Genomics Proteomics Bioinformatics 12 (2014) 1–7

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

REVIEW

Relative Specificity: All Substrates Are Not Created Equal

Yan Zeng *

Department of Medicine, Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455, USA

Received 26 November 2013; revised 21 December 2013; accepted 7 January 2014
Available online 31 January 2014

Handled by Jun Yu

KEYWORDS

Relative specificity;
Biochemical activity;
Substrates;
Biological systems

Abstract A biological molecule, *e.g.*, an enzyme, tends to interact with its many cognate substrates, targets, or partners differentially. Such a property is termed relative specificity and has been proposed to regulate important physiological functions, even though it has not been examined explicitly in most complex biochemical systems. This essay reviews several recent large-scale studies that investigate protein folding, signal transduction, RNA binding, translation and transcription in the context of relative specificity. These results and others support a pervasive role of relative specificity in diverse biological processes. It is becoming clear that relative specificity contributes fundamentally to the diversity and complexity of biological systems, which has significant implications in disease processes as well.

Introduction

Relative specificity is defined as the characteristic whereby in a biochemical system, a molecule, symbolized as E, interacts with its numerous substrates, targets or partners (collectively symbolized as {S}) differentially, thereby impacting them distinctively depending on the identity of individual substrates, targets or partners [1]. E can be a protein, RNA or any other biological molecule, capable of interacting with other molecules, *i.e.*, {S}, through binding and/or catalysis. Some examples are hemoglobin binding to O_2, CO_2 and a few other molecules, a receptor capturing different ligands, a cytochrome P450 enzyme metabolizing diverse chemicals, an RNA-binding protein associating with its RNA targets, a protein kinase phosphorylating substrates, a protein chaperone contacting unfolded or partially folded proteins, RNA polymerases transcribing genes, and the ribosome translating mRNAs. In many cellular systems, {S} can number in the hundreds, thousands or more. Research, however, has traditionally been centered on determining whether a molecule is the substrate or not of an E of interest. For various reasons, that an E might not target all of its {S} equally in complex systems, *i.e.*, relative specificity, is rarely treated as default, and what physiological consequences relative specificity may incur is even less investigated [1].

Evidence does exist to suggest that relative specificity is functionally relevant in complex biochemical systems. For example, the RNase Drosha (E) cleaved hundreds of human primary microRNA transcripts ({S}) with different efficiencies *in vitro*, which correlated with the expression of mature microRNAs *in vivo*, and such specificity was partially explainable by the structural properties of {S} [2]. The functionality of

* Corresponding author.
E-mail: zengx033@umn.edu (Zeng Y).

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.gpb.2014.01.001

relative specificity was also detected in systems involving a transcription factor and a protein kinase in budding yeast and an RNA-binding protein in humans [1]. The phenomena were generalized to formulate the relative specificity hypothesis, which has a number of features and implications. Firstly, it focuses on complex systems where an E acts on many, *e.g.*, hundreds or thousands of substrates, since such systems are abundant in nature, yet their relative specificity has been largely ignored partially due to technical limitations. Testing the hypothesis requires that we examine and compare the interactions between an E and its numerous substrates and then correlate the preferential interactions with a phenotype, in order to filter out the effects from factors other than the E of interest and make credible references to how the E's relative specificity contributes to a biological outcome. This is critical also because an observed biochemical property, *e.g.*, the binding of an E to a target, does not automatically equate to any biological function *in vivo*.

Secondly, the hypothesis does not stipulate the nature or origin of relative specificity in myriad biochemical systems or consider the abundance and subcellular localization of E and {S} to a first approximation. An E can bind to {S} with different on and off rates, different affinities, *etc*. It can bind to {S} to induce distinct conformational changes to selectively impact downstream signaling. It can also bind to {S} before efficient or inefficient enzymatic reactions. Any of these modes and their combinations could underlie the mechanism of relative specificity.

Thirdly, biological processes frequently mandate several biochemical activities in succession or in parallel. Analogous to rate-limiting reactions, even if relative specificity is exhibited by multiple Es in a biological process, the process may still be determined largely by the specificity from one of these Es. As an example, RNases Drosha and Dicer function in the same microRNA biogenesis pathway and were shown to cleave their respective substrates of microRNA transcripts preferentially *in vitro*, yet only the selectivity by Drosha significantly correlated with mature microRNA expression *in vivo* [2,3].

Lastly, the hypothesis promotes a reevaluation of certain concepts. For example, the literature contains ample statements like this: protein X has a high specificity. What it ultimately means is that X does not have many substrates. But it is likely that X still has more than one substrate, and X does not treat them equally. Conversely, if protein Y has a low specificity, then Y has many substrates, but again, Y still differentially interacts with these {S}. Furthermore, consider the following two schemes. In the first, an E acts identically on many substrates, which have different, sometime even opposing functions, but a specific biologic outcome nevertheless results, *e.g.*, cancer. In the second, the E acts on the same {S} differentially, again leading to a specific outcome, *e.g.*, cancer. Hence, a similar phenotype might originate from two distinct mechanisms. The key to distinguish between the two possibilities is to reveal whether the E reacts with {S} differently and the phenotypes can be partially explained by this relative specificity.

Below, I will review several recent studies of diverse biological systems in the context of relative specificity: (1) Hsp90–client interactions, (2) protein phosphorylation by the mechanistic target of rapamycin complex 1 (mTORC1), (3) RNA stabilization by RNA binding protein, fox-1 homolog (RBFOX1), (4) the impact of N-terminal codons on translation and (5) genome-wide transcription. I will then discuss a number of issues raised by the relative specificity hypothesis.

Hsp90–client interactions

Hsp90 is a molecular chaperone that associates with a large number of client proteins ({S}) to facilitate their folding. To study how Hsp90 recognizes {S}, Taipale et al. used a reporter assay to quantify the interactions between Hsp90 and hundreds of potential clients systematically in cell cultures [4]. Hsp90 was shown to interact with the majority of human kinases. However, the interaction was not binary, *i.e.*, substrate *vs.* non-substrate, but rather, as a sign of relative specificity, varied over a 100-fold range in strength, according to the reporter readouts. Cdc37, a co-chaperone of Hsp90, also selectively interacted with human kinases in a manner highly correlated with Hsp90. Mechanistically, the thermal stability of the kinases, with still poorly defined but both localized and broadly distributed components, was proposed to be the major determinant of how Hsp90 selects and discriminates {S}.

What are the functional consequences of differential Hsp90–client interactions? Taipale et al. found a modest but significant, negative correlation between the strength of Hsp90 interaction and recombinant kinase expression (R^2 = 0.15; [4]). Experimentally, the stronger the interaction, the larger the extent to which a recombinant client protein might be destabilized in cell cultures when Hsp90 was inhibited pharmacologically. In addition, weak human Hsp90 client kinases were more readily overexpressed than strong clients in bacteria, which lack Hsp90. These data suggest that Hsp90 selectivity might buffer protein folding; without Hsp90, intrinsically unstable proteins would have been even less stable and expressed at a lower level.

Protein phosphorylation by mTORC1

That relative specificity is functionally relevant can be easily rationalized if it correlates with differential gene expression. For the budding yeast Cdk1, its relative specificity *in vitro* positively correlated with substrate phosphorylation during mitosis, but whether the mere fact that Cdk1 phosphorylates {S} to varying degrees would have a biological consequence is not straightforward to address [1]. This potentially novel, global form of regulation has been at least partially tackled in the mTORC1 system [5].mTORC1 is a protein kinase that controls metabolism and growth in response to many stimuli, and its activity is altered by aging and in human disease such as cancers and can be inhibited by the drug rapamycin. Kang et al. performed *in vitro* kinase assays with recombinant mTORC1 and short peptides corresponding to mTORC1 phosphorylation sites in various substrates and their mutants [5]. Short peptides were used as the proxy for endogenous proteins because sequences immediately surrounding the phosphorylation sites in native substrates contain the most critical information necessary for kinase recognition and phosphorylation. mTORC1 was shown to phosphorylate some peptides/substrates more readily than others, indicative of relative specificity, or substrate quality as termed by Kang et al. [5]. mTORC1 activity depended partially on substrate binding affinity. Kang et al. then used a number of tests to demonstrate functional relevance [5]. For example, rapamycin, a pharmacologically