ABSTRACT The movement of the myosin motor along an actin filament involves a directed conformational change within the cross-bridge formed between the protein and the filament. Despite the structural data that has been obtained on this system, little is known of the mechanics of this conformational change. We have used existing crystallographic structures of three conformations of the myosin head, containing the motor domain and the lever arm, for structural comparisons and mechanical studies with a coarse-grained elastic network model. The results enable us to define structurally conserved domains within the protein and to better understand myosin flexibility. Notably they point to the role of the light chains in stiffening the lever arm and to changes in flexibility as a consequence of nucleotide binding. Proteins 2004;54:384-393.

© 2003 Wiley-Liss, Inc.

Key words: motor proteins; Gaussian network model; structural blocks; B-factors

INTRODUCTION

Myosin is an enzyme that converts the chemical energy resulting from the hydrolysis of ATP into directed mechanical movement along an actin filament. The actomyosin system is involved in numerous cell processes including vesicle trafficking, determinant partitioning, cell motility, neurosensory function, and muscle contraction. Although considerable crystallographic data have been gathered on this system, many questions concerning the molecular mechanisms underlying myosin mobility remain unanswered.

Myosin II, so-called conventional myosin, forms filaments and constitutes large assemblies of noncooperative motors within muscular tissues. It is an important member of a diverse family of myosin motor proteins. Various mechanisms have been proposed for myosin movement. The majority of biophysicists explain muscle contraction by the movement of the myosin lever arm, but other evidence has pointed to a biased Brownian ratchet mechanism and to the possibility of multiple myosin steps per ATP-driven cycle. It may however be possible to reconcile these apparently conflicting viewpoints. A part of the mechanism proposed by Houdusse et al. based on insights from X-ray structures, cryo-electron microscopy, and kinetic studies is presented in Figure 1.

The strong binding of myosin to actin (rigor state) weakens with ATP binding. This conformation is termed the near rigor state. The detached state, where myosin releases the actin filament may prevent a reverse power stroke and increase the lifetime of the prehydrolysis state. After hydrolysis of ATP in the myosin motor, phosphate binding stabilizes the so-called transition state until actin binds. This is followed by force generation and ADP release returning the system to its rigor state.

In the present article, we use theoretical methods to study myosin II, in an attempt to better understand the mechanics of its conformational changes. Because the myosin head is a large system (1147 amino acids, 130 kDa) and, moreover, undergoes large conformational changes, it is not easy to use conventional all-atom molecular mechanics or dynamics methods. We have thus chosen to study the problem with an anisotropic network model and also via a rigid block decomposition method. Both of these methods are coarse-grained and only use a single point, Cα, to represent each amino acid residue. The anisotropic network model provides data on the large-scale collective modes of vibration by converting the protein structure into a set of coupled springs between neighboring residues and carrying out a normal mode style analysis. It has been shown to provide data in excellent agreement with more refined all-atom approaches and with crystallographic temperature factors. The rigid block decomposition method is based on a comparison of inter-Cα distances between two structures of the same protein and the identification of blocks based on virtually constant inter-residue distances.

Together, these methods enable us to identify the rigid and flexible domains within the myosin structure and highlight the respective roles of the light chains and of nucleotide binding.
MATERIALS AND METHODS

1. Structures

Three structures of different states of the scallop myosin S1 head form the basis for our calculations. These structures have been determined by X-ray crystallography and are available in the Protein Data Bank (PDB)\textsuperscript{18} with the codes 1DFK, 1DFL, and 1B7T. 1DFK corresponds to scallop myosin S1 without an adenosine nucleotide and is believed to be the near rigor state (NR). 1DFL is the myosin head structure in the presence of ADP.VO\textsubscript{4}, corresponding to the transition state (TS). Finally, 1B7T is scallop myosin S1 complexed with ADP, which has been interpreted as a detached ATP state (DS). Hereafter, we refer to these three structures using the functional codes NR, TS, and DS, respectively (see Fig. 1).

Each of these structures is composed of three polypeptide chains: the main chain (835 residues), divided into the N-terminal motor domain and the \(\alpha\)-helical C-terminal lever arm; the essential light chain (ELC, 156 residues); and the regulatory light chain (RLC, 156 residues). Two of the three available structures of myosin, NR and TS, are rather poorly resolved (4.20 Å) and some residues are missing, limiting the head domain to 1019 and 1059 \(\alpha\)-carbons, respectively. DS, is better resolved (2.50 Å) and the structure is comprised of 1057 residues. These data are summarized in Table I.

![Fig. 1. Schematic view of the myosin cycle showing the DS, NR, and TS states (adapted from Houdusse and Sweeney\textsuperscript{10}). Note that this cycle corresponds to the thermal ratchet interpretation of myosin action. The main steps of the cycle are as follows: in the absence of nucleotide, myosin binds tightly to actin. This is termed the rigor state. The binding of ATP induces weakening of the binding via the so-called near-rigor state (NR). The near-rigor state is in equilibrium with a completely detached state (DS). As a result of ATP hydrolysis, myosin moves one step along the actin filament, leading to the so-called transition state (TS). Phosphate release puts myosin in the weakly binding NR state and the release of ADP returns it to the rigor state. For more details see Ref. 10.](image-url)

<table>
<thead>
<tr>
<th>Code</th>
<th>NR</th>
<th>TS</th>
<th>DS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supposed position in the cycle\textsuperscript{10} (Fig. 1)</td>
<td>Near rigor state</td>
<td>Transition state</td>
<td>Detached state</td>
</tr>
<tr>
<td>PDB entry</td>
<td>1DFK</td>
<td>1DFL</td>
<td>1B7T</td>
</tr>
<tr>
<td>Resolution</td>
<td>4.20 Å</td>
<td>4.20 Å</td>
<td>2.50 Å</td>
</tr>
<tr>
<td>Nucleotide</td>
<td>None</td>
<td>MgADP-VO\textsubscript{4}</td>
<td>MgADP</td>
</tr>
<tr>
<td>Experimental B-factors</td>
<td>Not available</td>
<td>Not available</td>
<td>Available</td>
</tr>
<tr>
<td>Number of residues</td>
<td>1019</td>
<td>1059</td>
<td>1057</td>
</tr>
<tr>
<td>Number of atoms</td>
<td>5031</td>
<td>5230</td>
<td>8249</td>
</tr>
</tbody>
</table>

2. Anisotropic Network Model

This coarse-grained elastic model begins by reducing a protein to a set of \(\alpha\)-carbons and takes no account of amino acid side chains or other peptide backbone atoms. The protein structure is taken into account by creating springs between spatially neighboring residues (whether or not these residues are sequential within the peptide chain). The vibrations of such a system can be analyzed either by assuming that all fluctuations are isotropic, as in the so-called Gaussian Network Model (GNM),\textsuperscript{19–23} or by taking into account anisotropy via the directions of movement of each residue, in the so-called Anisotropic Network Model (ANM).\textsuperscript{12} We have used the latter approach because...
it yields the structural changes associated with collective vibrational modes, rather than simply their overall magnitudes. The potential energy $V$ of a structure with $N$ residues is expressed within the ANM as a Gaussian form,

$$V = \frac{\gamma}{2} \Delta \mathbf{R}^T \mathbf{H} \Delta \mathbf{R}$$

where $\gamma$ is the spring constant, $\Delta \mathbf{R}$ is a $3N$-dimensional vector of the fluctuations $\Delta \mathbf{R}_i$ in the position vectors $\mathbf{R}_i$ of all sites ($1 \leq i \leq N$), $\Delta \mathbf{R}^T$ is its transpose, and $\mathbf{H}$ the Hessian matrix composed of the second derivatives of the potential energy. Thus, $V$ can also be written as follows:

$$V = \frac{\gamma}{2} \sum_i \sum_j h(r_{ij} - R_{ij})(R_{ij} - R_{ij})^2$$

The summations are performed over all interaction sites. $h(x)$ is the Heaviside step function [$h(x) = 1$, if $x \geq 0$, and zero otherwise], $R_{ij}$ is the distance between sites $i$ and $j$ in the protein structure, $R_{ij}$ is the same distance after fluctuation, and $r_c$ is the cutoff distance defining the interacting residue pairs for which Gaussian springs are created. $\mathbf{H}$ is expressed as a function of $N^2$ submatrices $H_{ij}$ of the following form:

$$H_{ij} = \frac{1}{2} \begin{bmatrix}
\frac{\partial^2 V}{\partial X_i \partial Y_j} & \frac{\partial^2 V}{\partial X_i \partial Z_j} & \frac{\partial^2 V}{\partial X_i \partial Z_j} \\
\frac{\partial^2 V}{\partial Y_i \partial X_j} & \frac{\partial^2 V}{\partial Y_i \partial Y_j} & \frac{\partial^2 V}{\partial Y_i \partial Z_j} \\
\frac{\partial^2 V}{\partial Z_i \partial X_j} & \frac{\partial^2 V}{\partial Z_i \partial Y_j} & \frac{\partial^2 V}{\partial Z_i \partial Z_j}
\end{bmatrix}$$

with $X_i$, $Y_i$, and $Z_i$ being the components of $\mathbf{R}_i$. Note that,

$$\left( \frac{\partial^2 V}{\partial X_i \partial Y_j} \right)_0 = -\left( \frac{\partial^2 V}{\partial X_i \partial Y_j} \right)_0 = -\frac{\gamma}{r_{ij}^2} (X_i^0 - X_j^0)(Y_i^0 - Y_j^0)$$

and

$$\left( \frac{\partial^2 V}{\partial X_i \partial Z_j} \right)_0 = \gamma \sum_j \frac{(X_i^0 - X^0_j)(Y_i^0 - Y^0_j)}{R_{ij}^2}$$

The correlations between the fluctuations at sites $i$ and $j$ are as follows:

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = \frac{1}{Z} \langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle \exp[-V/kT]d\{\Delta \mathbf{R}\}$$

$$= (3k_B T/\gamma) \text{tr}[\mathbf{H}^{-1}]_{ij}$$

where $k_B$ is the Boltzmann constant, $Z$ is the configurational partition function, and $\text{tr}[\mathbf{H}^{-1}]_{ij}$ is the trace of the $ij$th submatrix $[\mathbf{H}^{-1}]_{ij}$ of $\mathbf{H}^{-1}$, $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle$ can be expressed as a sum over the contributions $[\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j]_k$ of the $3N - 6$ individual internal fluctuation modes as $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = \sum_k [\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j]_k$. The contribution of the $k$th mode is explicitly given by

$$[\Delta \mathbf{R}_i, \Delta \mathbf{R}_j]_k = (3k_B T/\gamma) \text{tr}[\lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T]_{ij}$$
where $\lambda_k$ is the $k$th nonzero eigenvalue of $H$ and $u_k$ is the corresponding eigenvector. The eigenvalues are related to the frequencies of individual modes, and the eigenvectors describe the effect of each mode on the positions of the $N$ residues constituting the structure. The eigenvalues are usually organized in ascending order (after removing the six zero eigenvalues corresponding to overall translation and rotation), so that $\lambda_1$ denotes the lowest frequency and $[\Delta R_i \cdot \Delta R_j]_1$ is the correlation for this mode of motion separately. Likewise, $[(\Delta R_i)^2]_1$ is the mean-square fluctuation in the position of site $i$ for mode 1. The slowest vibrational modes usually dominate the collective dynamics of the structure and are particularly relevant to biological function.

3. Determination of Rigid Blocks

Blocks of residues that move together in a coupled manner can be determined by the comparison of two structures of the same protein. This analysis requires the construction of a symmetric matrix termed $D$, whose elements $D_{ij}$ are equal to 1 if the difference $\epsilon_{ij}$ of the distances between two residues $i$ and $j$ in the two protein structures studied is below a specified cutoff and is otherwise set to zero.

$$\Delta_{ij} = |dA(i,j) - dB(i,j)| \quad \text{and} \quad D_{ij} = h(r_{d} - \Delta_{ij})$$

where $dA(i,j)$ is the distance between residues $i$ and $j$ in structure A, $dB(i,j)$ is the distance between residues $i$ and $j$ in structure B, and $h(x)$ is the Heaviside step function $[h(x) = 1 \text{ if } x \geq 0, \text{ and zero otherwise}]$. $D$ has dimensions $N \times N$ for an $N$ residue protein. The value of the cutoff, $r_{d}$, is adjusted so that the analysis yields a reasonable number of blocks (see below).

As the resulting matrix is still complicated, it has to be refined in order to clearly delimit the underlying blocks. This procedure involves starting with the first residue and constituting a block with all consecutive residues $j$, as long

Fig. 3. Top: Ribbon diagrams of the DS, NR, and TS myosin head structures, color-coded on the basis of the calculated B-factors (the colors range from blue to red corresponding to increasing fluctuations). Bottom: Detailed view of the part of the motor domain showing the relay structure on the left and the nucleotide binding site on the right. Note that the color scale has been adapted to indicate changes within this fragment of the overall myosin structure.
as $D(1, j)$ is equal to 1. If $D(1, i)$ is equal to 0, a new block is started with the criteria $D(i, j) = 1$. Diagonal blocks are created this way. Two diagonal blocks A and B then become part of a single block if the matrix element $D(i_A, i_B)$ is equal to 1, where $i_A$ and $i_B$ are the central residues within blocks A and B, respectively (see Fig. 7). The final matrix $D$ is again a binary matrix, with $D(i, j) = 1$ if $i$ and $j$ belong to the same block.

**RESULTS AND DISCUSSION**

**Flexible Regions within the Myosin Head**

Starting from our ANM analysis of the three available structures of the myosin head, it is possible to calculate the overall fluctuations of each amino acid residue in the form of the B-factors commonly used in analyzing crystallographic structures:

$$B_i = \frac{8\pi^2}{3} \langle \Delta R_i \cdot \Delta R_i \rangle.$$

Figure 2 shows plots of these fluctuations for the DS, NR, and TS structures. All calculation of $R_i$ were performed with $r_c = 11$ Å following the study of Atilgan et al. It is noted that excellent agreement between such B-factors and crystallographic data has already been demonstrated for other proteins. We can only make such comparisons in the case of the better resolved DS structure, where the experimental values are available. The comparison with the theoretical results is presented in Figure 2 and shows a good overall agreement, with the exception of residues belonging to the lever arm (residues 775–835) and the RLC. These exceptions are most probably due to the interactions that exist between the myosin lever arms.
within the crystal lattice, but are naturally absent in our calculations. Because the spring constant $\gamma$ is the only remaining parameter of our calculations, its value can be determined by matching the areas under the experimental and theoretical B-factor curves. This has been done for the residues in the zone 1–800 and leads to a value of 1.3 kcal/(Å² mol). This value is comparable to the values found for other proteins.12

We can now compare the DS, NR, and TS states of myosin. All three structures show rather similar overall fluctuations. Each indicates a significant difference between the motor domain (residues 1–775), which is rigid, and the lever arm (residues 820–835), which is flexible. The regulatory light chain, which is located at the end of the lever arm structure, is also very flexible, in contrast to the essential light chain. It should be recalled that these results refer to an isolated myosin head, truncated at residue 835, and do not take into account the effects of interactions with the actin filament or between neighboring myosin motors.

It is also recalled that the myosin head structures we use are incomplete, and the absence of residues in some domains is the cause of significant local differences between the three states, which can be seen in Figure 2. This is notably the case for the peaks observed near residue 410 in the DS and TS structures and near residue 320 in the NR structure. There are however some mechanically significant differences between the three states, most notably for the contact region between the lever arm and the motor domain that are different in DS compared to either the NR or TS structures. This change shows up in Figure 2 as the peak in fluctuations of residues 48–56, which is only seen for NR and TS, whereas only the DS structure shows a peak for residues 508–510. The first peak can be easily explained by the fact that the residues 48–56, belonging to the SH3 $\beta$-barrel, are distant from the lever arm in the NR and TS structures, but close in DS. The second peak is coupled to the fact that the separations between the $\beta$-strand and $\alpha$-helix elements of the so-called “relay” structure from one another are larger in DS than in either NR or TS (the elements of the relay are visible of the left-hand side of the detailed views in the lower part of Fig. 3).

In order to link these results more easily to the three-dimensional structure of myosin, we use color-coded ribbon models (where increasing fluctuations are indicated with a blue to red gradation). The results shown in the upper part of Figure 3 again stress the overall similarity of the fluctuations for the three myosin structures. They also emphasize the flexibility of the loops that compose the actin binding domain at the top of the S1 domain and the, probably artifactual flexibility of the end of the lever arm, compared with the stiffer region near the essential light chain. Fluctuations are also seen to be more important at the surface of the motor domain and in the lever arm, whereas the buried ATP site is a relatively rigid zone. Because it is not easy to see the changes occurring with the motor domain in the full structure, we have added detailed views in the lower part of Figure 3. In addition to the changes in the relay discussed above, these views show that the most rigid region corresponds to switch II (the strand linking the central $\beta$-sheet to the $\alpha$-helix of the relay) in NR and TS, but instead to the ATP binding site in DS. This is in agreement with the remarks of Houssusse et al.,6 indicating that there is a stronger interaction between the elements linked by switch II in the former structures.

The RLC and ELC light chains are known to play an important role biologically, and they can be expected to modify the flexibility of the long $\alpha$-helices that constitute the lever arm. Their effect can be tested theoretically by comparing ANM calculations on the full myosin head with calculations on structures where the light chains have been removed. The results of these calculations are shown in Figures 4 and 5. Removing the light chains is seen to have a dramatic effect. As might be expected, in the absence of these proteins, there is a significant increase in the fluctuations within the lever arm. However, it is also interesting to note that although the more flexible parts of the motor domain (colored in orange in Fig. 5) are still located on the surface of the structure, they do not occur in the same zones. Notably, in the absence of the light chains, the loops near the actin-binding site become less flexible,
although the reason for this long-range coupling is not obvious. Overall, while maintaining the value for the spring constant $\gamma$, the structure without RLC and ELC becomes four times more flexible.

**Structurally Coherent Blocks of Residues**

The crystallographic data available for the DS, NR, and TS structures of the myosin head enables us to study flexibility from another point of view, by asking which blocks of residues move in a coherent, coupled manner as myosin undergoes the conformational changes linked to its motor cycle. We have carried out the rigid block analysis described in the methodology section for the three possible pairs of structures: DS-TS, DS-NR, and TS-NR. The limit distance $r_d$, which determines whether two residues are considered as part of the same block was chosen as 0.1 Å, following the preliminary studies illustrated in Figure 6. These show three representations of the matrix $\Delta$, where $\Delta_{i,j} = |dA(i,j) - dB(i,j)|$. The data shown refer to the case $A = TS$ and $B = DS$. The color of a point within the matrix is red if $\Delta_{i,j} = r_d$ and graduated from red to blue in terms of decreasing distance if $\Delta_{i,j} < r_d$. If $r_d = 10$ Å [Fig. 6(a)], we

![Fig. 7. Binary representation of the matrix $D_{ij}$ where 1's are colored in black and 0's in white. (a,c,e) The comparisons DS-TS, DS-NR, and NR-TS before refinement of the structural blocks (see Methods); (b, d, f) the same comparisons after refinement.](image-url)
obtain only two blocks which correspond, not surprisingly, to the myosin motor domain and the lever arm. By decreasing $r_d$ [Fig. 6(b,c)], a finer distinction of movement is obtained and more blocks appear. The selected limit of $r_d = 0.1$ Å leads to roughly 20 structural blocks after the refinement procedure described in Methods and is a reasonable limit given the limited resolution of the experimental data.

Figure 7 shows the $D$ matrix with $r_d = 0.1$ Å before and after refinement for the DS-TS, DS-NR, and TS-NR pairs. The resulting blocks can be linked to the three-dimensional structure of myosin, again using color-coded ribbon models (Fig. 8). Note that isolated residues and two-residue blocks have been colored gray.

These results are in agreement with the division into four subdomains connected by flexible regions suggested by Houdusse et al., although the subdivisions shown in Figure 8 are somewhat finer. The results for the three pairs of structures analyzed show overall similarity. There are however some notable differences. In particular, the helix at the top of the motor domain [colored tan in Fig. 8(a), residues 416–446] belongs to a single block for the TS-DS pair of structures, but is divided into three blocks [colored tan-yellow-orange in Fig. 8(b,c)] when the structure NR is involved in the comparison. Given the position of these residues, this change may well be related to the fact that the nucleotide binding pocket is occupied in the structures DS and TS, but empty in NR.

In fact, the presence of a nucleotide in the binding pocket seems to lead to larger structural blocks in several regions. Thus, the zone formed by residues 231–243 [shown as ice blue in Fig. 8(a)] forms a single block only when the nucleotide pocket is occupied and a similar result is found for the residues 216–230 and 244–356 [shown in orange in (Fig. 8(a)]. A similar distinction is found within the lever arm and light chains, where the three blocks observed in the presence of a bound nucleotide [Fig. 8(a)], become four blocks when the comparison involves an empty nucleotide pocket. It is also important to note that this analysis clearly shows the “pliant point” within the region 775–780 [indicated by an arrow between the yellow and red blocks in Fig. 8(c)] reported by Houdusse et al.10

**Links between Collective Vibrations and Structural Blocks**

In order to test whether the results obtained by our rigid block analysis are related to the ANM collective vibration analysis, we have repeated the B-factor calculations using a modified spring model of myosin. The modification involves using two different spring constants to mimic the existence of structural blocks. While maintaining the
usual spring constant between residues belonging to different blocks, we increase the spring constant by a factor of 100 for residue pairs within a single block. If the block analysis can be related to rigidity within blocks and flexibility between blocks, the modified spring constants would not be expected to significantly change the calculated B-factors. As a control, we have also carried out B-factor calculations with modified spring constants based on artificially constructed blocks, which cross the block boundaries we have actually determined. Note that the cutoff distance for forming inter-residue springs is kept at 11 Å for all these studies.

Figure 9 displays the modified B-factors calculated with two spring constants for the DS structure, taking into account the structural blocks obtained from the DS-TS comparison. The B-factors calculated with the standard spring constant of 1.3 kcal/(Å² · mol) are shown for comparison. Note that the total area under the two curves has been made equal. It can be seen that the modified B-factors are nearly identical to those calculated with a single spring constant. Minor differences occur for residues 475–525 and residues 650–690, which do not belong to structural blocks and are found to be a little more flexible than with the previous calculation.

We have repeated this analysis for the three available myosin structures, using either of the rigid block definitions involving the structure in question. This leads to a total of six different B-factor curves that can be compared with the single spring constant result. In all cases, the minor changes observed support the compatibility of the rigid block and the ANM analyses.

In contrast, if we use artificially constructed blocks bridging the principal boundaries between the true rigid blocks, much more significant changes in the B-factor curves are found. Compared with the reference B-factor curve, the mean relative error found with the artificial blocks is 22%, compared with only 5% with the correctly formed blocks. We can therefore conclude that there is indeed a close relation between the ANM calculations and the rigid block analysis.

CONCLUSIONS

By combining coarse-grained methods with available crystallographic data, we have been able to study the flexibility of myosin motor protein, a system involving almost 1000 amino acid residues. We have used two approaches to obtain information: first, calculating residue fluctuations using the ANM elastic model and, second, defining rigid structural blocks by an analysis of conformational changes. Good agreement is found with available experimental data.

These two approaches, which have been shown to yield compatible results, enable us to distinguish and to quantify the rigid and flexible domains within the myosin structure. Although, the basic mechanics of myosin seems to be preserved amongst its various known conformations, changes have been detected in the flexibility at the motor domain-lever arm interface and also linked to the presence or absence of a ligand within the nucleotide binding pocket. We have also been able to show that the regulatory and essential light chains play a significant role in determining the rigidity of the myosin lever arm.
REFERENCES