

POVZETEK

V diplomskem delu so predstavljeni matematični modeli in algoritmi zvijanja proteinov. Opisana je zgradba proteinov ter metode in orodja za določevanje molekulske strukture z najnižjo energijo. Celotna energija molekule je opisana s približnimi rešitvami Schrödingerjeve enačbe in jo je mogoče določiti tako z *ab initio* metodami kot z molekulsko mehaniko. Ravnovesje sil in dinamiko molekule opisuje diferencialna enačba, ki se imenuje Langevinova dinamika. Pri nizkih temperaturah je celotna energija sistema potencialna. Geometrija molekule pa je definirana z minimizatorjem površine potencialne energije. Le ta je polna lokalnih minimumov, tako lahko molekula prehaja v različna stanja. S pomočjo harmoničnih aproksimacij je opisano visokofrekvenčno gibanje molekul. Ker se algoritmi, ki iščejo rešitve diferencialnih enačb, izvajajo prepočasi, v diferencialno enačbo vpeljemo omejitve. Natančneje so opisani algoritmi za omejeno gibanje molekule in postopek za modeliranje potenciala. Podrobneje so obrazložena vsa orodja in rezultati programskega paketa CHARMM. Na koncu pa je prikazan še primer njegovega delovanja na proteinu BPTI.

ABSTRACT

In my diploma I focused on mathematical models and algorithms for protein folding. I described the structure of proteins and the methods and tools for determination of molecular structure with lowest energy. Total energy is described as an approximate solution of Schrödinger's equation. We can determine this solution with *ab initio* methods or with molecular mechanics. Force balance and molecular dynamics are described with differential equation called Langevin dynamics. The total energy of the system is approximately equal to the potential energy when temperatures are low. The geometry is given as an minimizer of the potential energy surface. It is full of minima, so the molecule can transfer into different states. Harmonic approximation describes the high frequency motions. To speed the algorithms for searching solution of differential equation we impose certain constraints. I described algorithms for constrained molecular dynamics and potential modeling procedure. I specified all the tools and results of program CHARMM. I also showed an example how CHARMM works on protein BPTI.

Math. Subj. Class. (2010): 92E10

Ključne besede: računska kemija, zgradba proteina, zvijanje proteina, molekulska mehanika, energija, aproksimacija, minimizacija, omejeno gibanje, iteracija SHAKE, modeliranje potenciala, programski paket CHARMM, energijska funkcija CHARMM.

Keywords: computational chemistry, protein structure, protein folding, molecular mechanics, energy, approximation, minimization, constrained dynamics, SHAKE iteration, potential modeling, computer program CHARMM, CHARMM energy function.

prikažejo različne vrednosti povprečnih temperatur in temperaturnih nihanj.

Tabela 1 poda rezultate dinamike, kjer je toleranca 10^{-6} in časovni koraki $h = 0.0005, 0.001, 0.004$ ps.

časovni korak		SHAKE	SOR
0.0005	its	19.0	9.25
	reset	4.62 m.	2.25 m.
	total	30.6 m.	28.2 m.
	setup	0.07 s.	0.07 s.
	omega	1.00	1.17
0.001	its	21.02	10.4
	reset	2.8 m.	1.34 m.
	total	16.3 m.	14.8 m.
	setup	0.07 s.	0.07 s.
	omega	1.00	1.19
0.004	its	23.06	15.3
	reset	0.872 m.	0.463 m.
	total	4.98 m.	3.76 m.
	setup	0.05 s.	0.07 s.
	omega	1.00	1.14

TABELA 1. Primerjava metod SHAKE in SOR za BPTI, $\text{tol} = 10^{-6}$.

Vrstica z imenom setup prikazuje čas ponovnega generiranja matričnih podatkovnih struktur. Celoten čas reševanja nelinearnih enačb (ponastavljanje koordinat) med potekom simulacije je podan v vrstici reset, celoten čas delovanja CHARMM je podan v vrstici total. Povprečno število iteracij na časovni korak je podano v vrstici its. Za shemo SOR je podan tudi končni relaksacijski parameter ω . Največje razlike glede na velikost časovnega koraka se pojavijo v vrstici reset.

Vsi algoritmi, ki uporabljajo omejene dolžine, podajo skoraj enake trajektorije. Razlike se pojavijo zaradi različnega tolerančnega parametra. Numerični rezultati kažejo, da se z omejitvami pri danem časovnem koraku zmanjša nihanje energije (tj. izboljša ohranitev energije) in poveča nihanje temperature oz. nihanje kinetične energije. Razmerje med nihanjem celotne in kinetične energije lahko uporabimo pri določitvi časovnega koraka za željeno stopnjo natančnosti simulacij.

LITERATURA

- [1] E. Barth, *PhD Thesis*, University of Kansas, 1994.
- [2] E. Barth, K. Kuczera, B. Leimkuhler, R. D. Skeel, *Algorithms for constrained molecular dynamics*, J. Comp. Chem. 16, 1192-1209, 1995.
- [3] Anže Lošdorfer Božič, *Uvod v zvijanje proteinov za fizike*, seminar, Fakulteta za matematiko in fiziko, 21. 03. 2008, zajeto oktober 2011.
- [4] H. J. C. Berendsen, G. Ciccotti, J. P. Ryckaert, *Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes.*, Journal of Computational Physics 23 (3), 327-341, 1977.

- [5] B. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, M. Karplus, *CHARMM: a program for macromolecular energy minimization and dynamics calculations.*, J. of Computational Chemistry 4, 187-217, 1983.
- [6] B. R. Brooks, J. W. Chu, B. L. Trout, *A super-linear minimization scheme for the nudged elastic band method.*, J. Chem. Phys., 119(24), 12708-12717, 2003.
- [7] A. T. Brunger, M. Karplus, *Proteins: Structure, Function, Bioinformatics*, 4, 148-156, 1988.
- [8] M. M. Cox, D. L. Nelson, *Lehninger Principles of Biochemistry, fourth edition*, WH Freeman, 116-147, 2005.
- [9] J. M. Cooley, J. W. Turkey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comput. 19, 297, 1965.
- [10] T. E. Creighton, *Proteins. Structures and Molecular Properties.*, WH Freeman and Company, New York, 1993.
- [11] J. Deisenhofer, O. Epp, H. Formanek, R. Huber, D. Kukla, A. Ruehlmann, W. Steigemann, *Shema BPTI v Brookhaven Protein Data Bank.*, <http://service.pdbj.org/mine/Display?-PDBID=4PTI&ARCHIVE=LegacyHeader>, 1982, zajeto december 2011.
- [12] D. B. Dix, *Mathematical models of protein folding*, University of South Carolina, <http://www.math.sc.edu/~dix/fold.pdf>, zajeto september 2011.
- [13] D. Eisenberg, L. Wesson, *Atomic solvation parameters applied to molecular dynamics of proteins in solution*, Protein Sci. 1, 227-235, 1992.
- [14] R. Fletcher, C. M. Reeves, *Function minimization by conjugate gradients*, Computer J., vol. 7, 149-154, 1964.
- [15] C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, England Cliff, NJ, 1971.
- [16] Iztok Grilc, *Modeli atomov skozi zgodovino*, seminar, Fakulteta za matematiko in fiziko, 2008.
- [17] M. H. Hao, S. C. Harvey, *Statistical termodinamics of protein folding: sequence dependence*, J. Phys. Chem. 98, 9882-9893, 1994.
- [18] http://en.wikipedia.org/wiki/History_of_molecular_biology#Protein_folding_and_first_structural_models, zajeto oktober 2011.
- [19] [http://sl.wikipedia.org/wiki/Računska_kemija](http://sl.wikipedia.org/wiki/Ra%C4%8Dunska_kemija), zajeto oktober 2011.
- [20] <http://www.pdb.org/pdb/home/home.do>, zajeto december 2011.
- [21] Peter Karlson, *Biokemija*, Ljubljana, Državna založba Slovenije, 36-63, 1980.
- [22] M. Levitt, S. Lifson, *Refinement of Protein Conformations Using a Macromolecular Energy Minimization Procedure*, J. Mol. Biol. 46, 269-279, 1969.
- [23] A. Neumaier, *Molecular modeling of proteins and mathematical prediction of protein structure*, SIAM review, Vol. 39, 407-460, 1997.
- [24] J. M. Ortega, W. C. Rheinboldt, *Iterative Solution od Nonlinear Equations in Several Variables*, Academic Press, 1970.

- [25] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N. J., 1980.
- [26] Steven Schultz, *Math helps explain protein folding*, <http://www.princeton.edu/prpwb/99/-0927/math.shtml>, zajeto september 2011.
- [27] R. Varga, *Matrix Iterative Analysis*, Prentice-Hall, 1962.
- [28] L. Verlet, *Computer Experiments on Classical Fluids*, Phys. Rev. Vol. 159, No. 98, 1967.
- [29] www.charmm.org, zajeto december 2011.