



wwPDB X-ray Structure Validation Summary Report ⓘ

May 10, 2022 – 10:06 pm BST

PDB ID : 7PQN
Title : Catalytic fragment of MASP-2 in complex with ecotin
Authors : Harmat, V.; Fodor, K.; Heja, D.
Deposited on : 2021-09-17
Resolution : 2.40 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.28.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.28.1

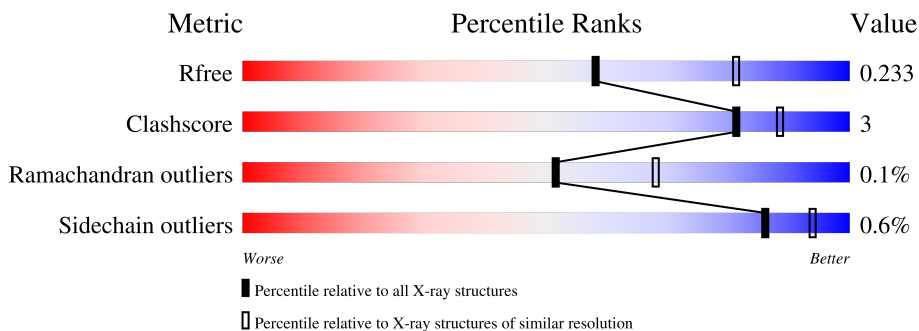
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	C	162	
1	D	162	
2	A	86	
2	aa	86	
3	B	242	
3	bb	242	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 6887 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 Ecotin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	141	1072	686	175	205	6	0	1	0
1	D	139	1042	670	171	195	6	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	78	557	353	84	114	6	0	0	0
2	aa	76	545	342	86	111	6	0	0	0

There are 8 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
aa	359	ALA	-	expression tag	UNP O00187
aa	360	SER	-	expression tag	UNP O00187
aa	361	MET	-	expression tag	UNP O00187
aa	362	THR	-	expression tag	UNP O00187

- Molecule 3 is a protein called Mannan-binding lectin serine protease 2 B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	238	1798	1148	297	343	10	0	0	0
3	bb	234	1739	1115	286	326	12	0	2	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 6 3 3	0	0
4	bb	1	Total C O 6 3 3	0	0
4	bb	1	Total C O 6 3 3	0	0


- Molecule 5 is water.

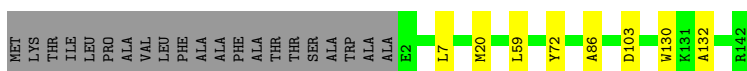
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	19	Total O 19 19	0	0
5	D	14	Total O 14 14	0	0
5	A	4	Total O 4 4	0	0
5	B	53	Total O 53 53	0	0
5	bb	26	Total O 26 26	0	0

3 Residue-property plots [i](#)


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. 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. 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.

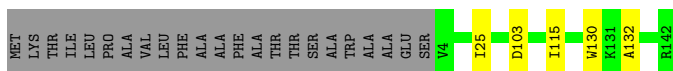
- Molecule 1: Ecotin

Chain C:  82% 5% 13%




- Molecule 1: Ecotin

Chain D:  83% 14%




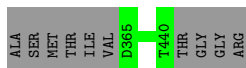
- Molecule 2: Mannan-binding lectin serine protease 2 A chain

Chain A:  88% 9%



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

Chain aa:  88% 12%



- Molecule 3: Mannan-binding lectin serine protease 2 B chain

Chain B:  89% 8% ..



- Molecule 3: Mannan-binding lectin serine protease 2 B chain

Chain bb:  95% ..

I445	A492	I567	Y602	GLU	LYS	PRO	PRO	TYR	PRO	ARG	GLY	S611	S633	F666
------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	99.56Å 102.75Å 109.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.83 – 2.40 29.83 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.7 (29.83-2.40) 99.7 (29.83-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.38 (at 2.39Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.192 , 0.233 0.192 , 0.233	Depositor DCC
R_{free} test set	2217 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å ²)	48.4	Xtrriage
Anisotropy	0.348	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.023 for k,h,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6887	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% 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: GOL

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	C	0.38	0/1097	0.59	0/1497
1	D	0.34	0/1064	0.55	0/1452
2	A	0.36	0/571	0.52	0/785
2	aa	0.34	0/558	0.53	0/765
3	B	0.41	0/1842	0.62	1/2508 (0.0%)
3	bb	0.37	0/1788	0.56	0/2439
All	All	0.38	0/6920	0.58	1/9446 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	575	LEU	CB-CG-CD1	-6.39	100.14	111.00

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	C	1072	0	1012	7	0
1	D	1042	0	983	5	0
2	A	557	0	489	1	0
2	aa	545	0	484	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1798	0	1687	13	0
3	bb	1739	0	1624	0	0
4	B	6	0	8	2	0
4	bb	12	0	16	0	0
5	A	4	0	0	0	0
5	B	53	0	0	0	0
5	C	19	0	0	0	0
5	D	14	0	0	0	0
5	bb	26	0	0	0	0
All	All	6887	0	6303	21	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 3.

The worst 5 of 21 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:C:103:ASP:HB3	1:D:103:ASP:HB3	1.64	0.80
3:B:559:GLU:HG2	3:B:673:ILE:HD13	1.72	0.72
3:B:537:LYS:NZ	3:B:686:PHE:OXT	2.23	0.69
1:C:86:ALA:HB3	3:B:467:THR:HB	1.80	0.63
3:B:515:GLU:HB2	3:B:539:ASN:HB3	1.83	0.60

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	C	140/162 (86%)	136 (97%)	4 (3%)	0	100 100
1	D	137/162 (85%)	134 (98%)	3 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	A	76/86 (88%)	75 (99%)	1 (1%)	0	100	100
2	aa	74/86 (86%)	71 (96%)	3 (4%)	0	100	100
3	B	234/242 (97%)	228 (97%)	6 (3%)	0	100	100
3	bb	232/242 (96%)	225 (97%)	6 (3%)	1 (0%)	34	48
All	All	893/980 (91%)	869 (97%)	23 (3%)	1 (0%)	51	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	bb	492	ALA

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	C	108/139 (78%)	107 (99%)	1 (1%)	78	90
1	D	102/139 (73%)	102 (100%)	0	100	100
2	A	57/72 (79%)	57 (100%)	0	100	100
2	aa	56/72 (78%)	56 (100%)	0	100	100
3	B	179/194 (92%)	178 (99%)	1 (1%)	86	94
3	bb	171/194 (88%)	169 (99%)	2 (1%)	71	85
All	All	673/810 (83%)	669 (99%)	4 (1%)	86	94

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	20	MET
3	B	633	SER
3	bb	567	ILE
3	bb	633	SER

Sometimes 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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

3 ligands are 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	GOL	bb	701	-	5,5,5	0.37	0	5,5,5	0.22	0
4	GOL	bb	702	-	5,5,5	0.35	0	5,5,5	0.43	0
4	GOL	B	700	-	5,5,5	0.35	0	5,5,5	0.57	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	bb	701	-	-	2/4/4/4	-
4	GOL	bb	702	-	-	2/4/4/4	-
4	GOL	B	700	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	700	GOL	O1-C1-C2-C3
4	bb	701	GOL	O1-C1-C2-C3
4	B	700	GOL	O1-C1-C2-O2
4	bb	701	GOL	O1-C1-C2-O2
4	B	700	GOL	C1-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	700	GOL	2	0

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.