Rapid Communication
Reply to Comment on: "Fast Determination of the Optimal Rotational Matrix for Macromolecular Superpositions"

PU LIU,[1] DIMITRIS K. AGRAFIOTIS,[1] DOUGLAS L. THEOBALD[2]
[1]*Johnson and Johnson Pharmaceutical Research and Development, L.L.C., 665 Stockton Drive, Exton, Pennsylvania 19341*
[2]*Department of Biochemistry, Brandeis University, 415 South Street, Waltham, Massachusetts 02454-9110*

Received 18 May 2010; Accepted 23 May 2010
DOI 10.1002/jcc.21606
Published online in Wiley Online Library (wileyonlinelibrary.com).

© 2010 Wiley Periodicals, Inc. J Comput Chem 00:000–000, 2010

Key words: rotational matrix; superposition; quaternion; adjoint matrix

The rotational matrix that optimally superpositions two molecular structures can be obtained by solving a symmetric 4 × 4 matrix constructed from the structural coordinates. This is traditionally accomplished by a computationally expensive inversion or decomposition of this matrix. In our recent communication, a simple and robust algorithm was proposed to rapidly determine the optimal rotation by using instead a Newton-Raphson quaternion-based method and an adjoint matrix. After determination of the 4 × 4 matrix, our method is more than an order of magnitude more efficient than traditional inversion/decomposition methods.

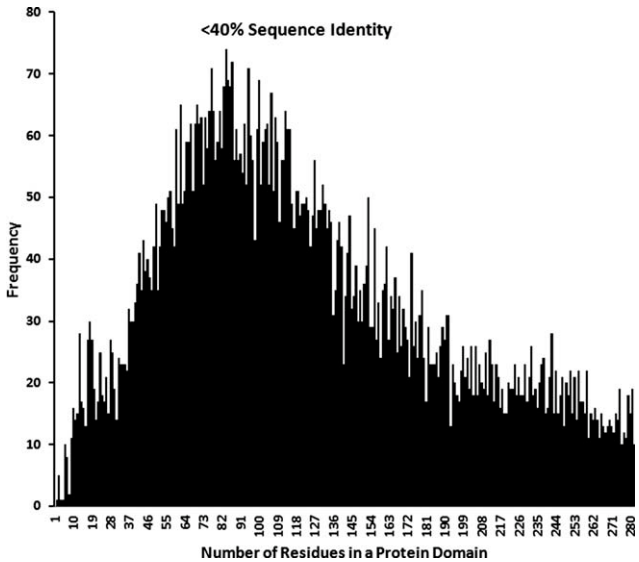

Figure 1. Frequency distribution as a function of the number of residues for all protein domains with less than 40% identity in SCOP domain database (as of May 15, 2010).[1,2] The unique mode of the distribution is roughly 83 residues. The majority of domains are less than 150 residues.

Kneller correctly points out that, for relatively large molecular systems, the time to construct the 4 × 4 matrix can outweigh the diagonalization of this matrix. In particular, he determined that the crossover point is about 100 atoms (e.g., 112 atoms from Kneller's empirical equation 7). The time in constructing the matrix scales linearly with the system size, while the time for solving this matrix is constant. Thus, the former will unavoidably dominate the latter when the system is large.

Our method is nevertheless useful in many, perhaps most, common applications. First, structural biologists conventionally superposition proteins using only the alpha carbons (one per residue), which significantly reduces the number of atoms in the computation by roughly an order of magnitude. Furthermore, many practical superposition analyses use only a single domain, an important factor as a large fraction of domains are less than 100 residues (see Fig. 1). In fact, from Kneller's equation 7, our algorithm nearly halves the time to calculate the rotational matrix for proteins with 150 atoms. Second, as mentioned in the first paragraph of our original article, many high-throughput

Correspondence to: P. Liu; e-mail: pliu24@its.jnj.com

© 2010 Wiley Periodicals, Inc.

analyses of molecular conformations use a very small number of atoms (e.g., 5–30), including fragment-assembly protein structure prediction, conformational sampling for small molecule drug design, and aligned-fragment pair multiple protein alignment. For these applications the matrix diagonalization is the dominant factor, and here our method greatly speeds the calculation. Third, there is no reason not to use our algorithm. For proteins of more than 150 atoms the computational time will always be less than the time using more traditional methods, and our code is freely available and readily incorporated into other packages. Hence our algorithm will be useful in many important practical superposition problems found throughout structural biology, bioinformatics, and chemoinformatics.

References

1. Brenner, S. E.; Koehl, P.; Levitt, M. Nucleic Acids Res 2000, 28, 254.
2. Chandonia, J. M.; Hon, G.; Walker, N. S.; Lo Conte, L.; Koehl, P.; Levitt, M.; Brenner, S. E. Nucleic Acids Res 2004, 32, D189.