Full wwPDB NMR Structure Validation Report (i)

Oct 11, 2021 – 02:26 PM EDT

PDB ID	:	2ORL
Title	:	Solution structure of the cytochrome c- para-aminophenol adduct
Authors	:	Assfalg, M.; Bertini, I.; Del Conte, R.; Giachetti, A.; Turano, P.
Deposited on	:	2007-02-03

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

We welcome your comments at *validation@mail.wwpdb.org*
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:

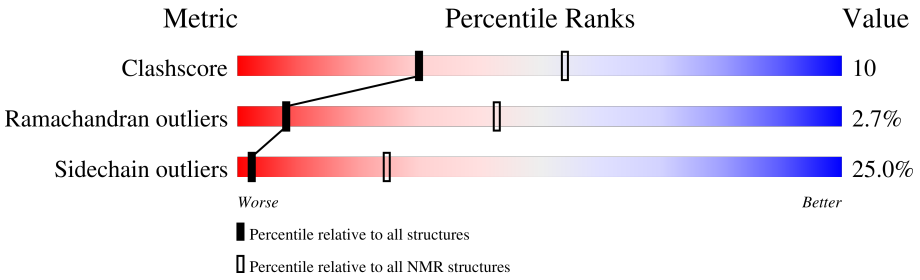
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.23.2
Ideal geometry (proteins)	:	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 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	108	61% 31% 6% •

2 Ensemble composition and analysis (i)

This entry contains 25 models. Model 12 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:-3-A:-1, A:1-A:103 (106)	0.24	12

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

NmrClust was unable to cluster the ensemble.

Error message: 1 Unexpected character

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1789 atoms, of which 890 are hydrogens and 0 are deuteriums.

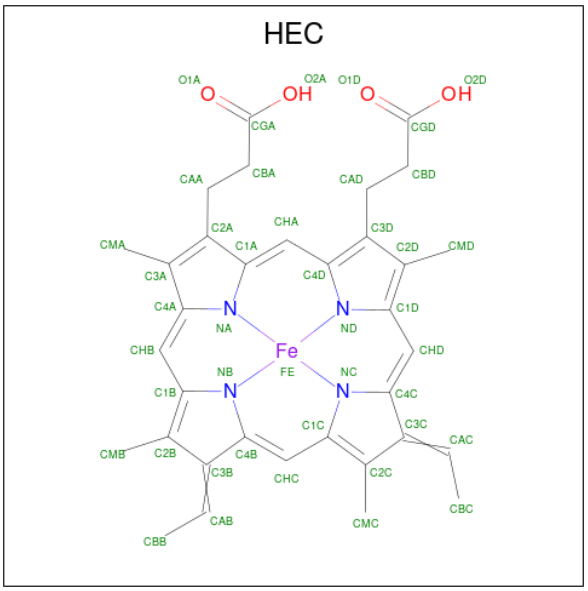
- Molecule 1 is a protein called Cytochrome c iso-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	108	1699	535	851	151	158	4	0

There is a discrepancy between the modelled and reference sequences:

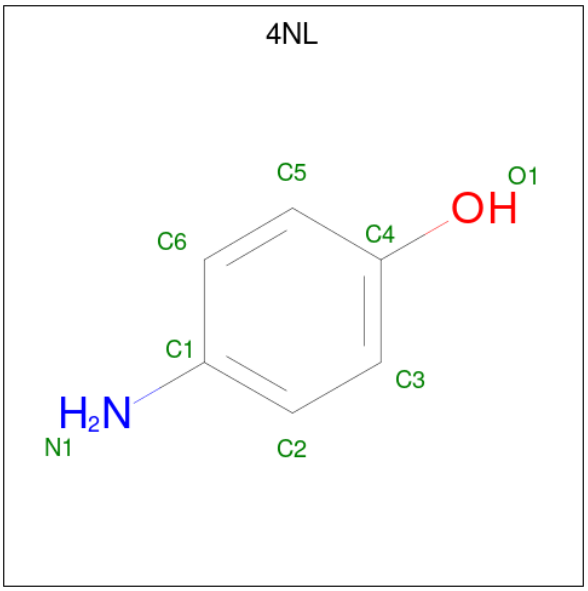
Chain	Residue	Modelled	Actual	Comment	Reference
A	102	THR	CYS	engineered mutation	UNP P00044

- Molecule 2 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).



Mol	Chain	Residues	Atoms					
			Total	C	Fe	H	N	O
2	A	1	75	34	1	32	4	4

- Molecule 3 is 4-AMINOPHENOL (three-letter code: 4NL) (formula: C_6H_7NO).



Mol	Chain	Residues	Atoms				
			Total	C	H	N	O
3	A	1	15	6	7	1	1

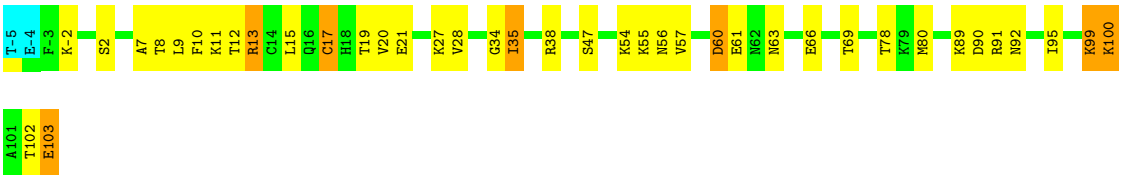
4 Residue-property plots (i)

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: Cytochrome c iso-1

Chain A: 61% 31% 6% •



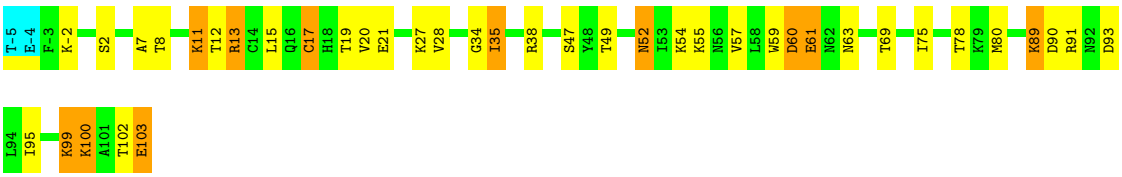
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Cytochrome c iso-1

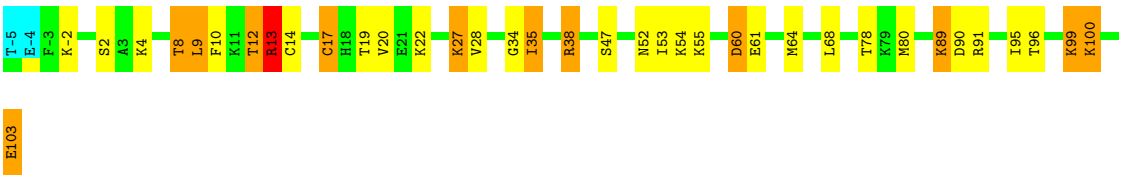
Chain A: 61% 27% 10% •



4.2.2 Score per residue for model 2

• Molecule 1: Cytochrome c iso-1

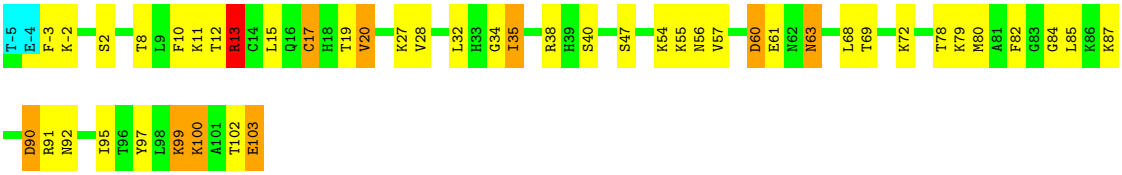
Chain A: 64% 22% 11% • •



4.2.3 Score per residue for model 3

- Molecule 1: Cytochrome c iso-1

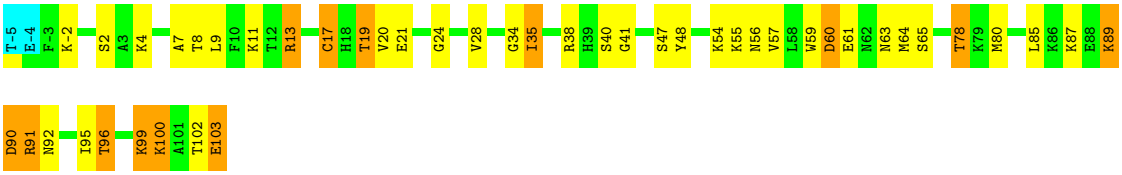
Chain A:



4.2.4 Score per residue for model 4

- Molecule 1: Cytochrome c iso-1

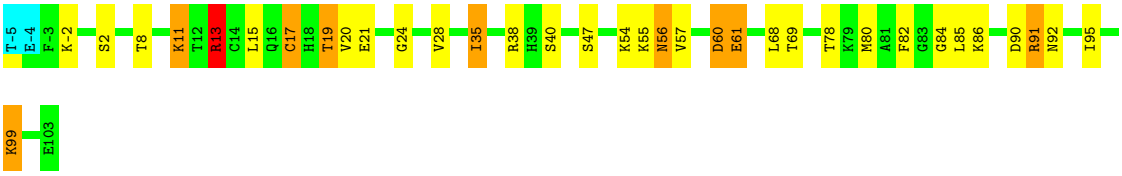
Chain A:



4.2.5 Score per residue for model 5

- Molecule 1: Cytochrome c iso-1

Chain A:



4.2.6 Score per residue for model 6

- Molecule 1: Cytochrome c iso-1

Chain A: 56% 33% 9% •

T-5 E-4 F-3 K-2 S2 A3 K4 A7 T8 L9 F10 K11 T12 R13 C14 L15 Q16 C17 H18 T19 V20 E21 G24 V28 H33 G34 I35 F36 G37 R38 S47 K54 K55 N56 V57 L58 W59 D60 E61 N62 N63 E66 M80 A81 F82 K87 E88 K89 D90 R91

N92 I95 T96 Y97 L98 K99 K100 A101 T102 E103

4.2.7 Score per residue for model 7

- Molecule 1: Cytochrome c iso-1

Chain A: 59% 31% 8% •

T-5 E-4 F-3 K-2 S2 A7 T8 L9 F10 K11 T12 R13 C14 L15 Q16 C17 H18 T19 V20 E21 V28 I35 R38 H39 A43 S47 Y48 T49 K54 K55 N56 V57 L58 W59 D60 E61 E66 T78 K79 F82 L85 D90 R91 N92 D93 L94 I95

L98 K99 K100 E103

4.2.8 Score per residue for model 8

- Molecule 1: Cytochrome c iso-1

Chain A: 61% 29% 8% •

T-5 E-4 F-3 K-2 S2 A7 T8 L9 F10 K11 T12 R13 C14 L15 Q16 C17 H18 T19 V20 E21 V28 H33 G34 I35 R38 H39 S47 K54 K55 N56 V57 D60 E61 N62 N63 E66 T69 T78 K79 M80 A81 F82 K86 D90 R91 N92

I95 K99 K100 A101 T102 E103

4.2.9 Score per residue for model 9

- Molecule 1: Cytochrome c iso-1

Chain A: 62% 27% 9% •

T-5 E-4 F-3 K-2 S2 A7 T8 L9 F10 K11 T12 R13 C14 L15 Q16 C17 H18 T19 V20 E21 K27 V28 I35 R38 H39 S40 S47 K54 K55 N56 V57 L58 W59 D60 E61 E66 T78 K79 M80 L85 K86 K87 D90 R91 N92 D93 L94 I95

L98 K99 K100 E103

4.2.10 Score per residue for model 10

- Molecule 1: Cytochrome c iso-1

Chain A: 61% 28% 9% •

T-5 E-4 F-3 K-2 S2 T8 L9 F10 R13 C14 C17 H18 T19 V20 E21 V28 H33 G34 I35 F36 G37 R38 H39 S40 G41 S47 Y48 K54 K55 N56 V57 D60 E61 S65 E66 T78 K79 M80 L85 K89 D90 R91 N92 I95 T96 Y97

L98 K99 T102 E103

4.2.11 Score per residue for model 11

- Molecule 1: Cytochrome c iso-1

Chain A: 65% 24% 9% •

T-5 E-4 F-3 K-2 S2 T8 L9 F10 K11 T12 R13 C14 C17 H18 T19 V20 E21 V28 G29 P30 H33 G34 I35 F36 G37 R38 S47 K54 K55 N56 V57 D60 E61 T78 K79 M80 K87 E88 K89 D90 R91 I95 K99 K100 A101 T102 E103

4.2.12 Score per residue for model 12 (medoid)

- Molecule 1: Cytochrome c iso-1

Chain A: 58% 31% 8% • •

T-5 E-4 S2 A7 T8 L9 F10 K11 T12 R13 C17 H18 T19 V20 E21 G24 K27 V28 G34 I35 R38 G41 S47 Y48 T49 K54 K55 N56 V57 D60 E61 N62 N63 M64 S65 E66 T78 K79 M80 K87 D90 R91 N92 D93 L94 I95

T96 Y97 L98 K99 K100 E103

4.2.13 Score per residue for model 13

- Molecule 1: Cytochrome c iso-1

Chain A: 63% 31% 5% •

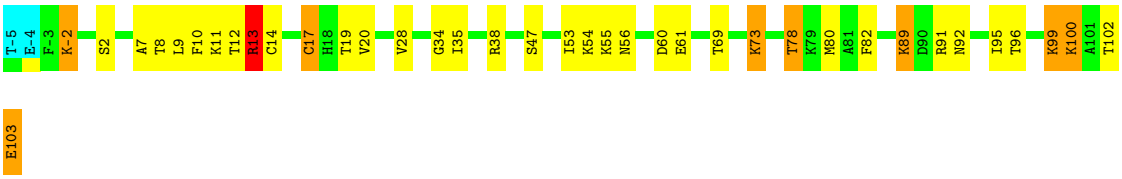
T-5 E-4 F-3 K-2 K4 K5 T8 K11 T12 R13 C14 L15 Q16 C17 H18 T19 V20 E21 K27 V28 I35 F36 G37 R38 H39 S40 S47 Y48 T49 K54 K55 N56 V57 D60 E61 N62 N63 M64 S65 E66 T78 K79 M80 D90 R91 N92 Y97 L98

T102 E103

4.2.14 Score per residue for model 14

- Molecule 1: Cytochrome c iso-1

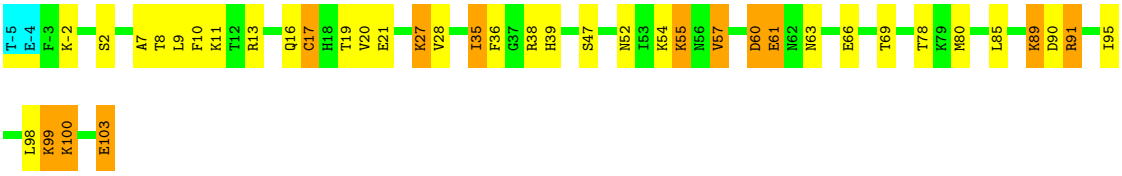
Chain A: 63% 27% 7%



4.2.15 Score per residue for model 15

- Molecule 1: Cytochrome c iso-1

Chain A: 61% 26% 11%



4.2.16 Score per residue for model 16

- Molecule 1: Cytochrome c iso-1

Chain A: 55% 33% 10%

4.2.17 Score per residue for model 17

- Molecule 1: Cytochrome c iso-1

Chain A: 67% 23% 8%

4.2.18 Score per residue for model 18

- Molecule 1: Cytochrome c iso-1

Chain A: 60% 30% 8% •

T-5 E-4 F-3 K-2 S2 A3 K4 A7 T8 L9 F10 K11 T12 R13 C14 L15 Q16 C17 H18 T19 V20 G24 V28 H33 G34 I35 F36 G37 R38 H39 S40 S47 K54 K55 N56 V57 D60 E61 T69 T78 K79 M80 K89 D90 R91 N92 I95

L98 K99 K100 A101 T102 E103

4.2.19 Score per residue for model 19

- Molecule 1: Cytochrome c iso-1

Chain A: 58% 31% 9% •

T-5 E-4 F-3 K-2 S2 A7 T8 L9 F10 K11 T12 R13 C14 L15 Q16 C17 H18 T19 V20 E21 K27 V28 H33 G34 I35 R38 S47 N52 I53 K54 K55 N56 V57 D60 E61 E66 Y67 L68 I75 T78 K79 M80 G84 K89 D90 R91 N92

I95 K99 K100 E103

4.2.20 Score per residue for model 20

- Molecule 1: Cytochrome c iso-1

Chain A: 56% 31% 10% • •

T-5 E-4 F-3 K-2 S2 A7 T8 L9 F10 K11 T12 R13 C14 C17 H18 T19 V20 E21 G24 K27 V28 H33 G34 I35 R38 G41 S47 Y48 I53 K54 K55 N56 V57 D60 E61 M64 S65 L68 P71 K72 T78 K79 M80 K87

D90 R91 I95 T96 Y97 L98 K99 K100 A101 T102 E103

4.2.21 Score per residue for model 21

- Molecule 1: Cytochrome c iso-1

Chain A: 57% 30% 11% •

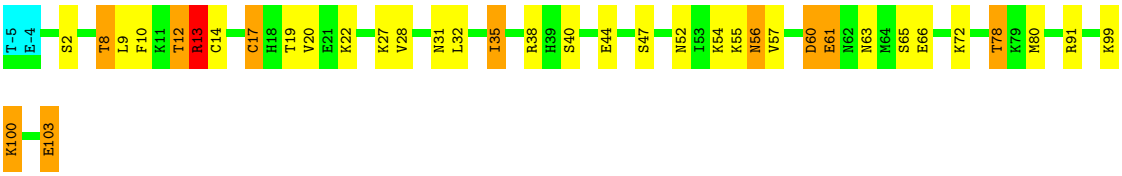
T-5 E-4 F-3 S2 A3 K4 A7 T8 L9 F10 K11 T12 R13 Q16 C17 H18 T19 V20 E21 G24 K27 V28 G34 I35 F36 G37 R38 S47 K54 K55 N56 V57 D60 E61 L68 K72 T78 K79 M80 A81 F82 G83 G84 L85 K89 D90 R91

N92 I95 L98 K99 K100 A101 T102 E103

4.2.22 Score per residue for model 22

- Molecule 1: Cytochrome c iso-1

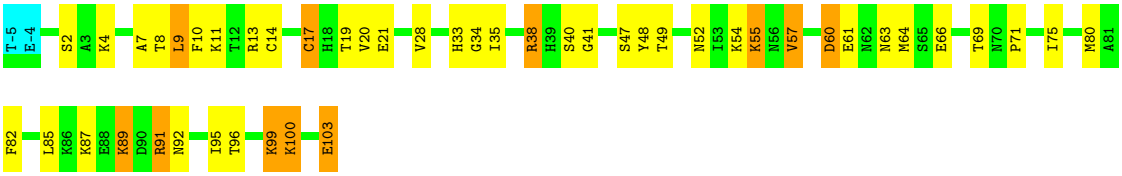
Chain A: 64% 24% 9% • •



4.2.23 Score per residue for model 23

- Molecule 1: Cytochrome c iso-1

Chain A: 55% 33% 10% •



4.2.24 Score per residue for model 24

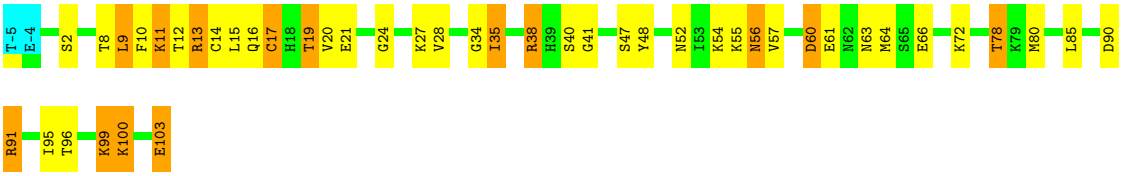
- Molecule 1: Cytochrome c iso-1

Chain A: 59% 28% 10% • •

4.2.25 Score per residue for model 25

- Molecule 1: Cytochrome c iso-1

Chain A: 56% 29% 13% •



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *combined docking and simulated annealing with NMR restraints.*

Of the 100 calculated structures, 25 were deposited, based on the following criterion: *structure with the lowest interaction energy.*

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

Software name	Classification	Version
AutoDock	structure solution	
XPLOR-NIH	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: HEC, 4NL

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	0.99±0.00	5±0/850 (0.6± 0.1%)	1.32±0.00	8±0/1136 (0.7± 0.0%)
All	All	0.99	134/21250 (0.6%)	1.32	200/28400 (0.7%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.2±0.4
All	All	0	6

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	91	ARG	CD-NE	-9.35	1.30	1.46	9	25
1	A	38	ARG	CD-NE	-9.20	1.30	1.46	15	25
1	A	91	ARG	NE-CZ	8.04	1.43	1.33	18	25
1	A	13	ARG	NE-CZ	5.94	1.40	1.33	14	25
1	A	38	ARG	CZ-NH2	-5.87	1.25	1.33	18	25
1	A	91	ARG	CZ-NH2	-5.18	1.26	1.33	13	9

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	38	ARG	NE-CZ-NH2	20.64	130.62	120.30	9	25
1	A	91	ARG	NE-CZ-NH2	13.82	127.21	120.30	12	25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	38	ARG	NE-CZ-NH1	-13.13	113.73	120.30	7	25
1	A	91	ARG	NE-CZ-NH1	-12.83	113.88	120.30	9	25
1	A	13	ARG	NE-CZ-NH2	9.80	125.20	120.30	16	25
1	A	91	ARG	CD-NE-CZ	9.68	137.16	123.60	22	25
1	A	38	ARG	CD-NE-CZ	8.35	135.29	123.60	23	25
1	A	13	ARG	NE-CZ-NH1	-7.09	116.75	120.30	9	25

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	13	ARG	Sidechain	6

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	832	836	836	17±3
2	A	43	32	30	3±1
3	A	8	7	7	0±0
All	All	22075	21875	21825	437

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 10.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:17:CYS:HB3	1:A:28:VAL:HB	0.85	1.47	17	25
1:A:100:LYS:NZ	1:A:100:LYS:HA	0.73	1.98	25	17
1:A:100:LYS:HA	1:A:100:LYS:NZ	0.72	1.99	23	5
1:A:69:THR:HG23	1:A:86:LYS:HA	0.71	1.61	8	2
1:A:82:PHE:HZ	1:A:85:LEU:HG	0.70	1.47	3	3
1:A:78:THR:HG23	1:A:80:MET:H	0.70	1.46	25	9
1:A:69:THR:HG22	1:A:86:LYS:HA	0.69	1.62	17	1
1:A:35:ILE:HD12	2:A:104:HEC:HMA3	0.67	1.64	15	14

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:11:LYS:HG3	1:A:15:LEU:HD22	0.67	1.67	3	10
1:A:80:MET:HB2	2:A:104:HEC:C1D	0.66	2.20	6	21
1:A:19:THR:HG21	1:A:24:GLY:HA3	0.66	1.67	20	9
1:A:99:LYS:O	1:A:103:GLU:HG3	0.65	1.92	23	22
1:A:10:PHE:HA	1:A:14:CYS:SG	0.65	2.31	22	7
1:A:34:GLY:HA2	1:A:102:THR:HG23	0.65	1.69	20	10
1:A:10:PHE:CD1	1:A:14:CYS:HB2	0.64	2.27	19	2
1:A:-3:PHE:HA	1:A:92:ASN:OD1	0.64	1.92	3	6
1:A:10:PHE:CZ	1:A:20:VAL:HB	0.63	2.28	6	6
1:A:13:ARG:CZ	1:A:13:ARG:HA	0.62	2.25	14	4
1:A:63:ASN:O	1:A:66:GLU:HG2	0.59	1.97	15	3
1:A:100:LYS:HA	1:A:100:LYS:HZ2	0.58	1.56	21	10
1:A:68:LEU:O	1:A:84:GLY:HA2	0.58	1.98	3	4
1:A:36:PHE:HB2	1:A:102:THR:HG21	0.58	1.75	10	5
1:A:82:PHE:CE2	2:A:104:HEC:HBB2	0.57	2.35	8	2
1:A:80:MET:HB2	2:A:104:HEC:ND	0.56	2.15	20	20
1:A:82:PHE:CZ	1:A:85:LEU:HG	0.56	2.33	3	6
1:A:10:PHE:CD2	1:A:14:CYS:HB2	0.56	2.36	24	1
1:A:7:ALA:O	1:A:11:LYS:HB2	0.55	2.00	21	15
1:A:34:GLY:O	1:A:38:ARG:HD3	0.55	2.02	23	7
1:A:85:LEU:O	1:A:91:ARG:HD3	0.54	2.03	4	5
1:A:17:CYS:O	1:A:27:LYS:HB3	0.53	2.03	3	3
1:A:-2:LYS:O	1:A:92:ASN:HB3	0.53	2.03	16	9
1:A:13:ARG:HA	1:A:13:ARG:NE	0.53	2.19	10	11
1:A:59:TRP:CZ2	2:A:104:HEC:HAA1	0.52	2.39	1	3
1:A:95:ILE:O	1:A:99:LYS:HB3	0.52	2.04	20	23
1:A:70:ASN:HB3	1:A:73:LYS:HB3	0.52	1.81	17	1
1:A:41:GLY:HA2	1:A:48:TYR:CG	0.50	2.40	20	4
1:A:20:VAL:HG21	1:A:97:TYR:CE2	0.50	2.42	12	6
1:A:8:THR:O	1:A:12:THR:HB	0.50	2.07	22	2
1:A:87:LYS:HB2	1:A:90:ASP:HB2	0.50	1.82	20	5
1:A:63:ASN:HA	1:A:66:GLU:CD	0.50	2.27	8	7
1:A:69:THR:HG22	1:A:86:LYS:HG2	0.49	1.84	24	1
1:A:52:ASN:HA	1:A:75:ILE:HD13	0.48	1.85	19	3
1:A:79:LYS:HG2	2:A:104:HEC:O2D	0.48	2.08	16	2
1:A:64:MET:HE3	1:A:95:ILE:HG12	0.48	1.85	12	2
1:A:79:LYS:HG3	2:A:104:HEC:O2D	0.47	2.10	18	7
1:A:64:MET:O	1:A:68:LEU:HG	0.47	2.10	20	2
1:A:11:LYS:HA	1:A:15:LEU:HB2	0.47	1.87	13	3
1:A:13:ARG:NE	1:A:13:ARG:HA	0.47	2.24	20	2
1:A:39:HIS:HA	1:A:57:VAL:O	0.47	2.09	8	5

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:71:PRO:HB3	1:A:80:MET:HE3	0.47	1.86	23	1
1:A:60:ASP:CG	1:A:61:GLU:N	0.46	2.69	22	22
1:A:89:LYS:O	1:A:92:ASN:HB2	0.46	2.11	23	1
1:A:-2:LYS:HA	1:A:-2:LYS:HE3	0.46	1.86	14	1
1:A:27:LYS:HZ3	1:A:27:LYS:HB2	0.46	1.69	2	1
1:A:27:LYS:HG2	1:A:28:VAL:H	0.46	1.69	24	3
1:A:13:ARG:HH11	1:A:82:PHE:HE1	0.46	1.54	14	1
1:A:99:LYS:HB2	3:A:105:4NL:O1	0.45	2.12	4	3
1:A:57:VAL:HG13	1:A:63:ASN:OD1	0.45	2.11	3	2
1:A:59:TRP:HA	1:A:63:ASN:OD1	0.45	2.12	16	4
1:A:41:GLY:HA2	1:A:48:TYR:CD2	0.44	2.48	23	1
1:A:103:GLU:HG2	3:A:105:4NL:H2	0.44	1.87	14	1
1:A:52:ASN:O	1:A:55:LYS:HB3	0.44	2.13	16	3
1:A:71:PRO:HG3	1:A:80:MET:HE2	0.44	1.89	20	1
1:A:27:LYS:HB2	1:A:27:LYS:HE3	0.43	1.59	12	4
1:A:36:PHE:HB3	3:A:105:4NL:C3	0.43	2.43	13	2
1:A:73:LYS:HE2	1:A:73:LYS:HA	0.43	1.90	14	1
1:A:100:LYS:HA	1:A:100:LYS:HZ3	0.43	1.72	19	1
1:A:91:ARG:O	1:A:95:ILE:HG13	0.43	2.13	21	2
1:A:41:GLY:HA2	1:A:48:TYR:CD1	0.43	2.49	10	2
1:A:20:VAL:HG21	1:A:97:TYR:HE2	0.43	1.73	13	1
1:A:10:PHE:HD1	1:A:14:CYS:HB2	0.43	1.73	10	3
1:A:96:THR:O	1:A:100:LYS:HB2	0.42	2.14	4	1
1:A:10:PHE:CZ	1:A:32:LEU:HD12	0.42	2.50	3	1
1:A:36:PHE:HE2	1:A:98:LEU:HB3	0.42	1.74	15	2
1:A:87:LYS:HD2	1:A:90:ASP:OD1	0.42	2.14	6	1
1:A:10:PHE:CE2	1:A:32:LEU:HD12	0.42	2.50	22	1
1:A:18:HIS:CE1	1:A:30:PRO:HD2	0.41	2.50	11	1
1:A:9:LEU:HD13	1:A:9:LEU:HA	0.41	1.73	7	3
1:A:35:ILE:HD12	2:A:104:HEC:CMA	0.41	2.42	15	1
1:A:87:LYS:HD2	1:A:90:ASP:CG	0.41	2.36	12	3
1:A:27:LYS:HB2	1:A:27:LYS:NZ	0.41	2.30	2	1
1:A:53:ILE:O	1:A:55:LYS:N	0.41	2.53	2	3
1:A:11:LYS:HG2	1:A:15:LEU:HD22	0.41	1.93	17	2
1:A:43:ALA:H	1:A:48:TYR:HE2	0.41	1.57	7	1
1:A:80:MET:O	1:A:80:MET:CG	0.40	2.69	20	1
1:A:89:LYS:H	1:A:89:LYS:HG2	0.40	1.56	4	1
1:A:60:ASP:HA	3:A:105:4NL:H4	0.40	1.93	17	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	105/108 (97%)	90±2 (86±2%)	12±2 (11±2%)	3±1 (3±1%)	8 43
All	All	2625/2700 (97%)	2257 (86%)	297 (11%)	71 (3%)	8 43

All 6 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	54	LYS	25
1	A	56	ASN	18
1	A	9	LEU	15
1	A	33	HIS	9
1	A	27	LYS	3
1	A	44	GLU	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	87/89 (98%)	65±2 (75±3%)	22±2 (25±3%)	2 25
All	All	2175/2225 (98%)	1632 (75%)	543 (25%)	2 25

All 49 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	8	THR	25
1	A	19	THR	25
1	A	20	VAL	25
1	A	35	ILE	25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	47	SER	25
1	A	60	ASP	25
1	A	2	SER	24
1	A	17	CYS	23
1	A	78	THR	23
1	A	99	LYS	23
1	A	103	GLU	22
1	A	57	VAL	21
1	A	90	ASP	20
1	A	100	LYS	20
1	A	21	GLU	19
1	A	55	LYS	17
1	A	61	GLU	14
1	A	89	LYS	13
1	A	-2	LYS	12
1	A	12	THR	12
1	A	11	LYS	11
1	A	40	SER	10
1	A	69	THR	9
1	A	4	LYS	9
1	A	98	LEU	9
1	A	49	THR	7
1	A	96	THR	7
1	A	65	SER	7
1	A	93	ASP	6
1	A	56	ASN	6
1	A	13	ARG	5
1	A	72	LYS	5
1	A	10	PHE	5
1	A	66	GLU	5
1	A	52	ASN	4
1	A	27	LYS	4
1	A	16	GLN	4
1	A	9	LEU	2
1	A	22	LYS	2
1	A	64	MET	2
1	A	91	ARG	2
1	A	87	LYS	2
1	A	63	ASN	1
1	A	80	MET	1
1	A	85	LEU	1
1	A	102	THR	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	5	LYS	1
1	A	73	LYS	1
1	A	31	ASN	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)

2 ligands are 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	HEC	A	104	1	26,50,50	1.06±0.00	2±0 (7±0%)
3	4NL	A	105	-	8,8,8	0.99±0.01	0±0 (0±0%)

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	HEC	A	104	1	18,82,82	1.12±0.01	1±0 (5±0%)
3	4NL	A	105	-	10,10,10	0.50±0.01	0±0 (0±0%)

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
3	4NL	A	105	-	-	-	0±0,1,1,1
2	HEC	A	104	1	-	0±0,6,54,54	-

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	104	HEC	C4D-ND	2.79	1.41	1.36	12	25
2	A	104	HEC	C1D-ND	2.14	1.40	1.36	18	25

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	A	104	HEC	CMD-C2D-C1D	2.65	124.39	128.46	8	25

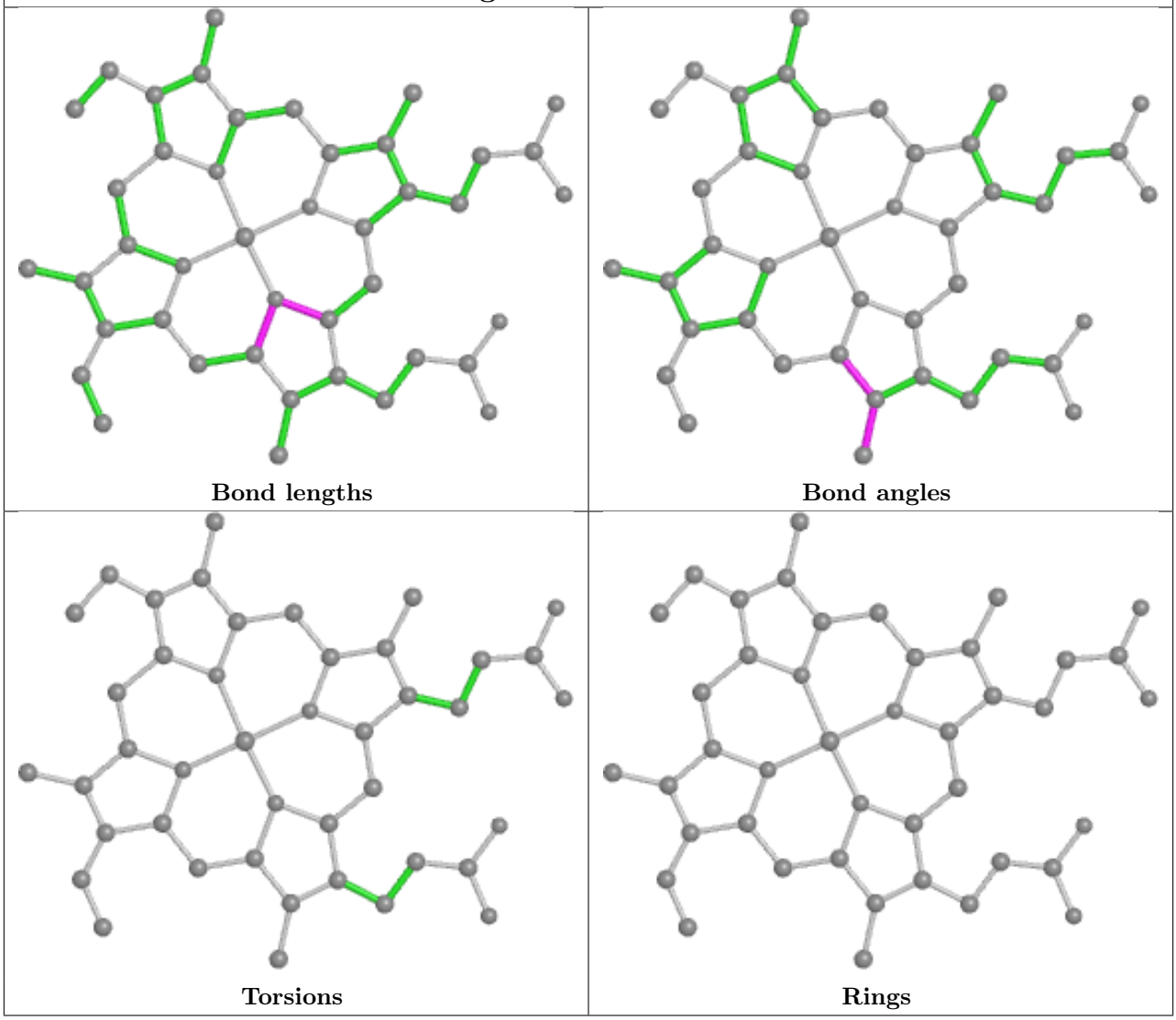
There are no chirality outliers.

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 HEC A 104



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