Reconstructing the Regulatory Kinase Pathways of Myogenesis from Phosphopeptide Data*[S]

Lawrence G. Puente‡‖, Sébastien Voisin§¶, Robin E. C. Lee‡‖, and Lynn A. Megeney‡‖**

Multiple kinase activities are required for skeletal muscle differentiation. However, the mechanisms by which these kinase pathways converge to coordinate the myogenic process are unknown. Using multiple phosphoprotein and phosphopeptide enrichment techniques we obtained phosphopeptides from growing and differentiating C2C12 muscle cells and determined specific peptide sequences using LC-MS/MS. To place these phosphopeptides into a rational context, a bioinformatics approach was used. Phosphorylation sites were matched to known site-specific and to site non-specific kinase-substrate interactions, and then other substrates and upstream regulators of the implicated kinases were incorporated into a model network of protein-protein interactions. The model network implicated several kinases of known relevance to myogenesis including AKT, GSK3, CDK5, p38, DYRK, and MAPKAPK2 kinases. This combination of proteomics and bioinformatics technologies should offer great utility as the volume of protein-protein and kinase-substrate information continues to increase. ***Molecular & Cellular Proteomics 5:2244–2251, 2006.***

The process of skeletal muscle formation, or myogenesis, has been intensely studied because of the medical significance of muscle self-repair and as a model for cellular differentiation in general. During embryonic development, secreted factors including Shh and Wnts drive the commitment and formation of myogenic progenitor cells (1). Myoblasts, proliferating myogenic progenitors that express MyoD and Myf5, differentiate into non-proliferating myocytes that express myogenin and MRF4 (2). Subsequently these myocytes fuse into multinucleated syncytia that mature into contractile myofibers. This process of myoblast proliferation, cell cycle arrest, and myocyte fusion is recapitulated in injury repair and in cell culture models of myoblast differentiation such as the murine C2C12 muscle cell line. Transcriptional control of myogenesis has been extensively studied, firmly establishing the essential role of the myogenic regulatory factors (MRFs)[1] Myf5, MyoD, myogenin, and MRF4. Accordingly recent interest has focused on defining the kinase signal transduction mechanisms that also modulate this process.

The best characterized kinase signaling factors in myogenesis are the p38 mitogen-activated protein kinases. Inhibitors of p38 signaling block muscle-specific gene expression and myotube formation (3, 4). There is evidence that p38 acts through multiple mechanisms including phosphorylation of the MRF cofactor myocyte enhancer factor 2 (4–6), phosphorylation of the MyoD cofactor E47, and enhanced recruitment of the SWI/SNF chromatin-remodeling complex to MRF-targeted promoters (7). Another key signaling kinase in muscle is AKT1. Stimulation of AKT1 by IGF-1 or insulin induces muscle hypertrophy, whereas inhibition of AKT1 by glucocorticoids promotes muscle atrophy. Hypertrophy is associated with AKT-mediated mTOR (mammalian target of rapamycin) activation and GSK3B repression (8). Conversely AKT1 inhibition leads to dephosphorylation and activation of FOXO transcription factors resulting in atrophy (9). Although some kinase activities are required for normal muscle development, others must be suppressed. JNK1 is normally inactive during myogenesis, and its activation leads to muscle pathology (10). In addition to these kinases numerous other kinase proteins have been implicated as regulators of the differentiation process. For example CDK5 expression and activity increases during early myogenesis, and expression of dominant-negative forms of CDK5 inhibits muscle formation (11, 12). However, an important unresolved question is how these different

From the ‡Ottawa Health Research Institute, Molecular Medicine Program, Ottawa Hospital, Ottawa, Ontario K1H 8L6, Canada, §Institute for Biological Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada, and ‖Department of Cellular and Molecular Medicine and Centre for Neuromuscular Disease, Faculty of Medicine, University of Ottawa, Ottawa, Ontario K1H 8M5, Canada

Received, April 12, 2006, and in revised form, August 3, 2006

Published, MCP Papers in Press, September 13, 2006, DOI 10.1074/mcp.M600134-MCP200

[1] The abbreviations used are: MRF, myogenic regulatory factor; 2D-GE, two-dimensional gel electrophoresis; CDK, cyclin-dependent kinase; CRHSP, calcium-regulated heat stable protein; CRMP, collapsin response mediator protein; dH_2O, deionized water; DYRK, dual specificity tyrosine (Y) phosphorylation-regulated kinase; ERK, extracellular signal-regulated kinase; FOXO, forkhead box (transcription factor); GSK, glycogen synthase kinase; IGF, insulin-like growth factor; JNK, c-Jun NH_2-terminal kinase; MAPKAPK, mitogen-activated protein kinase-activated protein kinase; PEA, phosphoprotein enriched in astrocytes; PKC, protein kinase C; STAT, signal transducer and activator of transcription; HNRNPK, heterogeneous nuclear ribonucleoprotein K; CK, casein kinase.

pathways are integrated into an overall network of kinase-substrate interactions to control myogenic differentiation.

Previously we profiled changes to protein phosphorylation during myogenesis on a proteome-wide scale by using phosphoprotein enrichment and comparative two-dimensional gel electrophoresis (13). Here we extend those findings by mapping sequence-specific sites of protein phosphorylation in growing and differentiating C2C12 cells. To place these results in a systems context, we took advantage of recent advances in the availability of protein-protein interaction databases. Databases of protein-protein interactions have been constructed from yeast, fruit fly, worm, and human cells using yeast two-hybrid, tag/pulldown, and literature search approaches (14). Subsets of the total protein-protein interaction set such as kinase-substrate interaction maps have also been constructed (15). To date, however, relatively few studies have leveraged this type of information to assist in the interpretation of results. Here we used kinase-substrate interaction databases to bioinformatically reconstruct a kinase signaling network based upon our experimentally identified phosphorylation events. This approach yielded a model kinase-substrate network that predicted several known features of myogenic signaling and revealed the potential relevance of newly identified phosphorylation events in relation to known muscle-related biochemical pathways.

EXPERIMENTAL PROCEDURES

Cells and Sample Preparation—C2C12 myoblasts were cultured in growth medium (10% fetal bovine serum, Dulbecco's modified Eagle's medium), and whole cell lysates were collected either 0 or 24 h after a change to differentiation medium (2% horse serum, Dulbecco's modified Eagle's medium) as described previously (13).

Phosphoprotein Enrichment—Phosphoprotein enrichment was performed by affinity column purification of total myoblast proteins using PhosphoProtein Purification kits (Qiagen) as described previously (13).

Analysis of Putative Phosphoproteins by One-dimensional Gel Electrophoresis—Phosphoprotein-enriched samples (typically 100–150 μg) were washed in water and concentrated using Amicon 10,000 molecular weight cutoff ultrafiltration columns (Millipore). Aliquots of retentate were separated by SDS-PAGE (12.5% acrylamide "mini" format). Gels were stained with Bio-Safe colloidal G-250 Coomassie Blue (Bio-Rad). 23 1-mm gel sections were excised from each lane, and in-gel digest and extraction of proteins was performed.

In-gel Digest—Gel pieces cut into 1-mm cubes were destained in 100 mM ammonium bicarbonate (NH_4HCO_3) in 30% ACN. Destained gel pieces were washed in dH_2O and shrunk in ACN. Gel pieces were incubated for 1 h at 56 °C in 10 mM DTT, 50 mM NH_4HCO_3; washed in 50 mM NH_4HCO_3; and incubated for 1 h at room temperature in 55 mM iodoacetamide, 50 mM NH_4HCO_3. Gel pieces were shrunk with ACN and reswollen in 50 mM NH_4HCO_3 with Promega modified trypsin (10 ng/μl). Sufficient 50 mM NH_4HCO_3 was added to cover the gel pieces, and sealed tubes were incubated overnight at 37 °C.

Preparation of Samples for IMAC—Approximately 15 μg of phosphoprotein-enriched protein extract was suspended in 50 mM NH_4HCO_3 containing Promega modified trypsin (10 ng/μl). After overnight incubation at 37 °C digest solutions were concentrated to ~40 μl and desalted through 20-μl GELoader tips (Eppendorf) packed with 20 μl of OligoR3 resin (Applied Biosystems) as described previously (16). Peptides were washed once by 20 μl of 1% acetic acid and eluted by 20 μl of 50% ACN, 1% acetic acid. Elution solvent was removed by SpeedVac and replaced by 40 μl of the appropriate IMAC binding buffer (see below).

Preparation of Metal-chelating Resin for IMAC—A 50% slurry of POROS 20MC resin (Applied Biosystems) or Propac-IMAC resin (Dionex) was prepared in dH_2O. Resin was washed twice in water, stripped by 50 mM EDTA, washed in dH_2O, and then washed in 1% acetic acid. Resin was incubated in 100 mM iron chloride, 1% acetic acid for 15 min with agitation. Activated resin was rinsed in 30% ACN, 1% acetic acid and then in 1% acetic acid.

IMAC Enrichment Using GELoader Tips—30 μl of prepared IMAC resin was added to the peptide sample in binding buffer (100 mM NaCl in 1% acetic acid) and agitated for 15 min before loading into a GELoader tip. Resin was rinsed once by 20 μl of a solution of 100 mM NaCl in 30% acetonitrile and 1% acetic acid and eluted in two steps by 250 mM ammonium phosphate, pH 9. In the first step, 5 μl was pushed into the resin and left for 5 min to neutralize the remaining acetic acid. In the second step, 15 μl of the elution buffer was pushed through the resin. The 20 μl of phosphopeptide-carrying eluate was evaporated by SpeedVac, and the phosphopeptides were suspended in 20 μl of 5% acetonitrile in 0.1% acetic acid.

IMAC Enrichment Using Pierce Phosphopeptide Isolation Kit—Gallium-chelated columns were used according to the manufacturer's recommendation with the following modifications. Using 40 μl of a solution of 5% acetic acid as binding buffer, the tryptic peptides were transferred to a Pierce column, left for 15 min at room temperature, and then centrifuged for 1 min at 3000 $\times$ *g*. The Gallium-chelated disc was rinsed twice with 50 μl of 100 mM NaCl in 1% acetic acid, twice with 30% acetonitrile in 1% acetic acid, and once with dH_2O and then eluted by three additions of 20 μl of 250 mM ammonium phosphate, pH 9. Eluates were pooled and evaporated by SpeedVac, and phosphopeptides were suspended in 20 μl of 5% ACN, 0.1% acetic acid before MS analysis.

Q-TOF Analyses—A fraction of each tryptic digest (typically 5–10 μl) was analyzed by nano-LC-MS/MS using a Q-TOF Ultima hybrid quadrupole time-of-flight mass spectrometer coupled to a CapLC capillary HPLC system (Waters, Milford, MA). Digests were separated on a 75-μm-inner diameter $\times$ 150-mm Inertsil ODS 3, 5-μm nano-HPLC column (Dionex/LC Packings, Sunnyvale, CA) using the following gradient conditions: 5–60% ACN, 0.2% formic acid in 30 min, and 60–90% ACN in 5 min. The mass spectrometer was set to automatically acquire MS/MS spectra on doubly, triply, and quadruply charged ions. The MS survey scan range was 400–1600 *m/z*; the MS/MS range was 50–2000. After the first LC-MS/MS run, an exclusion file was created listing all previously analyzed ions that appeared to represent non-phosphorylated peptides, and a second fraction of each tryptic digest was injected with the mass spectrometer set to ignore the ions selected during the first run.

Data Processing—Database searching was carried out using Mascot Daemon™ (Matrix Science) against the National Center for Biotechnology Information (NCBI) genome sequence database. In addition to carbamidomethylcysteines and oxidized methionine, Mascot was asked to search for the potential presence of phosphate groups on serine, threonine, or tyrosine residues by looking in the MS/MS fragmentation spectra for the mass of a HPO_3 group on these residues or for ions resulting from a neutral loss of HPO_3 or H_3PO_4 (79.9663 and 97.9769 Da, respectively). All putative phosphopeptide MS/MS spectra identified by Mascot were manually verified for both the presence of a phosphate group and the peptide sequence. In many cases, the MS/MS spectrum allowed precise identification of the modified residue.

Identification of Phosphopeptides in Existing LC-MS/MS Data—Previously phosphoprotein-enriched samples were separated by 2D-

TABLE I
Peptides sequenced by MS/MS

Gene symbol indicates the protein identity that best matches the indicated peptide. Mod. indicates the identity and sequence position of the modified residue(s) in the context of the matching protein. In the deduced peptide sequences the sites of phosphorylation are indicated with the letter "p," and missed trypsin cleavage sites are indicated with a hyphen. ND, not determined; CAMK, calcium/calmodulin-dependent protein kinase.

Gene symbol	Mod.	Peptide	Rpts[a]	Notes[b]	Potential kinases[c]
1200015F23Rik	Thr-117	TLDSP**pT**HLR	1	c, f	
Acsl3 (*Facl3*)	Tyr-591	LQAGE**pY**VSLGK	1	b, f	
Btf3	Ser-105	QLTEMLP**pS**ILNQLGADSLTSLR	1	b, d	
Btf3	Thr-100	QL**pT**EMLPSILNQLGADSLTSLR	1	b, d	
Calr	Tyr-150	VHVIFN**pY**K	1	c, d	
Canx	Ser-582	AEEDEILNR-**pS**PR	2	a, b, e, f	
Carhsp1	Ser-2	**pS**SEPPPPPLQPPTHQTSVGLLDTPR	1	b, d	
Carhsp1	Ser-31 + Ser-33	DR**pS**P**pS**PLR	1	c, f	DYRK2
Carhsp1	Ser-42	GNVVP**pS**PLPTR	7 (4)	a, b, c, d, e, f	DYRK2
Carhsp1	Ser-53	TF**pS**ATVR	1	a, c, f	AKT1, AKT3
Cfl1	Ser-3	A**pS**GVAVSDGVIK + acetyl	3	a, b, c, d	LIMK1
Crmp4	Ser-522	G**pS**PTRPNPPVR	1	c, e	DYRK2
Csda	ND	SRPLNAVSQDGK	1	b, f	
D111	Ser-942	**pS**LVK-GMITAPK	1	b, f	PKC-δ*
Dap1	Ser-51	DK-DDQEWEST**pS**PPKPTVFISGVIAR	1 (2)	b, e	
Ddx21	Ser-99	EIITEEP**pS**EEEADMPKPK	1	b, e	
G3bp	Ser-231	ST**pS**PAPADVAPAQEDLR	3 (2)	a, b, c, d, e	
gi|63578930	Thr-110	VLDFEHFLPMLQ**pT**VAK	1	c, d	
Hdgf	Ser-165	AGDVLED**pS**PK	3	a, b, c, d, f	
Hnrpk	Ser-284	DYDDM**pS**PR	1	c, f	ERK1/ERK2
Hspb1	Ser-15	SP**pS**WEPFR	1 (2)	a, b, e	MAPKAPK2
Hspb1	Ser-86	QL**pS**SGVSEIR	4 (4)	a, b, c, d, e, f	MAPKAPK2
Hspca	Ser-262	ESDDKPEIEDVG**pS**DEEEEEK-K	2	a, b, c, f	CK2
Hspcb	Ser-225	EKEI**pS**DDEAEEEKGEK	1	a, c, d	CK2
Hspcb	Ser-254	IEDVG**pS**DEEDDSGK	3	a, b, c, d, f	CK2*
Lmna/c	Ser-22	SGAQASSTPL**pS**PTR	1	a, b, f	
Lmna/c	Ser-390	LR-L**pS**PSPTSQR	1	a, b, e	
Lmna/c	Ser-390 + Ser-392	LR-L**pS**P**pS**PTSQR	1	a, b, e	
Marcks	Ser-163	LSGF**pS**FK-K	1 (3)	a, b, e	PRK1, PKC
Mlp	Ser-103	LSGL**pS**FK-R	1 (2)	b, e	
Naca	Ser-132	IEDL**pS**QQAQLAAAEK	1	b, d	
Nedd5	Ser-218	IYHLPDAE**pS**DEDEDFKEQTR	1	c, f	CK2*
Npm1	Ser-82	M**pS**VQPTVSLGGFEITPPVVLR	1	b, d	
Pea15[d]	Ser-116	QP**pS**EEEIIK	3	a, b, d	CAMK2
Pea15	Thr-6	AEYG**pT**LLQDLTNNITLEDLEQLK	2	b, d	
Psma6	Ser-63	LLD**pS**STVTHLFK	3	b, c, d	
Psma6	Ser-64	LLDS**pS**TVTHLFK	1	c, d	
Psmb3	Ser-181	DAV**pS**GMGVIVHVIEK	1	b, d	
Rplp2	Ser-102 + Ser-105	K-EE**pS**EE**pS**DDDMGFGLFD	1	a, b, e	BARK1/GRK2
Rplp2	Ser-6	YVA**pS**YLLAALGGNSSPSAK	3	b, c, d	
Rplp2	Ser-6 + Tyr-7	YVA**pS**p**Y**LLAALGGNSSPSAK	2	b, c, d	
Rplp2	Ser-64	LA**pS**VPAGGAVAVSAAPGSAAPAAGSAPAAAEEK	3	b, c, d	
Rplp2	Tyr-3	**pY**VASYLLAALGGNSSPSAK	2	b, c, d	
Rps3	Thr-221	DEILPT**pT**PISEQK	2	a, b, c, f	

GE, and 190 proteins were identified by LC-MS/MS (13). Many of these proteins showed evidence of differential phosphorylation between the undifferentiated and differentiating states (13). These data was extensively reanalyzed against the latest available NCBInr (nonredundant) database using Mascot server software version 2.0.04 (Matrix Science) allowing up to two missed cleavages and the variable modifications protein N-terminal acetylation, carbamidomethylation, deamidation (NQ), oxidation (Met), pyro-Glu (N-terminal Gln), and phosphorylation (YST). All putative phosphopeptide MS/MS spectra were manually reviewed.

Reconstructing a Pathway Map from Phosphopeptides—For each experimentally determined phosphorylation site, any kinase(s) known to phosphorylate the site *in vivo* or *in vitro* was identified using the bioinformatics tools KinaSource (A. Knebel, Dundee, Scotland, UK) and PhosphoSite (Cell Signaling Technology, Danvers, MA). In a small number of cases where indicated, the relevant kinase was predicted using ScanSite (17). Then KinaSource was queried for other known substrates of each kinase. The process of mapping substrates to kinases was then repeated for all proteins in the model, including the kinases themselves. Database-derived results were verified and ex-

TABLE I—*continued*

Gene symbol	Mod.	Peptide	Rpts[a]	Notes[b]	Potential kinases[c]
Stmn1	Ser-24	ASGQAFELIL**pS**PR	2 (3)	a, b, c, d, e	GSK3, CAMK, CDC2
Stmn1	Ser-37	ESVPDFPL**pS**PPK	2 (3)	a, b, c, d, e	GSK3, CAMK, CDC2
Tebp	Ser-113	DWEDD**pS**DEDMSNFDR	3 (2)	a, b, c, d, e	CK2*
Tpm1	Ser-174	KLVIIE**pS**DLER	1	c, d	
Tpm1	Ser-45	QLEDELV**pS**LQK	1	c, d	
Tuba2	Ser-275	AVCML**pS**NTTAIAEAWAR	1	c, d	
Tubb5	Thr-108	MAV**pT**FIGNSTAIQELFK	2	c, d	
Unnamed	ND	**pS**GIVVK-DVSIK	1	b, e	
XP_489094	Ser-27 + Thr-30	**pS**EE**pT**VELR	1	b, f	

[a] Rpts indicates the number of independently prepared phosphoprotein-enriched samples that produced spectra matching this phosphopeptide sequence. Numbers in parentheses represent the number of repeated observations of the peptide in independent IMAC experiments done on the same protein sample.

[b] Notes: a, phosphorylation of this site has previously been reported; b, this phosphopeptide was detected in undifferentiated C2C12 cell samples; c, this phosphopeptide was detected in differentiating C2C12 cell samples; d, this phosphopeptide was detected after phosphoprotein enrichment and two-dimensional gel electrophoresis; e, this phosphopeptide was detected after phosphoprotein enrichment and IMAC; f, this phosphopeptide was detected after phosphoprotein enrichment and one-dimensional SDS-PAGE.

[c] Potential kinases are those that have been shown to phosphorylate the specific site in other studies or that are predicted to phosphorylate the target site; predicted interactions are annotated with an asterisk (*).

[d] This phosphopeptide was reported previously by us elsewhere (13) but is included here for completeness.

tended by examination of the primary research literature. The identified substrate-kinase interactions were then diagrammed using the bioinformatics tool HubView (18).

RESULTS AND DISCUSSION

53 phosphopeptides were identified from C2C12 cell cultures undergoing growth or differentiation (Table I; see Supplements 1 and 2 for additional details). 17 phosphopeptides were identified after IMAC, 17 after one-dimensional gel electrophoresis (Fig. 1), and 30 after 2D-GE and LC-MS/MS. Only nine phosphopeptides were detected by multiple methods suggesting that each method introduces a bias into the selection of peptides. The apparently greater success of the 2D-GE method may indicate a benefit from greater sample fractionation or may simply reflect the larger number of such experiments performed. All phosphopeptides listed in Table I were matched to known or predicted proteins, and in 43 cases additional non-phosphorylated peptides were identified that supported the identification. Some phosphorylations were detected only in either growing or differentiating cells (Table II). Of the 55 identified phosphopeptides, at least 21 contained phosphorylation sites that have been reported previously.

A major challenge in systems biology is to transform large datasets into biologically relevant interpretations. To address this problem, we chose to use a novel bioinformatics approach. We reasoned that because phosphopeptides are the product of protein kinase activity we could construct a proposed kinase-substrate interaction network based upon our data. To perform this task, 19 experimentally detected phosphorylations for which the responsible kinase was known or strongly predicted were used as the "bootstrap" to create an initial network of kinase-substrate interactions (Fig. 2*A*; see also Supplement 3). Then other known substrates of those kinases, as well as interactions upstream of each kinase, were determined and incorporated into the model (Fig. 2*B*). For clarity, the abundant and promiscuous kinases PKC and CK2 were excluded from the model, although PKC was included in cases where phosphorylation was attributed to a specific PKC isoform.

The generally accepted view of myogenic signaling is that AKT, p38, MAPKAPK2, and CDK5 activity promote cell cycle exit and myotube formation, whereas ERK1/2 and GSK3 activity are associated with cell cycling and maintenance of the undifferentiated state. Of 114 kinases listed in KinaSource, only 21 appeared in our extrapolated network, and satisfyingly, this group prominently featured AKT, CDK5, GSK3, and MAPKAPK2 kinases (Fig. 2*C*). Furthermore some kinase-substrate interactions known to be relevant in myogenic signaling such as CDK5-myocyte enhancer factor 2 and AKT-FOXO1A were inferred by the model. Whereas the list of phosphopeptides alone provided little biological context, the extrapolated kinase-substrate network suggested a number of features of interest, such as the implication of DYRK kinases as promyogenic factors.

DYRK—DYRK interactions were prominent in the model network. Induction of DYRK family kinases is required for muscle differentiation (19) at least in part via DYRK-mediated phosphorylation and regulation of histone deacetylases (20). Here we found that the putative DYRK substrates CRHSP24 and CRMP4 are also phosphorylated in C2C12 (Table I and Fig. 3). Bioinformatics modeling revealed five more DYRK-substrate interactions (p27kip, p21waf, FOXO1A, glycogen

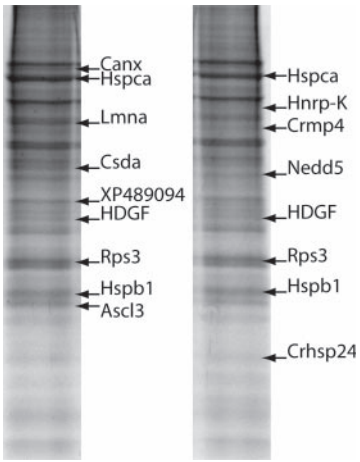

FIG. 1. **After phosphoprotein enrichment and one-dimensional SDS-PAGE, proteins were detected by G-250 Coomassie Blue stain.** *Labeled arrows* indicate identified proteins from which phosphopeptide MS/MS spectra were obtained.

TABLE II
Undifferentiated versus differentiating cells

Phosphopeptides that were detected in samples from only either undifferentiated or differentiated C2C12 but not under both conditions are listed separately. In some cases different phosphopeptides from the same protein were detected depending on the cell differentiation state, and the specific site of phosphorylation is given in these instances. Kinases proposed to be responsible for each phosphorylation are given in parentheses, MARCKS, myristoylated alanine rich protein kinase C substrate; MLP, MARCKS-like protein; NACA, nascent polypeptide-associated complex; CANX, calnexin; CALR, calreticulin; HSPCB, heat shock protein class b; CSDA, cold shock domain protein A.

Undifferentiated	Differentiating
ACSL3	CALR
BTF3	CRHSP24 Ser-31 (DYRK2)
CANX	CRHSP24 Ser-33 (DYRK2)
CRHSP24 Ser-2	CRHSP24 Ser-53 (AKT)
CSDA	CRMP4 (DYRK2)
D111 (PKC-δ)	HNRNPK (ERK1/ERK2)
DAP1	HSPCB Ser-225 (CK2)
DDX21	NEDD5 (CK2)
HSPB1 Ser-15 (MAPKAPK2)	Tropomyosin
Lamin	Tubulin
MARCKS (PKC, PRK1)	
MLP (PKC)	
NACA	
NPM1	
PEA-15 (CAMK2)	
PSMB3	
RPLP2 (GRK2)	
XP_489094	

synthase, and STAT3), suggesting additional mechanisms by which DYRK may contribute to myogenesis. Two of the phosphorylation sites ascribed to DYRK2 activity are thought to be priming sites for GSK3. In CRMP4 (Dpysl3, Drp3) Ser-522 appears to represent the priming site for a series of GSK3-mediated phosphorylations at serines 518, 514, and 509 (21). In C2C12 cells both site-specific (Table I) and general phosphorylation (13) of CRMP4 was detected only in differentiating cells, whereas Ser-522 is constitutively phosphorylated in brain. Although CRMP4 has been studied almost exclusively in the context of neurons, these data and the observation that the CRMP4 promoter contains a MyoD/myogenin binding site (22) indicate that CRMP proteins likely play an important role in muscle differentiation. eIF2B (eukaryotic translation initiation factor 2B) is also a substrate of both DYRK and GSK3. DYRK-mediated phosphorylation at Ser-539 is thought to permit subsequent phosphorylation of Ser-535 by GSK3 leading to repression of translation (23). Because muscle hypertrophy is associated with GSK3 inhibition and eIF2 derepression, this suggests that promyogenic functions of DYRK could be blocked by GSK3 activity. These observations are consistent with the model that DYRK activity is increased and GSK3 is repressed during myogenesis. In agreement with this principle, several putative GSK3 substrate proteins that were not DYRK substrates, pyruvate dehydrogenase and CRMP2, showed evidence of phosphorylation only in non-differentiated myoblasts, although specific phosphorylation sites were not determined (13).

CRHSP24—CaRHSP24/CRHSP24 is phosphoprotein of unknown function. DYRK2 can phosphorylate CRHSP24 at serines 31 and 33 *in vitro*, and Ser-43 (equivalent to Ser-42 in mouse) is phosphorylated in human embryonic kidney 293 cells (24). We detected phosphorylation at serines 31, 33, and 53 in differentiating C2C12 and constitutive phosphorylation at Ser-42 (Table I). Phosphorylation of CRHSP24 at Ser-53, ascribed to AKT activity, was also detected. CRHSP24 may play an important role in myogenesis. IGF-1 treatment, which stimulates muscle hypertrophy, stimulates phosphorylation of the Ser-53 site in human embryonic kidney 293 cells. CRHSP24 is a calcineurin substrate (25). Calcineurin activity is important in normal muscle development, and augmented calcineurin activation has been shown to limit the pathogenesis of muscular dystrophies (26).

PEA-15—Our previous observation of PEA-15 phosphorylation in C2C12 (13) was confirmed and extended with evidence of a second phosphorylation site (Table I). Ser-116 phosphorylated PEA-15 was repeatedly observed in growing cells but never in differentiating cells. This observation is consistent with the recent report that the Ser-104/Ser-116 dephosphorylated form of PEA-15 promotes cell cycle arrest via sequestration of ERK (27), and indeed, reduced ERK activity is a prerequisite for myoblast differentiation (28).

PKC-δ—A potential role for PKC-δ was inferred in the model (Fig. 2*C*), and the potential PKC-δ substrates 14-3-3, HMGA1 (high mobility group AT-hook 1), and HNRNPK were detected in phosphoprotein enrichment experiments (13). Although site-specific evidence of PKC-δ-mediated phosphorylation is lacking, these observations are interesting in light of a recent report that constitutively active PKC-δ promotes whereas kinase-inactive PKC-δ inhibits IGF-1-mediated pro-

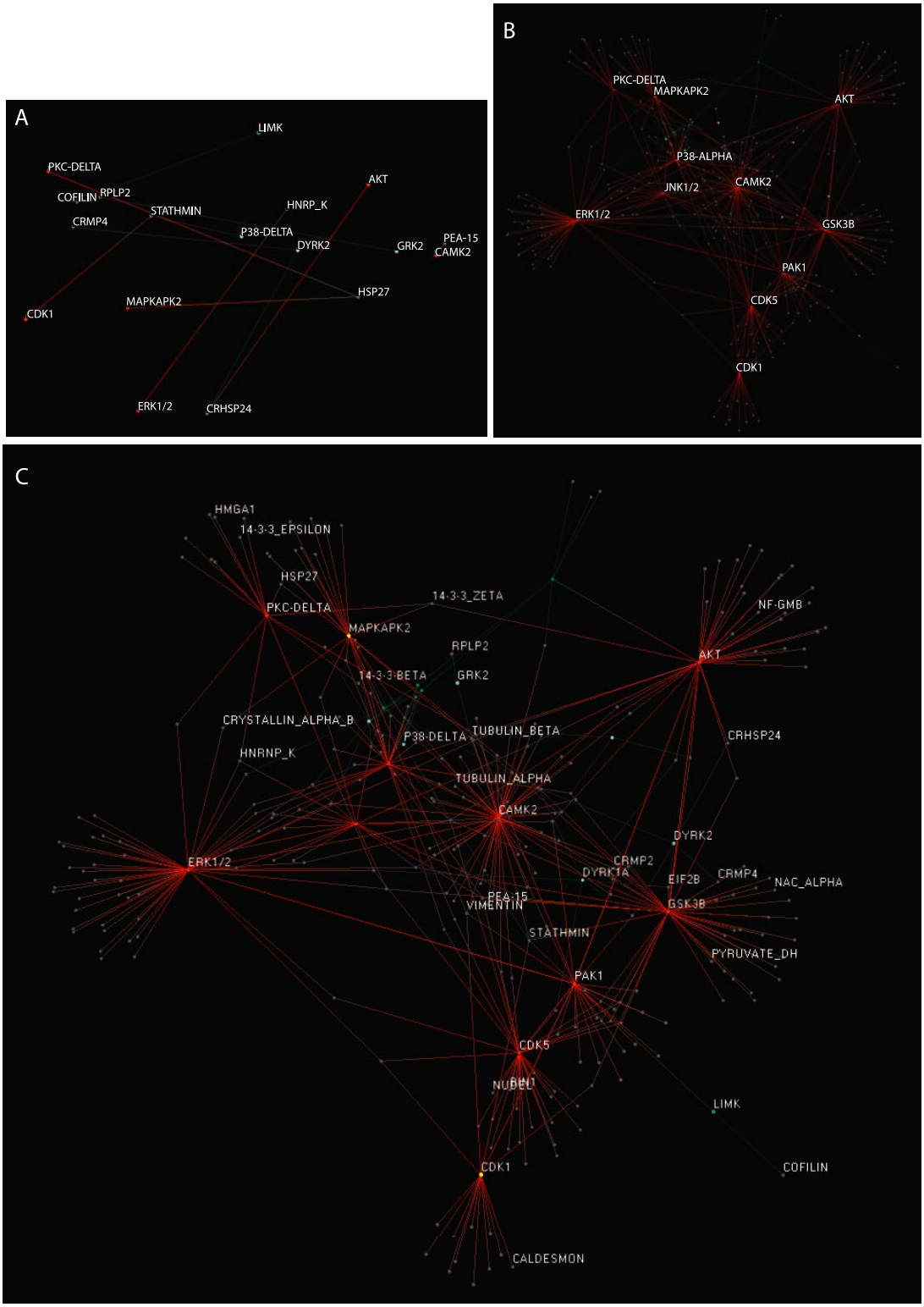

FIG. 2. **Graphical representations of kinase-substrate interaction pathways made using the bioinformatics visualization tool HubView (18).** *A*, experimentally detected phosphopeptides were matched to the corresponding kinases based on database and literature searches. Kinases are shown in *blue* (<15 substrates) or *orange* (>15 substrates), and non-kinase proteins are colored *gray*. *B*, the model network extended to include all substrates and upstream regulators of each implicated kinase. Major hub kinases are labeled. *C*, the model network with experimentally detected phosphoproteins from Table I and previous work (13) and their upstream kinases labeled. Additional visualizations can be found in Supplement 3. LIMK, LIM domain containing, protein kinase; NACA, nascent polypeptide-associated complex α; DH dehydrogenase.

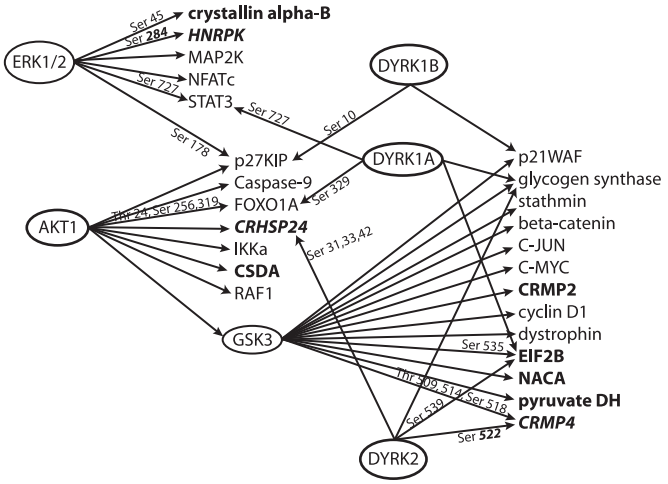

FIG. 3. **Detailed view of DYRK interactions.** Specific sites of phosphorylation are indicated where known. Proteins experimentally determined to be phosphorylated (13) are given in *bold type*, and proteins experimentally determined to be phosphorylated at specific sites (Table I) are given in *italic type*. The current model for kinase signaling in early myogenesis is that AKT is activated, GSK3 is consequently repressed, ERK activity is low, and DYRK expression and activity is increased. IKKa, inhibitor of κB kinase α; NFATc, nuclear factor of activated T cells; CSDA, cold shock domain protein A.

liferation of myoblasts (29). PKC-δ, as well as DYRK1a and ERK1/ERK2, can also phosphorylate STAT3, which is reported to promote myoblast proliferation and inhibit MyoD function and muscle differentiation (30–32). Because the aforementioned substrates contain multiple phosphorylation sites, additional studies will be required to determine the precise role of PKC.

Limitations to Application and Interpretation—Because the experimental data cover only a very small proportion of the phosphoproteome, it is important to consider the validity of extrapolating a model kinase network under these circumstances. The strategy used relies on three assumptions: that substrates have a limited number of kinases, that kinases have a large number of substrates, and that the overall topology of the kinase-substrate network is scale-free. First, if a substrate has many potential kinases then inference of specific kinase activities is not possible. Second, if a kinase has few substrates then the odds that one of those substrates will be found experimentally is extremely low. The last assumption determines whether the model is likely to encompass major features of the true *in vivo* network or if it reveals only local details. In a non-scale-free network a small data sample will show only local features, much as mapping a few roads in a city would reveal nothing about the overall layout of the traffic flow. However in a scale-free network, any two points on the network are connected by a very small number of interactions, and the overall picture emerges quickly, just as a sampling of commercial airline flights would quickly reveal the existence of traffic hubs at the major airports. Work on kinase-substrate interactions in yeast supports these three assumptions. An exhaustive evaluation of 1325 phosphoproteins and 87 kinases by Ptacek *et al.* (15) found that 73% of substrates were associated with only one or two kinases but that most kinases had multiple substrates. Analysis of the kinase-substrate subset of the extensive protein-protein interaction data available for yeast confirmed that this network is scale-free (18).

Another consideration was protein abundance. If myogenesis resulted in altered protein phosphorylation in only a subset of very low abundance proteins against a large backdrop of abundant invariant phosphoproteins then no relevant results would have been obtained. However, we found previously that the subset of the phosphoproteome that is sufficiently abundant for silver stain detection on gels does in fact undergo extensive alteration during early myogenesis (13). The use of data obtained from two-dimensional gels was especially valuable in retaining a biologically relevant focus. Comparative analysis of gels prepared from growing *versus* differentiating myocytes allowed us to focus the MS analysis on proteins that exhibited evidence of altered post-translational modification during myogenesis. All of these factors contributed to the ability to produce a model network that exhibited relevant features of interest. However, a number of caveats remain. First, because proteome coverage is incomplete, lack of detection of phosphorylation cannot be taken as strong evidence of dephosphorylation. Second, the results are biased toward the detection of relatively abundant peptides. Third, kinase-substrate databases are far from complete, although the amount of protein-protein interaction data available is increasing rapidly. We anticipate that increased coverage should not invalidate or contradict the results presented here. Because kinase networks appear to be scale-free (18), increased data coverage would likely increase the number of substrates more rapidly than the number of kinases, limiting any changes in the overall topology of the network. However, this remains to be experimentally demonstrated. Another challenge remaining to be addressed is the role of regulated phosphatases in the network. At this time inclusion of a comprehensive phosphatase dataset is not possible given the limited experimental observations in the literature combined with the absence of robust screening tools for phosphatase activity or dephosphorylation events.

Conclusion—Experimental phosphopeptide data and bioinformatics databases of kinase-substrate interactions were used to produce a model signaling network for C2C12 myogenesis. From a relatively small number of phosphopeptides, a complex network was readily constructed that inferred the existence of several known features of myogenic signaling. In addition, places for newly described phosphorylations in signaling cascades were predicted. New protein-protein interaction datasets including kinase-substrate datasets are being produced at an increasing rate, and the utility of this approach is certain to increase as more data become available.

Acknowledgments—We thank John F. Kelly for invaluable advice and discussions and Douglas J. Borris, Wen Ding, and Luc Tessier for expert technical assistance.

* This work was supported in part by grants from the Heart and Stroke Foundation of Canada and the Muscular Dystrophy Association (United States) (to L. A. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

¶ Supported by funding from the Genomic and Health Initiative.

** Holds the Mach-Gaennelsen Chair in Cardiac Research. To whom correspondence should be addressed: Ontario Genomics Innovation Centre, Ottawa Health Research Institute, 501 Smyth Rd., Ottawa, Ontario K1H 8L6, Canada. Tel.: 613-737-8899 (ext. 78618); Fax: 613-737-8803; E-mail: lmegeney@ohri.ca.

REFERENCES

1. Huh, M. S., Smid, J. K., and Rudnicki, M. A. (2005) Muscle function and dysfunction in health and disease. *Birth Defects Res.* **75,** 180–192
2. Chargé, S. B., and Rudnicki, M. A. (2004) Cellular and molecular regulation of muscle regeneration. *Physiol. Rev.* **84,** 209–238
3. Cuenda, A., and Cohen, P. (1999) Stress-activated protein kinase-2/p38 and a rapamycin-sensitive pathway are required for C2C12 myogenesis. *J. Biol. Chem.* **274,** 4341–4346
4. Zetser, A., Gredinger, E., and Bengal, E. (1999) p38 mitogen-activated protein kinase pathway promotes skeletal muscle differentiation. Participation of the Mef2c transcription factor. *J. Biol. Chem.* **274,** 5193–5200
5. de Angelis, L., Zhao, J., Andreucci, J. J., Olson, E. N., Cossu, G., and McDermott, J. C. (2005) Regulation of vertebrate myotome development by the p38 MAP kinase-MEF2 signaling pathway. *Dev. Biol.* **283,** 171–179
6. Han, J., Jiang, Y., Li, Z., Kravchenko, V. V., and Ulevitch, R. J. (1997) Activation of the transcription factor MEF2C by the MAP kinase p38 in inflammation. *Nature* **386,** 296–299
7. Lluis, F., Perdiguero, E., Nebreda, A. R., and Munoz-Canoves, P. (2006) Regulation of skeletal muscle gene expression by p38 MAP kinases. *Trends Cell Biol.* **16,** 36–44
8. Rommel, C., Bodine, S. C., Clarke, B. A., Rossman, R., Nunez, L., Stitt, T. N., Yancopoulos, G. D., and Glass, D. J. (2001) Mediation of IGF-1-induced skeletal myotube hypertrophy by PI(3)K/Akt/mTOR and PI(3)K/Akt/GSK3 pathways. *Nat. Cell Biol.* **3,** 1009–1013
9. Hoffman, E. P., and Nader, G. A. (2004) Balancing muscle hypertrophy and atrophy. *Nat. Med.* **10,** 584–585
10. Kolodziejczyk, S. M., Walsh, G. S., Balazsi, K., Seale, P., Sandoz, J., Hierlihy, A. M., Rudnicki, M. A., Chamberlain, J. S., Miller, F. D., and Megeney, L. A. (2001) Activation of JNK1 contributes to dystrophic muscle pathogenesis. *Curr. Biol.* **11,** 1278–1282
11. Lazaro, J. B., Kitzmann, M., Poul, M. A., Vandromme, M., Lamb, N. J., and Fernandez, A. (1997) Cyclin dependent kinase 5, cdk5, is a positive regulator of myogenesis in mouse C2 cells. *J. Cell Sci.* **110,** 1251–1260
12. Philpott, A., Porro, E. B., Kirschner, M. W., and Tsai, L. H. (1997) The role of cyclin-dependent kinase 5 and a novel regulatory subunit in regulating muscle differentiation and patterning. *Genes Dev.* **11,** 1409–1421
13. Puente, L. G., Carriere, J. F., Kelly, J. F., and Megeney, L. A. (2004) Comparative analysis of phosphoprotein-enriched myocyte proteomes reveals widespread alterations during differentiation. *FEBS Lett.* **574,** 138–144
14. Bork, P., Jensen, L. J., von Mering, C., Ramani, A. K., Lee, I., and Marcotte, E. M. (2004) Protein interaction networks from yeast to human. *Curr. Opin. Struct. Biol.* **14,** 292–299
15. Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., Guo, H., Jona, G., Breitkreutz, A., Sopko, R., McCartney, R. R., Schmidt, M. C., Rachidi, N., Lee, S. J., Mah, A. S., Meng, L., Stark, M. J., Stern, D. F., De Virgilio, C., Tyers, M., Andrews, B., Gerstein, M., Schweitzer, B., Predki, P. F., and Snyder, M. (2005) Global analysis of protein phosphorylation in yeast. *Nature* **438,** 679–684
16. Stensballe, A., Andersen, S., and Jensen, O. N. (2001) Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics* **1,** 207–222
17. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* **31,** 3635–3641
18. Lee, R. E., and Megeney, L. A. (2005) The yeast kinome displays scale free topology with functional hub clusters. *BMC Bioinformatics* **6,** 271
19. Deng, X., Ewton, D. Z., Pawlikowski, B., Maimone, M., and Friedman, E. (2003) Mirk/dyrk1B is a Rho-induced kinase active in skeletal muscle differentiation. *J. Biol. Chem.* **278,** 41347–41354
20. Deng, X., Ewton, D. Z., Mercer, S. E., and Friedman, E. (2005) Mirk/dyrk1B decreases the nuclear accumulation of class II histone deacetylases during skeletal muscle differentiation. *J. Biol. Chem.* **280,** 4894–4905
21. Cole, A. R., Knebel, A., Morrice, N. A., Robertson, L. A., Irving, A. J., Connolly, C. N., and Sutherland, C. (2004) GSK-3 phosphorylation of the Alzheimer epitope within collapsin response mediator proteins regulates axon elongation in primary neurons. *J. Biol. Chem.* **279,** 50176–50180
22. Matsuo, T., Stauffer, J. K., Walker, R. L., Meltzer, P., and Thiele, C. J. (2000) Structure and promoter analysis of the human unc-33-like phosphoprotein gene. E-box required for maximal expression in neuroblastoma and myoblasts. *J. Biol. Chem.* **275,** 16560–16568
23. Woods, Y. L., Cohen, P., Becker, W., Jakes, R., Goedert, M., Wang, X., and Proud, C. G. (2001) The kinase DYRK phosphorylates protein-synthesis initiation factor eIF2B[cepsilon] at Ser539 and the microtubule-associated protein tau at Thr212: potential role for DYRK as a glycogen synthase kinase 3-priming kinase. *Biochem. J.* **355,** 609–615
24. Auld, G. C., Campbell, D. G., Morrice, N., and Cohen, P. (2005) Identification of calcium-regulated heat-stable protein of 24 kDa (CRHSP24) as a physiological substrate for PKB and RSK using KESTREL. *Biochem. J.* **389,** 775–783
25. Groblewski, G. E., Yoshida, M., Bragado, M. J., Ernst, S. A., Leykam, J., and Williams, J. A. (1998) Purification and characterization of a novel physiological substrate for calcineurin in mammalian cells. *J. Biol. Chem.* **273,** 22738–22744
26. St-Pierre, S. J., Chakkalakal, J. V., Kolodziejczyk, S. M., Knudson, J. C., Jasmin, B. J., and Megeney, L. A. (2004) Glucocorticoid treatment alleviates dystrophic myofiber pathology by activation of the calcineurin/NF-AT pathway. *FASEB J.* **18,** 1937–1939
27. Krueger, J., Chou, F. L., Glading, A., Schaefer, E., and Ginsberg, M. H. (2005) Phosphorylation of phosphoprotein enriched in astrocytes (PEA-15) regulates extracellular signal-regulated kinase-dependent transcription and cell proliferation. *Mol. Biol. Cell* **16,** 3552–3561
28. Dorman, C. M., and Johnson, S. E. (1999) Activated Raf inhibits avian myogenesis through a MAPK-dependent mechanism. *Oncogene* **18,** 5167–5176
29. Czifra, G., Toth, I. B., Marincsak, R., Juhasz, I., Kovacs, I., Acs, P., Kovacs, L., Blumberg, P. M., and Biro, T. (2006) Insulin-like growth factor-I-coupled mitogenic signaling in primary cultured human skeletal muscle cells and in C2C12 myoblasts. A central role of protein kinase Cδ. *Cell. Signal.* **18,** 1461–1472
30. Jo, C., Kim, H., Jo, I., Choi, I., Jung, S. C., Kim, J., Kim, S. S., and Jo, S. A. (2005) Leukemia inhibitory factor blocks early differentiation of skeletal muscle cells by activating ERK. *Biochim. Biophys. Acta* **1743,** 187–197
31. Kami, K., and Senba, E. (2002) In vivo activation of STAT3 signaling in satellite cells and myofibers in regenerating rat skeletal muscles. *J. Histochem. Cytochem.* **50,** 1579–1589
32. Kataoka, Y., Matsumura, I., Ezoe, S., Nakata, S., Takigawa, E., Sato, Y., Kawasaki, A., Yokota, T., Nakajima, K., Felsani, A., and Kanakura, Y. (2003) Reciprocal inhibition between MyoD and STAT3 in the regulation of growth and differentiation of myoblasts. *J. Biol. Chem.* **278,** 44178–44187