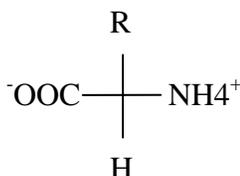


## Lecture #2 – Review of Protein Chemistry, Enzyme Specificity

Handouts: (1) Structures of amino acid side chains, (2) Sample bioconversions to examine specificity

### 1. Review of protein chemistry, enzyme specificity

Recall that proteins have a *primary* structure that is the string of amino acids as encoded by the DNA. There are twenty amino acids, of the general structure:



where R is the side chain. Side chains come in three “flavors”: hydrophobic, hydrophilic, and “other.” The extent to which side chains are hydrophilic or hydrophobic varies with each group (refer to handout). For example, the hydrophobic group contains both the branched chain aa’s valine and leucine, as well as the cyclic side chains of tryptophan and phenylalanine. The hydrophilic group contains the polar aa’s serine and threonine, as well as charged aa’s aspartic acid, glutamic acid, and lysine. The two classified as neither hydrophobic nor hydrophilic are cysteine (sulfur-containing) and proline.

Proteins also have *secondary*, *tertiary* and *quaternary* structures. The *secondary* structure is defined by the local structure of the linear (primary) string of amino acids. Secondary structures are general in the form of  $\alpha$ -helices,  $\beta$ -sheets, or flexible. The *tertiary* structure represents the higher-order folding of the chain into its final three-dimensional structure, while *quaternary* structures are formed from the interaction of two or more individually-folded chains.

Proteins take on their predestined structure based on the nature of the amino acids of which they are comprised. They are held together in an active (or inactive) conformation through various, mostly weak forces, including hydrophobic interactions, electrostatic (ionic) forces, and hydrogen bonding. Keep in mind as well that proteins are typically found in aqueous environments, so the presence of water is a prime determinant in the ultimate configuration of any polypeptide chain.

*Why do we care about the nature of the protein structure and the forces that hold it together?* Because the conformation of the protein and the composition of the side chains

in the backbone affect the *specificity* and *activity* of the enzyme. The folded conformation of an enzyme forms a binding pocket for a substrate to be converted, and the side chains that are present in and around that pocket determine which substrates will “fit” in the pocket.

Example: The Serine Proteases (EC Class 3 enzymes)

Trypsin, chymotrypsin, and elastase are all serine proteases, so-called because of a catalytically active serine residue in the active site that initiates hydrolysis of the peptide chain. All three enzymes react through the same mechanism, all three contain the “catalytic triad” of serine, histidine and aspartic acid that allows distribution of ionic charges to facilitate conversion.

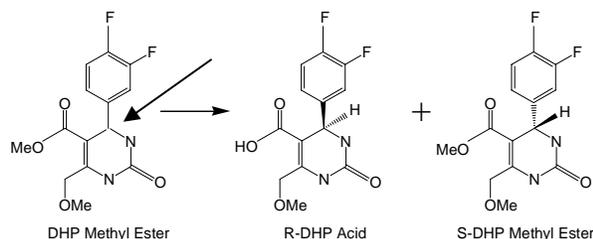
However, the specificity of each is different: *Note – present this in reverse order. Tell students we’re looking at serine proteases as an example, then ask which amino acids they think a protease would be specific for if the binding pocket contains a charged residue like aspartate. Then ask which amino acids might be in the pocket if the specificity is for the bulky hydrophobics.*

<b>Enzyme</b>	<b>Specificity</b>	<b>Binding site aa’s (in pocket)</b>
Trypsin	Lysine, arginine	Aspartate (ionic interactions with Lys, Arg)
Chymotrypsin	Phenylalanine, tyrosine, tryptophan	Serine, 2 glycines (allows large hydrophobic groups)
Elastase	Small hydrophobics, eg, alanine	Valine and threonine (allows only small side chains)

The various forces that are involved in establishing the structure of enzymes are also clearly involved in determining their specificity (*ie*, ionic interactions with trypsin, hydrophobic interactions with chymotrypsin and elastase). Only substrates that will fit into the binding pocket can be effectively converted.

The substrate specificity of an enzyme is important for its successful use as a biocatalyst. We will want to convert substrates that are analogous to the natural substrates, but a good biocatalyst will display activity against a good range of substrates. Consider again the serine proteases. These enzymes work to hydrolyze peptide bonds (amides), but are also active against esters.

For example, the enzyme subtilisin is a serine protease from a different family than that consisting of trypsin/chymotrypsin/elastase, but the same catalytic triad is present. This enzyme is widely used commercially in detergents; however, it can also catalyze the following bioconversion (Courtesy of Merck & Co., Inc. Used with permission.):



In this case, the R group looks nothing like a peptide, yet the enzyme is still active. The product here is an intermediate in the production of a pharmaceutical compound that was investigated for the treatment of benign prostate enlargement.

This example raises two issues about enzyme specificity. First, it is typically impossible to tell if an enzyme will absolutely have activity against a “new” compound, *ie*, one which it has not seen in nature. In this case, it was known that the protease was active against esters, and that in general, it had broad substrate specificity. Thus, this was a natural choice as a biocatalyst. However, it would not have been surprising if the enzyme was not active against this compound either. Structural protein chemistry has made significant advances towards determining the nature of binding pockets, and software programs do exist to model substrate binding. But, it’s still better and far more definitive to test catalytic activity experimentally in a wet lab than *in silico*. Having said that, one can look at the structures of substrates known to be converted to gain clues as to which enzymes should be tested on new substrates. More on that in a minute.

The second thing to take home from this example is that the substrate in this case was a racemic mixture of two enantiomers (see arrow for chiral center). Yet, the enzyme only acted on one of the enantiomers to convert it to an acid. (The S-ester was the desired product). This emphasizes the *enantioselectivity* of the enzyme, that is, its ability to selectively convert one of two enantiomers. The enantioselectivity of an enzyme can be quantified using a value called the *enantiomeric excess*, or EE, defined as follows:

$$EE(S) = \frac{S - R}{S + R} \times 100\%$$

A closer look at this equation reveals that the EE is simply the fraction of product in the desired conformation minus the fraction of product in the undesired conformation:

$$EE(S) = \frac{S}{S + R} - \frac{R}{S + R} \times 100\% = \frac{S - R}{S + R} \times 100\%$$

The total of both enantiomers equals 100%, so, for example, if a reaction produces 95% product in the S-form and 5% in the R-form, the EE is 90% with respect to S. For bioconversions, we usually like EE to be >95%, which means the yield must be >97.5%, a very high level of selectivity. Also the starting substrates here were a mixture of the two enantiomers, one could also start with a pro-chiral (symmetric) substrate that is selectively converted to a chiral product with high EE. [Examples from Lecture 1: lactate dehydrogenase, transaminase]

In this example, the starting substrate is a racemic mixture, so the EE is 0%. The biocatalyst is used to selectively hydrolyze one of the enantiomers, so the maximum yield is 50%. Such a reaction is called a *resolution*, because it resolves the mixture into a high-EE product of one versus the other enantiomer. Lipases are one group of the hydrolases that perform these reactions, and a variety of these enzymes are available for conversion. There is, however, a large disadvantage if the maximum achievable yield is 50%. We'll come back to this issue later in the course.

## **2. Substrate range/enzyme specificity**

Now, let's look at some examples of naturally-occurring enzymes with their native substrates, then consider other substrates that have been converted by these enzymes.

Examples (see handout):