



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:03 PM GMT

PDB ID : 1N8Z
Title : Crystal structure of extracellular domain of human HER2 complexed with Herceptin Fab
Authors : Cho, H.-S.; Mason, K.; Ramyar, K.X.; Stanley, A.M.; Gabelli, S.B.; Denney Jr., D.W.; Leahy, D.J.
Deposited on : 2002-11-21
Resolution : 2.52 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

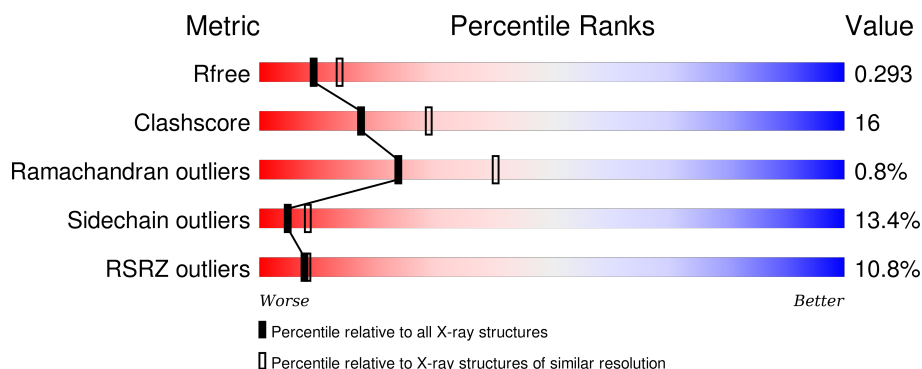
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.52 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	4241 (2.54-2.50)
Clashscore	102246	4968 (2.54-2.50)
Ramachandran outliers	100387	4873 (2.54-2.50)
Sidechain outliers	100360	4875 (2.54-2.50)
RSRZ outliers	91569	4253 (2.54-2.50)

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	214	<div> <div>11%</div> <div>63%</div> <div>30%</div> <div>7%</div> </div>
2	B	220	<div> <div>8%</div> <div>70%</div> <div>25%</div> <div>•</div> </div>
3	C	607	<div> <div>12%</div> <div>61%</div> <div>29%</div> <div>5%</div> <div>•</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NAG	C	766	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7890 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 Herceptin Fab (antibody) - light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	214	Total	C	N	O	S	0	0	0
			1645	1029	276	334	6			

- Molecule 2 is a protein called Herceptin Fab (antibody) - heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	220	Total	C	N	O	S	0	0	0
			1642	1036	276	324	6			

- Molecule 3 is a protein called Receptor protein-tyrosine kinase erbB-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	581	Total	C	N	O	S	0	0	0
			4491	2796	805	838	52			

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		

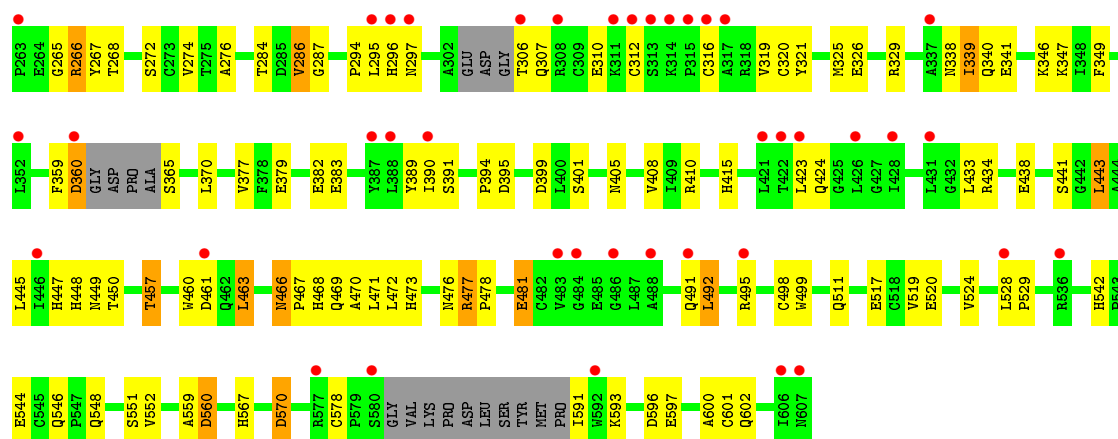
- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	19	Total 19	O 19	0	0
6	B	13	Total 13	O 13	0	0
6	C	47	Total 47	O 47	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	66.17Å 109.77Å 175.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.52 27.84 – 2.52	Depositor EDS
% Data completeness (in resolution range)	92.9 (20.00-2.52) 89.9 (27.84-2.52)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.26 (at 2.54Å)	Xtriage
Refinement program	REFMAC 5.1.19	Depositor
R, R_{free}	0.222 , 0.284 0.249 , 0.293	Depositor DCC
R_{free} test set	1886 reflections (4.74%)	DCC
Wilson B-factor (Å ²)	43.9	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 26.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 42444 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	7890	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NAG, 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	0.75	0/1682	0.91	3/2288 (0.1%)
2	B	0.72	0/1684	0.91	4/2298 (0.2%)
3	C	0.74	0/4596	0.95	13/6249 (0.2%)
All	All	0.74	0/7962	0.93	20/10835 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	C	0	1

There are no bond length outliers.

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	215	ASP	CB-CG-OD2	7.89	125.40	118.30
3	C	560	ASP	CB-CG-OD2	7.29	124.86	118.30
3	C	22	ASP	CB-CG-OD2	6.78	124.40	118.30
1	A	151	ASP	CB-CG-OD2	6.36	124.03	118.30
3	C	477	ARG	NE-CZ-NH2	-6.30	117.15	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	180	SER	Peptide

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	1645	0	1589	50	0
2	B	1642	0	1584	45	0
3	C	4491	0	4304	156	0
4	C	28	0	26	3	0
5	C	5	0	0	0	0
6	A	19	0	0	1	0
6	B	13	0	0	0	0
6	C	47	0	0	3	0
All	All	7890	0	7503	246	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 16.

The worst 5 of 246 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)
3:C:457:THR:HG22	3:C:498:CYS:O	1.49	1.10
3:C:266:ARG:HH11	3:C:266:ARG:HG2	1.20	1.03
3:C:124:THR:HG22	3:C:220:ALA:HB1	1.41	1.02
2:B:207:HIS:CD2	2:B:210:SER:HB2	1.98	0.97
3:C:457:THR:CG2	3:C:498:CYS:O	2.15	0.94

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	212/214 (99%)	190 (90%)	17 (8%)	5 (2%)	7	11
2	B	218/220 (99%)	210 (96%)	8 (4%)	0	100	100
3	C	571/607 (94%)	518 (91%)	50 (9%)	3 (0%)	34	55
All	All	1001/1041 (96%)	918 (92%)	75 (8%)	8 (1%)	24	40

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	151	ASP
3	C	167	SER
1	A	32	ALA
1	A	152	ASN
3	C	51	PHE

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	188/189 (100%)	157 (84%)	31 (16%)	3	4
2	B	179/181 (99%)	156 (87%)	23 (13%)	5	9
3	C	503/523 (96%)	440 (88%)	63 (12%)	6	10
All	All	870/893 (97%)	753 (87%)	117 (13%)	5	8

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	210	SER
3	C	83	THR
3	C	466	ASN
3	C	2	GLN
3	C	27	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	227	HIS
3	C	253	ASN
3	C	542	HIS
3	C	238	HIS
2	B	204	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](#)

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](#)

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	$\# Z > 2$
5	SO4	C	1001	-	4,4,4	0.31	0	6,6,6	0.68	0
4	NAG	C	738	3	14,14,15	0.48	0	15,19,21	1.39	3 (20%)
4	NAG	C	766	3	14,14,15	0.80	0	15,19,21	2.08	6 (40%)

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
5	SO4	C	1001	-	-	0/0/0/0	0/0/0/0
4	NAG	C	738	3	-	0/6/23/26	0/1/1/1
4	NAG	C	766	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	766	NAG	C2-N2-C7	-3.31	118.78	123.04
4	C	766	NAG	C1-O5-C5	-3.07	108.35	112.25
4	C	766	NAG	C3-C4-C5	-2.23	106.30	110.20
4	C	738	NAG	O4-C4-C3	-2.03	105.76	110.34
4	C	738	NAG	C4-C3-C2	2.38	114.92	111.23

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	766	NAG	3	0

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, 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	214/214 (100%)	0.75	23 (10%) 8 8	7, 16, 24, 33	0
2	B	220/220 (100%)	0.50	17 (7%) 16 18	8, 16, 24, 35	0
3	C	581/607 (95%)	0.68	70 (12%) 6 6	2, 16, 25, 38	0
All	All	1015/1041 (97%)	0.66	110 (10%) 8 8	2, 16, 24, 38	0

The worst 5 of 110 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	214	CYS	7.7
1	A	212	GLY	6.4
3	C	201	GLY	5.7
3	C	101	LEU	5.6
3	C	315	PRO	5.6

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(\AA^2)	Q<0.9
4	NAG	C	766	14/15	0.78	0.24	2.78	48,55,58,59	0
5	SO4	C	1001	5/5	0.94	0.21	0.72	59,60,61,62	0
4	NAG	C	738	14/15	0.89	0.15	0.11	33,43,46,47	0

6.5 Other polymers [i](#)

There are no such residues in this entry.