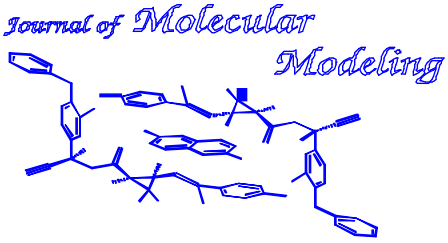
J. Mol. Model. **1996**, *2*, 390 – 398



Anionic Tetrahedral Complexes as Serine Protease Inhibitors

Michael Shokhen[‡]* and Dorit Arad[$]*

The George S. Wise Faculty of Life Sciences, Tel-Aviv University, Ramat Aviv, Israel 39040, Tel: 972-3-6408723, FAX: 972-3-6409407

[‡] Department of Biochemistry (shokhen@etgar.tau.ac.il)

[$] Department of Molecular Biology and Biotechnology (dorit@argaman.tau.ac.il)

Received: 4 April 1996 / Accepted: 10 September 1996 / Published: 11 October 1996

Abstract

Potent inhibitors of proteases are constantly sought because of their potential as new therapeutic lead compounds. In this paper we report a simple computational methodology for obtaining new ideas for functional groups that may act as effective inhibitors. We relate this study to serine proteases. We have analyzed all of the factors that operate in the enzyme–substrate interactions and govern the free energy for the transformation of the Michaelis complex (*MC)* to the anionic covalent tetrahedral complex (*TC)*. The free energy of this transformation ($\Delta\Delta G_{MC\text{-}TC}$) is the quantitative criterion that differentiates between the catalytic and inhibitory processes in proteases. The catalytic *TC* is shifted upwards ($\Delta\Delta G_{MC\text{-}TC} > 0$) relative to the *MC* in the free energy profile of the reaction, whereas the inhibitory tetrahedral species is shifted downward ($\Delta\Delta G_{MC\text{-}TC} < 0$). Therefore, the more stable the *TC*, the more effective it should be as an inhibitor. We conclude that the dominant contribution to the superstabilization of an anionic *TC* for transition state analog inhibitors originates from the formation of a σ-covalent bond between the reactive centers of the enzyme and its inhibitor. This energetic effect is a quantitative value obtained in *ab initio* calculations and provides an estimate as to whether a functional group is feasible as potent inhibitor or not. To support our methodology, we describe several examples where good agreement is shown between modeled *ab initio* quantum chemical calculations and experimental results extracted from the literature.

Keywords: Protease, methodology of inhibitor design, MO ab initio calculations, free energy profile.

Introduction

Serine proteases play an important role in regulating a wide variety of biological activities. Specific protease inhibitors thus serve as targets for many therapeutic applications [1-3]. Transition state analog inhibitors were first introduced by Wolfenden [4], who suggested that stable analogs of the transition state (*TS)* structure of any enzyme-substrate catalytic process should inhibit enzyme activity. The TS analog concept has been widely used over the last two decades for designing *TS*-analog protease inhibitors [5-8]. The *TS*-ana-

* *To whom correspondence should be addressed*

log inhibitor is a chemical structure that contains a functional group (for example aldehyde or ketone), which, in contrast to the native substrate, cannot be transformed further by the chemical machinery of the enzyme. The common feature of *TS*-analog inhibitors is that they mimic the shape of the native substrate in its *TS* state [5-8]. In this paper we analyze a specific type of TS analogs in which the inhibitory effect results from the intrinsic stability of the covalent complex, forming a stable charged species, rather than an intermediate on the reaction potential surface.

The first step in hydrolysis catalyzed by serine proteases is the conversion of the non-covalent enzyme-substrate (E-S) Michaelis complex (*MC*) to an anionic covalent tetrahedral complex (*TC*) – $MC \Rightarrow TC$ [9], where the new covalent bond is formed between the attacking nucleophilic atom in the active site of the enzyme and the electrophilic center of the substrate. On the reaction potential surface of enzymatic catalysis, the *TC* is located in a very shallow local minimum, close to the saddle point corresponding to the *TS* [10-14] and cannot accumulate [15]. Consequently, according to Hammond's postulate [16], the structure of the *TC* in the reaction with a native substrate should be close to that of the *TS*. The reactive center of native substrates of proteases contains a carbonyl group. We suggest that an appropriate structural variation of the reaction center of a native substrate can lead to an anionic superstable *TC*, which will be lower in energy than the *MC* or the products. Thus, an anionic inhibitor is located in the bottom of a potential well rather than on the top of the hill on the free energy profile of enzymatic reaction path. This type of inhibition implements thermodynamic control on the enzymatic catalysis. Anionic *TC*s have been detected experimentally by direct NMR measurements [17, 18]. Stabilization of the tetrahedral intermediate has been established to be an important factor for increasing the in-vitro potency of inhibitors [18]. Substitution of the location alpha to the reaction center carbonyl with electron withdrawing substituents as fluorine, trifluormethyl and heterocyclic substituted aromatic rings, contribute to *TS* stabilization, and derivatives bearing these substituents are found to be potent inhibitors [18]. The effect of substituents on chemical reactivity in general is also highlighted in a number of papers [19-22], and good correlation is found between calculated values (semi-empirical and *ab initio*) and experimental values. However, a theory that can analyze and quantify the experimental values and build an easy to follow general concept for obtaining new derivatives that may provide novel charged superstable inhibitors is missing. We now report one approach to such a theory.

Which forces contribute to the formation of a superstable *TC* in inhibitory processes in contrast to the unstable *TS* formed during the catalysis of native substrates? Wolfenden's theory [4], predicting the principal ability of the *TS*-analog inhibitors neither assumes the specific nature of the attractive forces involved nor requires that the enzyme be rigid or flexible. Opinions in the literature that favor the notion that in enzymatic catalysis noncovalent forces are responsible for stabilizing either the *TS* or the *TC* arose from the following concepts: (a) formation of hydrogen bonds with the oxyanion hole (reviewed by Ménard, & Storer [23]); (b) electrostatic stabilization [18]; (c) environmental effects of the active site [12,25,26]; and (d) hydrophobic effects [27].

In the present work we have estimated the relative contributions from different effects (proton transfer, hydrogen bond formation and etc.) accompanying the $MC \Rightarrow TC$ transformation to the total free energy of this process. By analyzing factors that control the $MC \Rightarrow TC$ step in catalysis, we conclude that the formation of a new covalent σ-bond during the nucleophilic attack of the protease nucleophilic center on the carbonyl group or its analog makes the main contribution to the energy stabilization of the *TC*. The superstabilization

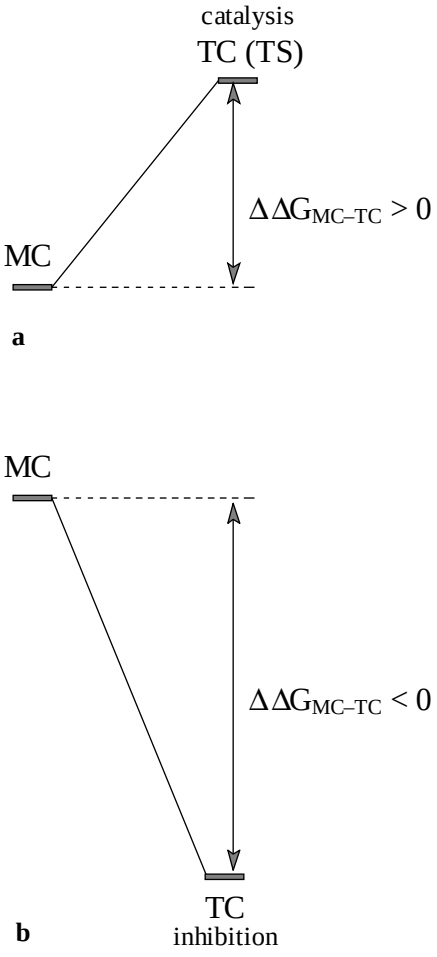


Figure 1. *(a) For a normal catalytic process, a tetrahedral complex (TC) is the transition state–the saddle point on a reaction free-energy profile. The TC is shifted upward from the energy of the Michaelis complex (MC).*
(b) For the inhibition of proteases, the TC is a thermodynamically very stable structure corresponding to the potential minimum on the reaction pass, which is shifted downward from the energy of the MC.

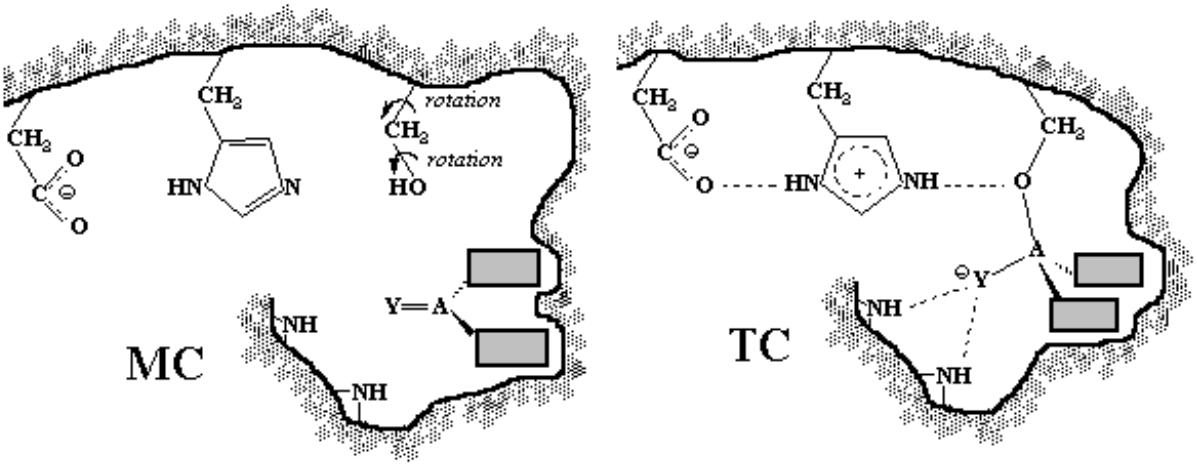


Figure 2. *Generalized schematic presentation of the structural change in the active site of a serine protease during the conversion of MC ⇒ TC. The rectangles designate the subsites of a substrate (inhibitor) belonging to the recognition site (RS) which are bound to the enzyme by noncovalent interactions. The two NH groups depict an oxyanion hole. Formation of the new covalent bond, O-A, in a TC between the reactive centers results in the creation of ring skeleton, which contains a Ser residue, an inhibitor molecule, and part of the backbone of the enzyme. Hydrogen bonds in the TC are depicted by dotted lines to stress that they are more rigid than in the MC. Additional binding between the enzyme and the inhibitor results in a loss of degrees of internal rotation.*

of the *TC* depends on the strength of the new forming σ-bond. This contribution can easily be estimated by model calculations, and thus provide a tool for a semi-quantitative choice of substituents needed in order to stabilize the *TC*. We demonstrate that simple routine quantum-chemical calculations of a small model system–the analog of the *TC* can be applied successfully for the design of new classes of protease inhibitors. Using examples of known inhibitors from the literature, we show that the effect of inhibition results in the formation of the superstable anionic *TC*. Our model calculations on these systems show excellent agreement between the predicted values and the experimental results.

Models and Methods

Structural analysis

In a normal catalytic pathway that corresponds to the nucleophilic attack of the enzyme nucleophile on the substrate carbonyl, which already locates as the *MC* in the active site, the *TC* lies on a hill higher in energy than the *MC*. If we succeed in stabilizing the *TC* to such an extent that it lies in a potential energy well lower than the *MC*, we achieve inhibition.

We define an unstable *TC* as the *TS* corresponding to the saddle point on a normal proteolytic reaction path. Conversely, we define here an inhibitory tetrahedral species as a thermodynamically stable enzyme-inhibitor complex that corresponds to the potential minimum on the reaction path. The free energy effect of the transformation $MC \Rightarrow TC$, $\Delta\Delta G_{MC\text{-}TC}$, is the quantitative criterion that differentiates between the catalytic and inhibitory processes. The catalytic *TC* is shifted upwards ($\Delta\Delta G_{MC\text{-}TC} > 0$) relative to the *MC* in the energy profile of the reaction (Fig. 1a), whereas the inhibitory tetrahedral species is shifted downward ($\Delta\Delta G_{MC\text{-}TC} < 0$), see Fig. 1b. Therefore, the more stable the *TC*, the more effective it should be as an inhibitor.

We limit our consideration of E-S interactions to the region of enzyme active site (*AS*) only, so that the Gibbs free-energy effect of the $MC \Rightarrow TC$ transformation can be approximated by the value of $\Delta\Delta G(AS)$ only.

The *AS* is divided into two structural regions: the catalytic (reactive) site (*CS*), where the chemical interaction between the enzyme and substrate occurs, accompanied by the formation and cleavage of covalent bonds, and the recognition site (*RS*), which binds the substrate by non-covalent interactions. No chemical transformations occur in the *RS*. Usually the *CS* fragment is much smaller than the *RS*.

$$AS = RS \cup CS \qquad (1)$$

The model introduced here for the subdivision of an E-S complex into its component parts is a trivial application of the approach that is widely used in molecular mechanics and quantum chemistry-structural-additivity analysis of molecular total energy**.** We assume that the *RS* fragment of an E-S complex conserves its 3D molecular structure in the $MC \Rightarrow TC$ path: $\Delta G_{MC}(RS) \approx \Delta G_{TC}(RS)$. Hence, we can use Eq. 1, to rewrite the expression for the Gibbs free-energy difference between the $MC \Rightarrow TC$ [$\Delta\Delta G(AS)$] for a reaction profile of the same substrate (or inhibitor):

$$\Delta\Delta G(AS) = \Delta G_{TC}(AS) - \Delta G_{MC}(AS) \approx \Delta\Delta G(CS) =$$

$= \Delta\Delta H(CS) - T\Delta\Delta S_{vib}\ (CS)$ (2)

where we have neglected the contribution from the *RS* region to $\Delta\Delta G(AS)$ in Eq. 3. The overall translational and rotational entropy, $T\Delta\Delta S\ (AS)_{tr,rot}$, does not contribute to $\Delta\Delta G(AS)$ because in both the *MC* and the *TC* states the substrate is bound to the enzyme in one complex. $S\ (AS)_{tr,rot}$ is determined by the mass and geometrical parameters of the complex [28], where the enzyme, which does not change its shape in the $MC \Rightarrow TC$ transformation, dominates, so that $\Delta\Delta S(AS)_{tr,rot} = 0$.

The E-S interactions in the *RS* ("recognition site") provide the driving force for aligning the substrate (or inhibitor) in the active site of an enzyme, which strongly influences the efficacy of enzymatic catalysis or inhibition. The chemical transformation of the substrate (or inhibitor) in the reaction path $MC \Rightarrow TC$ can proceed in the initial stage of the *MC* only if all the atoms of the CS are positioned in very specific, optimal interatomic distances for the desired chemical process. The *RS* fragment contributes the dominant part of the noncovalent binding between the enzyme and the substrate. The alignment can be characterized quantitatively by the values of $\Delta G_{TC}\ (AS)$ and $\Delta G_{MC}\ (AS)$. Although the noncovalent E-S interactions in the *RS* strongly influence the values of $\Delta G_{TC}\ (AS)$ and $\Delta G_{MC}\ (AS)$, according to Eq. 2, such influence is canceled in the value of their Gibbs free-energy difference, $\Delta\Delta G(AS)$: $\Delta\Delta G(RS) \approx 0$. Thus, we conclude that the free energy contribution from the $CS\ [\ \Delta\Delta G(CS)]$ determines the overall difference between the E-S binding energy during the transformation , the noncovalent E-S interactions in the *RS* play a minor role here.

In a related paper, Menger suggested the "Split-Site" model [29], where an active site is divided into a "binding region" and a "reaction region". The theoretical findings of Menger [29] are based on his idea that only *destabilizing* interactions between the enzyme and the substrate occur in the *CS* region (Menger's reactive site). Menger also accepted that the 3D structure of the "binding site" (analogous to the *RS* in our analysis) remains constant during enzymatic transformations. He postulated that (a) in the reactive site (*CS*), destabilization of the *TS* is much stronger than that of the *MC* and (b) in the binding site (*RS*), only noncovalent interactions can stabilize E-S complexes. We believe that covalent binding in the *CS* contribute to the stabilization effect. The quantitative picture is demonstrated by calculations.

Free energy components in the CS

In this section we discuss the contribution to $\Delta\Delta G(CS)$ from different processes accompanying the $MC \Rightarrow TC$ transformation. For simplicity we designate $\Delta G(CS)$ instead of $\Delta\Delta G(CS)$.

Reactivity center. The dominant process that occurs in the transformation of $MC \Rightarrow TC$ is the nucleophilic attack of the catalytic site of the enzyme on the electrophilic center of a substrate or an inhibitor occurring in the *CS*. In this analysis we separate a fragment in the *CS* – the "reactivity center", defined as the fragment that contains the reactive centers of the enzyme and substrate and the atoms of their valent surrounding. Most structural changes of the substrate in the reaction step are located in the reactivity center. ΔG_{rc} expresses the contribution of the reactivity center to the value of $\Delta\Delta G(CS)$.

The electrophilic center of the native substrate is either an amide or an ester carbonyl. The object of the nucleophilic attack in an inhibitor is a polarized double bond of the electrophilic site, *A=Y*, where *A* is an electrophilic center and *Y* is usually an electronegative substituent [6, 30]. The nucleophile atom *D* (*D* = *O* or *S*) forms the new bond *D-A*, which is partially covalent and partially ionic [31]. The transformation is accompanied by an $sp_2 \Rightarrow sp_3$ rehybridization, which causes pyramidalization at the A center and *A=Y* bond elongation. The ΔG_{rc} builds up mainly from a considerable stabilization energy, which results from the formation of a new covalent σ-bond D-A, (see Fig. 2). This strong stabilization effect is partially compensated by the destabilizing energy resulting from reducing the π character of the A=Y bond to a single bond. The net result of the transformation, however, is always $\Delta G_{rc} < 0$.

Proton Transfer. Serine proteases facilitate nucleophilic attack on a substrate by transferring a proton from the nucleophile [9]. The proton transfer occurs simultaneously with *TC* formation in reactions catalyzed by serine proteases. The process contributes [designated as $\Delta G_{pt}\ (CS)$] $\Delta G(CS)$, of the transformation to the total free energy effect. Warshel & Russell [12], using empirical procedures, estimated that the energy change involved in the proton transfer Ser-His in trypsin is 14 kcal/mol, so $\Delta G_{pt}\ (CS) > 0$.

Environmental Effects. It has been established that the main source of *TS* stabilization lies in the environmental effects in the active site of the enzyme (designated as $\Delta G_{env}\ (CS)$). The oxyanion binding site in serine proteases is one example of a widely studied environmental factor [23]. Such factors are often considered the main sources of anionic *TC* stabilization [32,33]. The structure of chymotrypsin-trifluoromethylketone inhibitor complexes provides an experimental demonstration of the contribution of environmental effects [34]. The negatively charged oxygen atom attached to the tetrahedral carbon of a hemiketal adduct (*TC*) is hydrogen-bonded to Ser195 and Gly193 amides in the oxyanion hole. In the oxyanion hole, the hydrogen bonding for a negatively charged oxygen atom in the ionized hemiketal in the *TC* is stronger than the carbonyl oxygen binding of the neutral *MC* [12, 35,36]. The superstabilization of the anionic *TC* relative to the *MC* is ascribed to this differential binding strength.

McMurray, and Dyckeys [37] studied trypsin inhibition using the series of model peptide ketones Lys-Ala-LysCH_2X. From a Hammett plot of -log K_i vs. σ_I; the authors concluded that the strength of binding of the hemiketals to the AS of the enzyme increases with the electron-withdrawing

ability of the varied substituent, X. The tightest binding was determined for the fluoromethylketone X = F. In our analysis, *ab initio* calculations were used to obtain a quantitative estimate on the energetic contribution of this effect. We considered the small molecules $(CF_3)HCO$ and $(NH_2)HCO$ to compare the CF_3 substituent with NH_2, which is the model for a native amide substrate. To simulate the hydrogen bonds in the oxyanion hole, we used one water molecule**.** The energy of hydrogen bonds was calculated by Gaussian 92 [38] at the 6-31+G*/3-21G level according to the following reaction equations:

MC: $(NH_2)HCO + HOH \Rightarrow (NH_2)HCO$ —HOH,
E_{MC} (NH_2) = -4.0 kcal/mol (3a)

TC: $HO(NH_2)HCO^- + HOH \Rightarrow HO(NH_2)HCO^-$–HOH,
E_{TC} (NH_2) = -12.4 kcal/mol (3b)

MC: $(CF_3)HCO + HOH \Rightarrow (CF_3)HCO$—HOH,
E_{MC} (CF_3) = -2.6 kcal/mol (3c)

TC: $HO(CF_3)HCO^- + HOH \Rightarrow HO(CF_3)HCO^-$—HOH,
E_{TC} (CF_3) = -14.7 kcal/mol (3d)

Our simple model for hydrogen bonds in the oxyanion hole provides an estimate of tighter binding for the *TC*:

$$\Delta E(NH_2) = E_{TC}(NH_2) - E_{MC}(NH_2) = -8.4 \text{ kcal/mol} \quad (4a)$$

$$\Delta E(CF_3) = E_{TC}(CF_3) - E_{MC}(CF_3) = -12.1 \text{ kcal/mol} \quad (4b)$$

$$\Delta\Delta E = \Delta E(CF_3) - \Delta E(NH_2) = -3.7 \text{ kcal/mol} \quad (4c)$$

The values are quantitatively close to experimental values [36] and to the theoretical estimations of Warshel et al. [12, 39, 40]. We can conclude that the hydrogen bonds in the oxyanion hole contribute not only to the additional stabilization of a *TC* but also to the superstabilization of a *TC* when *TS*-analog inhibitors with a strong electron-withdrawal substituent are involved. The value of the effect is small, however, when compared to the gain in energy that is due to the formation of a covalent bond in a *TC* (see the "Validity" section in Results and Discussion below).

Hwang and Warshel [39] proposed that electrostatic effects are key factors in serine protease catalysis. Enzymes provide the proper environment of polar-group dipoles to complement the changes in charge distribution from the *MC* to the *TS* [12,39]. The estimated energy value of *TS* stabilization is -7 kcal/mol [12]. Another theory claims that in water solution, substrate-enzyme binding is driven by a hydrophobic effect [27]. The most reasonable conclusion is that the environmental effects cause the *TC* to be more stable than the *MC*[ΔG_{env} *(CS)* < *0*].

Vibrational Entropy. The last component of free energy considered here is the contribution of vibrational entropy ΔS_{vib} *(CS)*. In a *TC* the nucleophilic center of the enzyme is covalently bound to the electrophile. Topologically, the formation of this bond may be considered a ring closure (see Fig. 2), meaning a loss of at least two degrees of internal rotation. The entropy difference between linear and cyclic compounds gives an entropy loss of 2.7 to 4.3 kcal/mol per internal rotation [41]. In addition, the skeleton of hydrogen bonds should be more rigid in the *TC* than in the *MC*. Such phenomena generally reduce the integral vibrational entropy of the *TC* relative to the *MC*. A quantitative estimate of the value of ΔS_{vib} *(CS)* can be calculated only by a sophisticated computational procedure because the system under consideration is extremely complex. For a qualitative picture, however, we may accept that ΔS_{vib} *(CS)* < *0*.

In summary, we can write the following qualitative expression reflecting the additivity of all the components of the main contributors to the $\Delta G(CS)$:

$$\Delta G(CS) = [\Delta G_{rc} < 0] - [T\Delta S_{vib}(CS) < 0] +$$

$$+ [\Delta G_{pt}(CS) > 0] + [\Delta G_{env}(CS) < 0] + Rest \quad (5)$$

where *Rest* = the free energy contribution from all other effects. We assume that the *Rest* effects are minor and principally do not change the general energetic picture.

Results and Discussion

Equation 5 provides the guideline for design of superstable anionic *TC*s because it summarizes the trends of influence of the main factors governing the thermodynamic stability of the *TC* in the *CS* region of the E-S complex. We concluded that the first term (ΔG_{rc}), which relates to the formation of a covalent σ-bond (*D-A*), should be the dominant component of $\Delta G(CS)$. The results of our computations in the following section support this conclusion. Actually, we only have one tool to design an effective inhibitor forming the superstable *TC*–the variation of the molecular structure of the desired candidate. Therefore, one must query which of the possible contributors to $\Delta G(CS)$ (reactivity center interactions, environmental effects, proton transfer, and vibrational entropy) are the most susceptible to these variations. The answer is that the variation of the *CS* fragment of a substrate influences only ΔG_{rc}, and the other terms in Eq. 5 are not relevant. Indeed, any change in ΔS_{vib} *(CS)* relating to the loss of degrees of internal rotation on the amino-acid residues of the enzyme is specific to a concrete type of enzyme (Fig. 2). Basic catalysis that accelerates reactivity of its nucleophilic group is also an intrinsic feature of an enzyme (for example the catalytic triad in serine proteases). Because catalysis proceeds between nonvariable structural fragments (the amino-acid residues of enzyme), the value of ΔG_{pt} *(CS)* is independent of substrate variation. On the other hand, the value

of ΔG_{env} *(CS)* depends on the variation in the molecular structure of the substrate. Because ΔG_{env} *(CS)* involves only noncovalent interactions, which are one or two orders of magnitude lower than the energy contribution from covalent forces, ΔG_{rc} makes the dominant contribution to $\Delta G(CS)$. We conclude, therefore, that the formation of a new, covalent D-A bond between the reactive centers of the enzymatic nucleophile and its substrate (inhibitor) is the main source of the tighter binding of a *TC* compared with a *MC*.

Methodology of design

The results of the thermodynamic analysis summarized in Eq. 5 and the refinement of the most susceptible region for the variations in the inhibitor structure lead us to the conclusion that the reaction center, which is a very small fragment of the real *TC*, can be used to model the relative stability between *MC* and *TC*.

The dominant structural factor – the new covalent bond formed during the $MC \Rightarrow TC$ transformation determines the choice of the computational method. This property–the formation of a covalent bond can be separated from the rest of the factors, and followed quantitatively. Only quantum chemical calculations can correctly take into consideration the covalent interactions and the electronic effects of substituents on the stability of the *TC*. The system is small enough to allow the use of *ab-initio* quantum mechanical calculations with large basis sets.

The design of the most potent inhibitor in a series for any serine protease can be reformulated as the task of the choice in a series of inhibitors. According to our model, a potential protease inhibitor I_k is simulated by the simplest molecule A_k, which contains the reactivity center of the I_k. The enzyme's nucleophilic site is also simulated by the simplest molecule or anion *D*.

We take into account only the values of ΔG_{rc} *(k)* for each inhibitor I_k. The modeled reaction scheme of $MC \Rightarrow TC$ can be presented as follows:

$$D + A_k \Rightarrow DA_k \qquad (6)$$

For the nucleophile *D* conserved in the series, the stability of the DA_k depends only on the electronic features of the atom of the electrophilic center in A_k and its covalently bound substituents. Because we calculate the actual total energies of molecules, we calculate the values of ΔE_{rc} *(k)* rather than those of ΔG_{rc} *(k)*. This approach is based on the suggestion of Dewar [42] that the variation in the entropy component of the free energy of reaction for a reaction series of analogous reagents should be insignificant. When designing potential inhibitors, the species A_k that form the most stable DA_k in the series should be the optimal candidates to construct the reactivity center of inhibitors.

Validity of the model

In this section we shall present examples to prove our model for the design of functional groups that can act as serine protease inhibitors. Our guideline was to design small molecules that mimic the reactivity centers of virtual substrates and inhibitors. All examples were selected from the literature to show that we can predict experimentally observed tendencies for the influence of the substituents on the catalytic/inhibition constants. Quantum-chemical calculations have shown that neutral nucleophiles as H_2O and CH_3OH cannot mediate nucleophilic addition to the carbonyl group without supporting basic or acidic catalysis [43-45]. Therefore, OH^- anion was used as the simplest model for the serine nucleophile of the enzyme. Consequently, all the modeled products of nucleophilic addition are anions, as illustrated by reaction schemes in Figs. 1 to 3.

We examined the following two types of carbonyl group derivatives: (1) variation of substituents at the conserved electrophilic center, and (2) variation of the atom of the electrophilic center. The calculated stabilities of the anionic *TC*s are presented in Table 1. All quantum mechanical calculations were performed using the standard Gaussian 92 program [38]. Our general position was to simplify the computational procedure, keeping in the same time the level of accuracy of the molecular calculations that provides reproduction of the principal chemical picture of the observed experimental reaction and structural features of reagents and products. This route is convenient for practical application for computer assisted drug design, where a large series of lead compounds needs to be calculated in a reasonable time. In our computational experiments we found that semi-empirical quantum chemical procedures like AM1 or PM3 (see the review article by Stewart about these methods [46]) failed to reproduce geometries and energies of anionic *TC* structures. In some cases, especially for the anionic *TCs* containing the second row elements, semi-empirical methods gave unrealistic geometries and energies for the $MC \Rightarrow TC$ transformation. For example, the high level *ab initio* calculations demonstrated the inability of the HS^- anion to form *TC* with formaldehyde [31, 47]. In contrast, semi-empirical procedures predicted the formation of a stable *TC*. Thus, we used the non expensive variant of *ab initio* calculations, which reproduce reasonable molecular geometries and avoids basis set superposition error (BSSE) in the reaction energy estimations [48]. Molecular geometries were optimized in the 3-21G* basis set and then total energies were recalculated in the basis set 6-31+G*. This is a commonly used computational approach that gives good results for anionic species [49].

The first set of calculations include a series of carbonyl derivatives where the substituents examined have different electron-donor and electron-acceptor properties. Figure 3 shows the list of varied substituents and the general scheme of *TC* formation for this series.

Table 1. *Ab initio 6-31+G*/3-21G* calculated stabilization energies E_{st} of Tetrahedral Complexes*

Substrate	E_{st} kcal/mol
$H(NH_2)C=O$	-22.7
H(OH)C=O	-32.6
$H(CF_3)C=O$	-59.9
H(F)C=O	-50.5
H(CN)C=O	-62.2
$HB(OH)_2$	-53.2
$F(OH)_2P=O$	-61.7 [a]

[a] *The E_{st} for the $F(OH)_2P=O$ was estimated from the reaction equation:*
$HO^- + F(OH)_2P=O \Rightarrow (HO)_2PO_2^- + HF$

The first example provides a detailed analysis of standard free energy for a-chymotrypsin-catalyzed hydrolysis of N-acetyl-L-tryptophan methyl ester and N-acetyl-L-tryptophan amide. Bender et al. [50] showed that the activation barrier for the acylation step measured from the *MC* is 13.5 kcal/mol for the ester substrate and 19.6 kcal/mol for the amide substrate. For the deacylation step, the activation barrier of 17.8 kcal/mol is the same for both cases because their acylenzymes are identical. The physical meaning is that the rate-determining step is acylation for hydrolyzing amides and deacylation for hydrolyzing esters. The free energy of stabilization of the *MC* compared with that of the separated enzyme and substrate is similar for both substrates:

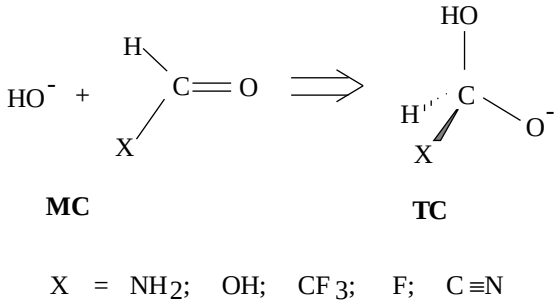


Figure 3. *Scheme of TC formation by the series of carbonyl derivatives where the varied substituents have different electron-donor and electron-acceptor properties.*

-4 kcal/mol for the ester and -3 kcal/mol for the amide. Hence, the molecular structure of the acylation step *TS* is about 6 kcal/mol lower in energy for the ester than for the amide, which causes the barrier difference. The calculations agree well with the experimental result: the model *TC* is about 10 kcal/mol less stable for an NH_2 substituent than for an OH substituent (see Table 1 and Figure 3).

The second example relates to the polyfluoroketones. That these moieties are strong *TS*-analog inhibitors for many members of the serine hydrolase class is well established [32, 33, 51]. The new bond forming between the electrophilic center of a carbonyl substrate and an attacking nucleophile in a *TC* should be more sensitive to the variation of the substituents than the hydrogen bonds discussed previously. In the model calculations the molecules CF_3CHO and NH_2CHO that mimic the reactivity center of substrates were subjected to nucleophilic addition of an HO^- anion. The reaction scheme presented in Figure 3 and the estimated values of the *TC* stability shown in Table 1 demonstrate that when compared with the NH_2 substituent, the CF_3 substituent superstabilizes the *TC* by -37.2 kcal/mol. This value is higher by one order-of-magnitude than the energetic effect of hydrogen bonds (environmental effect) considered above (see Eq. 4c). Therefore, the results confirm our hypothesis that the energetic contribution from the covalent bond between the reactive centers plays the leading role in superstabilizing the anionic *TC* for *TS*-analog protease inhibitors. To strengthen our case, we calculated additional model molecules with strong electron-withdrawal substituents attached to the carbonyl carbon – a fluorine substituent [FCHO] and cyano group [(CN)CHO]. The results of the calculated *TC* stability, presented in Table 1, show that superstabilization relative to the reference substituent, NH_2, is -27.8 kcal/mol for the fluorine and -39.5 kcal/mol for the cyano substituent. Thus, we can predict that peptidyl aldehyde derivatives containing $F(CO)^-$ and $NC(CO)^-$ fragments should be effective *TS*-analog inhibitors for serine proteases.

The following examples concern the variation of the atom of the electrophilic center of the substrate from a carbonyl carbon atom to the boron and phosphorous atoms. At nanomolar concentrations, peptide boronic acids are very

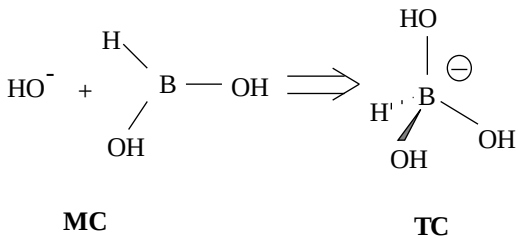
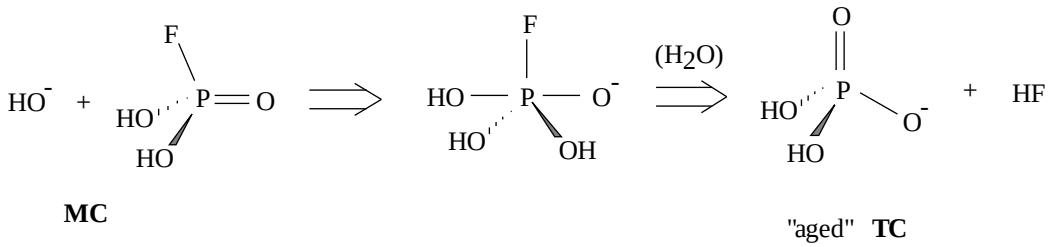


Figure 4. *Scheme of TC formation by the $HB(OH)_2$ – the simplest molecule simulating a boronic acid moiety of the inhibitor.*



MC "aged" **TC**

Figure 5. *The reaction center of DFP is modeled by $F(HO)_2PO$. Here TC is the aged tetrahedral anion, $O(HO)_2PO^-$.*

effective *TS*-analog inhibitors in a wide range of serine proteases [52, 53]. The active-site serine forms a covalent tetrahedral adduct *TC* with the boronic acid moiety of the inhibitor. We used the simplest molecule [$HB(OH)_2$] that forms a *TC* [$HB(OH)_3^-$] with OH^- anion (see the reaction scheme in Figure 4 and the stabilization energy of a *TC* in Table 1).

Many serine proteases are inhibited by organophosphorus compounds [8]. Diisopropylphosphofluoridate (DFP), which stoichiometrically and irreversibly inactivates serine proteases [54, 55] is used extensively as a diagnostic test for the presence of the active serine in an enzyme. Phosphorylated serine hydrolases are susceptible to an "aging" reaction, which involves hydrolysis of one phosphate bond [56]. The tetrahedral structure of an aged enzyme-inhibitor complex has a covalent bond between the serine oxygen and phosphorous atom [57]. In our calculations, the reactivity center of DFP was modeled by $F(HO)_2PO$. The aged product of the nucleophilic addition of the HO^- anion is the tetrahedral anion, $O(HO)_2PO^-$, as shown in Figure 5.

The calculated stabilization energies of boronic and phosphorous *TC*s are presented in Table 1. The superstabilization energies of these *TC*s relative to the *TC* of the amide ($HONH_2CO^-$) are -30.5 kcal/mol for boronic and -39.0 kcal/mol for phosphorous inhibitor. As for the carbonyl-based inhibitors, the *TC*s are more stable with the inhibitor than with the native substrate, and indeed, are found as selective inhibitors.

All energy values calculated here – the superstabilization of the inhibitor *TC*s relative to the *TC* for the modeled substrate – emerge from the formation of a s-covalent bond between the reactive centers. The mean energetic value of these effects is -30 kcal/mol, exceeding by one order of magnitude the relevant value of the hydrogen bonds in the oxyanion hole, which we estimated in Eq. 4c as -3.7 kcal/mol. This result confirms our theoretical prediction that the dominant contribution to the superstabilization of a *TC* for *TS*-analog inhibitors originates from the formation of a σ-covalent bond between the reactive centers.

Summary

We analyzed the mechanism of serine protease inhibition in a single catalytic step–the chemical transformation of *MC* ⇒ *TC*. We demonstrated that the dominant contribution to the superstabilization of a *TC* originates from the formation of a σ-covalent bond between the reactive centers in the E-S complex. We suggest a simple computational methodology for designing functional groups that can serve as the reactivity center for new classes of serine protease inhibitors–superstable anionic *TC*. To support our theoretical findings, we described several examples where excellent agreement exists between model *ab initio* quantum chemical calculations and well-known experimental results extracted from the literature.

Acknowledgments We are grateful to Mr. Yehuda Bronitsky from Ormat Industries for supporting the work of Dr. M. Shokhen. We also thank the Israeli Ministry for Absorption of New Immigrants for the Giladi scholarship and the Ministry of Science and Education for their help in supporting in this research.

References

1. Tapparelli, C; Metternich, R.; Ehrhardt, C.; Cook, N. S. *Trends Pharmacol. Sci.* **1993**, *14*, 366-376.
2. Wilhelm, O.; Reuning, U.; Jaenicke, F.; Schmitt, M.; Graeff, H. *Onkologie* **1994**, *17*, 358-366.
3. Claeson, G. *Blood Coagulation & Fibrinolysis* **1994**, *5*, 411-436.
4. Wolfenden, R. *Acc. Chem. Res.* **1972**, *5*, 10-18.
5. Wolfenden, R. *Ann. Rev. Biophys. Bioeng.* **1976**, *5*, 271-306.
6. Tomasselli, A. G.; Howe, W. J.; Sawyer T. K.; Wlodawer, A.; Heinrikson, R. L. *Chimicaoggi- Chemistry Today* **1991**, *9*, 6-27.
7. Dunn, B. M. *Advances in Detailed Reaction Mechanisms*; Vol.2; JAI Press, Inc, 1992; pp. 213-241.
8. Powers, J. C.; Harper, J. W. In *Research Monographs in Cell and Tissue Physiology*; Vol.12; *Proteinase Inhibitors* Barrett, A. J.; Salvesen, G., (Eds.); Elsevier: Amsterdam, 1988; pp. 55-152

9. Pòlgar, L. *Biol. Chem. Hoppe-Seyler* **1990**, *371, Suppl.*, 327-331.
10. Bender, M. L.; Kezdy, F. J.; Gunter, C. R. *J. Am. Chem. Soc.* **1964**, *86,* 3714-3721.
11. Longo, E.; Stamato, F. M. L.; Ferreira, R.; Tapia, O. *J. Theor. Biol.* **1985**, *112*, 783-798.
12. Warshel, A.; Russel, S. *J. Am. Chem. Soc.* **1986**, *108*, 6569-6579.
13. Daggett, V.; Schroder, S.; Kollman, P. *J. Am. Chem. Soc.* **1991**, *113,* 8926-8935.
14. Mulholand, A. J.; Grant, G. H.; Richards, W. G. *Protein Eng.* **1993**, *6,* 133-147.
15. Mackenzie, N. E.; Malthouse, J. P.; Scott, A. I. *Science* **1984**, *225,* 883-889.
16. Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77,* 334.
17. Abeles, R. H.; Alston, T. A. *J. Biol. Chem.* **1990**, *265,* 16705-16708.
18. Edwards, P.D.; Wolanin, D.J.; Andisik, D.W; Davis, M.W. *J. Med. Chem.* **1995**, *38*, 76–85.
19. Jursic, B.S.; Zdravkovski, Z. *Theochem* **1994**, *303*, 177.
20. Cork, D.G.; Hayashi, N. *Bull.Chem.Soc.Jpn.* **1993**, *66*, 1583.
21. Gilliom, R.D. *J.Compt.Chem.* **1985**, *6*, 437.
22. Pietrzycki, W.; Tomasik, P. *J.Chem.Phys.* **1993**, *90*, 2055.
23. Ménard, R.; Storer, A. C. *Biol. Chem. Hoppe-Seyler* **1992**, *373,* 393-400.
24. Warshel, A.; Naray-Szabo, G., Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28,* 3629-3637.
25. Dewar, M. J. S.; Storch, D. M. *Proc. Natl. Acad. Sci. USA* **1985**, *82,* 2225-2229.
26. Dewar, M. J. S. *Enzyme* **1986**, *36,* 8-20.
27. Wescott, C. R.; Klibanov, A. M. *Biochim. Biophys. Acta* **1994**, *1206,* 1-9.
28. Pitzer, K.S.; Brewer, L. *Thermodynamics*; McGraw-Hill: New York, 1961.
29. Menger, F. M. *Biochemistry* **1992**, *31,* 5368-5373.
30. Demuth, H-U. *J. Enzyme Inhibition*, **1990**, *3,* 249-278.
31. Shokhen, M.; Arad, D. *J. Mol. Model.* **1996**, *2*, 399-409.
32. Liang, T.-C.; Abeles, R. H. *Biochemistry* **1987**, *26,* 7603-7608.
33. Brady, K.; Liang, T.-C.; Abeles, R. H. *Biochemistry* **1989**, *28,* 9066-9070.
34. Brady, K.; Wei, A.; Ringe, D.; Abeles, R. H. Biochemistry **1990**, 29, 7600-7607.
35. Carter, P.; Wells, J. A. *Proteins: Struct., Funct., Genet.* **1990**, *7*, 335 - 342.
36. Braxton, S.; Wells, J. A. *J. Biol. Chem.* **1991**, *266,* 11797-11800.
37. McMurray, J. S.; Dyckes, D. F. *Biochemistry* **1986**, *25,* 2298-2301.
38. Frisch, M.J.; Head-Gordon, M.; Trucks,G.W.; Head-Gordon, M.; Gill, P.M.W.; Wong, M.W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Pople, J. A. *Gaussian 92, Revision A,* Gaussian, Inc., Pittsburgh PA, 1992.
39. Hwang, J.-K.; Warshel, A. *Biochemistry* **1987**, *26,* 2669-2673.
40. Warshel, A.; Aqvist, J. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267-284.
41. Page, M. I. *Chem. Soc. Revs.* **1973**, *2*, 295-323.
42. Dewar, M. J. S. *The Molecular Orbital Theory of Organic Chemistry*; McGraw-Hill: New York, 1969.
43. Williams, I. H.; Spangler, D.; Femec, D. A.; Maggiora, G. M.; Schowen, R. L. *J. Am. Chem. Soc.* **1983**, *105,* 31-40.
44. Williams, I. H. *J. Am. Chem. Soc.* **1987**, *109,* 6299-6307.
45. Madura, J. D.; Jorgensen W. I. *J. Am. Chem. Soc.* **1986**, *108,* 2517-2527.
46. Stewart, J.J.P. in *Reviews of Computational Chemistry*, Lipkovitz, K.; Boyd, D.B., Eds. VCM Publishers, New York, 1990, 45 - 81.
47. Howard, A. E. ; Kollman, P. A. *J. Am. Chem. Soc.* **1988**, 110, 7195 - 7200.
48. Feller, D.; Davidson, E.R. in *Reviews of Computational Chemistry*, Lipkovitz, K.; Boyd, D.B., Eds. VCH Publishers, New York, 1990, 1 – 43.
49. Here, J.W. Practical Strategies for Electronic Structure Calculations, Wavefunction (Irvine, California, USA) 1995.
50. Bender, M. L.; Kezdy, F. J.; Gunter, C. R. *J. Am. Chem. Soc.* **1964**, *86*, 3714-3721.
51. Abeles, R. H.; Alston, T. A. *J. Biol. Chem.* **1990**, *265*, 16705-16708.
52. Kettner, C.; Shenvi, A. B. *J. Biol. Chem.* **1984**, *259,* 15106-15114.
53. Bone, R.; Shenvi, A. B.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1987**, *26,* 7609-7614.
54. Jansen, E. J.; Nutting, M. D. F.; Jang, R.; Ball, A. K. *J. Biol. Chem.* **1949**, *179*, 189-199.
55. Jansen, E. J.; Nutting, M. D. F.; Ball, A. K. *J. Biol. Chem.* **1949**, *179*, 201-204.
56. Bender, M. L; Wedler, F. C. *J. Am. Chem. Soc.* **1972**, *94,* 2101-2109.
57. Harel, M.; Su, C.-T.; Frolow, F.; Ashani, Y.; Silman, I.; Sussman, J. L. *J. Mol. Biol.* **1991**, *221,*909-918.