Investigating the cyclization of enediyne analogs using density functional theory

Presented by Dylan Jay Estep

In partial fulfillment of the requirements for graduation with the Dean's Scholars Honors Degree in Biochemistry

_____ _____

Sean M. Kerwin, Ph.D. Date

Supervising Professor

_____ _____

John F. Stanton, Ph.D. Date

Honors Advisor in Chemistry and Biochemistry

Investigating the cyclization of enediyne analogs using density functional theory

Presented by Dylan Jay Estep[1] in partial fulfillment of the requirements for graduation with the Dean's Scholars Honors Degree in Biochemistry

Supervised by Sean M. Kerwin, Ph.D[2].

[1]Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, TX 78712
[2]College of Pharmacy, The University of Texas at Austin, Austin, TX 78712

Abstract

Enediynes are organic molecules that readily undergo a thermal rearrangement, now commonly known as the Bergman cyclization, to a cyclic *para* diradical form. Interest in this rearrangement was renewed when it was found to be crucial to the mechanism of cytotoxicity in a variety of natural products containing the enediyne structural moiety. Cyclization of these molecules leads to DNA strand scission and ultimately cell death. Recent efforts by medicinal chemists to discover therapeutically relevant enediyne derivatives have been complemented by computational approaches, which seek to compute energies and energetic barriers to cyclization that can accurately predict the behavior of these molecules *in vivo*. Here we demonstrate this approach for *cis*-hex-3-ene-1,5-diyne and two of its analogs using density functional theory, discuss the validity of its predictions, and investigate the effect of basis set on the description of these molecules' reactivity.

1. Introduction

1.1 The Bergman cyclization

Research into enediyne molecules emerged from chemists' persistent efforts to understand a chemical intermediate known as didehydrobenzene. For decades more commonly known in the literature as benzyne, the elusive molecular species, with molecular formula C_6H_4, is generated by the removal of two hydrogen atoms from benzene, C_6H_6 (Figure 1). Due to the symmetry of the parent benzene molecule, didehydrobenzene may have three isomers: the *ortho* isomer, in which the removed hydrogen atoms are bonded to adjacent carbon atoms on the ring (Figure 2a); the *meta* isomer, in which they are separated by one intervening carbon atom (Figure 2b); and the *para* isomer, in which they are separated by two intervening carbon atoms, located directly across the six-membered ring from one another (Figure 2c). These isomers have held the interest of many

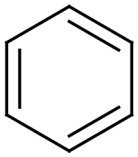


Figure 1. One of organic chemistry's classic molecular structures: benzene.

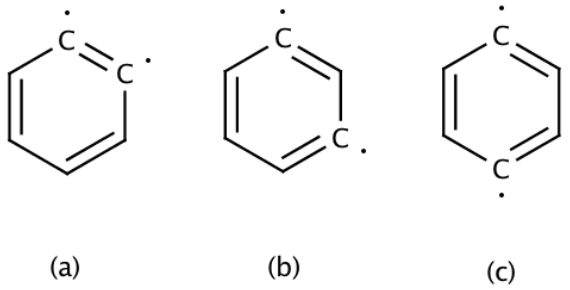


Figure 2. The three isomers of didehydrobenzene (as diradical structures): (a) *ortho*; (b) *meta*; (c) *para*

organic and physical chemists for the greater part of the 20th century; understanding the molecules and even determining whether they exist was of great interest to scientists interested in the mechanisms of organic reactions, synthetic pathways, and even bonding theory.

The very existence of *any* of the didehydrobenzene isomers was first postulated in 1927, when Bachmann and Clarke found the reaction of chlorobenzene and sodium to yield, in addition to expected products, some species that they reasoned could only be generated by a free radical mechanism. Both mono- and diradical species were proposed in their paper, with the diradical species taking the form of the *ortho* isomer.[1] The hypothesis was further advanced by Georg Wittig, who in 1942 predicted the intermediate when he observed the rates of reaction of phenyllithium with various halobenzenes to produce biphenyl, which ran contrary to what would have been expected had the reaction proceeded under a different mechanism.[2] Due to an observation of the reactivity of two fluoroanisole isomers from a previous study he had performed,[3] he predicted the intermediate took the form of an asymmetric zwitterion (Figure 3a), explicitly rejecting symmetric intermediates like the triple-bonded aryne form (Figure 3b); only in 1955 were these observations found to be relics of experimental error upon replication.[4] By this time, the structure of the intermediate had been more definitively resolved by John D. Roberts's classic study, in which chlorobenzene synthesized with one ^{14}C atom bonded to the chlorine atom was reacted with potassium amide. The observed 1:1 product ratio of aniline-1-^{14}C and aniline-2-^{14}C clearly demonstrated the intermediate didehydrobenzene was electrically neutral and symmetrical.[5] The gradual adoption of resonance theory and molecular orbital theory in organic chemistry during this time led contemporary chemists to consider the *ortho* didehydrobenzene as a hybrid of diradical, aryne, and even to some extent zwitterionic resonance contributors.

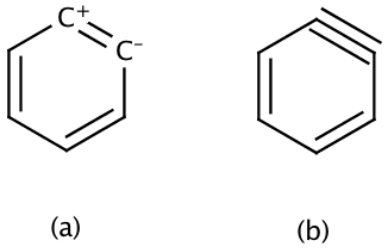


Figure 3. Two proposed structural forms of the *ortho* didehydrobenzene: (a) an asymmetric, zwitterionic form; (b) an electrically neutral and symmetric aryne form.

The *meta* and *para* isomers remained largely a mystery at this time, though some tenacious scientists continued to pursue them. Physical organic chemist Robert Bergman surmised that the *para* isomer might be generated transiently by provoking a thermal rearrangement of a molecule called *cis*-hex-3-ene-1,5-diyne ("enediyne,"

Figure 4a), whose synthesis and isolation had been published throughout the 1960s.[6,7] Due to the symmetry of the molecule and the likely rapidness of the conversion, Bergman modified the synthesis to incorporate deuterium (^{2}H) at both its acetylenic positions so that upon rearrangement, it would equilibrate with the structural isomer with deuterium at both its vinyl positions (Figure 4b). Analysis of the mixture produced after heating the compound showed a 1:1 ratio of the two structural isomers he expected with no other patterns of deuterium distribution detected, convincing Bergman that the rearrangement indeed proceeded through a species with the same symmetry as the suspected *para* didehydrobenzene.[8] However, this information didn't clarify the structure of the species – and the structural "contenders" for this didehydrobenzene isomer were even more numerous than there were for the *ortho* isomer – nor did it clarify whether it was accessed as an intermediate or merely a transition state.

To resolve the first issue, the enediyne was heated in a number of different reagents that might intercept the intermediate species; for example, heating it in a particular hydrocarbon resulted in the rapid generation of benzene and loss of starting material. The results put forth strong evidence for the diradical structure of the *para* didehydrobenzene, as radical molecules were virtually the only ones known to remove hydrogen atoms from hydrocarbons. Estimates of heat of formation gleaned by group equivalent techniques suggested that the didehydrobenzene was approximately only 14 kcal mol^{-1} higher in energy than the enediyne. Bergman noted that, if the species were a transition state, the rearrangement would proceed at room temperature if the

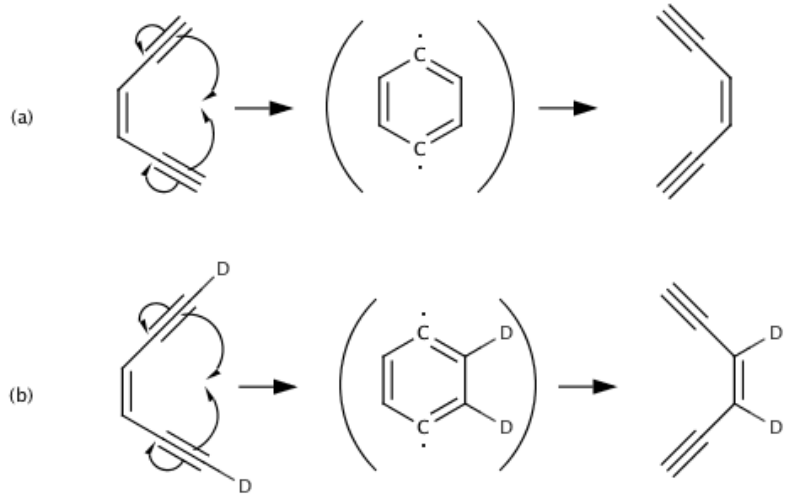


Figure 4. The Bergman cyclization. (a) According to Bergman's proposed mechanism, the enediyne molecule would rearrange upon heating to transiently form a *para* didehydrobenzene. Note that the enediyne structures on the left and right are equivalent. (b) Incorporating deuterium, indicated here as D, at the acetylenic positions of the enediyne allowed Bergman to test whether this mechanism was reasonable; note that by making this modification, the structures on the left and right are now distinct.

energetic barrier were only this high; because this did not occur, he concluded that the *para* didehydrobenzene was not a transition state, but an unstable intermediate. Impressively, he also guessed, in the absence of proper kinetic studies, that the transition state leading to the intermediate would be about 32 kcal mol^{-1} higher in energy than the enediyne, a prediction that turned out to be somewhat accurate.[9] Bergman's studies on the rearrangement leading from the enediyne molecule to the transient *para* didehydrobenzene species and back again were published in 1972; that rearrangement is now widely known as the Bergman cyclization.

1.2 Discovery of enediyne antibiotics and their mechanism of cytotoxicity

It took us until the 20th century to find enediynes and the *para* didehydrobenzene, but for all our brilliance, it turned out that some other organisms – bacteria, in fact – had already beaten us to the punch. The surprising biological relevance of enediynes was discovered in the midst of a program that sought to discover powerful natural products. In the early 1980s, a semi-automated colorimetric assay involving genetically engineered *E. coli* was developed to detect DNA damage in hopes of finding novel cytotoxic molecules. Of the 10724 fermentation broths tested, about 1% demonstrated significant DNA-damaging activity.[10] Some of the fermentation broths that tested positive came from the bacterium *Micromonospora echinospora* ssp *calichensis*, named after the caliche (chalky soil) sample in which it was found in Texas. From these broths, the calicheamicin family of molecules was isolated (Figure 5). The potency of these molecules was remarkable; in mouse models, they were approximately 1000 times as powerful as adriamycin, a cancer drug already in use at the time.

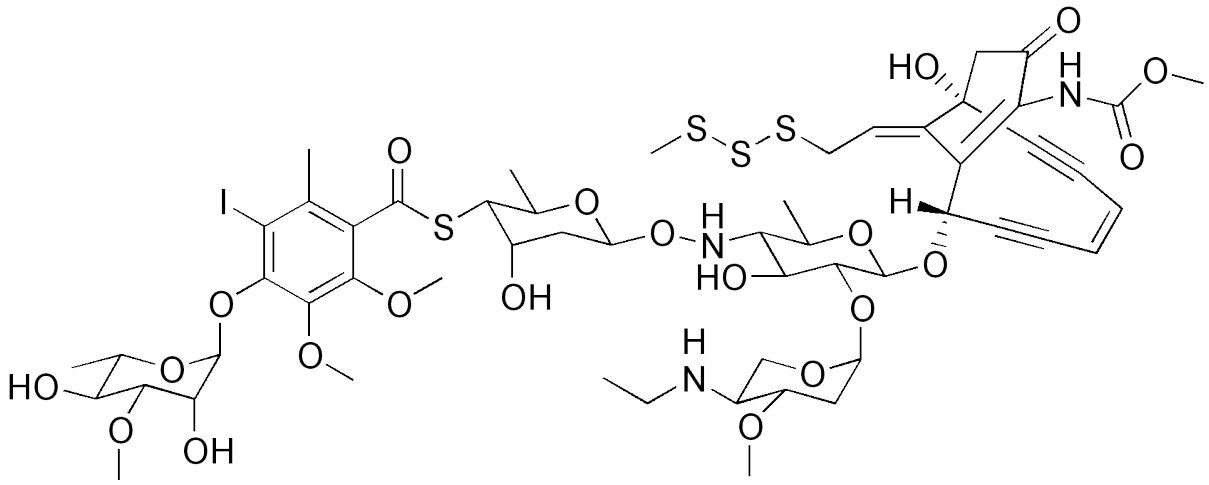


Figure 5. Calicheamicin γ'$_1$, one of the natural enediyne antibiotics discovered in the 1980s.

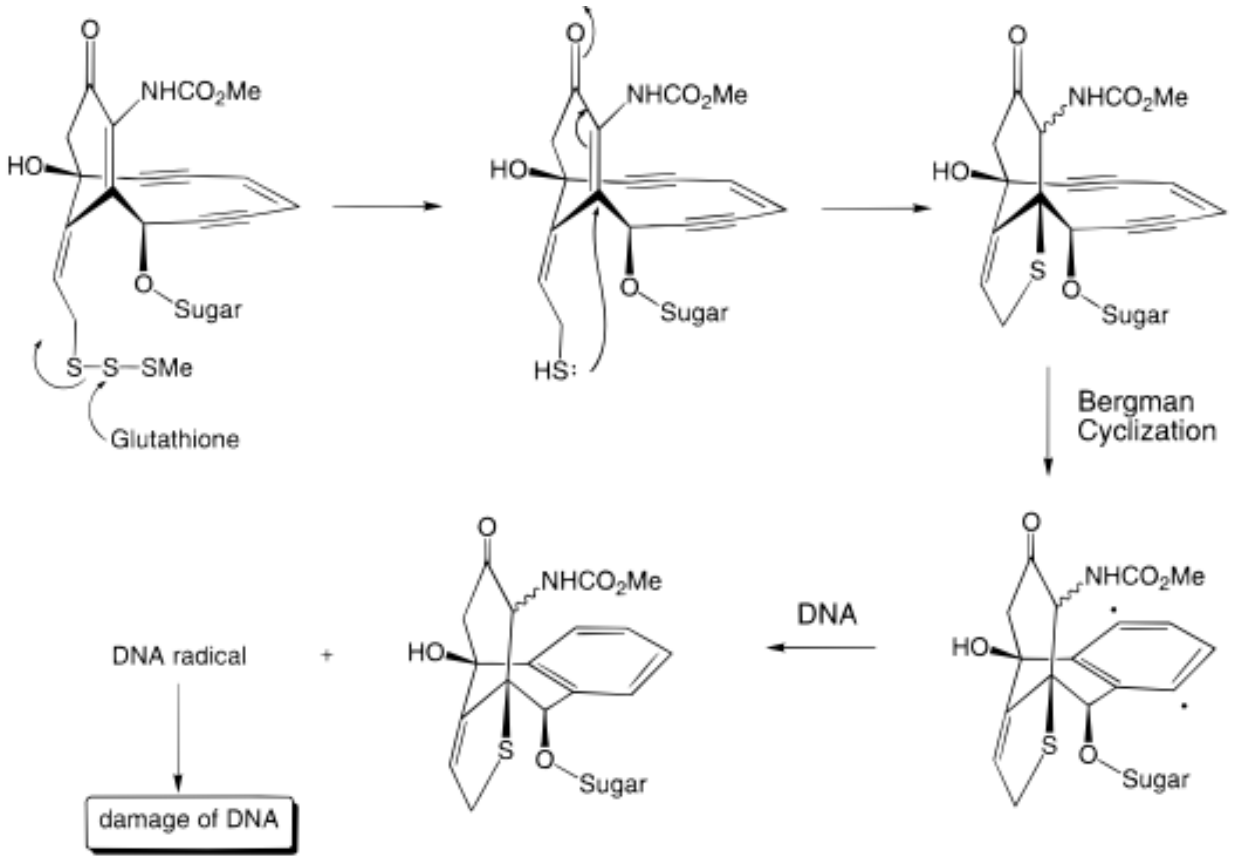


Figure 6. Mechanism by which calicheamicin γ′$_1$ effects DNA cleavage. An external nucleophile such as glutathione activates a "trigger" group, causing structural changes that promote the Bergman cyclization of the enediyne moiety. The diradical that is generated in this process then can cleave DNA. Adapted from E. Kraka, D. Cremer, *J. Am. Chem. Soc.* **2000,** *122,* 8245.

The structures of the calicheamicins, determined by extensive spectroscopic investigation and published in their final revised form in 1989 (after a preliminary 1987 publication), revealed an interesting surprise: the presence of the enediyne moiety embedded in a ten-membered ring.[11] Moreover, evidence emerging in the subsequent years suggested a truly awe-inspiring mechanism of action (figure 6): first, the calicheamicin molecule binds to the minor groove of double helical DNA in a sequence-specific manner, with proper recognition mediated by its oligosaccharide tail;[12] next, attack of the central sulfur atom in the trisulfide group occurs by an external nucleophile, prompting an internal Michael addition at an unsaturated carbon atom, contracting the ring structure and thus triggering Bergman cyclization of the enediyne moiety.[13] Isotope labeling studies have revealed the mechanisms by which the generated diradical species cleaves DNA: the pathways are numerous and varied, occuring at 1', 4' and 5' positions of the deoxyribose sugar.[14,15] In general, abstraction of a hydrogen atom from the deoxyribose sugar at these positions leads to the formation of unstable species, often causing double-stranded cleavage.

That biosynthetic pathways had produced enediyne-containing molecules, presumably as a defense for the organism, seemed remarkable. It wasn't unique to *Micromonospora*'s molecules, either: enediynes kept emerging in other solved molecular structures around the same time, including esperamicin,[16] dynemicin,[17] and the neocarzinostatin chromophore,[18] whose structure was actually solved first but received little attention until its similarity to the calicheamicins was noted.[19] These molecules and others like them are now collectively known as the enediyne antibiotics.

1.3 Design of enediyne analogs for use in cancer therapy

The discovery of nature's ingenious and formidable use of the enediyne structural moiety has inspired chemists to design new enediyne derivatives with tailored properties. In general, slight modifications to the enediyne core produce molecules with very similar chemistry. For instance, take the 3-aza-enediyne ("aza-enediyne," Figure 7), which is produced by the replacement of one carbon atom in the enediyne by a nitrogen atom. This molecule may also undergo a Bergman-type cyclization to yield a diradical species; however, its retro-Bergman cyclization, the rearrangement in which the ring opens at the side opposite to that which initially "closed," yields a distinct product from the starting material, unlike the enediyne (Figure 1a). While the rearrangements are expected to proceed by equivalent mechanisms to the rearrangements that the enediyne undergoes, the energetic barriers to the rearrangements are anticipated to be different, thus leading to a different distribution of products when the aza-enediyne is heated.

Incorporating a ring structure behind the double-bonded portion of the structural moiety gives rise to even more interesting chemistry. Consider the 1,2-dialkynylimidazole (Figure 8): retro-Bergman cyclization leads to a curious-looking nine-membered cumulene ring structure, and further rearrangement may lead to a carbene. Between each of these rearrangements exists a transition state, and thus an energetic barrier. The different energies of the various forms of the molecule and transition states dictates, again, which forms exist when the initial starting material is heated at a sufficient temperature.

The form of enediyne-based molecules known to exert its DNA-damaging and cytotoxic activity is the diradical form. A therapeutic enediyne derivative, then, must be able to access this

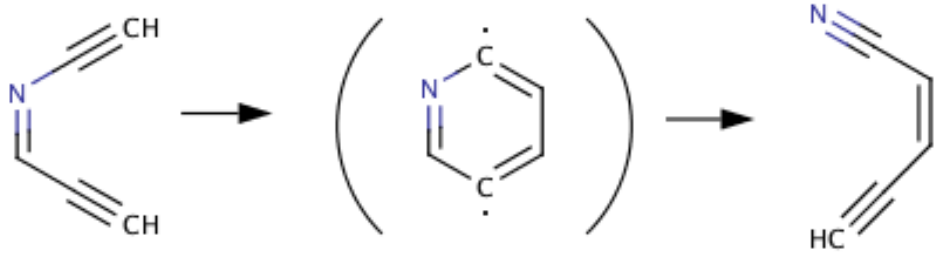


Figure 7. Bergman cyclization of aza-enediyne, proceeding through a diradical intermediate and producing an enynenitrile.

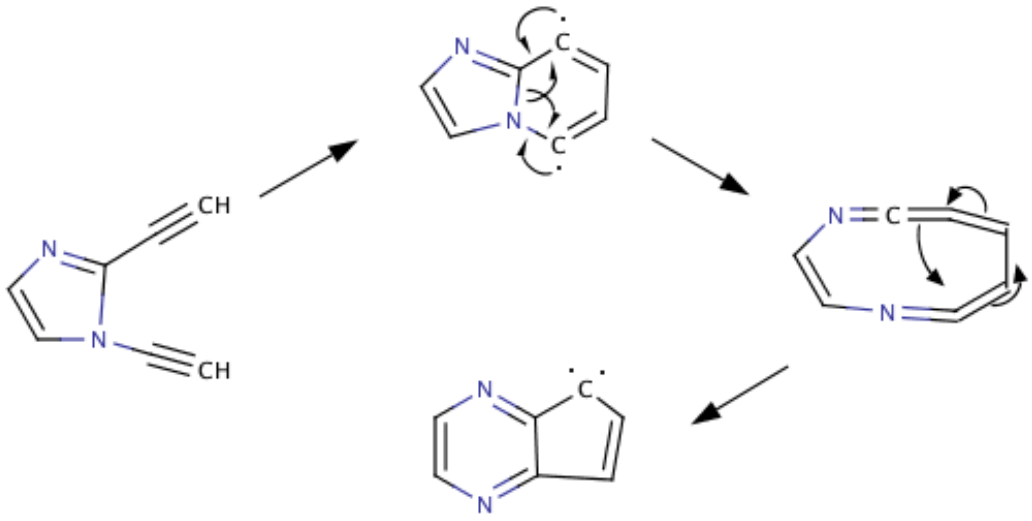


Figure 8. Rearrangement of 1,2-dialkynylimidazole. This species may undergo a Bergman cyclization; the unusual cumulene ring structure produced by the retro-Bergman rearrangement may further rearrange, giving rise to a carbene species.

form in order to damage the DNA of a disease-causing offender and eliminate it. But of course, these offenders reside within the human body, whose own DNA would be susceptible to such attack. How can this dilemma be resolved?

Assuming the DNA of both target and normal cells is accessible to an administered therapeutic enediyne derivative, the answer is simply that the two types of cells must differ sufficiently in environment such that the molecule assumes a harmful form in the target cell but a benign form in normal cells. This difference in conditions would represent a vulnerability to the molecule that the target cells have but that healthy human cells do not. Such vulnerabilities are often exploited in the molecular eradication of disease; for instance, an antibiotic may target a particular protein structure present in the bacterial form of an essential enzyme but not present in the human form, thus efficiently killing bacteria but leaving human cells unharmed.

Cancers are notoriously difficult diseases to treat because the differences in vulnerabilities between tumor and normal cells are often not very pronounced. Since tumor cells arise from many minor mutations to normal cells, they often have very similar vulnerabilities to the cells from which they came; a molecule effective in killing a tumor cell, then, often also does extensive collateral damage to normal cells. Nonetheless, research has revealed some consistent differences between normal and tumor cells that are being pursued in hopes of making cancer treatment more targeted and selective. Among these is the significant difference in pH of many tumor cells relative to their normal counterparts. Typical

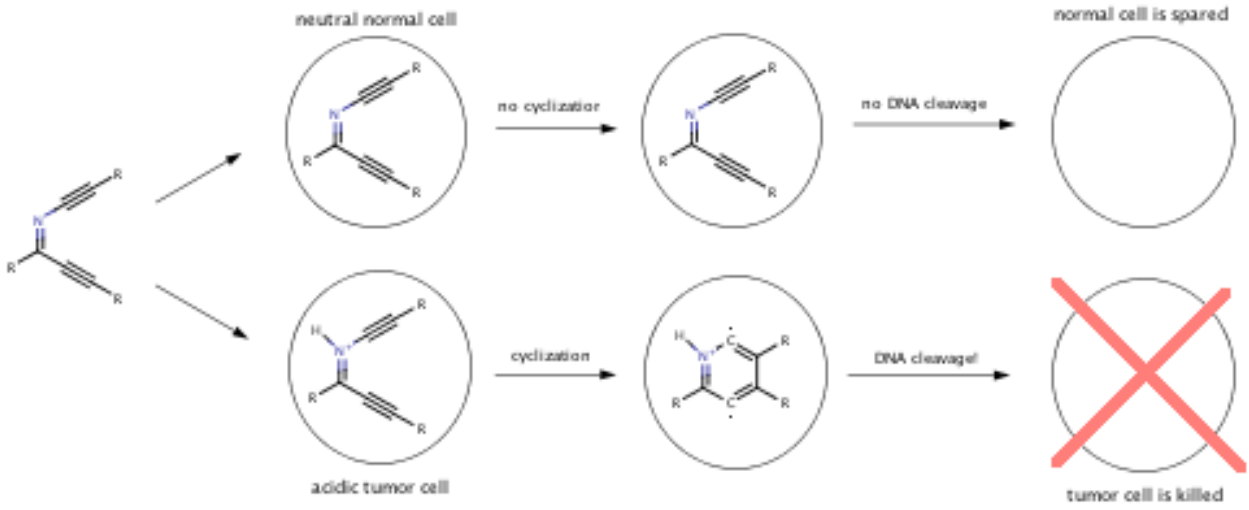


Figure 9. An example of a hypothetical, ideal enediyne derivative with selective antitumor activity mediated by the pH difference between normal and tumor cells. Under normal conditions, the molecule does not cyclize, thus sparing normal cells from DNA damage. The protonation of the molecule in the tumor cell's acidic environment promotes its cyclization to the diradical form, causing death of the tumor cell. Note that under normal conditions, cyclization is permissible as long as the molecule rapidly converts to any form other than the toxic diradical.

physiological pH is around 7.5; tumor cells, by contrast, are often found to be significantly more acidic, with pH 6.2 to 6.6.[20] This difference likely arises because the vasculature of tumors cannot supply enough oxygen and other nutrients for the entire growing population of cells, leading many tumor cells to produce excess lactic acid under anaerobic conditions. Even more acidic conditions can be promoted in tumors by treatment with ionophores, bringing pH as low as 5.5.[21]

Thus, with this difference of conditions in mind, the goal of the rational design of an enediyne-based antitumor molecule becomes clear: design one molecule that will take the form of the DNA-damaging diradical under acidic conditions but that will take any other form under normal physiological conditions (Figure 9). This is a taller order than it may appear; satisfaction of the following criteria devised by Kraka and Cremer would produce a highly desirable molecule:[22]

(1) the form of the molecule that it assumes under acidic conditions must have an energetic barrier to cyclization of less than 24 kcal mol^{-1} in order to permit formation of the diradical at physiological temperature
(2) all energetic barriers leading away from the diradical under acidic conditions should be much larger than the barrier to abstraction of hydrogen from DNA, or about 14 kcal mol^{-1}
(3) the diradical under acidic conditions should have high hydrogen-abstraction ability, as predicted by a small energetic singlet-triplet gap

(4) criteria 2 and 3 should not be satisfied by the molecule under neutral pH (normal physiological conditions); energies and energetic barriers must either prevent formation of a diradical altogether or only permit transient formation of a diradical before rapid conversion to a stabler, non-toxic form

While this rationally based approach to designing enediyne therapies with as limited toxicity as possible has emerged recently, other approaches have been undertaken as well. Success has been limited. Neocarzinostatin, one of the natural enediyne antibiotics mentioned above, has been approved in Japan for treatment of cancers of digestive organs.[23] Mylotarg, a monoclonal antibody linked to a calicheamicin-family molecule, was approved for treatment of certain acute myelogenous leukemia patients under the Food and Drug Administration's accelerated approval program, which seeks to bring potential therapies for particularly life-threatening diseases to the market more quickly, then follow up with more clinical trials to validate the approval. Mylotarg did not hold up to this subsequent scrutiny; follow-up trials had to be halted early when no benefit was detected and higher death rate was observed in the experimental group, and Mylotarg was therefore pulled from the market by the FDA. Both neocarzinostatin and Mylotarg are highly toxic.

1.4 Density functional theory

One of the main goals of computational chemistry is to predict the structure and properties of a molecule or material of interest. Acquiring information in this manner is particularly appealing when it is unfeasible, costly, or even impossible to obtain it directly through experiment. Given the elusive nature of didehydrobenzene molecules and their derivatives experimentally, it is hardly a surprise that these molecular species have been studied using theoretical methods. By the end of the 1990s, several high-level computational studies had been completed which accurately predicted the three didehydrobenzene isomers' relative stabilities and singlet ground states.[24]

Electronic structure methods typically compute the energy of a particular configuration of atoms in space – a molecular geometry – based on the principles of quantum mechanics. This is accomplished by solving the famous Schrödinger equation $\mathcal{H}\Psi = E\Psi$ to determine the wavefunction Ψ, from which various properties including the ground-state energy can be obtained. Even when nuclei are treated as stationary, for all but the very simplest atoms and molecules, this is a many-body problem requiring the simultaneous solution of so many differential equations that it is not feasible for modern computers. A large variety of methods exist in which different approximations to solution are employed to overcome this problem. Some, called semi-empirical methods, use parameters with values obtained from experiment in order to simplify this process and save computational cost, but many others, called *ab initio* methods, only use quantum mechanical principles in addition to some fundamental physical constants. The most well-known *ab*

initio method, the Hartree-Fock method, constructs a many-electron wavefunction using single-electron wavefunctions in a manner that fulfills the Pauli exclusion principle, which an earlier method devised by Hartree failed to take into account. The Hartree-Fock method is based upon the variational theorem, so energies it obtains must be higher than the true ground-state energy of the molecule of interest. However, a fundamental limit to how close Hartree-Fock calculations can get to the true energy exists due to the method's neglect of a phenomenon called electron correlation. This shortcoming makes energies obtained by Hartree-Fock decent approximations, but also ones that ultimately fail to make valid predictions for many systems.

Many methods, generally called post-Hartree-Fock methods, have been devised to compensate for Hartree-Fock theory's neglect of correlation. However, these are generally extremely computationally costly. A new method emerged in the 1960s called density functional theory (DFT), which greatly reduced computational cost and could in principle address both exchange and correlation effects. It arose from two important theorems proved by Hohenberg and Kohn: first, that the energy of a system was a function of a function (functional) of the electron density; and second, that the electron density which minimizes the energy is the true electron density.[25] DFT methods proceed by an iterative procedure whose goal is to solve self-consistent equations called Kohn-Sham equations;[26] for instance, an algorithm may begin by defining a trial electron density function from which the Kohn-Sham equations can be solved, then using the single-electron wavefunctions obtained to compute the electron density and check for convergence and proceed to a new iteration with an updated function if convergence is not achieved. The reduced computational cost of DFT arises from the significant reduction in the number of spatial variables in the equations by defining the electron density.

Computing the energy of a molecular geometry using electronic structure methods like Hartree-Fock and DFT can be complemented with methods of minimization so as to locate the geometry with the lowest energy (within the proximity of an initially provided geometry). Some of the more well known of these methods include steepest-descent and the Newton-Raphson method. Geometries optimized in this manner are taken to be the geometry that the molecule would assume in reality with the corresponding computed energy. Modified versions of these methods may also be used to find local energy maxima rather than minima, allowing for transition state geometries and energies to be computed. Comparing how much higher a transition state's energy is relative to the initial reactant molecule(s) allows for prediction of energetic barriers for the reaction.

2. Materials and Methods

2.1 Optimization of structural geometries and absolute energy determination

The computational chemistry packages GAMESS[27] and Gaussian 09[28] were used to determine the structures of the

enediyne and two of its derivatives, aza-enediyne and 1,2-dialkynylimidazole, as well as the products of their rearrangements (see section 1.3), using DFT. Because geometry optimization methods attempt to locate local minima in the neighborhood of an initially provided geometry, a roughly optimized geometry was first generated by building the molecules in a molecular editing program and subjecting them to energy minimization methods based upon molecular mechanics force fields. These geometries were then used as the initial inputs for the DFT optimizations.

The optimization algorithms terminate based on whether the computed gradient of the energy is within some very small value (effectively zero), indicating that a stationary point on the potential energy surface has been reached. To further clarify that the correct type of stationary point has been reached, namely a local minimum, geometry optimizations were followed by frequency calculations, which compute the eigenvalues to the matrix of second partial derivatives of energy. The presence of only real frequencies validated that the algorithm successfully located a local minimum.

For some optimizations – namely, the entire aza-enediyne series as well as all 1,2-dialkynylimidazole series species under acidic conditions besides the diradical – Cs molecular symmetry was enforced within the calculation by Gaussian 09. All others, however, including those obtained by GAMESS, employed no symmetry constraints.

In the cases of aza-enediyne and 1,2-dialkynylimidazole, the effect of the choice of functional on the description of the molecules' reactivity was examined by computing geometries using both the popular B3LYP and BPW91 functionals. In addition, for the aza-enediyne, the effect of basis set on the description was examined by performing computations using several Pople and correlation-consistent basis sets.

2.2 Locating transition state geometries and computing energetic barriers

The procedure employed to obtain transition states was a bit more involved than for other species. For each transition state, the molecule's geometry was optimized with the length of the bond formed or broken during the rearrangement fixed at several different values, first by molecular mechanics and then by DFT. The geometry that produced the highest energy was taken to be the one "closest" to an actual transition state (in some cases verified by frequency calculations) and was next subjected to an appropriate algorithm to further optimize the structure. The presence of one imaginary frequency following a frequency calculation was used to verify that a saddle point of the potential energy surface was reached, corresponding to a transition state.

2.3 Description of diradical species

The theoretical description of diradical species, including transition states with high diradical character, has long been acknowledged as a formidable challenge due to the multiconfigurational nature of these species. Calculations of diradical intermediates and transition states with high

	1	TS1-2	2S	2T	TS2-3	3
3-21G	-245.494	-245.458 22.8	-245.496 -1.3	-245.478 9.8	–	-245.559 -41.1
6-31G(d)	-246.857	-246.822 21.8	-246.863 -4.0	-246.849 4.6	-246.852 6.6	-246.916 -37.2
6-31G(d,p)	-246.857	-246.822 21.9	-246.868 -6.9	-246.854 1.7	-246.858 -0.5	-246.921 -40.3
cc-pVDZ	-246.870	-246.836 21.5	-246.881 -6.9	-246.992 2.1	-246.871 -0.8	-246.933 -39.8
cc-pVTZ	-246.949	–	-246.950 -0.8	-246.935 8.4	-246.942 3.8	-247.010 -38.8

Table 1. Calculated energies (hartree) of the aza-enediyne series using the B3LYP functional and the indicated basis sets. 1 represents the starting material, 2 the diradical, and 3 the product resulting from retro-Bergman cyclization. All energies were obtained by optimizations in which Cs symmetry was enforced. Failure to optimize the structure and compute the corresponding energy is denoted by an emdash (–). For all species but **1**, a relative energy compared to **1** is also provided in kcal mol^{-1}. Calculations were performed using Gaussian 09.

diradical character employed an unrestricted broken spin formalism. Although contamination of higher spin states can removed by a sum-correction method using the single-point triplet state energy at the broken-spin singlet geometry of the transition state,[29] this method was not employed in this work. The energies of diradical species in both singlet and triplet states will be presented, and transition state energies from singlet state calculations will be presented.

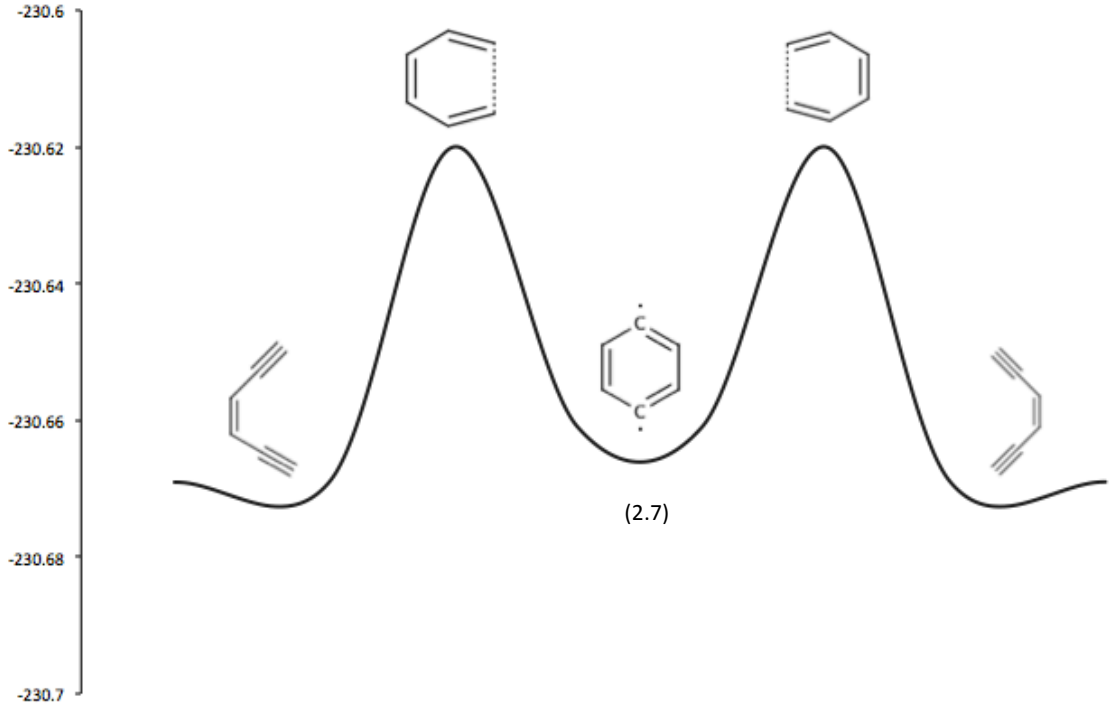
3. Results

The aza-enediyne series was selected as the one on which the effect of basis set on the description of reactivity would be tested. All calculations employed the widely used B3LYP functional; energies were computed as the sum of the electronic energy and a zero-point correction calculated from the vibrational frequency analysis. With two exceptions, all optimized structures and their corresponding energies were successfully obtained. The results of this investigation are given in Table 1.

Next, the enediyne and 1,2-dialkynylimidazole series were similarly characterized using the B3LYP functional and the Pople 6-31G(d,p) basis set, in addition to these series under acidic conditions. All optimized structures and their corresponding energies were successfully determined except for transition states under acidic conditions. Data are given in Table 2 while schematic diagrams representing the data for the series under neutral conditions are given in Figures 10, 11, and 12.

		1	TS1-2	2S	2T				
enediyne		-230.669	-230.620 30.7	-230.661 5.4	-230.656 8.1				
		1	**TS1-2**	**2S**	**2T**	**TS2-3**	**3**		
aza-enediyne	*neutral*	-246.857*	-246.822* 21.9	-246.868* -6.9	-246.854* 1.7	-246.858* -0.5	-246.921* -40.3		
	acidic	-247.189*	–	-247.195* -3.9	-247.190* -0.7	–	-247.236* -29.7		
		1	**TS1-2**	**2S**	**2T**	**TS2-3**	**3**	**TS3-4**	**4(T)**
1,2-dialkynylimidazole	*neutral*	-378.403	-378.361 25.9	-378.387 9.7	-378.378 15.4	-378.384 11.7	-378.422 -12.4	-378.396 3.9	-378.422 -12.3
	acidic	-378.765*	–	-378.753 7.7	-378.747 11.7	–	-378.782* -10.6	–	-378.770* -2.9

Table 2. Calculated energies (hartree) of the enediyne, aza-enediyne, and 1,2-dialkynylimidazole series under both neutral and acidic conditions using the B3LYP functional and the Pople 6-31G(d,p) basis set. **1**, **2**, and **3** represent the corresponding species as in Table 1; for the 1,2-dialkynylimidazole series, **4(T)** represents the carbene resulting from the rearrangement of **3** in its triplet state**.** Failure to optimize the structure and compute the corresponding energy is denoted by an emdash (–). Energies marked with an asterisk (*) were obtained by optimizations in which Cs symmetry was enforced; all others employed no symmetry constraints. For all species but **1**, a relative energy compared to **1** is also provided in kcal mol^{-1}. Except for the enediyne calculations, which were performed using GAMESS, all calculations were performed using Gaussian 09.



(2.7)

Figure 10. Reaction profile for the enediyne series. Energies, shown in hartree, were computed by B3LYP/6-31G(d,p) using GAMESS. Energy shown for the diradical is for its singlet state; singlet-triplet splitting value is given in parentheses in kcal mol^{-1}.

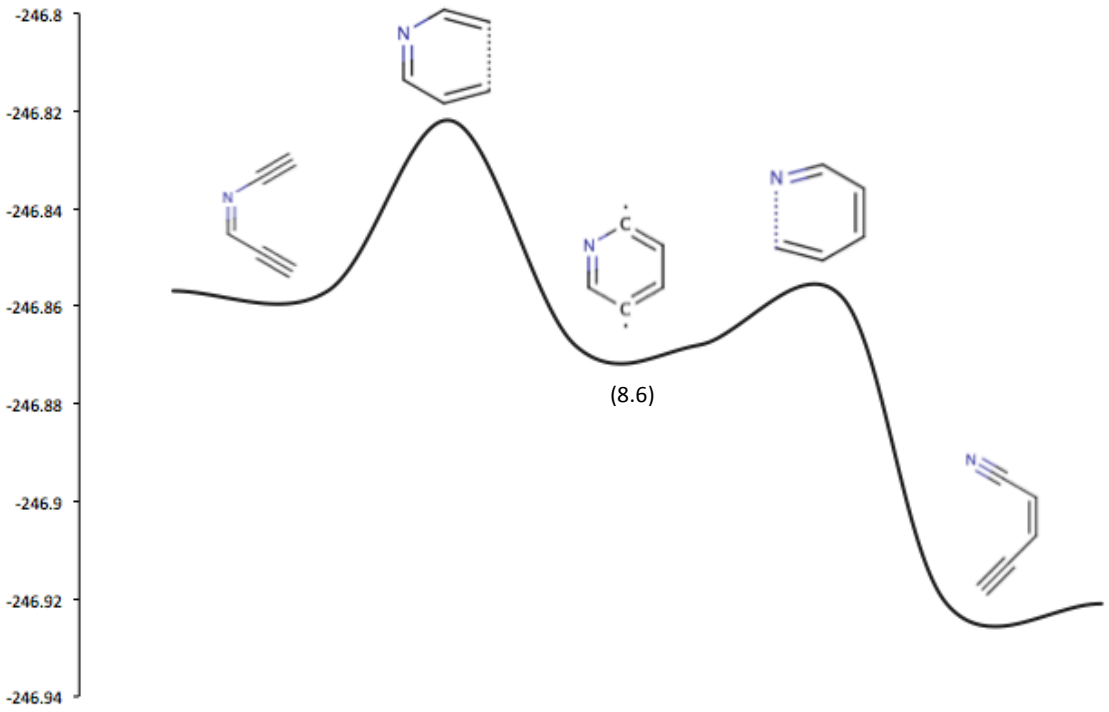


Figure 11. Reaction profile for the aza-enediyne series. Energies, shown in hartree, were computed by B3LYP/6-31G(d,p) using Gaussian 09. Energy shown for the diradical is for its singlet state; singlet-triplet splitting value is given in parentheses in kcal mol^{-1}.

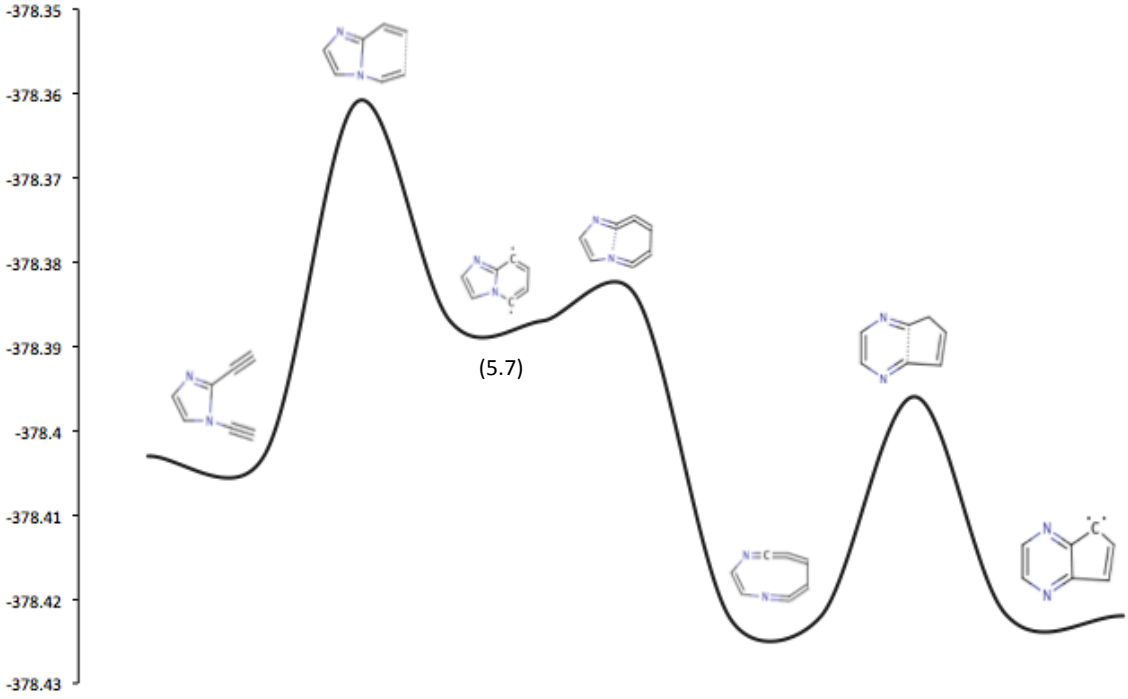


Figure 12. Reaction profile for the 1,2-dialkynylimidazole series. Energies, shown in hartree, were computed by B3LYP/6-31G(d,p) using Gaussian 09. Energy shown for the diradical is for its singlet state while the energy for the carbene is for its triplet state; singlet-triplet splitting value of the diradical is given in parentheses in kcal mol^{-1}.

		1	TS1-2	2S	2T	TS2-3	3		
aza-enediyne	3-21G	-245.346 *92.6*	-245.310 22.8 *92.6*	-245.348 -1.2 *92.7*	-245.331 9.8 *92.6*	–	-245.412 -41.1 *92.6*		
	6-31G(d)	-246.704 *95.9*	-246.669 21.8 *95.9*	–	-246.702 1.3 *92.6*	-246.705 -0.6 *92.7*	-246.768 -40.8 *92.6*		
	6-31G(d,p)	-246.709 *92.6*	-246.675 21.8 *92.6*	-246.720 -6.9 *92.6*	-246.707 1.7 *92.6*	-246.710 -0.4 *92.6*	-246.774 -40.4 *92.6*		
	cc-pVDZ	-246.722 *92.6*	-246.688 21.5 *92.5*	–	-246.719 2.1 *92.6*	-246.724 -0.7 *92.7*	-246.786 -39.1 *92.6*		
	cc-pVTZ	-246.801 *92.6*	-246.762 24.6	–	-246.787 8.5 *92.6*	–	-246.863 -38.8 *92.6*		
		1	**TS1-2**	**2S**	**2T**	**TS2-3**	**3**	**TS3-4**	**4(T)**
1,2-dialkynylimidazole	6-31G(d,p)	-378.181 *139.3*	-378.140 25.8 *139.3*	–	-378.156 15.3 *139.2*	–	-378.201 -12.5 *139.3*	-378.174 4.0 *139.3*	-378.200 -12.4 *139.2*

Table 3. Calculated energies (hartree) of the aza-enediyne and 1,2-dialkynylimidazole series using the B3LYP functional and the indicated basis sets. All calculations were performed by GAMESS and employed no symmetry restrictions. Failure to optimize the structure and compute the corresponding energy is indicated by an emdash (–). For all species but **1**, a relative energy compared to **1** is also provided in kcal mol^{-1}. For all species, a relative energy compared to that provided by the equivalent calculation using Gaussian 09 Is also provided in kcal mol^{-1} (italicized).

A number of the aforementioned calculations were also performed using both GAMESS and Gaussian. Discrepancies in the absolute energy values obtained by the two packages were noted early, so both absolute and relative energies were compared to investigate whether the package used to perform the computations affected the description of the reaction. These data are given in Table 3.

Finally, the aza-enediyne and 1,2-dialkynylimidazole systems were characterized using the BPW91 functional and the cc-pVDZ basis set. Data are provided in Table 4.

		1	TS1-2	2S	2T	TS2-3	3		
aza-enediyne	*neutral*	-246.848*	-246.823* 15.3	-246.869* -13.6	-246.846* 0.8	-246.860* -7.7	-246.908* -37.9		
	acidic	-247.175*	–	-247.184* -5.7	-247.177* -1.4	–	-247.221* -29.2		

		1	TS1-2	2S	2T	TS2-3	3	TS3-4	4
1,2-dialkynylimidazole	*neutral*	-378.384	–	-378.382 1.2	-378.368 10.0	–	-378.412 -17.6	–	-378.408 -15.3
	acidic	-378.750*	–	-378.744 3.7	-378.735 9.1	–	-378.771* -13.1	–	-378.759* -5.9

Table 4. Calculated energies (hartree) of the aza-enediyne and 1,2-dialkynylimidazole series under both neutral and acidic conditions using the BPW91 functional and the cc-pVDZ basis set. All calculations were performed by Gaussian 09. Energies marked with an asterisk (*) were obtained by optimizations in which Cs symmetry was enforced; all others employed no symmetry constraints. Failure to optimize the structure and compute the corresponding energy is indicated by an emdash (–). For all species but **1**, a relative energy compared to **1** is also provided in kcal mol^{-1}.

4. Discussion

4.1 DFT predicts cyclization barriers in good agreement with available experimental data

Experiment gives an energetic barrier to cyclization of 28.7 kcal mol^{-1} for the enediyne.[30] The value obtained by DFT using the B3LYP functional and 6-31G(d,p) basis set differed little: at 30.7 kcal mol^{-1}, it is only 2.0 kcal mol^{-1} higher than the experimentally obtained value.

Theoretical values for the barrier to the aza-enediyne's cyclization ranged from 21.5 to 24.6 kcal mol^{-1} when computed using the B3LYP functional. Nearly without exception, these values are in excellent agreement with the experimentally derived value of 22.5 ± 1.5 kcal mol^{-1} (for a close analog of the aza-enediyne).[31] Computing this barrier using the BPW91 functional provided poorer agreement, at only 15.3 kcal mol^{-1}. Because transition states for the protonated aza-enediyne had not been found at the time of this writing, an energetic barrier for this molecule under sufficiently acidic conditions could not be predicted. Experimental data for the barrier to 1,2-dialkynylimidazole cyclization was not found, though some exists for a similar molecule with a large substituent at one acetylenic position, whose energetic barrier to cyclization is 30.0 kcal mol^{-1}.[32] Intuitively, the 1,2-dialkynylimidazole lacking this substituent would have a lower energetic barrier to cyclization; indeed, the barrier predicted by DFT is 25.9 kcal mol^{-1}.

4.2 The DFT description of the aza-enediyne does not depend strongly on the choice of basis set

In general, the choice of basis set did not have a major effect on the description of the aza-enediyne's reactivity when all calculations were performed using the same functional. For instance, all basis sets predicted a barrier to cyclization within 1

kcal mol^{-1} of one another, and all basis set choices yielded the same predicted reactivity based on the criteria of Kraka and Cremer. However, performing the calculations using the BPW91 functional provided very different energetic barriers; though the trends in stability remained the same as those predicted by B3LYP, relative energies differed greatly. Furthermore, as noted previously, the energetic barriers predicted by BPW91 computations were lower than what the experimental literature suggests.

4.3 Predicted fates of enediyne analogs *in vivo*

The predicted energetic barrier to enediyne cyclization, at 30.7 kcal mol^{-1}, is significantly higher than 24 kcal mol^{-1}. Assuming the criteria of Kraka and Cremer hold, this species is therefore predicted not to cyclize under physiological conditions at neutral pH. Because the molecule lacks atoms or functional groups that would be readily protonated, under acidic conditions its reactivity is predicted to be similar.

Based on the aforementioned energetic barriers obtained by DFT, the aza-enediyne would be predicted to cyclize to its diradical form under neutral conditions. However, all calculations predict barriers far below 14 kcal mol^{-1} for the rearrangement of the diradical species to its retro-Bergman cyclization product, an enynenitrile. Furthermore, due to its fairly high singlet-triplet splitting value – consistently predicted to be over 8 kcal mol^{-1} – the diradical species is predicted to be unreactive in hydrogen abstraction reactions, in accordance with the proposed idea that low singlet-triplet splitting corresponds to high hydrogen abstraction ability.[33] Therefore, under neutral conditions, the aza-enediyne is predicted to cyclize to a fairly unreactive diradical form then be rapidly converted to its stable enynenitrile form. The absence of computed transition state energies for the protonated aza-enediyne precludes a full prediction of the molecule's behavior under acidic conditions; however, it may be noted that the relative stabilities of the protonated species follow the same trend as for the unprotonated ones. Furthermore, the singlet-triplet gap reduces significantly, to around 4 kcal mol^{-1}; the protonated diradical, then, is predicted to have greater hydrogen abstraction ability than its unprotonated counterpart.

DFT predicted the cyclization barrier for 1,2-dialkynylimidazole under neutral conditions to be 25.9 kcal mol^{-1}, just above 24 kcal mol^{-1}. By the criteria proposed by Kraka and Cremer, this barrier is slightly too high for significant cyclization of 1,2-dialkynylimidazole to occur under physiological conditions. The diradical species, if generated, would likely be fleeting, as the barrier for the retro-Bergman rearrangement is significantly lower than 14 kcal mol^{-1}; moreover, as evidenced by its high singlet-triplet gap, it would be unreactive. The relative stability of the nine-membered cumulene ring formed by the retro-Bergman rearrangement at first seems odd due to its significant strain; however, this seemingly surprising observation may be rationalized by noting that one of the molecule's resonance contributors is aromatic, conferring it great stability. Again, the failure to compute transition state geometries and energies for the protonated

form of this molecule precludes a full prediction of its behavior under acidic conditions, but as with the aza-enediyne, the same general trend of stability follows as with the unprotonated molecule, and protonation significantly reduces the singlet-triplet gap of the diradical species.

4.4 Absolute energies, but not relative energies, depend strongly on the software package used to compute them

The unexpected discrepancy between absolute energies computed by GAMESS and by Gaussian 09 should be noted. Table 3 highlights this observation; although the relative energies between species computed by GAMESS are nearly always very close (within tenths of a kcal mol^{-1}) to those computed by Gaussian (Table 1), they differ in absolute energy by a significant margin. Curiously, the discrepancy is consistent; for all aza-enediyne calculations, it rarely deviated more than 0.1 kcal mol^{-1} from 92.6 kcal mol^{-1}, and the same held true for 1,2-dialkynylimidazole calculations, though with a consistent difference in 139.2 kcal mol^{-1}. The reasons for these differences are unclear (to this naïve author), but because they are of such a significant magnitude, they nonetheless emphasize the importance of noting the source of theoretically obtained energetic data, whether comparing to prior studies in the literature or one's own previous studies.

5. Acknowledgments

I am deeply grateful to the Texas Advanced Computing Center and the San Diego Supercomputer Center for allowing me to perform many of my calculations on their supercomputers (Ranger and Trestles, respectively) and for providing invaluable technical support along the way. My old laptop, which overheated often and is now long gone, also thanks these superior computers for relieving it of (most of) this duty.

Special thanks to two excellent professors: Dr. Kerwin for teaching me to run my first electronic structure calculations patiently and without condescension, giving me my first foray into the difficult but enlightening world of theoretical chemistry; Dr. Stanton for teaching terrific quantum chemistry and scientific computation courses (you know, the notes I have from these classes are still better than anything I've seen in a book).

Thank you to the GAMESS Google group, which I stalked incessantly while trying to figure out various issues and obstacles I encountered throughout the process of using this strange software. I was curiously reassured to find that many people much smarter and more advanced than I am have struggled with the same sorts of problems that I was facing – and even some problems I had already surmounted. Thanks also for the laughs: I'm not sure I will ever see anything funnier than theoretical chemists passionately arguing online about the appropriate choice of DFT functionals.

Great thanks to my other professors who have guided me in a wide variety of capacities through my undergraduate career, in particular Dr. Cline, Dr. Laude, Dr. Keatinge-Clay, and Dr. Beckham.

Most sincere thanks to my family – particularly Mom, Dad, Mooie, Val, Eliot, and Mike – who always understood me even when they didn't understand the science. Thanks for cheering me along and reminding me that the best way to enjoy what you love is to share it.

Last, but not least, are my friends who've helped me in ways great and small, who are too numerous to name here. While it's true that the privileged overlook the role of small organic molecules in our health, perhaps even more often forgotten is the simple joy of laughter. Special thanks to Caitlin Taylor for singing "Torn" by Natalie Imbruglia with me in her apartment that one night.

Thanks to everyone for being patient with me.

6. References

[1] W. E. Bachmann, H. T. Clarke, *J. Am. Chem. Soc.* **1927,** *49,* 2089.

[2] G. Wittig, *Naturwissenschaften* **1942,** *30,* 696.

[3] G. Wittig, G. Fuhrmann, *Bw.* **1940,** *73,* 1197.

[4] R. Huisgen, H. Rist, *Naturwissenschaften* **1955,** *41,* 358.

[5] J. D. Roberts, H. E. Simmons, L. A. Carlsmith, C. W. Vaughan, *J. Am. Chem. Soc.* **1953,** *75,* 3290.

[6] F. Sondheimer, Y. Amiel, Y. Gaoni, *J.Amer. Chem. Soc.* **1961,** *84,* 270.

[7] W. H. Okamura, F. Sondheimer, *J. Amer. Chem. Soc.* **1967,** *89,* 5991.

[8] R. R. Jones, R. G. Bergman, *J. Amer. Chem. Soc.* **1972,** *94,* 660.

[9] R. G. Bergman, *Acc. Chem. Res.* **1973,** *6,* 25.

[10] R. K. Elmpuru, R. J. White, *Cancer Res.* **1983,** *43,* 2819.

[11] M. D. Lee, J. K. Manning, D. R. Williams, N. A. Kuck, R. T. Testa, D. B. Borden, *J. Antibiot.* **1989,** *42,* 1070.

[12] J. J. DeVoss, C. A. Townsend, W.-D. Ding, G. O. Morton, G. A. Ellestad, N. Zein, A. B. tabor, S. L. Schreiber, *J. Am. Chem. Soc.* **1990,** *112,* 9669.

[13] G. A. Ellestad, P. R. Hamann, N. Zein, G. O. Morton, M. M. Siegel, M. Pastel, D. B. Borders, W. J. McGahren, *Tetrahedron Lett.* **1989,** *30,* 3033.

[14] L. Yu, J. Golik, R. Harrison, P. Dedon, *J. Am. Chem. Soc.* **1994,** *116,* 9733.

[15] J. J. Hangeland, J. J. Devoss, J. A. Heath, C. A. Townsend, W. D. Ding, J. S. Ashcroft, G. A. Ellestad, *J. Am. Chem. Soc.* **1992,** *114,* 9200.

[16] J. Golik, J. Clardy, G. Dubay, G. Groenewold, H. Kawaguchi, M. Konishi, B. Krishnan, H. Ohkuma, K. Saitoh, T. W. Doyle, *J. Am. Chem. Soc.* **1987,** *109,* 3461.

[17] M. Konishi, H. Ohkuma, K. Matsumoto, T. Tsuno, H. Kanei, T. Miyaki, T. Oki, H. Kawaguchi, G. D. VanDuyne, J. Clardy, *J. Antibiot.* **1989,** *42,* 1449.

[18] K. Edo, M. Mizugaki,, Y. Koide, H. Seto, K. Furihata, N. Otake, N. Ishida, *Tetrahedron Lett.* **1985,** *26,* 331.

[19] K. C. Nicolaou, W.-M. Dai, *Angew. Chem. Int. Ed. Engl.* **1991,** *30,* 1387.

[20] I. F. Tannock, D. Rotin, *Cancer Res.* **1989,** *49,* 4373.

[21] S. Osinsky, L. Bubnovskaya, *Arch. Geschwulstforsch.* **1984,** *54,* 463.

[22] E. Kraka, D. Cremer, *J. Am. Chem. Soc.* **2000,** *122,* 8245.

[23] Maeda, H. In *Neocarzinostatin: The Past, Present, and Future of an Anticancer Drug;* Maeda, H., Edo, K., Ishida, N., Eds.; Springer: New York, 1997.

[24] W. Sander, *Acc. Chem. Res.* **1999,** *32,* 669.

[25] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964,** *136,* B864.

[26] W. Kohn, L. J. Sham, *Phys. Rev.* **1965,** *140,* A1133.

[27] M. W. Schmidt, et. al. *J. Comput. Chem.* **1993**, *14,* 1347.

[28] M. J. Frisch, et. al. Gaussian, Inc., Wallingford CT, 2009.

[29] E. C. Sherer, K. N. Kirschner, F. C. Pickard, C. Rein, S. Feldgus, G. C. Shields, *J. Phys. Chem. B* **2008,** *112,* 16917.

[30] W. R. Roth, H. Hopf, C. Horn, *Chem. Ber.* **1994,** *127,* 1765.

[31] J. H. Hoffner, M. J. Schottelius, D. Feichtinger, P. Chen, *J. Am. Chem. Soc.* **1998,** *120,* 376.

[32] C. Laroche, J. Li, C. Golzales, W. M. David, S. M. Kerwin, *Org. Biomol. Chem.* **2010,** *8,* 1535.

[33] C. F. Logan, P. Chen, *J. Am. Chem. Soc.* **1996,** *118,* 2113.