



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 13, 2024 – 01:34 PM EST

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 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
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

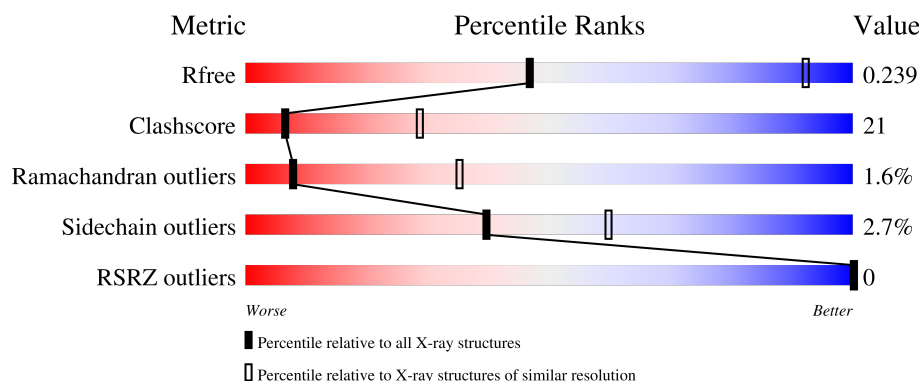
1 Overall quality at a glance

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}	164625	1261 (3.80-3.52)
Clashscore	180529	1328 (3.80-3.52)
Ramachandran outliers	177936	1306 (3.80-3.52)
Sidechain outliers	177891	1303 (3.80-3.52)
RSRZ outliers	164620	1260 (3.80-3.52)

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 $\leq 5\%$. 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
1	A	1122	Total	C	N	O	S	Se	0	0	0
			8699	5497	1505	1677	1	19			

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

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.

[illegible]

M1634	V1548	S1428	L1301	P1184	R1088	T976	V892
Y1635	D1549	N1429	K1302	Y1185	A1089	L977	N893
V1636	L1550	F1432	A1303	L1188	P1091	Q980	V896
P1637	L1551		S1304	L1195	E1097	G981	
Q1638		G1443	K1305	L1198	T1098	V885	V899
V1639	E1556	K1444	Y1311	S1198	Q1099	T986	L904
R1640	L1557	W1445		L1199	S1103	D987	N905
A1641	E1558	Q1446	A1326	Y1200	G1104	D991	I906
	N1559	T1457	L1327	T1201	L1105	T1004	T907
L1649	Q1560		R1328	N1202	A1106		D908
I1650	N1561	K1460	E1329	Q1205	L1107		Y909
V1651	L1562	A1461	I1330	L1206	E1110	N1011	W915
R1652	ALA	Q1462	W1331	L1209	T1111	E1012	
P1653	ASN	N1463	D1336	L1209	V1112	V1015	Q921
LEU	GLY	S1464	A1337	G1210	A1113	T1016	K922
GLU	N1463		A1338	T1211	D1114	V1017	ARG
HIS	L1569	L1471	S1339	K1212		A1018	TYR
HIS	E1570	L1474	G1340	K1218	K1120		GLY
HIS	Q1571		L1341		I1121	V1021	ALA
HIS	S1572	S1477	L1343	S1222	R1124	I1022	ASP
HIS		N1478	L1344	V1223	P1125	M1027	ILE
	E1575	S1479	Q1345	G1226	T1131	M1031	TYR
	V1576	P1483	L1350	T1227	V1132	A1032	ASP
	Q1577	W1485	K1351	S1228	N1133	S1033	ILE
	N1578	L1486	D1355	R1299	A1137	T1036	GLY
	L1579	R1487	R1358	L1230	L1138	L1039	GLN
	L1580	R1488	R1359	L1231	Q1139	T1044	GLY
	N1581	A1490	L1365	M1233	P1140	N1045	ARG
	Q1582	S1491		A1241	E1142		LEU
		N1503	R1371	M1258	I1146	L1046	A943
	E1592	V1504	I1377	V1263	P1147	T1047	
	F1593	L1505	W1378	R1263	S1155	D1048	R946
	V1599	Q1506	L1379	V1263		Q1051	F947
	E1606		G1383	T1273	G1161	T1058	G948
	Y1607	R1509	R1387	D1274	Q1162	L1063	G949
	Q1608	G1513	N1399	I1276	G1167	V1066	D950
		G1516	K1400	L1283	K1168	P1170	G951
	L1615	K1519	L1401	L1284	P1169	A1071	R956
	R1617	L1524	L1410	R1285	L1171	A1076	G957
	A1618	W1533	L1413	L1287	I1172	P1077	G958
	V1619	L1534	Q1416	M1292	A1174	V1078	K959
	T1620	Q1535	A1417	M1293	R1175	G1079	P960
	P1621	S1541	E1420	Y1296	I1177	R1080	P961
	P1627	V1542		Y1297	K1178		N963
		A1545		A1298	E1179	P1086	Q971
	M1630					V1087	A972
	V1631						V975
	E1632						
	S1633						

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) 99.9 (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)	Xtriage
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	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 118.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.037 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	8699	wwPDB-VP
Average B, all atoms (Å ²)	144.0	wwPDB-VP

Xtriage'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%)	8	35

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%)	40 60

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

Mol	Chain	Res	Type
1	A	1274	ASP
1	A	1378	TRP
1	A	1630	MSE
1	A	1336	ASP
1	A	1477	SER

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	1139	GLN
1	A	1172	ASN
1	A	1470	GLN
1	A	1590	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

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%)	-0.90	0 100 100	85, 139, 205, 267	0

There are no RSRZ outliers to report.

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.