Gibbs Free Energy of Hydrolytic Water Molecule in Acyl-Enzyme Intermediates of a Serine Protease: A Potential Application for Computer-Aided Discovery of Mechanism-Based Reversible Covalent Inhibitors

Yosuke Masuda,* Noriyuki Yamaotsu, and Shuichi Hirono*

School of Pharmaceutical Sciences, Kitasato University; 5–9–1 Shirokane, Minato-ku, Tokyo 108–8641, Japan.
Received May 23, 2017; accepted July 6, 2017

In order to predict the potencies of mechanism-based reversible covalent inhibitors, the relationships between calculated Gibbs free energy of hydrolytic water molecule in acyl-trypsin intermediates and experimentally measured catalytic rate constants ($k_{cat}$) were investigated. After obtaining representative solution structures by molecular dynamics (MD) simulations, hydration thermodynamics analyses using WaterMap™ were conducted. Consequently, we found for the first time that when Gibbs free energy of the hydrolytic water molecule was lower, logarithms of $k_{cat}$ were also lower. The hydrolytic water molecule with favorable Gibbs free energy may hydrolyze acylated serine slowly. Gibbs free energy of hydrolytic water molecule might be a useful descriptor for computer-aided discovery of mechanism-based reversible covalent inhibitors of hydrolytic enzymes.

Key words hydration thermodynamics analysis; Gibbs free energy; hydrolytic water molecule; mechanism-based reversible covalent inhibitor; catalytic rate constant; WaterMap

Covalent inhibitors are a reemerging class of drugs with several advantages including (1) high potency in small doses, (2) prolonged duration and less-frequent dosing, and (3) potency for anti-drug resistance.¹ Mechanism-based reversible covalent inhibitors (one type of covalent inhibitors) present reversible covalent inhibition via catalytic mechanisms of the target enzyme. They are expected to have lower risks of toxicity because drug–protein adducts should not exceed levels to induce a host immune response due to the small dose and the nature of reversible inhibition.² Historically, most covalent inhibitors including aspirin, penicillin, and omeprazole have been discovered serendipitously.³ Therefore, a rational method to predict selectivity and potency in silico is needed. Covalent docking protocols have been developed to predict the selectivity of covalent inhibitors.⁴ However, an effective method to predict potency of covalent inhibitors is yet to be reported. Potency of covalent inhibitors mainly depends on covalent bond formation/cleavage kinetics rather than general, non-covalent equilibrium binding kinetics⁵,⁶, therefore, the method to be developed should be able to estimate the reaction rate.

Serine proteases (EC 3.4.21) are involved in various biological processes and have been drug targets. A quantum mechanics calculation has revealed the energetics of serine protease hydrolysis,⁷ although it demands complex procedures and a long computation time. Meanwhile, hydration thermodynamics analysis, which can predict locations for hydration on protein surfaces and relative energy at the locations, has been a hot topic. In several protocols, WaterMap™ is based on restricted molecular dynamics (MD) simulation and inhomogeneous fluid theory,⁸ and has simple procedures and a reasonably short computation time. In WaterMap™, enthalpy ($\Delta H$) and entropy ($\Delta S$) are calculated from the force field and degree of ordering of water molecules, respectively, and then Gibbs free energy of a water molecule relative to bulk solvent ($AG$) is calculated.⁹ WaterMap™ has been applied to various targets, and demonstrated that displacement of a water molecule with unfavorable $AG$ improves binding affinity.¹⁰ Also, it has been shown that WaterMap™ is possibly useful in estimating binding kinetics of non-covalent ligands.¹¹

Hence, we conducted hydration thermodynamics analyses using WaterMap™ to predict the rates of covalent bond cleavage in serine protease catalysis of mechanism-based reversible covalent inhibitors. In serine proteases, trypsin (EC 3.4.21.4) has provided an important model for understanding catalytic mechanisms of an enzyme. Trypsin hydrolyzes substrates through two reaction steps: acylation and deacylation. In the deacylation step, a hydrolytic water molecule deprotonated by catalytic histidine nucleophilically attacks the carbonyl carbon of the acyl group in an acyl-trypsin intermediate, and a tetrahedral intermediate is formed. Then, cleavage of the C–O bond in the acyl group occurs, and hydrolysis is completed. Most mechanism-based reversible covalent inhibitors of trypsin include esters as the hydrolysis-susceptive group, and in ester hydrolysis, the deacylation reaction is generally the rate-determining step.¹²

Our study focused on $AG$ of hydrolytic water molecule in the acyl-trypsin intermediate, and relationships to catalytic rate constants ($k_{cat}$) were investigated. To our knowledge, our study for the first time extended hydration thermodynamics analysis from binding affinity and binding kinetics toward reaction kinetics of enzyme hydrolysis, and aimed to develop an effective application for the discovery of mechanism-based reversible covalent inhibitors.

First, crystal structures of acyl intermediates of bovine pancreatic trypsin were obtained from Protein Data Bank (PDB)¹⁶–¹⁸ (Table 1). These structures were prepared beforehand for MD simulations. The protonation state of ionizable amino acids at pH 8.0 (the optimum pH for enzymatic activity of trypsin)¹⁵ was determined with PROPKA.¹⁶ His57 $N_ε$ atom was not protonated because it works as a general base to deprotonate the hydrolytic water molecule. After restrained energy minimization using the Impref module of Impact¹⁹ 6.8, MD simulations using Desmond²⁰ 4.3 were performed with TIP3P water models²¹ and the OPLS force field²² to generate an ensemble of energetically accessible structures in water solution. The production simulation was performed for 4 to 5 ns at constant pressure (1 atm) and temperature (300K). The trajectories were extracted during 1 ns—4.0–5.0 (2AH4), 3.0–4.0 (1GBT), 3.5–4.5 (3VPK), 3.0–4.0 (2AGG), and 4.0–5.0 ns (2AGE), respectively—while the systems were equilibrated. Three representative solution structures were then selected by...
the clustering algorithm based on root-mean-squared distances (RMSD) of heavy atoms within 4 Å from the covalently modified serine.

The hydration thermodynamics analyses using WaterMap™ were then performed with TIP4P water models and the OPLS3 force field. All water molecules were deleted beforehand from the structures. The acyl-trypsin intermediates were truncated according to the standard protocol of WaterMap to decrease the computation time. After the 2-ns simulation, several water sites were located around the catalytic site. Among these sites, the hydrolytic water site had the smallest sum of distances $d$—given by $d = d_1 + d_2$—where $d_1$ is the distance between His57 N$_{\epsilon}$ atom and a water site around the catalytic site and $d_2$ is the distance between the carbonyl carbon atom of the acyl group in the catalytic serine modified with a covalent ligand and the water site—because the hydrolytic water site should be near both atoms. Then, the relationship between the calculated $\Delta G$ of hydrolytic water molecule—averaged in three representative structures of a covalent ligand ($\Delta G_{\text{wat}}$) and $k_{\text{cat}}$ was investigated.

All $\Delta G_{\text{wat}}$ exhibited positive values, which indicates that hydrolytic water molecule would obtain unfavorable $\Delta G$ in acyl-trypsin intermediates. When logarithms of $k_{\text{cat}}$ were lower, $\Delta G_{\text{wat}}$ values were also lower (Fig. 1). Moreover, covalent in-

Table 1. Covalent Ligands Used and Related Experimental Data

| Number | Ligand* | Type          | PDB ID | $k_{\text{cat}}$ (s$^{-1}$) |
|--------|---------|---------------|--------|-----------------------------|
| 1      | Nafamostat | Inhibitor  | 2AH4   | $3.2 \times 10^{-5}$ (ref. 10) (ref. 13) |
| 2      |          |              | 1GBT   | (ref. 11)                  |
| 3      | Gabexate | Inhibitor    | 3VPK   | $8.0 \times 10^{-4}$ ** (ref. 12) |
| 4      | suc-AAPK-pNA | Substrate | 2AGG   | 41 (ref. 10) (ref. 10) |
| 5      | suc-AAPR-pNA | Substrate | 2AGE   | 91 (ref. 10) (ref. 10) |

Abbreviations: suc-AAPK-pNA, succinyl-alanyl-alanyl-prolyl-lysyl-para-nitroanilide; suc-AAPR-pNA, succinyl-alanyl-alanyl-prolyl-arginyl-para-nitroanilide. * Inside the dotted rectangles: leaving groups. ** $k_{\text{cat}}$ value of gabexate was estimated from the experimental catalytic constant ($k_{\text{cat}}$) of nafamostat and the ratio of the experimental half-life ($t_{1/2}$) of nafamostat and gabexate in human plasma.
Inhibitors and substrates can be discriminated by the calculatedΔ$G_{\text{wat}}$, which indicates that Δ$G_{\text{wat}}$ would be an appropriate metric to predict whether a covalent inhibitor candidate would present inhibitor-equivalent $k_{\text{cat}}$ or substrate-equivalent $k_{\text{cat}}$.

When each covalent ligand was investigated (Fig. 2), there seems to be two strategies to design mechanism-based reversible covalent inhibitors: (1) to decrease Δ$H_{\text{wat}}$ and (2) to decrease $-T\Delta S_{\text{wat}}$. In the nafamostat-trypsin intermediate, negative Δ$H_{\text{wat}}$ should contribute to the small Δ$G_{\text{wat}}$. Meanwhile, in the gabexate-trypsin intermediate, small $-T\Delta S_{\text{wat}}$ may contribute to the small Δ$G_{\text{wat}}$. In addition, a trend was noticed—when averaged number of hydrogen bonds between acyl-trypsin and hydrolytic water molecule was higher, the enthalpy of hydrolytic water molecule (Δ$H_{\text{wat}}$) was lower. Hydrogen bonds should be a primary factor to stabilize the hydrolytic water molecule and decrease enthalpy. Also, as occupancy of the hydrolytic water site was lower, $-T\Delta S_{\text{wat}}$ was also lower. In the succinyl-alanyl-alanyl-prolyl-lysyl (suc-AAPK) trypsin intermediate, surrounding residues seemed to form a “cul-de-sac” for hydrolytic water and prevent water escaping from the catalytic site—this would increase occupancy and $-T\Delta S_{\text{wat}}$.

Conversely, in the gabexate-trypsin intermediate, there are no “cul-de-sac” residues, which should decrease occupancy and $-T\Delta S_{\text{wat}}$ (Fig. 3). Relationships between Δ$H_{\text{wat}}$, $-T\Delta S_{\text{wat}}$ and structural properties may provide useful clues for the optimization of a covalent ligand candidate to decrease Δ$G_{\text{wat}}$.

Studying the relationship between $k_{\text{cat}}$ and Δ$G_{\text{wat}}$, we hypothesized that unfavorable Δ$G_{\text{wat}}$ in the substrate-trypsin intermediates afforded from surrounding residues might accelerate deacylation in substrate catalysis. Conversely, favorable Δ$G_{\text{wat}}$ in the inhibitor-trypsin intermediates would slow down deacylation in inhibitor catalysis; in other words, stable or escapable hydrolytic water molecules should rarely transform to the hydroxyl group of the tetrahedral intermediate. This mechanism might possibly be universal in serine proteases because catalytic residues around the hydrolytic water molecule are highly preserved.

In this study, $k_{\text{cat}}$ of covalent ligands can be discriminated using Δ$G_{\text{wat}}$ as the primary factor. There might be two rea-
sons: (1) all hydrolysis-susceptive moiety in acyl-trypsin intermediates in this study includes the congeneric acyl group, and (2) energy necessary for the deprotonation of hydrolytic water molecule by catalytic histidine must be almost the same throughout the ligands in this study because the position of the hydrolytic water molecule is almost the same throughout all acyl-trypsin intermediates.

Few ligands were investigated in this study. Therefore, research including more ligands as well as other hydrolytic enzymes is planned in the future. Subsequently, $\Delta G_{\text{wat}}$ is expected to become a good metric for predicting potency of the mechanism-based reversible covalent inhibitors of hydrolytic enzymes.

**Conflict of Interest** The authors declare no conflict of interest.

**References**

1) Singh J., Petter R. C., Baillie T. A., Whitty A., *Nat. Rev. Drug Discov.* **10**, 307–317 (2011).
2) Bauer R. A., *Drug Discov. Today* **20**, 1061–1073 (2015).
3) Zhu K., Borrelli K. W., Greenwood J. R., Day T., Abel R., Farid R., S., Harder E., *J. Chem. Inf. Model.* **54**, 1932–1940 (2014).
4) Topf M., Richards W. G., *J. Am. Chem. Soc.* **126**, 14631–14641 (2004).
5) Lazaridis T., *J. Phys. Chem. B* **102**, 3542–3550 (1998).
6) Young T., Abel R., Kim B., Berne B. J., Friesner R. A., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 808–813 (2007).
7) Beuming T., Farid R., Sherman W., *Protein Sci.* **18**, 1609–1619 (2009).
8) Pearlstein R. A., Sherman W., Abel R., *Prot. Struct. Funct. Bioinf.* **81**, 1509–1526 (2013).
9) Hedstrom L., *Chem. Rev.* **102**, 4501–4524 (2002).
10) Radisky E. S., Lee J. M., Lu C. J. K., Koshland D. E. Jr., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6835–6840 (2006).
11) Mangel W. F., Singer P. T., Cyr D. M., Umland T. C., Toledo D. L., Stroud R. M., Pfugrath J. W., Sweet R. M., *Biochemistry* **29**, 8351–8357 (1990).
12) Masuda Y., Nitanai Y., Mizutani R., Noguchi S., *Prot. Struct. Funct. Bioinf.* **81**, 526–530 (2013).
13) Ramjee M. K., Henderson I. M. J., McLoughlin S. B., Padova A., *Thromb. Res.* **98**, 559–569 (2000).
14) Takeda K., *J. Pancreas* **8**, 526–532 (2007).
15) Northrop J. H., Kunitz M., *J. Gen. Physiol.* **16**, 295–311 (1932).
16) Olsson M. H. M., Søndergaard C. R., Rostkowski M., Jensen J. H., *J. Chem. Theory Comput.* **7**, 525–537 (2011).
17) Banks J. L., Beard H. S., Cao Y., Cho A. E., Damm W., Farid R., Felts A. K., Haigren T. A., Mainz D. L., Maple J. R., Murphy R., Philipp D. M., Repasky M. P., Zhang L. Y., Berne B. J., Friesner R. A., Gallicchio E., Levy R. M., *J. Comput. Chem.* **26**, 1752–1760 (2005).
18) Bowers K. J., Chow E., Xu H., Dror R. O., Eastwood M. P., Gregersen B. A., Klepeis J. L., Kolossvary I., Moraes M. A., Sacerdoti F. D., Salmon J. K., Shan Y., Shaw D. E., "Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)," Tampa, Florida, November 11–17 (2006).
19) Jorgensen W. L., Chandrasekhar J., Madura J. D., Impey R. W., Klein M. L., *J. Chem. Phys.* **79**, 926–935 (1983).
20) Harder E., Damm W., Maple J., Wu C., Reboul M., Yang J. Y., Wang L., Lupyan D., Dahleren M. K., Knight J. L., Kaus J. W., Cerutti D. S., Krilov G., Jorgensen W. L., Abel R., Friesner R. A., *J. Chem. Theory Comput.* **12**, 281–296 (2016).