



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 04:08 AM GMT

PDB ID : 4LXM  
Title : Crystal Structure of Human Beta Secretase in Complex with compound 12a  
Authors : Rondeau, J.M.; Bourgier, E.  
Deposited on : 2013-07-30  
Resolution : 2.30 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>

---

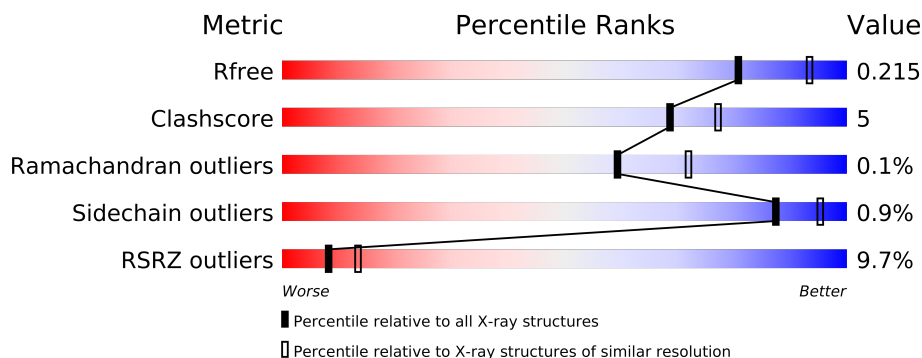
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.30 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	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (2.30-2.30)

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	402	
1	B	402	
1	C	402	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9581 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.

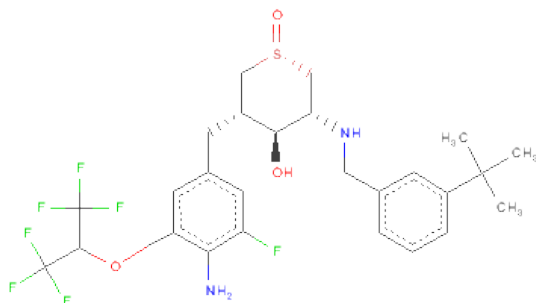
- Molecule 1 is a protein called Beta-secretase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	377	Total	C	N	O	S	0	0	0
			2966	1898	493	561	14			
1	B	377	Total	C	N	O	S	0	0	0
			2966	1898	493	561	14			
1	C	381	Total	C	N	O	S	0	1	0
			3001	1922	500	565	14			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33P	GLY	-	EXPRESSION TAG	UNP P56817
A	34P	PRO	-	EXPRESSION TAG	UNP P56817
B	33P	GLY	-	EXPRESSION TAG	UNP P56817
B	34P	PRO	-	EXPRESSION TAG	UNP P56817
C	33P	GLY	-	EXPRESSION TAG	UNP P56817
C	34P	PRO	-	EXPRESSION TAG	UNP P56817

- Molecule 2 is (1S,3S,4S,5R)-3-{4-AMINO-3-FLUORO-5-[(1,1,1,3,3,3-HEXAFLUOROPROPAN-2-YL)OXY]BENZYL}-5-[(3-TERT-BUTYLBENZYL)AMINO]TETRAHYDRO-2H-THIOPYRAN-4-OL1-OXIDE (three-letter code: 1YU) (formula: C<sub>26</sub>H<sub>31</sub>F<sub>7</sub>N<sub>2</sub>O<sub>3</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	S	0	0
			39	26	7	2	3	1		
2	B	1	Total	C	F	N	O	S	0	0
			39	26	7	2	3	1		
2	C	1	Total	C	F	N	O	S	0	0
			39	26	7	2	3	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	185	Total	O	0	0
			185	185		
3	B	160	Total	O	0	0
			160	160		
3	C	186	Total	O	0	0
			186	186		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.29Å 104.45Å 100.49Å 90.00° 104.30° 90.00°	Depositor
Resolution (Å)	16.67 – 2.30 16.67 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (16.67-2.30) 99.8 (16.67-2.30)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.69 (at 2.30Å)	Xtriage
Refinement program	CNX 2002	Depositor
R, $R_{free}$	0.194 , 0.218 0.191 , 0.215	Depositor DCC
$R_{free}$ test set	3719 reflections (5.38%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.5	Xtriage
Anisotropy	0.435	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 47.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 72839 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9581	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.03% 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: 1YU

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	0.31	0/3041	0.47	0/4133
1	B	0.31	0/3041	0.47	0/4133
1	C	0.31	0/3081	0.47	0/4188
All	All	0.31	0/9163	0.47	0/12454

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	2966	0	2875	23	0
1	B	2966	0	2875	30	0
1	C	3001	0	2912	34	0
2	A	39	0	31	0	0
2	B	39	0	31	0	0
2	C	39	0	31	0	0
3	A	185	0	0	3	0
3	B	160	0	0	4	0
3	C	186	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9581	0	8755	86	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 5.

The worst 5 of 86 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:267:LEU:HD23	1:B:267:LEU:H	1.45	0.81
1:C:267:LEU:H	1:C:267:LEU:HD23	1.46	0.80
1:A:267:LEU:HD23	1:A:267:LEU:H	1.44	0.79
1:B:192:PRO:HG2	1:B:288:MET:CE	2.22	0.69
1:C:309:VAL:HG11	1:C:321:LYS:HG3	1.74	0.68

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	373/402 (93%)	364 (98%)	9 (2%)	0	100	100
1	B	373/402 (93%)	361 (97%)	11 (3%)	1 (0%)	50	60
1	C	378/402 (94%)	369 (98%)	9 (2%)	0	100	100
All	All	1124/1206 (93%)	1094 (97%)	29 (3%)	1 (0%)	59	72

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	364	GLU

### 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	322/342 (94%)	319 (99%)	3 (1%)	87	95
1	B	322/342 (94%)	319 (99%)	3 (1%)	87	95
1	C	325/342 (95%)	322 (99%)	3 (1%)	87	95
All	All	969/1026 (94%)	960 (99%)	9 (1%)	87	95

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	197	TRP
1	C	267	LEU
1	C	125	GLU
1	A	267	LEU
1	B	267	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	278	ASN
1	B	293	ASN
1	C	114	ASN
1	B	114	ASN
1	C	92	ASN

### 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 ⓘ

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$
2	1YU	A	501	-	41,41,41	1.59	6 (14%)	63,63,63	1.34	7 (11%)
2	1YU	B	501	-	41,41,41	1.61	6 (14%)	63,63,63	1.36	9 (14%)
2	1YU	C	501	-	41,41,41	1.59	6 (14%)	63,63,63	1.36	7 (11%)

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	1YU	A	501	-	-	0/31/47/47	0/3/3/3
2	1YU	B	501	-	-	0/31/47/47	0/3/3/3
2	1YU	C	501	-	-	0/31/47/47	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	1YU	C25-C23	4.06	1.56	1.52
2	C	501	1YU	C25-C23	3.99	1.56	1.52
2	A	501	1YU	C25-C23	3.87	1.56	1.52
2	A	501	1YU	C19-C21	3.02	1.57	1.53
2	B	501	1YU	C19-C21	2.99	1.56	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	1YU	C8-C9-C10	4.57	119.55	115.94
2	A	501	1YU	C8-C9-C10	4.54	119.53	115.94
2	C	501	1YU	C8-C9-C10	4.31	119.35	115.94
2	A	501	1YU	C35-N33-C23	4.08	121.54	114.94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	1YU	C35-N33-C23	3.85	121.16	114.94

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, 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	377/402 (93%)	0.27	33 (8%) 10 16	28, 43, 82, 117	0
1	B	377/402 (93%)	0.40	39 (10%) 7 11	26, 44, 86, 119	0
1	C	381/402 (94%)	0.30	38 (9%) 8 13	27, 43, 85, 116	0
All	All	1135/1206 (94%)	0.33	110 (9%) 8 13	26, 44, 85, 119	0

The worst 5 of 110 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	314	THR	12.4
1	A	314	THR	11.3
1	B	361	VAL	11.2
1	B	312	VAL	10.0
1	C	312	VAL	9.7

### 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
2	1YU	C	501	39/39	0.13	-	33,38,51,52	0
2	1YU	A	501	39/39	0.14	-	33,38,52,53	0
2	1YU	B	501	39/39	0.12	-	33,37,51,53	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.