WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Apr 28, 2024 – 06:55 am BST

PDB ID	:	4W9M
Title	:	AMPPNP bound Rad50 in complex with dsDNA
Authors	:	Rojowska, A.; Lammens, K.
Deposited on	:	2014-08-27
Resolution	:	2.70 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	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
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)	:	2.36.2

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 2.70 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.
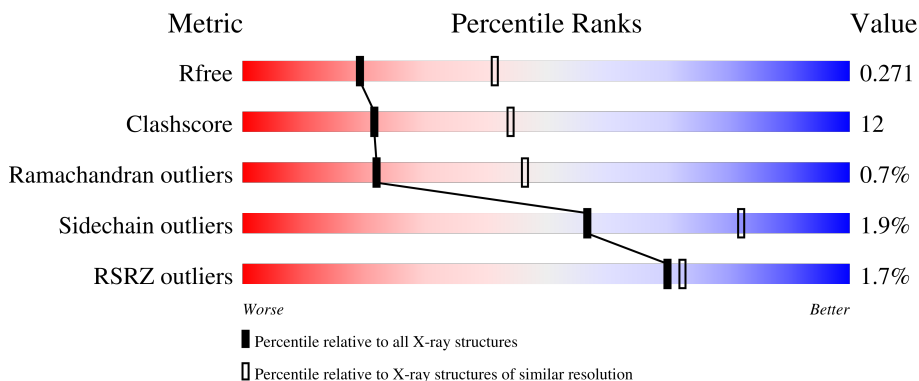
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	Y	15	27% / 60% / 33% / 7%
1	Z	15	27% / 87% / 13%
2	A	15	20% / 47% / 47% / 7%
2	G	15	27% / 47% / 40% / 13%
3	C	365	70% / 25% / · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	365	73% 22%
3	I	365	78% 18%
3	K	365	77% 19%
4	D	37	95%
4	F	37	73% 24%
4	J	37	73% 22%
4	L	37	68% 27% 5%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 14169 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 DNA chain called DNA (5'-D(*GP*GP*TP*CP*GP*GP*TP*GP*AP*CP* CP*GP*AP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Y	15	Total 306	C 145	N 59	O 88	P 14	0	0	0
1	Z	15	Total 306	C 145	N 59	O 88	P 14	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*GP*GP*TP*CP*GP*GP*TP*CP*AP*CP* CP*GP*AP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	15	Total 303	C 144	N 57	O 88	P 14	0	0	0
2	G	15	Total 303	C 144	N 57	O 88	P 14	0	0	0

- Molecule 3 is a protein called Probable DNA double-strand break repair Rad50 ATPase,Probable DNA double-strand break repair Rad50 ATPase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	352	Total 2824	C 1799	N 490	O 531	S 4	256	0	0
3	E	351	Total 2815	C 1794	N 489	O 528	S 4	215	0	0
3	I	352	Total 2824	C 1799	N 490	O 531	S 4	211	0	0
3	K	352	Total 2821	C 1797	N 490	O 530	S 4	239	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	676	SER	-	linker	UNP Q9X1X1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	677	SER	-	linker	UNP Q9X1X1
C	678	GLY	-	linker	UNP Q9X1X1
C	679	GLY	-	linker	UNP Q9X1X1
C	680	ALA	-	linker	UNP Q9X1X1
C	681	GLY	-	linker	UNP Q9X1X1
C	682	GLY	-	linker	UNP Q9X1X1
C	683	ALA	-	linker	UNP Q9X1X1
C	684	GLY	-	linker	UNP Q9X1X1
C	685	GLY	-	linker	UNP Q9X1X1
C	686	SER	-	linker	UNP Q9X1X1
E	676	SER	-	linker	UNP Q9X1X1
E	677	SER	-	linker	UNP Q9X1X1
E	678	GLY	-	linker	UNP Q9X1X1
E	679	GLY	-	linker	UNP Q9X1X1
E	680	ALA	-	linker	UNP Q9X1X1
E	681	GLY	-	linker	UNP Q9X1X1
E	682	GLY	-	linker	UNP Q9X1X1
E	683	ALA	-	linker	UNP Q9X1X1
E	684	GLY	-	linker	UNP Q9X1X1
E	685	GLY	-	linker	UNP Q9X1X1
E	686	SER	-	linker	UNP Q9X1X1
I	676	SER	-	linker	UNP Q9X1X1
I	677	SER	-	linker	UNP Q9X1X1
I	678	GLY	-	linker	UNP Q9X1X1
I	679	GLY	-	linker	UNP Q9X1X1
I	680	ALA	-	linker	UNP Q9X1X1
I	681	GLY	-	linker	UNP Q9X1X1
I	682	GLY	-	linker	UNP Q9X1X1
I	683	ALA	-	linker	UNP Q9X1X1
I	684	GLY	-	linker	UNP Q9X1X1
I	685	GLY	-	linker	UNP Q9X1X1
I	686	SER	-	linker	UNP Q9X1X1
K	189	SER	-	linker	UNP Q9X1X1
K	677	SER	-	linker	UNP Q9X1X1
K	678	GLY	-	linker	UNP Q9X1X1
K	679	GLY	-	linker	UNP Q9X1X1
K	680	ALA	-	linker	UNP Q9X1X1
K	681	GLY	-	linker	UNP Q9X1X1
K	682	GLY	-	linker	UNP Q9X1X1
K	683	ALA	-	linker	UNP Q9X1X1
K	684	GLY	-	linker	UNP Q9X1X1
K	685	GLY	-	linker	UNP Q9X1X1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	686	SER	-	linker	UNP Q9X1X1

- Molecule 4 is a protein called Exonuclease, putative.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	36	Total 319	C 209	N 50	O 60	81	0	0
4	F	36	Total 321	C 210	N 50	O 61	82	0	0
4	J	36	Total 319	C 209	N 50	O 60	83	0	0
4	L	35	Total 313	C 206	N 49	O 58	102	0	0

- Molecule 5 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).
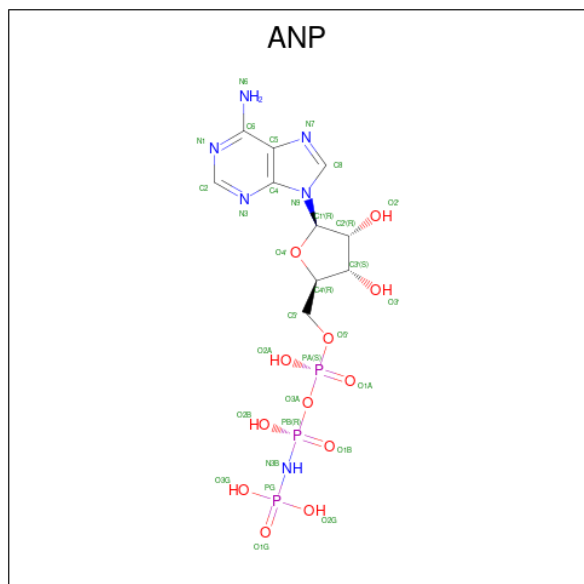
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	C	1	Total 31	C 10	N 6	O 12	P 3	0	0
5	E	1	Total 31	C 10	N 6	O 12	P 3	0	0
5	I	1	Total 31	C 10	N 6	O 12	P 3	0	0
5	K	1	Total 31	C 10	N 6	O 12	P 3	0	0

• Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	1	Total 1	Mg 1	0	0
6	E	1	Total 1	Mg 1	0	0
6	I	1	Total 1	Mg 1	0	0
6	K	1	Total 1	Mg 1	0	0

• Molecule 7 is water.

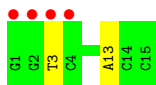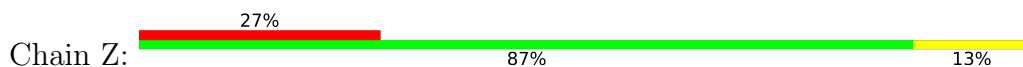| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	Y	2	Total 2	O 2	0	0
7	C	64	Total 64	O 64	0	0
7	D	1	Total 1	O 1	0	0
7	E	63	Total 63	O 63	0	0
7	F	1	Total 1	O 1	0	0
7	G	1	Total 1	O 1	0	0
7	I	67	Total 67	O 67	0	0
7	J	2	Total 2	O 2	0	0
7	K	63	Total 63	O 63	0	0
7	L	3	Total 3	O 3	0	0

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: DNA (5'-D(*GP*GP*TP*CP*GP*GP*TP*GP*AP*CP*CP*GP*AP*CP*C)-3')

Chain Y:
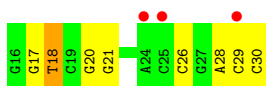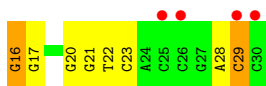

• Molecule 1: DNA (5'-D(*GP*GP*TP*CP*GP*GP*TP*GP*AP*CP*CP*GP*AP*CP*C)-3')

Chain Z:

• Molecule 2: DNA (5'-D(*GP*GP*TP*CP*GP*GP*TP*CP*AP*CP*CP*GP*AP*CP*C)-3')
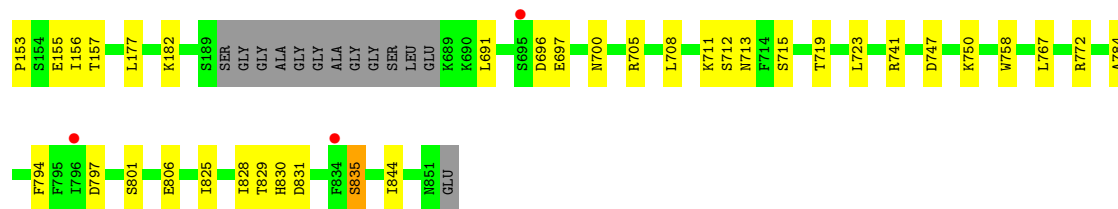
Chain A:

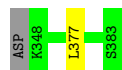• Molecule 2: DNA (5'-D(*GP*GP*TP*CP*GP*GP*TP*CP*AP*CP*CP*GP*AP*CP*C)-3')

Chain G:

• Molecule 3: Probable DNA double-strand break repair Rad50 ATPase,Probable DNA double-strand break repair Rad50 ATPase

Chain C:

M1 R2 P3 L6 R9 N10 F11 S23 G24 V27 V28 E29 G33 S37 S38 L39 F40 N50 G51 I52 R53 Y54 P55 Y60 T69 E77 R78 G79 G80 K81 L92 K108 I112 K115 S118 Q121 E122 K125 I129 R132

V138 F139 L140 P141 Q142 G143 E144 K147 L148 L149 I150 S151 P152 P153 S154 E155 I156 I160 S161 F164 Q165 E168 L177 M181 I188 SER GLY GLY ALA GLY GLY ALA GLY SER LEU E688 K692 D696 N713 L723 L727 K731 L737

D743 G752 F753 I754 I755 K756 R765 G766 L767 I777 L783 A784 E785 V786 R790 L791 D792 A793 F794 F795 I796 D797 E798 G799 F800 S801 S802 L803 D804 K808 I811 V814 L815 K823 V824 I828 T829 R832 E833 F834 S835 L842 N851 GLU

• Molecule 3: Probable DNA double-strand break repair Rad50 ATPase,Probable DNA double-strand break repair Rad50 ATPase

Chain E:

M1 V8 L14 V17 P31 N32 G33 A34 S37 E41 Y54 P55 Y60 E77 R78 K81 R82 R94 A98 K99 L100 S101 E102 I103 A110 A111 I112 K115 P116 T117 Q121 K125 I126 L127 T137 V138 F139 L140 P141 K147 L148 L149

I150 S151 P152 P153 T157 V163 E168 E174 K175 L176 L177 E185 N186 E187 I188 SER SER GLY GLY ALA GLY GLY ALA GLY SER LEU GLU K689 K690 L691 M694 E697 Y698 L703 L704 R705 F709 D710 K711 S712 N713 T719 L723 K731 T738

F742 D743 I744 D745 F753 K756 R765 S768 R772 I775 L783 R790 L791 D797 E798 G799 F800 L803 D804 T805 E806 N807 K808 E809 K810 I811 L821 I828 T829 V849 V850 N851 GLU

• Molecule 3: Probable DNA double-strand break repair Rad50 ATPase,Probable DNA double-strand break repair Rad50 ATPase

Chain I:

M1 R9 L12 I25 G33 S38 F48 G51 I52 R53 Y54 Y60 T69 R78 Y83 I89 L92 N97 G107 K108 K109 I112 E122 V123 E124 K125 I126 L127 I135 V138 F139 Q142 F164 Q165 S166 K167 E168 T169

L173 L176 L177 K178 E179 K180 M181 K182 K183 L184 I188 SER SER GLY GLY ALA GLY GLY ALA GLY SER LEU E688 L704 N713 R721 V722 V726 R729 T730 K731 A732 Y733 T738 R741 F742 D743 I744 D747 K750 I755 K756 D757 G766 L767

L774 I777 M781 S782 V786 L791 F795 I796 D797 E798 S801 I811 L821 I825 I828 N851 GLU

• Molecule 3: Probable DNA double-strand break repair Rad50 ATPase,Probable DNA double-strand break repair Rad50 ATPase

Chain K:

M1 E4 F11 L12 G13 L14 V27 V28 E29 S37 S38 L39 L47 R53 Y54 P55 N56 D59 Y60 R63 F74 Q75 R82 L100 I103 A113 A114 K115 S118 L127 G128 I129 R132 T133 F134 V138 D146 K147 I150 S151 P152

P153 S154 E155 I156 T157 L177 K182 S189 SER GLY GLY ALA GLY GLY ALA GLY GLY SER LEU GLU K689 K690 L691 S695 D696 E697 N700 R705 L708 K711 S712 N713 F714 S715 T719 L723 R741 D747 K750 W758 L767 R772 A784

F794 F795 I796 D797 S801 E806 I825 I828 T829 H830 D831 F834 S835 I844 N851 GLU

- Molecule 4: Exonuclease, putative

Chain D:

95%

- Molecule 4: Exonuclease, putative

ASP K348 L377 S383

Chain F:

73% 24%

D347 F352 E353 L354 F355 K356 L359 E363 L369 I376 L377 D378 E379 V380 K381 K382 SER

- Molecule 4: Exonuclease, putative

Chain J:

73% 22%

ASP K348 L349 K360 E363 E364 N365 H366 E367 K368 L369 I372 S383

- Molecule 4: Exonuclease, putative

Chain L:

68% 27% 5%

ASP K348 F355 L359 R362 E363 E364 N365 K368 I372 L377 K381 K382 SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	49.71Å 96.39Å 105.93Å 90.01° 89.22° 81.67°	Depositor
Resolution (Å)	47.69 – 2.70 47.68 – 2.70	Depositor EDS
% Data completeness (in resolution range)	96.3 (47.69-2.70) 92.4 (47.68-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 2.69Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.215 , 0.270 0.217 , 0.271	Depositor DCC
R_{free} test set	2573 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	43.5	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 29.7	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.368 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	14169	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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 (i)

5.1 Standard geometry (i)

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	Y	0.55	0/343	1.18	3/528 (0.6%)
1	Z	0.61	0/343	1.23	2/528 (0.4%)
2	A	0.54	0/339	1.26	2/521 (0.4%)
2	G	0.58	0/339	1.34	3/521 (0.6%)
3	C	0.49	1/2865 (0.0%)	0.65	5/3840 (0.1%)
3	E	0.61	5/2856 (0.2%)	0.77	6/3828 (0.2%)
3	I	0.64	2/2865 (0.1%)	0.87	4/3840 (0.1%)
3	K	0.54	0/2862	0.66	0/3836
4	D	0.29	0/323	0.42	0/427
4	F	0.31	0/325	0.49	0/430
4	J	0.31	0/323	0.45	0/427
4	L	0.31	0/317	0.54	0/419
All	All	0.56	8/14100 (0.1%)	0.80	25/19145 (0.1%)

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
3	E	0	3
3	I	0	1
All	All	0	4

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	108	LYS	C-N	-20.86	0.86	1.34
3	I	109	LYS	C-N	-7.37	1.17	1.34
3	E	54	TYR	CE1-CZ	-7.04	1.29	1.38
3	E	711	LYS	C-N	-5.65	1.21	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	28	VAL	C-N	-5.48	1.21	1.34

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	I	109	LYS	O-C-N	-29.63	75.29	122.70
3	I	109	LYS	CA-C-N	19.99	161.17	117.20
3	E	712	SER	O-C-N	-13.39	101.28	122.70
3	I	109	LYS	C-N-CA	12.67	153.38	121.70
3	E	711	LYS	O-C-N	-12.37	102.91	122.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	E	711	LYS	Mainchain,Peptide
3	E	799	GLY	Peptide
3	I	109	LYS	Mainchain

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	Y	306	0	169	5	0
1	Z	306	0	169	0	0
2	A	303	0	169	4	1
2	G	303	0	169	5	0
3	C	2824	0	2891	82	5
3	E	2815	0	2886	57	4
3	I	2824	0	2891	61	0
3	K	2821	0	2891	57	0
4	D	319	0	332	1	0
4	F	321	0	331	5	0
4	J	319	0	332	7	0
4	L	313	0	327	7	0
5	C	31	0	13	2	0
5	E	31	0	13	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	I	31	0	13	2	0
5	K	31	0	13	1	0
6	C	1	0	0	0	0
6	E	1	0	0	0	0
6	I	1	0	0	0	0
6	K	1	0	0	0	0
7	C	64	0	0	7	0
7	D	1	0	0	0	0
7	E	63	0	0	8	0
7	F	1	0	0	0	0
7	G	1	0	0	0	0
7	I	67	0	0	8	0
7	J	2	0	0	0	0
7	K	63	0	0	9	0
7	L	3	0	0	0	0
7	Y	2	0	0	0	0
All	All	14169	0	13609	286	5

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 12.

The worst 5 of 286 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)
3:I:108:LYS:C	3:I:109:LYS:CA	2.11	1.18
3:I:108:LYS:O	3:I:109:LYS:N	1.77	1.16
3:C:142:GLN:HG3	3:C:798:GLU:CB	1.79	1.13
3:I:798:GLU:CD	7:I:1001:HOH:O	1.86	1.10
3:I:108:LYS:CA	3:I:109:LYS:N	2.14	1.08

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:833:GLU:CA	3:E:94:ARG:NH1[1_455]	1.33	0.87
3:C:833:GLU:N	3:E:94:ARG:NH2[1_455]	1.87	0.33
3:C:833:GLU:N	3:E:94:ARG:NH1[1_455]	1.93	0.27
2:A:26:DC:OP2	3:C:108:LYS:NZ[1_655]	2.00	0.20
3:C:833:GLU:N	3:E:94:ARG:CZ[1_455]	2.10	0.10

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	
3	C	348/365 (95%)	327 (94%)	17 (5%)	4 (1%)	14	34
3	E	347/365 (95%)	331 (95%)	14 (4%)	2 (1%)	25	50
3	I	348/365 (95%)	329 (94%)	18 (5%)	1 (0%)	41	66
3	K	348/365 (95%)	325 (93%)	19 (6%)	4 (1%)	14	34
4	D	34/37 (92%)	32 (94%)	2 (6%)	0	100	100
4	F	34/37 (92%)	32 (94%)	2 (6%)	0	100	100
4	J	34/37 (92%)	32 (94%)	2 (6%)	0	100	100
4	L	33/37 (89%)	29 (88%)	4 (12%)	0	100	100
All	All	1526/1608 (95%)	1437 (94%)	78 (5%)	11 (1%)	22	46

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	713	ASN
3	K	712	SER
3	C	51	GLY
3	C	713	ASN
3	K	713	ASN

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	
3	C	306/311 (98%)	302 (99%)	4 (1%)	69	87

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	E	305/311 (98%)	297 (97%)	8 (3%)	46	75
3	I	306/311 (98%)	300 (98%)	6 (2%)	55	81
3	K	306/311 (98%)	300 (98%)	6 (2%)	55	81
4	D	36/37 (97%)	36 (100%)	0	100	100
4	F	36/37 (97%)	35 (97%)	1 (3%)	43	73
4	J	36/37 (97%)	35 (97%)	1 (3%)	43	73
4	L	35/37 (95%)	35 (100%)	0	100	100
All	All	1366/1392 (98%)	1340 (98%)	26 (2%)	57	82

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	89	ILE
3	I	123	VAL
3	K	835	SER
3	I	122	GLU
3	I	135	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	50	ASN
3	E	75	GLN
3	E	830	HIS
4	L	365	ASN

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 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	ANP	I	901	6	29,33,33	1.87	2 (6%)	31,52,52	1.50	5 (16%)
5	ANP	K	901	6	29,33,33	1.65	5 (17%)	31,52,52	1.29	4 (12%)
5	ANP	C	901	6	29,33,33	2.22	5 (17%)	31,52,52	0.96	2 (6%)
5	ANP	E	901	6	29,33,33	1.62	6 (20%)	31,52,52	1.28	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	ANP	I	901	6	-	1/14/38/38	0/3/3/3
5	ANP	K	901	6	-	5/14/38/38	0/3/3/3
5	ANP	C	901	6	-	1/14/38/38	0/3/3/3
5	ANP	E	901	6	-	1/14/38/38	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	901	ANP	PG-O1G	8.72	1.60	1.46
5	C	901	ANP	PG-O1G	7.94	1.58	1.46
5	C	901	ANP	PB-O1B	6.36	1.56	1.46
5	K	901	ANP	PB-O1B	5.85	1.55	1.46
5	E	901	ANP	PB-O1B	5.77	1.55	1.46

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	I	901	ANP	O1B-PB-N3B	-4.60	105.00	111.77
5	I	901	ANP	O1G-PG-N3B	3.42	116.81	111.77

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	K	901	ANP	O3G-PG-O1G	-3.25	105.29	113.45
5	E	901	ANP	O2G-PG-O1G	-3.19	105.43	113.45
5	E	901	ANP	O3G-PG-O1G	-3.13	105.57	113.45

There are no chirality outliers.
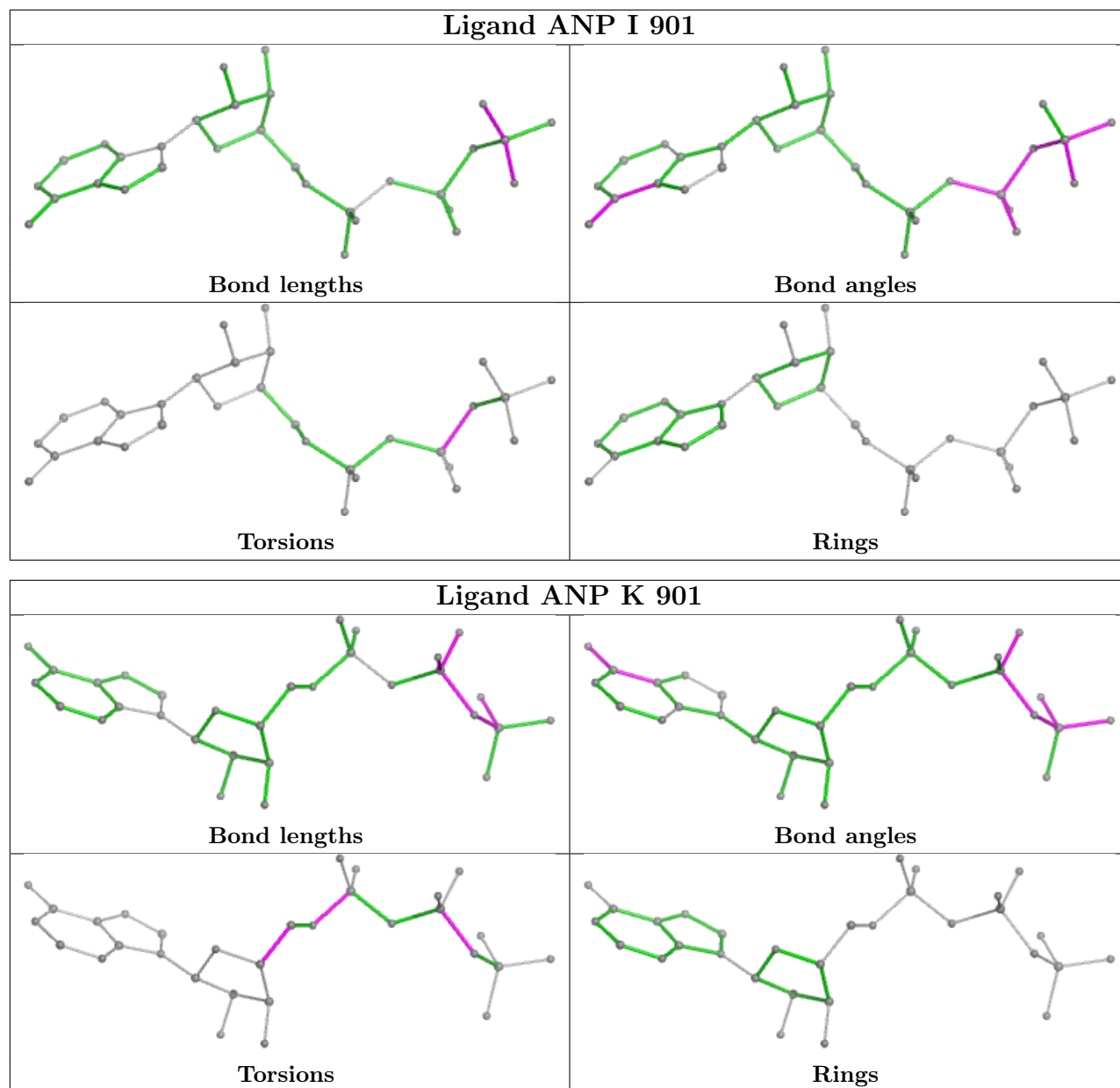
5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	901	ANP	PA-O3A-PB-O1B
5	E	901	ANP	PB-N3B-PG-O1G
5	I	901	ANP	PG-N3B-PB-O1B
5	K	901	ANP	PG-N3B-PB-O3A
5	K	901	ANP	C3'-C4'-C5'-O5'

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	I	901	ANP	2	0
5	K	901	ANP	1	0
5	C	901	ANP	2	0
5	E	901	ANP	3	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 I 901

Bond lengths

Bond angles

Torsions

Rings

Ligand ANP K 901

Bond lengths

Bond angles

Torsions

Rings

Ligand ANP C 901

Bond lengths

Bond angles

Torsions

Rings

Ligand ANP E 901

Bond lengths

Bond angles

Torsions

Rings
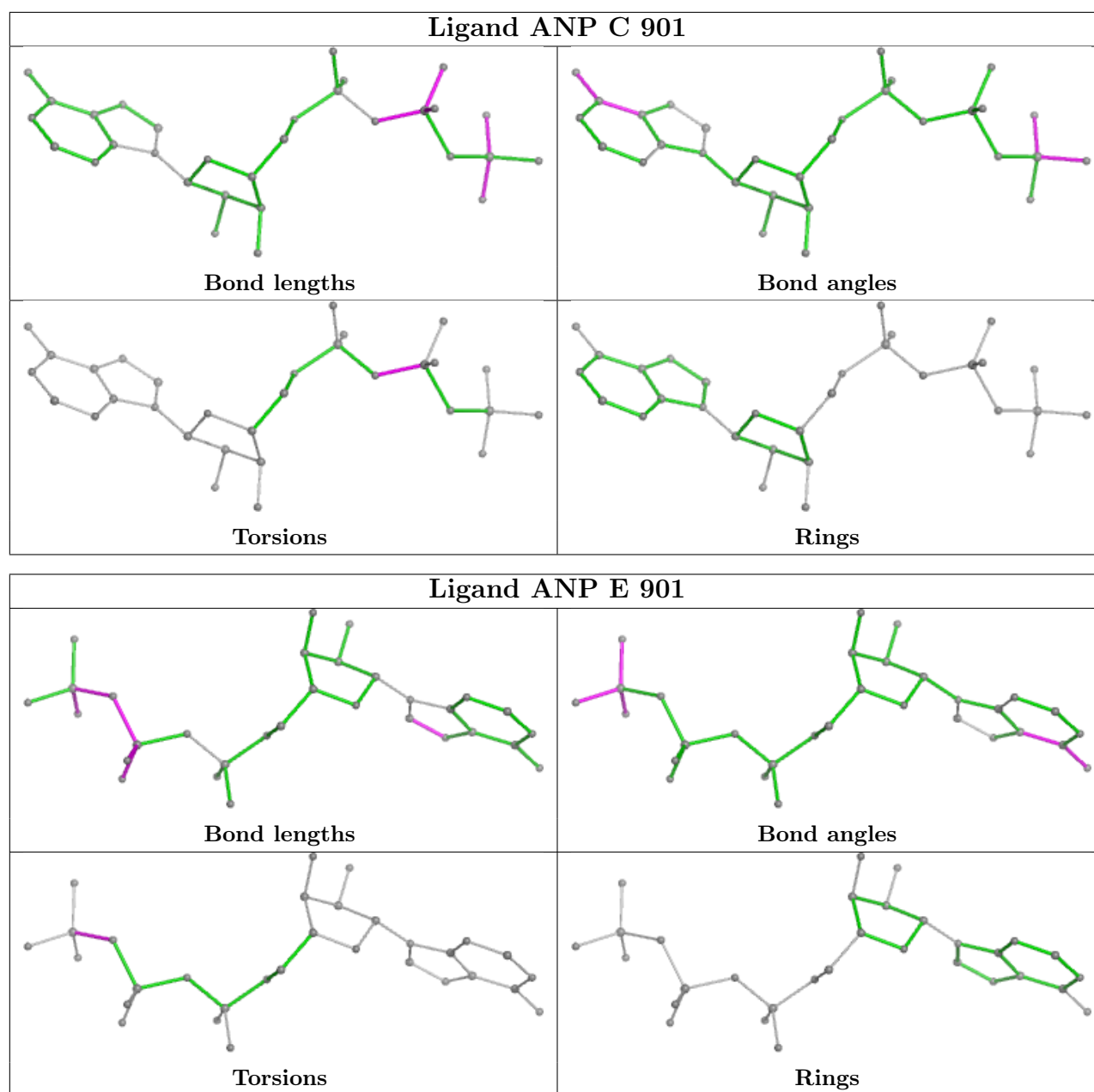
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	I	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	I	109:LYS	C	110:ALA	N	1.17
1	I	108:LYS	C	109:LYS	N	0.86

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	Y	15/15 (100%)	1.77	4 (26%)	0	0	136, 158, 186, 194	0
1	Z	15/15 (100%)	1.86	4 (26%)	0	0	122, 146, 183, 183	0
2	A	15/15 (100%)	1.02	3 (20%)	1	0	75, 140, 176, 177	0
2	G	15/15 (100%)	1.01	4 (26%)	0	0	64, 132, 167, 167	0
3	C	352/365 (96%)	-0.19	0	100	100	17, 29, 62, 82	94 (26%)
3	E	351/365 (96%)	-0.15	3 (0%)	84	85	17, 28, 59, 73	78 (22%)
3	I	352/365 (96%)	-0.15	5 (1%)	75	77	18, 29, 55, 89	67 (19%)
3	K	352/365 (96%)	-0.13	4 (1%)	80	82	16, 27, 57, 80	78 (22%)
4	D	36/37 (97%)	-0.41	0	100	100	41, 57, 72, 75	23 (63%)
4	F	36/37 (97%)	-0.45	0	100	100	43, 61, 74, 79	22 (61%)
4	J	36/37 (97%)	-0.45	0	100	100	36, 51, 65, 67	20 (55%)
4	L	35/37 (94%)	-0.34	0	100	100	45, 66, 80, 82	27 (77%)
All	All	1610/1668 (96%)	-0.12	27 (1%)	70	72	16, 30, 75, 194	409 (25%)

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	Z	1	DG	6.1
1	Y	2	DG	6.0
1	Y	1	DG	4.7
1	Z	2	DG	4.7
1	Z	4	DC	4.3

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 (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(Å^2)	Q<0.9
6	MG	C	902	1/1	0.96	0.13	15,15,15,15	0
6	MG	I	902	1/1	0.96	0.14	16,16,16,16	0
5	ANP	I	901	31/31	0.97	0.17	15,19,24,24	0
5	ANP	K	901	31/31	0.97	0.16	11,19,23,26	0
5	ANP	C	901	31/31	0.97	0.15	14,19,22,23	0
6	MG	E	902	1/1	0.97	0.14	14,14,14,14	0
5	ANP	E	901	31/31	0.97	0.15	11,19,23,24	0
6	MG	K	902	1/1	0.99	0.21	13,13,13,13	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 I 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
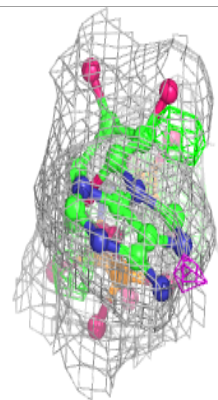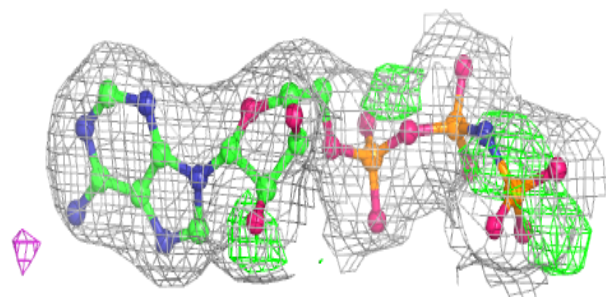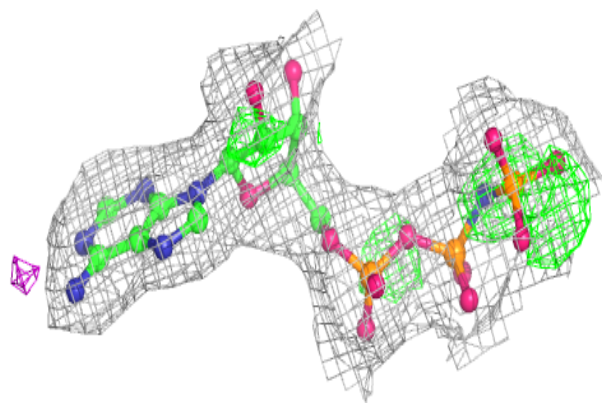
Electron density around ANP K 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
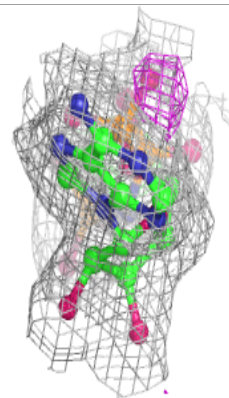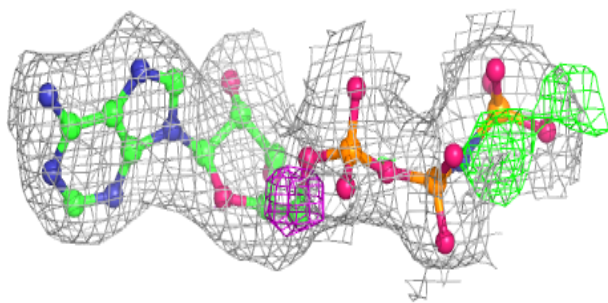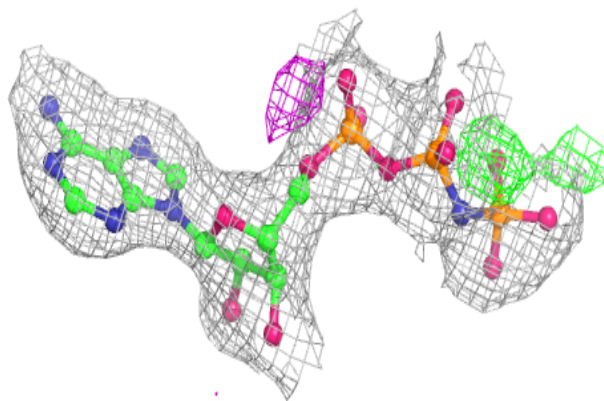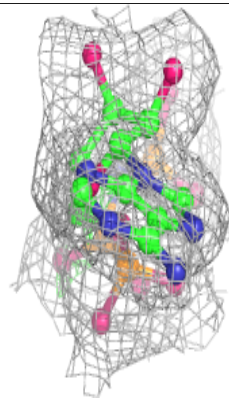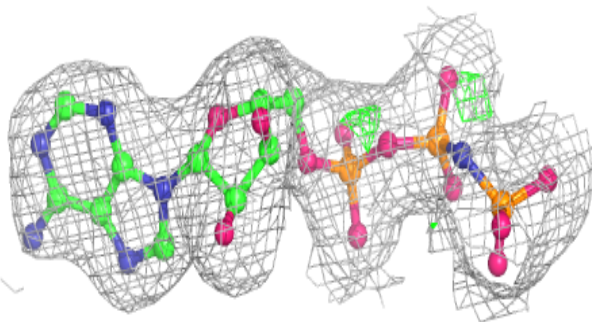mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
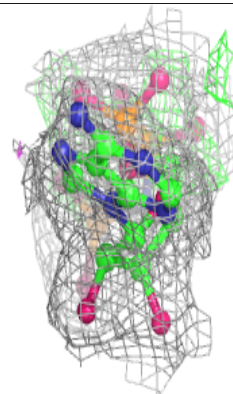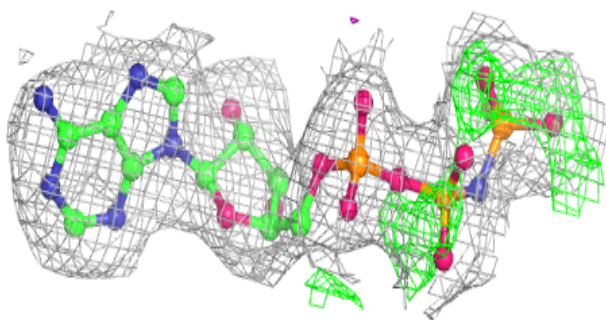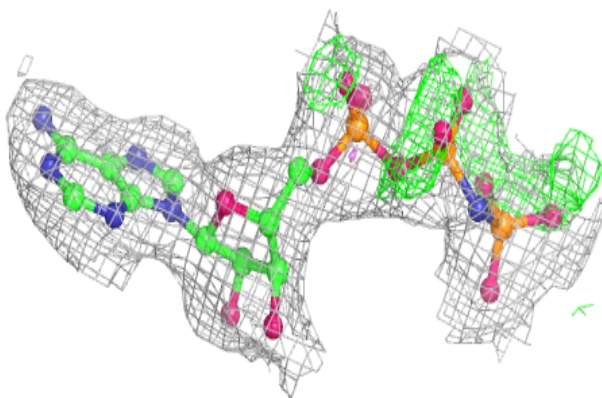
Electron density around ANP C 901:

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

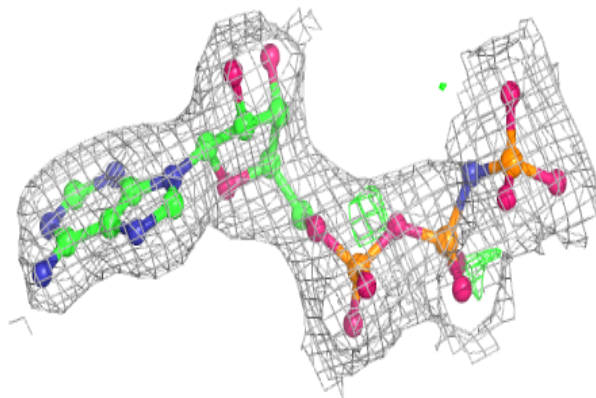**Electron density around ANP E 901:**

$2mF_o-DF_c$ (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.