



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 25, 2023 – 04:18 pm BST

PDB ID : 8AX8  
Title : Human Apolipoprotein E4 (ApoE4) N-terminal domain (space group P3121)  
Authors : Nemergut, M.; Marek, M.  
Deposited on : 2022-08-31  
Resolution : 1.55 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.34  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 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) : 2.34

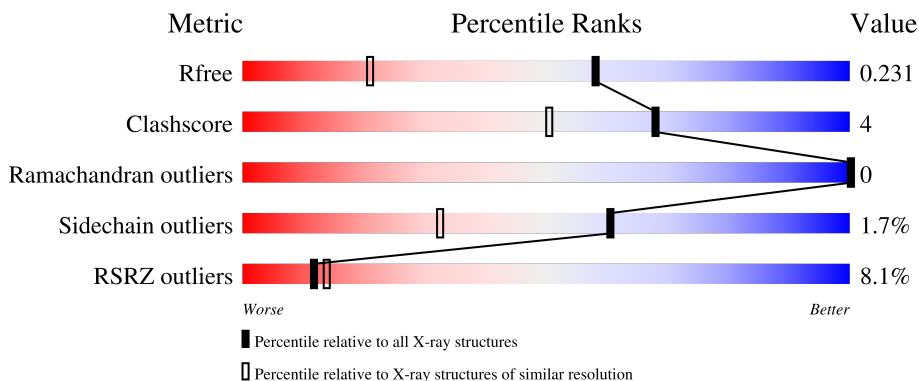
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.55 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))
$R_{free}$	130704	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	675	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1260 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 Maltose/maltodextrin-binding periplasmic protein, Apolipoprotein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	135	1155	715	221	214	5	0	4	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-375	GLY	ALA	engineered mutation	UNP P0AEX9
A	-374	PRO	LEU	engineered mutation	UNP P0AEX9
A	-373	MET	ALA	engineered mutation	UNP P0AEX9
A	-291	ALA	ASP	engineered mutation	UNP P0AEX9
A	-290	ALA	LYS	engineered mutation	UNP P0AEX9
A	-134	ALA	LYS	engineered mutation	UNP P0AEX9
A	-14	ALA	GLU	engineered mutation	UNP P0AEX9
A	-11	ALA	LYS	engineered mutation	UNP P0AEX9
A	-10	ALA	ASP	engineered mutation	UNP P0AEX9
A	-6	ASN	-	linker	UNP P0AEX9
A	-5	ALA	-	linker	UNP P0AEX9
A	-4	ALA	-	linker	UNP P0AEX9
A	-3	ALA	-	linker	UNP P0AEX9
A	-2	ALA	-	linker	UNP P0AEX9
A	-1	HIS	-	linker	UNP P0AEX9
A	0	MET	-	linker	UNP P0AEX9
A	112	ARG	CYS	engineered mutation	UNP P02649
A	257	ALA	PHE	engineered mutation	UNP P02649
A	264	ARG	TRP	engineered mutation	UNP P02649
A	269	ALA	VAL	engineered mutation	UNP P02649
A	279	GLN	LEU	engineered mutation	UNP P02649
A	287	GLU	VAL	engineered mutation	UNP P02649

- Molecule 2 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
2	A	104	Total 105	O 105	0	1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.69Å 47.69Å 104.77Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.42 – 1.55 38.43 – 1.55	Depositor EDS
% Data completeness (in resolution range)	99.9 (38.42-1.55) 99.9 (38.43-1.55)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 1.55Å)	Xtrriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, $R_{free}$	0.198 , 0.231 0.198 , 0.231	Depositor DCC
$R_{free}$ test set	975 reflections (4.72%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.3	Xtrriage
Anisotropy	0.378	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.046 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	1260	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.19% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

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

## 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	0.62	0/1172	0.83	4/1567 (0.3%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	134	ARG	NE-CZ-NH2	-5.49	117.55	120.30
1	A	126	LEU	CB-CG-CD1	-5.23	102.11	111.00
1	A	125[A]	MET	CB-CG-SD	-5.12	97.04	112.40
1	A	125[B]	MET	CB-CG-SD	-5.12	97.04	112.40

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	1155	0	1176	9	0
2	A	105	0	0	2	0
All	All	1260	0	1176	9	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 4.

All (9) 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:92:ARG:HB2	1:A:92:ARG:CZ	2.26	0.66
1:A:92:ARG:HB2	1:A:92:ARG:NH2	2.10	0.66
1:A:68:MET:O	1:A:72:LYS:HG3	1.98	0.62
1:A:114[A]:ARG:CZ	1:A:117:GLN:OE1	2.50	0.60
1:A:32:ARG:NH1	2:A:301:HOH:O	2.23	0.57
1:A:114[A]:ARG:NH1	1:A:117:GLN:OE1	2.42	0.53
1:A:83:THR:OG1	1:A:84:PRO:HD2	2.15	0.46
1:A:131:GLU:OE2	1:A:134:ARG:NH1	2.40	0.44
1:A:24:GLN:N	2:A:308:HOH:O	2.50	0.44

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	135/675 (20%)	133 (98%)	2 (2%)	0	100 100

There are no Ramachandran outliers to report.

### 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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	122/537 (23%)	120 (98%)	2 (2%)	62 35

All (2) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	64	MET
1	A	90	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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<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	135/675 (20%)	0.38	11 (8%) <b>12</b> <b>14</b>	17, 26, 59, 104	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	162	TYR	4.8
1	A	82	LEU	4.7
1	A	83	THR	4.6
1	A	81	GLN	4.4
1	A	93	LEU	4.1
1	A	91	ALA	4.0
1	A	84	PRO	3.9
1	A	92	ARG	3.6
1	A	161	VAL	2.8
1	A	163	GLN	2.7
1	A	90	ARG	2.7

### 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.