



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 25, 2024 – 09:42 AM EDT

PDB ID : 6MR3  
Title : Crystal structure of the competence-damaged protein (CinA) superfamily protein from *Streptococcus mutans*  
Authors : Stogios, P.J.; Cuff, M.; Xu, X.; Cui, H.; Di Leo, R.; Yim, V.; Chin, S.; Savchenko, A.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2018-10-11  
Resolution : 2.05 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

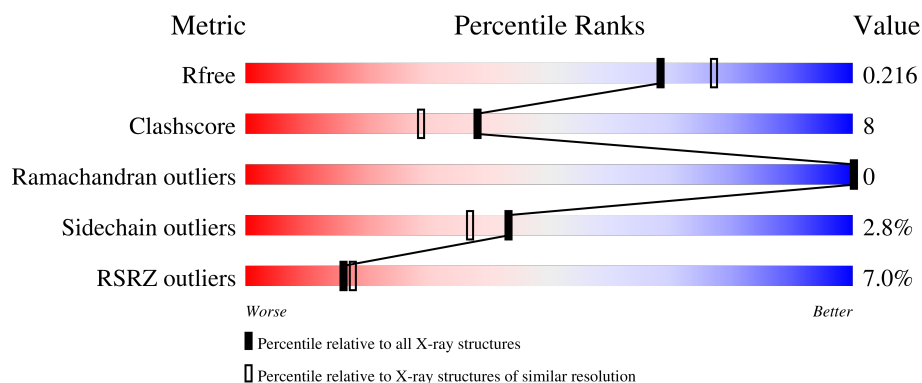
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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 of 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	418	 2% 51% 7% 41%
1	B	418	 3% 49% 9% 41%
1	C	418	 7% 46% 9% 43%

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
2	CL	A	502	-	-	X	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6204 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 Putative competence-damage inducible protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	245	Total	C	N	O	Se	0	1	0
			1902	1214	326	359	3			
1	B	245	Total	C	N	O	Se	0	0	0
			1897	1209	326	359	3			
1	C	238	Total	C	N	O	Se	0	0	0
			1841	1175	316	347	3			

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Cl	0	0
			2	2		
2	B	2	Total	Cl	0	0
			2	2		

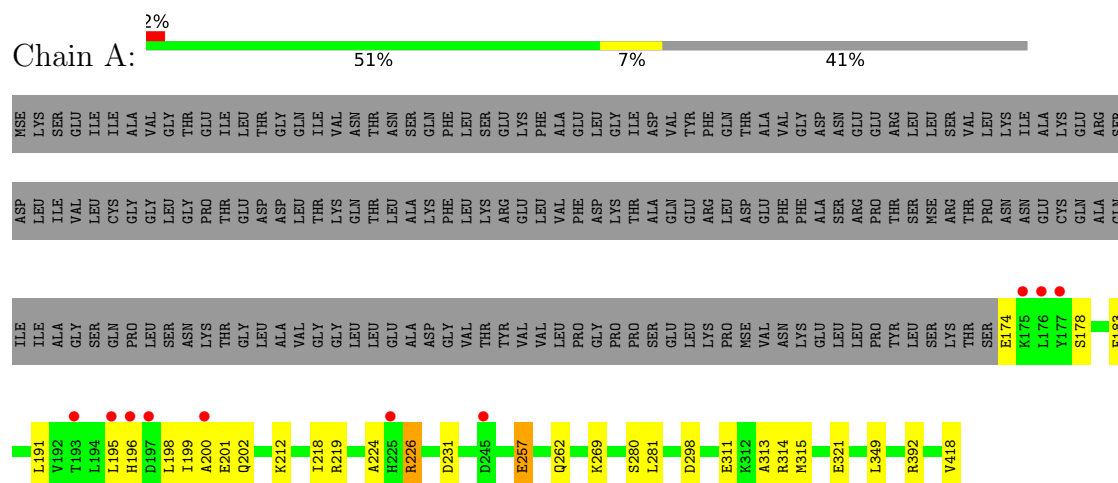
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	205	Total	O	0	8
			213	213		
3	B	260	Total	O	0	14
			274	274		
3	C	71	Total	O	0	1
			73	73		

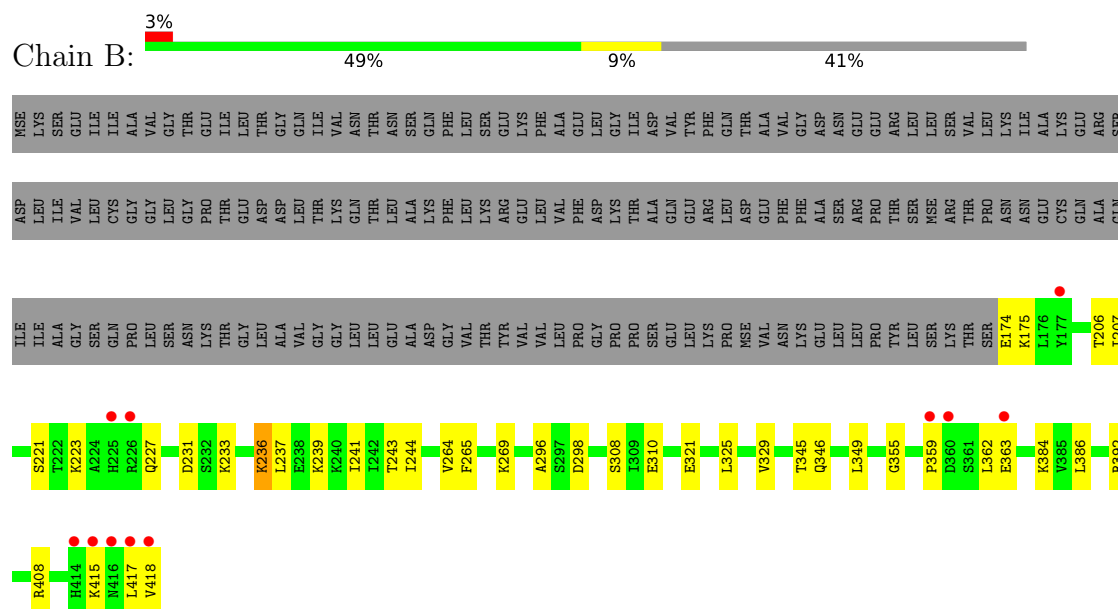
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 ( $RSRZ > 2$ ). 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: Putative competence-damage inducible protein



- Molecule 1: Putative competence-damage inducible protein



- Molecule 1: Putative competence-damage inducible protein



E339	E188	L191	V192	T193	L194	L195	H196	D197	L198	I199	ALA	GLU	GLN	T203	D204	P205	T206	T207	A208	P209	Y210	R219	L220	S221	T222	R226	Q227	K228	E229	K233	L234	D235	K236	L237	E238	K239	K240	I241	I242	T243	I244	D245	N246	D249	G255	E256	E311	M315	E321	E339	R340	Q343	L353	V356	L362	E363	G364	R379	L412	S413	H414	LYS	ASN	LEU	VAL		
ASP	LYS	ILE	VAL	LEU	VAL	CYS	ILE	ALA	VAL	GLY	LEU	THR	ILE	GLU	ASP	GLY	THR	GLN	THR	LEU	ASN	GLN	LYS	THR	VAL	LEU	VAL	ALA	LEU	ILE	THR	ALA	GLN	LYS	ARG	LEU	THR	ASP	GLU	VAL	PHE	GLY	ASP	LEU	GLY	ILE	THR	ARG	LEU	PRO	ASN	LYS	ILE	ALA	LYS	GLN	ALA	ARG	SER								
ILE	ILE	ALA	GLY	SER	GLN	PRO	LEU	SER	ASN	LYS	THR	GLY	LEU	ALA	VAL	GLY	LYS	LEU	GLU	ALA	ASP	ASP	VAL	GLY	THR	VAL	LEU	PRO	GLY	PRO	SER	GLU	D235	K236	L237	E238	K239	K240	I241	I242	T243	I244	D245	N246	D249	G255	E256	E311	M315	E321	E339	R340	Q343	L353	V356	L362	E363	G364	R379	L412	S413	H414	LYS	ASN	LEU	VAL	
ASP	LEU	ILE	VAL	LEU	CYS	ILE	ALA	VAL	GLY	LEU	THR	ILE	GLU	ASP	GLY	THR	GLN	THR	LEU	ASN	GLN	LYS	THR	VAL	LEU	VAL	ALA	LEU	PRO	GLY	PRO	SER	GLU	D235	K236	L237	E238	K239	K240	I241	I242	T243	I244	D245	N246	D249	G255	E256	E311	M315	E321	E339	R340	Q343	L353	V356	L362	E363	G364	R379	L412	S413	H414	LYS	ASN	LEU	VAL

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	152.03Å 54.51Å 111.01Å 90.00° 125.38° 90.00°	Depositor
Resolution (Å)	41.21 – 2.05 41.21 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.0 (41.21-2.05) 97.4 (41.21-1.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	18.96 (at 1.95Å)	Xtriage
Refinement program	PHENIX (1.14_3260)	Depositor
R, $R_{free}$	0.186 , 0.215 0.186 , 0.216	Depositor DCC
$R_{free}$ test set	2000 reflections (3.78%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.1	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 56.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6204	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	0.27	0/1932	0.49	0/2599
1	B	0.27	0/1924	0.50	0/2588
1	C	0.30	0/1867	0.55	0/2511
All	All	0.28	0/5723	0.51	0/7698

There are no bond length outliers.

There are no bond angle outliers.

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	1902	0	1955	27	0
1	B	1897	0	1944	27	0
1	C	1841	0	1885	37	0
2	A	2	0	0	2	0
2	B	2	0	0	0	0
3	A	213	0	0	5	0
3	B	274	0	0	1	0
3	C	73	0	0	0	0
All	All	6204	0	5784	87	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 8.

The worst 5 of 87 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:321:GLU:HG2	1:B:321:GLU:OE2	1.80	0.80
1:C:175:LYS:HE3	1:C:177:TYR:OH	1.92	0.69
1:C:205:PRO:HD2	1:C:237:LEU:HD22	1.74	0.69
1:B:207:ILE:HD11	1:B:237:LEU:HD11	1.75	0.67
1:C:226:ARG:HG3	1:C:229:GLU:H	1.60	0.66

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	244/418 (58%)	236 (97%)	8 (3%)	0	100	100
1	B	243/418 (58%)	237 (98%)	6 (2%)	0	100	100
1	C	234/418 (56%)	224 (96%)	10 (4%)	0	100	100
All	All	721/1254 (58%)	697 (97%)	24 (3%)	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	203/345 (59%)	198 (98%)	5 (2%)	47	40
1	B	202/345 (59%)	200 (99%)	2 (1%)	76	75
1	C	196/345 (57%)	186 (95%)	10 (5%)	24	15
All	All	601/1035 (58%)	584 (97%)	17 (3%)	43	37

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	353	LEU
1	C	414	HIS
1	C	175	LYS
1	C	188	GLU
1	C	197	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	242/418 (57%)	-0.11	10 (4%) 37 40	16, 30, 94, 131	0
1	B	242/418 (57%)	-0.06	11 (4%) 33 35	16, 31, 90, 140	0
1	C	235/418 (56%)	0.36	29 (12%) 4 3	26, 53, 124, 151	0
All	All	719/1254 (57%)	0.06	50 (6%) 16 18	16, 38, 110, 151	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	414	HIS	9.4
1	C	194	LEU	9.4
1	C	198	LEU	8.6
1	C	414	HIS	6.8
1	B	416	ASN	6.6

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CL	A	502	1/1	0.91	0.15	45,45,45,45	0
2	CL	B	501	1/1	0.92	0.08	31,31,31,31	1
2	CL	A	501	1/1	0.96	0.06	33,33,33,33	1
2	CL	B	502	1/1	0.99	0.10	30,30,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.