



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 4, 2016 – 01:18 PM EDT

PDB ID : 5B86
Title : Crystal structure of M-Sec
Authors : Yamashita, M.; Sato, Y.; Yamagata, A.; Fukai, S.
Deposited on : 2016-06-12
Resolution : 3.02 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

<http://wwpdb.org/validation/2016/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 : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

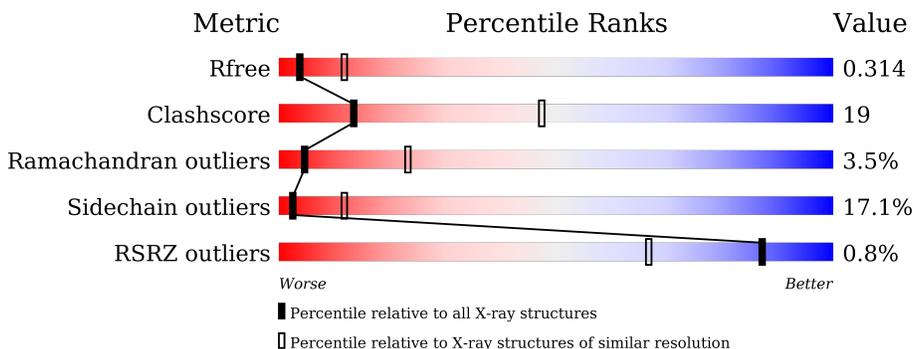
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.02 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}	91344	1773 (3.04-3.00)
Clashscore	102246	2117 (3.04-3.00)
Ramachandran outliers	100387	2050 (3.04-3.00)
Sidechain outliers	100360	2053 (3.04-3.00)
RSRZ outliers	91569	1788 (3.04-3.00)

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. 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	600	 % 52% 38% 6% . .
1	B	600	 % 49% 39% 9% . .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9312 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 Tumor necrosis factor alpha-induced protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	579	4638	2948	814	859	8	9	0	0	0
1	B	578	4628	2942	811	858	8	9	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	51	GLY	-	expression tag	UNP Q61333
A	52	PRO	-	expression tag	UNP Q61333
B	51	GLY	-	expression tag	UNP Q61333
B	52	PRO	-	expression tag	UNP Q61333

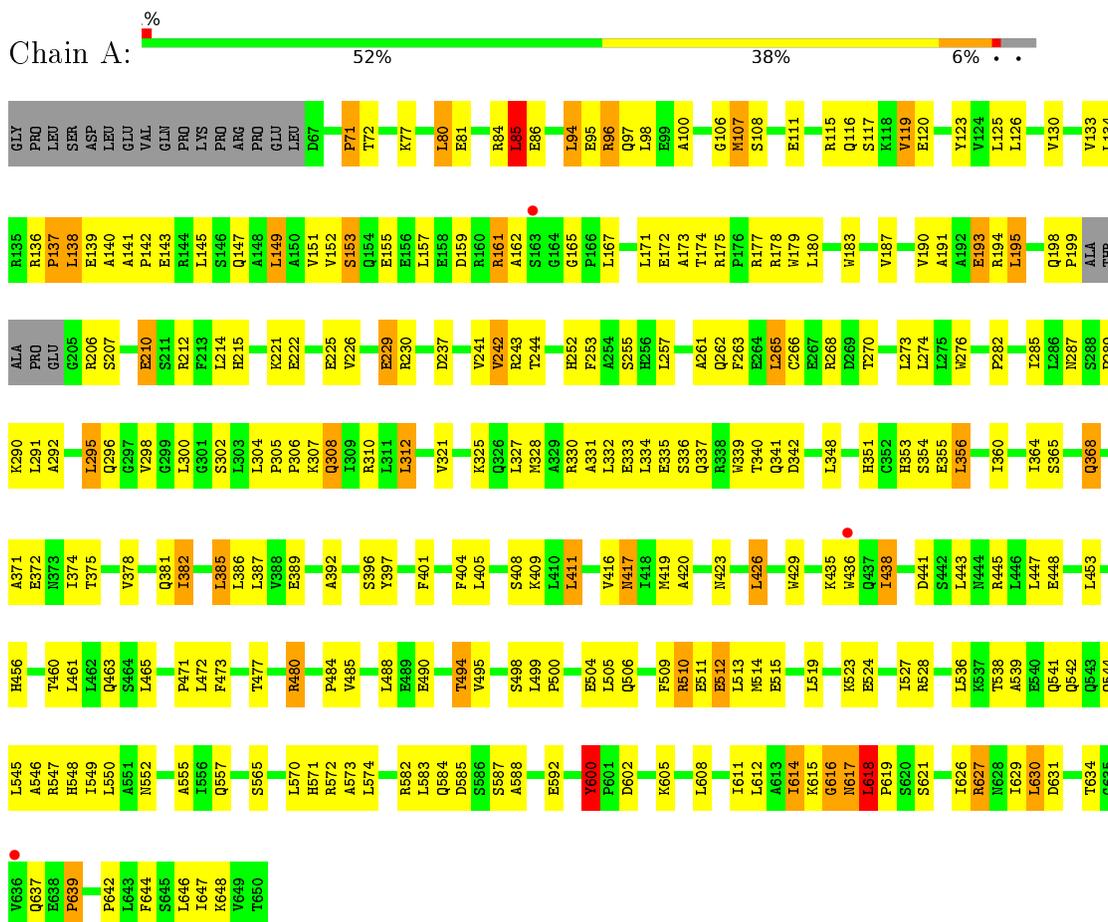
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	16	Total	O	0	0
			16	16		
2	B	30	Total	O	0	0
			30	30		

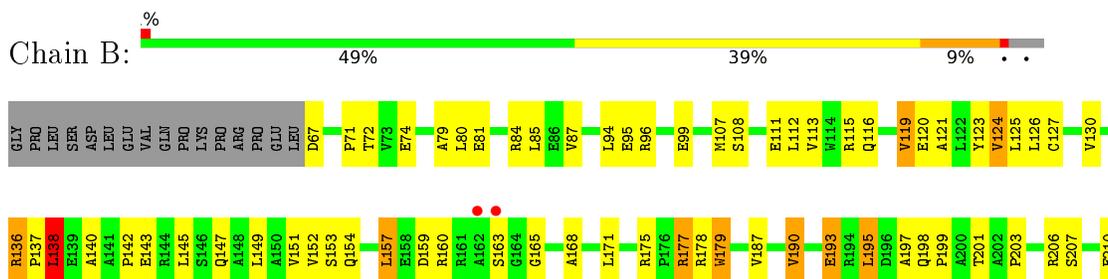
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Tumor necrosis factor alpha-induced protein 2



- Molecule 1: Tumor necrosis factor alpha-induced protein 2



S211	Y281	I358	R445	L519	G616
R212	F282	D359	L446	K523	M617
F213	M283	I360	L447	E524	L618
L214	D284	L361	F448	P449	P619
H215	I285	Q362	L450	I527	S620
M216	L286	I363	R528	R529	S621
Q217	M287	I364	L453	L529	E622
R218	S288	Q368	K454	R633	V623
T219	P289	E372	F458	L634	R624
K221	K290	E373	L461	V535	S625
L224	E294	N373	L462	L636	I626
V227	L303	T375	Q463	K637	R627
V228	L304	D377	S464	A539	N628
E229	P305	V378	L465	E540	I629
R230	P306	Q381	F466	Q541	D631
L231	K307	I382	L467	Q542	ILE
L234	Q308	I383	D468	Q543	ASN
F235	R310	Q384	L469	Q544	THR
P236	L311	Q385	K470	L550	GLY
E237	L312	L386	P471	A551	VAL
D238	L312	L387	L472	N552	GLN
F239	E313	L393	F473	Q557	E638
M240	F316	L394	K474	G558	P639
V241	L317	R395	F476	F559	P642
V242	S318	Q398	T477	L478	L643
R243	N319	S408	Q478	Q479	F644
T244	F320	K409	R480	R481	S645
E247	V321	L410	W481	T487	L646
Y251	S323	L411	L488	L489	L647
H252	V324	R415	E489	F490	R648
F253	Q326	W416	L491	I491	V649
A254	L327	M419	T492	T493	T650
S255	M328	A420	T494	Y495	
H256	A329	W423	L494	S496	
L257	R330	W424	Y495	E501	
C258	L334	C425	F502	S503	
L260	E335	L426	F427	E504	
F263	W339	F428	W429	L505	
E264	D342	F428	W436	B506	
C266	V343	W429	Q437	C508	
D269	A344	W436	I438	F509	
L272	P345	I436	S439	R510	
L273	G350	H440	H440	L513	
L274	H351	D441	D441	R514	
L275	C352	S442	S442	E515	
L276	H353	L356	L443	K615	
W279	S354	A357	R444		
L280	L356				
	A357				

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	91.41Å 107.83Å 229.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.44 – 3.02 46.44 – 3.02	Depositor EDS
% Data completeness (in resolution range)	86.1 (46.44-3.02) 87.3 (46.44-3.02)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 3.01Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.229 , 0.292 0.250 , 0.314	Depositor DCC
R_{free} test set	2148 reflections (4.97%)	DCC
Wilson B-factor (Å ²)	42.5	Xtrriage
Anisotropy	0.231	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 64.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	9312	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% 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	0/4717	0.71	2/6380 (0.0%)
1	B	0.46	0/4708	0.69	2/6369 (0.0%)
All	All	0.48	0/9425	0.70	4/12749 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	138	LEU	CA-CB-CG	7.45	132.44	115.30
1	B	265	LEU	CA-CB-CG	6.26	129.70	115.30
1	A	618	LEU	CA-CB-CG	6.03	129.17	115.30
1	A	85	LEU	CA-CB-CG	5.41	127.74	115.30

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	4638	0	4693	166	1
1	B	4628	0	4679	183	1
2	A	16	0	0	3	0
2	B	30	0	0	9	0
All	All	9312	0	9372	349	1

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 19.

The worst 5 of 349 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:B:605:LYS:HE3	1:B:630:LEU:HB2	1.52	0.92
1:B:179:TRP:NE1	2:B:701:HOH:O	2.00	0.91
1:A:608:LEU:HD23	1:A:626:ILE:HG12	1.51	0.91
1:A:328:MSE:HE2	1:A:360:ILE:HG23	1.53	0.90
1:B:615:LYS:HB3	1:B:618:LEU:HD21	1.55	0.89

All (1) 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:480:ARG:NH1	1:B:264:GLU:OE2[4_455]	2.17	0.03

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	575/600 (96%)	483 (84%)	67 (12%)	25 (4%)	3 18
1	B	574/600 (96%)	496 (86%)	63 (11%)	15 (3%)	7 31
All	All	1149/1200 (96%)	979 (85%)	130 (11%)	40 (4%)	4 23

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	71	PRO
1	A	463	GLN
1	A	639	PRO
1	B	286	LEU
1	B	616	GLY

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	504/513 (98%)	429 (85%)	75 (15%)	4 16
1	B	502/513 (98%)	405 (81%)	97 (19%)	2 9
All	All	1006/1026 (98%)	834 (83%)	172 (17%)	2 12

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	94	LEU
1	B	212	ARG
1	B	543	GLN
1	B	116	GLN
1	B	149	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	552	ASN
1	B	541	GLN
1	B	256	HIS
1	A	520	HIS
1	B	250	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	570/600 (95%)	-0.06	3 (0%) 91 76	28, 74, 126, 185	0
1	B	569/600 (94%)	-0.05	6 (1%) 82 57	36, 77, 132, 184	0
All	All	1139/1200 (94%)	-0.06	9 (0%) 87 67	28, 76, 130, 185	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	636	VAL	7.0
1	B	162	ALA	4.4
1	B	415	ARG	3.6
1	B	354	SER	3.1
1	A	163	SER	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.