# Development of Chemical Bond based Elastic Network Model and its application in identifying functional motions in H5N1 highly pathogenic avian influenza viruses

Yunho Jang and Xiu-Feng Wan*

Department of Basic Sciences,
College of Veterinary Medicine,
Mississippi State University,
Mississippi State, MS 39762, USA
E-mail: jang.yunho@gmail.com
E-mail: wan@cvm.msstate.edu
E-mail: wanhenry@yahoo.com
*Corresponding author

**Abstract:** Here we proposed a Chemical Bond based Elastic Network Model (CB-ENM) method, which was demonstrated with a better performance in representing functional movements, especially the local motions, than the conventional normal mode analysis. CB-ENM was applied in characterising functional motions in surface glycoprotein Haemagglutinin (HA) of H5N1 highly pathogenic avian influenza virus. Eleven local peaks were identified amongst HA fluctuation spectrum, and eight of these peaks were correlated with the reported five antigenic binding sites and Receptor Binding Sites (RBS) in influenza viruses. Our study showed that CB-ENM has a potential in identifying functional residues in protein structure.

**Keywords:** CB-ENM; chemical bond based elastic network model; normal mode analysis; functional motion; influenza virus; H5N1 highly pathogenic avian influenza virus; haemagglutinin; receptor binding site; antigenic binding site.

**Reference** to this paper should be made as follows: Jang, Y. and Wan, X-F. (2010) 'Development of Chemical Bond based Elastic Network Model and its application in identifying functional motions in H5N1 highly pathogenic avian influenza viruses', *Int. J. Bioinformatics Research and Applications*, Vol. 6, No. 1, pp.1–11.

**Biographical notes:** Yunho Jang received his PhD Degree in Mechanical Engineering from University of Massachusetts Amherst in 2008. His current research interests lie in computational structural biology and bioinformatics using various computational models.

Xiu-Feng (Henry) Wan received his PhD Degree in Veterinary Science and Master Degree in Computer Science from Mississippi State University in 2002. He had his postdoc training in Oak Ridge National Laboratory and University of Missouri-Columbia. Currently, he is an Assistant Professor in systems biology at Mississippi State University. Before joining Mississippi State University, he was a senior service fellow in the Influenza Division of the Centres for Disease Control and Prevention, Atlanta, GA.

His career goal is to understand ecology, evolution, microbe-microbe, and environment/host-microbe interaction by developing/applying systems biology approaches. His current research focuses primarily on influenza A virus.

## 1 Introduction

Normal Mode Analysis (NMA) is a harmonic analysis tool for molecular dynamic simulation in illustrating mechanical vibrations. In NMA, vibrational frequencies and directions of corresponding motions in the target system are calculated by solving the generalised eigenvaules problem of the Hessian matrix. NMA has been widely used to compute thermal fluctuations of a macromolecule around its energy equilibrium conformation (Brooks and Karplus, 1983). To reduce computational burden, coarse-grained NMA has been developed (Atilgan et al., 2001; Bahar et al., 1997). On the other hand, to simply the coarse-grained NMA, conventional studies usually assign the weights of the connections between the atoms to be an equal value (e.g., a constant value of 1). However, this assumption does not reflect the practical situations since the biophysical energy strengths may be very different among these atoms, for instance, the energy of C = C ($\Delta G$ = 147 kcal/mol) is almost twice that of C–C ($\Delta G$ = 83 kcal/mol). In addition, other chemical forces, for instance, hydrogen bonds, may also affect molecular dynamics. Jeong et al. (2006) has considered some of such effects, specifically disulfide bonds, hydrogen bonds, salt-bridges and *van der Waals* forces, in $\alpha$-carbon coarse-grained NMA, and their studies showed that these four types of chemical forces play an important role in representing protein dynamics with infinitesimal motions set by normal modes.

Nevertheless, an optimal NMA should have at least two requirements: (1) need to consider all atoms but with a high computational efficiency; (2) need to consider all types of chemical bonds inside the model. To achieve this goal, a more realistic and robust network connection, so called Chemical Bond based Elastic Network Model (CB-ENM), is proposed in this study. CB-ENM is an all-atom based model, and it considers chemical forces including both covalent bonds and secondary interactions. This proposed model is expected to reflect finer and more realistic motions of protein molecules, especially those local motions, which could be misrepresented in coarse-grained ENM.

Influenza A virus is a negative-stranded RNA virus that belongs to the *Ortho myxoviridae* family (Flint et al., 2004). Influenza A virus has eight genomic segments (segment 1–8) with varying lengths from about 890 to 2341 nucleotides which encode at least 11 proteins: PB2 by segment 1, PB1 and PB1-F by 2, PA by 3, haemagglutinin (HA) by 4, nucleoprotein (NP) by 5, neuraminidase (NA) by 6, membrane protein M1 and M2 by 7, and nonstructural protein NS1 and NS2 by 8. Influenza caused by influenza A viruses has posted a great threat to our public health. In last century, three influenza pandemics occurred in 1918, 1957, and 1968, respectively. Millions of lives have been lost due to influenza and influenza related complications. During the past decades, a number of different subtypes of influenza A viruses have emerged in human. The H5N1 highly pathogenic avian influenza viruses and recent swine-origin H1N1 influenza virus are two examples of these newly emerged viruses. To the date of September 24, 2009, the H5N1 highly pathogenic avian influenza viruses have caused 442 laboratory confirmed human cases, 262 of which are fatal. During the past five years, this disease

has been spread from Asia to Middle East, Europe and Africa. The swine-origin H1N1 influenza A viruses emerged in North America and causes the ongoing influenza pandemic. As of November 20, 2009, at least 6750 deaths are caused by this novel influenza.

Analysis of the molecular dynamics of influenza proteins, such as surface glycoproteins HA and NA, will shed some lights on the protein ligand interaction thus provide useful information for drug design and vaccine preparation towards disease prevention and control. Isin et al. (2002) applied $\alpha$-carbon coarse-grained *Gaussian* network model in studying the global motions of HA, and their results illustrated the global bending and twisting motions, which were suggested to be involved in viral fusion and recognition of host receptors, respectively. However, local motions for HA are still not characterised, and these local motions may be useful in identifying functional residues in HA.

In this study, we first developed an all-atom CB-ENM, and then compared this method with the conventional Constant Value based ENM (CV-ENM). Our comparisons were performed using five sets of proteins, which have two different conformations (open and closed). Our results showed that CB-ENM had a better performance than CV-ENM. We applied CB-ENM in characterising functional motions in HA of H5N1 highly pathogenic avian influenza viruses, and we identified 11 local peaks amongst the fluctuation spectrum, eight of which were related to the reported antigenic binding sites (A–E) and RBS in HA of influenza viruses.

## 2 Methodology

### *2.1 Normal mode analysis for harmonic fluctuation*

Given a set of coordinates of $n$ representative atoms, the global mass matrix and the global stiffness (i.e., Hessian) matrix can be derived. The position of the $i$th atom at time $t$ is denoted as

$$\vec{x}_i(t) = \left[x_i(t),\, y_i(t),\, z_i(t)\right]^T \in R^3. \tag{1}$$

The total kinetic energy has the form

$$T = \frac{1}{2}\sum_{i=1}^{n} m_i \left\|\dot{\vec{x}}(t)\right\|^2. \tag{2}$$

If we define $\vec{\delta}_i(t)$ as a vector of small displacement from the initial position $\vec{x}_i(0)$ such that

$$\vec{x}_i(t) = \vec{x}_i(0) + \vec{\delta}_i(t) \tag{3}$$

then equation (3) can be rearranged such that

$$T = \frac{1}{2}\dot{\vec{\delta}}^T M \dot{\vec{\delta}} \tag{4}$$

where

$$\vec{\delta} = \left[ \vec{\delta}_1^T, \ldots, \vec{\delta}_n^T \right]^T \in R^{3n}. \tag{5}$$

In the present case, the global mass matrix $M$ is diagonal. The total potential energy has the form

$$T = \frac{1}{2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} k_{ij} \left\{ \left\| \vec{x}_i(t) - \vec{x}_j(t) \right\| - \left\| \vec{x}_i(0) - \vec{x}_j(0) \right\| \right\}^2 \tag{6}$$

and

$$k_{ij} = \begin{cases} 1 & \text{if} \quad \left\| \vec{x}_i - \vec{x}_j \right\| \le d \\ 0 & \text{if} \quad \left\| \vec{x}_i - \vec{x}_j \right\| > d \end{cases} \tag{7}$$

where $d$ is a cutoff distance between $C_\alpha$ atoms and $k_{ij}$ is the ($i$th, $j$th) element of $k$ called the linking matrix being unity for all contacting pairs and zero for pairs not in contact. For the small deformations, it can be approximated in a classical quadratic form

$$V = \frac{1}{2} \vec{\delta}^T K \vec{\delta} \tag{8}$$

where $K$ is called the stiffness matrix. This matrix is nothing more than the second derivatives of the harmonic potential function with respect to the generalised coordinates.

Finally, the equations of motion for a coarse-grained protein model can be obtained as

$$M \ddot{\vec{\delta}} + K \vec{\delta} = \vec{0} \tag{9}$$

The fluctuation dynamics of the structure can be obtained by solving equation (9). Eigenvalues of the weighted stiffness matrix are the squared natural frequencies of the harmonic motions and the corresponding eigenvectors reflect the mode shapes.

## 2.2 Stiffness ration in NMA

In order to reflect the realistic interactions between atoms in the network, the stiffness ration used in CB-ENM is directly computed from the chemical bond dissociation energies (kcal/mol, CRC reference). Only atoms within the distance of 5Å are considered with *Van der Waals* forces. In CV-ENM, we used the same constant value of 1 for all connections. Table 1 represents stiffness ration used in CB-ENM and CV- ENM.

**Table 1** Stiffness ration of CV-ENM and CB-ENM

| *Type of bond* | *CV-ENM* | *CB-ENM* |
|---|---|---|
| H–N | 1 | 93 |
| H–C | 1 | 99 |
| H–O | 1 | 111 |
| C–C | 1 | 83 |
| C–N | 1 | 73 |
| C–O | 1 | 85 |

**Table 1** Stiffness ration of CV-ENM and CB-ENM (continued)

| *Type of bond* | *CV-ENM* | *CB-ENM* |
|---|---|---|
| C–S | 1 | 65 |
| C=O | 1 | 180 |
| C=C | 1 | 146 |
| C=N | 1 | 147 |
| Hydrogen bond | 1 | 10 |
| Disulphide bond | 1 | 50 |
| Ion bond | 1 | 10 |

### *2.3 Hydrogen atoms and secondary bonds*

The addition of hydrogen atoms into the protein structure will allow the elastic network to measure the hydrogen bonds and other hydrogen derived secondary forces during normal mode analysis. In this study, Python Molecule Viewer (PMV) (Sanner, 1999) was used for adding hydrogen atoms, and HBPlus (McDonald and Thornton, 1994) was employed for recovering secondary bonds such as hydrogen bonds, disulfide bond, and ion bond.

### *2.4 Measurements of overlap values to assess the performance of CB-ENM and CV-ENM*

An overlap value represents functional relation and similarity between two meta-stable structures. The overlap value is usually utilised to evaluate performance of NMA. An overlap value varies from 0 to 1, and a larger the overlap value demonstrates a better performance. In this study, two categories of overlap values, $O_{\text{exp}}$, and $O_{\text{eni}}$ are employed to compare the performance between CB-ENM and CV-ENM. The $O_{\text{exp}}$ represents functional relation and similarity between two end conformations, e.g., open and closed forms; $O_{\text{eni}}$ represents the best pathway during conformational changes, which is not detected by the directional vector from two end conformations. For instance, during hinge bending and shear motions, conformational pathways may not be the same.

The value of $O_{\text{CV exp}}$ is computed by eigenvector $\overline{v}_i$ from CV-ENM and the directional vector of two end conformations (open and closed forms) $\Delta\overline{x}_{\text{exp}} = \overline{x}_o - \overline{x}_c$ such that

$$O_{\text{CV exp}} = \frac{|\,\overline{v}_i \cdot \Delta\overline{x}_{\text{exp}}\,|}{(\overline{v}_i \cdot \overline{v}_i)^{1/2}(\Delta\overline{x}_{\text{exp}} \cdot \Delta\overline{x}_{\text{exp}})^{1/2}}. \tag{10}$$

Likewise, $O_{\text{CB exp}}$ is estimated by eigenvector $\overline{v}_{\text{CB}i}$ from CB-ENM and the directional vector used for $O_{\text{exp}}$ such that

$$O_{\text{CB exp}} = \frac{|\,\overline{v}_{\text{CB}i} \cdot \Delta\overline{x}_{\text{exp}}\,|}{(\overline{v}_{\text{CB}i} \cdot \overline{v}_{\text{CB}i})^{1/2}(\Delta\overline{x}_{\text{exp}} \cdot \Delta\overline{x}_{\text{exp}})^{1/2}}. \tag{11}$$

The $O_{\text{CBeni}}$ is compared by eigenvector $\overline{v}_{\text{CB}i}$ from CB-ENM and the directional vector, $\Delta\overline{x}_{\text{eni}} = \overline{x}_o - \overline{x}_1$, from elastic network interpolation (Kim et al., 2003), which can generate more feasible pathways between end conformations.

$$O_{\text{CBeni}} = \frac{|\,\overline{v}_i \cdot \Delta\overline{x}_{\text{eni}}\,|}{(\overline{v}_{\text{CB}i} \cdot \overline{v}_{\text{CB}i})^{1/2}(\Delta\overline{x}_{\text{eni}} \cdot \Delta\overline{x}_{\text{eni}})^{1/2}}. \tag{12}$$

### *2.5 Datasets*

In order to compare the performance of CB-ENM with that of CV-ENM in terms of physical reality, we have applied them to five sets of protein structures downloaded from PDB database (www.pdb.org), each of which has two significantly different conformations (e.g., open and closed forms). In addition, conformational transitions between two meta-stable structures of the same protein in these datasets were experimentally demonstrated to be involved in important biological processes, such as enzymatic catalysis regulation, transportation, and aggregation. These structures include calmodulin (PDB ID: 1cfd and 1cll), HIV-protease (1hhp and 1ajx), LAO binding protein (2lao and 1lst), triglycerid lipase (3tgl and 4tgl), and tyrosine phosphatase (1ypt and 1yts).

To characterise the functional motions of surface glycoprotein HA of H5N1 highly pathogenic avian influenza viruses, the HA tertiary structure of A/Vietnam/1203/2004 (PDB ID: 2fk0) is utilised in this study for CB-ENM analysis.

## 3 Results

### *3.1 Comparison of CB-ENM with CV-ENM*

To perform normal mode analyses, we conducted experiments using both CB-ENM and CV-ENM for the cutoff distance between 4Å and 12Å. The harmonic motions were not stable in CV-ENM with the cutoff distance of 4Å, but both models showed stable conditions at the cutoff of 5Å. Thus, we applied 5Å as the distance cutoff for both CE-ENM and CV-ENM in our comparisons. Each model performs with open forms since generally open forms are better defined for their conformational changes (Tama and Sanejouand, 2001). We did three types of comparisons: the largest overlap values over the first 30 modes; the overlap values for low modes; and the overlap values for high modes.

#### *3.1.1 Comparison of CB-ENM with CV-ENM for the first 30 modes*

We calculated the $O_{\text{exp}}$ values from the first 30 modes for both CB-ENM and CV-ENM. Table 2 shows the largest overlap values among these analyses, which reflects the conformational change with directional vector of $\Delta\overline{x}_{\text{exp}}$ and $\Delta\overline{x}_{\text{eni}}$. Our results show that CB-ENM performs better in simulating conformational changes in all cases. For instance, in the case of tyrosine phosphatase, $O_{\text{exp}}$ for CB-ENM is 0.50 while that for CV-ENM is 0.42. This higher $O_{\text{exp}}$ value from CB-ENM represents a more realistic conformational change with better harmonic movements.

**Table 2** The largest overlap values for the first 30 modes

| *Protein name* | *PDB ID* | | *Number of residues* | $O_{CVexp}$ | $O_{CBexp}$ | $O_{CBeni}$ |
|---|---|---|---|---|---|---|
| | *Open conformation* | *Closed conformation* | | | | |
| Calmodulin | 1cfd | 1cll | 144 | 0.69(1)* | 0.70(1) | 0.84(1) |
| HIV-protease | 1hhp | 1ajx | 99 | 0.79(8) | 0.81(8) | 0.81(8) |
| LAO binding protein | 2lao | 1lst | 238 | 0.93(1) | 0.95(1) | 0.96(1) |
| Triglycerid lipase | 3tgl | 4tgl | 265 | 0.32(12) | 0.34(12) | 0.37(12) |
| Tyrosine phosphatase | 1ypt | 1yts | 278 | 0.42(2) | 0.50(2) | 0.55(2) |

*The number in the parenthesis indicates the mode number with the largest overlap values.

Since $O_{\mathrm{CB\,exp}}$ is better than $O_{\mathrm{CV\,exp}}$, we only measure the overlap values of CB-NMA ($O_{\mathrm{CBeni}}$) for elastic network interpolation. The high $O_{\mathrm{CBeni}}$ values demonstrate that CB-ENM is fit for elastic network interpolation.

### *3.1.2 Comparisons of CB-ENM with CV-ENM at low and high modes*

In normal mode analysis, higher frequency motions are more likely to be localised in low mode. To the best of our knowledge, there has been no clear boundary between low and high modes. In this study, we set 12th mode as a boundary after analysing the fluctuation spectrum and the overall value in each mode. In order to further examine how CB-ENM excels CV-ENM (e.g., at low modes or high modes), we further separate the first 30 modes into two categories: low modes (1st–12th) and high modes (13th–30th). The overlap values for both low and high modes were computed. Table 3 shows the norm of 12 low modes and that of 18 high modes.

**Table 3** Overlap values for low modes and high modes

| *Protein name** | *Low modes (1–12th)* | | | *High modes (13–30th)* | | |
|---|---|---|---|---|---|---|
| | $\sqrt{\sum_{i=1}^{12}(O^{i}_{\mathrm{CV\,exp}})^2}$ | $\sqrt{\sum_{i=1}^{12}(O^{i}_{\mathrm{CB\,exp}})^2}$ | $\sqrt{\sum_{i=1}^{12}(O^{i}_{\mathrm{CBeni}})^2}$ | $\sqrt{\sum_{i=1}^{12}(O^{i}_{\mathrm{CV\,exp}})^2}$ | $\sqrt{\sum_{i=1}^{12}(O^{i}_{\mathrm{CB\,exp}})^2}$ | $\sqrt{\sum_{i=1}^{12}(O^{i}_{\mathrm{CBeni}})^2}$ |
| Calmodulin | 0.78 | 0.78 | 1.00 | 0.31 | 0.34 | 0.34 |
| HIV-protease | 0.84 | 0.85 | 0.85 | 0.78 | 0.78 | 0.78 |
| LAO binding protein | 0.97 | 0.98 | 0.99 | 0.23 | 0.23 | 0.24 |
| Triglycerid lipase | 0.31 | 0.33 | 0.53 | 0.59 | 0.70 | 0.78 |
| Tyrosine phosphatase | 0.66 | 0.70 | 0.89 | 0.61 | 0.63 | 0.63 |

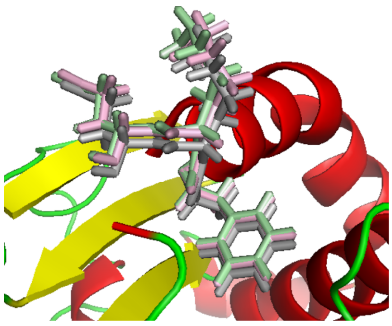
*The PDB ID was shown in Table 2.

In both low mode and high mode, CB-ENM has better performance than CE-ENM. From the analysis of triglycerid lipase, CB-ENM showed that much better performance in characterising high modes, which represent local harmonic motions. In low modes, the norm of the overlap value from CB-ENM (0.33) is only slightly better than that from CV-ENM (0.31); in high modes, the norm of the overlap value from CB-ENM (0.70) is much higher than that from CV-ENM (0.59).

The norms of the overlap values of CB-NMA ($O_{\text{CBeni}}$) for elastic network interpolation are high at both low modes and high modes. Thus, CB-NMA is fit for both low mode and high mode analysis for elastic network interpolation.

Since CB-ENM utilises the strength of chemical bonds as network connection, some of covalent bonds are expected to perform a better description in local dynamic motions at atomic level. In order to evaluate this local dynamic motion, we compared local fluctuations of 12th mode from CV-ENM and CB-ENM for triglycerid lipase. Figure 1 depicts the difference of local dynamic motions from CV-ENM and CB-ENM, which are affected by chemical bonds. These results demonstrated that it is critical to consider chemical bonds in order to represent local motions in real situation.

**Figure 1** The difference between local dynamic motions from CV-NMA and CB-ENM. The grey backbone (long side chain) represents the initial conformation as an open form. The local atom positions (green) by CB-NMA are different from those by CV-NMA (pink) (see online version for colours)



### *3.2 Application of CB-NMA in identifying functional motions from HA of H5N1 highly pathogenic avian influenza virus*

Surface glycoprotein HA of influenza virus is involved in both attachment to host cell receptors and viral fusion. The host cell RBS are located around 220 loop of HA protein structure. In addition, HA and another surface protein NA are the primary targets for host immune systems. During the past several decades, a large amount of work has been conducted in mapping antigen binding sites in influenza proteins, especially HA. Five antigen binding sites, named A–E, have been mapped in influenza A viruses (Wilson and Cox, 1990) (Figure 2). Antigen binding site A is located around 140 loop, B around 190 loop, and D around 220 loop. Antigen binding site D and RBS are overlapped.

To better our understanding the HA functions of H5N1 highly pathogenic avian influenza viruses, we applied this newly developed all atoms CB-ENM in characterising functional motions of H5 tertiary structure (PDB ID: 2fk0). We simplified the representation of each residue by averaging the fluctuations of all atoms for each residue. We compared the global and local motions from CV-NMA and CB-NMA, and also analysed the biological functions for the residues with low fluctuations.

**Figure 2** Antigenic binding sites A–E of A/H3N2 influenza A virus (see online version for colours)

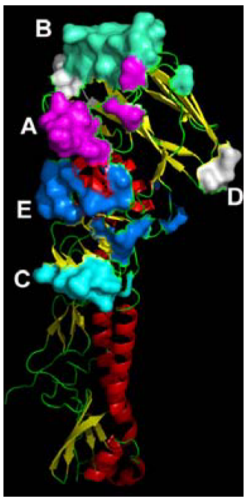



Figure 3 represents the comparison for the fluctuations of HA1 monomer in the lowest mode from CV-NMA and CB-NMA, which represents global motions. Our results showed that there is no significant difference between the results from these two models. However, our results demonstrated different magnitude of fluctuations at the peak areas between local motions (the 4th and 5th mode) from CV-NMA and CB-NMA (Figure 4).

**Figure 3** The global fluctuations of HA1 in H5N1 highly pathogenic avian influenza virus (the 1st mode). There are at least 11 local fluctuation peaks in this fluctuation spectrum. Most of these positions with peaks (low fluctuations) are correlated with reported Receptor Binding Sites (A–E) and Receptor Binding Sites (RBS). *x*-axis and *y*-axis are residue numbers and fluctuations, respectively (see online version for colours)

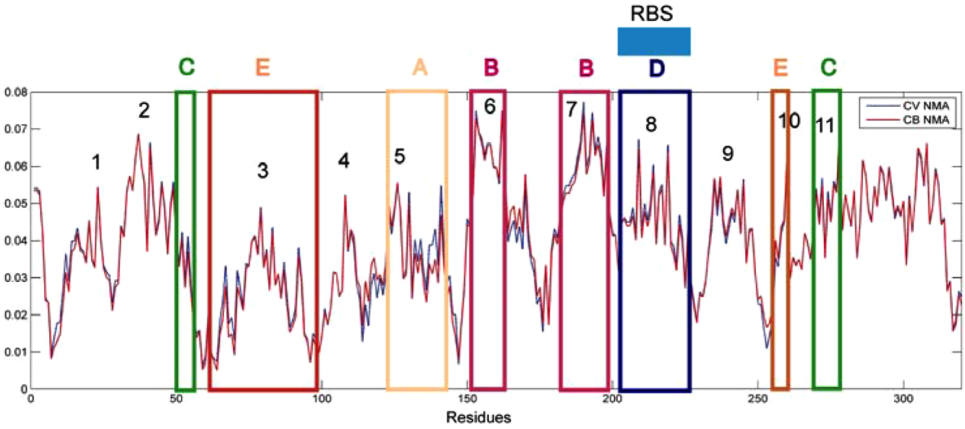
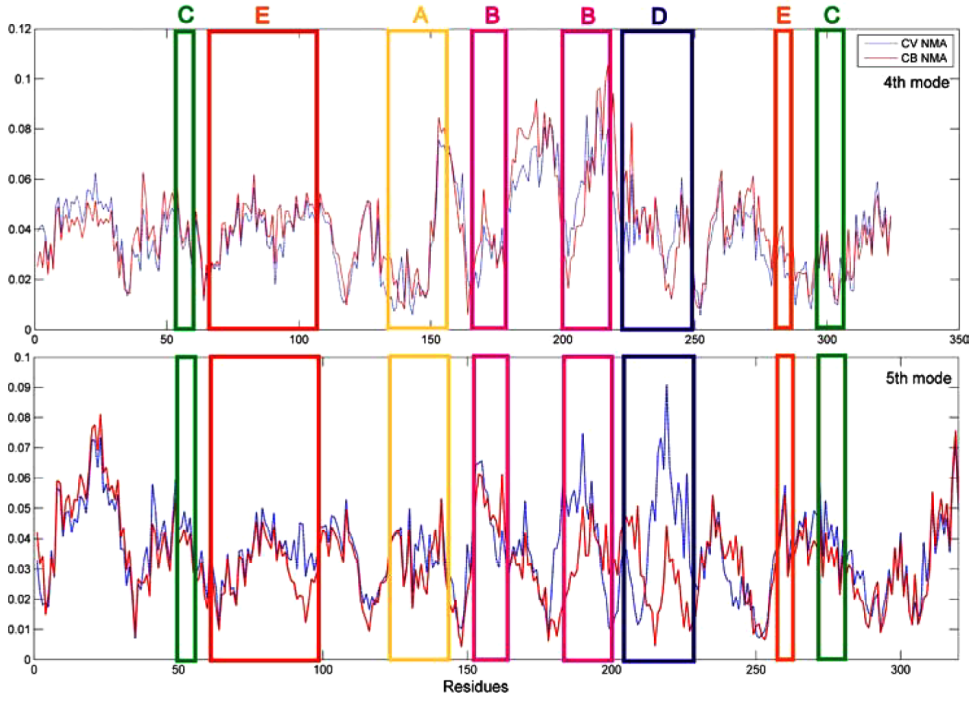


**Figure 4** The local fluctuations of HA1 from the 4th and 5th modes. Blue dash is CV-NMA and red is CB-NMA. *x*-axis and *y*-axis are residue numbers and fluctuations, respectively (see online version for colours)




As shown in Figure 3, our results demonstrated 11 local peaks among the fluctuation spectrum of HA1. All antigen binding sites A (peak 5), B (peak 6 and 7), C (peak 2 and 11), D (peak 8), E (peak 3 and 10) and RBS (peak 8) have low fluctuations. This demonstrated that high peaks (with local fluctuations) reflect the sites with important biological functions. It will be interesting to identify the biological functions for those residues with high peaks (with local fluctuations) but without any reported functions, for instance, the residues around position 240 (peak 9).

## 4 Discussion

In this study, an all-atom CB-ENM was developed. Since its stiffness ration was based on the strength of chemical bonds, CB-ENM reflects more realistic connection and movements. CB-ENM is expected to outperform conventional NMA for the atoms with special bonds, for instance, C=C and C=N in the aromatic amino acids. Through comparison with CV-ENM, our results showed that CB-ENM has an overall better performance than CV-ENM, especially in characterisation of local functional motions.

Our results showed that CB-ENM was more stable than conventional NMA. CB-ENM was able to construct the network using a distance cutoff of 3Å, with which instead the CV-ENM did not complete the analysis. Although the analysis with a 3Å distance cutoff was slightly worse than that with 5Å, no statistical significance was

identified between these two analyses. Thus, such a short distance cutoff could save the computational time and memory consumption in NMA.

Our application of CB-ENM in HA of H5N1 viruses showed that CB-ENM is useful in characterisation of local functional motions in protein structures. Thus, CB-ENM has a potential in identifying functional residues in protein macromolecules. So this newly developed method could be used not only to characterise a biological process for those sites with known biological functions but also to identify new sites with potential biological functions, which may facilitate the hypothesis generation and design of experiments for validation. Our future work will apply CB-ENM in exploring the details about HA-antibody interaction and HA-glycan interaction.

## References

Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O. and Bahar, I. (2001) 'Anisotropy of fluctuation dynamics of proteins with an elastic network model', *Biophys J.*, Vol. 80, pp.505–515.

Bahar, I., Atilgan, A.R. and Erman, B. (1997) 'Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential', *Fold Des.*, Vol. 2, pp.173–181.

Brooks, B. and Karplus, M. (1983) 'Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor', *Proc. Natl. Acad. Sci., USA*, Vol. 80, pp.6571–6575.

Flint, S.J., Enquist, L.W., Racaniello, V.R. and Skalka, A.M. (2004) *Principles of Virology: Molecular Biology, Pathogenesis, and Control of Animal Viruses*, ASM Press, Washington, DC.

Isin, B., Doruker, P. and Bahar, I. (2002) 'Functional motions of influenza virus hemagglutinin: a structure-based analytical approach', *Biophys J.*, Vol. 82, pp.569–581.

Jeong, J.I., Jang, Y. and Kim, M.K. (2006) 'A connection rule for alpha-carbon coarse-grained elastic network models using chemical bond information', *J. Mol. Graph Model*, Vol. 24, pp.296–306.

Kim, M.K., Li, W., Shapiro, B.A. and Chirikjian, G.S. (2003) 'A comparison between elastic network interpolation and MD simulation of 16S ribosomal RNA', *J. Biomol. Struct. Dyn.*, Vol. 21, pp.395–405.

McDonald, I.K. and Thornton, J.M. (1994) 'Satisfying hydrogen bonding potential in proteins', *Journal of Molecular Biology*, Vol. 238, pp.777–793.

Sanner, M.F. (1999) 'Python: a programming language for software integration and development', *J. Mol. Graph Model*, Vol. 17, pp.57–61.

Tama, F. and Sanejouand, Y.H. (2001) 'Conformational change of proteins arising from normal mode calculations', *Protein Eng.*, Vol. 14, pp.1–6.

Wilson, I.A. and Cox, N.J. (1990) 'Structural basis of immune recognition of influenza virus hemagglutinin', *Annu. Rev. Immunol.*, Vol. 8, pp.737–771.