Full wwPDB NMR Structure Validation Report (i)

Mar 6, 2022 – 02:55 PM EST

PDB ID	:	2KOJ
Title	:	Solution structure of mouse Par-3 PDZ2 (residues 450-558)
Authors	:	Volkman, B.F.; Tyler, R.C.; Peterson, F.C.; Center for Eukaryotic Structural Genomics (CESG)
Deposited on	:	2009-09-23

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/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:

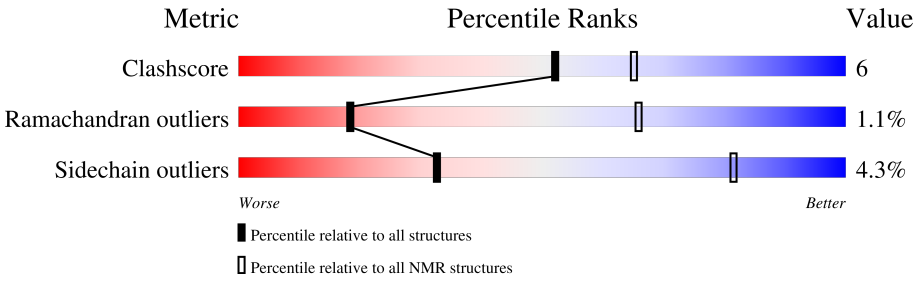
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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	111	66% 6% • 27%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:458-A:478, A:487-A:546 (81)	0.40	2

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 5 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1730 atoms, of which 885 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Partitioning defective 3 homolog.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	111	1730	527	885	155	161	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	448	GLY	-	expression tag	UNP Q99NH2
A	449	SER	-	expression tag	UNP Q99NH2

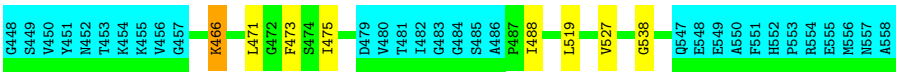
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: Partitioning defective 3 homolog

Chain A:



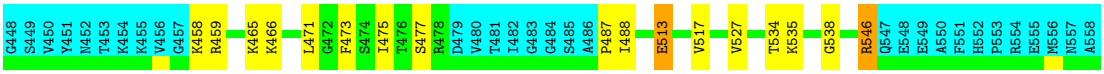
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Partitioning defective 3 homolog

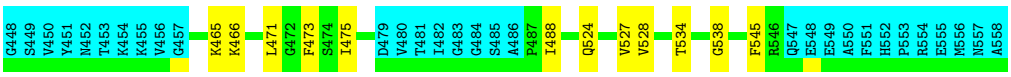
Chain A:



4.2.2 Score per residue for model 2 (medoid)

• Molecule 1: Partitioning defective 3 homolog

Chain A:



4.2.3 Score per residue for model 3

• Molecule 1: Partitioning defective 3 homolog

Chain A: 63% 10% 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K458 K466 L471 G472 F473 S474 I475 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 D502 G503 R504 V527 M536 E537 G538 Q547 E548 E549 A550 F551 H552 P553 R554 E555 M556 N557 A558

4.2.4 Score per residue for model 4

• Molecule 1: Partitioning defective 3 homolog

Chain A: 54% 18% • 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K458 R459 L464 K465 K466 L471 G472 F473 S474 I475 T476 S477 R478 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 Y489 V490 L505 R510 E526 L530 T534 G538 T539 L543 Q547 E548 E549 A550 F551 H552 P553

R554 E555 M556 N557 A558

4.2.5 Score per residue for model 5

• Molecule 1: Partitioning defective 3 homolog

Chain A: 60% 11% • 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K458 R459 K466 L471 G472 F473 S474 I475 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 K506 A507 L519 Q524 V527 M536 F545 R546 Q547 E548 E549 A550 F551 H552 P553 R554 E555 M556 N557 A558

4.2.6 Score per residue for model 6

• Molecule 1: Partitioning defective 3 homolog

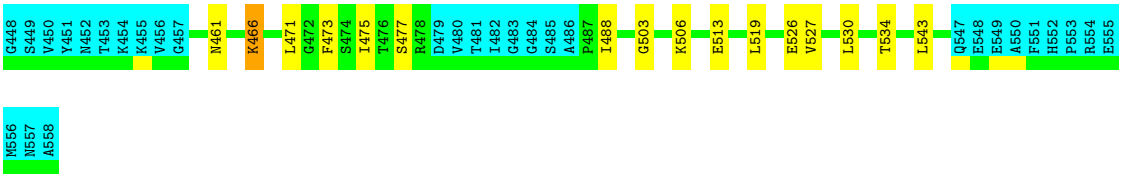
Chain A: 59% 14% 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K458 K466 L471 G472 F473 S474 I475 T476 S477 R478 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 G503 K506 D509 L519 S523 V527 T534 R546 Q547 E548 E549 A550 F551 H552 P553 R554 E555 M556 N557 A558

4.2.7 Score per residue for model 7

• Molecule 1: Partitioning defective 3 homolog

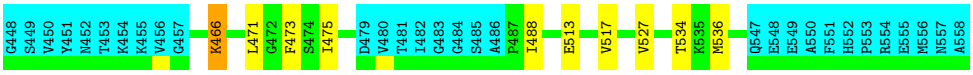
Chain A: 59% 14% • 27%



4.2.8 Score per residue for model 8

- Molecule 1: Partitioning defective 3 homolog

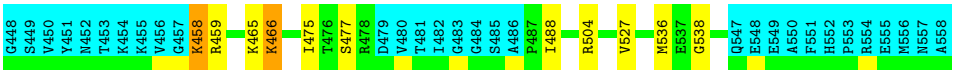
Chain A:



4.2.9 Score per residue for model 9

- Molecule 1: Partitioning defective 3 homolog

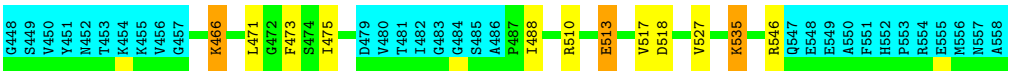
Chain A:



4.2.10 Score per residue for model 10

- Molecule 1: Partitioning defective 3 homolog

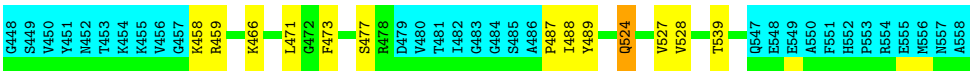
Chain A:



4.2.11 Score per residue for model 11

- Molecule 1: Partitioning defective 3 homolog

Chain A:



4.2.12 Score per residue for model 12

- Molecule 1: Partitioning defective 3 homolog

Chain A: 59% 14% 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K458 R459 K465 K466 L471 G472 F473 S474 I475 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 E513 L519 K522 V527 K535 G538 L543 V544 Q547 E548 E549 A550 F551 H552 P553 R554 E555 M556 N557 A558

4.2.13 Score per residue for model 13

- Molecule 1: Partitioning defective 3 homolog

Chain A: 61% 11% • 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K458 K466 E469 R478 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 R496 R504 L519 E526 V527 L530 M536 E537 G538 Q547 E548 E549 A550 F551 H552 P553 R554 E555 M556 N557 A558

4.2.14 Score per residue for model 14

- Molecule 1: Partitioning defective 3 homolog

Chain A: 58% 14% • 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K465 K466 E469 I475 T476 S477 R478 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 R496 G497 Q501 D502 G503 K506 D509 I512 L519 T534 G538 R546 Q547 E548 E549 A550 F551 H552 P553 R554 E555

M556 N557 A558

4.2.15 Score per residue for model 15

- Molecule 1: Partitioning defective 3 homolog

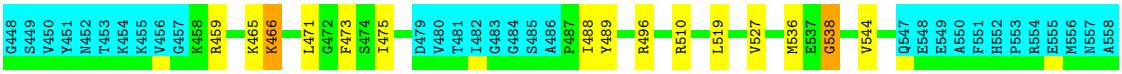
Chain A: 62% 9% • 27%

G448 S449 V450 Y451 N452 T453 K454 K455 V456 G457 K466 L471 G472 F473 S477 R478 D479 V480 T481 I482 G483 G484 S485 A486 P487 I488 R496 R504 L505 K506 V514 E526 L530 S541 Q547 E548 E549 A550 F551 H552 P553 R554 E555 M556 N557 A558

4.2.16 Score per residue for model 16

- Molecule 1: Partitioning defective 3 homolog

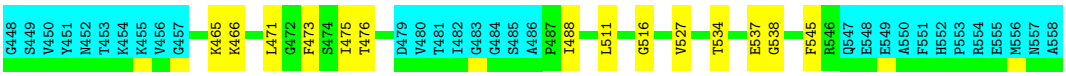
Chain A: 59% 12% • 27%



4.2.17 Score per residue for model 17

- Molecule 1: Partitioning defective 3 homolog

Chain A:



4.2.18 Score per residue for model 18

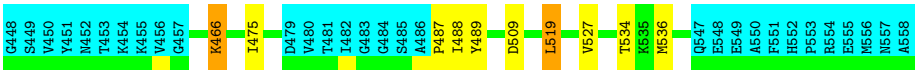
- Molecule 1: Partitioning defective 3 homolog

Chain A:

4.2.19 Score per residue for model 19

- Molecule 1: Partitioning defective 3 homolog

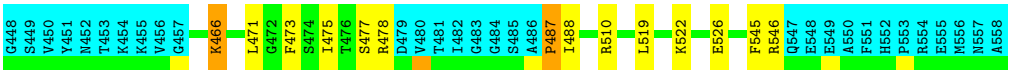
Chain A:



4.2.20 Score per residue for model 20

- Molecule 1: Partitioning defective 3 homolog

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *AUTOMATED METHODS WERE USED FOR BACKBONE CHEMICAL SHIFT ASSIGNMENT AND ITERATIVE NOE REFINEMENT. FINAL STRUCTURES WERE OBTAINED BY MOLECULAR DYNAMICS IN EXPLICIT SOLVENT.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
Xplor-NIH	refinement	2.9.3

No chemical shift data was provided.

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	621	671	671	7±2
All	All	12420	13420	13420	147

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:466:LYS:N	1:A:466:LYS:HE3	0.59	2.12	19	6
1:A:475:ILE:HB	1:A:488:ILE:CG2	0.57	2.29	17	12
1:A:488:ILE:HG13	1:A:527:VAL:HG21	0.56	1.78	19	13
1:A:466:LYS:NZ	1:A:536:MET:HA	0.55	2.17	9	3
1:A:477:SER:HB2	1:A:488:ILE:HD13	0.54	1.78	18	3
1:A:478:ARG:O	1:A:487:PRO:HD2	0.54	2.02	4	1
1:A:471:LEU:HB3	1:A:473:PHE:CE1	0.53	2.38	12	15
1:A:477:SER:HB3	1:A:488:ILE:HD13	0.53	1.81	6	6
1:A:476:THR:O	1:A:488:ILE:HG23	0.53	2.04	17	1
1:A:497:GLY:O	1:A:501:GLN:HG2	0.51	2.04	14	1
1:A:465:LYS:HA	1:A:538:GLY:O	0.51	2.05	12	8
1:A:509:ASP:HB3	1:A:546:ARG:NH1	0.51	2.20	6	1
1:A:503:GLY:O	1:A:506:LYS:HE3	0.49	2.07	14	3
1:A:504:ARG:HA	1:A:504:ARG:NE	0.49	2.23	13	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:487:PRO:HG3	1:A:510:ARG:NE	0.48	2.24	20	1
1:A:475:ILE:HB	1:A:488:ILE:HG23	0.48	1.86	18	13
1:A:475:ILE:HB	1:A:488:ILE:HG21	0.47	1.86	17	1
1:A:519:LEU:HA	1:A:522:LYS:HG2	0.47	1.87	12	1
1:A:546:ARG:HD3	1:A:546:ARG:O	0.47	2.10	1	1
1:A:506:LYS:HD2	1:A:507:ALA:N	0.47	2.25	5	1
1:A:513:GLU:HB2	1:A:517:VAL:O	0.46	2.10	10	3
1:A:502:ASP:OD2	1:A:504:ARG:HG2	0.45	2.11	3	1
1:A:487:PRO:HB2	1:A:489:TYR:CE1	0.45	2.47	11	1
1:A:466:LYS:HE3	1:A:535:LYS:O	0.45	2.11	10	1
1:A:488:ILE:HG13	1:A:527:VAL:CG2	0.45	2.42	10	8
1:A:514:VAL:HA	1:A:541:SER:O	0.45	2.12	15	1
1:A:488:ILE:HD12	1:A:527:VAL:HG23	0.45	1.88	17	1
1:A:526:GLU:O	1:A:530:LEU:HG	0.44	2.11	4	4
1:A:459:ARG:HB3	1:A:543:LEU:HD21	0.44	1.90	12	2
1:A:519:LEU:HB2	1:A:527:VAL:HG11	0.44	1.89	19	1
1:A:458:LYS:HD2	1:A:459:ARG:N	0.44	2.28	5	1
1:A:477:SER:HB3	1:A:488:ILE:HG12	0.44	1.89	4	1
1:A:519:LEU:HB2	1:A:527:VAL:CG1	0.44	2.43	5	1
1:A:464:LEU:O	1:A:539:THR:HA	0.43	2.12	18	2
1:A:487:PRO:HB3	1:A:510:ARG:CZ	0.43	2.43	4	1
1:A:469:GLU:HG3	1:A:496:ARG:HB2	0.43	1.90	14	1
1:A:513:GLU:CB	1:A:518:ASP:HA	0.42	2.44	10	1
1:A:509:ASP:HB3	1:A:546:ARG:HD2	0.42	1.92	14	1
1:A:459:ARG:HA	1:A:544:VAL:O	0.42	2.15	12	2
1:A:488:ILE:O	1:A:510:ARG:HA	0.41	2.16	10	1
1:A:504:ARG:HA	1:A:504:ARG:HE	0.41	1.75	9	1
1:A:489:TYR:HA	1:A:509:ASP:O	0.41	2.15	19	1
1:A:524:GLN:O	1:A:528:VAL:HG23	0.41	2.15	2	2
1:A:519:LEU:HD13	1:A:527:VAL:HG11	0.41	1.92	6	2
1:A:466:LYS:HE3	1:A:466:LYS:H	0.41	1.76	19	2
1:A:513:GLU:CG	1:A:543:LEU:HB3	0.41	2.46	7	1
1:A:489:TYR:CE2	1:A:510:ARG:HB2	0.41	2.51	16	1
1:A:506:LYS:HA	1:A:506:LYS:HE3	0.41	1.93	15	1
1:A:469:GLU:HG3	1:A:496:ARG:HB3	0.41	1.93	13	1
1:A:512:ILE:O	1:A:519:LEU:HG	0.41	2.16	14	1
1:A:504:ARG:HE	1:A:504:ARG:N	0.41	2.14	15	1
1:A:536:MET:HG2	1:A:536:MET:O	0.40	2.16	8	1
1:A:490:VAL:HG21	1:A:505:LEU:HD11	0.40	1.92	4	1
1:A:466:LYS:HE3	1:A:538:GLY:O	0.40	2.17	16	1
1:A:522:LYS:HD2	1:A:526:GLU:HG3	0.40	1.93	20	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:513:GLU:HA	1:A:519:LEU:HG	0.40	1.94	12	1

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	81/111 (73%)	77±1 (95±2%)	3±1 (3±2%)	1±1 (1±1%)	18	66
All	All	1620/2220 (73%)	1547 (95%)	55 (3%)	18 (1%)	18	66

All 5 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	534	THR	10
1	A	538	GLY	3
1	A	487	PRO	2
1	A	535	LYS	2
1	A	516	GLY	1

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	68/91 (75%)	65±1 (96±2%)	3±1 (4±2%)	33	81
All	All	1360/1820 (75%)	1302 (96%)	58 (4%)	33	81

All 18 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	466	LYS	20
1	A	458	LYS	7
1	A	519	LEU	6
1	A	459	ARG	4
1	A	546	ARG	3
1	A	536	MET	3
1	A	513	GLU	2
1	A	524	GLN	2
1	A	496	ARG	2
1	A	535	LYS	1
1	A	506	LYS	1
1	A	461	ASN	1
1	A	539	THR	1
1	A	478	ARG	1
1	A	504	ARG	1
1	A	511	LEU	1
1	A	537	GLU	1
1	A	491	LYS	1

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 monosaccharides 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