Software News and Updates
Carma: A Molecular Dynamics Analysis Program

NICHOLAS M. GLYKOS
Department of Molecular Biology and Genetics, Democritus University of Thrace, Dimitras 19, 68100 Alexandroupoli, Greece

Received 23 August 2005; Revised 11 December 2005; Accepted 11 March 2006
DOI 10.1002/jcc.20482
Published online 17 August 2006 in Wiley InterScience (www.interscience.wiley.com).

Abstract: A computer program has been developed to aid the analysis of molecular dynamics trajectories. The program is tuned for macromolecular large-scale problems and supports features such as removal of global translations-rotations of the solute, calculation of average distance maps and their corresponding standard deviations, calculation of the variance-covariance and cross-correlation matrices, and principal component analysis of trajectories with the added ability to create artificial trajectories based on selected eigenvectors. Limited graphics (trajectory viewing) capabilities are also available.

 J Comput Chem 27: 1765–1768, 2006

Key words: molecular dynamics; principal component analysis; essential dynamics; trajectory viewer

Introduction

The majority of molecular dynamics analysis programs available today usually come bundled with much larger simulation and visualisation packages. By being attached to specific suites of programs, their usability (in the form of stand-alone programs) is usually reduced. Additionally, several of these programs have a large memory footprint, making the analysis of large systems (with tens of thousands atoms) rather difficult. The design principle for the program *carma* presented in this communication was to produce a free open-source program based on a compiled language and with (i) a small memory footprint (mainly by reading the trajectory frame-by-frame and not as a whole), (ii) a simple command-line interface, (iii) the ability to be used as a stand-alone program supporting one of the most common trajectory formats (the DCD format), and (iv) the ability to graphically display a trajectory using relatively low level graphics calls (so that even an X-terminal can be used as a viewer).

Description of the Program

Carma is written in C and its only extraneous library dependencies are LAPACK (which is used for the eigenvalue calculations of the principal component analysis[1–6]) and the graphics library Ygl. On a GNU/Linux machine the stand-alone executable occupies only 115 kB. To further reduce its memory requirements, the program accesses the trajectory files frame-by-frame. The result of this choice is that for some of the calculations supported by the program up to three passes through the trajectory file are necessary (but note that if the physical memory is sufficient for holding the whole trajectory, the second and third passes may not need to access the disk, reading the data from the memory cache).

The current version of *carma* can process PDB, PSF, and binary DCD (trajectory) files. All interaction with the program is via command-line options and flags (with the exception of the program's graphics mode which is interactive). Depending on the calculation performed, the program produces formatted ASCII files containing numerical results, encapsulated postscript files containing images, or binary DCD files. Some of the calculations supported by the program are

- Removal of global rotations and translations of selected atoms or segments and calculation of the evolution of the rms deviation from a user-defined reference frame. This is achieved through the application of the least-squares superposition algorithm described by Kabsch.[7]
- Calculation and graphical representation of distance maps and average distance maps (and their corresponding rms deviations) for selected atoms and frames.
- Calculation of the variance–covariance matrix of the fluctuations of each of selected atoms from their average. The program can optionally produce the normalised (cross-correlation) matrix.
- Calculation of the eigenvalues and eigenvectors of the variance–covariance matrix.

Correspondence to: N. M. Glykos; e-mail: glykos@mbg.duth.gr

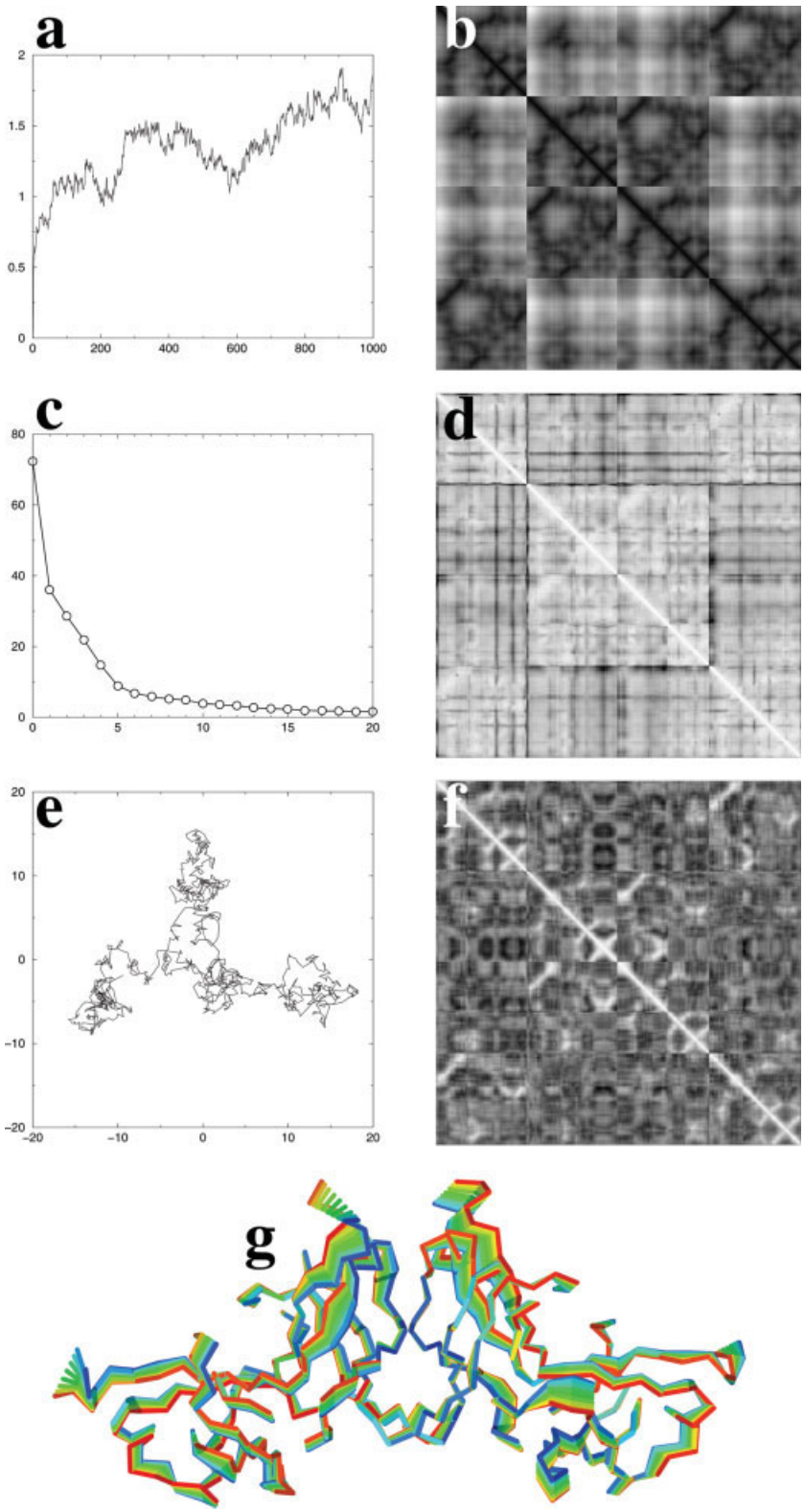

Figure 1. Examples of some of the results that can be obtained from the application of the program. (a) Evolution of the rms deviation (versus the starting structure); (b) Average C_α-C_α

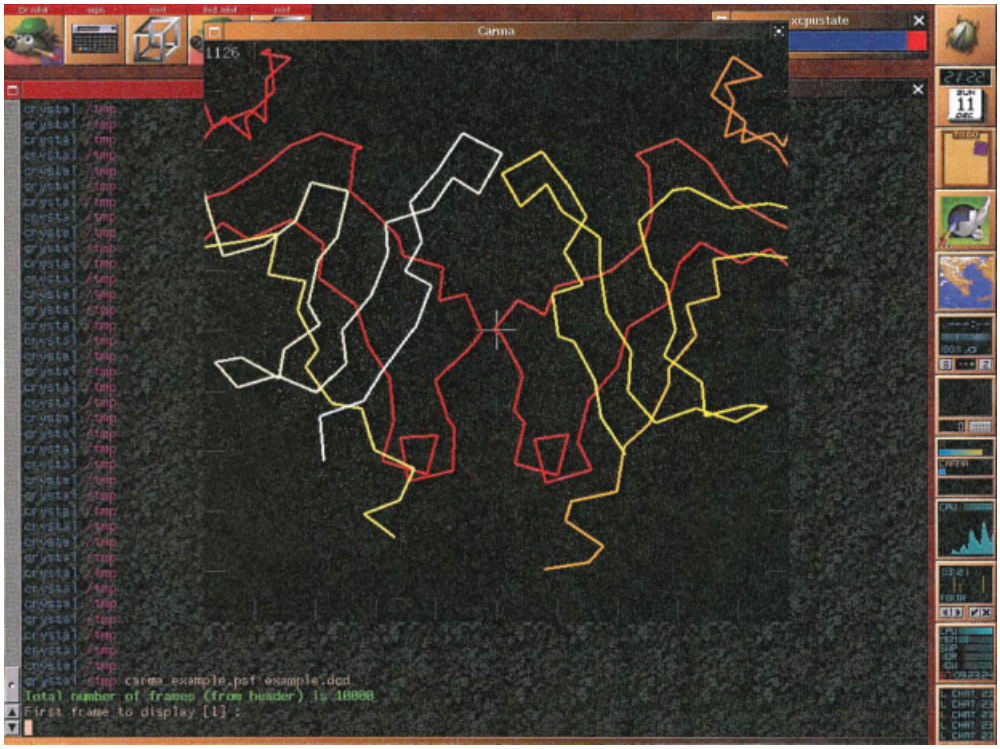

Figure 2. Screenshot from a GNU/Linux machine executing *carma* in graphics mode.

- Calculation of the projections of the atomic fluctuations along selected eigenvectors.
- Calculation of an artificial trajectory based on the projection of the input trajectory on selected eigenvectors.
- Calculation of an artificial trajectory containing a smooth representation of the motion due to one (selected) eigenvector.
- Construction of a trajectory file containing a set of structures that are derivable from a set of eigenvectors and eigenvalues.

Figure 1 shows a graphical representation of some of the results that can be obtained from the application of the program to the analysis of a macromolecular molecular dynamics trajectory. The results shown are from the analysis of a 12 ns molecular dynamics simulation of the HrcQb tetramer from *Pseudomonas syringae pv. phaseolicola* in explicit water (with a total of 59,428 atoms of which 4356 belong to the protein component). The most time consuming of the calculations performed by *carma* is the eigen problem of the principal component analysis, which may require up to three passes of the trajectory file. To give an indication of the program's reduced hardware requirements, an eigenvalue calculation for HrcQb was performed on a relatively old laptop computer equipped with a 650 MHz PIII CPU and 64 MB of physical memory. The whole calculation (for the 12 ns trajectory) took 15.3 min and the maximum physical memory consumption was only 6.3 MB.

Figure 2 is a snapshot from a GNU/Linux machine running *carma*, which illustrates the program's graphics capabilities.

Program Availability

Carma is free open-source software. The program's distribution includes source code, documentation, example scripts, and standalone executables for Irix, GNU/Linux, and windows (with graphics

Figure 1. (*continued*) distance map calculated over selected trajectory frames; (c) A graph depicting the magnitudes of the twenty largest eigenvalues; (d) A map of the rms deviations from the average C_α-C_α distances [compare with panel (b)]; (e) Projection of the C_α motion of the simulation on the plane of the two eigenvectors corresponding to the two largest eigenvalues; (f) Cross-correlation matrix; (g) A set of structures depicting the motion due to the eigenvector corresponding to the largest eigenvalue. Graphs (b), (d), and (f) have been directly produced by *carma*. Graphs (a), (c), and (e) were plotted with the program *xmgr*. The molecular graphics shown in panel (g) were produced with the program *rasmol*. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

support only for the GNU/Linux executable). The distribution is immediately available for download from the following world-wide web address: `http://www.mbg.duth.gr/~glykos/`.

References

1. Garcia, A. E. Phys Rev Lett 1992, 68, 2696.
2. Amadei, A.; Linssen, A. B. M; Berendsen, H. J. C. Protein Struct Funct Genet 1993, 17, 412.
3. Mongan, J. J Comput Aid Mol Des 2004, 18, 433.
4. Vriend, G. J Mol Graph 1990, 8, 52.
5. Lindahl, E.; Hess, B.; van der Spoel D. J Mol Model 2001, 7, 306.
6. Kendall, R. A.; Apra, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupius, M.; Fann, G. I.; Harrison, R. J.; Ju, J.; Nichols, J. A.; Nieplocha, J.; Straatsma, T. P.; Windus, T. L.; Wong, A. T. Comput Phys Commun 2000, 128, 260.
7. Kabsch, W. Acta Cryst 1976, A32, 922.