



# wwPDB NMR Structure Validation Summary Report ⓘ

Apr 27, 2016 – 05:44 AM BST

PDB ID : 2RP4  
Title : Solution Structure of the oligomerization domain in Dmp53  
Authors : Ou, H.D.; Doetsch, V.  
Deposited on : 2008-04-30

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

---

The following versions of software and data (see [references ⓘ](#)) 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

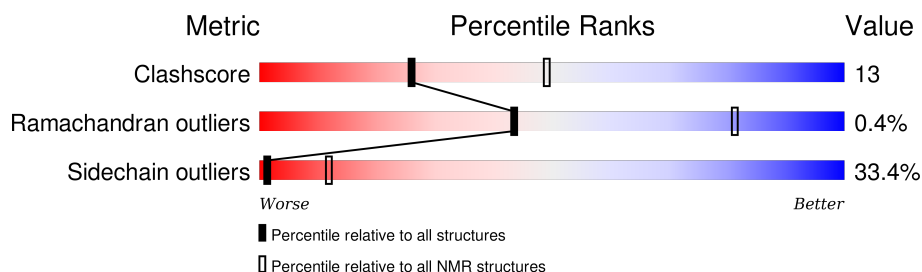
# 1 Overall quality at a glance

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	76	
1	B	76	
1	C	76	
1	D	76	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 18 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:320-A:326, A:330-A:384, B:320-B:326, B:330-B:384, C:320-C:326, C:330-C:384, D:320-D:325, D:330-D:384 (247)	0.55	18

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 3 single-model clusters were found.

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 4576 atoms, of which 2284 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Transcription factor p53.

Mol	Chain	Residues	Atoms						Trace
1	A	71	Total	C	H	N	O	S	0
			1144	355	571	106	110	2	
1	B	71	Total	C	H	N	O	S	0
			1144	355	571	106	110	2	
1	C	71	Total	C	H	N	O	S	0
			1144	355	571	106	110	2	
1	D	71	Total	C	H	N	O	S	0
			1144	355	571	106	110	2	

There are 20 discrepancies between the modelled and reference sequences:

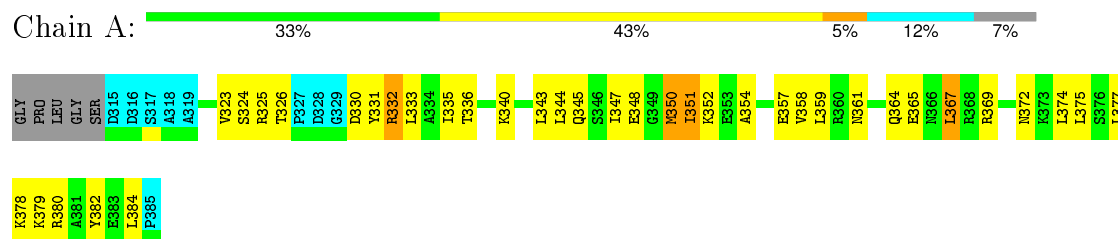
Chain	Residue	Modelled	Actual	Comment	Reference
A	310	GLY	-	EXPRESSION TAG	UNP Q9N6D8
A	311	PRO	-	EXPRESSION TAG	UNP Q9N6D8
A	312	LEU	-	EXPRESSION TAG	UNP Q9N6D8
A	313	GLY	-	EXPRESSION TAG	UNP Q9N6D8
A	314	SER	-	EXPRESSION TAG	UNP Q9N6D8
B	310	GLY	-	EXPRESSION TAG	UNP Q9N6D8
B	311	PRO	-	EXPRESSION TAG	UNP Q9N6D8
B	312	LEU	-	EXPRESSION TAG	UNP Q9N6D8
B	313	GLY	-	EXPRESSION TAG	UNP Q9N6D8
B	314	SER	-	EXPRESSION TAG	UNP Q9N6D8
C	310	GLY	-	EXPRESSION TAG	UNP Q9N6D8
C	311	PRO	-	EXPRESSION TAG	UNP Q9N6D8
C	312	LEU	-	EXPRESSION TAG	UNP Q9N6D8
C	313	GLY	-	EXPRESSION TAG	UNP Q9N6D8
C	314	SER	-	EXPRESSION TAG	UNP Q9N6D8
D	310	GLY	-	EXPRESSION TAG	UNP Q9N6D8
D	311	PRO	-	EXPRESSION TAG	UNP Q9N6D8
D	312	LEU	-	EXPRESSION TAG	UNP Q9N6D8
D	313	GLY	-	EXPRESSION TAG	UNP Q9N6D8
D	314	SER	-	EXPRESSION TAG	UNP Q9N6D8

## 4 Residue-property plots

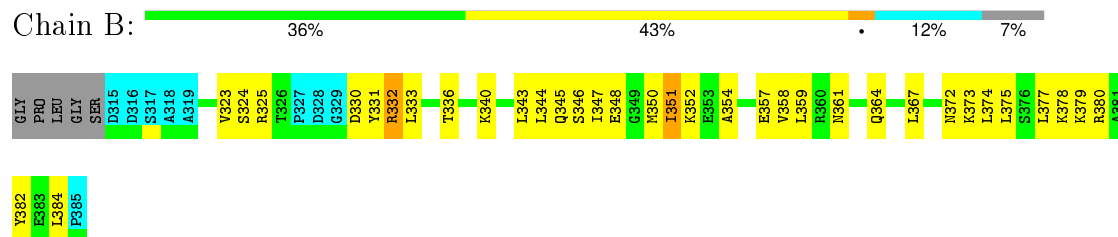
### 4.1 Average score per residue in the NMR ensemble

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.

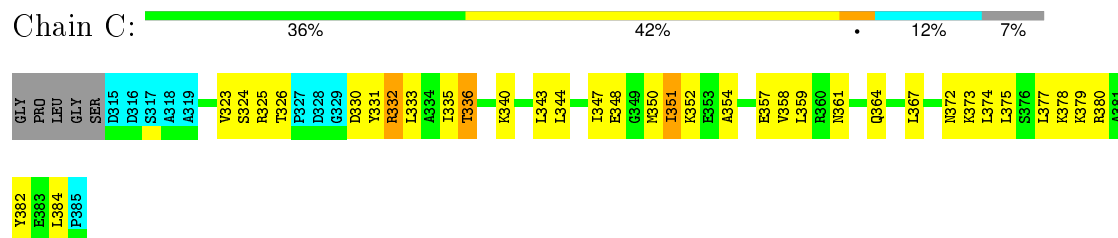
- Molecule 1: Transcription factor p53



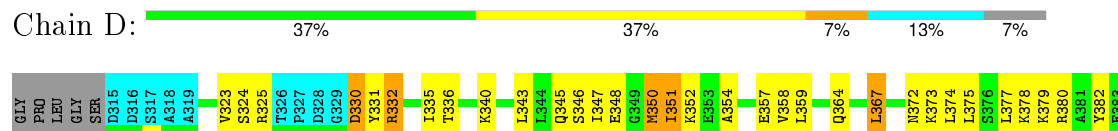
- Molecule 1: Transcription factor p53



- Molecule 1: Transcription factor p53



- Molecule 1: Transcription factor p53



L384  
P385

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

The representative model is number 18. Colouring as in section 4.1 above.

- Molecule 1: Transcription factor p53

Chain A: 37% 39% 12% 7%

GLY PRO LEU GLY SER D315 D316 S317 A318 A319 E320 N321 N322 N323 S324 R325 T326 P327 D328 D329 D330 Y331 R332 R333 L333 P338 N339 K340 E341 N342 S346 I347 N350 K351 K352 E353 E357 Q364 E365 N366 N367 A371 N372 K373 L374 L375 S376 L377 K380 N381 Y382 E383 L384 P385

- Molecule 1: Transcription factor p53

Chain B: 38% 37% 5% 12% 7%

GLY PRO LEU GLY SER D315 D316 S317 A318 A319 E320 N321 S324 R325 T326 P327 D328 D329 D330 Y331 R332 R333 L333 A334 T335 T336 C337 P338 E341 N342 L343 L344 T347 N350 I351 E357 V358 L359 R360 N361 P362 N363 Q364 E365 N366 L367 A371 N372 K373 L374 L375 S376 L377 R380

A381  
Y382  
E383  
L384  
P385

- Molecule 1: Transcription factor p53

Chain C: 37% 32% 12% 12% 7%

GLY PRO LEU GLY SER D315 D316 S317 A318 A319 E320 V323 S324 R325 T326 P327 D328 D329 D330 Y331 R332 R333 L333 T336 K340 L347 N350 I351 E357 V358 L359 R360 N361 Q364 E365 N366 L367 R368 R369 H370 A371 N372 K373 L374 L375 S376 L377 K378 K379 R380 A381 Y382 E383 L384

P385

- Molecule 1: Transcription factor p53

Chain D: 33% 38% 9% 13% 7%

GLY PRO LEU GLY SER D315 D316 S317 A318 A319 E320 V323 S324 R325 T326 P327 D328 D329 D330 Y331 R332 R333 L333 A334 T335 T336 C337 N342 Q345 S346 I347 E348 G349 N350 I351 K352 E353 A354 E357 V358 L359 Q364 L367 R368 A371 N372 K373 L374 L375 S376 L377 K378 K379 R380

A381  
Y382  
E383  
L384  
P385

## 5 Refinement protocol and experimental data overview ⓘ

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

Of the 100 calculated structures, 20 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
ARIA	refinement	1.2
TALOS	geometry optimization	

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	514	527	526	17±4
1	B	514	527	526	17±4
1	C	514	527	526	17±3
1	D	507	520	519	17±3
All	All	40980	42020	41940	1110

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 13.

5 of 354 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:C:347:ILE:HG12	1:C:377:LEU:HD23	0.84	1.48	18	20
1:C:346:SER:HA	1:D:350:MET:HB3	0.80	1.51	12	5
1:A:350:MET:HG3	1:A:374:LEU:HD22	0.80	1.52	12	2
1:B:352:LYS:HB2	1:C:352:LYS:HD3	0.79	1.53	10	1
1:C:350:MET:HG3	1:C:374:LEU:HD22	0.78	1.56	19	3



## 6.3 Torsion angles

### 6.3.1 Protein backbone

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	62/76 (82%)	59±1 (95±2%)	3±1 (5±2%)	0±0 (0±1%)	43	81
1	B	62/76 (82%)	59±1 (95±1%)	3±1 (5±1%)	0±0 (0±1%)	43	81
1	C	62/76 (82%)	59±1 (95±2%)	3±1 (4±2%)	0±0 (0±1%)	43	81
1	D	61/76 (80%)	58±1 (95±2%)	3±1 (5±1%)	0±1 (0±1%)	38	79
All	All	4940/6080 (81%)	4697 (95%)	222 (4%)	21 (0%)	43	81

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	B	330	ASP	3
1	D	330	ASP	3
1	A	330	ASP	3
1	C	330	ASP	2
1	C	384	LEU	2

### 6.3.2 Protein sidechains

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	55/64 (86%)	37±3 (67±6%)	18±3 (33±6%)	1	12
1	B	55/64 (86%)	37±3 (67±5%)	18±3 (33±5%)	1	12
1	C	55/64 (86%)	37±3 (68±5%)	18±3 (32±5%)	1	13
1	D	54/64 (84%)	35±1 (66±3%)	19±1 (34±3%)	1	11
All	All	4380/5120 (86%)	2918 (67%)	1462 (33%)	1	12

5 of 171 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	D	324	SER	20
1	D	382	TYR	20
1	D	330	ASP	20
1	B	332	ARG	20
1	D	332	ARG	19

### 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

No chemical shift data were provided