Abstract
Serine proteinase inhibitors (Serpins) are irreversible suicide inhibitors of proteases that regulate diverse physiological processes such as coagulation, fibrinolysis, complement activation, angiogenesis, apoptosis, inflammation, neoplasia and viral pathogenesis. The molecular structure and physical properties of serpins permit these proteins to adopt a number of variant conformations under physiological conditions including the native inhibitory form and several inactive, non-inhibitory forms, such as complexes with protease or other ligands, cleaved, polymerised and oxidised. Alterations of a serpin which affect its structure and/or secretion and thus reduce its functional levels may result in pathology. Serpin dysfunction has been implicated in thrombosis, emphysema, liver cirrhosis, immune hypersensitivity and mental disorders. SERPIN BASE in its current form shall be a trusted resource for researchers working on the superfamily and for biologists with a yearn to learn.

Keywords: Serpin, phylogeny, Umbrella sampling, Docker in repeats

Introduction:
Serpins (serine protease inhibitors or classified inhibitor family) are the largest and most broadly distributed superfamily of protease inhibitors (1-5). Serpin-like genes have been identified in animals, poxviruses, plants, bacteria and archaea, and over 1,500 members of this family have been identified to date. Analysis of the available genomic data reveals that all multicellular eukaryotes have serpins: humans, Drosophila, Arabidopsis thaliana and Caenorhabditis elegans have 36, 13, 29, and about 9 serpin-like genes, respectively. In contrast, serpins in prokaryotes are sporadically distributed and most serpin-containing prokaryotes have only a single serpin gene (6-15). The majority of serpins inhibit serine proteases, but serpins that inhibit caspases and papain-like cysteine proteases have also been identified. Rarely, serpins perform a noninhibitory function; for example, several human serpins function as hormone transporters and certain serpins function as molecular chaperones or tumor suppressors (16-20, 40-42). A phylogenetic study of the superfamily has segregated the eukaryotic serpins into 16 ‘clades’ (termed A-P). The proteins are named SERPIN Xy, where X is the clade and y is the number within that clade; many serpins also have alternative names from before this classification was proposed specifically or inhibitory function. Here, we summarize the evolution, structure and mechanism of serpin function and dysfunction.

*Corresponding Author
Sayak Ganguli
DBT Centre for Bioinformatics
Presidency University, Kolkata, India.
Implementation:
The database includes analytical information about the various conserved domains found in the serpin protein family. These conserved domains were identified using the umbrella sampling approach: It was performed to select the specific conservation scores with the least free energy. This method scores over other methods since it can be used both with Monte Carlo and Molecular Dynamics simulations and the modifications of the potential function can be written as a perturbation:

\[ V'(r^n) = W(r^n) + W(r^0) \]

Where \( W(r^n) \) is a weighting function, which can be expressed as a quadratic form:

\[ W(r^n) = kw(r^n - r_0)^2 \]

The results of the analysis can be viewed by users using the slideshow option. Phylogenetic analysis was performed using PHYLIP and several different sister groups were obtained. The tree(s) obtained can be viewed by the user at the "EVOLUTIONARY RELATIONSHIP AMONG THE SERPIN" segment. All the trees present are neighbour-joining trees based on Kimura calculations.

Results:
The non-parametric bootstrap technique was used to eliminate relationships with poor support; as a result, several branches have been found to radiate from the base of the trees due to uncertainty in their relationships(21-25). As the base of the tree indicates the lowest point at which groupings can be reliably determined, it does not reflect a fixed point in time. However, clades containing human serpin genes significantly are limited to sequences from vertebrates; furthermore, non-vertebrate clades are conspicuous by the absence of vertebrate sequences. Thus the phylogenetic analyses reveal that serpins of individual members are unique to the group they represent and there is no major event of divergence among the members (26 – 27). Their closeness in association is due to the similarity in domain acquisition, as evident from their common conserved domains across several groups of organisms and thus they probably follow the modular assembly path of protein evolution (28 – 39). The most common conserved domains that were identified were CDOO172 and the dockerin repeat domains.

Conclusion
SERPINBASE is a unique database where the user is able to access data regarding the evolutionary relationships and processes of evolution of serpins. It provides information regarding the modular evolutionary cascade of the group and also takes into account the conserved domains in the homologues. Future additions to the database would be simulated 3D structures of the important members and a section where the codon bias would also be included.

AVAILABILITY
http://bioinfopresidencycollegekolkata.edu.in/serpin.html

References


