A theoretical study of anthrax lethal factor inhibition by a set of novel carbamimidolyl-aryl-vinyl-carboxamidines: A possible mechanism involving zinc-ligation by amidine

Tam Luong Nguyen a, Rekha G. Panchal a, Igor A. Topol b, Douglas Lane a, Tara Kenny a, James C. Burnett a, Ann R. Hermone a, Connor McGrath a, Stanley K. Burt b, Rick Gussio c, Sina Bavari d,*

a Target Structure-Based Drug Discovery Group, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD 21702, USA
b Advanced Biomedical Computing Center, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD 21702, USA
c Target Structure-Based Drug Discovery Group, Information Technology Branch, Developmental Therapeutics Program, National Cancer Institute, Frederick, MD 21702, USA
d US Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, MD 21702, USA

Received 3 May 2007; received in revised form 2 July 2007; accepted 5 July 2007
Available online 20 July 2007

Abstract

A congeneric set of carbamimidolyl-aryl-vinyl-carboxamidines from the National Cancer Institute (NCI) open chemical repository were identified as potent inhibitors of anthrax lethal factor (LF), a zinc-dependent metalloprotease that plays a critical role in potentiating Bacillus anthracis infection. Surprisingly, these compounds exhibited no differential change in activity with concentration. Docking studies revealed that the indole-attached amidine substituents of these inhibitors were positioned in close proximity to the biological zinc atom and could potentially function as transition-state mimetics. This broaches the stunning possibility that the dose independence of these inhibitors is linked to zinc-ligation. Because the amidine functionality is highly basic and cationic, it is generally not considered a viable zinc-binding motif. However, quantum chemical calculations on small-molecule models predicted a marked decrease in the pKₐ of the amidine functionality when it is in close proximity to zinc, thus allowing for the formation of a robust zinc-amidine bond.

Keywords: Three-dimensional search query; Density functional theory; Docking study; Molecular dynamics; Zinc coordination

1. Introduction

Anthrax lethal factor (LF) is a zinc-dependent endopeptidase and a constituent of the anthrax toxin secreted by Bacillus anthracis. LF cleaves near the amino terminus of mitogen-activated protein kinase kinases (MAPKK), resulting in disruptions in key cellular signaling pathways and ultimately cell death [1]. A significant number of small-molecule LF inhibitors have been described in the literature [2–10], and not surprisingly, most of these inhibitors are centered to zinc-binding groups such as hydroxamate [2] and thiazolidine [3]. However, not every LF inhibitor has its mechanism of inhibition depicted in its chemical structure [8,9,11]. Among these are a set of guanidinylated compounds [8,9], which were shown to be potent LF inhibitors but do not possess the prototypical zinc-binding groups [12].

In this paper, we report that a second set of cationic compounds are potent LF inhibitors. These compounds are carbamimidolyl-aryl-vinyl-indole-carboxamidines and are chemically similar to reported LF inhibitors [11]. The biological activity of these compounds was shown to be statistically dose independent. A detailed computational investigation revealed an unexpected potential mechanism of action for these highly basic and cationic compounds.
2. Computational methodology

The geometries of the [Zn(ac)(im)2(am)] model and the XIZZOV crystal structure were optimized with the Gaussian-03 package (Gaussian, Inc., Wallingford, CT) using the DFT B3LYP approach. The conventional 6-31+G(d,p) basis sets on all atoms (except Zn) and the recently published 6-31G(f) basis set on Zn [13] were utilized. The pKa values of the amidine/guanidine NH’s were calculated using a previously-described protocol [14] but in this instance, the hydration energy of the proton (\(\Delta G_{\text{hydr}}(\text{H}^+)\)) was taken as \(-264.0 \text{ kcal/mol}\), which is the most recently obtained value [15]. The resulting pKa values were corrected by the addition of 0.18 pKa units based on the result that the calculated pKa of the arginine-guanidine NH using this method was 12.32 whereas the textbook experimental pKa value is 12.5.

Docking studies were performed using the InsightII and Cerius2 programs on a Silicon Graphics Octane 2 workstation. The X-ray structure of LF in complex with FSPA (FSPA = (2R)-2-[(4-fluoro-3-methylphenyl)sulfonylamino]-N-hydroxy-2-(tetrahydro-2H-pyran-4-yl)acetamide) (PDB code 1YQY) [2] was selected as the template and optimized using a published tethered-minimization protocol [16]. The ligands were initially docked into the catalytic site using the Ligand Docking module in Cerius2 and their conformations were refined using iterative cycles of hydropathic analysis, manual adjustment, and molecular mechanics simulations [16].

3. Results and discussion

3.1. Empirical basis for LF inhibition by diaminides

We had previously developed a pharmacophore for LF inhibition [11]. Here we virtually screened the NCI chemical repository for compounds that fit our pharmacophore criteria. The conventional 6-31+G(d,p) basis sets on all atoms (except Zn) and the recently published 6-31G(f) basis set on Zn [13] were utilized. The pKa values of the amidine/guanidine NH’s were calculated using a previously-described protocol [14] but in this instance, the hydration energy of the proton (\(\Delta G_{\text{hydr}}(\text{H}^+)\)) was taken as \(-264.0 \text{ kcal/mol}\), which is the most recently obtained value [15]. The resulting pKa values were corrected by the addition of 0.18 pKa units based on the result that the calculated pKa of the arginine-guanidine NH using this method was 12.32 whereas the textbook experimental pKa value is 12.5.

Docking studies were performed using the InsightII and Cerius2 programs on a Silicon Graphics Octane 2 workstation. The X-ray structure of LF in complex with FSPA (FSPA = (2R)-2-[(4-fluoro-3-methylphenyl)sulfonylamino]-N-hydroxy-2-(tetrahydro-2H-pyran-4-yl)acetamide) (PDB code 1YQY) [2] was selected as the template and optimized using a published tethered-minimization protocol [16]. The ligands were initially docked into the catalytic site using the Ligand Docking module in Cerius2 and their conformations were refined using iterative cycles of hydropathic analysis, manual adjustment, and molecular mechanics simulations [16].

A HPLC-based enzymatic assay [11] was used to determine the activity of 1–4 against LF. From triplicate results, compounds 1–4 had average %LF inhibitions of 89–92% at 20 \(\mu\text{M}\), by 59–88% at 10 \(\mu\text{M}\), and by 21–69% at 5 \(\mu\text{M}\) (Table 1). At first glance, it appeared that the activity of these compounds was dose independent. This is exemplified by compound 1, which maintained \(\sim75\%\) of its activity, despite a fourfold decrease in concentration. Subsequent statistical analyses confirmed the dose independence of compounds 1–4 (see Supplementary Information). This typical pattern of activity warranted additional study of compounds 1–4.

The apparent K_i’s of 1–4 were determined (Table 1). Compound 1 had a K_i of 2.5 ± 0.07 \(\mu\text{M}\), compound 2 a K_i of 1.9 ± 0.07 \(\mu\text{M}\), compound 3 a K_i of 12 ± 0.78 \(\mu\text{M}\), and compound 4 a K_i of 3.5 ± 0.78 \(\mu\text{M}\). These K_i values are in line with the potencies of inhibitors containing zinc-binding groups. For instance, acetohydroxyacid, which has a simple molecular structure and contains a strong zinc-binding motif [12], is characterized by IC_{50} values of 15–40 mM [17] for different matrix metalloproteases. Dixon plot analysis [18] of the kinetic data of 1–4 revealed linear relationships between the inhibitor concentrations and the inverse of the maximum velocity, V (Supplementary Information). This is consistent with a model of compounds 1–4 binding at the LF active site.

3.2. Potential zinc-binding motif in the amidine functionality

The experimental results presented here suggest that compounds 1–4 bind at the LF active and are dose independent in their activities. This is in contrast to most drugs, which are typically reversible inhibitors that exhibit a dose dependence in their biological effect. The direct relationship between dose and the biological response for these drugs is due to the principle of mass action, that is, the greater the number of inhibitor molecules that are present at the active site of a receptor, the greater is the biological effect. In contrast, since the inhibitory activities of 1–4 do not directly correlate with their concentrations in solution, it would appear that compounds 1–4 do not necessarily follow the principle of mass action. This indicates that compounds 1–4 may not be reversible inhibitors. The irreversible binding of 1–4 at the LF active site would explain their dose independence, since a single molecule

<table>
<thead>
<tr>
<th>Compound</th>
<th>K_i ((\mu\text{M}))</th>
<th>%LF_i (20 (\mu\text{M}))</th>
<th>%LF_i (10 (\mu\text{M}))</th>
<th>%LF_i (5 (\mu\text{M}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.5 ± 0.07</td>
<td>92 ± 5.6</td>
<td>84 ± 3.5</td>
<td>69 ± 1.4</td>
</tr>
<tr>
<td>2</td>
<td>1.9 ± 0.07</td>
<td>92 ± 2.8</td>
<td>88 ± 1.4</td>
<td>67 ± 1.4</td>
</tr>
<tr>
<td>3</td>
<td>12 ± 0.78</td>
<td>89 ± 4.6</td>
<td>59 ± 4.2</td>
<td>21 ± 4.2</td>
</tr>
<tr>
<td>4</td>
<td>3.5 ± 0.78</td>
<td>90 ± 6.5</td>
<td>70 ± 0.7</td>
<td>49 ± 2.8</td>
</tr>
</tbody>
</table>

The %LF inhibition (%LF_i) data of 1–4 are the averages ± standard deviation for four replicates at a compound concentration of 20 \(\mu\text{M}\), and three replicates at 5 and 10 \(\mu\text{M}\). Additionally, the HPLC-based assay was used for enzyme kinetic studies. The K_i values of 1–4 were determined using at least six different concentrations of the inhibitor ranging from 3 to 30 \(\mu\text{M}\), and the standard deviations were determined from at least duplicate results.

Fig. 1. Chemical structures of compounds 1–4 with their NSC numbering given in parentheses.
would definitely inactive the enzyme and abrogate the direct relationship between inhibitor concentration and biological response. However, while this hypothesis is reasonable, compounds 1–4 do not possess identifiable zinc-binding groups [12].

Instead compounds 1–4 are characterized by a common indole-attached amidine moiety. While both the indole and amidine functionalities are not considered viable zinc-binding groups, this moieties is likely responsible for the potent activity of these compounds. There are several potential mechanisms for rationalizing the dose independence of 1–4. The indole-attached amidine group may be chemically labile and in solution, the amidine functionality is in equilibrium with a product that is a stronger Lewis base. Another possibility is that the amidine functionality itself may possess unrealized zinc-binding potential. To the authors’ knowledge, the Zn–N (amidine) coordinate bond has not been experimentally determined. However, there are examples in the literature of guanidine–metal ion bonding. There are 150 structures in the Cambridge Structural Database (CSD) that depict the guanidine–metal ion interaction [19]. One hundred and forty-six of these CSD structures involve a diguanidine moiety engaged in a metal chelation interaction via its anti-oriented sp2 Nη lone pair electrons. The four exceptions are the CSD structures RIP-DID [20], RIPDOJ [20], ALOSUP [21], and XIZZOV [22], which have an individual guanidine functionality coordinated to metal ions via its syn-oriented sp2 Nη lone pair electrons. These four structures may offer a paradigm for the arginine–metal ion interaction in biological systems. In fact, three protein crystal structures have been determined that depict the 1-arginine–guanidine–metal ion coordinate interaction (PDB codes 1R30 [23], 1JV0 [24], and 3CEV [25]). In the 1R30, 1JV0, and 3CEV structures, the 1-arginine-guanidine moieties are coordinated to Fe2+, Mn2+, respectively. Since guanidine is characterized by a higher pKa than amidine, the guanidine–zinc interactions in these small-molecule complexes and biological systems provide a sound structural and chemical basis for amidine–zinc bonding.

3.3. Zn–guanidine bonding as the archetype

An archetype for the Zn–amidine bond is the small-molecule model composed of Zn(II) bound to Cy* (Cy* = (2-guanidinyl)ethyl-cyclen) investigated by Aoki et al. (CSD code XIZZOV) [22]. In a comprehensive biochemical and crystallographic study, Aoki et al. [22] showed that the pendant guanidine group ligated Zn to form a robust Zn–N(guanidine) bond in a neutral pH aqueous solution. The zinc complex was crystallized at pH 7.5 and its crystal structure was determined to an R-factor of 2.94 [22]. The crystal structure revealed a five-coordinate zinc atom that was bonded to the Nη1(eta) atom of guanidine with a bond distance of 1.95 Å. Shown in Fig. 2a, a molecular model was created from the XIZZOV (Zn(II)–Cy*) crystal structure, and its geometric parameters were optimized using the DFT B3LYP approach. As evident by an r.m.s. deviation of only 0.07 Å, the DFT-refined XIZZOV model and the starting crystal structure are structurally similar. Notably, the Zn–N(guanidine) bond distance in the DFT-optimized structure is 1.98 Å, which is only 0.03 Å longer than that in the XIZZOV crystal structure. This incremental change in the Zn–N(guanidine) bond distance during the DFT simulations is consistent with the robustness of the Zn–N(guanidine) bonding experimentally demonstrated by Aoki et al. [22].

Concomitantly, a molecular model of the Zn–amidine bond was constructed. The XIZZOV crystal structure was used as a template for the Zn–amide model. Since an indole-attached amidine moiety is common to compounds 1–4, this small-molecule model system consisted of Zn(II) coordinated to acidic acid, two imidazoles, and 1H-Indole-6-carboxamidine. This model is termed [Zn(ac)(im)2(am)]. Similarly, the geometry of the [Zn(ac)(im)2(am)] model was refined using DFT simulations at the B3LYP level (Fig. 2b). The resulting [Zn(ac)(im)2(am)] structure exhibited a strong structural correlation to the XIZZOV DFT model (Fig. 2). In the two structures, the amidine and guanidine groups have similar Zn–N–C bond angles of 130.6° and 138.8°, respectively. These bond angles are indicative of strong Zn–N(amidine or guanidine) interactions. Additionally, the amidine and guanidine have characteristic short bond distances to the zinc atom, which is again consistent with strong Zn–N(amidine or guanidine) interactions. The Zn–N(amidine) and Zn–N(guanidine) bond distances are 2.02 and 1.98 Å, respectively.
3.4. Novel chemical property of amidine in the presence of zinc

While these quantum chemical calculations predicted a robust Zn–N(amide) bond, the question remains as to how a cationic and highly basic group such as amidine could ligate zinc? Amidines exist predominantly in the protonated amidinium form and therefore are poor Lewis bases. The answer may be provided by experiments performed on the Zn–Cy* model system [22]. Aoki et al. [22] had observed that the pKₐ of unbound guanidine decreased from 12.4 to 5.9 when it was zinc-bound. This pointed to a facile deprotonation of the guanidinium ion in the presence of zinc and allowed for zinc coordination [22].

Postulating that zinc-ligation by the amidine functionality may occur by a similar mechanism, we calculated the amidine pKₐ values in two atomic models. The first model was [Zn(ac)(im)₂(am)] and the second was the 1H-indole-6-carboxamidine constituent by itself. The calculated amidine pKₐ's of the unligated 1H-indole-6-carboxamidine constituent and of the [Zn(ac)(im)₂(am)] model were 12.47 and 8.37, respectively. These calculations indicate a decrease of 4.1 pKₐ units when the amidine group is zinc-bound. Additionally, since it has been established that the hydrophobic environment of an enzyme can contribute to a pKₐ decrease of 2–3 units [26], the amidine pKₐ's of 1–4 may have effectively been decreased by 6.1–7.1 pKₐ units in the LF active. This marked decrease in pKₐ units indicates a facile deprotonation of the amidinium ion in the LF active site, thus establishing a mechanistic basis for zinc-ligation by amidine.

3.5. Structural basis for dose independence

While quantum chemical methods are powerful analytical tools, in this instance, they can hardly be used to delineate the structural basis for zinc coordination by the amidine groups in the full model system. As a consequence, docking studies and molecular dynamics simulations were performed on compounds 1–4 to elucidate in atomic detail their mechanism of inhibition. To establish a template for these docking studies, a molecular model of LF complexed with the KPVLPA sequence, which represents the P5–P1’ site of the native substrate MEK2 [27], was constructed. The MEK2 sequence was modeled in a transition state [28] with the Pro10-Ala11 amide cleavage site [27] positioned for nucleophilic attack by the zinc-bound water (Fig. 3a). In this conformation, the Pro10 carbonyl is zinc polarized and forms a hydrogen bond to the Tyr728 phenol. Additionally, hydrogen bonds are formed between: (1) the backbones of LF Gly657 and MEK2 Ala11, (2) the backbones of LF Tyr659 and MEK2 Leu9, and (3) the side chains of LF Glu662 and MEK2 Lys6.

Compounds 1–4 were docked in the LF active in their protonated amidinium forms. Fig. 3b shows the binding model of 3, which is representative of the set. The docked poses of 1–4 exhibit a strong structural correlation with the proposed binding mode of MEK2. In each instance, the indole-attached amidinium ion is a bioisostere of the Pro10-Ala11 amide cleavage site, the indole aromatic ring itself is a bioisostere of the Pro10 side chain, the second aromatic ring is a bioisostere of the Pro7 side chain, and the second amidinium group is a bioisostere of the Lys6 side chain amino. The chemical structures of 1–4 are characterized by a distance range between the two amidinium groups of 14.0–15.7 Å, and as revealed by the binding models, this metric could be mapped onto two electronegative regions on the LF surface, specifically the separate cationic sinks formed by the zinc-bound hydroxide and the carboxylate of Glu662.

To delineate the chemical feasibility of zinc-ligation by 1–4 in the LF active site, zinc-coordination models of 1–4 were generated from the nonbonded binding models. During molecular dynamics simulations, the atoms of LF were held fixed in Cartesian space, and the distances between the
inhibitor amidine nitrogen atom and the zinc atom were incrementally decreased using distance constraints to 2.0 Å, which is the coordinate covalent bond distance observed in the B3LYP level models. Fig. 3c shows the coordination model of compound 3, which is representative of the set. The hydropathic quality of the initial nonbonded binding model and the coordinate Zn–amidine model were compared. Since the ratios of favorable to unfavorable intermolecular interactions in the two models were similar, this suggests that the stereoelectronic features of the LF active site are complementary to zinc-ligation by the amidine groups.

3.6. Putative molecular mechanism of inhibition

Based on these atomic models, we can propose a molecular mechanism of inhibition involving the amidinium ion (Scheme 1). Binding of 1–4 to LF results in the formation of a hydrogen bonding network between the ligands and the catalytic engine. The indole-attached amidinium ion is hydrogen bonded to the phenol of Tyr728, and the zinc-bound water molecule, which itself is hydrogen bonded to the carboxylate of Glu687 (Scheme 1, panel I). The zinc-bound water molecule is deprotonated by Glu687 to give the reactive hydroxide ion. In lieu of the nucleophilic attack of the substrate, the hydroxide anion encounters and deprotonates the amidinium ion (Scheme 1, panel II). The resulting neutral amidine functionality is a markedly better Lewis base than the protonated amidinium ion, and is capable of ligating the zinc atom. In a potentially concerted mechanism, the neutral amidine functionality displaces the catalytic water molecule by mass action and coordinates the zinc atom (Scheme 1, panel III). The end result is a stable Zn–amidine complex with the indole-attached amidinium ion hydrogen bonded to the phenol of Tyr728 (Scheme 1, panel IV).

4. Conclusion

Four new potent LF inhibitors have been identified from the NCI chemical repository. The LF activities of these carbamimidolyl-aryl-vinyl-indole-carboxamidines were shown to be statistically dose independent. Because this dose independence may be a manifestation of zinc-ligation by these inhibitors, we employed a detailed computational study to elucidate their molecular mechanism of inhibition. The Zn–guanidine bond determined by Aoki et al. [22] was selected as the archetype of the Zn–amidine bond. Quantum chemical calculations predicted a stable Zn–N(amidine) bond with an atom–atom distance of 2.02 Å and Zn–N–C bond angle of 138.8°. Additionally, a calculated pKₐ decrease of 4.1–5.1 units for the amidine functionality in the presence of zinc established the mechanistic basis for zinc-ligation by a highly basic group such amidine that would otherwise to considered inert to zinc. Concomitantly, docking studies and molecular dynamics simulations delineated in atomic detail the structural basis for zinc-ligation by the amidines of 1–4 in the full LF active site.

Acknowledgments

The research described herein was sponsored by the US Army Medical Research and Material Command Research Plan #02-4-3U-057 and IAA #Y3-CM-100505 (MRMC and NCI). We acknowledge the National Cancer Institute.
for providing the compounds and for the allocation of computing time and staff support at the Advanced Biomedical Computing Center. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract N01-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. This research was supported in part by the Developmental Therapeutics Program in the Division of Cancer Treatment and Diagnosis of the National Cancer Institute.

Appendix A. Supplementary data

Statistical analysis and Dixon plots of assay data. Coordinates of the binding models of LF complexed with MEK2 and 3 NSC294200 before and after zinc coordination as well as coordinates for the small molecule models of XIZZOV and [Zn(ac)(im)2(am)].

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.theochem.2007.07.009.

References