



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 16, 2023 – 07:16 AM EDT

PDB ID : 1ZJK  
Title : Crystal structure of the zymogen catalytic region of human MASP-2  
Authors : Gal, P.; Harmat, V.; Kocsis, A.; Bian, T.; Barna, L.; Ambrus, G.; Vegh, B.;  
Balczer, J.; Sim, R.B.; Naray-Szabo, G.; Zavodszky, P.  
Deposited on : 2005-04-29  
Resolution : 2.18 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

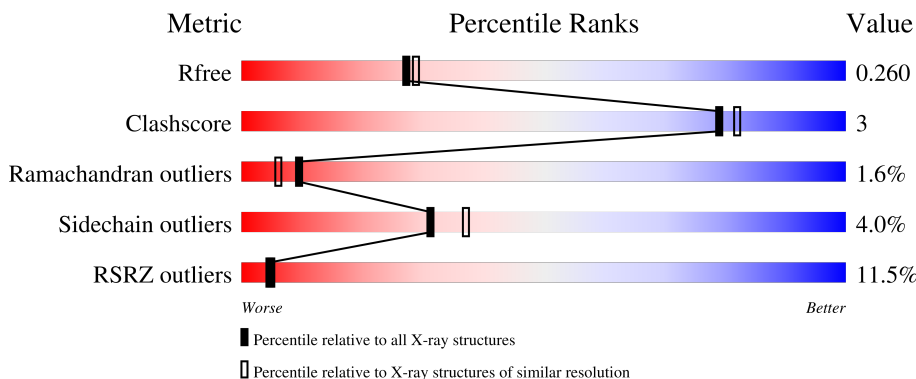
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.18 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	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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  $\geq 3$ , 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	403	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2994 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	390	2910	1855	481	552	22	1	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	284	ALA	-	cloning artifact	UNP O00187
A	285	SER	-	cloning artifact	UNP O00187
A	286	MET	-	cloning artifact	UNP O00187
A	444	GLN	ARG	engineered mutation	UNP O00187

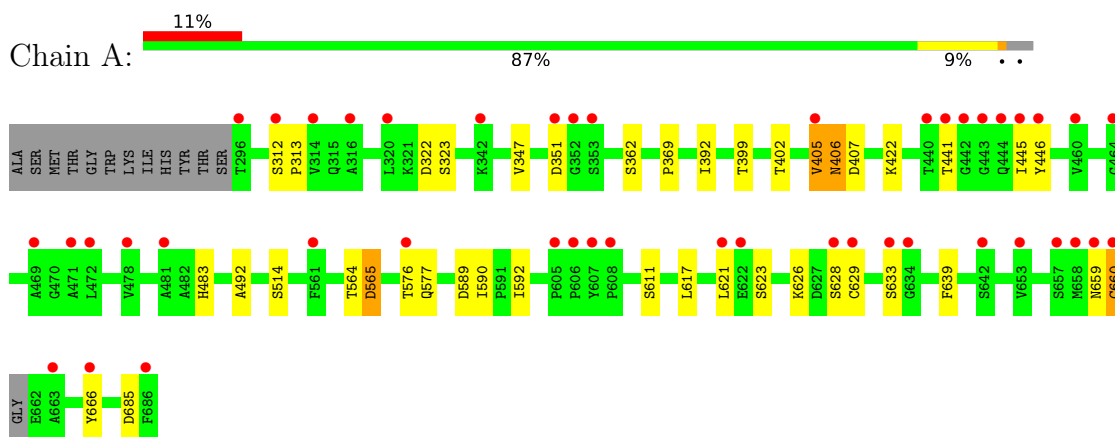
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	84	Total	O	0	0
			84	84		

### 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 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: Mannan-binding lectin serine protease 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.66Å 72.69Å 110.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.90 – 2.18 28.90 – 2.18	Depositor EDS
% Data completeness (in resolution range)	88.2 (28.90-2.18) 88.3 (28.90-2.18)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.45 (at 2.18Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.205 , 0.253 0.211 , 0.260	Depositor DCC
$R_{free}$ test set	938 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.2	Xtrriage
Anisotropy	0.161	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 36.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2994	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.87% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.49	1/2997 (0.0%)	0.65	8/4088 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	445	ILE	C-O	21.30	1.63	1.23

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	445	ILE	O-C-N	10.01	138.72	122.70
1	A	322	ASP	CB-CG-OD2	5.37	123.14	118.30
1	A	589	ASP	CB-CG-OD2	5.27	123.04	118.30
1	A	445	ILE	CA-C-O	-5.20	109.18	120.10
1	A	407	ASP	CB-CG-OD2	5.19	122.97	118.30
1	A	565	ASP	CB-CG-OD2	5.16	122.94	118.30
1	A	351	ASP	CB-CG-OD2	5.13	122.91	118.30
1	A	685	ASP	CB-CG-OD2	5.04	122.84	118.30

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	2910	0	2715	16	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	84	0	0	1	0
All	All	2994	0	2715	16	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.

All (16) 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:492:ALA:O	1:A:514:SER:OG	2.00	0.79
1:A:312:SER:OG	1:A:313:PRO:HD3	1.90	0.71
1:A:312:SER:OG	1:A:313:PRO:CD	2.50	0.58
1:A:422:LYS:NZ	2:A:80:HOH:O	2.40	0.55
1:A:483:HIS:NE2	1:A:633:SER:OG	2.33	0.55
1:A:564:THR:O	1:A:565:ASP:HB2	2.08	0.54
1:A:590:ILE:HG22	1:A:621:LEU:HG	1.90	0.53
1:A:483:HIS:CE1	1:A:633:SER:HG	2.23	0.53
1:A:312:SER:CB	1:A:313:PRO:HD3	2.40	0.51
1:A:405:VAL:O	1:A:406:ASN:HB3	2.13	0.49
1:A:312:SER:CB	1:A:313:PRO:CD	2.92	0.48
1:A:369:PRO:HB3	1:A:392:ILE:HD13	1.96	0.46
1:A:659:ASN:O	1:A:660:CYS:C	2.55	0.45
1:A:483:HIS:CE1	1:A:633:SER:OG	2.70	0.44
1:A:576:THR:O	1:A:576:THR:HG22	2.19	0.42
1:A:592:ILE:HD11	1:A:639:PHE:CZ	2.54	0.42

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	387/403 (96%)	352 (91%)	29 (8%)	6 (2%)	9 6

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	405	VAL
1	A	446	TYR
1	A	623	SER
1	A	626	LYS
1	A	577	GLN
1	A	406	ASN

### 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	298/330 (90%)	286 (96%)	12 (4%)	31 37

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	323	SER
1	A	347	VAL
1	A	362	SER
1	A	399	THR
1	A	402	THR
1	A	441	THR
1	A	611	SER
1	A	617	LEU
1	A	628	SER
1	A	629	CYS
1	A	660	CYS
1	A	666	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:



Mol	Chain	Res	Type
1	A	315	GLN
1	A	459	GLN
1	A	540	ASN

### 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	390/403 (96%)	0.51	45 (11%) <b>4</b> <b>5</b>	7, 26, 46, 63	1 (0%)

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	606	PRO	7.4
1	A	444	GLN	6.4
1	A	441	THR	6.1
1	A	440	THR	4.5
1	A	443	GLY	4.2
1	A	605	PRO	3.9
1	A	352	GLY	3.8
1	A	686	PHE	3.7
1	A	660	CYS	3.6
1	A	445	ILE	3.6
1	A	442	GLY	3.6
1	A	312	SER	3.3
1	A	622	GLU	3.1
1	A	405	VAL	3.1
1	A	629	CYS	3.0
1	A	658	MET	3.0
1	A	608	PRO	3.0
1	A	469	ALA	2.9
1	A	663	ALA	2.9
1	A	478	VAL	2.8
1	A	353	SER	2.8
1	A	351	ASP	2.7
1	A	633	SER	2.7
1	A	628	SER	2.6
1	A	464	GLY	2.6
1	A	657	SER	2.5
1	A	653	VAL	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	460	VAL	2.5
1	A	296	THR	2.5
1	A	659	ASN	2.4
1	A	634	GLY	2.4
1	A	471	ALA	2.4
1	A	561	PHE	2.4
1	A	666	TYR	2.4
1	A	472	LEU	2.3
1	A	316	ALA	2.3
1	A	642	SER	2.3
1	A	342	LYS	2.2
1	A	446	TYR	2.2
1	A	621	LEU	2.1
1	A	576	THR	2.1
1	A	481	ALA	2.1
1	A	607	TYR	2.1
1	A	320	LEU	2.0
1	A	314	VAL	2.0

## 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 monosaccharides 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.