



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 04:43 AM GMT

PDB ID : 3W2A  
Title : Crystal structure of VirB core domain complexed with the cis-acting site upstream icsp promoter  
Authors : Gao, X.P.; Waltersperger, S.; Wang, M.T.; Cui, S.  
Deposited on : 2012-11-27  
Resolution : 2.77 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>

---

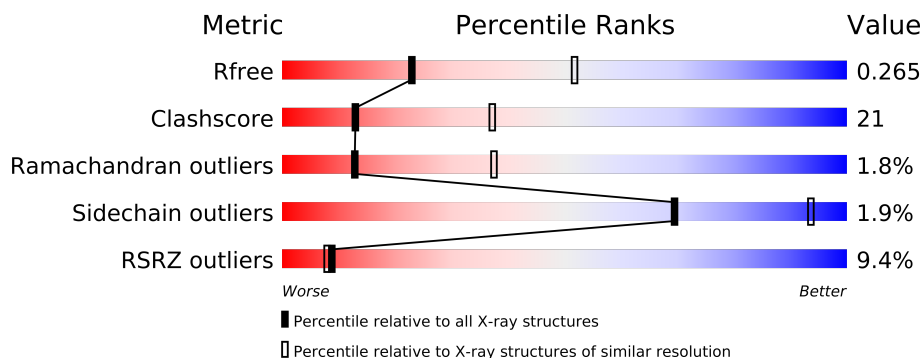
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.77 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	2193 (2.80-2.76)
Clashscore	79885	2751 (2.80-2.76)
Ramachandran outliers	78287	2699 (2.80-2.76)
Sidechain outliers	78261	2701 (2.80-2.76)
RSRZ outliers	66119	2196 (2.80-2.76)

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	143	
2	B	31	
3	C	31	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3382 atoms, of which 1478 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 Virulence regulon transcriptional activator VirB.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	114	Total	C	H	N	O	Se	0	1	0
			1887	595	967	150	173	2			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	108	MSE	-	EXPRESSION TAG	UNP P0A247
A	109	GLY	-	EXPRESSION TAG	UNP P0A247
A	110	SER	-	EXPRESSION TAG	UNP P0A247
A	111	SER	-	EXPRESSION TAG	UNP P0A247
A	112	HIS	-	EXPRESSION TAG	UNP P0A247
A	113	HIS	-	EXPRESSION TAG	UNP P0A247
A	114	HIS	-	EXPRESSION TAG	UNP P0A247
A	115	HIS	-	EXPRESSION TAG	UNP P0A247
A	116	HIS	-	EXPRESSION TAG	UNP P0A247
A	117	HIS	-	EXPRESSION TAG	UNP P0A247
A	118	SER	-	EXPRESSION TAG	UNP P0A247
A	119	SER	-	EXPRESSION TAG	UNP P0A247
A	120	GLY	-	EXPRESSION TAG	UNP P0A247
A	121	LEU	-	EXPRESSION TAG	UNP P0A247
A	122	VAL	-	EXPRESSION TAG	UNP P0A247
A	123	PRO	-	EXPRESSION TAG	UNP P0A247
A	124	ARG	-	EXPRESSION TAG	UNP P0A247
A	125	GLY	-	EXPRESSION TAG	UNP P0A247
A	126	SER	-	EXPRESSION TAG	UNP P0A247
A	127	HIS	-	EXPRESSION TAG	UNP P0A247
A	128	MSE	-	EXPRESSION TAG	UNP P0A247

- Molecule 2 is a DNA chain called DNA (31-mer).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	23	Total	C	H	N	O	P	0	0	0
			727	225	262	78	139	23			

- Molecule 3 is a DNA chain called DNA (31-mer).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	23	Total	C	H	N	O	P	0	0	0
			727	229	249	89	137	23			

- Molecule 4 is water.

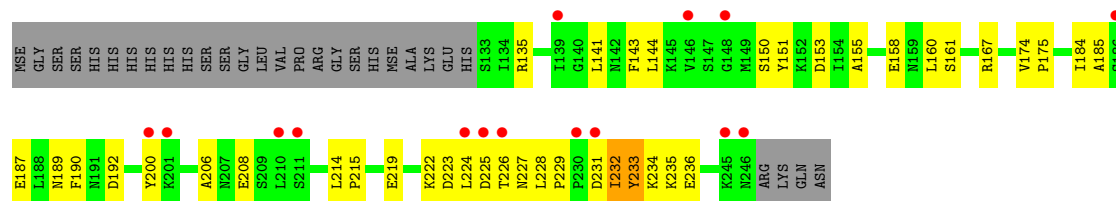
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	18	Total	O	0	0
			18	18		
4	B	12	Total	O	0	0
			12	12		
4	C	11	Total	O	0	0
			11	11		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Virulence regulon transcriptional activator VirB

Chain A: 



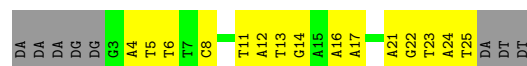
- Molecule 2: DNA (31-mer)

Chain B: 



- Molecule 3: DNA (31-mer)

Chain C: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	176.32Å 39.83Å 73.49Å 90.00° 105.79° 90.00°	Depositor
Resolution (Å)	48.25 – 2.77 48.24 – 2.78	Depositor EDS
% Data completeness (in resolution range)	97.8 (48.25-2.77) 98.6 (48.24-2.78)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 2.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.241 , 0.262 0.243 , 0.265	Depositor DCC
$R_{free}$ test set	630 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	58.1	Xtriage
Anisotropy	0.719	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 42.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	7 of 12618 reflections (0.055%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3382	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.22% of the height of the origin peak. No significant pseudotranslation is detected.*

<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	0.24	0/933	0.43	0/1251
2	B	0.47	0/519	1.13	0/797
3	C	0.52	0/537	1.17	1/828 (0.1%)
All	All	0.40	0/1989	0.91	1/2876 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	21	DA	O4'-C1'-N9	7.09	112.97	108.00

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	920	967	0	33	0
2	B	465	262	0	21	0
3	C	478	249	0	17	0
4	A	18	0	0	5	0
4	B	12	0	0	6	0
4	C	11	0	0	8	0
All	All	1904	1478	0	69	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 21.

The worst 5 of 69 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:135:ARG:NH2	1:A:174:VAL:O	1.91	1.02
3:C:14:DG:OP2	4:C:102:HOH:O	1.93	0.86
1:A:225:ASP:N	4:A:306:HOH:O	2.16	0.77
2:B:22:DA:N7	4:B:105:HOH:O	2.16	0.77
1:A:150:SER:OG	1:A:153:ASP:OD1	2.09	0.69

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	113/143 (79%)	102 (90%)	9 (8%)	2 (2%)	13	37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	232	ILE
1	A	233	TYR

### 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	104/125 (83%)	102 (98%)	2 (2%)	69	93

All (2) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	161	SER
1	A	223	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	198	ASN
1	A	207	ASN

### 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	114/143 (79%)	0.80	15 (13%) 4 4	40, 69, 123, 153	0
2	B	23/31 (74%)	0.04	0 100 100	54, 86, 224, 244	0
3	C	23/31 (74%)	-0.17	0 100 100	56, 99, 202, 229	0
All	All	160/205 (78%)	0.55	15 (9%) 9 8	40, 72, 184, 244	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	146	VAL	8.9
1	A	246	ASN	7.6
1	A	211	SER	4.2
1	A	226	THR	4.1
1	A	225	ASP	3.0

### 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.