Proteins aggregation and human diseases

Chin-Kun Hu
Institute of Physics of Academia Sinica, Taipei 11529, Taiwan
E-mail: huck@phys.sinica.edu.tw

Abstract. Many human diseases and the death of most supercentenarians are related to protein aggregation. Neurodegenerative diseases include Alzheimer’s disease (AD), Huntington’s disease (HD), Parkinson’s disease (PD), frontotemporal lobar degeneration, etc. Such diseases are due to progressive loss of structure or function of neurons caused by protein aggregation. For example, AD is considered to be related to aggregation of Aβ40 (peptide with 40 amino acids) and Aβ42 (peptide with 42 amino acids) and HD is considered to be related to aggregation of polyQ (polyglutamine) peptides. In this paper, we briefly review our recent discovery of key factors for protein aggregation. We used a lattice model to study the aggregation rates of proteins and found that the probability for a protein sequence to appear in the conformation of the aggregated state can be used to determine the temperature at which proteins can aggregate most quickly. We used molecular dynamics and simple models of polymer chains to study relaxation and aggregation of proteins under various conditions and found that when the bending-angle dependent and torsion-angle dependent interactions are zero or very small, then protein chains tend to aggregate at lower temperatures. All atom models were used to identify a key peptide chain for the aggregation of insulin chains and to find that two polyQ chains prefer anti-parallel conformation. It is pointed out that in many cases, protein aggregation does not result from protein mis-folding. A potential drug from Chinese medicine was found for Alzheimer’s disease.

1. Introduction

Up to 2014, there are 8 Nobel Prize winners in sciences with Chinese background, including Chen-Ning Yang (physics prize in 1957, born in 1922 in Hefei, China), Tsung-Dao Lee (physics prize in 1957, born in 1926 in Shanghai, China), Samuel Chao Chung Ting (physics prize in 1976, born on 27 January 1936 in USA), Yuan-Tseh Lee (chemistry prize in 1986, born on 19 November 1936 in Hsinchu, Taiwan), Steven Chu (physics prize in 1997, born on 28 February 1948 in St. Louis, MO, USA), Daniel Chee Tsui (physics prize in 1998, born on 28 February 1939 in Henan, China), Roger Yonchien Tsien (chemistry prize in 2008, born on 1 February 1952 in New York, USA), and Sir Charles Kuen Kao (physics prize in 2009, born on 4 November 1933 in Shanghai, China). Among these 8 Nobel Prize winners, Sir Charles Kuen Kao was identified to have Alzheimer’s disease (AD) in early 2003. In tradition, every Nobel Prize Winner should give a Nobel lecture. In 2009, Dr. Kao already lost the ability to present a Nobel lecture. Dr. Kao’s friends and family members helped him to prepare and present the Nobel lecture. Professor Steven Chu’s father Ju-Chin Chu (14 December 1919 - 15 November 2000) was an academician of Academia Sinica in Taipei and had AD. Thus among 8 Nobel Prize winners in sciences with Chinese background, two families had been identified with AD. It is quite a high percentage. It has been estimated that 45% of USA citizens with age higher than 85 have AD. AD and other human diseases become more serious in longevity societies.
Alois Alzheimer (14 June 1864 - 19 December 1915) was a German psychiatrist and neuropathologist. In 1901, he worked at Frankfurt Asylum and found that Mrs. Auguste Deter, 51-year-old patient at the Frankfurt Asylum, had strange behavioral symptoms, including a loss of short-term memory. In 1903, Alzheimer moved to Munich to work at Anatomical Laboratory at the Royal Psychiatric Clinic of the University of Munich. In April 1906, Mrs Deter died and her brain was sent to Alzheimer’s lab. in Munich. Alzheimer and two Italian technicians used the staining techniques to identify amyloid plaques and neurofibrillary tangles in Mrs. Auguste Deter’s brain. On 3 November 1906, Alzheimer reported this case study at the 37th meeting of the Society of Southwest German Psychiatrists in Tubingen, Germany the clinical and neuropathological characteristics of the disease [1, 2]. In 1910, Emil Kraepelin called the disease identified by Alois Alzheimer as Alzheimer disease (AD) [3].

The 3 November 2006 issue of Science magazine used AD as the cover story and contains two review papers about AD: (1) “A Century of Alzheimer’s Disease” by Michel Goedert and Maria Grazia [4], (2) “100 Years and Counting: Prospects for Defeating Alzheimer’s Disease” by Erik D. Roberson and Lennart Mucke [5]. It is widely believed that AD is mainly due to aggregation of Aβ40 [6] and Aβ42 [7] in AD patient’s brain and such Aβ40 and Aβ42 form hairpin structures [6, 7].

On 15 February 1872, George Huntington (9 April 1850 - 3 March 1916) read before the Meigs and Mason Academy of Medicine at Middleport, Ohio, USA a paper “On Chorea”. In the paper, he wrote: “Chorea is essentially a disease of the nervous system. The name chorea is given to the disease on account of the dancing propensities of those who are affected by it, and it is a very appropriate designation. The disease, as it is commonly seen, is by no means a dangerous or serious affection, however distressing it may be to the one suffering from it, or to his friends. Its most marked and characteristic feature is a clonic spasm affecting the voluntary muscles” [8]. The disease is then called “Huntington diseases (HD)” [9].

HD, spinobulbar muscular atrophy, dentatorubral-pallidoluysian atrophy, and six kinds of spinocerebellar ataxias are well known as trinucleotide repeat disorder of inherited neurodegenerations. Theses diseases are considered to be caused by abnormal polyglutamine (polyQ) repeat encoded by expanded CAG repeats in genomic sequence [10, 11, 12, 13, 14]. HD is the most common among these diseases which has prevalence estimated about 4 in 10,000 people of Western European ancestry [15].

In 1817, James Parkinson (11 April 1755 - 21 December 1824) published An Essay on the Shaking Palsy [16]. The Essay contains five chapters. In Chapter 1, Parkinson wrote: “Involuntary tremulous motion, with lessened muscular power, in parts not in action and even when supported; with a propensity to bend the trunk forwards, and to pass from a walking to a running pace: the senses and intellects being uninjured. The term Shaking Palsy has been vaguely employed by medical writers in general. By some it has been used to designate ordinary cases of Palsy, in which some slight tremblings have occurred; whilst by others it has been applied to certain anomalous affections, not belonging to Palsy.” The disease described by James Parkinson in [16] was later known as Parkinson disease (PD) [17, 18]. It has been proposed that PD can be related to aggregation of α-synuclein: the protein which has 140 amino acids, and is encoded by the SNCA gene [18].

A supercentenarian (or written as super-centenarian) is someone who has lived to or passed his/her 110th birthday. This age is achieved by about one in 1,000 centenarians (lived to or passed his/her 100th birthday). Senile systemic amyloidosis was determined to be the primary cause of death for 70% of supercentenarians who have been autopsied [19].

Thus it is important to understand the mechanism of protein aggregation. In the present paper, we briefly review the application of statistical physics to understand protein aggregation related to human diseases.

The basic content of statistical physics was formulated from late 19th century to early 20th
century by Ludwig Eduard Boltzmann (20 February 1844 - 5 September, 1906) [20] and Josiah Willard Gibbs (11 February 1839 - 28 April 1903) [21]. The objective of statistical physics is to understand the properties of a macroscopic system from interactions of the molecules or atoms of the macroscopic system. One important research subfield of statistical physics is the study of phase transitions and critical phenomena [22], for a recent review please read [23]. Universality and scaling are two key concepts developed from the study of phase transitions and critical phenomena [22, 23]. In 1971, Ken G. Wilson developed renormalization group (RG) theory [24, 25, 26] for critical phenomena, which provides a framework for understanding universality and scaling in critical systems, and can be used to calculate critical quantities of critical systems. For this contribution, Ken G. Wilson won the Nobel Prize in physics in 1982 and his Nobel Prize Lecture “The renormalization group and critical phenomena” was published in [27].

In the development of statistical physics, it has often been found that a simple model system can be useful for understanding collective or critical behavior [22, 23] of a complicated system consisting of a large number of atoms or molecules. For example, in 1945 Guggenheim [28] reported that in the $T/T_c$ vs $\rho/\rho_c$ plane ($T$, $T_c$, $\rho$ and $\rho_c$ are temperature, critical temperature, density, and critical density of the substance, respectively) the coexistence curves of $N_e$, $A_e$, $X_e$, $N_2$, $O_2$, CO, and CH$_4$ are very consistent with each other, and near the critical point the order parameter [22] $|\rho - \rho_c| \sim |1 - (T/T_c)|^\beta$, with the critical exponent $\beta = 1/3$. In 1952, Yang and Lee proposed that the critical behavior of the gas-liquid system can be represented by a lattice-gas model [29], which is equivalent to the Ising model [30], in which each atom or molecule on a lattice site is assigned a variable which can be either +1 or -1. In 1995-1996, Blöte and collaborators [31, 32] used Monte Carlo simulations to find that the critical exponent $\beta$ of the spontaneous magnetization and $\nu$ of the correlation length [22] of a three-dimensional (3D) Ising model are 0.3269(6) [32] and 0.6301(8) [31], respectively. In 2009, Sengers and Shanks [33] reviewed the experimental data for liquid-gas critical systems and reported that the order parameter and the correlation length have critical exponents $\beta = 0.3245$ and $\nu = 0.629\pm 0.003$, respectively. In 2012, Watanabe, Ito and Hu [34] used molecular dynamics simulations to find that $\beta$ and $\nu$ of a 3D Lennard-Jones (L-J) model system [35, 36] are 0.3285(7) and 0.63(4), respectively. The values of $\beta$ and $\nu$ obtained for simple model systems reported in Refs. [31, 32, 34] are highly consistent with experimental data reported in [28, 33]. Using Monte Carlo methods [37, 38, 39, 40] and analytic methods [41, 42] it has been found that many critical systems can have very nice universal and scaling behaviors [43, 44, 45, 46, 47, 48, 49, 50].

On the basis of molecular biology, it is of interest to know whether we can apply ideas and methods of statistical physics to understand protein aggregation related to human diseases. The purpose of this paper is to address this problem. In the next section, we first review the usage of lattice models to understand some factors which influence protein aggregation. In section 3, we review the application of an off-model of polymer chains to understand the influence of bending-angle dependent and torsion-angle dependent interactions on the aggregation of proteins. In section 4, we review some developments in the application of all-atom models to understand some problems related to protein aggregation. In section 5, we report the discovery of a potential drug for AD. Finally, in section 6, we summarize and discuss our result.

2. Lattice models for aggregation of proteins

Lattice Ising model has been used to describe critical phenomena of gas-liquid systems very successfully [23, 31, 32, 33, 51]. Thus one would like to use simple lattice models to understand folding and aggregation of proteins, which are polymer. As the temperature of the dilute polymer solution decreases, the polymer will undergo a collapse transition [52, 53]. The interacting self-avoiding walk (ISAW) [54, 55, 56, 57, 58, 59] on the simple cubic (SC) lattice is the simplest lattice model for homogenous polymers in dilute solution. The model has a collapse transition at a higher temperature and a freezing transition at a low temperature.
Figure 1. (Color online) (a) Spectrum of energies and associated low energy structures of the monomer sequence +HHPPHH-. H, P, + and - are in green, yellow, blue, and red, respectively. We set $E_{HH} = -1$ and $E_{+} = -1.4$. There are a total 1831 possible conformations that are spread among 17 possible energy values. The conformations in the first excited state represent the ensemble of $N^*$ structures and the $N^*$ conformation that coincides with the peptide state in the fibril [see Figure 2(a)] is enclosed in a black colored box. (b) The probability $P_{N^*}$ of populating the structure in the box in (a) as a function of $T$ for $E_{+} = 0, -0.3, 0.6, -1$ and -1.4 keeping $E_{HH} = -1$. The arrow indicates $T^*$, where $P_{N^*} = P_{N^*}^{max}$. Dependence of $P_{N^*}^{max}$ on $E_{+}$ for $E_{HH} = -1$ (c), and on $E_{HH}$ for $E_{+} = -1.4$ (d). Reproduced from Fig. 1 of [60].

Chen, Hsieh, and Hu (CHH) [59] developed a novel method to analyze phase behavior of lattice models based on partition function zeros by considering both loci of partition function zeros and thermodynamical functions associated with them. With this method, the first pair of complex conjugate zeros (first zeros) can be defined without ambiguity and the critical point of a small system can be defined as the peak position of the heat capacity component associated with the first zeros. For the system with two phase transitions, two pairs of first zeros corresponding to two phase transitions can be identified and two overlapping phase transitions can be well separated. This method was applied to the interacting self-avoiding walk (ISAW) of homopolymer with $N$ monomers on the simple cubic lattice. The exact partition functions
Figure 2. (Color online) (a) The lowest energy fibril structure for $E_{+-} = -1.4$ and $E_{HH} = -1$. (b) Same as in (a) but with $E_{+-} = 0$. (c) Double layer structure for $E_{HH} = -0.4$ but keeping $E_{+-} = -1.4$. (d) For $E_{+-} = -1.4$ and $E_{HH} = -0.3$ the fibril structure is entirely altered. (e) Temperature dependence of $\tau_{fib}$ for $E_{+-} = -1.4$ (circles) and $E_{+-} = -0.6$ (triangles). $N = 6$ and $E_{HH} = -1$. Arrows show the temperatures at which the fibril formation is fastest. Reproduced from Fig. 2 of [60].

$Z_N$ with $N$ up to 27 were calculated and CHH approach gives a clear scenario for the collapse and the freezing transitions [59]. CHH’s conformation number for the ISAW on the $3 \times 3 \times 3$ SC lattice is consistent with that of [54], which only studied the ISAW on the $3 \times 3 \times 3$ SC lattice.

In 2010, Li et al. studied a lattice model for protein aggregation [60]. In this model, an amino acid is represented by a bead on a site of a simple cubic lattice. The bead can be H (hydrophobic), P (polar), +1 (positively charged), or -1 (negatively charged) with interaction parameters: $E(PP) = E(P+) = E(P-) = -0.2$, $E(PH) = E(H+) = E(H-) = 0.2$, $E(+-)$ and $E(HH)$ are varying parameters, $E(++) = E(--) = -E(+-)/2$. $E(+-)$ is between -1.4 and 0 and $E(HH)$ is between -1 and 0. Each chain has $M = 8$ beads and there are $4^M = 65536$ different chains. Among such chains, 5950 chains have a unique ground state and are protein-like. For $E(+-) = -1.4$ and $E(HH) = -1$, the ground state, the first, the second, and the third
excited states are shown in Figure 1(a). When 6 peptide chains were used to simulate aggregation of such chains, one obtained stable conformations of Figures 2(a), 2(b), 2(c), and 2(d) with parameters (a) $E(+-) = -1.4$ and $E(HH) = -1$, (b) $E(+-) = 0$ and $E(HH) = -1$, (c) $E(+-) = -1.4$ and $E(HH) = -0.4$, and (d) $E(+-) = -1.4$ and $E(HH) = -0.3$, respectively. Please note that for $E(+-) = -1.4$ and $E(HH) = -1$, the system used the hairpin structure of the first excited state in Figure 1(a), called $N^*$, to form the aggregated state of Figure 2(a).

The number of conformations of a sequence of $M = 8$ beads, that satisfy self-avoidance is only 1831. The probability $P_{N^*}$ of populating the structure in the box in Figure 1(a) as a function of $T$ for $E_{+-} = 0$, -0.3, -0.6, -1, and -1.4 keeping $E_{HH} = -1$ had been calculated exactly and plotted in Fig. 1(b), in which arrow indicates $T^*$, where $P_{N^*} = P_{N^*}^{max}$. Dependence of $P_{N^*}^{max}$ on $E_{+-}$ for $E_{HH} = -1$, and on $E_{HH}$ for $E_{+-} = -1.4$ were shown in Figures 1(c) and 1(d), respectively.
Monte Carlo (MC) method was used to calculate the temperature dependence of the fibril formation time $\tau_{\text{fib}}$ for $E_{+-} = -1.4$ (circles) and $E_{+-} = -0.6$ (triangles) with $N = 6$ and $E_{HH} = -1$, and calculated results were shown in Figure 2(e), in which arrows show the temperatures at which the fibril formation is fastest. Figure 2(b) shows that for $E_{+-} = -1.4$ and $E_{HH} = -1$, $P_N$ has a maximum at $T = 0.55$, at the same temperature the fibril formation is fastest with the same parameters $E_{+-} = -1.4$ and $E_{HH} = -1$ in Figure 2(e). Thus Li et al. [60] found that the probability of a protein sequence appearing in an aggregated conformation can be used to determine the temperature at which the protein can aggregate most quickly. They also found a correlation between the time taken for the protein to aggregate and the strength of interactions between charged amino acids, which is consistent with previous experimental observations [61].

Co, Hu, and Li used the model of [60] to develop a lattice model for studying the aggregation of polypeptide chains in the presence of crowders [62]. Figure 3(a) shows an example of random conformations of peptide chains in a confined box. Figure 3(b) shows a snapshot of crowders (grey cubes) and polypeptide chains (color). The peptide chains can not enter the lattice sites occupied by crowders and there is no other interactions between peptide chains and crowders. Monte Carlo method was used to study the influence of various factors on the fibril formation time $\tau_{\text{fib}}$. One MC step consists of 9 local moves of peptide chains and one global move of a chain; 5 MC moves follow by one global move of all crowders.

Figure 3(c) shows the dependence of $\tau_{\text{fib}}$ on the crowder concentration $\phi_c$ for various sizes of crowders. Orange and red refer to non-cubic (rectangular parallelepiped) crowders with dimensions $2 \times 1 \times 1$ and $2 \times 2 \times 1$. Other crowder sizes are also shown. The numbers of polypeptide chains $N$ are 6, 10, and 24 for top, middle, and bottom panels, respectively. $\tau_{\text{fib}}$ of Fig. 3(c) decreases with the concentration of crowders if crowder sizes are large enough, but the growth is observed for crowders of small sizes. Figure 3(d) shows the dependence of $\tau_{\text{fib}}$ on the total surface area of crowders at three values of crowder concentration $\phi_c$ with $T = 0.65$, 0.69, and 0.78 for the number of chains $N = 6$, 10, and 24, respectively. The results are averaged over 50 to 200 MD trajectories for each set of parameters. Such results can be used to explain the recent experimental observation on the dual effect of crowding particles on fibril growth of proteins that for a fixed crowder concentration the fibrillation kinetics is fastest at intermediate values of total surface of crowders [63]. It becomes slow at either small or large coverage of cosolutes.

Figure 3(e) shows the U-shape dependence of $\tau_{\text{fib}}$ on the sizes of confining cubic box. The results obtained from 50 to 200 MC trajectories for each box size. Error bars come from averaging over MC runs. Due to competition between the energetics and entropic effects, the dependence of $\tau_{\text{fib}}$ on the size of confined space is described by a parabolic function [62].

3. Relaxation and aggregation of polymer chains

Besides protein aggregation mentioned in Section 1, there is another interesting biological problem: How a biological system can maintain in non-equilibrium state? In particular, why ancient seeds can maintain in a viable non-equilibrium state for about 2000 years [64] or even about 30,000 years [65]? This section will review some clues for answering these questions [51].

Many important biological macromolecules, such as DNA, RNA, proteins, etc, are polymers. To understand why a biological system can maintain in a non-equilibrium state for a very long time and why proteins aggregate, one can try to get relevant information from the relaxation and aggregation of a simple polymer model. In this section, we briefly review the work by Ma and Hu [66, 67, 68, 69, 70] about relaxation and aggregation of polymer chains.
Figure 4. (Color online) Physical quantities versus time $t$ in the unit of $\tau$. (a) Temperature-like quantities $T_0^*$ (red line), $T_P^*$ (black line) and $T_F^*$ (blue line) in unit $\epsilon$ for the whole system, for the polymer sites and for the fluid particles, respectively, in a single step; (b) $q$; (c) $\beta^{-1}$; (d) bond-pair correlation and (e) local helicity-pair correlation, for index distance $k = 1, 2, 4, 8$, in purple, green, blue and red colors, respectively. Data are plotted every 100 steps for (a), and 10000 steps for (b), (c), (d) and (e). The vertical dotted lines mark the change in integration time step $\delta t$ from $(\delta t)_1 \equiv 5.0 \times 10^{-6}\tau$ to $(\delta t)_2 \equiv 2.5 \times 10^{-5}\tau$ at time $t = 13.4\tau$, dividing the relaxation stage into $R_1$ and $R_2$; the intermediate transients, $G_1$ (from $t = 18.4750$ to $t = 29.2925$), $G_2$ ($t = 81.6725$ to $t = 89.9700$) and $G_3$ ($t = 169.315$ to $t = 176.650$) that are followed by the quasi-static stages $S_1$, $S_2$ and $S_3$, respectively. In $S_3$, the integration step is reduced to $(\delta t)_3 \equiv 5 \times 10^{-7}\tau$. The velocity distributions at those steps marked by inverted triangles on top of plot (a) are analyzed in Fig. 5. Reproduced from Fig. 1 of [66].

3.1. Model System with monomers connected by rigid bonds

In this subsection, we consider that polymer molecules are chains of rigidly connected monomers, interacting via L-J potentials between non-bonded monomer pairs and bending and torsion potentials among consecutive sites along the chains. The polymers were modified from the united atom model [71] where the bond length and the interaction potentials describe polyethylene. The L-J potential [35, 36] between non-bonded monomer pairs at sites $i$ and $j$ with the distance $r_{ij}$ is

$$v_{ij}(r_{ij}) = 4\epsilon\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6},$$

where $\sigma$ and $\epsilon$ were used as units for length and energy strength, respectively. The bond length between two neighboring monomers in a polymer chain was fixed at $b_0 = 0.357\sigma$.

For consecutive sites $i, j = i + 1, k = i + 2$, and $l = i + 3$ along a polymer chain, the unit bond vector $\vec{b}_i$ defines the direction for two consecutive sites indexed by $i$ and $j = i + 1$ along
Figure 5. (Color online) (a) Velocity distributions and (b) scaled velocity distributions, for the system described in Fig. 4. The curves in (a) are the fitted functions (Eq. (4)) for the simulated data (in symbols) at the twelve time spots marked in Fig. 1(a). The corresponding time \( t \) (in \( \tau \)) and fitted values \( q \) and \( \beta^{-1} \) (Fig. 1(b) and (c)) for those curves are listed. The distributions in (b) are rescaled according to Eq. (6), where the solid line corresponds to \((q, \beta) = (1.174, 1.0)\) and the dotted line represents the Maxwell-Boltzmann (MB) distribution with \((q, \beta) = (1.0, 1.0)\). We also plot the MB distribution with \((q, \beta) = (1.0, 0.554)\) as the dashed line, which has the same second moment \(<\tilde{v}^2>\) (or \(T_p^*)\) as that of solid line. The inset displays the simulation data in log scales. Note that the deviations from the main curves at time \( t = 23.9, 85.9 \) and 172.9 are contributed by the locally agitated events over the transient periods \( G_1, G_2 \) and \( G_3 \), respectively. Reproduced from Fig. 2 of [66].

the chain. The unit local helicity vector \( \tilde{u}_i \) is defined by

\[
\tilde{u}_i = \frac{\tilde{b}_i \times \tilde{b}_{i+1}}{|\tilde{b}_i \times \tilde{b}_{i+1}|}.
\]

The bending and torsion angles are determined by

\[
\cos \theta_{ijk} = \tilde{b}_i \cdot \tilde{b}_j, \quad \& \quad \cos \phi_{ijkl} = \tilde{u}_i \cdot \tilde{u}_j,
\]

respectively. \( 180^\circ - \theta_{ijk} \) and \( \phi_{ijkl} \) are, respectively, the valence angle and torsional angle considered in [72]. We set the bending potential
and the torsional potential

$$\Theta(\theta_{ijk}) = K_b 1097\epsilon (\cos \theta_{ijk} - \cos \theta_0)^2,$$

(2)

and the potential

$$\Phi(\phi_{ijkl}) = K_t \sum_{i=0}^{3} a_i (\cos \phi_{ijkl})^i,$$

(3)

with $a_0 = 18.635\epsilon$, $a_1 = 38.16\epsilon$, $a_2 = 10.3\epsilon$ and $a_3 = -67.075\epsilon$, as in [71], $\theta_0$ is either $\theta_0^{(I)} = 112.813^\circ$ or $\theta_0^{(II)} = 67.187^\circ$. We considered $K_b$ and $K_t$ and $\theta_0$ as the tuning parameters related to the polymer properties at longer time and length scales. A cutoff of $2.3\sigma$ was used for pair L-J interactions.

The length $\sigma_F$ and strength $\epsilon_F$ of the L-J interactions between fluid molecules were expressed in units of $\sigma$ and $\epsilon$, respectively. All fluid molecules and monomers were assumed to have the same mass $m$. The composite quantity $\tau = \sqrt{m\sigma^2/\epsilon}$ was used as time unit. The purely repulsive interaction potential between a monomer and a fluid molecule takes only the -12 power term in Eq. (1). In such homopolymer system, the thermodynamic conditions were changed by varying the volume (concentration) and the temperature. Heterogeneous conditions were introduced into the system by selecting a small portion of monomer sites, imposing different inter-site interactions. By mixing the polymers with spherical Lennard-Jones fluid molecules, we let the interaction potentials of fluid particles with the ordinary polymer sites be purely repulsive and with that small portions of specially selected sites (‘linker sites’) on polymers be the same as fluid-fluid potentials. Such setup mimics heterogeneities in the hydrophobic-hydrophilic effects in systems of biopolymers and help to stabilize the system. The shape and the strengths of all those interaction potentials are tunable as the means to alter the physical properties of the simulated system.

It is well known that the dynamic properties are related to the detailed numerical implements in solving the equations of motion in MD simulations [73]. In the simplified model, it should be emphasized that the well-tracked formulation of the discretized equations of motion that are treated as iterative dynamic systems. This aspect is appropriate in searching for microscopic origins of $q$-statistics [74]. The equations of motions in Lagrangian formulations were numerically solved using a velocity Verlet algorithm; the bonding between neighboring monomers was realized via the RATTLE scheme [75] to include the holonomic constraints for fixed bond lengths.

### 3.2. Relaxation of polymers with monomers in a chain connected by rigid bonds

We had chosen $K_b = K_t = 0.1$, corresponding to one tenth of those for polyethylene, so that the simulations can cover longer range of time scales. To start, we filled $N_P = 40$ polymer chains (each chain has $n = 100$ monomers) and $N_F/4 = 1500$ fluid molecules in a large enough cube of volume $V = L^3$ with periodic boundary conditions so that polymer chains can take some arbitrary random initial configurations. We took $\theta_0$ in Eq. (2) to be $\theta_0^{(I)} = 112.813^\circ$ in all chains, selected $\sigma_F = \sigma/4$ and let a small $N_L (=199)$ randomly selected ‘linker sites’ to have hydrophilic-like contacts with fluid particles.

We controlled the pressure and the kinetic energy [76] to adjust the volume $V$ until it reached some value $V \approx (32.5\sigma)^3$. Then we replaced each fluid molecule with four closely packed molecules of zero kinetic energies and allowed the new system (of $N_P$ polymers and $N_F$ fluid molecules) to relax and reached finally a volume $V = L^3$ with $L = 33.3875\sigma$.

The NPT processes designed for equilibrium simulations [76] had been used in this preparation stage, where we controlled externally the pressure ($P$) and the temperature ($T$) to drive the system of fixed number of particles ($N$) to the target density. The process also fine-tuned the conformations of the polymeric chains that helped the effective performance of RATTLE-algorithm in the subsequent NVE simulations.
We then started the production simulations with the integration time step $(\delta t)_1 \equiv 5.0 \times 10^{-6} \tau$ and focus on the NVE relaxation processes with constant volume $V$ and total energy $E$ without any external controlling. To speed up the simulation process, we increased the integration time step $\delta t$ from $(\delta t)_1 \equiv 5.0 \times 10^{-6} \tau$ by five times to $(\delta t)_2 \equiv 2.5 \times 10^{-5} \tau$ at time $t = 13.4 \tau$.

The production simulations from $t = 0$ continued for about 180 $\tau$ (at the final few $\tau$ the integration time step was reduced, see below). The results were shown in Figs. 1 and 2 of [66], reproduced as Figs. 4 and 5 in the present paper. Figure 4(a) shows that the time evolution of the temperature-like quantities, $T_p^0$, $T_p^1$ and $T_0^0$, consist of several stages, where $T_p^0$, $T_p^1$, and $T_0^0$ are twice of the mean kinetic energies per degrees of freedom for polymer chains, fluid molecules, and the whole system, respectively. For polymer chains, $T_p^0 \equiv f^{-1} \sum_{\text{site}_i} m_i v_i^2$, where $f = N_P(2n + 1)$ is total number of degrees of freedom for polymer chains when $(n - 1)$ holonomic constraints for each polymer chain were taken into account.

In the first stage $R_1$, the local conformation relaxes with decreasing correlations in bond-pairs (Fig. 4(d)) and local helicity pairs (Fig. 4(e)). We used inner product $\vec{b}_i \cdot \vec{b}_{i+k}$ and $\vec{u}_i \cdot \vec{u}_{i+k}$ for indices $i$ and $i + k$ along a chain to define the correlations of index distance $k$.

The increase of the integration time step at $t = 13.4 \tau$ would certainly affect the route of time evolution at microscopic level. Our focus here is whether the value of $q$ changes. The subsequent relaxation stage $R_2$ continues before the system becomes quasi-static with $T_p^1 \approx T_p^0$. The system is found subject to intermediate transients before and after reaching a quasi-static stage; there are three quasi-static stages, $S_1$, $S_2$ and $S_3$, following the transient stages $G_1$, $G_2$ and $G_3$, respectively. Entering the static stage $S_3$, we reduced the integration time step by a factor fifty to $(\delta t)_3 \equiv 5.0 \times 10^{-7} \tau$, to understand its effect on $q$. From simulation data, we estimated the diffusion constant for monomers as $D \approx 3.3e^2/\tau$. The time spans of $S_1$ and $S_2$ are large enough for a monomer to cross a distance ranging from $L/2$ to $L$.

The vertical dotted lines in Fig. 4 mark the change in integration time step $\delta t$ from $(\delta t)_1 \equiv 5.0 \times 10^{-6} \tau$ to $(\delta t)_2 \equiv 2.5 \times 10^{-5} \tau$ at time $t = 13.4 \tau$, dividing the relaxation stage into $R_1$ and $R_2$: the intermediate transients, $G_1$ (from $t = 18.4750$ to $t = 29.2925$), $G_2$ ($t = 81.6725$ to $t = 89.9700$) and $G_3$ ($t = 169.315$ to $t = 176.650$) that are followed by the quasi-static stages $S_1$, $S_2$ and $S_3$, respectively. The velocity distributions at those steps marked by inverted triangles on top of Fig. 4(a) were analyzed in Fig. 5.

Figure 5(a) shows that the velocity distributions of monomers at several typical time steps (marked by inverted triangles at the top line of Fig. 4) can be well described by Tsallis $q$-statistics [77] with a generalized Maxwell-Boltzmann distribution (GMBD) [78, 79, 80]

$$4\pi v^2 F(v) = A_q \beta^{3/2} 4\pi v^2 \times [e_q^{-\frac{\beta}{2} mv^2}]$$

(4)

where $e_q^x \equiv [1 + (1 - q)x^{1-q}]^{1/(1-q)}$, $A_q = \int_0^{\infty} 4\pi v^2 e_q^{-\frac{\beta}{2} mv^2} dv^{-1}$, and $1 \leq q < 7/5$ [66]. In the limit $q \rightarrow 1$, Eq. (4) becomes the Maxwell-Boltzmann (MB) distribution

$$4\pi v^2 \left(\frac{m \beta}{2\pi} \right)^{\frac{3}{2}} e^{-\frac{\beta}{2} mv^2}.$$  

(5)

The parameter $\beta$ in Eq. (4) can be absorbed to scale the velocity variable $v$,

$$\beta^{-\frac{1}{2}} 4\pi v^2 F(v) = A_q 4\pi v^2 e_q^{-\frac{\beta}{2} v^2},$$

(6)

where $\tilde{v} = \beta^\frac{1}{2} v$ and the right-hand side (RHS) of Eq. (6) is free of $\beta$, i.e., a family of distributions with different $\beta$ but same $q$ can collapse into one master curve.

Figure 5(b) shows that the velocity distributions in Fig. 5(a) collapse into a single curve, after being scaled by using Eq. (6). The parameter $\beta^{-1}$ increases linearly with the instantaneous
Figure 6. Snapshots of the configurations before (a) and after (b) the formation of polymer aggregation, during the phase separation process for a system with $N_P = 40$ chains, each with $n = 100$ rigidly bonded monomers, mixed with a fluid of $N_F = 6000$ spherical atoms (shown as grey dots). The monomers along a chain are also subject to bending and torsion (intra-chain) potentials in very weak strengths. Periodic boundary condition is applied to each of the three directions for the $33.3876\sigma \times 33.3876\sigma \times 33.3876\sigma$ cell. The instantaneous temperatures for polymer and fluid are $T^*_P = 18.21$ and $T^*_F = 19.95$ for (a); and $T^*_P = 6.335$ and $T^*_F = 8.24$ for (b), respectively. Reproduced from Fig. 1 of [68].

The temperature of monomers $T^*_P$ [66]. The result in Fig. 5(b) suggests that Eq. (4) should not be considered merely as a class of fitting functions, it relates the microscopic states over a wide time span as a single non equilibrium state characterized by the family of distributions of virtually the same $q$. Indeed, the reducing of integration time step $\delta t$ can change the heights and locations of regions $G_1$, $G_2$ and $G_3$ in Fig. 4(a), but does not change $q$, see e.g. $S_3$ in Fig. 4(a) and 4(b). As for the physical origin of such GMBD, it is useful to consider the thought of Beck and Cohen [81].

Note that each distribution in Fig. 5 is over the monomers in the system for one single step. If it is also a probability distribution for an individual monomer, then the $x$-component of the velocity of particle $i$ satisfies

$$< v_{ix}^2 > = C_{q, \beta} \int v_{ix}^2 e^{-\beta K} \prod_{j=1}^{N} dv_{jx} dv_{jy} dv_{jz},$$

where $K$ is the total kinetic energy of the polymers. Using $(e^{x^2}_q)' = e^{q x^2}_q$ or $(e^x_q)' = \frac{e^x_q}{1+(1-q)x}$ and integration by parts, we derived that the mean kinetic energy of a monomer is

$$< \frac{1}{2}mv^2 > = \frac{3}{2(2-q)} \frac{1}{\beta} - (1-q) < K >.$$

Equation (7) states a modified version of equipartition theorem; it becomes the conventional equipartition theorem only when $q = 1$. The conditions for the validity of Eq. (7), especially the role of heterogeneity [82], deserves further analysis.

3.3. Aggregation of polymer chains connected by rigid-bonds
We take a system at the end of simulation of Fig. 4 ($t = 180\tau$) as the initial system and reduce $K_b$ and $K_t$ to $1.0 \times 10^{-4}$. With quenching and heating alternatively, each for about $50\tau$, we let
Figure 7. Snapshots before (a) and after (b) the formation of aggregation, for a system similar to that in Fig. 6, except the fluid atoms are absent and the polymer sites are identical and subject to neither bending nor torsion potentials. The instantaneous temperature is $T_P^* = 26.78$ for (a) and $T_P^* = 6.78$ for (b), respectively. Reproduced from Fig. 2 of [68].

Figure 8. Snapshots before (a) and after (b) the formation of aggregation, for a system similar to that in Fig. 7, but the polymer chains are subject to the same bending and torsion angle potentials as those in the mixed system plotted in Fig. 6. The polymer temperature for (a) is $T_P^* = 21.7$ and for (b) is $T_P^* = 6.42$. Reproduced from Fig. 3 of [68].

the system free from possible residual local structures favored by the original angle potentials. We then let the system ready to relax. We also consider the case of pure polymers with zero ($K_b = K_t=0$) or very small ($K_b = K_t = 10^{-4}$) angle potentials in absence of fluid. It is found that, in all cases, very small or zero angle potentials render the systems to undergo slow quenching processes, as a dynamical system driven by a low temperature attractor caused by the numerical sink. Such scenario is also true for other cases considered in this subsection, as soon as the strengths of angle potentials are tiny.

Since integration of time step $\delta t$ is relevant in determining the time evolution of the aggregation process, we will consistently use two major integration time steps $(\delta t)_2 = 2.5 \times 10^{-5}$ and $(\delta t)_4 = 1 \times 10^{-4}$. Over the beginning a change (on the strengths of angle potential, or on the lengths of polymer chains, etc...) being made, we use the smaller one $(\delta t)_2$ and turn to the large
one afterward. While the quenching rate depends on the choice of $\delta t$, the features displayed on the aggregated states would not.

We examine the snapshots of the three cases: the mixed system of polymers and fluids (Fig. 6), the pure system where the monomers at linker sites are switched into ordinary monomers and $K_b$ and $K_t$ are set to 0 (Fig. 7), and the pure system where the monomers at linker sites are switched into ordinary monomers and $K_b$ and $K_t$ are kept at $10^{-4}$ (Fig. 8) [68]. In all of three cases, polymer chains aggregate as $K_b$ and $K_t$ become 0 or very small.

3.4. Polymer chains with monomers connected by springs

To check whether the obtain $q$-statistics with $q > 1$ is an artifact due to the RATTLE scheme [75] or rigid-bond constraints between neighboring monomers, we used the method of ordinary MD simulations to study the relaxation process in a system of polymer chains and Lennard-Jones molecules, in which each polymer chain consists of $n$ monomers labelled by integer $i = 1, \ldots, n$ from one end to the other; the monomers at the sites $i$ and $j = i + 1$ ($1 \leq i < n$) are connected by harmonic potentials

$$12k_{spring}(r_{ij} - r_0)^2.$$ (8)

The coupling constant $k_{spring}$ is $k_s = 10^s \times k_0$, with $s = 0, 1, 2, 3, 4$ for the five systems, respectively. Please note that the interactions between neighboring monomers of the form of Eq. (8) has been used in the Go-like model for proteins [83] and also in a RNA model [84]. L-J interactions, and bending-angle and torsional-angle dependent interactions are the same as those described in previous subsections.

The instantaneous temperatures of the monomers in polymer chains, $T_p^*$, and the fluid, $T_F^*$, were initially different, but they vary in time and their ratio $\Gamma^* = T_p^*/T_F^*$ approached to a constant during the relaxation process. The velocity distributions of monomers in the relaxation process follow $q$-statistics [77] with $q \geq 1$. The value of $q$ and the limiting ratio of $\Gamma^*$ depend on $k_{spring}$; in the strong strength limit, they approach those for the system, in which each polymer chain consists of monomers connected by rigid bonds; in the weak strength limit, $\Gamma^*$ and $q$ approach 1 corresponding to MB distribution for velocities of monomers. Typical simulation results are shown in Figs. 5, 6 and 7 of [67]. Such results indicate that there is a continuous increase of $q$ as one increases the strength of springs (between monomers) from the smaller to the larger values, and the result $q > 1$ in previous subsections is not an artifact of the RATTLE scheme [75] or the rigid-bond constraints.

It was found that when neighboring monomers of the spring-connected polymer chains have small or zero bending-angle and torsion-angle dependent potentials, then the polymer chains also aggregate [69]. When $K_b = K_t = 0.1$, polymer chains do not aggregate even at low temperatures. Thus $K_b$ and $K_t$ play important role for the aggregation of polymer chains.

In summary, the molecular dynamics was used to simulate various systems of polymer chains and L-J molecules; the neighboring monomers along a polymer chain are connected by rigid bonds [66] or spring of strength $k_{spring}$ [67]. It was found that during the relaxation process the velocity distributions of monomers in a wide range of simulation time can be well described by Tsallis $q$-statistics [77] with a generalized Maxwell-Boltzmann distribution (GMBD) [78, 79, 80] and a single scaling function. The value of $q$ is related to the conformation constraining potential, the interactions with background fluid, the destruction of chain homogeneity or the value of $k_{spring}$. Such study is useful for understanding microscopic origin of $q$-statistics [81, 74, 85]. The results suggest that complex non-equilibrium systems can still have simple scaling behavior and inspire many interesting problems for further studies.

The system studied in [66] has slow relaxation, similar to that of a spin-glass model at low temperatures [86]. The glassy behavior has also been found in the collagen fibril [87, 88]. It seems
that the glassy behavior of bio-polymers is the key factor for a biological system to maintain in a non-equilibrium state [23, 51].

It was also found that when neighboring monomers of a polymer chain have small or zero bending-angle and torsion-angle dependent potentials, then the polymer chains tend to aggregate [68, 69, 70]. The aggregation of proteins in human brains can result in neurodegenerative diseases, such as Alzheimer’s disease [89, 6, 7], Huntington’s disease [8], Parkinson’s disease [16, 17], etc. Thus it is of interest to know relaxation and aggregation behaviors of polymer chains in non-equilibrium situation.

4. All atom models for protein aggregation
In this section, we briefly review recent developments for simulations of all atom models for protein aggregation.
4.1. Oligomerization of peptides LVEALYL and RGFFYT and their binding affinity to insulin.

Recently it has been proposed a model for fibrils of human insulin in which the fibril growth proceeds via stacking LVEALYL (fragment 11-17 from chain B of insulin) into pairs of tightly interdigitated β-sheets. The experiments have also shown that LVEALYL has high propensity to self-assembly and binding to insulin. This necessitates study of oligomerization of LVEALYL and its binding affinity to full-length insulin. Using the all-atom simulations with Gromos96 43a1 force field and explicit water it is shown that LVEALYL can aggregate. Theoretical estimation of the binding free energy of LVEALYL to insulin by the molecular mechanic Poisson-Boltzmann surface area method reveals its strong binding affinity to chain B, implying that, in agreement with the experiments, LVEALYL can affect insulin aggregation via binding mechanism. We predict that, similar to LVEALYL, peptide RGFFYT (fragment B22-27) can self-assemble and bind to insulin modulating its fibril growth process. The binding affinity of RGFFYT is shown to be comparable with that of LVEALYL [90].

4.2. MD simulations of one and two polyglutamine peptides in explicit water molecules

Aggregation of polyglutamine peptides with β-sheet structures is related to some important neurodegenerative diseases such as Huntington's disease. However, it is not clear how polyglutamine peptides form the β-sheets and aggregate. To understand this problem, we performed all-atom replica-exchange molecular dynamics simulations of one and two polyglutamine peptides with 10 glutamine residues in explicit water molecules. Conformations and the transformation pathways of the global- and local-minimum free-energy states A to I of the the system with one polyQ-peptide and 1611 water molecules was shown in Fig. 9(a); Conformations and the transformation pathways of the global- and local-minimum free-energy states A to N of the system with two polyQ-peptides and 3129 water molecules was shown in Fig. 9(b).

Our results show that two polyglutamine peptides mainly formed helix or coil structures when they are separated, as in the system with one-polyglutamine peptide. As the interpeptide distance decreases, the intrapeptide β-sheet structure sometimes appear as an intermediate state, and finally the interpeptide β-sheets are formed. We also find that the polyglutamine dimer tends to form the antiparallel β-sheet conformations rather than the parallel β-sheet, which is consistent with previous experiments and a coarse-grained molecular dynamics simulation [15].

Table 1 shows that in the one polyQ-peptide system, the random-coil has the highest probability of 64.9±8.9, the helix has the second probability of 29.4±4.6, and intra-β-sheet has probability of only 6.3±5.1 Table 2 shows that in the two polyQ-peptide system, inter-β-sheet has the highest probability of 50.5±10.2, random-coil has the probability of 25.3±10.7, the helix has the probability of 23.1±3.2, and intra-β-sheet has the probability of 11.8±5.7. Thus the most stable conformation of two polyQ-peptide system, i.e. inter-β-sheet, is different from the most stable conformation of one polyQ-peptide system, i.e. the random-coil.
### Table 2. Probability and free energy difference of each secondary structure in the two polyQ-peptide system.

| structure     | probability(%) | $\Delta F$ (kcal/mol) |
|---------------|---------------|-----------------------|
| helix         | 23.1±3.2      | 0.47±0.15             |
| inter-β-sheet | 50.5±10.2     | -                     |
| intra-β-sheet | 11.8±5.7      | 0.87±0.31             |
| random-coil   | 25.3±10.7     | 0.41±0.28             |

4.3. MD approach to effect of Taiwan mutation (D7H) on structures of amyloid peptides

Recent experiments have shown that the Taiwan mutation (D7H) slows the fibril formation of amyloid peptides Aβ40 and Aβ42 [92]. Motivated by this finding, we have studied the influence of D7H mutation on structures of Aβ peptide monomers using the replica exchange molecular dynamics simulations with OPLS force field and implicit water model. Our study reveals that the mechanism behind modulation of aggregation rates is associated with decrease of $\beta$-content and dynamics of the salt bridge D23-K28. Estimating the bending free energy of this salt bridge, we have found that, in agreement with the experiments, the fibril formation rate of both peptides Aβ40 and Aβ42 is reduced about two times by mutation [93].

5. Searching potential drugs for AD

From database of about 32364 natural compounds [94], the Lipinskis rule of five [95] was applied to reduce the number of compounds from 32364 to 3699 which was further studied by the docking method.

In order to study the binding affinity to mature fibrils, two typical structures of 6Aβ39-40 obtained from Prof. R. Tycko [96] and 5Aβ17-42 (PDB ID: 2BEG [97]) were used.

A ranking of ligands by their binding energies to two receptors [96, 97] was made and the 10 top leads were listed on Table S1 of the Supporting Information (SI) in [98]. Among such 10 top leads only Dihydrochalcone could be purchased in the market.

The binding affinity of this ligand (Dihydrochalcone) to amyloid beta (Aβ) fibril has been thoroughly studied by computer simulation and experiment. Using the Thioflavin T (ThT) assay we had obtained the inhibition constant IC50≈ 2.46$\mu$M. This result is in good agreement with the estimation of the binding free energy obtained by the molecular mechanic-Poisson Boltzmann surface area method and all-atom simulation with the force field CHARMM 27 and water model TIP3P. Cell viability assays indicated that Dihydrochalcone could effectively reduce the cytotoxicity induced by Aβ. Thus, both in silico and in vitro studies show that Dihydrochalcone is a potential drug for the Alzheimers disease [98].

Dihydrochalcone is a compound from Chinese medicine Xuejie (English Dragon’s Blood; Latin Draconis Sanguis). According to Chinese historical record, Xuejie was already used as a medicine in Ming Dynasty. Dihydrochalcone is also the active ingredient of a traditional herbal tea in Africa.

6. Summary and discussion

It is pointed out in Section 1 that universality and scaling are key concepts developed from the study of phase transitions and critical phenomena in physical critical systems. Using computing packages developed by us [99], we have found a universal volume to surface area for proteins in Protein Data Bank [100]. We have used simple lattice models and polymer chain model to understand some key factors for protein aggregation. In the next step, it is of interest to develop universal theory for protein aggregation.
Figures 4 and 5 show that velocity distributions of monomers in polymer chains deviate from Maxwell-Boltzmann distribution. Since the data of Figures 4 and 5 were obtained by molecular dynamics simulation and the simulated systems still do not reach equilibrium. It is of interest to use other simulation methods, such as parallel tempering which can allow the simulated systems to reach equilibrium more quickly, to study the same polymer systems to check whether in such cases the velocity distributions of monomers in polymer chains still deviate from Maxwell-Boltzmann distribution.

It had been proposed that protein aggregation is due to protein misfolding [101, 102]. However, Figures 1(a) and 2(a) show that peptide chains aggregate with the conformation of the first excited state of the one chain system rather than the ground state conformation. Figure 9, Table 1, and Table 2 show that most stable conformation of one chain system is different from the most stable conformation of two-chains system. In such cases, protein aggregation is not due to protein mis-folding.

We have found Dihydrochalcone as a potential drug for AD [98] from in silico and in vitro study. It is of interest to use animal model to study the effect of Dihydrochalcone to prevent AD. Table S1 in SI of [98] list other 9 compounds besides Dihydrochalcone. These compounds are also potential drugs for AD. It is of interest to study the effect of such compounds to reduce protein aggregation via in vitro and in vivo study.

7. Acknowledgments
The author would like to thank collaborators, M. S. Li, N. T. Co, G. Reddy, J. E. Straub, D. Thirumalai, W.-J. Ma, H. -L. Chiang, C. -J. Chen, M. H. Viet, C.-Y. Chen, Y.-R. Chen, H. Okumura, P. M. Truong, M. H. Viet, P. H. Nguyenand, et al. for the work reported in this brief review paper. The author would like to thank Wei Li for inviting him to give a talk at The 4th International Workshop on Statistical Physics and Mathematics for Complex Systems (SPMCS2014), in Yichang, China on 13-16 October 2014; this paper was based on the talk presented at SPMCS2014. This work was supported by Grant MOST 103-2112-M-001 -016.

References
[1] Alzheimer A 1907 Allg. Z. Psychiatr. 64 146
[2] Alzheimer A 1911 Z. Ges. Neurol. Psychiat. 4 356
[3] Kraepelin E 1910 Psychiatrie. Ein Lehrbuch fur Studierende und AL rzte. II. Band (Barth Verlag, Leipzig)
[4] Goedert M and Grazi M 2006 Science 314 777-781
[5] Roberson ED and Mucke L 2006 Science 314 781-784
[6] Petkova AT, et al. 2002 Proc. Natl. Acad. Sci. 99 16742
[7] Luhrs T, et al. 2005 Proc. Natl. Acad. Sci. 102 17342
[8] Huntington G 1872 Medical and Surgical Reporter of Philadelphia 26 317-321
[9] MacDonald ME, Gusella JF 1996 Curr. Opin. Neurobiol. 6 638
[10] Koshy BT, Zoghbi HY 1997 Brain Pathol. 7 927
[11] Kakizuka A 1998 Trends Genet. 14 396
[12] Perutz MF 1999 Trends Biochem. Sci. 24 58
[13] Paulson HL 1999 Am. J. Hum. Genet. 64 339
[14] Riley BE and Orr HT 2006 Genes Dev. 20 2183
[15] Chiang HL, Chen CJ, Okumura H and Hu CK 2014 J. Comp. Chem. 35 1430-1437
[16] Parkinson J 1817 An Essay on the Shaking Palsy (Sherwood, Neely, and Jones, London)
[17] Polymerepoulos MH, et al. 1997 Science 276 2045-2047
[18] Irvine GB, El-Agnaf OM, Shankar GM and Walsh DM 2008 Mol. Med. 14 451-464
[19] Leslie M 2008 Science 321 1764-1765
[20] Boltzmann L. (1896, 1898) Lectures on gas theory translated by Stephen G. Brush (1964) Berkeley: University of California Press; (1995) New York: Dover
[21] Gibbs JW 1902 Elementary principles in statistical mechanics (New York), reprinted (Woodbridge, CT, Ox Bow Press, 1981)
[22] Stanley HE 1971 Introduction to Phase Transitions and Critical Phenomena (Oxford Univ. Press, New York)
[23] Hu CK 2014 Chinese J. Phys. 52 1-76
Lima JAS and Silva R 2005 Physica A J. Stat. Phys.
Tsallis C 1988 Computer Simulation of Liquids (Oxford Univ. Press, Oxford, 1987).
Guggenheim EA 1945 J. Chem. Phys. 13 253
Yang CN and Lee TD 1952 Phys. Rev. 87 404; Lee TD and Yang CN 1952 Phys. Rev. 87 410
Ising E 1925 Zeitschrift fur Physik 31 253
Blöte HWJ, Luijten E and Heringa JR 1995 J. Phys. A: Math. Gen. 28 6289-6313
Talapov AL and Blöte HWJ 1996 J. Phys. A: Math. Gen. 29 5727
Sengers JV and Shanks JG 2009 J. Stat. Phys. 137 857
Sekar H, Ito N, and Hu CK 2012 J. Chem. Phys. 136 204102
Jones JE 1924 Proc. R. Soc. London A 106 463
Lennard-Jones JE 1931 The Proceedings of the Physical Society 43 part 5
Metropolis N and Ulam S 1949 Journal of American Statistical Association 44 335
Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH and Teller E 1953 J. Chem. Phys. 21 1087
Hu CK 1992 Phys. Rev. B 46 6592-6595
Hu CK 1992 Phys. Rev. Lett. 69 2739-2742
Wu MC and Hu CK 2002 J. Phys. A. 35 5189-5206
Ivashkevich EV, Izmailian NS and Hu CK 2002 J. Phys. A. 35 5543-5561
Hu CK 1994 J. Phys. A 27 L813-L820
Hu CK, Lin CY and Chen JA 1995 Phys. Rev. Lett. 75 193-196
Hu CK and Lin CY 1996 Phys. Rev. Lett. 77 8-11
Wang FG and Hu CK 1997 Phys. Rev. E 56 2310-2313
Okabe Y, Kaneda K, Kikuchi M and Hu CK 1999 Phys. Rev. E 59 1585-1588
Tomita Y, Okabe Y and Hu CK 1999 Phys. Rev. E 60 2716-2720
Izmailian NS and Hu CK 2001 Phys. Rev. Lett. 86 5160-5163
Wu MC, Hu CK and Izmailian NS 2003 Phys. Rev. E 67 065103
Hu CK 2013 AIP Conf. Proc. 1518 541
Flory PJ 1949 J. Chem. Phys. 17 1347-1348
Flory PJ and Krigbaum WR 1949 J. Chem. Phys. 18 1086-1094
Shakhnovich E and Gutin A 1990 J. Chem. Phys. 93 5967
Vogel T, Bachmann M and Janke W 2007 Phys. Rev. E 76 061803
Lee JH, Kim SY and Lee J 2010 J. Chem. Phys. 133 114106
Lee JH, Kim SY and Lee J 2012 Phys. Rev. E 86 011802
Bedini A, Owczarek AL and Prellberg T 2013 J. Phys. A: Math. Gen. 46 085001
Chen CN, Hsieh YH and Hu CK 2013 EPL 104 20005
Li MS, Co NT, Reddy G, Hu CK, Straub JE and Thirumalai D 2010 Phys. Rev. Lett. 105 218101
Chiti F, Stefani M, Taddei N, Ramponi G and Dobson CM (2003) Nature 424 805
Co NT, Hu CK and Li MS 2013 J. Chem. Phys. 138 185101
Cabraile-Lago C, Quinlan-Pluck F, Lynch I, Dawson KA and Linse S 2010 ACS Chem. Neurosci. 1 279-287
Sallon S, et al. (2008) Science 320 1464
Yashina S, et al. 2012 Proc. Nat. Acad. SCI. USA 109 4008-4013
Ma WJ and Hu CK 2010 J. Phys. Soc. Japan 79 054001
Ma WJ and Hu CK 2010 J. Phys. Soc. Japan 79 104002
Hu CK and Ma WJ 2010 Prog. Theor. Phys. Suppl. 184 369
Steele D 1985 Chem. Soc. Faraday Trans. 2 81 1077
Karatassos K, Adolf AB and Hotston S 2000 J. Chem. Phys. 112 8695
Schlick T 2002 Molecular Modelling and Simulation: an Interdisciplinary Guide (Springer, New York)
Mayoral E and Robledo A 2005 Phys. Rev. E 72 26209; Grassberger P 2005 Phys. Rev. Lett. 95 140601 (2005); Robledo A 2006 Physica A 370 449
H. C Andersen., J. Comput. Phys. 52, 24 (1983).
M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids (Oxford Univ. Press, Oxford, 1987).
Tsallis C 1988 J. Stat. Phys. 52 479
Lima JAS and Silva R 2005 Physica A 338 272
Silva R Jr, Plastino AR and Lima JAS 1998 Phys. Lett. A 249 401
Douglas P, Bergamini S and Renzoni F 2006 Phys. Rev. Lett. 96 110601
Beck C and Cohen EGD 2003 Physica A 322 267
[82] Vignat C and Plastino A 2005 Phys. Lett. A 343 411
[83] Go N 1983 Ann. Rev. Biophys. Bioeng. 12 183
Kouza M, et al. 2005 Biophysical J. 89 3353
Li MS, Hu CK, Klimov DK and Thirumalai D 2006 Proc. Natl. Acad. Sci. U.S.A. 103 93
Li MS, Kouza M and Hu CK 2007 Biophysical J. 92 547
Kouza M, Hu CK and Li MS 2008 J. Chem. Phys. 128 045103
Fang HJ, et al. 2009 Biophysical J. 96 1892

[84] Mamasakhlisov YS, Hayryan S, Morozov VF and Hu CK 2007 Phys. Rev. E 75 061907
[85] Pluchino A, Rapisarda A and Tsallis C 2007 EPL 80 26002
[86] Dasgupta C, Ma SK and Hu CK 1979 Phys. Rev. B 20 3837
[87] Gevorkian SG, Allahverdyan AE, Gevorgyan DS and Hu CK 2011 EPL 95 23001
[88] Gevorkian SG, Allahverdyan AE, Gevorgyan DS, Simonian AL and Hu CK 2013 Plos One 8 e78526
[89] Goate A, et al. 1991 Nature 349 704
[90] Chiang HL, Ngo ST, Chen CJ, Hu CK and M. S. Li 2013 Plos One 8 65358
[91] Sayle RA and Milner-White EJ 1995 Trends Biochem. Sci. 20 374
[92] Chen WT et al. 2012 PLoS One 7 e35807
[93] Truong PM, Viet MH, Nguyenand PH, Hu CK and Li MS 2014 J. Phys. Chem. B 118 8972
[94] Chen CYC 2011 PLoS ONE 6 e15939
[95] Lipinski A, Lombardo F, Dominy BW and Feeney PJ 1997 Adv Drug Del Rev 23 3-25
[96] Petkova AT, Yau WM and Tycko R 2006 Biochemistry 45 408-512
[97] Luhrs T, Ritter C, Adrian M, Riek-Loher D, Bohrmann B, et al. 2005 Proc. Natl. Acad. Sci. U.S.A. 102 17342-17347
[98] Viet MH, Chen CY, Hu CK, Chen YR and Li MS 2013 Plos One 8 79151
[99] Hayryan S, Hu CK, Skrivanek J, Hayryan E and Pokorny I 2005 J. Compu. Chem. 26 334-343
Busa J, et al. 2005 Compu. Phys. Commu. 165 59-96
Busa JJ, Hayryan S, Wu MC, Busa J and Hu CK 2012 Compu. Phys. Commu. 183 2494-2497
[100] Wu MC, Li MS, Ma WJ, Kouza M and Hu CK 2011 EPL 96 68005
[101] Dobson CM 2003 Nature 426 884
[102] Chiti F and Dobson CM 2006 Annu. Rev. Biochem. 75 333