UNIWERSYTET MIKOŁAJA KOPERNIKA
INSTYTUT FIZYKI
ul. Grudziądzka 5/7 87-100 TORUŃ
http://www.phys.uni.torun.pl/
Tel. centr. (48 56) 611 33 10 Fax (48 56) 622 53 97
Sekretariat: (48 56) 622 63 70 e-mail: ifiz@phys.uni.torun.pl

Data: Toruń, February 14th, 2009
L. Dz.

Dr hab. Wiesław Nowak, PhD, DSc
Professor of N. Copernicus University
Theoretical Molecular Biophysics Group

Opinion on Dissertation
submitted to the Institute of Physics Polish Academy of Sciences, Warsaw
for the degree of Doctor of Philosophy (Science)
by MAKSIM KOUZA

Revolution in computer technology brings new opportunities in science. Complex calculations may be currently performed using much cheaper and better computers than just a few years ago. At the same time progress in scientific software development brings new advanced research tools to every laboratory. Numerical simulations of various physical and biological processes have become standard research methods. However, the computational studies of certain objects, such as biomolecules are far from being trivial. For example, a proper modeling of proteins is an area of vigorous research, due to their complexity and a subtle interplay of various factors governing the physics of such soft matter. Particularly interesting are numerical simulations of processes that may be observed experimentally. Protein folding is an excellent example of such phenomenon.

Presented to my review the PhD Thesis by Mr. Maksim Kouza "Numerical Simulation of Folding and Unfolding Proteins" is devoted to modern, important and general problem of theoretical interpretation of Atomic Force Microscope (AFM) experiments performed on single molecules. Since the famous paper by Gaub et al. published in Science in 1994, hundreds of experiments were performed in order to catch basic features of mechanical unfolding of proteins. The experiments provide us with force-extension curves which require a proper molecular interpretation. Computational approach seems to be the most promising method to create accurate mechanistic picture of proteins. The mechanical properties of proteins, scrutinized recently by Prof. M. Cieplak and Dr. J. Sułkowska [Biophys. J. **94** (2008) 6], are critically important for their biological and medical function. In his Thesis Mr. Kouza studied folding and unfolding process of several carefully selected model proteins. He used several variants of so called coarse grained and all-atoms molecular dynamics simulations methods. He successfully developed a new variant of replica exchange method, which is better suited for this type of studies than the standard procedure.

The dissertation consists of 122 pages. The material has been published (or will be published soon) in 6 papers (co-authored by M. Kouza) in high-impact worldwide known scientific journals (2x Biophys. J., 2x J. Chem. Phys., J. Phys. Chem. A). It contains Introduction (Chapter 1, with a list of publications), Conclusions, References (240 items). The material is logically divided into nine chapters. In the next part I will present the content of these chapters together with my comments and critical remarks.

Chapters 2 and 3 (20 pages) present basic concepts and computational tools. In general it is a well written, informative (and nicely coupled with the rest of the Thesis) part. However, I found several minor mistakes or omissions:

(a) The statement "*Helices are one-dimensional structures*" is wrong (page 11);

(b) The statement "*Basically, the conformational entropy of NS is zero*" is left without a comment nor any definition of this "conformational entropy");

(c) Explanations to formula (5) page 22 does not correspond to the symbols used in this formula;

(d) The definition of folding time, denoted as τ^i_u in page 31 (next to formula (31)) seems to be wrong.

In the Chapter 4 effects of finite size of protein on cooperativity and rate of protein fording are studied. The author used LMSC (lattice model with side chain) and a coarse-grained off-lattice model (defined earlier in the Thesis, but developed by other authors) to perform Monte Carlo and Langevin dynamics simulations for 23 different proteins. For estimation of thermodynamics properties the multiple histogram method was used. The extent of cooperativity of the transition to conformations close to the native ones (NBA) from an ensemble of unfolded states was measured using a parameter Ω_c. The interesting result obtained in this chapter is scaling of Ω_c with the number of aminoacids N: $\Omega_c = N^{\zeta}$, where ζ=2.2-2.4 (for strong first order transition expected value of ζ is 2.0). In my opinion the statement presented on page 38, that "*Since the barriers to global unfolding [are] relatively small it follows that there must be large conformational fluctuations...*" is poorly justified by the presented data. The shape of the potential well and not the height of the energy barriers to unfolding is the main factor determining the amplitude of conformational fluctuations.
The numerically estimated dependence of the folding free energy barrier on the number of aminoacids (dG ~ SQRT (N)) is the same as indicated by experimental data. This is a very valuable and informative result, supporting usage of simplified physical models of proteins.

The Chapter 5 presents mixed experimental/theoretical study of a small α domain hbSBD protein (52 AA). Thermal folding-unfolding transitions have been studied using circular dichroism (CD) spectroscopy. Folding temperature was determined as well as the transition enthalpy. Extensive numerical simulations confirmed that this protein is two state folder. Calculated folding time is at variance with the experimental one by a few orders of magnitude. Folding-unfolding transitions measured by CD looks "sharper" than the estimated form the Go-off-lattice model. In

general I consider this chapter an interesting and solid piece of scientific investigation. However, I have a real difficulty in estimation to what extent Mr. M. Kouza participated in all related studies. The chapter 5 is almost 1:1 replica of the Biophysical Journal paper **89** (2005) 3353, which has 7 authors (with MK as the first author). The Chapter 5 reproduces even some English misusage (*de-grate*, instead of *degrade*, page 45), but on the other hand some figures gain colors in the Thesis. I presume that the theoretical part of this chapter was prepared by the candidate. Thus, the often usage of the pronoun "our" is in this chapter misleading. More clear distinction between the work of the PhD Thesis author and other collaborators' contribution should be done in this chapter.

Chapters 6, 7 and 8 have a common denominator – the object of studies presented is single or three domain ubiquitin (Ub) protein. This is small modular protein extensively studied due to its huge physiological role. Ubitiquination marks proteins for degradation in proteosome, but Lys63-C attachment induces DNA repair and other useful processes. Mechanical stability of characteristic immunoglobulin like modules of Ubiquitin was studied by numerous authors. Thus, applying well developed numerical simulation Go-like models to this particular system is a very good idea, perhaps for the first time explored by P. Cieplak and P. Marszałek [J. Phys. Chem. 123 (2005) 194903 – not mentioned in the Thesis]. In these there chapters Mr. Kouza presents many new, detailed and important results on Ub phase diagrams, refolding under a quenched force, free energy landscapes and others. In my opinion the most valuable result is a new variant of the famous replica exchange method (REM) of simulations. This method, with an idea coined back in 1991 by Geyer, allows one to save a lot of computer time. In the paper published in J. Chem. Phys. 2008 by Kouza, Hu and Li, the authors propose to use force and not temperature, as a criterion for the replica exchange. This variant is much better suited for numerical simulations of AFM experiments. It is worth noting, that the similar approach of using REM has been proposed earlier in 2006 by Lu et al. (*Simulating Force-Induced Conformational Transitions in Polysaccharides with the SMD Replica Exchange Method; Z. Lu, H. Hu, W. Yang, and P. E. Marszalek; Biophys J. (2006); **91**(6): L57–L59)*. It is a pity that this

paper was omitted by the author from otherwise very extensive and competent literature review.

In the description of RE method presented in the section 6.4 I suspect that some problems in notation have occurred: the equation (48) does not follow from the detailed balance condition presented in the equation (46) (in my opinion the nominator and the denominator on RHS of Eq. (48) should be interchanged).

The statement on page 51, Chapter 6, that "*the external force increases unfolding barriers*" is perhaps not correct and is at variance with the statement presented on page 55 that "*the external force lowers the unfolding barrier*" (as may be seem from Fig. 19a, indeed).

The Chapters 6, 7 and 8, again presenting material already published in J. Chem. Phys. 2008 by Kouza, Hu and Li with only minor modifications, contain a lot of original results. The T-f (temperature vs. a fraction of native contacts) for different forms of Ub are particularly interesting. The observation that a three domain Ub may be a two state folder is also important, though in my opinion it needs further verification since the predicted by Go-like model value of energy barrier ~1 kcal /mol is clearly close to kT ~ 0.6 kcal/ mol (at 300K) and thus this value is too low to affect biological systems mechanics. There are also numerous data on mechanism of thermal and mechanical unfolding of ubiquitins presented. The analysis of folding sequencing is very clear, nicely presented and easy to understand. The observation that the mechanical unfolding pathways sometimes depend on a protein end selected for fixation is important, though not new, and it deserves more systematic study on a wider class of proteins. The research and the presented data on dependence of folding times on the quench force are also very useful (for example Fig. 30, Fig. 34) – such studies help a lot in establishing efficient and adequate methodology of numerical simulations of single molecule AFM experiments. I want to stress that Mr. Kouza presents comparisons of his modeling studies with experimental results everywhere it is possible. In general, his findings are in agreement with the experimental data.

The Chapters 9 and 10 are devoted to numerical studies of DDFLN4 protein, a biomolecule related to the organization of cellular cytoskeleton via its participation in the acting binding system. DDFLN4 has been studied experimentally by Sweiger in

2004 and computationally by Li et al. in 2008. In the Chapter 9 Mr. Kouza presents results of low speed mechanical unfolding of DDFLN4 and shows that a very low speed of dragging force is necessary to obtain a reasonable agreement between theoretical and experimental unfolding pathways. However, the existence of the third unfolding peak at $\Delta R \sim 22$ nm was not confirmed in the Go-like model simulation. Thus an extensive all-atom modeling of this process using GROMOS force field had been performed and the results of these studied are presented in the Chapter 10. The most interesting results is that this 22 nm peak in the force-extension curve is related to breaking of 5 non-native hydrogen bonds in the intermediate state of the studied protein. Thus, these two chapters together present a very informative case study on the methodology of numerical simulation of protein unfolding. The conclusion is that the simplified theoretical models, based solely on native contacts, are not always adequate and expensive but more realistic all-atoms models had too be used as well.

The Thesis is written in rather good English, but the language is very concise and more appropriate for a journal paper style than to a PhD thesis. One should be aware that in this Thesis there are large parts of the material already published in BJ or JCP. Quite often some terms are introduced with the reference to the literature and not explained, even shortly, in the Thesis. Numerous figures and plots are of very good quality. Conclusions are (in general) well founded in the presented data. The author has shown his excellent orientation in the scientific literature related to his research. There are some minor mistakes and omissions, mainly of technical nature, some examples are listed in the Appendix to this review. The extensive and original results obtained by Maksim Kouza, related to very modern topics of biological physics, have been published (or submitted for publication) in at least 5 world class journals. I expect that papers published by Maksim Kouza will be quoted often and for a long time in the scientific literature.

Despite some really minor and as always necessary criticism **I think that <u>this is a very solid thesis</u>, it is fully compliant with the Polish (and other, known to me, international) standards. Therefore, I recommend proceeding with further steps of the PhD degree procedure for Mr. Maksim Kouza.**

Depending on the outcome of the final Thesis exam, should the Scientific Council propose a distinction, I will be in favor.

Wiesław Nowak,

W. Nowak

PhD, DSc., Prof. NCU

<u>Appendix</u>

Minor remarks:

1. At least 16 acronyms used in the Thesis were not included in the list of abbreviations prepared by the author.
2. Typographical errors were noticed in the following references: [3], [6], [22], [34], [36], [105]. [108], [130], [153], [155], [159], [163], [164], [165], [171], [180,181],
3. Page 7 "was deciphered" → were deciphered
4. Page 22: **r** symbol is used for the description of the quantity used later as **R** (end-to-end vector)
5. Page 26: No comment of the meaning of Γ in Eq. (16), in such formula Γ_i should be rather presented
6. Page 44: "does not applied"?
7. Page 81: "force filed" → force field
8. Page 99: Table caption : reference [231] → [230]?
9. Page 100: ("our previous results" [321] – in [231] M Kouza is not a co-author, why "our"?
10. Page 101, 110: "now a days" → nowadays
11. Page 102: "steepest decent" → descent