

Apart from biological functions, peptides are of uttermost importance as models for unfolded, denatured or disordered state of the proteins. Similarly, mini-proteins such as Trp-cage have proven their role as simple models of both experimental and theoretical studies of protein folding. Molecular dynamics and computer simulations can provide an unique insight on processes at atomic level. However, simulations and inadequate conformational sampling. Both principal issues were tackled in this study. We demonstrated the inability of the used force fields to predict consistently individual conformational preferences of individual amino acids in the form of dipeptides and ensembles and shift propensities of individual amino acids to particular conformers.

In order to shed light on the nature of conformational ensembles under various denaturing conditions, we studied host-guest AXAXA peptides. The simulations revealed that thermal and chemical denaturation by urea produces qualitatively different results and shift propensities of individual amino acids to particular conformers. The problem of insufficient conformational sampling was dealt by introducing gyration- and inertia-tensor based collective coordinates to metadynamics. A novel miniprotein were designed by idea of retro transformation of protein sequence. The resulting retro Trp-cage molecule does not fold in water but the structure emerges upon addition of a cosolvent—2,2-trifluoroethanol (TFE) into buffer. We further developed a novel model of TFE based on generalized amber force field model of TFE were questioned.

However, this behavior was not observed in simulations and therefore the force field model of TFE were questioned.

We exhaustively optimized a field parameter of TFE and behaves realistically in TFE/water mixtures as we investigated by means of Kirkwood-Buff theory of solutions.

- Vymetá, J.; Slabý, I.; Špáha, A.; Vondrášek, J.; and Lyngstadas, S. P. Bioluminescence and molecular modeling of human ameloblastin suggest a two-domain intrinsically unstructured calcium-binding protein, *European Journal of Oral Sciences* **2008**, *116*(2), 124-134.
- Vymetá, J.; Slabý, I.; Špáha, A.; Vondrášek, J.; and Lyngstadas, S. P. Bioluminescence and molecular modeling of human ameloblastin suggest a two-domain intrinsically unstructured calcium-binding protein, *European Journal of Oral Sciences* **2008**, *116*(2), 1082-1092.
- Kumprecht, L.; Buděšínský, M.; Vondrášek, J.; Černý, J.; Cisárová, I.; Bryndá, J.; Herzig, V.; Koumrik, P.; Závada, J.; and Kraus, T. Rigid Duplec alpha-Cyclodextrin Reversibly Connected With Disulfide Bonds. Study of Alanine Dipeptide and Free-energy Space of Peptides—the Alanine Dipептиde Case *The Journal of Physical Chemistry B* **2010**, *114* (16), 5632-5642.
- Vymetá, J.; and Vondrášek, J. Metadynamics as a tool for mapping the conformational and free-energy space of peptides—the free energy profile of alanine dipepitiide. *Chemical Physics Letters* **2011**, *503*(4-6), 301-304.
- Vymetá, J.; and Vondrášek, J. The DF-LCSD(T) correction of the ϕ/ψ force field dihedral parameters significantly influences the free energy profile of alanine dipepitiide. *Chemical Physics Letters* **2011**, *503*(4-6), 301-304.
- Vymetá, M.; Vymetá, J.; Hejda, J.; Vondrášek, J.; and Jurgová, P. Like-Charge Guanidinium Pairing from Molecular Dynamics and Ab Initio Calculations. *The Journal of Physical Chemistry A* **2011**, *115* (41), 11193-11201.
- Vymetá, J.; and Vondrášek, J. Gyration- and Inertia-Tensor-Based Collective Coordinates for Metadynamics. Application on the Conformational Behavior of Polyamine Peptides and Trp-Cage Foldings. *The Journal of Physical Chemistry A* **2011**, *115* (41), 11455-11465.
- Vymetá, J.; and Vondrášek, J. Short Peptide Test Case. *Journal of Chemical Theory and Computation* **2013**, *9* (1), 441-451.
- Vymetá, J.; and Vondrášek, J. Critical Assessment of Current Force Fields. *Proteins (submitted)*.
- Towsé, C.-L.; Vymetá, J.; Vondrášek, J.; and Daggett, V. Potential for understanding residual structure in denatured states and intrinsically disordered proteins (*submitted*). *of Folding Similar to the Original Only upon TFE Addition (submitted)*.
- Vymetá, J.; Hejda, B.; Černý, J.; Zádек, L.; Chaloupková, R.; Sklenář, V.; and Vondrášek, J. From a Structure to Disorder. Retrospective Operation on the Trp-cage Miniprotein Sequence Produces an Unstructured Molecule Capable and Vondrášek, J. Unique Insights Face Two cardinal problems—Inaccuracy of force fields and Von Drášek, J. and Daggett, J.; and Vymetá, J. Short Peptide Test Case. *Journal of Chemical Theory and Computation* **2013**, *9* (1), 441-451.
- Vymetá, J.; and Vondrášek, J. Critical Assessment of Current Force Fields. *Proteins (submitted)*.
- Vymetá, J.; and Vondrášek, J. Perimetal and calculated properties of pure as well as water mixed solutions. *on Generalized Amber Force Field provides realistic agreement between experimental and calculated properties of pure as well as water mixed solutions.*

Abstract

Selected publications

Abstrakt

Peptidy, kromě své biologické funkce, představují také důležité modely nesbalených, denaturovaných nebo nestrukturovaných proteinů. Poborně důležitými modely pro experimentální i teoretické studium sbalování proteinů jsou miniproteiny, jako např. Trp-cage. Chování peptidů i proteinů lze studovat v počítačových simulacích pomocí metod molekulární dynamiky, které umožňují sledovat děje v atomistickém rozlišení. Tyto metody však čelí však dvěma zásadním problémům – přesnosti používaných energetických funkcí a nedostatečnému vzorkování konformačních stavů. V této disertaci jsem se zabýval oběma okruhy problémů.

Vliv rozdílných, běžně používaných energetických funkcí („force fields“) byl testován na modelu aminokyselinových dipeptidů. Žádná sada parametrů však nedokázala konzistentně reprodukovat konformační preference jednotlivých aminokyselin. Výsledky simulací byly mezi sebou srovnány a byly hledány příčiny jejich vzájemných odlišností.

Abychom odhalili, jakým způsobem různé podmínky ovlivňují konformační stavы peptidů, zkoumali jsme vlastnosti aminokyselin v AAXAA peptidech. Simulace odhalily zásadní rozdíl ve vlivu tepelné a chemické denaturace (močovinou) na charakter a zastoupení konformací peptidů, stejně jako konformačních preferencí jednotlivých aminokyselin.

K problematice vzorkování konformačního prostoru jsem přispěl zavedením kolektivních souřadnic pro metadynamiku odvozených z gyračního tenzoru a tenzoru setrvačnosti. Efektivita těchto kolektivních souřadnic popisujících velikost a tvar molekul byla testována v simulacích alaninových polypeptidů a Trp-cage miniproteinu. V těchto simulacích bylo úspěšně dosaženo reprodukovatelného nalezení nativní konformace miniproteinu a podstatného zlepšení ve vzorkování konformačního prostoru flexibilních polyalaninových peptidů.

Zcela nový miniprotein byl vytvořen obrácením sekvence Trp-cage. Tento umělý konstrukt však narozdíl od Trp-cage nevytváří stabilní třídimenzionální strukturu v běžných pufrech, ale strukturuje se až po přidání 2,2,2-trifluorethanolu (TFE). Stabilita a další vlastnosti molekuly retro Trp-cage byly studovány v MD simulacích, ale nepodařilo se nalézt strukturu indukující efekt TFE. Proto se stalo TFE předmětem našeho dalšího zájmu.

Nové parametry pro TFE, založené na předchozím modelu, byly optimalizovány, aby lépe a kvalitněji popsali vlastnosti nejen samotného TFE, ale i jeho vodných roztoků. Tento nový model realističtěji zachycuje chování směsných roztoků v rámci Kirkwoodovy-Buffovy teorie.

Curriculum vitae

Personal information

Name Mgr. Jiří Vymětal
Date of birth 19.9.1984, Kroměříž
Nationality Czech

Contact

Institute of Organic Chemistry and Biochemistry AS CR
Flemingovo nám. 2.
166 10 Praha 6
Czech Republic
tel: +420 220 183 294
e-mail: jiri.vymetal@uochb.cas.cz

Education

- 2009–present: Doctoral program Modeling of Chemical Properties of Nano- and Bio-structures, Faculty of Science, Charles University in Prague and Institute of Organic Chemistry and Biochemistry AS CR. Thesis: Computational study of short peptides and miniproteins in different environments.
- 2009: Master of Chemistry. Thesis: Theoretical Study of the Conformational Behavior and Structural Preferences of Amino Acids in Peptides by Metadynamics.
- 2007–2009: Master's program Modeling of Chemical Properties of Nano- and Bio-structures, Faculty of Science, Charles University in Prague.
- 2007: Bachelor of Biology. Thesis: Structural proteins involved in mineralization of tooth tissues of mammals.
- 2004–2007: Bachelor's program Molecular Biology and Biochemistry, Faculty of Science, Charles University in Prague.

and thus facilitates experiments in solutions.¹⁰ A mechanism how TFE stabilizes their formation. Additionally, TFE improves solubility of peptides and proteins buffers containing TFE (up to 50% v/v) usually stabilize the α -helices or induce

β -sheet structure e.g. trifluoroethanol (TFE) and hexafluoro-2-propanol (HFIP).¹⁰ Compounds are routinely utilized for their stabilizing effect on the protein and peptide presence of denaturant agents.⁹ Apart from these natural osmolytes some other can be naturally found in cells and organisms as a response to stress or adaptation state and functionality of the protein under non-physiological conditions. They the effect of protecting osmolytes is the opposite. They help in stabilizing of the na-

loses its compactness—this particular state is referred as *molten globule*. still present under the mild denaturation conditions but the global tertiary structure pact with significant residual structure.⁷ The secondary structure elements can be achieved only exceptionally and the denatured molecules stay much more compact with denaturation. In the fact, later study proved that the behavior of random coils can of quaternium expand their hydrodynamic radii to the values expected for random coils.⁶ These result have been often misinterpreted and generalized for any kind of urea and guanidinium hydrochloride belonging to the most commonly used denatu-

ralled denaturants and those with the opposite effect are *protecting osmolytes*.⁵ The chemical compounds in solution (*cosolvents*) that destabilize the folded state are unfolded state can be also influenced by chemical composition of the environment. as for example by temperature and pressure. The equilibrium between folded and however, the equilibrium can be shifted by change of thermodynamic conditions is usually strongly biased toward the folded one under physiochemical conditions. The folded proteins exist in continuous equilibrium with its unfolded state that

structured proteins but populate a range of isoenergetic conformers.⁴ that disordered proteins/regions do not manifest funnel-like energy landscape charge density due to the abundance of charged residues. It is generally accepted amino acids. Such sequences possess statistically lower hydrophobicity and higher and enriched in disorder-promoting (Ala, Arg, Gly, Ser, Glu, Lys and Pro) features—depletion in order-promoting (Leu, Val, Tyr, Trp, Phe, Cys and Asn) dictated.³ The amine-acid composition of unstructured proteins/regions has typical Disorder as well as a structure is coded in primary protein sequence and can be pre-

and translation. On the other hand, disorder is rarely exhibited by enzymes.^{1,2} However, proteins of eukaryotes are estimated to contain more disorder regions (10–45%) than the prokaryotes. The high content of disorder correlates highly with regulation and signaling functions, especially for proteins involved in transcription and replication. Proteins lacking well formed tertiary structures was found in all kingdoms of life. Disorder has been recognized as a vital property of many biologically active proteins.

1 Introduction

- [51] Kim, D. E.; Chivian, D.; and Baker, D. *Nucleic acids research* **2004**, *32*, W526–531.
- [52] Maupeut, J.; Derréumaux, P.; and Tuffery, P. *Nucleic acids research* **2009**, *37*,
- [53] Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; and Kollman, P. *Journal of computational chemistry* **2003**, *24*, 1999–2012.
- [54] Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.; L.; Dror, R. O.; and Shaw, D. E. *Proteins* **2010**, *78*, 1950–8.
- [55] Mackerell, A. D.; Feig, M.; and Brooks, C. L. *Journal of computational chemistry* **2004**, *25*, 1400–15.
- [56] Karpluski, G. A.; Friesner, R. A.; Tropiano-Rives, J.; and Jorgensen, W. L. *The Journal of physical chemistry B* **2001**, *105*, 6474–6487.
- [57] Swope, W. C.; Hori, H. W.; and Rice, J. E. *The journal of physical chemistry* **2010**, *114*, 8621–30.
- [58] Best, R. B.; Buchete, N.-V.; and Hummer, G. *Bioophysical journal* **2008**, *95*, 107–9.

peptides and proteins has not been completely elucidated. The proposed hypothesis expect either direct or indirect effects on polypeptide in TFE/water mixture. The direct effects involve preferential binding on surface of the molecule, local increase of TFE concentration in vicinity of the protein and stabilizing hydrogen bonding of backbone due to the lower effective dielectric constant of the mixed solvent.^{11,12} The indirect mechanism influences the thermodynamics of folding by chaotropic effect on solvation layers of folded and unfolded states.¹³ Furthermore, the activity of TFE can be pronounced in both cases by transient self-aggregation of TFE molecules.¹⁴

Short peptides are used traditionally as experimental and theoretical models of the unfolded, denatured or intrinsically disordered states of proteins.^{15,16} Various studies proved that the sampling of conformational space by peptide backbone is neither random nor uniform as could be inferred from random coil model¹⁷ employed in the last decades. However, the conformational states preferred by short flexible peptides cannot be directly and unambiguously determined by any experimental method. The individual intrinsic conformational preferences of amino acids are assumed to play a role in protein folding and residual structure of unfolded proteins.¹⁶ Different trends in conformational preferences can be observed by spectroscopic methods in various host-guest peptides such as GGXGG,¹⁸ GXG,¹⁹ AXA²⁰ or GPPXPPGY.²¹ However, amino acids themselves in the form of dipeptides manifest distinguishable intrinsic backbone preferences.^{22,23}

The structure of folded proteins has unique features—network of intramolecular hydrogen bonds maintained by secondary structure elements and native contacts between side chains of amino acids. The planar peptide bonds can act both as hydrogen bond donors and acceptors but the order of donor/acceptor pattern is governed by direction of polypeptide chain. Interestingly, almost symmetrical image of the three-dimensional structure of a protein can be constructed from natural L-amino acids, but the following criteria must be taken into account:²⁴

1. The direction of the backbone is reversed, i.e. the protein sequence is reversed from N-terminal to C-terminal.
2. The position of N-H and C=O groups is swapped. As a consequence the direction of hydrogen bonds is reversed.
3. Backbone torsion angles must be transformed by rules: $\phi \rightarrow -\psi$ and $\psi \rightarrow -\phi$.

All these conditions can be easily fulfilled by α -helical structures considering that the transformation of torsions projects them back to the same helical region on Ramachandran plot.

The effect of reverse protein sequence has been already tested experimentally. The “retro transformation” can lead to different outputs. The reverse protein sequences can either loss their ability to fold^{25,26} or the same (or topologically similar) fold can be preserved.²⁷⁻²⁹

- [28] Haack, T.; Sánchez, Y. M.; González, M. J.; and Giralt, E. *Journal of peptide science : an official publication of the European Peptide Society* **1997**, *3*, 299–313.
- [29] Pan, P. K.; Zheng, Z. F.; Lyu, P. C.; and Huang, P. C. *European journal of biochemistry / FEBS* **1999**, *266*, 33–9.
- [30] Rizzuti, B.; and Daggett, V. *Archives of biochemistry and biophysics* **2013**, *531*, 128–35.
- [31] Piana, S.; Klepeis, J. L.; and Shaw, D. E. *Current opinion in structural biology* **2014**, *24*, 98–105.
- [32] Piana, S.; Lindorff-Larsen, K.; and Shaw, D. E. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 17845–50.
- [33] Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; and Shaw, D. E. *Journal of the American Chemical Society* **2012**, *134*, 3787–91.
- [34] Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; and MacKerell, A. D. *Journal of Chemical Theory and Computation* **2012**, *8*, 3257–3273.
- [35] Cho, S. S.; Levy, Y.; and Wolynes, P. G. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 586–91.
- [36] Das, P.; Moll, M.; Stamatil, H.; Kavraki, L. E.; and Clementi, C. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 9885–90.
- [37] Beck, D. a. C.; and Daggett, V. *Biophysical journal* **2007**, *93*, 3382–91.
- [38] Toofanny, R. D.; Jonsson, A. L.; and Daggett, V. *Biophysical journal* **2010**, *98*, 2671–81.
- [39] Piana, S.; and Laio, A. *The Journal of Physical Chemistry B* **2007**, *111*, 4553–4559.
- [40] Vymětal, J.; and Vondrášek, J. *Chemical Physics Letters* **2011**, *503*, 301–304.
- [41] Laio, A.; and Parrinello, M. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 12562–6.
- [42] Laio, A.; and Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. 2008.
- [43] Šolc, K.; and Stockmayer, W. H. *Journal Of Chemical Physics* **1971**, *54*, 2756–&.
- [44] Minato, T.; and Hatano, A. *Macromolecules* **1981**, *14*, 1035–1038.
- [45] Theodorou, D. N.; and Suter, U. W. *Macromolecules* **1985**, *18*, 1206–1214.
- [46] Kirkwood, J. G.; and Buff, F. P. *The Journal of Chemical Physics* **1951**, *19*.
- [47] Chitra, R.; and Smith, P. E. *The Journal of Physical Chemistry B* **2002**, *106*, 1491–1500.
- [48] Vymětal, J.; and Vondrášek, J. *The journal of physical chemistry. B* **2010**, *114*, 5632–42.
- [49] Vymětal, J.; and Vondrášek, J. *Journal of Chemical Theory and Computation* **2013**, *9*, 441–451.
- [50] Vymětal, J.; and Vondrášek, J. *The journal of physical chemistry. A* **2011**, *115*, 11455–65.

The most detailed view of protein folding, denaturation and disorder can be derived by computer simulations.³⁰ However, the quality of a force field commonly used for fields depends critically on the quality of a force field. The commonality of computer simulations has been shown to catch the process of folding and describe its reversibility for model proteins in extensive simulations.³¹ However, results may differ in provided mechanistic and thermodynamic predictions.³² Simulations of protein disorder by sampling of the ensemble needs at least time scales of microseconds.³³ However, classic molecular dynamics are very computationally demanding because the proper sampling in force fields towards different secondary structure elements can produce like metadynamics allow to accelerate rare events on the reaction coordinate very efficiently. However, the focus of the problem is then shifted toward a choice of proper reaction coordinates for the complex events. The quest for universal protein folding coordinates has succeeded only partially. It was demonstrated that many native folding/unfolding processes can be represented on one-dimensional coordinate given by multidimensional reduction or embedding of relevant descriptors.³⁵⁻³⁸ However, such characteristics always have to involve the detailed knowledge of the particular folded state and its native contacts. Bias exchange metadynamics³⁹ offers an alternative way how to avoid complex reaction coordinates.

Simulations of complex biomolecular phenomena can be very difficult task for classical molecular dynamics because of intrinsic barriers of these processes. Methods like metadynamics allow to accelerate rare events on the reaction coordinate very efficiently. However, the focus of the problem is then shifted toward a choice of proper reaction coordinates for the complex events. The quest for universal protein folding coordinates has succeeded only partially. It was demonstrated that many native folding/unfolding processes can be represented on one-dimensional coordinate given by multidimensional reduction or embedding of relevant descriptors.³⁵⁻³⁸ However, such characteristics always have to involve the detailed knowledge of the particular folded state and its native contacts. Bias exchange metadynamics³⁹ offers an alternative way how to avoid complex reaction coordinates.

2 Aims of the study

The computational studies of peptides and miniproteins are challenging tasks for two principal reasons: i) uncertainty or unsatisfactory precision of the force fields and ii) high demands on the conformational sampling in the computer simulations. In the present thesis I attempted to deal with the following challenges:

1. Assessment of force fields for simulation of peptides.
2. Investigation of conformational preferences of amino acids in water environment and in solvents.
3. Design and testing of collective coordinates for metadynamics of peptides.
4. Characterization and reproduction of experimental properties of designed peptides.
5. Thorough parameterization of a cosolvent for correct water-mixtures properties and reliable usage for simulations of proteins and peptides.
6. Each topic 1–5 was elaborated as an individual study published or submitted to peer-reviewed scientific journal. All related published articles or manuscripts were attached to the thesis as Appendices A–F.

- [3] He, B.; Wang, K.; Liu, Y.; Xue, B.; Uversky, V., N.; and Dunker, A. K. *Cell research* **2009**, *19*, 929–949.
- [4] Uversky, V. N. *Biochimica et biophysica acta* **2013**, *1834*, 932–51.
- [5] Canchi, D. R.; and Gracica, A. E. *Annual review of physical chemistry* **2013**,
- [6] Tamford, C.; Kawahara, K.; and Lapanje, S. *Journal of the American Chemical Society* **1967**, *89*, 729–736.
- [7] Dill, K. A.; and Shortle, D. *Annual review of biochemistry* **1991**, *60*, 795–825.
- [8] Ohgushi, M.; and Wada, A. *FEBS letters* **1983**, *164*, 21–4.
- [9] Yamada, P. H.; Clark, M. E.; Hand, S. C.; Bowlin, B. D.; and Somero, G. N. *Science (New York, N.Y.)* **1982**, *217*, 1214–22.
- [10] Buck, M. *Quarterly reviews of biophysics* **1998**, *31*, 297–355.
- [11] Cammers-Goodwin, A.; Allen, T. J.; Oslie, S. L.; McClure, K. F.; Lee, J. H.; and Kemp, D. S. *Journal of the American Chemical Society* **1996**, *118*, 3082–
- [12] Luo, P.; and Baldwin, R. L. *Biochemistry* **1997**, *36*, 8413–21.
- [13] Wallace, R.; Lee, T. C.; and Cammers-Goodwin, A. *Journal of the American Chemical Society* **1998**, *120*, 5073–5079.
- [14] Hong, D.-P.; Hoshiba, M.; Kuboi, R.; and Goto, Y. *Journal of the American Chemical Society* **1999**, *121*, 8427–8433.
- [15] Shi, Z.; Chen, K.; Liu, Z.; and Kallenbach, N. R. *Chemical reviews* **2006**, *106*, 1877–97.
- [16] Schweitzer-Stennet, R. *Molecular biosystems* **2012**, *8*, 122–33.
- [17] Brant, D. A.; and Flory, P. J. *Journal of the American Chemical Society* **1965**,
- [18] Shi, Z.; Chen, K.; Liu, Z.; Ng, A.; Bracken, W. C.; and Kalleinbach, N. R. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 2791–2800.
- [19] Schweitzer-Stennet, R.; Hagerman, A.; Measey, T. J.; Mattheiu, D.; and Brown, A. M.; and Zondlo, N. J. *Biochemistry* **2012**, *132*, 540–51.
- [20] Eker, F.; Gribenow, K.; Cao, X.; Nahe, L.; and Schweizer-Stennet, R. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 10054.
- [21] Brown, A. M.; and Zondlo, N. J. *Biochemistry* **2012**, *51*, 5041–51.
- [22] Avbelj, F.; Gradaščik, S. G.; Gradaščik, J.; and Baldwin, R. L. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 1272–7.
- [23] Gradaščik, J.; Mohacek-Groshev, V.; Baldwin, R. L.; and Avbelj, F. *Procedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 5–14.
- [24] Gupta-Sarma, P. *FEBS letters* **1992**, *310*, 205–10.
- [25] Lacroix, E.; Vigueras, A. R.; and Seranno, L. *Foldinig & design* **1998**, *3*, 79–85.
- [26] Olszewski, K. A.; Kołomicki, A.; and Skołomicki, J. *Protein engineering* **1996**, *9*, 1794–8.
- [27] Witte, K.; Skołomicki, J.; and Wong, C.-H. *Journal of the American Chemical Society* **1998**, *120*, 13042–13045.

3 Methods

The potential energy functions (force fields) used in this work for description of proteins, peptides and solvents employ the most simplistic mathematical forms. They include a harmonic potential for simulation of bond stretching and angle bending and a variant of trigonometric series for representation of torsional terms. The non-bonded interactions are uniformly modeled by Lennard-Jones 6-12 potential and electrostatics using partial atomic charges interacting by Coulomb's law. The complete energy function follows typically:

$$E = \sum_{bonds} \frac{K_B}{2} (r - r_0)^2 + \sum_{angles} \frac{K_A}{2} (\theta - \theta_0)^2 + \sum_{torsions} \sum_n \frac{V_n}{2} (1 + \cos(n\phi - \delta_n)) \\ + \sum_{i < j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon r_{ij}} \right), \quad (1)$$

where r_0 and θ_0 stands for equilibrium bond lengths and angles, respectively and K_B and K_A for corresponding force constants. Torsion potential is usually expanded in cosine terms with different multiplicities n , phase offsets δ_n and magnitudes V_n . The atoms separated by more than two covalent bonds are allowed to influence each other via nonbonded interactions prescribed for each pair by partial atomic charges q or Lennard-Jones parameters A_{ij} and B_{ij} . The common practice is to treat non-bonded interactions between atoms separated by 3 bonds (1–4 interaction) specially. They are often scaled down because of their strength.

Despite of simple forms such force fields approximate the potential energy surface sufficiently well and in a robust manner what can be judged by the number of successful applications, e.g. protein folding simulations. Although different force fields share more or less the same energy function forms, they may differ substantially in parameters for individual types of interactions and the way how they were obtained.

As the extensive simulations of proteins began to be feasible due to the enormous progress in computational hardware the need for better calibrated force field intensified. Firstly, the force field were improved by more advanced *ab initio* quantum calculations of backbone conformers or φ/ψ profiles. This way led to many improvements, however, only the expensive high level *ab initio* methods can produce precise reference values capable of matching experimental data.⁴⁰ The current force fields still do not reach the precision needed for qualitative and quantitative reproduction of experimental thermodynamic or structural data on protein folding in solution. To address this problem directly the experimental data such as NMR shifts or coupling constants started to be utilized in parametrization process.³⁴

5 Závěry

Tato disertace se zabývala tématy, které zahrnují simulace peptidů a miniproteinů ve vodním prostředí nebo ve směsných solventech. Zde bych shrnul nejdůležitější výsledky:

- Energetické funkce (*force fields*) použité v naší studii nejsou schopny spolehlivě předpovědět jemné konformační preference aminokyselin. Rozdílné výsledky ze simulací se spíše zdají být artefakty nefyzikálních 1–4 interakcí a parciálních atomových nábojů.
- Teplotní a chemická denaturace indukuje rozdílné soubory konformací modelových peptidů, které však nejsou ani náhodné, ani stejnородé. Zároveň se však odlišují od referenčního stavu za fyziologických podmínek. Chemické posuny mohou být nedostatečně citlivé na lokální konformace aminokyselin, a proto mohou poskytovat zavádějící představu o povaze denaturowaného nebo nesbaleného stavu proteinů.
- Kolektivní souřadnice pro metadynamiku odvozené od gyračního tenzoru a tenzoru setrvačnosti mohou být použity k efektivnímu vzorkování v *bias exchange* metadynamice. Tyto souřadnice byly schopné rozsáhlého vzorkování konformerů alaninových polypeptidů i opakovaného sbalení miniproteinu Trp-cage.
- Obrácená sekvence miniproteinu Trp-cage se nesbaluje v běžném pufru, ale nabývá strukturu v 30% roztoku 2,2,2-trifluoretanolu (TFE). Tato struktura sice připomíná Trp-cage, ale liší se ve sbalení jádra miniproteinu. Robetta a PEP-FOLD sice správně předpovědely základní charakteristiky molekuly retro Trp-cage, ale selhaly v modelování postranních řetězců v jádře. Stabilita miniproteinu v simulacích velmi závisí na volbě parametrů energetických funkcí. Pokus více se přiblížit experimentálním podmírkám pomocí standartního modelu TFE selhal, protože nedokázal napodobit strukturu-indukující efekt.
- Důkladně jsme optimalizovali parametry modelu TFE pro lepší popis kapalné fáze a směsných roztoků s vodou. Vylepšení těchto vlastností bylo možné především díky implicitní polarizační korekci, použitou k výpočtu výparné entalpie během parametrizace.

References

- [1] Wright, P. E.; and Dyson, H. J. *Journal of molecular biology* **1999**, *293*, 321–31.
- [2] Uversky, V. N. *The international journal of biochemistry & cell biology* **2011**, *43*, 1090–103.

only one of many descriptors derived from gyration tensor. Gyration tensor (\mathbf{S}) and effective size is commonly described by radius of gyration, however, this property is shape. This approach has a long history in polymer physics and chemistry.^{43–45} The flexible molecule such as polypeptide can be naturally characterized by its size and facilitates the sampling in each of them.

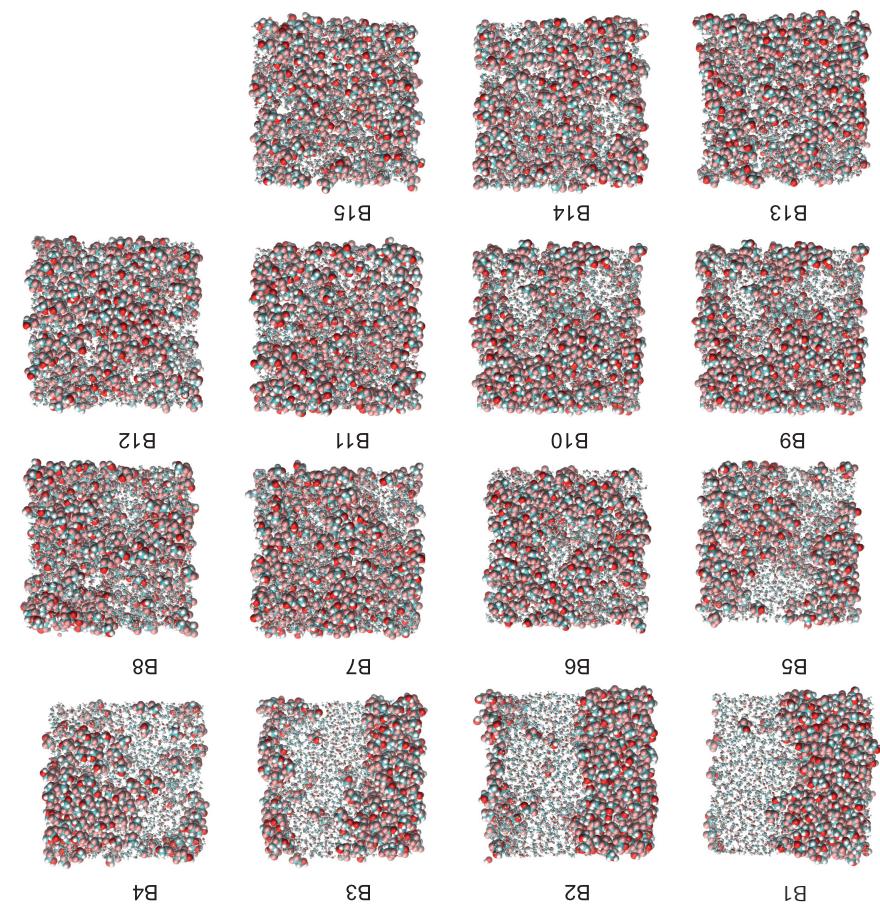
replica according to the utilized bias. The exchange of systems between replicas different collective coordinates. Therefore, different barriers may be crossed in each namics³⁹ uses replicas which are simulated at the same temperature but biased in names to sample the unbiased „transverse coordinates”. The bias exchange method helps to combine metadynamics with other approaches such as replica exchange that hence that depends on the starting conditions. These limitations can be overcome by combining simulations with slow processes, a hysteresis in the set of CVs. If they cannot capture an important slow mode only for the biased simulations appears. As a direct consequence, free energy profiles show poor correlations particularly if the CVs correspond to the slow modes are omitted. Metadynamics is capable of acceleration of rare events and extended sampling only for the biased set of CVs. The inappropriate choice of collective variables causes the most serious problems,

- Number of used CVs should not exceed the limit given by unfavorable scaling of metadynamics for each additional dimension.
- CVs should describe important slow modes, which limit the rate of studied events
- CVs must be continuous and differentiable function of atomic coordinates reactants, products or intermediates of the reactions or transitions.
- CVs should distinguish all important states, conformers, structures, phases, functionalities of metadynamics and its successful application. For the best performance the following properties and requirements on CVs must be fulfilled:⁴²

A choice of suitable collective coordinates (CVs) is a necessary condition for correct description of the system and force the transitions between one state to another one. The updates of the bias potential destabilize immediately the current configuration and force the algorithm to find a new one. The built continuously in an adaptive fashion in order to explore unvisited regions. The bias potential acting on collective coordinates significantly improves sampling. The selected descriptors are efficiently free energy profile is needed. After wards, the selected descriptors are efficiently free energy profile is needed. After no a priori knowledge about the underlying free energy profile is needed. Nevertheless, this descriptors must be specified and provided on the beginning of the simulation but are called collective coordinates (variables) or order parameters. Nevertheless, this access or system of interest can be described by small number of parameters which fields. The heart of the algorithm is dimensionality reduction. It assumes that the problem, namely,⁴¹ metadynamics has been successfully applied on various problems and rare events in molecular dynamics simulations. Since it was proposed by Liao and Parimille,⁴¹ metadynamics is a method for calculation of free energy profiles and accelerating of

velikosti. Rozdilne parametry TFE vykazovaly ruzne stupne agregacie. Uplna separace fazni byla pozorovana pro modely B1–B3. Modely B4–B15 vytvareli silnky setupne

Obrázek 5: Schémová TFE molekula.



tensor of inertia (**I**) describe the distribution of mass in the molecule determined by position of atoms:

$$\mathbf{S} = \frac{1}{N} \begin{pmatrix} \sum x_i^2 & \sum x_i y_i & \sum x_i z_i \\ \sum x_i y_i & \sum y_i^2 & \sum y_i z_i \\ \sum x_i z_i & \sum y_i z_i & \sum z_i^2 \end{pmatrix}, \quad (2)$$

$$\mathbf{I} = \begin{pmatrix} \sum m_i(y_i^2 + z_i^2) & \sum -m_i x_i y_i & \sum -m_i x_i z_i \\ \sum -m_i x_i y_i & \sum m_i(x_i^2 + z_i^2) & \sum -m_i x_i z_i \\ \sum -m_i x_i z_i & \sum -m_i y_i z_i & \sum m_i(x_i^2 + y_i^2) \end{pmatrix}. \quad (3)$$

All summations are performed over N atoms with individual masses m_i . The cartesian coordinates x_i , y_i and z_i must be related to the geometrical center and center of mass of the molecule for gyration and inertia tensor, respectively. Since both tensors are symmetric and positive semidefinite they can be diagonalized with non-negative eigenvalues. The corresponding eigenvectors determine important axes of the object—the geometric axes of an ellipsoid approximating the shape for gyration tensor or principal axes of inertia for tensor of inertia. The three eigenvalues S_1 , S_2 , S_3 (sorted in descending order) are related to the length of individual elliptic semi axes and I_1 , I_2 , I_3 represent principal moments of inertia around corresponding principal axes. Moments of inertia provide information on distribution of mass in perpendicular directions to the corresponding axes. For more or less homogeneous objects it also reflects the effective size that can be better expressed by radius of gyration around particular axis:

$$r_g^{ax} = \sqrt{\frac{I_{ax}}{m}}. \quad (4)$$

The shape descriptors proposed by Theodorou and Suter⁴⁵—Asphericity b and cylindricity c express obviously the deviation from spherical and cylindrical symmetry. Interpretation of the last descriptor relative anisotropy κ^2 also respects symmetry. It reaches 0 for spherical objects and approaches 1 for linear shapes. Anisotropy of planar objects converges to $\frac{1}{4}$.

The microscopic structure of liquid mixtures can be analyzed in framework of Kirkwood–Buff theory.⁴⁶ The unique feature of this approach is the fact, that the corresponding values of Kirkwood–Buff integrals (KBI) can be evaluated from MD simulations as well as calculated from experimental thermodynamic quantities. Kirkwood–Buff theory therefore represents a valuable link between theory and experiment.⁴⁷

The Kirkwood–Buff integrals can be obtained from simulations by using radial distribution functions:

Tabulka 1: Vlastnosti kapalného TFE při 298.15K and 100 kPa. Srovnání experimentálních data a hodnot ze simulací.

Hustota ρ , difuzní koeficient D , výparná entalpie ΔH_{vap} , zastoupení *trans* konformeru, koeficient tepelné roztažnosti α , objemová stlačitelnost β_T , viskozita η , dielektrická konstanta ϵ , průměrný diplový moment $\langle \mu \rangle$ a hydratační volná energie ΔG_{solv} jsou porovnány s experimentálními údaji.

vlastnost	simulace	expt ^a
$\rho[\text{kg.m}^{-3}]$	1382.5	1382.4
$D[10^9 \text{m}^2.\text{s}^{-1}]$	0.70	0.68
$\Delta H_{vap}[\text{kJ.mol}^{-1}]$	42.9(47.5) ^b	42.9
% <i>trans</i>	39.9	40
$\alpha[10^3 \text{K}^{-1}]$	1.27	1.28
$\beta_T[\text{GPa}^{-1}]$	1.32	1.23
$\eta[10^3 \text{kg.m}^{-1}.\text{s}^{-1}]$	1.52	1.72
ϵ_r	34	27
$\langle \mu \rangle[\text{Debye}]$	3.69	2.46
$\Delta G_{solv}[\text{kJ.mol}^{-1}]$	-16.8 (-21.4) ^{b,c}	-18.02
$\Delta G_{solv}[\text{kJ.mol}^{-1}]$	-16.6 (-21.2) ^{b,d}	-18.02

^aReference k původním článkům uvedeny v rukopisu přiloženém k disertaci.

^bHodnoty v závorkách jsou spočteny bez polarizační korekce.

^cSimulováno s TIP4P/Ew modelem vody.

^dSimulováno s TIP4P/2005 modelem vody.

vazby, tím nižší hodnota Kirkwoodova–Buffova integrálu pro TFE molekuly G_{TT} , který také odráží míru jejich shlukování. Struktura směsi a samoshlukování TFE molekul je demonstrováno na obr. 5. Hodnoty Kirkwoodových–Buffových integrálů vypočtené z experimentálních dat a simulací se shodovaly pouze pro ty nejpolárnější modely.

Model, který byl nakonec vybrán, byl dále blíže charakterizován. I další vlastnosti kapalné fáze, které nebyly optimalizovány, byly popsány velmi dobře – viskozita, koeficient teplotní roztažnosti, objemová stlačitelnost, dielektrická konstanta i hydratační volná energie (viz tab. 1). Kalibrace TFE parametrů pro použití s pokročilejšími modely pro vodu byla možná jen díky polarizační korekci. Ta nám uvožnila využít sílu interakcí TFE–TFE, TFE–voda a voda–voda tak, aby výsledné roztoky lépe splňovaly experimentální data.

4 Results and discussion

4.1 Intrinsic conformational preferences of amino acids

where g_{ij} is the radial distribution function between particles i and j .

$$(5) \quad G_{ij} = \int_{\infty}^0 4\pi r^2 (g_{ij}(r) - 1) dr,$$

We already tested the performance of force fields on the model molecule of alanine dipeptide⁴⁸ and demonstrated that high level *ab initio* calculation as reference for parameterization can improve agreement with experiment.⁴⁹ In order to point out the differences between force fields we compared comparatively backbone and side chain conformations of alanine dipeptides in four common force fields.⁴⁹ Additionally, our interest in propensities under different conditions resulted in a collaborative computational study that mapped the effect of denaturation on backbone propensities significantly differently from alanine.

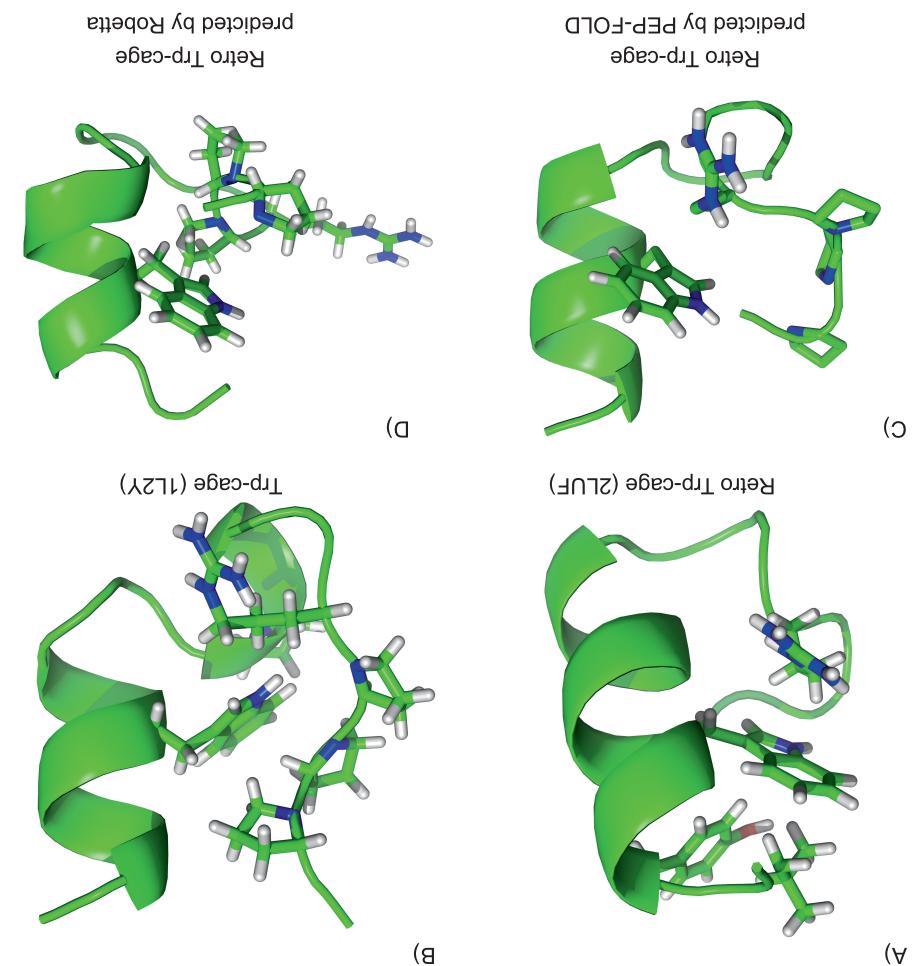
One of the important results of our study is presented in Fig. 1. The preferences

of amino acids were related to the alanine dipeptide in each force field in order to highlight underlying trends. Results obtained by different force fields manifested no consensus between any pair of force fields that was confirmed by very low correlation coefficients.

The fact that very dissimilar trends in backbone propensities were revealed even for the same dipole field such as F90 and F99SB suggested further examination. We prepared an artificial hybrid force field that mixed F99SB with partial atom charges from F90. The resulting backbone propensities correlated strongly with those obtained from F90. The role of the partial charges in the parental F99SB is apparent if their direct influence on torsional potentials *via* 1–4 electrostatic interactions is realized. These interactions are usually scaled by factors without any treatment of 1–4 electrostatics can be a plausible explanation for disagreement of physical justification. Therefore, the different charge distributions and inconsistencies in predictions of retro Trp-cage and Robetta are likely due to the difference in the way they treat 1–4 electrostatics.

We investigated conformational preferences of amino acids on X position in AAAA host peptide under different conditions—physiological, thermal (498K) and chemical (8M urea) denaturation. The AAAA host peptides were chosen as a model of unfolded or denatured protein sequence. The small side chain of alanine restricts the flexibility in respect to the established GXGCG host peptides and thus may provide more realistic picture of real protein chains.

Obrázek 4: Experimentální struktura retro Trp-cage (A), Trp cage (B) and modely retro Trp-cage predikované PEP-FOLDem (C) a Robetteou (D).



The obtained conformational preferences showed clear bias to certain regions of conformational space rather than random sampling under all conditions. However, the conformational bias for individual amino acids was modulated by specific conditions of denaturation. The increased temperature shifted the preferences toward more uniform sampling of all regions for all amino acids. However, the effect of urea on propensities did not result in uniform response. The population of ppII conformers assumed as a prevalent in unfolded or denatured states increased or decreased for particular amino acids. Nevertheless, the ppII regions were never sampled exclusively by any amino acid.

The simulated MD ensembles were validated by chemical shifts that provided a close agreement with experimental $^1\text{H}_\text{N}$ and $^{15}\text{N}_\text{H}$ chemical shifts obtained from all AAXAA peptides in urea at pH 3. The calculated chemical shifts from simulations exhibited bi-modal or multi-modal distributions for each amino acids in apparent accordance with the presence of sampled conformational basins. However, the detailed analysis revealed that there was no correspondence of regions on Ramachandran plot to particular values of chemical shifts. Therefore, it is difficult to elucidate local conformational propensities on basis of “random coil” chemical shifts. This statement was further supported by artificial increasing of ppII populations in the input data. The resulting calculated chemical shifts manifested very low sensitivity on ppII content.

4.2 Metadynamics in gyration-tensor-based collective coordinates

From a general point of view a flexible molecule such as polypeptide can be naturally characterized by its size and shape. This approach has a long history in polymer physics and chemistry.^{43–45} The effective size is commonly described by radius of gyration, however, this property is only one of many descriptors derived from gyration tensor. The others are able to express not only size but also proportions and shape of the molecule. In this study we implemented the size and the shape descriptors—components of gyration tensor (S'_i), principal radii of gyration (r_{gi}), asphericity (b'), acylindricity (c') and relative shape anisotropy (κ^2) in a computer code for metadynamics and tested their performance on model peptides and miniproteins.⁵⁰

Firstly, the performance of newly introduced collective coordinates was tested on polyalanine peptides of four different lengths (3, 6, 9, and 12 residues).

The one-dimensional free energy profiles as function of the collective coordinates in question from bias exchange metadynamics are presented in Fig.2. The smooth and often featureless character of all profiles reflects clearly the flexible nature of alanine peptides. Only the shortest Ala₃ peptide provided profiles with 3 distinguishable minima that corresponds to different conformational families. The other peptides favored α -helix as the most stable conformer that always correspond to the profound minimum on the free energy profiles. The increasing length of the helix due to

stabilizující charakter interakcí přítomných v experimentální struktuře. Parametry ff03 and charmm22/CMAP, známé svou preferencí helikálních struktur, tvořily v simulacích více stabilní helix než ff99SB-ILDN a OPLS-AA/L. Všechny však preferovaly relativně kompaktní struktury.

Protože je struktura retro Trp-cage stabilní pouze v přítomnosti TFE, snažili jsme se simulovat i tento efekt na strukturu miniproteinu. Standartně připravené parametry TFE ovlivnily především chování proteinu v ff99SB-ILDN, kde došlo rozvolnění a denaturaci struktury, které se projevilo zvýšením C_α -RMSD a poloměru gyrase. Tento efekt byl v protikladu s experimentálním pozorováním. Není jisté, co takové chování vyvolalo. Možným vysvětlením je nekompatibilita parametrů pro protein, vodního modelu a modelu pro TFE. Zlepšení modelu TFE jsme si vytýčili jako další cíl.

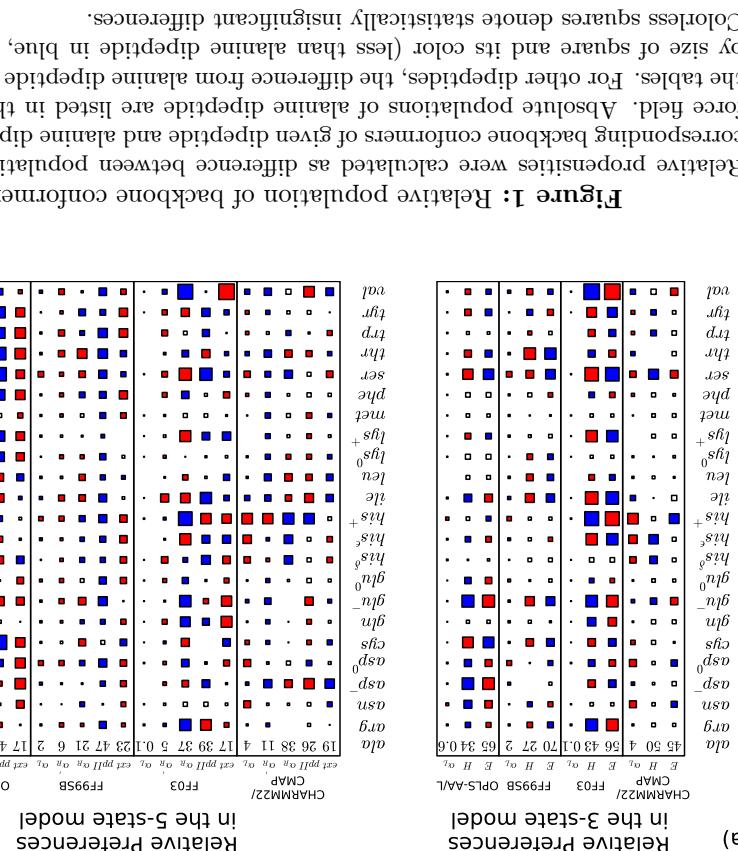
4.4 Optimalizace parametrů 2,2,2-trifluoretanolu

Parametrisace byla rozdělena do 3 fází. V té první jsme určili parametry vhodné pro optimalizaci. Jednotlivé parametry pro Lennard-Jonesovu a elektrostatickou interakci byly mírně pozměněny a jejich odezva na vlastnosti kapalného TFE byla testována v simulacích. Mezi vlastnosti, které jsme sledovali patřila hustota, výparná entalpie, difuzní koeficient a poměr *trans* a *cis* konformeru. Pro následnou optimalizaci bylo vybráno 5 parametrů – Lennard-Jonesovi parametry σ a ϵ pro fluor, parciální náboje fluoru a hydroxylového vodíku, a jeden parametr potenciálu pro torzi C–C–O–H.

V další fázi jsme testovali různé kombinace hodnot jednotlivých parametrů v simulacích kapalného TFE. Pro tento účel jsme vyvinuli protokol, který nám umožnil současně optimalizovat všechny 5 parametry. Závislosti všech 4 sledovaných vlastností byly reprezentovány pomocí fitu analytické funkce, která umožnila hladkou interpolaci a předpovědi výsledků pro jednotlivé parametry. Tento přístup umožnil navrhnut a otestovat v simulacích parametry splňující sledované vlastnosti TFE.

Zvláštní pozornost byla věnována výpočtu výparné entalpie. Použili jsme jednak klasický přístup využitý k získání parametrů původního modelu, ale i verzi výpočtu korigovanou⁵⁷ na polarizační efekty v plynné fázi. Obě varianty postupu vedly k parametrům, které poskytly shodu s experimentálními daty pro 4 výše zmíněné vlastnosti. Pokud byla zahrnuta i polarizační korekce, bylo nalezeno více nejednoznačných řešení, které stejně a kvalitně souhlasily s experimentálními daty. Toto zjištění bylo umožněno právě zvolenou optimalizační metodou. Protože nebylo zřejmé, které parametry budou nejlépe vyhovovat simulacím roztoků, vybrali jsme 28 z nich a ty posléze dálé testovali v simulacích směsi TFE–voda.

Vlastnosti roztoků jednotlivých parametrů se velmi lišily. Ačkoliv je TFE volně mísitelné s vodou, v řadě simulací došlo k oddělení obou kapalin. V rozpustnosti a mísitelnosti obou složek byl pozorován jasné trend. Čím vyšší polarita O–H

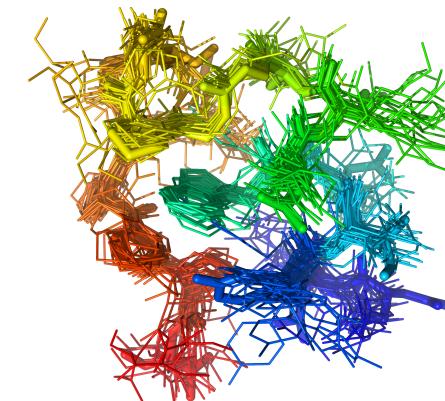


Abychom být charakterizováti povahu nestrukturovaného stavu retro-Trip-cage minimálně výčetem experimentální struktury z experimentální struktury miniproteínu a tak i soubor schopností zachytit nezávisle simulace pro každý set. Všechny sady parametrů struktury však ale spolu přechodné rozdíly v rozdílných chování. Veličiny rozdíl výsledků výsledků struktury ještě poskytují robustní pořadí klasifikaci podle významnosti výsledku. Všechny použité výčetky vykazují značnou mohutnost robustní struktury (2LUF), ale využívají různé parametry, abychom mohli posoudit výsledky. Všechny použité výčetky z experimentální struktury (2LUF), provedené s provedenou sítí, prověřily jisté soubory strukturální dynamiky.

a spíše polohy technického postavení, které je dle alternativních orientací. Nenapodobovatelnými výsledky struktury modelu trypofoládu jsou vlastnosti proteinu znamenající, že typlfoládu s argininem záchrannou pozici ještě model z PEPP-FOLDU, model však ani spojedlou poststrukturu výsledku záchrannou pozici ještě proteinu. Nativní interakce mezi všemi pozůstatky na obr. 4. Na rozdíl od α -helixu však záchrannou model nepospasal správné jeho funkci vlastnosti, které jsou vlastnosti retrovariatu, jak je například výsledkem spojení struktury Tripioninála s PEPP-FOLDU. Oba nastojí významně souběžně s jejich ohodnoceními. Před výsledkem experimentální struktury se pokusily předpovědět 3-dimen-

zionální strukturu retro-Trip-cage pomocí 2 modelovacích nástrojů – Robotky⁵¹ a PEPP-FOLDU.⁵² Obě nastojí významně za rozdílem určené struktury, která byla posleze uložena v PDB NMR experimentu křivku. Struktura-mimikující efekt TFE možnou provedení týpfoládu denaturace křivku. CD spektra ohlášila nařízenství obšálu helikálmich struktur a tímé poskytla týpfoládu struktury podobné nativnímu stavu ze simulací jíson porovnány s experimentální strukturou miniproteínu (tlusté).

Obrazek 3: Projekce struktur miniproteínu Trip-cage.



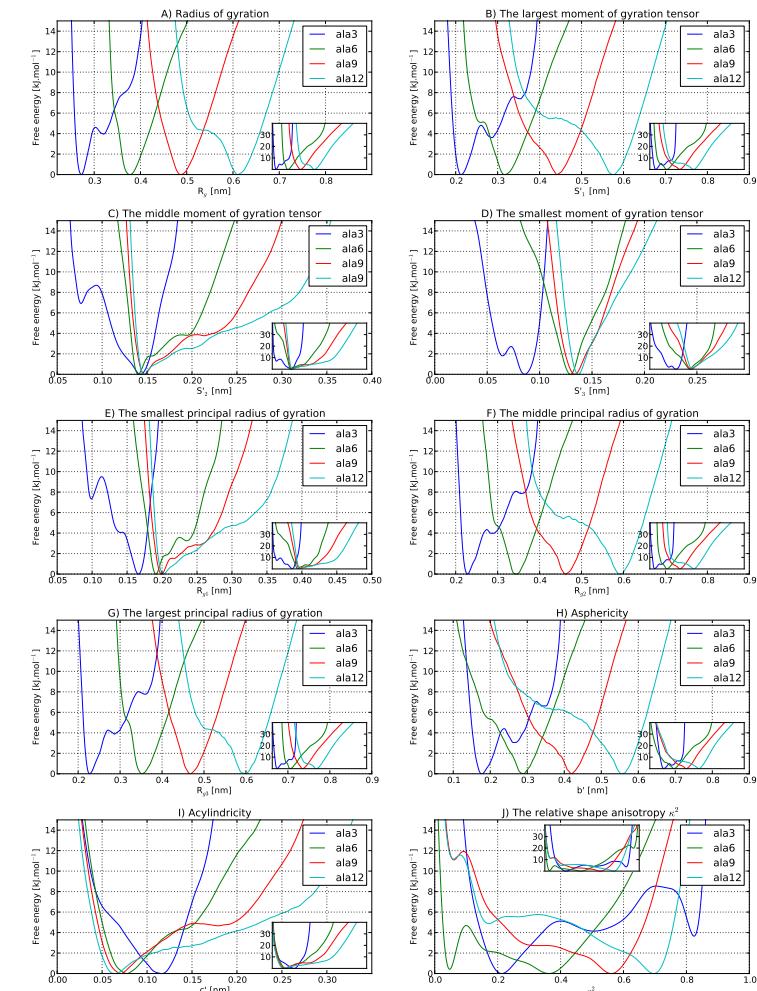
elongation of sequence for Ala₆₋₁₂ is reflected by shift of the minima in R_g , S'_1 , r_{g1} , r_{g2} , b' and κ^2 in panels A, B, F, G, H and J in Fig.2. On the other hand, the cylindrical symmetry of the helix must result in constant (length-independent) values of S'_2 , r_{g3} and c' as it is correctly depicted in panels C, D, E and I. We did not observe any difference in performance of protocols involving various combination of these collective coordinates. The rationale is that all employed CVs are internally linked and always represent a distinct transformation of 3 unique eigenvalues of gyration or inertia tensors. However, such transformations are still very useful for interpretation of molecular shapes.

The more stringent test was attempted to examine a folding of Trp-cage miniprotein using only gyration-tensor-based CVs. It succeeded in almost each simulation for simulations 200 ns long. Once the native state appeared it had been captured by neutral replica and usually remained folded to the end of the simulation. The illustrative picture of the folded Trp-cage molecules is provided in Fig.3. Apart from this structures with low backbone RMSD (<2 Å) from the experimental reference, simulations produced also relatively stable compact structures with non-native core packing and higher RMSD (2–4 Å). It is questionable if they have any relevance to folding or are only artifacts of chosen CVs.

4.3 Computational study of retro Trp-cage miniprotein

The amino acid sequence of retro Trp-cage was obtained by reversing sequence of Trp-cage, synthesized and studied by NMR and CD spectroscopy. Both methods confirmed the unstructured character of the molecule in aqueous buffer. However, changes were observed upon addition of 2,2,2-trifluoroethanol (TFE). CD spectroscopy revealed an increase of helical content and provided typical reversible heat denaturation curve with melting temperature about 32°C. The structure promoting effect of TFE allowed to conduct all NMR experiments necessary for determination of 3D structure of the miniprotein. The resulting structure was deposited in the Protein Data Bank under the accession code 2LUF.

We used the online version of both structure prediction methods—Robetta⁵¹ and PEP-FOLD⁵² to obtain three-dimensional model before the experimental structure was solved. Both services provided sets of models accompanied by their ranking. Robetta and PEP-FOLD succeeded in prediction of C-terminal helix. The overall fold of all models resembled the structure of original Trp-cage (1L2Y) as well as its retro variant (2LUF) which is apparent from Fig. 4. However, no model correctly assembled the core of the miniprotein. The distorted tryptophan–arginine motif was predicted only once by PEP-FOLD. Similarly, the canonical tryptophan–proline motif known from Trp-cage molecule was not found by any of the predicting methods. In all cases, non-native arrangements of tryptophan and proline side chains were proposed.



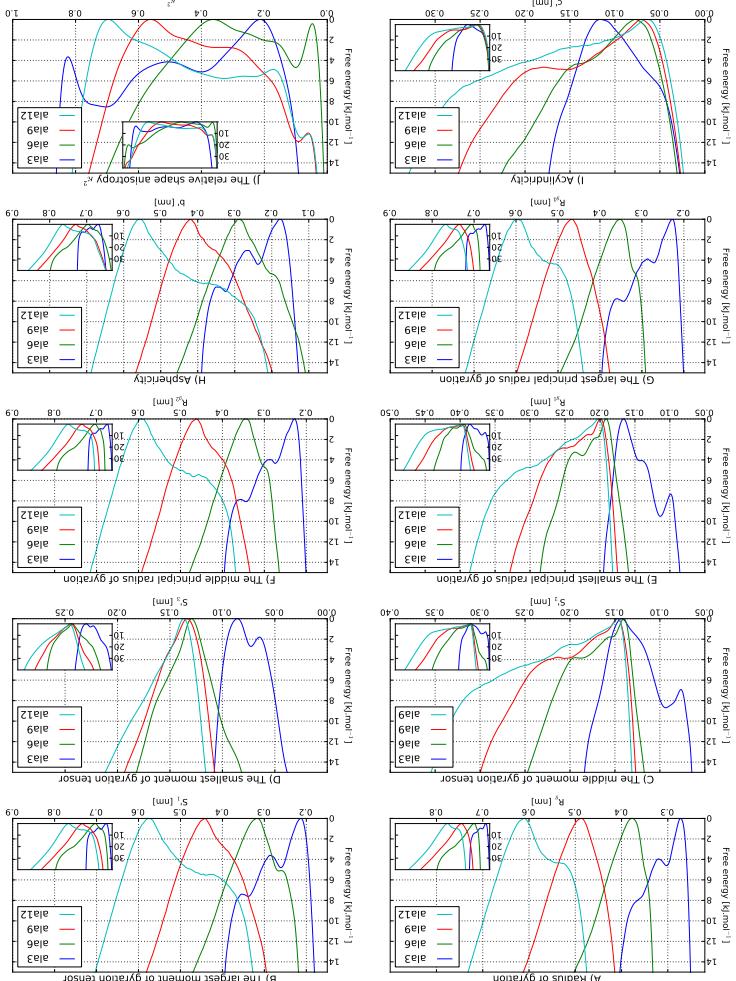
Obrázek 2: Profily volné energie alaninových peptidů (Ala₃₋₁₂).

Profily volné energie pro polylaninové peptidy Ala₃, Ala₆, Ala₉ and Ala₁₂. Jednotlivé grafy ukazují volnou energie jako funkci R_g (A), S'_1 (B), S'_2 (C), S'_3 (D), r_{g3} (E), r_{g2} (F), r_{g1} (G), b' (H), c' (I) and κ^2 (J).

$S_1^z(C)$, $S_3^z(D)$, $r_{g3}^z(E)$, $r_{g2}^z(F)$, $r_{g1}^z(G)$, $b(H)$, $c(I)$ and $f(J)$.
Ala12. The individual plots show the free energy as a function of $R_g^z(A)$, $S_1^z(B)$,
The free energy profiles obtained for polyalanine peptides Ala3, Ala6, Ala9 and
Ala12.

Figure

2: Free energy profiles of alanine peptides (Ala3-12).



Aminokyselínová sekvence retro Trp-cage miniproteínu býla vytvořena obrácením sekvence Trp-cage, syntetizovaná a dalej pak charakterizována násimi solupravovoulným spektroskopie NMR a CD spektroskopie. Oba metody potvrzily, že retro Trp-cage je esterifikovaný za bezúčich podmínek ve standardním pufu. Změny ve spektru

o aromaticitě Použitých metod nebo kohesiálních součiných o diskutabilní, zda mají násaký význam pro solupravování Trp-cage nebo se jedná cisté struktury s mimořádně slaboujm jádrem měničem RMD (2-4A). Je dobré 3. Kromě takto podobných struktur s nízkým RMD bývá často nazýván i podobností takto získaných konformací s experimentální strukturou je vedená na aktivní stav miniproteínu, který je stabilní až do konce simulačce. Uzavření téměř ve všechn simulacích, které byly limitovány délku 200ns, jakmile došlo ze začátku kohesiálnich součiných. To se však ukázalo jako ištěše, neboť bylo dosaženo Mohem následující testem bylo solupravování miniproteínu Trp-cage pouze za pomoci interpretovat různé tvary a symetrii.

Tenzorů nebo tenzoru extravacnosti. Přesto jsou všechny dleky své schopnosti provázané a představují pouze odlišné transformace tritiatučného navazání továry lepší výsledky než jiné. Dívalo dem je fakt, že jsou ve vše podstatně gyračního Nepozorovali jsme, že by kombinace těchto kohesiálních součiných Peptidu, které symetrii helixu konstantní, jak lze správně vidět v panelech C, D, E a I. Peptidu, jak je patrné na obr. 2 v panelech A, B, F, G, H a j pro R_g^z , S_1^z , r_{g3}^z a c musí zůstat dleky cylin-rg2, b a rg2. Na druhou stranu však S_2^z , S_3^z , r_{g1}^z , r_{g2}^z , r_{g3}^z a b se projevila v nekterých případech posunem minimu pro jednotlivé delky helixu se rozpoznatelného minimum na jednotlivých profilech. Rosouci délka stavuje jediné preferovány alélie jako svýs nestabilníjší konformér, který prodepeptidy jasné minima odpovídají jednoduchým rodičím konformermi. Ostatní výkazovaly jasné minima odpovídají jednoduchým rodičím konformermi. Peptidy výkazovaly flexibilitu peptidu. Pouze nejkratší Alas peptid poskytuje kritický, které poukazuje na obr. 2. Hlídáte a nepeptid informativní průběhy profilu jasné jak je už zaznamenalo prohly vole energie pro jednotlivé peptidy a kohesiální součiné, dimenzionální profilem. Bias exchange metody namísto poskytla jedno-dleky -3, 6, 9 a 12 aminokyseliny. Nasim prvními testovacími molekulami býla sada polyalanomyčch peptidů různých délky -3, 6, 9 a 12 aminokyseliny. Bius exchange metody namísto poskytla jedno-

a miniproteínu. V této studii jsme do počítače kódů pro metody namísto přidali kohesiální součinu (S_1^z , S_3^z , r_{g1}^z , r_{g2}^z , r_{g3}^z , b , c , f) a relativní an-

isotropie (f_z). Tento kohesiální součinice jsme poté testovali v simulaci peptidunice popisující všechny dleky, jmenovitě komponenty gyračního tenzoru

4.2 Metody namíka v kohesiálnich součinách odvozených od gyračního tenzoru a tenzoru extravacnosti.

V této studii jsme do počítače kódů pro metody namísto přidali kohesiální součinu

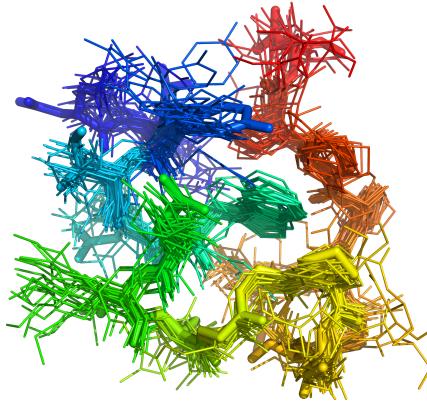
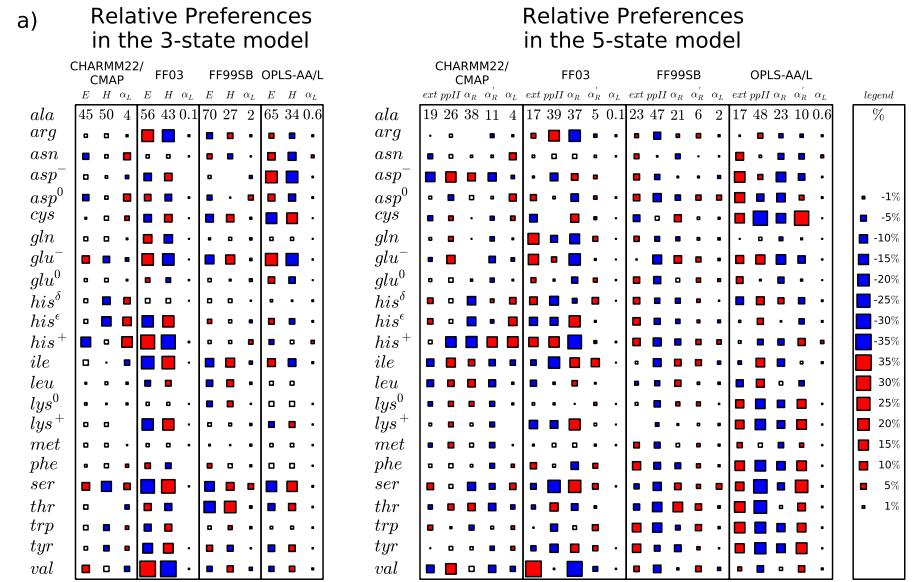


Figure 3: Superposition of Trp-cage structures.

The native-like structure obtained from simulations are compared with experimental structure of Trp-cage (thick).

In order to shed light on stability and character of disordered state of retro Trp-cage in water environment, we conducted a series of equilibrium MD simulations. All simulations started from the experimental structure (2LUF) but used different force fields in order to assess reproducibility of results. All four force fields involved in our study (ff03⁵³, ff99SB-ILDN⁵⁴, charmm22/CMAP⁵⁵ and OPLS-AA/L⁵⁶) provided different results in the simulations. Four independent runs for each force field manifested high diversity in maintaining structural features of retro Trp-cage and related properties. Nevertheless, all force field were able to keep, at least transiently, the fold of the miniprotein in water as proved by C_α-RMSD. This fact indicates that force fields are able to capture the stabilizing interactions found in experimental NMR structure. The helix favoring ff03 and charmm22/CMAP force fields kept higher helical content than ff99SB-ILDN and OPLS-AA/L, but all force fields provided similarly compact structures.

Since retro Trp-cage forms a stable structure only upon addition of TFE in buffer we tested the effect of this cosolvent on stability of miniprotein in simulations. The structures simulated by ff99SB-ILDN were dramatically affected by TFE model parametrized in general amber force field by usual procedure. The corresponding C_α-RMSD from experimental structure increased as well as radii of gyration. This changes indicated the rather denaturing effect of TFE on protein structure in clear disagreement with experiment. It is not obvious what caused the incorrect behavior in the solvent mixture. The probable explanations could be incompatibility of force field parameters for protein, water model and TFE cosolvent, which we decided to



Obrázek 1: Relativní četnost konformerů.

Relativní četnosti byly spočteny (v %) jako rozdíl mezi populacemi, které vykazovala daná aminokyselina a alaninem modelovaným stejnými parametry. Absolutní četnosti pro alanin dipeptid jsou uvedeny v prvním řádku. Pro ostatní aminokyseliny je rozdíl od alanin dipeptidu ukázán graficky jako obsah čtverečku, pokles jako modré a navýšení červeně. (bezbarvě neprůkazné hodnoty)

only for parameters with the most polar partial charges. Values of calculated $G_{T^{\star}}$ comparable to the experimental counterparts were found for the visual demonstration of self-aggregation in Fig. 5. Nevertheless, the Kirkwood-Buff integral calculation of O-H bonds, the lower values of $G_{T^{\star}}$ molecules were observed—the higher polarity of O-H bonds, the lower values of both liquids and no meaningful KBLs could be calculated from such simulations. The clear trends were observed on hydrogen atom transfer separation of both liquids with low partial charge on hydrogen atom transfer behavior. This TFE models experiments, several parameters showed the opposite behavior. These TFE models their properties. Although TFE is fully miscible with water in all concentrations in TFE/water mixtures modeled by candidate TFE parameters differed significantly in TFE/water mixtures modeled by candidate TFE parameters selected by a structure.

Both ways to calculate vaporization enthalpy led to parameters that were able to and tested in simulation of mixtures with water. Candidate parameters from all promising regions of parameter space were selected three were no unique parameters which could be chosen preferentially. Therefore 28 could not be clearly drawn without the extensive protocol we employed. This conclusions were considered, ambiguous and more precise solutions were found. This describes simultaneously all optimized properties. However, it the polarization costs describe simultaneously all optimized properties. However, it the polarization costs race. Zyxent teploj vedi k vice rovnomernemu vozrozhdeni konformacii pro-

terin. ⁵⁷ (enthalpy of vaporization, free energy of hydration) can be corrected by an additional energy costs for this pre-polarization, they are not transerable between phases. The condense phase charge distribution can be estimated and the affected properties polarization enthalpy. Since the partial atomic charges of ambe force fields minima field parameterization strategy we employed more precise approach to calculate very likely the same subsequently in simulations. In addition to the classical amber force field to predict the values of parameters in the best agreement with experiment and analytical forms and evaluated by an objective function. The analytic fit allowed 5 parameters. The dependence of 4 optimized properties was iteratively fitted in terms. ⁵⁷

The different combinations of parameters were tested exhaustively in simulations of pure TFE liquid. We developed a protocol for simultaneous optimization of all fluorine and hydrogen in hydroxyl group and a parameter of C-O-H torsion. Optimizatior-Lennard-Jones parameters (a and e) of fluorine, partial charges on liquid TFE in short simulations, namely—density, heat of vaporization, self-diffusion and Lennard-Jones parameters of atoms and tested their impact on properties of optimization were selected. We introduced perturbations to different partial charges 3 stages. Firstly, the parameters suitable for parameterization were scheduled to stages. Finally, the parameters with the most polar partial charges were selected to 3 stages. Finally, the parameters suitable for

4.4 Optimization of force field parameters for 2,2-trifluoroethanol

attempt as our next goal.

Necessitate jakehojli trenutni naprijed pouzivanymi parametry a fakt, ze je neps-

Pripavili jisme zcela arteficiální parameterizační, která výcházela z F99SB ale par-

kytý i velmi podobně variabilní F03 a F99SB se stala predmetem dalsho zkoumání.

Jeden z důležitých výsledků naší studie je záhy cen na obr. 1. Konformacii prefer-

minimální, jak potvrzil j jehoeficience.

není. Nicméně výsledky jednotlivých parameterizačí se velmi lišily a jich shoda je

funkcí poukazují na významné rozdíly konformacii preferenční pouze významy-

jepe výniky jednotlivé třídy. Tyto výsledky ziskané pomoci rizuny energetických

ence jednodílných amínoxyselin jsou vztazeny k preferenciám alginu dlepeptidu, aby

struktu v simulaci, nemělo výrazný vliv na hodnoty chemických posunů.

Simulace molekulární struktury, které produkovala býly ověřeny po-

baleném dienaturovaném stavu. V simulaci nízky neziskal naprostou prevalu.

polypropylene konformericu PP, o kterém se předpokládá, že preferovaná v nes-

dulkovála podobou obecnou odpověd. Nevedla ani k postnatálnu zvýšení obsahu

toru a peptidy tak významný vice výrovaře preferencie. Mimořádná výška nei-

race. Zyxent teploj vedi k vice rovnomernemu vozrozhdeni konformacii pros-

lozené náhodné pro zádušnu amínoxyselinu. Přitom výška závisí na způsobu denatu-

raci. Konformacii preferenční podmínek ziskané ze simulací nějso u v-

teinovým sekvenčím.

Konformacii stavu proteinovoho reťezce než glycin, tak se vice blízí reálnym pro-

teiny než bezužívající GCGC. Postřímaní reťezce alginu daleko vice meziú-

zvolený protot, že představují model pro nešablane nebo denaturované pro-

i tepelné (498K) a chemický (8M močovina) denaturující. AXAXA peptidy býly

modelovány především AXAXA za rizuny podmínek – jedná fyzioligických, tak

V další studii jsme sledovali konformacii preferenční amínoxyselinu na pozici X v

rozdílu v konformacii preferenční.

a nekonzistentní I-4 metrakce se tak nazývá jako možné vysvetlit pozorovaných

není, stejně jako ježíšskálová rizuny konfidenčnosti. Rozdíly nazývají distančne

I-4 elektrostatické interakce. Tyto interakce výška nejsou zcela fyzičké oddíval-

erance. Rozdíly nazývají mohou vlivem pozitivního tozini potenciály prostejnictvím

z F03. Tím stejné jednodušně prokazat vlivu parciálních nazývají na výsledné pre-

ciální nazývají převazal z F03. Výsledné preferenční silné korelaci vlastnosti zemí ziskanými

fyziky i větší podobné variabilní F03 a F99SB se stala predmetem zkoumání.

Nedostatečné jakehojli trenutni napříed používanými parametry a fakt, že je

se sice jednotlivé třídy. Tyto výsledky ziskané konformacii preferenční rizuny energetických

ence jednodílných amínoxyselin jsou vztazeny k preferenciám alginu dlepeptidu, aby

struktu v simulaci, nemělo výrazný vliv na hodnoty chemických posunů.

Simulace molekulární struktury, které produkovala býly ověřeny po-

baleném dienaturovaném stavu. V simulaci nízky neziskal naprostou prevalu.

polypropylene konformericu PP, o kterém se předpokládá, že preferovaná v nes-

dulkovála podobou obecnou odpověd. Nevedla ani k postnatálnu zvýšení obsahu

toru a peptidy tak významný vice výrovaře preferencie. Mimořádná výška nei-

race. Zyxent teploj vedi k vice rovnomernemu vozrozhdeni konformacii pros-

lozené náhodné pro zádušnu amínoxyselinu. Přitom výška závisí na způsobu denatu-

raci. Konformacii preferenční podmínek ziskané ze simulací nějso u v-

teinovým sekvenčím.

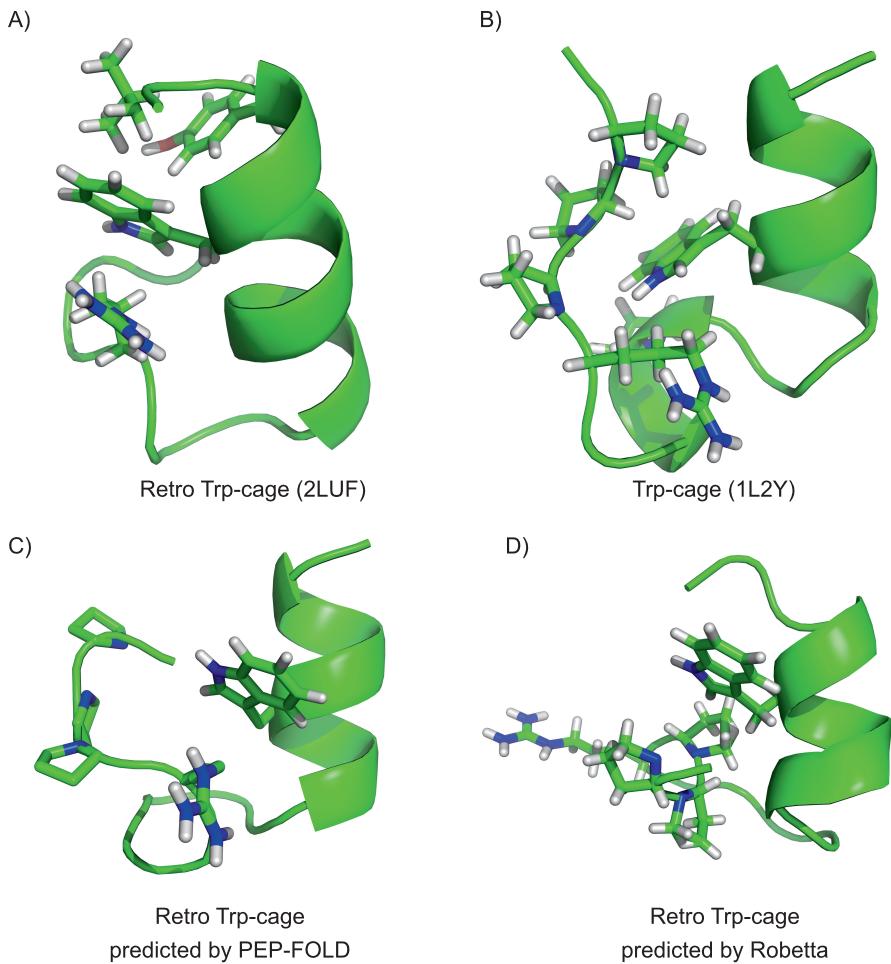


Figure 4: The experimental structure of retro Trp-cage (A), Trp cage (B) and retro Trp-cage models (C, D).

s nezápornými vlastními hodnotami. Vlastní vektory pak odpovídají důležitým osám molekul, jako jsou osy opisujícího elipsoidu nebo soustava os bez deviačních momentů setrvačnosti. Tři vlastní hodnoty S_1, S_2, S_3 se poté vztahují k délce příslušných elliptických poloos a I_1, I_2, I_3 představují momenty setrvačnosti kolem těchto os. Jednotlivé momenty setrvačnosti popisují rozmístění hmoty ve směrech kolmých k dané ose, což pro dostatečně homogenní tělesa vypovídá i o jejich velikosti. Ta je však lépe zachycena poloměrem setrvačnosti:

$$r_g^{ax} = \sqrt{\frac{I_{ax}}{m}}. \quad (9)$$

Tvar popisující veličiny, navržené Theodorouem a Suterem⁴⁵ – asphericita b a cylindricita c vyjadřují odchylku od symetrie koule a válce. Interpretace další veličiny – relativní anizotropie κ^2 také zohledňuje symetrii. Má hodnotu 0 pro vysoce symetrická prostorová tělesa a dosahuje 1 pro lineární útvary. Plochá tělesa poskytují hodnoty blízké $\frac{1}{4}$.

Mikroskopická struktura kapalných roztoků může být analyzována prostřednictvím Kirkwoodovy-Buffovy teorie.⁴⁶ Její unikátní vlastností je fakt, že hodnoty Kirkwoodových-Buffových integrálů mohou být vypočteny z počítačových simulací, stejně jako z makroskopických termodynamických veličin. Proto tato teorie představuje cenný most mezi atomistickými modely a experimenty.⁴⁷ Kirkwoodovy-Buffovy integrály mohou být získány ze simulací integrací radiálních distribučních funkcí:

$$G_{ij} = \int_0^\infty 4\pi r^2 (g_{ij}(r) - 1) dr, \quad (10)$$

kde g_{ij} je právě radiální distribuční funkce mezi částicí i částicemi j .

4 Výsledky a diskuze

4.1 Konformační preference aminokyselin

Chování různých parametrů energetických funkcí na modelové molekule alaninu dipeptidu bylo cílem našeho zájmu již dříve.⁴⁸ Ukázali jsme, že parametrisace založená na kvalitních *ab initio* výpočtech zlepšuje shodu s experimentem.⁴⁰ Později jsme náš zájem rozšířili na všechny aminokyselinové dipeptidy s cílem zhodnotit rozdíly ve čtyřech používaných sadách parametrů.⁴⁹ Náš zájem o konformační preferenci aminokyselin za různých podmínek vyústil ve spolupráci se skupinou prof. Daggett a společnou studii která zkoumala vliv denaturace na preferenci aminokyselin v modelových peptidech.

The properties of finally selected model were thoroughly characterized. The pure liquid TFE properties were reproduced excellently. Not only these involved in parameterization but also shear viscosity, coefficient of thermal expansion, isothermal compressibility, static dielectric constant and free energy of hydration were predicted by TFE model in good agreement with 2 advanced water models (TIP4P/Ew and TIP4P/2005) showed clearly the importance of polarization correction in course of parameterization process. The implicit correction to polarization cost in gas phase significantly improved reproduction of liquid properties of TFE. More importantly, it allowed to balance TFE-TFE, TFE-water and water-water interactions and thus reduce the self-aggregation of TFE molecules at level supported by experimental data.

Obtained with TIP4P/2005 water model, $\Delta G_{\text{sol},\text{a}}^{\text{exp}}$ and ΔH_{exp} represent a version without correction on polarization costs in gas phase. The value in parentheses represents a version without correction on polarization manuscript.

References to the original experimental studies are provided in the enclosed manuscript.

property	simulation	exp ^a
$D[10^9 \text{m}^2 \cdot \text{s}^{-1}]$	0.70	0.68
$p[\text{kPa.m}^{-3}]$	1382.5	1382.4
$\Delta H_{\text{exp}} [\text{kJ.mol}^{-1}]$	42.9(47.5) ^b	42.9
$\%trans$	39.9	40
$a[10^3 \text{K}^{-1}]$	1.27	1.28
$\beta_{\text{T}}[\text{GPa}^{-1}]$	1.32	1.23
$\eta[10^3 \text{kg.m}^{-1}. \text{s}^{-1}]$	1.52	1.72
$\Delta G_{\text{sol},\text{a}}^{\text{exp}} [\text{kJ.mol}^{-1}]$	42.9(47.5) ^b	42.9
$\Delta H_{\text{exp}} [\text{kJ.mol}^{-1}]$	42.9(47.5) ^b	42.9

The density p , self-diffusion coefficient D , enthalpy of vaporization ΔH_{exp} , ratio of the molecular volume (η) and free energy of hydration $\Delta G_{\text{sol},\text{a}}$ are compared for predicted trans conformer in liquid, coefficient of thermal expansion α , isothermal compressibility β_{T} , shear viscosity η , static dielectric constant ϵ , average dipole moment of the molecule (a) and free energy of hydration $\Delta G_{\text{sol},\text{a}}$.

Table 1: The properties of pure liquid TFE at 298.15K and 100 kPa. Comparison of the predicted and experimental values.

Všechy součty jsou provedeny přes N atomů s homotomostí m_i . Kartezské souřadnice x_i , y_i až z_i musí být vztaveny ke geometrickému středu nebo težišti molekuly. Pro-
tože jsou oba tenzory symetrické a pozitivní definitní, mohou být diagonalizovány

$$\mathbf{I} = \left(\sum m_i(y_i^2 + z_i^2) \quad \sum m_i(x_i^2 + z_i^2) \quad \sum m_i(x_i^2 + y_i^2) \\ \sum -m_i x_i y_i \quad \sum m_i(x_i^2 + z_i^2) \quad \sum m_i(x_i^2 + y_i^2) \right) . \quad (8)$$

$$\mathbf{S} = \frac{1}{N} \left(\sum x_i^2 \quad \sum y_i^2 \quad \sum z_i^2 \\ \sum x_i y_i \quad \sum y_i z_i \quad \sum x_i z_i \right) , \quad (7)$$

Gyraciální tenzor (\mathbf{S}) a tenzor sekvacenosti (\mathbf{I}) popisují rozložení hmoty v molekule, které je uvedeno polohou atomů:

pozice jedinu z možných deskriptoriů odvozených z gyračního tenzoru.
meru.⁴³⁻⁴⁵ Efektiivní věklosti lze poslat pomocí gyračního polomeru, ten je však moct jít jen výškou věklosti a tvaru. Ten to přistup má dlehou tradiční ve fyzičce poly-
Váribelní a flexibilní molekuly, kterými jsou i polypeptidy, lze charakterizovat po-
tvrz v každe z nich.

biás exchance metadyynamika³⁹ využívá replik, které se liší v použitých kolokviumech pros-

sobě mětodaři, jako je např. výměna replik (replica exchange). Takto vzniklá kvalita výsledků. Tyto nedostatky mohou být potlačeny kombinací mětody dynamicky dokáže metadyynamika urychlit, což se projeví jako hystrereze nepřizpůsobné ovlivňující pokud něj současně poskytuje všechny rozhodující pomalé módy. V tomto případě je ne-
Nevhodná volba kolokviumich součadnic může způsobit všechno problém, především dimenzionální.

- Jelikož poset by měl být malý, neboť efektivita metadyynamiky klesá s jejích rozdílností, reaktnit, produkty a meziproduktu reakci konformity, struktury, fázové, reaktnit, produkty a meziproduktu reakci konformity, struktury,
- Musí poslat dležité pomalé módy, ktere limitují rychlosť studovaných jevů.
- Musí být schopné rozložit dležité stavů, jako různé konformer, struktury,
- Musí být spojitu a diferencovatelnou funkci součadnic.

Spřavná volba kolokviumich součadnic je nezbytná pro funkčnost metadyynamiky a její uspěšné nasazení na daný problém. Kolokvium součadnice by měly splňovat tyto kritéria:⁴²

• Spřavná volba kolokviumich součadnic je nezbytná pro funkčnost metadyynamiky a podporí tak návštěvu ostatních oblastí kolokviumich součadnic.

• Spřavná volba kolokviumach součadnic je nezbytná pro funkčnost metadyynamiky a adaptivní upravován tak, aby desabilizoval aktuální stav sys-
temu a podporil tak návštěvu ostatních oblastí kolokviumich součadnic.

• Spřavná volba kolokviumach součadnic je nezbytná pro funkčnost metadyynamiky a potenciál, který je adaptivní přidomousti přidomouše potenciálu (bias potential), tyto sily jsou využívají elektivity vztahování konfor-

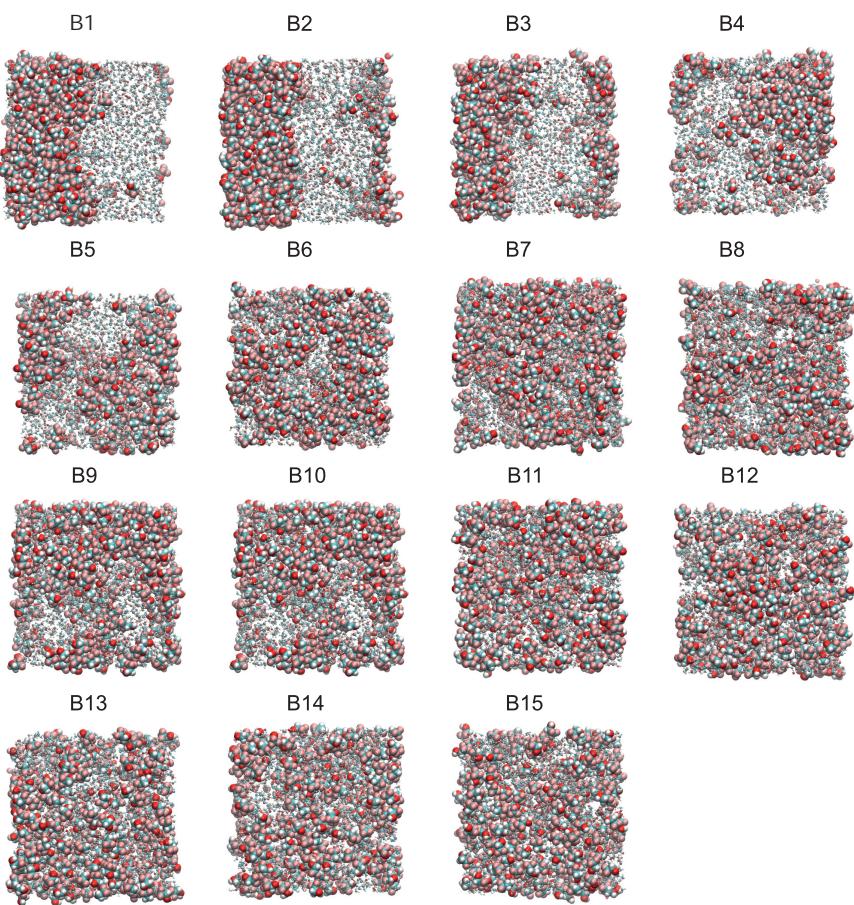


Figure 5: Self-aggregation of TFE molecules.

The different TFE force field parameters manifested distinct self-aggregation tendencies. The complete separation of both liquid phases was observed for models B1-B3. Models B4-B15 produced TFE clusters of decreasing size.

3 Metody

Funkce popisující potenciální energii (*force fields*) užité v této práci pro popis proteinů, peptidů a solventů využívají jednoduché matematické tvary. Je to především harmonický potenciál pro popis vibrací vazeb a deformací vazebných úhlů mezi atomy. Rotace kolem jednoduchých vazeb jsou popsány torzním potenciálem, který je dán některou formou kosínového rozvoje. Nevazebné interakce mezi atomy jsou popsány Lennard-Jonesovým 6-12 potenciálem a Coulombovým zákonem mezi parciálními náboji atomů. Kompletní forma potenciálu vypadá obvykle takto:

$$E = \sum_{bonds} \frac{K_B}{2} (r - r_0)^2 + \sum_{angles} \frac{K_A}{2} (\theta - \theta_0)^2 + \sum_{torsions} \sum_n \frac{V_n}{2} (1 + \cos(n\phi - \delta_n)) \\ + \sum_{i < j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon r_{ij}} \right), \quad (6)$$

kde r_0 a θ_0 představují rovnovážné délky vazeb a vazebních úhlů, K_B a K_A jim odpovídající silové konstanty. Torzní potenciál je definován složkami rozvoje s násobností n , fázovým posunem δ_n a velikostí V_n . Atomy oddělené více jak dvěma kovalentními vazbami na sebe působí nevazebnými interakcemi popsanými parciálními náboji q a Lennard-Jonesovými parametry A_{ij} and B_{ij} . Je běžné rozlišovat interakce mezi atomy oddelenými 3 vazbami (1-4 interakce) a ty částečně tlumit kvůli jejich síle a vlivu na rotační bariéry.

Navzdory své jednoduchosti, tyto matematické formy popisují dostatečně a robustně energetickou hyperplochu proteinů a peptidů, čemuž je důkazem např. jejich úspěšné využití při studiu sbalování proteinů. Ačkoliv energetické funkce sdílejí funkční tvary, mohou se znatelně lišit v jejich parametrech a způsobu, jakým byly získány.

Ruku v ruce s rostoucími možnostmi počítačových simulací díky rozvoji výpočetní techniky, zesílil i tlak na zpřesnění energetických funkcí. Ten se projevil nejdříve využitím *ab initio* výpočtů kvantové chemie. Nicméně jen ty nejkalitnější (a nejnáročnější) z nich mohou sloužit jako referenční hodnoty, které se vyrovnejí experimentálním datům.⁴⁰ Proto se v poslední době klade více důraz na přímé použití experimentálních dat jako NMR chemické posuny nebo vazebné konstanty pro kalibraci parametrů.³⁴

Metadynamika⁴¹ je metoda k výpočtu profilů volné energie, dokáže však také efektivně urychlovat pomalé procesy v molekulární dynamice. Podstatou celého algoritmu dimenzionální redukce. Metadynamika předpokládá, že cílový proces nebo systém může být popsán malým počtem parametrů – kolektivních souřadnic. Ty však musí být předem známy. Naopak nejsou vyžadovány žádné předpovědi profilu, který má metadynamika vytvořit. V průběhu simulace působí na vybrané

This thesis deals with several topics highly relevant for simulations of peptides and their most important results.

- The force fields involved in this study are not able to reproduce fine intrinsinc properties of amino acids. The propensities obtained from simulations and conditions of thermal and chemical denaturation generate different conformations. They can be efficiently used for extensive sampling in bias exchange metadynamics.
- Reverse sequence of Trp-cage does not fold in water but acquire a structure in 30% solution of 2,2,2-trifluoroethanol (TFE). The resulting structure resembles the structure of Trp-cage but possesses different packing amino acids in the core. Robetta and PEP-FOLD tools succeeded in predicting the helix in the structure but failed to model native interactions within the core. The stability of interpretation was interpreted differently upon a choice of force field. However, the attempt to mimic closely the experimental conditions by standard GAF model of TFE led to behavior incompatible with experiments by standard GAF model of TFE.
- We exhaustively optimized force field parameters of 2,2,2-trifluoroethanol for larization costs of molecule in gas phase.

5 Conclusions

- We exhaustively optimized force field parameters of 2,2,2-trifluoroethanol for larization costs of molecule in gas phase.
- Better description of pure liquid properties and realistic behavior in TFE/water mixtures. We demonstrated that the improvement over the standard GAF better describes the properties and realistic behavior in TFE/water mixtures.
- We exhaustively optimized force field parameters of 2,2,2-trifluoroethanol for larization costs of molecule in gas phase.

and therefore they might provide incorrect picture of denatured or unfolded states. Local backbone conformations and shifts were shown to be less sensitive to the local backbone conformations and shifted from the reference state under physiological condition. The chemical shift of peptide ensembles of model peptides which are neither random nor uniform than protein residues used to determine sampling of alanine ires. This collective coordinates were capable of exhaustive sampling of alanine ires.

- The gyration- and inertia-tensor based collective coordinates for metadynamics.
- Reverse sequence of Trp-cage does not fold in water but acquire a structure in 30% solution of 2,2,2-trifluoroethanol (TFE). The resulting structure resembles the structure of Trp-cage but possesses different packing amino acids in the core. The helix in the structure but failed to model native interactions within the core. The stability of interpretation was interpreted differently upon a choice of force field. However, the attempt to mimic closely the interpretation was interpreted differently upon a choice of force field. The helix in the structure but failed to model native interactions within the core. The stability of interpretation was interpreted differently upon a choice of force field.

The structure of TFE is different from that of alkanes and it has some specific properties. It is able to dissolve some organic compounds and it is miscible with water. The solvation of peptides in TFE is different from that in alkanes and it is more pronounced in polar solvents. The solvation of peptides in TFE is different from that in alkanes and it is more pronounced in polar solvents.

- We exhaustively optimized force field parameters of 2,2,2-trifluoroethanol for larization costs of molecule in gas phase.
- Better description of pure liquid properties and realistic behavior in TFE/water mixtures. We demonstrated that the improvement over the standard GAF better describes the properties and realistic behavior in TFE/water mixtures.
- We exhaustively optimized force field parameters of 2,2,2-trifluoroethanol for larization costs of molecule in gas phase.

kyjy prilozeny k disertaci jako prilohy A-F. Kzade z temat • 1-5 bylo zpracovano jako samostatna vedecka studie publicovana nebo zaslane do impaktniho vedeckeho periodika. Vecchny tyto claneky a rukopisy

5. Optimalizovat parametry solventu pro polohu popis jeho smesnych vod-

4. Charakterizovat a reprodukovat vlastnosti noveho mimoportelu pomoc výpočet-

3. Navrhnut a otestovat kolektivni soutaznice pro metody namiknu peptidu.

2. Zjistit jak jsou popsaty konformaci preferenice amimokyselin a jak se tisk v roznych podminkach protized.

1. Srovnat parametry energetickych funkci pouzivanych pro peptidy a proteiny.

Výpočetni studie peptidu a mimoportelu jsou narodené predevšim kvůli dřema při- činám: i) nejisté nebo nedostatečné přesnosti energetickych funkci a ii) vysokym závadavkem na vztahu konformaciálnemu prostoru. Čílem práce bylo

2 Cíle práce

Simulace komplexních proteinů by mohou být náročné z hlediska jejich následné rychlosti konformace. Alternativními cestami, která se vzhledem použití takových komplikovaných konstrukcí, prozrazuje musí obsahovat detali, které současně stávají a jež ho nemůže zastavit, redukuje relevantních parameterů. 35-38 Ty jsou však casto specifické dimenzionální redukcí relevantních součástí, ta však musí být ziskána multidi- rodnitá pro sbalování proteinů, však upravily jen částech. Tedy proces sice mo- kuď poskytuje jedno-dimensionální součásti, které dokázaly tyto komplexní děje řešitne popsat. Mnoho sňazy bývá využíváno na hledání reakčních základních charakteristik takových součástí, které dokázaly tyto se však težšíte problematickou k nálezení takových součástí, které dokázaly tyto namísto, která dokáže elektrické vlny vylepit děje na určené reakční koordinátu. Tím dané pohyb významná je mimoportel pomocí metoda jake je metady-

protize pro správne vztahovati tecto souborū je treba simluovat mikroskopickou-e- kazové skaly.³³ Ale i přes to mohou být výsledky znehnadoceny nepresnosti ener- getickych funkci a ježich záhy mnedostatkum, např. prilisnou stabilizaci jistých konformací.³⁸

1 Úvod

Mnoho důležitých proteinů v živých organismech je nestrukturovaných. Proteiny, které postrádají pevnou terciální strukturu, byly nalezeny ve všech taxonomických říších. Nicméně se odhaduje, že jich obecně eukarytické organismy obsahují více (10-45%) než prokaryota. Nestrukturované jsou nejčastěji proteiny s regulační a signální funkcí, především spojené s translací a transkripcí. Na druhou stranu, nejméně nestrukturovanosti vykazují enzymy.^{1,2}

Nestrukturovanost stejně jako struktura jsou zakódovány v proteinové sekvenci a tudíž mohou být předpovídány.³ Aminokyselinové složení nestrukturovaných proteinů je typicky obohaceno o Ala, Arg, Gly, Gln, Ser, Glu, Lys and Pro a naopak ochuzeno o Ile, Leu, Val, Tyr, Trp, Phe, Cys and Asn. Sekvence nestrukturovaných proteinů nebo regionů vykazují statisiticky méně hydrofobních aminokyselin a přebytek nabitých. Obecně se usuzuje, že tyto sekvence nemají energetickou hyperplochu s hlubokým minimem jako strukturované proteiny, ale spíše populují řadu isoenergetických stavů.⁴

Sbalené proteiny se nachází v neustálé rovnováze s nesbaleným stavem, ale ta je za fyziologických podmínek nakloněna k tomu sbalenému. Tato rovnováha však může být změněna za jiných termodynamických podmínek, jako je teplota nebo tlak. Podobný efekt může mít přítomnost dalších chemických látek v prostředí. Látky, které destabilizují sbalený stav, se nazývají *denaturanty* a ty s opačným účinkem jako *ochranné osmololyty*.⁵

Močovina a guanidinium chlorid patří k neznámějším denaturantům. Tanfordovy ranné výzkumy ukázaly, že proteiny v koncentrovaném roztoku guanidia zvětšují svůj hydrodynamický poloměr takovým způsobem, že by mohly mít náhodnou konformaci (*random coils*).⁶ Tyto výsledky však byly nevhodně zobecněny na jakýkoliv denatuovaný stav. Ve skutečnosti se ukázalo, že tyto víceméně náhodné konformace jsou velmi vyjímečné a denatuované proteiny spíše zůstavají kompaktní s velkým množstvím zbytkové struktury.⁷ Za slabých denaturačních podmínek může dokonce zůstat i sekundární struktura proteinu – tento stav se pak nazývá *molten globule*.⁸

Ochranné osmololyty mají přesně opačný účinek. Pomáhají stabilizovat nativní stav a zachovávat funkčnost proteinu i za nefyziologických podmínek. Vyskytují se přirozeně u organismů, pokud jsou vystaveny stresu nebo působení denaturačních látek.⁹ Mimo přírodní osmololyty jsou v laboratorní praxi používány syntetické látky s podobnými účinky, jako je např. trifluorethanol (TFE) and hexafluor-2-propanol (HFIP).¹⁰

Pufry obsahující TFE (až do 50% v/v) obvykle stabilizují α -helixy a nebo přímo indukují jejich tvorbu. TFE také zlepšuje rozpustnost proteinů a peptidů, a proto je využíván v experimentech.¹⁰ Mechanismus, jakým TFE stabilizuje strukturu, není úplně známý. Navržené hypotézy se opírají buď o přímé nebo nepřímé efekty v roztocích. Mezi ty přímé patří vazba na povrch peptidu nebo proteinu, která

může lokálně zvýšit koncentraci a stabilizovat vodíkové vazby v proteinech skrze snížení efektivního dielektrika prostředí.^{11,12} Nepřímými mechanismy mohou být chaotropní efekt na solvatační obal a různá síla tohoto efektu ve sbaleném a nesbaleném stavu.¹³ V každém případě může hrát roli samoshlukování molekul TFE.¹⁴

Krátké peptidy se tradičně používají jako experimentální a teoretické modely pro nesbalený, denatuovaný či nestrukturovaný stav proteinu.^{15,16} Různé studie prokázaly, že konformace těchto peptidů nejsou náhodné nebo uniformní, jak se po léta předpokládalo díky modelu *random coil*¹⁷. Bohužel však nemohou být tyto konformační stavy plně a bez rozporu charakterizovány žádnou současnou experimentální metodou. O těchto stavech se však předpokládá, že mohou hrát důležité role v procesu sbalování proteinu a také tvořit reziduální strukturu nesbalených proteinů.¹⁶ Různé trendy a preference však mohou být pozorovány spektroskopicky pro různé aminokyseliny v modelových peptidech jakou jsou GGXGG,¹⁸ GXG,¹⁹ AXA²⁰ nebo GPPXPPGY.²¹ Ukazuje se však, že i samotné aminokyseliny ve formě dipeptidu vykazují různé strukturální preferenze.^{22,23}

Struktura sbalených proteinů má pevná pravidla. Zahrnují síť vodíkových vazeb mezi prvky sekundární struktury a také nativní kontakty mezi postranními řetězci aminokyselin. Peptidová vazba může vystupovat jako donor i akceptor vodíkové vazby, ale toto pořadí je pevně dánou směrem polypeptidového řetězce. Nicméně téměř dokonalý symetrický obraz proteinu může být vytvořen z klasických L-amino-kyselin, pokud:²⁴

1. Směr polypeptidového řetězce je obrácen. To znamená, že sekvence proteinu je čtena z C-konce místo N-konce.
2. Pozice N-H a C=O skupiny peptidové vazby je prohozena. Tím pádem je i směr vodíkových vazeb.
3. Torzní úhly proteinové páteře se transformují dle pravidla: $\phi \rightarrow -\psi$ and $\psi \rightarrow -\phi$.

Tyto podmínky snadno splní α -helix, když si uvědomíme, že zmíněné torzní transformace jsou pro něj téměř invariантní (jak se lze přesvědčit na Ramachandranově diagramu).

Obrácení proteinové sekvence již bylo experimentálně testováno. Tato „retro transformace“ buď vedla ke ztrátě schopnosti se balit,^{25,26} nebo byla zachována a vznikla analogická (případně topologicky podobná) struktura.^{27–29}

Nejpodrobnejší pohled na procesy sbalování proteinu, denaturace nebo nestrukturované stavby může být získán počítačovými simulacemi.³⁰ Jejich spolehlivost však významně závisí na použitých energetických funkcích (*force fields*) a jejich parametrech. Současné běžně užívané parametry jsou schopny popsat reverzibilní sbalování a rozbalování modelových proteinů v simulacích molekulové dynamiky,³¹ bohužel se však často liší v detailech předpovězených mechanismů a termodynamických veličinách.³² Obzvláště simulace nestrukturovaných proteinů jsou velmi náročné,