



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 05:48 pm BST

PDB ID : 5Z2L
Title : Crystal structure of BdcA in complex with NADPH
Authors : Yang, W.S.; Hou, Y.J.; Li, D.F.; Wang, D.C.
Deposited on : 2018-01-03
Resolution : 1.70 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 ⓘ symbol.

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 : 2.11
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

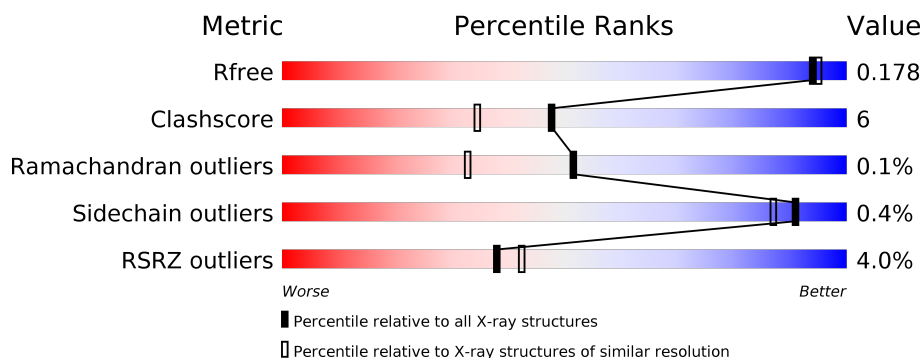
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 $\leq 5\%$. 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	245	<div> <div>3%</div> <div> <div></div> <div>90%</div> <div>7%</div> <div>.</div> </div> </div>
1	B	245	<div> <div>3%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div>.</div> </div> </div>
1	C	245	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>12%</div> <div>..</div> </div> </div>
1	D	245	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>9%</div> <div>.</div> </div> </div>
1	E	245	<div> <div>3%</div> <div> <div></div> <div>86%</div> <div>12%</div> <div>.</div> </div> </div>
1	F	245	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div>10%</div> <div>.</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	245	
1	H	245	
1	I	245	
1	J	245	
1	K	245	
1	L	245	

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
7	PG4	F	301	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 25111 atoms, of which 404 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 Cyclic-di-GMP-binding biofilm dispersal mediator protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	239	Total	C	N	O	S	0	10	0
			1784	1111	330	332	11			
1	D	240	Total	C	N	O	S	0	7	0
			1776	1103	333	329	11			
1	B	243	Total	C	N	O	S	0	11	0
			1823	1130	349	333	11			
1	C	241	Total	C	N	O	S	0	17	0
			1842	1143	345	342	12			
1	E	239	Total	C	N	O	S	0	6	0
			1772	1100	335	328	9			
1	J	242	Total	C	N	O	S	0	12	0
			1815	1125	340	339	11			
1	H	239	Total	C	N	O	S	0	5	0
			1757	1094	324	328	11			
1	K	244	Total	C	N	O	S	0	19	0
			1882	1167	360	343	12			
1	I	241	Total	C	N	O	S	0	9	0
			1786	1112	330	333	11			
1	F	239	Total	C	N	O	S	0	8	0
			1768	1103	326	328	11			
1	G	239	Total	C	N	O	S	0	9	0
			1769	1102	325	331	11			
1	L	239	Total	C	N	O	S	0	11	0
			1777	1103	331	332	11			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	238	LEU	-	expression tag	UNP P39333
A	239	GLU	-	expression tag	UNP P39333
A	240	HIS	-	expression tag	UNP P39333
A	241	HIS	-	expression tag	UNP P39333
A	242	HIS	-	expression tag	UNP P39333

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	243	HIS	-	expression tag	UNP P39333
A	244	HIS	-	expression tag	UNP P39333
A	245	HIS	-	expression tag	UNP P39333
D	238	LEU	-	expression tag	UNP P39333
D	239	GLU	-	expression tag	UNP P39333
D	240	HIS	-	expression tag	UNP P39333
D	241	HIS	-	expression tag	UNP P39333
D	242	HIS	-	expression tag	UNP P39333
D	243	HIS	-	expression tag	UNP P39333
D	244	HIS	-	expression tag	UNP P39333
D	245	HIS	-	expression tag	UNP P39333
B	238	LEU	-	expression tag	UNP P39333
B	239	GLU	-	expression tag	UNP P39333
B	240	HIS	-	expression tag	UNP P39333
B	241	HIS	-	expression tag	UNP P39333
B	242	HIS	-	expression tag	UNP P39333
B	243	HIS	-	expression tag	UNP P39333
B	244	HIS	-	expression tag	UNP P39333
B	245	HIS	-	expression tag	UNP P39333
C	238	LEU	-	expression tag	UNP P39333
C	239	GLU	-	expression tag	UNP P39333
C	240	HIS	-	expression tag	UNP P39333
C	241	HIS	-	expression tag	UNP P39333
C	242	HIS	-	expression tag	UNP P39333
C	243	HIS	-	expression tag	UNP P39333
C	244	HIS	-	expression tag	UNP P39333
C	245	HIS	-	expression tag	UNP P39333
E	238	LEU	-	expression tag	UNP P39333
E	239	GLU	-	expression tag	UNP P39333
E	240	HIS	-	expression tag	UNP P39333
E	241	HIS	-	expression tag	UNP P39333
E	242	HIS	-	expression tag	UNP P39333
E	243	HIS	-	expression tag	UNP P39333
E	244	HIS	-	expression tag	UNP P39333
E	245	HIS	-	expression tag	UNP P39333
J	238	LEU	-	expression tag	UNP P39333
J	239	GLU	-	expression tag	UNP P39333
J	240	HIS	-	expression tag	UNP P39333
J	241	HIS	-	expression tag	UNP P39333
J	242	HIS	-	expression tag	UNP P39333
J	243	HIS	-	expression tag	UNP P39333
J	244	HIS	-	expression tag	UNP P39333

Continued on next page...

Continued from previous page...

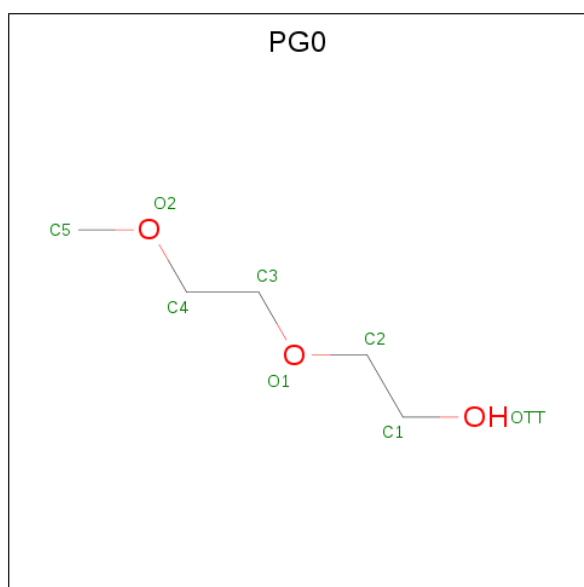
Chain	Residue	Modelled	Actual	Comment	Reference
J	245	HIS	-	expression tag	UNP P39333
H	238	LEU	-	expression tag	UNP P39333
H	239	GLU	-	expression tag	UNP P39333
H	240	HIS	-	expression tag	UNP P39333
H	241	HIS	-	expression tag	UNP P39333
H	242	HIS	-	expression tag	UNP P39333
H	243	HIS	-	expression tag	UNP P39333
H	244	HIS	-	expression tag	UNP P39333
H	245	HIS	-	expression tag	UNP P39333
K	238	LEU	-	expression tag	UNP P39333
K	239	GLU	-	expression tag	UNP P39333
K	240	HIS	-	expression tag	UNP P39333
K	241	HIS	-	expression tag	UNP P39333
K	242	HIS	-	expression tag	UNP P39333
K	243	HIS	-	expression tag	UNP P39333
K	244	HIS	-	expression tag	UNP P39333
K	245	HIS	-	expression tag	UNP P39333
I	238	LEU	-	expression tag	UNP P39333
I	239	GLU	-	expression tag	UNP P39333
I	240	HIS	-	expression tag	UNP P39333
I	241	HIS	-	expression tag	UNP P39333
I	242	HIS	-	expression tag	UNP P39333
I	243	HIS	-	expression tag	UNP P39333
I	244	HIS	-	expression tag	UNP P39333
I	245	HIS	-	expression tag	UNP P39333
F	238	LEU	-	expression tag	UNP P39333
F	239	GLU	-	expression tag	UNP P39333
F	240	HIS	-	expression tag	UNP P39333
F	241	HIS	-	expression tag	UNP P39333
F	242	HIS	-	expression tag	UNP P39333
F	243	HIS	-	expression tag	UNP P39333
F	244	HIS	-	expression tag	UNP P39333
F	245	HIS	-	expression tag	UNP P39333
G	238	LEU	-	expression tag	UNP P39333
G	239	GLU	-	expression tag	UNP P39333
G	240	HIS	-	expression tag	UNP P39333
G	241	HIS	-	expression tag	UNP P39333
G	242	HIS	-	expression tag	UNP P39333
G	243	HIS	-	expression tag	UNP P39333
G	244	HIS	-	expression tag	UNP P39333
G	245	HIS	-	expression tag	UNP P39333
L	238	LEU	-	expression tag	UNP P39333

Continued on next page...

Continued from previous page...

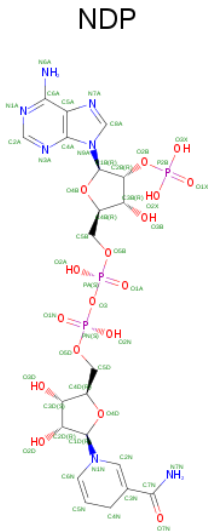
Chain	Residue	Modelled	Actual	Comment	Reference
L	239	GLU	-	expression tag	UNP P39333
L	240	HIS	-	expression tag	UNP P39333
L	241	HIS	-	expression tag	UNP P39333
L	242	HIS	-	expression tag	UNP P39333
L	243	HIS	-	expression tag	UNP P39333
L	244	HIS	-	expression tag	UNP P39333
L	245	HIS	-	expression tag	UNP P39333

- Molecule 2 is 2-(2-METHOXYETHOXY)ETHANOL (three-letter code: PG0) (formula: $C_5H_{12}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			20	5	12	3		

- Molecule 3 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: $C_{21}H_{30}N_7O_{17}P_3$) (labeled as "Ligand of Interest" by author).



Id	Chain	Residues	Atoms					ZeroOcc	AltConf	
3	A	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	D	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	B	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	C	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	E	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	J	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	H	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	K	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	I	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	F	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	G	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0
3	L	1	Total 74	C 21	H 26	N 7	O 17	P 3	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $\text{C}_2\text{H}_6\text{O}_2$).



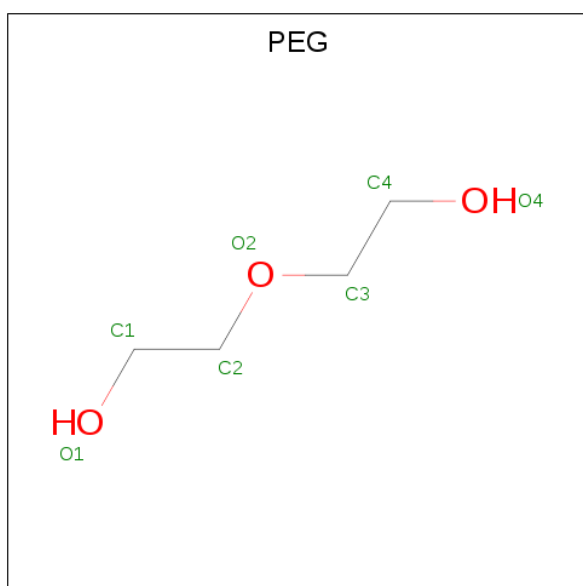
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	H	O	0	0
			10	2	6	2		
4	C	1	Total	C	H	O	0	0
			10	2	6	2		
4	C	1	Total	C	H	O	0	0
			10	2	6	2		
4	E	1	Total	C	H	O	0	0
			10	2	6	2		
4	I	1	Total	C	H	O	0	0
			10	2	6	2		
4	G	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



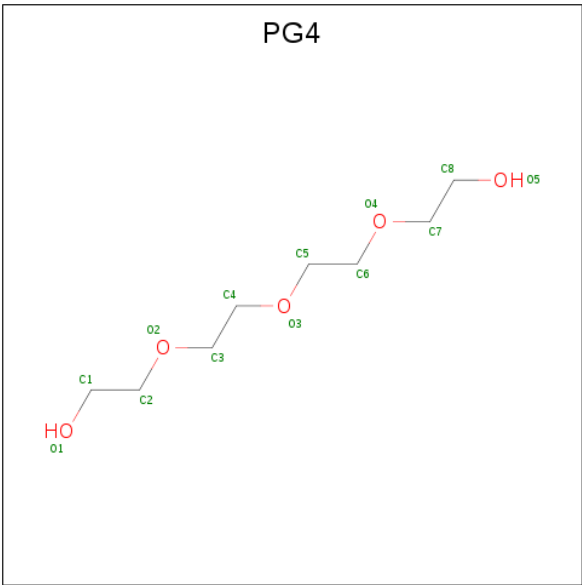
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total	C	H	O	0	0
			14	3	8	3		
5	E	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	K	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 7 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	F	1	Total	C	H	O	0	0
			31	8	18	5		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	155	Total	O	0	0
			155	155		
8	D	215	Total	O	0	0
			215	215		
8	B	185	Total	O	0	0
			185	185		
8	C	245	Total	O	0	0
			245	245		
8	E	185	Total	O	0	0
			185	185		
8	J	274	Total	O	0	0
			274	274		
8	H	118	Total	O	0	0
			118	118		
8	K	297	Total	O	0	0
			297	297		
8	I	265	Total	O	0	0
			265	265		
8	F	189	Total	O	0	0
			189	189		
8	G	94	Total	O	0	0
			94	94		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	L	294	Total 294	O 294	0	0

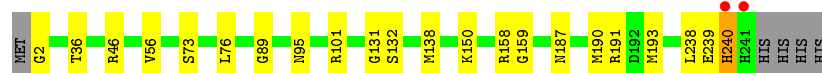
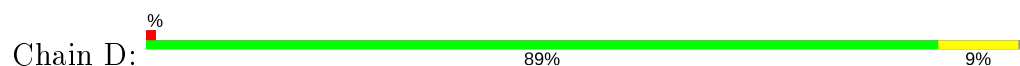
3 Residue-property plots [i](#)

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.

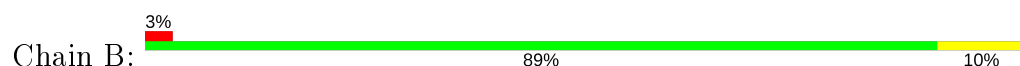
- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein



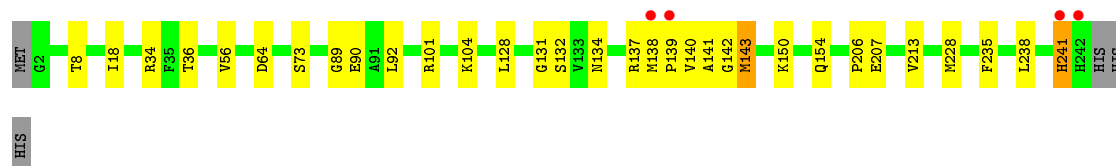
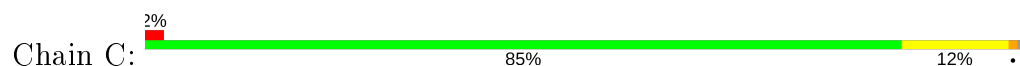
- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein



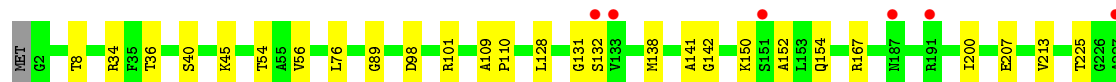
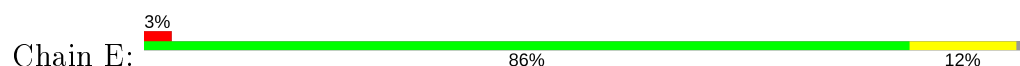
- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein

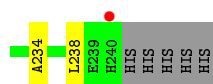


- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein

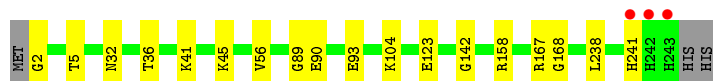
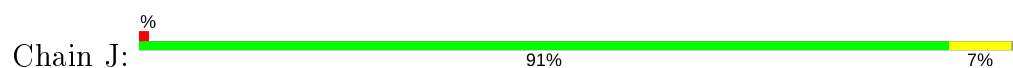


- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein

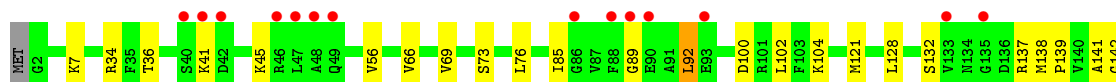
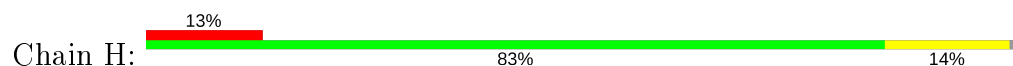




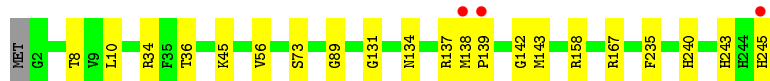
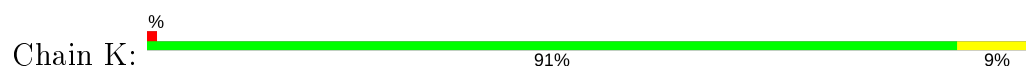
- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein



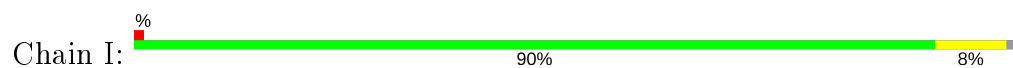
- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein



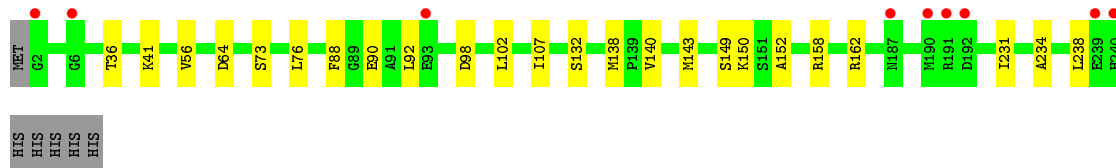
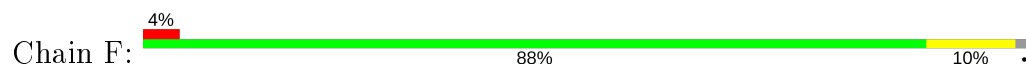
- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein



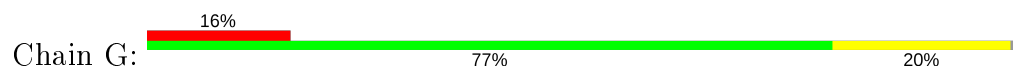
- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein

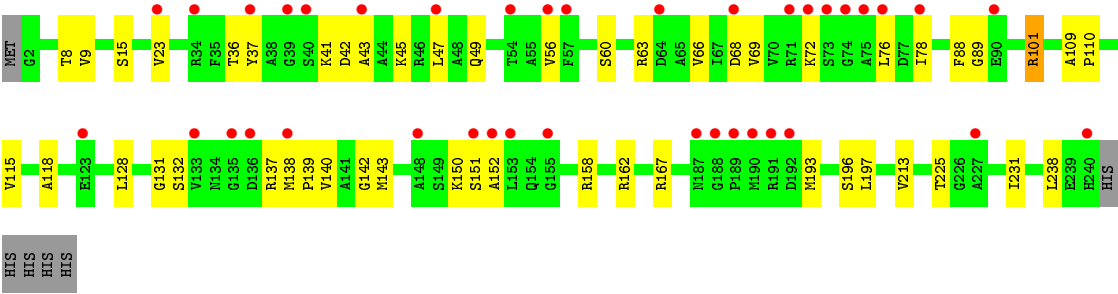


- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein



- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein





- Molecule 1: Cyclic-di-GMP-binding biofilm dispersal mediator protein

Chain L:

91%

7%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	65.83Å 72.28Å 172.32Å 99.22° 86.78° 107.70°	Depositor
Resolution (Å)	43.73 – 1.70 43.73 – 1.70	Depositor EDS
% Data completeness (in resolution range)	92.3 (43.73-1.70) 92.3 (43.73-1.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.99 (at 1.70Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.151 , 0.179 0.151 , 0.178	Depositor DCC
R_{free} test set	2009 reflections (0.66%)	wwPDB-VP
Wilson B-factor (Å ²)	19.6	Xtriage
Anisotropy	0.199	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 48.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.012 for -h,-k,k+l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	25111	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: GOL, EDO, PG4, PG0, NDP, PEG

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	0.36	0/1862	0.55	0/2516
1	B	0.36	0/1918	0.53	0/2590
1	C	0.44	0/1936	0.63	0/2614
1	D	0.40	0/1847	0.57	0/2496
1	E	0.37	0/1841	0.56	0/2487
1	F	0.36	0/1840	0.57	0/2486
1	G	0.31	0/1847	0.51	0/2498
1	H	0.33	0/1814	0.52	0/2453
1	I	0.44	0/1867	0.60	0/2526
1	J	0.46	0/1911	0.60	0/2583
1	K	0.49	0/1990	0.66	0/2688
1	L	0.50	0/1869	0.65	0/2525
All	All	0.41	0/22542	0.58	0/30462

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	1784	0	1796	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1823	0	1815	22	0
1	C	1842	0	1835	40	0
1	D	1776	0	1768	19	0
1	E	1772	0	1778	24	0
1	F	1768	0	1780	25	0
1	G	1769	0	1769	47	0
1	H	1757	0	1755	31	0
1	I	1786	0	1778	20	0
1	J	1815	0	1801	17	0
1	K	1882	0	1865	22	0
1	L	1777	0	1765	19	0
2	A	8	12	12	0	0
3	A	48	26	26	0	0
3	B	48	26	26	0	0
3	C	48	26	26	1	0
3	D	48	26	26	1	0
3	E	48	26	26	1	0
3	F	48	26	26	0	0
3	G	48	26	26	1	0
3	H	48	26	26	0	0
3	I	48	26	26	1	0
3	J	48	26	26	0	0
3	K	48	26	26	1	0
3	L	48	26	26	1	0
4	B	4	6	6	2	0
4	C	8	12	12	2	0
4	E	4	6	6	0	0
4	G	4	6	6	0	0
4	I	4	6	6	0	0
5	C	6	8	8	0	0
5	E	6	8	8	0	0
6	K	7	10	10	0	0
7	F	13	18	18	0	0
8	A	155	0	0	2	1
8	B	185	0	0	2	0
8	C	245	0	0	11	2
8	D	215	0	0	6	2
8	E	185	0	0	3	0
8	F	189	0	0	5	0
8	G	94	0	0	7	0
8	H	118	0	0	3	1
8	I	265	0	0	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	J	274	0	0	6	1
8	K	297	0	0	5	4
8	L	294	0	0	7	1
All	All	24707	404	21909	270	6

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 270 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:E:45:LYS:NZ	8:E:401:HOH:O	1.82	1.13
1:D:95:ASN:ND2	8:D:401:HOH:O	1.96	0.97
1:A:167:ARG:NH1	8:A:401:HOH:O	1.98	0.94
1:I:81:VAL:HB	1:I:129[B]:ILE:HD13	1.51	0.90
4:C:303:EDO:O2	8:C:401:HOH:O	1.91	0.88

The worst 5 of 6 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)
8:C:606:HOH:O	8:K:542:HOH:O[1_445]	1.90	0.30
8:K:634:HOH:O	8:L:630:HOH:O[1_655]	1.90	0.30
8:D:402:HOH:O	8:C:587:HOH:O[1_455]	1.99	0.21
8:D:587:HOH:O	8:K:575:HOH:O[1_445]	2.11	0.09
8:A:546:HOH:O	8:H:495:HOH:O[1_455]	2.12	0.08

5.3 Torsion angles ⓘ

5.3.1 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	247/245 (101%)	241 (98%)	6 (2%)	0	100	100
1	B	252/245 (103%)	247 (98%)	5 (2%)	0	100	100
1	C	256/245 (104%)	245 (96%)	9 (4%)	2 (1%)	19	6
1	D	245/245 (100%)	237 (97%)	7 (3%)	1 (0%)	34	18
1	E	244/245 (100%)	238 (98%)	6 (2%)	0	100	100
1	F	245/245 (100%)	239 (98%)	6 (2%)	0	100	100
1	G	246/245 (100%)	238 (97%)	8 (3%)	0	100	100
1	H	242/245 (99%)	233 (96%)	9 (4%)	0	100	100
1	I	248/245 (101%)	243 (98%)	5 (2%)	0	100	100
1	J	252/245 (103%)	244 (97%)	8 (3%)	0	100	100
1	K	261/245 (106%)	252 (97%)	9 (3%)	0	100	100
1	L	248/245 (101%)	240 (97%)	8 (3%)	0	100	100
All	All	2986/2940 (102%)	2897 (97%)	86 (3%)	3 (0%)	51	33

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	143[A]	MET
1	C	143[B]	MET
1	D	240	HIS

5.3.2 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	181/177 (102%)	176 (97%)	5 (3%)	43	25
1	B	186/177 (105%)	186 (100%)	0	100	100
1	C	188/177 (106%)	187 (100%)	1 (0%)	88	83
1	D	179/177 (101%)	179 (100%)	0	100	100
1	E	178/177 (101%)	178 (100%)	0	100	100
1	F	179/177 (101%)	179 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	180/177 (102%)	178 (99%)	2 (1%)	73	63
1	H	176/177 (99%)	174 (99%)	2 (1%)	73	63
1	I	182/177 (103%)	182 (100%)	0	100	100
1	J	186/177 (105%)	185 (100%)	1 (0%)	88	83
1	K	193/177 (109%)	193 (100%)	0	100	100
1	L	182/177 (103%)	180 (99%)	2 (1%)	73	63
All	All	2190/2124 (103%)	2177 (99%)	13 (1%)	91	80

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

Mol	Chain	Res	Type
1	C	241	HIS
1	J	241	HIS
1	G	101[B]	ARG
1	A	163	ASP
1	G	101[A]	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

23 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	NDP	K	302	-	45,52,52	1.00	2 (4%)	53,80,80	1.35	8 (15%)
4	EDO	E	302	-	3,3,3	0.52	0	2,2,2	0.22	0
3	NDP	L	301	-	45,52,52	0.96	2 (4%)	53,80,80	1.28	6 (11%)
2	PG0	A	301	-	7,7,7	0.49	0	6,6,6	0.25	0
3	NDP	A	302	-	45,52,52	0.96	3 (6%)	53,80,80	1.25	5 (9%)
5	GOL	C	301	-	5,5,5	0.36	0	5,5,5	0.19	0
3	NDP	D	301	-	45,52,52	0.97	3 (6%)	53,80,80	1.17	5 (9%)
7	PG4	F	301	-	12,12,12	0.50	0	11,11,11	0.32	0
3	NDP	I	302	-	45,52,52	1.04	4 (8%)	53,80,80	1.29	7 (13%)
3	NDP	E	303	-	45,52,52	1.00	3 (6%)	53,80,80	1.26	5 (9%)
3	NDP	J	301	-	45,52,52	1.10	4 (8%)	53,80,80	1.23	6 (11%)
4	EDO	I	301	-	3,3,3	0.48	0	2,2,2	0.26	0
3	NDP	C	304	-	45,52,52	0.94	1 (2%)	53,80,80	1.32	8 (15%)
4	EDO	C	302	-	3,3,3	0.48	0	2,2,2	0.27	0
4	EDO	C	303	-	3,3,3	0.38	0	2,2,2	0.37	0
4	EDO	G	301	-	3,3,3	0.45	0	2,2,2	0.40	0
6	PEG	K	301	-	6,6,6	0.44	0	5,5,5	0.29	0
3	NDP	H	301	-	45,52,52	0.95	3 (6%)	53,80,80	1.20	5 (9%)
3	NDP	F	302	-	45,52,52	0.92	2 (4%)	53,80,80	1.26	6 (11%)
4	EDO	B	301	-	3,3,3	0.42	0	2,2,2	0.39	0
3	NDP	B	302	-	45,52,52	1.03	3 (6%)	53,80,80	1.23	6 (11%)
3	NDP	G	302	-	45,52,52	0.95	2 (4%)	53,80,80	1.18	4 (7%)
5	GOL	E	301	-	5,5,5	0.32	0	5,5,5	0.30	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	NDP	K	302	-	-	5/30/77/77	0/5/5/5
4	EDO	E	302	-	-	0/1/1/1	-
3	NDP	L	301	-	-	6/30/77/77	0/5/5/5
2	PG0	A	301	-	-	3/5/5/5	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NDP	A	302	-	-	3/30/77/77	0/5/5/5
5	GOL	C	301	-	-	0/4/4/4	-
3	NDP	D	301	-	-	4/30/77/77	0/5/5/5
7	PG4	F	301	-	-	3/10/10/10	-
3	NDP	I	302	-	-	4/30/77/77	0/5/5/5
3	NDP	E	303	-	-	5/30/77/77	0/5/5/5
3	NDP	J	301	-	-	6/30/77/77	0/5/5/5
4	EDO	I	301	-	-	0/1/1/1	-
3	NDP	C	304	-	-	4/30/77/77	0/5/5/5
4	EDO	C	302	-	-	0/1/1/1	-
4	EDO	C	303	-	-	0/1/1/1	-
4	EDO	G	301	-	-	0/1/1/1	-
6	PEG	K	301	-	-	2/4/4/4	-
3	NDP	H	301	-	-	6/30/77/77	0/5/5/5
3	NDP	F	302	-	-	6/30/77/77	0/5/5/5
4	EDO	B	301	-	-	0/1/1/1	-
3	NDP	B	302	-	-	4/30/77/77	0/5/5/5
3	NDP	G	302	-	-	3/30/77/77	0/5/5/5
5	GOL	E	301	-	-	2/4/4/4	-

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	303	NDP	C6N-C5N	3.36	1.39	1.33
3	B	302	NDP	C6N-C5N	3.14	1.38	1.33
3	I	302	NDP	C6N-C5N	3.07	1.38	1.33
3	C	304	NDP	C6N-C5N	2.97	1.38	1.33
3	G	302	NDP	C6N-C5N	2.91	1.38	1.33

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	301	NDP	N3A-C2A-N1A	-4.35	121.88	128.68
3	D	301	NDP	N3A-C2A-N1A	-4.16	122.18	128.68
3	F	302	NDP	N3A-C2A-N1A	-4.11	122.25	128.68
3	K	302	NDP	N3A-C2A-N1A	-4.06	122.33	128.68
3	E	303	NDP	N3A-C2A-N1A	-3.86	122.64	128.68

There are no chirality outliers.

5 of 66 torsion outliers are listed below:

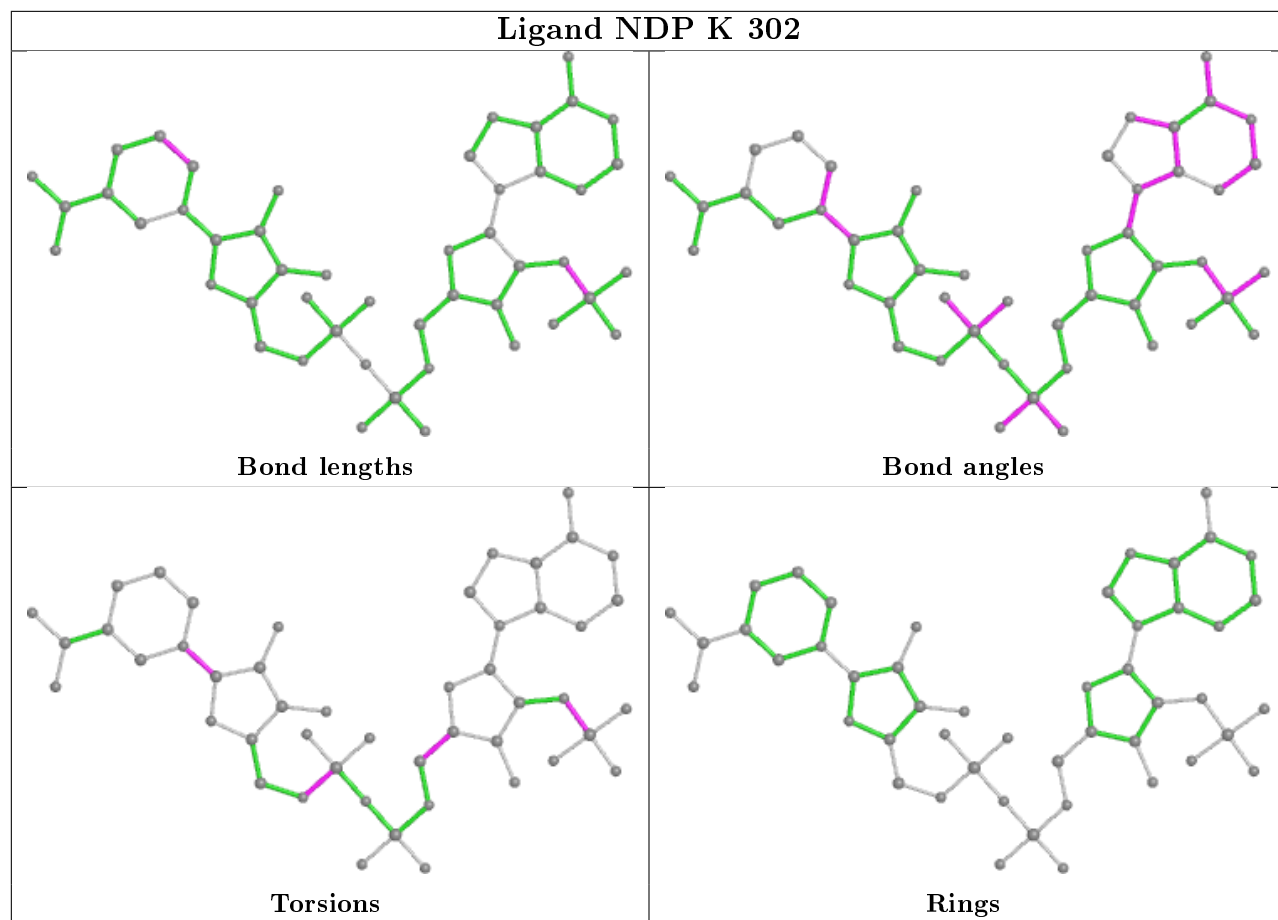
Mol	Chain	Res	Type	Atoms
3	K	302	NDP	C5D-O5D-PN-O2N
3	L	301	NDP	C5D-O5D-PN-O2N
3	E	303	NDP	C5D-O5D-PN-O2N
3	J	301	NDP	C5D-O5D-PN-O2N
3	C	304	NDP	C2B-O2B-P2B-O2X

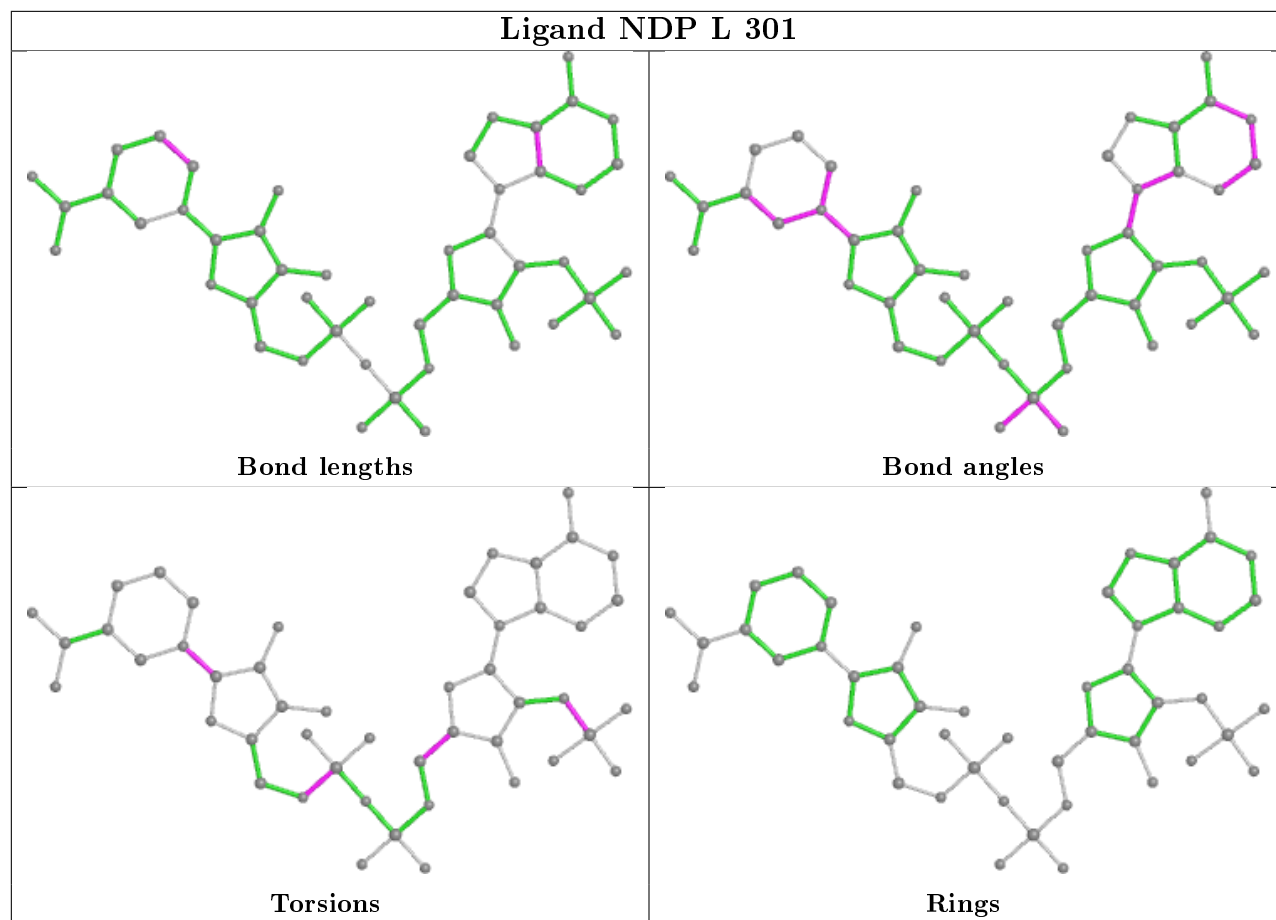
There are no ring outliers.

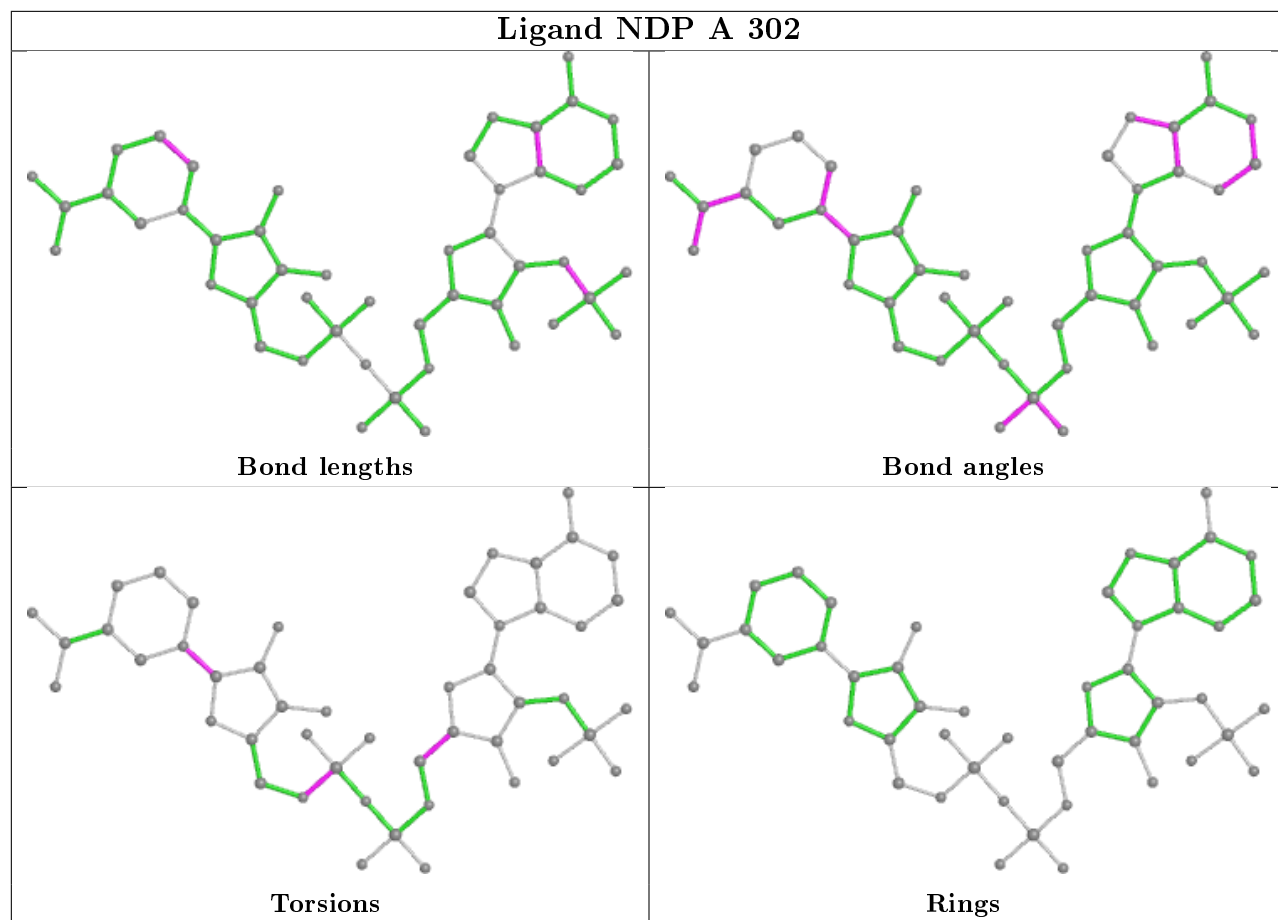
9 monomers are involved in 11 short contacts:

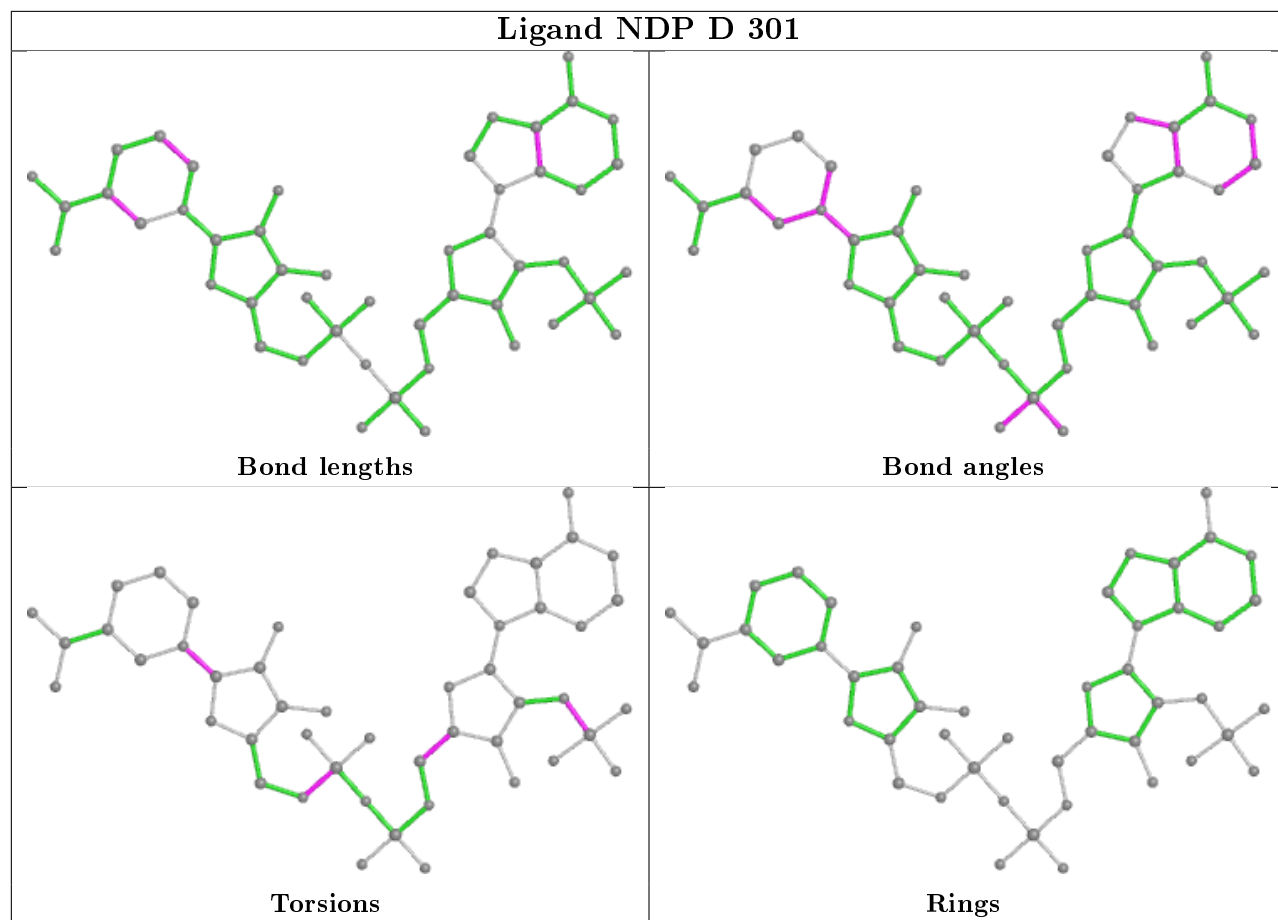
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	K	302	NDP	1	0
3	L	301	NDP	1	0
3	D	301	NDP	1	0
3	I	302	NDP	1	0
3	E	303	NDP	1	0
3	C	304	NDP	1	0
4	C	303	EDO	2	0
4	B	301	EDO	2	0
3	G	302	NDP	1	0

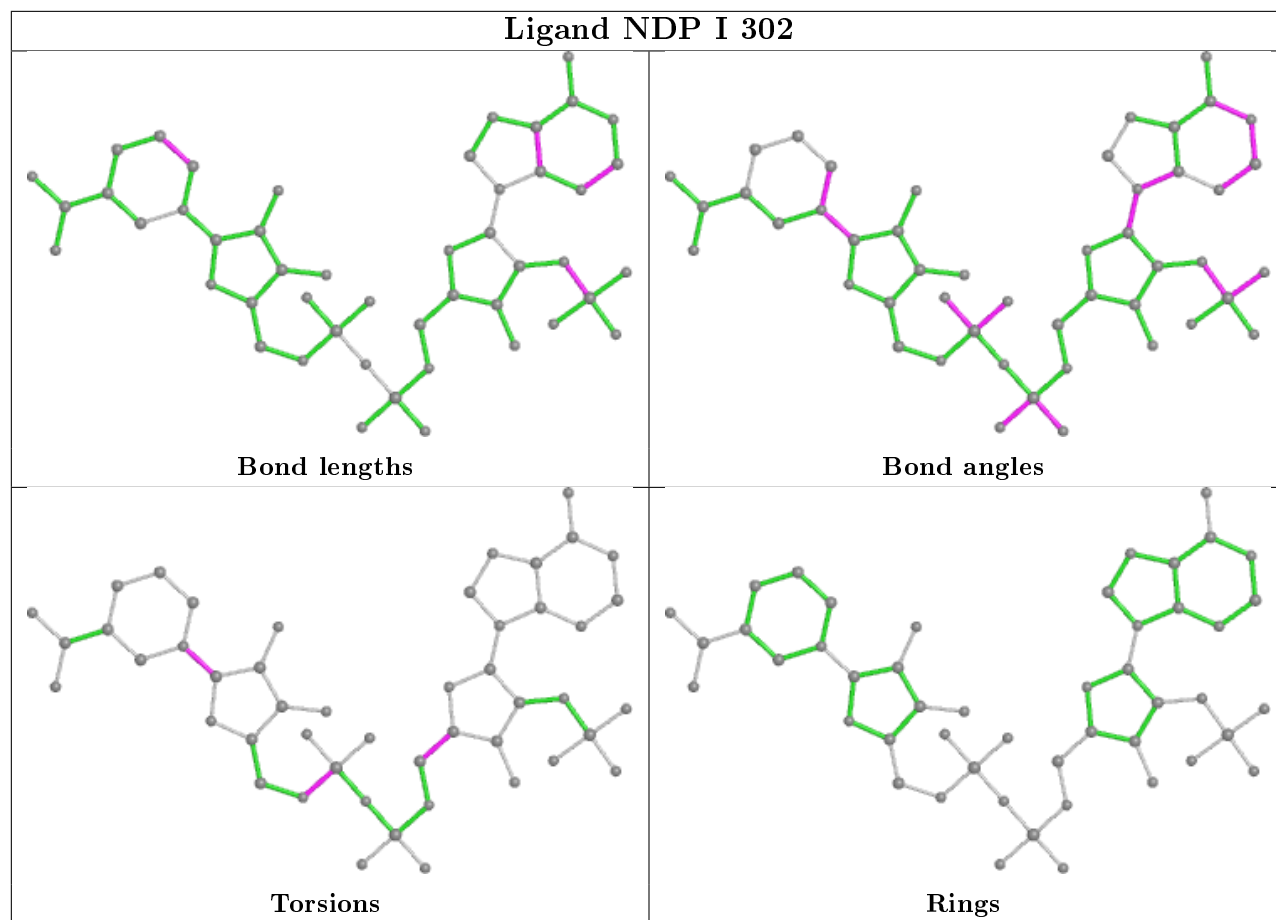
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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.

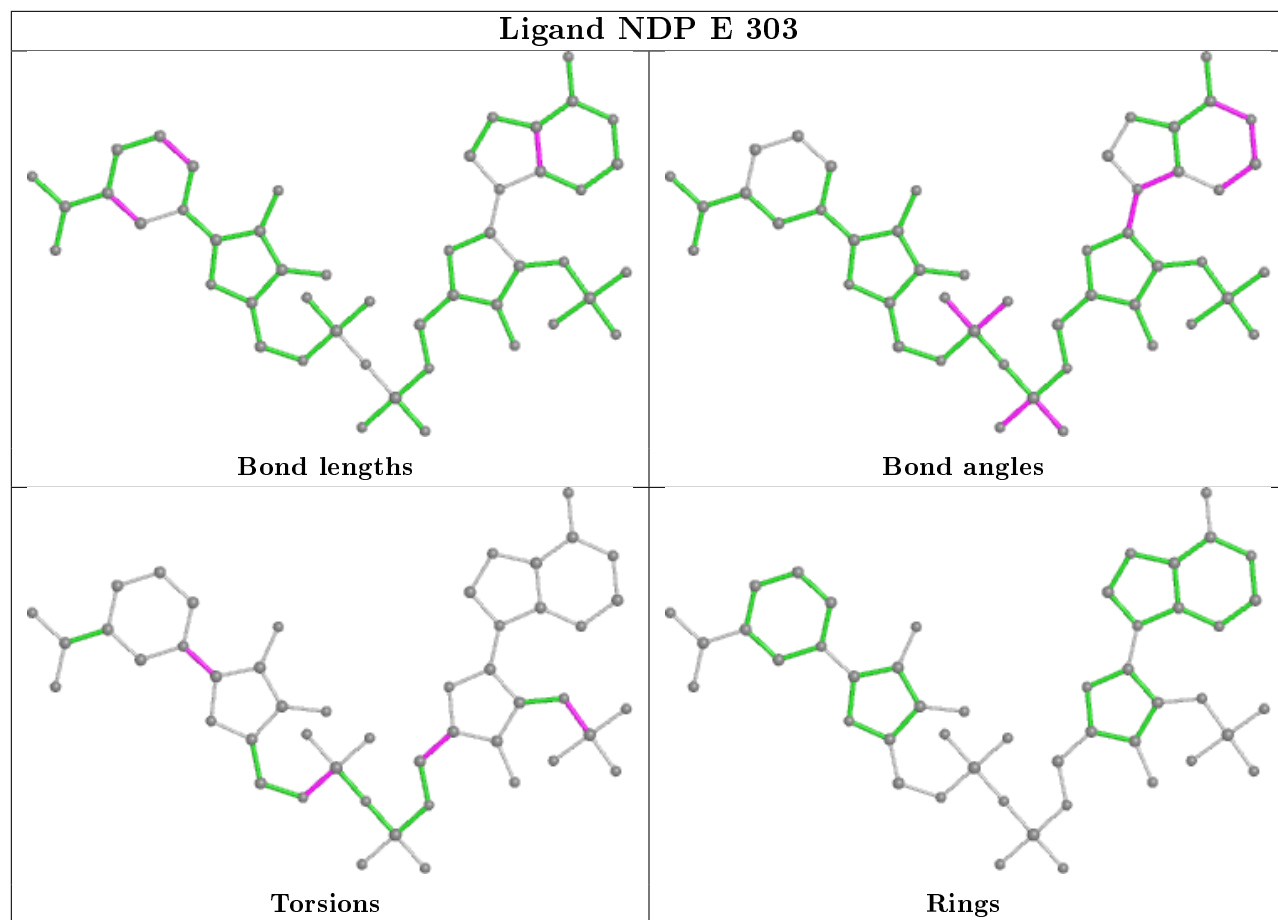


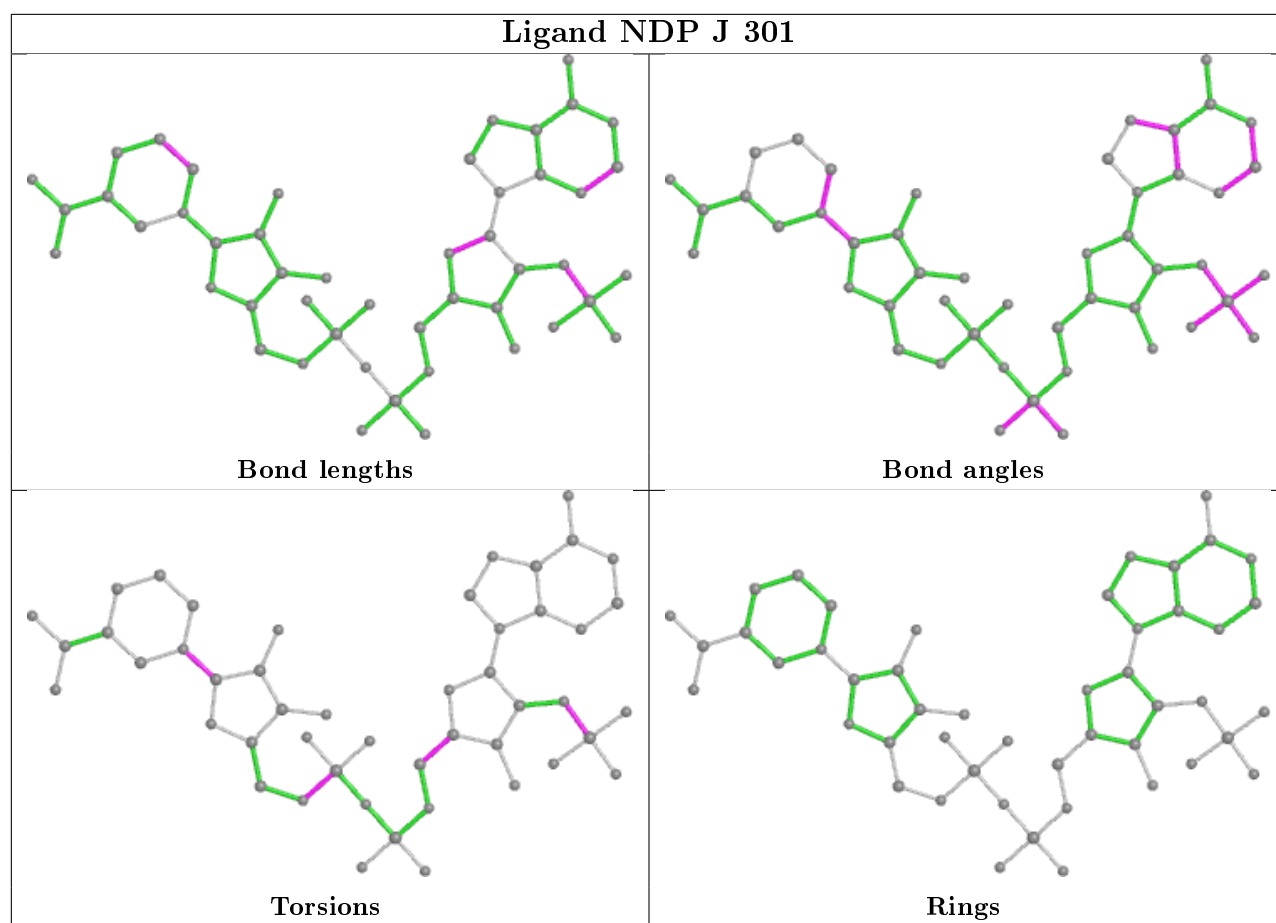


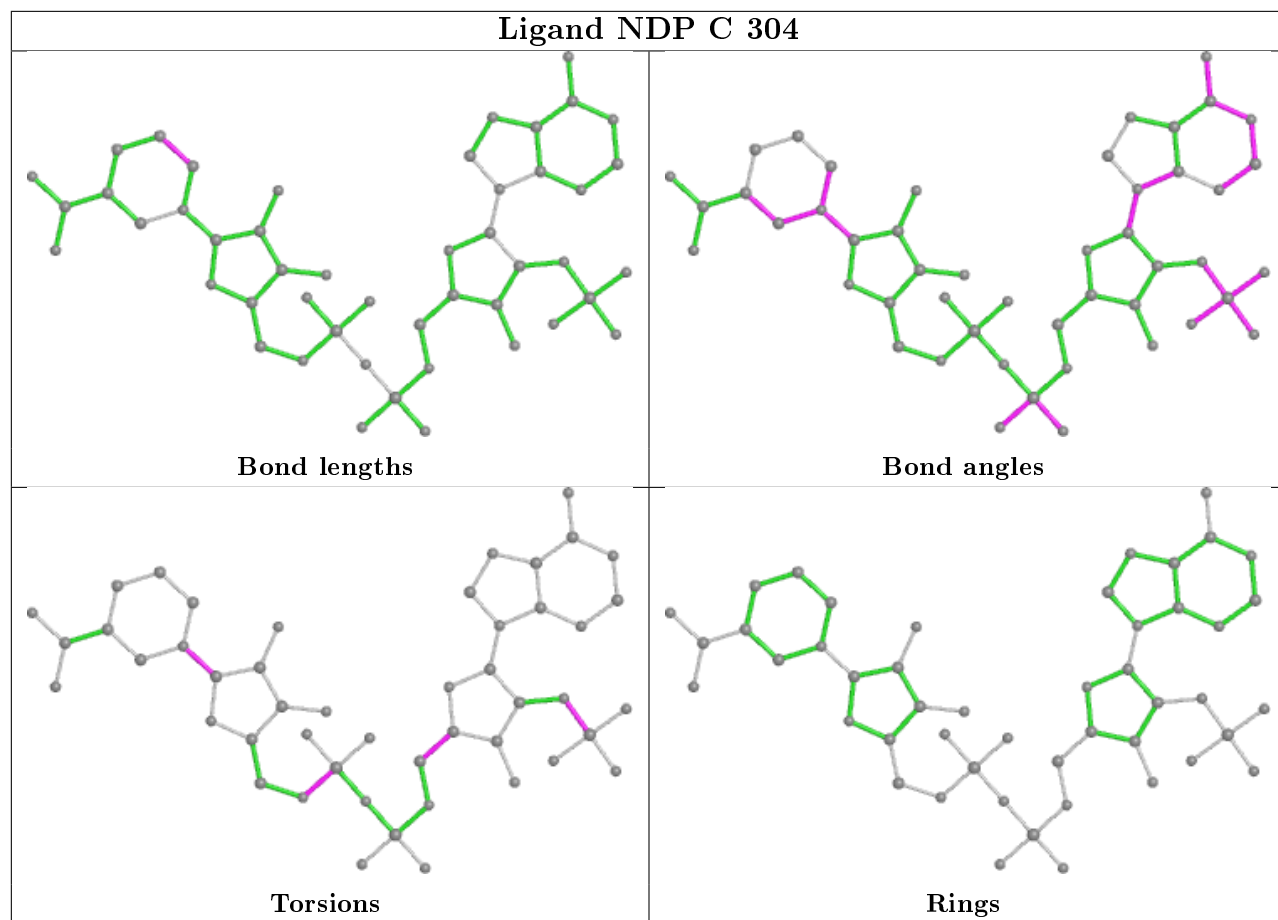


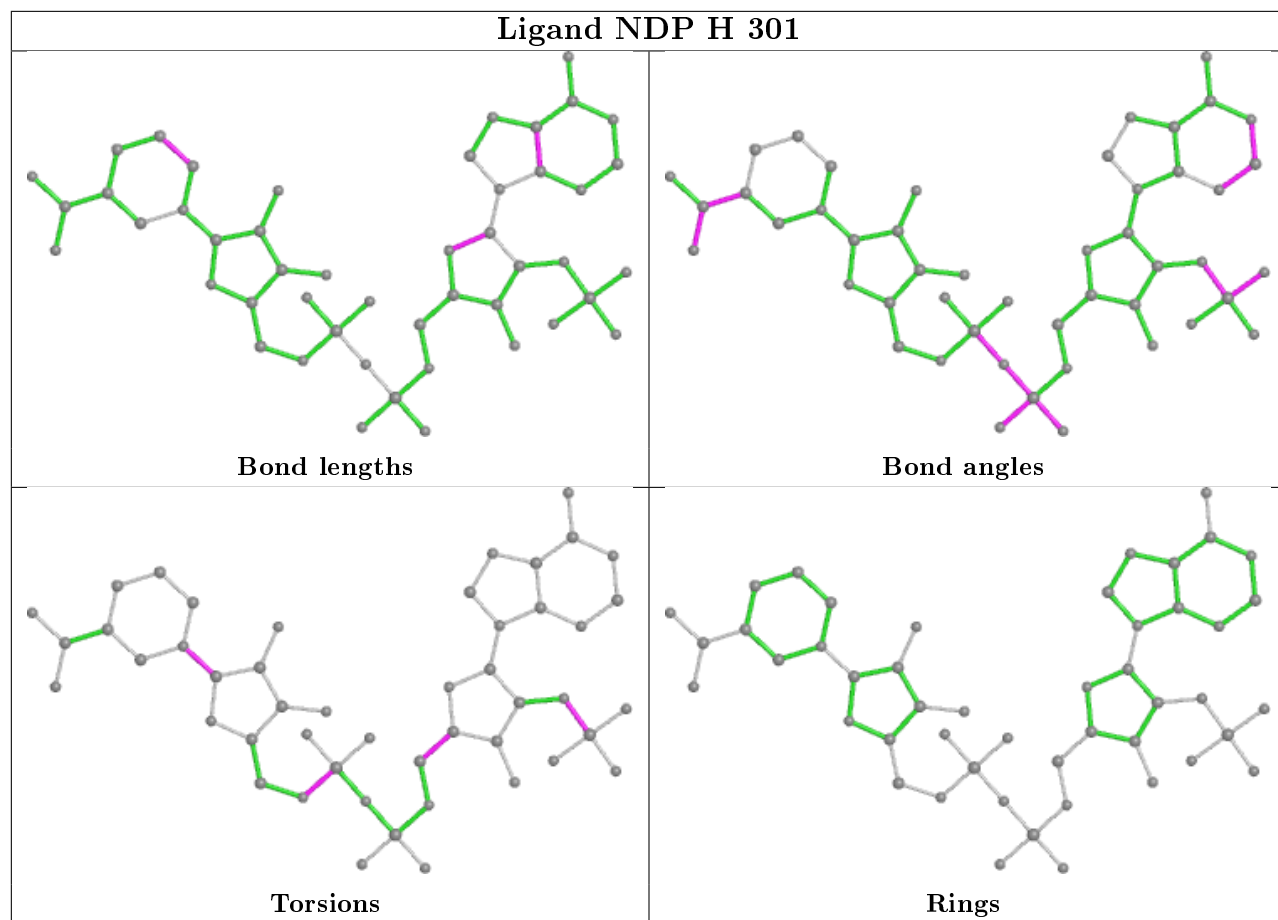


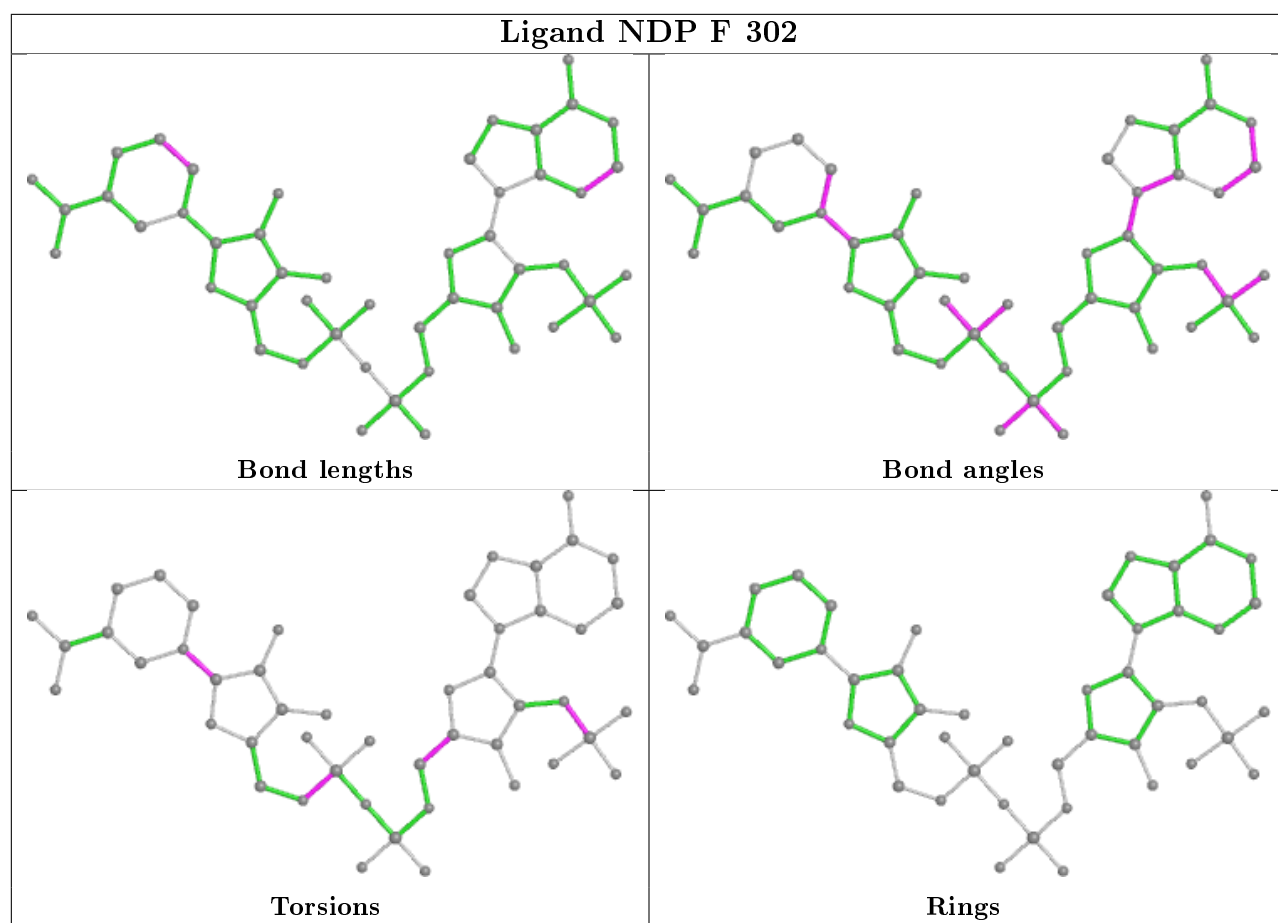


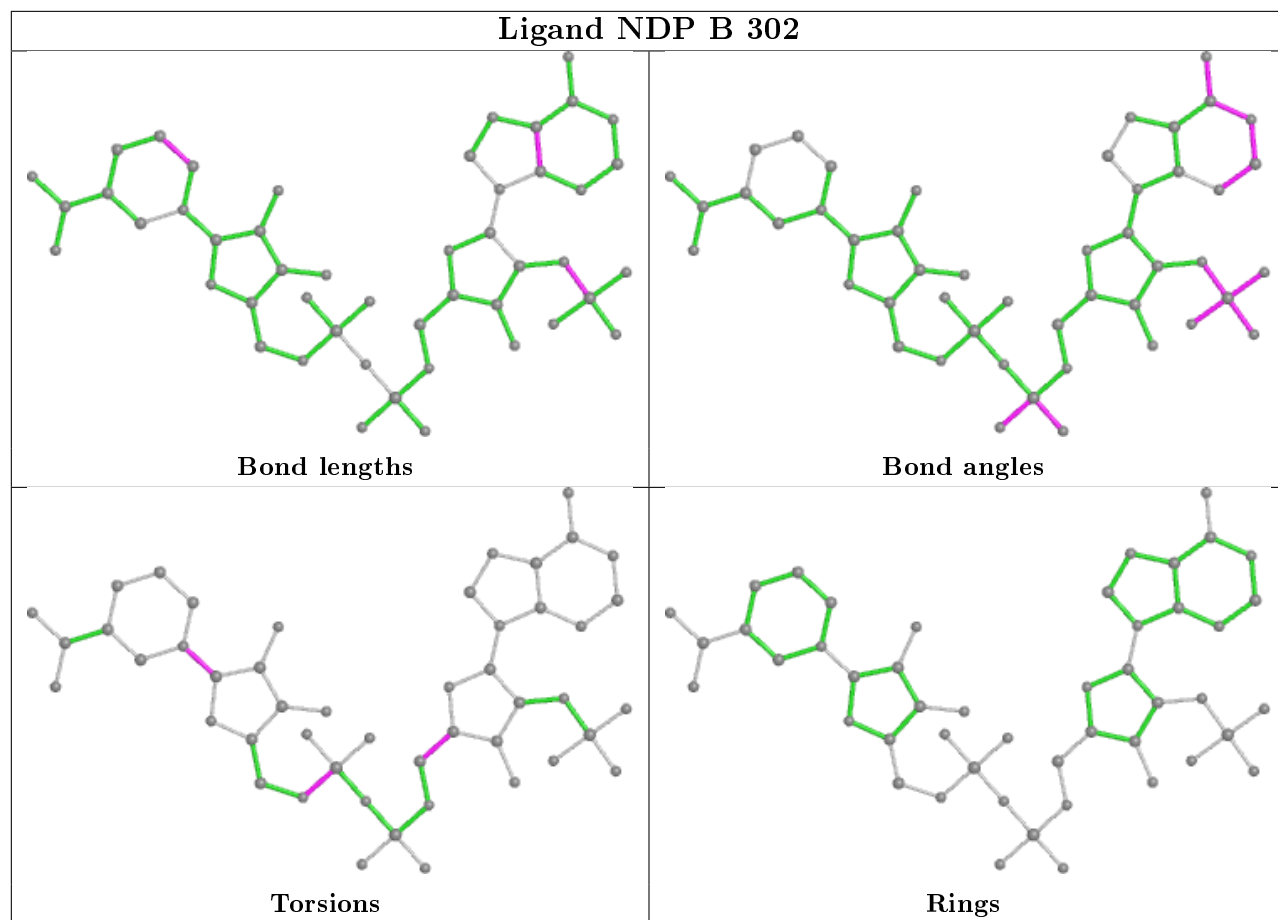


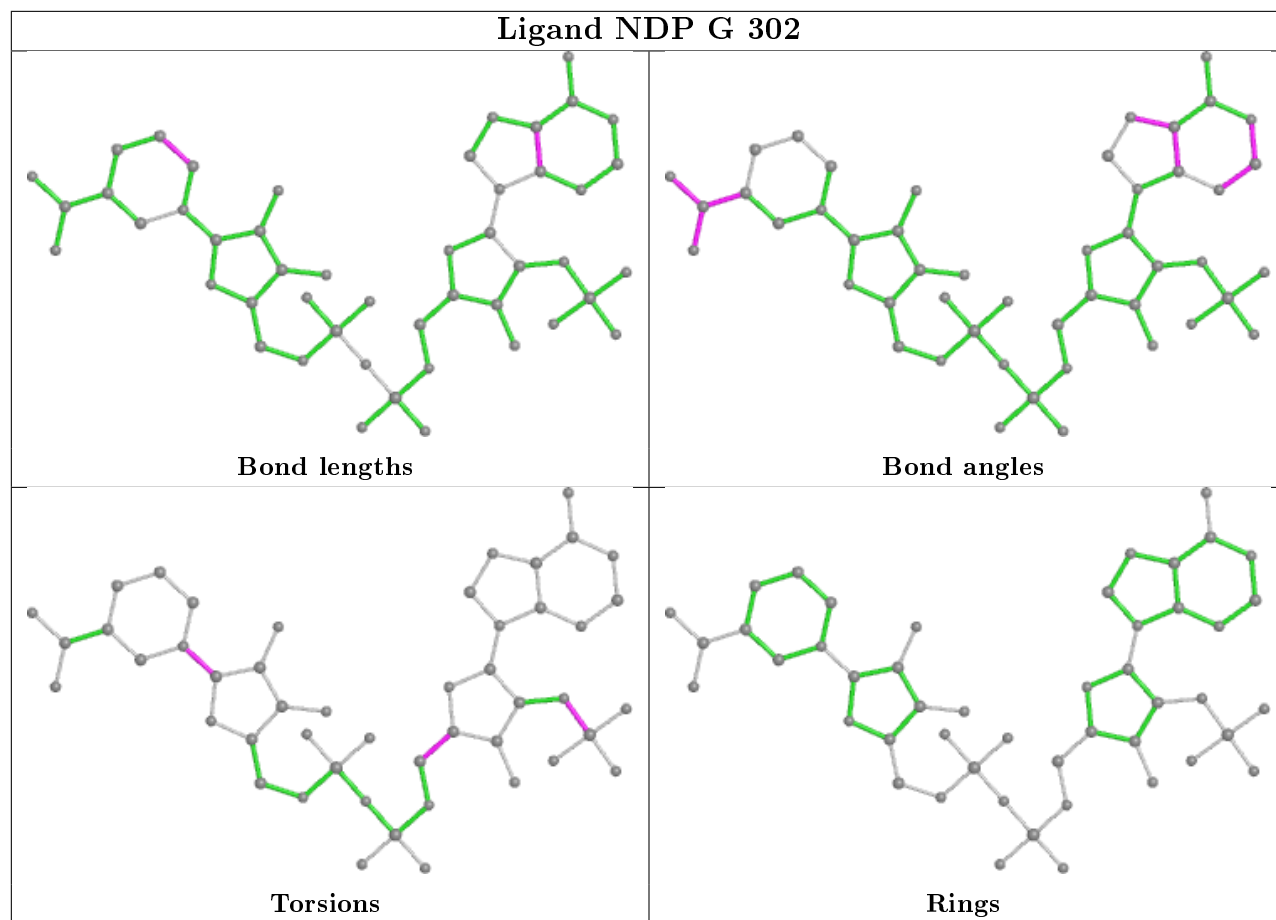












5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	239/245 (97%)	0.08	8 (3%) 46 51	21, 33, 50, 59	0
1	B	243/245 (99%)	-0.14	7 (2%) 51 56	18, 30, 46, 66	0
1	C	241/245 (98%)	-0.37	4 (1%) 70 74	14, 20, 34, 64	0
1	D	240/245 (97%)	-0.24	2 (0%) 86 88	13, 24, 37, 63	0
1	E	239/245 (97%)	0.00	7 (2%) 51 56	21, 29, 43, 63	0
1	F	239/245 (97%)	0.02	9 (3%) 40 45	21, 32, 49, 79	0
1	G	239/245 (97%)	0.88	38 (15%) 1 2	36, 54, 68, 73	0
1	H	239/245 (97%)	0.55	32 (13%) 3 3	26, 42, 64, 77	0
1	I	241/245 (98%)	-0.33	2 (0%) 86 88	13, 20, 34, 60	0
1	J	242/245 (98%)	-0.27	3 (1%) 79 82	13, 20, 34, 62	0
1	K	244/245 (99%)	-0.24	3 (1%) 79 82	10, 17, 30, 64	0
1	L	239/245 (97%)	-0.41	0 100 100	11, 17, 29, 45	0
All	All	2885/2940 (98%)	-0.04	115 (3%) 38 42	10, 26, 56, 79	0

The worst 5 of 115 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	241	HIS	6.0
1	D	241	HIS	5.8
1	H	89	GLY	5.8
1	F	240	HIS	5.6
1	J	242	HIS	5.2

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 carbohydrates in this entry.

6.4 Ligands ⓘ

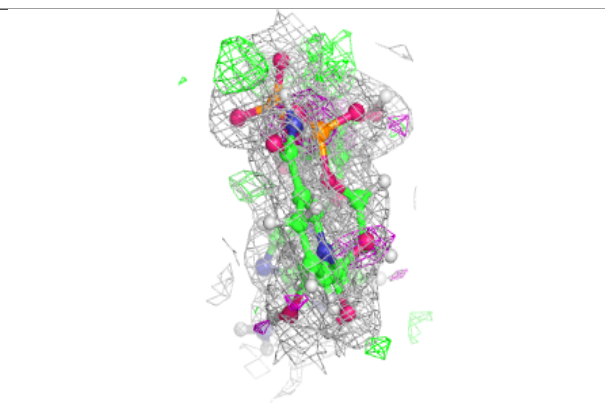
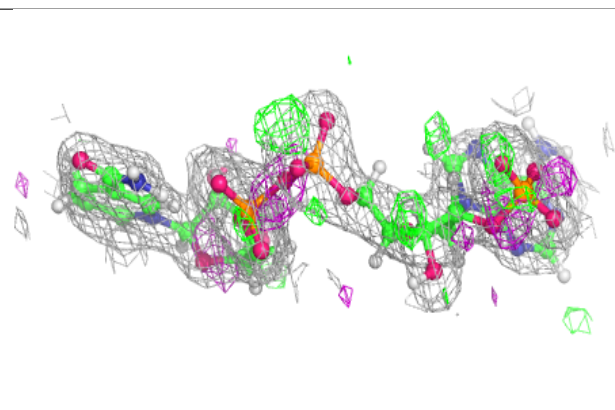
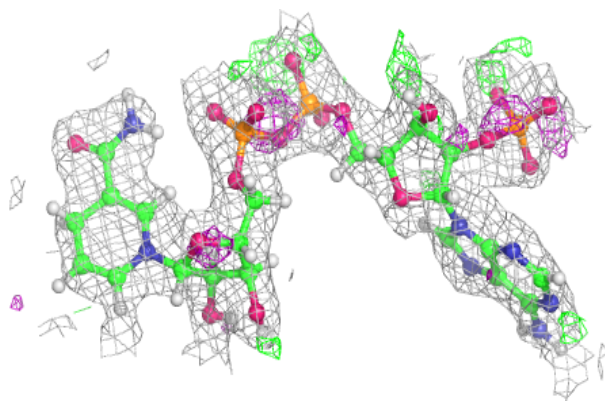
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	EDO	E	302	4/4	0.34	0.25	67,80,82,82	0
4	EDO	C	302	4/4	0.40	0.30	79,94,96,97	0
7	PG4	F	301	13/13	0.51	0.44	99,119,124,124	0
5	GOL	E	301	6/6	0.70	0.19	61,73,80,81	0
4	EDO	G	301	4/4	0.74	0.28	72,87,88,90	0
2	PG0	A	301	8/8	0.74	0.15	73,88,91,93	0
4	EDO	I	301	4/4	0.79	0.25	55,66,71,71	10
4	EDO	C	303	4/4	0.79	0.33	34,41,47,49	10
5	GOL	C	301	6/6	0.80	0.18	41,49,56,59	14
6	PEG	K	301	7/7	0.86	0.10	34,49,65,65	17
3	NDP	G	302	48/48	0.89	0.11	29,48,68,71	0
4	EDO	B	301	4/4	0.93	0.30	59,71,74,75	10
3	NDP	H	301	48/48	0.94	0.13	30,39,48,49	0
3	NDP	A	302	48/48	0.96	0.06	22,31,38,41	0
3	NDP	B	302	48/48	0.97	0.07	22,29,36,39	0
3	NDP	F	302	48/48	0.97	0.09	18,26,34,39	0
3	NDP	E	303	48/48	0.97	0.06	19,24,32,37	0
3	NDP	I	302	48/48	0.98	0.05	15,20,26,31	0
3	NDP	D	301	48/48	0.98	0.05	18,24,30,34	0
3	NDP	J	301	48/48	0.98	0.05	16,20,24,30	0
3	NDP	K	302	48/48	0.98	0.06	11,15,19,22	0
3	NDP	C	304	48/48	0.98	0.06	14,19,25,26	0
3	NDP	L	301	48/48	0.99	0.05	11,17,21,23	0

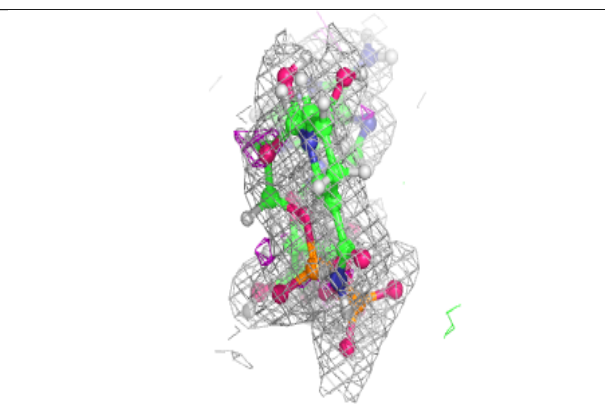
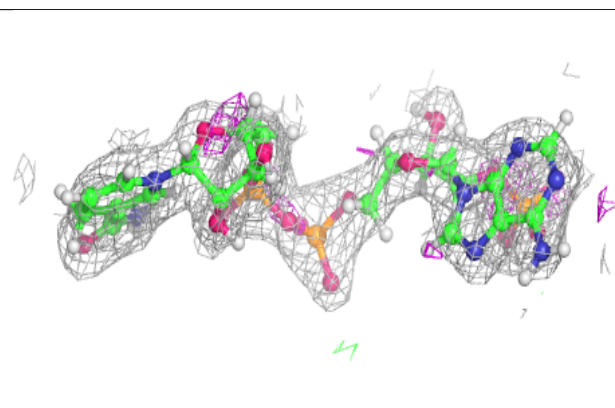
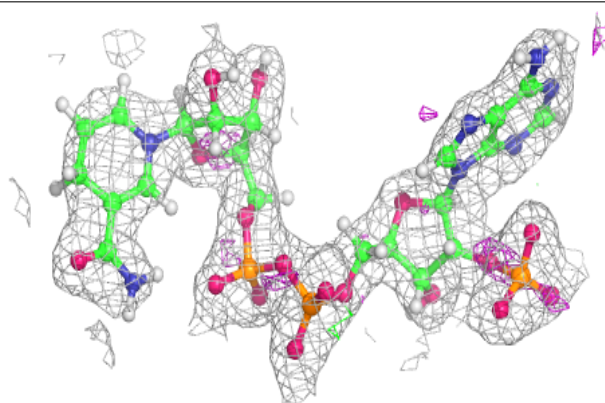
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around NDP G 302:

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

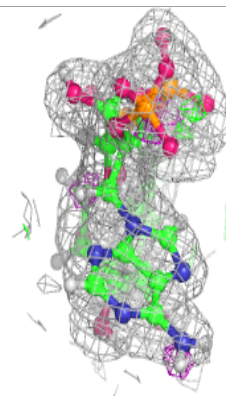
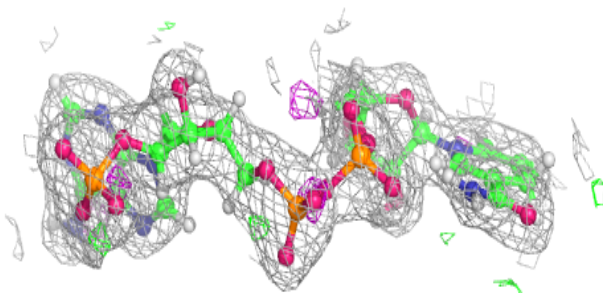
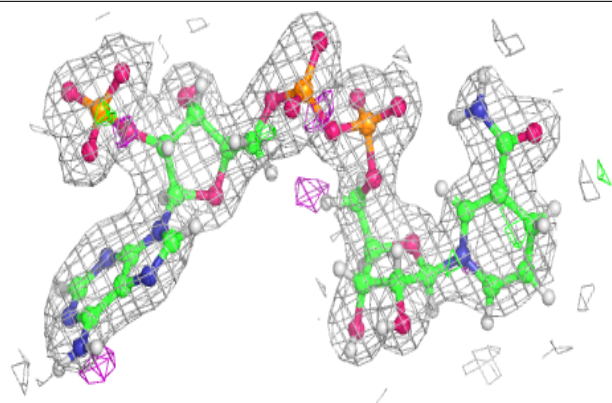
**Electron density around NDP H 301:**

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

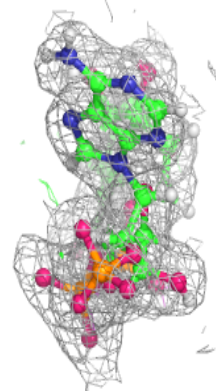
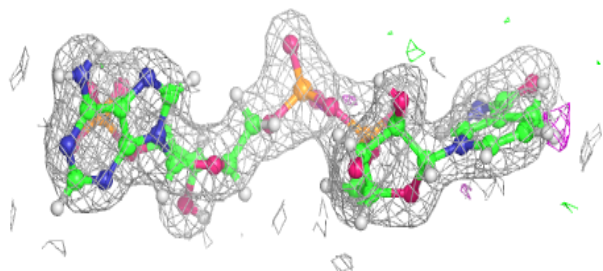
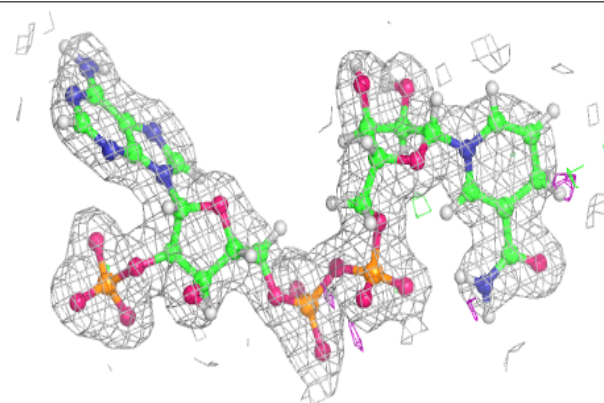


Electron density around NDP A 302:

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

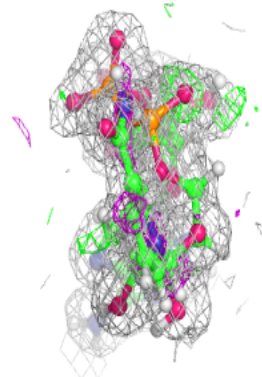
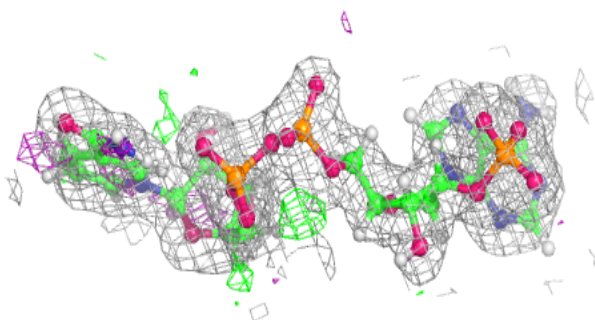
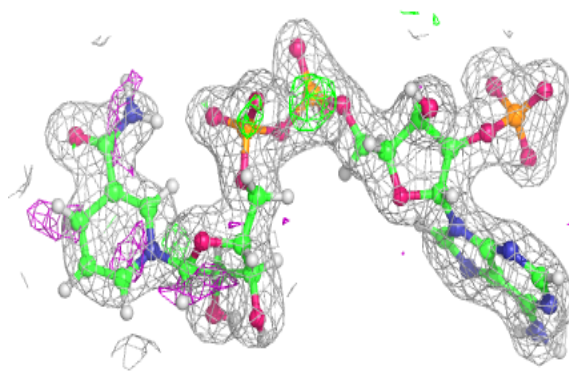
**Electron density around NDP B 302:**

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

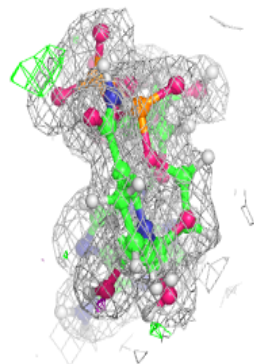
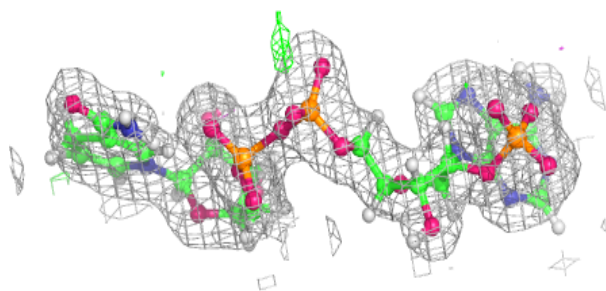
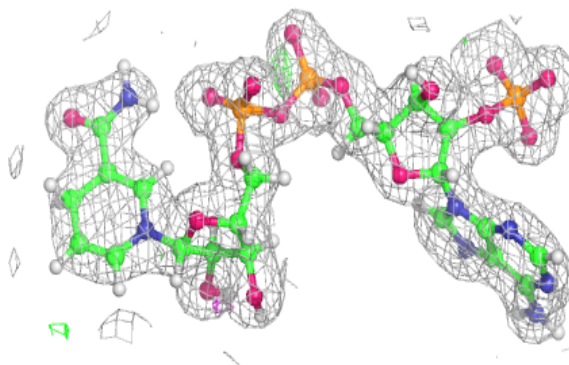


Electron density around NDP F 302:

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

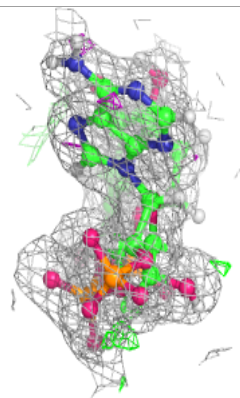
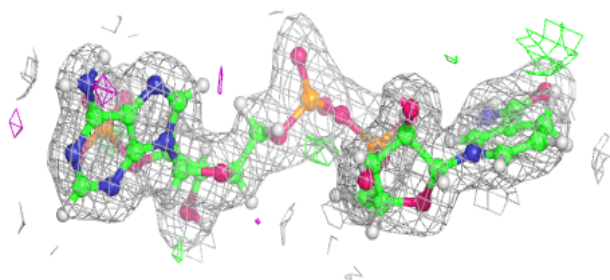
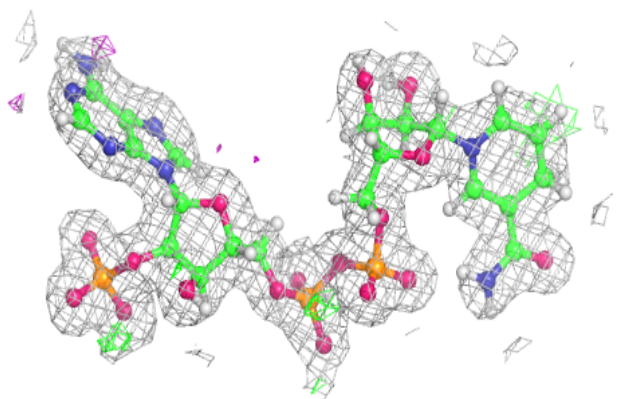
**Electron density around NDP E 303:**

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

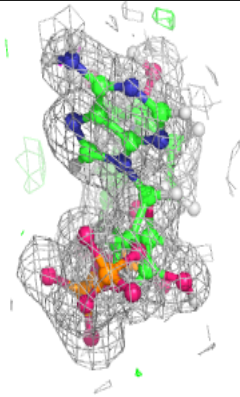
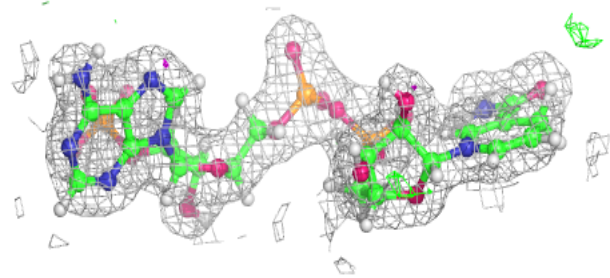
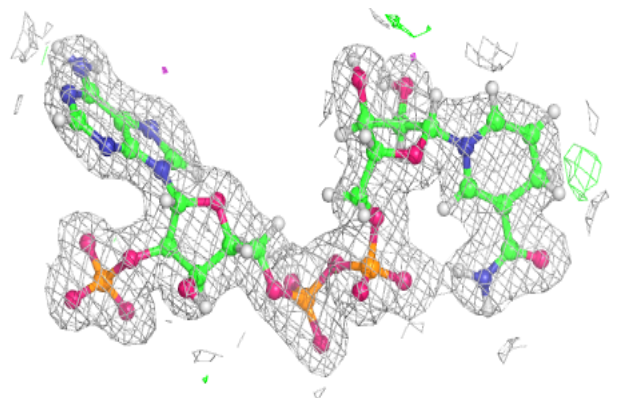


Electron density around NDP I 302:

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

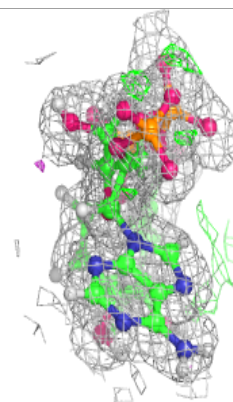
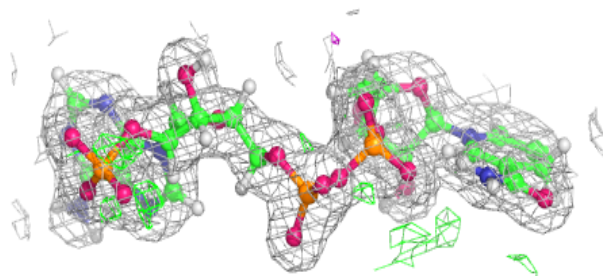
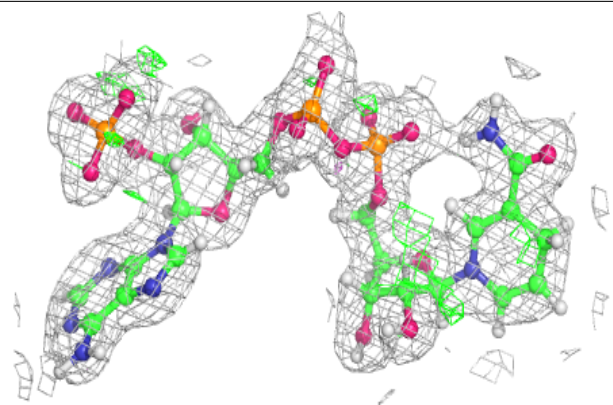
**Electron density around NDP D 301:**

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

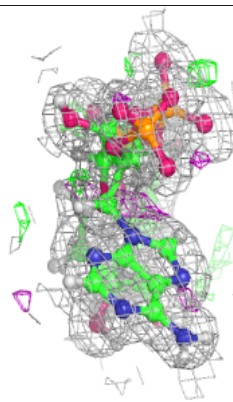
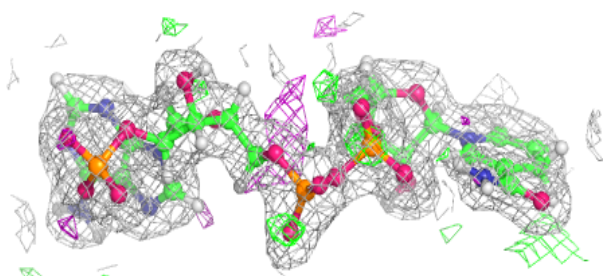
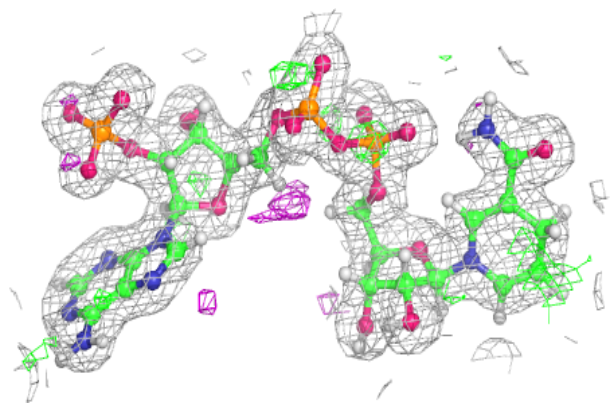


Electron density around NDP J 301:

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

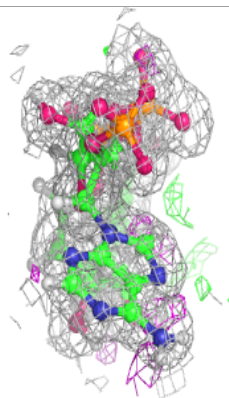
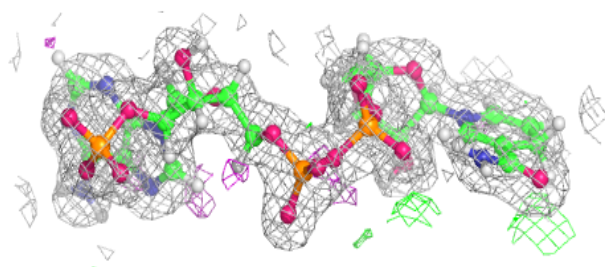
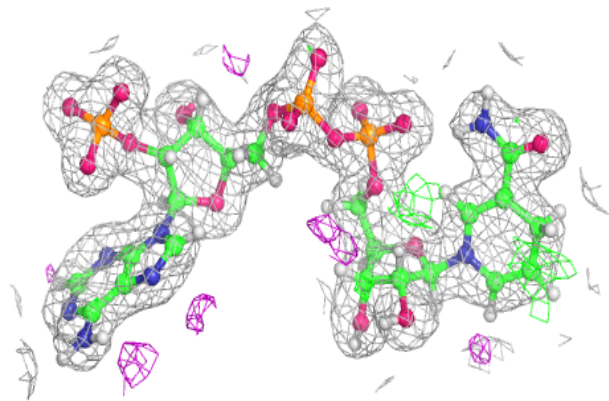
**Electron density around NDP K 302:**

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

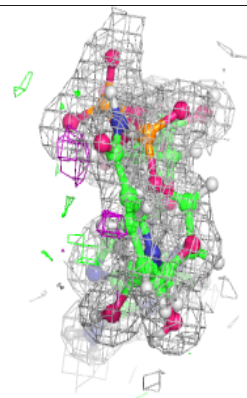
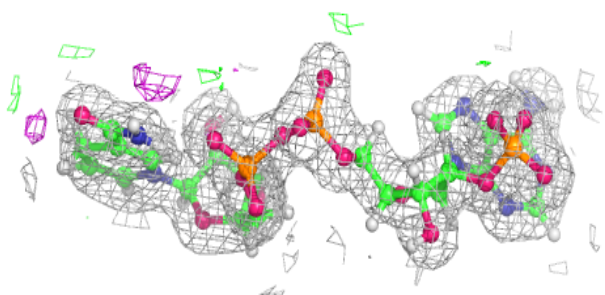
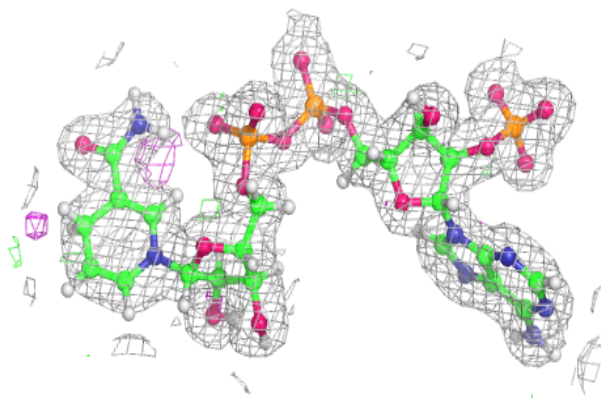


Electron density around NDP C 304:

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

**Electron density around NDP L 301:**

$2mF_o-DF_c$ (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.