



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 01:08 am BST

PDB ID : 5UPL  
Title : CDC42 binds PAK4 via an extended GTPase-effector interface - 2 peptide:  
PAK4FL, CDC42 - UNREFINED  
Authors : Ha, B.H.; Boggon, T.J.  
Deposited on : 2017-02-03  
Resolution : 3.00 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 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 : 2.11  
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

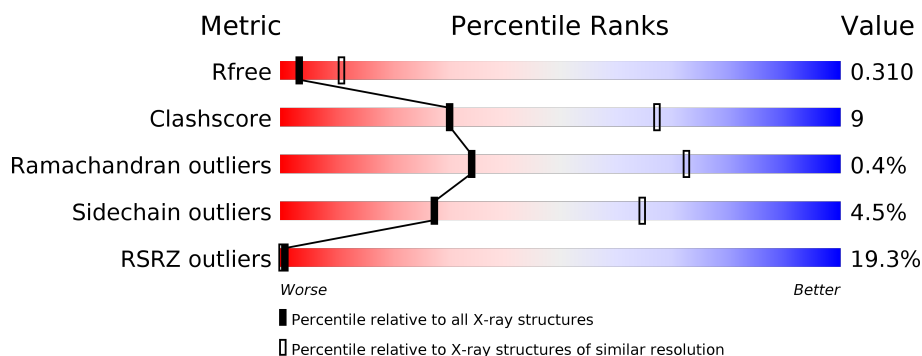
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	453	
2	B	185	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3639 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 Serine/threonine-protein kinase PAK 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	291	2302	1467	406	414	1	14	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	139	MET	-	initiating methionine	UNP O96013
A	140	GLY	-	expression tag	UNP O96013
A	141	SER	-	expression tag	UNP O96013
A	142	SER	-	expression tag	UNP O96013
A	143	HIS	-	expression tag	UNP O96013
A	144	HIS	-	expression tag	UNP O96013
A	145	HIS	-	expression tag	UNP O96013
A	146	HIS	-	expression tag	UNP O96013
A	147	HIS	-	expression tag	UNP O96013
A	148	HIS	-	expression tag	UNP O96013
A	149	SER	-	expression tag	UNP O96013
A	150	SER	-	expression tag	UNP O96013
A	151	GLY	-	expression tag	UNP O96013
A	152	LEU	-	expression tag	UNP O96013
A	153	VAL	-	expression tag	UNP O96013
A	154	PRO	-	expression tag	UNP O96013
A	155	ARG	-	expression tag	UNP O96013
A	156	GLY	-	expression tag	UNP O96013
A	157	SER	-	expression tag	UNP O96013
A	158	HIS	-	expression tag	UNP O96013
A	159	MET	-	expression tag	UNP O96013
A	160	GLU	-	expression tag	UNP O96013
A	161	ASN	-	expression tag	UNP O96013
A	162	LEU	-	expression tag	UNP O96013
A	163	TYR	-	expression tag	UNP O96013
A	164	PHE	-	expression tag	UNP O96013
A	165	GLN	-	expression tag	UNP O96013

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	166	GLY	-	expression tag	UNP O96013

- Molecule 2 is a protein called Cell division control protein 42 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	176	Total	C	N	O	S	0	0	0
			1337	865	209	257	6			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	178	LEU	-	expression tag	UNP P60953
B	179	GLU	-	expression tag	UNP P60953
B	180	HIS	-	expression tag	UNP P60953
B	181	HIS	-	expression tag	UNP P60953
B	182	HIS	-	expression tag	UNP P60953
B	183	HIS	-	expression tag	UNP P60953
B	184	HIS	-	expression tag	UNP P60953
B	185	HIS	-	expression tag	UNP P60953



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	141.65Å 141.65Å 62.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.64 – 3.00 43.64 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (43.64-3.00) 99.8 (43.64-3.00)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.55 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.10_2155	Depositor
R, $R_{free}$	0.257 , 0.310 0.257 , 0.310	Depositor DCC
$R_{free}$ test set	726 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	94.6	Xtriage
Anisotropy	0.012	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 53.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.038 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3639	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	128.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% 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](#)

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

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	0.57	0/2339	0.81	2/3167 (0.1%)
2	B	0.61	0/1367	0.68	1/1869 (0.1%)
All	All	0.59	0/3706	0.77	3/5036 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	537	ASN	N-CA-CB	7.21	123.58	110.60
1	A	589	ARG	N-CA-C	5.53	125.92	111.00
2	B	149	LEU	N-CA-C	5.39	125.56	111.00

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	2302	0	2360	48	0
2	B	1337	0	1308	27	0
All	All	3639	0	3668	68	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 9.

The worst 5 of 68 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:370:MET:HB3	1:A:381:MET:HB3	1.32	1.11
1:A:370:MET:CB	1:A:381:MET:HB3	2.09	0.81
1:A:370:MET:HB3	1:A:381:MET:CB	2.14	0.77
1:A:379:VAL:HG22	1:A:398:LEU:HD11	1.67	0.75
2:B:13:ALA:O	2:B:15:GLY:N	2.23	0.72

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	288/453 (64%)	270 (94%)	18 (6%)	0	100	100
2	B	174/185 (94%)	155 (89%)	17 (10%)	2 (1%)	14	50
All	All	462/638 (72%)	425 (92%)	35 (8%)	2 (0%)	34	72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	14	VAL
2	B	106	PRO

### 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	253/394 (64%)	241 (95%)	12 (5%)	26	63
2	B	144/165 (87%)	138 (96%)	6 (4%)	30	66
All	All	397/559 (71%)	379 (96%)	18 (4%)	27	64

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	516	PHE
1	A	522	LYS
2	B	65	ASP
1	A	415	GLU
1	A	488	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	365	ASN
2	B	103	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	SEP	A	474	1	8,9,10	0.68	0	8,12,14	0.92	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	SEP	A	474	1	-	2/5/8/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	474	SEP	N-CA-CB-OG
1	A	474	SEP	CA-CB-OG-P

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	474	SEP	1	0

## 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	290/453 (64%)	-0.14	3 (1%) 82 59	41, 70, 107, 133	0
2	B	176/185 (95%)	2.74	87 (49%) 0 0	223, 223, 223, 223	0
All	All	466/638 (73%)	0.95	90 (19%) 1 0	41, 87, 223, 223	0

The worst 5 of 90 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	136	PRO	17.0
2	B	61	GLN	11.3
2	B	67	LEU	10.9
2	B	70	LEU	9.8
2	B	85	VAL	9.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
1	SEP	A	474	10/11	0.95	0.31	57,68,74,75	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.