Journal of Computational Physics 229 (2010) 7309–7316

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Efficient perturbation analysis of elastic network models – Application to acetylcholinesterase of *T. californica*

K. Hamacher *

TU Darmstadt, Bioinformatics & Theo. Biology Group, Schnittspahnstr. 10, 64287 Darmstadt, Germany

ARTICLE INFO

Article history:
Received 1 October 2009
Received in revised form 9 June 2010
Accepted 9 June 2010
Available online 17 June 2010

Keywords:
Elastic network models
Molecular biophysics
Sensitivity analysis
Molecular evolution
Linear algebra

ABSTRACT

Elastic network models in their different flavors have become useful models for the dynamics and functions of biomolecular systems such as proteins and their complexes. Perturbation to the interactions occur due to randomized and fixated changes (in molecular evolution) or designed modifications of the protein structures (in bioengineering). These perturbations are modifications in the topology and the strength of the interactions modeled by the elastic network models. We discuss how a naive approach to compute properties for a large number of perturbed structures and interactions by repeated diagonalization can be replaced with an identity found in linear algebra. We argue about the computational complexity and discuss the advantages of the protocol. We apply the proposed algorithm to the acetylcholinesterase, a well-known enzyme in neurobiology, and show how one can gain insight into the "breathing dynamics" of a structural funnel necessary for the function of the protein. The computational speed-up was a 60-fold increase in this example.

1. Introduction

The understanding of the structure and the dynamics of biomolecular systems is of great importance for molecular biophysics, biomolecular engineering, molecular biology, and molecular evolution. The direct simulation by sophisticated, atomistic molecular dynamics simulations is the most advanced technique available. Due to the high computational costs this approach does not offer the opportunity to deal with more than just a few variants of a molecular system. However, consideration of large number of variants is of great importance in e.g. biomolecular design, where several molecular setups have to be compared to each other, or in molecular evolution, where hundreds to thousands of sequence mutants need to be superimposed onto a molecular structure and the ramifications of sequence changes to be quantified.

Several approximation schemes exist to reduce the computational resource demand. The most basic protocols consist of a harmonic approximation of the physical potential and the avoidance of integrating equations of motion. These models are known as elastic network models (ENMs).

ENMs for biomolecular systems and proteins in particular were introduced in biophysics some twenty years ago by the seminal work of Tirion [1]. Since then numerous applications [2–4] and improvements [5–8] were suggested in the physics and chemistry literature. The most desirable property – low computational complexity – of these models made them a prime tool to understand fluctuations around the native state. At the same time this simplicity puts serious restrictions on the applicability of these models [9]: insight into folding and refolding events are in general not possible.

* Tel.: +49 6151 16 5318.
E-mail address: hamacher@bio.tu-darmstadt.de
URL: http://www.kay-hamacher.de

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.06.015

The application of these models range from annotation of molecular evolution [10] to questions in bioengineering [11]. Conceptually ENMs are reduced variants to the well-known normal mode analysis in molecular dynamics packages [12–14].

Here we want to use ENMs to quantify the importance of individual contacts within a protein structure. To this end we introduce a simple ENM. The impact the existence or the disappearance of a residue–residue contact has on the dynamics around the native state can be investigated in a thought-experiment: we can compute the difference between the original system and a system, in which the particular contact under scrutiny is 'switched off', by computing in both set-ups the full dynamics and compare them by e.g. matrix norms. For thousands of contacts this procedure becomes, however, time consuming and therefore we set out to increase the efficiency of such massive thought experiments.

As a demonstration of the power of (a) the general idea of 'switched off' contacts and (b) the computational efficiency of the proposed algorithm we apply the method to the acetylcholinesterase (AChE), an important enzyme in e.g. neurobiology, for which a detailed understanding of the molecular evolution and its ramification in physical space is still lacking.

2. Elastic network models (ENMs)

Elastic network models typically introduce two reductions in describing the molecular mechanics: (1) amino acids are represented as beads placed at the respective C_α positions of the experimental protein structure; (2) the interactions are assumed to be local only and of harmonic origin. These assumption can be justified in the region of small fluctuations around the native state. Here a Taylor expansion of any interaction potential would necessary lead to leading harmonic terms. The locality - which is implemented by allowing amino acids to interact only within a cutoff distance - follows from the observation that local packing density determines protein fluctuations to a large extent. The reduced representation assumes that effects from side-chain atoms etc. can be "integrated out", that is, absorbed in effective force constants [15].

2.1. Gaussian network model

We focus here on the Gaussian Network Model (GNM) by Haliloglu et al. [16], Micheletti [17], and Erman [18], with a potential energy given by

$$H_{GNM} = \frac{\gamma}{2}\left[\sum_{i,j}^{N} \Gamma_{ij} \underbrace{(\Delta R_j - \Delta R_i)^2}_{=:\Delta R_{ij}}\right] \tag{1}$$

where the ΔR_i are displacements from the equilibrium positions in the native state and consequently ΔR_{ij} is the relative distance the two amino acids i and j are separated in a given configuration. γ is the uniform interaction strength and Γ_{ij} is the Kirchhoff matrix

$$\Gamma_{ij} = \begin{cases} 0; & \text{if } i \neq j \quad \text{and } R_{ij} > R_c \\ -1; & \text{if } i \neq j \text{ and } R_{ij} \leqslant R_c \\ -\sum\limits_{j=1, j\neq i}^{N} \Gamma_{ij}; & \text{if } i = j \end{cases}$$

where the cutoff distance R_c was set to 13Å throughout this study.

2.2. Observables from ENMs

We first need to note that the model set-up provides for translational symmetry. This effectively leads to a singular Kirchhoff matrix Γ with one eigenvalue $\lambda_N = 0$. As we only need to deal with a physically meaningful sub-space of movements we need to omit this particular eigenvalue/-vector from further consideration. The eigensystem of Γ must therefore be obtained by a singular value decomposition [19] (SVD) and not by direct eigensystem methods.

The covariance of residue motions can be computed as

$$C_{ij} := \langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_B T}{\gamma} \widetilde{\Gamma}^{-1}_{ij}$$

where $\widetilde{\Gamma}^{-1}$ is the Moore–Penrose [20,21] pseudo-inverse of Γ. The pseudo-inverse needs to be computed due to the above mentioned singularity of the original Kirchhoff matrix Γ.

Assuming the eigensystem of Γ to be sorted in descending order according to the singular values λ_k then we obtain for the covariance matrix entries

$$C_{ij} = \frac{3k_B T}{\gamma} \sum_{k=1}^{N-1} \frac{1}{\lambda_k} [\vec{u}_k \vec{u}_k^T]_{ij}$$

here $[\vec{u}_k\vec{u}_k^T]_{ij}$ is the (i,j) entry of the matrix composed of the singular vectors $\vec{u}_k$ of Γ and their respective transposed. The summation of a reduced spectral decomposition implements then the Moore–Penrose pseudo-inverse $\widetilde{\Gamma}^{-1}$ of Γ.

k_B and T are the Boltzmann constant and the temperature, respectively. In the following we will set $\frac{3k_BT}{\gamma} = 1$ for convenience. The crystallographic temperature factor for residue i is given as $B_i = \frac{8\pi^2 k_BT}{\gamma}(\widetilde{\Gamma}^{-1})_{ii} = \frac{8\pi^2}{3}\mathcal{C}_{ii}$.

3. Perturbation of ENMs

What happens if we want to investigate the influence of a modified interaction onto the overall dynamics? The full information ENMs can provide on the dynamics of systems, they are applied to, is contained in the covariance matrix $\mathcal{C}$. Temperature factors and thermodynamic properties follow solely from knowledge of $\mathcal{C}$ within this framework.

Therefore we can restrict our analysis to the changes induced in $\mathcal{C}$ upon changes in the Kirchhoff matrix Γ. Such modifications in the Kirchhoff matrix might be due to conscious processes such as molecular design efforts in drug development. They might also occur during evolutionary changes such as point mutations that result in a (locally) changed amino acid sequence.

To quantify changes in the covariance matrix $\mathcal{C}$ we use – as in a previous study [10] – a matrix norm between the original molecular $\mathcal{C}_{\text{orig}}$ and the modified system $\mathcal{C}_{\text{mod}}$. We will restrict our analysis to the well-known Frobenius norm $\|A\|_F := \sqrt{\sum_{ij} A_{ij}^2}$ for a matrix A. In our case the Frobenius norm upon molecule modification $f_{\text{orig}\to\text{mod}}$ becomes

$$f_{\text{orig}\to\text{mod}} = \sqrt{\sum_{i,j=1}^{N} [\mathcal{C}_{\text{orig}} - \mathcal{C}_{\text{mod}}]_{ij}^2}$$

In principle this can be easily achieved by (1) changing the Kirchhoff matrix, (2) applying a SVD, and (3) reconstructing the pseudo-inverse as above. However, if we have a large number of changes to investigate or an approximated continuum of interaction strengths this approach becomes prohibitively expensive. This conceptual problem cannot be solved by increased efficiency in the SVD procedure [22] alone. Instead we propose a new approach.

3.1. An efficient implementation

The effect of a change in the original Kirchhoff matrix Γ_{orig} to the covariance matrix of the modified system $\mathcal{C}_{\text{mod}}$ is easily obtained from the Woodbury matrix identity [23]:

$$(A + UWV)^{-1} = A^{-1} - A^{-1}U(W^{-1} + VA^{-1}U)^{-1}VA^{-1}$$

for general matrices A, U, W, V with sizes $N \times N$, $N \times M$, $M \times M$, and $M \times N$, respectively.

We define $\Delta := \Gamma_{\text{mod}} - \Gamma_{\text{orig}}$ and find for physical forces without loss of generality always a representation $\Delta = UWV$ for appropriately sized matrices U, V, W. Then we obtain from the Woodbury matrix identity for the covariance matrix of the modified system

$$\mathcal{C}_{\text{mod}} = \frac{3k_BT}{\gamma}\widetilde{\Gamma}_{\text{mod}}^{-1} = \frac{3k_BT}{\gamma}\left[\Gamma_{\text{orig}} + \Delta\right]^{-1} = \frac{3k_BT}{\gamma}\left[\widetilde{\Gamma}_{\text{orig}}^{-1} - \widetilde{\Gamma}_{\text{orig}}^{-1}U(W^{-1} + V\widetilde{\Gamma}_{\text{orig}}^{-1}U)^{-1}V\widetilde{\Gamma}_{\text{orig}}^{-1}\right] \tag{2}$$

Here a matrix $\widetilde{\Gamma}^{-1}$ is the Moore–Penrose pseudo-inverse of Γ. If now the size M of the matrix W is just a fraction the size of Γ, that is the number of amino acids N, it is computationally less expensive to apply this approach compared to finding the eigensystem every single time from scratch upon changes in the Kirchhoff matrix as in the naive approach. The diagonalization/inversion of the smaller matrices is much less costly: typically the readily available implementation of the SVD scale approximately with the third power of the system size $O(N^3)$. If we repeat this step n times, the overall CPU time scales like $t_{\text{naive}} \sim O(n\,N^3)$.

Using Eq. (2) we apply the full SVD once ($O(N^3)$) and $n-1$ times we apply the more efficient scheme which involves some matrix-multiplications and SVDs of a much smaller matrix. Each of these steps has complexity $O(M^3 + M^2 + 2\cdot M^2N + 4\cdot N^2M)$. Thus we observe

$$t_{\text{new}} \sim O\left(N^3 + nM^3 + nM^2 + 2n\cdot M^2N + 4n\cdot N^2M\right)$$

in a slightly adopted O-notation and the approximation $n \approx n-1$ for large n. Clearly we see that for M as just a fraction of N this algorithms allows for much larger n for a restricted computational resource. The estimate on the computational effort requires the "correct" order of matrix operations. If applied in the wrong order, the algorithm can perform worse than the naive approach. We have to store the reference matrix $\widetilde{\Gamma}_{\text{orig}}^{-1}$, thus the memory demands are always higher in this algorithm.

The Woodbury matrix identity is a rigorous result from linear algebra. Therefore our protocol does not introduce any approximation beyond the one of the GNM model. In fact the results of repeated SVDs and the application of the Woodbury identity are exactly the same. For the application test – described in the next section – we found the method also to be numerical robust. The changes in computed observables (temperature factors and covariance matrix entries) between repeated SVDs and the Woodbury-based algorithm was always smaller than 10^{-6} in the example.

3.2. A special case

In the case that U and V in Eq. (2) are of size $N \times 1$, and thus $M = 1$, then the Woodbury matrix identity reduces to the better known Sherman–Morrison formula [24]. This situation occurs if we only change the interaction between one and only one pair of amino acids. Then only the respective entries in the symmetric Kirchhoff must be changed and the diagonal entry updated. It follows for the vectors $\vec{u}$ and $\vec{v}$ (which are the columns of U and V):

$$\mathcal{C}_{\text{mod}} = \frac{3k_BT}{\gamma}\widetilde{\Gamma}^{-1}_{\text{mod}} = \frac{3k_BT}{\gamma}[\Gamma_{\text{orig}} + \Delta]^{-1} = \frac{3k_BT}{\gamma}\left[\widetilde{\Gamma}^{-1}_{\text{orig}} + \frac{\widetilde{\Gamma}^{-1}_{\text{orig}}\vec{u}\vec{v}^T\widetilde{\Gamma}^{-1}_{\text{orig}}}{1 + \vec{v}^T\widetilde{\Gamma}^{-1}_{\text{orig}}\vec{u}}\right] \tag{3}$$

If we want to modify the interaction for a pair of residues i and j by an amount δ then the difference in the Kirchhoff matrix is

$$\Delta_{kl} = \begin{cases} -\delta; & \text{for } (k = i \wedge l = j) \vee (k = j \wedge l = i) \\ -\delta; & \text{for } k = l = i \vee k = l = j \\ 0; & \text{otherwise} \end{cases} \tag{4}$$

This can be e.g. achieved by the vectors $\vec{u} = (\dots, \delta, \dots, \delta, \dots)^T$ and $\vec{v} = (\dots, 1, \dots, -1, \dots)^T$ with the entries at positions i and j.

Here – adopting the notion of above – we have $t_{\text{new}} \sim O(N^3 + 5nN^2 + nN)$. The break even point for application of the proposed method is therefore below any substantial protein size. It is therefore almost always better to implement our more elaborated scheme. From $\mathcal{C}_{\text{orig}}$ we can then compute all other values as described above. These analysis steps are independent of the algorithm used to obtain $\mathcal{C}_{\text{orig}}$ and therefore out of the focus of our present work.

3.3. Overview of the protocol

For a GNM potential $H_{GNM} = \frac{\gamma}{2}\left[\sum_{i,j}^{N}\Gamma_{ij}(\underbrace{\Delta R_j - \Delta R_i}_{=:\Delta R_{ij}})^2\right]$ with ΔR_i displacements from the equilibrium positions, ΔR_{ij} the relative distance the two amino acids i and j and Γ_{ij} the Kirchhoff matrix, one computes the covariance matrix entries $\mathcal{C}_{ij}$ by a Moore–Penrose pseudo-inverse $\mathcal{C}_{ij} := \langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_BT}{\gamma}\sum_{k}^{N-1}\frac{1}{\lambda_k}\left[\vec{u}_k\vec{u}_k^T\right]_{ij}$ where the eigenvectors $\vec{u}_k$ and the eigenvalues λ_k are taken from an initial singular value decomposition of the original, non-perturbed system.

In our protocol we compute the importance a contact between residues (i,j) has for the dynamics of the molecule by the Frobenius norm between the covariance matrices of the full and a set-up, which lacks this particular contact $f_{\text{orig}\to\text{mod}} = \sqrt{\sum_{i,j=1}^{N}\left[\mathcal{C}_{\text{orig}} - \mathcal{C}_{\text{mod}}\right]^2_{ij}}$.

To compute these covariance matrices one can employ the Woodbury matrix identity and start from the covariance matrix for the full system: $\mathcal{C}_{\text{mod}} = \frac{3k_BT}{\gamma}\left[\widetilde{\Gamma}^{-1}_{\text{orig}} + \frac{\widetilde{\Gamma}^{-1}_{\text{orig}}\vec{u}\vec{v}^T\widetilde{\Gamma}^{-1}_{\text{orig}}}{1+\vec{v}^T\widetilde{\Gamma}^{-1}_{\text{orig}}\vec{u}}\right]$. For a contact (i,j) we set here $\vec{u} := -\hat{e}_i$ and $\vec{v} := \hat{e}_j$, with $\hat{e}_i$ and $\hat{e}_j$ unit vectors with an entry of one in position i and j, respectively, and all other entries set to zero.

4. An example: AChE

In most eukaryotes acetylcholinesterase (AChE, EC 3.1.1.7) degrades by hydrolytic activity the neurotransmitter acetylcholine with a high turnover of some 1000–10,000 reactions per second near the diffusion limit [25]. Inhibitors of AChE include snake venom and chemical weapons [26] such as organophosphates. These substances undermine the enzymatic activity of AChE and thus acetylcholine accumulates in the brain with fatal consequences up to asphyxiation. At the same time patient of e.g. Alzheimer's disease are treated with lower doses of AChE-inhibitors [27]. This makes AChE an interesting protein for studying its structure and mechanics.

The enzymatic action occurs in a very narrow gorge, therefore substantial breathing motions of the gorge forming residues are necessary for entering of substrates and leaving of products.

Owing its importance several entries for AChE in the protein data bank exist. We have chosen the entry 2c5g [28] for *T. californica*, which contains also N-acetyl-D-glucosamine, triethylene-glycol, and 2-(trimethylammonium)ethyl-thiol.

We extracted the C_α positions and selected always the first alternative location if two were present. The structure has 532 amino acids and can thus be categorized in the class of large proteins for which a detailed insight by atomistic molecular dynamics for hundreds of variations – even with modern compute clusters – cannot be gained. It constitutes therefore a prime example for our new method.

4.1. Quality of B-factors in the GNM

To get a first impression of the reliability and accuracy of the GNM for AChE we applied the full method to the pdb-structure 1MAH and compare to results from a previous full atomistic molecular dynamics study [29].[1] We found for the computed

[1] sample values extracted from Fig. 4 in ref. [29].

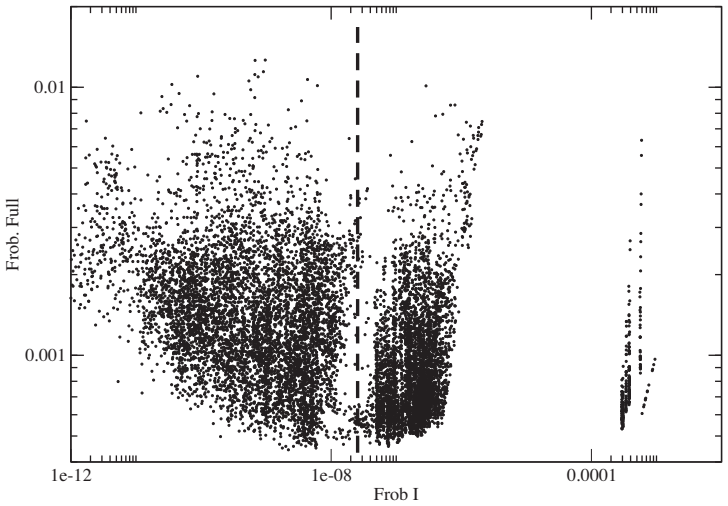

Fig. 1. Scatter plot of the Frobenius norms for all 10,872 switched-off contacts for the AChE. On the *x* axis we plot the Frobenius norm for the difference of the covariances matrices $\mathcal{C}$ restricted to the index set *I*, while the *y* axis shows the Frobenius norm for the full 532 amino acid covariances matrices. The gap along the *x* axis and the overall magnitude of 10^{-8} for the position of the broken line suggests that values to the left are below numerical accuracy and must be regarded as zero – thus no observable change in the molecular mechanics.

B-factors of reference [29] a correlation to the experimental B-factors in the pdb-file of just $r_P = 0.017$ (Pearson) and $r_S = 0.008$ (Spearman), while with the GNM approach we obtained $r_P = 0.49$ and $r_S = 0.49$. At least for the temperature factors these findings suggest a higher quality. Most likely these results will not hold for contemporary molecular dynamics studies with larger sampling times. Most molecular dynamics studies are done, however, with additional ligands and binding partners [30–33] – a situation to which the GNM is most likely not applicable. We therefore refrained from comparing to these studies.

4.2. Fluctuation changes in the active site of AChE

The substrate acetylcholine (ACh) interacts with the amino acids Trp84, Glu199, and Phe330 and the serine-hydrolase catalytic triad of Ser200–His440–Glu327. In our analysis we will extract just the covariances for these residues, eventually filtering $\mathcal{C}$ for these indices.

To quantify the changes in such covariances sub-matrices, computing the Frobenius norm of $\mathcal{C}^{(I)}_{\text{mod}}$ with respect to one of the starting, reference configuration $\mathcal{C}_{\text{orig}}$ was previously suggested [10]. Here $\mathcal{C}^{(I)}_{\text{mod}}$ stand for the covariances matrix of a modified set-up and then reduced to the indices in *I*.

We will stick to this idea and compute first $\mathcal{C}^{(I)}_{\text{orig}}$ as a reference matrix. *I* is the index set {84, 199, 200, 327, 330, 440} from above. We then modify the Kirchhoff matrix according to Eqs. (3) and (4) and compute $\mathcal{C}_{\text{mod}}$ by these equations. Finally we compute the Frobenius norm for the restriction set *I*

$$f(I) := \left\| \mathcal{C}^{(I)}_{\text{orig}} - \mathcal{C}^{(I)}_{\text{mod}} \right\|_F = \sqrt{\sum_{i \in I} |\mathcal{C}_{\text{orig},i} - \mathcal{C}_{\text{mod},i}|^2}$$

We decided to restrict our illustration here to a "switch-off" scenario: we switch off one contact after another and quantify the influence of such an interaction on the residues in *I* by the respective *f*. To this end we set $\delta = -1$ in the definition of $\vec{u}$ for each amino acid pair (i,j) within the distance $R_c = 13$Å one at a time, thus "switching" the contact "off".

Therefore the approach consists of a two-folded loop over all contacting amino acid pairs (i,j). Within the loop we use a special case of the Woodbury identity (Section 3.1), namely the special case of Eq. (4), that can be reduced to the special case for just one contacting pair (i,j) with $\vec{u} := -\hat{e}_i$ and $\vec{v} := \hat{e}_j$ (Section 3.3). For each contact pair we then compute the Frobenius norm between the original, full system and the modified one.

At this R_c we found 10,872 contacts that need to be sampled. In Fig. 1 we show the Frobenius norms $\|\mathcal{C}_{\text{orig}} - \mathcal{C}_{\text{mod}}\|_F$ for the full system of all amino acids and the one restricted to *I*. As is clear from the scattered points the Frobenius norms show no dependency. This points to need to formulate the research objective stringent: the overall change within the full molecule is much different than for the restricted "window of interest" *I*. Focusing on the full molecule might mask the relevant changes in the gorge region (*I*).

We will now focus on the Frobenius norm restricted to *I*. We filtered the most important influences by extracting those interacting amino acid pairs, that showed upon switch-off the largest Frobenius norm with respect to the reference system. The most important eight interactions are shown in Table 1. We superimpose these pairings onto the structure of AChE and show this in Fig. 2. Interestingly all of them involve the amino acid no. 84, which is also a member of *I*. The biggest influence

Table 1
The ids of most important interacting amino acid pairs in the AChE structure. The importance reveals itself by comparable larger values of the Frobenius norm f between the full system and the one in which this contact is switched off. We have filtered for those Frobenius norms f with $f > 5.7 \cdot 10^{-4}$. We omitted every pair that included two elements of I as changes due to interactions within I are obviously relevant for the dynamics of members of I.

73 ↔ 84	74 ↔ 84	75 ↔ 84	76 ↔ 84	77 ↔ 84	78 ↔ 84	79 ↔ 84	82 ↔ 84

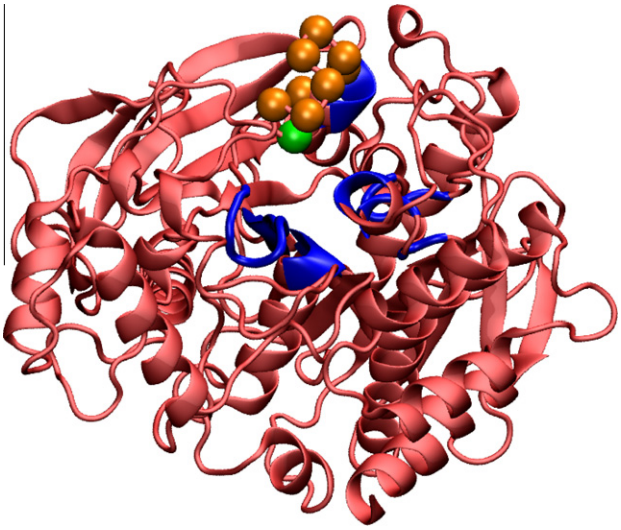

Fig. 2. The structure of acetylcholinesterase (pdb code 2c5g) in red. The gorge of the enzymatic center is shown in blue. The orange spheres represent the C_α positions of the residues contacting residue no. 84 (green), which were found to be most important in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stems therefore from the flexible loop at the entrance of the gorge. This loop is part of the bigger Omega loop, which was shown to be important for inhibitor binding in the AChE of *M. musculus* [34].

Interestingly we found in a preliminary BLAST [35] search an unusually high sequence variability in this region among different species – ranging from *C. carpio* (common carp), over *Danio rerio* (zebrafish), to *X. laevis* (African frog) and eventually *M. musculus* (mouse).

This portion of the gorge entrance plays some role – as can be seen in the structure 2c5g – in the transport of the product 2-(trimethylammonium)ethyl-thiol. Whether this is related to the above mentioned sequence variability or other reasons for the evolutionary interesting observation on high sequence variability will be discussed in a future paper, which will not focus on methodology.

4.3. Timings

On a Mac OS X 10.5 with an Intel 2.4 GHz Duo core processors with 4 GB RAM the full sampling for the 10,872 switch-off settings took just 143 s. The full sampling by the naive algorithm with repeated pseudo-inverse computation took considerably longer, 8480 s. Although absolute timings are just indications on performance improvements, the ~60 fold speed-up is striking.

5. Conclusions

In this study we have implemented an efficient algorithm to investigate changes in topology and interaction strength of elastic network models for biomolecular systems.

The speed-up compared to a naive approach for sampling a range of interactions in the acetylcholinesterase was close to 60. Theoretical arguments on the scaling of the computational effort point in the same direction.

In a first application we scanned the contacts within the acetylcholinesterase structure of *T. californica* for important contributions to the flexibilities and fluctuations in the gorge region of this molecule. We found a signal for a portion of the so called Omega loop, which is known for its contribution to inhibitor binding. This observation prompts for the speculation, that upon binding the inhibitor the AChE undergoes changes in its dynamics, which in turn might prevent the necessary "breathing" motion of the gorge to accommodate its native substrate. Detailed investigations on this will follow.

Generally speaking, with the new approach the design of particular dynamical and functional features by small, local modifications of the contact topology and interaction strength seems within reach. Combining this approach for the assessment of functional and dynamical ramifications of protein modifications with efficient optimization schemes [36–39] is therefore the next step we will undertake.

For general modifications of the Kirchhoff matrix using the Woodbury identity might not lead to any improvements in the computational efficiency. However, global changes in the Kirchhoff matrix are typically not done as this would imply a very different protein structure (as the contact map directly determines with the Kirchhoff matrix). In these cases with substantial modifications, where the number of effect residues is of the same order as the overall number of residues in the molecule, it is questionable whether elastic network models make sense at all. Most likely a significant change in the contact map corresponds to a protein structure, which is not even a local minimum of its free energy and thus the harmonic approximation implicitly presupposed by elastic network models is invalid.

One might, however, need such global changes when investigating the effects of mutations in traditional normal mode analysis using force-fields like AMBER or CHARMM. For an exact mathematical treatment the above procedure would not lead to an increase in performance as all matrix entries need to be modified. Therefore the matrices to be dealt with are as large as the starting matrix. However, in the event that the resulting entries in the respective Hessians do not deviate too much it might be possible to focus only on changes up to certain threshold, effectively dealing only with small sized matrices V, U, and W in the Woodbury identity of Eq. (2) again. This will be investigated in a forthcoming study.

Acknowledgments

KH is grateful to the Fonds der chemischen Industrie which supported this study through a grant for junior faculty. The study was performed by means of the BioPhysConnectoR-package [40] for the statistical software R [41]. The molecular structure was rendered with VMD [42]. The author is most grateful for the helpful comments of five anonymous referees, who helped to improve the style of the presentation considerably.

References

[1] M. Tirion, Large amplitude elastic motions in proteins from a single-parameter, atomic analysis, Phys. Rev. Lett. 77 (9) (1996) 1905–1908.
[2] K. Hamacher, J. Trylska, J.A. McCammon, Dependency map of proteins in the small ribosomal subunit, PLoS Comp. Biol. 2 (2006) e10.
[3] I. Bahar, R. Jernigan, Cooperative fluctuations and subunit communication in tryptophan synthase, Biochemistry 38 (1999) 3478–3490.
[4] I. Bahar, R.L. Jernigan, Vibrational dynamics of transfer RNAs. Comparison of the free and enzyme-bound forms, J. Mol. Biol. 281 (1998) 871–884.
[5] K. Hamacher, J.A. McCammon, Computing the amino acid specificity of fluctuations in biomolecular systems, J. Chem. Theory Comput. 2 (3) (2006) 873–878.
[6] P. Doruker, R. Jernigan, I. Bahar, Dynamics of large proteins through hierarchical levels of coarse-grained structures, J. Comp. Chem. 23 (2002) 119–127.
[7] F. Tama, I.C.L. Brooks, The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus, J. Mol. Biol. 318 (3) (2002) 733–747.
[8] O. Miyashita, J.N. Onuchic, P.G. Wolynes, Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins, Proc. Nat. Acad. Sci. 100 (22) (2003) 12570–12575.
[9] K. Hamacher, A. Hübsch, J.A. McCammon, A minimal model for stabilization of biomolecules by hydrocarbon cross-linking, J. Chem. Phys. 124 (16) (2006) 164907.
[10] K. Hamacher, Relating sequence evolution of HIV1-protease to its underlying molecular mechanics, Gene 422 (2008) 30–36.
[11] I. Bahar, A.R. Atilgan, B. Erman, Direct evaluation of thermal fluctuations in protein using a single parameter harmonic potential, Fold. Des. 2 (1997) 173–181.
[12] B. Brooks, M. Karplus, Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor, Proc. Nat. Acad. Sci. 80 (1983) 6571–6575.
[13] B. Brooks, D. Janezic, M. Karplus, Harmonic analysis of large systems. I. methodology, J. Comp. Chem. 16 (12) (1995) 1522–1542.
[14] B. Roux, M. Karplus, The normal modes of the gramicidin-A dimer channel, Biophys. J. 53 (3) (1988) 297–309.
[15] K. Hamacher, Temperature dependence of fluctuations in HIV1-Protease, Eur. Biophys. J. 39 (2010) 1051.
[16] T. Haliloglu, I. Bahar, B. Erman, Gaussian dynamics of folded proteins, Phys. Rev. Lett. 79 (16) (1997) 3090.
[17] C. Micheletti, P. Carloni, A. Maritan, Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and gaussian models, Proteins 55 (2004) 635.
[18] B. Erman, The Gaussian network model: precise prediction of residue fluctuations and application to binding problems, Biophys. J. 91 (10) (2006) 3589–3599.
[19] W.H. Press et al, Numerical Recipies in C, Cambridge University Press, Cambridge, 1995.
[20] E.H. Moore, Bull. Am. Math. Soc. 26 (1920) 394–395.
[21] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc. 51 (1955) 406–413.
[22] G.H. Golub, C.F.V. Loan, Matrix Computations, third ed., The Johns Hopkins University Press, Baltimore, 1996.
[23] W.W. Hager, SIAM Rev. 31 (1989) 221–239.
[24] J. Sherman, W.J. Morrison, Adjustment of an inverse matrix corresponding to a change in one element of a given matrix, Ann. Math. Stat. 21 (1) (1950) 124–127.
[25] I. Silman, J.L. Sussman, Acetylcholinesterase: 'classical' and 'non-classical' functions and pharmacology, Curr. Opin. Pharmacol. 5 (3) (2005) 293–302.
[26] C. Millard, G. Kryger, A. Ordentlich, H. Greenblatt, M. Harel, M. Raves, Y. Segall, D. Barak, A. Shafferman, I. Silman, J. Sussman, Crystal structures of aged phosphonylated acetylcholinesterase: nerve agent reaction products at the atomic level, Biochemistry 38 (22) (1999) 7032–7039.
[27] H. Greenblatt, H. Dvir, I. Silman, J. Sussman, Acetylcholinesterase: a multifaceted target for structure-based drug design of anticholinesterase agents for the treatment of Alzheimer's disease, J. Mol. Neurosci. 20 (3) (2003) 369–383.
[28] J.-P. Colletier, D. Fournier, H.M. Greenblatt, J. Stojan, J.L. Sussman, G. Zaccai, I. Silman, M. Weik, Structural insights into substrate traffic and inhibition in acetylcholinesterase, EMBO J. 25 (12) (2006) 2746–2756.
[29] K. Tai, T. Shen, U. Börjesson, M. Philippopoulos, J.A. McCammon, Analysis of a 10-ns molecular dynamics simulation of mouse acetylcholinesterase, Biophys. J. 81 (2001) 715–724.
[30] T. Shen, K. Tai, R.H. Henchman, J.A. McCammon, Molecular dynamics of acetylcholinesterase, Acc. Chem. Res. 35 (2002) 332–340.
[31] S. Tara, V. Helms, T.P. Straatsma, J.A. McCammon, Molecular dynamics of mouse acetylcholinesterase complexed with huperzine A, Biopolymers 50 (1999) 347–359.

[32] S. Tara, T.P. Straatsma, J.A. McCammon, Mouse acetylcholinesterase unliganded and in complex with huperzine A: a comparison of molecular dynamics simulations, Biopolymers 50 (1999) 34–43.
[33] S.T. Wlodek, T.W. Clark, L.R. Scott, J.A. McCammon, Molecular dynamics of acetylcholinesterase dimer complexed with tacrine, J. Am. Chem. Soc. 119 (1997) 9513–9522.
[34] J. Shi, Z. Radic, P. Taylor, Inhibitors of different structure induce distinguishing conformations in the Omega loop, Cys69–Cys96, of mouse acetylcholinesterase, J. Biol. Chem. 277 (45) (2002) 43301–43308.
[35] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[36] K. Hamacher, Adaptive extremal optimization by detrended fluctuation analysis, J. Comp. Phys. 227 (2) (2007) 1500–1509.
[37] K. Hamacher, Energy landscape paving as a perfect optimization approach under detrended fluctuation analysis, Physica A 378 (2) (2007) 307–314.
[38] K. Hamacher, Adaptation in stochastic tunneling global optimization of complex potential energy landscapes, Europhys. Lett. 74 (6) (2006) 944–950.
[39] K. Hamacher, W. Wenzel, The scaling behaviour of stochastic minimization algorithms in a perfect funnel landscape, Phys. Rev. E 59 (1) (1999) 938–941.
[40] F. Hoffgaard, P. Weil, K. Hamacher, BioPhysConnectoR: connecting sequence information and biophysical models, BMC Bioinformatics 11 (2010) 199.
[41] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (2008). URL <http://www.R-project.org>.
[42] W. Humphrey, A. Dalke, K. Schulten, VMD – Visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.