



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 11:44 PM GMT

PDB ID : 1YFL  
Title : T4Dam in Complex with Sinefungin and 16-mer Oligonucleotide Showing Semi-specific and Specific Contact and Flipped Base  
Authors : Horton, J.R.; Liebert, K.; Hattman, S.; Jeltsch, A.; Cheng, X.  
Deposited on : 2005-01-03  
Resolution : 3.09 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

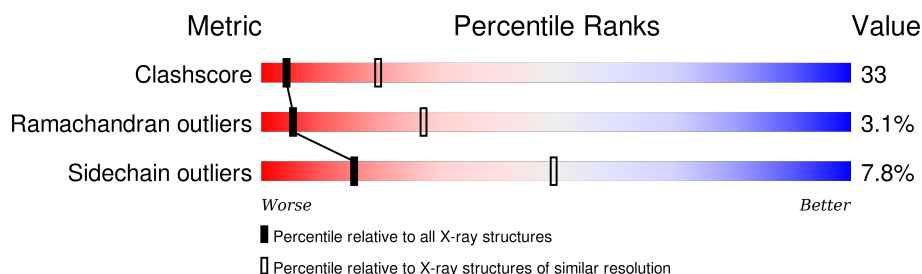
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.09 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	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)

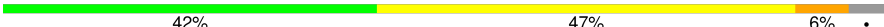
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	F	16	<div> <div>25%</div> <div>69%</div> <div>6%</div> </div>
1	G	16	<div> <div>25%</div> <div>69%</div> <div>6%</div> </div>
1	H	16	<div> <div>56%</div> <div>38%</div> <div>6%</div> </div>
1	I	16	<div> <div>38%</div> <div>56%</div> <div>6%</div> </div>
2	A	259	<div> <div>44%</div> <div>46%</div> <div>5%</div> <div>5%</div> </div>
2	B	259	<div> <div>44%</div> <div>46%</div> <div>5%</div> <div>5%</div> </div>
2	D	259	<div> <div>46%</div> <div>45%</div> <div>6%</div> <div>.</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	259	 A horizontal bar chart showing the quality of chain E. The bar is divided into three segments: green (42%), yellow (47%), and orange (6%). A small black dot is visible at the end of the bar.

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9379 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 5'-D(P\*TP\*CP\*AP\*CP\*AP\*GP\*GP\*AP\*TP\*CP\*CP\*TP\*GP\*TP\*GP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	16	Total	C	N	O	P	0	0	0
			329	156	60	97	16			
1	G	16	Total	C	N	O	P	0	0	0
			329	156	60	97	16			
1	H	16	Total	C	N	O	P	0	0	0
			329	156	60	97	16			
1	I	16	Total	C	N	O	P	0	0	0
			329	156	60	97	16			

- Molecule 2 is a protein called DNA adenine methylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	246	Total	C	N	O	S	0	0	0
			1969	1266	333	365	5			
2	B	246	Total	C	N	O	S	0	0	0
			1969	1266	333	365	5			
2	D	249	Total	C	N	O	S	0	0	0
			2010	1293	339	373	5			
2	E	248	Total	C	N	O	S	0	0	0
			2005	1290	338	372	5			

There are 12 discrepancies between the modelled and reference sequences:

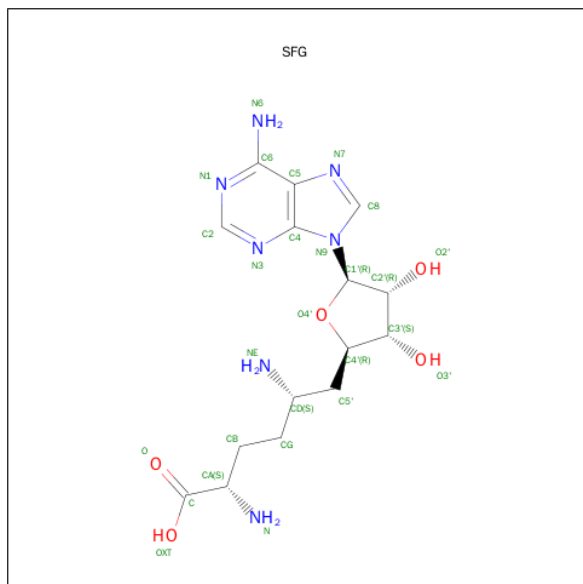
Chain	Residue	Modelled	Actual	Comment	Reference
A	139	ARG	GLN	SEE REMARK 999	UNP P04392
A	140	PHE	TYR	SEE REMARK 999	UNP P04392
A	209	LEU	GLN	SEE REMARK 999	UNP P04392
B	139	ARG	GLN	SEE REMARK 999	UNP P04392
B	140	PHE	TYR	SEE REMARK 999	UNP P04392
B	209	LEU	GLN	SEE REMARK 999	UNP P04392
D	139	ARG	GLN	SEE REMARK 999	UNP P04392

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	140	PHE	TYR	SEE REMARK 999	UNP P04392
D	209	LEU	GLN	SEE REMARK 999	UNP P04392
E	139	ARG	GLN	SEE REMARK 999	UNP P04392
E	140	PHE	TYR	SEE REMARK 999	UNP P04392
E	209	LEU	GLN	SEE REMARK 999	UNP P04392

- Molecule 3 is SINEFUNGIN (three-letter code: SFG) (formula:  $C_{15}H_{23}N_7O_5$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			27	15	7	5		
3	B	1	Total	C	N	O	0	0
			27	15	7	5		
3	D	1	Total	C	N	O	0	0
			27	15	7	5		
3	E	1	Total	C	N	O	0	0
			27	15	7	5		

- Molecule 4 is water.

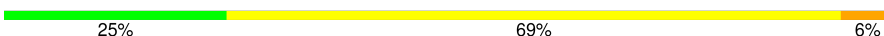
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	1	Total	O	0	0
			1	1		
4	E	1	Total	O	0	0
			1	1		

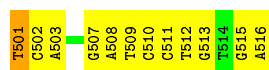
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.


Note EDS was not executed.

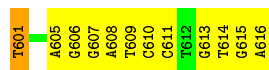
- Molecule 1: 5'-D(P\*TP\*CP\*AP\*CP\*AP\*GP\*GP\*AP\*TP\*CP\*CP\*TP\*GP\*TP\*GP\*A)-3',

Chain F: 



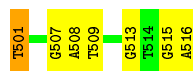
- Molecule 1: 5'-D(P\*TP\*CP\*AP\*CP\*AP\*GP\*GP\*AP\*TP\*CP\*CP\*TP\*GP\*TP\*GP\*A)-3',

Chain G: 



- Molecule 1: 5'-D(P\*TP\*CP\*AP\*CP\*AP\*GP\*GP\*AP\*TP\*CP\*CP\*TP\*GP\*TP\*GP\*A)-3',

Chain H: 



- Molecule 1: 5'-D(P\*TP\*CP\*AP\*CP\*AP\*GP\*GP\*AP\*TP\*CP\*CP\*TP\*GP\*TP\*GP\*A)-3',

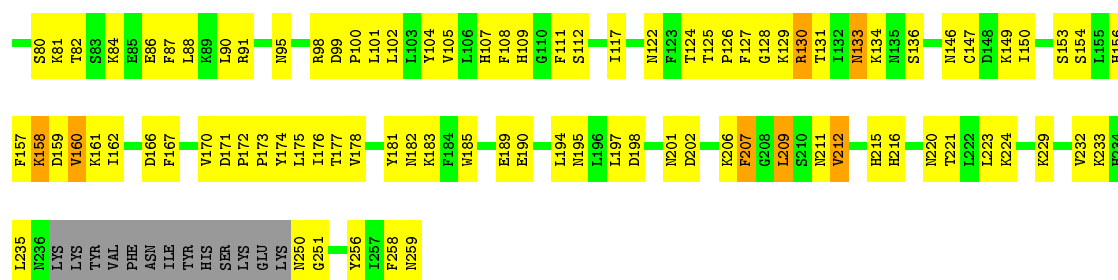
Chain I: 



- Molecule 2: DNA adenine methylase

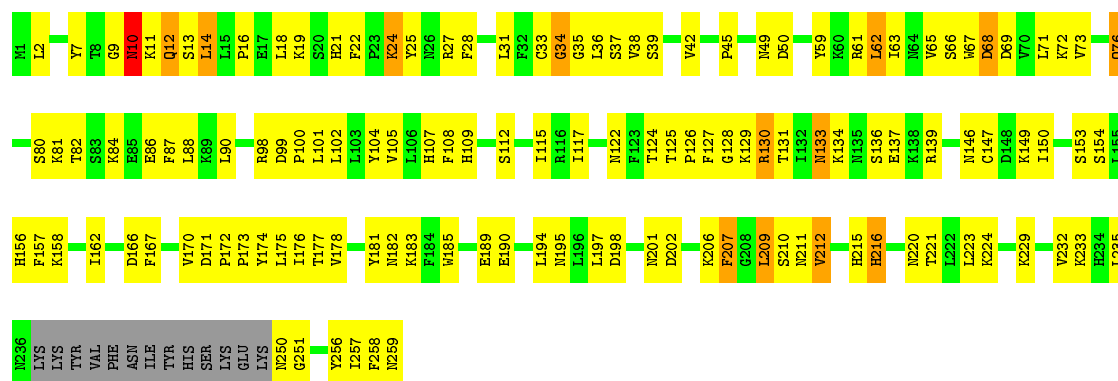
Chain A: 





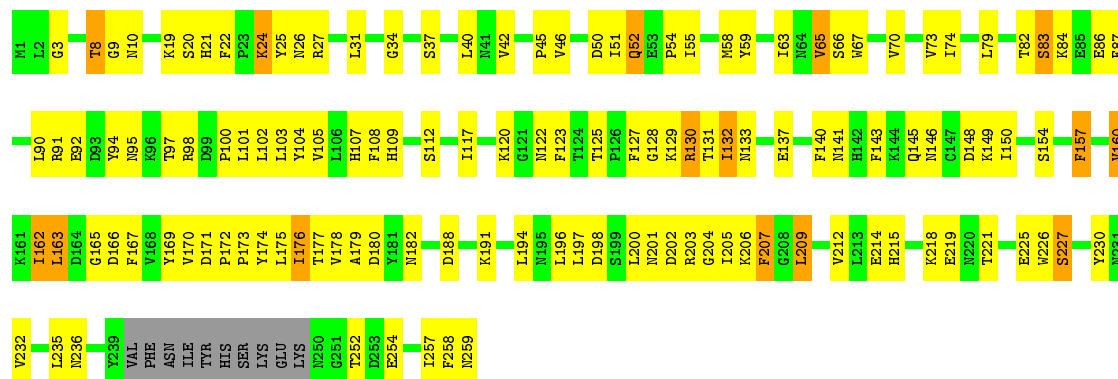
- Molecule 2: DNA adenine methylase

Chain B: 44% 46% 5% 5%



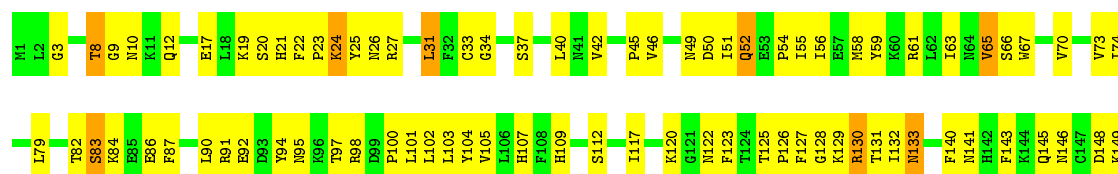
- Molecule 2: DNA adenine methylase

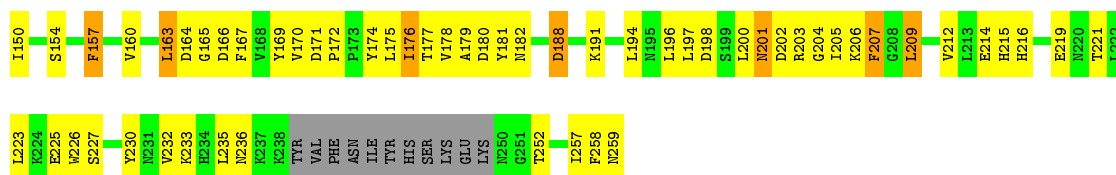
Chain D: 46% 45% 6%



- Molecule 2: DNA adenine methylase

Chain E: 42% 47% 6%







## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.87Å 117.65Å 87.86Å 90.00° 90.02° 90.00°	Depositor
Resolution (Å)	29.40 – 3.09	Depositor
% Data completeness (in resolution range)	97.9 (29.40-3.09)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.214 , 0.265	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9379	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	F	0.92	1/368 (0.3%)	0.90	0/564
1	G	0.85	1/368 (0.3%)	0.79	0/564
1	H	1.02	2/368 (0.5%)	1.06	1/564 (0.2%)
1	I	0.93	1/368 (0.3%)	0.90	0/564
2	A	0.51	0/2016	0.73	0/2727
2	B	0.52	0/2016	0.75	0/2727
2	D	0.54	0/2057	0.78	1/2780 (0.0%)
2	E	0.54	0/2052	0.78	0/2773
All	All	0.61	5/9613 (0.1%)	0.79	2/13263 (0.0%)

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	H	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	601	DT	OP3-P	-7.16	1.52	1.61
1	G	601	DT	OP3-P	-6.87	1.52	1.61
1	F	501	DT	OP3-P	-6.41	1.53	1.61
1	H	501	DT	OP3-P	-5.77	1.54	1.61
1	H	501	DT	P-O5'	5.57	1.65	1.59

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	H	501	DT	O5'-P-OP1	-8.75	97.83	105.70
2	D	160	VAL	CB-CA-C	-5.29	101.34	111.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	513	DG	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	F	329	0	181	23	0
1	G	329	0	181	23	0
1	H	329	0	181	13	0
1	I	329	0	181	33	0
2	A	1969	0	1877	131	0
2	B	1969	0	1877	130	0
2	D	2010	0	1938	135	0
2	E	2005	0	1936	141	0
3	A	27	0	22	3	0
3	B	27	0	22	0	0
3	D	27	0	22	4	0
3	E	27	0	22	5	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
All	All	9379	0	8440	590	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 33.

The worst 5 of 590 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:616:DA:H5"	2:A:8:THR:HG23	1.26	1.10

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:179:ALA:H	2:D:182:ASN:ND2	1.51	1.08
2:D:24:LYS:HE3	2:D:25:TYR:H	1.20	1.07
2:E:179:ALA:H	2:E:182:ASN:ND2	1.51	1.06
1:F:515:DG:H2"	1:F:516:DA:H5"	1.10	1.05

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	
2	A	242/259 (93%)	193 (80%)	41 (17%)	8 (3%)	5	26
2	B	242/259 (93%)	194 (80%)	42 (17%)	6 (2%)	7	32
2	D	245/259 (95%)	209 (85%)	27 (11%)	9 (4%)	4	23
2	E	244/259 (94%)	205 (84%)	32 (13%)	7 (3%)	6	29
All	All	973/1036 (94%)	801 (82%)	142 (15%)	30 (3%)	5	27

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	62	LEU
2	A	81	LYS
2	A	160	VAL
2	B	62	LEU
2	B	81	LYS

### 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	
2	A	211/242 (87%)	195 (92%)	16 (8%)	16	51
2	B	211/242 (87%)	193 (92%)	18 (8%)	13	45
2	D	218/242 (90%)	203 (93%)	15 (7%)	19	55
2	E	218/242 (90%)	200 (92%)	18 (8%)	14	46
All	All	858/968 (89%)	791 (92%)	67 (8%)	16	49

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	212	VAL
2	D	107	HIS
2	E	201	ASN
2	B	216	HIS
2	D	8	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	215	HIS
2	D	95	ASN
2	E	182	ASN
2	D	64	ASN
2	D	109	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

4 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	SFG	A	300	-	21,29,29	0.94	1 (4%)	19,42,42	3.28	8 (42%)
3	SFG	B	301	-	21,29,29	1.10	1 (4%)	19,42,42	3.17	6 (31%)
3	SFG	D	302	-	21,29,29	1.07	1 (4%)	19,42,42	2.87	3 (15%)
3	SFG	E	303	-	21,29,29	1.10	1 (4%)	19,42,42	2.87	4 (21%)

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	SFG	A	300	-	-	0/9/33/33	0/3/3/3
3	SFG	B	301	-	-	0/9/33/33	0/3/3/3
3	SFG	D	302	-	-	0/9/33/33	0/3/3/3
3	SFG	E	303	-	-	0/9/33/33	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	302	SFG	C2-N3	2.61	1.36	1.32
3	E	303	SFG	C2-N3	2.61	1.36	1.32
3	A	300	SFG	C2-N3	2.84	1.37	1.32
3	B	301	SFG	C2-N3	2.98	1.37	1.32

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	300	SFG	N3-C2-N1	-10.37	120.96	128.89
3	B	301	SFG	N3-C2-N1	-10.23	121.06	128.89
3	D	302	SFG	N3-C2-N1	-10.23	121.06	128.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	E	303	SFG	N3-C2-N1	-10.19	121.09	128.89
3	D	302	SFG	C1'-N9-C4	-3.59	121.53	126.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	300	SFG	3	0
3	D	302	SFG	4	0
3	E	303	SFG	5	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.