Published in final edited form as:
J Med Chem. 2008 September 25; 51(18): 5594–5607. doi:10.1021/jm800325y.

Inhibition of Geranylgeranyl Diphosphate Synthase by Bisphosphonates: A Crystallographic and Computational Investigation

Cammy K.-M. Chen[†,‡,°], **Michael P. Hudock**[||,°], **Yonghui Zhang**[¶,°], **Rey-Ting Guo**[‡,§], **Rong Cao**[||], **Joo Hwan No**[||], **Po-Huang Liang**[†,‡], **Tzu-Ping Ko**[‡], **Tao-Hsin Chang**[‡,§], **Shiou-chi Chang**[¶], **Yongcheng Song**[¶], **Jordan Axelson**[¶], **Anup Kumar**[¶], **Andrew H.-J. Wang**[†,‡,§,*], and **Eric Oldfield**[||,¶,*]

† *Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan*

‡ *Institute of Biological Chemistry, Academia Sinica, Taipei 106, Taiwan*

|| *Center for Biophysics and Computational Biology, 607 South Mathews Avenue, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

§ *Core Facility for Protein Crystallography, Academia Sinica, Taipei 115, Taiwan*

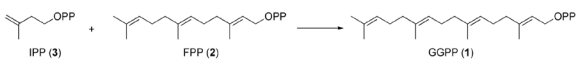
¶ *Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801*

Abstract

We report the x-ray structures of several bisphosphonate inhibitors of geranylgeranyl diphosphate synthase, a target for anti-cancer drugs. Bisphosphonates containing unbranched sidechains bind to either the farnesyl diphosphate (FPP) substrate site, the geranylgeranyl diphosphate (GGPP) product site, and in one case, both sites, with the bisphosphonate moiety interacting with 3 Mg^{2+} that occupy the same position as found in FPP synthase. However, each of three “V-shaped” bisphosphonates binds to both the FPP and GGPP sites. Using the Glide program, we reproduced the binding modes of 10 bisphosphonates with an RMS error of 1.3Å. Activities of the bisphosphonates in GGPPS inhibition were predicted with an overall error of 2x, using a comparative molecular similarity analysis, based on a docked-structure alignment. These results show that some GGPPS inhibitors can occupy both substrate and product site, and that binding modes as well as activity can be accurately predicted, facilitating the further development of GGPPS inhibitors as anti-cancer agents.

Introduction

Geranylgeranyl diphosphate synthase (GGPPS, EC 2.5.1.30) catalyzes the formation of geranylgeranyl diphosphate (**1**) from one molecule of farnesyl diphosphate (**2**) and one molecule of isopentenyl diphosphate (**3**)[1]:



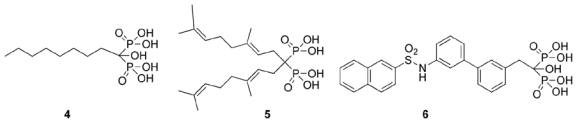
*Correspondence should be addressed to A.H.-J. W. ahjwang@gate.sinica.edu.tw, Tel +886-2-2788-1981, Fax +886-2-2788-2043 or E.O. eo@chad.scs.uiuc.edu, Tel 217-333-3374, Fax 217-244-0997.

° These authors contributed equally to this work.

Crystal structure coordinates have been deposited in the Protein Databank and will be released upon publication (2z4w, 2z4z, 2z78, 2z4x, 2z4y, 2z50, 2z52, 2z4v, 2z7I).

The GGPP product is used in the biosynthesis of many natural products, such as taxanes and gibberellins, and is also used to prenylate proteins such as Rho, Rap and Rac, involved in cell signaling pathways[2, 3], Figure 1. It can be further elongated by some polyprenyl synthases[4] to produce the long chain isoprenoids used in quinone biosynthesis, and in plants and some bacteria, two GGPP molecules can condense to form phytoene, the precursor for many carotenoids[5]. GGPPS is inhibited by a variety of bisphosphonates[6–9], and is of current interest in the context of the development of anti-cancer drugs[7, 8] which function by inhibiting protein prenylation, cell signaling and cell survival pathways, Figure 1.

In earlier work[6], we found that n-alkyl bisphosphonates such as **4**:

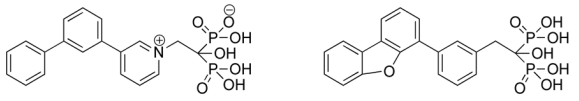


4 **5** **6**

had quite potent activity against GGPPS, and more recently, Weimer et al. have reported[7, 8] that novel diprenyl methylenebisphosphonates, such as digeranyl methylene bisphosphonate (**5**), have potent activity against GGPPS, as well as against a K562 tumor cell line, but the structures of neither the n-alkyl nor any dialkenyl bisphosphonate inhibitor-GGPPS complexes have been reported. The structure of human GGPPS is now known, however, with in recent work, Kavanagh et al.[10] finding the presence of the “isoprene fold” found in other prenyl synthases, such as farnesyl diphosphate synthase (FPPS, EC 2.5.1.10)[11–14]. These workers also showed[10] that the GGPP product bound to a central “inhibitor” binding site, and in more recent work[9], we have found that other GGPPS inhibitors such as **6** (which is too large to inhibit FPPS) also bind to this site, and are potent inhibitors of GGPPS activity. We also found that GGPPS substrates and diphosphate and bisphosphonate inhibitors can bind in four distinct ways to GGPPS, with their polar (diphosphate, bisphosphonate) groups binding to either the FPP or IPP diphosphate binding sites, and their more hydrophobic fragments binding to the (human) GGPP (inhibitor) site, or to the FPP (substrate) site[9]. Here, we report the first structures of a series of n-alkyl and dialkenyl bisphosphonates bound to GGPPS. We also show that the binding modes seen crystallographically can be well predicted computationally, facilitating the development of quantitative structure-activity models. Given the widespread use of bisphosphonates in treating bone resorption diseases and the current interest in them as anti-cancer agents[15–17], these results are of broad general interest since they lay the foundation for the further development of, in particular, the novel disubstituted bisphosphonates.

Results and Discussion

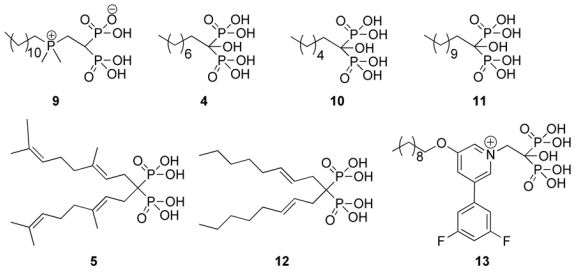
GGPPS is a highly α-helical protein and diphosphates (IPP, FPP and GGPP) as well as bisphosphonates such as **7**, **8** have previously been shown[9] to bind to GGPPS in four distinct ways.



7 **8**

In the first, the polar (diphosphate or bisphosphonate) groups (of e.g. **7**) bind in the F***PP*** substrate or GG***PP*** product binding site, with the large hydrophobic sidechain occupying the ***GG***PP sidechain site first reported by Kavanagh et al.,[10] Figure S1A. In the second mode, the F***PP*** substrate site is occupied (by e.g. FPP, zoledronate, minodronate or **8**) with the long sidechain (when present) occupying the ***F***PP sidechain site, Figure S1B. In the third binding

mode, seen only so far with **8**, the I***PP*** polar site is occupied by the polar bisphosphonate, and the large hydrophobic sidechain site resides in the hydrophobic **GG**PP (human) site, Figure S1C. And in the fourth binding mode, seen so far only with GGPP in the *Saccharomyces cerevisae* (yeast) GGPPS structure, the polar diphosphate of GGPP occupies the polar I***PP*** site, while the hydrophobic sidechain occupies the ***F***PP (substrate) hydrophobic sidechain site, Figure S1D. These observations clearly pose a challenge for structure based inhibitor design, since it is by no means clear how a given, novel inhibitor, might bind. We elected, therefore, to investigate the structures of the following series of inhibitors:



which, when combined with previous structures,[9] might provide a data base with which to test the results of computational docking calculations. If successful, it might then be possible to use docked structures of a much larger range of inhibitors (whose crystallographic structures are not known), to develop quantitative structure activity relationships using, in particular, the comparative molecular similarity index analysis (CoMSIA) method[18]. For the simpler bisphosphonates, it seemed possible that several of the four binding poses seen previously (Figure S1) might be possible, while for the branched chain species (**5,12,13**), it seemed likely that more than one site might be needed, in order to accommodate these "V-shaped" molecules.

To investigate the numerous possibilities for n-alkyl bisphosphonate binding, we first determined the structure of **4** bound to GGPPS, obtaining two structures. In the first, we find that **4** (yellow) binds to the FPP site (F***PP***-***F***PP)(Figures 2A,B; PDB File 2z4x). The bisphosphonate fragment coordinates to three Mg^{2+} (yellow), Figure 2B, which in turn are coordinated to the DDXXD repeats (Figure 2A). The C8 side-chain is closely aligned to the *thiolo*-farnesyl diphosphate (FsPP) side-chain seen in the FsPP-IPP-GGPPS structure (PDB File 2e8t) and, as can be seen in the superposition shown in Figure 2C, the three Mg^{2+} are closely aligned to the three Mg^{2+} seen in many FPPS structures[11–14, 19], with a ~0.32 Å RMSD between the Mg^{2+} positions in the GGPPS (PDB File 2z4x) and FPPS/IPP/zoledronate (PDB File 2f8z) structures. Essentially the same binding pattern is seen with the longer chain species **11**, Figure S2. These structures present the first observation of 3 Mg^{2+} bound to GGPPS and, as can be seen in Figure 2C, the similarity in metal binding in GGPPS and FPPS is clear. So, at least for medium-length side-chains (about the length of GPP), the FPP-binding site is an important target for bisphosphonates, as is binding to Mg^{2+}. Surprisingly, however, in a second structure of **4**, we find evidence for *two* binding sites for **4**, in chain A (1 in chain B). In this structure (PDB File 2z4y, chain 1A), Figure 2D, one inhibitor molecule binds to the FPP site while a second binds to the I***PP***-***GG***PP product/inhibitor site, the same binding pattern as seen previously with **8**[9]. There is, therefore, some variability in how bisphosphonates bind to GGPPS and indeed, this variability is also seen with binding of the GGPP product where in earlier work[9], we found GGPP bound to the I***PP***/***F***PP site (Figure S1D), in contrast to the F***PP***/***GG***PP binding mode seen with the human enzyme[10]. However, in a new structure (PBD 2z4v) we find the same binding pattern as seen in human GGPPS (Figure 2E, pink), due presumably to the slight changes in crystallization conditions (higher [Mg^{2+}]) employed in the current investigation. Ligand interaction diagrams for all of the new structures are shown in Figures S3. We find that that bisphosphonate groups of both **4** ($P2_1$) and **11** coordinate to three

Mg^{2+}, which in turn are coordinated to the DDXXD repeats. The bisphosphonate backbones form H bonds with Arg84, Lys169 and Lys233. In a second structure of **4**, **4** can bind to two sites; one in the FPP site, interacting with the side chains of Arg84, Lys169, Gln206 and Lys233, the other one in the GGPP product site, interacting with Arg39, His68, Arg84 and Arg85. In the case of **10**, there is no Mg^{2+} observable, and the bisphosphonate backbones interact with the side chains of Asp75, Arg84 and Lys233 via hydrogen bonds.

The wide range of binding modes observed experimentally suggests the possibility that two sites might be occupied by the 1,1-disubstituted bisphosphonates, such as **5**,**12**, which could bind in a "V-shaped" conformation that could easily occupy both the F***PP***-***F***PP and F***PP***-***GG***PP sites (Figure S1A,B) or the I***PP***-***GG***PP and I***PP***-***F***PP sites (Figure S1C,D). The former turns out to be the case as can be seen in Figure 3A where we see the location of the two geranyl chains in **5** (plus Mg^{2+}) bound to the GGPPS dimer, illustrating binding to the FPP and GGPP sites. Close-up views of **5,12** and **13** (PDB Files 2z4w, 2z4z and 2z78) superimposed on the FsPP and GGPP ligands in GGPPS (PDB Files 2e8t, 2z4v) are shown in Figures 3B-D (ligand interaction plots are provided in Figures S4A-C) and clearly indicate that each of these "V-shaped" molecules bind to both the FPP and GGPP (substrate, inhibitor) sites. A major difference between the three structures is, however, that the number of Mg^{2+} varies, (Figures 3B-D). The 2 Mg^{2+} (green) that are seen in the **5** structure occupy the same positions as those seen in the 3 Mg^{2+} structures (Figures 2B,C), but there are no Mg^{2+} seen in the two other structures, which might, however, simply reflect the relatively low pH (~4.8) required for crystallization, making Mg^{2+}-binding relatively weak. Data collection and refinement statistics for **5, 12** and **13** bound to GGPPS are shown in Table 4.

In addition to these structures, we also determined the structures of a shorter (C_6) n-alkyl inhibitor (**10**) as well as that of a potent long chain phosphonium inhibitor (**9**), bound to GGPPS, Figure S5A,B. Data collection are refinement statistics are shown in Table S1 in the Supporting Information. We find that the long chain phosphonium bisphosphonate binds to the GGPP inhibitor site while the short chain n-alkyl species binds to the FPP site, as illustrated in Figures S5A,B. Electron densities for all structures reported here are shown in Figure S6.

These results are clearly of interest in the context of deducing structure-activity relationships for GGPPS inhibition, not least since there are clearly several different types of binding that might be anticipated for a given compound, which would be expected to complicate the use of purely ligand-based predictions of activity. We thus next consider how these crystallographic results might be used to guide QSAR investigations.

Bisphosphonate Activities and QSAR

We investigated the activities in GGPPS inhibition of the 60 bisphosphonates (plus 1 phosphosulfonate, **61**) shown in Figure 4. The most potent inhibitors (Table 1) all contained long alkyl chains, with short chain species – conventional bisphosphonates such as pamidronate (**62**), risedronate (**64**) and alendronate (**65**) having essentially no activity (> 100 μM). Two phosphonate groups were essential for activity (e.g. **8**, IC_{50} = 4 μM; the phosphonosulfonate analog **61** had no detectable activity), but unlike FPPS inhibition, a cationic center was not essential for good activity, e.g. **15**, IC_{50} = 0.3 μM; **19**, IC_{50} = 0.6 μM. Likewise, the neutral sidechain species **5** had an IC_{50} = 1 μM. So, the cationic species are not acting as carbocationic transition state/reactive intermediate analogs, as they do in FPPS inhibition[20].

In order to interpret the activity results (Table 1), it seemed likely that it would be necessary to deduce the most probable binding site for each inhibitor, since the activity of a given compound would be expected to depend on which site it occupied. To do this, we first investigated the docking poses (using Glide[21]) of 12 inhibitors whose crystallographic structures were known, in order to validate this approach. Using a superposition of 12

bisphosphonate-GGPPS complexes (Figure 5A), we selected five different structures to serve as target receptors. They contained zero, one, two (2 structures) or all three Mg^{2+} ions, and had a variety of protein sidechain (Leu138, Arg39 and Arg85) conformations (Figure 5B and Table S2). Of the five structures, the GGPPS containing compound **8** yielded the best docking results (Table S2), most likely because it had two large bound ligands, with the protein (PDB File: 2e93) sidechain orientations (Figure 5B, cyan) permitting binding to many different inhibitors.

Twelve inhibitors whose active site conformations were known crystallographically were then docked into each target receptor in a cross docking approach, specifying no constraints, using Glide[21]. The results obtained were then compared with the known crystallographic results, with the GGPPS **8** "Mg^{2+}_{BC}" structure (Figure 5B, PDB File: 2e93) providing the best overall results (Table S2) Using this structure we found, on average, a 1.3 Å RMSD between the x-ray and docked poses, with the correct binding site predicted in all cases. We observed slightly worse results for the other receptor containing 2 Mg^{2+}, and very poor results for each of the other receptors (full details shown in Supporting Tables S2 and S3). Interestingly, the GGPPS **8** structure (which gave the best docking results) contains two bulky bisphosphonates, and can presumably accommodate ligands having diverse structures.

We next docked 51 additional GGPPS inhibitors, having a large range of activity (~0.1–150 μM, Figure 5 and Table 1) to the GGPPS **8** structure. A surface representation of all x-ray and predicted binding poses is presented in Figure 6A, and clearly shows the presence of two different binding sites. In all cases, the lowest energy poses of each inhibitor were those that might be deduced based on chemical similarity to known inhibitors. However, despite the good predicted binding poses, the docking score (G Score in Glide) was only weakly correlated with GGPPS inhibition ($r^2 = 0.3$, Figure 6B). A possible explanation for this is that although the calculations were performed using the highest level of accuracy available in Glide 4.5[21] (extra-precision mode), the receptor is still treated rigidly,[22] so the accuracy of the scoring function might be compromised. To improve the docking score, we next investigated the use of the linear interaction approximation (LIE)[23] method in the Liaison[24] program. This performs molecular mechanics simulations on the bound and free states of the enzyme-ligand complex using a continuum-solvent model,[24] and takes into account receptor flexibility. However, improvements were small: r^2=0.4, Figure 6C. Finally, we constructed a re-parameterized energy function, based on the experimentally determined activities, computed energies and the lipophilicity molecular descriptor, SlogP[25]. This resulted in a further improvement with experiment: $r^2 = 0.6$, $q^2 = 0.6$, F = 85, Figure 6D, Table 1, and using a training and test set approach, we found that IC_{50} values could be predicted within a factor of 3x (over a range of 4000x in activity). Full output results are shown in Table S4. So, while promising, these results are clearly worse than those we have obtained previously using CoMSIA or CoMFA (comparative molecular field analysis) methods for bisphosphonate inhibition of a variety of enzymes in which only a single site is occupied[16, 26, 27].

We thus next investigated the use of CoMSIA methods to predict activity. Structurally similar compounds were aligned to their closest x-ray structure, followed by a flexible common feature superposition, as described previously[16]. The alignment and fields, Figure 7A-C, resulted in a $q^2 = 0.59$, $r^2 = 0.8$, F = 73, n = 61, Figure 7D. Full results are shown in Tables S5 and S6. Although these results are quite promising, it seemed logical to see to what extent they might be improved upon by using a receptor-guided alignment[28] based on the Glide docking results discussed above. The CoMSIA fields so obtained are shown in Figure 7E-G and using this receptor-guided alignment found that all three parameters, q^2, r^2 and F, improved. The q^2 parameter increased from $q^2 = 0.6$ to $q^2 = 0.7$; r^2 from 0.8 to 0.9, and F from 73 to 166, Figure 7H, with predictions now within a factor of 2x (over a 4000x range in activity). Full results are shown in Tables 2 and S7. Also of importance, the receptor-guided alignment could be

generated rapidly, and in a uniform manner, for all compounds. Despite the slight differences in the predictivity of the models obtained from the two alignments, the CoMSIA fields produced are very similar, with a steric-favorable region (green) coincident with the end of the ***F***PP-***GG***PP sites in the protein, consistent with the observation that long alkyl species typically provide maximum inhibitory effects. Additionally, the hydrophobic-favorable CoMSIA fields (orange) are very prominent, especially in the receptor-guided alignment (Figure 7G), consistent again with the high potency of the long alkyl (and dialkyl) bisphosphonates in GGPPS inhibition.

Conclusions

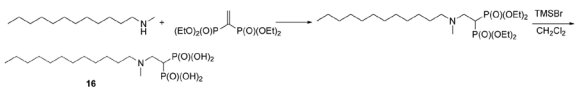
The results we have obtained above are of interest since they show how n-alkyl and dialkenyl bisphosphonates bind to, and inhibit, the GGPPS enzyme. All three n-alkyl bisphosphonates bind to the FPP site. Surprisingly, however, in one case we find that the GGPP site is also occupied, and in two structures (PDB Files 2z4x and 2z52), of the 24 GGPPS structures now investigated, we see for the first time the presence of three Mg^{2+}, located in essentially the same position as found in many FPPS-bisphosphonate structures[10–13]. The dialkenyl bisphosphonates investigated both bind with their polar groups in the F***PP***/GG***PP*** polar binding site, but one side-chain occupies the ***F***PP hydrophobic site, while the second occupies the ***GG***PP inhibitor site, basically as seen with the doubly-occupied n-alkyl bisphosphonate structure. A similar arrangement is seen with another, biphenyl-containing, “V-shaped” inhibitor. A computational investigation produced models with varying degrees of predictive utility. Docking scores performed worst (r^2 0.3), a linear interaction approximation (LIE) method had moderate predictivity (q^2 0.6, r^2 0.6), while a receptor-guided CoMSIA alignment method performed best overall (q^2 0.7, r^2 0.9). Given the selective GGPPS inhibition of the dialkenyl bisphosphonates, and their activity in cell-based assays, the availability of these new crystal structures is of broad general interest in the context of the development of potent and specific tumor cell growth inhibitors, where GGPPS inhibition offers a potentially interesting alternative to conventional FPPS-based cell growth inhibition by bisphosphonates, not least because such hydrophobic species are expected to bind only weakly to bone mineral[29].

Experimental Section

Chemicals

All reagents used were purchased from Aldrich (Milwaukee, WI). The purities of all compounds were routinely monitored by using ^1H and ^{31}P NMR spectroscopy at 400 or 500 MHz on Varian (Palo Alto, CA) Unity spectrometers using, in some instances, absolute spin-count quantitative analyses. The elemental analysis results for all new compounds are provided in the Supporting Information (Table S8). Samples of bisphosphonates **7, 11, 21, 23, 25, 28, 31, 37, 39, 43, 45, 46, 48, 49, 50, 52, 53, 55, 56, 57, 58, 59, 60, 62, 64, 65** were available from previous work[16]. Pyridinium-1-yl bisphosphonates **13, 15, 17, 18, 20, 24, 26, 27, 29, 34, 35, 36, 42, 44, 54, 63** were synthesized based on our published procedures[27].

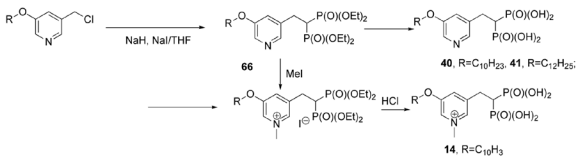
2-(N-dodecyl, N-methylaminoethyl)ethylidene-1,1-bisphosphonic acid (16) was prepared using the following scheme:



Dodecyl methylamine (1 g, 5 mmol) was mixed with tetraethyl vinylidene bisphosphonate (1.5 g, 5 mmol) and stirred for 7 days. CH_2Cl_2 (5 mL) was added, followed by TMSBr (6 g, 40 mmol). Water workup afforded **16** as a white powder. Anal. ($C_{15}H_{33}NNa_2O_6P_2 \cdot NaBr$) C, H,

N. ^{1}H NMR (D_2O, 400 MHz): δ 2.86–3.12 (4H, m), 2.57 (3H, s), 1.95–2.20 (1H, m), 1.41–1.49(m, 2H), 0.91–1.18 (18H, m), 0.65 (3H, t, *J* = 6.5 Hz); ^{31}P NMR (D_2O, 162M Hz): δ 19.9.

2-(5-Decyloxypyridin-3-yl)ethylidene-1,1-bisphosphonic acid (40) was prepared using the following scheme:

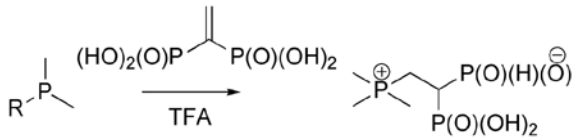


To a suspension of NaH (480 mg, 60 %, 12 mmol) in THF at 0 °C was added tetraethyl methylene bisphosphonate (2.88 g, 10 mmol). The mixture was allowed to stir at room temperature for 30 minutes. Then 3-decyloxy-5-chloromethyl-pyridine (2.83 g, 10 mmol) followed by NaI (12 mmol) were added, and the reaction mixture stirred for 2 hours at 80 °C. The mixture was quenched with saturated aq. NH_4Cl. After extraction with ether, the pyridine-bisphosphonate was purified by flash chromatography as a colorless oil (50 % yield). Direct dealkylation with TMSBr afforded **40** as a white powder. Anal. ($C_{17}H_{31}NO_7P_2 \cdot H_2O$) ^{1}H NMR ($D_2O$, 400 MHz): δ 7.91 (1H, s), 7.82 (1H, s), 7.28 (1H, s), 3.95 (3H, t, *J* = 6.4 Hz), 2.90 (2H, td, *J* = 14.8Hz, 6.8Hz), 1.93 (1H, tt, *J* = 21Hz, 6.8Hz), 1.55–1.63 (2H, m), 1.10–1.23 (2H, m), 1.00–1.11(12H, m), 0.63 (3H, t, *J* = 6.0Hz). ^{31}P NMR (D_2O, 162M Hz): δ 19.6.

3-(2, 2-bisphosphono-ethyl)-5-ethoxy-1-methyl-pyridinium iodide (14)—The pyridine-bisphosphonate **66** (1 mmol) was treated with MeI (5 mmol) in ether (5 mL) overnight. Upon removal of the solvent, the residue was hydrolyzed in refluxing concentrated HCl (37 %). After removal of the volatile solvent, the concentrated residue was washed with hexane, ether and then acetone to afford **14** as a grey powder. ^{1}H NMR (D_2O, 400 MHz): δ 8.17 (1H, s), 8.00 (1H, s), 7.83 (1H, s), 4.17 (3H, s), 3.95 (3H, t, J = 6.4 Hz), 3.01 (2H, td, J = 15Hz, 6.8Hz), 1.89 (1H, tt, J = 21Hz, 6.8Hz), 1.54–1.64 (2H, m), 1.10–1.22 (2H, m), 1.00–1.11(12H, m), 0.63 (3H, t, J = 6.0Hz). ^{31}P NMR (D_2O, 162M Hz): δ 18.6.

2-(5-Dodecyloxypyridin-3-yl)ethylidene-1,1-bisphosphonic acid (41) was prepared in the same way as for **40.** Anal. ($C_{19}H_{35}NO_7P_2 \cdot H_2O$), C, H, N. ^{1}H NMR ($D_2O$, 400 MHz): δ 7.98 (1H, s), 7.80 (1H, s), 7.38 (1H, s), 3.94 (3H, t, *J* = 6.4 Hz), 2.90 (2H, td, *J* = 14.8Hz, 6.8Hz), 1.99 (1H, tt, *J* = 21Hz, 6.8Hz), 1.55–1.63 (2H, m), 1.10–1.23 (2H, m), 1.00–1.11(14H, m), 0.63 (3H, t, *J* = 6.0Hz). ^{31}P NMR (D_2O, 162M Hz): δ 20.6.

Phosphonium-1-yl bisphosphonates (**9, 22**) were synthesized from the corresponding dimethylalkylphosphines via Michael addition in TFA:



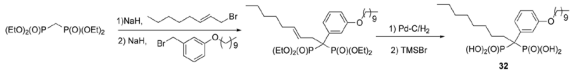
22, R=C_8H_{17}, **9**, R=$C_{12}H_{25}$,

2-(Dodecyl-dimethyl phosphonium-1-yl)ethylidene-1,1-bisphosphonic acid (9)—Dodecyl dimethylphosphine (1.2 mmol) and vinylidene-1, 1-diphosphonic acid (1.1 mmol) were dissolved in TFA (5 mL) and refluxed overnight under N_2. Upon removal of the solvent under reduced pressure, the residue was washed with hexane, ether, and then acetone to afford **9** as a white powder. Anal. ($C_{16}H_{35}Na_2O_6P_3$) C, H. ^{1}H NMR (D_2O, 400 MHz): δ 2.36–2.45

(2H, m), 2.00–2.18 (2H, m), 1.68 (6H, d, *J* = 13.5Hz), 1.65–1.69 (m, 1H), 1.36–1.45 (2H, m), 1.25–1.27 (2H, m), 1.00–1.20 (16H, m) 0.66 (3H, t, *J* = 6.5Hz). ^{31}P NMR (D_2O, 162M Hz): δ 31.1 (t, *J* = 22Hz), 17.9 (d, *J* = 22Hz).

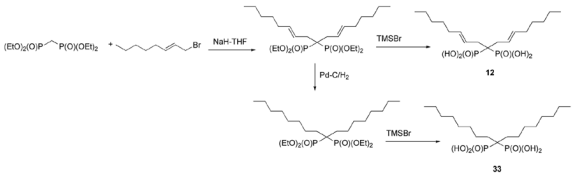
2-(Octyl-dimethyl phosphonium-1-yl)ethylidene-1,1-bisphosphonic acid (22)— **22** was prepared in the same way as **9**. Anal. ($C_{12}H_{29}O_6P_3$) C, H. ^{1}H NMR (D_2O, 500 MHz): δ 2.36–2.46 (2H, m), 2.00–2.18 (2H, m), 1.68 (6H, d, *J* = 14Hz), 1.68–1.79 (m, 1H), 1.36–1.47 (2H, m), 1.25–1.27 (2H, m), 1.00–1.20 (16H, m) 0.70 (3H, t, *J* = 6.5Hz). ^{31}P NMR (D_2O, 202M Hz): δ 32.1 (t, *J* = 23Hz), 17.9 (d, *J* = 23Hz).

1-Octyl-1-(3-decyloxybenzyl)-1, 1-bisphosphonic acid (BPH-804) was made using the following scheme:



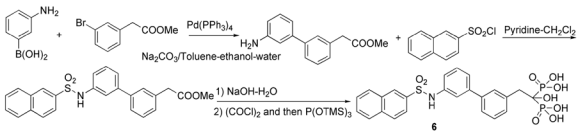
To a suspension of NaH (480 mg, 60 %, 12 mmol) in THF at 0 °C was added tetraethyl methylenebisphosphonate (2.88 g, 10 mmol). The mixture was allowed to stir at room temperature for 30 minutes. Then, 1-bromo-2-octane (1.9 g, 10 mmol) was added, dropwise. After 6 hrs, more NaH (480 mg, 60%, 12 mmol) was added, followed by 1-bromoethyl-3-decyloxy-benzene (3.59g, 11 mmol). The mixture was stirred overnight and quenched with aqueous NH_4Cl. Chromatography (5 % methanol in ethyl acetate) afforded the di-substituted intermediate as a colorless syrup (35 %). Hydrogenation in methanol in the presence of 10 % Pd/C, followed by dealkylation with TMSBr (8 equivalent) provided **32** as a white powder. Anal ($C_{26}H_{50}Na_4O_{10}P_2$·NaBr) C, H. ^{1}H NMR ($D_2O$, 400 MHz) 6.6–7.0 (4H, m), 3.75 (3H, t, *J* = 6.8Hz), 2.81–3.00 (2H, m), 0.69–1.60 (36H, m), ^{31}P NMR (D_2O, 162M Hz): δ 24.25.

1, 1-(2-octenyl)ethylidene-1,1-bisphosphonic acid (12) was made using the following scheme:



To a suspension of NaH (1 g, 60 %, 25 mmol) in THF (15 mL) at 0 °C was added tetraethyl methylenebisphosphonate (2.88 g, 10 mmol) and the mixture was allowed to stir at room temperature for 30 minutes. 1-Bromo-oct-2-enyl (5.7g, 30 mmol) was then added, dropwise, the mixture stirred overnight, then quenched with aqueous NH_4Cl. Chromatography (2% methanol in ethyl acetate) afforded the dialkylated intermediate in 65% yield. Direct dealkylation with TMSBr provided **12** as a white powder. Anal ($C_{17}H_{32}Na_2O_6P_2$·0.25H_2O) ^{1}H NMR (D_2O, 500 MHz): δ 5.5–5.7 (2H, m), 5.2–5.35 (2H, m), 2.18–2.32 (4H, m), 1.82–1.95 (4H, m), 0.98–1.21 (12H, m), 0.64 (6H, t, *J* = 6.0Hz). ^{31}P NMR (D_2O, 202M Hz): δ 23.7. Hydrogenation of the tetraethyl ester in methanol in the presence of Pd/C (10%), followed by dealkylation with TMSBr, gave **1, 1-octylethylidene-1,1-bisphosphonic acid (33)** as a white power. ^{1}H NMR (D_2O, 400 MHz): δ 1.40–1.60 (4H, m), 1.18–1.35 (4H, m), 0.98–1.18 (20H, m), 0.64 (6H, t, *J* = 6.5Hz). ^{31}P NMR (D_2O, 162M Hz): δ 25.2.

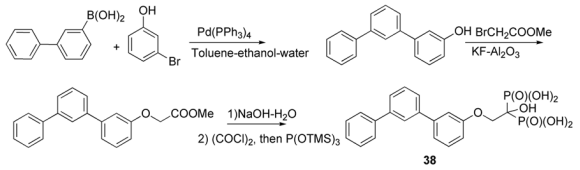
1-Hydroxy-2-[3′-(naphthalene-2-sulfonylamino)-biphenyl-3-yl]ethylidene-1,1-bisphosphonic acid (6) was made as follows:



6

3-Aminophenyl boronic acid (6 mmol), methyl 3-bromophenyl acetate (5 mmol), Na_2CO_3 (15 mmol), $Pd(PPh_3)_4$ (50 mg) in toluene (20 mL), H_2O (3 mL) and ethanol (3 mL) were heated at 110 °C under N_2 overnight. After extraction with diethyl ether, the (3′-amino-biphenyl-3-yl)-acetic acid methyl ester product was purified by column chromatography (56% yield). (3′-Amino-biphenyl-3-yl)-acetic acid methyl ester (5 mmol) and 2-naphthalene sulfonyl chloride (5 mmol) were then dissolved in anhydrous CH_2Cl_2 (10 mL), followed by addition of pyridine (10 mmol), dropwise. After washing with HCl (3N, 3 mL), the sulfonamide was isolated after chromatography as a syrup (78%). The ester (1 mmol) was then hydrolyzed with 3 N NaOH (1 mL) in methanol (5 mL) at room temperature for 1 h. After acidification with 2 N HCl, methanol was removed, and the resulting carboxylic acid filtered, then washed with water. The dried acid was dissolved in benzene (5 mL) and oxalyl chloride (2 mmol) added, followed by one drop of DMF. The reaction mixture was then stirred for 1 h. Upon removal of solvent, the crude acid chloride so obtained was dissolved in dry THF (5 mL) and $P(OTMS)_3$ (2 mmol) was added. After 3 h at room temperature, solvent was removed, methanol-H_2O (2 mL, 1:1) added, and the mixture stirred for 30 min. Concentrated aqueous NaOH was then added to precipitate the target compound, which was washed thoroughly with methanol, then ether, and dried to afford the bisphosphonic acid as its sodium salt. Anal. ($C_{24}H_{20}NO_9P_2SNa_3$) C, H, N. ^{1}H NMR (D $_2$O, 400 MHz): δ 8.18 (1H, s), 6.72–7.80 (14H, m), 3.16 (2H, t, J = 12.5 Hz). ^{31}P NMR (D_2O, 162M Hz): δ 20.3.

1-Hydroxy-2-([1′,1′;3′1″]terphenyl-3″-yloxy)-ethylidene-1,1-bisphosphonic acid (38) was made according to the following scheme:



38

3-Biphenyl boronic acid (6 mmol), 3-bromophenol (5 mmol), K_2CO_3 (15 mmol), and Pd $(PPh_3)_4$ (50 mg) in toluene (10 mL) and H_2O (3 mL) were heated at 110 °C under N_2, overnight. Upon extraction with diethyl ether, [1, 1′; 3′, 1″] terphenyl-3″-ol was purified by column chromatography as a white powder (78% yield). This compound (1 mmol) and $BrCH_2COOMe$ (1.2 mmol), K_2CO_3 (2 mmol) were then dissolved in acetone (5 mL) and refluxed overnight. Upon filtration, ([1,1′;3′,1″]terphenyl-3″-yloxy)-acetic acid methyl ester was isolated after chromatography (75 % yield). After hydrolysis with NaOH-H_2O and bisphosphorylation with $P(OTMS)_3$, as described above for **6**, **38** was obtained as a white powder. Anal. ($C_{20}H_{18}Na_2O_8P_2{\cdot}0.5H_2O$) C, H. ^{1}H NMR ($D_2O$, 400 MHz): δ 7.8 (1H, s), 7.6 (1H, d, J = 8.5 Hz), 6.9–7.7 (12H, m), 4.8 (2H, t, J = 12 Hz); ^{31}P NMR (D_2O, 162M Hz): δ 17.0.

1-Hydroxy-2-guanidino-ethylidene-bisphosphonic (51)—51 was prepared from 2-guanidine acetic acid by bisphosphorylation using the PCl_3-H_3PO_3-pyridine system, as reported previously[16]. Anal. ($C_3H_9N_3Na_2O_7P_2{\cdot}0.5CH_3OH$) C, H, N. ^{1}H NMR ($D_2O$, 400 MHz): δ 3.56 (2H, t, J = 11Hz). ^{31}P NMR (D_2O, 162M Hz): δ 16.5.

1-Hydroxy-2-(3-decyloxyphenyl)ethylidene-1,1-bisphosphonic acid. (19)—19 was prepared from (3-decyloxy-phenyl)-acetic acid methyl ester, by hydrolysis with NaOH and bisphosphorylation with $P(OTMS)_3$, as described for **6**. Anal. ($C_{18}H_{31}NaO_8P_2{\cdot}0.5H_2O$)

C, H, N. ^{1}H NMR (D_2O, 500 MHz): δ 7.07 (1H, t, J = 7.2Hz), 6.80–6.95 (2H, m), 6.68 (1H, d, 8Hz), 3.89 (2H, t, J = 6.0Hz), 3.05 (2H, tt, *J*= 22.5Hz, 12Hz), 1.55–1.62 (2H, m), 1.00–1.29 (14H, m), 0.64 (3H, t, J = 6.4Hz). ^{31}P NMR (D_2O, 202M Hz): δ 18.8.

Protein Expression and Purification—Yeast and human GGPPS were expressed and purified as described previously[6, 30]. The molecular weights of the purified enzymes were verified by mass spectrometry, and purities (>95%) determined by SDS/PAGE.

Crystallization and data collection for GGPPS complexes—Native GGPPS crystals for soaking were obtained by using the hanging drop method (Hampton Research; Laguna Niguel, CA) by mixing 2 μL of a GGPPS solution (5–10 mg/mL GGPPS in 25 mM Tris-HCl, pH 7.5 and 150 mM NaCl) with 2 μL of mother liquor (0.08 M CH_3COONa, pH 4.6, 16% PEG 4000, 6–10% glycerol, and 6–10% 1,2-propanediol), and equilibrating with 500 μL of the mother liquor. Crystals grew to 0.5 × 0.2 × 0.2 mm in 7 days, at room temperature, then were soaked in cryoprotectant solution containing 2.5 mM $MgCl_2$, 2.5 mM GGPP product or bisphosphonate (**4, 5, 10, 11, 12**), 0.08 M CH_3COONa, pH 4.6, 20% PEG 4000, 10% glycerol, and 10% 1,2-propanediol, for 3–12 h.

For preparing GGPPS-Mg-**4** ($P2_1$) crystals, 2 μL of GGPPS-substrate solution (5–10 mg/ml GGPPS, 2.5 mM Tris-HCl, 150 mM NaCl, 2.5 mM $MgCl_2$, 2.5 mM BPH252, pH 7.5) were mixed with 2 μL mother liquor (0.08 M CH_3COONa, pH 4.6, 12–16% PEG 4000, 8–10% glycerol, and 10% 1,2-propanediol) then were equilibrated with 500 μL mother liquor by using the hanging drop method, at room temperature. Crystals appeared in 5–7 days, and grew to 0.5 × 0.5 × 0.1 mm, then were soaked with cryoprotectant (containing 2.5 mM $MgCl_2$, 0.08 M CH_3COONa, pH 4.6, 20% PEG 4000, 10% glycerol, and 10% 1,2-propanediol) for 3 s, then frozen in liquid nitrogen. GGPPS-Mg-**4** ($P2_1$) was the only structure obtained by co-crystallization. All other crystals were obtained by soaking the native crystals with the bisphosphonates (using the same soaking method as described above).

X-ray diffraction data were collected at beamline BL13B1 of the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan) and Taiwan Contract BL12B2 station at Spring-8 (Hyogo, Japan). Diffraction data were processed and scaled by using the program HKL2000[31]. GGPPS-Mg-GGPP and eight GGPPS-bisphosphonate crystals belonged to the $P2_12_12_1$ space group and had typical unit cell parameters of a = 46–48 Å, b = 116–119 Å, and c = 126–130 Å. The monoclinic crystal (GGPPS-Mg-BPH252-p21) had unit cell parameters of a = 82 Å, b = 48 Å, c = 92 Å and b = 111°. Each asymmetric unit contained a dimeric GGPPS molecule. Prior to use in structural refinements, 5% randomly selected reflections were set aside for calculating R_{free} as a quality monitor[32].

Structure determination and refinement—The structures of the GGPPS-complexes were determined by using the native GGPPS structure (2DH4) solved previously, since the new crystals were isomorphous. For GGPPS, the 2Fo-Fc difference Fourier map showed clear electron densities for most amino acid residues, including those in the substrate binding site (s), but several loops and the C-terminal segments were disordered. Most product and bisphosphonate electron densities were obvious. Subsequent refinement with incorporation of the cofactors and water molecules at a 1.0σ map level yielded R and R_{free} values of 0.17–0.21 and 0.21–0.26, respectively, at 1.86–2.45 Å resolution. Statistics for the final models are listed in Tables 3,4 and S1. All manual modifications of the models were performed on an SGI Fuel (Silicon Graphics, Mountain View, CA) computer using the XtalView[33] program. Structure refinements, which included maximal likelihood and simulated-annealing protocols, were carried out by using CNS[34]. The PyMol (http://pymol.sourceforge.net/) program was used in creating figures.

GGPPS inhibition

GGPPS inhibitor screening was carried out basically as described previously[6, 9].

Molecular docking

Docking calculations were performed for all crystallographically determined protein-ligand complexes using a cross-docking approach to a series of GGPPS-bisphosphonate complex structures (**5**, PDB: 2z4w; **8**, PDB: 2e93; **9**, PDB: 2z71; **10,** PDB: 2z50**11**, PDB: 2z52). The target proteins were prepared using the protein preparation wizard in Maestro 8.0[35]. Hydrogen atoms were added and a +2 charge assigned to Mg ions present in the active site. In one case, Mg ions were modeled into the active site by initially placing them into locations observed in other GGPPS-bisphosphonate complexes. A full energy minimization in MacroModel[36] was then run on the entire protein to optimize the geometry using default protein parameters. For all other structures, hydrogen bonds were optimized to the default value. A receptor grid large enough to encompass all crystallographically observed binding sites was then generated from the prepared target protein. Constraints for the Mg^{2+} ions were created, but not actually used subsequently. Water and hetero-atoms > 5Å from the active site region were removed. Geometry optimized ligands were prepared using LigPrep[37], specifying a target pH 7.0 with tautomer and stereoisomer generation. For the docking calculations, standard-precision (SP) was specified for preliminary calculations, and the extra-precision (XP) mode specified for the final calculations. Crystallographically determined ligand poses from each structure were then compared with the top 5 poses obtained from Glide; the RMS errors are reported in the Text and in Tables S2 and S3. Docking poses were exported to Sybyl[38] for CoMSIA analysis and also into Liason,[24] for scoring function parameterization.

Docking Scoring Function Parameterization

Predicted docked poses of all ligands investigated (generated above) were imported into Liaison[24] for molecular-mechanics energy calculations. Default options were specified including a minimization sampling method using a truncated Newton algorithm. Ensemble averages of van der Waals, electrostatic and cavity (solvent exposed ligand surface area) energies were computed for the ligand-bound and ligand-free states using an implicit solvation model. The computed energies for each inhibitor complex and SlogP[25] (computed in MOE) were then imported into Strike,[39] where partial-least-squares (PLS) and multiple linear regression (MLR) methods were applied, to construct a linear equation representing binding affinity. The optimal number of components was automatically selected, and outliers identified. All molecules were used to construct an initial training set, then five test sets were selected at random from the dataset. Each test set compound was removed from the subsequent training set, with binding affinities being predicted by using the constructed linear equation. Coefficients for each energy term, fitting statistics and predictions were reported, and are shown in Tables S2 and S5.

Receptor-guided alignment

Based on the best prediction of crystallographic binding poses among all complexes, the target protein receptor (GGPPS-**8**, PDB: 2e93) was selected as the target receptor, and prepared as described above using the protein preparation wizard in Maestro 8.0[35]. Inhibitor structures were prepared as described above using LigPrep[37]. The prepared receptor grid and inhibitors were then used in a Glide 4.5[21] docking calculation, specifying extra-precision (XP).

Flexible Common Feature Alignment

Inhibitor structures were imported into MOE 2006.08,[40] where three-dimensional structures were generated using a course energy minimization protocol and the MMFF94 force field[41]. An alignment was constructed by using crystal structures of bound inhibitors as a template.

Additional inhibitors were added in a sequential manner to the template structures, employing the flexible alignment protocol implemented in MOE. Default gradient, RMSD and similarity terms were used, with the refine option selected. The final alignment was then exported to Sybyl for CoMSIA analysis.

CoMSIA Analysis

The aligned structures were exported from Maestro into Sybyl 7.3[38] and CoMSIA[18] fields computed for the aligned structures, using default grid spacing and probe atom types. Partial-least-square (PLS) regression was used to assign coefficients to grid points based on the experimentally determined GGPPS pIC_{50} values (where $pIC_{50} = -\log_{10}(IC_{50}, [M])$. The SAMPLS[42] method was used to determine the optimum number of components in the regression models. The final model was selected based on cross-validated r^2 (q^2), r^2, error and number of components, such that a statistically robust model could be generated with a minimum number of components. Test-set calculations were performed using 5 iterations of a leave 20% out approach. We used the progressive scrambling routine in Sybyl 7.3 to assess the stability of the model, by applying random activity value perturbations to structurally similar molecules within the dataset. Predictive ability, as measured by change in q^2 with respect to perturbed dependent variables, should remain constant through the series of perturbations, with the optimum value of 1, for stable models.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank H. Sagami (Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai, Japan) for providing the human GGPPS expression system. Portions of this research were carried out at the National Synchrotron Radiation Research Center, a national user facility supported by the National Science Council of Taiwan, Republic of China. This work was supported by Academia Sinica and the National Core Facility of High-throughput Protein Crystallography Grant NSC95-3112-B-001-015-Y (to A.H.-H.W.) and by the U.S. Public Health Service (NIH Grants GM-65307 and GM-073216, to E.O.). Y.Z. was supported by an American Heart Association, Midwest Affiliate, Postdoctoral Fellowship. Y.S. was supported by a Leukemia and Lymphoma Society Special Fellowship.

ABBREVIATIONS

GGPPS
: geranylgeranyl diphosphate synthase

GGPP
: geranylgeranyl diphosphate

FPPS
: farnesyl diphosphate synthase

FPP
: farnesyl diphosphate

CoMSIA
: comparative molecular similarity index analysis

QSAR
: quantitative structure activity relationship

RMSD
: root mean square deviation

CoMFA
: comparative molecular field analysis

SP
: standard precision

XP
: extra precision

RMS
: root mean square

PLS
: partial-least-squares

MLR
: multiple linear regression

DMF
: dimethylformamide

TFA
: trifluoroacetic acid

TMS
: tetramethylsilane

References

1. Kuzuguchi T, Morita Y, Sagami I, Sagami H, Ogura K. Human geranylgeranyl diphosphate synthase. cDNA cloning and expression. J Biol Chem 1999;274(9):5888–94. [PubMed: 10026212]
2. Park HJ, Kong D, Iruela-Arispe L, Begley U, Tang D, Galper JB. 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors interfere with angiogenesis by inhibiting the geranylgeranylation of RhoA. Circ Res 2002;91(2):143–50. [PubMed: 12142347]
3. Connor AM, Berger S, Narendran A, Keystone EC. Inhibition of protein geranylgeranylation induces apoptosis in synovial fibroblasts. Arthritis Res Ther 2006;8(4):R94. [PubMed: 16774691]
4. Ferella M, Montalvetti A, Rohloff P, Miranda K, Fang J, Reina S, Kawamukai M, Bua J, Nilsson D, Pravia C, Katzin A, Cassera MB, Aslund L, Andersson B, Docampo R, Bontempi EJ. A solanesyl-diphosphate synthase localizes in glycosomes of Trypanosoma cruzi. J Biol Chem 2006;281(51): 39339–48. [PubMed: 17062572]
5. Dogbo O, Laferriere A, D'Harlingue A, Camara B. Carotenoid biosynthesis: Isolation and characterization of a bifunctional enzyme catalyzing the synthesis of phytoene. Proc Natl Acad Sci U S A 1988;85(19):7054–7058. [PubMed: 16578835]
6. Szabo CM, Matsumura Y, Fukura S, Martin MB, Sanders JM, Sengupta S, Cieslak JA, Loftus TC, Lea CR, Lee HJ, Koohang A, Coates RM, 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. [PubMed: 12014956]
7. Shull LW, Wiemer AJ, Hohl RJ, Wiemer DF. Synthesis and biological activity of isoprenoid bisphosphonates. Bioorg Med Chem 2006;14(12):4130–6. [PubMed: 16517172]
8. Wiemer AJ, Tong H, Swanson KM, Hohl RJ. Digeranyl bisphosphonate inhibits geranylgeranyl pyrophosphate synthase. Biochem Biophys Res Commun 2007;353(4):921–5. [PubMed: 17208200]
9. Guo RT, Cao R, Liang PH, Ko TP, Chang TH, Hudock MP, Jeng WY, Chen CK, Zhang Y, Song Y, Kuo CJ, Yin F, Oldfield E, Wang AH. Bisphosphonates target multiple sites in both *cis*- and *trans*-prenyltransferases. Proc Natl Acad Sci U S A 2007;104(24):10022–7. [PubMed: 17535895]

10. Kavanagh KL, Dunford JE, Bunkoczi G, Russell RG, 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):22004–22012. [PubMed: 16698791]
11. Hosfield DJ, Zhang Y, Dougan DR, Broun A, Tari LW, Swanson RV, Finn J. Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. J Biol Chem 2004;279(10):8526–9. [PubMed: 14672944]
12. Kavanagh KL, Guo K, Dunford JE, Wu X, Knapp S, Ebetino FH, Rogers MJ, Russell RG, Oppermann U. The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs. Proc Natl Acad Sci U S A 2006;103(20):7829–34. [PubMed: 16684881]
13. Rondeau JM, Bitsch F, Bourgier E, Geiser M, Hemmig R, Kroemer M, Lehmann S, Ramage P, Rieffel S, Strauss A, Green JR, Jahnke W. Structural basis for the exceptional in vivo efficacy of bisphosphonate drugs. ChemMedChem 2006;1(2):267–73. [PubMed: 16892359]
14. Gabelli SB, McLellan JS, Montalvetti A, Oldfield E, Docampo R, Amzel LM. Structure and mechanism of the farnesyl diphosphate synthase from Trypanosoma cruzi: implications for drug design. Proteins 2006;62(1):80–8. [PubMed: 16288456]
15. Goffinet M, Thoulouzan M, Pradines A, Lajoie-Mazenc I, Weinbaum C, Faye JC, Seronie-Vivien S. Zoledronic acid treatment impairs protein geranyl-geranylation for biological effects in prostatic cells. BMC Cancer 2006;6:60. [PubMed: 16539721]
16. Zhang Y, Hudock MP, Krysiak K, Cao R, Bergan K, Yin F, Leon A, Oldfield E. Activity of sulfonium bisphosphonates on tumor cell lines. J Med Chem 2007;50(24):6067–6079. [PubMed: 17963374]
17. Clezardin P, Ebetino FH, Fournier PG. Bisphosphonates and cancer-induced bone disease: beyond their antiresorptive activity. Cancer Res 2005;65(12):4971–4. [PubMed: 15958534]
18. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 1994;37(24):4130–46. [PubMed: 7990113]
19. Tarshis LC, Yan M, Poulter CD, Sacchettini JC. Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-A resolution. Biochemistry 1994;33(36):10871–7. [PubMed: 8086404]
20. Martin MB, Arnold W, Heath HT 3rd, Urbina JA, Oldfield E. Nitrogen-containing bisphosphonates as carbocation transition state analogs for isoprenoid biosynthesis. Biochem Biophys Res Commun 1999;263(3):754–8. [PubMed: 10512752]
21. Glide 4.5. Schrodinger, LLC; New York, NY: 2007.
22. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem 2006;49(21):6177–96. [PubMed: 17034125]
23. Aqvist J, Medina C, Samuelsson JE. New method for predicting binding-affinity in computer-aided drug design. Protein Engineering 1994;7(3):385–391. [PubMed: 8177887]
24. Liaison 4.5. Schrodinger, LLC; New York, NY: 2007.
25. Wildman SA, Crippen GM. Prediction of physicochemical parameters by atomic contributions. Journal of Chemical Information and Computer Sciences 1999;39(5):868–873.
26. Sanders JM, Gomez AO, Mao J, Meints GA, Van Brussel EM, Burzynska A, Kafarski P, Gonzalez-Pacanowska D, Oldfield E. 3-D QSAR investigations of the inhibition of *Leishmania major* farnesyl pyrophosphate synthase by bisphosphonates. J Med Chem 2003;46(24):5171–83. [PubMed: 14613320]
27. Sanders JM, Song Y, Chan JM, Zhang Y, Jennings S, Kosztowski T, Odeh S, Flessner R, Schwerdtfeger C, Kotsikorou E, Meints GA, Gomez AO, Gonzalez-Pacanowska D, Raker AM, Wang H, van Beek ER, Papapoulos SE, Morita CT, Oldfield E. Pyridinium-1-yl bisphosphonates are potent inhibitors of farnesyl diphosphate synthase and bone resorption. J Med Chem 2005;48(8):2957–63. [PubMed: 15828834]
28. Kamath S, Buolamwini JK. Receptor-guided alignment-based comparative 3D-QSAR studies of benzylidene malonitrile tyrphostins as EGFR and HER-2 kinase inhibitors. J Med Chem 2003;46 (22):4657–68. [PubMed: 14561085]
29. Hirabayashi H, Sawamoto T, Fujisaki J, Tokunaga Y, Kimura S, Hata T. Relationship between physicochemical and osteotropic properties of bisphosphonic derivatives: rational design for osteotropic drug delivery system (ODDS). Pharm Res 2001;18(5):646–51. [PubMed: 11465420]

30. Chang TH, Guo RT, Ko TP, Wang AH, Liang PH. 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. [PubMed: 16554305]

31. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A 1997;276:307–326.

32. Brunger AT. Assessment of Phase Accuracy by Cross Validation - the Free R-Value - Methods and Applications. Acta Crystallographica Section D-Biological Crystallography 1993;49:24–36.

33. McRee DE. XtalView Xfit - A versatile program for manipulating atomic coordinates and electron density. Journal of Structural Biology 1999;125(2–3):156–165. [PubMed: 10222271]

34. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54(Pt 5):905–21. [PubMed: 9757107]

35. Maestro 8.0. Schrodinger, LLC; New York, NY: 2007.

36. MacroModel 9.5. Schrodinger, LLC; New York, NY: 2007.

37. LigPrep 2.1. Schrodinger, LLC; New York, NY: 2007.

38. Sybyl 7.3. Tripos, Inc; St. Louis, MO:

39. Strike 1.5. Schrodinger, LLC; New York, NY: 2007.

40. Molecular Operating Environment (MOE) 2006.08. Chemical Computing Group, Inc; Montreal, Quebec: 2006.

41. Halgren TA. Merck molecular force field .1. Basis, form, scope, parameterization, and performance of MMFF94. Journal of Computational Chemistry 1996;17(5–6):490–519.

42. Bush BL, Nachbar RB. Sample-Distance Partial Least-Squares - PLS optimized for many variables, with application to CoMFA. Journal of Computer-Aided Molecular Design 1993;7(5):587–619. [PubMed: 8294948]

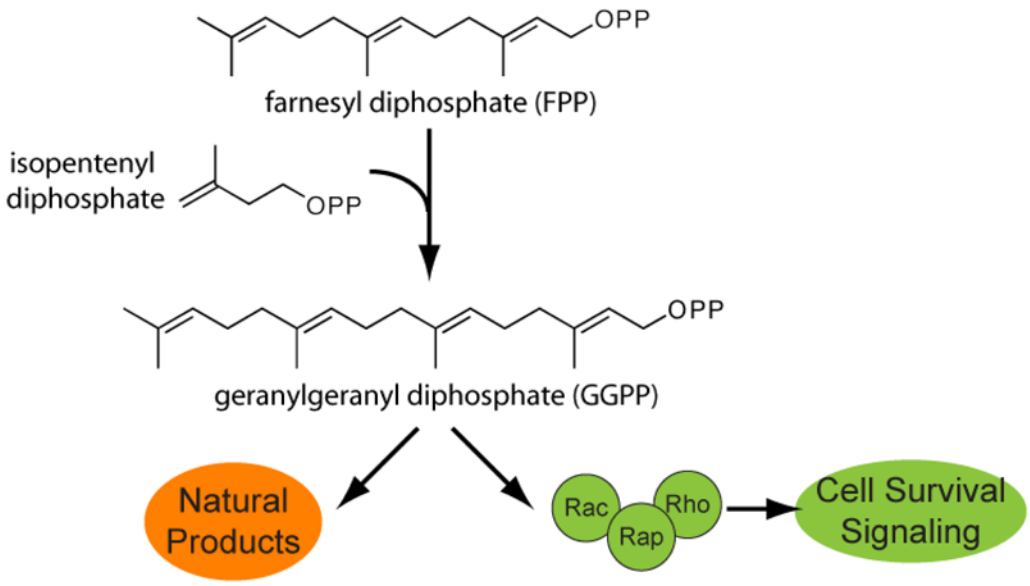


Figure 1.
GGPP biosynthesis pathway. GGPP is formed by condensation of FPP and IPP by the enzyme GGPPS. The GGPPS product can then be used prenylate cell signaling proteins such as Ras, Rac, and Rap and is also the precursor of many other isoprenoids.

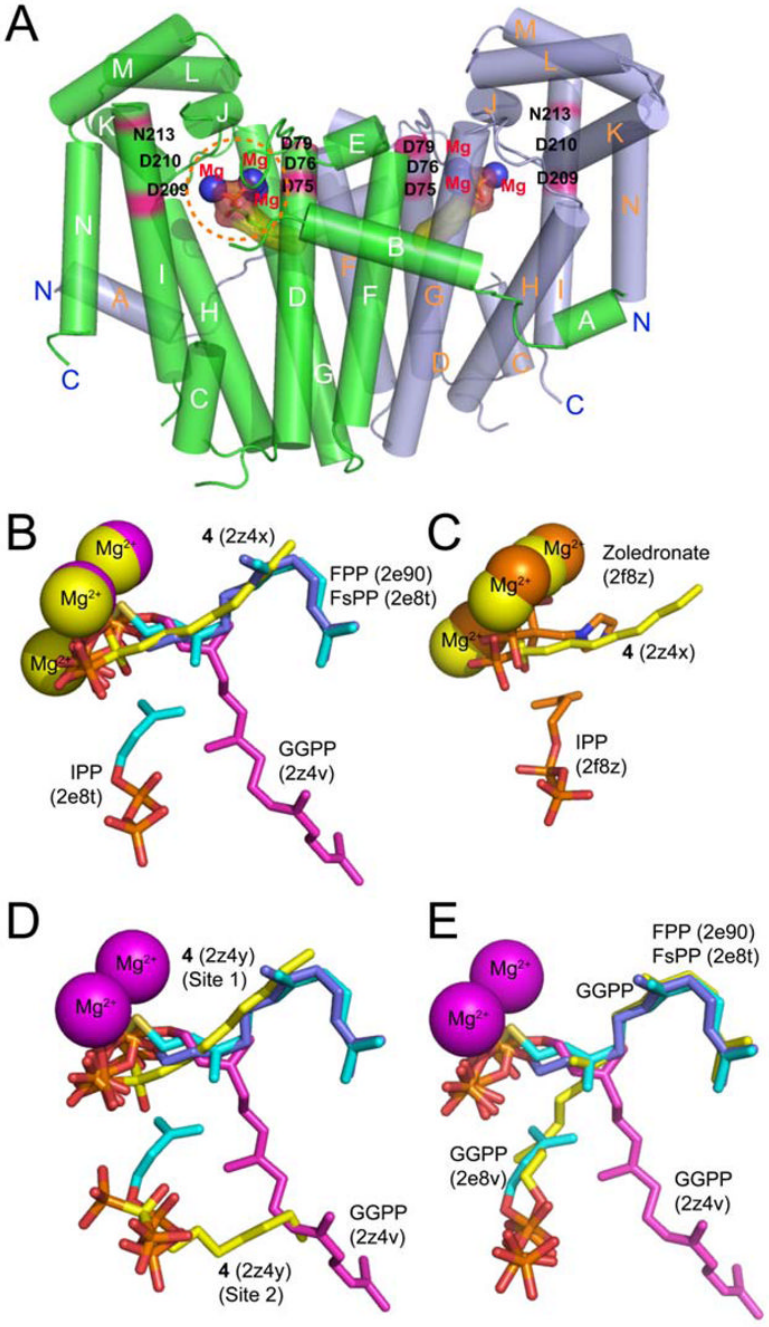


Figure 2.

GGPPS dimer structure and binding motifs. A) GGPPS dimer structure with **4** bound to Asp-rich active site (PDB File: 2z4x). B) **4** (yellow, PDB File: 2z4x) shown bound in the FPP site chelating 3 Mg^{2+} ions (yellow). Also shown are IPP + FsPP (PDB File: 2e8t), FPP (PDB File: 2e90) and GGPP (PDB File: 2z4v). C) Close overlap is observed between locations of Mg^{2+} ions in FPPS and GGPPS structures (PDB Files: 2f8z and 2z4x). D) **4** (yellow, PDB File: 2z4y) observed in two sites in the absence of Mg^{2+}. E) GGPP can bind in two distinct orientations, one (reported earlier, PDB File: 2e8v) structure extending from the IPP site to the FPP site (yellow), the second orientation (magenta, PDB File: 2z4v) is that reported in this work, and is the same as found with human GGPPS (PDB File: 2q80).

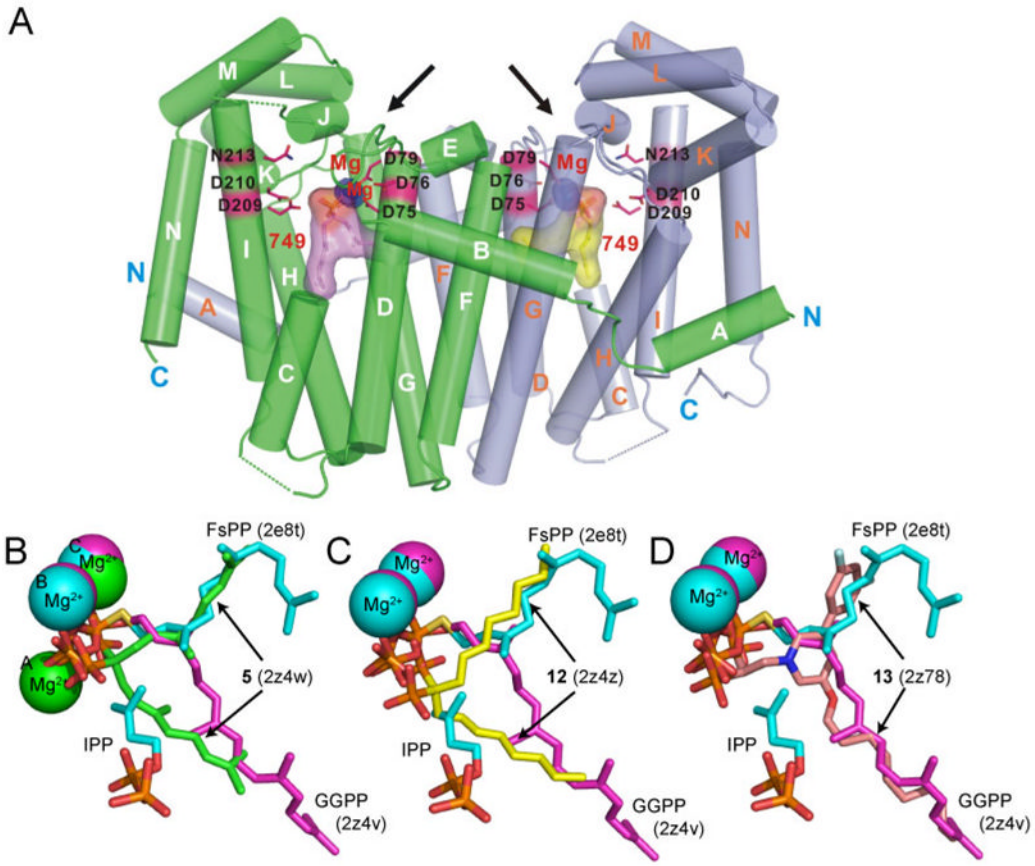


Figure 3.

GGPPS dimer structure and binding motifs of branched bisphosphonates. A) GGPPS dimer structure with **5** bound (PDB File: 2z4w). B–C) Disubstituted of "V-shape" bisphosphonate GGPPS inhibitors. The three inhibitors (**5, 12, 13**) bind in a very similar manner, despite the varying numbers of Mg^{2+} ions present. The hydrophobic tails extend into both the FPP as well as the GGPP sites. B) **5** (green) binds with two Mg^{2+} ions (green) (PDB File: 2z4w). C) **12** (yellow) binds without any Mg^{2+} apparent (PDB File: 2z4z). D) **13** (pink) likewise binds to both the FPP and GGPP sidechain sites and no Mg^{2+} is apparent (PDB File: 2z78).

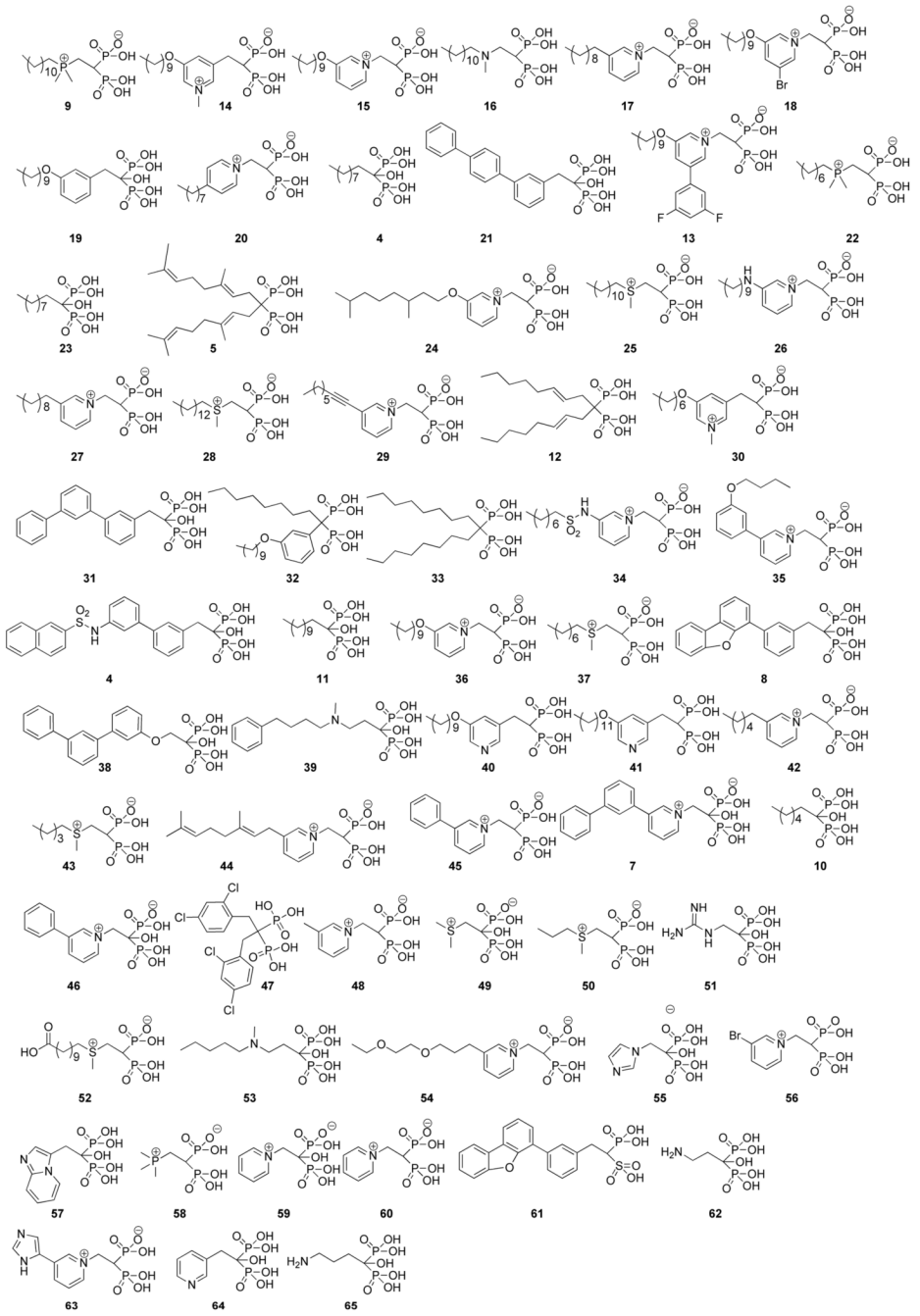


Figure 4.
GGPPS inhibitors investigated, rank-ordered in terms of decreasing activity from top left to bottom right.

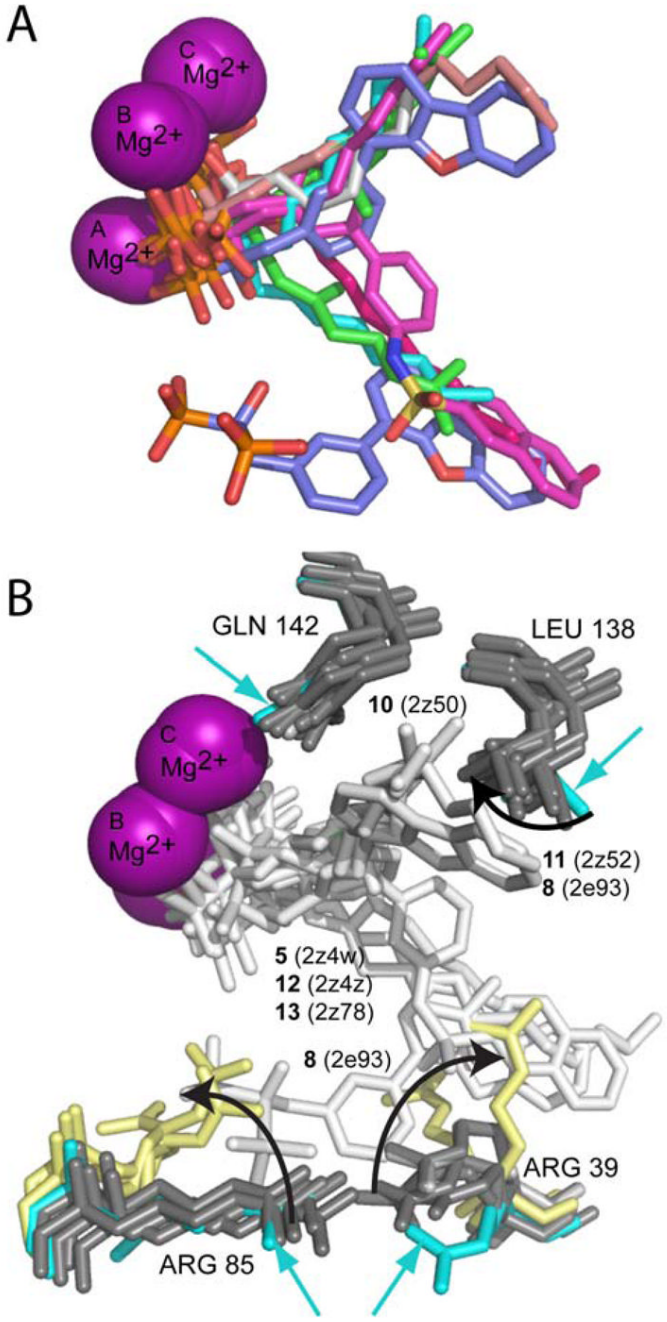


Figure 5.

GGPPS binding site organization. A) Superimposition of bisphosphonate inhibitors showing chelation of up to three Mg^{2+} ions, and several distinct binding modes. B) Highly variable residues surrounding the GGPPS inhibitors bound in the active site. Yellow residues have large deviations in sidechain conformation; grey residues are representative of those seen other all other structures. The cyan shading and arrows indicate the sidechain conformations that gave the best docking results.

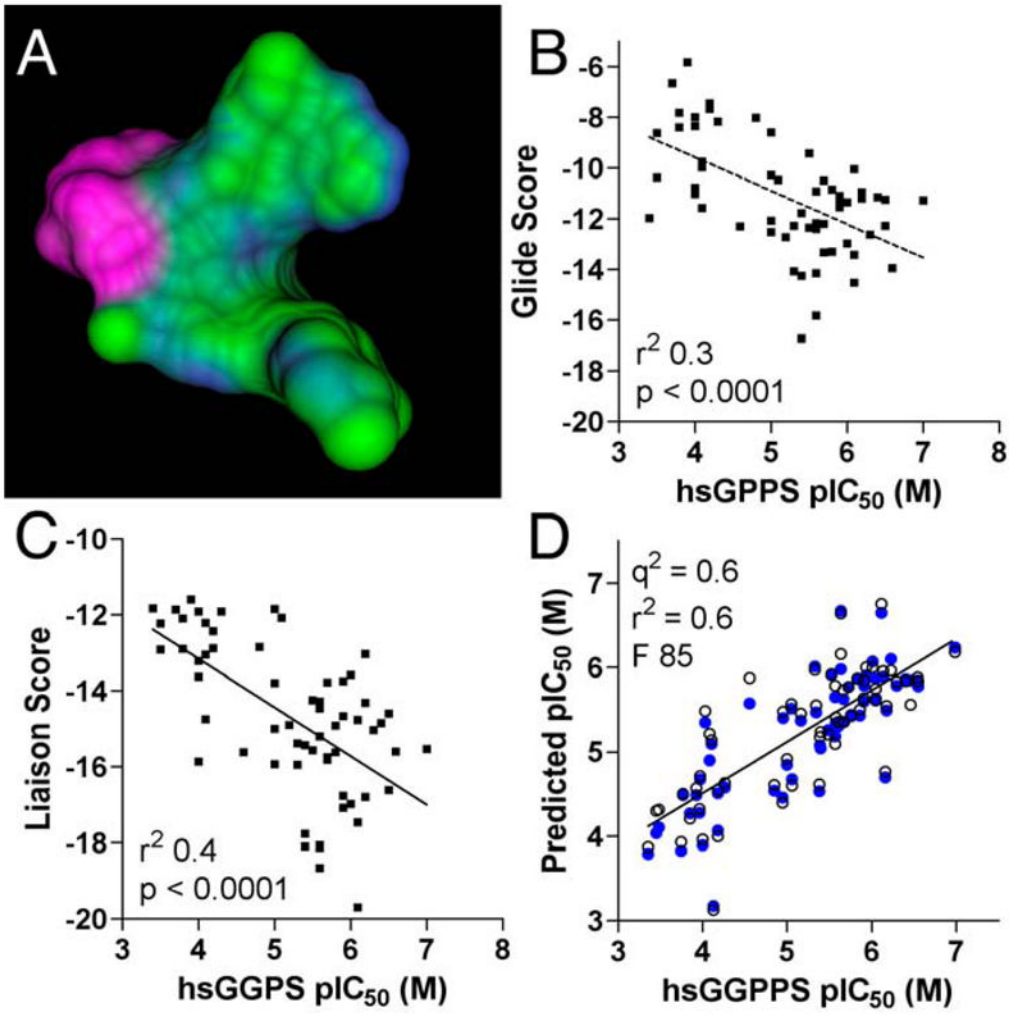


Figure 6.

Docking and Linear Interaction Approximation Results. A) van der Waals surface representation of the docked poses for all of the GGPPS inhibitors investigated, showing two distinct hydrophobic binding pockets, but only a single polar (bisphosphonate) binding domain (magenta). B) Correlation of Glide score with GGPPS pIC_{50} showing $r^2 = 0.3$. C) Correlation plot of Liaison score against GGPPS pIC_{50} showing $r^2 = 0.4$. D) GGPPS pIC_{50} experiment vs. predicted activity for reparameterized scoring function showing $q^2 = 0.6$, $r^2 = 0.6$, F = 85.

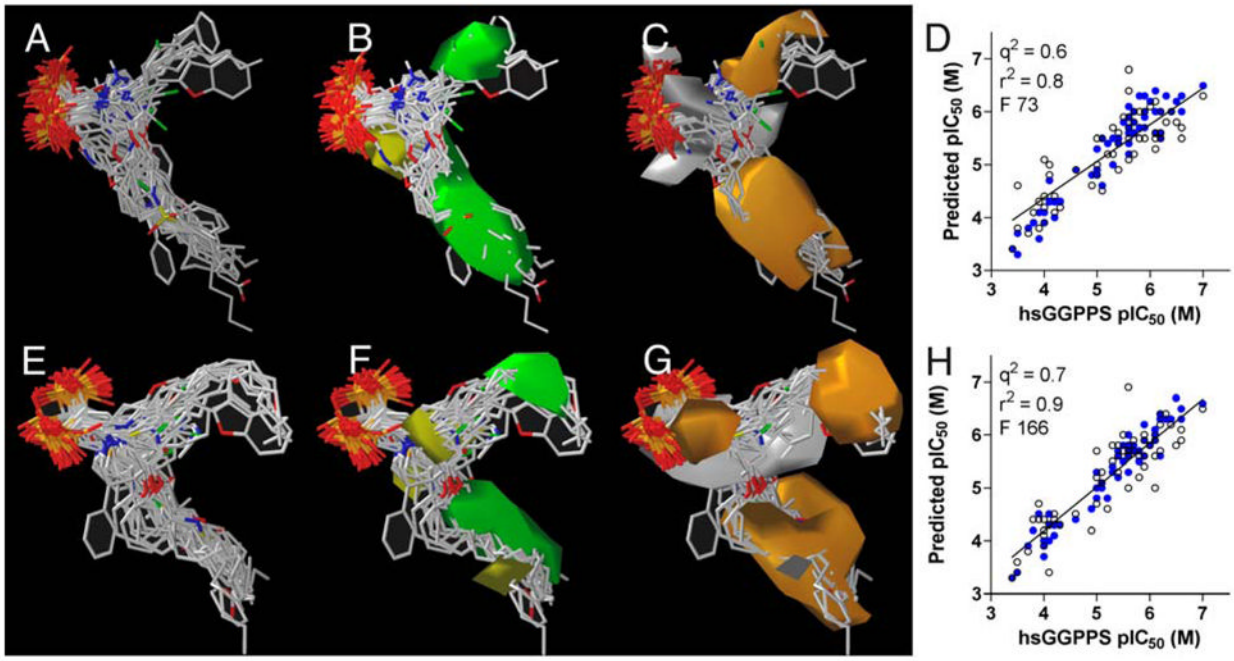


Figure 7.

CoMSIA fields and results for common feature and receptor-guided alignments. A–D) Common feature alignment, fields and prediction statistics. A) Common feature alignment. B) CoMSIA steric fields showing favorable regions in green and disfavored in yellow. C) CoMSIA hydrophobic fields showing favorable regions in orange and disfavored in white. D) Correlation plot of experimental GGPPS pIC_{50} vs. predicted activity, showing $q^2 = 0.6$, $r^2 = 0.8$ and F = 73. E–H) Receptor-guided alignment, fields and prediction statistics. E) Receptor-guided alignment. F) CoMSIA steric fields showing favorable regions in green and disfavored in yellow. G) CoMSIA hydrophobic fields showing favorable regions in orange and disfavored in white. H) Correlation plot of experimental GGPPS pIC_{50} vs. predicted activity, showing $q^2 = 0.7$, $r^2 = 0.9$ and F = 166.

Table 1

Glide, Liaison and Scoring Function Training and Test Set Predictions for GGPPS Inhibition.

			Computed Values							Test Set pIC50 Predictions [b]						
CPD[a]	IC50 (μM)	pIC50 (M)	Glide Score	Liaison Score	Liaison <Uvdw>	Liaison <Ucav>	Liaison <Uele>	SlogP	Training	1	2	3	4	5	Pred	Residual
9	0.1	7.0	−11.26	−15.52	25.97	−0.92	−588.38	0.20	6.2	5.9	6.2	6.3	6.3	6.2	6.2	0.8
14	0.3	6.6	−13.94	−15.59	23.47	0.97	−550.66	−1.05	5.8	5.7	5.8	5.8	5.8	**5.8**	5.8	0.7
15	0.3	6.5	−12.25	−16.60	21.84	0.72	−582.70	−1.23	5.8	5.6	**5.9**	5.8	5.9	5.9	5.9	0.7
16	0.3	6.5	−11.24	−14.59	30.16	−1.64	−541.97	−2.56	5.8	**5.2**	**5.9**	5.8	6.1	5.9	5.6	0.9
17	0.4	6.4	−11.13	−14.84	27.86	−0.16	−541.31	−1.06	5.9	5.6	5.9	5.9	5.9	5.9	5.8	0.6
18	0.5	6.3	−12.60	−15.02	23.07	1.49	−498.16	−0.39	5.8	5.9	5.8	5.9	5.7	5.8	5.8	0.5
19	0.6	6.2	−11.18	−16.79	10.20	1.35	−499.39	−0.92	6.1	**6.0**	6.2	6.0	6.2	6.2	6.0	0.3
20	0.7	6.2	−11.11	−14.31	33.90	0.27	−554.07	−1.84	5.5	5.3	5.5	5.5	**5.5**	5.5	5.5	0.6
4	0.7	6.2	−10.94	−13.03	52.57	0.86	−391.39	−2.93	4.7	5.1	4.6	4.9	4.6	4.6	4.8	1.4
21	0.7	6.1	−13.40	−17.46	14.85	1.62	−476.31	−1.10	5.9	5.9	**6.0**	5.9	6.0	6.0	6.0	0.2
13	0.8	6.1	−14.50	−19.69	3.81	1.07	−622.40	1.26	6.7	6.4	6.7	6.6	6.7	**6.8**	6.8	−0.6
22	0.9	6.1	−10.04	−14.75	37.02	−0.61	−524.48	−1.36	5.6	5.5	5.6	5.7	5.7	5.6	5.6	0.4
23	0.9	6.0	−11.34	−13.58	14.00	0.73	−317.90	−2.54	5.9	5.8	6.0	**5.8**	6.1	6.1	5.8	0.3
5	1.0	6.0	−12.97	−16.97	26.97	0.87	−546.14	1.54	6.1	6.3	**6.0**	6.2	5.9	**6.0**	6.0	0.0
24	1.2	5.9	−11.54	−17.07	19.71	0.49	−541.89	−1.52	5.9	5.6	6.0	5.9	6.0	**6.0**	6.0	−0.1
25	1.2	5.9	−11.08	−14.66	29.11	0.23	−565.45	−0.83	5.8	5.6	5.8	**5.8**	5.8	5.8	5.8	0.1
26	1.3	5.9	−11.43	−16.75	21.15	2.17	−551.14	−1.19	5.6	5.7	5.6	**5.6**	5.6	5.7	5.6	0.3
27	1.4	5.9	−11.32	−13.74	30.78	0.56	−473.26	−2.52	5.4	5.3	5.4	5.4	5.5	**5.5**	5.5	0.4
28	1.5	5.8	−13.30	−14.90	29.51	0.33	−552.59	−0.05	5.9	5.9	5.8	6.0	5.8	5.8	5.9	0.0
29	1.8	5.8	−10.83	−15.60	22.78	1.79	−529.85	−2.42	5.4	5.4	5.5	**5.4**	5.5	5.6	5.4	0.3
12	1.9	5.7	−12.18	−15.74	34.93	0.39	−557.13	0.42	5.8	5.9	5.7	5.9	5.6	5.7	5.8	0.0
30	2.1	5.7	−10.50	−13.77	31.69	0.90	−513.18	−2.40	5.4	**5.3**	5.4	5.4	**5.4**	5.4	5.4	0.3
31	2.1	5.7	−13.32	−15.81	23.22	1.77	−326.88	−1.10	5.6	6.1	5.6	5.7	5.6	5.7	5.7	−0.1
32	2.3	5.6	−15.81	−18.12	17.53	0.13	−631.08	2.88	6.7	6.6	6.6	6.8	6.5	6.6	6.6	−1.0
33	2.3	5.6	−12.15	−14.46	36.16	−0.58	−496.32	0.87	6.0	6.1	5.9	**6.2**	5.9	5.9	6.2	−0.5
34	2.5	5.6	−10.92	−18.05	16.50	3.32	−548.30	−2.64	5.3	5.4	5.4	5.3	**5.4**	5.5	5.4	0.2
35	2.7	5.6	−12.40	−15.19	27.07	2.08	−600.40	−1.90	5.3	5.3	5.3	**5.4**	5.3	**5.4**	5.4	0.2
6	2.7	5.6	−14.14	−18.65	1.95	5.59	−534.20	−0.22	5.6	6.2	5.7	5.7	5.5	5.9	5.8	−0.2
11	2.7	5.6	−12.24	−14.27	43.79	0.50	−507.04	−1.76	5.2	5.3	**5.1**	5.3	5.1	5.1	5.1	0.5
36	3.0	5.5	−12.34	−15.57	25.75	0.37	−565.66	−0.45	5.9	5.8	5.9	6.0	**5.9**	5.9	5.9	−0.4
37	3.3	5.5	−9.40	−14.24	38.18	0.45	−517.54	−2.39	5.3	**5.2**	5.2	5.3	5.3	5.3	5.2	0.3
8	4.0	5.4	−16.72	−18.09	47.42	1.40	−467.20	−0.87	5.0	**5.6**	**4.9**	5.3	4.8	4.9	5.2	0.2
38	4.1	5.4	−14.23	−17.76	32.46	3.16	−484.01	−1.27	5.1	5.6	5.0	5.2	4.9	5.1	5.2	0.2
39	4.2	5.4	−11.76	−15.42	40.03	2.22	−438.34	−4.75	4.5	4.7	4.5	4.6	4.6	4.7	4.6	0.8
40	4.5	5.3	−12.27	−15.37	34.33	1.11	−573.12	−0.84	5.5	**5.6**	5.4	5.6	5.4	5.5	5.6	−0.2
41	4.7	5.3	−14.05	−15.94	24.00	0.41	−635.98	−0.06	6.0	5.8	6.0	6.1	6.0	6.0	6.0	−0.6
42	6.8	5.2	−12.71	−14.89	34.08	0.26	−571.84	−2.62	5.4	5.1	5.4	5.4	**5.5**	**5.4**	5.5	−0.3
43	8.8	5.1	−10.46	−12.08	48.46	1.10	−488.47	−3.56	4.7	4.8	**4.6**	4.8	4.7	4.7	4.6	0.5
44	8.9	5.0	−12.06	−15.93	27.08	1.46	−579.42	−1.51	5.5	5.5	5.5	**5.6**	5.5	5.6	5.6	−0.5
45	10	5.0	−10.26	−13.80	42.22	1.52	−478.65	−3.08	4.8	**5.0**	4.8	4.9	**4.8**	**4.9**	4.9	0.1
7	11	5.0	−12.50	−14.99	17.59	3.07	−444.31	−2.09	5.4	5.7	5.5	**5.4**	5.4	**5.6**	5.5	−0.5
10	11	5.0	−8.60	−11.85	56.25	0.98	−504.43	−3.71	4.5	4.7	4.3	4.6	**4.4**	4.4	4.4	0.5
46	14	4.8	−8.02	−12.84	44.04	2.38	−476.51	−3.76	4.5	4.9	4.5	4.6	4.5	4.6	4.6	0.2
47	28	4.6	−12.29	−15.61	38.97	0.97	−526.22	0.57	5.6	**5.9**	5.4	**5.8**	5.4	5.5	5.9	−1.3
48	54	4.3	−8.17	−11.91	47.88	1.04	−478.59	−4.44	4.6	**4.6**	4.5	4.6	4.6	4.6	4.6	−0.4
49	66	4.2	−7.66	−12.43	54.44	1.73	−443.12	−5.80	4.1	4.3	**4.0**	4.1	4.1	4.1	4.0	0.2
50	66	4.2	−7.43	−12.88	49.63	1.19	−500.59	−4.34	4.5	4.6	4.5	4.6	4.5	4.5	4.5	−0.4
51	74	4.1	−9.73	−12.21	63.30	2.65	−484.91	−9.02	3.2	3.2	**3.1**	3.2	3.3	3.3	3.1	1.0
52	79	4.1	−9.97	−14.74	34.87	0.96	−500.12	−3.49	5.1	5.0	**5.1**	**5.1**	**5.2**	5.2	5.1	−1.0
53	83	4.1	−11.57	−13.04	43.21	−0.90	−515.34	−5.59	4.9	4.3	4.9	4.8	**5.2**	5.0	5.2	−1.1
54	94	4.0	−10.77	−15.85	22.98	1.21	−548.60	−3.76	5.3	5.0	**5.4**	5.3	5.6	**5.5**	5.5	−1.5

			Computed Values							Test Set pIC50 Predictions [b]						
CPD[a]	IC50 (μM)	pIC50 (M)	Glide Score	Liaison Score	Liaison <Uvdw>	Liaison <Ucav>	Liaison <Uele>	SlogP	Training	1	2	3	4	5	Pred	Residual
55	100	4.0	−8.33	−11.90	59.35	2.13	−461.76	−5.52	3.9	4.2	3.8	4.0	3.9	3.9	4.0	0.0
56	107	4.0	−7.98	−13.63	44.16	1.51	−505.70	−3.91	4.7	4.8	4.6	4.7	4.7	4.7	4.7	−0.7
57	108	4.0	−11.02	−13.21	50.27	2.06	−483.23	−4.74	4.3	4.5	4.2	**4.4**	**4.3**	4.3	4.3	−0.4
58	118	3.9	−5.81	−11.60	52.98	0.48	−404.20	−4.77	4.5	4.6	4.4	4.6	**4.6**	4.5	4.6	−0.6
59	141	3.8	−8.40	−12.09	54.71	0.88	−445.35	−5.43	4.3	4.3	**4.2**	4.3	4.4	4.3	4.2	−0.4
60	169	3.8	−7.82	−12.89	51.06	0.81	−506.09	−4.75	4.5	4.5	4.4	4.6	4.6	4.5	4.5	−0.7
62	180	3.7	−6.64	−11.86	53.71	2.12	−455.27	−7.05	3.8	**3.9**	3.8	3.9	4.0	**3.9**	3.9	−0.2
63	329	3.5	−10.36	−12.91	46.65	3.31	−508.64	−5.04	4.1	**4.4**	4.0	4.2	4.1	**4.2**	4.3	−0.8
64	350	3.5	−8.62	−12.23	58.53	1.87	−433.67	−5.04	4.0	**4.4**	3.9	**4.2**	4.0	4.0	4.3	−0.8
65	440	3.4	−11.97	−11.82	64.49	1.00	−637.28	−6.66	3.8	3.6	3.7	3.9	**3.9**	3.8	3.9	−0.5

[a] structures of all inhibitors shown in Figure 1

[b] bold values indicated compounds not included in training set.

Table 2

Receptor-Guided Alignment CoMSIA training and Test Set Predictions for GGPPS Inhibition.

	Experimental		GGPPS pIC50 Predictions							
				Test Sets						
CPD	IC50 (μM)	pIC50 (M)	Training	1	2	3	4	5	Predicted	Residual
9	0.10	7.0	6.6	6.4	6.7	6.5	6.3	6.5	6.5	0.5
14	0.28	6.6	6.5	**6.1**	6.6	6.5	6.5	6.5	6.1	0.4
15	0.28	6.6	6.3	6.0	6.4	6.3	6.1	**5.9**	5.9	0.7
16	0.35	6.5	6.7	6.3	6.5	6.7	**5.8**	6.7	5.8	0.6
17	0.40	6.4	6.3	6.3	**6.2**	6.3	6.2	6.4	6.2	0.2
18	0.51	6.3	6.3	6.6	6.3	**6.4**	6.4	6.3	6.4	−0.1
19	0.59	6.2	6.4	6.3	6.5	6.4	6.5	6.4	6.4	−0.2
20	0.66	6.2	6.3	**6.2**	6.2	6.3	6.2	6.3	6.2	0.0
4	0.71	6.2	5.6	5.6	5.7	5.7	5.8	5.6	5.7	0.5
21	0.72	6.1	6.0	6.0	**5.9**	6.0	6.1	6.0	5.9	0.2
13	0.76	6.1	6.1	6.3	6.1	6.1	6.2	**6.0**	6.0	0.1
22	0.89	6.1	5.9	5.9	6.0	5.8	**5.6**	5.9	5.6	0.5
23	0.89	6.1	6.0	**5.0**	6.0	6.1	6.0	6.0	5.0	1.0
5	0.98	6.0	5.8	5.6	6.0	5.8	5.7	5.9	5.8	0.2
24	1.15	5.9	5.9	5.8	5.9	6.0	6.1	**6.0**	6.0	−0.1
25	1.23	5.9	5.6	5.6	5.9	**5.6**	5.5	5.5	5.6	0.4
26	1.26	5.9	6.2	5.9	**5.9**	6.1	6.3	6.2	5.9	0.0
27	1.38	5.9	5.9	5.7	6.0	5.9	**5.4**	5.7	5.4	0.4
28	1.48	5.8	5.7	5.9	5.6	5.8	5.7	**5.6**	5.6	0.2
29	1.74	5.8	5.5	**5.2**	5.5	5.6	5.6	5.5	5.2	0.5
12	1.86	5.7	5.8	6.0	5.9	5.9	5.9	5.8	5.9	−0.2
30	2.14	5.7	5.7	5.5	5.8	**5.7**	5.5	5.8	5.7	−0.1
31	2.14	5.7	5.7	5.9	5.8	**5.7**	5.6	5.7	5.7	−0.1
32	2.34	5.6	5.7	5.5	5.9	5.7	**5.8**	5.7	5.8	−0.2
33	2.51	5.6	6.0	6.3	**6.9**	5.9	6.2	6.0	6.9	−1.3
34	2.51	5.6	5.8	5.8	5.6	5.8	5.7	5.8	5.7	−0.1
35	2.51	5.6	5.3	**5.0**	5.4	5.3	4.9	5.4	5.0	0.6
6	2.69	5.6	5.3	**5.8**	5.5	5.2	5.5	5.3	5.8	−0.3
11	2.69	5.6	5.6	5.3	5.8	5.6	5.7	**5.6**	5.6	−0.1
36	3.02	5.5	5.5	5.8	5.5	5.5	**5.7**	5.5	5.7	−0.1
37	3.24	5.5	5.8	5.5	**6.0**	5.8	5.5	5.7	6.0	−0.6
8	3.98	5.4	5.2	4.9	**5.3**	5.3	5.2	5.3	5.3	0.1
38	4.07	5.4	5.6	5.8	5.5	5.6	5.5	**5.6**	5.6	−0.2
39	4.17	5.4	5.7	5.7	5.5	**5.8**	5.5	5.6	5.8	−0.5
40	4.57	5.3	5.3	5.8	5.4	5.3	5.8	5.3	5.5	−0.2
41	5.01	5.3	5.4	**5.8**	5.4	5.5	5.6	5.4	5.8	−0.5
42	6.31	5.2	4.8	4.6	4.7	4.9	**4.6**	5.0	4.6	0.6
43	8.71	5.1	5.1	4.9	5.1	**5.1**	4.9	5.0	5.1	−0.1
44	8.91	5.1	5.0	5.2	**5.3**	5.0	5.0	5.0	5.3	−0.2
45	10.00	5.0	5.0	4.9	4.8	5.1	4.8	**5.2**	5.2	−0.2
7	11.22	5.0	5.3	5.2	5.1	**5.7**	5.1	5.3	5.7	−0.8
10	11.22	5.0	4.8	4.6	5.0	4.8	4.6	**4.7**	4.7	0.3
46	14.13	4.9	4.6	**4.2**	4.6	4.6	4.4	4.7	4.2	0.7
47	27.54	4.6	4.4	4.6	4.6	4.5	4.4	4.4	4.5	0.1
48	53.70	4.3	4.3	4.0	4.4	4.3	**4.3**	4.2	4.3	0.0
49	66.07	4.2	4.3	4.3	4.4	4.3	4.4	**4.5**	4.5	−0.3
50	66.07	4.2	4.1	4.0	**4.4**	4.1	4.1	4.2	4.4	−0.2
51	74.13	4.1	4.0	3.9	4.0	**3.4**	4.0	4.1	3.4	0.7
52	79.43	4.1	4.5	4.7	**6.0**	4.4	4.6	4.4	6.0	−1.9
53	79.43	4.1	4.3	4.2	4.1	4.2	**4.3**	4.2	4.3	−0.2

J Med Chem. Author manuscript; available in PMC 2008 September 25.

CPD	Experimental		GGPPS pIC50 Predictions							
				Test Sets						
	IC50 (μM)	pIC50 (M)	Training	1	2	3	4	5	Predicted	Residual
54	93.33	4.0	3.7	**4.4**	3.8	3.8	3.8	3.8	4.4	−0.4
55	100.00	4.0	3.7	4.0	3.7	3.9	**3.9**	3.8	3.9	0.1
56	107.15	4.0	4.0	4.0	4.1	4.1	4.0	**4.1**	4.1	−0.1
57	107.15	4.0	3.9	4.2	4.0	**4.2**	4.1	3.9	4.2	−0.2
58	117.49	3.9	4.5	4.4	**4.7**	4.6	4.4	4.4	4.7	−0.8
59	141.25	3.9	4.5	4.3	4.4	4.4	4.4	4.5	4.4	−0.5
60	169.82	3.8	4.2	4.1	4.2	**4.4**	4.3	4.1	4.4	−0.6
62	181.97	3.7	3.9	3.7	3.8	3.9	3.7	3.8	3.8	0.0
63	316.23	3.5	3.4	3.4	3.6	3.4	3.4	3.3	3.4	0.1
64	331.13	3.5	3.4	**3.6**	3.5	3.5	3.6	3.4	3.6	−0.2
65	436.52	3.4	3.3	3.6	3.3	3.2	3.2	3.3	3.3	0.0
q^2	0.68	0.66	0.78	0.67	0.64	0.62				
r^2	0.93	0.89	0.96	0.94	0.91	0.94				
N	5	3	5	4	4	5				
F	166	122	207	148	117	146				
n	61	51	51	51	51	51				
%Steric	0.22	0.25	0.24	0.21	0.22	0.20				
%Hydrophobic	0.36	0.38	0.38	0.35	0.35	0.35				
%Donor	0.23	0.20	0.22	0.23	0.22	0.24				
%Acceptor	0.20	0.18	0.17	0.20	0.21	0.21				

N = number of components; n = Number of training set compounds

[a] structures of all inhibitors shown in Figure 1

[b] bold values indicated compounds not included in training set.

Table 3

Data Collection and Refinement Statistics for GGPPS-Bisphosphonate Crystals.

Names	**GGPPS-Mg-4**	**GGPPS-Mg-11**	**GGPPS-Mg-4**	**GGPPS-Mg-GGPP**
PDB number	**2z4x**	**2z52**	**2z4y**	**2z4v**
		Data Collection		
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Resolution (Å)[a]	30–1.90 (1.97–1.90)	30–2.13 (2.21–2.13)	30–2.10 (2.18–2.10)	30–1.86 (1.93–1.86)
Unit Cell Dimensions				
a (Å)	82.41	46.09	45.95	47.16
b (Å)	47.87	116.29	116.21	115.94
c (Å)	91.8	128.41	126.02	128.4
β(°)	110.86			
No. of Reflections				
Observed	166199 (15414)	264434 (25207)	221099 (20940)	313204 (30826)
Unique	51962 (5138)	39542 (3878)	40023 (3951)	59802 (5928)
Completeness (%)	98.4 (98.4)	99.9 (99.7)	99.4 (99.7)	99.4 (99.9)
R_{merge} (%)	6.1 (41.8)	6.9 (32.5)	5.7 (45.3)	3.5 (10.2)
I/s(I)	18.8 (3.1)	26.0 (7.6)	29.1 (4.3)	42.4 (17.9)
		Refinement		
No. of reflections	49646 (4064)	38397 (3627)	38785 (3538)	59160 (5770)
R_{work} (%)	19.6 (28.5)	17.2 (17.9)	18.1 (19.7)	17.3 (18.6)
R_{free} (%)	24.6 (31.8)	21.0 (23.0)	23.2 (25.6)	21.7 (23.4)
Geometry deviations				
Bond lengths (Å)	0.018	0.019	0.019	0.021
Bond angles (°)	1.7	1.7	1.7	1.8
No. of all protein atoms	4659	4848	4670	5129
Mean B-values (Å^2)	32.5	28.7	37	27
No. of all cofactor atoms	42	48	55	63
Mean B-values (Å^2)	23.1	28.2	48.7	28.5
No. of water molecules	667	547	391	819
Mean B-values (Å^2)	52	43.3	49.3	45.5
Ramachandran plot (%)				
Most favored	95.5	96.8	95.7	96.5
Additionally allowed	4.5	3.2	4.3	3.5
Generously allowed	0	0	0	0

[a] Values in the parenthesis are the highest resolution shells.

Table 4

Data collection and refinement statistics for GGPPS-bisphosphonate crystals.

Names	**GGPPS-Mg-5**	**GGPPS-Mg-12**	**GGPPS-Mg-13**
PDB number	**2z4W**	**2z4z**	**2z78**
	Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Resolution (Å)[a]	30–2.45 (2.54–2.45)	30–2.09 (2.16–2.09)	30–2.10 (2.18–2.10)
Unit cell dimensions			
a (Å)	47.52	47.61	48.33
b (Å)	116.55	116.34	116.64
c (Å)	129.07	129.03	129.78
No. of reflections			
Observed	134955 (13120)	224260 (19223)	197140 (18388)
Unique	26937 (2624)	42400 (4090)	42730 (4179)
Completeness (%)	99.6 (99.6)	97.8 (95.3)	97.5 (97.9)
R_{merge} (%)	10.3 (53.2)	4.4 (27.4)	11.5 (33.6)
I/s(I)	14.8 (2.9)	40.6 (6.9)	16.3 (6.5)
	Refinement		
No. of reflections	24802 (2066)	40792 (3658)	41719 (3940)
R_{work} (%)	20.5 (33.3)	18.7 (24.6)	18.1 (21.9)
R_{free} (%)	26.0 (37.6)	23.7 (26.9)	23.7 (28.0)
Geometry deviations			
Bond lengths (Å)	0.01	0.015	0.018
Bond angles (°)	1.4	1.6	1.7
No. of all protein atoms	4816	5104	5214
Mean B-values ($Å^2$)	39.6	40.7	29.9
No. of all cofactor atoms	62	51	70
Mean B-values ($Å^2$)	36.9	47	48.4
No. of water molecules	314	619	711
Mean B-values ($Å^2$)	44.4	55	47.3
Ramachandran plot (%)			
Most favored	94.7	95.2	94.1
Additionally allowed	4.9	4.8	5.9
Generously allowed	0.4	0	0

[a]Values in the parenthesis are the highest resolution shells.