Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Mef2 and the skeletal muscle differentiation program

Michael V. Taylor\textsuperscript{1} and Simon M. Hughes\textsuperscript{2}

1. School of Biosciences, Sir Martin Evans Building, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK.
2. Randall Division of Cell and Molecular Biophysics, New Hunt's House, Guy's Campus, King's College London, London SE1 1UL, UK.

Abstract

Mef2 is a conserved and significant transcription factor in the control of muscle gene expression. In cell culture Mef2 synergises with MyoD-family members in the activation of gene expression and in the conversion of fibroblasts into myoblasts. Amongst its \textit{in vivo} roles, Mef2 is required for both \textit{Drosophila} muscle development and mammalian muscle regeneration. Mef2 has functions in other cell-types too, but this review focuses on skeletal muscle and surveys key findings on Mef2 from its discovery, shortly after that of MyoD, up to the present day. In particular, \textit{in vivo} functions, underpinning mechanisms and areas of uncertainty are highlighted. We describe how Mef2 sits at a nexus in the gene expression network that controls the muscle differentiation program, and how Mef2 activity must be regulated in time and space to orchestrate specific outputs within the different aspects of muscle development. A theme that emerges is that there is much to be learnt about the different Mef2 proteins (from different paralogous genes, spliced transcripts and species) and how the activity of these proteins is controlled.

Keywords

Skeletal muscle development, Mef2, gene expression network, differentiation, regeneration, alternative splicing, protein interactions

Contents

1. Discovery of the conserved Mef2 family
2. Mef2 gene expression
3. What does Mef2 do in muscle?
   3.1 Mammalian cell culture
   3.2 Mef2 function in \textit{Drosophila}
   3.3 Function of zebrafish Mef2
   3.4 Function of mouse Mef2
   3.5 Potential functions in muscle-related tissues
   3.6 Mef2 in muscle regeneration
   3.7 Mef2 and muscle fibre character
4. Mechanism of Mef2 action: Transcriptional targets and networks
1. Discovery of the conserved Mef2 family

We write on the thirtieth anniversary of the publication of the paper describing the discovery of MyoD and its ability to transform a fibroblast into a myoblast ([1] Lassar, this special issue). It is difficult to overstate the impact of this discovery both to the field of muscle differentiation and also beyond, to cell differentiation programs in general and to cellular reprogramming. In mammals, MyoD was quickly identified as one of four structurally and functionally related transcription factors, the Myogenic Regulatory Factors (MRFs)[2]. They bind to specific DNA sequences found in the cis-regulatory modules (CRMs) that control the expression of many muscle genes, and together the MRFs have a critical role in muscle differentiation during animal development [3]. However, there is more to making muscle than the MRFs. For example, other DNA sequences in CRMs together with their cognate transcription factors are required for muscle gene expression. Indeed, MRFs can induce gene expression through DNA sequences that do not contain their binding site [4, 5]. Mef2 quickly emerged as another significant player in the transcription of muscle genes.

Mef2 was first identified using mammalian cultured cells [6]. It was described as a protein activity that bound to a DNA sequence, distinct from the MRF-binding site, found in a CRM of the muscle creatine kinase (MCK) gene that is responsible for the expression of MCK during muscle cell differentiation. This Mef2-binding DNA sequence was subsequently found to be in the CRMs of nearly every known muscle-specific gene [5, 7]. Molecular cloning of Mef2 [8, 9] revealed that Mef2 is a member of the evolutionarily ancient MADS family of transcription factors [10, 11]. This family also contains SRF, which itself has a role in the regulation of muscle gene expression [12] and references cited therein; [13].

Mef2 is highly conserved (Fig. 1). There are single Mef2 genes in Drosophila, C. elegans and sea urchins, for example [7]. In contrast in the vertebrate genome, as with the MyoD family of MRFs, there are four closely-related Mef2 genes: Mef2a, -b, -c and -d [7]. Different transcripts are generated from each gene by alternative splicing [7, 9] (see 5.1 for functional significance). The encoded proteins bind DNA with the
consensus sequence CTA(A/T)C TAG and can form both homo- and hetero-dimers [8]. It is unknown whether different Mef2 isoforms bind to variants of the Mef2 consensus sequence [14]. Key functional domains include the N-terminal MADS box and the adjacent Mef2 domain, that together are necessary and sufficient for DNA-binding and dimerisation, and these regions show extensive sequence similarity across all the Mef2 proteins from different species [7]. Point mutations in the conserved MADS box of Drosophila Mef2 result in lethality [15]. The C-terminal region is required for transcriptional activation [16], but the sequence has diverged considerably in different Mef2 proteins. Mef2 protein binds DNA on high affinity binding [17], and while the MADS box confers DNA-binding specificity [8], the Mef2 domain modulates the DNA-binding affinity [16]. Both crystal and solution structures for a homodimerised highly conserved MADS/Mef2 domain of MEF2A bound to DNA have been determined [18-20]. Many aspects of the Mef2 structure/function relationship remain to be analysed.

2. Mef2 gene expression

An initial step in the investigation of the possible role of Mef2 in muscle differentiation was to determine its pattern of expression during embryonic development. Moreover, the expression pattern of individual genes, and indeed isoforms of each gene, continues to aid the interpretation of phenotypes and yield insights into the role of individual gene products. We will focus on Mef2 in skeletal muscle although, and in contrast to the skeletal muscle-specific MRFs, Mef2s are expressed in cardiac and smooth muscle, endothelium, brain neurons and many other tissues [21].

Invertebrates have a single Mef2 gene. In Drosophila, Mef2 is initially expressed throughout the mesoderm at gastrulation and subsequently almost exclusively in muscle. Both progenitors and differentiated cells of the somatic muscle (the equivalent of vertebrate skeletal muscle), heart and visceral muscle express Mef2 [22]. The situation is more complex in vertebrate species with their separate Mef2a, -b, -c and -d genes. Mouse was the first species in which their expression patterns were described during development.

Murine Mef2 genes are expressed at significant levels in a wide variety of tissues, but several show highest expression in muscle. Mef2c mRNA was first detected in the developing heart and shortly after in somitic muscle, and remained abundant in developing skeletal muscle at e14.5, being particularly concentrated at points of muscle attachment [23, 24]. Mef2d showed a similar although weaker pattern in somitic muscle [23]. Mef2a is also highly expressed in somites containing differentiated muscle fibres [23]. Mef2b has the most divergent structure, the least obvious muscle-related expression and the least evidence for involvement in muscle development. The expression pattern of the Mef2 genes during limb and head myogenesis is, to our knowledge, poorly described.

A striking feature of chick Mef2 expression is that it is biphasic. Mef2b and Mef2c show broad transient expression in nascent neck somites, prior to MRF expression [25]. In contrast later, in nascent trunk somites, Mef2c mRNA accumulates strongly prior to other Mef2 mRNAs in regions where both Myf5 and MyoD have become detectable [25]. Mef2d and Mef2a mRNAs are detected in each somite, a few hours after Mef2c [25, 26]. At slightly later stages, when terminally differentiated fibres are present in significant numbers in somites, the location of Mef2c and Mef2d mRNAs appear distinct, with Mef2c mRNA located at the dermomyotomal lips where new myogenesis is on-going and Mef2d mRNA located centrally within the myotome where more mature fibres are located [25]. Thus, Mef2c and Mef2d display a spatial and
temporal succession. This description reveals potentially interesting relationships both between the different Mef2 genes and to the MRFs, but the functional significance of this expression is at present unknown.

Zebrafish have six Mef2 genes, both *mef2a* and *mef2c* have two versions, and several are expressed during skeletal myogenesis [27]. Zebrafish *mef2d* is the first Mef2 gene detectably expressed, closely following the initiation of MRF expression in several distinct muscle precursor cell populations [28, 29]. *Mef2ca* mRNA accumulates later, in parallel with myosin expression, in several distinct myocyte populations as each begins to elongate and assemble myofibrils [29]. *mef2aa* and *mef2cb* are expressed in skeletal muscle after *mef2d* [30-36], but are not fully described. Expression patterns of *mef2b* and *mef2ab* have not been described, although neither was detected in somitic muscle at 2 days of age [35]. Mef2 immunoreactivity is particularly abundant in striated muscle [27, 29, 37-40]. Overall, it is clear that at least four Mef2 genes are sequentially expressed at high level during terminal differentiation of skeletal muscle in teleost fish.

Taken together, these studies show prominent and regulated expression of Mef2 genes during muscle development in a range of species. However, a lack of consensus across species in the order and location of expression of the four vertebrate Mef2 genes suggests that either underlying common patterns have been missed, or that gene usage has changed during evolution. Future studies at higher spatial and temporal resolution may resolve precisely when and where each Mef2 protein is available to function.

3. What does Mef2 do in muscle?

The prominent position of Mef2 in the field of muscle differentiation derives both from two early findings in mammalian cell culture and during development of *Drosophila melanogaster*, and from later work in a variety of model systems.

3.1 Mammalian cell culture

The first of these influential early studies uncovered a functional interaction between Mef2 proteins and MRFs [41]. The characteristic and striking property of the MRFs is that they can convert fibroblasts into myoblasts. This ability is greatly enhanced when MRFs are co-transfected with Mef2 [41]. Specifically, MyoD synergised with each of Mef2a, Mef2c and Mef2d, and Myogenin with Mef2c. Members of these two families of transcription factors also synergistically activated reporter gene expression [41]. This functional cooperation required direct physical interaction between the DNA-binding domains of both Mef2 and the MRF. Strikingly, it could occur through multiple copies of the binding sites for either Mef2 or MRF in the reporter construct. The study indicated that the two proteins could activate gene expression in two ways: (1) through binding to separate DNA sites; or (2) at a single site through a proposed Mef2/MRF physical interaction. Although muscle gene CRMs often contain nearby sites for the two factors [7, 42-45], the latter finding means that either factor could also activate transcription through the other’s binding site. This offers an explanation for how MRFs can induce muscle gene expression through DNA sequences that do not contain an MRF-binding site [4, 5].

Mutational analyses of MRFs had revealed that an alanine and a threonine in their DNA-binding domain are required for muscle gene activation [46, 47]. However, these amino acids are not required for DNA-binding, and so it was proposed that they are necessary for an interaction with a co-regulatory factor
required for muscle gene activation. Strikingly, the synergistic interaction between Mef2 and MRFs in both myogenic conversion of fibroblasts and expression of reporter genes requires these two specific amino acids, suggesting that Mef2 factors are the proposed co-regulators for the MRFs [41]. However, definitive in vivo evidence for such direct interactions is still lacking.

This research was complemented by a loss-of-function study that used a ‘dominant negative’ Mef2 construct to show that Mef2 family function is required by MRFs for the myogenic conversion of fibroblasts [48]. The same construct also inhibits formation of myotubes from C2C12 myoblasts. More recently, an alternative approach of knocking down Mef2a expression using a Mef2a-specific shRNA also impaired C2C12 myotube formation [49]. Together these cell culture studies point to an intriguing function of Mef2, but it was critical to probe the role of Mef2 in the much more complex in vivo situation, i.e. during the muscle differentiation program of gene expression in animal development. However, such analyses in vertebrates have been complicated by the aforementioned four closely-related Mef2 genes, with their overlapping expression patterns in multiple tissues.

3.2 Mef2 function in Drosophila

The second influential early finding came from the analysis of Mef2 function in Drosophila. Its single Mef2 gene encodes a highly conserved MADS/Mef2 domain and is also alternatively spliced, in common with other species [50] (Fig. 1). With the efficient genetics available, Drosophila was used to show that Mef2 is required for muscle development during embryogenesis. Mef2 mutant embryos lack differentiated somatic muscle [22, 51, 52]. Although somatic muscle progenitor cells are produced in a substantially unaffected pattern, they fail to fuse and differentiate further. The dominant conclusion from this period of Mef2 research was that it affects myogenesis relatively late, during the overt differentiation phase. We will return to this point in Section 4.1. Although the focus of this review is on skeletal muscle, it is of interest that Mef2 is also required for proper differentiation of both the heart and visceral muscle, and again early steps in their development appear substantially unaffected. Thus, Drosophila Mef2 is required for the differentiation of all the major muscle-types. As many muscle genes are expressed in multiple muscle-types, it is noteworthy that Mef2 might regulate some of the same genes in each, something that the MRFs cannot do.

There are two phases of myogenesis in Drosophila and Mef2 is required for each. The first phase produces the larval musculature during embryogenesis, and the second produces the adult musculature during metamorphosis. Drosophila adult somatic muscles are physiologically and functionally diverse and, unlike the larva, some contain multiple fibres, characteristics that are in common with vertebrate skeletal muscles [53-56]. The largest muscles include the dorso-lateral direct flight muscles (DLMs). The Mef2 mutants used to establish a function in the first phase of myogenesis are embryonic lethal, but other Mef2 hypomorphic mutants were found to have a reduced number of abnormally patterned DLMs [15, 57, 58]. More recently, temporally and spatially controlled expression of RNAi to knock-down Mef2 function only in adult myogenesis has shown that Mef2 is required for a diverse range of adult muscle types, including the DLMs and also the structurally distinct jump and leg muscles [59, 60]. DLMs are formed by a type of regeneration in a remodelling process during metamorphosis. Early in pupation, template larval muscles fuse with muscle progenitor cells that are maintained in an undifferentiated state in the larva. In early pupae in the RNAi Mef2 loss-of-function condition, unfused myoblasts are found around the muscle templates, remodelling is inhibited, and at the end of pupation DLMs are completely absent. Of note, Mef2 functions at
several distinct times in this muscle differentiation program (see 4.1). Temperature sensitive control of the Mef2 RNAi expression showed that Mef2 is required first for normal remodelling and then subsequently in promoting muscle maintenance [59]. Remodelling and maintenance are important aspects of tissue biology, and the functions for Mef2 in these processes may have implications for muscle repair.

3.3 Function of zebrafish Mef2

Mef2 gene function in zebrafish skeletal muscle has been examined by a combination of loss- and gain-of-function analyses. With the exception of mef2ca mutants, which show delayed cardiomyogenesis and reduced somitic myotome volume and muscle fibre size, strong skeletal muscle phenotypes of single Mef2 gene mutants or antisense morpholino knockdowns have not been reported [29, 33-35]. mef2ca;mef2cb double mutants have a severe lack of cardiomyogenesis, but no additional defect in skeletal myogenesis was described [34]. Nevertheless, knockdown of several Mef2 proteins with either a single broad-spectrum anti-Mef2 morpholino or a combination of mef2ca loss-of-function mutation with a specific Mef2d morpholino yielded a failure of proper sarcomere assembly and reduction in mRNA levels for a variety of sarcomeric proteins [29]. In contrast, mef2cb single mutants failed to show the cardiac-specific phenotype observed by two independent labs in Mef2cb morphants [33, 34]. In our view, therefore, results from morpholino studies should be treated with caution until either replicated by clean genetic analyses or a molecular mechanism is demonstrated explaining compensatory changes that rescue mutants [61]. Loss-of-function mutations in the other Mef2 genes are yet to be described. Taken together, these functional analyses suggest that a combination of Mef2 genes contribute to the overall Mef2 activity in skeletal muscle that is needed for proper myogenesis downstream of the MRF proteins and muscle growth, at least in part through effects on sarcomere assembly.

Although Mef2s normally act downstream of MRFs in skeletal myogenesis, once expressed Mef2s may also play a role in promoting MRF activity (see section 4.4). Over-expression of Mef2cb in zebrafish embryos leads to myod mRNA accumulation and slow myosin heavy chain expression in significant regions of cranial mesoderm [34, 36]. Over-expression of Mef2ca, on the other hand, does not induce these muscle genes, but instead appeared to dorsalize embryos, perhaps through up-regulation of BMP inhibitors such as chordin within the somite [36]. These data and others on the differential effects of splice variants (see 5.1) indicate that much remains to be learned about Mef2 function in specific times, places and types of skeletal myogenesis.

3.4 Function of mouse Mef2

Loss-of-function alleles of Mef2a, Mef2c and Mef2d have been generated by deleting the second coding exon [14]. There is no obvious skeletal muscle phenotype in Mef2a and Mef2d mutants. Mutants in Mef2c are early embryonic lethal shortly after somite formation, presumed due to cardiovascular defects. Nevertheless, a defect in the earliest stages of skeletal muscle development was suggested based on loss of expression of a reporter transgene in the somites [62]. The effect of Mef2c has been further explored through conditional knock out studies. Using a skeletal muscle specific Cre-recombinase line active from stage E8.5 to delete a floxed Mef2c allele, it was found that the mice die at post-natal day 1 with severely disorganized and fragmented skeletal muscle [63]. However, skeletal muscle differentiated and formed
normal myofibers during embryogenesis, but such myofibers subsequently and rapidly deteriorated with disorganized sarcomeres. It was concluded that Mef2c plays a key role in the maintenance of sarcomere integrity [63].

It should be noted that in another study using a very similar approach to the conditional loss-of-function of Mef2c in skeletal muscle, the mice were viable and did not show perinatal lethality [64]. Moreover, both sarcomere organisation and muscle ultrastructure were grossly normal, although there was a reduction in slow muscle fibres (see 3.7), and a function for muscle Mef2c was also identified in normal growth and glucose metabolism. The mice in this study had a slightly different Cre line in a different genetic background, and a fuller discussion of the possible impact of these two aspects is presented in [64]. Such differing findings suggest the role of Mef2c may be more complex than hitherto understood, perhaps due to differing functions in several muscle-related cell types.

3.5 Potential functions in muscle-related tissues

Chick Mef2c is expressed early and broadly in mesoderm, suggesting it may affect development of various tissues [25]. Over-expression of zebrafish Mef2cs in heterologous systems may influence osteo- or chondrogenesis through Sox9/BMP [65]. After initial myogenesis, zebrafish mef2ca mRNA becomes localised to fibre ends, whereas Xenopus Mef2c has been suggested to function in early tendon cells [29, 66]. Similarly, early expression of Mef2d in Xenopus may promote dermomyotome formation in addition to specific aspects of myogenesis [67, 68]. These observations caution that potential roles of Mef2 in muscle-related tissues could impact on muscle phenotypes observed and suggest that more research is required to define the cell autonomous roles of the Mef2 genes in muscle development in vivo.

3.6 Mef2 in muscle regeneration

Muscle regeneration in response to injury is a type of myogenesis related to, but distinct from, that occurring during development. It is dependent on satellite cells, a population of stem cells found underneath the basal lamina of muscle fibres (Zammit, this special issue). These cells generate myoblasts in post-natal skeletal muscle for both growth and for regeneration following injury, and become depleted in degenerative muscle diseases. The potential involvement of Mef2 factors in muscle regeneration has been investigated. Loss of Mef2a slows regeneration ([49], but see [69] for discussion of cell autonomy). Conditional deletion of the individual Mef2a, -c or -d genes in mouse satellite cells in vivo has no effect, but the combined conditional deletion of all three genes results in a block to regeneration [69]. Analysis suggests this is due to the Mef2 genes affecting the myoblast differentiation aspect of the regeneration process. This is the clearest demonstration to date of a function for Mef2 genes in mammalian muscle differentiation in vivo and echoes the role of Mef2 in the remodelling of Drosophila DLMs, a process with parallels to mammalian muscle growth and repair (see 3.2). As redundant roles for three Mef2 genes were revealed in murine muscle regeneration, it is possible that there might be a similar situation in embryonic muscle development that has prevented phenotypes being observed in single Mef2 gene knockouts.
3.7 Mef2 and muscle fibre character

Mef2 proteins have been implicated in the control of various aspects of muscle fibre character, either through developmental roles or by direct action in the adult fibre. Muscle fibres differ in size, contractile speed and metabolic properties based on the electrical activity imposed by the nerve that confers specific physiological characteristics to force production (reviewed in [70]). Such characteristics are frequently summarized by the term ‘fibre type’. Evidence exists that Mef2 regulates both muscle fibre type and size, but how the two effects are discriminated remains unclear. On some genetic backgrounds, muscle-specific deletion of mouse Mef2c can reduce muscle growth and sarcomere assembly [63]. Similarly, both embryonic zebrafish and mouse postnatal muscle growth was reduced after Mef2c knockdown and, in mouse, remaining muscle had altered contractile and metabolic properties [35, 64]. Conversely, in adult zebrafish, exercise increased muscle fibre size, mef2aa and mef2d mRNAs, but decreased mef2ca mRNA [71]. Thus, the relationship between Mef2 and fibre character appears complex.

Analysis of adult mice in which floxed Mef2c or Mef2d alleles were recombined in skeletal muscle by Cre driven from early-muscle transgenes has implicated Mef2 activity in the regulation of murine muscle fibre type [64, 72]. In vitro analyses show Mef2 can cooperate with a variety of factors, including Sp1 [73] on enhancers from myoglobin (an oxidative fibre gene) and MCK (which is preferentially expressed in fast fibres). Mef2 activity was first suggested to promote the slow oxidative fibre type based on: (i) the finding that Mef2a can cooperate with activated calcineurin protein phosphatase to activate enhancer elements that are preferentially expressed in slow/oxidative soleus compared to fast EDL muscle; (ii) the ability of slow-pattern electrical activity to activate calcineurin/Nfat signaling; and (iii) the preferential expression of an artificial enhancer consisting of three Mef2 binding sites in slow soleus muscle [74-76]. Although the levels of Mef2 protein appear similar in slow and fast muscles [72], muscle-specific deletion of either Mef2c or Mef2d, but not Mef2a, reduced the proportion of slow fibres, whereas over-expression of a Mef2c-VP16 fusion increased slow fibre content [72]. The persistence of significant numbers of each fibre type indicates, however, that either several Mef2 proteins cooperate or that other regulators of fibre type exist; generation of muscle fibres lacking several Mef2s would be of interest.

Muscle fibre type is not determined solely by electrical activity. In chick, cell autonomous intrinsic differences between early myoblast clones have been suggested to establish fibre type pattern, which is then confirmed and refined by innervating motorneurons [77, 78]. Tests have failed to find support for the hypotheses that increase in Mef2a or Nfatc activity can drive slow myosin expression in fast myoblasts or that Mef2 binding sites operate preferentially in myotubes derived from clones that express slow myosin [79]. Moreover, in zebrafish, in which slow and fast fibres are generated before motor innervation, levels of Mef2 protein appear similar [29]. Thus, evidence is lacking of an early role of Mef2 leading to persistent change in fibre character.

4. Mechanism of Mef2 action: Transcriptional targets and networks

Having reviewed the known function of Mef2 genes in a range of aspects of muscle differentiation, we will now consider the mechanisms of Mef2 action that underpin Mef2 function: how do Mef2s do what they do? To understand the place of a transcription factor such as Mef2 in a cell differentiation program, in this Section we consider gene expression. How is Mef2 expression regulated and what are the Mef2 target
genes, the products of which are the effector and realisator molecules that link the transcription factor with the cellular (and ultimately the animal) phenotype (see Fig. 2)? In Section 5, we will consider the Mef2 proteins and how they are modulated to produce specific outputs.

4.1 Mef2 target genes in the context of development

Analysis of Mef2 target genes indicates the role of Mef2 in the coordination of the gene expression programs that underpin muscle differentiation. Much initial progress in understanding the control of muscle gene expression came from single gene analysis in cultured muscle cell lines. This type of approach showed CRMs of most of the limited number of known muscle genes contained binding sites for Mef2 that are functional in cultured cells [7]. However, here we will focus on the physiological differentiation program that occurs during animal development. Initial investigations in Drosophila development again using the single gene approach, showed that ectopic expression of Mef2 in the ectodermal derivatives induces the ectopic expression of a range of muscle genes, including myosin, tropomyosin, β3-tubulin, nautilus, αPS2, muscle-blind, and muscle LIM [80-83]. In vivo analysis of Mef2 sites in CRMs and in vitro DNA-binding assays argued that a number of these genes are direct targets of Mef2, that is Mef2 binds to CRMs of the muscle gene to influence its expression [84-88]. However, the relative significance of individual targets to overall Mef2 function remained unclear.

Understanding was transformed by the application of genomic approaches, particularly Chromatin Immuno-Precipitation (ChIP) to map where Mef2 binds to genomic DNA. ChIP identified a significant number of potential Mef2 gene targets in both the mammalian C2C12 muscle cell line [43, 89] and in Drosophila embryos [90], the latter in conjunction with computational genome scanning for Mef2 sites. A pioneering study then used genomic ChIP in combination with gene expression profiling in Mef2 mutant Drosophila embryos to identify Mef2 target genes throughout normal development [91]. A series of time-points that spanned the period of muscle development was assayed. From this analysis we highlight three significant areas:

4.1.1 Number of Mef2 target genes

Approximately two hundred direct target genes were identified on the basis of having Mef2 binding in the vicinity of the transcription unit and expression being affected by Mef2 [91]. This is many more than had been anticipated from the pre-genomic era. Some of these Mef2-binding regions were shown to function as muscle CRMs. The total number of Mef2 targets could actually be double that described, because only half the genome was covered by the tiling arrays used for the ChIP, and could correspond to most muscle-specific genes.

4.1.2 Types of Mef2 target genes

The identified Mef2 target genes were expressed at different times during the differentiation program and have diverse functions [90, 91]. Many have roles in different aspects of terminal muscle differentiation, including muscle structure, adhesion, fusion, and neuromuscular junction assembly, which supported the widely held view of Mef2 as a regulator of the later steps of muscle differentiation [90, 91]. However, there had been prior indications of an effect of Mef2 on muscle gene expression earlier than this [86, 92], and a striking finding from the genomic approach was that many target genes are expressed early. These genes
have roles in the specification and diversification of muscle progenitor cells. The number and diversity of Mef2 target genes expanded the view of Mef2, suggesting a broad role for Mef2 in orchestrating successive steps in the muscle differentiation program in vivo during development.

4.1.3 Mef2 target gene specificity

The list of Mef2 target genes identified in Drosophila suggests that the activity of Mef2 is regulated in both space and time. Mef2 is expressed in the entire myogenic lineage from gastrulation to the end of embryogenesis and yet some target genes are expressed early, others late, and yet others in subsets of muscle cells [91]. How does this occur?

One property of Mef2 that is regulated is its binding to chromatin. There are three temporal profiles of CRM occupancy: continuously bound by Mef2; bound early but not late; and bound late but not early [91]. Moreover, binding of Mef2 to a gene’s CRM is correlated with the onset of that gene’s expression, with the correlation particularly strong for the late-bound CRMs [91]. Dynamic Mef2 binding was also found during C2C12 differentiation [89]. It is not known what underpins these observations. One explanation is cooperative chromatin binding with a differentially expressed factor. Sandmann et al [91] suggest co-operative binding with the transiently expressed bHLH protein Twist as a possible explanation for transient Mef2 binding to the early group of CRMs [91]. Other possible, non-exclusive, explanations include structural changes to chromatin, as suggested for the p38 MAPK pathway in a cell culture model of muscle differentiation [42], and on modifications of Mef2 (see 5.2).

Further indications that Mef2 is not an all-or-none regulator came from another genomic approach using expression profiling of an allelic series of Mef2 mutant embryos [93]. This revealed that target genes expressed early in muscle development require lower levels of Mef2 activity than those expressed later. The idea that levels of Mef2 activity control differential expression of muscle genes had also been suggested from expression of different levels of a Mef2 construct in a Mef2-null background [82]. This showed that distinct levels of Mef2 are required in Drosophila for particular properties of a muscle cell.

4.2 Vertebrate genomic studies and the emerging gene expression network

Equivalent genomic studies in vertebrate embryonic development to those in Drosophila are challenging and have not yet been undertaken. However, genomic approaches have been applied to good effect in cultured cells. Mef2-regulated genes have been identified in satellite cell-derived myoblasts using microarrays for expression profiling [69]. Microarray analysis of genes regulated by Mef2a, Mef2b, Mef2c and Mef2d in the C2C12 muscle cell line yielded the significant finding that individual Mef2 family members regulate specific subsets of the entire cohort of Mef2-regulated genes [94]. This was not anticipated because of the longstanding idea of functional redundancy in the vertebrate Mef2 family, which arose from the overlapping patterns of expression and lack of strong single mutant skeletal muscle phenotypes. ChIP has also been used to identify potential Mef2 target genes. In C2 muscle cells, ChIP-Seq identified genes bound by Mef2d [95], whereas ChIP-exo revealed both common and distinct Mef2a genomic targets in skeletal and cardiac muscle cells [96]. These studies, although presenting an incomplete picture, make it apparent that understanding how each Mef2 regulates all its targets in a co-ordinated manner is central to understanding the network of gene expression that controls the differentiation program of each muscle cell.
4.3 Downstream of Mef2: feedback and feed forward

To reveal the organisation of the gene regulatory network, understanding of Mef2 target genes that themselves act as regulators is essential. These include: (1) other transcription factors; (2) microRNAs (Mok et al, this special issue); and (3) modulators of Mef2 activity. Here, we highlight some illustrative examples (Fig. 2).

Mef2 targets that encode transcription factors with roles in muscle development have been identified in *Drosophila* embryogenesis [91] and in C2C12 cell differentiation [43]. Each of these transcription factors will have its own cohort of target genes. These may include *Mef2* itself (see 4.4), or targets that function with Mef2 in part of a feed-forward mechanism [91]. The latter echoes a feed-forward mechanism from MyoD via Mef2 and the p38 MAPK in mammalian cultured myoblasts that produces a delay in MyoD target gene expression [42]. The intrinsic delay of such feed-forward processes can be part of timing mechanisms controlling sequential gene activation in a differentiation program (see 4.4).

MicroRNAs, which also modulate gene expression, are among the direct transcriptional targets of Mef2. Examples include miR-1 and miR-133, and the largest known mammalian microRNA cluster, the Gtl2-Dio3 locus [49, 97, 98]. The latter is implicated in muscle regeneration. A documented example of a microRNA target that feeds back negatively on *Mef2* is miR-92b, which can interact with the 3'UTR of *Mef2* to fine-tune its expression [87].

Lastly, other Mef2 targets feed back on Mef2 to modulate its activity. These include HDAC9, a class II HDAC, which can associate with Mef2 proteins and suppress their transcriptional activity [99], and *Drosophila* Him, another inhibitor of Mef2 activity [91] (Taylor MV, unpublished). How such negative feedback loops operate to achieve progress and stasis in the muscle lineage *in vivo* remains to be determined.

4.4 Closing the Loop: Regulation of Mef2 expression

In vertebrates, MRFs can induce Mef2 mRNA, protein and DNA-binding activity [5, 9, 42, 100]. The specific DNA sequence in the mouse Mef2c gene required for its direct activation *in vivo* by MRFs has been defined [101, 102], as have other CRMs for other regulatory inputs [103]. There is also auto-activation of Mef2 expression, and this may be direct or indirect. One indirect route is through Mef2 proteins activating the expression of MRF genes, which in turn activate Mef2 genes. Such positive feedback has been described for *myogenin* in both cell culture and developing mouse embryos, and for *Myod* in satellite cells [104-107]. A direct route to auto-activation has been suggested for mouse Mef2c [101], but others have not found this [102].

There are significant parallels in the transcriptional regulation of the *Drosophila* Mef2 gene. It also has multiple CRMs, which indicates the capacity to integrate many regulatory inputs [108]. Notable amongst these is the bHLH transcription factor Twist, which activates *Mef2* expression [50, 109]. Twist has a major role in propelling cells down the myogenic pathway in *Drosophila* embryogenesis, and its role recalls that of the MRF family of bHLH proteins in vertebrates [110, 111]. Mef2 also acts on its own CRM, it auto-activates [112]. Other defined inputs include Tinman, which is related to the Nkx2 group of vertebrate transcription factors, and the steroid hormone ecdysone that functions in the coordination of metamorphosis [113, 114].

The cross-activation between Mef2 and the myogenic bHLH factors coupled with the observed auto-activation, in both *Drosophila* and vertebrates, points to evolutionarily conserved positive feedback gene
transcription loops as a mechanism for cell commitment. These loops can amplify and maintain the expression of these key transcription factors with the potential to maintain muscle differentiation. This is a central feature to incorporate into models of the regulatory transcriptional network that underpins muscle differentiation (Fig. 2).

4.5 Network features

Taken together, the above information indicates how the network will include both feed-forward and feed-back loops and how the activity of nodes like Mef2 in the muscle network might operate. Another aspect to incorporate into any model, as emphasised by Cunha et al [115], is the combinatorial and temporal features of the regulation of a gene at any node in the network. Initial steps to assemble this type of information have been taken. For example, Twist co-occupies 42% of all Mef2 bound CRMs in the early steps of *Drosophila* mesoderm development [116], and Lmd, a Gli-family transcription factor, modulates Mef2 transcriptional output when co-occupying CRMs [115].

In vertebrates, distinct properties of the different Mef2 and MRF family members also need to be considered. For example, there are aspects of the transcriptional inter-relationship between MRFs and Mef2s that are specific to particular members of the two families of transcription factors [117]. More generally, there are non-overlapping gene sets in the population of genes regulated by each of the four different mouse Mef2 genes, and only a small percentage of genes are co-regulated by all four [94]. Lastly, for any specific family member, relationships may not be the same in all situations. For example, Twist and Mef2 are not co-expressed at all stages of *Drosophila* muscle development. Clearly, although we can highlight key aspects of the network in muscle, much is not understood currently. Nevertheless, the prize for this substantial task of defining the myogenic network will be insights into how robustness of the transcription program is achieved and how its activity can respond to different conditions.

5. Mechanism of Mef2 action: post-transcriptional events

Multiple findings indicate that to comprehend the roles of Mef2 one must understand how its activity is regulated. Thus, Mef2 muscle target genes are expressed in both a spatially and temporally regulated manner during development (see 4.1). For example, Mef2 protein is expressed in *Drosophila* muscle progenitor cells very many hours before differentiation is underway and many target genes are expressed [118]. Although outside the scope of this review, Mef2 also functions in different cell-types to produce distinct outputs [21]. Combining these two characteristics, Mef2 may fulfil different functions in time and space within the developing and mature skeletal muscle system. Insight into how Mef2 is able to undertake its range of roles will require understanding of its post-transcriptional regulation. Here we highlight alternative splicing, protein modifications and protein interactions. Each was identified in the early years of Mef2 research, but much has been discovered since.

5.1 Alternative splicing

One way to diversify gene function is through alternative RNA splicing. Mef2 proteins across the metazoa show many small conserved regions throughout their length, in addition to the highly conserved N-
terminal MADS/Mef2 domain (Fig. 1). Furthermore, evolution has varied Mef2 gene products through alternative splicing, which appears to focus in specific regions, presumably those most forgiving of alteration, yet controlling function. For example, although *Drosophila Mef2* shows little conservation beyond the conserved N-terminal regions, the locations of exon borders and alternative splice sites are similar to those in vertebrate Mef2 genes. Two splice variants map to the region of the vertebrate alternatively spliced γ exon, despite lacking obvious sequence homology (designated exon 8 and γ in Fig. 1) [50, 82]. In vertebrates, specific alternative splice variants in these regions are generated in muscle, but *Drosophila* Mef2 isoforms have yet to reveal distinct function [82]. Even diploblasts, e.g. jellyfish and sea anemone, appear to require full length Mef2 for their mesendoderm-derived myogenesis and show good conservation to vertebrate Mef2s, with regions of alternative splicing matching a subset of those in vertebrates [119, 120]. Despite this conservation, individual species have sometimes lost domains. Downstream of the N-terminus C. *elegans Mef2* lacks most homologous regions, except for a C-terminal stretch matching the γ-exon, but it is not essential in myogenesis [121]. Such changes suggest that specific Mef2 sub-functions may only be retained in Mef2 genes of some species.

The multiplication of vertebrate Mef2 genes has permitted gene diversification. *Mef2b* genes show significantly greater divergence and lack two exons, obvious muscle-specific splice variants and several conserved domains [122]. In contrast, *Mef2a, -c* and -d genes have generally maintained alternative splicing, suggesting that each retains distinct functions in their splice variants (Fig. 1). Specific functions have been addressed in cell culture studies. Muscle-specific splicing of exon 3 occurs in *Mef2a, -c* and -d [9, 123-126]. In *Mef2d*, this seems to regulate responsiveness to PKA signaling that can control myoblast differentiation, whereas in *Mef2c* it controls responsiveness to kinases regulating the cell cycle and growth [95, 123, 127]. Splicing at other locations may modulate the efficacy of Mef2 in myogenesis by controlling its interaction with MyoD, responsiveness to signals and transcriptional activity [36, 128, 129].

Of potential clinical relevance, inclusion of exon 3α2 in *Mef2c* is diminished in rhabdomyosarcoma cells and replacement can reverse transformation and promote their terminal differentiation [130]. Moreover, altered Mef2 splicing may contribute to myotonic dystrophy [131]. Mechanistic understanding of Mef2 alternative splicing is rudimentary, and whether similar splice variants show common functional differences in various tissues or between genes remains unclear [67, 132]. Further studies are therefore essential to understand the regulation of alternative splicing and the functional roles of each isoform produced in muscle tissue in vivo.

### 5.2 Mef2 modification

Since the demonstration that phosphorylation of the MADS/Mef2 domain Mef2 in cultured cells could increase its DNA-binding activity [133], many different types of modification in different domains of Mef2 have been reported. These include the following illustrative examples in cultured muscle cell lines. p38 MAPK phosphorylates Mef2 at a number of sites and Thr293 phosphorylation of Mef2c is linked to enhanced transcriptional activity [134]. p38 also regulates the formation of DNA-binding Mef2/MyoD complexes and RNA PolII recruitment [42]. Mef2d is phosphorylated by PKA at Ser121 and Ser190 and this is linked to repressed transcriptional activity [135]. Transcriptional activity of Mef2D is also repressed by methylation at Lys267 [136] and by sumoylation at Lys439 in the C-terminal activation domain [137]. Another modification linked to Mef2 function is acetylation. In Mef2c this occurs at Lys4 [138] and at multiple sites in the activation
domain [139]. In both cases, modification is differentiation-dependent and DNA-binding activity is enhanced. In general, much remains unknown concerning the effects of these modifications and their significance in vivo during muscle development. Nevertheless, one example with intriguing in vivo findings is p38 MAPK, which is required for Mef2 activity and for muscle differentiation in mouse development [140]. In Drosophila, p38 activity in muscles regulates motor function and was shown to work through Mef2 in the regulation of life span [141]. This brief survey indicates that Mef2 proteins can be covalently modified in a variety of ways and that these modifications are a route to modulating Mef2 protein function.

5.3 Mef2 protein interactions

Shortly after its discovery it was apparent that protein interactions were significant to Mef2 function: it both dimerised and interacted with the MRFs [8, 41, 142]. Protein interactions can produce specific outcomes, and one largely unexplored aspect is whether the potentially large number of different Mef2 dimers have different functions. Most Mef2 protein interaction studies have been in vitro or in cell culture, and have been recently reviewed [143]. Here we highlight some examples with significance to Mef2 function in muscle.

Mef2 proteins interact with other transcription factors in addition to the MRFs. These include Smad proteins, which can bind to Mef2 in muscle cells and functionally co-operate in transcription assays [144]. The latter is dependent on p38 phosphorylation sites (see 5.2). Klf5 is a zinc finger transcription factor required for muscle regeneration in vivo and muscle differentiation in cell culture [145]. Klf5 directly associates with MyoD and Mef2 in vitro, and ChIP shows overlap with its chromatin sites and those bound by Mef2 and MyoD. The transcriptional adapter nTRIP6 can bind to Mef2c and repress target gene expression in myoblasts [146]. Intriguingly, TRIP6 can interact with Nup210 in a nuclear pore complex implicated in muscle gene expression control [147]. In the context of temporal control, the Nfix transcription factor that regulates fetal-specific muscle gene transcription binds to and enhances PKC-mediated phosphorylation of Mef2a and target gene expression [148]. In Drosophila, Mef2 can bind the myogenic bHLH protein Twist [149], paralleling the vertebrate MRF finding, and Twist co-occupies 42% of all Mef2-bound CRMs [116]. Mef2 can also bind two other transcriptional regulators with roles in muscle development, Sd and Vg, related to vertebrate TEAD and Vgl-2 respectively [149, 150].

Notable amongst the Mef2-interacting proteins are the class IIa HDACs. They can interact with Mef2 proteins and inhibit their activity, and over-expression can inhibit differentiation of muscle cell lines, a function linked to cytoplasmic/nuclear shuttling [151-153]. In vivo, class IIa HDACs suppress the formation of slow muscle fibres through repression of Mef2 activity [72]. Conversely, the acetyltransferase p300 can bind to Mef2c and thereby augment its transcriptional activity [154]. Lastly, the proline isomerase Pin1, which inhibits muscle differentiation in culture, can bind Mef2c isoforms containing the 3α1 exon upon phosphorylation of two Ser-Pro sites [155]. Over-expression of Pin1 decreases Mef2c stability and activity, likely through an induced conformational change.

The above examples illustrate how some Mef2 interactions with other transcriptional regulators might operate in the context of CRMs that contain cognate binding sites for each transcription factor. Others, for example Vg/Vgl-2, which has no DNA-binding domain, might require just one site, or be independent of DNA-binding. They also illustrate how protein modification and protein interaction are closely inter-related. In
order to trigger modification there must be at least a transient interaction, and some interactions may be dependent on prior protein modifications.

6. Coda: Conclusions and future directions

The family of Mef2 genes is well-characterised and a lot is known about the roles of the encoded Mef2 transcription factors in muscle cells. However, there are significant gaps to fill in linking findings in cell culture to the \textit{in vivo} muscle differentiation program. First, what do the different Mef2 genes do? Second, Mef2 genes produce different proteins by alternative splicing, and these proteins can undergo post-translational modifications and interact with diverse partner proteins. Understanding how these features impact on Mef2 protein activity to produce specific transcriptional outputs will be pivotal to a full understanding of Mef2 function within the muscle differentiation program. Moreover, although we have focussed on Mef2 in skeletal muscle, Mef2 has roles in other cell-types [14, 21, 156-158], and therefore, similar analyses in different tissues may illuminate Mef2 roles in muscle.

An unknown is the ancestral function of Mef2. As mentioned above, in some species, Mef2 is neither expressed in muscle, nor required for myogenesis [121, 159]. In these species it is possible that another MADS protein, an SRF family homologue, fulfils a Mef2-like role [160]. SRF and Mef2 homologues exist in yeast and regulate cell shape and stress-resistance [161]. As, in animals, SRF is a fundamental actin sensor [162] and muscle evolved as an extreme adaptation of the actomyosin cytoskeleton [163], modification of an actin-sensor system involving a MADS protein may have participated in the evolution of muscle. Such thinking indicates that in considering the various extant functions of Mef2 proteins, it will be helpful to identify conserved ancestral Mef2 function(s) and the evolutionary clades in which specific new roles evolved.

A core area for future work is the transcription network that underpins the muscle phenotype. The first steps to understand this on a genomic scale during development have been taken in \textit{Drosophila}. Linking \textit{in vivo} functional networks to the regulatory layers of chromatin and nuclear organisation will be illuminating. For example, p38 signalling to Mef2D can recruit a chromatin modifier to muscle gene CRMs [164, 165], and spatial organisation of the CRMs of muscle genes is reported to contribute to temporal regulation of expression [166]. In a similar vein, a Mef2c-containing complex has been described assembled at the nuclear pore to regulate muscle gene expression, development and maintenance [147]. Integration of these additional regulatory layers with more conventional gene expression studies will produce an in-depth understanding of the muscle differentiation program.

Acknowledgement

SMH is a Medical Research Council Scientist with Programme Grant MR/N021231/1 support. We apologise that, due to space limitations, in some areas we have had to quote previous reviews rather than the primary literature. We thank Massimo Ganassi and Yaniv Hinits for comments on the manuscript.
Figure legends

Figure 1 Alignment of Mef2 Structure and Functions.
Top row: named protein domains and their sizes. HJURP-C Holliday Junction Recognition Protein C-terminal Domain, TAD Transactivation domain, NLS nuclear localisation sequence.
Third row: conservation of domains is indicated by colour intensity (lengths not to scale). Species showing conserved regions are listed beneath: Hs Homo sapiens, Dr Fish Danio rerio, Brachi Brachiopod Terebratalia transversa, Ce Nematode Caenorhabditis elegans, Dm Fly Drosophila melanogaster, Pc Jellyfish Podocoryne carnea, Nv Sea anemone Nematostella vectensis.
Fourth to sixth rows: conserved exon structure and splice locations in Dm and Nv and lesser central conservation in Ce. Exon length differences are indicative, but not to scale.

Figure 2. Schematic model of Mef2 at a nexus in the network of gene expression that underpins the muscle differentiation program
1. Mef2 transcription is stimulated by the muscle bHLH proteins and by direct and indirect auto-regulatory loops (in red) and other inputs.
2. Mef2 transcripts are alternatively spliced.
3. Distinct Mef2 proteins are produced from different RNAs and different paralogous genes. They form dimers with the potential for distinct functions
4. Mef2 proteins activate a series of targets (arrows indicate the possibility of shared and distinct targets between different Mef2 proteins). The effectors include other transcription factors, microRNAs and other regulators (e.g. HDACs), which can feedback onto Mef2 though different routes (in blue). Mef2 plus these transcription factors can feed-forward (7) onto realisator genes (6).
5. The activity of Mef2 protein is modulated by modifications (X and Y), interaction with other proteins, including transcription factors and other regulators e.g. HDACs.
6. Mef2 proteins in different contexts/forms activate a range of realisator genes over space and time to produce the muscle phenotype.
7. Some realisator genes may be activated by a feed-forward mechanism.
Figure 1

[Diagram of protein domains and conservation across different species]
Figure 2
References


145. Hayashi, S., et al., Klf5 regulates muscle differentiation by directly targeting muscle-specific genes in cooperation with MyoD in mice. Elife, 2016. 5.


