Computational methods are used to predict the most favorable site of temozolomide towards attack by a water molecule. The energetics of the various complexes are presented as well as their geometries, including perturbations of each subunit caused by the presence of the other. Molecular electrostatic potential and Natural Bond Orbital (NBO) data are used to understand the interactions which conclude the terminal amide group is the preferred attack site where water can act as simultaneous proton donor and acceptor. Other potential proton acceptor N atoms within the aromatic ring structure represent weaker binding sites. Some of the less strongly bound structures include a CH···O H-bond.

Keywords: Hydrogen bonds, regioselectivity, CH--O H-bond, NBO
Introduction

The temozolomide (TMZ) molecule continues to attract the attention of numerous research groups. This molecule, of general formula C₆H₆N₆O₂,3,4-dihydro-3-methyl-4-oxoimidazo[5,1-d][1,2,3,5]tetrazine-8-carboxamide (Fig. 1), characterized by a fused pair of heterocyclic rings containing a total of five N atoms, as well as a carbonyl group and an amide group, has been reported to exhibit antitumor activity[1-3]. TMZ has found applications in medicine such as cytotoxic pro-drug for the treatment of astrocytoma, aggressive brain tumor, glioblastoma multiform, and melanoma [4]. In preclinical testing, TMZ developed by Stevens and co-workers has shown a broad spectrum of antineoplastic activity [4-6]. While its mechanism is largely unknown, one mechanism of action has been proposed [5], in which TMZ spontaneously hydrolyzes to methyltriazen-1-yl imidazole-4-carboxamide above pH 7 which then rapidly degrades to a highly reactive methyldiazonium ion. The nascent cation, an active methylating agent, is vulnerable to instantaneous nucleophilic attack by electron donors within DNA nucleotides, causing the transfer of a methyl group from the ion to form a methylated-DNA adduct.

Laboratory studies and clinical trials are investigating whether the anticancer potency of TMZ might be augmented by combining it with other pharmacologic agents. Some reported clinical trials have indicated that the addition of chloroquine is suspected to be beneficial for the treatment of gliomapatients [7-8]. Likewise, investigations on the effects of TMZ and quercetin on cell death in the human astrocytoma cell line MOGGCCM indicate that quercetin acts in concert with TMZ when used in combination rather than in separate pharmacological applications [9-10]. Along the same line, a theoretical study of the physicochemical (refractivity, polarizability, hydration energy, dipole moment, and Gibbs free energy of solvation) and geometrical properties of TMZ and four other anti-cancer drugs (methotrexate, carmustine, tamoxifen, hydroxifen) was performed using density functional theory to highlight their biological and anti-cancer activities [3]. Nevertheless, in spite of the biological importance of this molecule, as far we know, there has been no experimental nor theoretical work predicting the regioselectivity of the TMZ molecule in interaction with a proton donor. However, it is worth mentioning that the hydrogen bond (HB) propensity method of the Cambridge Structural Database (CSD) of TMZ shows an intermolecular N-H···O=C distance of 1.943 Å, suggesting a strong interaction between the carbonyl and amine groups [11]. Lowe et al. [12] reported a crystal structure of TMZ characterized by lattice parameters of 17.332, 7.351, and 13.247 Å, and an angle between the unit cell edges of 109.56°. The latter belongs to the P21/c space group. A doubly H-bonded dimer constitutes the asymmetric unit.
One carboxamide group forms an additional intermolecular NH···O HB; in both molecules the carboxamide group is coplanar with the heterocycle and its NH$_2$ group interacts with the imidazole nitrogen atom.

Because of the ubiquitous role of hydrogen bonding in many biological processes [13-23] and as TMZ is a polyfunctional heterocyclic base carrying two carbonyl groups and six nitrogen atoms as possible proton acceptor sites, a complementary way to probe the reactivity of this molecule is to analyze its HB propensity. Therefore, it appears useful to investigate in this first step the interaction between TMZ and water (as a prototypical proton donor) which is of course a major constituent of cells. All possible sites of attachment are examined, and a rank order of stability of various possible sites is obtained. Geometrical parameters of the complexes are calculated, and the results are analyzed via electrostatic potentials and Natural Bond Orbital (NBO) treatment [24-28].

**Computational Methods**

The equilibrium geometries of the complexes formed between water and TMZ were obtained using B3LYP/6-31+G(d,p), MP2/6-31+G(d,p), B3LYP/6-311+G** and B3LYP-D3/6-311+G** methods [23, 29-30]. After consideration of various different sites, seven water-TMZ conformers were obtained. Vibrational analysis showed each structure to be a true minimum.

The interaction energy of each complex was computed as the difference between the energy of the complex and the sum of the energies of the isolated TMZ and water monomers. These results were corrected for the basis set superposition error (BSSE) [31] using the Boys–Bernardi counterpoise correction [32].

The Molecular Electrostatic Potential (MEP), the interaction of a unit positive charge at a given point $\mathbf{r}'$ with reference to a molecular charge distribution, is evaluated by [33]:

$$
\nu(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}'
$$

where $\rho(\mathbf{r})$ is the molecular electron density function; the summation over A runs over all nuclei with charge $Z_A$ and distance $\mathbf{R}_A$. All calculations were performed using the Gaussian 09 software package [34-35]. Atomic charges and charge transfer energies were assessed by NBO as implemented in Gaussian 09. GaussView and Chemcraft programs were used for visualization [34].
Results and Discussion

The NBO charges on the atoms of the isolated TMZ molecule are shown in Figure 2, and the optimized geometries of the various water-TMZ complexes are illustrated in Figure 3. Details of the internal geometries of TMZ as an isolated molecule, and within each complex, are reported in Table S1 of the Supplementary Information.

Table 1 lists the interaction energy \( \Delta E \) of each of the seven complexes, along with the second-order NBO perturbation energies \( E(2) \) which are directly related to the individual HBs. With regard to the interaction energies, there are some consistent patterns evident in Table 1. The B3LYP-D3/6-311+G** method yields the strongest binding, 2-7 kJ/mol higher than MP2/6-31+G**, and B3LYP/6-31+G(d,p) the weakest. Even with a slightly larger basis set, B3LYP/6-311+G** predicts binding to be a bit weaker than MP2/6-31+G**. It is worth noting that Grimme et al. [36-38] have argued that these quantum chemical methods such as B3LYP-D3 based on London dispersion corrected density functional theory (DFT-D3) provide accurate interaction energies for large complexes. The typical DFT-D3 error for the basic gas-phase interaction energy has been estimated to be < 5% in several studies [39-41].

Most importantly, however, all four levels of theory indicate the same energetic ordering of the minima. The most strongly bound is complex W-TMZ1 in which the water molecule serves as both proton donor and acceptor, in a pair of OH···O and NH···O HBs. The structure binds the water to the TMZ with an energy of roughly 40 kJ/mol. Table 1 reports \( E(2) \) for the associated transfer of charge from the lone pair of the proton acceptor to the \( \sigma^* \) antibonding orbital of the donor to be 13.67 and 7.11 kcal/mol, respectively, which correspond to fairly strong HBs. This strength is verified by their shortness, with intermolecular distances of 1.875 and 2.064 Å, as displayed in Fig 3.

The next most stable complex is W-TMZ2 which resembles the global minimum in that the water engages in a cyclic structure. In this case, the water donates a proton to N of the heterocyclic ring, rather than a carbonyl O, while again accepting a proton from the terminal amino group. Despite these similarities, W-TMZ2 is less stable than W-TMZ1 by roughly 10 kJ/mol. The NBO parameters by themselves do not support this energetic ordering, as the NH···O HB has a larger \( E(2) \) in W-TMZ2 than in W-TMZ1. Both of these two HBs are slightly shorter than 2 Å. Another factor that contributes to this observation will be discussed below.

W-TMZ3 and W-TMZ4 are only slightly less stable than W-TMZ2; both are cyclic. Like W-TMZ1, W-TMZ3 also contains a HB to the carbonyl O17, but its second HB also involves
proton donation from the water to N2, leading to negative cooperativity. The second of these two HBs is significantly longer, at 2.408 Å, agreeing with the much smaller value of E(2) in Table 1. W-TMZ4 again donates a proton to a carbonyl O (this time O8), but its second HB is of the weaker CH···O variety, with R(H···O)=2.166 Å. These sorts of HBs have been extensively studied over the last few years [42-45] and are more common than earlier believed. CH···O HBs are also involved in W-TMZ5 and W-TMZ6. These CH···O HBs are weaker than in W-TMZ4, as judged by both a smaller E(2) and 0.2 Å longer intermolecular distance. This weakening is a natural consequence of the better proton-donating ability of the heteroaromatic ring vs a methyl group. The only minimum with a single HB is W-TMZ7 where charge is donated by the lone pair of N18 of the amino group. It may be noted from Table 1 that W-TMZ3 and the succeeding structures have much smaller values of E(2) than do the two most stable dimers, a prime factor in their lesser stability.

The formation of a AH···B HB will typically stretch the AH covalent bond, and the amount of this strength scales as the HB strength. The principal exception is a CH donor, for which the direction of bond length change, whether stretch or contraction, is not easily predictable, as there are a number of blue-shifting CH···O HBs in the literature where the CH bond is shortened. When the proton acceptor is a carbonyl group, the C=O bond will usually stretch. The changes undergone by the relevant bond lengths in the various complexes are compiled in Table 2. As may be seen in the first column, the NH bonds of TMZ stretch by 0.008 Å in W-TMZ1 and W-TMZ2 where they engage in NH···O HBs with the water. Consistent with their weaker nature, the CH bonds of complexes W-TMZ4-6 stretch by a smaller amount. Note also that the aromatic H of W-TMZ4 stretches by more than do the methyl C-H bonds of W-TMZ5 and W-TMZ6. The pertinent OH bond of water elongates in all cases, especially in the two most stable complexes, where the stretch exceeds 0.01 Å. It is worth pointing out that the C=O bonds stretch in all cases when they act as proton acceptors, by amounts between 0.003 and 0.012 Å.

Comparison of the stabilities of the various complexes leads to several observations. The amino group is a potent proton donor, forming a corresponding HB in the two most stable minima. As a secondary HB within the available cyclic structures carbonyl O17 acts as a stronger proton acceptor than does the N12 of the heterocyclic ring. The NBO data of Table 1 provide information about the charge transfer aspects of these HBs. But an equally important contributor arises from the electrostatic interactions. One way of considering these Coulombic forces may be gleaned by examination of the natural atomic charges on the various atoms. As may be seen in Fig 2, the two H atoms of the terminal amino group are the
most positively charged H atoms in the molecule, supporting the idea that this group is a strong proton donor. The carbonyl O17 has a high negative charge, more so than N12 of the ring, again consistent with preference of the HOH to donate a proton to the former atom. It may be noted as well that the CH protons bear a significant positive charge of roughly 0.25, aiding them in forming CH·O HBs, albeit weaker than NH·O. The other carbonyl O8 is also quite negative and capable of forming a strong OH·O HB, but the only proton donors in its vicinity are the weaker CH groups, weakening any such cyclic complexes.

Of course, atomic charges are an arbitrary quantity, sensitive to the means used to calculate them. In contrast, the molecular electrostatic potential (MEP) surrounding the TMZ molecule offers a non-arbitrary and more rigorous picture of what an approaching water molecule might see. The MEP has been particularly useful as an indicator of the sites or regions of molecules to which an approaching electrophile is attracted [46], and it has also been applied successfully to the study of interactions that involve a certain optimum relative orientation of the reactants, such as a drug and its cellular receptor [47-48]. This potential is displayed in Fig 4 where blue and red regions indicate, respectively, positive and negative regions of the potential. This diagram buttresses the ideas based on the atomic charges above. The terminal NH₂ group offers a positive region attractive to the O of a water molecule, and negative area surrounding the carbonyl O17 is highly favorable for the water H atom. Note that this red region is far more extensive than the small negative area near to N12. It is also worth pointing out that the red region that begins near O17 extends to the right and encompasses the lower region of the heteroatomic ring, including N1 and N2. It is to this area that the second H atom of water is attracted in W-TMZ3 or in W-TMZ6.

It is here that the MEP and atomic charge pictures part ways. The red negative region that extends to N1 and N2 are at odds with the charges on these atoms of +0.01 and -0.16. Another potentially misleading characteristic of the atomic charges are N atoms N9 and N12. Although they are assigned substantially negative charges of about -0.45, the MEP is positive near the former, and only barely negative around the latter. With regard to the region above the plane of the TMZ molecule, there are no negative regions lying above any of the ring atoms, either N or C, which helps to explain why there are no minima on the potential energy surface where the water molecule is located out of the TMZ plane. Another issue which adds to this absence of out-of-plane water positions is that the p-orbitals of the N atoms that participate in the aromatic rings are delocalized into the π cloud and hence unavailable to engage in a HB.
**Conclusions**

Water prefers to bind to TMZ in such a way that it can act as both proton donor and acceptor simultaneously. The most potent proton donor site on TMZ is the terminal amino group to which the water preferentially binds. The water can then donate a proton to the neighboring carbonyl O, which is preferred over the N atom of the heteroatomic ring by some 8-9 kJ/mol. The water is bound to TMZ by some 40 kJ/mol in this global minimum. There are other, weaker, binding sites identified, most of which involve a cyclic structure, with water simultaneously acting as both proton donor and acceptor. Some of these minima involve a CH-O H-bond. The least stable minimum is characterized by a single H-bond between the OH of water and a lone pair of the amino N atom. With specific regard to N proton acceptors, the N12 atom of the five-membered ring is the strongest proton acceptor. The observed trends are understandable and predictable based upon the charge distribution within the TMZ molecule.

**Acknowledgments**

O.E.K. would like to thank Snow College and in particular President Scott Wyatt and Dean Dan Black for their support for a study visit in Snow College, Ephraim, Utah, USA wherea version of this paper was completed. Computer, storage, and other resources from the Division of Research Computing in the Office of Research and Graduate Studies at Utah State University are gratefully acknowledged. This work was supported by NSF-CHE-1026826.
References


[48] R.J. Xavier, P. Dinesh, Spectroscopic (FTIR, FT-Raman, $^{13}$C and $^{1}$H NMR) investigation, molecular electrostatic potential, polarizability and first-order hyperpolarizability, FMO and NBO analysis of 1-methyl-2-imidazolethiol, Spectrochimica acta A 118 (2014) 999-1011.