wwPDB X-ray Structure Validation Summary Report (i)

May 21, 2020 – 10:31 pm BST

PDB ID	:	1CPC
Title	:	ISOLATION, CRYSTALLIZATION, CRYSTAL STRUCTURE ANALYSIS AND REFINEMENT OF CONSTITUTIVE C-PHYCOCYANIN FROM THE CHROMATICALLY ADAPTING CYANOBACTERIUM FREMYELLA DIPLOSIPHON AT 1.66 ANGSTROMS RESOLUTION
Authors	:	Duerring, M.; Schmidt, G.B.; Huber, R.
Deposited on	:	1990-10-11
Resolution	:	1.66 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

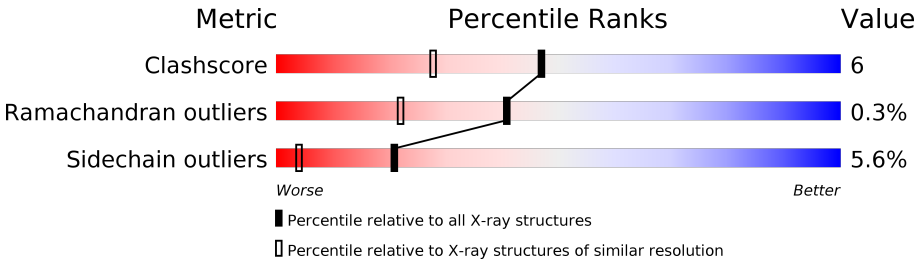
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.66 Å.

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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	162	83% 14%
1	K	162	81% 17%
2	B	172	84% 13%
2	L	172	78% 19%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5434 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called C-PHYCOCYANIN (ALPHA SUBUNIT).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	162	Total	C	N	O	S	11	0	0
			1214	766	203	241	4			
1	K	162	Total	C	N	O	S	14	0	0
			1214	766	203	241	4			

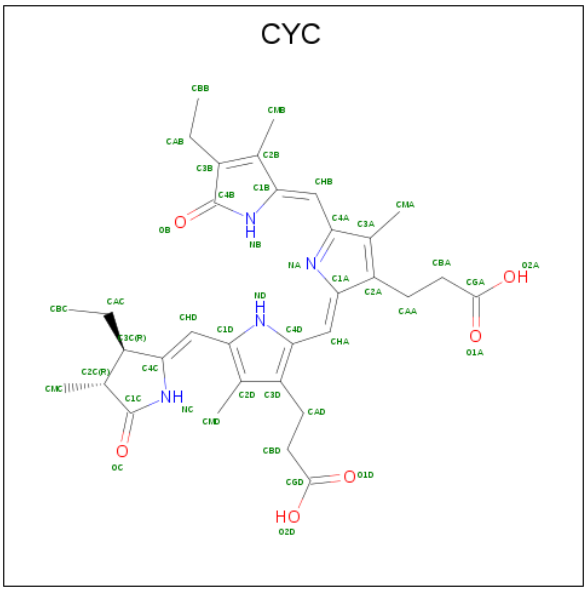
- Molecule 2 is a protein called C-PHYCOCYANIN (BETA SUBUNIT).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	23	0	0
			1253	775	220	249	9			
2	L	172	Total	C	N	O	S	26	0	0
			1252	775	220	248	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	171	SER	SER	CONFLICT	UNP P07119

- Molecule 3 is PHYCOCYANOBILIN (three-letter code: CYC) (formula: $C_{33}H_{40}N_4O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 43 33 4 6	0	0
3	B	1	Total C N O 43 33 4 6	0	0
3	B	1	Total C N O 43 33 4 6	0	0
3	K	1	Total C N O 43 33 4 6	0	0
3	L	1	Total C N O 43 33 4 6	0	0
3	L	1	Total C N O 43 33 4 6	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	75	Total O 75 75	0	0
4	B	58	Total O 58 58	0	0
4	K	57	Total O 57 57	0	0
4	L	53	Total O 53 53	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-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 outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

- Molecule 1: C-PHYCOCYANIN (ALPHA SUBUNIT)

Chain A:

- Molecule 1: C-PHYCOCYANIN (ALPHA SUBUNIT)

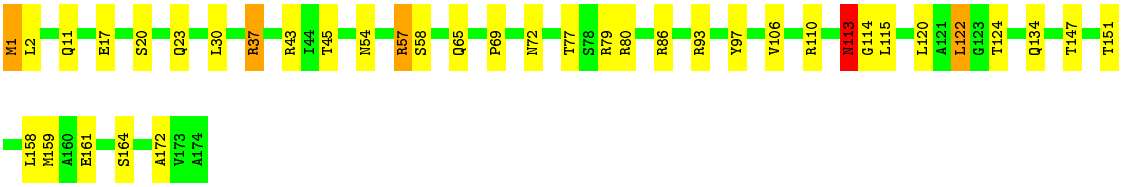
Chain K:

- Molecule 2: C-PHYCOCYANIN (BETA SUBUNIT)

Chain B:

- Molecule 2: C-PHYCOCYANIN (BETA SUBUNIT)

Chain L:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	180.26Å 180.26Å 61.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	(Not available) – 1.66	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-1.66)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	EREF	Depositor
R, R_{free}	0.181 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5434	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CYC, MEN

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 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	1.08	2/1238 (0.2%)	1.46	13/1680 (0.8%)
1	K	1.03	2/1238 (0.2%)	1.37	11/1680 (0.7%)
2	B	1.01	0/1255	1.54	12/1694 (0.7%)
2	L	0.99	0/1254	1.41	11/1693 (0.6%)
All	All	1.03	4/4985 (0.1%)	1.45	47/6747 (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 outliers	#Planarity outliers
1	A	0	3
1	K	0	2
2	L	1	8
All	All	1	13

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	164	SER	CB-OG	-6.12	1.34	1.42
1	K	128	TRP	NE1-CE2	-5.82	1.29	1.37
1	A	128	TRP	NE1-CE2	-5.69	1.30	1.37
1	K	120	ARG	CZ-NH1	5.49	1.40	1.33

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	120	ARG	NE-CZ-NH1	14.71	127.65	120.30
2	B	110	ARG	NE-CZ-NH1	14.35	127.48	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	K	120	ARG	NE-CZ-NH2	-13.31	113.64	120.30
2	B	152	ARG	NE-CZ-NH1	12.64	126.62	120.30
1	A	120	ARG	NE-CZ-NH2	-11.90	114.35	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	L	124	THR	CB

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	113	GLY	Mainchain
1	A	120	ARG	Sidechain
1	A	15	GLN	Mainchain
1	K	35	SER	Mainchain
1	K	93	ARG	Mainchain

5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1214	0	1190	19	0
1	K	1214	0	1190	18	0
2	B	1253	0	1257	15	1
2	L	1252	0	1254	12	1
3	A	43	0	37	1	0
3	B	86	0	74	5	0
3	K	43	0	37	1	0
3	L	86	0	74	3	0
4	A	75	0	0	1	1
4	B	58	0	0	0	0
4	K	57	0	0	0	1
4	L	53	0	0	0	0
All	All	5434	0	5113	63	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including

hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 63 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:15:GLN:HG3	1:A:17:ARG:HD3	1.24	1.19
1:K:73:ASN:HD22	1:K:73:ASN:H	1.17	0.88
2:B:20:SER:H	2:B:23:GLN:HE21	1.24	0.85
1:A:73:ASN:HD22	1:A:73:ASN:H	1.22	0.84
2:L:20:SER:H	2:L:23:GLN:HE21	1.38	0.71

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:57:ARG:NH1	4:A:216:HOH:O[3_555]	2.05	0.15
2:L:57:ARG:NH2	4:K:221:HOH:O[2_555]	2.15	0.05

5.3 Torsion angles (i)

5.3.1 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 X-ray entries followed by that with respect to entries of similar resolution.

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	160/162 (99%)	159 (99%)	1 (1%)	0	100 100
1	K	160/162 (99%)	157 (98%)	3 (2%)	0	100 100
2	B	169/172 (98%)	165 (98%)	3 (2%)	1 (1%)	25 8
2	L	169/172 (98%)	167 (99%)	1 (1%)	1 (1%)	25 8
All	All	658/668 (98%)	648 (98%)	8 (1%)	2 (0%)	41 22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	77	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	L	77	THR

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	125/125 (100%)	122 (98%)	3 (2%)	49 23
1	K	125/125 (100%)	118 (94%)	7 (6%)	21 4
2	B	126/126 (100%)	120 (95%)	6 (5%)	25 6
2	L	125/126 (99%)	113 (90%)	12 (10%)	8 1
All	All	501/502 (100%)	473 (94%)	28 (6%)	21 4

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	47	LYS
1	K	127	SER
2	L	147	THR
1	K	73	ASN
1	K	92	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	57	ASN
1	K	69	GLN
2	L	54	ASN
1	K	25	GLN
1	K	33	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or 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 length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	MEN	B	72	2	7,8,9	1.00	1 (14%)	6,9,11	0.71	0
2	MEN	L	72	2	7,8,9	0.84	0	6,9,11	2.16	3 (50%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MEN	B	72	2	-	2/7/8/10	-
2	MEN	L	72	2	-	1/7/8/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	72	MEN	CB-CA	2.09	1.58	1.53

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	L	72	MEN	OD1-CG-CB	-3.00	117.10	121.50
2	L	72	MEN	CB-CA-C	2.97	117.03	111.47
2	L	72	MEN	CB-CG-ND2	2.93	119.42	115.48

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	72	MEN	CA-CB-CG-OD1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	B	72	MEN	CA-CB-CG-ND2
2	L	72	MEN	CA-CB-CG-OD1

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	72	MEN	1	0
2	L	72	MEN	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or 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 length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	CYC	B	177	2	36,46,46	1.91	9 (25%)	44,67,67	1.82	9 (20%)
3	CYC	B	176	2	36,46,46	2.06	6 (16%)	44,67,67	2.08	7 (15%)
3	CYC	L	177	2	36,46,46	2.18	11 (30%)	44,67,67	1.78	11 (25%)
3	CYC	A	175	1	36,46,46	1.96	7 (19%)	44,67,67	1.72	7 (15%)
3	CYC	L	176	2	36,46,46	1.84	8 (22%)	44,67,67	2.05	8 (18%)
3	CYC	K	175	1	36,46,46	1.94	8 (22%)	44,67,67	1.87	15 (34%)

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	CYC	B	177	2	-	6/21/74/74	0/4/4/4
3	CYC	B	176	2	-	7/21/74/74	0/4/4/4
3	CYC	L	177	2	-	8/21/74/74	0/4/4/4
3	CYC	A	175	1	-	10/21/74/74	0/4/4/4
3	CYC	L	176	2	-	6/21/74/74	0/4/4/4
3	CYC	K	175	1	-	7/21/74/74	0/4/4/4

The worst 5 of 49 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	177	CYC	CHA-C1A	8.79	1.42	1.35
3	B	176	CYC	CHA-C1A	7.50	1.41	1.35
3	A	175	CYC	CHA-C1A	7.29	1.41	1.35
3	K	175	CYC	CHA-C1A	7.27	1.41	1.35
3	L	176	CYC	CHA-C1A	6.31	1.40	1.35

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	L	176	CYC	OC-C1C-C2C	-6.55	120.96	126.17
3	B	176	CYC	C2C-C1C-NC	6.51	113.89	108.27
3	B	176	CYC	OC-C1C-C2C	-6.27	121.19	126.17
3	L	176	CYC	C2C-C1C-NC	5.86	113.33	108.27
3	B	177	CYC	C2C-C1C-NC	5.15	112.71	108.27

There are no chirality outliers.

5 of 44 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	177	CYC	ND-C4D-CHA-C1A
3	B	177	CYC	NA-C4A-CHB-C1B
3	B	177	CYC	C3A-C4A-CHB-C1B
3	B	177	CYC	C2C-C3C-CAC-CBC
3	B	176	CYC	ND-C4D-CHA-C1A

There are no ring outliers.

6 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	177	CYC	1	0

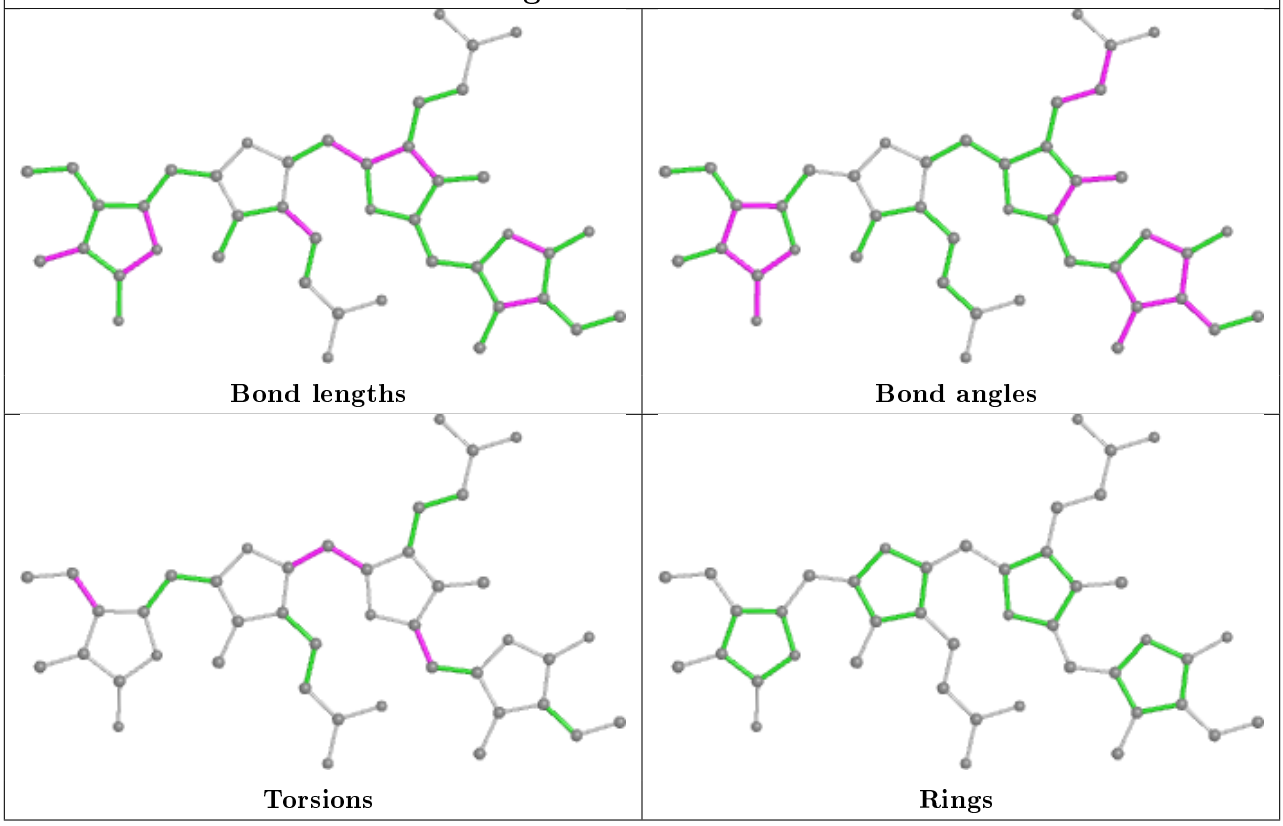
Continued on next page...

Continued from previous page...

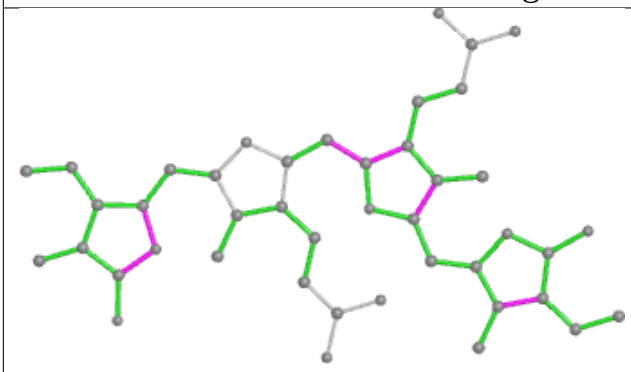
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	176	CYC	4	0
3	L	177	CYC	2	0
3	A	175	CYC	1	0
3	L	176	CYC	1	0
3	K	175	CYC	1	0

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.

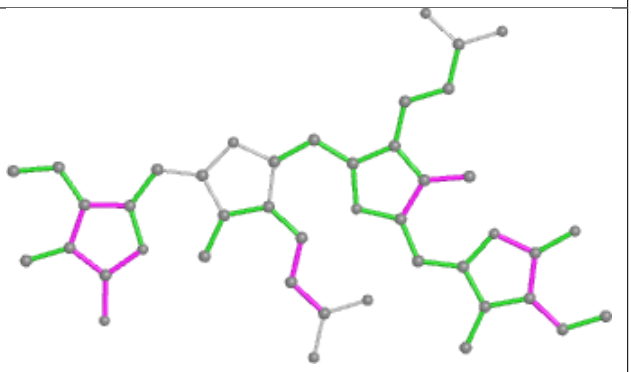
Ligand CYC B 177



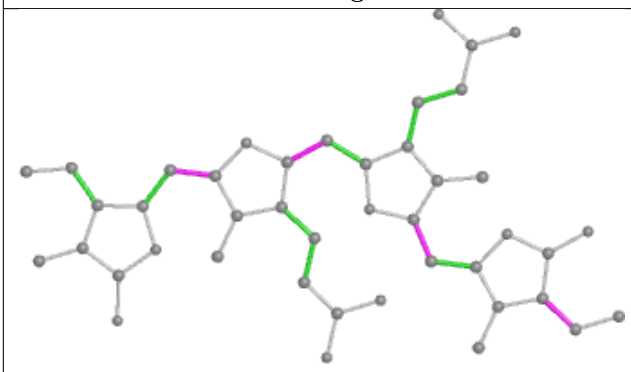
Ligand CYC B 176



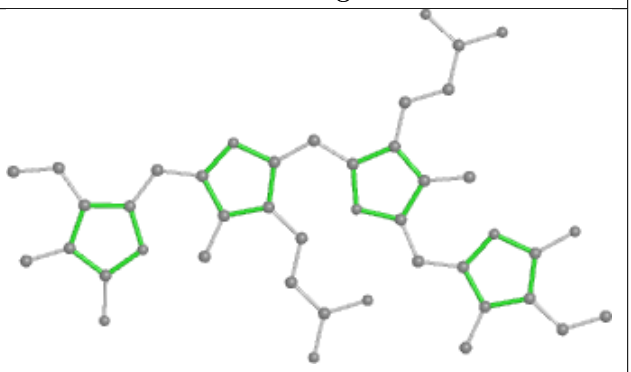
Bond lengths



Bond angles

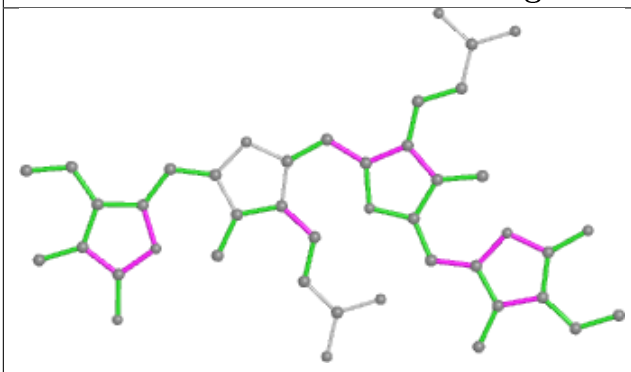


Torsions

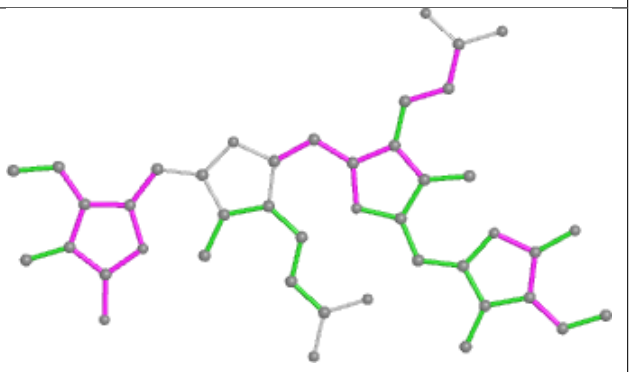


Rings

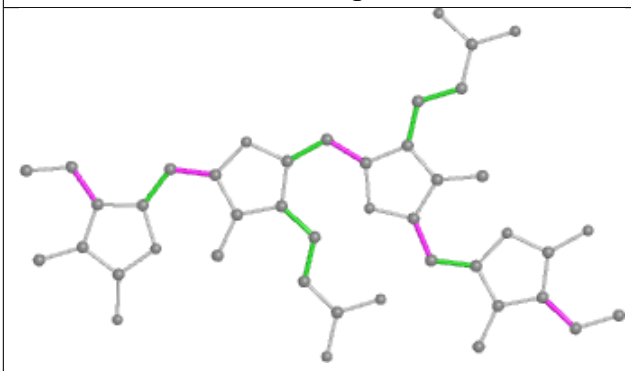
Ligand CYC L 177



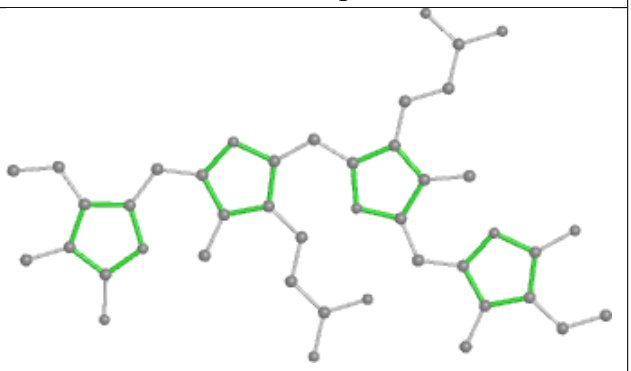
Bond lengths



Bond angles

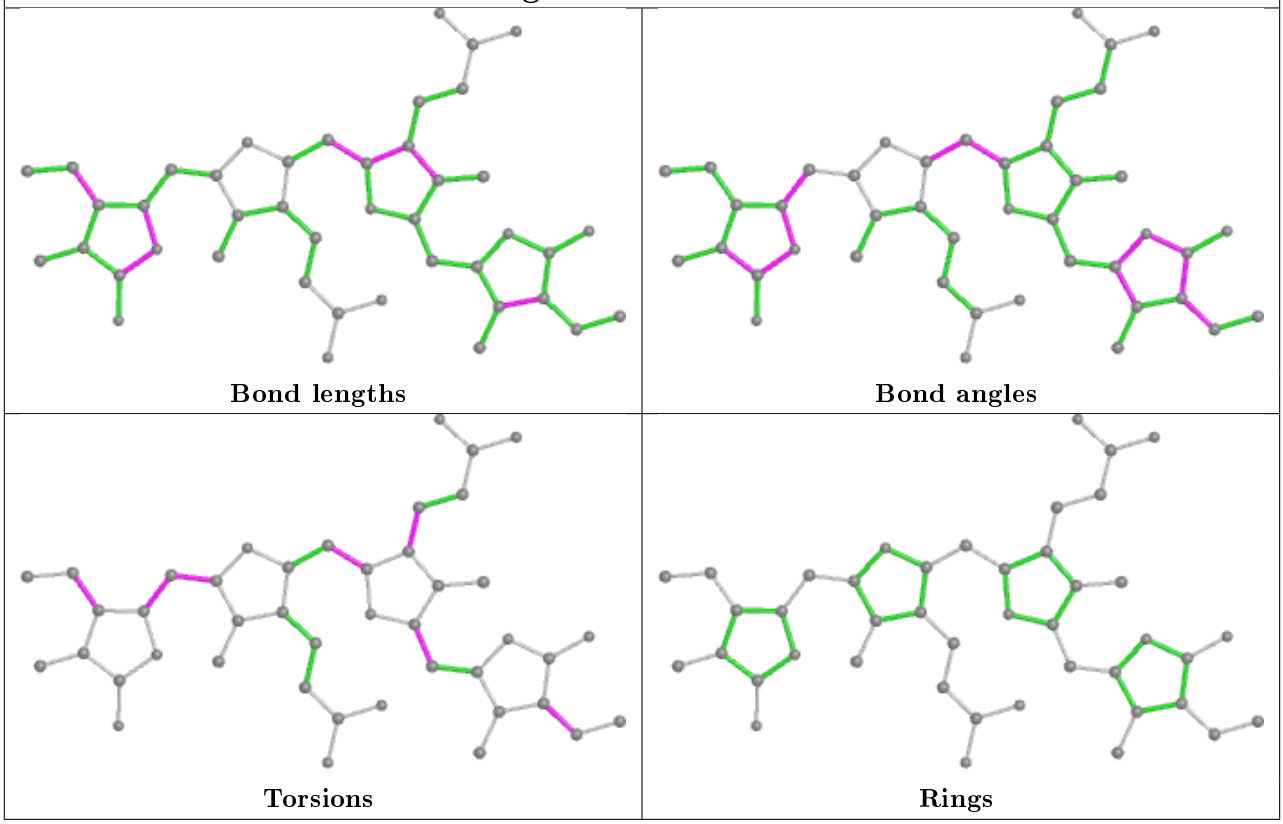


Torsions

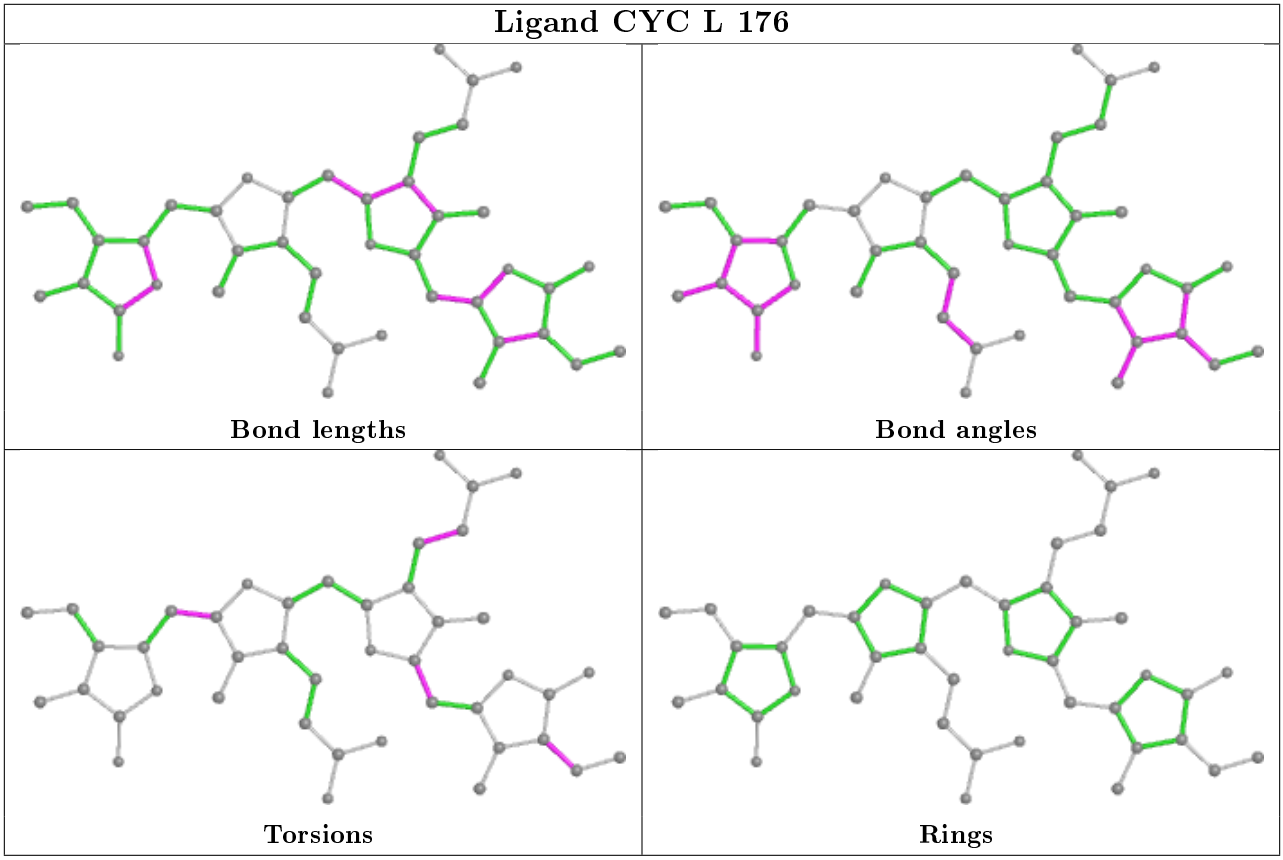


Rings

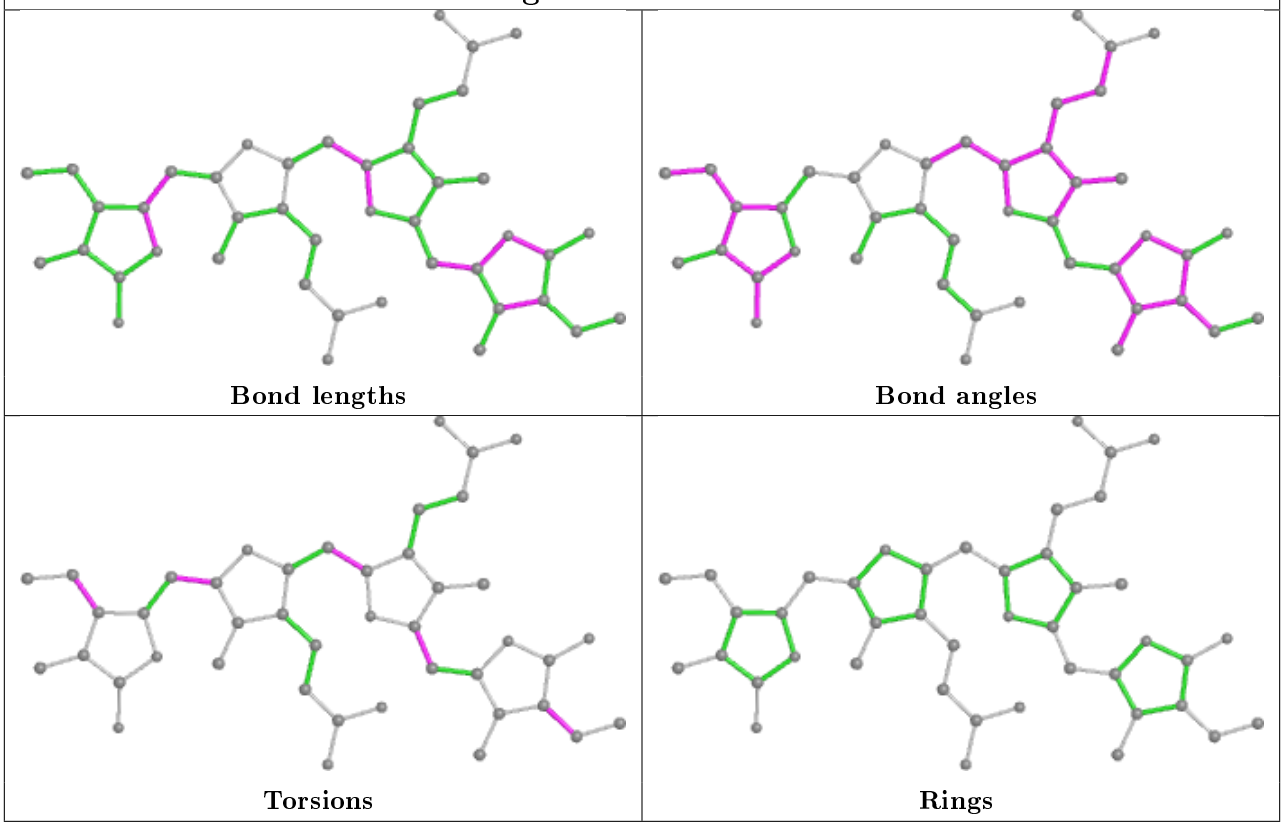
Ligand CYC A 175



Ligand CYC L 176



Ligand CYC K 175



5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.