wwPDB X-ray Structure Validation Summary Report (i)

Sep 12, 2023 – 07:22 PM EDT

PDB ID	:	4OGR
Title	:	crystal structure of P-TEFb complex with AFF4 and Tat
Authors	:	Schulze-Gahmen, U.; Alber, T.
Deposited on	:	2014-01-16
Resolution	:	3.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

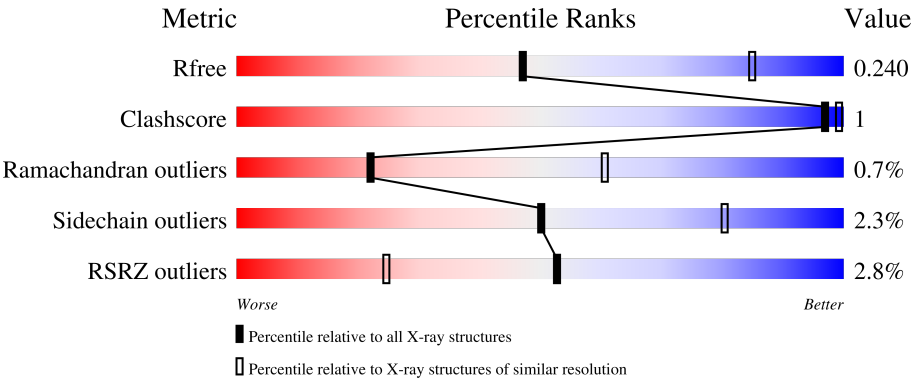
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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.



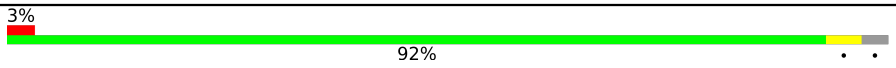
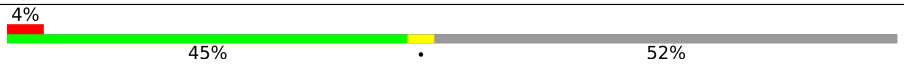
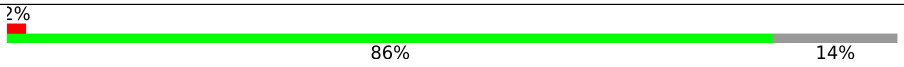
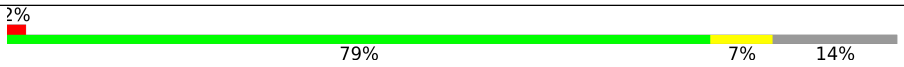
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	332	2% 91% · 5%
1	E	332	6% 87% 6% 6%
1	I	332	2% 91% 5% ·
2	B	264	% 92% 5% ·
2	F	264	4% 93% · · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	K	264	
3	C	75	
3	G	75	
3	L	75	
4	D	58	
4	H	58	
4	M	58	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 32026 atoms, of which 15872 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 9.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	315	Total	C	H	N	O	P	S	0	0	0
			5137	1628	2596	438	459	1	15			
1	E	311	Total	C	H	N	O	P	S	0	0	0
			5052	1604	2550	429	453	1	15			
1	I	319	Total	C	H	N	O	P	S	0	0	0
			5141	1636	2588	439	462	1	15			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P50750
A	0	HIS	-	expression tag	UNP P50750
E	-1	GLY	-	expression tag	UNP P50750
E	0	HIS	-	expression tag	UNP P50750
I	-1	GLY	-	expression tag	UNP P50750
I	0	HIS	-	expression tag	UNP P50750

- Molecule 2 is a protein called Cyclin-T1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	255	Total	C	H	N	O	S	0	0	0
			4081	1320	2020	351	380	10			
2	F	255	Total	C	H	N	O	S	0	0	0
			4067	1312	2014	352	379	10			
2	K	255	Total	C	H	N	O	S	0	0	0
			4065	1315	2007	353	380	10			

- Molecule 3 is a protein called AF4/FMR2 family member 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	55	Total	C	H	N	O	S	0	0	0
			829	265	391	83	87	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	G	36	Total	C	H	N	O	S	0	0	0
			488	167	220	42	57	2			
3	L	54	Total	C	H	N	O	S	0	0	0
			785	254	363	79	86	3			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	SER	-	expression tag	UNP Q9UHB7
C	0	ASN	-	expression tag	UNP Q9UHB7
C	1	ALA	-	expression tag	UNP Q9UHB7
G	-1	SER	-	expression tag	UNP Q9UHB7
G	0	ASN	-	expression tag	UNP Q9UHB7
G	1	ALA	-	expression tag	UNP Q9UHB7
L	-1	SER	-	expression tag	UNP Q9UHB7
L	0	ASN	-	expression tag	UNP Q9UHB7
L	1	ALA	-	expression tag	UNP Q9UHB7

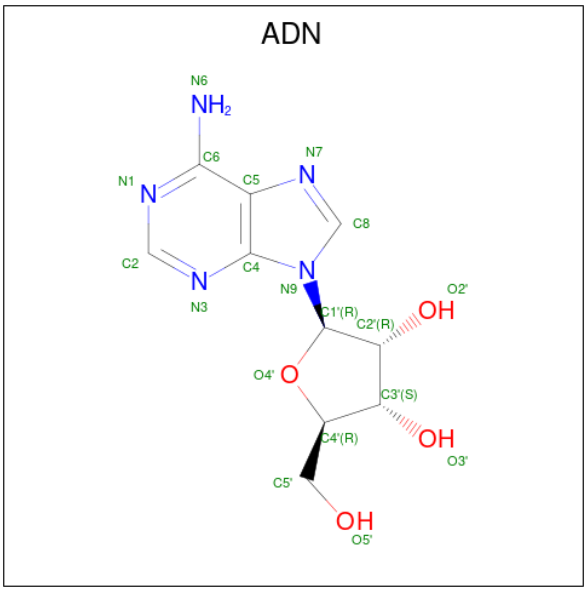
- Molecule 4 is a protein called Protein Tat.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	50	Total	C	H	N	O	S	0	0	0
			766	246	378	68	66	8			
4	H	50	Total	C	H	N	O	S	0	0	0
			749	243	367	65	66	8			
4	M	50	Total	C	H	N	O	S	0	0	0
			766	246	378	68	66	8			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	ACE	-	expression tag	UNP P69698
H	0	ACE	-	expression tag	UNP P69698
M	0	ACE	-	expression tag	UNP P69698

- Molecule 5 is ADENOSINE (three-letter code: ADN) (formula: $C_{10}H_{13}N_5O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O 19 10 5 4	0	0
5	E	1	Total C N O 19 10 5 4	0	0
5	I	1	Total C N O 19 10 5 4	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	2	Total Zn 2 2	0	0
6	H	2	Total Zn 2 2	0	0
6	M	2	Total Zn 2 2	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	9	Total O 9 9	0	0
7	B	5	Total O 5 5	0	0
7	C	1	Total O 1 1	0	0

Continued on next page...

Continued from previous page...

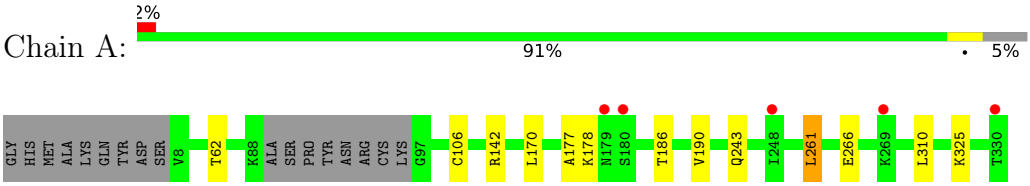
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	E	2	Total O 2 2	0	0
7	F	6	Total O 6 6	0	0
7	I	4	Total O 4 4	0	0
7	K	6	Total O 6 6	0	0
7	L	1	Total O 1 1	0	0
7	D	1	Total O 1 1	0	0
7	M	2	Total O 2 2	0	0

3 Residue-property plots (i)

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.

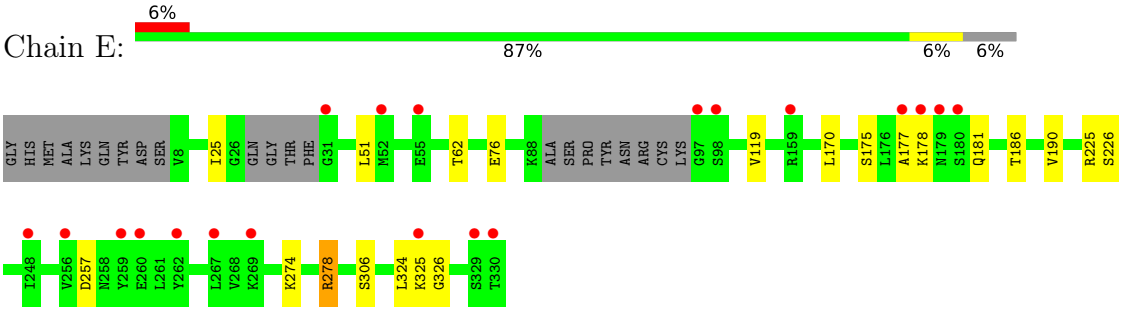
• Molecule 1: Cyclin-dependent kinase 9

Chain A:



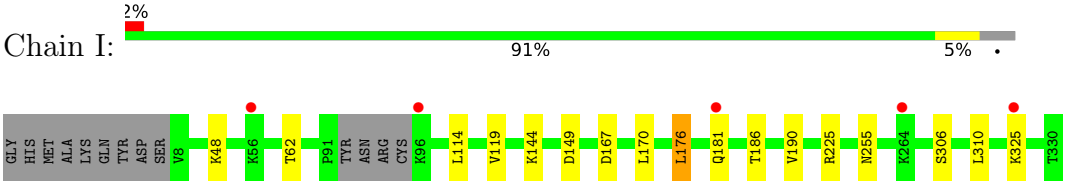
• Molecule 1: Cyclin-dependent kinase 9

Chain E:



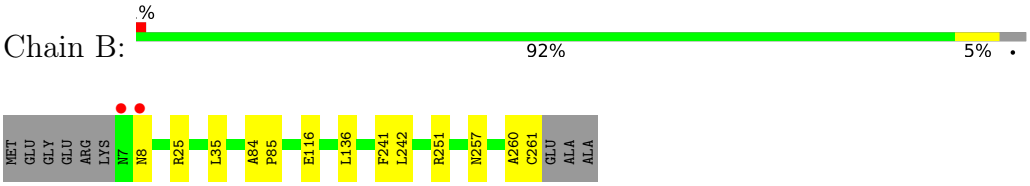
• Molecule 1: Cyclin-dependent kinase 9

Chain I:



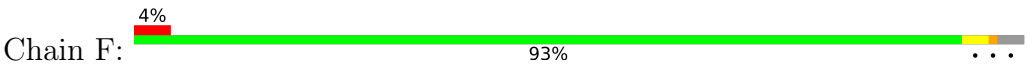
• Molecule 2: Cyclin-T1

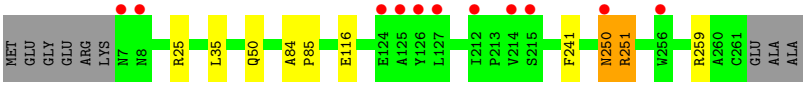
Chain B:



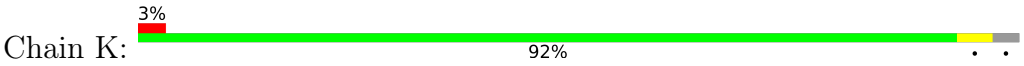
• Molecule 2: Cyclin-T1

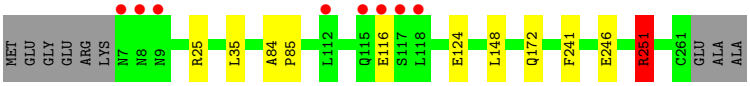
Chain F:



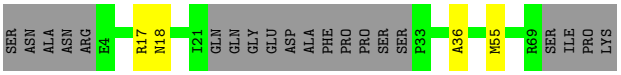


• Molecule 2: Cyclin-T1

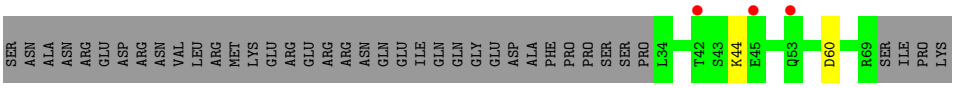




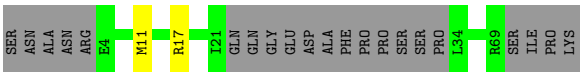
• Molecule 3: AF4/FMR2 family member 4



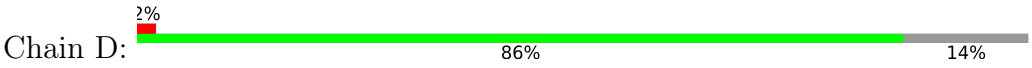
• Molecule 3: AF4/FMR2 family member 4

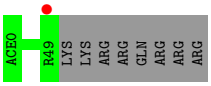


• Molecule 3: AF4/FMR2 family member 4

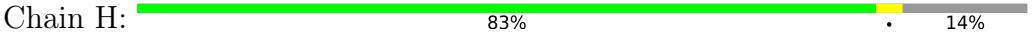


• Molecule 4: Protein Tat

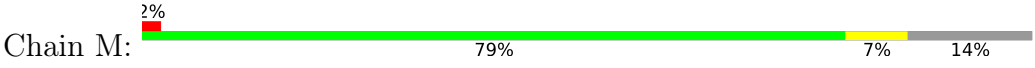


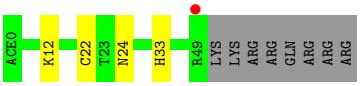


• Molecule 4: Protein Tat



• Molecule 4: Protein Tat




ACE0
K12
C22
T23
N24
H33
R49
LYS
LYS
ARG
ARG
GLN
ARG
ARG
ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	184.91Å 184.91Å 360.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.02 – 3.00 49.01 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.02-3.00) 100.0 (49.01-3.00)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 3.01Å)	Xtriage
Refinement program	PHENIX dev_1419	Depositor
R, R_{free}	0.205 , 0.232 0.214 , 0.240	Depositor DCC
R_{free} test set	2406 reflections (3.28%)	wwPDB-VP
Wilson B-factor (Å^2)	77.9	Xtriage
Anisotropy	0.339	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 58.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	32026	wwPDB-VP
Average B, all atoms (Å^2)	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.43% 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: ACE, ZN, ADN, TPO

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	0.23	0/2580	0.43	0/3482
1	E	0.23	0/2539	0.43	0/3427
1	I	0.23	0/2593	0.42	0/3504
2	B	0.23	0/2115	0.42	0/2888
2	F	0.23	0/2105	0.43	0/2873
2	K	0.22	0/2112	0.40	0/2885
3	C	0.23	0/442	0.42	0/589
3	G	0.23	0/272	0.36	0/367
3	L	0.23	0/425	0.42	0/567
4	D	0.24	0/398	0.39	0/538
4	H	0.24	0/392	0.40	0/531
4	M	0.24	0/398	0.39	0/538
All	All	0.23	0/16371	0.42	0/22189

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
2	F	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	250	ASN	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	2541	2596	2589	6	0
1	E	2502	2550	2543	6	0
1	I	2553	2588	2581	6	0
2	B	2061	2020	2009	2	0
2	F	2053	2014	2004	5	0
2	K	2058	2007	1996	3	0
3	C	438	391	391	4	0
3	G	268	220	220	0	0
3	L	422	363	363	2	0
4	D	388	378	377	0	0
4	H	382	367	366	2	0
4	M	388	378	377	2	0
5	A	19	0	13	1	0
5	E	19	0	13	0	0
5	I	19	0	13	0	0
6	D	2	0	0	0	0
6	H	2	0	0	0	0
6	M	2	0	0	0	0
7	A	9	0	0	0	0
7	B	5	0	0	0	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
7	E	2	0	0	0	0
7	F	6	0	0	0	0
7	I	4	0	0	0	0
7	K	6	0	0	1	0
7	L	1	0	0	0	0
7	M	2	0	0	0	0
All	All	16154	15872	15855	30	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 1.

The worst 5 of 30 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)
4:M:22:CYS:SG	4:M:33:HIS:CE1	2.93	0.61
1:E:119:VAL:O	1:E:225:ARG:NH2	2.36	0.57
2:F:251:ARG:NH2	4:H:26:TYR:O	2.37	0.57
1:E:324:LEU:O	1:E:326:GLY:N	2.40	0.55
1:I:114:LEU:O	1:I:225:ARG:NH1	2.39	0.54

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	310/332 (93%)	292 (94%)	15 (5%)	3 (1%)	15 53
1	E	304/332 (92%)	288 (95%)	14 (5%)	2 (1%)	22 60
1	I	314/332 (95%)	299 (95%)	12 (4%)	3 (1%)	15 53
2	B	253/264 (96%)	244 (96%)	7 (3%)	2 (1%)	19 57
2	F	253/264 (96%)	247 (98%)	5 (2%)	1 (0%)	34 72
2	K	253/264 (96%)	247 (98%)	5 (2%)	1 (0%)	34 72
3	C	51/75 (68%)	49 (96%)	2 (4%)	0	100 100
3	G	34/75 (45%)	31 (91%)	2 (6%)	1 (3%)	4 24
3	L	50/75 (67%)	48 (96%)	2 (4%)	0	100 100
4	D	48/58 (83%)	47 (98%)	1 (2%)	0	100 100
4	H	48/58 (83%)	48 (100%)	0	0	100 100
4	M	48/58 (83%)	47 (98%)	1 (2%)	0	100 100
All	All	1966/2187 (90%)	1887 (96%)	66 (3%)	13 (1%)	22 60

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	251	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	325	LYS
1	A	190	VAL
2	B	260	ALA
1	E	190	VAL

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	277/292 (95%)	273 (99%)	4 (1%)	67	88
1	E	272/292 (93%)	262 (96%)	10 (4%)	34	70
1	I	275/292 (94%)	270 (98%)	5 (2%)	59	85
2	B	228/239 (95%)	220 (96%)	8 (4%)	36	71
2	F	227/239 (95%)	222 (98%)	5 (2%)	52	81
2	K	227/239 (95%)	220 (97%)	7 (3%)	40	75
3	C	43/69 (62%)	43 (100%)	0	100	100
3	G	25/69 (36%)	24 (96%)	1 (4%)	31	68
3	L	40/69 (58%)	40 (100%)	0	100	100
4	D	44/52 (85%)	44 (100%)	0	100	100
4	H	43/52 (83%)	43 (100%)	0	100	100
4	M	44/52 (85%)	43 (98%)	1 (2%)	50	80
All	All	1745/1956 (89%)	1704 (98%)	41 (2%)	50	80

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	62	THR
2	K	116	GLU
1	I	170	LEU
1	I	306	SER
2	K	148	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

3 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	I	186	1	8,10,11	1.13	0	10,14,16	1.68	1 (10%)
1	TPO	E	186	1	8,10,11	1.15	0	10,14,16	1.64	1 (10%)
1	TPO	A	186	1	8,10,11	1.09	0	10,14,16	1.64	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	I	186	1	-	0/9/11/13	-
1	TPO	E	186	1	-	0/9/11/13	-
1	TPO	A	186	1	-	0/9/11/13	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	I	186	TPO	P-OG1-CB	-4.64	109.18	123.21
1	A	186	TPO	P-OG1-CB	-4.53	109.51	123.21
1	E	186	TPO	P-OG1-CB	-4.51	109.58	123.21

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 6 are monoatomic - leaving 3 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	ADN	E	401	-	18,21,21	1.00	1 (5%)	18,31,31	1.52	3 (16%)
5	ADN	I	401	-	18,21,21	1.00	1 (5%)	18,31,31	1.57	3 (16%)
5	ADN	A	401	-	18,21,21	0.99	1 (5%)	18,31,31	1.55	3 (16%)

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	ADN	E	401	-	-	2/2/22/22	0/3/3/3
5	ADN	I	401	-	-	2/2/22/22	0/3/3/3
5	ADN	A	401	-	-	2/2/22/22	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	401	ADN	C6-N6	3.34	1.46	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	401	ADN	C6-N6	3.30	1.46	1.34
5	A	401	ADN	C6-N6	3.28	1.46	1.34

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	401	ADN	N3-C2-N1	-4.26	122.02	128.68
5	I	401	ADN	N3-C2-N1	-4.18	122.15	128.68
5	E	401	ADN	N3-C2-N1	-4.06	122.34	128.68
5	E	401	ADN	C4-C5-N7	-3.74	105.50	109.40
5	I	401	ADN	C4-C5-N7	-3.74	105.50	109.40

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	401	ADN	C3'-C4'-C5'-O5'
5	E	401	ADN	O4'-C4'-C5'-O5'
5	A	401	ADN	C3'-C4'-C5'-O5'
5	A	401	ADN	O4'-C4'-C5'-O5'
5	I	401	ADN	O4'-C4'-C5'-O5'

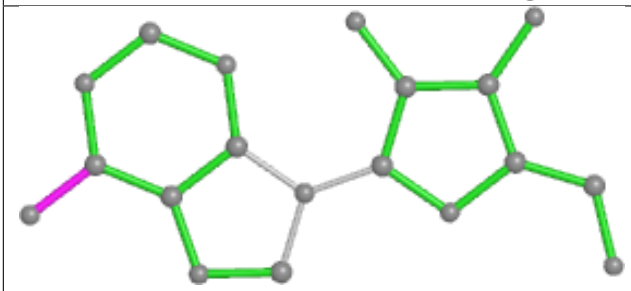
There are no ring outliers.

1 monomer is involved in 1 short contact:

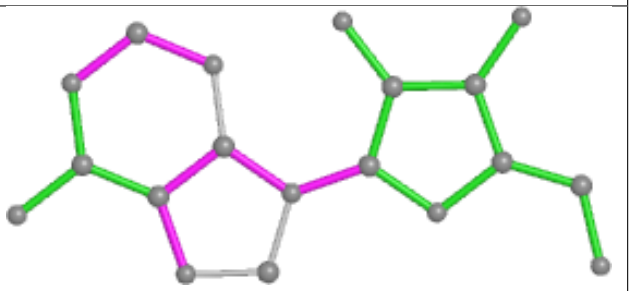
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	401	ADN	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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.

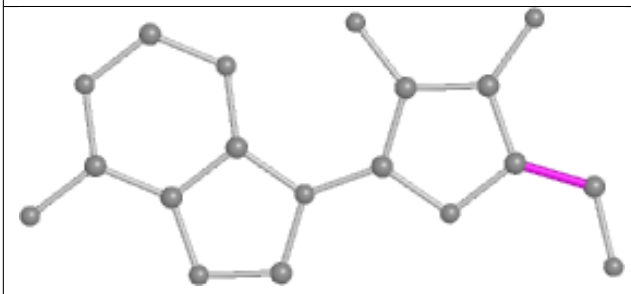
Ligand ADN E 401



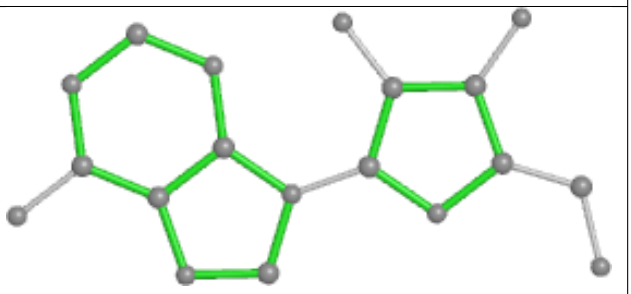
Bond lengths



Bond angles

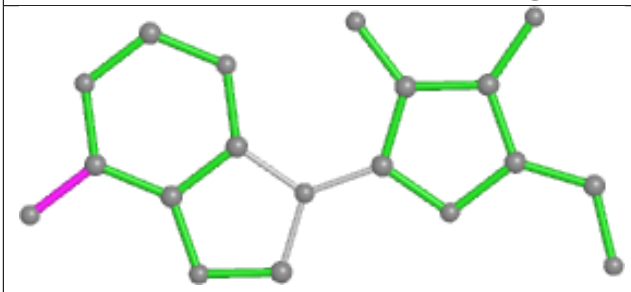


Torsions

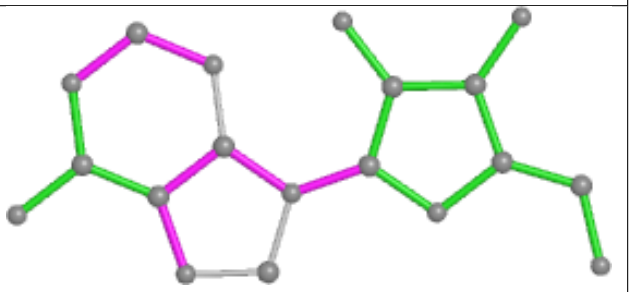


Rings

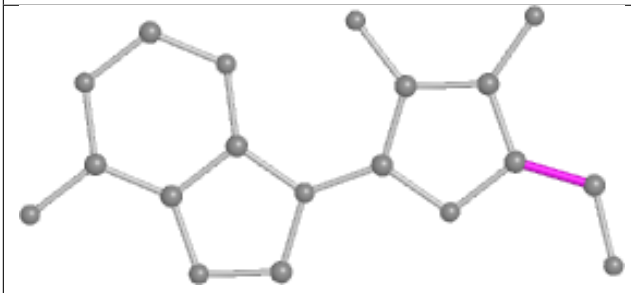
Ligand ADN I 401



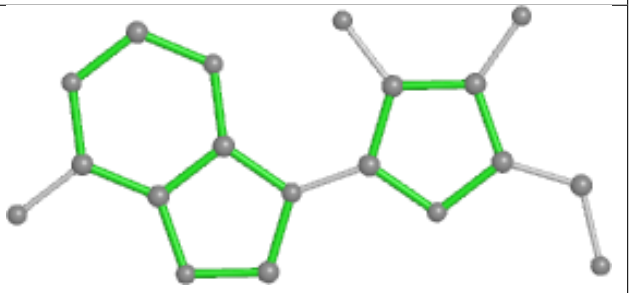
Bond lengths



Bond angles

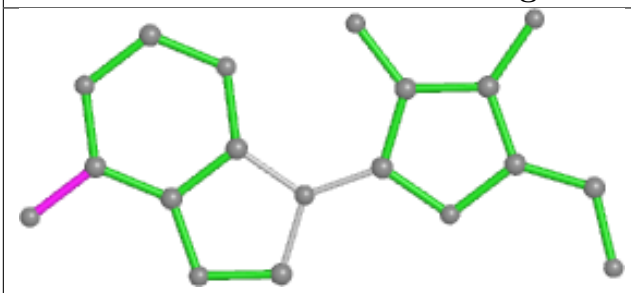


Torsions

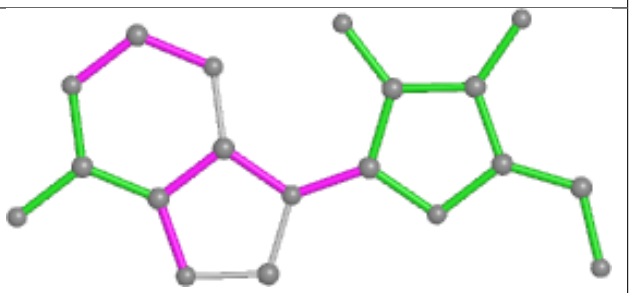


Rings

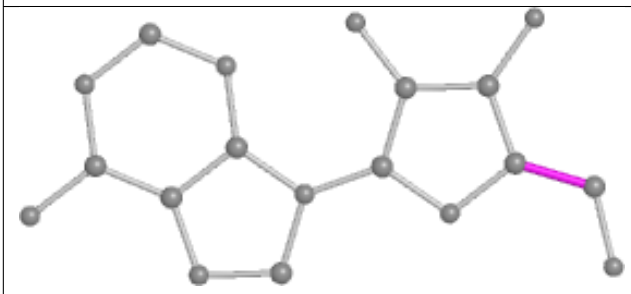
Ligand ADN A 401



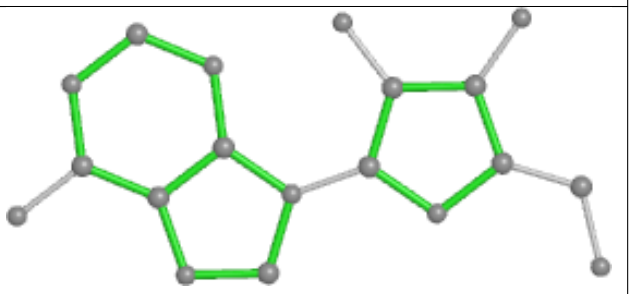
Bond lengths



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, 95^{th} 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	314/332 (94%)	0.18	5 (1%) 72 44	46, 66, 135, 210	0
1	E	310/332 (93%)	0.38	20 (6%) 18 5	50, 85, 142, 228	0
1	I	318/332 (95%)	0.26	5 (1%) 72 44	57, 78, 151, 215	0
2	B	255/264 (96%)	0.21	2 (0%) 86 65	48, 69, 134, 212	0
2	F	255/264 (96%)	0.36	11 (4%) 35 13	51, 75, 136, 185	0
2	K	255/264 (96%)	0.26	8 (3%) 49 21	61, 87, 165, 241	0
3	C	55/75 (73%)	0.04	0 100 100	55, 82, 123, 149	0
3	G	36/75 (48%)	0.79	3 (8%) 11 3	77, 106, 157, 179	0
3	L	54/75 (72%)	0.02	0 100 100	53, 88, 144, 155	0
4	D	49/58 (84%)	0.02	1 (2%) 65 36	53, 75, 104, 144	0
4	H	49/58 (84%)	0.27	0 100 100	58, 76, 102, 118	0
4	M	49/58 (84%)	0.17	1 (2%) 65 36	75, 98, 132, 154	0
All	All	1999/2187 (91%)	0.26	56 (2%) 53 25	46, 79, 143, 241	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	8	ASN	8.2
2	K	7	ASN	7.2
1	A	179	ASN	6.9
1	E	179	ASN	5.9
2	B	8	ASN	5.8

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, 95^{th} 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
1	TPO	A	186	11/12	0.94	0.22	47,62,74,75	0
1	TPO	E	186	11/12	0.95	0.20	56,71,85,86	0
1	TPO	I	186	11/12	0.96	0.19	72,82,99,99	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

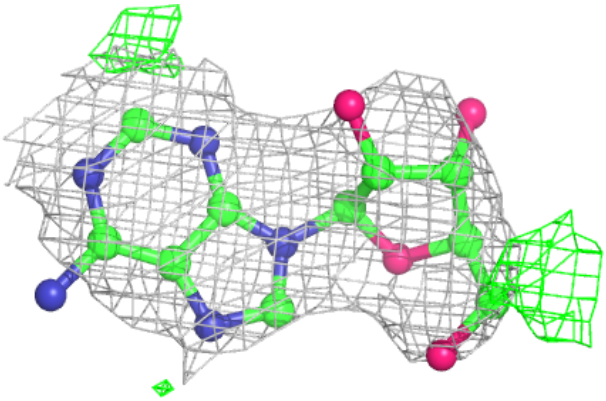
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

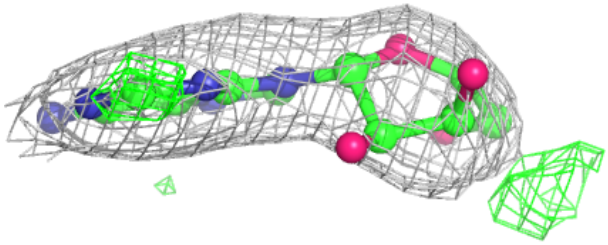
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	ADN	E	401	19/19	0.84	0.28	113,114,115,115	0
5	ADN	I	401	19/19	0.85	0.31	96,98,102,102	0
5	ADN	A	401	19/19	0.89	0.27	87,91,95,95	0
6	ZN	D	102	1/1	0.93	0.17	115,115,115,115	0
6	ZN	M	101	1/1	0.95	0.17	106,106,106,106	0
6	ZN	M	102	1/1	0.95	0.08	159,159,159,159	0
6	ZN	D	101	1/1	0.97	0.17	71,71,71,71	0
6	ZN	H	102	1/1	0.97	0.17	95,95,95,95	0
6	ZN	H	101	1/1	0.99	0.20	71,71,71,71	0

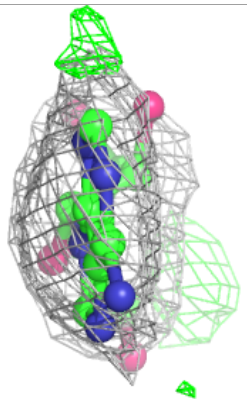
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around ADN E 401:

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

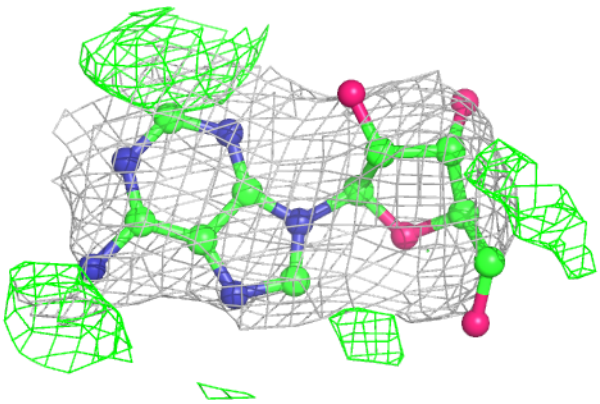


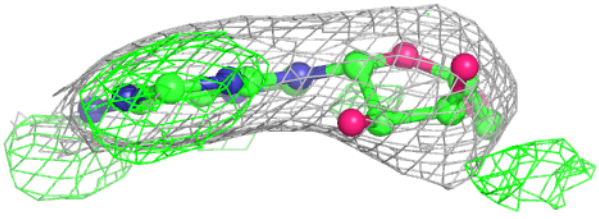


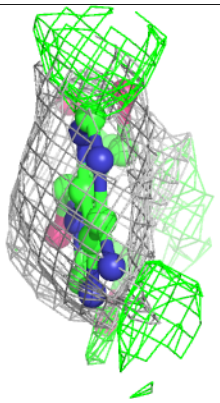


Electron density around ADN I 401:

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

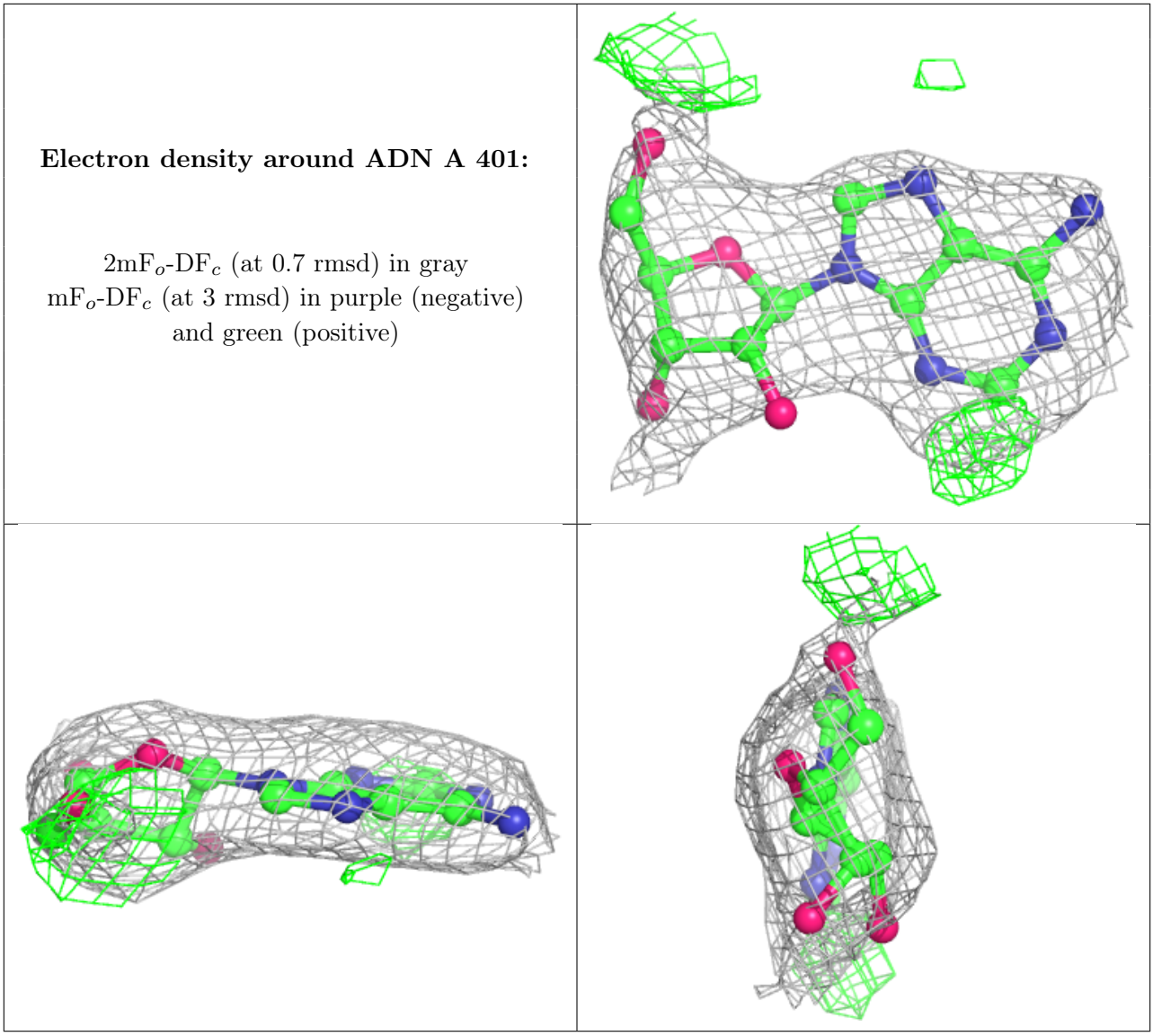






Electron density around ADN A 401:

$2mF_o$-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.