Full wwPDB X-ray Structure Validation Report ⓘ

Sep 26, 2017 – 10:05 PM EDT

PDB ID	:	4XBU
Title	:	In vitro Crystal Structure of PAK4 in complex with Inka peptide
Authors	:	Baskaran, Y.; Ang, K.C.; Anekal, P.V.; Chan, W.L.; Grimes, J.M.; Manser, E.; Robinson, R.C.
Deposited on	:	unknown
Resolution	:	2.06 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030345

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.06 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		3
Ramachandran outliers		0
Sidechain outliers		3.4%
RSRZ outliers		2.3%

Worse ← → Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
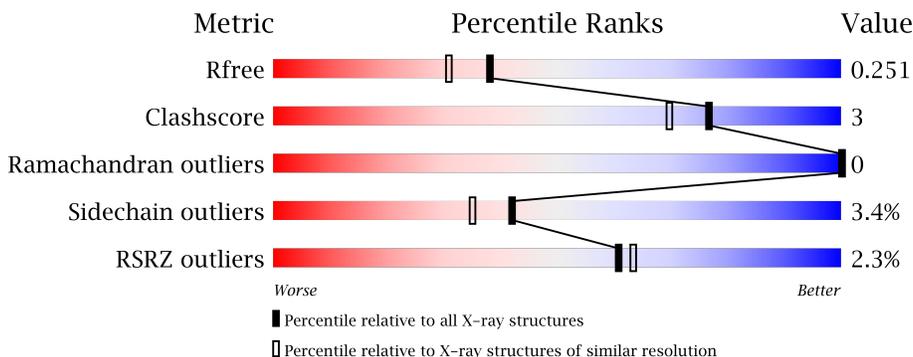
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	2028 (2.08-2.04)
Clashscore	112137	2143 (2.08-2.04)
Ramachandran outliers	110173	2126 (2.08-2.04)
Sidechain outliers	110143	2126 (2.08-2.04)
RSRZ outliers	101464	2035 (2.08-2.04)

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	326	2% 81% 8% · 10%
2	B	24	50% · 46%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2569 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	293	2322	1478	412	417	1	14	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	266	MET	-	expression tag	UNP O96013
A	267	ALA	-	expression tag	UNP O96013
A	268	GLU	-	expression tag	UNP O96013
A	269	GLU	-	expression tag	UNP O96013
A	270	HIS	-	expression tag	UNP O96013
A	271	HIS	-	expression tag	UNP O96013
A	272	HIS	-	expression tag	UNP O96013
A	273	HIS	-	expression tag	UNP O96013
A	274	HIS	-	expression tag	UNP O96013
A	275	HIS	-	expression tag	UNP O96013
A	276	HIS	-	expression tag	UNP O96013
A	277	HIS	-	expression tag	UNP O96013
A	278	LEU	-	expression tag	UNP O96013
A	279	GLU	-	expression tag	UNP O96013
A	280	VAL	-	expression tag	UNP O96013
A	281	LEU	-	expression tag	UNP O96013
A	282	PHE	-	expression tag	UNP O96013
A	283	GLN	-	expression tag	UNP O96013
A	284	GLY	-	expression tag	UNP O96013
A	285	PRO	-	expression tag	UNP O96013

- Molecule 2 is a protein called Protein FAM212A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	13	97	57	21	18	1	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	139	Total 139	O 139	0	0
3	B	11	Total 11	O 11	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: Serine/threonine-protein kinase PAK 4
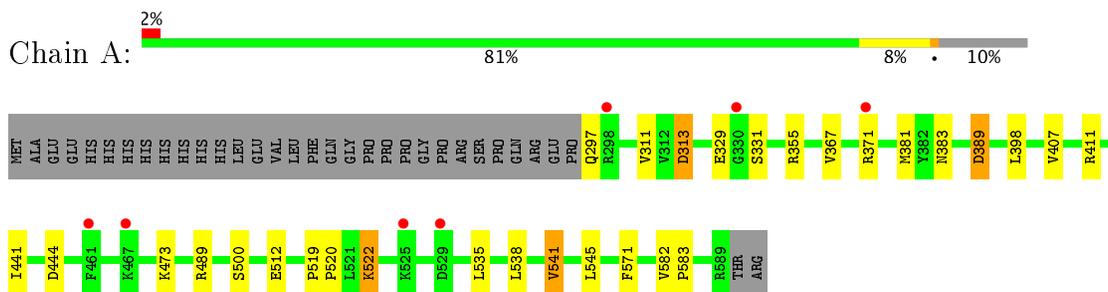
• Molecule 2: Protein FAM212A

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	65.62Å 65.62Å 184.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.81 – 2.06 32.81 – 2.06	Depositor EDS
% Data completeness (in resolution range)	99.9 (32.81-2.06) 100.0 (32.81-2.06)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.75 (at 2.06Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.199 , 0.243 0.206 , 0.251	Depositor DCC
R_{free} test set	1262 reflections (5.14%)	DCC
Wilson B-factor (Å^2)	36.7	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 43.0	EDS
L-test for twinning[2]	$< |L| > = 0.48,\ < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2569	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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)

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	0.92	0/2359	0.99	3/3193 (0.1%)
2	B	0.95	0/97	0.98	1/129 (0.8%)
All	All	0.92	0/2456	0.99	4/3322 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	398	LEU	CA-CB-CG	-6.71	99.87	115.30
1	A	389	ASP	CB-CG-OD2	-5.77	113.11	118.30
1	A	313	ASP	CB-CG-OD2	5.47	123.23	118.30
2	B	179	ARG	NE-CZ-NH1	5.13	122.86	120.30

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	2322	0	2381	17	0
2	B	97	0	97	0	0
3	A	139	0	0	4	0
3	B	11	0	0	0	0
All	All	2569	0	2478	17	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 3.

All (17) 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:313:ASP:HB3	3:A:700:HOH:O	1.76	0.85
1:A:512:GLU:OE1	3:A:707:HOH:O	2.09	0.71
1:A:313:ASP:OD2	1:A:383:ASN:ND2	2.21	0.61
1:A:313:ASP:CG	1:A:383:ASN:HD22	2.04	0.60
1:A:411:ARG:NH1	3:A:720:HOH:O	2.39	0.55
1:A:311:VAL:O	1:A:371:ARG:NH2	2.40	0.54
1:A:545:LEU:HD13	1:A:571:PHE:CZ	2.49	0.47
1:A:355:ARG:NH1	1:A:389:ASP:OD2	2.46	0.46
1:A:371:ARG:NH1	3:A:674:HOH:O	2.48	0.46
1:A:582:VAL:N	1:A:583:PRO:CD	2.81	0.44
1:A:441:ILE:HB	1:A:500:SER:HB2	2.01	0.43
1:A:473:LYS:HG2	1:A:489:ARG:O	2.19	0.42
1:A:519:PRO:HA	1:A:520:PRO:HD3	1.98	0.42
1:A:371:ARG:HA	1:A:381:MET:CE	2.50	0.41
1:A:311:VAL:HB	1:A:371:ARG:HH22	1.85	0.41
1:A:522:LYS:HB3	1:A:522:LYS:HE2	1.89	0.41
1:A:535:LEU:HD13	1:A:541:VAL:HG11	2.04	0.40

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	290/326 (89%)	279 (96%)	11 (4%)	0	100	100
2	B	11/24 (46%)	10 (91%)	1 (9%)	0	100	100
All	All	301/350 (86%)	289 (96%)	12 (4%)	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	255/285 (90%)	246 (96%)	9 (4%)	41	33
2	B	11/19 (58%)	11 (100%)	0	100	100
All	All	266/304 (88%)	257 (97%)	9 (3%)	42	35

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

Mol	Chain	Res	Type
1	A	297	GLN
1	A	329	GLU
1	A	331	SER
1	A	367	VAL
1	A	407	VAL
1	A	444	ASP
1	A	522	LYS
1	A	538	LEU
1	A	541	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3   RNA (i)

There are no RNA molecules in this entry.

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

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

       Full wwPDB X-ray Structure Validation Report

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	9,9,10	1.39	1 (11%)	9,12,14	2.33	2 (22%)

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	-	0/5/8/10	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	474	SEP	CA-C	3.16	1.54	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	474	SEP	OG-P-O1P	-2.56	99.29	106.47
1	A	474	SEP	OG-CB-CA	5.75	113.83	108.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	292/326 (89%)	0.06	7 (2%)	59	62	24, 38, 59, 88	0
2	B	13/24 (54%)	0.21	0	100	100	36, 43, 55, 56	0
All	All	305/350 (87%)	0.06	7 (2%)	61	63	24, 38, 58, 88	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	525	LYS	3.1
1	A	529	ASP	3.1
1	A	298	ARG	2.9
1	A	330	GLY	2.8
1	A	371	ARG	2.2
1	A	467	LYS	2.2
1	A	461	PHE	2.1

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	LLDF	B-factors(Å^2)	Q<0.9
1	SEP	A	474	10/11	0.98	0.11	-	32,37,40,41	0

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.