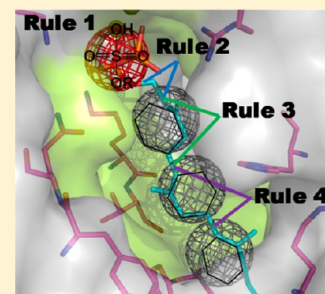


# Moiety-Linkage Map Reveals Selective Nonbisphosphonate Inhibitors of Human Geranylgeranyl Diphosphate Synthase

Shih-Hsun Chen,<sup>†</sup> Sheng-Wei Lin,<sup>‡</sup> Shen-Rong Lin,<sup>§</sup> Po-Huang Liang,<sup>\*,‡</sup> and Jinn-Moon Yang<sup>\*,†,§</sup><sup>†</sup>Department of Biological Science and Technology, National Chiao Tung University, Hsinchu 30050, Taiwan<sup>‡</sup>Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan<sup>§</sup>Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsinchu 30050, Taiwan

## S Supporting Information

**ABSTRACT:** Bisphosphonates are potent inhibitors of farnesyl pyrophosphate synthase (FPPS) and geranylgeranyl diphosphate synthase (GGPPS). Current bisphosphonate drugs (e.g., Fosamax and Zometa) are highly efficacious in the treatment of bone diseases such as osteoporosis, Paget's disease, and tumor-induced osteolysis, but they are often less potent in blood and soft-tissue due to their phosphate moieties. The discovery of nonbisphosphonate inhibitors of FPPS and/or GGPPS for the treatment of bone diseases and cancers is, therefore, a current goal. Here, we propose a moiety-linkage-based method, combining a site-moiety map with chemical structure rules (CSRs), to discover nonbisphosphonate inhibitors from thousands of commercially available compounds and known crystal structures. Our moiety-linkage map reveals the binding mechanisms and inhibitory efficacies of 51 human GGPPS (hGGPPS) inhibitors. To the best of our knowledge, we are the first team to discover two novel selective nonbisphosphonate inhibitors, which bind to the inhibitory site of hGGPPS, using CSRs and site-moiety maps. These two compounds can be considered as a novel lead for the potent inhibitors of hGGPPS for the treatment of cancers and mevalonate-pathway diseases. Moreover, based on our moiety-linkage map, we identified two key residues of hGGPPS, K202, and K212, which play an important role for the inhibitory effect of zoledronate ( $IC_{50} = 3.4 \mu M$  and  $2.4 \mu M$ , respectively). This result suggests that our method can discover specific hGGPPS inhibitors across multiple prenyltransferases. These results show that the compounds that highly fit our moiety-linkage map often inhibit hGGPPS activity and induce tumor cell apoptosis. We believe that our method is useful for discovering potential inhibitors and binding mechanisms for pharmaceutical targets.



## INTRODUCTION

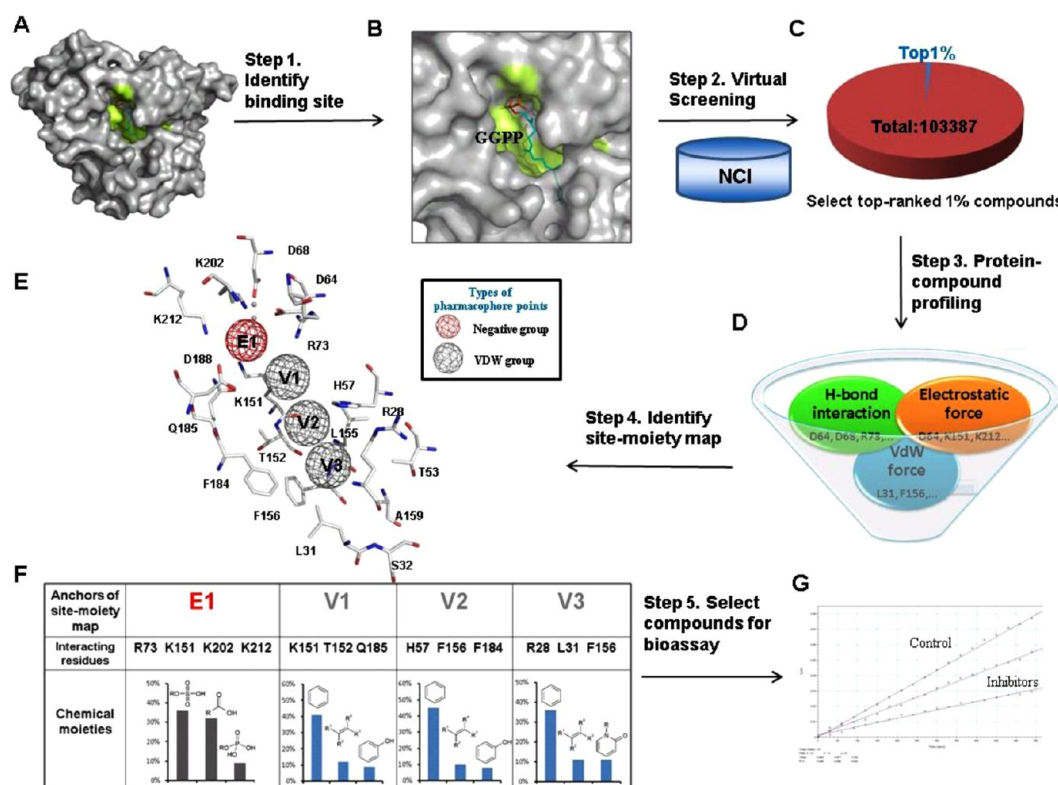
Geranylgeranyl diphosphate synthase (GGPPS) is a member of the *trans*-prenyltransferase family, which a *trans*-double bond is formed during the  $C_5$  isopentenyl diphosphate (IPP) condensation with  $C_{15}$  farnesyl diphosphate (FPP). The members in this family—which synthesize the final products of  $C_{15}$  by FPP synthase (FPPS),  $C_{20}$  by GGPPS,<sup>1–3</sup>  $C_{25}$  by farnesylgeranyl diphosphate synthase (FGPPS),  $C_{30}$  by hexaprenyl diphosphate synthase,  $C_{35}$  by heptaprenyl diphosphate synthase,  $C_{40}$  by octaprenyl diphosphate synthase (OPPS),  $C_{45}$  by solanesyl diphosphate synthase, and  $C_{50}$  decaprenyl diphosphate synthase—share sequence homology and possess similar 3D structures consisting of 15  $\alpha$ -helices connected by loops.<sup>4–10</sup> The  $C_{20}$  product of GGPP is the precursor of biomolecules such as chlorophylls,  $\alpha$ -tocopherol, or longer prenyl diphosphates used in quinine biosynthesis,<sup>11</sup> ent-kaurene, taxadiene,<sup>12</sup> and phytoene.<sup>13</sup> Moreover, GGPP and FPP are ligands for protein prenylation, a post-translational modification for signaling proteins (e.g., Ras, Rho, Rab, and Rac). FPPS and GGPPS have been identified as pharmaceutical targets for the treatment of cancers and bone resorption diseases such as osteoporosis, hypercalcemia, and metastatic bone disease.<sup>14,15</sup>

Zoledronate (Zometa), a potent bisphosphonate inhibitor of human FPPS (hFPPS)<sup>15,16</sup> but not human GGPPS

(hGGPPS),<sup>17</sup> is directed against protein farnesylation.<sup>14</sup> However, the dose in blood and soft-tissue is very low after oral administration, because approximately half of the drug is rapidly adsorbed to bone and the remainder is rapidly excreted unchanged via the kidney.<sup>18</sup> Several strategies—e.g., reducing bone affinity by deleting the hydroxyl group<sup>19</sup> and increasing side-chain lipophilicity for enhancing cellular uptake<sup>20</sup>—have been proposed to optimize the pharmacokinetic properties of bisphosphonates. Although these approaches have shown promising results *in vitro* and in animal models, the physicochemical properties of the resulting molecules are still dominated by the bisphosphonate moiety. Recently, the first nonbisphosphonate FPPS inhibitor that binds to an allosteric site on FPPS was identified.<sup>18</sup> However, there is no nonbisphosphonate GGPPS specific inhibitor so far. In considering the poor pharmacokinetic properties of bisphosphonates and no GGPPS specific inhibitors, we attempted to develop nonbisphosphonate inhibitors of GGPPS.

Here, we propose a moiety-linkage-based method to identify nonbisphosphonate hGGPPS inhibitors. Our method integrates site-moiety map (SiMMap)<sup>21</sup> with chemical structure rules (CSRs) strategy to derive pharmacophore rules of hGGPPS for

Received: April 18, 2013



**Figure 1.** Overview of moiety linkage-based method for hGGPPS inhibitors. (A) Identify protein surface and (B) binding site (in green) of human GGPPS using cocrystal structure (PDB code 2Q80). (C) Identify potential inhibitors for human GGPPS from NCI compound library using the in-house tool GEMDOCK. (D) Generate residue-compound interaction profiles, including H-bonds, electrostatic forces, and vdW force interactions, by selecting the top-ranked 1% (~1000) of compounds based on the piecewise linear potentials (PLP) of GEMDOCK. (E) Identify site-moiety map and (F) moiety compositions of the human GGPPS inhibitory-site based on conserved residue-compound interactions between binding pocket and chemical moieties. (G) Perform *in vitro* bioassay for compound candidates using rank-based consensus scores by combining pharmacophore-based and CSR-based scores (P+R scores).

identifying new nonbisphosphonate hGGPPS inhibitors from a large compound-library. Descriptive statistics and model performance revealed that the scores of CSRs and SiMMap are highly correlated to the half-maximal inhibitory concentration ( $IC_{50}$ ) for hGGPPS. These results illustrate the feasibility of using our derived hGGPPS pharmacophore rules and scoring methods to identify adapted hGGPPS inhibitors. In addition, the nonbisphosphonate compound was the specific hGGPPS inhibitor; conversely, it was not the inhibitor for human FPPS, *E. coli* UPPS, and some family members (such as yeast GGPPS and *E. coli* OPPS). Further, we explored what makes the inhibitor specifically binding to hGGPPS and found that K202 and K212 in hGGPPS are likely the key residues for providing such unique binding mechanism. The site-directed mutagenesis on these two Lys residues gained the inhibitory activity of zoledronate against hGGPPS at low micromolar concentrations. Our method and moiety-linkage maps are useful to discover potential inhibitors for the treatment of bone diseases and cancers and binding mechanisms for pharmaceutical targets.

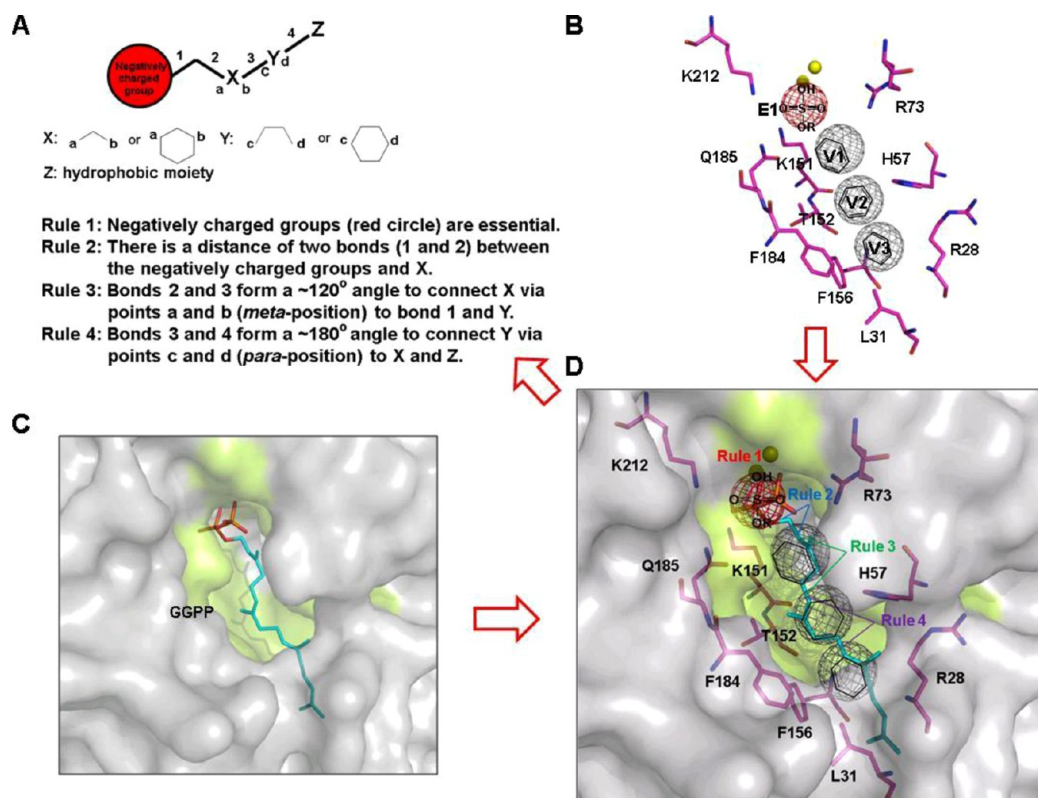
## MATERIALS AND METHODS

**Overview of Moiety Linkage-Based Method for hGGPPS Inhibitors.** Figure 1 illustrates the major steps of the moiety-linkage-based method for identifying hGGPPS inhibitors according to the site-moiety map (SiMMap) and chemical structure rules (CSR). First, we identified the binding site of hGGPPS using the pocket of the product and inhibitor

of the crystal structure (PDB code 2Q80\_chain A;<sup>9</sup> Figures 1A and 1B) in Protein Data Bank (PDB<sup>22</sup>). Our in-house docking tool (GEMDOCK<sup>23,24</sup>) was utilized to screen compound libraries (e.g., >100,000 compounds), and then we selected the top-ranked 1% of compounds (~1000 compounds; Figure 1C) based on the piecewise linear potential (PLP) score of GEMDOCK for generating the residue-compound interacting profiles (Figure 1D). After the interacting profiles were generated, we inferred the site-moiety map (Figure 1E), which consists of several anchors (Figure 1F), according to statistically significant interaction residues and moieties.<sup>21</sup> Finally, we selected top-ranked compounds for the bioassay by combining SiMMap and CSR scores (Figure 1G).

**Virtual Screening.** To discover the small-molecule inhibitors of hGGPPS, the hGGPPS cocrystal structure with the geranylgeranyl diphosphate (PDB code 2Q80;<sup>9</sup> Figures 1A and 1B) was downloaded from the PDB. The inhibitor binding site of hGGPPS was defined as the residues within a 10 Å radius sphere centered around the inhibitory product GGPP in FPP-GGPP site.<sup>15</sup> The coordinates of the atoms in the binding pocket were obtained from the PDB. Compound libraries used for virtual screening included the compounds from the National Cancer Institute (NCI) and Sigma-Aldrich (St. Louis, MO). The compounds were selected for docking if their molecular weights ranged between 200 and 700 Da. Finally, we selected 103,387 compounds.

We then used GEMDOCK to screen these compounds and selected the candidate compounds that fit into the binding site



**Figure 2.** Chemical structure rules (CSRs) and moiety linkage map of hGGPPs inhibitors. (A) Four chemical structure rules (CSRs) are derived from the moiety linkage map of the hGGPPs inhibitory site. These four CSRs are described in Supplemental Figures S2–S4 in detail. (B) The highest-percentage moiety and the respective conserved interacting residues of each anchor in our hGGPPs SiMMap. (C) Co-crystal structure of hGGPPs (PDB code 2Q80) and its inhibitor. (D) Superimposed GGPP binding mode combining the hGGPPs SiMMap and four CSRs.

of hGGPPs. GEMDOCK is an in-house program and easy to use. In addition, GEMDOCK used piecewise linear potential (PLP) that is a simple scoring function and was comparable to some scoring functions for estimating binding affinities.<sup>25–27</sup> Our previous studies revealed that GEMDOCK has similar performance to other docking methods such as DOCK,<sup>28</sup> FlexX,<sup>29</sup> and GOLD.<sup>24,30,31</sup> Furthermore, GEMDOCK has been successfully applied to identify the inhibitors and binding sites for some targets.<sup>32–35</sup> We then evaluated the accuracy of GEMDOCK by docking several known GGPPs inhibitors, GGPP (PDB code 2Q80), BPH-675 (PDB code 2E95<sup>15</sup>), and BPH-628 (PDB code 2E9A<sup>15</sup>), into the binding site because their structures are diverse. The docked conformations of three known compounds with the lowest scoring values were compared with the crystal structures based on the root-mean square deviation (RMSD) of heavy atoms. The average RMSD from GEMDOCK was less than 2.5 Å for these three compounds. Thus, the derived GEMDOCK parameters were adapted to screen thousands of commercially available compounds to identify nonbisphosphonate inhibitor candidates. After the docking procedure, we selected top-ranked 1% of compounds ( $\sim 1000$ ) based on the PLP scores of to construct three residue-compound interaction profiles, including electrostatic (E), hydrogen-bonding (H), and van der Waals (V), for inferring the site-moiety map<sup>21</sup> (Figures 1C and 1D).

**Site-Moiety Map.** Based on the residue-compound interaction profiles, we constructed the site-moiety map (SiMMap)<sup>21</sup> with several anchors describing the interaction preferences between protein pockets and compound moieties (Figures 1E and 1F). The SiMMap, using thousands of docked

poses generated by virtual screening tools (e.g., GEMDOCK, GOLD,<sup>36</sup> or AutoDock<sup>37</sup>), elucidates protein–ligand binding mechanisms and enriches the screening accuracy for post-screening analysis.<sup>38–40</sup> An anchor, which can be considered as a binding environment or a pharmacophore spot, possesses three essential elements: (i) conserved interacting residues of a binding subsite (i.e., a part of the binding site); (ii) moiety preference of this binding subsite; and (iii) an interaction type (i.e., electrostatic (E), hydrogen-bonding (H) or van der Waals (V)) between the moieties and the binding subsite. Our previous works show that SiMMAPs can present moiety preferences and physicochemical properties of binding sites and decipher the binding mechanisms.<sup>38,41</sup>

An interaction profile consists of  $L$  representative compounds (here,  $L = 1000$ ) and  $R$  interacting residues of the FPP-GGPP site. For the E or H profiles, an entry ( $C_{ij}$ ) was set to 1 if the compound  $i$  forms electrostatic or hydrogen-bonding interactions with the residue  $j$ ; otherwise,  $C_{ij}$  was set to 0. For the V profile, an entry ( $C_{ij}$ ) was set to 1 if the van der Waals interaction was less than  $-4$  kcal/mol. After the generation of the interaction profiles, we calculated the  $z$ -score value ( $Z_j$ ) of the interacting residue  $j$  to measure the interaction conservation between residue  $j$  and  $L$  representative compounds. The interacting conservation is treated as a binomial distribution, which is approximated as a normal distribution when either  $p \leq 0.5$  and  $Lp > 5$  or  $p > 0.5$  and  $L(1-p) > 5$ , where  $L$  is the number of selected compounds, and  $p$  is the probability of a compound forming an interaction. Theoretically, at least 500 compounds (here, we selected 1,000 compounds) should be selected to construct a target-



compound interaction profile. The  $Z_j$  is given as  $Z_j = f_j - \mu/\sigma$ , where  $f_j = \sum_{i=1}^L C_{ij}/L$ , and standard deviation ( $\sigma$ ) and mean ( $\mu$ ) are derived from 1,000 randomly shuffled profiles. Interactions between compounds and the residue  $j$  with a Z-score  $\geq 1.645$ , a commonly used statistical threshold, were regarded as consensus interactions. The spatially neighboring consensus residues with their interaction moieties form an anchor.

#### Moiety-Linkage Map and Chemical Structure Rules.

To discover nonbisphosphonate chemical structures with high binding affinity in the binding site of hGGPPS, we faced this challenge without known nonbisphosphonate inhibitor structures. To address this issue, we inferred the chemical structure rules to design the inhibitor candidates by connecting the preferred moiety compositions of each anchor in the hGGPPS site-moiety map (Figure 2A). We first superimposed the docked preferred moieties of three anchors (V1, V2, and V3; Figure 2B) and the product (or inhibitor), GGPP, of hGGPPS based on crystal structure (PDB code 2Q80;<sup>9</sup> Figure 2C). The compositions show that the preferred moieties of three hydrophobic anchors (benzene rings) and the hydrophobic part of GGPP were fitted well (Figure 2D). According to docked poses of the top-ranked 1,000 candidates, we selected well-fitted docked compounds and observed the consensus moiety linkages between any two moieties located in these three anchors. These results allowed us to derive four chemical structure rules (CSRs) and moiety linkages to construct the scaffold of potential inhibitor candidates for drug targets. We inferred the moiety-linkage map of the inhibitory site for discovering the nonbisphosphonate inhibitors for hGGPPS (Figure 2A and 2D).

**Chemical Structure Rule Scores.** After the moiety-linkage map was generated, we calculated the CSR score for a compound  $x$  as follows:  $S(x) = \sum_{c=1}^n A_c(x)W_c$  where  $A_c(x)$  is the CSR score of compound  $x$  in the anchor  $c$ ;  $W_c$  is the weight of the anchor  $c$ ; and  $n$  is the number of anchors in a binding site.  $A_c(x)$  is 1 if the moiety of compound  $x$  fits the related CSR and is located in the anchor  $c$ ;  $A_c(x)$  is 0.5 when the moiety of compound  $x$  only conforms to one of these two events.  $W_c$  is based on the importance of each CSR. In this study,  $W_c$  was 3, 3, 1, and 3 for the anchor E1, the linkage of E1 and V1, the connection of V1 and V2, and linkage between V2 and V3, respectively. The CSR score can be used to obtain new rankings of screened compounds and known bisphosphonate inhibitors.

**Materials for Bioassays.** *PfuTurbo*, the plasmid miniprep kit, DNA gel extraction kit, and Ni-NTA resin were purchased from Qiagen. The protein expression kit (including the pET-28a vector and competent JM109 and BL21 cells) was obtained from Novagen. The QuikChange site-directed mutagenesis kit was obtained from Stratagene. FPP, IPP, FOH, and GGOH were obtained from Sigma. All commercial buffers and reagents were of the highest grade.

**Expression and Purification of Wild-Type and Mutated Human GGPPS.** The wild-type and mutated hGGPPS were expressed in *E. coli* and purified using NiNTA chromatography as previously described.<sup>8</sup> The constructed plasmids were used to transform *Escherichia coli* JM109 competent cells at the same time, and the transformed cells were streaked on a Luria–Bertanini (LB) agar plate containing 100  $\mu\text{g/mL}$  ampicillin. Ampicillin-resistant colonies were selected from the agar plate and grown in 5 mL of LB culture containing 100  $\mu\text{g/mL}$  ampicillin overnight at 37 °C. The

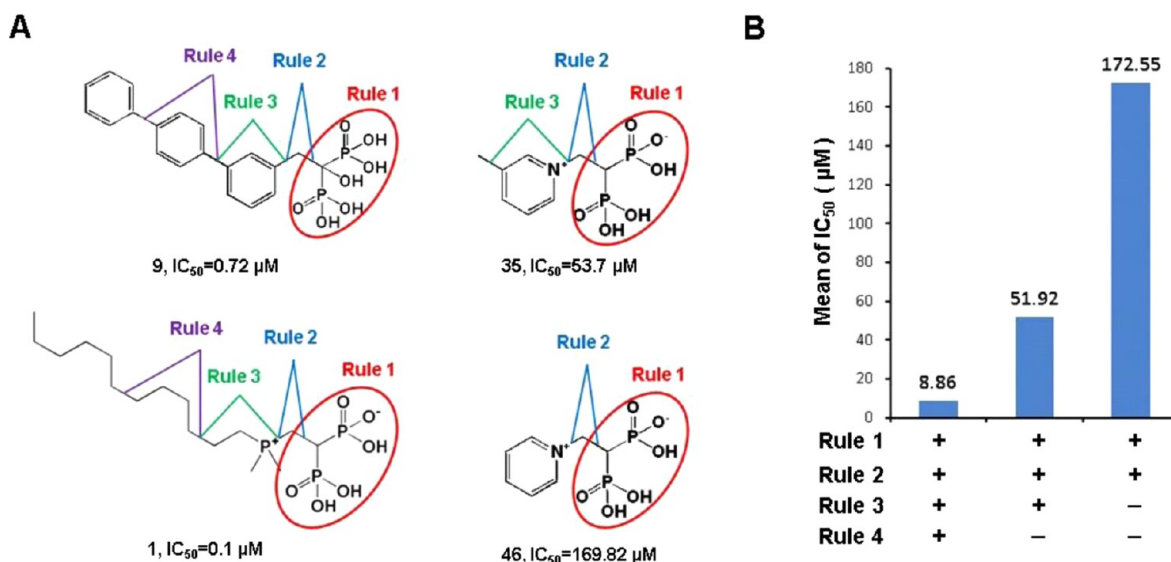
correct constructs, as confirmed by sequencing, were transformed into *E. coli* BL21(DE3) for protein expression.

A 20-mL overnight culture of a single transformant was used to inoculate 2 L of fresh LB medium containing 100  $\mu\text{g/mL}$  ampicillin. The cells were grown to  $A_{600} = 0.6$  and induced with 1 mM IPTG at 20 °C. After 16 h, the cells were harvested by centrifugation at 6000  $\times g$  for 20 min to collect the cell pellet. The enzyme purification was conducted at 4 °C. The cell pellet was suspended in 20 mL of lysis buffer containing 25 mM Tris-HCl, 150 mM NaCl, and 0.5 mM TCEP at pH 7.5. Cell lysate was prepared with a French pressure cell press (AIM-AMINCO Spectronic Instruments) and centrifuged at 20000  $\times g$  for 40 min to remove cell debris. The cell-free extract was then loaded onto the NiNTA column, which had previously been equilibrated with lysis buffer. The column was washed with 10 mM imidazole followed by 30 mM imidazole-containing buffer. His-tagged hGGPPS was eluted with 300 mM imidazole, dialyzed twice against 4 L of lysis buffer, subjected to FXa protease digestion to remove the His-tags, and then reloaded into another NiNTA column to collect the unbound flow-through. SDS-PAGE analysis was used to check the purity of hGGPPS.

**Expression and Purification of Human FPPS, *E. coli* OPPS, *E. coli* UPPS, and *S. cerevisiae* GGPPS.** Preparation of recombinant human FPPS, *E. coli* OPPS, *E. coli* UPPS, and *S. cerevisiae* GGPPS followed the methods of previous references.<sup>8,42–44</sup>

**Enzymatic-Activity Assays.** For enzymatic-activity measurements, 1  $\mu\text{M}$  human GGPPS was used. The reaction was initiated in 200  $\mu\text{L}$  solution containing 1 mM TCEP, 200  $\mu\text{M}$  MESG, 10  $\mu\text{M}$  FPP, 20  $\mu\text{M}$  IPP, 50 mM Tris, 2 mM  $\text{MgCl}_2$ , 0.2% Tween20, PNP (0.2 U), pyrophosphatase (0.006 U), DMSO (or inhibitors), and  $\text{ddH}_2\text{O}$  at 25 °C. For human FPPS, yeast GGPPS, *E. coli* OPPS, and *E. coli* UPPS activity assays, 1  $\mu\text{M}$  enzyme was used in the reaction mixtures containing 200  $\mu\text{M}$  MESG, 10  $\mu\text{M}$  FPP (or GPP), 20  $\mu\text{M}$  IPP, 100 mM Hepes–KOH (pH 7.5), 0.5 mM  $\text{MgCl}_2$ , 50 mM KCl, 0.1% Triton X-100, PNP (0.2 U), pyrophosphatase (0.006 U), DMSO (or inhibitors), and  $\text{ddH}_2\text{O}$  at 25 °C. The enzyme concentration used in all experiments was determined from its absorbance at 280 nm ( $\epsilon = 20,340 \text{ M}^{-1} \text{ cm}^{-1}$ ). The quantification of inorganic phosphate consumed in the reaction at 360 nm was detected by using ELISA. The  $\text{IC}_{50}$  values were obtained by fitting the inhibition data to the dose–response curve.

**Tumor Cell Growth Inhibition Assays.** Two human tumor cell lines, MCF-7 and MDA-MB-231, were cultured in MEM medium supplemented with 10% fetal bovine serum and 1.5 g of  $\text{NaHCO}_3$  and DMEM medium supplemented with 10% fetal bovine serum and 3.7 g of  $\text{NaHCO}_3$ , respectively. MCF-7 and MDA-MB-231 cells were plated in a 96-well plate at 100  $\mu\text{L}$  of medium per well and incubated (37 °C, 5%  $\text{CO}_2$ ) overnight to allow the cells to attach to the wells. Inhibitors dissolved in DMSO were added to each well to make various concentrations (NSC351204: from 12.5  $\mu\text{M}$  to 0.7 mM; zoledronate: from 0.0625  $\mu\text{M}$  to 0.7 mM) when the growing area of the cells covered 70% of the wells. Mixed samples were incubated for 4 days (37 °C, 5%  $\text{CO}_2$ ) to allow the drug to take effect. Ten  $\mu\text{L}$  of MTT ([3-(4,5-dimethylthiazole-2-yl)-2,5-diphenyltetrazolium bromide]) solution at 5 mg/mL in DPBS was then added to each well. After thoroughly mixing the MTT into the media, the mixture was then incubated (37 °C, 5%  $\text{CO}_2$ ) for 4 h to allow the MTT to be metabolized. The media



**Figure 3.** Relationship between CSRs and IC<sub>50</sub> values of 51 hGGPPS inhibitors. (A) The most potent hGGPPS ring-type (9, IC<sub>50</sub>: 0.72 μM) and linear-type (1, IC<sub>50</sub>: 0.1 μM) inhibitors match the four CSRs well. In addition, the IC<sub>50</sub> of compound 35, which fits rules 1–3 (53.7 μM), is better than that of compound 46, which fits only rules 1 and 2 (169.82 μM). (B) The average IC<sub>50</sub> (8.86 μM) of 26 compounds that fit four CSRs is much better than the average of those compounds fitting three (13 compounds, 51.92 μM) and two (10 compounds, 172.55 μM) CSRs.

was removed, and formazan (MTT metabolic product) was resuspended in 50 μL of DMSO in each well. The mixture's optical density was read at 540 nm by ELISA. The EC<sub>50</sub> values were obtained by fitting the inhibition data to the dose–response curve. For the “rescue” experiments, the requisite amounts of FOH and GGOH were added to the incubation media to produce a fixed 20 μM concentration.

## RESULTS AND DISCUSSION

**Site-Moiety Map of the Inhibitory Site in hGGPPS.** The site-moiety map of the inhibitory site (FPP-GGPP site)<sup>15</sup> in hGGPPS consists of one electrostatic force anchor (E1, red) and three van der Waals (vdW) anchors (V1, V2, and V3, gray; Figure 1E). To construct a site-moiety map of hGGPPS, we obtained the protein–ligand interacting profile by screening large compound-libraries using GEMDOCK.<sup>24</sup> After the site-moiety map was generated, a set of entities was identified for each anchor that was chemically related, supporting the concept that a given hotspot shares a chemical-physical binding environment with conserved binding residues (Figure 1F). Among the top-ranked ~1000 compounds, 180 formed electrostatic forces with residues R73, K151, K202, and K212 as well as chelating with magnesium in the anchor E1, and 752 formed steric interactions with residues R28, L31, and F156 in the V3 anchor (Figure 1D).

The anchor of a site-moiety map can describe the interacting preference between a binding pocket (forming by several residues) and its preferred moieties. For example, the anchor E1 is an electrostatic pocket with four residues (R73, K151, K202, and K212) that often form hydrogen bonds and electrostatic forces with three main moiety types (i.e., sulfate, carboxylate, and phosphate; Figure 1F). In addition, benzene, alkene, and toluene are the major functional groups that associate with vdW anchors V1, V2, and V3. For example, the V1 anchor consists of a pocket with conserved interactions between residues K151, T152, and Q185 and preferred bulky moieties, such as aromatic rings and toluene. V2 is situated between V1 and V3 and forms a steric pocket via residues H57,

F156, and F184. V3, which is adjacent to V2, constitutes a vdW environment with residues R28, L31, and F156.

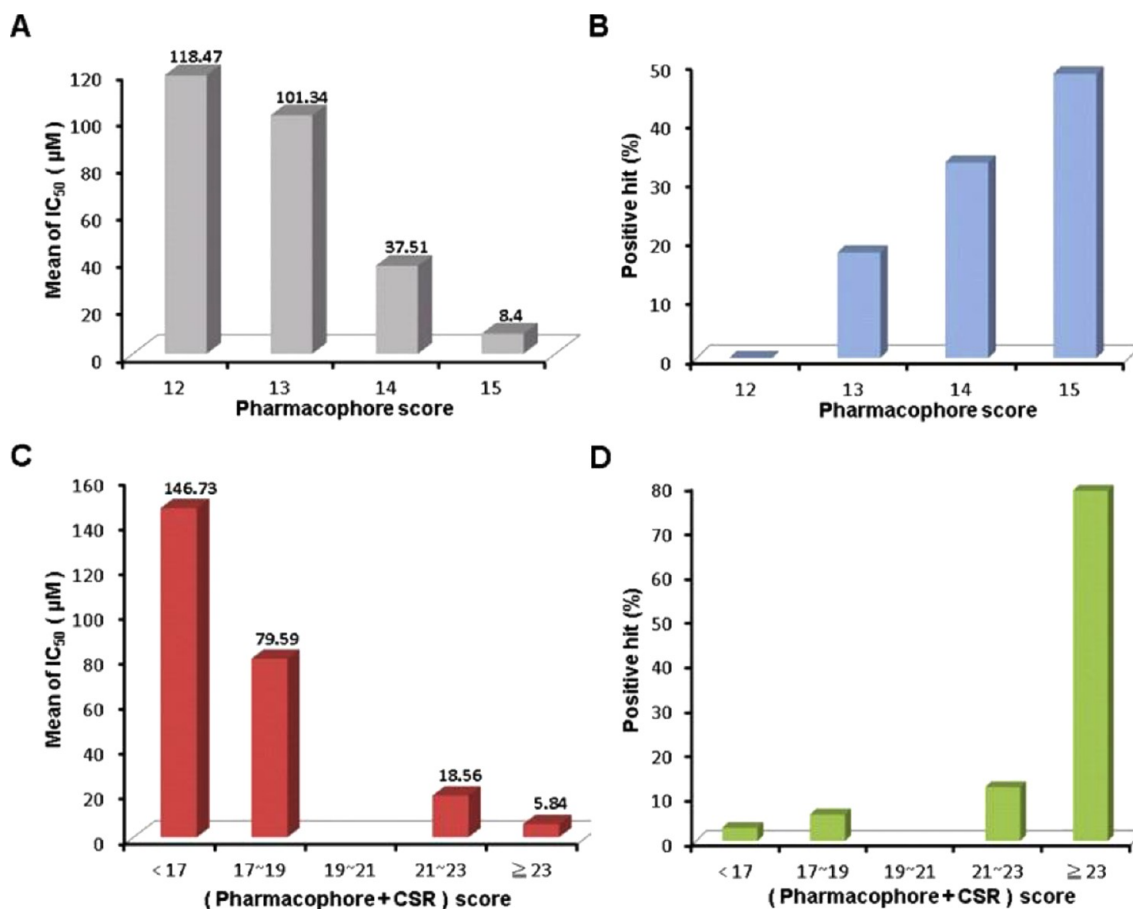
**Chemical Structure Rules for hGGPPS Inhibitors.** Chemical structure rules (CSRs), the moiety-based linkage rules which link several anchors of the hGGPPS SiMMap, can help to outline the essential chemical structures of potent inhibitors (Figure 2). To explore the moiety linker and the scaffolds of preferred compounds in the binding site, we superimposed the most-preferred moieties of anchors in the hGGPPS site-moiety map (Figures 1F and 2B) and the binding pose of product GGPP in the crystal structure (Figure 2C). The efficiently superimposed moieties of anchors V1–V3 and the hydrophobic parts of the product GGPP (Figure 2D) provided clues for linking the four anchors of the hGGPPS site-moiety and comprised the moiety-linkage map of the inhibitory site in hGGPPS. We were then able to identify four CSRs (Figure 2A) linking the functional groups to describe probable chemical structures of hGGPPS inhibitors based on the moiety-linkage map. Rule 1 states that a negatively charged group (i.e., sulfate, carboxylate, or phosphate) is essential in anchor E1, based on the site-moiety map; Rule 2 indicates that anchor E1 prefers a distance of two bonds to optimally connect with anchor V1, especially if the moiety in anchor V1 is a ring-type (Supplementary Figure S2); Rule 3 states that anchor V1 is linked to anchors E1 and V2 by two bonds that form a ~120° angle to connect the moiety in V1 via positions a and b (*meta*-position) to neighboring moieties (Supplementary Figure S3); Rule 4 indicates that the linkage between anchor V2 and neighboring anchors (V1 and V3) should have two bonds that form a ~180° angle (linear bond) to connect the moiety in V2 via positions c and d (*para*-position) to neighboring functional groups in V1 and V3, respectively (Supplementary Figure S4).

**Evaluation of the Chemical Structure Rules Score.** To verify the relationship between the four CSRs and inhibitory efficiency, we collected 51 hGGPPS bisphosphonate inhibitors that are known from previous work to bind to the FPP-GGPP site under investigation<sup>45</sup> (Figure 3 and Supplementary Figure S1). Among these 51 inhibitors, all of the inhibitors fit Rule 1

Table 1. CSR Statistics Using Mann–Whitney Tests Based on 51 Known Compounds

test rule	Event 1			Event 2			$p$ -value <sup>d</sup>
	fitting rules	$n_1^a$	$U_1^c$	fitting rules	$n_2^b$	$U_2^c$	
Rule 3	Rules 1–3	13	116	Rules 1–2	10	14	$1.56 \times 10^{-3}$
Rule 4	Rules 1–4	26	286.5	Rules 1–3	13	51.5	$4.64 \times 10^{-4}$
Rules 3 + 4	Rules 1–4	26	256	Rules 1–2	10	4	$8.60 \times 10^{-6}$

<sup>a</sup>The number of compounds fitting the rules of Event 1. <sup>b</sup>The number of compounds fitting the rules of Event 2. <sup>c</sup> $U_i = n_1 n_2 + (n_1(n_1+1)/2) - \sum R_i$ ,  $\sum R_i$  is the sum of the ranks for the event  $i$ . <sup>d</sup>The  $p$ -value is calculated between Events 1 and 2 (the confidence level is 99%,  $\alpha = 0.01$ ).



**Figure 4.** Performances of pharmacophore and ranked-based (pharmacophore+CSR, P+R) scores for 51 hGGPPS inhibitors. Among these 51 inhibitors, 33 compounds ( $IC_{50} \leq 10 \mu M$ ) were considered to be positive compounds. The relationships between the pharmacophore scores and (A) means of  $IC_{50}$  values as well as (B) positive hits show that high pharmacophore scores imply low  $IC_{50}$  values and that 48.5% (16/33) of positive compounds have a pharmacophore score of 15. (C) and (D) display the relationships of P+R scores with  $IC_{50}$  means and with positive hits, respectively. High P+R scores produce low  $IC_{50}$  values and can discriminate positive (~79%, 26/33) and nonpositive compounds. The P+R score significantly outperforms the pharmacophore score alone for identifying hGGPPS inhibitors.

and 50 inhibitors (the exception being compound 51) fit Rule 2. In addition, 33 compounds ( $IC_{50} \leq 10 \mu M$ ) were considered to be positive matches for potent hGGPPS inhibition in this study. 70% of positive inhibitors (23/33), such as the most potent inhibitors of ring-type (9,  $IC_{50} = 0.72 \mu M$ ) and linear-type (1,  $IC_{50} = 0.1 \mu M$ ), fit all four CSRs. Conversely, the  $IC_{50}$  values of compounds 35 and 46 were  $53.7 \mu M$  (fitting Rules 1–3) and  $169.82 \mu M$  (fitting Rules 1–2), respectively (Figure 3A). We then clustered 51 inhibitors into three groups by using CSRs and calculated the average  $IC_{50}$  values of each group. The average  $IC_{50}$  of the first group, compounds fitting all four CSRs, was the best ( $8.86 \mu M$ ) and significantly better than the second ( $51.92 \mu M$ ) and the third groups ( $172.55 \mu M$ ), which consist of compounds that fit only Rules 1–3 and Rule 1–2, respectively (Figure 3B). Furthermore, we utilized a statistical

method—i.e., the Mann–Whitney U Test—to confirm the effects of Rule 3, Rule 4, and combined Rules 3–4 on  $IC_{50}$  values. The  $p$ -values of comparing Rules 1–4 with Rules 1–3 and with Rules 1–2 were  $4.64 \times 10^{-4}$  and  $8.6 \times 10^{-6}$ , respectively. The  $p$ -value is  $1.56 \times 10^{-3}$  for the comparison of Rules 1–3 and Rules 1–2. These statistical results imply that Rule 3 and Rule 4 are significantly involved in the inhibition of hGGPPS (Table 1) for these 51 known inhibitors, results that are consistent with previous work.<sup>15</sup> They indicate that compounds with large and hydrophobic moieties in anchors V2 and V3 can increase the potency of hGGPPS inhibition. These results demonstrate that the CSRs are indeed related to inhibitory efficiency and useful for the identification of potent hGGPPS inhibitors.



Table 2. Summary of hGGPPS Inhibitors by Enzyme Assay Results, Chemical Structure Rules, P+R Scores, and Site-Moiety Maps

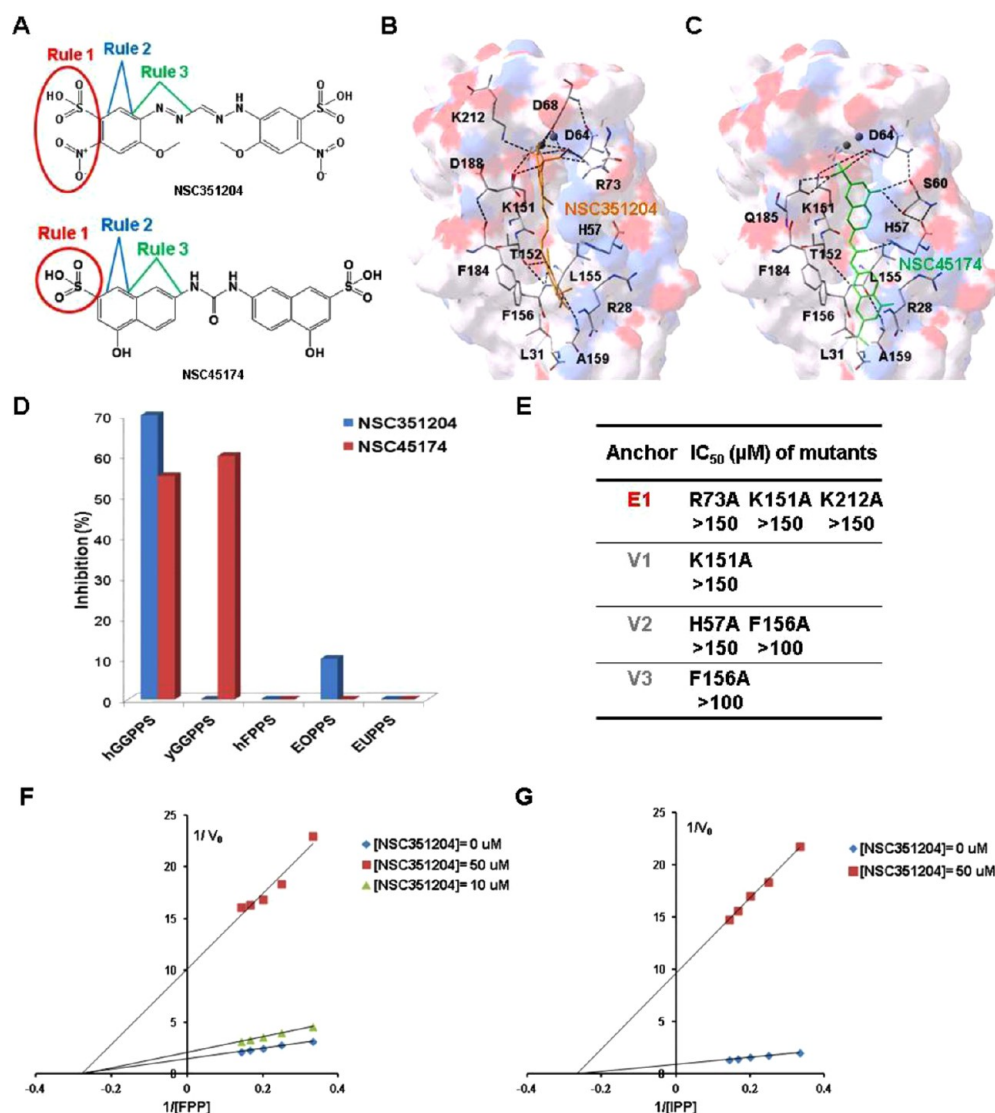
Compound ID	IC <sub>50</sub> (μM)	Rule 1	Rule 2	Rule 3	Rule 4	P+R score	Compound structure and Site-moiety map
1	0.1	●	●	●	●	24	
36	66.07	●	●	●		15	
44	117.49	●	●			13	
9	0.72	●	●	●	●	25	
34	14.13	●	●	●		18	
45	141.25	●	●			15	
NSC 351204	31.41	●	●	●		16	
NSC 45174	48.3	●	●	●		14	

**Evaluation of Pharmacophore Score.** To evaluate the pharmacophore spots (i.e., anchors) and pharmacophore scores to provide biological insights and guide the drug discovery process, we used the pharmacophore scores of 51 known ligands based on interacting residues with statistically significant Z-scores ( $\geq 1.645$ ) in protein small-molecule interaction profiles (Figure 4). Based on pharmacophore score ( $S$ ), we grouped these 51 inhibitors into four clusters: cluster 1 ( $S \geq 15$ ), cluster 2 ( $S = 14$ ), cluster 3 ( $S = 13$ ), and cluster 4 ( $S \leq 12$ ). For each cluster, we computed the average IC<sub>50</sub> and found that the pharmacophore score is highly correlated to IC<sub>50</sub> value potency. The mean IC<sub>50</sub> of the compounds in highest-scoring cluster 1 was 8.4 μM and was significantly more potent than any of the other clusters (mean IC<sub>50</sub> = 37.51 μM, 101.34 μM, and 118.47 μM, in descending order; Figure 4A). 48.5% (16/33) and 33.3% (11/33) of positive inhibitors were in clusters 1 and 2, respectively (Figure 4B). Conversely, 66.7% (12/18) of compounds with IC<sub>50</sub> > 10 μM were in clusters 3 and 4 ( $S \leq 13$ ; Supplementary Figure S5A). The results indicate that our proposed pharmacophore scoring is useful in discriminating the potency of the positive hits from among 51 inhibitors. However, the pharmacophore score may bias toward bisphosphonate inhibitors, because these 51 known inhibitors possess bisphosphonate moieties that contributed the bulk of their scores.

**Evaluation of Combining the Site-Moiety Map with CSRs.** Following the site-moiety map analysis and the compound-anchor-residue profile (Figure 1D), we used a novel postscreening analysis to rescore docked compounds by combining pharmacophore (P) and CSR (R) scores (Table 2). We utilized this strategy to improve the accuracy of identifying the adaptive chemical structures for hGGPPS inhibitors. After

CSRs and the SiMMap were identified for hGGPPS, the CSR score was used to refine adaptive chemical structures of compound candidates based on the extent of compound moieties matching the pharmacophore spots (anchors) of SiMMap. To avoid the individual biases of the pharmacophore- and CSR-scoring methods, a combined scoring method (P+R) was utilized for selecting candidate compounds from the virtual screening for use in bioassays (Figure 1G). Based on the (P+R) scores for 51 known inhibitors, the mean IC<sub>50</sub> value of inhibitors in the top-scoring group (combined score  $\geq 23$ ) was 5.84 μM (Figure 4C). The high P+R scores imply the potent inhibitory efficiency based on mean IC<sub>50</sub> values for hGGPPS. Additionally, the positive rate of that highest score group was 79% (26/33; Figure 4D), an increase of 30% against pharmacophore scores alone.

On the other hand, the P+R scores are low for ~89% (16/18) of nonpositive inhibitors can be effectively separated from 33 positive hits by our strategies (score  $\leq 19$ ; Supplementary Figure S5B and Figure 4D). The only two exceptions were compounds 25 (IC<sub>50</sub> = 79.43 μM) and 26 (IC<sub>50</sub> = 99.33 μM). These two inhibitors possess charged groups, matching anchor E1, but their long C–C chains with polar atoms or charged groups form poor hydrophobic interactions in anchors V2 and V3. These results demonstrate that combining the site-moiety map with CSRs strategy not only removed the bias of the pharmacophore, but also greatly advanced the performance of recognizing the positive inhibitors. For example, the P+R scores of compounds 8 (IC<sub>50</sub> = 0.71 μM) and 47 (IC<sub>50</sub> = 181.97 μM) are 23 and 13.5, respectively, while their pharmacophore scores are identical (i.e., 13). After the combined (P+R) scoring method was established, we applied it to identify potential



**Figure 5.** Characterization of nonbisphosphonate hGGPPS inhibitors by moiety linkage map, enzyme assay, site-directed mutagenesis, and molecular docking. (A) Chemical structures of two nonbisphosphonate inhibitors, NSC351204 and NSC45174, identified by our model. They fit CSRs 1, 2, and 3. The docked conformations of (B) NSC351204 and (C) NSC45174 are orange and green, respectively. The hydrogen bonds are indicated by black dashed lines, and atoms within the residues are displayed using the CPK model (oxygen in red, nitrogen in blue, and carbon in gray). (D) Enzyme-activity experiments showed that NSC351204 and NSC45174 are selective hGGPPS inhibitors as they are both inactive for other drug targets (human FPPS and *E. coli* UPPS) and *E. coli* OPPS. Notably, NSC351204 is also inactive for yGGPPS. (E) Comparisons of relative inhibitory effects of five hGGPPS mutants based on the significant interacting residues of the moiety linkage map. The Lineweaver–Burk plots show that NSC351204 is a noncompetitive hGGPPS inhibitor that does not compete with the two substrates, (F) FPP and (G) IPP. Our model reveals selective nonbisphosphonate inhibitors in the hGGPPS inhibitory site.

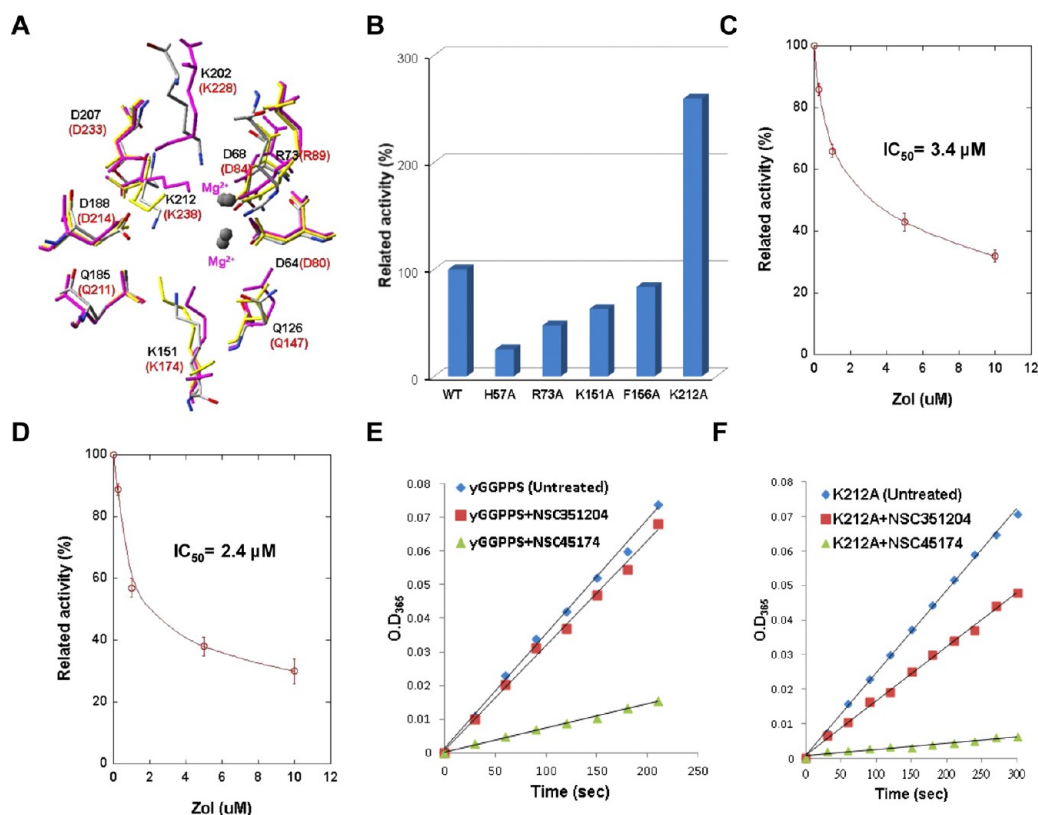
nonbisphosphonate inhibitors from screening available compounds on a large scale.

**Identification of Selective Nonbisphosphonate hGGPPS Inhibitors.** After screening the compounds in the NCI library using GEMDOCK and inferring the site-moiety map (pharmacophore scores) of the inhibitory site (FPP-GGPP site)<sup>15</sup> of hGGPPS, we used the P+R scoring method to select 10 available compounds for evaluation by bioassay. Among these 10 compounds, we discovered two nonbisphosphonate inhibitors of hGGPPS, NSC351204 (IC<sub>50</sub> = 31.41 μM) and NSC45174 (IC<sub>50</sub> = 48.3 μM) (Figure 5A and Table 2). The two compounds consistently docked into the FPP-GGPP site of hGGPPS (Figures 5B and 5C) because they are selective inhibitors for hGGPPS but not human FPPS (Figures 5D and 7C), *E. coli* UPPS, and *E. coli* OPPS (Figure 5D). Site-directed

mutagenesis of five interacting residues (R73A, K151A, K212A, H57A, and F156A) identified by SiMMap shows that these nonbisphosphonate compounds should occupy the hGGPPS inhibitory site (Figure 5E). Furthermore, experimental results also demonstrate that they are noncompetitive with the two substrates FPP and IPP (Figures 5F and 5G).

We then evaluated the binding and pharmacophore models of NSC351204 and NSC45174 (Figures 5B and 5C and Table 2). The docked pose of compound NSC351204 fits well with the four hotspots of the FPP-GGPP binding site; conversely, the docked pose of compound NSC45174 possesses only one negatively charged group (sulfate) which forms weak electrostatic interactions with interacting residues R73, K202, and K212. This difference of electrostatic interactions may be responsible for the slightly lower inhibitory effects of



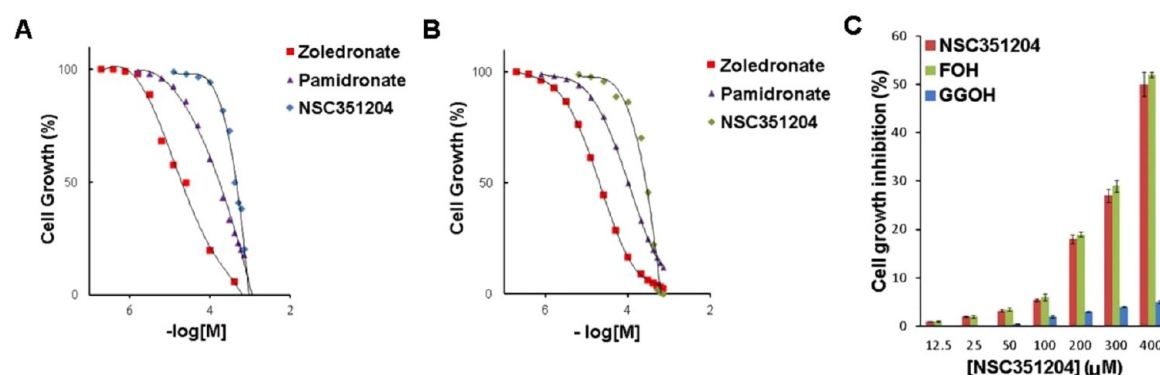


**Figure 6.** Residues K202 and K212 for hGGPPS-specific inhibition. (A) Superimposition of crystal structures of human GGPPS (PDB code 2Q80, gray) and two yeast GGPPS (PDB code 2E91, yellow and PDB code 2E8V, magenta). Two main differences between hGGPPS and yGGPPS are as follows: 1) K202 in hGGPPS is more stable than K228 in yGGPPS because K228 is missing in 90% of cocrystal structures of yGGPPS; 2) The side-chains of K212 (human) and K238 (yeast) orient in different directions. (B) The catalytic assay results of hGGPPS (WT) and five mutants. All mutants exhibit lower activity compared to WT except K212A. K202G also increases catalytic activity similar to K212A (not shown). These observations suggest that K202 and K212 may be key residues for selectively inhibiting prenyltransferases. Site-directed mutagenesis and enzymatic-activity results produced  $IC_{50}$  values for zoledronate on K202G (C) and K212A (D) of 3.4  $\mu$ M and 2.4  $\mu$ M, respectively. The inhibitory efficacy of zoledronate on these two mutated hGGPPS is increased 29- to 42-fold compared with the wild-type hGGPPS ( $IC_{50}$  = 100  $\mu$ M). (E-F) NSC45174—which does not interact with K202 or K212—inhibits both yGGPPS and hGGPPS (K212A), but NSC351204—which interacts with K212—efficiently inhibits only hGGPPS.

NSC45174 (Table 2). In addition, both compounds NSC351204 and NSC45174 do not follow Rule 4 linking anchors V2 and V3. They also lack large hydrophobic moieties that form strong vdW interactions with residues H57 and F156 in anchor V2. These two compounds, which are more potent than a conventional bisphosphonate against hGGPPS (i.e., zoledronate; compound **42**,  $IC_{50}$  = 100  $\mu$ M; Supplementary Figure S1), are a new scaffold of GGPPS inhibitors and provide an opportunity for designing new leads for the bone and related mevalonate-pathway diseases. These results reveal that the negatively charged groups and hydrophobic moieties of compounds that form powerful electrostatic (anchor E1) and strong vdW forces (anchors V2 and V3) are both important for hGGPPS inhibition. It is feasible that the potency of these two compounds could be improved to make novel candidates for hGGPPS inhibitors by following the P+R strategy.

**Site-Directed Mutagenesis of Significant Interacting Residues in the hGGPPS Inhibitory Site.** An anchor of a binding pocket possesses conserved interacting residues that constitute a specific physicochemical property, and it often engages in enzymatic functions. We investigated the roles of these conserved anchor residues of the site-moiety map in the inhibitory site of hGGPPS (Figures 1F and 2B). Each of the selected residues was replaced with alanine and expressed in

*Escherichia coli*. After purification using NiNTA affinity chromatography, all mutants migrated to a major band of an apparent molecular mass of ~35 kDa in SDS-PAGE (Supplementary Figure S6). We first investigated mutants of three residues (R73, K151, and K212) of anchor E1 forming electrostatic interactions with negatively charged moieties (i.e., phosphate) of inhibitors. Enzymatic-inhibition experiments revealed that mutating these positively charged residues remarkably reduced the inhibitory efficacy ( $IC_{50}$  > 150  $\mu$ M; Figure 5C), suggesting their contributions to inhibitor binding are major. In addition, the results of enzymatic-activity experiments showed that R73A and K151A mutants reduced activity by 53% and 40% compared with the wild type, respectively. Residues R73 and K151 of hGGPPS, which correspond to R84 and K169 of yeast GGPPS, have been proven to be the primary residues involved in binding diphosphates.<sup>9,15</sup> Notably, the replacement of the amino acid K212 with Ala increased the catalytic activity (Figure 6B). Interestingly, the corresponding mutation (K235L) in *E. coli* OPPS causes the  $k_{cat}$  value to be approximately  $10^3$ -fold smaller than in the wild type.<sup>46</sup> This result implies that K212 may be associated with the selective inhibitory mechanisms of prenyltransferases.



**Figure 7.** Breast cancer cell growth inhibitions. (A) MCF-7 cell and (B) MDA-MB-231 cell growth inhibitions by two FDA bisphosphonate drugs (zoledronate and pamidronate) and NSC351204 measured by MTT assay. For the MCF-7 cell assay,  $\text{EC}_{50}$  values of zoledronate, pamidronate, and NSC351204 are  $\sim 24 \mu\text{M}$ ,  $\sim 151 \mu\text{M}$ , and  $\sim 400 \mu\text{M}$ , respectively. For the MDA-MB-231 cell assay,  $\text{EC}_{50}$  values of zoledronate, pamidronate, and NSC351204 are  $\sim 20 \mu\text{M}$ ,  $\sim 100 \mu\text{M}$ , and  $\sim 300 \mu\text{M}$ , respectively. NSC351204 is the first nonbisphosphonate compound inhibiting hGGPPS in the mevalonate pathway to cause tumor cell apoptosis. (C) FOH and GGOH rescue of MCF-7 cell growth inhibition.

For the interaction residues of vdW anchors V1 (K151), V2 (H57 and F156), and V3 (F156), the mutation K151A exhibited less than 50% detectable inhibitory efficacy when treated with  $150 \mu\text{M}$  NSC351204 ( $\text{IC}_{50} > 150 \mu\text{M}$ ; Figure 5C) because of the loss of not only its electrostatic force/H-bond interactions with negatively charged inhibitor moieties but also the partial loss of the hydrophobic force in anchor V1. The residue H57 often forms hydrogen bonds with the phosphate group of IPP, facilitating IPP binding and catalysis. Furthermore, our model suggests that residues H57 and F156 play a critical role by forming hydrophobic environments in anchors V2 and V3 (Figures 1E and 1F), which are preferred by inhibitors with large hydrophobic moieties. As expected, the mutation H57A showed only 25% of wild type enzymatic activity and caused a  $>5$ -fold increase in the  $\text{IC}_{50}$  value ( $\text{IC}_{50} > 150 \mu\text{M}$ ; Figures 5E and 6B). When F156 was replaced with alanine (A156), this mutation exhibited a 48% inhibitory effect at a concentration of  $100 \mu\text{M}$  NSC451204 ( $\text{IC}_{50} > 100 \mu\text{M}$ ; Figure 5C). However, F156A retained  $\sim 88\%$  of its catalytic activity because F156 is not directly associated with the catalytic reaction (Figure 6B). These site-directed mutagenesis studies revealed that these interacting residues of SiMMap participate in biological functions and inhibitor binding in the inhibitory site of hGGPPS.

**Identification of Key Residues for hGGPPS-Specific Inhibition.** To explore the binding mechanisms of hGGPPS, we analyzed the differences between hGGPPS and yGGPPS based on their sequences, X-ray structures, and the inhibitor zoledronate. Their sequence identity and similarity are 44% and 60%, respectively, using BLASTP.<sup>47</sup> The compound zoledronate is potent for the prenyltransferase family (i.e.,  $\text{IC}_{50}$  for yGGPPS:  $0.66 \mu\text{M}$ )<sup>15</sup> with the exception of hGGPPS ( $\text{IC}_{50}$  for hGGPPS:  $100 \mu\text{M}$ ).<sup>45</sup> To address this issue, we performed structure-based studies and site-directed mutagenesis experiments on the proteins of the prenyltransferase family. According to the structure-based studies, the superimposition of three crystal structures, including hGGPPS-GGPP (PDB code: 2Q80),<sup>9</sup> yGGPPS-zoledronate (PDB codes: 2E91),<sup>15</sup> and yGGPPS-GGPPS (PDB code: 2E8V),<sup>15</sup> showed that most of the residues were well aligned except for the residue K228 (yGGPPS), which corresponds to K202 in hGGPPS (Figure 6A). Among 22 crystal structures of yGGPPS in PDB,<sup>48</sup> the residue K228 missing in two structures (PDB codes: 2E8V and 2E93)<sup>15</sup> implied that K228 may be not stable. Conversely,

K202 is more stable in hGGPPS, and its side-chain orientation is significantly different from those of K228 in yGGPPS (Figure 6A). These observations suggest that K202 may play a key role in the different inhibitory effects of zoledronate for hGGPPS and yGGPPS. To verify this hypothesis, K202 was replaced with Gly202 to mimic the structural environment of yGGPPS and tested using zoledronate. Indeed, the inhibition against hGGPPS<sup>K202G</sup> of zoledronate ( $\text{IC}_{50} = 3.4 \mu\text{M}$ ) regained 29-fold more inhibitory efficacy compared to wild type (Figure 6C).

According to our model, enzymatic-activity assays, and structure-based studies, K212 in hGGPPS is a significant interacting residue for inhibitor binding, and it differs from corresponding residues in yGGPPS and *E. coli* OPPS. For example, the side-chains of K238 in yGGPPS were often not consistent with K212 in hGGPPS (Figure 6A). The enzymatic activity showed the reverse results when replacing K212 with A212 in hGGPPS and K235 with L235 in *E. coli* OPPS. These results also suggest that K212 may play an important role in selective inhibition between hGGPPS and other prenyltransferases. To confirm this idea, K212A was also tested by zoledronate. The  $\text{IC}_{50}$  of zoledronate on the hGGPPS with the mutation K212A was  $2.4 \mu\text{M}$  (Figure 6D), and mutating K212 in hGGPPS successfully induced a 42-fold relative inhibitory efficacy of zoledronate against hGGPPS, which demonstrates that K212 causes the selectivity between hGGPPS and other prenyltransferases.

To further support our hypothesis, we tested two nonbisphosphonate inhibitors, NSC351204 and NSC45174, identified from our computational strategies against K212A and yGGPPS. From our predicted binding models for these two compounds, NSC351204 interacts with K212 (Figure 5B), but NSC45174 does not interact with K202 or K212 (Figure 5C). Moreover, site-directed mutagenesis and catalytic activity assays were consistent with our predicted model and binding modes (Figure 5E and 6B). The results show that NSC351204 can exhibit 85% detectable inhibitory effect against hGGPPS (WT) but has low inhibition against yGGPPS and K212A when used at a concentration of  $100 \mu\text{M}$ . However, NSC45174 can efficiently inhibit both yGGPPS and K212A (Figures 6E and 6F). These results lend support to our hypotheses and suggest that K202 and K212 in hGGPPS are key residues for hGGPPS-specific inhibition.

**Tumor Cell Growth Inhibition by Nonbisphosphonate hGGPPS Inhibitors.** The potential of bisphosphonates as

antitumor agents has been suggested by several *in vitro* and *in vivo* preclinical studies.<sup>49</sup> For example, BPH-675 has good activity against GGPPS and three tumor cell lines: MCF-7 (breast cancer), NCI-H460 (nonsmall cell lung cancer), and SF-268 (human glioblastoma).<sup>20</sup> The zoledronate (+endocrine) therapy was found to significantly reduce disease progression in premenopausal breast cancer patients.<sup>50</sup> We tested the activity of NSC351204 (Figure 5A)—our nonbisphosphonate inhibitor with the best inhibitory effect for hGGPPS—in killing tumor cells, using MCF-7 (breast cancer) and MDA-MB-231 (breast adenocarcinoma) cell lines. The results of cell growth inhibition clearly show that NSC351204 induced tumor cell death though its EC<sub>50</sub> values in these two breast cell lines was not more potent than those of zoledronate and pamidronate: in MCF-7 cells, the EC<sub>50</sub> of NSC351204 is ~400  $\mu$ M, and the EC<sub>50</sub> of zoledronate and pamidronate are ~24  $\mu$ M and ~151  $\mu$ M, respectively; in MDA-MB-231 cells, the EC<sub>50</sub> of NSC351204 is ~300  $\mu$ M, and the EC<sub>50</sub> of zoledronate and pamidronate are ~20  $\mu$ M and ~100  $\mu$ M, respectively (Figures 7A and 7B). Additionally, we performed the rescue assay<sup>20</sup> to verify whether our nonbisphosphonate inhibitor can target GGPPS to affect the geranylgeranylation of downstream proteins (i.e., Ras and RhoA)<sup>51</sup> using cell-based assay. There was almost rescue (>90%) from growth inhibition by NSC351204 in MCF-7 cell by addition of 20  $\mu$ M GGOH (Figure 7C) but not by FOH. This result suggests that NSC351204 is a selective GGPPS inhibitor, which is consistent with our *in vitro* data (Figure 5D). According to our model, in anchor V3 (Figure 1E and Table 2), the sulfate and nitro moieties in NSC351204 might be the primary cause of the reduced potency in the enzyme-activity and cell-based assays. Although pharmaceutical properties of NSC351204 like cellular permeability remain to be improved, information on its new chemotype, *in vitro* data, and lack of bone affinity indicates the high potential of nonbisphosphonate GGPPS inhibitors for binding at the inhibitory pocket. Combining a site-moiety map with CSRs in the context of a moiety-linkage-based virtual screening provided a novel advance for the discovery of nonbisphosphonate hGGPPS inhibitors to be used as antitumor agents.

**The Issues Regarding to Moiety-Linkage Map Method.** The anchors of the site-moiety map often contain conserved interactions between key pockets and their preferred moieties. A compound that agrees with the anchors is often able to activate or inhibit the target protein. Therefore, a site-moiety map can refine known inhibitors of hGGPPS by pharmacophore score based on the moiety preferences of each anchor (Figure 1F). For example, for the anchor E1, new scaffold moieties (i.e., sulfate, carboxylate, etc.) might represent a new chemical class of GGPPS inhibitors for guiding the optimization of known inhibitors. In addition, the performance of a scoring function for structure-based virtual screening is often limited by inadequate prediction of the true binding affinity of a small compound at a binding site. For example, the energy-based scoring methods would be biased toward high molecular weight and charged polar compounds.<sup>52,53</sup> Our P+R score function is useful in reducing these ill-effects of energy-based and pharmacophore-scoring methods to help recognize positive hits from thousands of screening compounds (Figures 4C and 4D).

In addition, compound **50**, which fit CSRs 1, 2, and 4 (P+R score: 24.5), showed a good inhibitory effect (IC<sub>50</sub> = 0.66  $\mu$ M) against hGGPPS despite not meeting CSR 3. One possible

reason is that compound **50** possesses a long C–C chain moiety, which forms hydrophobic interactions with conserved residues R28, L31, F156, and F184 of anchors V2 and V3. Moreover, compound **51** only matches CSRs 1 and 3 (P+R score: 17.5), but it is nonetheless a potent hGGPPS inhibitor (IC<sub>50</sub> = 4.07  $\mu$ M). The hydrophobic moieties (the second and third rings) of compound **51** occupy the anchors V2 and V3 well due to two single bonds on the second ring forming a ~120° angle by connecting neighboring moieties via the *meta*-position (Supplementary Figure 3A).

**The Potential of Nonbisphosphonate Inhibitors in Disease Therapies.** The nonbisphosphonate inhibitors of GGPPS may not be associated with high affinity to bone minerals. GGPPS and FPPS are essential enzymes in the mevalonate pathway and are downstream of HMG-CoA reductase. This result suggests that nonbisphosphonate GGPPS inhibitors could lower cholesterol in conjunction with or as an alternative to statins and FPPS inhibitors. The mevalonate pathway has also been implicated in other diseases such as Alzheimer's disease and has been explored in some clinical trials to elucidate the benefits of statins.<sup>54</sup> Bisphosphonates have been shown to be effective against Chagas disease, malaria, cryptosporidiosis, and leishmaniasis resulting from parasitic protozoa,<sup>55</sup> but nonbisphosphonate GGPPS inhibitors may be better suited for such applications, owing to their lack of affinity for bone tissue.

## CONCLUSIONS

To search for nonbisphosphonate hGGPPS inhibitors, we screened large libraries by using GEMDOCK. Two nonbisphosphonate inhibitors, NSC351204 and NSC45174, were selected based on the pharmacophore- and CSR-strategy (P+R) from top-ranking candidates (Figure 5A and Table 2). Though their inhibitory effects are not potent ( $\leq 10$   $\mu$ M) due to a lack of strong electrostatic forces and not meeting CSR 4, they are the beginning of a new scaffold of GGPPS inhibitors and drug leads for bone and related mevalonate-pathway diseases. It is feasible that the potency of these two compounds could be improved to make novel candidates for hGGPPS inhibitors by following the P+R strategy. Finally, we believe that combining a site-moiety map with CSRs is a useful general framework for identifying small-molecule inhibitors that can regulate activity of target proteins.

## ASSOCIATED CONTENT

### Supporting Information

Table S1, Scheme 1, and Supplemental Figures S1–S6. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone: 886-3-5712121 ext.56942. Fax: 886-3-5729288. E-mail: [moon@faculty.nctu.edu.tw](mailto:moon@faculty.nctu.edu.tw) (J.M.Y.). Phone: 886-2-33664069. Fax: 886-2-23635038. E-mail: [phliang@gate.sinica.edu.tw](mailto:phliang@gate.sinica.edu.tw) (P.H.L.).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors are grateful to the Core Facility for Protein Structural Analysis supported by the National Core Facility



Program for Biotechnology. Dr. Yang was supported by the National Science Council and partially supported by the Ministry of Education and National Health Research Institutes (NHRI-EX100-10009PI).

## REFERENCES

- (1) Sacchettini, J. C.; Poulter, C. D. Creating isoprenoid diversity. *Science* **1997**, *277* (5333), 1788–9.
- (2) Ogura, K.; Koyama, T. Enzymatic aspects of isoprenoid chain elongation. *Chem. Rev.* **1998**, *98* (4), 1263–1276.
- (3) Liang, P. H.; Ko, T. P.; Wang, A. H. Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* **2002**, *269* (14), 3339–54.
- (4) Poulter, C. D.; Rilling, H. C. Prenyltransferase: the mechanism of the reaction. *Biochemistry* **1976**, *15* (5), 1079–83.
- (5) Hosfield, D. J.; Zhang, Y.; Dougan, D. R.; Broun, A.; Tari, L. W.; Swanson, R. V.; Finn, J. Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. *J. Biol. Chem.* **2004**, *279* (10), 8526–9.
- (6) Sun, H. Y.; Ko, T. P.; Kuo, C. J.; Guo, R. T.; Chou, C. C.; Liang, P. H.; Wang, A. H. Homodimeric hexaprenyl pyrophosphate synthase from the thermoacidophilic crenarchaeon *Sulfolobus solfataricus* displays asymmetric subunit structures. *J. Bacteriol.* **2005**, *187* (23), 8137–48.
- (7) Gabelli, S. B.; McLellan, J. S.; Montalvetti, A.; Oldfield, E.; Docampo, R.; Amzel, L. M. Structure and mechanism of the farnesyl diphosphate synthase from *Trypanosoma cruzi*: implications for drug design. *Proteins* **2006**, *62* (1), 80–8.
- (8) Chang, T. H.; Guo, R. T.; Ko, T. P.; Wang, A. H.; Liang, P. H. Crystal structure of type-III geranylgeranyl pyrophosphate synthase from *Saccharomyces cerevisiae* and the mechanism of product chain length determination. *J. Biol. Chem.* **2006**, *281* (21), 14991–5000.
- (9) Kavanagh, K. L.; Dunford, J. E.; Bunkoczi, G.; Russell, R. G.; Oppermann, U. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J. Biol. Chem.* **2006**, *281* (31), 22004–12.
- (10) Kloer, D. P.; Welsch, R.; Beyer, P.; Schulz, G. E. Structure and reaction geometry of geranylgeranyl diphosphate synthase from *Sinapis alba*. *Biochemistry* **2006**, *45* (51), 15197–204.
- (11) Rudiger, W.; Benz, J.; Guthoff, C. Detection and partial characterization of activity of chlorophyll synthetase in etioplast membranes. *Eur. J. Biochem.* **1980**, *109* (1), 193–200.
- (12) Jiang, Y.; Proteau, P.; Poulter, D.; Ferro-Novick, S. BTS1 encodes a geranylgeranyl diphosphate synthase in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **1995**, *270* (37), 21793–9.
- (13) Olszewski, N.; Sun, T. P.; Gubler, F. Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell* **2002**, *14* (Suppl), S61–80.
- (14) Goffinet, M.; Thoulouzan, M.; Pradines, A.; Lajoie-Mazenc, I.; Weinbaum, C.; Faye, J. C.; Seronie-Vivien, S. Zoledronic acid treatment impairs protein geranyl-geranylation for biological effects in prostatic cells. *BMC Cancer* **2006**, *6*, 60.
- (15) Guo, R. T.; Cao, R.; Liang, P. H.; Ko, T. P.; Chang, T. H.; Hudock, M. P.; Jeng, W. Y.; Chen, C. K.; Zhang, Y.; Song, Y.; Kuo, C. J.; Yin, F.; Oldfield, E.; Wang, A. H. Bisphosphonates target multiple sites in both cis- and trans-prenyltransferases. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (24), 10022–7.
- (16) Dunford, J. E.; Thompson, K.; Coxon, F. P.; Luckman, S. P.; Hahn, F. M.; Poulter, C. D.; Ebetino, F. H.; Rogers, M. J. Structure-activity relationships for inhibition of farnesyl diphosphate synthase in vitro and inhibition of bone resorption in vivo by nitrogen-containing bisphosphonates. *J. Pharmacol. Exp. Ther.* **2001**, *296* (2), 235–42.
- (17) Szabo, C. M.; Matsumura, Y.; Fukura, S.; Martin, M. B.; Sanders, J. M.; Sengupta, S.; Cieslak, J. A.; Loftus, T. C.; Lea, C. R.; Lee, H. J.; Koohang, A.; Coates, R. M.; Sagami, H.; Oldfield, E. Inhibition of geranylgeranyl diphosphate synthase by bisphosphonates and diphosphates: a potential route to new bone antiresorption and antiparasitic agents. *J. Med. Chem.* **2002**, *45* (11), 2185–96.
- (18) Jahnke, W.; Rondeau, J. M.; Cotesta, S.; Marzinzik, A.; Pelle, X.; Geiser, M.; Strauss, A.; Gotte, M.; Bitsch, F.; Hemmig, R.; Henry, C.; Lehmann, S.; Glickman, J. F.; Roddy, T. P.; Stout, S. J.; Green, J. R. Allosteric non-bisphosphonate FPPS inhibitors identified by fragment-based discovery. *Nat. Chem. Biol.* **2010**, *6* (9), 660–6.
- (19) Simoni, D.; Gebbia, N.; Invidiata, F. P.; Eleopra, M.; Marchetti, P.; Rondanin, R.; Baruchello, R.; Provera, S.; Marchioro, C.; Tolomeo, M.; Marinelli, L.; Limongelli, V.; Novellino, E.; Kwaasi, A.; Dunford, J.; Buccheri, S.; Caccamo, N.; Dieli, F. Design, synthesis, and biological evaluation of novel aminobisphosphonates possessing an in vivo antitumor activity through a gamma delta-T lymphocytes-mediated activation mechanism. *J. Med. Chem.* **2008**, *51* (21), 6800–7.
- (20) Zhang, Y.; Cao, R.; Yin, F.; Hudock, M. P.; Guo, R. T.; Krysiak, K.; Mukherjee, S.; Gao, Y. G.; Robinson, H.; Song, Y.; No, J. H.; Bergan, K.; Leon, A.; Cass, L.; Goddard, A.; Chang, T. K.; Lin, F. Y.; Van Beek, E.; Papapoulos, S.; Wang, A. H.; Kubo, T.; Ochi, M.; Mukkamala, D.; Oldfield, E. Lipophilic bisphosphonates as dual farnesyl/geranylgeranyl diphosphate synthase inhibitors: an X-ray and NMR investigation. *J. Am. Chem. Soc.* **2009**, *131* (14), 5153–62.
- (21) Chen, Y. F.; Hsu, K. C.; Lin, S. R.; Wang, W. C.; Huang, Y. C.; Yang, J. M. SIMMap: a web server for inferring site-moiety map to recognize interaction preferences between protein pockets and compound moieties. *Nucleic Acids Res.* **2010**, *38* (Web Server issue), W424–30.
- (22) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (23) Hsu, K. C.; Chen, Y. F.; Lin, S. R.; Yang, J. M. iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. *BMC Bioinf.* **2011**, *12* (Suppl 1), S33.
- (24) Yang, J. M.; Chen, C. C. GEMDOCK: a generic evolutionary method for molecular docking. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 288–304.
- (25) Wang, R.; Lu, Y.; Wang, S. Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem.* **2003**, *46* (12), 2287–303.
- (26) Thomsen, R.; Christensen, M. H. MolDock: a new technique for high-accuracy molecular docking. *J. Med. Chem.* **2006**, *49* (11), 3315–21.
- (27) Yang, J. M.; Chen, Y. F.; Shen, T. W.; Kristal, B. S.; Hsu, D. F. Consensus scoring criteria for improving enrichment in virtual screening. *J. Chem. Inf. Model.* **2005**, *45* (4), 1134–46.
- (28) Ewing, T. J.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411–428.
- (29) Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins* **1999**, *37*, 228–241.
- (30) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (31) Yang, J. M.; Shen, T. W. A pharmacophore-based evolutionary approach for screening selective estrogen receptor modulators. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 205–220.
- (32) Yang, J. M.; Chen, Y. F.; Tu, Y. Y.; Yen, K. R.; Yang, Y. L. Combinatorial computational approaches to identify tetracycline derivatives as flavivirus inhibitors. *PLoS One* **2007**, *2* (5), e428.
- (33) Chin, K. H.; Lee, Y. C.; Tu, Z. L.; Chen, C. H.; Tseng, Y. H.; Yang, J. M.; Ryan, R. P.; McCarthy, Y.; Dow, J. M.; Wang, A. H.; Chou, S. H. The cAMP receptor-like protein CLP is a novel c-di-GMP receptor linking cell-cell signaling to virulence gene expression in *Xanthomonas campestris*. *J. Mol. Biol.* **2010**, *396* (3), 646–62.
- (34) Hung, H. C.; Tseng, C. P.; Yang, J. M.; Ju, Y. W.; Tseng, S. N.; Chen, Y. F.; Chao, Y. S.; Hsieh, H. P.; Shih, S. R.; Hsu, J. T. Aurintricarboxylic acid inhibits influenza virus neuraminidase. *Antiviral Res.* **2009**, *81* (2), 123–31.
- (35) Yang, M.-C.; Guan, H.-H.; Yang, J.-M.; Ko, C.-N.; Liu, M.-Y.; Lin, Y.-H.; Chen, C.-J.; Mao, S. J. T. Rational design for crystallization

of beta-lactoglobulin and vitamin D-3 complex: revealing a secondary binding site. *Cryst. Growth Des.* **2008**, *8*, 4268–4276.

(36) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.

(37) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 293–304.

(38) Chen, Y. F.; Hsu, K. C.; Lin, S. R.; Wang, W. C.; Huang, Y. C.; Yang, J. M. SiMMap: a web server for inferring site-moiety map to recognize interaction preferences between protein pockets and compound moieties. *Nucleic Acids Res.* **2010**, *38*, W424–W430.

(39) Yang, L.; Wang, K. J.; Chen, J. A.; Jegga, A. G.; Luo, H.; Shi, L. M.; Wan, C. L.; Guo, X. Z.; Qin, S. Y.; He, G. A.; Feng, G. Y.; He, L. Exploring off-targets and off-systems for adverse drug reactions via chemical-protein interactome - clozapine-induced agranulocytosis as a case study. *PLoS Comput. Biol.* **2011**, *7*, e1002016.

(40) Sawatdichaikul, O.; Hannongbua, S.; Sangma, C.; Wolschann, P.; Choowongkomon, K. In silico screening of epidermal growth factor receptor (EGFR) in the tyrosine kinase domain through a medicinal plant compound database. *J. Mol. Model.* **2012**, *18*, 1241–1254.

(41) Hsu, K. C.; Cheng, W. C.; Chen, Y. F.; Wang, H. J.; Li, L. T.; Wang, W. C.; Yang, J. M. Core site-moiety maps reveal inhibitors and binding mechanisms of orthologous proteins by screening compound libraries. *PLoS One* **2012**, *7*, e32142.

(42) Guo, R. T.; Kuo, C. J.; Chou, C. C.; Ko, T. P.; Shr, H. L.; Liang, P. H.; Wang, A. H. Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic *Thermotoga maritima* and mechanism of product chain length determination. *J. Biol. Chem.* **2004**, *279* (6), 4903–12.

(43) Pan, J. J.; Chiou, S. T.; Liang, P. H. Product distribution and pre-steady-state kinetic analysis of *Escherichia coli* undecaprenyl pyrophosphate synthase reaction. *Biochemistry* **2000**, *39* (35), 10936–42.

(44) Kavanagh, K. L.; Guo, K.; Dunford, J. E.; Wu, X.; Knapp, S.; Ebetino, F. H.; Rogers, M. J.; Russell, R. G.; Oppermann, U. The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (20), 7829–34.

(45) C, K. M. C.; Hudock, M. P.; Zhang, Y.; Guo, R. T.; Cao, R.; No, J. H.; Liang, P. H.; Ko, T. P.; Chang, T. H.; Chang, S. C.; Song, Y.; Axelson, J.; Kumar, A.; Wang, A. H.; Oldfield, E. Inhibition of geranylgeranyl diphosphate synthase by bisphosphonates: a crystallographic and computational investigation. *J. Med. Chem.* **2008**, *51* (18), 5594–607.

(46) Chang, K. M.; Chen, S. H.; Kuo, C. J.; Chang, C. K.; Guo, R. T.; Yang, J. M.; Liang, P. H. Roles of amino acids in the *Escherichia coli* octaprenyl diphosphate synthase active site probed by structure-guided site-directed mutagenesis. *Biochemistry* **2012**, *51* (16), 3412–9.

(47) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215* (3), 403–10.

(48) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–42.

(49) Stresing, V.; Daubine, F.; Benzaid, I.; Monkkonen, H.; Clezardin, P. Bisphosphonates in cancer therapy. *Cancer Lett* **2007**, *257* (1), 16–35.

(50) Gnant, M.; Mlineritsch, B.; Schippinger, W.; Luschin-Ebengreuth, G.; Postlberger, S.; Menzel, C.; Jakesz, R.; Seifert, M.; Hubalek, M.; Bjelic-Radisic, V.; Samonigg, H.; Tausch, C.; Eidtmann, H.; Steger, G.; Kwasny, W.; Dubsky, P.; Fridrik, M.; Fitzal, F.; Stierer, M.; Rucklinger, E.; Greil, R.; Marth, C. Endocrine therapy plus zoledronic acid in premenopausal breast cancer. *N. Engl. J. Med.* **2009**, *360* (7), 679–91.

(51) Houten, S. M.; Schneiders, M. S.; Wanders, R. J.; Waterham, H. R. Regulation of isoprenoid/cholesterol biosynthesis in cells from mevalonate kinase-deficient patients. *J. Biol. Chem.* **2003**, *278* (8), 5736–43.

(52) Yang, J. M.; Shen, T. W. A pharmacophore-based evolutionary approach for screening selective estrogen receptor modulators. *Proteins* **2005**, *59* (2), 205–20.

(53) Pan, Y.; Huang, N.; Cho, S.; MacKerell, A. D., Jr. Consideration of molecular weight during compound selection in virtual target-based database screening. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (1), 267–72.

(54) Davignon, J.; Leiter, L. A. Ongoing clinical trials of the pleiotropic effects of statins. *Vasc Health Risk Manage.* **2005**, *1* (1), 29–40.

(55) Docampo, R.; Moreno, S. N. Bisphosphonates as chemotherapeutic agents against trypanosomatid and apicomplexan parasites. *Curr. Drug Targets: Infect. Disord.* **2001**, *1* (1), 51–61.