

wwPDB X-ray Structure Validation Summary Report (i)

Feb 15, 2024 – 12:44 AM EST

PDB ID : 3NYL

Title : The X-ray structure of an antiparallel dimer of the human amyloid precursor

protein E2 domain

Authors: Ha, Y.; Hu, J.; Lee, S.; Liu, X.; Wang, Y.

Deposited on : 2010-07-15

Resolution : 2.80 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

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity : 4.02b-467 Xtriage (Phenix) : 1.13

henix) : 1.13 EDS : 2.36

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

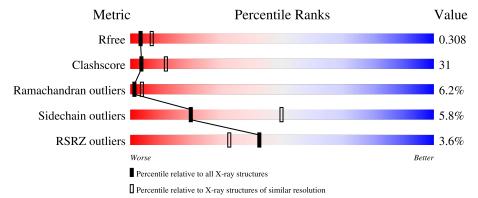
Validation Pipeline (wwPDB-VP) : 2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X- $RAY\ DIFFRACTION$

The reported resolution of this entry is 2.80 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	Similar resolution
Metric	$(\# ext{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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.

3%		Quali	Length	Chain	Mol
1 A 210 53% 34%	6% 7%		910	٨	1



2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1473 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 Amyloid beta (A4) protein (Peptidase nexin-II, Alzheimer disease), isoform CRA b.

\mathbf{Mol}	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
1	A	196	Total 1435	C 903	N 262	O 265	S 5	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	342	GLY	-	expression tag	UNP D3DSD2
A	343	SER	-	expression tag	UNP D3DSD2
A	344	HIS	-	expression tag	UNP D3DSD2
A	345	MET	-	expression tag	UNP D3DSD2

• Molecule 2 is water.

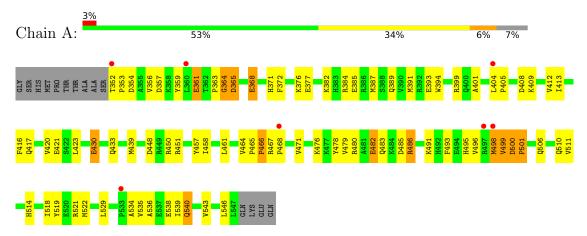
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	38	Total O 38 38	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 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.

• Molecule 1: Amyloid beta (A4) protein (Peptidase nexin-II, Alzheimer disease), isoform CRA b





4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants	83.59Å 83.59Å 147.34Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 - 2.80	Depositor
rtesolution (A)	29.75 - 2.80	EDS
% Data completeness	(Not available) (40.00-2.80)	Depositor
(in resolution range)	90.8 (29.75-2.80)	EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) > 1$	5.02 (at 2.80Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
D D.	0.258 , 0.322	Depositor
R, R_{free}	0.267 , 0.308	DCC
R_{free} test set	666 reflections (10.85%)	wwPDB-VP
Wilson B-factor (Å ²)	87.4	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.32, 85.2	EDS
L-test for twinning ²	$ < L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	1473	wwPDB-VP
Average B, all atoms $(Å^2)$	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 7.83% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical values of <|L|>, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

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).

Mal	Chain	Boı	nd lengths	Bond	angles
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	A	0.60	2/1461~(0.1%)	0.67	0/1989

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	Observed(A)	$\operatorname{Ideal}(ext{\AA})$
1	A	538	GLU	CB-CG	-5.24	1.42	1.52
1	A	377	GLU	CB-CG	-5.14	1.42	1.52

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	1435	0	1269	85	0
2	A	38	0	0	1	0
All	All	1473	0	1269	85	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 31.

The worst 5 of 85 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.



Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)
1:A:356:VAL:HG21	1:A:430:GLU:HG2	1.43	0.97
1:A:464:VAL:O	1:A:466:PRO:HD3	1.68	0.93
1:A:500:ASP:CB	1:A:501:PRO:HD3	2.06	0.86
1:A:466:PRO:O	1:A:468:PRO:HD3	1.78	0.84
1:A:372:PHE:CZ	1:A:376:LYS:HD2	2.16	0.80

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.

\mathbf{Mol}	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	194/210 (92%)	159 (82%)	23 (12%)	12 (6%)	1 4

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	365	ASP
1	A	368	GLU
1	A	498	MET
1	A	499	VAL
1	A	500	ASP

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.

Continued on next page...



Continued from previous page...

	•				
Mol	Chain	Analysed	Rotameric	Outliers	Percentiles

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	120/186 (64%)	113 (94%)	7 (6%)	20 50

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	482	GLU
1	A	486	ARG
1	A	546	LEU
1	A	491	LYS
1	A	458	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	371	HIS
1	A	483	GLN
1	A	495	HIS
1	A	514	HIS

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 monosaccharides 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^{th} 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	196/210 (93%)	0.08	7 (3%) 42 32	25, 53, 97, 189	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	498	MET	8.2
1	A	352	THR	7.0
1	A	533	PRO	3.4
1	A	497	ARG	3.3
1	A	468	PRO	2.3

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 monosaccharides 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.

