



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 07:14 am GMT

PDB ID : 3N4X  
Title : Structure of Csm1 full-length  
Authors : Corbett, K.D.; Harrison, S.C.  
Deposited on : 2010-05-23  
Resolution : 3.41 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
Xtriage (Phenix)	:	1.13
EDS	:	trunk31020
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)	:	trunk31020

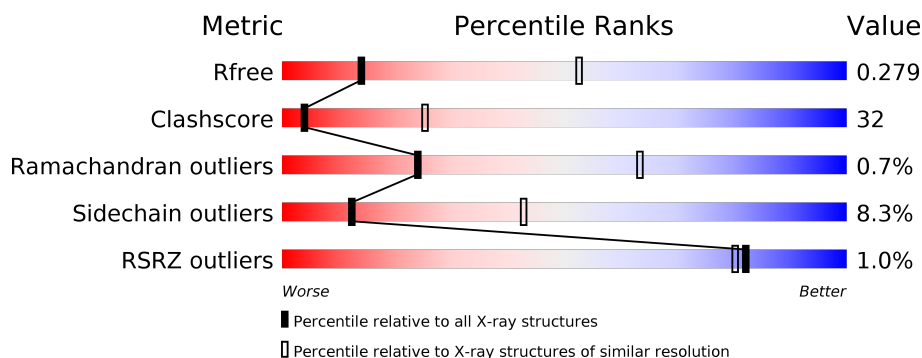
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.41 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	1928 (3.50-3.30)
Clashscore	122126	2051 (3.50-3.30)
Ramachandran outliers	120053	2006 (3.50-3.30)
Sidechain outliers	120020	2006 (3.50-3.30)
RSRZ outliers	108989	1827 (3.50-3.30)

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  $\leq 5\%$ . 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	190	<div> <div>%</div> <div> <div></div> <div>40%</div> <div>33%</div> <div>5%</div> <div>22%</div> </div> </div>
1	B	190	<div> <div>2%</div> <div> <div></div> <div>35%</div> <div>39%</div> <div>5%</div> <div>21%</div> </div> </div>
1	C	190	<div> <div></div> <div> <div>37%</div> <div>39%</div> <div>•</div> <div>19%</div> </div> </div>
1	D	190	<div> <div>%</div> <div> <div></div> <div>33%</div> <div>47%</div> <div>•</div> <div>16%</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 4949 atoms, of which 0 are hydrogens and 0 are deuteriums.

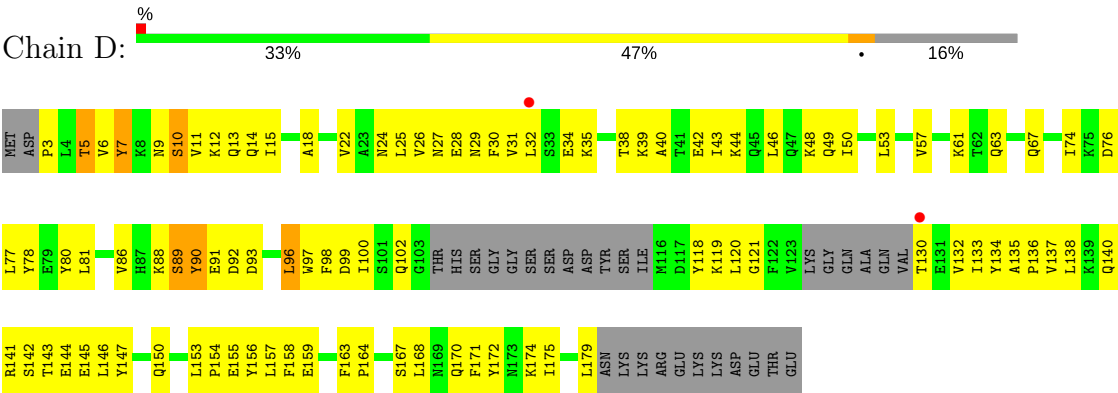
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Monopolin complex subunit CSM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total	C	N	O	S	0	0	0
			1202	770	190	240	2			
1	B	151	Total	C	N	O	S	0	0	0
			1229	789	193	245	2			
1	C	153	Total	C	N	O	S	0	0	0
			1232	787	197	246	2			
1	D	159	Total	C	N	O	S	0	0	0
			1286	822	204	258	2			



● Molecule 1: Monopolin complex subunit CSM1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.27Å 105.27Å 234.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.04 – 3.41 49.35 – 3.41	Depositor EDS
% Data completeness (in resolution range)	78.6 (35.04-3.41) 78.5 (49.35-3.41)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.37 (at 3.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.238 , 0.282 0.232 , 0.279	Depositor DCC
$R_{free}$ test set	825 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	125.8	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 83.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.086 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	4949	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	164.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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](#)

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	0.59	0/1219	0.78	3/1646 (0.2%)
1	B	0.56	0/1247	0.71	2/1683 (0.1%)
1	C	0.49	0/1250	0.68	0/1686
1	D	0.46	0/1305	0.62	0/1762
All	All	0.53	0/5021	0.70	5/6777 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	25	LEU	CB-CG-CD1	-7.51	98.24	111.00
1	B	20	LEU	CB-CG-CD1	6.20	121.53	111.00
1	A	25	LEU	CA-CB-CG	6.07	129.26	115.30
1	B	149	LEU	CA-CB-CG	-5.67	102.26	115.30
1	A	20	LEU	CB-CG-CD1	-5.08	102.37	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	1202	0	1197	71	0
1	B	1229	0	1225	84	0
1	C	1232	0	1228	98	0
1	D	1286	0	1281	98	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4949	0	4931	318	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 32.

The worst 5 of 318 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:81:LEU:HD11	1:B:172:TYR:HB2	1.29	1.11
1:C:172:TYR:HB2	1:D:81:LEU:HD11	1.42	1.01
1:C:81:LEU:HD11	1:D:172:TYR:HB2	1.45	0.95
1:C:54:ASN:O	1:C:58:LYS:HG2	1.71	0.89
1:B:96:LEU:H	1:B:96:LEU:HD12	1.34	0.88

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	143/190 (75%)	134 (94%)	8 (6%)	1 (1%)	24	63
1	B	143/190 (75%)	129 (90%)	13 (9%)	1 (1%)	24	63
1	C	147/190 (77%)	136 (92%)	10 (7%)	1 (1%)	24	63
1	D	153/190 (80%)	141 (92%)	11 (7%)	1 (1%)	24	63
All	All	586/760 (77%)	540 (92%)	42 (7%)	4 (1%)	24	63

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	93	ASP

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	93	ASP
1	D	93	ASP
1	C	93	ASP

### 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	137/176 (78%)	125 (91%)	12 (9%)	11	40
1	B	142/176 (81%)	130 (92%)	12 (8%)	12	42
1	C	141/176 (80%)	130 (92%)	11 (8%)	14	45
1	D	148/176 (84%)	136 (92%)	12 (8%)	13	44
All	All	568/704 (81%)	521 (92%)	47 (8%)	12	43

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	131	GLU
1	C	33	SER
1	D	142	SER
1	B	159	GLU
1	C	52	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	83	ASN
1	C	102	GLN
1	D	83	ASN
1	B	67	GLN
1	D	13	GLN

### 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 [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	149/190 (78%)	-0.09	1 (0%) 87 86	106, 152, 220, 262	0
1	B	151/190 (79%)	-0.04	3 (1%) 65 63	97, 161, 222, 268	0
1	C	153/190 (80%)	-0.18	0 100 100	98, 161, 217, 248	0
1	D	159/190 (83%)	-0.15	2 (1%) 77 75	108, 169, 227, 267	0
All	All	612/760 (80%)	-0.11	6 (0%) 82 80	97, 162, 222, 268	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	139	LYS	3.0
1	D	130	THR	2.7
1	A	157	LEU	2.5
1	B	146	LEU	2.4
1	B	147	TYR	2.3

### 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.