An efficient approach for the calculation of frequencies in macromolecules

An Ghysels

Supervisor(s): Michel Waroquier

I. Introduction

Conformational changes of macromolecules are essential in the understanding of e.g. proteins and drug design. The theoretical prediction is far from trivial, especially for large molecules. In many cases, collective motions are present which occur on a timescale ($\sim$ ms) that is too long to be accessible through molecular dynamics simulations. Normal mode analysis (NMA) has been proven succesful in exploring the potential energy surface (PES) within the harmonic oscillator approximation. The lowest frequency modes contribute the most to a conformational change. This paper presents a computationally attractive method that selects modes from the lower spectrum.

II. The Mobile Block Hessian approach

Until recently, NMA was limited to small proteins, because the required energy minimization is a computationally exhausting task. Secondly NMA requires the diagonalization of a $3N_a \times 3N_a$ matrix with N_a the number of atoms. As the system becomes larger, this too becomes an expensive calculation. A series of simplified models has been proposed. In particular the RTB (rotation-translation blocks) method deserves attention [1], [2]. It makes use of the concept that a peptide chain or protein can be seen as a subsequent set of rigid components, i.e. the peptide units. A peptide chain is thus divided into rigid blocks with six degrees of freedom each.

Recently we developed the Mobile Block Hessian (MBH) method [3], [4], which in a sense corresponds very much to the RTB method. The main difference is that MBH was developed to deal with partially optimized systems. The position/orientation of each block is optimized while the internal geometry is kept fixed at a plausible – but not necessarily optimized – geometry. This reduces the computational cost of the energy minimization. Applying the standard NMA on a partially optimized structure however results in imaginary frequencies, coordinate dependence and the wrong number of invariances of the Hessian (matrix with second derivatives of the PES). The MBH avoids these unphysical effects by taking into account energy gradient corrections. Moreover the number of variables is reduced, which facilitates the diagonalization of the Hessian.

In the original implementation of MBH, atoms could only be part of one rigid block. The MBH is now extended to the case where atoms can be part of two or more blocks. Two basic linkages can be realized: (1) blocks connected by one link atom, or (2) by two link atoms. The latter is referred to as the hinge type connection. The extension to linkages is implemented using a restraint technique and dummy atoms. This has the advantage that, by first duplicating the link atom (dummy atoms) and imposing a high coupling, the former MBH equations can be used without change.

A. Ghysels is Aspirant at the FWO and is with the Center for Molecular Modeling, Ghent University (UGent), Proeftuinstraat 86, Gent, Belgium. E-mail: An.Ghysels@UGent.be

III. Validation

A. Small test systems

The MBH method with linked blocks is tested on some basic test examples such as dimethylether and alkanes. The square overlap of the modes indicates in how far the system with reduced dimensionality (with blocks) can reproduce the original benchmark (full Hessian) frequencies and associated normal modes. The low frequency modes are reproduced by 87%-100%, which is largely sufficient.

B. Biological systems

In order to select an appropriate partitioning scheme for proteins, several block choices are tested as depicted in figure 1. The labels $[C_\alpha]^k$ indicate to how many blocks the central carbon atom belongs. For instance, in block choice $[C_\alpha]^2$ the carbon atom belongs to two blocks. Crambin, a polypeptide with 46 residues and 646 atoms, is taken as a test case.

The peptide bond (C=O-NH) is known to be fairly rigid, so peptide units are consistently treated as blocks. Block choice $[C_\alpha]^1$ was proposed by Tama et al.[1]; it has no linked blocks. The MBH results with linked blocks, as in block choices $[C_\alpha]^2$, $[C_\alpha]^3$, and $[C_\alpha]^4$ systematically overestimate the benchmark frequencies. This is explained by the fact that a constrained motion experiences higher resistance. The modes on the other hand are very well reproduced: the cumulative square overlap between MBH and benchmark modes shows that the lowest 50 modes are reproduced by over 90%, which is an improvement with respect to the original scheme $[C_\alpha]^1$ of Tama et al. Application of the hinge type connection, as in block choice [dih], selects the internal rotations around the N-C_α and C_α-C bonds (dihedral angles). Despite the rude approximation, the 20 lowest modes are reproduced by over 80%. As in practice mode following is only applied on the lowest five modes, this is again a satisfying result.

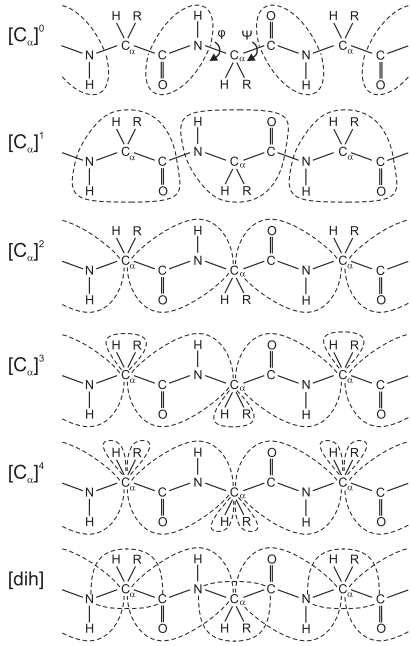

Figure 1. Block choices for proteins.

IV. Conclusion

The extension of the MBH to linked blocks allows to reproduce the low frequency modes of macromolecules while reducing the computational load.

Acknowledgments

This work is supported by the Fund for Scientific Research - Flanders.

References

[1] F. Tama, F. X. Gadea, O. Marques, and Y. H. Sanejouand, *Prot. Struct. Funct. Gen.*, vol. 41, no. 1, pp. 1–7, 2000.

[2] P. Durand, G. Trinquier, and Y. H. Sanejouand, *Biopol.*, vol. 34, no. 6, pp. 759–771, 1994.

[3] A. Ghysels, D. Van Neck, V. Van Speybroeck, T. Verstraelen, and M. Waroquier, *J. Chem. Phys.*, vol. 126, no. 22, pp. 224102, 2007.

[4] A. Ghysels, D. Van Neck, and M. Waroquier, *J. Chem. Phys.*, vol. 127, pp. 164108, 2007.