wwPDB NMR Structure Validation Summary Report (i)

May 29, 2020 – 01:45 am BST

PDB ID	:	2XKS
Title	:	Prion-like conversion during amyloid formation at atomic resolution
Authors	:	Eichner, T.; Kalverda, A.P.; Thompson, G.S.; Radford, S.E.; Homans, S.W.
Deposited on	:	2010-07-12

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
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

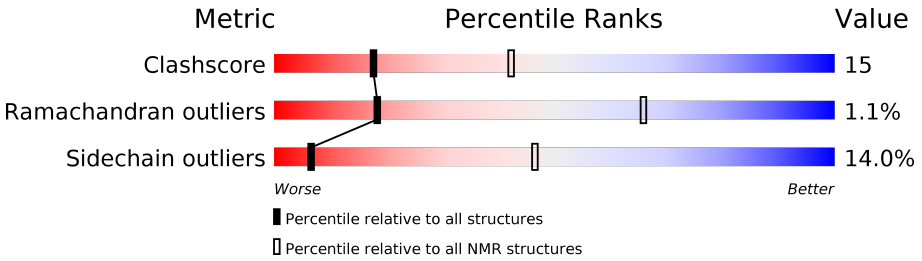
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 91%.

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	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	99	52% 31% • 15%

2 Ensemble composition and analysis (i)

This entry contains 30 models. Model 19 is the overall representative, medoid model (most similar to other models).

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:3-A:52, A:62-A:95 (84)	0.24	19

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 3 single-model clusters were found.

Cluster number	Models
1	3, 4, 5, 12, 19, 20, 21, 23, 24, 26, 29, 30
2	7, 10, 11, 13, 14, 16, 17, 18, 28
3	1, 6, 9, 15, 25, 27
Single-model clusters	2; 8; 22

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1618 atoms, of which 791 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called BETA-2-MICROGLOBULIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	99	1618	528	791	140	156	3	0

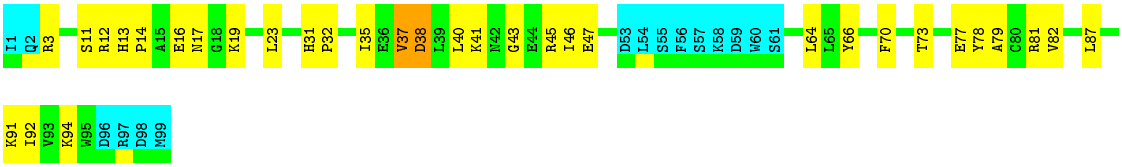
4 Residue-property plots (i)

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: BETA-2-MICROGLOBULIN

Chain A:

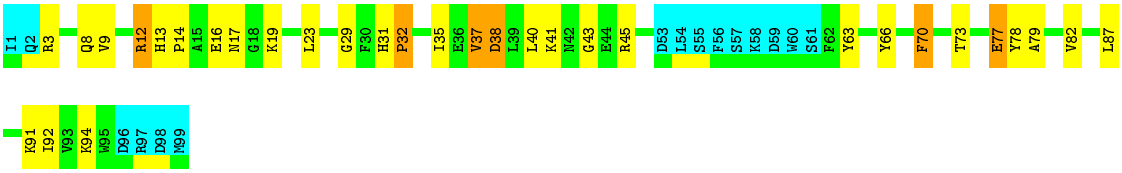


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 19. Colouring as in section 4.1 above.

- Molecule 1: BETA-2-MICROGLOBULIN

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *NONE*.

Of the 30 calculated structures, 30 were deposited, based on the following criterion: *LOWEST ENERGY*.

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

Software name	Classification	Version
CcpNmr Analysis	refinement	1.0
PASD ALGORITHM	structure solution	1.0
NMRView	structure solution	5.22
NMRPipe	structure solution	2.5
ARIA	structure solution	2.1
CcpNmr Analysis	structure solution	2.1
Xplor-NIH	structure solution	2.17
VNMR	structure solution	6.1C

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)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	1268
Number of shifts mapped to atoms	1260
Number of unparsed shifts	0
Number of shifts with mapping errors	8
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	91%

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)

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 (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	700	677	677	21±4
All	All	21000	20310	20310	620

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 15.

5 of 149 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:16:GLU:HB3	1:A:19:LYS:HB3	0.96	1.38	28	17
1:A:39:LEU:HD22	1:A:80:CYS:SG	0.85	2.12	23	1
1:A:41:LYS:HE2	1:A:46:ILE:HG12	0.82	1.51	22	5
1:A:87:LEU:HD13	1:A:91:LYS:HB2	0.76	1.58	9	29
1:A:13:HIS:HB3	1:A:14:PRO:HD2	0.75	1.58	21	28

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	84/99 (85%)	76±2 (91±2%)	7±2 (8±2%)	1±1 (1±1%)	18	66
All	All	2520/2970 (85%)	2289 (91%)	204 (8%)	27 (1%)	18	66

5 of 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	32	PRO	20
1	A	33	SER	3
1	A	47	GLU	1
1	A	86	THR	1
1	A	62	PHE	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	79/94 (84%)	68±2 (86±2%)	11±2 (14±2%)	6	46
All	All	2370/2820 (84%)	2038 (86%)	332 (14%)	6	46

5 of 42 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	70	PHE	30
1	A	37	VAL	30
1	A	23	LEU	29
1	A	38	ASP	24
1	A	35	ILE	23

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 (i)

The completeness of assignment taking into account all chemical shift lists is 91% for the well-defined parts and 89% for the entire structure.

7.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

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	1268
Number of shifts mapped to atoms	1260
Number of unparsed shifts	0
Number of shifts with mapping errors	8
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. First 5 (of 8) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	1	MET	CB	34.188	0.056	1
A	1	MET	HA	3.851	0.001	1
A	1	MET	HG3	2.467	0.001	1
A	1	MET	CG	31.208	0.043	1
A	1	MET	CA	55.221	0.065	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	100	0.15 ± 0.13	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	97	-0.35 ± 0.16	None needed (< 0.5 ppm)
$^{13}C'$	89	0.70 ± 0.11	Should be applied

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
^{15}N	89	-0.84 ± 0.50	None needed (imprecise)

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 91%, i.e. 987 atoms were assigned a chemical shift out of a possible 1090. 13 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	401/410 (98%)	162/163 (99%)	162/168 (96%)	77/79 (97%)
Sidechain	485/556 (87%)	302/329 (92%)	177/201 (88%)	6/26 (23%)
Aromatic	101/124 (81%)	54/66 (82%)	46/53 (87%)	1/5 (20%)
Overall	987/1090 (91%)	518/558 (93%)	385/422 (91%)	84/110 (76%)

7.1.4 Statistically unusual chemical shifts (i)

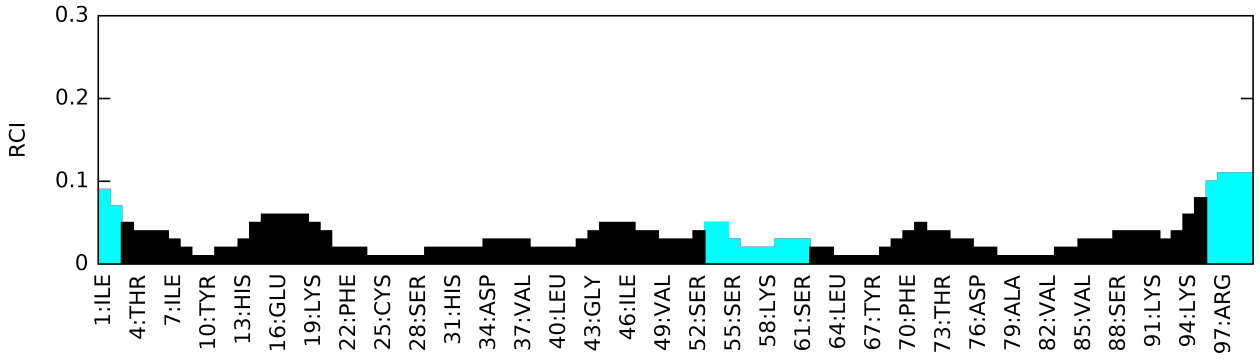
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	33	SER	HB3	1.76	5.25 – 2.45	-7.5
1	A	23	LEU	HB3	-0.84	3.34 – -0.26	-6.6
1	A	89	GLN	CD	171.37	186.62 – 172.72	-6.0
1	A	50	GLU	HB3	0.75	3.10 – 0.90	-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.

Random coil index (RCI) for chain A:


RCI
0.3
0.2
0.1
0
1:ILE
4:THR
7:ILE
10:TYR
13:HIS
16:GLU
19:LYS
22:PHE
25:CYS
28:SER
31:HIS
34:ASP
37:VAL
40:LEU
43:GLY
46:ILE
49:VAL
52:SER
55:SER
58:LYS
61:SER
64:LEU
67:TYR
70:PHE
73:THR
76:ASP
79:ALA
82:VAL
85:VAL
88:SER
91:LYS
94:LYS
97:ARG