Scholar Commons
USF University of South Florida

University of South Florida
Scholar Commons

Chemistry Faculty Publications | Chemistry

7-24-2014

Web-Based Computational Chemistry Education with CHARMMing II: Coarse-Grained Protein Folding

Frank C. Pickard IV
Laboratory of Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health

Benjamin T. Miller
Laboratory of Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health

Vinushka Schalk
New College of Florida

Michael G. Lerner
Earlham College

H. Lee Woodcock III
University of South Florida, hlw@usf.edu

See next page for additional authors

Follow this and additional works at: http://scholarcommons.usf.edu/chm_facpub

Scholar Commons Citation

Pickard, Frank C. IV; Miller, Benjamin T.; Schalk, Vinushka; Lerner, Michael G.; Woodcock, H. Lee III; and Brooks, Bernard R., "Web-Based Computational Chemistry Education with CHARMMing II: Coarse-Grained Protein Folding" (2014). *Chemistry Faculty Publications.* Paper 13.
http://scholarcommons.usf.edu/chm_facpub/13

This Article is brought to you for free and open access by the Chemistry at Scholar Commons. It has been accepted for inclusion in Chemistry Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

Authors
Frank C. Pickard IV, Benjamin T. Miller, Vinushka Schalk, Michael G. Lerner, H. Lee Woodcock III, and Bernard R. Brooks

This article is available at Scholar Commons: http://scholarcommons.usf.edu/chm_facpub/13

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Education

Web-Based Computational Chemistry Education with CHARMMing II: Coarse-Grained Protein Folding

Frank C. Pickard IV[1☯], Benjamin T. Miller[1☯], Vinushka Schalk[2], Michael G. Lerner[3], H. Lee Woodcock III[4], Bernard R. Brooks[1]*

1 Laboratory of Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, United States of America, **2** Department of Natural Sciences, New College of Florida, Sarasota, Florida, United States of America, **3** Department of Physics and Astronomy, Earlham College, Richmond, Indiana, United States of America, **4** Department of Chemistry, University of South Florida, Tampa, Florida, United States of America

Abstract: A lesson utilizing a coarse-grained (CG) Gō-like model has been implemented into the CHARMM INterface and Graphics (CHARMMing) web portal (www.charmming.org) to the Chemistry at HARvard Macromolecular Mechanics (CHARMM) molecular simulation package. While widely used to model various biophysical processes, such as protein folding and aggregation, CG models can also serve as an educational tool because they can provide qualitative descriptions of complex biophysical phenomena for a relatively cheap computational cost. As a proof of concept, this lesson demonstrates the construction of a CG model of a small globular protein, its simulation via Langevin dynamics, and the analysis of the resulting data. This lesson makes connections between modern molecular simulation techniques and topics commonly presented in an advanced undergraduate lecture on physical chemistry. It culminates in a straightforward analysis of a short dynamics trajectory of a small fast folding globular protein; we briefly describe the thermodynamic properties that can be calculated from this analysis. The assumptions inherent in the model and the data analysis are laid out in a clear, concise manner, and the techniques used are consistent with those employed by specialists in the field of CG modeling. One of the major tasks in building the Gō-like model is determining the relative strength of the nonbonded interactions between coarse-grained sites. New functionality has been added to CHARMMing to facilitate this process. The implementation of these features into CHARMMing helps automate many of the tedious aspects of constructing a CG Gō model. The CG model builder and its accompanying lesson should be a valuable tool to chemistry students, teachers, and modelers in the field.

Introduction

To function properly, most proteins must fold [1]. Determining the structure and understanding the mechanisms responsible for folding are active areas of biophysical research, as gleaning this information may provide critical insights towards fighting diseases that have been linked to protein structure, such as Alzheimer's [2,3]. Experimental determinations of protein structure are typically performed using X-ray diffraction of crystallized proteins or using NMR spectroscopy. Both techniques provide important information about a protein's native folded structure, yet both methods are not without their drawbacks. The process of crystallizing a protein is labor intensive, and structural information from a crystallized sample comes from a nonbiological environment. NMR studies on an aqueous sample yield time-averaged results and are unable to resolve many important dynamic details. Considering the limitations of these techniques, computer simulations are an important tool to supplement and interpret information provided through direct experiment. One way theoretical studies may lead to a better understanding of experimental results is by providing simple models with verifiable results.

Computer models can be constructed at a variety of scales or resolutions (Figure 1). Many simulation techniques represent each atom as a single interaction center. The interactions between beads can then be evaluated using a force field. However, it is possible to construct a coarse-grained (CG) model, in which multiple atoms are represented by a single center (as in the right-hand panel of Figure 1). These types of models can be useful theoretical tools because they can be specifically designed for a system or process of interest. By providing a simplified view of a complex molecular process, important physical details are retained in the model, while superfluous details are ignored. The resulting CG model distills the essence of the biophysical process into its most important physical details, allowing the computational scientist a fundamental understanding of the process.

The success of CG models in representing a variety of biological phenomena is widely acknowledged. Starting from the earliest protein folding simulations [4], to more recent studies on membranes [5] and transmembrane proteins [6], CG models have been applied to diverse phenomena such as protein aggregation [7], vesicle fusion [8], protein structure refinement [9], and the thermodynamics of RNA folding [10], among many other topics. For a thorough discussion of the many diverse applications of CG models, please see the recent review literature [11–13].

Besides their simplification of complex biological phenomena, another attractive feature of CG models is their inexpensiveness relative to all-atom (AA) models. The computational cost of performing a classi-

Citation: Pickard FC IV, Miller BT, Schalk V, Lerner MG, Woodcock HL III, et al. (2014) Web-Based Computational Chemistry Education with CHARMMing II: Coarse-Grained Protein Folding. PLoS Comput Biol 10(7): e1003738. doi:10.1371/journal.pcbi.1003738

Editor: Qiang Cui, University of Wisconsin-Madison, United States of America

Published July 24, 2014

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: This research was supported by the Intramural Research Program of the National Heart, Lung, and Blood Institute, NIH. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: brb@nih.gov

☯ These authors contributed equally to this work.

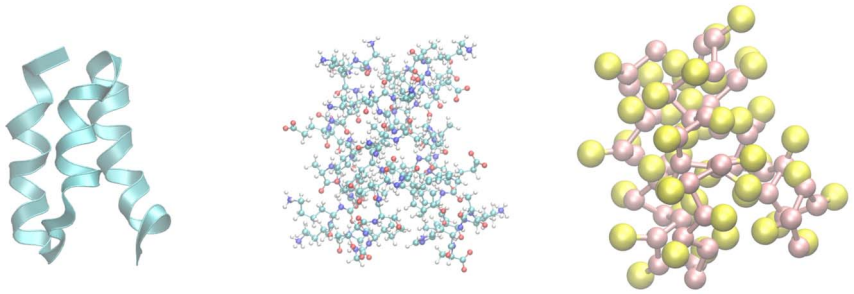

Figure 1. The native structure of the GA module of an albumin binding domain. Three representations of the structure are shown. The left panel shows the backbone of the native fold as a ribbon, which highlights the helical nature of the secondary structure. The center panel shows the AA structure, which is used to build the KT Gō model (right panel).
doi:10.1371/journal.pcbi.1003738.g001

cal molecular mechanics simulation is typically dominated by the calculation of nonbonded forces (electrostatic and van der Waals interactions), as the number of interaction pairs is proportional to the square of the number of particles present in the model system. For example, one of the most widely-used AA water models, TIP3P [14], represents a water molecule as three interaction sites. By comparison, the MARTINI model [5] represents four water molecules using a single CG site, an effective reduction of 12 sites into one. If a ten-to-one mapping is assumed for an entire CG model, the number of required nonbonded interactions to be computed is reduced by roughly two orders of magnitude.

One commonly used CG model is the Gō-like model [15]. In this type of model, an AA crystal structure is used not only to build the coordinates and topology of the CG model (as is the case with most CG models) but also to build the parameters for the non-bonded interactions. The assumption underlying the Gō model is that the native state below the melting temperature (T_m) is the global free energy minimum. We know that the protein folds into a given crystal structure; therefore, we assign parameters to always reproduce the experimentally known fold. Because we know the native state is stable, we know that native contacts are more stable than non-native contacts. Therefore we assign an attractive interaction potential to native contacts, and a short range repulsive interaction potential to non-native contacts. While the Gō model is not based on first principles, nor is it transferable (like the AA protein CHARMM force field) [16], these simulations reach equilibrium (for small globular proteins), allowing for direct comparison between simulation and experiment [17–20].

Gō-like models have been used to describe the kinetic features of protein folding [21]. Among the systems studied are the Trp-cage fast folder [22], the Villin headpiece [23], ribosomal protein S6 [24], and c-src SH3 and CI2 [25]. Various physical insights have been gained from these studies; however, the performance of a particular model depends upon how accurately it can reproduce AA kinetics. Some investigation of optimal parameters has been made [26]. An accurate Gō-like model can provide these insights at a much smaller computational cost than its AA counterpart at the expense of transferability and fine-grained insights into structural behavior.

A consequence of coarse-graining the molecular model is the smoothing of the interaction potential. This CG potential has two important implications: larger integration time step and accelerated dynamics. In the reductive mapping process, many lighter particles are lumped together to form massive interaction centers; this effectively integrates out the fastest moving degrees of freedom, allowing the usage of a longer integration time without incurring integration errors (phantom heating). For example, in AA simulations the hydrogen stretching motions occur at $\sim 3000\text{cm}^{-1}$, dictating an integration time step of 1 or 2 fs. By removing this high-frequency motion, one may safely use a longer time step. This allows a total dynamics simulation time one order of magnitude longer than that of an AA model over the same number of energy and force evaluations. The smoother interaction potential also indirectly accelerates MD by removing degrees of freedom, causing energetic barriers between conformations to disappear. This effect can accelerate the rate of biophysical processes by another factor of two [12]. Taken together, these effects can facilitate the simulation of many biological processes, such as protein folding, that would be impossible to rapidly simulate using AA dynamics. CG models are therefore very attractive in an educational environment, as they retain the qualitative features of their AA counterparts and can be performed cheaply and rapidly, leading to a far more interactive experience with the students.

In this lesson, we will use a Gō-like model in which each amino acid is mapped into two CG sites, one located at the alpha carbon (C_α) and one located at the side-chain center of mass (SC). Many Gō-like models only provide for one interaction site per residue. One such popular model has been developed by Karanicolas and Brooks [27,28]. A web server that generates CHARMM input files for this model is available. Users are not able to perform test runs of their models through the server itself, and there are no interactive lessons, making it less comprehensive than the current CHARMMing implementation. However, some tutorial materials are available from the Multiscale Modeling Tools for Structural Biology (MMTSB) website [29].

In the remainder of this work we describe some of the details underlying both the bonded and non-bonded forces in the Gō model potential. We also discuss the details of the dynamics simulations performed by CHARMMing [30] (www.charmming.org) using the CHARMM simulation package [31]. A procedure for extracting thermochemical values from the raw dynamic trajectories is also given. We then outline step-by-step directions for setting up, performing, and analyzing the dynamics simulation using CHARMMing. This manuscript will give a basic description of the procedure suitable for undergraduate students. New sections have been added to the CHARMM tutorial available at www.charmmtutorial.org with more in-depth descriptions of the methodology, and these are referenced within this work. Finally, we discuss the utility of this lesson both as an educational tool and as a research aid.

Methods

Two CG protein models have been implemented into CHARMMing [30]. The one described in this manuscript is a two-site Gō model based on the work of Klimov and Thirumalai and described further in this section [18]. The second model characterizes the chemical properties of residues and defines interactions accordingly. For the Gō model, the bonded and nonbonded parameters are explicitly constructed to bias the protein towards the experimental crystal structure. However, the inclusion of a SC particle

serves to hinder the progress down the folding funnel and frustrate the folding process. This creates a better balance between the folded and unfolded states than is typical in one-site Gō models such as the model of Karanicolas and Brooks referenced above, allowing for a more accurate description of folding thermodynamics. Further modifications to the standard Gō model allow the CG model to incorporate hydrogen bonding and secondary structure into the parameterization process.

The bonded and nonbonded force field terms of the two-site Gō model use the same functional form as those used in the AA CHARMM force field [16]. However, the model must be re-parameterized to account for the fact that each amino acid is only represented by two CG beads. This parameterization exploits the nontransferability of the model; parameters are designed to reproduce the secondary and tertiary structure of the AA system. The default parameters for the model are described in detail on charmmtutorial.org; interested readers are encouraged to consult this page (http://charmmtutorial.org/index.php/Coarse_Grained_Go_Models) for more information and default values of these parameters. Those readers wishing to learn about the functional form of the CHARMM force field may visit http://charmmtutorial.org/index.php/The_Energy_Function.

The bonded parameters for adjacent backbone atoms are modified based on the secondary structure (as determined by the STRIDE program [32]) of the AA model. However, the bond strength is somewhat weaker than would be found between backbone atoms of an AA model since these "bonds" actually represent more flexible linkages between residues. In addition, because the entirety of the backbone is represented by a single bead per amino acid, there is little to prevent 180 degree rotations of the side-chain that, in an AA representation, would correspond to a chiral flip. This is a rare occurrence in nature because of the energy barrier involved, and therefore, the model adds a strong improper dihedral term to mimic this barrier.

All nonbonded interactions between sites on non-adjacent amino acids are described by a 12–6 Lennard-Jones (LJ) potential. There are no direct electrostatic components to the energy, however electrostatic effects are incorporated indirectly as will be discussed below. The LJ parameters used for SC particles depend on whether they form a native contact within the crystal structure, defined as occurring when their positions are within 4.5 Å of one another. The objective of any Gō-like model is to preserve such contacts while discouraging others. So in this case the LJ potential is attractive, with the minimum energy occurring when the SC beads are at the same distance as the SC centers of mass in the original crystal structure. The strength of the attraction is determined by an experimentally derived contact potential, the most widely used of which was developed by Miyazawa and Jernigan [33] (however, CHARMMing supports other contact potentials as well). This contact potential is how electrostatics are taken indirectly into account. All other SC–SC interactions are modeled with a non-attractive LJ potential, following the fundamental assumption of Gō -like models that the native structure is most stable. All C_α–$C_\alpha$and C_α–SC interactions are also modeled non-attractively to account for volume exclusion, unless they take part in hydrogen bonding, in which case, the potential is slightly attractive. This method of simulating hydrogen bonds has been shown to produce qualitatively correct results for CG water [5].

One of the parameters of the Klimov-Thirumalai Gō-like model is the relative strength of the nonbonded versus the bonded interactions. The motivation behind this parameter, as described in the supporting information of [20], is that the strength of the native contacts must be scaled in order to provide a physiologically realistic melting temperature; however, the scale factor (nScale) is not known a priori. In the CHARMMing implementation, the user may specify nScale or ask the interface to estimate it. Our method of estimating nScale is to run temperature replica exchange [34], which gives a Boltzmann ensemble of the structure at a range of different temperatures. From these ensembles, a fraction of native contacts can be calculated, which can be used to plot a melting curve (see below), allowing the melting temperature (T_m) to be estimated. If T_m is close to a physiological value for the system of interest, then nScale is considered correct. Otherwise, it is strengthened or weakened depending on whether T_m is too low or too high. A full discussion of temperature replica exchange is beyond the scope of this manuscript, but interested readers may consult http://charmmtutorial.org/index.php/Temperature_replica_exchange. If a user asks CHARMMing to determine nScale for them, the software takes an initial guess (the default being 1, no scaling), runs a brief temperature replica exchange simulation, and adjusts nScale upwards or downwards based on the T_m calculated from that simulation. The process is repeated until an nScale is found that yields a physiological T_m.

Electrostatic interactions are not explicitly present in this model (they are accounted for implicitly through the LJ potential); commonly used implicit solvent schemes (e.g., Generalized Born) are fundamentally incompatible with this implementation of the Gō model. Solvation effects are incorporated into this model via Langevin Dynamics (LD) [35]. Both friction and knocking effects are approximately accounted for in this manner. CHARMMing allows the user to specify a collision frequency (damping coefficient), but $5ps^{-1}$ is employed in our lesson so as not to inhibit conformational transitions. The parameter files produced by CHARMMing set nonbonded cutoffs to 23 Å, with the standard CHARMM switching function operating beyond 18 Å [36].

Data Analysis

When analyzing trajectory data, an important consideration is the choice of reaction coordinate employed in data analysis. As illustrated in Table 1, the choice of reaction coordinate can have an

Table 1. 1PRB thermochemistry.

coord.	T_m	ΔH	ΔC_p
R_g	356.5	177.0	2.50
Δ	356.0	171.2	2.73
Q	355.2	170.0	2.94
Expt. [44]	345.3	170	1.1 ± 0.1
Expt. [45]	366.2	269	3.35

A comparison of three reaction coordinates, radius of gyration (R_g), backbone root mean squared deviation (Δ), and fraction of native contacts (Q), with experimental results [44,45]. All results computed using nScale = 0.91.

doi:10.1371/journal.pcbi.1003738.t001

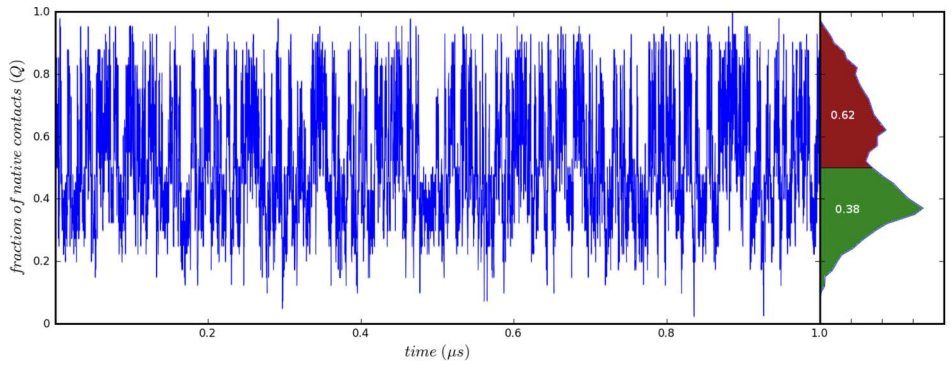

Figure 2. An example trajectory. The fraction of native contacts are plotted as a function of time. By inspecting the histogram at the right, we observe that this trajectory was simulated below T_m, as the protein is folded for 62% of the simulation.
doi:10.1371/journal.pcbi.1003738.g002

uncomfortably large effect upon the computed properties of the simulation. The choice of a reaction coordinate is very complex and depends heavily on what scientific questions are being asked about the system under study. Users are encouraged to consult the literature broadly when choosing a reaction coordinate for a new system. In this lesson, we will consider the fraction of native contacts (Q), a reaction coordinate which has been shown to be robust [37]. This reaction coordinate is often used when mapping the thermodynamic landscape of various folding pathways of a protein. When $0.5 \leq Q \leq 1$, a protein is considered folded; when $0 \leq Q < 0.5$, a protein is considered unfolded. The T_m of a protein occurs when it is equally likely to be folded as unfolded. Figure 2 shows an example trajectory of a protein below its melting point. From the relative frequencies of folded versus unfolded structures, we can calculate ΔG_{fold}. Furthermore, by considering how ΔG_{fold} changes with respect to temperature, we can plot a melting curve and apply the Gibbs-Helmholtz equation to determine the protein's heat capacity (C_p) and enthalpy of fusion (ΔH) (see equation below). Figure 3 gives an illustrative example of a computed melting curve.

$$\Delta G = \left[\Delta H_m (1 - \frac{T}{T_m})\right] - \Delta C_p \left[(T_m - T) + T \ln(\frac{T}{T_m})\right]$$

The Coarse-Grained Lesson

In this lesson, we consider the protein G related albumin-binding domain (PDB code: 1PRB). This is a small globular protein that has been extensively studied, both experimentally and theoretically [37–39]. Experiment has shown that this protein folds near the semi-empirical speed limit of $\sim 1\mu s$ [40]. In solution, 1PRB forms a three-helix bundle (Figure 1); it readily interconverts between its folded and unfolded states, allowing rapid convergence of its thermodynamic properties. These features, taken together, make it an ideal candidate to demonstrate the efficacy of CG methodologies in an educational environment. The sequence of steps that are carried out by the lesson, along with their inputs and outputs, are shown in Figure 4.

Steps 1 and 2: Upload a Protein to CHARMMing and Build the CG Structure

First CHARMMing must be directed to obtain an AA crystal structure for our protein of interest. The user must navigate to the "Submit Structure" page from the main menu. He or she selects "Retrieve a PDB using a PDB ID", and enters the PDB code "1PRB" into the text box. For "What Lesson is this structure associated with?", "Lesson 5" should be selected from the drop-down menu. The user then submits the structure. CHARMMing will now redirect to the page entitled "Build/Select Working Structure."

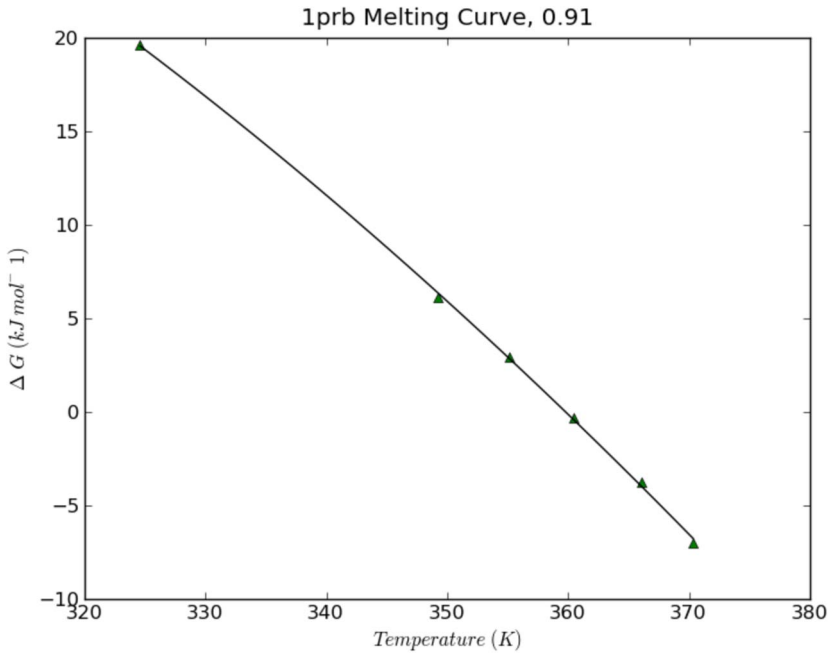

Figure 3. An example melting curve. Each point on the plot represents ΔG_{fold} calculated from a full trajectory. T_m occurs when $\Delta G_{fold} = 0$, and the curvature of the plot is related to C_p.
doi:10.1371/journal.pcbi.1003738.g003

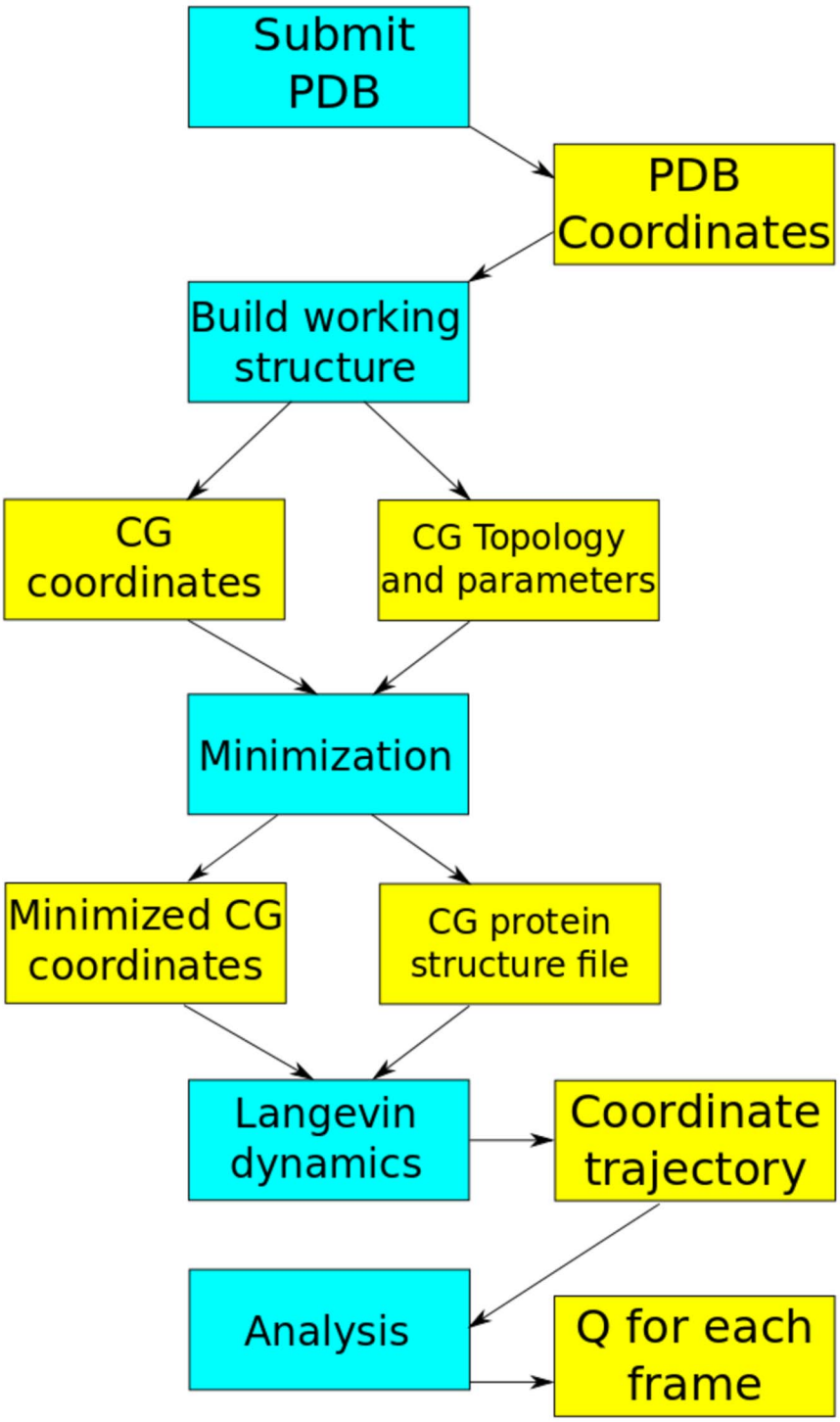

Figure 4. The workflow of the KT Gō model lesson. Blue boxes represent the steps described in the text. Yellow boxes show the inputs and outputs of each step. Note that the CHARMM protein structure file is only built when a calculation is performed (minimization, in this case). *Q* is the fraction of native contacts as defined in the text.
doi:10.1371/journal.pcbi.1003738.g004

On the "Build/Select Working Structure" page, an arbitrary name may be provided at the user's discretion to label the working structure. Next, the user must select what type of structure CHARMMing should build, select the radio button next to "A Klimov-Thirumalai style Gō model" and indicate that the "a-pro" segment should be used to build the model by selecting the appropriate box. These options tell CHARMMing to construct a Gō model with the appropriate CG topology, representing 1PRB. Also on this page, the user must tell CHARMMing how to generate parameters for the model system. We will use the Miyazawa-Jernigan statistical potential as the basis for generating the strengths of the native contact interactions. Under the "Contact types" heading, the "MJ" radio button should be selected. Next, the nScale parameter must be assigned; for 1PRB the optimal value is 0.91. It is important to note that the user would normally need to tune this value manually, as described above. The user can leave all other parameters at their default values and click "Submit."

Step 3: Light Minimization

Although the native crystal structure is determined to be at the free-energy minimum, we still ask the user to do a short energy minimization. This serves two purposes: firstly, it removes any minor issues with the geometry of the constructed structure, such as suboptimal non-native contacts or imperfect secondary structure, and secondly, it allows the user to verify that the model has been successfully built. In order to accomplish this, the user must go to the "Minimization" page and request ten steps of steepest descent minimization and 100 steps of adopted basis Newton-Raphson minimization. The process of minimization is described more extensively in the first article of this series and on charmmtutorial.org. Because a coarse-grained model has no high frequency motions, SHAKE, an algorithm used to constrain rapid bond and angle vibration [41], may be turned off.

Step 4: Run the Langevin Dynamics Simulation

Once minimization is done, the user must go to the "Langevin Dynamics" page. As described above, Langevin dynamics mimics the frictional effects of solvent. CHARMMing is limited to 1,000 steps of dynamics, as described in the first article of this series. However, users have the option of downloading all generated inputs and running them locally, assuming that they have a CHARMM license. When submitting the Langevin dynamics calculation, the user is instructed to set the "FBETA" value to five. This controls the simulated collision frequency used in the dynamics, which simulates the jostling effects of solvent and also serves to couple the system to a heat bath (in CHARMMing, this is always set at 300 K). CHARMMing uses a 1 fs time step for all dynamics calculations; for a production run this may be increased for the reasons given above.

Step 5: Perform Data Analysis

One of the most reliable metrics for gauging whether a protein is in a folded or unfolded state is the fraction of native contacts that are present (Q—see above for details). We have developed a native contacts calculator for Gō models in CHARMMing, and as the final step of the lesson, the user must use it to see how the fraction changes. Because the simulation is so short, native contacts are unlikely to change much, if at all; however, this is still a useful exercise since the same analysis can be applied to production simulations. The user is directed to visit the "KT-Go Native Contacts" page. Once he or she navigates to it, a table showing Q every 50 steps is displayed immediately; in the present version, no form needs to be filled in and submitted. At this point, all calculations for the lesson have been completed, and the user should have gained a very basic understanding of the methods and results.

Discussion and Conclusion

CG models are widely used in the study of biophysics because of their ability to efficiently simulate biomolecules and qualitatively test hypotheses. The same attributes that make them suitable as research tools (their speed and their qualitative nature) also make them suitable as an education tool. While building and running CG models in simulation software previously required the expertise of a specialist in the field, this is no longer true, as the CHARMMing web portal to CHARMM now automates many of the tasks required to build a Gō model, allowing this class of CG models to be used in the classroom.

CG Gō models and other reduced lattice models also have tremendous applications for research into the folding of small globular proteins, as they can correctly predict their thermodynamic properties [37], qualitative folding pathways [42], and ensembles of partially folded structures in the presence of various denaturing agents such as heat and chemicals [20]. The current work allows for the automation of the tedious process of parameter generation and model building for Gō models. Analogous tools for AA simulation have proven very valuable to the modeling community [43].

References

1. Fersht AR (1999) Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. 2nd edition. New York: Freeman.
2. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem 75: 333–366.
3. Thirumalai D, Klimov D, Dima R (2003) Emerging ideas on the molecular basis of protein and peptide aggregation. Curr Opin Struct Biol 13: 146–159.
4. Levitt M, Warshel A (1975) Computer simulation of protein folding. Nature 253: 694–698.
5. Marrink SJ, de Vries AH, Mark AE (2004) Coarse grained model for semiquantitative lipid simulations. J Phys Chem B 108: 750–760.
6. Bond PJ, Sansom MSP (2006) Insertion and assembly of membrane proteins via simulation. J Am Chem Soc 128: 2697–2704.
7. Bieler NS, Knowles TPJ, Frenkel D, Vácha R (2012) Connecting macroscopic observables and microscopic assembly events in amyloid formation using coarse grained simulations. PLoS Comput Biol 8: e1002692.
8. Marrink SJ, Mark AE (2003) The mechanism of vesicle fusion as revealed by molecular dynamics simulations. J Am Chem Soc 125: 11144–11145.
9. Delarue M, Dumas P (2004) On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models. Proc Natl Acad Sci U S A 101: 6957–6962.
10. Denesyuk NA, Thirumalai D (2013) Coarse-grained model for predicting rna folding thermodynamics. J Phys Chem B 117: 4901–4911.
11. Tozzini V (2005) Coarse-grained models for proteins. Curr Opin Struct Biol 15: 144–150.
12. Saunders MG, Voth GA (2013) Coarse-graining methods for computational biology. Annu Rev Biophys 42: 73–93.
13. Chan HS, Zhang Z, Wallin S, Liu Z (2011) Cooperativity, local-nonlocal coupling, and nonnative interactions: Principles of protein folding from coarse-grained models. Annu Rev Phys Chem 62: 301–326.
14. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79: 926–935.
15. Ueda Y, Taketomi H, Gō N (1978) Studies on protein folding, unfolding, and fluctuations by computer simulation. ii. a. three-dimensional lattice model of lysozyme. Biopolymers 17: 1531–1548.
16. MacKerell AD Jr (2005) Chapter 7: Empirical Force Fields for Proteins: Current Status and Future Directions. In: Spellmeyer DC, editor. Annual Reports in Computational Chemistry. Volume 1. Atlanta: Elsevier. pp. 91–102. Available: http://www.sciencedirect.com/science/article/pii/S1574140005010078. Accessed 25 June 2014. doi:http://dx.doi.org/10.1016/S1574-1400(05)01007-8.
17. Veitshans T, Klimov D, Thirumalai D (1997) Protein folding kinetics: timescales, pathways and energy landscapes in terms of sequence-dependent properties. Fold Des 2: 1–22.
18. Klimov DK, Thirumalai D (1998) Linking rates of folding in lattice models of proteins with underlying thermodynamic characteristics. J Chem Phys 109: 4119–4125.
19. Klimov DK, Thirumalai D (2000) Mechanisms and kinetics of β-hairpin formation. Proc Natl Acad Sci U S A 97: 2544–2549.
20. O'Brien EP, Ziv G, Haran G, Brooks BR, Thirumalai D (2008) Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. Proc Natl Acad Sci U S A 105: 13403–13408.
21. Hills RD, Brooks III CL (2009) Insights from coarse-grained Gōmodels for protein folding and dynamics. Int J Mol Sci 10: 889–905.
22. Linhananta A, Boer J, MacKay I (2005) The equilibrium properties and folding kinetics of an all-atom Gōmodel of the trp-cage. J Chem Phys 122: 114901.
23. Faccioli P, Sega M, Pederiva F, Orland H (2006) Dominant pathways in protein folding. Phys Rev Lett 97: 108101.
24. Wu L, Zhang J, Wang J, Li WF, Wang W (2007) Folding behavior of ribosomal protein S6 studied by modified Gō-like model. Phys Rev E 75: 031914.
25. Cho SS, Levy Y, Wolynes PG (2006) P versus Q: Structural reaction coordinates capture protein folding on smooth landscapes. Proc Natl Acad Sci U S A 103: 586–591.
26. Sulkowska JI, Cieplak M (2008) Selection of optimal variants of Gō-like models of proteins through studies of stretching. Biophys J 95: 3174–3191.
27. Karanicolas J, Brooks III CL (2002) The origins of asymmetry in the folding transition states of protein L and protein G. Protein Sci 11: 2351–2361.
28. Karanicolas J, Brooks III CL (2003) Improved Gō-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. J Mol Biol 334: 309–325.
29. Feig M, Karanicolas J, Brooks III CL (2004) MMTSB tool set: enhanced sampling and multiscale modeling methods for applications in structural biology. J Mol Graph Model 22: 377–395.
30. Miller BT, Singh RP, Klauda JB, Hodošček M, Brooks BR, et al. (2008) CHARMMing: A New, Flexible Web Portal for CHARMM. J Chem Inf Model 48: 1920–1929.
31. Brooks BR, Brooks III CL, Mackerell AD, Nilsson L, Petrella RJ, et al. (2009) CHARMM: The biomolecular simulation program. J Comput Chem 30: 1545–1614.
32. Heinig M, Frishman D (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res 32: W500–W502.
33. Miyazawa S, Jernigan RL (1996) Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. J Mol Biol 256: 623–644.
34. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 314: 141–151.
35. Leach AR (2001) Molecular Modelling: Principles and Applications. 2nd edition. Harlow, Essex: Pearson Education Limited. 388–389 pp.
36. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, et al. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4: 187–217.
37. Best RB, Hummer G (2010) Coordinate-dependent diffusion in protein folding. Proc Natl Acad Sci U S A 107: 1088–1093.
38. Takada S (2001) Protein folding simulation with solvent-induced force field: Folding pathway ensemble of three-helix-bundle proteins. Proteins 42: 85–98.
39. Wang T, Zhu Y, Gai F (2004) Folding of a three-helix bundle at the folding speed limit. J Phys Chem B 108: 3694–3697.
40. Hagen SJ, Hofrichter J, Szabo A, Eaton WA (1996) Diffusion-limited contact formation in unfolded cytochrome c: estimating the maximum rate of protein folding. Proc Natl Acad Sci U S A 93: 11615–11617.
41. Ryckaert JP, Ciccotti G, Berendsen HJ (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23: 327–341.

42. Kmiecik S, Kolinski A (2008) Folding pathway of the B1 domain of Protein G explored by multiscale modeling. Biophys J 94: 726–736.
43. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, et al. (2010) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J Comput Chem 31: 671–690.
44. Rozak DA, Orban J, Bryan PN (2005) G148-ga3: A streptococcal virulence module with atypical thermodynamics of folding optimally binds human serum albumin at physiological temperatures. BBA-Proteins Proteom 1753: 226–233.
45. Zhu Y, Fu X, Wang T, Tamura A, Takada S, et al. (2004) Guiding the search for a protein's maximum rate of folding. Chem Phys 307: 99–109.