Diffusion Coefficients of Endogenous Cytosolic Proteins from Rabbit Skinned Muscle Fibers

Brian E. Carlson,[†] Jim O. Vigoreaux,[‡§] and David W. Maughan[‡*]

[†]Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, Michigan; [‡]Department of Molecular Physiology and Biophysics, Health Science Research Facility, University of Vermont College of Medicine, Burlington, Vermont; and [§]Department of Biology, University of Vermont, Burlington, Vermont

ABSTRACT Efflux time courses of endogenous cytosolic proteins were obtained from rabbit psoas muscle fibers skinned in oil and transferred to physiological salt solution. Proteins were separated by gel electrophoresis and compared to load-matched standards for quantitative analysis. A radial diffusion model incorporating the dissociation and dissipation of supramolecular complexes accounts for an initial lag and subsequent efflux of glycolytic and glycogenolytic enzymes. The model includes terms representing protein crowding, myofilament lattice hindrance, and binding to the cytomatrix. Optimization algorithms returned estimates of the apparent diffusion coefficients, $D(r,t)$, that were very low at the onset of diffusion (~10^{-10} cm^2 s^{-1}) but increased with time as cytosolic protein density, which was initially high, decreased. $D(r,t)$ at later times ranged from 2.11 $\times$ 10^{-7} cm^2 s^{-1} (parvalbumin) to 0.20 $\times$ 10^{-7} cm^2 s^{-1} (phosphofructose kinase), values that are 3.6- to 12.3-fold lower than those predicted in bulk water. The low initial values are consistent with the presence of complexes in situ; the higher later values are consistent with molecular sieving and transient binding of dissociated proteins. Channeling of metabolic intermediates via enzyme complexes may enhance production of adenosine triphosphate at rates beyond that possible with randomly and/or sparsely distributed enzymes, thereby matching supply with demand.

INTRODUCTION

Glycolysis is a multistep anaerobic chemical process that converts glucose or one glucosyl moiety of glycogen into two pyruvate molecules. The free energy released assists in the formation of the high-energy compounds ATP and NADH used in cellular metabolism. Glycolysis occurs in nearly all organisms, and although specific aspects of the metabolic pathways may vary, the ubiquity of this metabolic process reflects both its evolutionary age and importance in cellular homeostasis.

Glycolysis is of particular importance to the function of vertebrate fast-twitch muscle fibers. These fibers are rich in glycogen granules that give the muscle its whitish tint. By rapidly breaking down glycogen and glucose, the linked processes of glycogenolysis and glycolysis represent an immediate source of ATP, complementing the aerobic pathways mediated by the mitochondria. The mechanical work and power produced by ATP-consuming actomyosin cross-bridges is limited, however, both by the amount of available substrate (one glycosyl unit yields three molecules of ATP and one glucose unit yields two molecules of ATP) and by the rate at which ATP is formed at the intermediate steps of glycolysis. The level of this anaerobic capacity (1) is important to the whole organism and establishes the level of success for a variety of motile activities; e.g., from the fight-and-flight response crucial to our early survival as a species to the performance of an elite sprinter in the modern Olympics. Furthermore, although aerobic pathways are the main conduit of ATP delivery, in many multisystem diseases that exhibit underlying mitochondrial dysfunction (e.g., myalgic encephalomyelitis) (2), glycolysis may be an important compensatory pathway for generating ATP.

Previous studies have established that within the muscle cell, the glycolytic enzymes are located in the fluid cytoplasm (3), specifically, within the I-band region of the sarcomere (4). Indirect evidence suggests that within the crowded interior of the I-band, glycolytic enzymes aggregate and form complexes. The close proximity of the enzymes would enable channeling of metabolic substrate intermediates (3–11). Although the expected enhancement in ATP production may help explain how ATP supply matches demand during high power output (1,12), direct evidence establishing the existence of putative complexes in situ, upon which the concept of metabolic channeling is predicted, remains elusive.

In a previous report (3), we identified and estimated the concentrations of the 15 most abundant cytosolic proteins in rabbit psoas (fast twitch) fibers, including 10 glycolytic enzymes. The glycolytic enzymes were present at the predicted stoichiometric ratios, where the mole fraction of each enzyme in the triose (three-carbon sugar) section of the pathway was roughly twice that of each enzyme in the hexose (six-carbon sugar) section, consistent with a mechanism of metabolic channeling. Here, we examine the rate at which these glycolytic enzymes and other cytosolic proteins diffuse out of skinned psoas fibers after being transferred from oil to physiological salt solution. We looked for evidence of whether the glycolytic enzymes are dissociated and freely diffusible under physiological conditions or exist

Submitted June 11, 2013, and accepted for publication December 31, 2013.

*Correspondence: dmaughan@uvm.edu

Editor: David Wolf.

0006-3495/14/02/0780/13 $2.00

http://dx.doi.org/10.1016/j.bpj.2013.12.044

as organized complexes that may assist their functional role. We reasoned that if the glycolytic enzymes exist as complexes, then the complexed enzymes should diffuse together with the same characteristic diffusion coefficient (which could be very small if the size of the complexed enzymes prevented movement through the actin-myosin lattice structure), and the rate of diffusion should be considerably less than that of the individual molecules in the same aqueous medium. If, however, the glycolytic enzymes are not complexed, the diffusion coefficients should scale more or less according to molecular weight and hydrodynamic radius (for example, smaller molecules like enolase would diffuse faster than larger molecules like pyruvate kinase). We based the model fits on measurements of the radial efflux time courses for 11 of the 15 cytosolic proteins examined previously. The results of the modeling are consistent with the hypothesis that supramolecular complexes exist for both glycogenolytic and glycolytic enzymes in the crowded intracellular environment of the living muscle cell. In this interpretation, complexes rapidly break apart under conditions where the cytosol becomes depleted of diffusible proteins. The apparent diffusion coefficients of the dissociated proteins, reduced by steric hindrance and transient binding in the meshlike myofilament lattice, increase with time as the hindering effect of cytosolic protein crowding diminishes.

MATERIALS AND METHODS

Experimental preparation and solutions

Muscle fiber segments from the psoas (fast-twitch fiber type IID) of New England white rabbits were used. Long segments of intact fibers were isolated in water-saturated mineral oil (7°C), peeled of their sarcolemma using a sharp needle, and cut into ~6-mm-long segments. Each segment was transferred through adjacent drops of relaxing solution (also at 7°C unless otherwise indicated) according to the procedures described below and illustrated in Fig. 1 *A*. Details of the method are given elsewhere (13). The width and length of each segment were measured using a digital micrometer both before and after transfer of the segment to relaxing solution. Sarcomere spacing was ~2.2 μm (i.e., at slack length, untethered). The standard relaxing solution contained (as in Maughan et al. (3)) 6 mM $MgCl_2$, 5 mM Na_2-ATP, 20 mM imidazole, 5 mM EGTA, 45 mM K_2 creatine phosphate, 50 mM sucrose, 4% w/v Dextran T500, and 174 mM ionic strength, pH 7.0 at 20°C (pH 7.3 at 7°C). In some cases, pH and temperature were varied as indicated in the text. As reported previously (14,15), 4% w/v of the high-molecular-weight polymer Dextran T500 ($M_w = 465$ kDa) osmotically constrains the otherwise swollen skinned fiber and its myofilament lattice to approximately its in vivo width (i.e., the fiber width measured initially in oil, ranging from 38.9 μm to 65.7 μm) and lattice spacing (i.e, the center-to-center thick-filament distance measured by small angle x-ray diffraction).

Protein radial diffusion protocols

Intracellular protein diffusion rates were measured by transferring the skinned fiber segments through a series of two or three drops (each 10 μL) of relaxing solution on the floor of a glass-bottomed dish containing 2–3 mL water-saturated mineral oil (Fig. 1 *A*). The volume into which the proteins diffused (10 μL) was very much larger than the fiber volume (3.7 ± 1.2 nL), thereby allowing us to neglect back diffusion in the diffusion models. Stainless steel tweezers were used to clasp one end of a fiber and to quickly transfer it from one drop to another. Fluid adhering to the surface of the fiber between drops was stripped off as the fiber was drawn through the oil phase. Transfers were accomplished within 1 s, unless otherwise indicated. Fibers remained in each drop for 2–1440 s; e.g., 2 s in the first drop, 30 s in the second drop, and 600 s in the last drop; or 35 s in the first drop, 80 s in the second drop, and 1140 s in the last drop, as in Fig. 1 *A*. (Very few experiments were run with a time of 2–4 s in the first drop; most were run with a time of 20–30 s in the first drop.) Each drop, except the last, was stirred by slowly moving the fiber in a circle or figure eight. The last drop was stirred occasionally. Incubation in the last drop was terminated by transferring the fiber (i.e., the cytomatrix fraction) to sodium dodecyl sulfate (SDS) sample buffer for electrophoretic separation of proteins (see below). In most cases, 5 μL aliquots of each drop containing part of the cytosolic fraction were combined with 2.5 μL of SDS sample buffer and 2.5 μL 50% glycerol, loaded onto the gel, and subjected to electrophoresis alongside diluted samples of the cytomatrix fractions (0.5 nL fiber volume per μL).

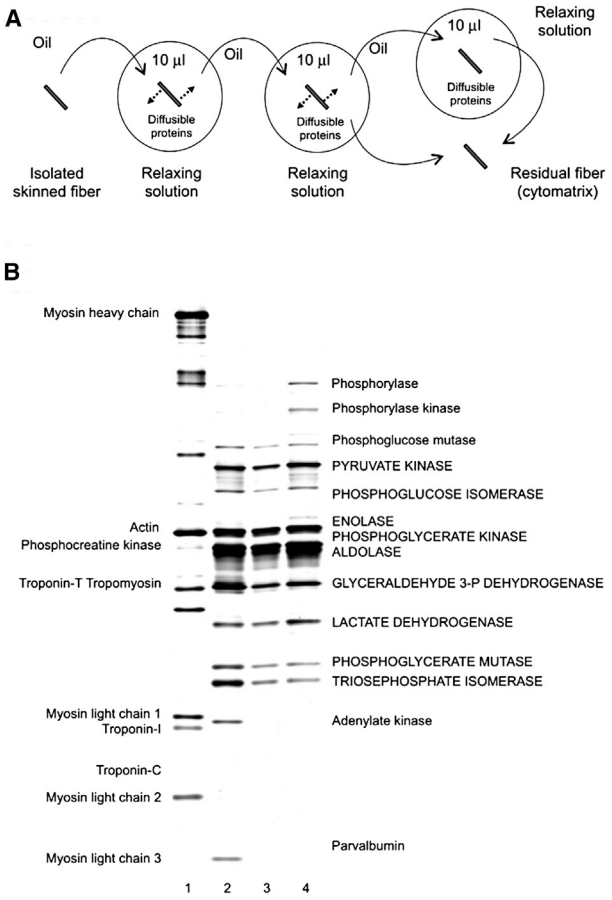

FIGURE 1 Experimental setup and radial diffusion of endogenous cytosolic proteins into relaxing solution from skinned fiber segments of rabbit psoas muscle. (*A*) The experimental protocol, where each fiber is placed sequentially in either two or three drops of relaxing solution (~10 μL each) for a specified period of time ranging from 2 s to 20 min (see text). (*B*) Proteins separated by SDS polyacrylamide gel electrophoresis with silver-stained gel. Proteins are identified as in Maughan et al. (3). Lanes 2–4, cytosolic protein fractions obtained by sequentially incubating the skinned fiber for 35 s (lane 2, load 5.4 μL = 51% of total volume), 80 s (lane 3, load 7.2 μL = 68% of total), and 1440 s (lane 4, load 10 μL = 90% of total) in 10 μL relaxing solution drops. Lane 1, cytomatrix fraction (residual fiber, load 5 μL = 14% of total). Fiber radius, 29.5 μm; fiber length, 6.56 mm.

SDS polyacrylamide gel electrophoresis

Proteins were separated by SDS polyacrylamide gel electrophoresis, and gels were analyzed as described elsewhere (3). Sample volumes were adjusted to achieve an average load of ~5 ng per protein band (determined by trial and error). The sample volumes, thus established, also set the maximum number of drops through which the fibers were transferred (nominally three) that allowed the gel analysis to remain within the lower limit of the sensitivity of the assay (1 ng, with a linear range of 1–25 ng). The dilution factors were used to correct band densities during subsequent analysis.

Modeling and parameter optimization

The skinned fiber segment was modeled as a long cylinder bathed in a relatively large, well-stirred volume of solution, as follows: Net movement of protein is directed radially outward, with net outward diffusion of individual species (coupled to dissociation of protein complexes, if present) beginning the moment the cylindrical fiber is transferred from oil to solution. Immediately before transfer, the bath is free of protein and the diffusible proteins and/or protein complexes within specific repeating subsegments of the fiber (defined later) are assumed to be uniformly distributed. Transfer times between drops of solution are neglected, since the time it took to transfer a fiber from one drop to another was short (<1 s) compared to the time a fiber remained in each drop. Taking these conditions into consideration for a specific fiber segment, diffusion in its simplest form is represented by the one-dimensional radial diffusion equation (16):

$$\frac{\partial C(r,t)}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[Dr\frac{\partial C(r,t)}{\partial r}\right] \quad (1)$$

and is subject to the following boundary and initial conditions:

$$\begin{aligned} \text{Boundary } C(a,t) &= 0;\ \frac{\partial C(0,t)}{\partial t} = 0 \\ \text{Initial } C(r,0) &= C_0 \text{ for } r<a, \end{aligned} \quad (2)$$

where C is the concentration of the specified cytosolic protein, r is the radial location inside the fiber (ranging between 0 and 20.6 μm in our simulations), and t is the time measured from the moment the fiber is initially transferred into the relaxing solution, D is the corresponding diffusion coefficient, a is the radius of the muscle fiber, and C_0 represents the initial concentration of the specified protein, assumed to be distributed uniformly throughout the specified region (or segment) of the skinned fiber.

We considered two models, the Constant D model, in which D is constant for all radial positions and times, and the Variable D model, in which D varies with radial position and time due to changes in local cytosolic protein density. Both models were evaluated by solving Eqs. 1 and 2 numerically, as described below. We determined the diffusional characteristics of 11 cytosolic proteins for which efflux time course data from experiments were available. These proteins included seven glycolytic enzymes (phosphoglucose isomerase, phosphofructose kinase, triosephosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate mutase, enolase, and pyruvate kinase) and four other abundant proteins in rabbit muscle cytosol (parvalbumin, adenylate kinase, phosphoglucose mutase, and glycogen phosphorylase, of which the latter two are glycogenolytic enzymes).

Previous immunohistochemical studies show that the glycolytic enzymes (as well as diffusible phosphocreatine kinase) reside in the I-band region of the sarcomere (17,18). Localization to the I-band is consistent with these enzymes being part of a supramolecular complex, the size of which would restrict diffusion into the more congested A-band (3). However, since we do not know a priori whether the enzymes are initially complexed or not, we referred all concentrations to total cytosolic volume. These concentrations included glycolytic enzyme concentrations as well as those of proteins more broadly distributed in the cytoplasm.

In both models, the initial conditions are established by the cytosolic diffusible protein density at the onset of diffusion ($t = 0$) and the fixed parameters listed in Table 1. We assumed, as an approximation, that the total initial diffusible protein density at $t = 0$ is equal to the sum of the 15 most abundant cytosolic proteins in rabbit psoas muscle, 1121 μmol/L cytosolic fluid volume (Tables II and III in Maughan et al. (3)). We neglected other cytosolic proteins and polypeptides, since their contribution to the total is relatively small (3,19).

In the Constant D model, the diffusion coefficients are independent of protein density, fiber radius, and time; therefore, the diffusion of one protein is uncoupled from that of another, allowing the normalized efflux time course to be solved for each protein species independently of all others. As shown later (Results), the Constant D model does not adequately fit the time course data, especially at short times.

In the Variable D model, the diffusion coefficients depend on radial and temporal variations in protein density, thereby introducing a conceptual basis for a more accurate model fit to the diffusion time course data. Included in this model are four mechanisms that contribute to the time-varying reduction of the diffusion coefficients (20–22).

Obstructions and tortuosity due to cytomatrix proteins

The actin and myosin filaments in a muscle fiber form a regular lattice structure that impedes free diffusion of proteins. The reduced diffusion coefficient, D_h, resulting from cytoskeletal hindrance is determined from the expression for a hexagonal arrangement of myofilaments (23) and the assumption that a protein of hydrodynamic radius, R_h, diffuses through a lattice structure at the same rate as a point particle through an altered structure reduced in spacing by R_h (20):

$$D_h = D_w\frac{1-[(R_h+R_o)/L]^2}{1+[(R_h+R_o)/L]^2}, \quad (3)$$

where D_w is the diffusion coefficient of a given protein in aqueous solution, R_o is the mean radius of the myofilaments, and L is the half-mean center-to-center spacing of the myofilaments. Although the structural differences between A- and I-band lattices may differentially affect the radial efflux of protein from each region, we assume, for simplicity, nominal values of R_o and L that reflect those of the A-band (Table 1). We assume this because 1), I-band values roughly approximate those of the A-band if the presence of titin, which runs alongside the thin filaments in the I-band (24) and nebulin, another actin-binding protein (25), are taken into account; and 2), at the sarcomere length at which the present experiments were conducted (2.2 μm), the A-/I-band volume ratio is 3:1 (3), so it is likely that the radial efflux of diffusible proteins from the A-band exceeds that from the I-band. Even complexed proteins initially confined to the I-band, once dissociated, will likely contribute to the radial efflux of proteins from the A-band, since the equilibration times associated with local axial fluxes between flanking I- and A-band regions of the 2.2 μm sarcomere will be short compared to equilibration times associated with the radial efflux of protein between the core and periphery of the ~20 μm radius skinned fiber.

Excluded volume due to cytosolic proteins

The diffusion rate of a protein in solution is reduced in the presence of a high concentration of other nonreactive background proteins, because the diffusing proteins are excluded from the space occupied by the other proteins (22,26). In this case, the reduced diffusion coefficient, D_c, is given by (22)

TABLE 1 Fixed and free parameters and defined variables used in models

Fixed parameters

Symbol	Description	Value	Units	Reference
a	Average radius of fibers	20.6 ± 4.3	μm	Fig. 2 legend
M_o	Oligomeric molecular weight of protein[a]	11.9–340.3	kDa	Table 2
D_w	Diffusion in dilute aqueous solution of protein[a]	(2.48–7.75) × 10^{-7}	$cm^2\ s^{-1}$	Table 2
R_h	Hydrodynamic radius of protein[a]	1.83–5.69	nm	Table 2
V_h	Hydrodynamic volume of protein[a]	0.0129–0.0137	dL/g	Interpolated from (22)
R_{thin}	Radius of thin filament	4	nm	(66)
R_{thick}	Radius of thick filament	6.5	nm	(66)
R_o	Average radius of thick and thin filaments	5.25	nm	(66)
L	Half center-to-center thick-thin filament spacing	12.85	nm	(15)
n_s	Supramolecular binding cooperativity exponent[b]	5[c]	unitless	—
k	Boltzmann constant	1.381 × 10^{-23}	J K^{-1}	
η	Viscosity of water at 7°C, at 39.4°C	1.45, 0.65	cP	(37)
T	Temperature of experimental conditions	280, 312.4	K	

Free parameters

Symbol	Description	Optimization range[d]	Optimum Values	Units
φ	Ease of vacancy, hydrodynamic volume product	0.01–1.00	0.12 ± 0.01	dL g^{-1}
k_b	Apparent binding constant of proteins to lattice	0–15	0.96 ± 0.17	unitless
k_s^{max}	Maximal supramolecular complex binding constant	10–1500	628 ± 111	unitless
ρ_{50}	Total cytosolic protein density at 50% binding	0–50	21.96 ± 0.90	mmol L^{-1}
D	Optimized diffusion coefficient (Constant D)	(0–1.00) × 10^{-7}	(0.08–0.68) × 10^{-7}	$cm^2\ s^{-1}$

Variables

Symbol	Description	Returned value range[e]	Units
$\rho(r,t)$	Total cytosolic protein density in a radial annulus	0–1.121	mmol L^{-1}
$D(r,t)$	Variable diffusion coefficient (Variable D)[f]	0.20–1.60 × 10^{-7}	$cm^2\ s^{-1}$

[a]Values depend on protein species.
[b]Index of the steepness of the transition from maximal to minimal supramolecular binding, defined in Eq. 6.
[c]For the simulations, the value was fixed at 5, corresponding to an arbitrarily high value for cooperativity of the supramolecular binding and dissociation (see text for details).
[d]Range of values over which optimum value is determined.
[e]Range of values returned from simulations. The values tabulated here for the optimized parameters are the mean ± SD over the 20 best fits to the data. The RMS error between simulation and data of the 20 best fits is 1.66 ± 0.02%.
[f]Evaluated at $t/a^2 = 0.035$.

$$D_c(r,t) = D_w \exp\left[\frac{-\gamma V_h \rho(r,t)}{1 - V_h^b \rho(r,t)}\right], \tag{4}$$

where γ is a unitless constant that represents the ease-of-vacancy formation of the protein, V_h is the average protein exclusion volume, which is equated to the hydrodynamic volume of the protein in accordance with O'Leary (22), $V_h{}^b$ is the average hydrodynamic volume of the background proteins in the cytosol, and $\rho(r,t)$ is the sum of the product of the concentrations of each protein (mmol/dL) and their respective molecular weights (g/mmol) present in the cytosol within a specific radial annulus at a given time, hereafter referred to as the total cytosolic protein density within that annulus.

Binding to cytomatrix proteins

There is ample evidence that some diffusible proteins bind transiently to structural proteins of the cytoskeleton (6,7,27–30). Gershon et al. (21) investigated the generic properties of protein binding to the cytoplasmic matrix in eukaryotic cells, which lead to a simple expression for the reduced diffusion coefficient, D_b:

$$D_b = \frac{D_w}{1 + k_b}, \tag{5}$$

where k_b is the apparent binding constant of diffusible proteins to the cytomatrix, indicating the ratio of bound to free protein molecules.

Supramolecular complexes

From thermodynamic principles (31,32), it is likely that enzymes in the crowded cytosol of the I-band and other regions of the myofibril form supramolecular complexes, the diffusion of which is severely restricted due to steric hindrance and binding to the myofilaments, as noted above. Clearly, enzymes within the complex must dissociate before they are free to diffuse. To factor in this circumstance, we adopt a form of the reduced diffusion coefficient similar to that in Eq. 5, incorporating a variable supramolecular binding term, k_s:

$$k_s = k_s^{\max} \frac{[\rho(r,t)/\rho_{50}]^{n_s}}{1 + [\rho(r,t)/\rho_{50}]^{n_s}} \tag{6}$$

where $k_s{}^{max}$ is the maximal supramolecular complex binding constant (i.e., the maximum ratio of complexed to free protein molecules), ρ_{50} is the total cytosolic protein density at half-maximum binding, and n_s is the supramolecular binding cooperativity parameter (i.e., an index of the steepness of the transition from maximal to minimal supramolecular binding). For simplicity, we assume that $k_s{}^{max}$, ρ_{50}, and n_s are the same for all complexes

(whether glycolytic or glycogenolytic) and are independent of position and time.

We combined all factors described by Eqs. 3–6 to yield an expression for an apparent diffusion coefficient of a given cytosolic protein in the muscle fiber, $D(r,t)$, that is independent of axial position but depends on radial position and time through the variation of total cytosolic protein density, $\rho(r,t)$:

$$D(r,t) = D_w \frac{\dfrac{1-[(R_h+R_o)/L]^2}{1+[(R_h+R_o)/L]^2}\exp\left[\dfrac{-\gamma V_h \rho(r,t)}{1-V_h^b \rho(r,t)}\right]}{1+k_b+k_s^{\max}\dfrac{[\rho(r,t)/\rho_{50}]^{n_s}}{1+[\rho(r,t)/\rho_{50}]^{n_s}}}, \quad (7)$$

where, as defined in Eqn. 4, $\rho(r,t)=\sum_i^N C_i(r,t)M_{o,i}$, where N is the total number of cytosolic proteins accounted for in this model (14), $C_i(r,t)$ is the concentration of protein i at radius r and time t, and $M_{o,i}$ is the molecular weight of protein i. The apparent diffusion coefficient is different for each protein and increases with time at any given position as $\rho(r,t)$ decreases. The formulation for $D(r,t)$ in the Variable D model does not account for specific protein-cytomatrix binding, protein-protein binding, or the competitive inhibition between these kinetic processes. Accounting for these possible interactions would either require detailed biochemical experiments to quantify the associated binding kinetics or would greatly increase the number of free parameters in this model. Units for all model parameters (fixed and free) and variables, assumed values for fixed parameters and the value ranges for each of the free parameters are given in Table 1.

The 15 partial differential equations, representing diffusion for each diffusible protein included in the Variable D model, are coupled through the protein crowding term (Eq. 4) and the supramolecular binding term (Eq. 6), both functions of $\rho(r,t)$. Therefore, the equations must be solved simultaneously. Of the 10 glycolytic enzymes whose diffusional characteristics are provisionally assumed to be governed by Eqs. 1, 2, and 7, experimental efflux data were available for only seven: phosphoglucose isomerase, phosphofructose kinase, triose-phosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, enolase, pyruvate kinase, and phosphoglycerate mutase. The diffusional characteristics of the remaining four (aldolase, phosphoglycerate kinase, lactate dehydrogenase, and phosphocreatine kinase) were specified by the global fit to the efflux time courses for the seven for which data are available, as was the diffusion of phosphocreatine kinase, since diffusible phosphocreatine kinase has been shown to bind to phosphofructose kinase and aldolase in the I-band of skeletal muscle (33). Two glycogenolytic enzymes for which experimental efflux data were available (phosphoglucose mutase and glycogen phosphorylase) are closely associated with particles of aggregated glycogen located in the extramyofibrillar space (34) and were therefore provisionally included in the Variable D model as complexed enzymes also governed by Eqs. 1, 2, and 7. Finally, in applying the Variable D model to adenylate kinase and parvalbumin (broadly distributed small proteins involved in cell metabolism and Ca^{2+} handling), we omitted terms in Eq. 7 that correspond to supramolecular equilibrium binding (Eq. 6), because both of these proteins have been shown to diffuse freely in skeletal muscle (10,13).

The data sets for each protein were composed from experiments run on multiple muscle fibers of different fiber radii. The time associated with the protein diffusion for each fiber was normalized by the square of the fiber radius. This normalized time value is denoted as t/a^2 and is scaled with a different a determined by the specific radius of each fiber. Data were pooled from multiple fibers to obtain data sets that were sufficiently large for analysis. Pooling was necessary because it was not possible to obtain more than one or two data points per fiber and because free parameters defined from model fits to a small number of data points have limited value (especially those from the Variable D model). When fitting a model to the data, we used the average fiber radius in the data set ($\bar{a}$ = 20.6 μm; Fig. 2 legend); thus, for both the Constant D and Variable D model simulations, the normalized time is equivalent to $t/\bar{a}^2$.

From the simulation solution, the amount of each protein that diffuses out of the fiber as a function of time, $M(t/a^2)$, was determined by integrating the concentration profile at a given time across the fiber cross section (yielding the total amount of diffusible protein retained in the myofiber) and then subtracting that amount from the initial amount. The initial amount was assumed to be equivalent to the total amount of protein that diffused out of the fiber after a long (infinite) time, M_∞. This value was used to normalize $M(t/a^2)$ to give data points of $M(t/a^2)/M_\infty$ ranging from 0 to 1, which were used to calculate the simulation curves of $M(t/a^2)/M_\infty$.

O'Leary (22) examined the diffusion rates of human and earthworm hemoglobin at different concentration levels. In modeling his results O'Leary noted that the product of the ease-of-vacancy formation (γ) and the hydrodynamic volume (V_h) is roughly constant across both proteins, despite their extreme difference in V_h. Thus, for our simulation, we assumed $\varphi = \gamma V_h$ is constant across all 15 proteins, whose V_h values lie almost entirely between the two extremes. Because there are no literature values for the hydrodynamic volumes of the proteins examined here, we used a linear fit to the plot of hydrodynamic volume versus molecular weight for human and earthworm hemoglobin reported by O'Leary (22) to obtain an interpolated hydrodynamic volume for each protein in this study (see Section B in the Supporting Material). As a check, we used these estimated hydrodynamic volumes and the Stokes-Einstein relationship to calculate the diffusion coefficients of the 15 proteins in bulk water. These calculated diffusion coefficients were slightly larger (by 7–10%) compared to those calculated from a general equation (35) based on experimental values obtained from other proteins. In addition to the parameter φ, we assumed for simplicity that the following parameters are constant across all proteins: the apparent binding constant of diffusible protein to cytomatrix (k_b), the binding constant of maximal supramolecular complex formation ($k_s^{\max}$), and the total cytosolic protein density at half-maximum complex formation (ρ_{50}). In applying Eqs. 1, 2 and 7, we returned optimized values for these four free parameters that yielded the best (simultaneous) fit to the seven glycolytic and four cytosolic proteins for which we had diffusion time course data. Preliminary optimizations were performed with the steepness index of supramolecular binding (equivalent to the cooperativity of binding), n_s, included as a free parameter. However, optimization of fit was found to be relatively insensitive to this parameter. Because values of $n_s > 5$ appear to make very little difference in the transition from maximal to minimal supramolecular binding, optimizations were therefore performed with $n_s = 5$ in the Variable D model. Minimizing the error between simulation and data returned optimum values for the four free parameters using a hybrid simulated annealing/gradient-based method. The same optimization method was used to specify the apparent diffusion coefficient in the Constant D model. Simulations and optimizations were performed on a MATLAB computational cluster utilizing 32 dual-core AMD Opteron-based servers (HP DL145 G2, 4GB RAM). The MATLAB Parallel Computing Toolbox was incorporated to distribute the computational tasks across the cores of the cluster to speed up the optimization procedure. Complete details of the hydrodynamic volume estimation method and calculation of error between simulation and data used during optimization are included in the Supporting Material.

RESULTS

Radial diffusion of proteins from skinned fiber

We measured the diffusion coefficients of 11 cytosolic protein species by assessing the fractional amount of each that diffused from a skinned muscle fiber. The fiber was skinned under oil and subsequently transferred through drops of relaxing solution. Comparison of skinned-fiber dimensions

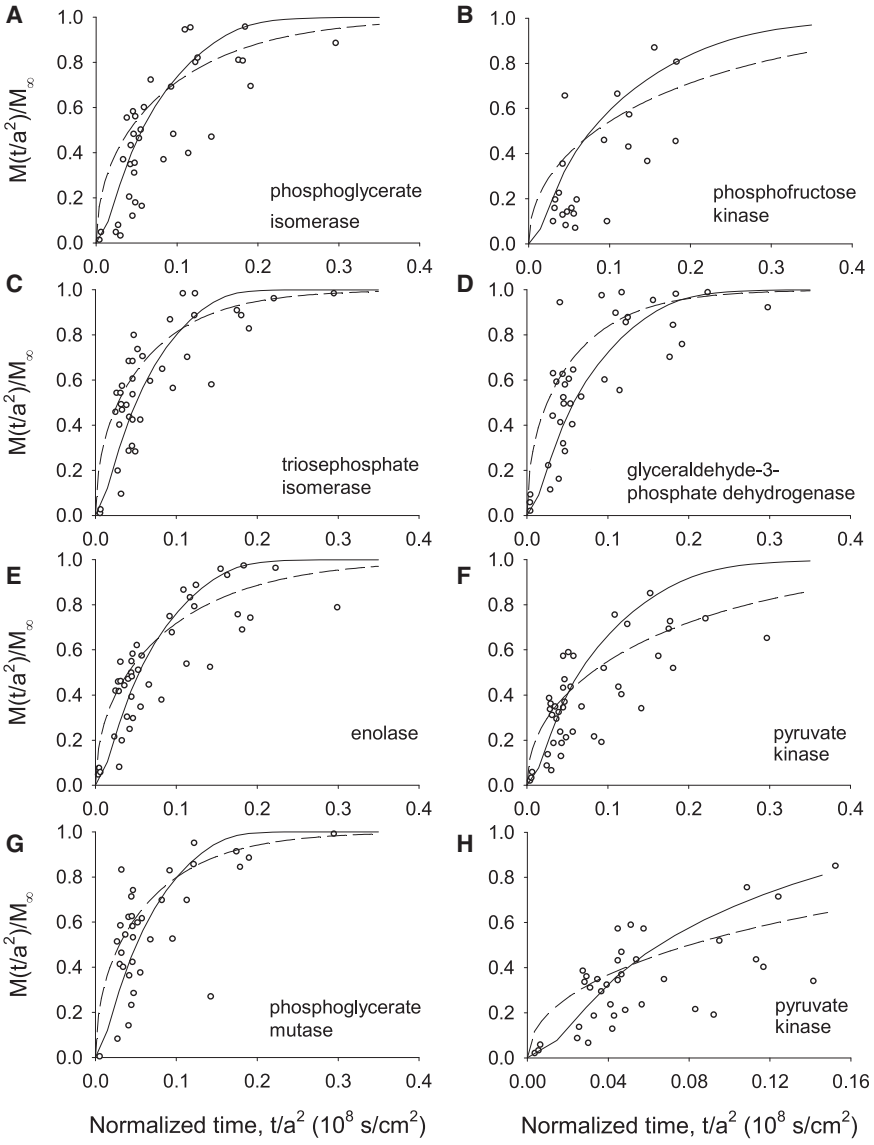

FIGURE 2 (*A–G*) Radial diffusion time courses for seven glycolytic enzymes. Ordinate, amount of protein, $M(t/a^2)$, relative to the total, M_∞, that diffuses from a segment of a skinned psoas fiber; abscissa, elapsed time fiber remains in each drop divided by the square of the fiber radius. Experiments were performed according to the following protocol. Skinned fibers were transferred sequentially from oil to drop 1, where they remained for variable periods of time, then through oil (<1 s) to drop 2, where they remained for 600 s. Average length of fiber segments, ~6 mm; average radius, 20.6 ± 4.3 μm (mean ± SD) In a subset of fibers (the three-drop experiment), fibers remained for variable periods in drops 1 and 2 before being transferred to drop 3 (600 s). Curves are generated from numerical solutions of the Constant D model (*hatched lines*) and Variable D model (*solid lines*). (*H*) Duplicate of *F* (pyruvate kinase) at expanded timescale. The equilibration time for endogenous pyruvate kinase is comparable to that reported for exogenous Rh-pyruvate kinase (55). See text for details.

before and after transfer from oil to solution confirmed that the fiber dimensions remained virtually unaltered due to the physiological colloidal pressure replicated by the osmotic agent, Dextran T500, in the relaxing solution. Our ability to separate the proteins on polyacrylamide gels and to measure the relative density of each band from aliquots from each drop allowed us to quantify the efflux of each protein. Fig. 1 *B* demonstrates the expected outcome that small proteins diffuse out of the fiber more rapidly than large proteins. For example, monomeric parvalbumin (11.9 kDa) and adenylate kinase (21.6 kDa) are present only in the first drop (lane 2, efflux 0–35 s) whereas oligomeric glycogen phosphorylase (194.3 kDa) and to a lesser degree pyruvate kinase (231.7 kDa) are most conspicuous in the third drop (lane 4, efflux 80–1440 s). Fig. 1 *B* indicates that only myofibrillar proteins remain in the fiber at the end of the diffusion time. Preliminary (not shown) and previous experiments (3) confirmed that 10- to 20-min incubations were sufficient to clear the fiber of diffusible cytosolic proteins.

Fig. 2, *A–G*, summarizes the diffusion time courses for seven glycolytic enzymes. Data from fibers of different radii were pooled by normalizing the elapsed time (t) with the square of the fiber radius (a^2). The fractional amount of protein in the solution drop, $M(t/a^2)/M_\infty$ (where $M(t/a^2)$ is the total amount diffused out at time t and M_∞ is the total amount diffused out at infinite time for a specific protein) is plotted against t/a^2 to yield the efflux trajectories. The solid lines represent the best optimized fit of the Variable D model to the data. Constant D model fits (Fig. 2, *dashed lines*) are shown for comparison. Least-squares analysis and visual inspection of the model fit to the data confirm that the Variable D model represents the data more accurately than does the Constant D model (RMS error = 0.0158 vs. 0.0169, respectively). Although this small difference in

RMS error is not by itself sufficient to distinguish between the two models, other criteria (such as Akaike or Bayesian information criteria (see Supporting Material) strongly favor the Variable D model. Furthermore, fitting the Constant D model to the 11 data sets requires 11 adjustable parameters, i.e., one diffusion coefficient for each of the 11 proteins, whereas fitting the Variable D model requires only four adjustable parameters that apply globally to all the proteins. In addition, the adjustable parameters of the Variable D model give us insight into mechanisms that could be contributing to the noted reduction in diffusion of proteins within these fibers. Columns 6 and 5 of Table 2 list values for the diffusion coefficients returned from the best fits to the Constant D and Variable D models, respectively. The optimized parameter average and standard deviation values for the 20 best fits of the Variable D model to the experimental data are given in Table 1.

Fig. 3, *A–D*, summarizes the results for four other diffusible cytosolic proteins. Fits of the Constant D model (Fig. 3, *solid lines*) and Variable D model (Fig. 3, *dashed lines*) were generated as in Fig. 2. Like the glycolytic enzymes, the diffusion time courses for the glycogenolytic enzymes glycogen phosphorylase and phosphoglucose mutase are better represented by the Variable D model than by the Constant D model. In contrast, the diffusion time courses for adenylate kinase and parvalbumin generated by the Variable D model could not be distinguished from those of the Constant D model given the available data.

Although most experiments were carried out at 7°C and pH 7.3, some were conducted at a temperature and intracellular pH that were closer to those in vivo (39.4°C and pH 6.8 (3,36). An approximation of the mean radial diffusion coefficient at a specific t/a^2, denoted as D_{t/a^2}, was calculated by applying the Constant D model to drops 1 and 2, given the value of $M(t/a^2)/M_\infty$ measured at the time the fiber was withdrawn from drop 1. For the case where ~2 s had elapsed, i.e., where $t/a^2 = 0.049\ (\pm\ 0.0032) \times 10^8$ s cm^{-2}), $D_{t/a^2=0.049}$ at 39.4°C was found to be approximately double that at 7°C ($D_{t/a^2=0.049}{}^{39.4°C}/D_{t/a^2=0.049}{}^{7°C} = 1.93 \pm 1.16$). The average Q_{10} of $D_{t/a^2=0.049}$ (calculated according to the formula $Q_{10}{}^{(39.4°C-7.4°C)/10} = D_{t/a^2=0.049}{}^{39.4°C}/D_{t/a^2=0.049}{}^{7°C}$) was 1.21 ± 0.19 for 9 of 11 proteins listed in Table 2 (the Q_{10}s of parvalbumin and adenylate kinase were not calculated, because the rates were too rapid at 39.4°C to be measured accurately). The value of 1.21 for Q_{10} is consistent with that for passive diffusion (1.32) (36,37). Because enzymatic reactions generally have Q_{10}s between 3 and 4, we conclude that metabolic processes do not play a significant role in determining the diffusion coefficients.

Oil or unstirred layer artifact?

Inspection of the data points of Fig. 2 reveals an apparent delay in diffusion of the order of 10 s after the onset. Although we provisionally attribute this to a delayed break-up of a putative supramolecular complex, it is possible that the lag reflects the presence of oil or an unstirred layer of polarized water (38) that dissipates with stirring. However, because we started stirring immediately upon transfer from oil to solution, it is unlikely that a stagnant layer of oil or polarized water, even if present initially, would be sustained for more than a second or two. The results of a three-drop experiment (presented in Table S2 in the Supporting Material) are consistent with this view. A freshly skinned fiber was transferred from oil to a drop of relaxing solution and after 2 s was returned to oil. During this brief period in solution, 3–9% of the diffusible protein

TABLE 2 Physical constants and radial diffusion coefficients for cytosolic proteins of rabbit psoas muscle fibers.

1	2	3	4	5	6	7
Protein	M_o (kDa)	R_h (nm)	D_w (× 10^{-7} cm^2/s)	$D(r,t)$ (× 10^{-7} cm^2/s)	D (× 10^{-7} cm^2/s)	n
Phosphoglucose isomerase	125.2	4.0	3.51	0.57	0.15	31
Phosphofructose kinase	340.3	5.7	2.48	0.20	0.08	19
Triosephosphate isomerase	53.3	3.0	4.69	0.99	0.23	32
Glyceraldehyde-3-phosphate dehydrogenase	144.7	4.2	3.34	0.50	0.25	29
Phosphoglycerate mutase	57.0	3.1	4.59	0.96	0.21	30
Enolase	93.9	3.7	3.87	0.70	0.16	38
Pyruvate kinase	231.7	5.0	2.84	0.33	0.08	39
Glycogen phosphorylase	194.3	4.7	3.02	0.39	0.08	23
Phosphoglucose mutase	61.4	3.2	4.47	0.91	0.16	28
Adenylate kinase	21.6	2.2	6.35	1.6	0.50	21
Parvalbumin	11.9	1.8	7.75	2.1	0.68	12

Column 1 lists the diffusible protein, with glycolytic enzymes in bold type. Column 2 lists the oligomeric molecular weight, M_o (sources listed in Table I of Maughan et al. (3)). Column 3 lists the hydrodynamic radius, R_h, calculated from $R_h{}^3 = (3/4\pi)\ V_h$, where V_h is the hydrodynamic volume of each protein interpolated from the hydrodynamic volume versus molecular weight data of O'Leary (22). See text and Supporting Materials for details. Column 4 lists the estimated value of the diffusion coefficient of the protein in bulk water, D_w (47), calculated from $D_w = (kT)/(6\pi\eta R_h)$, where k is the Boltzmann constant, $T = 280$ K, $\eta = 1.45$ cP at 7°C, and R_h is from column 3. Column 5 lists values for the Variable D model radial diffusion coefficients $D(r,t)$ at $t/a^2 = 0.35$ and $r \sim a$, calculated from an optimized global fit to the experimental data. Column 6 lists values for the Constant D model radial diffusion coefficients (D, constant over the diffusional time course). Column 7 lists the number (n) of fibers sampled. Temperature, 7 ± 1°C; pH 7.3.

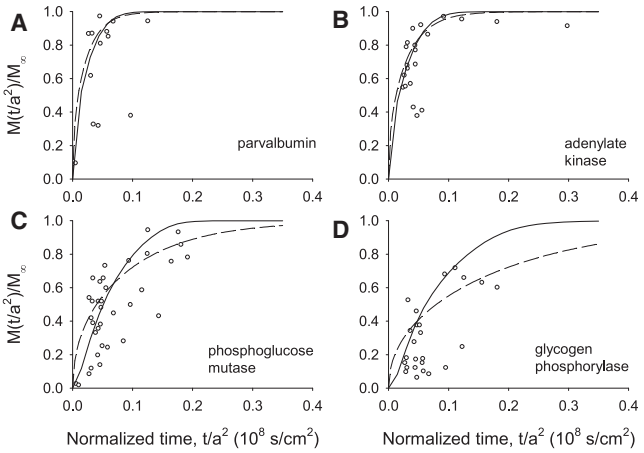

FIGURE 3 (*A–D*) Radial diffusion time courses for four cytosolic proteins. Ordinate, abscissa, and procedures are as in Fig. 2. Curves are generated from numerical solutions of the Constant D model (*hatched lines*) and Variable D model (*solid lines*). Note that the initial diffusional lag is evident in *C* and *D*. Fiber dimensions are as in Fig. 2. See text for details.

left the fiber, the exact amount depending on the protein species and fiber diameter. After a 60 s equilibration period in oil, the fiber was transferred to a second drop of relaxing solution for 2 s. That is, the initial conditions were reestablished and the procedure repeated, during which the fiber was depleted of a slightly greater amount of diffusible protein. Finally, the fiber was transferred rapidly through oil to a third drop of relaxing solution for a final washout period of 10 min. Values of D_{t/a^2} at $t/a^2 = 0.049$ (i.e., D of the Constant D model evaluated at $t/a^2 = 0.049$) were calculated for a representative set of proteins by applying the Constant D model to drops 1–3 and to drops 2–3, respectively, as explained in the Supporting Material. D_{t/a^2} at $t/a^2 = 0.049$ derived from the second transfer exceeded those from the first by a factor of 2.8 ± 0.6. If oil or polarized water were responsible for the lag, and it dissipated similarly for both transfers, roughly similar values of D_{t/a^2} at $t/a^2 = 0.049$ would be expected, and this was not the case. However, if a diffusion lag occurred because a protein complex (which takes time to dissociate) is initially present in the first transfer but not in the second, then a significantly higher value of $D_{t/a^2} = 0.049$ would be expected, and this is what was observed.

The relative contributions of viscosity and steric hindrance to reduced protein diffusivity

Fig. 4 *A* shows how the apparent diffusion coefficient returned from either the Constant D model (*open circles*) or the Variable D model (*solid circles*) is inversely proportional to molecular weight. This reciprocity can be at least partially explained by a Stokes-Einstein relationship (*dotted line*), where diffusion in water, D_w, is given by (39,40)

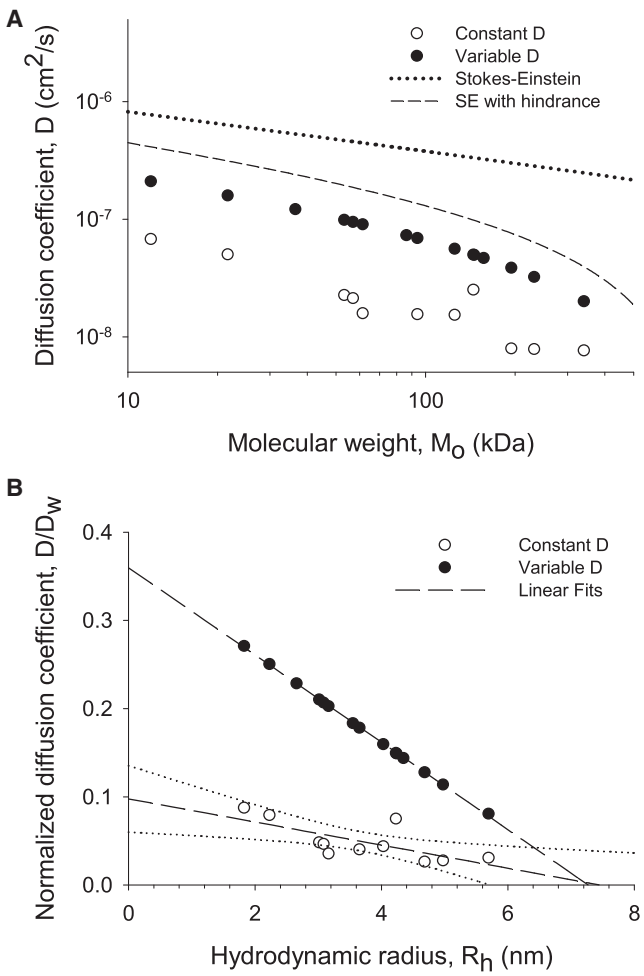

FIGURE 4 (*A*) D (Constant D model) and $D(r,t)$ (Variable D model) evaluated at later times versus protein molecular weight contrasted against predicted Stokes-Einstein relationships, with and without steric hindrance. Open and solid circles represent diffusion coefficients determined with the Constant D model, and those determined with the Variable D model ($D(r,t)$ evaluated at $t/a^2 = 0.35$ where $r = a$, i.e., the fiber radius), respectively. The dotted and dashed lines give the Stokes-Einstein relations using diffusion coefficients determined experimentally by Young et al. (35), without and with, respectively, the lattice hindrance term. The slopes of the double log plots are about −1, consistent with an inverse relationship between diffusion coefficient and hydrodynamic radius. (*B*) D/D_w (Constant D model) and $D(r,t)/D_w$ (Variable D model) at $t/a^2 = 0.35$ and $r \sim a$ as functions of the hydrodynamic radius. The Variable D curve is plotted according to Eq. 9, where values of $D(r,t)$, D_w, and R_h are from Table 2. Straight dashed lines are linear regression fits, and curved dotted lines are 95% confidence limits. Although the general form of Eq. 9 is nonlinear, the resulting plot (*solid circles*) is, to the naked eye, essentially linear for the parameter values used ($R_o = 5.25$ nm, $L = 12.85$ nm, and $R_h = 1.83$–5.69 nm), yielding an x axis intercept (i.e., the sieving radius) of 7.6 nm, i.e., a value close to that obtained by the linear fit.

$$D_w = \frac{kT}{6\pi\eta_w R_h}, \tag{8}$$

where k is Boltzmann's constant, T is temperature, η_w is the viscosity of water and R_h is the hydrodynamic radius of the

diffusing protein calculated from the linear-fit estimate of hydrodynamic volume described earlier.

Note, however, the substantial reduction in the magnitude of the diffusion coefficients beyond that predicted from a Stokes-Einstein relationship. The marked discrepancy indicates the presence of other retardation factors. As noted above, an important contributor is steric hindrance due to the lattice structure (Eq. 3) which, when factored in, further attenuates the diffusion coefficients (Fig. 4 *A*, *dashed line*):

$$D_h = \left[\frac{kT}{6\pi\eta_w R_h}\right]\left\{\frac{1-[(R_h+R_o)/L]^2}{1+[(R_h+R_o)/L]^2}\right\}. \quad (9)$$

In generating curves from Eq. 9, values of k, T, η_w, R_o, and L are given in Table 1, and values of R_h for each protein are given in Table 2.The first bracketed term accounts for the Stokes-Einstein relationship, and the second bracketed term accounts for steric hindrance, as in Eq. 3. Steric hindrance appears to moderately attenuate the diffusion coefficients returned from the Constant D model (51–87%; *open circles*) or the Variable D model at long diffusional times (55–91%; *solid circles*). However, steric hindrance alone does not explain the much lower initial diffusion rates, implying that additional factors, as considered in the Variable D model above, come into play.

Fig. 4 *B* recasts Fig. 4 *A* by showing the ratio of the apparent diffusion coefficients to their values in bulk water (D_w) plotted against the hydrodynamic radius, R_h. Both D/D_w (Constant D model, *open circles*) and $D(r,t)/D_w$ evaluated at long times, i.e., $t/a^2 = 0.35$ (Variable D model, *solid circles*) decline linearly with R_h, as would be expected of a molecular sieve of impenetrable rods (discussed in more detail later).

DISCUSSION

Glycolytic and other cytosolic enzymes that serve anaerobic metabolic pathways are abundant in fast skeletal muscle fibers. At the high concentrations found (3,4,41) and references therein), it is expected from thermodynamic and other considerations that multiprotein complexes would form (31,32,42). This study measures the diffusion coefficients of nine representative metabolic enzymes to assess whether their values are very low and not significantly different from one another, thereby supporting the existence of large complexes, or whether their values are higher and scale inversely with their hydrodynamic radii, thereby supporting a dissociated state. The results for seven glycolytic (Fig. 2) and two glycogenolytic (Fig. 3) enzymes show features of both scenarios, namely, an initial lag in diffusion with correspondingly low values of $D(r,t)$, suggesting that the enzymes exist as complexes in vivo, but with higher values at later times and a dependency of $D(r,t)$ on oligomeric size, suggesting that the enzymes dissociate upon immersion of the skinned fiber in relaxing solution. Adenylate kinase and parvalbumin do not appear to be complexed, since the initial lag in diffusion is absent (Fig. 3). The ability of the Variable D model to more accurately reproduce the general features of the diffusion time course for each species supports this interpretation.

Origin of diffusion lag

We considered several alternative mechanisms that could generate a diffusion lag, including one in which a barrier (a layer of oil or polarized water) is present initially but dissipates with time. As mentioned earlier, it is unlikely that a stagnant layer of oil or polarized water could be sustained because of continuous stirring, a conclusion bolstered by results of the three-drop experiment (Table S2). In this experiment, significantly lower values of D_{t/a^2} were associated with the initial transfer of the fiber from oil to solution compared to the values associated with reequilibration of the fiber in oil and a second transfer to solution. A mechanism in which an envelope of oil or polarized (highly structured) water reforms and dissipates during the second reequilibration in oil and transfer to solution would be expected to produce similar coefficients, which was not the case. Substantial experimental evidence suggests that an aqueous phase resembling bulk water constitutes nearly all (>96%) of the total water in the muscle (43–46), although this conclusion is somewhat controversial (38,47). The putative remainder (<4%) strongly absorbed to cytomatrix proteins or highly structured is probably of insufficient quantity to appreciably retard diffusion of cytosolic proteins.

We suggest a more realistic possibility, that native complexes of glycolytic and glycogenolytic enzymes break apart as the pool of monomers in equilibrium with the complexes becomes rapidly depleted. Cytosolic protein crowding and dissociation of such complexes, whose constituent molecules occupy a large volume fraction of the I-band in particular, provide a reasonable mechanism for the dependency of $D(r,t)$ on total protein density, reflected in the diffusional time lag. With dissociation of enzymes at reduced protein density, lattice hindrance and binding become the primary contributors to the reduced diffusivity at a later time. This interpretation is supported by the better fit of the Variable D model to the data (Figs. 2 and 3). Table 1 lists the average optimized parameter values with standard deviations for the 20 best fits to the data. The relatively narrow range of optimized values for parameters specifying the ease of vacancy-hydrodynamic volume product (φ), lattice binding equilibrium (k_b), and half dissociation activity density (ρ_{50}) indicates that these parameters are more sensitively prescribed by the data presented here than is the maximal complex dissociation equilibrium constant (k_s^{max}), which has a relatively wide range of optimized values. Despite these differences in parameter sensitivity, the

overall goodness of fit achieved with the optimized parameter values lends credence to the Variable D model.

There is considerable circumstantial evidence that glycolytic enzymes form complexes (7–9,17,47), a notion consistent with the results of the Variable D model. Further, concentration measurements of the glycolytic enzymes in related experiments with rabbit psoas muscle fibers (3) indicate a 1:2 stoichiometric relationship of enzymes in the hexose/triose pathways. Together, these results suggest the presence of a complex that effectively channels substrate intermediates through a potential bottleneck at the hexose/triose interface, enabling two pyruvate molecules to synchronously exit the complex for every glucose molecule that enters.

Diffusion after the lag

Our results show that the diffusion constants evaluated at a later time (after most of the putative complexes break up) scale roughly with molecular weight and hydrodynamic radius (Fig. 4). Obstructions and tortuosity appear to contribute most to the reduced diffusivity. Linear regressions through the data points intercept the hydrodynamic radius axis at 7.4 nm for the Constant D model and 7.3 nm for the Variable D model, indicating the radius at which particles are unable to diffuse out of the myofibril. The intercept from these models (and Eq. 9; see Fig. 4 *B*, *legend*) corresponds to a mesh size ($2x$) of 14.6–14.8 nm, and at 2.2 μm sarcomere spacing this range of sizes is within the range of surface-to-surface distances between adjacent myofilaments (14 nm in the A-band and 20 nm in the I-band (Table 2)). Thus, in regions of the sarcomere where cytosolic proteins are localized or into which they may diffuse, the myofilament lattice is of the appropriate size to be the molecular sieve.

The linear regression for the Variable D model in Fig. 4 *B* (*solid circles* and *dashed line*) gives $D(r,t)/D_w = 0.36$ (at $r = a$ and t/a^2 approaching 0.35×10^{-8} s cm^{-2}) for a point particle in the matrix ($R_h = 0$), indicating that the diffusion coefficients of small particles, such as metal ions or metabolites, are also attenuated by tortuosity and molecular sieving. The 64% reduction in radial diffusion coefficients for small particles inferred in this study is greater than the 50% reduction in longitudinal diffusion coefficients reported elsewhere for ionic potassium, sodium, sulfate, and adenosine triphosphate in vertebrate skeletal muscle (43). The greater reduction in radial diffusion compared to longitudinal diffusion is surprising in light of additional obstructions to longitudinal diffusion (Z-, M-, and A-band structures, including myosin heads, C-proteins, and myosin light chain extensions), but the diffusion anisotropy is consistent with that observed in previous experimental (48,49) or theoretical (50) work, where diffusion of water or small molecules in skeletal muscle was reported to be slower in the radial than in the longitudinal direction.

Electrostatic interactions with other proteins also appear to contribute to the reduced diffusion coefficients, but to a lesser extent. Muscle glycolytic enzymes have been shown to bind to one another differentially and to actin filaments under physiological ionic conditions (29,51–54). Indeed, the Variable D model returns an average apparent binding constant, k_b, of 0.96, implying that an appreciable fraction ($\sim 0.49 = k_b/(1 + k_b)$ (21)) of the diffusible molecules is bound to the cytomatrix at any moment. The corresponding binding free energy, ΔG, of that bound fraction is 0.10 kJ mol^{-1} ($= -RT \ln k_b$ (21)).

Relationship to other studies

In complementary studies, Kraft and colleagues (55) used confocal microscopy to examine the influx of exogenous fluorescently labeled molecules into chemically skinned fibers from rabbit psoas. Equilibration times varied widely, from minutes (e.g., rhodamine (Rh)-labeled pyruvate kinase) to days (e.g., Rh-labeled NEM-myosin S1), with large differences in equilibration times attributed to a combination of myofilament lattice obstruction and binding to fixed structures. Although diffusion coefficients were, for technical reasons, not reported and therefore cannot be compared with those listed in Table 2, the timescales over which equilibration occurs appear to be in rough accord. For example, the 8- to 10-min equilibration time for exogenous Rh-pyruvate kinase (Fig. 3 of Kraft et al. (55)) is comparable to that of endogenous pyruvate kinase (Fig. 2, *F* and *H*).

Other studies have been conducted in a variety of cell types to assess the relative contributions of geometrical hindrance and of binding to cytomatrix proteins to the reduced diffusion coefficient of cytosolic proteins. In most cases, a combined effect of interference by other proteins in the cytosol and obstruction by fixed cytomatrix proteins accounts for most of the reduced diffusivity (21). In some cases, steric hindrance by the myofilament lattice alone appears to account for most of the reduced diffusivity, as, for example, the 10-fold decrease in diffusion coefficients of native myoglobin observed in intact frog (56) and rat (57) skeletal muscle fibers. In a follow-up study, Papadopoulos and colleagues (58) found that the molecular sieve property of the filament lattice progressively reduced the diffusion coefficients of the microinjected proteins in relation to their size, from ~1/6 (cytochrome *c*, 3.1 nm diameter) to ~1/63 (ferritin, 12.2 nm) of their value predicted in water, with the exact values depending on fiber type. The marked dependence of the diffusion coefficients on protein size was explained by the high viscosity of the myoplasm (due to the high soluble protein content) and steric hindrance (due to the presence of the myofilament lattice and other structural barriers). These results also indicate that about half the number of otherwise freely diffusible cytosolic proteins are bound to the cytomatrix at any given moment, with a surprisingly low average binding free energy (0.10 kJ

mol^{-1}) compared to those reported for other cell systems (21).

The low initial diffusivity and apparent dependence of the diffusion coefficients on total cytosolic protein concentration reported here in skinned fibers are consistent with the low diffusion coefficient values reported elsewhere in intact muscle fibers (noted above). However, only at the outset of diffusion, when protein associations are most pronounced and soluble protein content (and cytosol viscosity) is highest, would one expect diffusion coefficients to be as low as those observed in intact cells, i.e., a tenth or less of that measured for a particular species in water. Only after the putative associations disperse and soluble protein content diminishes do individual molecules diffuse out more rapidly and in an ever-changing temporal pattern according to Eqs. 1, 2, and 7.

As noted previously, there is considerable circumstantial evidence, backed by strong thermodynamic arguments (31,59), that glycolytic enzymes form complexes in the crowded cytosol of muscle cells (3,9,32,47,60,61). Recently, Brownian dynamic simulations have contributed to our understanding of the critical residues involved in the molecular interactions between specific glycolytic enzyme species (53) and between glycolytic enzymes and globular and filamentous actin (53,54,62,63) in the formation of the putative complexes. The troponin-tropomyosin complex may help seed cytosolic protein aggregates, since binding of glycolytic enzymes to actin is enhanced by the presence of thin-filament-associated proteins (29). Aldolase, in particular, displays a preferred binding to thin filaments at troponinlike intervals of 38 nm, similar to that of the so-called I-bridges that appear to bridge neighboring thin filaments in the I-band of fast skeletal muscle fibers (64). These studies and others led to the hypothesis that glycolytic complexes (or aggregates of enzymes) are involved in metabolic channeling of substrate intermediates, which could elevate the rate of ATP production beyond that achievable with dissociated enzymes (8,9,53). The complexes are probably unstable, since the binding forces appear to be relatively weak in light of their propensity to dissociate readily. The complexes must also be dynamic, since myosin heads slide easily into the I-band region during rapid shortening.

Similar thermodynamic arguments can be applied to glycogenolytic enzymes associated with glycogen particles, which together with the bound enzymes constitute the glycosome (59). Glycogen phosphorylase and phosphoglucose mutase, which bind to these extramyofibrillar particles, catalyze the breakdown of glycogen (65). We speculate that exposing the skinned segment to the bathing medium releases the enzymes from the glycogen particle, resulting in the slight lag in efflux (Fig. 3, *C* and *D*).

CONCLUSION

The results of this study are consistent with a primary mechanism of reduced protein diffusivity in which the compact myofilament lattice of the I-band restricts diffusion of complexes (or aggregates) of glycolytic enzymes located within that region, whereas the more compact lattice of the A-band (containing the myosin heads and other structural proteins) excludes the complexes (3). When muscle fibers are stripped of their sarcolemma and immersed in a physiological bathing solution, the complexes break up into monomers normally in equilibrium with the complexed enzymes, and the monomers diffuse out of the fiber into the bath. The diffusion coefficient of individual species is reduced significantly at the outset due to associations between proteins when the cytosol is crowded, but as the total protein content in the cytosol falls, the complexes dissociate and the individual proteins diffuse out at rates that are inversely proportional to their size. Collisions with cytomatrix proteins (molecular sieving and tortuosity) hinder diffusion, eventually rendering diffusion coefficients that are a tenth to a fifth of those in bulk water.

SUPPORTING MATERIAL

Three tables, five figures, and Supporting Methods are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)00069-1.

We are grateful to Elizabeth Wegner, Janet Hurley, and Bill Barnes for their technical help. D.W.M. and B.E.C. thank Professor Jim Bassingthwaighte for his support and encouragement.

This work was supported by grants from the National Institutes of Health (DK33833, R01 AR38980, and R01 HL680343 to D.W.M. and T32 EB001650 and NIGMS P50-GM094503-02 to B.E.C.) and the National Science Foundation (IOS-0718417 to J.O.V. and CBET 1133260 to B.E.C.).

D.W.M. dedicates this article to the memory of Halcyon Delight and Orlo Harrison Maughan.

REFERENCES

1. Kawai, M., and R. Candau. 2009. Muscle contraction and supplying ATP to muscle cells. *In* Handbook of Exercise Physiology: From a Cellular to an Integrative Approach. P. Connes, O. Hue, and S. Perrey, editors. IOS Press, Amsterdam, pp. 1–22.
2. Booth, N. E., S. Myhill, and J. McLaren-Howard. 2012. Mitochondrial dysfunction and the pathophysiology of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *Int. J. Clin. Exp. Med.* 5:208–220.
3. Maughan, D. W., J. A. Henkin, and J. O. Vigoreaux. 2005. Concentrations of glycolytic enzymes and other cytosolic proteins in the diffusible fraction of a vertebrate muscle proteome. *Mol. Cell. Proteomics.* 4:1541–1549.
4. Pette, D. 1975. Some aspects of supramolecular organization of glycogenolytic and glycolytic enzymes in muscle. *Acta Histochem. Suppl.* 14:47–68.
5. Arnold, H., and D. Pette. 1968. Binding of glycolytic enzymes to structure proteins of the muscle. *Eur. J. Biochem.* 6:163–171.
6. Clarke, F. M., and C. J. Masters. 1973. Letter: Multi-enzyme aggregates: new evidence for an association of glycolytic components. *Biochim. Biophys. Acta.* 327:223–226.
7. Clarke, F. M., and C. J. Masters. 1975. On the association of glycolytic enzymes with structural proteins of skeletal muscle. *Biochim. Biophys. Acta.* 381:37–46.

8. Ottaway, J. H., and J. Mowbray. 1977. The role of compartmentation in the control of glycolysis. *Curr. Top. Cell. Regul.* 12:107–208.
9. Kurganov, B. I., N. P. Sugrobova, and L. S. Mil'man. 1985. Supramolecular organization of glycolytic enzymes. *J. Theor. Biol.* 116:509–526.
10. Maughan, D. W., and C. Lord. 1988. Protein diffusivities in skinned frog skeletal muscles. *Adv. Exp. Med. Biol.* 226:75–84.
11. Knull, H. R., and J. L. Walsh. 1992. Association of glycolytic enzymes with the cytoskeleton. *Curr. Top. Cell. Regul.* 33:15–30.
12. Maughan, D. W., and J. O. Vigoreaux. 1999. An integrated view of insect flight muscle: genes, motor molecules, and motion. *News Physiol. Sci.* 14:87–92.
13. Maughan, D. W., and R. E. Godt. 1999. Parvalbumin concentration and diffusion coefficient in frog myoplasm. *J. Muscle Res. Cell Motil.* 20:199–209.
14. Godt, R. E., and D. W. Maughan. 1981. Influence of osmotic compression on calcium activation and tension in skinned muscle fibers of the rabbit. *Pflugers Arch.* 391:334–337.
15. Kawai, M., J. S. Wray, and Y. Zhao. 1993. The effect of lattice spacing change on cross-bridge kinetics in chemically skinned rabbit psoas muscle fibers. I. Proportionality between the lattice spacing and the fiber width. *Biophys. J.* 64:187–196.
16. Crank, J. 1970. Mathematics of Diffusion. Oxford University Press, London.
17. Dölken, G., E. Leisner, and D. Pette. 1975. Immunofluorescent localization of glycogenolytic and glycolytic enzyme proteins and of malate dehydrogenase isozymes in cross-striated skeletal muscle and heart of the rabbit. *Histochemistry.* 43:113–121.
18. Wallimann, T., T. Schnyder, …, A. F. Quest. 1989. Subcellular compartmentation of creatine kinase isoenzymes, regulation of CK and octameric structure of mitochondrial CK: important aspects of the phosphoryl-creatine circuit. *Prog. Clin. Biol. Res.* 315:159–176.
19. Godt, R. E., and D. W. Maughan. 1988. On the composition of the cytosol of relaxed skeletal muscle of the frog. *Am. J. Physiol.* 254:C591–C604.
20. Blum, J. J., G. Lawler, …, I. Shin. 1989. Effect of cytoskeletal geometry on intracellular diffusion. *Biophys. J.* 56:995–1005.
21. Gershon, N. D., K. R. Porter, and B. L. Trus. 1985. The cytoplasmic matrix: its volume and surface area and the diffusion of molecules through it. *Proc. Natl. Acad. Sci. USA.* 82:5030–5034.
22. O'Leary, T. J. 1987. Concentration dependence of protein diffusion. *Biophys. J.* 52:137–139.
23. Rorschach, H. E., D. C. Chang, …, B. L. Nichols. 1973. The diffusion of water in striated muscle. *Ann. N. Y. Acad. Sci.* 204:445–452.
24. Granzier, H. L., and K. Wang. 1993. Passive tension and stiffness of vertebrate skeletal and insect flight muscles: the contribution of weak cross-bridges and elastic filaments. *Biophys. J.* 65:2141–2159.
25. McElhinny, A. S., S. T. Kazmierski, …, C. C. Gregorio. 2003. Nebulin: the nebulous, multifunctional giant of striated muscle. *Trends Cardiovasc. Med.* 13:195–201.
26. Muramatsu, N., and A. P. Minton. 1988. Tracer diffusion of globular proteins in concentrated protein solutions. *Proc. Natl. Acad. Sci. USA.* 85:2984–2988.
27. Clarke, F. M., and D. J. Morton. 1976. Aldolase binding to actin-containing filaments. Formation of paracrystals. *Biochem. J.* 159:797–798.
28. Stewart, M., D. J. Morton, and F. M. Clarke. 1980. Interaction of aldolase with actin-containing filaments. Structural studies. *Biochem. J.* 186:99–104.
29. Walsh, T. P., D. J. Winzor, …, D. J. Morton. 1980. Binding of aldolase to actin-containing filaments. Evidence of interaction with the regulatory proteins of skeletal muscle. *Biochem. J.* 186:89–98.
30. Stephan, P., F. Clarke, and D. Morton. 1986. The indirect binding of triose-phosphate isomerase to myofibrils to form a glycolytic enzyme mini-complex. *Biochim. Biophys. Acta.* 873:127–135.
31. Minton, A. P., and J. Wilf. 1981. Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry.* 20:4821–4826.
32. Ellis, R. J., and A. P. Minton. 2006. Protein aggregation in crowded environments. *Biol. Chem.* 387:485–497.
33. Kraft, T., T. Hornemann, …, T. Wallimann. 2000. Coupling of creatine kinase to glycolytic enzymes at the sarcomeric I-band of skeletal muscle: a biochemical study in situ. *J. Muscle Res. Cell Motil.* 21:691–703.
34. Meyer, F., L. M. Heilmeyer, Jr., …, E. H. Fischer. 1970. Control of phosphorylase activity in a muscle glycogen particle. I. Isolation and characterization of the protein-glycogen complex. *J. Biol. Chem.* 245:6642–6648.
35. Young, M. E., P. A. Carroad, and R. L. Bell. 1980. Estimation of diffusion coefficients of proteins. *Biotechnol. Bioeng.* 22:947–955.
36. Roos, A., and W. F. Boron. 1981. Intracellular pH. *Physiol. Rev.* 61:296–434.
37. Holmes, M. J., N. G. Parker, and M. J. W. Povey. 2011. Temperature dependence of bulk viscosity in water using acoustic spectroscopy. *J. Phys. Conf. Ser.* 269:012011.
38. Zheng, J. M., and G. H. Pollack. 2003. Long-range forces extending from polymer-gel surfaces. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 68:031408.
39. Einstein, A. 1905. About the molecular kinetic theory of thermal motion of particles suspended in liquids at rest. *Ann. Phys.* 322:549–560.
40. Sutherland, W. 1905. LXXV. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin. *Philos. Mag.* 9:781–785.
41. Scopes, R. K. 1973. Studies with a reconstituted muscle glycolytic system. The rate and extent of creatine phosphorylation by anaerobic glycolysis. *Biochem. J.* 134:197–208.
42. Parsegian, V. A., R. P. Rand, and D. C. Rau. 2000. Osmotic stress, crowding, preferential hydration, and binding: A comparison of perspectives. *Proc. Natl. Acad. Sci. USA.* 97:3987–3992.
43. Kushmerick, M. J., and R. J. Podolsky. 1969. Ionic mobility in muscle cells. *Science.* 166:1297–1298.
44. Maughan, D. W., and R. E. Godt. 2001. Protein osmotic pressure and the state of water in frog myoplasm. *Biophys. J.* 80:435–442.
45. Hill, A. V. 1930. The state of water in muscle and blood and the osmotic behaviour of muscle. *Proc. R. Soc. Lond. B Biol. Sci.* 106:477–505.
46. Rorschach, H. E., D. W. Bearden, …, R. M. Nicklow. 1988. Quasi-elastic scattering studies of water diffusion. *Scan. Microsc.* 2:2043–2049.
47. Fulton, A. B. 1982. How crowded is the cytoplasm? *Cell.* 30:345–347.
48. Cleveland, G. G., D. C. Chang, …, H. E. Rorschach. 1976. Nuclear magnetic resonance measurement of skeletal muscle: anisotrophy of the diffusion coefficient of the intracellular water. *Biophys. J.* 16:1043–1053.
49. Engel, J., M. Fechner, …, A. Stier. 1994. Anisotropic propagation of Ca^{2+} waves in isolated cardiomyocytes. *Biophys. J.* 66:1756–1762.
50. Shorten, P. R., and J. Sneyd. 2009. A mathematical analysis of obstructed diffusion within skeletal muscle. *Biophys. J.* 96:4764–4778.
51. Walsh, T. P., C. J. Masters, …, F. M. Clarke. 1981. The reversible binding of glycolytic enzymes in ovine skeletal muscle in response to tetanic stimulation. *Biochim. Biophys. Acta.* 675:29–39.
52. Gerlach, G., and H. W. Hofer. 1986. Interaction of immobilized phosphofructokinase with soluble muscle proteins. *Biochim. Biophys. Acta.* 881:398–404.
53. Ouporov, I. V., H. R. Knull, …, K. A. Thomasson. 2001. Brownian dynamics simulations of aldolase binding glyceraldehyde 3-phosphate dehydrogenase and the possibility of substrate channeling. *Biophys. J.* 80:2527–2535.
54. Lowe, S. L., C. Adrian, …, K. A. Thomasson. 2003. Brownian dynamics simulations of glycolytic enzyme subsets with F-actin. *Biopolymers.* 70:456–470.

55. Kraft, T., M. Messerli, …, B. Brenner. 1995. Equilibration and exchange of fluorescently labeled molecules in skinned skeletal muscle fibers visualized by confocal microscopy. *Biophys. J.* 69:1246–1258.

56. Baylor, S. M., and P. C. Pape. 1988. Measurement of myoglobin diffusivity in the myoplasm of frog skeletal muscle fibres. *J. Physiol.* 406:247–275.

57. Jürgens, K. D., T. Peters, and G. Gros. 1994. Diffusivity of myoglobin in intact skeletal muscle cells. *Proc. Natl. Acad. Sci. USA.* 91:3829–3833.

58. Papadopoulos, S., K. D. Jürgens, and G. Gros. 2000. Protein diffusion in living skeletal muscle fibers: dependence on protein size, fiber type, and contraction. *Biophys. J.* 79:2084–2094.

59. Chebotareva, N. A. 2007. Effect of molecular crowding on the enzymes of glycogenolysis. *Biochemistry* (*Mosc*). 72:1478–1490.

60. Maughan, D., and E. Wegner. 1989. On the organization and diffusion of glycolytic enzymes in skeletal muscle. *Prog. Clin. Biol. Res.* 315:137–147.

61. Zimmerman, S. B., and A. P. Minton. 1993. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* 22:27–65.

62. Ouporov, I. V., H. R. Knull, …, K. A. Thomasson. 2001. Interactions of glyceraldehyde-3-phosphate dehydrogenase with G- and F-actin predicted by Brownian dynamics. *J. Mol. Recognit.* 14:29–41.

63. Lowe, S. L., D. M. Atkinson, …, K. A. Thomasson. 2002. Brownian dynamics of interactions between aldolase mutants and F-actin. *J. Mol. Recognit.* 15:423–431.

64. Trombitás, K., P. H. Baatsen, and G. H. Pollack. 1988. I-bands of striated muscle contain lateral struts. *J. Ultrastruct. Mol. Struct. Res.* 100:13–30.

65. DiMauro, S., W. Trojaborg, …, L. P. Rowland. 1971. Binding of enzymes of glycogen metabolism to glycogen in skeletal muscle. *Arch. Biochem. Biophys.* 144:413–422.

66. Huxley, H. E., A. Stewart, …, T. Irving. 1994. X-ray diffraction measurements of the extensibility of actin and myosin filaments in contracting muscle. *Biophys. J.* 67:2411–2421.

Diffusion Coefficients of Endogenous Cytosolic Proteins from Rabbit Skinned Muscle Fibers

Brian E. Carlson,[†] Jim O. Vigoreaux,[‡§] and David W. Maughan[‡]*

[†]Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, Michigan; [‡]Department of Molecular Physiology and Biophysics, Health Science Research Facility, University of Vermont College of Medicine, Burlington, Vermont; and [§]Department of Biology, University of Vermont, Burlington, Vermont

SUPPORTING MATERIAL

A. Table of Symbols

Symbol	Description	Units
C	Protein concentration	µM
D	Diffusion coefficient of protein	$cm^2\ s^{-1}$
r	Radial position in fiber	µm
t	Elapsed time of diffusion	s
a	Outer fiber radius	µm
C_0	Initial concentration of protein	µM
D_w	Diffusion coefficient of protein in aqueous solution	$cm^2\ s^{-1}$
D_h	Diffusion coefficient of protein with steric hindrance	$cm^2\ s^{-1}$
R_h	Hydrodynamic radius of protein	nm
R_o	Average myofilament radius	nm
L	Average half center to center spacing of myofilaments	nm
D_c	Diffusion coefficient of protein in a crowded cytosol	$cm^2\ s^{-1}$
γ	Ease of vacancy of protein	unitless
V_h	Hydrodynamic volume of protein	$dL\ g^{-1}$
V_h^b	Average hydrodynamic volume of background proteins in cytosol	$dL\ g^{-1}$
ρ	Total protein density in the cytosol	$g\ dL^{-1}$
D_b	Diffusion coefficient of protein considering binding to cytomatrix	$cm^2\ s^{-1}$
k_b	Average apparent binding constant of proteins to cytomatrix	unitless
k_s	Average apparent binding constant of supramolecular protein complex	unitless
k_s^{max}	Maximal apparent binding constant of supramolecular protein complex	unitless
ρ_{50}	Total protein density at half maximal binding of the protein complex	$g\ dL^{-1}$
n_s	Supramolecular protein complex cooperativity parameter	unitless
M	Amount of protein diffused out of fiber	µmoles
M_∞	Total amount of diffusible protein in the fiber	µmoles
t/a^2	Diffusion time in fiber scaled by a cross-sectional area parameter	s/cm^2
$\overline{a}$	Average fiber radius over all experimental fibers	µm
φ	Ease of vacancy – hydrodynamic volume product	$dL\ g^{-1}$
k	Boltzmann's constant	$J\ °K^{-1}$
T	Experimental temperature	°K
η_w	Viscosity of water at 7°C	cP
ΔG	Free energy of protein binding to cytomatrix	$kJ\ mol^{-1}$

Table S1. Complete table of symbols, description and units used.

B. Test for Oil Dissipation Artifact

A freshly skinned fiber was transferred from oil to a drop of relaxing solution and, after two seconds, returned to oil. During this brief period in solution 3-9% (0.03-0.09: column 5) of the diffusible protein left the fiber, the exact amount depending on the protein species and fiber diameter. Following a 60 s equilibration period in oil, the fiber was then transferred to a second drop of relaxing solution for two seconds. That is, the initial

conditions were re-established and the procedure repeated, during which the fiber was depleted of a slightly greater amount of diffusible protein (0.04-0.12: column 6). Finally, the fiber was transferred rapidly through oil to a third drop of relaxing solution for a final washout period of ten minutes. Values of $D_{t/a2=0.049}$ (i.e., D of the Constant D model evaluated at $t/a^2=0.049 \times 10^8$) were calculated for a representative set of proteins by applying the Constant D model to drops 1-3 and to drops 2-3, respectively, as explained in the table caption. $D_{t/a2=0.049}$ derived from the second transfer (column 3) exceeded those from the first (column 2) by a factor of 2.8±0.6 (column 4). The fact that a significantly higher value of $D_{t/a2=0.049}$ was observed after a reduction in protein density is consistent with the breakup and dissipation of a protein complex and not with a diffusional delay due to dissipation of a layer of oil or polarized water (see main text)..

1	2	3	4	5	6
	D^1	D^2	D^2/D^1	$M^1/(M^1+M^2+M^3)$	$M^2/(M^2+M^3)$
Phosphoglucose isomerase	1.2 (1.0)	2.2 (1.9)	2.7 (2.0)	0.03 (0.02)	0.04 (0.03)
Triosephosphate isomerase	1.9 (1.3)	5.0 (4.7)	2.5 (0.7)	0.04 (0.02)	0.07 (0.05)
Pyruvate kinase	4.8 (5.8)	6.9 (3.9)	3.5 (2.3)	0.07 (0.05)	0.09 (0.04)
Enolase	2.1 (1.6)	8.5 (9.5)	3.8 (2.3)	0.04 (0.02)	0.09 (0.07)
Glyceraldehyde-3-phosphate dehydrogenase	4.1 (3.5)	10.0 (8.6)	3.5 (1.7)	0.06 (0.04)	0.11 (0.07)
Phosphoglycerate mutase	1.3 (1.4)	2.3 (2.5)	2.4 (2.2)	0.03 (0.03)	0.04 (0.03)
Phosphoglucose mutase	2.9 (2.6)	5.3 (5.7)	3.0 (2.2)	0.05 (0.04)	0.07 (0.06)
Adenylate kinase	1.8 (1.4)	3.5 (3.4)	2.0 (1.3)	0.04 (0.02)	0.06 (0.04)
Parvalbumin	5.3 (2.8)	12.0 (8.9)	2.2 (0.8)	0.09 (0.03)	0.12 (0.06)

Table S2. Comparison of diffusion coefficient $D_{t/a2}$ calculated before and after re-equilibration period in oil. Column 2 and 3: Radial diffusion coefficients ($D = D_{t/a2}$): $\times 10^{-8}$ cm^2 s^{-1} (means ± SD), 4 fibers, 7°C. Protocol: Skinned fiber transferred sequentially from oil to drop 1 (2 s), oil (60 s re-equilibration period), drop 2 (2 s), oil (<1 s), drop 3 (600 s), oil, fiber removed. Drop and fiber samples analyzed by SDS-PAGE. Diffusion coefficients calculated from equation [1] and [2] in text (Constant D model). Amount of each protein in each drop indicated by letter M; drop number indicated by superscript. D^1 calculated using the fraction $M^1/(M^1+M^2+M^3)$; D^2 calculated using the fraction $M^2/(M^2+M^3)$. Average radius of fiber subset, 25.5±3.7 μm. Phosphofructose kinase and glycogen phosphorylase not analyzed.

C. Estimation of Hydrodynamic Volume

Equation 4 and 7 in our simulation requires values for the hydrodynamic volume of each protein (as an estimate of the protein exclusion volume), as well as the average hydrodynamic volume of the background proteins in the cytosol. We estimated the hydrodynamic volume of each protein, V_h, for all 15 proteins tracked in this simulation, by means of a linear fit to the hydrodynamic volume of human and earthworm hemoglobin as a function of molecular weight from O'Leary (1). This method of estimating V_h was used because experimental values of the hydrodynamic volume are not available in the literature for these proteins. To incorporate the dependence of hydrodynamic volume on molecular weight a linear fit of hydrodynamic volume as a function of molecular weight is given by expression 6 in Figure 1S below. Of the 15 proteins in this study all but 6 of the proteins (TPI – 53.25 kDa, PGM – 57 kDa, PAR – 11.93 kDa, ADK – 21.64 kDa, PGO – 61.43 kDa and PGK – 36.5 kDa) fall in the molecular weight range spanned by human (65 kDa) and earthworm (3700 kDa) hemoglobin and of those 6 proteins, 3 are within 15 kDa of human hemoglobin (TPI,

Protein	Mo (Da)	Dw (cm2/s)		Rh (nm)			Vh (nm^3)			Vh (dL/g)		
		Young et al.	Linear Fit - SE	Young et al	Stokes-Einstein	Linear Fit	Young et al.	Stokes-Einstein	Linear fit	Young et al	Stokes-Einstein	Linear fit
PGI	125230	3.21899E-07	3.5127E-07	3.309	4.394	4.026	151.807	355.341	273.442	0.00730	0.01709	0.01315
PFK	340288	2.30679E-07	2.4842E-07	4.618	6.131	5.694	412.505	965.570	773.112	0.00730	0.01709	0.01368
TPI	53250	4.28071E-07	4.6927E-07	2.489	3.304	3.014	64.551	151.097	114.696	0.00730	0.01709	0.01297
GPD	144700	3.06761E-07	3.3435E-07	3.473	4.611	4.230	175.409	410.587	317.113	0.00730	0.01709	0.01320
ENO	93874	3.54357E-07	3.8746E-07	3.006	3.991	3.650	113.796	266.368	203.765	0.00730	0.01709	0.01307
PVK	231664	2.62222E-07	2.8427E-07	4.062	5.394	4.976	280.828	657.348	515.979	0.00730	0.01709	0.01341
PGM	57000	4.1847E-07	4.5863E-07	2.546	3.380	3.084	69.097	161.738	122.861	0.00730	0.01709	0.01298
PAR	11933	7.04757E-07	7.7462E-07	1.512	2.007	1.826	14.465	33.860	25.500	0.00730	0.01709	0.01287
ADK	21638	5.77941E-07	6.3484E-07	1.843	2.447	2.228	26.230	61.398	46.325	0.00730	0.01709	0.01289
PGO	61427	4.08166E-07	4.4721E-07	2.610	3.465	3.163	74.463	174.300	132.515	0.00730	0.01709	0.01299
PPB	194316	2.78048E-07	3.0212E-07	3.831	5.087	4.682	235.554	551.373	429.811	0.00730	0.01709	0.01332
ADO	156844	2.9863E-07	3.2524E-07	3.567	4.736	4.349	190.130	445.046	344.510	0.00730	0.01709	0.01323
PGK	36500	4.85503E-07	5.3279E-07	2.194	2.913	2.655	44.246	103.569	78.367	0.00730	0.01709	0.01293
LDH	145732	3.06035E-07	3.3353E-07	3.481	4.622	4.241	176.660	413.516	319.436	0.00730	0.01709	0.01320
PCK	86224	3.64541E-07	3.9879E-07	2.922	3.880	3.547	104.523	244.661	186.889	0.00730	0.01709	0.01305
Human Hb	65000	4.00545E-07	4.3876E-07	2.660	3.531	3.224	78.794	184.438	140.319	0.00730	0.01709	0.01300
Eworm Hb	3700000	1.04125E-07	9.5714E-08	10.231	13.584	14.777	4485.221	10498.775	13517.104	0.00730	0.01709	0.02200

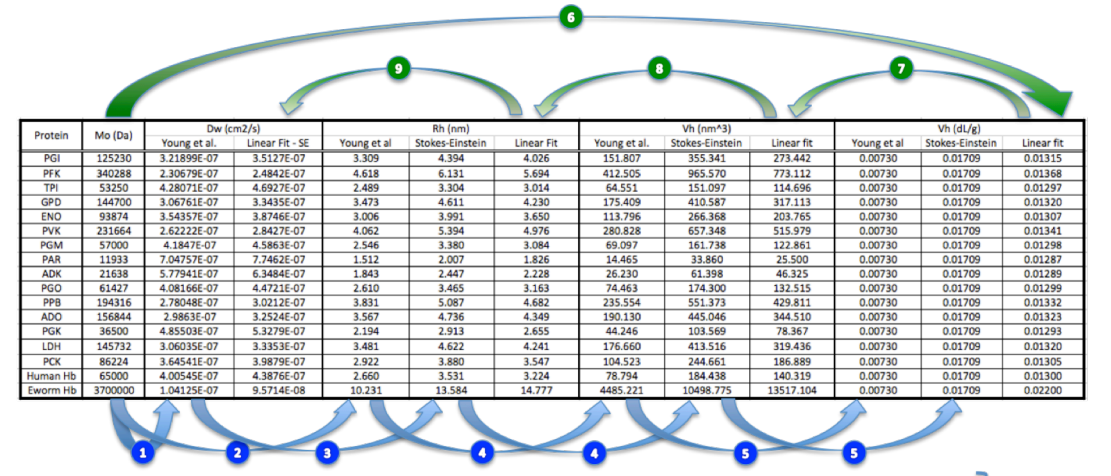

Figure S1. Comparison between several methods of calculating, diffusion of various proteins in an aqueous solution D_w, hydrodynamic radius R_h, and hydrodynamic volume V_h. Values of diffusion in water and hydrodynamic radius of each protein using linear estimation of V_h as a function of molecular weight (expression 6) is contrasted against that obtained from empirical expressions derived from Young et al. (expressions 1 and 2) using the Stokes-Einstein relation (expressions 3 and 9) to perform intermediate calculations.

PGM and PGO). A comparison of the estimated diffusion in water of each protein from Young et al. (2), shown in Expression 1, Figure 1S, with that obtained from the linear fit (Expressions 6-9, Figure 1S) shows between 7.5 to 10% higher rate of diffusion is predicted from the linear fit. In addition, the estimated hydrodynamic radius using the linear fit lies between that estimated directly from Young et al. (Expression 2, Figure 1S) and that obtained estimating D_w from Young et al. and then calculating R_h using the Stokes-Einstein expression (Expressions 1 and 3, Figure 1S). These observations suggest that the estimates of hydraulic volume used in our simulation fall within the range of the values that would be expected experimentally.

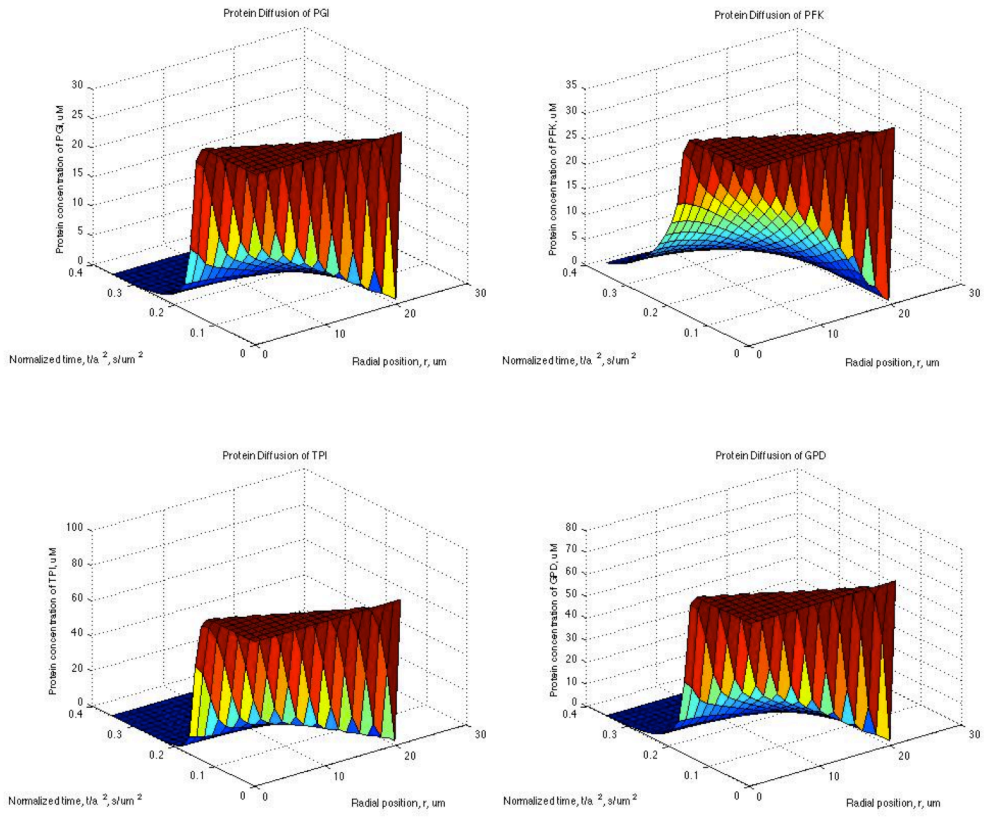

Figure S2. Surface plots of protein concentration across the fiber radius as a function of time for phosphoglucose isomerase (PGI), phosphofructose kinase (PFK), triosephosphate isomerase (TPI), and glyceraldehyde-3-phosphate dehydrogenase (GPD).

D. Surface plots of protein concentration across myofibril radius as a function time

In Figures 2S, 3S, 4S and 5S the simulation profiles of protein concentration as a function of radial position and time are shown for the optimized parameters of the Variable D model given in Table 1 in the text. It can be observed from these figures that parvalbumin and adenylate kinase (Figure 4S), which were not included in the

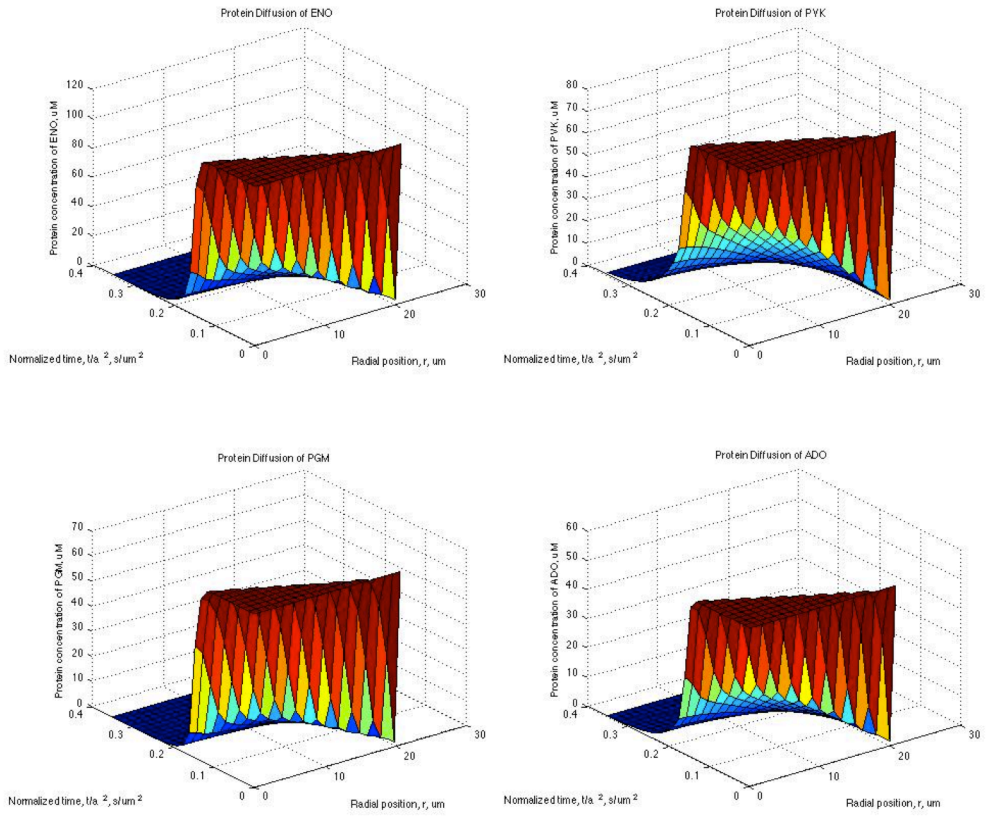

Figure S3. Surface plots of protein concentration across the fiber radius as a function of time for enolase (ENO), pyruvate kinase (PVK), phosphoglycerate mutase (PGM), and aldolase (ADO).

supramolecular protein complex, diffuse out of the myofibril rapidly and have nearly completely diffused out of the myofibril at t/a^2 of 0.1 s/μm^2. The remaining proteins take much longer to diffuse out, with concentrations approaching zero across the myofibril at $t/a^2 > 0.2$ s/μm^2. Phosphofructose kinase takes the longest to diffuse out of the myofibril, with near complete diffusion out of the myofibril at $t/a^2 \sim 0.4$ s/μm^2 (Figure 2S). In all proteins participating in the supramolecular complex the steep drop in the concentration that progresses along the radial direction from the outer fiber radius with time is a direct result of the complex dissociating as local total protein concentration decreases.

E. Simulation Model Code and Optimization Notes

All code and datasets used to simulate the best fits of the two different model versions to the experimental data are available upon request from the authors. Constant D and Variable D simulation codes are very similar but files in each folder differ even though they may be named the same. Therefore each simulation should be run out of its respective folder. All code is written in MATLAB and requires the Partial Differential Equation toolbox to run. Actual optimization code used to obtain these best fits is also

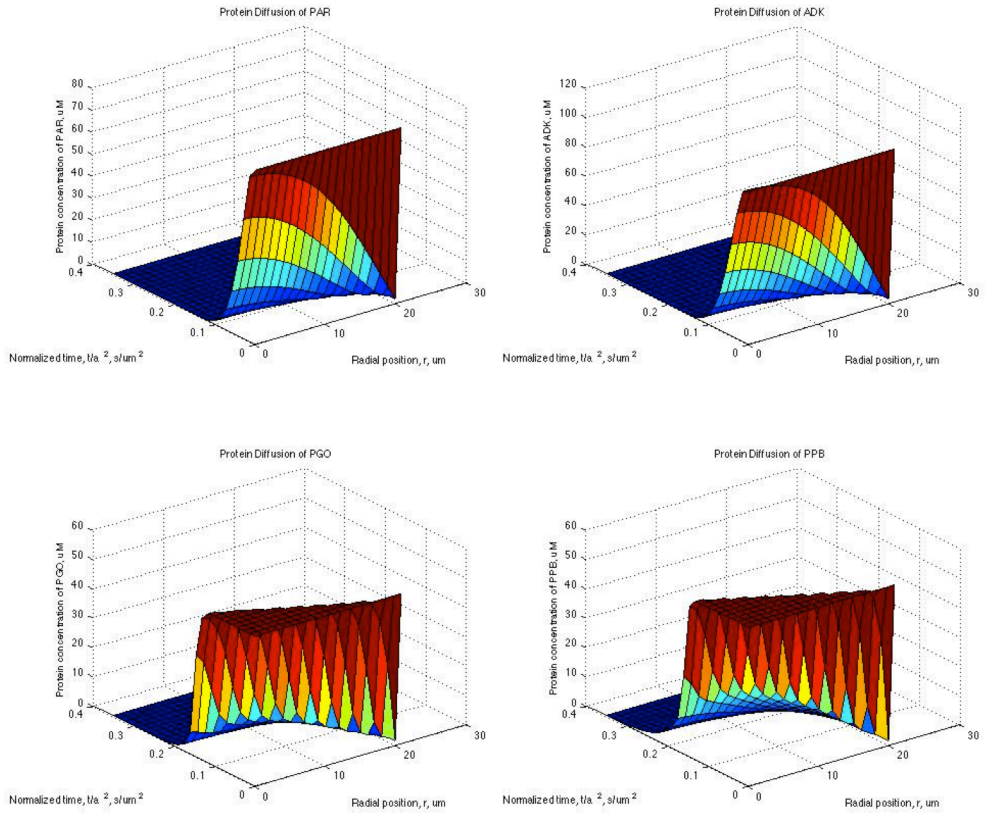

Figure S4. Surface plots of protein concentration across the fiber radius as a function of time for parvalbumin (PAR), adenylate kinase (ADK), phosphoglucose mutase (PGO), and glycogen phosphorylase (PPB).

available upon request and requires the additional Optimization toolbox to run locally on a single core and the Parallel Computing Toolbox to run the optimization in a parallel manner either locally on a multicore processor or on a computational cluster. These file are extensively documented and could be rewritten in other programming languages. In MATLAB help can be obtained for all scripts and functions by typing "help" followed by the script or function name. A brief one-line description of all the scripts and functions in a folder can be obtained by typing "help" followed by the folder name. This documentation and attached model code aims at facilitating model transparency and model replication.

There are three major differences from typical optimization methods, which are worth noting but are not covered in the original paper.

1) The optimization is performed in two stages. First a simulated annealing method explores the parameter space as defined by the upper and lower bounds on the parameters giving us a more global perspective on the minimization. This method randomly selects parameter values so is independent of the initial guess. The exact

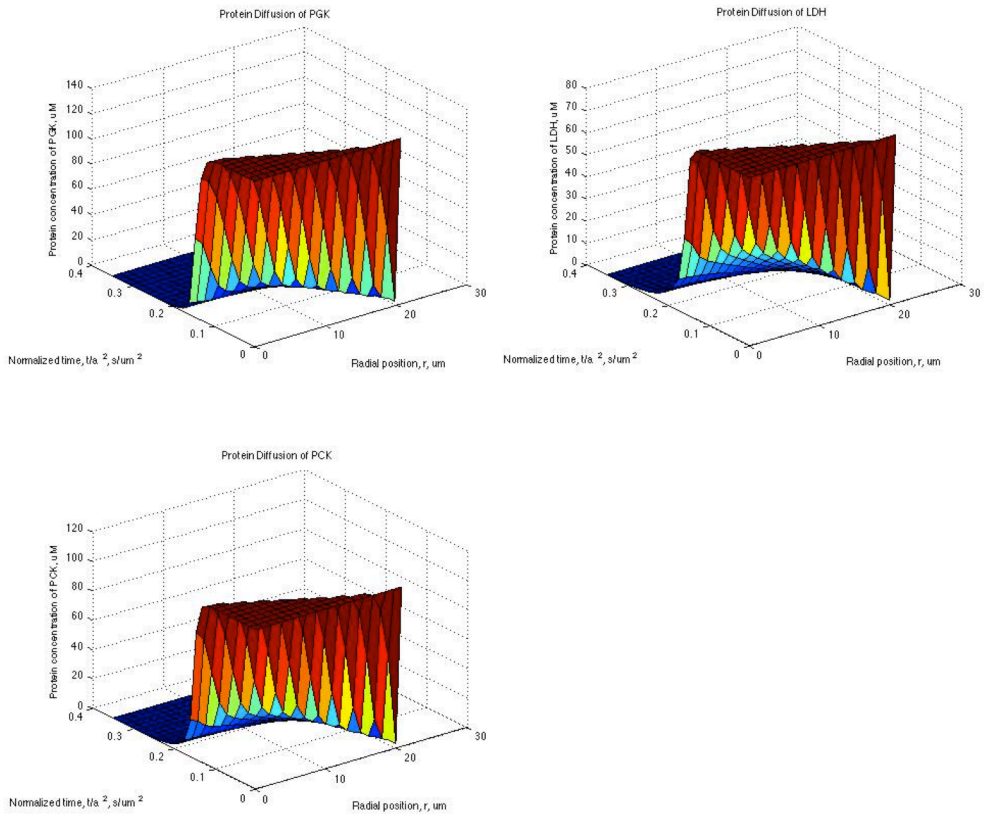

Figure S5. Surface plots of protein concentration across the fiber radius as a function of time for phosphoglycerate kinase (PGK), lactate dehydrogenase (LDH), and glycogen phosphocreatine kinase (PCK).

method is very similar to that developed by Boyan (3) which utilizes a Modified Lam annealing schedule and does not impose a neighborhood function that is dependent on the annealing temperature for each random selection of parameter values. Therefore this simulated annealing code is relatively model-independent and does not require a good initial guess; however, it does require the user to define the minimal bounded parameter space to search. Another modification made is that the user can specify the number of parameters between two iterations that are allowed to change. The exact parameters that do change between iterations is randomly selected and these parameters are allowed to change anywhere within their bounds. After the simulated annealing optimization returns a value we use *fmincon*, a gradient-based minimization function included in the MATLAB optimization toolbox, to find the local minima in the region identified by the global simulated annealing method. For more details on *fmincon* see MATLAB documentation (4).

2) The form of the least-squares minimization used to determine the residual error between simulation and data is not the traditional least-squares error in the y variable given by:

$$RE = \frac{\sqrt{\sum_{i=1}^{N}\left(y_{sim}\left(x_i\right)-y_i\right)^2}}{Ny_{\max}}$$

where RE is the residual error between the simulation evaluated at x_i, $y_{sim}(x_i)$, and y_i where x_i and y_i are the values of x and y for data point i, $y_{\max}$ is the maximal value of the y variable and N is the total number of data points. If this residual error was to be used the sigmoidal distribution of the data points would lead to an optimization that would tend to fit the large number of points at long t/a^2 ($t/a^2 > 0.10 \times 10^8$ s/cm^2) and would not try to minimize the error generated by the dramatic increase in diffusion around $t/a^2 \sim 0.05 \times 10^8$ s/cm^2. This type of data requires a residual error function that approximates the perpendicular distance of the data points to the simulation curve. This is given by:

$$RE = \frac{\sqrt{\sum_{i=1}^{N}\frac{\left(x_{sim}\left(y_i\right)-x_i\right)^2\left(y_{sim}\left(x_i\right)-y_i\right)^2}{\left(\frac{x_{sim}\left(y_i\right)-x_i}{x_{\max}}\right)^2+\left(\frac{y_{sim}\left(x_i\right)-y_i}{y_{\max}}\right)^2}}}{Nx_{\max}y_{\max}}$$

where $x_{sim}(y_i)$ is the simulation x value at y_i and $x_{\max}$ is the maximum value of the x variable. This approximate perpendicular distance residual error function resulted in much better fits to the sigmoidal shape of the data.

3) The optimization was run on either a multi-node remote cluster or a single/multi core local workstation. In the cluster configuration the computational jobs are distributed across the cores in an "embarrassingly parallel" manner, meaning that each job executes to completion without communicating with other jobs run at the same time. This is the simplest of cluster configurations and requires no message passing interface (MPI). However if no cluster is available to the user the optimization code (available on request) can be run on a multi core workstation using the Parallel Computing Toolbox.

F. Akaike and Bayesian Information Criteria and Residual Error

Akaike and Bayesian information criteria (AIC and BIC) were performed on the simulations that most closely replicated the data for both the Constant D and Variable D models. Both the AIC and BIC employ a penalty term that is a function of the number of parameters in each model in order to assess whether a better fit is justified in a model with more adjustable parameters. The expression for AIC is:

$$AIC = n\log\left(\sigma^2\right)+2K$$

where n is the number of datapoints, σ^2 is the sum of the residual error squared divided by the number of datapoints and K is the number of adjustable parameters in the model.

The expression for BIC is similar but has a larger penalty on additional free parameters and is given by:

$$BIC = n\log\left(\sigma^2\right) + K\log n$$

In the Variable D model there are 4 adjustable parameters, which determine the varying diffusion coefficients for all 11 proteins while in the Constant D model each proteins diffusion coefficient is determined separately yielding a model with 11 adjustable parameters. In addition the Variable D model yields a smaller residual error so it is evident that the AIC and BIC evaluations will favor the Variable D model. Despite this we have evaluated these criteria and they are given in Table 3S below and support the use of the Variable D over the Constant D model. When comparing the AIC and BIC values calculated here the more negative values indicate better fits. A difference of over 10 between the two models suggest that the model with the lower AIC or BIC value is highly preferred.

In addition we have evaluated the residual error in two regions of the data, at t/a^2 less than and greater than 0.125×10^8 s cm^{-2}, to compare how closely each model represents the data in these regions. It can be seen from Table 3S below that the Variable D model fits the data at the shorter time scales much more successfully than the Constant D model and almost equally at the longer time scales.

	Variable D Model	**Constant D Model**
Akaike Information Criteria	-1493	-1398
Bayesian Information Criteria	-1477	-1355
Delta AIC	0	95
Delta BIC	0	122
Residual Error	0.0204	0.0229
Residual Error $t/a^2 < 0.125$	0.0186	0.0232
Rsidual Error $t/a^2 > 0.125$	0.0784	0.0679

Table S3. Akaike and Bayesian Information Criteria and residual error comparisons between the Variable D and Constant D models. For both the AIC and BIC values a larger negative number is preferred. These values are negative because the residual errors were normalized by the maximal values on each axis to represent percentage error. In addition residual error was calculated over the data points where t/a^2 is less than and greater than 0.125×10^8 s cm^{-2}. Note that the residual error at short time and long times is calculated in a slightly different manner than the overall residual error leading to larger predicted error in the Constant D model for both short and long times than is found overall.

SUPPORTING REFERENCES

1) O'Leary TJ. Concentration dependence of protein diffusion. *Biophysical Journal* **52**:137-139, 1987.

2) Young ME, Carroad PA and Bell RL. Estimation of diffusion coefficients of proteins. *Biotechnology and Bioengineering* 22: 947-955, 1980.

3) Boyan JA. Learning evaluation functions for global optimization. Ph.D, Thesis, Carnegie Mellon University, August 1998

4) MATLAB documentation. http://www.mathworks.com/access/helpdesk/help/helpdesk.html