wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 05:11 PM BST

PDB ID	:	1U7Q
Title	:	THE SOLUTION STRUCTURE OF THE NUCLEOTIDE BINDING DOMAIN OF KDPB
Authors	:	Haupt, M.; Bramkamp, M.; Coles, M.; Altendorf, K.; Kessler, H.
Deposited on	:	2004-08-04

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

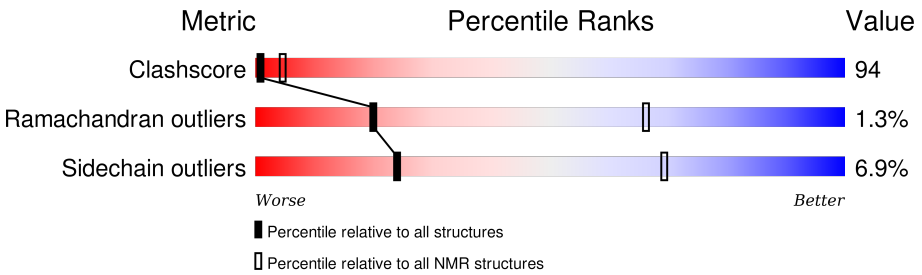
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

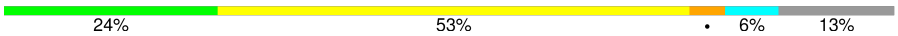
The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	156	

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 9 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:318-A:378, A:382-A:446 (126)	0.11	9

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 2 single-model clusters were found.

Cluster number	Models
1	3, 4, 5, 7, 8, 9, 12, 19, 20
2	2, 6, 13, 16, 17
3	10, 14, 15, 18
Single-model clusters	1; 11

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2098 atoms, of which 1057 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Potassium-transporting ATPase B chain.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	136	2098	642	1057	198	199	2	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	296	MET	-	EXPRESSION TAG	UNP P03960
A	297	GLY	-	EXPRESSION TAG	UNP P03960
A	298	HIS	-	EXPRESSION TAG	UNP P03960
A	299	HIS	-	EXPRESSION TAG	UNP P03960
A	300	HIS	-	EXPRESSION TAG	UNP P03960
A	301	HIS	-	EXPRESSION TAG	UNP P03960
A	302	HIS	-	EXPRESSION TAG	UNP P03960
A	303	HIS	-	EXPRESSION TAG	UNP P03960
A	304	HIS	-	EXPRESSION TAG	UNP P03960
A	305	HIS	-	EXPRESSION TAG	UNP P03960
A	306	HIS	-	EXPRESSION TAG	UNP P03960
A	307	HIS	-	EXPRESSION TAG	UNP P03960
A	308	SER	-	EXPRESSION TAG	UNP P03960
A	309	SER	-	EXPRESSION TAG	UNP P03960
A	310	GLY	-	EXPRESSION TAG	UNP P03960
A	311	HIS	-	EXPRESSION TAG	UNP P03960
A	312	GLY	-	EXPRESSION TAG	UNP P03960
A	313	GLY	-	EXPRESSION TAG	UNP P03960
A	314	ARG	-	EXPRESSION TAG	UNP P03960
A	315	HIS	-	EXPRESSION TAG	UNP P03960

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Potassium-transporting ATPase B chain

Chain A: 24% 53% • 6% 13%

MET GLY HIS HIS HIS HIS HIS HIS HIS HIS HIS HIS SER SER GLY HIS GLY GLY ARG HIS N316 R317 Q318 A319 S320 E321 F322 I323 P324 A325 Q326 G327 V328 D329 E330 L333 A334 D335 A336 A337 Q338 L339 A340 S341 L342 A343 D344 G349 R350 S351 I352 V353 I354 L355 A356 K357 Q358 R359

F360 N361 L362 R363 E364 R365 D366 V367 Q368 S369 L370 H371 A372 T373 F374 F377 T378 A379 Q380 S381 R382 M383 S384 G385 I386 N387 I388 D389 N390 R391 M392 I393 R394 K395 G396 S397 V398 D399 A400 I401 R402 R403 H404 V405 N408 F412 P413 V416 D417 Q418 K419 V420 V423 Q426

G427 A428 T429 P430 L431 V432 V433 V434 S437 R438 V439 L440 G441 V442 I443 A444 L445 K446 D447 I448 V449 K450 G451

4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 9. Colouring as in section 4.1 above.

- Molecule 1: Potassium-transporting ATPase B chain

Chain A: 26% 50% • 6% 13%

MET GLY HIS HIS HIS HIS HIS HIS HIS HIS HIS HIS SER SER GLY HIS GLY GLY ARG HIS N316 R317 Q318 A319 S320 E321 F322 I323 P324 A325 Q326 G327 V328 D329 E330 L333 A334 D335 A336 A337 Q338 L339 A340 S341 L342 A343 D344 G349 R350 S351 I352 V353 I354 L355 A356 K357 Q358 R359

F360 N361 L362 R363 E364 R365 D366 V367 Q368 S369 L370 H371 A372 T373 F374 F377 T378 A379 Q380 S381 R382 M383 S384 G385 I386 N387 I388 D389 N390 R391 M392 I393 R394 K395 G396 S397 V398 D399 A400 I401 R402 R403 H404 N408 F412 P413 T414 D415 V416 D417 Q418 K419 V420 D421 Q422 V423

Q426 G427 A428 T429 P430 L431 V432 V433 V434 V439 L440 G441 V442 I443 A444 L445 K446 D447 I448 V449 K450 G451

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *SIMULATED ANNEALING*.

Of the 50 calculated structures, 20 were deposited, based on the following criterion: *STRUCTURES WITH THE LEAST RESTRAINT VIOLATIONS*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR	refinement	NIH-2.9.3
XWINNMR	structure solution	3.5
X-PLOR	structure solution	NIH-2.9.3

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality (i)

6.1 Standard geometry (i)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	965	980	980	182±7
All	All	19300	19600	19600	3638

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 94.

5 of 403 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:339:LEU:HD11	1:A:393:ILE:HD12	1.10	1.24	6	20
1:A:398:VAL:HG22	1:A:432:VAL:HG23	1.03	1.26	3	20
1:A:370:LEU:HD12	1:A:388:ILE:HD11	1.03	1.24	1	1
1:A:322:PHE:CD1	1:A:355:LEU:HD21	1.03	1.87	11	20
1:A:420:VAL:HG23	1:A:442:VAL:HG21	1.02	1.32	1	20

6.3 Torsion angles (i)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	126/156 (81%)	121±0 (96±0%)	4±1 (3±1%)	2±0 (1±0%)	20	66
All	All	2520/3120 (81%)	2417 (96%)	71 (3%)	32 (1%)	20	66

All 2 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	362	LEU	20
1	A	389	ASP	12

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	103/127 (81%)	96±1 (93±1%)	7±1 (7±1%)	24	69
All	All	2060/2540 (81%)	1918 (93%)	142 (7%)	24	69

5 of 15 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	339	LEU	20
1	A	355	LEU	20
1	A	369	SER	20
1	A	395	LYS	19
1	A	445	LEU	17

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates (i)

There are no carbohydrates in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided