wwPDB X-ray Structure Validation Summary Report (i)

Oct 23, 2021 – 01:26 PM EDT

PDB ID	:	1RTD
Title	:	STRUCTURE OF A CATALYTIC COMPLEX OF HIV-1 REVERSE TRANSCRIPTASE: IMPLICATIONS FOR NUCLEOSIDE ANALOG DRUG RESISTANCE
Authors	:	Chopra, R.; Huang, H.; Verdine, G.L.; Harrison, S.C.
Deposited on	:	1998-08-26
Resolution	:	3.20 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:

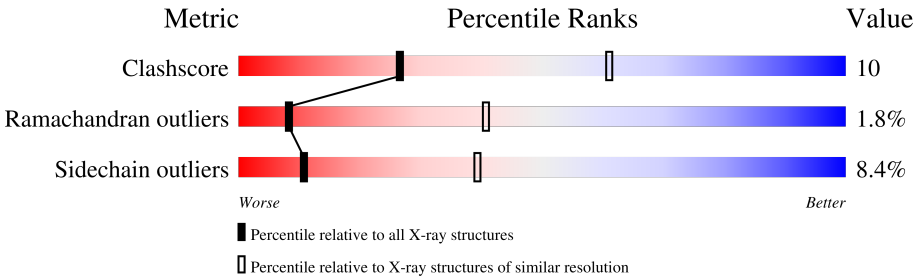
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.20 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	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	E	27	7% 41% 44% 7%
1	G	27	· 30% 59% 7%
2	F	21	5% 33% 62%
2	H	21	48% 52%
3	A	554	78% 19% ·
3	C	554	77% 20% ·
4	B	440	78% 14% · 6%
4	D	440	78% 14% · 6%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 17784 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 DNA chain called DNA TEMPLATE FOR REVERSE TRANSCRIPTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	25	Total	C	N	O	P	0	0	0
			512	242	100	146	24			
1	G	25	Total	C	N	O	P	0	0	0
			512	242	100	146	24			

- Molecule 2 is a DNA chain called DNA PRIMER FOR REVERSE TRANSCRIPTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	21	Total	C	N	O	P	0	0	0
			424	202	77	125	20			
2	H	21	Total	C	N	O	P	0	0	0
			424	202	77	125	20			

- Molecule 3 is a protein called PROTEIN (REVERSE TRANSCRIPTASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	554	Total	C	N	O	S	0	0	0
			4510	2917	751	833	9			
3	C	554	Total	C	N	O	S	0	0	0
			4508	2915	751	833	9			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	LYS	PRO	engineered mutation	UNP P03366
A	172	ARG	LYS	conflict	UNP P03366
A	258	CYS	GLN	engineered mutation	UNP P03366
A	471	ASP	ASN	conflict	UNP P03366
A	478	GLN	GLU	engineered mutation	UNP P03366
A	512	GLU	LYS	conflict	UNP P03366
C	1	LYS	PRO	engineered mutation	UNP P03366
C	172	ARG	LYS	conflict	UNP P03366

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	258	CYS	GLN	engineered mutation	UNP P03366
C	471	ASP	ASN	conflict	UNP P03366
C	478	GLN	GLU	engineered mutation	UNP P03366
C	512	GLU	LYS	conflict	UNP P03366

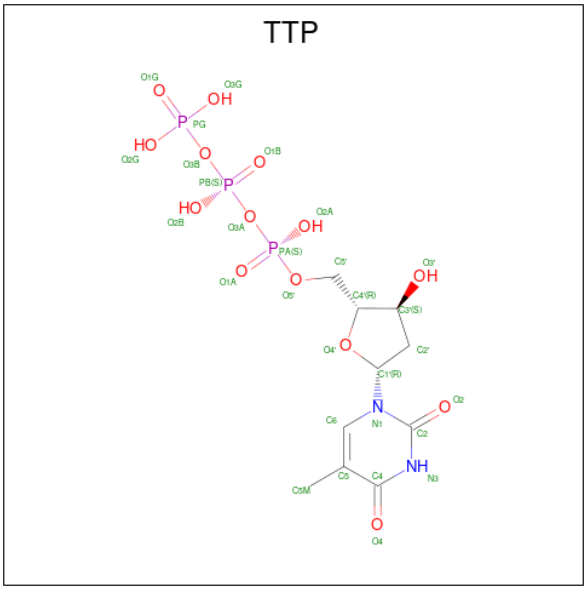
- Molecule 4 is a protein called PROTEIN (REVERSE TRANSCRIPTASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	414	Total	C	N	O	S	0	0	0
			3415	2221	567	620	7			
4	D	414	Total	C	N	O	S	0	0	0
			3415	2221	567	620	7			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	4	Total	Mg	0	0
			4	4		
5	C	2	Total	Mg	0	0
			2	2		

- Molecule 6 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula: $C_{10}H_{17}N_2O_{14}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			29	10	2	14	3		
6	C	1	Total	C	N	O	P	0	0
			29	10	2	14	3		

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. 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: DNA TEMPLATE FOR REVERSE TRANSCRIPTASE

Chain E:

• Molecule 1: DNA TEMPLATE FOR REVERSE TRANSCRIPTASE

Chain G:

• Molecule 2: DNA PRIMER FOR REVERSE TRANSCRIPTASE

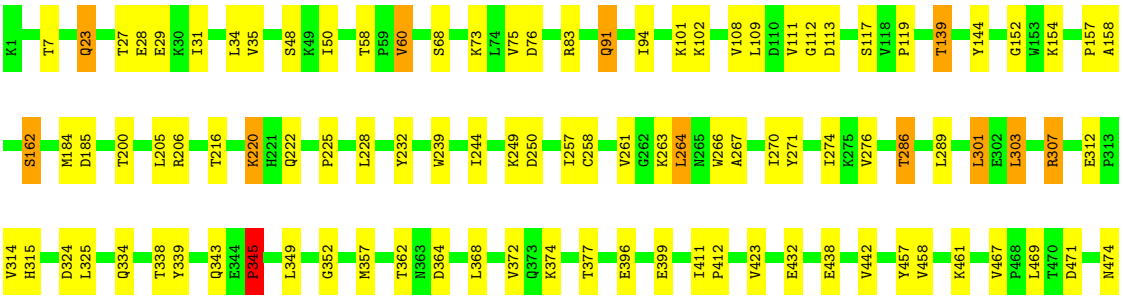
Chain F:

• Molecule 2: DNA PRIMER FOR REVERSE TRANSCRIPTASE

Chain H:

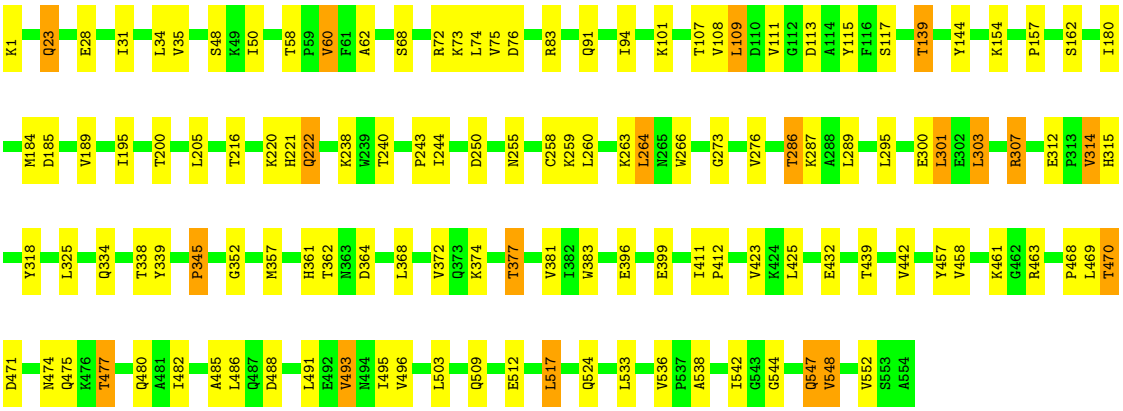
• Molecule 3: PROTEIN (REVERSE TRANSCRIPTASE)

Chain A:



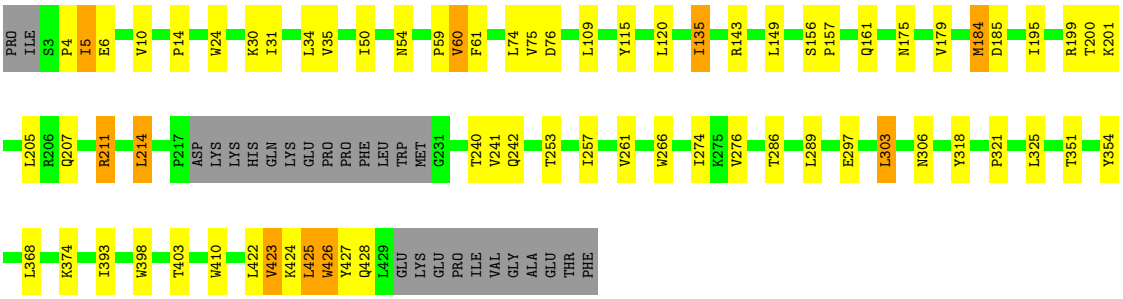
• Molecule 3: PROTEIN (REVERSE TRANSCRIPTASE)

Chain C:



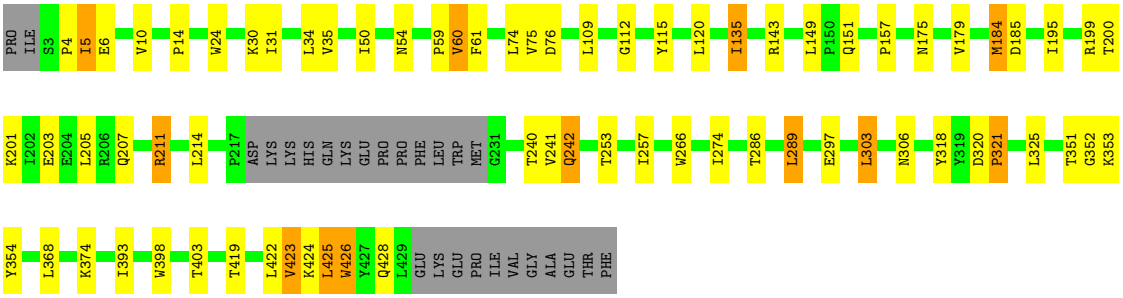
• Molecule 4: PROTEIN (REVERSE TRANSCRIPTASE)

Chain B:



• Molecule 4: PROTEIN (REVERSE TRANSCRIPTASE)

Chain D:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	78.84Å 150.70Å 280.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 3.20	Depositor
% Data completeness (in resolution range)	92.2 (12.00-3.20)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.224 , 0.298	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	17784	wwPDB-VP
Average B, all atoms (Å^2)	68.0	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: MG, TTP

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	E	0.87	0/575	1.40	12/886 (1.4%)
1	G	0.76	0/575	1.53	20/886 (2.3%)
2	F	0.94	0/474	1.59	17/729 (2.3%)
2	H	0.69	0/474	1.48	13/729 (1.8%)
3	A	0.73	1/4627 (0.0%)	0.92	8/6286 (0.1%)
3	C	0.75	2/4624 (0.0%)	0.91	6/6281 (0.1%)
4	B	0.72	0/3512	0.89	3/4774 (0.1%)
4	D	0.72	0/3512	0.90	2/4774 (0.0%)
All	All	0.74	3/18373 (0.0%)	1.00	81/25345 (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 outliers	#Planarity outliers
3	C	0	1
4	B	0	1
4	D	0	1
All	All	0	3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	512	GLU	CD-OE2	6.29	1.32	1.25
3	C	461	LYS	CE-NZ	-6.09	1.33	1.49
3	A	271	TYR	CD1-CE1	5.09	1.47	1.39

The worst 5 of 81 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	20	DG	N9-C1'-C2'	10.02	131.63	112.60
1	G	16	DA	N9-C1'-C2'	9.39	130.45	112.60
1	G	18	DC	N1-C1'-C2'	8.63	129.00	112.60
2	F	8	DC	N1-C1'-C2'	8.07	127.94	112.60
2	H	15	DA	N9-C1'-C2'	8.04	127.88	112.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	B	354	TYR	Sidechain
3	C	318	TYR	Sidechain
4	D	354	TYR	Sidechain

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	E	512	0	280	62	0
1	G	512	0	280	58	0
2	F	424	0	235	50	0
2	H	424	0	235	48	0
3	A	4510	0	4566	53	0
3	C	4508	0	4560	58	0
4	B	3415	0	3448	30	0
4	D	3415	0	3448	33	0
5	A	4	0	0	0	0
5	C	2	0	0	0	0
6	A	29	0	13	1	0
6	C	29	0	13	2	0
All	All	17784	0	17078	359	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 10.

The worst 5 of 359 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)
2:F:3:DA:H2'	2:F:4:DG:C8	1.91	1.05
1:E:24:DC:H2"	1:E:25:DT:O5'	1.55	1.05
1:G:25:DT:H2"	1:G:26:DG:O5'	1.59	1.03
1:G:24:DC:H2"	1:G:25:DT:O5'	1.56	1.02
1:E:5:DA:H2"	1:E:6:DC:H5'	1.46	0.97

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	
3	A	552/554 (100%)	514 (93%)	30 (5%)	8 (1%)	11	46
3	C	552/554 (100%)	509 (92%)	33 (6%)	10 (2%)	8	41
4	B	410/440 (93%)	384 (94%)	18 (4%)	8 (2%)	7	38
4	D	410/440 (93%)	383 (93%)	19 (5%)	8 (2%)	7	38
All	All	1924/1988 (97%)	1790 (93%)	100 (5%)	34 (2%)	8	41

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	345	PRO
3	A	412	PRO
3	A	538	ALA
4	B	5	ILE
4	B	240	THR

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
3	A	495/495 (100%)	446 (90%)	49 (10%)	8 30
3	C	494/495 (100%)	443 (90%)	51 (10%)	7 29
4	B	376/400 (94%)	352 (94%)	24 (6%)	17 52
4	D	376/400 (94%)	353 (94%)	23 (6%)	18 54
All	All	1741/1790 (97%)	1594 (92%)	147 (8%)	11 39

5 of 147 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	474	ASN
4	D	368	LEU
3	C	493	VAL
4	D	135	ILE
3	A	548	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	255	ASN
3	C	464	GLN
3	C	447	ASN
3	C	509	GLN
3	A	464	GLN

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)

Of 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
6	TTP	A	700	5	23,30,30	1.78	3 (13%)	29,47,47	2.93	5 (17%)
6	TTP	C	705	5	23,30,30	1.23	4 (17%)	29,47,47	2.85	5 (17%)

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
6	TTP	A	700	5	-	5/19/34/34	0/2/2/2
6	TTP	C	705	5	-	1/19/34/34	0/2/2/2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	700	TTP	PG-O3G	6.66	1.80	1.54
6	C	705	TTP	C4-N3	3.37	1.38	1.33
6	A	700	TTP	C4-N3	3.27	1.38	1.33
6	A	700	TTP	C6-C5	-2.37	1.33	1.40
6	C	705	TTP	C4-C5	2.32	1.46	1.41

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	700	TTP	C4-N3-C2	14.20	127.13	115.14
6	C	705	TTP	C4-N3-C2	13.86	126.84	115.14
6	A	700	TTP	O3G-PG-O2G	3.50	121.02	107.64
6	A	700	TTP	O2G-PG-O3B	-2.90	94.90	104.64
6	C	705	TTP	PB-O3B-PG	2.86	142.63	132.83

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	700	TTP	C5'-O5'-PA-O1A
6	A	700	TTP	C5'-O5'-PA-O3A
6	C	705	TTP	PB-O3A-PA-O1A
6	A	700	TTP	C5'-O5'-PA-O2A
6	A	700	TTP	PA-O3A-PB-O2B

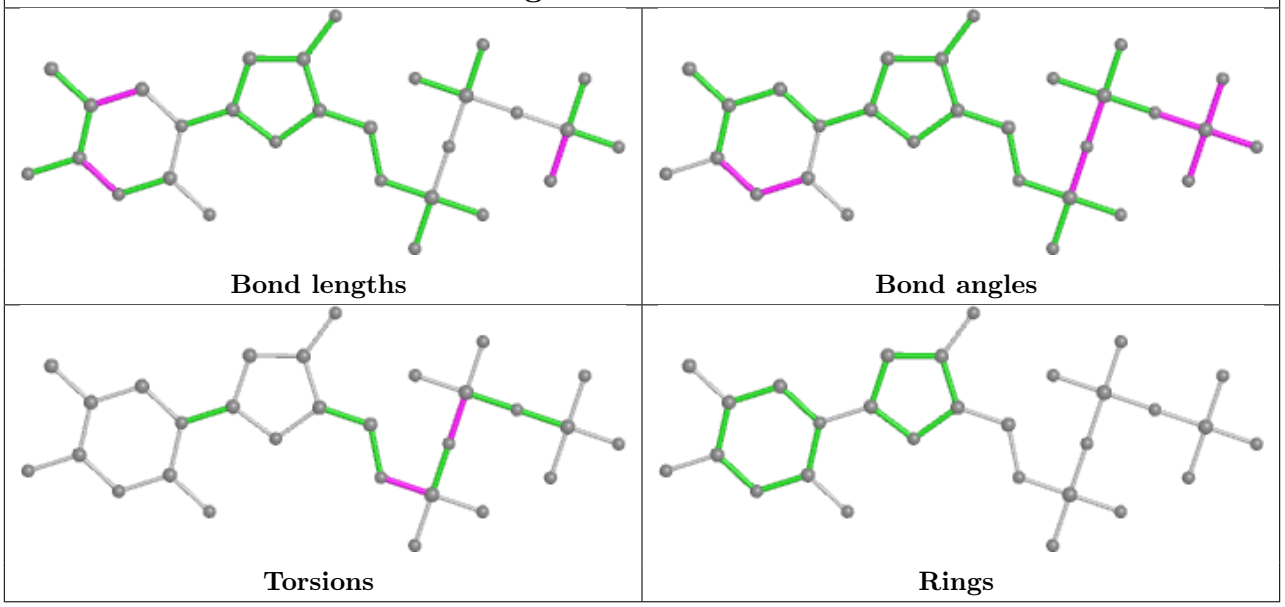
There are no ring outliers.

2 monomers are involved in 3 short contacts:

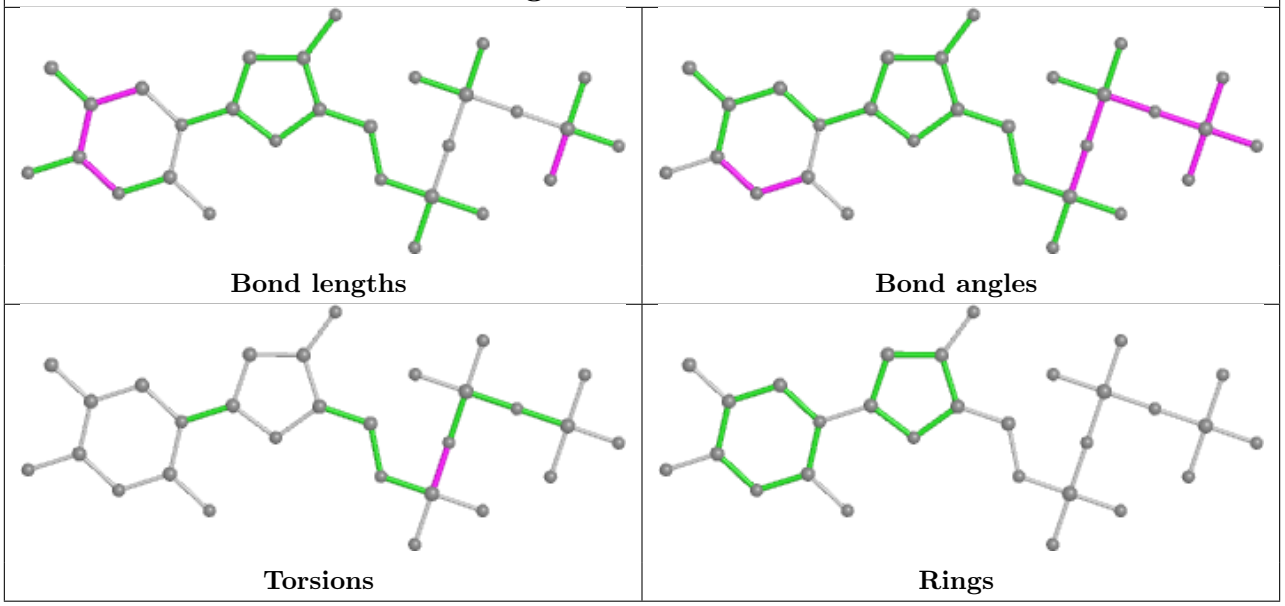
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	700	TTP	1	0
6	C	705	TTP	2	0

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 TTP A 700



Ligand TTP C 705



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

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.