Acta Crystallographica Section C
Crystal Structure Communications
ISSN 0108-2701

Bis[(1S)-1,4-azanediyl-1-(9-deaza-adenin-9-yl)-1,4-dideoxy-5-methyl-sulfanyl-D-ribitol] tetrakis(hydro-chloride) monohydrate: structure, DFT energy and ligand docking results of a potent methylthioadenosine phosphorylase inhibitor found in different molecular conformations

Graeme J. Gainsford,[a]* Gary B. Evans,[a] Karen A. Johnston[a] and Michael Seth[b]

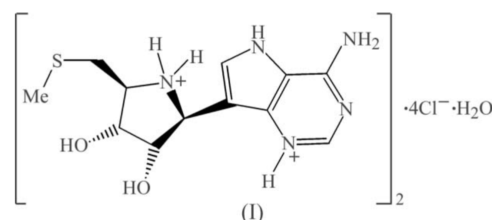
[a]Industrial Research Limited, PO Box 31-310, Lower Hutt, New Zealand, and [b]University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada
Correspondence e-mail: g.gainsford@irl.cri.nz

Received 25 August 2010
Accepted 16 September 2010
Online 8 October 2010

The title compound, abbreviated as 5′ThiomethylImmA, is a potent inhibitor of methylthioadenosine phosphorylase [Singh *et al.* (2004). *Biochemistry*, **43**, 9–18]. The synchrotron study reported here shows that the hydrochloride salt crystallizes with two independent, nearly superimposable, dications as a monohydrate with formula $2C_{12}H_{19}N_5O_2S^{2+}\cdot4Cl^{-}\cdot H_2O$. Hydrogen bonding utilizing the H atoms of the dication is found to favour certain molecular conformations in the salt, which are significantly different from those found as bound in the enzyme. Ligand docking studies starting from either of these dications or related neutral structures successfully place the conformationally revised structures in the enzyme active site but only under particular hydrogen-bonding and molecular flexibility criteria. Density functional theory calculations verify the energy similarity of the independent cations and confirm the significant energy cost of the required conformational change to the enzyme bound form. The results suggest that using crystallographically determined free ligand coordinates as starting parameters for modelling may have serious limitations.

Comment

The title compound, (I), a potent inhibitor of methylthioadenosine phosphorylase (MTAP), was prepared by Evans *et al.* (2004) as part of a synthesis program aimed at anticancer lead compounds. The structure and stereochemistry were established prior to its subsequent definition in the active site of the enzyme [Protein Data Bank (PDB) ID code 1K27 used hereafter), as reported by Singh *et al.* (2004)]. This report notes the details of the crystal structure of the double-protonated (dication) chloride salt, which includes a water of crystallization, (I). Additional information about the computational docking of the compound in the enzyme using the 1K27 coordinates has been determined using the docking program *GOLD* (Jones *et al.*, 1997; Nissink *et al.*, 2002). The results are discussed in the light of the different molecular conformations found for the two independent molecules as compared with that observed bound in the MTAP enzyme structure and the relative energies calculated by DFT (density functional theory) methods (Scientific Computing & Modelling, 2009).



(I)

The asymmetric unit contains two independent dications, four chloride anions and one water molecule of crystallization. These two molecular conformations (molecules 1 and 2 hereafter) have atom labels related by the numerical addition of 10 (*e.g.* N3 and N13, see Fig. 1).

The crystal packing can be described as a one-dimensional array built along the *b* screw axis, with some links along the *c* axis, utilizing a network of hydrogen bonds, *viz.* N—H···Cl, (C)O—H···Cl, O(water)—H···Cl, O(water)—H···N and one O—H···S (Table 1 and Fig. 2). There are no strong links between the lattices, with the methylsulfanyl (*e.g.* C6′, S1 and

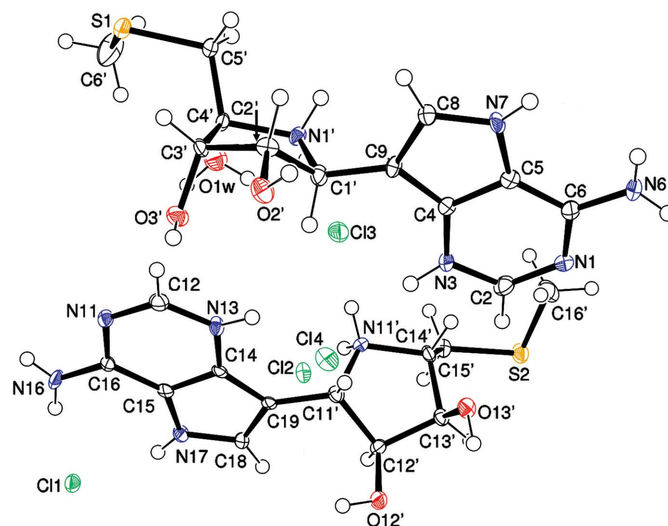


Figure 1
View of the asymmetric unit contents of (I), including two independent cations, shown with 50% displacement ellipsoids (Farrugia, 1997). H atoms have arbitrary radii.

 © 2010 International Union of Crystallography

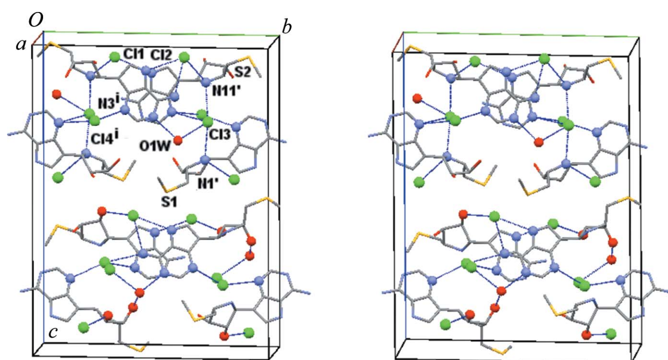


Figure 2
A *Mercury* (Macrae *et al.*, 2008) stereo-packing view of the cell highlighting the separated lattices along the *c* axis. H atoms have been omitted for clarity. Some of the atoms involved in hydrogen bonds are shown in ball mode (see Table 1). [Symmetry code: (i) $1 - x, y - \frac{1}{2}, \frac{1}{2} - z$.]

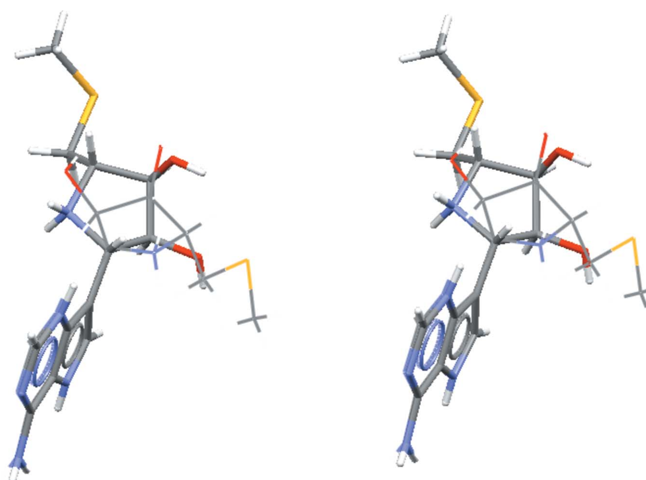


Figure 4
9-Deazaadenin-9-yl-overlapped stereoview of molecule 1 (thick bonds) and the final docked molecule DiMo1 (Table 2, thin bonds), illustrating the conformational differences (*Mercury*; Macrae *et al.*, 2008).

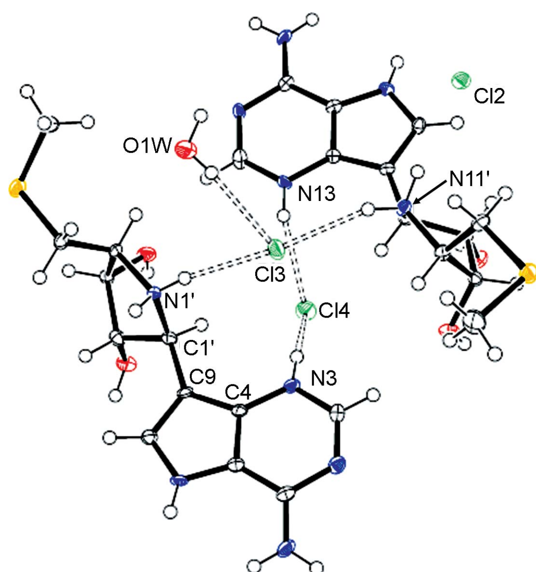


Figure 3
Key hydrogen-bonding interactions in (I), see *Comment*; the asymmetric unit is shown with selected hydrogen bonds as dotted lines (Farrugia, 1997). H atoms have arbitrary radii, while other atoms are shown with 50% displacement ellipsoids.

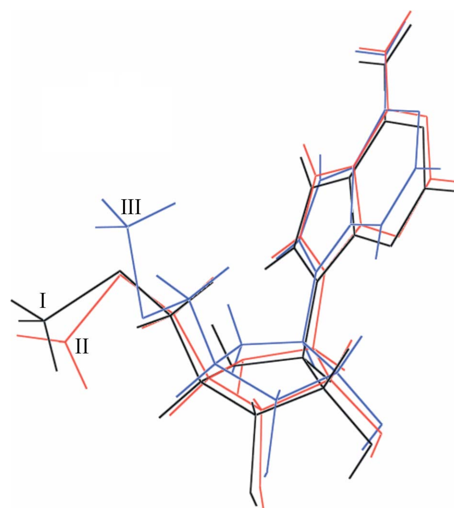


Figure 5
Overlapped views of *GOLD* CHEMSCORE docking models (see *Comment*). Colour key for the electronic version of the paper: black (denoted 'I'), ligand *A* (as in 1K27 with added H atoms); red (denoted 'II'), docked result starting with ligand *A* positions; blue (denoted 'III'), starting with molecule 1 (DiMo1).

C5′) interactions defining the spacing up the *a* cell axes. The basic building blocks consist of $D(2)$ moiety types (Bernstein *et al.*, 1995) which build to larger D_2^2 forms mainly; there are only a few examples of other types, *e.g.* bifurcated binding which generate the $R_2^1(7)$ (entries 7 and 8, Table 1) and $R_2^1(8)$ (entries 5 and 10, Table 1) motifs. The key hydrogen bonds that promote the observed conformations in both molecules, as distinct from that found in the enzyme (Table 2), involve the additional H atoms that give the overall 2+ charge (namely on atoms N3, N13, N1′ and N11′). These latter atoms form interlinking bonds (see entries 4 and 14, and 5 and 12 in Table 1), as shown in Fig. 3, which could not form if the conformations were as found in the enzyme.

It is common practice to take X-ray-diffraction-derived molecular parameters as starting models for ligand docking studies. In this case, it was of some interest to determine if the docking computations using the software package *GOLD* could rotate the two rings ~130° about the C1′—C9 bond to reproduce the observed conformation in the enzyme (see Table 2 and Fig. 4). The initial test starting with the found 1K27 ligand coordinates ('ligand *A*'; with H atoms in calculated positions) without using any constraints was successful. By contrast, neither the dication nor the neutral molecule based on molecules 1 and 2 gave any correctly docked solutions. Only when the pyrrolidine ring was allowed to flex, using the flip ring corners option, did any correct solutions appear; both the dication-based coordinates and the 'ligand *A*' sets gave some good matches but not with the highest docking scores.

To obtain the closest matches with the observed ligand *A* in 1K27, coupled with the highest docking scores for all the models, required three constraints: (i) specifying which enzyme atoms had hydrogen bonds (but not which ligand atoms were involved), (ii) flipping the pyramidal N atom in the pyrrolidine ring and (iii) exploring the ring conformations. Using *GOLD*'s CHEMSCORE and CHEMPLP scoring regimes gave the highest numerical score for the (starting coordinate) 'ligand *A*' set, the neutral and dication coordinates based on molecule 1 scoring somewhat less, but correctly in terms of matching the ligand conformation and position in 1K27. There was some variation in final positioning of the terminal methyl group in the dication run (Table 2, DiMo1 column), as indicated in Fig. 5 and Table 2.

These results led us to calculate the molecular energies using DFT methods (Scientific Computing & Modelling, 2009). The BLYP functional formed from the combination of the local density approximation (LDA) parameterization by Vosko *et al.* (1980), VWN5, with the B88 exchange gradient correction (Becke, 1988) and the LYP correlation gradient correction (Lee *et al.*, 1988) were used. The basis set used was triple-ζ quality for all valence orbitals and included two sets of polarization functions on each atom (Van Lenthe & Baerends, 2003). The polarization functions are '1p1d' on the H atoms and '1d1f' on the C, N and O atoms. The positions of the heavy atoms were frozen at the values taken from the crystal or docked structures but the coordinates of the H atoms were allowed to relax. The energy differences found, from the specified starting models, relative to the docked conformation are shown in Table 2 (final DFT-calculated geometries are not given).

The similarity of results for the molecule 1 and 2 conformations in geometry and energy [difference = 6 kJ mol^{-1} (1.24 kcal mol^{-1})] gives credence to the calculations, which show that the enzyme-bound 'ligand *A*' dication conformation has much higher energy ($\sim$253 kJ mol^{-1}). The energy required to distort the molecule to its docking geometry must be made up by energy gained through favorable interactions with the enzyme. A phosphate species is well placed to interact strongly with the cation. Furthermore, the protein structure and docking studies suggest that when in place the cation will form several hydrogen bonds with its host: four such P—O···H—O,N interactions are observed. Together, these bonds should be able to stabilize the cation in its docked conformation. In the case of the neutral molecule, the docked conformation is found to be lower in energy by 62 kJ mol^{-1} than the conformation suggested by the 1K27 ligand *A* crystal structure.

These latter results strongly suggest that docking of the free ligand molecules 1 and 2 in the 1K27 enzyme (and by analogy for other similar cases) must include hydrogen-bonding interactions. This was also indicated by the final (successful) criteria applied in the *GOLD* docking runs. A sobering corollary to this finding is that crystallographically determined free ligand conformations may not be reproduced in enzyme binding sites, which may limit their potential as starting models for drug discovery.

Experimental

The title compound was prepared as reported for compound 2 by Evans *et al.* (2004).

Crystal data

$2C_{12}H_{19}N_5O_2S^{2+}\cdot4Cl^-\cdot H_2O$	$Z = 4$
$M_r = 754.58$	Synchrotron radiation
Orthorhombic, $P2_12_12_1$	$\lambda = 0.92014$ Å
$a = 7.0080$ (14) Å	$\mu = 1.05$ mm^{-1}
$b = 18.717$ (4) Å	$T = 100$ K
$c = 24.825$ (5) Å	0.23 × 0.20 × 0.15 mm
$V = 3256.3$ (11) Å^3	

Data collection

ADSC Quantum CCD detector diffractometer	3864 independent reflections
3864 measured reflections	3821 reflections with $I > 2\sigma(I)$
	$\theta_{max} = 29.6°$

Refinement

$R[F^2 > 2\sigma(F^2)] = 0.027$	H atoms treated by a mixture of independent and constrained refinement
$wR(F^2) = 0.066$	$\Delta\rho_{max} = 0.40$ e Å^{-3}
$S = 1.06$	$\Delta\rho_{min} = -0.24$ e Å^{-3}
3864 reflections	Absolute structure: Flack (1983), 1604 Friedel pairs
454 parameters	Flack parameter: −0.03 (3)
12 restraints	

The NSLS beamline used was X7B. The N—H distances of two types (*viz*. N3/N7/N13/N17 and N6/N16) were restrained to be equal [*SHELXL97* (Sheldrick, 2008) SADI function with an effective standard uncertainty of 0.03 Å]. The methyl H atoms were constrained to an ideal geometry (C—H = 0.98 Å), but were allowed to rotate freely about the adjacent C—C bond. The hydroxy H atoms were constrained to an ideal geometry (O—H = 0.84 Å), with U_{iso}(H) = 1.2U_{eq}(O), but were allowed to rotate freely about the adjacent C—O bond. All other H atoms were placed in geometrically idealized positions and constrained to ride on their parent atoms, with C—H distances of 1.00 (primary), 0.99 (methylene) or 0.95 Å (aromatic). All H atoms were refined with U_{iso}(H) = 1.2U_{eq}(parent atom).

Data collection: *DENZO* (Otwinowski & Minor, 1997); cell refinement: *DENZO*; data reduction: *DENZO* and *SCALEPACK*

Table 1
Hydrogen-bond geometry (Å, °).

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
O2′—H2O′···Cl1^{i}	0.84	2.27	3.082 (2)	163
O3′—H3O′···O1*W*^{ii}	0.84	1.93	2.724 (4)	158
N1′—H1N′···Cl1^{iii}	0.89 (3)	2.32 (3)	3.199 (3)	169 (3)
N3—H3N···Cl4	0.85 (3)	2.31 (3)	3.097 (3)	154 (3)
N1′—H2N′···Cl3	0.84 (3)	2.29 (3)	3.120 (3)	171 (3)
N6—H6*A*···Cl4^{i}	0.83 (3)	2.62 (3)	3.223 (3)	131 (2)
N6—H6*B*···Cl2^{iii}	0.84 (3)	2.38 (3)	3.196 (3)	165 (3)
N7—H7N···Cl2^{iii}	0.83 (3)	2.42 (3)	3.232 (3)	170 (3)
O1*W*—H1*WA*···N1^{iv}	0.96 (3)	1.87 (4)	2.804 (4)	162 (3)
O1*W*—H2*WA*···Cl3	0.80 (4)	2.35 (4)	3.143 (3)	173 (3)
N11′—H11*A*···Cl2	0.92 (3)	2.35 (3)	3.268 (3)	171 (3)
N11′—H11*B*···Cl3	0.89 (3)	2.34 (3)	3.146 (3)	151 (3)
O12′—H12O···Cl2^{ii}	0.84	2.34	3.145 (2)	162
N13—H13N···Cl4	0.87 (3)	2.48 (3)	3.228 (3)	145 (3)
O13′—H13O···S2^{ii}	0.84	2.57	3.364 (2)	157
N17—H17N···Cl1	0.84 (3)	2.39 (3)	3.197 (3)	163 (3)
N16—H61*A*···Cl1	0.85 (3)	2.51 (3)	3.351 (3)	171 (3)
N16—H61*B*···Cl3^{iv}	0.82 (3)	2.62 (3)	3.311 (3)	143 (3)
C2—H2···O13′	0.95	2.46	3.383 (4)	164

Symmetry codes: (i) $-x+2, y+\frac{1}{2}, -z+\frac{1}{2}$; (ii) $x+1, y, z$; (iii) $-x+1, y+\frac{1}{2}, -z+\frac{1}{2}$; (iv) $-x+1, y-\frac{1}{2}, -z+\frac{1}{2}$.

Table 2
Comparison of torsion angles (°) and other parameters between the two molecules in (I), ligand *A* in 1K27 and docked models (see *Comment*).

	Molecule 1	Molecule 2	Ligand *A* in 1K27†	NeMo1‡	DiMo1§
N1′—C1′—C9—C8	88.6 (4)	88.7 (4)	−54	−56.2	−61.1
C2′—C1′—C9—C4	154.4 (3)	157.9 (3)	11	−0.1	−7.1
N1′—C1′—C2′—C3′	44.0 (3)	46.1 (3)	−0.4	1.8	1.8
C1′—C2′—C3′—C4′	−42.2 (3)	−39.8 (3)	−13	−15.7	−15.7
N1′—C4′—C5′—S1	175.0 (2)	176.5 (2)	50	53.7	173.7
C3′—C4′—C5′—S1	−68.5 (3)	−66.8 (3)	170	170.1	−69.9
C4′—C5′—S1—C6′	−84.8 (3)	−61.1 (3)	76	61.4	172.5
Pyrrole ring¶	Env (C2′)	Twist C11′—C12′	Env (C4′)	Env (C4′)	Env (C4′)
Envelope Δ†† (Å)	0.663 (3)	NA	0.34	0.37	0.37
Q(2) (Å)	0.442 (3)	0.453 (3)	0.22	0.23	0.23
φ(2) (°)	66.1 (4)	58.0 (4)	144	136	136
Relative energy‡‡	−256	−250	62	0	0

† Ligand *A*, as the generated dication (see *Comment*). ‡ NeMo1: docked neutral molecule 1 (without the N3 and one N1′ H atom). § DiMo1: docked from molecule 1 start, in dication form. ¶ Env (C*n*) = Envelope with flap atom C*n*. †† Δ From four-atom envelope plane. ‡‡ Relative energy in kJ mol^{-1} for DFT-optimized coordinates from starting models above (see *Comment*).

(Otwinowski & Minor, 1997); program(s) used to solve structure: *SHELXS97* (Sheldrick, 2008); program(s) used to refine structure: *SHELXL97* (Sheldrick, 2008); molecular graphics: *ORTEP-3* (Farrugia, 1997); software used to prepare material for publication: *SHELXL97* and *PLATON* (Spek, 2009).

We thank Drs J. Hanson and Z. Dauter of the Brookhaven National Laboratory, Long Island, New York, for their assistance, and the National Synchrotron Light Source for financial support.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: FG3191). Services for accessing these data are described at the back of the journal.

References

Becke, A. D. (1988). *Phys. Rev. A*, **38**, 3098–3100.
Bernstein, J., Davis, R. E., Shimoni, L. & Chang, N.-L. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1555–1573.
Evans, G. B., Furneaux, R. H., Schramm, V. L., Singh, V. & Tyler, P. C. (2004). *J. Med. Chem.* **47**, 3275–3281.
Farrugia, L. J. (1997). *J. Appl. Cryst.* **30**, 565.
Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. (1997). *J. Mol. Biol.* **267**, 727–748.
Lee, C., Yang, W. & Parr, R. G. (1988). *Phys. Rev. B*, **37**, 785–789.
Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
Nissink, J. W. M., Murray, C., Hartshorn, M., Verdonk, M. L., Cole, J. C. & Taylor, R. (2002). *Proteins*, **49**, 457–471.
Otwinowski, Z. & Minor, W. (1997). *Methods in Enzymology*, Vol. 276, *Macromolecular Crystallography*, Part A, edited by C. W. Carter Jr & R. M. Sweet, pp. 307–326. New York: Academic Press.
Scientific Computing & Modelling (2009). *ADF User's Guide*. Scientific Computing & Modelling NV, Vrije Universiteit, Theoretical Chemistry, Amsterdam, The Netherlands, http://www.scm.com.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Singh, V., Shi, W., Evans, G. B., Tyler, P. C., Furneaux, R. H., Almo, S. C. & Shramm, V. L. (2004). *Biochemistry*, **43**, 9–18.
Spek, A. L. (2009). *Acta Cryst.* D**65**, 148–155.
Van Lenthe, E. & Baerends, E. J. (2003). *J. Comput. Chem.* **24**, 1142–1156.
Vosko, S. H., Wilk, L. & Nusair, M. (1980). *Can. J. Phys.* **58**, 1200–1211.