A Comparative Analysis of the Electronic and Structural Characteristics of Polyphenolic Inhibitors of PAI-1

Patrick Spoutz
Eastern Michigan University

Follow this and additional works at: http://commons.emich.edu/honors

Recommended Citation
http://commons.emich.edu/honors/217

This Open Access Senior Honors Thesis is brought to you for free and open access by the Honors College at DigitalCommons@EMU. It has been accepted for inclusion in Senior Honors Theses by an authorized administrator of DigitalCommons@EMU. For more information, please contact lib-ir@emich.edu.
A Comparative Analysis of the Electronic and Structural Characteristics of Polyphenolic Inhibitors of PAI-1

Abstract
Polyphenolic compounds containing two gallate groups have been shown to be effective inhibitors of Plasminogen Activator Inhibitor-1 (PAI-1). In this work we use Density Functional calculations to examine energetically-accessible configurations of two digallate compounds exhibiting disparate inhibitory activity toward PAI-1. In addition, we examine the partial charges of potentially acidic protons and the relative stability of the respective conjugate bases of a select group of PAI-1 inhibitors. This analysis was carried out to determine whether such characteristics can be used to distinguish between effective and less effective inhibitors. An examination of the stable configurations of CDE-008 and CDE-056 indicates that the range of distances between gallate rings for the poor inhibitor (CDE-008) does not overlap with the range of distances for the better inhibitor (CDE-056). However, no clear differences were observed in the partial charge distribution or relative stability of conjugate bases for the group of inhibitors examined here.

Degree Type
Open Access Senior Honors Thesis

Department
Chemistry

Subject Categories
Chemistry

This open access senior honors thesis is available at DigitalCommons@EMU: http://commons.emich.edu/honors/217
A COMPARATIVE ANALYSIS OF THE ELECTRONIC AND STRUCTURAL CHARACTERISTICS OF POLYPHENOLIC INHIBITORS OF PAI-1

by

Patrick Spoutz

A Senior Thesis Submitted to the Eastern Michigan University Honors Program

In Partial Fulfillment of the Requirement for Graduation with Honors in Chemistry

Approved at Ypsilanti, Michigan on this date, April 19th, 2010

Supervising Instructor – Dr. Maria C. Milletti, Professor of Chemistry

Honors Advisor – Dr. Heather Holmes, Associate Professor of Chemistry

Department Head – Dr. Ross Nord, Professor of Chemistry

Honors College Director – Dr. James Knapp
Abstract

Polyphenolic compounds containing two gallate groups have been shown to be effective inhibitors of Plasminogen Activator Inhibitor-1 (PAI-1). In this work we use Density Functional calculations to examine energetically-accessible configurations of two digallate compounds exhibiting disparate inhibitory activity toward PAI-1. In addition, we examine the partial charges of potentially acidic protons and the relative stability of the respective conjugate bases of a select group of PAI-1 inhibitors. This analysis was carried out to determine whether such characteristics can be used to distinguish between effective and less effective inhibitors. An examination of the stable configurations of CDE-008 and CDE-056 indicates that the range of distances between gallate rings for the poor inhibitor (CDE-008) does not overlap with the range of distances for the better inhibitor (CDE-056). However, no clear differences were observed in the partial charge distribution or relative stability of conjugate bases for the group of inhibitors examined here.
# Table of Contents

I. Chapter 1 – Introduction..........................................................p. 4

II. Chapter 2 – Theoretical Background........................................ p. 7
   A. The Schrödinger equation.................................................. p. 7
   B. The Born-Oppenheimer approximation................................. p. 7
   C. The variation method...................................................... p. 8
   D. The molecular orbital approximation.................................... p. 9
   E. Basis sets................................................................. p. 10
   F. Hartree-Fock theory...................................................... p. 11
   G. Density functional theory................................................ p. 15
   H. Natural bond orbitals..................................................... p. 17
   I. Details of the calculations............................................... p. 18

III. Chapter 3 – Results and Discussion...................................... p. 19
   A. Conformational studies.................................................... p. 19
   B. Acidity of phenolic hydrogens......................................... p. 32
   C. Conclusions.............................................................. p. 38

IV. References........................................................................... p. 39

V. Appendix A........................................................................ p. 43
Introduction

PAI-1 (plasminogen activator inhibitor-1) is a serine protease inhibitor that plays a critical role in lessening the rate at which blood clots are dissolved and preventing hemorrhaging by inhibiting the fibrinolytic system. [1] Specifically, it inhibits urokinase plasminogen activator (uPA) and tissue plasminogen activator (tPA), which, collectively, are important in activating plasmin, an enzyme whose principal purpose is to digest coagulated tissue at the site of injury. [2] By inhibiting uPA and tPA, PAI-1 allows blood clots to persist. However, in some cases, it is desirable to increase the rate at which blood clots are dissolved. These situations include a variety of cardiovascular diseases, such as atherosclerosis and thromboembolic diseases. Therefore, effective inhibitors of PAI-1 are needed. [1] Individuals with low levels of PAI-1 show abnormal bleeding, suggesting an important role for PAI-1 in blood clot stabilization. [3] Since PAI-1 is only involved in blood clot fibrinolysis and not coagulation, its inhibition has no discernable effect on blood clot formation, only on blood clot dissolution. Finding ways to inhibit PAI-1 and decreasing its blood concentration may be useful means of battling cardiovascular disease. [1]

Dr. Emal at Eastern Michigan University and his collaborators at the University of Michigan (Drs. Dan Lawrence and Mark Warnock) are attempting to produce a more active inhibitor that fits the active site most efficiently in order to better understand the inhibition process. Tannic acid, shown below in Figure 1.a, has been shown to be a very effective inhibitor. Unfortunately, tannic acid is an unsuitable candidate for drug development because of its toxicity. [4] Dr. Emal’s work focuses on finding inhibitors that have similar characteristics to tannic acid but have physicochemical characteristics
appropriate for a drug. Clearly, the distinctive feature in the structure of tannic acid is the presence of several gallate groups. However, gallic acid (shown in Figure 1.b) has been proven ineffective, a sign that one gallate group is insufficient for effective inhibition. Compounds containing two gallates have been shown to be much more effective, consequently these are the types of compounds (digallates) that Dr. Emal’s group is pursuing. [5]

![Figure 1: (a) Tannic Acid (b) Gallic acid](image)

In this work we examine a collection of digallate compounds with varying activities toward PAI-1 inhibition using Density Functional methods to determine the qualities that separate effective from less effective inhibitors.
Project Goals

The overall goal of this project is to explore the relationship between electronic and structural factors such as the distance between polyphenolic (gallate) rings, the relative acidity of gallate protons, and the inhibitory activity of a series of polyphenolic PAI-1 inhibitors. Any such trends may be able to point us toward new methods for identifying promising inhibitor candidates. Specifically, we will determine the relative thermodynamic stability of different conformations of a variety of digallate compounds in an attempt to understand which conformations lead to increased activity.

We also compare the partial charges of the gallate hydroxyl protons as calculated by an NBO analysis, as well as the relative stability of the respective conjugate bases. Taken together, these two factors will indicate whether there are differences in the pKa of the gallate protons of various inhibitors and whether these differences can be related to the activity of the inhibitor.

By repeated analysis of this kind, we hope to contribute to the search for a suitable candidate for the conformation of the active site.
Theoretical Background [6,7]

The Schrödinger Equation

Quantum mechanics is the branch of physics that describes the behavior of matter at the smallest level. The most important tool available to the quantum theorist is the Schrödinger wave equation, shown below in its time-independent form; this allows the theorist to predict the location of an electron in space probabilistically.

$$\hat{H}\Psi = E\Psi \quad (1)$$

The Hamiltonian operator, $\hat{H}$, is the energy operator on the wavefunction, $\Psi$. When $\Psi$ is an eigenfunction of the Hamiltonian, the scalar eigenvalue, $E$, represents the total energy of the system. The square of the wavefunction determines the probability of finding an electron in a given region of space. When the probability density distribution for a given electron in the system is examined, the resulting volume is known as an orbital.

The Born-Oppenheimer Approximation

Given the complexity inherent in modeling systems with many atoms, many approximations, including the Born-Oppenheimer approximation, are needed to lessen the computational load. The molecular Hamiltonian is shown in equation 2, where the first term represents the kinetic energy of the nuclei, the second term operates to determine the kinetic energy of the electrons, the third term represents the nuclear-nuclear repulsion energy, and the fifth term represents the total potential energy due to the repulsion of each electron-electron pair.
$$\hat{H} = -\left(\frac{\hbar^2}{2m}\right)\sum_{\alpha} \frac{1}{m_{\alpha}} \nabla_{\alpha}^2 - \left(\frac{e^2}{4\pi \epsilon_0}\right) \sum_{i} \nabla_i^2 - \sum_{\alpha} \sum_{\beta > \alpha} \frac{z_{\alpha} z_{\beta} e^{r_{\alpha\beta}}}{r_{\alpha\beta}} - \sum_{i} \sum_{j>i} \frac{e^{r_{ij}}}{r_{ij}}$$

(2)

The Born-Oppenheimer Approximation is typically the first approximation utilized. It allows the theorist to separate the nuclear and electronic motion. Since the nucleus is many orders of magnitude more massive than the electrons, it can be assumed to be stationary with respect to electronic motion with little loss of accuracy. With this assumption, the molecular Hamiltonian is simplified so that it is based entirely on the electronic component (equation 3). Coulombic interactions between the nucleus and electrons are still included in the second term, only the nuclear kinetic component falls away and the nuclear-nuclear repulsion term becomes a constant.

$$\hat{H}_{el} = -\left(\frac{\hbar^2}{2m_e}\right) \sum_i \nabla_i^2 - \sum_{\alpha} \sum_{i} \frac{z_{\alpha} e^{r_{\alpha i}}}{r_{i\alpha}} + \sum_{i} \sum_{j>i} \frac{e^{r_{ij}}}{r_{ij}}$$

(3)

Even with this simplified operator, there is no analytical solution to the Schrödinger equation. The conventional approach to finding an approximate solution to the equation for molecules is discussed below.

The Variation Principle

There is no analytical (or exact) solution to the Schrödinger equation when the system under study contains two or more electrons. This is because the last term in the molecular Hamiltonian (see equation 2), representing electron-electron repulsion, prevents the equation from being separated into individual terms. In cases where the wavefunction is not an eigenfunction of the electronic Hamiltonian, no energy eigenvalue can be obtained. At best, we can arrive at an approximate, or expectation, value for the
energy of the system. In these cases, the variation principle allows us to establish an upper bound for the true eigenvalue. From this principle, it is established that if $\Psi$ is any normalized and well behaved wavefunction, then the following relationship holds true

$$\int \Psi^* \hat{H} \Psi \, d\tau \geq E_{gs}$$

(4)

where the expression on the left represents the expectation (or average) value of the energy, $E_{gs}$ is the true ground state energy and $\hat{H}$ is the molecular Hamiltonian operator.

If the exact wavefunction ($\Psi$) is not known, a trial wavefunction can be substituted. Typically, a trial wavefunction will contain one or more parameters that can be varied computationally in order to find the wavefunction with the lowest energy. In practice, the expectation value, calculated as shown in equation 4, is minimized with respect to any variational parameter. This minimized energy can then be said to be a good approximation to the true energy eigenvalue for the system at hand.

The Molecular Orbital Approximation

One common way to introduce variational parameters in the trial wavefunction is to express the function that represents each Molecular Orbital as a Linear Combination of Atomic Orbital functions (LCAO-MO). This is shown in equation 5, where $\Phi_\mu$ are the atomic orbitals and $c_{\mu\ell}$ are coefficients in the linear expansion.

$$\Phi_\ell = \sum_{\mu=1}^{N} c_{\mu\ell} \chi_\mu$$

(5)

The molecular wavefunction can then be optimized by setting each partial derivative of the energy expectation value with respect to the set of coefficients equal to zero.

Using the linear variational parameters illustrated above leads to a description of molecular orbitals as a linear combination (mixing) of atomic orbitals, thus spanning the
entire molecule. The Pauli Exclusion Principle states that a wavefunction describing a system of fermions (electrons are fermions) must be antisymmetric with respect to the exchange of any two particles. To preserve the anti-symmetric nature of the wavefunction, we write the molecular wavefunction as a Slater determinant of molecular orbital functions, as shown in equation 6. In this way, each electron occupies in turn each molecular orbital, which keeps us good with the Pauli Principle. In fact, if two electrons with the same spin were to inhabit the same orbital, two columns of the determinant would become identical and the determinant would vanish. It is in this way that the Pauli Exclusion Principle is obeyed.

\[ \Psi_{el} = \frac{1}{\sqrt{n!}} \begin{bmatrix} \Phi_1(1) & \cdots & \Phi_n(1) \\ \vdots & \ddots & \vdots \\ \Phi_1(n) & \cdots & \Phi_n(n) \end{bmatrix} \] (6)

**Basis Sets**

A basis set is defined by the set of atomic functions used in the linear expansion of molecular orbitals (LCAO-MO). These are modified hydrogen-like functions, derived from the exact solution to the Schrödinger equation for the hydrogen atom. In practice, however, hydrogen-like functions are time-consuming to integrate. When a computer program is used to calculate the integrals necessary to solve the Hartree-Fock equations described below, Gaussian type functions are employed to describe the atomic orbitals. Gaussian functions, shown in equation 7 in polar coordinates, are computationally less expensive to integrate and differentiate.

\[ G_{nlm} = N_n r^{n-1} e^{-\alpha r^2} Y_l^m(\theta, \phi) \] (7)
In equation 7, \( n \) represents the principal quantum number, \( N_n \) is the normalization constant, \( \alpha \) is a parameter that accounts for the orbital diffuseness and \( Y_{l}^{m}(\theta, \phi) \) is the spherical harmonics portion of the wavefunction.

However, Gaussian functions are not a good representation of hydrogen-like atomic functions. Consequently, a linear combination of several Gaussian functions (called primitive or contracted Gaussian functions) is used to describe each atomic orbital. Basis sets are labeled in a manner that indicates the number of primitive Gaussian functions used to describe each atomic orbital. Specifically, the labels take the form \( n-xyzG \), where \( n \) is the number of primitive Gaussians used in the linear expansion of core atomic orbitals, \( x, y \) and \( z \) denote the number of functions used in the description of valence orbitals, and \( G \) specifies the use of Gaussian type functions. For example, in this work we often use the 6-311G basis set, where six primitive Gaussian functions are combined to represent the core orbitals. The valence orbitals are modeled using three sets of primitive Gaussian functions; in this case the sets are composed of three, one and one Gaussians, respectively. This is termed a ‘split valence’ basis set, where each set describes different portions of the valence orbital. Increasing the number of primitive Gaussian functions used in a basis set will generally lead to superior (i.e., more accurate) results, though it increases the computational intensity of the calculations.

Hartree-Fock Theory [8]

Among the many modern computational techniques that can be used to obtain the electronic wavefunction, the Hartree-Fock Self-Consistent Field (HF-SCF) approach is the one most commonly used. In this method, as described below, a set of homogeneous
second order linear differential equations are derived and then solved using an iterative approach.

The electronic Hamiltonian shown in equation 3 can be separated into one and two electron components. The one-electron operator, shown in equation 8, represents the kinetic energy of a single electron, as well as its attraction to each nucleus in the molecule.

\[
\hat{\mathcal{H}}_I = -\frac{1}{2} \nabla_i^2 - \sum_\alpha \frac{Z_\alpha}{r_{i\alpha}}
\]  

(8)

In addition, we can define the two electron operator, shown in equation 9, which represents the electron-electron repulsion.

\[
\hat{\mathcal{G}}_{ij} = \frac{1}{r_{ij}}
\]  

(9)

We can then rewrite the electronic Hamiltonian as follows, using the operators defined in equations 8 and 9.

\[
\hat{\mathcal{H}}_{el} = \sum_i \hat{\mathcal{H}}_I + \sum_i \sum_{j \neq i} \hat{\mathcal{G}}_{ij}
\]  

(10)

Recall that the left-hand side of equation 4 represents the expectation value of the energy, which is an upper bound to the true ground state energy. If we arbitrarily label this value as the Hartree-Fock energy, we obtain the expression in equation 11.

\[
E_{HF} = \int \Psi^* \hat{\mathcal{H}} \Psi \, d\tau
\]  

(11)

Using the more compact bra-ket notation, equation 11 becomes

\[
E_{HF} = < \Psi_{el} | \hat{\mathcal{H}}_{el} | \Psi_{el} >
\]  

(12)

Here \( \Psi \) is \( \Psi_{el} \) because we are using a Slater determinant to represent the molecule’s wavefunction, as described above. Expanding each of the expressions in equation 12 produces a sum of terms in which the operators in equations 8 and 9 are operating on products of molecular orbital wavefunctions from the Slater determinant.
We define two such terms as the Coulomb \((J_{ij})\) and exchange \((K_{ij})\) integrals. These are shown in equations 13 and 14, respectively.

\[
J_{ij} = \langle \Phi_i(1)\Phi_i(2)|\frac{1}{r_{12}}|\Phi_j(1)\Phi_j(2) \rangle
\]

\[
K_{ij} = \langle \Phi_i(1)\Phi_j(2)|\frac{1}{r_{12}}|\Phi_i(2)\Phi_j(1) \rangle
\]

The Coulomb integral has a relatively straightforward physical interpretation: it represents the electrostatic repulsion between electrons. Unfortunately, the exchange integral has no immediate classical interpretation. Its name comes from the fact that the two electrons (1 and 2) exchange wavefunctions \((\Phi_i \text{ and } \Phi_j)\) during the interaction described in \(K_{ij}\). However, the exchange integral makes a nontrivial contribution to the energy of the molecule, and as such cannot be ignored.

Each molecular function in the above integrals can be expressed as the linear expansion of atomic orbitals as shown in equation 5. After this substitution is made, we can employ the variational principle by taking the derivative of equation 12 with respect to each coefficient in the linear expansion and then setting each derivative equal to zero. This results in a set of equations, called the Hartree-Fock equations. We first define the Fock operator as

\[
\hat{F} = \hat{H}^{core}(1) + \sum_{j=1}^{n/2} [2\hat{j}_j(1) - \hat{R}_j(1)]
\]

where \(\hat{j}\) and \(\hat{R}\) represent the Coulomb and exchange operators, respectively, and \(\hat{H}^{core}\) represents the sum of the kinetic energy of the electrons and the electron-nuclear attraction.
Considering a closed shell configuration containing $n$ electrons occupying $n/2$ orbitals and using the definition of the Fock operator given above, the equation representing the total energy of the system can be rewritten as

$$E_{HF} = 2 \sum_{i=1}^{n/2} <\Phi_i(1)|\hat{F}_i|\Phi_i(1)> + \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij})$$  \hspace{1cm} (16)$$

After the appropriate LCAO linear expansion is substituted for each one of the molecular orbitals, the Variational Principle is employed and partial derivatives are taken with respect to each coefficient in the linear expansion and set to zero. This gives rise to a set of homogeneous linear differential equations, each describing one of the electrons in the system. The general form of these equations, in matrix representation, is shown in equation 17.

$$\sum_s c_{si}(F_{rs} - \varepsilon_i S_{rs}) = 0$$  \hspace{1cm} (17)$$

These are the Hartree-Fock-Roothaan equations, where $F_{rs}$ is the $rs^{th}$ element in the Fock matrix, $S_{rs}$ is the $rs^{th}$ element in the overlap matrix, and $c_{si}$ is the $s^{th}$ coefficient in the linear expansion of the $i^{th}$ molecular orbital.

It now becomes possible to use the methods of linear algebra to solve for each coefficient in each of the Hartree-Fock-Roothaan equations. However, this requires an initial guess at the wavefunction, labeled $\Psi_0$, because we do not know the set of $c_{si}$ coefficients until we have solved the above equations, yet the coefficients are necessary to build the Fock matrix. Given this guess, the set of HF-SCF equations are solved. The resulting set of coefficients is used to generate a new wavefunction, designated $\Psi_1$. This process is repeated in an iterative fashion, until the wavefunction and eigenvalues show no noticeable change from one iteration to another; the system is then said to be “self-consistent.” This is the so-called Self-Consistent Field Method.
Density Functional Theory [9]

The Hartree Fock approach described above is termed an ab initio method because it makes no approximations in the solution to the Schrödinger equation, other than those previously discussed. In contrast, Density Functional Theory is not ab initio because its functionals are not exact, rather they incorporate one or more parameters.

Density functional theory (DFT) solves for the electron probability density, designated $\rho$, instead of the wavefunction, as in HF theory. Its strongest asset is that it produces more accurate results than HF-SCF with roughly the same computational work.

The key to the success of DFT is a reduction in the overall number of variables under observation. Instead of calculating $3n+n$ variables (corresponding to the spatial coordinates of every electron and spin, respectively) we only need to use $3 (x,y,z)$. The Hohenberg-Kohn theorems and Kohn-Sham methods form the core of DFT (introduced computationally in 1995 in Professor J.A. Pople’s Gaussian software). [10] The Hohenberg-Kohn Theorem is formulated as follows: for a molecule with a nondegenerate ground state, the ground state molecular energy, wavefunction, and all other properties are uniquely determined by the ground-state electron probability density $\rho_0(x,y,z)$. If we can determine the electron density, $\rho_0(x,y,z)$, we can determine all other variables of interest.

The ground-state electron probability density is, in turn, a function dependent on its coordinates in space. We can then state the ground-state energy is a functional of $\rho_0$, given in equation 18,

$$E_0 = E_\psi[\rho_0]$$  (18)
Where \( v \) represents the dependence of the ground state energy on the external potential (the interaction between the electron and the nuclei), denoted \( v(\mathbf{r}_i) \). We now form an analogous equation for the ground state energy using the individual terms of the molecular Hamiltonian (each dependent upon \( \rho_0 \)), as shown in 19.

\[
E_0 = E_v[\rho_0] = \bar{T}[\rho_0] + \bar{V}_{Ne}[\rho_0] + \bar{V}_{ee}[\rho_0]
\]  

(19)

These terms correspond to the molecular Hamiltonian’s terms for the average kinetic energy, electron-nucleus potential, and electron-electron repulsion, respectively. \( \bar{T} \) and \( \bar{V}_{ee} \) are more difficult to calculate than \( \bar{V}_{Ne} \), since the electron-nucleus potential can be calculated directly for each electron as a function of distance, and charge. Therefore, for the former two terms we define a new functional,

\[
F[\rho_0] = \bar{T}[\rho_0] + \bar{V}_{ee}[\rho_0]
\]  

(20)

such that equation 20 can be rewritten as

\[
E_0 = F[\rho_0] + \bar{V}_{Ne}[\rho_0] + \bar{V}_{ee}[\rho_0]
\]  

(21)

Hohenberg and Kohn also produced a modified version of the variational principle to work with DFT, formulated as follows:

\[
E_0 \leq E_v[\rho_{trial}]
\]  

(22)

The Hohenberg-Kohn theorem does not provide a method for the calculation of this functional, notably \( \rho_0 \). We must use the Kohn-Sham method to solve for \( F[\rho_0] \) and to find solutions for all elements of equation 21. The key to KS (as Kohn-Sham will hereafter be known) is to use a reference system (labeled as s) consisting of non-interacting electrons in the external potential, \( v(\mathbf{r}_i) \), designated s. As a result of this, there will be no \( \bar{V}_{ee} \) term. Consequently,

\[
\rho_s(\mathbf{r}) = \rho_0(\mathbf{r})
\]  

(23)
where $\rho_s(\mathbf{r})$ is equivalent to the reference system’s ground-state density. We now define $ar{T}[\rho_0]$ and $\bar{V}_{ee}[\rho_0]$ as the difference between the reference system and the real system.

$$
\Delta \bar{T}[\rho_0] = \bar{T}[\rho_0] - T_s[\rho_0] \tag{24}
$$

$$
\Delta \bar{V}_{ee}[\rho_0] = \bar{V}_{ee}[\rho_0] - \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} \, dr_1 \, dr_2 \tag{25}
$$

Equation 25 can be rewritten as,

$$
E_v[\rho_0] = \int \rho(r)\nu(r) \, dr + \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} \, dr_1 \, dr_2 + T_s[\rho_0] + \Delta \bar{T}[\rho_0] + \Delta \bar{V}_{ee}[\rho_0] \tag{26}
$$

where $\int \rho(r)\nu(r) \, dr$ is the nuclei-electron attraction. The only terms left undefined are $\Delta \bar{T}[\rho_0]$ and $\Delta \bar{V}_{ee}[\rho_0]$. We define their sum to be the exchange-correlation energy functional, as follows.

$$
E_{xc}[\rho] = \Delta \bar{T}[\rho_0] + \Delta \bar{V}_{ee}[\rho_0] \tag{27}
$$

From this exchange-correlation energy functional, we use methods similar to the self-consistent field approach and a basis set expansion, such as 6-31G, to solve for $\rho_0$. A variety of functionals is available. For this work, we use B3LYP, which combines Becke’s three-parameter exchange functional with Yang, Lee, and Parr’s correlation functional in a hybrid HF-DFT functional.

Natural Bond Orbitals \cite{11,12}

The use of natural bond orbitals (NBO) is a form of electron population analysis that corresponds very closely to the Lewis picture of localized bonds and lone pairs. NBO uses natural orbitals (valence shell atomic orbitals derived from the eigenfunctions of the electron density matrix) to compute atomic charges and assign each electron to a bonding, antibonding, core, or Rydberg orbital.
The use of NBO, compared with the more traditional Mulliken population analysis (where the products of LCAO coefficients are used as a measure of the electron density assigned to each nucleus) is advantageous in several respects. Most importantly, Mulliken analysis has difficulty characterizing charge distribution and highly ionic bonds. Mulliken analysis can correct for these deficiencies with the use of a higher basis set, but that can be computationally intensive. Since NBO’s natural orbitals are already in close correspondence with the traditional Lewis structure, few corrections are needed, leading to satisfactory results at a lower basis set than would be required in Mulliken analysis.

Details of the Calculations

The Gaussian 03 software package was used for all calculations [13]. Structures were optimized to a minimum using the Berny algorithm [14]. Resulting force constants were used to calculate vibrational frequencies. All calculations were carried out using density functional theory. Specifically, the hybrid method B3LYP was used, which includes Becke’s three parameter exchange-correlation hybrid functionals and the correlation functional of Lee, Yang, and Parr [15]. The 6-31G(d,p) basis set [16] was used for all optimizations, except for CDE-011 which was also optimized using 6-311G(d,p) [17]. Given that the resulting structures showed practically no difference between the two basis sets, 6-31G(d,p) was used for subsequent calculations given its superior calculational efficiency. The NBO population analysis was carried out at the same level of theory [18].
Results and Discussion

Conformational Studies

Dr. Emal and his group have obtained data as to the activity of a series of digallate compounds as inhibitors to PAI-1 [5]. Figure 2 shows some examples of digallate compounds with various ‘linkers’ between gallate groups and their respective IC₅₀ values.

![Chemical structures of digallate compounds](image)

**Figure 2**: A series of digallate compounds of varying chain lengths and their associated IC₅₀ Clockwise from upper left, these are molecules CDE-008 (n=1), CDE-009, CDE-010/CDE-012, and CDE-011.
In this section we consider two compounds with opposite characteristics: CDE-008 has a short, flexible linker and is very active (IC$_{50}$ is 0.365), while CDE-056 has a very constrained linker and is a poorer inhibitor (IC$_{50}$ is 2.87). A comparison between these two molecules will allow us to test the validity of our approach to predicting each compound’s accessible conformations, as well as its activity.

The general computational approach we used is comprised of several steps: initially, the molecule’s structure is optimized to a minimum at the B3LYP/6-31G* level of theory. This structure is not necessarily the global minimum, rather the closest structure to the starting conformation that lies at the bottom of a potential energy well. Starting from this optimized structure, a series of relaxed scan calculations are performed. In a relaxed scan, one of the molecule’s structural variables (bond length, angle, or dihedral angle) is fixed and varied in a step-wise fashion; the remainder of the molecule is then optimized at each step and the energy calculated. We performed the scan calculations discussed here by rotating single bonds in the molecule, thus exploring all possible conformers. In practice, single bonds are rotated by varying the appropriate dihedral angles. The dihedral angles scanned for each of the inhibitors and the resulting molecular energies are available upon request.\(^1\) It is expected that only the more stable rotamers will exist in solution in significant amounts. Therefore any conformer with an energy more than 2.71 kcal/mol\(^2\) higher than the lowest conformer for a given molecule is eliminated from further consideration.

\(^1\) This file is too large to be included here.
\(^2\) This limit is based on a table of Gibbs’s free energy and equilibrium constants and it ensures that no more than 1% of the active conformations are eliminated.
Once a set of low-energy conformers is established, we need to determine any relationship between conformation and activity. To this end, we need to identify unique and relevant markers of a specific conformation so they can be related to the inhibitor’s activity. It is most likely that the phenyl rings of the gallate groups and their hydroxyl substituents will be the portion of each inhibitor interacting with PAI-1. This is borne out by the fact that such groups are at the periphery of the molecule and thus will be ‘seen’ by PAI-1 first, as well as experimental evidence indicating that two gallate groups are necessary for inhibition [5]. The structural markers we will use are the distance between the centers of the two phenyl rings of the gallate groups and their relative orientation in space. The latter is calculated by placing one ring at the center of the Cartesian coordinate system and then finding the normal vector to the plane of the other ring. Since the normal vector is defined by three components, there are four markers for each conformation. Once these four variables are calculated for each energy-accessible conformation of both inhibitors, we will plot them to determine any correlations with the activity of the inhibitors. In other words, are there values or ranges of values for these variables that occur in the active inhibitor but not in the poor inhibitor? If so, this is a strong indication that such values are necessary for inhibition and from these we can draw conclusions about the structural features that make a good inhibitor.

When the initial set of relaxed scan calculations was performed on the two inhibitors, we noted that a hydrogen bond formed between the hydroxyl group of one gallate and that of the other gallate. This bond is fairly strong and once formed persists throughout the scan calculation, causing the two rings to remain somewhat ‘tied’ as the backbone

---

3 See Appendix A for details on how the four markers are calculated.
rotates to take advantage of the stabilizing effect of the hydrogen bond. As this interaction would be prevented by the presence of solvent molecules in an aqueous environment, we added an aqueous solvent cage to the model using the polarizable continuum (PCM) model. In this method the solvent is represented as a structureless polarizable medium characterized by its dielectric constant [19]. Unfortunately, the calculations did not reach convergence in several of the steps since this particular solvent model encounters difficulties when the molecule assumes a somewhat ‘folded’ configuration and this occurs often as the bonds are rotated. To solve this problem, we decided to substitute the gallate hydroxyl oxygens with sulfur atoms; this prevents hydrogen bonding without significantly changing the structural or electronic characteristics of the molecule. Since the structural variables we will use to analyze the inhibitors are not affected by the groups at the periphery of the rings, the results will also not be affected.

There are three single bonds in CDE-008 that can affect the configuration of the molecule, as shown in Figure 3. All were rotated 360° in 20° increments.

Figure 3: Bonds rotated in the relaxed scans for CDE-008.
In the first scan calculation the linker (middle) bond was rotated starting from the initial optimized structure. See Figure 4 and Table I for a list of relative energies at each step of this scan.

![Energies Associated with CDE-008 Central Linker Rotation](image)

**Figure 4:** Plot of the values shown in Table I.

For each of the angle conformations of the linker bond that lead to a relative energy no higher than 2.71 kcal/mol test (highlighted in bold in Table I), a nested scan calculation was performed. In these nested scan calculations, the linker bond was frozen in the position indicated in Table I and the other two bonds were rotated in a way that includes all possible combinations of dihedral values (360° in 20° increments). The results of such nested scans are available upon request.4

---

4 This file is too large to be included here.
**Table I:** Relative Energies as a Function of Rotation of the Linker Bond in CDE-008. This data provides the basis for subsequent scans of adjacent dihedral angles. Energies listed are relative to the lowest energy conformation. Bolded values were used in subsequent scans.

<table>
<thead>
<tr>
<th>Value of Dihedral Angle</th>
<th>Relative Energy (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>175.3841</td>
<td>0.29</td>
</tr>
<tr>
<td>195.3841</td>
<td>1.55</td>
</tr>
<tr>
<td>215.3841</td>
<td>4.59</td>
</tr>
<tr>
<td>235.3841</td>
<td>7.68</td>
</tr>
<tr>
<td>255.3841</td>
<td>4.75</td>
</tr>
<tr>
<td>275.3841</td>
<td>2.20</td>
</tr>
<tr>
<td>295.3841</td>
<td>0.43</td>
</tr>
<tr>
<td>315.3841</td>
<td>0.44</td>
</tr>
<tr>
<td>335.3841</td>
<td>1.76</td>
</tr>
<tr>
<td>355.3841</td>
<td>2.91</td>
</tr>
<tr>
<td>375.3841</td>
<td>2.62</td>
</tr>
<tr>
<td>395.3841</td>
<td>1.50</td>
</tr>
<tr>
<td>415.3841</td>
<td>0.86</td>
</tr>
<tr>
<td>435.3841</td>
<td>1.51</td>
</tr>
<tr>
<td>455.3841</td>
<td>3.06</td>
</tr>
<tr>
<td>475.3841</td>
<td>3.96</td>
</tr>
<tr>
<td>495.3841</td>
<td>2.77</td>
</tr>
<tr>
<td>515.3841</td>
<td>0.75</td>
</tr>
<tr>
<td>535.3841</td>
<td>0.00</td>
</tr>
</tbody>
</table>
A similar set of nested scan calculations was carried out for CDE-056, which was selected as the “poor” inhibitor. Since the linker bond cannot rotate freely because of the benzene ring, scan calculations were conducted on the two free rotating bonds shown in Figure 5. All other ‘single’ bonds cannot undergo truly free rotation because of $\pi$ delocalization between the gallate phenyl rings and carboxyl groups. The list of relative energies as a function of dihedral angle values is available upon request.

**Figure 5:** Bonds rotated in the relaxed scans for CDE-056.

The next step in the data analysis is to determine the structural characteristics (markers, as defined above) for the conformers of each molecule that are within 2.71 kcal/mol of the global minimum. For each of these conformations, Cartesian coordinates were used to calculate the distance between the centers of the two gallate rings, as well as the direction of a normal vector from the center of one ring with respect to the other (See Appendix A for details).
It was hoped that this normal vector would provide insight into the relative orientation of the gallate rings when interacting with the active site of PAI-1. However, this information could only be useful if the range of distances for the good and the poor inhibitors showed significant overlap. If the possible gallate-gallate distances for the two molecules overlapped, then the range of possible relative orientation of the gallate rings could be used as the distinguishing factor between poor and good inhibitors. As shown in Figure 6, CDE-056 and CDE-008 showed no overlap of possible gallate-gallate distances and further analysis of ring orientations in these two molecules would not prove helpful. Such analysis may prove useful in other sets of molecules that show significant gallate-gallate distance overlap.

![Graph showing range of distances between rings for CDE-056 and CDE-008](image)

**Figure 6:** Range of distances between rings for CDE-056 and CDE-008

---

5 The results of calculations, distance, and normal vector values are available upon request.
All other digallate molecules under consideration have longer, more complex linkers. After a brief review, it became apparent to us that all other candidates have many more degrees of freedom than CDE-008 and CDE-056. To obtain additional data similar to that in Figure 6, conformational scans for all relevant degrees of freedom would need to be carried out and consequently this approach is extremely calculation intensive. Therefore, it would be desirable to find alternative characteristics that are also related to activity but can be calculated more quickly. To this end, we consider here the electronic characteristic of a set of inhibitors. Specifically, the goal is to use a Natural Bond Order (NBO) analysis to determine whether differences in the pKa of gallate protons are responsible for the observed variations in activity in a series of PAI-1 inhibitors. We consider it possible that differences in the pKa’s of these molecules will allow for interactions with the active site of PAI-1 that can determine the affinity of this inhibitor for the enzyme. It is possible that it is the ionized form of the inhibitor that reacts with PAI-1, and as a result, it is possible that the stability of the conjugate bases may play a major role in inhibition activity. To this end, we investigate the stability of conjugate bases of a variety of inhibitors of varying effectiveness, seeking trends that can tell us something useful about the qualities of effective inhibitors.

We begin by determining the optimized structures for the compounds of interest: CDE-008, CDE-011, CDE-127, CDE-010, and CDE-013. These are shown as line diagrams in Figure 7 and as ball and stick models in Figures 8-12.
Figure 7: The structure of digallate compounds considered in this work [2]. The compounds are labeled as follows: (a) CDE-008, (b) CDE-010, (c) CDE-011, (d) CDE-013, and (e) CDE-127.
Figure 8: CDE-010 optimized structure at the B3LYP/6-311G(d,p) level of theory.

Figure 9: CDE-127 optimized structure at the B3LYP/6-311G(d,p) level of theory.
Figure 10: CDE-013 optimized structure at the B3LYP/6-311G(d,p) level of theory.

Figure 11: CDE-011 optimized structure at the B3LYP/6-311G(d,p) level of theory.
However, we know from the results detailed above that these compounds have many low energy conformations that can be expected to be present in solution. Therefore, the particular conformations shown in Figures 8-12 above are not necessarily representative of the molecule’s structural characteristics. In fact, an analysis of the distance between gallate rings reveals no direct correlation with activity, as shown in Table II.

**Table II**: Optimized distances between gallate rings and activities for the five inhibitors under examination.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Distances</th>
<th>IC$_{50}$ value (µM, a measure of activity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDE-011</td>
<td>20.83</td>
<td>338</td>
</tr>
<tr>
<td>CDE-010</td>
<td>15.58</td>
<td>196</td>
</tr>
<tr>
<td>CDE-013</td>
<td>6.42</td>
<td>9.64</td>
</tr>
<tr>
<td>CDE-127</td>
<td>13.13</td>
<td>7.88</td>
</tr>
<tr>
<td>CDE-008</td>
<td>7.78</td>
<td>0.558</td>
</tr>
</tbody>
</table>
Acidity of phenolic hydrogens

We now turn our attention to electronic characteristics: specifically, the electron density distribution around the hydroxyl protons on the gallate rings. We conducted an NBO population analysis of CDE-008, CDE-011, CDE-127, CDE-013, and CDE-010 and found that partial charges on potentially acidic protons were very similar. Across all molecules studied, partial charges were between 0.46 and 0.49. These values do not vary enough to be significant predictors of PAI-1 effectiveness. We also compared the electron density distribution of the molecules by generating electrostatic potential surfaces mapped onto the electron density distribution. The resulting surfaces are shown in Figures 13-17. Although certain hydrogen atoms clearly carry more positive partial charge, as evidenced by the degree of “blueness,” no clear pattern that can be related to activity is apparent.

Figure 13: Electrostatic potential mapped onto total density for CDE-127. Blue regions indicate positive charge and red regions indicate negative charge. Yellow and green colors indicate neutral regions.
**Figure 14:** Electrostatic potential mapped onto total density for CDE-013. Blue regions indicate positive charge and red regions indicate negative charge. Yellow and green colors indicate neutral regions.

**Figure 15:** Electrostatic potential mapped onto total density for CDE-010. Blue regions indicate positive charge and red regions indicate negative charge. Yellow and green colors indicate neutral regions.
Figure 16: Electrostatic potential mapped onto total density for CDE-011. Blue regions indicate positive charge and red regions indicate negative charge. Yellow and green colors indicate neutral regions.

Figure 17: Electrostatic potential mapped onto total density for CDE-008. Blue regions indicate positive charge and red regions indicate negative charge. Yellow and green colors indicate neutral regions.
Finally, we turn our attention to the stability of the conjugate bases that would be formed by ionization of the gallate protons. Stability of the resulting conjugate base, in combination with partial charge of the acidic proton, should give an indication of the relative pKa of such protons. We removed each hydroxyl proton on the gallate rings in turn and optimized the structure of the resulting anion. We carried out this analysis on CDE-008 and CDE-011 only, since we wanted to determine whether this was a valid approach and these two molecules exhibit the largest difference in activity. Since CDE-008 is symmetrical with respect to the two gallate rings, for this molecule we remove only the protons on one of the two rings. The relative energies of the optimized anions are displayed in Tables III and IV.

**Table III:** Total energy and relative stability of conjugate bases for CDE-008.

<table>
<thead>
<tr>
<th>Ionized proton</th>
<th>Energy (a.u.)</th>
<th>Relative Energy (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>-1370.190294</td>
<td>16.1258</td>
</tr>
<tr>
<td>38</td>
<td>-1370.215725</td>
<td>0.1657</td>
</tr>
<tr>
<td>40</td>
<td>-1370.215989</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**Table IV:** Total energy and relative stability of conjugate bases for CDE-011.

<table>
<thead>
<tr>
<th>Ionized proton</th>
<th>Energy (a.u.)</th>
<th>Relative Energy (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>-1831.842765</td>
<td>0.0000</td>
</tr>
<tr>
<td>55</td>
<td>-1831.842124</td>
<td>0.4020</td>
</tr>
<tr>
<td>59</td>
<td>-1831.841786</td>
<td>0.6139</td>
</tr>
<tr>
<td>61</td>
<td>-1831.824101</td>
<td>11.7136</td>
</tr>
<tr>
<td>57</td>
<td>-1831.817154</td>
<td>16.0732</td>
</tr>
<tr>
<td>51</td>
<td>-1831.817581</td>
<td>15.8055</td>
</tr>
</tbody>
</table>
The acidic hydrogens under consideration and the relative stability of the respective conjugate bases are shown in Figures 18 and 19.

**Figure 18:** Relative stability of conjugate bases of CDE-008.

**Figure 19:** Relative stability of conjugate bases of CDE-011.
In CDE-008 the *ortho* hydroxyl group on the same side as the carbonyl of the ester group appears to be much less likely to ionize than the other two hydroxyl groups. However, in CDE-011 the two rings show a different pattern of stability: in the first ring one hydroxyl group is less likely to ionize than the other two (*the ortho* hydroxyl group on the opposite side from the carbonyl of the ester group) while in the second ring both *ortho* hydroxyls are less likely to ionize. This may suggest that significant differences in the pKa of gallate protons could be related to activity.
Conclusions

The chief result of the conformational studies is that the distance between gallate rings in all low-energy conformations of CDE-056 (the poor inhibitor) spans the range between 5 and 7Å, while for CDE-008 (the good inhibitor) that range is between 8 and 16Å. This suggests that the distance between gallate phenyl rings may correlate with inhibitor effectiveness and may provide some hints about the structure of the PAI-1 active site. Following this line of reasoning, CDE-056 is a poor inhibitor because its ‘arms’ cannot open to be at least 8Å apart. Analysis of similar compounds, with a variety of linkers will be necessary to further support this hypothesis.

Five digallate compounds with varying inhibitory activity toward PAI-1 have also been studied with respect to acidity of gallate hydroxyl hydrogens. While the partial charges of the acidic protons are very similar within each compound and across all molecules, we observe unexpected differences in the pattern of stability of conjugate bases between CDE-008 and CDE-011. In addition, we confirm the results of earlier work indicating that each compound has a large number of conformations of similar energy that are likely to exist in solution and therefore a single optimized structure gives no useful information in terms of potential activity.
References


Appendix A

Summary of Data Analysis Calculations

Finding the center of one gallate ring, called "Ring O."

\[
X_0 = \frac{X_4 + X_1}{2} \quad Y_0 = \frac{Y_4 + Y_1}{2} \quad Z_0 = \frac{Z_4 + Z_1}{2}
\]

Subtracting \(X_0\), \(Y_0\), and \(Z_0\) from the coordinates of every atom in the molecule, the molecule is moved such that Ring O is centered at the origin. From this position, the center of the other gallate, "Ring A" is calculated, using similar methods as above. The distance between the two rings can then easily be calculated using familiar methods, shown below.

\[
Distance = \sqrt{(X_0 - X_A)^2 + (Y_0 - Y_A)^2 + (Z_0 - Z_A)^2}
\]

In order to calculate the normal vector, three atoms in Ring A are selected; In this case, C1, C4, and C5 are chosen. The two vectors in the plane of Ring A, are as follows.

\[
u =<(X_4 - X_1), (Y_4 - Y_1), (Z_4 - Z_1)>
\]

\[
v =<(X_5 - X_1), (Y_5 - Y_1), (Z_5 - Z_1)>
\]

\[
u \times v = \begin{bmatrix} i & j & k \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{bmatrix} = i * \begin{bmatrix} u_y & u_z \\ v_y & v_z \end{bmatrix} - j * \begin{bmatrix} u_x & u_z \\ v_x & v_z \end{bmatrix} + k * \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix}
\]

\[
u \times v = i(u_y v_z - u_z v_y) - j(u_x v_z - u_z v_x) + k(u_x v_y - u_y v_x)
\]

This is the normal vector.