wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Mar 2, 2017 – 12:24 pm GMT

PDB ID : 5FN3
EMDB ID: : EMD-3238
Title : Cryo-EM structure of gamma secretase in class 1 of the apo- state ensemble
Authors : Bai, X.C.; Rajendra, E.; Yang, G.H.; Shi, Y.G.; Scheres, S.H.W.
Deposited on : 2015-11-10
Resolution : 4.10 Å(reported)

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.10 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	Percentile Ranks	Value
Clashscore		20
Ramachandran outliers		4.0%
Sidechain outliers		10.8%

Worse *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all EM structures
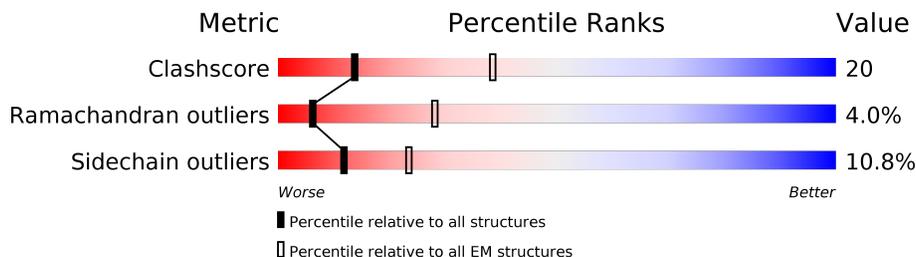
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	709	69% 21% • 6%
2	B	467	28% 25% 7% • 39%
3	C	265	62% 27% • 8%
4	D	101	78% 19% • •
5	G	24	38% 33% 25% •

2 Entity composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 10291 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 NICASTRIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	665	Total 5222	C 3312	N 888	O 1001	S 21	0	0

- Molecule 2 is a protein called PRESENILIN-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	286	Total 2234	C 1523	N 335	O 363	S 13	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	256	THR	TYR	CONFLICT	UNP P49768

- Molecule 3 is a protein called GAMMA-SECRETASE SUBUNIT APH-1A.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	243	Total 1868	C 1252	N 299	O 313	S 4	0	0

- Molecule 4 is a protein called GAMMA-SECRETASE SUBUNIT PEN-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	100	Total 847	C 579	N 133	O 134	S 1	0	0

- Molecule 5 is a protein called POLY ALA CHAIN.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	G	24	Total 120	C 72	N 24	O 24	0	0

# 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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

- Molecule 1: NICASTRIN

- Molecule 2: PRESENILIN-1

ALA GLN ARG ASP SER HIS LEU GLY PRO HIS ARG SER THR PRO GLU SER ARG ALA ALA VAL GLN GLU LEU SER SER SER ILE LEU ALA GLY GLU GLU ASP PRO GLU GLU ARG **G378** L381 G382 L383 F386 S390 V391 L392 V393 D403 W404 N405 T406 T407 I408 A409 C410 F411 V412 A413

I414 L415 I416 G417 L418 C419 L420 T421 L422 K429 K430 A434 L435 P436 I437 S438 L443 V444 F445 Y446 T449 D450 Y451 L452 V453 Q454 P455 D458 Q459 Q464 F465 Y466 I467
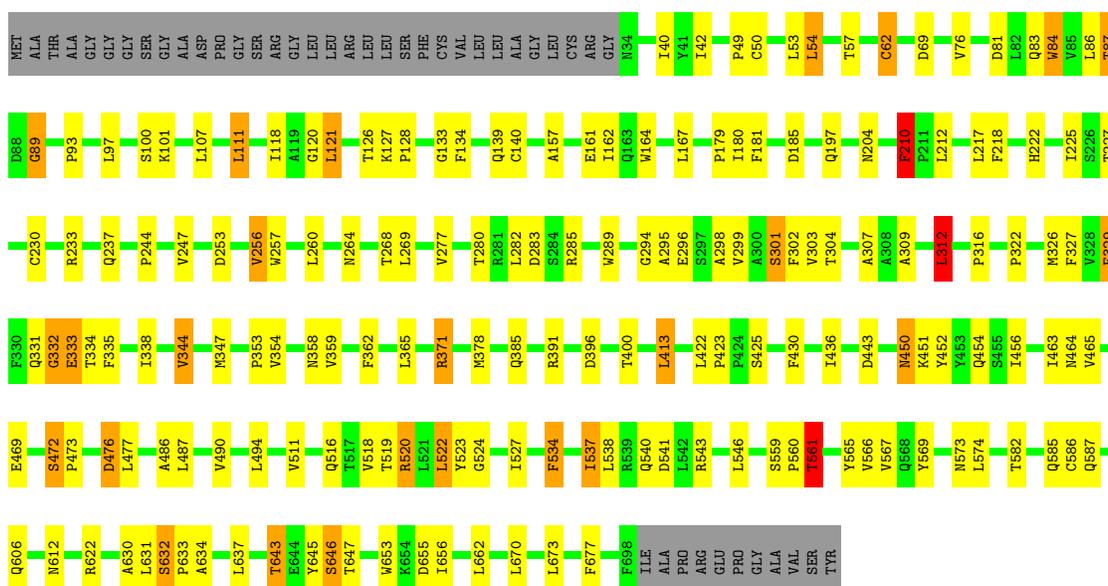
● Molecule 3: GAMMA-SECRETASE SUBUNIT APH-1A

Chain C: 62%   27%   ·   8%

MET **G2** F7 A17 T24 D28 I34 L35 V36 F40 F41 W42 L43 V44 S45 L46 L47 S50 V51 L56 D61 Q68 Y69 F74 V78 S79 V80 Y90 Y91 K92 L93 L94 L104 D107 I112 Q116 M117 A118 Y119 V120 L123 S124

I127 I128 S129 G130 V131 V134 I137 V147 G148 I149 H150 G151 D152 Y156 F157 L158 T159 S160 A161 F162 L163 I168 L169 L170 H171 T172 F173 W174 G175 V176 V177 F178 F179 C182 L190 S196 H197 L198 L199 T200 L203 T204 F205 L206 N207 Y210 L214 L215

P216 I217 Y218 G225 L236 R237 S238 I239 L244 CYS ARG ARG GLN GLU ASP SER ARG VAL MET VAL TYR SER ALA LEU ARG ILE PRO PRO GLU ASP

● Molecule 4: GAMMA-SECRETASE SUBUNIT PEN-2

Chain D: 78%   19%   · ·

MET **N2** K17 L26 P27 F28 I34 F35 R39 K54 G55 W58 A61 F66 W67 S73 W74 I75 F78 G86 F94 T95 I96 P97 L98 P101

● Molecule 5: POLY ALA CHAIN

Chain G: 38%   33%   25%   ·

A5 A8 A9 A10 A11 A12 **A13** A14 A15 A16 A17 A18 A21 A22 A23 A24 A25 A26 A27 A28
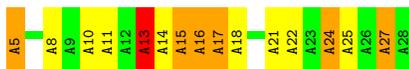
4   Experimental information ⓘ

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	63873	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	38	Depositor
Minimum defocus (nm)	700	Depositor
Maximum defocus (nm)	3200	Depositor
Magnification	35714	Depositor
Image detector	GATAN K2 QUANTUM (4K X 4K)	Depositor

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	#$|Z|$ >2	RMSZ	#$|Z|$ >2
1	A	0.52	1/5345 (0.0%)	0.81	2/7284 (0.0%)
2	B	0.66	1/2291 (0.0%)	0.98	5/3130 (0.2%)
3	C	0.65	0/1920	0.95	2/2619 (0.1%)
4	D	0.66	0/880	0.82	0/1201
5	G	0.20	0/119	0.34	0/165
All	All	0.59	2/10555 (0.0%)	0.87	9/14399 (0.1%)

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
1	A	0	5
2	B	0	15
5	G	0	8
All	All	0	28

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	257	TRP	CB-CG	-5.53	1.40	1.50
2	B	244	TRP	CB-CG	5.17	1.59	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	117	PRO	CA-N-CD	-7.03	101.66	111.50
3	C	47	LEU	CA-CB-CG	6.85	131.04	115.30
3	C	69	TYR	CA-CB-CG	6.14	125.07	113.40
2	B	116	THR	C-N-CD	5.95	140.89	128.40
1	A	312	LEU	CA-CB-CG	5.46	127.87	115.30

There are no chirality outliers.

5 of 28 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	256	VAL	Peptide
1	A	335	PHE	Peptide
1	A	57	THR	Peptide
1	A	62	CYS	Peptide
1	A	622	ARG	Peptide

5.2   Too-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 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	5222	0	5120	91	0
2	B	2234	0	2344	261	0
3	C	1868	0	1907	54	0
4	D	847	0	836	11	0
5	G	120	0	119	59	0
All	All	10291	0	10326	412	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 20.

The worst 5 of 412 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)
5:G:15:ALA:C	5:G:17:ALA:HB3	1.31	1.42
2:B:112:GLN:CD	5:G:15:ALA:HB2	1.61	1.20
5:G:15:ALA:O	5:G:17:ALA:CB	1.89	1.19
2:B:146:MET:CE	5:G:25:ALA:H	1.56	1.18
2:B:146:MET:HE1	5:G:25:ALA:N	1.58	1.17

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 PDB entries followed by that with respect to all EM entries.

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	663/709 (94%)	534 (80%)	99 (15%)	30 (4%)	3	30
2	B	280/467 (60%)	248 (89%)	16 (6%)	16 (6%)	2	25
3	C	241/265 (91%)	221 (92%)	16 (7%)	4 (2%)	11	51
4	D	98/101 (97%)	89 (91%)	8 (8%)	1 (1%)	18	60
5	G	22/24 (92%)	20 (91%)	1 (4%)	1 (4%)	3	30
All	All	1304/1566 (83%)	1112 (85%)	140 (11%)	52 (4%)	6	33

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	157	ALA
1	A	333	GLU
1	A	473	PRO
1	A	643	THR
2	B	116	THR

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 PDB entries followed by that with respect to all EM entries.

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	583/612 (95%)	530 (91%)	53 (9%)	11	43
2	B	240/408 (59%)	206 (86%)	34 (14%)	4	26
3	C	192/214 (90%)	170 (88%)	22 (12%)	6	33
4	D	87/89 (98%)	77 (88%)	10 (12%)	6	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1102/1323 (83%)	983 (89%)	119 (11%)	12 36

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	80	LYS
2	B	184	GLU
4	D	17	LYS
2	B	87	VAL
2	B	133	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	454	GLN
1	A	606	GLN
2	B	464	GLN
1	A	450	ASN
2	B	454	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

There are no ligands in this entry.

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.