Enveloped viruses understood via multiscale simulation: computer-aided vaccine design

Z. Shreif · P. Adhangale · S. Cheluvaraja · R. Perera · R. Kuhn · P. Ortoleva

Originally published in the journal Sci Model Simul, Volume 15, Nos 1–3, 363–380.
DOI: 10.1007/s10820-008-9101-5

Abstract Enveloped viruses are viewed as an opportunity to understand how highly organized and functional biosystems can emerge from a collection of millions of chaotically moving atoms. They are an intermediate level of complexity between macromolecules and bacteria. They are a natural system for testing theories of self-assembly and structural transitions, and for demonstrating the derivation of principles of microbiology from laws of molecular physics. As some constitute threats to human health, a computer-aided vaccine and drug design strategy that would follow from a quantitative model would be an important contribution. However, current molecular dynamics simulation approaches are not practical for modeling such systems. Our multiscale approach simultaneously accounts for the outer protein net and inner protein/genomic core, and their less structured membranous material and host fluid. It follows from a rigorous multiscale deductive analysis of laws of molecular physics. Two types of order parameters are introduced: (1) those for structures wherein constituent molecules retain long-lived connectivity (they specify the nanoscale structure as a deformation from a reference configuration) and (2) those for which there is no connectivity but organization is maintained on the average (they are field variables such as mass density or measures of preferred orientation). Rigorous multiscale techniques are used to derive equations for the order parameters dynamics. The equations account for thermal-average forces, diffusion coefficients, and effects of random forces. Statistical properties of the atomic-scale fluctuations and the order parameters are co-evolved. By combining rigorous multiscale techniques and modern supercomputing, systems of extreme complexity can be modeled.

Keywords Enveloped viruses · Structural transitions · All-atom multiscale analysis · Multiscale computation · Liouville equation · Langevin equations

Z. Shreif · P. Adhangale · S. Cheluvaraja · P. Ortoleva (✉)
Center for Cell and Virus Theory, Department of Chemistry,
Indiana University, Bloomington, IN 47405, USA
e-mail: ortoleva@indiana.edu

R. Perera · R. Kuhn
Department of Biological Sciences, Purdue University,
West Lafayette, IN 47907, USA

S. Yip & T. Diaz de la Rubia (eds.), *Sci Model Simul*, DOI: 10.1007/978-1-4020-9741-6_19

1 Introduction

Deriving principles of microbial behavior from laws of molecular physics remains a grand challenge. While one expects many steps in the derivation can be accomplished based on the classical mechanics of an N-atom system, it is far from clear how to proceed in detail due to the extreme complexity of these supra-million atom systems. Most notably, molecular dynamics (MD) codes are not practical for simulating even a simple bionanosystem of about 2 million atoms (e.g. a nonenveloped virus) over biologically relevant time periods (i.e. milliseconds or longer). For example, the efficient MD code NAMD, run on a 1024-processor supercomputer [1], would take about 3000 years to simulate a simple virus over a millisecond; the largest NAMD simulation published to date is for a ribosome system of approximately 2.64 million atoms over few nanoseconds only [2].

We hypothesize that a first step in the endeavor to achieve a quantitative, predictive virology is to establish a rigorous intermediate scale description. Due to their important role in human health, complex structure, and inherent multiscale nature, enveloped viruses provide an ideal system for guiding and testing this approach. Experimental evidence suggests that an enveloped virus manifests three types of organization:

- an outer protein net that can display several well-defined structures; for each structure, proteins maintain long-lived connectivity with specific nearest-neighbor units (Fig. 1);
- a sea of membranous material below the protein net; this subsystem consists of phospholipid molecules whose nearest-neighbors are continuously changing but for which there is long-lived structure on-the-average. Also, biological membranes display liquid-crystal transitions [3] either autonomously or as promoted by proteins traversing the membranous subsystem; and
- genomic RNA or DNA complexed with proteins in which there is long-lived connectivity between nucleotides but which, as evidenced by cryo-electron microscopy and X-ray diffraction data, often lack well-defined structure [4].

We hypothesize that these three types of organization and the interplay of their stochastic dynamics is the essence of principles governing the structural transitions and stability of enveloped viruses.

Developing a quantitative understanding of enveloped viruses is of great importance for global health. Human pathogenic enveloped viruses include Dengue (Fig. 1) and HIV. Understanding the mechanisms underlying virus entry and hijacking of host cell processes is a main step in preventing the often fatal virus infections. The aim is not only to be able to attack the virus in question (or simply prevent its docking proteins from binding to the host cell receptors), but also to use viruses for therapeutic delivery. Since viruses have a natural ability to find and penetrate host cells, using them as a means to deliver genes, drugs, and other therapeutic agents holds great promise in medical advancement. Enveloped viruses provide a natural choice as they are able to entirely fuse inside the cell before delivering their payload [7]. Our strategy is to develop a predictive whole-virus model that serves as a basis of a computer-aided antiviral vaccine and drug design capability. Furthermore, this predictive model could be a key element of a system for forecasting the potential pandemic of a given strain via an assessment of computer-generated mutants, and similarly for testing the effects of a library of potential antiviral drugs.

To achieve these practical goals, and to arrive at a fundamental understanding of complex bionanosystems, we suggest starting from the laws of molecular physics. These, as considered here, are Newton's equations for an N-atom system. For an enveloped virus, N is on the order of 10^8. However, the conceptual framework within which one discusses viral phenomena does

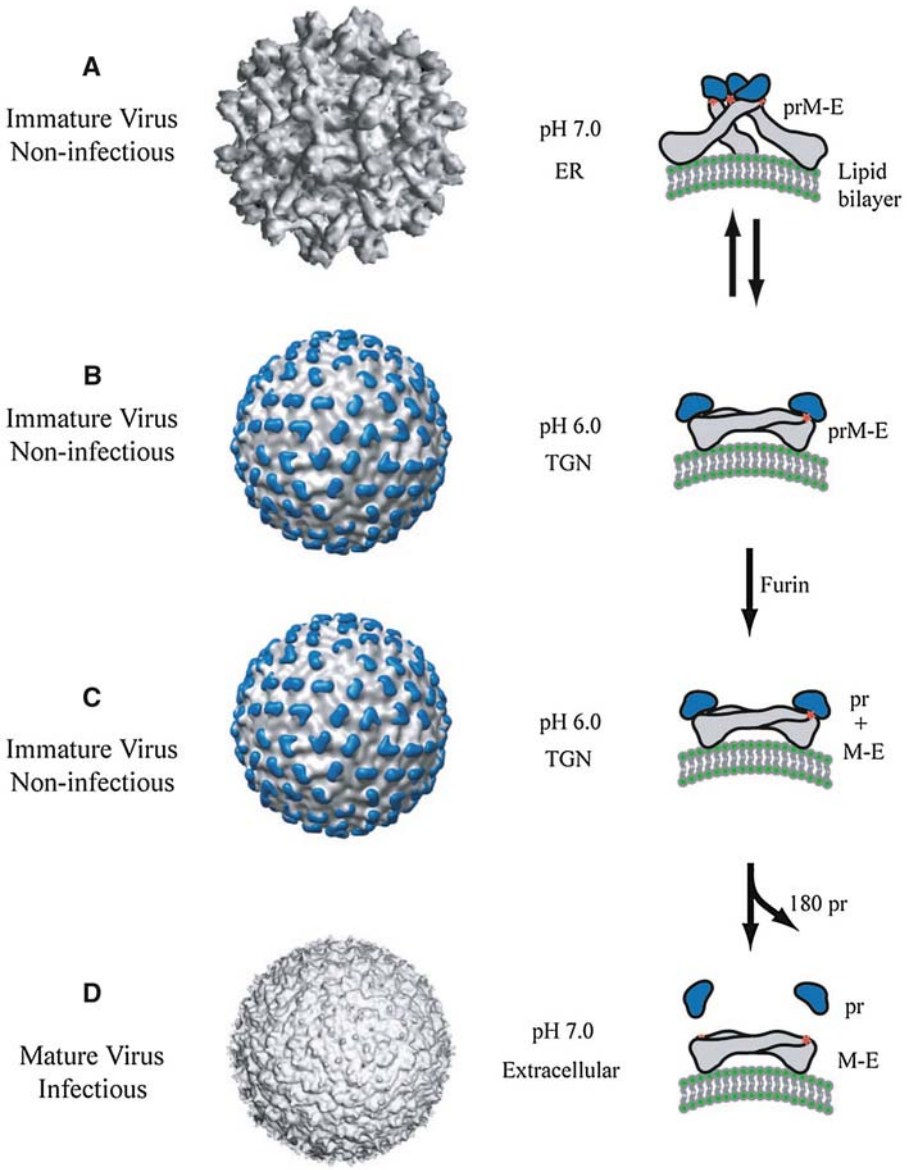

Fig. 1 **A** Cryo-EM reconstruction of the immature Dengue virion at neutral pH. **B** Cryo-EM reconstruction of the immature virion at low pH. **C** Cleavage of the prM protein into its 'pr' peptide and M protein by the host endoprotease, furin. **D** The cryo-EM reconstruction of the mature virion. From Refs. [5,6]

not involve keeping track of the positions and momenta of all the atoms. Nonetheless, an all-atom description is required to derive the principles of enveloped viral behavior from laws of molecular physics. Processes involved in viral behavior include the 10^{-14} second atomic vibration and collisions and the millisecond or longer overall structural transitions. In addition, various size scales are involved: The scale of the nearest-neighbor atom distance (a few angstroms) to the diameter of the enveloped virus and a closely associated aqueous

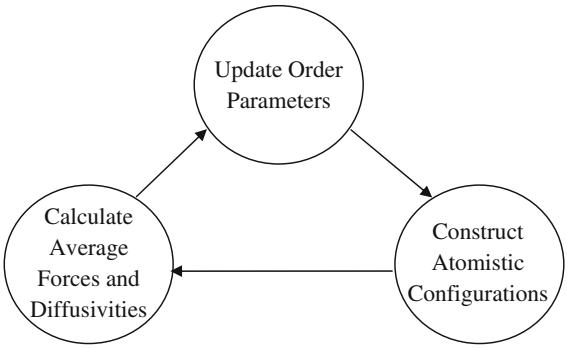

Fig. 2 Order parameters characterizing nanoscale features affect the relative probability of the atomistic configurations which, in turn, mediates the forces driving order parameter dynamics. This feedback loop is central to a complete multiscale understanding of nanosystems and the true nature of their dynamics. This loop is also the schematic workflow of our AMA/ACM computational strategy.

layer (several thousand angstroms). As the short scale phenomena affect the larger scale ones, and conversely (see Fig. 2), viruses have a strongly multiscale character.

Multiscale techniques have been discussed extensively in the literature [8–31]. These studies start with the Liouville equation and arrive at a Fokker-Planck or Smoluchowski type equation for the stochastic dynamics of a set of slowly evolving variables (order parameters). Of particular relevance to the present study are recent advances [20–31] wherein it was shown one could make the hypothesis that the N-atom probability density is a function of the $6N$ atomic positions and momenta both directly and, via a set of order parameters, indirectly. It was shown that both dependencies could be reconstructed when there is a clear separation of timescales, and that such an assumed dual dependence is not a violation of the number ($6N$) of classical degrees of freedom. Technical advances were also introduced which facilitated the derivation of the stochastic equations of the order parameter dynamics and allowed for an atomistic description of the entire system (i.e. within the host medium and the nanostructures). Furthermore, it was shown in earlier work [30,31] how to transcend the conceptual gap between continuum and all-atom theories; starting from the more fundamental all-atom formulation and, via a deductive multiscale analysis, equations coupling field variables with selected facets of the all-atom description were derived.

For the present study, we suggest that a mixed continuum/all-atom approach is a natural framework for achieving an integrated theory of enveloped viruses. As there are several types of molecular organization to account for, we introduce distinct types of order parameters to characterize them. Details on these order parameters are provided in Sects. 2 and 3. Common features of both types of order parameters are as follows:

- they are expressed in terms of the all-atom state (i.e. the $6N$ positions/momenta of the N atoms) in terms of which the basic laws of molecular physics are posed;
- they evolve on timescales relevant to virology (i.e. microseconds or longer, and not 10^{-14} seconds as for atomic vibration/collisions);
- they characterize the nanoscale features of interest to virology (e.g. outer protein net structure, genomic structure, and liquid-crystal order in the membranous subsystem); and
- they form a complete set, i.e. they do not couple to other slow variables not considered in the analysis.

In Sect. 4, we develop a multiscale theory of viral dynamics and, in the process, clarify the need for the above criteria on the nature of order parameters. We derive rigorous equations for their stochastic dynamics, notably structural fluctuations and transitions. In Sect. 5, we discuss our multiscale simulation method and present results for the STMV virus to illustrate the feasibility of our approach. Applications and conclusions are discussed in Sect. 6.

2 Order parameters for connected structures

Order parameters for subsystems of connected atoms have been constructed as generalized Fourier amplitudes [23,24]. They represent the degree to which a structure is a deformation of a reference configuration (e.g. a cryoTEM reconstruction). For Fourier analysis one uses sine and cosine basis functions. In our approach, other basis functions are introduced as follows.

The system is embedded in a volume Ω. Orthogonal basis functions $U_{\underline{\ell}}\left(\vec{r}\right)$ for point $\vec{r}$ in Ω with triplet labeling index $\underline{\ell}$ are introduced. The basis functions are periodic if computations are carried out using periodic boundary conditions to approximate a large system by minimizing boundary effects or to handle Coulomb forces. According to our earlier method [32], a virus and its microenvironment are embedded in a 3-D space where a point $\vec{r}$ is considered a displacement from an original point $\vec{r}^{\,0}$ in the undeformed space. The deformation of space taking any point $\vec{r}^{\,0}$ to the position after deformation $\vec{r}$ and the basis functions are used to introduce a set of order parameters $\vec{\Phi}_{\underline{\ell}}$ via

$$\vec{r} = \sum_{\underline{\ell}} \vec{\Phi}_{\underline{\ell}} U_{\underline{\ell}}\left(\vec{r}^{\,0}\right). \tag{2.1}$$

As the $\vec{\Phi}_{\underline{\ell}}$ change, a point $\vec{r}^{\,0}$ is moved to $\vec{r}$ and thus, to connect the $\vec{\Phi}_{\underline{\ell}}$ to the physical system, the nanostructure embedded in the space is deformed. The $\vec{\Phi}_{\underline{\ell}}$ are interpreted to be a set of vector order parameters that serve as the starting point of an AMA (All-Atom Multiscale Analysis) approach. In what follows, we show how to use (2.1) for a finite set of basis functions and the associated order parameters, and prove that the latter are slowly evolving for appropriately chosen basis functions.

Each atom in the system is moved via the above deformation by evolving the $\vec{\Phi}_{\underline{\ell}}$. However, given a finite truncation of the $\underline{\ell}$-sum, there will be residual displacement of individual atoms above that due to the continuous deformation generated by the order parameters. Denoting the residual of atom i as $\vec{\sigma}_i$, we write its position $\vec{r}_i$ as

$$\vec{r}_i = \sum_{\underline{\ell}} \vec{\Phi}_{\underline{\ell}} U_{\underline{\ell}}\left(\vec{r}^{\,0}\right) + \vec{\sigma}_i. \tag{2.2}$$

The size of $\vec{\sigma}_i$ can be controlled by the choice of basis functions, the number of terms in the $\underline{\ell}$ sum, and the way the $\vec{\Phi}_{\underline{\ell}}$ are defined. Conversely, imposing a permissible size threshold for the residuals allows us to determine the number of basis functions needed to minimize the $\vec{\sigma}_i$, and hence order parameters, to include in the analysis.

A concrete definition of the order parameters is developed as follows. Define the mass-weighted mean-square residual S via

$$S = \sum_{i=1}^{N} m_i \left|\vec{\sigma}_i\right|^2 \Theta_i^c, \tag{2.3}$$

where N is the total number of atoms in the system, m_i is the mass of atom i, and Θ_i^c is one when atom i belongs to the connected subsystem and zero otherwise. With (2.2), this yields

$$S = \sum_{i=1}^{N} m_i \Theta_i^c \left| \vec{r}_i - \sum U_{\underline{\ell}} \left(\vec{r}_i^{\,0} \right) \vec{\Phi}_{\underline{\ell}} \right|^2. \tag{2.4}$$

We assert that the optimal order parameters are those which minimize S, i.e. those containing the maximum amount of information so that the $\vec{\sigma}_i$ are, on the average, the smallest. Thus, we obtain the relationship between $\vec{\Phi}_{\underline{\ell}}$ and $\Gamma_r = \left\{ \vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N \right\}$ via minimizing S with respect to the $\vec{\Phi}_{\underline{\ell}}$ keeping Γ_r constant [22]. This implies

$$\sum_{\underline{\ell}'} B_{\underline{\ell}\,\underline{\ell}'} \vec{\Phi}_{\underline{\ell}'} = \sum_{i=1}^{N} m_i \vec{r}_i U_{\underline{\ell}} \left(\vec{r}_i^{\,0} \right) \Theta_i^c, \tag{2.5}$$

$$B_{\underline{\ell}\,\underline{\ell}'} = \sum_{i=1}^{N} m_i U_{\underline{\ell}} \left(\vec{r}_i^{\,0} \right) U_{\underline{\ell}'} \left(\vec{r}_i^{\,0} \right) \Theta_i^c. \tag{2.6}$$

Orthogonality of the basis functions implies that the B matrix is nearly diagonal. Hence, the order parameters can easily be computed numerically in terms of the atomic positions by solving (2.5).

To proceed, we must be more precise regarding the normalization of the basis functions. For the function $U_{\underline{\ell}}$ with $\underline{\ell} = (0, 0, 0)$, we take $U_{\underline{0}} = 1$. Thus, $B_{\underline{0}\,\underline{0}}$ is the total mass of the atoms in the connected structure. Furthermore, if the B matrix is diagonal, one can show that $\vec{\Phi}_{\underline{0}}$ is the center-of-mass (CM). From earlier studies [20–29], this implies that $\vec{\Phi}_{\underline{0}}$ is slowly varying. If $\varepsilon = m/B_{\underline{0}\,\underline{0}}$ for typical atomic mass m, then the eigenvalues of B are large, i.e. $\mathrm{O}\left(\varepsilon^{-1}\right)$.

A necessary condition for a variable to satisfy a Langevin equation (i.e. to be an order parameter in our multiscale formulation) is that it evolves on a timescale much longer than that of the vibration/collisions of individual atoms. It can be shown that for diagonal B matrix $\mathrm{d}\vec{\Phi}_{\underline{0}}/\mathrm{d}t = \varepsilon \vec{\Pi}_{\underline{0}}$, where $\vec{\Pi}_{\underline{\ell}}$ is the conjugate momentum of $\vec{\Phi}_{\underline{\ell}}$. Taking $\mathrm{d}\vec{\Phi}_{\underline{\ell}}/\mathrm{d}t = \varepsilon \vec{\Pi}_{\underline{\ell}}$, and applying Newton's equations ($\mathrm{d}\vec{\Phi}_{\underline{\ell}}/\mathrm{d}t = -\mathcal{L}\vec{\Phi}_{\underline{\ell}}$ for Liouville operator $\mathcal{L}$), yields

$$\varepsilon \sum_{\underline{\ell}'} \vec{\Pi}_{\underline{\ell}'} B_{\underline{\ell}\,\underline{\ell}'} = \sum_{i=1}^{N} \vec{p}_i U_{\underline{\ell}} \left(\vec{r}_i^{\,0} \right) \Theta_i^c. \tag{2.7}$$

Inclusion of m_i in the above expressions gives the order parameters the character of generalized CM variables.

While the $\vec{\Phi}_{\underline{\ell}}$ qualify as order parameters from the above perspective, they do not suffice as a way to characterize the aqueous microenvironment or the membranous material of the enveloped virus. This follows because the molecules in these two subsystems do not maintain nearest-neighbor connectivity. Thus, a second type of order parameters is required for enveloped virus modeling, as developed in the next section.

3 Order parameter fields for disconnected subsystems

For disconnected systems, one needs an all-atom/continuum multiscale (ACM) approach [31]. The starting point of our ACM theory is a set of field variables that change across the system. These field variables must be related to the atomistic description to achieve a rigorous formulation. The membranous material of an enveloped virus contains a subsystem composed of continuously changing molecules. Application of ACM theory to such systems is illustrated as follows. Let the membranous continuum be comprised of N_t types of molecules labeled $k = 1, 2, \ldots, N_t$. Each type is described by a mass density field variable Ψ_k at spatial point $\vec{R}$:

$$\Psi_k\left(\vec{R}, \Gamma_r\right) = \sum_{i=1}^{N} m_i \delta\left(\vec{R} - \varepsilon_d \vec{r}_i\right) \Theta_i^k, \tag{3.1}$$

where δ is the Dirac delta function centered at $\vec{0}$, ε_d is a smallness parameter, and $\Theta_i^k = 1$ when atom i belongs to a molecule of type k in the disconnected subsystem, and zero otherwise. As these field variables are intended to be coherent in character, $\vec{R}$ is scaled such that it undergoes a displacement of about one unit as several nanometers are traversed. In contrast, the $\vec{r}_i$ undergo a displacement of about one unit as a typical nearest-neighbor interatomic distance for a condensed medium (a few angstroms) is traversed. Thereby, it is natural to scale $\vec{r}_i$ to track the fine-structural details in the system. With this, we let ε_d be the ratio of the typical nearest-neighbor interatomic distance to the size of a nanocomponent (e.g. a viral capsomer). This length scale ratio characterizes the multiscale nature of the enveloped virus. As $\varepsilon_d \ll 1$, it provides a natural expansion parameter for solving the equations of molecular physics, i.e. the Liouville equation for the N-atom probability density. Newton's equations imply

$$\frac{\mathrm{d}\Psi_k}{\mathrm{d}t} = -\mathcal{L}\Psi_k = -\varepsilon_d \vec{\nabla} \vec{G}_k\left(\vec{R}, \Gamma\right) \equiv \varepsilon_d J_k\left(\vec{R}, \Gamma\right) \tag{3.2}$$

$$\vec{G}_k\left(\vec{R}, \Gamma\right) = \sum_{i=1}^{N} \vec{p}_i \delta\left(\vec{R} - \varepsilon_d \vec{r}_i\right) \Theta_i^k, \tag{3.3}$$

where $\vec{G}_k$ is the momentum density of molecules of type k, $\vec{\nabla}$ is the $\vec{R}$-gradient, J_k is the divergence (defined by (3.2)), $\vec{p}_i$ is the momentum of atom i, and $\Gamma = \left\{\Gamma_r, \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N\right\}$ is the set of $6N$ atomistic state variables. For quasi-equilibrium conditions, the average momentum $\left\langle \vec{p}_i \right\rangle$ is small. Thus, the momenta of the atoms in the expression for $\vec{G}_k$ tend to cancel each other. This suggests that $\vec{G}_k$ are of order $\mathrm{O}\left(\varepsilon_d^0\right)$, and thus Ψ_k are slowly evolving, at a rate of $\mathrm{O}\left(\varepsilon_d\right)$.

Order parameter fields like Ψ_k are indexed by $\vec{R}$ which varies continuously across the system. Thus, with Γ_r dependency being understood, we sometimes use the notation $\Psi_k(\vec{R})$ to reflect this parameterization, i.e. to label Ψ_k as the $\vec{R}$-associated order parameter, much like $\vec{r}_i$ is the i-associated position variable (although $\vec{r}_i$ is not an order parameter as it has predominantly 10^{-14} second timescale dynamics). That for each molecule type k there is a Ψ_k for every point $\vec{R}$ in the system suggests there is an uncountable infinity of slow field

variables, $\Psi_k(\vec{R})$. Finally, in order to connect the smallness parameter, ε, of Sect. 2 with that of this section, ε_d, we suggest that ε_d is proportional to ε, and thus take $\varepsilon_d = \varepsilon$ for simplicity.

4 Multiscale integration for enveloped virus modeling

Integration of the multiple types of order parameters (Sects. 2 and 3) needed for enveloped virus modeling is achieved in a self-consistent fashion as follows. We hypothesize that the N-atom probability density ρ for the composite virus/microenvironment system has multiscale character and can thus be rewritten to express its dependency on both the set of atomic positions and momenta Γ and the order parameters. The reduced probability density W is defined as a function of the set of order parameters describing the connected subsystem $\underline{\vec{\Phi}} = \vec{\Phi}_{\underline{\ell}}$ for all $\underline{\ell}$ included, and a functional of the set of order parameter fields $\underline{\Psi}(\vec{R}) = \{\Psi_1(\vec{R}), \ldots, \Psi_{N_t}(\vec{R})\}$ describing the disconnected subsystem. By definition, W takes the form

$$W\left[\underline{\vec{\Phi}}, \underline{\Psi}, t\right] = \int d^{6N}\Gamma^* \prod_{\underline{\ell}} \delta\left(\vec{\Phi}_{\underline{\ell}} - \vec{\Phi}^*_{\underline{\ell}}\right) \prod_{k=1}^{N_t} \Delta\left(\Psi_k - \Psi_k^*\right) \rho\left(\Gamma^*, t\right), \tag{4.1}$$

where Δ is a continuum product of δ-functions for all positions $\vec{R}$, Γ^* is the N-atom state over which integration is taken, and $\vec{\Phi}^*_{\underline{\ell}}$ and Ψ_k^* are the order parameters for state Γ^*. With this and the fact that the N-atom probability density ρ satisfies the Liouville equation $\partial\rho/\partial t = \mathcal{L}\rho$, W is found to satisfy the conservation equation

$$\begin{aligned}\frac{\partial W}{\partial t} = &-\varepsilon \sum_{k=1}^{N_t} \int \frac{\mathrm{d}^3 R}{v_c} \frac{\delta}{\delta \Psi_k\left(\vec{R}\right)} \int \mathrm{d}^{6N}\Gamma^* \prod_{\underline{\ell}} \delta\left(\vec{\Phi}_{\underline{\ell}} - \vec{\Phi}^*_{\underline{\ell}}\right) \\ &\times \prod_{k=1}^{N_t} \Delta\left(\Psi_k - \Psi_k^*\right) J_k^*\left(\vec{R}\right) \rho\left(\Gamma^*, t\right) \\ &-\varepsilon \sum_{\underline{\ell}} \frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}}} \int \mathrm{d}^{6N}\Gamma^* \prod_{\underline{\ell}} \delta\left(\vec{\Phi}_{\underline{\ell}} - \vec{\Phi}^*_{\underline{\ell}}\right) \prod_{k=1}^{N_t} \Delta\left(\Psi_k - \Psi_k^*\right) \vec{\Pi}^*_{\underline{\ell}} \rho\left(\Gamma^*, t\right)\end{aligned} \tag{4.2}$$

where v_c is the minimal volume for which it is reasonable to speak of a field variable, and the superscript * for any variable indicates evaluation at Γ^*.

We hypothesize that to reflect the multiscale character of the system, ρ should be written in the form

$$\rho\left(\Gamma, \underline{\vec{\Phi}}, \underline{\Psi}, t_0, \underline{t}; \varepsilon\right). \tag{4.3}$$

The time variables $t_n = \varepsilon^n t$, $n = 0, 1, \ldots$ are introduced to track processes on timescales $\mathrm{O}\left(\varepsilon^{-n}\right)$ for t_n. The set $\underline{t} = \{t_1, t_2, \ldots\}$ tracks the slow processes of interest to viral dynamics, i.e. much slower than those on the 10^{-14} second scale of atomic vibration/collisions. In contrast, t_0 tracks the fast atomistic processes. The ansatz (4.3) is not a violation of the

number ($6N$) of degrees of freedom, but a recognition that ρ depends on Γ in two ways (i.e. both directly and, via $\underline{\vec{\Phi}}$ and $\underline{\Psi}$, indirectly).

With this and the discrete and field order parameters, the chain rule implies the Liouville equation takes the multiscale form

$$\frac{\partial \rho}{\partial t} = \left(\mathcal{L}_0 + \varepsilon \mathcal{L}_1\right) \rho \tag{4.4}$$

$$\mathcal{L}_0 = -\sum_{i=1}^{N} \left(\frac{\vec{p}_i}{m_i} \frac{\partial}{\partial \vec{r}_i} + \vec{F}_i \frac{\partial}{\partial \vec{p}_i} \right) \tag{4.5}$$

$$\mathcal{L}_1 = \mathcal{L}_\Phi + \mathcal{L}_\Psi \tag{4.6}$$

$$\mathcal{L}_\Phi = -\sum_{\underline{\ell}} \vec{\Pi}_{\underline{\ell}} \frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}}} \tag{4.7}$$

$$\mathcal{L}_\Psi = -\sum_{k=1}^{N_t} \int \frac{\mathrm{d}^3 R}{\upsilon_c} J_k \left(\vec{R}\right) \frac{\delta}{\delta \Psi_k \left(\vec{R}\right)} \tag{4.8}$$

The operator $\mathcal{L}_1$ involves derivatives with respect to $\underline{\vec{\Phi}}$ and functional derivatives with respect to $\underline{\Psi}$ at constant Γ, and conversely for $\mathcal{L}_0$. By mapping the Liouville problem to a higher dimensional descriptive variable space (i.e. $6N$ plus the number of variables in $\underline{\vec{\Phi}}$ and the function space of the order parameter fields $\underline{\Psi}$), our strategy as suggested by our earlier studies [20–31] is to solve the Liouville equation in the higher dimensional representation, and then use the solution to obtain an equation of stochastic $\underline{\vec{\Phi}}$,$\underline{\Psi}$-dynamics.

The development continues with the perturbation expansion

$$\rho = \sum_{n=0}^{\infty} \varepsilon^n \rho_n, \tag{4.9}$$

and examining the multiscale Liouville equation at each order in ε. We hypothesize the lowest order behavior of ρ is slowly varying in time since the phenomena of interest vary on the millisecond or longer, and not the 10^{-14} second time scale. Thus, we assume the lowest order solution ρ_0 is independent of t_0 and, furthermore, is quasi-equilibrium in character.

To $\mathrm{O}\left(\varepsilon^0\right)$, the above assumptions imply $\mathcal{L}_0 \rho_0 = 0$ so that ρ_0 is in the null space of $\mathcal{L}_0$ but is otherwise unknown. We determine ρ_0 by adopting an entropy maximization procedure with canonical constraint of fixed average energy as discussed earlier in the context of nanosystems [24]; this is equivalent to taking the system to be isothermal. With this, we obtain

$$\rho_0 \left[\Gamma; \underline{\vec{\Phi}}, \underline{\Psi}, \underline{t}\right] = \hat{\rho} \left[\Gamma; \underline{\vec{\Phi}}, \underline{\Psi}\right] W_0 \left[\underline{\vec{\Phi}}, \underline{\Psi}, \underline{t}\right] \tag{4.10}$$

$$\hat{\rho} = \frac{e^{-\beta H}}{Q\left[\underline{\vec{\Phi}}, \underline{\Psi}\right]}, \tag{4.11}$$

where β is the inverse temperature, H is the Hamiltonian,

$$H(\Gamma) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V(\Gamma_r), \tag{4.12}$$

for N-atom potential V, and Q is the partition function which is a function of $\vec{\underline{\Phi}}$ and a functional of $\underline{\Psi}$ given by

$$Q\left[\vec{\underline{\Phi}}, \underline{\Psi}\right] = \int \mathrm{d}^{6N}\Gamma^* \prod_{\underline{\ell}} \delta\left(\vec{\Phi}_{\underline{\ell}} - \vec{\Phi}^*_{\underline{\ell}}\right) \prod_{k=1}^{N_t} \Delta\left(\Psi_k - \Psi_k^*\right) e^{-\beta H^*}. \tag{4.13}$$

To O (ε), the multiscale Liouville equation implies

$$\left(\frac{\partial}{\partial t_0} - \mathcal{L}_0\right)\rho_1 = -\frac{\partial \rho_0}{\partial t_1} + \mathcal{L}_1\rho_0. \tag{4.14}$$

This admits the solution

$$\rho_1 = -\int_0^{t_0} \mathrm{d}t_0' e^{\mathcal{L}_0\left(t_0 - t_0'\right)} \left\{\frac{\partial \rho_0}{\partial t_1} - \mathcal{L}_1\rho_0\right\}, \tag{4.15}$$

where the initial first order distribution was taken to be zero as suggested earlier [25,26] to ensure that the system is initially in equilibrium. As a consequence, the final stochastic equation is closed in W.

Inserting (4.6), (4.7), (4.8), and (4.10) in (4.15) yields

$$\begin{aligned}\rho_1 = -t_0\hat{\rho}\frac{\partial W_0}{\partial t_1} - \hat{\rho}\int_0^{t_0} \mathrm{d}t_0' e^{\mathcal{L}_0\left(t_0 - t_0'\right)} \left\{\sum_{\underline{\ell}} \vec{\Pi}_{\underline{\ell}} \left(\frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}}} - \beta\left\langle \vec{f}_{\underline{\ell}}\right\rangle\right)\right. \\ \left. + \sum_{k=1}^{N_t} \int \frac{\mathrm{d}^3 R}{v_c} J_k\left(\vec{R}\right)\left(\frac{\delta}{\delta\Psi_k\left(\vec{R}\right)} - \beta\left\langle h_k\left(\vec{R}\right)\right\rangle\right)\right\} W_0\end{aligned} \tag{4.16}$$

where the thermal-average forces are given by

$$\left\langle \vec{f}_{\underline{\ell}}\right\rangle = -\frac{\partial F}{\partial \vec{\Phi}_{\underline{\ell}}} \tag{4.17}$$

$$\left\langle h_k\left(\vec{R}\right)\right\rangle = -\frac{\partial F}{\partial \Psi_k\left(\vec{R}\right)} \tag{4.18}$$

and F is the free energy related to Q via $Q = e^{-\beta F}$.

Using the Gibbs hypothesis, imposing the condition that ρ_1 be finite as $t_0 \to \infty$, and using the fact that the thermal-averages of J_k and $\vec{\Pi}_{\underline{\ell}}$ are zero (since the weighing factor $\hat{\rho}$ is even in the $\vec{p}_i$, while J_k and $\vec{\Pi}_{\underline{\ell}}$ are odd in them), we find W_0 to be independent of t_1. With

this, (4.16) becomes

$$\rho_1 = -\hat{\rho}\int_0^{t_0} \mathrm{d}t_0' e^{\mathcal{L}_0\left(t_0-t_0'\right)}\left\{\sum_{\underline{\ell}} \vec{\Pi}_{\underline{\ell}}\left(\frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}}} - \beta\left\langle \vec{f}_{\underline{\ell}}\right\rangle\right)\right.$$
$$\left. + \sum_{k=1}^{N_t}\int \frac{\mathrm{d}^3R}{v_c} J_k\left(\vec{R}\right)\left(\frac{\delta}{\delta\Psi_k\left(\vec{R}\right)} - \beta\left\langle h_k\left(\vec{R}\right)\right\rangle\right)\right\} W_0 \tag{4.19}$$

Inserting (4.10) and (4.19) in the conservation equation (4.2) yields

$$\frac{\partial W}{\partial t} = \varepsilon^2 \sum_{\underline{\ell}}\sum_{\underline{\ell}'} \frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}}} \overrightarrow{\vec{D}}^{\Phi}_{\underline{\ell}\underline{\ell}'}\left(\frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}'}} - \beta\left\langle \vec{f}_{\underline{\ell}'}\right\rangle\right) W$$
$$+ \varepsilon^2 \sum_{k=1}^{N_t}\sum_{\underline{\ell}}\int \frac{\mathrm{d}^3R}{v_c}\frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}}} \vec{D}_{\underline{\ell}k}\left(\frac{\delta}{\delta\Psi_k\left(\vec{R}\right)} - \beta\left\langle h_k\left(\vec{R}\right)\right\rangle\right) W$$
$$+ \varepsilon^2 \sum_{k=1}^{N_t}\sum_{\underline{\ell}}\int \frac{\mathrm{d}^3R}{v_c}\frac{\delta}{\delta\Psi_k\left(\vec{R}\right)} \vec{D}_{k\underline{\ell}}\left(\frac{\partial}{\partial \vec{\Phi}_{\underline{\ell}}} - \beta\left\langle \vec{f}_{\underline{\ell}}\right\rangle\right) W$$
$$+ \varepsilon^2 \sum_{k=1}^{N_t}\sum_{k'=1}^{N_t}\int \frac{\mathrm{d}^3R}{v_c}\frac{\mathrm{d}^3R'}{v_c}\frac{\delta}{\delta\Psi_k\left(\vec{R}\right)} D^{\Psi}_{kk'}\left(\frac{\delta}{\delta\Psi_{k'}\left(\vec{R}'\right)} - \beta\left\langle h_{k'}\left(\vec{R}'\right)\right\rangle\right) W \tag{4.20}$$

where $\overrightarrow{\vec{D}}^{\Phi}_{\underline{\ell}\underline{\ell}'}$, $\vec{D}_{\underline{\ell}k}$, $\vec{D}_{k\underline{\ell}}$, and $D^{\Psi}_{kk'}$ are diffusion coefficients defined as

$$\overrightarrow{\vec{D}}^{\Phi}_{\underline{\ell}\underline{\ell}'} = \int_{-\infty}^{0} \mathrm{d}t_0 \vec{\Pi}_{\underline{\ell}} e^{-\mathcal{L}_0 t_0}\vec{\Pi}_{\underline{\ell}'} \tag{4.21}$$

$$\vec{D}_{\underline{\ell}k}\left(\vec{R}\right) = \int_{-\infty}^{0} \mathrm{d}t_0 \vec{\Pi}_{\underline{\ell}} e^{-\mathcal{L}_0 t_0} J_k\left(\vec{R}\right) \tag{4.22}$$

$$\vec{D}_{k\underline{\ell}}\left(\vec{R}\right) = \int_{-\infty}^{0} \mathrm{d}t_0 J_k\left(\vec{R}\right) e^{-\mathcal{L}_0 t_0}\vec{\Pi}_{\underline{\ell}} \tag{4.23}$$

$$D^{\Psi}_{kk'}\left(\vec{R},\vec{R}'\right) = \int_{-\infty}^{0} \mathrm{d}t_0 J_k\left(\vec{R}\right) e^{-\mathcal{L}_0 t_0} J_{k'}\left(\vec{R}'\right) \tag{4.24}$$

As in Ref. [31], there are symmetry rules relating the cross-diffusion coefficients. The set of Langevin equations equivalent to this generalized Smoluchowski equation (4.20), provides

a practical approach to the simulation of enveloped virus systems as outlined in the next section.

5 Multiscale computations and the NanoX platform

The all-atom multiscale approach of the previous section can be implemented as a platform for the simulation of nanosystems. In the flow chart of Fig. 3 it is seen how order parameters are co-evolved with the statistics of the atomic fluctuations in our NanoX platform. Interscale coupling as in Fig. 2 is manifested through the thermal-average forces and diffusion coefficients constructed via short-time ensemble/molecular dynamics computations in the indicated modules. We have implemented this Fig. 3 workflow, creating the NanoX simulator.

At this writing, NanoX is built on the order parameters of Sect. 2. For it to be practical, the diffusion coefficients and thermal-average forces must not be excessively demanding on CPU time. In Fig. 4a we show an example of an order parameter trajectory for the STMV (nonenveloped) virus in vacuum. This system has been studied elsewhere using classic MD [33]. Our simulation began from the crystal cryo-structure which was equilibrated for 1 ns at various temperatures. The order parameter trajectory (here we choose the z component of the 001 mode) is shown for the first 20 ps and is seen to change slowly except at the beginning where the system escapes from its potential energy-minimized unphysical structure. The velocity of the order parameters are obtained by differentiating equation (2.1). After thermalization, the order parameters hardly change for several picoseconds (Fig. 4b) whereas the velocity (not shown) appears to fluctuate about zero. These fluctuations are not highly correlated in time as shown by the rapid decay of the velocity auto-correlation function in Fig. 4c. This demonstrates that only short MD runs are required to calculate the diffusion coefficients via (4.21). The behavior of the auto-correlation function was studied at various temperatures to check consistency with the notion that the diffusion coefficient decreases with temperature. The auto-correlations were normalized by dividing by their starting values which are provided in Table 1. If the velocity autocorrelation function does not approach zero exponentially, then the friction-dominated Smoluchowski limit of Sect. 4 is not valid and the order parameter velocities should be added to the list of order parameters, i.e. a Fokker-Planck limit is appropriate [24,25]. This does not appear to be the case in this system. To add artificial intelligence, NanoX will automatically determine if a Smoluchowski or

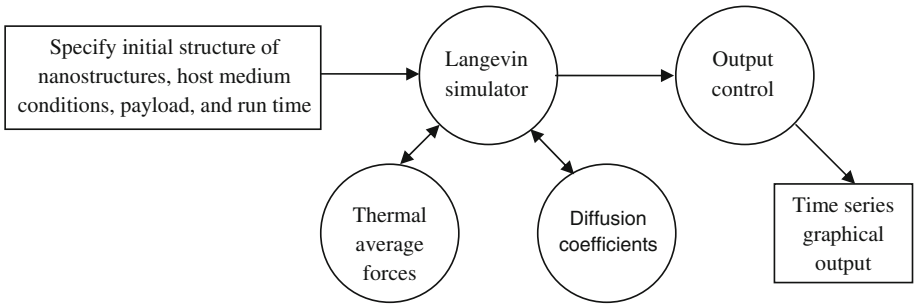

Fig. 3 Schematic NanoX workflow indicating that thermal-average forces and diffusion coefficients are computed "on the fly" since they co-evolve with the order parameters describing overall features of a bionanosystem

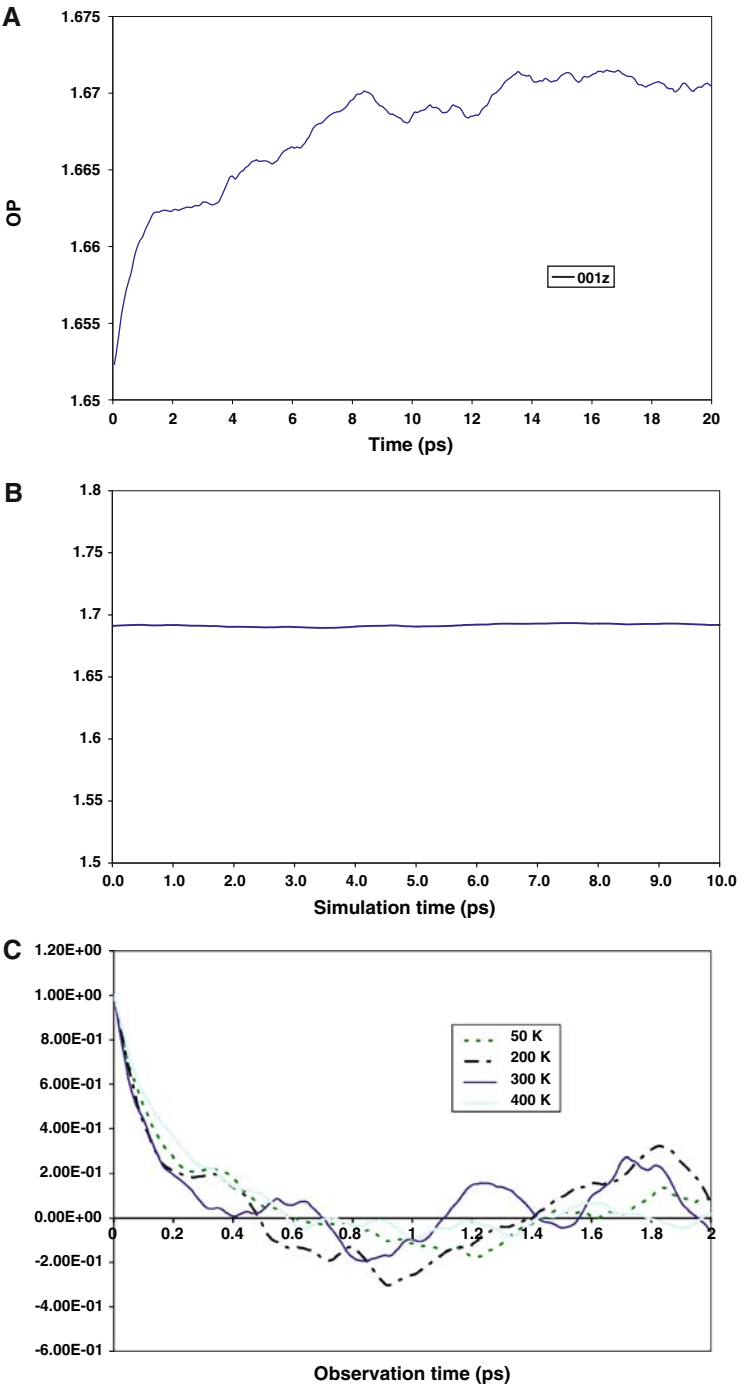

Fig. 4 **A** Order parameter 001z as a function of time. **B** Order parameter 001z as a function of time after equilibration (300 K). **C** Normalized order parameter velocity auto-correlation function for component 001z as a function of observation time for different temperatures

Table 1 Initial order parameter velocity auto-correlation function for component 001z at different temperatures

Temperature (K)	Initial order parameter velocity auto-correlation function
50	1.49E-09
200	7.08E-09
300	9.33E-09
400	2.08E-08

Fokker-Planck approach is required. In a similar way, NanoX will include additional order parameters if necessary, a capability enabled by our automated order parameter generation scheme of Sect. 2. This feature gives our approach a degree of self-consistency; for example, if the order parameter velocity autocorrelation functions are poorly behaved, then additional order parameters are needed to complete the theory.

The thermal-average forces are calculated using two methods. In the first method, the integral in (4.17) is calculated by a Monte Carlo approach by sampling a fixed order parameter ensemble. Random sampling generates unfavourable configurations and yields high energy configurations with negligible Boltzmann weight. Rather, we use a sampling method whereby only order parameter components orthogonal to the fixed ones are varied. We use this high frequency order parameter ensemble to calculate order parameter velocities and forces. In the second method, a short MD run is performed and Gibbs hypothesis is used to calculate the averages. This approximation is valid since, as discussed above, the order parameters vary slowly in time and the only limitation is the sampling of configuration space by the MD trajectory. We hypothesize that a necessary condition for self-consistency is that the order parameter velocity auto-correlation time is much less than the characteristic time of evolution of the order parameters as driven by the thermal-average forces. Once the thermal-average forces and diffusion coefficients are calculated, the feedback loop in Fig. 3 underlying our multiscale approach is accounted for via the flowchart. The system evolution is carried out via a sequence starting with short MD runs followed by a Langevin evolution timestep for the order parameters. This requires the thermal-average forces and diffusion constants. After the Langevin evolution timestep, the atomic configurations are regenerated by using equation (2.2). Constrained energy minimization and annealing of high energy contacts generated by this "fine graining" is needed to ensure sensible atomic scale configurations. This completes the evolution cycle of Fig. 3. We have implemented this workflow for the STMV virus in our NanoX simulator.

The calculation of the thermal-average forces, diffusion constants, and the energy minimization are all CPU intensive requiring optimizations at different stages. Our simulations were run on the 768 node IBM JS21 cluster at Indiana University using NAMD and we built our simulation modules to effectively utilize existing NAMD resources without introducing unnecessary overhead. Based on preliminary estimates, the thermal-average force module takes about 30 minutes to sample and analyze 2000 configurations, a 10 ps NAMD run for STMV in vacuum takes 15 minutes on 64 processors. Further reductions in CPU time will be possible with greater integration into the NAMD source code.

6 Applications and conclusions

Because of the impracticality of straightforward all-atom models, phenomenological approaches have mainly been used for systems biology. These approaches require recalibration

with each new application, a main impediment to progress in computer simulations of biological systems. This difficulty is compounded by the complexity of the systems of interest, leading to ambiguities created by over-calibration, i.e. arriving at the right answer for the wrong reason. To transcend this impediment, we derived principles of microbiology from the laws of molecular physics as outlined for bionanosystems as in Sects. 2 to 4.

To demonstrate that this is possible with extreme complexity, we focused on enveloped viruses. We integrated several types of order parameters into a self-consistent framework and showed that this can be accomplished via rigorous multiscale analysis. This approach overcomes computational difficulties associated with the great number of atoms involved, and accounts for the existence of subsystems with distinct types of physics. Enveloped viruses with their self-assembly, maturation scenario, structural transitions, and richness of phenomena appear to be a prime candidate.

Many systems can be modeled if a self-consistent ACM theory was developed. Examples include

- local sites such as receptors or pores in a cell membrane;
- an enveloped virus involving a composite of a lipid membrane and protein/DNA or RNA structures;
- a nanocapsule with its surrounding cloud of leaking drug molecules;
- an intrabacterial energy-storage granule; and
- a nanodroplet with embedded macromolecule(s).

The fundamental science of these systems and their potential importance for medicine and nanotechnology make the development of ACM approaches of great interest [31]. For example, an ACM approach could provide the basis of a computer-aided strategy for the design of antiviral vaccines or nanocapsules for the delivery of drugs, genes, or siRNA to diseased cells.

Open questions about enveloped viruses include the following.

- What is the structure of the genome-protein core complex deep within the phospholipid subsystem, i.e. is it well organized or highly fluctuating?
- Does the interaction of the phospholipid subsystem with the outer protein net, traversing protein or the genome-protein core, induce liquid-crystal order in the phospholipid subsystem?
- What factors restrict the size, structure, and stability of the ghost particles that are devoid of the genome-protein core?
- What are the ranges of pH, salinity, concentrations of selected ions, and other conditions in the microenvironment that favor a given structure of the protein outer net (see Fig. 1)?
- What is the effect of an applied electric field, or stress applied through an AFM tip, on viral structure and stability?
- Can the virus be grown around, or injected with, a magnetic or fluorescent nanoparticle probe?
- Does the structural transition in Dengue's outer protein net (Fig. 1) involve bond cleavage or formation?
- Can chemical labels pass through the outer protein net and selectively bond to the genome-protein inner core and provide structural information on the core structure [24]?

An integrated multiple order parameter model to address these questions is suggested in Fig. 5.

We propose a model composed of four order parameters accounting for each of the subsystems of Fig.5. Starting from the aqueous microenvironment, we proceed inward to the

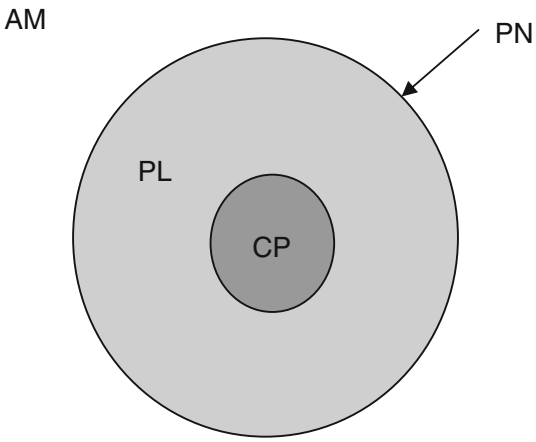

Fig. 5 Schematic four subsystem model: AM (aqueous microenvironment), PN (outer protein net), PL (phospholipid zone), and CP (inner genome-protein core particle)

genome-protein core. The aqueous microenvironment is described by a set of order parameter fields Ψ_q^{aq} specifying the mass-weighted position-orientation density for the water molecules (as in Sect. 3). The outer protein net is considered to be connected over the time period of the structural transition. Thus, it is described via a set of structural order parameters $\vec{\Phi}_{\underline{\ell}}^{PN}$ from the set of order parameters introduced in Sect. 2 that specify the CM, orientation, and details of the conformation of the outer protein net. The membranous zone contains at least two major components (e.g. phospholipids and Glycoproteins) denoted by A and B here. These are characterized by the order parameter fields Ψ_A and Ψ_B giving the spatial distribution of their CM. The genome-protein core is, for simplicity here, assumed to consist of one connected object described by a set of order parameters, as in Sect. 2. More complex models would also include multiple parts of the core, each of which is connected internally, which require their individual set of order parameters and, therefore, could be suited to study core self-assembly/maturation within the overall assembled enveloped system. Proteins bridging the outer protein net and the core could also be accounted for via such structural order parameters. Finally, order parameter fields could be used to track the exchange of small molecules between the microenvironment and the various inner subsystems.In the integrated model, one solves the Langevin equations as outlined in Sect. 5, but for all the order parameters. All order parameters in the model are coupled in two ways. The thermal-average force for any one order parameter depends on others, thereby accounting for a variety of thermodynamic interactions. Furthermore, the diffusion coefficients for the order parameters provide cross-frictional effects among them. Those associated with the field variables introduce nonlocal effects so that the thermal-average force at $\vec{R}$ affects order parameter fields at $\vec{R}'$.

While the solution of the Langevin equations presents no major computational challenges, the construction of thermal-average forces and diffusion coefficients does. However, recent results of Sect. 5 suggest that the correlation functions for our order parameters have a characteristic time of around a picosecond. Thus, only short MD simulations are needed to estimate them.

A simple system for testing the multiscale model of Sect. 4 is to use the structural transitions in the outer protein net of Dengue virus (Fig. 1). In particular, Dengue "ghosts" consist only of an outer protein net and inner phospholipid material, i.e. they are devoid of the inner protein-genomic core assembly. Thus, the model of a ghost consists of the protein-net, as described

via the set of discrete structural parameters introduced in Sect. 2, while the surrounding aqueous medium and inner phospholipid subsystems can be described via order parameter fields. With this, we conclude that with the integration of rigorous multiscale analysis and supercomputing, complex bionanosystems can be modeled, principles of microbiology can be derived, and practical benefits for nanotechnology and biomedicine can be achieved.

Acknowledgements This project was supported in part by the National Institute of Health (NIBIB), the National Science Foundation (CCRC Program), and Indiana University's college of Arts and Sciences through the Center for Cell and Virus Theory.

References

1. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., Schulten, K.: Scalable molecular dynamics with NAMD. J. Comput. Chem. **26**, 1781–1802 (2005)
2. Sanbonmatsu, K.Y., Tung, C.S.: High performance computing in biology: multimillion atom simulations of nanoscale systems. J. Struct. Biol. **157**, 470–480 (2007)
3. Stewart, G.T.: Liquid crystals in biology. I. Historical, biological and medical aspects. Liquid. Cryst. **30**, 541–557 (2003)
4. Zhang, Y., Kostyuchenko, V.A., Rossman, M.G.: Structural analysis of viral nucleocapsids by subtraction of partial projections. J. Struct. Biol. **157**, 356–364 (2007)
5. Zhang, Y., Zhang, W., Ogata, S., Clements, D., Strauss, J.H., Baker, T.S., Kuhn, R.J., Rossmann, M.G.: Conformational changes of the flavivirus E glycoprotein. Structure **12**, 1607–1618 (2004)
6. Zhang, Y., Corver, J., Chipman, P.R., Zhang, W., Pletnev, S.V., Sedlak, D., Baker, T.S., Strauss, J.H., Kuhn, R.J., Rossman, M.G.: Structures of immature flavivirus particles. EMBO J. **22**, 2604–2613 (2003)
7. Klasse, P.J., Bron, R., Marsh, M.: Mechanisms of enveloped virus entry into animal cells. Adv. Drug. Deliv. Rev. **34**, 65–91 (1998)
8. Chandrasekhar, S.: Stochastic problems in physics and astronomy. Rev. Mod. Phys. **15**, 1–89 (1943)
9. Bose, S., Ortoleva, P.: Reacting hard sphere dynamics: Liouville equation for condensed media. J. Chem. Phys. **70**, 3041–3056 (1979)
10. Bose, S., Ortoleva, P.: A hard sphere model of chemical reaction in condensed media. Phys. Lett. A **69**, 367–369 (1979)
11. Bose, S., Bose, S., Ortoleva, P.: Dynamic Padé approximants for chemical center waves. J. Chem. Phys. **72**, 4258–4263 (1980)
12. Bose, S., Medina-Noyola, M., Ortoleva, P.: Third body effects on reactions in liquids. J. Chem. Phys. **75**, 1762–1771 (1981)
13. Deutch, J.M., Oppenheim, I.: The concept of Brownian motion in modern statistical mechanics. Faraday Discuss. Chem. Soc. **83**, 1–20 (1987)
14. Shea, J.E., Oppenheim, I.: Fokker-Planck equation and Langevin equation for one Brownian particle in a nonequilibrium bath. J. Phys. Chem. **100**, 19035–19042 (1996)
15. Shea, J.E., Oppenheim, I.: Fokker-Planck equation and non-linear hydrodynamic equations of a system of several Brownian particles in a non-equilibrium bath. Phys. A **247**, 417–443 (1997)
16. Peters, M.H.: Fokker-Planck equation and the grand molecular friction tensor for combined translational and rotational motions of structured Brownian particles near structures surface. J. Chem. Phys. **110**, 528–538 (1998)
17. Peters, M.H.: Fokker-Planck equation, molecular friction, and molecular dynamics for Brownian particle transport near external solid surfaces. J. Stat. Phys. **94**, 557–586 (1999)
18. Coffey, W.T., Kalmykov, Y.P., Waldron, J.T.: The Langevin Equation with Applications to Stochastic Problems in Physics Chemistry and Electrical Engineering. World Scientific Publishing Co, River Edge (2004)
19. Ortoleva, P.: Nanoparticle dynamics: a multiscale analysis of the Liouville equation. J. Phys. Chem. **109**, 21258–21266 (2005)
20. Miao, Y., Ortoleva, P.: All-atom multiscaling and new ensembles for dynamical nanoparticles. J. Chem. Phys. **125**, 044901 (2006)
21. Miao, Y., Ortoleva, P.: Viral structural transitions: an all-atom multiscale theory. J. Chem. Phys. **125**, 214901 (2006)
22. Shreif, Z., Ortoleva, P.: Curvilinear all-atom multiscale (CAM) theory of macromolecular dynamics. J. Stat. Phys. **130**, 669–685 (2008)

23. Miao, Y., Ortoleva, P.: Molecular dynamics/OP eXtrapolation (MD/OPX) for bionanosystem simulations. J. Comput. Chem. (2008). doi:10.1002/jcc.21071
24. Pankavich, S., Miao, Y., Ortoleva, J, Shreif, Z., Ortoleva, P.: Stochastic dynamics of bionanosystems: multiscale analysis and specialized ensembles. J. Chem. Phys. **128**, 234908 (2008)
25. Pankavich, S., Shreif, Z., Ortoleva, P.: Multiscaling for classical nanosystems: derivation of Smoluchowski and Fokker-Planck equations. Phys. A **387**, 4053–4069 (2008)
26. Shreif, Z., Ortoleva, P.: Multiscale derivation of an augmented Smoluchowski. Phys. A (2008, accepted)
27. Shreif, Z., Ortoleva, P.: Computer-aided design of nanocapsules for therapeutic delivery. Comput. Math. Methods Med. (2008, to appear)
28. Pankavich, S., Shreif, Z., Miao, Y., Ortoleva, P.: Self-assembly of nanocomponents into composite structures: derivation and simulation of Langevin equation. J. Chem. Phys. (2008, accepted)
29. Pankavich, S., Ortoleva, P.: Self-assembly of nanocomponents into composite structures: multiscale derivation of stochastic chemical kinetic models. ACS Nano (2008, in preparation)
30. Pankavich, S., Ortoleva, P.: Multiscaling for systems with a broad continuum of characteristic lengths and times: structural transitions in nanocomposites. (2008, in preparation)
31. Shreif, Z.,Ortoleva, P.: All-atom/continuum multiscale theory: application to nanocapsule therapeutic delivery. Multiscale Model. Simul. (2008, submitted)
32. Jaqaman, K., Ortoleva, P.: New space warping method for the simulation of large-scale macromolecular conformational changes. J. Comput. Chem. **23**, 484–491 (2002)
33. Freddolino, P.L., Arkhipov, A.S., Larson, S.B., McPherson, A., Schulten, K.: Molecular dynamics simulations of the complete satellite tobacco mosaic virus. Structure **14**, 437–449 (2006)