Full wwPDB NMR Structure Validation Report (i)

Dec 13, 2023 – 08:23 PM EST

PDB ID	:	2JNP
BMRB ID	:	15120
Title	:	Solution structure of matrix metalloproteinase 3 (MMP-3) in the presence of N-isobutyl-N-[4-methoxyphenylsulfonyl]glycyl hydroxamic acid (NNGH)
Authors	:	Alcaraz, L.A.; Banci, L.; Bertini, I.; Cantini, F.; Donaire, A.; Gonnelli, L.
Deposited on	:	2007-01-30

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:

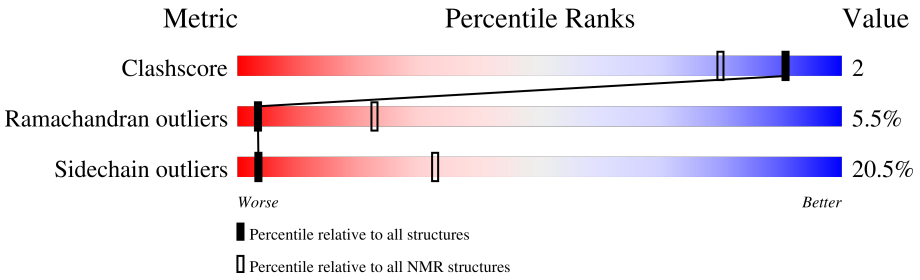
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 is 64%.

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	161	80% 18%

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	253	9	-

2 Ensemble composition and analysis (i)

This entry contains 25 models. Model 14 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:89-A:247 (159)	1.05	14

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, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25
2	8, 13
Single-model clusters	24

3 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2533 atoms, of which 1230 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Matrix metalloproteinase-3.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	161	2489	818	1211	213	245	2	0

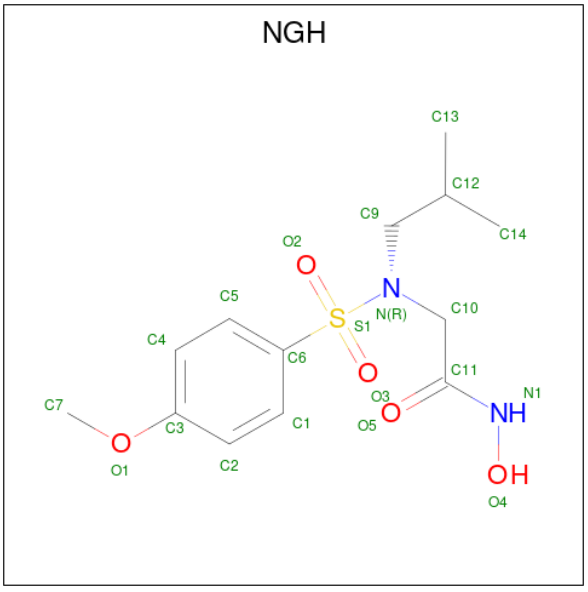
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	2	2	2

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	
			Total	Ca
3	A	2	2	2

- Molecule 4 is N-ISOBUTYL-N-[4-METHOXYPHENYLSULFONYL]GLYCYL HYDROXAMIC ACID (three-letter code: NGH) (formula: $C_{13}H_{20}N_2O_5S$).



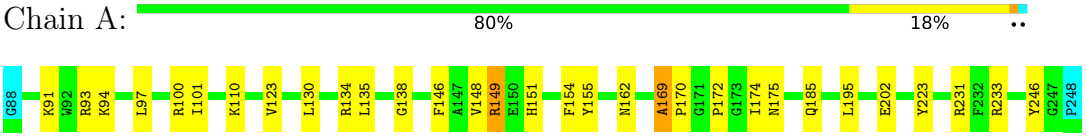
Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	S
4	A	1	40	13	19	2	5	1

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-3

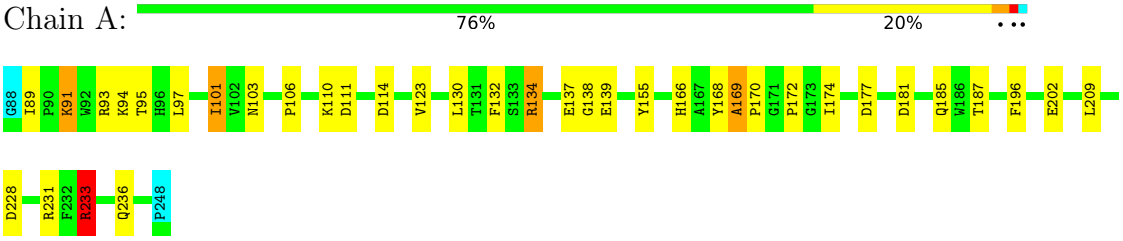


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

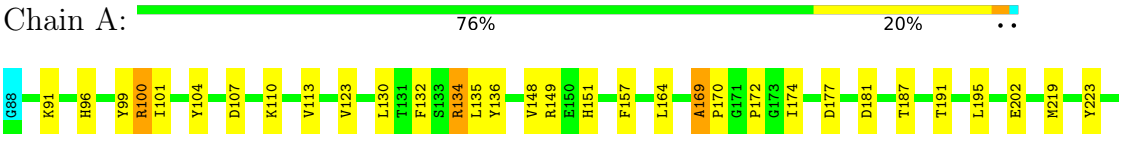
4.2.1 Score per residue for model 1

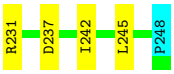
• Molecule 1: Matrix metalloproteinase-3



4.2.2 Score per residue for model 2

• Molecule 1: Matrix metalloproteinase-3

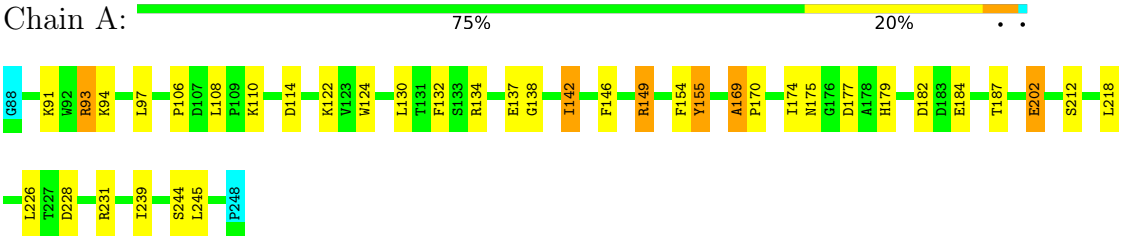




4.2.3 Score per residue for model 3

• Molecule 1: Matrix metalloproteinase-3

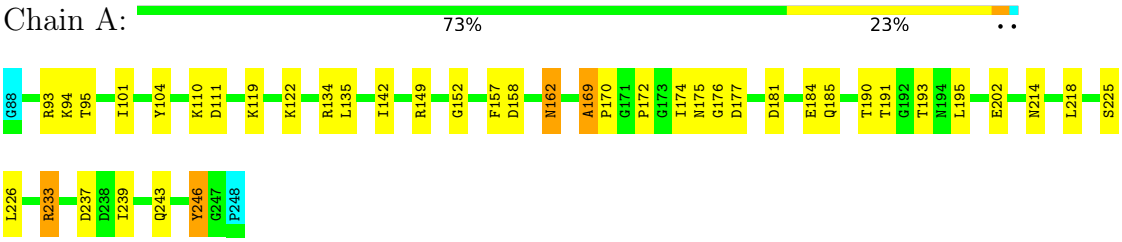
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: Matrix metalloproteinase-3

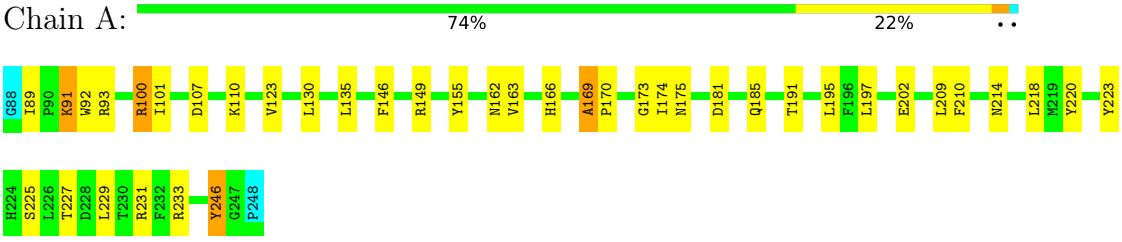
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: Matrix metalloproteinase-3

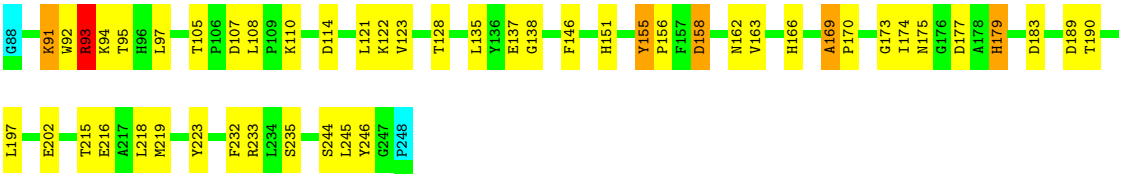
Chain A:



4.2.6 Score per residue for model 6

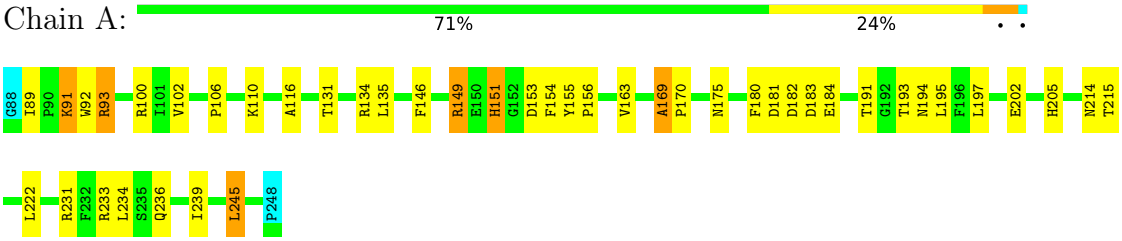
• Molecule 1: Matrix metalloproteinase-3

Chain A:



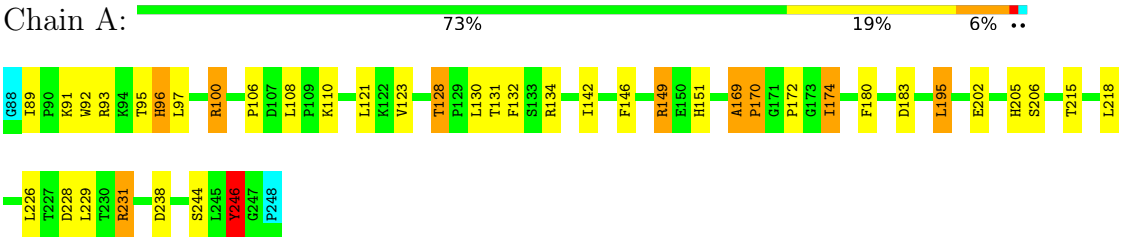
4.2.7 Score per residue for model 7

• Molecule 1: Matrix metalloproteinase-3



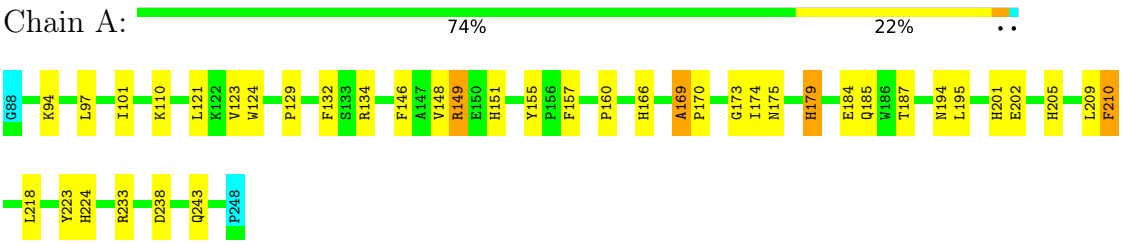
4.2.8 Score per residue for model 8

• Molecule 1: Matrix metalloproteinase-3



4.2.9 Score per residue for model 9

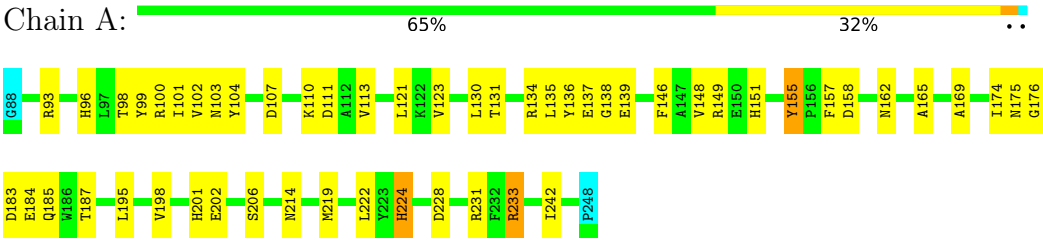
• Molecule 1: Matrix metalloproteinase-3



4.2.10 Score per residue for model 10

• Molecule 1: Matrix metalloproteinase-3

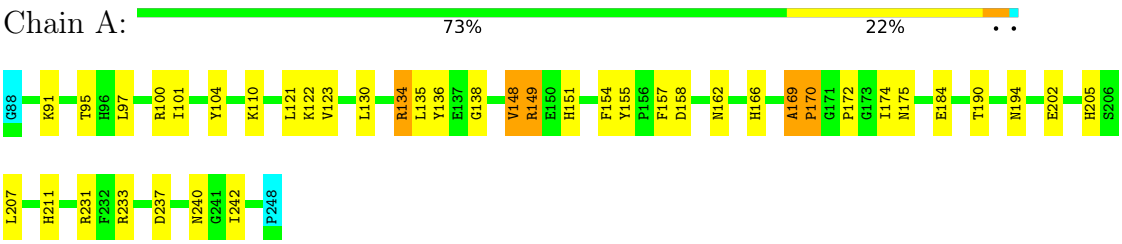
Chain A:



4.2.11 Score per residue for model 11

- Molecule 1: Matrix metalloproteinase-3

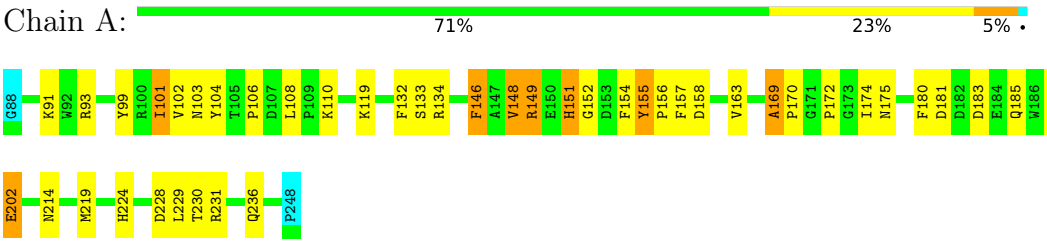
Chain A:



4.2.12 Score per residue for model 12

- Molecule 1: Matrix metalloproteinase-3

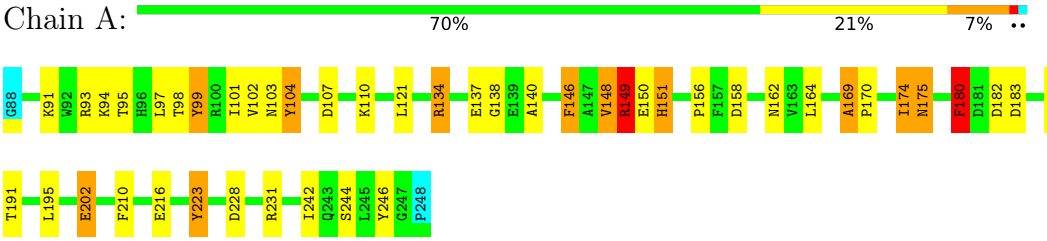
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: Matrix metalloproteinase-3

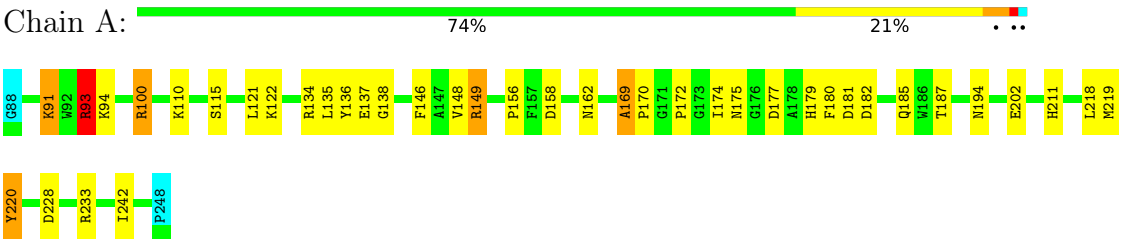
Chain A:



4.2.14 Score per residue for model 14 (medoid)

• Molecule 1: Matrix metalloproteinase-3

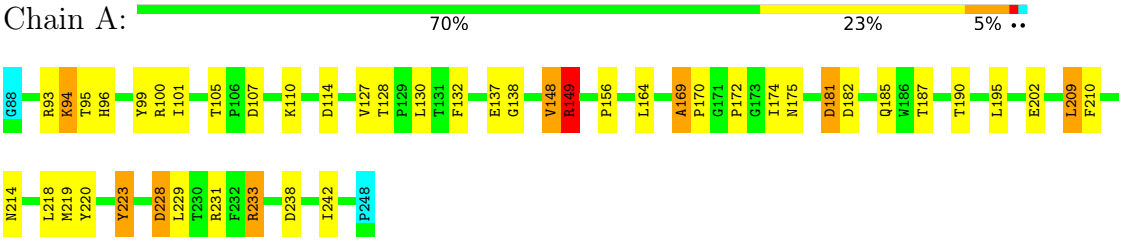
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Matrix metalloproteinase-3

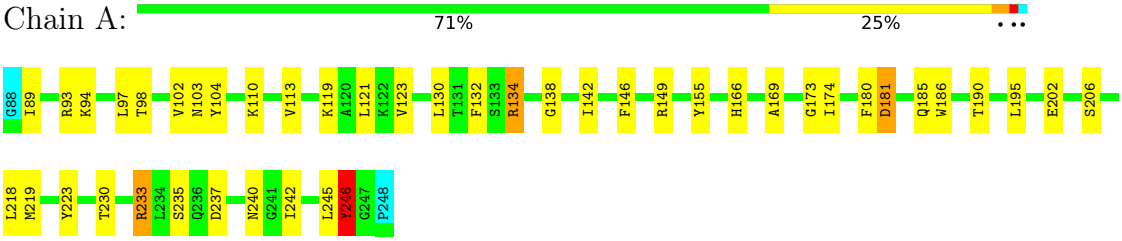
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Matrix metalloproteinase-3

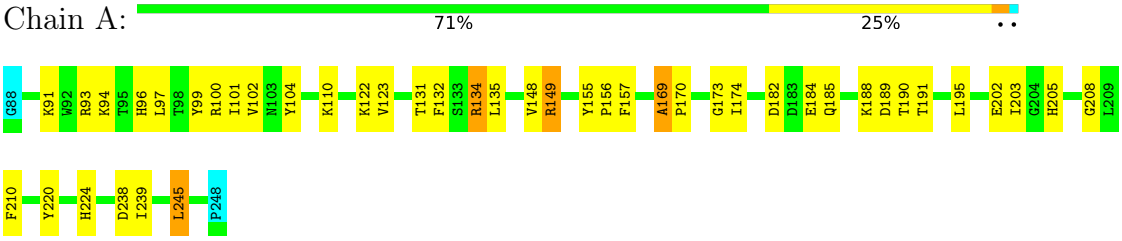
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: Matrix metalloproteinase-3

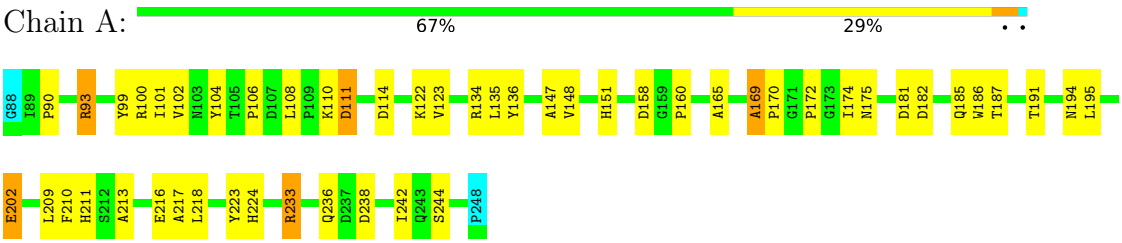
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: Matrix metalloproteinase-3

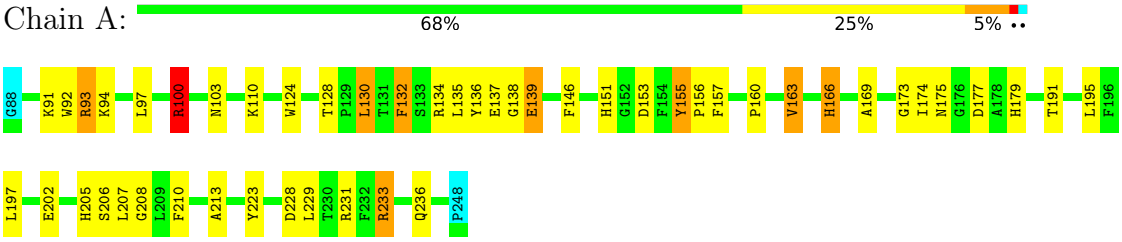
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: Matrix metalloproteinase-3

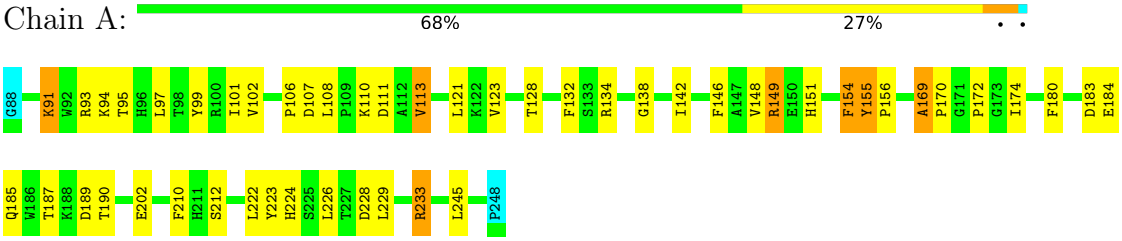
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Matrix metalloproteinase-3

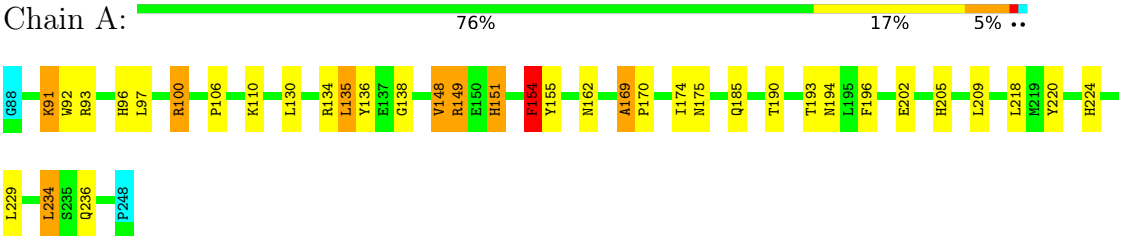
Chain A:



4.2.21 Score per residue for model 21

• Molecule 1: Matrix metalloproteinase-3

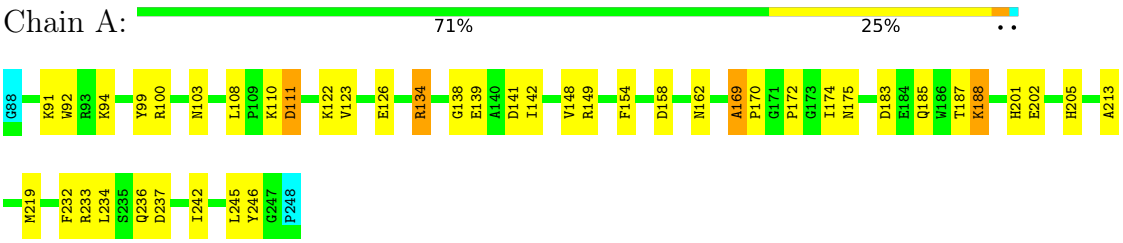
Chain A:



4.2.22 Score per residue for model 22

• Molecule 1: Matrix metalloproteinase-3

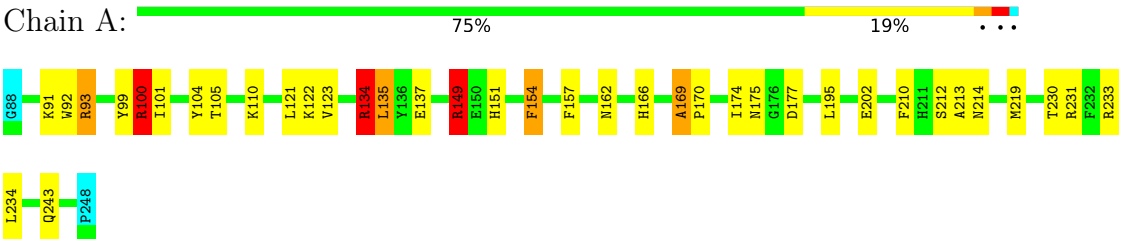
Chain A:



4.2.23 Score per residue for model 23

• Molecule 1: Matrix metalloproteinase-3

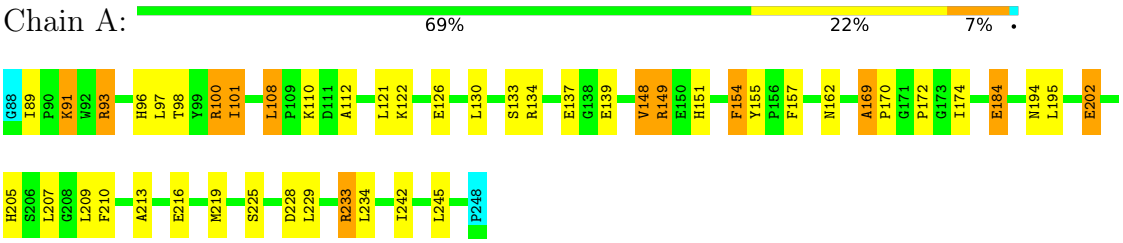
Chain A:



4.2.24 Score per residue for model 24

• Molecule 1: Matrix metalloproteinase-3

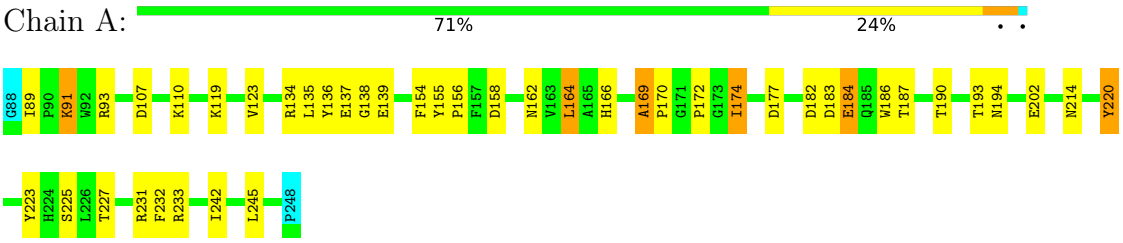
Chain A:



4.2.25 Score per residue for model 25

• Molecule 1: Matrix metalloproteinase-3

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 400 calculated structures, 25 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	
Amber	refinement	8
CSI	structure solution	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	1374
Number of shifts mapped to atoms	1374
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	64%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NGH, CA, ZN

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	0.81±0.01	1±0/1305 (0.1± 0.0%)	1.24±0.04	6±3/1783 (0.3± 0.2%)
All	All	0.81	25/32625 (0.1%)	1.24	147/44575 (0.3%)

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	2.7±1.6
All	All	0	67

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	202	GLU	CD-OE2	9.45	1.36	1.25	19	25

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	231	ARG	NE-CZ-NH1	12.81	126.70	120.30	8	11
1	A	158	ASP	CB-CG-OD1	11.45	128.60	118.30	11	3
1	A	228	ASP	CB-CG-OD1	11.20	128.38	118.30	1	2
1	A	100	ARG	NE-CZ-NH1	10.37	125.48	120.30	21	9
1	A	134	ARG	NE-CZ-NH1	9.71	125.16	120.30	18	12
1	A	231	ARG	NE-CZ-NH2	-9.11	115.75	120.30	7	3
1	A	233	ARG	NE-CZ-NH1	8.91	124.76	120.30	11	12
1	A	149	ARG	NE-CZ-NH1	8.83	124.72	120.30	8	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	93	ARG	NE-CZ-NH1	8.82	124.71	120.30	23	11
1	A	111	ASP	CB-CG-OD2	-8.77	110.41	118.30	22	1
1	A	220	TYR	CB-CG-CD2	-8.40	115.96	121.00	25	3
1	A	246	TYR	CB-CG-CD2	-7.72	116.37	121.00	8	2
1	A	177	ASP	CB-CG-OD2	7.64	125.18	118.30	19	4
1	A	139	GLU	OE1-CD-OE2	-7.61	114.16	123.30	19	1
1	A	134	ARG	NE-CZ-NH2	-7.58	116.51	120.30	4	3
1	A	170	PRO	N-CA-CB	7.44	112.23	103.30	8	1
1	A	233	ARG	NE-CZ-NH2	-7.43	116.58	120.30	1	2
1	A	184	GLU	OE1-CD-OE2	-7.01	114.88	123.30	24	6
1	A	231	ARG	CD-NE-CZ	6.91	133.28	123.60	8	3
1	A	160	PRO	N-CD-CG	6.89	113.54	103.20	9	1
1	A	100	ARG	NE-CZ-NH2	6.77	123.68	120.30	19	5
1	A	146	PHE	CB-CG-CD1	6.76	125.53	120.80	13	1
1	A	220	TYR	CB-CG-CD1	6.46	124.88	121.00	25	2
1	A	237	ASP	CB-CG-OD1	6.46	124.11	118.30	4	1
1	A	158	ASP	CB-CG-OD2	6.36	124.02	118.30	6	3
1	A	177	ASP	CB-CG-OD1	6.28	123.95	118.30	1	4
1	A	246	TYR	CB-CG-CD1	6.25	124.75	121.00	8	1
1	A	228	ASP	CB-CG-OD2	6.19	123.87	118.30	24	2
1	A	160	PRO	N-CA-CB	6.07	110.58	103.30	18	1
1	A	111	ASP	CB-CG-OD1	6.07	123.76	118.30	18	2
1	A	181	ASP	CB-CG-OD1	6.03	123.73	118.30	15	5
1	A	237	ASP	CB-CG-OD2	-5.99	112.91	118.30	4	1
1	A	238	ASP	CB-CG-OD2	-5.96	112.94	118.30	8	2
1	A	146	PHE	CB-CG-CD2	-5.90	116.67	120.80	13	1
1	A	99	TYR	CB-CG-CD2	-5.88	117.47	121.00	2	1
1	A	149	ARG	NE-CZ-NH2	-5.65	117.48	120.30	23	1
1	A	113	VAL	CA-CB-CG1	5.64	119.36	110.90	20	1
1	A	157	PHE	CB-CG-CD2	-5.60	116.88	120.80	23	1
1	A	183	ASP	CB-CG-OD1	5.55	123.30	118.30	20	1
1	A	180	PHE	CB-CG-CD2	-5.54	116.92	120.80	8	2
1	A	246	TYR	CA-CB-CG	5.44	123.73	113.40	16	1
1	A	174	ILE	CA-CB-CG1	5.39	121.25	111.00	25	1
1	A	155	TYR	CB-CG-CD2	-5.22	117.87	121.00	6	2
1	A	231	ARG	NH1-CZ-NH2	-5.22	113.66	119.40	8	1
1	A	209	LEU	CB-CA-C	5.21	120.10	110.20	15	1
1	A	154	PHE	CB-CG-CD2	-5.13	117.21	120.80	21	1
1	A	180	PHE	CB-CG-CD1	5.08	124.36	120.80	8	1
1	A	148	VAL	CA-CB-CG2	5.07	118.50	110.90	24	1
1	A	142	ILE	CA-CB-CG1	5.04	120.57	111.00	3	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	181	ASP	CB-CG-OD2	-5.01	113.79	118.30	14	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	223	TYR	Sidechain	7
1	A	134	ARG	Sidechain	7
1	A	100	ARG	Sidechain	6
1	A	155	TYR	Sidechain	6
1	A	233	ARG	Sidechain	5
1	A	93	ARG	Sidechain	5
1	A	246	TYR	Sidechain	4
1	A	151	HIS	Sidechain	4
1	A	149	ARG	Sidechain	4
1	A	104	TYR	Sidechain	3
1	A	99	TYR	Sidechain	3
1	A	231	ARG	Sidechain	2
1	A	179	HIS	Sidechain	2
1	A	232	PHE	Sidechain	2
1	A	211	HIS	Sidechain	2
1	A	168	TYR	Sidechain	1
1	A	180	PHE	Sidechain	1
1	A	181	ASP	Sidechain	1
1	A	220	TYR	Sidechain	1
1	A	166	HIS	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	1266	1199	1198	4±2
4	A	21	19	20	0±0
All	All	32275	30450	30447	101

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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:169:ALA:HB3	1:A:170:PRO:HD3	0.59	1.74	9	20
1:A:91:LYS:HE2	1:A:169:ALA:HB1	0.56	1.77	20	3
1:A:100:ARG:NH2	1:A:135:LEU:HD12	0.56	2.14	21	1
1:A:91:LYS:CE	1:A:169:ALA:HB1	0.56	2.31	20	2
1:A:176:GLY:HA2	1:A:206:SER:HB2	0.53	1.80	10	1
1:A:169:ALA:HB3	1:A:170:PRO:CD	0.53	2.34	14	11
1:A:164:LEU:HD21	1:A:193:THR:HG23	0.53	1.81	25	1
1:A:130:LEU:HD21	1:A:132:PHE:CD2	0.52	2.39	19	1
1:A:146:PHE:CZ	1:A:180:PHE:CD2	0.52	2.97	13	1
1:A:165:ALA:HB1	1:A:202:GLU:CD	0.52	2.25	18	1
1:A:148:VAL:HG22	1:A:149:ARG:H	0.51	1.66	11	5
1:A:234:LEU:HD23	1:A:234:LEU:H	0.51	1.66	23	1
1:A:101:ILE:HD12	1:A:101:ILE:H	0.50	1.66	4	7
1:A:209:LEU:HD11	1:A:238:ASP:HA	0.47	1.85	9	1
1:A:100:ARG:HH22	1:A:135:LEU:HD12	0.47	1.69	21	1
1:A:146:PHE:CD1	1:A:180:PHE:HB2	0.47	2.45	13	1
1:A:108:LEU:HD13	1:A:112:ALA:CB	0.46	2.40	24	1
1:A:104:TYR:HA	1:A:113:VAL:HG11	0.46	1.88	2	2
1:A:128:THR:HB	1:A:246:TYR:CE1	0.46	2.46	8	1
1:A:116:ALA:HB2	1:A:195:LEU:HD13	0.46	1.88	7	1
1:A:208:GLY:HA3	1:A:245:LEU:HD11	0.46	1.88	17	1
4:A:253:NGH:H133	4:A:253:NGH:H102	0.46	1.88	20	1
1:A:100:ARG:HH21	1:A:135:LEU:C	0.45	2.15	21	1
1:A:99:TYR:CE2	1:A:121:LEU:HD22	0.44	2.48	13	1
1:A:134:ARG:C	1:A:135:LEU:HD23	0.44	2.32	23	1
1:A:146:PHE:CD2	1:A:186:TRP:CH2	0.44	3.06	13	1
1:A:234:LEU:H	1:A:234:LEU:CD2	0.44	2.26	24	1
1:A:94:LYS:HE3	1:A:96:HIS:O	0.43	2.13	15	1
1:A:116:ALA:HB2	1:A:195:LEU:CD1	0.43	2.44	7	1
1:A:213:ALA:HA	1:A:217:ALA:HB3	0.43	1.90	18	1
1:A:103:ASN:HB2	1:A:146:PHE:CD1	0.43	2.48	16	1
1:A:101:ILE:HD12	1:A:101:ILE:N	0.43	2.29	17	1
1:A:146:PHE:CE1	1:A:180:PHE:CG	0.42	3.06	13	1
1:A:103:ASN:HB2	1:A:146:PHE:CD2	0.42	2.49	12	1
1:A:169:ALA:HB2	1:A:205:HIS:O	0.42	2.14	17	2
1:A:169:ALA:CB	1:A:170:PRO:CD	0.42	2.98	23	3
1:A:234:LEU:H	1:A:234:LEU:HD22	0.42	1.74	21	1
1:A:197:LEU:CD2	4:A:253:NGH:H71	0.42	2.44	6	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:98:THR:HG22	1:A:133:SER:HB3	0.42	1.92	24	1
1:A:165:ALA:HB3	1:A:198:VAL:CG1	0.42	2.44	10	1
1:A:188:LYS:HA	1:A:188:LYS:HE2	0.42	1.90	22	1
1:A:101:ILE:H	1:A:101:ILE:HD12	0.42	1.74	24	1
1:A:174:ILE:O	1:A:175:ASN:C	0.42	2.58	13	1
1:A:216:GLU:CD	1:A:216:GLU:H	0.42	2.18	24	1
1:A:211:HIS:CD2	1:A:211:HIS:N	0.41	2.88	18	1
1:A:101:ILE:N	1:A:101:ILE:HD12	0.41	2.31	11	1
1:A:147:ALA:HB1	1:A:151:HIS:CG	0.41	2.51	18	1
1:A:108:LEU:HD12	1:A:195:LEU:CD2	0.41	2.46	8	1
1:A:100:ARG:HH21	1:A:139:GLU:C	0.40	2.19	19	1
1:A:103:ASN:HB2	1:A:146:PHE:CG	0.40	2.52	13	1
1:A:216:GLU:N	1:A:216:GLU:CD	0.40	2.75	6	1
1:A:96:HIS:HA	1:A:131:THR:O	0.40	2.17	8	1
1:A:169:ALA:HB2	1:A:206:SER:HA	0.40	1.94	8	1
1:A:99:TYR:CE2	1:A:132:PHE:CG	0.40	3.09	12	1
1:A:142:ILE:HG23	1:A:176:GLY:O	0.40	2.16	4	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	159/161 (99%)	126±4 (79±3%)	24±4 (15±3%)	9±2 (6±1%)	3	23
All	All	3975/4025 (99%)	3146 (79%)	609 (15%)	220 (6%)	3	23

All 45 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	169	ALA	25
1	A	175	ASN	16
1	A	149	ARG	15
1	A	91	LYS	14
1	A	138	GLY	14

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	148	VAL	14
1	A	172	PRO	13
1	A	156	PRO	10
1	A	154	PHE	9
1	A	106	PRO	8
1	A	190	THR	7
1	A	137	GLU	6
1	A	162	ASN	6
1	A	173	GLY	6
1	A	163	VAL	5
1	A	139	GLU	4
1	A	213	ALA	4
1	A	245	LEU	3
1	A	93	ARG	3
1	A	89	ILE	3
1	A	228	ASP	3
1	A	224	HIS	3
1	A	194	ASN	3
1	A	152	GLY	2
1	A	225	SER	2
1	A	229	LEU	2
1	A	210	PHE	2
1	A	218	LEU	1
1	A	191	THR	1
1	A	174	ILE	1
1	A	129	PRO	1
1	A	157	PHE	1
1	A	195	LEU	1
1	A	140	ALA	1
1	A	182	ASP	1
1	A	185	GLN	1
1	A	90	PRO	1
1	A	216	GLU	1
1	A	160	PRO	1
1	A	208	GLY	1
1	A	223	TYR	1
1	A	212	SER	1
1	A	219	MET	1
1	A	183	ASP	1
1	A	186	TRP	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	134/135 (99%)	107±4 (79±3%)	27±4 (21±3%)	3 33
All	All	3350/3375 (99%)	2663 (79%)	687 (21%)	3 33

All 116 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	110	LYS	25
1	A	174	ILE	24
1	A	123	VAL	15
1	A	155	TYR	14
1	A	135	LEU	14
1	A	151	HIS	14
1	A	195	LEU	14
1	A	94	LYS	13
1	A	97	LEU	13
1	A	185	GLN	13
1	A	130	LEU	12
1	A	187	THR	12
1	A	93	ARG	12
1	A	233	ARG	11
1	A	242	ILE	11
1	A	146	PHE	11
1	A	91	LYS	10
1	A	132	PHE	10
1	A	134	ARG	10
1	A	122	LYS	10
1	A	218	LEU	10
1	A	121	LEU	10
1	A	166	HIS	9
1	A	219	MET	9
1	A	210	PHE	9
1	A	95	THR	8
1	A	101	ILE	8
1	A	107	ASP	8
1	A	136	TYR	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	157	PHE	8
1	A	245	LEU	8
1	A	214	ASN	8
1	A	92	TRP	8
1	A	102	VAL	8
1	A	236	GLN	7
1	A	191	THR	7
1	A	108	LEU	7
1	A	182	ASP	7
1	A	158	ASP	7
1	A	162	ASN	7
1	A	183	ASP	7
1	A	180	PHE	7
1	A	205	HIS	7
1	A	209	LEU	6
1	A	96	HIS	6
1	A	104	TYR	6
1	A	184	GLU	6
1	A	246	TYR	6
1	A	149	ARG	6
1	A	229	LEU	6
1	A	114	ASP	5
1	A	223	TYR	5
1	A	137	GLU	5
1	A	142	ILE	5
1	A	179	HIS	5
1	A	228	ASP	5
1	A	111	ASP	5
1	A	100	ARG	5
1	A	181	ASP	5
1	A	128	THR	5
1	A	154	PHE	5
1	A	194	ASN	5
1	A	99	TYR	5
1	A	224	HIS	5
1	A	89	ILE	4
1	A	103	ASN	4
1	A	164	LEU	4
1	A	237	ASP	4
1	A	202	GLU	4
1	A	226	LEU	4
1	A	239	ILE	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	244	SER	4
1	A	119	LYS	4
1	A	124	TRP	3
1	A	193	THR	3
1	A	243	GLN	3
1	A	197	LEU	3
1	A	105	THR	3
1	A	189	ASP	3
1	A	215	THR	3
1	A	131	THR	3
1	A	222	LEU	3
1	A	234	LEU	3
1	A	201	HIS	3
1	A	98	THR	3
1	A	207	LEU	3
1	A	230	THR	3
1	A	188	LYS	3
1	A	220	TYR	3
1	A	196	PHE	2
1	A	177	ASP	2
1	A	212	SER	2
1	A	190	THR	2
1	A	225	SER	2
1	A	227	THR	2
1	A	235	SER	2
1	A	153	ASP	2
1	A	240	ASN	2
1	A	113	VAL	2
1	A	186	TRP	2
1	A	206	SER	2
1	A	238	ASP	2
1	A	126	GLU	2
1	A	175	ASN	1
1	A	231	ARG	1
1	A	170	PRO	1
1	A	133	SER	1
1	A	150	GLU	1
1	A	216	GLU	1
1	A	115	SER	1
1	A	127	VAL	1
1	A	203	ILE	1
1	A	163	VAL	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	141	ASP	1
1	A	232	PHE	1
1	A	139	GLU	1

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

Of 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	253	2	21,21,21	1.68±0.09	4±1 (20±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	253	2	27,29,29	1.44±0.25	5±2 (17±6%)

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	253	2	1±0,1,5,5	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	253	NGH	O4-N1	5.06	1.27	1.40	9	25
4	A	253	NGH	C11-N1	3.81	1.36	1.32	23	21
4	A	253	NGH	O3-S1	3.63	1.47	1.43	4	18
4	A	253	NGH	O2-S1	3.47	1.47	1.43	24	23
4	A	253	NGH	C6-S1	3.42	1.81	1.76	21	20

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
4	A	253	NGH	C6-S1-N	7.04	115.69	107.30	5	17
4	A	253	NGH	O5-C11-N1	5.33	116.72	123.27	19	5
4	A	253	NGH	C12-C9-N	4.65	117.80	112.41	20	8
4	A	253	NGH	O2-S1-C6	4.10	102.86	108.05	20	13
4	A	253	NGH	O3-S1-O2	3.85	113.28	119.52	13	9
4	A	253	NGH	O3-S1-N	3.82	103.20	106.69	5	7
4	A	253	NGH	O3-S1-C6	3.73	103.33	108.05	16	17
4	A	253	NGH	O2-S1-N	3.33	103.65	106.69	25	8
4	A	253	NGH	C10-C11-N1	3.20	122.29	115.25	11	1
4	A	253	NGH	O4-N1-C11	3.12	124.41	119.79	21	10
4	A	253	NGH	C1-C6-S1	3.05	122.97	119.76	9	4
4	A	253	NGH	C7-O1-C3	3.03	124.09	117.51	9	15
4	A	253	NGH	C9-N-S1	2.54	123.10	117.52	17	4
4	A	253	NGH	O5-C11-C10	2.51	116.68	121.08	11	1
4	A	253	NGH	C5-C6-S1	2.11	117.53	119.76	9	1

All unique chiral outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
4	A	253	NGH	N	9

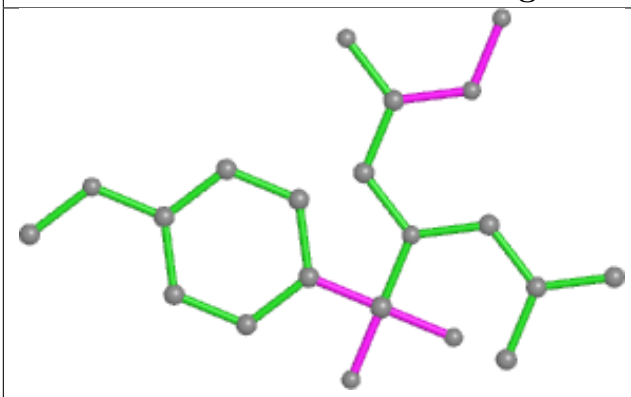
All unique torsion outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Models (Total)
4	A	253	NGH	C10-C11-N1-O4	1
4	A	253	NGH	O5-C11-N1-O4	1

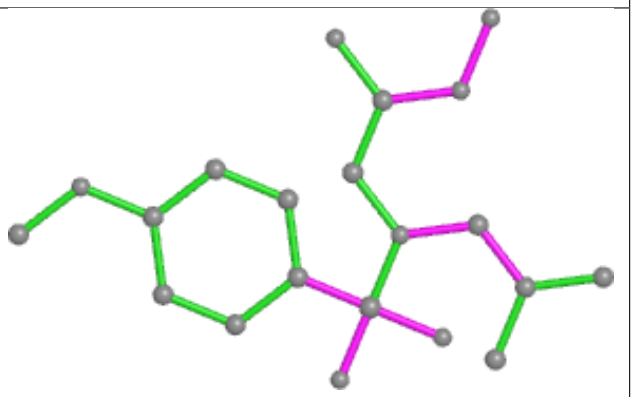
There are no ring outliers.

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 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.

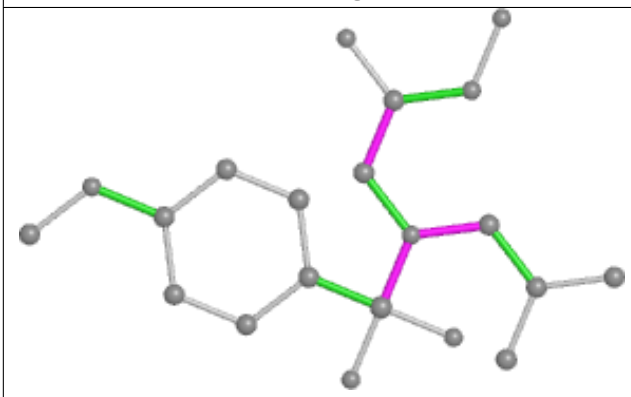
Ligand NGH A 253



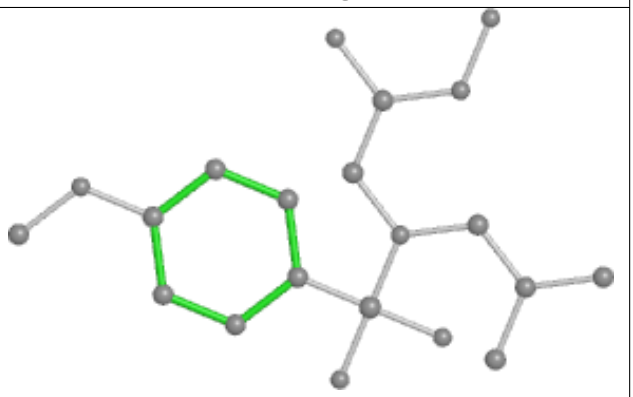
Bond lengths



Bond angles



Torsions



Rings

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

The completeness of assignment taking into account all chemical shift lists is 64% for the well-defined parts and 64% 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

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	1374
Number of shifts mapped to atoms	1374
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	14

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	121	2.58 ± 0.13	Should be applied
$^{13}C_{\beta}$	110	2.77 ± 0.20	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	146	0.17 ± 0.65	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 64%, i.e. 1372 atoms were assigned a chemical shift out of a possible 2132. 0 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	570/789 (72%)	303/321 (94%)	121/318 (38%)	146/150 (97%)
Sidechain	765/1099 (70%)	544/717 (76%)	215/349 (62%)	6/33 (18%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	37/244 (15%)	34/122 (28%)	0/111 (0%)	3/11 (27%)
Overall	1372/2132 (64%)	881/1160 (76%)	336/778 (43%)	155/194 (80%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 64%, i.e. 1374 atoms were assigned a chemical shift out of a possible 2150. 0 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	571/798 (72%)	304/325 (94%)	121/322 (38%)	146/151 (97%)
Sidechain	766/1108 (69%)	545/723 (75%)	215/352 (61%)	6/33 (18%)
Aromatic	37/244 (15%)	34/122 (28%)	0/111 (0%)	3/11 (27%)
Overall	1374/2150 (64%)	883/1170 (75%)	336/785 (43%)	155/195 (79%)

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	208	GLY	CA	55.80	38.93 – 51.79	8.1
1	A	105	THR	HB	1.94	2.57 – 5.77	-7.0
1	A	93	ARG	CD	37.30	38.57 – 47.75	-6.4
1	A	152	GLY	CA	53.30	38.93 – 51.79	6.2
1	A	91	LYS	CE	47.07	37.57 – 46.21	6.0
1	A	94	LYS	CG	18.85	19.35 – 30.45	-5.5
1	A	119	LYS	CD	23.17	23.50 – 34.42	-5.3
1	A	91	LYS	CG	19.06	19.35 – 30.45	-5.3
1	A	202	GLU	HB2	0.96	1.00 – 3.05	-5.2
1	A	193	THR	HG21	0.05	0.08 – 2.19	-5.1
1	A	193	THR	HG22	0.05	0.08 – 2.19	-5.1
1	A	193	THR	HG23	0.05	0.08 – 2.19	-5.1
1	A	122	LYS	CE	37.48	37.57 – 46.21	-5.1
1	A	218	LEU	H	11.40	5.09 – 11.34	5.1

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: