

The Increasing Complexity of Deciphering the Molecular Basis of Substrate Specificity

Marcy Hernick*

Department of Pharmaceutical Sciences, Appalachian College of Pharmacy, Oakwood, USA

Enzyme substrate specificity has been widely examined as a topic of research. Work in this area ranges from studies that simply seek to define enzyme substrate preferences to those that seek to identify the molecular basis of substrate specificity, and are carried out using a combination of biophysical and biochemical approaches. Information gained regarding substrate specificity can be used to identify the physiological substrates of enzymes, provide insights into enzyme function and/or regulation, and can be used for drug development. While studies aimed at defining enzyme substrate preferences are relatively straightforward, those seeking to identify the molecular basis of substrate specificity are complicated when there are side chains with multiple functions in catalysis (i.e., ligand binding, chemistry), side chains with synergistic function, and/or protein dynamics/conformational changes that take place.

Since enzymes can often catalyze their respective reactions with more than one substrate, studies aimed at probing substrate specificity often begin by examining the enzyme's ability to accept structurally distinct molecules via enzyme activity measurements with various substrates under steady-state conditions and comparison of the second-order rate constants (k_{cat}/K_M values), which describe the rates of free enzymes and free substrates reacting to form products. The parameter k_{cat}/K_M is preferred over k_{cat} because it includes contributions from substrate binding through the first irreversible step, often times chemistry, whereas the parameter k_{cat} simply reflects the rate of the slowest step in the catalytic cycle (e.g., chemistry, substrate binding, product release) and can vary between enzymes (i.e., wild-type, mutants) and substrates. The resulting information is used to identify features on substrate(s) that are critical for catalytic turnover and/or to define structural limitations of the ligand for tolerance by the enzyme. For example, *N*-acetyl-1-D-myo-inositol-2-amino-2-deoxy- α -D-glucopyranoside deacetylase (MshB) is a metal-dependent deacetylase that catalyzes the deacetylation of *N*-acetyl-1-D-myo-inositol-2-amino-2-deoxy- α -D-glucopyranoside (GlcNAc-Ins), a key step in the biosynthesis of mycothiol [1]. Experiments examining the MshB substrate preferences by measuring its activity with various substrates were used to establish GlcNAc-Ins as the physiological substrate for MshB, [1] including the stereochemical requirement for the 1-D-inositol-containing isomer [2]. While the natural substrate for MshB is GlcNAc-Ins, these studies also demonstrated the ability of MshB to catalyze the deacetylation (e.g., GlcNAc) and deamidation (e.g., AcCySmB-GlcN-Ins) of number of substrates, albeit with different efficiencies [1,3].

The molecular basis of the observed substrate specificity preferences is probed using a combination of biochemical and biophysical techniques. For example, important ligand-side chain interactions can be identified from structures of ligand-enzyme complexes and their importance in catalysis confirmed using a combination of site-directed mutagenesis (i.e., Ala) and kinetics experiments, while in the absence of structural data sequence alignment results can be used to identify conserved side chains to be probed biochemically. The magnitude of the decreases in activity varies widely depending on the nature of the interaction and can be used to gain insights into role of the side chain in the catalytic mechanism. The crystal structure of MshB complexed

with β -octylglucoside (BOG), whose glucoside moiety is proposed to mimic the binding of the GlcNAc moiety of the natural substrate to MshB, suggests a number of important hydrogen bonding interactions between side chains in the active site and substrate that can be tested biochemically [4]. For example, the MshB•BOG crystal structure suggests that the side chain of His144 shares a hydrogen bond with one of the glucose hydroxyl groups [4]. Mutation of His144 to Ala results in an 100-fold decrease in k_{cat}/K_M (substrate = GlcNAc) confirming the importance of this side chain in substrate recognition/catalysis [5]. Once important side chains have been confirmed biochemically, additional mutations can be used define size, hydrophobicity, charge, and hydrogen bonding potential requirements of the side chain as a determinant of substrate specificity.

In addition to mutagenesis experiments, additional studies may be required to more precisely define the nature of the molecular interactions involved in substrate specificity. This may include ligand-binding studies to separate contributions from binding and chemistry, as well as pH variation experiments to determine if there are ionizations involved in substrate recognition. For example, the contributions of binding and chemistry towards the substrate specificity preferences of the metal-dependent deacetylase UDP-3-O-(R-3-hydroxymyristoyl)-*N*-acetylglucosamine deacetylase (LpxC) were probed using product binding studies.[6]Results from these studies indicate that the fatty acid and UDP moieties make important contributions to the molecular recognition of substrate (4-5 kcal/mol and 2.9 kcal/mol, respectively), while the acetate moiety makes only a minor contribution on ligand recognition. Furthermore, pH variation studies identified that two ionizations are required for maximal binding of product to LpxC. To gain insights into the molecular basis of ligand recognition, results from crystallography and docking studies were used to identify side chains on LpxC that may contribute to the recognition of the UDPGlcNAc moiety (e.g., Phe192, Asp197, Lys239), and mutagenesis studies were used in combination with product binding studies to confirm that these side chains make important interactions with ligand for molecular recognition by LpxC [6].

Defining the molecular basis of substrate specificity is made increasingly complex when there are side chains with multiple functions in catalysis (i.e., ligand binding, chemistry), side chains with synergistic function, as well as factors such as protein dynamics/conformational changes. LpxC can be used as an example to highlight the complexity

*Corresponding author: Marcy Hernick, Department of Pharmaceutical Sciences, Appalachian College of Pharmacy, Oakwood, USA, Tel. 276-498-5244; E-mail: mhernick@acp.edu

Received August 12, 2013; Accepted August 14, 2013; Published August 17, 2013

Citation: Hernick M (2013) The Increasing Complexity of Deciphering the Molecular Basis of Substrate Specificity. Biochem Anal Biochem 2: e142. doi: 10.4172/2161-1009.1000e142

Copyright: © 2013 Hernick M. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

of having side chains with multiple and synergistic functions. The side chains of Glu78 and His265 function as a general acid-base catalyst pair in the chemical reaction, and work synergistically to bind substrate/product [6,7]. Consequently, examination of the kinetic k_{cat}/K_M values for the single E78A and H265A mutants underestimates the contribution of the individual side chains to catalysis, and a combination of mutagenesis (i.e., double E78A/H265A mutant), pH variation, and solvent isotope effect experiments were required to ascertain the contributions and functions of these side chains to catalysis [7]. Interestingly, the Glu78 and His265 side chains also function synergistically in ligand binding, thus requiring examination of the single (i.e., E78A, H265A) and double E78A/H265A mutant in the product binding studies.

While crystal structures provide valuable insights into the molecular recognition of ligands by enzymes, it is important to remember that these structures only offer snapshots into the enzyme-ligand binding process and may not capture all forms of the enzyme-ligand complex when there are protein dynamics/conformational changes that take place. This complexity can be illustrated using MshB as an example. Initial crystal structures of Zn-MshB [8] and MshB•BOG [4] did not suggest a potential role for Tyr142 in ligand binding or catalysis; however, results from biochemical studies indicated that Tyr142 is a dynamic side chain that participates throughout the catalytic cycle and a model describing how the Tyr142 side chain could move to facilitate substrate binding, chemistry, and product release was proposed [5].

This model is supported by a more recent crystal structure of MshB containing bound acetate and glycerol molecules that captures the Tyr142 side chain in the proposed chemistry position [9]. While these studies have advanced our understanding of molecular recognition by MshB, additional studies are needed to better understand the role of protein dynamics in ligand recognition, including identifying other side chains that aid the movements that occur during the catalytic cycle. This will require an interdisciplinary approach that uses computational tools such as molecular dynamics simulations and docking studies, as well as additional biophysical (e.g., MshB crystal structures) and biochemical (e.g., mutagenesis, kinetics, binding studies) experiments. Research aimed at understanding the molecular basis of enzyme substrate specificity is not trivial, and requires an interdisciplinary

approach with a combination of biophysical and biochemical studies to fully elucidate the enzyme-substrate interactions that define substrate specificity. This will continue to be an important avenue of research as results from studies that define the molecular basis of substrate specificity preferences have important implications for understanding the biological function of enzymes and for the development of potent and specific inhibitors.

References

1. Newton GL, Av-Gay Y, Fahey RC (2000) N-Acetyl-1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase (MshB) is a key enzyme in mycothiol biosynthesis. *J Bacteriol* 182: 6958-6963.
2. Nicholas GM, Eckman LL, Kovac P, Otero-Quintero S, Bewley CA (2003) Synthesis of 1-D- and 1-L-myo-Inositol 2-N-Acetamido-2-deoxy-a-D-glucopyranoside Establishes Substrate Specificity of the Mycobacterium tuberculosis Enzyme AcGI Deacetylase. *Bioorganic & Medicinal Chemistry* 11: 2641-2647.
3. Newton GL, Ko M, Ta P, Av-Gay Y, Fahey RC (2006) Purification and characterization of Mycobacterium tuberculosis 1D-myo-inositol-2-acetamido-2-deoxy-alpha-D-glucopyranoside deacetylase, MshB, a mycothiol biosynthetic enzyme. *Protein Expr Purif* 47: 542-550.
4. McCarthy AA, Peterson NA, Knijff R, Baker EN (2004) Crystal structure of MshB from Mycobacterium tuberculosis, a deacetylase involved in mycothiol biosynthesis. *J Mol Biol* 335: 1131-1141.
5. Huang X, Hernick M (2012) Examination of mechanism of N-acetyl-1-D-myo-inositol-2-amino-2-deoxy-1- β -D-glucopyranoside deacetylase (MshB) reveals unexpected role for dynamic tyrosine. *J Biol Chem* 287: 10424-10434.
6. Hernick M, Fierke CA (2006) Catalytic mechanism and molecular recognition of *E. coli* UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase probed by mutagenesis. *Biochemistry* 45: 15240-15248.
7. Hernick M, Gennadios HA, Whittington DA, Rusche KM, Christianson DW, et al. (2005) UDP-3-O-((R)-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase functions through a general acid-base catalyst pair mechanism. *J Biol Chem* 280: 16969-16978.
8. Maynes JT, Garen C, Cherney MM, Newton G, Arad D, et al. (2003) The crystal structure of 1-D-myo-inositol-2-acetamido-2-deoxy-alpha-D-glucopyranoside deacetylase (MshB) from Mycobacterium tuberculosis reveals a zinc hydrolase with a lactate dehydrogenase fold. *J Biol Chem* 278: 47166-47170.
9. Broadley SG, Gumbart JC, Weber BW, Marakalala MJ, Steenkamp DJ, et al. (2012) A new crystal form of MshB from Mycobacterium tuberculosis with glycerol and acetate in the active site suggests the catalytic mechanism. *Acta Crystallogr D Biol Crystallogr* 68: 1450-1459.