



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 10:38 AM GMT

PDB ID : 3MM4  
Title : Crystal structure of the receiver domain of the histidine kinase CKI1 from *Arabidopsis thaliana*  
Authors : Marek, J.; Klumpler, T.; Pekarova, B.; Triskova, O.; Horak, J.; Zidek, L.; Dopitova, R.; Hejatk, J.; Janda, L.  
Deposited on : 2010-04-19  
Resolution : 2.00 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  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

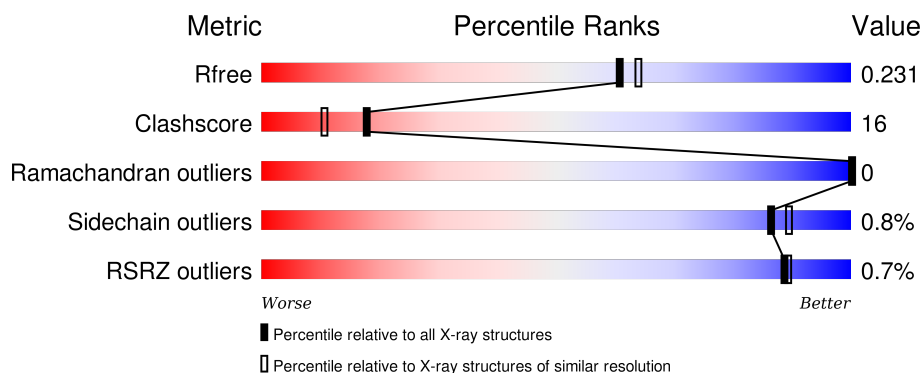
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	206	 48% 21% .. 30%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1293 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 Histidine kinase homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	145	Total	C	N	O	S	0	7	0
			1161	723	202	229	7			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	925	GLY	-	EXPRESSION TAG	UNP O22267
A	926	SER	-	EXPRESSION TAG	UNP O22267
A	927	SER	-	EXPRESSION TAG	UNP O22267
A	928	HIS	-	EXPRESSION TAG	UNP O22267
A	929	HIS	-	EXPRESSION TAG	UNP O22267
A	930	HIS	-	EXPRESSION TAG	UNP O22267
A	931	HIS	-	EXPRESSION TAG	UNP O22267
A	932	HIS	-	EXPRESSION TAG	UNP O22267
A	933	HIS	-	EXPRESSION TAG	UNP O22267
A	934	SER	-	EXPRESSION TAG	UNP O22267
A	935	SER	-	EXPRESSION TAG	UNP O22267
A	936	GLY	-	EXPRESSION TAG	UNP O22267
A	937	LEU	-	EXPRESSION TAG	UNP O22267
A	938	VAL	-	EXPRESSION TAG	UNP O22267
A	939	PRO	-	EXPRESSION TAG	UNP O22267
A	940	ARG	-	EXPRESSION TAG	UNP O22267
A	941	GLY	-	EXPRESSION TAG	UNP O22267
A	942	SER	-	EXPRESSION TAG	UNP O22267
A	943	HIS	-	EXPRESSION TAG	UNP O22267
A	1123	LEU	-	EXPRESSION TAG	UNP O22267
A	1124	GLU	-	EXPRESSION TAG	UNP O22267
A	1125	HIS	-	EXPRESSION TAG	UNP O22267
A	1126	HIS	-	EXPRESSION TAG	UNP O22267
A	1127	HIS	-	EXPRESSION TAG	UNP O22267
A	1128	HIS	-	EXPRESSION TAG	UNP O22267
A	1129	HIS	-	EXPRESSION TAG	UNP O22267
A	1130	HIS	-	EXPRESSION TAG	UNP O22267

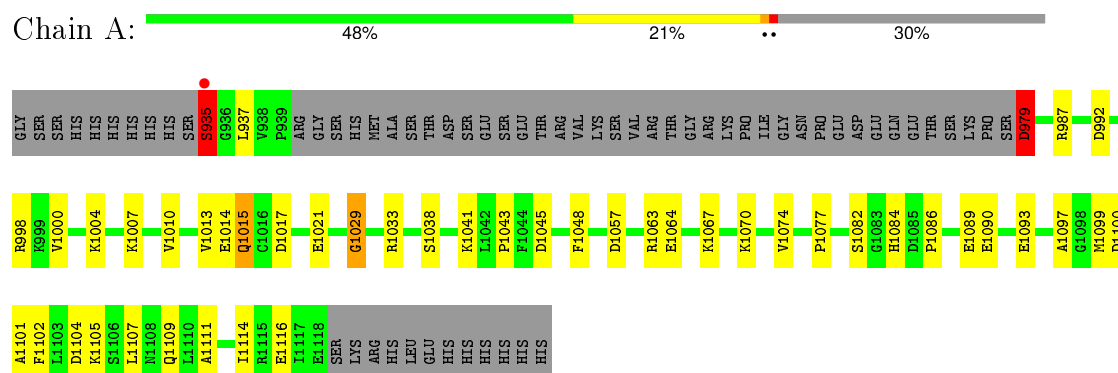
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	132	Total 132	O 132	0	0

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

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Histidine kinase homolog



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.46Å 99.83Å 79.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	5.00 – 2.00 5.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (5.00-2.00) 99.9 (5.00-2.00)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.26 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.176 , 0.229 0.176 , 0.231	Depositor DCC
$R_{free}$ test set	706 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtriage
Anisotropy	0.086	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.96 , 102.3	EDS
Estimated twinning fraction	0.019 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.035 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 14063 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1293	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows:

<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, 0.375 respectively for untwinned datasets, and 0.333, 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	1.49	9/1206 (0.7%)	1.19	9/1615 (0.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1064	GLU	CG-CD	6.21	1.61	1.51
1	A	1029	GLY	N-CA	6.19	1.55	1.46
1	A	1104	ASP	CB-CG	-5.70	1.39	1.51
1	A	1000	VAL	CB-CG1	5.38	1.64	1.52
1	A	1015	GLN	CG-CD	5.32	1.63	1.51
1	A	1101	ALA	CA-CB	5.29	1.63	1.52
1	A	1111	ALA	CA-CB	5.27	1.63	1.52
1	A	1099	MET	CB-CG	5.16	1.67	1.51
1	A	1074	VAL	CB-CG2	5.11	1.63	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	979	ASP	CB-CG-OD2	8.54	125.98	118.30
1	A	979	ASP	N-CA-C	6.48	128.48	111.00
1	A	1057	ASP	CB-CG-OD1	6.09	123.78	118.30
1	A	1057	ASP	CB-CG-OD2	-5.99	112.91	118.30
1	A	992	ASP	CB-CG-OD2	-5.66	113.20	118.30
1	A	1045	ASP	CB-CG-OD1	-5.51	113.34	118.30
1	A	1017	ASP	CB-CG-OD1	5.30	123.07	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	935	SER	N-CA-C	5.24	125.15	111.00
1	A	998	ARG	NE-CZ-NH1	-5.12	117.74	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	935	SER	Peptide

## 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	1161	0	1172	37	0
2	A	132	0	0	11	0
All	All	1293	0	1172	37	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 16.

All (37) 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:1038:SER:O	1:A:1041[B]:LYS:NZ	1.99	0.94
1:A:1105:LYS:HE2	2:A:60:HOH:O	1.67	0.93
1:A:1010[A]:VAL:HG11	1:A:1013:VAL:HG22	1.61	0.80
1:A:1004:LYS:O	1:A:1007[B]:LYS:HG3	1.83	0.78
1:A:1086:PRO:HD3	1:A:1102:PHE:CZ	2.20	0.76
1:A:1014:GLU:HB2	2:A:72:HOH:O	1.92	0.69
1:A:1004:LYS:O	1:A:1007[B]:LYS:CG	2.41	0.68
1:A:987[B]:ARG:NH1	1:A:1043:PRO:O	2.26	0.68
1:A:1084:HIS:HE1	1:A:1090:GLU:OE2	1.78	0.66
1:A:979:ASP:OD1	2:A:112:HOH:O	2.15	0.65
1:A:937:LEU:HD11	1:A:1116:GLU:HB3	1.80	0.64
1:A:987[B]:ARG:NH1	1:A:1043:PRO:C	2.51	0.63

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1086:PRO:HD3	1:A:1102:PHE:CE1	2.36	0.60
1:A:1105:LYS:HD3	2:A:17:HOH:O	2.05	0.56
1:A:1041[A]:LYS:HD2	2:A:91:HOH:O	2.06	0.55
1:A:1010[A]:VAL:HG11	1:A:1013:VAL:CG2	2.34	0.54
1:A:1089:GLU:O	1:A:1093:GLU:HG3	2.08	0.53
1:A:1014:GLU:HG3	1:A:1015:GLN:N	2.23	0.52
1:A:1063:ARG:HG3	1:A:1097:ALA:HA	1.91	0.51
1:A:987[B]:ARG:HH12	1:A:1043:PRO:C	2.12	0.51
1:A:937:LEU:CD1	1:A:1116:GLU:HB3	2.42	0.50
1:A:1107:LEU:HD12	2:A:85:HOH:O	2.11	0.49
1:A:1084:HIS:CE1	1:A:1090:GLU:OE2	2.63	0.48
1:A:1004:LYS:HA	1:A:1007[B]:LYS:HG2	1.96	0.48
1:A:1102:PHE:CE1	2:A:116:HOH:O	2.66	0.47
1:A:1048:PHE:HZ	1:A:1114:ILE:HD11	1.81	0.45
1:A:1067:LYS:HA	1:A:1070:LYS:HD2	1.99	0.45
1:A:1077:PRO:HA	1:A:1100:ASP:OD2	2.17	0.44
1:A:1105:LYS:NZ	2:A:7:HOH:O	2.50	0.44
1:A:1029:GLY:O	1:A:1033:ARG:HG2	2.18	0.44
1:A:1082:SER:HB3	1:A:1102:PHE:CZ	2.53	0.43
1:A:935:SER:N	2:A:81:HOH:O	2.51	0.43
1:A:1014:GLU:CD	2:A:72:HOH:O	2.57	0.43
1:A:1004:LYS:O	1:A:1007[B]:LYS:HG2	2.19	0.42
1:A:987[A]:ARG:NH1	1:A:1043:PRO:O	2.52	0.42
1:A:935:SER:HB2	1:A:1109:GLN:CB	2.50	0.41
1:A:1021:GLU:OE2	2:A:86:HOH:O	2.21	0.40

There are no symmetry-related clashes.

## 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	148/206 (72%)	147 (99%)	1 (1%)	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	132/181 (73%)	131 (99%)	1 (1%)	86 89

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	979	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	A	1084	HIS

### 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

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	145/206 (70%)	-0.45	1 (0%) 89 89	12, 25, 38, 41	1 (0%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	935	SER	2.9

### 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.