



wwPDB NMR Structure Validation Summary Report ⓘ

Feb 13, 2017 – 12:24 am GMT

PDB ID : 2LSP
Title : solution structures of BRD4 second bromodomain with NF-kB-K310ac peptide
Authors : Zhang, G.; Liu, R.; Zhong, Y.; Plotnikov, A.N.; Zhang, W.; Rusinova, E.; Gerona-Nevarro, G.; Moshkina, N.; Joshua, J.; Chuang, P.Y.; Ohlmeyer, M.; He, J.; Zhou, M.-M.
Deposited on : 2012-05-03

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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

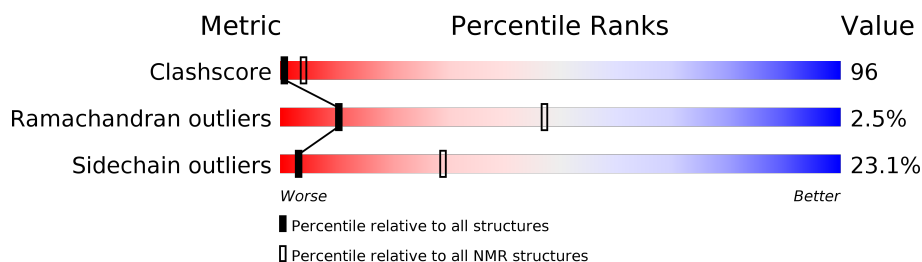
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 82%.

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 ≥ 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	13	100%
2	B	128	16% 45% 18% 20%

2 Ensemble composition and analysis

This entry contains 20 models. Model 17 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	B:353-B:454 (102)	0.10	17

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 4 single-model clusters were found.

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

3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2294 atoms, of which 1141 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called NF-kB-K310ac peptide.

Mol	Chain	Residues	Atoms						Trace
1	A	13	Total	C	H	N	O	S	0
			239	74	123	19	22	1	

- Molecule 2 is a protein called Bromodomain-containing protein 4.

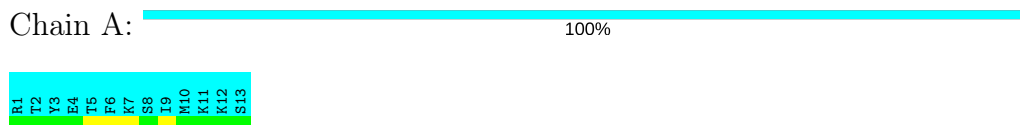
Mol	Chain	Residues	Atoms						Trace
2	B	128	Total	C	H	N	O	S	0
			2055	658	1018	175	193	11	

4 Residue-property plots

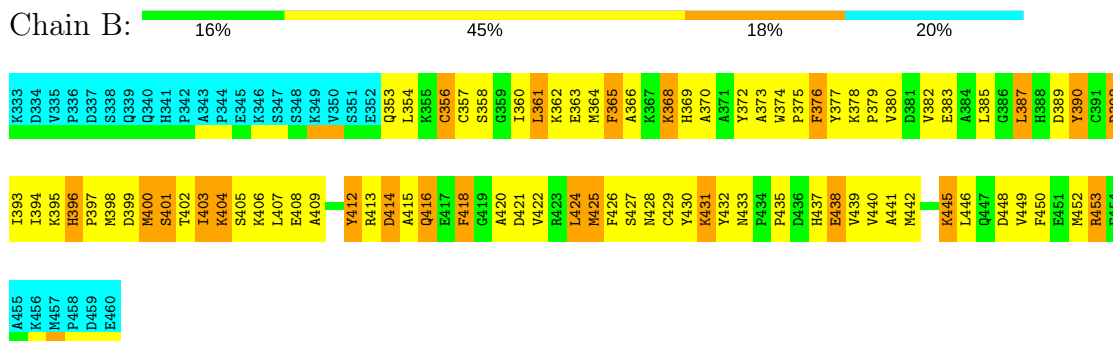
4.1 Average score per residue in the NMR ensemble

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: NF-kB-K310ac peptide



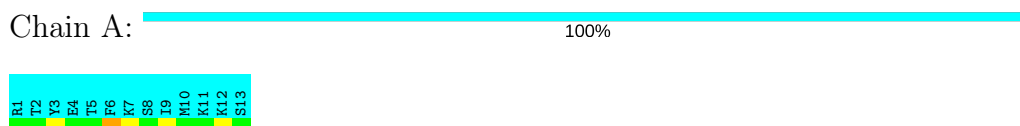
- Molecule 2: Bromodomain-containing protein 4



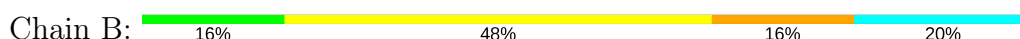
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 17. Colouring as in section 4.1 above.

- Molecule 1: NF-kB-K310ac peptide



- Molecule 2: Bromodomain-containing protein 4



K333	I393	K333
D334	I394	D334
V335	K395	V335
P336	H396	P336
D337	P397	D337
S338	M398	S338
Q339	D399	Q339
Q340	M400	Q340
H341	S401	H341
P342	T402	P342
A343	I403	A343
P344	K404	P344
E345	S405	E345
K346	K406	K346
S347	L407	S347
S348	E408	S348
K349	A409	K349
V350		V350
S351	Y412	S351
E352	R413	E352
Q353	D414	Q353
L354	A415	L354
K355	Q416	K355
C356	E417	C356
C357	F418	C357
S358	G419	S358
G359	A420	G359
L360	D421	L360
L361	V422	L361
K362	R423	K362
E363	L424	E363
K364	M425	K364
F365	F426	F365
A366	S427	A366
K367	N428	K367
C368	C429	C368
H369	Y430	H369
A370	K431	A370
A371	Y432	A371
Y372	N433	Y372
A373	P434	A373
W374	P435	W374
P375	D436	P375
F376	H437	F376
Y377	E438	Y377
K378	V439	K378
P379	V440	P379
V380	A441	V380
D381	M442	D381
Y382	A443	Y382
E383	R444	E383
A384	K445	A384
L385	L446	L385
G386	Q447	G386
L387	D448	L387
H388	V449	H388
D389	F450	D389
Y390		Y390
C391	R453	C391
D392	F454	D392

K455
K456
K457
P458
D459
E460

5 Refinement protocol and experimental data overview

The models were refined using the following method: *DGSA-distance geometry simulated annealing, torsion angle 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
ARIA	refinement	2.3
TALOS	geometry optimization	3.70F1
CNS	structure solution	1.21
CNS	refinement	

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)	2lsp_cs.str
Number of chemical shift lists	1
Total number of shifts	1435
Number of shifts mapped to atoms	1435
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	82%

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](#)

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

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
2	B	0.85±0.01	0±0/860 (0.0±0.0%)	0.93±0.01	3±0/1157 (0.2±0.0%)
All	All	0.85	0/17200 (0.0%)	0.93	56/23140 (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
2	B	412	TYR	CB-CG-CD2	7.52	125.51	121.00	13	20
2	B	412	TYR	CB-CG-CD1	-7.17	116.69	121.00	13	20
2	B	418	PHE	CB-CG-CD1	-5.86	116.70	120.80	5	16

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	0	0	0	0±0
2	B	837	825	821	159±6
All	All	16740	16500	16420	3187

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 96.

5 of 338 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:373:ALA:HB2	2:B:446:LEU:HD21	1.11	1.22	16	20
2:B:372:TYR:HA	2:B:442:MET:SD	0.92	2.03	17	20
2:B:372:TYR:CG	2:B:445:LYS:HD2	0.92	1.98	8	7
2:B:372:TYR:CD2	2:B:445:LYS:HG2	0.91	2.00	12	4
2:B:357:CYS:HB3	2:B:418:PHE:CD1	0.91	2.01	5	20

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	0	-	-	-	-	
2	B	102/128 (80%)	85±1 (83±1%)	15±1 (15±1%)	3±1 (3±1%)	10	47
All	All	2040/2820 (72%)	1691 (83%)	297 (15%)	52 (3%)	10	47

All 5 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	387	LEU	20
2	B	390	TYR	12
2	B	414	ASP	12
2	B	373	ALA	7
2	B	367	LYS	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	0	-	-	-
2	B	89/113 (79%)	68±2 (77±2%)	21±2 (23±2%)	3 29
All	All	1780/2500 (71%)	1368 (77%)	412 (23%)	3 29

5 of 37 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)
2	B	403	ILE	20
2	B	368	LYS	20
2	B	361	LEU	20
2	B	365	PHE	20
2	B	392	ASP	20

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
1	ALY	A	7	1	11,11,12	0.95±0.08	0±0 (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
1	ALY	A	7	1	9,12,14	1.11±0.08	0±0 (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
1	ALY	A	7	1	-	0±0,8,10,12	0±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 [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

The completeness of assignment taking into account all chemical shift lists is 82% for the well-defined parts and 78% for the entire structure.

7.1 Chemical shift list 1

File name: 2lsp_cs.str

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

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	1435
Number of shifts mapped to atoms	1435
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	19

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	127	-0.46 ± 0.20	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	123	0.33 ± 0.14	None needed (< 0.5 ppm)
$^{13}\text{C}'$	10	—	None (insufficient data)
^{15}N	114	0.42 ± 0.31	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 82%, i.e. 1084 atoms were assigned a chemical shift out of a possible 1318. 13 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	397/500 (79%)	192/199 (96%)	110/204 (54%)	95/97 (98%)
Sidechain	582/672 (87%)	367/400 (92%)	213/242 (88%)	2/30 (7%)

Continued on next page...

Continued from previous page...

	Total	¹ H	¹³ C	¹⁵ N
Aromatic	105/146 (72%)	68/76 (89%)	36/61 (59%)	1/9 (11%)
Overall	1084/1318 (82%)	627/675 (93%)	359/507 (71%)	98/136 (72%)

7.1.4 Statistically unusual chemical shifts [i](#)

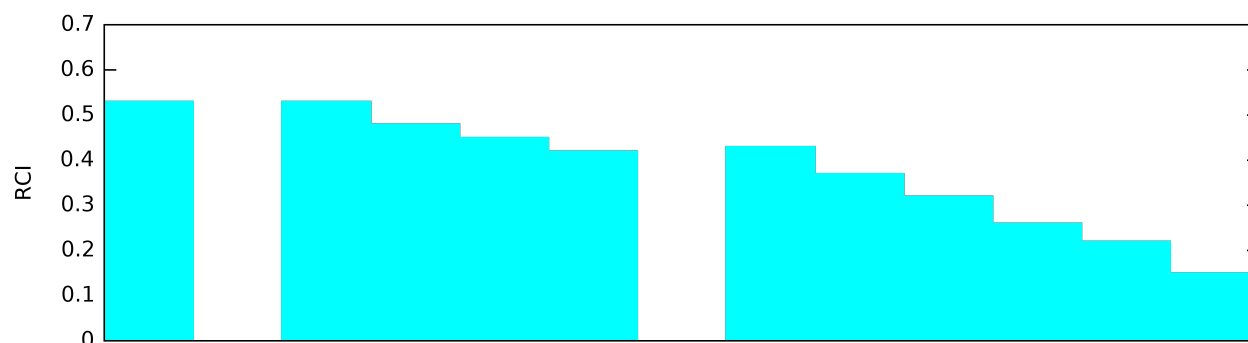
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
2	B	347	SER	C	8.45	183.48 – 165.88	-94.4
2	B	386	GLY	C	7.90	183.33 – 164.53	-88.3
2	B	346	LYS	C	8.51	186.58 – 166.78	-84.9
2	B	385	LEU	C	7.48	187.06 – 166.96	-84.3
2	B	430	TYR	C	7.61	185.42 – 165.42	-83.9
2	B	447	GLN	C	11.57	186.20 – 166.50	-83.6
2	B	388	HIS	C	8.30	185.27 – 165.27	-83.5
2	B	437	HIS	C	9.41	185.27 – 165.27	-82.9
2	B	384	ALA	C	7.86	188.57 – 166.97	-78.7
2	B	366	ALA	C	9.11	188.57 – 166.97	-78.1
2	B	375	PRO	HG2	-0.72	3.48 – 0.38	-8.6
2	B	447	GLN	HB2	0.07	3.30 – 0.80	-7.9
2	B	375	PRO	HB2	-0.35	3.82 – 0.32	-6.9
2	B	357	CYS	HB2	0.35	5.20 – 0.70	-5.8
2	B	422	VAL	HG11	-0.67	2.13 – -0.47	-5.8
2	B	422	VAL	HG13	-0.67	2.13 – -0.47	-5.8
2	B	422	VAL	HG12	-0.67	2.13 – -0.47	-5.8
2	B	375	PRO	HD2	1.62	5.45 – 1.85	-5.6
2	B	439	VAL	H	11.70	11.69 – 4.89	5.0

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

The images below report *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.

Random coil index (RCI) for chain A:



Random coil index (RCI) for chain B:

