WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Sep 25, 2023 – 01:09 pm BST

PDB ID	:	2C2T
Title	:	Human Dihydrofolate Reductase Complexed With NADPH and 2,4-Diamino-5-((7,8-dicarbaundecaboran-7-yl)methyl)-6-methylpyrimidine, a novel boron containing, nonclassical Antifolate
Authors	:	leung, A.K.W.; reynolds, R.C.; riordan, J.M.; borhani, D.W.
Deposited on	:	2005-09-30
Resolution	:	1.50 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.50 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.
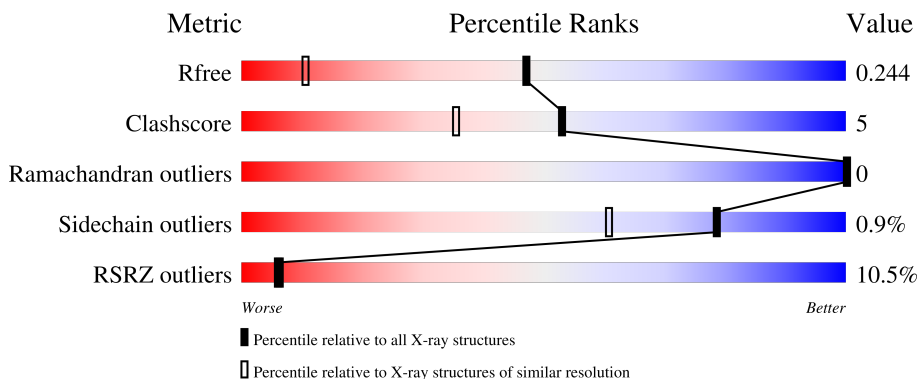
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	186	10% / 87% / 12% ·
1	B	186	11% / 91% / 9% ·

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 criteria:

| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	39B	A	1189[B]	X	-	-	-
3	39B	B	1189[B]	X	-	-	-

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 3925 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.

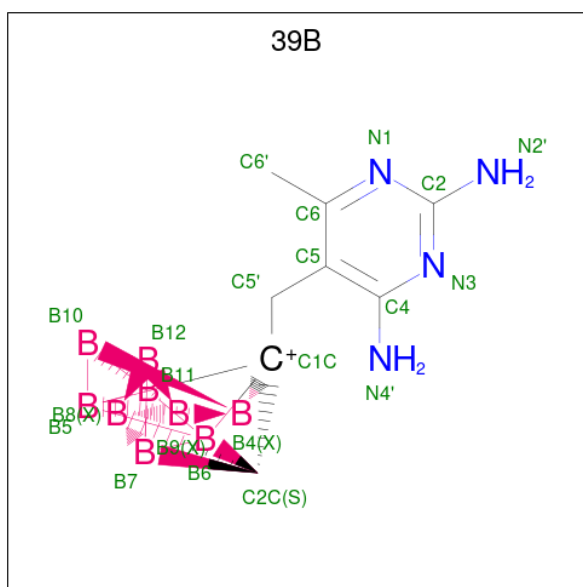- Molecule 1 is a protein called DIHYDROFOLATE REDUCTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	186	Total 1582	C 1022	N 264	O 289	S 7	15	17	0
1	B	186	Total 1550	C 994	N 257	O 291	S 8	27	11	0

- Molecule 2 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: $C_{21}H_{30}N_7O_{17}P_3$).
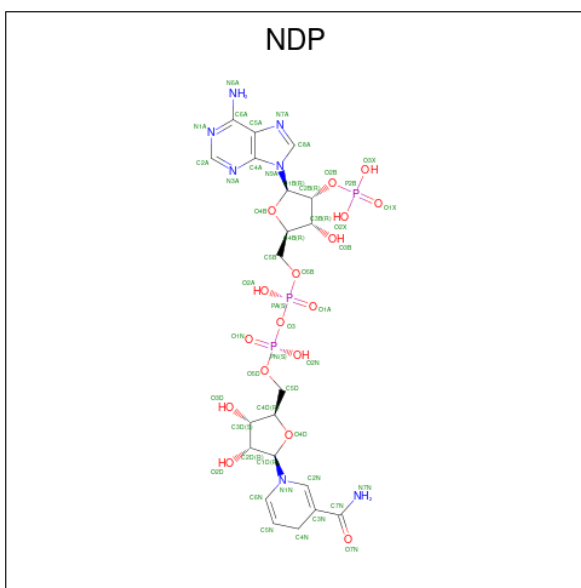
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 48	C 21	N 7	O 17	P 3	0	0
2	B	1	Total 48	C 21	N 7	O 17	P 3	0	0

- Molecule 3 is (S)-2,4-DIAMINO-5-((7,8-DICARBAUNDECABORAN-7-YL)METHYL)-6-METHYLPYRIMIDINE (three-letter code: 39B) (formula: $C_8H_9B_9N_4$).

39B

N1

C6'

C2 NH2 N2'

C6

C5

C5'

N N3

B10 B12

B B B11 C4

B B B8(X) B5 B B B B B4(X) C1C

NH2 N4'

B B9(X) B6

B7 C2C(S)

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	B	C	N	0	1
			21	9	8	4		
3	A	1	Total	B	C	N	0	1
			21	9	8	4		
3	B	1	Total	B	C	N	0	1
			21	9	8	4		
3	B	1	Total	B	C	N	0	1
			21	9	8	4		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

C3

OH O3

C1 C2

HO OH

O1 O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 6	C 3	O 3	0	0

• Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	297	Total 297	O 297	0	8
5	B	310	Total 310	O 310	0	7

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: DIHYDROFOLATE REDUCTASE
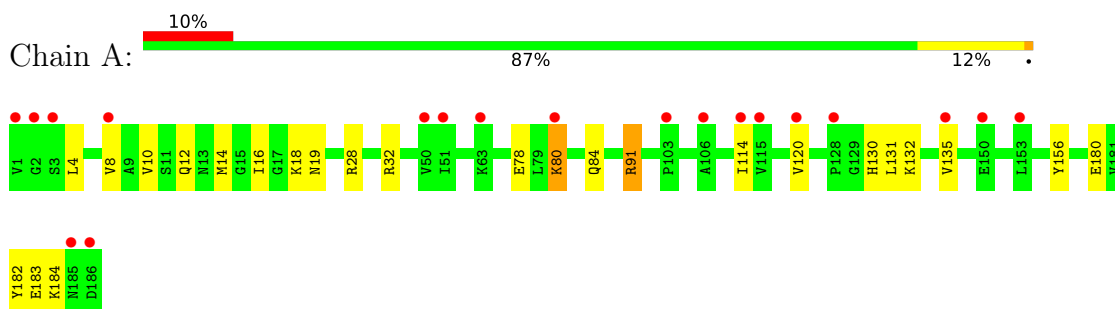
• Molecule 1: DIHYDROFOLATE REDUCTASE
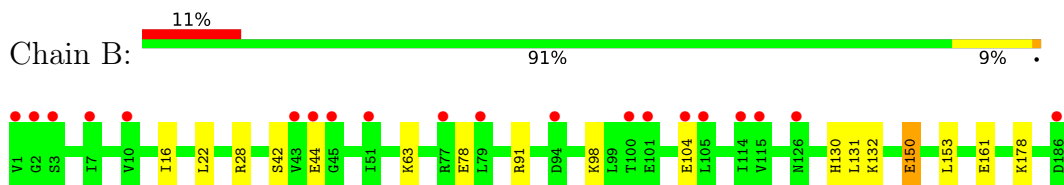
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	87.63Å 94.35Å 96.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.50 19.93 – 1.50	Depositor EDS
% Data completeness (in resolution range)	98.1 (20.00-1.50) 98.1 (19.93-1.50)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.71 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.158 , 0.213 0.194 , 0.244	Depositor DCC
R_{free} test set	3168 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	17.7	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 58.7	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3925	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.99% 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: GOL, NDP, 39B

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	0.93	4/1679 (0.2%)	0.75	4/2264 (0.2%)
1	B	1.67	6/1623 (0.4%)	1.00	7/2190 (0.3%)
All	All	1.34	10/3302 (0.3%)	0.88	11/4454 (0.2%)

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	1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	44	GLU	CB-CG	48.21	2.43	1.52
1	B	28	ARG	CD-NE	-24.66	1.04	1.46
1	B	150	GLU	CB-CG	-23.52	1.07	1.52
1	A	91	ARG	CD-NE	-19.00	1.14	1.46
1	B	63	LYS	CB-CG	-17.81	1.04	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	150	GLU	CB-CG-CD	19.94	168.03	114.20
1	B	28	ARG	CG-CD-NE	17.91	149.42	111.80
1	B	28	ARG	CD-NE-CZ	17.67	148.34	123.60
1	B	150	GLU	CA-CB-CG	10.78	137.11	113.40
1	B	63	LYS	CA-CB-CG	8.77	132.68	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	91	ARG	Sidechain

CLOSE-CONTACTS INFOmissingINFO

5.2 Torsion angles ⓘ

5.2.1 Protein backbone ⓘ

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	201/186 (108%)	198 (98%)	3 (2%)	0	100	100
1	B	195/186 (105%)	193 (99%)	2 (1%)	0	100	100
All	All	396/372 (106%)	391 (99%)	5 (1%)	0	100	100

There are no Ramachandran outliers to report.

5.2.2 Protein sidechains ⓘ

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	185/168 (110%)	185 (100%)	0	100	100
1	B	179/168 (106%)	176 (98%)	3 (2%)	60	33
All	All	364/336 (108%)	361 (99%)	3 (1%)	78	66

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	42	SER
1	B	131	LEU
1	B	150	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	127	HIS
1	A	130	HIS
1	A	140	GLN
1	B	29	ASN

5.2.3 RNA ⓘ

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.4 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.5 Ligand geometry ⓘ

7 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	#$|Z| > 2$
4	GOL	B	1190	-	5,5,5	0.70	0	5,5,5	0.37	0
3	39B	A	1188[A]	-	24,36,36	1.32	2 (8%)	62,109,109	0.79	2 (3%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	NDP	A	1187	-	45,52,52	1.19	3 (6%)	53,80,80	1.27	5 (9%)
3	39B	A	1189[B]	-	24,36,36	1.31	2 (8%)	62,109,109	0.76	2 (3%)
3	39B	B	1188[A]	-	24,36,36	1.16	2 (8%)	62,109,109	1.01	4 (6%)
3	39B	B	1189[B]	-	24,36,36	1.16	2 (8%)	62,109,109	1.01	4 (6%)
2	NDP	B	1187	-	45,52,52	1.38	4 (8%)	53,80,80	1.26	5 (9%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	B	1190	-	-	4/4/4/4	-
3	39B	A	1188[A]	-	-	2/5/300/300	0/1/16/16
2	NDP	A	1187	-	-	4/30/77/77	0/5/5/5
3	39B	A	1189[B]	-	1/1/11/11	2/5/300/300	0/1/16/16
3	39B	B	1188[A]	-	-	1/5/300/300	0/1/16/16
3	39B	B	1189[B]	-	1/1/11/11	1/5/300/300	0/1/16/16
2	NDP	B	1187	-	-	3/30/77/77	0/5/5/5

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1187	NDP	O7N-C7N	5.14	1.36	1.24
3	A	1188[A]	39B	C6-N1	4.80	1.41	1.34
3	A	1189[B]	39B	C6-N1	4.80	1.41	1.34
2	A	1187	NDP	O7N-C7N	4.58	1.35	1.24
2	B	1187	NDP	C2A-N3A	4.30	1.39	1.32

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1187	NDP	N3A-C2A-N1A	-5.63	119.89	128.68
2	B	1187	NDP	N3A-C2A-N1A	-5.15	120.63	128.68
3	B	1188[A]	39B	C2-N1-C6	4.50	120.46	116.79
3	B	1189[B]	39B	C2-N1-C6	4.50	120.46	116.79
2	B	1187	NDP	C3N-C7N-N7N	3.57	124.00	117.67

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	1189[B]	39B	C2C
3	B	1189[B]	39B	C2C

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1189[B]	39B	C2C-C1C-C5'-C5
3	A	1189[B]	39B	B5-C1C-C5'-C5
3	B	1189[B]	39B	C2C-C1C-C5'-C5
4	B	1190	GOL	O1-C1-C2-C3
4	B	1190	GOL	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1188[A]	39B	3	0
2	A	1187	NDP	2	0
3	A	1189[B]	39B	1	0
3	B	1188[A]	39B	4	0
3	B	1189[B]	39B	1	0
2	B	1187	NDP	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.
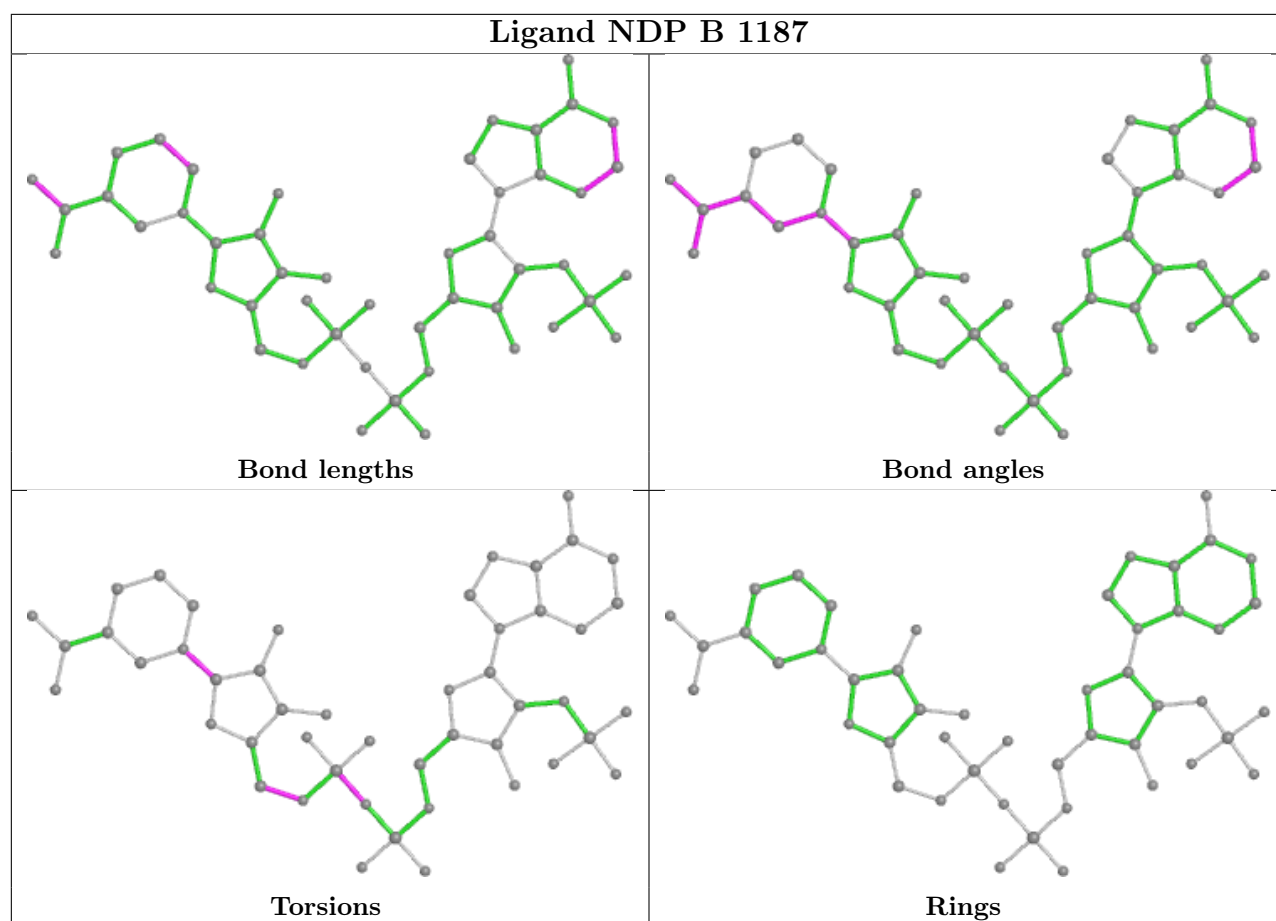
Ligand 39B A 1188 (A)

Bond lengths

Bond angles

Torsions

Rings

Ligand NDP A 1187

Bond lengths

Bond angles

Torsions

Rings

Ligand 39B A 1189 (B)

Bond lengths

Bond angles

Torsions

Rings

Ligand 39B B 1188 (A)

Bond lengths

Bond angles

Torsions

Rings

Ligand 39B B 1189 (B)

Bond lengths

Bond angles

Torsions

Rings

Ligand NDP B 1187

Bond lengths

Bond angles

Torsions

Rings
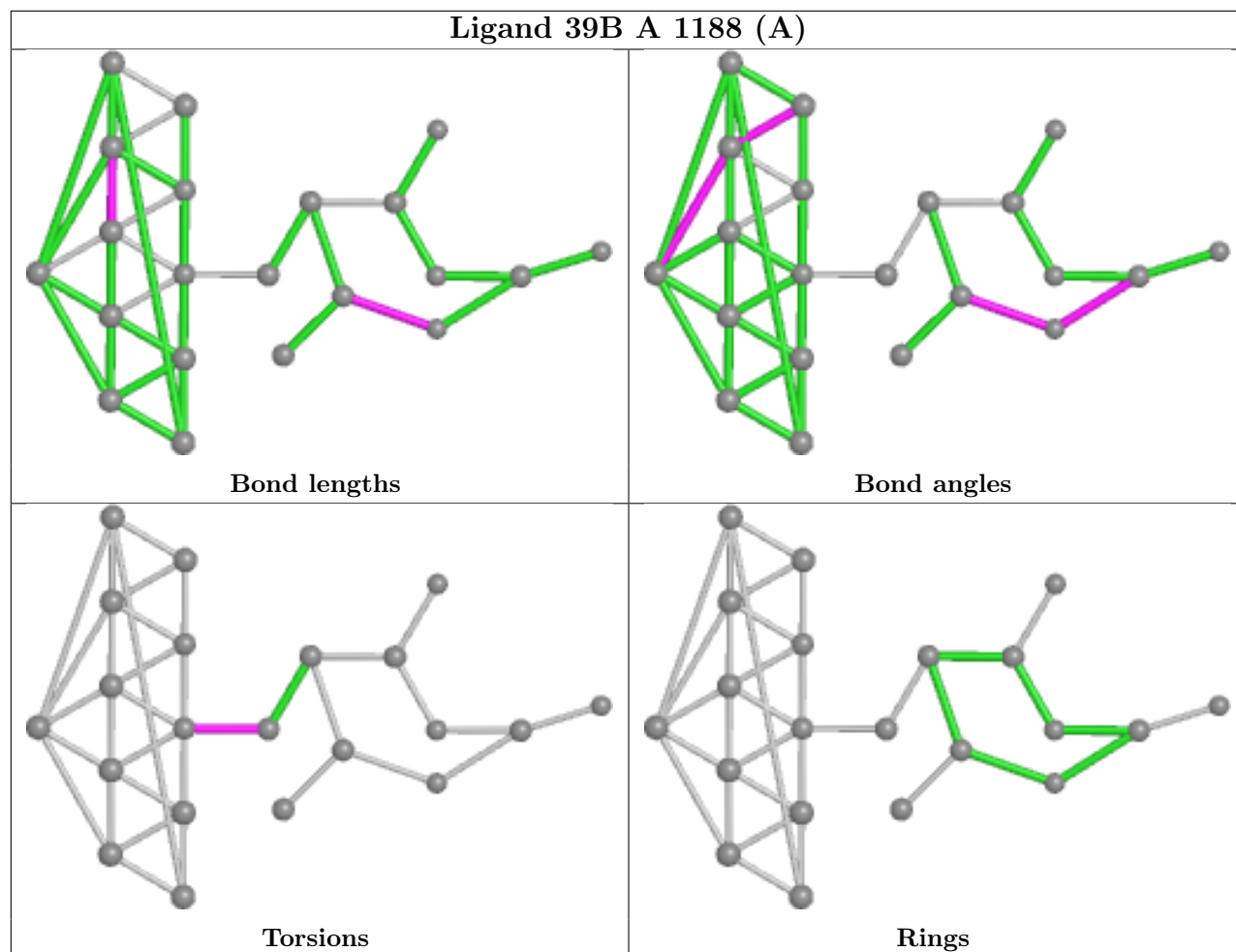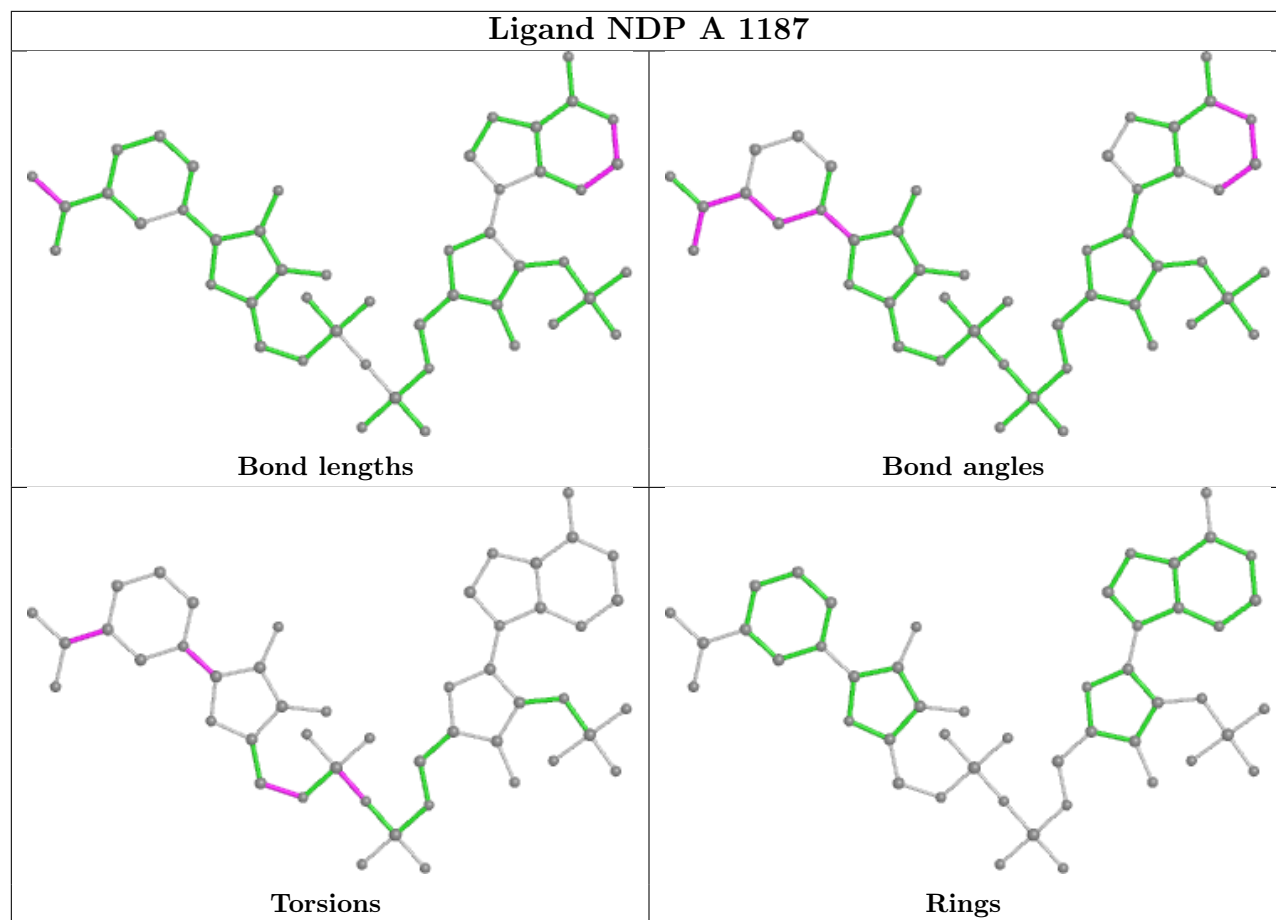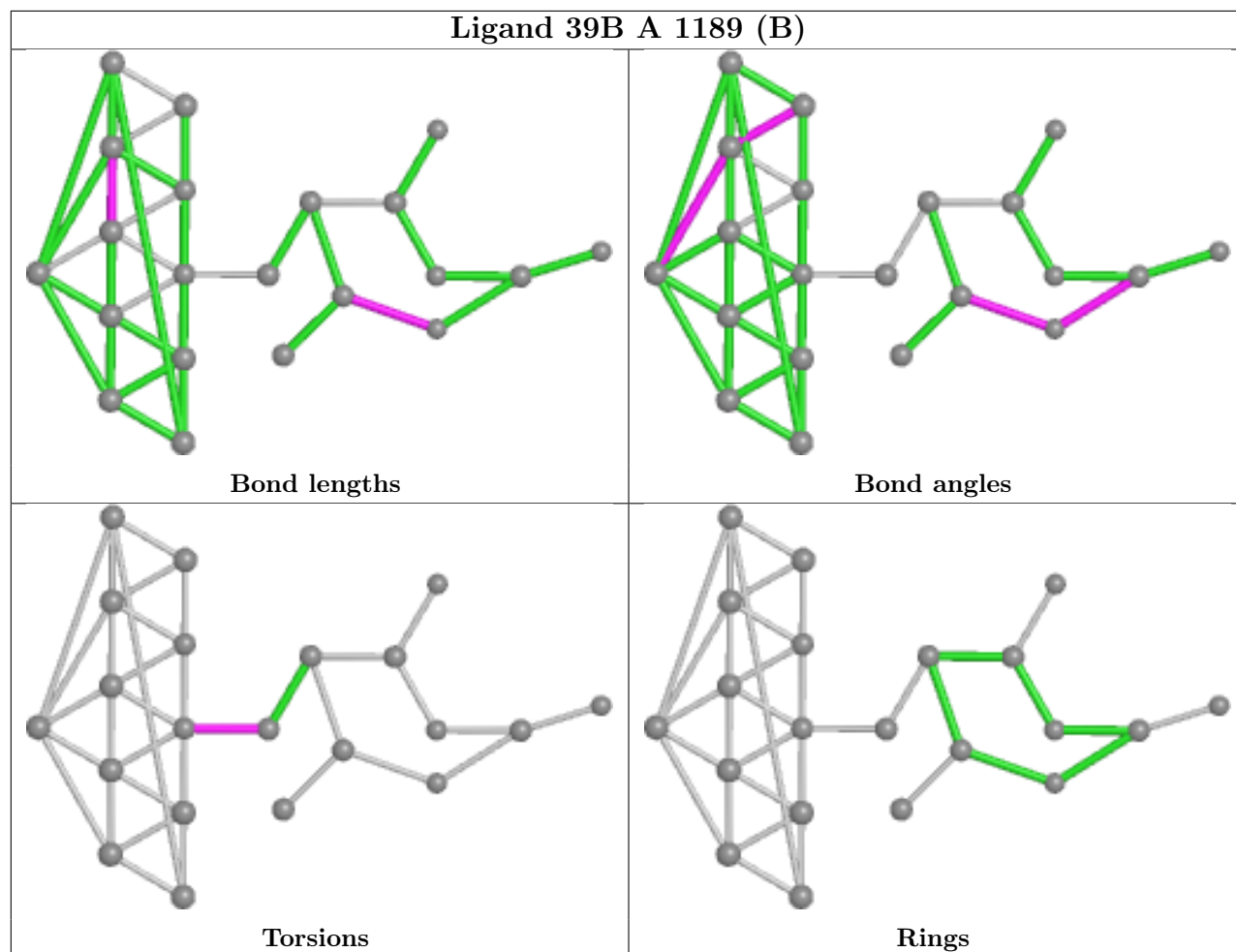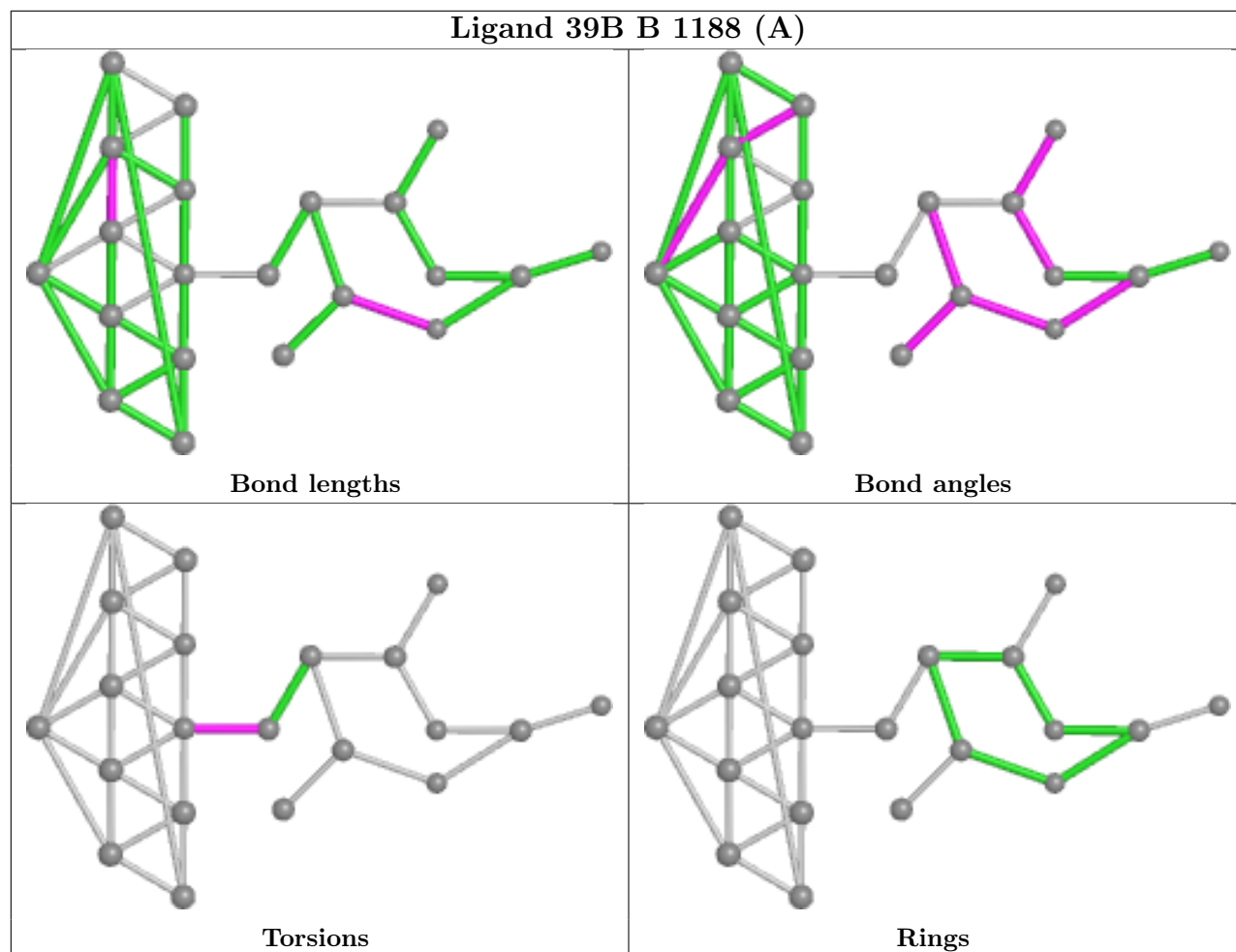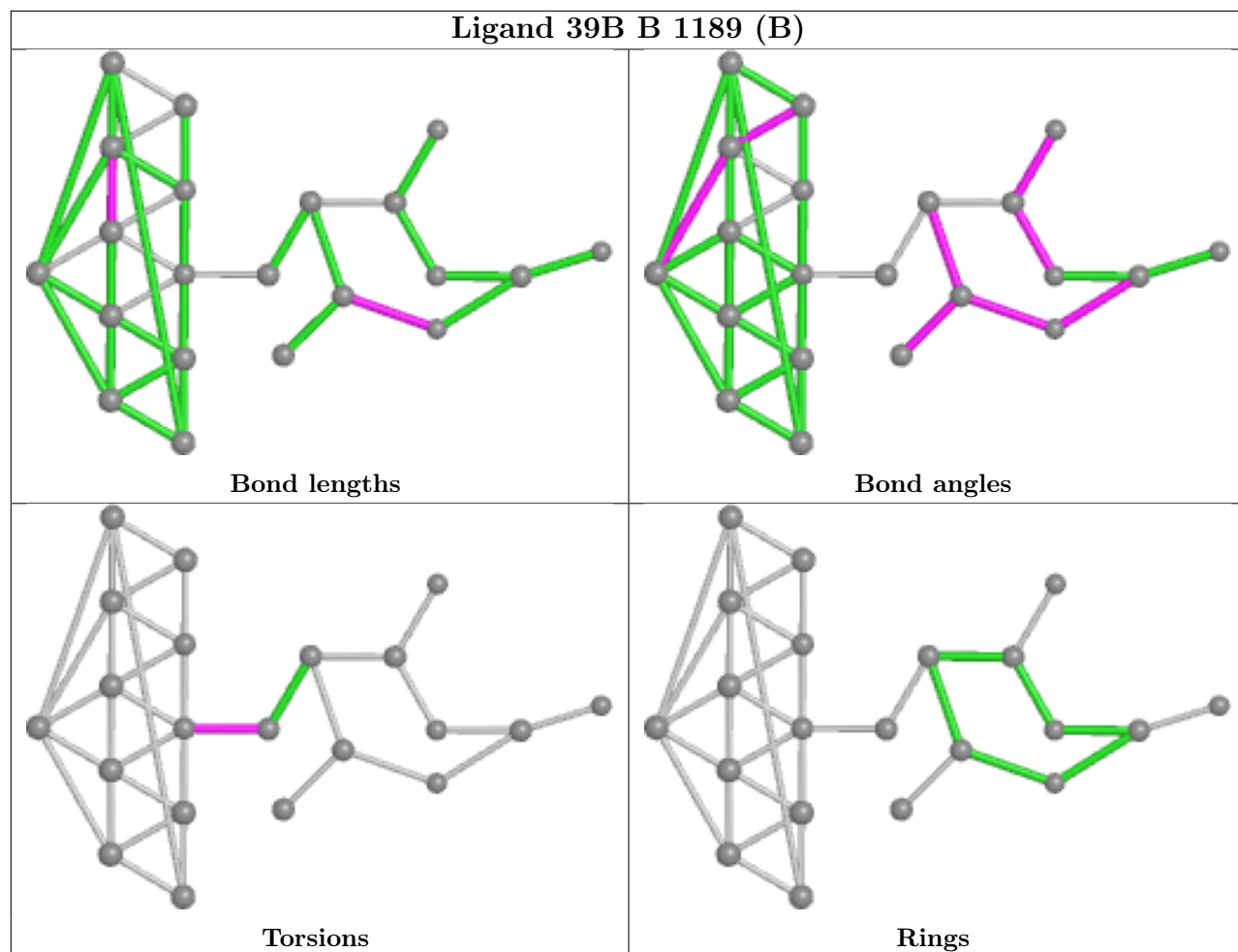
5.6 Other polymers ⓘ

There are no such residues in this entry.

5.7 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, 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	186/186 (100%)	0.62	19 (10%)	6	7	14, 20, 30, 37	4 (2%)
1	B	186/186 (100%)	0.73	20 (10%)	5	5	15, 22, 31, 36	7 (3%)
All	All	372/372 (100%)	0.68	39 (10%)	6	6	14, 21, 31, 37	11 (2%)

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1	VAL	9.3
1	A	1	VAL	8.0
1	A	103	PRO	6.9
1	B	2	GLY	6.2
1	A	3	SER	5.4

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

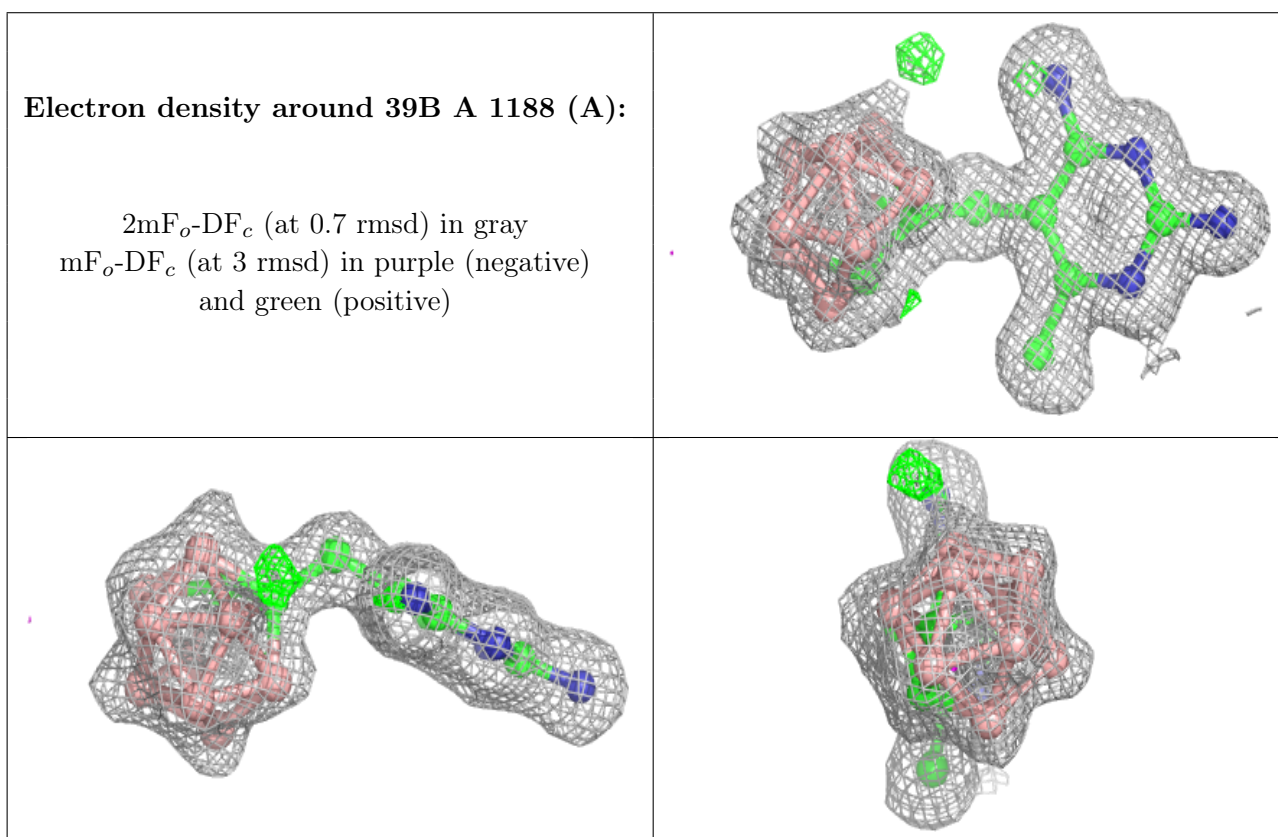There are no non-standard protein/DNA/RNA residues in this entry.

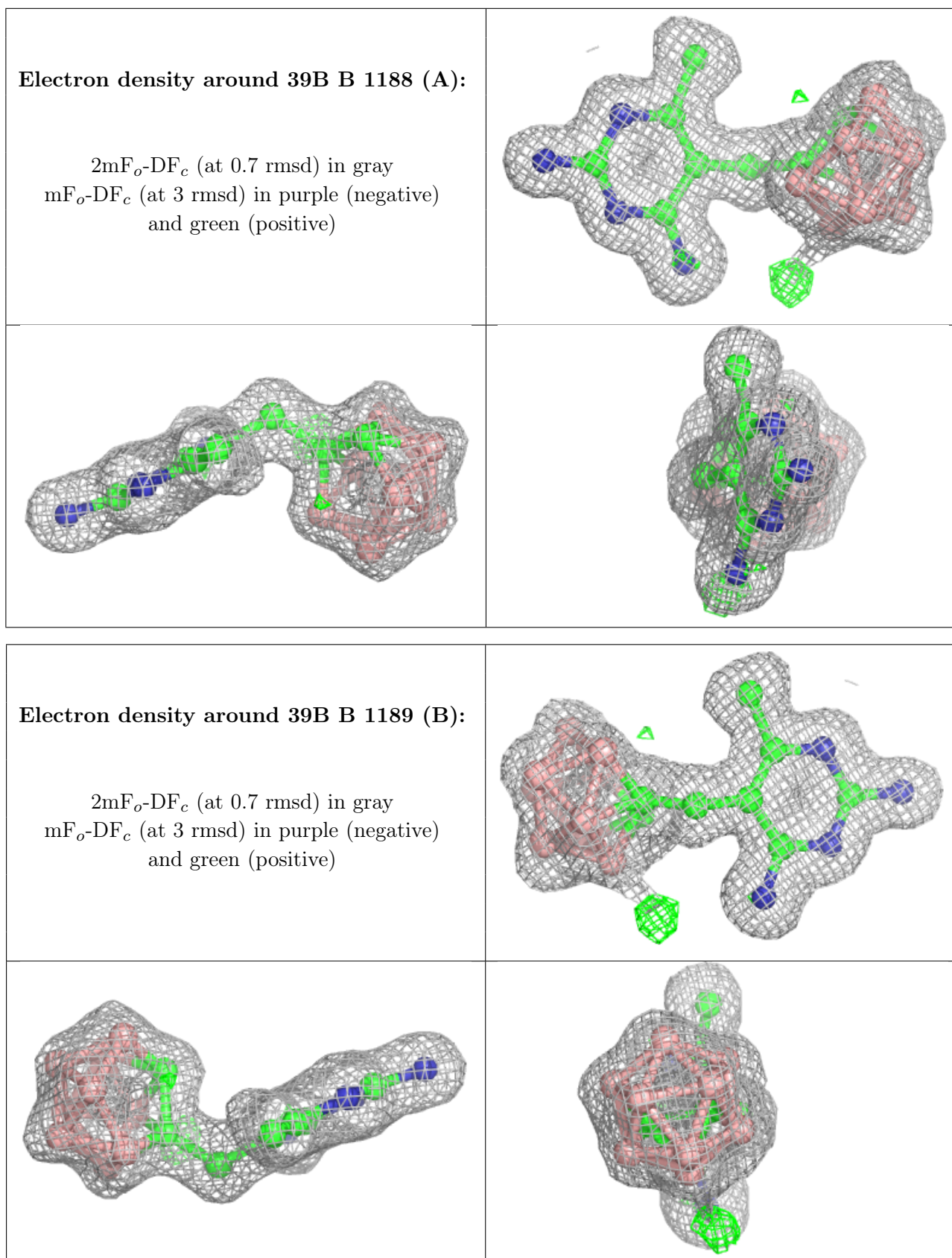## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GOL	B	1190	6/6	0.73	0.34	35,36,38,38	0
3	39B	A	1188[A]	21/21	0.94	0.12	16,19,21,22	21
3	39B	B	1188[A]	21/21	0.94	0.15	19,21,22,23	21
3	39B	B	1189[B]	21/21	0.94	0.14	19,21,22,22	21
2	NDP	B	1187	48/48	0.94	0.10	15,20,26,29	0
3	39B	A	1189[B]	21/21	0.95	0.12	16,19,21,22	21
2	NDP	A	1187	48/48	0.96	0.08	15,19,23,25	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.
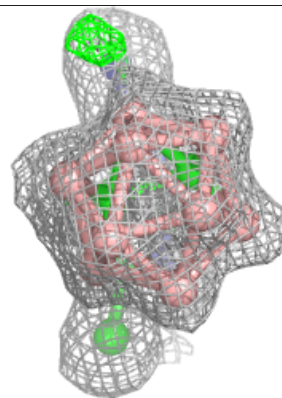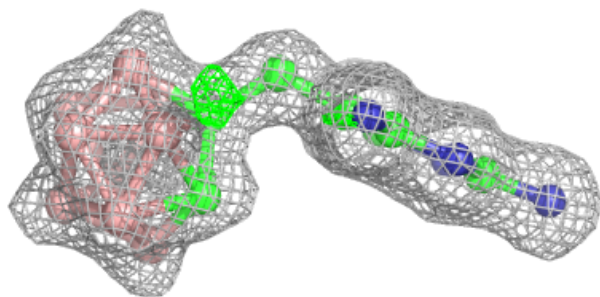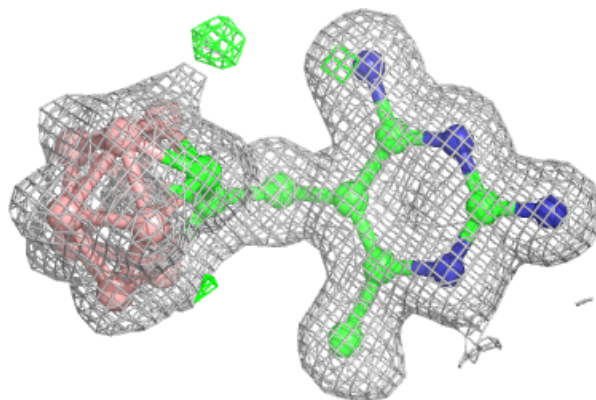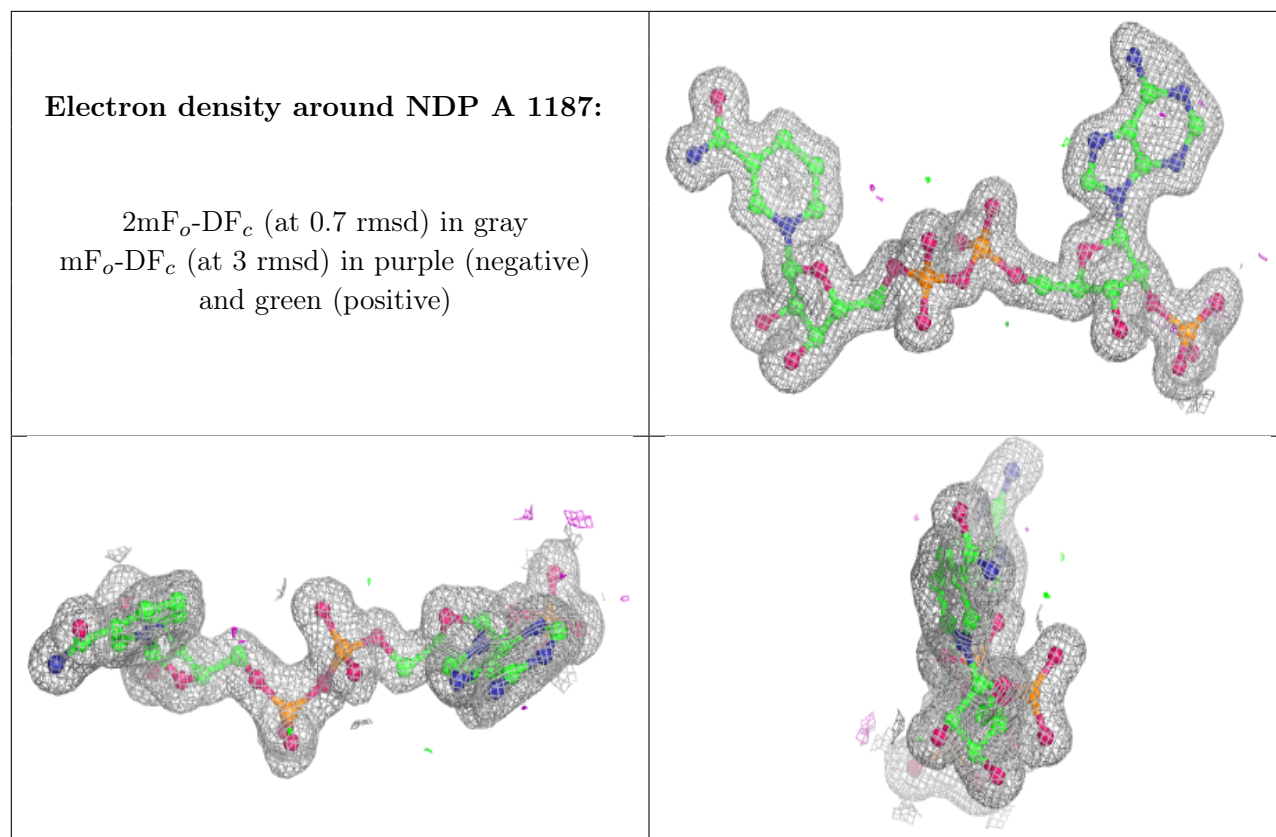

Electron density around 39B A 1188 (A):

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

Electron density around 39B B 1188 (A):

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

Electron density around 39B B 1189 (B):

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

Electron density around NDP B 1187:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
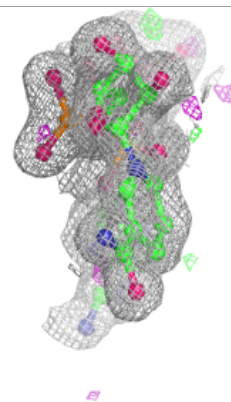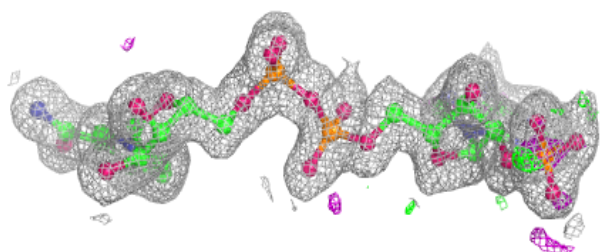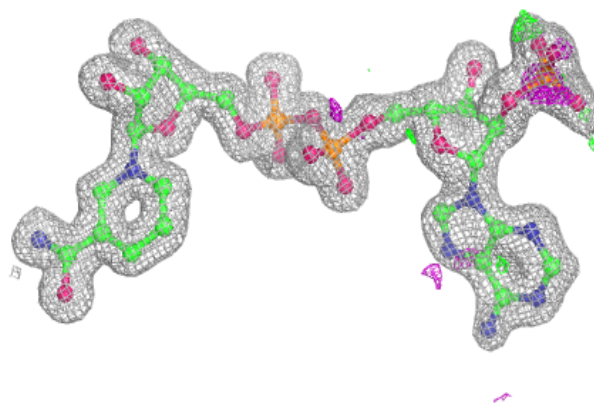
Electron density around 39B A 1189 (B):

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

Electron density around NDP A 1187:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.