



# Full wwPDB X-ray Structure Validation Report ⓘ

May 18, 2020 – 07:07 am BST

PDB ID : 4JPB  
Title : The structure of a ternary complex between CheA domains P4 and P5 with CheW and with an unzipped fragment of TM14, a chemoreceptor analog from *Thermotoga maritima*.  
Authors : Li, X.; Bayas, C.; Bilwes, A.M.; Crane, B.R.  
Deposited on : 2013-03-19  
Resolution : 3.19 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

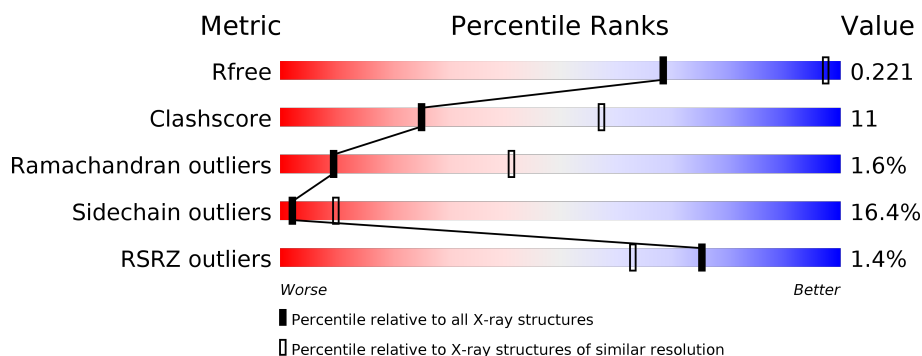
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.19 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	1467 (3.20-3.16)
Clashscore	141614	1599 (3.20-3.16)
Ramachandran outliers	138981	1574 (3.20-3.16)
Sidechain outliers	138945	1573 (3.20-3.16)
RSRZ outliers	127900	1423 (3.20-3.16)

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	319	<div> <div>24%</div> <div>13%</div> <div>•</div> <div>59%</div> </div>
2	W	151	<div> <div>%</div> <div>60%</div> <div>28%</div> <div>5%</div> <div>7%</div> </div>
3	B	90	<div> <div>68%</div> <div>22%</div> <div>•</div> <div>7%</div> </div>
3	C	90	<div> <div>4%</div> <div>64%</div> <div>24%</div> <div>6%</div> <div>6%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3468 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 CheA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	131	Total	C	N	O	S	0	1	0
			1027	661	172	193	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	353	GLY	-	EXPRESSION TAG	UNP Q56310
A	354	SER	-	EXPRESSION TAG	UNP Q56310
A	355	HIS	-	EXPRESSION TAG	UNP Q56310
A	?	-	PHE	DELETION	UNP Q56310

- Molecule 2 is a protein called Chemotaxis protein CheW.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	W	140	Total	C	N	O	S	0	0	0
			1113	716	184	211	2			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	-3	GLY	-	EXPRESSION TAG	UNP Q56311
W	-2	SER	-	EXPRESSION TAG	UNP Q56311
W	-1	HIS	-	EXPRESSION TAG	UNP Q56311
W	0	MET	-	EXPRESSION TAG	UNP Q56311

- Molecule 3 is a protein called Methyl-accepting chemotaxis protein.

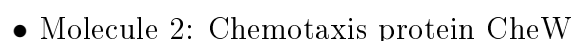
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	84	Total	C	N	O	S	0	2	0
			667	407	125	133	2			
3	C	85	Total	C	N	O	S	0	0	0
			661	403	121	135	2			

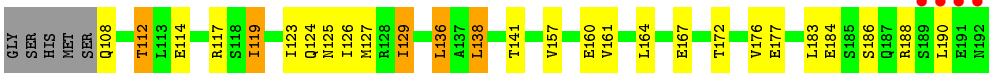
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	103	GLY	-	EXPRESSION TAG	UNP Q7DFA3
B	104	SER	-	EXPRESSION TAG	UNP Q7DFA3
B	105	HIS	-	EXPRESSION TAG	UNP Q7DFA3
B	106	MET	-	EXPRESSION TAG	UNP Q7DFA3
C	103	GLY	-	EXPRESSION TAG	UNP Q7DFA3
C	104	SER	-	EXPRESSION TAG	UNP Q7DFA3
C	105	HIS	-	EXPRESSION TAG	UNP Q7DFA3
C	106	MET	-	EXPRESSION TAG	UNP Q7DFA3



- Molecule 1: Chemotaxis protein CheA





## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	213.59Å 213.59Å 208.84Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.16 – 3.19 46.16 – 3.19	Depositor EDS
% Data completeness (in resolution range)	99.4 (46.16-3.19) 93.3 (46.16-3.19)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.49 (at 3.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.196 , 0.220 0.200 , 0.221	Depositor DCC
$R_{free}$ test set	2000 reflections (6.55%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	91.6	Xtriage
Anisotropy	0.367	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 23.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3468	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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	0.60	0/1040	0.76	0/1407
2	W	0.51	0/1124	0.66	0/1512
3	B	0.47	0/675	0.60	0/904
3	C	0.51	0/661	0.63	0/887
All	All	0.53	0/3500	0.67	0/4710

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 [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	1027	0	1088	29	1
2	W	1113	0	1177	26	1
3	B	667	0	691	19	0
3	C	661	0	679	15	0
All	All	3468	0	3635	76	1

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 11.

All (76) 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)
3:B:157:VAL:HG23	3:C:136:LEU:HD12	1.65	0.78
2:W:110:ASP:HB2	2:W:127:LYS:HE2	1.67	0.77
1:A:560:ILE:HD12	1:A:632:GLN:HE21	1.57	0.69
3:C:114:GLU:OE1	3:C:117:ARG:NH2	2.26	0.68
1:A:569:ILE:HD13	1:A:583:ILE:HD13	1.77	0.67
1:A:569:ILE:HG21	1:A:583:ILE:HG21	1.77	0.67
3:B:128[B]:ARG:NH1	3:B:132:GLU:OE1	2.27	0.67
1:A:544:ILE:HD11	1:A:559:PRO:HB3	1.81	0.62
1:A:619:ASN:N	1:A:619:ASN:OD1	2.33	0.61
1:A:565:THR:OG1	1:A:566:ILE:N	2.32	0.61
2:W:46:ARG:NH2	2:W:145:ILE:O	2.34	0.61
2:W:105:THR:O	2:W:107:ASN:N	2.34	0.60
1:A:598:VAL:HG11	1:A:670:ILE:HG23	1.84	0.60
1:A:568:SER:OG	3:B:131[A]:ARG:NH2	2.32	0.59
3:B:178:LYS:HB3	3:C:119:ILE:HG12	1.84	0.59
3:C:125:ASN:O	3:C:129:ILE:HG13	2.04	0.57
2:W:26:ASP:OD2	2:W:135:TYR:OH	2.20	0.57
1:A:560:ILE:HD11	3:B:142:ILE:HG12	1.86	0.57
1:A:598:VAL:HG12	1:A:599:LEU:HD13	1.86	0.56
3:C:157:VAL:O	3:C:161:VAL:HG23	2.05	0.56
3:B:142:ILE:HG22	3:B:146:ARG:HH21	1.71	0.56
3:B:123:ILE:O	3:B:127:MET:HG3	2.06	0.55
2:W:142:ILE:O	2:W:146:THR:HG23	2.07	0.54
1:A:644:PHE:O	1:A:647:VAL:HG23	2.08	0.53
3:B:182:ILE:CD1	3:C:119:ILE:HG13	2.39	0.53
1:A:661:ILE:HD11	2:W:54:ASN:HB2	1.89	0.53
1:A:658:ASP:OD2	1:A:658:ASP:N	2.42	0.53
3:B:182:ILE:HD11	3:C:119:ILE:HG13	1.91	0.52
2:W:118:PHE:O	2:W:121:LYS:HB2	2.09	0.52
3:B:112:THR:O	3:B:116:ILE:HG13	2.11	0.50
2:W:121:LYS:HD2	2:W:144:GLU:OE2	2.13	0.49
1:A:590:ILE:HG23	1:A:622:TYR:HA	1.93	0.49
2:W:86:THR:HG21	2:W:139:ASP:HA	1.95	0.49
2:W:17:GLU:O	2:W:96:ASP:HB2	2.13	0.49
3:C:108:GLN:O	3:C:112:THR:HG22	2.12	0.49
2:W:121:LYS:HG2	2:W:141:ILE:HG12	1.96	0.48
2:W:49:VAL:HG12	2:W:63:VAL:HG12	1.96	0.47
2:W:142:ILE:O	2:W:145:ILE:HG22	2.14	0.47
3:B:164:LEU:HD13	3:C:129:ILE:HG22	1.96	0.46
2:W:19:ASP:O	2:W:20:GLU:HG2	2.16	0.46
2:W:131:ARG:C	2:W:131:ARG:HD2	2.36	0.46
2:W:105:THR:C	2:W:107:ASN:H	2.18	0.45

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:W:105:THR:HG23	2:W:108:GLN:HG2	1.98	0.45
3:C:138:LEU:HD13	3:C:138:LEU:HA	1.78	0.45
1:A:642:LYS:HB2	1:A:642:LYS:HE3	1.85	0.45
1:A:564:ASP:OD2	1:A:565:THR:HG22	2.17	0.45
3:C:186:SER:O	3:C:190:LEU:HB2	2.16	0.45
3:B:185:SER:HA	3:C:108:GLN:HG2	1.99	0.45
2:W:9:LYS:HD2	2:W:10:GLU:H	1.81	0.45
1:A:580[B]:ARG:HG3	1:A:581:ASP:N	2.32	0.44
2:W:146:THR:OG1	2:W:147:VAL:N	2.48	0.44
3:C:183:LEU:HD23	3:C:183:LEU:HA	1.84	0.44
1:A:596:TRP:CD1	1:A:597:GLU:HG3	2.53	0.44
2:W:63:VAL:HG11	2:W:145:ILE:HD11	2.00	0.44
1:A:570:SER:HB3	1:A:609:GLU:OE1	2.18	0.44
1:A:541:LEU:HD13	1:A:543:ILE:HG12	2.00	0.43
1:A:600:GLN:NE2	2:W:44:LYS:HE3	2.34	0.43
3:B:128[B]:ARG:HH22	3:C:167:GLU:CD	2.22	0.42
1:A:629:LEU:HD13	1:A:629:LEU:HA	1.66	0.42
3:B:177:GLU:O	3:B:181:GLU:HG3	2.20	0.42
2:W:11:PHE:HB3	2:W:104:ILE:HG13	2.01	0.42
2:W:88:ASP:N	2:W:88:ASP:OD1	2.49	0.42
2:W:36:LYS:HD2	2:W:94:LEU:HD11	2.02	0.42
1:A:568:SER:CB	3:B:131[A]:ARG:HH22	2.31	0.42
3:B:121:LYS:O	3:B:124:GLN:HB2	2.19	0.42
3:B:119:ILE:O	3:B:123:ILE:HG22	2.20	0.41
1:A:604:LYS:O	1:A:606:GLU:N	2.53	0.41
3:B:108:GLN:NE2	3:C:188:ARG:HD2	2.34	0.41
2:W:73:PHE:HE1	2:W:75:GLU:HG2	1.85	0.41
1:A:670:ILE:HG22	1:A:670:ILE:O	2.21	0.41
2:W:34:ILE:HG12	2:W:81:ILE:HB	2.02	0.41
3:B:146:ARG:HE	3:B:146:ARG:HB2	1.50	0.41
1:A:579:ASP:OD1	1:A:579:ASP:N	2.43	0.41
1:A:580[B]:ARG:HD3	1:A:593:TYR:CE1	2.57	0.40
1:A:543:ILE:HG22	1:A:544:ILE:N	2.36	0.40
1:A:549:VAL:HG23	1:A:629:LEU:HD13	2.03	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:577:VAL:O	2:W:21:GLN:NE2[3_555]	2.11	0.09

## 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	130/319 (41%)	109 (84%)	18 (14%)	3 (2%)	6	32
2	W	138/151 (91%)	125 (91%)	9 (6%)	4 (3%)	4	26
3	B	84/90 (93%)	82 (98%)	2 (2%)	0	100	100
3	C	83/90 (92%)	76 (92%)	7 (8%)	0	100	100
All	All	435/650 (67%)	392 (90%)	36 (8%)	7 (2%)	9	41

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	605	GLU
2	W	106	GLU
1	A	645	SER
2	W	19	ASP
2	W	78	MET
1	A	603	HIS
2	W	9	LYS

### 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	115/279 (41%)	88 (76%)	27 (24%)	1	3
2	W	128/136 (94%)	113 (88%)	15 (12%)	5	22
3	B	73/76 (96%)	66 (90%)	7 (10%)	8	30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	72/76 (95%)	56 (78%)	16 (22%)	1	4
All	All	388/567 (68%)	323 (83%)	65 (17%)	2	9

All (65) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	541	LEU
1	A	544	ILE
1	A	545	GLN
1	A	547	LEU
1	A	549	VAL
1	A	560	ILE
1	A	565	THR
1	A	567	LEU
1	A	569	ILE
1	A	570	SER
1	A	571	LYS
1	A	573	ASP
1	A	576	ARG
1	A	580[A]	ARG
1	A	580[B]	ARG
1	A	583	ILE
1	A	586	ARG
1	A	599	LEU
1	A	619	ASN
1	A	620	ARG
1	A	628	ASP
1	A	629	LEU
1	A	634	ASP
1	A	636	VAL
1	A	640	LEU
1	A	642	LYS
1	A	643	VAL
2	W	8	LEU
2	W	11	PHE
2	W	12	GLU
2	W	25	PHE
2	W	26	ASP
2	W	27	VAL
2	W	50	GLU
2	W	55	LEU
2	W	105	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	W	126	VAL
2	W	131	ARG
2	W	134	ILE
2	W	140	LYS
2	W	142	ILE
2	W	145	ILE
3	B	111	GLU
3	B	123	ILE
3	B	131[A]	ARG
3	B	131[B]	ARG
3	B	160	GLU
3	B	164	LEU
3	B	168	THR
3	C	112	THR
3	C	119	ILE
3	C	123	ILE
3	C	124	GLN
3	C	126	ILE
3	C	127	MET
3	C	129	ILE
3	C	136	LEU
3	C	138	LEU
3	C	141	THR
3	C	160	GLU
3	C	164	LEU
3	C	172	THR
3	C	176	VAL
3	C	177	GLU
3	C	184	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	545	GLN
1	A	600	GLN
1	A	632	GLN
3	B	108	GLN

### 5.3.3 RNA ⓘ

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	131/319 (41%)	-0.17	1 (0%) 86 77	6, 25, 64, 83	0
2	W	140/151 (92%)	-0.17	1 (0%) 87 81	8, 31, 78, 93	0
3	B	84/90 (93%)	-0.09	0 100 100	14, 30, 65, 89	0
3	C	85/90 (94%)	0.19	4 (4%) 31 18	12, 35, 106, 132	0
All	All	440/650 (67%)	-0.09	6 (1%) 75 63	6, 30, 78, 132	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	192	ASN	6.2
1	A	541	LEU	5.6
3	C	190	LEU	5.6
3	C	189	SER	5.1
2	W	130	GLY	3.3
3	C	191	GLU	3.1

### 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.