



# wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 11:20 PM BST

PDB ID : 2KLR  
Title : Solid-state NMR structure of the alpha-crystallin domain in alphaB-crystallin oligomers  
Authors : Jehle, S.; Rajagopal, P.; Markovic, S.; Bardiaux, B.; Kuehne, R.; Higman, V.A.; Klevit, R.E.; van Rossum, B.; Oschkinat, H.  
Deposited on : 2009-07-08

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

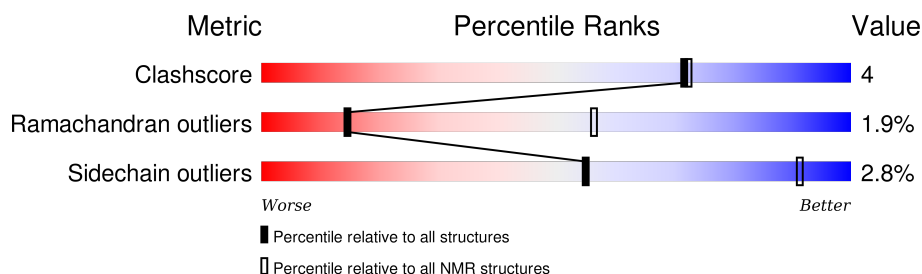
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLID-STATE 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	175	<div> <div style="width: 39%; background-color: green;"></div> <div style="width: 6%; background-color: yellow;"></div> <div style="width: 53%; background-color: grey;"></div> <div style="width: 6%; background-color: cyan;"></div> </div> <div>39% 6% 53%</div>
1	B	175	<div> <div style="width: 40%; background-color: green;"></div> <div style="width: 5%; background-color: yellow;"></div> <div style="width: 53%; background-color: grey;"></div> <div style="width: 5%; background-color: cyan;"></div> </div> <div>40% 5% 53%</div>

## 2 Ensemble composition and analysis ⓘ

This entry contains 10 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:73-A:150, B:73-B:150 (156)	0.80	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 2 clusters and 2 single-model clusters were found.

Cluster number	Models
1	2, 4, 8, 10
2	1, 3, 6, 9
Single-model clusters	5; 7

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

There is only 1 type of molecule in this entry. The entry contains 2670 atoms, of which 1338 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Alpha-crystallin B chain.

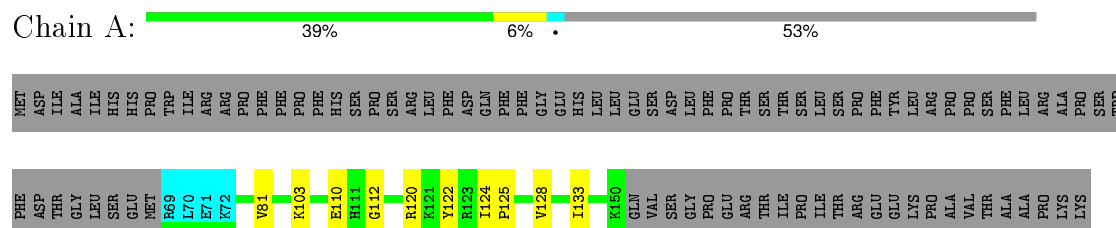
Mol	Chain	Residues	Atoms					Trace
1	A	82	Total	C	H	N	O	0
			1335	417	669	123	126	
1	B	82	Total	C	H	N	O	0
			1335	417	669	123	126	

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

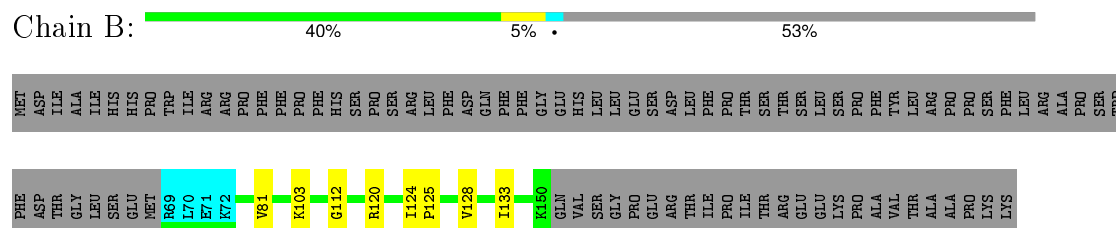
### 4.1 Average score per residue in the NMR ensemble

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: Alpha-crystallin B chain



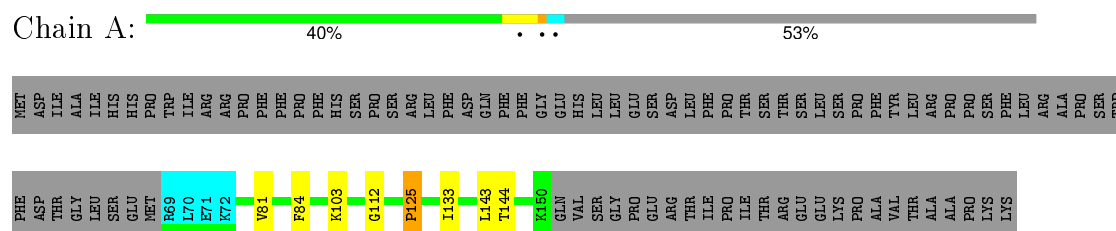
- Molecule 1: Alpha-crystallin B chain



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

The representative model is number 10. Colouring as in section 4.1 above.

- Molecule 1: Alpha-crystallin B chain





MET	ASP	ILE	ALA	ILE	HIS	HIS	PRO	TRP	ILE	ARG	ARG	PRO	PHE	PHE	PHE	HIS	SER	PRO	SER	ARG	LEU	PHE	ASP	GLN	PHE	PHE	GLY	GLU	HIS	LEU	LEU	GLU	SER	ASP	LEU	PHE	PRO	THR	SER	SER	THR	THR	LEU	LEU	SER	PRO	PHE	TYR	LEU	ARG	PRO	PRO	SER	PHE	LEU	ARG	ALA	PRO	SER	TRP
PHE	ASP	THR	GLY	LEU	SER	GLU	MET	R69	L70	E71	K72	V61	F64	K103	K121	P125	I133	L143	T144	K150	GLN	VAL	SER	GLY	PRO	GLU	ARG	THR	ILE	PRO	ILE	THR	ARG	GLU	GLU	LYS	PRO	ALA	VAL	THR	THR	ALA	ALA	PRO	LYS	LYS														

## 5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 10 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	structure solution	
ARIA	structure solution	2.2
ARIA	refinement	2.2
CNS	structure solution	1.2
CNS	refinement	1.2
SOLARIA	structure solution	1

No chemical shift data was provided. Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

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

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	0.97±0.04	0±0/642 (0.0±0.0%)	0.76±0.03	0±0/865 (0.0±0.0%)
1	B	0.97±0.05	0±0/642 (0.0±0.0%)	0.76±0.03	0±0/865 (0.0±0.0%)
All	All	0.97	2/12840 (0.0%)	0.76	0/17300 (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	0.3±0.5
1	B	0.0±0.0	0.4±0.7
All	All	0	7

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	112	GLY	N-CA	-5.79	1.37	1.46	2	1
1	B	112	GLY	N-CA	-5.76	1.37	1.46	2	1

There are no bond-angle outliers.

There are no chirality outliers.

5 of 6 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	B	116	ARG	Sidechain	2
1	B	107	ARG	Sidechain	1
1	B	120	ARG	Sidechain	1
1	A	116	ARG	Sidechain	1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group	Models (Total)
1	A	107	ARG	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	629	626	621	6±2
1	B	629	626	621	6±2
All	All	12580	12520	12420	108

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 4.

5 of 79 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:83:HIS:HA	1:B:140:ASP:O	0.66	1.91	8	4
1:A:83:HIS:HA	1:A:140:ASP:O	0.64	1.91	8	4
1:A:122:TYR:CD2	1:B:112:GLY:HA3	0.59	2.33	3	3
1:B:98:ILE:HG23	1:B:122:TYR:HB2	0.58	1.75	5	1
1:A:98:ILE:HG23	1:A:122:TYR:HB2	0.58	1.75	5	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	77/175 (44%)	70±1 (91±2%)	5±2 (7±2%)	2±1 (2±1%)	14	56
1	B	77/175 (44%)	70±1 (91±2%)	5±2 (7±2%)	1±1 (2±1%)	15	58
All	All	1540/3500 (44%)	1407 (91%)	104 (7%)	29 (2%)	14	56

5 of 10 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	125	PRO	8
1	A	125	PRO	8
1	A	110	GLU	4
1	B	110	GLU	3
1	A	81	VAL	1

### 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	72/161 (45%)	70±2 (97±2%)	2±2 (3±2%)	55	91
1	B	72/161 (45%)	70±2 (97±2%)	2±2 (3±2%)	55	91
All	All	1440/3220 (45%)	1400 (97%)	40 (3%)	55	91

5 of 26 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	B	103	LYS	3
1	A	103	LYS	3
1	B	84	PHE	3
1	A	84	PHE	3
1	A	134	THR	2

### 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