



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 11:14 pm BST

PDB ID : 1WMX  
Title : Crystal Structure of Family 30 Carbohydrate Binding Module  
Authors : Horiguchi, Y.; Kono, M.; Suzuki, A.; Yamane, T.; Arai, M.; Sakka, K.; Omiya, K.  
Deposited on : 2004-07-21  
Resolution : 2.00 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](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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

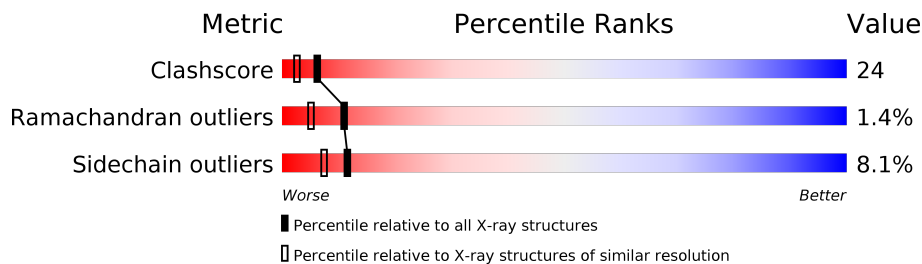
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	205	
1	B	205	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3148 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 COG3291: FOG: PKD repeat.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	173	Total 1384	C 892	N 221	O 268	S 3	7	0	0
1	B	195	Total 1556	C 1002	N 252	O 299	S 3	9	0	0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	85	Total 85	O 85	0	0

*Continued on next page...*

*Continued from previous page...*

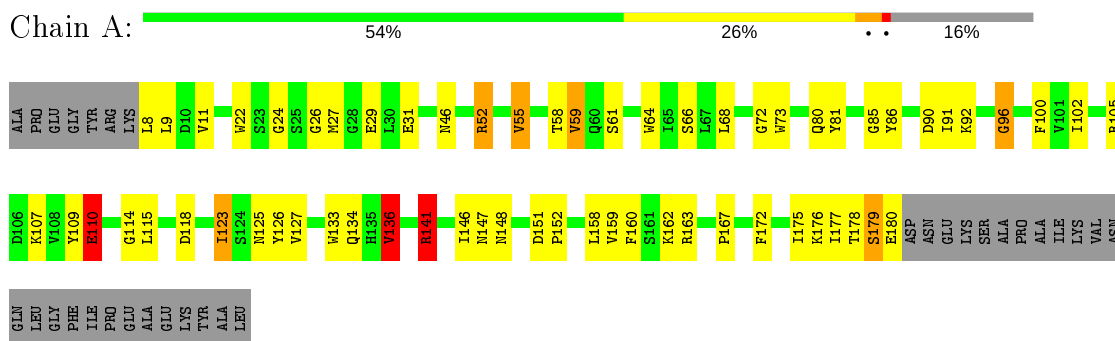
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	B	118	Total 118	O 118	0	0

### 3 Residue-property plots [i](#)

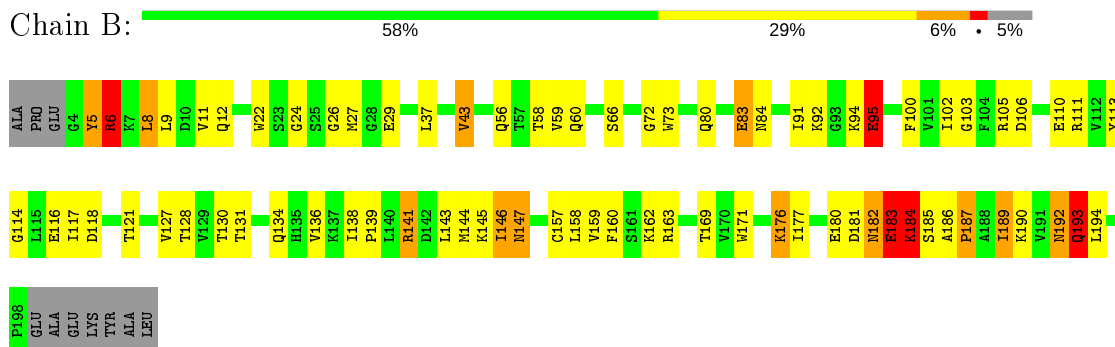
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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.

Note EDS was not executed.

- Molecule 1: COG3291: FOG: PKD repeat



- Molecule 1: COG3291: FOG: PKD repeat



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.72Å 81.86Å 86.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	22.29 – 2.00	Depositor
% Data completeness (in resolution range)	(Not available) (22.29-2.00)	Depositor
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.224 , 0.233	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	3148	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

## 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	0.83	3/1418 (0.2%)	0.96	3/1935 (0.2%)
1	B	1.17	12/1594 (0.8%)	1.36	17/2172 (0.8%)
All	All	1.02	15/3012 (0.5%)	1.19	20/4107 (0.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	B	0	3
All	All	0	6

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	95	GLU	C-N	-15.24	1.05	1.33
1	B	139	PRO	C-N	-13.87	1.02	1.34
1	B	193	GLN	C-N	-12.18	1.06	1.34
1	A	96	GLY	C-N	11.93	1.54	1.33
1	B	6	ARG	CD-NE	-9.36	1.30	1.46

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	6	ARG	NE-CZ-NH1	22.32	131.46	120.30
1	B	141	ARG	NE-CZ-NH1	16.33	128.46	120.30
1	B	141	ARG	CD-NE-CZ	13.46	142.45	123.60
1	B	6	ARG	NE-CZ-NH2	-13.26	113.67	120.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	141	ARG	NE-CZ-NH2	-9.10	115.75	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	136	VAL	Mainchain
1	A	141	ARG	Sidechain
1	A	96	GLY	Mainchain
1	B	60	GLN	Mainchain
1	B	95	GLU	Mainchain

## 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	1384	0	1350	55	0
1	B	1556	0	1522	90	1
2	B	5	0	0	0	0
3	A	85	0	0	5	0
3	B	118	0	0	3	0
All	All	3148	0	2872	141	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 24.

The worst 5 of 141 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:190:LYS:H	1:B:193:GLN:NE2	1.10	1.42
1:B:190:LYS:N	1:B:193:GLN:NE2	1.85	1.25
1:B:190:LYS:N	1:B:193:GLN:HE21	1.45	1.12
1:B:5:TYR:HA	1:B:8:LEU:HD22	1.14	1.11
1:B:5:TYR:HA	1:B:8:LEU:CD2	1.93	0.99

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-



metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:116:GLU:O	1:B:190:LYS:NZ[3_656]	1.89	0.31

### 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	171/205 (83%)	155 (91%)	15 (9%)	1 (1%)	25	19
1	B	193/205 (94%)	175 (91%)	14 (7%)	4 (2%)	7	2
All	All	364/410 (89%)	330 (91%)	29 (8%)	5 (1%)	11	5

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	146	ILE
1	B	184	LYS
1	B	147	ASN
1	B	114	GLY
1	A	114	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	157/182 (86%)	146 (93%)	11 (7%)	15	10
1	B	175/182 (96%)	159 (91%)	16 (9%)	9	5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	332/364 (91%)	305 (92%)	27 (8%)	11 7

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	8	LEU
1	B	56	GLN
1	B	189	ILE
1	B	37	LEU
1	A	110	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	54	ASN
1	B	56	GLN
1	B	148	ASN
1	B	46	ASN
1	B	125	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 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
2	SO4	B	300	-	4,4,4	0.26	0	6,6,6	0.07	0

There are no bond length outliers.

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	6
1	A	1

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	86:TYR	C	87:LEU	N	1.17
1	B	157:CYS	C	158:LEU	N	1.17
1	B	183:GLU	C	184:LYS	N	1.14
1	B	187:PRO	C	188:ALA	N	1.13
1	B	193:GLN	C	194:LEU	N	1.06

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.