WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 16, 2023 – 12:14 pm GMT

PDB ID	:	4CXA
Title	:	Crystal structure of the human CDK12-cyclin K complex bound to AMPPNP
Authors	:	Dixon Clarke, S.E.; Elkins, J.M.; Pike, A.C.W.; Nowak, R.; Goubin, S.; Mahajan, R.P.; Kopec, J.; Froese, S.; Tallant, C.; Carpenter, E.P.; Mackenzie, A.; Faust, B.; Burgess-Brown, N.; von Delft, F.; Arrowsmith, C.; Edwards, A.M.; Bountra, C.; Bullock, A.
Deposited on	:	2014-04-04
Resolution	:	3.15 Å(reported)

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

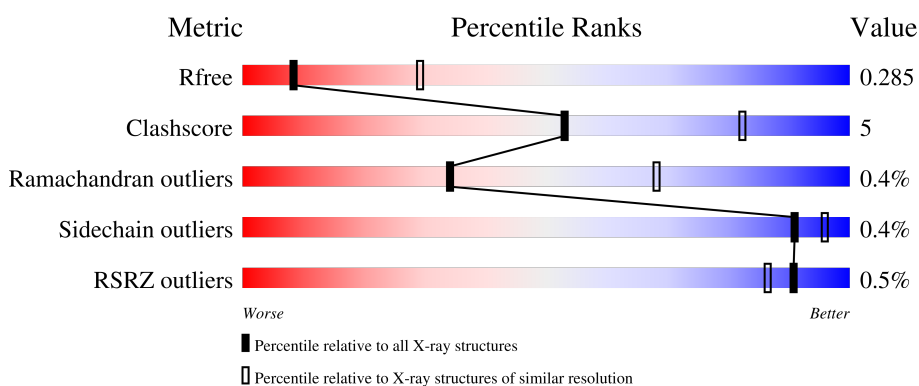|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.15 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	Percentile Ranks	Value
Rfree		0.285
Clashscore		5
Ramachandran outliers		0.4%
Sidechain outliers		0.4%
RSRZ outliers		0.5%

Worse ... *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

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	340	78% ... 19% ... •
1	C	340	86% ... 11% ... • •
2	B	259	83% ... 12% ... 5%
2	D	259	85% ... 8% ... 7%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 8681 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 CYCLIN-DEPENDENT KINASE 12.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	330	Total	C	N	O	P	S	0	0	0
			2466	1595	405	448	1	17			
1	C	331	Total	C	N	O	P	S	0	0	0
			2471	1595	404	455	1	16			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	713	SER	-	expression tag	UNP Q9NYV4
A	714	MET	-	expression tag	UNP Q9NYV4
C	713	SER	-	expression tag	UNP Q9NYV4
C	714	MET	-	expression tag	UNP Q9NYV4

- Molecule 2 is a protein called CYCLIN-K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	246	Total	C	N	O	S	0	0	0
			1936	1258	310	355	13			
2	D	241	Total	C	N	O	S	0	0	0
			1777	1161	279	326	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	9	SER	-	expression tag	UNP O75909
B	10	MET	-	expression tag	UNP O75909
D	9	SER	-	expression tag	UNP O75909
D	10	MET	-	expression tag	UNP O75909

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

ANP
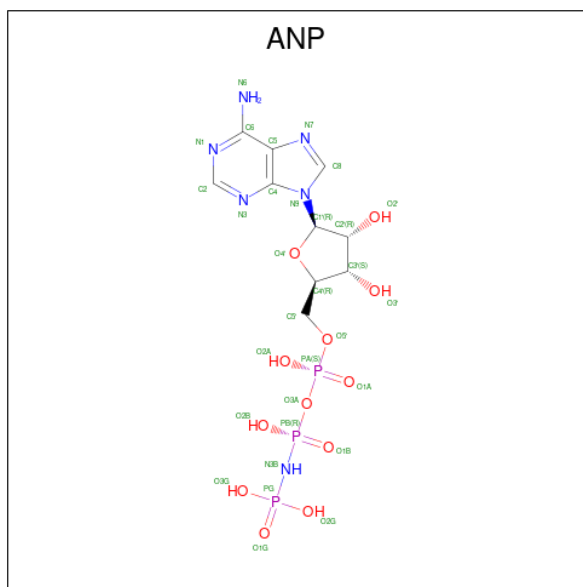
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 31	C 10	N 6	O 12	P 3	0	0

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: CYCLIN-DEPENDENT KINASE 12
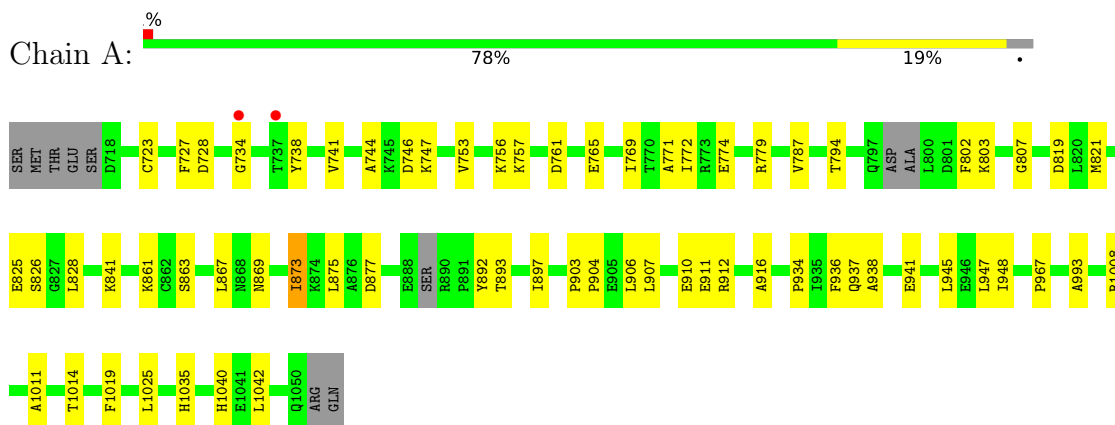
• Molecule 1: CYCLIN-DEPENDENT KINASE 12
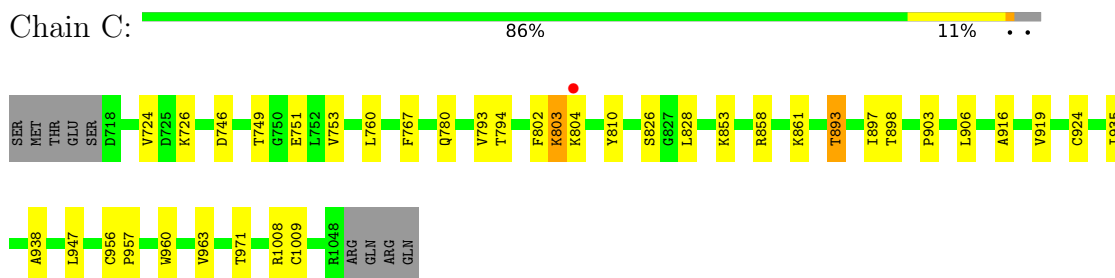
• Molecule 2: CYCLIN-K

• Molecule 2: CYCLIN-K

SER MET SER VAL THR SER ALA ASN LEU ASP HIS THR LYS P22 W26 L31 Q37 L41 E46 D59 G74 H79 F90 P91 L123 L124 N125 Q128 F129 F132 Q156 I178 L199 Q200 W201 W225 R233 E236 Y258 K262

GLN GLN MET PRO HIS
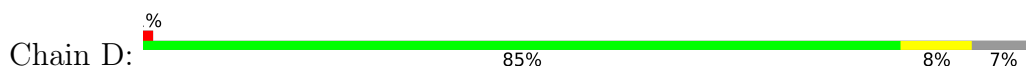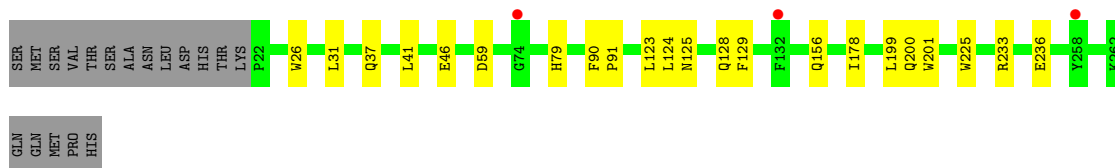
4 Data and refinement statistics ⓘ

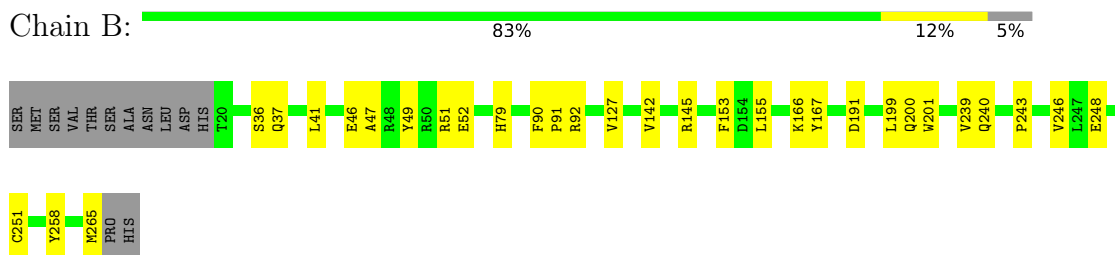| Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	69.84Å 138.79Å 71.89Å 90.00° 104.99° 90.00°	Depositor
Resolution (Å)	30.53 – 3.15 30.53 – 3.15	Depositor EDS
% Data completeness (in resolution range)	99.8 (30.53-3.15) 99.9 (30.53-3.15)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 3.18Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.221 , 0.279 0.228 , 0.285	Depositor DCC
R_{free} test set	1070 reflections (4.67%)	wwPDB-VP
Wilson B-factor (Å^2)	72.1	Xtriage
Anisotropy	0.536	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 32.8	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.067 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8681	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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: TPO, 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	0/2514	0.48	0/3430
1	C	0.26	0/2520	0.49	0/3444
2	B	0.25	0/1991	0.44	0/2721
2	D	0.24	0/1827	0.46	0/2511
All	All	0.25	0/8852	0.47	0/12106

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 ⓘ

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	2466	0	2254	38	0
1	C	2471	0	2254	22	0
2	B	1936	0	1775	20	0
2	D	1777	0	1553	14	0
3	A	31	0	13	2	0
All	All	8681	0	7849	89	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 5.

The worst 5 of 89 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:C:802:PHE:O	1:C:804:LYS:N	2.13	0.81
1:A:803:LYS:HA	2:B:142:VAL:HG11	1.68	0.76
1:A:867:LEU:HD13	1:A:873:ILE:HG12	1.74	0.69
2:B:52:GLU:OE1	2:B:92:ARG:NH1	2.28	0.66
1:A:938:ALA:HB2	1:A:947:LEU:HD12	1.76	0.66

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	323/340 (95%)	301 (93%)	20 (6%)	2 (1%)	25	62
1	C	328/340 (96%)	302 (92%)	24 (7%)	2 (1%)	25	62
2	B	244/259 (94%)	238 (98%)	6 (2%)	0	100	100
2	D	239/259 (92%)	230 (96%)	9 (4%)	0	100	100
All	All	1134/1198 (95%)	1071 (94%)	59 (5%)	4 (0%)	34	68

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	803	LYS
1	A	897	ILE
1	C	897	ILE
1	A	802	PHE

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	235/307 (76%)	234 (100%)	1 (0%)	91	96
1	C	237/307 (77%)	236 (100%)	1 (0%)	91	96
2	B	194/233 (83%)	193 (100%)	1 (0%)	88	95
2	D	162/233 (70%)	162 (100%)	0	100	100
All	All	828/1080 (77%)	825 (100%)	3 (0%)	91	96

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	873	ILE
2	B	127	VAL
1	C	753	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1035	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)

2 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	TPO	A	893	1	8,10,11	1.13	0	10,14,16	1.80	1 (10%)
1	TPO	C	893	1	8,10,11	1.18	0	10,14,16	1.74	1 (10%)

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	TPO	A	893	1	-	2/9/11/13	-
1	TPO	C	893	1	-	4/9/11/13	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	893	TPO	P-OG1-CB	-5.09	107.82	123.21
1	C	893	TPO	P-OG1-CB	-5.01	108.08	123.21

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	893	TPO	C-CA-CB-CG2
1	C	893	TPO	O-C-CA-CB
1	C	893	TPO	CB-OG1-P-O1P
1	A	893	TPO	CB-OG1-P-O2P
1	C	893	TPO	CB-OG1-P-O3P

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	893	TPO	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	$\#|Z| > 2$
3	ANP	A	2051	-	29,33,33	1.88	5 (17%)	31,52,52	1.06	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 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	ANP	A	2051	-	-	1/14/38/38	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2051	ANP	PG-O1G	8.03	1.58	1.46
3	A	2051	ANP	PG-N3B	2.69	1.70	1.63
3	A	2051	ANP	PG-O3G	-2.45	1.50	1.56
3	A	2051	ANP	PB-O1B	2.45	1.50	1.46
3	A	2051	ANP	PB-N3B	2.10	1.68	1.63

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	2051	ANP	PB-O3A-PA	-3.28	121.08	132.62
3	A	2051	ANP	C5-C6-N6	2.27	123.80	120.35

There are no chirality outliers.

All (1) torsion outliers are listed below:

| Mol	Chain	Res	Type	Atoms
3	A	2051	ANP	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2051	ANP	2	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 2051

Bond lengths

Bond angles

Torsions

Rings
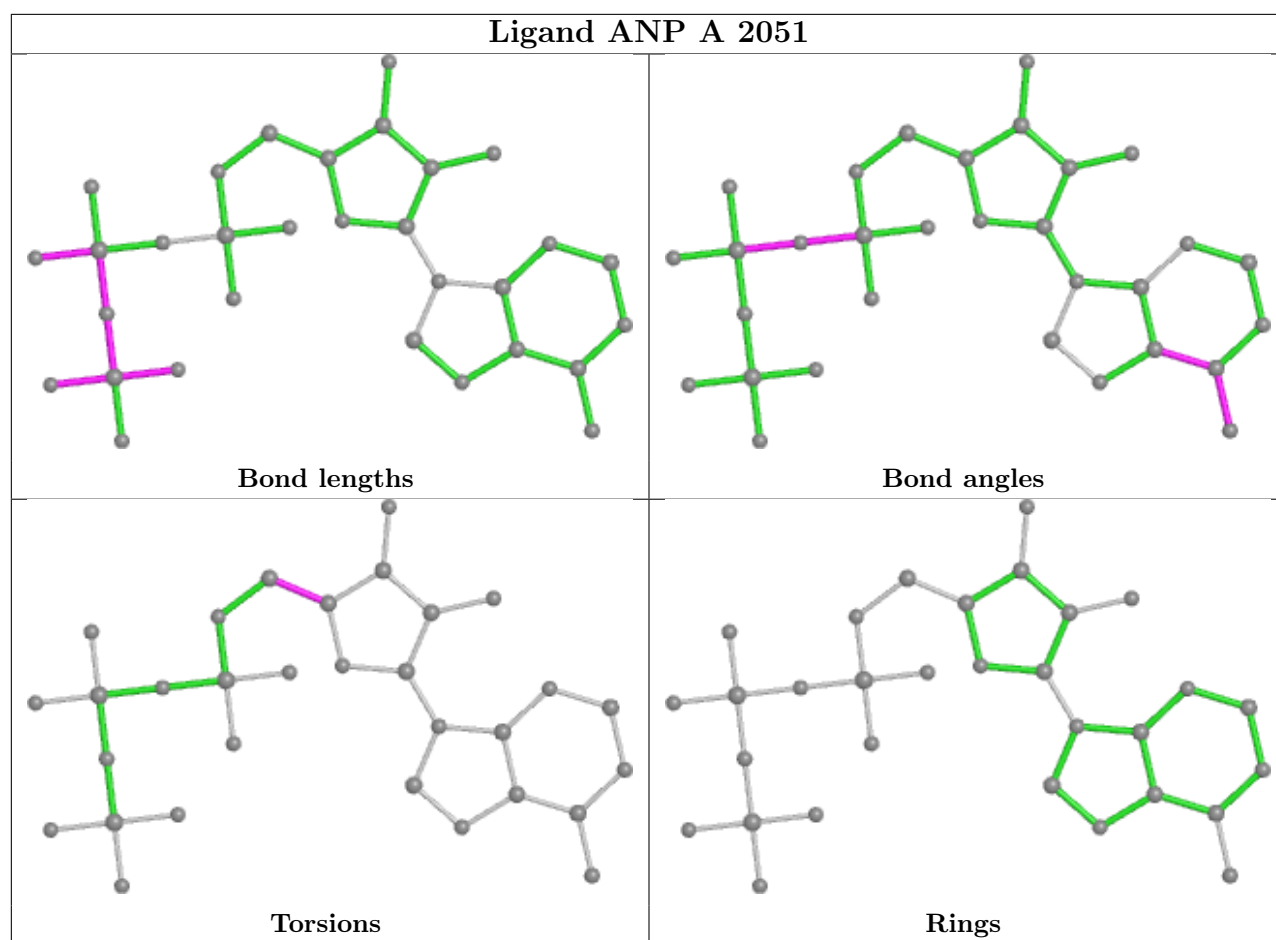
5.7 Other polymers (i)

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 ⓘ

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(Å²)	Q<0.9
1	A	329/340 (96%)	-0.33	2 (0%)	89	84	46, 67, 96, 114	0
1	C	330/340 (97%)	-0.28	1 (0%)	94	92	42, 67, 107, 134	0
2	B	246/259 (94%)	-0.33	0	100	100	49, 69, 92, 110	0
2	D	241/259 (93%)	-0.12	3 (1%)	79	68	59, 89, 118, 126	0
All	All	1146/1198 (95%)	-0.27	6 (0%)	91	86	42, 72, 107, 134	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	74	GLY	3.1
1	A	734	GLY	2.7
1	C	804	LYS	2.4
1	A	737	THR	2.3
2	D	258	TYR	2.3

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

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(Å²)	Q<0.9
1	TPO	A	893	11/12	0.88	0.20	80,89,98,101	0
1	TPO	C	893	11/12	0.88	0.22	79,90,105,114	0

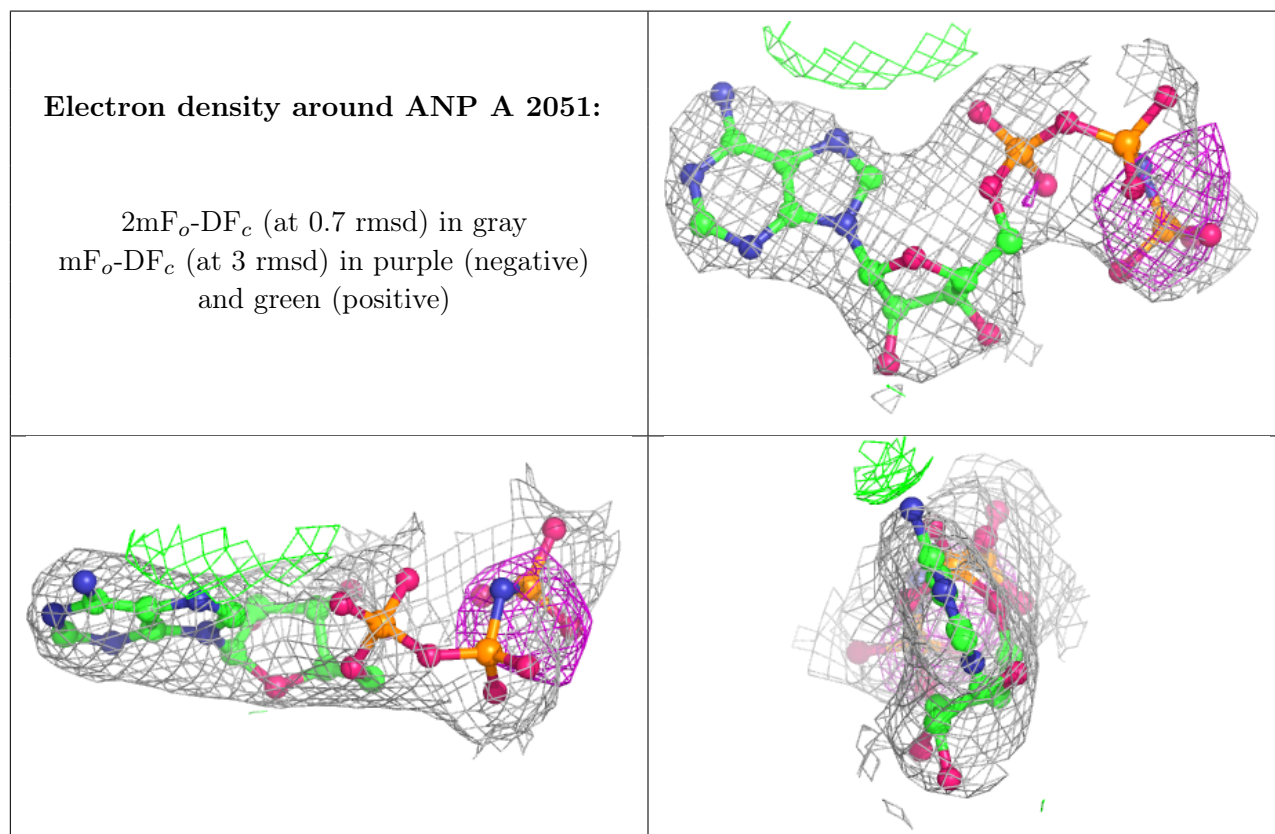## 6.3 Carbohydrates ⓘ

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	ANP	A	2051	31/31	0.93	0.15	53,68,103,113	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 2051:

$2mF_o$-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.