wwPDB X-ray Structure Validation Summary Report (i)

Jan 13, 2024 – 10:32 pm GMT

PDB ID	:	6TIJ
Title	:	Structure of A. niger Fdc WT in complex with indol-2-carboxylic acid
Authors	:	Leys, D.
Deposited on	:	2019-11-22
Resolution	:	1.12 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:

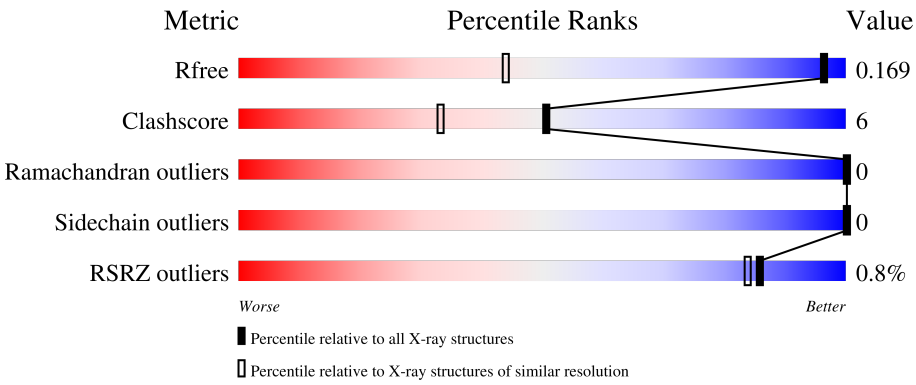
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

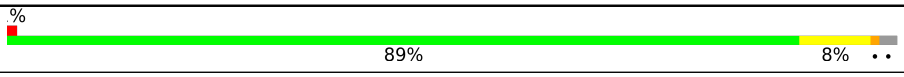
The reported resolution of this entry is 1.12 Å.

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	1168 (1.14-1.10)
Clashscore	141614	1205 (1.14-1.10)
Ramachandran outliers	138981	1168 (1.14-1.10)
Sidechain outliers	138945	1165 (1.14-1.10)
RSRZ outliers	127900	1146 (1.14-1.10)

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	AAA	508	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	ICB	AAA	605[A]	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4843 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 Ferulic acid decarboxylase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	500	4123	2608	709	776	30	0	34	1

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	499B	LEU	-	expression tag	UNP A2QHE5
AAA	499C	GLU	-	expression tag	UNP A2QHE5
AAA	502	HIS	-	expression tag	UNP A2QHE5
AAA	503	HIS	-	expression tag	UNP A2QHE5
AAA	504	HIS	-	expression tag	UNP A2QHE5
AAA	505	HIS	-	expression tag	UNP A2QHE5
AAA	506	HIS	-	expression tag	UNP A2QHE5
AAA	507	HIS	-	expression tag	UNP A2QHE5

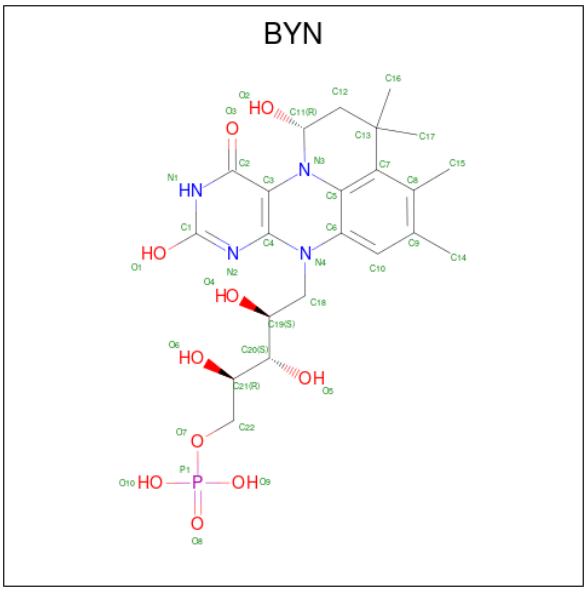
- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mn		
2	AAA	1	1	1	0	0

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

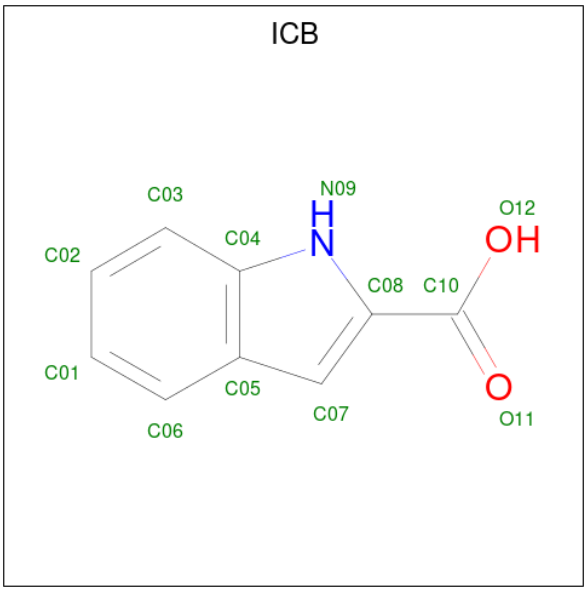
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	K		
3	AAA	2	2	2	0	0

- Molecule 4 is hydroxylated prenyl-FMN (three-letter code: BYN) (formula: $C_{22}H_{31}N_4O_{10}P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	AAA	1	37	22	4	10	1	0	0

- Molecule 5 is 1H-indole-2-carboxylic acid (three-letter code: ICB) (formula: $C_9H_7NO_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	AAA	1	12	9	1	2	0	1
5	AAA	1	12	9	1	2	0	0

- Molecule 6 is water.

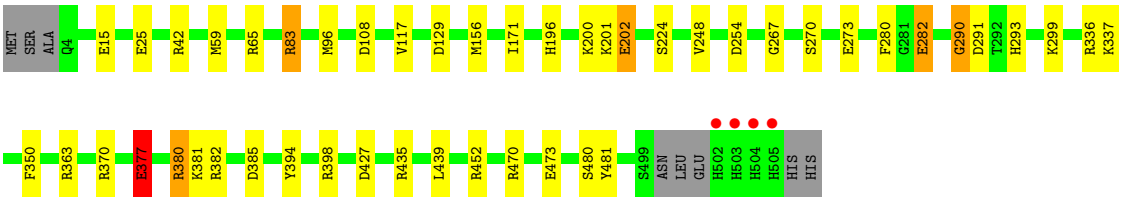
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	AAA	656	656	656	0	4

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: Ferulic acid decarboxylase 1

Chain AAA:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	95.78Å 63.73Å 87.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	64.69 – 1.12 64.68 – 1.12	Depositor EDS
% Data completeness (in resolution range)	98.5 (64.69-1.12) 98.5 (64.68-1.12)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.78 (at 1.12Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.140 , 0.169 0.140 , 0.169	Depositor DCC
R_{free} test set	10013 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	9.3	Xtriage
Anisotropy	0.274	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 46.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	4843	wwPDB-VP
Average B, all atoms (Å^2)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.82% 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: K, BYN, ICB, MN

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	AAA	1.04	14/4225 (0.3%)	1.15	25/5740 (0.4%)

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	AAA	0	3

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	282[A]	GLU	CD-OE2	12.59	1.39	1.25
1	AAA	282[B]	GLU	CD-OE2	12.59	1.39	1.25
1	AAA	282[C]	GLU	CD-OE2	12.59	1.39	1.25
1	AAA	377[A]	GLU	CD-OE1	9.09	1.35	1.25
1	AAA	377[B]	GLU	CD-OE1	9.09	1.35	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	65	ARG	NE-CZ-NH1	-20.54	110.03	120.30
1	AAA	65	ARG	NE-CZ-NH2	10.87	125.73	120.30
1	AAA	156	MET	CA-CB-CG	9.22	128.97	113.30
1	AAA	452	ARG	NE-CZ-NH2	-8.47	116.06	120.30
1	AAA	435	ARG	NE-CZ-NH2	-8.36	116.12	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	377[A]	GLU	Sidechain
1	AAA	380	ARG	Sidechain
1	AAA	83	ARG	Sidechain

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	AAA	4123	0	3997	42	0
2	AAA	1	0	0	0	0
3	AAA	2	0	0	0	0
4	AAA	37	0	0	5	0
5	AAA	24	0	0	11	0
6	AAA	656	0	0	20	3
All	All	4843	0	3997	47	3

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 6.

The worst 5 of 47 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:AAA:439[A]:LEU:HD21	5:AAA:605[A]:ICB:C07	1.67	1.22
1:AAA:282[A]:GLU:OE2	6:AAA:701:HOH:O	1.67	1.11
1:AAA:439[A]:LEU:CD2	5:AAA:605[A]:ICB:C07	2.29	1.11
4:AAA:604:BYN:O2	5:AAA:605[A]:ICB:C08	1.99	1.11
1:AAA:381:LYS:HG3	6:AAA:716:HOH:O	1.59	1.03

All (3) 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)
6:AAA:784:HOH:O	6:AAA:935:HOH:O[2_655]	1.61	0.59
6:AAA:735:HOH:O	6:AAA:1175:HOH:O[3_555]	2.18	0.02
6:AAA:819:HOH:O	6:AAA:1187:HOH:O[2_555]	2.19	0.01

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	AAA	531/508 (104%)	520 (98%)	11 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	AAA	445/430 (104%)	445 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

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)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 3 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	$\#\|Z\| > 2$
4	BYN	AAA	604	2,3	37,40,40	2.51	5 (13%)	49,64,64	1.44	6 (12%)
5	ICB	AAA	605[A]	-	11,13,13	1.87	4 (36%)	14,18,18	2.03	5 (35%)
5	ICB	AAA	606	-	11,13,13	2.10	3 (27%)	14,18,18	1.70	3 (21%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	BYN	AAA	604	2,3	-	2/18/33/33	0/4/4/4
5	ICB	AAA	605[A]	-	-	0/2/4/4	0/2/2/2
5	ICB	AAA	606	-	-	0/2/4/4	0/2/2/2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	AAA	604	BYN	C12-C11	-12.05	1.41	1.52
4	AAA	604	BYN	C5-C6	-5.86	1.33	1.40
5	AAA	606	ICB	C07-C08	5.02	1.48	1.39
4	AAA	604	BYN	O2-C11	3.65	1.47	1.41
4	AAA	604	BYN	C3-N3	3.01	1.42	1.37

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	AAA	604	BYN	C13-C7-C5	-4.78	116.26	121.48
5	AAA	605[A]	ICB	O12-C10-C08	3.77	123.27	114.69
4	AAA	604	BYN	C5-C6-N4	3.67	123.58	118.92
4	AAA	604	BYN	C6-C5-C7	3.22	125.83	120.75
5	AAA	606	ICB	O12-C10-C08	3.16	121.88	114.69

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	AAA	604	BYN	C19-C20-C21-C22
4	AAA	604	BYN	C21-C22-O7-P1

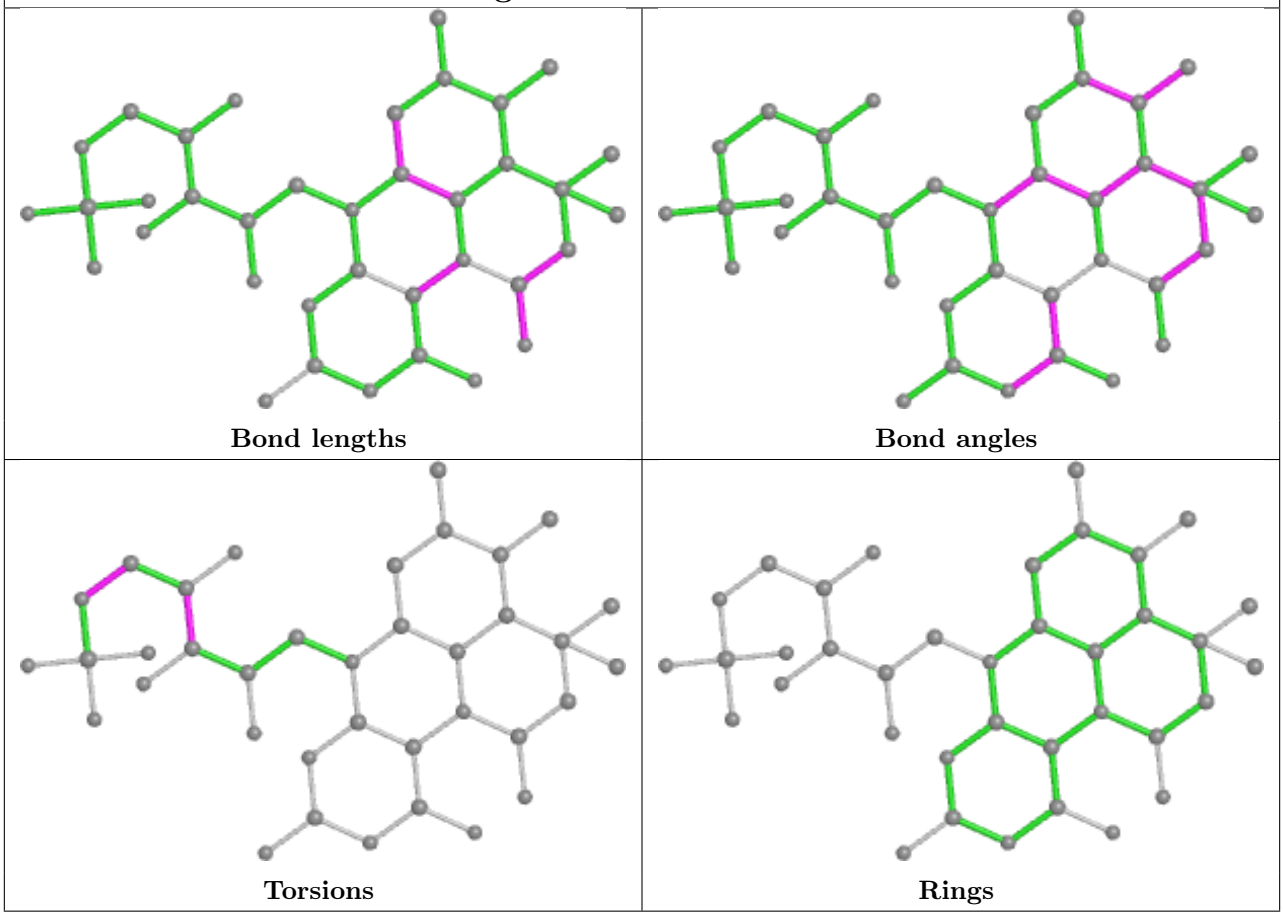
There are no ring outliers.

2 monomers are involved in 12 short contacts:

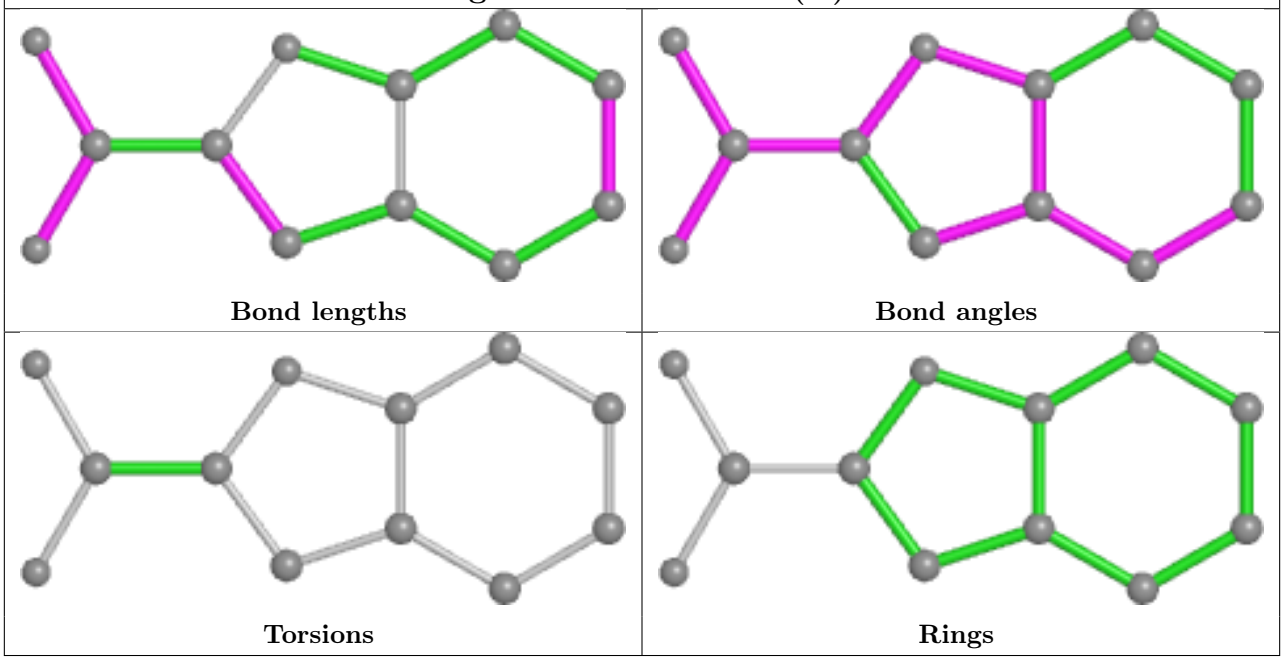
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	AAA	604	BYN	5	0
5	AAA	605[A]	ICB	11	0

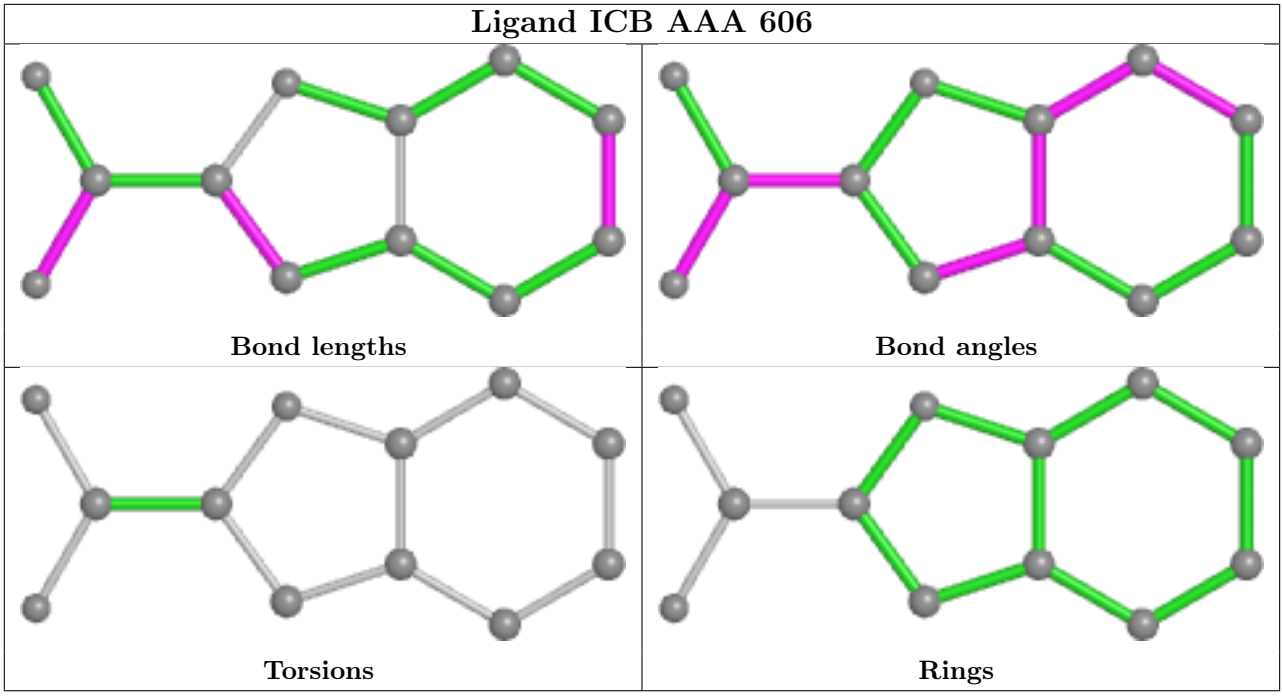
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 BYN AAA 604



Ligand ICB AAA 605 (A)





5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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	AAA	500/508 (98%)	-0.42	4 (0%) 86 84	6, 10, 23, 73	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	503	HIS	5.5
1	AAA	504	HIS	4.5
1	AAA	502	HIS	4.2
1	AAA	505	HIS	4.1

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

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
5	ICB	AAA	606	12/12	0.92	0.19	22,26,40,49	0
5	ICB	AAA	605[A]	12/12	0.97	0.11	11,13,23,37	12
4	BYN	AAA	604	37/37	0.99	0.06	6,7,10,11	1

Continued on next page...

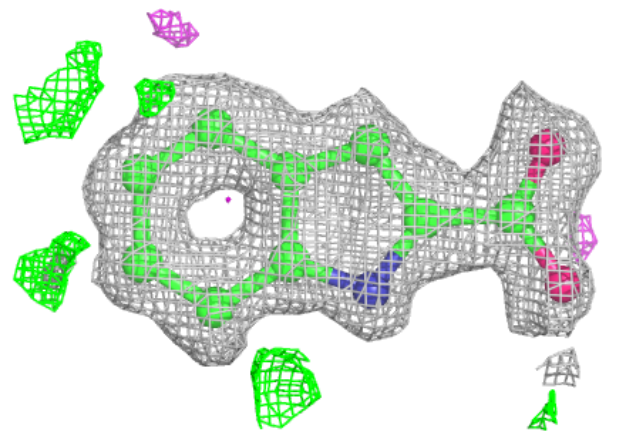
Continued from previous page...

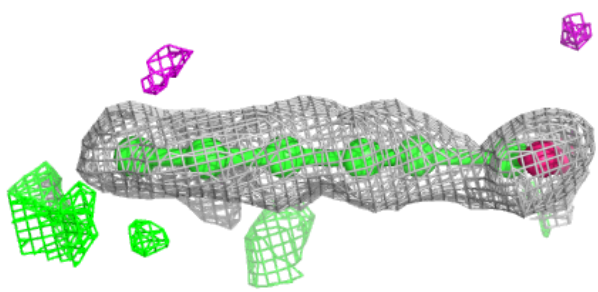
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MN	AAA	601	1/1	1.00	0.04	7,7,7,7	0
3	K	AAA	602	1/1	1.00	0.04	8,8,8,8	0
3	K	AAA	603	1/1	1.00	0.04	8,8,8,8	0

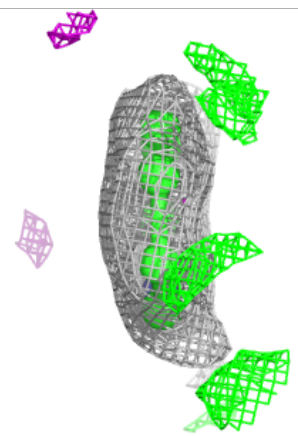
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ICB AAA 606:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

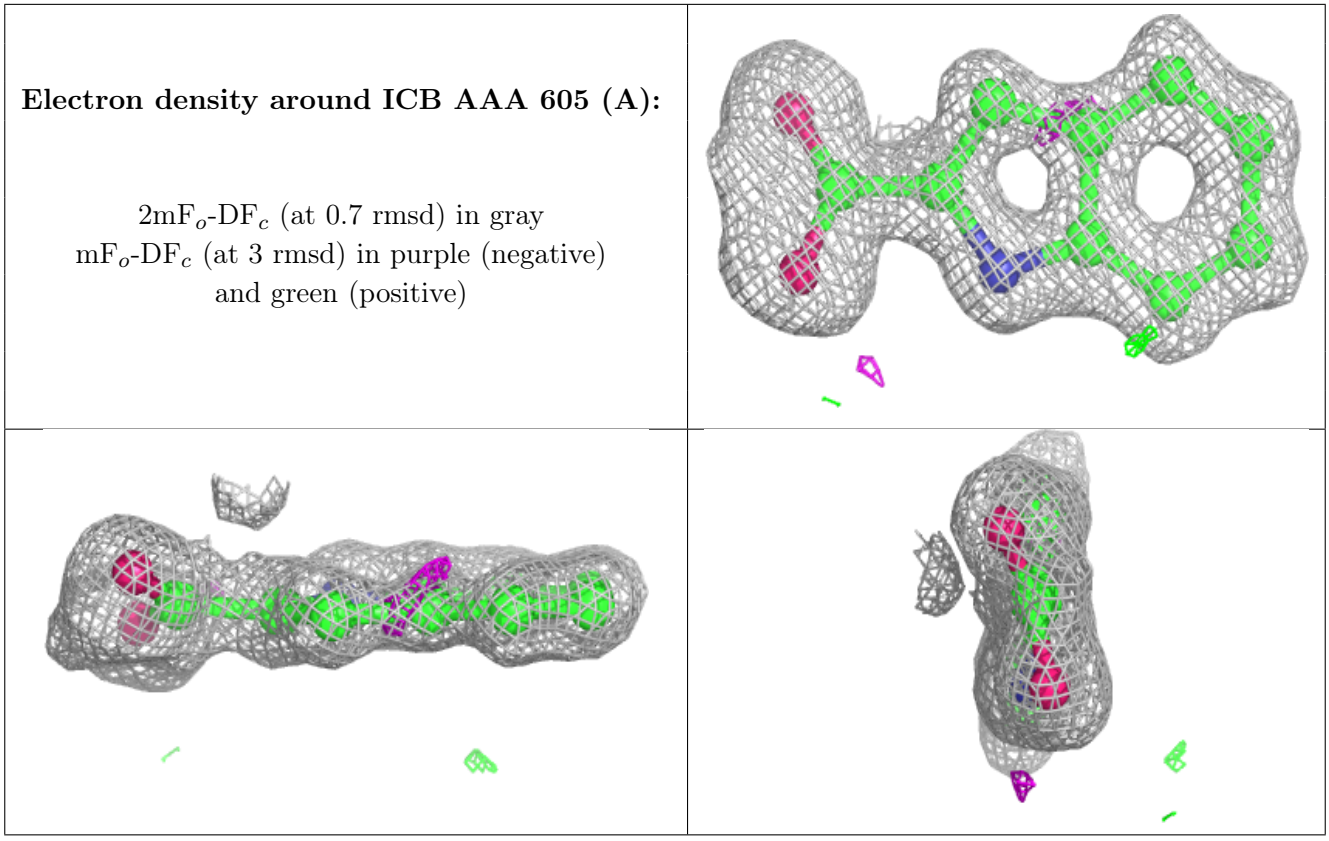






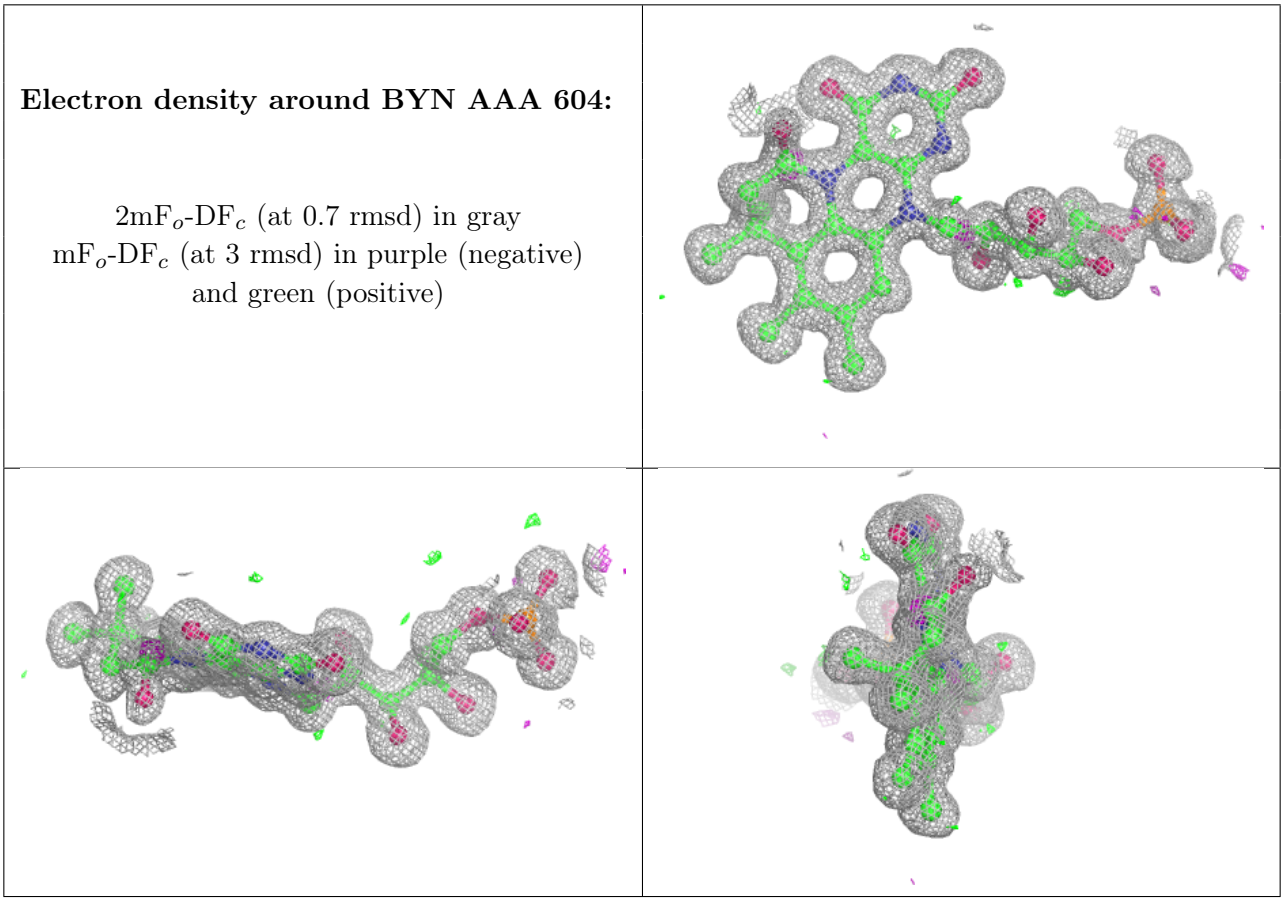
Electron density around ICB AAA 605 (A):

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



Electron density around BYN AAA 604:

$2mF_o$-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.