



Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 06:29 pm BST

PDB ID : 4U59
Title : Crystal structure of Salmonella alpha-2-macroglobulin reacted with methy-
lamine
Authors : Wong, S.G.; Dessen, A.
Deposited on : 2014-07-24
Resolution : 3.09 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

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

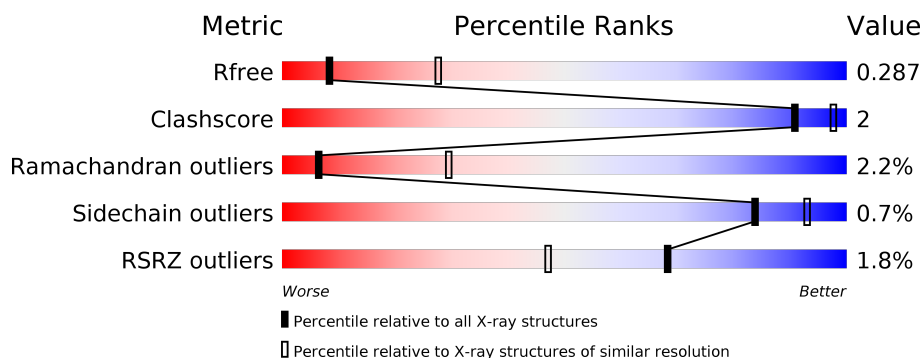
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.09 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	1647	<div> <div>2%</div> <div>87%</div> <div>7%</div> <div>7%</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 23680 atoms, of which 11798 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 alpha-2-macroglobulin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	1538	Total	C	H	N	O	S	0	0	0
			23680	7488	11798	2091	2280	23			

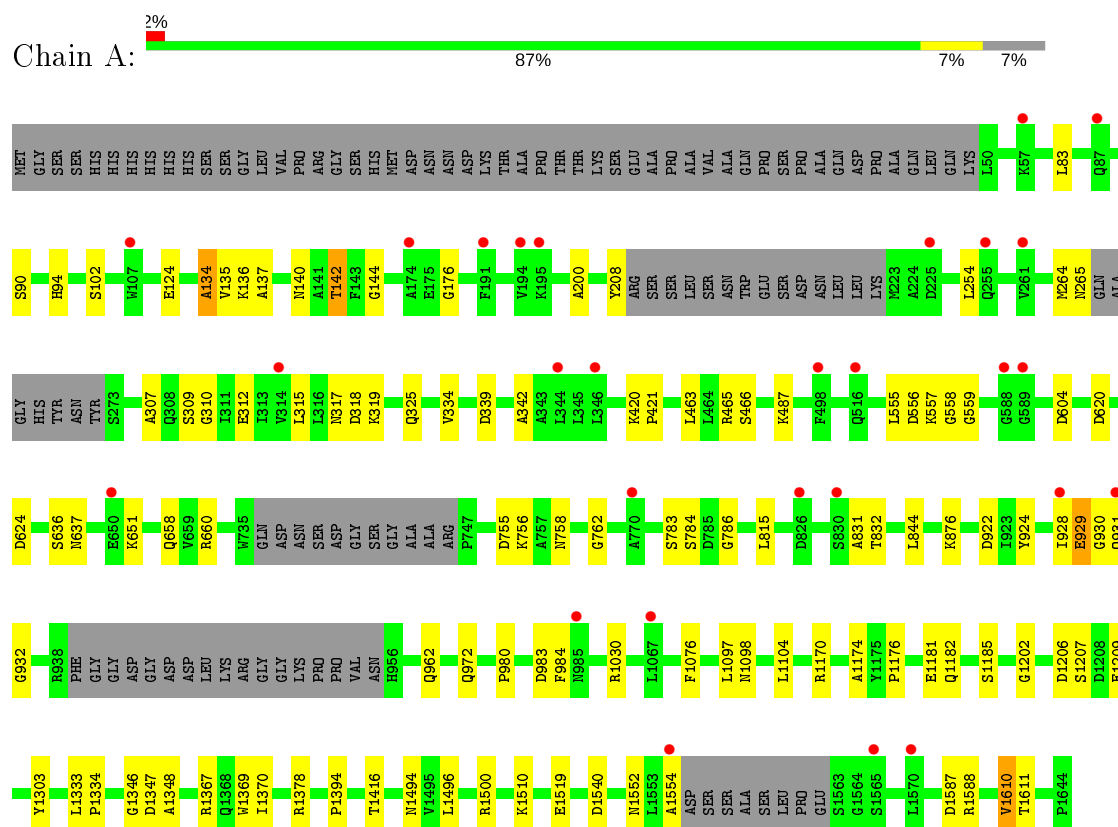
There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	initiating methionine	UNP Q8ZN46
A	-1	GLY	-	expression tag	UNP Q8ZN46
A	0	SER	-	expression tag	UNP Q8ZN46
A	1	SER	-	expression tag	UNP Q8ZN46
A	2	HIS	-	expression tag	UNP Q8ZN46
A	3	HIS	-	expression tag	UNP Q8ZN46
A	4	HIS	-	expression tag	UNP Q8ZN46
A	5	HIS	-	expression tag	UNP Q8ZN46
A	6	HIS	-	expression tag	UNP Q8ZN46
A	7	HIS	-	expression tag	UNP Q8ZN46
A	8	SER	-	expression tag	UNP Q8ZN46
A	9	SER	-	expression tag	UNP Q8ZN46
A	10	GLY	-	expression tag	UNP Q8ZN46
A	11	LEU	-	expression tag	UNP Q8ZN46
A	12	VAL	-	expression tag	UNP Q8ZN46
A	13	PRO	-	expression tag	UNP Q8ZN46
A	14	ARG	-	expression tag	UNP Q8ZN46
A	15	GLY	-	expression tag	UNP Q8ZN46
A	16	SER	-	expression tag	UNP Q8ZN46
A	17	HIS	-	expression tag	UNP Q8ZN46
A	18	MET	-	expression tag	UNP Q8ZN46
A	98	ALA	LYS	engineered mutation	UNP Q8ZN46
A	99	ALA	LYS	engineered mutation	UNP Q8ZN46

3 Residue-property plots [i](#)

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 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: alpha-2-macroglobulin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	254.96 Å 82.23 Å 98.92 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.73 – 3.09 56.65 – 3.09	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.73-3.09) 100.0 (56.65-3.09)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.08 (at 3.07 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.231 , 0.282 0.236 , 0.287	Depositor DCC
R_{free} test set	1952 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	75.3	Xtriage
Anisotropy	0.622	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 44.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	23680	wwPDB-VP
Average B, all atoms (Å ²)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MEQ

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	0.21	0/12112	0.39	0/16501

There are no bond length outliers.

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	11882	11798	11780	40	2
All	All	11882	11798	11780	40	2

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 2.

All (40) 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:922:ASP:OD2	1:A:924:TYR:N	2.25	0.69
1:A:1182:MEQ:HE2	1:A:1416:THR:HG21	1.80	0.64
1:A:134:ALA:O	1:A:136:LYS:N	2.29	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:317:ASN:O	1:A:319:LYS:N	2.35	0.60
1:A:325:GLN:OE1	1:A:325:GLN:N	2.38	0.56
1:A:962:GLN:NE2	1:A:980:PRO:O	2.37	0.55
1:A:930:GLY:O	1:A:932:GLY:N	2.40	0.55
1:A:1510:LYS:NZ	1:A:1519:GLU:OE2	2.31	0.54
1:A:463:LEU:O	1:A:465:ARG:NH1	2.43	0.52
1:A:755:ASP:OD1	1:A:756:LYS:N	2.43	0.51
1:A:784:SER:O	1:A:786:GLY:N	2.44	0.50
1:A:1494:ASN:O	1:A:1496:LEU:N	2.45	0.49
1:A:1170:ARG:NH2	1:A:1202:GLY:O	2.46	0.49
1:A:312:GLU:N	1:A:312:GLU:OE1	2.46	0.49
1:A:1207:SER:O	1:A:1209:GLU:N	2.41	0.49
1:A:972:GLN:N	1:A:972:GLN:OE1	2.46	0.48
1:A:555:LEU:O	1:A:557:LYS:N	2.46	0.48
1:A:1182:MEQ:HB2	1:A:1182:MEQ:HE3	1.96	0.47
1:A:1587:ASP:OD2	1:A:1588:ARG:NH1	2.47	0.47
1:A:636:SER:OG	1:A:637:ASN:N	2.46	0.46
1:A:1346:GLY:O	1:A:1348:ALA:N	2.50	0.45
1:A:557:LYS:O	1:A:559:GLY:N	2.50	0.45
1:A:1030:ARG:NH1	1:A:1076:PHE:CE1	2.85	0.45
1:A:1500:ARG:NH1	1:A:1540:ASP:OD1	2.48	0.44
1:A:1610:VAL:HG23	1:A:1611:THR:H	1.82	0.44
1:A:928:ILE:HG23	1:A:929:GLU:H	1.81	0.44
1:A:312:GLU:OE2	1:A:334:VAL:HG11	2.18	0.44
1:A:1174:ALA:HB3	1:A:1182:MEQ:HE1	1.99	0.43
1:A:604:ASP:OD1	1:A:604:ASP:N	2.47	0.43
1:A:312:GLU:OE2	1:A:334:VAL:HG21	2.19	0.42
1:A:783:SER:HB3	1:A:815:LEU:HD23	2.01	0.42
1:A:420:LYS:HB2	1:A:421:PRO:HD2	2.01	0.42
1:A:309:SER:OG	1:A:310:GLY:N	2.53	0.42
1:A:620:ASP:O	1:A:624:ASP:N	2.50	0.41
1:A:1181:GLU:OE2	1:A:1185:SER:OG	2.33	0.41
1:A:658:GLN:OE1	1:A:660:ARG:NE	2.49	0.41
1:A:83:LEU:HA	1:A:137:ALA:HA	2.03	0.41
1:A:200:ALA:HB1	1:A:315:LEU:HD21	2.03	0.41
1:A:1333:LEU:HB3	1:A:1334:PRO:HD3	2.02	0.41
1:A:1552:ASN:O	1:A:1554:ALA:N	2.47	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:124:GLU:OE2	1:A:466:SER:HG[2_555]	1.48	0.12
1:A:102:SER:OG	1:A:339:ASP:OD2[2_555]	2.10	0.10

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	1525/1647 (93%)	1374 (90%)	117 (8%)	34 (2%)	6	29

All (34) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	135	VAL
1	A	929	GLU
1	A	1370	ILE
1	A	90	SER
1	A	176	GLY
1	A	342	ALA
1	A	558	GLY
1	A	1098	ASN
1	A	1104	LEU
1	A	1347	ASP
1	A	1367	ARG
1	A	144	GLY
1	A	307	ALA
1	A	487	LYS
1	A	983	ASP
1	A	1176	PRO
1	A	1206	ASP
1	A	134	ALA
1	A	140	ASN
1	A	254	LEU
1	A	556	ASP
1	A	651	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	758	ASN
1	A	762	GLY
1	A	831	ALA
1	A	832	THR
1	A	984	PHE
1	A	1394	PRO
1	A	264	MET
1	A	318	ASP
1	A	1369	TRP
1	A	142	THR
1	A	876	LYS
1	A	931	GLN

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 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	1259/1347 (94%)	1250 (99%)	9 (1%)	84 93

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	94	HIS
1	A	142	THR
1	A	208	TYR
1	A	265	ASN
1	A	844	LEU
1	A	1097	LEU
1	A	1303	TYR
1	A	1378	ARG
1	A	1610	VAL

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

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
1	MEQ	A	1182	1	8,9,10	0.89	0	5,10,12	1.09	1 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MEQ	A	1182	1	-	6/8/9/11	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1182	MEQ	CG-CD-NE2	2.11	119.22	116.29

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	1182	MEQ	C-CA-CB-CG
1	A	1182	MEQ	CG-CD-NE2-CE
1	A	1182	MEQ	OE1-CD-NE2-CE
1	A	1182	MEQ	CA-CB-CG-CD
1	A	1182	MEQ	OE1-CD-CG-CB

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	1182	MEQ	NE2-CD-CG-CB

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1182	MEQ	3	0

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	1537/1647 (93%)	0.16	28 (1%)	68 47	35, 78, 121, 148	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	588	GLY	6.5
1	A	830	SER	3.4
1	A	87	GLN	3.2
1	A	174	ALA	3.2
1	A	195	LYS	3.1
1	A	57	LYS	3.0
1	A	589	GLY	2.9
1	A	516	GLN	2.9
1	A	225	ASP	2.8
1	A	985	ASN	2.8
1	A	194	VAL	2.7
1	A	931	GLN	2.7
1	A	344	LEU	2.6
1	A	1554	ALA	2.6
1	A	498	PHE	2.5
1	A	314	VAL	2.4
1	A	770	ALA	2.3
1	A	928	ILE	2.3
1	A	107	TRP	2.3
1	A	826	ASP	2.3
1	A	1565	SER	2.3
1	A	346	LEU	2.3
1	A	650	GLU	2.1
1	A	261	VAL	2.1
1	A	255	GLN	2.1
1	A	191	PHE	2.0
1	A	1570	LEU	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1067	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
1	MEQ	A	1182	10/11	0.94	0.29	51,77,98,117	0

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.