Differential geometry, topological invariant and machine learning approaches to virus dynamics

Yang Wang & Guowei Wei

Department of Mathematics, Michigan State University

Yiying Tong

Computer Science & Engineering, Michigan State University

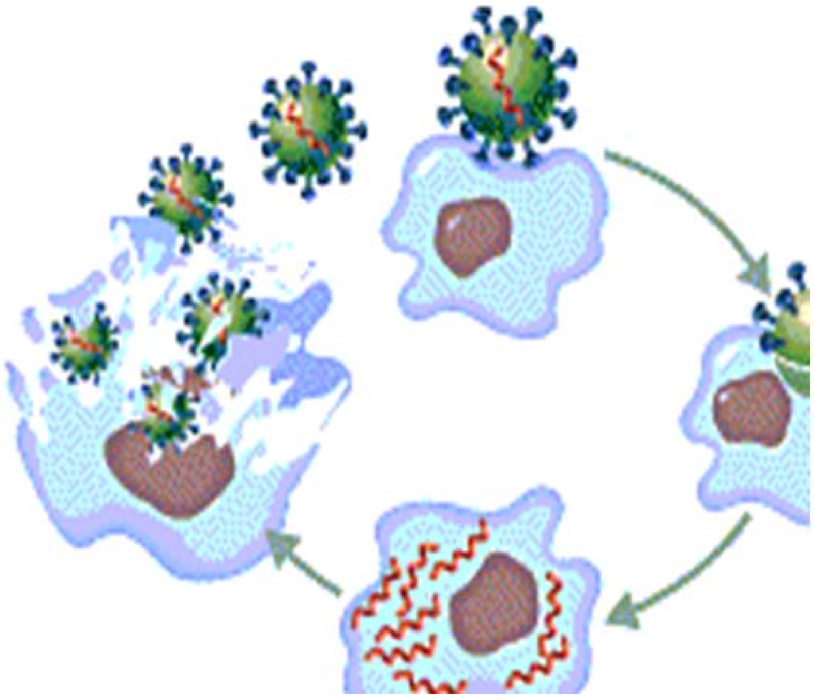
Haomin Zhou

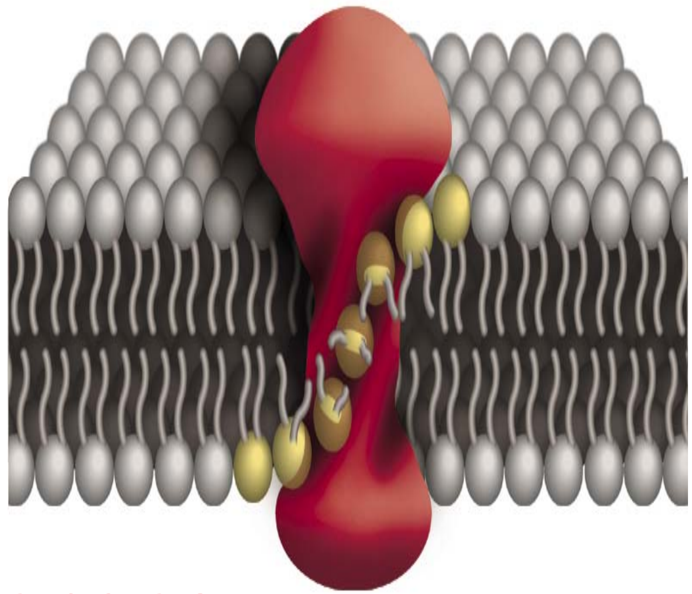
School of Mathematics, Georgia Institute of Technology

http://virology.wisc.edu/virusworld/

Differential geometry, topological invariant and machine learning approaches to

- Understand molecular mechanism of virus life circles
- Develop visual-analytic methods for virus infection prevention
- Extract biological functions and properties from dynamic data





Challenges

High dimension: ~ 10 million dimensions

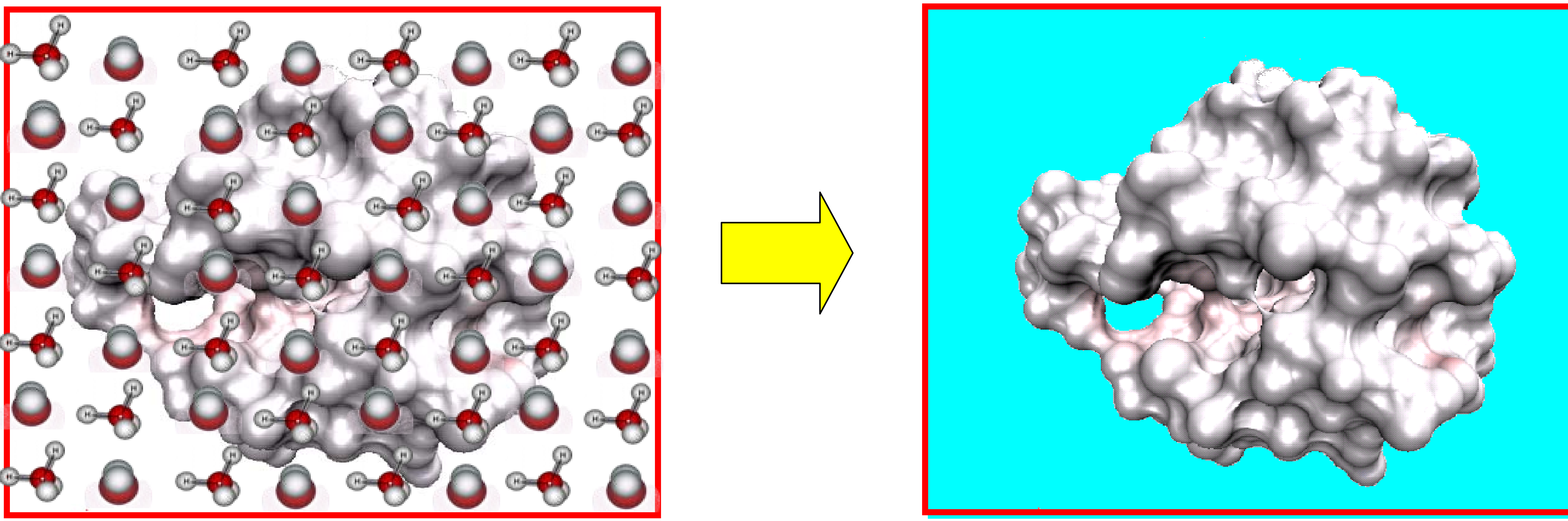
Massive data sets: ~ 10^{18} data points

No viable physical / Mathematical models

MICHIGAN STATE UNIVERSITY

Dimension reduction by multiscale analysis

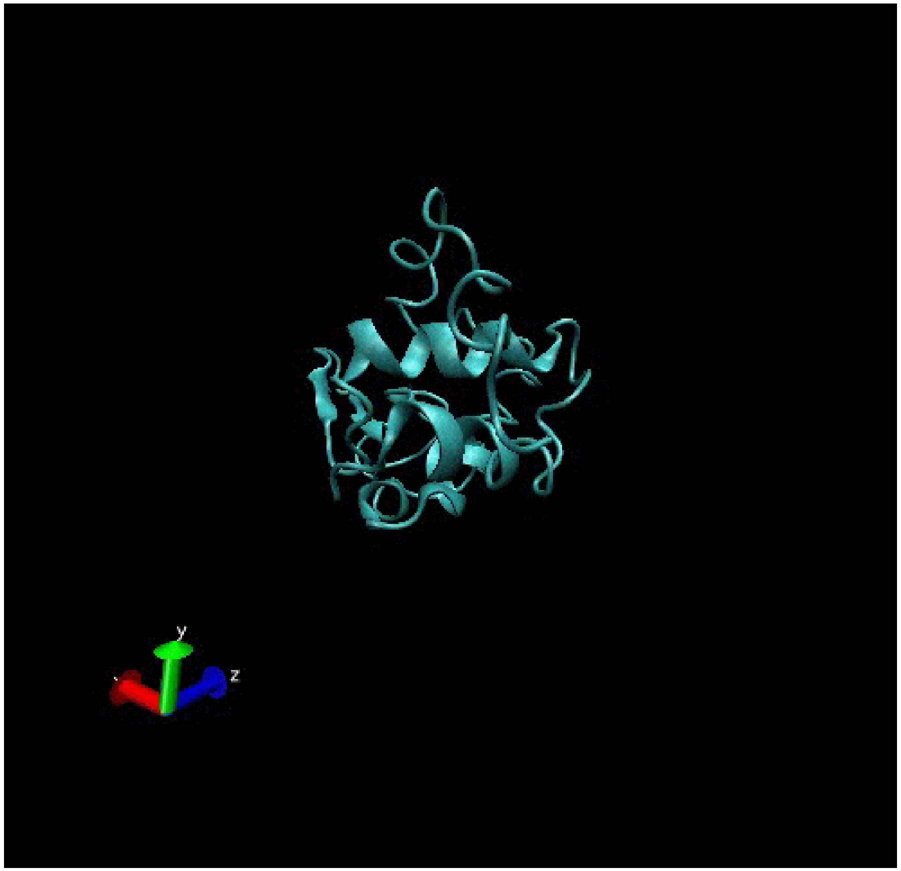
- Solvent is described by continuum models
- Viruses are described by discrete models

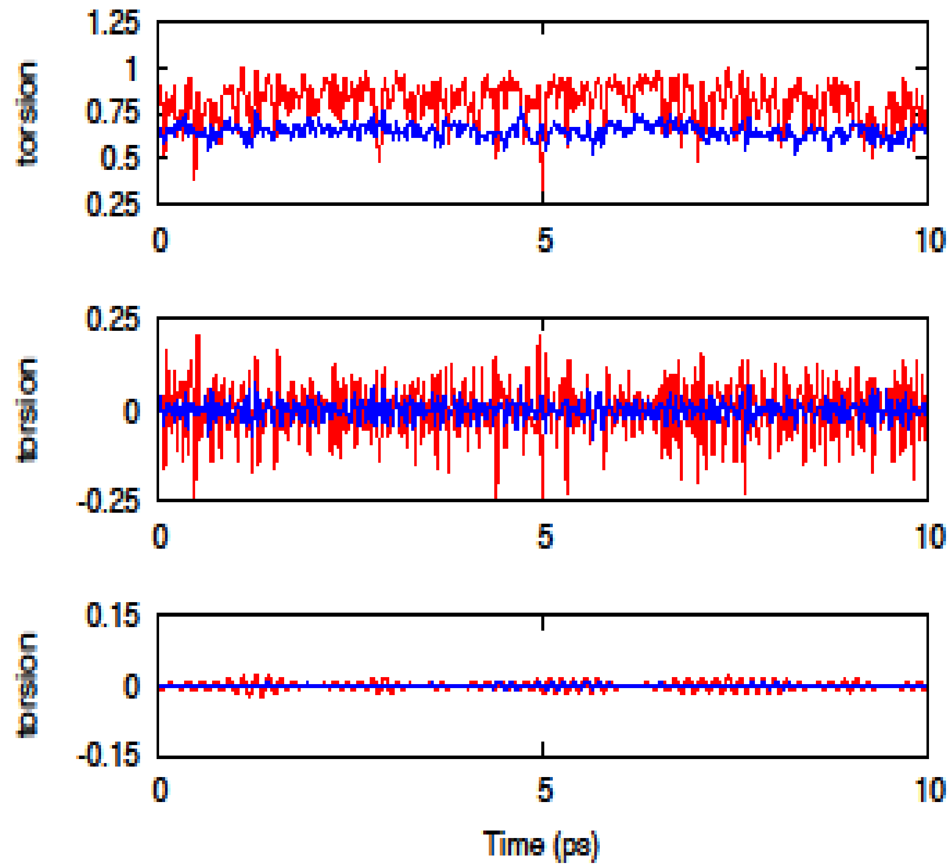


The interface between the discrete and the continuum is described by differential geometry theory of surfaces

(Reduce the dimension by about one order)

Coarse-grained dynamic model based on persistently stable manifolds characterized by the time series of Frenet – Serret frames, torsion angles and curvatures





Machine learning approach to further reduce the dimension by 1 to 3 orders

(Tong, Wang, Wei, Zhou, 2010)

Differential geometry based multiscale free energy functional for excessively large data size reduction of virus systems

$$\min G = \min \iiint \{Geometric + Electro + Fluid + MM\} dxdzdt$$

$$G_{Geometric} = \gamma |\nabla S| + Sp + (1-S)\rho_s u$$

$$G_{Electro} = S\left[\rho_m \phi - \frac{\varepsilon_m}{2}|\nabla \phi|^2\right] + (1-S)\left[-\frac{\varepsilon_s}{2}|\nabla \phi|^2 - k_B T \sum c_j \left(e^{-q_j \phi / K_B T} - 1\right)\right]$$

$$G_{Fluid} = -(1-S)\left[\rho_s \frac{v^2}{2} - p + \frac{\mu}{8}\int^t \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right)^2 dt'\right]$$

$$G_{MM} = -S\sum\left[\rho_j \frac{\dot{z}_j^2}{2} - U(z)\right]$$

(Wei, J Math Biol 2010)

Generalized Navier-Stokes Equation for fluid flow

$$\rho_s\left(\frac{\partial v}{\partial t}+v\cdot\nabla v\right)=-\nabla p+\frac{1}{1-S}\nabla\bullet(1-S)T+F$$

$$F=\frac{S}{1-S}\left(-\nabla p-\frac{1-S}{S}\nabla(\rho_s u)+\frac{\rho_m}{S}\nabla(S\phi)\right)$$

$$\nabla\cdot v=0$$

Generalized Newton equation for molecular dynamics

$$\rho_j\frac{d^2 z_j}{dt^2}=f_{SSI}^{j}+f_{RF}^{j}+f_{PI}^{j}$$

$$f_{SSI}^{j}=-\frac{1-S}{S}\nabla_j(\rho_s u)$$

$$f_{RF}^{j}=\frac{\rho_m}{S}\nabla_j(S\phi)$$

$$f_{PI}^{j}=-\nabla_j U(z)$$

MICHIGAN STATE UNIVERSITY

Generalized Poisson-Boltzmann Equation for electrostatics

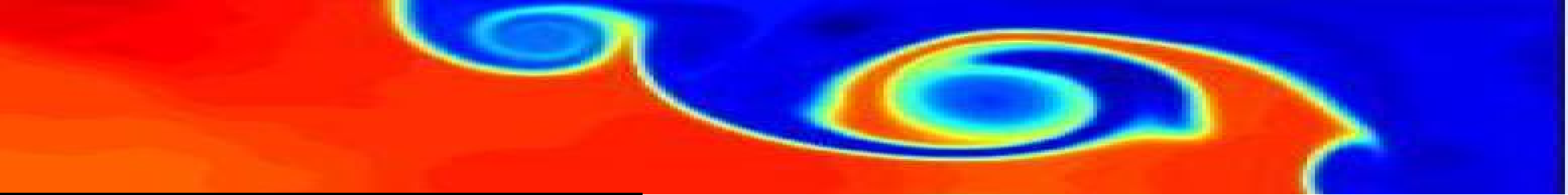
$$-\nabla \bullet \varepsilon(S)\nabla\phi = S\rho_m + (1-S)\sum q_j c_j e^{-q_j\phi/k_BT}$$

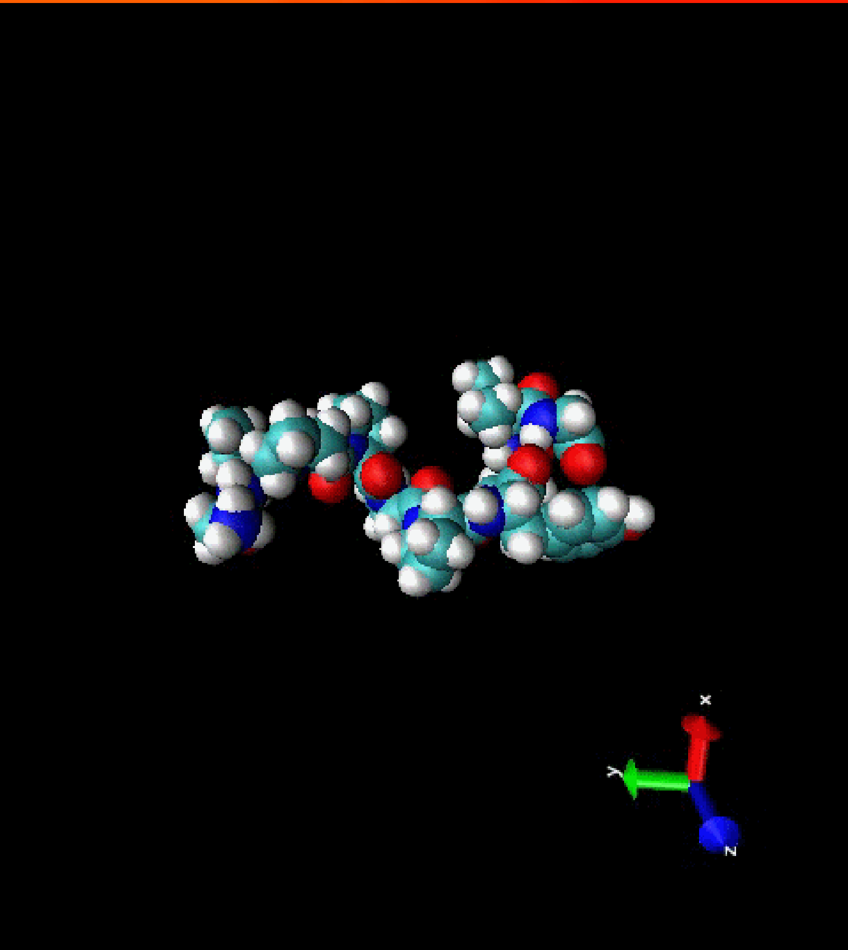
$$\varepsilon(S) = S\varepsilon_m + (1-S)\varepsilon_s$$

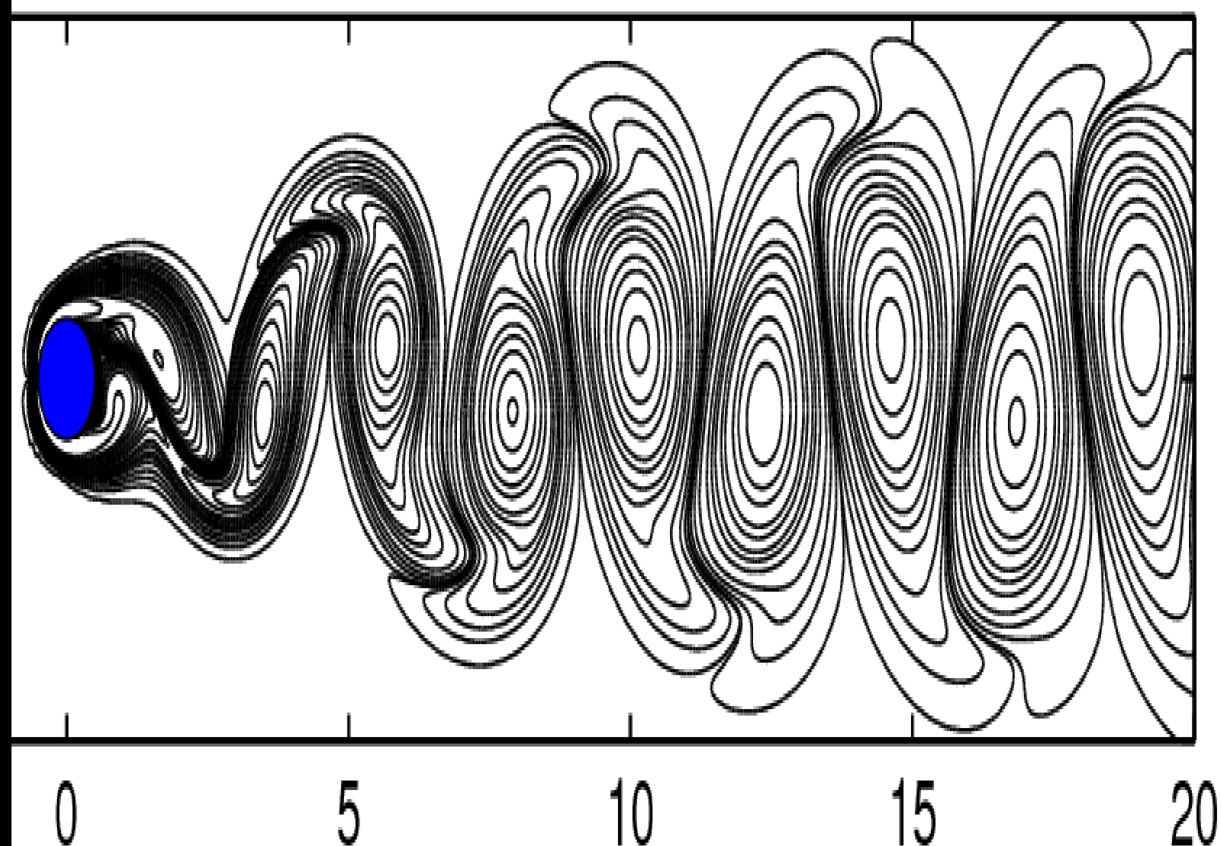
Generalized Laplace-Beltrami Equation for surface dynamics

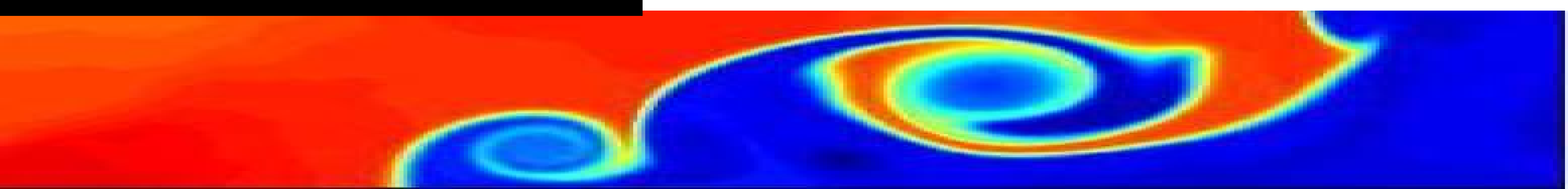
$$\frac{\partial S}{\partial t} = |\nabla S| \left\{ \begin{array}{l} \nabla \bullet \frac{\gamma\nabla S}{|\nabla S|} - p + \rho_s u - \rho_m\phi + \frac{\varepsilon_m}{2}|\nabla\phi|^2 \\ -\frac{\varepsilon_s}{2}|\nabla\phi|^2 - k_BT\sum c_j\left(e^{-q_j\phi/K_BT} - 1\right) \\ -\left[\rho_s\frac{v^2}{2} - p + \frac{\mu}{8}\int^t\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right)^2 dt'\right] \\ +\sum\left[\rho_j\frac{\dot{z}_j^2}{2} - U(z)\right] \end{array} \right\}$$

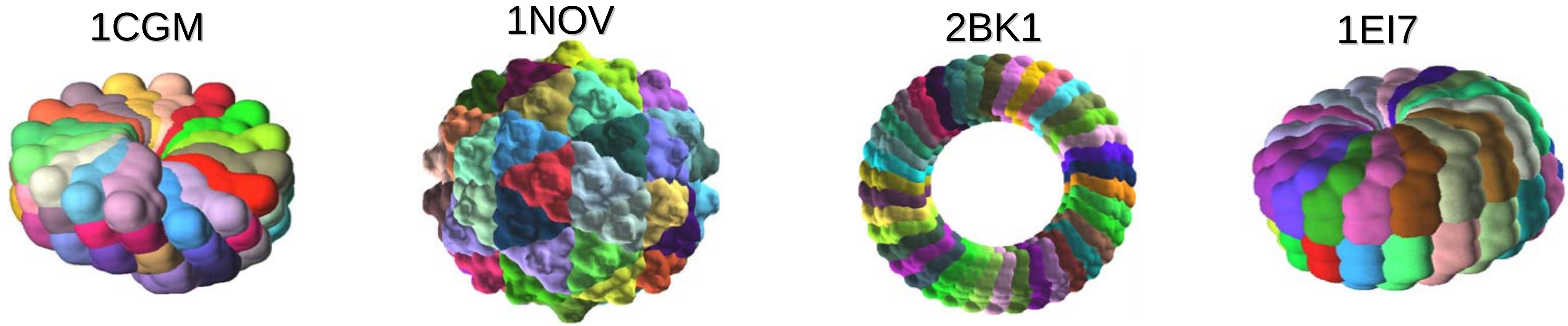
MICHIGAN STATE UNIVERSITY



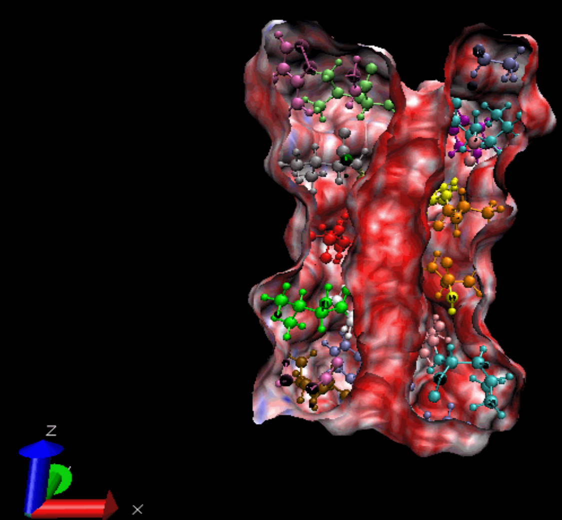


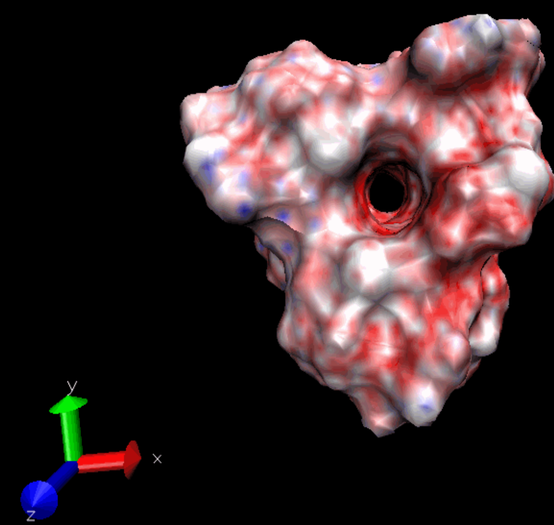

0
5
10
15
20

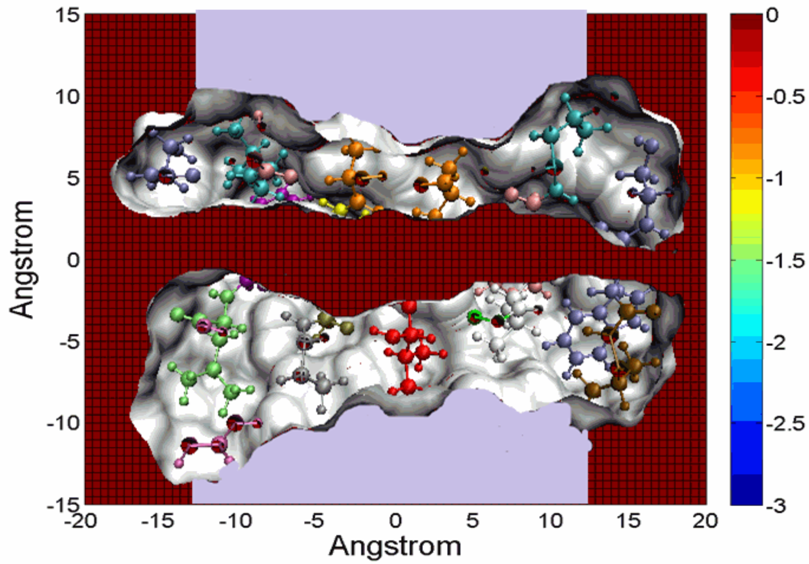


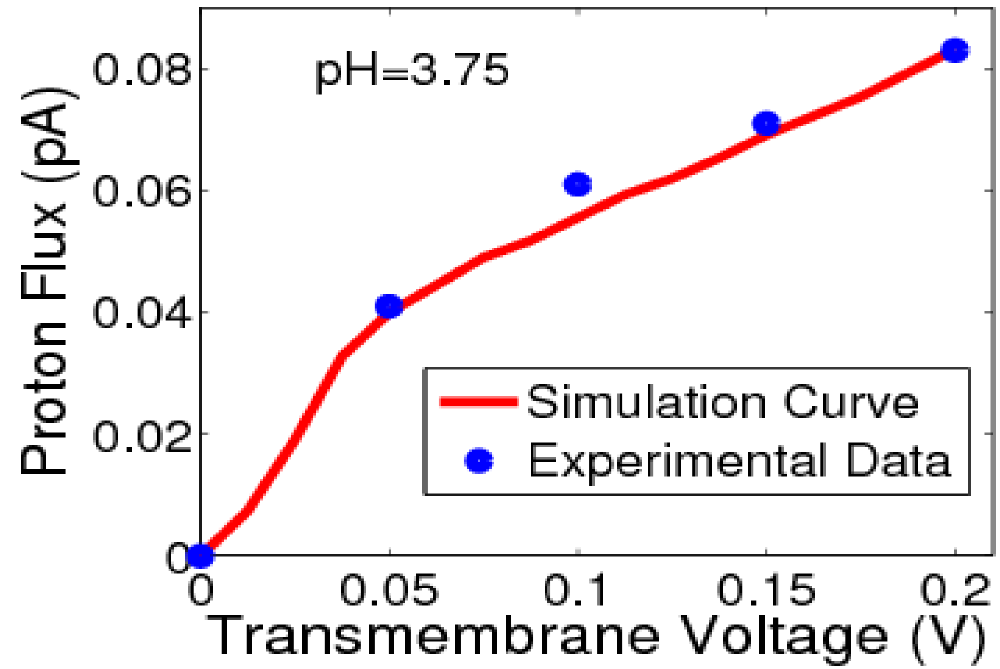


Virus morphology and virus ion channel



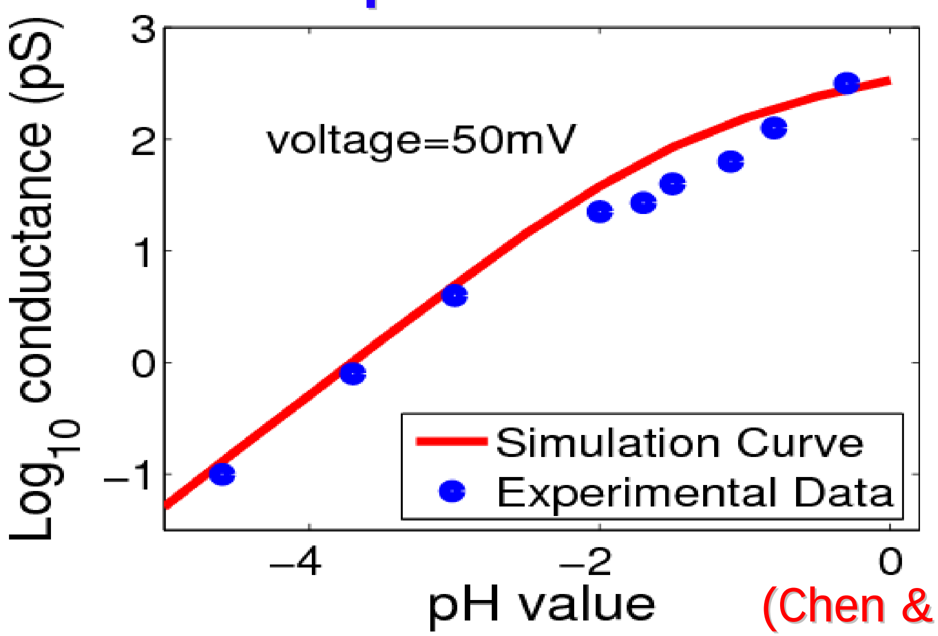




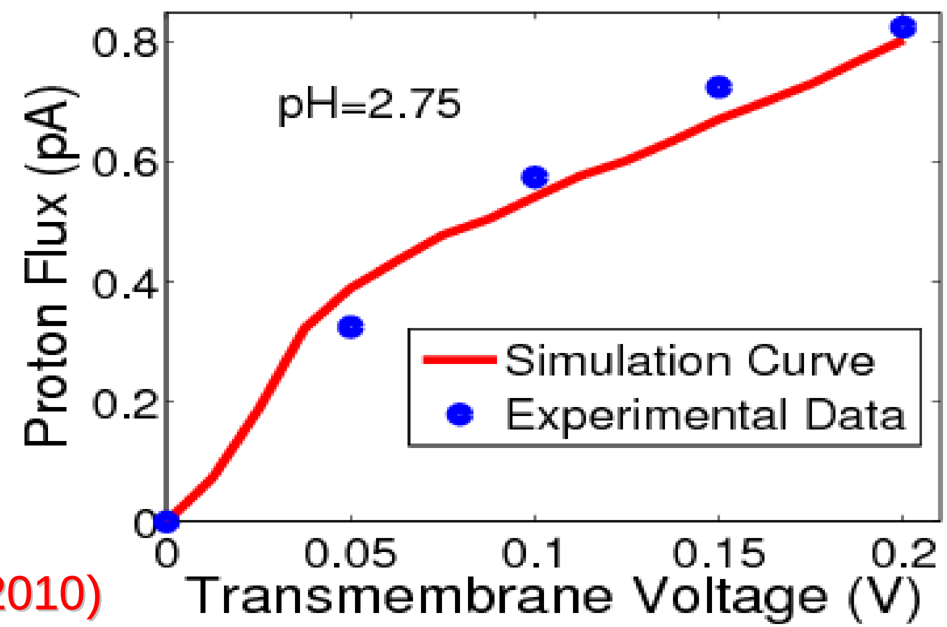


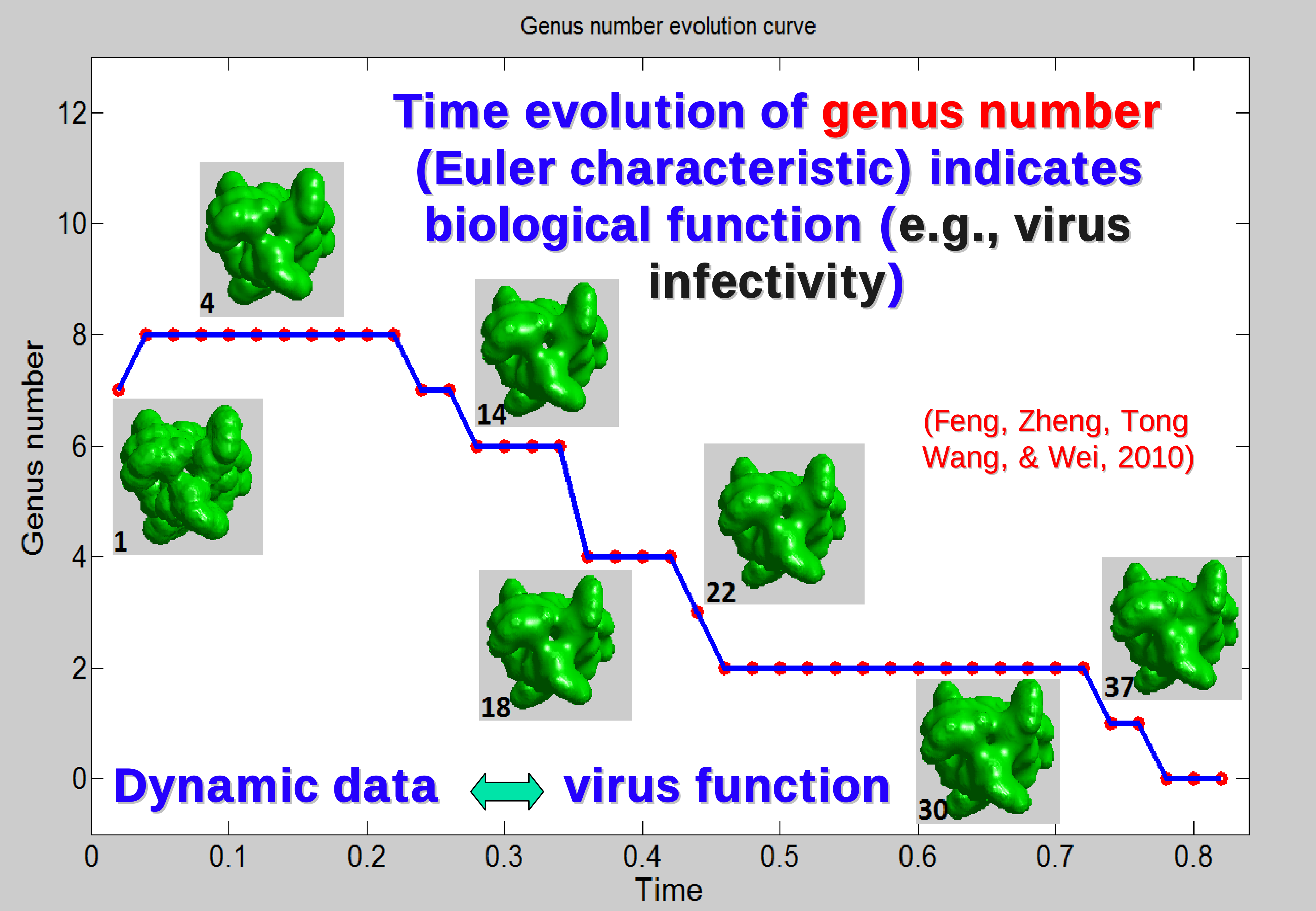
Proton transport of Gramicidin A

(Expl: Eisenman et al., 1980)



(Chen & Wei, 2010)




Genus number evolution curve
Time evolution of genus number (Euler characteristic) indicates biological function (e.g., virus infectivity)
(Feng, Zheng, Tong Wang, & Wei, 2010)
Dynamic data
virus function
Genus number
Time
1
4
14
18
22
30
37
0
2
4
6
8
10
12
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

1. Qiong Zheng and Guo-Wei Wei, Implicit Poisson Nernst-Planck model for ion channels. (2010)
2. Yang Wang, Guo-Wei Wei and Siyang Yang, Model decomposing evolution equations, (38 pages) (2010).
3. Zhan Chen and G. W. Wei, Differential geometry based salvation model III: Quantum formulation (35 pages). (2010).
4. Kelin Xia, Meng Zhan and Guo-Wei Wei,The matched interface and boundary (MIB) method for multi-domain elliptic interface problems. (2010)
5. Yang Wang, Guo-Wei Wei and Siyang Yang, Empirical Model Decomposition using Local Spectral Evolution Kernels, (38 pages) (2010).
6. Qiong Zheng, DuanChen and Guo-Wei Wei, Second-order convergent Poisson Nernst-Planck solver for ion channels. (2010)
7. Duan Chen and Guo-Wei Wei, Quantum dynamics in continuum model for proton channel transport. (2010)
8. Sigal Gottlieb, Guo-Wei Wei, and Shan Zhao, A unified discontinuous Galerkin framework for timeintegration (46 pages), (2010)
9. Zhan Chen, Nathan Baker and G.W. Wei, Differential geometry based solvation model II: Lagrangian formulation (53 pages) submitted, (2010)
10. Weihua Geng, and G. W. Wei, Multiscale molecular dynamics via the matched interface and boundary (MIB) method, Journal of Computational Physics, 230, 435-457 (2010).
11. Duan Chen, Zhan Chen, Changjun Chen, Weihua Geng and Guo-Wei Wei, MIBPB: A software package for electrostatic analysis, Journal of Computational Chemistry, published online: 15 SEP (2010).
12. Zhan Chen, Nathan Baker and G. W. Wei, Differential geometry based solvation model I: Eulerian formulation, Journal of Computational Physics, 229, 8231-8258 (2010).
13. Duan Chen and Guo-Wei Wei, Modeling and simulation of nano-electronic devices, Journal of Computational Physics, 229, 4431-4460, (2010).
14. Changjun Chen, Rishu Saxena, and Guo-Wei Wei, Differential geometry based multiscale model for virus capsid dynamics, Int. J. Biomed Imaging, Volume 2010, Article ID 308627, 9 pages (2010)
15. Guo-Wei, Wei, Differential geometry based multiscale models, Bulletin of Mathematical Biology, volume 72, 1562-1622, (2010).

Summary

- *Differential geometry based multiscale models for viruses*
- *Topological invariants for virus function characterization*
- *Stable manifolds for identifying coarse-grain clusters*
- *Machine learning methods for dimension reduction*
- *Theoretical prediction agrees with experimental data*