



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2017 – 04:12 PM EST

PDB ID : 1KGY  
Title : Crystal Structure of the EphB2-ephrinB2 complex  
Authors : Himanen, J.P.; Rajashankar, K.R.; Lackmann, M.; Cowan, C.A.; Henkemeyer, M.; Nikolov, D.B.  
Deposited on : 2001-11-28  
Resolution : 2.70 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

<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	:	unknown
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

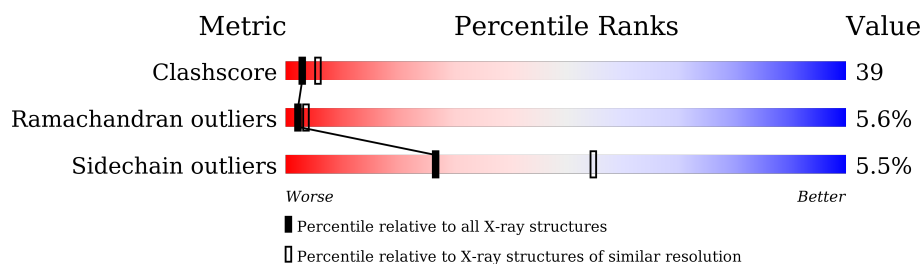
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	181	
1	B	181	
1	C	181	
1	D	181	
2	E	138	
2	F	138	
2	G	138	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	138	<div><div></div><div>43%</div><div>51%</div><div>6%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10260 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 EPHRIN TYPE-B RECEPTOR 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	181	Total	C	N	O	S	0	0	0
			1463	925	252	274	12			
1	B	181	Total	C	N	O	S	0	0	0
			1463	925	252	274	12			
1	C	181	Total	C	N	O	S	0	0	0
			1463	925	252	274	12			
1	D	181	Total	C	N	O	S	0	0	0
			1463	925	252	274	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	ALA	VAL	CONFLICT	UNP P54763
B	227	ALA	VAL	CONFLICT	UNP P54763
C	427	ALA	VAL	CONFLICT	UNP P54763
D	627	ALA	VAL	CONFLICT	UNP P54763

- Molecule 2 is a protein called EPHRIN-B2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	138	Total	C	N	O	S	0	0	0
			1102	706	179	210	7			
2	F	138	Total	C	N	O	S	0	0	0
			1102	706	179	210	7			
2	G	138	Total	C	N	O	S	0	0	0
			1102	706	179	210	7			
2	H	138	Total	C	N	O	S	0	0	0
			1102	706	179	210	7			

There are 4 discrepancies between the modelled and reference sequences:

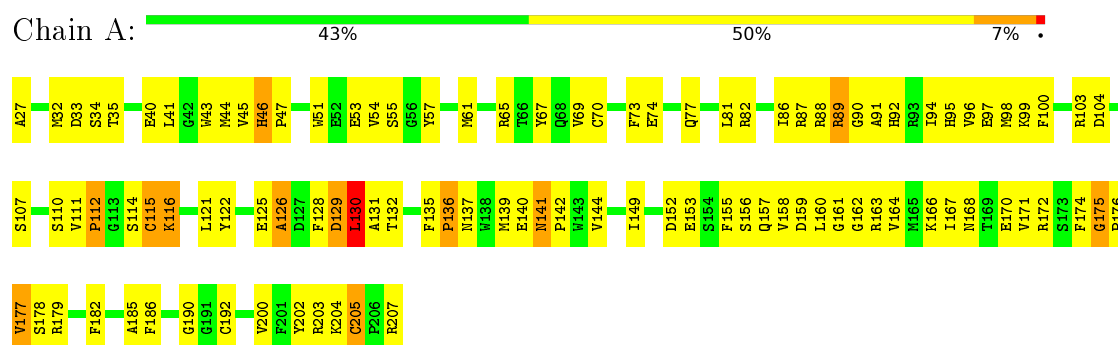
Chain	Residue	Modelled	Actual	Comment	Reference
E	1049	GLY	GLN	CONFLICT	UNP P52800
F	1249	GLY	GLN	CONFLICT	UNP P52800
G	1449	GLY	GLN	CONFLICT	UNP P52800
H	1649	GLY	GLN	CONFLICT	UNP P52800

### 3 Residue-property plots

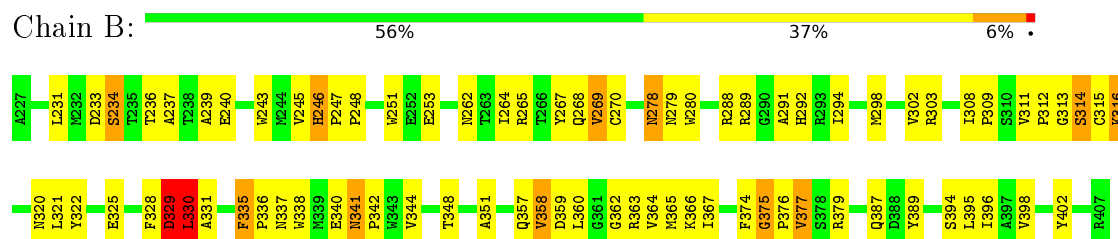
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

Note EDS was not executed.

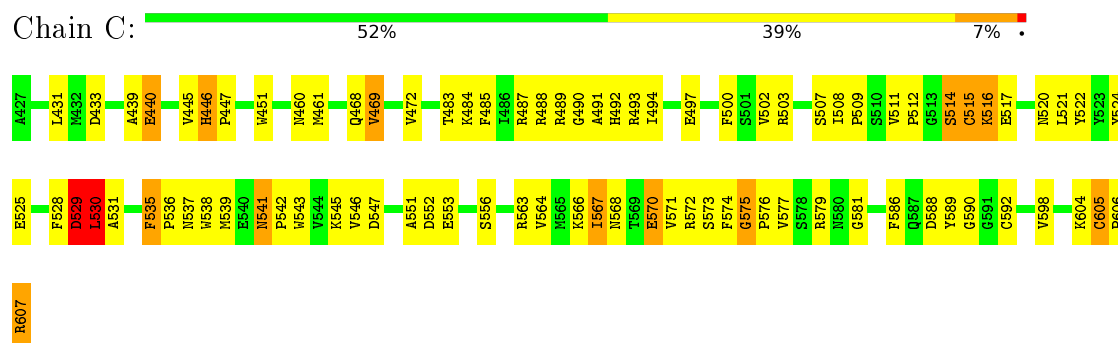
#### • Molecule 1: EPHRIN TYPE-B RECEPTOR 2



#### • Molecule 1: EPHRIN TYPE-B RECEPTOR 2

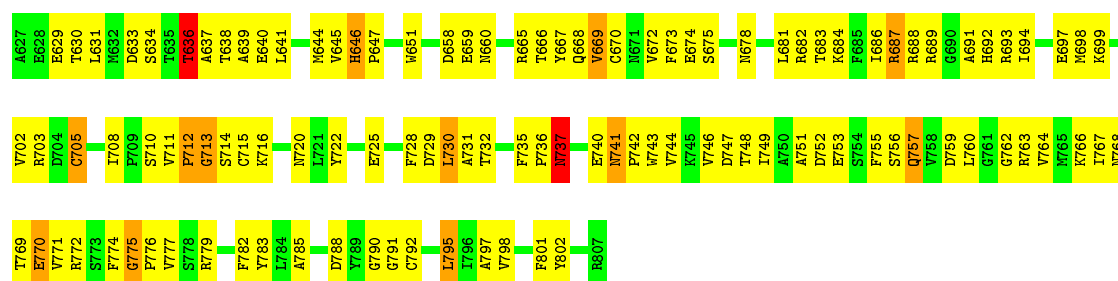


#### • Molecule 1: EPHRIN TYPE-B RECEPTOR 2



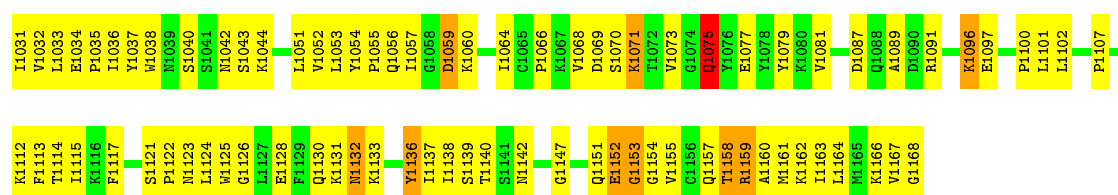
#### • Molecule 1: EPHRIN TYPE-B RECEPTOR 2

Chain D: 



• Molecule 2: EPHRIN-B2

Chain E: 



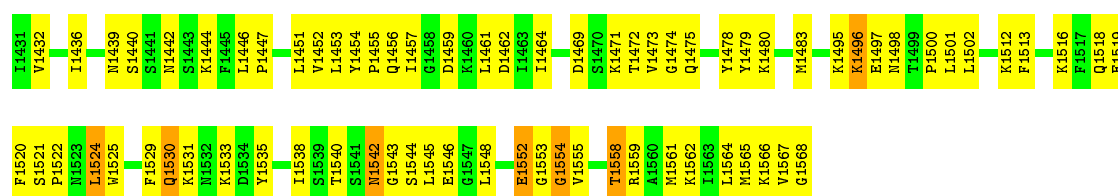
• Molecule 2: EPHRIN-B2

Chain F: 



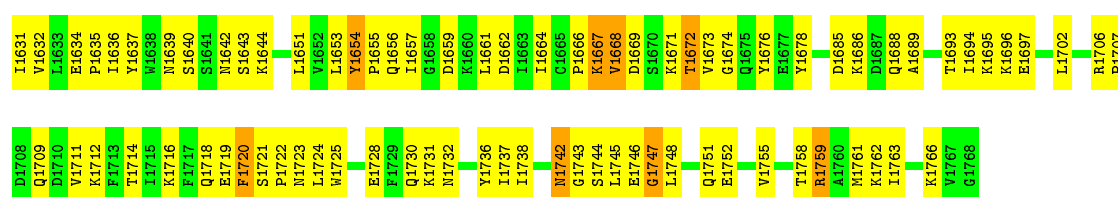
• Molecule 2: EPHRIN-B2

Chain G: 



• Molecule 2: EPHRIN-B2

Chain H: 



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.00 Å   78.00 Å   78.00 Å 69.00°   75.00°   69.00°	Depositor
Resolution (Å)	25.00 – 2.70	Depositor
% Data completeness (in resolution range)	(Not available) (25.00-2.70)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.243 , 0.296	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10260	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP



## 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	0.43	0/1502	0.74	0/2035
1	B	0.46	0/1502	0.72	0/2035
1	C	0.44	0/1502	0.74	0/2035
1	D	0.41	0/1502	0.69	0/2035
2	E	0.42	0/1125	0.75	0/1519
2	F	0.42	0/1125	0.74	0/1519
2	G	0.47	0/1125	0.79	0/1519
2	H	0.43	0/1125	0.75	0/1519
All	All	0.43	0/10508	0.74	0/14216

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 [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	1463	0	1389	206	0
1	B	1463	0	1389	93	0
1	C	1463	0	1389	87	0
1	D	1463	0	1389	112	0
2	E	1102	0	1096	137	0
2	F	1102	0	1096	67	0
2	G	1102	0	1096	78	0
2	H	1102	0	1096	154	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	10260	0	9940	790	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 39.

The worst 5 of 790 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:128:PHE:CZ	2:H:1639:ASN:HB3	1.84	1.12
1:A:128:PHE:HB3	2:H:1637:TYR:CD2	1.86	1.10
1:A:89:ARG:HA	2:H:1671:LYS:CG	1.83	1.09
1:A:89:ARG:N	2:H:1671:LYS:NZ	2.02	1.08
1:A:135:PHE:CE2	2:H:1634:GLU:OE2	2.08	1.06

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	179/181 (99%)	144 (80%)	22 (12%)	13 (7%)	1	1
1	B	179/181 (99%)	135 (75%)	32 (18%)	12 (7%)	1	2
1	C	179/181 (99%)	144 (80%)	24 (13%)	11 (6%)	2	3
1	D	179/181 (99%)	138 (77%)	31 (17%)	10 (6%)	2	3
2	E	136/138 (99%)	112 (82%)	18 (13%)	6 (4%)	3	6
2	F	136/138 (99%)	110 (81%)	23 (17%)	3 (2%)	8	22
2	G	136/138 (99%)	112 (82%)	17 (12%)	7 (5%)	2	4
2	H	136/138 (99%)	110 (81%)	18 (13%)	8 (6%)	2	3
All	All	1260/1276 (99%)	1005 (80%)	185 (15%)	70 (6%)	2	3

5 of 70 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	34	SER
1	A	130	LEU
1	A	175	GLY
1	B	309	PRO
1	B	375	GLY

### 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	160/160 (100%)	155 (97%)	5 (3%)	47	78
1	B	160/160 (100%)	153 (96%)	7 (4%)	35	65
1	C	160/160 (100%)	152 (95%)	8 (5%)	30	60
1	D	160/160 (100%)	146 (91%)	14 (9%)	12	28
2	E	124/124 (100%)	118 (95%)	6 (5%)	31	62
2	F	124/124 (100%)	115 (93%)	9 (7%)	17	39
2	G	124/124 (100%)	115 (93%)	9 (7%)	17	39
2	H	124/124 (100%)	120 (97%)	4 (3%)	46	77
All	All	1136/1136 (100%)	1074 (94%)	62 (6%)	27	55

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	741	ASN
2	E	1059	ASP
2	G	1564	LEU
1	D	770	GLU
2	E	1075	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	741	ASN
2	E	1109	GLN
2	H	1703	ASN
1	D	768	ASN
2	E	1118	GLN

### 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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.