# The Power of STB CLOUD in Remote AI Drug Discovery

April 27, 2023    by Sabina Lee | Sponsored by **Syntekabio**

STB CLOUD is Syntekabio's latest cloud-based, fully automated AI platform. It's designed for developers to engage in their drug discovery and development processes to find drug candidates easier and faster.

Early-stage drug development involves technical difficulty tiers. For example, rapid virtual screening is an absolute must for handling billions of compounds. While optimal poses by artificial intelligence deep learning methods are required, optimal electrostatic and proton arrangements between protein pockets and ligands are still required to apply sensitive free energy or empirical binding energy. After all, MD simulation is just such a process. Therefore, ranking according to the exact binding energy in the optimal pose state is possible, which leads to improvement of AI hit success rate after in vitro validation.

That's not all. It should be able to address pocket diversity according to different protein pockets in terms of shape, position and topology, including shallow pockets, deep pockets, sub-pockets and pockets in multi-chains. After ADME/Tox, one can finally discuss whether or not it's a blockbuster drug candidate.

Traditionally, this level of early-stage drug development takes five years or more. However, now with the help of the AI ■■drug platform, many parts can be performed with automation in an environment with minimal human error. Then only verification experiments are required for the final confirmation. In addition, even if one or two failures occur during the course, it is possible to reach the pre-clinical stage in two years.
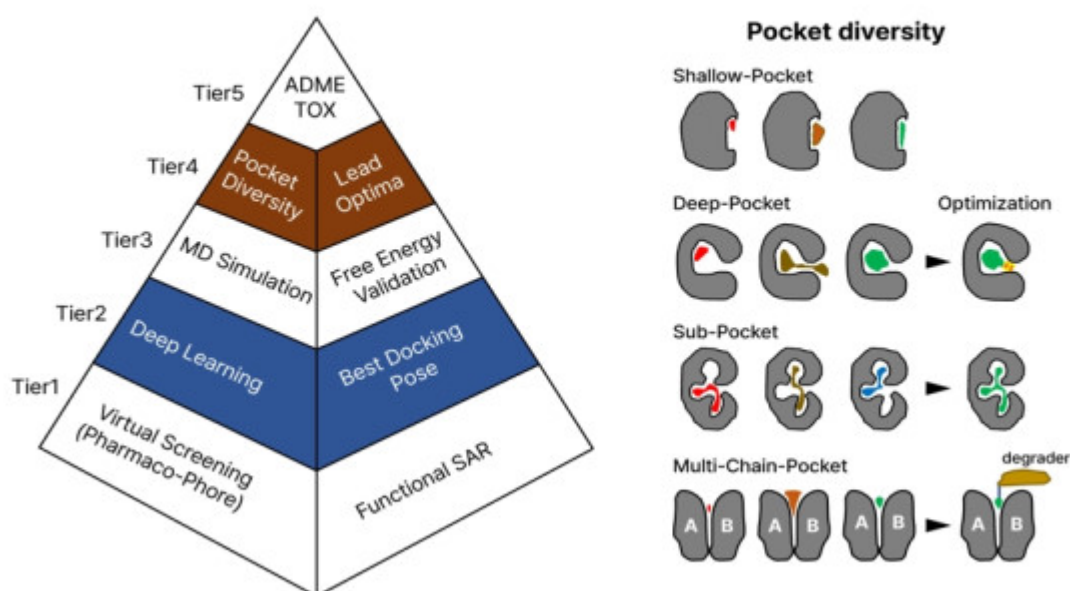
*Figure 1. AI-Drug Platform Technical Difficulty Tiers*

The figure shows the five technical difficulties and development order of the AI ■■drug platform. Here, virtual screening in Tier 1 is a step of screening 1 billion compounds in an efficient way suitable for a given protein pocket, and AI-deep learning in Tier 2 is to find the most suitable pose by applying the learned effect between the protein pocket and ligand. MD simulation at Tier3 is to create protonation, electrostatic arrangement and equilibrium pose between protein and ligand, at Tier4 pocket diversity is the selection aspect in terms of effectiveness of drugs in the ranking, and Tier 5, ADME/Tox relates to absorption, distribution, metabolism, excretion and toxicity.

# Convergence of BIO and IT technology in One: STB CLOUD

STB CLOUD is a platform that automatically integrates and operates heterogeneous hardware and hundreds of bio software. The system prepares a solution to handle bio big data with a slum system to

simultaneously operate supercomputing hardware and parallel distribution hardware. Then Kubernetes, an open-source container orchestration system originally designed by Google, manages multiple operational tasks of containerized applications that are run via a workflow environment, such as container distribution and operation management, including running container environment and docker for container operation.

Bio software is designed to handle specific bio data but it cannot be a solution service by itself. It can function as a service platform only when it is integrated with various heterogeneous servers—CPU, GPU, etc.—and IT technology. In this synergistic fusion of bio software and bio big data, STB CLOUD works as a complete solution service driven by its core product 'DeepMatcher™-Hit.'

DeepMatcher™ is a three-step process. It screens 1 billion compounds for each target protein within hundreds of CPUs and GPUs that operate concurrently, first in the virtual pre-screening (DMC-Pre), then the deep screening step (DMC-SCR) involving the optimal protein-ligand creating a pose. Lastly, hundreds of GPUs are operated for a week in the molecular dynamics verification (DMC-MD) stage to derive active substances.

Composed of hundreds of scripts running fully automatically, DeepMatcher™ utilizes cutting-edge IT technologies such as Kubernetes, containers and docker, which together effectively functions as the administrator, operator and executor all in one.

STB CLOUD has 3,000 GPUs and CPUs bundled in the cloud with DeepMatcher™ as its core processor. It automatically generates a report of new drug candidate substances 'AI-Hit' at the click of a button. AI-Hit, a drug active substance known for a high success rate in the experiment, is verified and processed automatically for drug optimization to produce an 'AI-Lead.'
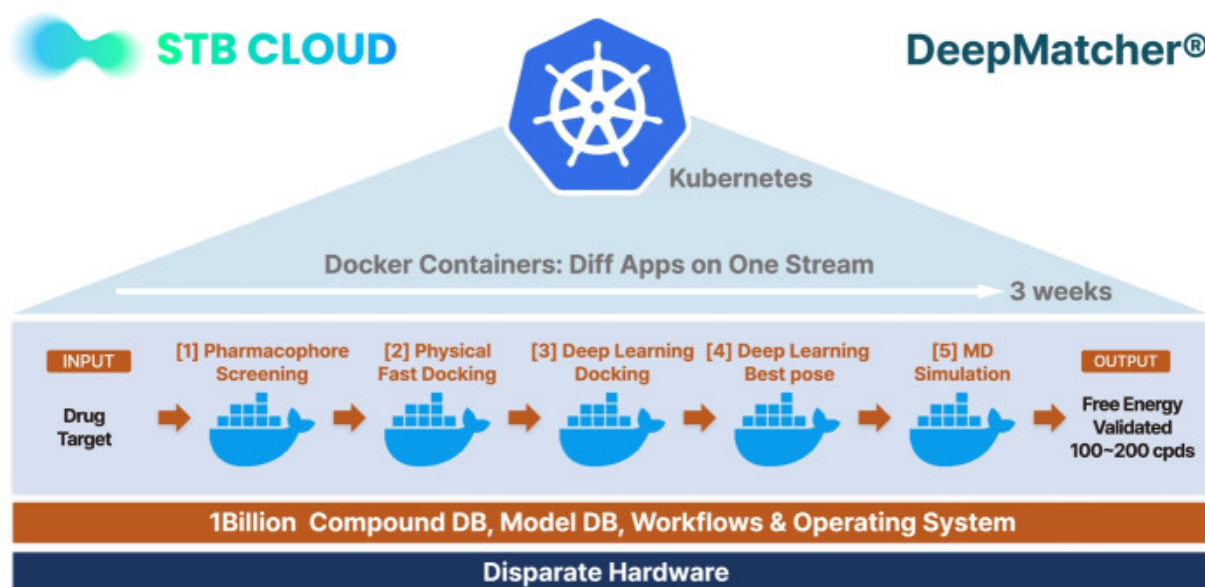
*Figure 2. Docker container-based DeepMatcher platform.*

STB CLOUD mainly builds standard Kubenetes based on the supercomputing processing environment, and converts the DeepMatcher drug discovery platform script into an image (pod) in Docker to achieve distributed processing and continuity with supercomputing technology using the Kubernetes scheduler. In particular, 5 huge Dockers were made to operate in the form of each container, and the whole is in one stream. Another feature of these Docker images is that each proceeds while communicating to GPU/GPU and storage, etc., and operates without errors on hundreds of nodes under the management of Kubernetes.

# STB CLOUD Difference

Existing AI drug discovery platforms so far provide a software-centric service, which requires a hardware—a supercomputer—to run on and trained specialized personnel to operate. However, STB CLOUD operates as a fully automated, cloud-based AI drug development platform without needing software or hardware to execute the process.

Software as a service (SaaS) is the dominant form of AI drug platforms most commonly available in the marketplace. In SaaS, researchers usually use software provided by an AI drug developer to calculate

interactions between target proteins and compounds. Although the SaaS environment doesn't demand much resources from the client, it comes with a caveat: researchers operating the SaaS platform must be trained experts in software, data servers and data libraries with serious drug design capabilities.

This is where STB CLOUD, unlike SaaS, performs better as a simplified standard solution, allowing anyone to manage the full cycle of deriving active substance candidates conveniently and accurately. This enables users to find a candidate for a small molecular drug active substance—Hit—that has binding affinity at the level of micro molar to sub nano molar. Within three weeks, the user can select the disease-related target protein by selecting the protein of their choice.

Anyone with a computer connected to Internet can now develop new drugs in a few simple clicks, anytime, anywhere. STB CLOUD gives pharmaceutical companies and researchers freedom from having to be at specific locations and a huge advantage over the new drug development cost and time, while minimizing uncertainty of the process outcome. A new era of drug development possibilities seems limitless as biotech companies can quickly expand new drug pipelines through STB CLOUD, including innovative first-in-class drugs.

Built on its own AI supercomputing infrastructure with 3,000 units of high-performance servers, STB CLOUD is powered by deep learning through DeepMatcher™, the culmination of Syntekabio's software, hardware and middleware technologies.

# Evolution of STB CLOUD for AI drug discovery

With commercialization of STB CLOUD, the neo-antigen discovery platform 'NEO-ARS™' is planned to be installed on the cloud in addition to DeepMatcher™. NEO-ARS™ can predict neo-antigens, which are tumor-specific mutant peptides obtained from cells or blood samples of cancer patients. Unlike other algorithms, it is a platform that predicts immunogenicity by predicting not only the binding affinity between MHC proteins and neo-antigens, but also the binding environment between T-cell receptors and neo-antigens.

Syntekabio has various platforms to be used throughout the entire life cycle of new drug development, such as DeepMatcher Auto-Lead-Opt, Tox/ADME, and AI-based antibody drug platform. These platforms, which are currently being serviced in the localized environment, will soon be loaded sequentially onto

STB CLOUD.

DeepMatcher™ Auto-Lead-Opt, NEO-ARS™ and new antibody drug platforms are together expected to demonstrate high-quality performance even beyond DeepMatcher™. Genome platforms such as PGM-ARS and NGS-ARS will be installed in the cloud later in 2023 to help with personalized precision medicine and clinical research, including biomarker discovery.

Standardization, simplification and automation of services crucial for efficient drug development were once available only in the R&D area in a limited capacity. Lately cloud services are seen as an indicator for company's competitiveness in technology and an essential component for major platforms' expansion of services. Today, drug discovery and development are no longer restricted geographically and can be utilized by those without the backing of big capital. Transforming the cloud-based AI drug development, STB CLOUD is enabling innovations across the globe leading the way firmly into the future.

- Syntekabio