Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 01:33 AM EDT

PDB ID	:	2LUD
BMRB ID	:	18517
Title	:	Solution structure of a conformational mutant of the adhesion protein delta-Bd37 from Babesia divergens
Authors	:	Murciano, B.; Barthe, P.; Delbecq, S.; Roumestand, C.
Deposited on	:	2012-06-12

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:

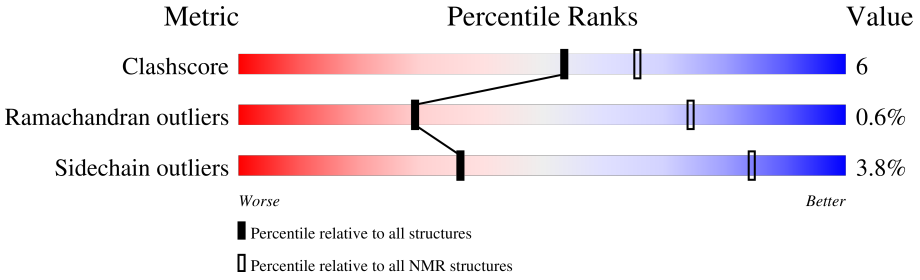
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 81%.

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	224	82% 10% 8%

2 Ensemble composition and analysis (i)

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: *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:71-A:159, A:170-A:287 (207)	1.15	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 3 clusters and 9 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3505 atoms, of which 1781 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Glycosylphosphatidylinositol-anchored merozoite surface protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	224	3505	1097	1781	283	338	6	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	248	ALA	LYS	engineered mutation	UNP Q8T117
A	279	ALA	GLU	engineered mutation	UNP Q8T117
A	283	ALA	ASP	engineered mutation	UNP Q8T117

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: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 82% 10% 8%

V69 K70 R83 F86 I90 P96 I99 R107 Y132 E144 A145 A146 L150 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 F181 Q189 L212 F216 S230 G231 Y232 N266 L270 V271 I275 R278 M281 M282 P288 I289

G290 H291 H292

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: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 76% 14% • 8%

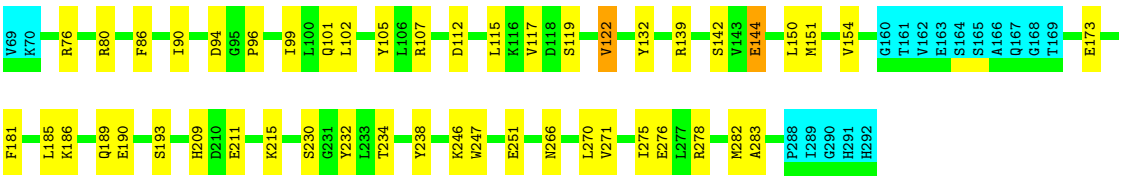
V69 K70 T71 L72 R76 F86 I90 P96 I99 R107 D110 T111 L114 K128 Y132 A146 L150 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 F181 K186 Q189 G197 L201 M205 H209 L212 T225 S226 E227

D228 I229 S230 L233 T234 M242 W247 N266 L270 L274 I275 E276 M281 M282 P288 I289 G290 H291 H292

4.2.2 Score per residue for model 2

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

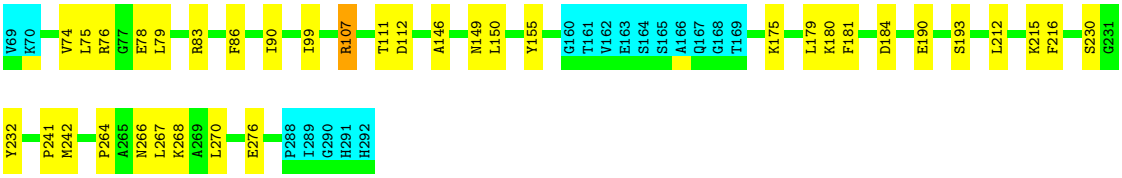
Chain A: 71% 21% • 8%



4.2.3 Score per residue for model 3

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

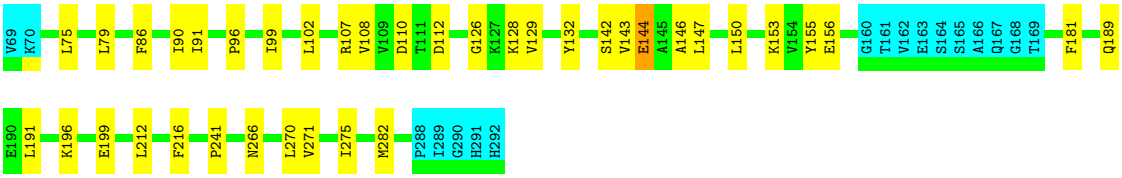
Chain A: 76% 16% 8%



4.2.4 Score per residue for model 4

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

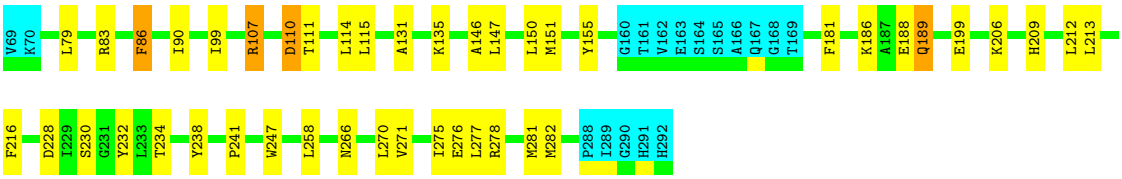
Chain A: 75% 17% 8%



4.2.5 Score per residue for model 5

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 73% 18% • 8%



4.2.6 Score per residue for model 6 (medoid)

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 73% 19% 8%

V69 K70 V74 L75 E78 L79 R83 F86 I90 D94 G95 P96 I99 L115 D118 K121 A131 Y132 K135 A146 L150 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 D170 K175 K180 F181 D184 E188 Q189 E199

K206 L212 F216 T234 Y238 M242 W247 G256 S260 N266 L267 K268 A269 L270 V271 L274 I275 M281 P288 I289 G290 H291 H292

4.2.7 Score per residue for model 7

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 77% 15% 8%

V69 K70 E84 A85 F86 I90 P96 I99 T111 L114 D118 K121 A131 Y132 K135 S142 V143 E144 L147 D148 N149 L150 M151 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 D170 L174 F181 L185 E188 Q189 L212 S226

E227 D228 I229 S230 G231 Y232 N266 L270 V271 L274 I275 P288 I289 G290 H291 H292

4.2.8 Score per residue for model 8

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 76% 16% 8%

V69 K70 T71 L75 F86 I90 D94 R107 D118 K130 A131 Y132 K135 G140 E144 A145 A146 L147 D148 N149 L150 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 D170 S171 K175 L185 Q189 L201 M205 K206 T207 Q208 E211

S230 P236 E237 A240 P241 M242 N266 L270 V271 I275 M281 M282 P288 I289 G290 H291 H292

4.2.9 Score per residue for model 9

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 75% 17% 8%

V69 K70 T71 L75 R83 F86 I90 T98 V117 D118 L133 I136 G137 I138 E144 A145 A146 L150 V154 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 L174 A187 E188 Q189 E190 L191 G197 L201 M205 E211 K214

S226 E227 S230 G231 Y232 L233 T234 A245 K246 W247 A248 K249 P264 A265 N266 L270 V271 I275 M281 M282 P288 I289 G290 H291 H292

4.2.10 Score per residue for model 10

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 73% 18% • 8%

V69 K70 T71 L72 R76 R80 F86 I90 R107 D110 K116 V117 D118 S119 V122 V143 A146 L147 L150 V154 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 D170 S171 E172 E173 L174 K180 D184 H192 K196 E199 A200 L201

L202 M205 L212 K215 F216 T225 I229 P236 M242 H255 N266 A269 L270 V271 A272 E273 E276 M281 P288 I289 G290 H291 H292

4.2.11 Score per residue for model 11

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 73% 19% • 8%

V69 K70 T71 L75 L79 Q82 R83 E84 A85 E89 I99 L102 V103 D118 E124 K128 L133 I136 V143 A146 L147 D148 N149 L150 M151 V154 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 D170 S171 K175 L179 A187 E188

Q189 E190 L191 H192 L212 F216 S230 G231 Y232 E262 N266 L270 V271 I275 R278 M281 M282 A283 L284 L285 Y286 G287 P288 I289 G290 H291 H292

4.2.12 Score per residue for model 12

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 78% 13% • 8%

V69 K70 L75 E78 L79 R80 G81 Q82 R83 E84 A85 F86 I90 P96 R107 D110 Y132 A146 L147 D148 N149 L150 M151 I152 Y155 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 L201 M205 E211 L212 L213 K214 K215 F216 S230 G231 Y232

P241 N266 L270 E276 L277 M281 P288 I289 G290 H291 H292

4.2.13 Score per residue for model 13

- Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 75% 17% 8%

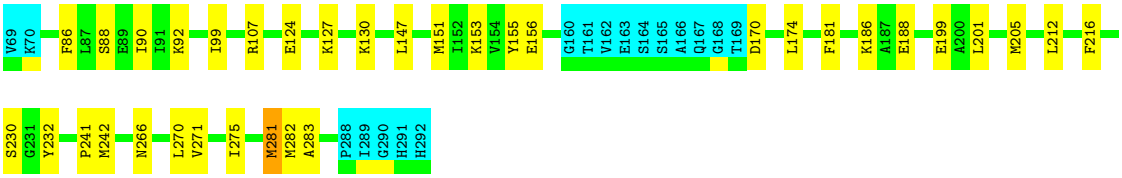
V69 K70 T71 L72 R76 F86 I90 D94 G95 P96 I99 L102 D112 Y132 L150 M151 I152 K153 V154 Y155 E156 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 K180 F181 D184 L201 M205 Q208 H209 D210 E211 L212 F216 S230

G231 Y232 L233 T234 W247 E251 I254 L258 N266 L270 V271 I275 E276 L277 R278 M281 M282 P288 I289 G290 H291 H292

4.2.14 Score per residue for model 14

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

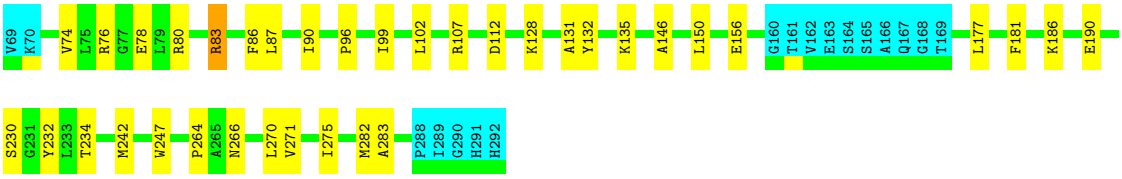
Chain A: 77% 15% 8%



4.2.15 Score per residue for model 15

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

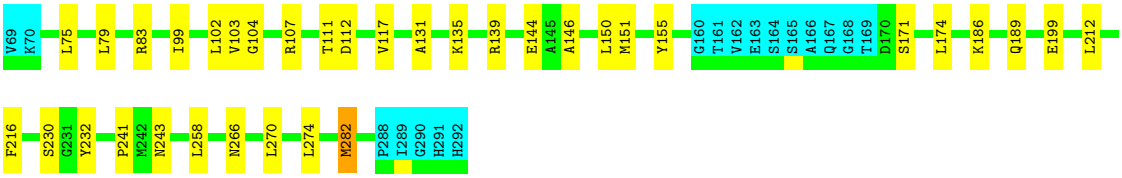
Chain A: 76% 16% 8%



4.2.16 Score per residue for model 16

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

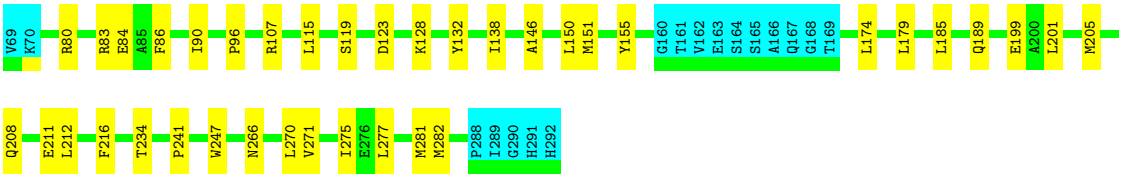
Chain A: 77% 15% 8%



4.2.17 Score per residue for model 17

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 75% 17% 8%



4.2.18 Score per residue for model 18

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 76% 16% 8%

V69 K70 T71 L75 R76 L79 Q82 R83 F86 I90 D94 G95 P96 I99 R107 D112 L113 L114 D118 S119 T120 A131 Y132 K135 L150 M151 V154 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 D170 L174 F181 L212

F216 T220 T234 W247 E251 I254 N266 L270 R278 M282 P288 I289 G290 H291 H292

4.2.19 Score per residue for model 19

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 74% 17% • 8%

V69 K70 T71 L72 R76 F86 E89 I90 I99 L100 Q101 L102 D112 K116 V117 T120 E124 K127 K128 I138 E144 A145 A146 L150 Y155 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 L174 F181 E199 L202 M205

E211 L212 L213 K214 K215 F216 T225 S226 E227 D228 I229 S230 G231 Y232 P241 H255 N266 L270 L274 R278 M281 M282 P288 I289 G290 H291 H292

4.2.20 Score per residue for model 20

• Molecule 1: Glycosylphosphatidylinositol-anchored merozoite surface protein

Chain A: 76% 15% 8%

V69 K70 L75 L79 R83 E84 A85 F86 I90 P96 L102 Y105 D112 K128 Y132 E144 A145 A146 N149 L150 G160 T161 V162 E163 S164 S165 A166 Q167 G168 T169 D170 S171 L174 Q189 H192 L212 F216 T234 M242

W247 A265 N266 L267 K268 A269 L270 V271 L274 I275 R278 M281 M282 Y286 G287 P288 I289 G290 H291 H292

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *molecular dynamics, simulated annealing.*

Of the 600 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy.*

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

Software name	Classification	Version
CYANA	structure solution	2.1
CNS	refinement	1.2

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	2411
Number of shifts mapped to atoms	2411
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	81%

6 Model quality (i)

6.1 Standard geometry (i)

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	0.72±0.01	0±0/1625 (0.0± 0.0%)	0.68±0.01	0±0/2190 (0.0± 0.0%)
All	All	0.72	1/32500 (0.0%)	0.68	2/43800 (0.0%)

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	0.6±0.8
All	All	0	11

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	262	GLU	C-N	-5.62	1.21	1.34	11	1

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	278	ARG	NE-CZ-NH1	6.73	123.66	120.30	19	2

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	107	ARG	Sidechain	4
1	A	76	ARG	Sidechain	2
1	A	83	ARG	Sidechain	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	278	ARG	Sidechain	2
1	A	80	ARG	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	1604	1664	1664	18±2
All	All	32080	33280	33279	361

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 6.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:205:MET:HG2	1:A:281:MET:SD	0.63	2.34	12	6
1:A:155:TYR:CE1	1:A:241:PRO:HA	0.60	2.31	16	5
1:A:230:SER:HA	1:A:232:TYR:CE2	0.59	2.32	13	11
1:A:170:ASP:HA	1:A:174:LEU:HD13	0.59	1.72	10	1
1:A:143:VAL:O	1:A:147:LEU:HG	0.58	1.97	11	2
1:A:278:ARG:O	1:A:282:MET:HG2	0.58	1.99	20	1
1:A:205:MET:HA	1:A:281:MET:SD	0.58	2.38	17	3
1:A:138:ILE:HD11	1:A:174:LEU:HG	0.57	1.76	19	2
1:A:155:TYR:CZ	1:A:241:PRO:HA	0.57	2.35	4	3
1:A:131:ALA:O	1:A:135:LYS:HG3	0.56	2.01	16	6
1:A:180:LYS:O	1:A:184:ASP:HB2	0.56	1.99	10	4
1:A:96:PRO:HA	1:A:132:TYR:CE1	0.56	2.36	6	11
1:A:189:GLN:HA	1:A:192:HIS:CD2	0.55	2.37	11	1
1:A:146:ALA:O	1:A:150:LEU:HG	0.55	2.02	20	14
1:A:271:VAL:O	1:A:275:ILE:HG13	0.55	2.02	5	11
1:A:266:ASN:O	1:A:270:LEU:HG	0.55	2.01	6	20
1:A:86:PHE:O	1:A:90:ILE:HG12	0.54	2.03	17	18
1:A:190:GLU:HA	1:A:193:SER:OG	0.54	2.03	3	2
1:A:212:LEU:HD13	1:A:274:LEU:HA	0.54	1.80	6	5
1:A:107:ARG:O	1:A:111:THR:HG22	0.54	2.03	5	1
1:A:234:THR:HG22	1:A:247:TRP:CZ3	0.54	2.38	9	9

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:170:ASP:HA	1:A:174:LEU:HD23	0.54	1.80	18	2
1:A:208:GLN:O	1:A:211:GLU:HB3	0.53	2.04	17	3
1:A:201:LEU:O	1:A:205:MET:HG3	0.53	2.04	14	8
1:A:83:ARG:NH2	1:A:117:VAL:HG11	0.53	2.19	9	2
1:A:72:LEU:O	1:A:76:ARG:HG2	0.52	2.03	19	2
1:A:271:VAL:O	1:A:275:ILE:HG12	0.52	2.04	20	2
1:A:104:GLY:O	1:A:107:ARG:HB3	0.52	2.05	16	1
1:A:99:ILE:HG21	1:A:181:PHE:O	0.52	2.05	15	11
1:A:124:GLU:O	1:A:127:LYS:HB3	0.52	2.04	19	2
1:A:185:LEU:O	1:A:189:GLN:HG3	0.51	2.05	7	2
1:A:247:TRP:HE1	1:A:278:ARG:NH1	0.51	2.03	13	1
1:A:74:VAL:O	1:A:78:GLU:HG2	0.51	2.06	6	2
1:A:90:ILE:HG13	1:A:105:TYR:CD1	0.50	2.42	20	2
1:A:79:LEU:O	1:A:83:ARG:HG3	0.50	2.07	5	1
1:A:251:GLU:O	1:A:254:ILE:HB	0.50	2.07	13	2
1:A:96:PRO:HA	1:A:132:TYR:CD1	0.50	2.41	7	5
1:A:230:SER:HA	1:A:232:TYR:CE1	0.50	2.41	9	1
1:A:107:ARG:O	1:A:110:ASP:HB3	0.50	2.07	10	3
1:A:152:ILE:O	1:A:156:GLU:HB2	0.49	2.06	13	1
1:A:130:LYS:HE2	1:A:140:GLY:O	0.49	2.08	8	1
1:A:272:ALA:O	1:A:276:GLU:HG2	0.49	2.06	10	1
1:A:234:THR:HA	1:A:242:MET:HE2	0.49	1.84	1	1
1:A:144:GLU:O	1:A:148:ASP:HB2	0.48	2.08	8	1
1:A:205:MET:HG2	1:A:281:MET:CG	0.48	2.38	10	1
1:A:107:ARG:HA	1:A:110:ASP:OD2	0.48	2.08	12	2
1:A:80:ARG:O	1:A:84:GLU:HG2	0.48	2.07	17	1
1:A:83:ARG:HD2	1:A:112:ASP:OD1	0.48	2.07	18	1
1:A:277:LEU:O	1:A:281:MET:HB2	0.48	2.09	12	3
1:A:186:LYS:O	1:A:189:GLN:HB3	0.48	2.09	1	3
1:A:107:ARG:NH1	1:A:283:ALA:HA	0.48	2.23	2	3
1:A:71:THR:O	1:A:75:LEU:HG	0.48	2.08	11	4
1:A:237:GLU:HB3	1:A:240:ALA:HB2	0.48	1.84	8	1
1:A:212:LEU:O	1:A:216:PHE:HD1	0.48	1.91	19	14
1:A:88:SER:O	1:A:92:LYS:HG2	0.48	2.07	14	1
1:A:277:LEU:O	1:A:281:MET:HG2	0.48	2.08	5	1
1:A:82:GLN:O	1:A:86:PHE:HB2	0.48	2.08	12	2
1:A:75:LEU:O	1:A:79:LEU:HG	0.48	2.09	4	8
1:A:99:ILE:HB	1:A:181:PHE:CG	0.48	2.43	18	5
1:A:116:LYS:HE3	1:A:255:HIS:CE1	0.47	2.44	10	1
1:A:128:LYS:HE2	1:A:128:LYS:HA	0.47	1.85	20	3
1:A:267:LEU:HD12	1:A:268:LYS:N	0.47	2.24	6	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:147:LEU:O	1:A:151:MET:HG2	0.47	2.09	14	4
1:A:216:PHE:CZ	1:A:274:LEU:HD12	0.47	2.45	20	1
1:A:146:ALA:O	1:A:149:ASN:HB3	0.47	2.10	11	2
1:A:189:GLN:HB3	1:A:238:TYR:CD1	0.47	2.45	2	2
1:A:107:ARG:NH1	1:A:111:THR:HB	0.47	2.25	3	1
1:A:86:PHE:CD2	1:A:108:VAL:HG21	0.47	2.45	4	1
1:A:281:MET:O	1:A:285:LEU:HB2	0.47	2.10	11	1
1:A:112:ASP:OD1	1:A:117:VAL:HB	0.46	2.09	19	2
1:A:142:SER:HB2	1:A:144:GLU:OE1	0.46	2.10	2	1
1:A:107:ARG:CZ	1:A:111:THR:HB	0.46	2.40	3	1
1:A:126:GLY:O	1:A:129:VAL:HB	0.46	2.10	4	1
1:A:133:LEU:O	1:A:136:ILE:HG12	0.46	2.10	9	2
1:A:282:MET:O	1:A:286:TYR:HD1	0.45	1.94	20	1
1:A:175:LYS:O	1:A:179:LEU:HD23	0.45	2.10	3	1
1:A:142:SER:HB3	1:A:144:GLU:OE2	0.45	2.11	7	1
1:A:128:LYS:HA	1:A:128:LYS:HE2	0.45	1.87	19	1
1:A:83:ARG:HD3	1:A:84:GLU:N	0.45	2.27	20	1
1:A:171:SER:H	1:A:174:LEU:HD22	0.45	1.72	16	1
1:A:78:GLU:O	1:A:82:GLN:HG2	0.45	2.12	12	1
1:A:227:GLU:HA	1:A:230:SER:OG	0.45	2.12	19	2
1:A:119:SER:O	1:A:122:VAL:HG12	0.45	2.13	10	2
1:A:202:LEU:HA	1:A:205:MET:SD	0.45	2.52	10	1
1:A:186:LYS:O	1:A:190:GLU:HG2	0.44	2.11	15	2
1:A:247:TRP:HE1	1:A:278:ARG:NH2	0.44	2.10	5	1
1:A:150:LEU:O	1:A:154:VAL:HG23	0.44	2.12	18	6
1:A:171:SER:O	1:A:175:LYS:HG3	0.44	2.13	8	1
1:A:278:ARG:HH11	1:A:278:ARG:HG3	0.44	1.72	19	1
1:A:107:ARG:HD2	1:A:282:MET:HB3	0.44	1.89	16	1
1:A:116:LYS:HE2	1:A:255:HIS:NE2	0.44	2.27	19	1
1:A:91:ILE:O	1:A:128:LYS:HE3	0.44	2.12	4	1
1:A:147:LEU:O	1:A:151:MET:HB2	0.44	2.13	11	1
1:A:99:ILE:HB	1:A:181:PHE:CD2	0.43	2.48	4	1
1:A:225:THR:HB	1:A:229:ILE:CD1	0.43	2.42	10	1
1:A:83:ARG:NH2	1:A:87:LEU:HD22	0.43	2.29	15	1
1:A:118:ASP:HB2	1:A:121:LYS:HG2	0.43	1.89	6	1
1:A:247:TRP:CE3	1:A:247:TRP:HA	0.43	2.47	9	2
1:A:153:LYS:HD3	1:A:156:GLU:OE1	0.43	2.14	4	2
1:A:132:TYR:O	1:A:135:LYS:HB3	0.43	2.13	8	1
1:A:211:GLU:O	1:A:214:LYS:HB3	0.43	2.14	12	3
1:A:111:THR:O	1:A:114:LEU:HD12	0.43	2.13	1	1
1:A:211:GLU:OE1	1:A:215:LYS:HE2	0.43	2.13	2	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:283:ALA:O	1:A:287:GLY:HA3	0.43	2.14	11	1
1:A:171:SER:O	1:A:175:LYS:HG2	0.43	2.14	11	1
1:A:115:LEU:HD23	1:A:115:LEU:O	0.42	2.14	5	1
1:A:118:ASP:CG	1:A:121:LYS:HG2	0.42	2.34	7	1
1:A:85:ALA:O	1:A:89:GLU:HG3	0.42	2.14	11	1
1:A:202:LEU:O	1:A:205:MET:HG2	0.42	2.13	19	1
1:A:233:LEU:O	1:A:242:MET:HG2	0.42	2.14	1	1
1:A:225:THR:OG1	1:A:229:ILE:HG12	0.42	2.14	1	2
1:A:269:ALA:O	1:A:273:GLU:HG2	0.42	2.15	10	1
1:A:99:ILE:HG21	1:A:185:LEU:HB2	0.42	1.92	2	1
1:A:192:HIS:CD2	1:A:196:LYS:HD2	0.42	2.49	10	1
1:A:209:HIS:CD2	1:A:278:ARG:HG3	0.42	2.50	2	1
1:A:232:TYR:HH	1:A:247:TRP:HE3	0.42	1.56	13	1
1:A:151:MET:CE	1:A:151:MET:HA	0.42	2.44	18	1
1:A:189:GLN:OE1	1:A:189:GLN:HA	0.42	2.14	20	1
1:A:80:ARG:O	1:A:84:GLU:HB2	0.42	2.15	12	1
1:A:185:LEU:O	1:A:189:GLN:HG2	0.41	2.15	17	1
1:A:99:ILE:O	1:A:103:VAL:HG22	0.41	2.15	16	2
1:A:72:LEU:O	1:A:76:ARG:HG3	0.41	2.15	13	1
1:A:206:LYS:HB3	1:A:206:LYS:NZ	0.41	2.30	5	1
1:A:143:VAL:O	1:A:147:LEU:HD13	0.41	2.15	4	1
1:A:148:ASP:O	1:A:152:ILE:HG13	0.41	2.16	12	1
1:A:171:SER:HB3	1:A:174:LEU:HD23	0.41	1.93	20	1
1:A:128:LYS:HA	1:A:128:LYS:CE	0.41	2.46	19	1
1:A:256:GLY:O	1:A:260:SER:HB2	0.41	2.16	6	1
1:A:170:ASP:HA	1:A:174:LEU:HD12	0.41	1.93	7	1
1:A:196:LYS:O	1:A:196:LYS:HD3	0.41	2.15	4	1
1:A:226:SER:O	1:A:229:ILE:HG13	0.41	2.15	7	1
1:A:86:PHE:O	1:A:89:GLU:HB3	0.41	2.16	19	1
1:A:114:LEU:HD13	1:A:275:ILE:HG12	0.41	1.92	7	1
1:A:138:ILE:HD11	1:A:174:LEU:HD22	0.41	1.93	9	1
1:A:76:ARG:O	1:A:80:ARG:HG3	0.41	2.16	15	1
1:A:76:ARG:HB2	1:A:276:GLU:OE2	0.41	2.15	1	1
1:A:234:THR:HA	1:A:242:MET:CE	0.41	2.46	20	1
1:A:189:GLN:HG2	1:A:238:TYR:CD1	0.40	2.51	5	1
1:A:171:SER:HB3	1:A:173:GLU:OE2	0.40	2.16	10	1
1:A:124:GLU:O	1:A:128:LYS:HG2	0.40	2.16	11	1
1:A:107:ARG:HD2	1:A:282:MET:CB	0.40	2.46	16	1
1:A:115:LEU:O	1:A:115:LEU:HD13	0.40	2.16	2	1
1:A:245:ALA:O	1:A:249:LYS:HG2	0.40	2.16	9	1
1:A:170:ASP:OD1	1:A:175:LYS:HG2	0.40	2.16	6	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:187:ALA:O	1:A:191:LEU:HG	0.40	2.16	9	2
1:A:74:VAL:O	1:A:78:GLU:HG3	0.40	2.17	15	1
1:A:142:SER:HB3	1:A:144:GLU:OE1	0.40	2.16	4	1
1:A:119:SER:O	1:A:123:ASP:HB2	0.40	2.16	17	1
1:A:112:ASP:HA	1:A:117:VAL:CG2	0.40	2.46	19	1
1:A:209:HIS:O	1:A:213:LEU:HG	0.40	2.17	5	1
1:A:209:HIS:CE1	1:A:234:THR:HB	0.40	2.52	13	1
1:A:155:TYR:CD1	1:A:241:PRO:HA	0.40	2.51	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	207/224 (92%)	197±3 (95±1%)	9±2 (4±1%)	1±1 (1±1%)	29	74
All	All	4140/4480 (92%)	3931 (95%)	186 (4%)	23 (1%)	29	74

All 11 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	118	ASP	5
1	A	94	ASP	4
1	A	264	PRO	3
1	A	197	GLY	2
1	A	236	PRO	2
1	A	227	GLU	2
1	A	230	SER	1
1	A	241	PRO	1
1	A	226	SER	1
1	A	242	MET	1
1	A	265	ALA	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	174/187 (93%)	167±2 (96±1%)	7±2 (4±1%)	36	84
All	All	3480/3740 (93%)	3347 (96%)	133 (4%)	36	84

All 53 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	282	MET	14
1	A	102	LEU	8
1	A	199	GLU	8
1	A	242	MET	7
1	A	144	GLU	7
1	A	112	ASP	6
1	A	83	ARG	5
1	A	151	MET	4
1	A	276	GLU	4
1	A	281	MET	4
1	A	114	LEU	3
1	A	227	GLU	3
1	A	189	GLN	3
1	A	258	LEU	3
1	A	225	THR	2
1	A	80	ARG	2
1	A	139	ARG	2
1	A	173	GLU	2
1	A	107	ARG	2
1	A	215	LYS	2
1	A	86	PHE	2
1	A	110	ASP	2
1	A	115	LEU	2
1	A	188	GLU	2
1	A	206	LYS	2
1	A	111	THR	2
1	A	179	LEU	2
1	A	278	ARG	2
1	A	120	THR	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	72	LEU	1
1	A	209	HIS	1
1	A	101	GLN	1
1	A	122	VAL	1
1	A	246	LYS	1
1	A	251	GLU	1
1	A	76	ARG	1
1	A	191	LEU	1
1	A	228	ASP	1
1	A	84	GLU	1
1	A	149	ASN	1
1	A	94	ASP	1
1	A	98	THR	1
1	A	82	GLN	1
1	A	285	LEU	1
1	A	130	LYS	1
1	A	186	LYS	1
1	A	177	LEU	1
1	A	243	ASN	1
1	A	128	LYS	1
1	A	220	THR	1
1	A	100	LEU	1
1	A	268	LYS	1
1	A	274	LEU	1

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)

There are no ligands in this entry.

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)

The completeness of assignment taking into account all chemical shift lists is 81% for the well-defined parts and 79% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	2411
Number of shifts mapped to atoms	2411
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	218	-0.70 ± 0.06	Should be checked
$^{13}C_{\beta}$	200	-0.21 ± 0.06	None needed (< 0.5 ppm)
$^{13}C'$	213	-0.49 ± 0.05	None needed (< 0.5 ppm)
^{15}N	204	0.61 ± 0.21	Should be applied

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 81%, i.e. 2284 atoms were assigned a chemical shift out of a possible 2837. 0 out of 46 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	1000/1039 (96%)	402/423 (95%)	405/414 (98%)	193/202 (96%)
Sidechain	1224/1661 (74%)	973/1087 (90%)	248/528 (47%)	3/46 (7%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	60/137 (44%)	59/67 (88%)	0/66 (0%)	1/4 (25%)
Overall	2284/2837 (81%)	1434/1577 (91%)	653/1008 (65%)	197/252 (78%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 79%, i.e. 2410 atoms were assigned a chemical shift out of a possible 3037. 0 out of 48 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	1061/1125 (94%)	426/459 (93%)	431/448 (96%)	204/218 (94%)
Sidechain	1289/1761 (73%)	1024/1153 (89%)	262/560 (47%)	3/48 (6%)
Aromatic	60/151 (40%)	59/75 (79%)	0/70 (0%)	1/6 (17%)
Overall	2410/3037 (79%)	1509/1687 (89%)	693/1078 (64%)	208/272 (76%)

7.1.4 Statistically unusual chemical shifts (i)

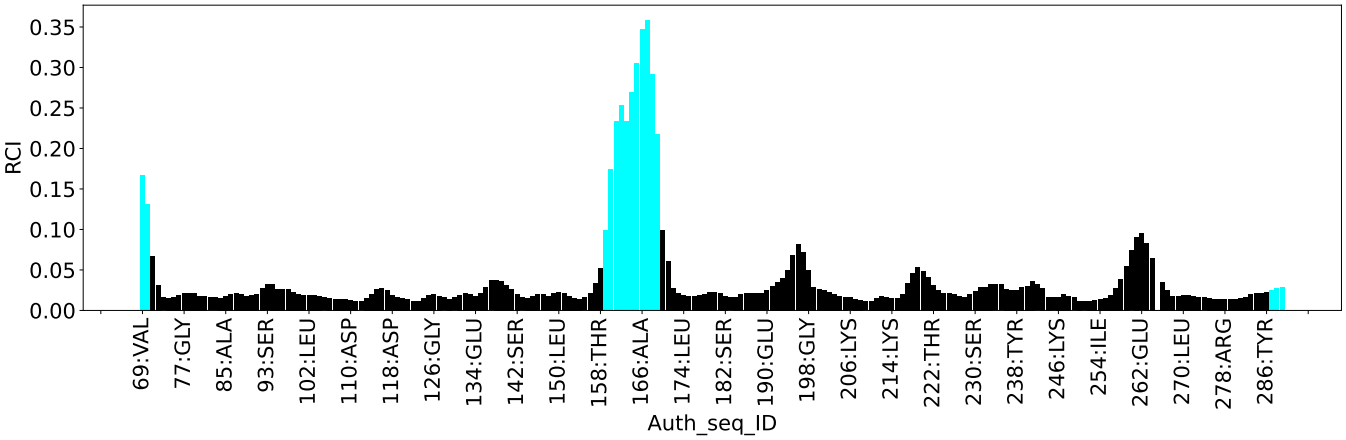
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	244	ALA	H	12.11	5.31 – 11.08	6.8
1	A	95	GLY	HA2	1.70	2.15 – 5.77	-6.2
1	A	285	LEU	HB2	-0.21	-0.07 – 3.30	-5.4

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:


RCI
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
69:VAL
77:GLY
85:ALA
93:SER
102:LEU
110:ASP
118:ASP
126:GLY
134:GLU
142:SER
150:LEU
158:THR
166:ALA
174:LEU
182:SER
190:GLU
198:GLY
206:LYS
214:LYS
222:THR
230:SER
238:TYR
246:LYS
254:ILE
262:GLU
270:LEU
278:ARG
286:TYR
Auth_seq_ID