ORIGINAL SCIENTIFIC PAPER

CROATICA CHEMICA ACTA

Croat. Chem. Acta **2025**, *98*(4), 231–245
Published online: December 14, 2025
DOI: 10.5562/cca4168

HKD 1926

A View of the Inhibition Specificity of Aurintricarboxylic Acid Toward Protein Tyrosine Phosphatases: Implications for Structure-Based Phosphatase Inhibitor Design

Petar M. Mitrasinovic

Center for Biophysical and Chemical Research, Belgrade Institute of Science and Technology, 11060 Belgrade, Serbia
Author's e-mail address: petar.mitrasinovic@sbb.rs

RECEIVED: March 19, 2025 ✶ REVISED: November 4, 2025 ✶ ACCEPTED: November 4, 2025

Abstract: Protein tyrosine phosphatases (PTPs) play a vital role in intracellular signaling pathways associated with cell transformation, growth, and proliferation. Among various carboxylic acid derivatives, aurintricarboxylic acid (ATA) has been identified as the binder with the highest affinity for *Yersinia* PTP (YopH). Herein, the underlying specificity of ATA, as a competitive inhibitor of several PTPs (YopH, PTP1B, TCPTP, HePTP, CD45, VHR, and Cdc25A), is explored. By observing each PTP as a primary target receptor and the rest of them as simultaneous secondary receptors, the specificity factor, as a function of free ligand concentration ([ATA]), shows at lower [ATA] (between 10^{-11} and 10^{-9} M) a 23 to 500-fold specificity in favor of YopH versus the other PTPs, indicating that the ATA-YopH interaction is the most specific. Near [ATA] = 10^{-8} M, the specificity factor ≈ 1 is associated with equal occupancy of the primary receptor YopH and the set of secondary receptors. When observing YopH as a primary receptor relative to every other single PTP as a secondary receptor at [ATA] ϵ (10^{-11}, 10^{-9} M), the specificity to YopH is larger nearly 6, 25, 25, 40, 100, and 120 times than that to PTP1B, HePTP, CD45, TCPTP, Cdc25A, and VHR, respectively. The relative contribution to the binding free energy of each individual YopH residue located in the active site (403-Cys-Arg-Ala-Gly-Val-Gly-Arg-Thr-410) and in the proximal WPD loop (352-Gly-Asn-Trp-Pro-Asp-Gln-Thr-Ala-Val-Ser-361) was estimated using alanine scanning mutagenesis. The so-called "hot spots" of binding affinity that dominate the interaction are mostly Trp354 and Thr358. Although YopH has no homolog in *Staphylococcus aureus*, off-target effects of ATA cannot be ruled out. Considering the ability of ATA to bind to the *Staphylococcus aureus* serine/threonine phosphatase STP1 directly, the specificity factor indicates a roughly 100-fold preference for YopH over STP1. Whereas the biggest issue for identifying phosphatase inhibitors is the lack of specificity, this work helps to conceptualize the ATA structure as a pharmacophore for designing new inhibitors capable of achieving specificity against phosphatases. Based on a two-layer QM/MM (SCC-DFTB-D/AMBER) approach, the conformational relationship between the L7 loop (384-Glu-Ser-Lys-Gly-Ser-Ser-Ala-Val-Ala-392) and the WPD loop in substrate (pTyr) recognition was characterized by constructing a binding isotherm log(*f*/1-*f*)-log[substrate concentration], where *f* denotes the fraction of the YopH receptors occupied by the substrate. The observed positive association suggests that L7 is an allosteric site in the YopH structure.

Keywords: aurintricarboxylic acid, inhibitor, pharmacophore, phosphatase, specificity.

INTRODUCTION

A PHOSPHATASE is a subcategory of hydrolases that catalyzes the hydrolysis of its substrate by using water to cleave a phosphoric acid monoester into a phosphate ion and an alcohol.[1] Phosphorylation by protein tyrosine kinases (PTKs) and dephosphorylation by protein tyrosine phosphatases play a variety of roles in cellular regulation and signaling.[2] In other words, PTKs and PTPs are in charge with a post-translational modification being essential to the cell's regulatory network.[3] Regulation of vital cellular processes, which control growth, proliferation, motility, apoptosis, metabolism, angiogenesis, inflammation, etc.,[4] is associated with reversible protein tyrosine phosphorylation catalyzed by the balanced action of PTKs[5–9] and PTPs.[10,11] Thus, the aetiology of some human diseases, including diabetes, cancer and immune dysfunctions, is associated with the perturbed balance between PTK[12,13]

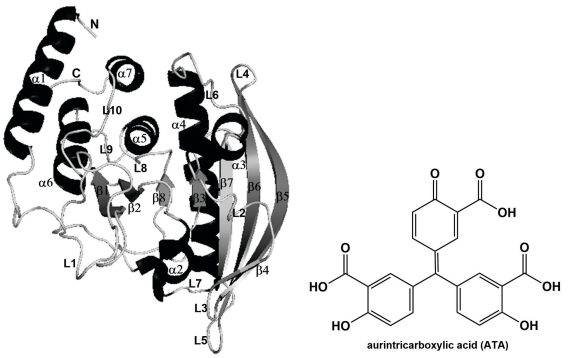

Figure 1. (Left) The tertiary structure of apo (ligand-free) YopH (PDB ID: 1YPT) consists of seven alpha helices (α1–α7, black), eight beta sheets (β1–β8, gray), and ten loops (L1–L10, white). (Right) The chemical structure of ATA. Among structurally diversified carboxylic acid derivatives, ATA has been identified using both experimental methods[28] and biomolecular modeling[29] as a direct binder of the largest affinity to YopH.

and PTP[14,15] activities. It means that signaling pathways regulated by protein tyrosine phosphorylation represent an abundant source of possible targets for novel drug design and development,[4] thereby making PTKs[12,16] and PTPs[17,18] attractive enzymes for pharmaceutical research.[19,20]

YopH is a very active PTP involved in cell transformation, growth, and proliferation. In the context of the virulence of the bacterium *Yersinia pestis*, which caused serious pandemics by infecting humans and some mammals in the past,[21,22] YopH is capable of breaking down signal transduction mechanisms in immune cells and inhibiting the immune response. Because vaccines[23,24] and antibiotics[25] are not effective treatments, search for new antibacterial agents that selectively inhibit the activity of YopH (Figure 1, Left) is of certain biomedical priority. Aurintricarboxylic acid (ATA, Figure 1, Right) is a small molecule with well-recognized ability to inhibit nucleases and nucleic acid-binding enzymes.[26,27] Among structurally diversified carboxylic acid derivatives, ATA was identified using both experimental methods[28] and biomolecular modeling[29] as a binder of the largest affinity to YopH, associated with a submicromolar IC_{50} value of 0.01 μM.[28] The biggest challenge for phosphatase inhibitor identification is the achievement of specificity.

The structural diversity of PTPs and their localization to different cellular compartments mean that each enzyme may be involved in different signaling pathways.[30] Numerous cloned cDNAs indicated a homologous (circa 260-residue) catalytic domain characteristic of PTPs, which includes the active site consensus sequence - HisCysXXGlyXGlyArgThr/Ser (loop L8, Figure 1, Left).[31] The underlying specificity issue of ATA, as a competitive inhibitor of several mammalian PTPs (YopH, PTP1B, TCPTP, HePTP, CD45, VHR, and Cdc25A), is addressed in this study. Noteworthy is that the direct binding of ATA with the *Staphylococcus aureus* serine/threonine phosphatase STP1 (IC_{50} = 1.03 μM) was reported.[32] Considering the lack of the *Yersinia pestis* YopH homolog in *Staphylococcus aureus*, the off-target effects of ATA cannot be neglected.[32] Therefore, ATA preference for YopH over STP1 is investigated here. The experimental identification of a second substrate binding site (loop L7, Figure 1, Left) highlighted the need for a more comprehensive understanding of the structure of YopH in substrate recognition,[33] which has been characterized by both experimental[34–36] and computational[37] means. Some progress, which is appropriate for the specific research context, is elaborated here.

METHODS

Theoretical Concepts

Theoretical concepts[38–40] used to characterize the specificity of ATA against phosphatases are briefly presented in this section.

The binding equilibrium between the ligand (L) and the primary receptor (PR) is given by:

$$\mathrm{PR{:}L} \leftrightarrow \mathrm{PR} + \mathrm{L} \tag{1}$$

The strength of this interaction is characterized by the dissociation constant, $K_{D,PR}$:

$$K_{D,PR} = \frac{[PR][L]}{[PR:L]} \tag{2}$$

If the ligand can bind to a set of N secondary receptors, SR_i, in addition to the primary one, PR, then the binding equilibrium for each of these interactions is given by:

$$SR_i : L \leftrightarrow SR_i + L \tag{3}$$

The dissociation constant for each of these interactions is given by:

$$K_{D,SR_i} = \frac{[SR_i][L]}{[SR_i:L]} \tag{4}$$

In structure-based inhibitor design and development, the presence of PR and SR_i in the same sample means that the free (equilibrium) ligand concentration, [L], is identical in equations 2 and 4 being satisfied simultaneously. Thus, the specificity factor, α, can be defined[39] as the ratio of the concentration of ligand bound to the primary receptor, PR, to the total concentration of ligand bound to all of the secondary receptors, SR_i:

$$\alpha = \frac{[PR:L]}{\sum_{i=1}^{N}[SR_i:L]} \tag{5}$$

It is convenient to express α in terms of the dissociation constants and the ligand concentration. For the primary receptor, PR:

$$[PR]_{total} = [PR] + [PR : L] \quad (6)$$

so:

$$[PR] = [PR]_{total} - [PR : L] \quad (7)$$

Rearranging equation 2 gives:

$$[PR] = \frac{K_{D,PR}[PR : L]}{[L]} \quad (8)$$

Substituting the expression for [PR] from equation 8 into equation 7 gives:

$$[PR : L] = \frac{[PR]_{total}}{1 + \frac{K_{D,PR}}{[L]}} \quad (9)$$

For the secondary receptor, SR_i:

$$[SR_i]_{total} = [SR_i] + [SR_i : L] \quad (10)$$

so:

$$[SR_i] = [SR_i]_{total} - [SR_i : L] \quad (11)$$

Rearranging equation 4 gives:

$$[SR_i] = \frac{K_{D,SR_i}[SR_i : L]}{[L]} \quad (12)$$

Substituting the expression for $[SR_i]$ from equation 12 into equation 11 gives:

$$[SR_i : L] = \frac{[SR_i]_{total}}{1 + \frac{K_{D,SR_i}}{[L]}} \quad (13)$$

Equation 13 is valid for each of the secondary receptor-ligand interactions separately.

Putting the relationships from equations 9 and 13 into equation 5 gives:

$$\alpha = \frac{\frac{[PR]_{total}}{1 + \frac{K_{D,PR}}{[L]}}}{\sum_{i=1}^{N}\left(\frac{[SR_i]_{total}}{1 + \frac{K_{D,SR_i}}{[L]}}\right)} \quad (14)$$

If the dissociation constants are known, the free ligand concentration and the total concentrations of PR and SR_i, equation 14 enables to calculate the specificity. A high value of α, which shows how much of the ligand is bound to a specific receptorb compared to others, suggests that the primary receptor is preferentially bound over the secondary receptors.[39]

The activity of a receptor is usually seen through whether a ligand is bound to it or not. It is useful to describe the specificity in terms of fractional occupancy. The fractional saturation or fractional occupancy, *f*, is the extent to which the binding sites on a receptor are filled with ligand. If it is assumed that each receptor molecule can bind to one ligand molecule, then *f* is the ratio of the number of receptor molecules that have ligand bound to them to the total number of receptor molecules.[38] For the primary receptor, PR, in terms of concentrations, *f* can be given by:

$$f_{PR} = \frac{[PR : L]}{[PR] + [PR : L]} = \frac{[PR : L]}{[PR]_{total}} \quad (15)$$

Based on equation 2, it follows:

$$[PR : L] = \frac{[PR][L]}{K_{D,PR}} \quad (16)$$

Substitution of the expression for [PR:L] from equation 16 into equation 15 gives:

$$f_{PR} = \frac{[PR]\,[L]}{K_{D,PR}\left([PR] + \frac{[PR]\,[L]}{K_{D,PR}}\right)} \quad (17)$$

Rearranging the right-hand side of equation 17 gives:

$$f_{PR} = \frac{\frac{[L]}{K_{D,PR}}}{1 + \frac{[L]}{K_{D,PR}}} \quad (18)$$

or

$$f_{PR} = \frac{[L]}{K_{D,PR} + [L]} \quad (19)$$

For the secondary receptor, SR_i, the fractional occupancy, *f*, can be expressed as:

$$f_{SR_i} = \frac{[SR_i : L]}{[SR_i] + [SR_i : L]} = \frac{[SR_i : L]}{[SR_i]_{total}} \quad (20)$$

Based on equation 4, it follows:

$$[SR_i : L] = \frac{[SR_i][L]}{K_{D,SR_i}} \quad (21)$$

Substitution of the expression for $[SR_i:L]$ from equation 21 into equation 20 gives:

$$f_{SR_i} = \frac{[SR_i]\,[L]}{K_{D,SR_i}\left([SR_i] + \frac{[SR_i]\,[L]}{K_{D,SR_i}}\right)} \quad (22)$$

Rearranging the right-hand side of equation 22 gives:

$$f_{SR_i} = \frac{\frac{[L]}{K_{D,SR_i}}}{1 + \frac{[L]}{K_{D,SR_i}}} \quad (23)$$

or

$$f_{SR_i} = \frac{[L]}{K_{D,SR_i} + [L]} \quad (24)$$

Based on equation 5, the specificity factor, α, can be expressed in terms of the fractional occupancies, f_{PR} and f_{SRi}, and the total concentrations, $[PR]_{total}$ and $[SR_i]_{total}$, by means of:

$$\alpha = \frac{f_{PR}\,[PR]_{total}}{\sum_{i=1}^{N} f_{SR_i}\,[SR_i]_{total}} \quad (25)$$

A large fractional occupancy of the target receptor, PR, and a low fractional occupancy of off-target receptors, SR_i, corresponds to high specificity.

Using equations 18 and 23, α can be expressed as:

$$\alpha = \frac{\frac{\frac{[L]}{K_{D,PR}}}{1 + \frac{[L]}{K_{D,PR}}}\,[PR]_{total}}{\sum_{i=1}^{N} \frac{\frac{[L]}{K_{D,SR_i}}}{1 + \frac{[L]}{K_{D,SR_i}}}\,[SR_i]_{total}} \quad (26)$$

Using equations 19 and 24, α can be written as:

$$\alpha = \frac{\frac{[L]}{K_{D,PR} + [L]}\,[PR]_{total}}{\sum_{i=1}^{N} \frac{[L]}{K_{D,SR_i} + [L]}\,[SR_i]_{total}} \quad (27)$$

If the dissociation constants and total concentrations PR and SR_i are known, equation 27 is convenient for examining the specificity factor, α, as a function of the free ligand concentration, [L], and for observing trends in specificity by comparing different ranges of [L] against the values of $K_{D,PR}$ and $K_{D,SRi}$.

A plot of fractional occupancy, f, as a function of ligand concentration, determined at constant temperature, is known as a binding isotherm or binding curve. The value of the dissociation constant, K_D, corresponds to the concentration of free ligand, [L], at which the receptor is half saturated. For a binding curve to be meaningful, all measurements have to be made at constant temperature. f at a given concentration of free ligand depends on K_D as given by equations 19 and 24. The value of K_D depends on temperature according to the relationship:

$$\Delta G^{\circ}_{bind} = R\,T \ln K_D \quad (28)$$

where ΔG°_{bind} is the standard free-energy change upon complex formation, that is, the change in free energy when 1 mole of receptor binds to 1 mole of ligand under standard conditions (1 molar solution of each), R is the gas constant (1.9872×10^{-3} kcal K^{-1} mol^{-1}), and T is the absolute temperature. It means that K_D is a constant only if T is maintained at a constant value. If T is allowed to fluctuate while making binding measurements, the results will not make sense.[38]

Computational Background

Alanine scanning mutagenesis is an effective tool for determining specific protein residues that make the most dominant contributions to ligand-protein binding interactions. To identify so-called "hot spots" of binding, residues in the ATA-YopH interface (active site and WPD loop) were individually mutated to alanine by applying the PyMol mutagenesis tool to the wild-type (wt) receptor in a backbone-dependent manner.[41] The experimental structure of wt YopH in the closed conformation was used (PDB ID: 1YTN, with the ligand removed). The universal force field (UFF) molecular mechanics (MM) method, implemented in the ArgusLab 4.0.1 software package,[42] was used for initially cleaning up the structure of every receptor. 1000 steepest descent minimization steps with a gradient convergence of 10^{-2} kcal mol^{-1} $Å^{-1}$ and MM non-bonded cutoffs of 8 Å (start) and 10 Å (end) were used in order to eliminate potential clashes and bad contacts. The relative change in binding free energy ($\Delta\Delta G$) was calculated for each single-point mutant relative to wt YopH ($\Delta\Delta G = \Delta G_{mutant} - \Delta G_{wt}$) using docking experiments. The geometry of ATA was optimized using Gaussian 98 at the B3LYP/6-31G level of theory.[43] Flexible docking of ATA in the receptor active site was performed using the AScore/ShapeDock protocol from ArgusLab 4.0.1.[42] AScore is an empirical scoring function based on the decomposition of the total free energy of protein-ligand binding into physically interpretable components. The various contributions account for van der Waals (VDW) interactions between ligand and protein, hydrophobic effects, hydrogen bonding between ligand and protein, hydrogen bonding involving charged donor and/or acceptor groups, deformation effects, and translational and rotational entropy loss effects in the binding process, respectively. The intra-ligand van der Waals energy is included in the total VDW term. The ShapeDock engine provides an approximate alternative to the complicated search problem and explores all energy minima. Flexible ligand docking is incorporated by describing the ligand as a torsion tree. Groups of bonded atoms that do not have rotating bonds are nodes, while torsions are the bonds between nodes. The topology of the torsion tree is a decisive factor for efficient docking. In the case of ATA, the

active site-adapted docking box had a size of 11.5 Å × 16.5 Å × 10 Å with a grid resolution of 0.25 Å. From each docking experiment, the physically relevant pose was selected, i.e. the one with the lowest binding free energy in the finally refined cluster of poses that bind to the correct binding site. This particular protocol has previously been shown to be consistent for docking small ligand molecules in crystal structures of viral proteins, yielding binding free energies that correlated well with experimental inhibitory concentrations.[44] The average thermochemical accuracy (2 kcal mol^{-1}) associated with molecular modeling of various systems[45–50] suggests that empirical scoring functions may be applicable to predictions of binding affinity in structure-based drug design.[51–54]

To investigate the involvement of the WPD loop and the L7 loop in substrate (phosphorylated tyrosine - pTyr) recognition, it was necessary to estimate the dissociation constants of the two substrate binding sites in the YopH structure. A two-layer quantum mechanics/molecular mechanics (QM/MM) approach was used to calculate the binding affinity of pTyr for YopH. This means that a small structural part of the biomolecular complex is treated with the higher-level QM method, while the rest of the structure is treated with the computationally less demanding MM method. This approach accounts for changes in the electronic structure in critical regions of enzyme function, thereby overcoming the limitations of MM simulations, which cannot capture interactions associated with charge transfer. The chosen approach is essentially the one previously used in the case of ATA binding, in which the QM region consists of the active site, the WPD loop, and ATA.[29] For the reported ΔG°_{bind} of –11.3 kcal mol^{-1},[29] equation 28 gives a K_D value of about 5.8 nM, which is in reasonable agreement with the experimental value of 5 nM.[28]

The geometry of pTyr was optimized using Gaussian 98 at the B3LYP/6-31G level of theory.[43] Initial structures of the complexes were generated by docking pTyr into the two substrate binding sites in the closed conformation of wt YopH (PDB ID: 1YTN) using ArgusLab 4.0.1.[42] The closed conformation of the WPD loop located near the active site in the YopH structure characterizes a more relevant receptor, in terms of quantitative correlation of calculated binding free energies with experimental data, than does the open one.[29] The method described below is quite insensitive to the specific choice of the experimental structure of the receptor in the closed conformation because the initial structure is subjected to a series of different minimization steps (steepest descent, conjugate gradient, and QM/MM), which ultimately yield almost the same result. The lower affinity binding site is the L7 loop, while the higher affinity binding site encompasses the active site and the WPD loop. Flexible docking of pTyr in the L7 loop was performed using the same protocol described above, with the only exception that the binding box adapted to the binding site had a size of 17.5 Å × 11.9 Å × 10 Å with a mesh resolution of 0.25 Å. Similarly, for flexible docking of pTyr in the active site and WPD loop, the binding box had a size of 22.7 Å × 16.6 Å × 17 Å with a mesh resolution of 0.25 Å. Before QM/MM calculations were performed, molecular complexes were prepared using the Amber 11 suite of programs.[55,56] The solute was included in the calculations using the Amber 11 utility, tLeap, with the macromolecular force field ff99sb.[57] Since Amber 11 does not recognize ligands as part of the solute by default, the atom types, charges, and connectivity in the pTyr structure were parameterized to include pTyr in the calculations. The atomic charges in the substrate were assigned using the restrained electrostatic potential (RESP) fitting technique,[58] implemented in AmberTools 1.5,[55,56] which was applied to the pTyr ESP obtained with Gaussian 98 at the HF/6-31G level of theory.[43] The assignment of other parameters was performed using the general Amber force field (GAFF),[59,60] which is compatible with the ff99sb force field used for YopH.[57] Solute solvation was simulated using a 10 Å pad of about 14000 TIP3P water molecules. A two-step geometric minimization was performed using the Sander module of Amber 11 to eliminate potential collisions and bad contacts.[55,56] The positions of the solute atoms were initially fixed, while the positions of the water atoms were minimized by reducing the initial harmonic constraint of 2 kcal mol^{-1} $Å^{-2}$ on all nonhydrogen, nonwater atoms using 2500 steepest descent and 2500 conjugate gradient minimization steps. Consequently, the entire system was minimized without constraints by performing a total of 10000 steps - 5000 steepest descent and 5000 conjugate gradient minimization steps.

The central part of the complex was treated using the QM method - the self-consistent charge density-functional tight-binding (SCC-DFTB) method[61,62] with dispersion correction (D),[63] as implemented in Amber 11.[64,65] An electron temperature of 100 K was set to accelerate the convergence of SCC in the SCC-DFTB-D calculations. An empirical correction that takes into account the dispersion energy has been added to SCC-DFTB to provide a more balanced description of interactions within the system.[66] Density functional theory (DFT) based on empirical density functionals has been recognized for its decent computational cost and inability to properly describe dispersive forces within unconventional biomolecular systems.[67,68] Extended DFT with dispersion correction (DFT-D) has enabled energy minimization and vibrational analysis of extended molecular systems with hundreds of atoms.[69,70] Being several orders of magnitude faster than DFT-D,[49,66,71–73] SCC-DFTB-D is an optimal choice for performing quantum chemical simulations of extended molecular systems.[66] Therefore, in the case of

the lower affinity state, the QM region encompassing the L7 loop and the substrate was treated using SCC-DFTB-D. In the case of the higher affinity state, the QM region containing the catalytic binding site, the WPD loop and the substrate was treated using SCC-DFTB-D. The remaining parts of the complexes were treated at the MM level using the Amber force field ff99sb.[57] Atoms in the MM layer bonded to atoms in the QM layer were replaced by hydrogen atoms during the higher-level part of the calculations. The free energy of binding was calculated as the difference in free energy between the bound and completely unbound states. Since enthalpy calculation is often impractical and difficult to converge,[74] enthalpy was estimated using the molecular mechanics-generalized Born solvent-accessible surface area (MM-GBSA) model[75] in single-point energy calculations on optimized geometries performed by the Sander module under Amber 11 for the complex, substrate, and receptor, respectively.[55,56] To determine the entropy and the different entropy contributions (translation, rotation and vibration), the normal mode analysis[76] was based on frequency calculations performed by the Nmode module of Amber 11.[55,56] All calculations were performed using a computer with a dual-core Intel Core i3 processor (2 GHz), a 64-bit Windows 10 Pro operating system, and 8 GB of RAM. The generated substrate binding conformations are shown in Figures S1 and S2, while all contributions to the binding free energy are given in the captions (Supplementary Information).

RESULTS AND DISCUSSION

Although it is tempting to define the specificity factor by considering all targets to which a ligand could bind, such an approach has some practical limitations: it is almost impossible to know all the biological targets to which a ligand could bind, the concentrations of the targets (secondary receptors), and the corresponding dissociation constants. The selection of specific mammalian PTPs for study here is based on available experimental values from the literature,[28] both the dissociation constants of the ATA:PTP complexes and the total concentrations of PTPs, which allows the analysis of the specificity factor, α, as a function of the free ligand concentration, [L], using equation 27. Although these mammalian PTPs are distinct proteins with different functions, and their expression and activity can be studied in different cellular and biochemical contexts, it is quite feasible and common to encounter them in the same experimental sample, either in a cellular lysate, a homogeneous tissue, or a purified protein mixture.

The specificity of ATA is examined here under the assumption of experimental conditions under which the binding process between ATA and PTP was previously examined.[28] All measurements were performed at pH 7 and constant temperature (303 K), using *p*-Nitrophenyl phosphate (*p*NPP) as a substrate. The total concentration of each PTP was kept at 10 nM. The IC_{50} value was measured at fixed substrate concentration that is equal to the experimentally determined value of the Michaelis constant (K_m) for each PTP. Although kinetics and enzymology were not the focus of that study, in the case of YopH, ATA was shown to be a competitive inhibitor.[28] Reported dissociation constants, K_D, are 5, 30.5, 125, 125, 200, 500, and 600 nM for YopH, PTP1B, CD45, HePTP, TCPTP, Cdc25A, and VHR, respectively.[28]

The specificity factor, α, as a function of the logarithm of free ATA concentration for each PTP being primary receptor, while all the other PTPs are simultaneous secondary receptors, is shown in Figure 2. At [ATA] between 10^{-11} and 10^{-9} M, α has an approximately constant value, which is about 3.5, 0.15, 0.03, 0.03, 0.02, 0.01 and 0.007 for YopH, PTP1B, CD45, HePTP, TCPTP, Cdc25A and

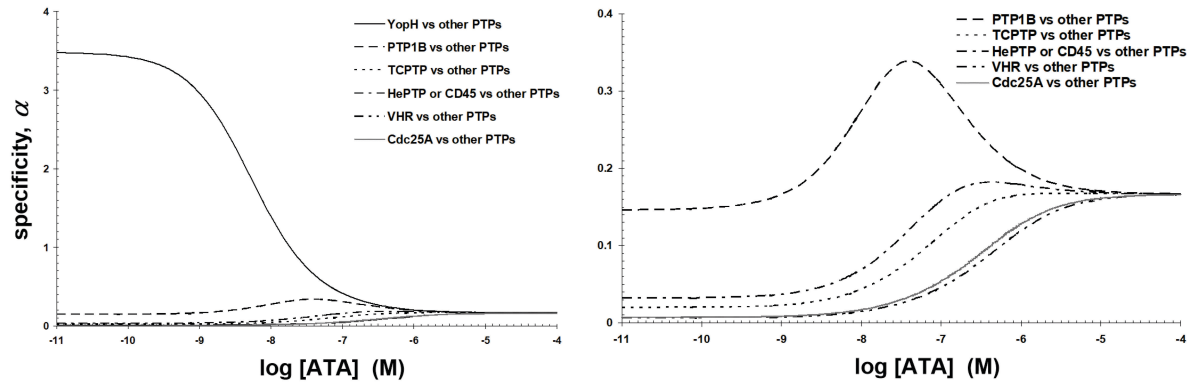

Figure 2. (Left) The specificity factor (α) vs. log [ATA] for each PTP as a primary receptor vs. all the other ones as simultaneous secondary receptors, indicating the greatest specificity of the ATA-YopH interaction. (Right) The adjusted scale of α makes the plots for PTP1B, TCPTP, HePTP, CD45, VHR, and Cdc25A visible.

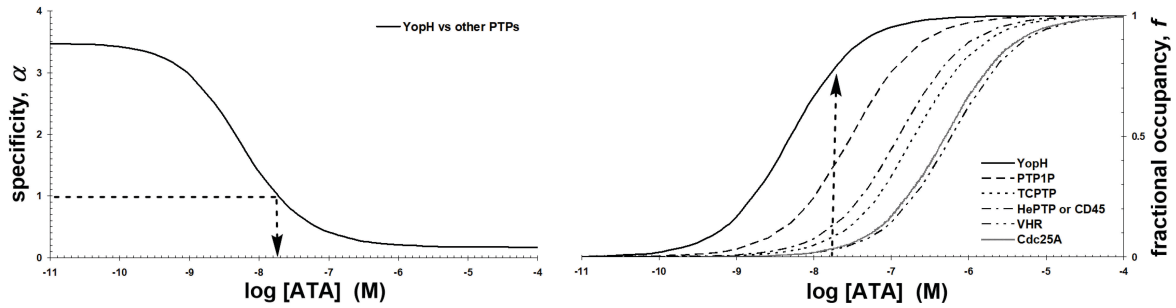

Figure 3. The specificity factor (α) ≈ 1 at [ATA] around 10^{-8} M (Left) is associated with equal occupancy of the primary receptor YopH and the set of secondary receptors (Right).

VHR as the primary receptor, respectively. This 23 to 500-fold specificity in favor of YopH versus the other PTPs indicates that the ATA-YopH interaction is the most specific. When [ATA] >> $K_{D,PR}$ and $K_{D,SRi}$, e.g. [ATA] > 10^{-5} M, equation 27 approximately gives $\alpha \approx [PR]_{total} \Big/ \sum_{i=1}^{N} [SR_i]_{total}$, which means that the specificity factor converges to 1/6 ≈ 0.167 for equal total concentration of each PTP (Figure 2).

The theoretical basis described in the previous section is based on two conditions that must be met for specific binding. First, the dissociation constant for the primary target receptor must be smaller than or equal to the ligand concentration. Second, the dissociation constants for secondary or off-target receptors must be larger than or equal to the ligand concentration.[39] Equation 27 shows that the value of α is large when the numerator is large but the denominator small. If the primary (PR) and secondary (SR_i) receptors are present at the same concentration, α will be large when $K_{D,PR} \leq [L] \leq K_{D,SRi}$. In this range PR is occupied substantially, but SR_i (i = 1,N) are not. This range of [ATA] is only satisfied by the dissociation constants when YopH is PR and the other PTPs are SR_i. As the ATA concentration increases above 10^{-9} M, an increasing amount of ATA binds to the secondary receptors, and the specificity of ATA against YopH decreases (Figure 3, Left). At [ATA] circa 10^{-8} M, the value of α of unity (Figure 3, Left) is associated with equal occupancy of the primary receptor YopH and the set of secondary receptors (Figure 3, Right).

To further dissect ATA specificity issue, the case in which there are just two receptors, YopH as PR and every other single PTP as SR_1, has been considered. The value of α as a function of free ATA concentration is graphed in Figure 4. At low ligand concentration, the primary receptor prevails in terms of fractional occupancy and the specificity is high. Note the values of α at [ATA] between 10^{-11} and 10^{-9} M: 6, 25, 25, 40, 100, and 120 for YopH vs. PTP1B, HePTP, CD45, TCPTP, Cdc25A, and VHR, respectively (Figure 4). These particular values are essentially the reciprocal values of those obtained via dividing the dissociation constant for PR by the dissociation constants for SR_1, respectively. An explanation is based on equation 27 that, in case of the presence of only two receptors - PR and SR_1, becomes:

$$\alpha = \frac{\dfrac{[L]}{K_{D,PR} + [L]}[PR]_{total}}{\dfrac{[L]}{K_{D,SR_1} + [L]}[SR_1]_{total}} \quad (29)$$

At the same total concentrations of PR and SR_1, $[PR]_{total}$ and $[SR_1]_{total}$, equation 29 gives:

$$\alpha = \frac{\dfrac{1}{K_{D,PR} + [L]}}{\dfrac{1}{K_{D,SR_1} + [L]}} \quad (30)$$

At low ligand concentration, [L] << $K_{D,SR1}$, α is approximately:

$$\alpha \approx \frac{K_{D,SR_1}}{K_{D,PR}} \quad (31)$$

Therefore, at [ATA] between 10^{-11} and 10^{-9} M, the dissociation constant for PR, $K_{D,PR}$ = $K_{D,YopH}$ = 5 nM divided by the dissociation constant for SR_1, $K_{D,SR1}$ (30.5 nM for PTP1B, 125 nM for HePTP, 125 nM for CD45, 200 nM for TCPTP, 500 nM for Cdc25A, and 600 nM for VHR), gives the reciprocal values of α, respectively (Figure 4). As the ligand concentration increases, an increasing amount of ATA binds to SR_1, and the specificity decreases.

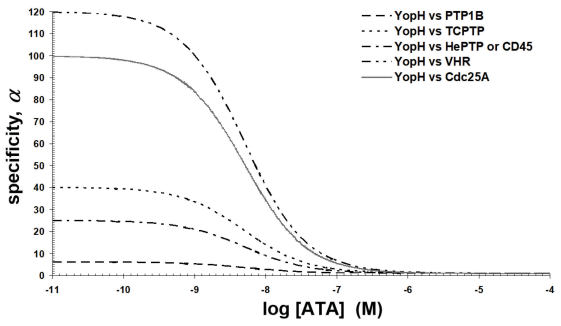

Figure 4. The specificity factor (α) vs. log [ATA] for YopH as a primary receptor vs. every other single PTP as a secondary receptor.

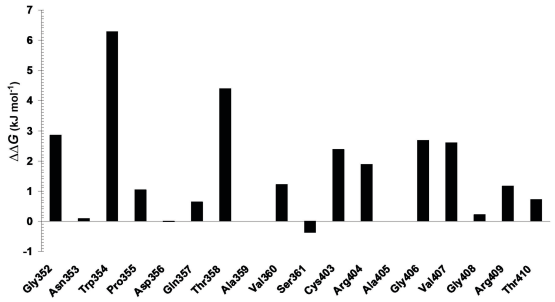

Figure 5. The free energy distribution of binding at the ATA-YopH interface (active site and WPD loop) was determined by alanine scanning mutagenesis. The usual cutoff value of 4.184 kJ mol^{-1} (1 kcal mol^{-1}) in $\Delta\Delta G$ indicates that Trp354 and Thr358 dominate the ATA-YopH interaction.

PTPs are, most likely, the first enzyme family known that executes the same reaction at highly different rates modulated by differences in their conformational dynamics, as shown by comparing the protein dynamics of YopH to that of PTP1B.[77] Molecular dynamics (MD) approach and related metrics revealed characterization of the YopH activity by the closure of the so-called "WPD loop" (L6: 352-Gly-Asn-Trp-Pro-Asp-Gln-Thr-Ala-Val-Ser-361, extending between β7 and α4, Figure 1, Left) upon ligand binding,[37,77] thereby speaking in favor of a previous indication that the Asp356-containing loop movement appears critical for YopH catalysis.[30] Since the co-crystal structure of ATA:YopH has not yet been determined, the binding mechanism was investigated using biomolecular modeling. MD simulations showed that ATA binds stronger to the active site (L8: 403-Cys-Arg-Ala-Gly-Val-Gly-Arg-Thr-410, extending between β8 and α5, Figure 1, Left) than does the natural substrate such as pTyr, indicating the ability of ATA to inhibit YopH activity competitively.[78,79] Sophisticated QM/MM insights into molecular basis of the binding mechanism confirmed that the ATA-induced open-closed conformational change of WPD loop is associated with YopH inhibition.[29] Therefore, herein, the relative binding free energy contributions of all residues, located in both the active site and its proximal WPD loop, are evaluated using alanine scanning mutagenesis.

In order to determine hot spots of binding free energy in the ATA-YopH interface, interfacial residues were mutated individually to alanine and the change in binding free energy was calculated for each single-point mutant relative to wt YopH as described in the Methods section. A cutoff value of 4.184 kJ mol^{-1} (1 kcal mol^{-1}) in $\Delta\Delta G$ is commonly used to distinguish between alanine mutations that significantly affect protein function or binding and those that do not. Figure 5 illustrates that the hot spots of binding affinity that dominate the ATA-YopH interaction are Trp354 and Thr358. This is in agreement with the nature and atomic details of the recognition between ATA and YopH.[29] The importance of the conformational flexibility of the Asp356-containing (or WPD) loop for YopH catalysis means that variability in its amino acid sequence might affect its dynamics influencing the catalytic rate.[80] Figure 5 shows that Ser361 from the WPD loop is slightly prone to destabilizing the ATA:YopH complex. Although the specific effect of Ser361 is energetically below 1 kcal mol^{-1}, this indication should not be underestimated, as the stability of the WPD loop in the closed, active conformation is essential for YopH function. Therefore, some experimental studies based on isothermal titration calorimetry could determine which residues of the WPD loop weaken the interaction, thereby proving whether Ser361 plays a vital role in ATA binding or not. Replacing Ser with another amino acid, such as Asp, Thr or Ala, can drastically change the structure and stability of the protein because these new residues have different chemical properties, including size, acidity and hydrogen bonding ability. In the case of Ser361, a Ser/Thr substitution might be most likely, as it occurs in the consensus sequence (402-HisCysXXGlyXGlyArgThr/Ser-410) of the PTP active site at position 410.[31]

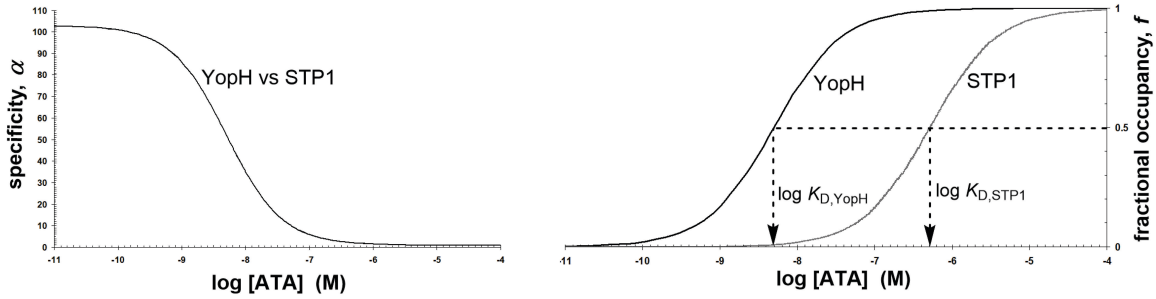

Figure 6. (Left) The specificity factor (α) vs. log [ATA] for YopH as a primary receptor vs. STP1 as a secondary receptor. At [ATA] between 10^{-11} and 10^{-9} M, α is about 103, which is the reciprocal value of that obtained by dividing the dissociation constant for YopH (5 nM) by the dissociation constant for STP1 (515 nM). (Right) The fractional occupancies of YopH and STP1 vs. log [ATA].

The biggest issue for phosphatase inhibitor identification is the lack of specificity. For example, by comparing the crystal structure of the *Yersinia pestis* protein Tyr phosphatase - YopH (PDB ID: 1YPT) to that of the *Staphylococcus aureus* Ser/Thr phosphatase - STP1 (PDB ID: 5F1M), there is no base to claim that ATA can inhibit STP1 because of its ability to inhibit YopH. Although YopH and STP1 are distinct proteins with different functions and origins, they can be present in the same sample if, for example, a sample is derived from a mixed culture of *Yersinia* and *Staphylococcus aureus* or if both proteins are purified and combined in a biochemical experiment, which is often one of the common procedures in the drug design and development process. ATA is a mixed-type inhibitor of PTP and STP1,[32] so caution is needed regarding the assumption that the IC_{50} value is twice the equilibrium dissociation constant of the ATA:PTP complex, which was not tested in the original study.[28] Given that ATA is capable of directly binding to STP1 and that the co-crystal structure of ATA:STP1 has not yet been determined,[32] the inhibitory mechanism has been studied at the molecular level using MD and related metrics that predict mechanistic details of the ATA-STP1 interaction[81] that are consistent with the mechanistic details of the ATA-YopH interaction.[29] Although there is no YopH homolog in *Staphylococcus aureus*, the off-target effects of ATA cannot be precluded.[32] Thus, at free ATA concentration between 10^{-11} and 10^{-9} M, the specificity factor, α, as a function of log [ATA] indicates a roughly 100-fold preference for YopH over STP1 (Figure 6, Left). At low free ATA concentration, most of ATA are associated with the primary receptor, YopH, and the specificity is high (Figure 6, Right). As free ATA concentration increases, an increasing amount of ATA binds to the secondary receptor, STP1, and the specificity decreases, reaching a value of 1 at high free ATA concentration. This can be seen through the ratio of $K_{D,YopH}$ to $K_{D,STP1}$, which is circa 10^{-2} (Figure 6, Right), meaning that the affinity of the primary receptor is much higher than that of the secondary receptor. If the difference between the affinities is not so large, that is, if the values of $K_{D,PR}$ and $K_{D,SRi}$ are of the similar order of magnitude, then it is not possible to observe high specificity, unless off-target receptors are only present at very low concentration.

This work aids in understanding the potential of structural changes in ATA. If such a structural modification improves the steric fit of ATA to the primary receptor, or enhances the charge complementarity to it, then the affinity of ATA for PR will increase, or $K_{D,PR}$ will become smaller. Such a change may result in an increase in specificity if the altered interactions do not also increase the affinity of ATA for the secondary receptors by an equal amount. The structural modifications of ATA that widen the gap between $K_{D,PR}$ and $K_{D,SRi}$ and enhance specificity are more preferable than those that reduce this gap and lower specificity. However, the feasibility of such structural changes is not likely, because ATA structure modifications presumably lead to reduction of inhibitory properties against YopH and higher cytotoxicity.[79] Caution is needed regarding these indications, as only two structural analogs of ATA, pararosaniline and eriochrome cyanine R, have been considered,[79] although they can be considered representative examples of structurally diverse ATA analogs to a large extent. Thus, the present study speaks in favor of the ability of ATA to display some structural features of phosphatase inhibitor that are recognized at receptor site and responsible for inhibitory activity. It means that ATA structure may be viewed as a pharmacophore for identification of phosphatase inhibitor candidates capable of achieving specificity.

Crystallographic experiments have shown that the L7 loop region in the YopH structure represents a second substrate binding site.[33] Experimental understanding of the basis of the L7 loop region's involvement in promoting substrate recognition at the catalytic active site has been hampered by the fact that the second substrate binding site is far from the catalytic binding site, and the underlying conformational changes could not be inferred from static crystal structures.[33] However, MD simulations revealed that the most extensive conformational changes occur in the L7 loop region extending to helix α4 (Figure 1, Left), and the ordered helix α4 was found to correlate with the closure of the WPD loop (L6 in Figure 1, Left), indicating a significant coupling of the conformational flexibility of the L7 loop with the movements of the WPD loop.[37] In other words, MD simulations suggest that substrate binding at the second substrate binding site may induce a more ordered shape in helix α4, thereby stabilizing the movement of the WPD loop in the closed, active conformation.[37]

To quantitatively characterize the association between the L7 and WPD loops in substrate (pTyr) recognition, the binding free energies for the two substrate binding sites were calculated using the QM/MM (SCC-DFTB-D/AMBER) strategy described in the previous section and converted to K_{D1} and K_{D2} values using equation 28. The calculated values of –4.9 and –10.1 kcal mol^{-1} (Supplementary Information) correspond to values of approximately K_{D1} = 0.3 mM and K_{D2} = 46.7 nM for the lower affinity site (L7 loop) and the higher affinity site (active site and WPD loop), respectively.

It is convenient to denote the higher-affinity site as the "relaxed" or R state of the system and the lower-affinity site as the "tense" or T state of the system. A binding isotherm is a plot of the fractional occupancy as a function of the free ligand concentration at constant temperature. The T- and R-state binding isotherms, based on equation

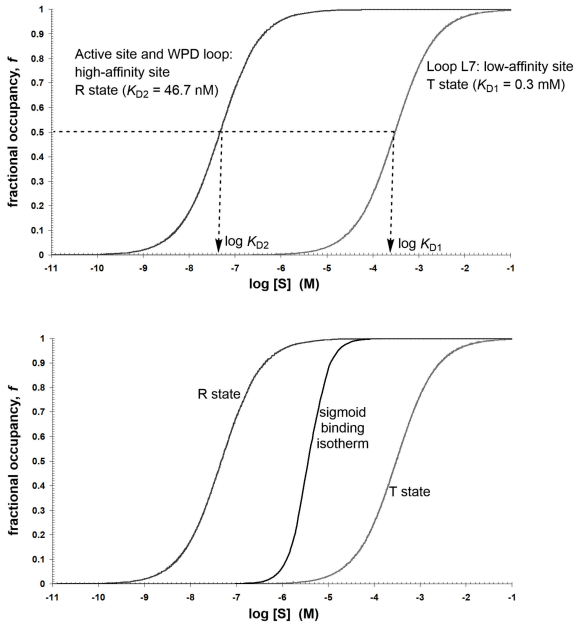

Figure 7. (Top) Binding isotherms for substrate (S: pTyr) recognition by YopH in the low-affinity or T state (loop L7: 384-Glu-Ser-Lys-Gly-Ser-Ser-Ala-Val-Ala-392) and high-affinity or R state (active site (L8: 403-Cys-Arg-Ala-Gly-Val-Gly-Arg-Thr-410) and WPD loop (L6: 352-Gly-Asn-Trp-Pro-Asp-Gln-Thr-Ala-Val-Ser-361)). (Bottom) The net sigmoid binding isotherm for allosteric system is a combination of the T- and R-state isotherms.

19, are shown in Figure 7 (Top). Substrate concentrations at which YopH is half saturated correspond to the values of K_{D1} and K_{D2}. If the binding of one ligand molecule in the T state does not affect the binding of the others in the R state, then the binding events are considered to be independent. However, when the T and R states do not follow a graded response associated with a simple binding equilibrium, but are rather prone to an ultrasensitive response, YopH occupancy f rises from 0.1 to 0.9 over a substrate concentrations range of less than 100-fold (Figure 7, Bottom). In other words, the T and R states tend to switch from the unbound - off to the bound - on state (from 10 % to 90 % of the maximum response) for a change in substrate concentration of less than 100-fold.

When the binding of a ligand to a receptor is ultrasensitive, the binding is cooperative. It means that, if more ligand molecules bind to the receptor, then the saturation of the receptor increases more sharply than would be expected for a normal binding event. A sigmoid rather than a hyperbolic binding curve results from cooperativity in binding events. Cooperative binding, in principle, does not always cause conformational changes in the receptor or the ligand.[40] A binding curve as a combination of the T- and R-state isotherms is based on the following model of two sequential equilibria:

$$PR + L \underset{K_{D1}}{\leftrightarrow} PR:L \tag{32}$$

$$PR:L + L \underset{K_{D2}}{\leftrightarrow} PR:L_2 \tag{33}$$

The fractional occupancy of protein is given by:

$$f_{PR} = \frac{[PR:L] + 2[PR:L_2]}{2([PR] + [PR:L] + [PR:L_2])} \tag{34}$$

Based on equation 32, it follows:

$$[PR:L] = \frac{[PR][L]}{K_{D1}} \tag{35}$$

Based on equation 33, it follows:

$$[PR:L_2] = \frac{[PR:L][L]}{K_{D2}} \tag{36}$$

Substitution of the expression for [PR:L] from equation 35 into equation 36 gives:

$$[PR:L_2] = \frac{[PR][L]}{K_{D1}} \frac{[L]}{K_{D2}} \tag{37}$$

By using equations 35 and 37, equation 34 becomes:

$$f_{PR} = \frac{\frac{[PR][L]}{K_{D1}} + 2\frac{[PR][L]}{K_{D1}}\frac{[L]}{K_{D2}}}{2([PR] + \frac{[PR][L]}{K_{D1}} + \frac{[PR][L]}{K_{D1}}\frac{[L]}{K_{D2}})} \tag{38}$$

Rearranging the right-hand side of equation 38 gives:

$$f_{PR} = \frac{\frac{[L]}{K_{D1}} + 2\frac{[L]}{K_{D1}}\frac{[L]}{K_{D2}}}{2 + 2\frac{[L]}{K_{D1}} + 2\frac{[L]}{K_{D1}}\frac{[L]}{K_{D2}}} \tag{39}$$

When PR = YopH, L = S (pTyr), K_{D1} = 0.3 mM and K_{D2} = 46.7 nM, equation 39 defines a sigmoid binding isotherm that is a combination of the T- and R-state isotherms (Figure 7, Bottom). The possibility of transmitting the conformational change from one binding site to another is an example of allostery in a receptor. The activity of an allosteric protein is modulated by interactions that occur at a distance from the active site. In addition to the catalytic center or active site (K_{D2}), an allosteric protein has another allosteric site (K_{D1}) where a ligand binds and modifies the protein structure through covalent interactions and/or other types of interactions. In this sense, the sigmoid binding curve shown in Figure 7 (Bottom) points to L7 as the allosteric site where pTyr binds and modifies

the YopH structure, mainly through electrostatic interactions (Figure S2, Supplementary Information).

The value of f usually changes from low to high over concentration range that is relatively narrow, so that the most representative data points are quite often crowded together on the left side of the f vs. [L] graph, and thus make it difficult to read off the value of f by visual inspection of the binding isotherm. To expand this region of the graph, it is convenient to observe the ratio of the fraction of protein that is bound to the fraction that is unbound. Using equation 19 gives:

$$\frac{\text{fraction of bound PR}}{\text{fraction of unbound PR}} = \frac{f_{PR}}{1-f_{PR}} = \frac{\dfrac{[L]}{K_{D,PR}+[L]}}{1-\left(\dfrac{[L]}{K_{D,PR}+[L]}\right)} = \frac{[L]}{K_{D,PR}} \quad (40)$$

The logarithm of both sides of equation 40 gives:

$$\log\left(\frac{f_{PR}}{1-f_{PR}}\right) = \log\left(\frac{[L]}{K_{D,PR}}\right) = \log[L] - \log K_{D,PR} \quad (41)$$

For L = S (pTyr), plotting $\log(f/1-f)$ as a function of $\log[S]$ yields a linear plot with a slope of unity. If the receptor is half saturated, then $(f/1-f) = 1$, that is, $\log(f/1-f) = 0$.

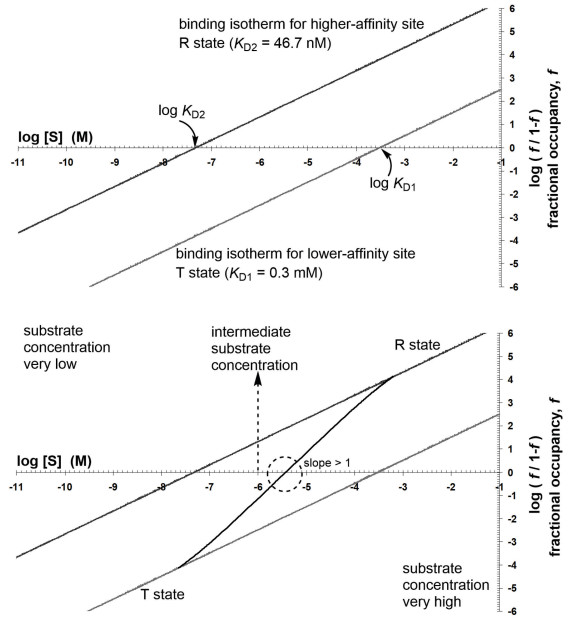

Figure 8. (Top) Binding isotherms for the T and R states are shown as a log (f/1–f) – log [S] plot, where f denotes the fraction of YopH receptors occupied by the substrate (S: pTyr). (Bottom) As the S concentration increases, the binding isotherm switches from the T state isotherm to the R state isotherm. When log (f/1–f) = 0 at intermediate substrate concentration, the slope of the resulting binding isotherm is > 1.

Therefore, the intercept of the straight line on the horizontal axis is equal to log K_D. The binding isotherms for the T and R states are shown as a log (f/1–f) – log[S] plot in Figure 8 (Top). The slope of the log-log binding isotherm at the point when the receptor is half saturated with ligand is known as the Hill coefficient. If the actual binding isotherm is not linear ($K_{D1} \neq K_{D2}$), or has a slope different from unity, then this is an indication of more complex binding process due to some other factors such as allostery. Based on equation 39, the combination of the T- and R-state binding curves gives:

$$\frac{\text{fraction of bound PR}}{\text{fraction of unbound PR}} = \frac{f_{PR}}{1-f_{PR}} = \frac{\dfrac{[L]}{K_{D1}} + 2\dfrac{[L]}{K_{D1}}\dfrac{[L]}{K_{D2}}}{2+\dfrac{[L]}{K_{D1}}} \quad (42)$$

The logarithm of both sides of equation 42 gives:

$$\log\left(\frac{f_{PR}}{1-f_{PR}}\right) = \log\left(\frac{\dfrac{[L]}{K_{D1}} + 2\dfrac{[L]}{K_{D1}}\dfrac{[L]}{K_{D2}}}{2+\dfrac{[L]}{K_{D1}}}\right) = \log\left(\frac{[L]}{K_{D1}} + 2\frac{[L]}{K_{D1}}\frac{[L]}{K_{D2}}\right) - \log\left(2+\frac{[L]}{K_{D1}}\right) \quad (43)$$

When PR = YopH, L = S (pTyr), K_{D1} = 0.3 mM and K_{D2} = 46.7 nM, equation 43 defines a resulting binding isotherm that is no longer linear (Figure 8, Bottom). As the S concentration increases, the binding isotherm switches from the T state isotherm to the R state isotherm. When $\log(f/1-f) = 0$ at intermediate substrate concentration (between 10^{-6} and 10^{-5} M), the slope (Hill coefficient) of the resulting binding isotherm is > 1. The expression that quantifies the Hill coefficient[40] is mathematically derived in the Supplementary Information and is given by:

$$n_H = \left.\frac{d\log\left(\dfrac{f}{1-f}\right)}{d\log[L]}\right|_{\substack{\text{evaluated at}\\ f=0.5}} = \frac{2}{1+\dfrac{1}{2}\sqrt{\dfrac{K_{D2}}{K_{D1}}}} \quad (44)$$

Therefore, the Hill coefficient for the resulting binding isotherm turns out to be about 1.98, indicating positive cooperation between the T and R states in substrate recognition.

The paucity of available experimental data on the catalytic efficiency of enzyme substrates often means that a more complete understanding of the actual substrate preferences is difficult to achieve. The present work, together with some previous studies,[29,71,82] illustrates how additional experiments can be stimulated by biomolecular modeling to better understand the kinetics of ligand/enzyme recognition for further research or therapeutic product development.[83]

CONCLUSIONS

Among structurally diversified carboxylic acid derivatives, aurintricarboxylic acid had been identified using both experimental methods[28] and biomolecular modeling[29] as a direct binder having the largest affinity to YopH. In this study, specificity issue, associated with the competitive inhibition of some protein Tyr phosphatases (YopH, PTP1B, TCPTP, HePTP, CD45, VHR, and Cdc25A) by ATA, was addressed.

By observing each PTP as a primary target receptor and the other ones as simultaneous off-target receptors, the specificity factor, as a function of free ATA concentration, demonstrated a 23 to 500-fold specificity in favor of YopH versus the other PTPs at [ATA] between 10^{-11} and 10^{-9} M, meaning that the ATA-YopH interaction is the most specific. At [ATA] around 10^{-8} M, the specificity factor ≈ 1 was lined up with equal occupancy of the primary receptor - YopH and the set of accessory receptors. When YopH was considered as a primary receptor with respect to every other single PTP as a secondary receptor at [ATA] ϵ (10^{-11}, 10^{-9} M), the specificity to YopH was larger nearly 6, 25, 25, 40, 100, and 120 times than that to PTP1B, HePTP, CD45, TCPTP, Cdc25A, and VHR, respectively.

By applying alanine scanning mutagenesis to YopH residues in the active site (403-Cys-Arg-Ala-Gly-Val-Gly-Arg-Thr-410) and in the proximal WPD loop (352-Gly-Asn-Trp-Pro-Asp-Gln-Thr-Ala-Val-Ser-361), it was found that the spots of binding that dominate the ATA-YopH interaction are Trp354 and Thr358 to a greatest extent.

The biggest issue for phosphatase inhibitor identification is the lack of specificity. For example, by comparing the crystal structures of the protein Tyr phosphatase YopH (PDB ID: 1YPT) and the Ser/Thr phosphatase STP1 (PDB ID: 5F1M), there is no base to claim that ATA can inhibit STP1 because of its inhibition of YopH. However, the direct binding of ATA to STP1 had been observed experimentally.[32] Thus, the specificity factor indicated an approximately 100-fold preference for YopH versus STP1. It speaks in favor of the ability of ATA to display some structural features of potential phosphatase inhibitor that are recognized at receptor site and responsible for inhibitory activity. In other words, the present study comes up with the ATA structure as a pharmacophore for identifying phosphatase inhibitor candidates capable of achieving specificity.

The Hill coefficient of the binding isotherm $\log(f/1–f) – \log[S]$, where f denotes the fraction of the YopH receptors occupied by the substrate, indicated a positive cooperation between the L7 loop and the WPD loop in pTyr recognition, suggesting that L7 is an allosteric site in the YopH structure.

Acknowledgments. Prof. David A. Case of Rutgers University is acknowledged for granting the author an academic license for using the molecular dynamics software package Amber 11 in combination with AmberTools 1.5. The very first working version of the present manuscript is available at ChemRxiv as two unpublished preprints.[84,85]

Conflict of Interest. The author declares that he has no conflict of interest.

Supplementary Information. Supporting information to the paper is attached to the electronic version of the article at: https://doi.org/10.5562/cca4168.

PDF files with attached documents are best viewed with Adobe Acrobat Reader which is free and can be downloaded from Adobe's web site.

REFERENCES

[1] A. Bairoch, *Nucleic Acids Res.* **2000**, *28*(1), 304–305. https://doi.org/10.1093/nar/28.1.304

[2] S. Liberti, F. Sacco, A. Calderone, L. Perfetto, M. Iannuccelli, S. Panni, E. Santonico, A. Palma, A. P. Nardozza, L. Castagnoli, G. Cesareni, *FEBS J.* **2013**, *280*(2), 379–387. https://doi.org/10.1111/j.1742-4658.2012.08712.x

[3] F. Sacco, L. Perfetto, L. Castagnoli, G. Cesareni, *FEBS Lett.* **2012**, *586*(17), 2732–2739. https://doi.org/10.1016/j.febslet.2012.05.008

[4] R. He, L. F. Zeng, Y. He, S. Zhang, Z. Y. Zhang, *FEBS J.* **2013**, *280*(2), 731–750. https://doi.org/10.1111/j.1742-4658.2012.08718.x

[5] P. Gupta, F. I. Khan, D. Ambreen, D. Lai, M. F. Alajmi, A. Hussain, A. Islam, F. Ahmad, M. I. Hassan, *Int. J. Biol. Macromol.* **2020**, *147*, 177–186. https://doi.org/10.1016/j.ijbiomac.2020.01.023

[6] P. Gupta, F. I. Khan, S. Roy, S. Anwar, R. Dahiya, M. F. Alajmi, A. Hussain, M. T. Rehman, D. Lai, M. I. Hassan, *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **2020**, *225*, 117453. https://doi.org/10.1016/j.saa.2019.117453

[7] F. I. Khan, P. Gupta, S. Roy, N. Azum, K. A. Alamry, A. M. Asiri, D. Lai, M. I. Hassan, *Int. J. Biol. Macromol.* **2020**, *161*, 1496–1505. https://doi.org/10.1016/j.ijbiomac.2020.07.280

[8] F. I. Khan, S. Ali, W. Chen, F. Anjum, A. Shafie, M. I. Hassan, D. Lai, *Curr. Top. Med. Chem.* **2021**, *21*(31), 2839–2850. https://doi.org/10.2174/1568026621666211105095731

[9] F. Naz, F. I. Khan, T. Mohammad, P. Khan, S. Manzoor, G. M. Hasan, K. A. Lobb, S. Luqman, A. Islam, F. Ahmad, M. I. Hassan, *Int. J. Biol. Macromol.* **2018**, *107*(Part B), 2580–2589. https://doi.org/10.1016/j.ijbiomac.2017.10.143

[10] F. Xie, H. Dong, H. Zhang, *Front. Immunol.* **2021**, *12*, 783370. https://doi.org/10.3389/fimmu.2021.783370
[11] H. Yang, L. Wang, C. Shigley, W. Yang, *Bone Res.* **2022**, *10*(1), 10. https://doi.org/10.1038/s41413-021-00181-x
[12] T. Mohammad, F. I. Khan, K. A. Lobb, A. Islam, F. Ahmad, M. I. Hassan, *J. Biomol. Struct. Dyn.* **2019**, *37*(7), 1813–1829. https://doi.org/10.1080/07391102.2018.1468282
[13] S. B. Syed, F. I. Khan, S. H. Khan, S. Srivastava, G. M. Hasan, K. A. Lobb, A. Islam, F. Ahmad, M. I. Hassan, *Int. J. Biol. Macromol.* **2018**, *111*, 208–218. https://doi.org/10.1016/j.ijbiomac.2017.12.164
[14] D. Kołodziej-Sobczak, Ł. Sobczak, K. Z. Łączkowski, *Int. J. Mol. Sci.* **2024**, *25*(13), 7033. https://doi.org/10.3390/ijms25137033
[15] C. L. Welsh, P. Pandey, L. G. Ahuja, *Adv. Cancer Res.* **2021**, *152*, 263–303. https://doi.org/10.1016/bs.acr.2021.06.001
[16] F. I. Khan, D. Lai, R. Anwer, I. Azim, M. K. A. Khan, *J. Enzyme Inhib. Med. Chem.* **2020**, *35*(1), 172–186. https://doi.org/10.1080/14756366.2019.1692828
[17] L. M. Kelam, V. Chhabra, S. Dhiman, D. Kumari, M. E. Sobhia, *Expert Opin. Ther. Pat.* **2024**, *34*(4), 187–209. https://doi.org/10.1080/13543776.2024.2362203
[18] Z. H. Yu, Z. Y. Zhang, *Chem. Rev.* **2018**, *118*(3), 1069–1091. https://doi.org/10.1021/acs.chemrev.7b00105
[19] B. Bodenmiller, S. Wanka, C. Kraft, J. Urban, D. Campbell, P. G. Pedrioli, B. Gerrits, P. Picotti, H. Lam, O. Vitek, M. Y. Brusniak, B. Roschitzki, C. Zhang, K. M. Shokat, R. Schlapbach, A. Colman-Lerner, G. P. Nolan, A. I. Nesvizhskii, M. Peter, R. Loewith, C. Von Mering, R. Aebersold, *Sci. Signal.* **2010**, *3*(153), rs4. https://doi.org/10.1126/scisignal.2001182
[20] X. Li, M. Wilmanns, J. Thornton, M. Köhn, *Sci. Signal.* **2013**, *6*(275), rs10. https://doi.org/10.1126/scisignal.2003203
[21] B. J. Hinnebusch, *J. Mol. Med.* **1997**, *75*(9), 645–652. https://doi.org/10.1007/s001090050148
[22] R. W. Titball, S. E. Leary, *Br. Med. Bull.* **1998**, *54*(3), 625–633. https://doi.org/10.1093/oxfordjournals.bmb.a011715
[23] A. M. Friedlander, S. L. Welkos, P. L. Worsham, G. P. Andrews, D. G. Heath, G. W. Jr. Anderson, M. L. M. Pitt, J. Estep, K. Davis, *Clin. Infect. Dis.* **1995**, *21*(Suppl 2), S178–S181. https://doi.org/10.1093/clinids/21.Supplement_2.S178
[24] R. W. Titball, E. D. Williamson, *Vaccine* **2001**, *19*(30), 4175–4184. https://doi.org/10.1016/s0264-410x(01)00163-3
[25] M. L. de la Puerta, A. G. Trinidad, M. del Carmen Rodríguez, J. Bogetz, M. S. Crespo, T. Mustelin, A. Alonso, Y. Bayón, *PLoS One* **2009**, *4*(2), e4431. https://doi.org/10.1371/journal.pone.0004431
[26] R. B. Hallick, B. K. Chelm, P. W. Gray, E. M. Jr. Orozco, *Nucleic Acids Res.* **1977**, *4*(9), 3055–3064. https://doi.org/10.1093/nar/4.9.3055
[27] Y. Benchokroun, J. Couprie, A. K. Larsen, *Biochem. Pharmacol.* **1995**, *49*(3), 305–313. https://doi.org/10.1016/0006-2952(94)00465-X
[28] F. Liang, Z. Huang, S.-Y. Lee, J. Liang, M. I. Ivanov, A. Alonso, J. B. Bliska, D. S. Lawrence, T. Mustelin, Z.-Y. Zhang, *J. Biol. Chem.* **2003**, *278*(43), 41734–41741. https://doi.org/10.1074/jbc.m307152200
[29] P. M. Mitrasinovic, *J. Biomol. Struct. Dyn.* **2023**, *41*(5), 1879–1894. https://doi.org//10.1080/07391102.2021.2025148
[30] H. L. Schubert, E. B. Fauman, J. A. Stuckey, J. E. Dixon, M. A. Saper, *Protein Sci.* **1995**, *4*(9), 1904–1913. https://doi.org/10.1002/pro.5560040924
[31] Z. Y. Zhang, J. E. Dixon, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1994**, *68*, l–36. https://doi.org/10.1002/9780470123140.ch1
[32] W. Zheng, X. Cai, M. Xie, Y. Liang, T. Wang, Z. Li, *Cell Chem. Biol.* **2016**, *23*(8), 1002–1013. https://doi.org/10.1016/j.chembiol.2016.06.014
[33] M. I. Ivanov, J. A. Stuckey, H. L. Schubert, M. A. Saper, J. B. Bliska, *Mol. Microbiol.* **2005**, *55*(5), 1346–1356. https://doi.org/10.1111/j.1365-2958.2005.04477.x
[34] J. Phan, K. Lee, S. Cherry, J. E. Tropea, T. R. Burke Jr, D. S. Waugh, *Biochemistry* **2003**, *42*(45), 13113–13121. https://doi.org/10.1021/bi030156m
[35] J. A. Stuckey, H. L. Schubert, E. B. Fauman, Z. Y. Zhang, J. E. Dixon, M. A. Saper, *Nature* **1994**, *370*(6490), 571–575. https://doi.org/10.1038/370571a0
[36] E. B. Fauman, C. Yuvaniyama, H. L. Schubert, J. A. Stuckey, M. A. Saper, *J. Biol. Chem.* **1996**, *271*(31), 18780–18788. https://doi.org/10.1074/jbc.271.31.18780
[37] X. Hu, C. Erec Stebbins, *Biophys. J.* **2006**, *91*(3), 948–956. https://doi.org/10.1529/biophysj.105.080259
[38] J. Kuriyan, B. Konforti, D. Wemmer, *The Molecules of Life: Physical and Chemical Principles*, Garland Science, Taylor & Francis Group LLC, New York and London, USA and UK, **2013**, pp. 531–580.
[39] J. Kuriyan, B. Konforti, D. Wemmer, *The Molecules of Life: Physical and Chemical Principles, Garland Science*, Taylor & Francis Group LLC, New York and London, USA and UK, **2013**, pp. 581–631.
[40] J. Kuriyan, B. Konforti, D. Wemmer, *The Molecules of Life: Physical and Chemical Principles*, Garland Science, Taylor & Francis Group LLC, New York and London, USA and UK, **2013**, pp. 633–670.
[41] W. L. DeLano, *PyMol*, *version 0.99*, DeLano Scientific LLC, South San Francisco, CA, USA, **2006**.

[42] M. A. Thompson, *ArgusLab*, *version 4.0.1*, Planaria Software, Seattle, WA, USA, **2004**.

[43] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. A. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople, *Gaussian 98*, *revision A.9*, Gaussian, Inc., Pittsburgh, PA, USA, **1998**.

[44] M. L. Mihajlovic, P. M. Mitrasinovic, *Mol. Simulation* **2009**, *35*(4), 311–324. https://doi.org/10.1080/08927020802430752

[45] P. M. Mitrasinovic, *J. Comput. Chem.* **2001**, *22*(13), 1387–1395. https://doi.org/10.1002/jcc.1097

[46] P. M. Mitrasinovic, *J. Phys. Chem. A* **2002**, *106*(30), 7026–7033. https://doi.org/10.1021/jp025554y

[47] P. M. Mitrasinovic, *Chem. Phys.* **2003**, *286*(1), 1–13. https://doi.org/10.1016/S0301-0104(02)00902-3

[48] P. M. Mitrasinovic, *Bioconjug. Chem.* **2005**, *16*(3), 588–597. https://doi.org/10.1021/bc0500490

[49] P. M. Mitrasinovic, *Curr. Org. Chem.* **2010**, *14*(2), 198–211. https://doi.org/10.2174/138527210790069857

[50] P. M. Mitrasinovic, *Curr. Org. Synth.* **2012**, *9*(2), 233–246. https://doi.org/10.2174/157017912799829058

[51] P. M. Mitrasinovic, P. T., Palakshan, S. Tripathi, A. N. Tripathi, *Med. Chem.* **2013**, *9*(2), 193–202. https://doi.org/10.2174/1573406411309020004

[52] P. M. Mitrasinovic, *Med. Chem.* **2014**, *10*(1), 46–58. https://doi.org/10.2174/15734064100113226122124

[53] P. M. Mitrasinovic, *Med. Chem.* **2014**, *10*(3), 252–270. https://doi.org/10.2174/1573406410031403041434442

[54] P. M. Mitrasinovic, *J. Biomol. Struct. Dyn.* **2018**, *36*(9), 2292–2302. https://doi.org/10.1080/07391102.2017.1358670

[55] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*(16), 1668–1688. https://doi.org/10.1002/jcc.20290

[56] D. A. Case, T. A. Darden, T. E. III. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. P. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, *Amber 11*, University of California, San Francisco, CA, USA, **2010**.

[57] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, *65*(3), 712–725. https://doi.org/10.1002/prot.21123

[58] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*(40), 10269–10280. https://doi.org/10.1021/j100142a004

[59] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **2006**, *25*(2), 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005

[60] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*(9), 1157–1174. https://doi.org/10.1002/jcc.20035

[61] M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, E. Kaxiras, *J. Chem. Phys.* **2001**, *114*(12), 5149–5155. https://doi.org/10.1063/1.1329889

[62] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, *Phys. Rev. B* **1998**, *58*(11), 7260–7268. https://doi.org/10.1103/PhysRevB.58.7260

[63] S. Grimme, J. Antony, T. Schwabe, C. Muck-Lichtenfeld, *Org. Biomol. Chem.* **2007**, *5*(5), 741–758. https://doi.org/10.1039/b615319b

[64] R. C. Walker, M. F. Crowley, D. A. Case, *J. Comput. Chem.* **2008**, *29*(7), 1019–1031. https://doi.org/10.1002/jcc.20857

[65] G. M. Seabra, R. C. Walker, M. Elstner, D. A. Case, A. E. Roitberg, *J. Phys. Chem. A* **2007**, *111*(26), 5655–5664. https://doi.org/10.1021/jp070071l

[66] T. Kubar, P. Jurecka, J. Cerny, J. Rezac, M. Otyepka, H. Valdes, P. Hobza, *J. Phys. Chem. A* **2007**, *111*(26), 5642–5647. https://doi.org/10.1021/jp068858j

[67] P. Hobza, J. Sponer, T. Reschel, *J. Comput. Chem.* **1995**, *16*(11), 1315–1325. https://doi.org/10.1002/jcc.540161102

[68] S. Kristyan, P. Pulay, *Chem. Phys. Lett.* **1994**, *229*(3), 175–180. https://doi.org/10.1016/0009-2614(94)01027-7

[69] P. Jurecka, J. Cerny, P. Hobza, D. R. Salahub, *J. Comput. Chem.* **2007**, *28*(2), 555–569. https://doi.org/10.1002/jcc.20570

[70] P. Jurecka, J. Sponer, J. Cerny, P. Hobza, *Phys. Chem. Chem. Phys. (PCCP)* **2006**, *8*(17), 1985–1993. https://doi.org/10.1039/b600027d

[71] P. M. Mitrasinovic, *J. Biomol. Struct. Dyn.* **2022**, *40*(20), 9713–9723. https://doi.org/10.1080/07391102.2021.1932603
[72] N. Ahmed, B. V. Babu, S. Singh, P. M. Mitrasinovic, *Heterocycles* **2012**, *85*(7), 1629–1653. https://doi.org/10.3987/COM-12-12494
[73] A. Pavlov, P. M. Mitrasinovic, *Curr. Org. Chem.* **2010**, *14*(2), 129–137. https://doi.org/10.2174/138527210790069866
[74] P. M. Mitrasinovic, *Croat. Chem. Acta* **2019**, *92*(1), 43–57. https://doi.org/10.5562/cca3456
[75] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, *Acc. Chem. Res.* **2000**, *33*(12), 889–897. https://doi.org/10.1021/ar000033j
[76] D. A. Case, *Curr. Opin. Struct. Biol.* **1994**, *4*(2), 285–290. https://doi.org/10.1016/S0959-440X(94)90321-2
[77] R. M. Crean, M. Biler, M. W. van der Kamp, A. C. Hengge, S. C. L. Kamerlin, *J. Am. Chem. Soc.* **2021**, *143*(10), 3830–3845. https://doi.org/10.1021/jacs.0c11806
[78] A. Kuban-Jankowska, K. K. Sahu, P. Niedzialkowski, M. Gorska, J. A. Tuszynski, T. Ossowski, M. Wozniak, *Oncotarget* **2015**, *6*(21), 18364–18373. https://doi.org/10.18632/oncotarget.4625
[79] A. Kuban-Jankowska, K. K. Sahu, M. Gorska, P. Niedzialkowski, J. A. Tuszynski, T. Ossowski, M. Wozniak, *World J. Microbiol. Biotechnol.* **2016**, *32*(10), 163. https://doi.org/10.1007/s11274-016-2123-3
[80] S. Ke, M. C. Ho, N. Zhadin, H. Deng, R. Callender, *J. Phys. Chem. B* **2012**, *116*(21), 6166–6176. https://doi.org/10.1021/jp3037846
[81] T. T. Liu, T. Yang, M. N. Gao, K. X. Chen, S. Yang, K. Q. Yu, H. L. Jiang, *Acta Pharmacol. Sin.* **2019**, *40*(6), 850–858. https://doi.org/10.1038/s41401-019-0216-x
[82] P. M. Mitrasinovic, *Acta Chim. Slov*. **2020**, *67*(3), 876–884. https://doi.org/10.17344/acsi.2020.5847
[83] P. M. Mitrasinovic, *Rev. Roum. Chim.* **2025**, *70*(3–4), 133–146. https://doi.org/10.33224/rrch.2025.70.3-4.01
[84] P. M. Mitrasinovic, *ChemRxiv* **2024**, 20246lb2g. https://doi.org/10.26434/chemrxiv-2024-6lb2g
[85] P. M. Mitrasinovic, *ChemRxiv* **2024**, 2024mvpt6. https://doi.org/10.26434/chemrxiv-2024-mvpt