Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 22 (2004) 473–478
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Combating bioterrorism with personal computers

W. Graham Richards*, Guy H. Grant, Karl N. Harrison

Central Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QH, UK

Accepted 4 March 2004

Available online 27 April 2004

Abstract

Using personal computers in a grid is permitting the in silico screening of millions of molecules to seek out potential inhibitors of agents that pose bioterror threats. Current projects are targeting anthrax and smallpox, but the approach has many attractions for investigating any known protein target and its inhibition.

Keywords: Personal computers; Bioterrorism; In silico screening

1. Introduction

In a much publicised project run by the National Foundation for Cancer Research Center for Computational Drug Design in Oxford, personal computers can utilise a screensaver to probe whether small molecules are potential inhibitors of proteins involved in cancer [1]. Started in April 2001, the user community has now grown to over two million PCs in some 225 countries. This is an enormous resource and fuller details can be found at http://www.chem.ox.ac.uk/curecancer.html.

In outline, participants download a screensaver that, while their computer is idle, takes molecules one at a time and tries them out as potential inhibitors of the protein, investigating binding into a pre-defined binding site on the target protein. The distributed grid and the central server, which sends out the test molecules and receives information about promising results 'hits' is being provided by United Devices Inc. of Austin, Texas [2]. In the cancer project a database of some 3.5 billion molecules, all of which satisfy Lipinski's criteria for drug properties [3], was screened using software produced by Keith Davies. The THINK program [4] works on the basis of pharmacophore matching: binding possibilities on protein and ligand are matched and scored, with the query permitting some flexibility in the protein and all possible conformations of the small molecule being considered by allowing up to 10 rotatable bonds and taking every conformation generated by considering 10° steps about each dihedral angle. Many millions of conformations may be considered for each molecule screened. A more detailed description of the THINK procedure is given in reference [4]. More recently, the hits from that first phase of the project have been run again using the docking software LigandFit [5] developed by Venkatachalam et al. of Accelrys Inc. Being a three-dimensional docking program, this software can take into account the question of chirality, which is obscured in THINK: the latter achieves its amazing speed by working with SMILE strings.

This technology and resource is now also being used to seek potential drug leads which could be used to combat perceived bioterrorist threats.

2. Anthrax

In any project of this type, the two scientific requirements are: first, a defined three-dimensional protein target binding site, and second, the database of small molecules.

The database of small drug-like molecules was built starting with publicly available catalogues of molecules on sale from commercial suppliers, which yields about 1.5 million compounds. Published literature on chemical libraries, derived from combinatorial chemistry, provides a further billion structures [6]. When filtered to ensure drug-like properties such as molecular weight in the correct range; solubility, and suitable lipid partition qualities, this reduces to some 35 million compounds. In the pharmacophore studies, each of these molecules had a hundred variants created

* Corresponding author.
E-mail address: graham.richards@chemistry.ox.ac.uk (W.G. Richards).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.03.007

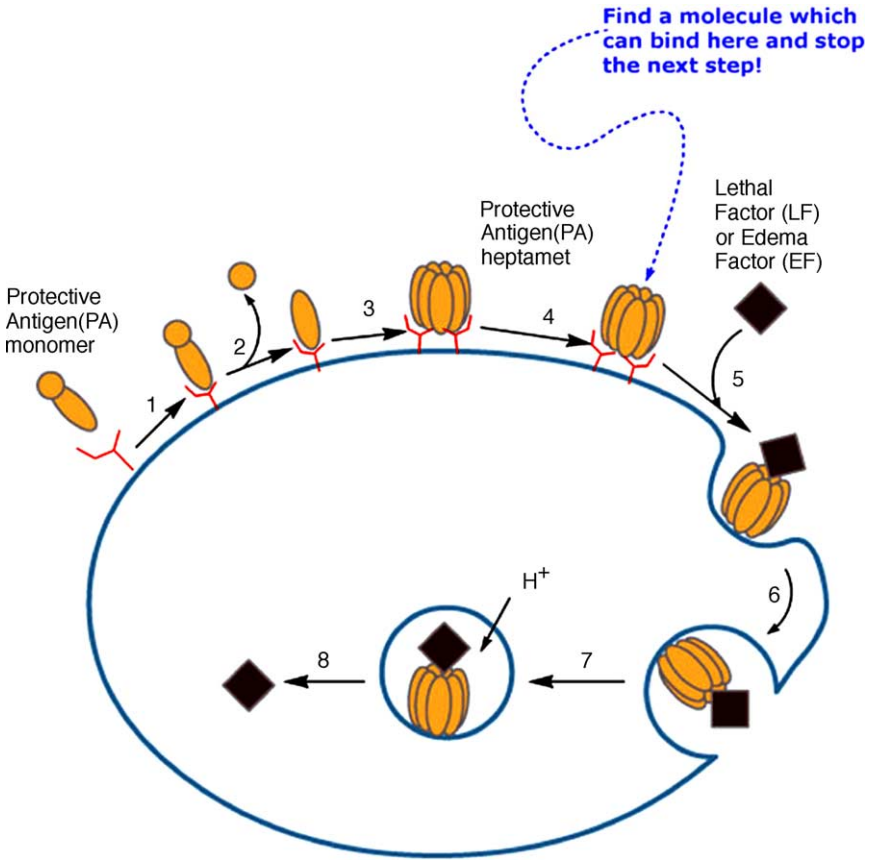

Fig. 1. Proteins forming the anthrax toxin and their mode of action.

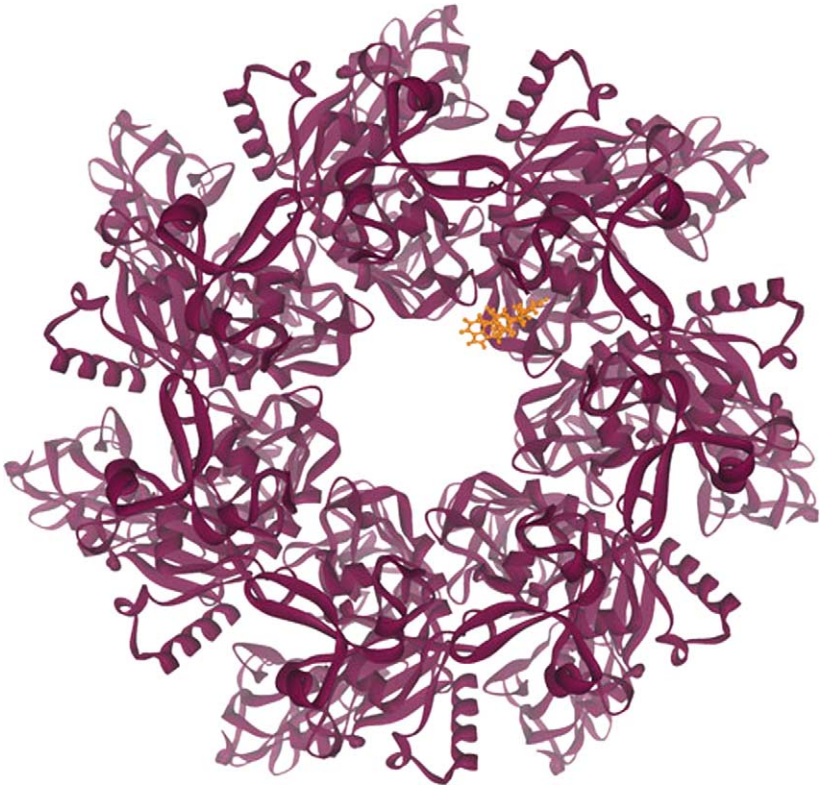

Fig. 2. Axial view of the binding site target on the protection antigen component of the anthrax toxin.

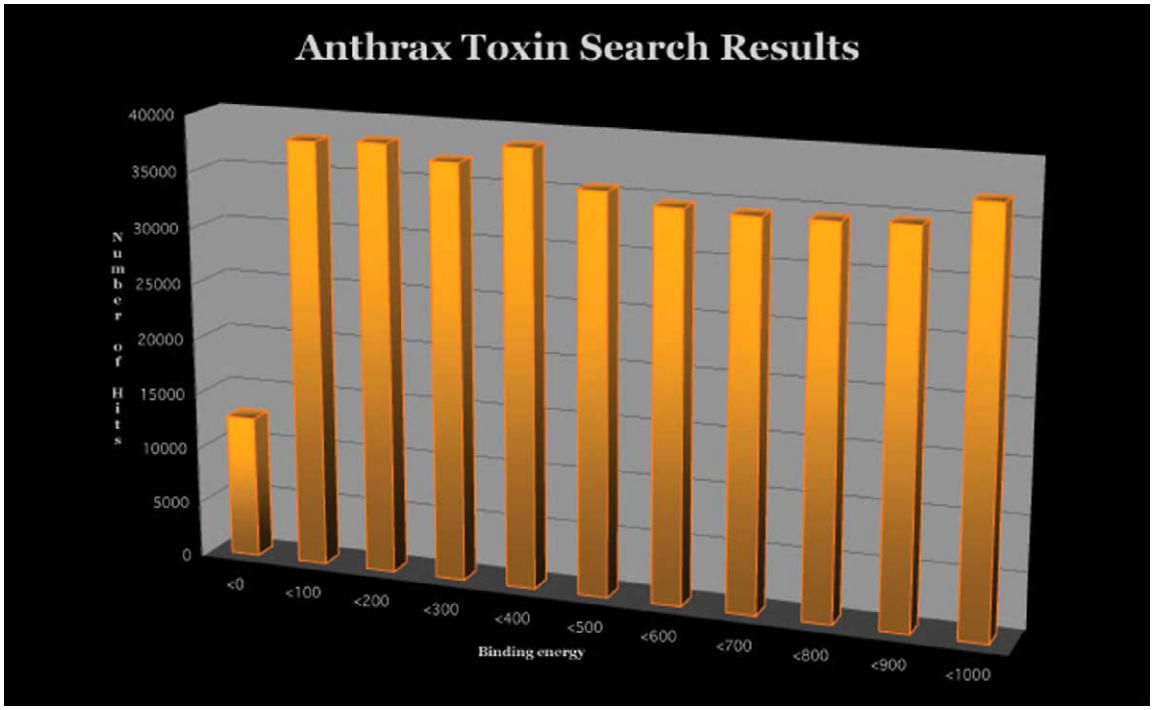

Fig. 3. Histogram of 'hits' for the anthrax target site. The scoring is arbitrary but represents a crude binding free energy with scores with negative values being most promising.

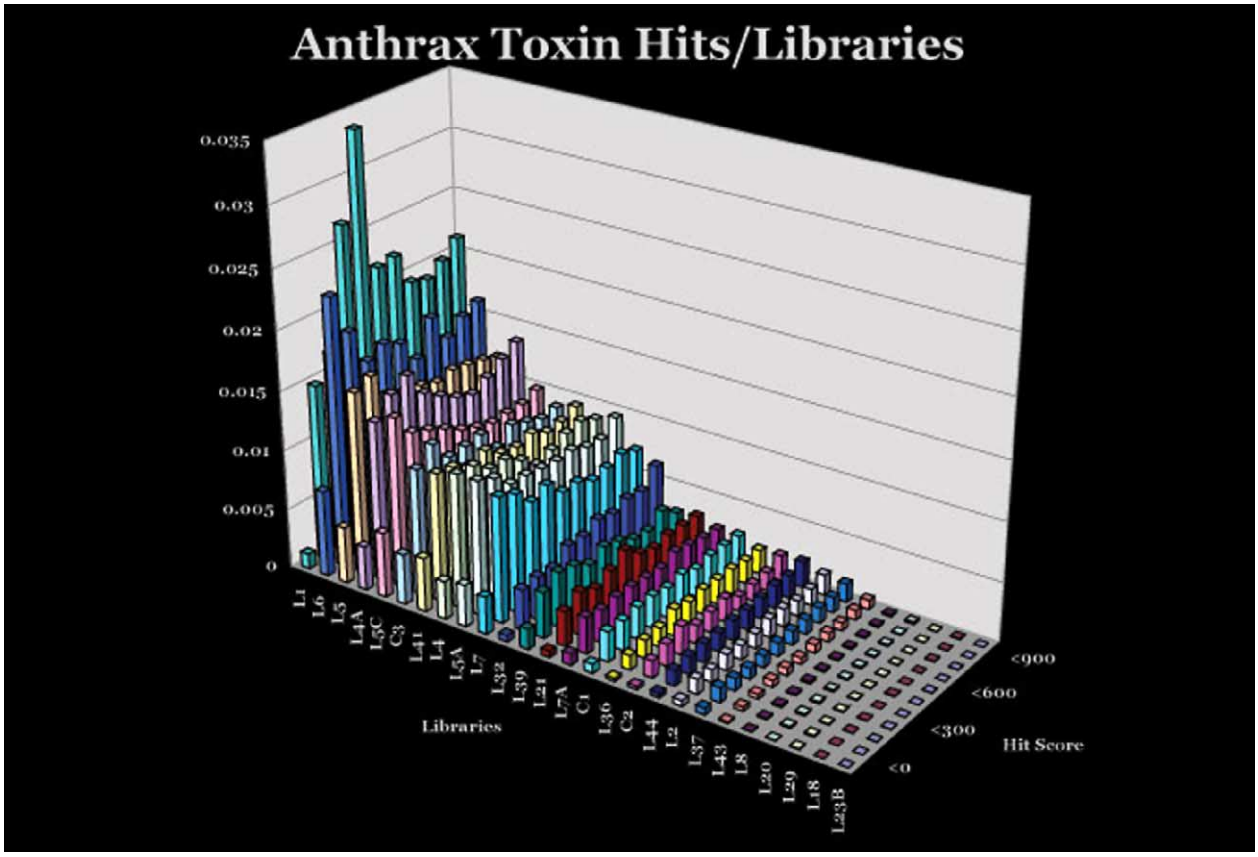

Fig. 4. Histogram of 'hits' for a variety of chemical libraries.

'on the fly' by substituting atoms by suitable groups using random number generators.

The target site in the case of anthrax was pinpointed [8] using software developed by Glick et al. [7] to probe the surface of the protection antigen, which is one of the three proteins that make up the toxin [9] (see Fig. 1) (PDB code 1acc; heptameric structure, C. Petosa, personal communication). Tellingly, the proteins are not toxic individually but only when the protective antigen forms a heptamer and the lethal factor and oedema factor also combine. The postulated binding site is situated at the junction between the protein units and just where the protective antigen would bind the so-called lethal factor (see Fig. 2). Significantly, the stoichiometry of lethal factor/oedema factor binding to the protective antigen is 3:7, implying that the binding site sits across a dimer interface [10].

The histogram of hits against this target is illustrated in Fig. 3. These hits are not evenly spread amongst the originating chemical libraries as shown in Fig. 4. Full details of all libraries can be found on our website.

The hits have been passed on to US Government Laboratories since only they have the resources to take forward the work which has no commercial attraction for a pharmaceutical company, despite the fact that in the event of an attack, any drug to combat anthrax might be more useful than mass vaccination.

3. Smallpox

During the course of this conference, we have launched a similar project in the hope of treating smallpox, the variola virus. Variola is the most virulent member of the orthopox genus of virus. It is specific for humans and has no other animal hosts. Smallpox was eliminated from the world in 1977, however, stocks of the variola virus are known to exist, and its use as a weapon of bioterrorism remains a frightening possibility. The availability of drugs to counter the virus would be a major defence, despite there being a vaccine.

The vaccine was very successful in eradicating the variola virus but the vaccine does have serious side effects. There are, on an average, two deaths per million vaccinations, and several hundred people could have serious side effects. Additionally, people whose immune system is not functioning correctly must not be given the vaccine. These would include young children, pregnant mothers and those

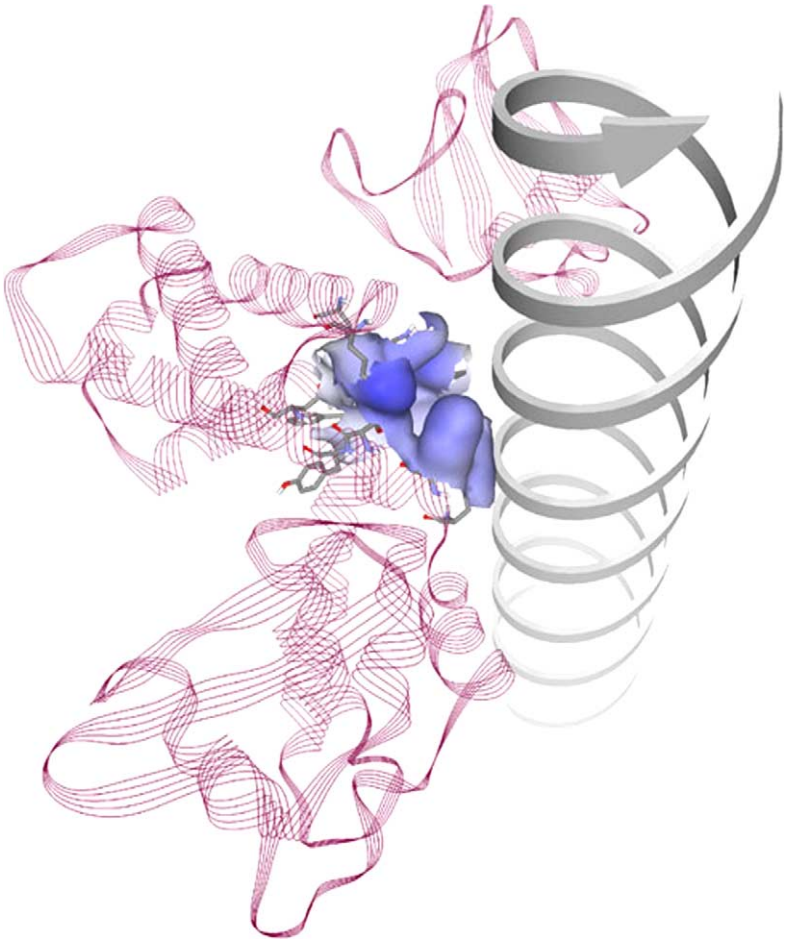

Fig. 5. Hypothetical model [11] of Type I topoisomerase, the target for potential anti-smallpox drugs, bound to DNA. Residues in the RKRH and SKxxY motifs are highlighted, together with other residues thought to be responsible for stabilising reaction intermediates.

on certain medications. It is also suggested that a "weapons grade" smallpox virus may be different from the natural variola virus. In fact, the US government is stock piling a different strain of the vaccinia virus compared with the one the UK government has started collecting. It takes time for a vaccine to activate the immune system via the production of antibodies; during that time the vaccine offers no protection. So a drug would be useful for an emergency situation. Vaccination is only prevention, not a cure.

A prototype for the orthopox viruses is vaccinia virus, which shows considerable sequence similarity to variola, yet does not lead to disease in humans. There are no X-ray crystal structures for the variola proteins, but such is the degree of amino acid similarity we can start with structures available for vaccinia and use homology modelling to prepare detailed target binding sites.

One such structure is that of the Type I topoisomerase (Fig. 5). This class of enzyme plays an important role in DNA replication and transcription, and has been proposed as a putative target for anticancer therapy. The principal role of Type I is the relaxation of supercoiled DNA through the cleavage of a single strand of a DNA double helix. Unwinding, followed by recombination, can then take place. While these enzymes are present in other organisms, they are particularly important for compact viral genomes, making them an ideal target for intervention.

The structure of vaccinia topoisomerase contains two domains but it has been shown that the isolated C-terminal domain (residues 81–314; PDB code 1A41; resolution 2.3 Å) [11] is catalytically competent. Sequence analysis of topoisomerases has also shown that certain motifs are conserved in the active site: these include an RHRW group that appears to be involved in stabilising reaction intermediates, and an SKxxY motif that is directly involved in forming the covalent linkage between the protein and DNA. The sequence alignment shown in Fig. 6 indicates that only three amino acids differ between the vaccinia and variola proteins over their entire length; only one of these is found in the C-terminal domain, and it doesn't interfere with the active site. Initially, we will screen compounds against the second site, as the first is not completely resolved in the vaccinia crystal structure. The missing fragment of seven residues, whose structure is likely to be templated by bound DNA, will be modelled using homology modelling and molecular dynamics on the active site.

The model building will involve the generation of many possible conformations for the missing region of structure. The best of these will then be checked for stability using molecular dynamics simulations before selection of the final model. The tool [7] will be used to predict the most likely area for drug molecules to bind, and this will be

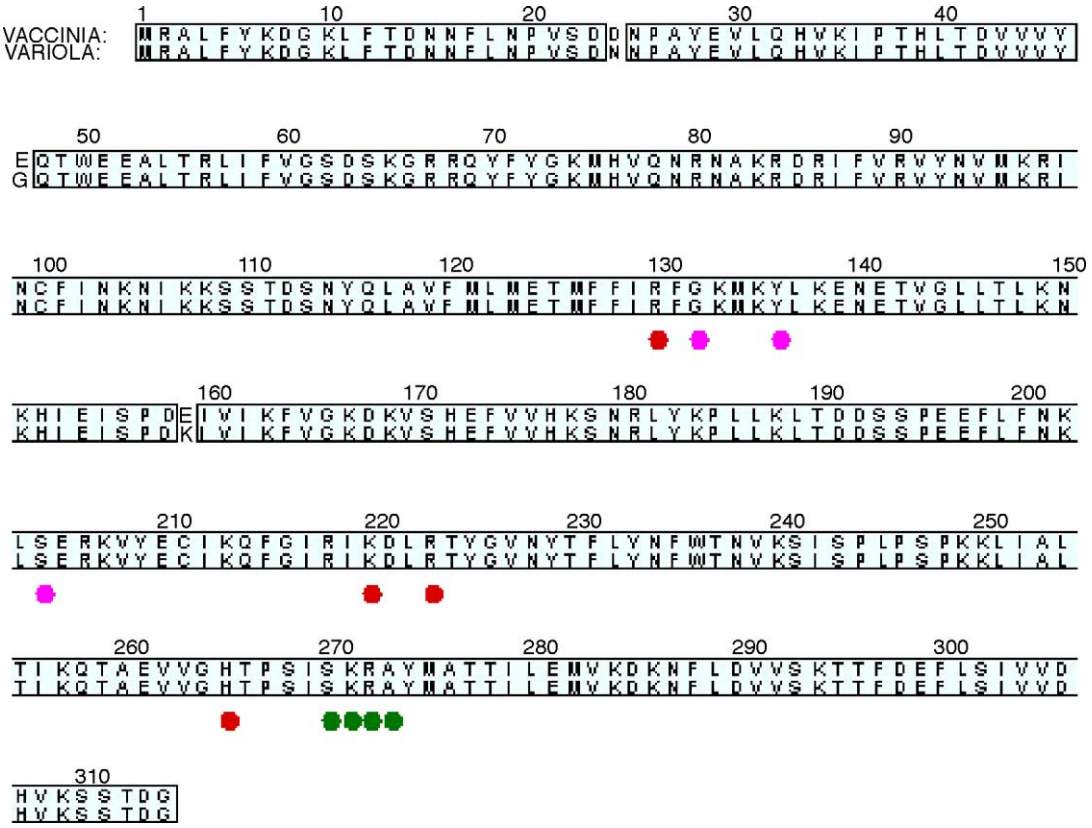

Fig. 6. Sequence alignment of the variola and vaccinia Type I topoisomerases. The RKRH motif is marked by red dots, the SKxxY motif with green dots and other residues responsible for chain cleavage with lilac dots.

taken forward into a screen using the grid of personal computers.

A second issue that will also be addressed is that of specificity for the variola/vaccinia protein over that of the human host. Crystal structures exist for the human Type I protein allowing secondary screening to be carried out. Only those molecules that fit the viral protein, but not that from humans, need to be considered further. A second advantage of such an approach is that the inhibitors of the human enzyme might show some promise as agents for cancer therapy. Thus, the screening project will identify not just potential agents against smallpox but also putative anti-cancer drugs.

4. Discussion

The sheer power of the approach using distributed computing over a grid of personal computers is quite staggering. In our project, we now have over 2 million PCs from more than one million individuals in some 225 countries, with an average 2200 machines joining each day. This has provided over 225,000 years of CPU time. Other popular examples of distributed computing are the SETI@home project [12] and the distributed.net RC5 project [13]. The computing power of these distributed computing projects is immense; it is estimated that our project has the processing power of a 60–100 teraflop machine. This is almost double the power of the largest supercomputer in the world, and more powerful than the top 10 computers in the world combined together [14].

This type of problem, where we have many molecules to consider, is perfect for grid computing: small batches of molecules can be sent to a separate personal computer to be processed, with results transmitted back to the central server over the internet. Fuller details of the smallpox project, together with a section of questions and answers can be found at http://www.chem.ox.ac.uk/smallpox.

The project has illustrated just how keenly the general public appreciates the opportunity to take part in a major scientific project.

Acknowledgements

The bulk of this work has been supported by the National Foundation for Cancer Research. Much of the data preparation was undertaken by Professor Christopher Reynolds and George Psaroudakis of Essex University, and protein modelling has been performed by EvotecOAI Ltd. Advice on the smallpox targets was provided by Dr. Grant McFadden of Roberts Research Institute, Canada.

References

[1] W.G. Richards, Virtual screening using GRID computing: the screensaver project, Nat. Rev. Drug Discov. 1 (2002) 551–555.
[2] United Devices (http://www.ud.com).
[3] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 23 (1997) 3–25.
[4] E.K. Davies, C.J. Davies, Ultra high throughput screening using THINK on the internet, Abstr. Pap. Am. Chem. Soc. 224 001-CINF, Part 1, 18 August 2002.
[5] C.M. Venkatachalam, T. Jiang, T. Oldfield, M. Waldman, LigandFit: a novel method for the shape directed rapid docking of ligands to protein active sites, J. Mol. Graph. Model. 21 (2003) 289–307.
[6] R.E. Dolle, Comprehensive survey of combinatorial library synthesis, J. Comb. Chem. 3 (2000) 477–517.
[7] M. Glick, D.D. Robinson, G.H. Grant, W.G. Richards, Identification of ligand binding sites on proteins using a multi-scale approach, J. Am. Chem. Soc. 124 (2002) 2337–2344.
[8] M. Glick, G.H. Grant, W.G. Richards, Pinpointing anthrax-toxin inhibitors, Nat. Biotechnol. 20 (2002) 118–119.
[9] C. Petosa, R.J. Collier, K.R. Klimpel, S.H. Leppla, R.C. Liddington, Crystal structure of the anthrax toxin protective antigen, Nature 385 (1997) 833–838.
[10] J. Mogridge, K. Cunningham, D.B. Lacy, M. Mourez, R.J. Collier, The lethal and edema factors of anthrax toxin bind only to oligomeric forms of the protective antigen, PNAS 99 (2002) 7045–7048.
[11] C. Cheng, P. Kussie, N. Pavletch, S. Shuman, Conservation of structure and mechanism between eukaryotic topoisomerase I and site-specific recombinases, Cell 92 (1998) 841–850.
[12] The SETI project, University of Berkeley (http://www.setiathome.ssl.berkeley.edu).
[13] Distributed.net project (http://www.distributed.net).
[14] Top500 supercomputers (http://www.top500.org/).