



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 7, 2023 – 07:50 pm BST

PDB ID : 7ZM5  
Title : Structure of Mossman virus receptor binding protein  
Authors : Stelfox, A.J.; Bowden, T.A.; Rissanen, I.; Harlos, K.  
Deposited on : 2022-04-19  
Resolution : 1.62 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  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (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

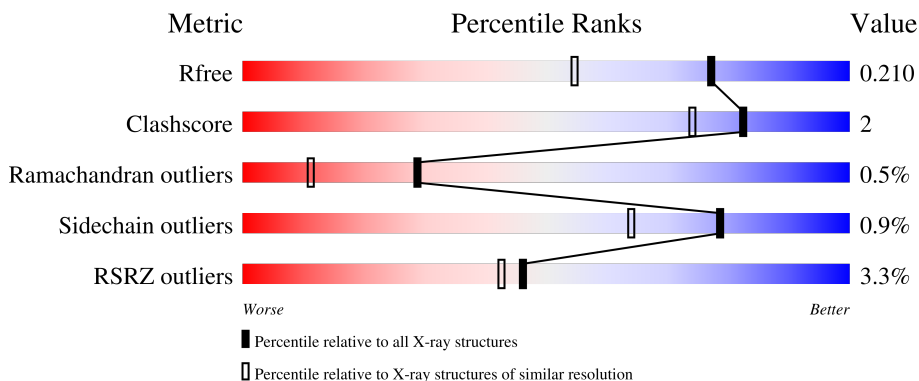
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.62 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	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (1.64-1.60)

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	632	 2% 63% 5% 32%
1	B	632	 3% 64% 5% 32%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 7472 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 Attachment glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	429	Total 3387	C 2131	N 589	O 647	S 20	0	0	0
1	B	428	Total 3379	C 2125	N 588	O 646	S 20	0	0	0

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	337	Total 337	O 337	0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	B	355	Total 355	O 355	0	0



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.20Å 83.93Å 122.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	52.38 – 1.62 69.17 – 1.62	Depositor EDS
% Data completeness (in resolution range)	99.7 (52.38-1.62) 99.7 (69.17-1.62)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 1.62Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.176 , 0.212 0.175 , 0.210	Depositor DCC
$R_{free}$ test set	5029 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.6	Xtrriage
Anisotropy	0.556	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 45.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.018 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	7472	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.33	0/3470	0.62	1/4728 (0.0%)
1	B	0.36	0/3462	0.61	0/4717
All	All	0.35	0/6932	0.61	1/9445 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	597	LEU	CA-CB-CG	5.14	127.13	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	3387	0	3278	21	0
1	B	3379	0	3268	13	0
2	A	14	0	13	0	0
3	A	337	0	0	3	1
3	B	355	0	0	1	1
All	All	7472	0	6559	33	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 2.

All (33) 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:267:MET:HE1	1:A:273:ASP:H	1.44	0.82
1:A:267:MET:HE2	1:A:270:THR:O	1.80	0.80
1:B:488:TYR:OH	1:B:490:LYS:NZ	2.22	0.72
1:A:267:MET:HE1	1:A:273:ASP:N	2.08	0.68
1:A:370:ASN:O	1:A:398:ARG:NH1	2.27	0.68
1:A:294:PRO:HG2	1:A:602:LEU:HD11	1.76	0.67
1:A:267:MET:CE	1:A:273:ASP:H	2.10	0.64
1:A:267:MET:HE3	1:A:273:ASP:HB3	1.79	0.64
1:B:429:ASP:HB3	1:B:430:PRO:CD	2.31	0.60
1:B:370:ASN:O	1:B:398:ARG:NH1	2.38	0.56
1:A:323:LEU:HD23	1:A:349:HIS:HB3	1.88	0.55
1:A:554:LEU:HG	1:A:555:ARG:HG2	1.93	0.48
1:A:487:TRP:CD2	1:A:527:PRO:HA	2.51	0.46
1:B:338:HIS:HD1	1:B:429:ASP:CG	2.18	0.46
1:A:208:ARG:NH2	3:A:814:HOH:O	2.49	0.46
1:A:456:VAL:HG22	1:A:502:VAL:HG11	1.98	0.44
1:A:365:ASN:ND2	3:A:811:HOH:O	2.47	0.43
1:B:273:ASP:OD2	1:B:276:HIS:HB2	2.17	0.43
1:B:245:ASN:OD1	1:B:614:ASN:HB3	2.18	0.43
1:B:572:HIS:CD2	1:B:572:HIS:C	2.91	0.43
1:A:229:LEU:HD11	1:A:624:TRP:CD1	2.55	0.42
1:A:267:MET:HE3	1:A:267:MET:HB3	1.77	0.42
1:A:608:THR:HB	1:A:617:PRO:HB2	2.02	0.42
1:B:419:LYS:HD2	1:B:423:TYR:CE2	2.55	0.42
1:B:487:TRP:CD2	1:B:527:PRO:HA	2.54	0.41
1:B:410:GLU:OE2	1:B:523:ARG:NH2	2.50	0.41
1:A:572:HIS:CD2	1:A:572:HIS:C	2.94	0.41
1:A:287:ARG:NH2	3:A:826:HOH:O	2.52	0.41
1:A:287:ARG:NH1	1:B:303:THR:OG1	2.53	0.41
1:A:225:GLN:HE21	1:A:227:THR:HG23	1.86	0.41
1:A:463:GLN:OE1	1:A:486:GLY:HA3	2.21	0.41
1:B:463:GLN:OE1	1:B:486:GLY:HA3	2.20	0.41
1:B:490:LYS:HE3	3:B:1009:HOH:O	2.22	0.40

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)
3:A:1108:HOH:O	3:B:806:HOH:O[1_655]	2.16	0.04

### 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	427/632 (68%)	411 (96%)	14 (3%)	2 (0%)	29	11
1	B	426/632 (67%)	407 (96%)	17 (4%)	2 (0%)	29	11
All	All	853/1264 (68%)	818 (96%)	31 (4%)	4 (0%)	29	11

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	429	ASP
1	B	430	PRO
1	A	559	THR
1	A	444	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	378/561 (67%)	374 (99%)	4 (1%)	73	56
1	B	377/561 (67%)	374 (99%)	3 (1%)	81	69
All	All	755/1122 (67%)	748 (99%)	7 (1%)	78	64

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	545	PHE
1	A	572	HIS
1	A	597	LEU
1	A	627	TYR
1	B	228	ASP
1	B	545	PHE
1	B	572	HIS

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](#)

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	$\# Z  > 2$
2	NAG	A	701	1	14,14,15	0.53	0	17,19,21	0.84	1 (5%)

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
2	NAG	A	701	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	701	NAG	C1-O5-C5	2.87	116.08	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	NAG	O5-C5-C6-O6
2	A	701	NAG	C4-C5-C6-O6

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

### 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	429/632 (67%)	-0.28	11 (2%) 56 53	18, 31, 58, 84	0
1	B	428/632 (67%)	-0.25	17 (3%) 38 34	18, 29, 50, 76	0
All	All	857/1264 (67%)	-0.27	28 (3%) 46 42	18, 30, 54, 84	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	501	GLN	8.1
1	A	448	PHE	6.8
1	B	429	ASP	5.2
1	A	447	TYR	4.8
1	B	219	THR	4.7
1	A	631	HIS	4.6
1	B	292	TRP	4.0
1	B	630	CYS	4.0
1	B	220	CYS	3.8
1	A	520	CYS	3.4
1	B	204	TYR	3.4
1	B	501	GLN	3.2
1	B	431	ASP	3.2
1	B	631	HIS	3.1
1	B	500	GLY	2.9
1	B	364	ARG	2.8
1	A	517	GLN	2.7
1	B	447	TYR	2.7
1	B	499	ASP	2.5
1	B	612	MET	2.5
1	A	499	ASP	2.5
1	B	430	PRO	2.4
1	A	569	TRP	2.3
1	A	431	ASP	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	446	GLU	2.2
1	A	399	SER	2.1
1	B	221	GLN	2.0
1	A	430	PRO	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](#)

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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	A	701	14/15	0.92	0.10	32,36,41,42	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.