WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

Nov 17, 2024 – 04:30 PM EST

|		
PDB ID	:	2MCY
BMRB ID	:	19458
Title	:	CR1 Sushi domains 2 and 3
Authors	:	Park, H.J.; Guariento, M.J.; Maciejewski, M.; Hauart, R.; Tham, W.; Cowman, A.F.; Schmidt, C.Q.; Martens, H.; Liszewski, K.M.; Hourcade, D.; Barlow, P.N.; Atkinson, J.P.
Deposited on	:	2013-08-27

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 28%.

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		18
Ramachandran outliers		4.8%
Sidechain outliers		3.0%

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures
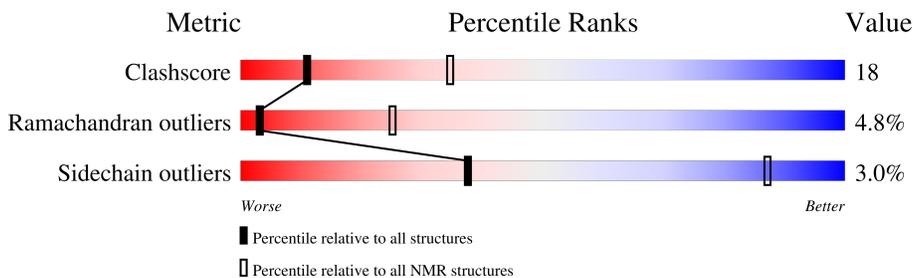
Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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	136	65% 32% • •

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 4 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:61-A:192 (132)	1.70	4

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 3 single-model clusters were found.

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

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 1957 atoms, of which 960 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Complement receptor type 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	132	1957	623	960	172	193	9	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	57	GLU	-	expression tag	UNP P17927
A	58	ALA	-	expression tag	UNP P17927
A	59	GLU	-	expression tag	UNP P17927
A	60	ALA	-	expression tag	UNP P17927
A	115	THR	ASN	conflict	UNP P17927

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: Complement receptor type 1
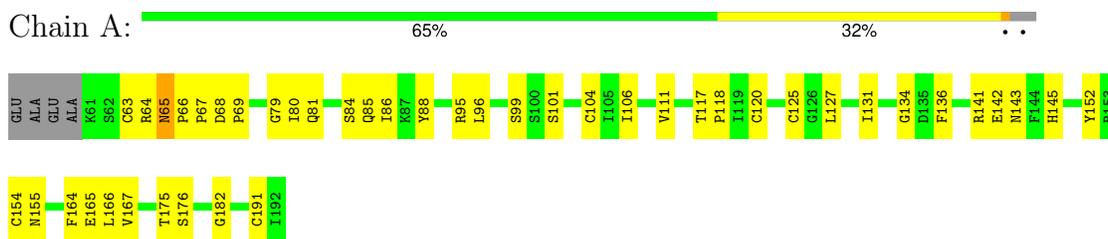
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

• Molecule 1: Complement receptor type 1
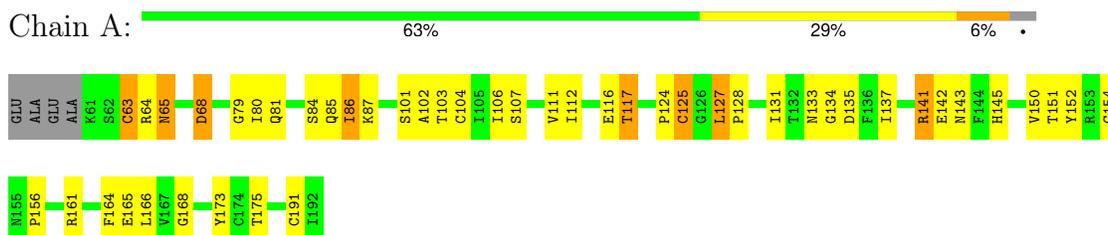
5 Refinement protocol and experimental data overview ⓘ

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

Of the 100 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
X-PLOR NIH	structure solution	
X-PLOR NIH	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	486
Number of shifts mapped to atoms	485
Number of unparsed shifts	0
Number of shifts with mapping errors	1
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	28%

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	0.82±0.04	0±0/1022 (0.0± 0.0%)	0.74±0.04	0±1/1392 (0.0± 0.0%)
All	All	0.82	0/20440 (0.0%)	0.74	6/27840 (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	0.2±0.4
All	All	0	4

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	88	TYR	CB-CG-CD1	-5.85	117.49	121.00	20	2
1	A	88	TYR	CB-CG-CD2	-5.58	117.65	121.00	15	1
1	A	104	CYS	CB-CA-C	-5.45	99.49	110.40	9	1
1	A	104	CYS	CA-CB-SG	5.30	123.55	114.00	9	2

There are no chirality outliers.

All 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	A	153	ARG	Sidechain	3
1	A	95	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	997	960	959	35±7
All	All	19940	19200	19180	705

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 18.

5 of 404 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:122:ARG:HB3	1:A:145:HIS:CE1	0.90	2.01	12	3
1:A:106:ILE:HG12	1:A:111:VAL:HG22	0.89	1.45	18	5
1:A:64:ARG:O	1:A:65:ASN:HB3	0.79	1.77	2	11
1:A:67:PRO:CG	1:A:117:THR:HG21	0.78	2.08	5	1
1:A:166:LEU:HD21	1:A:170:PRO:HB3	0.78	1.54	15	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	130/136 (96%)	113±2 (87±1%)	11±2 (8±2%)	6±2 (5±1%)	3	25
All	All	2600/2720 (96%)	2262 (87%)	212 (8%)	126 (5%)	3	25

5 of 34 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	117	THR	20
1	A	65	ASN	16
1	A	145	HIS	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	176	SER	8
1	A	66	PRO	7

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	115/117 (98%)	112±2 (97±2%)	4±2 (3±2%)	37	87
All	All	2300/2340 (98%)	2230 (97%)	70 (3%)	37	87

5 of 24 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	125	CYS	10
1	A	63	CYS	8
1	A	65	ASN	7
1	A	169	GLU	5
1	A	155	ASN	4

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 oligosaccharides 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 ⓘ

The completeness of assignment taking into account all chemical shift lists is 28% for the well-defined parts and 28% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

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	486
Number of shifts mapped to atoms	485
Number of unparsed shifts	0
Number of shifts with mapping errors	1
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	101

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	120	CYS	HG	1.7	.	.

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	75	0.53 ± 0.77	None needed (imprecise)
$^{13}C_{\beta}$	58	-2.38 ± 0.76	Should be checked
$^{13}C'$	0	—	None (insufficient data)
^{15}N	67	-1.08 ± 0.86	None needed (imprecise)

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 28%, i.e. 470 atoms were assigned a chemical shift out of a possible 1684. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	256/653 (39%)	122/268 (46%)	75/264 (28%)	59/121 (49%)
Sidechain	201/908 (22%)	130/591 (22%)	71/283 (25%)	0/34 (0%)
Aromatic	13/123 (11%)	9/60 (15%)	4/59 (7%)	0/4 (0%)
Overall	470/1684 (28%)	261/919 (28%)	150/606 (25%)	59/159 (37%)

7.1.4   Statistically unusual chemical shifts   ⓘ

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	180	GLN	CD	50.89	173.59 − 185.85	-105.1
1	A	136	PHE	CD2	13.65	125.53 − 137.61	-97.6
1	A	136	PHE	CD1	13.65	125.33 − 137.83	-94.3
1	A	152	TYR	CD1	25.83	125.84 − 139.60	-77.7
1	A	152	TYR	CD2	25.83	125.28 − 140.14	-71.9
1	A	143	ASN	CG	25.14	164.52 − 188.90	-62.2
1	A	184	TRP	CG	25.21	101.31 − 120.62	-44.4
1	A	152	TYR	CG	27.97	112.42 − 146.96	-29.4
1	A	118	PRO	CB	65.42	26.06 − 37.61	29.1
1	A	188	ALA	CB	68.71	10.19 − 27.75	28.3
1	A	156	PRO	CB	64.01	26.06 − 37.61	27.9
1	A	158	SER	CB	24.48	56.28 − 71.32	-26.1
1	A	178	ASP	CG	27.50	149.18 − 208.82	-25.4
1	A	127	LEU	HD11	7.39	-0.61 − 2.12	24.3
1	A	127	LEU	HD12	7.39	-0.61 − 2.12	24.3
1	A	127	LEU	HD13	7.39	-0.61 − 2.12	24.3
1	A	132	THR	CB	31.32	61.12 − 78.27	-22.4
1	A	129	PRO	CG	51.66	21.69 − 32.72	22.2
1	A	152	TYR	HD1	0.55	5.49 − 8.39	-22.0
1	A	152	TYR	HD2	0.55	5.48 − 8.39	-21.9
1	A	127	LEU	HD21	6.79	-0.65 − 2.13	21.8
1	A	127	LEU	HD22	6.79	-0.65 − 2.13	21.8
1	A	127	LEU	HD23	6.79	-0.65 − 2.13	21.8
1	A	192	ILE	HD11	6.75	-0.72 − 2.09	21.6

Continued on next page...

Continued from previous page...

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	192	ILE	HD12	6.75	-0.72 – 2.09	21.6
1	A	192	ILE	HD13	6.75	-0.72 – 2.09	21.6
1	A	151	THR	CB	33.35	61.12 – 78.27	-21.2
1	A	148	SER	CB	31.97	56.28 – 71.32	-21.2
1	A	185	SER	CB	32.02	56.28 – 71.32	-21.1
1	A	117	THR	CB	34.04	61.12 – 78.27	-20.8
1	A	130	THR	CB	34.72	61.12 – 78.27	-20.4
1	A	136	PHE	HD2	0.90	5.52 – 8.61	-19.9
1	A	136	PHE	HD1	0.90	5.51 – 8.60	-19.9
1	A	163	VAL	CB	63.74	23.86 – 41.50	17.6
1	A	145	HIS	CB	65.99	19.76 – 40.75	17.0
1	A	175	THR	CB	41.19	61.12 – 78.27	-16.6
1	A	138	SER	CB	39.43	56.28 – 71.32	-16.2
1	A	166	LEU	CB	71.06	33.11 – 51.34	15.8
1	A	140	ASN	CB	64.05	30.50 – 46.89	15.5
1	A	155	ASN	CB	64.01	30.50 – 46.89	15.4
1	A	126	GLY	CA	64.67	38.93 – 51.79	15.0
1	A	176	SER	CB	41.23	56.28 – 71.32	-15.0
1	A	173	TYR	CB	70.02	28.67 – 49.81	14.6
1	A	184	TRP	HD1	2.43	5.46 – 8.81	-14.1
1	A	114	ASP	CB	19.10	32.98 – 48.76	-13.8
1	A	147	GLY	CA	62.92	38.93 – 51.79	13.7
1	A	186	GLY	CA	61.92	38.93 – 51.79	12.9
1	A	159	GLY	CA	61.27	38.93 – 51.79	12.4
1	A	143	ASN	HD21	1.62	4.94 – 9.72	-11.9
1	A	143	ASN	HD22	1.62	4.69 – 9.61	-11.2
1	A	168	GLY	CA	59.77	38.93 – 51.79	11.2
1	A	128	PRO	CA	46.89	55.85 – 70.84	-11.0
1	A	153	ARG	HD2	0.80	1.97 – 4.26	-10.1
1	A	161	ARG	HD2	0.87	1.97 – 4.26	-9.8
1	A	184	TRP	HB3	-0.38	1.31 – 4.93	-9.7
1	A	187	PRO	HD2	0.48	1.93 – 5.38	-9.2
1	A	158	SER	HB2	1.56	2.61 – 5.13	-9.2
1	A	157	GLY	CA	56.89	38.93 – 51.79	9.0
1	A	153	ARG	HD3	0.80	1.81 – 4.39	-8.9
1	A	181	VAL	CB	48.06	23.86 – 41.50	8.7
1	A	161	ARG	HD3	0.87	1.81 – 4.39	-8.7
1	A	158	SER	HB3	1.56	2.49 – 5.20	-8.4
1	A	187	PRO	HD3	0.48	1.76 – 5.48	-8.4
1	A	142	GLU	HG3	0.49	1.20 – 3.30	-8.4
1	A	148	SER	HB2	1.75	2.61 – 5.13	-8.4
1	A	185	SER	HB3	1.57	2.49 – 5.20	-8.4

Continued from previous page...

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	187	PRO	CB	41.52	26.06 − 37.61	8.4
1	A	134	GLY	CA	55.83	38.93 − 51.79	8.1
1	A	142	GLU	HG2	0.64	1.24 − 3.30	-7.9
1	A	122	ARG	CD	50.39	38.57 − 47.75	7.9
1	A	148	SER	HB3	1.71	2.49 − 5.20	-7.9
1	A	175	THR	HB	1.70	2.57 − 5.77	-7.7
1	A	160	GLY	CA	55.23	38.93 − 51.79	7.7
1	A	130	THR	HB	1.78	2.57 − 5.77	-7.4
1	A	183	ILE	CD1	25.25	5.18 − 21.60	7.2
1	A	139	THR	CA	44.39	49.41 − 75.05	-7.0
1	A	170	PRO	CB	39.83	26.06 − 37.61	6.9
1	A	124	PRO	CA	53.08	55.85 − 70.84	-6.8
1	A	184	TRP	HB2	0.89	1.51 − 4.87	-6.8
1	A	132	THR	HB	2.01	2.57 − 5.77	-6.8
1	A	185	SER	HB2	2.18	2.61 − 5.13	-6.7
1	A	149	VAL	CA	43.62	48.38 − 76.73	-6.7
1	A	156	PRO	HB2	4.31	0.37 − 3.78	6.5
1	A	169	GLU	HB2	3.36	1.00 − 3.05	6.5
1	A	161	ARG	CB	42.21	21.74 − 39.52	6.5
1	A	142	GLU	CB	40.88	21.56 − 38.37	6.5
1	A	178	ASP	CB	30.65	32.98 − 48.76	-6.5
1	A	116	GLU	CB	19.41	21.56 − 38.37	-6.3
1	A	141	ARG	CG	34.72	21.24 − 33.19	6.3
1	A	129	PRO	CA	54.41	55.85 − 70.84	-6.0
1	A	162	LYS	HD2	0.40	0.58 − 2.64	-5.8
1	A	128	PRO	N	104.64	108.67 − 162.11	-5.8
1	A	162	LYS	HD3	0.40	0.54 − 2.65	-5.7
1	A	189	PRO	CA	55.11	55.85 − 70.84	-5.5
1	A	156	PRO	HB3	3.90	0.25 − 3.76	5.4
1	A	114	ASP	HB2	1.32	1.41 − 4.01	-5.3
1	A	176	SER	HB3	2.41	2.49 − 5.20	-5.3
1	A	188	ALA	CA	63.21	43.52 − 62.81	5.2
1	A	169	GLU	HB3	3.09	0.95 − 3.05	5.2
1	A	170	PRO	CA	55.80	55.85 − 70.84	-5.0
1	A	171	SER	CB	71.38	56.28 − 71.32	5.0

7.1.5   Random Coil Index (RCI) plots ⓘ

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: