



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2020 – 12:19 PM BST

PDB ID : 4JTW
Title : Crystal structure of HCV NS5B polymerase in complex with compound 1
Authors : Coulombe, R.
Deposited on : 2013-03-24
Resolution : 3.00 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 : 2.13.1
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) : 2.13.1

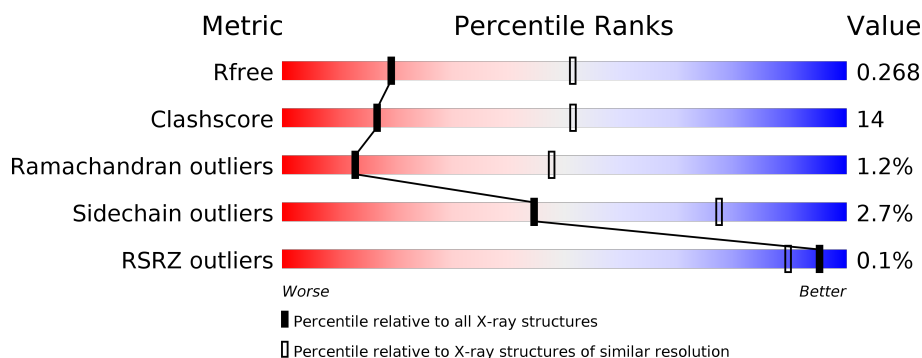
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	576	
1	B	576	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9122 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 Genome polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	560	Total	C	N	O	S	0	0	0
			4362	2747	771	812	32			
1	B	560	Total	C	N	O	S	0	0	0
			4362	2747	771	812	32			

There are 12 discrepancies between the modelled and reference sequences:

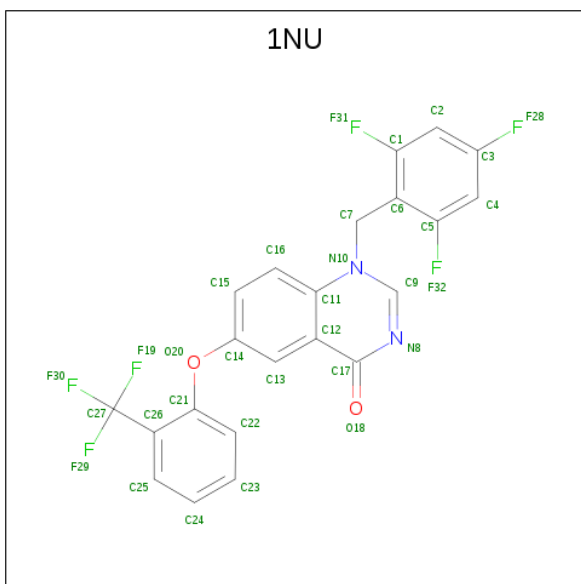
Chain	Residue	Modelled	Actual	Comment	Reference
A	571	HIS	-	EXPRESSION TAG	UNP O92972
A	572	HIS	-	EXPRESSION TAG	UNP O92972
A	573	HIS	-	EXPRESSION TAG	UNP O92972
A	574	HIS	-	EXPRESSION TAG	UNP O92972
A	575	HIS	-	EXPRESSION TAG	UNP O92972
A	576	HIS	-	EXPRESSION TAG	UNP O92972
B	571	HIS	-	EXPRESSION TAG	UNP O92972
B	572	HIS	-	EXPRESSION TAG	UNP O92972
B	573	HIS	-	EXPRESSION TAG	UNP O92972
B	574	HIS	-	EXPRESSION TAG	UNP O92972
B	575	HIS	-	EXPRESSION TAG	UNP O92972
B	576	HIS	-	EXPRESSION TAG	UNP O92972

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 3 is 1-(2,4,6-trifluorobenzyl)-6-[2-(trifluoromethyl)phenoxy]quinazolin-4(1H)-one (three-letter code: 1NU) (formula: C₂₂H₁₂F₆N₂O₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	F	N	O	0	0
			32	22	6	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	F	N	O	0	0
			32	22	6	2	2		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	Mg	0	0
			2	2		
4	A	2	Total	Mg	0	0
			2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	169	Total	O	0	0
			169	169		
5	B	149	Total	O	0	0
			149	149		

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.

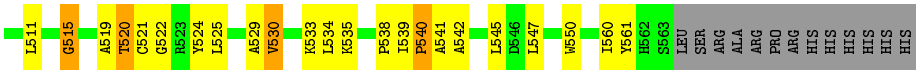
Chain A: 73% 23%

The sequence logo for Chain A displays the conservation of amino acids at each of the 100 positions. The y-axis lists the amino acids, and the x-axis lists the positions. The color scale at the top indicates the conservation level, ranging from 73% (green) to 23% (red). The logo shows that positions 1-100 are highly conserved, with many positions having a single dominant amino acid. Positions 1-100 are highly conserved, with many positions having a single dominant amino acid. Positions 1-100 are highly conserved, with many positions having a single dominant amino acid.

Chain B:

62% 33%

Parameter	Log-Likelihood (approx.)	Category
W408	0.5	Green
W409	0.2	Green
I412	0.8	Green
I419	0.3	Green
R422	0.1	Green
W423	0.4	Green
W426	0.6	Green
W427	0.2	Green
S431	0.1	Green
L434	0.3	Green
L439	0.5	Green
C445	0.2	Green
C451	0.4	Green
I454	0.6	Green
E455	0.3	Green
P456	0.1	Green
L457	0.4	Green
D458	0.2	Green
L459	0.5	Green
P460	0.3	Green
I463	0.6	Green
E464	0.4	Green
L465	0.2	Green
S470	0.5	Green
L474	0.3	Green
Y477	0.1	Green
E481	0.4	Green
W485	0.6	Green
C488	0.3	Green
L489	0.5	Green
W494	0.2	Green
P496	0.4	Green
L497	0.1	Green
R498	0.3	Green
R501	0.6	Green
R505	0.2	Green
S506	0.4	Green
W507	0.1	Green
W509	0.3	Green
W410	0.7	Red
C311	0.5	Red
T312	0.3	Red
L313	0.1	Red
L314	0.4	Red
W315	0.6	Red
L320	0.2	Red
W321	0.5	Red
W322	0.3	Red
L323	0.1	Red
C324	0.4	Red
A327	0.6	Red
E331	0.2	Red
P339	0.5	Red
Y346	0.3	Red
S347	0.1	Red
D352	0.4	Red
L353	0.6	Red
P354	0.2	Red
D359	0.5	Red
L360	0.3	Red
L362	0.1	Red
S368	0.4	Red
R369	0.6	Red
W370	0.2	Red
L474	0.5	Red
Y477	0.3	Red
E481	0.1	Red
W485	0.4	Red
C488	0.6	Red
L489	0.2	Red
W494	0.5	Red
P496	0.3	Red
L497	0.1	Red
R498	0.4	Red
R501	0.6	Red
R505	0.2	Red
S506	0.5	Red
W507	0.3	Red
W509	0.1	Red
W411	0.8	Red
C312	0.6	Red
T313	0.4	Red
L314	0.2	Red
W315	0.5	Red
L320	0.3	Red
W321	0.1	Red
W322	0.4	Red
L323	0.6	Red
C324	0.2	Red
A327	0.5	Red
E331	0.3	Red
P339	0.1	Red
Y346	0.4	Red
S347	0.6	Red
D352	0.2	Red
L353	0.5	Red
P354	0.3	Red
D359	0.1	Red
L360	0.4	Red
L362	0.6	Red
S368	0.2	Red
R369	0.5	Red
W370	0.3	Red
L474	0.1	Red
Y477	0.4	Red
E481	0.6	Red
W485	0.2	Red
C488	0.5	Red
L489	0.3	Red
W494	0.1	Red
P496	0.4	Red
L497	0.6	Red
R498	0.2	Red
R501	0.5	Red
R505	0.3	Red
S506	0.1	Red
W507	0.4	Red
W509	0.6	Red
W412	0.9	Red
E418	0.7	Red
L34	0.5	Red
L10	0.3	Red
I11	0.1	Red
T12	0.4	Red
P13	0.6	Red
G14	0.2	Red
S19	0.5	Red
E143	0.3	Red
F144	0.1	Red
C146	0.4	Red
P149	0.6	Red
L15	0.2	Red
L159	0.5	Red
I160	0.3	Red
L165	0.1	Red
R168	0.4	Red
P169	0.6	Red
E171	0.2	Red
K172	0.5	Red
M173	0.3	Red
D177	0.1	Red
T181	0.4	Red
L182	0.6	Red



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	106.43Å 108.00Å 134.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.00 39.19 – 2.99	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.00) 97.1 (39.19-2.99)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.51 (at 3.01Å)	Xtriage
Refinement program	CNX 2002	Depositor
R, R_{free}	0.200 , 0.274 0.197 , 0.268	Depositor DCC
R_{free} test set	3118 reflections (9.97%)	wwPDB-VP
Wilson B-factor (Å ²)	52.0	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 45.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.008 for k,h,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	9122	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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: 1NU, MG, GOL

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	0.37	0/4457	0.63	1/6049 (0.0%)
1	B	0.34	0/4457	0.60	0/6049
All	All	0.35	0/8914	0.61	1/12098 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	351	GLY	N-CA-C	-5.18	100.15	113.10

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	4362	0	4374	104	0
1	B	4362	0	4374	145	0
2	A	6	0	8	0	0
2	B	6	0	8	0	0
3	A	32	0	12	1	0
3	B	32	0	12	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	169	0	0	7	0
5	B	149	0	0	12	0
All	All	9122	0	8788	247	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 14.

The worst 5 of 247 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:B:515:GLY:HA2	1:B:519:ALA:HB2	1.48	0.92
1:B:314:LEU:HB3	1:B:321:VAL:HG13	1.52	0.90
1:A:18:GLU:HG2	1:A:401:ARG:NH2	1.84	0.90
1:B:233:ILE:HD12	1:B:262:ILE:HA	1.59	0.84
1:A:394:ARG:O	1:A:398:GLU:HG3	1.77	0.84

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	556/576 (96%)	515 (93%)	38 (7%)	3 (0%)	29 68
1	B	556/576 (96%)	499 (90%)	47 (8%)	10 (2%)	8 37
All	All	1112/1152 (96%)	1014 (91%)	85 (8%)	13 (1%)	13 48

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	15	ALA
1	B	33	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	525	LEU
1	B	530	VAL
1	B	541	ALA

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	477/491 (97%)	463 (97%)	14 (3%)	42	76
1	B	477/491 (97%)	465 (98%)	12 (2%)	47	79
All	All	954/982 (97%)	928 (97%)	26 (3%)	44	77

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	505	ARG
1	B	34	HIS
1	B	521	CYS
1	A	544	GLN
1	A	547	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	562	HIS
1	B	34	HIS
1	B	446	GLN
1	A	544	GLN
1	B	251	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	GOL	B	601	-	5,5,5	0.43	0	5,5,5	0.18	0
3	1NU	A	3002	-	34,35,35	2.09	11 (32%)	46,52,52	1.58	8 (17%)
3	1NU	B	602	-	34,35,35	2.44	14 (41%)	46,52,52	1.77	11 (23%)
2	GOL	A	3001	-	5,5,5	0.45	0	5,5,5	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
2	GOL	B	601	-	-	0/4/4/4	-
3	1NU	A	3002	-	-	0/14/14/14	0/4/4/4
3	1NU	B	602	-	-	0/14/14/14	0/4/4/4
2	GOL	A	3001	-	-	0/4/4/4	-

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	602	1NU	C12-C11	5.88	1.46	1.41
3	B	602	1NU	C11-N10	5.18	1.46	1.39

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	3002	1NU	C12-C11	4.80	1.45	1.41
3	A	3002	1NU	C11-N10	4.32	1.45	1.39
3	B	602	1NU	C25-C26	4.14	1.44	1.39

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	602	1NU	C16-C11-N10	4.56	126.68	121.55
3	A	3002	1NU	C4-C5-C6	-4.34	121.55	124.59
3	B	602	1NU	C4-C5-C6	-3.99	121.79	124.59
3	B	602	1NU	C2-C1-C6	-3.86	121.88	124.59
3	A	3002	1NU	C16-C11-N10	3.73	125.75	121.55

There are no chirality outliers.

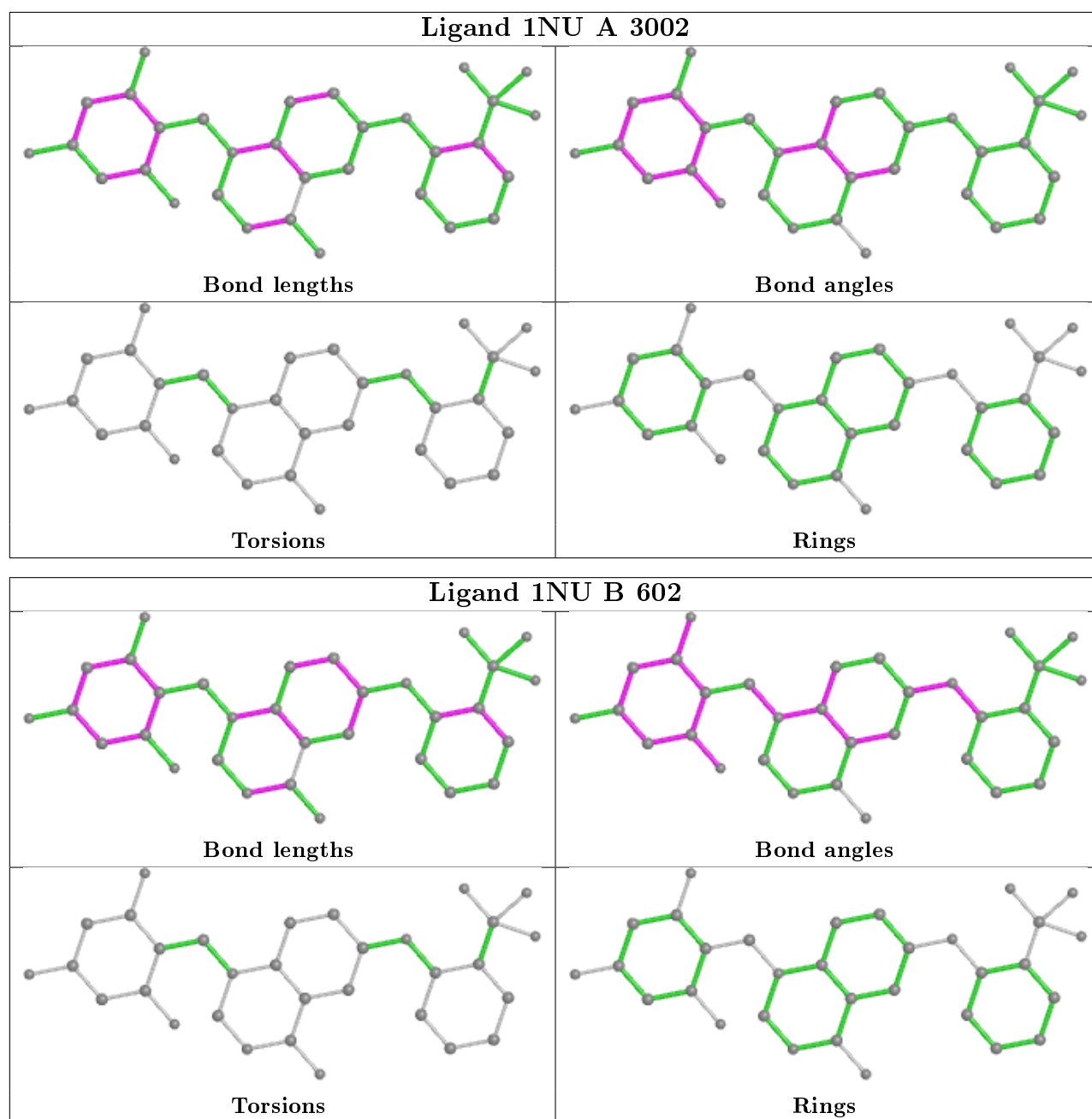
There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	3002	1NU	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, 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	560/576 (97%)	-0.66	0 100 100	16, 36, 59, 79	0
1	B	560/576 (97%)	-0.40	1 (0%) 95 87	19, 50, 91, 103	0
All	All	1120/1152 (97%)	-0.53	1 (0%) 95 89	16, 42, 80, 103	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	24	ASN	2.3

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	MG	B	604	1/1	0.70	0.40	56,56,56,56	0
4	MG	A	3003	1/1	0.76	0.24	46,46,46,46	0
4	MG	A	3004	1/1	0.87	0.24	44,44,44,44	0
3	1NU	B	602	32/32	0.91	0.24	57,64,70,72	0

Continued on next page...

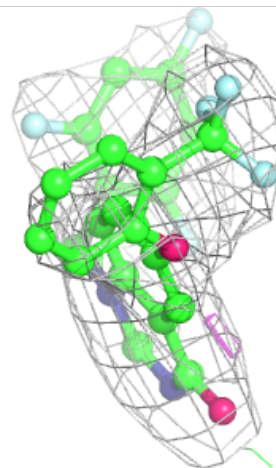
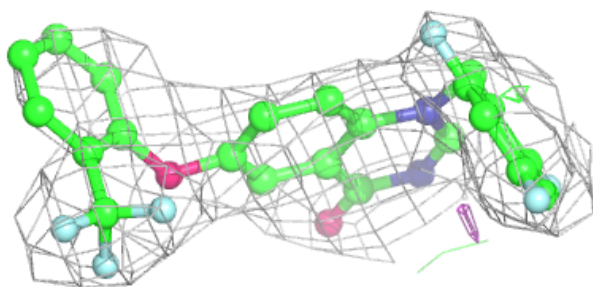
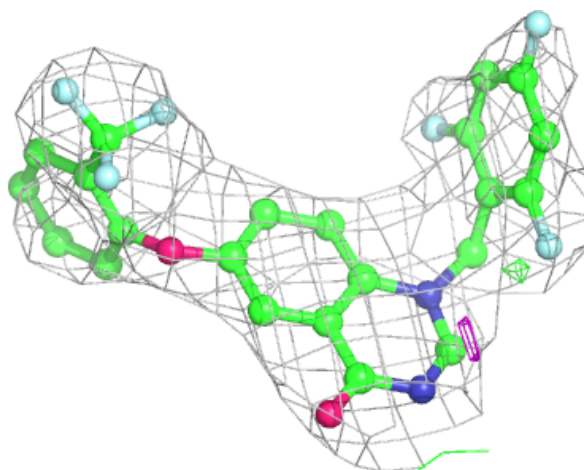
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	GOL	B	601	6/6	0.92	0.26	44,46,47,48	0
2	GOL	A	3001	6/6	0.92	0.22	50,52,53,53	0
4	MG	B	603	1/1	0.93	0.27	46,46,46,46	0
3	1NU	A	3002	32/32	0.96	0.18	37,43,52,54	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.

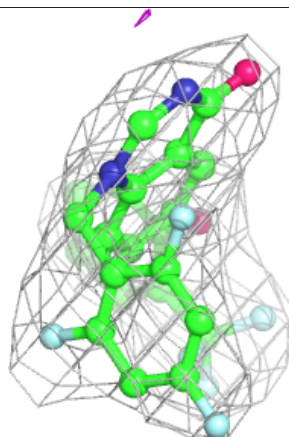
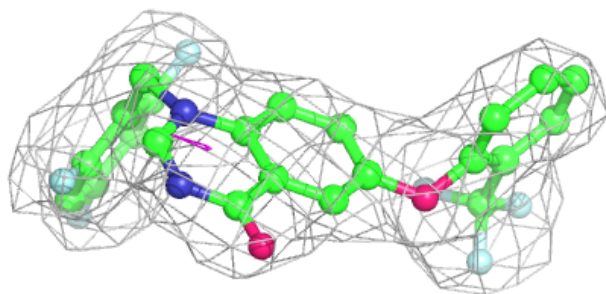
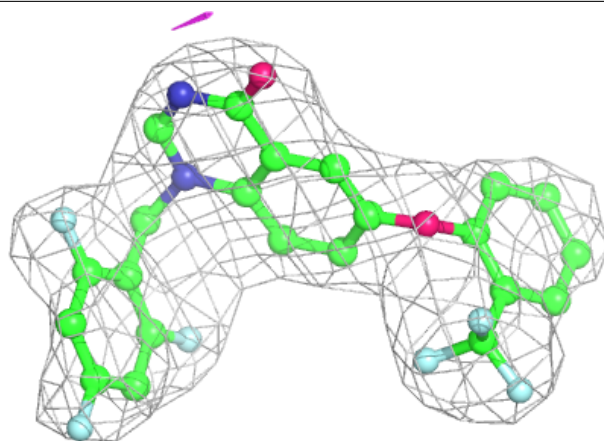
Electron density around 1NU B 602:

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 1NU A 3002:

$2mF_o-DF_c$ (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.