ACS chemical biology

ARTICLES

pubs.acs.org/acschemicalbiology

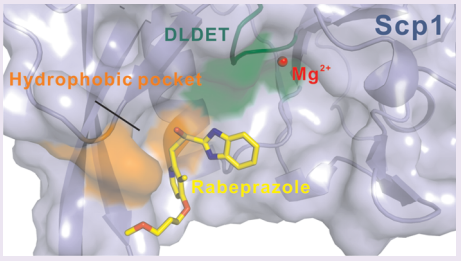
Selective Inactivation of a Human Neuronal Silencing Phosphatase by a Small Molecule Inhibitor

Mengmeng Zhang,[†] Eun Jeong Cho,[‡] Gayle Burstein,[†] Dionicio Siegel,[†] and Yan Zhang[†,*]

[†]Department of Chemistry and Biochemistry and [‡]The Texas Institute for Drug and Diagnostic Development, University of Texas at Austin, Austin, Texas 78712, United States

Supporting Information

ABSTRACT: The unstructured C-terminal domain (CTD) of eukaryotic RNA polymerase II dynamically regulates the process of transcription by recruiting different factors to nascent mRNA through its multiple phosphorylation patterns. A newly discovered class of phosphatases, the human small C-terminal domain phosphatases (Scp's), specifically dephosphorylates phosphorylated Ser_5 (phospho.Ser_5) of the tandem heptad repeats of the CTD of RNA polymerase II. Scp's also function as transcription regulators that epigenetically silence the expression of specific neuronal genes, whose inactivation leads to neuronal stem cell differentiation. Small molecule inhibitors of Scp's will be valuable for elucidating their mechanism in nervous system development and can possibly offer new strategies to treat diseases related to neurodegeneration. Despite the difficulty in developing selective inhibitors of protein phosphatases, we have recognized a characteristic hydrophobic binding pocket adjacent to the active site in Scp's that may facilitate selective inhibition. In the present study, we successfully identified the first selective lead compound, rabeprazole, for the Scp/TFIIF-interacting CTD phosphatase (Fcp) family. The high-resolution crystal structure of rabeprazole-bound Scp1 showed that the compound indeed binds to the hydrophobic binding pocket. We further confirmed that rabeprazole only targets Scp's but not its close family members Fcp1 and Dullard or bacteriophage λ Ser/Thr phosphatase. Such specificity may prove important for *In Vivo* studies since accidental inhibition of Fcp1 or Dullard would result in cell malfunctions and even cell death.



The C-terminal domain (CTD) of eukaryotic RNA polymerase II serves as a binding scaffold for numerous regulatory factors involved in transcription, RNA processing, histone modificationm, and various other cellular events.[1] During transcription, the long, unstructured, and largely repetitive CTD undergoes dynamic phosphorylation and dephosphorylation cycles, creating different phosphorylation patterns (termed the "CTD code") that temporally and spatially orchestrate the recruitment of those regulatory factors. Though inherently disordered, the CTD is evolutionarily conserved from yeast to human.[2] The CTD usually consists of 26−52 tandem heptapeptide repeats with the consensus sequence $Y_1S_2P_3T_4S_5P_6S_7$.[2] Phosphorylation of the CTD (with serines at positions 2 and 5 being the primary phosphorylation sites) is a major mechanism by which cells regulate gene expression. The CTD specific kinases and phosphatases function as house-keeping regulatory factors for global transcription.[3] Recently, it has been shown that certain CTD regulatory factors can also epigenetically modulate the expression level of a specific group of genes.[4−6]

As an example, a newly discovered class of phosphatases, the human small C-terminal domain phosphatases (Scp's), specifically dephosphorylates phosphorylated Ser_5 (phospho.Ser_5) of the tandem heptad repeats of the CTD.[7] Interestingly, Scp's have also been shown to epigenetically silence the expression of a specific set of neuronal genes in neuronal stem cells and non-neuronal cells by acting as co-repressors in REST/NRSF complex.[4] Inhibition of Scp's in P19 stem cells by the dominant negative mutants (which retain the overall structure but have abolished phosphatase activity) or microRNA miR-124 (which directly targets the untranslated region of Scp genes and suppresses their expression) allows neuronal gene expression and induces neuronal differentiation.[4,8] Given the demonstrated role of Scp's in limiting inappropriate expression of neuronal specific genes in pluripotent cells and the fact that their down-regulation leads to neuronal differentiation, Scp's serve as promising new targets for small molecule inhibitors to regulate neuronal stem cell development and to promote neuronal differentiation.

However, one of the greatest challenges associated with phosphatase inhibitor identification is the cross inhibition of other phosphatases due to poor selectivity,[9] which usually stems from small, uncharacteristic active sites of phosphatases. For Scp inhibitors, selectivity is of great concern as two close family members in the Scp/Fcp family, Fcp1[10] and Dullard,[11] play essential roles in cell survival as well as proper development.[11−13]

The crystal structures of Scp/Fcp family members Fcp1 and Scp1 were initially solved by Ghosh *et al.*[14] and Kamenski *et al.*[15] Unlike the traditional cysteine-based or dimetal-dependent phosphatases, the Scp/Fcp family members belong to a unique family of phosphatases that rely on the DxDx(T/V) motif and Mg^{2+} to catalyze the phosphoryl-transfer,[7,11] as found in the

Received: November 8, 2010
Accepted: February 8, 2011
Published: February 09, 2011

haloacid dehydrogenase (HAD) superfamily.[16] In fact, the overall core fold of Scp's resembles the core domains of other HAD family members, despite low sequence similarity.[17]

The HAD superfamily can be further divided into three subfamilies (Supplementary Figure S1) according to the presence and the location of a second domain known as the "cap domain".[18,19] Both type I and II subfamily members utilize the cap domain to shield the active site from bulk solvent and to achieve substrate recognition.[16] The type III HAD superfamily, including the Scp/Fcp family, do not have a cap domain, requiring specificity achieved through alternative strategies, such as recruitment of other regulatory proteins. This feature poses an additional challenge for inhibitor design. In fact, no specific inhibitors have been reported to date for type III HAD family members.

We have previously solved the complex structure of Scp1 bound to its substrate peptide,[20] captured snapshots of the phosphoryl-transfer reaction at each step and the formation of the phospho-aspartyl intermediate using X-ray crystallography, and established the reaction mechanism of Scp's.[17] These crystallographic and biochemical studies not only provided insights into the catalytic mechanism of Scp/Fcp but also hinted a novel strategy of specific recognition of substrates by Scp's.

The complex structure of Scp1 bound to its substrate peptide revealed a hydrophobic binding pocket that is specific for Scp's, suggesting that a specific Scp inhibitor might be obtained through targeting this pocket.[20] In the present study, we exploited this possibility and identified rabeprazole as the first reported lead compound for Scp inhibition (K_i = 5 ± 1 μM). This small molecule shows no inhibition toward Fcp1 or Dullard, nor toward bacteriophage λ Ser/Thr phosphoprotein phosphatase (λPPase). This extraordinary selectivity can be explained through analysis of our high resolution structure of Scp1 complexed with the compound, which shows, as expected, the compound binds specifically to the unique hydrophobic binding pocket of Scp's. The structure highlights the chemical functional groups that make essential contributions to binding. To the best of our knowledge, this is the first selective lead compound for the Scp/Fcp phosphatase family, as well as the type III HAD family.

RESULTS AND DISCUSSION

Human Scp's as Target for Inhibitor Identification. Phosphatases have historically proven to be difficult targets for inhibitor design. The difficulty stems from three major challenges in inhibitor design: (1) affinity: the binding pockets of phosphatases tend to be small and therefore limit potential molecular interactions; (2) specificity: the substrate-specificities of phosphatases tend to overlap, creating possibilities for cross inhibition; (3) permeability: the mimicry of the charged phosphate group, which is key to recognition by phosphatases, often prevents the compounds from penetrating cell membranes. Owing to these difficulties, one of the best phosphatases inhibitors, I-C11, inhibits its target lymphoid-specific tyrosine phosphatase (Lyp) with a K_i of 2.9 ± 0.5 μM and shows some degree of promiscuity.[9] With hundreds of Ser/Thr PPases and protein tyrosine phosphatases identified, the specific inhibitors reported are very limited. The most potent inhibitor for any phosphatase is a small molecule reported recently that targets T cell protein tyrosine phosphatase with a K_i of 4.3 ± 0.2 nM and a minor degree of cross inhibition.[21]

The newly discovered Scp's might prove to be viable targets for inhibitor design. First, the structure of Scp1 bound to its natural substrate phosphorylated CTD peptide reveals a spacious substrate-binding area compared to that of other phosphatases (Figure 1a). Most notably, there is a hydrophobic pocket unique to Scp's that specifically recognizes the Pro_3 of the CTD and is about 7 Å away from the active site where the phosphate group binds (Figure 1a). Compounds that target this hydrophobic pocket are expected to have high binding affinities due to ample opportunities of making molecular contacts. Second, cross inhibition of close family members might be prevented by targeting this hydrophobic pocket since it is unique to Scp's. The hydrophobic pocket is located between the three-stranded β sheet insertion domain and the core domain. Many of the hydrophobic residues lining a part of the hydrophobic pocket are from the insertion domain, which is unique to the Scp/Fcp family and is highly diversified within Scp/Fcp family members (Figure 1b). There is no sequence similarity between Scp, Fcp, Dullard, and other family members in the insertion domain (Figure 1c). Finally, the affinity provided by binding to the hydrophobic pocket might mitigate the necessity of having a charged moiety to mimic phosphate, making the prospect compound more membrane-permeable. By considering these unique advantages of Scp's, we reasoned that Scp's might be feasible targets for selective and high affinity inhibitor design.

Identification of Scp Inhibitors. Scp's include three highly similar homologous family members in the human: Scp1, Scp2, and Scp3.[7] Since no differences are found in the catalytic activities or biological functions of Scp1−3, they are thought to be functionally redundant.[4] Our structural studies of Scp1, 2, and 3 show little overall structural difference and no observable difference at the active site (Y. Zhang, unpublished data). Furthermore, they display identical kinetic characteristics against *p*-nitrophenyl phosphate (*p*NPP) and phosphorylated CTD peptide (Y. Kim and J. E. Dixon, personal communication). Therefore, we chose the best-characterized family member, Scp1, in the following studies.

To identify the inhibitors for Scp's, we screened a pilot library of the NIH clinical collection (~400 compounds) and spectrum collection (2000 compounds) for their ability to impede phosphatase activity of Scp1. The screening assay was performed using *p*NPP as the substrate at a concentration comparable to its K_m (~6 mM). It was confirmed that Scp1 was stable under the assay conditions and can tolerate up to 10% DMSO. A Z′ factor of 0.87 was obtained using the optimized screening protocol. Thirty-nine compounds showed greater than 70% inhibition when screened at a 50 μM inhibitor concentration.

To eliminate the false positives, we compared our initial hits with the collaborative drug discovery database (CDD) and identified eight compounds that are sufficiently soluble and less likely to be false positives. To further confirm the inhibitory effect of the compounds, we used a secondary assay to monitor phosphatase activity, where the CTD-derived phosphorylated peptide ($S_{a5}P_{a6}S_{a7}Y_{b1}S_{b2}P_{b3}T_{b4}S_{b5}P_{b6}S_{b7}Y_{c1}S_{c2}P_{c3}T_{c4}$phospho.$S_{c5}P_{c6}S_{c7}$) was used as the substrate and malachite green reagent was used to capture the released phosphate and in doing so produce a colorimetric signal. Five of the eight compounds showed inhibition for Scp1 when using the natural substrate in the assay. Rabeprazole (Figure 2a), which showed the strongest inhibition, was further characterized to exhibit an IC_{50} of 4 ± 0.7 μM in the *p*NPP assay and 9 ± 3 μM in the malachite green assay, where the concentration of *p*NPP and phosphorylated

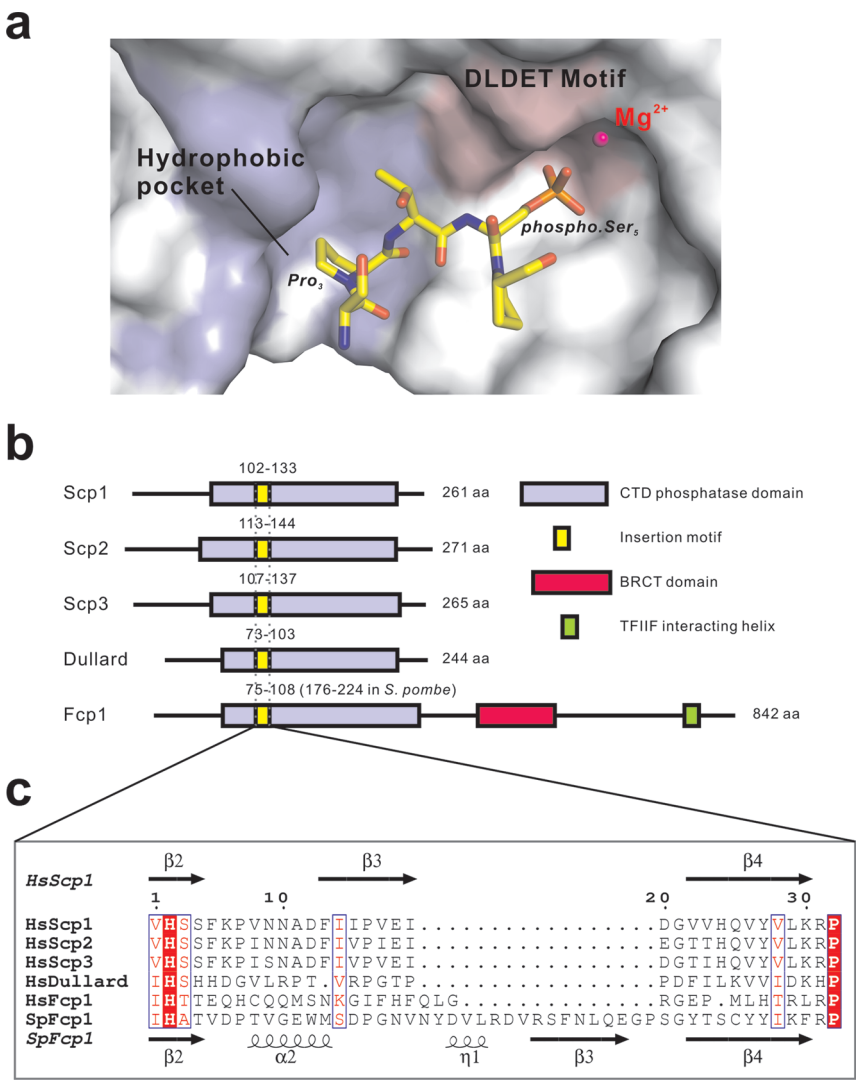


Figure 1. The hydrophobic binding pocket, located at the insertion domain, is unique to Scp's. (a) The complex structure of Scp1 and a CTD-derived peptide (PDB code: 2ght). The active site DxDxT motif is shaded with pink color, and the hydrophobic pocket residues are shaded with blue color. (b) Domain construct of human CTD phosphatases Scp1 (NCBI accession number: AAH12977), Scp2 (AAH65920), Scp3 (NP_005799), Dullard (AAH09295), Fcp1 (AAC64549) and *S. pombe* Fcp1 (NP_594768). Each domain is represented by a colored block. The catalytic domain is colored in light blue. The insertion domain, which is partially conserved, is colored in yellow. (c) The zoom-in figure of the insertion domain alignment. The upper panel shows the secondary structure of human Scp1 insertion domain, and the lower panel shows the secondary structure of *S. pombe* Fcp1 insertion domain. α indicates α helix, β indicates β sheet, and η indicates 3_{10} helix.

peptide were 6 mM and 30 μM, respectively, close to their respective K_m values (Figure 2b). To eliminate the possibility that the inhibitory effect of the compound was caused by protein denaturation, the thermal stability of Scp1 in the absence or presence of rabeprazole was tested in differential scanning fluorimetry assay.[22] Under the measurement condition, the melting temperature of Scp1 is 57.8 and 59.1 °C in the absence and presence of the compound, respectively. This result excludes protein destabilization as the mechanism for the observed inhibitory effect of rabeprazole on Scp1 (Supplementary Figure S2).

To probe the mechanism of inhibition, the steady-state kinetics using both the *p*NPP and malachite green assays were determined. Kinetic analysis revealed that the compound is a mixed inhibitor for Scp1 against *p*NPP. The inhibition constants K_i and K_i' of the compound were determined to be $5 \pm 2\,\mu$M and $10 \pm 2\,\mu$M using *p*NPP as the substrate. Interestingly, when the CTD-derived phosphorylated peptide was used as the substrate, rabeprazole exhibited characteristics of a competitive inhibitor: while the K_i was determined to be $5 \pm 1\,\mu$M, the K_i' was at least 10 times higher than the K_i and could not be precisely determined (Figure 2c). Rabeprazole very likely directly competes with the natural substrate CTD peptide in binding to Scp1.

Complex Structure of Scp1 and Rabeprazole. To understand the inhibition mechanism of rabeprazole at the molecular level, we attempted to obtain the structure of Scp1 and rabeprazole by X-ray crystallography. Unfortunately, crystals derived from the published conditions proved to be too fragile upon compound soaking. We therefore identified new crystallization conditions for wildtype Scp1 that yielded Scp1 crystals in much higher quality (see Methods). Unlike the original conditions, with high concentration of ammonium sulfate as precipitant, the crystals grown from the new poly(ethylene glycol) (PEG)-based conditions are much more durable upon compound soaking. The

dx.doi.org/10.1021/cb100357t |*ACS Chem. Biol.* 2011, 6, 511–519

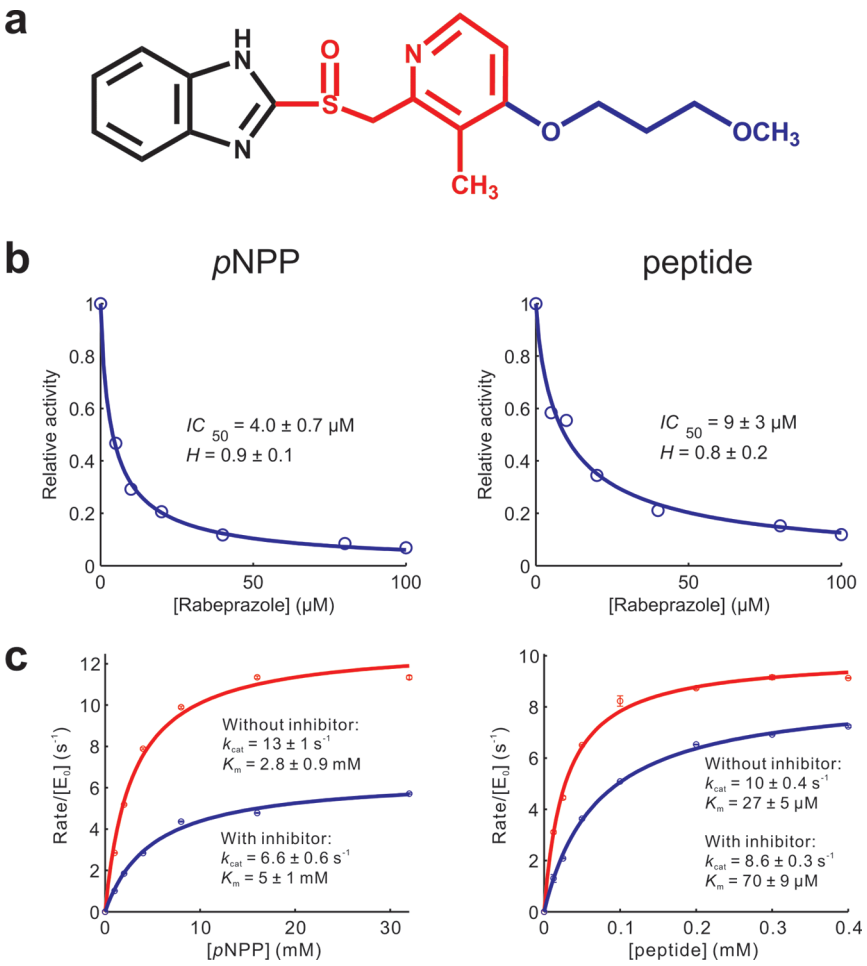


Figure 2. Inhibition of Scp1 by rabeprazole. (a) Structure of rabeprazole. The benzimidazole ring is in black. The methyl pyridine ring and methyl sulfinyl groups are in red. The methoxypropoxy tail is in blue. (b) Concentration−response of rabeprazole inhibition toward Scp1 tested by the *p*NPP assay and the malachite green assay. IC_{50} values were derived by fitting the data to the eq 1 in Supporting Information. (c) Steady-state kinetics of Scp1 in the presence (blue) and absence (red) of rabeprazole. Inhibition constants were derived by fitting the data to eq 2 in Supporting Information.

complex structures were obtained by soaking the Scp1 crystals in buffer containing 0.1−1 mM rabeprazole and incubating for various amounts of time at 25 °C. After optimization of soaking conditions, the crystals soaked in ~0.5 mM rabeprazole for 2−3 h were used in X-ray data collection with good diffraction quality (Table 1).

In all of the data collected from multiple crystals, a strong area of electron density was observed at the hydrophobic pocket for recognition of the Pro_3 of CTD substrate (Figure 3a). The final structure of the Scp1−rabeprazole complex was refined to 2.35 Å with the protein portion of the structure highly identical to apo Scp1 (Table 1) (Supplementary Figure S3). Between the two molecules in each asymmetric unit, molecule A shows strong and consistent positive density at the active site, indicating effective compound binding. In comparison to molecule A, the density at the active site of molecule B is much weaker. Our subsequent discussion of protein and ligand interaction focuses on molecule A in which a better model can be generated on the basis of the diffraction data.

The total buried contact area of Scp1 upon rabeprazole binding is about 279 $Å^2$. The shape complementarily (Sc) value is calculated to be 0.54, which indicates a good shape complementarity between the protein and ligand surfaces (an Sc value of 1 reflects a perfect fit) (see Supporting Information). The most distinguishing characteristic of the compound binding is the site of interaction of the methyl pyridine ring at the hydrophobic pocket surrounded by Tyr158, Phe106, Val118, Ile120, Val127, and Leu155 (Figure 3b and c). The distances between the methyl pyridine ring and surrounding key residues are within 3.5−4.5 Å. The binding to this hydrophobic pocket is noteworthy as this "insertion domain" is unique for Scp's and is believed to be the major structural element for the specific recognition of Pro_3 of the CTD.[20] Since the compound binds to the hydrophobic pocket of Scp1, it is likely to prevent Pro_3 of the substrate from binding to the protein, explaining why the compound resembled a competitive inhibitor to Scp1 when the peptide was used as the substrate (Figure 3c). To further validate the binding mode of rabeprazole binding to Scp1, mutagenesis study around the hydrophobic binding pocket was performed. Four mutants, including F106E, V127A, K157A, and L155A, were tested by the *p*NPP assay. The IC_{50} of rabeprazole when tested against each mutant was derived from the concentration−response curve (Supplementary Figure S4). It is clearly shown that the F106E mutant increased IC_{50} by at least 10-fold. The other mutants, V127A, K157A, and L155A, all have moderately increased IC_{50}, indicating some degree of loss of inhibition. The mutagenesis results show that the hydrophobic residues in the insertion domain are very important for inhibition of Scp1 by rabeprazole.

dx.doi.org/10.1021/cb100357t |*ACS Chem. Biol.* 2011, 6, 511–519

Table 1. Crystallographic Data for Scp1−Rabeprazole Complex

Data Collection	
space group	*C*2
cell dimensions:	
a, *b*, *c* (Å)	125.5, 78.3, 62.7
α, *β*, *γ* (deg)	90.0, 111.9, 90.0
resolution (Å)	50−2.35 (2.39−2.35)[a]
no. of unique reflections	21285
R_{sym} or R_{merge} (%)	10.0 (39.7)
$I/\sigma(I)$	11.4 (1.6)
completeness (%)	90.2 (47.3)
redundancy	3.1 (2.2)
Refinement	
resolution (Å)	64.95−2.35
no. of reflections (test set)	20188 (1097)
R_{work}/R_{free} (%)[b]	20.7/26.5
no. of atoms	
protein	2928
ion	2
ligand	25
water	112
B-factors (Å^2)	
protein	32.0
Mg^{2+}	28.0
ligand	52.5
water	34.6
rms deviations	
bond lengths (Å)	0.020
bond angles (deg)	1.977
Ramachandran plot (%)	
most favored	85.7
additionally allowed	13.7
generally allowed	0.3

[a] Highest resolution shell is shown in parentheses. [b] R_{free} is calculated with 5% of the data randomly omitted from refinement.

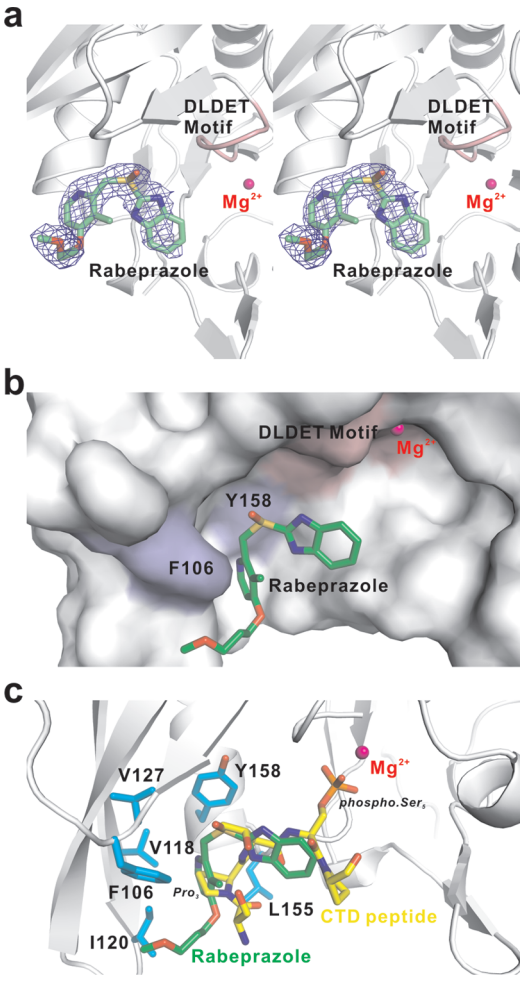


Figure 3. Complex structure of Scp1 and rabeprazole. (a) Structure of rabeprazole bound to Scp1 with the blue SIGMAA-weighted $2F_o - F_c$ electron-density map contoured at 1σ shown in stereo. (b) Surface representation of rabeprazole bound to the hydrophobic pocket of Scp1. (c) Superimposition of Scp1-rabperazole complex and Scp1-CTD peptide complex (PDB code: 2ght). The protein portion is identical. The hydrophobic pocket residues are shown in stick in cyan.

The interaction of the sulfoxide group of rabeprazole and Tyr158 side chain resembles a cation-π interaction where the electron-rich π cloud of the phenyl ring of Tyr158 interacts with the partially positively charged sulfur of rabeprazole. Cation-π interactions are not uncommon in protein-inhibitor interactions and have been recognized as an important noncovalent interaction in structural biology.[23−26] For instance, the horse liver alcohol dehydrogenase is bound by its inhibitor sulfoxides through a cation-π interaction between the inhibitor and the benzene ring of Phe93 of the enzyme.[27] In our structure, the sulfur atom locates directly over the phenyl ring of Tyr158. The distance between the sulfur atom and the center of phenyl ring is about 4 Å, which is within the range of typical distance for "amino-aromatic" interactions.[28] Indeed, when we eliminated this cation-π interaction by mutating Tyr158 to Ala, a loss of inhibition by at least 10-fold was observed when tested by the *p*NPP assay (Supplementary Figure S4).

It is likely that not all functional groups of rabeprazole contribute to Scp1 binding. Unlike pyridine ring and sulfoxide, the benzimidazole ring showed little if any electron density which suggests that this portion is flexible in our structure. Accordingly, the refinement of this part of the structure presents a higher thermal factor than the active site residues, and thus we rebuilt our model with the benzimidazole ring as 50% occupancy. The refinement puts the group in a position that extends it outward of the active site, contributing little to the recognition of the inhibitor. In addition, the methoxypropoxy "tail" portion also showed partial densities consistently in multiple crystals. Although there are two oxygen atoms on the "tail", no identifiable polar interactions were observed in our structure. It may also form hydrophobic interactions with the pocket, and the distances between the "tail" and surrounding hydrophobic residues are within 4−4.4 Å.

It should be noted that the rabeprazole used in our structural studies is a mixture of two enantiomers with the sulfur as the chiral center. However, only (*R*)-rabeprazole can be built into the electron density; the *S* enantiomer does not fit the density well (Supplementary Figure S5). We predict that only the *R* enantiomer binds the protein and exerts inhibitory effect.

Selectivity of Rabeprazole. To evaluate the specificity of rabeprazole as a Scp inhibitor, we tested its inhibition against

dx.doi.org/10.1021/cb100357t |*ACS Chem. Biol.* 2011, 6, 511–519

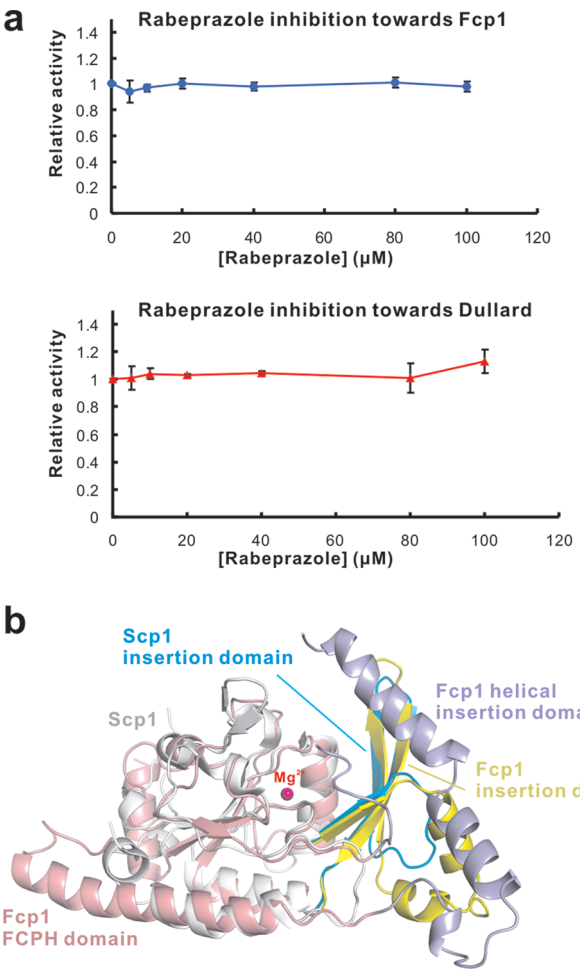


Figure 4. Inhibition of Fcp1 and Dullard by rabeprazole. (a) Concentration−response of rabeprazole inhibition toward *S. pombe* Fcp1 and human Dullard tested by the *p*NPP assay. (b) Superimposition of Scp1 (white) and *S. pombe* Fcp1 FCPH domain (pink, PDB code: 3ef0). The catalytic core folds of the two proteins are similar. The three-stranded β sheet insertion domain of Scp1 (cyan) is accessible for substrate recognition, whereas the insertion domain of Fcp1 (yellow) exhibits different secondary structure and is buried by an additional helical insertion domain (light blue).

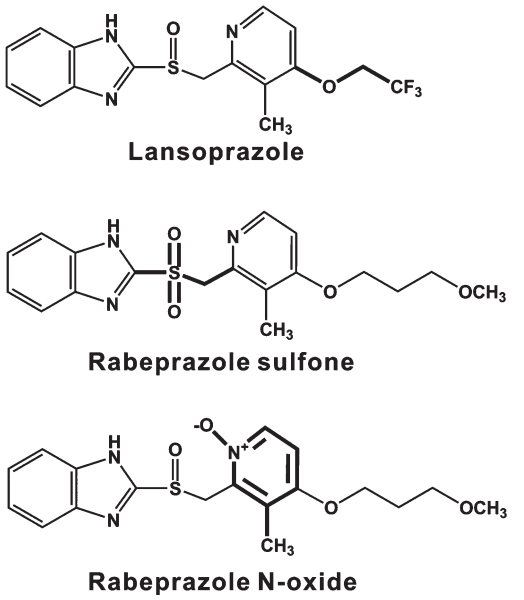


Figure 5. Structures of rabeprazole analogues. The differences between the analogues and rabeprazole are indicated by thick lines.

other Scp/Fcp family phosphatases including Fcp1 and Dullard, whose inactivation usually lead to cell death or improper development.[12,13,29,30] Such tests are not only important to evaluate the potential value of rabeprazole in biological applications but also crucial for confirming the mechanism of the inhibition. In the Scp/Fcp phosphatase family, all enzymes share the same signature motif DxDx(T/V) for phosphoryl transfer. They all belong to the HAD superfamily (also called DxDx(T/V) superfamily) type III subfamily which predominantly act on protein substrates.[16]

The possible off-target inhibitory effect of rabeprazole was tested on the closest homologues of Scp's: *Schizosaccharomyces pombe* (*S. pombe*) Fcp1[14,31] and human Dullard. His-tagged *S. pombe* Fcp1 and His-GST-tagged Dullard were purified to homogeneity and exhibit stable phosphatase activity (Supplementary Figure S6). The steady-state kinetic studies of these enzymes were performed. The K_m for Fcp1 and Dullard were determined to be 15 ± 3 mM and 10 ± 2 mM respectively, comparable with previously reported values.[11,31] The concentration−response of inhibition was tested for both proteins using substrate concentrations comparable to the K_m. No inhibition was observed for either enzyme when up to 0.1 mM rabeprazole was used (Figure 4a).

The observed specificity of rabeprazole can be explained structurally. The major interactions of rabeprazole with Scp are attributed to the insertion domain, which is unique to the Scp/Fcp family and poorly conserved among the family member (Figure 1c). For example, the three-stranded β sheet insertion domain in Scp1 is accessible for substrate recognition, whereas in Fcp1 it exhibits a different secondary structure and is buried by a helical insertion domain, suggesting a different binding interface between Fcp1 and the CTD with phospho.Ser_2 (Figure 4b). Even though the Dullard structure has yet to be determined, the secondary structure is predicted to be different from the three-stranded β sheet as observed in Scp1. It is highly likely that rabeprazole does not inhibit human Dullard due to a different insertion domain in Dullard.

Rabeprazole also exhibits no inhibition of λPPase, a Ser/Thr PPases outside the Scp/Fcp family, at up to 1 mM concentration (data not shown).

We next determined if rabeprazole analogues also exhibit inhibitory effect on Scp's (Figure 5). Rabeprazole sulfone, a metabolite of rabeprazole,[32] did not show any inhibition in our concentration−response assay (data not shown). The absence of inhibition might stem from the loss of cation-π interaction between the sulfoxide group and Tyr158 with the extra oxygen in rabeprazole sulfone. The result further demonstrates that the sulfoxide group is important for the Scp1-rabeprazole interaction and should be kept intact.

Rabeprazole *N*-oxide (Figure 5), a contaminate during the rabeprazole synthesis process,[33] also exhibited no inhibition toward Scp1. It is likely that the change of the methyl pyridine ring interrupts or weakens the van der Waals interactions originally existing between the hydrophobic pocket residues and rabeprazole.

Likewise, lansoprazole (marketed as Prevacid in U.S., Figure 5), which differs from rabeprazole only at the tail portion, showed no

inhibition, probably because the electrons of pyridine ring are dramatically influenced by converting the ether to the trifluoromethyl group. It is equally likely that the shorter tail and significant change in hydrophobicity disrupt important contacts. Since the pyridine ring is important in making hydrophobic interactions with the protein, the changed electron distribution or changed hydrophobicity will likely weaken the hydrophobic interactions.

Taken together, these results not only demonstrate the excellent specificity of rabeprazole but also validate our strategy of targeting the hydrophobic binding pocket adjacent to the active site. To the best of our knowledge, rabeprazole is the first reported selective inhibitor for the Scp/Fcp family proteins and also for the type III HAD subfamily proteins.

Considerations on the Clinical Application of Rabeprazole. Rabeprazole is the active ingredient of an FDA-approved antiulcer drug AcipHex, which is used to treat gastresophageal reflux disease (GERD). In an acidic environment such as the lumen of gastric parietal cells, rabeprazole inhibits its target H^+/K^+ ATPase by forming a covalent bond with the active site cysteine under acidic condition (pH around 1). Rabeprazole behaves as a pro-drug for GERD and only becomes activated in highly acidic condition through two protonations and a subsequent spontaneous rearrangement to form the active sulfenamide (Supplementary Figure S7).[34] However, we reason that rabeprazole does not inhibit Scp1 through this mechanism, since no cysteines participate in the catalysis of Scp1 and no covalent bond between rabeprazole and Scp1 was observed in the complex structure (crystals obtained at neutral pH). This reasoning is corroborated by the fact that lansoprazole, another a proton pump inhibitor that shares the same mechanism with rabeprazole in treating GERD, does not inhibit Scp1.

Further Development of Inhibitors with Higher Binding Affinity. Based on the complex structure of Scp1 and rabeprazole, further development of Scp's specific inhibitors with higher affinity should be attainable. The important groups of this lead compound contributing to the binding are the methyl pyridine ring and methyl sulfinyl moieties. The benzimidazole ring seems to play little positive role if not hindering the compound binding, and therefore, its replacement by other groups should be desirable. In our structure, no interaction of rabeprazole with the DxDx motif was observed. Further optimization can be explored by keeping the methyl pyridine ring and methyl sulfinyl groups of rabeprazole intact but changing other portions to explore interactions with the active site pocket.

In conclusion, we successfully identified the first selective lead compound for the inhibition of Scp's in our high-throughput screening efforts. Our kinetic and structural analyses further confirmed its inhibitory effect toward Scp1 but not its close family members, Fcp1 and Dullard. Thus, rabeprazole represents the first specific inhibitor for Scp/Fcp family phosphatases and also the first inhibitor for type III HAD family members. The complex structure of Scp1 bound with rabeprazole clearly shows the binding mode of the small molecule. As we expected, rabeprazole binds to the unique hydrophobic pocket of Scp's that, in natural context, binds to Pro_3 of the CTD peptide. Since the hydrophobic pocket is located at the insertion domain that is unique to Scp/Fcp family members and shows diverse sequence among those family members, cross inhibition of other family members is prevented. The pyridine ring and sulfoxide groups are essential for this compound binding to Scp1. Other groups that may not substantially contribute to the binding can be optimized to make more potent inhibitors. The present study provides a starting point for the development of new inhibitors of Scp's that can induce neuronal stem cell differentiation, which is of great interest in both basic and clinical research.

METHODS

Materials. *p*NPP was purchased from Fluka, Sigma-Aldrich. CTD peptides were purchased from Anaspec. SYPRO orange dye was purchased from Invitrogen. The *S. pombe* Fcp1 cDNA was purchased from National BioResource Project (NBRP)-Yeast Genetic Resource Center (Osaka City University, Osaka, Japan). Malachite green reagent was purchased from BIOMOL, Enzo Life Sciences. Rabeprazole and rabeprazole analogues were purchased from Toronto Research Chemicals, Inc., and the identity and purity of each compound was confirmed by our in-house mass spec facility.

Cloning, Protein Expression, and Purification. Wildtype or mutant human Scp1 (residues 77−256), wildtype *S. pombe* Fcp1 (residues 148−641), and wildtype human Dullard (residues 46−244) were expressed in *E.coli* BL21(DE3) strain and purified using Ni^{2+}-NTA affinity beads and, when necessary, gel filtration. The detailed cloning, expression, and purification procedures are provided in Supporting Information.

High-Throughput Screening. The high-throughput screening was performed in 384-well plates using *p*NPP as the substrate. We screened a library of NIH clinical collection (~400 compounds) and spectrum collection (2000 compounds) in our local high-throughput facility TI3D. The proteins were preincubated in the presence of 50 μM compounds in 0.8% DMSO at RT for 10 min. All the screening reactions (10 μL) were carried out at 37 °C in Master Buffer (50 mM Tris-acetate pH 5.5, 10 mM $MgCl_2$, 0.02% Triton X-100, 0.5% DMSO) with 25 ng of Scp1, and *p*NPP concentration at its K_m (6 mM). The reactions were quenched by addition of 40 μL of 0.25 N NaOH after 10 min of reaction. Release of *p*NP was determined by measuring absorbance at 410 nm.

***p*NPP assay.** The activity of each phosphatase toward *p*NPP in the absence or presence of inhibitor was measured in Assay Buffer (50 mM Tris-acetate pH 5.5, 10 mM $MgCl_2$, 0.02% Triton X-100, 1% DMSO) with an appropriate amount of protein (50 ng of Scp1 or Scp1 mutants, 5 μg of Dullard or Fcp1) at 37 °C in 20 μL volume. When inhibitor was included, protein and inhibitor were preincubated in Assay Buffer at RT for 10 min in 15 μL volume. After 10−15 min of reaction, the reactions were quenched by adding 80 μL of 0.25 N NaOH. Released *p*NP was quantified by measuring absorbance at 410 nm.

Malachite Green Assay. The activity of Scp1 toward 17-mer Ser_5-phosphorylated peptide ($S_{a5}P_{a6}S_{a7}Y_{b1}S_{b2}P_{b3}T_{b4}S_{b5}P_{b6}S_{b7}Y_{c1}S_{c2}P_{c3}T_{c4}$-phospho.$S_{c5}P_{c6}S_{c7}$) in the absence or presence of inhibitor was measured in Assay Buffer with 5 ng of Scp1 at 37 °C in 20 μL volume. When inhibitor was included, protein and inhibitor were preincubated in Assay Buffer at RT for 10 min in 15 μL volume. After 3 min of reaction, the reactions were quenched by adding 40 μL of malachite green reagent. The release of free inorganic phosphate was determined by measuring the absorbance at 620 nm.

Determination of Steady-State and Inhibition Constants. The steady-state constants (k_{cat} and K_m) were determined by varying substrate concentration (*p*NPP 0.5−50 mM; peptide 0.01−0.4 mM) in the absence of inhibitor. IC_{50} was determined by concentration−response assay, using a fixed concentration of substrate (around K_m) and various concentrations of the inhibitor (0−100 μM). Inhibition constants (K_i and K_i') were measured by varying substrate concentration in the absence of inhibitor and in the presence of inhibitor at IC_{50} concentration. The procedure of data processing is provided in Supporting Information.

Differential Scanning Fluorimetry. The method was based on a modified published protocol.[22] Wildtype Scp1 was mixed with

rabeprazole in the reaction containing 20 mM HEPES pH7.5, 50 mM NaCl, 1% DMSO, and 5× SYPRO orange dye. The final concentrations of protein and rabeprazole were 20 μM and 100 μM, respectively. The unfolding process of protein was monitored during the temperature increase from RT to 85 °C by LightCycler 480 machine (Roche). The protein only, protein with 1% DMSO, and protein with 1 mM peptide were used as controls. The denature process of the protein was fitted into a monophase graph, and the T_m was derived using the LightCycler 480 Protein Melting Software.

Crystallization and Structure Determination. Wildtype Scp1 was crystallized in 30% PEG 3350 and 0.2 M magnesium acetate. Crystals were transferred to a stabilizer consisting of the same buffer condition and 0.1−1 mM inhibitor. After soaking for 2−12 h, crystals were transferred to a cryo-protecting stabilizer containing 30% (v/v) glycerol, 25% (w/v) PEG 3350, and 0.2 M magnesium acetate. After a brief period of equilibration, crystals were frozen in nylon loops in liquid nitrogen and stored in liquid nitrogen prior to data collection. The detailed procedure for data collection and processing is described in Supporting Information.

■ ASSOCIATED CONTENT

Supporting Information. This material is available free of charge *via* the Internet at http://pubs.acs.org.

Accession Codes

Coordinates of the Scp1-rabeprazole complex structure have been deposited in the Protein Data Bank with the accession number 3PGL.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 512-471-8645. Fax: 512-471-9469. E-mail: jzhang@cm.utexas.edu.

■ ACKNOWLEDGMENT

We thank the staff of the Advanced Light Source (ALS), Berkeley, California for their assistance during data collection at beamlines 5.0.1 and 5.0.2. We are grateful to Drs. X. Chen, B. Iverson, G. Gill and A. Ellington for their helpful discussions. We also want to thank Yeast Genetic Resource Center (Osaka City University, Japan) for providing the cDNA of *S. pombe* Fcp1. This work was supported by University of Texas at Austin and NIH R03DA030556 (PI: Y. Zhang). The authors declare no competing financial interests.

■ REFERENCES

(1) Buratowski, S. (2009) Progression through the RNA polymerase II CTD cycle. *Mol. Cell 36*, 541–546.

(2) Corden, J. L. (1990) Tails of RNA polymerase II. *Trends Biochem. Sci. 15*, 383–387.

(3) Majello, B., and Napolitano, G. (2001) Control of RNA polymerase II activity by dedicated CTD kinases and phosphatases. *Front. Biosci. 6*, D1358–1368.

(4) Yeo, M., Lee, S. K., Lee, B., Ruiz, E. C., Pfaff, S. L., and Gill, G. N. (2005) Small CTD phosphatases function in silencing neuronal gene expression. *Science 307*, 596–600.

(5) Li, M., Phatnani, H. P., Guan, Z., Sage, H., Greenleaf, A. L., and Zhou, P. (2005) Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc. Natl. Acad. Sci. U.S.A. 102*, 17636–17641.

(6) Zhang, M., Gill, G. N., Zhang, Y. (2010) Bio-molecular architects: a scaffold provided by the C-terminal domain of eukaryotic RNA polymerase II *Nano Rev. 1*, 5502; DOI: 10.3402/nano.v1i0.5502.

(7) Yeo, M., Lin, P. S., Dahmus, M. E., and Gill, G. N. (2003) A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J. Biol. Chem. 278*, 26078–26085.

(8) Visvanathan, J., Lee, S., Lee, B., Lee, J. W., and Lee, S. K. (2007) The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes Dev. 21*, 744–749.

(9) Yu, X., Sun, J. P., He, Y., Guo, X., Liu, S., Zhou, B., Hudmon, A., and Zhang, Z. Y. (2007) Structure, inhibitor, and regulatory mechanism of Lyp, a lymphoid-specific tyrosine phosphatase implicated in autoimmune diseases. *Proc. Natl. Acad. Sci. U.S.A. 104*, 19767–19772.

(10) Cho, H., Kim, T. K., Mancebo, H., Lane, W. S., Flores, O., and Reinberg, D. (1999) A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev. 13*, 1540–1552.

(11) Kim, Y., Gentry, M. S., Harris, T. E., Wiley, S. E., Lawrence, J. C., Jr., and Dixon, J. E. (2007) A conserved phosphatase cascade that regulates nuclear membrane biogenesis. *Proc. Natl. Acad. Sci. U.S.A. 104*, 6596–6601.

(12) Archambault, J., Chambers, R. S., Kobor, M. S., Ho, Y., Cartier, M., Bolotin, D., Andrews, B., Kane, C. M., and Greenblatt, J. (1997) An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A. 94*, 14300–14305.

(13) Chambers, R. S., and Dahmus, M. E. (1994) Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. *J. Biol. Chem. 269*, 26243–26248.

(14) Ghosh, A., Shuman, S., and Lima, C. D. (2008) The structure of Fcp1, an essential RNA polymerase II CTD phosphatase. *Mol. Cell 32*, 478–490.

(15) Kamenski, T., Heilmeier, S., Meinhart, A., and Cramer, P. (2004) Structure and mechanism of RNA polymerase II CTD phosphatases. *Mol. Cell 15*, 399–407.

(16) Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold. *Trends Biochem. Sci. 29*, 495–503.

(17) Zhang, M., Liu, J., Kim, Y., Dixon, J. E., Pfaff, S. L., Gill, G. N., Noel, J. P., and Zhang, Y. (2010) Structural and functional analysis of the phosphoryl transfer reaction mediated by the human small C-terminal domain phosphatase, Scp1. *Protein Sci. 19*, 974–986.

(18) Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2000) The crystal structure of bacillus cereus phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. *Biochemistry 39*, 10385–10396.

(19) Zhang, G., Mazurkie, A. S., Dunaway-Mariano, D., and Allen, K. N. (2002) Kinetic evidence for a substrate-induced fit in phosphonoacetaldehyde hydrolase catalysis. *Biochemistry 41*, 13370–13377.

(20) Zhang, Y., Kim, Y., Genoud, N., Gao, J., Kelly, J. W., Pfaff, S. L., Gill, G. N., Dixon, J. E., and Noel, J. P. (2006) Determinants for dephosphorylation of the RNA polymerase II C-terminal domain by Scp1. *Mol. Cell 24*, 759–770.

(21) Zhang, S., Chen, L., Luo, Y., Gunawan, A., Lawrence, D. S., and Zhang, Z. Y. (2009) Acquisition of a potent and selective TC-PTP inhibitor via a stepwise fluorophore-tagged combinatorial synthesis and screening strategy. *J. Am. Chem. Soc. 131*, 13072–13079.

(22) Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc 2*, 2212–2221.

(23) Dougherty, D. A. (1996) Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science 271*, 163–168.

(24) Ma, J. C., and Dougherty, D. A. (1997) The Cation-π Interaction. *Chem. Rev. 97*, 1303–1324.

(25) Scrutton, N. S., and Raine, A. R. (1996) Cation-pi bonding and amino-aromatic interactions in the biomolecular recognition of substituted ammonium ligands. *Biochem. J. 319*, 1–8.

(26) Gallivan, J. P., and Dougherty, D. A. (1999) Cation-pi interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A. 96*, 9459–9464.
(27) Cho, H., Ramaswamy, S., and Plapp, B. V. (1997) Flexibility of liver alcohol dehydrogenase in stereoselective binding of 3-butylthiolane 1-oxides. *Biochemistry 36*, 382–389.
(28) Burley, S. K., and Petsko, G. A. (1986) Amino-aromatic interactions in proteins. *FEBS Lett. 203*, 139–143.
(29) Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999) An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in S. cerevisiae. *Mol. Cell 4*, 55–62.
(30) Satow, R., Chan, T. C., and Asashima, M. (2002) Molecular cloning and characterization of dullard: a novel gene required for neural development. *Biochem. Biophys. Res. Commun. 295*, 85–91.
(31) Hausmann, S., and Shuman, S. (2003) Defining the active site of Schizosaccharomyces pombe C-terminal domain phosphatase Fcp1. *J. Biol. Chem. 278*, 13627–13632.
(32) Klotz, U. (2000) Pharmacokinetic considerations in the eradication of Helicobacter pylori. *Clin. Pharmacokinet. 38*, 243–270.
(33) Reddy, G. M., Mukkanti, K., Bhaskar, B. V., and Reddy, P. P. (2009) Synthesis of Metabolites and Related Substances of Rabeprazole, an Anti-Ulcerative Drug. *Synth. Commun. 39*, 278–290.
(34) Roche, V. F. (2006) The chemically elegant proton pump inhibitors. *Am. J. Pharm. Educ. 70*, 101.