Full wwPDB NMR Structure Validation Report (i)

Jun 11, 2024 – 11:38 PM EDT

PDB ID	:	2K8D
Title	:	Solution structure of a zinc-binding methionine sulfoxide reductase
Authors	:	Carella, M.; Ohlenschlager, O.; Gorlach, M.
Deposited on	:	2008-09-05

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
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

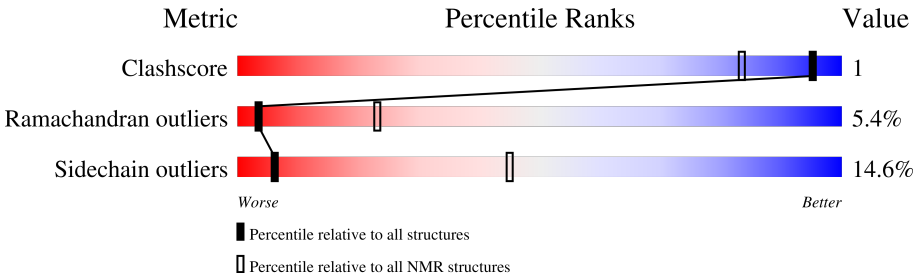
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	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	151	74% 14% • 11%

2 Ensemble composition and analysis

This entry contains 20 models. Model 6 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:7-A:47, A:62-A:154 (134)	1.06	6

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 3 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2367 atoms, of which 1154 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Peptide methionine sulfoxide reductase msrB.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	151	2366	756	1154	217	228	11	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	1	1	1

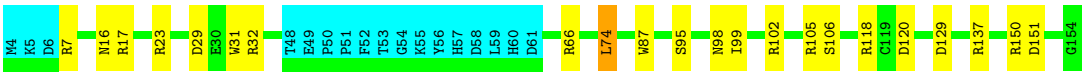
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, 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: Peptide methionine sulfoxide reductase msrB

Chain A:



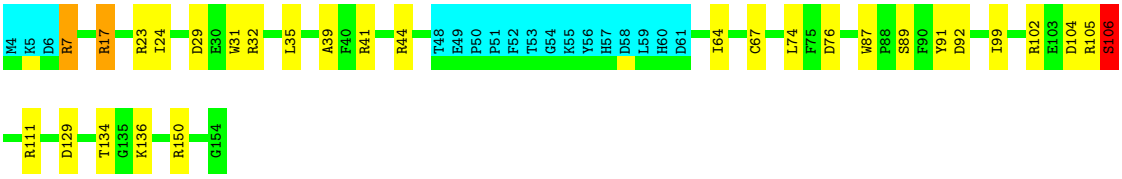
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Peptide methionine sulfoxide reductase msrB

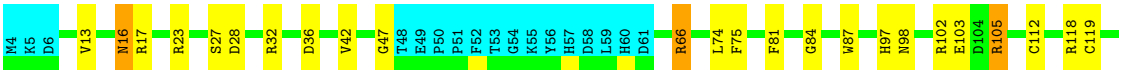
Chain A:

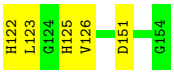


4.2.2 Score per residue for model 2

• Molecule 1: Peptide methionine sulfoxide reductase msrB

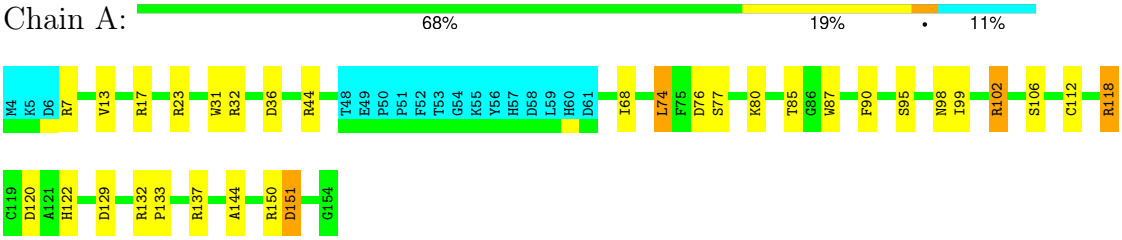
Chain A:





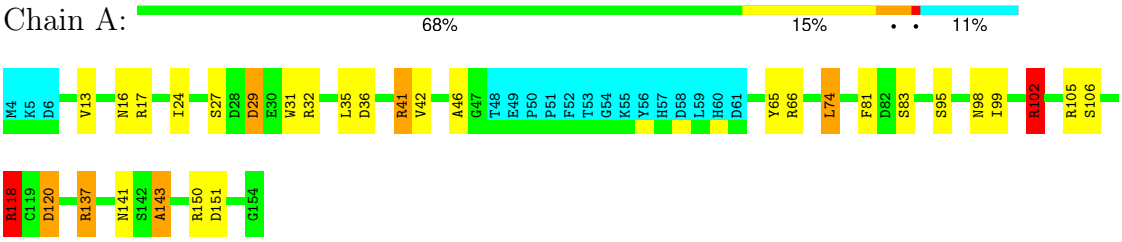
4.2.3 Score per residue for model 3

• Molecule 1: Peptide methionine sulfoxide reductase msrB



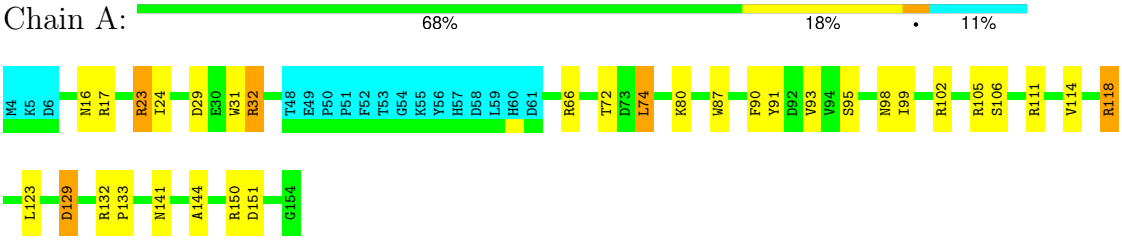
4.2.4 Score per residue for model 4

• Molecule 1: Peptide methionine sulfoxide reductase msrB



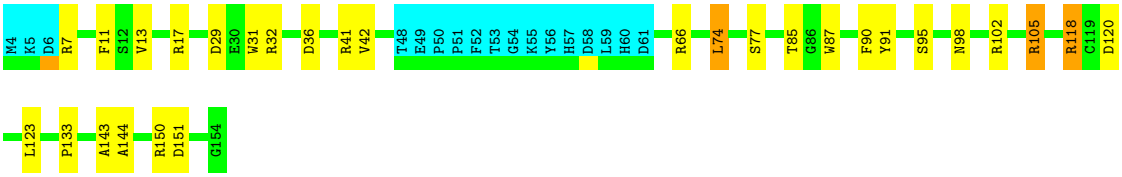
4.2.5 Score per residue for model 5

• Molecule 1: Peptide methionine sulfoxide reductase msrB



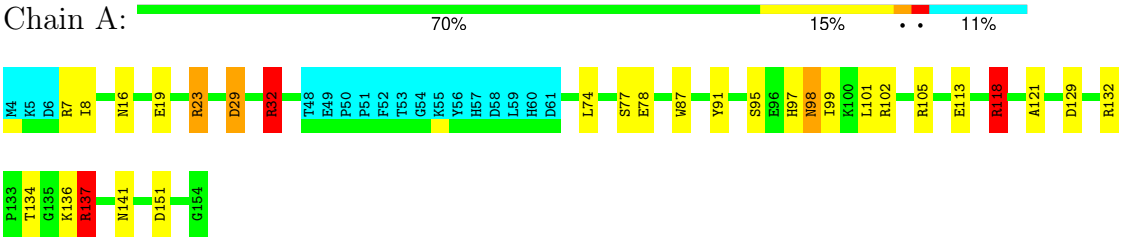
4.2.6 Score per residue for model 6 (medoid)

• Molecule 1: Peptide methionine sulfoxide reductase msrB



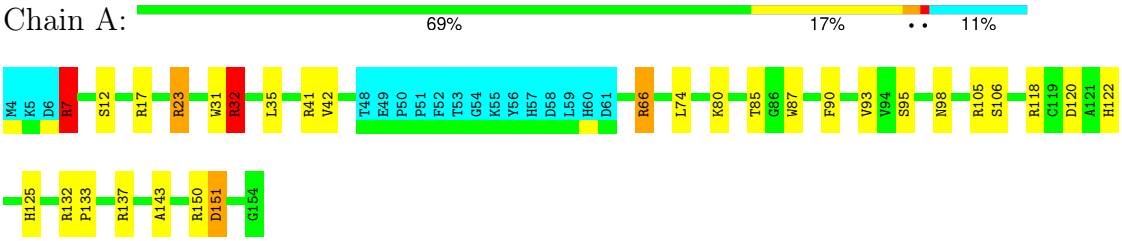
4.2.7 Score per residue for model 7

• Molecule 1: Peptide methionine sulfoxide reductase msrB



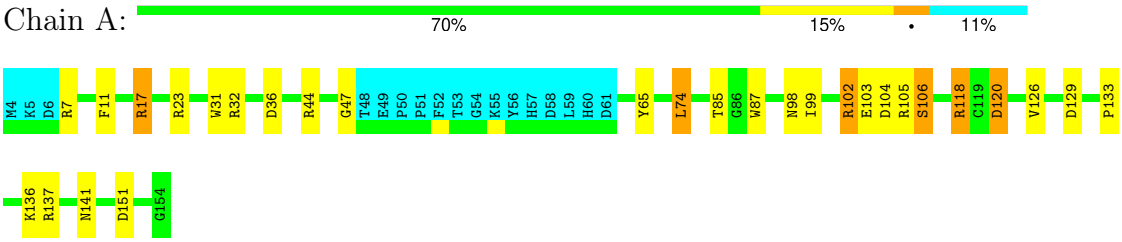
4.2.8 Score per residue for model 8

• Molecule 1: Peptide methionine sulfoxide reductase msrB



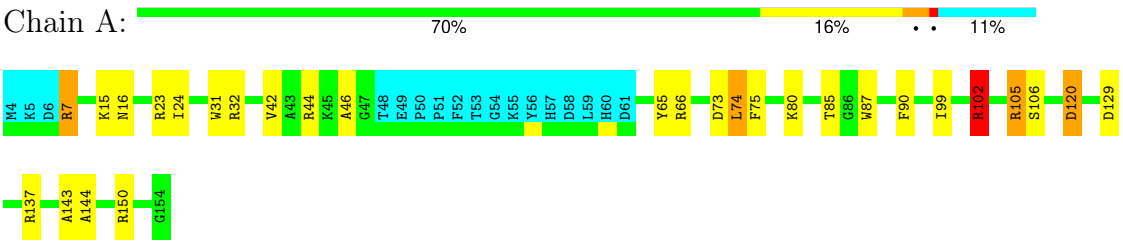
4.2.9 Score per residue for model 9

• Molecule 1: Peptide methionine sulfoxide reductase msrB



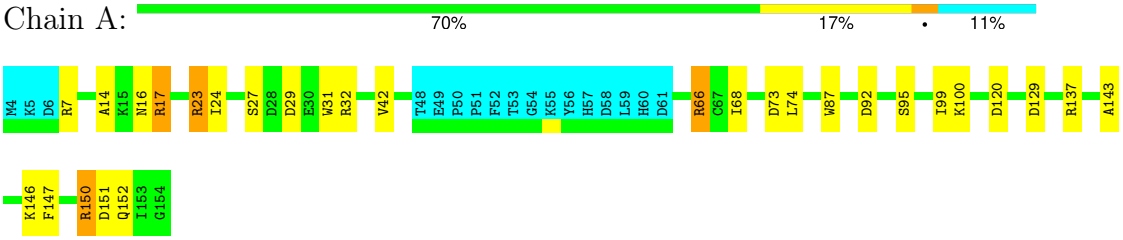
4.2.10 Score per residue for model 10

• Molecule 1: Peptide methionine sulfoxide reductase msrB



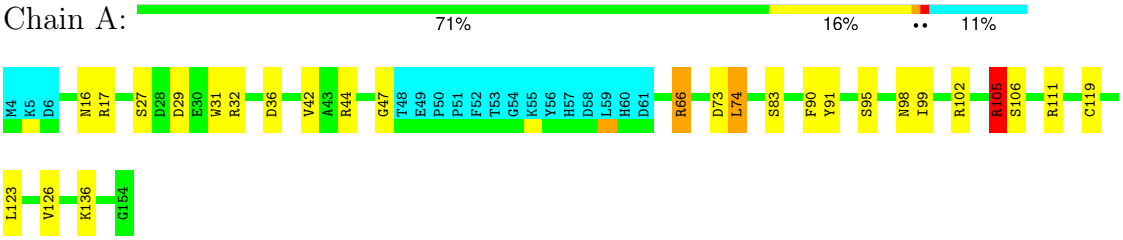
4.2.11 Score per residue for model 11

• Molecule 1: Peptide methionine sulfoxide reductase msrB



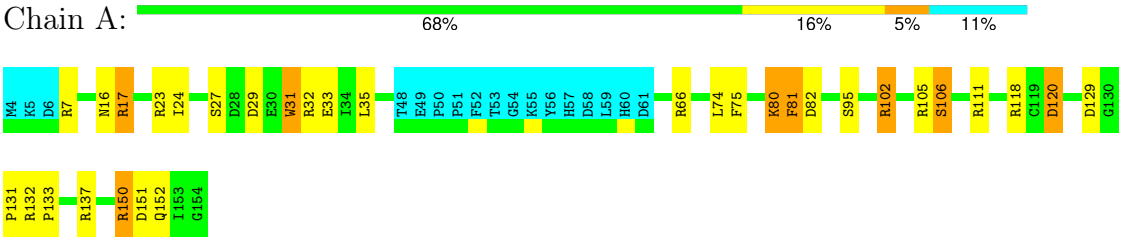
4.2.12 Score per residue for model 12

• Molecule 1: Peptide methionine sulfoxide reductase msrB



4.2.13 Score per residue for model 13

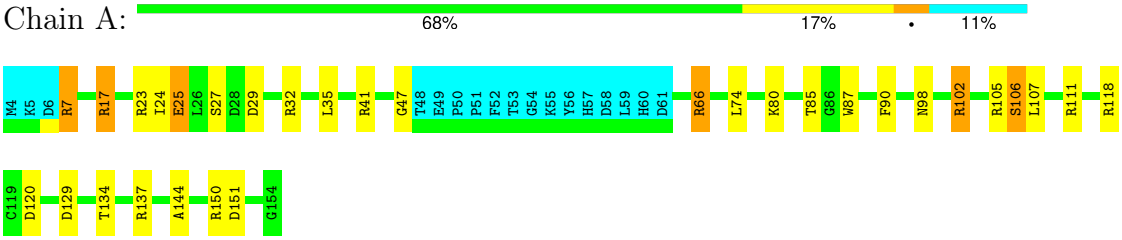
• Molecule 1: Peptide methionine sulfoxide reductase msrB



4.2.14 Score per residue for model 14

• Molecule 1: Peptide methionine sulfoxide reductase msrB

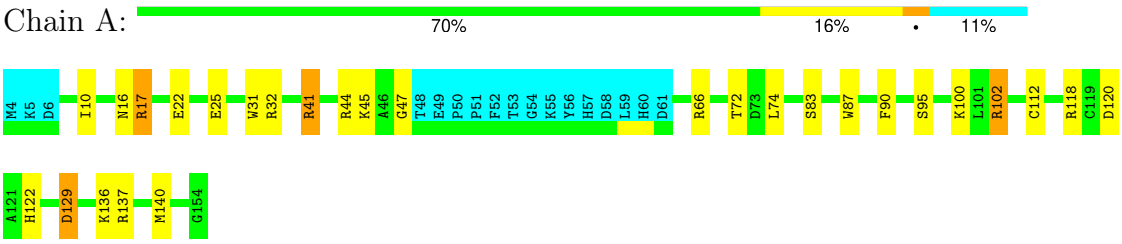
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Peptide methionine sulfoxide reductase msrB

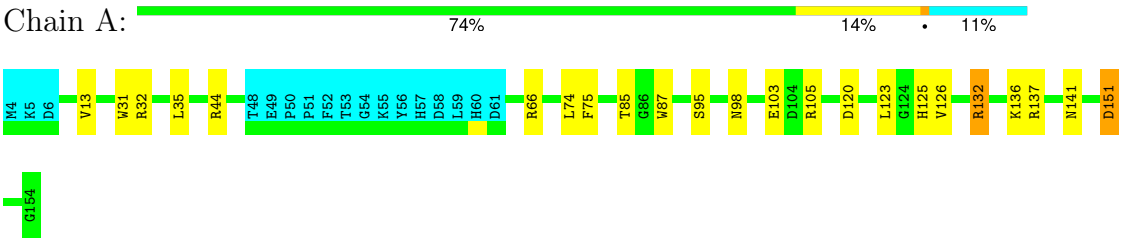
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Peptide methionine sulfoxide reductase msrB

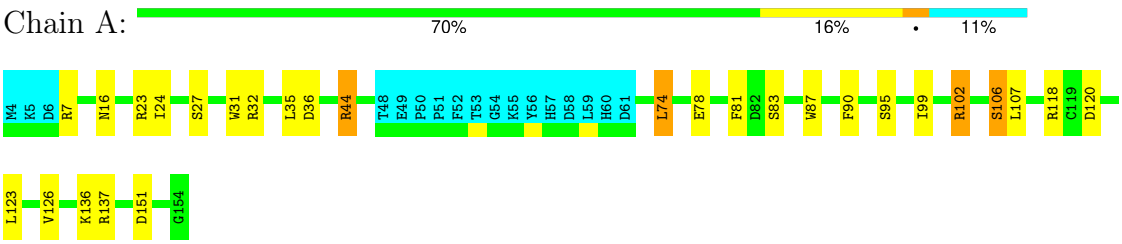
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: Peptide methionine sulfoxide reductase msrB

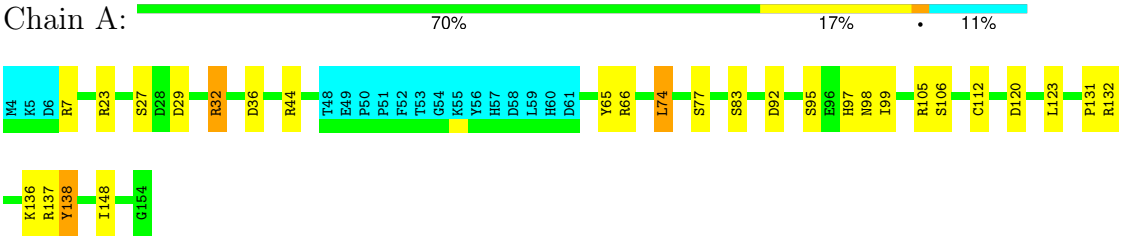
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: Peptide methionine sulfoxide reductase msrB

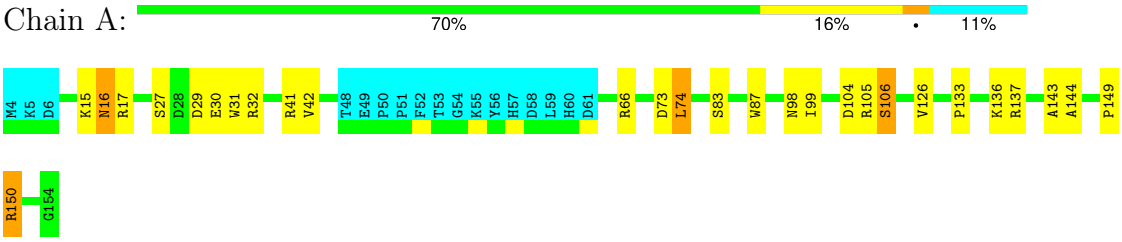
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: Peptide methionine sulfoxide reductase msrB

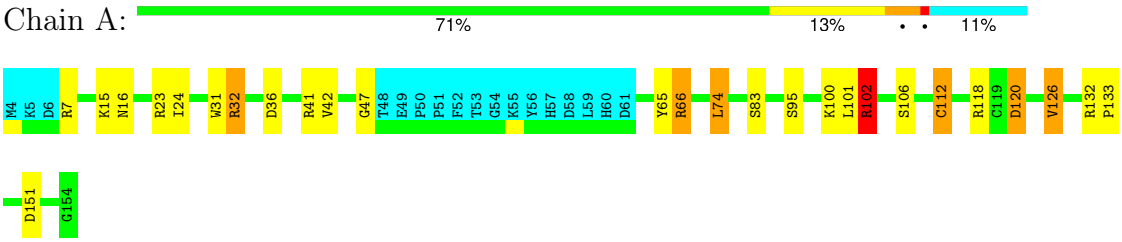
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Peptide methionine sulfoxide reductase msrB

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *DGSA-distance geometry simulated annealing, energy minimisation.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
DYANA	structure solution	1.5
OPAL	refinement	2.6

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

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

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	0.62±0.01	0±0/1092 (0.0± 0.0%)	1.17±0.03	5±2/1469 (0.3± 0.1%)
All	All	0.62	0/21840 (0.0%)	1.17	93/29380 (0.3%)

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	2.8±1.2
All	All	0	56

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	105	ARG	NE-CZ-NH1	9.97	125.29	120.30	18	2
1	A	137	ARG	NE-CZ-NH2	-9.57	115.52	120.30	14	9
1	A	32	ARG	NE-CZ-NH1	8.95	124.77	120.30	7	4
1	A	118	ARG	NE-CZ-NH1	8.38	124.49	120.30	5	5
1	A	7	ARG	NE-CZ-NH2	-8.17	116.22	120.30	8	6
1	A	23	ARG	CD-NE-CZ	7.90	134.66	123.60	11	1
1	A	32	ARG	NE-CZ-NH2	-7.85	116.37	120.30	7	1
1	A	17	ARG	NE-CZ-NH1	7.74	124.17	120.30	3	6
1	A	118	ARG	CD-NE-CZ	7.51	134.12	123.60	5	2
1	A	102	ARG	NE-CZ-NH2	-7.40	116.60	120.30	20	5
1	A	23	ARG	NE-CZ-NH1	7.36	123.98	120.30	11	3
1	A	17	ARG	NE-CZ-NH2	-6.99	116.81	120.30	1	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	105	ARG	NE-CZ-NH2	-6.96	116.82	120.30	10	3
1	A	23	ARG	NE-CZ-NH2	-6.68	116.96	120.30	11	4
1	A	41	ARG	NE-CZ-NH2	-6.64	116.98	120.30	15	2
1	A	111	ARG	NE-CZ-NH2	-6.50	117.05	120.30	14	5
1	A	118	ARG	NE-CZ-NH2	-6.44	117.08	120.30	4	4
1	A	41	ARG	NE-CZ-NH1	6.44	123.52	120.30	20	5
1	A	132	ARG	NE-CZ-NH2	-6.15	117.22	120.30	20	1
1	A	102	ARG	NE-CZ-NH1	6.09	123.35	120.30	15	1
1	A	150	ARG	NE-CZ-NH2	-6.03	117.28	120.30	13	1
1	A	137	ARG	NE-CZ-NH1	5.93	123.27	120.30	8	4
1	A	66	ARG	NE-CZ-NH2	-5.81	117.39	120.30	10	2
1	A	102	ARG	CD-NE-CZ	5.81	131.73	123.60	10	4
1	A	7	ARG	NE-CZ-NH1	5.56	123.08	120.30	10	2
1	A	105	ARG	CD-NE-CZ	5.53	131.35	123.60	18	1
1	A	7	ARG	CD-NE-CZ	5.49	131.28	123.60	6	1
1	A	132	ARG	CD-NE-CZ	5.43	131.20	123.60	20	1
1	A	66	ARG	CD-NE-CZ	5.36	131.11	123.60	10	1
1	A	44	ARG	NE-CZ-NH2	-5.31	117.64	120.30	18	1
1	A	32	ARG	CD-NE-CZ	5.24	130.94	123.60	7	1
1	A	150	ARG	NE-CZ-NH1	5.24	122.92	120.30	1	1
1	A	44	ARG	NE-CZ-NH1	5.12	122.86	120.30	17	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	7	ARG	Sidechain	8
1	A	66	ARG	Sidechain	8
1	A	118	ARG	Sidechain	6
1	A	102	ARG	Sidechain	6
1	A	23	ARG	Sidechain	5
1	A	17	ARG	Sidechain	5
1	A	91	TYR	Sidechain	4
1	A	137	ARG	Sidechain	3
1	A	41	ARG	Sidechain	3
1	A	32	ARG	Sidechain	3
1	A	105	ARG	Sidechain	2
1	A	65	TYR	Sidechain	1
1	A	150	ARG	Sidechain	1
1	A	138	TYR	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	1070	1029	1035	2±1
All	All	21420	20580	20730	43

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 1.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:42:VAL:HG12	1:A:143:ALA:HB1	0.67	1.66	11	5
1:A:10:ILE:HD12	1:A:72:THR:HG23	0.64	1.67	15	1
1:A:42:VAL:HG13	1:A:47:GLY:HA2	0.57	1.77	12	3
1:A:10:ILE:HD12	1:A:72:THR:CG2	0.56	2.30	15	1
1:A:68:ILE:HD12	1:A:146:LYS:HD2	0.55	1.77	11	1
1:A:42:VAL:CG1	1:A:143:ALA:HB1	0.53	2.33	6	2
1:A:68:ILE:HG22	1:A:144:ALA:HA	0.52	1.82	3	1
1:A:112:CYS:SG	1:A:126:VAL:HG13	0.51	2.45	20	1
1:A:112:CYS:HB3	1:A:126:VAL:HG13	0.50	1.83	2	1
1:A:14:ALA:HB1	1:A:150:ARG:HD2	0.49	1.84	11	1
1:A:126:VAL:HG21	1:A:136:LYS:HE2	0.48	1.85	17	5
1:A:99:ILE:HD11	1:A:114:VAL:HG13	0.46	1.87	5	1
1:A:80:LYS:HE2	1:A:81:PHE:H	0.46	1.71	13	1
1:A:92:ASP:HA	1:A:138:TYR:CE1	0.46	2.45	18	1
1:A:68:ILE:HD12	1:A:146:LYS:CD	0.46	2.40	11	1
1:A:29:ASP:HA	1:A:32:ARG:CZ	0.45	2.41	7	1
1:A:64:ILE:HG22	1:A:76:ASP:HA	0.45	1.88	1	1
1:A:31:TRP:CE3	1:A:35:LEU:CD1	0.44	3.01	13	1
1:A:102:ARG:HG2	1:A:113:GLU:CG	0.43	2.44	7	1
1:A:104:ASP:OD2	1:A:106:SER:OG	0.43	2.35	1	3
1:A:97:HIS:CG	1:A:98:ASN:N	0.42	2.88	18	3
1:A:35:LEU:HD22	1:A:39:ALA:CB	0.42	2.45	1	1
1:A:27:SER:OG	1:A:30:GLU:OE2	0.41	2.37	19	1
1:A:8:ILE:HD13	1:A:118:ARG:HD3	0.41	1.93	7	1
1:A:97:HIS:CG	1:A:98:ASN:H	0.41	2.33	18	2
1:A:14:ALA:HB1	1:A:150:ARG:CD	0.41	2.45	11	1
1:A:23:ARG:HH12	1:A:25:GLU:CD	0.41	2.19	14	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	133/151 (88%)	103±2 (78±2%)	22±3 (17±2%)	7±2 (5±2%)	3	23
All	All	2660/3020 (88%)	2069 (78%)	447 (17%)	144 (5%)	3	23

All 29 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	31	TRP	16
1	A	151	ASP	14
1	A	95	SER	14
1	A	120	ASP	13
1	A	74	LEU	11
1	A	106	SER	9
1	A	133	PRO	8
1	A	129	ASP	7
1	A	150	ARG	7
1	A	83	SER	5
1	A	144	ALA	5
1	A	75	PHE	4
1	A	16	ASN	3
1	A	132	ARG	3
1	A	27	SER	3
1	A	29	ASP	3
1	A	47	GLY	3
1	A	119	CYS	2
1	A	122	HIS	2
1	A	46	ALA	2
1	A	131	PRO	2
1	A	67	CYS	1
1	A	84	GLY	1
1	A	143	ALA	1
1	A	121	ALA	1
1	A	7	ARG	1
1	A	82	ASP	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	32	ARG	1
1	A	149	PRO	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	116/132 (88%)	99±2 (85±2%)	17±2 (15±2%)	6 45
All	All	2320/2640 (88%)	1981 (85%)	339 (15%)	6 45

All 68 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	74	LEU	20
1	A	32	ARG	18
1	A	87	TRP	15
1	A	105	ARG	14
1	A	102	ARG	12
1	A	99	ILE	11
1	A	16	ASN	11
1	A	66	ARG	11
1	A	98	ASN	11
1	A	24	ILE	9
1	A	29	ASP	9
1	A	106	SER	9
1	A	36	ASP	9
1	A	90	PHE	9
1	A	44	ARG	8
1	A	118	ARG	8
1	A	123	LEU	7
1	A	23	ARG	7
1	A	85	THR	7
1	A	17	ARG	6
1	A	80	LYS	6
1	A	120	ASP	6
1	A	13	VAL	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	27	SER	5
1	A	129	ASP	5
1	A	35	LEU	5
1	A	141	ASN	5
1	A	132	ARG	5
1	A	136	LYS	4
1	A	81	PHE	4
1	A	77	SER	4
1	A	112	CYS	4
1	A	137	ARG	4
1	A	65	TYR	4
1	A	73	ASP	4
1	A	134	THR	3
1	A	103	GLU	3
1	A	125	HIS	3
1	A	151	ASP	3
1	A	150	ARG	3
1	A	15	LYS	3
1	A	100	LYS	3
1	A	92	ASP	2
1	A	122	HIS	2
1	A	93	VAL	2
1	A	11	PHE	2
1	A	78	GLU	2
1	A	101	LEU	2
1	A	7	ARG	2
1	A	152	GLN	2
1	A	25	GLU	2
1	A	107	LEU	2
1	A	83	SER	2
1	A	41	ARG	1
1	A	89	SER	1
1	A	28	ASP	1
1	A	76	ASP	1
1	A	72	THR	1
1	A	19	GLU	1
1	A	91	TYR	1
1	A	12	SER	1
1	A	147	PHE	1
1	A	33	GLU	1
1	A	22	GLU	1
1	A	45	LYS	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	140	MET	1
1	A	148	ILE	1
1	A	126	VAL	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 monosaccharides 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 (i)

No chemical shift data were provided