 doi:10.20944/preprints202006.0350.v1

Mini-review

Antibodies Engineering by Computational Approach

Mujahed I. Mustafa

Department of Biotechnology, University of Bahri, Khartoum, Sudan.

Correspondence; mujahedibrahim44@gmail.com

Abstract:

In the pre era of synthetic antibodies, pharmaceutical companies depend on finding novel drugs from medicinal plants and other traditional resources; while in present, technological advances in biology, computer and robotics give the researchers the ability to rewrite and edit DNA in order to synthesize very large sets of drug candidates; these novel and improved candidates serves the basis for creating another library of drug candidates and so on until we find the right biomolecule for the disease of interest. all these technologies combined together to synthesize therapeutic antibodies for many types of cancer, autoimmune diseases, and infectious diseases, that can address diseases much more readily to very rapidly get therapeutics into patients so that we can potentially have an impact on disease. The antibodies mechanism is recognize and bind to disease cells and pinpoint the immune system to attack those cells effectively. Now a days, they dependent on computational approach to guide and accelerate the process of antibodies engineering by combination of selection system and use of high-throughput data acquisition and analysis to build and construct populations of next generation antibodies that are thermo-stable, non-immunogenic as possible, and to be administered to many humans as possible. In this review, I will discuss the latest *in silico* methods for antibodies engineering.

Keywords: Antibodies engineering; Computational approach; Novel drugs; Synthetic immunology; Next generation antibodies.

1. Introduction:

Synthetic immunology is the engineering synthetic systems that serve complex immunological role [1-3] such as next generation antibodies with estimated global pharmaceutical market of USD 140 billon by 2024 [4]. In recent years, antibody engineering by *in silico* methods are nearly worldwide accepted as critical approach to guide and accelerate the process of engineering antibodies by combination of selection

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

system and use of high-throughput data acquisition and analysis to build and construct populations of next generation antibodies that are thermo-stable, non-immunogenic as possible, and to be administered to many humans as possible [5]. Engineering antibodies by computational approach is mainly for coming of ways of combining high-throughput data accusation, high-through genomic sequencing and robotized high-throughput selection along with synthetic natural DNA sources to carefully consider designs to accelerate the rate of antibodies discovery as drugs [5] **(Figure 1a)**. The history of that has been initially a pioneering technique use high-throughput sequencing to investigate why is libraries don't produce more hits, and how to improve the quality of these hits, to make them better drugs, that are shelf-stable, thermo-stable, aggregation resistant and non-immunogenic [6-8].

It's well-known that antibody holds two chains (VL and VH), each of which is consist of several domains. The antigen-binding site is located in the 'variable' domains of each chain. The remainder of the variable domains is structurally well conserved at the backbone level. Therefore, a main focus of antibody design is dedicated for predicting the conformations of the CDR loops from their sequences [9, 10]. The increasing knowledge of sequence–structure co-relation in antibodies, and the advancement in *in silico* approach particularly in protein modeling, has facilitated growth of *in silico* methods that can aid in engineering antibodies for desired alterations [5, 11, 12]. The main target area of biological engineer's is complementarity-determining regions (CDRs) because antibodies function through it [13]. The recent advancement in synthetic biology have let to manipulates the amino acid sequences by enhancing the affinity of CDRs [14]. Nevertheless, the trade-off between binding affinity and other properties is a concern of engineering antibodies. To conquer such an issue, a standard approach to enhance the properties of antibody is random mutagenesis based on *in vitro* libraries [15].

Now a day, it's relatively an ease task to engineer binding affinities with other properties through such an *in vitro*, library-based method by elevating temperature and controlling solution conditions during the selection process. However, due to the advances in computational power and deep sequencing, and artificial intelligence, *in silico* strategies is becoming an alternate approach in engineering antibody [16]. Yet, such predictions are not completely pleasing for good two explanations, there is no straightforward protocol to design by, and the precision of *in silico* approach is not as good as that of *in vitro* libraries approach because the biophysical properties of biomolecules are not well understood [15, 17]. In this review, I discuss the frontier in *in silico* engineering of antibodies properties. Among several properties, I focused on the stability, viscosity, and immunogenicity of antibodies, all of which have garnered much devotion in antibodies engineering by computational approach **(Figure 1a)**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

2. Physicochemical and biological properties of antibodies:

2.1. Stability:

One of the most important properties in antibody drug discovery is protein stability, which there is 2 types, physical and chemical stabilities. Commonly, protein stability can be categorized into conformational stability and colloidal stability. Proteins are only slightly stable, and proteins in solution are in equilibrium between folded and unfolded conformations **(Figure 2)**. Many homology modeling based studies [18-20] has been sufficient for developability goals (excluding the cases when antibody present rare CDR in whichever heavy chain or light chain loops).

Protein aggregate is state occurring when a protein in in a folded-unfolded stability may accumulate into oligomeric condition which is irreversible process (Figure 2); on the protein surface, short hydrophobic fragments designated as aggregation pone regions (APRs), researchers have considered to dictates the aggregation tendencies of proteins, furthermore, any mutation in on this region can intensely affect the rate of aggregation

2.2. Viscosity:

Another important property in antibody drug discovery is antibody viscosity, due of its applied consequences with respect to formulation and administration. The behavior of the concentration-dependent viscosity of antibodies depends on pairwise interactions or self-association, which further leads to higher-order intermolecular interactions [15] **(Figure 2)**.

2.3. Immunogenicity

The term "immunogenicity" is refers to the patients' immune response against the proteins. Immunogenicity evaluation usually carries by animals testing, which is cost a lot of time, cost and effort. Therefore, use of *in silico* approach cut-off a lot of time and costs, by facilitating sequence alignments to check the amino acids similarities of antibodies and target of interest [21].

3. Prediction and designing of physicochemical properties of antibodies:

3.1. Overview of computational prediction and designing:

Several *in silico* approach for physicochemical properties prediction such as viscosity, protein stability and immunogenicity have been established to enable predictive protein engineering. The input and output of these tools are summarized in **(Figure 1b)**. Generally, these prediction models can be classified into 2 groups: statistical predictions and physics-based predictions. Physics-based approaches, the unique about these

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

approaches, do not rely on any experimental data to achieve predictions because it based on laws of physics. While statistical approaches depend on statistical data extracted from experimental information, and the predictions accuracy is proportional to the quality of artificial intelligence that used to train the prediction approaches [22, 23].

3.2. Prediction and designing of Viscosity:

Viscosity prediction has gathered far share of devotion as a target of interest in antibody design by *in silico* method. Some studies have revealed that high-antibody viscosities are better associated with negative than positive charges [24]. In some cases, aggregation form due to unusual viscosity behavior of an antibody. This behavior led some scientists to suggest that viscosity behaviors of antibodies are driven by their crossponding amino acid sequences. Keeping in mind that the constant domains of antibodies are highly conserved, the differences in behaviors are probably as a result of differences in the variable regions [25].

3.2. Prediction of colloidal stability and solubility:

One of the hot zone in research field is proteins aggregation, specifically regards to the capability to develop protein therapeutics. Theoretically, solubility and aggregation are totally different marvels, because they are reversible and irreversible processes, respectively [26, 27]. At present, many *in silico* tools are accessible to predict aggregation rate and aggregation pone regions.

3.3. Prediction of chemical stability:

In therapeutic antibodies prediction, many *in silico* approaches have been suggested to evaluate the chemical stability [6, 28, 29]. One of the most frequent degradation procedures is the chemical modification of Asparagine and Aspartic acid residues, which have the same degradation pathway [6]. There is no protocol to predict such degradation. Yet, there are statistical-based and physics-based approaches; the first one is experimental data dependent, combined with homology modeling [28]; while the second approach is superior to the first one by quantum mechanical calculations dependent without the need of experimental data [30].

3.4. Prediction and designing of immunogenicity of antibodies:

The prediction of high physicochemical properties such as high stability that a chief complete unfolding which is critical factor in therapeutic antibodies for good immune responses in patients. Due to the recent advances in computational immunology there are various *in silico* tools which facilitated to predict and reduce protein immunogenicity as much as possible [31-34] **(Table 3)**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

4. Perspectives:

In the last decade, antibodies engineering by computational approach is draws so much attention by guiding and accelerating the process of antibodies engineering by combination of selection system and use of high-throughput data acquisition and analysis to build and construct populations of next generation antibodies that are thermo-stable, non-immunogenic as possible, and to be administered to many humans as possible [6-8]. Which it's strongly suggested that present of homology modeling-based antibody techniques are dependable enough to be used in high-throughput, sequence-based computational analysis [35]. Yet, CDR-H3 structure prediction is still challenging. Since antibodies function has highlighted on CDR-H3diversity, to engineer a better functional antibody. CDR-H3 structure prediction approaches and antibody-antigen complexes must be developed. Another critical area is need to be improved in the line of antibodies engineering by computational approach is proteins flexibility.

To conclude, Despite of there is no *in silico* unified protocol for antibodies engineering, although computational approach it remain an indispensable method for antibodies designing. There is no doubt that the combination of *in silico and* library-based method will assist in production of antibodies therapeutics to assist the immune system in the battle against all types of life threatening and debilitating disorders.

Conflicts of Interest: The author declares that there are no conflicts of interest regarding the publication of this paper.

References:

[1] W. Si, C. Li, and P. Wei, "Synthetic immunology: T-cell engineering and adoptive immunotherapy," *Synthetic and Systems Biotechnology,* vol. 3, 09/01 2018.

[2] B. Geering and M. Fussenegger, "Synthetic immunology: Modulating the human immune system," *Trends in biotechnology,* vol. 33, 11/11 2014.

[3] S. Singh, *Systems and Synthetic Immunology*, 2020.

[4] D. Ecker, S. Jones, and H. Levine, "The Therapeutic Monoclonal Antibody Market," *mAbs,* vol. 7, 12/20 2014.

[5] J. Almagro, A. Teplyakov, J. Luo, R. Sweet, S. Kodangattil, F. Hernandez-Guzman*, et al.*, "Second antibody modeling assessment (AMA-II)," *Proteins,* vol. 82, 08/01 2014.

[6] S. Kumar, N. Plotnikov, J. Rouse, and S. Singh, "Biopharmaceutical Informatics: Supporting biologic drug development via molecular modelling and informatics," *Journal of Pharmacy and Pharmacology,* vol. 70, 02/01 2017.

[7] J. Datta, *MOLECULAR MODELLING & SIMULATION : A REVIEW FROM MOLECULAR INFORMATICS*, 2020.

[8] M. Viola, J. Sequeira, R. Seiça, F. Veiga, J. Serra, A. C. Santos*, et al.*, "Subcutaneous delivery of monoclonal antibodies: How do we get there?," *Journal of Controlled Release,* vol. 286, 08/01 2018.

[9] K. Locker and A. Herr, "Antibodies: Structure and Immune Effector Functions," ed, 2020.

[10] R. Stanfield and I. Wilson, "Antibody Structure," *Microbiology Spectrum,* vol. 2, 04/18 2014.

[11] J. Almagro, M. Beavers, F. Hernandez-Guzman, J. Maier, J. Shaulsky, K. Butenhof*, et al.*, "Antibody Modeling Assessment," *Proteins,* vol. 79, pp. 3050-66, 11/01 2011.

[12] A. Teplyakov, J. Luo, G. Obmolova, T. Malia, R. Sweet, R. Stanfield*, et al.*, "Antibody modeling assessment II: Structures and Models," *Proteins: Structure, Function, and Bioinformatics,* vol. 82, 08/01 2014.

[13] D. Kuroda, H. Shirai, M. Jacobson, and H. Nakamura, "Computer-aided antibody design," *Protein engineering, design & selection : PEDS,* vol. 25, pp. 507-22, 06/02 2012.

[14] T. Farhadi, A. Fakharian, and S. Hashemian, "Affinity Improvement of a Humanized Antiviral Antibody by Structure-Based Computational Design," *International Journal of Peptide Research and Therapeutics,* vol. 25, 12/11 2017.

[15] D. Kuroda and K. Tsumoto, "Engineering Stability, Viscosity, and Immunogenicity of Antibodies by Computational Design," *Journal of Pharmaceutical Sciences,* vol. 109, 01/17 2020.

[16] J. Zhao, R. Nussinov, W. J. Wu, and B. Ma, "In Silico Methods in Antibody Design," *Antibodies,* vol. 7, p. 22, 06/29 2018.

[17] T. Böldicke, *Antibody Engineering*, 2018.

[18] P. Nichols, L. Li, S. Kumar, P. Buck, S. Singh, S. Goswami*, et al.*, "Rational design of viscosity reducing mutants of a monoclonal antibody: Hydrophobic versus electrostatic inter-molecular interactions," *mAbs,* vol. 7, pp. 212-230, 01/03 2015.

[19] J. M. Perchiacca and P. M. Tessier, "Engineering Aggregation-Resistant Antibodies," *Annual Review of Chemical and Biomolecular Engineering,* vol. 3, pp. 263-286, 2012.

[20] J. Perchiacca, M. Bhattacharya, and P. Tessier, "Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions," *Protein engineering, design & selection : PEDS,* vol. 25, pp. 591-602, 07/27 2012.

[21] N. Pham and W. Meng, "Protein Aggregation and Immunogenicity of Biotherapeutics," *International Journal of Pharmaceutics,* vol. 585, p. 119523, 06/01 2020.

[22] B. Dahiyat and S. Mayo, "De Novo Protein Design: Fully Automated Sequence Selection," *Science (New York, N.Y.),* vol. 278, pp. 82-7, 11/01 1997.

[23] B. Kuhlman and D. Baker, "Native protein sequences are close to optimal for their structures," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 97, pp. 10383-8, 10/01 2000.

[24] R. Kramer, V. Shende, N. Motl, C. Pace, and J. M. Scholtz, "Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility," *Biophysical journal,* vol. 102, pp. 1907-15, 04/18 2012.

[25] L. Li, S. Kumar, P. Buck, C. Burns, J. Lavoie, S. Singh*, et al.*, "Concentration Dependent Viscosity of Monoclonal Antibody Solutions: Explaining Experimental Behavior in Terms of Molecular Properties," *Pharmaceutical research,* vol. 31, 06/07 2014.

[26] F. Agostini, M. Vendruscolo, and G. Tartaglia, "Sequence-Based Prediction of Protein Solubility," *Journal of molecular biology,* vol. 421, pp. 237-41, 12/09 2011.

[27] P. Sormanni, F. Aprile, and M. Vendruscolo, "The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility," *Journal of Molecular Biology,* vol. 427, 10/14 2014.

[28] J. Sydow, F. Lipsmeier, V. Larraillet, M. Hilger, B. Mautz, M. Mølhøj*, et al.*, "Structure-Based Prediction of Asparagine and Aspartate Degradation Sites in Antibody Variable Regions," *PloS one,* vol. 9, p. e100736, 06/24 2014.

[29] J. Aledo, F. R. Cantón, and F. Veredas, "A machine learning approach for predicting methionine oxidation sites," *BMC Bioinformatics,* vol. 18, 12/01 2017.

[30] N. Plotnikov, S. Singh, J. Rouse, and S. Kumar, "Quantifying Risks of Asparagine Deamidation and Aspartate Isomerization in Biopharmaceuticals by Computing Reaction Free Energy Surfaces," *The journal of physical chemistry. B,* vol. 121, 01/04 2017.

[31] K. R. Abhinandan and A. c. r. Martin, "Analyzing the “Degree of Humanness” of Antibody Sequences," *Journal of molecular biology,* vol. 369, pp. 852-62, 07/01 2007.

[32] S. Gao, K. Huang, H. Tu, and A. Adler, "Monoclonal antibody humanness score and its applications," *BMC biotechnology,* vol. 13, p. 55, 07/05 2013.

[33] P. Olimpieri, P. Marcatili, and A. Tramontano, "Tabhu: Tools for antibody humanization," *Bioinformatics (Oxford, England),* vol. 31, 10/09 2014.

[34] Y. Choi, D. Verma, K. Griswold, and C. Bailey-Kellogg, "EpiSweep: Computationally Driven Reengineering of Therapeutic Proteins to Reduce Immunogenicity While Maintaining Function." vol. 1529, ed, 2017, pp. 375-398.

[35] D. Kuroda, H. Shirai, M. Kobori, and H. Nakamura, "Structural classification of CDR-H3 revisited: A lesson in antibody modeling," *Proteins,* vol. 73, pp. 608-20, 11/15 2008.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

Table 1: Features Used in the Machine Learning Model DeepDDG for Predicting Thermo-stability:

Categories	Features
Sequence-based features	Amino acid types
	Protein design probability
	Position-specific scoring matrix
	Fitness score derived from a multiple sequence alignment
Structure-based features	Backbone dihedral angles
	Secondary structures
	Solvent-accessible surface area
	Number of hydrogen bonds
	Distance and orientation between the mutated residues and the neighboring residues

Table 2: Features used in the Machine Learning Model SOLart for Predicting Solubility:

Categories	Features
Sequence-based features	Amino acid compositions
	Protein length
	Secondary structures
Structure-based features	Solubility-dependent statistical potentials
	Secondary structures
	Solvent-accessible surface area

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

Table 3: *In silico* approaches to evaluate, predict, and reduce the immunogenicity of antibodies:

Assessment of Immunogenicity	Description	URL
T20 score analyzer	Humanness score based on sequence identity to the top 20 matched human antibody sequences	https://dm.lakepharma.com/bioinformatics/
SHAB	Humanness score based on sequence identity to human antibody sequences	http://www.bioinf.org.uk/abs/shab/
Humanization of antibodies Tabhu	Framework template search followed by CDR grafting with back mutations	http://www.biocomputing.it/tabhu

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

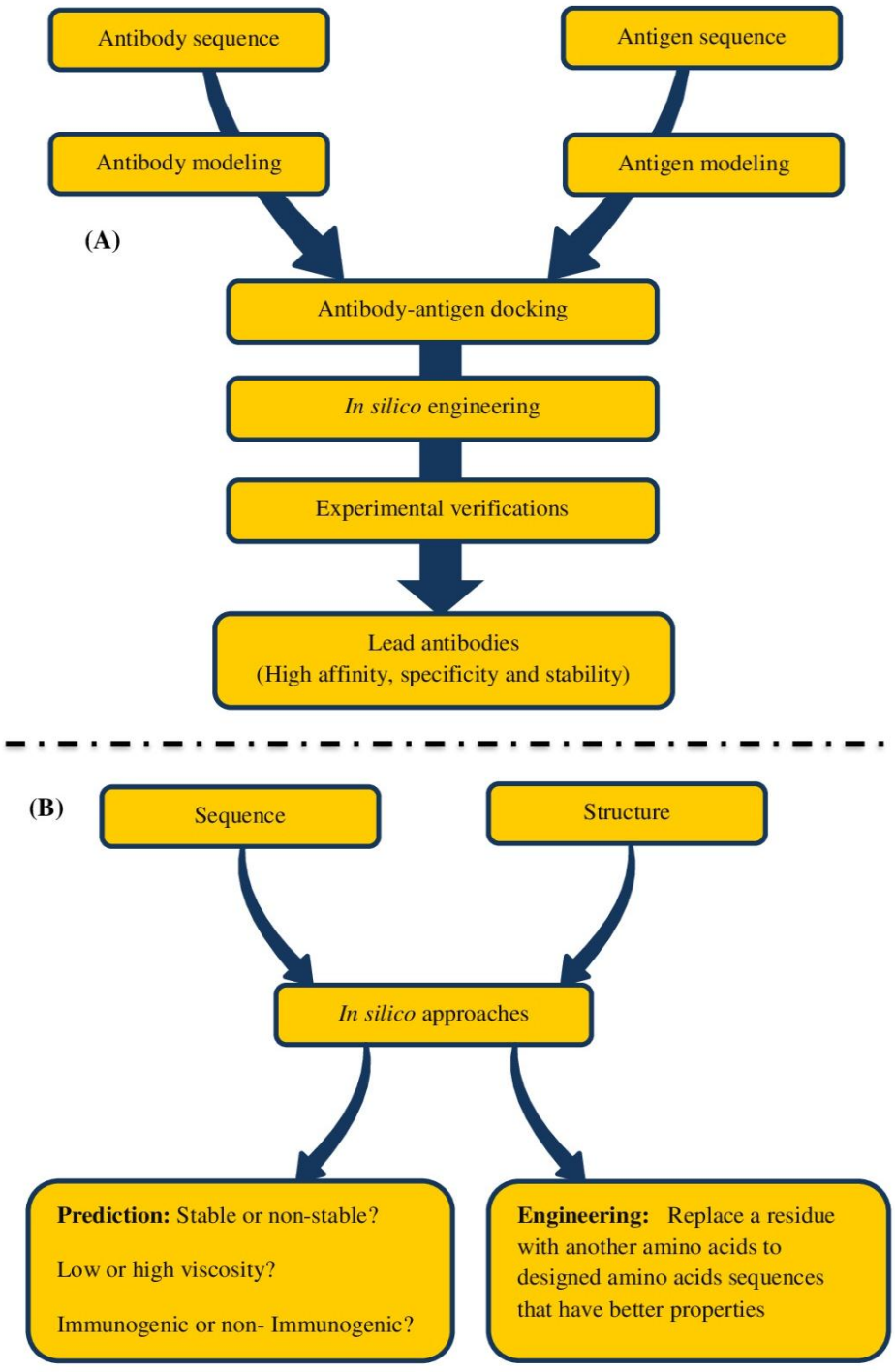

Figure 1: **(A)** Development procedure of antibodies engineering and the roles of computational approach.. **(B)** Main workflow of computational approach.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 June 2020 doi:10.20944/preprints202006.0350.v1

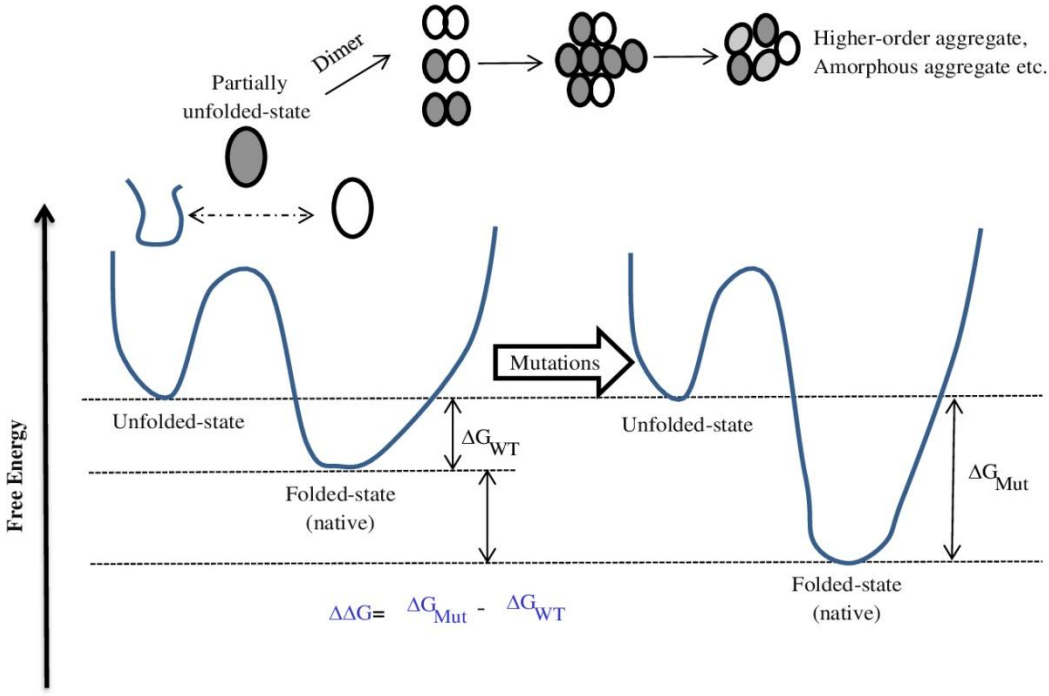

Figure 2: Equilibrium between folded and unfolded proteins and its relation to protein stability