W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

May 28, 2020 – 11:41 pm BST

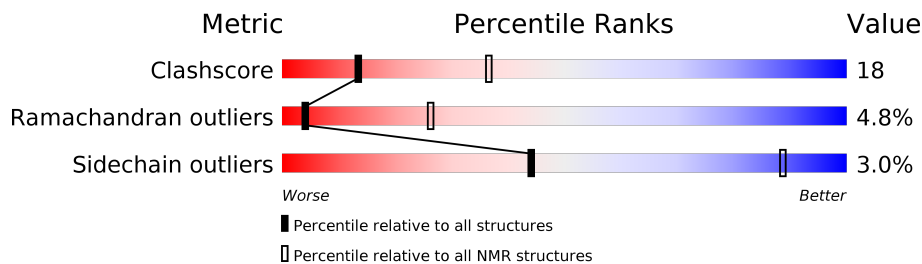|		
PDB ID	:	2MCY
Title	:	CR1 Sushi domains 2 and 3
Authors	:	Park, H.J.; Guariento, M.J.; Maciejewski, M.; Hauart, R.; Tham, W.; Cowman, A.F.; Schmidt, C.Q.; Martens, H.; Liszewski, K.M.; Hourcade, D.; Barlow, P.N.; Atkinson, J.P.
Deposited on	:	2013-08-27

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 29%.

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	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	136	64% 32% • •

2 Ensemble composition and analysis (i)

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)	0.98	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 (i)

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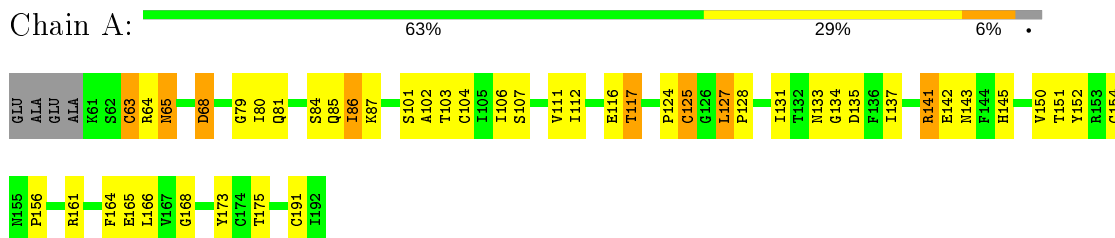
• Molecule 1: Complement receptor type 1
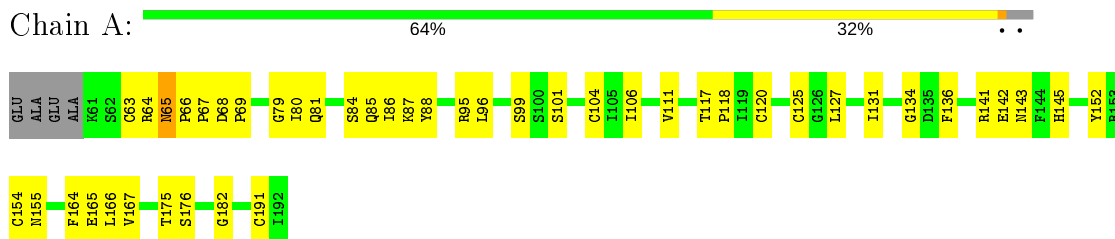
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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 (i)

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 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	486
Number of shifts mapped to atoms	486
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	29%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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 407 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

5.2 Torsion angles (i)

5.2.1 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%)	4	26
All	All	2600/2720 (96%)	2262 (87%)	212 (8%)	126 (5%)	4	26

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	176	SER	8
1	A	145	HIS	8
1	A	66	PRO	7

5.2.2 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	115/117 (98%)	112±2 (97±2%)	4±2 (3±2%)	44	89
All	All	2300/2340 (98%)	2230 (97%)	70 (3%)	44	89

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	175	THR	4

5.2.3   RNA ⓘ

There are no RNA molecules in this entry.

5.3   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.4   Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.5   Ligand geometry ⓘ

There are no ligands in this entry.

5.6   Other polymers ⓘ

There are no such molecules in this entry.

5.7   Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Chemical shift validation ⓘ

The completeness of assignment taking into account all chemical shift lists is 29% for the well-defined parts and 29% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

6.1.1 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	486
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	98

6.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	75	0.49 ± 0.32	None needed (< 0.5 ppm)
$^{13}C_\beta$	58	-2.30 ± 0.76	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	67	-1.06 ± 0.82	None needed (imprecise)

6.1.3 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 29%, i.e. 445 atoms were assigned a chemical shift out of a possible 1530. 1 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	255/638 (40%)	121/253 (48%)	75/264 (28%)	59/121 (49%)
Sidechain	177/778 (23%)	109/461 (24%)	68/283 (24%)	0/34 (0%)

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

	Total	**¹H**	**¹³C**	**¹⁵N**
Aromatic	13/114 (11%)	9/60 (15%)	4/50 (8%)	0/4 (0%)
Overall	445/1530 (29%)	239/774 (31%)	147/597 (25%)	59/159 (37%)

6.1.4   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.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	143	ASN	CG	25.14	183.77 − 169.77	-108.3
1	A	136	PHE	CD2	13.65	137.34 − 125.84	-102.6
1	A	136	PHE	CD1	13.65	137.63 − 125.43	-96.6
1	A	180	GLN	CD	50.89	186.62 − 172.72	-92.6
1	A	152	TYR	CD1	25.83	139.11 − 126.41	-84.2
1	A	178	ASP	CG	27.50	188.49 − 169.99	-82.0
1	A	152	TYR	CD2	25.83	140.11 − 125.31	-72.2
1	A	184	TRP	CG	25.21	118.98 − 101.98	-50.2
1	A	152	TYR	CG	27.97	141.74 − 116.84	-40.7
1	A	118	PRO	CB	65.42	37.79 − 25.89	28.2
1	A	188	ALA	CB	68.71	28.03 − 9.93	27.5
1	A	156	PRO	CB	64.01	37.79 − 25.89	27.0
1	A	158	SER	CB	24.48	71.24 − 56.34	-26.4
1	A	127	LEU	HD11	7.39	2.16 − -0.64	23.7
1	A	127	LEU	HD12	7.39	2.16 − -0.64	23.7
1	A	127	LEU	HD13	7.39	2.16 − -0.64	23.7
1	A	132	THR	CB	31.32	78.10 − 61.30	-22.8
1	A	129	PRO	CG	51.66	32.66 − 21.76	22.4
1	A	151	THR	CB	33.35	78.10 − 61.30	-21.6
1	A	127	LEU	HD22	6.79	2.14 − -0.66	21.6
1	A	127	LEU	HD21	6.79	2.14 − -0.66	21.6
1	A	127	LEU	HD23	6.79	2.14 − -0.66	21.6
1	A	148	SER	CB	31.97	71.24 − 56.34	-21.4
1	A	185	SER	CB	32.02	71.24 − 56.34	-21.3
1	A	152	TYR	HD2	0.55	8.44 − 5.44	-21.3
1	A	152	TYR	HD1	0.55	8.44 − 5.44	-21.3
1	A	117	THR	CB	34.04	78.10 − 61.30	-21.2
1	A	192	ILE	HD12	6.75	2.13 − -0.77	20.9
1	A	192	ILE	HD13	6.75	2.13 − -0.77	20.9
1	A	192	ILE	HD11	6.75	2.13 − -0.77	20.9
1	A	130	THR	CB	34.72	78.10 − 61.30	-20.8
1	A	136	PHE	HD2	0.90	8.56 − 5.56	-20.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	136	PHE	HD1	0.90	8.56 − 5.56	-20.5
1	A	163	VAL	CB	63.74	41.76 − 23.66	17.1
1	A	145	HIS	CB	65.99	40.69 − 19.69	17.0
1	A	175	THR	CB	41.19	78.10 − 61.30	-17.0
1	A	138	SER	CB	39.43	71.24 − 56.34	-16.3
1	A	166	LEU	CB	71.06	51.69 − 32.89	15.3
1	A	176	SER	CB	41.23	71.24 − 56.34	-15.1
1	A	140	ASN	CB	64.05	47.13 − 30.23	15.0
1	A	155	ASN	CB	64.01	47.13 − 30.23	15.0
1	A	126	GLY	CA	64.67	51.81 − 38.91	15.0
1	A	173	TYR	CB	70.02	50.05 − 28.55	14.3
1	A	147	GLY	CA	62.92	51.81 − 38.91	13.6
1	A	114	ASP	CB	19.10	49.06 − 32.66	-13.3
1	A	184	TRP	HD1	2.43	8.95 − 5.35	-13.1
1	A	186	GLY	CA	61.92	51.81 − 38.91	12.8
1	A	159	GLY	CA	61.27	51.81 − 38.91	12.3
1	A	143	ASN	HD21	1.62	9.74 − 4.94	-11.9
1	A	143	ASN	HD22	1.62	9.59 − 4.69	-11.3
1	A	168	GLY	CA	59.77	51.81 − 38.91	11.2
1	A	128	PRO	CA	46.89	71.13 − 55.53	-10.5
1	A	153	ARG	HD2	0.80	4.27 − 1.97	-10.1
1	A	161	ARG	HD2	0.87	4.27 − 1.97	-9.8
1	A	184	TRP	HB3	-0.38	4.94 − 1.34	-9.8
1	A	153	ARG	HD3	0.80	4.36 − 1.86	-9.3
1	A	161	ARG	HD3	0.87	4.36 − 1.86	-9.0
1	A	157	GLY	CA	56.89	51.81 − 38.91	8.9
1	A	158	SER	HB2	1.56	5.18 − 2.58	-8.9
1	A	187	PRO	HD2	0.48	5.45 − 1.85	-8.8
1	A	181	VAL	CB	48.06	41.76 − 23.66	8.5
1	A	142	GLU	HG3	0.49	3.31 − 1.21	-8.5
1	A	187	PRO	HD3	0.48	5.52 − 1.72	-8.3
1	A	122	ARG	CD	50.39	47.57 − 38.77	8.2
1	A	158	SER	HB3	1.56	5.25 − 2.45	-8.2
1	A	148	SER	HB2	1.75	5.18 − 2.58	-8.2
1	A	185	SER	HB3	1.57	5.25 − 2.45	-8.1
1	A	187	PRO	CB	41.52	37.79 − 25.89	8.1
1	A	134	GLY	CA	55.83	51.81 − 38.91	8.1
1	A	142	GLU	HG2	0.64	3.33 − 1.23	-7.8
1	A	160	GLY	CA	55.23	51.81 − 38.91	7.7
1	A	148	SER	HB3	1.71	5.25 − 2.45	-7.6
1	A	175	THR	HB	1.70	5.82 − 2.52	-7.5
1	A	130	THR	HB	1.78	5.82 − 2.52	-7.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	183	ILE	CD1	25.25	21.91 − 5.01	7.0
1	A	139	THR	CA	44.39	75.37 − 49.07	-6.8
1	A	170	PRO	CB	39.83	37.79 − 25.89	6.7
1	A	184	TRP	HB2	0.89	4.94 − 1.44	-6.6
1	A	124	PRO	CA	53.08	71.13 − 55.53	-6.6
1	A	185	SER	HB2	2.18	5.18 − 2.58	-6.6
1	A	132	THR	HB	2.01	5.82 − 2.52	-6.6
1	A	149	VAL	CA	43.62	76.93 − 48.03	-6.5
1	A	156	PRO	HB2	4.31	3.82 − 0.32	6.4
1	A	169	GLU	HB2	3.36	3.08 − 0.98	6.3
1	A	161	ARG	CB	42.21	39.81 − 21.51	6.3
1	A	142	GLU	CB	40.88	38.65 − 21.35	6.3
1	A	141	ARG	CG	34.72	33.23 − 21.23	6.2
1	A	178	ASP	CB	30.65	49.06 − 32.66	-6.2
1	A	116	GLU	CB	19.41	38.65 − 21.35	-6.1
1	A	129	PRO	CA	54.41	71.13 − 55.53	-5.7
1	A	189	PRO	CA	55.11	71.13 − 55.53	-5.3
1	A	156	PRO	HB3	3.90	3.81 − 0.21	5.3
1	A	162	LYS	HD2	0.40	2.76 − 0.46	-5.2
1	A	162	LYS	HD3	0.40	2.75 − 0.45	-5.2
1	A	114	ASP	HB2	1.32	4.07 − 1.37	-5.2
1	A	176	SER	HB3	2.41	5.25 − 2.45	-5.1
1	A	171	SER	CB	71.38	71.24 − 56.34	5.1
1	A	188	ALA	CA	63.21	63.07 − 43.27	5.1

6.1.5   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: