wwPDB NMR Structure Validation Summary Report (i)

Apr 20, 2024 – 11:17 PM EDT

PDB ID	:	2LTC
BMRB ID	:	18467
Title	:	Fas1-4, R555W
Authors	:	Underhaug, J.; Nielsen, N.; Runager, K.
Deposited on	:	2012-05-16

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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

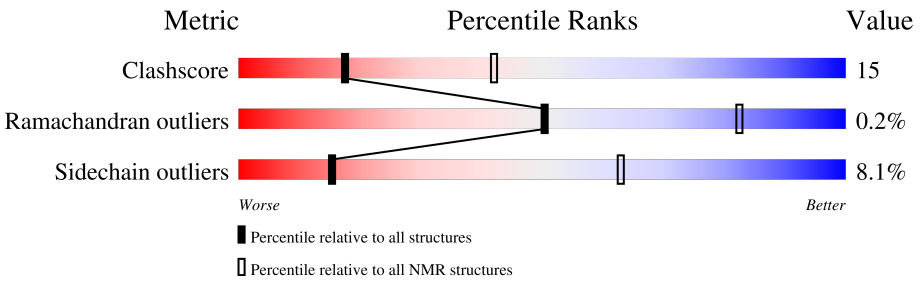
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 84%.

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	135	66% 30% .

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 2 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:503-A:632 (130)	0.59	2

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. No single-model clusters were found.

Cluster number	Models
1	1, 2, 5, 7, 8, 10
2	4, 9
3	3, 6

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2065 atoms, of which 1049 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Transforming growth factor-beta-induced protein ig-h3.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	135	2065	641	1049	177	194	4	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	500	ALA	-	expression tag	UNP Q15582
A	501	GLY	-	expression tag	UNP Q15582
A	555	TRP	ARG	engineered mutation	UNP Q15582
A	634	ALA	PRO	engineered mutation	UNP Q15582

4 Residue-property plots (i)

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: Transforming growth factor-beta-induced protein ig-h3

Chain A: 66% 30% •

A500 G501 M502 V505 M506 D507 V508 L509 F515 L518 Q523 E529 N532 V536 V539 F540 N544 L550 E554 W555 L558 L559 K563 E564 L565 L569 K570 Y571 H572 I573 I577 S580 I583 R588 L589 L592 L597 E598 L601

S607 V608 N609 V613 D617 H626 V627 I628 L632 Q633 A634

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

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: Transforming growth factor-beta-induced protein ig-h3

Chain A: 66% 25% 5% •

A500 G501 M502 V505 L509 F515 L518 I522 Q523 L527 T528 E529 T530 L531 N532 V539 F540 L550 E554 W555 L558 L559 G560 D561 A562 K563 E564 L565 L569 K570 Y571 H572 I573 S580 I583 R588 L589 L592 K596 L597 L601 V608

N609 V613 D617 H626 V627 I628 L632 Q633 A634

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *DGSA-distance geometry simulated annealing.*

Of the 100 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
ARIA	structure solution	
CNSSOLVE	structure solution	
ARIA	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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1501
Number of shifts mapped to atoms	1500
Number of unparsed shifts	0
Number of shifts with mapping errors	1
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	84%

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	984	1017	1017	30±4
All	All	9840	10170	10170	302

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 160 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:609:ASN:HB2	1:A:632:LEU:HD23	0.89	1.43	6	1
1:A:609:ASN:HD22	1:A:632:LEU:HD13	0.88	1.28	2	5
1:A:536:VAL:HG23	1:A:577:ILE:HG13	0.86	1.47	5	1
1:A:508:VAL:HG11	1:A:625:VAL:HG21	0.82	1.50	1	1
1:A:583:ILE:HG22	1:A:601:LEU:HB2	0.77	1.54	8	8

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	130/135 (96%)	123±2 (95±1%)	7±2 (5±1%)	0±0 (0±0%)	50 82
All	All	1300/1350 (96%)	1231 (95%)	66 (5%)	3 (0%)	50 82

All 2 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	585	ALA	2
1	A	617	ASP	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	108/110 (98%)	99±3 (92±2%)	9±3 (8±2%)	15 63
All	All	1080/1100 (98%)	992 (92%)	88 (8%)	15 63

5 of 29 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	555	TRP	7
1	A	559	LEU	7
1	A	563	LYS	7
1	A	628	ILE	7
1	A	523	GLN	6

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 monosaccharides 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 84% for the well-defined parts and 82% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *FAS1-4,_R555W,_shifts*

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	1501
Number of shifts mapped to atoms	1500
Number of unparsed shifts	0
Number of shifts with mapping errors	1
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

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

- No matching atom found in the structure. All 1 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	572	HIS	HD1	11.554	0.004	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
$^{13}C_{\alpha}$	133	-0.36 ± 0.08	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	119	0.14 ± 0.12	None needed (< 0.5 ppm)
$^{13}C'$	118	-0.10 ± 0.14	None needed (< 0.5 ppm)
^{15}N	124	0.52 ± 0.32	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 84%, i.e. 1474 atoms were assigned a chemical shift out of a possible 1760. 0 out of 32 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	628/651 (96%)	259/266 (97%)	247/260 (95%)	122/125 (98%)
Sidechain	790/1034 (76%)	549/679 (81%)	235/319 (74%)	6/36 (17%)
Aromatic	56/75 (75%)	30/37 (81%)	24/34 (71%)	2/4 (50%)
Overall	1474/1760 (84%)	838/982 (85%)	506/613 (83%)	130/165 (79%)

7.1.4 Statistically unusual chemical shifts (i)

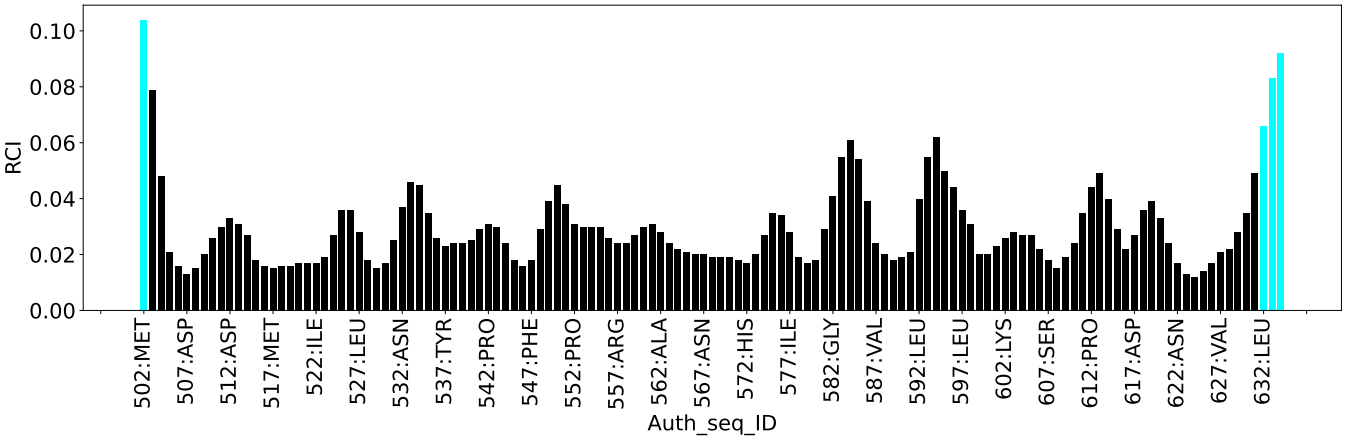
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	538	THR	HG1	7.16	0.08 – 2.19	28.5
1	A	543	THR	HG1	5.28	0.08 – 2.19	19.6
1	A	504	THR	HG1	5.05	0.08 – 2.19	18.5
1	A	528	THR	HG1	4.76	0.08 – 2.19	17.2
1	A	530	THR	HG1	4.76	0.08 – 2.19	17.2

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:


RCI
0.00
0.02
0.04
0.06
0.08
0.10
502:MET
507:ASP
512:ASP
517:MET
522:ILE
527:LEU
532:ASN
537:TYR
542:PRO
547:PHE
552:PRO
557:ARG
562:ALA
567:ASN
572:HIS
577:ILE
582:GLY
587:VAL
592:LEU
597:LEU
602:LYS
607:SER
612:PRO
617:ASP
622:ASN
627:VAL
632:LEU
Auth_seq_ID