



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 13, 2023 – 08:43 AM EDT

PDB ID : 4RIY
Title : Crystal structure of an EGFR/HER3 kinase domain heterodimer containing the cancer-associated HER3-E909G mutation
Authors : Littlefield, P.; Liu, L.; Jura, N.
Deposited on : 2014-10-07
Resolution : 2.98 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 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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35.1
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

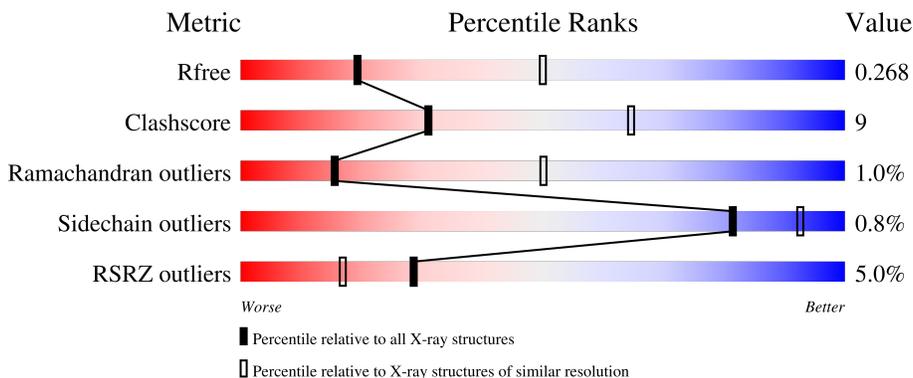
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.98 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	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

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 $\leq 5\%$. 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	326	 0% 73% 10% 17%
1	C	326	 2% 73% 11% 16%
2	B	345	 7% 69% 14% 15%
2	D	345	 7% 64% 20% 14%

2 Entry composition [i](#)

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 Receptor tyrosine-protein kinase erbB-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	269	2118	1371	369	365	13	0	0	0
1	C	275	2172	1407	373	379	13	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	676	GLY	-	expression tag	UNP P21860
A	677	ALA	-	expression tag	UNP P21860
A	678	GLY	-	expression tag	UNP P21860
A	909	GLY	GLU	engineered mutation	UNP P21860
C	676	GLY	-	expression tag	UNP P21860
C	677	ALA	-	expression tag	UNP P21860
C	678	GLY	-	expression tag	UNP P21860
C	909	GLY	GLU	engineered mutation	UNP P21860

- Molecule 2 is a protein called Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	293	2343	1510	397	421	15	0	0	0
2	D	297	2369	1525	402	427	15	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

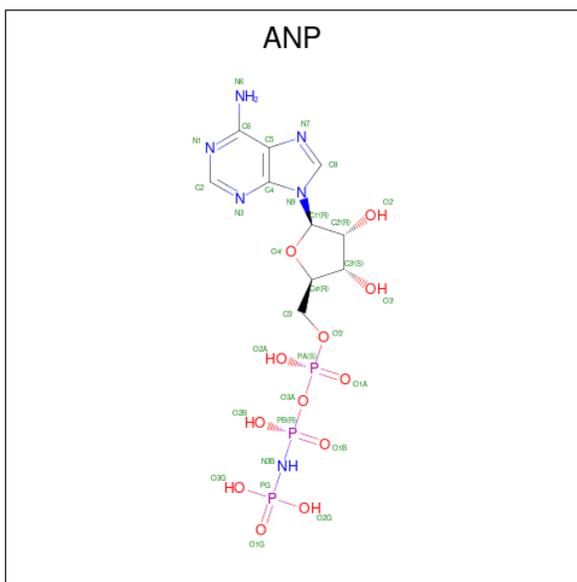
Chain	Residue	Modelled	Actual	Comment	Reference
B	654	GLY	-	expression tag	UNP P00533
B	655	ALA	-	expression tag	UNP P00533
B	656	MET	-	expression tag	UNP P00533
B	657	GLY	-	expression tag	UNP P00533

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	924	ARG	VAL	engineered mutation	UNP P00533
B	973	ALA	PHE	engineered mutation	UNP P00533
B	977	ALA	LEU	engineered mutation	UNP P00533
D	654	GLY	-	expression tag	UNP P00533
D	655	ALA	-	expression tag	UNP P00533
D	656	MET	-	expression tag	UNP P00533
D	657	GLY	-	expression tag	UNP P00533
D	924	ARG	VAL	engineered mutation	UNP P00533
D	973	ALA	PHE	engineered mutation	UNP P00533
D	977	ALA	LEU	engineered mutation	UNP P00533

- 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
			Total	C	N	O	P		
3	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	C	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

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

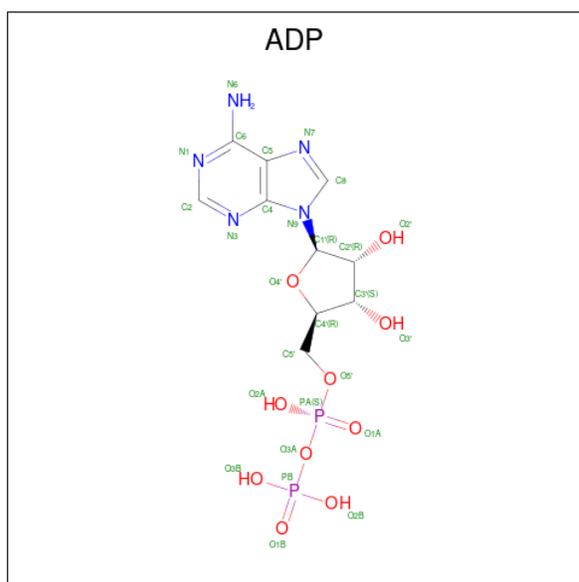
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Mg	0	0
			1	1		

Continued on next page...

Continued from previous page...

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

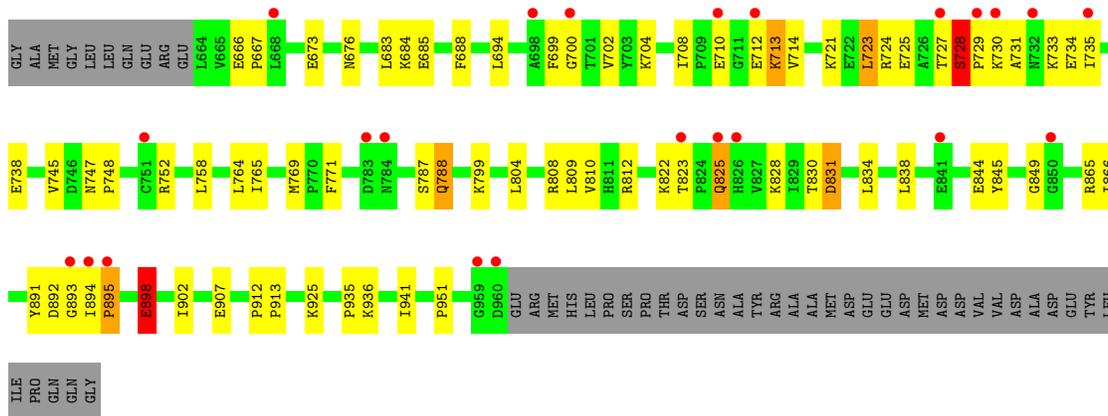
- Molecule 5 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C N O P 27 10 5 10 2	0	0
5	D	1	Total C N O P 27 10 5 10 2	0	0

D960	GLU
	ARG
	MET
	HIS
	LEU
	PRO
	SER
	PRO
	THR
	ASP
	SER
	ASN
	ALA
	TYR
	ARG
	ALA
	ALA
	ALA
	MET
	ASP
	GLU
	GLU
	ASP
	MET
	ASP
	ASP
	VAL
	VAL
	ASP
	ALA
	ASP
	GLU
	TYR
	LEU
	ILE
	ILE
	PRO
	GLN
	GLN
	GLY

• Molecule 2: Epidermal growth factor receptor



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	65.33Å 155.14Å 87.30Å 90.00° 110.93° 90.00°	Depositor
Resolution (Å)	56.20 – 2.98 56.79 – 2.98	Depositor EDS
% Data completeness (in resolution range)	97.6 (56.20-2.98) 93.2 (56.79-2.98)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.61 (at 2.96Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.216 , 0.267 0.220 , 0.268	Depositor DCC
R_{free} test set	3254 reflections (10.06%)	wwPDB-VP
Wilson B-factor (Å ²)	38.7	Xtrriage
Anisotropy	0.674	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 54.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.035 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	9122	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.52% 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: MG, ADP, 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	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.26	0/2169	0.45	0/2940
1	C	0.24	0/2224	0.46	0/3017
2	B	0.33	1/2395 (0.0%)	0.56	0/3243
2	D	0.41	1/2422 (0.0%)	0.64	1/3279 (0.0%)
All	All	0.32	2/9210 (0.0%)	0.54	1/12479 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1
2	B	0	1
2	D	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	898	GLU	CG-CD	-8.07	1.39	1.51
2	B	898	GLU	CD-OE2	-5.26	1.19	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	723	LEU	CB-CG-CD2	-13.27	88.44	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	751	CYS	Peptide
1	C	846	GLN	Peptide
2	D	728	SER	Peptide

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	2118	0	2171	26	0
1	C	2172	0	2222	28	0
2	B	2343	0	2402	43	0
2	D	2369	0	2428	75	0
3	A	31	0	13	0	0
3	C	31	0	13	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	B	27	0	12	1	0
5	D	27	0	12	2	0
All	All	9122	0	9273	167	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 9.

The worst 5 of 167 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)
2:B:751:CYS:HB2	2:B:828:LYS:HD3	1.57	0.84
2:D:734:GLU:HG2	2:D:735:ILE:N	1.91	0.84
2:D:723:LEU:HD21	2:D:734:GLU:OE2	1.77	0.83
1:C:945:GLU:HG2	1:C:948:ARG:NH1	1.96	0.80
2:B:751:CYS:SG	2:B:752:ARG:N	2.56	0.78

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	265/326 (81%)	259 (98%)	5 (2%)	1 (0%)	34	70
1	C	271/326 (83%)	265 (98%)	6 (2%)	0	100	100
2	B	289/345 (84%)	271 (94%)	13 (4%)	5 (2%)	9	36
2	D	295/345 (86%)	276 (94%)	14 (5%)	5 (2%)	9	36
All	All	1120/1342 (84%)	1071 (96%)	38 (3%)	11 (1%)	15	50

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	747	LEU
2	B	897	SER
2	B	898	GLU
2	D	713	LYS
2	B	672	GLY

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	228/277 (82%)	227 (100%)	1 (0%)	91	97
1	C	235/277 (85%)	235 (100%)	0	100	100
2	B	258/300 (86%)	255 (99%)	3 (1%)	71	89
2	D	260/300 (87%)	256 (98%)	4 (2%)	65	86
All	All	981/1154 (85%)	973 (99%)	8 (1%)	81	92

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	898	GLU
2	D	825	GLN
2	D	728	SER
2	B	843	LYS
2	D	788	GLN

Sometimes 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 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$
5	ADP	D	1001	4	24,29,29	0.97	1 (4%)	29,45,45	1.50	5 (17%)
5	ADP	B	1001	4	24,29,29	0.95	1 (4%)	29,45,45	1.43	4 (13%)
3	ANP	C	1101	4	29,33,33	1.15	5 (17%)	31,52,52	0.98	1 (3%)
3	ANP	A	1101	4	29,33,33	1.08	4 (13%)	31,52,52	1.13	4 (12%)

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
5	ADP	D	1001	4	-	1/12/32/32	0/3/3/3
5	ADP	B	1001	4	-	8/12/32/32	0/3/3/3
3	ANP	C	1101	4	-	4/14/38/38	0/3/3/3
3	ANP	A	1101	4	-	7/14/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1101	ANP	PG-N3B	2.69	1.70	1.63
5	D	1001	ADP	C5-C4	2.63	1.47	1.40
3	A	1101	ANP	PG-N3B	2.60	1.70	1.63
5	B	1001	ADP	C5-C4	2.51	1.47	1.40
3	C	1101	ANP	PG-O1G	2.47	1.50	1.46

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1001	ADP	C3'-C2'-C1'	3.62	106.43	100.98
5	D	1001	ADP	PA-O3A-PB	-3.58	120.56	132.83
5	B	1001	ADP	N3-C2-N1	-3.25	123.60	128.68
5	D	1001	ADP	N3-C2-N1	-3.10	123.84	128.68
5	D	1001	ADP	C3'-C2'-C1'	2.77	105.15	100.98

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

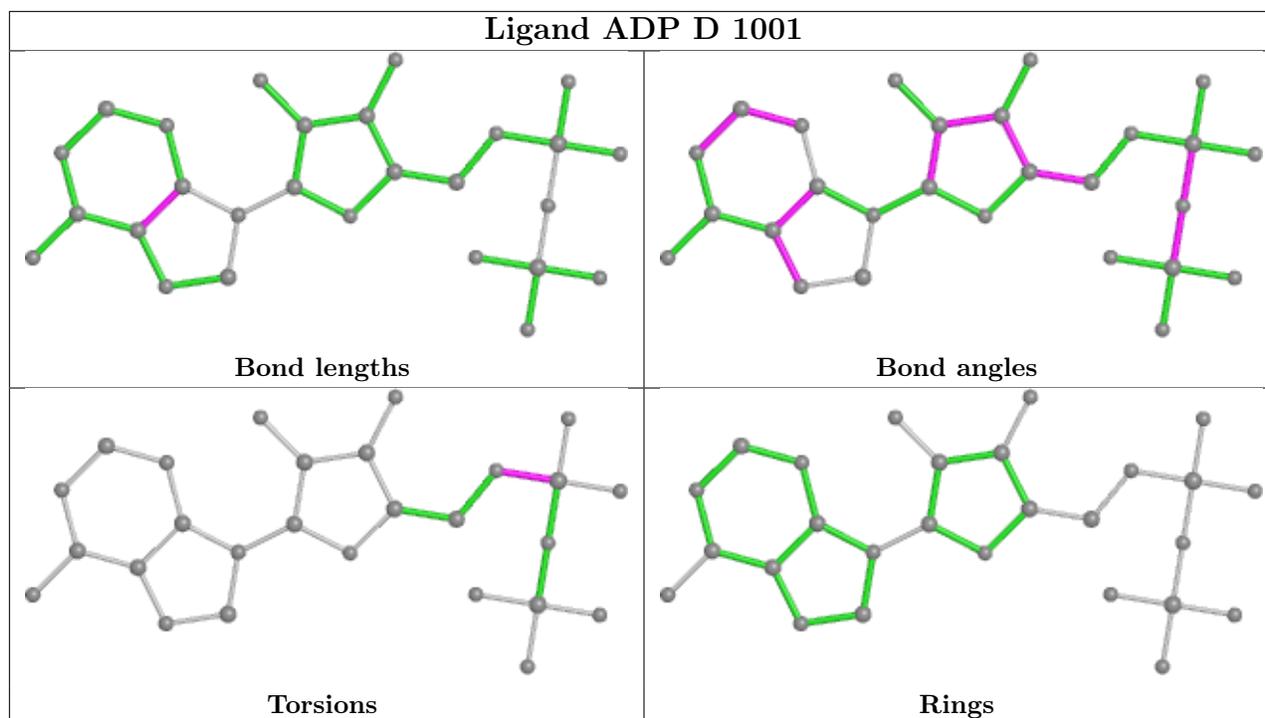
Mol	Chain	Res	Type	Atoms
3	A	1101	ANP	PA-O3A-PB-O1B
3	A	1101	ANP	PA-O3A-PB-O2B
3	A	1101	ANP	O4'-C4'-C5'-O5'
3	C	1101	ANP	PG-N3B-PB-O1B
3	C	1101	ANP	PA-O3A-PB-O1B

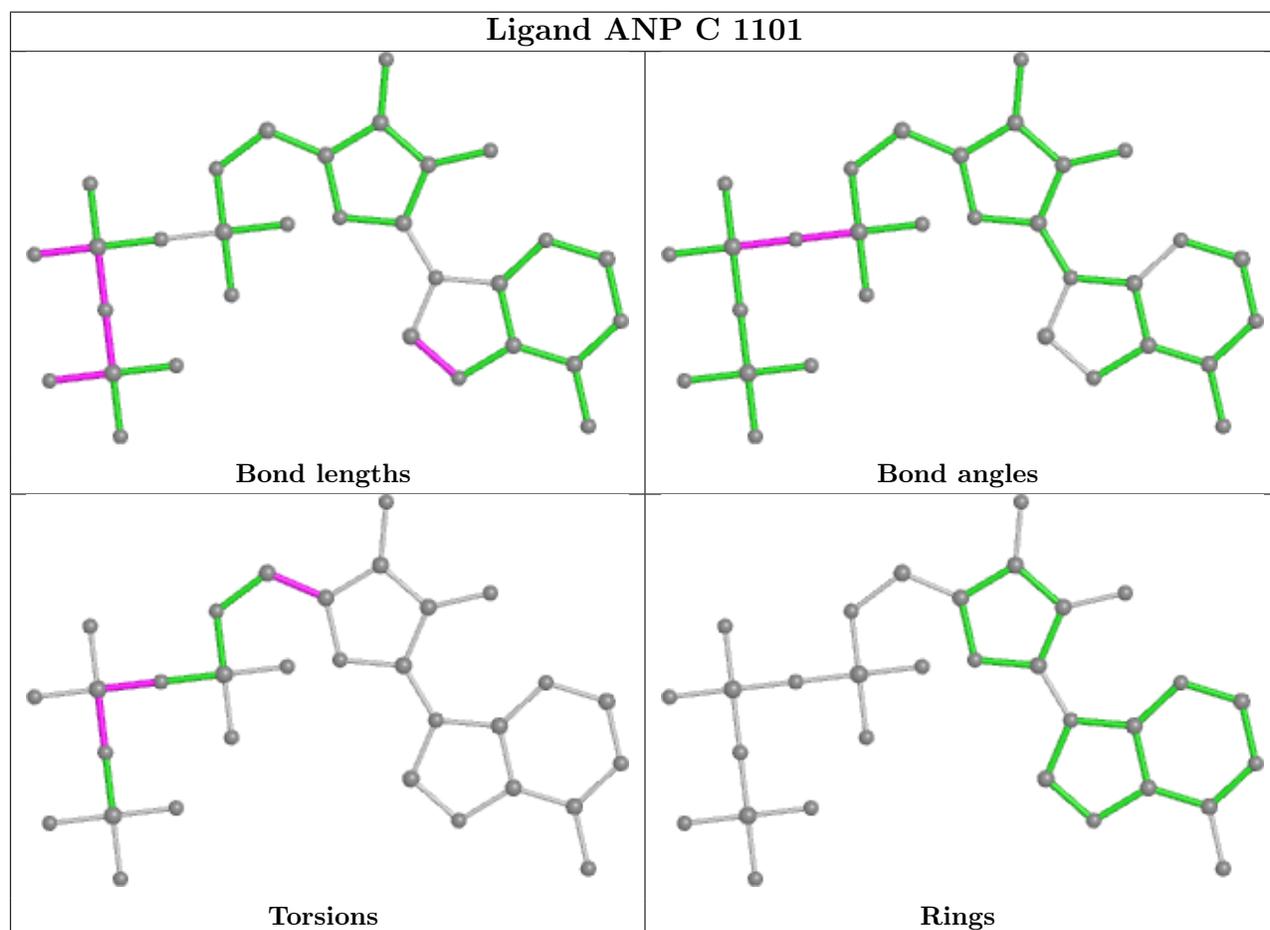
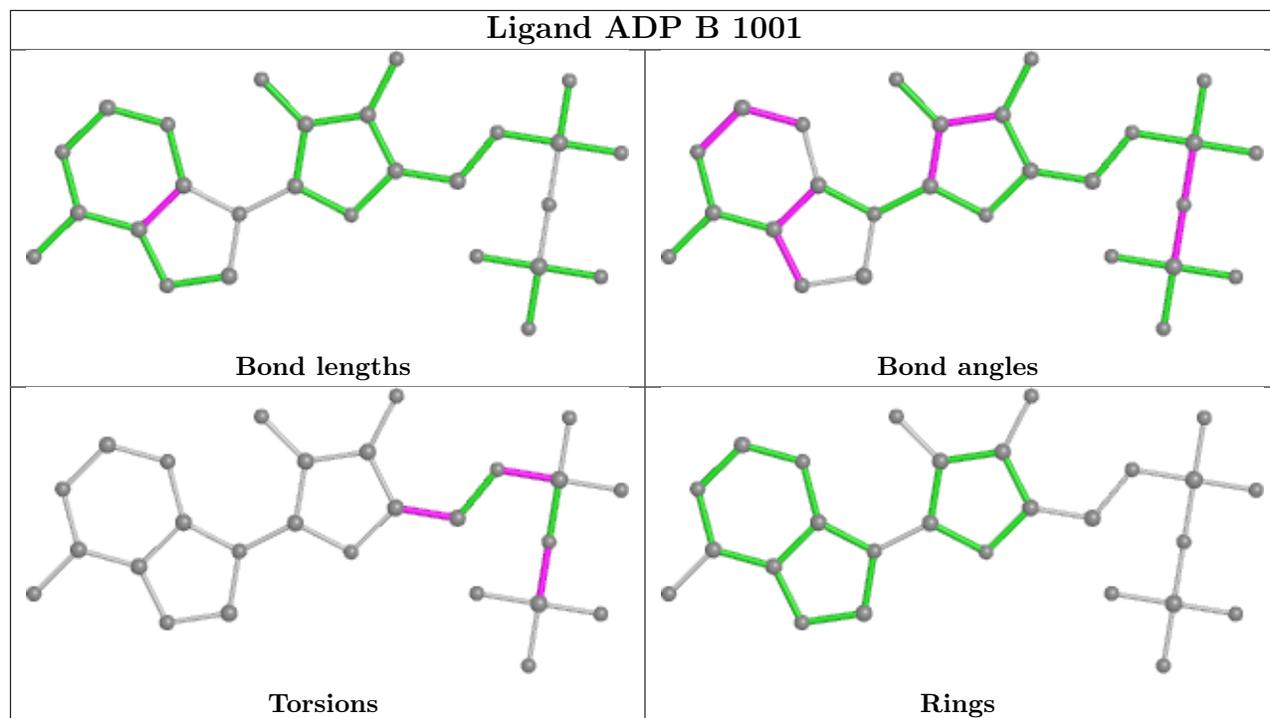
There are no ring outliers.

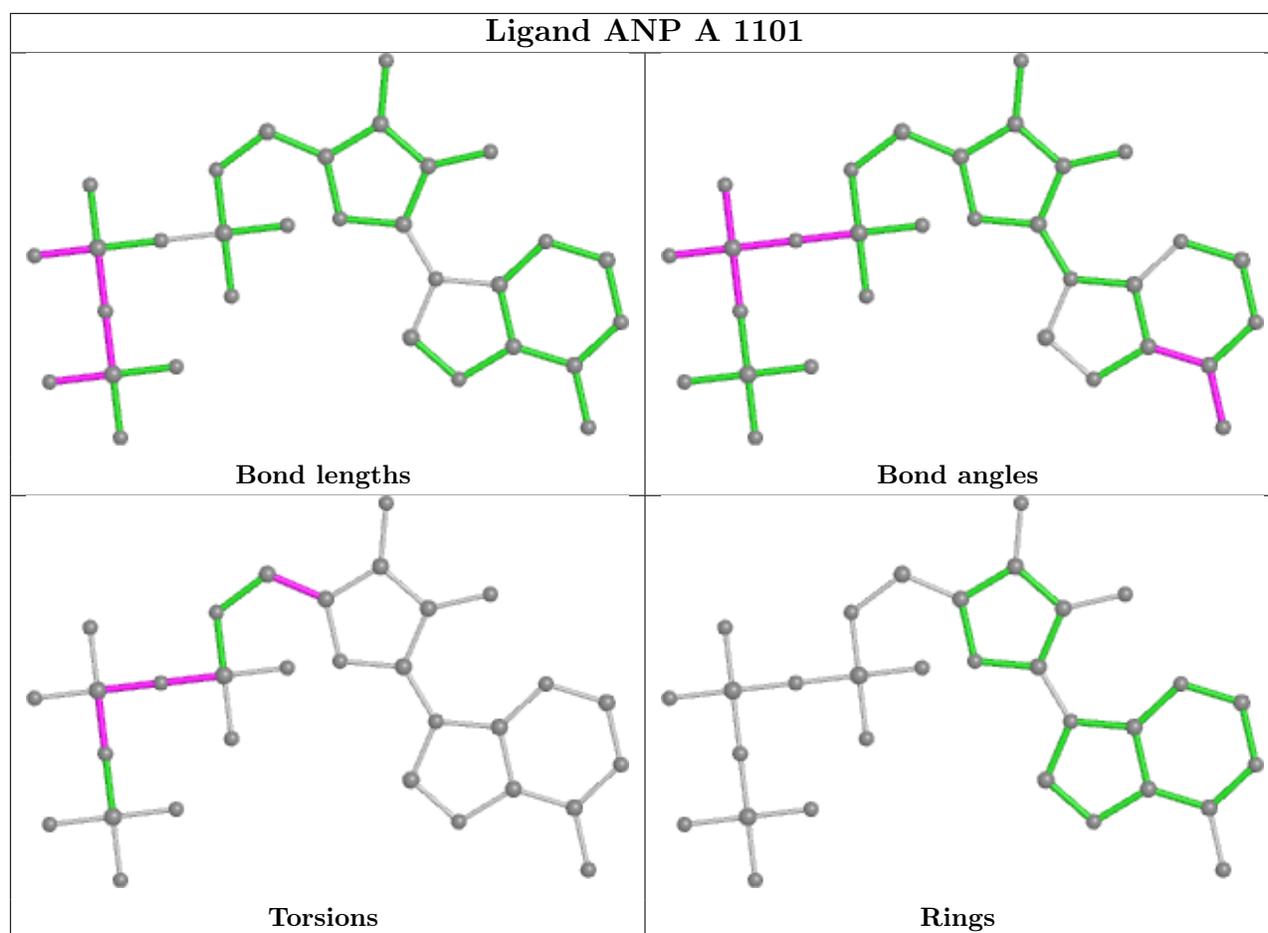
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	1001	ADP	2	0
5	B	1001	ADP	1	0
3	C	1101	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 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	269/326 (82%)	0.30	2 (0%) 87 74	18, 38, 67, 110	0
1	C	275/326 (84%)	0.28	8 (2%) 51 32	17, 37, 72, 107	0
2	B	293/345 (84%)	0.73	24 (8%) 11 5	20, 42, 98, 123	0
2	D	297/345 (86%)	0.68	23 (7%) 13 6	20, 44, 93, 128	0
All	All	1134/1342 (84%)	0.51	57 (5%) 28 17	17, 41, 88, 128	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	714	VAL	6.8
2	B	713	LYS	5.3
2	B	729	PRO	5.2
2	D	729	PRO	4.9
2	D	710	GLU	4.3

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](#)

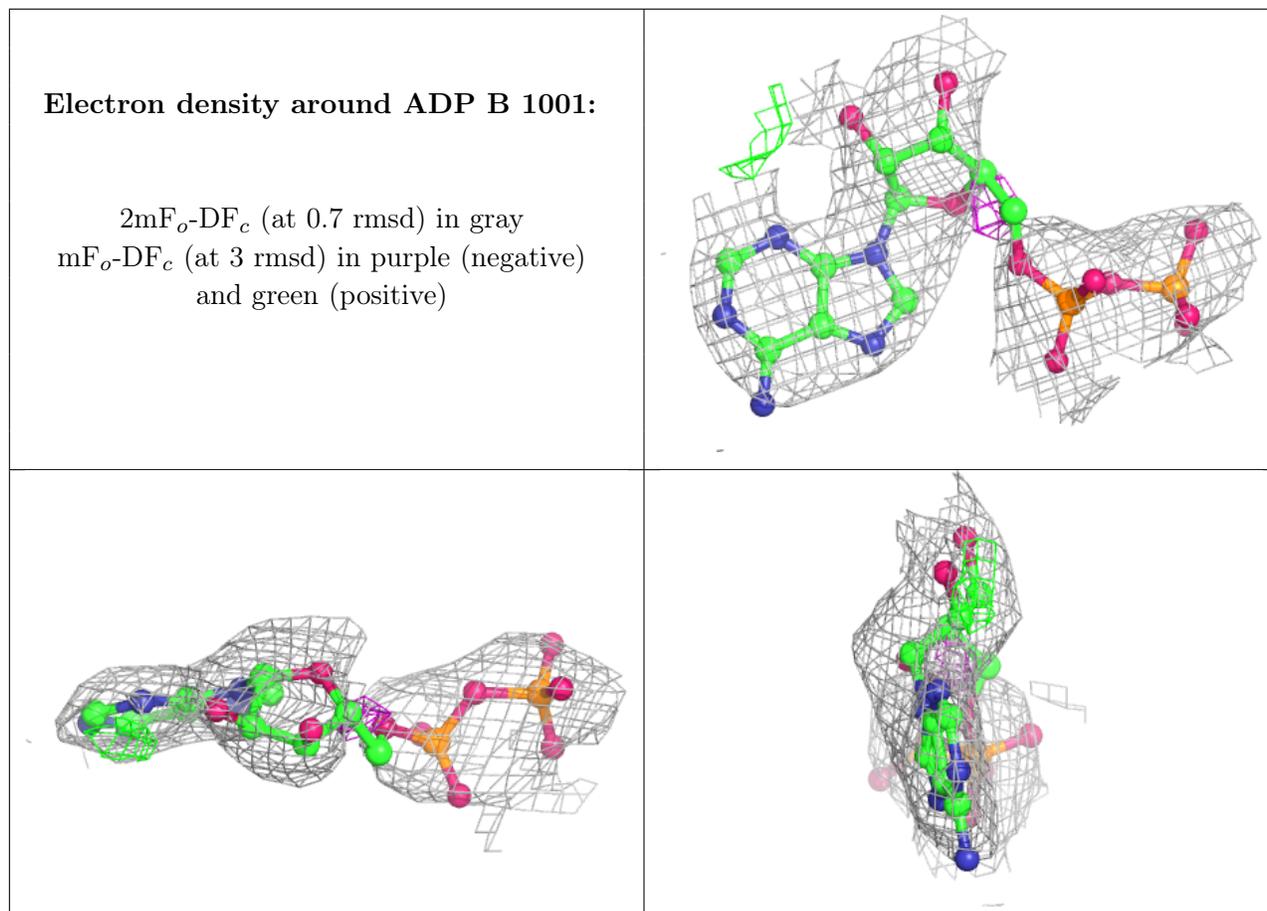
There are no monosaccharides in this entry.

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.

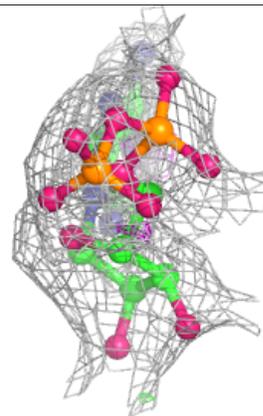
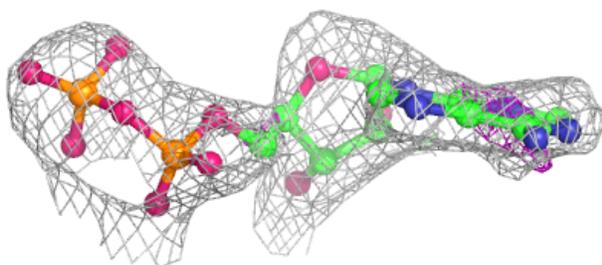
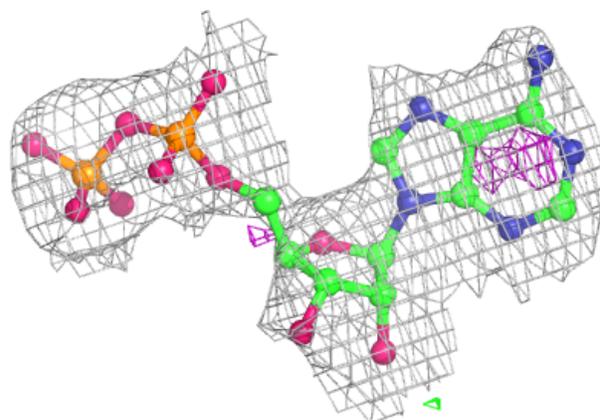
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	ADP	B	1001	27/27	0.73	0.32	62,74,145,146	0
4	MG	B	1002	1/1	0.77	0.34	60,60,60,60	0
5	ADP	D	1001	27/27	0.79	0.26	50,62,144,145	0
4	MG	D	1002	1/1	0.82	0.17	59,59,59,59	0
4	MG	A	1102	1/1	0.88	0.10	23,23,23,23	0
4	MG	C	1102	1/1	0.90	0.21	17,17,17,17	0
3	ANP	A	1101	31/31	0.95	0.20	25,42,62,89	0
3	ANP	C	1101	31/31	0.96	0.16	17,31,47,79	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.

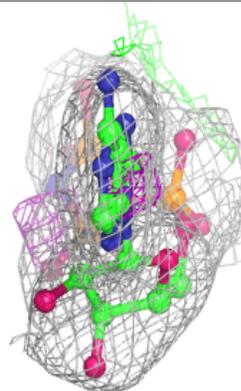
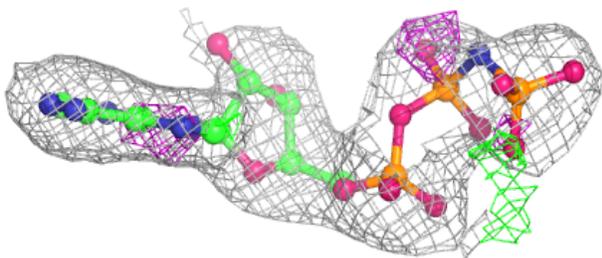
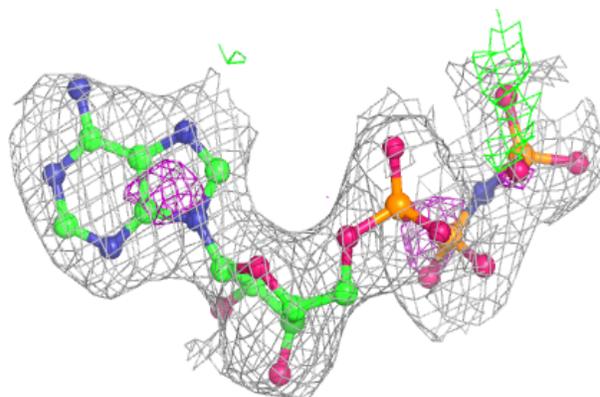


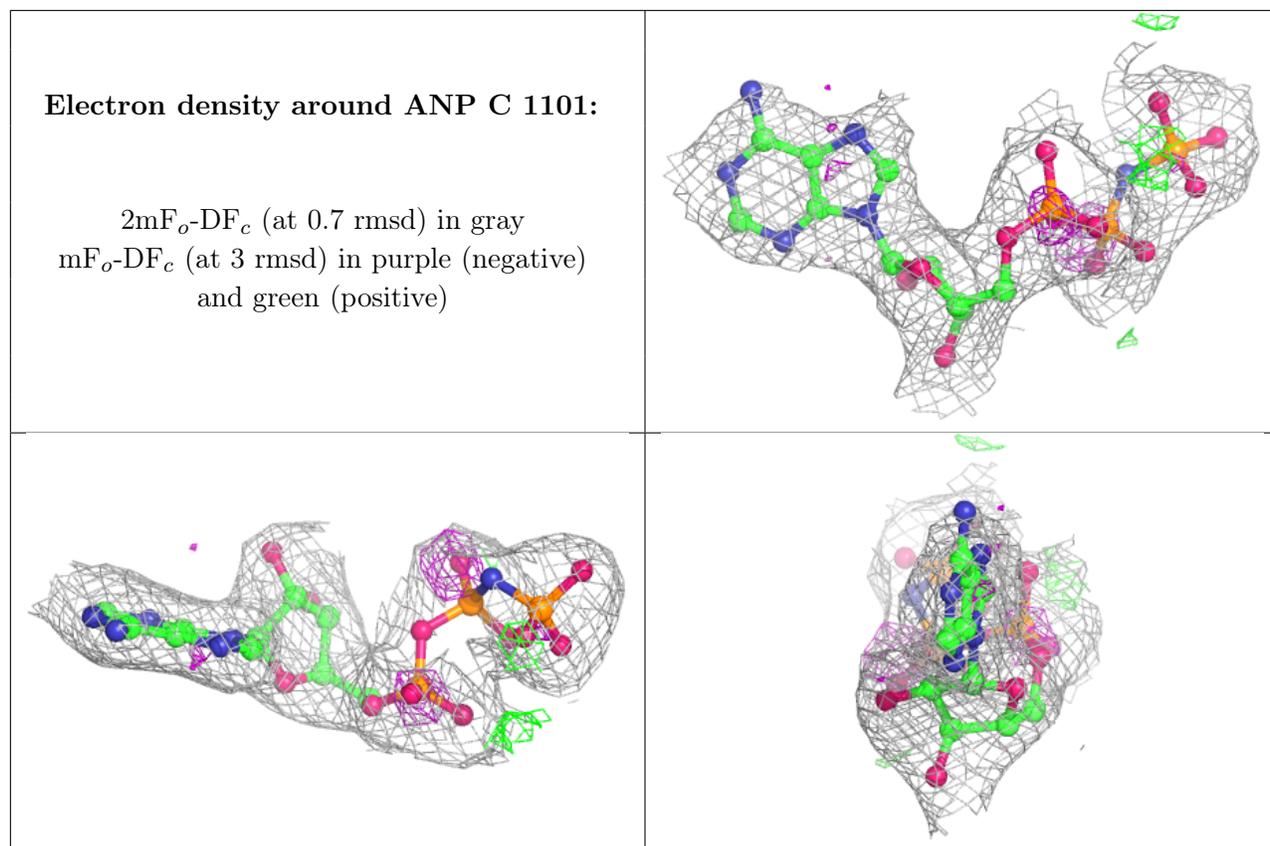
Electron density around ADP D 1001:

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

**Electron density around ANP A 1101:**

$2mF_o-DF_c$ (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.