



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 04:03 pm GMT

PDB ID : 4NEC
Title : Conversion of a Disulfide Bond into a Thioacetal Group during Echinomycin Biosynthesis
Authors : Hotta, K.; Keegan, R.M.; Ranganathan, S.; Fang, M.; Bibby, J.; Winn, M.D.; Sato, M.; Lian, M.; Watanabe, K.; Rigden, D.J.; Kim, C.-Y.
Deposited on : 2013-10-29
Resolution : 1.50 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

<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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

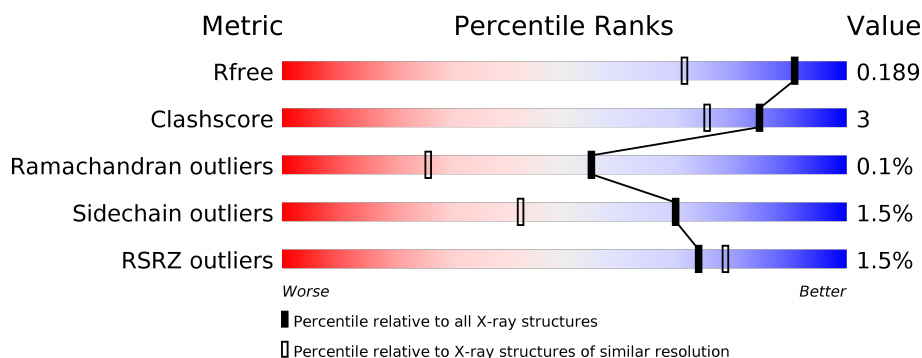
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 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}	100719	2279 (1.50-1.50)
Clashscore	112137	2503 (1.50-1.50)
Ramachandran outliers	110173	2445 (1.50-1.50)
Sidechain outliers	110143	2443 (1.50-1.50)
RSRZ outliers	101464	2305 (1.50-1.50)

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. 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.

Mol	Chain	Length	Quality of chain
1	A	264	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>5% • 10%</div> </div> </div>
1	B	264	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>5% • 8%</div> </div> </div>
1	C	264	<div> <div>%</div> <div> <div></div> <div>84%</div> <div>• 12%</div> </div> </div>
1	D	264	<div> <div></div> <div> <div></div> <div>85%</div> <div>6% 9%</div> </div> </div>
2	E	8	<div> <div></div> <div> <div></div> <div>75%</div> <div>25%</div> </div> </div>
2	F	8	<div> <div></div> <div> <div></div> <div>100%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	8	
2	H	8	

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
4	PGE	A	402	-	-	-	X
4	PGE	D	404	-	-	-	X
7	ACT	B	402	-	-	X	-
8	15P	D	402	-	-	-	X
8	15P	D	403	-	-	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 8028 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 Putative SAM-dependent methyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	238	Total	C	N	O	S	0	2	0
			1801	1129	318	349	5			
1	B	244	Total	C	N	O	S	0	1	0
			1858	1160	334	359	5			
1	C	233	Total	C	N	O	S	0	4	0
			1793	1122	320	346	5			
1	D	239	Total	C	N	O	S	0	1	0
			1799	1128	318	348	5			

There are 160 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q0X0A7
A	-18	GLY	-	EXPRESSION TAG	UNP Q0X0A7
A	-17	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	-16	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	-15	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	-14	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	-13	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	-12	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	-11	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	-10	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	-9	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	-8	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	-7	GLY	-	EXPRESSION TAG	UNP Q0X0A7
A	-6	LEU	-	EXPRESSION TAG	UNP Q0X0A7
A	-5	VAL	-	EXPRESSION TAG	UNP Q0X0A7
A	-4	PRO	-	EXPRESSION TAG	UNP Q0X0A7
A	-3	ARG	-	EXPRESSION TAG	UNP Q0X0A7
A	-2	GLY	-	EXPRESSION TAG	UNP Q0X0A7
A	-1	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	0	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	225	TRP	-	EXPRESSION TAG	UNP Q0X0A7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	226	ASN	-	EXPRESSION TAG	UNP Q0X0A7
A	227	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	228	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	229	SER	-	EXPRESSION TAG	UNP Q0X0A7
A	230	VAL	-	EXPRESSION TAG	UNP Q0X0A7
A	231	ASP	-	EXPRESSION TAG	UNP Q0X0A7
A	232	LYS	-	EXPRESSION TAG	UNP Q0X0A7
A	233	LEU	-	EXPRESSION TAG	UNP Q0X0A7
A	234	ALA	-	EXPRESSION TAG	UNP Q0X0A7
A	235	ALA	-	EXPRESSION TAG	UNP Q0X0A7
A	236	ALA	-	EXPRESSION TAG	UNP Q0X0A7
A	237	LEU	-	EXPRESSION TAG	UNP Q0X0A7
A	238	GLU	-	EXPRESSION TAG	UNP Q0X0A7
A	239	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	240	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	241	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	242	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	243	HIS	-	EXPRESSION TAG	UNP Q0X0A7
A	244	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	-19	MET	-	EXPRESSION TAG	UNP Q0X0A7
B	-18	GLY	-	EXPRESSION TAG	UNP Q0X0A7
B	-17	SER	-	EXPRESSION TAG	UNP Q0X0A7
B	-16	SER	-	EXPRESSION TAG	UNP Q0X0A7
B	-15	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	-14	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	-13	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	-12	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	-11	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	-10	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	-9	SER	-	EXPRESSION TAG	UNP Q0X0A7
B	-8	SER	-	EXPRESSION TAG	UNP Q0X0A7
B	-7	GLY	-	EXPRESSION TAG	UNP Q0X0A7
B	-6	LEU	-	EXPRESSION TAG	UNP Q0X0A7
B	-5	VAL	-	EXPRESSION TAG	UNP Q0X0A7
B	-4	PRO	-	EXPRESSION TAG	UNP Q0X0A7
B	-3	ARG	-	EXPRESSION TAG	UNP Q0X0A7
B	-2	GLY	-	EXPRESSION TAG	UNP Q0X0A7
B	-1	SER	-	EXPRESSION TAG	UNP Q0X0A7
B	0	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	225	TRP	-	EXPRESSION TAG	UNP Q0X0A7
B	226	ASN	-	EXPRESSION TAG	UNP Q0X0A7
B	227	SER	-	EXPRESSION TAG	UNP Q0X0A7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	228	SER	-	EXPRESSION TAG	UNP Q0X0A7
B	229	SER	-	EXPRESSION TAG	UNP Q0X0A7
B	230	VAL	-	EXPRESSION TAG	UNP Q0X0A7
B	231	ASP	-	EXPRESSION TAG	UNP Q0X0A7
B	232	LYS	-	EXPRESSION TAG	UNP Q0X0A7
B	233	LEU	-	EXPRESSION TAG	UNP Q0X0A7
B	234	ALA	-	EXPRESSION TAG	UNP Q0X0A7
B	235	ALA	-	EXPRESSION TAG	UNP Q0X0A7
B	236	ALA	-	EXPRESSION TAG	UNP Q0X0A7
B	237	LEU	-	EXPRESSION TAG	UNP Q0X0A7
B	238	GLU	-	EXPRESSION TAG	UNP Q0X0A7
B	239	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	240	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	241	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	242	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	243	HIS	-	EXPRESSION TAG	UNP Q0X0A7
B	244	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	-19	MET	-	EXPRESSION TAG	UNP Q0X0A7
C	-18	GLY	-	EXPRESSION TAG	UNP Q0X0A7
C	-17	SER	-	EXPRESSION TAG	UNP Q0X0A7
C	-16	SER	-	EXPRESSION TAG	UNP Q0X0A7
C	-15	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	-14	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	-13	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	-12	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	-11	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	-10	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	-9	SER	-	EXPRESSION TAG	UNP Q0X0A7
C	-8	SER	-	EXPRESSION TAG	UNP Q0X0A7
C	-7	GLY	-	EXPRESSION TAG	UNP Q0X0A7
C	-6	LEU	-	EXPRESSION TAG	UNP Q0X0A7
C	-5	VAL	-	EXPRESSION TAG	UNP Q0X0A7
C	-4	PRO	-	EXPRESSION TAG	UNP Q0X0A7
C	-3	ARG	-	EXPRESSION TAG	UNP Q0X0A7
C	-2	GLY	-	EXPRESSION TAG	UNP Q0X0A7
C	-1	SER	-	EXPRESSION TAG	UNP Q0X0A7
C	0	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	225	TRP	-	EXPRESSION TAG	UNP Q0X0A7
C	226	ASN	-	EXPRESSION TAG	UNP Q0X0A7
C	227	SER	-	EXPRESSION TAG	UNP Q0X0A7
C	228	SER	-	EXPRESSION TAG	UNP Q0X0A7
C	229	SER	-	EXPRESSION TAG	UNP Q0X0A7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	230	VAL	-	EXPRESSION TAG	UNP Q0X0A7
C	231	ASP	-	EXPRESSION TAG	UNP Q0X0A7
C	232	LYS	-	EXPRESSION TAG	UNP Q0X0A7
C	233	LEU	-	EXPRESSION TAG	UNP Q0X0A7
C	234	ALA	-	EXPRESSION TAG	UNP Q0X0A7
C	235	ALA	-	EXPRESSION TAG	UNP Q0X0A7
C	236	ALA	-	EXPRESSION TAG	UNP Q0X0A7
C	237	LEU	-	EXPRESSION TAG	UNP Q0X0A7
C	238	GLU	-	EXPRESSION TAG	UNP Q0X0A7
C	239	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	240	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	241	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	242	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	243	HIS	-	EXPRESSION TAG	UNP Q0X0A7
C	244	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	-19	MET	-	EXPRESSION TAG	UNP Q0X0A7
D	-18	GLY	-	EXPRESSION TAG	UNP Q0X0A7
D	-17	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	-16	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	-15	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	-14	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	-13	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	-12	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	-11	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	-10	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	-9	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	-8	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	-7	GLY	-	EXPRESSION TAG	UNP Q0X0A7
D	-6	LEU	-	EXPRESSION TAG	UNP Q0X0A7
D	-5	VAL	-	EXPRESSION TAG	UNP Q0X0A7
D	-4	PRO	-	EXPRESSION TAG	UNP Q0X0A7
D	-3	ARG	-	EXPRESSION TAG	UNP Q0X0A7
D	-2	GLY	-	EXPRESSION TAG	UNP Q0X0A7
D	-1	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	0	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	225	TRP	-	EXPRESSION TAG	UNP Q0X0A7
D	226	ASN	-	EXPRESSION TAG	UNP Q0X0A7
D	227	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	228	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	229	SER	-	EXPRESSION TAG	UNP Q0X0A7
D	230	VAL	-	EXPRESSION TAG	UNP Q0X0A7
D	231	ASP	-	EXPRESSION TAG	UNP Q0X0A7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	232	LYS	-	EXPRESSION TAG	UNP Q0X0A7
D	233	LEU	-	EXPRESSION TAG	UNP Q0X0A7
D	234	ALA	-	EXPRESSION TAG	UNP Q0X0A7
D	235	ALA	-	EXPRESSION TAG	UNP Q0X0A7
D	236	ALA	-	EXPRESSION TAG	UNP Q0X0A7
D	237	LEU	-	EXPRESSION TAG	UNP Q0X0A7
D	238	GLU	-	EXPRESSION TAG	UNP Q0X0A7
D	239	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	240	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	241	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	242	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	243	HIS	-	EXPRESSION TAG	UNP Q0X0A7
D	244	HIS	-	EXPRESSION TAG	UNP Q0X0A7

- Molecule 2 is a protein called Echinomycin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	8	Total	C	N	O	S	0	0	0
			53	33	8	10	2			
2	F	8	Total	C	N	O	S	0	0	0
			53	33	8	10	2			
2	G	8	Total	C	N	O	S	0	0	0
			53	33	8	10	2			
2	H	8	Total	C	N	O	S	0	0	0
			53	33	8	10	2			

- Molecule 3 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: C₁₄H₂₀N₆O₅S).



- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $\text{C}_6\text{H}_{14}\text{O}_4$).

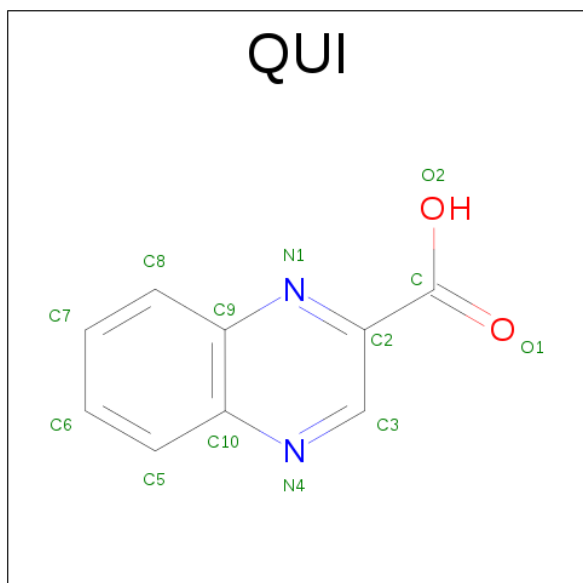


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			7	4	3		
4	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Na	0	0
			2	2		
5	C	1	Total	Na	0	0
			1	1		

- Molecule 6 is 2-CARBOXYQUINOXALINE (three-letter code: QUI) (formula: C₉H₆N₂O₂).



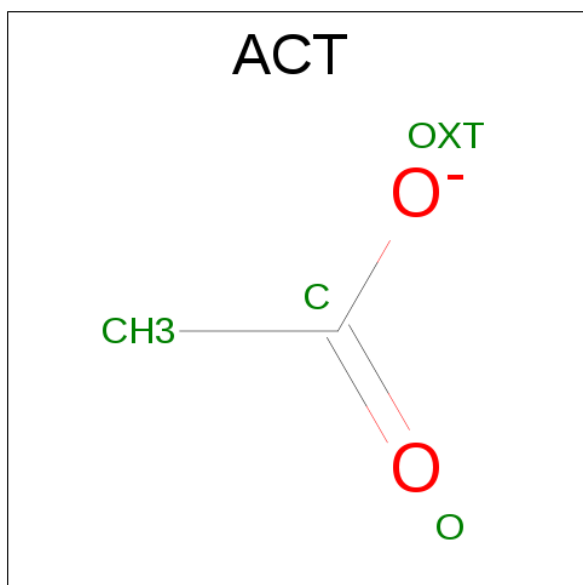
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	E	1	Total	C	N	O	0	0
			12	9	2	1		
6	E	1	Total	C	N	O	0	0
			12	9	2	1		
6	F	1	Total	C	N	O	0	0
			12	9	2	1		
6	F	1	Total	C	N	O	0	0
			12	9	2	1		
6	G	1	Total	C	N	O	0	0
			12	9	2	1		
6	G	1	Total	C	N	O	0	0
			12	9	2	1		

Continued on next page...

Continued from previous page...

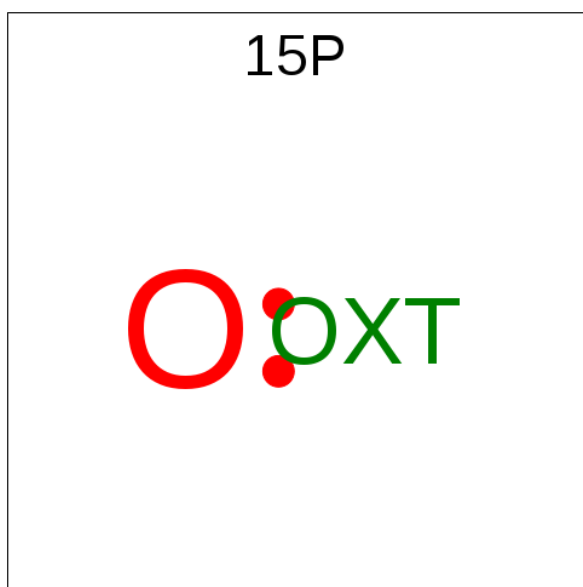
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	H	1	Total	C	N	O	0	0
			12	9	2	1		
6	H	1	Total	C	N	O	0	0
			12	9	2	1		

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is POLYETHYLENE GLYCOL (N=34) (three-letter code: 15P) (formula: $C_{69}H_{140}O_{35}$).

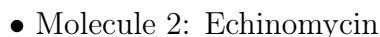
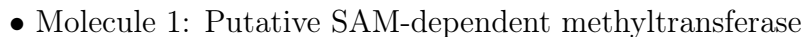
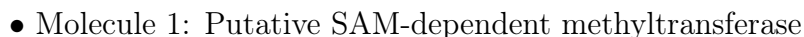
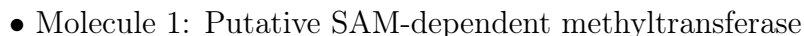


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	D	1	Total	C	O	0	0
			9	6	3		
8	D	1	Total	C	O	0	0
			18	13	5		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	54	Total	O	0	0
			54	54		
9	B	86	Total	O	0	0
			86	86		
9	F	2	Total	O	0	0
			2	2		
9	C	78	Total	O	0	0
			78	78		
9	G	2	Total	O	0	0
			2	2		
9	D	97	Total	O	0	0
			97	97		
9	H	1	Total	O	0	0
			1	1		

- Molecule 1: Putative SAM-dependent methyltransferase




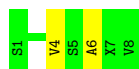
- Molecule 2: Echinomycin

Chain F:  100%

There are no outlier residues recorded for this chain.

- Molecule 2: Echinomycin

Chain G:  75% 25%



- Molecule 2: Echinomycin

Chain H:  63% 25% 13%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	58.19Å 146.88Å 58.19Å 90.00° 102.63° 90.00°	Depositor
Resolution (Å)	45.42 – 1.50 45.42 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.42-1.50) 99.7 (45.42-1.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.60 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.150 , 0.190 0.159 , 0.189	Depositor DCC
R_{free} test set	7590 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	15.7	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 19.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.447 for l,-k,h	Xtriage
Reported twinning fraction	0.530 for H, K, L 0.470 for L, -K, H	Depositor
Outliers	0 of 151251 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8028	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: NCY, SAH, 15P, N2C, NA, DSN, MVA, PGE, ACT, QUI

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	0.28	0/1842	0.50	0/2505
1	B	0.29	0/1904	0.53	0/2591
1	C	0.30	0/1836	0.51	0/2498
1	D	0.30	0/1841	0.54	0/2505
2	E	0.42	0/8	0.96	0/8
2	F	0.51	0/8	1.23	0/8
2	G	0.37	0/8	0.79	0/8
2	H	0.45	0/8	1.37	0/8
All	All	0.29	0/7455	0.52	0/10131

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
2	E	0	1
2	G	0	1
2	H	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	6	ALA	Peptide
2	G	6	ALA	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	H	6	ALA	Peptide

5.2 Too-close contacts ⓘ

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	1801	0	1740	9	0
1	B	1858	0	1787	12	0
1	C	1793	0	1714	6	0
1	D	1799	0	1737	9	0
2	E	53	0	54	1	0
2	F	53	0	54	0	0
2	G	53	0	54	2	0
2	H	53	0	54	2	0
3	A	26	0	19	0	0
3	B	26	0	19	0	0
3	C	26	0	19	0	0
3	D	26	0	19	1	0
4	A	7	0	9	0	0
4	D	4	0	4	5	0
5	A	2	0	0	0	0
5	C	1	0	0	0	0
6	E	24	0	10	0	0
6	F	24	0	10	0	0
6	G	24	0	10	0	0
6	H	24	0	10	0	0
7	B	4	0	3	2	0
8	D	27	0	26	1	0
9	A	54	0	0	0	0
9	B	86	0	0	1	0
9	C	78	0	0	0	0
9	D	97	0	0	1	0
9	F	2	0	0	0	0
9	G	2	0	0	0	0
9	H	1	0	0	0	0
All	All	8028	0	7352	41	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 3.

The worst 5 of 41 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:C:106[A]:ASP:OD1	1:C:134[A]:ARG:HD3	1.47	1.13
1:B:33:LEU:HD23	1:B:39:ILE:HD11	1.65	0.79
1:D:155:ARG:NH1	9:D:596:HOH:O	2.22	0.73
1:A:212:LEU:HD13	1:A:214:VAL:HG13	1.74	0.69
1:B:18[B]:ARG:NH2	1:B:22:ASP:O	2.26	0.68

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	236/264 (89%)	230 (98%)	6 (2%)	0	100	100
1	B	243/264 (92%)	236 (97%)	6 (2%)	1 (0%)	38	14
1	C	233/264 (88%)	227 (97%)	6 (3%)	0	100	100
1	D	236/264 (89%)	232 (98%)	4 (2%)	0	100	100
2	E	2/8 (25%)	2 (100%)	0	0	100	100
2	F	2/8 (25%)	2 (100%)	0	0	100	100
2	G	2/8 (25%)	2 (100%)	0	0	100	100
2	H	2/8 (25%)	2 (100%)	0	0	100	100
All	All	956/1088 (88%)	933 (98%)	22 (2%)	1 (0%)	55	26

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	119	GLY

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	179/204 (88%)	176 (98%)	3 (2%)	66	36
1	B	187/204 (92%)	185 (99%)	2 (1%)	78	56
1	C	177/204 (87%)	175 (99%)	2 (1%)	78	56
1	D	179/204 (88%)	175 (98%)	4 (2%)	57	24
All	All	722/816 (88%)	711 (98%)	11 (2%)	70	42

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	243	HIS
1	C	1	MET
1	D	40	GLN
1	B	115	HIS
1	D	3	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

24 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	DSN	E	1	2,6	5,5,6	0.40	0	1,5,7	2.20	1 (100%)
2	N2C	E	3	2	7,7,8	1.13	1 (14%)	6,7,9	2.32	2 (33%)
2	MVA	E	4	2	7,7,8	0.45	0	7,8,10	1.29	1 (14%)
2	DSN	E	5	2,6	5,5,6	0.54	0	1,5,7	2.48	1 (100%)
2	NCY	E	7	2	6,6,7	1.22	1 (16%)	6,6,8	1.57	1 (16%)
2	MVA	E	8	2	7,7,8	0.38	0	7,8,10	0.71	0
2	DSN	F	1	2,6	5,5,6	0.50	0	1,5,7	2.27	1 (100%)
2	N2C	F	3	2	7,7,8	0.96	0	6,7,9	1.44	2 (33%)
2	MVA	F	4	2	7,7,8	0.39	0	7,8,10	1.26	1 (14%)
2	DSN	F	5	2,6	5,5,6	0.49	0	1,5,7	2.16	1 (100%)
2	NCY	F	7	2	6,6,7	1.26	1 (16%)	6,6,8	1.07	1 (16%)
2	MVA	F	8	2	7,7,8	0.36	0	7,8,10	0.85	0
2	DSN	G	1	2,6	5,5,6	0.43	0	1,5,7	1.98	0
2	N2C	G	3	2	7,7,8	1.22	1 (14%)	6,7,9	2.10	2 (33%)
2	MVA	G	4	2	7,7,8	0.28	0	7,8,10	1.20	0
2	DSN	G	5	2,6	5,5,6	0.45	0	1,5,7	2.28	1 (100%)
2	NCY	G	7	2	6,6,7	1.21	1 (16%)	6,6,8	1.43	1 (16%)
2	MVA	G	8	2	7,7,8	0.39	0	7,8,10	0.78	0
2	DSN	H	1	2,6	5,5,6	0.49	0	1,5,7	2.49	1 (100%)
2	N2C	H	3	2	7,7,8	1.27	1 (14%)	6,7,9	1.48	1 (16%)
2	MVA	H	4	2	7,7,8	0.34	0	7,8,10	1.11	2 (28%)
2	DSN	H	5	2,6	5,5,6	0.43	0	1,5,7	2.03	1 (100%)
2	NCY	H	7	2	6,6,7	1.26	1 (16%)	6,6,8	1.52	1 (16%)
2	MVA	H	8	2	7,7,8	0.41	0	7,8,10	0.53	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
2	DSN	E	1	2,6	-	0/2/4/6	0/0/0/0
2	N2C	E	3	2	-	0/3/7/9	0/0/0/0
2	MVA	E	4	2	-	0/5/8/10	0/0/0/0
2	DSN	E	5	2,6	-	0/2/4/6	0/0/0/0
2	NCY	E	7	2	-	0/2/6/8	0/0/0/0
2	MVA	E	8	2	-	0/5/8/10	0/0/0/0
2	DSN	F	1	2,6	-	0/2/4/6	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	N2C	F	3	2	-	0/3/7/9	0/0/0/0
2	MVA	F	4	2	-	0/5/8/10	0/0/0/0
2	DSN	F	5	2,6	-	0/2/4/6	0/0/0/0
2	NCY	F	7	2	-	0/2/6/8	0/0/0/0
2	MVA	F	8	2	-	0/5/8/10	0/0/0/0
2	DSN	G	1	2,6	-	0/2/4/6	0/0/0/0
2	N2C	G	3	2	-	0/3/7/9	0/0/0/0
2	MVA	G	4	2	-	0/5/8/10	0/0/0/0
2	DSN	G	5	2,6	-	0/2/4/6	0/0/0/0
2	NCY	G	7	2	-	0/2/6/8	0/0/0/0
2	MVA	G	8	2	-	0/5/8/10	0/0/0/0
2	DSN	H	1	2,6	-	0/2/4/6	0/0/0/0
2	N2C	H	3	2	-	0/3/7/9	0/0/0/0
2	MVA	H	4	2	-	0/5/8/10	0/0/0/0
2	DSN	H	5	2,6	-	0/2/4/6	0/0/0/0
2	NCY	H	7	2	-	0/2/6/8	0/0/0/0
2	MVA	H	8	2	-	0/5/8/10	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	3	N2C	CB-SG	-2.95	1.77	1.80
2	H	3	N2C	CB-SG	-2.95	1.77	1.80
2	E	3	N2C	CB-SG	-2.26	1.78	1.80
2	H	7	NCY	CB-SG	-2.20	1.76	1.81
2	G	7	NCY	CB-SG	-2.09	1.77	1.81

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	7	NCY	O-C-CA	-3.35	117.34	125.15
2	E	7	NCY	O-C-CA	-3.15	117.81	125.15
2	G	3	N2C	O-C-CA	-3.01	118.14	125.15
2	G	7	NCY	O-C-CA	-2.85	118.51	125.15
2	E	3	N2C	O-C-CA	-2.74	118.76	125.15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	4	MVA	1	0
2	G	4	MVA	2	0
2	H	7	NCY	1	0
2	H	8	MVA	2	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 3 are monoatomic - leaving 17 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$
3	SAH	A	401	-	20,28,28	1.06	1 (5%)	20,40,40	1.75	2 (10%)
4	PGE	A	402	-	6,6,9	0.40	0	5,5,8	0.35	0
3	SAH	B	401	-	20,28,28	1.08	1 (5%)	20,40,40	1.74	2 (10%)
7	ACT	B	402	-	1,3,3	1.33	0	0,3,3	0.00	-
3	SAH	C	401	-	20,28,28	1.06	1 (5%)	20,40,40	1.86	3 (15%)
3	SAH	D	401	-	20,28,28	1.09	1 (5%)	20,40,40	1.76	4 (20%)
8	15P	D	402	-	8,8,103	0.49	0	7,7,102	0.15	0
8	15P	D	403	-	16,16,103	0.55	0	14,14,102	0.13	0
4	PGE	D	404	-	3,3,9	0.40	0	2,2,8	0.28	0
6	QUI	E	101	2	13,13,14	1.68	2 (15%)	15,17,19	2.07	6 (40%)
6	QUI	E	102	2	13,13,14	1.72	2 (15%)	15,17,19	1.81	4 (26%)
6	QUI	F	101	2	13,13,14	1.54	2 (15%)	15,17,19	1.83	6 (40%)
6	QUI	F	102	2	13,13,14	1.65	2 (15%)	15,17,19	2.10	5 (33%)
6	QUI	G	101	2	13,13,14	1.64	3 (23%)	15,17,19	2.16	6 (40%)
6	QUI	G	102	2	13,13,14	1.66	2 (15%)	15,17,19	2.13	6 (40%)
6	QUI	H	101	2	13,13,14	1.68	3 (23%)	15,17,19	2.23	7 (46%)
6	QUI	H	102	2	13,13,14	1.64	2 (15%)	15,17,19	1.95	5 (33%)

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	SAH	A	401	-	-	0/7/31/31	0/3/3/3
4	PGE	A	402	-	-	0/4/4/7	0/0/0/0
3	SAH	B	401	-	-	0/7/31/31	0/3/3/3
7	ACT	B	402	-	-	0/0/0/0	0/0/0/0
3	SAH	C	401	-	-	0/7/31/31	0/3/3/3
3	SAH	D	401	-	-	0/7/31/31	0/3/3/3
8	15P	D	402	-	-	0/6/6/101	0/0/0/0
8	15P	D	403	-	-	0/12/12/101	0/0/0/0
4	PGE	D	404	-	-	0/1/1/7	0/0/0/0
6	QUI	E	101	2	-	0/2/2/4	0/2/2/2
6	QUI	E	102	2	-	0/2/2/4	0/2/2/2
6	QUI	F	101	2	-	0/2/2/4	0/2/2/2
6	QUI	F	102	2	-	0/2/2/4	0/2/2/2
6	QUI	G	101	2	-	0/2/2/4	0/2/2/2
6	QUI	G	102	2	-	0/2/2/4	0/2/2/2
6	QUI	H	101	2	-	0/2/2/4	0/2/2/2
6	QUI	H	102	2	-	0/2/2/4	0/2/2/2

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	101	QUI	C3-N4	2.00	1.35	1.31
6	G	101	QUI	C3-N4	2.00	1.35	1.31
6	F	101	QUI	C2-N1	2.25	1.35	1.33
6	G	101	QUI	C2-N1	2.49	1.35	1.33
6	H	102	QUI	C2-N1	2.54	1.35	1.33

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	401	SAH	N3-C2-N1	-6.51	123.19	128.86
3	A	401	SAH	N3-C2-N1	-6.27	123.40	128.86
3	B	401	SAH	N3-C2-N1	-6.23	123.43	128.86
6	G	101	QUI	O1-C-C2	-5.82	119.28	124.32
3	D	401	SAH	N3-C2-N1	-5.77	123.83	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	402	ACT	2	0
3	D	401	SAH	1	0
8	D	402	15P	1	0
4	D	404	PGE	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](#)

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, 95th 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	238/264 (90%)	-0.45	6 (2%) 58 64	9, 18, 35, 58	0
1	B	244/264 (92%)	-0.48	5 (2%) 65 71	9, 17, 30, 61	1 (0%)
1	C	233/264 (88%)	-0.51	2 (0%) 84 87	8, 19, 31, 45	1 (0%)
1	D	239/264 (90%)	-0.62	1 (0%) 92 94	9, 15, 27, 44	1 (0%)
2	E	2/8 (25%)	-0.39	0 100 100	21, 21, 21, 22	0
2	F	2/8 (25%)	-0.78	0 100 100	16, 16, 16, 17	0
2	G	2/8 (25%)	-0.49	0 100 100	21, 21, 21, 23	0
2	H	2/8 (25%)	-0.93	0 100 100	12, 12, 12, 17	0
All	All	962/1088 (88%)	-0.52	14 (1%) 74 78	8, 17, 31, 61	3 (0%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	174	PRO	5.4
1	A	175	THR	4.0
1	B	167	LEU	3.9
1	B	168	PRO	3.8
1	B	0	HIS	3.8

6.2 Non-standard residues in protein, DNA, RNA chains [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, 95th 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(\AA^2)	Q<0.9
2	NCY	G	7	7/8	0.98	0.07	-	18,20,21,22	0
2	NCY	H	7	7/8	0.99	0.05	-	14,15,15,16	0
2	MVA	G	4	8/9	0.94	0.09	-	24,26,29,30	0
2	NCY	F	7	7/8	0.99	0.05	-	13,14,15,16	0
2	NCY	E	7	7/8	0.99	0.05	-	17,18,19,19	0
2	MVA	F	4	8/9	0.96	0.05	-	15,16,18,20	0
2	MVA	G	8	8/9	0.97	0.05	-	20,22,24,24	0
2	N2C	E	3	8/9	0.98	0.07	-	20,21,22,22	0
2	MVA	F	8	8/9	0.97	0.07	-	13,16,17,17	0
2	N2C	G	3	8/9	0.98	0.06	-	16,22,23,24	0
2	DSN	G	5	6/7	0.96	0.07	-	22,24,25,26	0
2	DSN	E	5	6/7	0.96	0.07	-	21,22,25,26	0
2	DSN	H	5	6/7	0.97	0.06	-	16,16,17,18	0
2	DSN	F	1	6/7	0.98	0.05	-	15,16,16,18	0
2	DSN	E	1	6/7	0.98	0.04	-	17,19,19,20	0
2	N2C	H	3	8/9	0.99	0.05	-	12,13,14,15	0
2	MVA	E	4	8/9	0.93	0.10	-	21,25,27,28	0
2	MVA	H	8	8/9	0.95	0.07	-	13,14,15,15	0
2	MVA	H	4	8/9	0.97	0.05	-	13,14,15,15	0
2	DSN	H	1	6/7	0.98	0.05	-	13,13,14,15	0
2	DSN	F	5	6/7	0.98	0.04	-	16,17,18,18	0
2	MVA	E	8	8/9	0.98	0.05	-	16,18,20,22	0
2	N2C	F	3	8/9	0.99	0.05	-	13,15,16,16	0
2	DSN	G	1	6/7	0.98	0.04	-	18,19,20,22	0

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, 95th 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(\AA^2)	Q<0.9
4	PGE	D	404	4/10	0.84	0.19	13.21	12,14,16,16	4
8	15P	D	402	9/104	0.91	0.17	8.31	16,28,33,36	0
8	15P	D	403	18/104	0.85	0.20	8.17	20,28,34,35	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	PGE	A	402	7/10	0.86	0.14	4.65	29,31,33,34	0
6	QUI	E	101	12/13	0.97	0.08	1.96	19,21,22,23	0
7	ACT	B	402	4/4	0.95	0.11	1.85	16,16,21,24	0
6	QUI	G	102	12/13	0.92	0.13	1.03	26,32,39,39	0
6	QUI	E	102	12/13	0.93	0.09	0.61	21,23,24,24	0
3	SAH	B	401	26/26	0.99	0.06	0.12	10,12,13,15	0
6	QUI	F	101	12/13	0.97	0.07	0.02	16,18,20,21	0
3	SAH	A	401	26/26	0.98	0.06	-0.03	12,14,16,17	0
6	QUI	F	102	12/13	0.96	0.09	-0.11	21,23,25,25	0
3	SAH	D	401	26/26	0.99	0.05	-0.31	9,11,14,15	0
6	QUI	H	102	12/13	0.97	0.06	-0.42	19,20,22,22	0
3	SAH	C	401	26/26	0.98	0.05	-0.56	15,16,18,19	0
5	NA	A	403	1/1	0.97	0.06	-0.57	26,26,26,26	0
6	QUI	G	101	12/13	0.97	0.06	-0.84	21,25,28,29	0
6	QUI	H	101	12/13	0.98	0.05	-0.94	13,16,18,19	0
5	NA	C	402	1/1	0.96	0.06	-1.38	45,45,45,45	0
5	NA	A	404	1/1	0.99	0.03	-1.68	29,29,29,29	0

6.5 Other polymers [i](#)

There are no such residues in this entry.