W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

Full wwPDB NMR Structure Validation Report ⓘ

Mar 5, 2022 – 03:04 PM EST

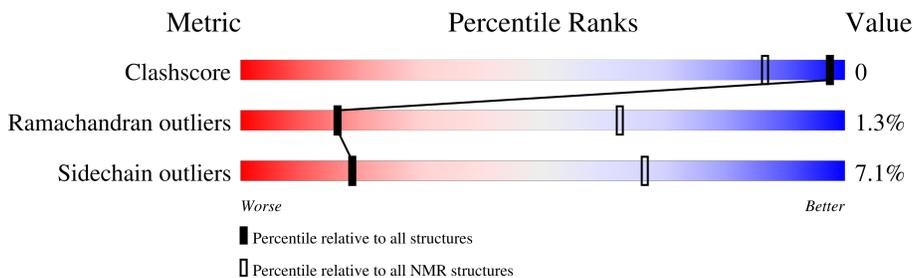|		
PDB ID	:	2K2W
Title	:	Second BRCT domain of NBS1
Authors	:	Xu, C.; Cui, G.; Botuyan, M.; Mer, G.
Deposited on	:	2008-04-14

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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		1.3%
Sidechain outliers		7.1%

Worse　　　　　*Better*

■ Percentile relative to all structures

▯ Percentile relative to all NMR structures

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	118	65%　5%　24%　6%

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 3 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:218-A:250, A:257-A:270, A:289-A:324 (83)	0.29	3

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 2 clusters and 16 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1754 atoms, of which 895 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Recombination and DNA repair protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	111	1754	546	895	143	165	5	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	207	MET	-	expression tag	UNP Q6EKW1
A	208	HIS	-	expression tag	UNP Q6EKW1
A	209	HIS	-	expression tag	UNP Q6EKW1
A	210	HIS	-	expression tag	UNP Q6EKW1
A	211	HIS	-	expression tag	UNP Q6EKW1
A	212	HIS	-	expression tag	UNP Q6EKW1
A	213	HIS	-	expression tag	UNP Q6EKW1
A	214	MET	-	expression tag	UNP Q6EKW1

WORLDWIDE PDB
PROTEIN DATA BANK

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.
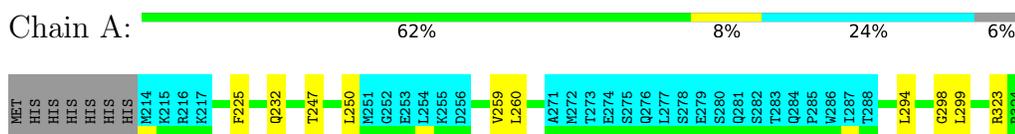
• Molecule 1: Recombination and DNA repair protein
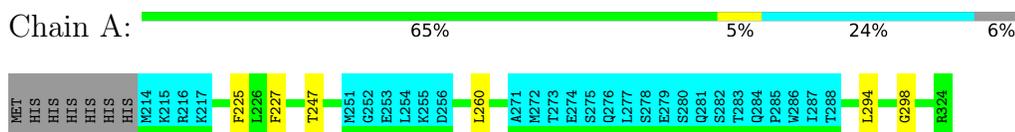
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1
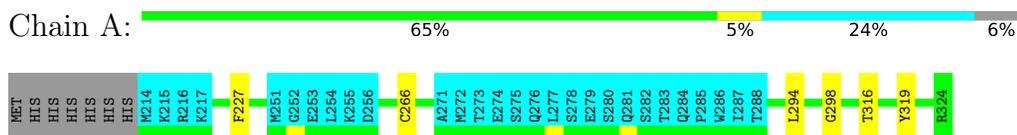
• Molecule 1: Recombination and DNA repair protein

4.2.2 Score per residue for model 2

• Molecule 1: Recombination and DNA repair protein

4.2.3 Score per residue for model 3 (medoid)

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.4 Score per residue for model 4

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.5 Score per residue for model 5

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.6 Score per residue for model 6

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.7 Score per residue for model 7
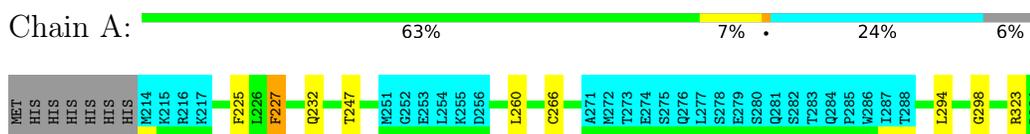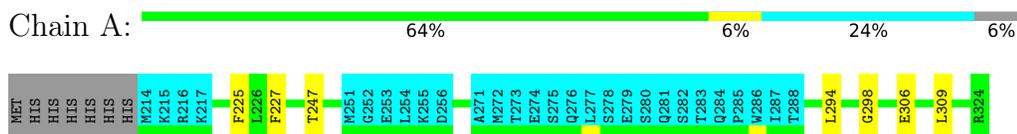
• Molecule 1: Recombination and DNA repair protein

Chain A:

MET HIS HIS HIS HIS HIS M214 K215 R216 K217 F225 T247 D248 L249 I250 M251 G252 E253 L254 K255 D256 L260 A271 M272 T273 E274 S275 Q276 L277 S278 E279 S280 Q281 S282 T283 Q284 P285 W286 I287 T288 G298 Y319 R324

4.2.8 Score per residue for model 8

• Molecule 1: Recombination and DNA repair protein

Chain A: 63%   8%   24%   6%

MET HIS HIS HIS HIS HIS M214 K215 R216 K217 F227 Q232 M251 G252 E253 L254 K255 D256 L260 C266 A271 M272 T273 E274 S275 Q276 L277 S278 E279 S280 Q281 S282 T283 Q284 P285 W286 I287 T288 S289 T290 G298 E304 A305 E306 Y319 R324

4.2.9 Score per residue for model 9

• Molecule 1: Recombination and DNA repair protein

Chain A: 64%   5% •   24%   6%

MET HIS HIS HIS HIS HIS M214 K215 R216 K217 F227 L228 M251 G252 E253 L254 K255 D256 L260 V267 A271 M272 T273 E274 S275 Q276 L277 S278 E279 S280 Q281 S282 T283 Q284 P285 W286 I287 T288 G298 E304 C320 R324

4.2.10 Score per residue for model 10

• Molecule 1: Recombination and DNA repair protein

Chain A: 62%   8%   24%   6%

MET HIS HIS HIS HIS HIS M214 K215 R216 K217 F225 L226 F227 Q232 T247 M251 G252 E253 L254 K255 D256 L260 D261 N262 C266 A271 M272 T273 E274 S275 Q276 L277 S278 E279 S280 Q281 S282 T283 Q284 P285 W286 I287 T288 Q295 G298 R323 R324

4.2.11 Score per residue for model 11

• Molecule 1: Recombination and DNA repair protein

Chain A: 64%   5% •   24%   6%

MET HIS HIS HIS HIS HIS M214 K215 R216 K217 S218 K223 F227 Q232 M251 G252 E253 L254 K255 D256 C266 A271 M272 T273 E274 S275 Q276 L277 S278 E279 S280 Q281 S282 T283 Q284 P285 W286 I287 T288 L294 G298 R324

4.2.12 Score per residue for model 12

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.13 Score per residue for model 13

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.14 Score per residue for model 14
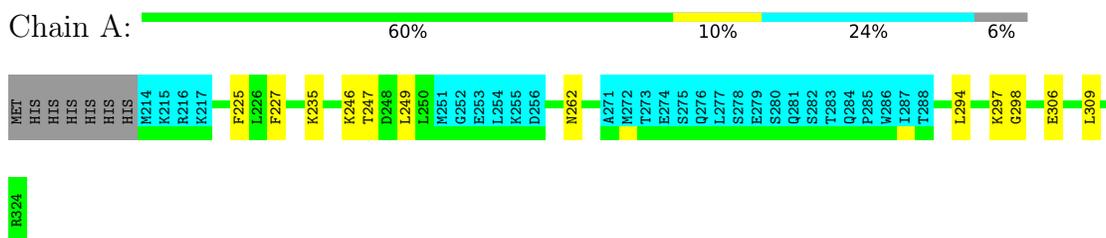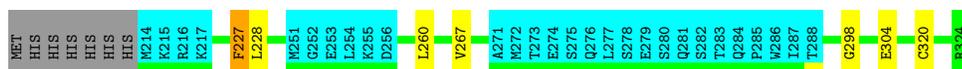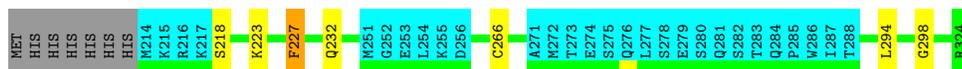
• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.15 Score per residue for model 15

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.16 Score per residue for model 16

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.17 Score per residue for model 17

• Molecule 1: Recombination and DNA repair protein

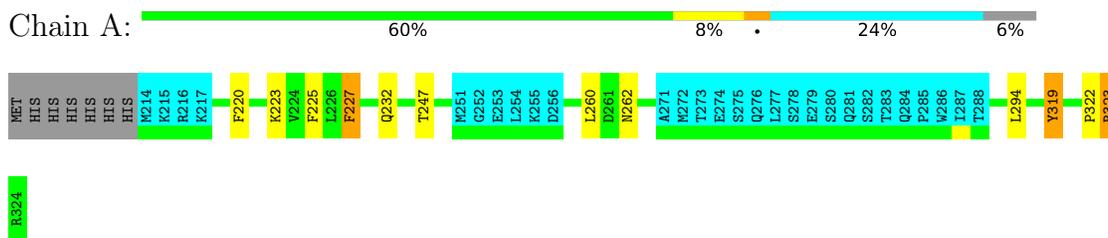Chain A:

4.2.18 Score per residue for model 18

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.19 Score per residue for model 19

• Molecule 1: Recombination and DNA repair protein

Chain A:

4.2.20 Score per residue for model 20

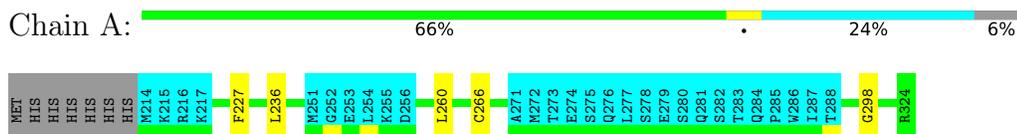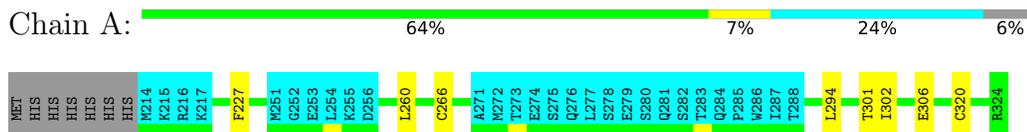• Molecule 1: Recombination and DNA repair protein

Chain A:
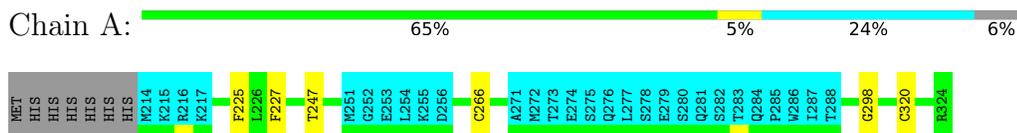
5 Refinement protocol and experimental data overview ⓘ

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

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 solution	
Amber	refinement	8.0

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

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	0.61±0.00	0±0/645 (0.0± 0.0%)	0.93±0.03	1±1/872 (0.1± 0.1%)
All	All	0.61	0/12900 (0.0%)	0.93	11/17440 (0.1%)

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(^{o})	Ideal(^{o})	Models Worst	Total
1	A	319	TYR	CB-CG-CD2	-5.45	117.73	121.00	12	3
1	A	227	PHE	CB-CG-CD2	-5.24	117.14	120.80	8	3
1	A	320	CYS	CA-CB-SG	-5.21	104.63	114.00	20	1
1	A	227	PHE	CB-CG-CD1	5.09	124.36	120.80	17	4

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	636	671	671	1±0
All	All	12720	13420	13420	13

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:225:PHE:CZ	1:A:247:THR:HG22	0.55	2.36	6	12
1:A:223:LYS:HE3	1:A:266:CYS:SG	0.43	2.53	11	1

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	82/118 (69%)	72±2 (88±2%)	9±2 (11±2%)	1±0 (1±1%)	16	63
All	All	1640/2360 (69%)	1442 (88%)	176 (11%)	22 (1%)	16	63

All 6 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	298	GLY	17
1	A	221	LYS	1
1	A	316	THR	1
1	A	218	SER	1
1	A	322	PRO	1
1	A	323	ARG	1

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	72/105 (69%)	67±2 (93±3%)	5±2 (7±3%)	18	67
All	All	1440/2100 (69%)	1338 (93%)	102 (7%)	18	67

All 30 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	227	PHE	13
1	A	260	LEU	11
1	A	294	LEU	11
1	A	266	CYS	8
1	A	232	GLN	7
1	A	323	ARG	6
1	A	319	TYR	6
1	A	306	GLU	6
1	A	262	ASN	4
1	A	309	LEU	3
1	A	320	CYS	3
1	A	249	LEU	2
1	A	304	GLU	2
1	A	295	GLN	2
1	A	302	ILE	2
1	A	241	LEU	2
1	A	250	LEU	1
1	A	259	VAL	1
1	A	299	LEU	1
1	A	235	LYS	1
1	A	246	LYS	1
1	A	297	LYS	1
1	A	290	THR	1
1	A	228	LEU	1
1	A	267	VAL	1
1	A	234	LYS	1
1	A	220	PHE	1
1	A	223	LYS	1
1	A	236	LEU	1
1	A	301	THR	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 monosaccharides 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