

# ANM 2.0 site documentation

Download the documentation pages in PDF

- **Theory**
- **Manual**
- **Submission page**
- **Applicability**
- **Software used**
- **Input parameters**
- **Advanced Input Options**
- **Output text files**
- **Output coordinate files**
- **Using the graphical interface**
- **Known problems and possible solutions**
- **References**

***Theory***

ANM (Anisotropic network model) is a simple NMA tool for analysis of vibrational motions in molecular systems introduced in 2000 (1,2). It uses *Elastic Network (EN)* methodology and represents the system in the residue level. The macromolecule is thus represented as a network, or graph. In the original model each protein node is the C $\alpha$  atom of a residue and the overall potential is simply the sum of harmonic potentials between interacting nodes. The network includes all interactions within a cutoff distance which is the only predetermined parameter in the model. Information about the orientation of each interaction with respect to the global coordinates system is considered within the Force constant matrix and allows prediction of anisotropic motions. The model was successfully applied for exploring the relation between function and dynamics for many proteins (e.g. 3,4,5). The force constant of the system can be described by a *Hessian* matrix:

$$\mathbf{H} = \begin{bmatrix} H_{1,1} & H_{1,2} & \dots & H_{1,N} \\ H_{2,1} & H_{2,2} & \dots & H_{2,N} \\ \vdots & \vdots & & \vdots \\ H_{N,1} & H_{N,2} & \dots & H_{N,N} \end{bmatrix} \quad (1)$$

Each element  $H_{ij}$  is a 3 $\times$ 3 matrix which holds the anisotropic informations regarding the orientation of nodes  $i,j$ . Each such sub matrix (or the "super element" of the *Hessian*) is defined as:

$$H_{i,j} = \begin{bmatrix} \partial^2 V / \partial X_i \partial X_j & \partial^2 V / \partial X_i \partial Y_j & \partial^2 V / \partial X_i \partial Z_j \\ \partial^2 V / \partial Y_i \partial X_j & \partial^2 V / \partial Y_i \partial Y_j & \partial^2 V / \partial Y_i \partial Z_j \\ \partial^2 V / \partial Z_i \partial X_j & \partial^2 V / \partial Z_i \partial Y_j & \partial^2 V / \partial Z_i \partial Z_j \end{bmatrix} \quad (2)$$

namely the second partial derivatives of the potential  $V$ . These partial derivatives comprise a simple matrix of cosines. In practice the off-diagonal super elements are calculated by:

$$H_{i,j} = \begin{bmatrix} -\gamma \frac{(X_j - X_i)(X_j - X_i)}{s_{i,j}^2} & -\gamma \frac{(X_j - X_i)(Y_j - Y_i)}{s_{i,j}^2} & -\gamma \frac{(X_j - X_i)(Z_j - Z_i)}{s_{i,j}^2} \\ -\gamma \frac{(Y_j - Y_i)(X_j - X_i)}{s_{i,j}^2} & -\gamma \frac{(Y_j - Y_i)(Y_j - Y_i)}{s_{i,j}^2} & -\gamma \frac{(Y_j - Y_i)(Z_j - Z_i)}{s_{i,j}^2} \\ -\gamma \frac{(Z_j - Z_i)(X_j - X_i)}{s_{i,j}^2} & -\gamma \frac{(Z_j - Z_i)(Y_j - Y_i)}{s_{i,j}^2} & -\gamma \frac{(Z_j - Z_i)(Z_j - Z_i)}{s_{i,j}^2} \end{bmatrix} \quad (3)$$

where  $\gamma$  is an interaction constant which taken as constant in the original ANM.  $s_{i,j}$  is the instantaneous distance between nodes  $i$  and  $j$ .

The diagonal super elements are calculated by:

$$H_{i,i} = - \sum_{j=1, j \neq i}^N H_{i,j} \quad (4)$$

Note that the force constant matrix  $\mathbf{H}$  holds information regarding the orientation of the nodes, but not

regarding the type of the interaction (such is whether the interaction is covalent or non-covalent, hydrophobic or non-hydrophobic, etc.). In addition, the distance between the interacting nodes is not considered directly. To account for the distance between the interactions we can weight each interaction between nodes  $i, j$  by the distance:

$$\gamma_{i,j} \propto \frac{1}{s_{i,j}^p} \quad (5)$$

where  $p$  is an empirical parameter. The new off-diagonal elements of the *Hessian* matrix take the form of:

$$H_{i,j} = -\frac{1}{s_{i,j}^{p+2}} \begin{bmatrix} (X_j - X_i)(X_j - X_i) & (X_j - X_i)(Y_j - Y_i) & (X_j - X_i)(Z_j - Z_i) \\ (Y_j - Y_i)(X_j - X_i) & (Y_j - Y_i)(Y_j - Y_i) & (Y_j - Y_i)(Z_j - Z_i) \\ (Z_j - Z_i)(X_j - X_i) & (Z_j - Z_i)(Y_j - Y_i) & (Z_j - Z_i)(Z_j - Z_i) \end{bmatrix} \quad (6)$$

The inverse of the *Hessian* matrix is the covariance matrix of  $3N$  multi-variant Gaussian distribution which hold the desired information about the fluctuations. The *Hessian*, however is not invertible, as its rank is  $3N-6$  (6 variables responsible to a rigid body motion). To obtain a pseudo inverse, a solution to the eigenvalue problem is obtained:

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{U}^T \quad (7)$$

The pseudo inverse is composed of the  $3N-6$  eigenvectors and their respective non-zero eigen values according to:

$$\tilde{\mathbf{H}}^{-1} = \sum_{i=1}^{3N} \frac{1}{\lambda_i} \mathbf{U}_i \mathbf{U}_i^T \quad (8)$$

Where  $\lambda_i$  are the eigenvalues of H sorted by their size from small to large and  $\mathbf{U}_i$  the corresponding eigenvectors. The eigenvectors (the columns of the matrix U) describe the vibrational direction and the relative amplitude in the different modes. The mean square fluctuations of individual residues can be obtained by summing the fluctuations in the individual modes as follows:

$$\langle \Delta \mathbf{R}_i^2 \rangle = \langle \Delta X_i^2 \rangle + \langle \Delta Y_i^2 \rangle + \langle \Delta Z_i^2 \rangle = \frac{k_B T}{\gamma} [\tilde{\mathbf{H}}_{3i-2,3i-2}^{-1} + \tilde{\mathbf{H}}_{3i-1,3i-1}^{-1} + \tilde{\mathbf{H}}_{3i,3i}^{-1}] \quad (9)$$

The cross correlation between different residues can be also determined from the inverse Hessian matrix:

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = \langle \Delta X_i \Delta X_j \rangle + \langle \Delta Y_i \Delta Y_j \rangle + \langle \Delta Z_i \Delta Z_j \rangle = \frac{k_B T}{\gamma} [\tilde{\mathbf{H}}_{3i-2,3j-2}^{-1} + \tilde{\mathbf{H}}_{3i-1,3j-1}^{-1} + \tilde{\mathbf{H}}_{3i,3j}^{-1}] \quad (10)$$

For more advanced reading about the theory behind the model and its applications see references 11,12.

[Back to the top of the documentation page](#)

[Description of the web site](#)

## Applicability

The new version (ANM 2.0) of this web site is currently built to be used with proteins and nucleic acids. Other ligands which exist in the coordinates file may also be included. For proteins and nucleic acids the site suggests default nodes (Ca for amino acids, P, C4', C2 for nucleic acids), while for non standard ligands the user need to manually chose nodes and interaction cutoffs if she wishes to include those in the model. For full functionality, the size of the protein should be up to 1500 residues. For such structures the calculation of 20 dominant modes is performed within seconds. The calculation of the theoretical B-factors based on all the modes is done within 2 minutes. The analysis pages could appear in a delay, depending on the communication speed and the memory of the client computer (downloading large result files to the user machine), but will not appear within more than a minute or two for such moderate-size proteins in any case. For larger proteins some problems with the graphics are expected. The calculation of the b-factors might take longer. The dominant modes calculation, however, will be performed and text files with the results will still be available within few minutes. The site was tested with proteins up to 2000 residues. For structures with more nodes expect severe functional problems. Note that the Jmol applet on the main interface page could be slow for large structures, as the coordinates file (which should include a frame for each normal mode) becomes huge and causing memory issues. The graphic pages that appears under the [create PDB files](#) option should behave better because only information about individual modes is saved in the memory at each time.

The site was successfully tested on WindowsXP and Windows7 using the following browsers:

- Firefox 32.0
- Internet Explorer 11.0
- Opera 8.53
- Chrome 37.0

On Linux, using the following browsers:

- Firefox 30.0 on Ubuntu 13.10
- Firefox 24.7 on CentOS 6.5
- Chrome 33.0 on Ubuntu 13.10

On MacOSX, using the following browsers:

- Mozilla Firefox

The site is not fully functional using IE and Safari on Macintosh.

## ***The eigensolvers***

This site uses [BLZPACK](#)(6) and [Matlab](#)(9) for calculating a subset of eigenmodes. The default value is 20 modes but this number can be changed in the input page. If the number of requested modes is different than 20, the user should choose matlab as the eigensolver. The site calculates and report the slowest modes. BLZPACK is a Fortran 77 implementation of the block Lanczos algorithm for the solution of the eigenvalue problems. It was originally intended to solve large, sparse, generalized problems which makes it very appropriate for the use here. In Matlab we use the eigs function which is the Matlab interface to the ARPACK package. The server calculates a subset of the eigenpairs and not all of them. This is sufficient since we are usually interested in a small set of eigenpairs corresponding to the smallest eigenvalues ("slow modes"). The eigen decomposers use extremely efficient algorithms (6). For molecules with 300 residues (900×900 ANM Hessian matrix); the solution is obtained within 2 seconds and for structures with thousands of nodes it is typically obtained within few minutes.

## ***Jmol - the molecular graphics tool***

[Jmol](#) (7) is a free code Java based molecular graphics tool. It was chosen to provide the graphics for this site as the code is actively and rapidly developed, and since it was originally designed to support visualization of vibrations. This support includes the option for animation of all vibrational modes in one session with only button click required to switch between the modes. Directional vectors can be presented for easy visualization of the direction of motion at each mode. Detailed documentation and description of Jmol capabilities are available on the Jmol site. In general, many options can be

controlled by the mouse from the pop up menu (right mouse click on the applet). Other options are available by command line scripting language which is highly compatible with the [RasMol](#) scripting language (open the "console" option from the pop up menu to execute commands). For several options which are directly related to the vibrations and for some important visualization features we built a control panel for the user as will be described.

We highly acknowledge the work done by the Jmol team that allow us (and other) to present sophisticated data in a convenient and high quality graphical environment.

Several security problems with Java applets gradually limit the accessibility of the jmol sessions in different platforms and browsers. We are aware to these issues and collected [here](#) solutions to some of the problems.

### ***Other software applied***

[Matlab](#) (9) is also used for matrix manipulations and a fast calculation of theoretical B-factors based on all the modes using the eig function. Perl modules are applied for web programming and preparations of graphs and figures. C code is used for parsing of structures and calculation of the Hessian matrix. The Matlab, Perl and the C codes are available upon request.

[Back to the top of the doumentation page](#)

### ***Input parameters***

The minimal required input is very simple. The user should only specify an input coordinates' file in a [PDB format](#). There are 2 ways to specify the file:

- Specification of a 4 letter code PDB ID. The user can choose by a checkbox if to use the original PDB coordinates (default) or the putative [biological unit](#) supplied by the PDB. The latter should reflect better the biological multimeric state of the protein.
- Uploading of a coordinates file from the user computer to the server.

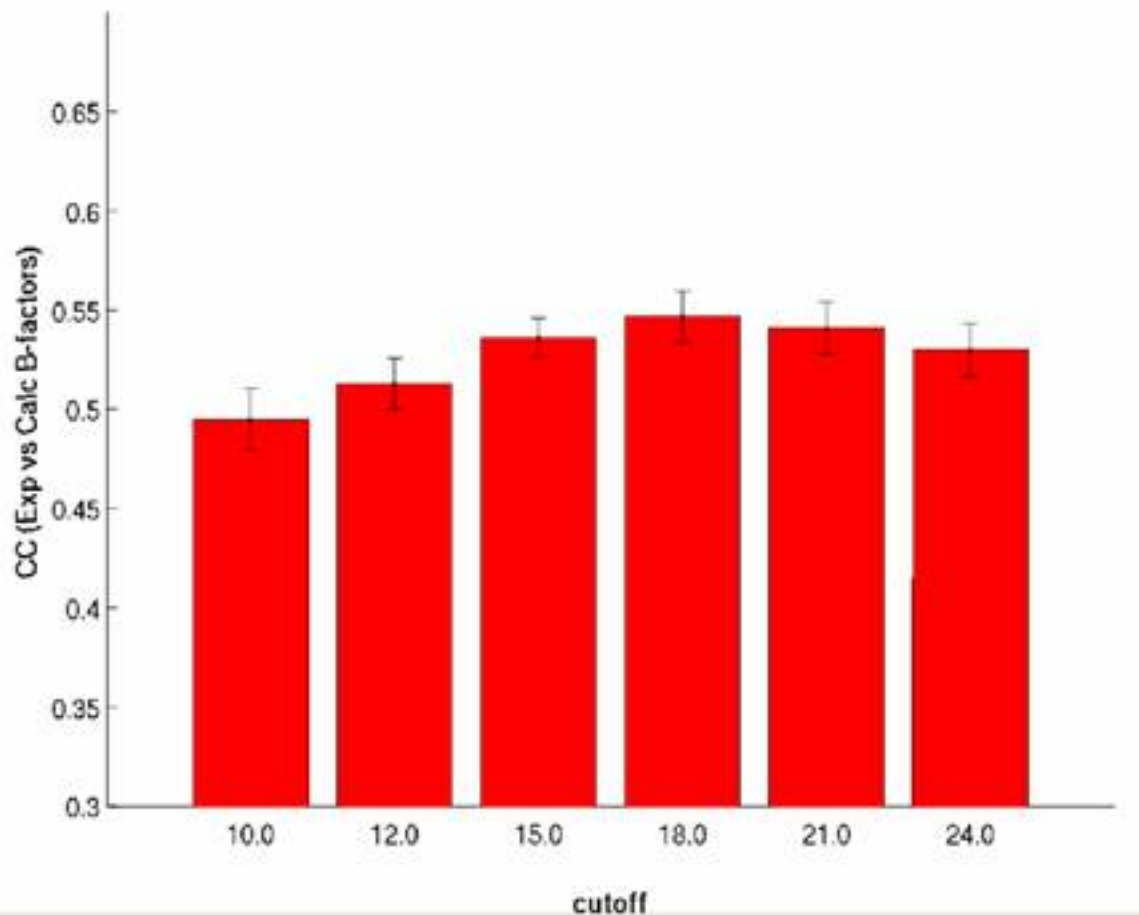


The other input parameters are usually not required for a successful run. Still, they might be important and relevant and influence the quality of the results:

The user should determine if she wishes to include all polypeptides chains in the calculation (default) or only individual chain. In the latter, a chain ID (usually a Latin letter A,B etc.) should be specified on the appropriate text field in the input page. An input chain symbols such as "\*", " ", "-", "\_", "0" will cause inclusion of all polypeptide chains in the calculation.

In addition, the user can define which models will be included in the calculation. Models are often (but not always) used to include alternative conformations in the file. The two most frequent situations in which the coordinates file include different models are NMR files and files with the biological unit information. In the first case the user should choose which of the models in the file (usually marked as integer numbers) should serve as the equilibrium state for the calculation. In the second case the user should probably specify "all" in the text field for model as all the coordinates in the file (even those found under different models) are part of one biological unit.

In then basic input page the user has control over two parameters in the model. The first is the cutoff which defines the maximum interaction range between C $\alpha$  atoms. Fig. 1 shows the performance of the model (in terms of correlation with the experimental fluctuations) for a range of cutoffs. It can be clearly seen that on average for globular proteins the best results are obtained for cutoffs of 15-21Å. For individual proteins, however, for example extended proteins or membrane proteins, smaller cutoffs maybe preferred. It should be also noted that calculation with larger cutoffs are slower (the Hessian matrix is less sparse).



*Fig 1. The agreement between the theoretical fluctuations in ANM and the experimental values*

The second parameter is the distance weighting factor (Eq. 5,6). Large scale statistical analysis indicates the best value for this parameter is 2.5. Using this value the correlation between theoretical and experimental b-factors are better by 4% comparing to the original ANM from about 55% to about 59%. When running with  $p=0$  (default), the original non-weighted ANM is executed.

### **Advanced Input Options**

Version 2.0 of the ANM website allows now more sophisticated control over the run parameters. These optional parameters should be provided from the "Proceed with advanced input options button" in the bottom of the input page. The user can control the identity of the atoms which will serve as nodes in the network. For proteins, the C $\alpha$  atoms are the default nodes, but the user can change to other backbone atoms if needed. For nucleic acids, the default setting is that each nucleotide will be represented by 3 nodes. By default each nucleotide is represented in the network by three nodes, positioned at the P atom of the phosphate group, C4' in the base and C2 in the sugar ring. This model has been shown (Yang et al., 2006) to have superior predictive abilities, based on the comparison with experimental B-factors using 64 oligonucleotide/protein complexes (13) and on case studies such as the application of ANM to the ribosome (14). The server provides the option of adopting different number and types of user-defined atoms for representing the nucleotide nodes. All backbone heavy atoms of the protein and all the nucleotide atom types are listed and can be selected as nodes. Small ligands can also be explicitly included in the ANM analysis. Ligands, if present in the PDB coordinates file, will be automatically parsed and listed and their atoms will appear in separate a list. Each atom from this list can also be selected by the user as a network node.

To support the above changes, the server applies now a new strategy that enables flexibility in the definition of network nodes and cutoff distances of interaction. The cutoff,  $rc$  depends on the identity of potentially interacting atoms  $i, j$ . For each atom type  $i$ , an associated distance range of interaction,  $t_i$ , is assigned. For each pair of atom types  $i, j$  the cutoff distance  $rc$  is defined as  $t_i+t_j$ . This definition implies that the number of parameters in the system is bound by the number of different atom types. In the simplest case where we want a unified threshold distance  $rc$  for all atom types, as in the traditional ANM, we simply assign  $t_i = rc/2$  for all atom types we would like to include. The server suggests  $rc=15$  as default setting.

In the advanced input page of version 2.0 the user has an additional option of importing external normal modes, e.g. submitting principal modes that have been calculated by another method, and take advantage of the powerful GUI of the server for visualizing the motions and the cross-correlations

between residues. The dimensions of these uploaded vectors must match the number of system nodes defined by the user. The user should upload 3 files, for the X,Y and Z components of the modes according to [this format](#)

## ***Output text files***

A series of text files which include the outcome of the calculation plus additional analyses are available for download:

- **Slow eigenvectors (oanm.slwevs)** : List the 20 eigen vectors which correspond to smallest 20 eigen values.
- **Slow Eigenvalues (oanm.eigvals)** : Lists the 36 smallest eigenvalues. The 6 smallest eigenvalues should be roughly zero. Starting from the seventh, the eigenvalues are related to the vibrational frequencies (proportional to the square root of the eigenvalue) and to the contribution to the magnitude of fluctuations (proportional to 1/eigenvalue).
- **X components of slow eigenvectors (oanm.slwX)** : List the X axis component of the 20 eigenvectors (a column to each vector) which correspond to smallest 20 non-zero eigenvalues.
- **Y components of slow eigenvectors (oanm.slwY)** : List the Y axis component of the 20 eigenvectors which correspond to smallest 20 non-zero eigenvalues.
- **Z components of slow eigenvectors (oanm.slwZ)** : List the Z axis component of the 20 eigenvectors which correspond to smallest 20 non-zero eigenvalues.
- **Normalized slow modes. Fluctuations of residues in each individual mode (oanm.slwmodes)** : The normalized self fluctuations (the norm of each vector is 1) for each residues in each mode (a column to each mode). The first column is the residue index.
- **Hessian matrix in a coordinates (sparse) format (oanm.hes)** : The Hessian calculated according to Eq. 3 or 6. The Hessian is brought in a format *i, j, value*. In the original matrix H we

have  $H[i][j]=value$ ,  $H[j][i]=value$ . Zero entries in Hessian are not listed in this file.

- **Theoretical and experimental b-factors (oanm.bfactors)** : Includes the theoretical and experimental B-factor (temperature factor). The first column is the residue index, the second is ANM calculated theoretical B-factors and the third is the x-ray crystallographic b-factors taken from the PDB file.
- **Correlation between theoretical and experimental b-factors (oanm.corr)** : Pearson's correlation coefficient between theoretical and experimental b-factors.
- **Estimated value of spring constant (oanm.gamma)** : Based on the difference between the experimental and calculated b-factors, it is feasible to calculate the structure-specific spring constant (gamma).
- **Gamess file (protein.out)** : See description in the section below.
- **XYZ file (protein.xyz)** : See description in the section below.

### ***Output coordinate files***

- **Coordinate files in PDB format for individual mode fluctuations** : PDB files for individual modes can be constructed. The files include set of models that describe the fluctuations in this mode in a linear sequence. These files can be downloaded and read by any application capable of showing show/animate PDB structures. The user controls the amplitude of the fluctuations and the number of frames in the animation (related to the smoothness of the motion in the animation). Note that the animation should be performed in a palindromic fashion and not in a loop fashion (the first frame and the last frame are in two extremes of the fluctuations).
- **Coordinate files in PDB format with anisotropic temperature factors**: The model predicts also the magnitude of the fluctuations in each individual axis and the covariance between the fluctuations in the different directions. These are directly related to the anisotropic temperature

factors sometimes reported in the protein data bank. The values of the anisotropic factors can be based purely on the theoretical results of ANM, or they can be scaled according to the experimental b-factors. This scaling improves significantly the prediction of the mean square fluctuations in the X, Y, Z directions and slightly the correlation between the motions in the different directions.

- **Gamess file** : This file includes the input coordinates and the most significant eigenvectors (20 by default). The format of these sections is according to [Gamess](#) (10) file format. This file is also readable by Jmol.
- **XYZ file** : This file includes the input coordinates and the most significant eigenvectors (20 by default). The format is: atom,x,y,z,x<sub>vib</sub>,y<sub>vib</sub>,z<sub>vib</sub> and the structure is therefore repeated according to the number of normal modes calculated.

### ***Using the main interface***

- **Modes**: Visualization of the slow modes calculated. The fluctuations of the selected mode appear and the molecule is colored according to the self-fluctuations. Red colors correspond to large fluctuations and blue colors to small fluctuations. By default, the vibrations in the first mode are animated and the residues are colored accordingly.
- **Frequencies of vibrations** : The default state of the analysis page presents the vibration frequencies of the different modes in the same proportions as the real frequencies. The user, however, can change the frequency for each mode using the pulldown menu for the purpose of visualization or analysis. In this case the proportions between the vibrational frequencies of the different modes are not realistic any more (the change is valid only to the mode where it was applied). The visualization frequencies (in Hertz units) are, of course, the vibrational frequencies as appear on the screen and not the real vibrational frequencies of the molecule, which are orders of magnitudes larger.

- **Amplitude of vibrations** : The amplitude of the vibrations can be changed by the user to emphasize the motion or to observe minor fluctuations that can not be distinguished otherwise. It is important to note the model does not always accurately represents the real size of the fluctuations but only the relative amplitude between the residues.
- **Vectors** : Vectors represent the direction of motion of each residue in each vibrational mode can be shown or hidden by checking on the vectors check box. The user can change the color of the vectors, their width and their length.
- **Display** : The size of individual atoms (space fill) and the width of the bonds (wire frame) can be controlled by the pulldown menus in the graphical user interface, in addition to the available built in options in Jmol. If for some reason the desired action is not performed, try to click on other value in the same pulldown menu and then press again on the initial one. Another possibility is that not all atoms are selected. From the Jmol menu (right click on the applet area) choose "select all" and then try again to perform the action.
- **Labels** : Residues can be labeled by their type and their number in the coordinates part of the PDB input file. Individual atoms can be labeled, as well as all atoms. The clear option in the menu remove all the labels. Note that the regular labeling options of Jmol will not be functional with this file format.
- **Colors** : The protein can be colored by the polypeptide chains, by the experimental temperature (b) factors and by the fluctuations in each individual mode. The coloring scale is red-white-blue, where red indicates larger fluctuations. The coloring mode chosen from this pulldown menu is disconnected of the display method or the vibrational mode, so potentially we can have the residues vibrate according to one mode but colored according to a different mode.
- **ANM model** : Shows the network of the interactions for 10.0Å cutoff.
- **Chain connectivity** : Color each polypeptide chain in different color and shows only the trace of the chains with no other non bonded interactions between sequentially distal residues. Because of the poor annotation in this file format, the Jmol built-in menu does not support equivalent

options. The chain connectivity option may take some time to be executed for large proteins.

- **Snapshot** : Since there is no direct way to save images from the Jmol applet, we add now (ANM 2.0) an option to easily capture the Jmol screen and save the image in JPG format.

### ***Additional visual analysis options***

- **Self fluctuations and b factors as a function of residue index** : The graph presents the fluctuations of individual residues according to experimental b-factors (or mean square fluctuations in NMR structures), according to the normalized values (in units of standard deviations from the mean), according to the calculated mean squared fluctuations or according to the fluctuations in each individual mode. For each polypeptide chain a separate graph is given. A single parameter or any combination of the above parameters can be chosen for the display. The graphs (in .png format) can be downloaded. The residue numbers on the X-axis are taken from the coordinates part of the PDB file.
- **Correlation analysis by two dimensional matrices** : The web site offers elaborated options to explore the correlations in fluctuations between residues. The user can choose to explore the correlations between residue pairs in each **individual** mode. In addition, a **range** of modes (within the first 20) can be given and then the correlation according to the sum of all the modes in this range is obtained. A correlation based on **all** modes can also be obtained. Initially, when linking to the covariance page from the main interface page, a correlation matrix based on the first mode is shown. Using the options in the lower-left frame the user can choose to analyze a different correlation matrix as described above. The structure in the upper-left Jmol frame vibrates in the same mode as that of the correlation matrix . If a range of modes is given, the structure vibrates in the lowest mode in this range.
- The color scheme of correlation matrix in the right frame is such that correlated residues are colored by red and anti-correlated residues are colored by blue. Weak correlations appear in light colors. Each cell in this matrix is a hot link to a 10×10 matrix which zoom in to this region in the matrix. For big proteins this allows an analysis of each pair, which is not feasible in the



resolution of the large matrix. Moving the mouse over of the cells of the small 10×10 matrix, boxes appear which give information about the correlation value as well as the distance between the residues (the C $\alpha$  atoms).&nbsp; This matrix is also "hot". Clicking on each cell will highlight the residues on top of the structure in the Jmol frame. In ANM 2.0 we added an option to navigate directly from this small matrix to the adjacent four neighbouring sub-matrices using arrows, without going through the large matrix. The colors of the residues in the Jmol frame are the same as those on the correlation matrix and indicate the degree of correlation between the residues. The correlation matrices are standard figures in a .png format and can be downloaded from this page.

- **Inter residue distance fluctuations and deformation energy** : The web site also offers options to explore the changes in distance between residues. The user can explore the square-distance fluctuations between residue pairs in each **individual** mode. Initially, when linking to the "distance fluctuations and deformation energy page" from the main interface page, a square-distance fluctuation matrix and a deformation energy graph based on the first mode are shown. Using the options in the lower-left frame the user can choose to analyze a different mode. The structure in the upper-left Jmol frame vibrates in the same mode as that of the matrix. The user can choose to analyse the vectorial difference, which considers also the orientation of the distance vector. The deformation energy is proportional to the sum of the square fluctuations with all interacting residues
- The color scheme of matrix is such that large inter-residue fluctuations appear as blue and small fluctuations as red. Each cell in this matrix is a hot link to a 10×10 matrix which zoom into a sub-region in the matrix. For big proteins this allows an analysis of each pair, which is not feasible in the resolution of the large matrix. Moving the mouse over of the cells of the small 10×10 matrix, boxes appear which give information about the fluctuation value as well as the equilibrium distances (between C $\alpha$  atoms). This matrix is also "hot". Clicking on each cell highlights the residues on top of the structure in the Jmol frame. In ANM 2.0 we added an option to navigate directly from this small matrix to the adjacent four neighbouring sub-matrices using arrows, without going through the large matrix. The colors of the residues in the Jmol frame are the same as those on the square distance fluctuations matrix. The matrices are figures in a .png

format and can be downloaded from this page.

A graph of deformation energy for each residue shows the potential energy of the residue considering all the interactions ("springs") in which it takes part. For each residue all the entries in its row/column matrix for which there is an interaction (based on the user's cutoff) are summed up. The number below the graph shows the overall energy of the system. According to the equipartition theorem the overall internal potential energy of the system in each independent mode is constant and equals to  $k_B T/2$  (around 0.3 Kcal/mol at 300K).

- **Anisotropic displacement parameters** : ANM can calculate anisotropic displacement parameters (ADP) for the atoms in the network. Since the vast majority of experimental crystal structures are not resolved with ADPs, the possibility to rapidly predict such parameters is an important added value of the ANM model. From the link "Anisotropic factors" in the lower panel, the user can obtain PDB files with ADPs to its input structure, static image of the ADPs, created using the Rastep program, presented as ellipsoids around the equilibrium coordinates, and interactive Jmol session in which the nodes are plotted as ellipsoids, whose shape and color indicate the shape of the 3D Gaussian probability distribution for the location of the atoms, but the user can control the size, shape, color and format of the ellipsoids using the Jmol scripting language. For each option (test file, image, Jmol session) the user can normalize the size of the ADPs/ellipsoids according to the values of the sum of the isotropic B-factors in the input file. ("purely computational"). Alternatively, the ADPs of each individual atom can be normalized by the B-factor of this atom, such that every atom has its own normalization constant ("refined by experimental b-factors")
- **Eigenvalues** : A figure with the distribution of the 20 dominant eigenvalues is shown. This allows for a quick visual inspection for the relative contribution of each mode and possible degeneracy between the modes.

### ***Known problems and possible solutions***

- A problem with the functionality of the site on Mac computers still exists. Some of the options,

involving interactions with the Jmol applet using a graphical user interface, are not working on most browsers. Other options are still functional and the text files holding the computational results are available in any case.

- Delays in the calculation of B-factors and correlation matrix based on all modes. Since these options essentially requires more heavy computations the results might appear in delay comparing to the other data.
- Options at the main graphical user interface are sometimes not responding well. If an action chosen from the pulldown windows is not performed, try to click on other value in the same pulldown menu and then press again on the initial one. It is also possible that not all atoms are selected. From the Jmol menu (right click on the applet area) choose "select all" and then try again to perform the action.
- If the motion is very slow and the applet does not responds well, most likely the molecule is very large. You may try to use the site with a computer with larger memory.

### ***Selected reading and references***

(1) *Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to alpha-amylase inhibitor.* Doruker, P, Atilgan, AR & Bahar, I. Proteins 40, 512-524, (2000).

(2) *Anisotropy of fluctuation dynamics of proteins with an elastic network model.* Atilgan, AR, Durrell, SR, Jernigan, RL, Demirel, MC, Keskin, O. & Bahar, I. Biophys. J. 80, 505-515, (2001).

(3) *Computational prediction of allosteric structural changes by a simple mechanical model: application to hemoglobin T to R transition.* Chunyan, X, Tobi, D & Bahar, I. J. Mol. Biol. 333, 153-168 (2003).

(4) *Global ribosome motions revealed with elastic network model.* Wang, Y, Rader, AJ, Bahar, I & Jernigan RL. Structure. 147, 302-314 (2004).

(5) *Relating molecular flexibility to function: A case study of tubulin.* Keskin, O, Durell, SR, Bahar, I, Jernigan, RL, Covell, DG. *Biophys J.* 83, 663-680 (2002).

(6) <http://crd.lbl.gov/~osni/marques.html#BLZPACK>

(7) *oGNM: Online Computation of Structural Dynamics Using the Gaussian Network Model.*

Yang, LW, Rader, AJ, Liu, X, Jursa, CJ, Ching, SC, Karimi, H & Bahar, I. *Nuc. A. Res.* In press (2006).

(8) <http://jmol.sourceforge.net/>

(9) The [MathWorks](#), Inc

(10) *General Atomic and Molecular Electronic Structure System.* Schmidt, MW, Baldrige, KK, Boatz, JA, Elbert, ST, Gordon, MS, Jensen, JH, Oseki, SK, Matsunaga, N, Nguyen, KA, Su SJ, Windus, TL, Dupuis, M & Montgomery, JA. *J.Comput.Chem.* 14, 1347-1363 (1993).

(11) *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems* (2006) Mathematical Biology Series, CRC Press.

(12) *Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies.* Chennubhotla, C, Rader, AJ., Yang, LW, and Bahar, I., *Phys. Biol.* 2, S173-S180 (2005).

(13) *oGNM: Online computation of structural dynamics using the Gaussian Network Model* Yang, LW, Rader, AJ., Liu Xiong, Christopher Jon Jursa, Chen, SC, Karimi, HA and Bahar, I., *Nucleic Acids Res.* 34, W24-W31 (2006).

(14) *Global ribosome motions revealed with elastic network model assemblies.* Wang, Y, Rader, AJ,

Bahar, I, Jernigan, RL., J Struct Biol. 147, 302-314 (2004).