Physiol Genomics 42: 406–419, 2010.
First published May 18, 2010; doi:10.1152/physiolgenomics.00055.2010.

Evolution of the regulatory control of vertebrate striated muscle: the roles of troponin I and myosin binding protein-C

Justin F. Shaffer[1] and Todd E. Gillis[2]

[1]*Department of Bioengineering, University of Washington, Seattle, Washington; and* [2]*Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada*

Submitted 22 March 2010; accepted in final form 14 May 2010

Shaffer JF, Gillis TE. Evolution of the regulatory control of vertebrate striated muscle: the roles of troponin I and myosin binding protein-C. *Physiol Genomics* 42: 406–419, 2010. First published May 18, 2010; doi:10.1152/physiolgenomics.00055.2010.—Troponin I (TnI) and myosin binding protein-C (MyBP-C) are key regulatory proteins of contractile function in vertebrate muscle. TnI modulates the Ca^{2+} activation signal, while MyBP-C regulates cross-bridge cycling kinetics. In vertebrates, each protein is distributed as tissue-specific paralogs in fast skeletal (fs), slow skeletal (ss), and cardiac (c) muscles. The purpose of this study is to characterize how TnI and MyBP-C have changed during the evolution of vertebrate striated muscle and how tissue-specific paralogs have adapted to different physiological conditions. To accomplish this we have completed phylogenetic analyses using the amino acid sequences of all known TnI and MyBP-C isoforms. This includes 99 TnI sequences (fs, ss, and c) from 51 different species and 62 MyBP-C sequences from 26 species, with representatives from each vertebrate group. Results indicate that the role of protein kinase A (PKA) and protein kinase C (PKC) in regulating contractile function has changed during the evolution of vertebrate striated muscle. This is reflected in an increased number of phosphorylatable sites in cTnI and cMyBP-C in endothermic vertebrates and the loss of two PKC sites in fsTnI in a common ancestor of mammals, birds, and reptiles. In addition, we find that His^{132}, Val^{134}, and Asn^{141} in human ssTnI, previously identified as enabling contractile function during cellular acidosis, are present in all vertebrate cTnI isoforms except those from monotremes, marsupials, and eutherian mammals. This suggests that the replacement of these residues with alternative residues coincides with the evolution of endothermy in the mammalian lineage.

phosphorylation; regulatory proteins; contractility

THROUGH EVOLUTIONARY PROCESSES the vertebrate heart has become increasingly complex. This is reflected in changes in anatomy as well as in the complexity of the mechanisms that regulate cardiac function. In early proto-vertebrates the heart was a simple muscular tube with pulsatile contraction (3). Through the development of valves and septa, the heart became a multichambered organ in fishes with a single atrium and a single ventricle, where cardiac output is primarily regulated by changes in stroke volume (6, 15, 54). Further morphological changes, as found in hearts of crocodilian and noncrocodilian reptiles, began to enable the separation of the pulmonary and systemic circulatory systems (33). In endothermic species, such as humans and chickens, the heart has four contractile chambers with complete isolation of pulmonary and systemic circulation and heart rates that can exceed 600 beats per minute (6, 15, 60). It is changes in heart rate, instead of stroke volume, that are the primary regulator of cardiac output in these endothermic animals (54). Despite these changes in function and morphology, neither the basic mechanism responsible for triggering muscle contraction nor the complement of contractile proteins responsible for this phenomenon have changed. However, contractile protein isoforms do differ in primary amino acid sequence, which translates into functional differences in the sarcomere and give rise to differences in the physiological capability between hearts of different species (18).

Evolutionary pressures have also driven the specialization of the different muscle types (fast skeletal, slow skeletal, and cardiac) within the same species. For example, one significant difference in the function of cardiac and fast skeletal muscle is in myocyte recruitment during contraction. Every time the heart beats each myocyte contracts while in fast skeletal muscle, myocytes are activated as needed to generate the required force (23). As every myocyte contracts with each heart beat, additional control mechanisms have evolved within the vertebrate heart that allow for the graded activation of the contractile reaction. This enables significant modulation of the rate and strength of cardiac contraction, and as a result, cardiac output. In mammals, these control mechanisms include the Frank-Starling response to sarcomere length, changes to the rate and magnitude of the Ca^{2+} transient, as well as the modulation of specific regulatory proteins via phosphorylation (4, 23, 37). While the Ca^{2+} sensitivity of fast skeletal muscle is affected by sarcomere length, function is not regulated to the same degree is as cardiac muscle (2).

Two proteins that are critical to the regulation of striated muscle function at the level of the sarcomere are troponin I (TnI), and myosin binding protein-C (MyBP-C). TnI is a component protein of the troponin (Tn) complex, the Ca^{2+} activated trigger of the contractile reaction. MyBP-C is found in the A-bands of striated muscle sarcomeres where it binds myosin and/or actin to regulate cross-bridge cycling (51, 57, 61). Comparison of the tissue-specific paralogs of these two regulatory proteins between different muscle types in humans reveals significant differences in their sequences as well as in the role that they play in regulating muscle function. For example, each of these proteins in the heart is a potential target for phosphorylation (42, 50, 55). Such posttranslational modification results in changes in the Ca^{2+} sensitivity and kinetics of force development and leads to changes in the rate and strength of cardiac contraction (31, 57, 62). Protein kinase A (PKA), activated by β-adrenergic stimulation, targets two residues on human cardiac (c) TnI (46) and four residues on human cardiac (c) MyBP-C (30). Protein kinase C (PKC), activated by α-adrenergic stimulation, can target five residues on human cTnI (42) as well as residues on cMyBP-C (38, 39,

Address for reprint requests and other correspondence: T. E. Gillis, Dept. of Integrative Biology, Univ. of Guelph, Guelph, ON, N1G-2W1 Canada (e-mail: tgillis@uoguelph.ca).

69). Interestingly, there is no evidence that the fast and slow skeletal isoforms of human TnI or MyBP-C are phosphorylated in vivo by either PKA or PKC (24, 58, 59, 65). This suggests that there are differences in how cardiac muscle is regulated by the phosphorylation of TnI and MyBP-C compared with fast and slow skeletal muscle.

The goal of this study is to characterize how the primary amino acid sequences of TnI and MyBP-C have changed during the evolution of vertebrate striated muscle, and how tissue-specific paralogs have adapted to different physiological conditions. In particular, we investigated how the regulatory control mechanisms of striated muscle have developed during vertebrate evolution and the evolution of different muscle types. This work was accomplished by integrating phylogenetic analysis of all known amino acid sequences of TnI and MyBP-C with known functional characteristics of these two proteins. Included within this analysis are protein sequences from tunicates (urochordates), fishes, amphibians, reptiles, birds, the platypus, *Ornithorhynchus anatinus*, (monotreme), an opossum, *Monodelphis domestica*, (marsupial), and a large number of eutherian mammals. This collection of sequences from all vertebrate lineages enabled a comprehensive analysis of how TnI and MyBP-C have evolved within the different muscle types during vertebrate evolution. Through this synthesis we have gained new insight into how the regulation of striated muscle has evolved during the specialization of different muscle types. In addition, by examining the sequence alignments of each protein from such a broad array of species, we have identified patterns in the conservation of specific residues across the different orthologs and paralogs of each protein. These sequence motifs and specific residues represent potential targets for future studies into molecular interactions that regulate the function of these proteins.

METHODS

All sequences used in this study were obtained from either the National Center for Biotechnology Information (Bethesda, MD) nonredundant protein database (http://www.ncbi.nlm.nih.gov/guide/proteins/), the Ensembl Genome Browser (Wellcome Trust Genome Campus, Hinxton, Cambridge UK) (http://uswest.ensembl.org/index.html), or GenBank (http://www.ncbi.nlm.nih.gov/genbank/). For this study we obtained 147 TnI sequences from 52 different species and 139 sequences of MyBP-C from 44 different species. All efforts were made to get paralogs from fast skeletal muscle, slow skeletal muscle, and cardiac muscle for each species. This was not always possible as not all have been identified or annotated. For a number of the fish species there were multiple tissue-specific paralogs. A single isoform was chosen for such species by aligning the multiple sequences one at a time to the closest relative and noting the percent identity: the sequence with the highest percent identity was selected and used for further analysis. Also, a number of sequences contained missing data (Xs in sequence) and were therefore excluded from analysis. The final number of sequences used was 99 for TnI from 51 different species and 62 for MyBP-C from 26 different species. A complete list of all the TnI and MyBP-C sequences used in this study are shown in Tables 1 and 2 (see Supplementary Data for FASTA amino acid sequences).[1]

Sequences were aligned and neighbor-joining trees (47) were generated using ClustalX2 (version 2.0.12) (36). Further phylogenetic analyses were performed with the PHYLIP package (version 3.6b; Joe Felsenstein, Department of Genome Sciences, University of Washington, Seattle, WA) (16). One hundred boot-strapped replicates of the data were created with Seqboot (PHYLIP). Maximum likelihood trees (using the PHYLIP program Proml) were created for all boot-strapped replicates with the Jones-Taylor-Thronton probability model, global rearrangements, and randomized order of sequence input (jumbled). Consensus trees were created with Consense (PHYLIP) using majority rule (extended) restrictions. All trees were rooted with midpoint rooting where the root is placed halfway between the two most distant taxa. Values at nodes represent the number of times (out of 100) that the specific branch point occurred. The program Figtree (version 1.2.2) (45) was used to examine and display all trees.

[1] The online version of this article contains supplemental material.

RESULTS

Phylogenetic analysis of TnI and MyBP-C. The topology of the phylogenetic tree created for TnI reveals that there are separate clades for the sequences from each tissue type (Fig. 1). In the TnI tree, cTnI is monophyletic with ssTnI, which both appear to have diverged from fast skeletal (fs) TnI, the ancestral form of TnI. The phylogenetic analysis of TnI demonstrates that there are three tissue-specific clades of TnI. The exception to this pattern is that cTnI from all fish species, including the zebrafish (*Danio rerio*), group with the ssTnI clade of the phylogenetic tree (Fig. 1). This suggests that in fish there has been less divergence in the genes that encode for these two tissue-specific paralogs. Comparison of cTnI from all fish to that of other vertebrate species reveals that the NH_2-terminal peptide found in most vertebrate cTnI is missing in cTnI from all fish species (see Fig. 3). It is this difference that likely makes fish cTnI isoforms more similar to slow skeletal (ss) TnI isoforms. Due to this similarity it is possible that some TnI isoforms that have been sequenced from the genomic DNA of fish species and identified as ssTnI may actually be expressed in the heart.

Phylogenetic analysis of MyBP-C sequences resulted in monophyletic clades for each of the three tissue-specific isoforms (Fig. 2). However, the MyBP-C tree suggests a different order for the separation of the tissue specific clades compared with the TnI tree (Fig. 1). Here, cMyBP-C paralogs are monophyletic, and the fast skeletal and slow skeletal paralogs cluster in a group that diverged from cMyBP-C, suggesting that cMyBP-C was the ancestral form of MyBP-C. One explanation for this difference, compared with the TnI tree, is that there were less MyBP-C sequences to be used in the tree building and these sequences did not cover the same range of vertebrate species. To test this we constructed new trees using only TnI and MyBP-C sequences that were found in the same species and tissues. While this reduced the number of sequences used in the analysis, the same result was found (data not shown).

Within each of the tissue-specific clades for both proteins, there is delineation between the sequence clusters for the different vertebrate groups. For example, within the mammals there is clear separation between cTnI and cMyBP-C from the platypus (monotreme), the opossum (marsupial), and eutherian mammals (Figs. 1 and 2). In addition, mammals are clearly separated from birds and other vertebrate groups (Figs. 1 and 2). The pattern by which the cTnI and cMyBP-C isoforms from each vertebrate group diverge map well onto the accepted vertebrate phylogeny (see Fig. 5) (29). These results illustrate that the sequence of the tissue-specific paralogs are well conserved in species of the same vertebrate group and that the

Table 1. *List of TnI sequences used in this analysis*

Name	Organism	Common Name	Taxonomy	Identifier	Database
c_Cow	*Bos taurus*	cow	EM	P08057.2	GenBank
c_Dog	*Canis familiaris*	dog	EM	ENSCAFP00000003805	Ensembl
c_Gray_Wolf	*Canis lupus*	gray wolf	EM	Q8MKD5.3	GenBank
c_Goat	*Capra hircus*	goat	EM	AAK56404.1	GenBank
c_Guinea_Pig	*Cavia porcellus*	guinea pig	EM	ENSCPOP00000016138	Ensembl
c_Horse	*Equus caballus*	horse	EM	Q5PYI0.3	GenBank
c_Cat	*Felis catus*	cat	EM	NP_001009237.1	GenBank
c_Human	*Homo sapiens*	human	EM	P19429.3	GenBank
c_Macaque	*Macaca mulatta*	rhesus macaque	EM	XP_001085820.1	GenBank
c_Mouse	*Mus musculus*	mouse	EM	ENSMUSP00000096458	Ensembl
c_Rabbit	*Oryctolagus cuniculus*	European rabbit	EM	P02646.2	GenBank
c_Chimpanzee	*Pan troglodytes*	chimpanzee	EM	XP_001134934.1	GenBank
c_Orangutan	*Pongo pygmaeus*	orangutan	EM	ENSPPYP00000011661	Ensembl
c_Flying_Fox	*Pteropus vampyrus*	flying fox	EM	ENSPVAP00000006881	Ensembl
c_Rat	*Rattus norvegicus*	Norway rat	EM	CAA41402.1	GenBank
c_Squirrel	*Spermophilus tridecemlineatus*	ground squirrel	EM	ENSSTOP00000004951	Ensembl
c_Pig	*Sus scrofa*	pig	EM	NP_001092069.1	GenBank
c_Opossum	*Monodelphis domestica*	gray-short tailed opossum	MS	XP_001381826.1	GenBank
c_Platypus	*Ornithorhynchus anatinus*	platypus	MT	XP_001508807.1	GenBank
c_Quail	*Coturnix japonica*	Japanese quail	bird	P27672.2	GenBank
c_Chicken	*Gallus gallus*	chicken	bird	NP_998735.1	GenBank
c_Turkey	*Meleagris gallopavo*	turkey	bird	AAS45405.1	GenBank
c_Anole_Lizard	*Anolis carolinensis*	anole lizard	reptile	ENSACAP00000016893	Ensembl
c_Toad	*Bufo marinus*	cane toad	AM	AAX69047.1	GenBank
c_Bullfrog	*Rana catesbeiana*	American bullfrog	AM	AAO33937.1	GenBank
c_Clawed_Frog	*Xenopus laevis*	African clawed frog	AM	NP_001011410.1	GenBank
c_Tropical_Frog	*Xenopus tropicalis*	Western clawed frog	AM	NP_001081378.1	GenBank
c_Zebrafish	*Danio rerio*	zebrafish	fish	NP_001008613.1	GenBank
c_Trout	*Oncorhynchus mykiss*	rainbow trout	fish	HM012798	GenBank
c_Salmon	*Salmo salar*	Atlantic salmon	fish	AAB53386.1	GenBank
Sea_Squirt	*Halocynthia roretzi*	sea squirt	Uro	BAB83811.1	GenBank
Vase_Tunicate	*Ciona intestinalis*	vase tunicate	Uro	AAL27686.2	GenBank
fs_Cow	*Bos taurus*	cow	EM	XP_582883.2	GenBank
fs_Titi	*Callicebus moloch*	dusky titi	EM	ACB21271.1	GenBank
fs_Marmoset	*Callithrix jacchus*	marmoset	EM	ABZ80199.1	GenBank
fs_Dog	*Canis familiaris*	dog	EM	ENSCAFP00000014808	Ensembl
fs_Gray_Wolf	*Canis lupus*	gray wolf	EM	XP_851068.1	GenBank
fs_Goat	*Capra hircus*	goat	EM	AAK56403.1	GenBank
fs_Guinea_Pig	*Cavia porcellus*	guinea pig	EM	AAK56402.1	GenBank
fs_Horse	*Equus caballus*	horse	EM	XP_001493020.	GenBank
fs_Human	*Homo sapiens*	human	EM	NP_003273.1	GenBank
fs_Elephant	*Loxodonta africana*	African elephant	EM	ENSLAFP00000010414	Ensembl
fs_Macaque	*Macaca mulatta*	rhesus macaque	EM	XP_001117040.1	GenBank
fs_Mouse	*Mus musculus*	mouse	EM	NP_033431.1	GenBank
fs_Rabbit	*Oryctolagus cuniculus*	European rabbit	EM	P02643.3	GenBank
fs_Baboon	*Papio anubis*	olive baboon	EM	ABX52214.1	GenBank
fs_Orangutan	*Pongo pygmaeus*	orangutan	EM	ENSPPYP00000003385	Ensembl
fs_Rat	*Rattus norvegicus*	Norway rat	EM	P27768.2	GenBank
fs_Bat	*Rhinolophus ferrumequinum*	horseshoe bat	EM	ACC68943.1	GenBank
fs_Squirrel	*Spermophilus tridecemlineatus*	ground squirrel	EM	ENSSTOP00000007004	Ensembl
fs_Pig	*Sus scrofa*	pig	EM	ACD70341.1	GenBank
fs_Platypus	*Ornithorhynchus anatinus*	platypus	MT	XP_001518819.1	GenBank
fs_Quail	*Coturnix japonica*	Japanese quail	bird	P68247.2	GenBank
fs_Chicken	*Gallus gallus*	chicken	bird	NP_990748	GenBank
fs_Zebra_Finch	*Taeniopygia guttata*	zebra finch	bird	ENSTGUP00000009687	Ensembl
fs_Anole_Lizard	*Anolis carolinensis*	Anole lizard	reptile	ENSACAP00000009373	Ensembl
fs_Bullfrog	*Rana catesbeiana*	American bullfrog	AM	AAW73073.1	GenBank
fs_Tropical_Frog	*Xenopus tropicalis*	Western clawed frog	AM	NP_001011110.1	GenBank
fs_Clawed_Frog	*Xenopus laevis*	African clawed frog	AM	NP_001079556.1	GenBank
fs_Herring	*Clupea hargenus*	Atlantic herring	fish	AAB05825.1	GenBank
fs_Zebrafish	*Danio rerio*	zebrafish	fish	AAH92995.1	GenBank
fs_Cod	*Gadus morhua*	Atlantic cod	fish	AF498091	GenBank
fs_Stickleback	*Gasterosteus aculeatus*	three-spined stickleback	fish	ENSGACP00000022899	Ensembl
fs_Trout	*Oncorhynchus mykiss*	rainbow trout	fish	BAG31345.1	GenBank
fs_Medaka	*Oryzias latipes*	medaka	fish	ENSORLP00000000672	Ensembl
fs_Halibut	*Paralichthys olivaceus*	bastard halibut	fish	AAP82940.1	GenBank
fs_Minnow	*Pimephales promelas*	fathead minnow	fish	ABG77304.1	GenBank
fs_Salmon	*Salmo salar*	Atlantic salmon	fish	ACI67113.1	GenBank
fs_Bream	*Sparus aurata*	gilt-head bream	fish	AAT44429.1	GenBank

Continued

Table 1.—*Continued*

Name	Organism	Common Name	Taxonomy	Identifier	Database
fs_Fugu	*Takifugu Rubripes*	Japanese pufferfish	fish	ENSTRUP00000042560	Ensembl
fs_Pufferfish	*Tetraodon nigroviridis*	spotted green pufferfish	fish	ENSTNIP00000005376	Ensembl
ss_Dog	*Canis familiaris*	dog	EM	ENSCAFP00000015880	Ensembl
ss_Goat	*Capra hircus*	goat	EM	AAK56402.1	GenBank
ss_Guinea_Pig	*Cavia porcellus*	guinea pig	EM	ENSCPOP00000015611	Ensembl
ss_Human	*Homo sapiens*	human	EM	AAC14461.1	GenBank
ss_Elephant	*Loxodonta africana*	African elephant	EM	ENSLAFP00000008481	Ensembl
ss_Macaque	*Macaca mulatta*	Rhesus macaque	EM	XP_001108591.1	GenBank
ss_Mouse_Lemur	*Microcebus murinus*	mouse lemur	EM	ENSMICP00000004100	Ensembl
ss_Mouse	*Mus musculus*	mouse	EM	NP_001106173.1	GenBank
ss_Rabbit	*Oryctolagus cuniculus*	European rabbit	EM	P02645.1	GenBank
ss_Bushbaby	*Otolemur garnetti*	bushbaby	EM	ENSOGAP00000012926	Ensembl
ss_Chimpanzee	*Pan troglodytes*	chimpanzee	EM	XP_525019.2	GenBank
ss_Orangutan	*Pongo pygmaeus*	orangutan	EM	ENSPPYP00000011902	Ensembl
ss_Hyrax	*Procavia capensis*	rock hyrax	EM	ENSPCAP00000003368	Ensembl
ss_Flying_Fox	*Pteropus vampyrus*	flying fox	EM	ENSPVAP00000007908	Ensembl
ss_Rat	*Rattus norvegicus*	Norway rat	EM	NP_058880.1	GenBank
ss_Squirrel	*Spermophilus tridecemlineatus*	ground squirrel	EM	ENSSTOP00000011514	Ensembl
ss_Pig	*Sus scrofa*	pig	EM	AAP37479.1	GenBank
ss_Opossum	*Monodelphis domestica*	gray-short tailed opossum	MS	XP_001368433.1	GenBank
ss_Chicken	*Gallus gallus*	chicken	bird	XP_419242.2	GenBank
ss_Zebra_Finch	*Taeniopygia guttata*	zebra finch	bird	P19237	GenBank
ss_Anole_Lizard	*Anolis carolinensis*	Anole lizard	reptile	ENSACAP00000003597	Ensembl
ss_Tropical_Frog	*Xenopus tropicalis*	Western clawed frog	AM	NP_988959.1	GenBank
ss_Pike	*Esox lucius*	Northern pike	fish	ACO14305.1	GenBank
ss_Stickleback	*Gasterosteus aculeatus*	three-spined stickleback	fish	ENSGACP00000012841	Ensembl
ss_Medaka	*Oryzias latipes*	medaka	fish	ENSORLP00000008781	Ensembl
ss_Salmon	*Salmo salar*	Atlantic salmon	fish	ACM09409.1	GenBank
ss_Fugu	*Takifugu Rubripes*	Japanese pufferfish	fish	ENSTRUP00000030689	Ensembl
ss_Puffer	*Tetraodon nigroviridis*	spotted green pufferfish	fish	ENSTNIP00000018086	Ensembl

TnI, troponin I; c, cardiac, fs; fast skeletal; ss, slow skeletal; EM, eutherian mammal; MS, marsupial; MT, monotreme; Uro, urochordate; AM, amphibian.

changes that have occurred in parallel with the evolution and divergence of the different vertebrate groups.

Sequence alignment analysis. Alignment of all sequences for TnI and MyBP-C reveal that there is a high degree of conservation across all sequences for each protein (Figs. 3 and 4). For example, two important sequence motifs in TnI are the inhibitory peptide (human residues 137–146) and switch peptide (147–162). When the muscle is relaxed the inhibitory peptide is bound to actin. This interaction prevents the movement of tropomyosin across the thin filament and the exposure of myosin binding sites. Following the Ca^{2+} activation of TnC, TnI changes position and the switch peptide binds to the exposed hydrophobic patch on the NH_2 terminus of TnC. This change in the position of TnI, away from actin, releases tropomyosin and enables the exposure of myosin binding sites (31). Comparison of the amino acid sequences of the inhibitory peptide and switch peptide across all known TnI isoforms demonstrates extremely high conservation (Fig. 3). For example, the inhibitory peptide consists of ten residues and only one of these significantly varies between all TnI paralogs. A similar degree of conservation is found between the switch peptides across all TnI paralogs. The switch peptide is 16 residues in length of which 12 are identical across all known TnI isoforms. The remaining four have extremely low variation between the sequences (Fig. 3). This conservation is remarkable considering that there is ~500 million years (My) of evolution separating humans and tunicates (70) and reflects the functional importance of these peptides to muscle function.

Similar levels of conservation are found when comparing MyBP-C sequences. The unique ~100 amino acid MyBP-C motif, or M-domain, is found in the NH_2 terminus of the protein between domains C1 and C2 (Fig. 4) and is essential for the regulation of actomyosin interactions. The M-domain binds myosin S2 (the hinge region of myosin connecting the S1 heads to the myosin rod) (24) as well as actin (52) to potentially limit cross-bridge formation. When phosphorylated by PKA, the M-domain no longer binds S2 (24) or actin (52), allowing cross bridges to form more readily and resulting in increases in cross-bridge cycling rates (32, 57). Even though specific amino acids in the M-domain have not been implicated in functional effects, there are multiple regions in the M-domain that are highly conserved (including residues 293–300 and 331–353, human numbering) that may be of functional importance. There are also regions of MyBP-C that are not well conserved and are completely unique to the cardiac paralog. All cMyBP-C isoforms contain an additional ~100 amino acid domain at the extreme NH_2 terminus (the C0 domain) which is not found in fsMyBP-C nor ssMyBP-C (Supplementary Fig. S1). In addition, all cMyBP-C isoforms contain an insert in the M-domain that contains a PKA phosphorylation site (residues 276 to 284, human numbering) and a 28 amino acid insert in the C5-domain (residues 687 to 714, human numbering). The insertion in the M-domain, while present in all sequences studied, is more conserved within mammalian species than lower vertebrates. A sequence rich in Pro and Ala residues (the Pro-Ala rich region) is found between the C0 and C1 domains in cMyBP-C and at the extreme NH_2 terminus of fsMyBP-C and ssMyBP-C (Supplementary Fig. S1) that has previously been shown to scale with heart rate in mammals (51, 53).

Table 2. *List of MyBP-C sequences used in this analysis*

Name	Organism	Common Name	Taxonomy	Identifier	Database
c_Cow	*Bos taurus*	cow	EM	NP_001070004.1	GenBank
c_Dog	*Canis familiaris*	dog	EM	NP_001041571.1	GenBank
c_Guinea_Pig	*Cavia porcellus*	guinea pig	EM	ENSCPOP00000012034	Ensembl
c_Horse	*Equus caballus*	horse	EM	XP_001491151.2	GenBank
c_Cat	*Felis catus*	cat	EM	ENSFCAP00000002329	Ensembl
c_Human	*Homo sapiens*	human	EM	NP_000247.2	GenBank
c_Macaque	*Macaca mulatta*	rhesus macaque	EM	ENSMMUP00000020472	Ensembl
c_Chimpanzee	*Pan troglodytes*	chimpanzee	EM	ENSPTRP00000056324	Ensembl
c_Orangutan	*Pongo pygmaeus*	orangutan	EM	ENSPPYP00000003794	Ensembl
c_Flying_Fox	*Pteropus vampyrus*	flying fox	EM	ENSPVAP00000006921	Ensembl
c_Rat	*Rattus norvegicus*	Norway rat	EM	NP_001099960.1	GenBank
c_Mouse	*Mus musculus*	mouse	EM	AF097333.1	GenBank
c_Opossum	*Monodelphis domestica*	gray-short tailed opossum	MS	ENSMODP00000024836	Ensembl
c_Platypus	*Ornithorhynchus anatinus*	platypus	MT	ENSOANP00000012560	Ensembl
c_Chicken	*Gallus gallus*	chicken	bird	NP_990447.1	GenBank
c_Zebra_Finch	*Taeniopygia guttata*	zebra finch	bird	ENSTGUP00000010777	Ensembl
c_Anole_Lizard	*Anolis carolinensis*	anole lizard	reptile	ENSACAP00000002655	Ensembl
c_Axolotl	*Ambystoma mexicanum*	axolotl	AM	MYPC3_AMBME	UnitProt
c_Clawed_Frog	*Xenopus laevis*	African clawed frog	AM	Q90X86_XENLA	UnitProt
c_Tropical_Frog	*Xenopus tropicalis*	Western clawed frog	AM	NP_001106379.1	GenBank
c_Zebrafish	*Danio rerio*	zebrafish	fish	NP_001037814.2	GenBank
c_Stickleback	*Gasterosteus aculeatus*	three-spined stickleback	fish	ENSGACP00000020475	Ensembl
c_Medaka	*Oryzias latipes*	medaka	fish	ENSORLP00000007637	Ensembl
c_Fugu	*Takifugu Rubripes*	Japanese pufferfish	fish	ENSTRUP00000044247	Ensembl
c_Puffer	*Tetraodon nigroviridis*	spotted green pufferfish	fish	ENSTNIP00000012278	Ensembl
fs_Dog	*Canis familiaris*	dog	EM	XP_533608.2	GenBank
fs_Guinea_Pig	*Cavia porcellus*	guinea pig	EM	ENSCPOP00000015829	Ensembl
fs_Horse	*Equus caballus*	horse	EM	ENSECAP00000018317	Ensembl
fs_Human	*Homo sapiens*	human	EM	NP_004524.3	GenBank
fs_Macaque	*Macaca mulatta*	rhesus macaque	EM	ENSMMUP00000008835	Ensembl
fs_Mouse	*Mus musculus*	mouse	EM	NP_666301.1	GenBank
fs_Orangutan	*Pongo pygmaeus*	orangutan	EM	ENSPPYP00000011513	Ensembl
fs_Rat	*Rattus norvegicus*	Norway rat	EM	NP_001099727.1	GenBank
fs_Dolphin	*Tursiops truncatus*	dolphin	EM	ENSTTRP00000002269	Ensembl
fs_Platypus	*Ornithorhynchus anatinus*	platypus	MT	ENSOANP00000006210	Ensembl
fs_Chicken	*Gallus gallus*	chicken	bird	NP_001038124.1	GenBank
fs_Anole_Lizard	*Anolis carolinensis*	anole lizard	reptile	ENSACAP00000012249	Ensembl
fs_Tropical_Frog	*Xenopus tropicalis*	Western clawed frog	AM	ENSXETP00000007358	Ensembl
fs_Stickleback	*Gasterosteus aculeatus*	three-spined stickleback	fish	ENSGACP00000010807	Ensembl
fs_Medaka	*Oryzias latipes*	medaka	fish	ENSORLP00000008812	Ensembl
fs_Fugu	*Takifugu Rubripes*	Japanese pufferfish	fish	ENSTRUP00000035083	Ensembl
fs_Puffer	*Tetraodon nigroviridis*	spotted green pufferfish	fish	ENSTNIP00000005801	Ensembl
ss_Cow	*Bos taurus*	cow	EM	NP_001104243.1	GenBank
ss_Dog	*Canis familiaris*	dog	EM	ENSCAFP00000010679	Ensembl
ss_Guinea_Pig	*Cavia porcellus*	guinea pig	EM	ENSCPOP00000016631	Ensembl
ss_Horse	*Equus caballus*	horse	EM	XP_001497234.2	GenBank
ss_Human	*Homo sapiens*	human	EM	NP_002456.2	GenBank
ss_Macaque	*Macaca mulatta*	rhesus macaque	EM	ENSMMUP00000027471	Ensembl
ss_Mouse	*Mus musculus*	mouse	EM	NP_780627.2	GenBank
ss_Chimpanzee	*Pan troglodytes*	chimpanzee	EM	ENSPTRP00000041533	Ensembl
ss_Orangutan	*Pongo pygmaeus*	orangutan	EM	ENSPPYP00000005564	Ensembl
ss_Rat	*Rattus norvegicus*	Norway rat	EM	ENSRNOP00000035289	Ensembl
ss_Opossum	*Monodelphis domestica*	gray-short tailed opossum	MS	ENSMODP00000003684	Ensembl
ss_Platypus	*Ornithorhynchus anatinus*	platypus	MT	ENSOANP00000008864	Ensembl
ss_Chicken	Gallus gallus	chicken	bird	XP_416332.2	GenBank
ss_Zebra_Finch	Taeniopygia guttata	zebra finch	bird	ENSTGUP00000011665	Ensembl
ss_Anole_Lizard	*Anolis carolinensis*	anole lizard	reptile	ENSACAP00000016527	Ensembl
ss_Zebrafish	*Danio rerio*	zebrafish	fish	NP_001007323.1	GenBank
ss_Stickleback	*Gasterosteus aculeatus*	three-spined stickleback	fish	ENSGACP00000026275	Ensembl
ss_Medaka	*Oryzias latipes*	medaka	fish	ENSORLP00000019750	Ensembl
ss_Fugu	*Takifugu Rubripes*	Japanese pufferfish	fish	ENSTRUP00000028945	Ensembl
ss_Puffer	*Tetraodon nigroviridis*	spotted green pufferfish	fish	ENSTNIP00000001343	Ensembl

MyBP-C, myosin binding protein-C.

Several amino acids were identified in MyBP-C that are conserved depending on species. There are eight residues that switch from nonpolar and neutral charge in mammalian cMyBP-C to charged amino acids in all other isoforms (Fig. 4 and Supplementary Fig. S1). For example, in human cMyBP-C, and all other mammalian cardiac isoforms, there is a Gly at residue number 354 (or its equivalent), while there is an Arg at the equivalent site in nonmammalian cardiac iso-

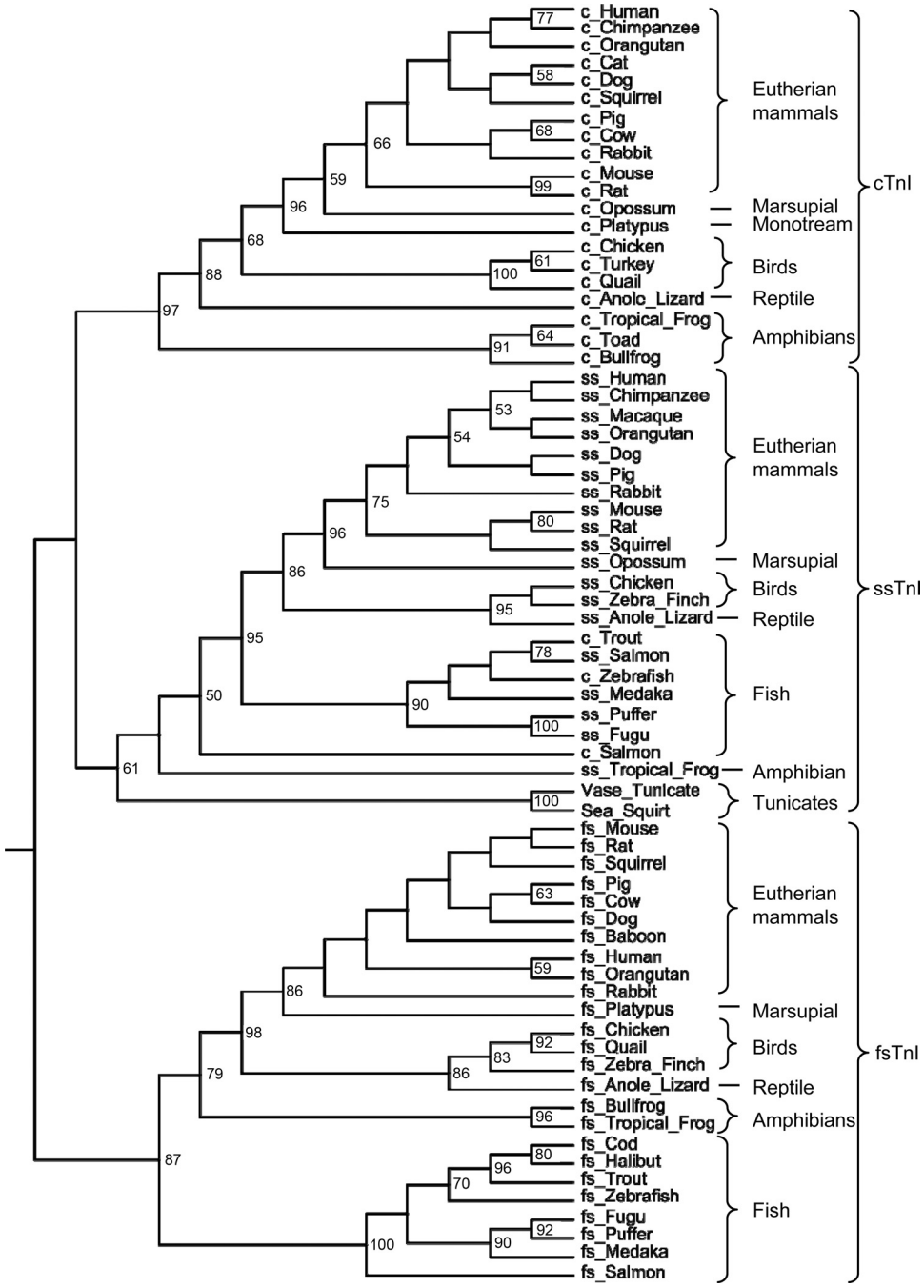

Fig. 1. A maximum likelihood (ML) consensus tree of the troponin I (TnI) family We included 99 sequences from 51 different species in this analysis; however, this tree has been trimmed to 69 sequences from 29 species to ease comprehension. The ML tree was constructed by resampling 100 datasets (bootstrap) with randomized input order using Seqboot in the PHYLIP software. The tree is midpoint rooted. Clear splits are observed between the clades for each of the tissue types [fast skeletal (fs), slow skeletal (ss), and cardiac (c)]. The splits between clades are the site of TnI duplication in vertebrates. Numbers indicate bootstrap values and represent the number of times out of 100 replicates that the same result was achieved for the respective partition.

forms as well as all fast skeletal and slow skeletal isoforms. This substitution pattern is similar to the evolutionary loss of the His button in TnI (described below). The functional importance of these amino acids is unknown.

Evolution of regulatory control via phosphorylation of TnI and MyBP-C. Phosphorylation of human cTnI and cMyBP-C by PKA and PKC represent important pathways for regulating cardiac output. The target residues enable the functional properties of the regulatory proteins to be directly modified. The results of the current study demonstrate that the number of phosphorylation targets in both cTnI and cMyBP-C increased during vertebrate evolution (Fig. 5). As a result, the regulation of cardiac function became increasing complex. In human cTnI, Ser^{23} and Ser^{24} are located in the cardiac specific NH_2-terminal peptide that is present in most cTnI isoforms. They are not present in any fsTnI or ssTnI isoforms. These two residues are targets for PKA and PKC (42). As shown in Fig. 3, this peptide is found in cTnI from all endothermic species, including eutherian mammals, marsupials, birds, as well as in the ectothermic amphibians and reptiles. It is not present, however, in cTnI from any fish species. There is an NH_2-terminal extension in TnI from the tunicates but the sequence of this displays no similarity to that found in other vertebrate TnI isoforms. It is not known if the NH_2-terminal peptide is present within monotremes, as only the COOH terminus of the cTnI gene in the platypus has been sequenced (residues 154–210). As the NH_2-terminal peptide is not present in either

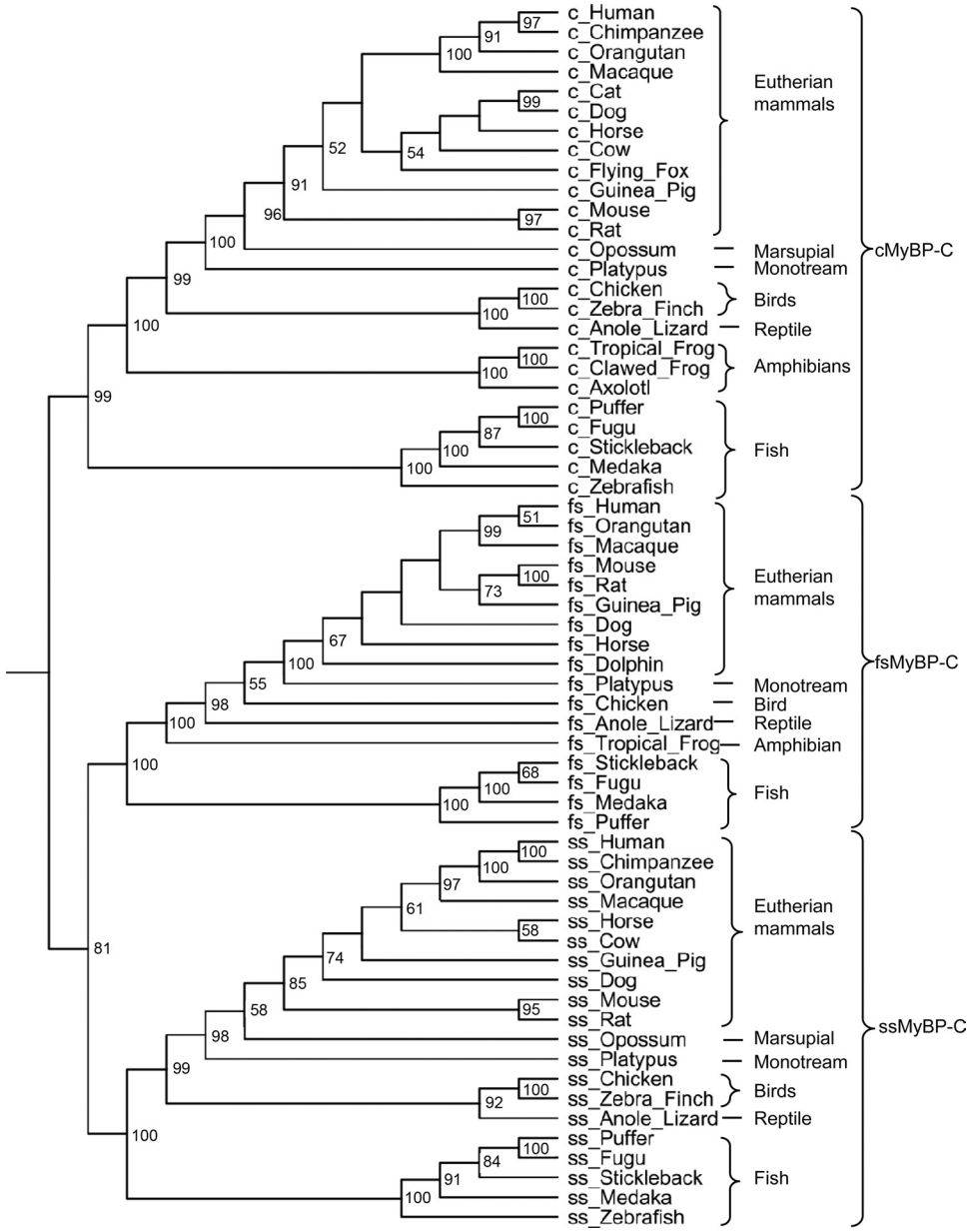

Fig. 2. An ML consensus tree of the myosin binding protein-C (MyBP-C) family. We included 62 sequences from 26 different species in this analysis. The ML tree was constructed by resampling 100 datasets (bootstrap) with randomized input order using Seqboot in the PHYLIP software. The tree is midpoint rooted. Clear splits are observed between the clades for each of the tissue types (fs, ss, and c). The splits between clades are the site of MyBP-C duplication in vertebrates. Numbers indicate bootstrap values and represent the number of times out of 100 replicates that the same result was achieved for the respective partition.

fsTnI or ssTnI, this suggests that it appeared in cTnI after the evolution of distinct muscle types.

In addition to Ser^{23} and Ser^{24}, there are three additional targets for PKC on human cTnI: Ser^{43}, Ser^{45}, and Thr^{144}. The current analysis demonstrates that all known cTnI isoforms contain a Ser residue at positions equivalent to 43 and 45 in human cTnI. Phosphorylation of these residues in human and rat cTnI decreases maximum tension and Ca^{2+} sensitivity of force generation, lowers affinity of myosin S1 for actin, and reduces thin filament sliding speed (7, 41). While there are potentially phosphorylatable residues (Ser or Thr) at positions equivalent to 43 and 45 in all ssTnI isoforms (Fig. 3) there is no evidence that they are phosphorylated in vivo (65). fsTnI from all fish and amphibian species contain putative PKC targets at positions equivalent to Ser^{43}, Ser^{45} in human cTnI (Fig. 3). These residues are however absent from the fsTnI of all mammal, bird, and reptile species (Fig. 3). This suggests that these putative PKC targets were present in the ancestral form of TnI that through duplication generated the cTnI isoform. There have been no studies to date, however, to determine if they are phosphorylated in vivo. While these residues have been retained in cTnI and ssTnI in all species they were lost in fsTnI in a common ancestor to reptiles, birds, and mammals. As mentioned earlier, residue 144 in human cTnI is also a target for PKC. The current analysis demonstrates that there is a phosphorylatable residue present at this position (or at its equivalent) in cTnI from all known eutherian mammals as well as from two fish species (zebra fish, trout) and one amphibian species (cane toad) (Fig. 3). However, no marsupial, bird, or reptile cTnI contains a phosphorylatable residue at this position nor is there a phosphorylatable residue at the equivalent position in ssTnI or fsTnI isoforms.

Four PKA phosphorylation sites (Ser^{275}, Ser^{284}, Ser^{304}, and Ser^{311}) have been identified in human cMyBP-C (30). Of these,

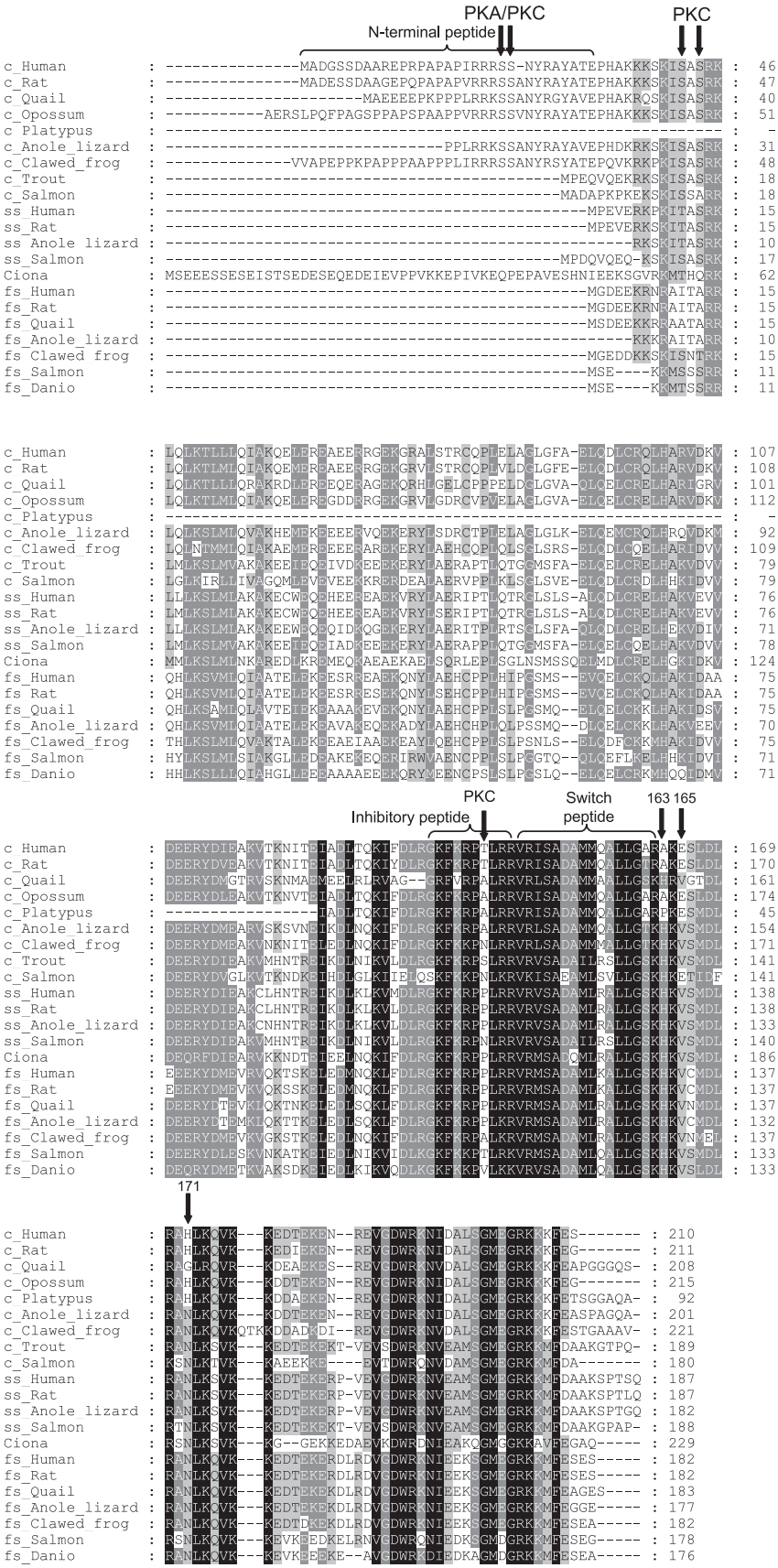

Fig. 3. Sequence alignment of TnI sequences from c, ss, and fs muscle from select vertebrate species. Human cTnI protein kinase A (PKA) and PKC targets are indicated, as are the position equivalents of residues 163, 165, and 171 in human cTnI. Location of inhibitory and switch peptides are also indicated on the sequence.

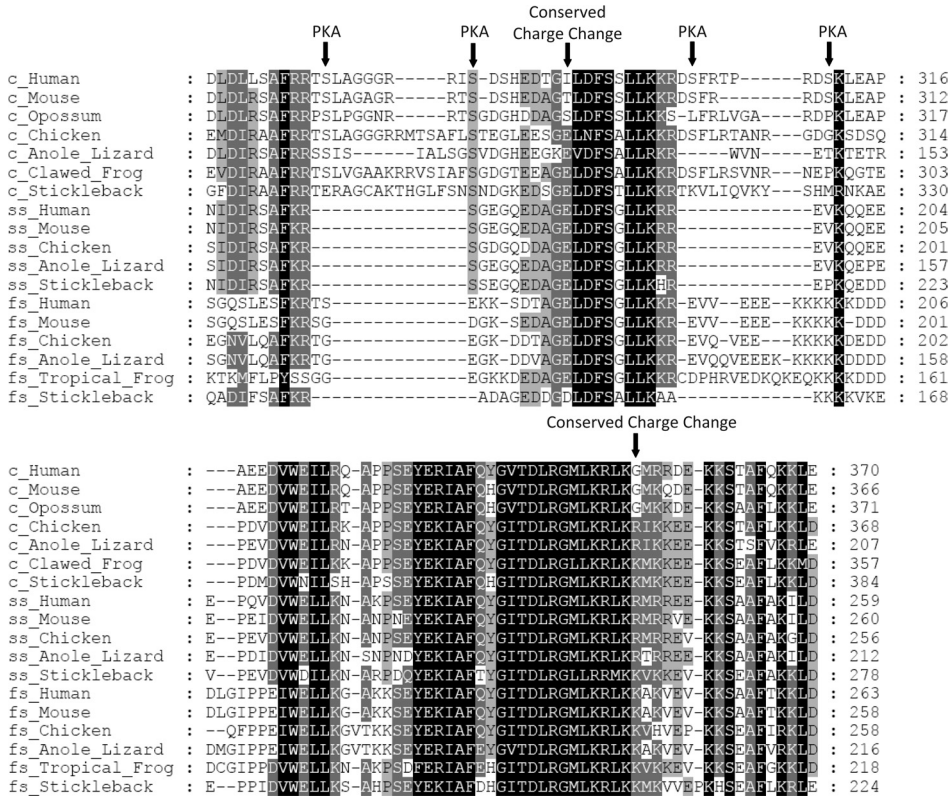

Fig. 4. Sequence alignment of the MyBP-C motif (M-domain) of c, ss, and fs MyBP-C isoforms from select vertebrate species. Human cMyBP-C PKA targets are indicated, as are residues where there is conserved charge change.

three are unique to cMyBP-C, while fsMyBP-C and ssMyBP-C each possess Ser^{275} (at its equivalent site, Fig. 4). However, only the cardiac isoforms of MyBP-C are phosphorylated by PKA in vitro (24). The distribution of phosphorylation sites in cardiac MyBP-C varies by species and the total number of sites increases with increasing vertebrate complexity (Figs. 4 and 5). Species from all of the major vertebrate classes (fish through mammals) possess Ser^{275} (except tropical frog, axolotl, stickleback, and medaka) and Ser^{304} (except fugu, tropical frog, and anole lizard), while Ser^{284} is not found in fish or reptiles, and Ser^{311} is found only in eutherian mammals (except pig and cow). In summary, cMyBP-C isoforms of vertebrate groups contain the following numbers of PKA phosphorylation sites: fish, 2; amphibians, 3; reptiles, 1; birds, 3; and mammals, 4. cMyBP-C sequence data for only one reptile (the anole lizard) were available, so conclusions regarding the number of PKA phosphorylation sites in reptilian cMyBP-C are not complete.

Residue changes within the COOH terminus of mammalian TnI. Examination of all TnI sequences reveals that every fsTnI and ssTnI isoform as well as TnI from the tunicate *Ciona intestinalis* contains a His, Val, and Asn at the positions equivalent to 132, 134, and 141 of human ssTnI. This pattern can be seen in Fig. 3. These positions correspond to 163, 165, and 172 in human cTnI where there is an Ala, Glu, and His. These three residues are at the same relative positions in cTnI from all known eutherian mammals and the opossum, a marsupial (Fig. 3). Work by Day et al. (13) suggests that the presence of His at position 132 in mammalian ssTnI helps to protect contractile function during cellular acidosis and have called this residue the "histidine button." Other studies have also suggested that the presence of Ala^{163} in human cTnI is responsible, at least in part, for the loss of Ca^{2+} sensitivity that occurs when the pH of the heart is decreased (63, 64, 66, 68). In addition, transgenic mouse hearts expressing A164H cTnI (mouse cTnI equivalent) have improved function when challenged by hypoxia, acidosis, ischemia, and chronic heart failure compared with wild-type hearts (13). Subsequent work by Westfall and Metzger (66) has demonstrated that Val^{134} and Asn^{141} in human ssTnI also reduce the impact of acidosis on contractile function.

Interestingly, as in human ssTnI, cTnI from all nonmammalian vertebrate species (amphibians, fishes, birds, etc.) contain a His at the equivalent position to 163 in human cTnI while the platypus (monotreme) has a Pro at this position (Fig. 3). Also as in human ssTnI there is a Val and an Asn at the positions equivalent to 165 and 172 in cTnI from most animals (Fig. 3). As mentioned above, the exceptions are the eutherian mammals and opossum as well as the platypus (Fig. 3). In addition, all bird cTnI isoforms contain a Gly at the position equivalent to Asn^{172}, while in cTnI from two amphibian species (*Bufo marinus* and *Rana catesbeina*) there is an Ala at the position equivalent to Val^{172}. There is also a Glu at the position equivalent to Val^{165} in Atlantic salmon. Together, these results suggest that the presence of His^{131}, Val^{134}, and Asn^{141} in ssTnI (and at the equivalent positions in fsTnI) is the ancestral state and that Ala, Glu, and His at the equivalent positions in mammalian cTnI is the derived state.

DISCUSSION

In this study, we have investigated the roles of TnI and MyBP-C in the evolution of regulatory function in vertebrate striated muscle. The major results are that TnI and MyBP-C isoforms form distinct clades according to muscle types (car-

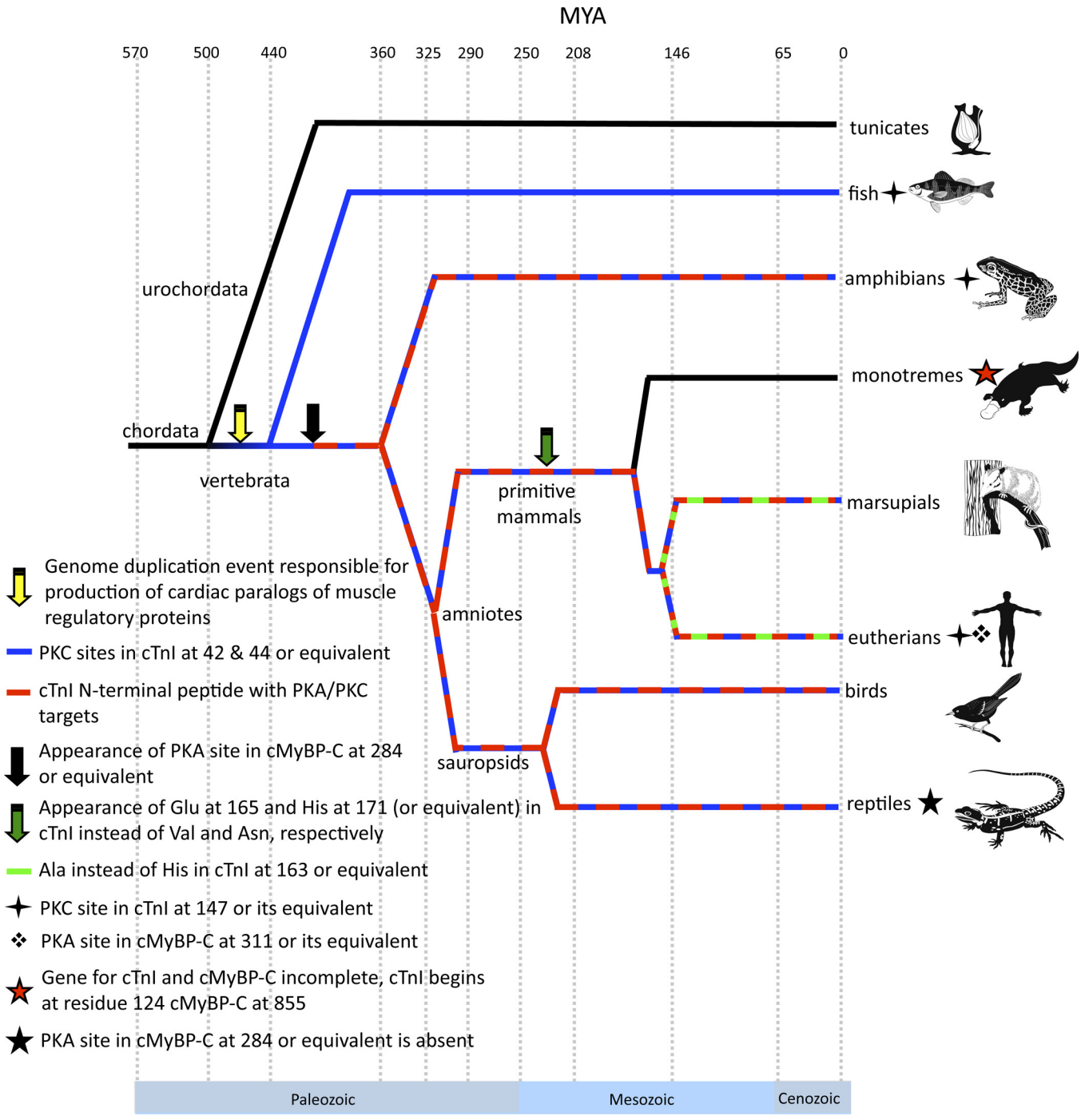

Fig. 5. A vertebrate phylogeny onto which critical changes in the sequence of cTnI and cMyBP-C have been mapped.

diac, fast skeletal, and slow skeletal) and that the evolution of tissue-specific isoforms mimics the evolution of vertebrates. Regulatory features (e.g., phosphorylation sites) of TnI and MyBP-C evolved as vertebrates became more complex, with the most number of regulatory features found in eutherian mammalian species. We now further understand how the functions of TnI and MyBP-C have changed with respect to vertebrate evolution and can begin to investigate the functional contributions of conserved amino acid residues in each protein.

Evolution of TnI and MyBP-C isoforms. The current study suggests that a fast skeletal-like ortholog (fsTnI) is the ancestral form of TnI and that the gene for fsTnI would then have been duplicated, producing an isoform that evolved into the cardiac paralog. Subsequently, a second genome duplication occurred and one of the replicated cTnI genes evolved into the slow skeletal paralog. This finding supports previous work by Oota and Saitou (43), who suggested, using thick and thin filament proteins, that the three muscle types found in vertebrates are the result of genome duplication events. Work by Chong and Jin (10), however, suggests that cTnI is the product of the duplication of the ssTnI gene that arose from the duplication of the fsTnI gene. These authors also suggest that

there are evolutionary relationships between TnI and TnT with the two proteins arising from a TnI-like ancestor protein. This previous study used monoclonal antibodies to characterize similarities in protein conformation between different isoforms and fragments of recombinantly produced cTnI and cTnT. The difference in results between the Chong and Jin (10) study and the current results in regards to the order of TnI isoform evolution is likely due to different methodological approaches (antibody recognition of tertiary structure vs. sequence analysis).

Recent studies using multiple gene families have established that a number of genome duplication events occurred after the divergence of the urochordates [500 million years ago (Mya)] but before the divergence of the fish lineage (440 Mya) from the vertebrate lineage (28, 34). Examination of the distribution of the TnI tissue paralogs present in the tunicates and the fish used in the present study suggests that the production of the three tissue-specific paralogs of TnI coincides with this time frame. Work by Cleto et al. (12) has identified a single gene for TnI in *C. intestinalis* that is expressed in the heart, the body wall, and the embryonic larval tail. The muscle in the body wall is homologous to vertebrate smooth muscle, while that in the heart and juvenile tail is homologous to vertebrate striated muscle (12, 40). This indicates that any gene duplication events that produced multiple isoforms of TnI in the vertebrates occurred after the urochordate lineage diverged from the vertebrate lineage. Multiple TnI isoforms have been found in a second tunicate (*Halocynthia roretzi*), but phylogenetic analysis indicates that these evolved after the tunicate lineage derived from the vertebrate lineage (72). Genome sequencing of the Atlantic salmon, *Salmo salar*, has identified separate genes for fsTnI, cTnI, and ssTnI. This indicates that the two genome duplication events responsible for the three tissue-specific paralogs occurred in a common ancestor of the bony fishes and all subsequent lineages. A number of species within the fishes have multiple isoforms of each of the tissue-specific paralogs. A similar finding has been reported for a variety of other cellular proteins and is the result of genome duplication events specific to the fish lineage that occurred after it diverged from the vertebrate lineage (1, 21, 22, 35, 48).

According to the results of this study, cardiac MyBP-C is the ancestral form of MyBP-C, while slow skeletal and fast skeletal isoforms diverged afterwards (Fig. 2). This pattern is different than that observed for TnI (Fig. 1) as well as other thick filament proteins (43). One possible explanation for the difference between the MyBP-C phylogenetic tree topology and that of TnI and other muscle proteins is that the sequence of the cMyBP-C orthologs is very unlike that of fsMyBP-C or ssMyBP-C and that the fsMyBP-C and ssMyBP-C orthologs are highly similar to each other. As a result, fsMyBP-C and ssMyBP-C are monophyletic. In addition, the most parsimonious evolutionary path has cMyBP-C as the ancestor protein and the fast skeletal and slow skeletal clades derived from this ancestor. However, the changes in the sequence of cMyBP-C could have easily occurred after a gene duplication event that produced the slow skeletal paralog from the cardiac paralog. The sequence differences in cMyBP-C are reflective of it becoming highly specialized, understood to have arisen as cardiac muscle function specialized from fast skeletal and slow skeletal muscle function.

Evolution of functional roles of TnI and MyBP-C. The influence of adrenergic stimulation on cardiac contractility is an important, and well characterized, regulatory mechanism of cardiac output (7, 9, 23, 31, 37, 55, 57, 71). At the level of the sarcomere, force generation and cross-bridge cycling is modulated via the phosphorylation of TnI and MyBP-C following α- or β-adrenergic stimulation. Multiple phosphorylation sites exist in each protein: in human cTnI there are five known target residues for phosphorylation and in cMyBP-C there are four. Each of these is specifically targeted by PKA and/or PKC. The phosphorylation of the NH_2-terminal peptide of human cTnI at Ser^{23} and Ser^{24} is thought to weaken the interaction between its NH_2-terminal domain and the NH_2-terminal domain of cTnC (9), thus decreasing the Ca^{2+} affinity of cTnC site II and resulting in a loss of Ca^{2+} sensitivity of myofilament activation (46) and an increase in the rate of relaxation (71). PKA phosphorylation of cMyBP-C increases cross-bridge cycling rates (31, 57, 61, 62) possibly through abolishment of interactions between the NH_2-terminal M-domain and actin (52) and/or myosin S2 (24). While the function of fast and slow skeletal muscle in mammals is modified following adrenergic stimulation the effect is thought to be via changes in the kinetics of Ca^{2+} cycling not in the contractile proteins (8).

The results of this study suggest that the TnI NH_2-terminal peptide, containing two phosphorylatable targets, appeared in the vertebrate lineage after the fishes diverged but before the amphibians diverged (360 Mya). This represents a gain of regulatory control and correlates with the colonization of land by vertebrates. As the energetic cost of locomotion is higher in terrestrial animals compared with aquatic animals (49), having the ability to match heart rate to aerobic requirement would increase the energy efficiency of the animal.

cTnI isoforms from nonmammalian species contain potential phosphorylation sites at Ser^{43}, Ser^{45}, and Thr^{144} (or equivalents). The computer program NetPhos (5) gives a high score (>97%) for each of these putative phosphorylatable residues: however, no studies have examined if these residues are phosphorylated under physiological conditions. There are two explanations for the presence/absence of a phosphorylatable residue at position 144 in cTnI from different species: it appeared independently in the different groups (eutherians, fish, and amphibians), or a phosphorylatable residue was present at this location in cTnI in a common vertebrate ancestor that was then lost in all groups/species other than those mentioned. The second explanation is the most parsimonious given that a phosphorylatable residue is not found at a position equivalent to 144 in any fsTnI, suggesting that this addition occurred after the evolution of the cTnI paralog.

The complexity of PKA phosphorylation of cMyBP-C has increased through evolutionary time, with endothermic vertebrates (mammals, birds) containing more PKA phosphorylation sites in the M-domain of cMyBP-C isoforms than ectothermic vertebrates (reptiles, amphibians, fish). Ser^{275} (human numbering) and Ser^{304} are present in all vertebrates, whereas Ser^{284} is only found in mammals, birds, and amphibians, and Ser^{311} is only present in eutherian mammals (Fig. 4). The fact that Ser^{284} (or its equivalent) is not found in cMyBP-C isoforms of reptiles and fish is interesting because this site has been hypothesized to be a "gateway" site towards a hierarchical phosphorylation of cMyBP-C. Gautel et al. (17) demonstrated that when Ser^{284} is mutated to an Ala (to abolish

phosphorylation) in human recombinant cMyBP-C, PKA phosphorylation levels of Ser^{275} and Ser^{304} were significantly decreased. Ser^{275} and Ser^{304} were only phosphorylated to maximal levels when Ser^{284} was able to be phosphorylated. These results led the authors to conclude that in vivo, Ser^{284} is phosphorylated by PKA first, resulting in a conformational change of the M-domain to allow further phosphorylation of the other sites. If this was the case in all vertebrates, then fish and reptile cMyBP-C would not be able to be phosphorylated since they lack an equivalent phosphorylation site to Ser^{284}. However, the three-dimensional conformation of the M-domain may differ between species, thus exposing other PKA sites and allowing phosphorylation to occur even if Ser^{284} (or its equivalent) is not present in a given isoform. The crystal structure for the M-domain has yet to be solved for any MyBP-C isoform, so comparisons between M-domain structures cannot be made at this time. Further work is also needed to determine if hierarchical phosphorylation of cMyBP-C is a common theme amongst all vertebrates.

Regulation of TnI function through changes in the COOH terminus of TnI. The replacement of His^{163}, Val^{165}, and Asn^{172} with an Ala, Glu and His, respectively, in cTnI from all eutherian mammals and marsupials represents a change that has been maintained for over 148 million years (My). This suggests that these mutations in cTnI benefit cardiac function. Westfall and Metzger (66) have shown that the insertion of any of these residues (Ala, Glu and His) at their respective positions into ssTnI reduces the Ca^{2+} sensitivity of myocyte contraction at physiological pH. We therefore hypothesize that it is this reduction in Ca^{2+} sensitivity caused by these replacements in mammalian and marsupial cTnI that is the benefit. This is relevant because previous studies have shown that the Ca^{2+} sensitivity of cardiac tissue isolated from trout, frog, rat and guinea pig increases to a similar degree as temperature is increased (11, 27). In addition, when compared at 7°C, the Ca^{2+} sensitivity of the trout heart and frog heart is ten-fold and eight-fold higher than that of the rat heart, respectively (11). 7°C is a physiological temperature of the trout, and the Ca^{2+} sensitivity of its heart at this temperature is similar to that of a rat heart at 37°C (20). This means that if the temperature of the trout heart were increased to 37°C, its Ca^{2+} sensitivity would increase and it would likely become locked in contracture. This would be due to Ca^{2+} remaining bound to TnC. To avoid this condition, it was therefore necessary for the Ca^{2+} sensitivity of the vertebrate heart to decrease as endothermy evolved and the physiological temperature of vertebrates increased. This is the reason why the Ca^{2+} sensitivity of the mammalian heart is ten-fold less than that of the trout when compared at the same temperature. One mechanism responsible for lower Ca^{2+} sensitivity is that the Ca^{2+} affinity of mammalian cTnC is two-fold lower that of trout cTnC (19).

We now propose that a second mechanism by which the Ca^{2+} sensitivity of the mammalian heart was reduced was via the mutation of mammalian cTnI to contain Ala^{163}, Glu^{165}, and His^{172}. This change in sequence correlates with the appearance of endothermy in mammalian evolution. As mentioned above, cTnI isoforms from a variety of bird species contain a His and a Gly at positions equivalent to 163 and 165. This suggests that the replacement of His^{163} with an Ala in eutherian cTnI occurred after the lineage that led to the birds diverged from the synapsid lineage that leads to eutherian mammals. However, the Ca^{2+} sensitivity of force generation by cardiac tissue from the turkey and rat are similar when measured at the same temperature (25, 44). This suggests that other mechanisms evolved to decrease the Ca^{2+} sensitivity of the bird heart to enable function at high temperatures. In addition, as the Ca^{2+} sensitivity of fast skeletal muscle is not as sensitive to temperature as cardiac muscle is (26), it was not necessary for its Ca^{2+} sensitivity to be altered as endothermy evolved. Finally, the protective effect of inserting a His at residue 163 of cTnI may only be functional in mammalian hearts as Churcott et al. (11) have shown that Ca^{2+} sensitivity of actin-activated myosin ATPase isolated from the trout heart is significantly affected by a 0.2 pH unit increase at 7.0. However, as these authors measured unloaded ATPase activity in solution, this result may not represent what would occur under more physiological conditions when contraction occurs against a load.

As a change in temperature alters the function of cellular proteins (56), it is also likely that there have been changes during the evolution of endothermy to the functional characteristics of the membrane channels and transporters that regulate intercellular Ca^{2+}. This includes the sarco/endoplasmic reticulum Ca^{2+} ATPase, the ryanodine receptor, and L-type Ca^{2+} channels. Work by Elias et al. (14) has shown that the activity of Na^{2+}/Ca^{2+} exchanger (NCX) cloned from the trout heart is temperature insensitive and, as a result, able to function at temperatures lower than canine NCX. Further work is needed to characterize how the other Ca^{2+} handling proteins have changed during the evolution of the vertebrate heart.

While the presence of Ala^{163}, Glu^{165}, and His^{173} in human cTnI makes the heart more susceptible to cellular acidosis, this condition only occurs in the adult mammalian heart under extreme conditions like ischemia. During embryonic development when cardiac acidosis is common, ssTnI is expressed in the heart (67). Cellular acidosis is however common in fast skeletal muscle following exhaustive anaerobic exercise. There is a distinct benefit to having the protective influence of a His at the equivalent position in fsTnI on myocyte contractility.

Conclusions and Perspectives

The results of this analysis demonstrate that tissue-specific paralogs of TnI and MyBP-C present in fast skeletal, slow skeletal, and cardiac muscle are the result of gene duplication events that occurred after the divergence of tunicates from the vertebrate lineage. Following these duplication events, each tissue-specific paralog began to change in sequence and conferred specialized function and capability to each muscle type. This includes the increased number of phosphorylation sites in both cTnI and cMyBP-C, the loss of PKC targets in fsTnI in an ancestor of all reptiles, birds, and mammals, as well as the replacement of His^{163}, Val^{165}, and Asn^{172} with Ala, Glu, and His, respectively, in mammalian cTnI. These changes in sequence altered the regulatory capacity of these proteins and coincided with the increased complexity of cardiac structure and function. By analyzing how TnI and MyBP-C have changed during vertebrate evolution we have provided new insight into how the functional characteristics of the different muscle types have evolved and have raised new questions into the regulatory control of contractile function.

ACKNOWLEDGMENTS

The authors acknowledge Courtney Deck for assisting with sequence compilation and Ian Smith for assisting with creation of Fig. 5. The authors also like to thank Sarah Alderman and Katharine Bell for editorial efforts.

GRANTS

This work was supported by a National Science Foundation (US) Graduate Research Fellowship to J. F. Shaffer and a National Sciences and Engineering Research Council Discovery (Canada) grant to T. E. Gillis.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

1. **Albalat R, Baquero M, Minguillon C.** Identification and characterisation of the developmental expression pattern of tbx5b, a novel tbx5 gene in zebrafish. *Gene Expr Patterns* 10: 24–30, 2010.
2. **Allen DG, Jewell BR, Murray JW.** The contribution of activation processes to the length-tension relation of cardiac muscle. *Nature* 248: 606–607, 1974.
3. **Anderson M.** Electrophysiological studies on initiation and reversal of the heart beat in Ciona intestinalis. *J Exp Biol* 49: 363–385, 1968.
4. **Bers DM.** Calcium cycling and signaling in cardiac myocytes. *Annu Rev Physiol* 70: 23–49, 2008.
5. **Blom N, Gammeltoft S, Brunak S.** Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* 294: 1351–1362, 1999.
6. **Burggren W, Farrell AP, Lillywhite H.** Vertebrate Cardiovascular Systems. In: *Handbook of Physiology* edited by Dantzler WH. New York: Oxford University Press, 1997, p. 215–308.
7. **Burkart EM, Sumandea MP, Kobayashi T, Nili M, Martin AF, Homsher E, Solaro RJ.** Phosphorylation or glutamic acid substitution at protein kinase C sites on cardiac troponin I differentially depress myofilament tension and shortening velocity. *J Biol Chem* 278: 11265–11272, 2003.
8. **Cairns SP, Dulhunty AF.** Beta-adrenergic potentiation of E-C coupling increases force in rat skeletal muscle. *Muscle Nerve* 16: 1317–1325, 1993.
9. **Chandra M, Dong WJ, Pan BS, Cheung HC, Solaro RJ.** Effects of protein kinase A phosphorylation on signaling between cardiac troponin I and the N-terminal domain of cardiac troponin C. *Biochemistry* 36: 13305–13311, 1997.
10. **Chong SM, Jin JP.** To investigate protein evolution by detecting suppressed epitope structures. *J Mol Evol* 68: 448–460, 2009.
11. **Churcott CS, Moyes CD, Bressler BH, Baldwin KM, Tibbits GF.** Temperature and pH effects on Ca^{2+} sensitivity of cardiac myofibrils: a comparison of trout with mammals. *Am J Physiol Regul Integr Comp Physiol* 267: R62–R70, 1994.
12. **Cleto CL, Vandenberghe AE, MacLean DW, Pannunzio P, Tortorelli C, Meedel TH, Satou Y, Satoh N, Hastings KE.** Ascidian larva reveals ancient origin of vertebrate-skeletal-muscle troponin I characteristics in chordate locomotory muscle. *Mol Biol Evol* 20: 2113–2122, 2003.
13. **Day SM, Westfall MV, Fomicheva EV, Hoyer K, Yasuda S, La Cross NC, D'Alecy LG, Ingwall JS, Metzger JM.** Histidine button engineered into cardiac troponin I protects the ischemic and failing heart. *Nat Med* 12: 181–189, 2006.
14. **Elias CL, Xue XH, Marshall CR, Omelchenko A, Hryshko LV, Tibbits GF.** Temperature dependence of cloned mammalian and salmonid cardiac Na^+/Ca^{2+} exchanger isoforms. *Am J Physiol Cell Physiol* 281: C993–C1000, 2001.
15. **Farmer CG.** Evolution of the vertebrate cardio-pulmonary system. *Ann Ref Physiol* 61: 573–592, 1999.
16. **Felsenstein J.** Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol* 266: 418–427, 1996.
17. **Gautel M, Zuffardi O, Freiburg A, Labeit S.** Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? *EMBO J* 14: 1952–1960, 1995.
18. **Gillis TE, Marshall CR, Tibbits GF.** Functional and evolutionary relationships of troponin C. *Physiol Genomics* 32: 16–27, 2007.
19. **Gillis TE, Marshall CR, Xue XH, Borgford TJ, Tibbits GF.** Ca^{2+} binding to cardiac troponin C: effects of temperature and pH on mammalian and salmonid isoforms. *Am J Physiol Regul Integr Comp Physiol* 279: R1707–R1715, 2000.
20. **Gillis TE, Tibbits GF.** Beating the cold: the functional evolution of troponin C in teleost fish. *Comp Biochem Physiol A Mol Integr Physiol* 132: 763–772, 2002.
21. **Gillis WQ, St John J, Bowerman B, Schneider SQ.** Whole genome duplications and expansion of the vertebrate GATA transcription factor gene family. *BMC Evol Biol* 9: 207, 2009.
22. **Good-Avila SV, Yegorov S, Harron S, Bogerd J, Glen P, Ozon J, Wilson BC.** Relaxin gene family in teleosts: phylogeny, syntenic mapping, selective constraint, and expression analysis. *BMC Evol Biol* 9: 293, 2009.
23. **Gordon AM, Homsher E, Regnier M.** Regulation of contraction in striated muscle. *Physiol Rev* 80: 853–924, 2000.
24. **Gruen M, Prinz H, Gautel M.** cAPK-phosphorylation controls the interaction of the regulatory domain of cardiac myosin binding protein C with myosin-S2 in an on-off fashion. *FEBS Lett* 453: 254–259, 1999.
25. **Gwathmey JK, Hajjar RJ.** Calcium-activated force in a turkey model of spontaneous dilated cardiomyopathy: adaptive changes in thin myofilament Ca^{2+} regulation with resultant implications on contractile performance. *J Mol Cell Cardiol* 24: 1459–1470, 1992.
26. **Harrison SM, Bers DM.** Modification of temperature dependence of myofilament Ca^{2+} sensitivity by troponin C replacement. *Am J Physiol Cell Physiol* 258: C282–C288, 1990.
27. **Harrison SM, Bers DM.** Temperature dependence of myofilament Ca^{2+} sensitivity of rat, guinea pig, and frog ventricular muscle. *Am J Physiol Cell Physiol* 258: C274–C281, 1990.
28. **Hufton AL, Groth D, Vingron M, Lehrach H, Poustka AJ, Panopoulou G.** Early vertebrate whole genome duplications were predated by a period of intense genome rearrangement. *Genome Res* 18: 1582–1591, 2008.
29. **Jacobs GH.** Evolution of colour vision in mammals. *Philos Trans R Soc Lond B Biol Sci* 364: 2957–2967, 2009.
30. **Jia W, Shaffer JF, Harris SP, Leary JA.** Identification of novel protein kinase A phosphorylation sites in the M-domain of human and murine cardiac myosin binding protein-C using mass spectrometry analysis. *J Proteome Res* 9: 1843–1853, 2010.
31. **Kobayashi T, Solaro RJ.** Calcium, thin filaments, and the integrative biology of cardiac contractility. *Ann Rev Physiol* 67: 39–67, 2005.
32. **Korte FS, McDonald KS, Harris SP, Moss RL.** Loaded shortening, power output, and rate of force redevelopment are increased with knockout of cardiac myosin binding protein-C. *Circ Res* 93: 752–758, 2003.
33. **Koshiba-Takeuchi K, Mori AD, Kaynak BL, Cebra-Thomas J, Sukonnik T, Georges RO, Latham S, Beck L, Henkelman RM, Black BL, Olson EN, Wade J, Takeuchi JK, Nemer M, Gilbert SF, Bruneau BG.** Reptilian heart development and the molecular basis of cardiac chamber evolution. *Nature* 461: 95–98, 2009.
34. **Kuraku S, Meyer A, Kuratani S.** Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol Biol Evol* 26: 47–59, 2009.
35. **Larhammar D, Nordstrom K, Larsson TA.** Evolution of vertebrate rod and cone phototransduction genes. *Philos Trans R Soc Lond B Biol Sci* 364: 2867–2880, 2009.
36. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948, 2007.
37. **Layland J, Solaro RJ, Shah AM.** Regulation of cardiac contractile function by troponin I phosphorylation. *Cardiovasc Res* 66: 12–21, 2005.
38. **Lim MS, Sutherland C, Walsh MP.** Phosphorylation of bovine cardiac C-protein by protein kinase C. *Biochem Biophys Res Commun* 132: 1187–1195, 1985.
39. **Mohamed AS, Dignam JD, Schlender KK.** Cardiac myosin-binding protein C (MyBP-C): identification of protein kinase A and protein kinase C phosphorylation sites. *Arch Biochem Biophys* 358: 313–319, 1998.
40. **Nevitt G, Gilly WF.** Morphological and physiological properties of non-striated muscle from the tunicate, Ciona intestinalis: parallels with vertebrate skeletal muscle. *Tissue Cell* 18: 341–360, 1986.
41. **Noland TA Jr, Guo X, Raynor RL, Jideama NM, Averyhart-Fullard V, Solaro RJ, Kuo JF.** Cardiac troponin I mutants phosphorylation by protein kinases C and A and regulation of Ca^{2+}-stimulated MgATPase of reconstituted actomyosin S-1. *J Biol Chem* 270: 25445–25454, 1995.
42. **Noland TA Jr, Raynor RL, Jideama NM, Guo X, Kazanietz MG, Blumberg PM, Solaro RJ, Kuo JF.** Differential regulation of cardiac

actomyosin S-1 MgATPase by protein kinase C isozyme-specific phosphorylation of specific sites in cardiac troponin I and its phosphorylation site mutants. *Biochemistry* 35: 14923–14931, 1996.
43. **Oota S, Saitou N.** Phylogenetic relationship of muscle tissues deduced from superimposition of gene trees. *Mol Biol Evol* 16: 856–867, 1999.
44. **Patrick SM, Hoskins AC, Kentish JC, White E, Shiels HA, Cazorla O.** Enhanced length-dependent Ca^{2+} activation in fish cardiomyocytes permits a large operating range of sarcomere lengths. *J Mol Cell Cardiol* 48: 917–924, 2010.
45. **Rambaut A.** *FigTree*, version 1.2.2, 2006.
46. **Robertson SP, Johnson JD, Holroyde MJ, Kranias EG, Potter JD, Solaro RJ.** The effect of troponin I phosphorylation on the Ca^{2+}-binding properties of the Ca^{2+}-regulatory site of bovine cardiac troponin. *J Biol Chem* 257: 260–263, 1982.
47. **Saitou N, Nei M.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406–425, 1987.
48. **Santini F, Harmon LJ, Carnevale G, Alfaro ME.** Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol Biol* 9: 194, 2009.
49. **Schmidt-Nielsen K.** Locomotion: energy cost of swimming, flying, and running. *Science* 177: 222–228, 1972.
50. **Scruggs SB, Hinken AC, Thawornkaiwong A, Robbins J, Walker LA, de Tombe PP, Geenen DL, Buttrick PM, Solaro RJ.** Ablation of ventricular myosin regulatory light chain phosphorylation in mice causes cardiac dysfunction in situ and affects neighboring myofilament protein phosphorylation. *J Biol Chem* 284: 5097–5106, 2009.
51. **Shaffer JF.** HS Species-specific differences in the Pro-Ala rich region of cardiac myosin binding protein-C. *J Muscle Res Cell Motil* 30: 303–306, 2010.
52. **Shaffer JF, Kensler RW, Harris SP.** The myosin-binding protein C motif binds to F-actin in a phosphorylation-sensitive manner. *J Biol Chem* 284: 12318–12327, 2009.
53. **Shaffer JF, Wong P, Bezold KL, Harris SP.** Functional differences between the N-terminal domains of mouse and human cardiac myosin binding protein-C. *J Biotechnology Biomedicine* Epub 2010 Apr 7.
54. **Shiels HA, White E.** The Frank-Starling mechanism in vertebrate cardiac myocytes. *J Exp Biol* 211: 2005–2013, 2008.
55. **Solaro RJ, Moir AJ, Perry SV.** Phosphorylation of troponin I and the inotropic effect of adrenaline in the perfused rabbit heart. *Nature* 262: 615–617, 1976.
56. **Somero GN.** Proteins and temperature. *Annu Rev Physiol* 57: 43–68, 1995.
57. **Stelzer JE, Patel JR, Walker JW, Moss RL.** Differential roles of cardiac myosin-binding protein C and cardiac troponin I in the myofibrillar force responses to protein kinase A phosphorylation. *Circ Res* 101: 503–511, 2007.
58. **Stull JT, High CW.** Phosphorylation of skeletal muscle contractile proteins in vivo. *Biochem Biophys Res Commun* 77: 1078–1083, 1977.
59. **Stull JT, Manning DR, High CW, Blumenthal DK.** Phosphorylation of contractile proteins in heart and skeletal muscle. *Fed Proc* 39: 1552–1557, 1980.
60. **Suzuki M, Sasaki N, Miki T, Sakamoto N, Ohmoto-Sekine Y, Tamagawa M, Seino S, Marban E, Nakaya H.** Role of sarcolemmal K(ATP) channels in cardioprotection against ischemia/reperfusion injury in mice. *J Clin Invest* 109: 509–516, 2002.
61. **Tong CW, Stelzer JE, Greaser ML, Powers PA, Moss RL.** Acceleration of crossbridge kinetics by protein kinase A phosphorylation of cardiac myosin binding protein C modulates cardiac function. *Circ Res* 103: 974–982, 2008.
62. **Weisberg A, Winegrad S.** Alteration of myosin cross bridges by phosphorylation of myosin-binding protein C in cardiac muscle. *Proc Natl Acad Sci USA* 93: 8999–9003, 1996.
63. **Westfall MV, Albayya FP, Metzger JM.** Functional analysis of troponin I regulatory domains in the intact myofilament of adult single cardiac myocytes. *J Biol Chem* 274: 22508–22516, 1999.
64. **Westfall MV, Albayya FP, Turner II, Metzger JM.** Chimera analysis of troponin I domains that influence Ca^{2+}-activated myofilament tension in adult cardiac myocytes. *Circ Res* 86: 470–477, 2000.
65. **Westfall MV, Borton AR.** Role of troponin I phosphorylation in protein kinase C-mediated enhanced contractile performance of rat myocytes. *J Biol Chem* 278: 33694–33700, 2003.
66. **Westfall MV, Metzger JM.** Single amino acid substitutions define isoform-specific effects of troponin I on myofilament Ca^{2+} and pH sensitivity. *J Mol Cell Cardiol* 43: 107–118, 2007.
67. **Westfall MV, Rust EM, Metzger JM.** Slow skeletal troponin I gene transfer, expression, and myofilament incorporation enhances adult cardiac myocyte contractile function. *Proc Natl Acad Sci USA* 94: 5444–5449, 1997.
68. **Westfall MV, Turner I, Albayya FP, Metzger JM.** Troponin I chimera analysis of the cardiac myofilament tension response to protein kinase A. *Am J Physiol Cell Physiol* 280: C324–C332, 2001.
69. **Xiao L, Zhao Q, Du Y, Yuan C, Solaro RJ, Buttrick PM.** PKCepsilon increases phosphorylation of the cardiac myosin binding protein C at serine 302 both in vitro and in vivo. *Biochemistry* 46: 7054–7061, 2007.
70. **Yandell M, Mungall CJ, Smith C, Prochnik S, Kaminker J, Hartzell G, Lewis S, Rubin GM.** Large-scale trends in the evolution of gene structures within 11 animal genomes. *PLoS Comput Biol* 2: e15, 2006.
71. **Yasuda S, Coutu P, Sadayappan S, Robbins J, and Metzger JM.** Cardiac transgenic and gene transfer strategies converge to support an important role for troponin I in regulating relaxation in cardiac myocytes. *Circ Res* 101: 377–386, 2007.
72. **Yuasa HJ, Kawamura K, Yamamoto H, Takagi T.** The structural organization of ascidian Halocynthia roretzi troponin I genes. *J Biochem (Tokyo)* 132: 135–141, 2002.