W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB NMR Structure Validation Report ⓘ

Feb 26, 2022 – 11:11 AM EST

PDB ID	:	2B1O
Title	:	Solution Structure of Ca2+-bound DdCAD-1
Authors	:	Lin, Z.; Sriskanthadevan, S.; Huang, H.B.; Siu, C.H.; Yang, D.W.
Deposited on	:	2005-09-16

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

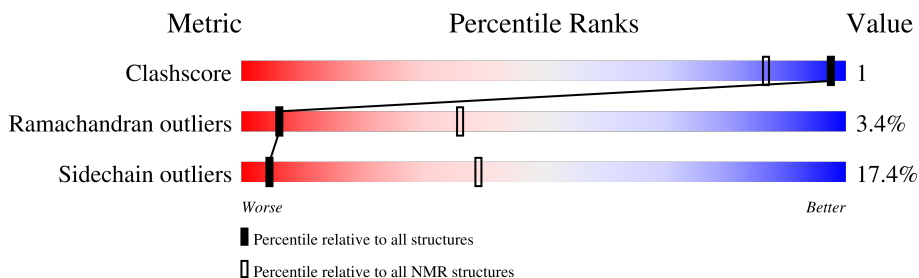|		
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al.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

1   Overall quality at a glance ⓘ

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		1
Ramachandran outliers		3.4%
Sidechain outliers		17.4%

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	212	86% 11% •

2 Ensemble composition and analysis ⓘ

This entry contains 10 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: *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:3-A:31, A:37-A:213 (206)	0.56	2

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 1 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 3, 4, 5, 6, 8, 9, 10
Single-model clusters	7

3 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 3040 atoms, of which 1361 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Calcium-dependent cell adhesion molecule-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	212	3037	1061	1361	276	332	7	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol	Chain	Residues	Atoms	
			Total	Ca
2	A	3	3	3

4 Residue-property plots ⓘ

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.

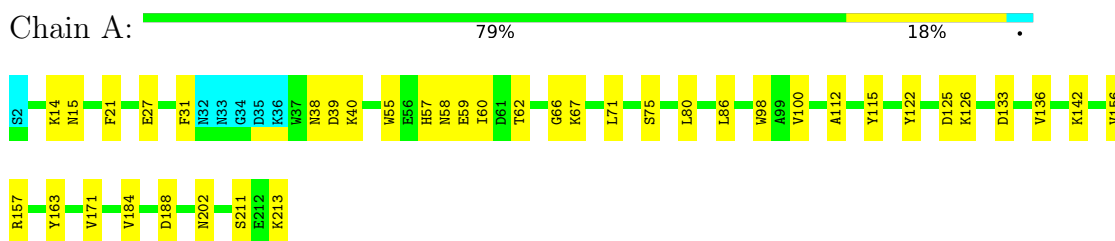- Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A:
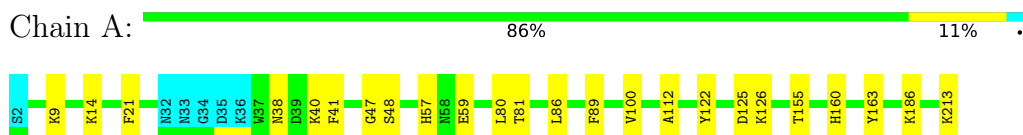
4.2 Scores per residue for each member of the ensemble

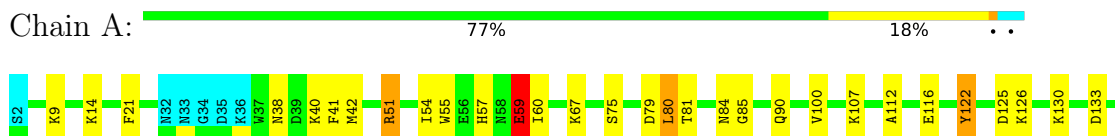Colouring as in section 4.1 above.
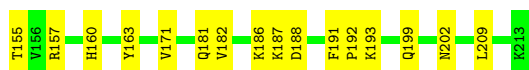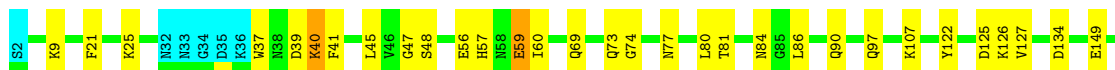
4.2.1 Score per residue for model 1

- Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A:

4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A:

L143 D147 T155 H160 Y163 D164 V184 Q199 D200 D201 N202 F205 L209 K213

4.2.3 Score per residue for model 3

• Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A: 75% 21%

S2 K9 F21 K25 N32 N33 G34 D35 K36 W37 N38 D39 K40 F41 L45 V46 G47 S48 E56 H57 N58 E59 I60 Q69 Q73 G74 N77 L80 T81 N84 G85 L86 Q90 Q97 K107 Y122 D125 K126 V127 D134 E149

T155 V156 R157 H160 Y163 V171 Q181 V182 K186 K187 D188 F191 P192 K193 Q199 N202 L209 K213

4.2.4 Score per residue for model 4

• Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A: 78% 17%

S2 V3 D4 V8 K9 F10 K14 F21 F31 N32 N33 G34 D35 K36 W37 N38 D39 K40 F41 L45 V46 G47 S48 R51 C52 S75 L80 N84 G85 L86 S87 K88 F89 L92 F96 Q97 V100 K107 Y122 D125 K126 V127 D133

T155 Q158 T159 H160 T161 P162 Y163 V184 I185 K186 K213

4.2.5 Score per residue for model 5

• Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A: 81% 16%

S2 V8 K9 K14 F21 E27 F31 N32 N33 G34 D35 K36 W37 N38 F41 M42 N49 H57 D79 L80 T81 F89 W98 A99 V100 D101 A112 Y115 K126 Q137 I140 D147 R157 H160 T161 P162 Y163 Q181 V182 T183

V184 I185 K186 K193 L209 K213
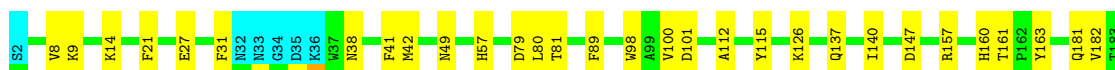
4.2.6 Score per residue for model 6

• Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A: 76% 20%

S2 F10 G13 K14 F21 V29 N32 N33 G34 D35 K36 K40 G47 S48 C52 W55 E56 H57 N58 E59 I60 D61 T62 K67 L80 T81 S82 I83 L86 F89 W98 A99 V100 S110 Y122 D125 K126 V127 D131 G132 D133

Q137 L138 L143 H153 H160 Y163 V182 K186 K193 N194 M195 S211 E212 K213
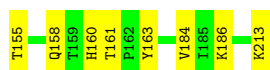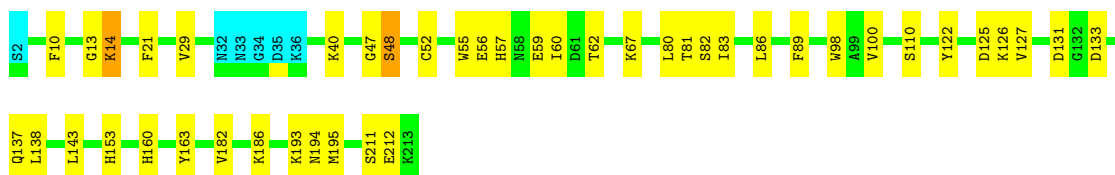
4.2.7 Score per residue for model 7

- Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A: 80% 17% •

S2 V8 K9 F12 K25 N32 N33 G34 D35 K36 K40 F41 G47 S48 H57 L80 F89 W98 A99 V100 K103 I104 K107 A112 E116 Y122 K126 D131 L138 K142 S148 E149 I150 T155 Q158 K174

V184 I185 K186 K187 P192 K193 S211 E212 K213
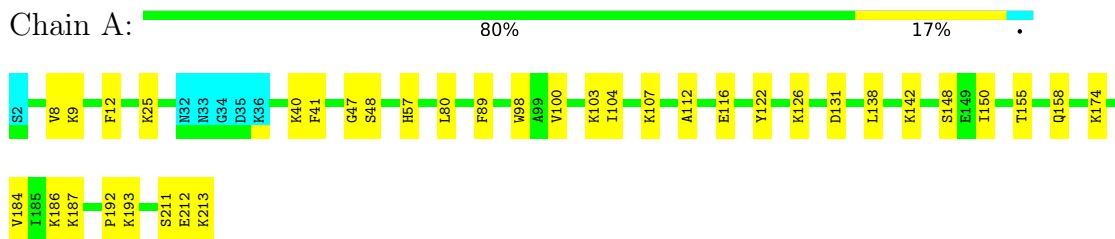
4.2.8 Score per residue for model 8

- Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A: 75% 21% • •

S2 K9 F12 G13 K14 E19 S20 F21 K25 R30 F31 N32 N33 G34 D35 K36 W37 N38 D39 K40 F41 G47 S48 R51 C52 N53 H57 N58 E59 I60 E70 S75 L80 T81 S82 I83 P93 F96 V100 V105 S110 T111 A112

D125 K126 V127 D131 G132 D133 Q137 L138 P139 K142 L143 L154 Q158 T159 H160 T161 K186 L209 N210 S211 E212 K213
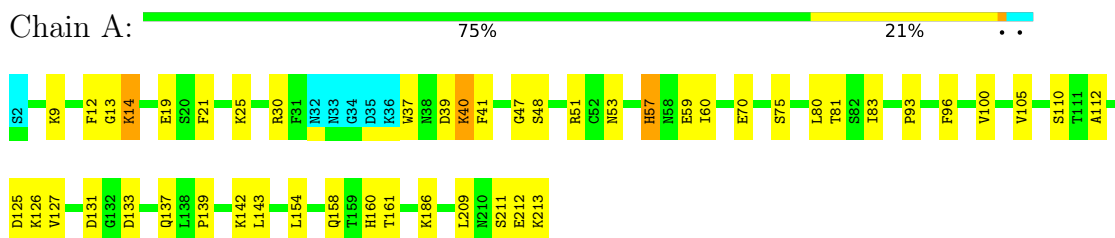
4.2.9 Score per residue for model 9

- Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A: 77% 18% • •

S2 K9 F10 K14 F31 N32 N33 G34 D35 K36 W37 N38 D39 K40 G47 S48 C52 H57 N58 E59 G66 K67 S75 T76 N77 K88 F89 K107 S110 T111 A112 E116 Y122 D125 K126 Q137 D147 S148 T155 V156 R157 D164

Y165 V166 S170 V171 K174 V182 K186 E189 T190 F191 S211 E212 K213
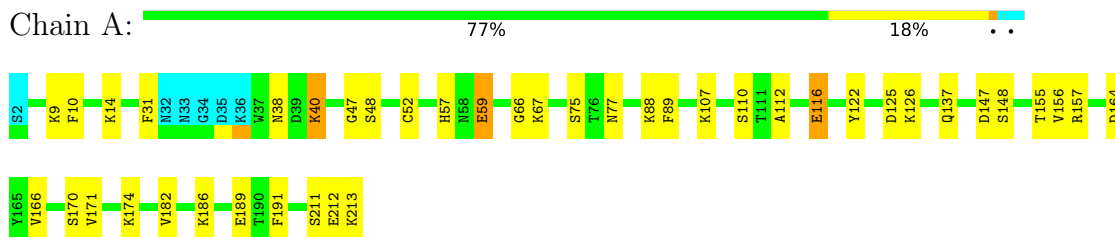
4.2.10 Score per residue for model 10

- Molecule 1: Calcium-dependent cell adhesion molecule-1

Chain A:

76%   20%

S2 F10 F11 F12 K25 N32 N33 G34 D35 K36 W37 N38 D39 K40 F41 L45 V46 G47 S48 R51 I54 W55 E56 H57 N58 E59 Q73 T81 L86 F89 L92 F96 Q97 W98 K103 S110 T111 A112 M117 Y122 Q123 V124 D125 K126

C129 K130 D131 D134 Q137 T155 V156 R157 H160 Y163 D164 Y165 V166 V182 K186 K187 M195 K213
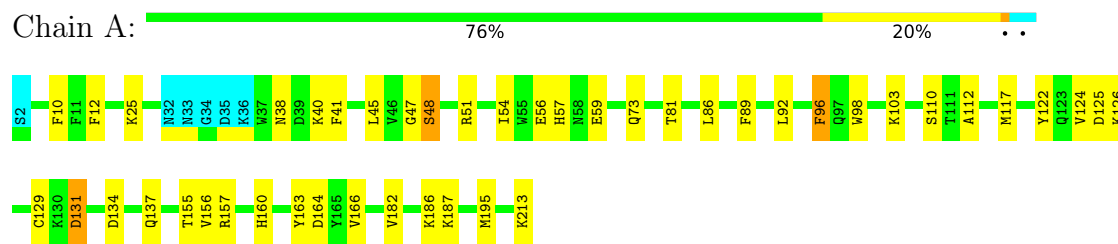
5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *torsion angle dynamics.*

Of the 100 calculated structures, 10 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
CYANA	structure solution	1.0.5
Amber	refinement	7.0

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: CA

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	0.71±0.01	0±0/1673 (0.0± 0.0%)	1.02±0.02	1±1/2275 (0.1± 0.1%)
All	All	0.71	0/16730 (0.0%)	1.02	13/22750 (0.1%)

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	59	GLU	OE1-CD-OE2	-6.21	115.85	123.30	2	4
1	A	51	ARG	NE-CZ-NH1	6.08	123.34	120.30	10	3
1	A	56	GLU	OE1-CD-OE2	-5.34	116.89	123.30	3	3
1	A	124	VAL	CA-CB-CG1	5.25	118.77	110.90	10	1
1	A	157	ARG	NE-CZ-NH1	5.15	122.87	120.30	1	2

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	1633	1332	1577	2±1
All	All	16360	13320	15770	17

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:92:LEU:HD13	1:A:96:PHE:CE1	0.64	2.27	10	1
1:A:92:LEU:HD13	1:A:96:PHE:CE2	0.50	2.41	4	1
1:A:117:MET:HG3	1:A:156:VAL:HG22	0.46	1.86	10	1
1:A:181:GLN:HG2	1:A:182:VAL:H	0.46	1.70	3	2
1:A:3:VAL:HG22	1:A:4:ASP:H	0.46	1.70	4	1
1:A:80:LEU:HD13	1:A:80:LEU:H	0.45	1.71	2	1
1:A:116:GLU:O	1:A:156:VAL:HG23	0.45	2.11	9	1
1:A:191:PHE:CD1	1:A:192:PRO:HD2	0.44	2.47	3	1
1:A:115:TYR:CZ	1:A:156:VAL:HG21	0.44	2.48	1	1
1:A:93:PRO:CG	1:A:96:PHE:CD2	0.43	3.01	8	1
1:A:184:VAL:HG13	1:A:205:PHE:CZ	0.42	2.49	2	1
1:A:156:VAL:HG22	1:A:166:VAL:CG2	0.42	2.44	9	1
1:A:54:ILE:HB	1:A:86:LEU:HD13	0.42	1.91	10	1
1:A:96:PHE:CD2	1:A:139:PRO:CB	0.42	3.03	8	1
1:A:98:TRP:CZ2	1:A:140:ILE:HG21	0.42	2.50	5	1
1:A:41:PHE:CE2	1:A:89:PHE:CE2	0.41	3.09	5	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	205/212 (97%)	173±3 (84±2%)	25±4 (12±2%)	7±2 (3±1%)	6	36
All	All	2050/2120 (97%)	1730 (84%)	251 (12%)	69 (3%)	6	36

All 25 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	125	ASP	8
1	A	112	ALA	7
1	A	122	TYR	7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	47	GLY	7
1	A	48	SER	7
1	A	84	ASN	3
1	A	40	LYS	3
1	A	131	ASP	3
1	A	39	ASP	2
1	A	60	ILE	2
1	A	66	GLY	2
1	A	59	GLU	2
1	A	13	GLY	2
1	A	14	LYS	2
1	A	110	SER	2
1	A	71	LEU	1
1	A	85	GLY	1
1	A	202	ASN	1
1	A	74	GLY	1
1	A	3	VAL	1
1	A	86	LEU	1
1	A	212	GLU	1
1	A	192	PRO	1
1	A	193	LYS	1
1	A	57	HIS	1

6.3.2 Protein sidechains ⓘ

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	188/193 (97%)	155±4 (83±2%)	33±4 (17±2%)	4	39
All	All	1880/1930 (97%)	1552 (83%)	328 (17%)	4	39

All 108 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	126	LYS	10
1	A	40	LYS	9
1	A	57	HIS	9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	80	LEU	8
1	A	186	LYS	8
1	A	14	LYS	7
1	A	21	PHE	7
1	A	100	VAL	7
1	A	163	TYR	7
1	A	9	LYS	7
1	A	160	HIS	7
1	A	38	ASN	6
1	A	213	LYS	6
1	A	41	PHE	6
1	A	81	THR	6
1	A	155	THR	6
1	A	75	SER	5
1	A	133	ASP	5
1	A	211	SER	5
1	A	59	GLU	5
1	A	107	LYS	5
1	A	89	PHE	5
1	A	137	GLN	5
1	A	31	PHE	4
1	A	67	LYS	4
1	A	86	LEU	4
1	A	98	TRP	4
1	A	184	VAL	4
1	A	209	LEU	4
1	A	25	LYS	4
1	A	127	VAL	4
1	A	10	PHE	4
1	A	55	TRP	3
1	A	142	LYS	3
1	A	171	VAL	3
1	A	116	GLU	3
1	A	143	LEU	3
1	A	147	ASP	3
1	A	164	ASP	3
1	A	45	LEU	3
1	A	60	ILE	3
1	A	157	ARG	3
1	A	187	LYS	3
1	A	193	LYS	3
1	A	8	VAL	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	52	CYS	3
1	A	158	GLN	3
1	A	161	THR	3
1	A	182	VAL	3
1	A	12	PHE	3
1	A	212	GLU	3
1	A	27	GLU	2
1	A	62	THR	2
1	A	188	ASP	2
1	A	202	ASN	2
1	A	42	MET	2
1	A	51	ARG	2
1	A	79	ASP	2
1	A	90	GLN	2
1	A	122	TYR	2
1	A	199	GLN	2
1	A	37	TRP	2
1	A	39	ASP	2
1	A	73	GLN	2
1	A	77	ASN	2
1	A	97	GLN	2
1	A	134	ASP	2
1	A	88	LYS	2
1	A	96	PHE	2
1	A	48	SER	2
1	A	83	ILE	2
1	A	131	ASP	2
1	A	138	LEU	2
1	A	195	MET	2
1	A	103	LYS	2
1	A	148	SER	2
1	A	174	LYS	2
1	A	110	SER	2
1	A	15	ASN	1
1	A	58	ASN	1
1	A	136	VAL	1
1	A	54	ILE	1
1	A	130	LYS	1
1	A	200	ASP	1
1	A	201	ASP	1
1	A	69	GLN	1
1	A	149	GLU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	84	ASN	1
1	A	49	ASN	1
1	A	101	ASP	1
1	A	115	TYR	1
1	A	29	VAL	1
1	A	82	SER	1
1	A	153	HIS	1
1	A	194	ASN	1
1	A	104	ILE	1
1	A	150	ILE	1
1	A	19	GLU	1
1	A	30	ARG	1
1	A	53	ASN	1
1	A	70	GLU	1
1	A	105	VAL	1
1	A	154	LEU	1
1	A	170	SER	1
1	A	189	GLU	1
1	A	191	PHE	1
1	A	129	CYS	1
1	A	166	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 3 ligands modelled in this entry, 3 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