



# Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 06:43 am BST

PDB ID : 3SAQ  
Title : Structure of D13, the scaffolding protein of vaccinia virus  
Authors : Coulibaly, F.  
Deposited on : 2011-06-03  
Resolution : 3.51 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

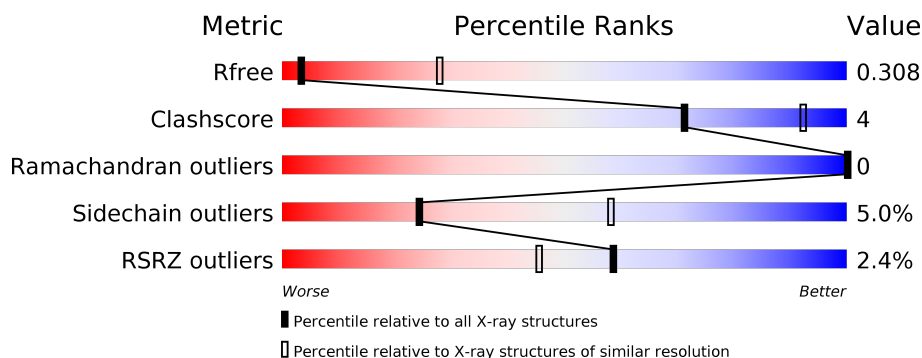
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.51 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	1161 (3.60-3.44)
Clashscore	141614	1244 (3.60-3.44)
Ramachandran outliers	138981	1206 (3.60-3.44)
Sidechain outliers	138945	1207 (3.60-3.44)
RSRZ outliers	127900	1080 (3.60-3.44)

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	576	
1	B	576	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7428 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 Rifampicin resistance protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	465	Total	C	N	O	S	0	0	0
			3612	2321	571	709	11			
1	B	492	Total	C	N	O	S	0	0	0
			3816	2446	608	751	11			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	MET	-	EXPRESSION TAG	UNP P68440
A	-23	SER	-	EXPRESSION TAG	UNP P68440
A	-22	TYR	-	EXPRESSION TAG	UNP P68440
A	-21	TYR	-	EXPRESSION TAG	UNP P68440
A	-20	HIS	-	EXPRESSION TAG	UNP P68440
A	-19	HIS	-	EXPRESSION TAG	UNP P68440
A	-18	HIS	-	EXPRESSION TAG	UNP P68440
A	-17	HIS	-	EXPRESSION TAG	UNP P68440
A	-16	HIS	-	EXPRESSION TAG	UNP P68440
A	-15	HIS	-	EXPRESSION TAG	UNP P68440
A	-14	ASP	-	EXPRESSION TAG	UNP P68440
A	-13	TYR	-	EXPRESSION TAG	UNP P68440
A	-12	ASP	-	EXPRESSION TAG	UNP P68440
A	-11	ILE	-	EXPRESSION TAG	UNP P68440
A	-10	PRO	-	EXPRESSION TAG	UNP P68440
A	-9	THR	-	EXPRESSION TAG	UNP P68440
A	-8	THR	-	EXPRESSION TAG	UNP P68440
A	-7	GLU	-	EXPRESSION TAG	UNP P68440
A	-6	ASN	-	EXPRESSION TAG	UNP P68440
A	-5	LEU	-	EXPRESSION TAG	UNP P68440
A	-4	TYR	-	EXPRESSION TAG	UNP P68440
A	-3	PHE	-	EXPRESSION TAG	UNP P68440
A	-2	GLN	-	EXPRESSION TAG	UNP P68440
A	-1	GLY	-	EXPRESSION TAG	UNP P68440
A	0	ALA	-	EXPRESSION TAG	UNP P68440

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-24	MET	-	EXPRESSION TAG	UNP P68440
B	-23	SER	-	EXPRESSION TAG	UNP P68440
B	-22	TYR	-	EXPRESSION TAG	UNP P68440
B	-21	TYR	-	EXPRESSION TAG	UNP P68440
B	-20	HIS	-	EXPRESSION TAG	UNP P68440
B	-19	HIS	-	EXPRESSION TAG	UNP P68440
B	-18	HIS	-	EXPRESSION TAG	UNP P68440
B	-17	HIS	-	EXPRESSION TAG	UNP P68440
B	-16	HIS	-	EXPRESSION TAG	UNP P68440
B	-15	HIS	-	EXPRESSION TAG	UNP P68440
B	-14	ASP	-	EXPRESSION TAG	UNP P68440
B	-13	TYR	-	EXPRESSION TAG	UNP P68440
B	-12	ASP	-	EXPRESSION TAG	UNP P68440
B	-11	ILE	-	EXPRESSION TAG	UNP P68440
B	-10	PRO	-	EXPRESSION TAG	UNP P68440
B	-9	THR	-	EXPRESSION TAG	UNP P68440
B	-8	THR	-	EXPRESSION TAG	UNP P68440
B	-7	GLU	-	EXPRESSION TAG	UNP P68440
B	-6	ASN	-	EXPRESSION TAG	UNP P68440
B	-5	LEU	-	EXPRESSION TAG	UNP P68440
B	-4	TYR	-	EXPRESSION TAG	UNP P68440
B	-3	PHE	-	EXPRESSION TAG	UNP P68440
B	-2	GLN	-	EXPRESSION TAG	UNP P68440
B	-1	GLY	-	EXPRESSION TAG	UNP P68440
B	0	ALA	-	EXPRESSION TAG	UNP P68440

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 ( $\text{RSRZ} > 2$ ). 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.

- Chain A:
- 

- Chain B:
- 
- 76% 8% 15%
- MET SER TYR THR HIS HIS HIS HIS ASP TYR ASP ASP ILE PRO THR THR GLU ASN ASN LEU TYR PHE GLN GLY ALA MET ASN ASN THR ILE ILE ASN ASN SER LEU LEU ILE GLY GLY ASP ASP SER SER ILE LYS ARG SER ASN VAL VAL PHE ALA VAL ASP ASP SER GLN ILE PRO THR L31 Y32 M33 V30
- G47 P40 ASP R50 S55 Y62 L69 E77 F84 Y89 R93 S129 S139 M145 I146 K149 F150 P151 S188 R193 L212 Q223 S227 F228 I229 T243 V249 L255 Y262 D293 E314 V315 V341 T352 A361
- LYS M355 S356 F357 K364 A367 I368 K377 I384 S385 H386 I391 D392 I431 D432 I439 R442 D450 Y453 L477 N480 P481 THR PRO LYS ILE PHE PHE ARG PRO THR ILE THR ALA ASN VAL VAL ARG ARG GLY LYS D501 K502 L503 M512 Q522
- K534 V535 D538 I543 I546 MET GLY ASP ASN

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.13Å 125.13Å 370.83Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	26.81 – 3.51 26.81 – 3.51	Depositor EDS
% Data completeness (in resolution range)	97.8 (26.81-3.51) 97.3 (26.81-3.51)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.73 (at 3.55Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, $R_{free}$	0.245 , 0.257 0.283 , 0.308	Depositor DCC
$R_{free}$ test set	699 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	83.3	Xtriage
Anisotropy	0.600	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 58.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	7428	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% 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	0.47	0/3691	0.67	1/5038 (0.0%)
1	B	0.48	0/3899	0.67	1/5322 (0.0%)
All	All	0.47	0/7590	0.67	2/10360 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	129	SER	C-N-CA	5.48	133.80	122.30
1	B	129	SER	C-N-CA	5.21	133.24	122.30

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	3612	0	3479	27	0
1	B	3816	0	3674	32	0
All	All	7428	0	7153	51	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 (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:A:262:TYR:HE1	1:B:450:ASP:HB3	1.20	1.05
1:A:262:TYR:CE1	1:B:450:ASP:HB3	1.98	0.98
1:A:262:TYR:HE1	1:B:450:ASP:CB	1.84	0.91
1:B:535:VAL:HG12	1:B:543:ILE:HG22	1.66	0.77
1:A:89:TYR:HA	1:A:129:SER:OG	1.92	0.69
1:B:89:TYR:HA	1:B:129:SER:OG	1.93	0.68
1:B:391:ILE:HD13	1:B:432:ASP:HB3	1.76	0.68
1:A:391:ILE:HD13	1:A:432:ASP:HB3	1.75	0.67
1:A:262:TYR:HE1	1:B:450:ASP:CG	1.98	0.66
1:B:249:VAL:HG11	1:B:503:LEU:HD13	1.80	0.63
1:B:384:ILE:HG22	1:B:386:HIS:HD2	1.65	0.62
1:A:384:ILE:HG22	1:A:386:HIS:HD2	1.65	0.61
1:A:249:VAL:HG11	1:A:503:LEU:HD13	1.82	0.60
1:A:262:TYR:CE1	1:B:450:ASP:CB	2.72	0.57
1:A:262:TYR:CE1	1:B:450:ASP:CG	2.78	0.55
1:A:255:LEU:HD22	1:A:503:LEU:HD23	1.90	0.53
1:B:255:LEU:HD22	1:B:503:LEU:HD23	1.90	0.53
1:A:391:ILE:CD1	1:A:431:ILE:HD12	2.39	0.52
1:B:384:ILE:CG2	1:B:386:HIS:HD2	2.23	0.52
1:A:384:ILE:CG2	1:A:386:HIS:HD2	2.22	0.52
1:B:391:ILE:CD1	1:B:431:ILE:HD12	2.40	0.51
1:A:391:ILE:HD12	1:A:431:ILE:HD12	1.91	0.51
1:B:84:PHE:HB3	1:B:148:ILE:HG21	1.92	0.51
1:B:391:ILE:HD12	1:B:431:ILE:HD12	1.93	0.49
1:B:315:VAL:HG22	1:B:341:VAL:HB	1.95	0.48
1:B:439:ILE:H	1:B:439:ILE:HD13	1.78	0.48
1:A:315:VAL:HG22	1:A:341:VAL:HB	1.96	0.48
1:A:439:ILE:H	1:A:439:ILE:HD13	1.78	0.47
1:B:367:ALA:HB3	1:B:385:SER:HB2	1.98	0.45
1:B:77:GLU:HB2	1:B:151:PRO:HD3	1.98	0.45
1:A:262:TYR:HB3	1:A:522:GLN:HB2	1.98	0.45
1:B:368:ILE:HG23	1:B:384:ILE:HD11	1.97	0.45
1:A:77:GLU:HB2	1:A:151:PRO:HD3	1.97	0.45
1:A:367:ALA:HB3	1:A:385:SER:HB2	1.98	0.45
1:A:368:ILE:HG23	1:A:384:ILE:HD11	1.97	0.44
1:B:364:LYS:HE3	1:B:392:ASP:OD1	2.18	0.44
1:B:512:MET:HG2	1:B:512:MET:H	1.57	0.44
1:A:364:LYS:HE3	1:A:392:ASP:OD1	2.18	0.43
1:B:453:TYR:CD2	1:B:477:LEU:HD21	2.53	0.43
1:B:139:SER:OG	1:B:149:LYS:HE2	2.18	0.43
1:B:262:TYR:HB3	1:B:522:GLN:HB2	2.00	0.43
1:A:42:VAL:HG23	1:A:55:SER:HB2	2.00	0.42

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:69:LEU:CD2	1:A:212:LEU:HD11	2.50	0.42
1:B:42:VAL:HG23	1:B:55:SER:HB2	2.02	0.42
1:A:235:VAL:HG13	1:B:442:ARG:CZ	2.50	0.41
1:B:89:TYR:O	1:B:93:LYS:HD2	2.20	0.41
1:A:453:TYR:CD2	1:A:477:LEU:HD21	2.56	0.40
1:B:69:LEU:CD2	1:B:212:LEU:HD11	2.51	0.40
1:A:262:TYR:CE1	1:B:450:ASP:OD1	2.74	0.40
1:B:534:LYS:HB2	1:B:546:ILE:HD11	2.03	0.40
1:A:302:PRO:HA	1:A:303:PRO:HD3	2.01	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	453/576 (79%)	437 (96%)	16 (4%)	0	100	100
1	B	484/576 (84%)	466 (96%)	18 (4%)	0	100	100
All	All	937/1152 (81%)	903 (96%)	34 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	407/528 (77%)	386 (95%)	21 (5%)	23	57
1	B	429/528 (81%)	408 (95%)	21 (5%)	25	59
All	All	836/1056 (79%)	794 (95%)	42 (5%)	24	58

All (42) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	33	MET
1	A	42	VAL
1	A	93	LYS
1	A	139	SER
1	A	145	ASN
1	A	188	SER
1	A	193	ARG
1	A	227	SER
1	A	229	ILE
1	A	243	THR
1	A	293	ASP
1	A	314	GLU
1	A	352	THR
1	A	356	SER
1	A	357	PHE
1	A	377	LYS
1	A	384	ILE
1	A	439	ILE
1	A	502	LYS
1	A	512	MET
1	A	513	ASP
1	B	33	MET
1	B	42	VAL
1	B	93	LYS
1	B	139	SER
1	B	145	ASN
1	B	188	SER
1	B	193	ARG
1	B	227	SER
1	B	229	ILE
1	B	243	THR
1	B	293	ASP
1	B	314	GLU
1	B	352	THR
1	B	356	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	357	PHE
1	B	377	LYS
1	B	384	ILE
1	B	439	ILE
1	B	502	LYS
1	B	512	MET
1	B	538	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	45	ASN

### 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](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

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	465/576 (80%)	0.25	19 (4%) 37 27	64, 110, 162, 201	0
1	B	492/576 (85%)	-0.20	4 (0%) 86 75	34, 84, 145, 192	0
All	All	957/1152 (83%)	0.02	23 (2%) 59 45	34, 99, 154, 201	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	355	ASN	5.2
1	B	223	GLN	3.3
1	A	310	ALA	3.2
1	A	512	MET	3.0
1	A	352	THR	2.9
1	A	327	ASP	2.8
1	A	62	TYR	2.7
1	A	351	GLY	2.7
1	A	226	PRO	2.6
1	A	145	ASN	2.5
1	A	140	ILE	2.5
1	A	40	SER	2.4
1	A	479	PHE	2.4
1	B	534	LYS	2.3
1	A	305	GLY	2.3
1	A	61	GLN	2.2
1	B	352	THR	2.2
1	A	176	SER	2.2
1	A	175	ASP	2.2
1	A	311	GLU	2.2
1	A	195	SER	2.1
1	A	32	TYR	2.1
1	B	62	TYR	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.