wwPDB NMR Structure Validation Summary Report (i)

Apr 20, 2024 – 08:23 PM EDT

PDB ID	:	2KR0
Title	:	Solution structure of the proteasome ubiquitin receptor Rpn13
Authors	:	Chen, X.; Lee, B.; Finley, D.; Walters, K.J.
Deposited on	:	2009-11-25

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
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

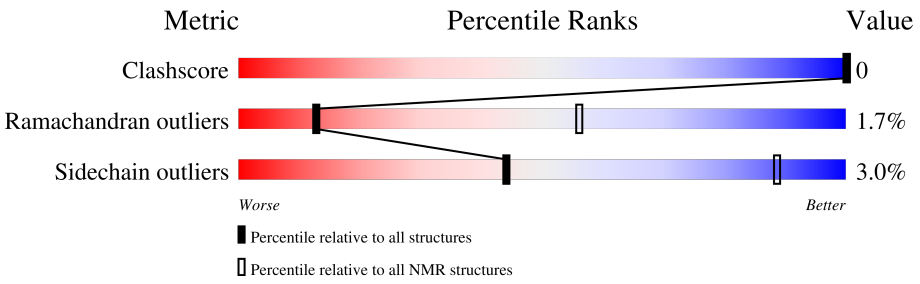
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	411	50% • 48% •

2 Ensemble composition and analysis (i)

This entry contains 31 models. Model 4 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:22-A:131, A:286-A:385 (210)	0.87	4

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 4 clusters and 1 single-model cluster was found.

Cluster number	Models
1	2, 6, 8, 11, 12, 13, 15, 17, 19, 22, 23, 27, 29
2	3, 4, 5, 9, 10, 18, 25, 26, 28
3	7, 16, 21, 24, 30, 31
4	1, 14
Single-model clusters	20

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 5853 atoms, of which 2901 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Proteasomal ubiquitin receptor ADRM1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	407	5853	1825	2901	506	604	17	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q16186
A	-2	PRO	-	expression tag	UNP Q16186
A	-1	GLY	-	expression tag	UNP Q16186
A	0	SER	-	expression tag	UNP Q16186

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: Proteasomal ubiquitin receptor ADRM1

Chain A: 50% 48%

GLY PRO GLY SER M1 T2 T3 S4 G5 A6 L7 F8 P9 S10 L11 V12 P13 G14 S15 R16 G17 A18 S19 N20 K21 D72 P132 M133 P134 G135 A136 L137 G138 A139 S140 G141 S142 S143 G144 H145 E146 L147 S148 A149 L150 G151 G152 E153 G154 G155 L156 Q157 S158 L159 L160 G161 N162 M163

S164 H165 S166 Q167 L168 M169 Q170 L171 I172 G173 P174 A175 G176 L177 G178 G179 L180 G181 G182 L183 G184 A185 L186 T187 G188 P189 G190 L191 A192 S193 L194 L195 G196 S197 S198 G199 P200 P201 G202 S203 S204 S205 S206 S207 S208 S209 R210 S211 Q212 S213 A214 A215 V216 T217 P218 S219 S220 T221 T222 S223

S224 T225 R226 A227 T228 P229 A230 P231 S232 A233 P234 A235 A236 A237 S238 A239 T240 S241 P242 S243 P244 A245 P246 S247 S248 G249 N250 G251 A252 S253 T254 A255 A256 S257 P258 T259 Q260 P261 I262 Q263 L264 S265 D266 L267 Q268 S269 I270 L271 A272 T273 M274 N275 V276 P277 A278 G279 P280 A281 G282 G283

Q284 Q285 S316 S333 G353 K386 P387 E388 Q389 K390 E391 G392 D393 T394 K395 D396 K397 K398 D399 E400 E401 E402 D403 M404 S405 L406 D407

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

The representative model is number 4. Colouring as in section 4.1 above.

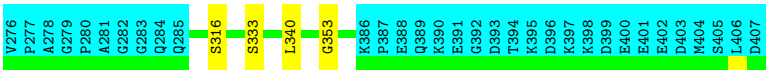
• Molecule 1: Proteasomal ubiquitin receptor ADRM1

Chain A: 49% 48%

GLY PRO GLY SER M1 T2 T3 S4 G5 A6 L7 F8 P9 S10 L11 V12 P13 G14 S15 R16 G17 A18 S19 N20 K21 R43 Q51 K62 D72 M109 P132 M133 P134 G135 A136 L137 G138 A139 S140 G141 S142 S143 G144 H145 E146 L147 S148 A149 L150 G151 G152 E153 G154 G155

L156 Q157 S158 L159 L160 G161 N162 M163 S164 H165 S166 Q167 L168 M169 Q170 L171 I172 G173 P174 A175 G176 L177 G178 G179 L180 G181 G182 L183 G184 A185 L186 T187 G188 P189 G190 L191 A192 S193 L194 L195 G196 S197 S198 G199 P200 P201 G202 S203 S204 S205 S206 S207 S208 S209 R210 S211 Q212 S213 A214 A215

V216 T217 P218 S219 S220 T221 T222 S223 S224 T225 R226 A227 T228 P229 A230 P231 S232 A233 P234 A235 A236 A237 S238 A239 T240 S241 P242 S243 P244 A245 P246 S247 S248 G249 N250 G251 A252 S253 T254 A255 A256 S257 P258 T259 Q260 P261 I262 Q263 L264 S265 D266 L267 Q268 S269 I270 L271 A272 T273 M274 N275


V276
P277
A278
G279
P280
A281
G282
G283
Q284
Q285
S316
S333
L340
G353
K386
P387
E388
Q389
K390
E391
G392
D393
T394
K395
D396
K397
K398
D399
E400
E401
E402
D403
M404
S405
L406
D407

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 50 calculated structures, 31 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
X-PLOR NIH	refinement	
X-PLOR NIH	structure solution	

No chemical shift data was provided.

6 Model quality (i)

6.1 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 (average) 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	0.71±0.00	0±0/1676 (0.0± 0.0%)	0.97±0.01	1±1/2274 (0.0± 0.0%)
All	All	0.71	0/51956 (0.0%)	0.97	21/70494 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.1±0.2
All	All	0	2

There are no bond-length outliers.

5 of 9 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	43	ARG	NE-CZ-NH2	-8.45	116.07	120.30	19	8
1	A	122	ARG	NE-CZ-NH2	-7.79	116.40	120.30	24	2
1	A	43	ARG	NE-CZ-NH1	6.72	123.66	120.30	15	2
1	A	309	ARG	NE-CZ-NH2	-6.53	117.03	120.30	11	1
1	A	92	ARG	NE-CZ-NH2	-6.26	117.17	120.30	6	1

There are no chirality outliers.

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

Mol	Chain	Res	Type	Group	Models (Total)
1	A	94	TYR	Sidechain	1
1	A	48	TYR	Sidechain	1

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	1640	1615	1615	0±0
All	All	50840	50065	50065	4

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:333:SER:H	1:A:334:PRO:CD	0.47	2.22	19	3
1:A:333:SER:H	1:A:334:PRO:HD2	0.45	1.72	19	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	210/411 (51%)	199±2 (95±1%)	8±2 (4±1%)	4±1 (2±0%)	13	56
All	All	6510/12741 (51%)	6157 (95%)	243 (4%)	110 (2%)	13	56

5 of 11 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	333	SER	31
1	A	353	GLY	31
1	A	316	SER	24
1	A	286	VAL	8
1	A	55	SER	6

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	179/324 (55%)	174±2 (97±1%)	5±2 (3±1%)	44 89
All	All	5549/10044 (55%)	5381 (97%)	168 (3%)	44 89

5 of 38 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	72	ASP	19
1	A	320	LEU	15
1	A	356	MET	11
1	A	358	GLN	11
1	A	340	LEU	9

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