



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 07:32 am BST

PDB ID : 2ZJB  
Title : Crystal structure of the human Dmc1-M200V polymorphic variant  
Authors : Hikiba, J.; Hirota, K.; Kagawa, W.; Ikawa, S.; Kinebuchi, T.; Sakane, I.; Takizawa, Y.; Yokoyama, S.; Mandon-Pepin, B.; Nicolas, A.; Shibata, T.; Ohta, K.; Kurumizaka, H.  
Deposited on : 2008-03-02  
Resolution : 3.50 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

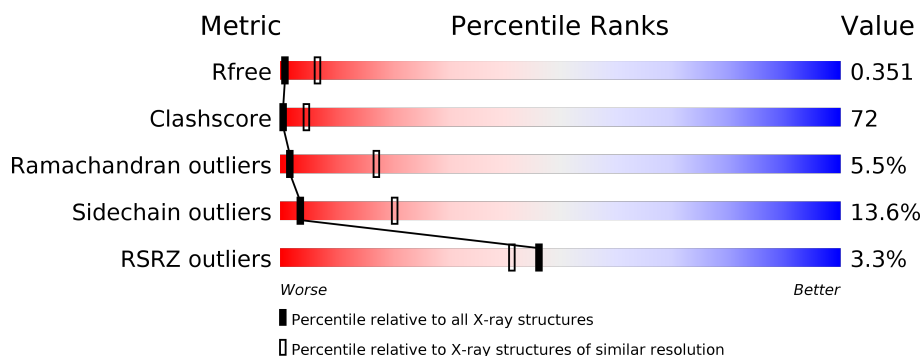
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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 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  $\leq 5\%$ . 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	343	<div> <div>15%</div> <div>44%</div> <div>9%</div> <div>30%</div> </div>
1	B	343	<div> <div>3%</div> <div>18%</div> <div>43%</div> <div>8%</div> <div>30%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3748 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.

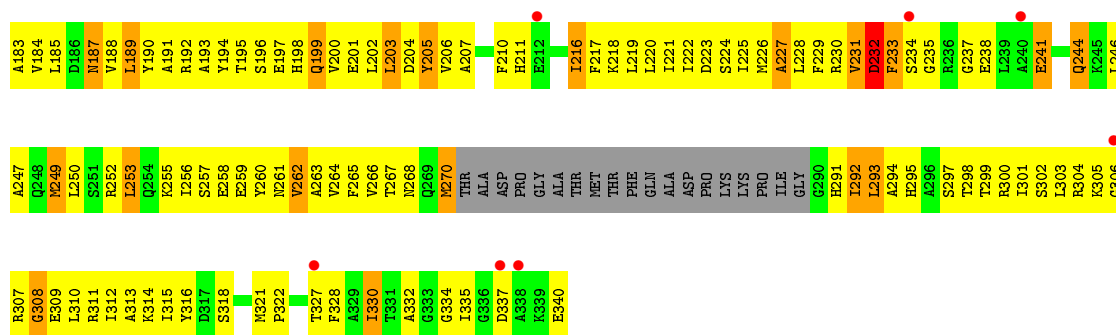
- Molecule 1 is a protein called Meiotic recombination protein DMC1/LIM15 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	239	Total	C	N	O	S	0	0	0
			1874	1186	325	356	7			
1	B	239	Total	C	N	O	S	0	0	0
			1874	1186	325	356	7			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP Q14565
A	-1	SER	-	EXPRESSION TAG	UNP Q14565
A	0	HIS	-	EXPRESSION TAG	UNP Q14565
A	200	VAL	MET	SEE REMARK 999	UNP Q14565
B	-2	GLY	-	EXPRESSION TAG	UNP Q14565
B	-1	SER	-	EXPRESSION TAG	UNP Q14565
B	0	HIS	-	EXPRESSION TAG	UNP Q14565
B	200	VAL	MET	SEE REMARK 999	UNP Q14565





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.09 Å   124.09 Å   217.87 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	20.00 – 3.50 46.28 – 3.48	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-3.50) 99.9 (46.28-3.48)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.53 (at 3.48 Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.294   ,   0.351 0.293   ,   0.351	Depositor DCC
$R_{free}$ test set	1138 reflections (10.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	92.9	Xtriage
Anisotropy	0.978	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 140.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	3748	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	119.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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 ⓘ

### 5.1 Standard geometry ⓘ

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	0.42	0/1907	0.66	0/2564
1	B	0.41	0/1907	0.65	0/2564
All	All	0.41	0/3814	0.66	0/5128

There are no bond length outliers.

There are no bond angle outliers.

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	1874	0	1856	285	0
1	B	1874	0	1856	265	0
All	All	3748	0	3712	537	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 72.

The worst 5 of 537 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:247:ALA:HB2	1:A:292:ILE:HD11	1.24	1.11
1:B:304:ARG:HG3	1:B:312:ILE:HB	1.45	0.99

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:141:VAL:O	1:A:145:LEU:HD22	1.63	0.98
1:A:119:MET:HG2	1:A:258:GLU:HA	1.45	0.96
1:A:247:ALA:HB2	1:A:292:ILE:CD1	1.96	0.95

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	A	235/343 (68%)	184 (78%)	36 (15%)	15 (6%)	1	14
1	B	235/343 (68%)	188 (80%)	36 (15%)	11 (5%)	2	20
All	All	470/686 (68%)	372 (79%)	72 (15%)	26 (6%)	2	17

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	148	ALA
1	A	226	MET
1	A	227	ALA
1	B	227	ALA
1	A	95	ARG

### 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	195/282 (69%)	167 (86%)	28 (14%)	3	18
1	B	195/282 (69%)	170 (87%)	25 (13%)	4	22
All	All	390/564 (69%)	337 (86%)	53 (14%)	3	20

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	249	MET
1	B	108	PHE
1	B	270	MET
1	A	250	LEU
1	A	293	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	291	HIS
1	B	291	HIS
1	B	198	HIS
1	A	187	ASN
1	B	181	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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 [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, 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	239/343 (69%)	-0.08	5 (2%) 63 58	32, 111, 173, 195	0
1	B	239/343 (69%)	0.05	11 (4%) 32 28	76, 119, 182, 198	0
All	All	478/686 (69%)	-0.02	16 (3%) 46 41	32, 114, 176, 198	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	340	GLU	5.6
1	B	327	THR	4.0
1	B	338	ALA	3.3
1	B	306	GLY	2.9
1	B	89	PHE	2.9

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.