



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2024 – 09:13 PM EST

PDB ID : 4HZS  
Title : Crystal structure of Ack1 kinase domain with C-terminal SH3 domain  
Authors : Gajiwala, K.S.  
Deposited on : 2012-11-15  
Resolution : 3.23 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 : 2.36  
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) : 2.36

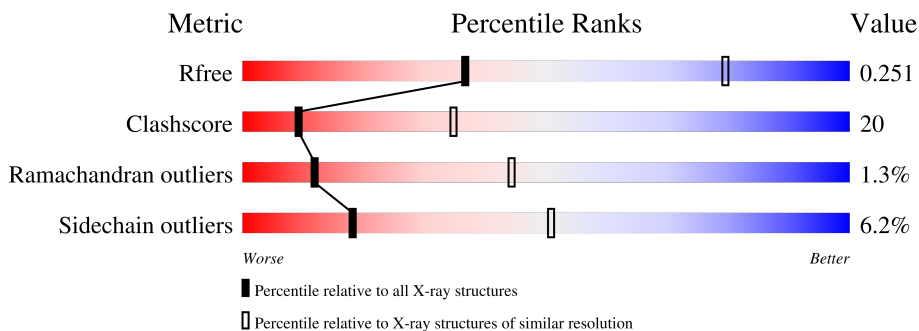
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.23 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	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)

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	341	58% 32% 6%
1	B	341	58% 32% 6%
1	C	341	58% 32% 6%
1	D	341	58% 32% 6%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 10196 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.

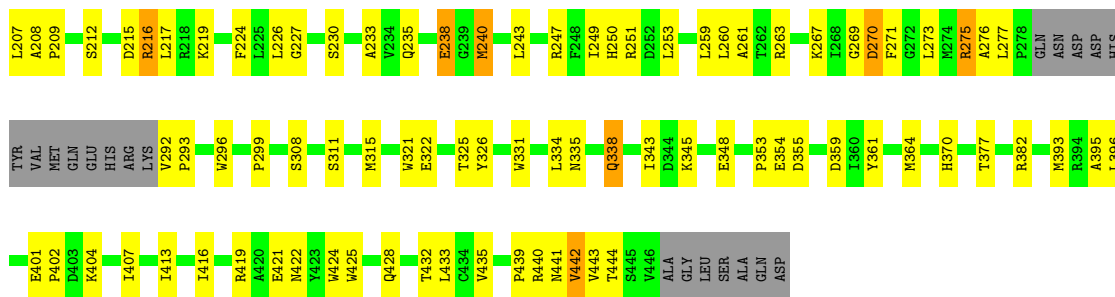
- Molecule 1 is a protein called Activated CDC42 kinase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	319	2549	1624	451	457	17	0	0	0
1	B	319	2549	1624	451	457	17	0	0	0
1	C	319	2549	1624	451	457	17	0	0	0
1	D	319	2549	1624	451	457	17	0	0	0

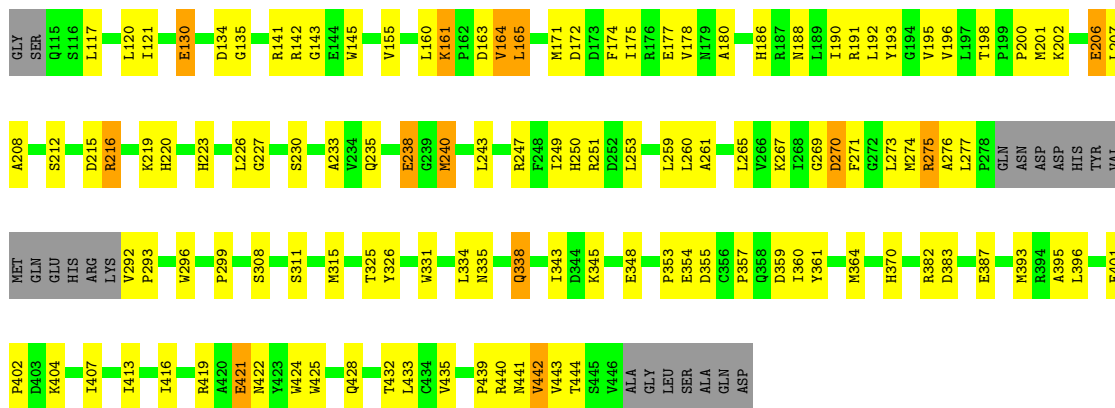
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	113	GLY	-	expression tag	UNP Q07912
A	114	SER	-	expression tag	UNP Q07912
B	113	GLY	-	expression tag	UNP Q07912
B	114	SER	-	expression tag	UNP Q07912
C	113	GLY	-	expression tag	UNP Q07912
C	114	SER	-	expression tag	UNP Q07912
D	113	GLY	-	expression tag	UNP Q07912
D	114	SER	-	expression tag	UNP Q07912





● Molecule 1: Activated CDC42 kinase 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	145.46Å 145.46Å 103.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	73.04 – 3.23 73.04 – 3.23	Depositor EDS
% Data completeness (in resolution range)	99.5 (73.04-3.23) 99.4 (73.04-3.23)	Depositor EDS
$R_{merge}$	0.19	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 3.26Å)	Xtrriage
Refinement program	CNX	Depositor
R, $R_{free}$	0.246 , 0.270 0.230 , 0.251	Depositor DCC
$R_{free}$ test set	1724 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.5	Xtrriage
Anisotropy	0.172	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , -7.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.168 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	10196	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.20% 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	0.45	0/2610	0.55	0/3539
1	B	0.45	0/2610	0.55	0/3539
1	C	0.45	0/2610	0.55	0/3539
1	D	0.44	0/2610	0.55	0/3539
All	All	0.45	0/10440	0.55	0/14156

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	2549	0	2540	103	0
1	B	2549	0	2540	99	0
1	C	2549	0	2540	110	0
1	D	2549	0	2540	105	0
All	All	10196	0	10160	405	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 20.

The worst 5 of 405 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:C:165:LEU:H	1:C:165:LEU:HD23	1.30	0.96
1:D:165:LEU:HD23	1:D:165:LEU:H	1.31	0.95
1:B:165:LEU:HD23	1:B:165:LEU:H	1.30	0.95
1:A:165:LEU:H	1:A:165:LEU:HD23	1.31	0.94
1:C:117:LEU:HD12	1:C:175:ILE:HG23	1.58	0.85

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	315/341 (92%)	283 (90%)	28 (9%)	4 (1%)	12	44
1	B	315/341 (92%)	283 (90%)	28 (9%)	4 (1%)	12	44
1	C	315/341 (92%)	282 (90%)	29 (9%)	4 (1%)	12	44
1	D	315/341 (92%)	283 (90%)	28 (9%)	4 (1%)	12	44
All	All	1260/1364 (92%)	1131 (90%)	113 (9%)	16 (1%)	12	44

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	206	GLU
1	B	206	GLU
1	C	206	GLU
1	D	206	GLU
1	B	130	GLU

#### 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	276/296 (93%)	259 (94%)	17 (6%)	18	51
1	B	276/296 (93%)	259 (94%)	17 (6%)	18	51
1	C	276/296 (93%)	258 (94%)	18 (6%)	17	49
1	D	276/296 (93%)	260 (94%)	16 (6%)	20	53
All	All	1104/1184 (93%)	1036 (94%)	68 (6%)	18	51

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	207	LEU
1	D	219	LYS
1	D	359	ASP
1	B	219	LYS
1	B	216	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	186	HIS
1	C	422	ASN
1	C	408	GLN
1	D	115	GLN
1	B	115	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 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.