

SOLUTION BRIEF

Cryo-Electron Microscopy (Cryo-EM)
Intel® Scalable System Framework



Cryo-Electron Microscopy: Revolutionizing Drug Discovery Pipelines with 3D Proteins Imaged at Near-Atomic Resolutions

Accelerating Drug Discovery with Nobel Prize-Winning Cryo-Electron Microscopy

Jacques Dubochet (Switzerland), Joachim Frank (US), and Richard Henderson (UK) were awarded the Nobel Prize in Chemistry 2017 for their work in cryo-electron microscopy. Cryo-EM is “decisive for both the basic understanding of life’s chemistry and the development of pharmaceuticals” said the Nobel Committee, “biochemistry is now facing an explosive development and is all set for an exciting future.”³ Cryo-electron microscopy “fills an important gap and extends the range of molecules that can be determined at atomic resolution” according to Nobel recipient Frank.⁴

Cryo-Electron Microscopy (Cryo-EM) is a rapidly emerging technique used to construct three-dimensional models of proteins and molecules at near-atomic resolutions. The technique captures 2D images of protein molecules frozen in a thin layer of ice under cryogenic conditions using a transmission electron microscope (TEM). With the help of high-performance computing systems and optimized codes, thousands of these 2D images are then aligned and combined to form 3D models.

The structure of a protein determines how it interacts with drugs and other substances in the living system. Discovering and modeling a protein structure is therefore essential in streamlining the drug discovery process. Unlike in x-ray crystallography, cryo-EM doesn’t require crystallization; a long, arduous process that, when possible, can take as long as several years and often locks the protein in an unnatural conformation. Single-particle cryo-EM enables the study of molecules in near-native environments, making it an essential tool for understanding how drugs and their target proteins may interact in the body. The technique can be applied to particles in a wide range of sizes, from hemoglobin (64 kDa) to particles of several megadaltons.¹

While cryo-EM was invented in 1968, recent advances in technology have enabled it to regain traction, accurately modeling 3D proteins to a higher resolution than most x-ray crystallography procedures can achieve.² Technical advancements have increased scientific excitement for cryo-EM in the recent years, leading it to win Nature Journal’s Method of the Year award in 2015² and the Nobel Prize for chemistry in 2017.^{1,3}

Big Data in Cryo-EM

Recent advances in transmission electron microscopes and data analysis software have exponentially increased the data storage and computation requirements in cryo-EM. This presents some challenges:

- **Burdens of Data Management:** Petabytes of research and clinical data are being created which needs to be stored, managed, shared, and analyzed effectively and securely
- **Converging Infrastructure:** Local and cloud-based HPC big-data analytics workloads are converging, requiring high-level orchestration
- **Compute limitations:** The traditional compute cluster itself, working as the backbone of cryo-EM analysis, faces three common challenges: system bottlenecks, divergent workloads, and barriers to extending usage.

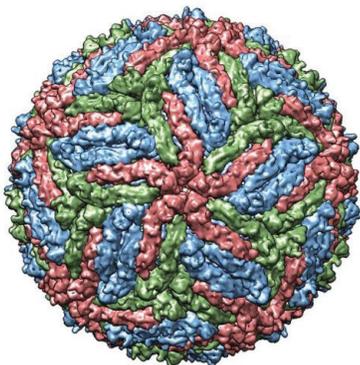


Image from the RCSB PDB (www.rcsb.org) of PDB ID 5IRE (Sirohi, D., Chen, Z., Sun, L., Kloese, T., Pierson, T., Rossmann, M., Kuhn, R.) (2016) The 3.8 angstrom resolution cryo-EM structure of Zika virus. Science 352: 467-470

Cryo-EM has fundamentally changed research and development in structural biology, and thus demands new HPC infrastructure



Each TEM generates approximately 1 PB of data annually



This amount requires between \$1-3M of new HPC infrastructure per TEM



Top research facilities are deploying 3-5 TEMs

...yet many institutions are uncertain how to size the HPC infrastructure required to support this innovative new field.



Reduce Time

Pre-validated Intel® SSF solutions reduce time to insight and remove constraints to life science research.



Optimal Performance

Intel® SSF solutions enable greater configurability and flexibility than contemporary supercomputers across storage, compute, fabric, and software. Intelligence at the hardware and software levels enables supercomputing sites to tune for optimal performance.



High Efficiency

With tight integration and high efficiency, Intel® SSF solutions deliver lower overall TCO by reducing server floorspace footprints, energy, and cooling costs.

Many workloads...One framework

Intel has been working on modernizing various 3D molecular reconstruction applications since 2013 as part of the Intel Parallel Computing Center Program (IPCC). The Intel® Scalable System Framework (Intel® SSF) is a holistic architectural approach that can help address key challenges faced by these improved cryo-EM applications. Intel® SSF is an HPC architecture that utilizes innovative Intel memory, fabric, storage, and system software technologies to increase performance and provide an overall system balance resulting in decreased time to complete the analysis. The balanced integration of system components and modernized code, developed collaboratively with scientists, makes Intel® SSF an ideal tool to help power innovative scientific research that is easy for enterprises to deploy. The framework is designed to improve cost models and scale efficiently as scientific workloads continue to grow. Intel® SSF for life sciences is designed to achieve optimized performance for a variety of workloads; from genomics to molecular imaging/dynamics to machine learning, deep learning, and AI. Intel® SSF also supports a heterogeneous compute environment and supports a scale-out model that enables the separate scaling of compute and storage.

“We are facing a revolution in biochemistry,” said Nobel Committee Chairman Sara Snogerup Linse during the announcement. “Now we can see the intricate details of the biomolecules in every corner of our cells, in every drop of our body fluids. We can understand how they are built and how they act and how they work together in large communities.” “Soon there are no more secrets,” she said.⁴

Break Free From Current Compute Infrastructure Constraints Limiting Your Research

Intel® SSF solutions that include optimized software applications provide an effective cryo-EM platform that will help scientists achieve scalable results faster, maximize the use of available infrastructure, and enable them to achieve new breakthroughs in 3D protein modeling.



¹ Brzezinski, Peter. “Scientific Background on the Nobel Prize in Chemistry 2017: The Development of Cryo-Electron Microscopy.” Nobel Prize, The Royal Swedish Academy of Sciences, 4 Oct. 2017. https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2017/advanced-chemistryprize2017.pdf

² “Method of the Year 2015.” Nature Journal, Springer Nature, 30 Dec. 2015. www.nature.com/nmeth/journal/v13/n1/full/nmeth.3730.html.

³ “The 2017 Nobel Prize in Chemistry - Press Release.” Nobelprize.org. Nobel Media AB 2014. Web. 9 Oct 2017. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2017/press.html

⁴ Vonberg, Judith. “Scientists Win Nobel for Work to Visualize Biomolecules.” CNN, Cable News Network, 6 Oct. 2017. www.cnn.com/2017/10/04/world/nobel-prize-chemistry-2017-cryo-electron-microscopy/index.html.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at www.intel.com.

Benchmark results were obtained prior to implementation of recent software patches and firmware updates intended to address exploits referred to as “Spectre” and “Meltdown.” Implementation of these updates may make these results inapplicable to your device or system. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to www.intel.com/benchmarks.

Optimization Notice: Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. Notice Revision #20110804

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document. You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

Intel, the Intel logo, Intel® Scalable System Framework (Intel® SSF) are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

© 2018 Intel Corporation. Printed in USA 0518/RA/MESH/PDF Please Recycle 337666-001US