Progress in Biophysics and Molecular Biology xxx (2014) 1–7

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

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

Original research

# Chemical specificity and conformational flexibility in proteinase–inhibitor interaction: Scaffolds for promiscuous binding

Sneha Vishwanath, Narayanaswamy Srinivasan*

*Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India*

ARTICLE INFO

*Article history:*
Available online xxx

*Keywords:*
Interface flexibility
Promiscuous binding
Proteinases
Proteinase inhibitors
Protein–protein interaction
Sequence variability

ABSTRACT

One of the most important roles of proteins in cellular milieu is recognition of other biomolecules including other proteins. Protein–protein complexes are involved in many essential cellular processes. Interfaces of protein–protein complexes are traditionally known to be conserved in evolution and less flexible than other solvent interacting tertiary structural surface. But many examples are emerging where these features do not hold good. An understanding of inter-play between flexibility and sequence conservation is emerging, providing a fresh dimension to the paradigm of sequence–structure–function relationship. The functional manifestation of the inter-relation between sequence conservation and flexibility of interface is exemplified in this review using proteinase–inhibitor protein complexes.



## 1. Introduction

Association between different biomolecules is the driving engine of all the cellular processes in a biological system. Biomolecules that participate in the recognition process can range from small metabolites like glycine to large macromolecules like proteins. Proteins interact with biomolecules such as DNA, RNA, other proteins, metabolites and carbohydrates to carry out different cellular processes. Among these, protein–protein interactions are the most extensively studied and common form of recognition. Protein–protein interactions are common in large number of diverse processes such as cell–cell recognition, signal transduction and metabolism. Importance of protein–protein interactions can be understood from the abnormalities and diseases associated with absence of appropriate protein–protein interactions or occurrence of undesired protein–protein interactions (Ryan and Matthews, 2005).

Protein–protein complexes are characterized by their interfaces, which are the surfaces of interaction between the proteins. Many studies have characterized various structural and sequence features of interfaces of protein–protein complexes. Broadly these features are planarity (Jones and Thornton, 1996), circularity (Jones and Thornton, 1996), conservation of interface residues compared to the non-interacting surface residues (Lichtarge and Sowa, 2002), propensity of amino acids (Jones and Thornton, 1996) etc. Among these, the most important defining feature is the conservation of interface residues. Residues at the interfaces are reasonably well conserved due to both structural and functional constraints. Interface conservation has been exploited to distinguish interface regions from non-interface surface regions (Capra and Singh, 2007; Chelliah et al., 2004; Lichtarge and Sowa, 2002). However interfaces of some of the protein–protein complexes have been reported to show hypervariability. Exon regions of some genes have also been reported to be evolving at a higher rate when compared to intron regions. For example extremely high rate of non-synonymous nucleotide substitutions in the active domain-coding region of wheat thionin genes (Castagnaro et al., 1992), accelerated amino acid substitutions in the mature protein-coding regions of phospholipase A2 isozyme from the venom gland of *Trimeresurus flavoviridis* (Nakashima et al., 1993) and major histocompatibilty complex (Bjorkman et al., 1987). These proteins show accelerated rates of substitutions under adaptive evolution towards a wide range of cognate interacting partners (Castagnaro et al., 1992; Nakashima et al., 1993).

In majority of studies on protein–protein complexes, including those mentioned above, the static representations of 3-D structures solved using X-ray crystallography are used. But in recent times, importance of internal motion of proteins has been realized and its role in dictating the structure and function as well as sequence variability is appreciated. Because of these observations the concept of sequence–structure–function–dynamics (Fig. 1) is

*Abbreviations:* CIN, conservation of interface nature; CIP, conservation of interaction pattern.

* Corresponding author.
*E-mail addresses:* snehav@mbu.iisc.ernet.in (S. Vishwanath), ns@mbu.iisc.ernet.in, srinivasan@ncbs.res.in (N. Srinivasan).

http://dx.doi.org/10.1016/j.pbiomolbio.2014.08.003
0079-6107/

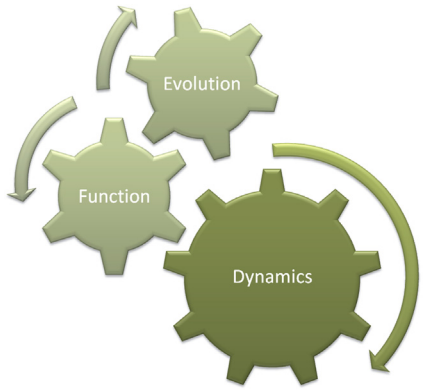


**Fig. 1.** Intricate relation between dynamics, function and evolution of proteins.

being understood better, where dynamics not only influences the structure and function but also the evolutionary dynamics perceived from sequence variations. One of the most common forms of motions observed in protein structures is the backrub motion, which essentially is the side chain coupled backbone motion (Davis et al., 2006; Friedland et al., 2009). Backrub motion is one of the important components for accommodating various amino acids in a protein fold without any change or subtle change in the fold and stability. Mutations in folds are accommodated through backrub motion by preserving the hydrogen bond pattern and the ideal geometry (Davis et al., 2006).

Friedland et al. reported reproduction of sequence variability observed in ubiquitin subfamily using only dynamic information for the three dimensional structure of ubiquitin (Friedland et al., 2009). For other proteins also, larger sequence space has been obtained using information only on the backbone flexibility (Mandell and Kortemme, 2009). The sequence space thus generated has been shown to bridge distant homologs better (Mandell and Kortemme, 2009). On comparing proteins across families and superfamilies, the backbone flexibility has been observed to be conserved (Maguid et al., 2006). Residues which diverge during evolution have been observed to be highly flexible (Liu and Bahar, 2012). This correlation between sequence divergence and mobility suggests that flexible regions in proteins can act as a scaffold for amino acid substitutions because of its loose packing leading to emergence of new functions with the retention of fold. Dellus-Gur et al. illustrated the role of dynamics in the evolution of new folds and novel functions with the examples of TIM barrel and DHFR fold (Dellus-Gur et al., 2013). TIM barrel fold, which has many promiscuous functions, has loosely packed active sites and high sequence divergence at and near active site. While in DHFR fold, which has only one function associated, the active site is rigid and shows lower sequence divergence (Dellus-Gur et al., 2013).

Some of the interface residues of protein–protein complexes are rigid especially if the residues are conserved and correspond to hot spots (Swapna et al., 2012; Yogurtcu et al., 2008). Rigidity of interface residues before complexation decreases the conformational entropy cost during protein–protein complexation (Yogurtcu et al., 2008). However flexible regions have also been implicated in the recognition process of proteins. For example, regions corresponding to RNA and protein binding have been reported to be highly flexible in *Saccharomyces cerevisiae* ribosomal protein L30 (Chao et al., 2003). Flexibility of RNA binding region in the protein region is crucial. Kaneko et al. illustrated the role of dynamic surface loops in broadening or narrowing the specificity and recognition of SH2, SH3 and PDZ domains (Kaneko et al., 2011). Dynamics of proteins has been reported to change after the formation of protein–protein complex (Grunberg et al., 2006).

Interface residues of proteinous proteinase inhibitors have been reported to be hypervariable. For a few examples, the interface residues have been observed to be flexible as well (Hubbard et al., 1991). Various arguments have been provided in the literature for the accelerated evolution of the interface of proteinase inhibitors. Some of the arguments are positive Darwinian selection (Creighton and Darby, 1989), hitchhiking (Hill and Hastie, 1987); allosteric sites (Pritchard and Dufton, 1999) and high gene copy (Jiang et al., 1994; Rheaume et al., 1994) among others. Accelerated mutations at the interface enable inhibitors to act on a number of exogenous proteinases thus increasing the fitness of the organism (Creighton and Darby, 1989). It is also argued that these mutations are selected due to their co-occurrence with other useful mutations – hitchhiking (Hill and Hastie, 1987). Existence of other important surface residues (allosteric sites) (Pritchard and Dufton, 1999), which regulates the inhibition, allows hypervariability at the interface of inhibitor. High gene copy number may account for hypervariability observed in paralogous inhibitors (Jiang et al., 1994; Rheaume et al., 1994).

In this review, characteristics of interfaces of proteinase – proteinase inhibitor protein complexes are revisited for a dataset larger than studied before. High sequence divergence and flexibility of interface residues, we believe, has led to promiscuous binding and better adaptability to different array of proteinases while low sequence divergence and low flexibility result in a specific binding as is the case with the proteinases which are specific towards their substrate. A better understanding of interplay of sequence conservation and mobility will allow us to design ligands with desired specificity and for developing therapeutics for the regulation of proteinases in diseased conditions.

## 2. Proteinases and inhibitor proteins

Proteinases are found in diverse taxa from archaea to prokaryotes to higher eukaryotes and perform diverse roles such as defense, digestion, blood coagulation and apoptosis. Based on the type of amino acid at the active center, proteinases have been classified into seven types namely aspartic proteinases, asparagine proteinases, cysteine proteinases, glutamic proteinases, metalloproteinases, serine proteinases, and threonine proteinases (Rawlings et al., 2012). Proteinases cleave diverse proteins; the specificity of substrates for proteinases is defined by the amino acid at the cleavage point (termed as P1) and the amino acid residues adjoining the cleavage point (termed P3′, P2′, P1′, P2, P3, P4 from N terminus to C-terminus (Schechter and Berger, 1968)). To maintain proteostasis, activity of proteinases has to be regulated. Many mechanisms exist for regulation of proteinases such as propeptides co-occurring with proteinase region, degradation of proteinases by other proteinases and inhibition by inhibitors.

Inhibitors of proteinases can range from small molecules to peptides to proteins. Because of the roles of proteinases in various physiological events, imbalance in the expression of the proteinases or mutations in proteinases can lead to many diseased conditions such as type 2 diabetes (Yoshida et al., 2012) and neural injuries and stroke (Huang and Wang, 2001). Examples of small molecule inhibitors include teneligliptin (Yoshida et al., 2012) which inhibits the dipeptidyl peptidase IV, 3-acetyl-2-aminoquinolin-4-one which inhibits calpain 1 (Kang et al., 2009) etc. Examples of peptide inhibitors are calpastatin which inhibits calpain (Todd et al., 2003), spinorphin which inhibits enkephalin-degrading enzymes like dipeptidyl peptidase III (Yamamoto et al., 2002) etc. In this review, the focus will be on complexes between proteinases and protein inhibitors. Like proteinases, the protein inhibitors (including pro-peptides) are found in diverse taxa and have been classified into 91 families (Rawlings et al., 2012). The classification of both proteinases and inhibitors is based on

Please cite this article in press as: Vishwanath, S., Srinivasan, N., Chemical specificity and conformational flexibility in proteinase–inhibitor interaction: Scaffolds for promiscuous binding, Progress in Biophysics and Molecular Biology (2014), http://dx.doi.org/10.1016/j.pbiomolbio.2014.08.003