W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 01:34 pm BST

PDB ID	:	3CEQ
Title	:	The TPR domain of Human Kinesin Light Chain 2 (hKLC2)
Authors	:	Zhu, H.; Shen, Y.; MacKenzie, F.; Arrowsmith, C.H.; Edwards, A.M.; Boun-tra, C.; Weigelt, J.; Bochkarev, A.; Park, H.; Structural Genomics Consortium (SGC)
Deposited on	:	2008-02-29
Resolution	:	2.75 Å(reported)

The following versions of software and data (see references (1)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
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 (i)

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

The reported resolution of this entry is 2.75 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.

■ Percentile relative to all X-ray structures

❘ Percentile relative to X-ray structures of similar resolution
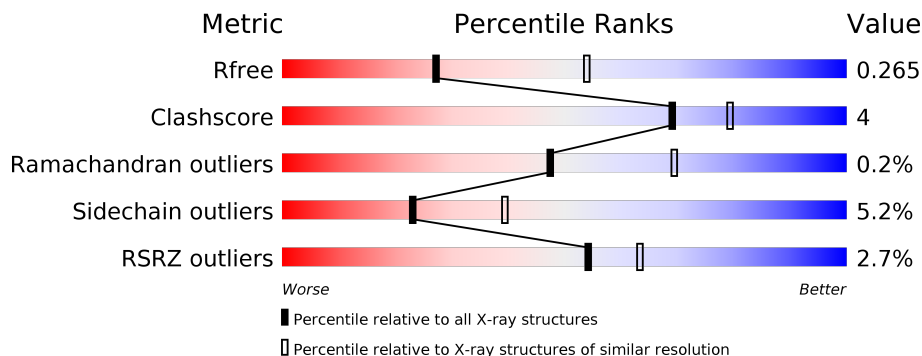
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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 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 <=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	283	2% ... 76% ... 11% · 12%
1	B	283	3% ... 81% ... 13% · 5%

2   Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4087 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 Kinesin light chain 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	250	Total 1961	C 1223	N 358	O 372	S 8	0	0	0
1	B	269	Total 2124	C 1325	N 386	O 405	S 8	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	198	MET	-	EXPRESSION TAG	UNP Q9H0B6
A	199	GLY	-	EXPRESSION TAG	UNP Q9H0B6
A	200	SER	-	EXPRESSION TAG	UNP Q9H0B6
A	201	SER	-	EXPRESSION TAG	UNP Q9H0B6
A	202	HIS	-	EXPRESSION TAG	UNP Q9H0B6
A	203	HIS	-	EXPRESSION TAG	UNP Q9H0B6
A	204	HIS	-	EXPRESSION TAG	UNP Q9H0B6
A	205	HIS	-	EXPRESSION TAG	UNP Q9H0B6
A	206	HIS	-	EXPRESSION TAG	UNP Q9H0B6
A	207	HIS	-	EXPRESSION TAG	UNP Q9H0B6
A	208	SER	-	EXPRESSION TAG	UNP Q9H0B6
A	209	SER	-	EXPRESSION TAG	UNP Q9H0B6
A	210	GLY	-	EXPRESSION TAG	UNP Q9H0B6
A	211	LEU	-	EXPRESSION TAG	UNP Q9H0B6
A	212	VAL	-	EXPRESSION TAG	UNP Q9H0B6
A	213	PRO	-	EXPRESSION TAG	UNP Q9H0B6
A	214	ARG	-	EXPRESSION TAG	UNP Q9H0B6
A	215	GLY	-	EXPRESSION TAG	UNP Q9H0B6
A	216	SER	-	EXPRESSION TAG	UNP Q9H0B6
B	198	MET	-	EXPRESSION TAG	UNP Q9H0B6
B	199	GLY	-	EXPRESSION TAG	UNP Q9H0B6
B	200	SER	-	EXPRESSION TAG	UNP Q9H0B6
B	201	SER	-	EXPRESSION TAG	UNP Q9H0B6
B	202	HIS	-	EXPRESSION TAG	UNP Q9H0B6
B	203	HIS	-	EXPRESSION TAG	UNP Q9H0B6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	204	HIS	-	EXPRESSION TAG	UNP Q9H0B6
B	205	HIS	-	EXPRESSION TAG	UNP Q9H0B6
B	206	HIS	-	EXPRESSION TAG	UNP Q9H0B6
B	207	HIS	-	EXPRESSION TAG	UNP Q9H0B6
B	208	SER	-	EXPRESSION TAG	UNP Q9H0B6
B	209	SER	-	EXPRESSION TAG	UNP Q9H0B6
B	210	GLY	-	EXPRESSION TAG	UNP Q9H0B6
B	211	LEU	-	EXPRESSION TAG	UNP Q9H0B6
B	212	VAL	-	EXPRESSION TAG	UNP Q9H0B6
B	213	PRO	-	EXPRESSION TAG	UNP Q9H0B6
B	214	ARG	-	EXPRESSION TAG	UNP Q9H0B6
B	215	GLY	-	EXPRESSION TAG	UNP Q9H0B6
B	216	SER	-	EXPRESSION TAG	UNP Q9H0B6

- Molecule 2 is water.

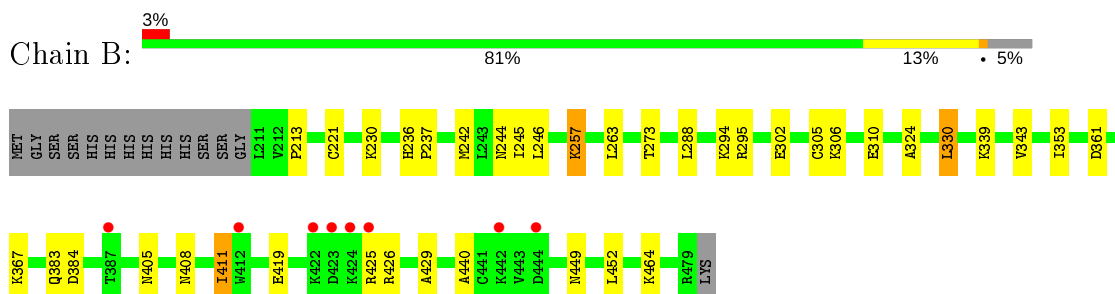| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total 2	O 2	0	0

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 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: Kinesin light chain 2

Chain A:

2%

76%   11%   •   12%

MET GLY SER SER HIS HIS HIS HIS HIS SER SER GLY L211 V212 P213 R214 G215 H234 D235 H236 P237 N244 I245 L246 L262 L263 L267 E271 K272 T273 L274 G275 L291 K297 Y298 E299 E300 R312 Q326 L332 N336 K339 V343 P360 D361

K378 Q383 R395 F401 N408 I411 E420 S421 LYS ASP LYS ARG ARG ASP SER ALA PRO TYR GLY GLU TYR GLY SER TRP TYR K439 K442 VAL ASP S445 R453 K464 S476 R479 LYS

- Molecule 1: Kinesin light chain 2

Chain B:

3%

81%   13%   •   5%

MET GLY SER SER HIS HIS HIS HIS HIS SER SER GLY L211 V212 P213 C221 K230 H236 P237 M242 L243 N244 I245 L246 K257 L263 T273 L288 K294 R295 E302 C305 K306 E310 A324 L330 K339 V343 I353 D361

K367 Q383 D384 T387 N405 N408 I411 W412 V413 E419 K422 D423 K424 R425 R426 A429 A440 C441 K442 V443 D444 N449 L452 K464 R479 LYS

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	70.42Å 99.94Å 103.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.75 29.51 – 2.75	Depositor EDS
% Data completeness (in resolution range)	98.4 (30.00-2.75) 98.4 (29.51-2.75)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.84 (at 2.76Å)	Xtriage
Refinement program	REFMAC refmac_5.2.0019	Depositor
R, R_{free}	0.234 , 0.271 0.230 , 0.265	Depositor DCC
R_{free} test set	979 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	85.8	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 70.3	EDS
L-test for twinning [2]	$<|L|> = 0.50, <L^2> = 0.33$	Xtriage
Estimated twinning fraction	0.004 for -h,l,k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4087	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.04% 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 ⓘ

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	0.31	0/1991	0.47	1/2682 (0.0%)
1	B	0.31	0/2162	0.48	1/2916 (0.0%)
All	All	0.31	0/4153	0.47	2/5598 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	213	PRO	N-CA-CB	5.89	110.37	103.30
1	A	213	PRO	N-CA-CB	5.88	110.36	103.30

There are no chirality outliers.

There are no planarity outliers.

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	1961	0	1956	13	0
1	B	2124	0	2107	17	0
2	B	2	0	0	0	0
All	All	4087	0	4063	30	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 4.

The worst 5 of 30 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:B:405:ASN:H	1:B:408:ASN:HB3	1.49	0.76
1:A:360:PRO:HB2	1:A:395:ARG:HD3	1.81	0.62
1:A:246:LEU:HB3	1:A:262:LEU:HD13	1.83	0.60
1:A:297:LYS:HE2	1:A:300:GLU:HG3	1.85	0.59
1:B:237:PRO:HD3	1:B:273:THR:HG21	1.86	0.56

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	244/283 (86%)	234 (96%)	9 (4%)	1 (0%)	34	53
1	B	267/283 (94%)	260 (97%)	7 (3%)	0	100	100
All	All	511/566 (90%)	494 (97%)	16 (3%)	1 (0%)	47	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	215	GLY

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	203/236 (86%)	193 (95%)	10 (5%)	25	43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	220/236 (93%)	208 (94%)	12 (6%)	21	37
All	All	423/472 (90%)	401 (95%)	22 (5%)	23	39

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	479	ARG
1	B	257	LYS
1	B	411	ILE
1	B	230	LYS
1	B	244	ASN

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	408	ASN

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 (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(Å²)	Q<0.9
1	A	250/283 (88%)	0.34	6 (2%)	59	68	71, 77, 83, 85	0
1	B	269/283 (95%)	0.30	8 (2%)	50	59	67, 77, 80, 84	0
All	All	519/566 (91%)	0.32	14 (2%)	54	63	67, 77, 82, 85	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	424	LYS	4.2
1	B	442	LYS	3.9
1	B	425	ARG	3.6
1	A	298	TYR	3.1
1	B	444	ASP	2.8

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.