



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2023 – 04:12 AM EDT

PDB ID : 2F9Z  
Title : Complex between the chemotaxis deamidase CheD and the chemotaxis phosphatase CheC from *Thermotoga maritima*  
Authors : Chao, X.; Park, S.Y.; Bilwes, A.M.; Crane, B.R.  
Deposited on : 2005-12-06  
Resolution : 2.40 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  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

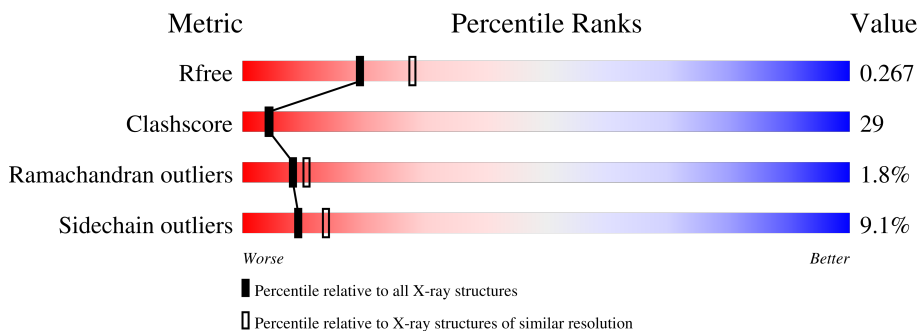
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	205	47% (green), 41% (yellow), 6% (orange), 6% (red), 0% (grey)
1	B	205	43% (green), 45% (yellow), 7% (orange), 0% (red), 5% (grey)
2	C	159	50% (green), 44% (yellow), 0% (orange), 0% (red), 6% (grey)
2	D	159	50% (green), 43% (yellow), 0% (orange), 0% (red), 7% (grey)

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5552 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 chemotaxis protein CheC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	193	1452	939	222	281	10	0	0	0
1	B	190	1433	930	224	268	11	0	0	0

- Molecule 2 is a protein called PROTEIN (chemotaxis methylation protein).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	154	1140	719	201	210	10	0	0	0
2	D	154	1135	715	202	209	9	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	ALA	-	cloning artifact	UNP Q9X005
C	0	HIS	-	cloning artifact	UNP Q9X005
D	-1	ALA	-	cloning artifact	UNP Q9X005
D	0	HIS	-	cloning artifact	UNP Q9X005

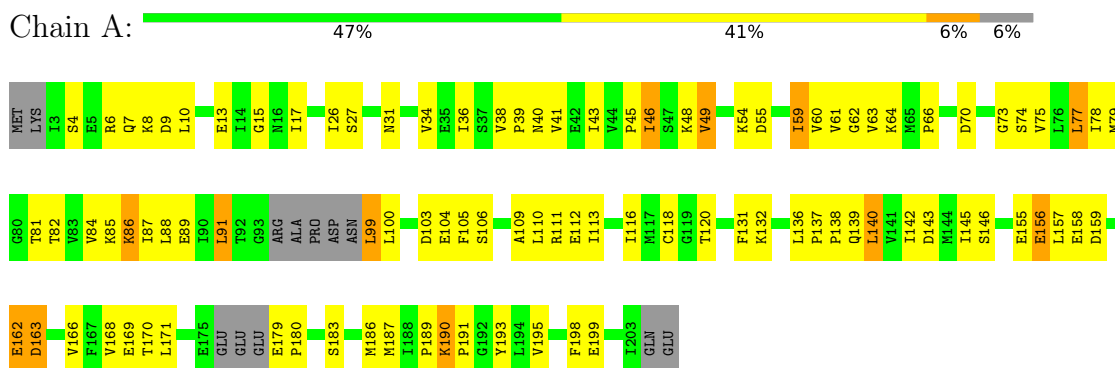
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	111	Total 111	O 111	0	0
3	B	100	Total 100	O 100	0	0
3	C	94	Total 94	O 94	0	0
3	D	87	Total 87	O 87	0	0

### 3 Residue-property plots [i](#)

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. 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. 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.

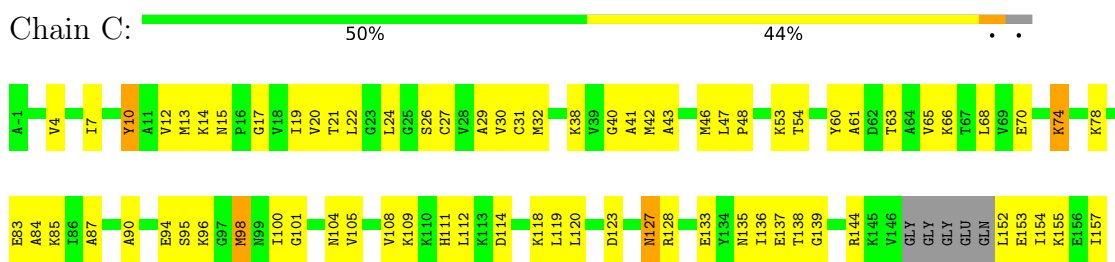
- Molecule 1: chemotaxis protein CheC



- Molecule 1: chemotaxis protein CheC

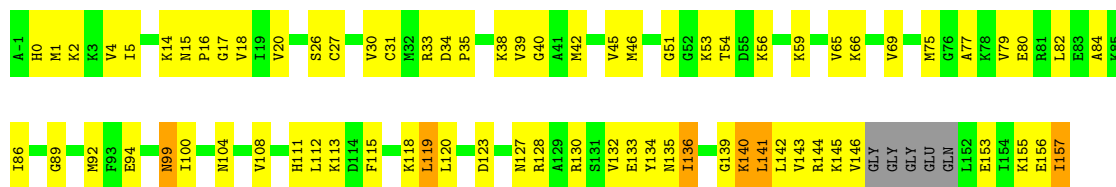


- Molecule 2: PROTEIN (chemotaxis methylation protein)



- Molecule 2: PROTEIN (chemotaxis methylation protein)

Chain D:  50% 43%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.08Å 66.08Å 161.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.40 28.61 – 2.40	Depositor EDS
% Data completeness (in resolution range)	93.4 (30.00-2.40) 53.0 (28.61-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.63 (at 2.39Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.211 , 0.275 0.211 , 0.267	Depositor DCC
$R_{free}$ test set	1635 reflections (9.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.9	Xtrriage
Anisotropy	0.193	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 33.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.37$ , $\langle L^2 \rangle = 0.20$	Xtrriage
Estimated twinning fraction	0.186 for -h,-k,l 0.319 for h,-h-k,-l 0.217 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5552	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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	0.36	0/1468	0.65	0/1987
1	B	0.42	0/1450	0.67	1/1961 (0.1%)
2	C	0.42	1/1149 (0.1%)	0.68	1/1531 (0.1%)
2	D	0.38	0/1145	0.69	1/1527 (0.1%)
All	All	0.39	1/5212 (0.0%)	0.67	3/7006 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	157	ILE	C-OXT	-5.57	1.12	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	114	ASP	CB-CA-C	-6.43	97.54	110.40
1	B	164	GLN	N-CA-C	-6.12	94.48	111.00
2	D	17	GLY	N-CA-C	5.54	126.95	113.10

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	1452	0	1504	90	0
1	B	1433	0	1518	97	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1140	0	1220	65	0
2	D	1135	0	1202	59	0
3	A	111	0	0	10	0
3	B	100	0	0	18	0
3	C	94	0	0	6	0
3	D	87	0	0	6	0
All	All	5552	0	5444	305	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 29.

The worst 5 of 305 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:45:PRO:HB3	1:A:163:ASP:HB3	1.33	1.11
1:A:190:LYS:HB3	1:A:191:PRO:HD2	1.41	1.00
1:A:66:PRO:HG2	1:A:136:LEU:HB2	1.49	0.94
2:D:39:VAL:HG21	2:D:75:MET:HG2	1.51	0.91
2:C:7:ILE:H	2:C:7:ILE:HD12	1.36	0.90

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	187/205 (91%)	169 (90%)	14 (8%)	4 (2%)	7	8
1	B	184/205 (90%)	159 (86%)	19 (10%)	6 (3%)	4	3
2	C	150/159 (94%)	131 (87%)	17 (11%)	2 (1%)	12	17
2	D	150/159 (94%)	137 (91%)	13 (9%)	0	100	100
All	All	671/728 (92%)	596 (89%)	63 (9%)	12 (2%)	8	10



5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	90	ILE
1	B	91	LEU
1	B	94	ARG
1	B	95	ALA
1	A	163	ASP

### 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	162/179 (90%)	146 (90%)	16 (10%)	8 11
1	B	161/179 (90%)	146 (91%)	15 (9%)	9 13
2	C	119/122 (98%)	112 (94%)	7 (6%)	19 32
2	D	117/122 (96%)	104 (89%)	13 (11%)	6 8
All	All	559/602 (93%)	508 (91%)	51 (9%)	9 14

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	112	GLU
2	C	96	LYS
2	D	141	LEU
1	B	139	GLN
2	C	78	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	135	ASN
2	D	99	ASN
2	D	104	ASN
2	C	104	ASN
2	C	15	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 monosaccharides 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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.