Full wwPDB NMR Structure Validation Report (i)

Apr 26, 2016 – 02:01 PM BST

PDB ID	:	1CUR
Title	:	REDUCED RUSTICYANIN, NMR
Authors	:	Botuyan, M.V.; Dyson, H.J.
Deposited on	:	1996-04-19

This is a Full wwPDB NMR Structure Validation 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 (i) symbol.

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

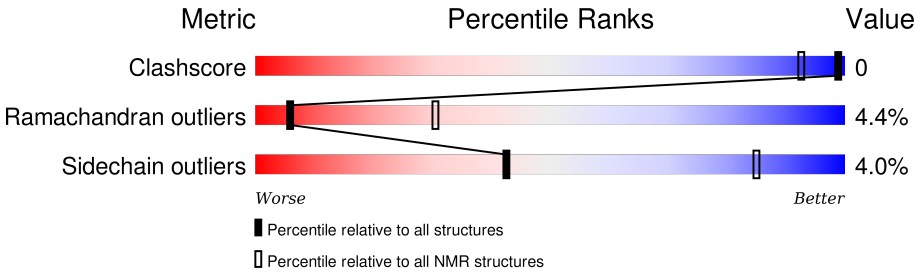
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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 was not calculated.

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	155	88% 7% • 5%

2 Ensemble composition and analysis

This entry contains 15 models. Model 12 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:5-A:47, A:51-A:155 (148)	0.48	12

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 4 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2345 atoms, of which 1174 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CU(I) RUSTICYANIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	155	2344	764	1174	186	216	4	0

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms	
			Total	Cu
2	A	1	1	1

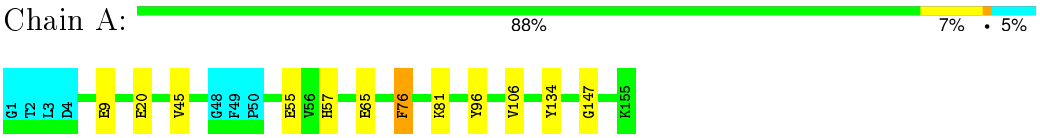
4 Residue-property plots (i)

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: CU(I) RUSTICYANIN

Chain A:



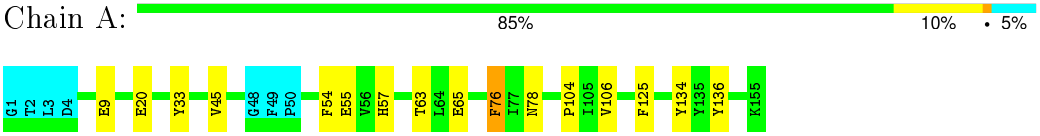
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: CU(I) RUSTICYANIN

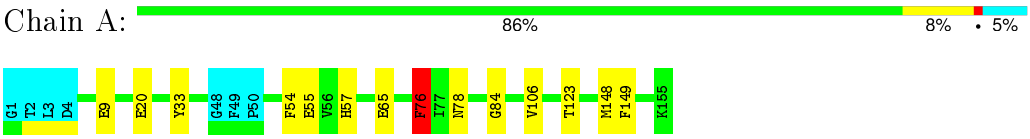
Chain A:



4.2.2 Score per residue for model 2

- Molecule 1: CU(I) RUSTICYANIN

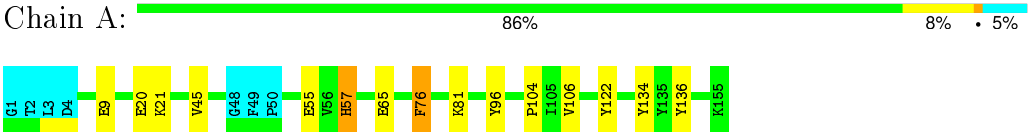
Chain A:



4.2.3 Score per residue for model 3

• Molecule 1: CU(I) RUSTICYANIN

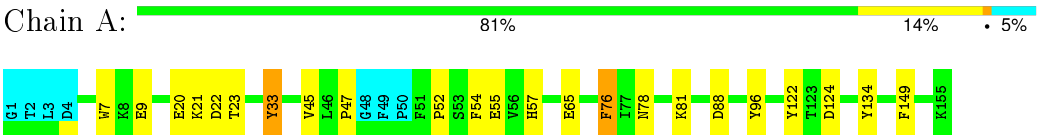
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: CU(I) RUSTICYANIN

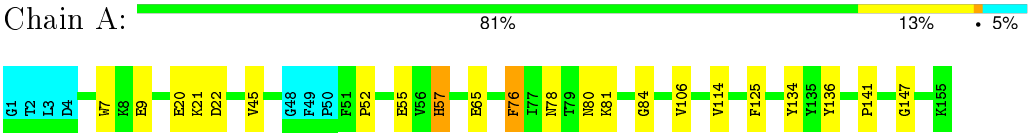
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: CU(I) RUSTICYANIN

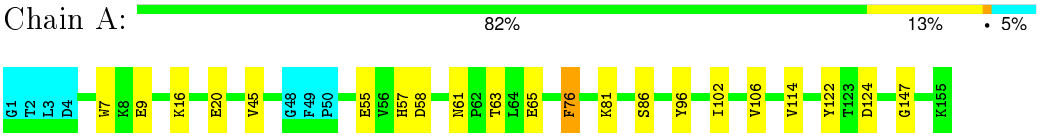
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: CU(I) RUSTICYANIN

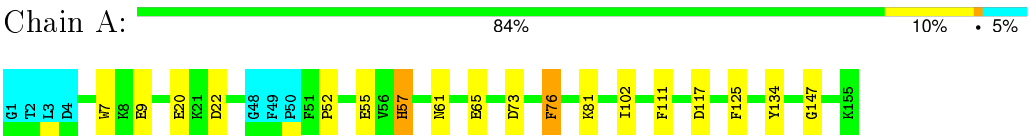
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: CU(I) RUSTICYANIN

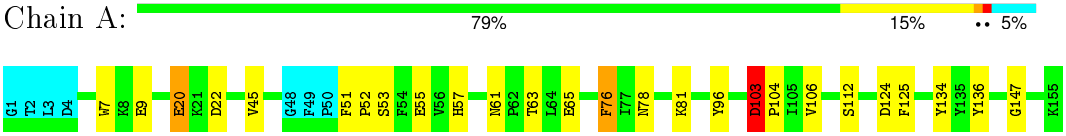
Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: CU(I) RUSTICYANIN

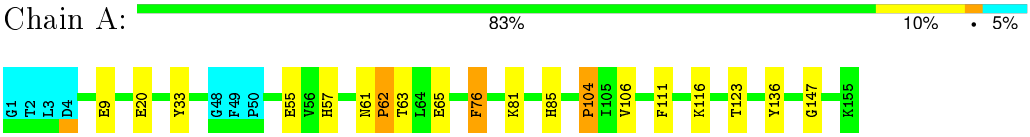
Chain A:



4.2.9 Score per residue for model 9

• Molecule 1: CU(I) RUSTICYANIN

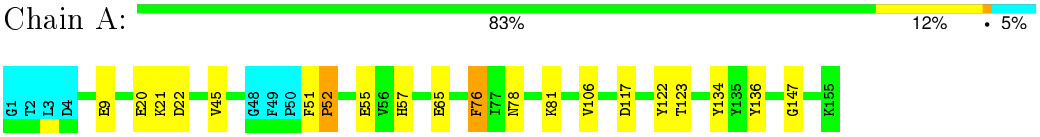
Chain A:



4.2.10 Score per residue for model 10

• Molecule 1: CU(I) RUSTICYANIN

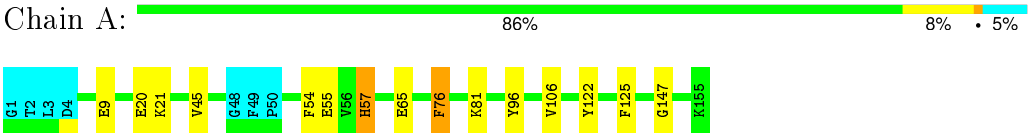
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: CU(I) RUSTICYANIN

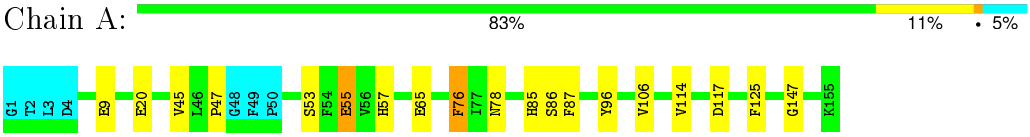
Chain A:



4.2.12 Score per residue for model 12 (medoid)

• Molecule 1: CU(I) RUSTICYANIN

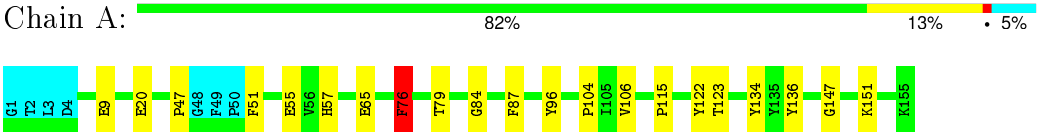
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: CU(I) RUSTICYANIN

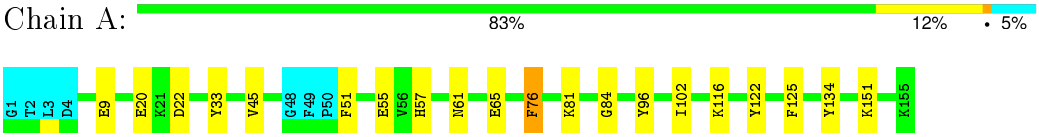
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: CU(I) RUSTICYANIN

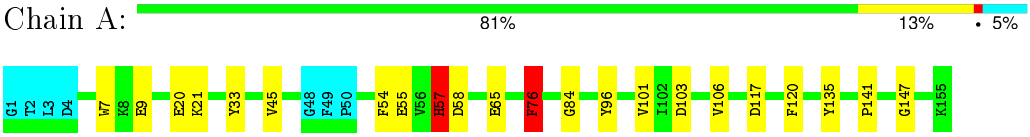
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: CU(I) RUSTICYANIN

Chain A:



5 Refinement protocol and experimental data overview

Of the ? calculated structures, 15 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
AMBER	refinement	

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality

6.1 Standard geometry

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

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	0.93±0.00	4±0/1157 (0.4±0.0%)	0.98±0.01	2±1/1581 (0.1±0.0%)
All	All	0.93	61/17355 (0.4%)	0.98	27/23715 (0.1%)

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	4.5±1.1
All	All	0	68

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	65	GLU	CD-OE2	10.20	1.36	1.25	7	15
1	A	55	GLU	CD-OE2	10.14	1.36	1.25	9	15
1	A	20	GLU	CD-OE2	9.99	1.36	1.25	1	15
1	A	9	GLU	CD-OE2	9.82	1.36	1.25	14	15
1	A	88	ASP	CG-OD2	5.01	1.36	1.25	4	1

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
1	A	76	PHE	CB-CG-CD2	-8.92	114.56	120.80	10	14
1	A	76	PHE	CB-CG-CD1	-7.32	115.68	120.80	11	11
1	A	148	MET	CG-SD-CE	-5.49	91.42	100.20	2	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	103	ASP	CB-CG-OD2	-5.26	113.57	118.30	8	1

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	76	PHE	Sidechain	15
1	A	96	TYR	Sidechain	9
1	A	134	TYR	Sidechain	9
1	A	136	TYR	Sidechain	7
1	A	125	PHE	Sidechain	7
1	A	33	TYR	Sidechain	6
1	A	54	PHE	Sidechain	5
1	A	122	TYR	Sidechain	5
1	A	120	PHE	Sidechain	1
1	A	51	PHE	Sidechain	1
1	A	87	PHE	Sidechain	1
1	A	135	TYR	Sidechain	1
1	A	111	PHE	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	1121	1129	1114	0±1
All	All	16830	16935	16710	3

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:76:PHE:CZ	1:A:87:PHE:HB2	0.46	2.45	13	1
1:A:57:HIS:CG	1:A:58:ASP:N	0.42	2.88	15	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:79:THR:HG23	1:A:115:PRO:HG2	0.40	1.93	13	1

6.3 Torsion angles

6.3.1 Protein backbone

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	147/155 (95%)	119±3 (81±2%)	21±3 (14±2%)	7±2 (4±1%)	6	30
All	All	2205/2325 (95%)	1789 (81%)	318 (14%)	98 (4%)	6	30

All 24 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	57	HIS	14
1	A	106	VAL	12
1	A	45	VAL	11
1	A	147	GLY	10
1	A	81	LYS	10
1	A	84	GLY	5
1	A	104	PRO	5
1	A	52	PRO	4
1	A	117	ASP	4
1	A	51	PHE	3
1	A	47	PRO	3
1	A	103	ASP	2
1	A	63	THR	2
1	A	141	PRO	2
1	A	85	HIS	2
1	A	62	PRO	1
1	A	111	PHE	1
1	A	102	ILE	1
1	A	114	VAL	1
1	A	86	SER	1
1	A	101	VAL	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	61	ASN	1
1	A	58	ASP	1
1	A	22	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	121/126 (96%)	116±2 (96±2%)	5±2 (4±2%)	42	85
All	All	1815/1890 (96%)	1742 (96%)	73 (4%)	42	85

All 30 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	78	ASN	7
1	A	57	HIS	6
1	A	21	LYS	6
1	A	7	TRP	6
1	A	22	ASP	5
1	A	61	ASN	4
1	A	123	THR	4
1	A	124	ASP	3
1	A	122	TYR	2
1	A	63	THR	2
1	A	53	SER	2
1	A	102	ILE	2
1	A	116	LYS	2
1	A	114	VAL	2
1	A	52	PRO	2
1	A	151	LYS	2
1	A	76	PHE	2
1	A	149	PHE	2
1	A	62	PRO	1
1	A	55	GLU	1
1	A	103	ASP	1
1	A	112	SER	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	20	GLU	1
1	A	23	THR	1
1	A	73	ASP	1
1	A	86	SER	1
1	A	80	ASN	1
1	A	16	LYS	1
1	A	33	TYR	1
1	A	104	PRO	1

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 carbohydrates in this entry.

6.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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

No chemical shift data were provided