



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 03:09 pm BST

PDB ID : 4HRM  
Title : Structural Basis for Eliciting a Cytotoxic Effect in HER2-Overexpressing Cancer Cells via Binding to the Extracellular Domain of HER2  
Authors : Jost, C.; Schilling, J.; Plueckthun, A.  
Deposited on : 2012-10-28  
Resolution : 3.20 Å(reported)

This is a wwPDB X-ray Structure 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

<https://www.wwpdb.org/validation/2017/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  
Xtrriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

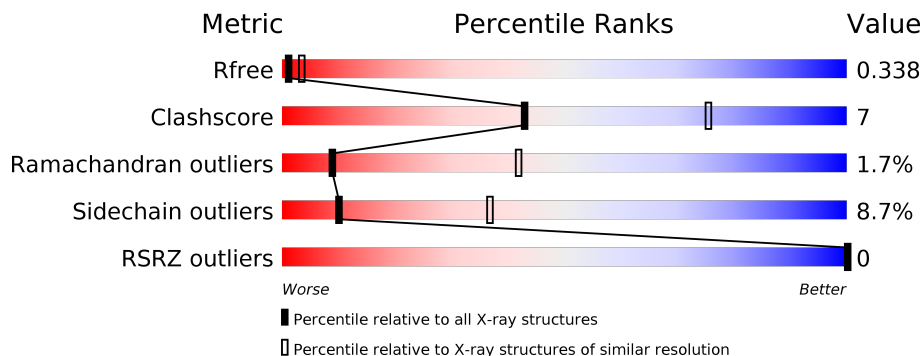
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 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  $\leq 5\%$ . 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	198	
1	C	198	
2	B	172	
2	D	172	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3517 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 Domain I of receptor tyrosine-protein kinase erbB-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	166	1123	708	192	215	8	0	0	0
1	A	139	905	570	158	173	4	0	0	0

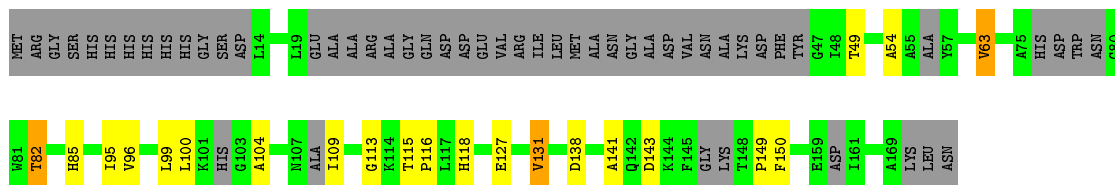
There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	HIS	-	EXPRESSION TAG	UNP P04626
C	46	ASP	ASN	ENGINEERED MUTATION	UNP P04626
C	102	ASP	ASN	ENGINEERED MUTATION	UNP P04626
C	165	ASP	ASN	ENGINEERED MUTATION	UNP P04626
C	198	ALA	-	EXPRESSION TAG	UNP P04626
A	1	HIS	-	EXPRESSION TAG	UNP P04626
A	46	ASP	ASN	ENGINEERED MUTATION	UNP P04626
A	102	ASP	ASN	ENGINEERED MUTATION	UNP P04626
A	165	ASP	ASN	ENGINEERED MUTATION	UNP P04626
A	198	ALA	-	EXPRESSION TAG	UNP P04626

- Molecule 2 is a protein called Designed Ankyrin Repeat Protein 9\_26.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	123	758	468	142	148	0	0	0
2	D	119	731	451	137	143	0	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	138.50Å 60.70Å 107.20Å 90.00° 118.90° 90.00°	Depositor
Resolution (Å)	43.12 – 3.20 43.12 – 3.20	Depositor EDS
% Data completeness (in resolution range)	97.5 (43.12-3.20) 97.6 (43.12-3.20)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.06 (at 3.19Å)	Xtrriage
Refinement program	PHENIX 1.8_1069	Depositor
R, $R_{free}$	0.314 , 0.339 0.315 , 0.338	Depositor DCC
$R_{free}$ test set	623 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.6	Xtrriage
Anisotropy	0.611	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 112.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.79	EDS
Total number of atoms	3517	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.08% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical 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 [\(i\)](#)

### 5.1 Standard geometry [\(i\)](#)

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	0.27	0/907	0.44	0/1233
1	C	0.27	0/1133	0.50	1/1549 (0.1%)
2	B	0.26	0/761	0.39	0/1043
2	D	0.30	0/735	0.40	0/1006
All	All	0.27	0/3536	0.44	1/4831 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	17	PRO	N-CA-CB	6.12	110.64	103.30

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	905	0	713	12	0
1	C	1123	0	959	16	0
2	B	758	0	602	6	0
2	D	731	0	574	10	0
All	All	3517	0	2848	43	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 7.

The worst 5 of 43 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)
2:B:49:THR:OG1	2:B:52:HIS:ND1	2.28	0.66
1:A:87:GLU:O	1:A:89:ASN:ND2	2.31	0.64
1:C:85:LEU:HD23	1:C:89:ASN:HA	1.82	0.62
1:A:14:PRO:O	1:A:16:SER:N	2.34	0.61
1:A:92:LEU:HB3	1:A:131:VAL:HG12	1.83	0.61

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	114/198 (58%)	94 (82%)	17 (15%)	3 (3%)	5	31
1	C	153/198 (77%)	131 (86%)	18 (12%)	4 (3%)	5	31
2	B	107/172 (62%)	103 (96%)	4 (4%)	0	100	100
2	D	103/172 (60%)	94 (91%)	8 (8%)	1 (1%)	15	54
All	All	477/740 (64%)	422 (88%)	47 (10%)	8 (2%)	9	42

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	17	PRO
1	C	119	GLN
1	A	120	LEU
1	C	73	PRO
1	C	84	GLN



### 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	68/176 (39%)	60 (88%)	8 (12%)	5	23
1	C	99/176 (56%)	95 (96%)	4 (4%)	31	66
2	B	50/131 (38%)	45 (90%)	5 (10%)	7	30
2	D	48/131 (37%)	42 (88%)	6 (12%)	4	21
All	All	265/614 (43%)	242 (91%)	23 (9%)	10	37

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	124	THR
2	B	19	LEU
2	D	131	VAL
1	A	140	CYS
2	B	63	VAL

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](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 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, 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	139/198 (70%)	-0.20	0 100 100	42, 56, 71, 92	0
1	C	166/198 (83%)	-0.28	0 100 100	21, 42, 53, 63	0
2	B	123/172 (71%)	-0.22	0 100 100	24, 44, 68, 76	0
2	D	119/172 (69%)	-0.23	0 100 100	36, 49, 64, 81	0
All	All	547/740 (73%)	-0.24	0 100 100	21, 49, 67, 92	0

There are no RSRZ outliers to report.

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.