



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 04:10 AM GMT

PDB ID : 4M3Q  
Title : Crystal structure of the catalytic domain of the proto-oncogene tyrosine-protein kinase MER in complex with inhibitor UNC1917  
Authors : Zhang, W.; Zhang, D.; Stashko, M.A.; DeRyckere, D.; Hunter, D.; Kireev, D.B.; Miley, M.; Cummings, C.; Lee, M.; Norris-Drouin, J.; Stewart, W.M.; Sather, S.; Zhou, Y.; Kirkpatrick, G.; Machius, M.; Janzen, W.P.; Earp, H.S.; Graham, D.K.; Frye, S.; Wang, X.  
Deposited on : 2013-08-06  
Resolution : 2.72 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

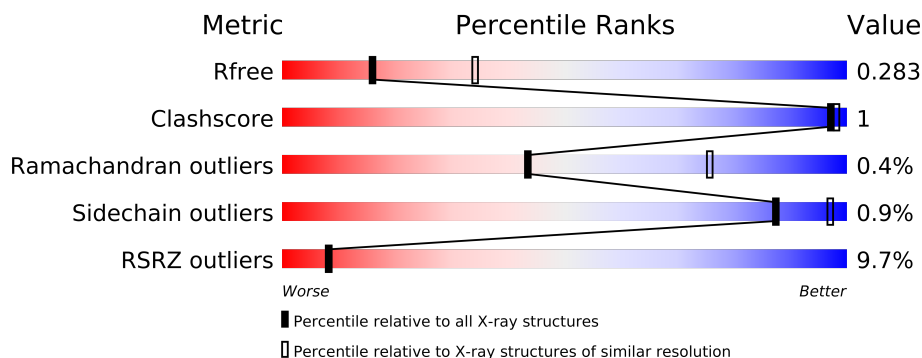
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.72 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))
$R_{free}$	66092	1770 (2.74-2.70)
Clashscore	79885	2183 (2.74-2.70)
Ramachandran outliers	78287	2147 (2.74-2.70)
Sidechain outliers	78261	2148 (2.74-2.70)
RSRZ outliers	66119	1772 (2.74-2.70)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	313	
1	B	313	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7899 atoms, of which 3975 are hydrogens and 0 are deuterium.

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 Tyrosine-protein kinase Mer.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	243	Total	C	H	N	O	S	0	0	0
			3937	1257	1980	326	355	19			
1	B	236	Total	C	H	N	O	S	0	0	0
			3844	1229	1941	317	338	19			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	552	MET	-	EXPRESSION TAG	UNP Q12866
A	553	GLY	-	EXPRESSION TAG	UNP Q12866
A	554	SER	-	EXPRESSION TAG	UNP Q12866
A	555	SER	-	EXPRESSION TAG	UNP Q12866
A	556	HIS	-	EXPRESSION TAG	UNP Q12866
A	557	HIS	-	EXPRESSION TAG	UNP Q12866
A	558	HIS	-	EXPRESSION TAG	UNP Q12866
A	559	HIS	-	EXPRESSION TAG	UNP Q12866
A	560	HIS	-	EXPRESSION TAG	UNP Q12866
A	561	HIS	-	EXPRESSION TAG	UNP Q12866
A	562	SER	-	EXPRESSION TAG	UNP Q12866
A	563	SER	-	EXPRESSION TAG	UNP Q12866
A	564	GLY	-	EXPRESSION TAG	UNP Q12866
A	565	LEU	-	EXPRESSION TAG	UNP Q12866
A	566	VAL	-	EXPRESSION TAG	UNP Q12866
A	567	PRO	-	EXPRESSION TAG	UNP Q12866
A	568	ARG	-	EXPRESSION TAG	UNP Q12866
A	569	GLY	-	EXPRESSION TAG	UNP Q12866
B	552	MET	-	EXPRESSION TAG	UNP Q12866
B	553	GLY	-	EXPRESSION TAG	UNP Q12866
B	554	SER	-	EXPRESSION TAG	UNP Q12866
B	555	SER	-	EXPRESSION TAG	UNP Q12866
B	556	HIS	-	EXPRESSION TAG	UNP Q12866
B	557	HIS	-	EXPRESSION TAG	UNP Q12866
B	558	HIS	-	EXPRESSION TAG	UNP Q12866

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	559	HIS	-	EXPRESSION TAG	UNP Q12866
B	560	HIS	-	EXPRESSION TAG	UNP Q12866
B	561	HIS	-	EXPRESSION TAG	UNP Q12866
B	562	SER	-	EXPRESSION TAG	UNP Q12866
B	563	SER	-	EXPRESSION TAG	UNP Q12866
B	564	GLY	-	EXPRESSION TAG	UNP Q12866
B	565	LEU	-	EXPRESSION TAG	UNP Q12866
B	566	VAL	-	EXPRESSION TAG	UNP Q12866
B	567	PRO	-	EXPRESSION TAG	UNP Q12866
B	568	ARG	-	EXPRESSION TAG	UNP Q12866
B	569	GLY	-	EXPRESSION TAG	UNP Q12866

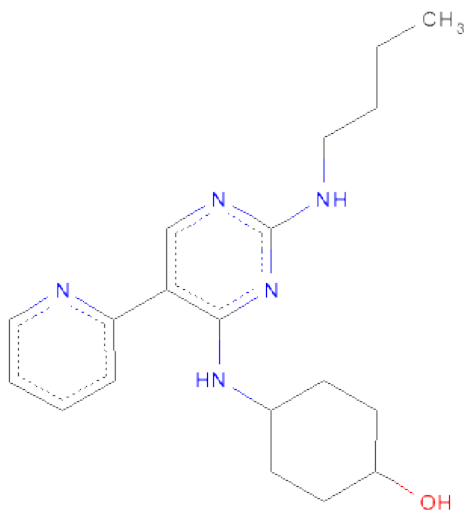
- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	3	Total Cl 3 3	0	0
2	A	4	Total Cl 4 4	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0

- Molecule 4 is TRANS-4-{[2-(BUTYLAMINO)-5-(PYRIDIN-2-YL)PYRIMIDIN-4-YL]AMINO}CYCLOHEXANOL (three-letter code: 24K) (formula: C<sub>19</sub>H<sub>27</sub>N<sub>5</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			52	19	27	5	1		
4	B	1	Total	C	H	N	O	0	0
			52	19	27	5	1		

- Molecule 5 is water.

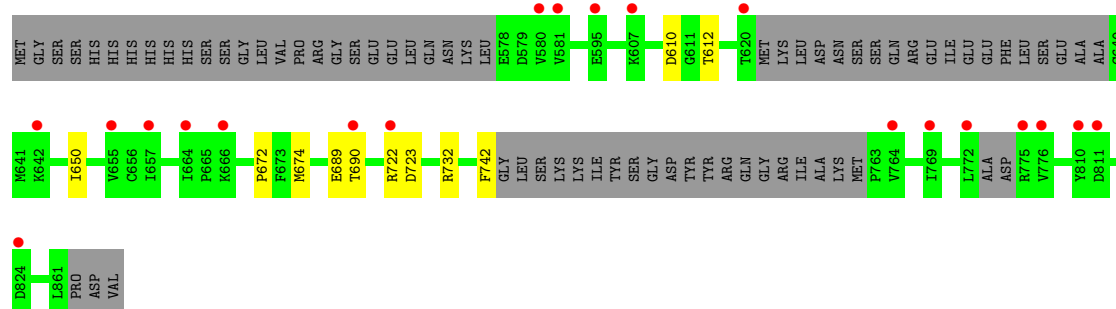
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	4	Total	O	0	0
			4	4		
5	B	2	Total	O	0	0
			2	2		

### 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 ( $RSRZ > 2$ ). 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.

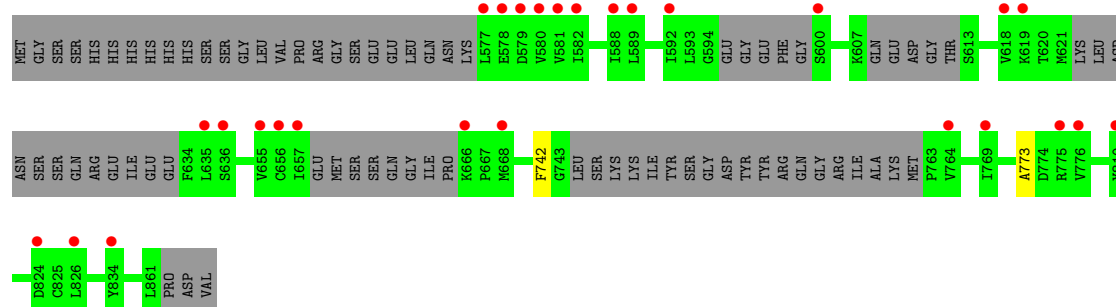
- Molecule 1: Tyrosine-protein kinase Mer

Chain A: 



- Molecule 1: Tyrosine-protein kinase Mer

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.92Å 91.12Å 69.70Å 90.00° 99.81° 90.00°	Depositor
Resolution (Å)	34.74 – 2.72 34.74 – 2.72	Depositor EDS
% Data completeness (in resolution range)	83.6 (34.74-2.72) 79.4 (34.74-2.72)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.71 (at 2.72Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1439)	Depositor
R, $R_{free}$	0.243 , 0.296 0.242 , 0.283	Depositor DCC
$R_{free}$ test set	1340 reflections (9.95%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.8	Xtriage
Anisotropy	0.391	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 18.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	1 of 14223 reflections (0.007%)	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	7899	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 24K, CL

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	0.21	0/1998	0.37	0/2698
1	B	0.21	0/1941	0.37	0/2619
All	All	0.21	0/3939	0.37	0/5317

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1957	1980	0	7	0
1	B	1903	1941	0	0	0
2	A	4	0	0	5	0
2	B	3	0	0	0	0
3	A	1	0	0	0	0
4	A	25	27	0	0	0
4	B	25	27	0	0	0
5	A	4	0	0	0	0
5	B	2	0	0	0	0
All	All	3924	3975	0	7	0



Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

The worst 5 of 7 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:732:ARG:NH1	2:A:904:CL:CL	2.55	0.75
1:A:689:GLU:O	1:A:690:THR:OG1	2.13	0.66
1:A:672:PRO:HG2	2:A:904:CL:CL	2.52	0.46
1:A:732:ARG:NH2	2:A:904:CL:CL	2.88	0.44
1:A:650:ILE:HA	2:A:904:CL:CL	2.56	0.42

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	235/313 (75%)	228 (97%)	6 (3%)	1 (0%)	43	76
1	B	224/313 (72%)	219 (98%)	4 (2%)	1 (0%)	43	76
All	All	459/626 (73%)	447 (97%)	10 (2%)	2 (0%)	43	76

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	723	ASP
1	B	773	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	220/280 (79%)	217 (99%)	3 (1%)	78	95
1	B	213/280 (76%)	212 (100%)	1 (0%)	94	99
All	All	433/560 (77%)	429 (99%)	4 (1%)	87	97

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	674	MET
1	A	722	ARG
1	A	742	PHE
1	B	742	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	648	ASN
1	B	821	GLN
1	B	728	ASN
1	A	728	ASN
1	B	807	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 8 are monoatomic - leaving 2 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	24K	A	906	-	27,27,27	1.94	2 (7%)	35,35,35	2.20	13 (37%)
4	24K	B	904	-	27,27,27	1.92	2 (7%)	35,35,35	2.00	10 (28%)

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	24K	A	906	-	-	0/13/23/23	0/3/3/3
4	24K	B	904	-	-	1/13/23/23	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	904	24K	C06-N05	7.33	1.46	1.34
4	A	906	24K	C06-N05	7.12	1.46	1.34
4	A	906	24K	C10-N11	6.51	1.45	1.35
4	B	904	24K	C10-N11	6.21	1.45	1.35

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	904	24K	N07-C06-N19	-7.61	120.13	126.68
4	A	906	24K	N07-C06-N19	-7.49	120.23	126.68
4	A	906	24K	C24-N25-C20	3.99	122.64	117.20
4	B	904	24K	C24-N25-C20	3.79	122.36	117.20
4	A	906	24K	C09-C08-N07	-3.65	120.96	125.23

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	904	24K	N19-C10-N11-C12

There are no ring outliers.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	243/313 (77%)	0.64	20 (8%) 12 13	17, 39, 76, 111	0
1	B	236/313 (75%)	0.68	27 (11%) 6 6	14, 37, 70, 110	0
All	All	479/626 (76%)	0.66	47 (9%) 8 8	14, 38, 73, 111	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	657	ILE	5.2
1	B	775	ARG	5.2
1	A	776	VAL	4.8
1	B	580	VAL	4.4
1	B	635	LEU	4.3

### 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 carbohydrates 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. 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, 95<sup>th</sup> 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	24K	B	904	25/25	0.28	1.22	41,52,65,66	0
2	CL	A	904	1/1	0.20	0.08	75,75,75,75	0
2	CL	B	903	1/1	0.16	0.02	66,66,66,66	0
4	24K	A	906	25/25	0.19	-0.08	12,32,48,51	0
2	CL	B	902	1/1	0.14	-1.12	44,44,44,44	0
3	MG	A	905	1/1	0.16	-1.16	18,18,18,18	0
2	CL	A	901	1/1	0.07	-3.70	28,28,28,28	0
2	CL	A	903	1/1	0.08	-3.76	46,46,46,46	0
2	CL	B	901	1/1	0.03	-5.21	26,26,26,26	0
2	CL	A	902	1/1	0.06	-5.37	30,30,30,30	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.