WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report ⓘ

Feb 13, 2017 – 02:04 am GMT

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
http://wwpdb.org/validation/2016/NMRValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

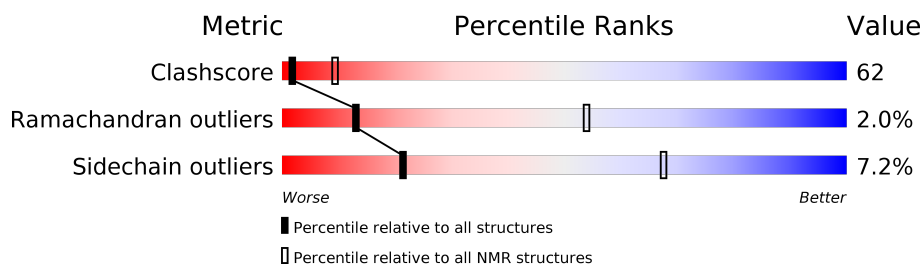|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk28760
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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	Percentile Ranks	Value
Clashscore		62
Ramachandran outliers		2.0%
Sidechain outliers		7.2%

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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 ⓘ

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 ⓘ

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

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
2	D	123	Total	C	H	N	O	S	0
			1901	598	964	155	178	6	

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

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

Continued on next page...

Continued from previous page...

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

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

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

4 Residue-property plots ⓘ

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.
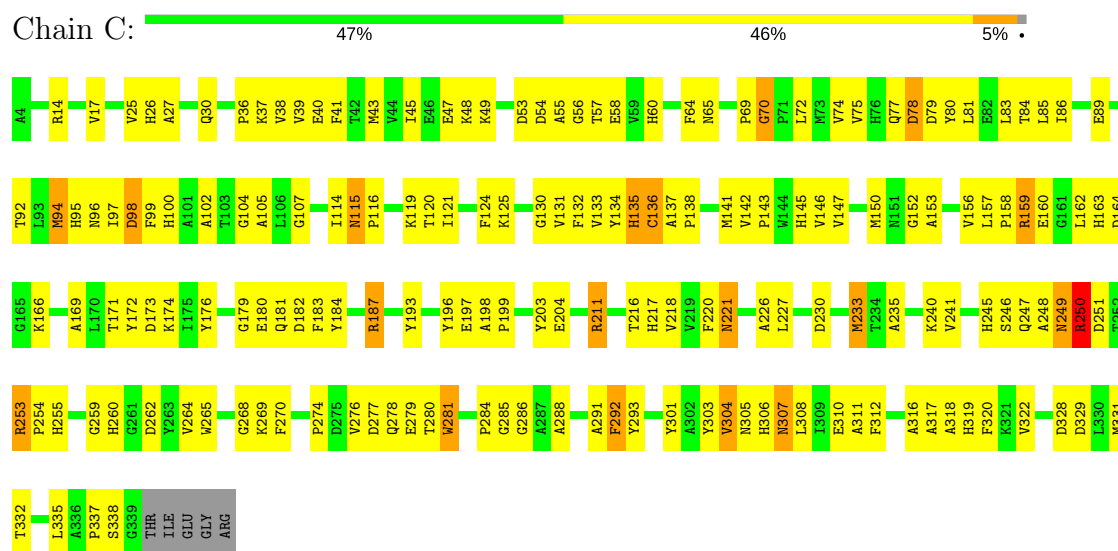
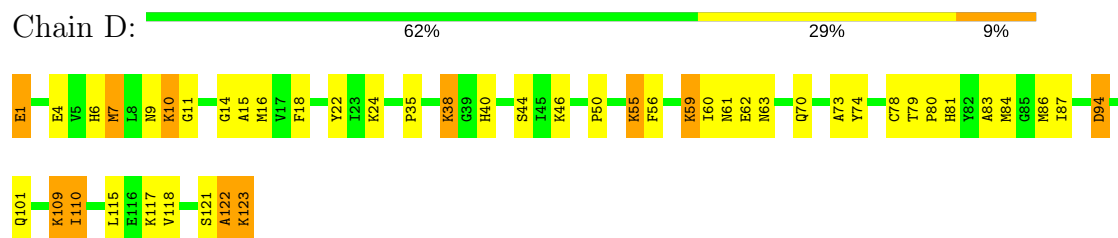• Molecule 1: Copper-containing nitrite reductase
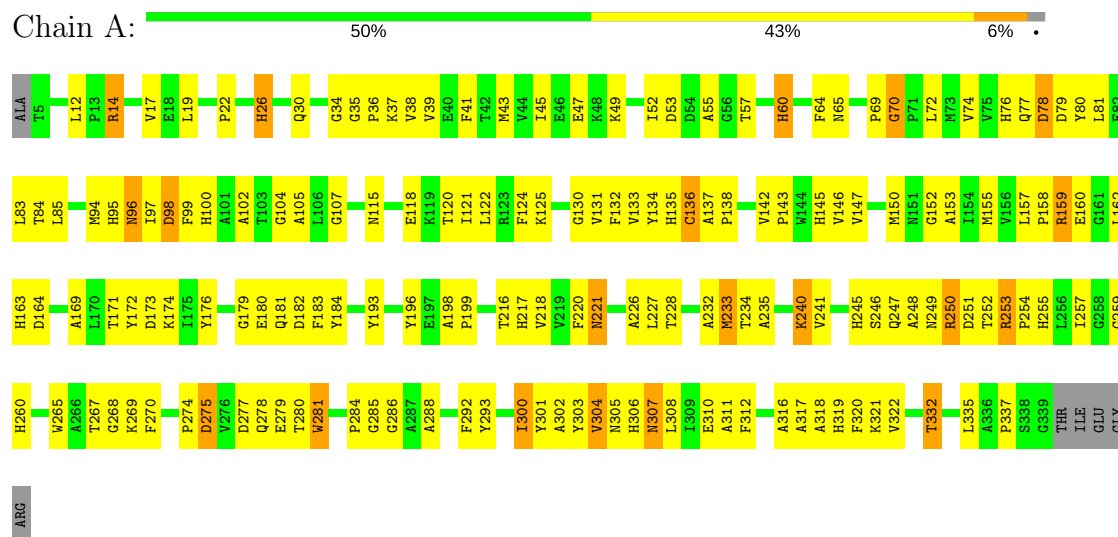
Chain A:

• Molecule 1: Copper-containing nitrite reductase

Chain B:

• Molecule 1: Copper-containing nitrite reductase

Chain C:
47%　　　　46%　　　　5% •

• Molecule 2: Pseudoazurin

Chain D:
62%　　　　29%　　　　9%
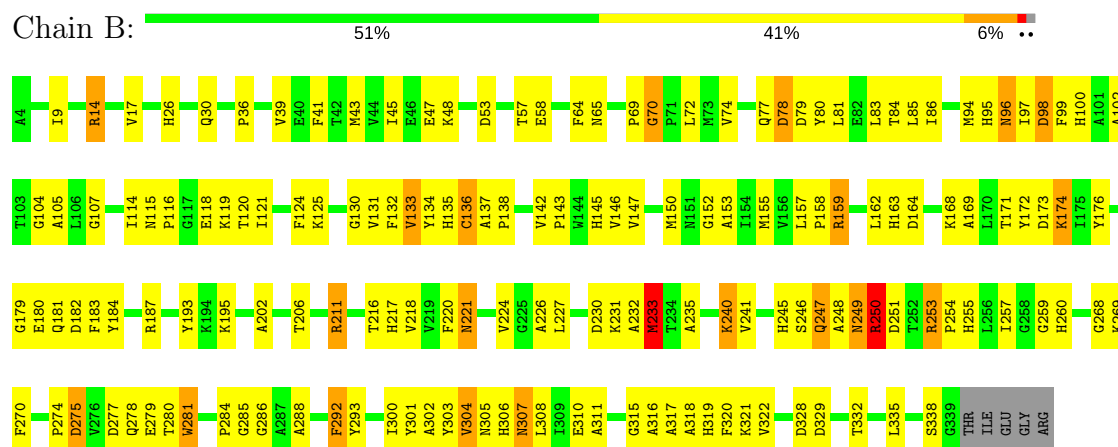
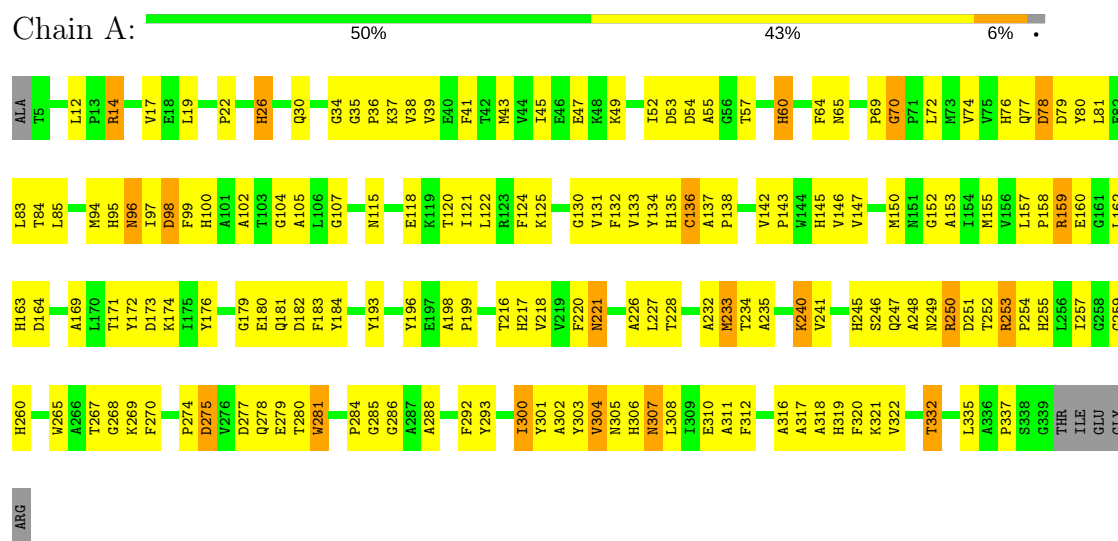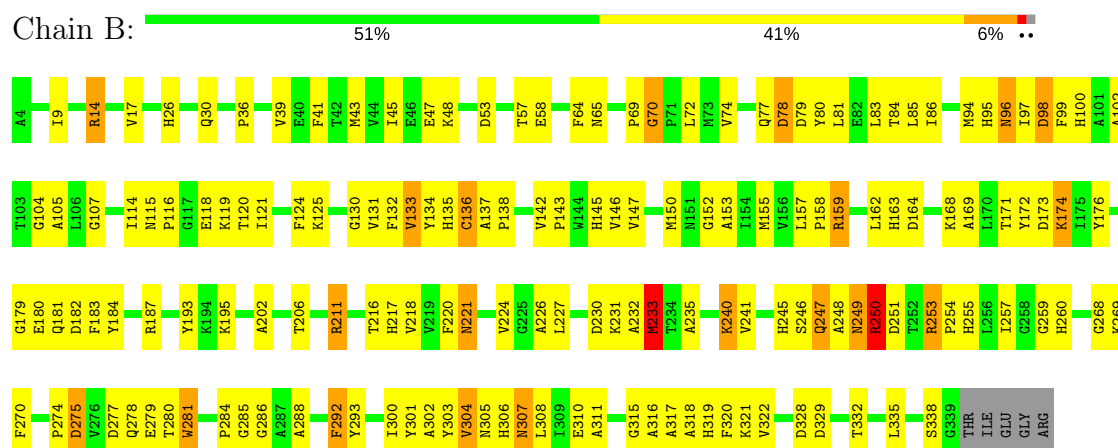## 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

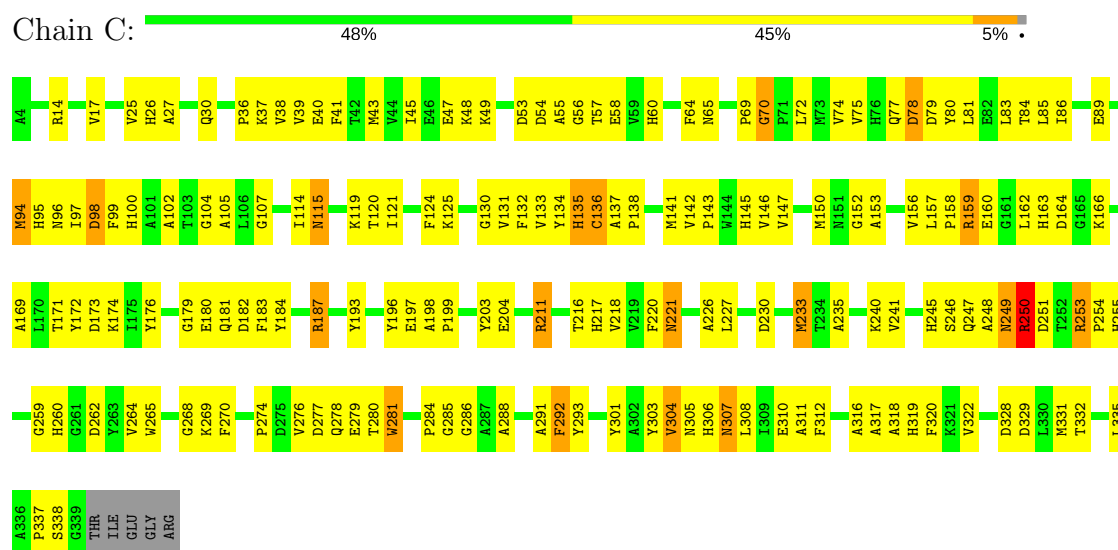The representative model is number 2. Colouring as in section 4.1 above.

• Molecule 1: Copper-containing nitrite reductase

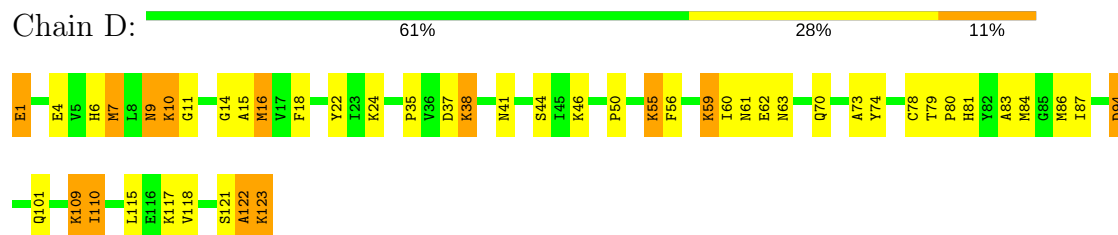Chain A:
50%　　　　43%　　　　6% •

• Molecule 1: Copper-containing nitrite reductase

Chain B:

51%　　　　41%　　　　6%

• Molecule 1: Copper-containing nitrite reductase

Chain C:

48%　　　　45%　　　　5%

• Molecule 2: Pseudoazurin

Chain D:

61%　　　　28%　　　　11%

5 Refinement protocol and experimental data overview ⓘ

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

WORLDWIDE
www**PDB**
PROTEIN DATA BANK

6 Model quality ⓘ

6.1 Standard geometry ⓘ

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 ⓘ

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 1737 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%)	11	48
1	B	334/341 (98%)	302±0 (90±0%)	24±0 (7±0%)	8±0 (2±0%)	11	48
1	C	334/341 (98%)	301±0 (90±0%)	27±0 (8±0%)	6±0 (2±0%)	14	57
2	D	121/123 (98%)	114±0 (94±0%)	6±0 (5±0%)	1±0 (1±0%)	27	73
All	All	22440/22920 (98%)	20340 (91%)	1640 (7%)	460 (2%)	13	54

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	C	250	ARG	20
1	B	233	MET	20
1	A	60	HIS	20
1	C	164	ASP	20
1	A	233	MET	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%)	33	79
1	B	264/268 (99%)	246±0 (93±0%)	18±0 (7±0%)	23	69
1	C	264/268 (99%)	245±1 (93±0%)	19±1 (7±0%)	21	67
2	D	102/102 (100%)	88±1 (86±1%)	14±1 (14±1%)	8	48
All	All	17880/18120 (99%)	16597 (93%)	1283 (7%)	21	67

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	136	CYS	20
1	B	253	ARG	20
1	B	133	VAL	20
1	B	233	MET	20
1	C	250	ARG	20

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 16 ligands modelled in this entry, 16 are 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 (i)

No chemical shift data were provided