



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2024 – 09:21 PM EDT

PDB ID : 2YHU  
Title : Trypanosoma brucei PTR1 in complex with inhibitor WHF30  
Authors : Nerini, E.; Dawson, A.; Hunter, W.N.; Costi, M.P.  
Deposited on : 2011-05-06  
Resolution : 2.01 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

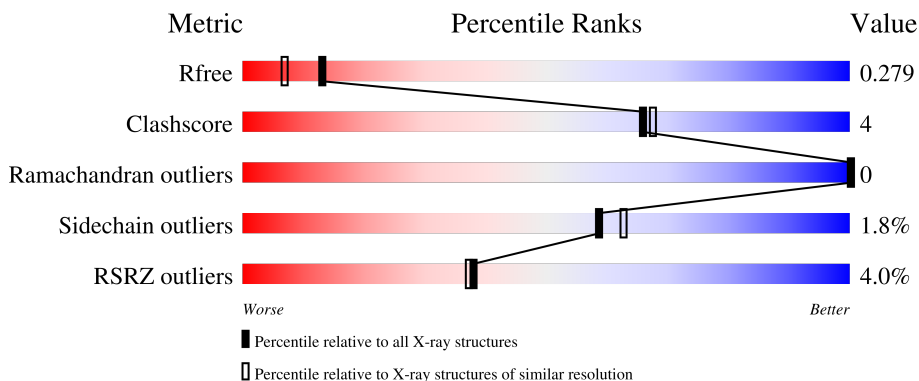
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.01 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	288	
1	B	288	
1	C	288	
1	D	288	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 8164 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 PTERIDINE REDUCTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	251	1886	1187	331	356	12	0	3	0
1	B	250	1882	1183	330	357	12	0	4	0
1	C	252	1888	1187	333	356	12	0	2	0
1	D	249	1876	1177	332	356	11	0	2	0

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP O76290
A	-18	GLY	-	expression tag	UNP O76290
A	-17	SER	-	expression tag	UNP O76290
A	-16	SER	-	expression tag	UNP O76290
A	-15	HIS	-	expression tag	UNP O76290
A	-14	HIS	-	expression tag	UNP O76290
A	-13	HIS	-	expression tag	UNP O76290
A	-12	HIS	-	expression tag	UNP O76290
A	-11	HIS	-	expression tag	UNP O76290
A	-10	HIS	-	expression tag	UNP O76290
A	-9	SER	-	expression tag	UNP O76290
A	-8	SER	-	expression tag	UNP O76290
A	-7	GLY	-	expression tag	UNP O76290
A	-6	LEU	-	expression tag	UNP O76290
A	-5	VAL	-	expression tag	UNP O76290
A	-4	PRO	-	expression tag	UNP O76290
A	-3	ARG	-	expression tag	UNP O76290
A	-2	GLY	-	expression tag	UNP O76290
A	-1	SER	-	expression tag	UNP O76290
A	0	HIS	-	expression tag	UNP O76290
B	-19	MET	-	expression tag	UNP O76290

*Continued on next page...*

*Continued from previous page...*

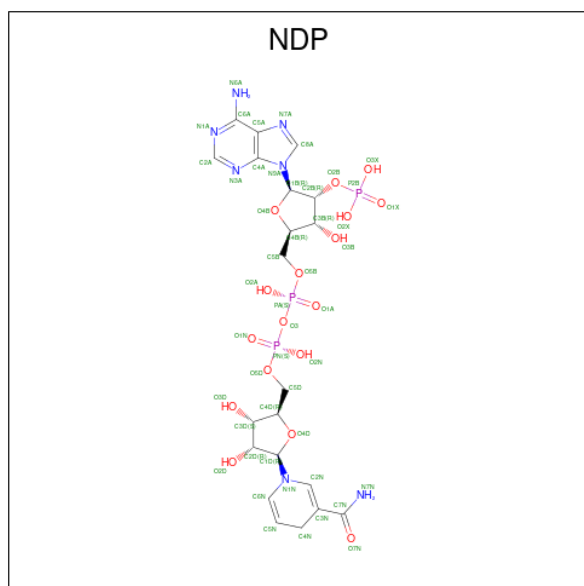
Chain	Residue	Modelled	Actual	Comment	Reference
B	-18	GLY	-	expression tag	UNP O76290
B	-17	SER	-	expression tag	UNP O76290
B	-16	SER	-	expression tag	UNP O76290
B	-15	HIS	-	expression tag	UNP O76290
B	-14	HIS	-	expression tag	UNP O76290
B	-13	HIS	-	expression tag	UNP O76290
B	-12	HIS	-	expression tag	UNP O76290
B	-11	HIS	-	expression tag	UNP O76290
B	-10	HIS	-	expression tag	UNP O76290
B	-9	SER	-	expression tag	UNP O76290
B	-8	SER	-	expression tag	UNP O76290
B	-7	GLY	-	expression tag	UNP O76290
B	-6	LEU	-	expression tag	UNP O76290
B	-5	VAL	-	expression tag	UNP O76290
B	-4	PRO	-	expression tag	UNP O76290
B	-3	ARG	-	expression tag	UNP O76290
B	-2	GLY	-	expression tag	UNP O76290
B	-1	SER	-	expression tag	UNP O76290
B	0	HIS	-	expression tag	UNP O76290
C	-19	MET	-	expression tag	UNP O76290
C	-18	GLY	-	expression tag	UNP O76290
C	-17	SER	-	expression tag	UNP O76290
C	-16	SER	-	expression tag	UNP O76290
C	-15	HIS	-	expression tag	UNP O76290
C	-14	HIS	-	expression tag	UNP O76290
C	-13	HIS	-	expression tag	UNP O76290
C	-12	HIS	-	expression tag	UNP O76290
C	-11	HIS	-	expression tag	UNP O76290
C	-10	HIS	-	expression tag	UNP O76290
C	-9	SER	-	expression tag	UNP O76290
C	-8	SER	-	expression tag	UNP O76290
C	-7	GLY	-	expression tag	UNP O76290
C	-6	LEU	-	expression tag	UNP O76290
C	-5	VAL	-	expression tag	UNP O76290
C	-4	PRO	-	expression tag	UNP O76290
C	-3	ARG	-	expression tag	UNP O76290
C	-2	GLY	-	expression tag	UNP O76290
C	-1	SER	-	expression tag	UNP O76290
C	0	HIS	-	expression tag	UNP O76290
D	-19	MET	-	expression tag	UNP O76290
D	-18	GLY	-	expression tag	UNP O76290
D	-17	SER	-	expression tag	UNP O76290

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-16	SER	-	expression tag	UNP O76290
D	-15	HIS	-	expression tag	UNP O76290
D	-14	HIS	-	expression tag	UNP O76290
D	-13	HIS	-	expression tag	UNP O76290
D	-12	HIS	-	expression tag	UNP O76290
D	-11	HIS	-	expression tag	UNP O76290
D	-10	HIS	-	expression tag	UNP O76290
D	-9	SER	-	expression tag	UNP O76290
D	-8	SER	-	expression tag	UNP O76290
D	-7	GLY	-	expression tag	UNP O76290
D	-6	LEU	-	expression tag	UNP O76290
D	-5	VAL	-	expression tag	UNP O76290
D	-4	PRO	-	expression tag	UNP O76290
D	-3	ARG	-	expression tag	UNP O76290
D	-2	GLY	-	expression tag	UNP O76290
D	-1	SER	-	expression tag	UNP O76290
D	0	HIS	-	expression tag	UNP O76290

- Molecule 2 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula:  $C_{21}H_{30}N_7O_{17}P_3$ ).



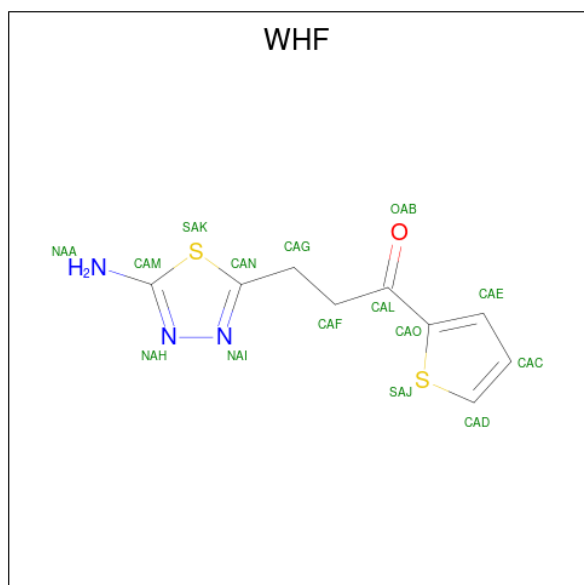
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
2	C	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	D	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 3 is 3-(5-AMINO-1,3,4-THIADIAZOL-2-YL)-1-THIOPHEN-2-YLPROPAN-1-ONE E (three-letter code: WHF) (formula: C<sub>9</sub>H<sub>9</sub>N<sub>3</sub>OS<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	A	1	Total	C	N	O	S	0	0
			15	9	3	1	2		
3	B	1	Total	C	N	O	S	0	0
			15	9	3	1	2		
3	C	1	Total	C	N	O	S	0	0
			15	9	3	1	2		
3	D	1	Total	C	N	O	S	0	0
			15	9	3	1	2		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0

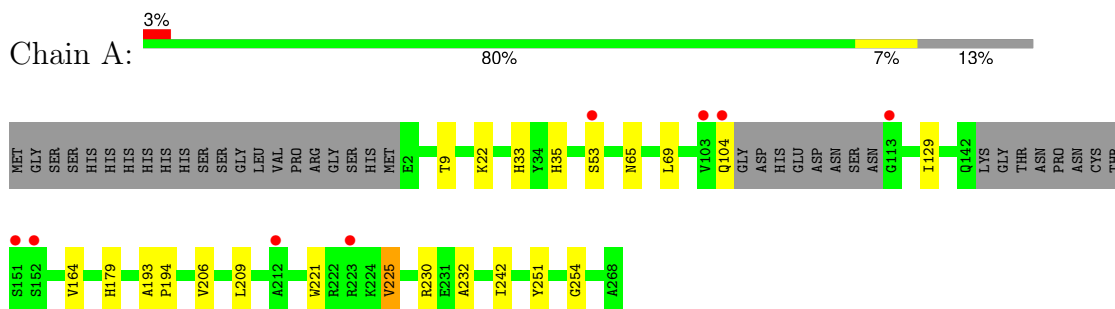
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	104	Total O 104 104	0	0
5	B	99	Total O 99 99	0	0
5	C	91	Total O 91 91	0	0
5	D	78	Total O 78 78	0	0

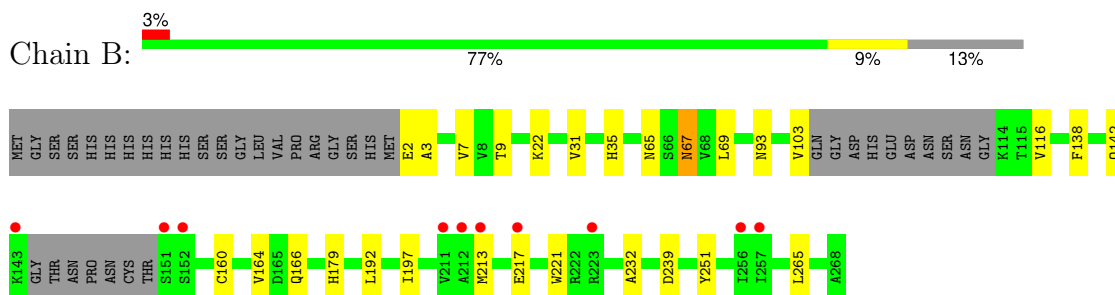
### 3 Residue-property plots i

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

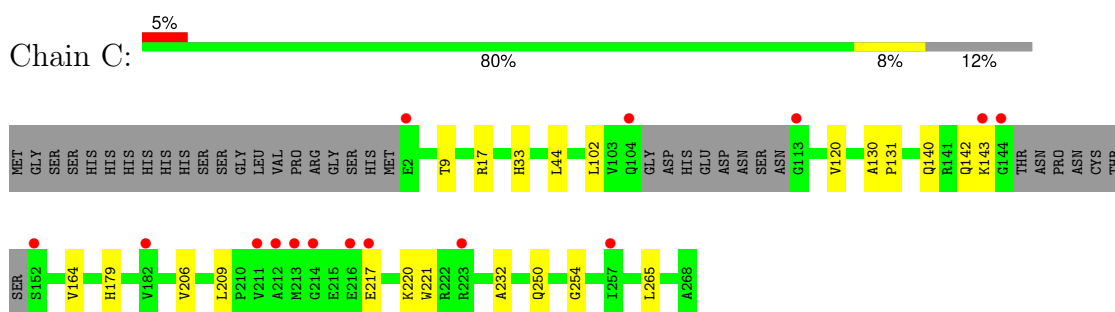
- Molecule 1: PTERIDINE REDUCTASE



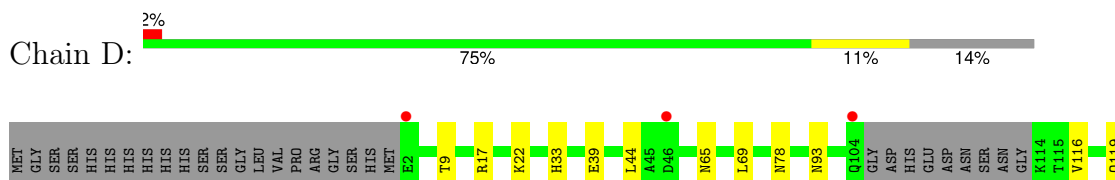
- Molecule 1: PTERIDINE REDUCTASE



- Molecule 1: PTERIDINE REDUCTASE



- Molecule 1: PTERIDINE REDUCTASE







## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.55Å 90.29Å 82.46Å 90.00° 115.43° 90.00°	Depositor
Resolution (Å)	40.85 – 2.01 40.85 – 2.01	Depositor EDS
% Data completeness (in resolution range)	97.1 (40.85-2.01) 97.1 (40.85-2.01)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.69 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.235 , 0.280 0.234 , 0.279	Depositor DCC
$R_{free}$ test set	3277 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.3	Xtrriage
Anisotropy	0.046	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 46.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.005 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	8164	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.98 % 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 5.3626e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: ACT, WHF, NDP, CSX

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.49	0/1914	0.55	0/2593
1	B	0.50	0/1913	0.57	0/2592
1	C	0.50	0/1913	0.55	0/2590
1	D	0.50	0/1898	0.54	0/2573
All	All	0.50	0/7638	0.55	0/10348

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	1886	0	1923	16	0
1	B	1882	0	1917	18	0
1	C	1888	0	1923	13	0
1	D	1876	0	1898	25	0
2	A	48	0	26	2	0
2	B	48	0	26	1	0
2	C	48	0	26	0	0
2	D	48	0	26	0	0
3	A	15	0	9	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	15	0	9	2	0
3	C	15	0	9	1	0
3	D	15	0	9	2	0
4	B	4	0	3	0	0
4	D	4	0	3	0	0
5	A	104	0	0	0	1
5	B	99	0	0	1	1
5	C	91	0	0	1	0
5	D	78	0	0	2	0
All	All	8164	0	7807	63	1

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 4.

All (63) 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:230[B]:ARG:HH11	1:D:230[B]:ARG:HG3	1.03	1.11
1:D:230[B]:ARG:HH11	1:D:230[B]:ARG:CG	1.65	1.09
1:D:230[B]:ARG:HG3	1:D:230[B]:ARG:NH1	1.82	0.90
1:C:206:VAL:HG13	1:C:209:LEU:HD21	1.70	0.73
1:D:230[B]:ARG:CG	1:D:230[B]:ARG:NH1	2.38	0.70
1:B:116:VAL:HG21	1:D:136:MET:HE1	1.77	0.66
1:A:9:THR:HA	1:A:33:HIS:HB3	1.79	0.64
1:D:78:ASN:OD1	1:D:141:ARG:NH1	2.31	0.63
1:B:22:LYS:HE3	1:B:239:ASP:OD1	2.00	0.62
1:A:104:GLN:HG2	1:C:140:GLN:HE22	1.66	0.60
1:A:221:TRP:CZ3	3:A:302:WHF:HAD	2.36	0.59
1:A:164:VAL:HG22	1:A:179:HIS:CD2	2.36	0.59
1:B:2:GLU:HG2	1:B:3:ALA:H	1.67	0.58
1:A:225:VAL:HG21	1:A:230:ARG:O	2.05	0.56
1:D:221:TRP:CZ3	3:D:302:WHF:HAD	2.41	0.56
1:B:67[A]:ASN:H	1:B:67[A]:ASN:HD22	1.56	0.54
1:B:221:TRP:CZ3	3:B:302:WHF:HAD	2.43	0.53
1:B:221:TRP:CH2	3:B:302:WHF:HAD	2.44	0.53
1:C:9:THR:HA	1:C:33:HIS:HB3	1.91	0.52
1:D:9:THR:HA	1:D:33:HIS:HB3	1.91	0.52
1:C:232:ALA:HB2	1:D:251:TYR:CE2	2.45	0.51
1:A:65:ASN:HA	1:A:69:LEU:HD22	1.92	0.51
1:C:221:TRP:CH2	3:C:302:WHF:HAD	2.48	0.49
1:D:206:VAL:HG22	1:D:263:LEU:HD22	1.93	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:160:CYS:HB3	5:B:2057:HOH:O	2.13	0.48
1:A:251:TYR:CE2	1:B:232:ALA:HB2	2.49	0.47
1:A:232:ALA:HB2	1:B:251:TYR:CE2	2.49	0.47
1:B:65:ASN:HA	1:B:69:LEU:HD22	1.97	0.47
1:D:221:TRP:CH2	3:D:302:WHF:HAD	2.49	0.47
1:A:22:LYS:HG2	1:A:242:ILE:HG12	1.97	0.46
1:C:265:LEU:HD11	1:D:254:GLY:HA3	1.96	0.46
1:D:116:VAL:HA	1:D:119:GLN:HE21	1.80	0.46
1:A:193:ALA:HB3	1:A:194:PRO:HD3	1.98	0.46
1:D:141:ARG:HG2	5:D:2047:HOH:O	2.16	0.45
1:D:164:VAL:HG22	1:D:179:HIS:CD2	2.51	0.45
1:B:9:THR:O	1:B:93:ASN:HB3	2.16	0.45
1:B:164:VAL:HG22	1:B:179:HIS:CD2	2.51	0.45
1:C:130:ALA:HB3	1:C:131:PRO:HD3	1.99	0.45
1:A:129[B]:ILE:HD13	1:C:120:VAL:HG11	1.99	0.45
1:A:221:TRP:CH2	3:A:302:WHF:HAD	2.51	0.45
1:D:65:ASN:HA	1:D:69:LEU:HD22	1.98	0.45
1:A:35:HIS:HB2	2:A:301:NDP:C2A	2.47	0.44
1:C:164:VAL:HG22	1:C:179:HIS:CD2	2.52	0.44
1:C:217:GLU:HG2	5:C:2064:HOH:O	2.18	0.44
1:A:254:GLY:HA3	1:B:265:LEU:HD11	2.00	0.44
1:D:17:ARG:HG3	1:D:44:LEU:HD22	2.00	0.44
1:B:35:HIS:HB2	2:B:301:NDP:C2A	2.48	0.43
1:B:192:LEU:HB3	1:B:197:ILE:HB	2.00	0.43
1:C:254:GLY:HA3	1:D:265:LEU:HD11	2.01	0.43
1:D:9:THR:O	1:D:93:ASN:HB3	2.18	0.43
1:D:220:LYS:O	1:D:223:ARG:HG3	2.18	0.43
2:A:301:NDP:H41N	3:A:302:WHF:HAF2	2.00	0.43
1:D:192:LEU:HB3	1:D:197:ILE:HB	2.01	0.43
1:B:138:PHE:O	1:B:142:GLN:HG2	2.20	0.42
1:B:7:VAL:HG22	1:B:31:VAL:HB	2.02	0.42
1:A:206:VAL:CG2	1:A:209:LEU:HD11	2.50	0.42
1:C:17:ARG:HG3	1:C:44:LEU:HD22	2.02	0.42
1:D:225:VAL:O	1:D:229:ARG:HD3	2.21	0.41
1:D:160:CYS:HB3	5:D:2052:HOH:O	2.21	0.41
1:D:22:LYS:HG2	1:D:242:ILE:HG13	2.03	0.41
1:A:221:TRP:CE3	3:A:302:WHF:HAD	2.56	0.40
1:B:103:VAL:HG21	1:D:195:TYR:CZ	2.56	0.40

All (1) 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)
5:A:2009:HOH:O	5:B:2079:HOH:O[2_545]	1.91	0.29

### 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	247/288 (86%)	239 (97%)	8 (3%)	0	100	100
1	B	247/288 (86%)	237 (96%)	10 (4%)	0	100	100
1	C	247/288 (86%)	237 (96%)	10 (4%)	0	100	100
1	D	244/288 (85%)	233 (96%)	11 (4%)	0	100	100
All	All	985/1152 (86%)	946 (96%)	39 (4%)	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	201/230 (87%)	199 (99%)	2 (1%)	76	81
1	B	202/230 (88%)	197 (98%)	5 (2%)	47	49
1	C	200/230 (87%)	195 (98%)	5 (2%)	47	49
1	D	199/230 (86%)	196 (98%)	3 (2%)	65	69
All	All	802/920 (87%)	787 (98%)	15 (2%)	59	61

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	SER
1	A	225	VAL
1	B	67[A]	ASN
1	B	67[B]	ASN
1	B	166	GLN
1	B	213	MET
1	B	217	GLU
1	C	102	LEU
1	C	142	GLN
1	C	143	LYS
1	C	220	LYS
1	C	250	GLN
1	D	39	GLU
1	D	206	VAL
1	D	215	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	B	179	HIS
1	C	179	HIS
1	D	119	GLN
1	D	179	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

4 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	CSX	D	168	1	3,6,7	0.53	0	1,6,8	1.34	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	CSX	C	168	1	3,6,7	0.44	0	1,6,8	0.03	0
1	CSX	B	168	1	3,6,7	0.54	0	1,6,8	0.41	0
1	CSX	A	168	1	3,6,7	0.67	0	1,6,8	0.18	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
1	CSX	D	168	1	-	0/2/5/7	-
1	CSX	C	168	1	-	0/2/5/7	-
1	CSX	B	168	1	-	0/2/5/7	-
1	CSX	A	168	1	-	0/2/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

10 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	ACT	D	1269	-	3,3,3	0.80	0	3,3,3	1.35	0
3	WHF	C	302	-	11,16,16	2.86	3 (27%)	7,21,21	2.51	1 (14%)
2	NDP	A	301	-	47,52,52	1.59	7 (14%)	61,80,80	1.92	8 (13%)
3	WHF	B	302	-	11,16,16	2.93	3 (27%)	7,21,21	2.51	1 (14%)
3	WHF	D	302	-	11,16,16	2.88	3 (27%)	7,21,21	2.27	1 (14%)
2	NDP	D	301	-	47,52,52	1.57	5 (10%)	61,80,80	1.85	6 (9%)
4	ACT	B	1269	-	3,3,3	0.91	0	3,3,3	1.35	0
2	NDP	C	301	-	47,52,52	1.52	6 (12%)	61,80,80	1.85	7 (11%)
3	WHF	A	302	-	11,16,16	2.92	3 (27%)	7,21,21	2.62	2 (28%)
2	NDP	B	301	-	47,52,52	1.59	5 (10%)	61,80,80	1.81	9 (14%)

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	WHF	C	302	-	-	1/4/9/9	0/2/2/2
2	NDP	A	301	-	-	1/30/77/77	0/5/5/5
3	WHF	B	302	-	-	1/4/9/9	0/2/2/2
3	WHF	D	302	-	-	0/4/9/9	0/2/2/2
2	NDP	D	301	-	-	1/30/77/77	0/5/5/5
2	NDP	C	301	-	-	1/30/77/77	0/5/5/5
3	WHF	A	302	-	-	0/4/9/9	0/2/2/2
2	NDP	B	301	-	-	1/30/77/77	0/5/5/5

All (35) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	302	WHF	NAH-NAI	-6.55	1.25	1.37
3	B	302	WHF	NAH-NAI	-6.39	1.25	1.37
3	D	302	WHF	NAH-NAI	-6.37	1.25	1.37
3	C	302	WHF	NAH-NAI	-6.32	1.25	1.37
3	D	302	WHF	CAG-CAN	5.88	1.53	1.49
3	A	302	WHF	CAO-SAJ	-5.44	1.65	1.72
2	C	301	NDP	C4N-C3N	-5.40	1.39	1.50
2	A	301	NDP	C4N-C3N	-5.23	1.40	1.50
2	B	301	NDP	C4N-C3N	-5.18	1.40	1.50
2	B	301	NDP	PA-O3	5.16	1.65	1.59
2	D	301	NDP	C4N-C3N	-5.10	1.40	1.50

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	302	WHF	CAG-CAN	4.87	1.52	1.49
3	B	302	WHF	CAO-SAJ	-4.86	1.66	1.72
3	C	302	WHF	CAG-CAN	4.68	1.52	1.49
3	C	302	WHF	CAO-SAJ	-4.63	1.66	1.72
2	D	301	NDP	P2B-O1X	4.48	1.64	1.50
2	A	301	NDP	P2B-O1X	4.42	1.64	1.50
3	A	302	WHF	CAG-CAN	3.98	1.52	1.49
2	C	301	NDP	P2B-O1X	3.83	1.62	1.50
2	D	301	NDP	C4N-C5N	-3.75	1.39	1.49
2	A	301	NDP	C4N-C5N	-3.67	1.39	1.49
2	A	301	NDP	P2B-O2B	3.62	1.65	1.59
2	C	301	NDP	C4N-C5N	-3.48	1.40	1.49
2	B	301	NDP	C4N-C5N	-3.37	1.40	1.49
2	C	301	NDP	O4B-C1B	3.07	1.44	1.40
3	D	302	WHF	CAO-SAJ	-2.93	1.68	1.72
2	D	301	NDP	P2B-O2B	2.48	1.63	1.59
2	C	301	NDP	C7N-C3N	2.41	1.53	1.48
2	D	301	NDP	C6N-C5N	2.37	1.40	1.33
2	A	301	NDP	C7N-C3N	2.30	1.53	1.48
2	A	301	NDP	O4B-C1B	2.24	1.43	1.40
2	A	301	NDP	C6N-C5N	2.23	1.39	1.33
2	B	301	NDP	P2B-O1X	2.14	1.57	1.50
2	B	301	NDP	P2B-O2B	2.13	1.63	1.59
2	C	301	NDP	C6N-C5N	2.08	1.39	1.33

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	NDP	O4B-C1B-N9A	8.31	119.77	108.75
2	D	301	NDP	C4B-O4B-C1B	-7.66	102.91	109.92
2	C	301	NDP	C4B-O4B-C1B	-7.41	103.14	109.92
2	B	301	NDP	C4B-O4B-C1B	-6.98	103.53	109.92
2	A	301	NDP	C4B-O4B-C1B	-6.87	103.63	109.92
2	D	301	NDP	O4B-C1B-N9A	6.79	117.75	108.75
3	A	302	WHF	CAC-CAD-SAJ	-6.43	107.76	112.98
3	C	302	WHF	CAC-CAD-SAJ	-6.31	107.86	112.98
2	B	301	NDP	O4B-C1B-N9A	6.28	117.08	108.75
3	B	302	WHF	CAC-CAD-SAJ	-6.17	107.97	112.98
2	C	301	NDP	O4B-C1B-N9A	6.13	116.88	108.75
3	D	302	WHF	CAC-CAD-SAJ	-5.83	108.25	112.98
2	D	301	NDP	N3A-C2A-N1A	-5.54	121.15	128.67
2	C	301	NDP	N3A-C2A-N1A	-5.52	121.18	128.67

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	NDP	N3A-C2A-N1A	-5.42	121.32	128.67
2	B	301	NDP	N3A-C2A-N1A	-5.24	121.56	128.67
2	C	301	NDP	C2B-C1B-N9A	-4.71	102.09	112.56
2	B	301	NDP	C2B-C1B-N9A	-4.34	102.93	112.56
2	D	301	NDP	C2B-C1B-N9A	-3.82	104.07	112.56
2	A	301	NDP	C2B-C1B-N9A	-3.66	104.43	112.56
2	C	301	NDP	C2D-C1D-N1N	-3.60	104.46	113.31
2	A	301	NDP	C2D-C1D-N1N	-3.06	105.77	113.31
2	A	301	NDP	C3N-C2N-N1N	-2.78	119.12	123.20
2	B	301	NDP	C3N-C2N-N1N	-2.68	119.27	123.20
2	B	301	NDP	C2D-C1D-N1N	-2.64	106.82	113.31
2	D	301	NDP	C3N-C2N-N1N	-2.44	119.62	123.20
2	A	301	NDP	O2B-P2B-O1X	-2.37	100.89	109.33
2	B	301	NDP	O3X-P2B-O2B	2.33	114.94	105.85
2	B	301	NDP	C1B-N9A-C4A	-2.32	122.56	126.64
2	D	301	NDP	C1B-N9A-C4A	-2.32	122.57	126.64
2	C	301	NDP	C3N-C2N-N1N	-2.15	120.05	123.20
2	B	301	NDP	O3X-P2B-O2X	2.14	115.83	107.80
3	A	302	WHF	CAF-CAG-CAN	-2.12	109.21	112.79
2	C	301	NDP	O2B-P2B-O1X	-2.03	102.09	109.33
2	A	301	NDP	O4B-C4B-C3B	-2.02	101.15	105.15

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	302	WHF	CAL-CAF-CAG-CAN
3	C	302	WHF	CAL-CAF-CAG-CAN
2	C	301	NDP	O4D-C1D-N1N-C6N
2	B	301	NDP	O4D-C1D-N1N-C6N
2	A	301	NDP	O4D-C1D-N1N-C6N
2	D	301	NDP	O4D-C1D-N1N-C6N

There are no ring outliers.

6 monomers are involved in 11 short contacts:

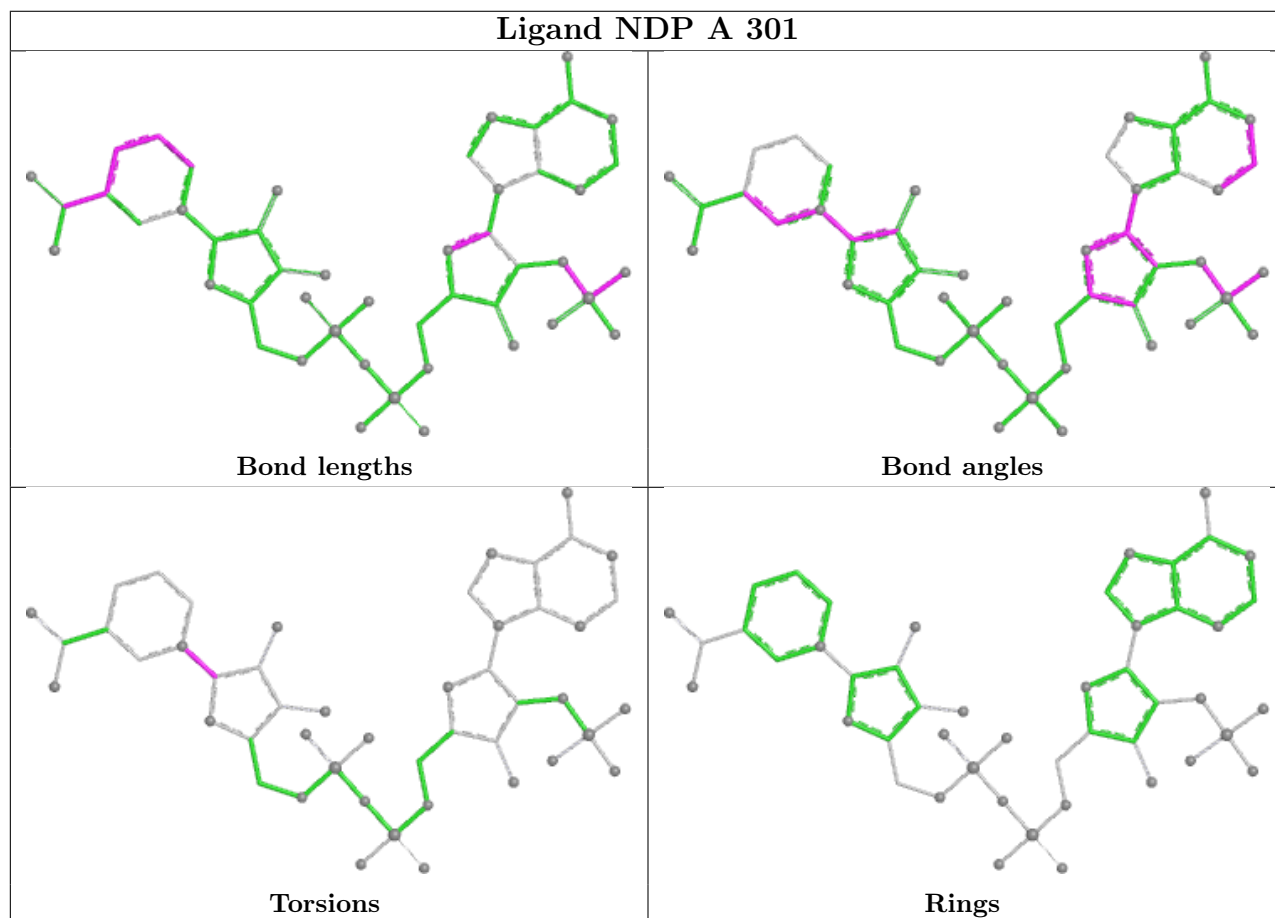
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	302	WHF	1	0
2	A	301	NDP	2	0
3	B	302	WHF	2	0
3	D	302	WHF	2	0

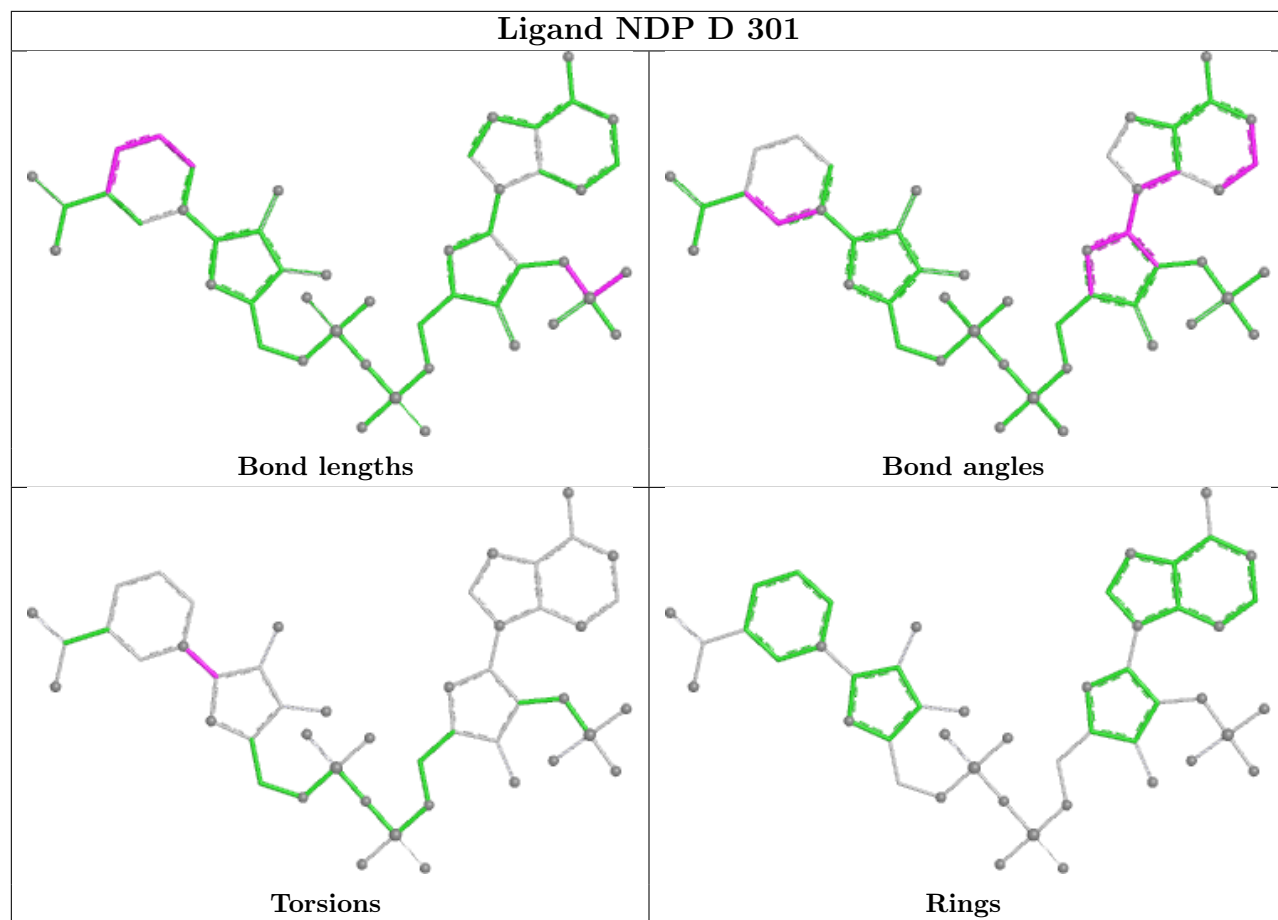
Continued on next page...

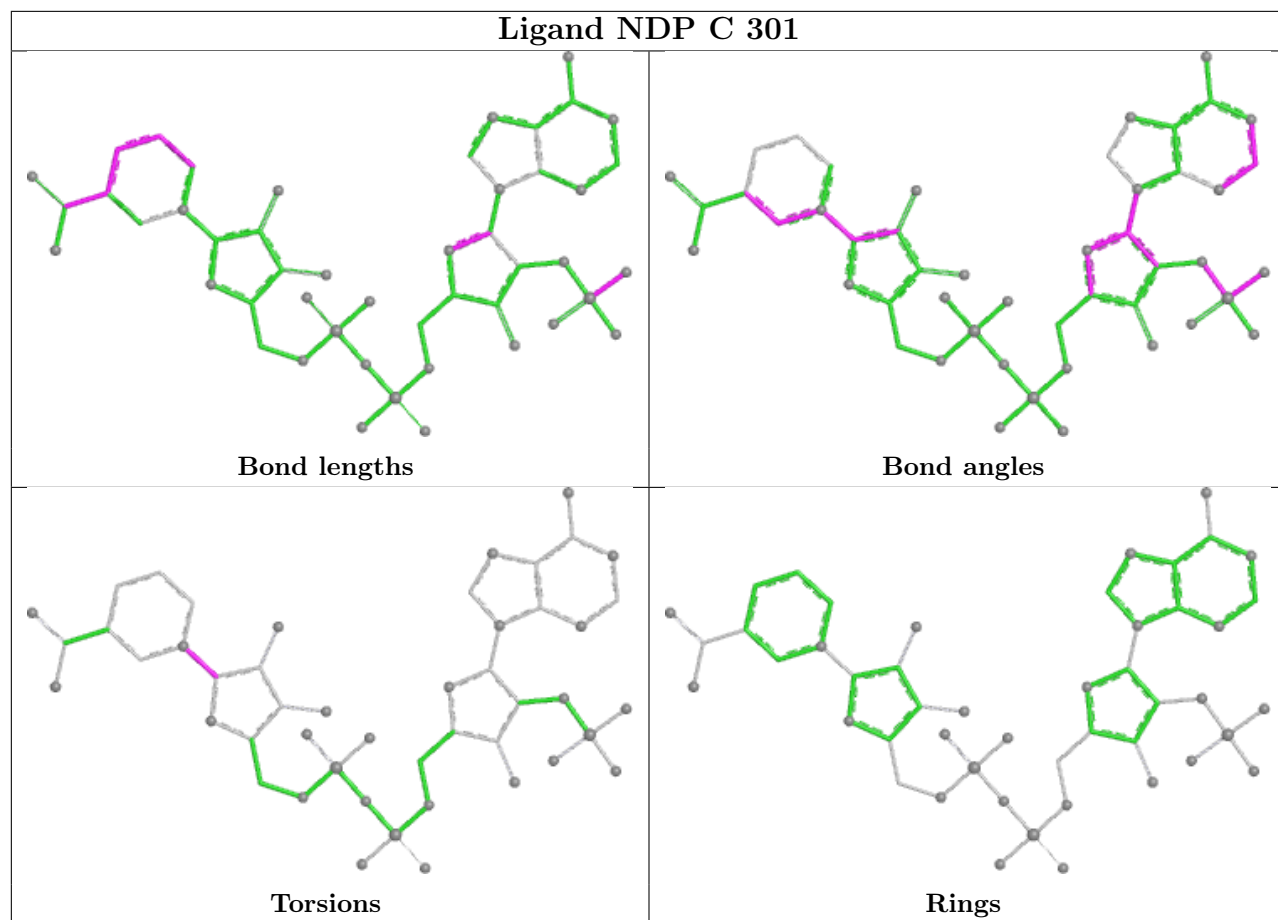
Continued from previous page...

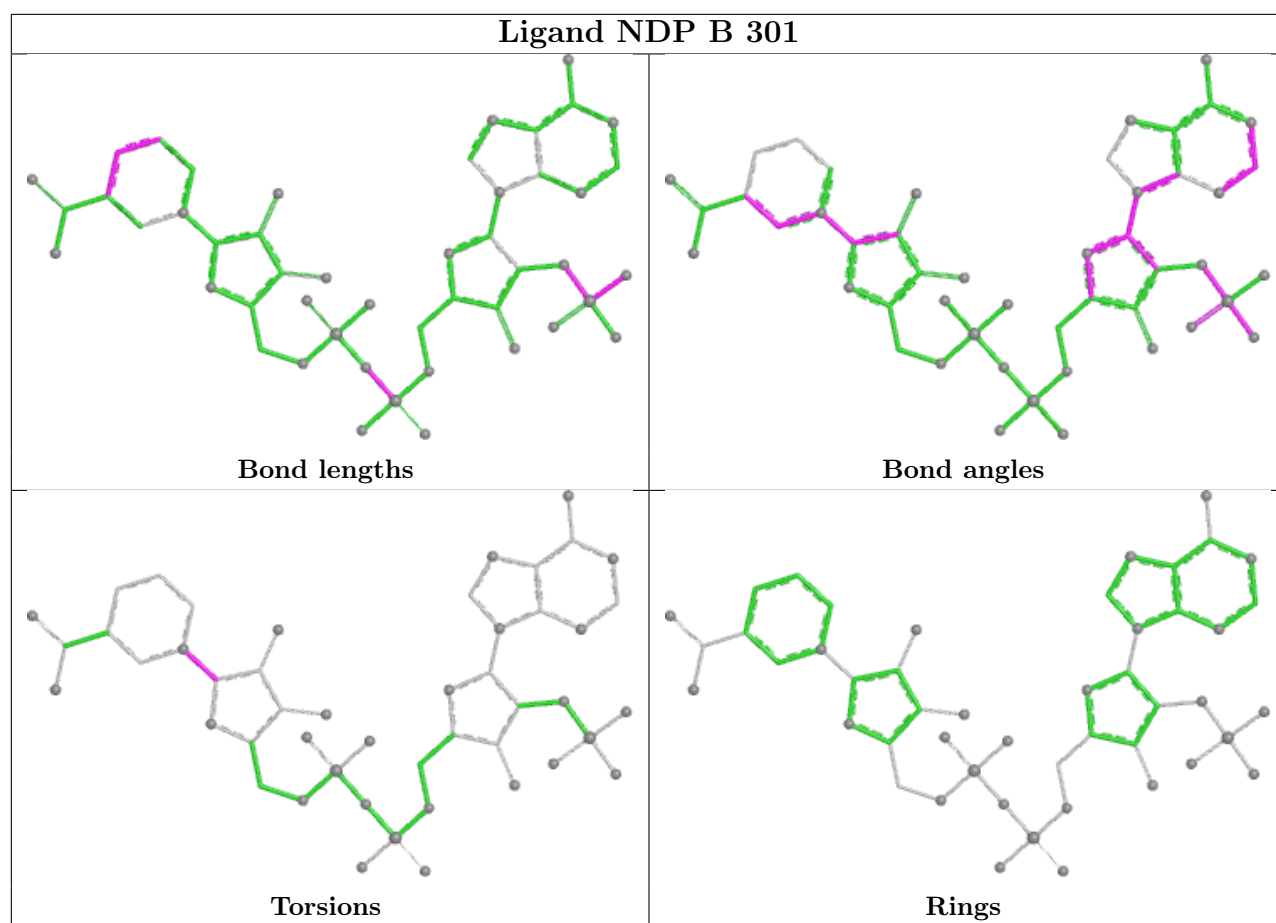
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	302	WHF	4	0
2	B	301	NDP	1	0

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	250/288 (86%)	0.28	8 (3%) 47 46	11, 18, 32, 39	2 (0%)
1	B	249/288 (86%)	0.33	10 (4%) 38 37	11, 18, 32, 43	0
1	C	251/288 (87%)	0.39	15 (5%) 21 20	10, 19, 37, 46	0
1	D	248/288 (86%)	0.24	7 (2%) 53 51	11, 19, 32, 39	0
All	All	998/1152 (86%)	0.31	40 (4%) 38 37	10, 18, 33, 46	2 (0%)

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	151	SER	9.6
1	C	144	GLY	8.9
1	B	152	SER	7.9
1	C	143	LYS	6.7
1	A	151	SER	6.4
1	C	113	GLY	5.8
1	A	113	GLY	5.4
1	B	143	LYS	5.3
1	C	152	SER	5.1
1	C	212	ALA	4.8
1	D	104	GLN	4.5
1	A	104	GLN	4.3
1	D	152	SER	4.1
1	D	2	GLU	4.0
1	B	212	ALA	3.6
1	B	211	VAL	3.5
1	C	211	VAL	3.5
1	A	152	SER	3.5
1	C	223	ARG	3.5
1	C	216	GLU	3.4
1	A	212	ALA	3.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	223	ARG	3.1
1	C	104	GLN	3.0
1	B	213	MET	2.6
1	C	214	GLY	2.5
1	D	216	GLU	2.5
1	B	223	ARG	2.3
1	B	257	ILE	2.3
1	C	257	ILE	2.3
1	B	217	GLU	2.2
1	A	223	ARG	2.2
1	A	103	VAL	2.1
1	D	141	ARG	2.1
1	C	182	VAL	2.1
1	D	46	ASP	2.1
1	C	217	GLU	2.1
1	C	213	MET	2.1
1	B	256	ILE	2.0
1	C	2	GLU	2.0
1	A	53	SER	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	CSX	C	168	7/8	0.90	0.09	22,22,25,29	0
1	CSX	A	168	7/8	0.92	0.10	23,23,29,31	0
1	CSX	D	168	7/8	0.93	0.09	20,20,26,28	0
1	CSX	B	168	7/8	0.94	0.09	21,22,28,29	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

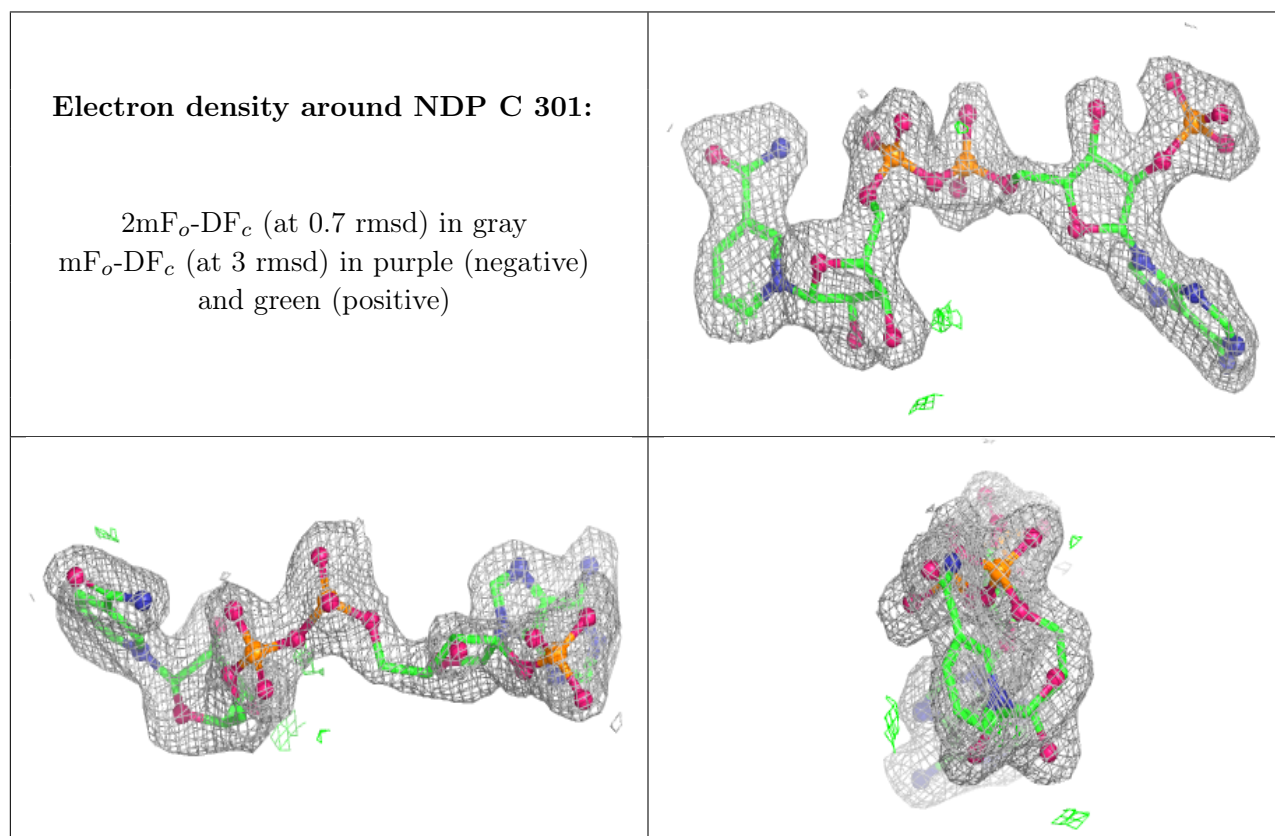
## 6.4 Ligands [i](#)

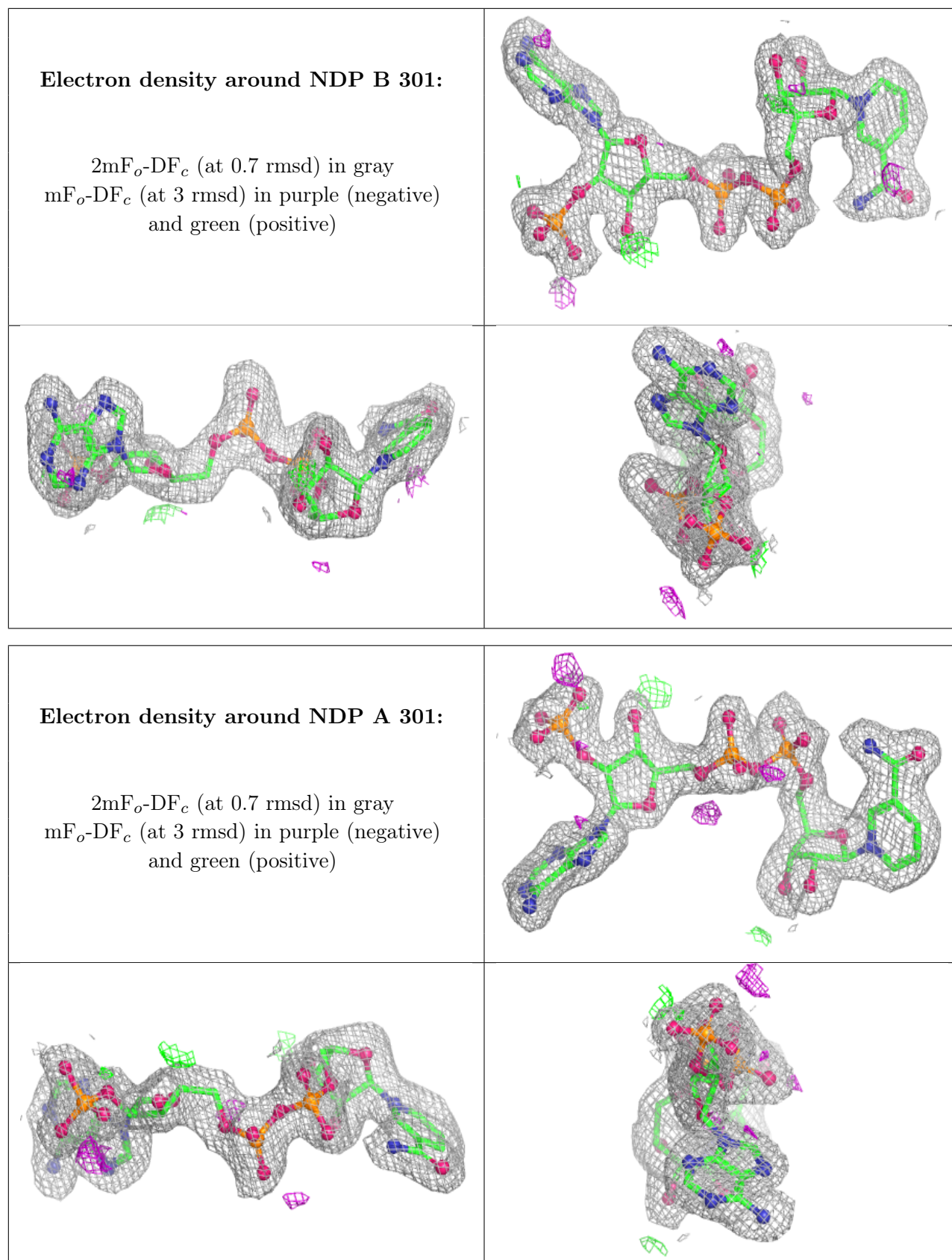
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

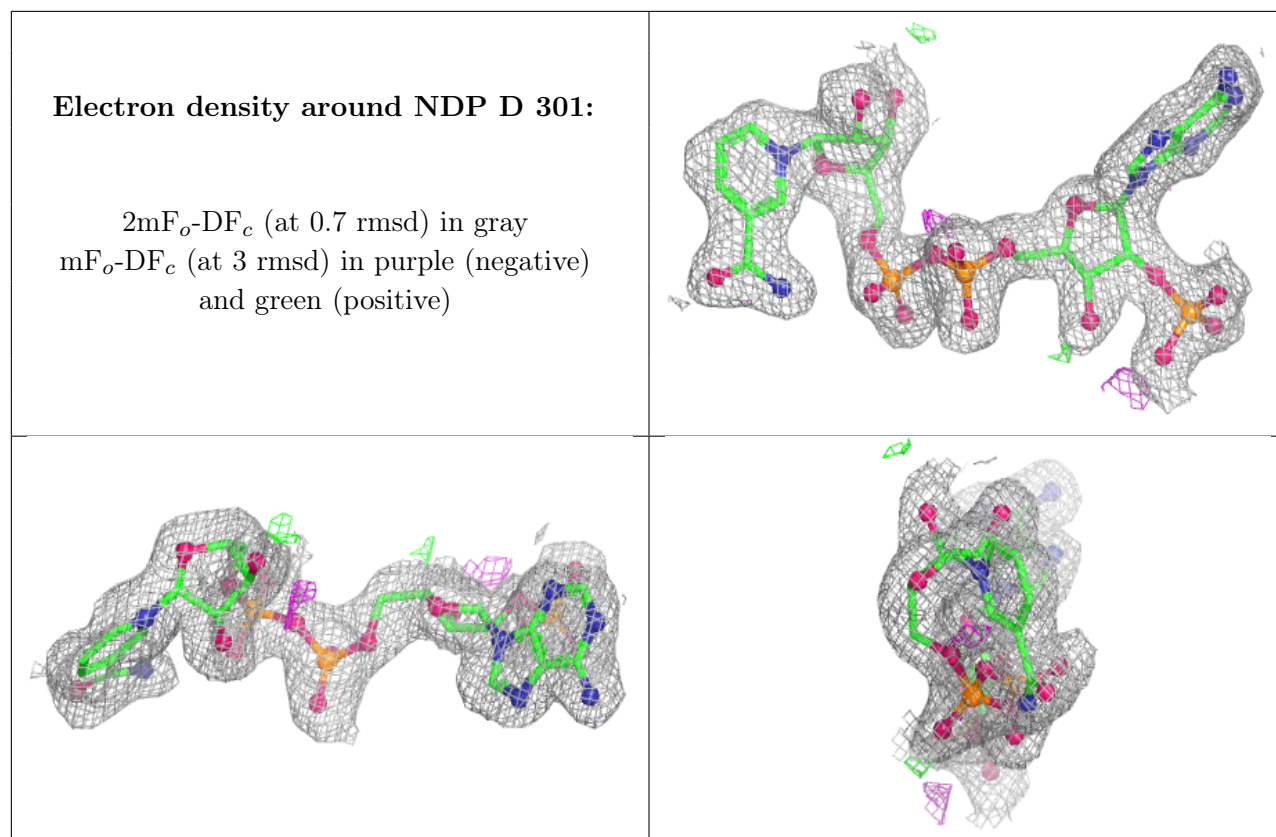
median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ACT	D	1269	4/4	0.62	0.27	66,66,66,66	0
3	WHF	B	302	15/15	0.80	0.18	30,36,45,45	0
3	WHF	C	302	15/15	0.82	0.17	46,47,53,53	0
4	ACT	B	1269	4/4	0.82	0.15	28,28,29,29	0
3	WHF	A	302	15/15	0.82	0.16	31,36,41,41	0
3	WHF	D	302	15/15	0.88	0.17	35,37,41,42	0
2	NDP	C	301	48/48	0.92	0.14	14,22,24,25	0
2	NDP	B	301	48/48	0.95	0.10	11,16,18,19	0
2	NDP	A	301	48/48	0.95	0.10	14,17,19,20	0
2	NDP	D	301	48/48	0.96	0.11	14,19,21,21	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.







## 6.5 Other polymers [i](#)

There are no such residues in this entry.