



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 05:22 pm BST

PDB ID : 4QQZ  
Title : Crystal structure of T. fusca Cas3-AMPPNP  
Authors : Ke, A.; Huo, Y.; Nam, K.H.  
Deposited on : 2014-06-30  
Resolution : 2.93 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](mailto: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 : 2.11  
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) : 2.11

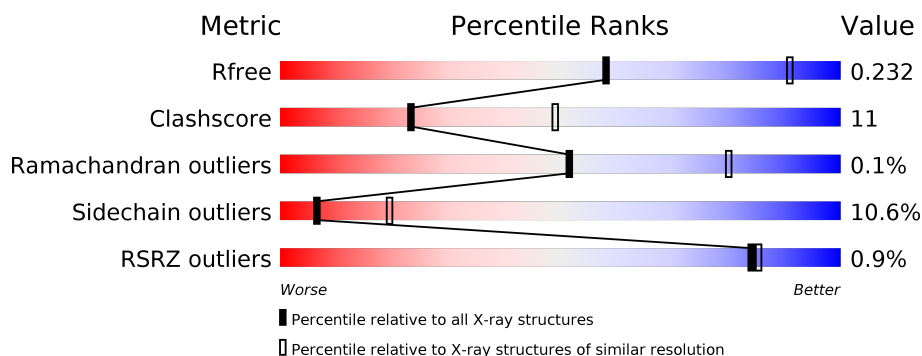
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.93 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	2969 (2.98-2.90)
Clashscore	141614	3218 (2.98-2.90)
Ramachandran outliers	138981	3122 (2.98-2.90)
Sidechain outliers	138945	3124 (2.98-2.90)
RSRZ outliers	127900	2902 (2.98-2.90)

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  $\geq 3$ , 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	964	<div> <div>2%</div> <div>68% 21% 6%</div> </div>
1	C	964	<div> <div>70% 20% 6%</div> </div>
1	E	964	<div> <div>70% 20% 7%</div> </div>
1	G	964	<div> <div>% 65% 23% 7%</div> </div>
2	B	12	<div> <div>8% 25% 50% 17% 8%</div> </div>
2	D	12	<div> <div>25% 42% 25% 8%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	12	<div><div></div><div>8%</div><div>33%</div><div>42%</div><div>17%</div><div>8%</div></div>
2	H	12	<div><div></div><div>33%</div><div>25%</div><div>33%</div><div>8%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 28851 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 CRISPR-associated helicase, Cas3 family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	903	Total	C	N	O	S	0	0	0
			7022	4460	1250	1285	27			
1	C	902	Total	C	N	O	S	0	0	0
			7001	4441	1243	1290	27			
1	E	899	Total	C	N	O	S	0	0	0
			6985	4434	1243	1281	27			
1	G	892	Total	C	N	O	S	0	0	0
			6923	4398	1230	1268	27			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	INITIATING METHIONINE	UNP Q47PJ0
A	-18	GLY	-	EXPRESSION TAG	UNP Q47PJ0
A	-17	SER	-	EXPRESSION TAG	UNP Q47PJ0
A	-16	SER	-	EXPRESSION TAG	UNP Q47PJ0
A	-15	HIS	-	EXPRESSION TAG	UNP Q47PJ0
A	-14	HIS	-	EXPRESSION TAG	UNP Q47PJ0
A	-13	HIS	-	EXPRESSION TAG	UNP Q47PJ0
A	-12	HIS	-	EXPRESSION TAG	UNP Q47PJ0
A	-11	HIS	-	EXPRESSION TAG	UNP Q47PJ0
A	-10	HIS	-	EXPRESSION TAG	UNP Q47PJ0
A	-9	SER	-	EXPRESSION TAG	UNP Q47PJ0
A	-8	SER	-	EXPRESSION TAG	UNP Q47PJ0
A	-7	GLY	-	EXPRESSION TAG	UNP Q47PJ0
A	-6	LEU	-	EXPRESSION TAG	UNP Q47PJ0
A	-5	VAL	-	EXPRESSION TAG	UNP Q47PJ0
A	-4	PRO	-	EXPRESSION TAG	UNP Q47PJ0
A	-3	ARG	-	EXPRESSION TAG	UNP Q47PJ0
A	-2	GLY	-	EXPRESSION TAG	UNP Q47PJ0
A	-1	SER	-	EXPRESSION TAG	UNP Q47PJ0
A	0	HIS	-	EXPRESSION TAG	UNP Q47PJ0
C	-19	MET	-	INITIATING METHIONINE	UNP Q47PJ0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-18	GLY	-	EXPRESSION TAG	UNP Q47PJ0
C	-17	SER	-	EXPRESSION TAG	UNP Q47PJ0
C	-16	SER	-	EXPRESSION TAG	UNP Q47PJ0
C	-15	HIS	-	EXPRESSION TAG	UNP Q47PJ0
C	-14	HIS	-	EXPRESSION TAG	UNP Q47PJ0
C	-13	HIS	-	EXPRESSION TAG	UNP Q47PJ0
C	-12	HIS	-	EXPRESSION TAG	UNP Q47PJ0
C	-11	HIS	-	EXPRESSION TAG	UNP Q47PJ0
C	-10	HIS	-	EXPRESSION TAG	UNP Q47PJ0
C	-9	SER	-	EXPRESSION TAG	UNP Q47PJ0
C	-8	SER	-	EXPRESSION TAG	UNP Q47PJ0
C	-7	GLY	-	EXPRESSION TAG	UNP Q47PJ0
C	-6	LEU	-	EXPRESSION TAG	UNP Q47PJ0
C	-5	VAL	-	EXPRESSION TAG	UNP Q47PJ0
C	-4	PRO	-	EXPRESSION TAG	UNP Q47PJ0
C	-3	ARG	-	EXPRESSION TAG	UNP Q47PJ0
C	-2	GLY	-	EXPRESSION TAG	UNP Q47PJ0
C	-1	SER	-	EXPRESSION TAG	UNP Q47PJ0
C	0	HIS	-	EXPRESSION TAG	UNP Q47PJ0
E	-19	MET	-	INITIATING METHIONINE	UNP Q47PJ0
E	-18	GLY	-	EXPRESSION TAG	UNP Q47PJ0
E	-17	SER	-	EXPRESSION TAG	UNP Q47PJ0
E	-16	SER	-	EXPRESSION TAG	UNP Q47PJ0
E	-15	HIS	-	EXPRESSION TAG	UNP Q47PJ0
E	-14	HIS	-	EXPRESSION TAG	UNP Q47PJ0
E	-13	HIS	-	EXPRESSION TAG	UNP Q47PJ0
E	-12	HIS	-	EXPRESSION TAG	UNP Q47PJ0
E	-11	HIS	-	EXPRESSION TAG	UNP Q47PJ0
E	-10	HIS	-	EXPRESSION TAG	UNP Q47PJ0
E	-9	SER	-	EXPRESSION TAG	UNP Q47PJ0
E	-8	SER	-	EXPRESSION TAG	UNP Q47PJ0
E	-7	GLY	-	EXPRESSION TAG	UNP Q47PJ0
E	-6	LEU	-	EXPRESSION TAG	UNP Q47PJ0
E	-5	VAL	-	EXPRESSION TAG	UNP Q47PJ0
E	-4	PRO	-	EXPRESSION TAG	UNP Q47PJ0
E	-3	ARG	-	EXPRESSION TAG	UNP Q47PJ0
E	-2	GLY	-	EXPRESSION TAG	UNP Q47PJ0
E	-1	SER	-	EXPRESSION TAG	UNP Q47PJ0
E	0	HIS	-	EXPRESSION TAG	UNP Q47PJ0
G	-19	MET	-	INITIATING METHIONINE	UNP Q47PJ0
G	-18	GLY	-	EXPRESSION TAG	UNP Q47PJ0
G	-17	SER	-	EXPRESSION TAG	UNP Q47PJ0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	-16	SER	-	EXPRESSION TAG	UNP Q47PJ0
G	-15	HIS	-	EXPRESSION TAG	UNP Q47PJ0
G	-14	HIS	-	EXPRESSION TAG	UNP Q47PJ0
G	-13	HIS	-	EXPRESSION TAG	UNP Q47PJ0
G	-12	HIS	-	EXPRESSION TAG	UNP Q47PJ0
G	-11	HIS	-	EXPRESSION TAG	UNP Q47PJ0
G	-10	HIS	-	EXPRESSION TAG	UNP Q47PJ0
G	-9	SER	-	EXPRESSION TAG	UNP Q47PJ0
G	-8	SER	-	EXPRESSION TAG	UNP Q47PJ0
G	-7	GLY	-	EXPRESSION TAG	UNP Q47PJ0
G	-6	LEU	-	EXPRESSION TAG	UNP Q47PJ0
G	-5	VAL	-	EXPRESSION TAG	UNP Q47PJ0
G	-4	PRO	-	EXPRESSION TAG	UNP Q47PJ0
G	-3	ARG	-	EXPRESSION TAG	UNP Q47PJ0
G	-2	GLY	-	EXPRESSION TAG	UNP Q47PJ0
G	-1	SER	-	EXPRESSION TAG	UNP Q47PJ0
G	0	HIS	-	EXPRESSION TAG	UNP Q47PJ0

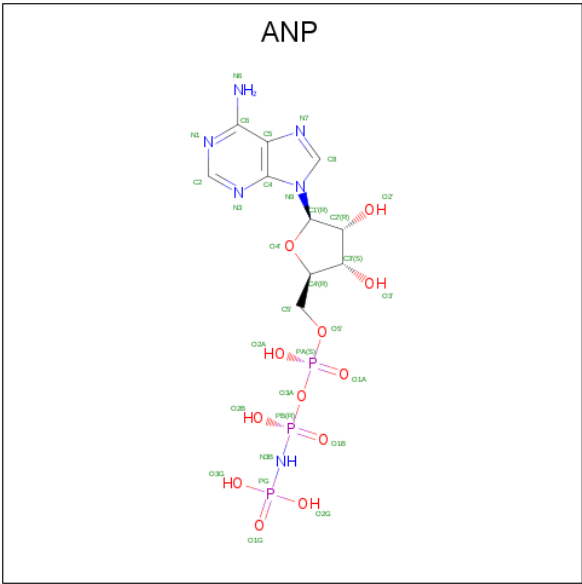
- Molecule 2 is a DNA chain called DNA (5'-D(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	11	Total	C	N	O	P	0	0	0
			197	90	45	51	11			
2	D	11	Total	C	N	O	P	0	0	0
			197	90	45	51	11			
2	F	11	Total	C	N	O	P	0	0	0
			197	90	45	51	11			
2	H	11	Total	C	N	O	P	0	0	0
			197	90	45	51	11			

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	2	Total	Fe	0	0
			2	2		
3	A	2	Total	Fe	0	0
			2	2		
3	C	2	Total	Fe	0	0
			2	2		
3	E	2	Total	Fe	0	0
			2	2		

- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>12</sub>P<sub>3</sub>).

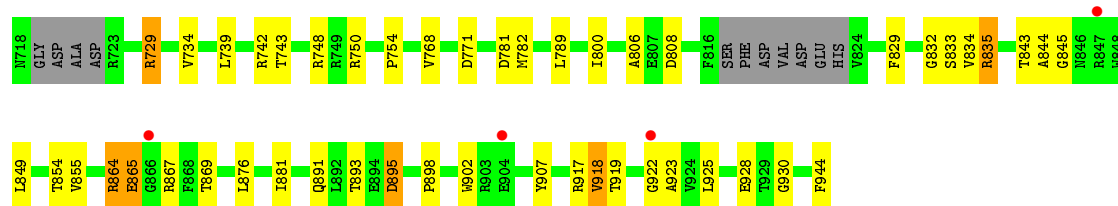




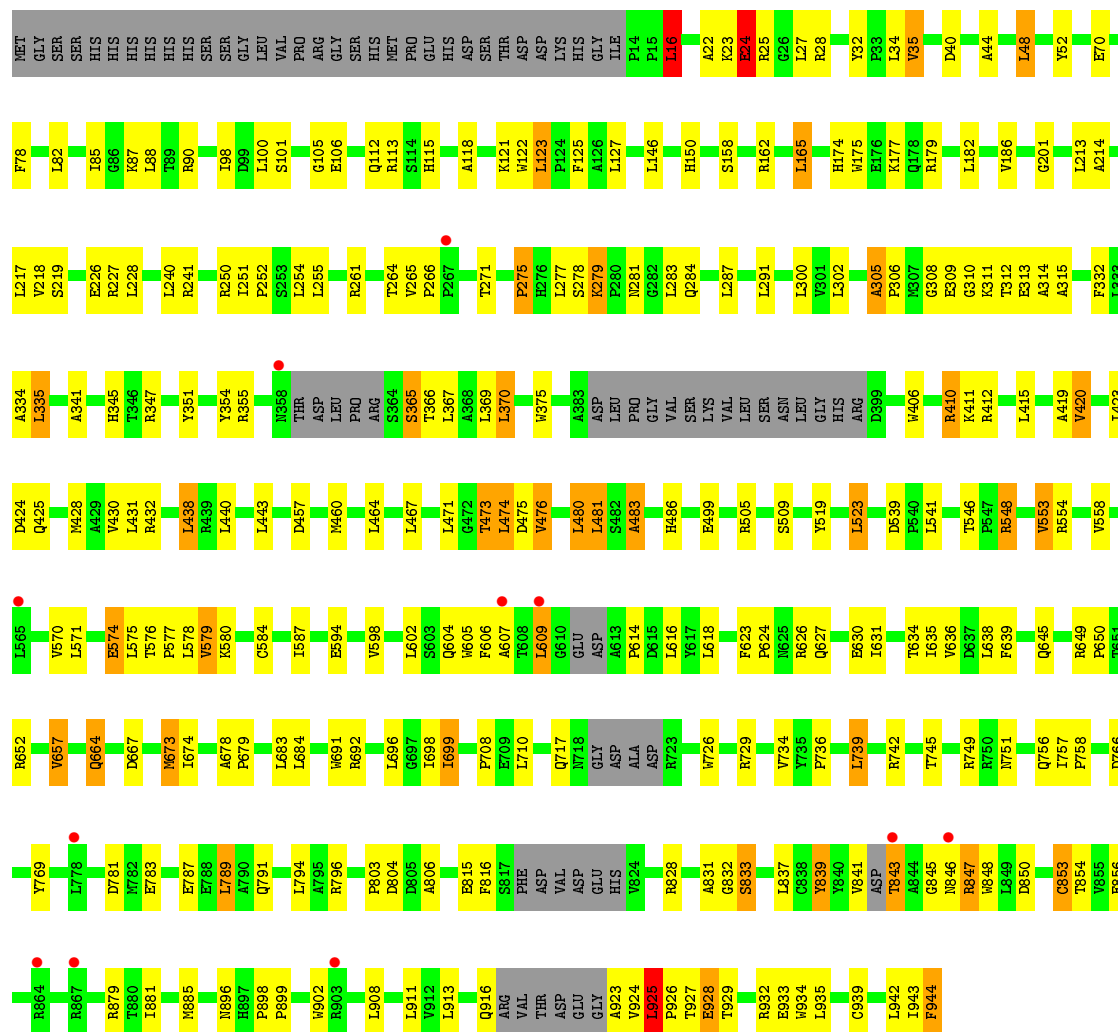




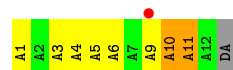
GLU ASP	Y459	A368 L369	L254	K121 K122	GLY SER	MET
A613	L464	L370	A258	L123	SER	
P614	R469	R375	R261	P124	HIS	
D615	W470	L376	W285	F125	HIS	
L616	L471	P377	P266	A126	HIS	
	G472	P378		L127	HIS	
B630	T473	A383	T271	Y132	HIS	
I631	L474	ASP			HIS	
T632	D475	LEU	L277	L137	SER	
A633	P476	PRO	S278	V138	GLY	
T634	P477	GLY	R279	T139	LEU	
I635		VAL	L283	M145	VAL	
V636	L480	VAL	L284	L146	PRO	
D637	L481	SER	L287	H150	ARG	
L638	L482	LVS			GLY	
	S482	VAL			SER	
R649	A483	LEU			SER	
P650	W506	LEU			HIS	
R652		SER	L291	F159	HIS	
	P513	ASN		P160	MET	
L656		LEU	T296	S161	PRO	
	E516	GLY		R162	GLU	
V662	V517	HIS	L300	M163	HIS	
	S518	ARG	T304	P164	ASP	
L666	D399			L165	SER	
D667	T404		G310	A166	THR	
L668	L407			E167	ASP	
	R534		A315	S172	ASP	
M673	L541	R410		P173	LVS	
L677	P542	K411	G326	K177	HIS	
A678	R548	R418	G327	Q178	GLY	
R679		A419	R328		ILE	P14
R680	V553	V420	F332	L182		P15
L682			L333			L16
L683	V558	L423	A334	V186	F20	
L684		D424	L335		R21	
	L165			G201	A22	
			T340		R23	
W691	V570	M428		L217	E24	
R692		A429	R345	V218	L34	
L693	T576	V430		S219	V35	
E694		L431	Y351			
H695	V579	L438	R355	L228	L48	
L696	K580	R439	V356	R241	W49	
G697	A566	L440	R357	A242		
L698		L443	R358		L53	
	E594	V447	THR	E245		
R701			ASP	T246	G56	
T702	V598	V450	LEU	S247		
Q703	Y599		F362	L248	S64	
			R363	R249		
R706	F606	A453	S364	R250	I85	
Q707			L251	I251		
R708			S365	R250	I96	
	L609	V456	T366	P252		
P715	GLY		L367	S368	F111	



- Molecule 1: CRISPR-associated helicase, Cas3 family



- Molecule 2: DNA (5'-D(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3')



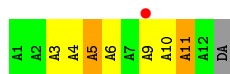
- Molecule 2: DNA (5'-D(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3')

Chain D: 




- Molecule 2: DNA (5'-D(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3')

Chain F: 



- Molecule 2: DNA (5'-D(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3')

Chain H: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.92Å 222.06Å 124.90Å 90.00° 104.29° 90.00°	Depositor
Resolution (Å)	50.11 – 2.93 50.11 – 2.61	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.11-2.93) 88.9 (50.11-2.61)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 2.61Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, $R_{free}$	0.182 , 0.234 0.184 , 0.232	Depositor DCC
$R_{free}$ test set	2000 reflections (1.47%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.2	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 37.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	28851	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.49% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<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.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	0.56	4/7199 (0.1%)	0.67	11/9817 (0.1%)
1	C	0.51	0/7176	0.64	4/9790 (0.0%)
1	E	0.49	0/7160	0.63	5/9766 (0.1%)
1	G	0.52	1/7095 (0.0%)	0.66	7/9675 (0.1%)
2	B	1.13	0/222	2.38	18/339 (5.3%)
2	D	1.09	0/222	2.36	14/339 (4.1%)
2	F	1.11	0/222	2.26	14/339 (4.1%)
2	H	0.93	0/222	2.13	14/339 (4.1%)
All	All	0.55	5/29518 (0.0%)	0.77	87/40404 (0.2%)

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	A	0	1
1	G	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	481	LEU	C-N	-15.87	0.97	1.34
1	A	614	PRO	N-CD	10.18	1.62	1.47
1	A	606	PHE	CE1-CZ	8.85	1.54	1.37
1	G	275	PRO	N-CD	8.81	1.60	1.47
1	A	484	THR	CB-CG2	-5.30	1.34	1.52

The worst 5 of 87 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	11	DA	O4'-C1'-N9	13.99	117.79	108.00
2	H	10	DA	O4'-C1'-C2'	-12.55	95.86	105.90
2	D	11	DA	O4'-C1'-N9	12.37	116.66	108.00
2	F	11	DA	O4'-C4'-C3'	12.25	113.35	106.00
2	B	9	DA	O4'-C1'-N9	11.96	116.37	108.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	482	SER	Mainchain
1	G	925	LEU	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	7022	0	6970	161	0
1	C	7001	0	6944	148	0
1	E	6985	0	6939	131	0
1	G	6923	0	6878	199	0
2	B	197	0	99	9	0
2	D	197	0	99	10	0
2	F	197	0	99	6	0
2	H	197	0	99	12	0
3	A	2	0	0	0	0
3	C	2	0	0	0	0
3	E	2	0	0	0	0
3	G	2	0	0	0	0
4	A	31	0	13	4	0
4	C	31	0	13	6	0
4	E	31	0	13	5	0
4	G	31	0	13	8	0
All	All	28851	0	28179	643	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 11.

The worst 5 of 643 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:G:618:LEU:CD1	1:G:657:VAL:HG12	1.71	1.21
1:G:587:ILE:HD11	1:G:657:VAL:CG2	1.76	1.14
1:G:310:GLY:N	4:G:1003:ANP:O1A	1.86	1.08
1:G:587:ILE:CD1	1:G:657:VAL:HG22	1.85	1.06
1:G:618:LEU:HD12	1:G:657:VAL:HG12	1.36	1.05

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	889/964 (92%)	853 (96%)	36 (4%)	0	100	100
1	C	892/964 (92%)	862 (97%)	29 (3%)	1 (0%)	51	80
1	E	887/964 (92%)	852 (96%)	35 (4%)	0	100	100
1	G	876/964 (91%)	830 (95%)	44 (5%)	2 (0%)	47	76
All	All	3544/3856 (92%)	3397 (96%)	144 (4%)	3 (0%)	51	80

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	605	TRP
1	C	483	ALA
1	G	853	CYS

### 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	744/796 (94%)	671 (90%)	73 (10%)	8	23
1	C	742/796 (93%)	665 (90%)	77 (10%)	7	20
1	E	740/796 (93%)	657 (89%)	83 (11%)	6	17
1	G	733/796 (92%)	651 (89%)	82 (11%)	6	17
All	All	2959/3184 (93%)	2644 (89%)	315 (11%)	6	19

5 of 315 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	817	SER
1	E	304	THR
1	G	652	ARG
1	C	881	ILE
1	E	146	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	281	ASN
1	G	811	ASN
1	G	664	GLN
1	E	358	ASN
1	G	717	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 carbohydrates in this entry.



## 5.6 Ligand geometry

Of 12 ligands modelled in this entry, 8 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
4	ANP	E	1003	-	29,33,33	2.12	11 (37%)	31,52,52	2.75	10 (32%)
4	ANP	G	1003	-	29,33,33	1.84	10 (34%)	31,52,52	3.28	14 (45%)
4	ANP	A	1003	-	29,33,33	1.22	4 (13%)	31,52,52	1.02	2 (6%)
4	ANP	C	1003	-	29,33,33	2.14	11 (37%)	31,52,52	2.53	13 (41%)

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	E	1003	-	-	6/14/38/38	0/3/3/3
4	ANP	G	1003	-	-	7/14/38/38	0/3/3/3
4	ANP	A	1003	-	-	6/14/38/38	0/3/3/3
4	ANP	C	1003	-	-	8/14/38/38	0/3/3/3

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	1003	ANP	C2'-C1'	-5.15	1.45	1.53
4	E	1003	ANP	PB-O2B	-4.90	1.43	1.56
4	E	1003	ANP	C4-N3	-3.88	1.30	1.35
4	G	1003	ANP	PB-O2B	-3.71	1.46	1.56
4	E	1003	ANP	C2'-C1'	-3.56	1.48	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	1003	ANP	O1G-PG-N3B	-9.97	97.09	111.77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	1003	ANP	O2B-PB-O1B	7.91	126.50	109.92
4	C	1003	ANP	O2B-PB-O1B	6.61	123.78	109.92
4	G	1003	ANP	O3'-C3'-C2'	-6.32	91.37	111.82
4	E	1003	ANP	O1G-PG-N3B	-6.22	102.62	111.77

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

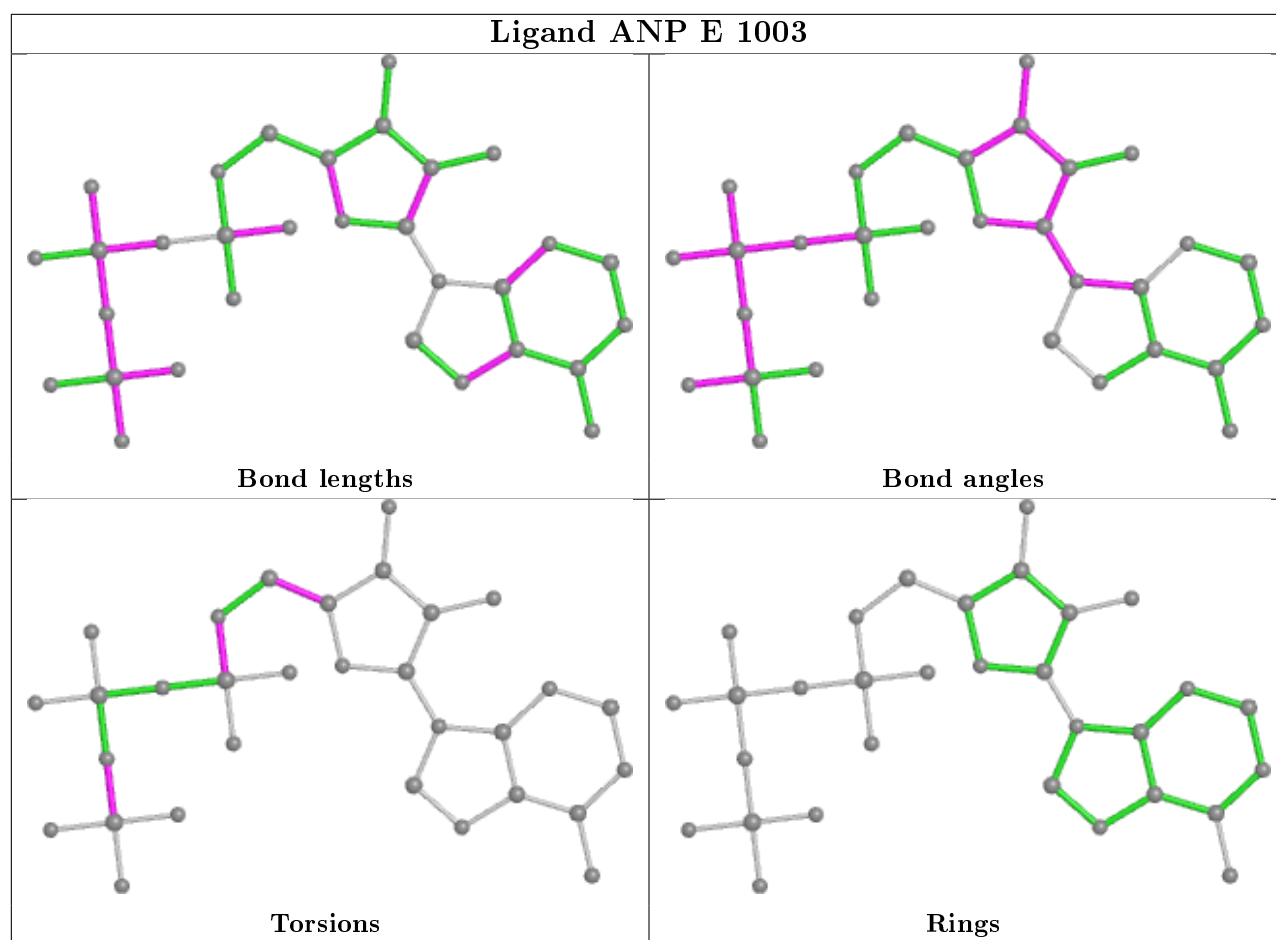
Mol	Chain	Res	Type	Atoms
4	G	1003	ANP	PB-N3B-PG-O1G
4	G	1003	ANP	PA-O3A-PB-O1B
4	G	1003	ANP	PA-O3A-PB-O2B
4	E	1003	ANP	PB-N3B-PG-O1G
4	E	1003	ANP	C5'-O5'-PA-O1A

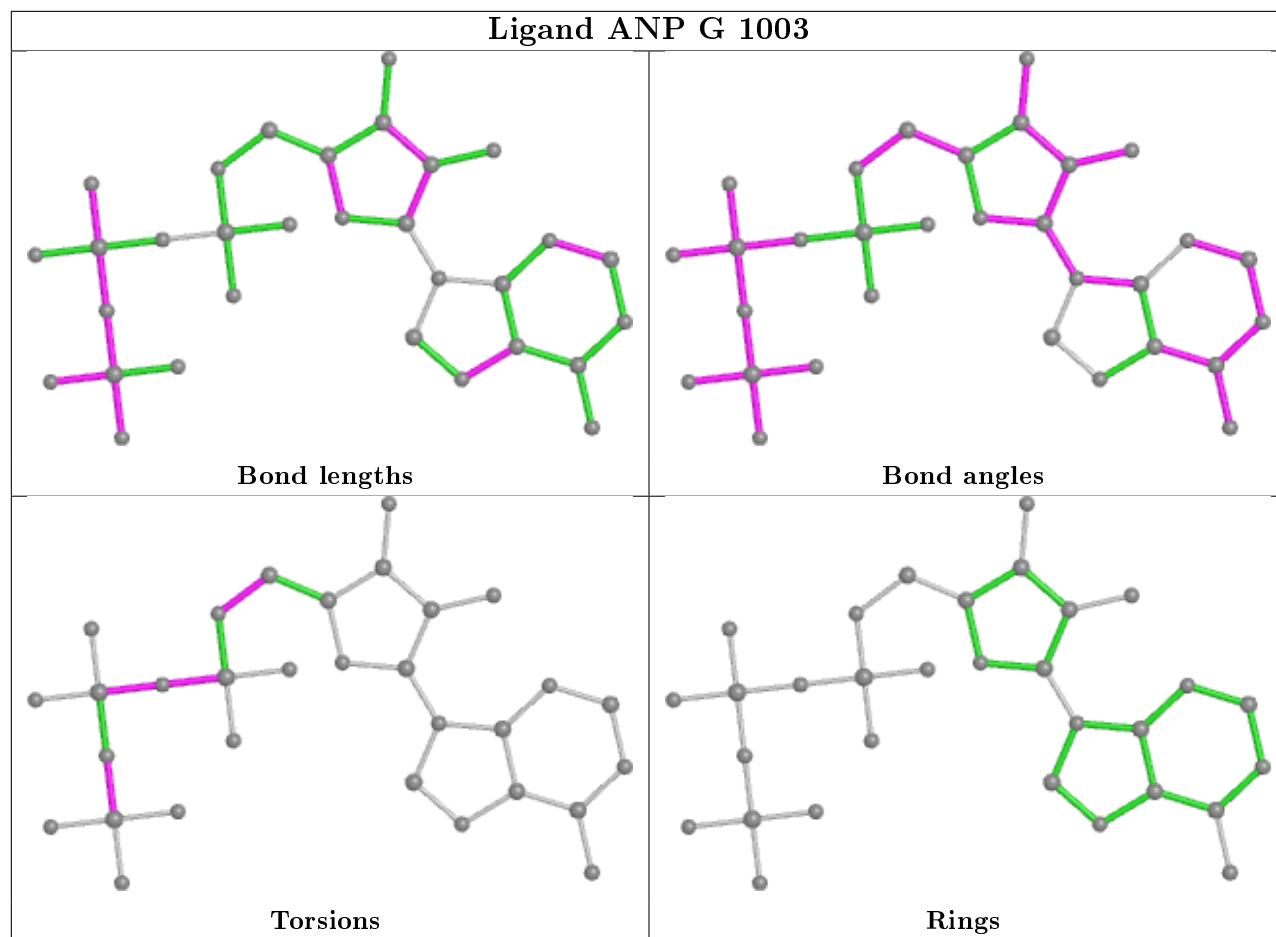
There are no ring outliers.

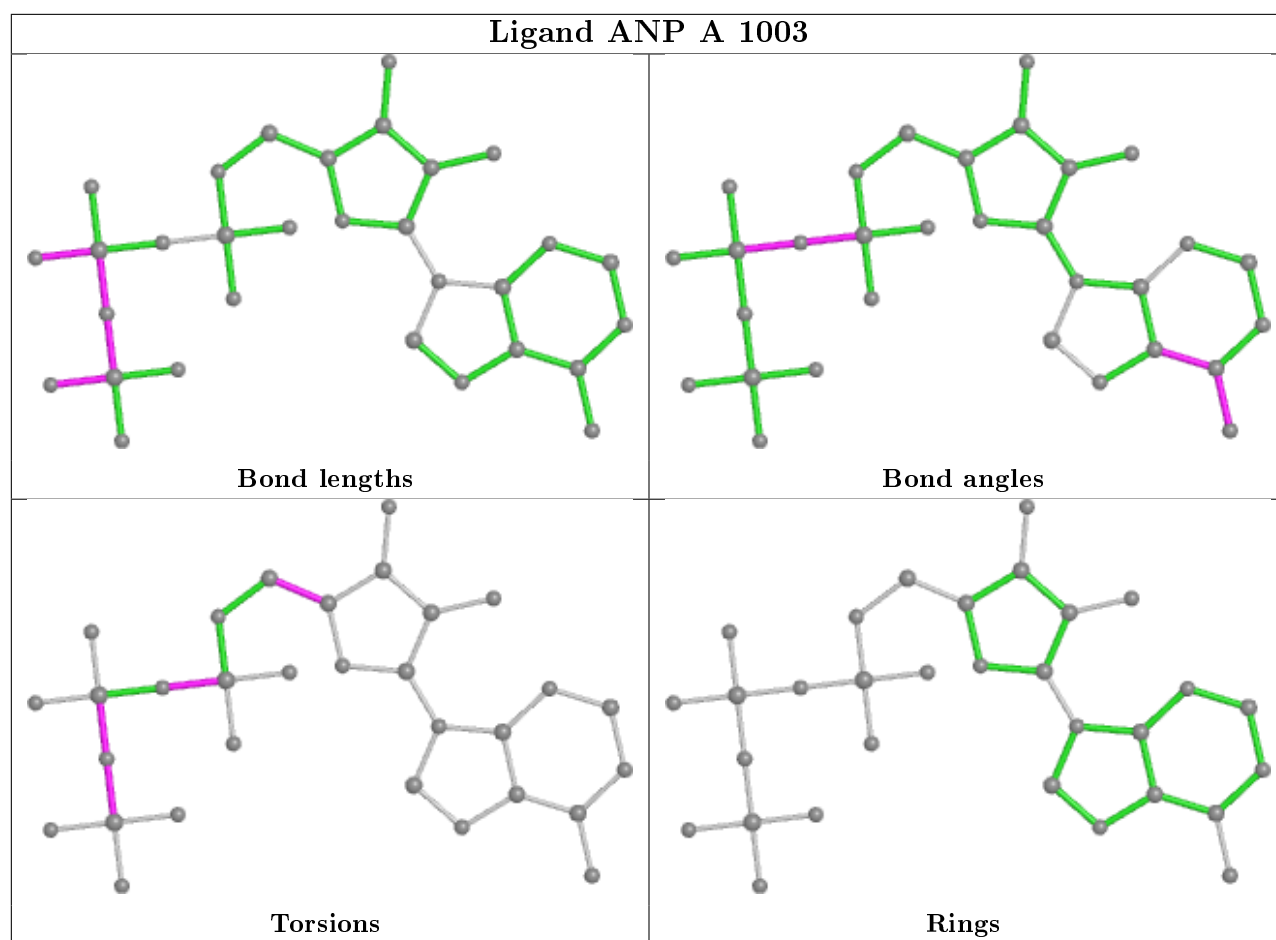
4 monomers are involved in 23 short contacts:

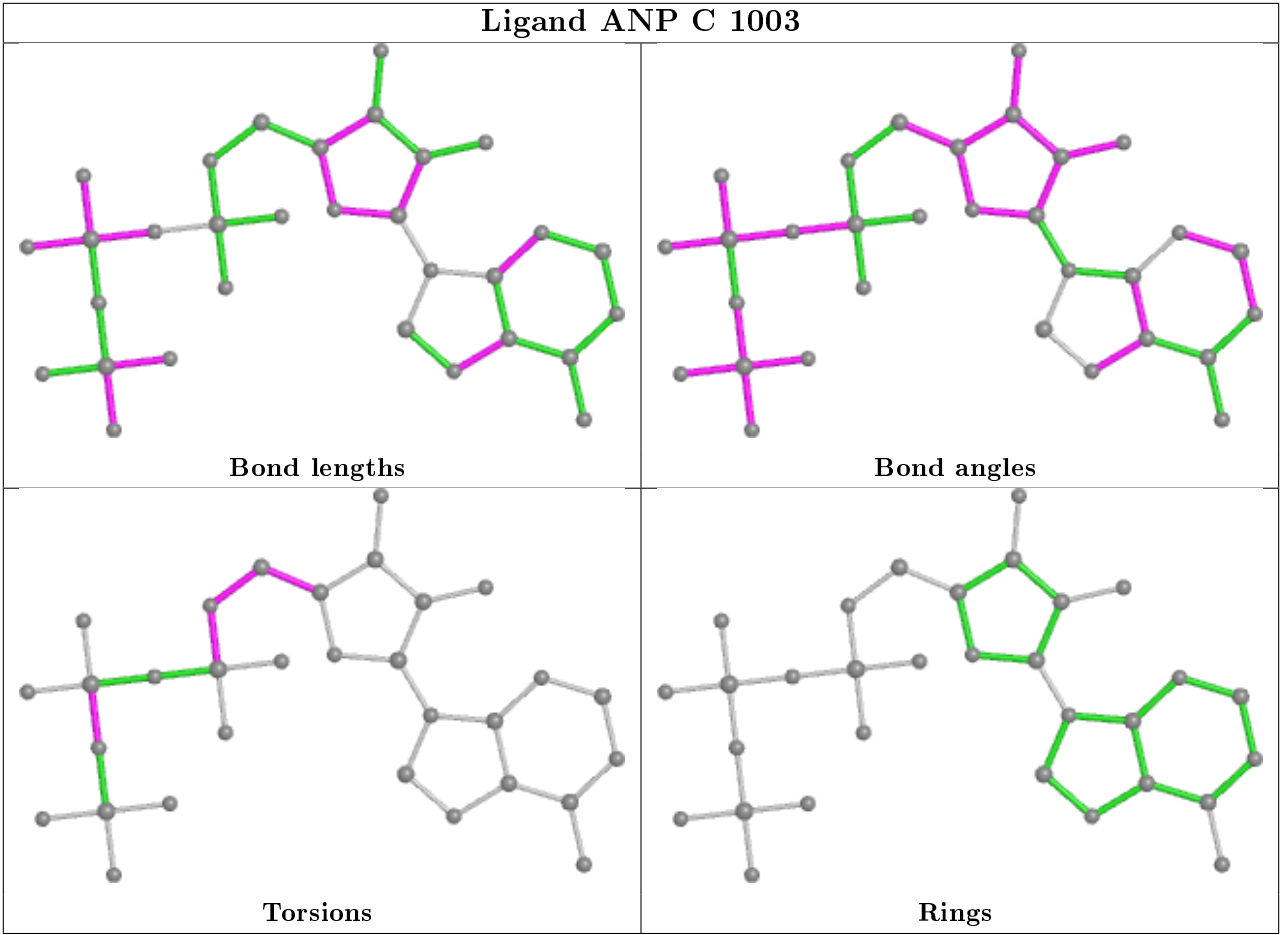
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	1003	ANP	5	0
4	G	1003	ANP	8	0
4	A	1003	ANP	4	0
4	C	1003	ANP	6	0

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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	481:LEU	C	482:SER	N	0.97

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	903/964 (93%)	-0.36	15 (1%) 70 71	18, 41, 83, 116	0
1	C	902/964 (93%)	-0.51	1 (0%) 95 96	22, 41, 66, 99	0
1	E	899/964 (93%)	-0.46	4 (0%) 92 93	17, 37, 70, 97	0
1	G	892/964 (92%)	-0.28	11 (1%) 79 80	23, 52, 89, 108	0
2	B	11/12 (91%)	0.22	1 (9%) 9 7	33, 44, 86, 103	0
2	D	11/12 (91%)	-0.09	0 100 100	35, 42, 83, 90	0
2	F	11/12 (91%)	0.34	1 (9%) 9 7	28, 39, 92, 94	0
2	H	11/12 (91%)	0.31	0 100 100	40, 56, 99, 101	0
All	All	3640/3904 (93%)	-0.40	33 (0%) 84 85	17, 42, 81, 116	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	846	ASN	5.1
1	A	864	ARG	3.3
1	A	868	PHE	3.1
1	G	867	ARG	3.1
1	A	923	ALA	3.1

### 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 carbohydrates 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, 95<sup>th</sup> 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.

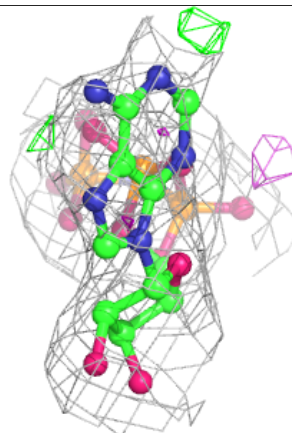
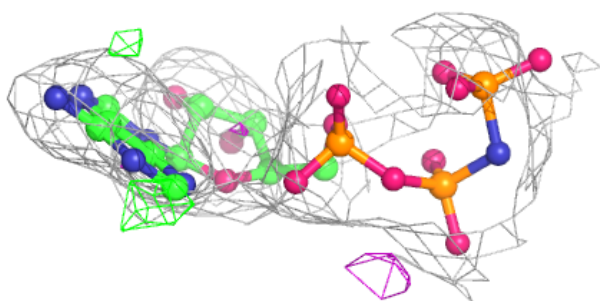
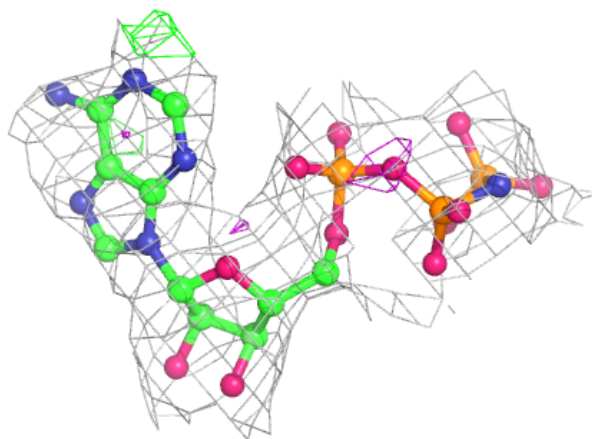
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	ANP	G	1003	31/31	0.79	0.18	70,94,121,141	0
4	ANP	E	1003	31/31	0.81	0.18	63,72,92,107	0
4	ANP	A	1003	31/31	0.90	0.15	51,67,83,115	0
4	ANP	C	1003	31/31	0.91	0.11	59,71,93,103	0
3	FE	A	1001	1/1	0.99	0.17	28,28,28,28	0
3	FE	G	1001	1/1	0.99	0.16	34,34,34,34	0
3	FE	C	1001	1/1	0.99	0.15	27,27,27,27	0
3	FE	E	1002	1/1	0.99	0.16	23,23,23,23	0
3	FE	G	1002	1/1	0.99	0.16	31,31,31,31	0
3	FE	A	1002	1/1	0.99	0.16	24,24,24,24	0
3	FE	E	1001	1/1	0.99	0.16	27,27,27,27	0
3	FE	C	1002	1/1	0.99	0.15	27,27,27,27	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.

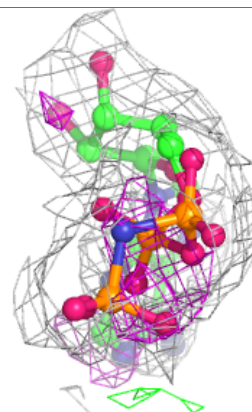
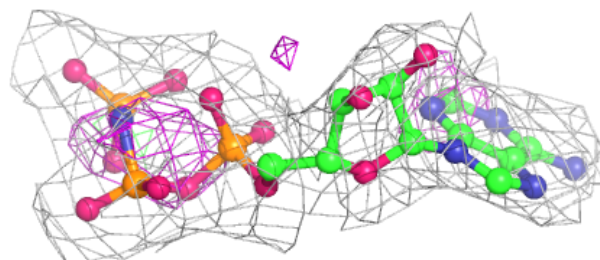
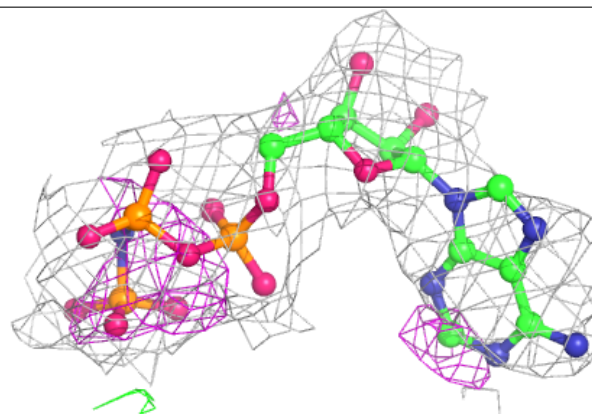


**Electron density around ANP G 1003:**

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

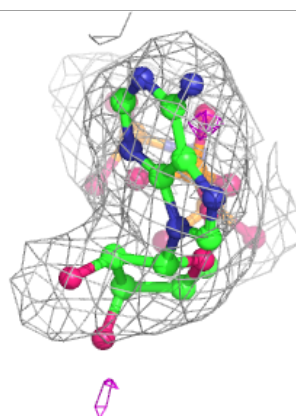
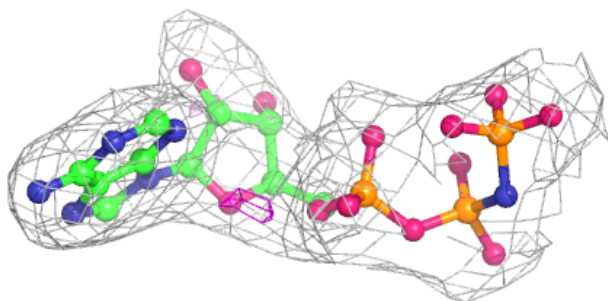
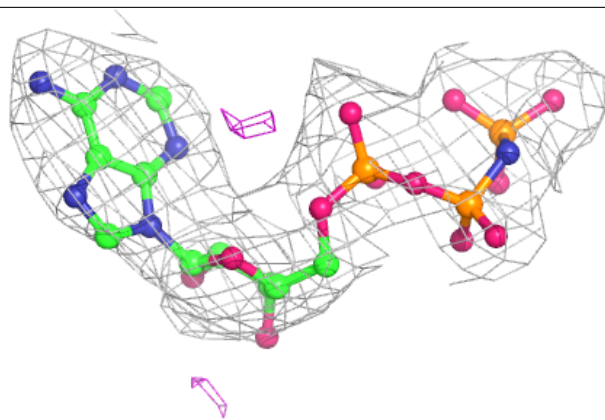
**Electron density around ANP E 1003:**

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

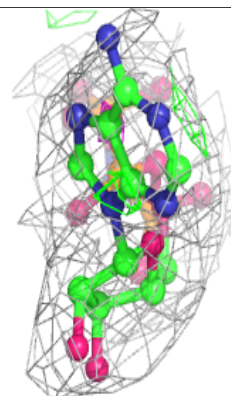
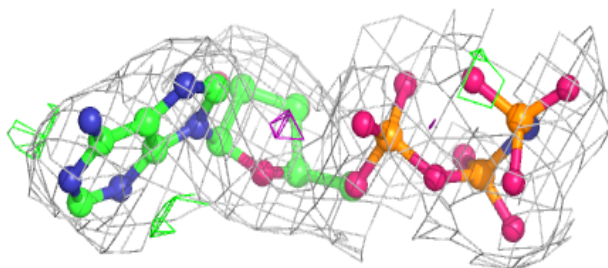
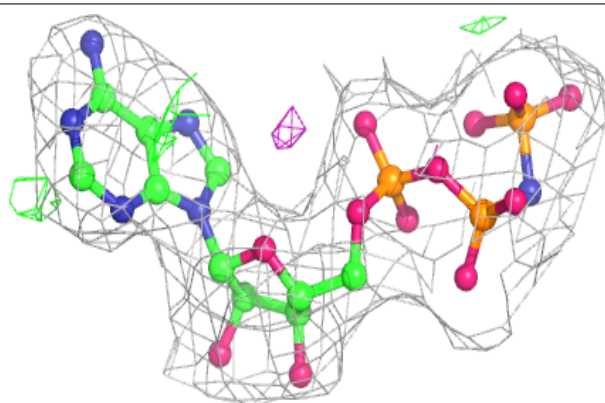


**Electron density around ANP A 1003:**

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

**Electron density around ANP C 1003:**

$2mF_o-DF_c$  (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.