Molecular Regulatory Mechanisms Controlling Skeletal Muscle Development

Stephen Konieczny*

This presentation of the research in my laboratory will be a little bit distinct from the speakers that you have heard earlier and that you will hear during the rest of this conference. When you think about muscle, what most of you think about is a very complex structure, the sarcomere. We know, of course, that the sarcomere is composed of all of the proteins discussed by the previous speakers this morning. The importance of the sarcomere is that it is the functional unit of muscles and is the unit that achieves muscle cell contractions. A lot of very interesting and important biochemical and structural questions can be addressed concerning the sarcomere and how it is actually assembled. However, there is another important question that I want to touch upon today. That is: How do muscle cells actually accumulate the contractile proteins? More specifically: How are the genes that encode these proteins regulated? How does a muscle cell know that it should turn these genes "on" and how do non-muscle cells know that they should keep these genes "off"? This is a very large, broad question. It involves a lot of research in many different laboratories. What I would like to do today is to give you an overview of our understanding of gene regulation. Because this is really an early developmental question, I want to move back in development. I want you to forget about the mature adult muscle and think about where these muscle cells come from in the embryo.

Early Development

We all, of course, start out our lives as a multi potential cell, a single fertilized egg. These cells eventually give rise to specific stem cells. In this case, we are talking about a mesodermal stem cell. This cell has a very important role during development and makes very important decisions. For instance, some mesodermal stem cells will undergo a conversion to a bone cell lineage. Other mesodermal stem cells will undergo conversion to a fat cell lineage. And, of course, some mesodermal stem cells will undergo conversion to a muscle cell lineage. Today, I will concentrate on the muscle cells and the conversion process which eventually produces skeletal myoblasts, which are myogenic precursor cells. Myoblasts, as I am sure most of you know, are proliferative muscle precursor cells; they are undifferentiated at this stage. However, later in development these cells proliferate and populate a particular area of the embryo. At some point in development, these cells undergo a differentiation process. The differentiation process is characterized by two classical events. The first is morphological differentiation in which the individual myoblast cells actually fuse with one another to form very large, multi-nucleated muscle fibers. The second, and just as important, stage of differentiation is biochemical. These cells now begin expressing the contractile protein genes. They begin accumulating contractile proteins and forming functional sarcomeres, thus allowing the cell to contract. There are a number of important questions we can address here this morning. One of the questions is whether we can identify the regulatory mechanisms by which this mesodermal stem cell makes a decision to go to a myogenic cell lineage versus a fat cell lineage. The second question I will try to address is: Can we identify the molecular regulatory mechanisms which control the expression of the contractile protein genes?

One of the reasons most of us work with muscle cells is that muscle is actually an excellent model system to study some of these more basic developmental questions. Myoblasts can be isolated right out of an embryo and maintained in culture. They have a very characteristic morphology. Most of the cells are bipolar in shape. At this stage in development, they are not expressing the contractile protein genes. However, if we allow this culture to reach a little higher cell density and alter the growth conditions of the culture simply by removing specific serum growth factors, we will see a very interesting phenomenon occurring; morphological differentiation. Here you have a single large muscle fiber that is composed of hundreds of individual nuclei which came from individual myoblasts. All of the myoblasts in this culture fuse to form one large muscle fiber. At this stage of development, a muscle cell begins expressing the contractile protein genes and thereby accumulating the contractile proteins themselves. We have been interested in trying to understand how these genes become transcriptionally activated during this developmental stage and why it occurs only during this developmental time and only in this particular cell type.

Cloning

One of the ways that most molecular biologists approach these questions is to try to clone various genes. In fact, we have cloned a number of contractile protein genes. This slide happens to be a schematic drawing of the fast troponin I gene isolated from quail. This particular piece of DNA contains about 500 bases of 5′ flanking sequences. This is

*S. Konieczny, Dept. of Biological Sciences, Purdue University, West Lafayette, IN 47907

Reciprocal Meat Conference Proceedings, Volume 44, 1991.

ED. NOTE: Transcribed from a recording of Dr. Konieczny's presentation.

where the promoter region of the gene is located. It contains eight exons shown in white and it contains about 1½ kbp of 3′ flanking DNA. There are two important properties of contractile protein genes: the first is that most contractile protein genes begin the protein coding region within the second or third exon. The first exon usually is comprised of non-protein coding nucleotides. The second important point for the contractile protein genes is that most contractile protein genes contain a relatively large first intron. In the case of this troponin I gene, this is about 1.5 kbp. In the case of the M creatine kinase gene, this is about 3,000 bp in length. I am mentioning this because I will come back to this in just a moment. We have an extra gene here and the question I want to address is, can we identify, within this piece of DNA, the regulatory sequences which control the expression of this gene?

If you think about gene expression in muscle cells or contractile protein gene expression, there are three major components to it. The first is that the gene has to be turned on at the correct developmental time, it has to be turned on in the correct tissue type and it has to be turned on to the correct quantitative level. That is, you have to make enough of this product to eventually produce enough of the protein product. In order to begin addressing whether we can identify regions on this gene which encode this type of regulatory information, we can take the quail gene and introduce it back into mouse myogenic cell lines. Then we isolate RNA from three different groups looking at myoblasts versus myofibers or undifferentiated cells versus differentiated cells. As you can see, the mouse alpha actin which is the endogenous normal actin gene in these cells is expressed at very low levels in myoblasts and increases in myofibers. Now notice that for the introduced quail gene in these cells, the regulation is exactly as predicted. Myoblasts have very low levels of troponin I, but as the cells differentiate troponin I expression increases. What this tells us is a couple of important things. The first is that quail genes and genes found in a mouse or in bovines or in almost any organism are regulated in a similar fashion. Therefore, there has been an evolutionary conservation in the regulatory mechanism. If you can understand how one type of gene works in one species, most likely it works in a similar fashion in different species. The second point is that the gene, such as the one we used in these studies, must contain all of the essential regulatory information to be expressed in an appropriate fashion. In fact, we can take the quail gene and introduce it into a mouse embryo and make transgenic mice and, again, the quail gene is expressed only in skeletal muscles and only at the correct developmental stage. After a number of years of work, we have identified the entire, or what we believe is the entire regulatory region that is responsible for correct expression of the troponin I gene. Finally, it is interesting that this DNA sequence happens to be located in the first intron. The troponin I gene is regulated by DNA sequences internal to the gene, not in the promoter region, where most genes usually are regulated. The second point is that this sequence contains three protein binding sites. Therefore, for the troponin I gene to be expressed requires a number of components, including at least five regulatory proteins. The first protein is referred to as a muscle regulatory factor and I will talk a lot more about this in just a minute. This protein must interact with another protein, referred to as E12, as a dimer complex. Then we have three other proteins; proteins 1, 2, and 3. Two other points I want to mention here; the first is that all five of these proteins are absolutely required for this gene to be expressed. If we knock out any of the single proteins or their binding sites, this gene will not be expressed. The second important point is that E12 and proteins 1, 2, and 3 are found in all cell types. Although the troponin I gene is expressed only in muscle cells, the proteins that help with regulating expression of this gene are found in all cell types. The only muscle-specific regulatory protein is the factor which is referred to as the muscle regulatory factor. I would like to concentrate a little bit more on what these muscle regulatory factors are and why we think they are important for the development of a normal muscle cell.

Muscle Regulatory Factors

I think most of you are probably beginning to hear more and more about these factors. They are actually referred to as the helix-loop-helix muscle regulatory factors of which there are four; MRF4, Myo D, myogenin and Myf-5. The helix-loop-helix portion is a structural motif which is common to these four proteins. Why do we think that these four proteins are important in the regulation of gene expression in muscle cells? The first piece of evidence which suggests that these proteins are important in muscle development comes from the fact that the genes encoding these proteins are expressed only in muscle cells. If we look at MRF4, Myo D or myogenin expression, we can see that these genes are all expressed in various skeletal muscles such as plantarus, gastrocnemius and soleus. They are not found in heart muscles or smooth muscle and they are certainly not found in non-muscle tissues. That is at least one indication that these factors may be important to skeletal muscles because they are expressed only in skeletal muscle tissues in both the developing embryo and in the adult. Another important question we can address is that while we know where the proteins are expressed and where the genes are expressed; where, in fact, are the protein products located? A double immunofluorescent photomicrograph of a muscle fiber stained with antibodies against myosin and antibodies against MRF4 indicates where these products are found. As expected, myosin is cytoplasmic whereas the muscle regulatory factor proteins are found in the nucleus. This would be the predicted location for these proteins if, in fact, they are involved in regulating expression of contractile protein genes. We can ask that question a little bit more directly. Proteins that are found in the nuclei of cells can have a number of different functions but one of those functions, of course, may be to bind a DNA sequence. So we simply asked whether the muscle regulatory factors, the proteins themselves, recognize the same regulatory sequences which control expression of all of the contractile proteins genes. In this very simplistic example, what we are using here is a DNA probe which is the same DNA sequence that I showed earlier. The regulatory region is that associated with the troponin I gene. When we mix the various muscle regulatory factors such as Myo D, myogenin, MRF4 or the associated factor, E12 with this DNA, we see no specific binding of the protein to that DNA. However, when we co-incubate Myo D with E12 or

myogenin with E12 or MRF4 with E12, we see a very tight and specific binding of these protein complexes. So the dimers are recognizing the regulatory region that controls expression of the contractile protein genes. It turns out that these proteins actually bind to the regulatory regions associated with almost all contractile protein genes that we have been able to look at.

Although DNA-protein interactions are suggestive that these proteins are involved in regulating the expression of a particular gene, this is not direct proof. What we have done is to do what is referred to as transactivation essays. In this case, we have taken an alpha actin gene which normally would be expressed only in skeletal muscle cells. We have introduced this alpha actin gene into a non-muscle cell and looked for expression. The base line represents the expression of the alpha actin gene in the non-muscle cell. Note, however, if we take alpha actin and co-introduce into the non-muscle cell the Myo D protein, the myogenin protein or the MRF4 protein, what you see are very high levels of expression of this particular gene. In fact, there is about 150 to 200 fold increase in expression in a non-muscle cell when we add these muscle regulatory factors. Again, this is using alpha actin. We can use troponin I and see an identical type of response or almost any other known contractile protein genes.

What I just demonstrated to you is that the contractile protein genes and at least troponin I are regulated by these various protein factors and at least one of these protein factors represent a family known as the muscle regulatory factors. Clearly, this is the way the contractile protein genes initially are expressed. However, there is another important aspect of these muscle regulatory factors. If we look at a fibroblast cell line and we stain it with an antibody against myosin heavy chain, you see absolutely no staining. However, if we take the same fibroblast, non-muscle cells, and introduce MRF4, Myo D or myogenin regulatory proteins and then stain the cultures for myosin heavy chain, you will see that some of these cells begin expressing this specific contractile protein gene. In another example, if we introduce into the fibroblasts the muscle regulatory factor MRF4 and then stain the cells with an antibody against myosin, you not only see cells staining, you actually can discern that they are multi-nucleated. In fact, if we look at this entire culture, we find that this culture is expressing all of the contractile protein genes and is making functional sarcomeres. The point here is that the muscle regulatory factors have an extremely powerful effect on cells. We can introduce these proteins into a non-muscle cell and can convert it to a muscle cell. In fact, we can introduce these proteins into fat cells and convert them into skeletal muscle cells. Therefore, these factors have a profound effect on the development of cells in culture and presumably a profound effect on the development of the muscle in the animal.

In Conclusion

Finally, there are two other thoughts I want to leave you with. There are a number of effects of various components of serum such as serum growth factors. Although the muscle regulatory factors are extremely potent inducers of skeletal muscle development, it turns out that they themselves are regulated by various serum components. If we do a similar experiment to that I showed you by taking a non-muscle cell and introducing Myo D, myogenin and MRF4 and maintain these cultures in media which lacks serum or, in this case, lacks fibroblast growth factor, we see very high levels of differentiation. If you take the same cultures but this time feed the cells media containing serum or containing fibroblast growth factor, the percent differentiation is reduced approximately 90%. Therefore, the environment plays a major role in regulating these factors and certainly serum plays a major role in regulating these factors.

I want to leave you with a number of important questions that our laboratory is studying and that I hope you come away from this conference thinking about. First, what are the individual roles of the basic helix-loop-helix muscle regulatory factors? Why, in fact, are there four different factors even though they all appear to play very similar roles in controlling muscle development? How do serum growth factors actually regulate the activities of these proteins? Finally, what are the roles that these muscle regulatory factors have in maintaining the adult musculature? All of the studies that I discussed this morning deal with embryonic cells. Since these factors are maintained in the adult and they are found in the adult skeletal muscle, they presumably play a very important role in maintaining fiber distribution in the adult animal. Understanding how these factors maintain the adult musculature may lead to a better understanding of how to optimize muscle production in a variety of animals.