



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 04:44 AM GMT

PDB ID : 3W2R
Title : EGFR Kinase domain T790M/L858R mutant with compound 4
Authors : Sogabe, S.; Kawakita, Y.; Igaki, S.
Deposited on : 2012-12-03
Resolution : 2.05 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

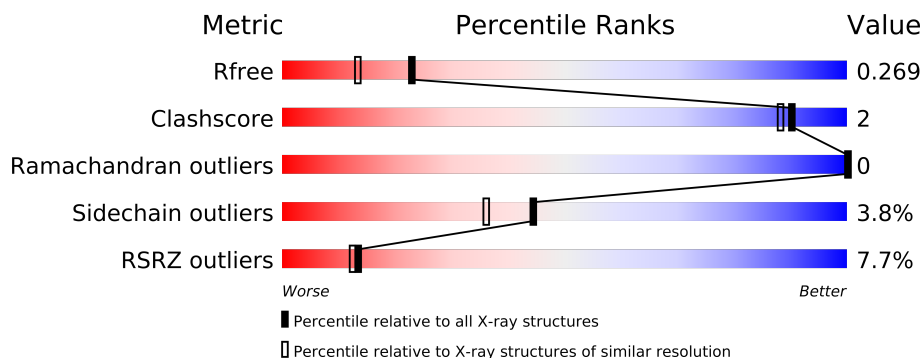
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.05 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	1380 (2.06-2.02)
Clashscore	79885	1577 (2.06-2.02)
Ramachandran outliers	78287	1565 (2.06-2.02)
Sidechain outliers	78261	1565 (2.06-2.02)
RSRZ outliers	66119	1381 (2.06-2.02)

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. 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	331	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2514 atoms, of which 0 are hydrogen 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.

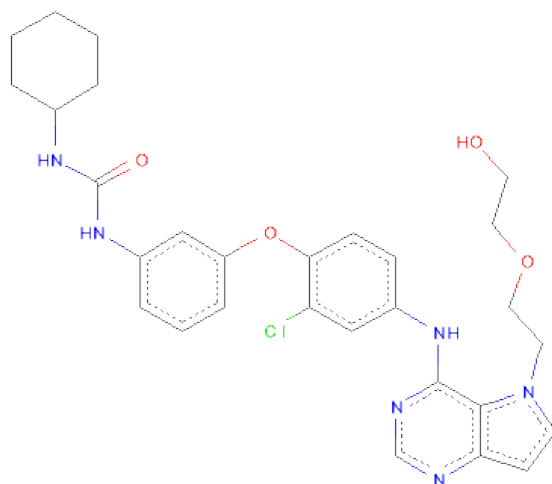
- Molecule 1 is a protein called Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	297	2394	1538	402	435	19	0	1	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	692	HIS	-	EXPRESSION TAG	UNP P00533
A	693	HIS	-	EXPRESSION TAG	UNP P00533
A	694	HIS	-	EXPRESSION TAG	UNP P00533
A	695	HIS	-	EXPRESSION TAG	UNP P00533
A	696	HIS	-	EXPRESSION TAG	UNP P00533
A	697	HIS	-	EXPRESSION TAG	UNP P00533
A	790	MET	THR	ENGINEERED MUTATION	UNP P00533
A	858	ARG	LEU	ENGINEERED MUTATION	UNP P00533

- Molecule 2 is 1-{3-[2-CHLORO-4-({5-[2-(2-HYDROXYETHOXY)ETHYL]-5H-PYRROLO[3,2-D]PYRIMIDIN-4-YL}AMINO)PHENOXY]PHENYL}-3-CYCLOHEXYLUREA (three-letter code: W2R) (formula: C₂₉H₃₃ClN₆O₄).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Cl	N	O		
2	A	1	40	29	1	6	4	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	4	2	2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	76	Total	O	0	0
			76	76		

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	61.16Å 61.16Å 169.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.05 30.29 – 2.05	Depositor EDS
% Data completeness (in resolution range)	97.9 (40.00-2.05) 98.0 (30.29-2.05)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.52 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.217 , 0.275 0.215 , 0.269	Depositor DCC
R_{free} test set	1066 reflections (5.44%)	DCC
Wilson B-factor (Å ²)	43.2	Xtriage
Anisotropy	0.451	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 43.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	2 of 20752 reflections (0.010%)	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2514	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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: W2R, EDO

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	0.50	0/2447	0.67	0/3309

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	2394	0	0	5	0
2	A	40	0	0	0	0
3	A	4	0	6	3	0
4	A	76	0	0	2	0
All	All	2514	0	6	5	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 2.

All (5) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:952:MET:CB	3:A:1102:EDO:H12	2.34	0.57
1:A:954:ASP:N	3:A:1102:EDO:HO1	2.05	0.55
1:A:852:LYS:NZ	4:A:2034:HOH:O	2.46	0.48
1:A:953:ILE:N	3:A:1102:EDO:HO2	2.15	0.44
1:A:982:GLN:N	4:A:2072:HOH:O	2.51	0.43

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	292/331 (88%)	282 (97%)	10 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	265/292 (91%)	255 (96%)	10 (4%)	44	36

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	732	ILE
1	A	784	SER
1	A	816	ASN
1	A	849	GLN
1	A	855	ASP
1	A	860	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	926	ILE
1	A	976	GLN
1	A	1001	LEU
1	A	1017	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

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	$\# Z > 2$
2	W2R	A	1101	-	44,44,44	0.86	2 (4%)	58,59,59	1.35	6 (10%)
3	EDO	A	1102	-	3,3,3	0.68	0	2,2,2	0.37	0

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	W2R	A	1101	-	-	0/22/30/30	0/3/5/5
3	EDO	A	1102	-	-	0/1/1/1	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1101	W2R	C18-N33	3.14	1.42	1.36
2	A	1101	W2R	C18-C12	-2.86	1.38	1.44

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1101	W2R	C10-N30-C11	-4.61	111.81	115.37
2	A	1101	W2R	C9-N32-C12	3.40	111.11	106.46
2	A	1101	W2R	C10-N31-C18	2.85	118.48	116.69
2	A	1101	W2R	O39-C27-C26	-2.67	103.36	108.76
2	A	1101	W2R	C14-N33-C18	-2.35	123.91	128.67
2	A	1101	W2R	N33-C18-N31	2.18	121.14	118.92

There are no chirality outliers.

There are no torsion outliers.

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, 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	297/331 (89%)	0.30	23 (7%) 13 12	29, 47, 92, 170	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1004	GLU	5.8
1	A	1006	ASP	5.5
1	A	735	GLY	4.7
1	A	1007	MET	4.3
1	A	1005	GLU	4.2
1	A	722	ALA	3.4
1	A	1016	TYR	3.2
1	A	1008	ASP	3.1
1	A	1003	ASP	3.0
1	A	983	GLY	3.0
1	A	982	GLN	2.9
1	A	807	ASP	2.8
1	A	734	GLU	2.8
1	A	1015	GLU	2.8
1	A	985	GLU	2.7
1	A	876	VAL	2.7
1	A	1009	ASP	2.4
1	A	861	LEU	2.3
1	A	699	PRO	2.1
1	A	984	ASP	2.1
1	A	1014	ASP	2.1
1	A	889	ARG	2.0
1	A	947	MET	2.0

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, 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	W2R	A	1101	40/40	0.13	0.22	29,36,46,53	0
3	EDO	A	1102	4/4	0.10	-0.37	36,42,44,45	0

6.5 Other polymers

There are no such residues in this entry.