Full wwPDB X-ray Structure Validation Report (i)

Aug 23, 2023 – 02:10 PM EDT

PDB ID	:	3CJG
Title	:	Crystal structure of VEGFR2 in complex with a 3,4,5-trimethoxy aniline containing pyrimidine
Authors	:	Nolte, R.T.
Deposited on	:	2008-03-12
Resolution	:	2.25 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 (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:

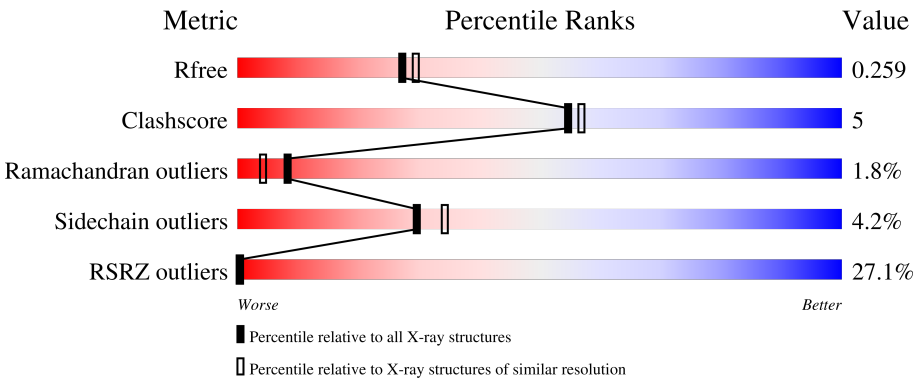
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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 2.25 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	309	25% 78% 13% • 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2460 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.

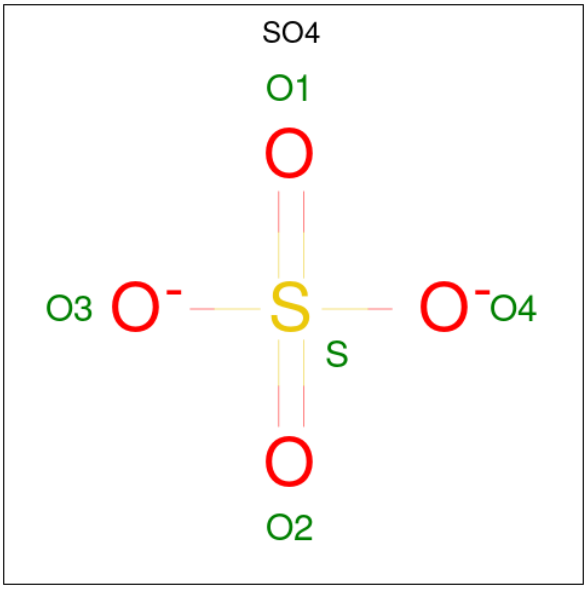
- Molecule 1 is a protein called Vascular endothelial growth factor receptor 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	282	2303	1468	401	416	18	0	6	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	833	ASN	LYS	See remark 999	UNP P35968

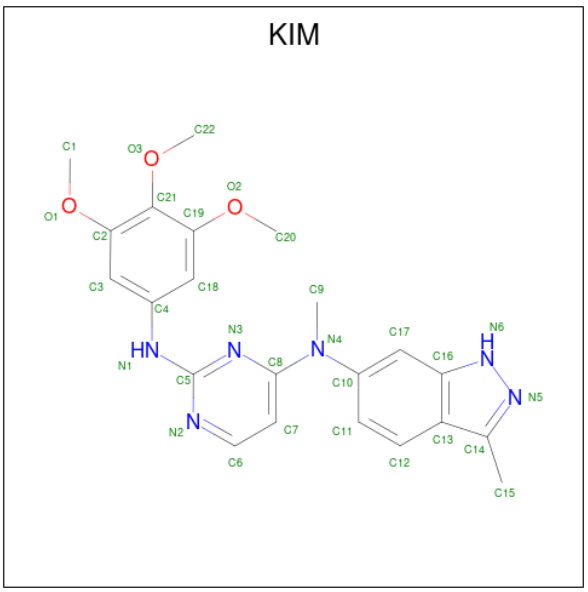
- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is N 4 -methyl-N 4 -(3-methyl-1H-indazol-6-yl)-N 2 -(3,4,5-trimethoxyphenyl)py

rimidine-2,4-diamine (three-letter code: KIM) (formula: $C_{22}H_{24}N_6O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	31	22	6	3	0	0

- Molecule 4 is water.

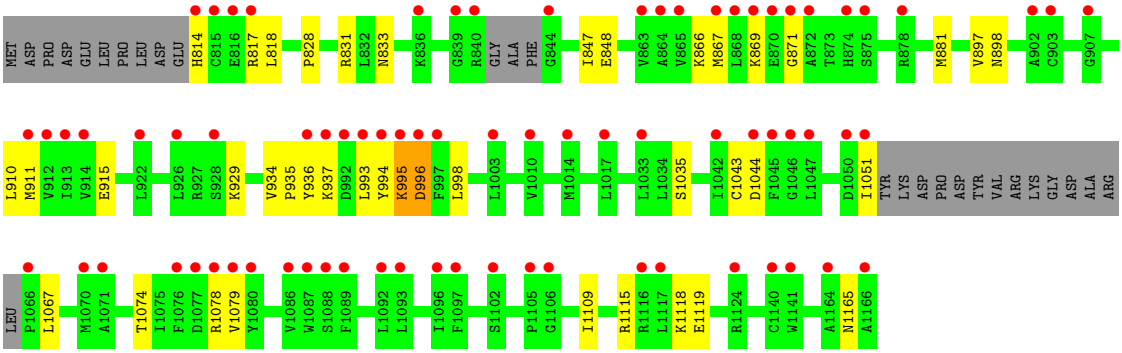
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	116	116	116	0	0

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: Vascular endothelial growth factor receptor 2

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	37.70Å 95.54Å 97.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.95 – 2.25 19.95 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.1 (19.95-2.25) 99.2 (19.95-2.25)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.03 (at 2.26Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.207 , 0.253 0.213 , 0.259	Depositor DCC
R_{free} test set	544 reflections (3.19%)	wwPDB-VP
Wilson B-factor (Å^2)	45.5	Xtriage
Anisotropy	0.455	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 59.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.017 for -h,l,k	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2460	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.46% 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: KIM, CSO, SO4

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	0.52	0/2332	0.61	0/3145

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 (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	2303	0	2273	23	0
2	A	10	0	0	0	0
3	A	31	0	24	3	0
4	A	116	0	0	3	0
All	All	2460	0	2297	24	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.

All (24) 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:1074:THR:HG22	1:A:1079:VAL:O	1.80	0.82
1:A:897:VAL:HG21	1:A:1043:CYS:HB3	1.83	0.61
1:A:1067:LEU:HD23	1:A:1109:ILE:CG2	2.32	0.59
1:A:934:VAL:HG22	1:A:937:LYS:HD2	1.85	0.57
1:A:833:ASN:ND2	4:A:1282:HOH:O	2.38	0.57
1:A:866:LYS:HB3	3:A:1167:KIM:H15A	1.86	0.56
1:A:1115:ARG:CZ	1:A:1119[A]:GLU:OE2	2.56	0.54
1:A:828:PRO:HG2	1:A:831:ARG:HD2	1.89	0.54
1:A:847:ILE:HG22	1:A:848:GLU:O	2.08	0.54
1:A:994:TYR:O	1:A:995:LYS:CB	2.56	0.53
1:A:881:MET:HE2	1:A:910:LEU:CD2	2.41	0.51
3:A:1167:KIM:N3	3:A:1167:KIM:H3	2.26	0.51
1:A:935:PRO:O	1:A:937:LYS:N	2.45	0.49
1:A:881:MET:HE2	1:A:910:LEU:HD21	1.94	0.49
1:A:867:MET:HB3	1:A:911:MET:HG2	1.95	0.48
1:A:897:VAL:CG2	1:A:1043:CYS:HB3	2.44	0.48
1:A:828:PRO:HG2	1:A:831:ARG:CD	2.44	0.47
1:A:866:LYS:CB	3:A:1167:KIM:H15A	2.44	0.47
1:A:1074:THR:CG2	1:A:1079:VAL:O	2.56	0.46
1:A:929:LYS:HE2	1:A:998:LEU:HD13	1.99	0.44
1:A:994:TYR:HA	4:A:1278:HOH:O	2.19	0.43
1:A:1115:ARG:NH1	1:A:1119[A]:GLU:OE2	2.52	0.42
1:A:818:LEU:HD22	4:A:1189:HOH:O	2.19	0.41
1:A:898:ASN:HB2	1:A:915:GLU:CD	2.41	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	277/309 (90%)	264 (95%)	8 (3%)	5 (2%)	8 4

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	995	LYS
1	A	996	ASP
1	A	936	TYR
1	A	1044	ASP
1	A	871	GLY

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	244/267 (91%)	233 (96%)	11 (4%)	27 31

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	814	HIS
1	A	817	ARG
1	A	869	LYS
1	A	993	LEU
1	A	996	ASP
1	A	1035	SER
1	A	1051	ILE
1	A	1078	ARG
1	A	1118[A]	LYS
1	A	1118[B]	LYS
1	A	1165	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1165	ASN

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	CSO	A	1022	1	3,6,7	0.56	0	0,6,8	-	-
1	CSO	A	860[B]	1	3,6,7	0.60	0	0,6,8	-	-
1	CSO	A	860[A]	1	3,6,7	0.62	0	0,6,8	-	-

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	CSO	A	1022	1	-	0/1/5/7	-
1	CSO	A	860[B]	1	-	0/1/5/7	-
1	CSO	A	860[A]	1	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

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)

3 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	$\#\|Z\| > 2$
2	SO4	A	1	-	4,4,4	0.12	0	6,6,6	0.27	0
2	SO4	A	2	-	4,4,4	0.15	0	6,6,6	0.14	0
3	KIM	A	1167	-	32,34,34	1.14	3 (9%)	40,48,48	2.17	12 (30%)

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	KIM	A	1167	-	-	2/18/18/18	0/4/4/4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1167	KIM	C15-C14	-3.23	1.48	1.50
3	A	1167	KIM	C17-C16	-2.64	1.37	1.41
3	A	1167	KIM	C5-N1	2.30	1.41	1.36

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1167	KIM	C10-C17-C16	-7.03	116.08	121.02
3	A	1167	KIM	C1-O1-C2	-6.34	107.96	117.53
3	A	1167	KIM	C20-O2-C19	-4.07	111.38	117.53
3	A	1167	KIM	N3-C8-N4	3.49	121.18	115.82
3	A	1167	KIM	C22-O3-C21	-3.15	106.15	114.78
3	A	1167	KIM	C10-N4-C8	3.13	127.39	120.80
3	A	1167	KIM	C11-C12-C13	-2.84	117.18	121.13
3	A	1167	KIM	C7-C8-N3	-2.46	119.52	123.53
3	A	1167	KIM	O1-C2-C21	2.32	119.24	115.16
3	A	1167	KIM	C17-C16-C13	2.15	123.93	121.16
3	A	1167	KIM	N2-C5-N3	-2.08	124.58	126.55
3	A	1167	KIM	O1-C2-C3	-2.03	120.62	124.12

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1167	KIM	C21-C2-O1-C1
3	A	1167	KIM	C3-C2-O1-C1

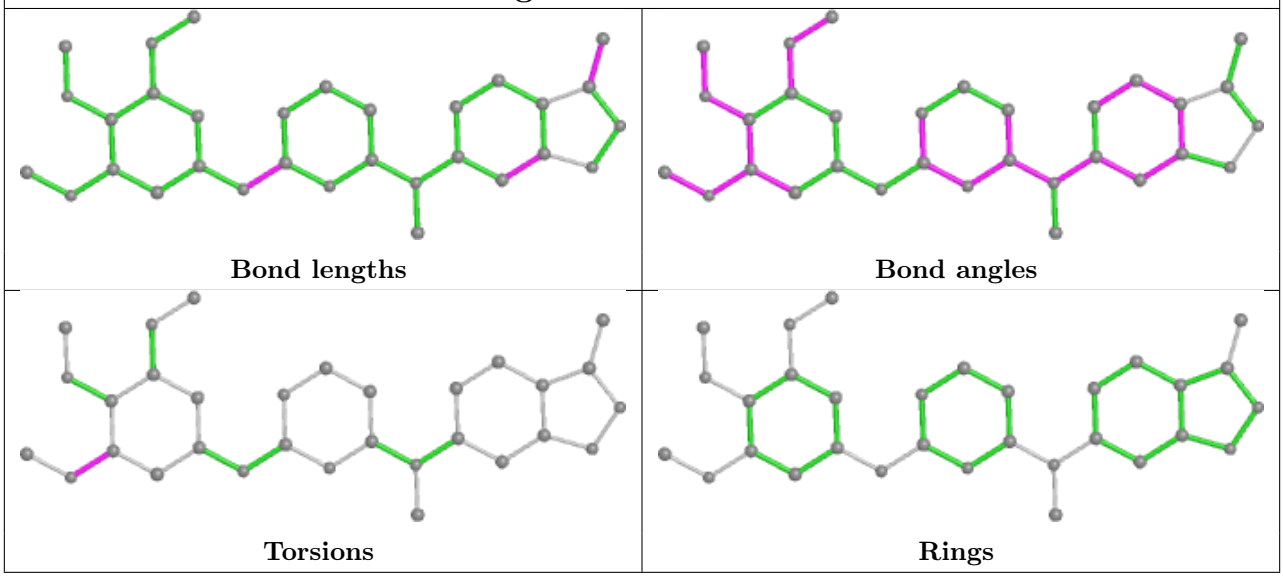
There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1167	KIM	3	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 KIM A 1167



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	937:LYS	C	992:ASP	N	5.83

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	A	280/309 (90%)	1.40	76 (27%) 0 0	47, 57, 71, 79	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	993	LEU	8.1
1	A	1106	GLY	7.7
1	A	936	TYR	7.3
1	A	1046	GLY	6.6
1	A	815	CYS	6.6
1	A	816	GLU	6.4
1	A	872	ALA	6.2
1	A	871	GLY	6.2
1	A	875	SER	6.1
1	A	1047	LEU	5.9
1	A	1078	ARG	5.7
1	A	865	VAL	5.7
1	A	814	HIS	5.6
1	A	1079	VAL	5.5
1	A	994	TYR	5.4
1	A	870	GLU	5.1
1	A	1166	ALA	5.0
1	A	1050	ASP	4.9
1	A	1077	ASP	4.9
1	A	913	ILE	4.9
1	A	912	VAL	4.3
1	A	864	ALA	4.3
1	A	1071	ALA	4.2
1	A	840	ARG	4.1
1	A	995	LYS	3.9
1	A	1164	ALA	3.9
1	A	1087	TRP	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	817	ARG	3.6
1	A	1066	PRO	3.6
1	A	836	LYS	3.6
1	A	1089	PHE	3.5
1	A	922	LEU	3.4
1	A	926	LEU	3.3
1	A	1096	ILE	3.3
1	A	1076	PHE	3.3
1	A	1117	LEU	3.3
1	A	911	MET	3.2
1	A	1051	ILE	3.2
1	A	1080	TYR	3.0
1	A	878	ARG	3.0
1	A	996	ASP	2.9
1	A	839	GLY	2.9
1	A	1141	TRP	2.9
1	A	1093	LEU	2.9
1	A	869	LYS	2.8
1	A	867	MET	2.7
1	A	1124	ARG	2.6
1	A	1044	ASP	2.6
1	A	868	LEU	2.6
1	A	1086	VAL	2.6
1	A	1092	LEU	2.5
1	A	1003	LEU	2.5
1	A	1033	LEU	2.5
1	A	1097	PHE	2.5
1	A	1105	PRO	2.4
1	A	1088	SER	2.4
1	A	1045	PHE	2.4
1	A	1140	CYS	2.3
1	A	1017	LEU	2.3
1	A	1042	ILE	2.3
1	A	914	VAL	2.3
1	A	1070	MET	2.2
1	A	992	ASP	2.2
1	A	1102	SER	2.2
1	A	1014	MET	2.1
1	A	928	SER	2.1
1	A	1010	VAL	2.1
1	A	997	PHE	2.1
1	A	844	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	907	GLY	2.1
1	A	1116	ARG	2.1
1	A	863	VAL	2.1
1	A	903	CYS	2.1
1	A	874	HIS	2.0
1	A	902	ALA	2.0
1	A	937	LYS	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, 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
1	CSO	A	1022	7/8	0.63	0.25	52,58,59,60	0
1	CSO	A	860[B]	7/8	0.80	0.19	53,54,56,56	7
1	CSO	A	860[A]	7/8	0.80	0.19	45,52,52,52	7

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

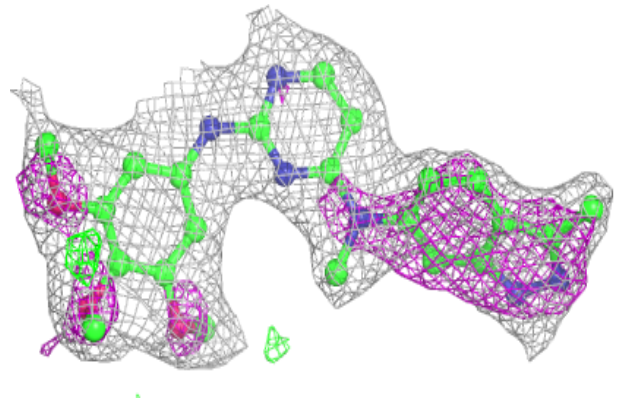
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

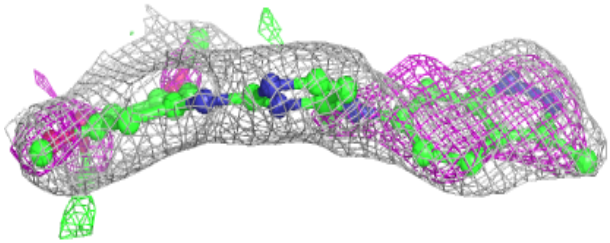
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	A	2	5/5	0.85	0.34	99,99,100,100	0
3	KIM	A	1167	31/31	0.86	0.19	45,51,53,54	0
2	SO4	A	1	5/5	0.95	0.43	94,95,95,95	0

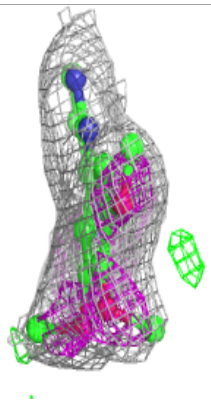
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around KIM A 1167:

$2mF_o$-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.