

# A Hierarchical Coarse-Grained (All-Atom-to-All-Residue) Computer Simulation Approach: Self-Assembly of Peptides

Ras B. Pandey<sup>1\*</sup>, Zhifeng Kuang<sup>2</sup>, Barry L. Farmer<sup>2</sup>

<sup>1</sup> Department of Physics and Astronomy, University of Southern Mississippi, Hattiesburg, Mississippi, United States of America, <sup>2</sup> Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright Patterson Air Force Base, Dayton, Ohio, United States of America

## Abstract

A hierarchical computational approach (all-atom residue to all-residue peptide) is introduced to study self-organizing structures of peptides as a function of temperature. A simulated residue-residue interaction involving all-atom description, analogous to knowledge-based analysis (with different input), is used as an input to a phenomenological coarse-grained interaction for large scales computer simulations. A set of short peptides P1 (<sup>1</sup>H <sup>2</sup>S <sup>3</sup>S <sup>4</sup>Y <sup>5</sup>W <sup>6</sup>Y <sup>7</sup>A <sup>8</sup>F <sup>9</sup>N <sup>10</sup>N <sup>11</sup>K <sup>12</sup>T) is considered as an example to illustrate the utility. We find that peptides assemble rather fast into globular aggregates at low temperatures and disperse as random-coil at high temperatures. The specificity of the mass distribution of the self-assembly depends on the temperature and spatial lengths which are identified from the scaling of the structure factor. Analysis of energy and mobility profiles, gyration radius of peptide, and radial distribution function of the assembly provide insight into the multi-scale (intra- and inter-chain) characteristics. Thermal response of the global assembly with the simulated residue-residue interaction is consistent with that of the knowledge-based analysis despite expected quantitative differences.

**Citation:** Pandey RB, Kuang Z, Farmer BL (2013) A Hierarchical Coarse-Grained (All-Atom-to-All-Residue) Computer Simulation Approach: Self-Assembly of Peptides. PLoS ONE 8(8): e70847. doi:10.1371/journal.pone.0070847

**Editor:** Eugene A. Permyakov, Russian Academy of Sciences, Institute for Biological Instrumentation, Russian Federation

**Received:** May 21, 2013; **Accepted:** June 24, 2013; **Published:** August 13, 2013

**Copyright:** © 2013 Pandey et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work is supported by the Air Force Research Laboratory. The funders had no role in 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.

\* E-mail: ras.pandey@usm.edu

## Introduction

Peptides are some of the most versatile constituents in designing advanced materials, from bio-functionalized nanoparticles [1–10] to modulating the kinetics of proteins in cells and beyond. Specificity of the amino acids in a short peptide chain is key to their selective binding (covalent and non-covalent) to substrates. Peptides have become a valuable constitutive component in both materials as well as drug design due to prolific conformational response with specific residue interactions. Understanding the unique interaction of a peptide is a challenge in itself. For example, the residue-residue interaction in a solution of free amino acids could be different from a residue-residue interaction in an isolated peptide (intra-chain) or in a peptide melt due to the interplay between the steric constraints of covalent peptide bonds and specific residue interactions. Including solvent, substrate, and other components enhances the complexity in understanding the effect of residue-residue interactions. To probe such systems, one has to start from the building blocks, i.e. amino acids, and develop a feasible method (e.g., bottom-up) to examine the consequences of residue interactions. In this article we introduce such an approach and address how peptides assemble. Since temperature competes directly with the interactions, it is a natural parameter to examine its response in peptide assembly and dispersion.

Residue-residue interaction [11–18] is critical in understanding the multi-scale equilibrium structure of large peptides and proteins where it is not feasible to incorporate atomic scale details, a

challenging issue in computational modeling. Some degree of coarse-graining and approximations are therefore unavoidable in order to carry out large-scale simulations [19–30]. Such procedures include developing effective interaction potentials among residues, exploring the phase space selectively, resorting to efficient and effective methods, etc. In modeling the structure of proteins, knowledge-based contact matrix [11–18] is extensively used to develop phenomenological residue-residue interactions. A number of knowledge-based contact potentials based on a growing ensemble of protein data base has been developed to understand the folding dynamics of proteins. We propose a simulated residue-residue interaction based on an all-atom Molecular Dynamics (MD) simulation. Analogous to knowledge-based phenomenological interaction [17,18], simulated residue-residue interaction matrix can be used as an input to phenomenological interaction in hierarchy to carry out large-scale computer simulations (see below). Such a coarse-grained approach has been recently used to understand the binding of peptides with a graphene substrate [1] where the simulated interaction (residue-substrate) matrix is relatively small. Results of all-atom approach, i.e., the relative binding of each residue are verified by the coarse-grained method before large-scale simulations were performed [1]. In this article, we focus on a residue-residue interaction matrix, which is much larger than the residue-substrate [1] interaction matrix and study the self-assembly of peptides, P1.

As mentioned above, solvent plays an important role in modulating the structure and assembly of peptides and proteins.

Understanding the self-assembly of unsolvated peptides (i.e. in vacuum) first is also important before investigating the effect of solvent. A number of neutron scattering experiments are recently performed on powder samples of proteins where simulations in vacuum are used to interpret the scattering data [31–33]; these studies also include the effect of solvent by identifying the differences. Thus, apart from the simplicity, there is a value in exploring the structure and dynamics of peptides in vacuum (appropriate for powder samples), i.e., constraining to residue-residue interactions alone before we incorporate the solvent.

## Model and Method

Simulations are carried out in two steps in hierarchy (bottom-up coarse-graining): (i) estimate the residue-residue interaction among 20 amino acids (210 interaction pairs of the 20×20 matrix) using an all-atom MD simulation, and (ii) use the simulated residue-residue interaction matrix as an input to a phenomenological interaction in the coarse-grained representation of peptide chains (see below).

### All-atom approach

The actual residue interaction inside proteins and between proteins depends on amino acid size, geometry, conformation and the local biochemical environment. It is a very difficult task to develop a portable force field to take all the effects into account. The extreme simplicity of the potential function is based on the hypothesis that a system fluctuates around an equilibrium reference configuration. To find the equilibrium reference configuration between two residues, we resort to molecular dynamics simulation in vacuum using the AMBER ff99SB force field.

A total of 210 residue pairs are simulated in the same protocol using NAMD2.9 simulation software. The initial backbone positions of two residues are the same. The mass centers of backbones are 1 nm apart. Each amino acid is capped by an acetyl beginning group, ACE, and an N-methylamine ending group, NME, to avoid strong terminal interactions and to mimic the bond connectivity. Due to bond connectivity and counter ionic effect in solution, like charged residues can appear side by side in a protein. For charged residues, necessary counter ions are added to neutralize the side chains. Each system is minimized for 2000 steps using a sophisticated conjugate gradient and line search algorithm, then heated up to 300 K increasing 30 K from 0 K every 1000 steps, and then equilibrated for 20 ns in vacuum setting a cutoff distance of 1.2 nm and a switching distance of 1.0 nm for both van der Waals and electrostatic interactions. Finally, another 10 ns production run is performed and trajectories are stored every 10 ps for each system.

The above MD simulations are independently repeated three times for each system. The obtained total system potentials in the last 10 ns for each of the three independent simulations are averaged along their own trajectories and compared. The trajectories corresponding to the lowest average total potential in the last 10 ns are used to calculate the interaction energy between two residues according to the X-PLOR van der Waals function

$$V_{ij} = \left\{ 4E_{ij} \left[ \left( \frac{\Sigma}{R_{ij}} \right)^{12} - \left( \frac{\Sigma}{R_{ij}} \right)^6 \right] + Q_i Q_j \frac{C}{R_{ij}} \right\} SW(R_{ij}, R_{on}, R_C), \quad (1)$$

$$R_{ij} < R_C$$

where  $R_{ij}$  is the distance between two atoms.  $Q_i$  and  $Q_j$  are atomic partial charges.  $C$  is a dimensional constant. The cutoff distance

$R_C$  is equal to 1.2 nm. The switching distance  $R_{on}$  is 1.0 nm. The switching function  $SW$  has the form

$$SW(R_{ij}, R_{on}, R_C) = \begin{cases} 0 & \text{if } R_{ij} > R_C \\ \frac{(R_{ij}^2 - R_C^2)^2 (R_C^2 + 2R_{ij}^2 - 3R_{on}^2)}{(R_C^2 - R_{on}^2)^2} & \text{if } R_C > R_{ij} > R_{on} \\ 1 & \text{if } R_{ij} < R_{on} \end{cases} \quad (2)$$

The interaction energy matrix is presented in Table 1. In our simulations, we assume that all residue pairs are exposed to each other. In reality, hydrophobic residues are buried inside a protein, whereas hydrophilic residues are exposed to the environment. We constrain here to residue-residue interactions alone for simplicity; solvent (explicit and implicit) could be incorporated to modulate the distribution of residues (hydrophobic or hydrophilic depending on the nature of solvent) and therefore the structure of the peptide accordingly.

### All-residue approach

The intra-molecular detail of each residue is ignored in a coarse-grained description of the peptide, which is a set of nodes tethered together by flexible peptide bonds on a cubic lattice [1]. A residue is represented by a node and its specific characteristics are captured by unique residue-residue interactions (see Table 1). We use a bond-fluctuation method as before [17,18] to exploit the efficiency of the discrete lattice with ample degrees of freedom. Peptide nodes interact with neighboring nodes with a generalized Lennard-Jones potential,

$$U_{ij} = \left[ \varepsilon_{ij} \left( \frac{\sigma}{r_{ij}} \right)^{12} + \varepsilon_{ij} \left( \frac{\sigma}{r_{ij}} \right)^6 \right] \quad r_{ij} < r_c \quad (3)$$

where  $r_{ij}$  is the distance between the residues at site  $i$  and  $j$ ,  $r_c = \sqrt{\delta}$  is the range of interaction and  $\sigma = 1$  in units of lattice constant. The residue-residue pair interaction (Table 1) is used for the coefficients  $\varepsilon_{ij}$  (a measure of the depth) of the generalized LJ potential.

We consider a cubic lattice of size  $L^3$ . Peptides (P1) of a volume fraction ( $C_p$ ) are then placed in the box in random configurations subject to excluded volume constraint. With the constraints on fluctuating bond length  $l$  ( $2 \leq l \leq \sqrt{10}$ ) in units of lattice constant) and excluded volume, each residue performs its stochastic motion with the Metropolis algorithm as follows. An attempt is made to move a randomly selected residue of a randomly selected peptide chain from its current position at site ( $i$ ) to a neighboring site  $j$  with the Boltzmann probability  $\exp(-\Delta E_{ij}/T)$ , where  $\Delta E_{ij}$  is the change in energy (Eq. 3) between its new ( $E_j$ ) and old ( $E_i$ ) configuration  $\Delta E_{ij} = E_j - E_i$  and  $T$  is the temperature in reduced units of the Boltzmann constant and the energy ( $\varepsilon_{ij}$ ). Attempts to move each residue in the simulation box once defines the unit Monte Carlo step (MCS). Simulations are performed for a sufficiently long time to identify the structural changes from small to large scales. A number of local and global physical quantities are evaluated during the simulations including the structural profile of each residue, the variations of the root mean square (RMS) displacement of the center of mass of each peptide and the radius of gyration with the time steps, structure factor and radial distribution function of the self-assembled structures. Simulations are performed at different temperatures at a peptide concentration  $C_p = 0.1$  with as many as 100 independent samples on a  $64^3$  lattice to estimate the average value of the physical quantities. We have carried out simulations with different peptide concentrations but

**Table 1.** The minimum pairwise interaction energy (kcal/mol) of 20 amino acids from all-atom MD simulation in vacuum (like charge pair interaction energy, i.e., D-D, D-E, E-E, R-R, R-K and K-K are positive all others are negative).

	A	G	V	L	I	P	F	M	W	C	D	E	R	K	H	N	Q	S	T	Y
A	6																			
G	3	7																		
V	5	7	3																	
L	5	8	4	4																
I	3	4	3	4	6															
P	2	2	2	2	5	1														
F	8	6	4	8	6	4	7													
M	7	8	5	8	5	4	6	8												
W	9	9	9	12	11	11	14	14	15											
C	7	10	4	4	8	9	6	8	8	5										
D	21	23	22	21	21	11	20	27	26	30	67									
E	22	21	21	22	21	15	22	27	11	29	68	78								
R	22	21	27	26	25	27	23	24	30	19	132	138	41							
K	21	23	23	26	25	29	31	6	39	22	138	149	45	43						
H	11	9	-9	-9	8	10	13	14	18	11	29	30	30	29	15					
N	11	14	11	10	9	6	12	14	15	13	22	25	19	24	16	18				
Q	12	12	11	11	10	11	13	12	12	12	37	30	31	30	15	19	23			
S	11	14	9	9	6	3	13	9	16	7	39	34	22	23	14	13	18	13		
T	10	12	6	10	6	3	12	8	12	11	39	39	19	22	17	13	18	12	10	
Y	10	11	7	7	5	9	10	12	11	10	25	24	35	29	14	10	19	11	10	12

doi:10.1371/journal.pone.0070847.t001

we will focus on the low concentration for clarity of the structural evolution with the temperature. Different size lattices are also used to assure that results on the qualitative trends of the physical quantities remain independent of the sample size.

## Results and Discussion

Peptides (P1) are randomly distributed initially in their random conformations in the simulation box. As residues perform stochastic motion with the Metropolis algorithm, each peptide moves and undergoes conformational changes with the time step. Peptide segments within the range of interaction may be bound due to non-covalent interactions and unbound due to thermal agitation. Distribution of peptides evolves with time and reaches a steady-state configuration. Time to reach the steady-state equilibrium depends on the temperature. Figure 1 shows a set of typical snapshots at the steady state at a range of low-to-high temperatures, which illustrates the differences in morphology due to distribution of peptides. Both, residue-residue interactions and temperature compete in the self-assembly of peptides and their dispersion. At low temperatures ( $T = 0.7, 0.8$ ), the residue-residue interaction is more dominant over the thermal energy. Peptides self-assemble into aggregates with relatively high density due to attractive residue-residue interactions. Peptides disperse at high temperature ( $T = 1.0$ ) (overcoming the non-covalent residue-residue interactions) with almost uniform low density throughout the lattice. The aggregation of peptides is interaction-driven at low temperature where system reaches the steady state rather quickly. In fact, at low temperatures some peptides may be trapped very quickly in the self-assembly process without exploring all possible conformations. Peptides explore conformations rather more thoroughly at high temperatures. Consequences of such a

difference in thermal response of peptides should be reflected in physical properties of peptides and their self-assembly (see below).

As mentioned above, we have analyzed both local and global physical quantities. The energy profile (energy of each residue in each peptide at equilibrium) of the peptide P1 is presented in Figure 2 at different temperatures ( $T = 0.5-1.0$ ). We track the energy of each residue and peptide during the course of the simulation. The energy  $E_n$  ( $n = 1, 2, \dots, 12$ ) of each residue node is the average value evaluated from all peptide chains and all independent samples using Eq. 3. The profile pattern remains nearly the same at low and high temperatures albeit with lower and higher energy values. Two residues with the highest and lowest energy are  $^{10}\text{N}$  and  $^{11}\text{K}$ , respectively, which shows that not only the specificity of a residue is important but also its sequential position (e.g., compare the energy of  $^9\text{N}$  and  $^{10}\text{N}$ ). The corresponding mobility profile of the peptide (see Figure S1) shows that minimum energy of a residue does not necessarily correspond to lowest mobility. The residues at the ends ( $^1\text{H}$ ,  $^{12}\text{T}$ ) are relatively more mobile than those in the interior due to constraints imposed by the peptide bonds. Residues  $^7\text{A}$  and  $^6\text{Y}$  appear to be the most mobile and are the second lowest and second highest energy, respectively (see Figure 2). Thus energy alone is not a measure of mobility of a residue in a peptide chain.

Let us examine how peptides move and conform. Figure 3 shows the variation of the average root mean square (RMS) displacement  $R_c$  of the center of mass of a peptide chain with the time step ( $t$ ) in temperature range  $T = 0.5-1.0$ . The asymptotic dependence of  $R_c$  on  $t$  can be described by a power-law,  $R_c \propto t^\gamma$ , where  $\gamma = 1/2$  describes the diffusive nature of the peptides' motion. At a high temperature ( $T = 0.5-1.0$ ),  $\gamma \approx 1/2$  (see Figure 4) but  $\gamma$  ( $\gamma \approx 0.3-0.4$ )  $< 1/2$  at low temperatures  $T = 0.5-0.8$ , which implies sub-diffusive dynamics of the peptide chain. The residue-residue

interaction dominates over the thermal energy at low temperatures. The peptide chains self-assemble into aggregates on encountering each other. Because of the self-organizing morphology, the mobility of the residues, and therefore the peptide chains, decreases at low temperatures leading to sub-diffusive dynamics. Peptide chains become free and diffuse as residues unbind from assembly at high temperatures. The overall dynamics of peptides is thus consistent with the visual inspection of the snapshots (Figure 1). Variation of the average radius of gyration ( $R_g$ ) of peptide chains with the time step (inset Figure 3) suggests that the conformations have reached equilibrium during the course of simulation at each temperature. How does the equilibrium value of  $R_g$  depend on the temperature? The plot of  $R_g$  with the temperature ( $T$ ) is included in Figure 4 (inset). We see that the radius of gyration increases monotonically on raising the temperature in the low-to-intermediate temperature range ( $T = 0.5-1.0$ ) and approaches a constant at high temperatures.

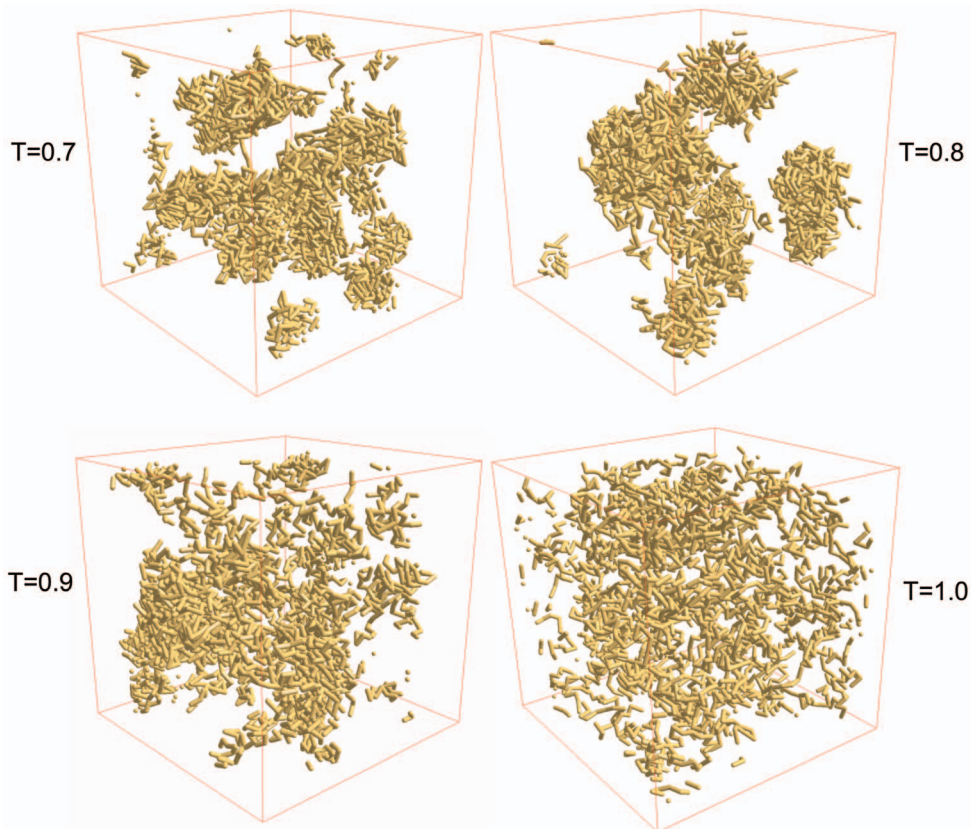
Large-scale structures resulting from the self-assembly of peptides can be studied by examining the radial distribution function (RDF), which is the average number of particles (residues) from the center. We have analyzed (figure S2) the spatial dependence of RDF in the temperature range  $T = 0.5-1.5$ . We find that the rapid assembly of peptides at low temperature ( $T = 0.5$ ) leads to an aggregate with high density at the center followed by a sharp decay with the distance ( $r$ ). The density of aggregates spreads on raising the temperature ( $T = 0.5-0.8$ ). Peptide chains disperse at high temperature where no aggregate develops.

The multi-scale morphology of the self-assembly of peptides can be studied by analyzing the structure factor  $S(q)$  (Figure 4),

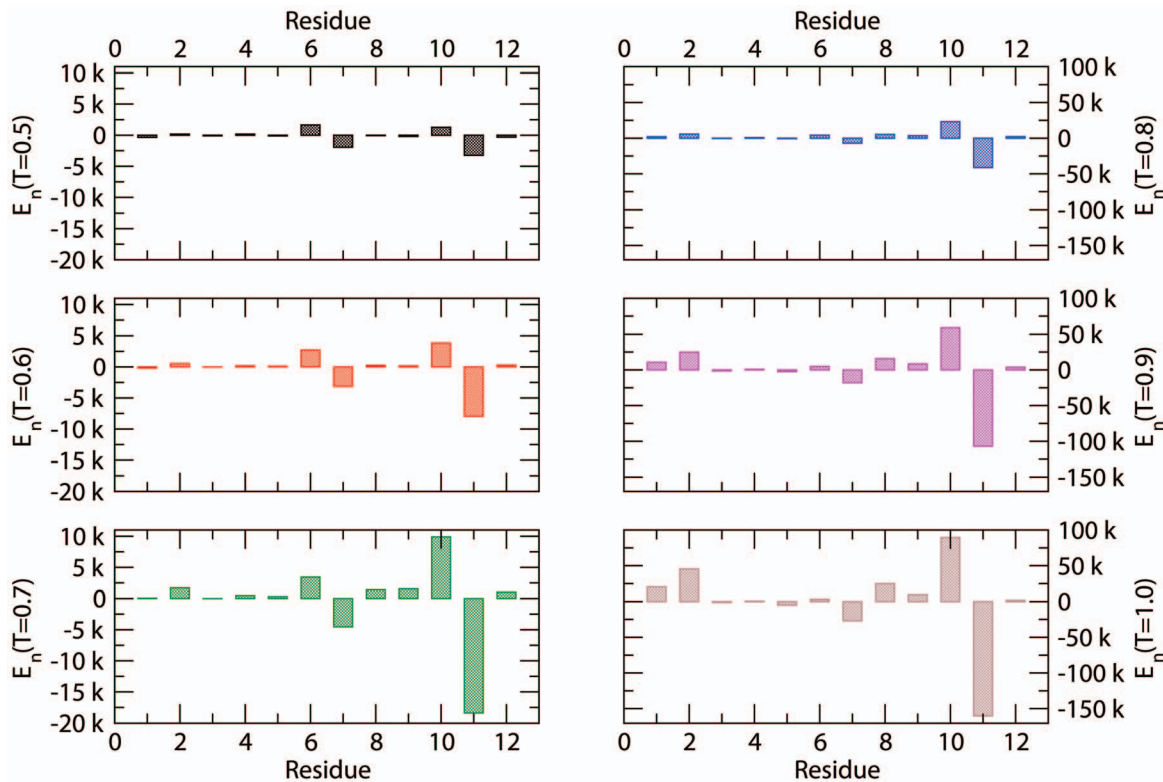
$$S(q) = \left\langle \frac{1}{N} \left| \sum_{j=1}^N e^{-i\vec{q}\cdot\vec{r}_j} \right|^2 \right\rangle_{|\vec{q}|} \quad (4)$$

where  $\vec{r}_j$  is the position of each residue and  $|\vec{q}| = 2\pi/\lambda$  is the wave vector of wavelength,  $\lambda$ . One can study the mass distribution of particles (residues) by estimating the exponent  $\nu$  in the power-law scaling  $S(q) \propto q^{-1/\nu}$ . Spatial scaling of mass ( $M$ ) with the radius of gyration ( $R_g \approx \lambda$ ) of the aggregate,  $M \propto R_g^D$  provides an estimate of its effective dimension  $D = 1/\nu$ . Higher value of  $D$ , e.g.,  $D = 3$  implies a solid while  $D = 2$  represents an ideal chain with a heterogeneous mass distribution on a cubic lattice;  $D > 3$  does not make sense and could be an artifact of fitting data in the wrong regions including crossover regimes.

Variation of the structure factor  $S(q)$  with  $q$  provides an insight into a rather rich structure at all length scales. Note that the wavelength  $\lambda \approx 5-15$  (in units of lattice constant) is comparable to spatial spread of a peptide aggregate and corresponds to wave vector  $q \approx 0.5-1.0$ ; the radius of gyration of peptide  $R_g \approx 3.2-4.2$  (see figure 3). At the low temperature ( $T = 0.5$ ), we see a rather solid morphology ( $q \approx 0.4-0.8$ ). Spreading of the solid-like morphology of self-assembly of peptides is clearly seen on raising the temperature ( $T = 0.5, 0.6$ ). Oscillation in  $S(q)$  sets at smaller scales (order of the chain length) at higher temperature ( $T = 0.8$ ) becomes persistent at all length scales at high temperatures ( $T \geq 1.0$ ) with dispersion of peptide chains. We would like to point out that the oscillatory nature of the



**Figure 1. Snapshots of peptides P1 at  $T = 0.7$  (top left),  $0.8$  (top right),  $0.9$  (bottom left),  $1.0$  (bottom right) at the end of  $5 \times 10^5$  MCS time on a  $64^3$  lattice with peptide concentration  $C_p = 0.1$ .**  
doi:10.1371/journal.pone.0070847.g001



**Figure 2. Energy ( $E_n$ ) of each residue of P1:  $^1\text{H}$   $^2\text{S}$   $^3\text{S}$   $^4\text{Y}$   $^5\text{W}$   $^6\text{Y}$   $^7\text{A}$   $^8\text{F}$   $^9\text{N}$   $^{10}\text{N}$   $^{11}\text{K}$   $^{12}\text{T}$  at temperature  $T=0.5-1.0$ . Simulations are performed on a  $64^3$  lattice with the peptide concentration  $C_p=0.1$  with as many as 100 independent samples to estimate the average. doi:10.1371/journal.pone.0070847.g002**

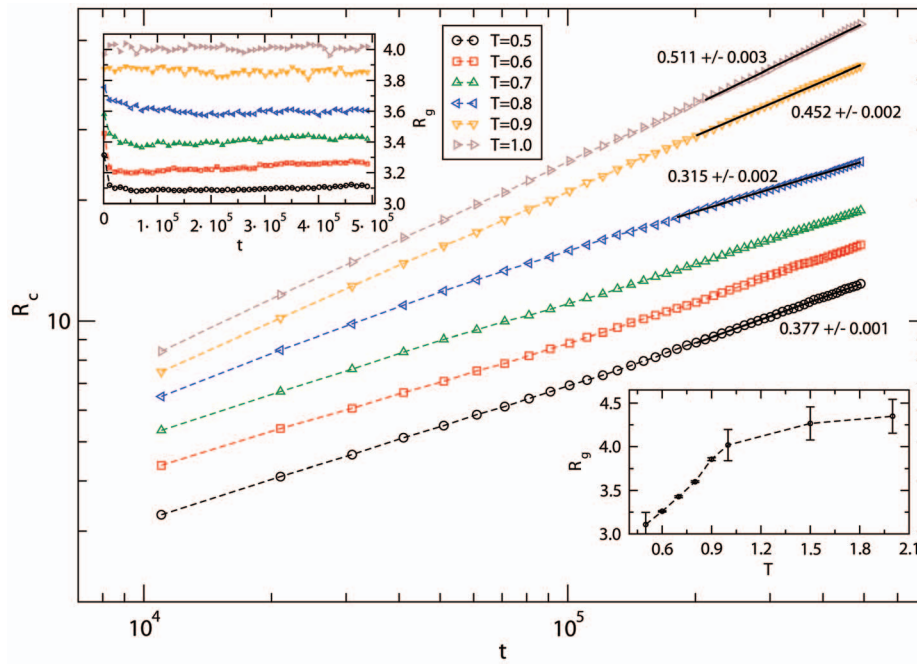
structure factor is not an artifact of cubic lattice considered here but due to thermodynamics of peptides at high temperatures. A general fit of oscillatory data at  $T=1.0$  shows that the mass distribution of peptide is heterogeneous in morphology of a random walk. Thus, the structural evolution of peptides, from a highly dispersed peptide chains at high temperatures (an expected thermodynamic behavior) to its aggregation on reducing the temperature can be studied by such hierarchical coarse-grained approach. Unfortunately, we are not aware of experimental data on such assembly at present.

Alternatively, one may consider other interactions such as knowledge-based interaction (previously used in study of protein folding [17,18]) which is derived from the ensemble of protein structures available in the protein data bank (PDB). X-ray crystallographic data in PDB represent snapshots of proteins' conformations in their native structures in unique solvent. Knowledge-based analysis [12] involves further assumptions and approximations to derive residue-residue contact maps among the residues. Thus, the knowledge-based interaction captures the essence of residue-residue interaction in a different environment which is much more complex than the simulated residue-residue interaction considered here. This does not mean that the knowledge-based interaction is superior than the simulated one as the prior method resorts to a number of assumptions and approximations. Nevertheless it is worth exploring what happens if we use the knowledge-based residue-residue interactions (previously used in study of protein folding [17,18]) in place of simulated interaction? We have carried out such simulation with the classic MJ interaction matrix [12]. Data for the variation of the structure factor  $S(q)$  with  $q$  is presented in figure 5. Note that the temperature scales are dramatically different from those used

with simulated residue-residue interaction matrix. This is primarily due to differences in magnitude of the MJ matrix elements [12] and the simulated interaction (table 1). We see that peptide chains assemble into aggregates with a rather solid density ( $D \approx 3$ ) at the low temperature ( $T=0.010$ ) which shows a dispersive trend of peptides on increasing the temperature (e.g.  $D \approx 2.6$  at  $T=0.012$ ). Despite the shift in temperature scale, the general self-organizing trend (from aggregation to dispersion) appears consistent with our hierarchical coarse-grained approach. Because of the differences in assumptions and approximation made in deriving the knowledge-based contact matrix [12] and the direct simulation of individual residue-residue interaction (i.e. with free residue unlike the residues as a part of protein), quantitative differences (figure 4 and 5) are not un-expected. Such coarse-grained approach with specific residue-residue interaction matrix provides an additional alternative to address complex problems in bio-inspired assembly.

## Conclusions

A hierarchical coarse-graining scheme introduced here thus provides a useful method to investigate multi-scale self-organizing structures of such complex constituents as peptides (e.g. P1). It involves all-atom MD simulations to estimate the residue-residue interactions in which twenty amino acids constitute 210 independent pairs, each of which has its unique interaction energy (Table 1). The simulated interaction matrix forms the basis for the residue-residue interaction similar to knowledge-based contact matrices [12–18] in an all-residue representation of the peptide chain. The structural evolutions are analyzed in detail by examining both local and global physical quantities spanning the entire scale.

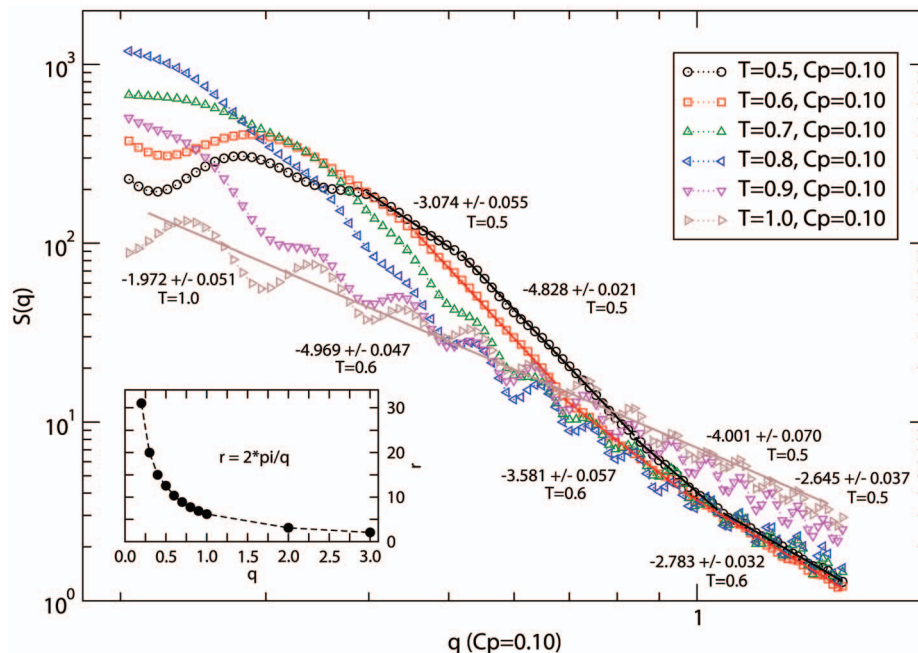


**Figure 3. Variation of the root mean square (RMS) displacement of the center of mass of peptides with the time steps on a log-log scale.** Inset figures show the variation of the radius of gyration with the time step (top left) and dependence of the equilibrium  $R_g$  on the temperature ( $T$ ). Simulations are performed on a  $64^3$  lattice with the peptide concentration  $C_p = 0.1$  with as many as 100 independent samples to estimate the average.

doi:10.1371/journal.pone.0070847.g003

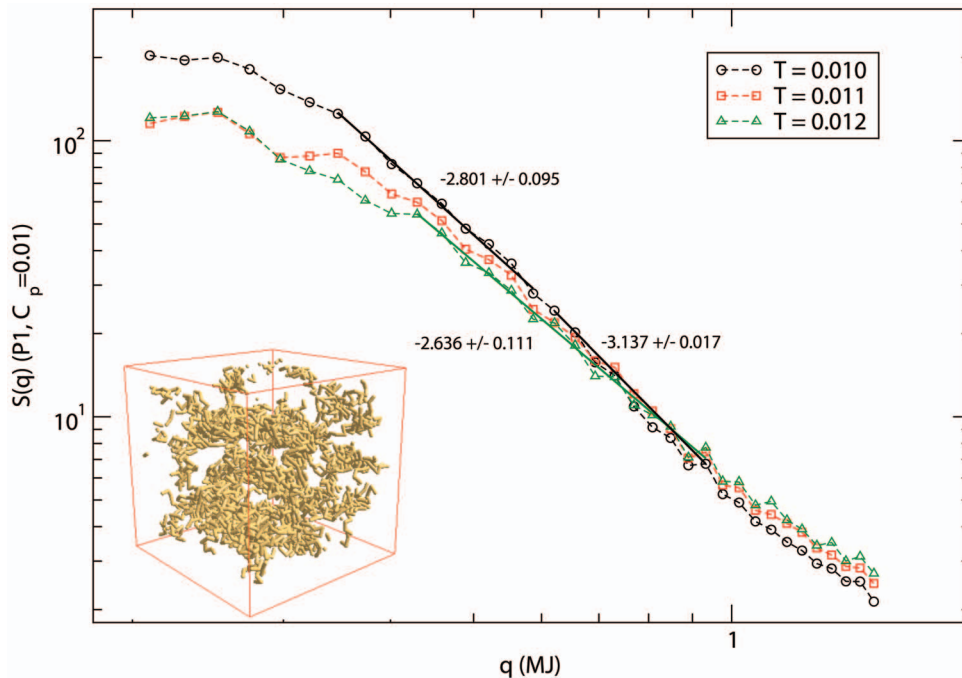
Visualizations of the self-organizing assembly clearly show a systematic change in morphology at a range of length scales with the temperature. We find that the energy profile of residues does not necessarily dictate their mobility, which would have been

expected for a simple system with its thermodynamics controlled primarily by interactions. Spatial distribution of residues within a peptide does respond to self-assembly of peptides. The dynamics of peptides as they perform their stochastic motion during the self-



**Figure 4. Variation of the structure factor  $S(q)$  with the wave vector  $q$  at temperature  $T=0.5-1.0$ .** A spatial scale of the wave vector  $q$  is included in the inset for a guide. Simulations are performed on a  $64^3$  lattice with the peptide concentration  $C_p = 0.1$  with as many as 100 independent samples to estimate the average.

doi:10.1371/journal.pone.0070847.g004



**Figure 5. Variation of the structure factor  $S(q)$  with the wave vector  $q$  at temperature  $T = 0.010, 0.011, 0.012$ .** Miyazawa-Jernigan (MJ) [12] residue-residue interaction is used for the coefficient  $\varepsilon_{ij}$  in coarse-grained potential (eq. 3). Simulations are performed on a  $64^3$  lattice with the peptide concentration  $C_p = 0.1$  with as many as 100 independent samples to estimate the average. Inset is a snapshot at  $T = 0.010$ . doi:10.1371/journal.pone.0070847.g005

assembly and their dispersion shows distinct but appropriate characteristics, i.e., sub-diffusion and diffusion. The radius of gyration of peptides responds linearly to temperature as the self-organized aggregates expand on raising the temperature before approaching saturation at high temperatures. The spatial variation of the radial distribution function reveals that the solid core of the aggregates softens as the density spreads with the temperature. The scaling of the structure factor with the wave vector provides valuable insight into the multi-scale structure of the assembly. For example, on raising the temperature from its low value, the relatively high density of the self-assembly expands over the length scale, which is sensitive to temperature, before reaching a well-dispersed distribution of an ideal chain at high temperature. We hope that this study will stimulate experiments with multi-scale resolution to verify or contradict our predictions. Such a hierarchical coarse-graining is not limited to Monte Carlo simulations with bond-fluctuation methods (used here) but could be extended to such approach as Molecular Dynamics.

As mentioned in the beginning, interaction of peptides in a solvent matrix (implicit or explicit) plays a critical role in modulating the structure and dynamics of peptides and its assembly. For example, we have shown [23], how the structure and dynamics of a protein is affected by the quality of solvent by incorporating the hydrophobicity of each amino acids and its unique interaction in an effective solvent medium. Structure of a histone is found to exhibit a non-monotonic response to solvent quality in a recent Monte Carlo simulation [34]. There are enormous opportunities to improve the simulated residue-residue interaction matrix as well as in incorporating such realistic factors

as solvent and substrate, some of which may be taken up in future efforts.

## Supporting Information

**Figure S1** Mobility ( $M_p$ ; average number of successful hops) of each residue of P1:  $^1\text{H}$   $^2\text{S}$   $^3\text{S}$   $^4\text{Y}$   $^5\text{W}$   $^6\text{Y}$   $^7\text{A}$   $^8\text{F}$   $^9\text{N}$   $^{10}\text{N}$   $^{11}\text{K}$   $^{12}\text{T}$  at temperature  $T = 0.5 - 1.0$ . Simulations are performed on a  $64^3$  lattice with the peptide concentration  $C_p = 0.1$  with as many as 100 independent samples to estimate the average. (TIF)

**Figure S2** Variation of the radial distribution function (RDF), a measure of the number of particles (residues) with distance  $r$  at temperature  $T = 0.5 - 1.5$ . Simulations are performed on a  $64^3$  lattice with the peptide concentration  $C_p = 0.1$  with as many as 100 independent samples to estimate the average. (TIF)

## Acknowledgments

We thank Diana Lovejoy for reading the manuscript and corrections.

## Author Contributions

Conceived and designed the experiments: RBP ZK BLF. Performed the experiments: RBP ZK. Analyzed the data: RBP ZK. Contributed reagents/materials/analysis tools: RBP ZK BLF. Wrote the paper: RBP ZK BLF. Designed the coarse-grained approach: RBP. Conducted all-atom MD simulation: ZK.

## References

- Pandey RB, Kuang Z, Farmer BL, Kim SS, Naik RR (2012) Stability of peptide (P1, P2) binding to a graphene sheet via an all-atom to all-residue coarse-grained approach. *Soft Matter* 8: 9101–9109.
- Tomasio SM, Walsh TR (2009) Modeling the binding affinity of peptides for graphitic surfaces. Influences of aromatic content and interfacial shape. *J Phys Chem* 113: 8778–8785.

3. Whaley SR, English DS, Hu EL, Barbara PF, Belcher AM (2000) Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly. *Nature* 405: 665–668.
4. Sarikaya M, Tamerler C, Jen AKY, Schulten K, Baneyx F (2004) Molecular biomimetics: nanotechnology through biology. *Nat Mater* 2: 577–585.
5. Fang Y, Wu Q, Dickerson MB, Cai Y, Shian S, et al. (2009) Protein-mediated layer-by-layer syntheses of freestanding microscale titania structures with biologically-assembled 3-D morphologies. *Chem Mater* 21: 5704–5710.
6. Pender MJ, Sowards LA, Hartgerink JD, Stone MO, Naik RR (2006) Peptide-Mediated Formation of Single-Wall Carbon Nanotube Composites. *Nano Lett* 6: 40–44.
7. Cui Y, Kim SN, Jones SE, Wissler LL, Naik RR, et al. (2010) Chemical functionalization of graphene enabled by phage displayed peptides. *Nano Lett* 10: 4559–4565.
8. Kim SN, Kuang Z, Slocik JM, Jones SE, Cui Y, et al. (2011) Preferential binding of peptides to graphene edges and planes. *J Am Chem Soc* 133: 14480–14483.
9. Pandey RB, Heinz H, Feng J, Farmer BL, Slocik JM, et al. (2009) Adsorption of peptides (A3, Flg, Pd2, Pd4) on gold and palladium surfaces by a coarse-grained Monte Carlo simulation. *Phys. Chem. Chem. Phys.* 11: 1989–2001.
10. Heinz H, Farmer BL, Pandey RB, Slocik JM, Patnaik SS, et al. (2009) Molecular interactions of peptides with gold, palladium, and Pd-Au bimetal surfaces in aqueous solution. *J. Am. Chem. Soc.* 131: 9704–9714.
11. Tanaka S, Scheraga HA (1976) Medium and long range interaction parameters between amino acids for predicting three dimensional structures of proteins. *Macromolecules* 9: 945–950.
12. Miyazawa S, Jernigan RL (1985) Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules* 18:534–552.
13. Miyazawa S, Jernigan RL (1996) Residue-residue potentials with a favorable contact pair term for simulation and treading. *J Mol Biol* 256: 623–644.
14. Betancourt MR, Thirumalai D (1999) Pair potentials for protein folding: choice of reference states and sensitivity of predicted native states to variations in the interaction schemes. *Protein Sci* 2:361–369.
15. Godzik A, Kolinski A, Skolnick J (1996) Knowledge-based potentials for protein folding: what can we learn from protein structures? *Proteins* 4: 363–366.
16. Huang S-Y, Xiaoqin Z (2011) Statistical mechanics-based method to extract atomic distance-dependent potentials from protein structures. *Proteins* 79:2648–2661.
17. Fritsche M, Pandey RB, Farmer BL, Heermann D (2012) Conformational temperature-dependent behavior of a histone h2ax: A coarse-grained Monte Carlo approach via knowledge-based interaction potentials. *PLoS one* 7: e32075–1 – e32075–8.
18. Pandey RB, Farmer BL (2012) Random coil to globular thermal response of a protein (H3.1) with three knowledge-based coarse-grained potentials. *PLoS one* 7: e49352–1 – e32075–9.
19. Lyubartsev AP, Laaksonen A (1995) Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. *Phys. Rev. E* 52: 3730–3737.
20. Zhou J, Chen S, Jiang S (2003) Orientation of adsorbed antibodies on charged surfaces by computer simulation based on a united-residue model. *Langmuir* 19: 3472–3478.
21. van Giessen AE, Straub JE (2005) Monte Carlo simulations of polyaniline using a reduced model and statistics-based interaction potential. *J. Chem. Phys.* 122: 0249041–0249049.
22. Reith D, M. Putz, F. Muller-Plathe, Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.* 24 (2003) 1624–1636.
23. Pandey RB, Farmer BL (2010) Global structure of a human immunodeficiency virus –1 protease (1DIFA dimer) in an effective solvent medium by a Monte Carlo simulation. *J. Chem. Phys.* 132: 125101–125106.
24. Liwo A, Czaplowski C, Oldziej S, Scheraga HA (2008) Computational techniques for efficient conformational sampling of protein. *Curr. Opin. Struct. Biol.* 18: 134–139.
25. Ercolessi F, Adams J (1994) Interatomic potentials from first-principle calculations: the force-matching method. *Europhys. Lett.* 26: 583–588.
26. Zhou J, Thorpe IF, Izvekov S (2007) Coarse-grained peptide modeling using a systematic multiscale approach. *G.A. Voth, Biophys. J.* 92: 4289–4303.
27. de Jong DH, Singh G, Drew Bennett WF, Arnarez C, Wassenar TA, et al. (2013) *J. Chem. Theory Comput.* 9: 687–697.
28. Sorensen J, Xavier P, Skeby KK, Marrink SJ, Schiott B (2011) Protofibrillar assembly towards the formation of amyloid fibrils. *J. Phys. Chem. Lett.* 2: 2385–2390.
29. Peng LX, Yu L, Howell SB, Gough DA (2011) Aggregation properties of a polymeric anticancer therapeutic: a coarse-grained modeling study. *J. Chem. Info. and Modeling* 51: 3030–3035.
30. Wu Z, Cui Q, Yethiraj A (2011) A new coarse-grained force field for membrane-peptide simulations. *J. Chem. Theory and Compu.* 7: 3793–3802.
31. Sakai VG, Khodadadi S, Cicerone MT, Curtis JE, Sokolov AP, Roh JH (2013) Solvent effects on protein fast dynamics: implications for biopreservation. *Soft Matter* 9: 5336–5340.
32. Miao Y, Yi Z, Glass DC, Hong L, Tyagi M, et al. (2012) Temperature-dependent dynamical transitions of different classes of amino acid residue in a globular protein. *J. Am. Chem. Soc.* 134: 19576–19579.
33. Hong L, Glass DC, Nickels JD, Petricaroli S, Yi Z, et al. (2013) Elastic and conformational softness of a globular protein. *Phys. Rev. Lett.* 110: 028104–1–028104–5.
34. Pandey RB, Farmer BL (2013) Conformational response to solvent interaction and temperature of a protein (histone H3.1) by a coarse-grained Monte Carlo simulation, preprint.