Full wwPDB X-ray Structure Validation Report ⓘ

Nov 13, 2024 – 03:21 pm GMT

PDB ID	:	5K3Y
Title	:	Crystal structure of AuroraB/INCENP in complex with BI 811283
Authors	:	Bader, G.; Zahn, S.K.; Zoephel, A.
Deposited on	:	2016-05-20
Resolution	:	1.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

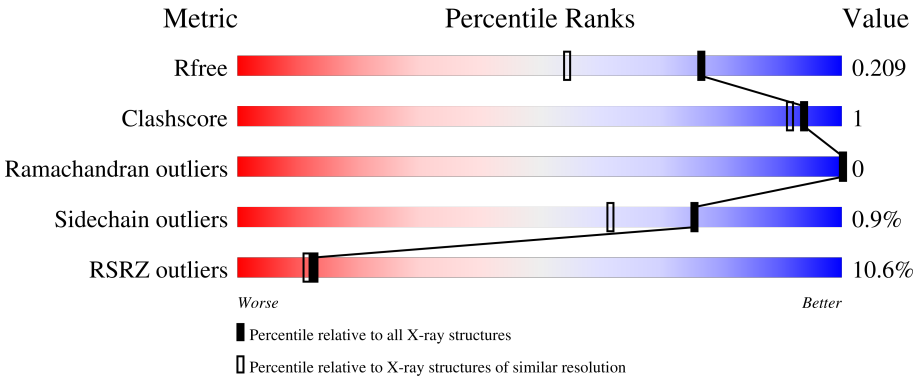
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.60 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}	164625	4274 (1.60-1.60)
Clashscore	180529	4682 (1.60-1.60)
Ramachandran outliers	177936	4583 (1.60-1.60)
Sidechain outliers	177891	4582 (1.60-1.60)
RSRZ outliers	164620	4272 (1.60-1.60)

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	280	5% 91% 5% .
1	B	280	8% 95% . .
2	C	59	19% 59% 41%
2	D	59	34% 59% . 39%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5878 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 Aurora kinase B-A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	268	Total	C	N	O	P	S	0	2	0
			2251	1447	402	388	1	13			
1	B	278	Total	C	N	O	P	S	0	5	0
			2346	1507	419	403	1	16			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	GLY	-	expression tag	UNP Q6DE08
A	96	VAL	GLY	conflict	UNP Q6DE08
B	77	GLY	-	expression tag	UNP Q6DE08
B	96	VAL	GLY	conflict	UNP Q6DE08

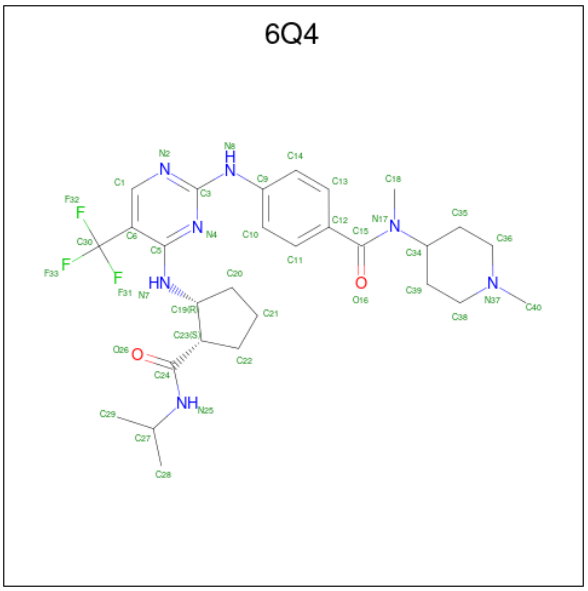
- Molecule 2 is a protein called Inner centromere protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	35	Total	C	N	O	S	0	0	0
			288	184	47	56	1			
2	D	36	Total	C	N	O	S	0	0	0
			301	192	50	58	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	789	MET	-	initiating methionine	UNP O13024
D	789	MET	-	initiating methionine	UNP O13024

- Molecule 3 is N-methyl-N-(1-methylpiperidin-4-yl)-4-{[4-({(1R,2S)-2-[(propan-2-yl)carbamoyl]cyclopentyl}amino)-5-(trifluoromethyl)pyrimidin-2-yl]amino}benzamide (three-letter code: 6Q4) (formula: $C_{28}H_{38}F_3N_7O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C F N O 40 28 3 7 2	0	0
3	B	1	Total C F N O 40 28 3 7 2	0	0

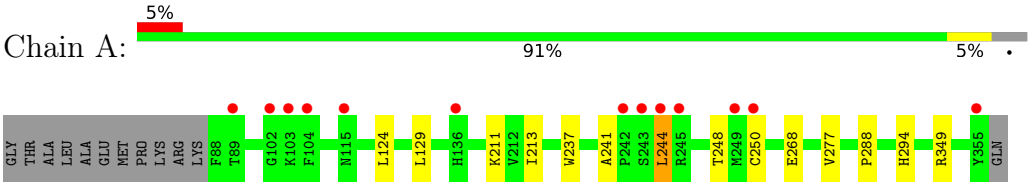
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	287	Total O 287 287	0	0
4	B	272	Total O 272 272	0	0
4	C	26	Total O 26 26	0	0
4	D	27	Total O 27 27	0	0

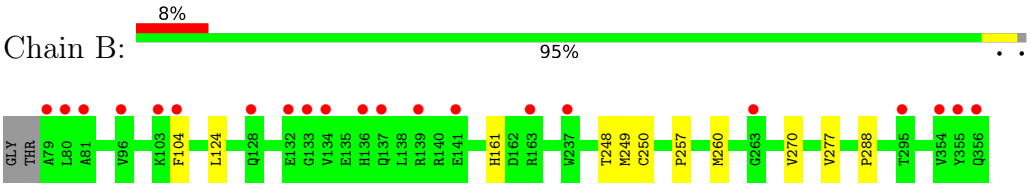
3 Residue-property plots (i)

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.

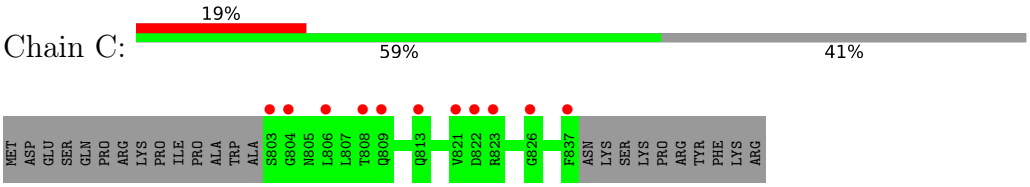
• Molecule 1: Aurora kinase B-A



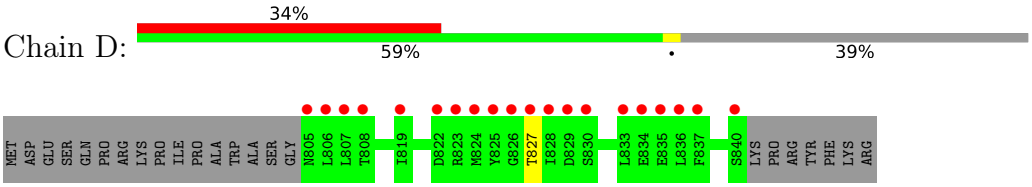
• Molecule 1: Aurora kinase B-A



• Molecule 2: Inner centromere protein A



• Molecule 2: Inner centromere protein A



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.91Å 67.30Å 116.62Å 90.00° 96.19° 90.00°	Depositor
Resolution (Å)	32.00 – 1.60 32.00 – 1.60	Depositor EDS
% Data completeness (in resolution range)	98.0 (32.00-1.60) 97.9 (32.00-1.60)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	1.98 (at 1.60Å)	Xtriage
Refinement program	BUSTER-TNT 2.11.6	Depositor
R, R_{free}	0.180 , 0.201 0.187 , 0.209	Depositor DCC
R_{free} test set	4543 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	18.9	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 46.5	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	0.96	EDS
Total number of atoms	5878	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.98% 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: 6Q4, 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	0.36	0/2301	0.56	0/3099
1	B	0.36	0/2403	0.56	0/3234
2	C	0.38	0/293	0.55	0/395
2	D	0.38	0/306	0.53	0/412
All	All	0.36	0/5303	0.55	0/7140

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	2251	0	2259	7	0
1	B	2346	0	2369	7	0
2	C	288	0	285	0	0
2	D	301	0	301	0	0
3	A	40	0	0	0	0
3	B	40	0	0	0	0
4	A	287	0	0	0	0
4	B	272	0	0	0	0
4	C	26	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	27	0	0	0	0
All	All	5878	0	5214	14	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.

All (14) 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:B:250:CYS:H	1:B:260[B]:MET:HE1	1.30	0.95
1:B:250:CYS:N	1:B:260[B]:MET:HE1	2.06	0.67
1:A:277:VAL:HG13	1:A:288:PRO:HD2	1.87	0.56
1:A:237:TRP:CD2	1:A:250:CYS:HB2	2.42	0.54
1:A:244[A]:LEU:HD11	1:A:268:GLU:HB3	1.90	0.54
1:A:213:ILE:HD11	1:A:244[A]:LEU:HD13	1.91	0.51
1:B:277:VAL:HG13	1:B:288:PRO:HD2	1.92	0.50
1:A:124:LEU:HB3	1:A:129:LEU:HD11	1.94	0.49
1:B:257:PRO:HG2	1:B:270[B]:VAL:HG13	1.99	0.45
1:B:104:PHE:HB2	1:B:124:LEU:HD23	1.99	0.45
1:B:249:MET:HA	1:B:260[B]:MET:HE1	2.01	0.43
1:A:211:LYS:HE2	1:A:244[B]:LEU:HD11	2.03	0.41
1:B:257:PRO:HD2	1:B:260[B]:MET:HB2	2.02	0.40
1:A:241:ALA:O	1:A:244[A]:LEU:HD22	2.22	0.40

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	267/280 (95%)	260 (97%)	7 (3%)	0	100	100
1	B	280/280 (100%)	272 (97%)	8 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	33/59 (56%)	33 (100%)	0	0	100	100
2	D	34/59 (58%)	31 (91%)	3 (9%)	0	100	100
All	All	614/678 (91%)	596 (97%)	18 (3%)	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	A	244/251 (97%)	240 (98%)	4 (2%)	58	37
1	B	255/251 (102%)	254 (100%)	1 (0%)	89	82
2	C	32/54 (59%)	32 (100%)	0	100	100
2	D	34/54 (63%)	33 (97%)	1 (3%)	37	15
All	All	565/610 (93%)	559 (99%)	6 (1%)	75	53

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	244[A]	LEU
1	A	244[B]	LEU
1	A	294	HIS
1	A	349	ARG
1	B	161	HIS
2	D	827	THR

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)

2 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	A	248	1	8,10,11	1.24	1 (12%)	10,14,16	1.15	1 (10%)
1	TPO	B	248	1	8,10,11	1.49	1 (12%)	10,14,16	1.34	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	A	248	1	-	0/9/11/13	-
1	TPO	B	248	1	-	0/9/11/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	248	TPO	P-OG1	-3.14	1.53	1.59
1	A	248	TPO	P-OG1	-2.63	1.54	1.59

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	248	TPO	P-OG1-CB	-3.20	113.55	123.21
1	A	248	TPO	P-OG1-CB	-2.48	115.72	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands 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	$\#\|Z\| > 2$
3	6Q4	B	401	-	43,43,43	0.62	1 (2%)	58,62,62	0.54	0
3	6Q4	A	401	-	43,43,43	0.67	1 (2%)	58,62,62	1.05	1 (1%)

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	6Q4	B	401	-	-	0/34/54/54	0/4/4/4
3	6Q4	A	401	-	-	5/34/54/54	0/4/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	401	6Q4	C34-N17	2.75	1.52	1.48
3	B	401	6Q4	C34-N17	2.52	1.52	1.48

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	401	6Q4	C18-N17-C34	-5.87	111.16	117.89

There are no chirality outliers.

All (5) torsion outliers are listed below:

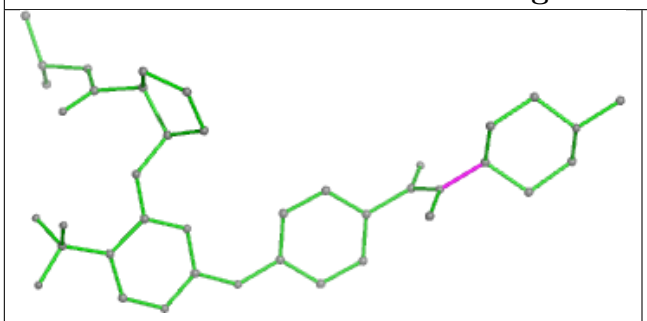
Mol	Chain	Res	Type	Atoms
3	A	401	6Q4	C35-C34-N17-C15
3	A	401	6Q4	C39-C34-N17-C15
3	A	401	6Q4	C35-C34-N17-C18
3	A	401	6Q4	C39-C34-N17-C18
3	A	401	6Q4	C29-C27-N25-C24

There are no ring outliers.

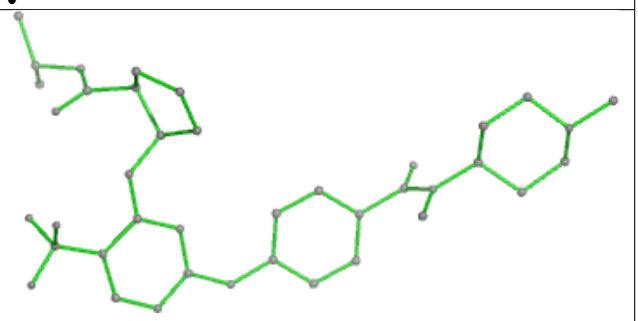
No monomer is involved in short contacts.

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.

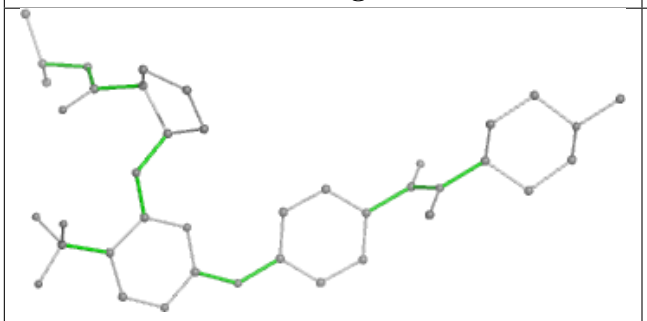
Ligand 6Q4 B 401



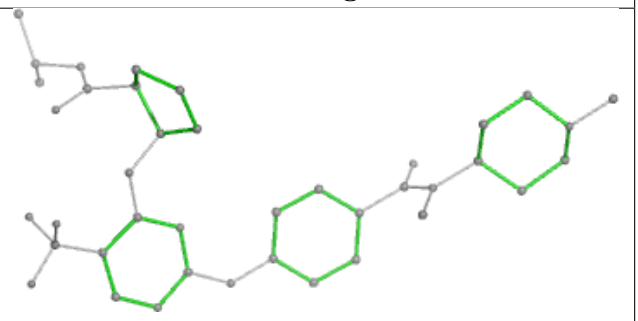
Bond lengths



Bond angles

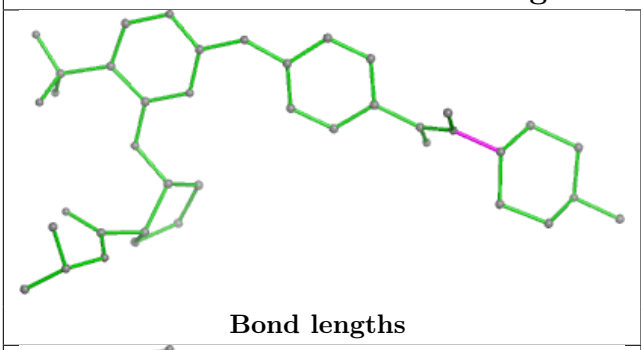


Torsions

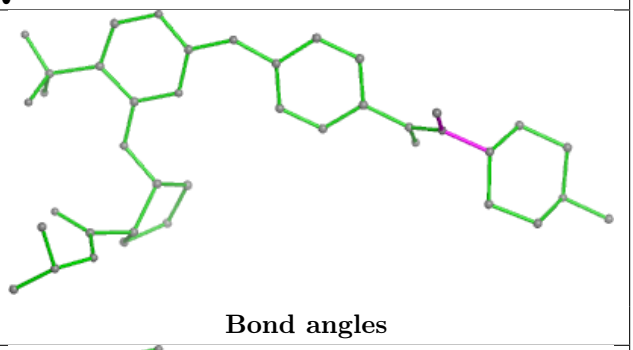


Rings

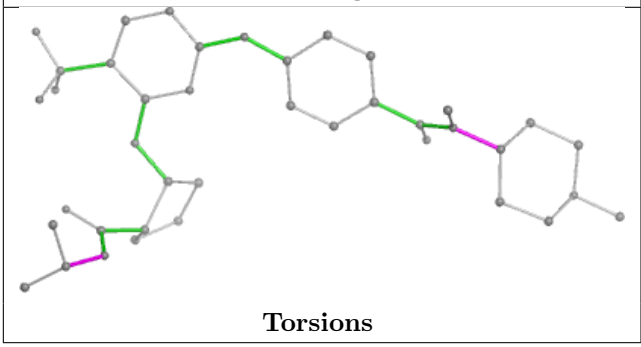
Ligand 6Q4 A 401



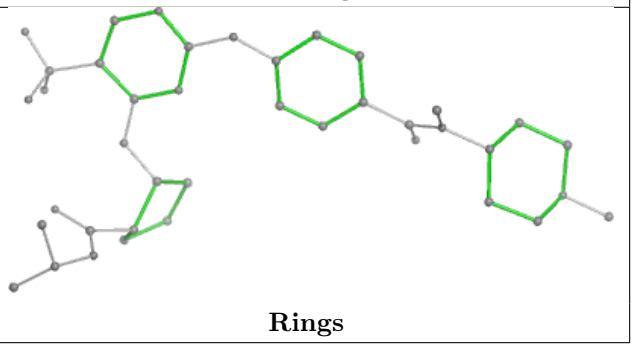
Bond lengths



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	267/280 (95%)	0.08	13 (4%)	36	35	8, 22, 45, 59	2 (0%)
1	B	277/280 (98%)	0.25	21 (7%)	21	20	6, 23, 54, 79	5 (1%)
2	C	35/59 (59%)	1.63	11 (31%)	1	1	28, 44, 60, 68	0
2	D	36/59 (61%)	1.99	20 (55%)	0	0	26, 51, 68, 76	0
All	All	615/678 (90%)	0.35	65 (10%)	13	11	6, 24, 55, 79	7 (1%)

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	104	PHE	6.8
1	B	79	ALA	5.0
2	C	837	PHE	5.0
1	A	244[A]	LEU	4.7
2	D	826	GLY	4.6
2	D	840	SER	4.3
1	A	355	TYR	4.2
2	D	827	THR	4.1
2	C	821	VAL	3.9
2	D	822	ASP	3.9
2	D	828	ILE	3.7
2	C	823	ARG	3.3
1	B	356	GLN	3.3
1	B	355	TYR	3.3
1	B	104	PHE	3.3
1	B	136	HIS	3.2
1	B	103	LYS	3.1
2	D	825	TYR	3.1
2	C	813	GLN	3.1
1	A	103	LYS	3.0
2	C	822	ASP	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	833	LEU	2.9
2	D	805	ASN	2.9
1	B	134	VAL	2.9
1	B	80	LEU	2.7
1	B	139	ARG	2.7
1	B	132	GLU	2.7
1	B	263	GLY	2.7
1	A	89	THR	2.7
2	D	808	THR	2.7
1	A	115	ASN	2.6
2	C	808	THR	2.6
1	B	237	TRP	2.5
1	B	163	ARG	2.5
2	D	823	ARG	2.5
2	C	826	GLY	2.5
1	B	81	ALA	2.5
2	D	824	MET	2.4
2	D	835	GLU	2.4
2	C	809	GLN	2.4
2	D	829	ASP	2.4
1	A	249	MET	2.3
2	C	806	LEU	2.3
1	A	136	HIS	2.3
1	B	141	GLU	2.3
1	B	96	VAL	2.3
1	B	295	THR	2.3
2	D	806	LEU	2.3
2	C	803	SER	2.2
2	D	830	SER	2.2
1	B	133	GLY	2.2
2	C	804	GLY	2.2
1	B	354	VAL	2.2
1	B	128	GLN	2.1
2	D	837	PHE	2.1
1	A	242	PRO	2.1
1	A	243	SER	2.1
2	D	836	LEU	2.1
1	A	245	ARG	2.1
2	D	834	GLU	2.1
1	A	102	GLY	2.1
1	A	250	CYS	2.1
2	D	807	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	137	GLN	2.0
2	D	819	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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
1	TPO	B	248	11/12	0.95	0.09	26,27,30,31	0
1	TPO	A	248	11/12	0.98	0.07	21,23,26,27	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

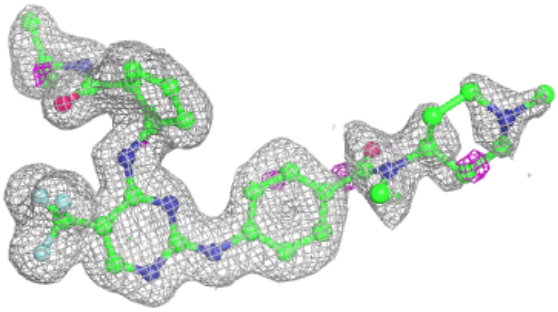
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

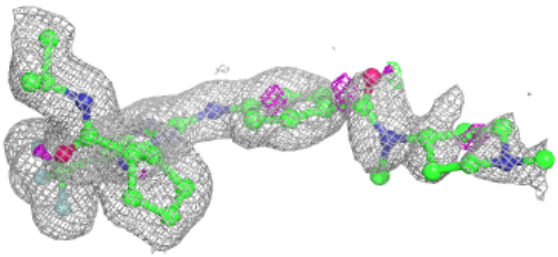
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	6Q4	B	401	40/40	0.93	0.10	18,28,49,50	0
3	6Q4	A	401	40/40	0.95	0.10	13,20,51,53	0

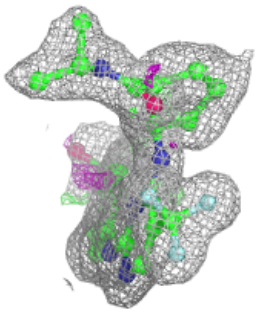
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around 6Q4 B 401:

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

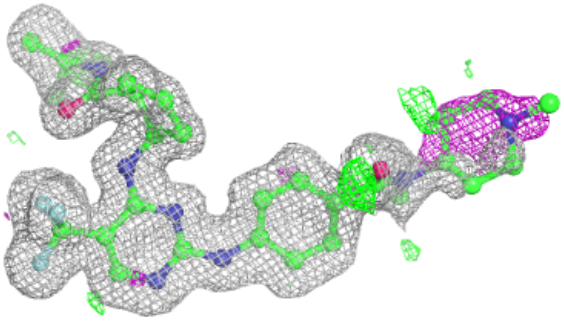


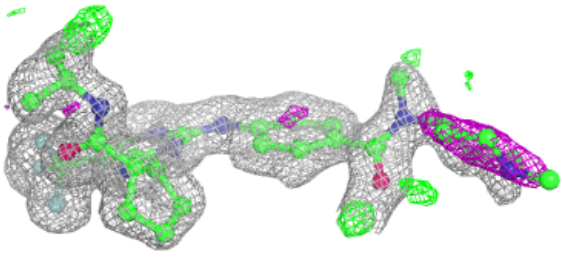


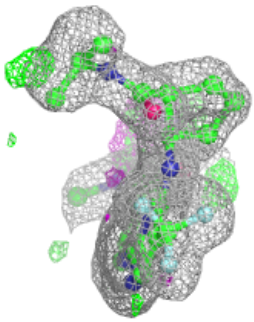


Electron density around 6Q4 A 401:

$2mF_o$-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.