★ During the next decade standard medical practice will change dramatically. **Peter Sloot,** of the ViroLab project, explains how personalised, targeted treatments will allow medical professionals to move beyond prevention towards adopting pre-emptive health strategies

ViroLab: from the molecule to the man

The human body is enormously complex, as are the many factors that can affect its healthy functioning. The complete relationship between the genome, proteome, metabolome and physiome and an individuals overall health involves multi-scale, multi-science systems, and crosses many orders of magnitude in both temporal and spatial scales.

Understanding, quantifying and handling this complexity is one of the biggest scientific challenges of our time.

It is the central assertion of the ViroLab project that computer science represents the ideal language to study and understand these systems. Comprised of some 14 partners from across Europe, as well as collaborators in both China and Russia, the ViroLab project aims to develop a virtual laboratory in order to improve the treatment of infectious diseases.

In this paper we will briefly discuss the ViroLab system itself – a Grid based decision support system – as well as look at some of the other major issues we face. One of the most pressing is HIV drug resistance, one of the few areas in medicine where genetic information is widely available. Indeed, it has already been used for a number of years. As a consequence, in addition to the clinical data, large numbers of complex genetic sequences are available.

Pushing, pulling and bridging the gaps

The change in paradigm to in-silico studies has brought with it an application pull from biomedicine.

Along with in-vivo and in-vitro studies, more and more details of biomedical processes are being simulated in the in-silico area.

Typical examples include both pre-operative simulation and visualisation of vascular surgery, and also expert systems for drug ranking. At the same time we have seen a technology push in terms of computing and data availability. In order to close the computational gap in systems biology, we need to construct, integrate and manage several different models.

A general scheme for conducting this kind of e-Science research is depicted in Figure one, below.

During the past ten years significant progress has been made in the treatment of patients infected with viral diseases. Almost 20 anti-retroviral drugs are now available to treat HIV, and in some cases these drugs have been successful in completely suppressing the virus. However, we must also recognise that in a large proportion of patients they have failed to achieve this goal, which has resulted in the rapid selection of drug-resistant viruses and loss of drug effectiveness.

Our vision is to provide researchers and medical doctors with a virtual laboratory for infectious diseases – ViroLab – so as to enable easy access to distributed

Figure 1

**A general architecture for conducting e-Science research**
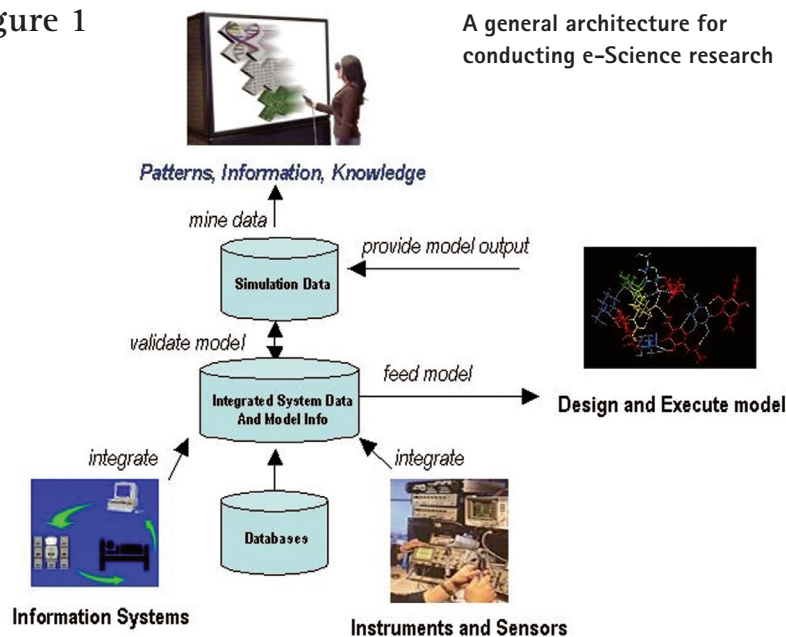
simulations as well as to facilitate the sharing, processing and analysing of virological, immunological, clinical and experimental data. The stages in the typical usage scenario we envisage are outlined below:

- A scientist from a laboratory for clinical and epidemiological virology in Utrecht will be able to securely access a virus sequence, amino acid or mutation data from a hospital's AIDS lab in Rome – using data Grid technology components running in Stuttgart.

- The scientist will then apply the quality indicators needed for data provenance tracking. This will be done using provenance server components running in Krakow

- This data is used as input for both the molecular dynamics and Immune system simulations which run on Grid-nodes at UCL (University College, London) and UvA, (University of Amsterdam).

- The virtualised DSS automatically derives meta-rules

- Intelligent system components from Amsterdam use first order logic to clean the rules, check the consistency of the logic, and identify conflicts and redundancy

- The scientist validates the new rules, which are then automatically uploaded into the virtualised DSS

- A new ranking is presented

At ViroLab we have decided to put an interpretation tool at the centre of a distributed virtual laboratory. One of these interpretation tools is Retrogram. Retrogram estimates the sensitivity of the available drugs by interpreting the genotype of a patient using mutational algorithms, algorithms which have been developed by experts on the basis of scientific literature. Furthermore, the system also supports Grid-based distributed data access and computation. Indeed, ViroLab includes advanced tools

Figure 2
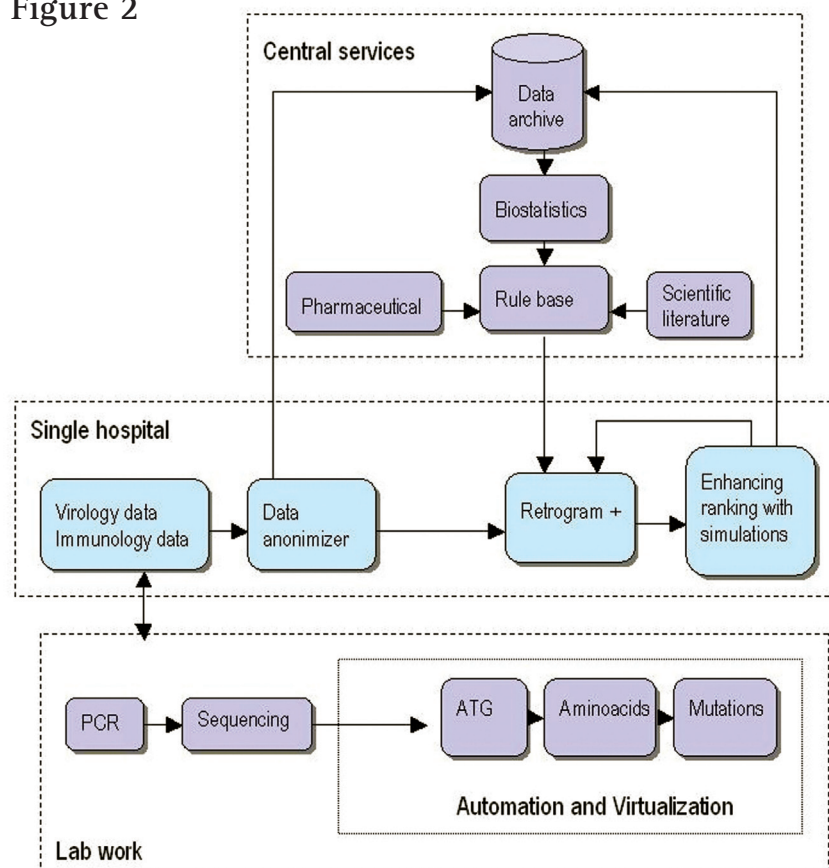
The ViroLab data and control flow.

for bio-statistical analysis, visualisation, modelling and simulation that enable prediction of the temporal virological and immunological response of viruses with complex mutation patterns for drug therapy, as seen in Figure two, above.

Preliminary results
In his 1982 research paper 'New Computer Architectures and Their Relationships to Physics', prominent American scientist W.D. Hillis wrote that, 'Progress in natural sciences comes from taking things apart; progress in computer science comes from bringing things together,' words that continue to ring true.

Indeed, within the ViroLab project we quickly recognised that we needed both statistical and immunological models if we were to be able to study the dynamics of the HIV populations. It also became clear that we would require molecular dynamics models in order to study drug affinities, in addition to rule and parameter-based decision support.

We have enhanced ViroLab by adding

We quickly recognised that we needed both statistical and immunological models if we were to be able to study the dynamics of the HIV populations

cellular automata (CA) and molecular dynamics modelling of the onset of both the HIV and AIDS infections. We also continue to study the progress of the disease through Complex Networks Simulation.

Figure 3
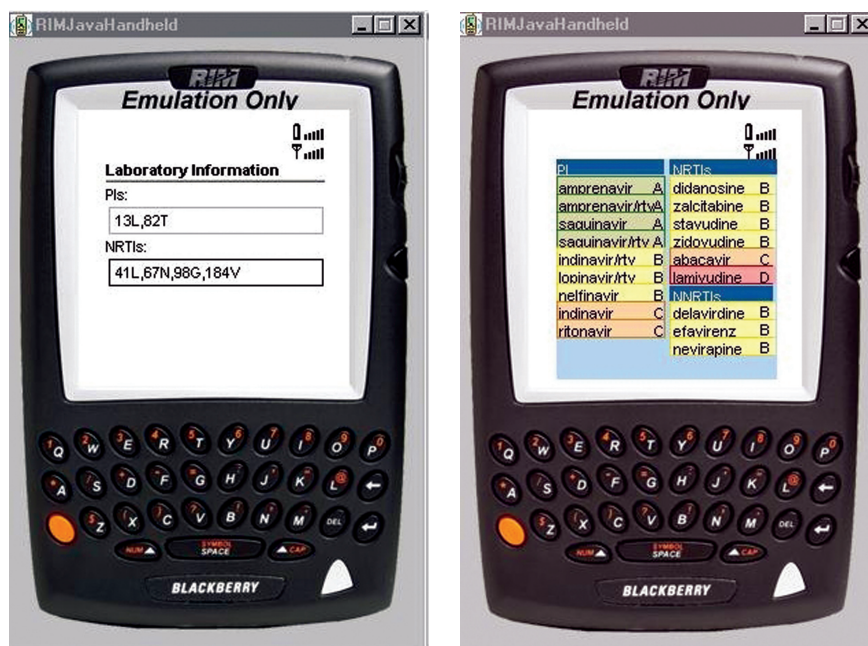
Drug ranking. Left: patient specific Reverse Transcriptase and Protease mutations, Right: resulting drug ranking

These approaches render the highly complex, multi-dimensional data much more accessible. ViroLab uses a mesoscopic model to study the evolution of the HIV infection and the onset of AIDS. The model takes into account the global features of the immune response to any pathogen, HIV's fast mutation rate, and a fair amount of spatial localisation, which may occur in the lymph nodes.

We developed a non-uniform CA model to study the dynamics of drug therapy of HIV infection, which simulates four-phases (acute, chronic, drug treatment response and onset of AIDS). Three different drug therapies (mono-therapy, combined drug therapy and highly active antiretroviral therapy) can also be studied using this model.

We currently use this simulation data together with complementary information from the Centre of Disease Control in Atlanta, USA, with the goal of accurately predicting the dynamics of the epidemic through complex sexual networks.

The ViroLab project is focussed on developing a method of predicting the overall drug sensitivity of the virus. In order to achieve this goal, the decision software interprets the genotype of a patient by using rules developed by experts on the basis of the available literature, rules which take into account the relationship between the genotype and phenotype. It must also be considered that it is based on the data available from clinical studies and on the relationship between the presence of genotype and the clinical outcome.

Figure three, above, shows a typical output of the ranking using the latter method.

At a glance

Objectives:

ViroLab's mission is to develop a virtual laboratory for infectious diseases. Genetic information is expected to become increasingly significant in many areas of medicine in the future.

Partners:

Universiteit van Amsterdam, The Netherlands; Universitair Medisch Centrum Utrecht, The Netherlands; Universität Stuttgart, Germany; Universita Cattolica del Sacro Cuore, Italy; Institute de recerca de la SIDA, IRSICAIXA Foundation, Spain; Universita Degli Studi di Brescia, Italy; Katholieke Universiteit Leuven, Belgium; Eötvös Loránd Tudományegetem, Hungary; Gridwisetech, Poland; ACK Cyfronet AGH, Poland; University College London, United Kingdom; Virology Education B.V., The Netherlands

Contact:

e: sloot@science.uva.nl
t: +31 205 25 7462
www.science.uva.nl/~sloot

Peter Sloot

Professor
Universiteit van Amsterdam

Peter Sloot has been a full professor in Computational Sciences at the Informatics Institute of the Faculty of Science at the Universiteit van Amsterdam since 2001. In his research he focuses on the theory and application of complex systems through distributed mesoscopic computer simulation; trying to understand how information progresses through various spatial and temporal scales.