



# Full wwPDB NMR Structure Validation Report ⓘ

Mar 6, 2026 – 09:12 PM UTC

PDB ID : 2JSD / pdb\_00002jsd  
Title : Solution structure of MMP20 complexed with NNGH  
Authors : Arendt, Y.; Banci, L.; Bertini, I.; Cantini, F.; Cozzi, R.; Del Conte, R.;  
Gonnelli, L.; Structural Proteomics in Europe (SPINE)  
Deposited on : 2007-07-03

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Buster-report : wwPDB partial adaption of 1.1.7 (2018)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
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) : 2.49

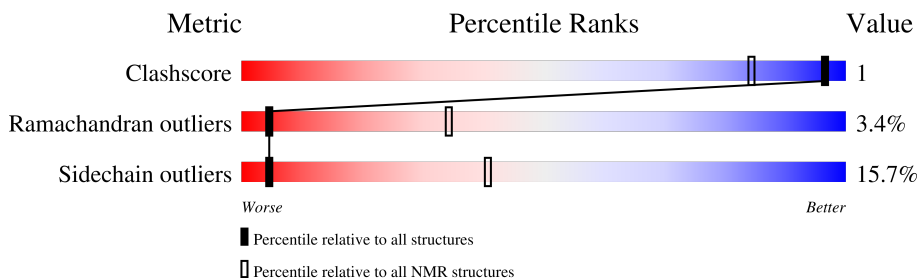
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	160	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA and RNA chains that are outliers for geometric criteria:

Mol	Chain	Compound	Res	Total models with violations	
				Chirality	Geometry
4	A	NGH	277	10	-

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 18 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:115-A:271 (157)	0.93	18

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. No single-model clusters were found.

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

### 3 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2446 atoms, of which 1185 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Matrix metalloproteinase-20.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	160	2402	790	1166	205	236	5	0

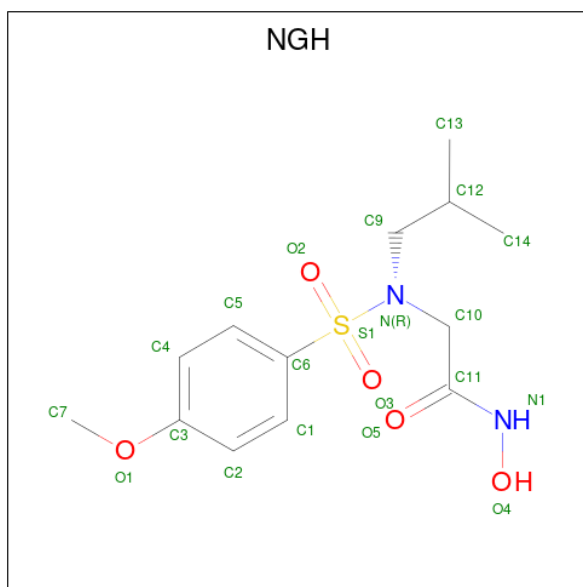
- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	
			Total	Ca
2	A	2	2	2

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
3	A	2	2	2

- Molecule 4 is N-ISOBUTYL-N-[4-METHOXYPHENYLSULFONYL]GLYCYL HYDROX-AMIC ACID (CCD ID: NGH) (formula: C<sub>13</sub>H<sub>20</sub>N<sub>2</sub>O<sub>5</sub>S).



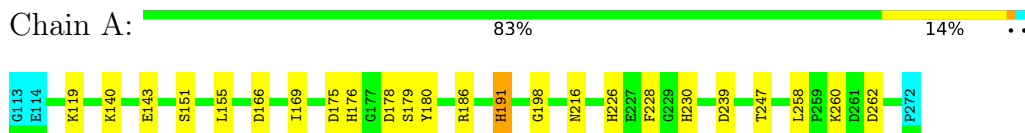
Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	S
4	A	1	40	13	19	2	5	1

## 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: Matrix metalloproteinase-20

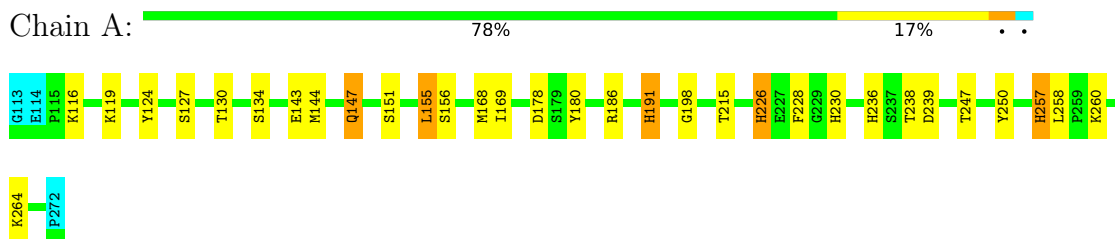


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

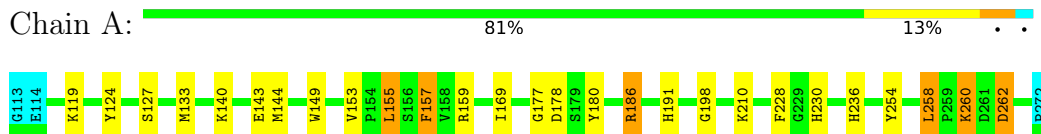
#### 4.2.1 Score per residue for model 1

- Molecule 1: Matrix metalloproteinase-20



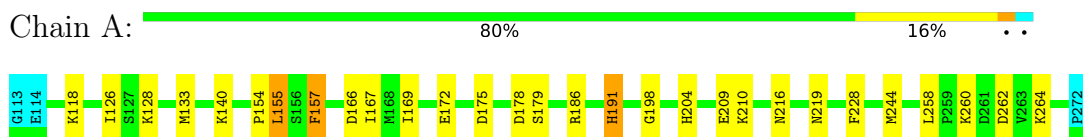
#### 4.2.2 Score per residue for model 2

- Molecule 1: Matrix metalloproteinase-20



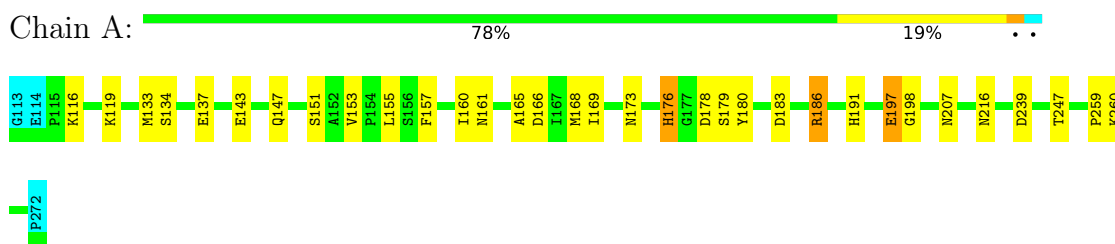
### 4.2.3 Score per residue for model 3

- Molecule 1: Matrix metalloproteinase-20



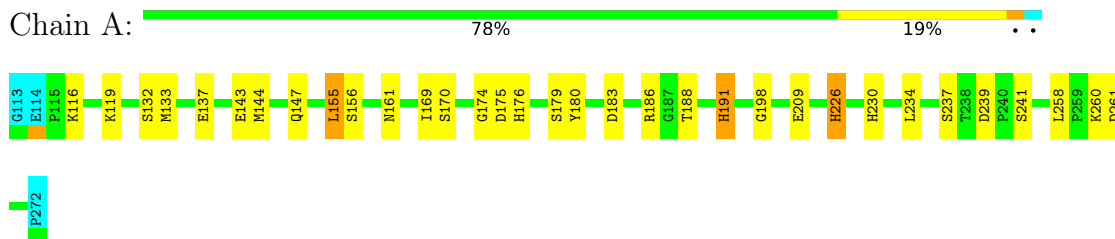
### 4.2.4 Score per residue for model 4

- Molecule 1: Matrix metalloproteinase-20



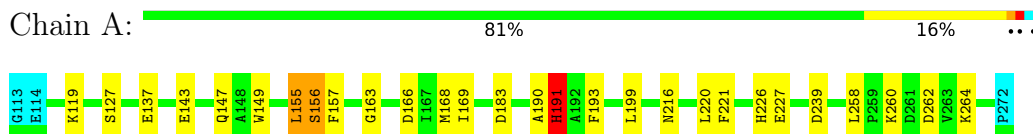
### 4.2.5 Score per residue for model 5

- Molecule 1: Matrix metalloproteinase-20



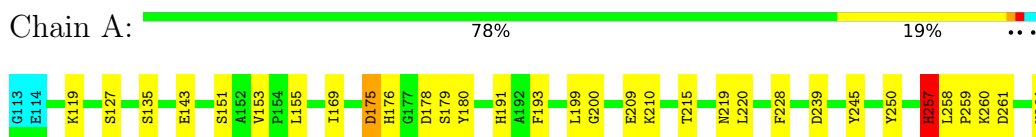
### 4.2.6 Score per residue for model 6

- Molecule 1: Matrix metalloproteinase-20



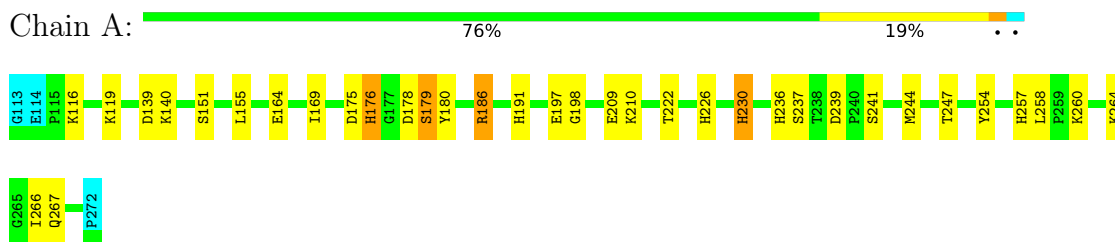
### 4.2.7 Score per residue for model 7

- Molecule 1: Matrix metalloproteinase-20



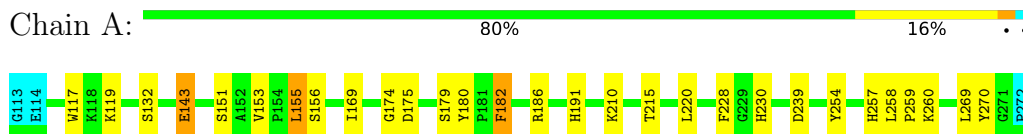
### 4.2.8 Score per residue for model 8

- Molecule 1: Matrix metalloproteinase-20



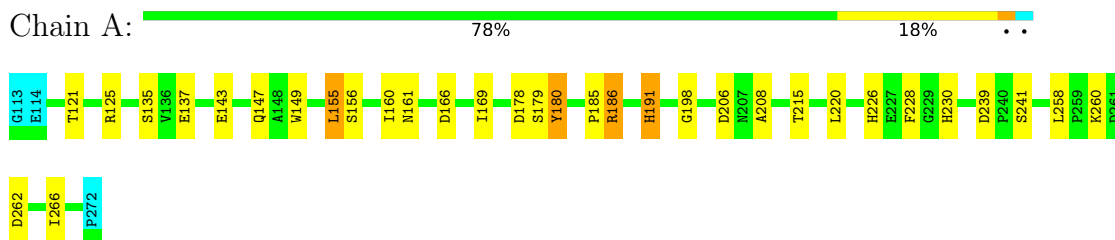
### 4.2.9 Score per residue for model 9

- Molecule 1: Matrix metalloproteinase-20



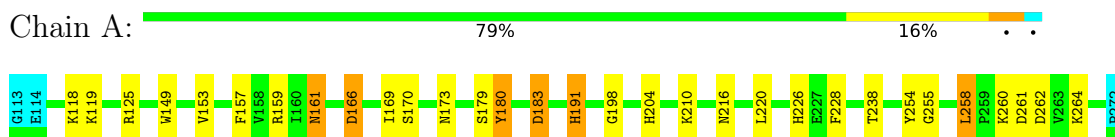
### 4.2.10 Score per residue for model 10

- Molecule 1: Matrix metalloproteinase-20



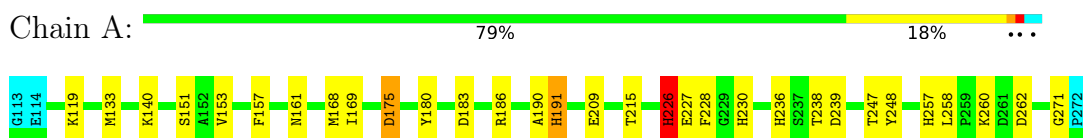
#### 4.2.11 Score per residue for model 11

- Molecule 1: Matrix metalloproteinase-20



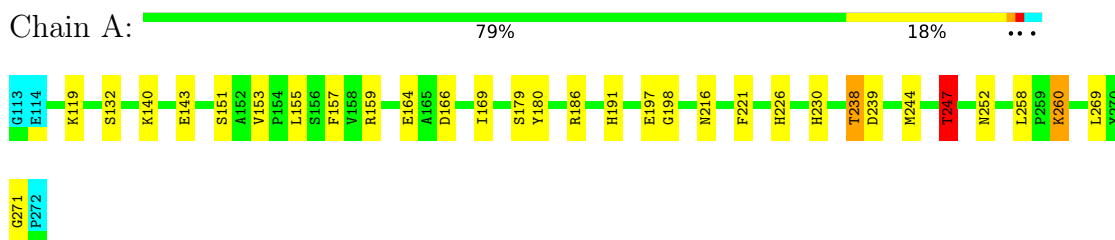
#### 4.2.12 Score per residue for model 12

- Molecule 1: Matrix metalloproteinase-20



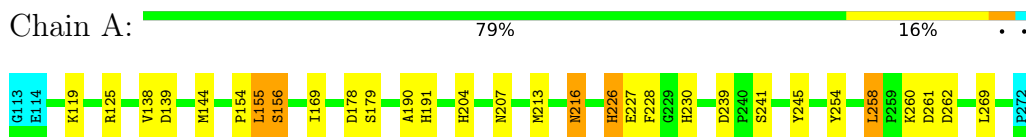
#### 4.2.13 Score per residue for model 13

- Molecule 1: Matrix metalloproteinase-20



#### 4.2.14 Score per residue for model 14

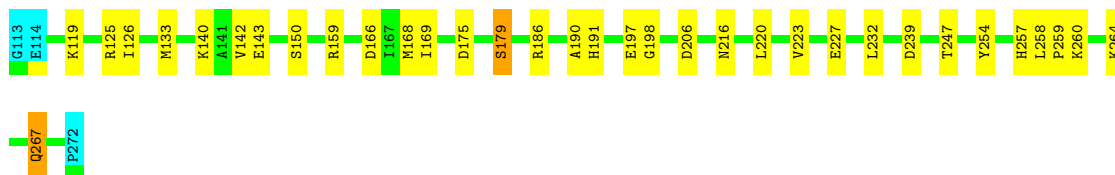
- Molecule 1: Matrix metalloproteinase-20



#### 4.2.15 Score per residue for model 15

- Molecule 1: Matrix metalloproteinase-20

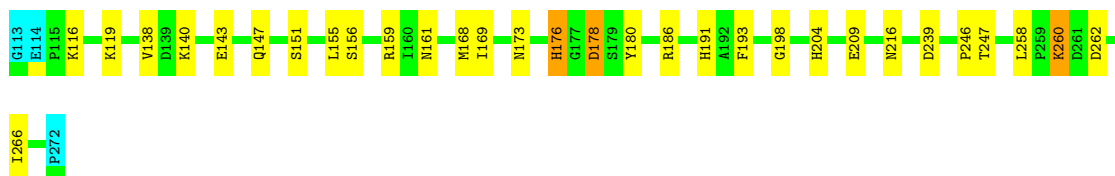




#### 4.2.16 Score per residue for model 16

- Molecule 1: Matrix metalloproteinase-20

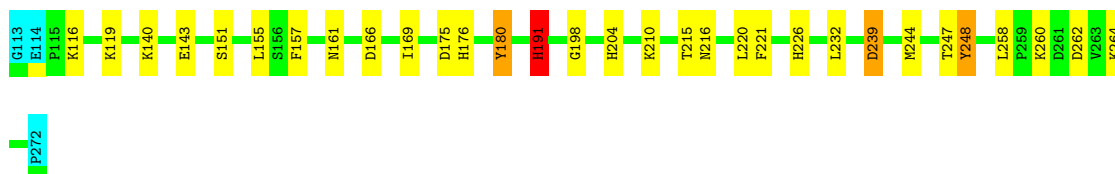
Chain A: 79% 18% ..



#### 4.2.17 Score per residue for model 17

- Molecule 1: Matrix metalloproteinase-20

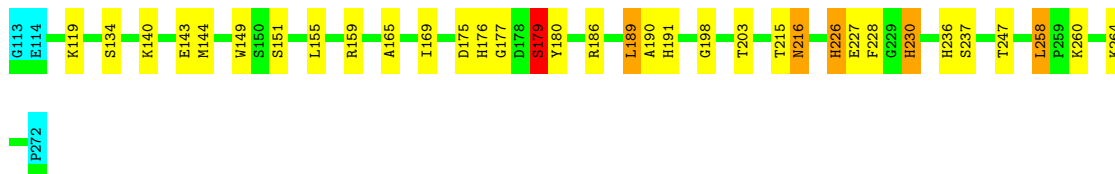
Chain A: 79% 17% ...



#### 4.2.18 Score per residue for model 18 (medoid)

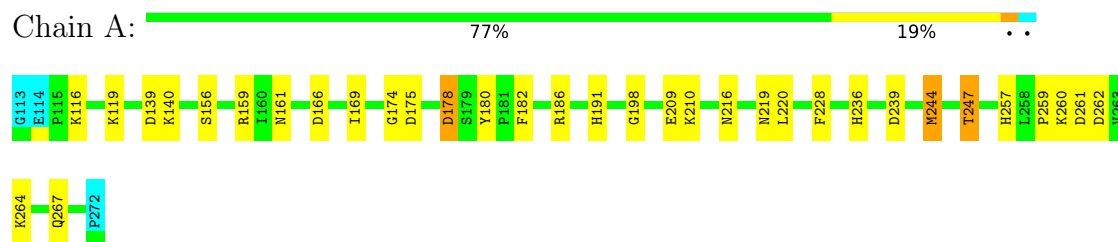
- Molecule 1: Matrix metalloproteinase-20

Chain A: 77% 18% ...



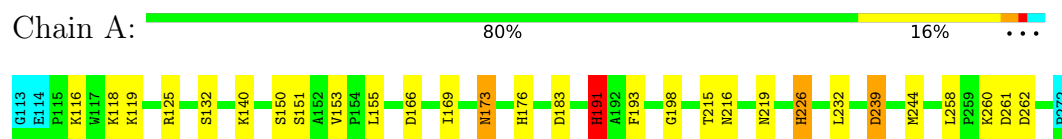
#### 4.2.19 Score per residue for model 19

- Molecule 1: Matrix metalloproteinase-20



#### 4.2.20 Score per residue for model 20

- Molecule 1: Matrix metalloproteinase-20



## 5 Refinement protocol and experimental data overview

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

Of the 400 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	structure solution	2.1
Amber	refinement	8.0
QUEEN	refinement	
AQUA	refinement	

No chemical shift data was provided.

## 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: CA, ZN, NGH

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	0.89±0.02	0±0/1255 ( 0.0± 0.0%)	1.54±0.03	8±3/1706 ( 0.5± 0.2%)
All	All	0.89	3/25100 ( 0.0%)	1.54	155/34120 ( 0.5%)

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	1.5±1.4
All	All	0	30

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
1	A	197	GLU	CA-C	5.91	1.55	1.52	8	1
1	A	178	ASP	CA-C	5.70	1.58	1.53	8	1
1	A	271	GLY	CA-C	5.45	1.55	1.51	13	1

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(°)	Ideal(°)	Models	
								Worst	Total
1	A	262	ASP	CA-CB-CG	8.87	121.47	112.60	2	2
1	A	182	PHE	CA-CB-CG	8.78	122.58	113.80	9	1
1	A	178	ASP	CA-CB-CG	8.19	120.79	112.60	8	8
1	A	226	HIS	CA-CB-CG	7.44	121.24	113.80	14	2
1	A	191	HIS	CA-CB-CG	7.12	120.92	113.80	3	1

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	204	HIS	CA-CB-CG	6.86	120.66	113.80	3	4
1	A	238	THR	CA-C-N	6.76	129.79	120.39	1	1
1	A	238	THR	C-N-CA	6.76	129.79	120.39	1	1
1	A	178	ASP	CA-C-N	6.74	134.40	121.54	4	2
1	A	178	ASP	C-N-CA	6.74	134.40	121.54	4	2
1	A	191	HIS	CB-CG-CD2	-6.73	122.45	131.20	20	10
1	A	228	PHE	CA-C-N	6.67	127.50	120.03	11	11
1	A	228	PHE	C-N-CA	6.67	127.50	120.03	11	11
1	A	139	ASP	CA-CB-CG	6.48	119.08	112.60	8	1
1	A	261	ASP	CA-C-N	6.45	129.24	120.54	20	2
1	A	261	ASP	C-N-CA	6.45	129.24	120.54	20	2
1	A	230	HIS	CB-CG-CD2	-6.42	122.86	131.20	5	5
1	A	157	PHE	CA-CB-CG	6.40	120.20	113.80	11	5
1	A	176	HIS	CA-CB-CG	6.24	120.04	113.80	8	2
1	A	186	ARG	CB-CA-C	-6.14	108.49	115.79	18	1
1	A	166	ASP	N-CA-C	-6.13	104.51	111.07	15	2
1	A	179	SER	CA-C-N	6.11	127.66	122.28	14	2
1	A	179	SER	C-N-CA	6.11	127.66	122.28	14	2
1	A	177	GLY	CA-C-N	6.00	133.00	121.54	2	1
1	A	177	GLY	C-N-CA	6.00	133.00	121.54	2	1
1	A	176	HIS	CB-CG-CD2	-5.95	123.47	131.20	16	1
1	A	257	HIS	CA-CB-CG	5.93	119.73	113.80	8	1
1	A	204	HIS	CB-CG-CD2	-5.92	123.50	131.20	16	1
1	A	241	SER	CA-C-N	5.90	132.81	121.54	14	1
1	A	241	SER	C-N-CA	5.90	132.81	121.54	14	1
1	A	246	PRO	CA-C-N	5.90	132.81	121.54	16	1
1	A	246	PRO	C-N-CA	5.90	132.81	121.54	16	1
1	A	236	HIS	CB-CG-CD2	-5.84	123.61	131.20	1	6
1	A	208	ALA	CA-C-N	5.78	129.79	122.16	10	1
1	A	208	ALA	C-N-CA	5.78	129.79	122.16	10	1
1	A	257	HIS	CB-CG-CD2	-5.77	123.70	131.20	15	3
1	A	267	GLN	OE1-CD-NE2	-5.73	116.87	122.60	15	1
1	A	125	ARG	CD-NE-CZ	5.72	132.40	124.40	11	2
1	A	163	GLY	CA-C-O	-5.66	117.71	122.29	6	1
1	A	247	THR	CA-CB-CG2	5.63	120.08	110.50	13	1
1	A	221	PHE	CA-CB-CG	5.62	119.42	113.80	6	3
1	A	186	ARG	N-CA-C	5.59	115.67	108.34	12	1
1	A	147	GLN	OE1-CD-NE2	-5.59	117.01	122.60	4	2
1	A	185	PRO	CA-C-N	5.56	132.16	121.54	10	1
1	A	185	PRO	C-N-CA	5.56	132.16	121.54	10	1
1	A	219	ASN	N-CA-C	5.55	117.79	108.96	7	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	234	LEU	N-CA-C	5.55	117.69	110.53	5	1
1	A	143	GLU	CA-CB-CG	5.52	125.14	114.10	9	1
1	A	175	ASP	CA-CB-CG	5.50	118.11	112.60	18	2
1	A	207	ASN	CA-CB-CG	5.41	118.01	112.60	4	1
1	A	186	ARG	CA-C-N	5.40	126.52	120.42	4	1
1	A	186	ARG	C-N-CA	5.40	126.52	120.42	4	1
1	A	199	LEU	CA-C-N	5.38	125.08	119.92	6	1
1	A	199	LEU	C-N-CA	5.38	125.08	119.92	6	1
1	A	163	GLY	O-C-N	5.31	129.05	122.78	6	1
1	A	230	HIS	CA-CB-CG	5.31	119.11	113.80	8	1
1	A	226	HIS	CB-CG-CD2	-5.30	124.31	131.20	12	1
1	A	239	ASP	CA-C-O	-5.29	113.11	119.67	20	1
1	A	173	ASN	CA-CB-CG	5.27	117.87	112.60	20	1
1	A	266	ILE	N-CA-C	5.26	115.99	110.62	8	1
1	A	255	GLY	CA-C-N	5.25	129.09	120.63	11	1
1	A	255	GLY	C-N-CA	5.25	129.09	120.63	11	1
1	A	166	ASP	CA-CB-CG	-5.23	107.37	112.60	11	2
1	A	219	ASN	CA-CB-CG	5.23	117.83	112.60	19	1
1	A	258	LEU	CB-CA-C	5.17	115.57	109.47	14	1
1	A	165	ALA	N-CA-C	5.17	117.97	110.17	18	1
1	A	153	VAL	CA-C-N	5.16	126.29	119.84	9	2
1	A	153	VAL	C-N-CA	5.16	126.29	119.84	9	2
1	A	167	ILE	N-CA-CB	-5.12	105.55	112.10	3	1
1	A	132	SER	N-CA-C	5.12	117.76	111.82	5	1
1	A	207	ASN	OD1-CG-ND2	-5.11	117.49	122.60	14	1
1	A	239	ASP	CA-C-N	5.09	124.81	119.82	17	1
1	A	239	ASP	C-N-CA	5.09	124.81	119.82	17	1
1	A	180	TYR	CA-C-N	5.05	126.16	119.84	11	1
1	A	180	TYR	C-N-CA	5.05	126.16	119.84	11	1
1	A	236	HIS	CA-CB-CG	-5.05	108.75	113.80	1	1
1	A	271	GLY	CA-C-O	-5.04	116.27	119.65	13	1
1	A	245	TYR	CA-C-N	5.02	126.11	119.84	14	1
1	A	245	TYR	C-N-CA	5.02	126.11	119.84	14	1
1	A	215	THR	CA-C-N	5.02	131.12	121.54	10	1
1	A	215	THR	C-N-CA	5.02	131.12	121.54	10	1
1	A	174	GLY	CA-C-N	5.00	131.09	121.54	19	1
1	A	174	GLY	C-N-CA	5.00	131.09	121.54	19	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	226	HIS	Sidechain	9
1	A	230	HIS	Sidechain	5
1	A	157	PHE	Sidechain	3
1	A	191	HIS	Sidechain	3
1	A	124	TYR	Sidechain	2
1	A	250	TYR	Sidechain	2
1	A	248	TYR	Sidechain	2
1	A	209	GLU	Sidechain	1
1	A	176	HIS	Sidechain	1
1	A	186	ARG	Sidechain	1
1	A	159	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	1215	1148	1148	2±2
4	A	21	19	19	0±1
All	All	24800	23340	23345	39

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:190:ALA:HB1	1:A:227:GLU:CD	0.62	2.19	14	3
1:A:189:LEU:H	1:A:189:LEU:HD22	0.61	1.56	18	1
1:A:149:TRP:CE3	1:A:258:LEU:HD21	0.58	2.33	6	3
1:A:149:TRP:CZ3	1:A:258:LEU:HD21	0.58	2.33	6	2
1:A:226:HIS:CD2	4:A:277:NGH:O4	0.52	2.61	20	3
1:A:226:HIS:CE1	4:A:277:NGH:O5	0.51	2.63	14	1
1:A:155:LEU:HD11	1:A:157:PHE:CE1	0.50	2.41	3	2
1:A:180:TYR:CE1	1:A:191:HIS:CE1	0.50	3.00	17	2
1:A:117:TRP:CD1	1:A:270:TYR:HH	0.50	2.24	9	1
1:A:155:LEU:HD12	1:A:156:SER:N	0.49	2.23	14	6
1:A:244:MET:N	1:A:244:MET:HE3	0.45	2.26	19	1
1:A:216:ASN:HD22	1:A:216:ASN:C	0.45	2.19	14	1

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:149:TRP:CE3	1:A:266:ILE:HD12	0.43	2.49	10	1
1:A:190:ALA:HB1	1:A:227:GLU:HG3	0.43	1.89	12	2
1:A:190:ALA:HB3	4:A:277:NGH:H5	0.43	1.91	14	1
1:A:226:HIS:CD2	1:A:226:HIS:C	0.42	2.97	18	1
1:A:155:LEU:CD2	1:A:266:ILE:HD11	0.42	2.45	7	1
1:A:160:ILE:HG12	1:A:165:ALA:HB2	0.42	1.90	4	1
1:A:178:ASP:C	1:A:180:TYR:H	0.41	2.23	19	1
1:A:155:LEU:HD12	1:A:155:LEU:C	0.40	2.41	9	3
1:A:226:HIS:C	1:A:226:HIS:CD2	0.40	3.00	12	2

## 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	157/160 (98%)	134±3 (85±2%)	18±4 (11±3%)	5±2 (3±1%)	4	34
All	All	3140/3200 (98%)	2671 (85%)	361 (11%)	108 (3%)	4	34

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	260	LYS	20
1	A	198	GLY	15
1	A	186	ARG	8
1	A	247	THR	7
1	A	179	SER	7
1	A	176	HIS	7
1	A	216	ASN	5
1	A	183	ASP	5
1	A	259	PRO	5
1	A	257	HIS	4
1	A	175	ASP	4
1	A	197	GLU	3
1	A	161	ASN	3

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	154	PRO	2
1	A	174	GLY	2
1	A	238	THR	2
1	A	178	ASP	1
1	A	188	THR	1
1	A	241	SER	1
1	A	200	GLY	1
1	A	164	GLU	1
1	A	215	THR	1
1	A	271	GLY	1
1	A	177	GLY	1
1	A	182	PHE	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	126/128 (98%)	106±2 (84±2%)	20±2 (16±2%)	4	41
All	All	2520/2560 (98%)	2125 (84%)	395 (16%)	4	41

All 80 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	169	ILE	20
1	A	119	LYS	18
1	A	191	HIS	17
1	A	258	LEU	17
1	A	239	ASP	16
1	A	155	LEU	15
1	A	180	TYR	14
1	A	143	GLU	13
1	A	151	SER	11
1	A	140	LYS	11
1	A	264	LYS	10
1	A	262	ASP	10
1	A	116	LYS	8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	210	LYS	8
1	A	216	ASN	8
1	A	220	LEU	8
1	A	166	ASP	7
1	A	179	SER	7
1	A	168	MET	6
1	A	186	ARG	6
1	A	215	THR	6
1	A	133	MET	6
1	A	153	VAL	6
1	A	159	ARG	6
1	A	254	TYR	6
1	A	209	GLU	6
1	A	244	MET	6
1	A	161	ASN	6
1	A	175	ASP	6
1	A	144	MET	5
1	A	147	GLN	5
1	A	247	THR	5
1	A	127	SER	4
1	A	137	GLU	4
1	A	173	ASN	4
1	A	261	ASP	4
1	A	156	SER	4
1	A	193	PHE	4
1	A	134	SER	3
1	A	260	LYS	3
1	A	118	LYS	3
1	A	237	SER	3
1	A	267	GLN	3
1	A	132	SER	3
1	A	269	LEU	3
1	A	125	ARG	3
1	A	232	LEU	3
1	A	126	ILE	2
1	A	219	ASN	2
1	A	170	SER	2
1	A	135	SER	2
1	A	257	HIS	2
1	A	241	SER	2
1	A	206	ASP	2
1	A	183	ASP	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	238	THR	2
1	A	138	VAL	2
1	A	139	ASP	2
1	A	150	SER	2
1	A	130	THR	1
1	A	128	LYS	1
1	A	172	GLU	1
1	A	197	GLU	1
1	A	199	LEU	1
1	A	245	TYR	1
1	A	222	THR	1
1	A	182	PHE	1
1	A	121	THR	1
1	A	160	ILE	1
1	A	164	GLU	1
1	A	252	ASN	1
1	A	213	MET	1
1	A	230	HIS	1
1	A	142	VAL	1
1	A	223	VAL	1
1	A	178	ASP	1
1	A	266	ILE	1
1	A	248	TYR	1
1	A	189	LEU	1
1	A	203	THR	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 oligosaccharides in this entry.

## 6.6 Ligand geometry

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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
4	NGH	A	277	-	21,21,21	1.61±0.07	3±1 (14±3%)

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
4	NGH	A	277	-	29,29,29	1.57±0.39	6±3 (19±8%)

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
4	NGH	A	277	-	-	0±0,24,24,24	0±0,1,1,1

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
4	A	277	NGH	O4-N1	6.44	1.24	1.40	13	20
4	A	277	NGH	O2-S1	3.25	1.47	1.43	8	20
4	A	277	NGH	O3-S1	2.68	1.46	1.43	11	16
4	A	277	NGH	C6-S1	2.46	1.79	1.76	15	2

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
4	A	277	NGH	C9-N	2.17	1.51	1.47	1	1
4	A	277	NGH	S1-N	2.02	1.66	1.63	1	1

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(°)	Ideal(°)	Models	
								Worst	Total
4	A	277	NGH	C6-S1-N	6.87	98.98	107.28	2	13
4	A	277	NGH	O3-S1-C6	6.75	116.49	108.10	2	15
4	A	277	NGH	O2-S1-C6	5.97	100.68	108.10	13	12
4	A	277	NGH	O5-C11-N1	5.87	116.06	123.27	2	10
4	A	277	NGH	C9-N-S1	4.40	126.92	117.32	8	13
4	A	277	NGH	O3-S1-N	4.02	110.47	106.69	13	3
4	A	277	NGH	C5-C6-S1	3.93	115.83	119.73	2	6
4	A	277	NGH	C1-C6-S1	3.82	123.53	119.73	2	10
4	A	277	NGH	O4-N1-C11	3.56	114.53	119.79	20	3
4	A	277	NGH	C12-C9-N	3.26	108.64	112.41	1	3
4	A	277	NGH	O2-S1-N	2.82	104.03	106.69	20	5
4	A	277	NGH	C7-O1-C3	2.77	123.43	117.50	7	9
4	A	277	NGH	O3-S1-O2	2.48	115.72	119.59	10	3
4	A	277	NGH	O5-C11-C10	2.38	125.18	121.02	14	4
4	A	277	NGH	C10-N-C9	2.32	110.90	116.35	20	1
4	A	277	NGH	C11-C10-N	2.03	110.78	116.43	20	1

All unique chiral outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
4	A	277	NGH	N	10

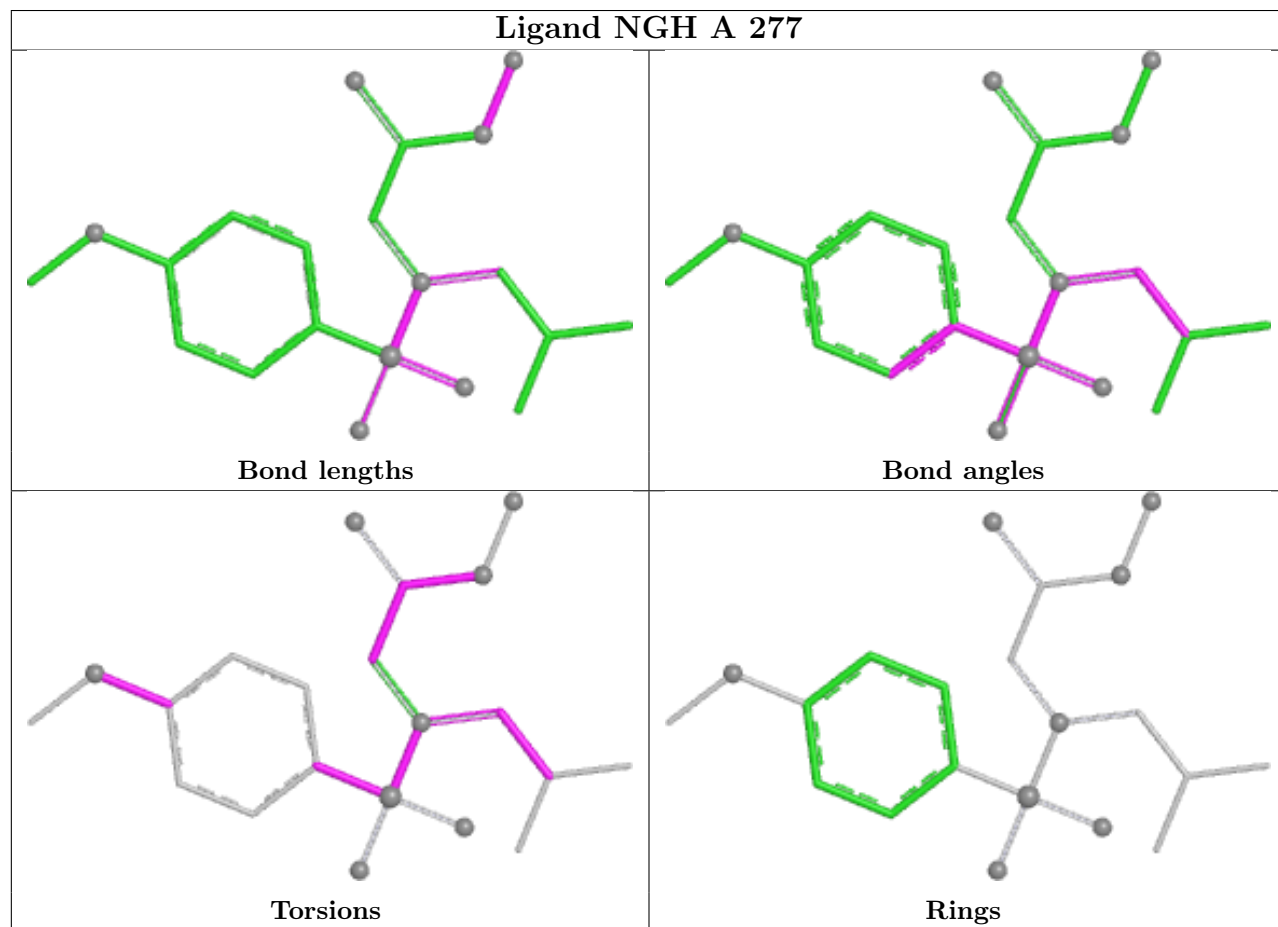
All unique torsion outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Models (Total)
4	A	277	NGH	C10-C11-N1-O4	2
4	A	277	NGH	O5-C11-N1-O4	1

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 than 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.



## 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

No chemical shift data were provided