Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 03:55 pm GMT

PDB ID : 4ZJ7

Title : Crystal structure of the karyopherin Kap121p bound to the extreme C-terminus of the protein phosphatase Cdc14p

Authors : Kobayashi, J.; Hirano, H.; Matsuura, Y.

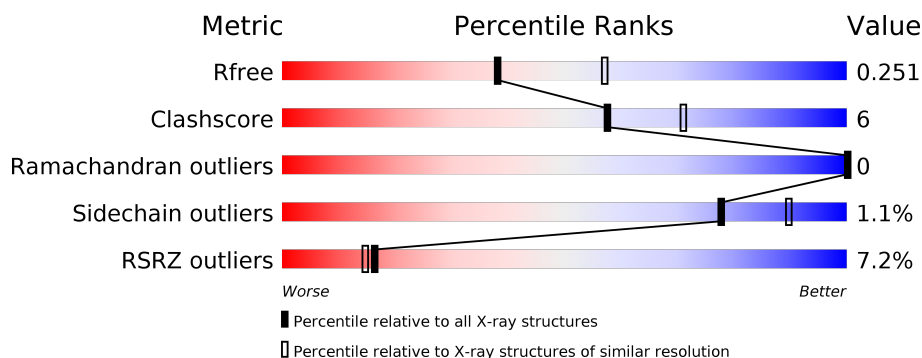Deposited on : 2015-04-29

Resolution : 2.40 Å(reported)

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	Percentile Ranks	Value
Rfree		0.251
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		1.1%
RSRZ outliers		7.2%

Worse ← → Better

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (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 on the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	1078	7% / 82% / 12% / 6%
2	B	35	11% / · / 86%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7947 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 Importin subunit beta-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1017	Total 7775	C 4997	N 1248	O 1495	S 35	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	PRO	deletion	UNP P32337
A	?	-	SER	deletion	UNP P32337
A	?	-	SER	deletion	UNP P32337
A	?	-	LYS	deletion	UNP P32337
A	?	-	LEU	deletion	UNP P32337
A	?	-	MET	deletion	UNP P32337
A	?	-	ILE	deletion	UNP P32337
A	?	-	MET	deletion	UNP P32337
A	?	-	SER	deletion	UNP P32337
A	?	-	LYS	deletion	UNP P32337
A	?	-	ASN	deletion	UNP P32337

- Molecule 2 is a protein called Tyrosine-protein phosphatase CDC14.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	5	Total 33	C 20	N 6	O 7	0	0	0

- Molecule 3 is water.

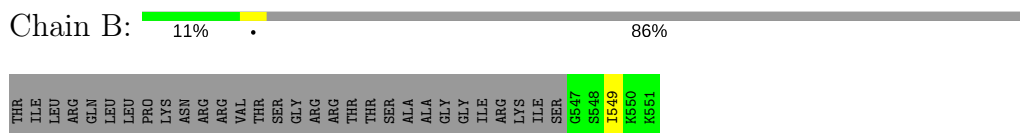| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	138	Total 138	O 138	0	0
3	B	1	Total 1	O 1	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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 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: Importin subunit beta-3

• Molecule 2: Tyrosine-protein phosphatase CDC14

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	77.26Å 127.76Å 131.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.81 – 2.40 26.25 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.4 (25.81-2.40) 99.4 (26.25-2.40)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.64 (at 2.39Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.219 , 0.252 0.221 , 0.251	Depositor DCC
R_{free} test set	2594 reflections (5.05%)	DCC
Wilson B-factor (Å^2)	46.5	Xtriage
Anisotropy	0.402	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 43.0	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.012 for -h,l,k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7947	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.23	0/7909	0.41	0/10774
2	B	0.19	0/32	0.31	0/39
All	All	0.23	0/7941	0.41	0/10813

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	7775	0	7668	87	0
2	B	33	0	33	1	0
3	A	138	0	0	3	0
3	B	1	0	0	0	0
All	All	7947	0	7701	87	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 6.

All (87) 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:929:GLN:HG3	1:A:973:LYS:HD2	1.38	1.03

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1009:TYR:HB3	1:A:1053:VAL:HG21	1.60	0.81
1:A:417:ILE:O	1:A:458:LYS:NZ	2.12	0.80
1:A:636:GLN:HE22	1:A:713:LEU:HB2	1.57	0.69
1:A:406:ILE:O	3:A:1101:HOH:O	2.12	0.67
1:A:206:LYS:NZ	1:A:248:GLU:OE2	2.28	0.66
1:A:946:LEU:HD23	1:A:949:ILE:HD12	1.79	0.64
1:A:430:ASN:HA	2:B:549:ILE:HD13	1.79	0.64
1:A:648:ASN:O	1:A:651:GLN:NE2	2.28	0.61
1:A:973:LYS:HE2	1:A:1007:PHE:CD2	2.36	0.59
1:A:683:GLN:NE2	3:A:1110:HOH:O	2.34	0.59
1:A:1066:SER:O	1:A:1070:MET:HG2	2.04	0.57
1:A:992:TRP:O	1:A:995:THR:HG22	2.04	0.57
1:A:973:LYS:HE2	1:A:1007:PHE:CE2	2.40	0.56
1:A:929:GLN:HG3	1:A:973:LYS:CD	2.25	0.56
1:A:706:LEU:HD21	1:A:750:ALA:HA	1.87	0.56
1:A:689:LEU:HB3	1:A:692:GLN:HB2	1.87	0.56
1:A:335:ASP:HB3	1:A:337:THR:H	1.72	0.54
1:A:1061:LEU:HD21	1:A:1072:ILE:HD12	1.89	0.54
1:A:142:GLN:O	1:A:146:GLU:HG3	2.09	0.53
1:A:947:VAL:HG13	1:A:991:ASN:HD21	1.73	0.53
1:A:872:GLU:O	1:A:876:VAL:HG23	2.09	0.53
1:A:943:LEU:HD22	1:A:974:ILE:HG21	1.90	0.53
1:A:857:LEU:HD23	1:A:886:LEU:HD21	1.90	0.53
1:A:747:TRP:HE1	1:A:785:CYS:HB2	1.74	0.53
1:A:883:ILE:HA	1:A:886:LEU:HD12	1.91	0.53
1:A:112:SER:O	1:A:119:ARG:NH2	2.42	0.52
1:A:709:LEU:HD12	1:A:753:LYS:HG2	1.92	0.52
1:A:92:THR:OG1	1:A:133:ASP:OD2	2.28	0.51
1:A:294:GLU:HG2	1:A:358:LYS:HE3	1.92	0.51
1:A:561:ALA:HB1	1:A:580:LEU:HD21	1.92	0.51
1:A:1013:SER:HB2	1:A:1053:VAL:HG13	1.92	0.50
1:A:244:GLU:HA	1:A:247:ILE:HD12	1.94	0.50
1:A:417:ILE:HG13	1:A:458:LYS:NZ	2.27	0.49
1:A:144:LEU:HD21	1:A:162:ILE:HD12	1.93	0.49
1:A:170:ILE:HG21	1:A:216:LEU:HD11	1.95	0.49
1:A:376:ILE:HD13	1:A:412:MET:HB3	1.93	0.48
1:A:641:ILE:HD12	1:A:666:ALA:HB2	1.95	0.48
1:A:908:CYS:HB3	1:A:916:ILE:HG22	1.95	0.48
1:A:1006:SER:HA	1:A:1045:LEU:HD12	1.94	0.48
1:A:747:TRP:CG	1:A:781:MET:HG3	2.49	0.48
1:A:417:ILE:HG13	1:A:458:LYS:HZ2	1.79	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:835:LEU:HD12	1:A:874:ILE:HG22	1.96	0.47
1:A:902:ILE:N	1:A:903:PRO:HD2	2.29	0.47
1:A:888:GLN:HA	1:A:930:TYR:CE1	2.50	0.47
1:A:170:ILE:HG23	1:A:178:ILE:HG12	1.97	0.46
1:A:999:ILE:HB	1:A:1042:GLU:HG2	1.97	0.46
1:A:657:VAL:HG12	1:A:666:ALA:HA	1.98	0.46
1:A:646:ALA:C	1:A:648:ASN:H	2.18	0.45
1:A:650:GLN:OE1	1:A:656:ASP:HA	2.16	0.45
1:A:641:ILE:HG12	1:A:642:GLU:H	1.82	0.45
1:A:718:ARG:NH1	1:A:761:GLU:OE2	2.50	0.45
1:A:647:ALA:HA	1:A:650:GLN:HB2	1.97	0.44
1:A:715:ASP:HB2	1:A:766:ILE:HD11	1.98	0.44
1:A:152:ASN:HD22	1:A:155:PHE:HD2	1.66	0.44
1:A:945:THR:O	1:A:949:ILE:HG13	2.18	0.44
1:A:246:LEU:O	1:A:250:VAL:HG23	2.18	0.44
1:A:775:VAL:HG22	1:A:841:SER:HA	2.00	0.43
1:A:918:GLN:HB2	1:A:963:SER:HA	2.00	0.43
1:A:152:ASN:HA	1:A:153:PRO:HD2	1.86	0.43
1:A:179:LEU:HB2	1:A:180:PRO:HD3	2.01	0.42
1:A:272:ILE:HG21	1:A:313:VAL:HG13	2.01	0.42
1:A:119:ARG:HD2	1:A:154:ASN:OD1	2.20	0.42
1:A:646:ALA:O	1:A:648:ASN:N	2.43	0.42
1:A:771:HIS:O	1:A:775:VAL:HG23	2.20	0.42
1:A:946:LEU:HB2	1:A:971:ILE:HD11	2.01	0.42
1:A:1060:LEU:HA	1:A:1060:LEU:HD12	1.90	0.42
1:A:511:GLU:HB3	1:A:556:LYS:HD2	2.02	0.42
1:A:967:ALA:O	1:A:971:ILE:HG12	2.20	0.42
1:A:869:LEU:HD23	1:A:904:LYS:HG2	2.01	0.42
1:A:646:ALA:C	1:A:648:ASN:N	2.73	0.42
1:A:1054:ILE:HG23	1:A:1088:PHE:CD2	2.55	0.42
1:A:203:GLY:HA2	1:A:206:LYS:HG2	2.02	0.42
1:A:923:ILE:O	1:A:926:VAL:HG12	2.21	0.41
1:A:237:ASP:OD1	1:A:237:ASP:N	2.53	0.41
1:A:349:ARG:NH1	3:A:1115:HOH:O	2.37	0.41
1:A:570:LYS:NZ	1:A:616:ASP:OD2	2.35	0.41
1:A:298:GLN:H	1:A:298:GLN:HG3	1.73	0.41
1:A:458:LYS:HD2	1:A:463:CYS:SG	2.61	0.41
1:A:667:ILE:HG21	1:A:713:LEU:HD13	2.02	0.41
1:A:874:ILE:HD12	1:A:874:ILE:H	1.86	0.41
1:A:1003:GLU:H	1:A:1003:GLU:CD	2.23	0.41
1:A:317:MET:HB2	1:A:348:ALA:HB2	2.03	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:957:LEU:O	1:A:961:ARG:N	2.54	0.41
1:A:133:ASP:N	1:A:133:ASP:OD1	2.54	0.40
1:A:971:ILE:HG22	1:A:975:LEU:HD12	2.04	0.40
1:A:495:LEU:O	1:A:499:LEU:HG	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	991/1078 (92%)	952 (96%)	39 (4%)	0	100	100
2	B	3/35 (9%)	3 (100%)	0	0	100	100
All	All	994/1113 (89%)	955 (96%)	39 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	834/937 (89%)	825 (99%)	9 (1%)	78	90
2	B	3/29 (10%)	3 (100%)	0	100	100
All	All	837/966 (87%)	828 (99%)	9 (1%)	78	90

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	GLU
1	A	291	VAL
1	A	313	VAL
1	A	403	ILE
1	A	667	ILE
1	A	930	TYR
1	A	968	SER
1	A	987	THR
1	A	1029	ILE

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	636	GLN
1	A	668	HIS
1	A	683	GLN
1	A	801	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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 ⓘ

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, 95th 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^2)	Q<0.9
1	A	1017/1078 (94%)	0.19	74 (7%)	16	14	26, 62, 118, 139	0
2	B	5/35 (14%)	-0.62	0	100	100	34, 42, 54, 58	0
All	All	1022/1113 (91%)	0.18	74 (7%)	16	15	26, 61, 117, 139	0

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	664	HIS	6.8
1	A	930	TYR	6.0
1	A	660	VAL	5.6
1	A	871	ASN	5.1
1	A	829	PHE	5.1
1	A	988	TYR	4.5
1	A	1089	ALA	4.5
1	A	658	VAL	4.4
1	A	133	ASP	4.3
1	A	665	ILE	4.1
1	A	931	ALA	4.1
1	A	977	ALA	4.1
1	A	900	ALA	4.1
1	A	987	THR	4.0
1	A	652	TYR	4.0
1	A	978	TYR	3.7
1	A	828	ASP	3.7
1	A	854	LEU	3.6
1	A	597	ALA	3.6
1	A	903	PRO	3.6
1	A	1067	SER	3.5
1	A	811	ASP	3.5
1	A	1016	ILE	3.4
1	A	1027	SER	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	26	ILE	3.3
1	A	649	PHE	3.3
1	A	907	GLU	3.2
1	A	98	VAL	3.1
1	A	666	ALA	3.1
1	A	663	LYS	2.9
1	A	596	ASP	2.9
1	A	986	ASP	2.9
1	A	659	GLN	2.9
1	A	985	VAL	2.9
1	A	1071	ALA	2.8
1	A	810	GLN	2.8
1	A	1063	PHE	2.8
1	A	1070	MET	2.7
1	A	827	GLU	2.7
1	A	661	GLN	2.7
1	A	1057	VAL	2.7
1	A	809	MET	2.7
1	A	1045	LEU	2.6
1	A	783	ASP	2.5
1	A	874	ILE	2.5
1	A	991	ASN	2.5
1	A	904	LYS	2.4
1	A	10	ARG	2.4
1	A	62	THR	2.4
1	A	973	LYS	2.4
1	A	789	ASP	2.4
1	A	957	LEU	2.4
1	A	873	PRO	2.3
1	A	1073	PHE	2.3
1	A	1028	ASN	2.3
1	A	435	ILE	2.3
1	A	970	ALA	2.3
1	A	852	HIS	2.2
1	A	990	ALA	2.2
1	A	901	PHE	2.2
1	A	1006	SER	2.2
1	A	340	GLU	2.1
1	A	431	VAL	2.1
1	A	1066	SER	2.1
1	A	662	GLY	2.1
1	A	76	LEU	2.1

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	328	ALA	2.1
1	A	792	ALA	2.1
1	A	132	ASP	2.1
1	A	656	ASP	2.0
1	A	436	SER	2.0
1	A	25	GLN	2.0
1	A	899	ASN	2.0
1	A	386	ALA	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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.