



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 06:51 am BST

PDB ID : 2J0S  
Title : The crystal structure of the Exon Junction Complex at 2.2 Å resolution  
Authors : Bono, F.; Ebert, J.; Lorentzen, E.; Conti, E.  
Deposited on : 2006-08-04  
Resolution : 2.21 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

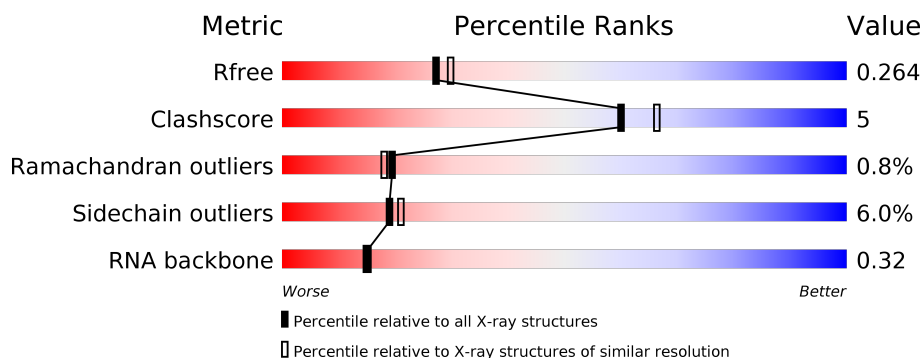
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.21 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	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RNA backbone	3102	1049 (2.64-1.80)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	410	82% 10% • 5%
2	C	146	85% 11% ...
3	D	89	87% 11% •
4	E	15	27% 13% 60%
5	T	150	23% 6% 71%

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 5906 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 ATP-DEPENDENT RNA HELICASE DDX48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	391	Total	C	N	O	S	0	8	0
			3138	1986	540	591	21			

- Molecule 2 is a protein called PROTEIN MAGO NASHI HOMOLOG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	143	Total	C	N	O	S	0	2	0
			1192	773	199	217	3			

- Molecule 3 is a protein called RNA-BINDING PROTEIN 8A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	89	Total	C	N	O	S	0	1	0
			710	453	116	138	3			

- Molecule 4 is a RNA chain called 5'-R(\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	6	Total	C	N	O	P	0	0	0
			121	54	12	49	6			

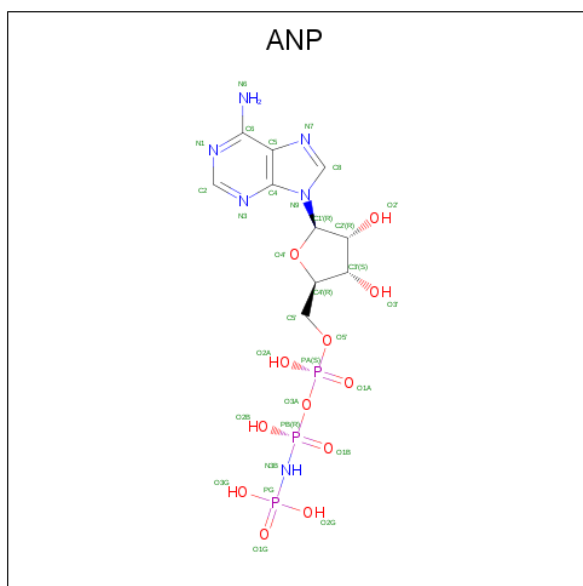
- Molecule 5 is a protein called PROTEIN CASC3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	T	44	Total	C	N	O	0	0	0
			369	228	69	72			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Mg	0	0
			1	1		

- Molecule 7 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
7	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	168	Total	O	0	0
			168	168		
8	C	117	Total	O	0	0
			117	117		
8	D	46	Total	O	0	0
			46	46		
8	E	4	Total	O	0	0
			4	4		
8	T	9	Total	O	0	0
			9	9		



ARG  
HIS  
GLN  
GLY

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	169.44 Å   169.44 Å   71.04 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	43.73 – 2.21 43.72 – 2.21	Depositor EDS
% Data completeness (in resolution range)	100.0 (43.73-2.21) 99.7 (43.72-2.21)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.39 (at 2.20 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.185 , 0.216 0.237 , 0.264	Depositor DCC
$R_{free}$ test set	2931 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.6	Xtriage
Anisotropy	0.173	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 51.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.030 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5906	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	A	0.62	0/3211	0.75	2/4340 (0.0%)
2	C	0.80	1/1228 (0.1%)	0.84	3/1655 (0.2%)
3	D	0.65	0/730	0.67	0/989
4	E	1.33	1/132 (0.8%)	1.77	3/200 (1.5%)
5	T	0.66	0/379	0.72	0/510
All	All	0.69	2/5680 (0.0%)	0.80	8/7694 (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
1	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	1	U	OP3-P	-9.43	1.49	1.61
2	C	131	CYS	CB-SG	-6.91	1.70	1.82

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	27	ARG	NE-CZ-NH1	7.26	123.93	120.30
1	A	48	ASP	CB-CG-OD1	6.39	124.05	118.30
4	E	1	U	O4'-C1'-N1	6.33	113.26	108.20
2	C	27	ARG	NE-CZ-NH2	-5.86	117.37	120.30
2	C	82	ARG	NE-CZ-NH2	5.72	123.16	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	177	ARG	Peptide

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3138	0	3166	35	0
2	C	1192	0	1173	8	0
3	D	710	0	662	6	0
4	E	121	0	61	0	0
5	T	369	0	308	4	0
6	A	1	0	0	0	0
7	A	31	0	13	0	0
8	A	168	0	0	0	0
8	C	117	0	0	2	0
8	D	46	0	0	0	0
8	E	4	0	0	0	0
8	T	9	0	0	0	0
All	All	5906	0	5383	50	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 50 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:A:121:GLN:HE22	1:A:342:ASP:H	1.29	0.80
1:A:121:GLN:HE21	1:A:324:ARG:HH22	1.35	0.72
1:A:275:LEU:HD23	1:A:348[B]:LEU:CD2	2.18	0.72
2:C:102[A]:SER:OG	8:C:2087:HOH:O	2.08	0.70
1:A:275:LEU:HD23	1:A:348[B]:LEU:HD21	1.74	0.68

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	397/410 (97%)	388 (98%)	6 (2%)	3 (1%)	19	18
2	C	143/146 (98%)	142 (99%)	1 (1%)	0	100	100
3	D	88/89 (99%)	86 (98%)	1 (1%)	1 (1%)	14	11
5	T	40/150 (27%)	36 (90%)	3 (8%)	1 (2%)	5	2
All	All	668/795 (84%)	652 (98%)	11 (2%)	5 (1%)	19	21

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	408	ALA
1	A	409	ASP
3	D	139	LEU
1	A	23	MET
5	T	216	GLY

### 5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	345/360 (96%)	324 (94%)	21 (6%)	18	20
2	C	130/134 (97%)	122 (94%)	8 (6%)	18	19
3	D	72/75 (96%)	67 (93%)	5 (7%)	15	15
5	T	35/133 (26%)	33 (94%)	2 (6%)	20	23
All	All	582/702 (83%)	546 (94%)	36 (6%)	19	19

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	337	TRP
1	A	407	VAL
3	D	138	ASP
1	A	348[B]	LEU
2	C	7	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	356	ASN
2	C	13	HIS
3	D	103	HIS
1	A	351	ASN
3	D	90	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	E	5/15 (33%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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
7	ANP	A	1413	6	29,33,33	1.97	9 (31%)	31,52,52	1.71	7 (22%)

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
7	ANP	A	1413	6	-	2/14/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1413	ANP	PB-O1B	5.42	1.54	1.46
7	A	1413	ANP	PB-N3B	3.62	1.72	1.63
7	A	1413	ANP	PG-O1G	3.44	1.51	1.46
7	A	1413	ANP	PG-O2G	-3.43	1.47	1.56
7	A	1413	ANP	O4'-C1'	3.19	1.45	1.41

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1413	ANP	O2B-PB-O1B	4.67	119.70	109.92
7	A	1413	ANP	N3-C2-N1	-4.07	122.31	128.68
7	A	1413	ANP	O3G-PG-O2G	3.37	116.62	107.64
7	A	1413	ANP	O1G-PG-N3B	-3.09	107.21	111.77
7	A	1413	ANP	C1'-N9-C4	-2.29	122.61	126.64

There are no chirality outliers.

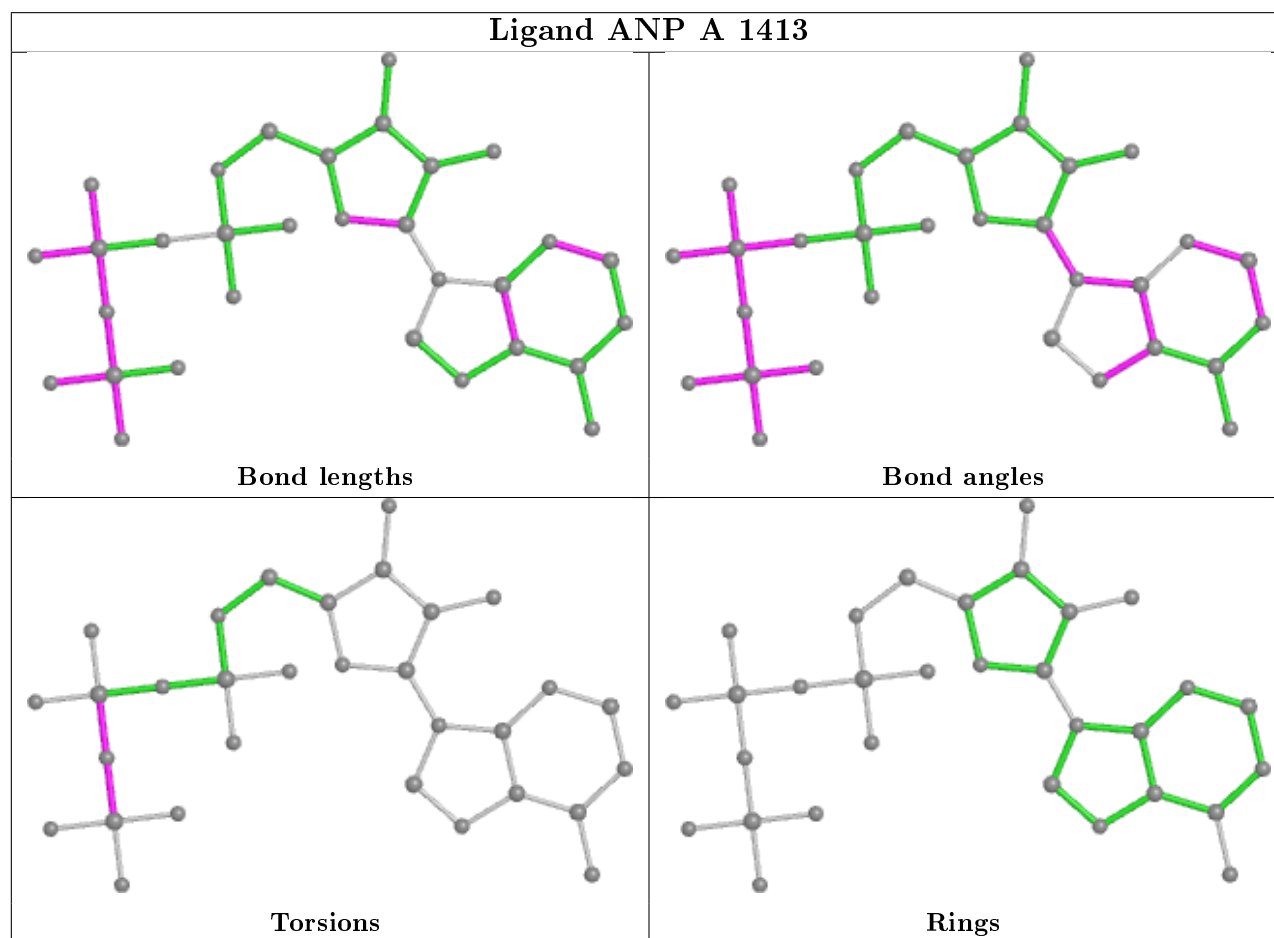
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	1413	ANP	PG-N3B-PB-O1B
7	A	1413	ANP	PB-N3B-PG-O1G

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

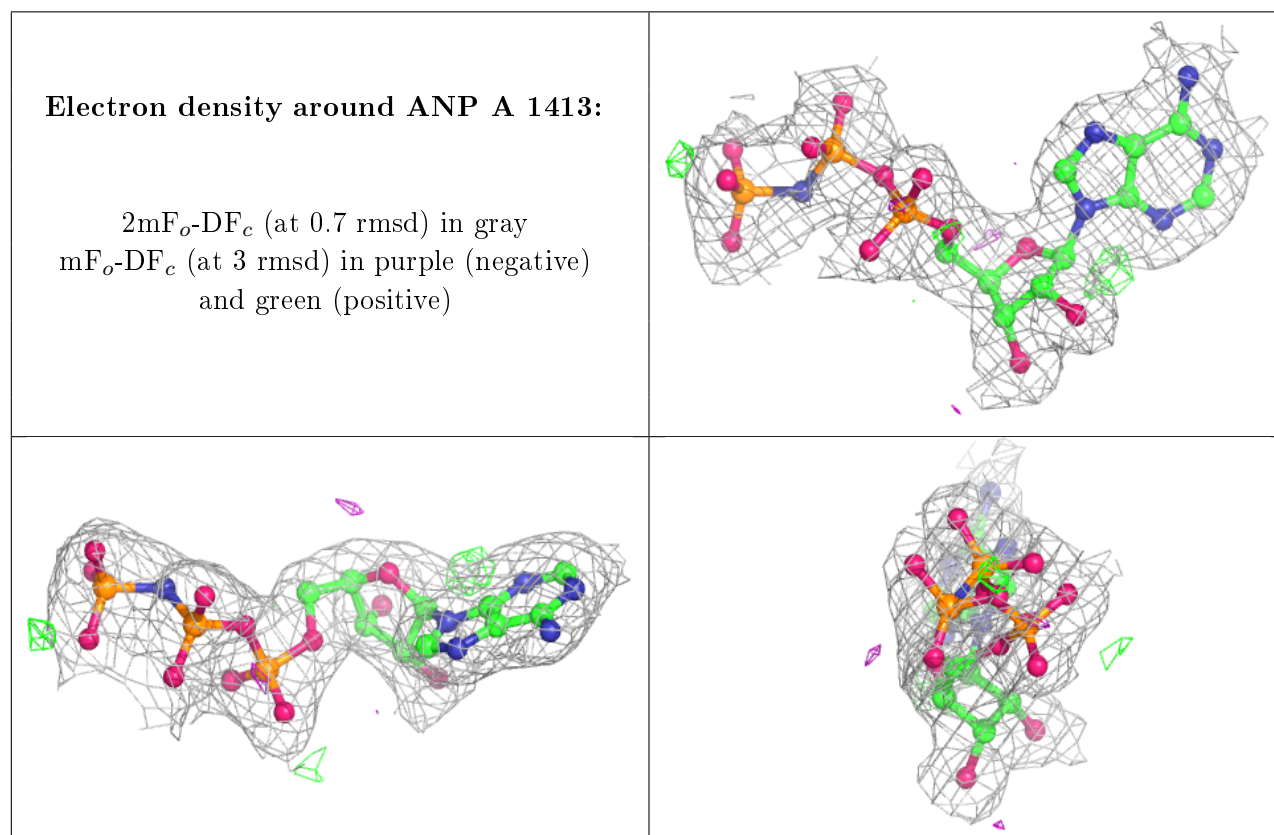
### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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.



## 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.