



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:24 PM GMT

PDB ID : 3TVJ
Title : Catalytic fragment of MASP-2 in complex with its specific inhibitor developed by directed evolution on SGCI scaffold
Authors : Heja, D.; Harmat, V.; Dobo, J.; Szasz, R.; Kekesi, K.A.; Zavodszky, P.; Gal, P.; Pal, G.
Deposited on : 2011-09-20
Resolution : 1.28 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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3450 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 Mannan-binding lectin serine protease 2 A chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	83	Total	C	N	O	S	5	5	0
			659	412	107	134	6			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	359	ALA	-	EXPRESSION TAG	UNP O00187
A	360	SER	-	EXPRESSION TAG	UNP O00187
A	361	MET	-	EXPRESSION TAG	UNP O00187
A	362	THR	-	EXPRESSION TAG	UNP O00187
A	413	ASP	GLU	CONFLICT	UNP O00187

- Molecule 2 is a protein called Mannan-binding lectin serine protease 2 B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	242	Total	C	N	O	S	5	15	0
			1969	1257	332	368	12			

- Molecule 3 is a protein called Protease inhibitor SGPI-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	35	Total	C	N	O	S	4	3	0
			283	169	49	59	6			

There are 8 discrepancies between the modelled and reference sequences:

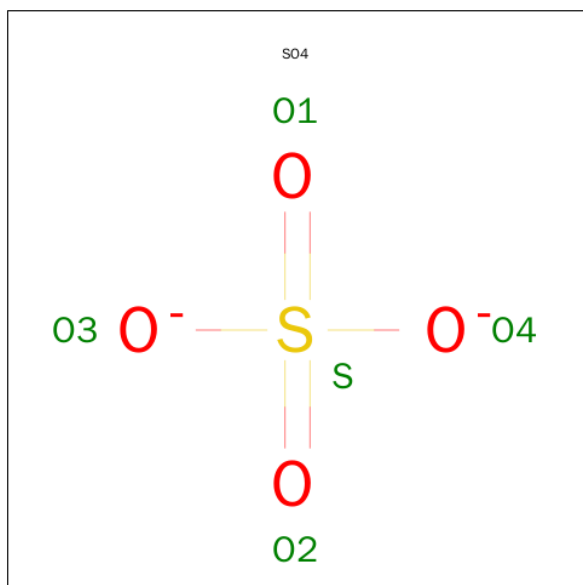
Chain	Residue	Modelled	Actual	Comment	Reference
I	-2	GLY	-	EXPRESSION TAG	UNP O46162
I	-1	SER	-	EXPRESSION TAG	UNP O46162
I	0	GLY	-	EXPRESSION TAG	UNP O46162
I	27	VAL	ALA	ENGINEERED MUTATION	UNP O46162
I	30	LYS	LEU	ENGINEERED MUTATION	UNP O46162

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	31	LEU	LYS	ENGINEERED MUTATION	UNP O46162
I	32	TRP	ALA	ENGINEERED MUTATION	UNP O46162
I	34	ASN	PRO	ENGINEERED MUTATION	UNP O46162

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total O S 5 4 1	0	0

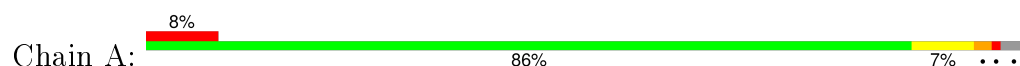
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	127	Total O 133 133	0	7
5	B	334	Total O 348 348	0	15
5	I	48	Total O 53 53	0	5

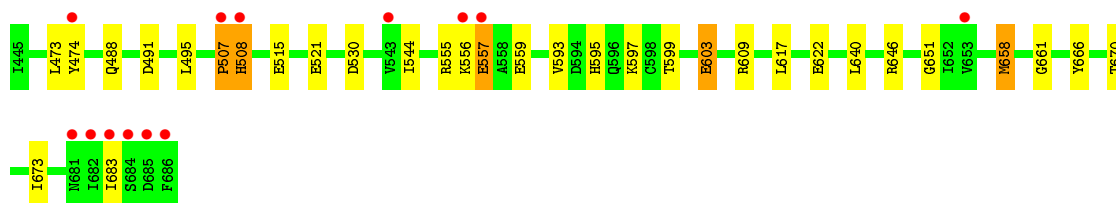
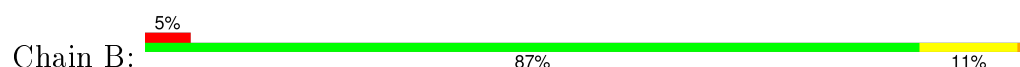
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ($\text{RSRZ} > 2$). 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.

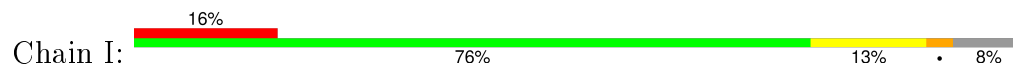
- Molecule 1: Mannan-binding lectin serine protease 2 A chain



- Molecule 2: Mannan-binding lectin serine protease 2 B chain



- Molecule 3: Protease inhibitor SGPI-2



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	50.92Å 62.47Å 114.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.00 – 1.28 38.00 – 1.28	Depositor EDS
% Data completeness (in resolution range)	95.9 (38.00-1.28) 95.9 (38.00-1.28)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.29 (at 1.28Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.156 , 0.205 0.165 , 0.211	Depositor DCC
R_{free} test set	4584 reflections (5.33%)	DCC
Wilson B-factor (Å ²)	11.2	Xtriage
Anisotropy	0.178	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 58.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 90574 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3450	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.94% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: SO4

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	1.28	2/679 (0.3%)	1.09	5/920 (0.5%)
2	B	1.20	5/2050 (0.2%)	1.11	7/2783 (0.3%)
3	I	1.20	1/291 (0.3%)	1.00	0/386
All	All	1.22	8/3020 (0.3%)	1.09	12/4089 (0.3%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	389	LYS	CD-CE	-14.89	1.14	1.51
3	I	1	GLU	N-CA	-7.52	1.31	1.46
2	B	508[A]	HIS	N-CA	6.25	1.58	1.46
2	B	508[B]	HIS	N-CA	6.25	1.58	1.46
2	B	603	GLU	CB-CG	-6.17	1.40	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	376[A]	ARG	NE-CZ-NH2	-7.97	116.31	120.30
1	A	376[B]	ARG	NE-CZ-NH2	-7.97	116.31	120.30
2	B	530	ASP	CB-CG-OD1	7.27	124.84	118.30
2	B	491	ASP	CB-CG-OD2	-5.98	112.92	118.30
2	B	507	PRO	C-N-CA	5.84	136.30	121.70

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	659	0	635	12	0
2	B	1969	0	1916	24	0
3	I	283	0	272	8	0
4	B	5	0	0	0	0
5	A	133	0	0	6	1
5	B	348	0	0	5	2
5	I	53	0	0	1	0
All	All	3450	0	2823	42	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 7.

The worst 5 of 42 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:404[B]:LYS:HG3	1:A:405[B]:VAL:HG23	1.43	1.00
1:A:376[B]:ARG:HD2	5:A:475:HOH:O	1.75	0.87
3:I:5[B]:GLU:HG2	3:I:8:THR:OG1	1.75	0.85
2:B:507:PRO:HG2	2:B:508[B]:HIS:CD2	2.13	0.84
1:A:376[B]:ARG:NH2	5:A:468[B]:HOH:O	2.16	0.77

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)
5:A:148:HOH:O	5:B:370:HOH:O[4_445]	1.93	0.27
5:B:310:HOH:O	5:B:419:HOH:O[3_445]	2.15	0.05

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	86/86 (100%)	82 (95%)	4 (5%)	0	100	100
2	B	255/242 (105%)	245 (96%)	10 (4%)	0	100	100
3	I	35/38 (92%)	32 (91%)	3 (9%)	0	100	100
All	All	376/366 (103%)	359 (96%)	17 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	75/72 (104%)	73 (97%)	2 (3%)	52	11
2	B	209/194 (108%)	208 (100%)	1 (0%)	92	71
3	I	34/32 (106%)	32 (94%)	2 (6%)	24	1
All	All	318/298 (107%)	313 (98%)	5 (2%)	82	31

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	404[A]	LYS
1	A	404[B]	LYS
2	B	557	GLU
3	I	35[A]	GLN
3	I	35[B]	GLN

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

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](#)

1 ligand 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
4	SO4	B	1	-	4,4,4	2.52	1 (25%)	6,6,6	0.87	0

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
4	SO4	B	1	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1	SO4	O2-S	4.70	1.63	1.47

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	83/86 (96%)	0.35	7 (8%)	14 9	10, 16, 27, 40	2 (2%)
2	B	242/242 (100%)	0.48	13 (5%)	29 24	8, 14, 27, 46	2 (0%)
3	I	35/38 (92%)	0.82	6 (17%)	2 2	10, 21, 35, 51	2 (5%)
All	All	360/366 (98%)	0.49	26 (7%)	18 13	8, 15, 30, 51	6 (1%)

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	686	PHE	6.9
1	A	405[A]	VAL	5.5
1	A	362	THR	5.5
3	I	7	GLY	4.8
2	B	684	SER	4.6

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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
4	SO4	B	1	5/5	0.96	0.09	-0.74	14,17,18,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.