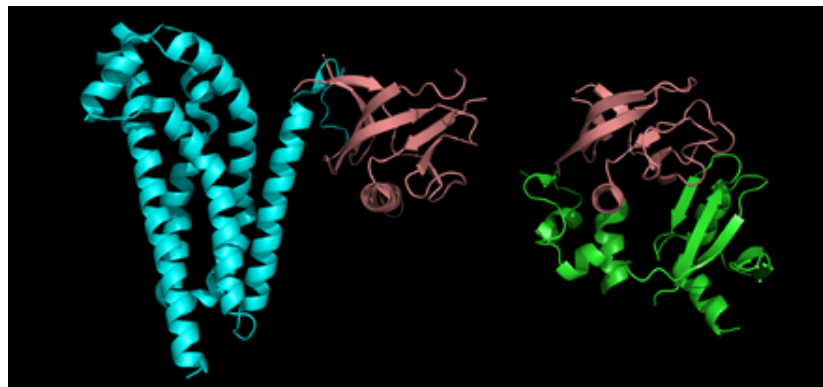


Next-generation sequencing technologies

Source: [Biochemical Society Transactions | Portland Press](#)

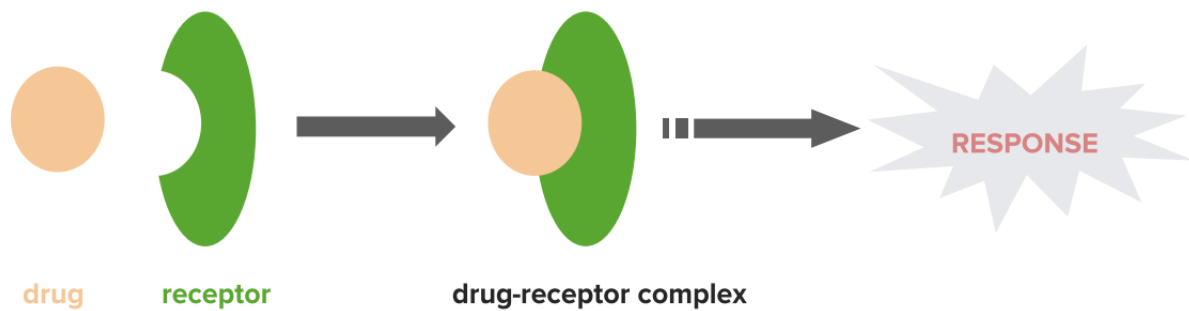
The application of UD-MIS for choosing non redundant representative sets of protein from sequence data sets can be served in a variety of applications. For instance, selecting representative sets of protein structure which is one of the biology problems in the world.



Two protein structure

Source: https://www.wikiwand.com/en/Protein_structure

Protein structure is the three-dimensional arrangement of atoms in an amino acid-chain molecule. It enables us to understand ligand protein binding and develop binding affinity techniques [1]. Binding affinity is the strength of the binding interaction between a single biomolecule (e.g. protein or DNA) to its ligand/binding partner (e.g. drug or inhibitor) [2].



Drug receptor

Source: <https://www.lecturio.com/magazine/biological-interaction/>

Business pitch.

Are you thinking how to increase your market share? Do you want to make more effective medicine with less side effects to the human body? Here is our reliable solution:

It becomes possible to firstly predict how and where a particular small molecule might interact with a protein and then to identify putative ligands for a specific protein site [1].

Targeting specific chosen protein to evaluate ligand affinity can be done by selecting and analyzing representative sets of protein structure. Target-oriented drug design aims to find a high affinity ligand that would bind the target protein in order to block its disease-associated function catalytic activity or interaction with other molecules.

Structural knowledge on the exact interactions of drugs with their target protein has been applied mainly to predict potency changes of chemically modified lead compounds in protein structure-based drug design industry [3][4].

Drug design industry overview:



Drug discovery industry

Source: <https://physicsworld.com/a/intelligent-drug-discovery/>

In recent years, the drug discovery technologies industry has seen remarkable developments and rapid progress. Biology, chemistry and computation provide an amalgamation that is reflective of the industry's future. In this industry, growth in the protein sequencing market is mainly driven by factors such as increasing focus on target-based drug development by pharmaceutical and biotechnology companies, advancements in clinical mass spectrometry and analytical techniques, and increasing public-private financial support for proteomic research [5].

Supply side of protein sequencing market is characterized by advancement in technology and technique. The demand side of this market is characterized by increasing use of proteomic research in drug design industry, patent expiry of major blockbuster drugs and increasing R&D expenditure of major pharmaceutical and biotechnology companies [5].



From structure to design drug

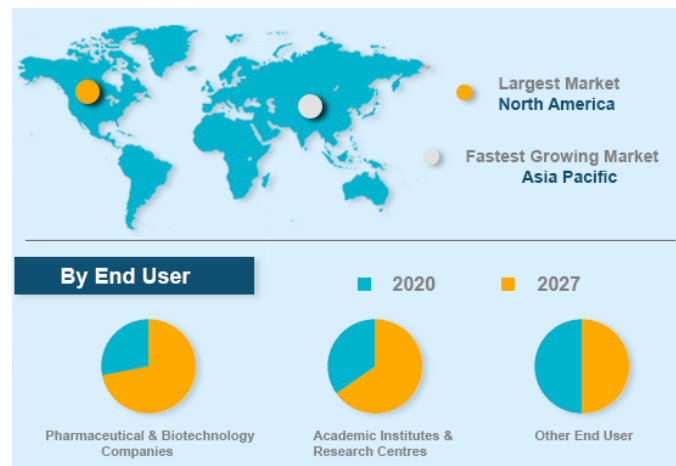
Source:

<https://www.ddw-online.com/target-discovery-and-drug-design-extracting-the-value-from-genomics-1421-200008/>

The 10 largest pharmaceutical companies now pay nearly \$80 billion a year to come up with fewer and fewer successful drugs. The reason is that drugs that are easiest to find and that safely and effectively treat common disorders have all been found; what is left is hunting for drugs that address problems with complex and elusive solutions and that would treat disorders affecting only tiny portions of the population and thus could return far less in revenue. Because finding new, successful drugs has become so much harder, the average cost of bringing one to market nearly doubled between 2003 and 2013 to \$2.6 billion. These same challenges have increased the lab-to-market timeline to 12 years, with 90 percent of drugs washing out in one of the phases of human trials. Many scientists in the field think that computer-aided techniques will ultimately improve drug development in several ways such as identifying more promising drug candidates and by speeding up the overall process.

Several biotechnology and pharmaceutical companies are making sizeable investments on utilizing protein sequencing in drug development. That's why Every one of the major pharmaceutical companies has announced a partnership with at least one of the AI-based drug-discovery start-ups. They raised more than \$1 billion in funding in 2018, and as of last September, they were on track to raise \$1.5 billion in 2019[6].

According to the market research reports, protein sequencing industry garnered \$5.39 in 2018 and it is estimated to reach \$9.92 billion by 2026, growing at a compound annual growth rate (CAGR) of 7.8% from 2019 to 2026 [7]. The global market for drug discovery should grow from \$69.8 billion in 2020 to \$110.4 billion by 2025 with a CAGR of 9.6% for the period of 2020-2025 [8]. Of all the regions, North America holds an enormous potential in the protein sequencing market. Several public and private organizations have committed large funds on proteomic research. Rising focus of players on structure-based drug design and the advent of technologically advanced analytical methods used in protein sequencing are bolstering the regional demand [9].



Market segmentation

Source:

<https://www.futurewiseresearch.com/healthcare-market-research/Protein-Sequencing-Market/3383>

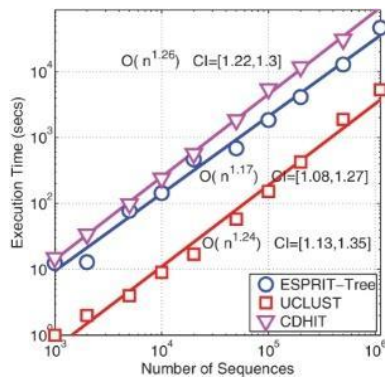
Benchmarking:

There are several conventional methods to select non-redundant representative subsets of protein sequence. Our technology has dramatically promoted sequencing technologies by offering low-cost and ultra-high-throughput sequencing. In addition to accuracy, computational complexity is an important issue.

Previous methods for this task, such as CD-HIT, PISCES and UCLUST, apply a heuristic threshold-based algorithm. Next-generation sequencing technologies are able to demonstrate a good speedup derived from solving UD-MIS problem quantumly compared with conventional famous methods based on the Threshold algorithm.

Table: Computational Complexity (time)

| Conventional Method | US-MIS Quantum |



Source: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3251834/>

Computational Costs: Our technologies to provide non-redundant representative subsets of protein sequence reduces computational complexity which result in a lower computational cost. It can cut the time spent in discovery from 4.5 years to as little as one year, reduces discovery costs by 80 percent and results in one-fifth the number of synthesized compounds as is normally needed to produce a single winning drug.

The “production” costs are accounted for [10]:

- Labor, administration, management, utilities, reagents, and consumables
- Sequencing instruments and other large equipment (amortized over three years)
- Informatics activities directly related to sequence production (e.g., laboratory information management systems and initial data processing)
- Indirect Costs as they relate to the above items.

Competitors:

Several top players are aimed at launching cutting-edge mass spectrometry platforms to gain unique competitive edge over their peers in the global protein sequencing market. Some prominent companies are also consolidating their overall presence by strengthening the sales of reagents and consumables. Top players operating in the market include Thermo Fisher Scientific, Charles River Laboratories, and Shimadzu Corporation [9].

| Company name | Country | R&D budget | Profit |

Company name: Thermo Fisher Scientific

Country: USA

R&D budget: \$ 1.18 billion - 2020

Profit: \$ 32.22 billion – 2020

Company name: Shimadzu Corporation

Country: Japan

R&D budget: NA

Profit: 3.710 million – 2020

Company name: Charles River Laboratories

Country: USA

R&D budget: \$ NA

Profit: \$ 571.2 million – 2020

Mapping our competitive position:

Mapping the competitive position of the company helps us to identify what we are charging for and what our customers are paying for. It shows us the primary benefit that our product offers and also our rival's core competitiveness [12]. Creating such a map involves three steps:

1. Define the market.

People's need is not only to cure their disease, but to get a medicine or drug which is more effective while it has fewer side effects. That is why people are very concern about their medications. To address their needs, pharmaceutical companies need to invest more to improve effectiveness and safety of their drugs. To this end, they require reliable and accurate solution from drug discovery industry. This is where protein sequencing technology companies could play vital role to tackle the problem.

2. Determine the primary benefit

The primary benefits of our product would be the accuracy of the protein structure identification and shorten the drug discovery process. Our product reduces discovery costs by 80 percent. Moreover, highly skilled engineer and scientists guarantee the reliability and validity of the product.

3. Plot position

As it is mentioned before, there are some competitors in the market including Thermo Fisher Scientific. SWOT analysis shows that this company has several strengths such as successful track record of developing new products, Reliable suppliers, strong dealer community, strong brand portfolio and High level of customer satisfaction [12]. On the other hand, there are some weaknesses such as, gaps in the product range sold by the company, days inventory is high compared to the other competitors, not very good at product demand forecasting [12]. It can be concluded that as a start-up company we might be threatened from the competitors' strengths such as strong brand portfolio and strong dealer community. However, the company can benefit from several business opportunities. For instance, our highly skilled engineer and scientists could forecast demand using latest techniques. This leads customer satisfaction improvement, reduction in operating cost and higher operating profit. Moreover, accuracy of our solution leads stronger brand portfolio in long term. In conclusion, the first five years would be very critical for our company although its growth rate could be considerable during this period if the resources are allocated precisely and efficiently towards the drugs that address problems with complex and elusive solutions.

Potential customer:

| Company name | Country | R&D budget | Profit |

Company name: Pfizer

Country: USA

R&D budget: \$ 9.4 billion - 2020

Profit: \$ 41.9 billion – 2020

Company name: AstraZeneca

Country: United Kingdom

R&D budget: \$ 6 billion - 2020

Profit: \$ 26.6 billion – 2020

Company name: Regeneron pharmaceuticals:

Country: USA

R&D budget: \$ 2.7 billion - 2020

Profit: \$ 8.497 billion – 2020

Company name: Sygnature Discovery

Country: United Kingdom

R&D budget: NA

Profit: \$ 68.9 million – 2020

[1]

<https://portlandpress.com/biochemsoctrans/article-abstract/46/5/1367/67796/Protein-structure-and-computational-drug-discovery?redirectedFrom=fulltext>

[2] [Binding Affinity | Dissociation Constant | Malvern Panalytical](#)

[3] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4236214/>

[4] [Frontiers | Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview | Pharmacology \(frontiersin.org\)](#)

[5]

<https://www.marketsandmarkets.com/Market-Reports/protein-sequencing-market-121629357.html>

[6] [Hunting for New Drugs with AI \(nature.com\)](#)

[7]

<https://www.medgadget.com/2020/08/protein-sequencing-market-to-reach-9-92-billion-by-2026-allied-market-research.html>

[8]

<https://www.bccresearch.com/market-research/biotechnology/drug-discovery-technologies-report.html>

[9]

<https://www.biospace.com/article/protein-sequencing-market-increasing-focus-on-target-based-drug-development-activities-to-drive-the-market/>

[10] <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>

[11] <https://hbr.org/2007/11/mapping-your-competitive-position>

[12] <http://fernfortuniversity.com/term-papers/swot/1433/164-thermo-fisher-scientific.php>