Total	O	0	2
			360	360		
4	B	388	Total	O	0	1
			388	388		
4	C	352	Total	O	0	0
			352	352		
4	D	344	Total	O	0	0
			344	344		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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
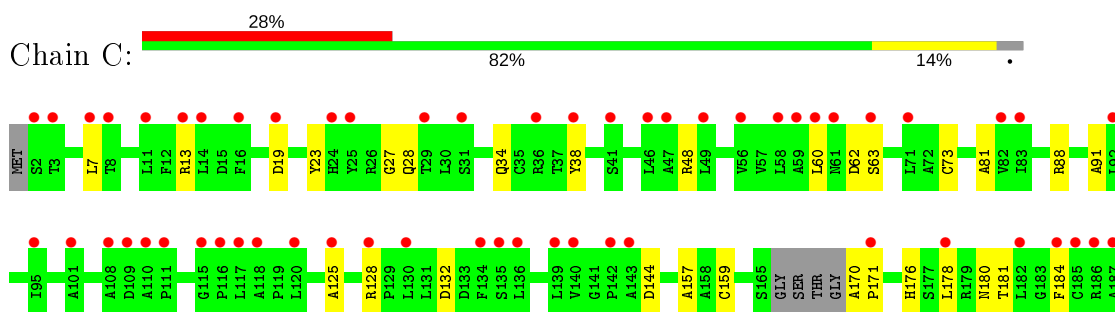sample, but not in the model, are shown in grey.
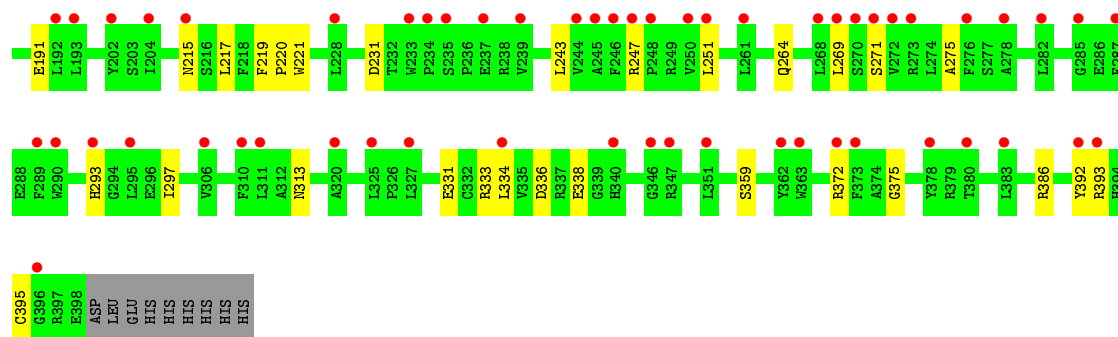
- Molecule 1: Anthranilate–CoA ligase
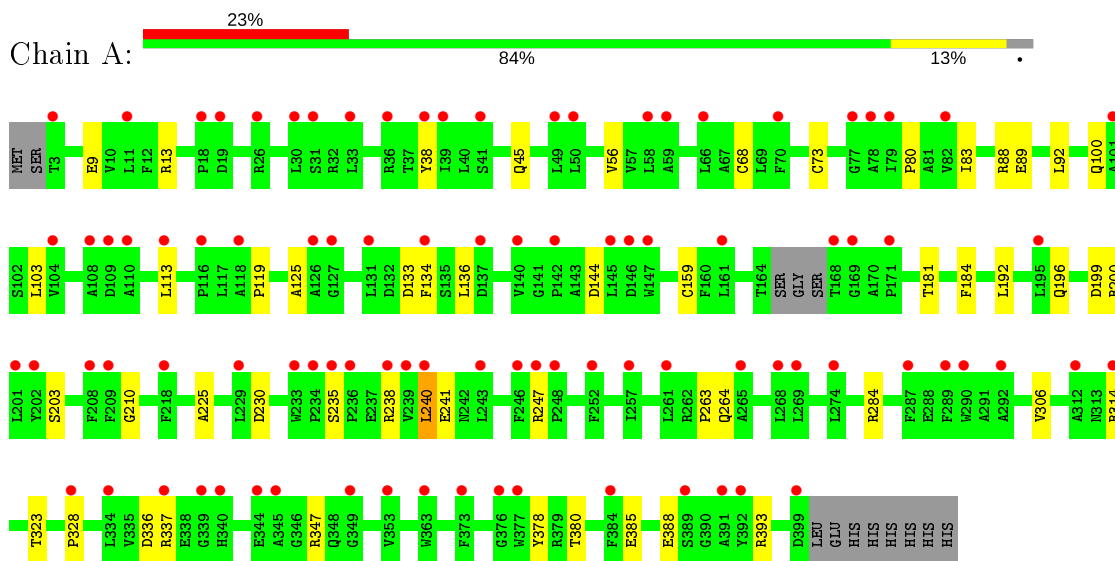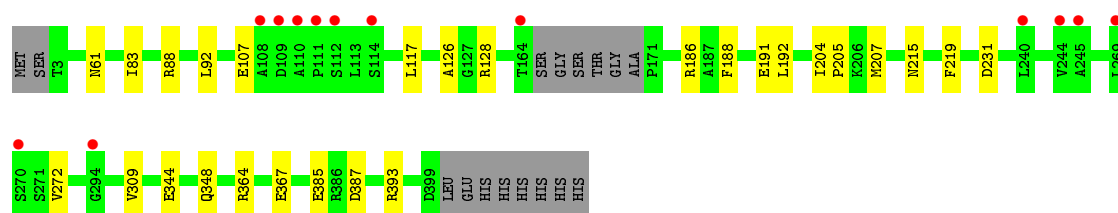
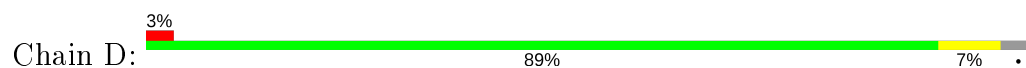Chain A:

- Molecule 1: Anthranilate–CoA ligase

Chain B:

- Molecule 1: Anthranilate–CoA ligase

Chain C:
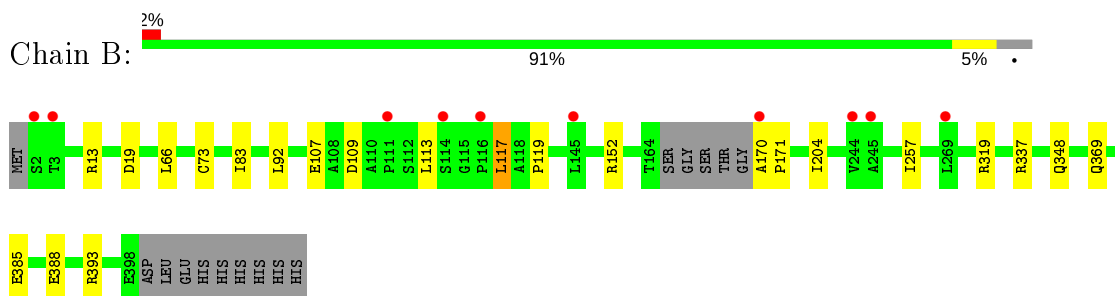
• Molecule 1: Anthranilate–CoA ligase

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	71.56Å　　84.13Å　　138.21Å 90.00°　　92.59°　　90.00°	Depositor
Resolution (Å)	46.02　–　1.67 46.02　–　1.67	Depositor EDS
% Data completeness (in resolution range)	99.6 (46.02-1.67) 87.2 (46.02-1.67)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.30 (at 1.67Å)	Xtriage
Refinement program	PHENIX 1.12rc2_2821	Depositor
R, R_{free}	0.270　,　0.313 0.270　,　0.313	Depositor DCC
R_{free} test set	9232 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	9.6	Xtriage
Anisotropy	0.561	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 41.5	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.119 for h,-k,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	25754	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.49 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.1411e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: PGE, 9SN

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	0.29	0/3156	0.50	0/4292
1	B	0.34	0/3136	0.55	0/4262
1	C	0.29	0/3112	0.49	0/4233
1	D	0.34	0/3096	0.56	0/4210
All	All	0.32	0/12500	0.53	0/16997

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	3056	3017	3016	47	1
1	B	3042	3012	3012	24	0
1	C	3021	2964	2964	62	1
1	D	3011	2959	2958	23	0
2	A	33	18	0	1	0
2	B	33	18	0	0	0
2	C	33	18	0	1	0
2	D	33	18	0	1	0
3	C	10	14	14	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	360	0	0	29	8
4	B	388	0	0	16	6
4	C	352	0	0	44	8
4	D	344	0	0	16	6
All	All	13716	12038	11964	160	19

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 7.

The worst 5 of 160 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:D:309:VAL:O	4:D:1002:HOH:O	1.85	0.94
1:A:89:GLU:OE1	4:A:1001:HOH:O	1.85	0.94
1:D:367:GLU:O	4:D:1001:HOH:O	1.85	0.93
1:A:314:ARG:NH1	4:A:1005:HOH:O	2.04	0.91
1:D:385:GLU:OE2	1:D:393:ARG:NH1	2.05	0.89

The worst 5 of 19 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:1028:HOH:O	4:D:1138:HOH:O[1_545]	1.75	0.45
4:B:1292:HOH:O	4:D:1214:HOH:O[2_746]	1.88	0.32
4:C:1221:HOH:O	4:C:1303:HOH:O[2_545]	1.91	0.29
4:C:1010:HOH:O	4:C:1191:HOH:O[2_545]	1.95	0.25
4:B:1268:HOH:O	4:B:1321:HOH:O[2_646]	1.96	0.24

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	398/407 (98%)	391 (98%)	7 (2%)	0	100	100
1	B	394/407 (97%)	388 (98%)	6 (2%)	0	100	100
1	C	394/407 (97%)	388 (98%)	6 (2%)	0	100	100
1	D	390/407 (96%)	383 (98%)	7 (2%)	0	100	100
All	All	1576/1628 (97%)	1550 (98%)	26 (2%)	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	315/320 (98%)	313 (99%)	2 (1%)	86	79
1	B	312/320 (98%)	311 (100%)	1 (0%)	92	89
1	C	308/320 (96%)	308 (100%)	0	100	100
1	D	307/320 (96%)	306 (100%)	1 (0%)	92	89
All	All	1242/1280 (97%)	1238 (100%)	4 (0%)	92	89

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	240	LEU
1	A	337	ARG
1	B	117	LEU
1	D	117	LEU

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

Mol	Chain	Res	Type
1	C	264	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 carbohydrates 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	#$|Z| > 2$
2	9SN	C	900	-	33,36,36	0.66	0	39,54,54	0.90	3 (7%)
2	9SN	D	900	-	33,36,36	0.61	0	39,54,54	0.72	1 (2%)
2	9SN	A	900	-	33,36,36	0.66	0	39,54,54	0.72	1 (2%)
2	9SN	B	900	-	33,36,36	0.66	0	39,54,54	0.68	1 (2%)
3	PGE	C	901	-	9,9,9	0.27	0	8,8,8	0.47	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
2	9SN	C	900	-	-	3/15/35/35	0/4/4/4
2	9SN	D	900	-	-	2/15/35/35	0/4/4/4
2	9SN	A	900	-	-	4/15/35/35	0/4/4/4

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	9SN	B	900	-	-	3/15/35/35	0/4/4/4
3	PGE	C	901	-	-	4/7/7/7	-

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	900	9SN	O14-C24-O26	2.97	124.52	122.13
2	A	900	9SN	C3-C2-N1	2.30	123.85	120.35
2	D	900	9SN	C3-C2-N1	2.25	123.77	120.35
2	B	900	9SN	C3-C2-N1	2.20	123.70	120.35
2	C	900	9SN	P12-O14-C24	2.11	127.69	123.04

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	901	PGE	O2-C3-C4-O3
3	C	901	PGE	O1-C1-C2-O2
3	C	901	PGE	O3-C5-C6-O4
2	D	900	9SN	C24-O14-P12-O11
2	A	900	9SN	C24-O14-P12-O11

There are no ring outliers.
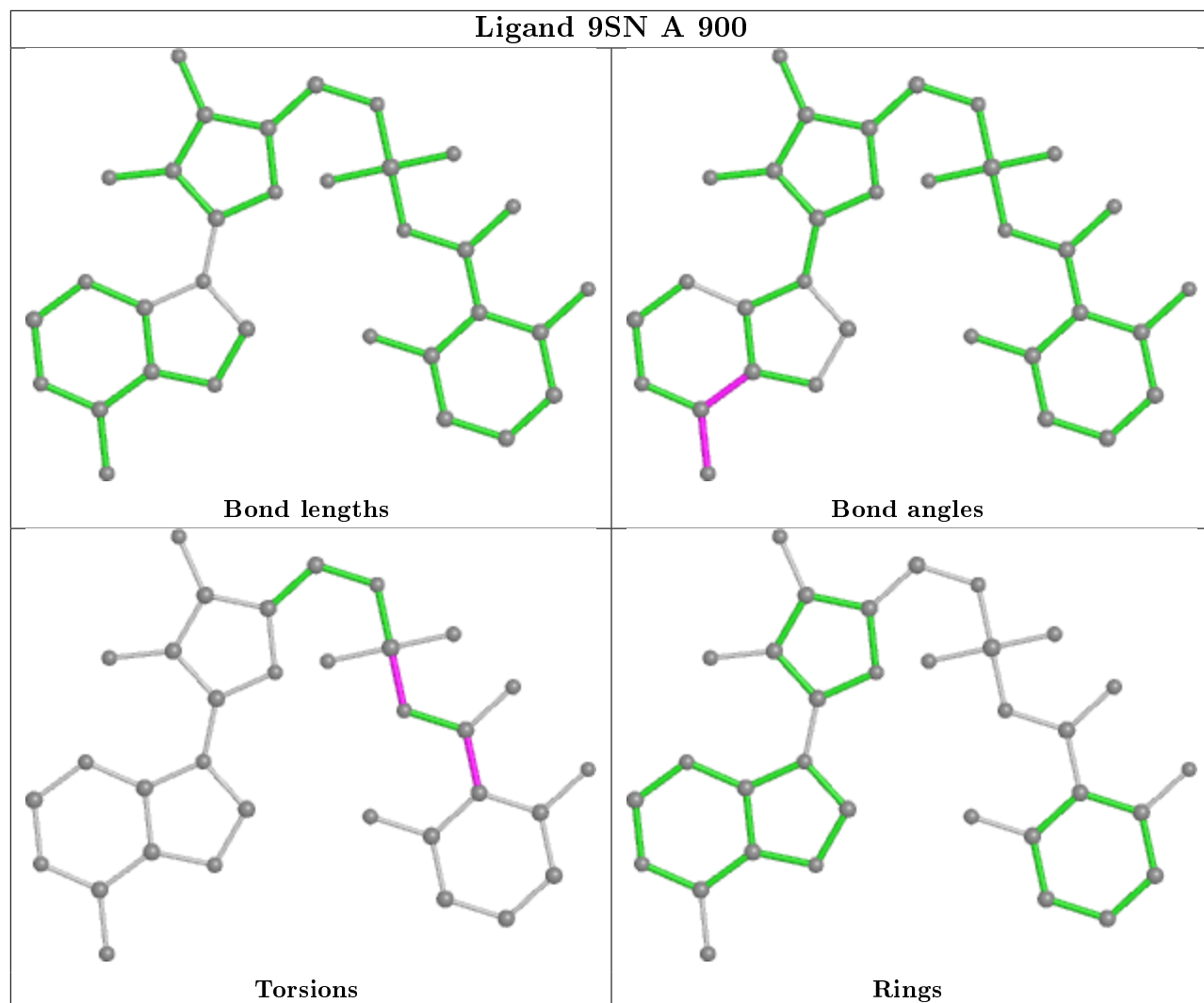
4 monomers are involved in 6 short contacts:

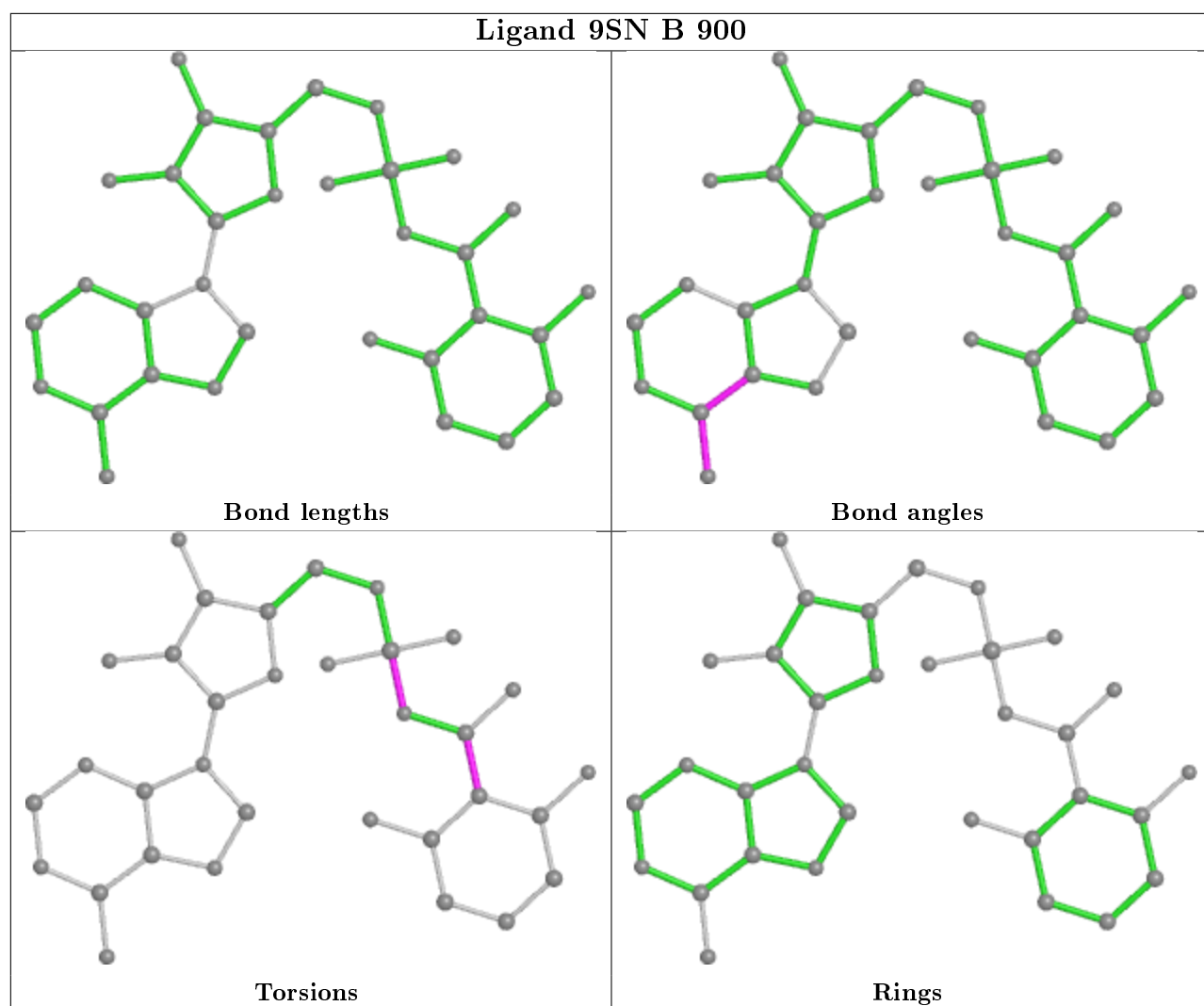| Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	900	9SN	1	0
2	D	900	9SN	1	0
2	A	900	9SN	1	0
3	C	901	PGE	3	0

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 9SN C 900

Bond lengths

Bond angles

Torsions

Rings

Ligand 9SN D 900

Bond lengths

Bond angles

Torsions

Rings

Ligand 9SN A 900

Bond lengths

Bond angles

Torsions

Rings

Ligand 9SN B 900

Bond lengths

Bond angles

Torsions

Rings
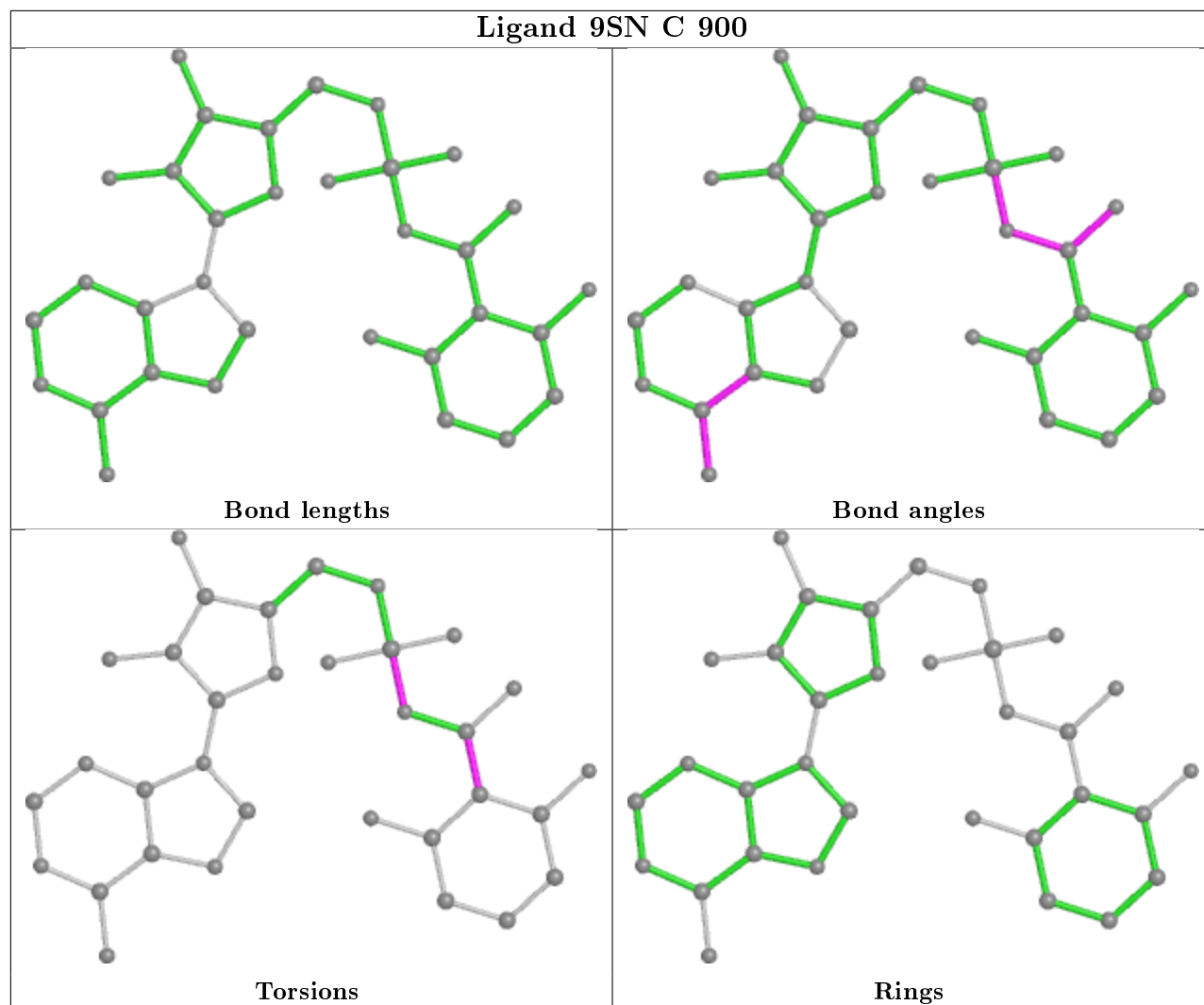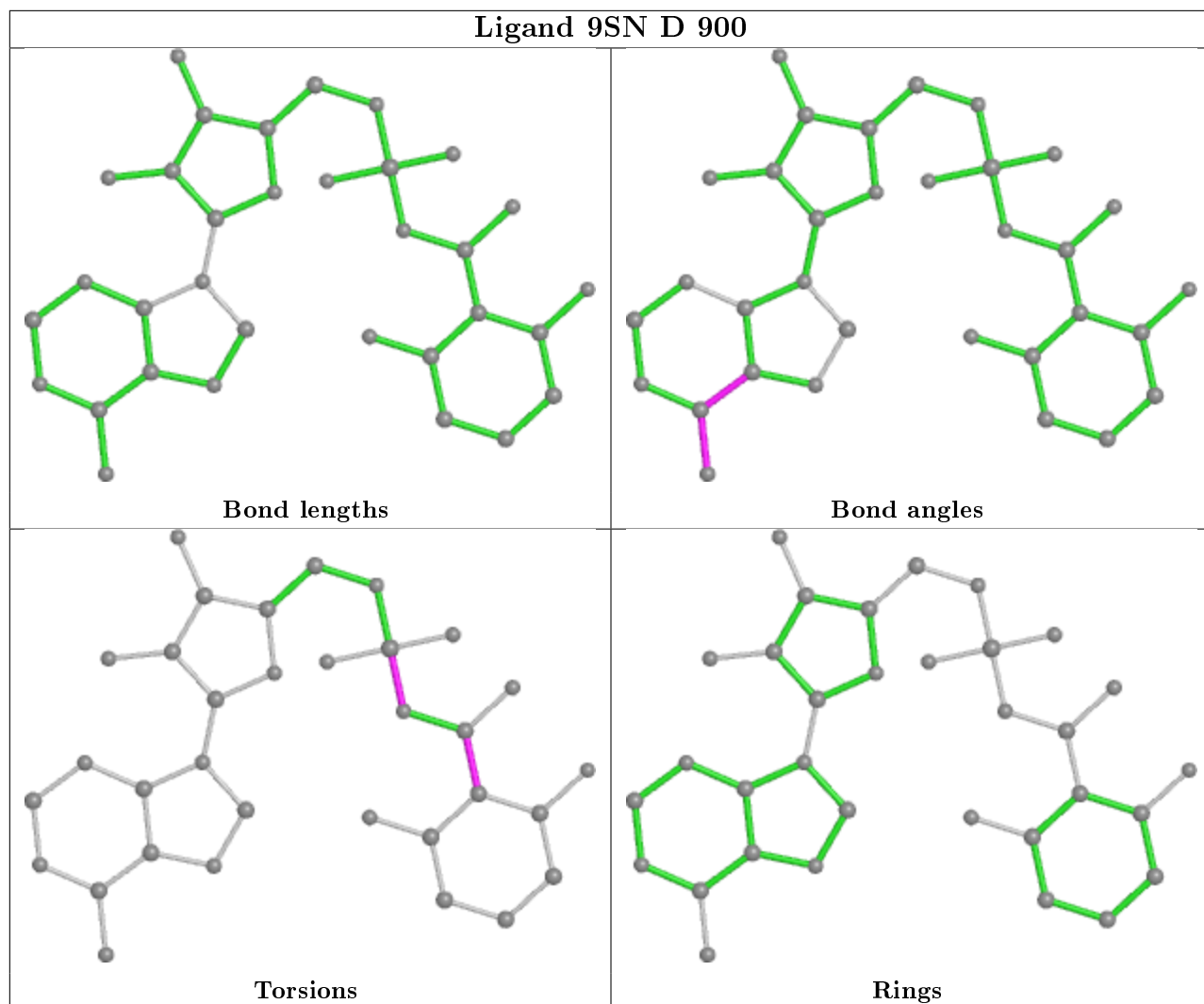
5.7 Other polymers ⓘ

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 (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(Å²)	Q<0.9
1	A	394/407 (96%)	1.44	93 (23%)	0	0	15, 24, 38, 54	0
1	B	392/407 (96%)	0.34	10 (2%)	56	58	3, 11, 24, 42	0
1	C	393/407 (96%)	1.59	112 (28%)	0	0	17, 26, 43, 57	0
1	D	391/407 (96%)	0.43	13 (3%)	46	49	3, 12, 27, 34	0
All	All	1570/1628 (96%)	0.95	228 (14%)	2	2	3, 21, 36, 57	0

The worst 5 of 228 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	118	ALA	6.5
1	A	233	TRP	5.7
1	A	209[A]	PHE	5.5
1	C	109	ASP	5.2
1	A	127	GLY	5.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.
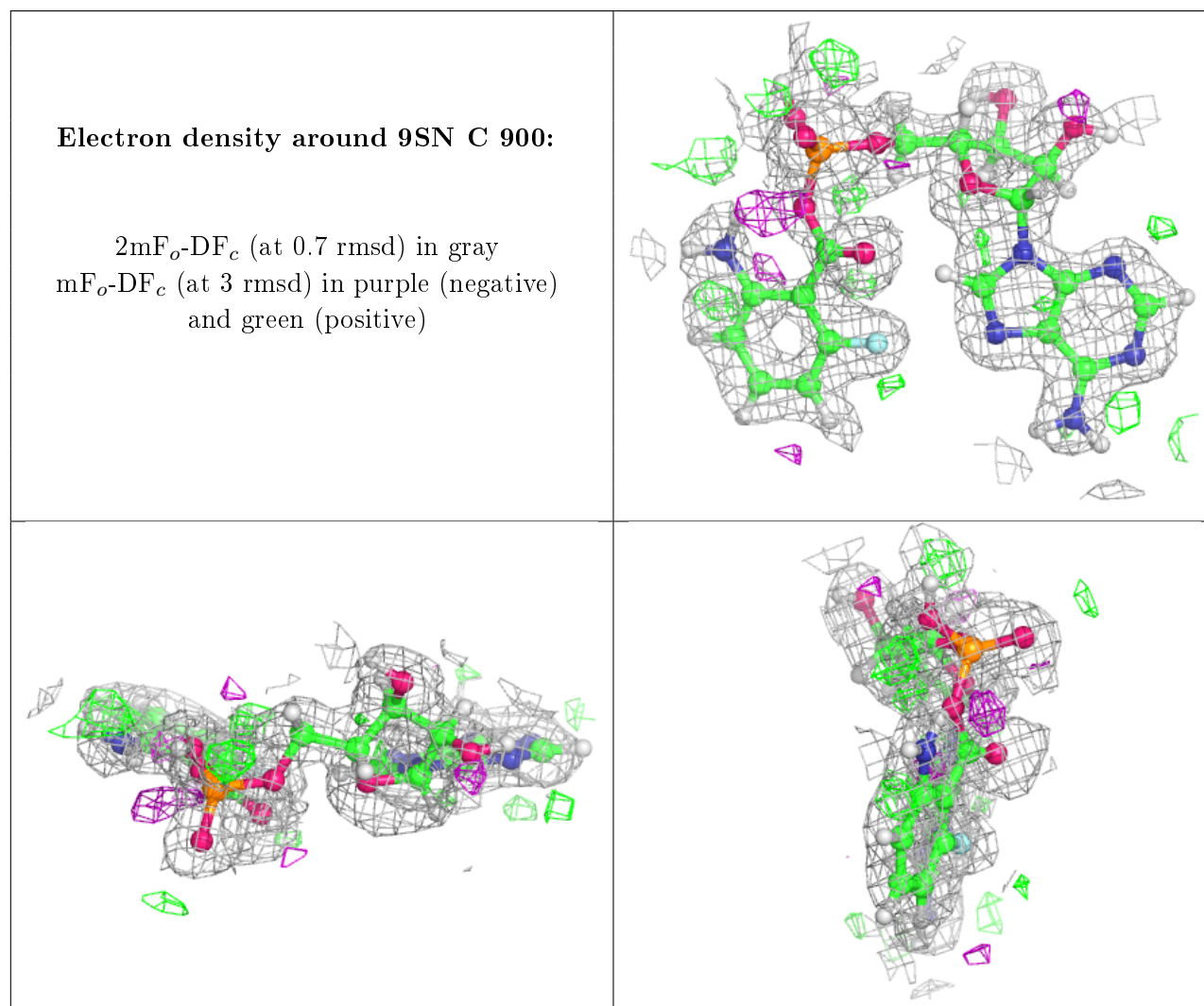
6.3 Carbohydrates (i)

There are no carbohydrates 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(Å²)	Q<0.9
3	PGE	C	901	10/10	0.83	0.21	23,35,42,50	0
2	9SN	C	900	33/33	0.87	0.13	5,15,24,34	0
2	9SN	A	900	33/33	0.93	0.11	8,19,25,28	0
2	9SN	D	900	33/33	0.94	0.11	8,8,11,23	0
2	9SN	B	900	33/33	0.96	0.09	8,8,10,13	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 9SN C 900:

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

Electron density around 9SN A 900:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
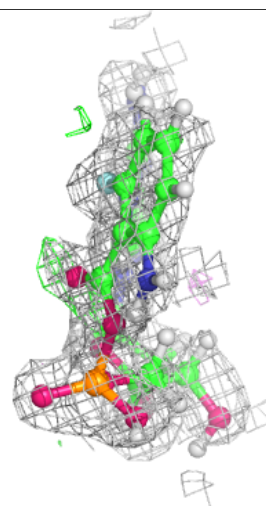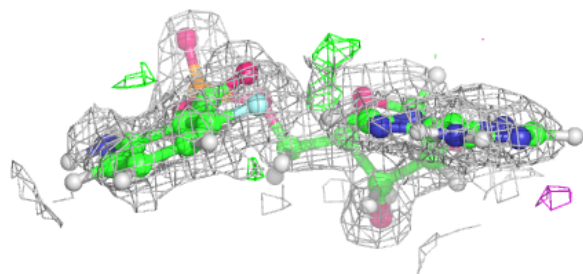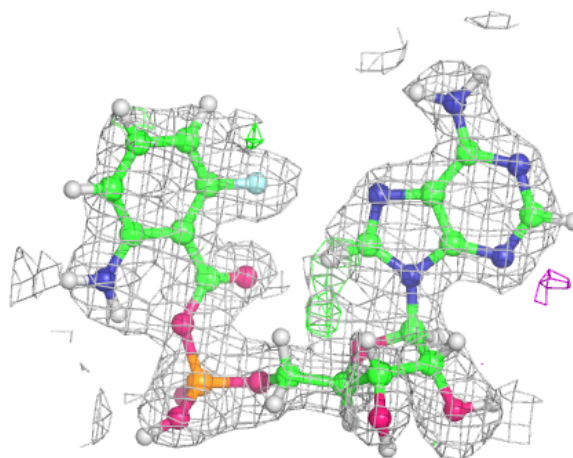
Electron density around 9SN D 900:

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

Electron density around 9SN B 900:

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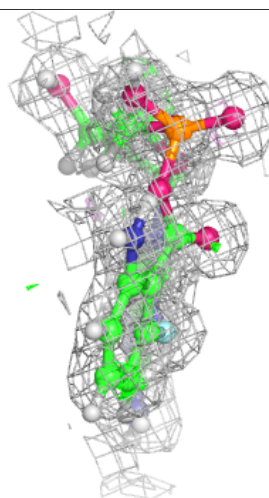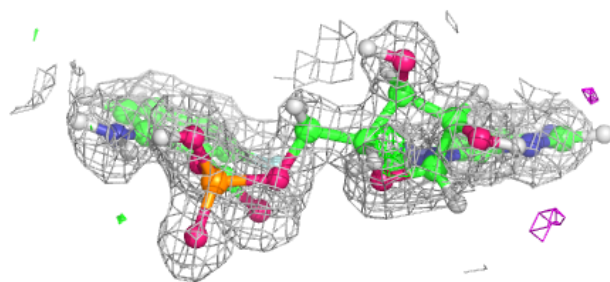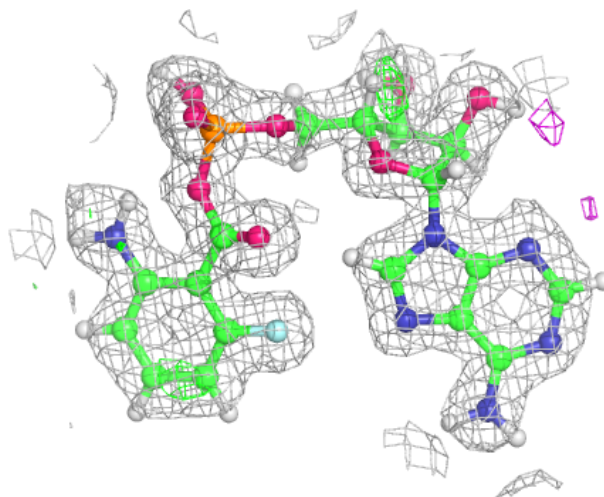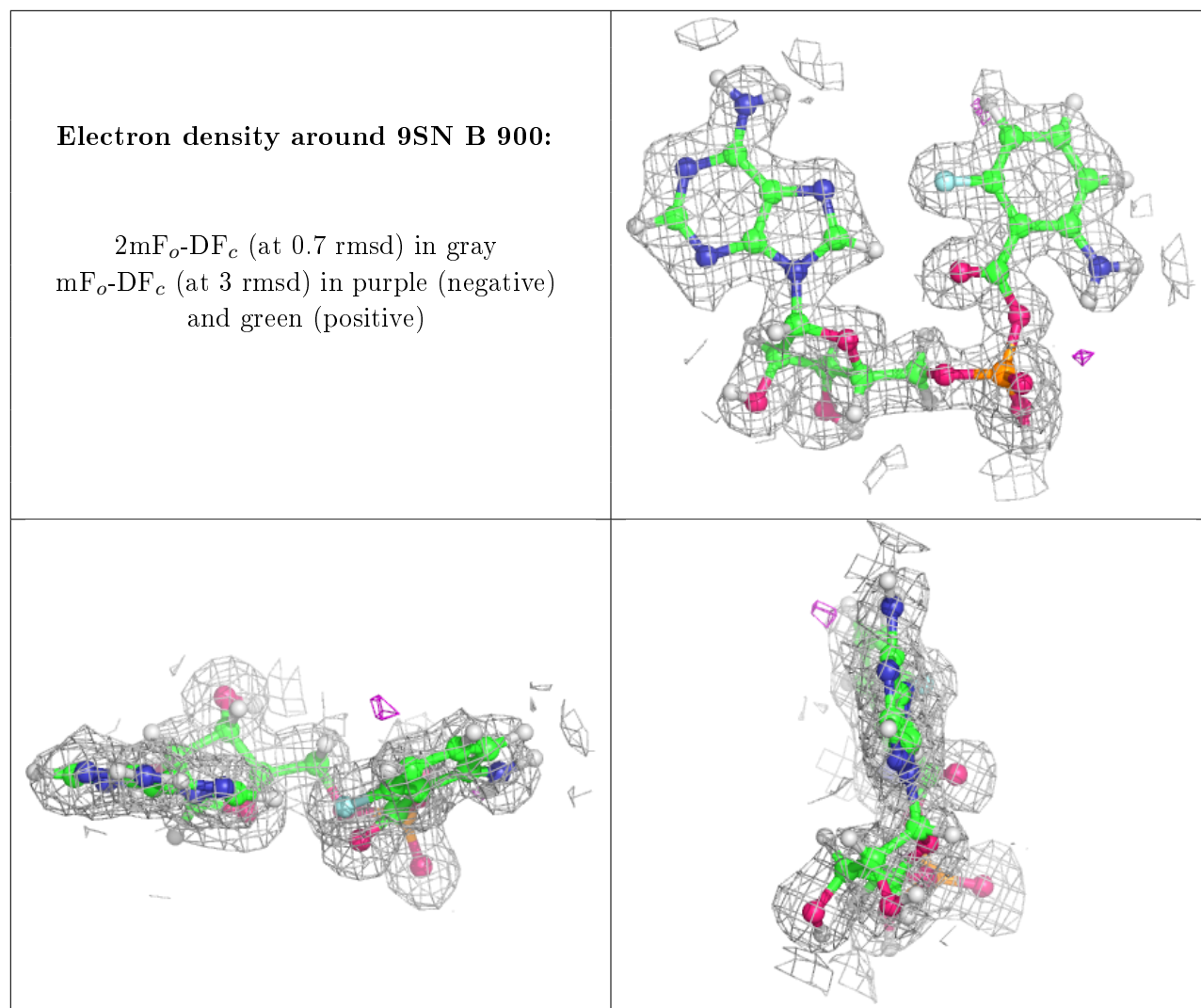
6.5 Other polymers (i)

There are no such residues in this entry.