Fragment-Based Drug Design Facilitated by Protein-Templated Click Chemistry: Fragment Linking and Optimization of Inhibitors of the Aspartic Protease Endothiapepsin


Abstract: There is an urgent need for the development of efficient methodologies that accelerate drug discovery. We demonstrate that the strategic combination of fragment linking/optimization and protein-templated click chemistry is an efficient and powerful method that accelerates the hit-identification process for the aspartic protease endothiapepsin. The best binder, which inhibits endothiapepsin with an IC\textsubscript{50} value of 43 \( \mu \text{m} \), represents the first example of triazole-based inhibitors of endothiapepsin. Our strategy could find application on a whole range of drug targets.

Despite recent developments in medicinal chemistry, there is a continuous need for the development of more efficient, rapid, and facile strategies to accelerate the drug-discovery process. In recent decades, fragment-based drug design (FBDD) has emerged as an effective and novel paradigm in drug discovery for numerous biological targets.[1–3] FBDD has higher hit rates and better coverage of the chemical space, enabling the use of smaller libraries than those used for high-throughput screening.[25] Since the first report of FBDD, it started to be more widely used in the mid-1990s[26] and has since expanded rapidly. Over the course of the past two decades, various pharmaceutical and biotechnology companies have used FBDD and developed more than 18 drugs that are currently in clinical trials.[27] Upon identification of a fragment,[28] it has to be optimized to a hit/lead compound and eventually to a drug candidate by fragment growing, linking, merging, or optimization. On the one hand, fragment growing has become the optimization strategy of choice,[22–24] even though it is time consuming because it requires synthesis and validation of the binding mode of each derivative in the fragment-optimization cycle. To overcome this hurdle, we have previously developed strategies in which we combined fragment growing with dynamic combinatorial chemistry (DCC) to render the initial stage of the drug-discovery process more effective.[13] Fragment linking, on the other hand, is very attractive because of its potential for superadditivity (an improvement of ligand efficiency (LE) and not just maintenance of LE), but challenging as it requires the preservation of the binding modes of the individual fragments in adjacent pockets and identification of the best linker with an ideal fit.[14,15] It is presumably due to these challenges that there are only few reports of fragment linking,[4,16] demonstrating the efficiency of linking low-affinity fragments to higher-affinity binders.[17–24] We have recently reported a combination of DCC and fragment linking/optimization, which reduces the risks associated with fragment linking.[22]

In addition to DCC, protein-templated click chemistry (PTCC) has emerged as a powerful strategy to design/optimize a hit/lead for biological targets and holds the potential to reduce the risks associated with fragment-linking.[26,27] PTCC relies on the bio-orthogonal 1,3-dipolar cycloaddition of azide and alkyne building blocks facilitated by the protein target.[28] This highly exothermic reaction produces 1,4- and 1,5-triazoles, which are extremely stable under acidic/basic pH as well as in harsh oxidative/reductive conditions. Furthermore, triazoles can participate in H-bonding, \( \pi-\pi \)-stacking, and dipole–dipole interactions with the target protein and are a bioisostere of amide bonds. In PTCC, the individual azide and alkyne fragments bind to adjacent pockets of the protein and if the functional groups are oriented in a proper manner, the protein “clicks” them together to afford its own triazole inhibitor (Figure 1). We have therefore envisaged that the potentially synergistic combination of fragment linking and PTCC would represent an efficient hit/lead identification/optimization approach in medicinal chemistry. Here, we have combined fragment linking and PTCC by designing flexibility into the linker and letting the protein select the best combination of building blocks to identify a new class of hits for endothiapepsin, belonging to the pepsin-like aspartic proteases.

Aspartic proteases are a family of enzymes that are widely found in fungi, vertebrates, and plants, as well as in HIV retro-
Owing to its high degree of similarity to aspartic proteases, endothiapepsin has served as a model enzyme for mechanistic studies as well as for the identification of inhibitors of renin and β-secretase. Endothiapepsin is a robust enzyme, is available in large quantities, crystallizes easily, and remains active at room temperature for more than three weeks, making this enzyme a convenient representative for aspartic proteases. All aspartic proteases consist of two structurally similar domains, which contribute an aspartic acid residue to the catalytic dyad that is responsible for the water-mediated cleavage of the substrate’s peptide bond. Although the linkage of two known inhibitors of acetylcholinesterase via a triazolyl linker using PTCC has been reported, the inhibitors that are linked do not qualify as fragments. To the best of our knowledge, there is no report of fragment linking using PTCC. Herein, we describe how we combined fragment linking/optimization and PTCC for the efficient fragment-to-hit optimization of inhibitors of the aspartic protease endothiapepsin.

We used X-ray crystal structures of endothiapepsin in complex with fragments 1 and 2 (Protein Data Bank (PDB) codes: 3PBZ and 3PLD, respectively, Figure 2), identified by Klebe and co-workers. Both 1 and 2 are engaged in strong H-bonding interactions with the catalytic dyad consisting of amino acid residues D35 and D219, using their hydrazide and amidine groups, respectively (Figure 2). Except for the number of H-bond acceptors (four) for 1, both fragments 1 and 2 obey Astex’s “rule of three,” with a molecular weight (Mw) of 207 and 201 Da, three H-bond donors, and two H-bond acceptors, two freely rotatable bonds and total polar surface areas (TPSAs) of 58.4 and 49.9 Å², respectively. At a concentration of 1 mM, fragments 1 and 2 display 89 and 84% inhibition of endothiapepsin, respectively. Considering their promising physicochemical properties, inhibitory potency, their small size (15 and 12 heavy atoms, respectively) and the fact that they bind to adjacent pockets of endothiapepsin, we chose them as a starting point for fragment linking/optimization into an inhibitor of endothiapepsin.

Figure 2. X-ray crystal structure of endothiapepsin in complex with fragments 1 and 2 (PDB code: 3PBZ and 3PLD, respectively) and a modeled potential triazole inhibitor in the active site. Color code: protein skeleton: C: gray, O: red, and N: blue; fragment skeleton: C: purple, yellow and green, N: blue, O: red, Cl: green. Hydrogen bonds below 3.0 Å are shown as black, dashed lines.

Figure 1. Schematic representation of protein-templated click chemistry leading to a triazole-based inhibitor starting from a library of azides and alkynes.

Viruses. This class of enzymes plays a causative role in several important diseases such as malaria, Alzheimer’s disease, hypertension, and AIDS. Owing to its high degree of similarity with these drug targets, endothiapepsin has served as a model enzyme for mechanistic studies as well as for the identification of inhibitors of renin and β-secretase. Endothiapepsin is a robust enzyme, is available in large quantities, crystallizes easily, and remains active at room temperature for more than three weeks, making this enzyme a convenient representative for aspartic proteases. All aspartic proteases consist of two structurally similar domains, which contribute an aspartic acid residue to the catalytic dyad that is responsible for the water-mediated cleavage of the substrate’s peptide bond.

Although the linkage of two known inhibitors of acetylcholinesterase via a triazolyl linker using PTCC has been reported, the inhibitors that are linked do not qualify as fragments. To the best of our knowledge, there is no report of fragment linking using PTCC. Herein, we describe how we combined fragment linking/optimization and PTCC for the efficient fragment-to-hit optimization of inhibitors of the aspartic protease endothiapepsin.

We used X-ray crystal structures of endothiapepsin in complex with fragments 1 and 2 (Protein Data Bank (PDB) codes: 3PBZ and 3PLD, respectively, Figure 2), identified by Klebe and co-workers. Both 1 and 2 are engaged in strong H-bonding interactions with the catalytic dyad consisting of amino acid residues D35 and D219, using their hydrazide and amidine groups, respectively (Figure 2). Except for the number of H-bond acceptors (four) for 1, both fragments 1 and 2 obey Astex’s “rule of three,” with a molecular weight (Mw) of 207 and 201 Da, three H-bond donors, and two H-bond acceptors, two freely rotatable bonds and total polar surface areas (TPSAs) of 58.4 and 49.9 Å², respectively. At a concentration of 1 mM, fragments 1 and 2 display 89 and 84% inhibition of endothiapepsin, respectively. Considering their promising physicochemical properties, inhibitory potency, their small size (15 and 12 heavy atoms, respectively) and the fact that they bind to adjacent pockets of endothiapepsin, we chose them as a starting point for fragment linking/optimization into an inhibitor of endothiapepsin.

Figure 2. X-ray crystal structure of endothiapepsin in complex with fragments 1 and 2 (PDB code: 3PBZ and 3PLD, respectively) and a modeled potential triazole inhibitor in the active site. Color code: protein skeleton: C: gray, O: red, and N: blue; fragment skeleton: C: purple, yellow and green, N: blue, O: red, Cl: green. Hydrogen bonds below 3.0 Å are shown as black, dashed lines.

Figure 1. Schematic representation of protein-templated click chemistry leading to a triazole-based inhibitor starting from a library of azides and alkynes.

3PBLZ and 3PLD, respectively, Figure 2), identified by Klebe and co-workers. Both 1 and 2 are engaged in strong H-bonding interactions with the catalytic dyad consisting of amino acid residues D35 and D219, using their hydrazide and amidine groups, respectively (Figure 2). Except for the number of H-bond acceptors (four) for 1, both fragments 1 and 2 obey Astex’s “rule of three,” with a molecular weight (Mw) of 207 and 201 Da, three H-bond donors, and two H-bond acceptors, two freely rotatable bonds and total polar surface areas (TPSAs) of 58.4 and 49.9 Å², respectively. At a concentration of 1 mM, fragments 1 and 2 display 89 and 84% inhibition of endothiapepsin, respectively. Considering their promising physicochemical properties, inhibitory potency, their small size (15 and 12 heavy atoms, respectively) and the fact that they bind to adjacent pockets of endothiapepsin, we chose them as a starting point for fragment linking/optimization into an inhibitor of endothiapepsin.

Figure 2. X-ray crystal structure of endothiapepsin in complex with fragments 1 and 2 (PDB code: 3PBZ and 3PLD, respectively) and a modeled potential triazole inhibitor in the active site. Color code: protein skeleton: C: gray, O: red, and N: blue; fragment skeleton: C: purple, yellow and green, N: blue, O: red, Cl: green. Hydrogen bonds below 3.0 Å are shown as black, dashed lines.

Figure 1. Schematic representation of protein-templated click chemistry leading to a triazole-based inhibitor starting from a library of azides and alkynes.
we synthesized alkyne... 2016 value of 43.

\[ m \] =

The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

insoluble –– (100 \( \% \))

no inhibition – – –

Structure of the triazoles (17–20) identified using PTCC, inactive triazole 21.

To investigate the biochemical activity of the binders identified by PTCC, we synthesized all four triazoles from their corresponding azide and alkyne precursors using the Cu\(^{1+}\)-catalyzed 1,3-cycloaddition (Schemes S3–S6 in the Supporting Information). In addition, we synthesized an inactive triazole 21 to demonstrate the efficiency of PTCC (Figure 3 and Scheme S7 in the Supporting Information). We determined their inhibitory activity using a fluorescence-based assay adapted from the HIV-protease assay.\(^{46}\)

The enzyme-activity assay confirmed the result of the PTCC experiment. Three out of the four triazoles indeed inhibit endo-thiapsin with IC\(_{50}\) values in the range of 43–121 \( \mu \)M (Figures S10–S12 in the Supporting Information). We were unable to determine the IC\(_{50}\) value of 20 because of its poor solubility even at 250 \( \mu \)M using the maximum possible DMSO concentration for the assay. The inactive triazole 21, which was not observed in the PTCC but synthesized as a control, did not show any activity in the enzyme-activity assay. The most potent triazole inhibitor 17 displays an IC\(_{50}\) value of 43 \( \mu \)M (Table 1). The experimental Gibbs free energies of binding (\( \Delta G \)) and ligand efficiencies (LE), derived from the experimental IC\(_{50}\) values using the Cheng–Prusoff equation,\(^{47}\) correlate with the calculated values using the scoring function HYDE in the LeadIT.

![Scheme 1](image-url)

**Scheme 1.** a) Structures and retrosynthetic analysis of the designed triazole inhibitors starting from fragments 1 and 2; b) structures of the azides 3–11 and the alkynes 12–15.

![Figure 3](image-url)

**Figure 3.** Structure of the triazoles (17–20) identified using PTCC, inactive triazole 21.

| Table 1. | The IC\(_{50}\) values, ligand efficiency (LE), calculated and experimental Gibbs free energies of binding (\( \Delta G \)) of triazole inhibitors. |
|---|---|---|---|---|
| Inhibitors | IC\(_{50}\) \([\mu \text{M]}\) | \( \Delta G_{\text{EXPT}} \) \([\text{kJ mol}^{-1}]\) | EC\(_{50}\) \([\text{kMol}^{-1}]\) | \( \Delta G_{\text{HYDE}} \) \([\text{kJ mol}^{-1}]\) |
| 17 | 43 ± 0 | −27 | 0.25 | −25 |
| 18 | 94 ± 18 | −25 | 0.26 | −19 |
| 19 | 121 ± 3 | −24 | 0.22 | −25 |
| 20 | insoluble | – | – | – |
| 7 | no inhibition | – | – | – |
| 14 | no inhibition | – | – | – |
| 12 | 142 ± 52 | – | – | – |

[a] 26 experiments were performed and only initial six experiments were considered to calculate the initial slope (\( n = 6 \)); 11 different concentrations of inhibitor were used, starting at 1 nM; each experiment was carried out in duplicate and the errors are given in standard deviations (SD), [b] The Gibbs free energy of binding (\( \Delta G_{\text{EXPT}} \)) and the ligand efficiencies (LEs) derived from the experimentally determined IC\(_{50}\) values, [c] Values indicate the calculated Gibbs free energy of binding (\( \Delta G_{\text{HYDE}} \)) calculated by the HYDE scoring function in the LeadIT suite. |
suite ($\Delta G_{\text{prot}}$) = $-25$ kJ mol$^{-1}$, Table 1). This correlation is also valid for the other triazole inhibitors (Table 1).

To validate the predicted binding mode from fragment linking, we tried to soak crystals of endothiapepsin with the most potent triazole inhibitor 17. Due to limited solubility, we were not able to obtain crystals of 17 with endothiapepsin. Based on the inhibitory potencies, replacement of $-\text{Cl}$ in 19 with a $-\text{OH}$ group in 17, leads to a decrease in $IC_{50}$ value from 121 to 43 $\mu$M. This result indicates that the $-\text{OH}$ group is involved in more favorable interactions than $-\text{Cl}$, which could be due to the H-bonding interaction with I300 in the S2 pocket, as illustrated by modeling studies (Figure 4a, and Figure S1 in the Supporting Information). Moreover, the alkyne 12 displays an $IC_{50}$ value of 142 $\mu$M (Figure S13 in the Supporting Information) and is present in both 17 and 19, two identified triazoles. Fragment 12 is a privileged fragment for endothiapepsin and most probably the binding mode of 12 is retained in both 17 and 19.

According to modeling and docking, as shown in Figure 4a, and Figure S1, respectively, both 17 and 19 address the catalytic dyad using their triazolyl linker to form direct H-bonds with D35. The NH group of both compounds is involved in several hydrophobic contacts with I300, I302, I304, F194, and I217, maintaining the binding mode of fragment 2.

Triazole 18 displays an $IC_{50}$ of 94 $\mu$M and (S)-18 addresses the catalytic dyad using its triazolyl linker to form a direct H-bond with D35, as indicated by modeling and docking studies (Figure 4b). The $-\text{NH}$ group of the triazole is engaged in H-bonding interactions with D33 and G221. Both phenyl substituents of the triazole (S)-18 occupy the S3 and S2 pockets and are involved in hydrophobic interactions with F116, I122, and L125 in the S3 pocket, and I300, I302, I304, F194, and I217 in the S2 pocket, which preserve the binding mode of fragments 1 and 2, respectively.

In conclusion, we have demonstrated for the first time that the strategic combination of fragment linking/optimization and PTCC is an efficient and powerful method that accelerates the hit-identification process for the aspartic protease endothiapepsin. We have exploited the sensitive UPLC-TOF-SIM method to identify the triazole binders templated by the protein. The best binder inhibits endothiapepsin with an $IC_{50}$ value of 43 $\mu$M. Due to the limited solubility of the triazoles identified, we were unable to obtain crystals of any triazole in complex with endothiapepsin. We have reported the first example of triazole-based inhibitors of endothiapepsin. The advantage of our approach is that, a catalytic amount of protein is sufficient to initiate and accelerate triazole formation from a sufficiently large library. Our strategic combination of methodologies proved to be very successful for the hit identification for the aspartic protease endothiapepsin and could be applied to a wide range of biological targets. It could be used in the early stages of drug development and holds the potential to greatly accelerate the drug-discovery process.

Acknowledgements

We would like to acknowledge T.D. Tiemersma-Wegman for help with the UPLC-TOF-SIM measurements, Prof. G. Klebe and N. Radeva for fruitful discussions and help with protein crystallization experiments, and Syncom for providing alkynes and PTCC is an efficient and powerful method that accelerates the hit-identification process for the aspartic protease endothiapepsin. Funding was granted by the Netherlands Organisation for Scientific Research (NWO-CW, ECHO-STIP grant to A.K.H.H.) and by the Dutch Ministry of Education, Culture, Science (gravitation program 024.001.035).

Keywords: click chemistry · drug design · enzymes · inhibitors · liquid chromatography

References


Received: June 23, 2016
Published online on September 7, 2016