

Precise spring constant assignment in elastic network model for identification of vibration frequency and modeshape[†]

Mingwen Hu¹, Sharad Raj¹, Byung Kim¹, Wing Kam Liu², Seunghyun Baik³,
 Taesung Kim³, Byeong-Soo Lim³ and Moon Ki Kim^{3,*}

¹Department of Mechanical and Industrial Engineering, University of Massachusetts; 160 Governors Dr., Amherst, MA 01003, USA

²Department of Mechanical Engineering, Northwestern University; 2145 Sheridan Rd., Evanston, IL 60208 USA

³School of Mechanical Engineering, Sungkyunkwan University; 300 Cheoncheon-dong, Suwon, 440-746, Korea

(Manuscript Received August 6, 2009; Revised March 28, 2010; Accepted June 16, 2010)

Abstract

We presented a feasible framework of studying dynamics of macromolecules by applying elastic network model (ENM) and vibration spectroscopy. We first identified the precise force constants of covalent bonds commonly observed in macromolecules by matching their reported Raman shifts data with predicted wavenumbers determined by normal mode analysis (NMA). Assigning the obtained spring constants to other small chemical compounds such as ethynyl isocyanide (C₃HN) and diacetylene (C₄H₂), we not only predicted their vibration wavenumbers precisely but also identified their individual mode shapes from NMA. We extensively tested this chemical information based ENM with one of amino acids, cysteine. Subsequent comparison of frequencies and modeshapes also yields a strong agreement between computed and experimental data. Consequently, the proposed method enables us to identify low frequency modeshapes that are in general functionally important collective motions of macromolecules but have hardly been revealed experimentally even using terahertz spectroscopy.

Keywords: Elastic network model; Normal mode analysis; Force constant calculation; Mode shape identification; Spectroscopy

1. Introduction

Although determination of force constants in macromolecules has been widely studied, it is still one of interesting topics in macromolecular dynamics study [1-5]. As it has been reported that atomic force constants depend on the assignment of the vibration spectrum, experimental data for vibration assignment can be utilized to identify bond specific spring constants that are analogous to force constants in macromolecules [6-10]. It is also noted that the covalent bonds of molecules are not rigid, but more like massless pseudo springs that can be stretched and bent. At ordinary temperatures, these bonds vibrate in a variety of ways and the vibration energy of molecules may be assigned to quantum levels in the same manner as are their electronic states. For example, vibration frequency of a diatomic molecule can be calculated by Eq. (1) which describes the major factors that influence the stretching frequency of a bond between two masses m_1 and m_2 , respectively [11].

$$\tilde{\nu} = \left(\frac{1}{2\pi c} \right) \left[\sqrt{\frac{f(m_1 + m_2)}{m_1 m_2}} \right] \quad (1)$$

where $\tilde{\nu}$ is a frequency (cm^{-1}), f is a force constant, and c is the speed of light. In the analogy of a mass-spring system, the force constant f corresponds to the spring's stiffness proportional to the strength of the bond between m_1 and m_2 . Although such a quantum mechanics enables us to determine precise force constants (or fields) by electronic structure calculation, it is limited to small chemical compounds due to its computational complexity. So it is not feasible to study macromolecular dynamics using this methodology.

As an alternative, elastic network model (ENM) has been developed for analyzing macromolecular dynamics in frequency domain [12-19]. In general, it treats a molecular structure as a mass-spring system in which each substituent atom is considered as a point mass and its atomic interactions with the surrounding atoms can be represented by massless linear springs in Cartesian coordinates (not in internal coordinates). The relationship between elastic network and internal coordinates such as bond angle and torsion angle was already discussed elsewhere [17]. In ENM, one can directly assign spring

[†]This paper was recommended for publication in revised form by Associate Editor Yong-Tae Kim

*Corresponding author. Tel.: +82 31 299 4840, Fax: +82 31 290 5889

E-mail address: mkkim@me.skku.ac.kr

© KSME & Springer 2010

constants obtained from Eq. (1) to covalent bonds while other stiffness values of virtual springs should be adjusted by comparison in frequency domain between the experimental data and the result from ENM based normal mode analysis (NMA).

Here we propose a mechanically precise and computationally efficient framework to model and predict vibration modes in macromolecules based on an amalgamation of experimentally obtained vibration frequency with the computational approach of NMA which incorporates atomistic details to calculate the virtual spring constants between various atoms. These computed modes can be assigned to the corresponding frequency spectrum obtained by spectroscopy experiment with which so far we have only detected vibration frequencies as a specific signature identifying the tested molecule itself (not its motions).

To validate the proposed method, we select two linear molecules (ethylene isocyanide and diacetylene) and one amino acid (cysteine), construct their ENMs by assigning computed spring constants and the appropriate mass values for atoms, respectively, calculate vibration frequencies and modes by using NMA, and finally compare those results with the experimental data.

2. Method

2.1 Chemical information based elastic network model

In the conventional ENMs, a macromolecule has been treated as a mass-spring system. That is, each atom is modeled as an individual point mass connected to other ones by "virtual" linear springs with a uniform dimensionless stiffness value [12-15]. However, in the chemical information based ENM proposed here, we assign precise spring constants with a "realistic" dimension obtained from Eq. (1) keeping in mind the nature of specific covalent bonds as well as non-bonded interactions as a function of the spatial distance between a pair of atoms.

In the scope of the undertaken modeling scheme and force field parameterization, the effect of non-bonded interactions on molecular vibration are more dominant than that of bonded interactions due to the conformational flexibility of macromolecules. Hence, a distance cutoff scheme shown in Eq. (4) is adopted to replicate the non-bonded interactions in this context.

The total kinetic energy T and potential energy V in a network of n point masses can be presented as

$$T = \frac{1}{2} \sum_{i=1}^n m_i \|\dot{\vec{x}}_i(t)\|^2 \quad (2)$$

$$V = \frac{1}{2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n k_{i,j} \left\{ \|\vec{x}_i(t) - \vec{x}_j(t)\| - \|\vec{x}_i(0) - \vec{x}_j(0)\| \right\}^2 \quad (3)$$

where $\vec{x}_i(t)$ is the position of the i^{th} atom at time t , m_i is the mass of the i^{th} atom, and $k_{i,j}$ is an element of the linking matrix which is defined as

$$k_{i,j} = \begin{cases} C & d_{i,j} \leq l \\ C \exp[-(d_{i,j} - l)] & l \leq d_{i,j} \leq u \\ 0 & u \leq d_{i,j} \end{cases} \quad (4)$$

where $d_{i,j}$ is the distance between the i^{th} and the j^{th} atoms and u is a upper-bound cutoff distance of 8\AA such that for a given atom, any atom which is farther than this distance is considered to have no connection. Additionally, l is a lower-bound cutoff distance of 2\AA , which is normally greater than the average covalent bond length. C is a non-bonded force constant of 6×10^3 dynes/cm to be assigned between any pair of atoms that are spatially proximate around the generally observed covalent bond length. When the value of $d_{i,j}$ is in between l and u , the force constant is designed to be exponentially decayed as $d_{i,j}$ increases. In addition to the above assignment for non-bonded interactions, bonded atoms are also sorted and then bond specific force constants are substantially assigned.

Let's define $\vec{\delta}(t)$ as a vector of small displacement and the global mass matrix M for the whole network system such that

$$\vec{x}_i(t) = \vec{x}_i(0) + \vec{\delta}_i(t) \quad (5)$$

$$T = \frac{1}{2} \dot{\vec{\delta}}^T M \dot{\vec{\delta}} \quad (6)$$

where $\vec{\delta} = [\vec{\delta}_1^T, \dots, \vec{\delta}_n^T]^T \in R^{3n}$. If we assume that the displacement is very small, V in Eq. (3) can be simplified as a classical quadratic potential energy function such that

$$V = \frac{1}{2} \dot{\vec{\delta}}^T K \dot{\vec{\delta}} \quad (7)$$

Here K is the stiffness matrix for the whole network. Finally, the equations of motion for ENM can be simply represented as follows:

$$M \ddot{\vec{\delta}} + K \vec{\delta} = 0 \quad (8)$$

From Eq. (8), we can compute the eigenvectors and the eigenvalues, which can be interpreted as vibration frequencies and corresponding modes of macromolecules, respectively.

2.2 Force constant parameterization using linear molecules

Force fields consist of bonded and non-bonded interactions. As the name implies, the non-bonded interactions (representing intra-molecular and inter-molecular forces) exist between atoms which are not linked by covalent bonds but spatially closed to each other. They dominantly depend on the conformation of a given molecule.

In a general 3-dimensional (3D) structure, non-bonded atoms interactions are more predominant due to close proximity

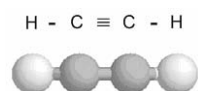


Fig. 1. Schematic of acetylene: a ball and stick representation of the ENM model setup using the appropriate values of masses and spatial coordinates.

of atoms in the spatial domain. Moreover, atoms in the same molecule can occasionally become very close to each other, leading to large values for the non-bonded energy and forces, especially Van der Waals, so special measures are sometimes needed to accommodate this effect. With the current state of study, though fundamentally decipherable, more rigorous mathematical models are essential for expressing non-bonded force constants as a function of the distance between two atoms involved. Therefore, linear molecules are favorable to our analysis for obtaining precise force constants because the effect of non-bonded interactions is minimal on modes involving simple stretching. This assumption confers with the theoretical findings since simple stretching can be considered as only a strong function of bonded force constants.

Here we present how the proposed methodology was applied to a linear molecule, acetylene (C_2H_2 , $H-C \equiv C-H$). Fig. 1 illustrates the molecular structure of acetylene showing the four atoms including two hydrogen and two carbon, their orientation, and the bonds among them. The balls represent the atoms with their appropriate mass values, while the sticks depict covalent bonds between atoms with the appropriate values of force constant. In order to setup the model elucidated above, the positional coordinates of acetylene were obtained from the Computational Chemistry Comparison and Benchmark Database in National Institute of Standards and Technology (NIST, <http://www.nist.gov>).

Once inputting acetylene's atomic masses and geometrical information into ENM, NMA computes eigenvalues and eigenvectors of the acetylene structure. To be more specific in this case, we have two different bonded force constants representing $C \equiv C$ triple bond and $C-H$ single bond, respectively. These two values will be identified by the following proposed optimization scheme. The computed eigenvalues and eigenvectors represent corresponding wavenumbers (i.e., vibration frequencies) and modeshapes of acetylene's vibration modes, respectively. The animation was utilized to represent the perturbed positions of the atoms in the given mode. For example, Fig. 2 represents all the three possible modeshapes in acetylene such as $C \equiv C$ stretching, $C-H$ symmetric and asymmetric stretching. Thereby, the proposed ENM based NMA enables us to correlate wavenumbers to their modeshapes, which has never been done by any experiment.

2.3 Optimization of computed force constants

According to experimental frequencies of acetylene [20], it has three fundamental vibration modes which are 1974 cm^{-1} for $C \equiv C$ triple bond stretching, 3289 cm^{-1} for $C-H$ single

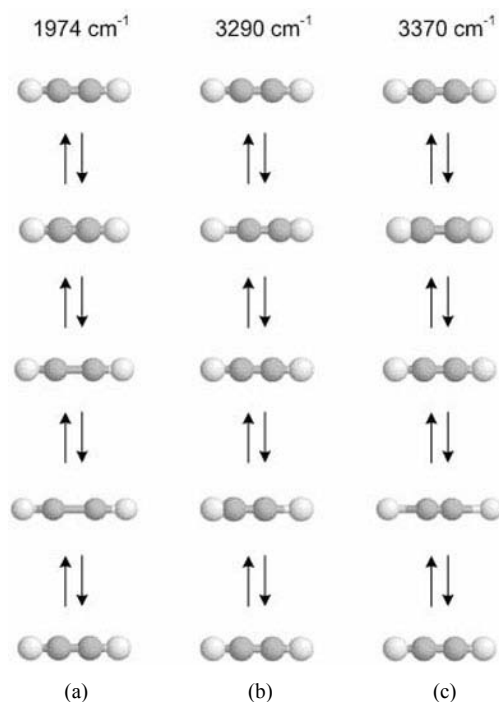


Fig. 2. Animations of acetylene's vibration modes: (a) $C \equiv C$ triple bond stretching mode (b) $C-H$ single bonds symmetric stretching mode (c) $C-H$ single bonds asymmetric stretching mode.

bond asymmetric stretching, and 3374 cm^{-1} for $C-H$ single bond symmetric stretching, respectively. Since the acetylene's modeshapes were identified from computed eigenvectors, the corresponding wavenumbers were also recorded. In order to match those vibration frequencies reported from NIST, an optimization scheme was adopted to minimize the error in the computed wavenumbers. As wavenumber is a function of force constants and masses involved in Eq. (1), the force constants can be determined by optimizing computed wavenumbers as close to the literature as possible.

To do this, we generated a 3D plot of percentage difference between computed and experimental data in wavenumber against the combination of the two spring constants for $C \equiv C$ triple bond and $C-H$ single bond in the case of acetylene. Fig. 3 shows an example plot corresponding to $C \equiv C$ stretching mode in which we can determine the final spring constants based on a minimum value of percentage difference in wavenumber. Two more similar plots were generated for $C-H$ symmetric and asymmetric stretching modes, respectively (not displayed here). Based on the inference from these plots, the values for force constants in acetylene were computed and tabulated in Table 1. In addition, the three fundamental vibration modes each of which includes its frequency and a unique modeshape were also obtained by NMA (Fig. 2). An identical methodology was then employed into use for cyanogen (C_2N_2 , $N \equiv C-C \equiv N$, experimental frequencies from reference [21]). Its results are also summarized in the Table 1. We then extended this parameterization process to other small chemical compounds to identify various force constants for the most

Table 1. A list of bond specific force constants. Comparison between experimental and predicted frequencies yields specific force constants of each pair of atoms. The obtained force constants can be assigned to appropriate spring constants in ENM.

Molecule	Bond	Experimental Frequency (1/cm) [21,22]	Predicted Frequency (1/cm)	Difference (%)	Computed Force Constants (10^5 dynes/cm)
Acetylene	C \equiv C	1974	1974	0.00	15.71
	C–H	3289 (Asym.)	3290	0.03	5.94
	C–H	3374 (Sym.)	3370	0.12	5.94
Cyanogen	C–C	846	846	0.00	6.99
	C \equiv N	2158 (Asym.)	2150	0.37	17.64
	C \equiv N	2330 (Sym.)	2424	4.03	17.64

Table 2. A list of computed force constants for various atom pairs to be used in amino acid modeling.

Bond	Force Constants (10^5 dynes/cm)
C–C	6.99
C=C	10.94
C–H	5.94
C–O	5.025
C=O	14.22
C–N	4.98
C=N	10.63
C–S	3.05
H–S	2.95
H–O	7.74
H–N	5.41
S–S	2.5

common covalent bonds that appear in protein structures. The results are listed in Table 2 and they will be used as a starting point to construct the chemical information based ENM for protein structures.

3. Simulation results

3.1 Linear molecules (C_3HN and C_4H_2)

To test the reliability of computed force constants in Table 1, we set up the ENMs for ethynyl isocyanide (C_3HN , $N \equiv C - C \equiv C - H$) and diacetylene (C_4H_2 , $H - C \equiv C - C \equiv C - H$) with those force constants and performed NMA to compute their vibration frequencies. The predicted data are compared with the experimental data [20, 22] in Table 3. It suggests that the computed force constants give a good prediction of the frequencies for various stretching modes in diacetylene and ethynyl isocyanide.

A substantially vast range of force constant values have

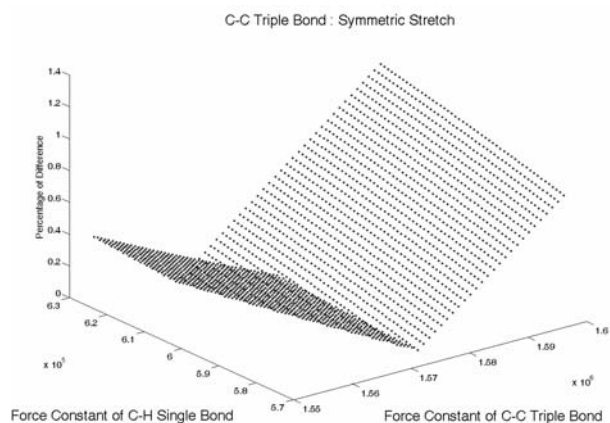


Fig. 3. Optimization of spring constants for C \equiv C triple bond symmetric stretching mode. The variation of percentage of difference between the computed and the experimental wavenumbers is represented along the Z axis with respect to the two bonds (i.e. spring constants of C \equiv C and C–H bonds along the X and Y axes, respectively).

been reported in these molecules. Wu and Shen [23] determined the value of stretching force constant for C–C single bond to be 3.58×10^5 dynes/cm. Similar low values have been also reported elsewhere [24, 25]. However, Turrell [26] reported a force constant of 7.83×10^5 dynes/cm for the same bond in diacetylene based on an observed (i.e., Raman active) symmetrical stretching mode of 874 cm^{-1} . Although these two extreme values are excluded, a range of values for C–C single bond have been reported, spanning from 4.5×10^5 dynes/cm to 6.7×10^5 dynes/cm depending on the compound containing this bond itself and its orientation.

Similarly, our NMA computes the similar force constant to what have been mentioned in the literature, but unlike some other contemporary studies, it can be stated that for a particular bond, the corresponding force constants are almost unique and some sort of invariance can be attributed to the surrounding atoms and the interactions among them. The closeness of the computed frequencies to those obtained from experimental data reaffirms this finding.

3.2 Amino acid (cysteine)

Cysteine has been a widely studied amino acid due to the availability of substantial crystallographic data, computationally generated low energy metastable states elucidating the structural details of its numerous possible conformations [27–34], and the relative simplicity of the orientation of the constituent atoms. Disparate studies discussing the vibration spectrum assignment in cysteine have been conducted and reported by several authors [35–38]. These efforts have been necessitated by the fact that vibration assignment is predominantly governed by the structure of a given amino acid, which in turn is dependent on the orientation of the constituent molecules.

In nature, like other amino acids, cysteine has been found in many different conformations. Assignment of a unique vibration spectrum to a generalized coordinate set could be consid-

Table 3. Comparison between experimental and predicted frequencies. Vibration modes in ethynyl isocyanide and diacetylene are predicted, respectively, using the computed bond specific force constants listed in Table 1 and then they are compared with nominal experimental values obtained experimentally.

Molecule	Bond	Experimental Frequency (1/cm) [20,22]	Predicted Frequency (1/cm)	Difference (%)
N≡C-C≡C-H (Ethynyl isocyanide)	C-H	3327	3334	0.21
	C≡C	2079	2089	0.48
	C-C	864	851	1.50
	C≡N	2274	2392	5.19
H-C≡C-C≡C-H (Diacetylene)	C-C	2020	2038	0.89
	C≡C	2184	2318	6.14
	C-H	3293	3319	0.79
	C-H	3329	3335	0.18

ered as a rather naive approach. On the other hand, with the structural complexities of the examined structures and the realization of a huge set of possible solutions coupled with the limited structural information from experimental techniques such as X-ray crystallography, accurate vibration assignment almost seems like a daunting task. It is also reported that a detailed study on the constituent force fields has been done in terms of internal coordinates for better representation of the intra-molecular and inter-molecular forces. They can be subsequently applied to modeling of complex biological systems and then simulations run for obtaining the results to be used for vibration spectrum assignment [39-43].

The low-frequency vibrations are typically dominated by non-covalent, inter-molecular interactions such as electrostatic, Van der Waals, and hydrogen bonds [44]. These lower modes seem to be much related to more global and collective motions among non-bonded constituent atoms even from a single amino acid to large macromolecules. In contrast, higher modes involve more of individual bond stretching between a pair or pairs of atoms. So they mainly depend on bonded force constants [45].

Using this approach of identifying and distinguishing bonded and non-bonded interactions, to facilitate a suitable and approximate replication of the naturally existing force fields, we examine cysteine to investigate its vibration spectra which provides us with an ordered set of frequencies and corresponding modeshapes. As can be seen in both Table 4 and Fig. 4, the first four modes predicted by NMA using the chemical information based ENM are almost identical to the ones reported in the literature.

In accordance with the hypothesis put forward in the previous section, it is indeed observed that the lower frequencies,

Table 4. The first four mode comparison between experimental and predicted frequencies for cysteine. A generalized force constant of 7×10^5 dynes/cm, a non-bonded force constant of 6×10^3 dynes/cm, and a cutoff distance of 8Å were applied when computing vibration frequencies.

Molecule	Experimental Frequency (1/cm) [46]	Predicted Frequency (1/cm)
L- Cysteine	46.79	46.80
	55.53	55.53
	71.64	71.65
	79.81	79.81

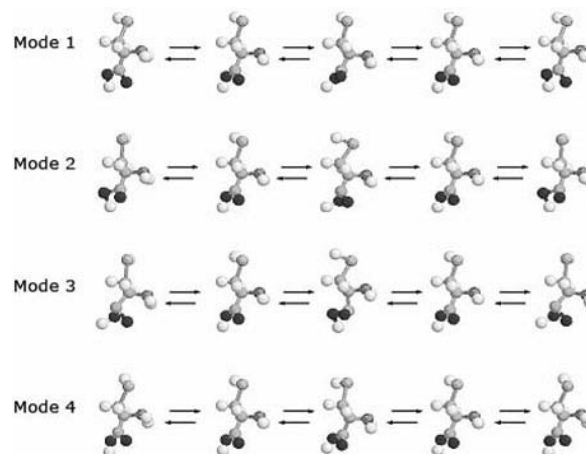


Fig. 4. Modeshape animations of cysteine for the first four computed modes.

more global modes, are strongly dependent on the connectivity of the model which can be determined by the scheme of force fields. Simulation results summarized graphically in Fig. 5 and numerically in Appendix A strongly support the fact that variation on non-bonded force constants under the same connectivity does not affect too much the low mode frequencies (not absolute values but normalized ones) and their modeshapes. One can also recognize that there is significant variation on higher modes due to the simplification of force constants for various covalent bonds as a generalized force constant. More precise force fields would be required to accurately determine the frequencies for higher modes with predominantly local vibrations.

Nonetheless, the observed unique mode number can be utilized to provide us with a sequentially arranged modeshapes. Namely, although the values of frequencies for certain higher modes might not be precise yet, the information on sequence of animations is very useful to give an ordered set of modeshapes assorted in ascending values of their corresponding frequencies. Consequently, the model itself can provide the user with a unique mode number below which swapping of modes cannot take place for the given force fields. Animations therefore can be suitably utilized for mode shape assignment.

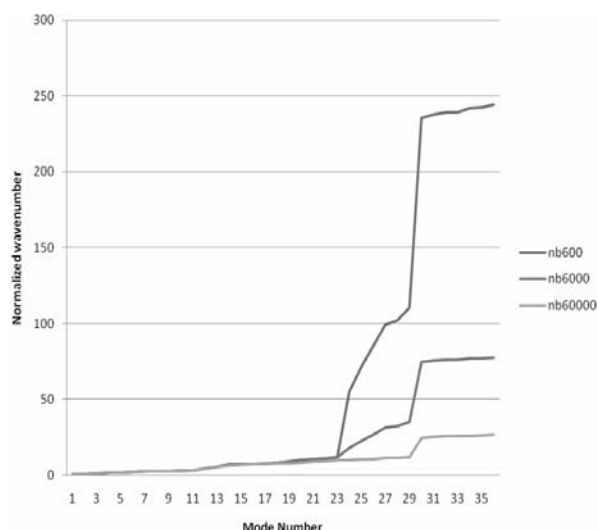


Fig. 5. A plot of normalized wavenumbers versus the mode number for the case of generalized force fields which elucidates the observed divergence at higher modes resulting from a variation in the non-bonded force constants from 6×10^2 dynes/cm to 6×10^4 dynes/cm. See Appendix A for more details.

4. Discussion

4.1 Force field parameterization

The robustness of the fundamental behavior of the modeling scheme adopted to generate the reported results has to be ascertained to safely assume repeatability and consistency in the outcome. A sensitivity analysis has been mandatorily done to understand the effects of variations in input parameters on the output. There are primarily three input variables; bonded and non-bonded force constants as well as the connectivity of the constituent atoms.

Bonded force constants were determined by performing NMA on linear molecules, so as to evaluate the values for the same by comparing the output for bond stretching from that observed from the reported values in the literature. For instance, force constants for acetylene and cyanogen have been reported in the Table 1. Similar methodology was adopted to compute the force constants for numerous other pairs of atoms as shown in Table 2.

Hence, the approach adopted in this paper was to run many simulations with different combinations of these parameters (i.e. to optimize the outcome based on the proposed hypothesis of using a generalized set of force fields resulting in Table 4). Firstly, ratios of bonded and non-bonded force constants were varied by several orders of magnitude. Subsequently, eigenvector sets for the perturbed conditions were compared to one another and animations for modeshapes were generated to observe any discrepancies. Also, the corresponding frequencies of these vibrations were normalized and the results were plotted in Fig. 5. In this context, normalized wavenumbers have been defined as the ratio of frequencies of modes to the slowest mode. That is, while the first mode is considered

as unity, all the subsequent modes can be expressed in terms of the first mode.

For instance, values for vibration frequencies for cysteine with different ratios of bonded and non-bonded force constants have been summarized in Appendix A. These values were obtained from both a generalized force constant of 7×10^5 dynes/cm and bond specific force constants, respectively, in combination with the following three different non-bonded force constants; 6×10^2 dynes/cm, 6×10^3 dynes/cm, and 6×10^4 dynes/cm. Although the values of normalized wavenumbers are of little significance since they are merely ratios of all the mode frequencies with respect to the first mode yet, the plots illustrated in Fig. 5 give us an insight into model's behavior. Such an analysis was required to substantiate the effect of variations in non-bonded force constants as well as to optimize the proposed simulation scheme. Moreover, animations suggest that these induced variations in the input parameters do not alter the modeshapes up to a certain critical mode number (i.e., 23 in the case of cysteine). This invariance in the major modeshapes can be attributed to the robustness of the proposed modeling technique which can catch the nature of system dynamics without loss of generality.

The above sensitivity analysis suitably explained the effects of variations in the input parameters on output behaviors and also facilitated the use of a generalized force field parameterization scheme in addition to the specified non-bonded force constants in form of an exponentially decreasing function.

4.2 Sensitivity to the cutoff distance

In addition to understanding the effect of force field parameterization, cutoff distance values were also varied to further examine the dependence of vibration behaviors of ENM on its connectivity matrix which replicates the interacting pairs of atoms. Since the construction of linking matrix totally depends upon a certain cutoff distance used, disparate tests with different cutoff values were performed and their results have been summarized in Appendix B.

To understand the effects of intra-molecular connectivity on the vibration behaviors, we not only computed the wavenumbers but also generated animations for the corresponding modeshapes. While a set of wavenumbers for cutoff values of 8Å and 5Å are almost identical, they are very different from the case with the cutoff value of 3Å. It can be attributed to the fact that global motions are dominantly dependent on non-bonded acting pairs, and in turn, on its linking matrix. As for the higher modes, the values tend to converge shown in both Fig. 6 and Appendix B because these vibrations are characterized as local vibrations pertaining to bond stretching and bending. Hence, this result is much in agreement with the proposed hypothesis.

Again, higher modes consisting of more local vibrations are less related to non-bonded intra-atomic interactions. Therefore, as expected, the sensitivity analysis with different values of cutoff distance yields disparate linking matrices so that the

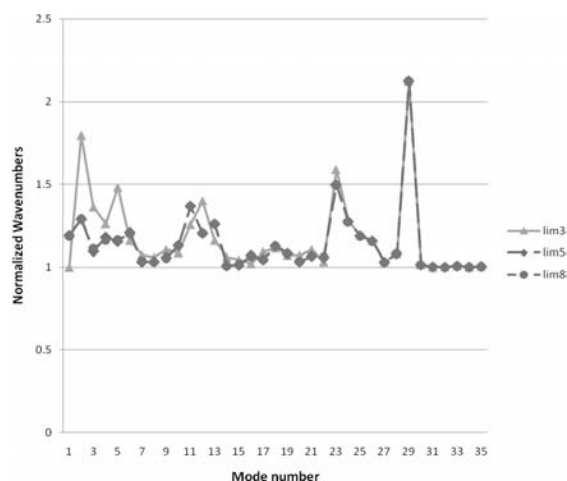


Fig. 6. Plot of the ratio between two consecutive wavenumbers ω_{i+1}/ω_i against mode number representing the variance of wavenumbers in terms of cutoff distance. Since the first non-zero mode with a cutoff of 3Å is unrealistically small, the computed value of ω_2/ω_1 was discarded and the arbitrary number of 1 is instead chosen to match scales. See Appendix B for more details.

modeshapes can be also different for the lower modes while subsequent animations suggest that the higher modes be identical.

5. Conclusions

The ENM based NMA has been successfully used to investigate vibration modeshapes of macromolecules which are strongly related to their functional behaviors. However, the traditional scheme of ENM has too much simplified the sophisticated force fields into an elastic network connected by dimensionless binary springs even neglecting mass weight. So the vibration frequency cannot be computed from the current methodology. On the other hand, vibration spectroscopy techniques such as Raman and infrared spectroscopies have been widely used as an important analytical tool for chemical and biological studies by detecting unique vibration frequencies of molecules as their "finger print". It also has a limitation to the identification of the tested compounds.

To assign modeshapes to corresponding vibration frequencies of macromolecules, we propose a chemical information based ENM embedding precise spring constants. It is designed to compute both realistic vibration frequencies and corresponding modeshapes. Comparing this theoretical output to experimentally obtained spectroscopy data, we can identify modeshapes of the wavenumbers of our interest.

We first determined bonded force constants by comparison between the computed and the measured wavenumbers of small molecules. Linear molecules have been used in this parameterization process to minimize the coupling effects caused by non-bonded interactions. Since all involved atoms line up in linear molecules, their stretching modes are predominantly affected by strong bonded interactions but the

effect of non-bonded interactions is negligible on both symmetric and asymmetric stretching vibration modes. As a result, the bonded force constants have been obtained and applied to yield the same values of frequency as the experimental data in the two other testing linear molecules such as ethynyl isocyanide and diacetylene.

The proposed method has been extensively applied to one of amino acid structures, cysteine. The simulation output was compared with the experiment data obtained from terahertz spectroscopy [46]. Both animations and the corresponding frequencies suggest a favorable trend in precise prediction of frequencies and modeshapes, especially for the first few slower but functionally very important modes. Moreover, it has been observed that the computed frequencies are indeed fundamental modes, and in fact, are not characterized by lattice vibrations as proposed elsewhere [46]. A deviation for higher modes was an expected outcome due to the use of approximate force fields. While a more accurate force field parameterization, accounting for both bonded and non-bonded interactions, will facilitate a better prediction of vibrational frequencies in amino acids in a manner elucidated for cysteine, the statistical data for macromolecules suggest that a majority of biological functions be observed in the low frequency domain and be associated with a conformational change. This would result in designing a more appropriate connectivity matrix because the connectivity is a crucial factor in the determination of low modes.

It is expected that this novel methodology can be employed as a new paradigm of vibration assignment scheme. Since the modeshapes as well as vibration frequencies can be computed by NMA with the chemical information based ENM, it will enable us to not only reveal vibration spectrum of a given macromolecule but also visualize the corresponding modeshapes which is not currently impossible for large macromolecules in either traditional ENM based simulations nor spectroscopy experiments. It is also expected that more complicated biological mechanism such as large conformational change in proteins and protein-protein interactions would be investigated by the proposed computational method in aid of spectroscopy experiment.

Acknowledgement

This work was equally supported by both WCU program (R33-2008-000-10079-0) under the Ministry of Education, Science and Technology and Energy Resources Technology R&D program (R200811041) under the Ministry of Knowledge Economy, Republic of Korea. The authors are also grateful for the support of a Micro Raman device (Ramboss 500i) provided by DongWoo Optron Co. Ltd.

References

- [1] D. J. Millen, Determination of stretching force constants of weakly bound dimers from centrifugal distortion constants,

- Can. J. Chem.* 63 (1985) 1477-1479.
- [2] P. Phillipson, Electronic Bases of Molecular Vibrations. I. General Theory for Diatomic Molecules, *J. Chem. Phys.* 39 (1963) 3010-3016.
- [3] B. K. Novosadov, Forces in molecules equations for the calculation of force constants, *J. Struct. Chem.* 16 (6) (1975) 891-894.
- [4] V. I. Bazhanov, Statistical method for determination of structural parameters of molecules from electron diffraction data, *J. Struct. Chem.* 27 (1) (1986) 30-35.
- [5] A. Navarro, M. P. Fernandez-Liencres, A. Ben Altabef, M. F. Gomez, J. J. L. Gonzalez and R. Escribano, A set of force constants common for SiH₃-X molecules, *J. Mol. Struct.* 482 (1999) 601-607.
- [6] K. Venkateswarlu and S. Sundaram, Evaluation of Force Constants from Raman Effect Data: Molecules, Radicals and Groups of Pyramidal XY₃ Type, *Proc. Phys. Soc. A.* 69 (1955) 180-183.
- [7] O. Alvarez-Bajo, M. Sanchez-Castellanos, C. A. Amezcua-Eccius and R. Lemus, Force constants and transition intensities in the U(v+1) model for molecular vibrational excitations, *J. Mol. Spectrosc.* 237 (2006) 247-258.
- [8] L. Bizzocchi, C. D. Esposti, A. Mazzavillani and F. Tamasia, Fourier transform infrared spectroscopy of the 2 nu(3) overtone band of different ICN isotopomers: an improved evaluation of the anharmonic force field of cyanogen iodide, *J. Mol. Spectrosc.* 221 (2003) 213-220.
- [9] A. E. Özel, Y. Büyükmurat and S. Akyüz, Infrared-spectra and normal-coordinate analysis of quinoline and quinoline complexes, *J. Mol. Struct.* 565 (2001) 455-462.
- [10] P. D. Kieng, A. B. Kovrikov and A. I. Komyak, Estimating force constants and calculating normal oscillations of the (U O₂C₁₄ · 2H₂O)²⁻ anionic complex in a rubidium uranyl chloride crystal, *J. App. Spec.* 18 (3) (1973) 336-340.
- [11] M.C. Gupta, *Atomic and Molecular Spectroscopy*, New Age International(P) Limited, New Delhi, (2001).
- [12] M. K. Kim, G. S. Chirikjian, R. L. Jernigan, Elastic models of conformational transitions in macromolecules, *J. Mol. Grap. Model.* 21 (2002) 151-160.
- [13] M. K. Kim, R. L. Jernigan, G. S. Chirikjian, Efficient generation of feasible pathways for protein conformational transitions, *Biophysical J.* 83 (2002) 1620-1630.
- [14] M. K. Kim, Li. W., B. A. Shapiro, G. S. Chirikjian, A comparison between elastic network interpolation and MD simulation for 16S Ribosomal RNA, *J. Biomol. Struct. and Dyn.* 21 (2003) 395-405.
- [15] M. K. Kim, R. L. Jernigan, G. S. Chirikjian, An elastic network model of HK97 capsid maturation, *J. Struct. Biol.* 143 (2003) 107-117.
- [16] M. K. Kim, R. L. Jernigan, G. S. Chirikjian, Rigid-cluster models of conformational transitions in macromolecular machines and assemblies, *Biophysical J.* 89 (2005) 43-55.
- [17] M. K. Kim, Y. Jang, J. I. Jeong, Using harmonic analysis and optimization to study macromolecular dynamics, *Int. J. Control Autom. Syst.* 4 (2006) 382-393.
- [18] Y. Jang, J. I. Jeong, M. K. Kim, UMMS: constrained harmonic and anharmonic analyses of macromolecules based on elastic network models, *Nucleic Acids Research* 34 (2006) W57-W62.
- [19] J. I. Jeong, Y. Jang, M. K. Kim, A connection rule for alpha-carbon coarse-grained elastic network model using chemical bond information, *J. Mol. Grap. Model.* 24 (2006) 296-306.
- [20] T. Shimanouchi, *Tables of Molecular vibrational Frequencies, Consolidated. Volume I*, Superintendent of Documents-U.S. Government Printing Office, (1972).
- [21] M. E. Jacox, Vibrational and Electronic Energy Levels of Polyatomic Transient Molecules, *Journal of Physical and Chemical Reference Data*, (1994).
- [22] O. V. Dorofeeva and L. V. Gurvich, Thermodynamic properties of linear carbon chain molecules with conjugated triple bonds: Part I. Polyacetylenes, H(C≡C)_nH (n = 2–6), cyanopolyacetylenes, H(C≡C)_nC≡N (n = 1–5), and dicyanopolyacetylenes, N≡C(C≡C)_nC≡N (n = 1–4), *Thermochimica Acta*, 12 (1991) 273-286.
- [23] T. Y. Wu and S. T. Shen, Chin, On the Force Constants and Fundamental Vibrations of Diacetylene, *J. Phys.* 2 (1936) 128.
- [24] S. M. Ferigle, F. F. Cleveland and A. G. Meister, Force Constants and Fundamental Frequencies of Dimethyldiacetylene and Diacetylene, *Phys. Rev.* 81 (1951) 302.
- [25] G. Herzberg. *Infrared and Raman spectra of polyatomic molecules*, Van Nostrand Reinhold, New York, USA, (1945).
- [26] G. C. Turrell, Infrared Spectra and Force Constants of Cyanoacetylene, *J. Chem. Phys.* 26 (1957) 1544.
- [27] J. Cz. Dobrowolski, J. E. Rode and J. Sadlej, Cysteine conformations revisited, *J. Mol. Struct.* 810 (2007) 129-134.
- [28] A. Fernandez-Ramosa, E. Cabaleiro-Lagoa, J. M. Hermida-Ramona, E. Martinez-Nuneza and A. Pena-Gallegoa, DFT conformational study of cysteine in gas phase and aqueous solution, *J. Mol. Struct.* 498 (2000) 191-200.
- [29] C. C. Wang, J. H. Chen, S. H. Yin and W. J. Chuang, Predicting the redox state and secondary structure of cysteine residues in proteins using NMR chemical shifts, *Proteins* 63 (2006) 219-226.
- [30] M. Noguera, L. Rodriguez-Santiago, M. Sodupe and J. Bertran, Protonation of glycine, serine and cysteine. Conformations, proton affinities and intrinsic basicities, *J. Mol. Struct.* 537 (2001) 307-318.
- [31] S. Z. Liu, H. Q. Wang, Z. Y. Zhou, X. L. Dong and X. L. Gong, Theoretical study of helical structure caused by chirality of cysteine dimer, *Int. J. Quant. Chem.* 105 (2005) 66-73.
- [32] K. A. Kerr, J. P. Ashmore and T. F. Koetzle, A neutron diffraction study of L-cysteine, *Acta Cryst. B* 31 (1975) 2022-2026.
- [33] P. Tarakeshwar and S. Manogaran, Vibrational frequencies of cysteine and serine zwitterions-an ab initio assignment, *J. Mol. Struct.* 305 (1994) 205-224.
- [34] S. Gronert and R. A. J. O'Hair, Ab Initio studies of amino acid conformations. 1. The conformers of alanine, serine,

- and cysteine, *J. Am. Chem. Soc.* 117 (1995) 2071-2081.
- [35] P. Tarakeshwar and S. Manogaran, Vibrational frequencies of cysteine and serine zwitterions-an ab initio assignment, *Spectrochimica Acta. Part A* 51 (1995) 925-928.
- [36] N. N. Brandt, A. Y. Chikishev, A. V. Kargovsky, M. M. Nazarov and O. D. Parashchuk, Terahertz time-domain and Raman spectroscopy of the sulfur-containing peptide dimers: Low-frequency markers of disulfide bridges, *Vib. Spec.* 47 (2008) 53-58.
- [37] A. Barth, The infrared absorption of amino acid side chains, *Prog. Biophys. Mol. Biol.* 74 (2000) 141-173.
- [38] N. Wright, The infra-red absorption spectra of the stereoisomers of cysteine, *J. Biol. Chem.* 120 (1937) 641-646.
- [39] V. Krishnakumar, N. Prabavathi and S. Muthunatesan, Density functional theory study of vibrational spectra, and assignment of fundamental vibrational modes of 1-bromo 4-fluoronaphthalene, *Spectrochimica Acta Part A* 70 (2008) 201-209.
- [40] V. Krishnakumar, R. John Xavier, FT Raman and FT-IR spectral studies of 3-mercapto-1,2,4-triazole. *Spectrochimica Acta Part A* 60 (2004) 709-714.
- [41] F. P. Ureña, M. F. Gómez, J. L. González and E. M. Torres, A new insight into the vibrational analysis of pyridine, *Spectrochimica Acta Part A* 59 (2003) 2815-2839.
- [42] V. P. Morozov, V. V. Belokopytov, G. D. Zegzhda and V. N. Moiseenko, Force constants and vibration frequencies and forms of the coordination center of cysteine-containing chromium trischelates, *J. Struct. Chem.* 39 (1998) 508-513.
- [43] M. E. Tuttolomondo, L. E. Fernández, A. Navarro, E. L. Varetti and A. Ben Altabef, Experimental and theoretical vibrational study of 2,2,2-trifluoroethyl trifluoromethanesulfonate, CF₃SO₂OCH₂CF₃, *Spectrochimica Acta Part A* 60 (2004) 611-619.
- [44] N. Go, T. Noguti and T. Nishikawa, Dynamics of a small globular protein in terms of low-frequency vibrational modes, *Proc. Natl Acad. Sci. USA* 80 (1983) 3696-3700.
- [45] A. Barth and C. Zscherp, What vibrations tell us about proteins, *Q. Rev. of Biophys.* 35 (2002) 369-430.
- [46] T. M. Korter, R. Balu, M. B. Campbell, M. C. Beard, S. K. Gregurick and E. J. Heilweil, Terahertz spectroscopy of solid serine and cysteine, *Chem. Phys. Lett.* 418 (2006) 65-70.

Appendix A

Computed wavenumbers based on both generalized and bond specific force constants in addition to the three different values of non-bonded force constants referred to as cases; A: 6×10^2 dynes/cm, B: 6×10^3 dynes/cm, C: 6×10^4 dynes/cm. The reported values of wavenumbers are in cm^{-1} .

Mode number	Generalized force fields			Bond specific force fields		
	Non-Bonded force constants			Non-bonded force constants		
	A	B	C	A	B	C
1	14.80	46.80	147.71	14.80	46.80	147.64
2	17.57	55.53	175.08	17.57	55.53	174.93

3	22.66	71.65	225.85	22.66	71.64	225.62
4	25.25	79.81	250.87	25.25	79.79	250.11
5	29.29	92.60	292.59	29.29	92.60	292.56
6	34.18	107.94	337.08	34.18	107.82	334.21
7	41.01	129.61	406.86	41.01	129.59	405.98
8	42.53	134.39	421.82	42.53	134.37	421.04
9	43.90	138.61	431.87	43.90	138.55	430.50
10	46.38	146.17	449.90	46.37	146.04	447.24
⋮	⋮	⋮	⋮	⋮	⋮	⋮
27	1469.86	1473.66	1683.45	1382.89	1391.36	1698.3
28	1512.73	1520.15	1724.99	1553.18	1560.40	1768.0
29	1639.04	1644.98	1787.33	1960.77	1963.80	1996.2
30	3485.08	3495.69	3613.04	2263.47	2280.15	2495.8
31	3521.27	3552.06	3757.09	3096.41	3131.75	3352.3
32	3537.21	3565.45	3804.38	3155.34	3172.11	3481.5
33	3541.40	3570.52	3807.82	3258.68	3294.68	3536.4
34	3578.32	3603.39	3850.76	3297.10	3332.88	3629.6
35	3588.72	3611.5	3884.92	3331.13	3349.6	3705.0
36	3615.82	3632.8	3963.33	3722.85	3745.7	3977.6

Appendix B

Predicted wavenumbers of cysteine based on different values of cutoff distance. Underestimation of non-bonded interactions with relatively short cutoff distance results in unrealistic lower frequencies while higher modes are less sensitive to non-bonded interactions.

Mode number	Wavenumbers*			Ratio (%) $\left(\frac{\text{Case 1}}{\text{Case 3}} \times 100\right)$
	Case 1(3Å)	Case 2(5Å)	Case 3(8Å)	
1	0.15	46.62	46.80	0.32
2	19.92	55.51	55.53	35.87
3	35.77	71.56	71.65	49.92
4	48.71	78.42	79.81	61.03
5	61.38	92.54	92.60	66.29
6	90.86	106.80	107.94	84.18
7	105.37	129.13	129.61	81.30
8	112.99	133.33	134.39	84.08
9	119.61	137.93	138.61	86.29
10	132.19	146.03	146.17	90.44
⋮	⋮	⋮	⋮	⋮
27	1472.98	1473.64	1473.66	99.95
28	1519.52	1520.15	1520.15	99.96
29	1644.22	1644.97	1644.98	99.95
30	3490.11	3495.09	3495.69	99.84
31	3546.75	3551.82	3552.06	99.85
32	3549.33	3564.79	3565.45	99.55
33	3565.68	3570.52	3570.52	99.86
34	3598.23	3603.13	3603.39	99.86

35	3604.03	3611.54	3611.54	99.79
36	3628.02	3632.86	3632.86	99.87

*The non-bonded force constant C is set to be 6×10^3 dynes/cm in this test.



Mingwen Hu received the B.S. degree in Mechanical Engineering from Yuan Ze University in 2002 and then the M.S. degree in Mechanical and Mechatronic Engineering from National Taiwan Ocean University in 2004. He is currently pursuing the Ph.D. degree in Mechanical and Industrial Engineering,

University of Massachusetts-Amherst. His researches focus on investigating macromolecules by experimental and theoretical approaches which include vibrational spectroscopy and elastic network model.



Byung Kim received the B.S. in Mechanical Engineering from UC Berkeley in 1978, M.S. and Ph.D. in Mechanical Engineering from MIT in 1980 and 1983, respectively. He has joined the faculty member in the department of Mechanical and Industrial Engineering at the University of Massachusetts, Amherst since 1983. His research interests are in the area of micro and nanomanufacturing, SERS, protein dynamics.



Wing Kam Liu obtained by B.S. with the highest honor from from the University of Illinois at Chicago Circle in 1976 and received his M.S in 1977 and Ph.D. in 1981, both from California Institute of Technology. He is the Walter P. Murphy Professor of Mechanical Engineering Department of Northwestern

University, Founding Director of the NSF Summer Institute on Nano Mechanics and Materials, and Founding Chairman of the prestigious ASME NanoEngineering Council. Multi-resolution continuum theory and immersed finite element method are his main research interests.



Moon Ki Kim received the B.S. and M.S. degrees in Mechanical Engineering from Seoul National University in 1997 and 1999, respectively, and the M.S.E. and Ph.D. degrees from The Johns Hopkins University in 2002 and 2004, respectively. He had been an Assistant Professor in the department of Mechanical

and Industrial Engineering at University of Massachusetts, Amherst from 2004 to 2008. He has worked for School of Mechanical Engineering at Sungkyunkwan University as an associate professor since 2009. His research interests are focused on computational structural biology based on robot kinematics, bioinstrumentations, and multiscale modeling and simulation.