



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 14, 2019 – 11:09 AM EDT

PDB ID : 2V5M  
Title : Structural basis for Dscam isoform specificity  
Authors : Meijers, R.; Puettmann-Holgado, R.; Skiniotis, G.; Liu, J.-H.; Walz, T.;  
Schmucker, D.; Wang, J.-H.  
Deposited on : 2007-07-06  
Resolution : 1.95 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 : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.4  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

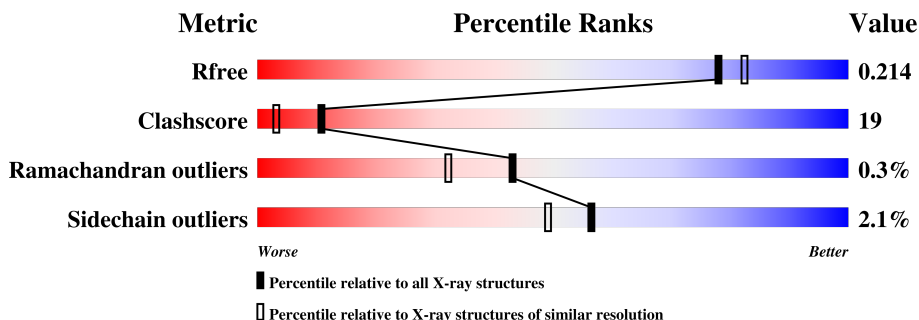
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 1.95 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))
$R_{free}$	111664	2220 (1.96-1.96)
Clashscore	122126	2333 (1.96-1.96)
Ramachandran outliers	120053	2314 (1.96-1.96)
Sidechain outliers	120020	2314 (1.96-1.96)

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. 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	388	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	601	-	X	-	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3915 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 DSCAM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	0	25	0
			3144	1980	545	606	13			

There are 20 discrepancies between the modelled and reference sequences:

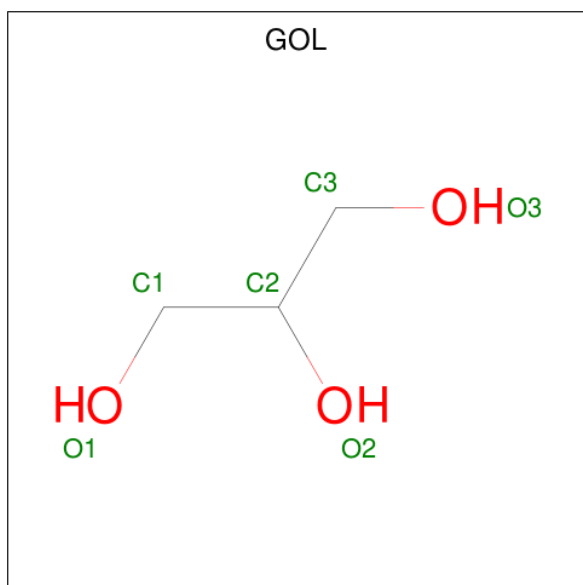
Chain	Residue	Modelled	Actual	Comment	Reference
A	207	ILE	VAL	variant	UNP Q9NBA1
A	211	ALA	PRO	variant	UNP Q9NBA1
A	213	ARG	LYS	variant	UNP Q9NBA1
A	214	THR	ILE	variant	UNP Q9NBA1
A	215	PRO	ASN	variant	UNP Q9NBA1
A	216	ALA	THR	variant	UNP Q9NBA1
A	218	VAL	THR	variant	UNP Q9NBA1
A	219	GLN	TYR	variant	UNP Q9NBA1
A	222	LEU	ASN	variant	UNP Q9NBA1
A	223	GLU	ILE	variant	UNP Q9NBA1
A	224	LEU	VAL	variant	UNP Q9NBA1
A	225	MET	GLU	variant	UNP Q9NBA1
A	226	VAL	SER	variant	UNP Q9NBA1
A	227	ALA	MET	variant	UNP Q9NBA1
A	228	HIS	ALA	variant	UNP Q9NBA1
A	229	THR	SER	variant	UNP Q9NBA1
A	230	ILE	THR	variant	UNP Q9NBA1
A	231	SER	ALA	variant	UNP Q9NBA1
A	232	LEU	ILE	variant	UNP Q9NBA1
A	239	PHE	TYR	variant	UNP Q9NBA1

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	709	Total	O	0	0
			709	709		

### 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.

- Molecule 1: DSCAM



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.17Å 99.17Å 163.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.95 29.71 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.6 (20.00-1.95) 98.6 (29.71-1.95)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.169 , 0.205 0.178 , 0.214	Depositor DCC
$R_{free}$ test set	3006 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.0	Xtriage
Anisotropy	0.377	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 50.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3915	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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: GOL, 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	0.67	2/3291 (0.1%)	1.35	24/4458 (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	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	386	LYS	C-O	5.77	1.34	1.23
1	A	387	LEU	C-O	5.43	1.33	1.23

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	193	ARG	NE-CZ-NH2	-22.43	109.08	120.30
1	A	193	ARG	NE-CZ-NH1	18.34	129.47	120.30
1	A	198	LYS	CD-CE-NZ	8.24	130.65	111.70
1	A	72	ARG	NE-CZ-NH1	-7.75	116.42	120.30
1	A	85	ARG	NE-CZ-NH2	-7.68	116.46	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	58[A]	ASP	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	3144	0	3111	120	0
2	A	56	0	50	3	0
3	A	6	0	6	1	0
4	A	709	0	0	66	6
All	All	3915	0	3167	122	6

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 19.

The worst 5 of 122 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:115:VAL:HB	4:A:2284:HOH:O	1.26	1.26
1:A:14:ILE:HG23	4:A:2037:HOH:O	1.50	1.10
1:A:235:PRO:CD	4:A:2482:HOH:O	2.00	1.07
1:A:57[B]:SER:O	1:A:58[B]:ASP:HB2	1.52	1.07
1:A:235:PRO:CG	4:A:2482:HOH:O	2.04	1.05

The worst 5 of 6 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)
4:A:2185:HOH:O	4:A:2211:HOH:O[8_665]	1.50	0.70
4:A:2400:HOH:O	4:A:2400:HOH:O[2_765]	1.74	0.46
4:A:2029:HOH:O	4:A:2450:HOH:O[5_755]	1.82	0.38
4:A:2429:HOH:O	4:A:2429:HOH:O[6_565]	2.00	0.20
4:A:2279:HOH:O	4:A:2316:HOH:O[5_755]	2.02	0.18

## 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	411/388 (106%)	398 (97%)	11 (3%)	2 (0%)	31 18

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	58[A]	ASP
1	A	58[B]	ASP

### 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	356/332 (107%)	349 (98%)	7 (2%)	58 51

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	193	ARG
1	A	353	SER
1	A	251	GLU
1	A	48	ASP
1	A	319	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	97	HIS
1	A	147	ASN
1	A	111	ASN
1	A	31	ASN
1	A	141	HIS

### 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](#)

5 ligands are 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	401	1,2	14,14,15	0.87	0	17,19,21	3.13	10 (58%)
2	NAG	A	402	2	14,14,15	0.59	0	17,19,21	2.31	8 (47%)
2	NAG	A	501	1,2	14,14,15	0.41	0	17,19,21	2.65	7 (41%)
2	NAG	A	502	2	14,14,15	0.50	0	17,19,21	1.24	2 (11%)
3	GOL	A	601	-	5,5,5	0.49	0	5,5,5	5.63	3 (60%)

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	401	1,2	-	4/6/23/26	0/1/1/1
2	NAG	A	402	2	-	2/6/23/26	0/1/1/1
2	NAG	A	501	1,2	-	2/6/23/26	0/1/1/1
2	NAG	A	502	2	-	0/6/23/26	0/1/1/1
3	GOL	A	601	-	-	4/4/4/4	-

There are no bond length outliers.

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	GOL	C3-C2-C1	-11.52	66.86	111.75
2	A	401	NAG	O6-C6-C5	-6.77	87.82	111.29
2	A	501	NAG	C1-O5-C5	5.76	120.03	112.20
2	A	401	NAG	C3-C4-C5	5.50	120.12	110.23
2	A	501	NAG	C6-C5-C4	-5.23	100.72	113.00

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	601	GOL	C1-C2-C3-O3
2	A	501	NAG	O5-C5-C6-O6
2	A	401	NAG	C4-C5-C6-O6
2	A	501	NAG	C4-C5-C6-O6
2	A	401	NAG	C8-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	NAG	2	0
2	A	502	NAG	1	0
3	A	601	GOL	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	55:ILE	C	56[B]:SER	N	2.25

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.