



# wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 03:56 PM GMT

PDB ID : 4B04  
Title : Crystal structure of the Catalytic Domain of Human DUSP26 (C152S)  
Authors : Won, E.-Y.; Lee, D.Y.; Park, S.G.; Yokoyama, S.; Kim, S.J.; Chi, S.-W.  
Deposited on : 2012-06-28  
Resolution : 2.21 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

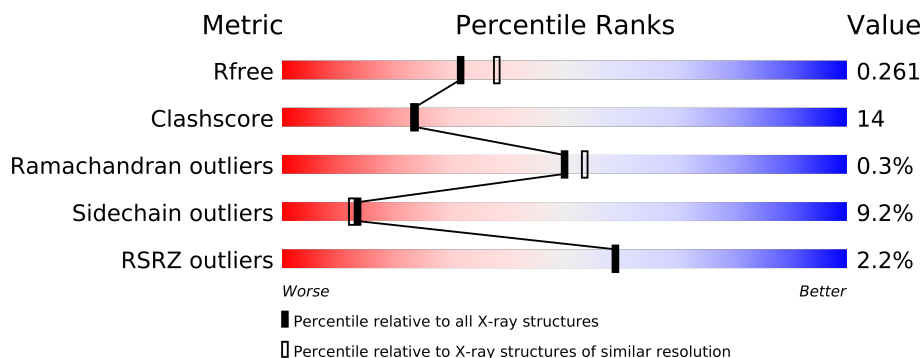
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.21 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))
$R_{free}$	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (2.20-2.20)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	160	
1	B	160	
1	C	160	
1	D	160	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9687 atoms, of which 4821 are hydrogens and 0 are deuterium.

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 DUAL SPECIFICITY PROTEIN PHOSPHATASE 26.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	152	Total	C	H	N	O	S	0	0	0
			2435	764	1225	232	210	4			
1	B	150	Total	C	H	N	O	S	0	0	0
			2408	756	1212	230	207	3			
1	C	150	Total	C	H	N	O	S	0	0	0
			2384	749	1200	223	208	4			
1	D	147	Total	C	H	N	O	S	0	0	0
			2347	735	1184	223	202	3			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	60	MET	-	EXPRESSION TAG	UNP Q9BV47
A	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
A	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
A	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	219	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47
B	60	MET	-	EXPRESSION TAG	UNP Q9BV47
B	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
B	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
B	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	219	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47
C	60	MET	-	EXPRESSION TAG	UNP Q9BV47

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
C	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
C	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	219	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47
D	60	MET	-	EXPRESSION TAG	UNP Q9BV47
D	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
D	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
D	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	219	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	44	Total O 44 44	0	0
2	B	22	Total O 22 22	0	0
2	C	27	Total O 27 27	0	0
2	D	20	Total O 20 20	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.60Å 82.78Å 91.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.12 – 2.21 40.12 – 2.21	Depositor EDS
% Data completeness (in resolution range)	98.0 (40.12-2.21) 98.0 (40.12-2.21)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.38 (at 2.20Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.200 , 0.266 0.191 , 0.261	Depositor DCC
$R_{free}$ test set	1576 reflections (4.97%)	DCC
Wilson B-factor (Å <sup>2</sup> )	28.0	Xtriage
Anisotropy	0.599	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 35.1	EDS
Estimated twinning fraction	0.008 for k,h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 31751 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9687	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.60 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.2858e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	1.08	0/1237	0.97	2/1670 (0.1%)
1	B	0.95	1/1223 (0.1%)	0.93	2/1653 (0.1%)
1	C	1.00	1/1208 (0.1%)	0.93	3/1631 (0.2%)
1	D	0.98	1/1186 (0.1%)	0.96	2/1598 (0.1%)
All	All	1.00	3/4854 (0.1%)	0.95	9/6552 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	166	TYR	CE2-CZ	5.83	1.46	1.38
1	D	177	GLU	CG-CD	5.64	1.60	1.51
1	B	166	TYR	CG-CD1	5.38	1.46	1.39

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	156	VAL	CB-CA-C	-6.89	98.31	111.40
1	D	186	ARG	NE-CZ-NH2	-6.44	117.08	120.30
1	A	202	ASP	CB-CG-OD1	6.11	123.80	118.30
1	A	73	ASP	CB-CG-OD1	5.59	123.33	118.30
1	C	60	MET	CB-CG-SD	-5.53	95.81	112.40

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1210	1225	0	46	0
1	B	1196	1212	0	40	0
1	C	1184	1200	0	38	0
1	D	1163	1184	0	33	0
2	A	44	0	0	2	0
2	B	22	0	0	0	0
2	C	27	0	0	1	0
2	D	20	0	0	1	0
All	All	4866	4821	0	138	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 14.

The worst 5 of 138 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:82:GLU:HA	1:B:82:GLU:OE1	1.67	0.92
1:A:74:GLN:HG2	1:A:153:ALA:HA	1.66	0.77
1:C:126:MET:HE3	1:C:162:LEU:HD22	1.68	0.76
1:B:128:ILE:HG21	1:B:129:HIS:CD2	2.21	0.75
1:A:74:GLN:CG	1:A:153:ALA:HA	2.16	0.75

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	150/160 (94%)	148 (99%)	2 (1%)	0	100	100
1	B	148/160 (92%)	135 (91%)	11 (7%)	2 (1%)	16	12
1	C	146/160 (91%)	141 (97%)	5 (3%)	0	100	100
1	D	141/160 (88%)	134 (95%)	7 (5%)	0	100	100
All	All	585/640 (91%)	558 (95%)	25 (4%)	2 (0%)	50	53



All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	99	TRP
1	B	155	GLY

### 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	124/132 (94%)	116 (94%)	8 (6%)	24	25
1	B	123/132 (93%)	109 (89%)	14 (11%)	8	7
1	C	122/132 (92%)	109 (89%)	13 (11%)	10	8
1	D	119/132 (90%)	109 (92%)	10 (8%)	16	15
All	All	488/528 (92%)	443 (91%)	45 (9%)	13	12

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	206	ARG
1	C	84	ARG
1	D	172	HIS
1	C	60	MET
1	C	85	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	185	HIS
1	B	191	ASN
1	D	74	GLN
1	B	96	HIS
1	B	143	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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	152/160 (95%)	-0.17	0 100 100	15, 30, 55, 71	0
1	B	150/160 (93%)	0.12	9 (6%) 21 21	18, 38, 74, 101	0
1	C	150/160 (93%)	-0.05	1 (0%) 84 86	19, 36, 63, 82	0
1	D	147/160 (91%)	0.10	3 (2%) 62 62	22, 42, 61, 83	0
All	All	599/640 (93%)	0.00	13 (2%) 59 59	15, 37, 65, 101	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	100	ARG	3.8
1	B	104	GLU	3.4
1	B	99	TRP	2.8
1	B	101	GLY	2.8
1	B	65	VAL	2.8

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.