Faster Drug Discovery Using Hardware Accelerated Protein Docking and Screening

The essential problem to be addressed in our innovation is to develop an affordable alternative to address the challenge of fully exploiting rapidly increasing protein and chemical structure libraries in the identification of new drugs. The volume of data for each of these libraries grows increasingly fast (doubling every 6 months) while processor speeds have leveled off. A new technology is required to prevent the need to build increasingly large and expensive clusters of computers to simply perform early stage drug candidate screening.

Affordable new therapies and pharmaceuticals are being increasing looked to as the solution for reducing the staggering growth of the nation's healthcare $1.6 trillion expenditure. Unfortunately, the cost of developing new pharmaceuticals is continuing to rise. The industrial response has been a reduction in numbers of new chemical entities, by one estimate, as low as 0.6 new drugs per pharmaceutical company per year. In the emerging era of personalized medicine, where a 'one size fits all' drug therapy will be a prescription of the past, the costs per treated patient will continue to rise. As solutions tailored to match specific biomarker profiles emerge, the need to characterize each drug's interactions will become increasingly paramount, increasing costs and causing a corresponding reduction in new drug treatment options.

Confounding the situation is the huge increase in the volume of information available to the pharmaceutical researcher. In the post genomic era, the quantity of new proteins identified and at least partially characterized continues to rise greatly. Researchers are awash in protein and molecular data from newly sequenced genomes, from combinatorial chemistry assays, and details of potential biomarkers indicating everything from disease to acute toxic exposure with no effective means to fully utilize it. By one estimate, the growth rate of new data is faster than doubling every 6 months, when accounting for derivative and proprietary databases.

The drug discovery pipeline relies heavily on increasingly ineffective technologies. The new volume of protein specific information, while an extremely valuable asset for identifying and characterizing new protein functions and interactions, has the unfortunate aspect of extending the time required to computationally screen new drugs using traditional algorithms and technologies. Adding to the volume of information is the increased role of robotically driven combinatorial chemistry and rapidly expanding candidate drug libraries.

Traditional techniques to identify candidate drugs among the combined datasets employ large numbers of general purpose processors (clusters), using off-the-shelf docking programs to identify candidates. As increases in processor speeds level off as the economically viable end of Moore's law is reached and demands for faster turnaround grow, a new approach is needed.

Accelerated Data Concepts (ADC) proposes an innovative technology where considerable commercial potential can be realized in both the drug discovery IP development as a licensed product and as a valuable service to provide integration to commercial database systems such as Oracle and to open source systems such as MySQL. This technology ultimately will lead to a high-speed chemical database search product based on a hardware approach of data searching, thus addressing the current challenges of searching complex chemical compound structures facing researchers today. By doing this, added cost is incurred by having to power these additional resources and to maintain them. By bringing in a ADC accelerated desktop solution, we not only lower the maintenance cost but we also lower the power requirement which will also provide a environmental friendly solution.

Scientific advantages will be realized by providing a seamless integration path between the application and the hardware. This will allow the scientist to focus on the science and drug discovery process rather than on the software and hardware interfacing and special tool development. At the moment there are no off the shelf high-speed drug screening solutions employing our proposed technology that are available where a researcher can simply plug in a complete package and start analyzing data sets.

Our technology will advance clinical research by providing the foundation for developing a desktop solution that can be adapted to real time analysis for a "Lab on a Chip" system based on the detection of nucleic acids. These lower cost desktop systems will find their way into areas of the world where state-of-the-art medical technology is not available but is desperately needed to help combat diseases such as AIDs and the latest bird flu epidemic.

By enabling scientist to do more with less overhead, further discoveries will be made that will enable Americas competitive edge in the market place through technology and meet further market needs in areas such as nanotechnology, supercomputing, and new drug development.

Market Potential

The target market for these types of products is estimated in the neighborhood of $2 billion annually. In recent report, the fully loaded cost for developing a new pharmaceutical was estimated at $897 million and increasing. With the major proportion of costs associated with drug development in later stage clinical trials (60%), improving productivity in the early stage of the pipeline (candidate identification) is critical. Faster identification expands the timeframe where a drug can be sold under patent. Better identification reduces the waste in failed clinical trials. Cheaper identification improves

profitability and shifts resources to improved early stage evaluations critical for safety evaluations.

By next year (2007), at least 35 blockbuster drugs representing $81 billion in annual sales will be coming off patent (according to Datamonitor, 2002). The combination of revenue potential for drugs while on patent, together with the depletion of the existing drug pipeline, makes the opportunity prime for solutions which can quickly identify new patentable drug candidates. Pharmaceutical companies, large and small, are motivated to keep their pipelines stocked with new, viable candidates identified as early as possible in the pipeline.

While the aggregate research investment in drug development increases, the pharmaceutical industry is undergoing transformation. As large pharmaceutical companies increasingly outsource their efforts, smaller companies are picking up the slack. The larger number of small companies greatly increases the market for the low-power, low cost ADC solution for high performance drug screening.

The present market is occupied commercially by vendors including Tripos, Accelrys, Schrodinger, and MDL delivering fully integrated solutions for drug discovery, including components for protein-ligand docking and scoring. These established vendors while potentially competitive, will also be developed as delivery channels for deploying the ADC based docking solution.

The ADC product will be delivered as a component to be integrated into solutions, both commercial and open source, at both the desktop and server level. The combination of component licensing with existing vendors, sales at the desktop level and sales of server integration components are expected to generate at least $1M within 18 months and $3M annually within 3 years.

Innovation – Accelerated protein molecule docking interaction

Among the most unpredictable molecules in nature are proteins. When genes are translated into proteins, this information becomes a three-dimensional problem of folds, helices, motifs, and sheets. This information holds much value for the chemist trying to create a drug for a specific target. For the chemist to make use of this information, an atom to atom three dimensional view is needed to ascertain how the ligand (drug) will dock into the target (protein) molecule.

Using ADC's innovative approach for this drug docking application allows the desktop computer system to simulate the in-silico drug discovery process as opposed to the wet science method used to find a potential new drug. The code, written in C, will run the simulation of the molecule interactions. These simulation runs will require algorithms to calculate the interaction energies of all the atoms of one molecule with all the atoms of the other molecule. This now becomes a massively parallel problem that is ideally suited for the ADC hardware. The algorithms specific to the atom to atom calculations that are required are done in hardware at a much higher speed than a standard CPU.