



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 07:37 pm BST

PDB ID : 5Y09  
Title : Crystal structure of Deg9 at 295 K  
Authors : Ouyang, M.; Li, X.Y.; Liu, L.; Zhang, L.X.  
Deposited on : 2017-07-15  
Resolution : 2.45 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 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



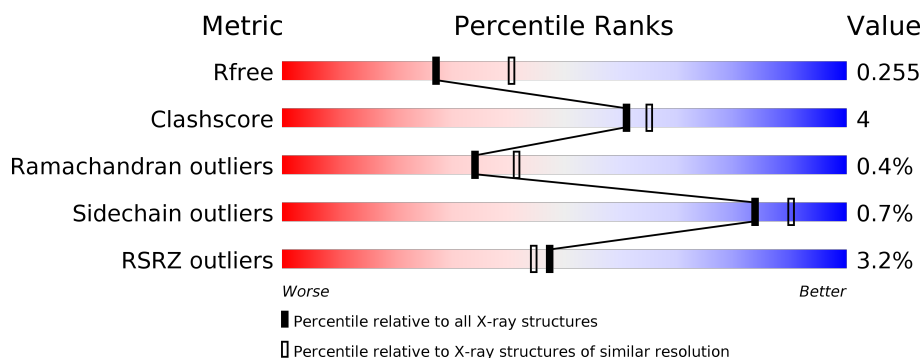
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)
RSRZ outliers	127900	1547 (2.46-2.42)

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	566	<div> <div>%</div> <div> <div></div> <div>73%</div> <div>8%</div> <div>18%</div> </div> </div>
1	B	566	<div> <div>4%</div> <div> <div></div> <div>57%</div> <div>8%</div> <div>35%</div> </div> </div>



## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 6685 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 Protease Do-like 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	465	Total	C	N	O	S	0	1	0
			3618	2307	606	691	14			
1	B	369	Total	C	N	O	S	0	0	0
			2893	1863	483	535	12			

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	MET	-	expression tag	UNP Q9FL12
A	28	GLY	-	expression tag	UNP Q9FL12
A	29	SER	-	expression tag	UNP Q9FL12
A	30	SER	-	expression tag	UNP Q9FL12
A	31	HIS	-	expression tag	UNP Q9FL12
A	32	HIS	-	expression tag	UNP Q9FL12
A	33	HIS	-	expression tag	UNP Q9FL12
A	34	HIS	-	expression tag	UNP Q9FL12
A	35	HIS	-	expression tag	UNP Q9FL12
A	36	HIS	-	expression tag	UNP Q9FL12
A	37	SER	-	expression tag	UNP Q9FL12
A	38	SER	-	expression tag	UNP Q9FL12
A	39	GLY	-	expression tag	UNP Q9FL12
A	40	LEU	-	expression tag	UNP Q9FL12
A	41	VAL	-	expression tag	UNP Q9FL12
A	42	PRO	-	expression tag	UNP Q9FL12
A	43	ARG	-	expression tag	UNP Q9FL12
A	44	GLY	-	expression tag	UNP Q9FL12
A	45	SER	-	expression tag	UNP Q9FL12
A	46	HIS	-	expression tag	UNP Q9FL12
A	47	MET	-	expression tag	UNP Q9FL12
A	48	ALA	-	expression tag	UNP Q9FL12
A	49	SER	-	expression tag	UNP Q9FL12
A	50	MET	-	expression tag	UNP Q9FL12
A	51	THR	-	expression tag	UNP Q9FL12

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	52	GLY	-	expression tag	UNP Q9FL12
A	53	GLY	-	expression tag	UNP Q9FL12
A	54	GLN	-	expression tag	UNP Q9FL12
A	55	GLN	-	expression tag	UNP Q9FL12
A	56	MET	-	expression tag	UNP Q9FL12
A	57	GLY	-	expression tag	UNP Q9FL12
A	58	ARG	-	expression tag	UNP Q9FL12
A	59	GLY	-	expression tag	UNP Q9FL12
A	60	SER	-	expression tag	UNP Q9FL12
A	61	GLU	-	expression tag	UNP Q9FL12
A	62	PHE	-	expression tag	UNP Q9FL12
A	63	GLU	-	expression tag	UNP Q9FL12
A	64	LEU	-	expression tag	UNP Q9FL12
B	27	MET	-	expression tag	UNP Q9FL12
B	28	GLY	-	expression tag	UNP Q9FL12
B	29	SER	-	expression tag	UNP Q9FL12
B	30	SER	-	expression tag	UNP Q9FL12
B	31	HIS	-	expression tag	UNP Q9FL12
B	32	HIS	-	expression tag	UNP Q9FL12
B	33	HIS	-	expression tag	UNP Q9FL12
B	34	HIS	-	expression tag	UNP Q9FL12
B	35	HIS	-	expression tag	UNP Q9FL12
B	36	HIS	-	expression tag	UNP Q9FL12
B	37	SER	-	expression tag	UNP Q9FL12
B	38	SER	-	expression tag	UNP Q9FL12
B	39	GLY	-	expression tag	UNP Q9FL12
B	40	LEU	-	expression tag	UNP Q9FL12
B	41	VAL	-	expression tag	UNP Q9FL12
B	42	PRO	-	expression tag	UNP Q9FL12
B	43	ARG	-	expression tag	UNP Q9FL12
B	44	GLY	-	expression tag	UNP Q9FL12
B	45	SER	-	expression tag	UNP Q9FL12
B	46	HIS	-	expression tag	UNP Q9FL12
B	47	MET	-	expression tag	UNP Q9FL12
B	48	ALA	-	expression tag	UNP Q9FL12
B	49	SER	-	expression tag	UNP Q9FL12
B	50	MET	-	expression tag	UNP Q9FL12
B	51	THR	-	expression tag	UNP Q9FL12
B	52	GLY	-	expression tag	UNP Q9FL12
B	53	GLY	-	expression tag	UNP Q9FL12
B	54	GLN	-	expression tag	UNP Q9FL12
B	55	GLN	-	expression tag	UNP Q9FL12

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	56	MET	-	expression tag	UNP Q9FL12
B	57	GLY	-	expression tag	UNP Q9FL12
B	58	ARG	-	expression tag	UNP Q9FL12
B	59	GLY	-	expression tag	UNP Q9FL12
B	60	SER	-	expression tag	UNP Q9FL12
B	61	GLU	-	expression tag	UNP Q9FL12
B	62	PHE	-	expression tag	UNP Q9FL12
B	63	GLU	-	expression tag	UNP Q9FL12
B	64	LEU	-	expression tag	UNP Q9FL12

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	103	Total O 103 103	0	0
2	B	71	Total O 71 71	0	0







## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.37Å 132.37Å 154.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.89 – 2.45 38.89 – 2.45	Depositor EDS
% Data completeness (in resolution range)	95.6 (38.89-2.45) 95.6 (38.89-2.45)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.10 (at 2.45Å)	Xtriage
Refinement program	PHENIX 1.8.1 _1168	Depositor
R, $R_{free}$	0.206 , 0.253 0.208 , 0.255	Depositor DCC
$R_{free}$ test set	2380 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.6	Xtriage
Anisotropy	0.099	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 5.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.129 for -h,k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6685	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% 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	0.27	0/3695	0.46	0/5012
1	B	0.27	0/2953	0.45	0/3999
All	All	0.27	0/6648	0.46	0/9011

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	3618	0	3605	28	0
1	B	2893	0	2901	30	0
2	A	103	0	0	0	0
2	B	71	0	0	0	0
All	All	6685	0	6506	58	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 4.

All (58) 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:163:ARG:HG2	1:A:203:LEU:HD11	1.65	0.77

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:334:GLY:N	1:A:369:GLU:OE1	2.18	0.77
1:A:368:PRO:HA	1:A:370:SER:N	2.00	0.76
1:B:549:LYS:HG2	1:B:559:VAL:HG22	1.73	0.70
1:A:368:PRO:HA	1:A:370:SER:H	1.56	0.70
1:B:281:PRO:HB2	1:B:282:ALA:HB2	1.72	0.70
1:A:350:GLY:HA3	1:A:424:LYS:HG2	1.75	0.68
1:B:146:ARG:HH22	1:B:448:PRO:HB3	1.59	0.66
1:B:281:PRO:HB3	1:B:293:ALA:H	1.60	0.66
1:A:367:ALA:O	1:A:370:SER:N	2.29	0.64
1:A:274:ASN:ND2	1:A:275:SER:O	2.34	0.58
1:B:396:ARG:HG2	1:B:399:GLU:HB2	1.85	0.58
1:A:356:LYS:HD3	1:A:388:ALA:HA	1.87	0.56
1:B:333:LEU:HB2	1:B:407:ILE:HD13	1.88	0.56
1:A:223:LEU:HD12	1:A:224:PRO:HD2	1.89	0.55
1:A:367:ALA:O	1:A:369:GLU:HB3	2.07	0.55
1:A:524:ALA:HB3	1:A:551:ASN:HB2	1.89	0.54
1:B:327:TYR:CZ	1:B:329:GLY:HA2	2.44	0.52
1:B:167:ASN:ND2	1:B:279:GLY:O	2.43	0.52
1:A:254:ILE:HG22	1:A:263:GLU:HB3	1.92	0.51
1:B:157:PHE:CE2	1:B:281:PRO:HG2	2.46	0.51
1:B:292:ILE:N	1:B:309:ILE:O	2.44	0.50
1:B:155:SER:HB2	1:B:280:GLY:HA3	1.93	0.50
1:A:266:GLY:HA2	1:A:310:PRO:HD3	1.94	0.50
1:A:248:VAL:H	1:A:270:ASP:HB3	1.76	0.49
1:B:350:GLY:HA3	1:B:424:LYS:HG2	1.95	0.49
1:B:265:LEU:HD13	1:B:401:ILE:HD11	1.94	0.49
1:A:363:ILE:HG21	1:A:370:SER:HA	1.95	0.48
1:B:129:VAL:N	1:B:156:GLY:O	2.47	0.48
1:A:327:TYR:CZ	1:A:329:GLY:HA2	2.49	0.48
1:A:382:PHE:CZ	1:A:407:ILE:HG12	2.49	0.48
1:B:157:PHE:HE2	1:B:281:PRO:HG2	1.78	0.48
1:B:310:PRO:HB2	1:B:312:PRO:HD2	1.96	0.47
1:A:237:ILE:H	1:A:237:ILE:HD12	1.80	0.46
1:A:501:VAL:HG12	1:A:521:GLN:HA	1.96	0.46
1:A:269:ILE:HD12	1:A:271:ALA:HB3	1.98	0.46
1:B:281:PRO:CB	1:B:292:ILE:HA	2.48	0.44
1:B:356:LYS:HE3	1:B:356:LYS:HB2	1.79	0.44
1:B:569:THR:O	1:B:573:LEU:HG	2.17	0.44
1:B:281:PRO:HB3	1:B:293:ALA:N	2.29	0.44
1:B:138:PRO:HA	1:B:146:ARG:HD3	2.00	0.43
1:B:311:THR:N	1:B:312:PRO:HD2	2.33	0.43

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:528:LYS:HA	1:B:529:PRO:HD3	1.86	0.43
1:A:335:ILE:HG22	1:A:363:ILE:HG12	2.00	0.43
1:B:292:ILE:HB	1:B:309:ILE:HB	2.00	0.43
1:A:471:TYR:CZ	1:A:480:PRO:HD3	2.54	0.42
1:A:310:PRO:HB2	1:A:312:PRO:HD2	2.01	0.42
1:A:146:ARG:NH2	1:A:448:PRO:HB3	2.35	0.42
1:B:178:LYS:HA	1:B:188:TYR:O	2.20	0.42
1:A:135:HIS:HB2	1:A:150:TYR:CE1	2.55	0.41
1:B:549:LYS:HE2	1:B:549:LYS:HB3	1.83	0.41
1:B:432:LYS:HE2	1:B:432:LYS:HB3	1.91	0.41
1:A:158:ILE:HG23	1:A:212:TRP:HZ3	1.85	0.41
1:B:364:GLU:O	1:B:370:SER:HB2	2.21	0.41
1:A:169:HIS:O	1:A:172:GLU:HG2	2.21	0.41
1:A:578:ILE:HA	1:A:579:PRO:HD3	1.89	0.41
1:B:281:PRO:HB3	1:B:292:ILE:HA	2.03	0.41
1:B:471:TYR:CZ	1:B:480:PRO:HD3	2.57	0.40

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	464/566 (82%)	453 (98%)	10 (2%)	1 (0%)	47	57
1	B	357/566 (63%)	345 (97%)	10 (3%)	2 (1%)	25	29
All	All	821/1132 (72%)	798 (97%)	20 (2%)	3 (0%)	34	41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	171	VAL
1	B	281	PRO

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	368	PRO

### 5.3.2 Protein sidechains ⓘ

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	402/490 (82%)	399 (99%)	3 (1%)	84	90
1	B	320/490 (65%)	318 (99%)	2 (1%)	86	91
All	All	722/980 (74%)	717 (99%)	5 (1%)	84	90

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	246	SER
1	A	356	LYS
1	A	405	TYR
1	B	163	ARG
1	B	415	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	465/566 (82%)	-0.06	3 (0%) 89 89	10, 23, 46, 82	0
1	B	369/566 (65%)	0.32	24 (6%) 18 15	9, 26, 60, 75	1 (0%)
All	All	834/1132 (73%)	0.11	27 (3%) 47 44	9, 24, 56, 82	1 (0%)

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	282	ALA	6.0
1	B	290	VAL	5.8
1	B	160	GLY	5.2
1	B	130	LYS	4.8
1	B	132	PHE	4.5
1	A	123	PRO	4.3
1	B	159	ILE	4.0
1	B	281	PRO	3.9
1	B	165	LEU	3.8
1	B	283	PHE	3.8
1	B	158	ILE	3.7
1	B	179	LEU	3.7
1	B	294	PHE	3.7
1	B	157	PHE	3.2
1	B	131	VAL	3.2
1	B	163	ARG	3.0
1	B	292	ILE	2.9
1	B	280	GLY	2.9
1	A	271	ALA	2.7
1	B	177	VAL	2.5
1	B	269	ILE	2.5
1	B	178	LYS	2.4
1	A	300	GLU	2.4
1	B	190	ALA	2.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	189	LEU	2.2
1	B	446	GLY	2.1
1	B	323	LYS	2.1

## 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.