Full wwPDB NMR Structure Validation Report (i)

Feb 26, 2022 – 03:50 PM EST

PDB ID	:	2BBY
Title	:	DNA-BINDING DOMAIN FROM HUMAN RAP30, NMR, 30 STRUCTURES
Authors	:	Groft, C.M.; Uljon, S.N.; Wang, R.; Werner, M.H.
Deposited on	:	1998-04-27

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:

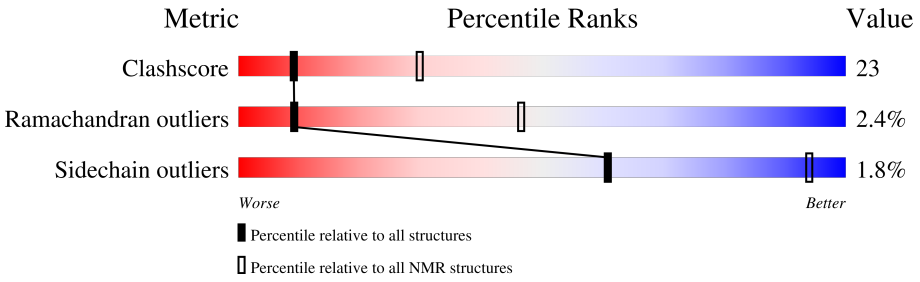
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)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

1 Overall quality at a glance (i)

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	69	54% 35% • 10%

2 Ensemble composition and analysis (i)

This entry contains 30 models. Model 4 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:179-A:240 (62)	0.14	4

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 2 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30
2	14, 17, 21
Single-model clusters	9

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1190 atoms, of which 599 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called RAP30.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	69	1190	383	599	104	103	1	0

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: RAP30

Chain A: 54% 35% • 10%

R175 A176 A176 R177 A178 D179 K180 Q181 H182 V183 F188 S189 A190 Y196 Y197 N198 V210 V211 L217 I220 G221 V222 Q223 N224 V225 K226 G227 I228 H229 K230 N231 T232 W233 E234 R240 H241 Y242 Q243

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: RAP30

Chain A: 54% 33% • 10%

R175 A176 A176 R177 A178 D179 V183 A190 K193 H194 Q195 Y196 Y197 N198 V210 V211 L217 I220 G221 V222 Q223 N224 V225 K226 G227 I228 N231 T232 W233 E234 L235 K236 R240 H241 Y242 Q243

4.2.2 Score per residue for model 2

• Molecule 1: RAP30

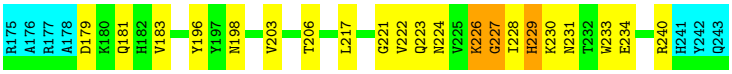
Chain A: 58% 32% 10%

R175 A176 A176 R177 A178 D179 V183 Y196 Q208 P209 V210 V211 L217 K218 E219 I220 G221 V222 Q223 N224 V225 K226 G227 I228 H229 T232 W233 E234 L235 K236 Y239 R240 H241 Y242 Q243

4.2.3 Score per residue for model 3

- Molecule 1: RAP30

Chain A: 59% 26% • 10%



4.2.4 Score per residue for model 4 (medoid)

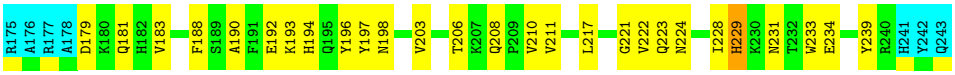
- Molecule 1: RAP30

Chain A: 51% 36% • 10%

4.2.5 Score per residue for model 5

- Molecule 1: RAP30

Chain A: 51% 38% • 10%



4.2.6 Score per residue for model 6

- Molecule 1: RAP30

Chain A: 49% 39% • 10%

4.2.7 Score per residue for model 7

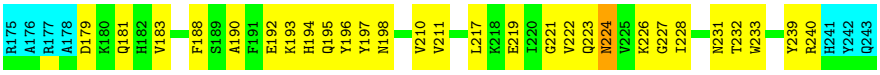
- Molecule 1: RAP30

Chain A: 42% 43% • 10%

4.2.8 Score per residue for model 8

- Molecule 1: RAP30

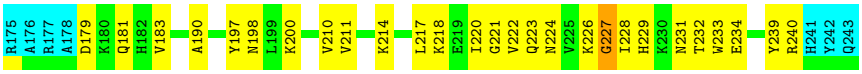
Chain A: 49% 39% • 10%



4.2.9 Score per residue for model 9

- Molecule 1: RAP30

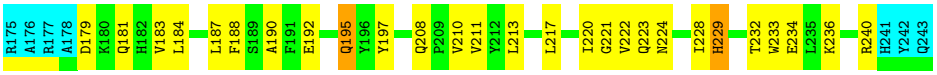
Chain A: 51% 38% • 10%



4.2.10 Score per residue for model 10

- Molecule 1: RAP30

Chain A: 51% 36% • 10%



4.2.11 Score per residue for model 11

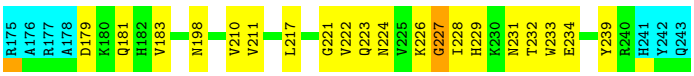
- Molecule 1: RAP30

Chain A: 48% 42% 10%

4.2.12 Score per residue for model 12

- Molecule 1: RAP30

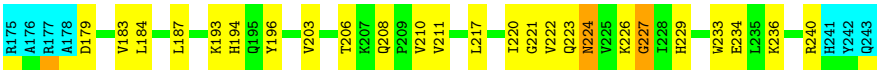
Chain A: 61% 28% • 10%



4.2.13 Score per residue for model 13

- Molecule 1: RAP30

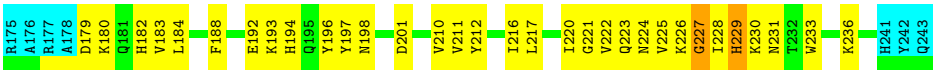
Chain A: 54% 33% • 10%



4.2.14 Score per residue for model 14

- Molecule 1: RAP30

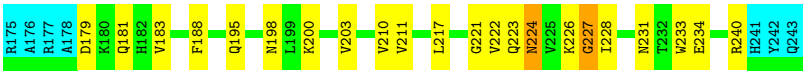
Chain A: 43% 43% • 10%



4.2.15 Score per residue for model 15

- Molecule 1: RAP30

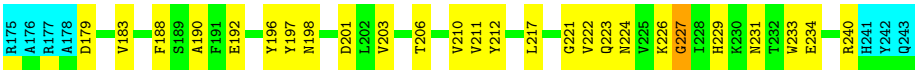
Chain A: 58% 29% • 10%



4.2.16 Score per residue for model 16

- Molecule 1: RAP30

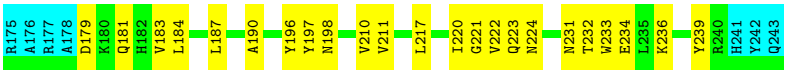
Chain A: 52% 36% • 10%



4.2.17 Score per residue for model 17

- Molecule 1: RAP30

Chain A: 57% 33% 10%



4.2.18 Score per residue for model 18

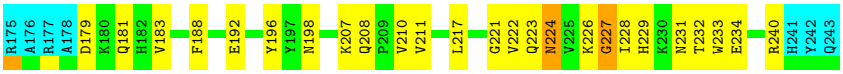
- Molecule 1: RAP30

Chain A: 58% 30% • 10%

4.2.19 Score per residue for model 19

- Molecule 1: RAP30

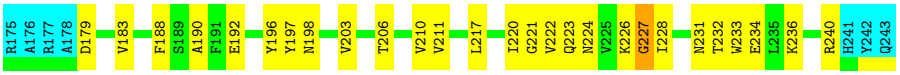
Chain A: 54% 33% • 10%



4.2.20 Score per residue for model 20

- Molecule 1: RAP30

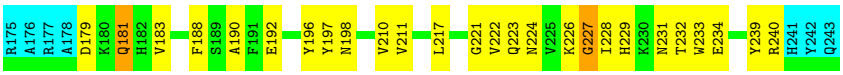
Chain A: 51% 38% • 10%



4.2.21 Score per residue for model 21

- Molecule 1: RAP30

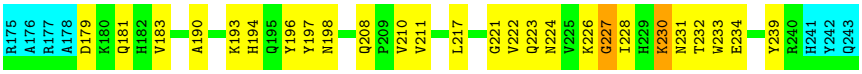
Chain A: 52% 35% • 10%



4.2.22 Score per residue for model 22

- Molecule 1: RAP30

Chain A: 52% 35% • 10%



4.2.23 Score per residue for model 23

- Molecule 1: RAP30

Chain A: 55% 32% • 10%

R175 A176 R177 A178 D179 K180 Q181 H182 V183 A190 K193 H194 Q195 Y196 Y197 N198 Q208 L217 K218 E219 I220 G221 V222 Q223 N224 V225 K226 G227 I228 H229 K230 N231 T232 W233 E234 Y239 R240 H241 Y242 Q243

4.2.24 Score per residue for model 24

- Molecule 1: RAP30

Chain A: 58% 30% • 10%

R175 A176 R177 A178 D179 V183 L184 L187 A190 K193 H194 Q195 Y196 Y197 V210 V211 Y212 L217 G221 V222 Q223 N224 G227 I228 H229 T232 W233 H241 Y242 Q243

4.2.25 Score per residue for model 25

- Molecule 1: RAP30

Chain A: 43% 42% • 10%

R175 A176 R177 A178 D179 K180 V183 L184 F188 S189 A190 F191 E192 K193 H194 Q195 Y196 Y197 N198 V210 V211 Y212 I216 L217 I220 G221 V222 Q223 N224 V225 K226 G227 I228 N231 T232 W233 E234 L235 K236 R240 H241 Y242 Q243

4.2.26 Score per residue for model 26

- Molecule 1: RAP30

Chain A: 54% 33% • 10%

R175 A176 R177 A178 D179 V183 M186 S189 Y196 Y197 N198 V210 V211 K214 L217 K218 E219 I220 G221 V222 Q223 N224 V225 K226 G227 I228 H229 K230 N231 T232 W233 E234 L235 K236 R240 H241 Y242 Q243

4.2.27 Score per residue for model 27

- Molecule 1: RAP30

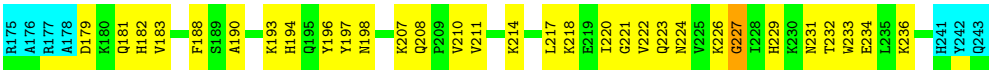
Chain A: 45% 43% • 10%

R175 A176 R177 A178 D179 K180 V183 F188 S189 A190 F191 E192 K193 H194 Q195 Y196 Y197 N198 V210 V211 Y212 I216 L217 I220 G221 V222 Q223 N224 V225 K226 G227 I228 H229 K230 N231 T232 W233 E234 L235 K236 Y239 R240 H241 Y242 Q243

4.2.28 Score per residue for model 28

- Molecule 1: RAP30

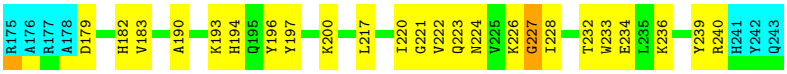
Chain A: 45% 43% • 10%



4.2.29 Score per residue for model 29

- Molecule 1: RAP30

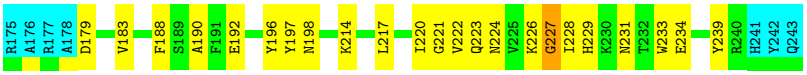
Chain A: 55% 33% • 10%



4.2.30 Score per residue for model 30

- Molecule 1: RAP30

Chain A: 57% 32% • 10%



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *DYNAMICAL SIMULATED ANNEALING*.

Of the 150 calculated structures, 30 were deposited, based on the following criterion: *LEAST RESTRAINT AND ANGULAR VIOLATION*.

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

Software name	Classification	Version
X-PLOR	refinement	3.851
X-PLOR	structure solution	3.851

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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	528	539	539	25±3
All	All	15840	16170	16170	737

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 23.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:179:ASP:O	1:A:183:VAL:HG23	0.76	1.81	21	28
1:A:223:GLN:HE21	1:A:224:ASN:H	0.69	1.31	12	27
1:A:223:GLN:NE2	1:A:224:ASN:H	0.68	1.85	12	30
1:A:217:LEU:O	1:A:221:GLY:N	0.65	2.28	29	30
1:A:223:GLN:HE22	1:A:232:THR:N	0.63	1.91	12	18
1:A:188:PHE:CZ	1:A:240:ARG:NH1	0.62	2.67	16	4
1:A:222:VAL:HG12	1:A:223:GLN:N	0.60	2.12	15	30
1:A:222:VAL:O	1:A:233:TRP:CE3	0.60	2.54	10	30
1:A:179:ASP:O	1:A:183:VAL:N	0.59	2.33	16	28
1:A:208:GLN:N	1:A:208:GLN:OE1	0.58	2.36	4	10
1:A:181:GLN:O	1:A:181:GLN:NE2	0.58	2.37	12	12
1:A:181:GLN:C	1:A:181:GLN:NE2	0.57	2.58	21	1
1:A:184:LEU:HD21	1:A:220:ILE:HD11	0.57	1.77	7	1
1:A:196:TYR:CE1	1:A:234:GLU:OE2	0.57	2.58	30	14

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:223:GLN:OE1	1:A:233:TRP:CE2	0.56	2.59	1	28
1:A:223:GLN:OE1	1:A:233:TRP:NE1	0.56	2.39	16	13
1:A:193:LYS:O	1:A:194:HIS:ND1	0.55	2.40	27	6
1:A:223:GLN:HE22	1:A:232:THR:H	0.55	1.44	19	9
1:A:196:TYR:CE2	1:A:234:GLU:CD	0.55	2.81	21	1
1:A:220:ILE:O	1:A:236:LYS:N	0.54	2.39	10	14
1:A:188:PHE:CE2	1:A:240:ARG:NH1	0.54	2.75	16	1
1:A:195:GLN:O	1:A:234:GLU:OE1	0.54	2.26	15	1
1:A:184:LEU:HD21	1:A:220:ILE:CD1	0.54	2.33	7	1
1:A:210:VAL:HG23	1:A:211:VAL:N	0.53	2.19	12	26
1:A:188:PHE:O	1:A:192:GLU:N	0.53	2.41	16	14
1:A:180:LYS:N	1:A:212:TYR:OH	0.53	2.41	27	3
1:A:190:ALA:O	1:A:197:TYR:CE1	0.53	2.61	21	21
1:A:224:ASN:HD21	1:A:234:GLU:CD	0.53	2.07	12	2
1:A:226:LYS:O	1:A:229:HIS:N	0.52	2.39	28	15
1:A:184:LEU:HD23	1:A:187:LEU:HD12	0.52	1.82	24	6
1:A:181:GLN:NE2	1:A:181:GLN:O	0.51	2.43	3	3
1:A:220:ILE:CD1	1:A:240:ARG:HH21	0.51	2.19	26	1
1:A:217:LEU:O	1:A:221:GLY:CA	0.51	2.59	1	29
1:A:224:ASN:ND2	1:A:234:GLU:OE2	0.50	2.44	9	2
1:A:224:ASN:ND2	1:A:234:GLU:CD	0.50	2.64	12	2
1:A:193:LYS:O	1:A:194:HIS:CD2	0.50	2.64	22	5
1:A:196:TYR:O	1:A:197:TYR:CD1	0.50	2.65	27	9
1:A:198:ASN:CB	1:A:231:ASN:O	0.50	2.60	9	25
1:A:193:LYS:O	1:A:194:HIS:CG	0.50	2.64	22	8
1:A:188:PHE:CE1	1:A:240:ARG:NH1	0.50	2.80	15	1
1:A:195:GLN:O	1:A:195:GLN:CD	0.50	2.50	8	6
1:A:196:TYR:CE1	1:A:234:GLU:CD	0.50	2.85	23	18
1:A:240:ARG:CG	1:A:240:ARG:HH11	0.50	2.20	25	3
1:A:179:ASP:O	1:A:183:VAL:CG2	0.49	2.59	9	25
1:A:195:GLN:O	1:A:195:GLN:NE2	0.49	2.44	18	2
1:A:219:GLU:N	1:A:219:GLU:OE1	0.49	2.46	7	1
1:A:222:VAL:HG12	1:A:223:GLN:H	0.49	1.68	12	11
1:A:228:ILE:C	1:A:229:HIS:CG	0.48	2.86	3	6
1:A:195:GLN:OE1	1:A:234:GLU:OE2	0.48	2.31	10	3
1:A:234:GLU:CD	1:A:234:GLU:N	0.48	2.66	12	2
1:A:222:VAL:CG1	1:A:223:GLN:N	0.48	2.76	9	18
1:A:195:GLN:OE1	1:A:234:GLU:OE1	0.48	2.31	18	2
1:A:233:TRP:O	1:A:234:GLU:OE2	0.48	2.32	15	1
1:A:227:GLY:O	1:A:229:HIS:N	0.48	2.45	9	1
1:A:226:LYS:O	1:A:227:GLY:C	0.48	2.52	22	24

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:190:ALA:O	1:A:197:TYR:CZ	0.48	2.67	21	5
1:A:240:ARG:CG	1:A:240:ARG:NH1	0.47	2.75	1	3
1:A:220:ILE:HG22	1:A:239:TYR:OH	0.47	2.10	27	5
1:A:228:ILE:O	1:A:229:HIS:CB	0.47	2.62	24	6
1:A:196:TYR:CE2	1:A:234:GLU:OE2	0.47	2.68	21	1
1:A:203:VAL:O	1:A:206:THR:O	0.46	2.33	3	6
1:A:223:GLN:HE21	1:A:224:ASN:N	0.46	2.08	26	5
1:A:223:GLN:NE2	1:A:232:THR:H	0.46	2.09	19	6
1:A:227:GLY:C	1:A:229:HIS:H	0.46	2.14	21	9
1:A:219:GLU:O	1:A:239:TYR:OH	0.46	2.34	2	3
1:A:214:LYS:O	1:A:218:LYS:N	0.45	2.45	6	5
1:A:228:ILE:O	1:A:229:HIS:CD2	0.45	2.69	24	2
1:A:213:LEU:HD22	1:A:213:LEU:N	0.45	2.26	10	1
1:A:220:ILE:CD1	1:A:240:ARG:NH2	0.45	2.80	9	2
1:A:215:GLU:O	1:A:219:GLU:OE2	0.45	2.35	7	1
1:A:210:VAL:CG2	1:A:211:VAL:N	0.44	2.79	25	10
1:A:200:LYS:O	1:A:203:VAL:HG22	0.44	2.12	15	1
1:A:212:TYR:CE2	1:A:216:ILE:HD11	0.44	2.48	27	2
1:A:240:ARG:HH11	1:A:240:ARG:CB	0.44	2.25	25	1
1:A:220:ILE:O	1:A:236:LYS:CB	0.44	2.66	20	3
1:A:227:GLY:O	1:A:229:HIS:CD2	0.44	2.70	3	1
1:A:233:TRP:O	1:A:234:GLU:CD	0.43	2.56	15	1
1:A:226:LYS:O	1:A:227:GLY:O	0.43	2.37	22	2
1:A:239:TYR:CD1	1:A:239:TYR:N	0.43	2.87	9	3
1:A:198:ASN:OD1	1:A:201:ASP:N	0.42	2.44	16	1
1:A:228:ILE:O	1:A:229:HIS:CG	0.42	2.71	24	2
1:A:227:GLY:C	1:A:229:HIS:N	0.42	2.72	14	1
1:A:211:VAL:HG23	1:A:212:TYR:N	0.42	2.30	16	2
1:A:198:ASN:OD1	1:A:201:ASP:CG	0.42	2.58	14	1
1:A:186:MET:O	1:A:189:SER:OG	0.42	2.38	26	1
1:A:207:LYS:N	1:A:208:GLN:OE1	0.41	2.54	11	2
1:A:220:ILE:HG21	1:A:240:ARG:CZ	0.41	2.46	6	1
1:A:229:HIS:N	1:A:229:HIS:CD2	0.41	2.88	12	1
1:A:184:LEU:HD21	1:A:216:ILE:CG2	0.41	2.45	25	2
1:A:183:VAL:HG22	1:A:208:GLN:HE22	0.41	1.76	2	1
1:A:224:ASN:O	1:A:230:LYS:O	0.41	2.39	22	1
1:A:214:LYS:HB3	1:A:218:LYS:HZ3	0.40	1.76	9	1
1:A:225:VAL:HG23	1:A:225:VAL:O	0.40	2.17	14	1
1:A:239:TYR:N	1:A:239:TYR:CD1	0.40	2.89	22	1
1:A:214:LYS:CB	1:A:214:LYS:NZ	0.40	2.85	30	1
1:A:196:TYR:CZ	1:A:234:GLU:OE2	0.40	2.75	11	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	62/69 (90%)	57±1 (92±1%)	3±1 (5±2%)	1±1 (2±1%)	9 46
All	All	1860/2070 (90%)	1715 (92%)	101 (5%)	44 (2%)	9 46

All 4 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	227	GLY	23
1	A	228	ILE	14
1	A	229	HIS	6
1	A	207	LYS	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	59/64 (92%)	58±1 (98±1%)	1±1 (2±1%)	61 94
All	All	1770/1920 (92%)	1738 (98%)	32 (2%)	61 94

All 9 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	240	ARG	12
1	A	224	ASN	7
1	A	230	LYS	4
1	A	200	LYS	3
1	A	226	LYS	2
1	A	208	GLN	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	219	GLU	1
1	A	195	GLN	1
1	A	181	GLN	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)

No chemical shift data were provided