WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 26, 2023 – 09:01 PM EDT

PDB ID	:	3GPL
Title	:	Crystal structure of the ternary complex of RecD2 with DNA and ADPNP
Authors	:	Saikrishnan, K.; Cook, N.; Wigley, D.B.
Deposited on	:	2009-03-23
Resolution	:	2.50 Å(reported)

The following versions of software and data (see references (i)) 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
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   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 2.50 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.
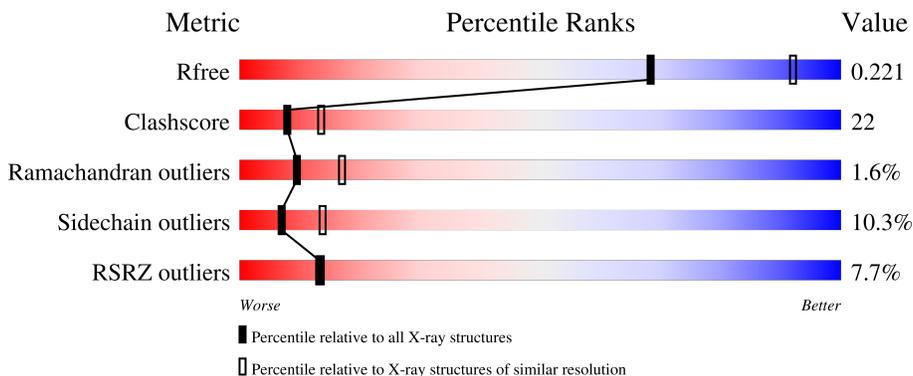
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	A	574	6% / 58% 26% 6% 10%
1	B	574	8% / 57% 27% 6% 10%
2	X	8	12% 75% 12%
2	Y	8	12% 75% 12%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 8322 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 Exodeoxyribonuclease V, subunit RecD, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	516	Total 3899	C 2447	N 720	O 722	S 10	0	2	0
1	B	516	Total 3899	C 2447	N 720	O 722	S 10	0	2	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	150	MET	-	expression tag	UNP Q9RT63
A	716	LEU	-	expression tag	UNP Q9RT63
A	717	GLU	-	expression tag	UNP Q9RT63
A	718	HIS	-	expression tag	UNP Q9RT63
A	719	HIS	-	expression tag	UNP Q9RT63
A	720	HIS	-	expression tag	UNP Q9RT63
A	721	HIS	-	expression tag	UNP Q9RT63
A	722	HIS	-	expression tag	UNP Q9RT63
A	723	HIS	-	expression tag	UNP Q9RT63
B	150	MET	-	expression tag	UNP Q9RT63
B	716	LEU	-	expression tag	UNP Q9RT63
B	717	GLU	-	expression tag	UNP Q9RT63
B	718	HIS	-	expression tag	UNP Q9RT63
B	719	HIS	-	expression tag	UNP Q9RT63
B	720	HIS	-	expression tag	UNP Q9RT63
B	721	HIS	-	expression tag	UNP Q9RT63
B	722	HIS	-	expression tag	UNP Q9RT63
B	723	HIS	-	expression tag	UNP Q9RT63

- Molecule 2 is a DNA chain called 5'-D(*T*TP*TP*TP*TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	7	Total 137	C 70	N 14	O 47	P 6	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Y	7	Total 137	C 70	N 14	O 47	P 6	1	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Mg 1	0	0
3	B	1	Total 1	Mg 1	0	0

- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).
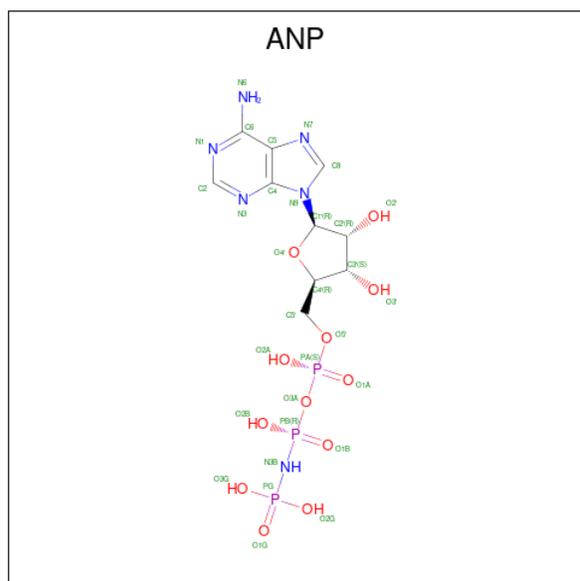
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 31	C 10	N 6	O 12	P 3	0	0
4	B	1	Total 31	C 10	N 6	O 12	P 3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	94	Total 94	O 94	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	76	Total 76	O 76	0	0
5	X	9	Total 9	O 9	0	0
5	Y	7	Total 7	O 7	0	0

3 Residue-property plots ⓘ

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: Exodeoxyribonuclease V, subunit RecD, putative

Chain A:

• Molecule 1: Exodeoxyribonuclease V, subunit RecD, putative

Chain B:

• Molecule 2: 5'-D(*T*TP*TP*TP*TP*TP*TP*T)-3'

Chain X:

• Molecule 2: 5'-D(*T*TP*TP*TP*TP*TP*TP*T)-3'

Chain Y:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	66.12Å　130.36Å　80.05Å 90.00°　107.34°　90.00°	Depositor
Resolution (Å)	43.27　–　2.50	Depositor
	43.27　–　2.50	EDS
% Data completeness (in resolution range)	98.4 (43.27-2.50)	Depositor
	98.5 (43.27-2.50)	EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$<I/\sigma(I)>$ [1]	3.00 (at 2.51Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.232　,　0.273	Depositor
	0.220　,　0.221	DCC
R_{free} test set	2200 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	38.5	Xtriage
Anisotropy	0.372	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 46.1	EDS
L-test for twinning[2]	$<|L|> = 0.45$, $<L^2> = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8322	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 4.14% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ANP, MG

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	0.34	0/3975	0.63	3/5395 (0.1%)
1	B	0.35	0/3975	0.63	3/5395 (0.1%)
2	X	0.39	0/150	1.02	0/230
2	Y	0.39	0/150	1.03	0/230
All	All	0.35	0/8250	0.65	6/11250 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	574	GLY	N-CA-C	5.81	127.61	113.10
1	A	574	GLY	N-CA-C	5.75	127.49	113.10
1	B	298	LEU	N-CA-C	-5.23	96.89	111.00
1	B	575	GLU	N-CA-C	5.14	124.88	111.00
1	A	298	LEU	N-CA-C	-5.11	97.20	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	3899	0	3959	172	0
1	B	3899	0	3959	174	0

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	X	137	0	86	7	0
2	Y	137	0	86	8	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	31	0	13	3	0
4	B	31	0	13	4	0
5	A	94	0	0	9	0
5	B	76	0	0	10	0
5	X	9	0	0	2	0
5	Y	7	0	0	0	0
All	All	8322	0	8116	350	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 22.

The worst 5 of 350 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:579:ARG:HB3	1:B:579:ARG:HH11	1.28	0.99
1:A:357:VAL:HB	1:A:485:THR:HG22	1.43	0.98
1:B:357:VAL:HB	1:B:485:THR:HG22	1.45	0.97
1:A:579:ARG:HH11	1:A:579:ARG:HB3	1.27	0.96
1:A:258:VAL:HB	1:A:262:GLN:HE21	1.35	0.92

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	512/574 (89%)	474 (93%)	30 (6%)	8 (2%)	9	17
1	B	512/574 (89%)	476 (93%)	28 (6%)	8 (2%)	9	17

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1024/1148 (89%)	950 (93%)	58 (6%)	16 (2%)	9	17

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	257	ARG
1	A	575	GLU
1	B	257	ARG
1	B	575	GLU
1	A	576	GLY

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	396/435 (91%)	354 (89%)	42 (11%)	6	13
1	B	396/435 (91%)	355 (90%)	41 (10%)	7	13
All	All	792/870 (91%)	709 (90%)	83 (10%)	7	13

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	432	LEU
1	B	579	ARG
1	B	444	LEU
1	B	521	ASN
1	B	637	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	502	GLN
1	B	625	ASN
1	B	521	ASN

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type
1	B	572	ASN
1	B	638	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
4	ANP	A	801	3	29,33,33	1.15	2 (6%)	31,52,52	1.77	7 (22%)
4	ANP	B	801	3	29,33,33	1.15	2 (6%)	31,52,52	1.73	6 (19%)

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
4	ANP	A	801	3	-	3/14/38/38	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	ANP	B	801	3	-	3/14/38/38	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	801	ANP	PB-O2B	-3.10	1.48	1.56
4	B	801	ANP	PB-O2B	-3.06	1.48	1.56
4	B	801	ANP	O4'-C1'	2.31	1.44	1.41
4	A	801	ANP	O4'-C1'	2.27	1.44	1.41

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	801	ANP	N3-C2-N1	-4.52	121.61	128.68
4	A	801	ANP	N3-C2-N1	-4.47	121.69	128.68
4	B	801	ANP	O2B-PB-O1B	4.04	118.39	109.92
4	A	801	ANP	O2B-PB-O1B	3.85	118.00	109.92
4	A	801	ANP	C5'-C4'-C3'	-3.85	100.76	115.18

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	801	ANP	PB-N3B-PG-O1G
4	A	801	ANP	PG-N3B-PB-O1B
4	B	801	ANP	PB-N3B-PG-O1G
4	B	801	ANP	PG-N3B-PB-O1B
4	B	801	ANP	PG-N3B-PB-O3A

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	801	ANP	3	0
4	B	801	ANP	4	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 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.

Ligand ANP A 801

Bond lengths

Bond angles

Torsions

Rings

Ligand ANP B 801

Bond lengths

Bond angles

Torsions

Rings
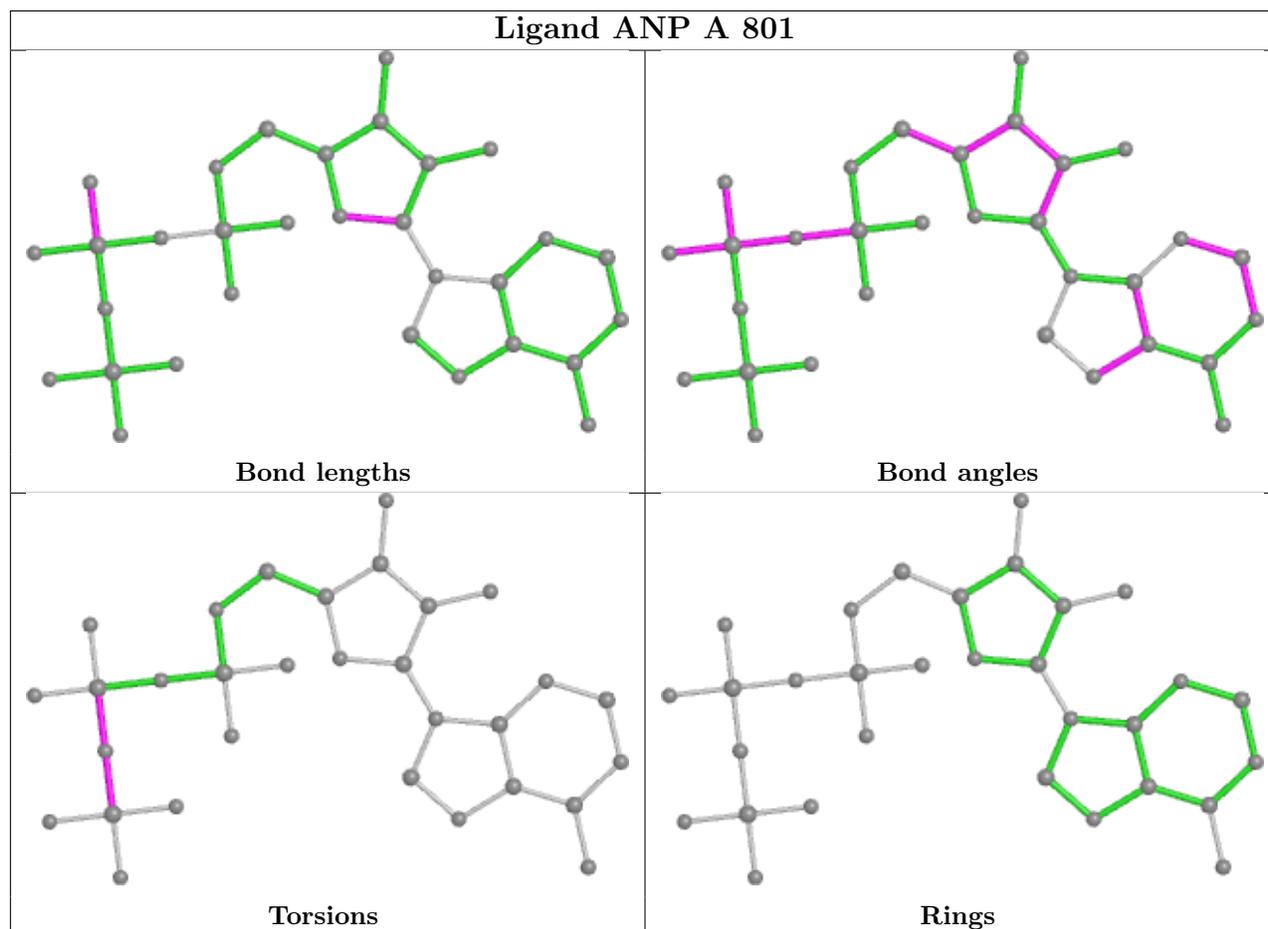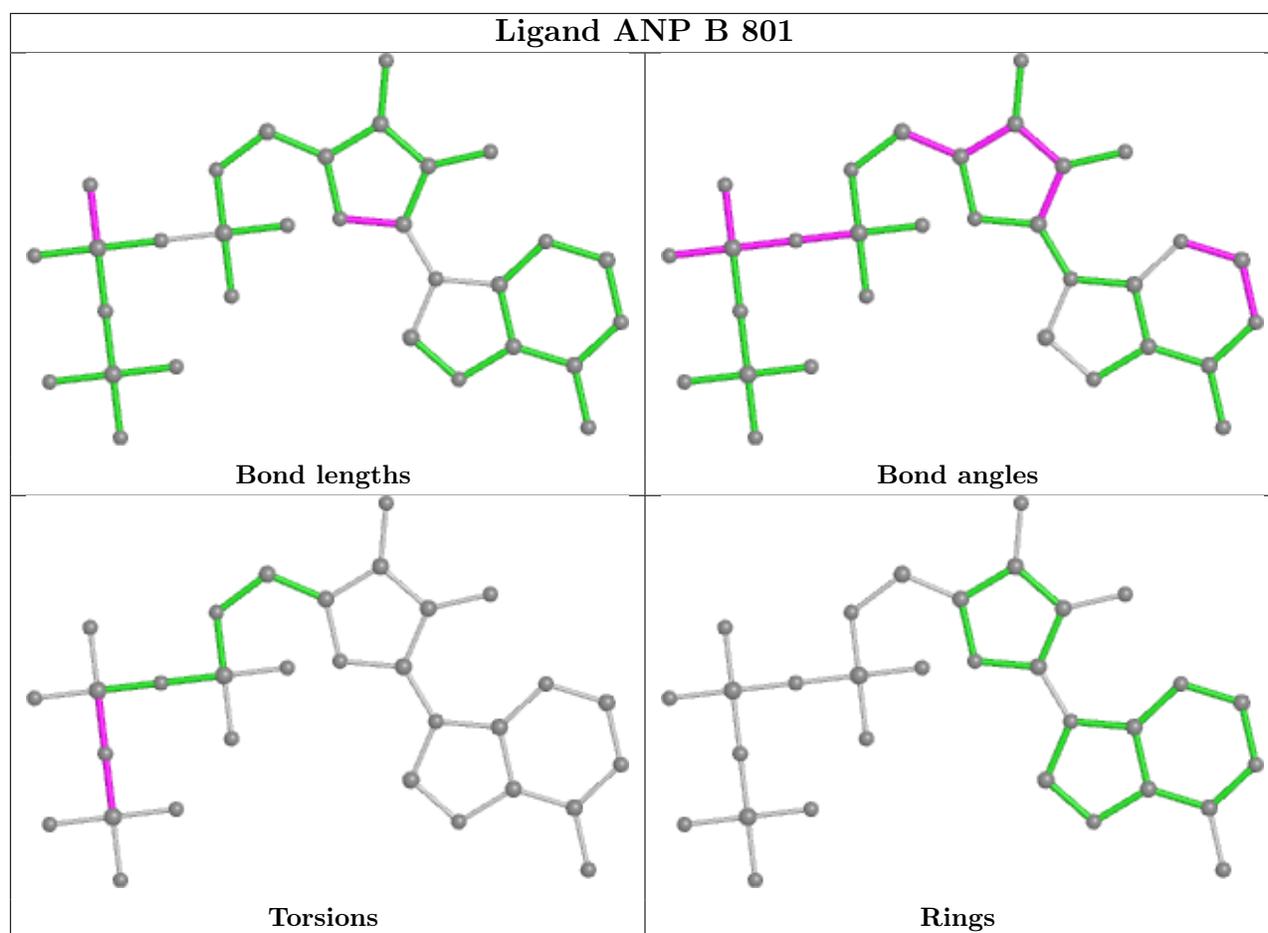
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6   Fit of model and data   ⓘ

6.1   Protein, DNA and RNA chains   ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	516/574 (89%)	0.50	37 (7%)	15	16	18, 39, 83, 112	2 (0%)
1	B	516/574 (89%)	0.45	44 (8%)	10	10	21, 41, 84, 115	2 (0%)
2	X	7/8 (87%)	-0.03	0	100	100	37, 40, 67, 85	0
2	Y	7/8 (87%)	0.08	0	100	100	37, 42, 71, 86	0
All	All	1046/1164 (89%)	0.47	81 (7%)	13	13	18, 40, 84, 115	4 (0%)

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	199	GLU	6.9
1	B	184	ALA	6.6
1	B	194	THR	6.5
1	A	188	LEU	6.3
1	B	195	LEU	6.2

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

There are no non-standard protein/DNA/RNA residues in this entry.
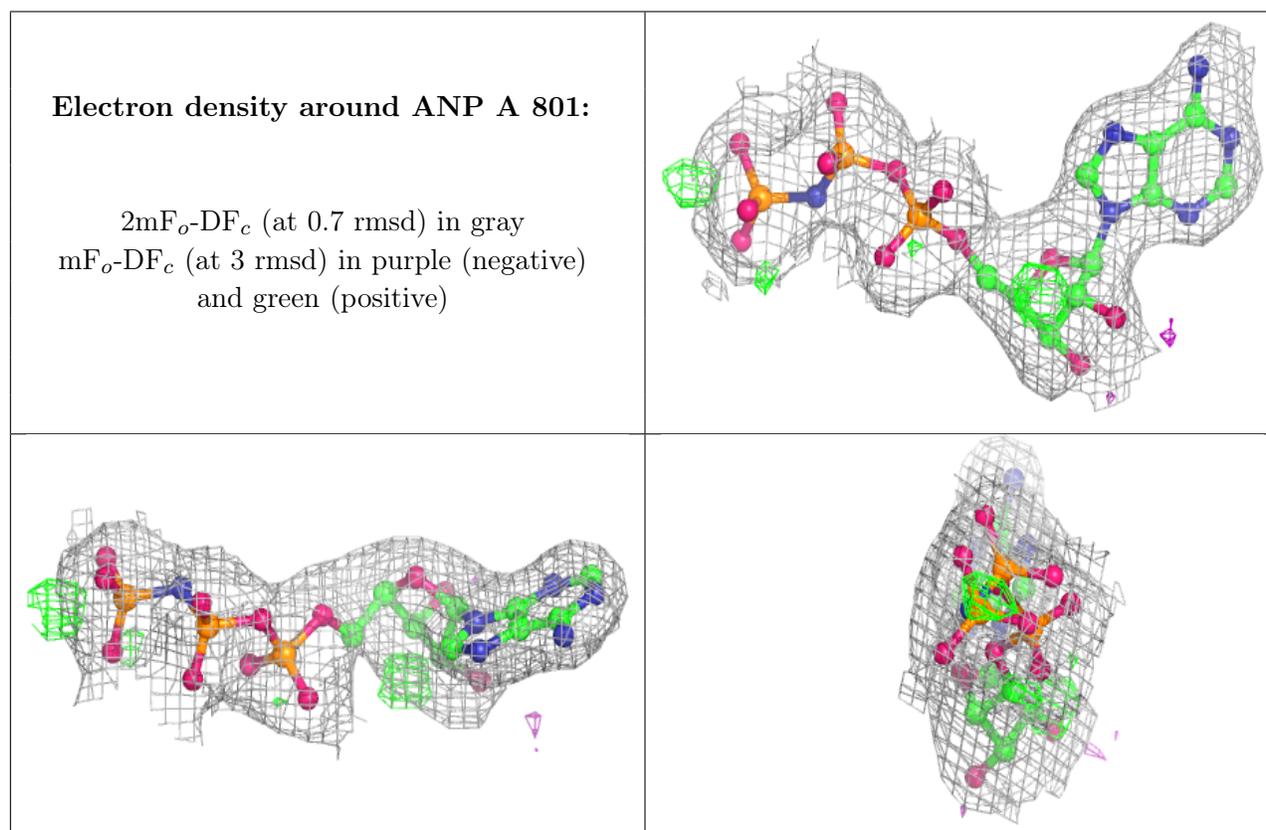
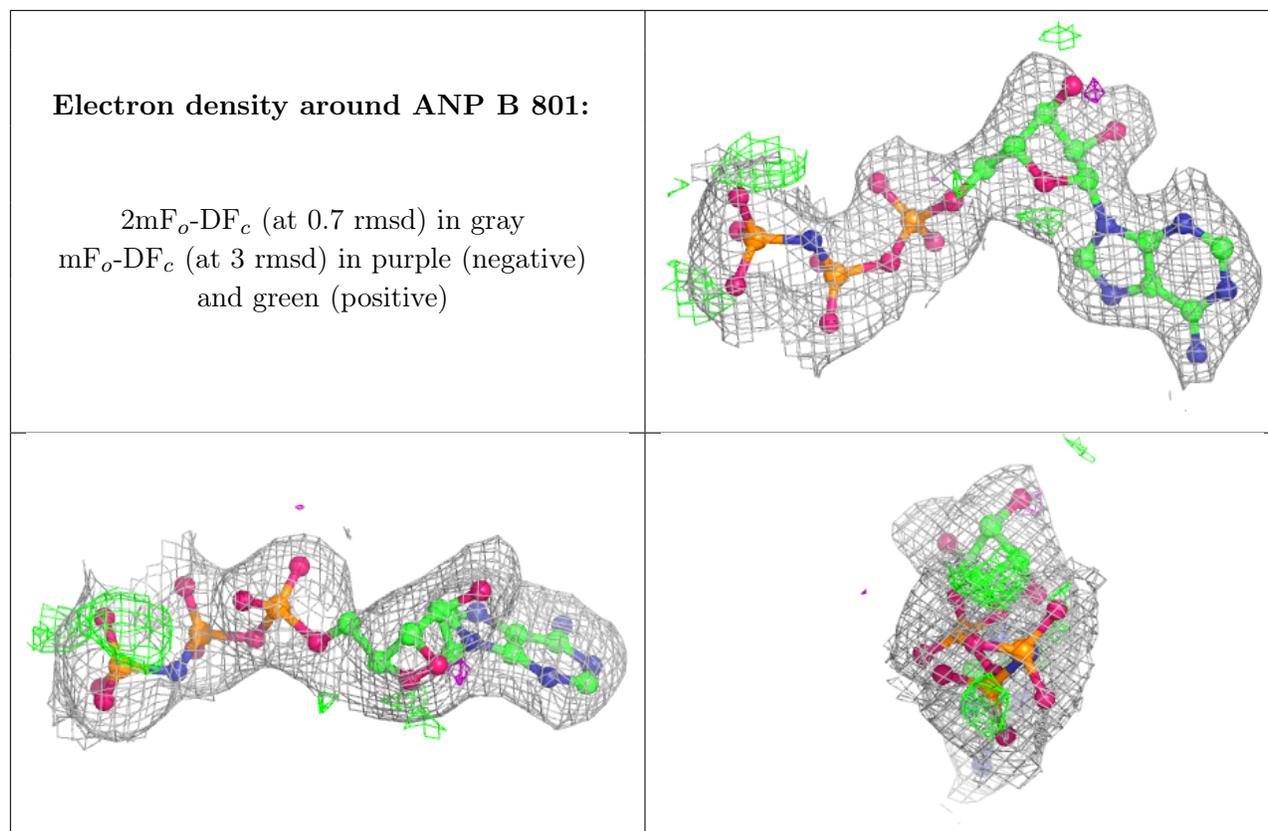## 6.3   Carbohydrates   ⓘ

There are no monosaccharides in this entry.

6.4   Ligands   ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MG	A	800	1/1	0.95	0.13	21,21,21,21	0
4	ANP	A	801	31/31	0.97	0.19	14,33,46,61	0
4	ANP	B	801	31/31	0.97	0.14	22,35,47,56	0
3	MG	B	800	1/1	0.99	0.06	20,20,20,20	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 ANP A 801:

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

Electron density around ANP B 801:

$2mF_o$-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 ⓘ

There are no such residues in this entry.