



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 08:51 am BST

PDB ID : 2HWY  
Title : Structure of PIN domain of human SMG5.  
Authors : Glavan, F.; Behm-Ansmant, I.; Izaurralde, E.; Conti, E.  
Deposited on : 2006-08-02  
Resolution : 2.75 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

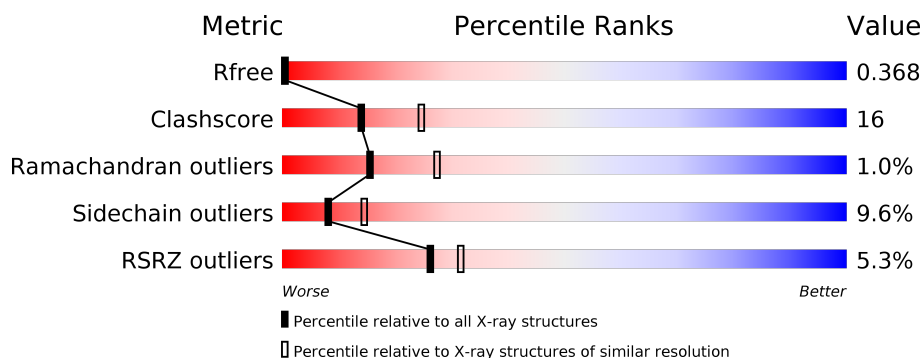
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	164	<div> <div>2%</div> <div> <div></div> <div>48%</div> <div>19%</div> <div>•</div> <div>29%</div> </div> </div>
1	B	164	<div> <div>5%</div> <div> <div></div> <div>45%</div> <div>18%</div> <div>• •</div> <div>32%</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1816 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.

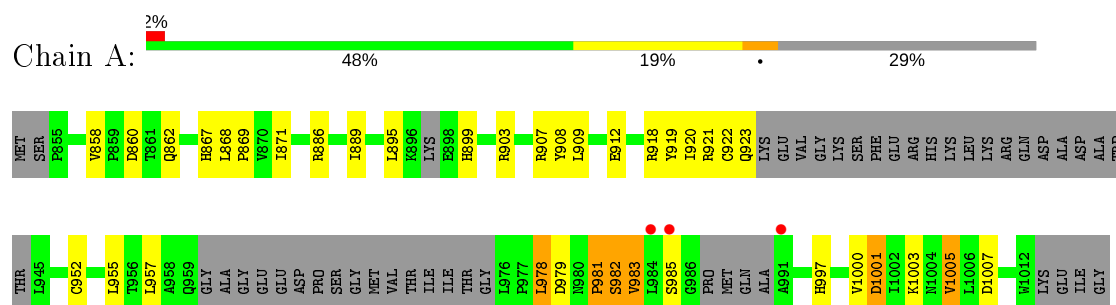
- Molecule 1 is a protein called Protein SMG5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	116	Total	C	N	O	S	76	0	0
			921	593	163	162	3			
1	B	111	Total	C	N	O	S	91	0	0
			895	575	159	158	3			

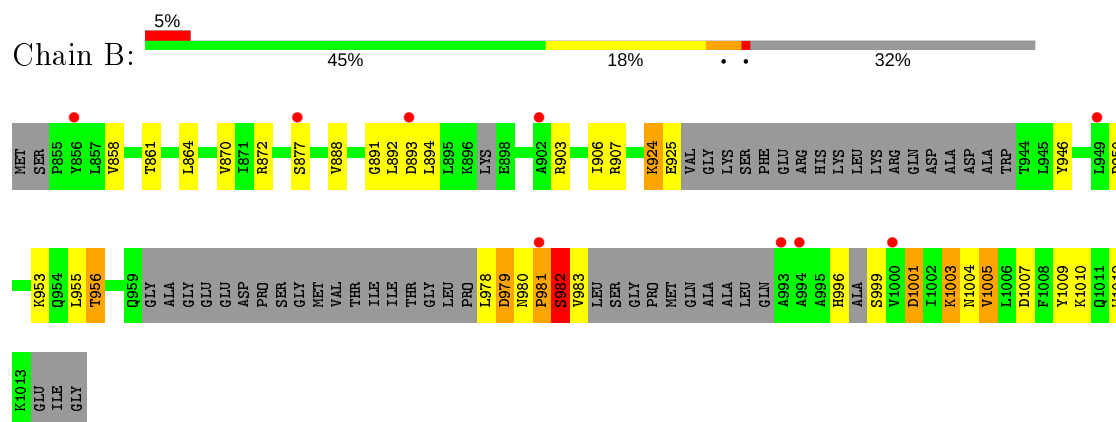
### 3 Residue-property plots [i](#)

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 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: Protein SMG5



#### • Molecule 1: Protein SMG5



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	45.59Å 42.53Å 65.10Å 90.00° 95.98° 90.00°	Depositor
Resolution (Å)	30.00 – 2.75 27.75 – 2.75	Depositor EDS
% Data completeness (in resolution range)	98.8 (30.00-2.75) 98.8 (27.75-2.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.78 (at 2.76Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.271 , 0.329 0.298 , 0.368	Depositor DCC
$R_{free}$ test set	495 reflections (7.56%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.8	Xtriage
Anisotropy	0.148	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 54.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	1816	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.33% 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	0.68	1/937 (0.1%)	0.82	5/1265 (0.4%)
1	B	0.52	0/909	0.73	4/1222 (0.3%)
All	All	0.61	1/1846 (0.1%)	0.78	9/2487 (0.4%)

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

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	922	CYS	CB-SG	-5.57	1.72	1.81

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	982	SER	CB-CA-C	-11.18	88.86	110.10
1	A	983	VAL	N-CA-C	6.06	127.36	111.00
1	A	983	VAL	N-CA-CB	5.93	124.56	111.50
1	B	999	SER	N-CA-C	5.35	125.45	111.00
1	B	982	SER	C-N-CA	5.26	134.86	121.70
1	B	950	ASP	CB-CG-OD2	5.24	123.02	118.30
1	A	1001	ASP	CB-CG-OD2	5.21	122.99	118.30
1	B	1001	ASP	CB-CG-OD2	5.19	122.97	118.30
1	A	1007	ASP	CB-CG-OD2	5.16	122.94	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	982	SER	Mainchain
1	A	985	SER	Peptide
1	A	997	HIS	Peptide

## 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	921	0	946	19	3
1	B	895	0	918	32	3
All	All	1816	0	1864	51	5

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 16.

All (51) 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:982:SER:O	1:B:983:VAL:HG22	1.33	1.22
1:B:924:LYS:O	1:B:925:GLU:HB2	1.37	1.08
1:B:982:SER:HA	1:B:1004:ASN:ND2	1.69	1.07
1:B:858:VAL:HG21	1:B:955:LEU:HD21	1.01	1.00
1:B:858:VAL:HG21	1:B:955:LEU:CD2	1.90	1.00
1:B:858:VAL:CG2	1:B:955:LEU:HD21	1.97	0.92
1:A:955:LEU:HD13	1:A:1000:VAL:HG11	1.54	0.90
1:B:924:LYS:O	1:B:925:GLU:CB	2.18	0.89
1:B:982:SER:HA	1:B:1004:ASN:HD21	1.34	0.86
1:B:983:VAL:H	1:B:1004:ASN:HD22	1.24	0.82
1:B:982:SER:O	1:B:983:VAL:CG2	2.24	0.81
1:B:983:VAL:H	1:B:1004:ASN:ND2	1.78	0.79
1:B:980:ASN:HB3	1:B:1005:VAL:HG13	1.70	0.72
1:B:982:SER:CA	1:B:1004:ASN:HD21	2.02	0.71
1:A:858:VAL:HG11	1:A:955:LEU:HD21	1.72	0.70
1:A:983:VAL:O	1:A:983:VAL:HG12	1.91	0.70
1:B:982:SER:CA	1:B:1004:ASN:ND2	2.52	0.70
1:B:864:LEU:HD13	1:B:892:LEU:HD11	1.75	0.69
1:B:978:LEU:CD2	1:B:1001:ASP:HB2	2.30	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:983:VAL:N	1:B:1004:ASN:ND2	2.46	0.61
1:B:978:LEU:HD23	1:B:1001:ASP:HB2	1.83	0.61
1:A:909:LEU:CD2	1:A:920:ILE:HD11	2.34	0.58
1:A:978:LEU:HD12	1:A:979:ASP:N	2.19	0.57
1:A:895:LEU:HD22	1:A:899:HIS:CE1	2.40	0.57
1:B:893:ASP:OD1	1:B:906:ILE:HG13	2.04	0.56
1:A:981:PRO:O	1:A:1005:VAL:HG22	2.05	0.56
1:B:979:ASP:N	1:B:979:ASP:OD2	2.41	0.54
1:A:858:VAL:HG11	1:A:955:LEU:CD2	2.36	0.53
1:A:903:ARG:O	1:A:907:ARG:HG3	2.10	0.52
1:A:858:VAL:HG21	1:A:955:LEU:HD21	1.94	0.50
1:B:980:ASN:CB	1:B:1005:VAL:HG13	2.41	0.50
1:A:858:VAL:CG1	1:A:955:LEU:HD21	2.40	0.50
1:A:858:VAL:HG23	1:A:858:VAL:O	2.13	0.48
1:B:983:VAL:N	1:B:1004:ASN:HD22	2.02	0.47
1:B:982:SER:C	1:B:983:VAL:HG22	2.24	0.46
1:B:953:LYS:O	1:B:956:THR:HG22	2.16	0.45
1:B:981:PRO:HD2	1:B:1003:LYS:O	2.17	0.44
1:A:908:TYR:CZ	1:A:912:GLU:HG3	2.53	0.44
1:A:868:LEU:N	1:A:869:PRO:CD	2.80	0.44
1:A:895:LEU:HD22	1:A:899:HIS:ND1	2.32	0.44
1:A:860:ASP:CB	1:A:981:PRO:HA	2.48	0.43
1:A:867:HIS:O	1:A:871:ILE:HD12	2.19	0.43
1:B:903:ARG:O	1:B:907:ARG:HG3	2.19	0.43
1:B:891:GLY:O	1:B:894:LEU:HB2	2.19	0.42
1:A:868:LEU:HD11	1:A:908:TYR:CD2	2.55	0.41
1:B:888:VAL:O	1:B:892:LEU:HG	2.20	0.41
1:A:981:PRO:HD2	1:A:1003:LYS:O	2.20	0.41
1:B:864:LEU:CD1	1:B:892:LEU:HD11	2.48	0.40
1:B:877:SER:HB3	1:B:1012:TRP:CH2	2.55	0.40
1:B:877:SER:HB2	1:B:1012:TRP:CZ2	2.56	0.40
1:B:870:VAL:CG1	1:B:1009:TYR:HB2	2.51	0.40

All (5) 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)
1:A:919:TYR:OH	1:A:957:LEU:CD2[2_556]	1.07	1.13
1:B:996:HIS:CD2	1:B:1010:LYS:NZ[2_655]	1.54	0.66
1:A:919:TYR:CZ	1:A:957:LEU:CD2[2_556]	2.04	0.16

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1001:ASP:OD2	1:B:877:SER:O[1_455]	2.10	0.10
1:B:996:HIS:NE2	1:B:1010:LYS:NZ[2_655]	2.17	0.03

## 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	106/164 (65%)	97 (92%)	8 (8%)	1 (1%)	17	31
1	B	99/164 (60%)	93 (94%)	5 (5%)	1 (1%)	15	27
All	All	205/328 (62%)	190 (93%)	13 (6%)	2 (1%)	15	27

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	981	PRO
1	B	981	PRO

### 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	100/138 (72%)	91 (91%)	9 (9%)	9	16
1	B	98/138 (71%)	88 (90%)	10 (10%)	7	12
All	All	198/276 (72%)	179 (90%)	19 (10%)	8	14

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	862	GLN
1	A	886	ARG
1	A	889	ILE
1	A	918	ARG
1	A	921	ARG
1	A	923	GLN
1	A	952	CYS
1	A	978	LEU
1	A	1005	VAL
1	B	861	THR
1	B	872	ARG
1	B	924	LYS
1	B	946	TYR
1	B	956	THR
1	B	979	ASP
1	B	982	SER
1	B	1003	LYS
1	B	1005	VAL
1	B	1007	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	862	GLN
1	B	1004	ASN

### 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 [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	116/164 (70%)	0.39	3 (2%) 56 65	20, 25, 32, 37	20 (17%)
1	B	111/164 (67%)	0.66	9 (8%) 12 14	20, 26, 37, 45	23 (20%)
All	All	227/328 (69%)	0.52	12 (5%) 26 31	20, 25, 34, 45	43 (18%)

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	993	ALA	4.0
1	B	856	TYR	3.3
1	B	994	ALA	3.0
1	B	902	ALA	2.6
1	B	949	LEU	2.6
1	A	985	SER	2.4
1	A	984	LEU	2.4
1	B	893	ASP	2.3
1	B	981	PRO	2.2
1	B	877	SER	2.1
1	A	991	ALA	2.1
1	B	1000	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 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.