wwPDB X-ray Structure Validation Summary Report (i)

Feb 22, 2024 – 11:11 AM EST

PDB ID	:	4Y49
Title	:	Crystal structure of yeast N-terminal acetyltransferase (ppGpp) NatE in complex with a bisubstrate
Authors	:	Dong, J.; Wang, S.; York, J.D.
Deposited on	:	2015-02-10
Resolution	:	3.95 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 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:

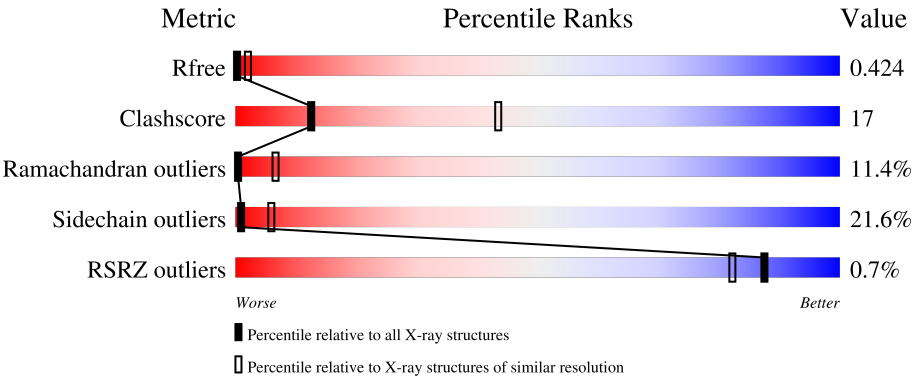
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.95 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))
R_{free}	130704	1025 (4.22-3.70)
Clashscore	141614	1085 (4.22-3.70)
Ramachandran outliers	138981	1047 (4.22-3.70)
Sidechain outliers	138945	1039 (4.22-3.70)
RSRZ outliers	127900	1013 (4.28-3.64)

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 of 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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	854	60% 23% 6% • 10%
1	G	854	60% 23% 7% • 10%
1	M	854	60% 23% 6% • 10%
2	B	238	40% 22% 6% • 31%
2	H	238	44% 21% • • 31%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	N	238	42% 23% 31%
3	C	176	% 63% 19% 6% 11%
3	I	176	58% 23% 6% 12%
3	O	176	66% 18% 13%
4	E	8	75% 25%
4	K	8	12% 75% 25%
4	Q	8	75% 25%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	G4P	G	901	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 20122 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 N-terminal acetyltransferase A complex subunit NAT1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	769	Total	C	N	O	S	0	0	0
			4636	2891	860	877	8			
1	G	770	Total	C	N	O	S	0	0	0
			4656	2910	855	881	10			
1	M	769	Total	C	N	O	S	0	0	0
			4624	2889	848	877	10			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	31	TYR	GLU	conflict	UNP P12945
G	31	TYR	GLU	conflict	UNP P12945
M	31	TYR	GLU	conflict	UNP P12945

- Molecule 2 is a protein called N-terminal acetyltransferase A complex catalytic subunit ARD1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	164	Total	C	N	O	S	0	0	0
			1024	640	188	188	8			
2	H	164	Total	C	N	O	S	0	0	0
			1025	635	190	192	8			
2	N	164	Total	C	N	O	S	0	0	0
			1025	636	189	192	8			

- Molecule 3 is a protein called N-terminal acetyltransferase A complex subunit NAT5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	156	Total	C	N	O	S	4	0	0
			888	555	168	163	2			
3	I	155	Total	C	N	O	S	3	0	0
			879	545	167	165	2			

Continued on next page...

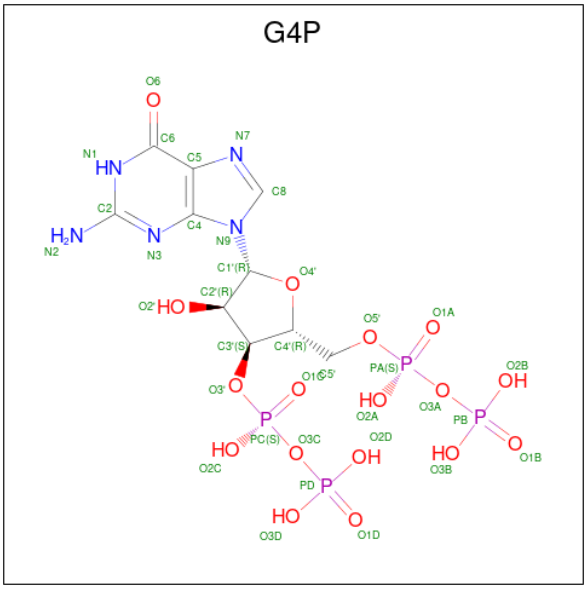
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	O	154	Total	C	N	O	S	1	1	0
			861	532	165	162	2			

- Molecule 4 is a protein called ALA-ALA-ALA-ALA-ALA-ALA.

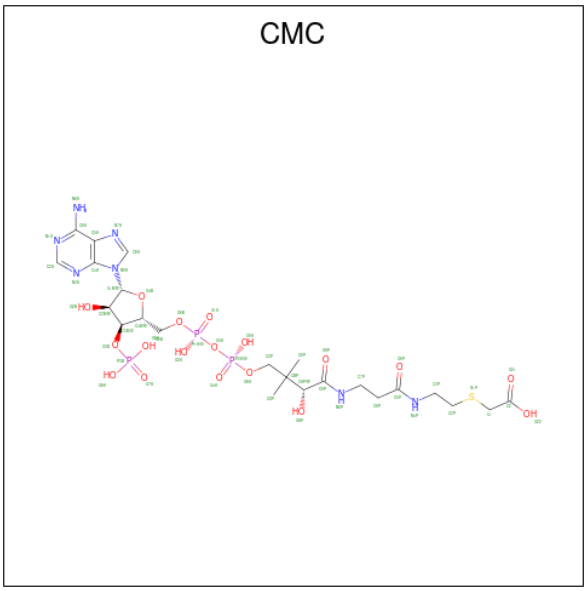
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	6	Total	C	N	O	0	0	0
			30	18	6	6			
4	K	6	Total	C	N	O	0	0	0
			30	18	6	6			
4	Q	6	Total	C	N	O	0	0	0
			30	18	6	6			

- Molecule 5 is GUANOSINE-5',3'-TETRAPHOSPHATE (three-letter code: G4P) (formula: $C_{10}H_{17}N_5O_{17}P_4$).



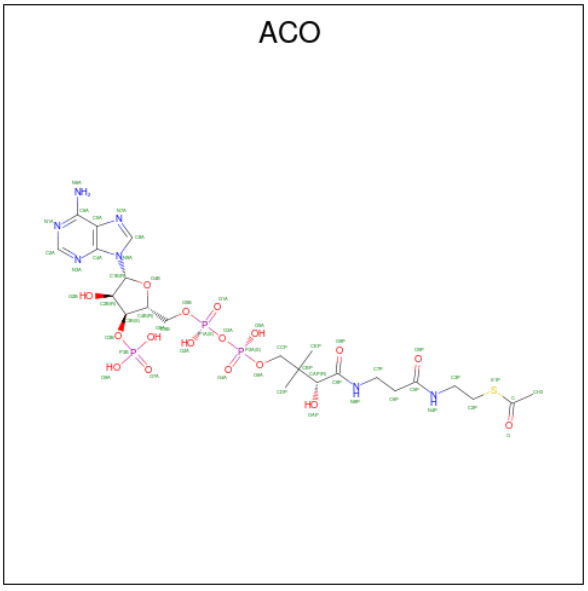
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			36	10	5	17	4		
5	G	1	Total	C	N	O	P	0	0
			36	10	5	17	4		
5	M	1	Total	C	N	O	P	0	0
			36	10	5	17	4		

- Molecule 6 is CARBOXYMETHYL COENZYME *A (three-letter code: CMC) (formula: $C_{23}H_{38}N_7O_{18}P_3S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	B	1	Total	C	N	O	P	S	51	0
			51	23	7	17	3	1		
6	H	1	Total	C	N	O	P	S	51	0
			51	23	7	17	3	1		
6	N	1	Total	C	N	O	P	S	51	0
			51	23	7	17	3	1		

- Molecule 7 is ACETYL COENZYME *A (three-letter code: ACO) (formula: $C_{23}H_{38}N_7O_{17}P_3S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
7	C	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

Continued on next page...

Continued from previous page...

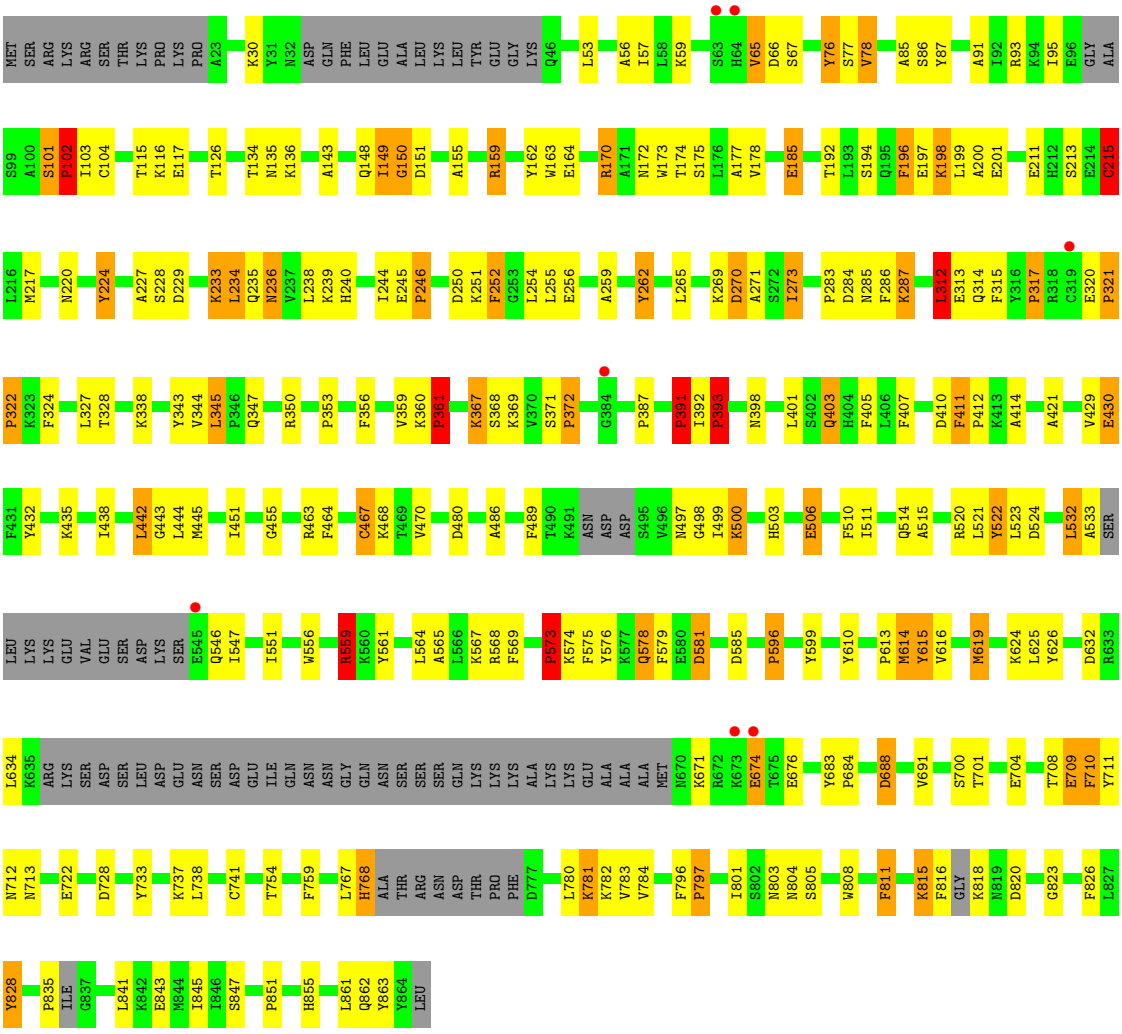
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
7	I	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
7	O	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

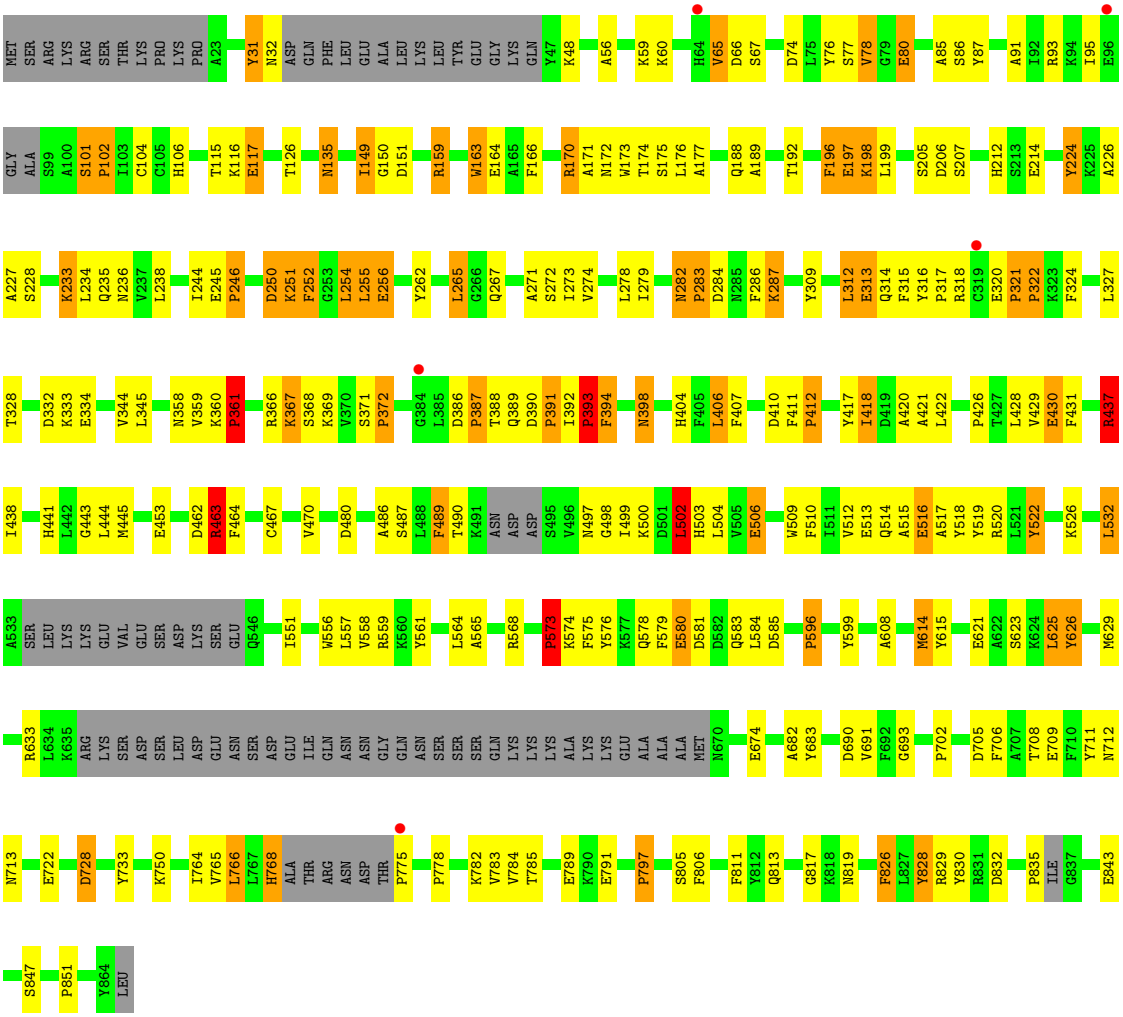
• Molecule 1: N-terminal acetyltransferase A complex subunit NAT1

Chain A:



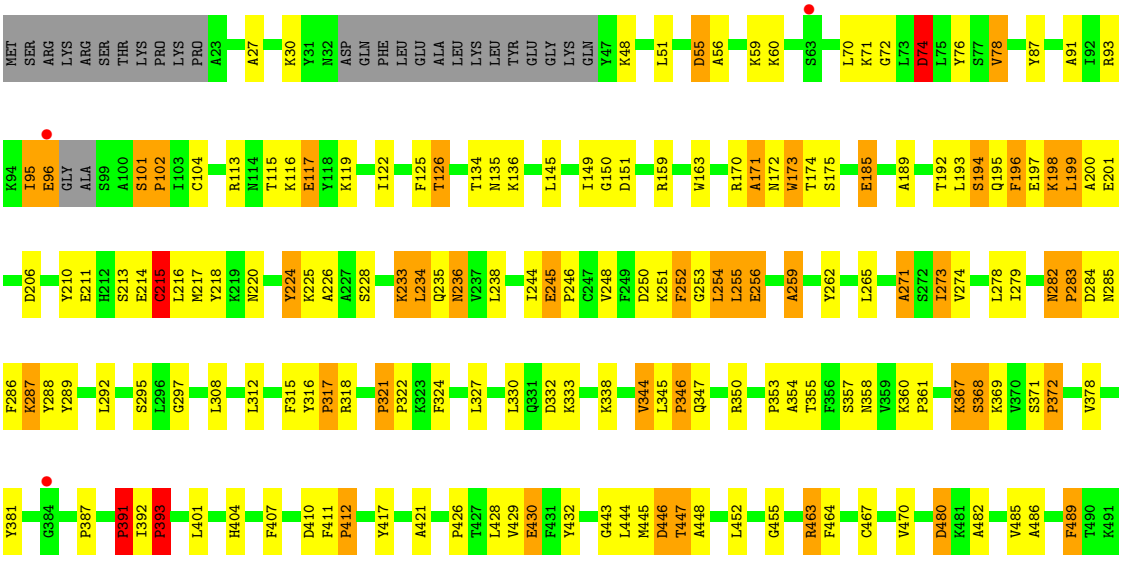
• Molecule 1: N-terminal acetyltransferase A complex subunit NAT1

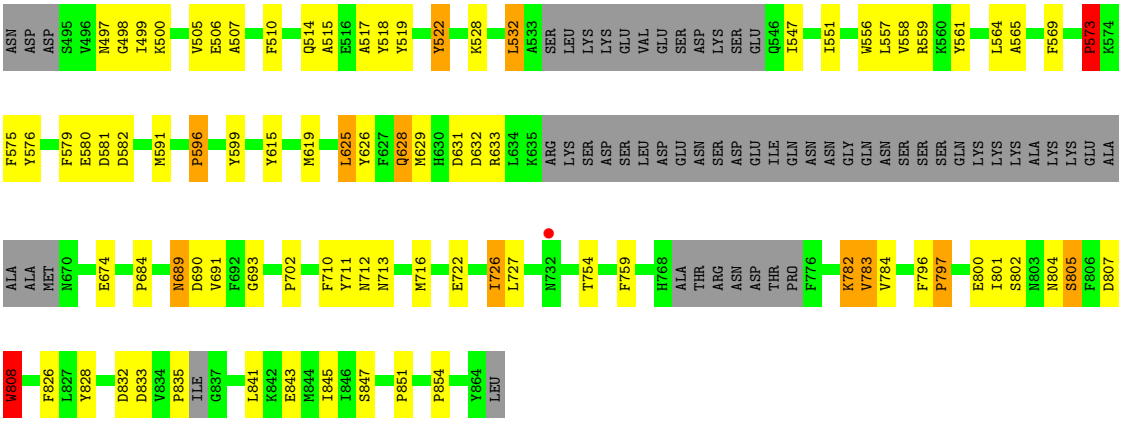
Chain G:



• Molecule 1: N-terminal acetyltransferase A complex subunit NAT1

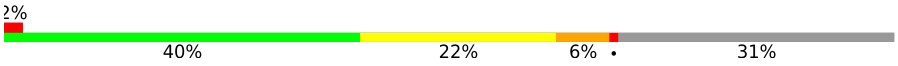
Chain M:

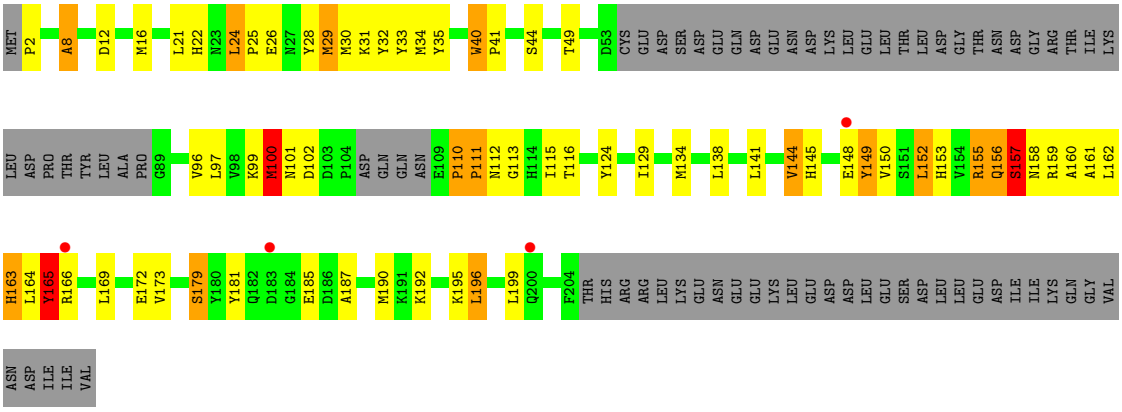




• Molecule 2: N-terminal acetyltransferase A complex catalytic subunit ARD1

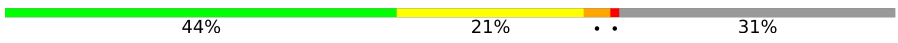
Chain B:

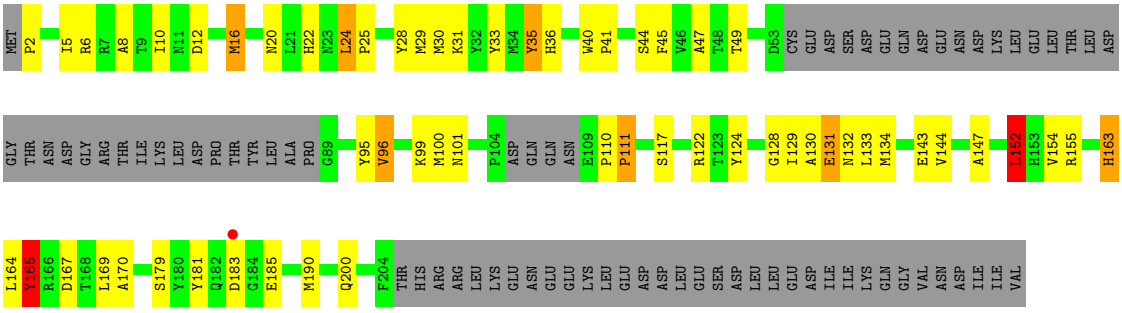




• Molecule 2: N-terminal acetyltransferase A complex catalytic subunit ARD1

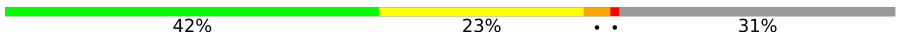
Chain H:

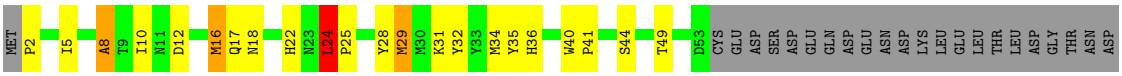


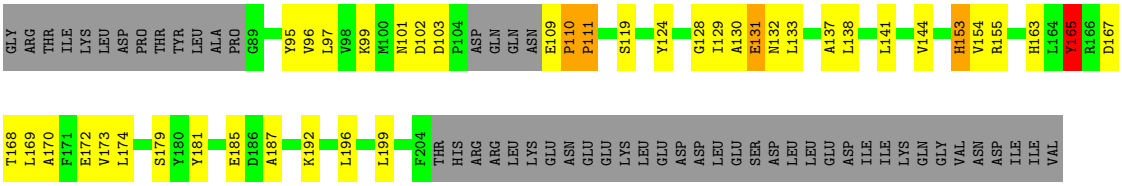


• Molecule 2: N-terminal acetyltransferase A complex catalytic subunit ARD1

Chain N:

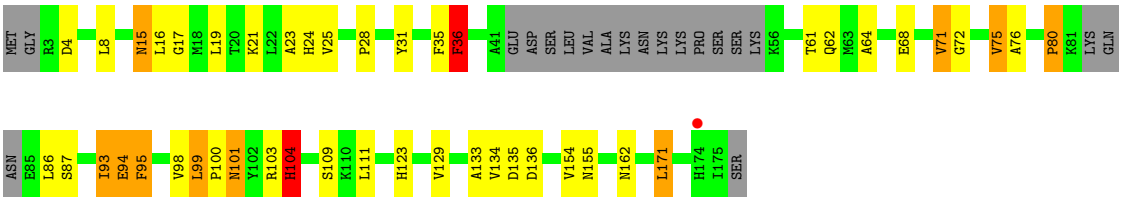






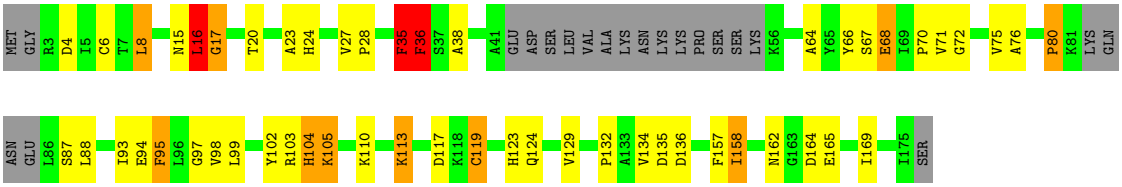
• Molecule 3: N-terminal acetyltransferase A complex subunit NAT5

Chain C:



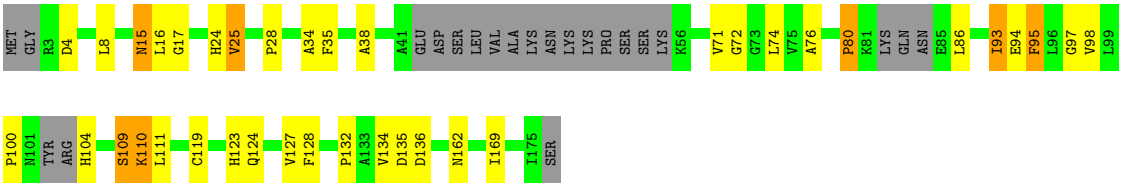
• Molecule 3: N-terminal acetyltransferase A complex subunit NAT5

Chain I:



• Molecule 3: N-terminal acetyltransferase A complex subunit NAT5

Chain O:



• Molecule 4: ALA-ALA-ALA-ALA-ALA-ALA

Chain E:

• Molecule 4: ALA-ALA-ALA-ALA-ALA-ALA

Chain K:

• Molecule 4: ALA-ALA-ALA-ALA-ALA-ALA

Chain Q:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	111.72Å 146.49Å 254.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.63 – 3.95 49.63 – 3.95	Depositor EDS
% Data completeness (in resolution range)	96.5 (49.63-3.95) 96.5 (49.63-3.95)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.03 (at 4.00Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.355 , 0.421 0.366 , 0.424	Depositor DCC
R_{free} test set	1799 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	35.4	Xtriage
Anisotropy	0.855	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.14 , 112.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.23$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	20122	wwPDB-VP
Average B, all atoms (Å^2)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 92.28 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.9860e-09. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ACO, G4P, CMC

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	1.26	19/4701 (0.4%)	0.90	20/6458 (0.3%)
1	G	1.04	13/4725 (0.3%)	0.95	21/6500 (0.3%)
1	M	0.95	10/4689 (0.2%)	0.90	22/6453 (0.3%)
2	B	1.13	7/1040 (0.7%)	1.06	7/1424 (0.5%)
2	H	0.93	2/1041 (0.2%)	0.99	5/1424 (0.4%)
2	N	0.79	1/1040 (0.1%)	0.92	5/1423 (0.4%)
3	C	1.30	4/900 (0.4%)	1.06	6/1243 (0.5%)
3	I	1.23	5/889 (0.6%)	1.09	9/1224 (0.7%)
3	O	0.70	0/870	0.86	5/1197 (0.4%)
4	E	1.17	0/29	1.07	0/39
4	K	1.18	0/29	0.99	0/39
4	Q	1.24	0/29	0.66	0/39
All	All	1.08	61/19982 (0.3%)	0.94	100/27463 (0.4%)

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	1
1	G	0	1
1	M	0	2
3	C	0	1
3	I	0	1
All	All	0	6

The worst 5 of 61 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	768	HIS	C-O	33.82	1.87	1.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	215	CYS	CB-SG	32.14	2.36	1.82
1	A	709	GLU	CD-OE1	28.81	1.57	1.25
1	A	768	HIS	C-O	26.80	1.74	1.23
1	A	709	GLU	CD-OE2	22.47	1.50	1.25

The worst 5 of 100 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	35	PHE	O-C-N	-16.32	96.58	122.70
3	I	36	PHE	N-CA-CB	13.10	134.18	110.60
3	C	35	PHE	CA-C-N	11.87	143.30	117.20
3	C	36	PHE	O-C-N	-9.46	107.57	122.70
2	H	134	MET	CG-SD-CE	-9.31	85.31	100.20

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	781	LYS	Peptide
3	C	36	PHE	Mainchain
1	G	806	PHE	Peptide
3	I	16	LEU	Mainchain
1	M	171	ALA	Peptide

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	4636	0	2969	127	0
1	G	4656	0	2966	129	0
1	M	4624	0	2960	137	0
2	B	1024	0	710	41	0
2	H	1025	0	695	25	0
2	N	1025	0	699	24	0
3	C	888	0	539	26	0
3	I	879	0	527	32	0
3	O	861	0	511	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	30	0	14	0	0
4	K	30	0	14	0	0
4	Q	30	0	14	0	0
5	A	36	0	11	1	0
5	G	36	0	11	3	0
5	M	36	0	11	0	0
6	B	51	0	31	0	0
6	H	51	0	31	0	0
6	N	51	0	31	0	0
7	C	51	0	34	10	0
7	I	51	0	34	9	0
7	O	51	0	34	5	0
All	All	20122	0	12846	548	0

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 17.

The worst 5 of 548 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:G:614:MET:CB	1:G:614:MET:CG	1.77	1.62
1:M:185:GLU:CD	1:M:185:GLU:CG	1.84	1.44
1:M:215:CYS:CB	1:M:215:CYS:SG	2.08	1.42
1:A:467:CYS:SG	1:A:467:CYS:CB	2.09	1.41
2:B:29:MET:CG	2:B:29:MET:SD	2.15	1.33

There are no symmetry-related clashes.

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	751/854 (88%)	487 (65%)	185 (25%)	79 (10%)	0 8
1	G	754/854 (88%)	494 (66%)	179 (24%)	81 (11%)	0 8
1	M	753/854 (88%)	504 (67%)	173 (23%)	76 (10%)	0 9
2	B	158/238 (66%)	94 (60%)	39 (25%)	25 (16%)	0 3
2	H	158/238 (66%)	91 (58%)	49 (31%)	18 (11%)	0 6
2	N	158/238 (66%)	93 (59%)	42 (27%)	23 (15%)	0 3
3	C	150/176 (85%)	94 (63%)	35 (23%)	21 (14%)	0 4
3	I	149/176 (85%)	89 (60%)	37 (25%)	23 (15%)	0 3
3	O	147/176 (84%)	90 (61%)	41 (28%)	16 (11%)	0 8
4	E	4/8 (50%)	4 (100%)	0	0	100 100
4	K	4/8 (50%)	3 (75%)	1 (25%)	0	100 100
4	Q	4/8 (50%)	4 (100%)	0	0	100 100
All	All	3190/3828 (83%)	2047 (64%)	781 (24%)	362 (11%)	0 7

5 of 362 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	VAL
1	A	101	SER
1	A	116	LYS
1	A	135	ASN
1	A	163	TRP

5.3.2 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 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	186/759 (24%)	142 (76%)	44 (24%)	1 5
1	G	189/759 (25%)	137 (72%)	52 (28%)	0 3
1	M	188/759 (25%)	155 (82%)	33 (18%)	2 12
2	B	51/216 (24%)	43 (84%)	8 (16%)	2 16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	H	51/216 (24%)	38 (74%)	13 (26%)	0 4
2	N	51/216 (24%)	40 (78%)	11 (22%)	1 6
3	C	29/153 (19%)	22 (76%)	7 (24%)	0 4
3	I	28/153 (18%)	25 (89%)	3 (11%)	6 27
3	O	25/153 (16%)	24 (96%)	1 (4%)	31 57
All	All	798/3384 (24%)	626 (78%)	172 (22%)	1 6

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	29	MET
1	M	430	GLU
2	H	117	SER
1	M	76	TYR
1	M	615	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	M	347	GLN
1	M	236	ASN
1	M	135	ASN
1	G	503	HIS
1	M	220	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

9 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	$\#\|Z\| > 2$
5	G4P	G	901	-	30,38,38	1.78	7 (23%)	42,61,61	2.29	14 (33%)
5	G4P	A	901	-	30,38,38	1.18	2 (6%)	42,61,61	1.40	7 (16%)
7	ACO	C	201	-	45,53,53	1.12	3 (6%)	56,79,79	1.74	10 (17%)
6	CMC	H	301	-	45,53,54	0.99	2 (4%)	55,78,80	1.50	7 (12%)
6	CMC	B	301	-	45,53,54	1.00	2 (4%)	55,78,80	1.50	8 (14%)
7	ACO	I	201	-	45,53,53	1.40	9 (20%)	56,79,79	2.04	19 (33%)
7	ACO	O	201	-	45,53,53	1.09	3 (6%)	56,79,79	1.43	6 (10%)
5	G4P	M	901	-	30,38,38	1.10	3 (10%)	42,61,61	1.50	7 (16%)
6	CMC	N	301	-	45,53,54	1.01	2 (4%)	55,78,80	1.53	9 (16%)

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
5	G4P	G	901	-	-	2/23/43/43	0/3/3/3
5	G4P	A	901	-	-	5/23/43/43	0/3/3/3
7	ACO	C	201	-	-	18/47/67/67	0/3/3/3
6	CMC	H	301	-	-	4/46/67/68	0/3/3/3
6	CMC	B	301	-	-	3/46/67/68	0/3/3/3
7	ACO	I	201	-	-	22/47/67/67	0/3/3/3
7	ACO	O	201	-	-	15/47/67/67	0/3/3/3
5	G4P	M	901	-	-	5/23/43/43	0/3/3/3
6	CMC	N	301	-	-	3/46/67/68	0/3/3/3

The worst 5 of 33 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	901	G4P	PD-O1D	4.96	1.66	1.50
5	G	901	G4P	O3'-C3'	3.82	1.58	1.44
7	C	201	ACO	O4B-C1B	3.71	1.46	1.41
7	I	201	ACO	C5P-N4P	3.65	1.41	1.33
7	I	201	ACO	C5A-C4A	3.43	1.50	1.40

The worst 5 of 87 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	G	901	G4P	PC-O3C-PD	-7.13	108.38	132.83
7	I	201	ACO	CAP-C9P-N8P	5.43	127.39	116.58
7	C	201	ACO	P2A-O3A-P1A	-5.26	114.77	132.83
7	O	201	ACO	P2A-O3A-P1A	-4.78	116.42	132.83
6	B	301	CMC	C2P-S1P-C1	4.77	109.68	101.71

There are no chirality outliers.

5 of 77 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	901	G4P	O4'-C4'-C5'-O5'
5	A	901	G4P	C3'-C4'-C5'-O5'
5	G	901	G4P	C4'-C3'-O3'-PC
5	M	901	G4P	C2'-C3'-O3'-PC
6	B	301	CMC	C5B-O5B-P1A-O1A

There are no ring outliers.

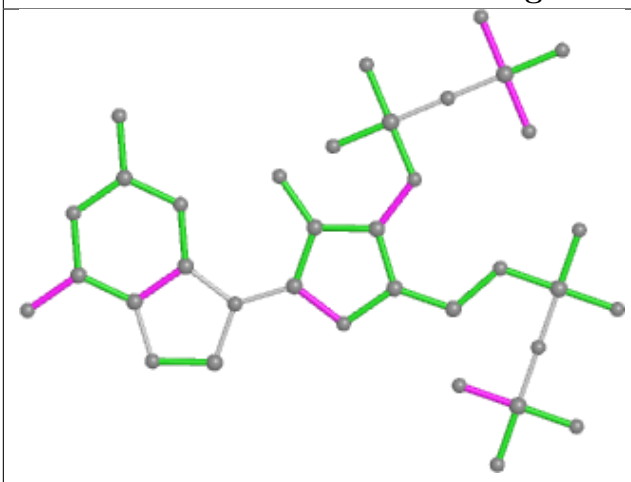
5 monomers are involved in 28 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	901	G4P	3	0
5	A	901	G4P	1	0
7	C	201	ACO	10	0
7	I	201	ACO	9	0
7	O	201	ACO	5	0

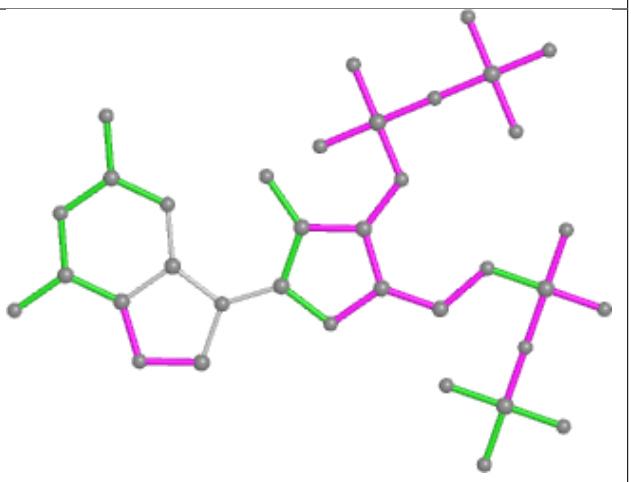
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

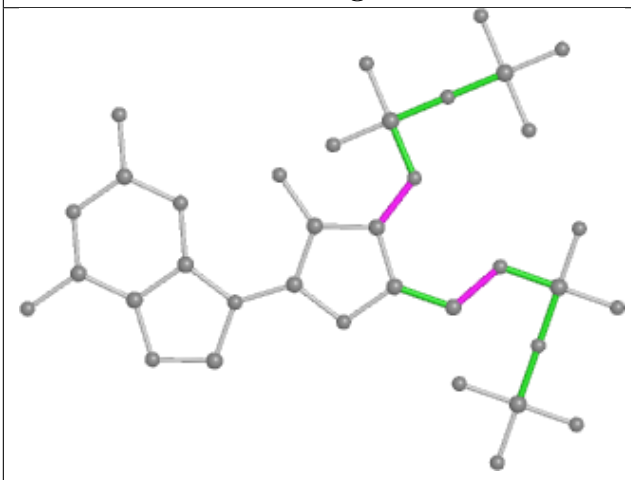
Ligand G4P G 901



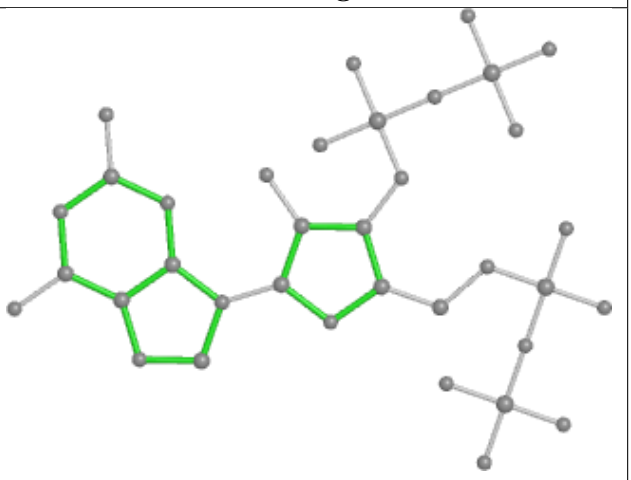
Bond lengths



Bond angles

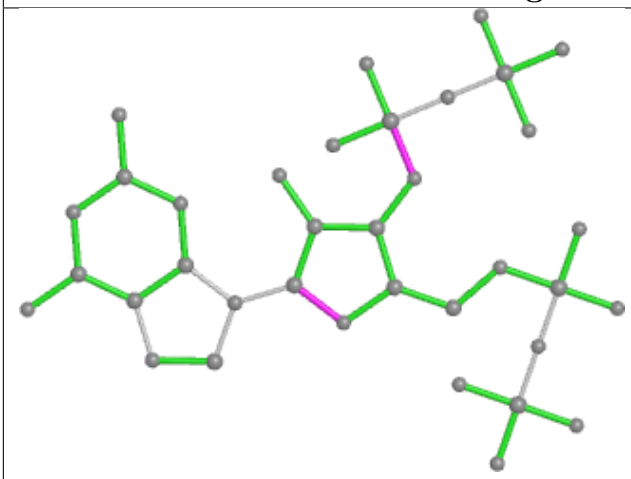


Torsions

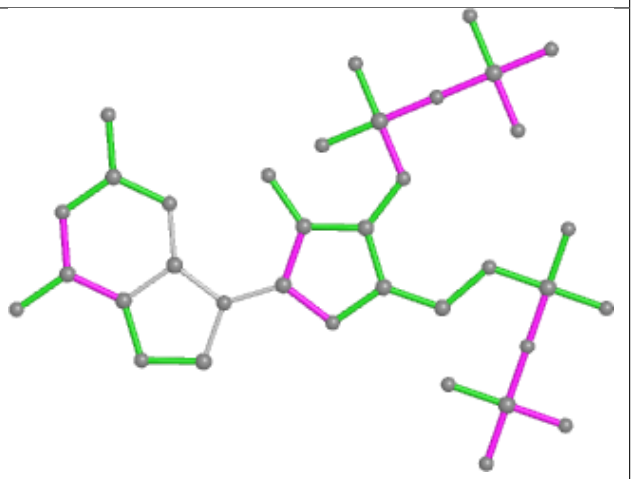


Rings

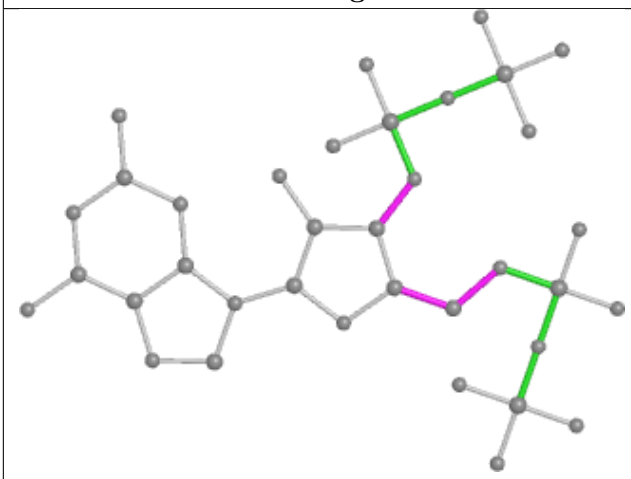
Ligand G4P A 901



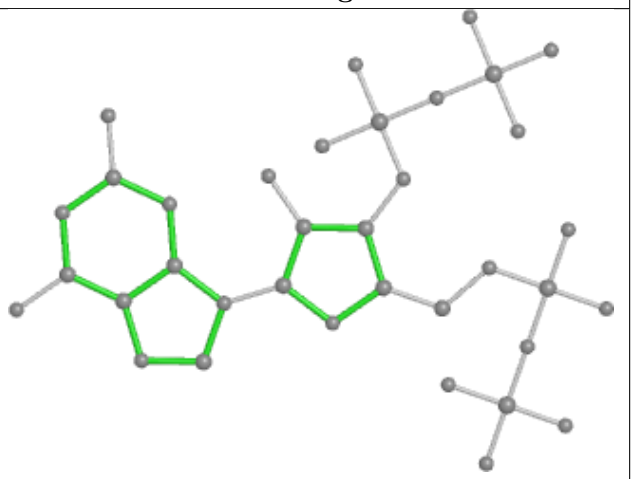
Bond lengths



Bond angles

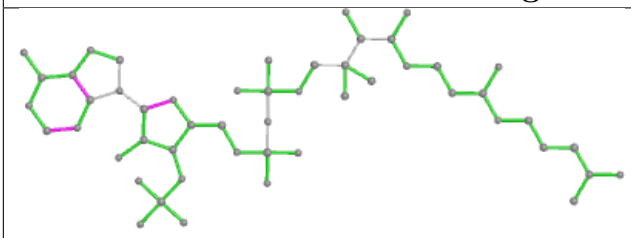


Torsions

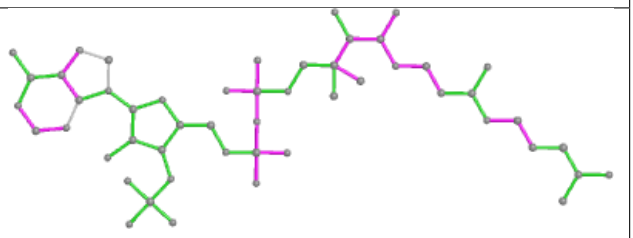


Rings

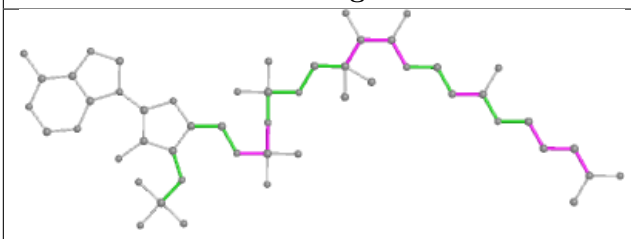
Ligand ACO C 201



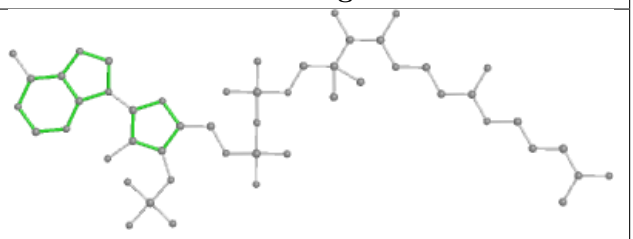
Bond lengths



Bond angles

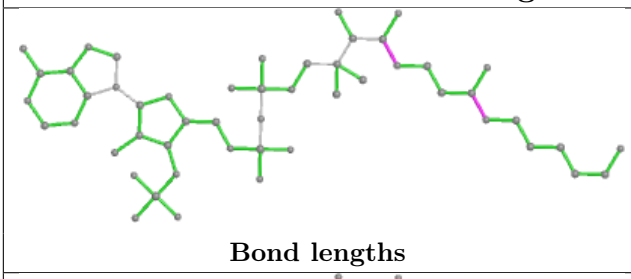


Torsions

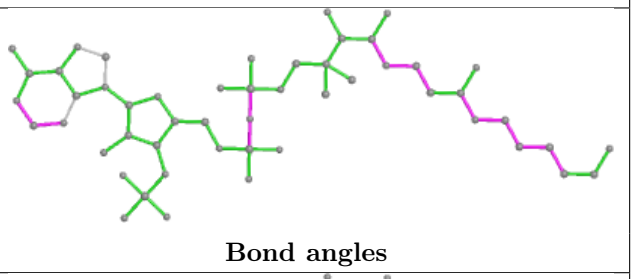


Rings

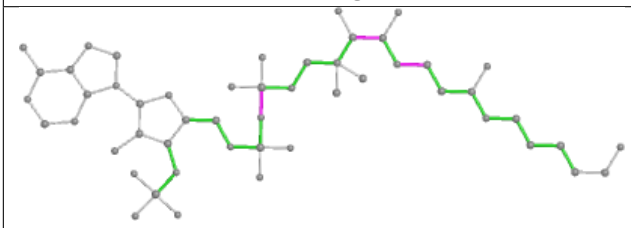
Ligand CMC H 301



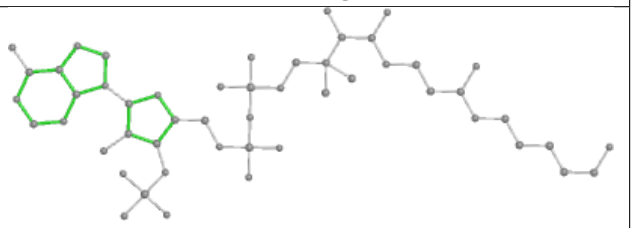
Bond lengths



Bond angles

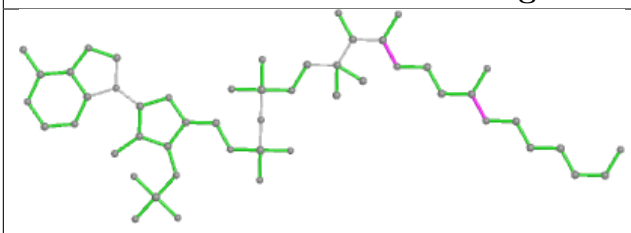


Torsions

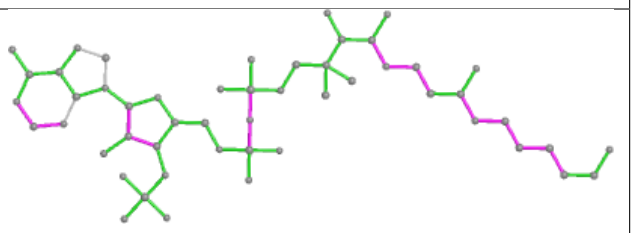


Rings

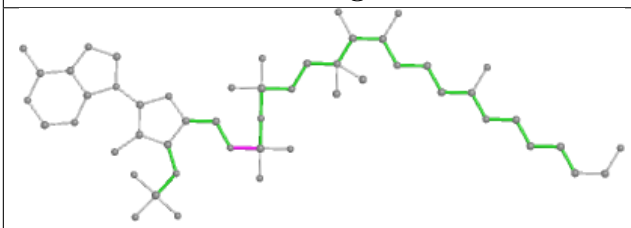
Ligand CMC B 301



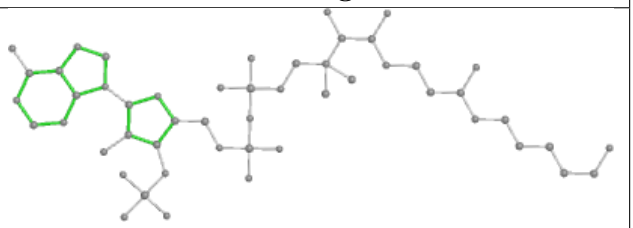
Bond lengths



Bond angles

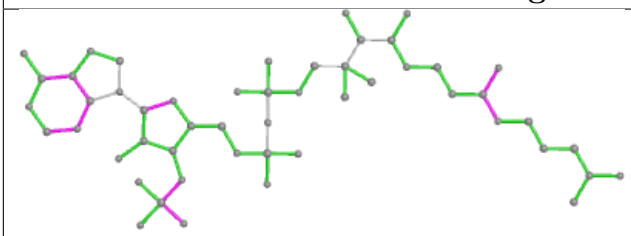


Torsions

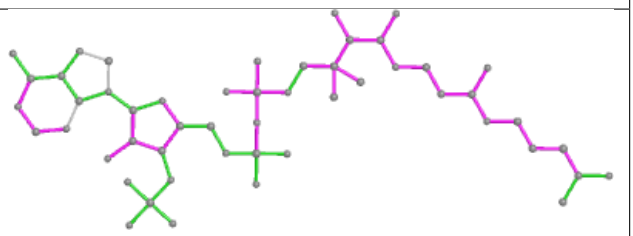


Rings

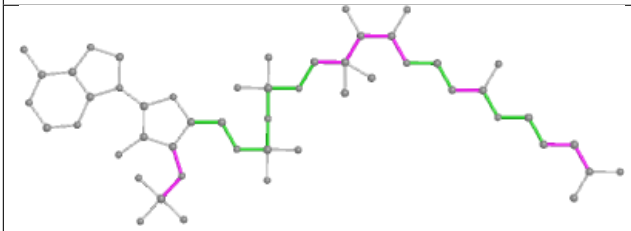
Ligand ACO I 201



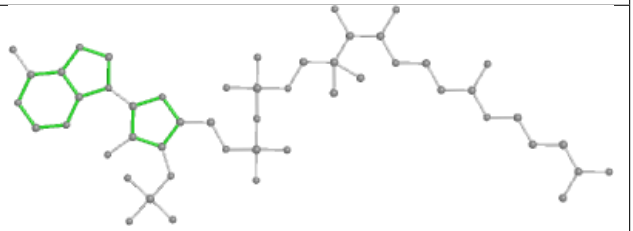
Bond lengths



Bond angles

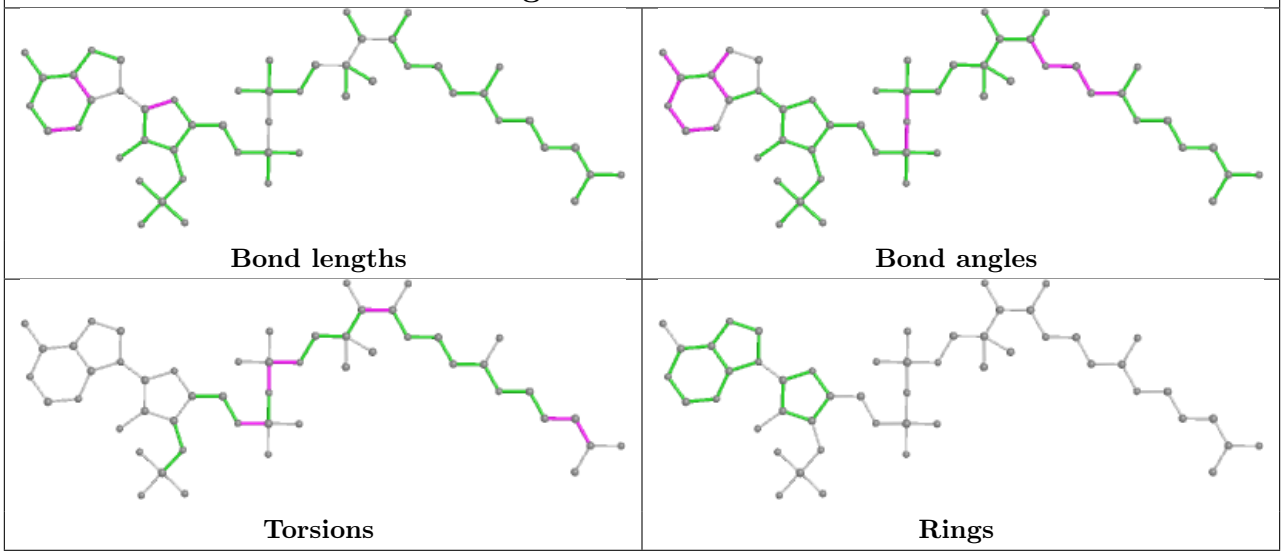


Torsions

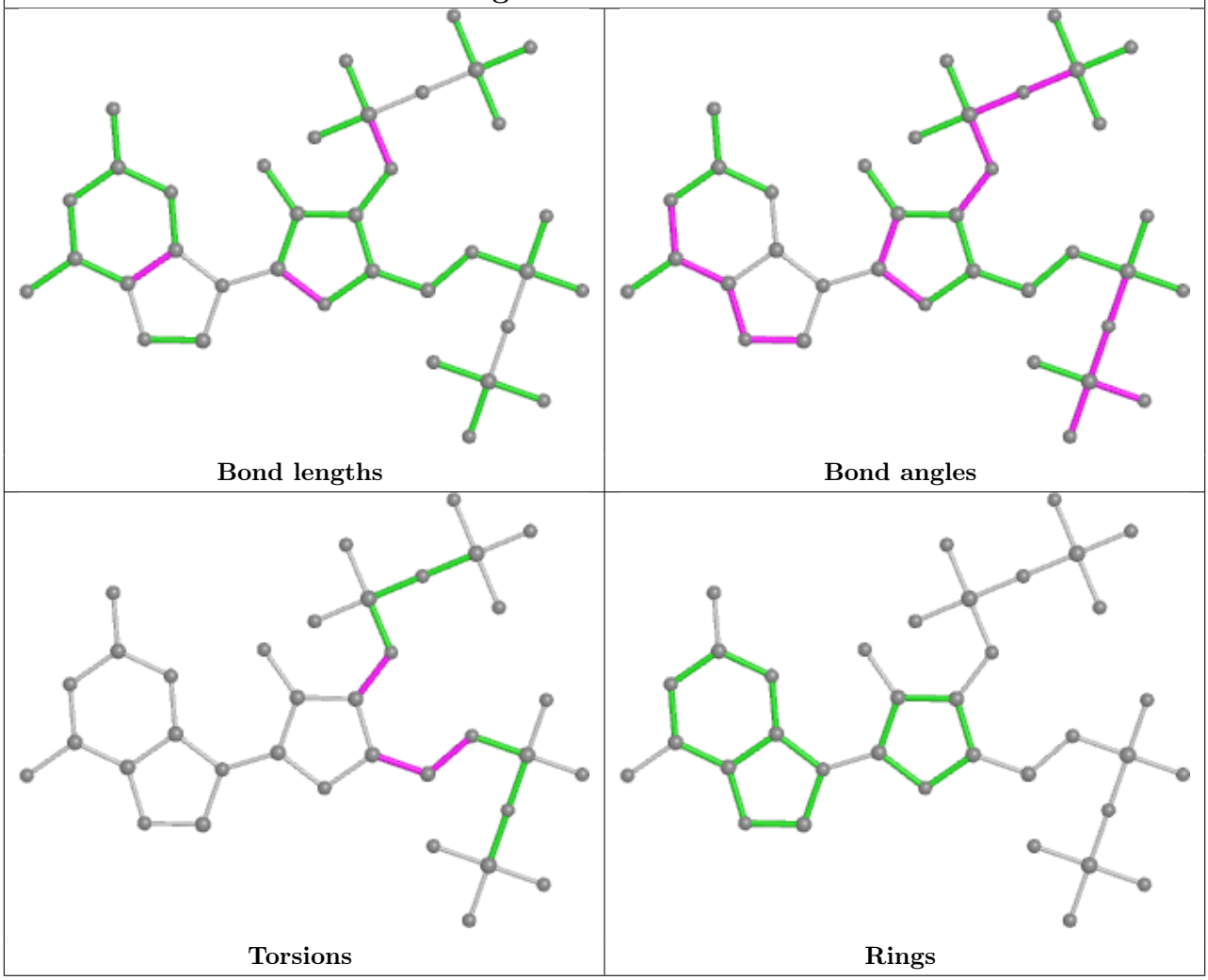


Rings

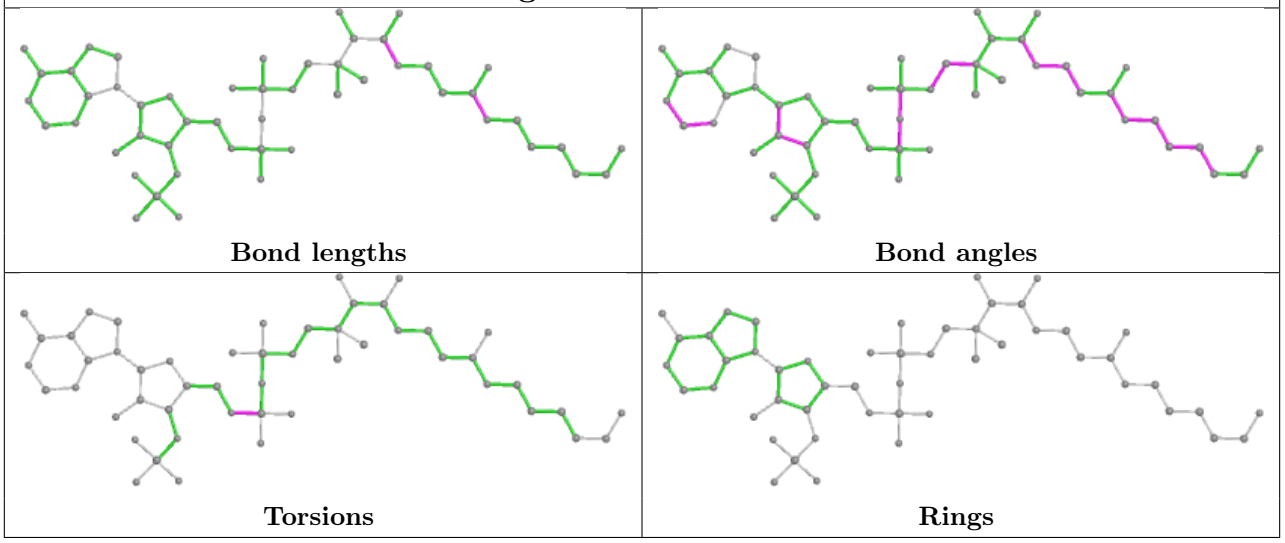
Ligand ACO O 201



Ligand G4P M 901



Ligand CMC N 301



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	35:PHE	C	36:PHE	N	0.90

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	769/854 (90%)	-0.12	7 (0%) 84 77	83, 86, 91, 94	0
1	G	770/854 (90%)	-0.08	5 (0%) 89 84	73, 85, 88, 88	0
1	M	769/854 (90%)	-0.14	4 (0%) 91 85	84, 86, 90, 92	0
2	B	164/238 (68%)	-0.13	4 (2%) 59 49	48, 84, 86, 88	0
2	H	164/238 (68%)	-0.18	1 (0%) 89 84	84, 85, 86, 87	0
2	N	164/238 (68%)	-0.18	0 100 100	54, 86, 87, 88	0
3	C	156/176 (88%)	-0.01	1 (0%) 89 84	84, 86, 88, 88	2 (1%)
3	I	155/176 (88%)	0.01	0 100 100	85, 87, 89, 90	2 (1%)
3	O	154/176 (87%)	-0.06	0 100 100	86, 90, 93, 93	1 (0%)
4	E	6/8 (75%)	0.48	0 100 100	83, 83, 84, 84	0
4	K	6/8 (75%)	0.33	1 (16%) 1 2	84, 84, 85, 85	0
4	Q	6/8 (75%)	0.32	0 100 100	85, 85, 85, 85	0
All	All	3283/3828 (85%)	-0.11	23 (0%) 87 82	48, 86, 90, 94	5 (0%)

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	545	GLU	4.0
1	G	775	PRO	3.5
1	A	64	HIS	3.1
1	G	384	GLY	2.9
1	M	732	ASN	2.9

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

LIGAND-RSR INFOmissingINFO

6.5 Other polymers (i)

There are no such residues in this entry.