wwPDB NMR Structure Validation Summary Report (i)

Mar 6, 2022 – 06:27 PM EST

PDB ID	:	2P80
Title	:	Solution structure of the complex between nitrite reductase and pseudoazurin from A. faecalis
Authors	:	Vlasie, M.D.; Ubbink, M.
Deposited on	:	2007-03-21

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 (i)) were used in the production of this report:

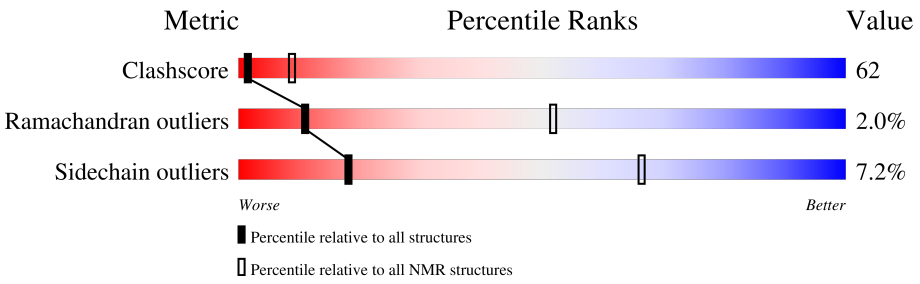
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 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	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	341	50% 43% 6% •
1	B	341	51% 41% 6% ••
1	C	341	47% 46% 5% •
2	D	123	62% 29% 9%

2 Ensemble composition and analysis (i)

This entry contains 20 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: *closest to the average*.

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:339, B:4-B:339, C:4-C:339 (1007)	-0.00	2
2	D:1-D:123 (123)	0.00	10

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

3 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 17072 atoms, of which 8444 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Copper-containing nitrite reductase.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	335	5045	1636	2490	429	479	11	0
1	B	336	5055	1639	2495	430	480	11	0
1	C	336	5055	1639	2495	430	480	11	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	341	ILE	-	expression tag	UNP P38501
A	342	GLU	-	expression tag	UNP P38501
A	343	GLY	-	expression tag	UNP P38501
A	344	ARG	-	expression tag	UNP P38501
B	341	ILE	-	expression tag	UNP P38501
B	342	GLU	-	expression tag	UNP P38501
B	343	GLY	-	expression tag	UNP P38501
B	344	ARG	-	expression tag	UNP P38501
C	341	ILE	-	expression tag	UNP P38501
C	342	GLU	-	expression tag	UNP P38501
C	343	GLY	-	expression tag	UNP P38501
C	344	ARG	-	expression tag	UNP P38501

- Molecule 2 is a protein called Pseudoazurin.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	D	123	1901	598	964	155	178	6	0

- Molecule 3 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms	
			Total	Cu
3	A	2	2	2
3	B	2	2	2

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	
			Total	Cu
3	C	2	2	2
			Total	Cu
3	D	1	1	1

- Molecule 4 is GADOLINIUM ATOM (three-letter code: GD) (formula: Gd).

Mol	Chain	Residues	Atoms	
			Total	Gd
4	A	3	3	3
			Total	Gd
4	B	3	3	3
			Total	Gd
4	C	3	3	3

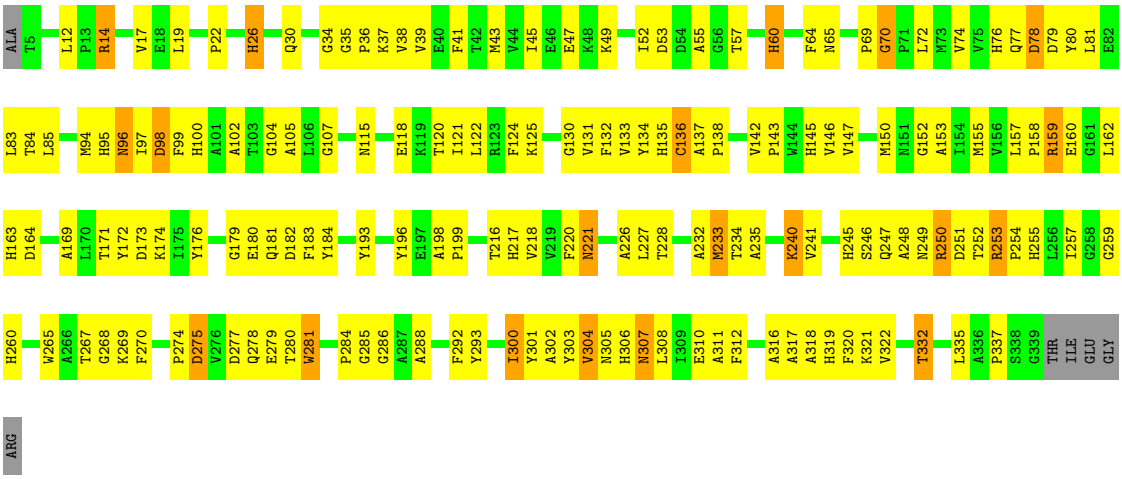
4 Residue-property plots (i)

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.

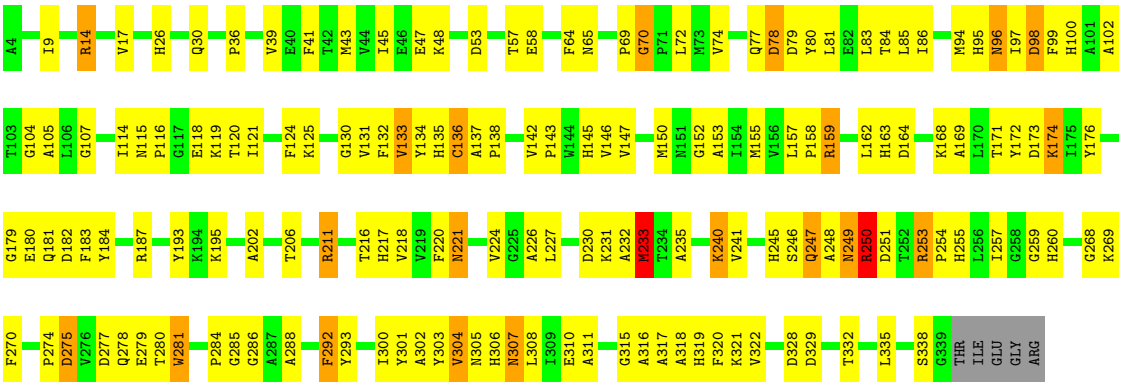
• Molecule 1: Copper-containing nitrite reductase

Chain A: 50% 43% 6% •



• Molecule 1: Copper-containing nitrite reductase

Chain B: 51% 41% 6% ••



• Molecule 1: Copper-containing nitrite reductase

Chain C: 47% 46% 5% •

A4 R14 V17 V25 H26 A27 Q30 P36 K37 V38 V39 E40 F41 T42 M43 Y44 I45 E46 E47 K48 K49 D53 D54 A55 G56 T57 E58 V59 H60 F64 N65 P69 G70 P71 L72 M73 V74 V75 H76 Q77 D78 D79 Y80 L81 E82 L83 T84 L85 I86 E89

T92 L93 M94 H95 N96 I97 D98 F99 H100 A101 A102 T103 G104 A105 L106 G107 I114 N115 P116 K119 T120 I121 F124 K125 G130 V131 F132 V133 Y134 H135 C136 A137 P138 M141 V142 P143 W144 H145 V146 V147 M150 N151 G152 A153 V156 L157 P158 R159 E160 G161 L162 H163 D164

G165 K166 A169 L170 T171 Y172 D173 K174 I175 Y176 G179 E180 Q181 D182 F183 Y184 R187 Y193 Y196 E197 A198 P199 Y203 E204 R211 T216 H217 V218 V219 F220 N221 A226 L227 D230 M233 T234 A235 K240 V241 H245 S246 Q247 A248 N249 R250 D251 T252

R253 P254 H255 G259 H260 G261 D262 Y263 V264 W265 G268 K269 F270 P274 D275 V276 D277 Q278 E279 T280 W281 P284 G285 G286 A287 A288 A291 F292 Y293 Y301 A302 Y303 V304 N305 H306 N307 L308 I309 E310 A311 F312 A316 A317 A318 H319 F320 K321 V322 D328 D329 L330 M331

T332 L335 A336 P337 S338 G339 THR ILE GLU GLY ARG

• Molecule 2: Pseudoazurin

Chain D: 62% 29% 9%

E1 E4 V5 H6 M7 L8 N9 K10 G11 G14 A15 M16 V17 F18 Y22 I23 K24 P35 K38 G39 H40 S44 I45 K46 P50 K55 F56 K59 I60 N61 E62 N63 Q70 A73 Y74 C78 T79 P80 H81 Y82 A83 M84 G85 M86 I87 D94

Q101 K109 I110 L115 E116 K117 V118 S121 A122 K123

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: Copper-containing nitrite reductase

Chain A: 50% 43% 6% •

ALA T5 L12 P13 R14 V17 E18 L19 P22 H26 Q30 G34 G35 P36 K37 V38 V39 E40 F41 T42 M43 V44 I45 E46 E47 K48 K49 I52 D53 D54 A55 G56 T57 H60 F64 N65 P69 G70 P71 L72 M73 V74 V75 H76 Q77 D78 D79 Y80 L81 E82

L83 T84 L85 M94 H95 N96 I97 D98 F99 H100 A101 A102 T103 G104 A105 L106 G107 N115 E118 K119 T120 I121 L122 R123 F124 K125 G130 V131 F132 V133 Y134 H135 C136 A137 P138 V142 P143 W144 H145 V146 V147 M150 N151 G152 A153 I154 M155 V156 L157 P158 R159 E160 G161 L162

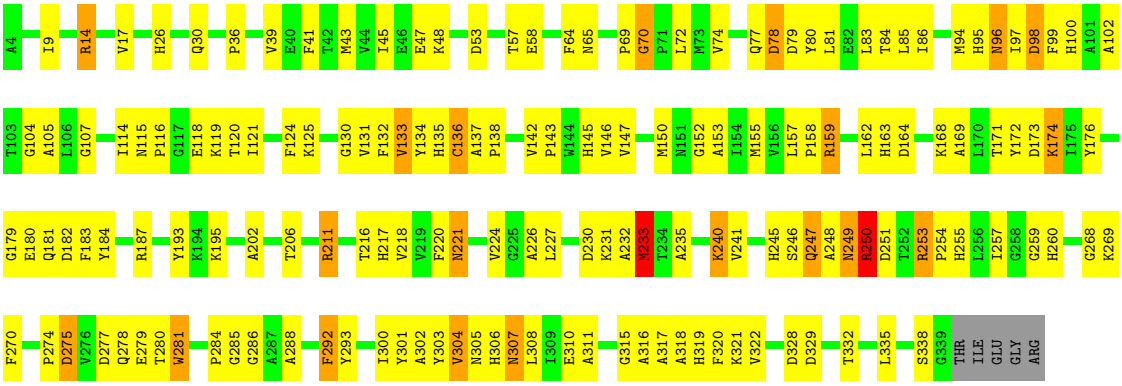
H163 D164 A169 L170 T171 Y172 D173 K174 I175 Y176 G179 E180 Q181 D182 F183 Y184 Y193 Y196 E197 A198 P199 T216 H217 V218 V219 F220 N221 A226 L227 T228 A232 M233 T234 A235 K240 V241 H245 S246 Q247 A248 N249 R250 D251 T252 R253 P254 H255 L256 I257 G258 G259

H260 W265 A266 T267 G268 K269 F270 P274 D275 V276 D277 Q278 E279 T280 W281 P284 G285 G286 A287 A288 F292 Y293 I300 Y301 A302 Y303 V304 N305 H306 N307 L308 I309 E310 A311 F312 A316 A317 A318 H319 F320 K321 V322 T332 L335 A336 P337 S338 G339 THR ILE GLU GLY

ARG

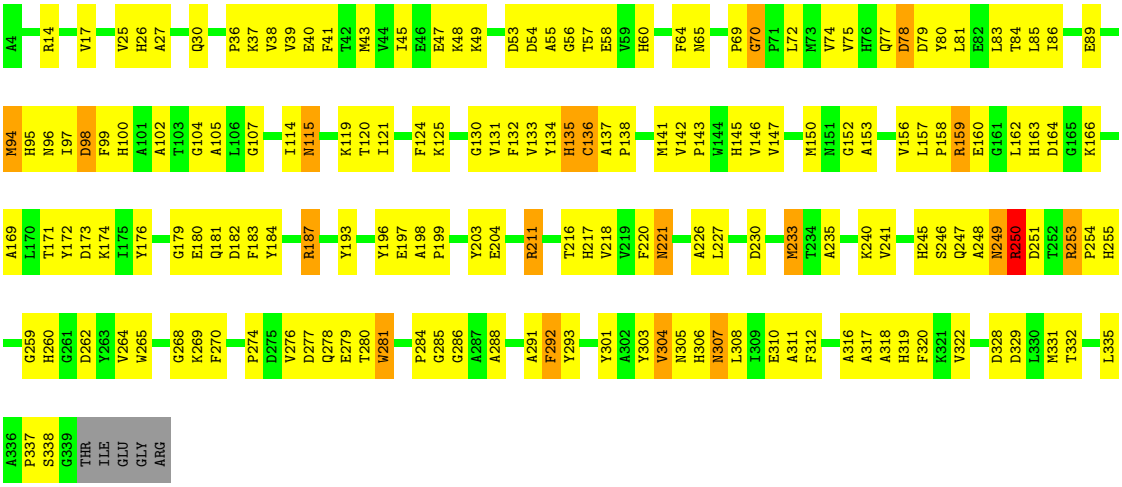
- Molecule 1: Copper-containing nitrite reductase

Chain B:



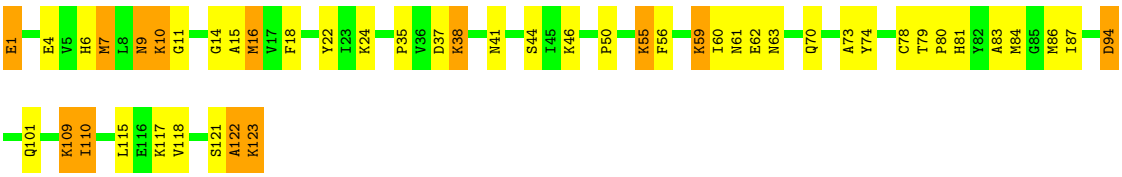
- Molecule 1: Copper-containing nitrite reductase

Chain C:



- Molecule 2: Pseudoazurin

Chain D:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *distance geometry, rigid body molecular dynamics, restrained side-chains energy minimization.*

Of the 134 calculated structures, 20 were deposited, based on the following criterion: *structures closest to the average.*

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

Software name	Classification	Version
X-PLOR	structure solution	2.9.9
X-PLOR	refinement	2.9.9

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

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	2555	2490	2480	345±2
1	B	2560	2495	2485	350±0
1	C	2560	2495	2485	396±30
2	D	937	964	963	105±28
All	All	172560	168880	168259	21257

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 62.

5 of 1631 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:C:94:MET:CB	2:D:83:ALA:HB1	1.43	1.43	7	9
1:C:203:TYR:CG	2:D:15:ALA:HB2	1.36	1.54	18	7
1:C:141:MET:SD	2:D:84:MET:SD	1.31	2.28	8	12
1:A:255:HIS:CE1	1:C:100:HIS:CE1	1.29	2.21	19	20
1:C:92:THR:O	2:D:110:ILE:CG2	1.27	1.83	17	2

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	333/341 (98%)	300±0 (90±0%)	25±0 (8±0%)	8±0 (2±0%)	9 46
1	B	334/341 (98%)	302±0 (90±0%)	24±0 (7±0%)	8±0 (2±0%)	9 46
1	C	334/341 (98%)	301±0 (90±0%)	27±0 (8±0%)	6±0 (2±0%)	12 54
2	D	121/123 (98%)	114±0 (94±0%)	6±0 (5±0%)	1±0 (1±0%)	24 71
All	All	22440/22920 (98%)	20340 (91%)	1640 (7%)	460 (2%)	11 52

5 of 23 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	26	HIS	20
1	A	60	HIS	20
1	A	70	GLY	20
1	A	78	ASP	20
1	A	96	ASN	20

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	264/268 (99%)	251±0 (95±0%)	13±0 (5±0%)	29 78
1	B	264/268 (99%)	246±0 (93±0%)	18±0 (7±0%)	19 68
1	C	264/268 (99%)	245±1 (93±0%)	19±1 (7±0%)	18 66
2	D	102/102 (100%)	88±1 (86±1%)	14±1 (14±1%)	7 47
All	All	17880/18120 (99%)	16597 (93%)	1283 (7%)	18 66

5 of 72 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	14	ARG	20
1	A	98	ASP	20
1	A	122	LEU	20
1	A	136	CYS	20
1	A	159	ARG	20

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)

Of 16 ligands modelled in this entry, 16 are monoatomic - leaving 0 for Mogul analysis.

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)

No chemical shift data were provided