

GENERATION AND SELECTION OF PROTEIN LIBRARY IN SILICO

ABSTRACT

The present invention provides a methodology for efficiently generating and screening protein libraries for optimized proteins with desirable biological functions, such as improved binding affinity towards biologically and/or therapeutically important target molecules. This process is carried out computationally in a high throughput manner by mining the ever-expanding databases of protein sequences of all organisms, especially human. In one embodiment, a method for constructing a library of designed proteins, comprising the steps of: providing an amino acid sequence derived from a lead protein, the amino acid sequence being designated as a lead sequence; comparing the lead sequence with a plurality of tester protein sequences; and selecting from the plurality of tester protein sequences at least two peptide segments that have at least 15% sequence identity with the lead sequence, the selected peptide segments forming a hit library; and forming a library of designed proteins by substituting the lead sequence with the hit library. The library of designed proteins can be expressed in vitro or in vivo to produce a library of recombinant proteins that can be screened for novel or improved function(s) over the lead protein, such as an antibody against therapeutically important target.

(no suitable figure)