



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:14 PM GMT

PDB ID : 4HP7  
Title : Trioxacarcin D517 as a product of guanine robbery from d(AACCGGTT)  
Authors : Smaltz, D.J.; Magauer, T.; Proepper, K.; Dittrich, B.; Myers, A.G.  
Deposited on : 2012-10-23  
Resolution : 1.09 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) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

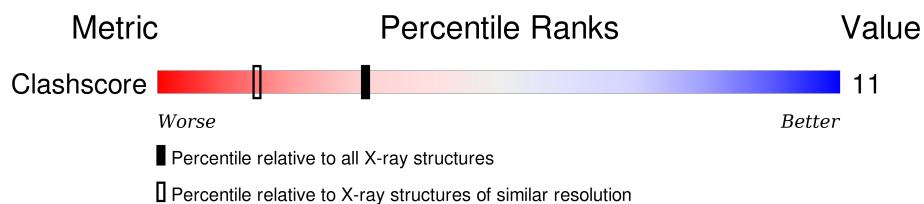
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.09 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.



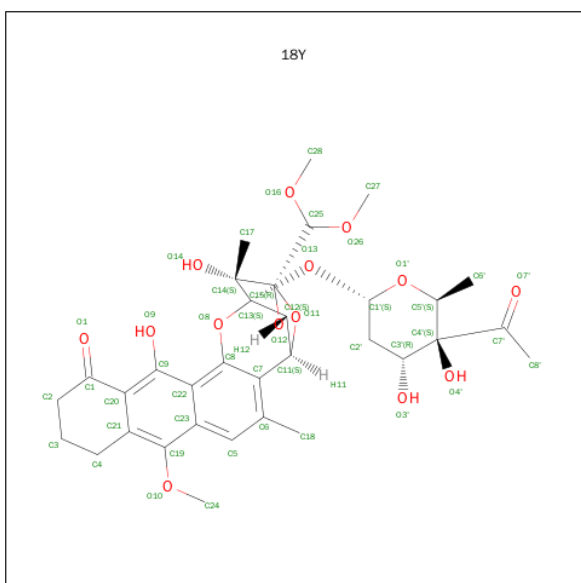
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	1055 (1.14-1.06)

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 451 atoms, of which 192 are hydrogens and 0 are deuteriums.

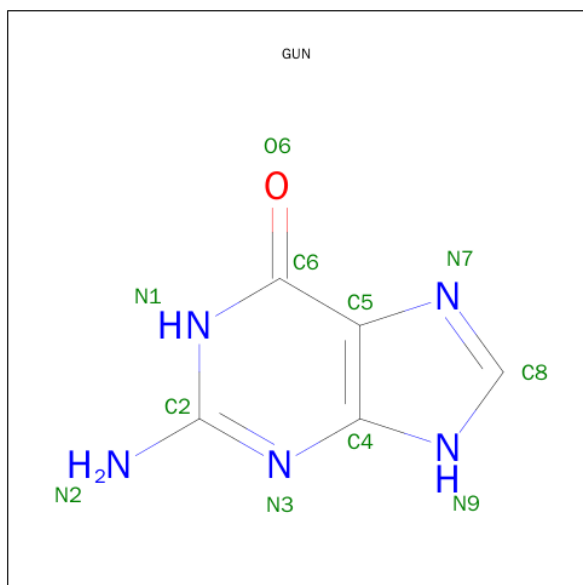
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is SUGAR (TRIOXACARCIN A ANALOGUE, BOUND FORM) (three-letter code: 18Y) (formula:  $C_{33}H_{40}O_{14}$ ).



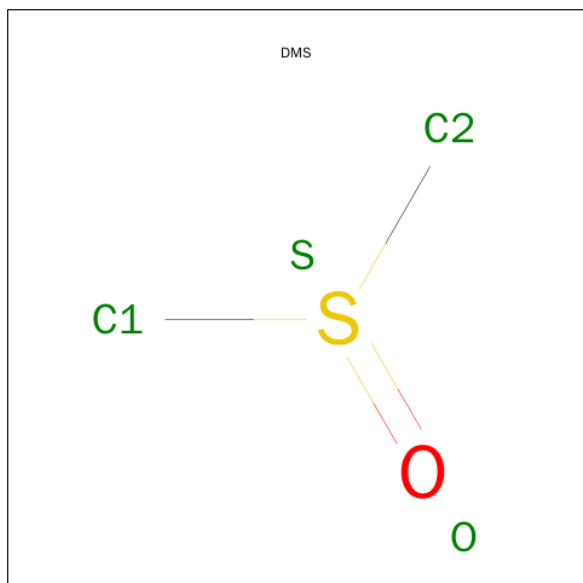
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
1	A	1	Total	C	H	O	0	1
			90	34	41	15		
1	B	1	Total	C	H	O	0	0
			84	33	37	14		
1	C	1	Total	C	H	O	0	0
			84	33	37	14		
1	D	1	Total	C	H	O	0	1
			86	33	37	16		

- Molecule 2 is GUANINE (three-letter code: GUN) (formula:  $C_5H_5N_5O$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	0	1
			15	5	4	5	1		
2	B	1	Total	C	H	N	O	0	0
			15	5	4	5	1		
2	C	1	Total	C	H	N	O	0	0
			15	5	4	5	1		
2	D	1	Total	C	H	N	O	0	0
			15	5	4	5	1		

- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula:  $C_2H_6OS$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	C	1	Total 10	C 2	H 6	O 1	S 1	0	0
3	C	1	Total 10	C 2	H 6	O 1	S 1	0	0
3	D	1	Total 20	C 4	H 12	O 2	S 2	0	1

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	7	Total 7	O 7	0	0

### 3 Residue-property plots

There is no protein, DNA or RNA chain in this entry to show sequence plots.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	24.09 Å   9.83 Å   35.09 Å 90.00°   92.32°   90.00°	Depositor
Resolution (Å)	40.00 – 1.09 35.06 – 1.09	Depositor EDS
% Data completeness (in resolution range)	96.6 (40.00-1.09) 97.0 (35.06-1.09)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.18 (at 1.09 Å)	Xtriage
Refinement program	SHELXL-2012	Depositor
R, $R_{free}$	0.153   ,   0.216 (Not available)   ,   (Not available)	Depositor DCC
$R_{free}$ test set	NotAvailable	DCC
Wilson B-factor (Å <sup>2</sup> )	7.7	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41   ,   104.4	EDS
Estimated twinning fraction	0.000 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.54$ , $\langle L^2 \rangle = 0.38$	Xtriage
Outliers	10 of 41897 reflections (0.024%)	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	451	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	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 61.51 % 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 1.2961e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 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: 18Y, DMS, GUN

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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	49	41	9	0	0
1	B	47	37	37	0	0
1	C	47	37	38	1	0
1	D	49	37	13	0	0
2	A	11	4	0	0	0
2	B	11	4	5	1	0
2	C	11	4	5	1	0
2	D	11	4	5	0	0
3	C	8	12	12	0	0
3	D	8	12	12	0	0
4	D	7	0	0	0	0
All	All	259	192	136	2	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 11.

All (2) 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:B:102:GUN:HN9	2:C:102:GUN:HN21	1.40	0.68
1:C:101:18Y:O3'	1:C:101:18Y:C8'	2.67	0.43

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

### 5.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

16 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
1	18Y	A	101[A]	-	44,51,53	4.09	22 (50%)	57,85,88	2.47	18 (31%)
1	18Y	A	101[B]	-	44,51,53	4.08	22 (50%)	57,85,88	2.46	17 (29%)
1	18Y	A	101[C]	-	44,51,53	4.08	22 (50%)	57,85,88	2.37	17 (29%)
2	GUN	A	102[A]	-	9,12,12	3.37	6 (66%)	5,17,17	9.28	4 (80%)
2	GUN	A	102[B]	-	9,12,12	3.37	6 (66%)	5,17,17	9.28	4 (80%)
1	18Y	B	101	2	45,53,53	2.76	23 (51%)	56,88,88	3.13	19 (33%)
2	GUN	B	102	1	9,12,12	1.00	0	5,17,17	5.51	2 (40%)
1	18Y	C	101	2	45,53,53	2.21	14 (31%)	56,88,88	2.01	15 (26%)
2	GUN	C	102	1	9,12,12	2.80	4 (44%)	5,17,17	3.27	3 (60%)
3	DMS	C	106	-	3,3,3	0.62	0	3,3,3	0.27	0
3	DMS	C	107	-	3,3,3	0.61	0	3,3,3	1.53	0
1	18Y	D	101[A]	-	45,53,53	3.04	16 (35%)	56,88,88	3.27	21 (37%)
1	18Y	D	101[B]	-	45,53,53	2.99	16 (35%)	56,88,88	3.06	22 (39%)
2	GUN	D	102	1	9,12,12	2.19	3 (33%)	5,17,17	3.93	3 (60%)
3	DMS	D	108[A]	-	3,3,3	0.68	0	3,3,3	1.00	0
3	DMS	D	108[B]	-	3,3,3	0.85	0	3,3,3	1.22	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	18Y	A	101[A]	-	-	0/16/94/100	0/1/7/7
1	18Y	A	101[B]	-	-	0/16/94/100	0/1/7/7
1	18Y	A	101[C]	-	-	0/16/94/100	0/1/7/7
2	GUN	A	102[A]	-	-	0/0/0/0	0/2/2/2
2	GUN	A	102[B]	-	-	0/0/0/0	0/2/2/2
1	18Y	B	101	2	-	0/20/100/100	0/1/7/7
2	GUN	B	102	1	-	0/0/0/0	0/2/2/2
1	18Y	C	101	2	-	0/20/100/100	0/1/7/7
2	GUN	C	102	1	-	0/0/0/0	0/2/2/2
3	DMS	C	106	-	-	0/0/0/0	0/0/0/0
3	DMS	C	107	-	-	0/0/0/0	0/0/0/0
1	18Y	D	101[A]	-	-	0/20/100/100	0/1/7/7
1	18Y	D	101[B]	-	-	0/20/100/100	0/1/7/7
2	GUN	D	102	1	-	0/0/0/0	0/2/2/2
3	DMS	D	108[A]	-	-	0/0/0/0	0/0/0/0
3	DMS	D	108[B]	-	-	0/0/0/0	0/0/0/0

The worst 5 of 154 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	101[A]	18Y	O1'-C5'	-14.96	1.23	1.44
1	A	101[C]	18Y	O1'-C5'	-14.96	1.23	1.44
1	A	101[B]	18Y	O1'-C5'	-14.96	1.23	1.44
1	A	101[A]	18Y	O3'-C3'	-11.41	1.24	1.43
1	A	101[C]	18Y	O3'-C3'	-11.41	1.24	1.43

The worst 5 of 145 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	101[A]	18Y	O26-C25-O16	-15.75	85.95	110.69
2	A	102[A]	GUN	C4-C5-N7	-9.16	101.05	109.48
2	A	102[B]	GUN	C4-C5-N7	-9.16	101.05	109.48
1	B	101	18Y	O26-C25-O16	-6.58	100.35	110.69
1	B	101	18Y	O11-C15-O12	-6.39	95.93	106.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	102	GUN	1	0
1	C	101	18Y	1	0
2	C	102	GUN	1	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](#)

There are no RSRZ outliers to report within protein, DNA, RNA chains in this entry.

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	18Y	A	101[A]	45/47	0.94	0.10	-	5,10,51,52	5
2	GUN	D	102	11/11	0.98	0.07	-	5,6,7,8	0
3	DMS	D	108[A]	4/4	0.92	0.39	-	15,22,23,23	10
1	18Y	A	101[C]	45/47	0.94	0.10	-	5,10,34,51	5
1	18Y	A	101[B]	45/47	0.94	0.10	-	5,10,51,52	5
1	18Y	C	101	47/47	0.96	0.08	-	6,11,28,31	0
1	18Y	D	101[A]	47/47	0.95	0.09	-	6,11,31,34	11
1	18Y	D	101[B]	47/47	0.95	0.09	-	6,11,31,34	11
2	GUN	C	102	11/11	0.98	0.06	-	4,6,8,9	0
3	DMS	C	107	4/4	0.97	0.09	-	15,23,29,29	0
2	GUN	A	102[A]	11/11	0.97	0.08	-	5,6,7,8	1
1	18Y	B	101	47/47	0.96	0.08	-	6,13,31,47	0
2	GUN	A	102[B]	11/11	0.97	0.08	-	5,6,7,8	1
3	DMS	D	108[B]	4/4	0.92	0.39	-	15,22,23,23	10
3	DMS	C	106	4/4	0.94	0.17	-	15,22,23,23	0
2	GUN	B	102	11/11	0.98	0.07	-	4,5,6,7	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.