

## wwPDB X-ray Structure Validation Summary Report (i)

#### Oct 3, 2023 – 08:42 PM EDT

:	6P1D
:	Crystal structure of EGFR with mutant-selective dihydrodibenzodiazepinone
	allosteric inhibitor
:	Heppner, D.E.; Eck, M.J.
:	2019-05-19
:	2.40 Å(reported)
	: : :

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

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

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

The following versions of software and data (see references (1)) were used in the production of this report:

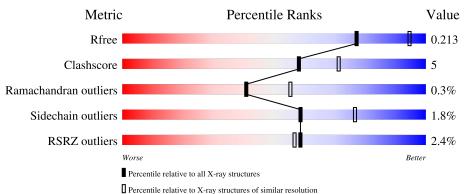
MolProbity	:	4.02b-467 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)		
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

## 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 2.40 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain	
1	А	327	% • 83%	11% 6%
1	В	327	<sup>2%</sup> 78%	13% • 8%
1	С	327	% • 78%	10% • 11%
1	D	327	4% 76%	14% 10%



## 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 10222 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	л	295	Total	С	Ν	0	$\mathbf{S}$	0	0	0
	D	295	2380	1530	404	427	19	0	0	0
1	Λ	309	Total	С	Ν	0	S	0	3	0
	А	309	2519	1616	428	455	20	0		0
1	В	300	Total	С	Ν	0	S	0	2	0
	D	300	2427	1558	412	438	19	0		0
1	С	292	Total	С	Ν	0	S	0	1	0
		292	2353	1515	399	421	18	0		0

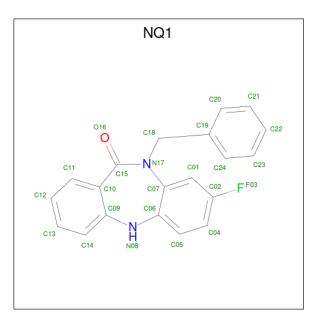
• Molecule 1 is a protein called Epidermal growth factor receptor.

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	790	MET	THR	conflict	UNP P00533
D	948	ARG	VAL	conflict	UNP P00533
A	790	MET	THR	conflict	UNP P00533
А	948	ARG	VAL	conflict	UNP P00533
В	790	MET	THR	conflict	UNP P00533
В	948	ARG	VAL	conflict	UNP P00533
С	790	MET	THR	conflict	UNP P00533
С	948	ARG	VAL	conflict	UNP P00533

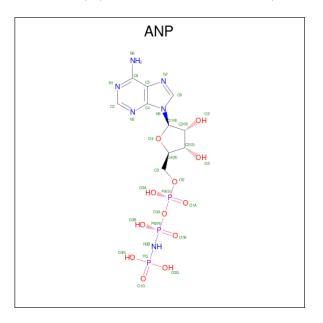
• Molecule 2 is 10-benzyl-8-fluoro-5,10-dihydro-11H-dibenzo[b,e][1,4]diazepin-11-one (three-letter code: NQ1) (formula:  $C_{20}H_{15}FN_2O$ ) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	Л	1	Total	С	F	Ν	0	0	0
	D	1	24	20	1	2	1	0	0
2	В	1	Total	С	F	Ν	Ο	0	0
	D	1	24	20	1	2	1	0	
2	С	1	Total	С	F	Ν	Ο	0	0
	U	1	24	20	1	2	1	0	0

• Molecule 3 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula:  $C_{10}H_{17}N_6O_{12}P_3$ ).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
3	Л	1	Total	С	Ν	Ο	Р	0	0	
5	D	T	31	10	6	12	3	0	0	
3	Δ	1	1 Total C N		Ν	Ο	Р	0	0	
5	D A		31	10	6	12	3	0	0	
3	В	1	Total	С	Ν	Ο	Р	0	0	
5	D	1	31	10	6	12	3	0	0	
3	С	1	Total	С	Ν	Ο	Р	0	0	
5	U	1	31	10	6	12	3	0	0	

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total Mg 1 1	0	0
4	А	1	Total Mg 1 1	0	0
4	В	1	Total Mg 1 1	0	0
4	С	1	Total Mg 1 1	0	0

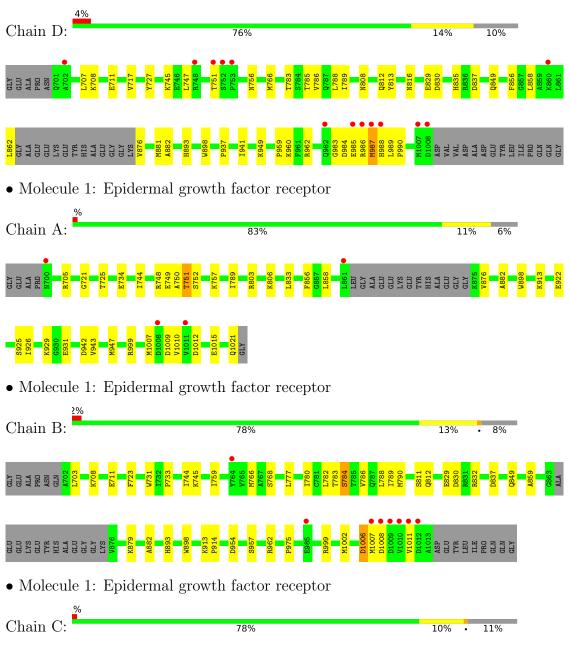
• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	82	TotalO8282	0	0
5	А	108	Total O 108 108	0	0
5	В	84	Total         O           84         84	0	0
5	С	69	Total O 69 69	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: Epidermal growth factor receptor



MET ASP ASP VAL VAL ASP ASP GLU TYR ILE ILEU ILE CLU GLN GLN GLN

# 



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	71.95Å 101.81Å 86.51Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $101.04^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	65.17 - 2.40	Depositor
Resolution (A)	65.21 - 2.40	EDS
% Data completeness	99.2 (65.17-2.40)	Depositor
(in resolution range)	96.5 (65.21-2.40)	EDS
R <sub>merge</sub>	0.10	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.64$ (at $2.40\text{\AA}$ )	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
D D.	0.195 , $0.213$	Depositor
$R, R_{free}$	0.195 , $0.213$	DCC
$R_{free}$ test set	2336 reflections $(4.91%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	29.4	Xtriage
Anisotropy	0.555	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.35 , $43.8$	EDS
L-test for twinning <sup>2</sup>	$ \langle L  \rangle = 0.49, \langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10222	wwPDB-VP
Average B, all atoms $(Å^2)$	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 8.08% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

## 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: NQ1, MG, ANP

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
	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	А	0.30	0/2576	0.49	0/3482
1	В	0.28	0/2482	0.50	0/3356
1	С	0.25	0/2408	0.42	0/3258
1	D	0.31	0/2432	0.50	0/3288
All	All	0.28	0/9898	0.48	0/13384

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	А	2519	0	2566	18	0
1	В	2427	0	2474	27	0
1	С	2353	0	2400	25	0
1	D	2380	0	2430	30	0
2	В	24	0	0	3	0
2	С	24	0	0	2	0
2	D	24	0	0	4	0
3	А	31	0	13	0	0
3	В	31	0	13	1	0

Continued on next page...



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	С	31	0	13	0	0
3	D	31	0	13	1	0
4	А	1	0	0	0	0
4	В	1	0	0	0	0
4	С	1	0	0	0	0
4	D	1	0	0	0	0
5	А	108	0	0	2	0
5	В	84	0	0	2	0
5	С	69	0	0	3	0
5	D	82	0	0	4	0
All	All	10222	0	9922	95	0

Continued from previous page...

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 5.

The worst 5 of 95 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:A:749:GLU:HB2	1:B:832:ARG:HG3	1.70	0.71
1:D:962:ARG:NH2	1:C:718:LEU:O	2.27	0.67
1:B:766:MET:HG3	2:B:1102:NQ1:C14	2.25	0.65
1:D:812:GLN:NE2	1:D:816:ASN:OD1	2.32	0.63
1:D:941:ILE:HD11	1:A:922:GLU:HG2	1.81	0.62

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	Perce	entiles
1	А	308/327~(94%)	301 (98%)	7 (2%)	0	100	100
1	В	298/327~(91%)	286 (96%)	9(3%)	3 (1%)	15	23

Continued on next page...



Mol	Chain	Analysed	Analysed Favoured		Outliers	Perce	entiles
1	$\mathbf{C}$	289/327~(88%)	286~(99%)	3~(1%)	0	100	100
1	D	291/327~(89%)	285~(98%)	5(2%)	1 (0%)	41	55
All	All	1186/1308~(91%)	1158 (98%)	24 (2%)	4 (0%)	41	55

Continued from previous page...

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	В	784	SER
1	В	1007	MET
1	В	1006	ASP
1	D	808	ASN

#### 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	А	280/287~(98%)	271~(97%)	9~(3%)	39	59	
1	В	269/287~(94%)	267~(99%)	2(1%)	84	92	
1	$\mathbf{C}$	260/287~(91%)	256~(98%)	4 (2%)	65	80	
1	D	264/287~(92%)	260~(98%)	4 (2%)	65	80	
All	All	1073/1148 (94%)	1054 (98%)	19 (2%)	59	76	

5 of 19 residues with a non-rotameric side chain are listed below:

Mol	Chain	Res	Type
1	В	1011	VAL
1	С	862	LEU
1	С	988	HIS
1	С	860	LYS
1	А	876	VAL

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such side chains are listed below:



Mol	Chain	Res	Type
1	А	849	GLN
1	В	849	GLN
1	С	812	GLN

#### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

#### 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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	Turne	ype Chain Res Link		Tink	Bo	Bond lengths			Bond angles		
10101	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2	
3	ANP	С	1102	4	29,33,33	1.09	4 (13%)	$31,\!52,\!52$	1.03	3 (9%)	
3	ANP	А	1101	4	29,33,33	1.10	4 (13%)	31,52,52	0.99	3 (9%)	
2	NQ1	D	1101	-	27,27,27	2.51	7 (25%)	37,38,38	1.36	5 (13%)	
2	NQ1	С	1101	-	27,27,27	2.52	8 (29%)	37,38,38	1.43	6 (16%)	
2	NQ1	В	1102	-	27,27,27	2.50	8 (29%)	37,38,38	1.59	7 (18%)	
3	ANP	В	1101	4	29,33,33	1.09	4 (13%)	31,52,52	0.94	3 (9%)	
3	ANP	D	1102	4	29,33,33	1.08	4 (13%)	31,52,52	0.95	2(6%)	

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



6P1I	)
------	---

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ANP	С	1102	4	-	5/14/38/38	0/3/3/3
3	ANP	А	1101	4	-	4/14/38/38	0/3/3/3
2	NQ1	D	1101	-	-	2/4/4/4	0/4/4/4
2	NQ1	С	1101	-	-	2/4/4/4	0/4/4/4
2	NQ1	В	1102	-	-	1/4/4/4	0/4/4/4
3	ANP	В	1101	4	-	4/14/38/38	0/3/3/3
3	ANP	D	1102	4	-	4/14/38/38	0/3/3/3

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
2	D	1101	NQ1	C15-N17	6.71	1.45	1.37
2	С	1101	NQ1	C15-N17	6.69	1.45	1.37
2	В	1102	NQ1	C15-N17	6.56	1.45	1.37
2	В	1102	NQ1	C06-N08	5.47	1.48	1.41
2	В	1102	NQ1	C09-N08	5.31	1.48	1.41

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$\mathbf{Ideal}(^{o})$
2	С	1101	NQ1	C07-N17-C15	4.21	129.45	124.96
2	D	1101	NQ1	C07-N17-C15	4.01	129.24	124.96
2	В	1102	NQ1	C07-N17-C15	4.00	129.23	124.96
2	В	1102	NQ1	C19-C18-N17	3.98	120.11	113.39
2	В	1102	NQ1	C18-N17-C07	-3.76	114.31	118.28

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	1102	ANP	PG-N3B-PB-O1B
3	D	1102	ANP	PA-O3A-PB-O1B
3	D	1102	ANP	PA-O3A-PB-O2B
3	А	1101	ANP	PG-N3B-PB-O1B
3	А	1101	ANP	PA-O3A-PB-O1B

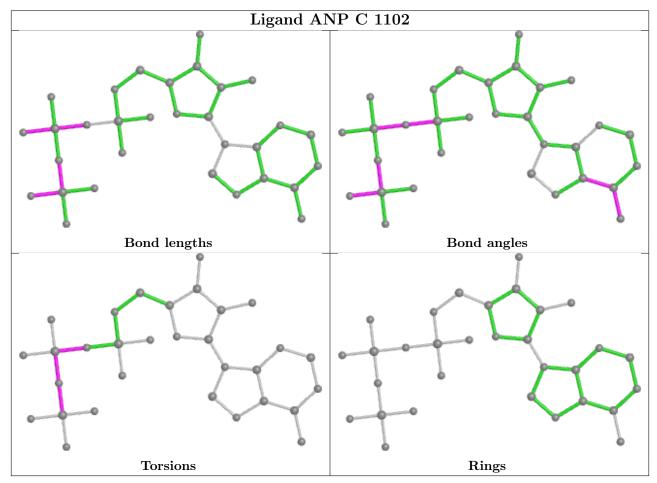
There are no ring outliers.

5 monomers are involved in 11 short contacts:



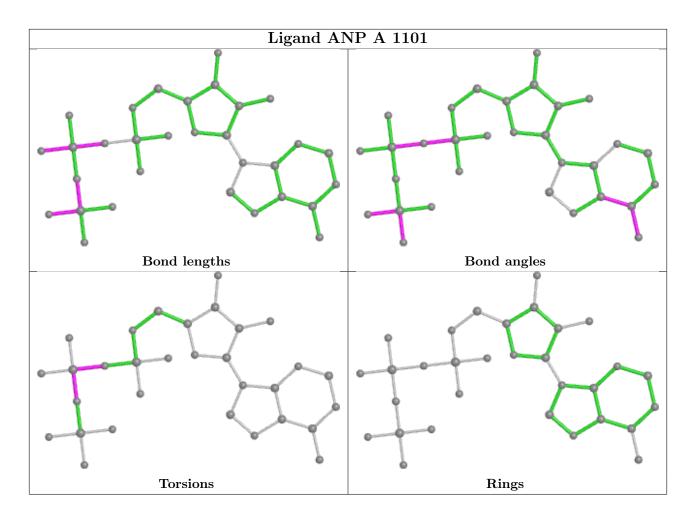
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1101	NQ1	4	0
2	С	1101	NQ1	2	0
2	В	1102	NQ1	3	0
3	В	1101	ANP	1	0
3	D	1102	ANP	1	0

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 then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and similar 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.

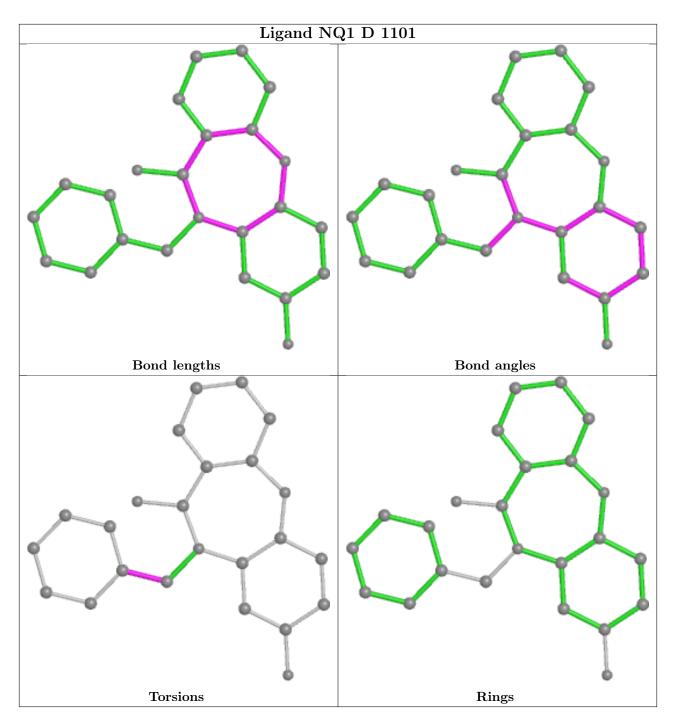




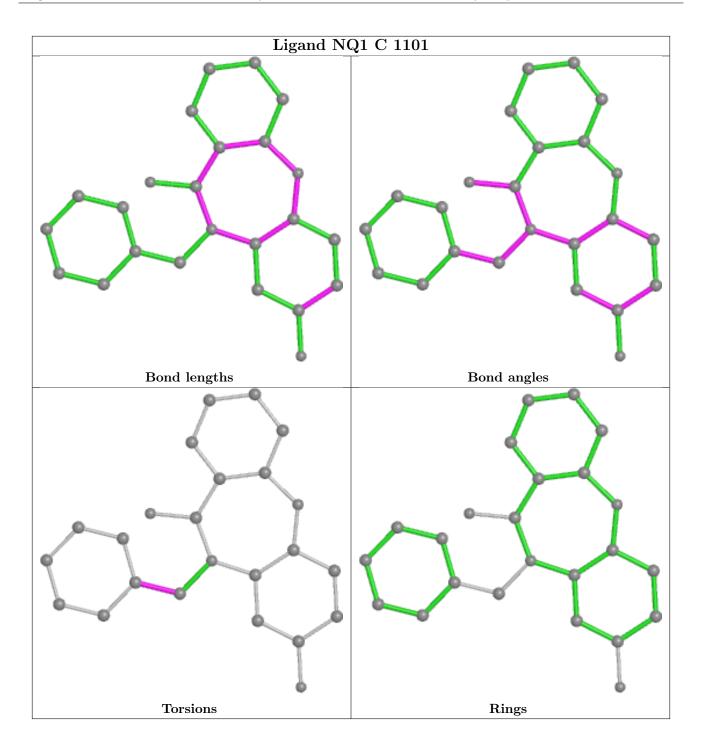




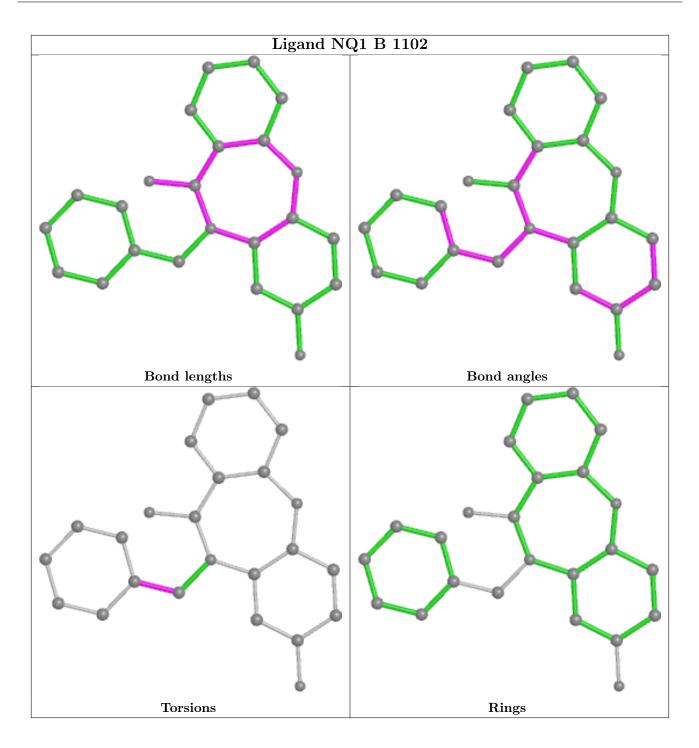






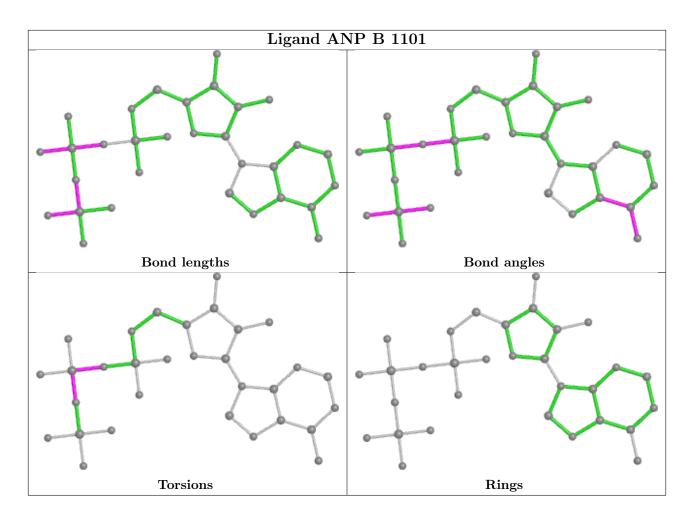






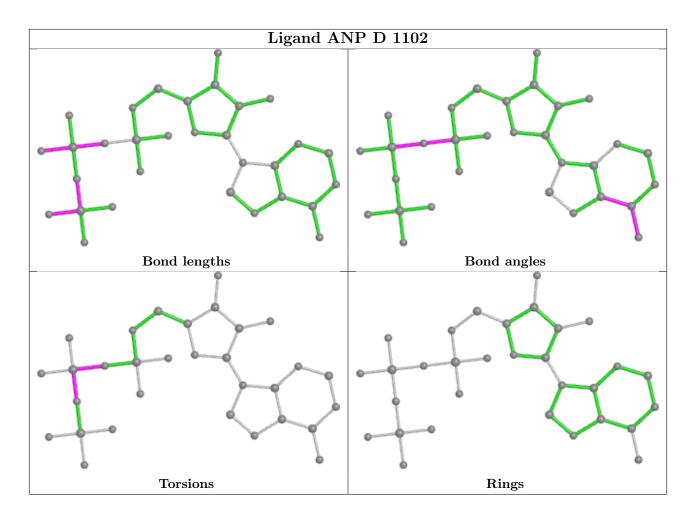












## 5.7 Other polymers (i)

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,  $95^{th}$  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	А	309/327~(94%)	-0.09	4 (1%) 77 75	20, 30, 53, 85	0
1	В	300/327~(91%)	-0.03	8 (2%) 54 52	18, 34, 54, 124	0
1	С	292/327~(89%)	-0.02	4 (1%) 75 73	20, 35, 56, 80	0
1	D	295/327~(90%)	-0.01	13 (4%) 34 33	19, 33, 63, 88	0
All	All	1196/1308 (91%)	-0.04	29 (2%) 59 57	18, 33, 57, 124	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type	RSRZ
1	В	1009	ASP	8.5
1	В	1010	VAL	6.2
1	А	1008	ASP	5.3
1	В	1008	ASP	5.2
1	В	1007	MET	5.0

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

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

### 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

### 6.4 Ligands (i)

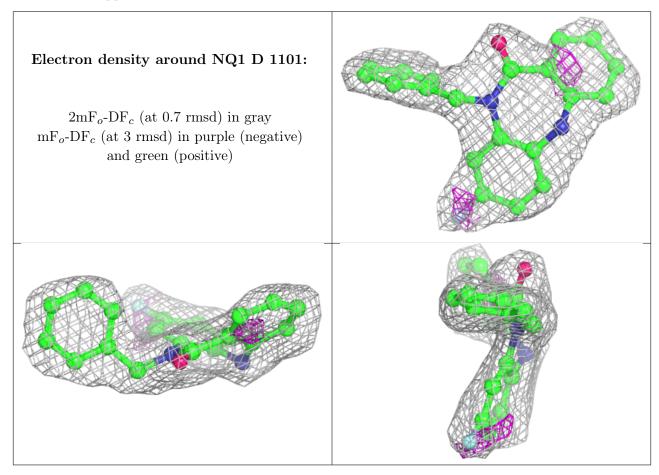
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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



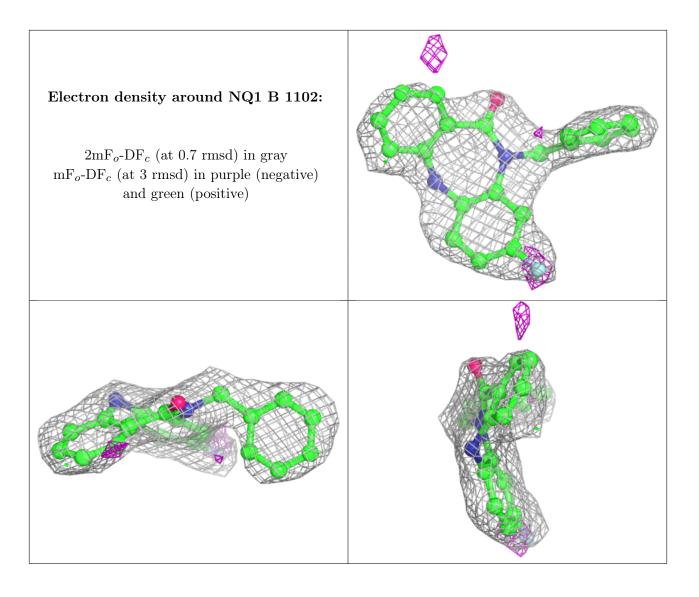
6P1	۱D
-----	----

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	MG	С	1103	1/1	0.89	0.11	30,30,30,30	0
4	MG	А	1102	1/1	0.90	0.08	28,28,28,28	0
2	NQ1	D	1101	24/24	0.92	0.23	27,31,39,41	0
2	NQ1	В	1102	24/24	0.92	0.26	$29,\!34,\!44,\!45$	0
4	MG	D	1103	1/1	0.94	0.07	21,21,21,21	0
2	NQ1	С	1101	24/24	0.95	0.17	23,29,38,39	0
3	ANP	С	1102	31/31	0.96	0.13	$25,\!32,\!37,\!41$	0
4	MG	В	1103	1/1	0.96	0.06	29,29,29,29	0
3	ANP	А	1101	31/31	0.96	0.12	$24,\!28,\!35,\!37$	0
3	ANP	В	1101	31/31	0.97	0.14	19,29,36,41	0
3	ANP	D	1102	31/31	0.97	0.14	19,25,32,35	0

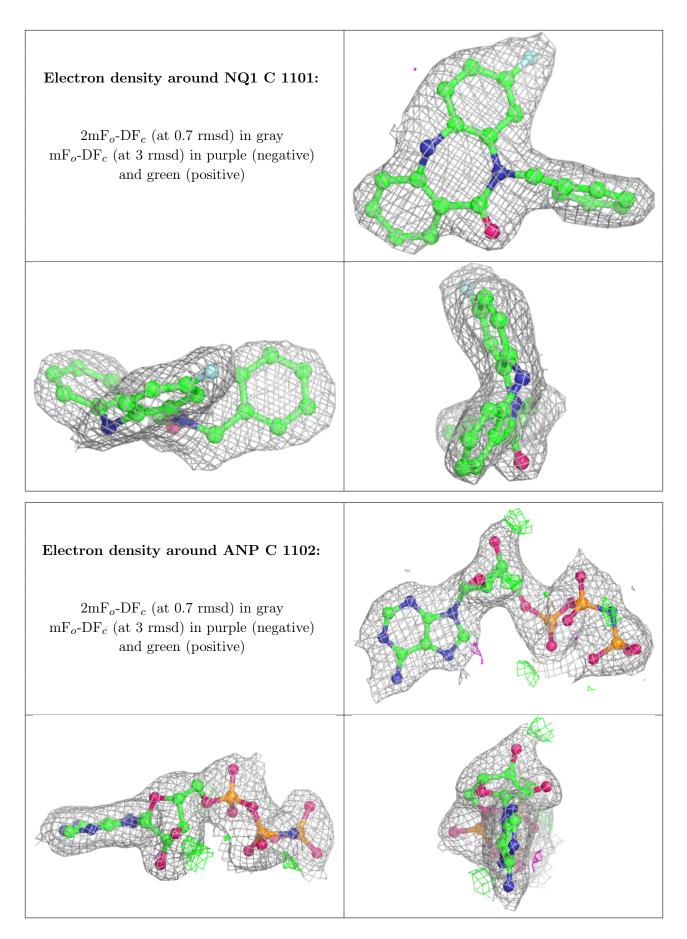
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



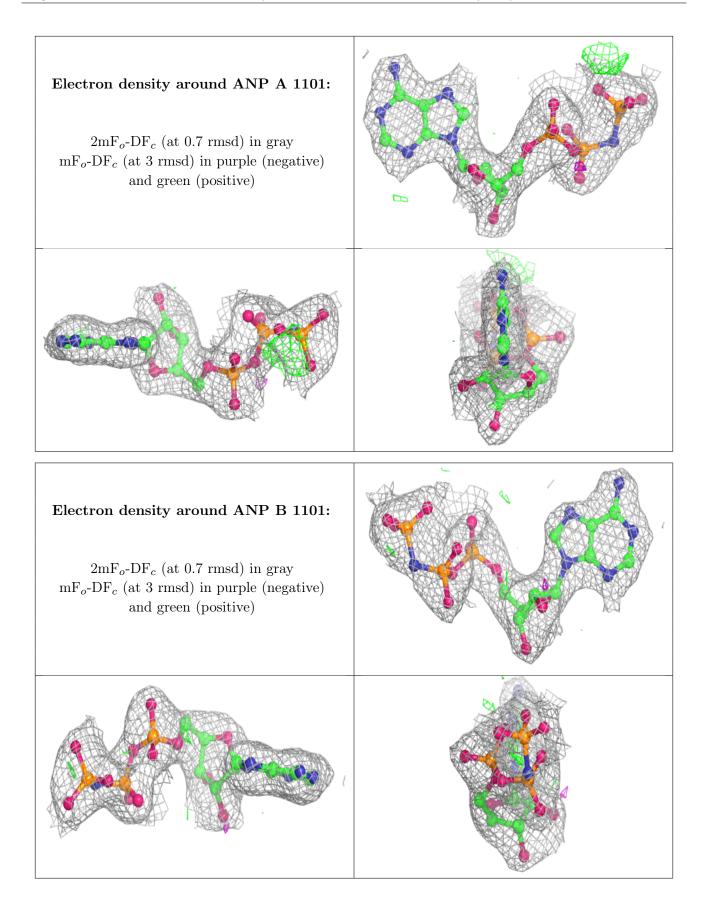




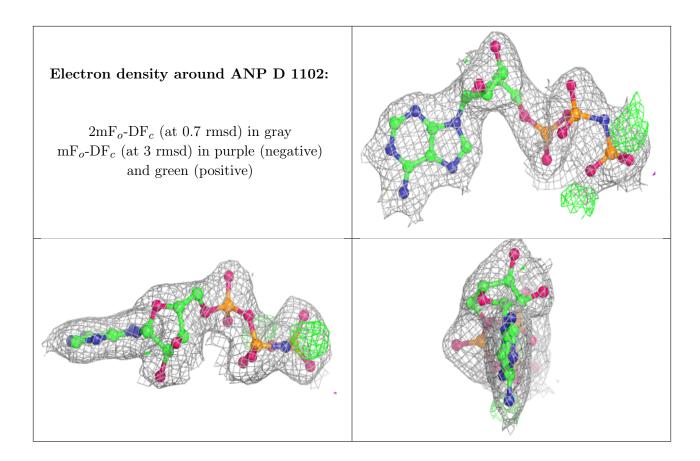












## 6.5 Other polymers (i)

There are no such residues in this entry.

