



Full wwPDB NMR Structure Validation Report ⓘ

Mar 5, 2026 – 05:16 AM UTC

PDB ID : 5HOU / pdb_00005hou
BMRB ID : 30002
Title : Solution Structure of p53TAD-TAZ1
Authors : Krois, A.S.; Ferreon, J.C.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E.
Deposited on : 2016-01-19

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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
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) : 2.49

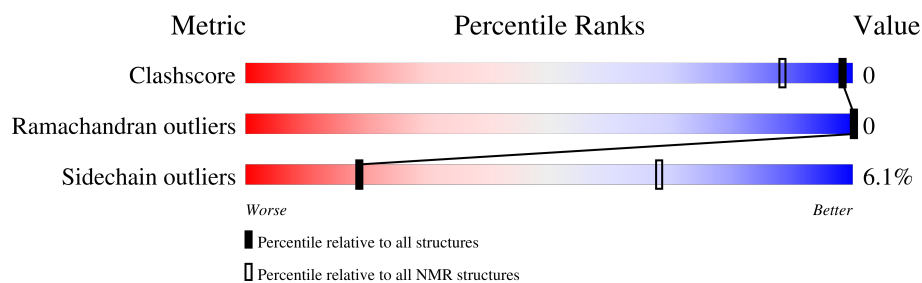
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 83%.

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	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	171	51% . 46%

2 Ensemble composition and analysis

This entry contains 20 models. Model 10 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:1018-A:1026, A:1043-A:1053, A:3349-A:3367, A:3383-A:3435 (92)	0.38	10

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 6 single-model clusters were found.

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2613 atoms, of which 1281 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cellular tumor antigen p53,CREB-binding protein fusion protein.

Mol	Chain	Residues	Atoms						Trace
1	A	171	Total	C	H	N	O	S	0
			2610	814	1281	245	255	15	

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP P04637
A	2	SER	-	expression tag	UNP P04637
A	3	HIS	-	expression tag	UNP P04637
A	2001	GLY	-	linker	UNP P04637
A	2002	SER	-	linker	UNP P04637
A	2003	CYS	-	linker	UNP P04637
A	2004	PHE	-	linker	UNP P04637
A	2005	ASN	-	linker	UNP P04637
A	2006	GLY	-	linker	UNP P04637
A	2007	THR	-	linker	UNP P04637

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

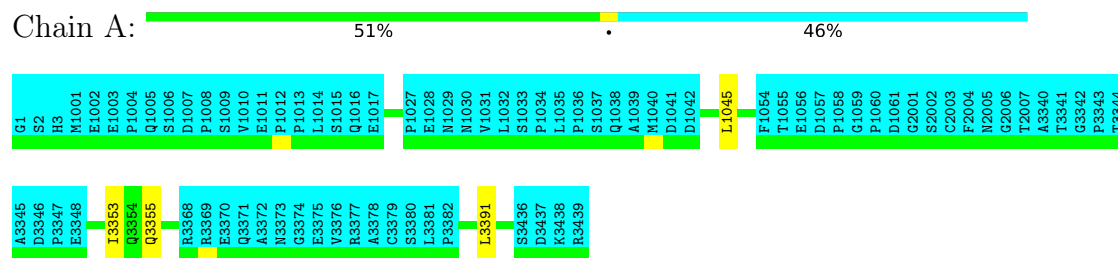
Mol	Chain	Residues	Atoms	
2	A	3	Total	Zn
			3	3

4 Residue-property plots

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: Cellular tumor antigen p53,CREB-binding protein fusion protein

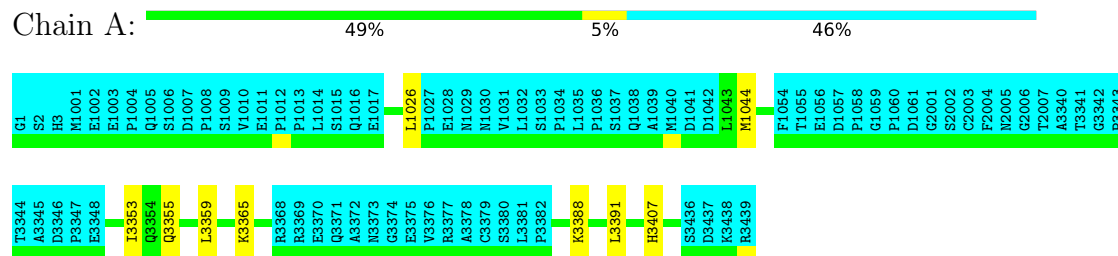


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

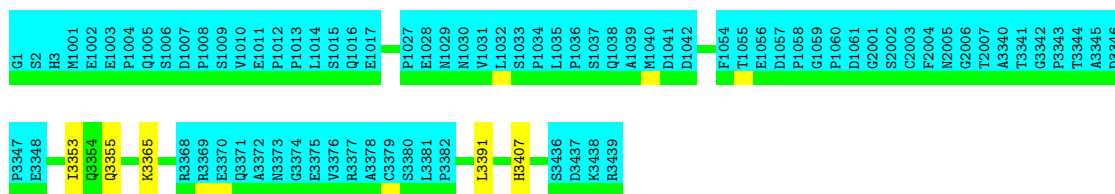
- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.2 Score per residue for model 2

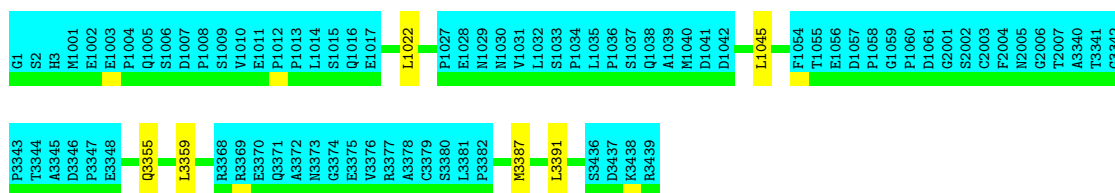
- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein





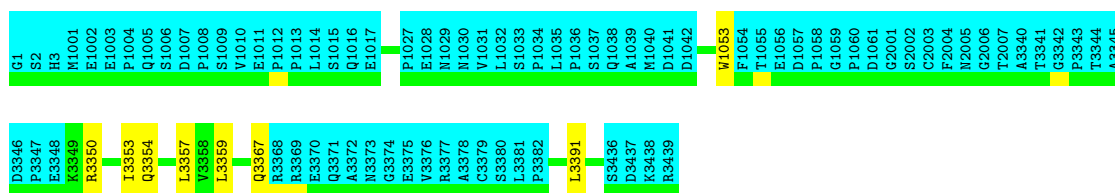
4.2.3 Score per residue for model 3

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



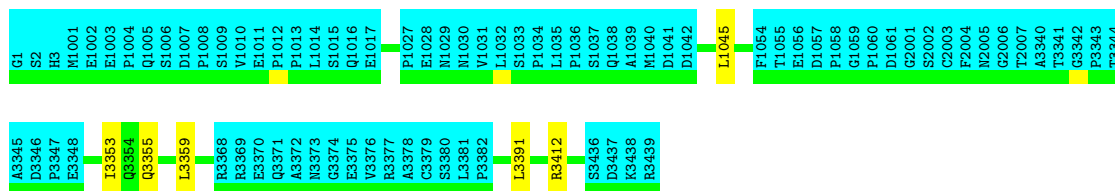
4.2.4 Score per residue for model 4

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.5 Score per residue for model 5

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



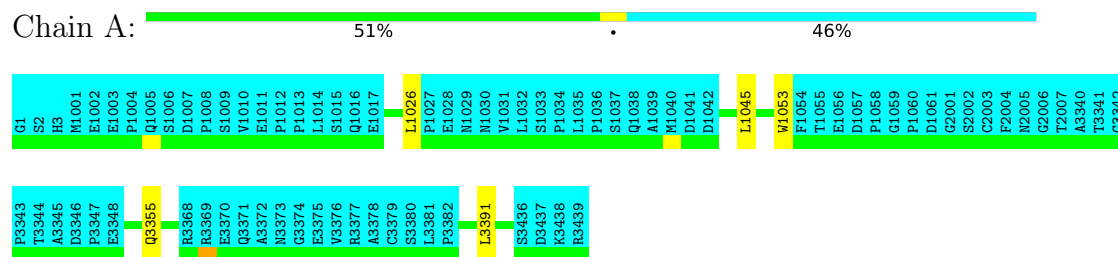
4.2.6 Score per residue for model 6

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



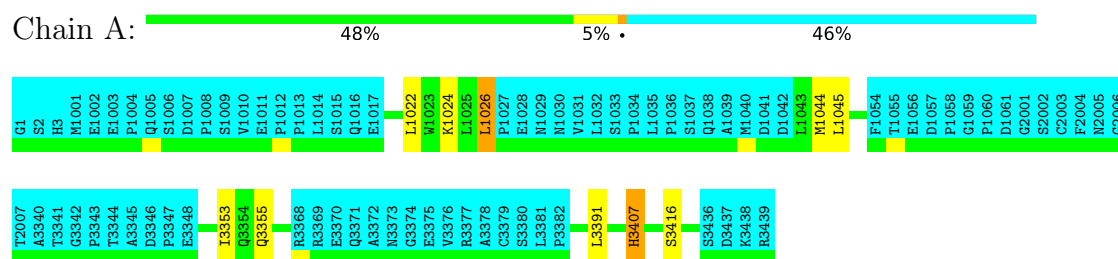
4.2.7 Score per residue for model 7

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



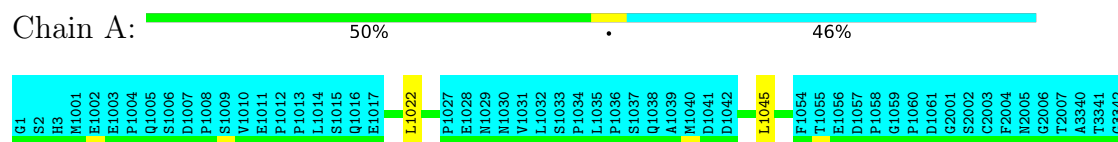
4.2.8 Score per residue for model 8

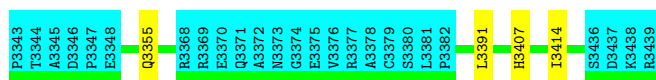
- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.9 Score per residue for model 9

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein

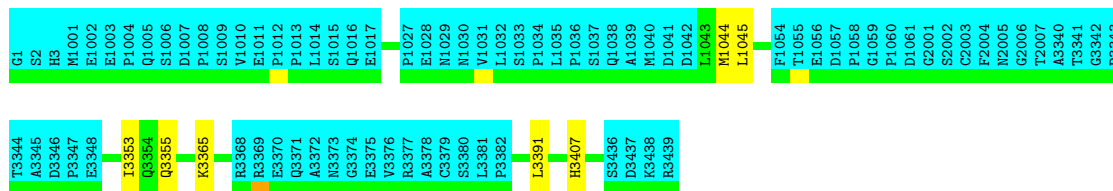




4.2.10 Score per residue for model 10 (medoid)

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein

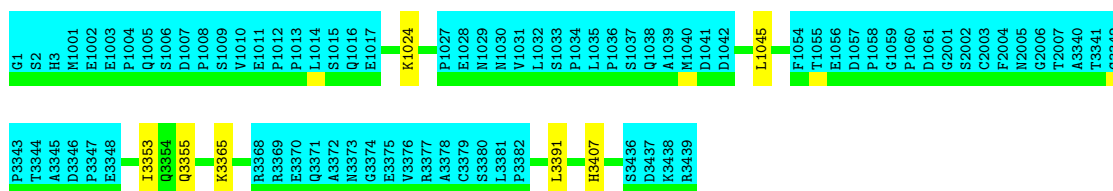
Chain A: 50% . 46%



4.2.11 Score per residue for model 11

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein

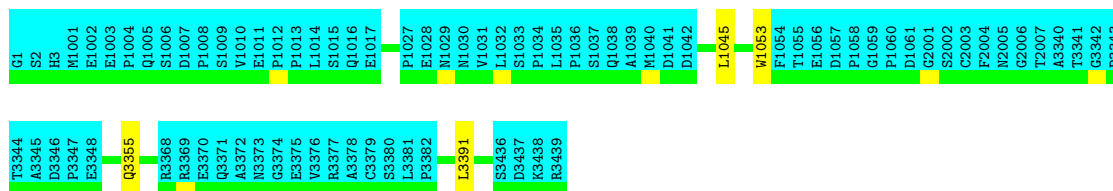
Chain A: 50% . 46%



4.2.12 Score per residue for model 12

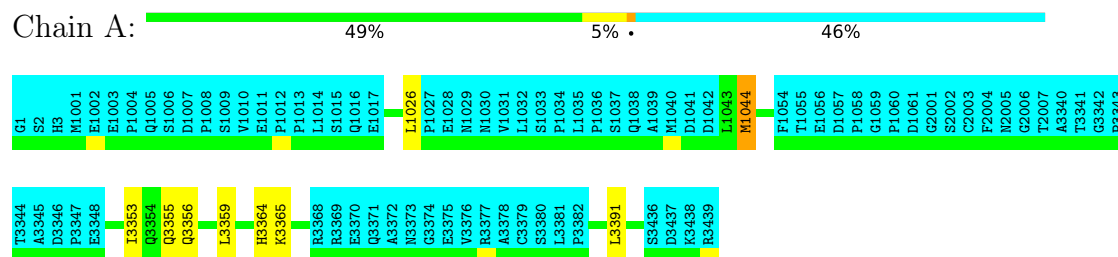
- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein

Chain A: 51% . 46%



4.2.13 Score per residue for model 13

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.14 Score per residue for model 14

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



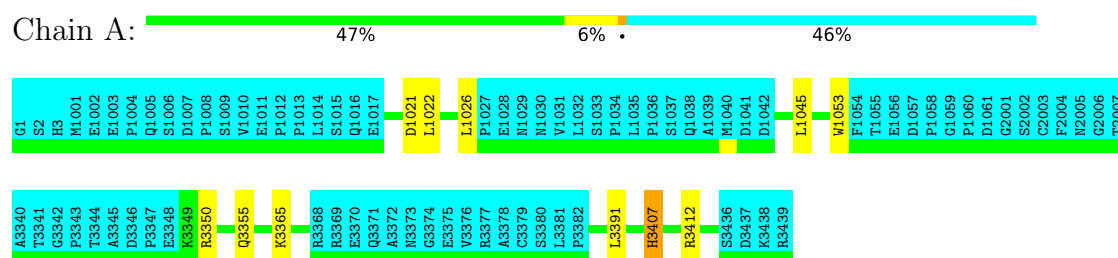
4.2.15 Score per residue for model 15

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.16 Score per residue for model 16

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.17 Score per residue for model 17

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.18 Score per residue for model 18

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



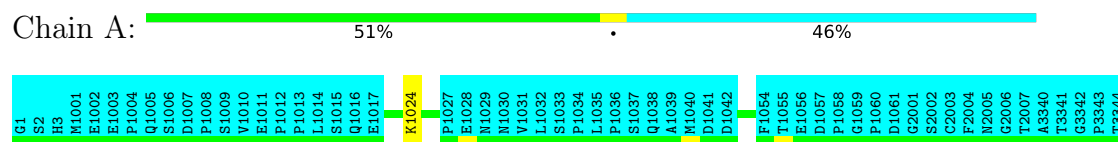
4.2.19 Score per residue for model 19

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



4.2.20 Score per residue for model 20

- Molecule 1: Cellular tumor antigen p53,CREB-binding protein fusion protein



A3345	I3353	R3368
P3346	Q3354	S3369
P3347	Q3355	E3370
E3348	Q3356	Q3371
		A3372
		N3373
		G3374
		S3375
		V3376
		R3377
		A3378
		C3379
		S3380
		L3381
		P3382
		L3391
		S3436
		D3437
		K3438
		R3439

5 Refinement protocol and experimental data overview

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

Of the 200 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
CYANA	structure calculation	2.1
Amber	refinement	12

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1842
Number of shifts mapped to atoms	1842
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	83%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.90±0.01	0±0/761 (0.0± 0.0%)	1.51±0.02	2±2/1030 (0.2± 0.2%)
All	All	0.90	0/15220 (0.0%)	1.51	32/20600 (0.2%)

There are no bond-length outliers.

All 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(°)	Ideal(°)	Models	
								Worst	Total
1	A	3365	LYS	CA-C-N	5.54	127.64	120.44	18	8
1	A	3365	LYS	C-N-CA	5.54	127.64	120.44	18	8
1	A	3407	HIS	CB-CG-CD2	-5.50	124.05	131.20	16	3
1	A	1026	LEU	N-CA-C	5.49	115.91	109.60	16	2
1	A	3367	GLN	OE1-CD-NE2	-5.48	117.12	122.60	4	1
1	A	1021	ASP	CA-C-N	5.39	128.28	120.79	16	1
1	A	1021	ASP	C-N-CA	5.39	128.28	120.79	16	1
1	A	3364	HIS	CB-CG-CD2	-5.29	124.32	131.20	13	1
1	A	3412	ARG	NE-CZ-NH2	5.16	123.85	119.20	16	3
1	A	1044	MET	N-CA-C	5.11	117.31	110.35	1	1
1	A	3350	ARG	NE-CZ-NH2	5.08	123.77	119.20	16	1
1	A	3356	GLN	OE1-CD-NE2	-5.02	117.58	122.60	20	1
1	A	3354	GLN	OE1-CD-NE2	-5.01	117.59	122.60	4	1

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	743	755	750	0±1
All	All	14920	15100	15000	8

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:1044:MET:HE3	1:A:3359:LEU:HD22	0.54	1.79	13	1
1:A:1045:LEU:CD1	1:A:1053:TRP:HE1	0.45	2.24	16	3
1:A:1022:LEU:HD23	1:A:1022:LEU:C	0.42	2.40	3	3
1:A:1044:MET:HE1	1:A:3356:GLN:OE1	0.42	2.14	13	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	92/171 (54%)	90±1 (98±1%)	2±1 (2±1%)	0±0 (0±0%)	100	100
All	All	1840/3420 (54%)	1809 (98%)	31 (2%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	85/153 (56%)	80±2 (94±2%)	5±2 (6±2%)	19 68
All	All	1700/3060 (56%)	1597 (94%)	103 (6%)	19 68

All 19 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	3391	LEU	20
1	A	3355	GLN	18
1	A	3353	ILE	13
1	A	1045	LEU	9
1	A	1026	LEU	7
1	A	3359	LEU	6
1	A	3407	HIS	6
1	A	1044	MET	4
1	A	3387	MET	3
1	A	1053	TRP	3
1	A	1024	LYS	3
1	A	1022	LEU	3
1	A	3350	ARG	2
1	A	3388	LYS	1
1	A	3357	LEU	1
1	A	1051	GLU	1
1	A	3416	SER	1
1	A	3414	ILE	1
1	A	3412	ARG	1

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

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](#)

The completeness of assignment taking into account all chemical shift lists is 83% for the well-defined parts and 82% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *p53taz1_correct_assignments.str*

7.1.1 Bookkeeping [i](#)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1842
Number of shifts mapped to atoms	1842
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	163	-0.29 ± 0.06	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	154	0.18 ± 0.09	None needed (< 0.5 ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
^{15}N	141	-0.39 ± 0.22	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 83%, i.e. 1083 atoms were assigned a chemical shift out of a possible 1310. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	347/455 (76%)	179/182 (98%)	89/184 (48%)	79/89 (89%)
Sidechain	662/748 (89%)	458/488 (94%)	193/228 (85%)	11/32 (34%)

Continued on next page...

Continued from previous page...

	Total	¹H	¹³C	¹⁵N
Aromatic	74/107 (69%)	39/55 (71%)	32/36 (89%)	3/16 (19%)
Overall	1083/1310 (83%)	676/725 (93%)	314/448 (70%)	93/137 (68%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 82%, i.e. 1842 atoms were assigned a chemical shift out of a possible 2255. 0 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹H	¹³C	¹⁵N
Backbone	629/832 (76%)	325/334 (97%)	163/342 (48%)	141/156 (90%)
Sidechain	1123/1288 (87%)	774/832 (93%)	331/403 (82%)	18/53 (34%)
Aromatic	90/135 (67%)	49/69 (71%)	38/48 (79%)	3/18 (17%)
Overall	1842/2255 (82%)	1148/1235 (93%)	532/793 (67%)	162/227 (71%)

7.1.4 Statistically unusual chemical shifts [i](#)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	3359	LEU	CB	31.83	33.11 – 51.34	-5.7

7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:

