JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

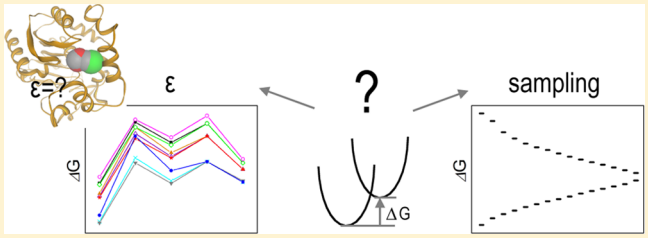
Challenges within the Linear Response Approximation When Studying Enzyme Catalysis and Effects of Mutations

Avital Sharir-Ivry,[†] Rajapandian Varatharaj,[†] and Avital Shurki*[,‡]

Institute for Drug Design, School of Pharmacy, The Lise Meitner-Minerva Center for Computational Quantum Chemistry, The Hebrew University of Jerusalem, Jerusalem 91120, Israel

Supporting Information

ABSTRACT: Various aspects of the linear response approximation (LRA) approach were examined when calculating reaction barriers within an enzyme and its different mutants. Scaling the electrostatic interactions is shown to slightly affect the absolute values of the barriers but not the overall trend when comparing wild-type and mutants. Convergence of the overall energetics was shown to depend on the sampling. Finally, the contribution of particular residues was shown to be significant, despite its small value.



1. INTRODUCTION

Understanding catalysis within enzymes as well as development and design of improved/new enzymes largely depends on our capabilities to obtain reliable and detailed description of the underlying energetics. This includes, for example, calculations of reaction profiles along with analysis of the contribution of specific amino acids. Electrostatics was shown to play a key role in enzyme catalysis, often providing the largest contribution.[1−4] As a result, various methods that evaluate both the overall electrostatic contribution and the electrostatic contribution of particular amino acids to further our understanding of enzymes were developed.[5−17]

These methods differ by various parameters that characterize them. For example, while some are based on classical calculations,[5−7] and are, thus, relatively easy to carry out, others are based on quantum calculations and are computationally more demanding.[8−12] Some utilize single structures,[13−17] usually the optimized structure of reactant state (RS) and transition state (TS),[9,10,13] while others are based on average of varying numbers of conformations.[8] Some utilize microscopic models, while others employ semimicroscopic models.[6,7] Finally some calculate the electrostatic interaction of a static system,[8−10,13] while others include the effect of reorganization resulting from the adjustment of the environment.[6,7,18]

The linear response approximation (LRA) is such a method.[6,7] It allows estimation of the free energy difference between any two states (e.g., RS and product state or RS and TS). It can account for the overall free energy, when considering all the interactions. Alternatively, it can account for only the electrostatic contribution. The method can be used either with explicit all-atom or semimicroscopic models. It is based on average over different conformations and accounts for the effect of reorganization. It involves calculation of only the two states of interest. As such it has an advantage over other methods such as the free energy perturbation (FEP)[19−21] or methods that follow a minimum path via fine scans,[22,23] where multiple points along the reaction coordinate should be calculated. Additional advantage of the LRA relates to the additive character of the approach which enables decomposition of the energy into the contributions of different components, such as separate protein residues (e.g., refs 18, 24, and 25). The accuracy of the LRA approach depends primarily on the validity of the approximation in the given system. When valid, however, the quality of the results still depends on the underlying force field (FF) and the simulation parameters. The latter involve two major factors. The first is the quality of the sampling. Multidimensional potential energy surfaces of proteins consist of many local minima and barriers. Yet, simulation time scale is finite, and the protein may be trapped in a local minimum during the time scale of the simulation. Proper sampling of the conformational space, which is essential to obtain a reliable average, is therefore, highly difficult. Often, trajectory averages, obtained from different starting conformations, differ considerably. Moreover, the LRA approach with microscopic models involves very large opposing contributions resulting with even worse convergence problems.[26] Various enhanced sampling algorithms, such as the smart walking[27] or the parallel tempering (also called replica exchange),[28−32] were developed. These methods facilitate barrier crossing, allowing, therefore, to sample regions of the conformational space, which would not have been reached using regular sampling. For example, high temperature systems easily undergo barrier-crossing events from one minimum to another. Thus, these methods exploit this fact to sample larger volumes of the phase space by exchanging conformations between the high and low temperature systems. These methods which are frequently used

Received: August 20, 2014

   dx.doi.org/10.1021/ct500751f | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

within protein folding were shown to considerably improve sampling.[28−32]

The second factor that affects the quality of the results within the LRA calculations relates to proper description of the interactions. When describing electrostatic interactions in proteins, Warshel and co-workers pointed out that although microscopic models provide rigorous formulation, a scaling factor, ε_p, is often required for quantitative results.[33−36] Despite enhanced sampling methods, convergence and proper description of reorganization effects are not fully realized.[35,36] In particular it is difficult to capture water penetration upon ionization of charged residues. Moreover, often, incomplete description of the interactions is exercised, e.g., when induced dipoles are not taken into account explicitly. Using effective scaling factor was shown to compensate for all these electrostatic interactions, which are not included explicitly in the model, and is often employed.[33,37] Yet, the value of such scaling factor was shown to be entirely dependent on the method and system used.[33,37]

In this work we explored the reliability of the LRA results to properly describe enzyme catalysis and the effects of mutations. We focused on a substitution reaction within haloalkane dehalogenase (DhlA), which is the first step in its conversion of haloalkanes into alcohols, and studied the catalytic effect of wild-type DhlA and its mutants. We first examined the effect of scaling electrostatic interactions, both on the absolute solvation energies of the reacting fragments in the protein surrounding and on the trends in the overall reaction barrier when comparing wild-type and different mutants. We then examined the convergence of the results when using enhanced sampling. Here, we explored the convergence of both the overall energetics as well as the contribution of particular residues.

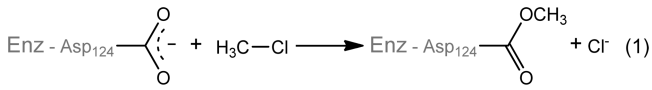
2. METHODS

2.1. Initial Protein Coordinates. Initial coordinates used in the simulations of wild-type (*wt*) DhlA correspond to the crystal structure of haloalkane dehalogenase from *Xanthobacter autotrophicus* complexed with 1,2-dichloroethane, taken from the 2DHC entry code in the Brookhaven Protein Data Bank (PDB).[38] For simplicity, 1,2-dichloroethane was replaced by methyl chloride. Initial coordinates used in the simulations of the W175Y-DhlA mutant correspond to the 1BEE PDB code which was crystallized at room temperature and pH = 6.[39] Those of the F172W mutant, on the other hand, correspond to the 1HDE PDB code.[40] The active site structure of these two mutants (including Asp124, His289, Trp125, and Trp175) was aligned with the active site structure of *wt*-DhlA. The coordinates of methyl chloride were then taken to be identical to those used in *wt*-DhlA. Finally, 3D structures of the remaining mutants (W175F, W125F) are not available, and the initial geometry was based on the *wt*-initial geometry, where the particular point mutation was introduced.

Hydrogen atoms were added using the MOLARIS program package.[41,42] For consistency, water molecules from the crystal structures were all removed, and new water molecules in and around the protein were introduced by the MOLARIS program, up to a radius of 18 Å from the reactive fragments. When compared to the X-ray structure, a water molecule close to the catalytic His289 in the active site was absent in the initial structure of both F172W and W175Y and was therefore introduced manually.

Geometry optimization, at present, is not possible at the valence bond (VB) and, thus, at the combined VB molecular mechanics (VB/MM) levels of calculation. Therefore, the geometry for the reacting fragments (methyl chloride and formate) was optimized at the HF/6-31G(d) level (see ref 43 for more details). Geometry at the TS was reoptimized at the ONIOM(HF/6-31G(d):UFF) level of calculation and was found very similar. The combined FEP approach along with the potential of mean force (PMF) protocol then followed using the two optimized geometries (RS and TS) as initial and final geometries (see the following sections). Finally, RS and TS geometries of the reactive part were chosen based on the highest and lowest energies along the VBSCF/MM - PMF/FEP free energy profile.

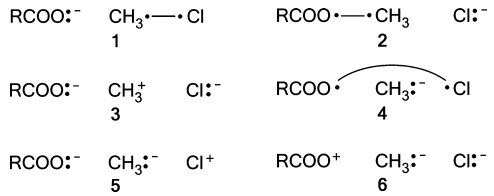
2.2. Ab Initio VB Calculations. The substrate, methyl chloride, as well as the attacking carboxylic group of Asp124 (highlighted by black color in eq 1) were considered as the reactive part and studied with ab initio VB. The substitution reaction studied (eq 1) is the first step in the conversion of haloalkanes into alcohols within the DhlA enzyme shown below:



(1)

This system involves four electrons that can occupy three different centers, the carboxylic group, the methyl group, and the chloride, resulting with six VB configurations **1**−**6** (Scheme 1). The contribution of configurations **5** and **6** was found

Scheme 1



earlier to be negligible in that system.[43,44] Thus, the calculations used configurations **1**−**4** that were found earlier to be sufficient.

The overall VB wave function therefore is given by

$$\psi = c_1\phi_1 + c_2\phi_2 + c_3\phi_3 + c_4\phi_4 \tag{2}$$

where ϕ_1, ϕ_2, ϕ_3, and ϕ_4 are the wave functions of structures **1**−**4**, respectively (Scheme 1), and c_i are the respective coefficients (see ref 45 for an overview).

Atoms of the quantum mechanical (QM) region are assigned with partial charges according to

$$Q_j^k = \sum_i w_i q_{i,j}^k \tag{3}$$

where k designates the k^{th} fragment in the reacting species (e.g, Cl, CH_3, or the RCOO group), Q_j^k is the partial charge on atom j within the k^{th} fragment, $q_{i,j}^k$ is the partial charge of the j^{th} atom in the i^{th} VB configuration, and w_i is the weight of the i^{th} VB configuration as described by the Coulson-Chirgwin formula:[46]

$$w_i = c_i^2 + \sum_j c_i c_j S_{ij} \tag{4}$$

here, c_i are the coefficients of the i^{th} *VB* configuration, and S_{ij} is the overlap between the i^{th} and j^{th} VB configurations.

Ab initio VB calculations of the system in vacuum, involved the VB self-consistent field (VBSCF) approach.[47] The inner core electrons and the out-of-plane (π-type) electrons were

dx.doi.org/10.1021/ct500751f | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

frozen at the Hartree−Fock level, while the remaining 24 valence electrons were explicitly included in the VB calculations. The calculations employed the 6-31G(d) basis set, and the VB orbitals were all strictly localized on the respective fragments (RCO_2, CH_3, and Cl). We note in this respect that a larger QM part was tested in our earlier studies of that system and demonstrated negligible differences when compared to the smaller system.[43] VB calculations were performed with the Xiamen VB (XMVB-0.1) program package,[48,49] while one electron integrals were obtained by the GAUSSIAN 03 program package.[50]

2.3. Classical Calculations. Classical calculations were used to sample protein conformations. Thus, at each point along the reaction coordinate, the coordinates of the initially optimized reactive part (methyl chloride and the carboxylate group of Asp124) were kept fixed, while the rest of the system was subjected to relaxation using molecular dynamics (MD) simulations. All relaxations were performed with the ENZYMIX module of the MOLARIS simulation package,[41,42] with the ENZYMIX FF to describe the protein environment. A detailed description of the FF parameters describing the nonbonded interactions of the reactive fragments with the environment, which were missing in the original ENZYMIX FF and were therefore added, is given in Table 1S and Figure 1S of the Supporting Information. This includes force field parameters as well as partial charges for the reacting fragments at the various VB states.

Spherical simulation systems were used. These systems included the reacting fragments and all the protein plus explicit water molecules in and around the protein up to a radius of 18 Å. The system's boundaries were restrained by the surface constrained all atom solvent model[41] for water molecules and by harmonic restraint to crystallographic positions for proteins. The rest of the system was represented by a bulk dielectric with a dielectric constant of 80. Long range electrostatic effects were treated by the local reaction field method.[51]

The calculations involved addition of two constraints which were found to be essential for proper results: 1. a constraint of 0.2 kcal mol^{-1} $Å^{-2}$ to keep a distance of 3.6 Å between the leaving chloride and the indole nitrogen of Trp125 and 2. a constraint of 0.2 kcal mol^{-1} $Å^{-2}$ to keep a distance of 3.2 Å between the chloride and the indole nitrogen of Trp175. The same constraints were therefore used also in calculations of the various mutants. In W175Y, since Trp175 is replaced by tyrosine, the second constraint involved the phenol oxygen of tyrosine instead of the indole nitrogen. Similarly, in W125F and W175F, where tryptophan is replaced by phenylalanine, the respective constraints (first and second, respectively, for W125F and W175F) involved the $C\varepsilon_1$ carbon of the phenyl ring, which is the closest to the chloride, instead of the indole nitrogen.

2.4. Procedure To Obtain VB/MM Wave Function. The VB/MM method involves an interface program that communicates between the ab initio VB calculations and the MM calculations within XMVB[48,49] and MOLARIS[41,42] packages, respectively. Detailed explanation of the VB/MM methodology can be found elsewhere;[52] here we only highlight some of the main points. The first step in any VB/MM calculation is the optimization of the VB wave function in the given environment. Thus, an initial guess wave function (set of weights for the VB structures) is assumed, and the system is propagated using this wave function. The surrounding system (protein plus explicit water molecules) is allowed to relax and accommodate to the wave function, which is kept fixed along the relaxation. Interactions between the surrounding atoms and the reacting fragments are calculated classically by mechanical embedding for each individual VB structure. Finally, the structures are mixed by solving the eigenvalue problem for the relaxed system, and a new wave function (set of new weights for the VB structures) is obtained along with the corresponding energy. This new wave function represents the reacting fragments' wave function which is polarized due to the surrounding.[53] The process is then repeated by utilizing the newly obtained weights. The iterative process continues until changes in the weights due to the environment are insignificant, suggesting that convergence is reached and the VB/MM energy as well as wave function are optimized.

The computational specifications to obtain the VBSCF/MM wave function in this work are as follows: for any geometry, the corresponding VBSCF gas-phase weights served as initial guess for *wt*-DhlA or aqueous solution, and wt weights served as initial guess for all mutants. The system was then relaxed for 100 ps using this initial guess for the wave function. 100 conformations for wt and 50 conformations for aqueous solution and mutants were then collected every 1 ps along additional 100/50 ps of simulation time, respectively, and a single point calculation using VBSCF/MM was carried out for each one of the sampled conformations. The resulting new weights and thus wave function obtained as an average of the wave functions of all conformations (100 for wt and 50 for aqueous solution and mutants, respectively) served as a new guess for additional relaxation. The process was repeated until the changes in the average weight of the various VB configurations were below 0.005. The calculations were carried out at 300 K and involved a 1 fs step size.

2.5. FEP/PMF Calculations. The reaction profiles of *wt*-DhlA and its different mutants at the VBSCF/MM level were calculated by following the combined FEP approach along with the PMF protocol.[21,22] The coordinates for the reacting fragments in aqueous solution, in wt, as well as in the various mutants, were fixed, while the coordinates of the remaining protein residues as well as the water molecules were relaxed. The reaction path was divided into 25 points starting from enzyme substrate complex all the way to the TS based on coordinates obtained at the optimization level for the two end points RS (enzyme substrate complex) and TS within the protein. For each point along the reaction path the VBSCF/MM wave function was calculated using the procedure described in section 2.4 with a fixed geometry for the reactive fragments. Once converged, each point was simulated for 50 ps in aqueous solution, in wt and in the different mutants, and energies were collected at the VBSCF/MM level every 1 ps. All the simulations were carried out at 300 K, with a 1 fs step size.

2.6. LRA Studies of Solvation. Solvation free energies were calculated as the change in free energy upon moving the reacting fragments, in a given state (e.g., RS or TS), from vacuum to the active site of the protein. Only electrostatic interactions were considered, and evaluation of the solvation free energy both in RS and in TS (ΔG_{sol}^{RS} and ΔG_{sol}^{TS}) was based on the LRA approach following eq 5:[7,21]

$$\Delta G_{sol}^{x} = \frac{1}{2}[\langle U_{ch} - U_{unch}\rangle_{ch}^{x} + \langle U_{ch} - U_{unch}\rangle_{unch}^{x}] \quad (5)$$

here $\langle\ \rangle_i$ designates an average over trajectories propagated on U_i. U_{ch} and U_{unch} represent the potential energy of the whole system (reacting fragments along with the protein and water

dx.doi.org/10.1021/ct500751f | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

environment) when the partial charges of the reacting fragments are considered, or set to zero, respectively, and x stands for the state for which solvation is studied (e.g., RS or TS). The VdW parameters within the uncharged state, U_{unch}, were kept identical to those within the charged state.

Owing to the additive character of the LRA and to the electrostatic character of the solvation, the overall solvation was estimated as a linear combination of weighted solvations of the various VB configurations, as described in eq 6:

$$\Delta G^{x}_{sol} = \Delta G^{x}_{sol}(\psi) = w^{x}_{1}\Delta G^{x}_{sol}(\phi_1) + w^{x}_{2}\Delta G^{x}_{sol}(\phi_2) + w^{x}_{3}\Delta G^{x}_{sol}(\phi_3) + w^{x}_{4}\Delta G^{x}_{sol}(\phi_4) \quad (6)$$

here, $\Delta G^{x}_{sol}(\psi)$ is the overall solvation of state x, $\Delta G^{x}_{sol}(\phi_i)$ is the solvation of the i^{th} VB configuration in state x, and w^{x}_{i} is the weight of the i^{th} VB configuration within state x, where x is RS or TS.

The VB/MM wave function for each state (RS and TS) was first defined by using the procedure described in section 2.4. Each state was than simulated for 50 ps (with fixed wave function and geometry of the reacting fragments as obtained in the first step) at a given temperature, T_i, both when the reactive fragments are polar or nonpolar (i.e., when their partial charges are set to zero). Fixing the wave function (and the geometry) here is equivalent to using an empirical VB (EVB) type of surface for the simulations, where the weights of the VB structures are taken from VB/MM calculations. This procedure largely reduces the computational cost which amounts to 48.0 min CPU time on a single Xeon(R) X5650 2.67 GHz processor for a simulation of 50 ps instead of 56.4 min CPU time without freezing the wave function (or even 143.6 min without freezing both geometry and wave function, while using an EVB surface as a reference potential and VB/MM only for the 50 sampled conformations). Relaxation of the system was then carried out at 300 K during 100 ps. Once relaxed, each state was simulated at 300 K for an additional 100 ps for *wt*-DhlA or 50 ps for aqueous solution and for all mutants. Different surrounding (protein and/or water) conformations were collected every 1 ps during these simulations. All simulations were done with a 1 fs step size. The success of the LRA results largely depends on the reliability of the sampling.

Thus, similar to parallel tempering molecular dynamics methods, the process was repeated at least 16 times for wt and its mutants and 5 times for aqueous solution, with different initial temperatures, T_i, ranging from 300 to 400 K ensuring proper sampling of the energy surface.

Finally, the electrostatic interaction between polar reactive fragments and the environment (protein and water) was calculated using both conformations taken from polar and nonpolar reactive fragments simulations. Each one of the terms in the LRA calculations thus involved averaging over at least 1600, 800, and 250 different conformations for *wt*-DhlA, the various mutants, and aqueous solution, respectively. The electrostatic interaction was calculated at the mechanical embedding level of calculation with each one of the VB structures separately and involved ionization of all ionizable residues. As pK_a calculations of residues within proteins are highly demanding (e.g., refs 54−56), ionization of the residues was based on experimental data when available (e.g., in the case of His289),[38,57−59] and the respective pK_a in solution otherwise. While exploring the effect of scaling the electrostatic contributions of the ionizable residues, various scaling schemes were used and the contributions were divided by the corresponding scaling factors. For all the remaining calculations, the contribution of all the inner ionizable residues (i.e., all the ionizable residues within the interior of the protein, which are not located on the protein surface) was then scaled by $\varepsilon = 4$ to account for various electrostatic interactions which are not included explicitly in the calculations.[36]

3. RESULTS AND DISCUSSION

3.1. Scaling Electrostatic Interactions − Effect on Solvation Free Energy. The inaccuracies in the description of electrostatic interactions arising from, for example, not considering explicitly induced dipoles or difficulties in capturing water penetration, were shown to relate mainly to the charged residues and scaling of these interactions was shown to be necessary for quantitative purposes.[33−37] Hence, to examine the effect of scaling the electrostatic interactions on the results, a scaling-factor was used to scale the electrostatic contribution of all ionizable residues. Moreover, earlier studies suggested that the values of the scaling factors normally range between 1 and 10, depending on the type of interactions missing in the model.[3,33,35,36] This range therefore was explored in our studies. Finally, earlier studies involved different scaling factors for different residues within the same system. Therefore, to explore the effect of such option, we classified the ionizable residues in two groups: 1. outer residues, including all ionizable residues which are located on the surface of the protein and are, thus, surrounded by solvent. 2. inner residues, including all the remaining ionizable residues which are surrounded by the interior of the protein. These two groups were scaled independently, in different manners. Namely, (ε_i, ε_o) represent, respectively, the scaling factor of the inner and outer residues.

Figure 1 presents the solvation energy of the reacting fragments within aqueous solution along with *wt*-DhlA and its

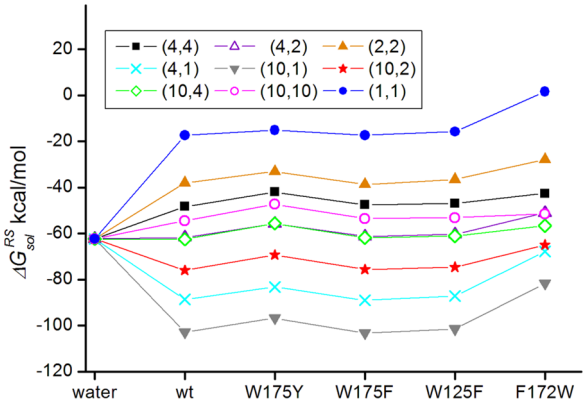


Figure 1. Solvation energy of the reactive fragments in RS, ΔG^{RS}_{sol}, using different scaling of the electrostatic interactions with the ionizable residues. Each graph corresponds to a different set of scaling factors, (ε_i, ε_o), for the inner and outer ionizable residues, respectively. Solvation energies in aqueous solution along with *wt*-DhlA and its mutants are presented.

mutants in the RS. The solvation at the TS is presented in Figure 2S in the Supporting Information. When no scaling factor is used ($\varepsilon_i = 1$, $\varepsilon_o = 1$), the solvation energy within the proteins (*wt*-DhlA or any of its mutants) is much lower (in absolute value) than solvation within aqueous solution. Enzymes, however, were shown to often solvate the reacting fragments better than aqueous solution and to a further extent TS relative to RS[4,60] (a principle, which guides rational TS

dx.doi.org/10.1021/ct500751f | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

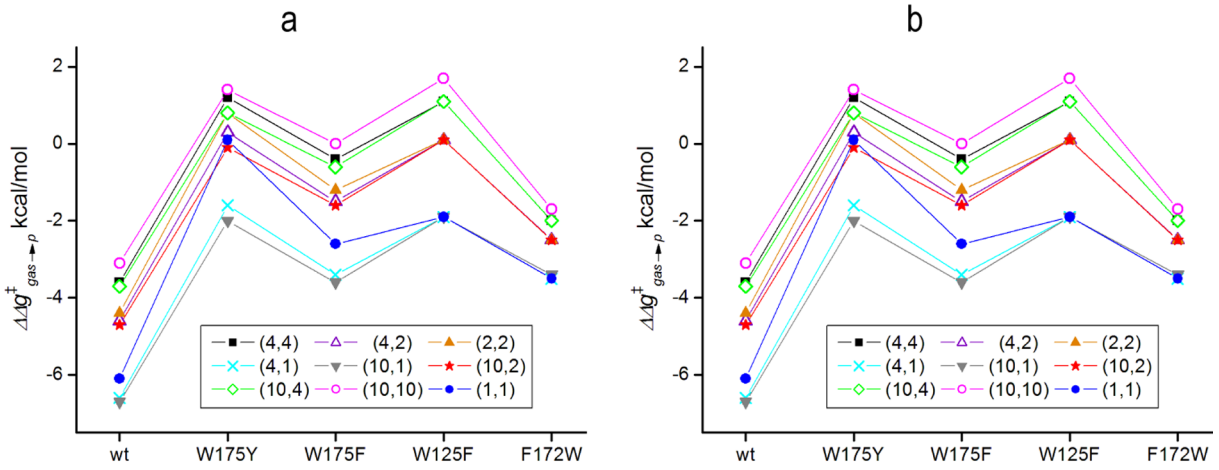


Figure 2. Effect of the scaling scheme of the electrostatic interactions with the ionizable residues on **a**. The overall change in the reaction barrier compared to that in vacuum, $\Delta\Delta g^{\ddagger}_{gas \rightarrow p}$, expressed as the differential solvation energy, $\Delta\Delta G^{RS \rightarrow TS}_{sol}$ **b**. The relative catalytic ability of the mutants, $\Delta\Delta g^{\ddagger}_{wt \rightarrow mut}$. Each graph corresponds to a different set of scaling factors, (ε_i, ε_o), for the inner and outer ionizable residues, respectively. Differential solvation energies in *wt*-DhlA as well as in its mutants are presented. Dashed line represents the observed experimental estimations.

analog design as inhibitor e.g., ref 61). This behavior (higher solvation in the enzyme) was shown to be true specifically for *wt*-DhlA,[62] suggesting that in this particular system, the description of the solvation is not sufficient without scaling the electrostatic interactions.

Different (ε_i, ε_o) schemes of scaling the electrostatic interactions are presented. The results clearly demonstrate that solvation free energies within the proteins are very sensitive to the scaling factors, as expected. Several trends can be viewed in Figure 1. Looking at the solvation energy obtained with a constant $\varepsilon_i = 10$ and varying ε_o, the following order is found: $|\Delta G^{RS}_{sol}(10,1)| > |\Delta G^{RS}_{sol}(10,2)| > |\Delta G^{RS}_{sol}(10,4)| > |\Delta G^{RS}_{sol}(10,10)|$. The same trend is observed when looking at the various solvation values obtained with a constant $\varepsilon_i = 4$. It is concluded therefore that when ε_i, the scaling factor of the inner residues, is kept constant (e.g., $\varepsilon_i = 4$ or $\varepsilon_i = 10$), the solvation energies grow as the scaling factor of the outer residues, ε_o, decreases. Considering the fact that the overall solvation obtained with no scaling is too low, these trends suggest that the contributions of outer residues to the solvation energy are properly captured and no significant corrections for the electrostatic interactions are required.

Opposite trends are observed when keeping ε_o constant with varying ε_i. Thus, for example, when $\varepsilon_o = 2$ and varying ε_i are used, the following trend is observed, $|\Delta G^{RS}_{sol}(10,2)| > |\Delta G^{RS}_{sol}(4,2)| > |\Delta G^{RS}_{sol}(2,2)|$. Similar trends are observed with $\varepsilon_o = 4$ or $\varepsilon_o = 1$ with different ε_i values. Namely, when the scaling of the outer residues, ε_o, is kept constant, the solvation energies grow as the scaling factor of the inner residues, ε_i, increases. Therefore, considering the fact that the overall solvation obtained with no scaling is too low, these opposite trends in that case suggest that contrary to the outer residues, the contributions of inner residues to the solvation energy are not fully captured and scaling is required. In other words, the growing solvation values with increased ε_i values suggest that most of the deficiencies in the description of the electrostatic interactions are within the protein interior. This result highlights, as expected, the difficulty to properly capture the effects of ionizable residues, especially in nonpolar sites in the interior of proteins.[36,63] We note in this respect that similar analysis performed for the solvation energies of the reacting fragments at the TS resulted with similar trends, leading to the same conclusions (see Figure 2S in the Supporting Information).

3.2. Scaling Electrostatic Interactions – Effect on Catalysis. The solvation energies seem to significantly change depending on the scaling scheme used. However, how do these changes in the solvation energy affect studies of catalysis?

The overall change in the reaction barrier compared to that in vacuum, $\Delta\Delta g^{\ddagger}_{gas \rightarrow p}$, can be derived from the difference in the solvation energies at the TS and the RS:

$$\Delta\Delta g^{\ddagger}_{gas \rightarrow p} = \Delta\Delta G^{RS \rightarrow TS}_{sol} = \Delta G^{TS}_{sol} - \Delta G^{RS}_{sol} \quad (7)$$

Figure 2a presents this differential solvation for *wt*-DhlA and its various mutants with different scaling schemes (ε_i, ε_o). The significant changes in the solvation energies, observed in Figure 1 ranging from 80 kcal/mol for *wt*-DhlA up to ~130 kcal/mol for mutants that involve ionizable residues, diminish into <5 kcal/mol respectively, when the differential solvation is considered. Therefore, the trends of the overall changes in the reaction barrier seem to be relatively indifferent to the scaling scheme used, in particular when compared with the great sensitivity of the absolute solvation energies.

This conclusion is also apparent when considering the catalytic effect. Comparison of the changes within the protein, $\Delta\Delta g^{\ddagger}_{gas \rightarrow p}$ with the changes within solution $\Delta\Delta g^{\ddagger}_{gas \rightarrow w}$ provides information on the catalytic effect, $\Delta\Delta g^{\ddagger}_{w \rightarrow p}$. Thus, the overall calculated catalysis of *wt*-DhlA, $\Delta\Delta g^{\ddagger}_{w \rightarrow wt}$, ranges between −8.3 and −12.0 kcal/mol, with different scaling schemes. These values are in agreement with an experimental estimation of ~−11.7 kcal/mol which is based on an observed overall barrier of 15.3 kcal/mol in *wt*-DhlA and estimated barrier of ~27 kcal/mol in solution.[62]

Finally, when considering the effect of the various mutations on the overall barrier, $\Delta\Delta g^{\ddagger}_{wt \rightarrow mut}$, as shown in Figure 2b, this conclusion becomes even clearer. $\Delta\Delta g^{\ddagger}_{wt \rightarrow mut}$ can be derived from the difference in the differential solvation of both mutant and *wt*-DhlA:

$$\begin{aligned}\Delta\Delta g^{\ddagger}_{wt \rightarrow mut} &= \Delta g^{\ddagger}_{mut} - \Delta g^{\ddagger}_{wt} \\ &= \Delta\Delta G^{RS \rightarrow TS}_{sol}(mut) - \Delta\Delta G^{RS \rightarrow TS}_{sol}(wt)\end{aligned} \quad (8)$$

dx.doi.org/10.1021/ct500751f | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

The figure therefore presents $\Delta\Delta g^{\ddagger}_{wt \rightarrow mut}$ calculated with different scaling schemes along with the experimental observation. $\Delta\Delta g^{\ddagger}_{wt \rightarrow mut}$ values with different scaling schemes obtained for W175Y and F172W, where the 3D structure was available, span a wider range of values than those obtained for W175F and W125F, where the structure was derived from *wt*. In other words, W175Y and F172W exhibit greater dependence in the scaling scheme, due to larger changes in the contributions of particular residues relative to *wt*. This suggests that our calculations may have not captured all the structural changes due to the mutations in W175F and W125F. Still, the energies, for all the mutants (W175Y, W175F, W125F, and F172W) with different scaling schemes, are nearly identical, differing by up to only 2 kcal/mol. This implies that the overall barrier of the mutant relative to that in *wt* (the total effect of the mutation) is practically indifferent to the choice of the scaling scheme in these cases. In addition, increased barrier relative to *wt* (an anticatalytic effect) is obtained for all the W175Y, W175F, W125F, and F172W mutants, as expected. It is concluded that despite the significant changes obtained in the solvation energies with different scaling schemes, the trends in catalysis are virtually unaffected. Furthermore, despite the fact that structural changes in the protein due to the mutation were not fully captured, the overall effect compared to wt was obtained.

The remaining of the results in this work will be presented with the (4,1) scaling scheme as it provides higher overall solvation energies and better accounts for the difference between the effect of the protein and the water when considering catalysis. We will stress, however, that with respect to the trends in catalysis, other schemes exhibit similar results and if used would lead to the same conclusions.

3.3. Convergence of the Solvation Energies. Methods such as the parallel tempering have demonstrated the usefulness of using higher temperatures to enhance sampling.[28−32] Reliability of the results, however, largely depends on proper convergence of such methods. The protein solvation energies for either RS or TS were based on averages over 1600 or 800 different conformations for wt or mutants, respectively. Our protein sampling involved simulating 16 replicas of the system (RS or TS) each at different temperature T_i. Conformations from the different higher temperatures were than simulated at the lower-room-temperature, and after relaxation conformations were collected (see Methods section 2.6 for more details). The final result can be considered, therefore, as average over 16 averages.

Thus, in order to examine the convergence of the results and the error range as a function of number of averages taken into account, the average over all possible subgroups of 2, 3, 4, ... 15 averages was calculated. Figure 3 presents all possible averages obtained for the solvation energy of the RS in *wt*-DhlA as a function of the number of subgroups included (varying from 2 to 15). The overall average within each subgroup along with two standard errors from that average are also presented.

It is clear from the figure that the range of possible solvation energy values in the RS decreases with the number of subgroups. The same is true also for the standard errors. It seems that for the current sampling scheme reliable results are obtained only when at least 15 subgroups are considered, where the range of possible averages diminishes to <1 kcal/mol. Similar behavior was observed for the solvation of the TS (see Figure 3S in the Supporting Information).

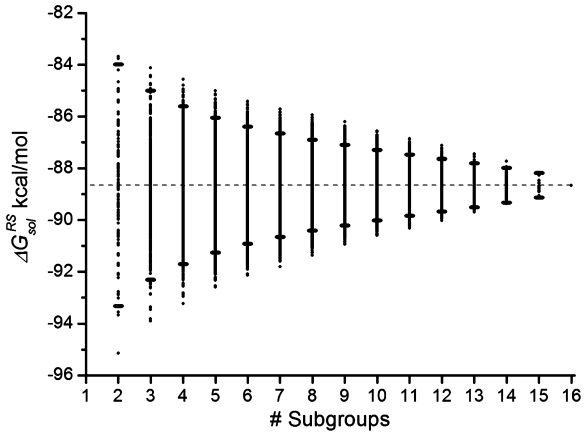


Figure 3. All possible solvation energies of *wt*-DhlA in RS, ΔG^{RS}_{sol} depending on the number of subgroups included to generate the average. Average from every possible combination of subgroups is presented by a dot. The overall average obtained from all 16 subgroups is presented by a dashed line, and plain horizontal lines represent two standard errors of each subgroup from that average. The (4,1) scaling scheme was used to scale electrostatic contributions of ionizable residues.

Figure 4 highlights the convergence of changes in the overall barrier of *wt*-DhlA compared to vacuum, $\Delta\Delta g^{\ddagger}_{gas \rightarrow p}$, as

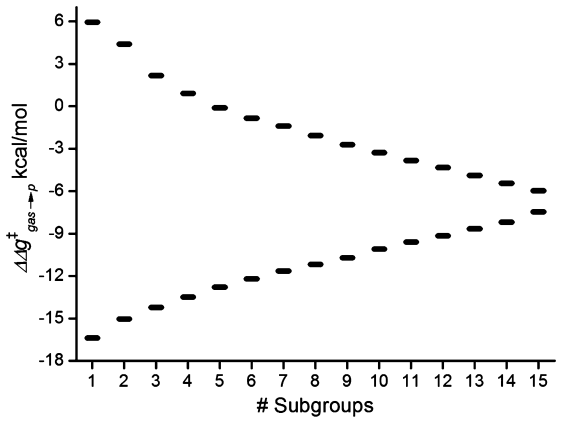


Figure 4. All possible differential solvation averages of *wt*-DhlA compared with vacuum, $\Delta\Delta g^{\ddagger}_{gas \rightarrow p}$, depending on the number of subgroups included to generate the average. Bold horizontal lines represent the maximum and minimum values obtained at each subgroup. The (4,1) scaling scheme was used to scale electrostatic contributions of ionizable residues.

obtained by the difference in the solvation of TS and RS, $\Delta\Delta G^{RS \rightarrow TS}_{sol}$ (see eq 7). The figure displays only the extreme possible values of differential solvation, maximum and minimum, obtained as a function of the number of subgroups considered. As in calculations of solvation free energy, the range of possible values is very high when only a few subgroups are used. For example when only 3 subgroups are used, the differential solvation ranges between −14.2 and 2.1 when the average is −6.6 kcal/mol. In other words the distance from the average can be higher than the value of the average itself resulting with an error which is >100% suggesting that the results are clearly not reliable. This range, however, significantly decreases with the higher number of subgroups considered. When averages of various combinations of 15 subgroups are considered, the differential solvation ranges between −7.5 to

dx.doi.org/10.1021/ct500751f | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

−6.0 kcal/mol. Most of the possible values, however, concentrate in a narrower range of ±0.3 kcal/mol around the overall average which is −6.6 kcal/mol. Again it is concluded that sampling obtained from 15 subgroups is sufficient yet required to obtain reliable results. This is especially true when the goal is to compare the barrier of different mutants, where very small differences are expected to begin with.

Figure 5 presents the maximum and minimum values along with the overall average of $\Delta\Delta g^{\ddagger}_{gas \rightarrow p}$ for *wt*-DhlA and various

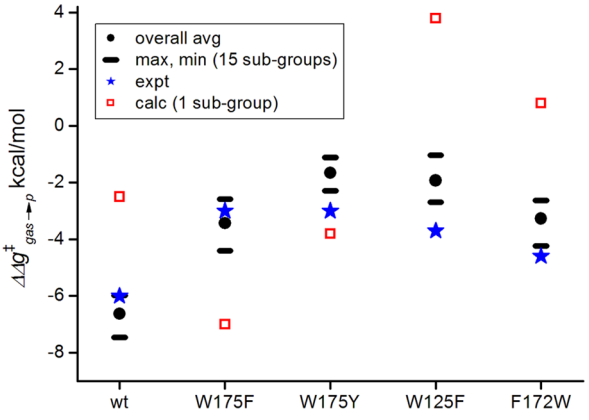


Figure 5. Changes in activation free energy compared with vacuum, $\Delta\Delta g^{\ddagger}_{gas \rightarrow p}$ for *wt*-DhlA and various of its mutants obtained by the differential solvation, $\Delta\Delta G^{RS \rightarrow TS}_{sol}$. Maximum and minimum values obtained from 15 subgroups are presented by horizontal bold lines and the overall average by circles. Values from only one subset without the use of high-temperatures are presented by squares. Experimental values are presented by stars. The overall average is presented by circles. The (4,1) scaling scheme was used to scale electrostatic contributions of ionizable residues.

of its mutants, resulting from 15 subgroups. Experimental values are also presented whenever available. For vacuum the G2 calculated value of 21.3 kcal/mol was considered as an estimation for the observed value.[64] From the figure is it clear that the various possible values obtained from 15 subgroups already produce the correct experimental trends with absolute values which are not too far from the experimental ones. Averages obtained in the framework of 16 subgroups, as was used in all remaining calculations in this work, are therefore clearly reliable. The figure also presents averages which are based on only one subgroup calculated at 300 K with no conformations from high temperatures involved. The results, in this case, present a considerably different picture from the experimentally observed picture. Here, in addition to the absolute values which are very different, the trends are also very different. Thus, for example, mutants that are known to be less efficient compared with *wt*-DhlA (e.g., W175F) are calculated to be more efficient. These results emphasize in agreement with common knowledge the importance of using proper sampling and the advantage of utilizing conformations from higher temperatures to enhance the sampling.

3.4. Convergence of per Residue Contributions to Changes in the Barrier. One of the advantages of using the LRA approach is its additive character. Namely contributions of particular amino acids to the differential solvation, $\Delta\Delta G^{RS \rightarrow TS}_{sol,i}$, can be calculated.

$$\Delta\Delta G^{RS \rightarrow TS}_{sol,i} = \Delta G^{TS}_{sol,i} - \Delta G^{RS}_{sol,i} \quad (9)$$

Here, $\Delta G^{TS}_{sol,i}$ and $\Delta G^{RS}_{sol,i}$ represent the contribution of the i^{th} residue to the solvation free energy of TS and RS, respectively. This differential solvation is a measure of the changes in the activation barrier compared to vacuum. The contributions of the particular residues can, hence, be identified as catalytic/anticatalytic. These specific contributions, however, are expected to be very small, and the question is how significant are the values obtained. It is therefore important to examine how well the per-residue contributions converge and what is the deviation of each contribution from its respective average.

Figure 6 presents the average contributions of particular residues in *wt*-DhlA to the change in the overall barrier,

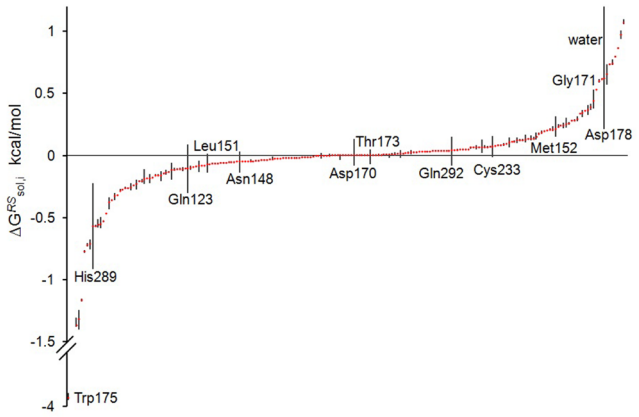


Figure 6. Average contributions of particular *wt*-DhlA residues to the overall change in the barrier, calculated as the contribution to the differential solvation, $\Delta\Delta G^{RS \rightarrow TS}_{sol,i}$, along with a sum of the average contribution of all the surrounding water molecules. Averages are obtained from all 16 subsets with error bars defined by the maximal and minimal averages over 15 possible subsets. For clarity, the residues are ordered by ascending values of average $\Delta\Delta G^{RS \rightarrow TS}_{sol,i}$ values. Residues with no contribution are excluded. Values correspond to the contribution from both the backbone and the side-chain. The (4,1) scaling scheme was used to scale electrostatic contributions of ionizable residues.

obtained from all 16 subsets. Error bars are dictated by the maximal and minimal contributions obtained when averages over only 15 possible subsets are taken into account. Due to the difficulty to label all the residues, the actual values along with the names of all residues can be found in Table 2S in the Supporting Information.

As seen from the figure the range of possible contributions for each residue is usually very narrow often smaller than 0.1 kcal/mol. In these cases the contributions appear to be fully converged and values are therefore valuable. That is, negative and positive average values signifying catalytic and anticatalytic contributions of the particular residue, respectively. Nineteen residues (most of which are labeled in the figure) involve a difference of maximum to minimum possible contribution which exceeds 0.1 kcal/mol. This implies that the particular contribution did not fully converge. While this could suggest that classification of the contributions as catalytic/anticatalytic is not valuable, as can be seen from the results this is not the case. That is, even when the range of possible residue contribution is larger than 0.1 kcal/mol, usually the entire contribution range involves either positive or negative values, and the residue can therefore be classified with high certainty as catalytic or anticatalytic. For example, the largest differences are observed for the contributions of both His289 and the water

dx.doi.org/10.1021/ct500751f | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

molecules. Still the contribution of His289 and the water molecules is negative and positive, respectively, throughout the entire range (namely even when considering the maximal and minimal values, respectively). Therefore, at least qualitatively, His289 and the water molecules can, in relatively high certainty, be considered catalytic/anticatalytic, respectively, according to the calculations. It is worth noting that the anticatalytic contribution of the water molecules in the enzyme is in agreement with the concept that this S_N2 reaction is much slower in water.

In a few cases, the maximal and minimal values are located at positive and negative values, respectively (in particular, Gln123, Asn148, Asp170, Thr173, and Gln292). Yet in most of these cases, the majority of the possible contribution range involves either positive or negative values (such as the anticatalytic Gln123 and Asn148). In fact, when standard error is considered, rather than the maximum-minimum average values usually the entire range is either positive or negative, with the exception of Asp170, Thr173, and Gln292. In the latter three cases, the overall average is very close to zero, and the classification of the residue as catalytic/anticatalytic is therefore insignificant. A full list of all residues with their calculated catalytic/anticatalytic effect can be found in the Supporting Information.

4. CONCLUDING REMARKS

Computationally quantifying the effects of mutations on overall reaction barriers within enzymes is highly important both for better understanding and for design purposes. In this work, we explored various aspects of the LRA approach to calculate solvation energies, changes in overall barriers due, for example, to mutations as well as the effects of particular residues.

The effect of scaling the electrostatic interactions was first examined. It was shown that suitable scaling is required to properly describe the catalytic effect. Yet, when comparing *wt*-enzyme with different mutants, the trends are usually kept regardless of the scaling scheme used. This, as expected, was found true especially in cases where the mutations do not involve ionizable residues. The effect of proper sampling on the results, and in particular the convergence, was also explored. It was shown that large sampling is highly important for proper convergence. In the spirit of methods such as the parallel tempering it was shown that exploiting conformations obtained from high temperature simulations is useful to enhance sampling.

Finally, the specific contributions of specific residues were also analyzed. Analyzing the contributions to the overall catalytic effect revealed that despite their very small size, the sign of the contribution is meaningful and can be used to understand the effect of the particular residues on catalysis.

■ ASSOCIATED CONTENT

Supporting Information

FF parameters for nonbonded van der Waals interactions, VB structures along with the corresponding atomic partial charges, solvation energies in the TS using different scaling schemes, convergence of ΔG_{sol}^{TS} and average contributions of particular residues in *wt*-DhlA to the change in the overall barrier. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +972-2-675-8696. Fax: +972-2-675-7076. E-mail: avitalsh@ekmd.huji.ac.il.

Present Address

‡Affiliated with the David R. Bloom Center for Pharmacy at the Hebrew University.

Author Contributions

†Equal contribution.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Arieh Warshel for many helpful discussions. This research was supported by the ISRAEL SCIENCE FOUNDATION (grant no. 1337/13). We also thank the Alex Grass Center for Drug Design and Synthesis of Novel Therapeutics for financial support. R.V. thanks The Planning and Budgeting Committee (PBC) Israel for the Postdoctoral Fellowship. A.S.I. thanks the Dan Meydan Foundation.

■ REFERENCES

(1) Warshel, A. Energetics of Enzyme Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 5250−5254.

(2) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M. Electrostatic basis for enzyme catalysis. *Chem. Rev.* **2006**, *106*, 3210−3235.

(3) Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W. Modeling electrostatic effects in proteins. *Biochim. Biophys. Acta, Proteins Proteomics* **2006**, *1764*, 1647−1676.

(4) Shurki, A.; Strajbl, M.; Schutz, C. N.; Warshel, A. Electrostatic Basis for Bioenergetics. *Methods Enzymol.* **2004**, *380*, 52−84.

(5) Carlsson, J.; Boukharta, L.; Åqvist, J. Combining Docking, Molecular Dynamics and the Linear Interaction Energy Method to Predict Binding Modes and Affinities for Non-nucleoside Inhibitors to HIV-1 Reverse Transcriptase. *J. Med. Chem.* **2008**, *51*, 2648−2656.

(6) Lee, F. S.; Chu, Z. T.; Bolger, M. B.; Warshel, A. Calculations of Antibody-Antigen Interactions: Microscopic and Semi-Microscopic Evaluation of the Free Energies of Binding of Phosphorylcholine Analogs to McPC603. *Protein Eng.* **1992**, *5*, 215−228.

(7) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. Examining Methods for Calculations of Binding Free Energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA Calculations of Ligands Binding to an HIV Protease. *Proteins* **2000**, *39*, 393−407.

(8) Perakyla, M.; Kollman, P. A. A Simulation of the Catalytic Mechanism of Aspartylglucosaminidase Using ab Initio Quantum Mechanics and Molecular Dynamics. *J. Am. Chem. Soc.* **1997**, *119*, 1189−1196.

(9) Lyne, P. D.; Mulholland, A. J.; Richards, W. G. Insights into Chorismate Mutase Catalysis From a Combined QM/MM Simulation of the Enzyme Reaction. *J. Am. Chem. Soc.* **1995**, *117*, 11345−11350.

(10) Szefczyk, B.; Mulholland, A. J.; Ranaghan, K. E.; Sokalski, W. A. Differential Transition-State Stabilization in Enzyme Catalysis: Quantum Chemical Analysis of Interactions in the Chorismate Mutase Reaction and Prediction of the Optimal Catalytic Field. *J. Am. Chem. Soc.* **2004**, *126*, 16148−16159.

(11) Funke, S. A.; Otte, N.; Eggert, T.; Bocola, M.; Jaeger, K.-E.; Thiel, W. Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. *Protein Eng., Des. Sel.* **2005**, *18*, 509−514.

(12) Liao, R.-Z.; Thiel, W. Convergence in the QM-Only and QM/MM Modeling of Enzymatic Reactions: A Case Study for Acetylene Hydratase. *J. Comput. Chem.* **2013**, *34*, 2389−2397.

(13) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. Computer Simulation and Analysis of the Reaction

Pathway of Triosephosphate Isomerase. *Biochemistry* **1991**, *30*, 5826−5832.
(14) Muegge, I.; Schweins, T.; Langen, R.; Warshel, A. Electrostatic Control of GTP and GDP Binding in the Oncoprotein p21 ras. *Structure* **1996**, *4*, 475−489.
(15) Warshel, A.; Papazyan, A.; Muegge, I. Microscopic and Semimacroscopic Redox Calculations: What Can and Cannot be Learned from Continuum Models. *J. Biol. Inorg. Chem.* **1997**, *2*, 143−152.
(16) Xiang, Y.; Oelschlaeger, P.; Florian, J.; Goodman, M. F.; Warshel, A. Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: Evaluating amino acid group contribution and allosteric coupling for ionized residues in human pol beta. *Biochemistry* **2006**, *45*, 7036−7048.
(17) Frushicheva, M. P.; Cao, J.; Chu, Z. T.; Warshel, A. Exploring challenges in rational enzyme design by simulating the catalysis in artificial kemp eliminase. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16869−16874.
(18) Florian, J.; Goodman, M. F.; Warshel, A. Theoretical Investigation of the Binding Free Energies and Key Substrute-Recognition Components of the Replication Fidelity of Human DNA Polymerase b. *J. Phys. Chem. B* **2002**, *106*, 5739−5753.
(19) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420.
(20) Valleau, J. P.; Torrie, G. M. *Modern Theoretical Chemistry*; Plenum Press: New York, 1977; Vol. *5*, pp 169−194.
(21) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.
(22) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(23) Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.
(24) Shurki, A.; Warshel, A. Why does the Ras Switch "Break" by Oncogenic Mutations? *Proteins* **2004**, *55*, 1−10.
(25) Luo, J.; van Loo, B.; Kamerlin, S. C. L. Catalytic promiscuity in Pseudomonas aeruginosa arylsulfatase as an example of chemistry-driven protein evolution. *FEBS Lett.* **2012**, *586*, 1622−1630.
(26) Sham, Y. Y.; Chu, Z. T.; Warshel, A. Consistent Calculations of pK_a's of Ionizable Residues in Proteins: Semi-microscopic and Macroscopic Approaches. *J. Phys. Chem. B* **1997**, *101*, 4458−4472.
(27) Zhou, R.; Berne, B. J. Smart walking: A new method for Boltzmann sampling of protein conformations. *J. Chem. Phys.* **1997**, *107*, 9185−9196.
(28) Swendsen, R. H.; Wang, J. S. Replica Monte Carlo simulation of spin glasses. *Phys. Rev. Lett.* **1986**, *57*, 2607−2609.
(29) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(30) Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.* **2002**, *116*, 9058.
(31) Earl, D. J.; Deem, M. W. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910−3916.
(32) Rosta, E.; Hummer, G. Error and efficiency of replica exchange molecular dynamics simulations. *J. Chem. Phys.* **2009**, *131*, 165102.
(33) Warshel, A.; Russell, S. T. Calculations of Electrostatic Interactions in Biological Systems and in Solutions. *Q. Rev. Biophys.* **1984**, *17*, 283−421.
(34) Sham, Y. Y.; Muegge, I.; Warshel, A. The Effect of Protein Relaxation on Charge-Charge Interactions and Dielectric Constants of Proteins. *Biophys. J.* **1998**, *74*, 1744−1753.
(35) Warshel, A.; Dryga, A. Simulating electrostatic energies in proteins: Perspectives and some recent studies of pK(a)s, redox, and other crucial functional properties. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 3469−3484.
(36) Schutz, C. N.; Warshel, A. What are the dielectric "constants" of proteins and how to validate electrostatic models. *Proteins* **2001**, *44*, 400−417.
(37) Dwyer, J. J.; Gittis, A. G.; Karp, D. A.; Lattman, E. E.; Spencer, D. S.; Stites, W. E.; Garcia-Moreno, E. B. High apparent dielectric constants in the interior of a protein reflect water penetration. *Biophys. J.* **2000**, *79*, 1610−20.
(38) Verschueren, K. H.; Seljee, F.; Rozeboom, H. J.; Kalk, K. H.; Dijkstra, B. W. Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature* **1993**, *363*, 693−698.
(39) Krooshof, G. H.; Ridder, I. S.; Tepper, A.; Vos, G. J.; Rozeboom, H. J.; Kalk, K. H.; Dijkstra, B. W.; Janssen, D. B. Kinetic analysis and X-ray structure of haloalkane dehalogenase with a modified halide-binding site. *Biochemistry* **1998**, *37*, 15013−15023.
(40) Schanstra, J. P.; Ridder, I. S.; Heimeriks, G. J.; Rink, R.; Poelarends, G. J.; Kalk, K. H.; Dijkstra, B. W.; Janssen, D. B. Kinetic characterization and X-ray structure of a mutant of haloalkane dehalogenase with higher catalytic activity and modified substrate range. *Biochemistry* **1996**, *35*, 13186−13195.
(41) King, G.; Warshel, A. A Surface Constrained All-Atom Solvent Model for Effective Simulations of Polar Solutions. *J. Chem. Phys.* **1989**, *91*, 3647−3661.
(42) Chu, Z. T.; Villa, J.; Strajbl, M.; Schutz, C. N.; Shurki, A.; Warshel, A. *MOLARIS version beta9.05*; University of Southern California: Los Angeles, 2004.
(43) Sharir-Ivry, A.; Shnerb, T.; Strajbl, M.; Shurki, A. VB/MM protein landscapes: a study of the S(N)2 reaction in haloalkane dehalogenase. *J. Phys. Chem. B* **2010**, *114*, 2212−2218.
(44) Song, L. C.; Wu, W.; Hiberty, P. C.; Shaik, S. Identity S(N)2 reactions X-+CH3X -> XCH3+X-(X=F,Cl,Br, and I) in vacuum and in aqueous solution: A valence bond study. *Chem.—Eur. J.* **2006**, *12*, 7458−7466.
(45) Shurki, A.; Derat, E.; Barrozoc, A.; Kamerlin, S. C. L. How Valence Bond Theory Can Help You Understand Your (Bio)chemical Reaction. *Chem. Soc. Rev.* **2014**, DOI: 10.1039/c0xx00000x.
(46) Chirgwin, H. B.; Coulson, C. A. *Proc. R. Soc. London, Ser. A* **1950**, *2*, 196.
(47) van Lenthe, J. H.; Balint-Kurti, G. G. The Valence Bond Self Consistent Field Method (VB-SCF): Theory and Test Calculations. *J. Chem. Phys.* **1983**, *78*, 5699−5713.
(48) Song, L.; Mo, Y.; Zhang, Q.; Wu, W. *XMVB 1.0: An Ab Initio Spin-free Valence Bond Program*; Xiamen University: Xiamen, 361005, China, 2003.
(49) Song, L. C.; Mo, Y. R.; Zhang, Q. N.; Wu, W. XMVB*: A program for ab initio nonorthogonal valence bond computations. *J. Comput. Chem.* **2005**, *26*, 514−521.
(50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision B.04; Gaussian, Inc.: Pittsburgh, PA, 2003.
(51) Lee, F. S.; Warshel, A. A Local Reaction Field Method for Fast Evaluation of Long-Range Electrostatic Interactions in Molecular Simulations. *J. Chem. Phys.* **1992**, *97*, 3100−3107.

(52) Shurki, A.; Crown, H. A. Hybrid ab initio VB/MM method - A valence bond ride through classical landscapes. *J. Phys. Chem. B* **2005**, *109*, 23638−23644.
(53) Sharir-Ivry, A.; Shurki, A. Valence-Bond Based Hybrid Quantum Mechanics Molecular Mechanics Approaches and Proper Inclusion of the Effect of the Surrounding. *Isr. J. Chem.* **2014**, *54*, 1189−1204.
(54) McCoy, R. S.; Braun-Sand, S. B. Semimicroscopic investigation of active site pKa values in peptidylarginine deiminase 4. *Theor. Chem. Acc.* **2012**, *131*, 1293−1299.
(55) Repic, M.; Purg, M.; Vianello, R.; Mavri, J. Examining Electrostatic Preorganization in Monoamine Oxidases A and B by Structural Comparison and pKa Calculations. *J. Phys. Chem. B* **2014**, *118*, 4326−4332.
(56) Borstnar, R.; Repic, M.; Kamerlin, S. C. L.; Vianello, R.; Mavri, J. Computational Study of the pKa Values of Potential Catalytic Residues in the Active Site of Monoamine Oxidase B. *J. Chem. Theory Comput.* **2012**, *8*, 3864−3870.
(57) Verschueren, K. H. G.; Franken, S. M.; Rozenboom, H. J.; Kalk, K. H.; Dijkstra, B. W. Refined X-ray structures of haloalkane dehalogenase at pH 6.2 and pH 8.2 and implications for the reaction mechanism. *J. Mol. Biol.* **1993**, *232*, 856−872.
(58) Pries, F.; Kingma, J.; Krooshof, G. H.; Jeronimusstratingh, C. M.; Bruins, A. P.; Janssen, D. B. Histidine-289 Is Essential for Hydrolysis of the Alkyl-Enzyme Intermediate of Haloalkane Dehalogenase. *J. Biol. Chem.* **1995**, *270*, 10405−10411.
(59) Krooshof, G. H.; Kwant, E. M.; Damborský, J.; Koca, J.; Janssen, D. B. Repositioning the catalytic triad aspartic acid of haloalkane dehalogenase: effects on stability, kinetics, and structure. *Biochemistry* **1997**, *36*, 9571−80.
(60) Shurki, A.; Warshel, A. Structure/Function Correlations of Protreins using MM, QM/MM and Related Approaches; Methods, Concepts, Pitfalls and Current Progress. *Adv. Protein Chem.* **2003**, *66*, 249−312.
(61) Roca, M.; Vardi-Kilshtain, A.; Warshel, A. Toward Accurate Screening in Computer Aided Enzyme Design. *Biochemistry* **2009**, *48*, 3046−3056.
(62) Olsson, M. H. M.; Warshel, A. Solute solvent dynamics and energetics in enzyme catalysis: the S(N)2 reaction of dehalogenase as a general benchmark. *J. Am. Chem. Soc.* **2004**, *126*, 15167−15179.
(63) Garcia-Moreno, B.; Dwyer, J. J.; Gittis, A. G.; Lattman, E. E.; Spencer, D. S.; Stites, W. E. Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophys. Chem.* **1997**, *64*, 211−224.
(64) Devi-Kesavan, L. S.; Gao, J. Combined QM/MM Study of the Mechanism and Kinetic Isotope Effect of the Nucleophilic Substitution Reaction in Haloalkane Dehalogenase. *J. Am. Chem. Soc.* **2003**, *125*, 1532−1540.