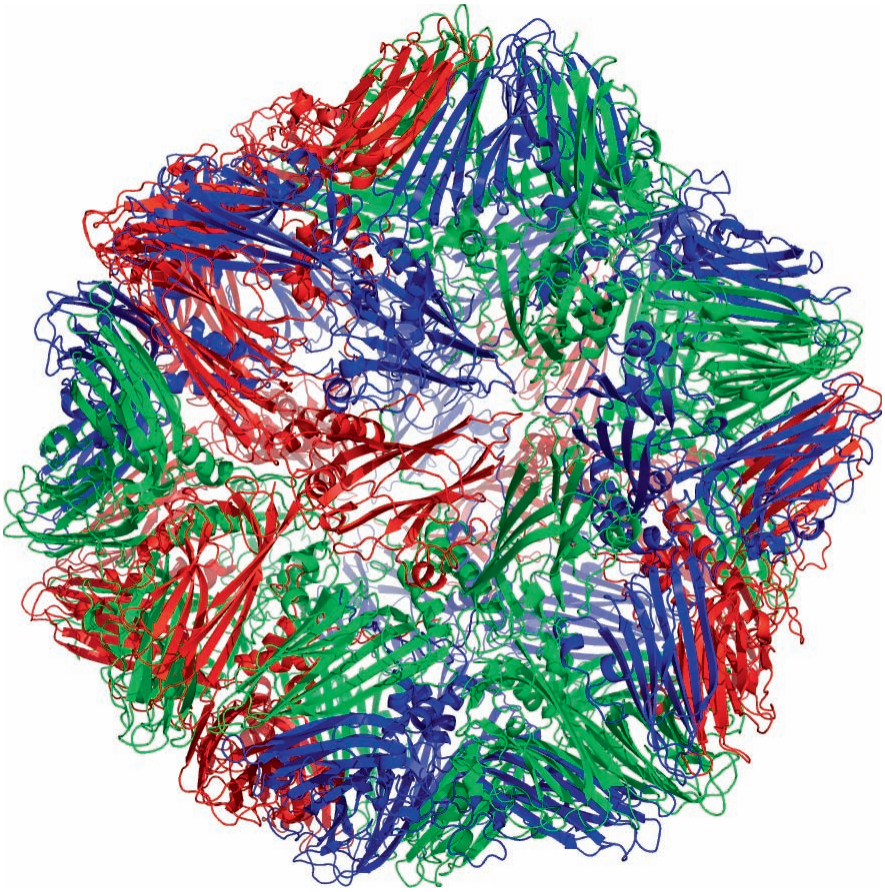
NSC News
National Supercomputer Centre at Linköpings universitet
2008:2



Virus in Neolith

Viral Capsid Dynamics

David van der Spoel is using the NSC Neolith system to make direct simulations of the satellite tobacco necrosis virus capsid.

Read further on page 3

BioGrid

BioGrid is a project to establish a Nordic grid infrastructure for bioinformatics to provide computational power for the increased needs within life sciences.

Read further on page 2

Spring time for Swedish HPC

Swedish high-performance computing (HPC) is currently getting important resources. From the Swedish Research Council, SNIC (Swedish National Infrastructure for Computing) this year receives 63.2 million SEK, which is an essential and long-awaited increase of about 30 %. Furthermore, the Knut and Alice Wallenberg Foundation generously provides additional large-scale resources devoted to selected scientific areas. During summer, one such system for climate and turbulence research will be installed and operated by NSC together with PDC, providing increased capacity to research groups at SMHI, KTH and Stockholm University. Recently, the second application round for similar systems, to be acquired during 2009, attracted large interest from many scientifically well-renowned groups. Furthermore, Sweden is via SNIC taking part in the planning phase for future "peta-scale" computing systems in the EU project "PRACE", Partnership for advanced computing in Europe.

Following the enlarged computational demands, there is also an increasing need for large-scale storage facilities, including long-term secure preservation of research data, with back-up and migration of data keeping pace with technical developments. This is one area of interest for NSC, where Tom Langborg is coordinating the SNIC storage expertise. The importance of this area is also reflected by that the annual international NSC conference this year will be devoted to storage issues. The NSC'08 conference will take place 14–15 October 2008 with satellite meetings specialised on network and security tasks (cf. separate announcement).

The new resources now made available open important possibilities for groundbreaking research in a number of scientific disciplines, including material physics, quantum chemistry, particle physics, fluid dynamics and bioinformatics, just to mention a few.

Thus, the near future for Swedish high-performance computing looks very bright and promising, and with that I wish you all a nice and pleasant summer vacation while the computers are running!

BENGT PERSSON, DIRECTOR OF NSC

BioGrid

Life sciences have undergone an immense transformation during the recent years, where advances in genomics, proteomics and other high-throughput techniques produce floods of raw data that need to be stored, analysed and interpreted in various ways. Bioinformatics is crucial to efficiently utilise these gold mines of data in order to better understand the roles of proteins and genes and to spark ideas for new experiments.

BioGrid is a project, supported by NDGF (Nordic DataGrid Facility), with aim to establish a Nordic grid infrastructure for bioinformatics. Many of the computational tasks in bioinformatics are especially suitable for the grid, since they often need a large number of computers for a short time, so called burst computing.

The widely used bioinformatic software packages BLAST and HMMer have already been gridified in an optimised way and allowing for multicore support. Furthermore, frequently used databases have been made available on the distributed and cached storage system within the Nordic grid. Further applications are currently in the pipeline to be gridified, including multiple sequence alignment, molecule dynamics and phylogeny. The project is coordinated by Bengt Persson at NSC, with steering group members, developers and users spread all over the Nordic countries.

New staff member: Kent Engström

I joined the NSC staff this spring. My first two projects here at NSC are to install and configure NSC's new network infrastructure and to help build and then manage the new SMHI cluster Bore/Gimle. I've been working at Linköping University since 1999 (mostly with IT security and networking) and also got my M.Sc. in Computer Science from Linköping University in 1995.

Viral Capsid Dynamics

All virus particles have a capsid consisting of protein molecules arranged in an ordered fashion (Fig. 1). In fact, the orientations of the coat proteins are related by simple symmetry operators. These proteins are held together by the same forces that determine the structure of a single protein: electrostatic forces (hydrogen bonds, salt-bridges), Van der Waals forces, and the effective driving force due to the hydrophobic effect. In other words, water forces the aliphatic and, to some extent, the aromatic groups "out of the way", and the particular structure that a protein assumes is then the result of combining this effective force with the direct forces. Assembly of multiple proteins to form multimers (dimers, trimers etc.) is a common process, and indeed many proteins are functional only in the multimeric state. A viral capsid is very similar in this respect: it consists of (at least) 60 proteins that aggregate in a specific manner to form a shell (Fig. 1) around the genetic material. By using one (or at most a few) proteins many times, the viral genome (RNA or DNA) can be encapsulated inside the virus particle with minimum evolutionary effort.

The structures of quite a few viral capsids have been elucidated by protein X-ray crystallography. Unfortunately it has hitherto not been possible to determine the structure of the RNA inside a virus capsid. It is actually not known at all whether the RNA does adopt a particular structure inside the capsid. An argument in favor of this is that the packing density can be very high (for some virus particles it is estimated that there is a net pressure inside the capsid of 100 bar). If there is a particular RNA structure, it has until now escaped detection by X-ray crystallography, predominantly because virus particles inside the crystal can be in many different (symmetry-related) orientations.

In a SNIC-sponsored project we are performing computer simulations of the molecular dynamics of an entire virus capsid. As a model system the satellite tobacco necrosis virus (STNV) was chosen, the structure of which was determined in Uppsala [1]. Molecular dynamics simulations form a branch of computational chemistry in which one computes the energies and forces due to an empirical force field on the atoms in a system, and integrates their positions using Newton's equtions of motion in small time steps. The maximum stepsize is determined by the fastest accelerations in the system and the limiting factor is the motion of the light hydrogen atoms. By constraining the fast vibrations in the system, in particular the degrees of freedom of the hydrogen atoms, we can utilize stepsizes of up to 5 fs length.

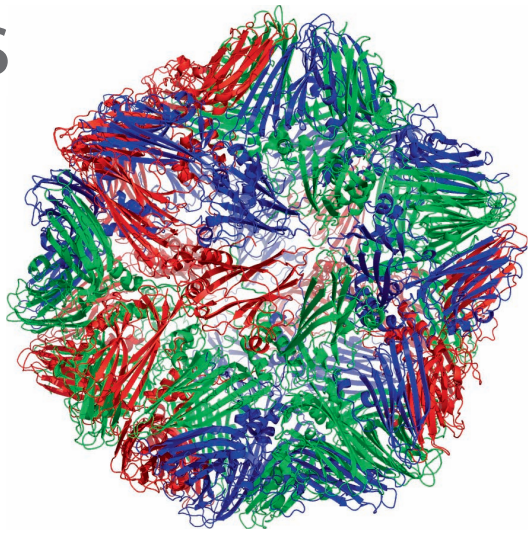


Fig. 1: Ribbon diagram of the STNV capsid. The fivefold symmetry axis can be seen at the front of the particle (comprising of one red, two blue and two green proteins).

In our simulations we used the OPLS force field [2], and the GROMACS software [3], which is the most efficient software for this purpose due in large to recent efforts in enhancing the parallel scaling. The system is divided in segments and local interactions are calculated on single cores and communication is limited to neighboring cells. From a parallel computing standpoint, the bottleneck is the calculation of long range electrostatic interactions which require global communication.

New staff member: Torben Rasmussen

Hi, I am here to try to fill the gap after Patrick Norman as the new applications expert in theoretical chemistry. I will also slowly edge into systems administration as I get more acquainted with everything at NSC. The past four years, I have been working as a researcher at the department of Biochemistry and Organic Chemistry at Uppsala University and before that I worked as a post doc at the department of Theoretical Chemistry at Lund University. Now, I hope to be a source of assistance and useful information for the users of NSC's supercomputer systems.

Dedicated nodes will perform these calculations to minimize the number of participating all-to-all communicating nodes.

The complete virus capsid with ions and water comprises of roughly 1 million atoms, which makes this system one of the largest biomolecular simulations to date. When utilizing about 100 cores at Neolith we can obtain roughly 6 ns (1.2×10^9 time steps) per day. It is also possible to use many more cores (we have used up to 700 cores for a single calculation), but in most cases this is not as efficient. With the increased availability of computer time and more efficient software [3], MD simulations are now slowly becoming a "normal scientific technique", in the sense that multiple "simulation experiments" are being performed, where previously only a single simulation was done in most cases. In addition we can now test different environmental conditions like salt concentration. This means that we in practice run many simulations on roughly 100 cores each, and get more throughput for our calculations.

The regular structure of a virus capsid is stabilized by the large interaction surface between proteins. In the case of STNV each individual coat protein has roughly 15 hydrogen bonds and 25 nm^2 hydrophobic contact area (which is about 20% of the total area) to neighboring proteins. Another stabilizing feature of many virus capsids is formed by the interactions with RNA. Indeed, many coat proteins have a particular stretch of amino acids, often one of the termini, that has a large positive charge, to compensate the large negative charge on the phosphate backbone in the RNA [4]. In STNV there is such an N-terminal arm of 24 residues, that has a charge of +8, which is protruding inside the capsid. The structure of the 14 outermost residues of this arm was unresolved in the crystal structure [1] because it is not identical in every copy of the protein. In our initial simulations of the STNV capsid we have not modeled in the RNA yet. Instead, we have used a high salt concentration (negative chloride ions) to compensate for the lack of RNA. In the beginning of the simulation we find that the ions very quickly redistribute themselves such that they bind to the N-terminal arms of the coat proteins. The reorientation of the arms and the equilibration of the entire capsid takes place on a time scale of 100 ns (Fig. 2a), which is an order of magnitude more than the time scale reached in the first-ever atomistic simulation of an entire virus capsid [5]. A slight expansion of the capsid can also be seen over the course of the simulation and might be due to the osmotic pressure from the chloride ions added on the inside of the system (Fig. 2b).

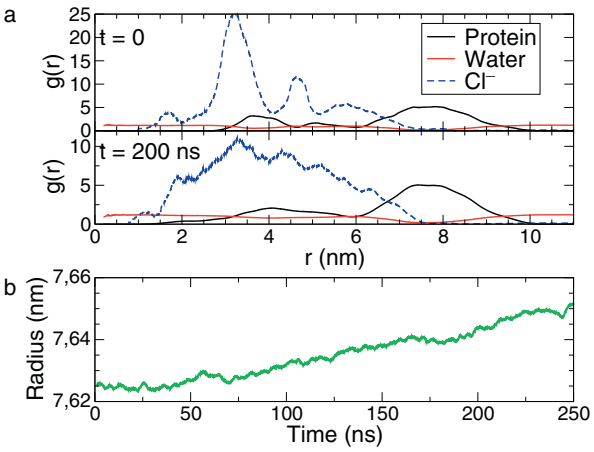


Fig. 2a: Radial distribution of atoms with respect to the center of mass of the particle at the start of the simulation and after 200 ns.

Fig. 2b: Radius of gyration of the virus particle as a function of time.

NSC'08 conference

The annual NSC conference this year will focus on large-scale storage issues. NSC'08 will gather people with experience of building and administrating/maintaining storage systems. We expect a number of international speakers reporting on recent developments, planning and technical status regarding large-scale storage. The NSC'08 conference starts at 10 a.m. on 14 Oct and finishes by lunch on 15 Oct.

Adjacent to the NSC'08 conference, SUSEC, which is an association for IT security staff at SUNET-connected organisations, will organise their SUSEC autumn workshop with a focus on security issues. Also adjacent to the NSC'08 conference, SUNET (Swedish University Network) arranges the TREFpunkt 19 with a focus on network issues.

More information regarding the three conferences will be available at: http://www.nsc.liu.se/nsc08.

We wish you all very welcome to these conferences in Linköping in October 2008!

The complexity involved with the simulation of an entire virus capsid makes for a formidable computational challenge. The time scale for any significant analysis of capsid dynamics must logically be 500 ns – 1 µs. Until recently, high-performance computing in Sweden was severely underfunded. With the advent of the Neotlith cluster, Swedish scientists can at last start competing internationally. Using Neolith we will be able to perform a series of long virus capsid simulations, which will allow us to address some of the most pertinent questions in virology.

In the long run we will use this kind of simulations to make predictions about novel inroads for anti-viral therapy.

DAVID VAN DER SPOEL &
DANIEL LARSSON

References

[1] T. Unge, L. Liljas, B. Strandberg, I. Vaara, K. K. Kannan, K. Fridborg, C. E. Nordman and P. J. Lentz Jr.: "Satellite tobacco necrosis virus structure at 4.0 A resolution", Nature 285, (1980) 373-377

[2] W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives: "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids", J. Amer. Chem. Soc. 118 (1996) 11225-11236

[3] B. Hess, C. Kutzner, D. van der Spoel and Erik Lindahl: "GROMACS 4.0: Algorithms for highly efficient, load-balanced, and scalable molecular simulation", J. Chem. Theor. Comp. 4 (2008) 435-447

[4] D. van der Spoel, K. A. Feenstra, M. A. Hemminga and H. J. C. Berendsen: "Molecular Modelling of the RNA Binding N-terminal part of CCMV coat protein in Solution with Phosphate Ions", Biophys. J. 71 (1996) 2920-2932

[5] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson and K. Schulten: "Molecular dynamics simulations of the complete satellite tobacco mosaic virus", Structure 14 (2006) 437-449

David van der Spoel got his Ph.D. in computational chemistry from the university of Groningen, The Netherlands in 1996. He is currently employed as an associate professor in the molecular biophysics group at the department of cell and molecular biology, Uppsala university.

NSC hosts HP-Cast NTIG in Linköping April 1 – 2, 2008

"Welcome to NSC and HP-Cast NTIG" was the greeting that met the participants of this event a sunny day in early April. With several very large HP-clusters being installed in the Nordic countries the idea to form a Nordic Technical Interest Group (NTIG) in high performance computing using HP clusters came up during last years Super-computing show SC2007. Organized under the umbrella of HP's HPC User Group, the HP Consortium for Advanced Scientific and Technical Computing (HP CAST), NSC was proud to be chosen as the first site for this event.

The meeting took place on the NSC premises as a lunch-to-lunch meeting April 1 – 2 this year. Approximately 30 people attended with representatives from computer centres from all the Nordic countries except Iceland as well as from HP and GoVirtual. Since this was the first meeting the program started with site presentations from the larger sites in order to get to know each other and what type of equipment is installed, experiences from using the systems as well as problems and solutions. Some sites have recent or very recent installations while other sites have had their systems for a while. The recent systems all have Infiniband and many of them are using blade servers. This type of information exchange was very much appreciated among the attendees. The second half of this first day was dedicated to updates from HP by Martin Walker and Frank Baetke regarding future road maps as well as what is going on in the PRACE project. A tour of NSC's Neolith system wrapped up the afternoon and in the evening a delicious dinner was served, kindly sponsored by HP.

The second day of the meeting had a more technical content with presentations on HPC tools by Ed Benson HP HPCD and an update on Service and Support Structures for HPC in the Nordic Region by Jeanette Sharkey, HP HPCD. Frequent questions and discussions showed that these topics were close to the heart of many of the user sites that participated in the meeting. The user sites then met to discuss and prioritize support issues that are important which were later discussed jointly with HP. Everyone that attended felt that this meeting was successful and very useful and would like to see a continuation next year. CSC graciously offered to host the next meeting March 24 – 25 in Helsinki.
See you then!

Participants in the first ever HP-Cast Nordic Technical Interest Group (NTIG) meeting gathered outside NSC's new computer building.

UPCOMING EVENTS

ISC'08; International Supercomputing Conference
June 17 – 20, 2008, Dresden, Germany.
http://www.supercomp.de/isc08/index.php5

HPDC'08; ACM/IEE International Symposium on High Performance Distributed Computing
June 23 – 27, 2008, Boston, USA.
http://www.hpdc.org

Joint EGEE and OSG Workshop on VO Management in Production Grids (held in conjunction with HPDC'08)
June 24, 2008, Boston, USA.
http://indico.cern.ch/conferenceDisplay.py?confId=27789

International Summer School on Grid Computing 2008
July 6 – 18, 2008, Hungary.
http://www.issgc.org

ICANN 2008; The 18th International Conference on Artificial Neural Networks
September 3 – 6, 2008, Prague, Czech Republic.
http://www.icann2008.org

EGEE'08; Enabling Grids for E-sciencE
September 22 – 26, 2008, Istanbul, Turkey.
http://www.eu-egee.org/egee08/home.html

NSC'08 conference with focus on storage issues
October 14 – 15, 2008, Linköping, Sweden.
http://www.nsc.liu.se/nsc08

SUSEC autumn workshop with a focus on security issues
October 14 – 15, 2008, Linköping, Sweden.
http://www.nsc.liu.se/nsc08

SUNET TREFpunkt 19 with focus on network issues
October 15 – 16, 2008, Linköping, Sweden.
http://www.nsc.liu.se/nsc08

E-science conference: "Global challenges – Regional opportunities: How can Research Infrastructure and eScience support Nordic competitiveness?"
November 12 – 13, 2008, Stockholm, Sweden.
http://www.vr.se

SC 08; International Conference for High Performance Computing, Networking, Storage and Analysis
November 15 – 21, 2008, Austin, Texas, USA.
http://sc08.supercomputing.org

PACT 2008; 7th International Conference on Parallel Architectures and Compilation Techniques
October 25 – 29, 2008, Toronto, Canada.
http://www.eecg.toronto.edu/pact/

ICPADS 2008; 14th International Conference on Parallel and Distributed Systems
December 8 – 10, 2008, Melbourne, Australia.
http://www.deakin.edu.au/conferences/icpads2008/

HiPC-08; International Conference on High Performance Computing
December 17 – 20, Bangalore, India.
http://www.hipc.org

PerComm-09; IEEE Intl Conference on Pervasive Computing and Communications
March 16 – 20, 2009, Dallas, USA.
http://www.percom.org

IPDPS-09; IEEE International Parallel and Distributed Processing Symposium
May 25 – 29, 2009, Rome, Italy.
http://ipdps.org

Linköpings universitet

National Supercomputer Centre, Linköpings universitet, 581 83 Linköping, Sweden

Tel: +46 (0) 13-28 26 18, fax: +46 (0) 13-28 25 35, e-mail: info@nsc.liu.se

www.nsc.liu.se