Prospects for Clinical Application of Molecular Dynamics in HIV Treatment

Matthew Caldwell
CoMPLEX, University College London

Supervisors: Peter Coveney, Matthew Davies

5008 Words

March 9, 2007

Contents

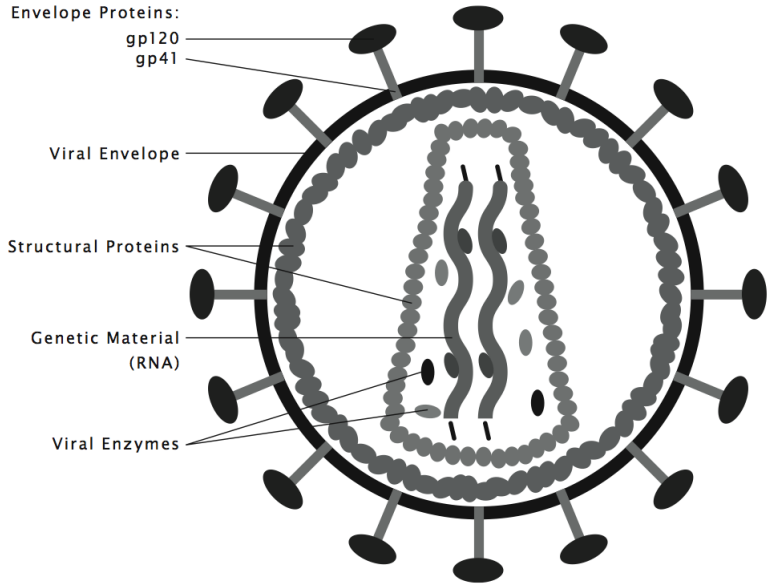

Figure 1: Structure of an HIV virion [9].

1 HIV

1.1 Lifecycle

The Human Immunodeficiency Virus, HIV, is a *lentivirus*, a particular form of *retrovirus*. Its infectious particles or virions consist of a lipid bilayer envelope containing several proteins needed for the viral lifecycle and two copies of the viral genome in the form of single-stranded RNA. (Figure 1.)

The virus principally targets the CD4+ T helper cells of the human adaptive immune system, although other cells such as macrophages may also be infected if they have CD4+ surface receptors. Membrane proteins in the viral envelope bind to these receptors to initiate entry to the cell. [6]

Once the virion membrane has fused with the host cell, the viral material is released into the cytoplasm. Here, the RNA genome is converted into double-stranded DNA by the viral enzyme reverse transcriptase, and this is then transplanted into the host's own genome by another viral enzyme, integrase. The integrated form is known as a *provirus*, and it may remain inactive in this state for a considerable time. If the cell divides, the provirus will be replicated along with the rest of the DNA and the daughter cells will also be infected.

Eventually, the provirus genes may get up-regulated, in which event the cell's own machinery transcribes the DNA into messenger RNA, some of which will be incorporated into new virus particles and some translated into three polyprotein precursor chains that are in turn used in the construction of new virions.[1] One of these chains—Env—is cleaved by a host enzyme into proteins needed for the virion envelope. The other two—Gag and Gag-Pol—are cleaved by the viral enzyme protease during the process of virion assembly, whereby a new virus particle is budded off from the host cell membrane. [19]

Infected cells are ultimately destroyed either by the viral activity itself, co-opting critical cell

[1]Both host and viral transcription factors can be involved in promoting viral expression. The host factors are also involved in triggering the T cell's immunological duties—so the cell is attacked from within just when it is needed to deal with attacks from without.

resources for the manufacture of virus particles, or by a host immune response to the distressed cells. As T cell numbers decline, so does the host's immune system. The person becomes susceptible to a range of 'opportunistic' secondary infections, leading to failing health and eventual death.

1.2 Drug Targets, Mutability and Resistance

Each of the main processes in the HIV lifecycle represents a potential target for drug therapy: if the virus can be prevented from going through these processes, the progress of infection will be arrested. Although some drugs exist or are under development to prevent viral entry into cells or proviral integration into the host DNA, the most successful inhibition targets to date have been reverse transcription and protease cleavage. [16]

Reverse transcriptase inhibitors were the first antiretroviral treatments to make any inroads into HIV disease progression and are still used extensively today. There are two basic classes: nucleoside/nucleotide analogue inhibitors (NRTIs), which pose as nucleotides and are incorporated by RT into the proviral DNA chain, causing premature termination; and non-nucleoside inhibitors (NNRTIs), which bind to RT at some location away from its active transcription site, inducing conformational changes that stop the active site from functioning.

Protease inhibitors (PIs) were introduced in the mid 1990s, and act by occupying the active site of the protease, thereby preventing it from performing the polyprotein cleavage necessary for new virus particles to be infective.

PIs and NRTIs are *competitive* inhibitors, which depend on the enzyme binding them in preference to its normal functional substrate. NNRTIs are non-competitive, but nevertheless need to bind their targets correctly. All of these drugs are therefore highly sensitive to the specific conformation of the corresponding viral enzymes.

Unfortunately, whereas forward transcription is a complex, high-fidelity cellular process with extensive error-checking and correction, reverse transcription is more haphazard and errors are common—some sources report rates as high as 1 in 1700 bases, although not uniformly distributed across the 9.2 Kb of the HIV genome [14]. While most such mutations will lead to loss of fitness, some will not, and if they happen to chance upon a competitive advantage they will multiply; the virus is, in effect, constantly exploring the fitness landscape of each patient's body.

In the presence of a strong selective pressure, such as a successful drug, mutations that result in structural changes which reduce drug efficacy will quickly win out, even if they would otherwise lead to reduced fitness.[2] As a result, HIV has a strong tendency to develop drug resistance over time within individual patients, and many resistant strains are now found in the population. Although they remain less common, new infections can occur with any of these strains.

1.3 Clinical Practice

In the immediate aftermath of infection with HIV, the virus multiplies rapidly and levels of viræmia are high. However, the body soon mounts an apparently-successful immune response and free virus can be almost eliminated. Large quantities of provirus remain in the CD4+ T cells, the majority inactive at any particular time. During this latent phase, which often lasts for many years, the patient remains asymptomatic, but his or her T cell population is gradually

[2] When drug regimens are discontinued, the virus will often revert to wild-type in a relatively short period, suggesting that the resistant mutants are indeed at some competitive disadvantage on a level playing field. [12]

whittled away. Eventually, the immune system is overwhelmed, viræmia returns to high levels and the infection becomes chronically symptomatic.

Although infection is most often diagnosed during the latent phase, drug treatment is usually not initiated until the natural latency is coming to an end: typical indications for commencement are a CD4+ count of less than 200 cells/mm^3 or progression to symptomatic immune suppression. Benefits of earlier intervention are debated, but in general are considered to be outweighed by the risks of selecting for resistance and exhausting treatment options that will be more important later on. [12]

HIV drugs are invariably used in combination to reduce the probability of resistance developing in the short term: while a patient is likely to have viral strains with mutations reducing susceptibility to any one drug, the chances of combined mutations conferring multiple resistance in a single strain are lower. However, the choice of combination in the context of each particular patient may be a delicate business. Resistance is not the only concern—patients may be more or less sensitive to toxicities or may have other conditions that affect the choice, such as hepatic damage—but it is a major ingredient in the decision.

Several forms of resistance testing are available for use both prior to treatment and for ongoing monitoring. The most common are genotypic assays in which some or all of the viral genome—including at least the reverse transcriptase and protease genes—is amplified by PCR and sequenced. Since a patient may carry several different viral strains, this process is not guaranteed to give an exact picture of the HIV infection, but it will usually provide a reasonable estimate of the predominant strain(s).

The viral sequence data is currently used in *knowledge-based* estimation of drug resistance. A number of specific mutations are known to be associated with resistance to particular treatments and these mutations are treated as *markers*: essentially, as keys into a lookup table. The knowledge may reside in computer databases or in the brains of experienced clinicians; it may derive from clinical trials, biochemical experimentation or prior molecular modelling. Information is constantly accruing, but it remains necessarily historical and reactive. In particular, such a system cannot predict the effect of novel mutations or combinations.

Because of the mutability of the virus, HIV treatment is inevitably a dynamic process. Correctly identifying drug susceptibility *once* is not sufficient. Treatment efficacy must be regularly monitored and drug regimens adjusted as circumstances change.

2 Molecular Dynamics

2.1 Overview

Molecular Dynamics constitutes an attempt to model physical, chemical and biochemical behaviour from the ground up, in terms of the interactions between individual atoms *over time*. The primacy of temporal evolution—dynamics—is what distinguishes MD from other molecular modelling endeavours, although there is otherwise considerable overlap. The MD rubric encompasses a number of approaches, but all are based primarily on classical Newtonian modelling; quantum mechanical models are generally intractable for even rather small many-body systems, let alone those on a biologically relevant scale.

Classical approaches are normally adequate for conformational and binding affinity calculations pertinent to biological systems, but they cannot account for changes in electronic configuration (which is to say, chemical reactions). Some hybrid models, in which the bulk of the system

is handled classically but quantum effects are included at active sites, can be applied to this problem.

Even within a purely classical framework, there are a number of variations on the MD theme. Systems of interest, such as enzymes and their substrates in explicit solvent, can involve a very large number of atoms with many interactions between them. In order to render the calculations computationally feasible, a whole arsenal of approximations may be brought to bear, and different MD applications will include different combinations of these. It is important to recognise that, although the constituent elements have some theoretical basis, there is a strong *empirical* component too: an approximation without a reasonable correspondence to observable reality is of little use to anyone.

2.2 Modelling Methods

A classical MD simulation considers only atomic *nuclei*: electrons contribute to the net charge of an atom and to the molecular bond geometry, but are not modelled in their own right; this is justified by the Born-Oppenheimer approximation, which asserts that electronic distribution is so much more sprightly than nuclear arrangement as to be relatively negligible.

The nuclei are treated as point masses in a classical many-body system, whose behaviour is governed by Newton's second law:

$$\mathbf{F} = m\frac{d^2\mathbf{r}}{dt^2}$$

The force $\mathbf{F}$ acting across the system depends on the spatial derivative of the potential energy function, V, which in turn depends on the interactions between the bodies at a particular instant in time.

The basic steps of a simulation are therefore:

- Determine the *force field* or *potential energy surface* acting on all atoms based on their positions, species, interactions and any specified constraints.
- Integrate this force field numerically over some small timestep to update the positions and momenta of all atoms.
- Repeat.

Force field estimation is the most expensive—and most contentious—stage of the process. Several standard definitions exist, in multiple implementations, and there are many custom variants according to the nature of the application and the available data.[3]

A typical functional form for such a force field is:

$$V(\mathbf{r}^N) = \sum_{bonds}\frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{angles}\frac{k_i}{2}(\theta_i - \theta_{i,0})^2 + \sum_{torsions}\frac{V_n}{2}(1 + cos(n\omega - \gamma))$$
$$+\sum_{i=1}^{N}\sum_{j=i+1}^{N}\left(4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right)$$

[3] Availability of several popular codes on high performance grid platforms—discussed in §4.1—will probably encourage greater standardisation in future.

The first three terms represent the intramolecular interactions of pairs of atoms separated by 1, 2 and 3 bonds respectively, while the fourth represents non-bonded interactions, in this case comprising a 12-6 Lennard-Jones potential for van der Waals interactions and Coulombic potential for electrostatic interactions.

The precise formulæ used to calculate each term will vary between force field implementations. The choices are in part determined by factors such as the particular atomic species involved, the intended purpose of the model and the number of parameters that must be experimentally determined for it, as well as by computational efficiency and numerical stability. For example, the repulsive r^{-12} term in the above Lennard-Jones potential is somewhat arbitrary—quantum mechanics would suggest an exponential form—but is computationally expedient. [10]

The numerical integration step can, similarly, be performed in many ways. A common choice is the Verlet algorithm, which has a higher order error term than, for example, Euler's method. Verlet's trick is to average changes at two time steps, leading to cancellation of terms in δt^3 from the positional calculation:

$$\mathbf{r}(t+\delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \delta t^2 \mathbf{a}(t) + \mathrm{O}(\delta t^4)$$

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\delta t) - \mathbf{r}(t-\delta t)}{2\delta t} + \mathrm{O}(\delta t^2)$$

This method has some disadvantages—notably a loss of numerical precision due to the different sizes of the terms in the calculation of $\mathbf{r}$—so a number of variations on the algorithm are sometimes substituted.

The maximum time step that can be used in a simulation without aliasing is governed by the fastest vibrations in the system, usually those of bonded interactions of hydrogen atoms. Strictly, these impose a limit of the order 0.5 fs. However, it is common to apply special constraints to these interactions—such as the SHAKE algorithm—which effectively treat the bonds as rigid units. This increases the allowable step size to the 1-2 fs range. [1]

By contrast, the kinds of biological processes we are interested in modelling with MD operate on timescales of tens of nanoseconds and up, essentially without limit. Clearly, that's *a lot* of simulation steps. Even with ongoing algorithmic improvements, increases in processor power and the use of distributed computing resources, running arbitrarily long simulations is unrealistic.

Instead, several techniques are employed to ensure that the initial conditions are close to the states of interest, making use of structural information determined experimentally via X-ray crystallography and NMR. Such data are not obtained under physiological conditions, so part of the initialisation process is to gradually warm up the system to a more relevant temperature by adding appropriately-distributed kinetic energy.

Changes in *free energy* are important for determining binding affinities between molecules, like those for the drug/enzyme interactions of concern here. Although differences can be estimated from explicit simulations—for example, via careful interpolation around *thermodynamic cycles*, as in the Free Energy Perturbation and Thermodynamic Integration techniques—such calculations are both computationally expensive and feasible only in relatively constrained circumstances.

A common alternative is to use an *implicit solvent* approximation, such as MM/PBSA (molecular mechanics/Poisson-Boltzmann surface area), in which the free energy is decomposed into

a molecular mechanics term for the solute molecules and a solvation term for interactions with a continuum of surrounding water:

$$\Delta G_{bind} = \Delta G_{MM} + \Delta G_{solv}$$

Here, ΔG_{MM} is calculated in terms of atomic interactions according to an explicit force field such as described above, while ΔG_{solv} is in turn broken down into polar and non-polar terms:

$$\Delta G_{solv} = \Delta G_{polar} + \Delta G_{nonpolar}$$

Polar solvation energy is derived from the Poisson-Boltzmann equation or some simplification thereof, such as the Generalised Born model; while the non-polar is taken to be proportional to the solvent accessible surface area [7]. This formulation fails to account for configurational entropy, so a further term is sometimes added, for example calculated by normal mode analysis. [15]

It's worth noting that, although MD is notionally deterministic, the range of methodologies in use, along with varying initialisation and equilibration protocols and parameterisations can sometimes give rise to substantively different results; an example is given in §3.4. In the case of such contradictions—or indeed, results at odds with experiment—MD must be treated with scepticism. However, it is appropriate to give credence to MD where it produces answers that are consistent across different implementations and in agreement with empirical measures.

3 Modelling Enzyme/Drug Interactions

3.1 Rationale

As described in §1.2, the operation of anti-retroviral drugs is specific to the precise conformation of their target enzymes. A drug must have a high probability of binding its target to be effective, and this binding affinity can be sensitive to small changes in enzyme sequence. There will be a strong evolutionary imperative for HIV to exploit any such sensitivity given the chance. It is therefore important to have a good understanding of the specifics of interactions between drugs and their targets, both in the design and development of new drugs and in the use of existing therapies in a clinical setting.

Since the interactions of interest are dynamic and occur on the molecular level, MD is an obvious tool to apply to the problem, and indeed MD techniques are often used in the pharmaceutical industry alongside more static molecular modelling methods. Application to understanding interactions with mutant enzymes and the mechanisms of resistance has been less widespread, but is gaining ground as the technology matures; some of this research is discussed below.

MD methods can, of course, be applied to the whole range of drugs for HIV, whether in common use, like RTIs, or more novel treatments targeting integrase and viral fusion. However, perhaps because the enzyme itself is relatively tractable, protease inhibitors have attracted the most attention, and we shall focus on those here.

3.2 HIV Protease Inhibitors

HIV protease is a symmetric homodimer with a characteristic 'bulldog' shape. The central active site is formed by the dimerisation interface, with a dyad of aspartic acid residues—one

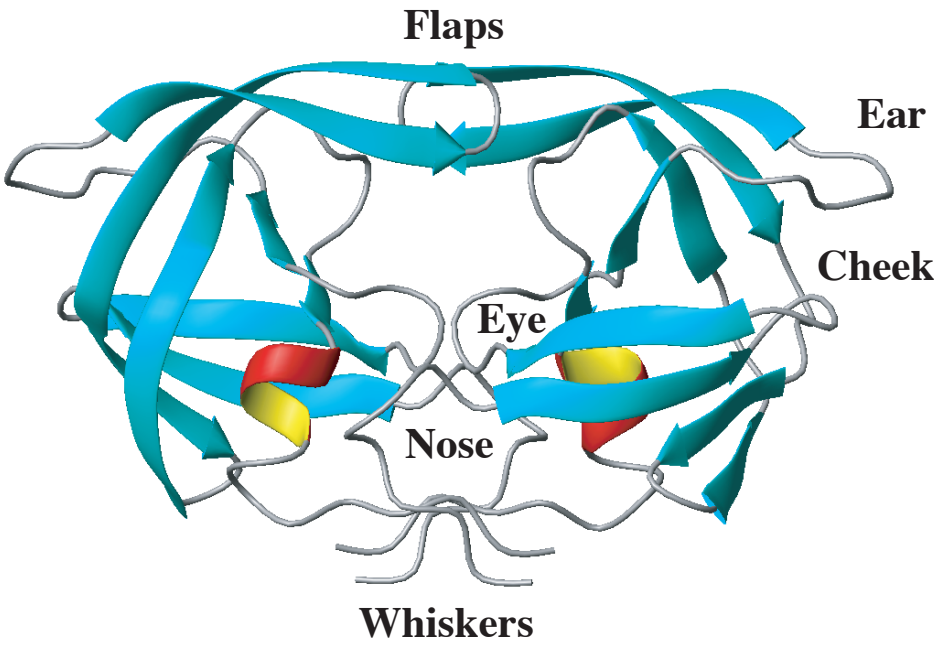

Figure 2: 'Bulldog' structure of HIV protease.

from each monomer—at its base which, in conjunction with a conserved water molecule to which both residues are hydrogen bonded, catalyses cleavage of the precursor polyproteins. A pair of flexible hairpin β-sheets, known as the *flaps*, surround the active site and control access to it. Other areas of the structure are named for their resemblance to features of the 'canine' face: *ears, cheeks, eyes, nose, whiskers.*[4] (Figure 2.)

For notational convenience, residues in the monomer chains are referred to by single letter amino acid code and sequence position: the active site aspartic acid is D25. A similar form is used for mutations, but with a suffix for the substituted amino acid: G48V indicates that the mutant has a valine at position 48 instead of the wild-type glycine.

There are currently ten FDA-approved protease inhibitor drugs.[5] All are competitive substrate analogues that occupy the active site in place of the Gag and Gag-Pol polyproteins. Most are peptides or peptidomimetic, though there are also some nonpeptidic PIs. [18]

Inhibitory effects depend on preferentially binding the active site, of course, but also on *remaining* bound—drugs with a high dissociation rate will still leave many opportunities for the polyproteins to be cleaved. Resistant mutant proteases appear to dissociate from inhibitors more readily. [11]

3.3 Investigating the Mechanics of PI Resistance

Perryman et al [13] applied MD to the binding behaviour of PI drugs complexed with both the wild-type protease and a double mutant strain, V82F/I84V, known to exhibit significant inhibitor resistance in patients. The mutations are not at the active site, and static conformational differences between the two enzymes—as determined by X-ray crystallography—are apparently

[4]This slightly fanciful terminology, due to Perryman et al, supercedes an earlier one that was prejudicially based on a particular functional interpretation.

[5]One is a prodrug version of another, so there are really only nine distinct agents.

small, whereas the binding affinities of V82F/I84V for all current PI drugs are much worse than for the wild-type. The primary purpose of the study, then, was to investigate qualitative changes in dynamic behaviour that might explain this difference.

Because of the role played by the flaps in managing access to the active site, particular attention was paid to the flap dynamics, and also to specific conformational changes in the flap *tips* (G48-G49-I50-G51-G52) that had previously been shown to correlate with flap opening.

The enzymes were simulated for 22 ns using an all atom setup with explicit water. The initial configuration was derived from existing crystallographic data of the enzyme in a bound complex with the drug tipranavir. Although the behavioural variations between the two systems over the simulation period were not dramatic, there was certainly some difference. In particular, the mutant showed both a greater tendency toward flap opening and also more of the flap tip 'curling' that precedes opening. The repeated relationship between curling and flap opening supports the earlier research suggesting a link between the two motions.

In addition, simulations indicated an anti-correlation between the flap opening and the separation between residues in the ear and cheek. This relationship could make the ear-cheek region a possible target for a new class of allosteric PIs, although no such have yet been developed.

3.4 Predicting Binding Affinities for Mutant Proteases

Given a selection of drug candidates and sequence data from the patient, it is desirable to rank the binding affinities—and hence the likely effectiveness—of each to the patient-specific enzymes. Recent papers by Hou & Yu [8] and by Stoica, Sadiq & Coveney [15] demonstrate that a combination of MD and MM/PBSA can generate such binding affinity data with good agreement to published results from enzyme inhibition assays.

Hou & Yu studied the wild-type protease and the V82F/I84V mutant used by Perryman et al, complexed with the inhibitor amprenavir and two related drugs currently under development. Stoica et al compared the wild-type with the L90M single mutant and the G48V/L90M double, complexed with saquinavir.

Although there were differences in methodology, both papers report a close correlation between predicted affinities and experimental results, and in particular a correct ranking of affinities in all cases. The analyses of the contributions of the different energy terms are also broadly similar: in both cases the van der Waals, electrostatic and non-polar solvation interactions favour binding, while polar solvation and entropic contributions oppose.

It is suggested that the decreased binding affinity in the mutant strains is principally caused by changes in the polar energy contributions due to destabilisation of electrostatic interactions at the binding site. This is in contrast to Perryman et al's assertion that increased flap dynamics are responsible, and indeed Hou & Yu's simulations show the opposite effect (they spend some time specifically addressing this contradiction). This underlines the need for validation of simulation-based results against other methods.

At present, the binding affinity rankings, which are essentially quantitative, appear to be more plausible than more qualitative predictions as to the molecular mechanisms of drug resistance. While the latter will doubtless improve, the former have more immediate potential for practical use in improving HIV treatment.

4 Requirements for Practical Application

Given the mutable nature of HIV, treatment regimens need to be carefully selected for each patient, and constantly maintained. While remarkable progress has been made in management of the disease over the last two decades, there is still plenty of room for improvement in current strategies. It should be clear from the foregoing discussion that MD and related methods such as MM/PBSA could have something to offer in this regard, since they can take into account patient-specific information that may have no place in current treatment lore. This is especially so when it comes to synergistic or compensatory interactions between mutations.

Nevertheless, significant obstacles remain before these methods are likely to be embraced beyond the research environment and accepted for routine medical use. Some of these obstacles are technical, but those are probably less important: they are subject to purely technological solutions, many of which are already in sight. Potentially more problematic are the cultural and administrative barriers.

4.1 Computational Tractability & the Grid

The sort of MD investigations described above have been major research projects, taking teams of scientists months or years and requiring enormous amounts of data preparation, computer time and novel analysis. This situation is improving, but the kind of ready availability and ease of use that would be needed for routine medical application is still some way off. However, the likely form of future developments that will make this a possibility can be discerned and the idea no longer seems far-fetched.

It is estimated that in order for a test to be of clinical use it must have a turnaround time of no more than two weeks.[6] Running multiple simulations of the interactions between different inhibitors and patient-specific mutant enzymes—and, moreover, doing so regularly and for many patients—will require very significant computer resources. Moore's Law notwithstanding, it will be many years before that sort of power is available on everyday workstations, but high performance resources do exist and the ongoing developments in textitgrid computing can provide access to them.

Grids—distributed computing systems that work transparently across administrative domains—are an active area of computational research, and several large and powerful systems are already in use [5]. To date, they have tended to be rather unwieldy and geared more towards exploring the subject of grid computing itself than providing supercomputing capabilities to scientific—let alone clinical—end users. The ideal of application transparency has been more or less nonexistent, and would-be clients often have to wrestle with heavyweight middleware and layers of administrative complexity that have nothing to do with the ostensible task at hand. [3]

But such teething problems are to be expected for an emerging technology and improvements are being made. One notable example is the RealityGrid/OMII Application Hosting Environment, a web-services based wrapper that hides much of the complexity of existing grid middlewares. AHE provides a unified, user-oriented interface with particular support for molecular dynamics: several major MD applications are hosted on the service, greatly simplifying the process of deploying simulations to the grid. It is already being used to model HIV drug/enzyme interactions with turnaround times on a clinically-relevant scale. [4]

[6] Phenotypic resistance assays take longer than this, which may be one reason they are less widely used than the genotypic ones.

4.2 Standardisation & Ease of Use

Even with the relatively standard MD applications available on AHE, setting up and running simulations is not a trivial undertaking. These programs are intended to be general purpose and, as such, require a problem to be specified in excruciating detail, with vast quantities of positional and topological data and many operational parameters.

While this makes sense in a research context, where every simulation is a unique adventure with its own idiosyncratic features, it isn't appropriate for routine diagnostic purposes. Instead, the breadth of research experience needs to be assimilated into a single body of input data and a single suite of tests that can be invoked automatically with nothing more than the patient's viral sequence data, generating a standard set of results in a form accessible to the clinician.

The field is probably not yet mature enough for a ready consensus to be reached on the constituents of such a system, but a few observations can be made:

- Given the relative expense of molecular dynamics, it will almost certainly be worth expending a lot of (internal) bioinformatics effort to optimise the nature and quantity of simulation.
- A large number of protease structures are already known and the system should be able to select the most appropriate one to use as a starting point, based on the mutations in the supplied sequence.
- For clinical credibility, it will almost certainly be necessary to perform multiple runs with varying initial conditions and evaluate the results statistically. If there is no consensus on optimum force field definition and job parameterisation, it may also be necessary to include variations of these details too.
- In theory, ongoing use combined with clinical outcome data should allow for monitoring and continuous refinement of such a system. However, this may raise significant questions of confidentiality and regulation.

While a program of this nature is not going to materialise overnight, there is nothing about it that is beyond the reach of what we already know. It is certainly more credible that these tasks will be comprehensively automated to the point where the end user takes them for granted than to suppose that busy doctors or lab technicians will ever immerse themselves in the nitty gritty of performing them by hand.

4.3 Integration with Clinical Practice

It seems evident that a system like the above, freed from the current arcane knowledge requirements of molecular modelling, could be added to existing testing procedures easily. Doctors would systematically receive binding affinity ranking estimates along with sequence data, T cell counts and viral load, using the results as they saw fit. There are, however, some obvious administrative obstacles.

The regulatory situation regarding such testing is unclear. In the US, it would appear to fall under the terms of the current work-in-progress FDA guidelines on 'In Vitro Diagnostic Multivariate Index Assays':

> [A]n IVDMIA is a test system that employs data, derived in part from one or more in vitro assays, and an algorithm that usually, but not necessarily, runs on software, to generate a result that diagnoses a disease or condition or is used in the cure, mitigation, treatment, or prevention of disease. [17]

It would therefore probably be subject to extensive regulatory approval procedures before being accepted for clinical use. This is a costly business and is most often undertaken by well-resourced companies in the expectation of future income streams from commercial exploitation of the tests. Exactly how that might square with a system that depends on high-performance grid computing resources—typically built and maintained by the public sector—remains to be seen. If the results are sufficiently beneficial, presumably some accommodation would be reached. Still, such administrative concerns are bound to play a part.

Another possible obstacle to acceptance of the MD-based testing is medical culture. Anecdotally, at least, many doctors have a mistrust of computer simulation methods, seeing them as unrealistic and unreliable. There is a common preference for solid clinical experience over the purely mechanical. Of course, once such a test entered routine use, it would be incorporated into clinical experience—if it produced consistently useful results, taking those results on board would become part of HIV treatment best practice, just as have existing analyses. In the meantime, however, some level of scepticism is only to be expected. [2]

5 Conclusions

Molecular Dynamics and other low-level simulation methods can provide a valuable understanding of the biophysical mechanisms by which diseases such as HIV work. In the light of the high mutation rates of the virus, and the consequent prevalence and variability of drug resistance, such insights can make valuable contributions to patient-specific treatments.

Fulfilment of this potential will depend on many factors, not least the ready availability of high performance computing resources as promised by the nascent technologies of the grid. Many obstacles remain before such approaches can be widely adopted, but none are insuperable given our current knowledge, and it is reasonable to suppose that things will get easier in years to come. The field is young.

While this discussion has focused on the particular case of protease inhibitors, the techniques apply more broadly. We should expect equivalent testing in relation to other drug classes to form a part of future systems for tailoring treatments to the particular spectrum of susceptibilities of each patient.

MD does not represent a silver bullet with respect to HIV/AIDS, nor to the matter of patient-specific treatments for other diseases, but it does offer another weapon in the armoury of techniques being brought to bear. Given the potential benefits, the value of further research and development in this direction is unmistakable.

References

[1] Oren M. Becker and Martin Karplus. *A Guide to Biomolecular Simulations*. Springer, 2006.

[2] Paul R. Billings. Three barriers to innovative diagnostics. *Nature Biotechnology*, 24:917–918, 2006.

[3] Jonathan Chin and Peter V. Coveney. Towards tractable toolkits for the Grid: a plea for lightweight, usable middleware. Technical report, UK e-Science Technical Report, 2004.

[4] P. V. Coveney, R. S. Saksena, S. J. Zasada, M. McKeown, and S. Pickles. The application hosting environment: Lightweight middleware for grid-based computational science. *Computer Physics Communications*, (176):406–418, 2007.

[5] Peter V. Coveney. Scientific grid computing. *Philosophical Transactions of the Royal Society A*, 363:1707–1713, 2005.

[6] Robert W. Doms and Didier Trono. The plasma membrane as a combat zone in the HIV battlefield. *Genes Dev.*, 14(21):2677–2688, 2000.

[7] Federico Fogolari, Alessandro Brigo, and Henriette Molinari. Protocol for MM/PBSA molecular dynamics simulations of proteins. *Biophysical Journal*, 85:159–166, July 2003.

[8] T. Hou and R. Yu. Molecular dynamics and free energy studies on the wild-type and double mutant HIV-1 protease complexed with amprenavir and two amprenavir-related inhibitors: Mechanism for binding and drug resistance. *Journal of Medicinal Chemistry*, 2007.

[9] Patricia Kahn, editor. *AIDS Vaccine Handbook: Global Perspectives (2nd Edition)*. AIDS Vaccine Advocacy Coalition, 2005.

[10] Andrew R. Leach. *Molecular Modelling: Principles and Applications, 2nd Edition*. Pearson Education, 2001.

[11] Barbara Maschera, Graham Darby, Giorgio Palu, Lois L. Wright, Margaret Tisdale, Richard Myers, Edward D. Blair, and Eric S. Furfine. Human Immunodeficiency Virus: Mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. *J. Biol. Chem.*, 271(52):33231–33235, 1996.

[12] Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Technical report, Department of Health and Human Services, October 2006.

[13] Alexander L. Perryman, Jung-Hsin Lin, and J. Andrew McCammon. HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: Possible contributions to drug resistance and a potential new target site for drugs. *Protein Science*, 13:1108–1123, 2004.

[14] JD Roberts, K Bebenek, and TA Kunkel. The accuracy of reverse transcriptase from HIV-1. *Science*, 242(4882):1171–1173, 1988.

[15] Ileana Stoics, S. Kashif Sadiq, and Peter V. Coveney. An accurate binding affinity ranking method for inhibitor-bound mutant HIV-1 proteases. *Prepress*, 2007.

[16] Zelalem Temesgen, David Warnke, and Mary J Kasten. Current status of antiretroviral therapy. *Expert Opinion on Pharmacotherapy*, 7(12):1541–1554, 2006.

[17] US Food and Drug Administration. Draft Guidance for Industry, Clinical Laboratories, and FDA Staff: In Vitro Diagnostic Multivariate Index Assays. *http://www.fda.gov/cdrh/oivd/guidance/1599.pdf*, September 2006.

[18] US Food and Drug Administration. Drugs used in the treatment of HIV infection. *http://www.fda.gov/oashi/aids/virals.html*, 2007.

[19] Wei-Kung Wang, Mao-Yuan Chen, Che-Yen Chuang, Kuan-Teh Jeang, and Li-Min Huang. Molecular biology of human immunodeficiency virus type 1. *Journal of Microbiology, Immunology and Infection*, (33):131–140, 2000.