European Journal of Medicinal Chemistry 96 (2015) 405–424

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Binding cooperativity between a ligand carbonyl group and a hydrophobic side chain can be enhanced by additional H-bonds in a distance dependent manner: A case study with thrombin inhibitors

Ahmed M. Said*, David G. Hangauer*

Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY 14260, USA

ARTICLE INFO

Article history:
Received 10 November 2014
Received in revised form
21 March 2015
Accepted 25 March 2015
Available online 27 March 2015

Keywords:
Cooperativity
Protein-ligand binding
Ligand binding affinity
Scoring functions
B-factor
Thrombin
Non-additivity

ABSTRACT

One of the underappreciated non-covalent binding factors, which can significantly affect ligand-protein binding affinity, is the cooperativity between ligand functional groups. Using four different series of thrombin inhibitors, we reveal a strong positive cooperativity between an H-bond accepting carbonyl functionality and the adjacent P3 hydrophobic side chain. Adding an H-bond donating amine adjacent to the P3 hydrophobic side chain further increases this positive cooperativity thereby improving the K_i by as much as 546-fold. In contrast, adding an amidine multiple H-bond/salt bridge group in the distal S1 pocket does not affect this cooperativity. An analysis of the crystallographic *B*-factors of the ligand groups inside the binding site indicates that the strong cooperativity is mainly due to a significant mutual reduction in the residual mobility of the hydrophobic side chain and the H-bonding functionalities that is absent when the separation distance is large. This type of cooperativity is important to encode in binding affinity prediction software, and to consider in SAR studies.

1. Introduction

Reversible ligand-protein binding involves multiple weak non-covalent interactions, such as hydrogen bonding and hydrophobic interactions between the ligand and its protein host. However, this process is complex in that it also involves enthalpy, entropy, desolvation and other important water structural changes as the ligand departs from the bulk solution and binds in the protein site [1–9]. Scoring functions have been widely used by medicinal and computational chemists within structure-based-design technologies in an attempt to predict the effect of these non-covalent interactions on ligand binding affinity. These scoring functions can be classified as physics-based, empirical, knowledge-based, or descriptor-based due to the differing overall methods employed for predicting ligand binding affinity [10–15]. A recent comparative assessment and an evaluation carried out by Yan Li et al. [16] on the accuracy of binding affinity predictions made by several scoring functions concluded that the reliability of these predictions is generally limited because they do not properly address the non-additive features among protein-ligand complexes in their algorithms. Stahl et al. [17] has recently reviewed a number of the factors that influence the binding free energy of the ligand to its target. A number of these factors are still poorly understood, yet are crucial to the ligand-macromolecule binding phenomenon. Factors in need of further investigation include desolvation/resolvation, the hydrophobic effect, enthalpy-entropy compensation, protein and ligand conformational changes, and the non-additivity/cooperativity between the ligand's functional groups. Many of these crucial factors are not included, or inadequately so, in the algorithms used in the scoring functions.

Cooperativity (non-additivity) was initially described by Williams and coworkers [18, 19] as a phenomenon whereby multiple guest-host binding interactions act together to produce a binding

Abbreviations: Me, methyl; Bn, benzyl; *Boc*, *tert*-butoxycarbonyl; ΔG, standard Gibbs free energy; Ki, inhibition constant; EDCI, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride; HOAT, 1-Hydroxy-7-azabenzotriazole; DIEA, diisoproplyethylamine; HATU, 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo [4,5-b]pyridinium3-oxid hexafluorophosphate; $LiAlH_4$, lithium aluminum hydride; tPa, tissue plasminogen activator; DMF, N,N-dimethylformamide; DCM, dichloromethane; DCE, dichloroethane; TEA, triethylamine; HPLC, high performance liquid chromatography; ^{1}H NMR, proton nuclear magnetic resonance; ^{13}C NMR, carbon nuclear magnetic resonance; HRMS, high resolution mass spectrometry; sd, standard deviation.

* Corresponding authors.
E-mail addresses: ahmedmoh@buffalo.edu (A.M. Said), hangauer@buffalo.edu (D.G. Hangauer).

http://dx.doi.org/10.1016/j.ejmech.2015.03.059
0223-5234/

affinity stronger than (positive cooperativity) or weaker than (negative cooperativity) the sum of these individual interactions. More recent studies have begun to elucidate the concept of cooperativity as applied to ligand-protein systems, and in significant depth [20–23]. Further studies of this type, with a number of ligand-protein systems, are needed in order to provide a range of experimental data and fundamental insights from which one can develop general rules that can be applied to the SAR and computational prediction of ligand-protein binding affinity. Recognizing the need to include cooperativity in scoring functions, Kuhn and coworkers [24] have started to address this significant deficiency. The value utilized across a number of scoring functions to the energy contribution of hydrophobic interactions per each square angstrom of the hydrophobic contact surface area is 0.1 [(kJ/mol)/Å^2] [25]. However this single value is not generally accurate because it ignores the effect of the neighboring groups. The range of Å^2 hydrophobic contributions in the current study was from 0.0367 to 0.2285 kJ/mol, about a 6-fold span. This variation in hydrophobic contribution to binding accumulates across the full side chain contact surface area, and consequently can be very significant, such that using a single value will lead to substantial error. Herein, we reveal that if a hydrophobic group is mutually reinforced by one adjacent H-bond, on either side, the contribution of each square angstrom hydrophobic contact surface area is 2–3 times greater than in absence of this H-bond. If the hydrophobic side chain is further reinforced by a second immediately adjacent H-bonding group the contribution of each square angstrom hydrophobic contact surface area is doubled again to be 6 times greater than in absence of these H-bonds. This can result in a substantial change in the K_d for the ligand. For example, suppose a hydrophobic side chain would contribute 5.7 kJ/mol to the ligand binding free energy in the absence of cooperativity with an adjacent H-bond, and the per Å^2 hydrophobic binding increases by 2-fold to 11.4 kJ/mol with the cooperativity included. The ligand binding affinity contribution of the hydrophobic side chain (i.e. K_d) is then greater by 10-fold than what would have been predicted without considering cooperativity. An increase per square Å^2 hydrophobic contribution by 3-fold to 17.1 kJ/mol would result in a 100-fold stronger binding than would have been predicted in the absence of cooperativity. The magnitude of these changes in the K_d's clearly needs to be considered in designing and interpreting SAR studies, as well as in developing more accurate scoring functions. The current study adds to the body of experimental data needed to continue developing computational binding affinity predictions, and provides qualitative fundamental insights for medicinal chemists to utilize in SAR ligand optimization programs.

Thermodynamic double functional group replacement cycles can be used to evaluate ligand functional group's cooperativity [26–31]. This can be done by comparing the binding free energy change when the ligand has the two new functional groups with the sum of the binding free energy changes occurring when each new group is present individually. If the difference is significantly more favorable (i.e. more negative free energy change) the simultaneous presence the two groups synergistically (positive cooperativity) improves the ligand binding affinity. On the other hand, if the difference is significantly less favorable then the two groups are engaged in negative cooperativity. If the difference is zero there is no cooperativity (additive concept) between the two groups.

The current study reveals new aspects of ligand functional group's cooperative contributions to the binding free energy. The serine protease thrombin is used as the protein host model system, mainly because of both of its structural and therapeutic properties. Thrombin functions in the coagulation cascade. One of these functions is to convert soluble fibrinogen into insoluble strands of fibrin. Some other thrombin functions are the activation of the coagulation factors and the activation of platelet aggregation. Due to its important physiological role, thrombin has been an interesting target for medicinal chemists and many thrombin inhibitors have been discovered over the last few decades. Thrombin has a rigid and well-defined binding site with numerous deposited x-ray crystal structures with a variety of bound inhibitors and is considered one of the best-characterized enzymes. The active site of the thrombin consists of three pockets [32]: (1) The S1 pocket which is deep and consists of a hydrophobic channel with the carboxylate of Asp189 and two backbone carbonyls at the bottom of the pocket. This carboxylate anion is responsible for strong ionic and bifurcated H-bonding interactions with positively charged ligand amidine containing P1 residues; (2) The S2 hydrophobic pocket which mainly consists of Tyr60A and Trp60D side chains of the 60-insertion loop as well as the isobutyl group of Leu99; and (3) The S3-pocket which is a well-defined hydrophobic pocket [33], consisting mainly of the side chains of Trp215, Ile174, and Leu99. The thrombin inhibitors described herein have structural similarities to melagatran, which is used as an anticoagulant, and thus provide further insights into this class of drug molecules.

This current study mainly focuses on answering two unique questions which have yet to be addressed in any study that we are aware of. First, can a ligand carbonyl group, acting as a hydrogen bond acceptor with the protein, engage in positive cooperativity with a nearby hydrophobic side chain? As a corollary to this question, if this is the case, what will be the effect of that on the magnitude of the energy contribution of each square angstrom of hydrophobic contact surface area, and how can these results be added to scoring functions and SAR studies so as to improve more accurately predict ligand binding affinity? Second, can cooperativity be further enhanced by an additional ligand H-bonding group and will this occur in a distance dependent manner?

To answer these questions, four series of structurally related thrombin inhibitors were designed. The structural features and the expected interactions of the ligands with the thrombin binding site are shown in Fig. 1.

As shown in Fig. 1, the scaffold for the designed thrombin inhibitors includes a proline which binds in the S2 pocket, fitting under the Tyr60A and Trp60D of the 60-loop (similar to the natural substrate) [34]. The difference between the two categories (Fig. 1a and b) of inhibitors is that series I and II (Fig. 1a) have a *m*-chlorobenzyl moiety binding in the S1 pocket whereas series III and IV (Fig. 1b) have a more firmly bound (salt bridge and three H-bonds) benzamidine moiety. Within each of these categories the size of the P3 hydrophobic side chain was gradually increased from a methyl to a benzyl side chain, in presence or absence of the carbonyl group (blue box) that is H-bonded to the amino group of Gly216 residue. In order to investigate the effect of the presence or absence of an additional adjacent hydrogen bonding moiety on the cooperativity between the carbonyl group and the hydrophobic side chain, analogous ligands wherein X = H or NH_2 (H-bonds with the protein carbonyl oxygen of the Gly216 residue) were also prepared and analyzed.

2. Results and discussions

2.1. Synthesis of thrombin inhibitors

The synthesis of thrombin inhibitors series IA (**3a**–**3f**), series IIA (**4a**–**4f**), series IIIA (**12a**–**12f**) and series IVA (**13a**–**13f**) was performed using a different synthetic route than what was previously published [20] and is included in the supplementary infomation.