

How did Bio-Product develop the best mutation AI prediction platform in the world?

3DM Engineering is worldwide the first AI platform that can predict which mutations should be combined (5-7 protein mutations) to massively increase protein performance. Between 10% to 40% of the predicted protein combinations outperform current expensive and laborious laboratory tests. Founder and CCO Henk-Jan Joosten gives his insight into the unique story of Bio-Product.

In 2000, Henk-Jan Joosten, founder and CCO of Bio-Product did a Masters Degree at the Bio-Informatics Department of The Radboud University. Together with two other students, he built a database that contained as much data they could find for the protein family of the nuclear hormone receptors. This was on request of the pharmaceutical company Organon. After this database was completed, it turned out to be very useful to have all data of one protein family together in one database.

Bio-Product; A company that enables researchers to build smart solutions with protein family databases (3DM):

As Dr. Joosten has been involved in the complete process of building this first protein family database, he realized that the whole process can be automated and therefore a program could be made that can automatically build databases for all existing protein families. The idea of a company that would sell licenses to protein family databases was born. The 5th of May 2008 Dr. Joosten started Bio-Product, a company from which pharmaceutical and biotech companies could license access to the, so called, 3DM databases for any protein target that the customer is working on.

Solving problems: the 3DM database and its unique 3D protein numbering scheme

The 3DM databases of Bio-Product solve a huge problem in protein research and engineering; in 3DM all proteins are linked to each other via a unified numbering scheme, called 3D numbers. Consequently, all sequences and protein structures of all organisms in a protein family are renumbered, based on the 3D positions of each residue, which hyperlinks all sequences, all structures, all underlying data, and all 3DM tools that have been developed to analyze all this curated data.

Until this day, no other company in the world has such a sophisticated protein numbering system, making 3DM the largest storage of hyperlinked high quality protein data in the world. Already with the first 3DM system, important discoveries have been made about the functioning of these nuclear receptors, which resulted in research papers published in peer reviewed high impact journals.

Used across multiple types of protein related research and engineering

At the start of Bio-Product it still took about 3 months to generate a 3DM database for one protein family with these systems containing so much useful data for scientist working in many different types of protein related R&D, they would become a valuable tool in the researcher's arsenal with Bio-Product becoming a valuable player in the world of protein research. First Bio-Product targeted the protein engineering market for which tools have been developed that use 3DM data.

In 2011 Bio-Product released a new revolutionary tool that can be used for designing smart mutant libraries. This tool has successfully been used in protein research resulting in dozens of papers and patents (www.bio-product.nl). In 2014 a tool for the analysis of patent data available for complete protein families was released. Because the process is automated, 3DM finds many more patents about one protein family, compared to manual searches done by patent attorneys. As all sequences in 3DM are synchronized via the 3D-numbers, including the once claimed in patents, patent attorneys can now transfer claimed mutations from all patented proteins in the superfamily to the target protein of a customer. Until today, this 3DM feature remains unique-in-the-world.

AI platforms Helix Pathogenicity and 3DM Engineering opening new markets

Since 2018 Bio-Product has built 3DM databases for all possible known structural protein families (~55.000 3DM databases) and targeted 3DM systems that cover the full-length of all human proteins (~100.000 3DM databases). This human protein targeted set was requested by pharmaceutical customers, giving them instant access to all important information that is

captured in all these 3DM systems for every single amino acid of all human proteins.

This immense data set of high-quality hyperlinked protein data enabled Bio-Product to train a revolutionary AI platform, called Helix, that can predict if a human mutation might cause a disease. After years of training Helix can predict if a mutation can cause a disease with 97% overall accuracy, much better than any other AI tool currently available in the market. It even outperforms the recently published Google's Alpha Missence when we used an independent set of 4000 mutations available in hospitals in the BRCA1 protein that can lead to breast cancer. As Google didn't have this set, we were able to blind test the performance of Alpha Missence and predicted just 79% of these mutations correctly.

In November 2024 Bio-Product released another revolutionary AI tool, called 3DM Engineering, that can accurately predict how to combine mutations to make big leaps in the performance of proteins. In a first pilot we compared the performance of 3DM Engineering to a very costly high throughput optimization project of a customer. This showed that with 3DM Engineering, they could have saved 98.5% of labwork, time and costs. Its performance was compared to the test using data of a benchmark of the Protein Engineering Tournament 2024 and 3DM Engineering outperformed all competitors. All projects done so far show that AI, when trained on massive amounts of highly organized data, as captured in 3DM systems, can design better performing proteins, even on already highly optimized proteins that customers have been optimizing for more than a decade. Bio-Product was recently awarded the **Global Excellence Award: "Best Mutation Prediction Company 2025"** given by GHP. Henk-Jan Joosten believes that with 3DM Engineering we are on the way to disrupt how protein related R&D is done today in fields such as protein engineering, drug design, and disease diagnostics.

