Full wwPDB NMR Structure Validation Report (i)

Oct 10, 2021 – 03:59 PM EDT

PDB ID	:	2K31
Title	:	Solution Structure of cGMP-binding GAF domain of Phosphodiesterase 5
Authors	:	Heikaus, C.C.; Stout, J.R.; Sekharan, M.R.; Eakin, C.M.; Rajagopal, P.; Brzovic, P.S.; Beavo, J.A.; Klevit, R.E.
Deposited on	:	2008-04-16

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

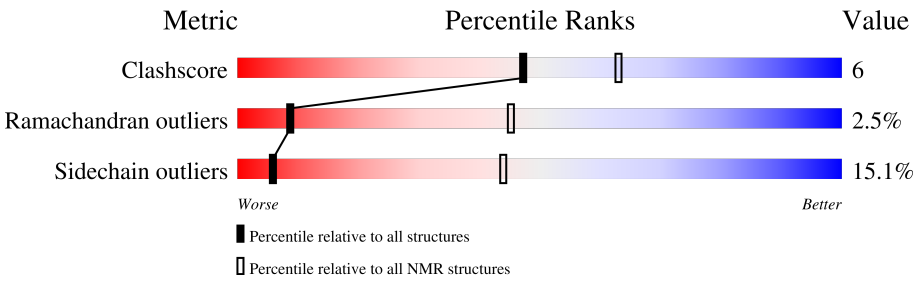
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.23.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 87%.

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	176	60% 15% • 8% 15%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:157-A:263, A:268-A:277, A:285-A:302 (135)	0.25	1

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 1 single-model cluster was found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2342 atoms, of which 1151 are hydrogens and 0 are deuteriums.

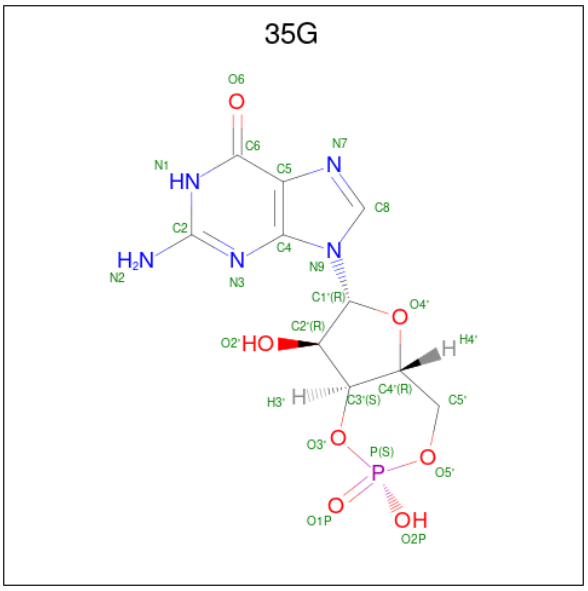
- Molecule 1 is a protein called Phosphodiesterase 5A, cGMP-specific.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	149	2308	738	1140	197	227	6	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	153	MET	-	initiating methionine	UNP Q0VBW0
A	295	GLU	ALA	engineered mutation	UNP Q0VBW0
A	302	GLU	ILE	engineered mutation	UNP Q0VBW0
A	321	LEU	-	expression tag	UNP Q0VBW0
A	322	GLU	-	expression tag	UNP Q0VBW0
A	323	HIS	-	expression tag	UNP Q0VBW0
A	324	HIS	-	expression tag	UNP Q0VBW0
A	325	HIS	-	expression tag	UNP Q0VBW0
A	326	HIS	-	expression tag	UNP Q0VBW0
A	327	HIS	-	expression tag	UNP Q0VBW0
A	328	HIS	-	expression tag	UNP Q0VBW0

- Molecule 2 is GUANOSINE-3',5'-MONOPHOSPHATE (three-letter code: 35G) (formula: $C_{10}H_{12}N_5O_7P$).



Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	P
2	A	1	34	10	11	5	7	1

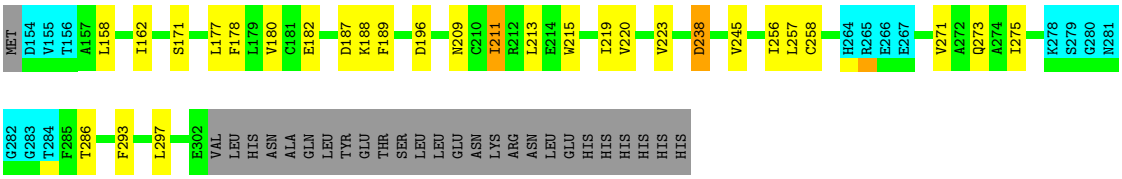
4 Residue-property plots (i)

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: Phosphodiesterase 5A, cGMP-specific

Chain A: 60% 15% • 8% 15%



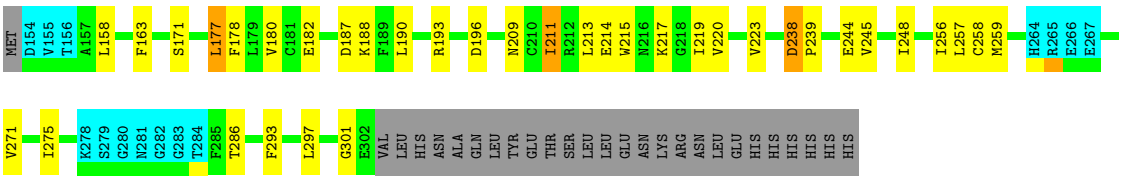
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

• Molecule 1: Phosphodiesterase 5A, cGMP-specific

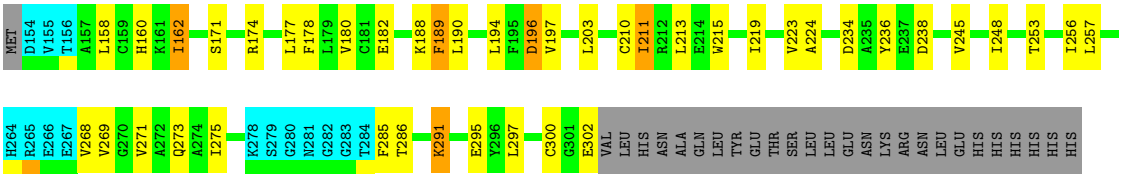
Chain A: 56% 19% • 8% 15%



4.2.2 Score per residue for model 2

• Molecule 1: Phosphodiesterase 5A, cGMP-specific

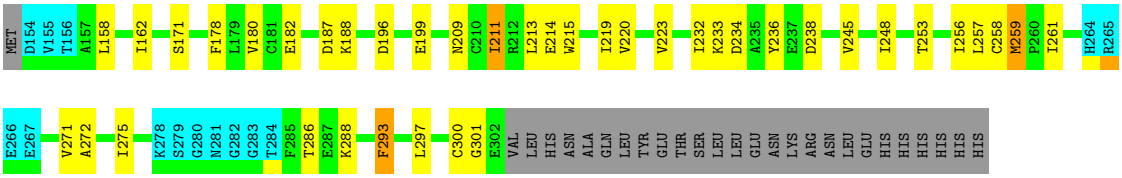
Chain A: 52% 22% • 8% 15%



4.2.3 Score per residue for model 3

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

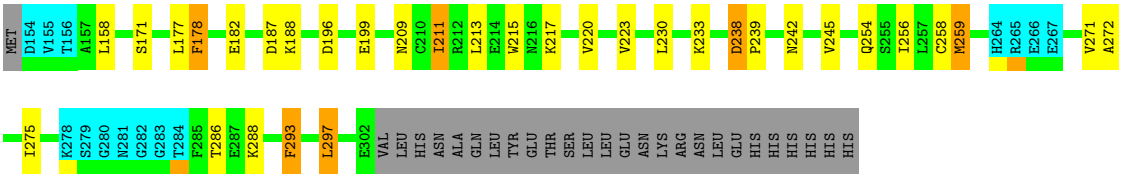
Chain A: 54% 21% • 8% 15%



4.2.4 Score per residue for model 4

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

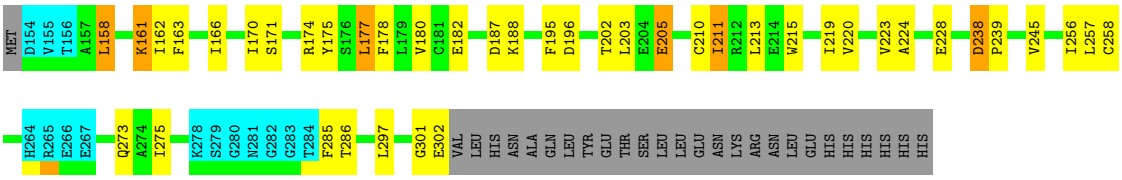
Chain A: 58% 15% • 8% 15%



4.2.5 Score per residue for model 5

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 53% 20% • 8% 15%



4.2.6 Score per residue for model 6

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 57% 17% • 8% 15%

MET D154 V155 T156 A157 L158 F163 S171 L177 F178 E182 D187 K188 F189 D196 N209 C210 I211 R212 L213 E214 W215 N216 K217 G218 I219 V220 V223 E228 Y236 E237 D238 V245 T253 I256 L257 C258 H264 R265 E266 E267 V268 V271 A272 Q273

A274 I275 K278 S279 G280 N281 G282 G283 T284 F285 T286 K291 L297 G301 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS HIS

4.2.7 Score per residue for model 7

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 56% 19% • 8% 15%

MET D154 V155 T156 A157 L158 C159 I162 F163 L164 S171 L177 F178 L179 V180 C181 E182 D187 K188 S192 T202 L203 E204 E205 N209 C210 I211 R212 L213 E214 W215 V220 V223 E228 D238 P239 V245 I256 L257 K262 N263 H264 R265 E266 E267

I275 N276 K277 K278 S279 G280 N281 G282 G283 T284 F285 T286 F293 L297 G301 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS HIS

4.2.8 Score per residue for model 8

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 56% 18% • 8% 15%

MET D154 V155 T156 A157 L158 C159 H160 K161 I162 F163 L164 H165 I166 I170 S171 Y175 S176 L177 F178 E182 D187 K188 F189 L194 F195 D196 V197 A206 N209 C210 I211 R212 L213 E214 W215 V220 V223 K233 D238 P239 V245 I256 L257 C258 K262

N263 H264 R265 E266 E267 Q273 A274 I275 K278 S279 G280 N281 G282 G283 T284 F285 T286 C300 G301 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS

4.2.9 Score per residue for model 9

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 60% 15% • 8% 15%

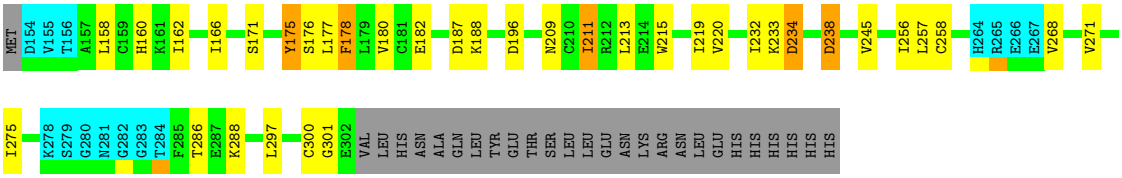
MET D154 V155 T156 A157 L158 C159 I162 S171 R174 Y175 F178 E182 D187 K188 D196 V197 A198 L203 N209 C210 I211 R212 L213 E214 W215 N216 K217 V220 V223 E228 R240 V245 I248 I256 L257 C258 H264 R265 E266 E267 I275

K278 S279 G280 N281 G282 G283 T284 F285 T286 C300 G301 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS HIS

4.2.10 Score per residue for model 10

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

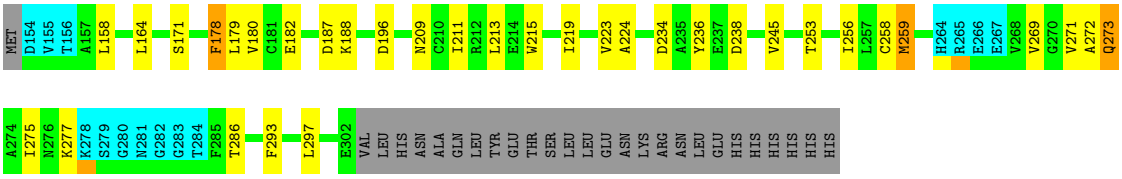
Chain A: 56% 18% • 8% 15%



4.2.11 Score per residue for model 11

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

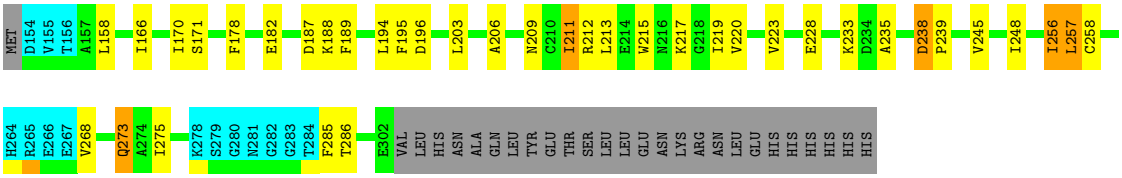
Chain A: 57% 18% • 8% 15%



4.2.12 Score per residue for model 12

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

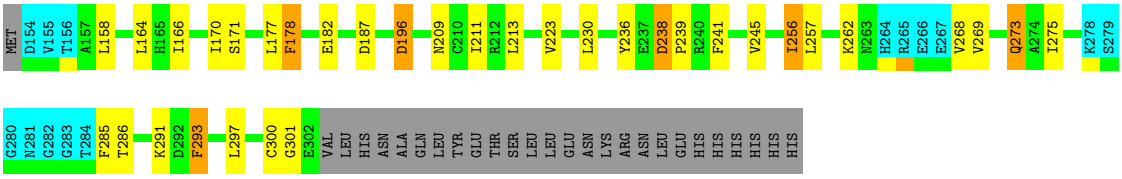
Chain A: 55% 19% • 8% 15%



4.2.13 Score per residue for model 13

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 57% 16% • 8% 15%



4.2.14 Score per residue for model 14

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 60% 15% • 8% 15%

MET D154 V155 T156 A157 L158 S171 S176 L177 F178 L179 V180 C181 E182 D187 K188 I191 D196 N209 C210 I211 R212 L213 E214 W215 V223 D234 D238 P239 V245 I248 I256 M259 H264 R265 E266 E267 V268 Q273 A274 I275 N276 K277 K278 S279

G280 N281 G282 G283 T284 F285 T286 E287 K291 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS HIS

4.2.15 Score per residue for model 15

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 56% 18% • 8% 15%

MET D154 V155 T156 A157 L158 C159 H160 K161 I162 I166 S171 A172 D173 R174 Y175 F178 E182 D187 K188 F189 L194 F195 D196 V197 A206 N209 C210 I211 R212 L213 E214 W215 V223 E228 P229 L230 D238 P239 R240 V245 I248 I256 L257 C258

K262 N263 H264 R265 E266 E267 I275 N276 K277 K278 S279 G280 N281 G282 G283 T284 F285 T286 L297 C300 G301 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS HIS

4.2.16 Score per residue for model 16

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 57% 18% • 8% 15%

MET D154 V155 T156 A157 L158 I162 F163 L164 S171 L177 F178 E182 D187 K188 F189 L190 T202 L203 E204 E205 C210 I211 R212 L213 E214 W215 I219 V220 V223 A224 E228 D234 D238 V245 I248 I256 L257 C258 H264 R265 E266 E267

V271 I275 K278 S279 G280 N281 G282 G283 T284 F285 T286 K291 C300 G301 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS HIS

4.2.17 Score per residue for model 17

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 58% 18% • 8% 15%

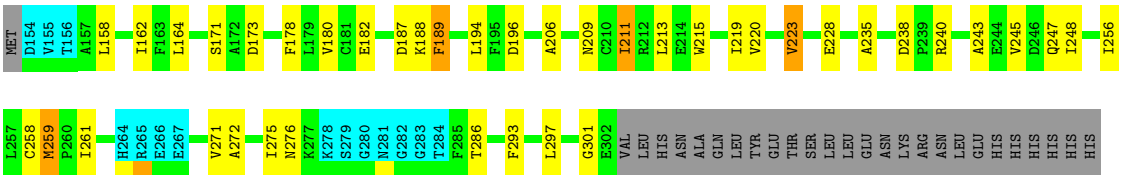
MET D154 V155 T156 A157 L158 S171 F178 L179 V180 C181 E182 D187 K188 I191 D196 N209 C210 I211 R212 L213 E214 W215 I219 V220 V223 D238 P239 N242 A243 E244 V245 I256 L257 C258 N263 H264 R265 E266 E267 V268 V269 A274 I275 K278 S279

G280 N281 G282 G283 T284 F285 T286 K291 E302 VAL LEU HIS ASN ALA GLN LEU TYR GLU THR SER LEU LEU GLU ASN LYS ARG ASN LEU GLU HIS HIS HIS HIS HIS HIS

4.2.18 Score per residue for model 18

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

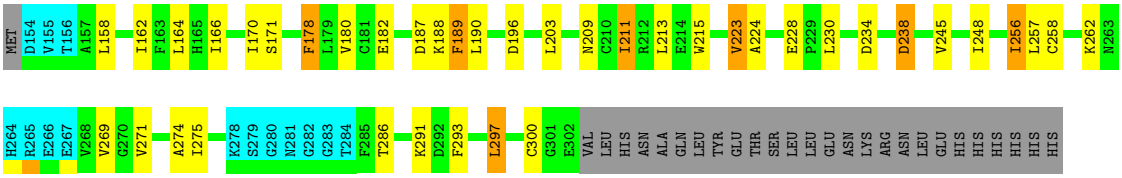
Chain A: 53% 21% • 8% 15%



4.2.19 Score per residue for model 19

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

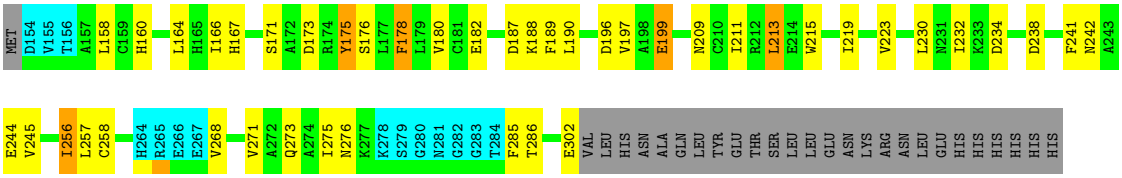
Chain A: 54% 19% • 8% 15%



4.2.20 Score per residue for model 20

- Molecule 1: Phosphodiesterase 5A, cGMP-specific

Chain A: 52% 22% • 8% 15%



5 Refinement protocol and experimental data overview (i)

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

Of the 100 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
CNS	structure solution	1.1
CYANA	structure solution	2.0
CYANA	refinement	2.0

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	1765
Number of shifts mapped to atoms	1765
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

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: 35G

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	1065	1047	1043	14±3
2	A	23	11	11	2±1
All	All	21760	21160	21080	275

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:273:GLN:HE22	2:A:1:35G:H2'	0.71	1.45	14	2
1:A:297:LEU:HA	1:A:300:CYS:SG	0.67	2.30	13	2
1:A:158:LEU:O	1:A:161:LYS:HG3	0.65	1.92	5	1
1:A:256:ILE:HG12	1:A:257:LEU:N	0.63	2.09	17	2
1:A:273:GLN:HE22	2:A:1:35G:H3'	0.62	1.55	8	2
1:A:256:ILE:HD11	1:A:273:GLN:HG2	0.61	1.73	14	2
1:A:162:ILE:HG21	1:A:300:CYS:SG	0.59	2.37	10	3
1:A:162:ILE:HD13	1:A:300:CYS:SG	0.59	2.36	10	3
1:A:219:ILE:HD12	1:A:258:CYS:SG	0.59	2.37	12	8
1:A:161:LYS:HD2	1:A:162:ILE:N	0.57	2.14	5	1
1:A:211:ILE:HG21	1:A:245:VAL:HG13	0.55	1.78	5	19

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:256:ILE:HG13	1:A:274:ALA:O	0.55	2.01	19	2
1:A:230:LEU:HB3	1:A:258:CYS:SG	0.54	2.42	20	4
1:A:191:ILE:HG23	1:A:212:ARG:HG2	0.54	1.79	17	2
1:A:162:ILE:HG12	1:A:300:CYS:SG	0.53	2.44	9	2
1:A:188:LYS:HE3	1:A:215:TRP:CD1	0.52	2.40	7	18
1:A:273:GLN:NE2	2:A:1:35G:H2'	0.51	2.20	14	2
1:A:178:PHE:CE2	1:A:211:ILE:HD11	0.50	2.41	11	1
1:A:223:VAL:HG21	1:A:258:CYS:SG	0.50	2.46	5	1
1:A:223:VAL:HG12	1:A:228:GLU:O	0.50	2.07	9	6
1:A:273:GLN:OE1	2:A:1:35G:H2'	0.50	2.06	12	1
1:A:245:VAL:HG21	2:A:1:35G:O4'	0.50	2.06	8	13
1:A:202:THR:OG1	1:A:205:GLU:HB2	0.50	2.07	7	3
1:A:173:ASP:N	1:A:276:ASN:HD22	0.50	2.05	20	2
1:A:219:ILE:HD11	1:A:273:GLN:NE2	0.49	2.22	20	2
1:A:161:LYS:HD2	1:A:161:LYS:C	0.49	2.28	5	1
1:A:188:LYS:HE3	1:A:215:TRP:NE1	0.49	2.22	15	15
1:A:219:ILE:HD11	1:A:273:GLN:HE22	0.49	1.67	5	2
1:A:194:LEU:HD21	1:A:203:LEU:HD12	0.49	1.83	12	1
1:A:258:CYS:HA	1:A:273:GLN:OE1	0.48	2.08	6	1
1:A:259:MET:SD	1:A:272:ALA:HB3	0.48	2.49	18	3
1:A:178:PHE:CE2	2:A:1:35G:H8	0.48	2.43	19	4
1:A:256:ILE:HD13	1:A:257:LEU:N	0.47	2.24	13	1
1:A:180:VAL:HG12	1:A:269:VAL:O	0.47	2.09	19	3
1:A:232:ILE:HG12	1:A:258:CYS:SG	0.47	2.50	10	2
1:A:257:LEU:HB2	1:A:285:PHE:CE2	0.47	2.45	20	2
1:A:194:LEU:HD11	1:A:203:LEU:HD11	0.47	1.86	2	1
1:A:220:VAL:HA	1:A:258:CYS:SG	0.47	2.49	8	3
1:A:235:ALA:HA	1:A:238:ASP:OD2	0.47	2.08	18	1
1:A:182:GLU:OE1	1:A:188:LYS:HE3	0.47	2.09	5	1
1:A:163:PHE:CE2	1:A:177:LEU:HD11	0.47	2.45	6	1
1:A:273:GLN:OE1	2:A:1:35G:H3'	0.47	2.10	2	2
1:A:257:LEU:HB2	1:A:285:PHE:CZ	0.47	2.45	5	3
1:A:166:ILE:HD12	1:A:175:TYR:CZ	0.46	2.46	5	5
1:A:234:ASP:O	1:A:238:ASP:HB2	0.46	2.09	16	7
1:A:259:MET:SD	1:A:297:LEU:HD11	0.46	2.51	3	1
1:A:238:ASP:OD2	1:A:239:PRO:HD2	0.46	2.11	5	10
1:A:293:PHE:O	1:A:297:LEU:HG	0.46	2.11	19	6
1:A:215:TRP:CH2	1:A:224:ALA:HB1	0.46	2.46	11	5
1:A:236:TYR:CD2	1:A:253:THR:HG22	0.46	2.46	6	4
1:A:194:LEU:HA	1:A:206:ALA:O	0.45	2.11	15	4
1:A:261:ILE:HD11	1:A:272:ALA:HB2	0.45	1.88	3	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:178:PHE:CD1	1:A:178:PHE:N	0.45	2.85	4	2
1:A:189:PHE:CD1	1:A:189:PHE:N	0.45	2.84	15	5
1:A:163:PHE:CZ	1:A:177:LEU:HD22	0.45	2.46	5	3
1:A:297:LEU:O	1:A:301:GLY:N	0.45	2.50	3	9
1:A:214:GLU:OE1	1:A:217:LYS:HG2	0.44	2.12	1	1
1:A:160:HIS:O	1:A:164:LEU:HD12	0.44	2.11	8	1
1:A:178:PHE:HE2	1:A:211:ILE:HD11	0.44	1.70	11	1
1:A:243:ALA:O	1:A:247:GLN:HG3	0.44	2.12	18	1
1:A:257:LEU:HD13	1:A:285:PHE:CG	0.44	2.47	20	1
1:A:173:ASP:HB2	1:A:276:ASN:OD1	0.44	2.12	18	1
1:A:263:ASN:HB2	1:A:269:VAL:HG13	0.44	1.90	17	1
1:A:167:HIS:CE1	1:A:199:GLU:HB2	0.43	2.48	20	1
1:A:257:LEU:HD13	1:A:285:PHE:CD1	0.43	2.49	2	1
1:A:178:PHE:CE1	1:A:211:ILE:HD11	0.43	2.49	6	1
1:A:160:HIS:NE2	1:A:197:VAL:HG13	0.43	2.28	8	2
1:A:257:LEU:HD23	1:A:285:PHE:CD1	0.43	2.48	12	1
1:A:232:ILE:HB	1:A:256:ILE:O	0.43	2.13	20	1
1:A:190:LEU:HD13	1:A:213:LEU:HD12	0.43	1.90	20	2
1:A:180:VAL:HG13	1:A:215:TRP:CZ3	0.43	2.49	10	9
1:A:195:PHE:CE1	1:A:211:ILE:HG12	0.43	2.48	5	3
1:A:174:ARG:NH1	1:A:198:ALA:HB2	0.43	2.28	9	1
1:A:235:ALA:HB2	1:A:256:ILE:HG21	0.43	1.90	12	1
1:A:233:LYS:O	1:A:254:GLN:O	0.43	2.37	4	1
1:A:245:VAL:HG11	2:A:1:35G:C1'	0.43	2.43	17	5
1:A:166:ILE:O	1:A:170:ILE:HG13	0.42	2.14	19	5
1:A:262:LYS:HA	1:A:269:VAL:HG22	0.42	1.91	13	1
1:A:160:HIS:CE1	1:A:197:VAL:HG13	0.42	2.50	20	2
1:A:223:VAL:HG13	1:A:228:GLU:O	0.42	2.15	16	3
1:A:287:GLU:O	1:A:291:LYS:HG2	0.42	2.15	14	1
1:A:219:ILE:O	1:A:223:VAL:HG23	0.41	2.16	2	1
1:A:159:CYS:O	1:A:162:ILE:HG13	0.41	2.16	7	1
1:A:263:ASN:OD1	1:A:302:GLU:HA	0.41	2.15	7	1
1:A:176:SER:OG	1:A:273:GLN:HB3	0.41	2.16	14	1
1:A:174:ARG:HD2	1:A:196:ASP:OD2	0.41	2.14	2	1
1:A:176:SER:CB	2:A:1:35G:N7	0.41	2.84	20	1
1:A:159:CYS:SG	1:A:300:CYS:SG	0.41	3.16	9	1
1:A:190:LEU:HD12	1:A:214:GLU:O	0.40	2.16	1	1
1:A:175:TYR:O	1:A:197:VAL:HG23	0.40	2.16	9	1
1:A:189:PHE:N	1:A:189:PHE:CD1	0.40	2.89	20	1
1:A:291:LYS:O	1:A:295:GLU:HG3	0.40	2.16	2	1
1:A:271:VAL:HG12	1:A:273:GLN:HE22	0.40	1.76	11	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:236:TYR:HA	1:A:241:PHE:CD2	0.40	2.52	13	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	134/176 (76%)	124±2 (92±2%)	7±2 (5±1%)	3±1 (2±1%)	9	45
All	All	2680/3520 (76%)	2477 (92%)	136 (5%)	67 (2%)	9	45

All 7 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	171	SER	20
1	A	209	ASN	17
1	A	196	ASP	15
1	A	268	VAL	8
1	A	210	CYS	3
1	A	285	PHE	3
1	A	234	ASP	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	115/152 (76%)	98±2 (85±2%)	17±2 (15±2%)	6	44
All	All	2300/3040 (76%)	1952 (85%)	348 (15%)	6	44

All 51 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	158	LEU	20
1	A	178	PHE	20
1	A	213	LEU	20
1	A	256	ILE	20
1	A	275	ILE	20
1	A	286	THR	20
1	A	182	GLU	19
1	A	187	ASP	19
1	A	223	VAL	15
1	A	211	ILE	14
1	A	238	ASP	14
1	A	220	VAL	12
1	A	248	ILE	10
1	A	271	VAL	10
1	A	177	LEU	9
1	A	164	LEU	8
1	A	189	PHE	7
1	A	257	LEU	6
1	A	259	MET	6
1	A	291	LYS	6
1	A	293	PHE	6
1	A	203	LEU	5
1	A	297	LEU	4
1	A	217	LYS	4
1	A	196	ASP	4
1	A	273	GLN	4
1	A	244	GLU	3
1	A	162	ILE	3
1	A	302	GLU	3
1	A	199	GLU	3
1	A	288	LYS	3
1	A	242	ASN	3
1	A	240	ARG	3
1	A	190	LEU	2
1	A	233	LYS	2
1	A	205	GLU	2
1	A	212	ARG	2
1	A	258	CYS	2
1	A	277	LYS	2
1	A	175	TYR	2
1	A	193	ARG	1
1	A	214	GLU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	161	LYS	1
1	A	174	ARG	1
1	A	192	SER	1
1	A	160	HIS	1
1	A	176	SER	1
1	A	179	LEU	1
1	A	230	LEU	1
1	A	234	ASP	1
1	A	241	PHE	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)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	35G	A	1	-	22,26,26	5.33±0.01	12±0 (54±1%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles

that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	35G	A	1	-	26,41,41	4.28±0.01	12±1 (45±1%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	35G	A	1	-	-	0±0,0,31,31	0±0,4,4,4

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
2	A	1	35G	C6-N1	13.63	1.56	1.33	13	20
2	A	1	35G	C4-N3	11.33	1.53	1.35	15	20
2	A	1	35G	P-O5'	9.29	1.68	1.57	15	20
2	A	1	35G	C6-C5	8.53	1.56	1.41	10	20
2	A	1	35G	C2-N1	7.96	1.49	1.35	13	20
2	A	1	35G	P-O3'	4.80	1.65	1.57	14	20
2	A	1	35G	O4'-C4'	4.35	1.35	1.45	7	20
2	A	1	35G	O3'-C3'	4.12	1.50	1.44	13	20
2	A	1	35G	C5'-C4'	4.11	1.58	1.51	4	20
2	A	1	35G	C2-N3	3.03	1.48	1.34	6	20
2	A	1	35G	C5-C4	2.95	1.48	1.40	17	20
2	A	1	35G	O5'-C5'	2.28	1.49	1.46	15	18

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
2	A	1	35G	N3-C2-N1	11.41	112.00	127.22	10	20
2	A	1	35G	C6-C5-C4	9.56	111.67	120.80	10	20
2	A	1	35G	C2-N3-C4	8.76	125.36	115.36	10	20
2	A	1	35G	O5'-P-O3'	7.13	95.85	105.68	15	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	A	1	35G	C6-N1-C2	5.41	124.53	115.93	14	20
2	A	1	35G	N2-C2-N1	4.74	124.62	117.25	3	20
2	A	1	35G	C1'-N9-C4	4.15	119.34	126.64	10	20
2	A	1	35G	O2P-P-O1P	4.13	121.68	108.73	4	20
2	A	1	35G	N2-C2-N3	3.48	123.46	117.79	13	20
2	A	1	35G	O4'-C4'-C5'	3.12	121.69	112.37	15	20
2	A	1	35G	O3'-C3'-C2'	2.73	118.28	115.61	14	19
2	A	1	35G	O3'-C3'-C4'	2.30	108.97	110.71	7	17

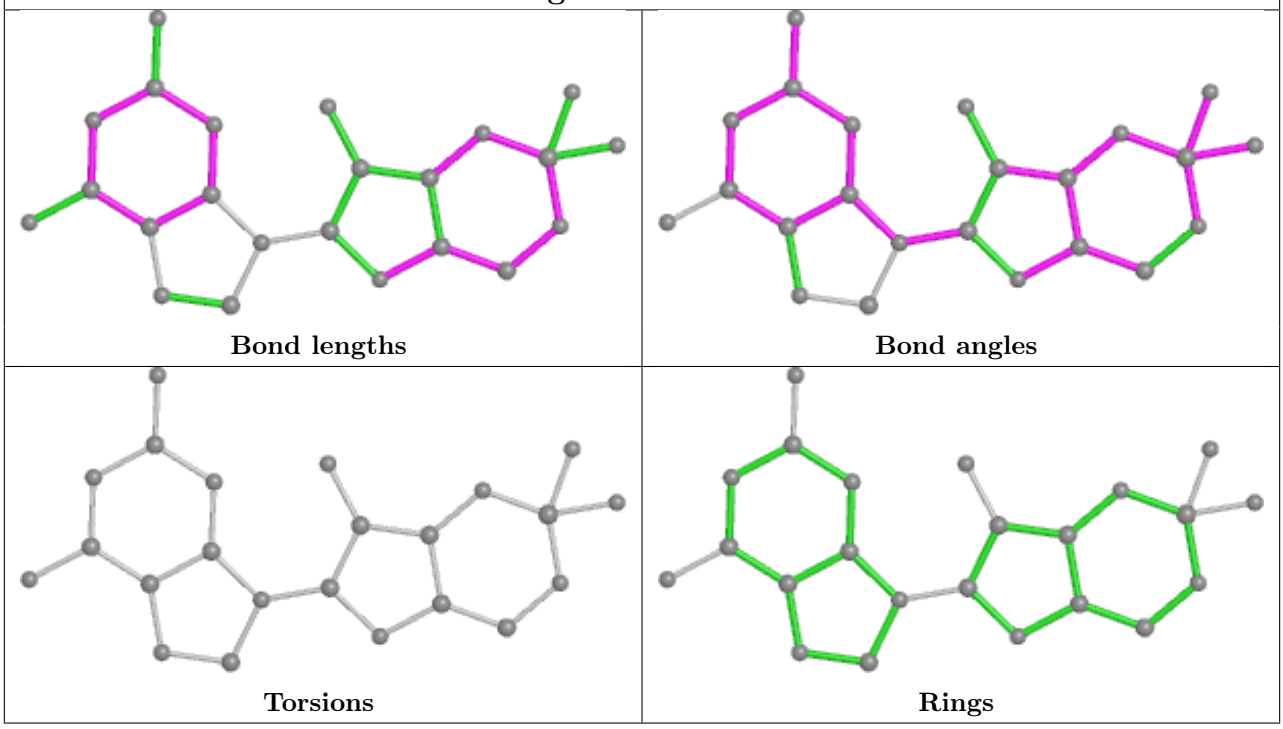
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand 35G A 1



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 87% for the well-defined parts and 86% 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	1765
Number of shifts mapped to atoms	1765
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	161	-0.11 ± 0.14	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	150	-0.12 ± 0.13	None needed (< 0.5 ppm)
$^{13}C'$	124	0.10 ± 0.12	None needed (< 0.5 ppm)
^{15}N	150	0.75 ± 0.31	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 87%, i.e. 1429 atoms were assigned a chemical shift out of a possible 1650. 19 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	636/669 (95%)	262/267 (98%)	245/270 (91%)	129/132 (98%)
Sidechain	683/824 (83%)	412/480 (86%)	271/312 (87%)	0/32 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	110/157 (70%)	58/83 (70%)	51/65 (78%)	1/9 (11%)
Overall	1429/1650 (87%)	732/830 (88%)	567/647 (88%)	130/173 (75%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 86%, i.e. 1548 atoms were assigned a chemical shift out of a possible 1804. 20 out of 21 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

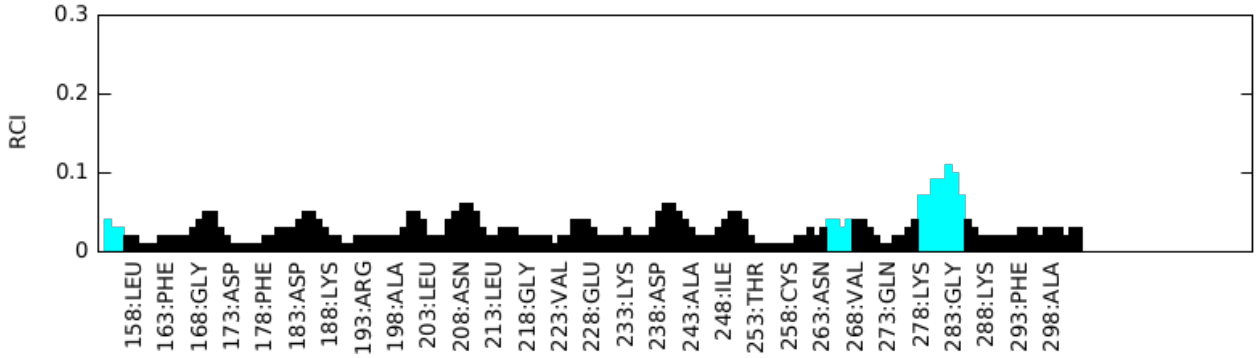
	Total	^{1}H	^{13}C	^{15}N
Backbone	698/739 (94%)	287/295 (97%)	270/298 (91%)	141/146 (97%)
Sidechain	737/900 (82%)	446/524 (85%)	291/339 (86%)	0/37 (0%)
Aromatic	113/165 (68%)	60/87 (69%)	52/67 (78%)	1/11 (9%)
Overall	1548/1804 (86%)	793/906 (88%)	613/704 (87%)	142/194 (73%)

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.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	188	LYS	HB2	0.52	3.03 – 0.53	-5.0

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.

Random coil index (RCI) for chain A: