IIT Guwahati

Lecture 18

Course BT 631

Protein Structure function and Crystallography

Prof. Arun Goyal

Dept. of Biosciences and Bioengineering

Relation between Fold and Function

Protein taxonomy

- **Structure determination of proteins requires much more work than the determination of DNA sequences and new structures are appearing at an increasing rate.**
- **With the growing database of known protein conformations there is a need to find ways of classifying protein structures, to define a structural *taxonomy* for proteins or protein domains.**
- **The enormous variability in natural amino acid sequences does not translate into a similar variability in folds.**
- **Many proteins have the similar folds, even in cases where they have no evolutionary relation.**

Relation between Fold and Function

Protein taxonomy

One common way of classification is to divide all proteins or protein domains into 5 major groups referring to their secondary structure content: all α, all β, α/β, α + β and a 5th group of proteins which has little secondary structure and do not fit in any of the other categories.

The α/β and α + β proteins differ in that the α/β proteins are mainly built up of parallel strands connected by helices (βαβ motifs), whereas the α + β proteins have β-strands and helices connected in a more irregular fashion.

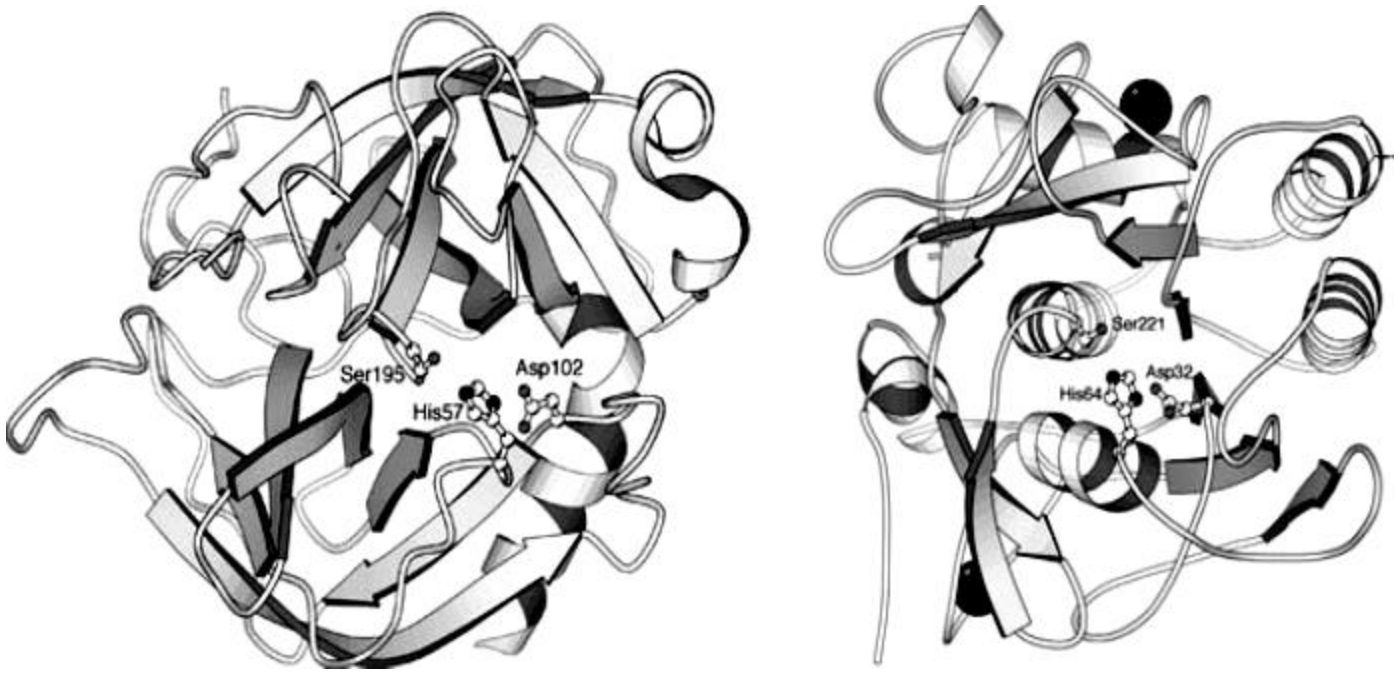
Within each class there are many different folds and each fold may have many variants.

Biochemical function: Convergence and divergence

- It is important to note that the classification of protein folds in structural classes is not linked to any functional classification.
- In general, a certain type of function is not restricted to a certain class of folds.
- This is illustrated by the serine proteases and related enzymes, where there are several distinct families of proteins with similar function but with different fold.

Biochemical function: Convergence and divergence

- The serine protease, chymotrypsin belongs to a large group of mostly eukaryotic proteolytic enzymes.
- Chymotrypsin is composed of two similar domains that have a β-barrel fold (Fig.).
- The active site and catalytic mechanism of chymotrypsin is very similar to that of subtilisin, which is a bacterial enzyme with a completely different fold (α+β). (Fig.)
- Chymotrypsin and subtilisin do not have a common origin.



The folds of the serine proteases Chymotrypsin (*left*) and Subtilisin (*right*) are very different, but the proteins have the same function and enzymatic mechanism. The Ser-His-Asp catalytic triad is indicated.

Biochemical function: Convergence and divergence

The three residues involved in the catalytic mechanism (the catalytic triad) Ser-His-Asp are the same but occur in different orders in the sequence. The catalytic center and the mechanism are quite similar. This is a case of *Convergent evolution*: the same function has evolved independently in two proteins. The serine protease example shows that the actual fold of a protein can be regarded as a means to form a stable structural scaffold. On this scaffold, active sites and other functional properties are molded.

The existing folds are the result of their evolutionary history, rather than the suitability of a certain type of fold for a certain type of function.

Convergent Evolution: Different structures but same function. (Chymotrypsin and Subtilisin)

Divergent Evolution: Similar structures but different functions.

Biochemical function: Convergence and divergence

Similar structures performing same function.

Still, many functions are connected to specific types of folds. One example is immunoglobulin: all immunoglobulins have several domains, all of them β-sandwiches with very similar topology. These types of domains are not only found in the proteins of our immune system but also in many cell surface receptors.

The cause of this similarity is probably that most or all of these proteins have evolved from a common origin.

Biochemical function: Convergence and divergence

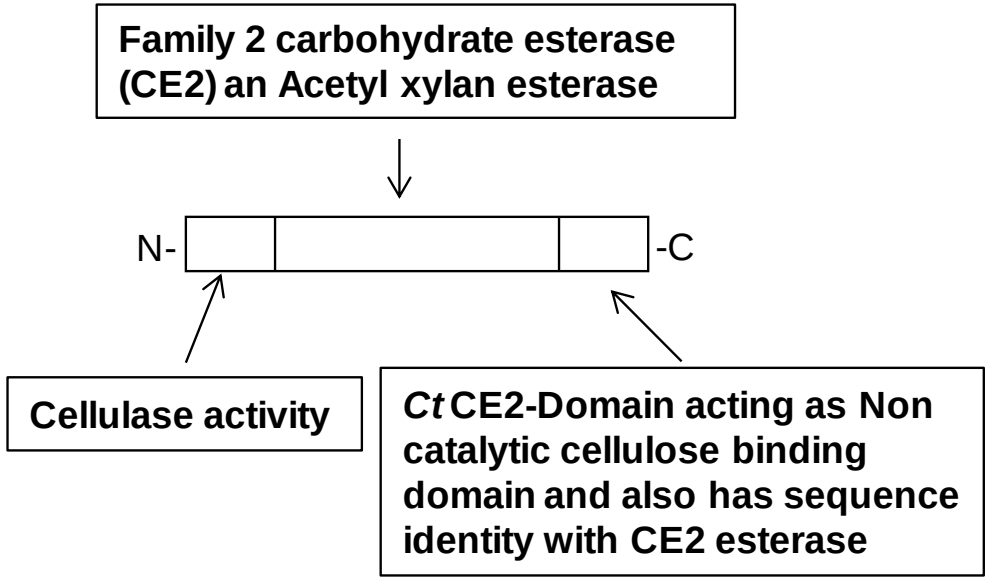
Divergent evolution

When two similar proteins have similar structure but perform different functions is called Divergent evolution.

Single protein domain with multifunction

A carbohydrate esterase (*Ct*Cel5C-CE2) from *Clostridium thermocellum* contains an N-terminal module that displays cellulase activity and a C-terminal module, *Ct*CE2, which exhibits a noncatalytic cellulose-binding function, but also shares the sequence identity with the family 2 carbohydrate esterases (CE2).

It was shown that, unlike other CE2 members, *the CtCE2 domain displays divergent catalytic esterase and noncatalytic carbohydrate binding functions.*

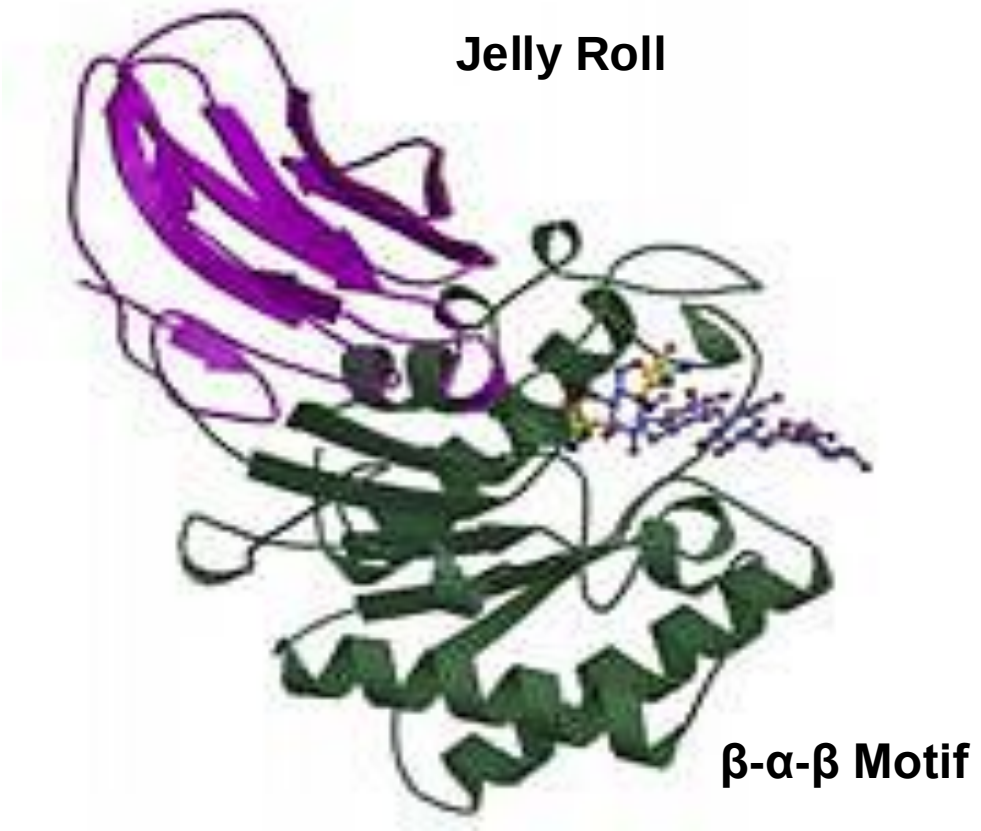


Biochemical function: Convergence and divergence

Single protein domain with multifunction

The 3D structures revealed a bidomain enzyme in which an N-terminal β-sheet "jelly roll" domain (around 130 residues) is linked to a C-terminal domain of approximately 220 residues.

The C-terminal domains possess a typical α/β hydrolase fold consisting of repeating β-α-β motifs that form a curved central five-stranded parallel β-sheet.



Structure of CE2, Acetyl xylan esterase