W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 27, 2024 – 06:24 PM EDT

PDB ID	:	9B9W
Title	:	Crystal structure of the ternary complex of DCAF1 and WDR5 with PROTAC, OICR-40792
Authors	:	Mabanglo, M.F.; Wilson, B.J.; Mamai, A.; Hoffer, L.; Al-awar, R.; Vedadi, M.
Deposited on	:	2024-04-03
Resolution	:	1.92 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:

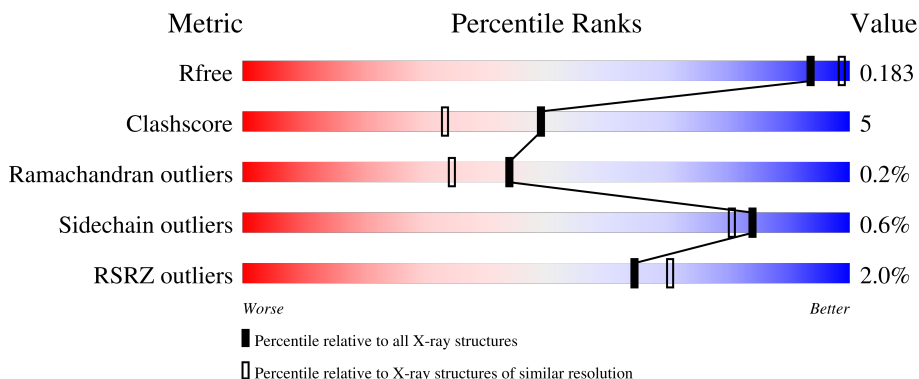|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 1.92 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.183
Clashscore		5
Ramachandran outliers		0.2%
Sidechain outliers		0.6%
RSRZ outliers		2.0%

Worse   *Better*

▎ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1028 (1.92-1.92)
Clashscore	180529	1100 (1.92-1.92)
Ramachandran outliers	177936	1087 (1.92-1.92)
Sidechain outliers	177891	1087 (1.92-1.92)
RSRZ outliers	164620	1028 (1.92-1.92)

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	329	82% 11% 8%
2	B	338	3% 78% 11% 11%

2   Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5354 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 WD repeat-containing protein 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	304	Total 2378	C 1518	N 395	O 454	S 11	0	4	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	6	MET	-	initiating methionine	UNP P61964
A	7	HIS	-	expression tag	UNP P61964
A	8	HIS	-	expression tag	UNP P61964
A	9	HIS	-	expression tag	UNP P61964
A	10	HIS	-	expression tag	UNP P61964
A	11	HIS	-	expression tag	UNP P61964
A	12	HIS	-	expression tag	UNP P61964
A	13	SER	-	expression tag	UNP P61964
A	14	SER	-	expression tag	UNP P61964
A	15	GLY	-	expression tag	UNP P61964
A	16	ARG	-	expression tag	UNP P61964
A	17	GLU	-	expression tag	UNP P61964
A	18	ASN	-	expression tag	UNP P61964
A	19	LEU	-	expression tag	UNP P61964
A	20	TYR	-	expression tag	UNP P61964
A	21	PHE	-	expression tag	UNP P61964
A	22	GLN	-	expression tag	UNP P61964
A	23	GLY	-	expression tag	UNP P61964

- Molecule 2 is a protein called DDB1- and CUL4-associated factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	301	Total 2408	C 1523	N 410	O 458	S 17	0	3	0

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1053	MET	-	initiating methionine	UNP Q9Y4B6
B	1054	GLY	-	expression tag	UNP Q9Y4B6
B	1055	SER	-	expression tag	UNP Q9Y4B6
B	1056	SER	-	expression tag	UNP Q9Y4B6
B	1057	HIS	-	expression tag	UNP Q9Y4B6
B	1058	HIS	-	expression tag	UNP Q9Y4B6
B	1059	HIS	-	expression tag	UNP Q9Y4B6
B	1060	HIS	-	expression tag	UNP Q9Y4B6
B	1061	HIS	-	expression tag	UNP Q9Y4B6
B	1062	HIS	-	expression tag	UNP Q9Y4B6
B	1063	SER	-	expression tag	UNP Q9Y4B6
B	1064	SER	-	expression tag	UNP Q9Y4B6
B	1065	GLY	-	expression tag	UNP Q9Y4B6
B	1066	ARG	-	expression tag	UNP Q9Y4B6
B	1067	GLU	-	expression tag	UNP Q9Y4B6
B	1068	ASN	-	expression tag	UNP Q9Y4B6
B	1069	LEU	-	expression tag	UNP Q9Y4B6
B	1070	TYR	-	expression tag	UNP Q9Y4B6
B	1071	PHE	-	expression tag	UNP Q9Y4B6
B	1072	GLN	-	expression tag	UNP Q9Y4B6
B	1073	GLY	-	expression tag	UNP Q9Y4B6
B	1074	SER	-	expression tag	UNP Q9Y4B6
B	1075	SER	-	expression tag	UNP Q9Y4B6
B	1076	ARG	-	expression tag	UNP Q9Y4B6
B	1077	ALA	-	expression tag	UNP Q9Y4B6
B	1078	SER	-	expression tag	UNP Q9Y4B6
B	1079	ALA	-	expression tag	UNP Q9Y4B6

- Molecule 3 is (4P)-N-[(1R)-3-amino-1-(3-chloro-4-fluorophenyl)-3-oxopropyl]-4-(4-chloro-2-fluorophenyl)-5-[(25E)-1-{(1P)-6-fluoro-3'-[4-fluoro-2-(trifluoromethyl)benzamido]-4'-[(3R,5S)-3,4,5-trimethylpiperazin-1-yl][1,1'-biphenyl]-3-yl}-1,24-dioxo-5,8,11,14,17,20-hexaoxa-2,23-diazahexacos-25-en-26-yl]-1H-pyrrole-3-carboxamide (three-letter code: A1ANN) (formula: $C_{65}H_{71}Cl_2F_7N_8O_{11}$) (labeled as "Ligand of Interest" by depositor).

A1ANN
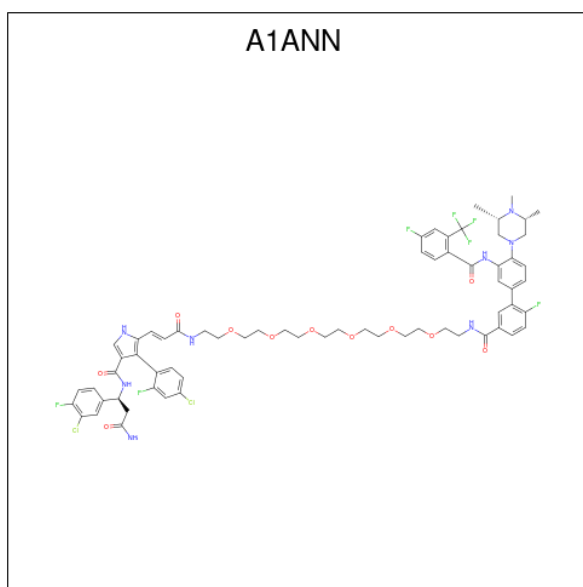
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Cl	F	N	O		
3	A	1	93	65	2	7	8	11	0	0

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	258	258	258	0	0
4	B	217	217	217	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.
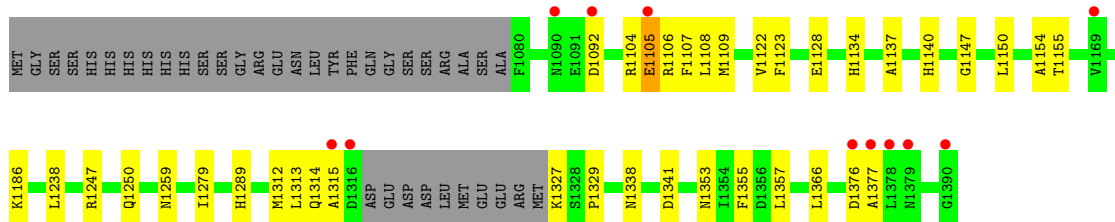
• Molecule 1: WD repeat-containing protein 5
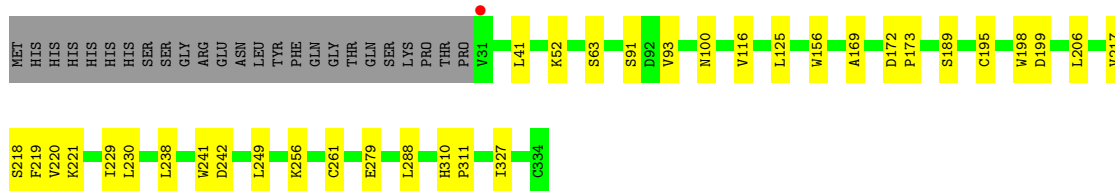
• Molecule 2: DDB1- and CUL4-associated factor 1

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	79.89Å 83.53Å 130.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.29 – 1.92 43.29 – 1.92	Depositor EDS
% Data completeness (in resolution range)	100.0 (43.29-1.92) 100.0 (43.29-1.92)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.86 (at 1.92Å)	Xtriage
Refinement program	PHENIX 1.21_5207	Depositor
R, R_{free}	0.172 , 0.183 0.172 , 0.183	Depositor DCC
R_{free} test set	3397 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	28.0	Xtriage
Anisotropy	0.130	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 36.4	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5354	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% 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: A1ANN

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	0.40	0/2440	0.58	0/3309
2	B	0.37	0/2469	0.60	0/3345
All	All	0.39	0/4909	0.59	0/6654

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2378	0	2367	22	0
2	B	2408	0	2322	29	0
3	A	93	0	0	0	0
4	A	258	0	0	2	0
4	B	217	0	0	7	0
All	All	5354	0	4689	51	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 51 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:100:ASN:O	1:A:116:VAL:HG12	1.71	0.90
2:B:1107:PHE:CE1	2:B:1128:GLU:HG3	2.24	0.73
2:B:1105:GLU:HA	4:B:1524:HOH:O	1.90	0.71
2:B:1155[B]:THR:HG21	4:B:1530:HOH:O	1.91	0.71
2:B:1357:LEU:HD13	2:B:1366:LEU:HD11	1.72	0.70

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	306/329 (93%)	294 (96%)	12 (4%)	0	100	100
2	B	300/338 (89%)	285 (95%)	14 (5%)	1 (0%)	37	26
All	All	606/667 (91%)	579 (96%)	26 (4%)	1 (0%)	44	34

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	1315	ALA

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	270/289 (93%)	270 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	268/297 (90%)	265 (99%)	3 (1%)	70	63
All	All	538/586 (92%)	535 (99%)	3 (1%)	84	80

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

Mol	Chain	Res	Type
2	B	1092	ASP
2	B	1105	GLU
2	B	1259	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1126	GLN
2	B	1140	HIS
2	B	1250	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the

expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
3	A1ANN	A	1401	-	98,99,99	2.35	19 (19%)	126,136,136	1.81	26 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	A1ANN	A	1401	-	-	11/72/92/92	0/7/7/7

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1401	A1ANN	C56-N55	-11.61	1.29	1.47
3	A	1401	A1ANN	C86-C85	-9.80	1.38	1.50
3	A	1401	A1ANN	C24-N23	6.24	1.47	1.33
3	A	1401	A1ANN	C31-C32	-6.01	1.38	1.49
3	A	1401	A1ANN	C65-C64	5.87	1.53	1.46

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1401	A1ANN	C65-C64-C63	11.17	141.16	124.18
3	A	1401	A1ANN	C72-N71-C69	-4.98	116.42	122.34
3	A	1401	A1ANN	C26-C24-N23	-4.05	108.71	117.12
3	A	1401	A1ANN	C57-C59-N51	3.63	116.23	110.86
3	A	1401	A1ANN	C67-N66-C65	3.48	109.95	104.56

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1401	A1ANN	N23-C24-C26-C27
3	A	1401	A1ANN	O25-C24-C26-C27
3	A	1401	A1ANN	N23-C24-C26-C61
3	A	1401	A1ANN	O17-C18-C19-O20
3	A	1401	A1ANN	O25-C24-C26-C61

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 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 A1ANN A 1401
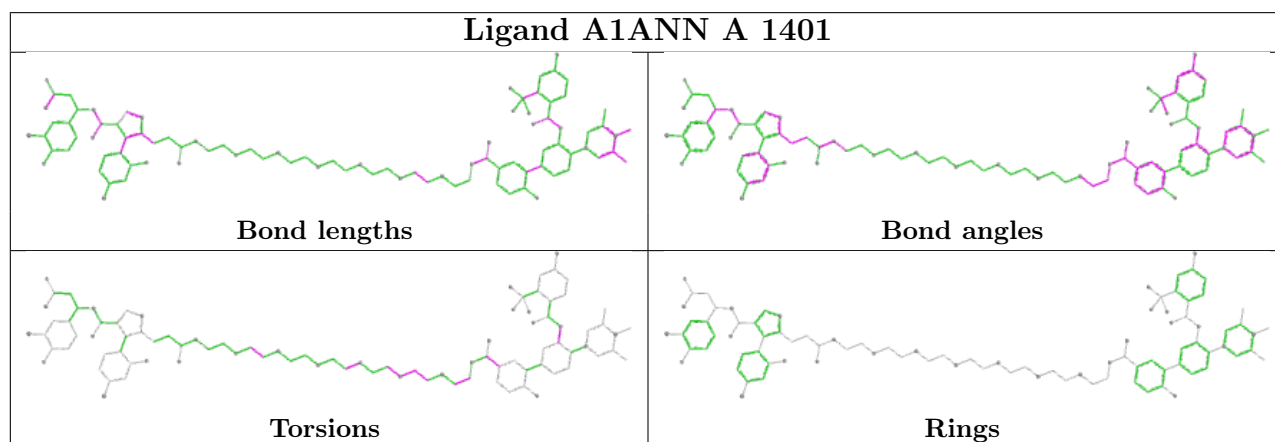
5.7 Other polymers ⓘ

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	304/329 (92%)	-0.47	1 (0%)	90	93	11, 26, 40, 50	4 (1%)
2	B	301/338 (89%)	-0.15	11 (3%)	45	52	12, 30, 51, 83	3 (0%)
All	All	605/667 (90%)	-0.31	12 (1%)	64	70	11, 28, 46, 83	7 (1%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1378	LEU	3.8
1	A	31	VAL	3.5
2	B	1390	GLY	3.3
2	B	1376	ASP	2.8
2	B	1379	ASN	2.6

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

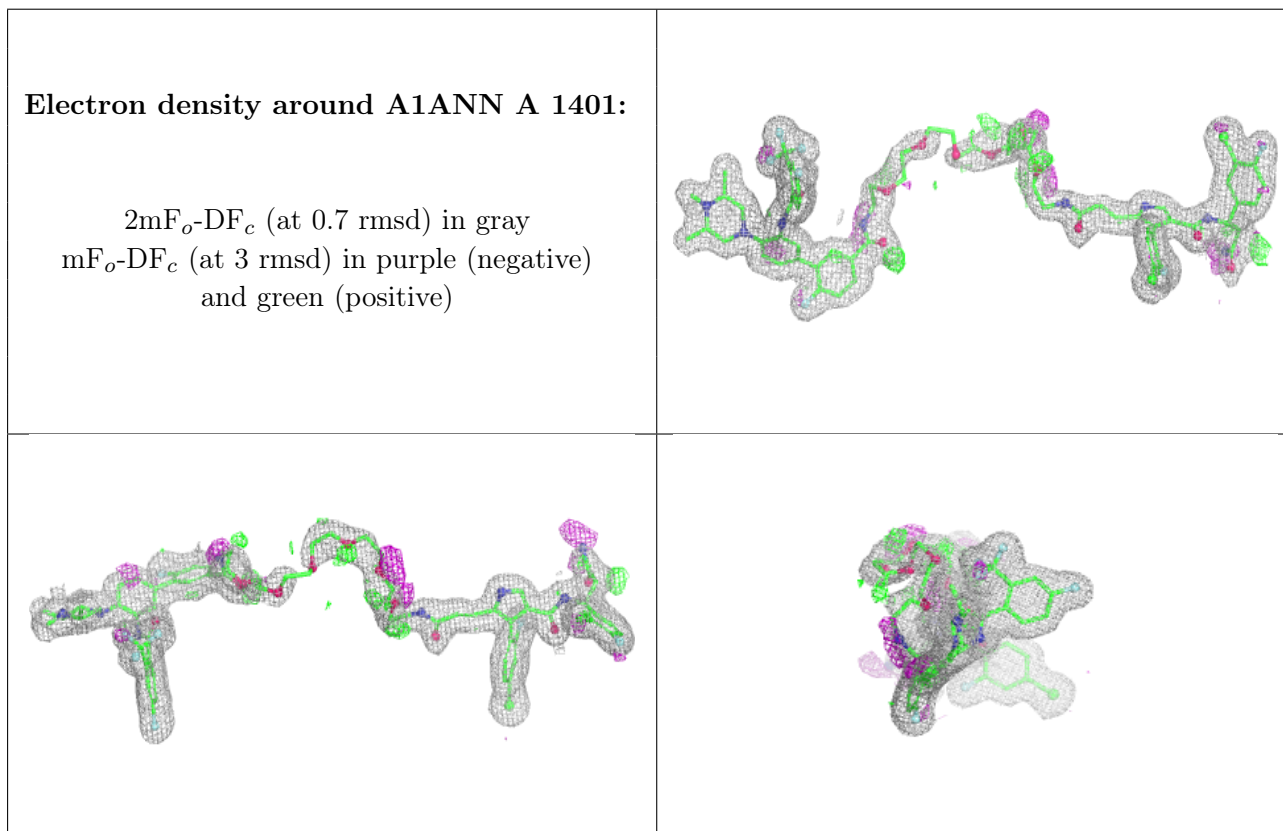There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	A1ANN	A	1401	93/93	0.95	0.10	16,26,64,72	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 A1ANN A 1401:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.