



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 12, 2017 – 06:28 pm GMT

PDB ID : 1J4P  
Title : NMR STRUCTURE OF THE FHA1 DOMAIN OF RAD53 IN COMPLEX  
WITH A RAD9-DERIVED PHOSPHOTHREONINE (AT T155) PEPTIDE  
Authors : Yuan, C.; Yongkiettrakul, S.; Byeon, I.-J.L.; Zhou, S.; Tsai, M.-D.  
Deposited on : 2001-10-22

This is a Full wwPDB NMR Structure Validation 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	:	1.7.2 (RC1), CSD as538be (2017)
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk28760
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

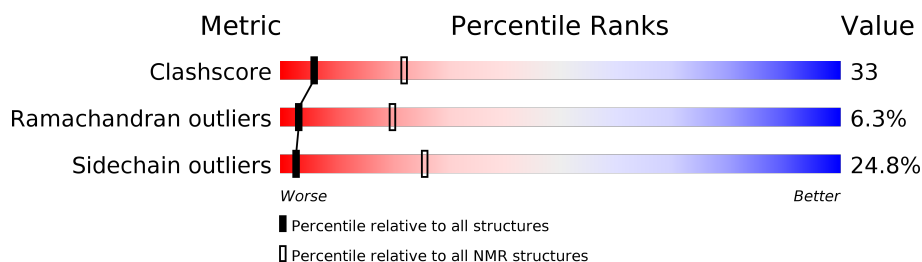
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

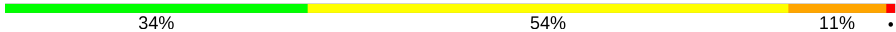
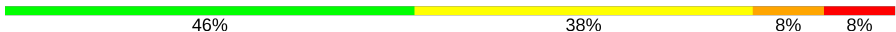
The overall completeness of chemical shifts assignment was not calculated.

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	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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	151	 34% 54% 11% .
2	B	13	 46% 38% 8% 8%

## 2 Ensemble composition and analysis ⓘ

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2644 atoms, of which 1333 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PROTEIN KINASE SPK1.

Mol	Chain	Residues	Atoms						Trace
1	A	151	Total	C	H	N	O	S	0
			2424	751	1224	215	230	4	

- Molecule 2 is a protein called DNA REPAIR PROTEIN RAD9.

Mol	Chain	Residues	Atoms						Trace
2	B	13	Total	C	H	N	O	P	S
			220	68	109	16	25	1	1

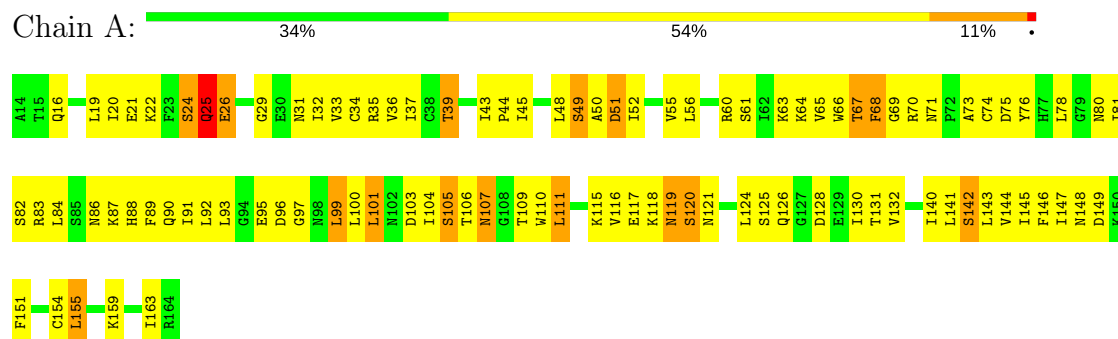
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	171	TPO	THR	MODIFIED RESIDUE	UNP P14737

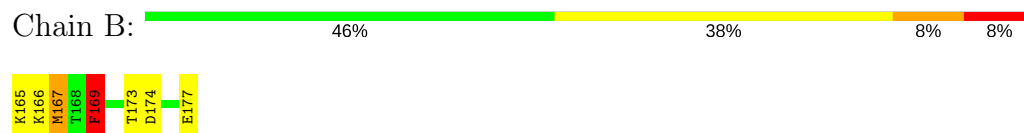
## 4 Residue-property plots [i](#)

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: PROTEIN KINASE SPK1



### • Molecule 2: DNA REPAIR PROTEIN RAD9



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *THE COMPLEX STRUCTURES ARE GENERATED USING A TOTAL OF 2438 RESTRAINTS. AMONG THEM, 3 ARTIFICIAL CONSTRAINTS, 192 TALOS- DERIVED DIHEDRAL ANGLE RESTRAINS, 78 RESTRAINTS FROM H-BOND, 16 INTERMOLECULAR DISTANCE CONSTRAINS, AND 2149 INTRA-FHA1 AND INTRA- PEPTIDE DISTANCE CONSTRAINTS..*

Of the ? calculated structures, 1 were deposited, based on the following criterion: ?.

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

Software name	Classification	Version
XWINNMR	structure solution	2.6
CNS	structure solution	1.0
CNS	refinement	1.0

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 ⓘ

### 6.1 Standard geometry ⓘ

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

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 ⓘ

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	1200	1224	1221	85
2	B	111	109	105	3
All	All	1311	1333	1326	86

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 33.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:33:VAL:HG23	1:A:49:SER:HA	0.89	1.45
1:A:110:TRP:CE2	1:A:115:LYS:HB2	0.78	2.13
1:A:101:LEU:HD21	1:A:130:ILE:HD11	0.72	1.59
1:A:141:LEU:HD23	1:A:142:SER:N	0.72	1.99
1:A:103:ASP:OD2	1:A:116:VAL:HG23	0.70	1.86
1:A:33:VAL:HG23	1:A:49:SER:CA	0.70	2.17
1:A:52:ILE:HD13	1:A:154:CYS:SG	0.69	2.27
1:A:43:ILE:HD13	1:A:78:LEU:HD22	0.68	1.64
1:A:37:ILE:HG23	1:A:45:ILE:HG12	0.65	1.67
1:A:101:LEU:CD1	1:A:124:LEU:HD13	0.64	2.23
1:A:43:ILE:CD1	1:A:78:LEU:HD22	0.64	2.22
1:A:71:ASN:OD1	1:A:73:ALA:HB3	0.63	1.94

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:24:SER:O	1:A:25:GLN:O	0.62	2.18
1:A:51:ASP:O	1:A:55:VAL:HG23	0.60	1.96
1:A:99:LEU:O	1:A:100:LEU:HD12	0.60	1.96
1:A:103:ASP:CG	1:A:118:LYS:HA	0.59	2.17
1:A:43:ILE:HD13	1:A:78:LEU:CD2	0.59	2.28
1:A:33:VAL:HG11	1:A:93:LEU:HD22	0.58	1.74
1:A:33:VAL:CG1	1:A:93:LEU:HD22	0.58	2.28
1:A:81:ILE:HD11	1:A:141:LEU:HD12	0.55	1.77
1:A:32:ILE:HA	1:A:49:SER:HB3	0.54	1.78
1:A:36:VAL:CG2	1:A:48:LEU:HD11	0.54	2.33
1:A:69:GLY:HA3	1:A:86:ASN:O	0.54	2.03
1:A:36:VAL:HG21	1:A:68:PHE:HE2	0.54	1.64
1:A:101:LEU:HD21	1:A:130:ILE:CD1	0.53	2.32
1:A:131:THR:HG22	1:A:140:ILE:CG2	0.53	2.33
1:A:81:ILE:CD1	1:A:141:LEU:HD12	0.53	2.34
1:A:43:ILE:HD12	1:A:143:LEU:HD21	0.53	1.79
1:A:68:PHE:CG	1:A:89:PHE:CE2	0.53	2.97
1:A:36:VAL:HG23	1:A:48:LEU:HD11	0.52	1.80
1:A:34:CYS:SG	1:A:99:LEU:HD22	0.52	2.44
1:A:68:PHE:N	1:A:68:PHE:CD1	0.52	2.78
1:A:106:THR:HG22	2:B:173:THR:HG22	0.51	1.81
1:A:106:THR:CG2	2:B:173:THR:HG22	0.50	2.37
1:A:65:VAL:HG13	1:A:91:ILE:O	0.50	2.06
1:A:88:HIS:CE1	1:A:105:SER:HB3	0.50	2.42
1:A:163:ILE:O	1:A:163:ILE:HG23	0.50	2.07
1:A:36:VAL:HG21	1:A:68:PHE:CE2	0.49	2.42
1:A:83:ARG:O	1:A:107:ASN:ND2	0.49	2.46
1:A:131:THR:CG2	1:A:140:ILE:CG2	0.48	2.92
1:A:96:ASP:O	1:A:155:LEU:HD12	0.48	2.09
1:A:101:LEU:HD11	1:A:124:LEU:HD13	0.48	1.85
1:A:88:HIS:CD2	1:A:132:VAL:CG1	0.48	2.97
2:B:169:PHE:CD1	2:B:169:PHE:N	0.48	2.80
1:A:43:ILE:CG2	1:A:44:PRO:HD2	0.47	2.39
1:A:33:VAL:CG2	1:A:50:ALA:N	0.47	2.77
1:A:100:LEU:HD23	1:A:121:ASN:HB3	0.47	1.86
1:A:65:VAL:HG12	1:A:90:GLN:HG2	0.47	1.86
1:A:68:PHE:CZ	1:A:91:ILE:HD12	0.46	2.45
1:A:103:ASP:O	1:A:119:ASN:N	0.46	2.46
1:A:21:GLU:O	1:A:25:GLN:HG3	0.46	2.10
1:A:68:PHE:CD2	1:A:76:TYR:CB	0.46	2.99
1:A:88:HIS:CD2	1:A:105:SER:OG	0.46	2.68
1:A:97:GLY:HA2	1:A:155:LEU:HD12	0.46	1.88

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:25:GLN:OE1	1:A:25:GLN:N	0.46	2.49
1:A:68:PHE:CD2	1:A:76:TYR:HB2	0.45	2.46
1:A:25:GLN:CD	1:A:26:GLU:N	0.45	2.69
1:A:39:THR:CG2	1:A:144:VAL:HG21	0.45	2.41
1:A:110:TRP:NE1	1:A:115:LYS:HB2	0.45	2.27
1:A:84:LEU:HD21	1:A:132:VAL:HG23	0.45	1.89
1:A:124:LEU:HD21	1:A:147:ILE:HG12	0.44	1.87
1:A:49:SER:O	1:A:64:LYS:CE	0.44	2.65
1:A:33:VAL:CG1	1:A:93:LEU:HD13	0.44	2.42
1:A:48:LEU:HB3	1:A:66:TRP:CE2	0.43	2.49
1:A:66:TRP:O	1:A:90:GLN:CG	0.43	2.66
1:A:65:VAL:HG13	1:A:91:ILE:C	0.43	2.34
1:A:111:LEU:HD13	1:A:128:ASP:OD1	0.43	2.14
1:A:84:LEU:CD2	1:A:132:VAL:HG23	0.43	2.43
1:A:68:PHE:O	1:A:88:HIS:N	0.43	2.52
1:A:109:THR:HG22	1:A:130:ILE:CG2	0.43	2.43
1:A:65:VAL:CG1	1:A:91:ILE:O	0.42	2.67
1:A:130:ILE:N	1:A:143:LEU:O	0.42	2.51
1:A:117:GLU:O	1:A:120:SER:HB2	0.42	2.15
1:A:87:LYS:HB2	1:A:104:ILE:HD12	0.42	1.92
1:A:107:ASN:N	1:A:107:ASN:OD1	0.42	2.52
1:A:91:ILE:HD11	1:A:145:ILE:HD11	0.41	1.91
1:A:103:ASP:HB3	1:A:109:THR:HB	0.41	1.92
1:A:151:PHE:CE2	1:A:155:LEU:HD23	0.41	2.50
1:A:70:ARG:CD	1:A:81:ILE:O	0.41	2.69
1:A:55:VAL:HG13	1:A:63:LYS:HG2	0.41	1.91
1:A:66:TRP:O	1:A:90:GLN:HA	0.41	2.16
1:A:67:THR:O	1:A:75:ASP:N	0.41	2.53
1:A:109:THR:CG2	1:A:130:ILE:CG2	0.41	2.98
1:A:67:THR:HG23	1:A:90:GLN:HG3	0.40	1.92
1:A:88:HIS:O	1:A:105:SER:N	0.40	2.54
1:A:111:LEU:HD12	1:A:111:LEU:C	0.40	2.36

## 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	149/151 (99%)	126 (85%)	17 (11%)	6 (4%)	6	33
2	B	10/13 (77%)	4 (40%)	2 (20%)	4 (40%)	0	0
All	All	159/164 (97%)	130 (82%)	19 (12%)	10 (6%)	3	20

All 10 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	25	GLN
1	A	119	ASN
1	A	29	GLY
1	A	31	ASN
2	B	167	MET
1	A	95	GLU
2	B	169	PHE
1	A	80	ASN
2	B	166	LYS
2	B	174	ASP

### 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	137/137 (100%)	104 (76%)	33 (24%)	3	27
2	B	12/12 (100%)	8 (67%)	4 (33%)	1	12
All	All	149/149 (100%)	112 (75%)	37 (25%)	3	26

All 37 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	25	GLN
2	B	165	LYS
1	A	26	GLU
1	A	68	PHE
1	A	92	LEU
1	A	111	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	49	SER
1	A	19	LEU
1	A	126	GLN
1	A	20	ILE
1	A	101	LEU
1	A	67	THR
1	A	99	LEU
1	A	22	LYS
1	A	149	ASP
1	A	56	LEU
1	A	60	ARG
1	A	125	SER
2	B	167	MET
1	A	148	ASN
1	A	51	ASP
1	A	82	SER
1	A	35	ARG
1	A	105	SER
1	A	120	SER
1	A	155	LEU
1	A	16	GLN
1	A	74	CYS
1	A	146	PHE
1	A	107	ASN
1	A	24	SER
1	A	39	THR
1	A	142	SER
1	A	159	LYS
2	B	169	PHE
1	A	61	SER
2	B	177	GLU

### 6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds

that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	TPO	B	171	2	9,10,11	1.20	0 (0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	TPO	B	171	2	10,14,16	1.05	0 (0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TPO	B	171	2	-	0,8,11,13	0,0,0,0

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 6.5 Carbohydrates

There are no carbohydrates in this entry.

## 6.6 Ligand geometry

There are no ligands in this entry.

## 6.7 Other polymers

There are no such molecules in this entry.

## 6.8 Polymer linkage issues

There are no chain breaks in this entry.

## 7 Chemical shift validation

No chemical shift data were provided