



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 07:32 pm GMT

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
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

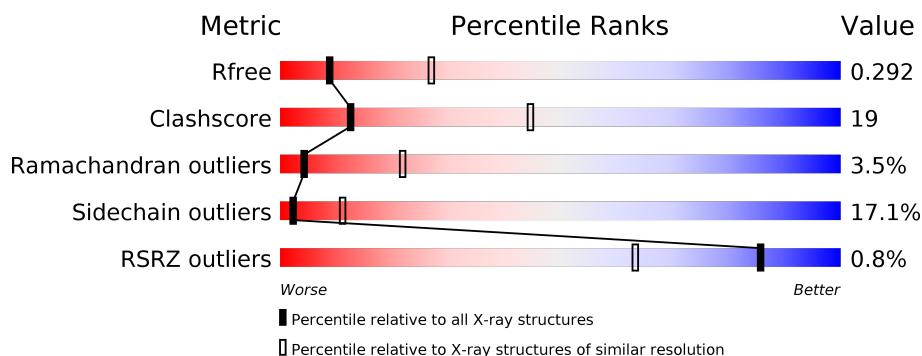
1 Overall quality at a glance ⓘ

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}	100719	1924 (3.04-3.00)
Clashscore	112137	2279 (3.04-3.00)
Ramachandran outliers	110173	2207 (3.04-3.00)
Sidechain outliers	110143	2210 (3.04-3.00)
RSRZ outliers	101464	1948 (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	<div> <div>%</div> <div> <div></div> <div>52%</div> <div>38%</div> <div>6%</div> <div>• •</div> </div> </div>
1	B	600	<div> <div>%</div> <div> <div></div> <div>49%</div> <div>39%</div> <div>9%</div> <div>• •</div> </div> </div>

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
1	A	579	Total	C	N	O	S	Se	0	0	0
			4638	2948	814	859	8	9			
1	B	578	Total	C	N	O	S	Se	0	0	0
			4628	2942	811	858	8	9			

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

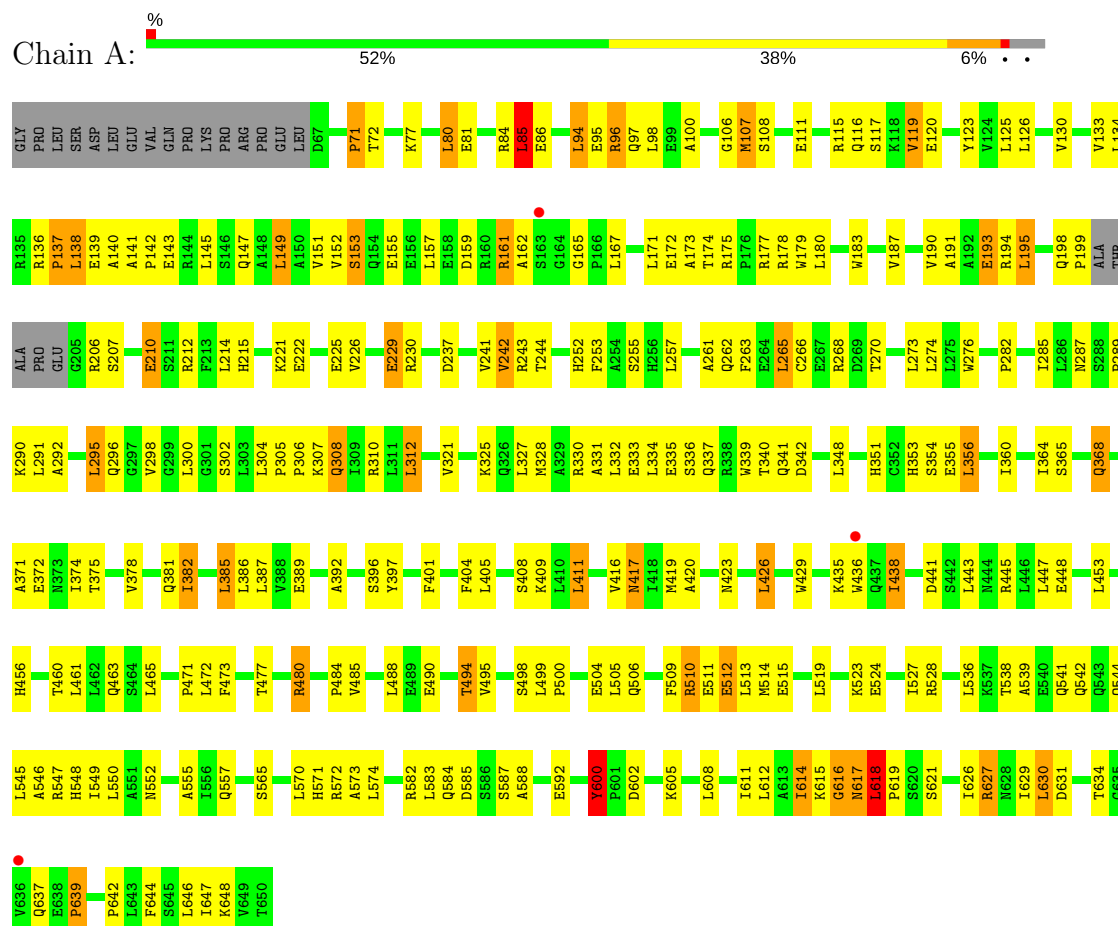
- Molecule 2 is water.

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

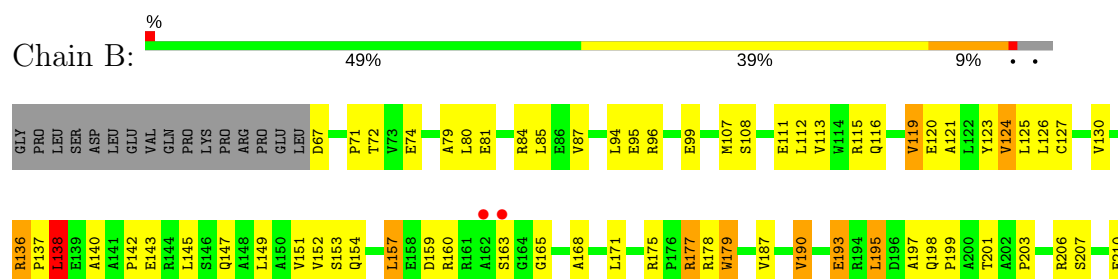
3 Residue-property plots

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



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



G616	G617	L618	P619	S620	S621	S622	V623	R624	S625	I626	R627	N628	I629	L630	D631	ILE	ASN	THR	GLY	VAL	GLN	E638	P639	P642	L643	F644	S645	L646	I647	K648	V649	T650																						
L519	R445	L446	L447	E448	P449	L450	L453	K454	F458	L461	L462	Q463	S464	L465	F466	L467	D468	L469	K470	P471	L472	F473	K474	Q557	G558	F559	E562	A567	T568	M576	E579	I580	I581	R582	L583	A588	I589	K590	V593	D602	K605	G606	H607	L608	L612	A613	I614	E515	K615					
I358	D359	I360	L361	Q362	P449	L450	L453	K454	F458	I374	I375	Q463	S464	L465	F466	L467	D468	L469	K470	P471	L472	F473	K474	Q393	L394	R395	Q398	S408	K409	L410	L411	R415	V416	M419	A420	M423	M424	C425	L426	F427	F428	M429	W436	Q437	I438	S439	H440	D441	S442	L443	H444			
Y281	F282	N283	D284	L285	L286	N287	S288	P289	K290	L291	E294	L303	L304	P305	P306	K307	Q308	I309	R310	L311	D312	E313	F316	L317	S318	N319	R243	F390	V321	T322	S323	V324	K325	Q326	L327	M328	A329	R330	L334	E335	W339	E264	D342	V343	A344	P345	G350	H351	C352	H353	S354	E355	L356	A357
S211	R212	F213	L214	H215	M216	G217	R218	T219	M220	K221	L224	V227	V228	E229	R230	L231	L234	F235	P236	D237	E238	F239	M240	V241	V242	R243	T244	E247	Y251	H252	F253	A254	S255	H256	L257	C258	A259	L260	F263	E264	L265	C266	D269	L272	L273	L274	L275	V276	N279	L280				

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	91.41 Å 107.83 Å 229.58 Å 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 3.01 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.229 , 0.292 0.225 , 0.292	Depositor DCC
R_{free} test set	3717 reflections (5.00%)	DCC
Wilson B-factor (Å ²)	42.5	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 64.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
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

Xtriage'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	17
1	B	574/600 (96%)	496 (86%)	63 (11%)	15 (3%)	6	29
All	All	1149/1200 (96%)	979 (85%)	130 (11%)	40 (4%)	4	22

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 ⓘ

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%)	3	16
1	B	502/513 (98%)	405 (81%)	97 (19%)	1	8
All	All	1006/1026 (98%)	834 (83%)	172 (17%)	2	11

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 ⓘ

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 [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%)	90 73	28, 74, 126, 185	0
1	B	569/600 (94%)	-0.05	6 (1%)	80 54	36, 77, 132, 184	0
All	All	1139/1200 (94%)	-0.06	9 (0%)	86 63	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.