Journal of Molecular Graphics and Modelling 38 (2012) 324–333

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Toward understanding the molecular basis for chemical allosteric modulator design

Qi Wang[a,b,1], Mingyue Zheng[c,1], Zhimin Huang[b], Xinyi Liu[b], Huchen Zhou[a], Yingyi Chen[b], Ting Shi[b], Jian Zhang[b,*]

[a] School of Pharmacy, Shanghai Jiao Tong University, Shanghai, 200240, China
[b] Department of Pathophysiology, Key Laboratory of Cell Differentiation and Apoptosis of Chinese Ministry of Education, Shanghai JiaoTong University, School of Medicine, Shanghai 200025, China
[c] Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

ARTICLE INFO

Article history:
Accepted 27 July 2012
Available online 8 August 2012

Keywords:
Allosteric regulation
Allosteric modulators
Structural rigidity
"Allosteric-like" rule

ABSTRACT

Among the regulation mechanisms of cellular function, allosteric regulation is the most direct, rapid and efficient. Due to the wider receptor selectivity and lower target-based toxicity, compared with orthosteric ligands, allosteric modulators are expected to play a larger role in pharmaceutical research and development. However, current difficulties, such as a low affinity and unknown structural features of potential allosteric small-molecules, usually obstruct the discovery of allosteric modulators. In this study, we compared known allosteric modulators with various compounds from different databases to unveil the structural and qualitative characteristics of allosteric modulators. The results show that allosteric modulators generally contain more hydrophobic scaffolds and have a higher structural rigidity, i.e., less rotatable bonds and more rings. Based on this analysis, an empirical rule was defined to determine the structural requirements for an allosteric modulator. It was found that a large proportion of allosteric modulators (80%) can be successfully retrieved by this "allosteric-like" filter, which shows good discriminatory power in identifying allosteric modulators. Therefore, the study provides deeper insight into the chemical properties of allosteric modulators and has a good potential for the design or optimization of allosteric compounds.

1. Introduction

With the advances in the techniques of molecular biology, our knowledge of the regulation mechanisms of cellular functions has expanded enormously in the past few decades [1]. Among these mechanisms, allostery, namely allosteric regulation, is regarded as the most direct, rapid and efficient regulatory mechanism [1]. It is defined as the regulation of protein function, structure and/or flexibility induced by the binding of a ligand or another protein at a site away from the orthosteric site [2]. This remote regulation effect is caused by the binding of the regulators to protein regions with a high conformational flexibility or by promoting the association or dissociation of oligomeric enzymes [3]. Allosteric regulation provides an effective mechanism for proteins to directly sense the alteration of the cellular milieu and to respond to maintain the homeostasis of the cell without a cellular energy cost [1]. In such regulation behaviors, the site other than the orthosteric site is defined as an allosteric site and is predominantly involved in the binding of metal ions and small molecules that are known as allosteric modulators.

The binding of an allosteric modulator can lead to the redistribution of protein conformation ensembles and can alter the rates of their interconversion, leading to positive or negative effects on protein function. Thus, allosteric binding sites have drawn increasing attention as novel targets for new drug development [2]. Chemical allosteric modulators provide several theoretical advantages over orthosteric ligands as potential therapeutic agents. For example, allosteric modulators have the potential to maintain their activity dependence and both temporal and spatial aspects of endogenous physiological signaling because of their quiescence in the absence of endogenous orthosteric activity. In addition, allosteric modulators can achieve a better selectivity due to the higher sequence divergence in allosteric sites, and the limited positive or negative cooperativity imposes a 'ceiling' on the magnitude of their allosteric effect. Therefore, allosteric modulators could be administered with a lower propensity toward target-based toxicity than orthosteric ligands [4].

Because of the potential advantages of allosteric modulators, allosteric sites are considered excellent drug targets, and the

* Corresponding author.
E-mail address: jian.zhang@sjtu.edu.cn (J. Zhang).
[1] Q.W. and M.Z. contributed equally to this work.

1093-3263/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2012.07.006

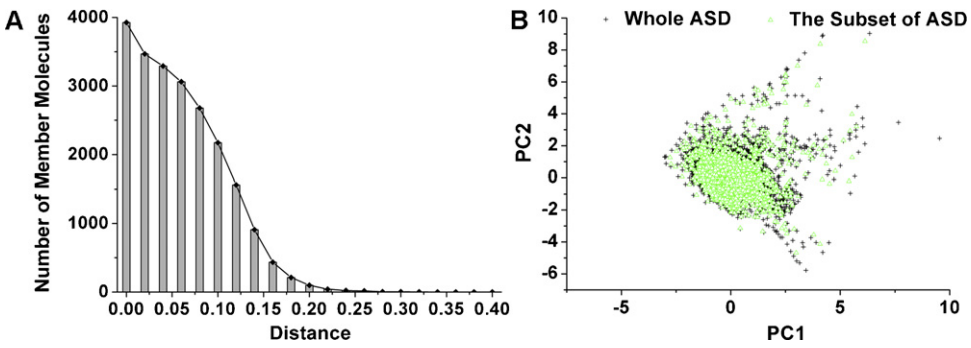

Fig. 1. (A) The number of compounds in the corresponding subset of the original ASD versus the distance parameter in the clustering. The distance represents a maximum dissimilarity calculated by "1 – Tanimoto similarity of molecular fingerprint FCFP4". (B) The distribution of compounds from the original ASD (7851 compounds, black plus sign) and ASD dataset (3916 compounds, green triangle) in the plot of the first two principal components extracted with the MOE descriptors.

discovery of more allosteric modulators could advance the development of novel drugs toward certain human diseases. The control of oxygen binding to hemoglobin by protons and the AMP-induced activation of glycogen phosphorylase are two of the earliest discovered cases of allosteric regulation [5,6]. As the investigation of this phenomenon moved forward, allosteric regulation aroused extensive interest and was defined by Jacob and Monod [7]. By the 1970s, approximately one dozen allosteric modulators had been discovered [8–10], and in recent years, a series of allosteric drugs targeting kinases, GPCRs and ion channels have been discovered, such as the development of Gleevec (allosteric inhibitor of Abl) [11], Cinacalcet (allosteric activator of a calcium-sensing receptor) [4] and Maraviroc (allosteric inhibitor of a chemokine receptor) [4]. Although much progress has been made in the development of allosteric modulators, we have not yet discovered a general method that can be widely used to deliberately discover allosteric modulators, other than the systematic discovery of genes and proteins that could be targeted. In the past, allosteric modulators were primarily found either by chance or by intuition followed by experimental verification [1]. Recently, a few allosteric modulators have been discovered by high-throughput screening (HTS). However, the allosteric small molecule "hits" typically have low affinities and unknown chemical compositions unsuitable for the discovery of allosteric modulators [1]. These problems could be alleviated by screening the potential for allosteric modulators in an "allosteric-like" small-molecule library and optimizing the structure of the hit according to the "allosteric-like" rule. Unfortunately, we have little knowledge about such rational design methods for chemical allosteric modulators.

To understand the inherent nature of allosteric modulators, the structural and physicochemical characteristics of allosteric modulators were systematically analyzed by comparing with various compounds from other chemical resources. Based on the analysis, an empirical rule was defined to determine the structural requirements for an allosteric modulator. Taken together, the study provides useful information for understanding the chemical properties of allosteric modulators and is a good guideline for the design or optimization of allosteric modulators.

2. Materials and methods

2.1. Databases

Six databases of compounds were collected for this research. As the primary research target, allosteric modulators were obtained from the AlloSteric Database (ASD), which is a recent database that provides a central resource for allosteric molecules [12]. In the ASD, 7851 organic allosteric modulators were manually curated from the literature. The number of reported allosteric modulators can differ significantly between allosteric proteins due to the different amount of research effort for different targets [12]. To avoid the potential bias caused by the inhomogeneous distribution of modulators, a refined subset with 3916 diverse allosteric modulators is provided to users of the ASD by the clustering method in Scitegic Pipeline Pilot v7.5 [13], as described below. The clustering analysis was performed with the "Cluster Molecules" component in the "Analysis and Statistics" module of Pipeline Pilot 7.5. In the component, the algorithm is a partitioning method, and a number of representative compounds are chosen as cluster centers to derive a subset from the original dataset by the maximum dissimilarity method. In each cluster, the distance between the center compound and member compounds is calculated by "one minus the Tanimoto similarity of molecular fingerprint FCFP4". The clustering begins by randomly choosing a compound as the first cluster center. The compound maximally distant from the first point is selected as the next cluster center. The compound maximally distant from both current points is selected after that. The non-selected objects are then assigned to the nearest cluster center to determine the cluster membership. The process repeats itself until there are the largest number of cluster centers and at least 2 compounds in each cluster. From this process, 3916 clusters were built, and the center compounds were extracted as the ASD subset. In every cluster, the maximum distance between the center and the members was 0.34 (Fig. 1A). To evaluate the final subset, descriptors of all compounds were calculated with the MOE 2008 modeling suite [14]. Twenty-seven common descriptors in drug discovery were calculated by MOE, including *weight* (molecular weight), *SlogP* (log of the octanol/water partition coefficient), *TPSA* (polar surface area), *a_aro* (number of aromatic atoms), *a_count* (number of atoms), *a_heavy* (number of heavy atoms), *a_nC* (number of carbon atoms), *a_nN* (number of nitrogen atoms), *a_nO* (number of oxygen atoms), *a_nP* (number of phosphorus atoms), *a_nS* (number of sulfur atoms), *a_nF* (number of fluorine atoms), *a_nCl* (number of chlorine atoms), *a_nBr* (number of bromine atoms), *a_nI* (number of iodine atoms), *b_ar* (number of aromatic bonds), *b_count* (number of bonds), *b_single* (number of single bonds), b_double (number of double bonds), *b_triple* (number of triple bonds), *b_heavy* (number of bonds between heavy atoms), *b_rotN* (number of rotatable bonds), *b_rotR* (fraction of rotatable bonds), *chiral* (number of chiral centers), *rings* (number of rings), *lip_acc* (number of O and N atoms) and *lip_don* (number of OH and NH atoms). In addition, principle component analysis (PCA) was used to analyze the relationship between the final subset and the original dataset, and the total variance represented within the first two principal components was 95%. As shown in Fig. 1B, the distribution of the subset is highly overlapped with the whole dataset, indicating that it is well balanced and is representative of the allosteric modulators. Therefore, the final subset (hereinafter referred to as the ASD dataset)