



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2016 – 11:12 AM EDT

PDB ID : 5EOB
Title : Crystal structure of CMET in complex with novel inhibitor
Authors : Liu, Q.; Chen, T.; Xu, Y.
Deposited on : 2015-11-10
Resolution : 1.75 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

<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

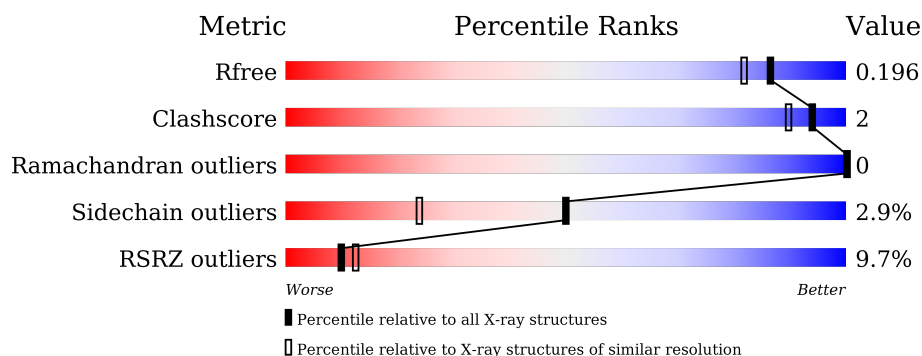
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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}	91344	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	319	<div> <div>9%</div> <div>84%</div> <div>5%</div> <div>10%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2408 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.

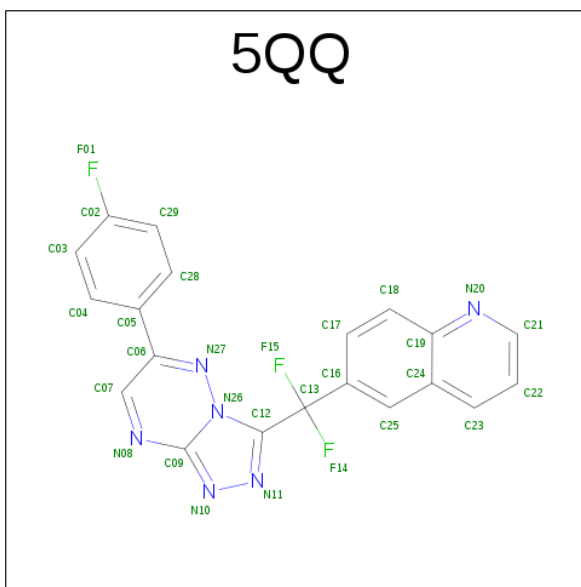
- Molecule 1 is a protein called Hepatocyte growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	288	2260	1460	382	403	15	0	1	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1028	MET	-	initiating methionine	UNP P08581
A	1029	GLY	-	expression tag	UNP P08581
A	1030	SER	-	expression tag	UNP P08581
A	1031	SER	-	expression tag	UNP P08581
A	1032	HIS	-	expression tag	UNP P08581
A	1033	HIS	-	expression tag	UNP P08581
A	1034	HIS	-	expression tag	UNP P08581
A	1035	HIS	-	expression tag	UNP P08581
A	1036	HIS	-	expression tag	UNP P08581
A	1037	HIS	-	expression tag	UNP P08581

- Molecule 2 is 6-[bis(fluoranyl)-[6-(4-fluorophenyl)-[1,2,4]triazolo[4,3-b][1,2,4]triazin-3-yl]methyl]quinoline (three-letter code: 5QQ) (formula: C₂₀H₁₁F₃N₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	F	N	0	0
			29	20	3	6		

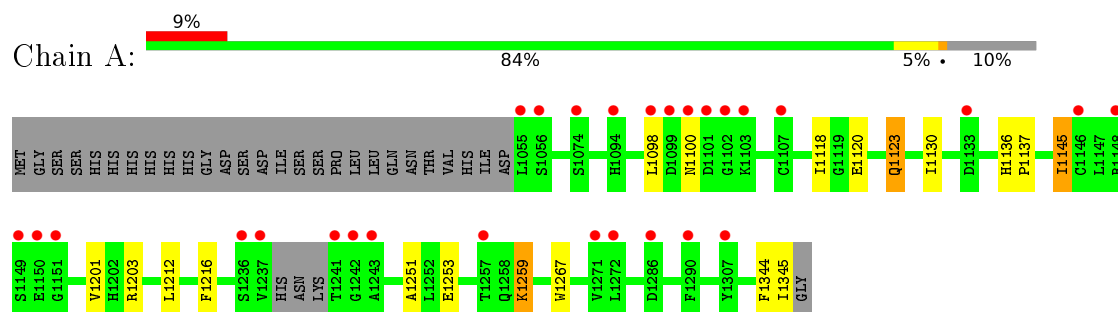
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	119	Total	O	0	0
			119	119		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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: Hepatocyte growth factor receptor



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	37.88Å 43.39Å 86.08Å 90.00° 92.68° 90.00°	Depositor
Resolution (Å)	23.91 – 1.75 23.92 – 1.75	Depositor EDS
% Data completeness (in resolution range)	94.1 (23.91-1.75) 98.1 (23.92-1.75)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.86 (at 1.75Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.173 , 0.197 0.173 , 0.196	Depositor DCC
R_{free} test set	835 reflections (3.00%)	DCC
Wilson B-factor (Å ²)	17.2	Xtriage
Anisotropy	0.396	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.47 , 72.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.042 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2408	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.01% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 5QQ

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	0.35	0/2317	0.53	0/3141

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	2260	0	2259	8	0
2	A	29	0	0	0	0
3	A	119	0	0	0	0
All	All	2408	0	2259	8	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 2.

All (8) 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:1253:GLU:HG3	1:A:1259:LYS:HE2	1.91	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1120:GLU:O	1:A:1123:GLN:HG3	2.15	0.46
1:A:1344:PHE:O	1:A:1345:ILE:HB	2.17	0.43
1:A:1212:LEU:HD22	1:A:1216:PHE:HA	2.00	0.42
1:A:1136:HIS:CG	1:A:1137:PRO:HD2	2.55	0.42
1:A:1201:VAL:O	1:A:1203:ARG:HD3	2.20	0.41
1:A:1251:ALA:HA	1:A:1267:TRP:CD2	2.55	0.41
1:A:1145:ILE:H	1:A:1145:ILE:HD13	1.86	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	285/319 (89%)	282 (99%)	3 (1%)	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	246/283 (87%)	239 (97%)	7 (3%)	51	25

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1098	LEU
1	A	1100	ASN
1	A	1118	ILE
1	A	1123	GLN
1	A	1130	ILE
1	A	1145	ILE
1	A	1259	LYS

Some 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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	$\# Z > 2$
2	5QQ	A	1401	-	28,33,33	1.85	6 (21%)	33,49,49	1.85	6 (18%)

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
2	5QQ	A	1401	-	-	0/12/16/16	0/5/5/5

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1401	5QQ	C05-C06	-5.73	1.39	1.48
2	A	1401	5QQ	F14-C13	-2.95	1.33	1.37
2	A	1401	5QQ	C24-C19	-2.58	1.38	1.42
2	A	1401	5QQ	F15-C13	-2.50	1.34	1.37
2	A	1401	5QQ	C07-N08	2.37	1.35	1.31
2	A	1401	5QQ	C09-N08	4.41	1.38	1.32

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1401	5QQ	C06-C07-N08	-4.26	119.64	123.70
2	A	1401	5QQ	F14-C13-C16	-3.25	107.61	109.92
2	A	1401	5QQ	C22-C21-N20	-2.11	120.58	123.94
2	A	1401	5QQ	C13-C12-N11	2.16	127.18	122.63
2	A	1401	5QQ	C21-N20-C19	2.57	121.29	116.88
2	A	1401	5QQ	F15-C13-F14	5.33	110.16	104.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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(Å ²)	Q<0.9
1	A	288/319 (90%)	0.50	28 (9%) 10 12	7, 16, 38, 62	5 (1%)

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1237	VAL	11.3
1	A	1241	THR	10.5
1	A	1242	GLY	10.2
1	A	1107	CYS	7.4
1	A	1101	ASP	6.9
1	A	1146	CYS	6.0
1	A	1102	GLY	5.5
1	A	1286	ASP	5.2
1	A	1100	ASN	4.7
1	A	1094	HIS	4.6
1	A	1243	ALA	4.2
1	A	1290	PHE	3.9
1	A	1055	LEU	3.6
1	A	1257	THR	3.6
1	A	1103	LYS	3.5
1	A	1307	TYR	3.3
1	A	1150	GLU	3.2
1	A	1074	SER	3.0
1	A	1272	LEU	2.9
1	A	1236	SER	2.7
1	A	1271	VAL	2.6
1	A	1149	SER	2.6
1	A	1151	GLY	2.5
1	A	1133	ASP	2.3
1	A	1056	SER	2.3
1	A	1099	ASP	2.3
1	A	1148	ARG	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1098	LEU	2.0

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
2	5QQ	A	1401	29/29	0.94	0.09	0.11	11,13,25,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.