



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 04:39 pm BST

PDB ID : 4RTD
Title : Escherichia coli alpha-2-macroglobulin activated by porcine elastase
Authors : Fyfe, C.D.; Grinter, R.; Roszak, A.W.; Josts, I.; Cogdell, R.J.; Walker, D.
Deposited on : 2014-11-14
Resolution : 3.65 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

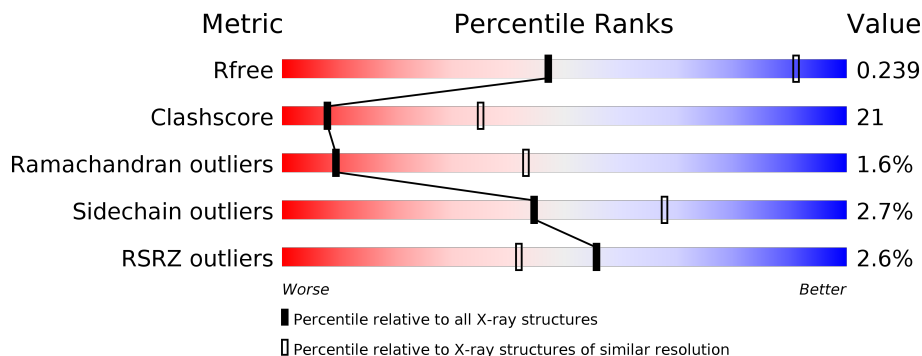
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.65 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	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)

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 ≥ 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	1639	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8699 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 Uncharacterized lipoprotein YfhM.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	1122	8699	5497	1505	1677	1	19	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	219	PRO	SER	ENGINEERED MUTATION	UNP P76578
A	606	ARG	GLN	ENGINEERED MUTATION	UNP P76578
A	1587	ASN	SER	ENGINEERED MUTATION	UNP P76578
A	1654	LEU	-	EXPRESSION TAG	UNP P76578
A	1655	GLU	-	EXPRESSION TAG	UNP P76578
A	1656	HIS	-	EXPRESSION TAG	UNP P76578
A	1657	HIS	-	EXPRESSION TAG	UNP P76578
A	1658	HIS	-	EXPRESSION TAG	UNP P76578
A	1659	HIS	-	EXPRESSION TAG	UNP P76578
A	1660	HIS	-	EXPRESSION TAG	UNP P76578
A	1661	HIS	-	EXPRESSION TAG	UNP P76578

4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	176.06Å 176.06Å 161.13Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.87 – 3.65 46.87 – 3.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.87-3.65) 100.0 (46.87-3.65)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.83 (at 3.66Å)	Xtrriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.177 , 0.238 0.183 , 0.239	Depositor DCC
R_{free} test set	1033 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	126.2	Xtrriage
Anisotropy	0.162	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 113.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.037 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8699	wwPDB-VP
Average B, all atoms (Å ²)	144.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.50	0/8863	0.75	3/12024 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	460	MSE	CG-SE-CE	7.75	115.94	98.90
1	A	1105	LEU	CA-CB-CG	5.31	127.52	115.30
1	A	1488	MSE	CA-CB-CG	-5.20	104.46	113.30

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	8699	0	8589	370	0
All	All	8699	0	8589	370	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 21.

The worst 5 of 370 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:639:ILE:HG21	1:A:738:TRP:O	1.73	0.89
1:A:396:LEU:HD22	1:A:906:ILE:HD11	1.56	0.87
1:A:1045:ASN:OD1	1:A:1047:THR:HG22	1.77	0.84
1:A:1504:VAL:HG21	1:A:1639:TRP:CD1	2.15	0.81
1:A:1106:ALA:HA	1:A:1113:ALA:HA	1.62	0.81

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	1112/1639 (68%)	989 (89%)	105 (9%)	18 (2%)	9 43

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	642	GLU
1	A	730	GLU
1	A	1169	PRO
1	A	1178	LYS
1	A	947	PHE

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	927/1340 (69%)	902 (97%)	25 (3%)	44 68

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1155	SER
1	A	1228	SER
1	A	1607	TYR
1	A	1198	SER
1	A	1274	ASP

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

Mol	Chain	Res	Type
1	A	1139	GLN
1	A	1172	ASN
1	A	1470	GLN
1	A	1590	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 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 [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, 95th 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	1103/1639 (67%)	0.05	29 (2%) 56 42	85, 139, 205, 267	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1108	PRO	4.9
1	A	943	ALA	4.0
1	A	616	GLY	3.3
1	A	615	PRO	3.2
1	A	651	VAL	3.1

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 carbohydrates 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.