An experimental study on randomly generated protein sequences can provide important insights into the origin and mechanism of secondary structure formation and protein folding. In this study we bring biophysical characterization of five protein sequences selected from the in silico generated library of random chains. The sequences were selected on the basis of bioinformatic analysis in order to find the candidates with the maximum potential to possess secondary structure. This study shows that the random polypeptide sequences form stable secondary structures and in some show the signs of tertiary structure, such as hydrophobic core formation and distinctive oligomerization pattern.

While the work presented in this thesis is work in progress on a larger study, the data already demonstrate that unevolved protein sequence space provides a lot of potential for secondary and tertiary structure formation that awaits its characterization.