# Mechanism and Functionality of Species-Specific *Camk2β* Alternative Splicing

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

Andreas Franz

2022

This work was carried out in the period of September 2017 to April 2022 under the supervision of Prof. Dr. Markus Wahl and Prof. Dr. Florian Heyd at the Institute of Chemistry and Biochemistry, Freie Universität Berlin, Germany.

First reviewer:
Prof. Dr. Markus Wahl
Institute for Chemistry and Biochemistry
Laboratory of Structural Biochemistry
Freie Universität Berlin
Takustraße 6
14195 Berlin, Germany

Second reviewer:
Prof. Dr. Florian Heyd
Institute for Chemistry and Biochemistry
Laboratory of RNA Biochemistry
Freie Universität Berlin
Takustraße 6
14195 Berlin, Germany

Date of defence: 27.06.2022

# Acknowledgements

First, I would like to thank my two supervisors Prof. Florian Heyd and Prof. Markus Wahl for giving me the opportunity to work on such an interesting and challenging project. I want thank Prof. Florian Heyd for developing this research idea, his constant support and extremely helpful and motivating discussions. I am especially thankful for the tremendous freedom in pursuing my own research ideas. I want to thank Prof. Markus Wahl for his steady support, organizing diverse research collaborations and his very helpful insights and discussions.

I want to further thank all the following people who have supported me during these last years: Dr. Marco Preußner for direct supervision during the initial phase of this project and his constant support and tremendously helpful recommendations. Dr. Tarek Hilal and Dr. Bernhard Loll for support with structural biology and cryo-EM. Dr. Eva Absmeier for support in establishing and troubleshooting protein purification strategies. Nicole Dimos for her very productive master thesis under my supervision, which lead to some extremely helpful insights. Esta Kempf for her master thesis under my supervision, her work on tau aggregation and proofreading this thesis. Ioana Weber for her support with everything mouse- and microscopy-related, her general support in the CaMKII project and proofreading this thesis. Lena Grass for valuable discussions on protein-purification. Claudia Alings, Nicole Holton and Clemens Langner for helping with everyday problems. Antje Grünwald and Bobby Viet Draegert for help with the mouse facility. Dr. Iva Lucic for very helpful discussions regarding the kinetics of CaMKII. Our collaboration partners from the research group of Prof. Henning Urlaub, especially Dr. Yanlong Ji, for help with the mass spectrometry analysis. Benno Kuropka and Chris Weise also for the help with everything regarding mass spectrometry. Our collaboration partners from the research group of Prof. Dietmar Schmitz (Charité Universitätsmedizin) especially Dr. Alexander Stumpf.

I want to thank all the current and former members of the AGs Wahl and Heyd for creating such a motivating and delightful working atmosphere.

Lastly, I want to thank all my family and friends for their support during these last years.

# Selbstständigkeitserklärung

Hierdurch versichere ich, Andreas Franz, dass ich die hier vorliegende Arbeit selbstständig verfasst und keine anderen als die von mir angegeben Quellen und Hilfsmittel verwendet habe. Ich versichere, dass ich meine Dissertation nicht schon einmal in einem anderen Promotionsverfahren eingereicht habe.

Berlin, den 27.04.2022 _____

# Table of Contents

# Abstract

Species-specific alternative splicing has been recognized for many years, but its regulation has remained enigmatic and its functionality is still being discussed. Calcium/calmodulin-dependent protein kinase II (CaMKII) is a key player in learning and memory in vertebrates and highly conserved across evolution. In mammals, CaMKII is encoded by four homologous genes and alternative splicing leads to multiple protein isoforms. Here, I characterize *Camk2β* alternative splicing and reveal several primate-specific splice isoforms, which are generated through exclusion of exon 16. These isoforms show altered kinetic properties, changes in substrate specificity and differential intracellular localization. Furthermore, I show that primate-specific *Camk2β* alternative splicing is *cis*-regulated, as a human minigene maintains exon skipping in different *trans*-acting environments. I demonstrate that weakening of the branch point during evolution is sufficient for primate-specific *Camk2β* exon 16 skipping, thus providing a paradigm for *cis*-directed species-specific alternative splicing regulation. Building on this observation, we used CRISPR/Cas9 and introduced the human intronic *cis*-acting element, containing the identified branch point sequence, into the mouse genome. This indeed results in a human-like *Camk2β* splicing pattern in the brain of mutant mice. Electrophysical characterization of CA1 neurons of mutant mice reveals a strong impairment of long-term potentiation (LTP), thus connecting species-specific alternative splicing with a fundamental function in learning and memory. The mutant mice exhibit various behavioural alterations, highlighting the importance of a correct balance of CaMKII isoforms for normal brain function. By studying species-specific *Camk2β* alternative splicing on a molecular, cellular and organismal level, this work contributes to the understanding of species-specific splicing regulation and reveals a role of alternative splicing in generating diverse, functional proteomes and species-specific functionality.

# Zusammenfassung

Spezies-spezifisches alternatives Spleißen ist seit vielen Jahren bekannt, aber seine Regulierung ist nach wie vor rätselhaft und die genaue Funktionalität wird noch immer diskutiert. Die Calcium/Calmodulin-abhängige Proteinkinase II (CaMKII) spielt eine Schlüsselrolle beim Lernen und Gedächtnis von Wirbeltieren und die grundlegende Funktionsweise ist in allen Spezies hoch konserviert. In Säugetieren wird CaMKII von vier homologen Genen kodiert, welche durch alternatives Spleißen eine Vielzahl an Protein-Isoformen bilden. In dieser Arbeit charakterisiere ich das alternative Spleißen von *Camk2β* in verschiedenen Spezies und entdecke mehrere Primaten-spezifische Spleiß-Isoformen, die durch den Ausschluss von Exon 16 entstehen. Diese Isoformen zeigen veränderte kinetische Eigenschaften, Änderungen in der Substratspezifität und eine unterschiedliche intrazelluläre Lokalisierung. Darüber hinaus zeige ich, dass das primatenspezifische alternative *Camk2β*-Spleißen in *cis* kontrolliert wird, da ein menschliches Minigen das Exon-Skipping in verschiedenen *trans*-wirkenden Umgebungen aufrechterhält. Ich zeige, dass die Schwächung des „Branch Points“ im Laufe der Evolution das Primaten-spezifische *Camk2β* Exon 16-Skipping ausreichend erklärt, und liefere damit ein Paradigma für die *cis*-gesteuerte Regulation von spezies-spezifischen alternativem Spleißen. Aufbauend auf diesen Beobachtungen haben wir mittels Crispr/Cas9 das menschliche intronische *cis*-wirkende Element in das Mausgenom eingeführt, was wie erwartet zu einem menschenähnlichen *Camk2β*-Spleißmuster im Gehirn der veränderten Mäuse führt. Diese Tiere zeigen eine massive Beeinträchtigung der Langzeitpotenzierung (LTP) in CA1-Neuronen, was auf eine Verbindung zwischen dem artspezifischen alternativen Spleißen und einer grundlegenden Funktion bei Lernen und Gedächtnis hindeutet. Die Mäuse weisen zudem verschiedene Verhaltensmerkmale auf, was die Bedeutung des korrekten Gleichgewichts der CaMKII-Isoformen für eine normale Gehirnfunktion unterstreicht. Durch die Untersuchung des spezies-spezifischen alternativen Spleißens von *Camk2β* auf molekularer und zellulärer Ebene, sowie im intakten Organismus, trägt diese Arbeit zum Verständnis der artspezifischen Regulation von alternativen Spleißen bei. Sie offenbart eine Rolle von alternativen Spleißen bei der Erzeugung unterschiedlicher, funktioneller Proteome, sowie artspezifischer Funktionalität.

# Introduction

## The mechanism of alternative splicing

Splicing was discovered in the late 1970s, with the observation that segments from adenovirus messenger RNA are removed during its maturation (Berget et al., 1977; Chow et al., 1977). Alternative splicing was reported soon after, when it was discovered that the same gene could encode for both membrane-bound and secreted antibodies (Alt et al., 1980; Early et al., 1980). Splicing describes the process in which non-coding elements, termed introns, are removed from precursor messenger RNA (pre-mRNA) and the resulting coding elements, termed exons, are joined together to form the final messenger RNA (mRNA). In alternative splicing, different sets of exons from a pre-mRNA are spliced together, leading to different mRNA isoforms and vastly amplifying the palette of translated proteins. In humans, almost every protein-coding gene contains introns. While exons are generally short, containing 50-300 nucleotides (nt), introns have a median length of 1-2 kilobases (kb), but can be considerably longer. The maturation process of pre-mRNA also includes the addition of a 7-methylguanosine (m7G) cap to the 5´-end of the transcript (Shatkin, 1976) and polyadenylation at the 3’ end (Colgan and Manley, 1997; Wahle and Keller, 1992). Both processes are essential for mRNA stability, nuclear export and initiation of translation. The poly(A)-tail further regulates mRNA levels via the interaction with a number of RNA-binding proteins (RNABPs) and other RNA species, such as micro RNAs (Tian and Manley, 2017).

Multiple modes of alternative splicing exist (Black, 2003; Ule and Blencowe, 2019) (Figure 1A): Cassette exons are among the most frequently observed events in humans and describe exons that are either included or excluded. Mutually exclusive exons are a subgroup of cassette exons that are only included individually. Alternative 5’ and 3’ splice sites can lengthen or shorten exons. Intron retention on the other hand describes a process in which originally non-coding elements are retained in the final mRNA. In contrast to cassette exons, these introns are themselves flanked by exons. Not strictly alternative splicing events themself, alternative transcription start sites or alternative polyadenylation sites can affect the 5’ or 3’-proximal exons.

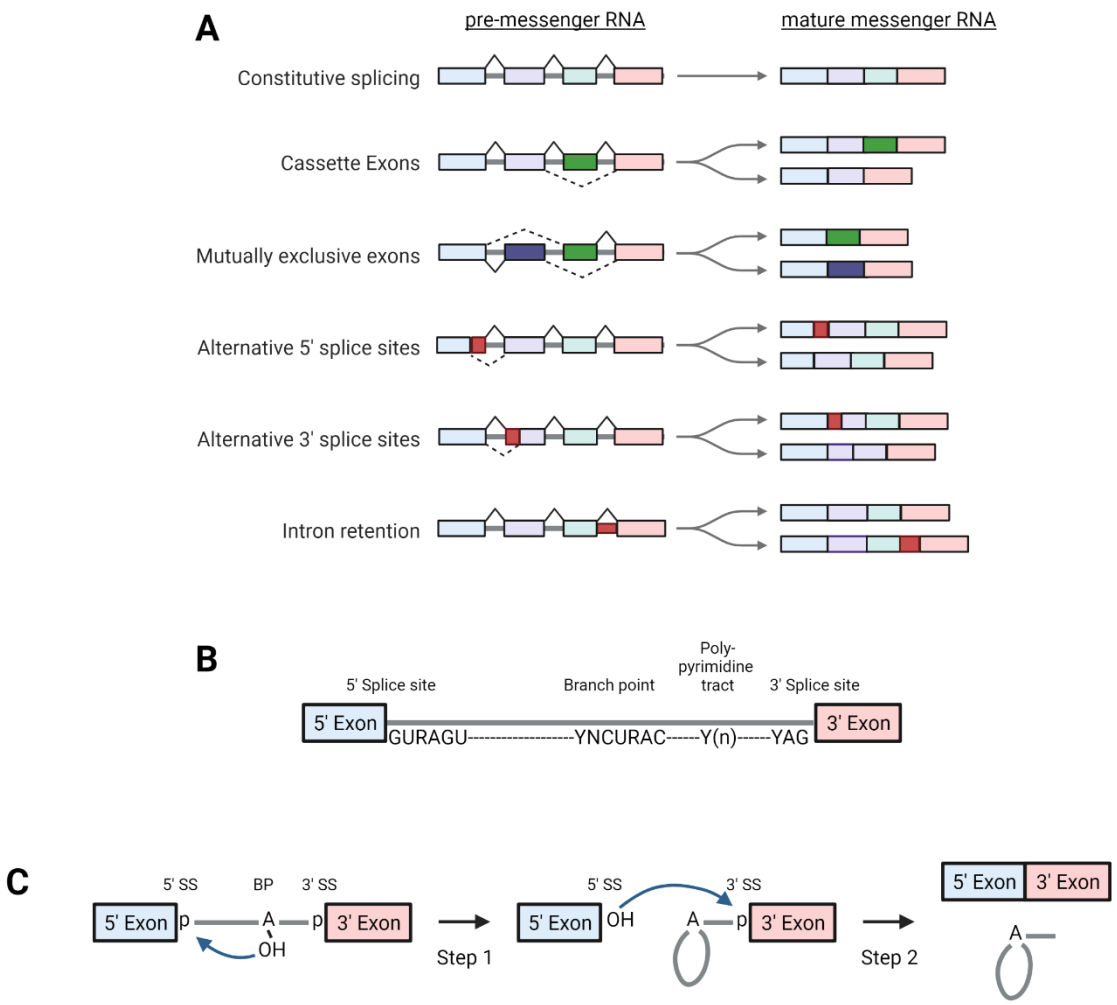



Figure 1: The principles of alternative splicing.

(A) Categories of alternative splicing events. Cassette exons can either be included or excluded in the final mRNA. Mutually exclusive exons are included only one at a time. Alternative 5' or 3' splice sites can lengthen or shorten an exon. Intron retention indicates the retaining of a normally noncoding sequence that itself is surrounded by exons. Adapted from (Black, 2003) (B) Conserved sequences that define intronic and exonic elements. These are the upstream 5' splice site, the branch point sequence containing the branch point adenine, the downstream polypyrimidine track and the 3' splice site. (C) Schematic of the splicing reaction. Introns are removed in two steps. In the first step, branching, the 3' OH group of the branch point adenine attacks the phosphodiester bond at the 5' splice site. In a second step, exon-ligation, the free 2' OH group of the upstream exon attacks the phosphodiester bond at the 3'SS, generating ligated exons and an intron lariat. Adapted from (Kastner et al., 2019; Plaschka et al., 2019).

Introns are defined by three short consensus sequences, the 5’ splice site (SS), the 3’SS and the branch point (BP) sequence. The BP is usually found 18-40 nt upstream of the 3’SS and often followed by a polypyrimidine track (PPT) (Figure 1B). In higher eukaryotes and especially for alternative exons, these sequences are often degenerated and deviate more from the consensus motif, leading to a lower affinity for core components of the splicing machinery. The reaction itself consists of two

consecutive $S_N2$-type transesterification reactions (Figure 1C). In a first step, the 2'-OH group of the BP adenosine carries out a nucleophilic attack on the phosphodiester bond at the 5'SS. This leads to a free 5' exon and an intron lariat-3'-exon intermediate. In the second step, the 3'-OH group of the 5' exon attacks the phosphodiester bond at the 3'SS, generating ligated exons and a lariat intron. Splicing is performed by the spliceosome, a highly dynamic, multi-megadalton RNA-protein machinery (Kastner *et al.*, 2019; Plaschka *et al.*, 2019; Shi et al., 2019; Wahl et al., 2009).

The major spliceosome consists of five small nuclear ribonucleoproteins (snRNPs, U1, U2, U4, U5, U6), each containing a small nuclear RNA (or two in the case of U4/U6), seven Sm proteins that form a ring (Lsm proteins for U6) and a variable number of additional proteins (Will and Lührmann, 2011). The spliceosome is less complex in simple eukaryotes such as yeast, containing around ~90 proteins (Fabrizio et al., 2009; Kastner *et al.*, 2019). In humans, the degeneration of consensus sequences allows for more plasticity and around 170 proteins are known to associate with the spliceosome (Wahl *et al.*, 2009). The spliceosome assembles on the pre-mRNA in a highly dynamic and stepwise manner on every substrate anew. Extensive structural and compositional rearrangements are needed to form the active site and carry out the two splicing reactions. At least ten distinct spliceosome states can be defined, including the E, A, pre-B, B, $B^{act}$, B*, C, C*, P, and intron lariat spliceosome (ILS) complexes (Kastner *et al.*, 2019; Plaschka *et al.*, 2019; Shi *et al.*, 2019). In a very short overview, the 5'SS is initially recognized by the U1 snRNP, whereas the BP is bound by the U2 snRNP. The U2-BP interaction is reinforced by the U2AF65/35 heterodimer, which bind the PPT. Additional interaction comes from the SF3a and SF3b complexes, contacting the intron upstream of the BP. Recruitment of the tri-snRNP U4/U6.U5 forms the pre-B complex, which is subsequently remodelled. U4 is displaced and additional proteins form the $B^{act}$ complex, later transformed into the catalytically competent B* complex. This complex carries out the first "branching" reaction, after which U1 and U4 are released, resulting in the C complex. After another structural rearrangement, the activated C* complex is formed and catalyses the second splicing reaction, termed "exon ligation". The 5' and 3' exons are end-joined, generating the P-complex. Release of the spliced mRNA results in the intron lariat spliceosome (ILS) complex, which after dissociation of the intron-lariat is disassembled and can be recycled for a new splicing reaction. Structural and compositional changes of the spliceosome are mainly driven by eight conserved

helicases, which trigger RNA-RNA and RNA-protein rearrangements and help in proofreading the selection of the correct intron sequences.

Introns can be exceedingly long and are sometimes removed in multiple steps, termed recursive splicing. This has initially been reported for a small number of long *Drosophila* (Burnette et al., 2005; Hatton et al., 1998) and human (Duff et al., 2015; Sibley et al., 2015) introns and is associated with increased splicing accuracy (Pai et al., 2018). A recent study found recursive splicing to be prevalent across the human genome, occurring in virtually all introns and being preferred for introns above 10 kb length (Wan et al., 2021). Similar to the stochastic nature of transcription, where short active periods (transcriptional “bursting”) are followed by inactive periods of variable length (Rodriguez et al., 2019), a model of stochastic splice site selection was suggested. The outcome of this process is still highly precise, and in some sense “deterministic”, but the exact sequence of events, namely the choice of recursive splice sites, is stochastic in nature.

## Splicing regulatory networks

Despite recent challenges towards the deterministic model of splicing, the question of how intronic and exonic sequences are recognized or “defined” by the spliceosome is of utmost importance (De Conti et al., 2013). A special focus are “bridging interactions”, which connect spliceosome components that recognize separate sequence elements of the pre-mRNA.

Apart from core splicing factors, interactions by additional, variable components contribute to the correct exon and intron definition and facilitate recruitment of the spliceosome. These additional factors often act early in the stepwise assembly of the splicing machinery. As such, splicing and especially alternative splicing is controlled by a network of so called *cis*-acting elements and *trans*-acting factors (Ule and Blencowe, 2019). *Cis*-acting elements are sequences on the pre-mRNA that regulate splicing of adjacent exons or introns. They are categorized based on their location and effect on the splicing reaction into exonic splicing enhancer (ESE), exonic splicing silencer (ESS), intronic splicing enhancer (ISE) or intronic splicing silencer (ISS) motifs. Most of these elements are found in close proximity to the exons they regulate and the majority of them are found within the first 300 exon-proximal nucleotides (Barash et al., 2010; Lee and Rio, 2015). These elements are bound by *trans*-acting proteins which in turn interact with components of the spliceosome, thereby facilitating the exon/ intron definition interactions (De Conti *et al.*, 2013). SR proteins and hnRNPs (heterogeneous nuclear ribonucleoproteins) are classic regulators of alternative splicing, but many additional factors exist. These often contain typical RNA-binding domains, such as RNA-recognition motifs (RRMs), K homology domains or zinc-finger domains (Ule and Blencowe, 2019). SR proteins contain RRMs and an eponymous RS-domain, which is intrinsically disordered and rich in arginine and serine residues. These domains mediate heteromeric interactions that promote the combinatorial assembly of the spliceosome, and therefore typically act as enhancers when binding between other components. SR proteins typically bind to exonic enhancer sequences, but can also bind to intronic sequences where they often counteract the suppressive effects of hnRNPs (Chen and Manley, 2009; Fu and Ares, 2014). Showcasing the flexibility of alternative splicing regulation, SR proteins have also been found to repress splicing when binding to intronic splicing silencer motifs downstream of a target exon (Erkelenz et al., 2013). hnRNPs typically repress splicing of nearby exons, sometimes

by sterically hindering the binding of core spliceosome components via overlapping binding sites. They often repress cryptic splice sites within introns (Sibley et al., 2016), which can act as decoys by competing with canonical splice sites (Parra et al., 2018). Similar to SR proteins, their effect on a given exon is strongly position-dependent (Fu and Ares, 2014; Ule and Blencowe, 2019). These regulatory principles also hold true for splicing regulators other than SR proteins and hnRNPs, such as NOVA and TIA, for which opposing effects have been reported depending on the exact location of their binding site (Ule et al., 2006; Wang et al., 2010).

## Species-specific alternative splicing

Position-dependent splicing rules are similar across many core splicing regulators and other RBPs and have furthermore been conserved across cell types and species (Ajith et al., 2016; Brooks et al., 2011; Ule and Blencowe, 2019; Witten and Ule, 2011). For instance, splicing rules inferred from mouse data can be used to successfully predict human splicing events (Barbosa-Morais et al., 2012). RBPs often operate in a condition-specific manner and regulate functionally coherent ‘‘networks’’ of exons and introns (Braunschweig et al., 2013). Consequently, alternative splicing is frequently regulated in a tissue-, developmental-, disease-, circadian-, temperature and also species-specific manner (Barbosa-Morais *et al.*, 2012; Preußner et al., 2017; Preußner et al., 2014; Ule and Blencowe, 2019). Splicing differences between individuals have also been reported (Graveley, 2008). On a functional level, alternative splicing has been shown to diversify protein interaction capabilities (Buljan et al., 2012; Ellis et al., 2012; Yang et al., 2016) and a role in establishing higher phenotypic complexity has long been known (Barbosa-Morais *et al.*, 2012; Hillier et al., 2009; Stolc et al., 2004).

Whereas the number of genes roughly correlates with the complexity of an organism for unicellular species such as *Escherichia Coli* or *Saccharomyces cerevisiae*, this does not hold true for higher eukaryotes. Already during early stages of the human genome project and similar efforts, it was revealed that the number of genes in vertebrates is far below the number anticipated necessary to infer the observed level of phenotypic complexity. Early predictions thus suggested transcriptome diversity, generated by alternative splicing, to be key in creating biological complexity (Ewing and Green, 2000). In general, the frequency of alternative splicing through exon skipping has increased during animal evolution, with the highest skipping frequencies detected in the primate nervous system (Barbosa-Morais *et al.*, 2012; Kim et al., 2006). This is in part mediated by an increase in intron density, the shortening of exons and the presence of heterogeneous splice sites in animal genomes (Csuros et al., 2011; Grau-Bové et al., 2018). The general increase in alternative splicing in mammals is strongly enriched in frame-preserving mutations, suggesting a functional relevance (Grau-Bové *et al.*, 2018). Interestingly, the process of gaining introns has not been a one-way path during evolution and the highest intron-density is predicted for the last common ancestor of animals (Csuros *et al.*, 2011) and not in currently living species, including humans.

New alternative exons can be generated by exon shuffling, exonization of intronic sequences, and the transition of constitutive to alternative exons (Keren et al., 2010). Exon shuffling describes the insertion of a pre-existing exon from one gene into another, or the duplication of a pre-existing exon of the same gene. As exons often correlate with protein domains, this enables the modular building of genes and proteins (Liu and Grigoriev, 2004). Exonization of intronic sequences often affects transposable elements such as Alu elements (Lev-Maor et al., 2003), which are more prevalent in primate genomes (Zhang and Chasin, 2006). These new exons are often alternative: whereas most constitutive exons in humans and mice have been present in their last common ancestor, 72% of alternative exons were created post-split (Modrek and Lee, 2003). Exon-loss also exists, but is much less frequent compared to exon-gain (Wang et al., 2015).

Overall, alternative splicing patterns have rapidly diverged between species (Modrek and Lee, 2003; Pan et al., 2004) and are now more similar between different organs within a species, than they are between the same organs of different species (Barbosa-Morais *et al.*, 2012; Merkin et al., 2012). The highest level of divergence has been found in neuronal tissue and the frequency of alternative splicing swiftly decreases with evolutionary distance from humans, even within the order of primates. Most of these splicing events appear *cis*-regulated, confirming the earlier hypothesis that the regulatory maps of RBPs have been largely conserved during evolution. The binding codes of these *trans*-acting factors seem largely invariant, whereas the regulatory modules and genes they affect are highly plastic and more likely to vary during evolution (Brooks *et al.*, 2011; Ule and Blencowe, 2019). These species-specific differences in splicing are not limited to animals, since similar observations have been made for various plant species (Kannan et al., 2018; Shi *et al.*, 2019). The phenomenon is also not restricted to splicing alone: species-specific conversions of regulatory elements, such as the transition from a transcription enhancer to a promotor sequence, have also been reported (Carelli et al., 2018). This can lead to species-specifically transcribed loci or 5’ exons. Gene expression patterns on the other hand are predominantly tissue- or organ-specific and have largely been conserved during vertebrate evolution (Barbosa-Morais *et al.*, 2012; Lin et al., 2014).

Many individual examples of species-specific alternative splicing events have been reported. The synaptic apolipoprotein E receptor 2 (*Apoer2*) gene shows mammalian-

specific exon gain, as well as mouse-specific exon skipping events (Gallo et al., 2022). Another clinically-relevant gene that undergoes species-specific regulation is *trem2*, a known risk-factor for Alzheimer's disease (Yanaizu et al., 2020). Strain-specific splicing of a gene involved in pigmentation was reported for a rapidly speciated fish species, potentially affecting mate choice (Terai et al., 2003). Modulation of a heat-sensitive channel (*Trpv1*) into sensing infrared radiation has been discovered in vampire bats (Gracheva et al., 2011). A human-specific splicing event has been reported for the *Prmt5* gene, controlling cell cycle arrest (Sohail and Xie, 2015). One of the most studied examples is the mammalian-specific splicing of *Ptbp1* exon 9, which is involved in neurogenesis (Barbosa-Morais *et al.*, 2012). Artificial induction of exon 9 skipping in chicken cells results in mammalian-like splicing changes (Gueroussov et al., 2015). Lastly, strain-specific splicing of *Camk2.1* has been indicated as a mechanism for natural adaptation of circadian timing in the marine midge *Clunio marinus* (Kaiser et al., 2016). The last example is of special interest, as the vertebrate orthologs of this gene have been strongly conserved during evolution and identified as key factors of neuronal plasticity. A potential species-specific regulation could thus have profound repercussions on establishing cognitive abilities in higher mammals.

## Calcium-calmodulin dependent protein kinase II

The calcium/calmodulin-dependent protein kinase II (CaMKII) is a unique serine/threonine-specific protein kinase that is involved in numerous regulatory pathways and plays an important role in learning and memory in vertebrates (Hell, 2014). In neuronal signalling, CaMKII plays a central role in the transduction of the cellular calcium influx into the phosphorylation of ion channels, a key mechanism underlying synaptic plasticity (Hudmon and Schulman, 2002). A unique feature of this kinase is the ability to not only respond to the amplitude, but also to the frequency of the activation signal. When the calcium frequency spike exceeds a characteristic threshold, the enzyme is able to adopt a calcium-independent activation state, which persists even in the absence of an extracellular activation signal (Chao et al., 2011; Meyer et al., 1992). This process is considered to be one of the fundamental mechanisms underlying long-term potentiation (LTP) (Malenka and Bear, 2004).

The CaMKII monomer consists of a highly conserved N-terminal Ser/Thr-specific kinase domain, followed by a regulatory segment that traps the enzyme in an autoinhibited state (Figure 2A). A linker domain of variable length connects the regulatory segment to a hub or association domain that enables oligomerization of the enzyme. The hub domains of multiple subunits assemble into two hexameric rings that form a dodecameric holoenzyme, in which the catalytic domains are in close proximity to facilitate rapid *trans*-autophosphorylation (Chao *et al.*, 2011; Rellos et al., 2010) (Figure 2B). The regulatory segment features three phosphorylation sites of functional importance and a calcium/calmodulin binding motif. Upon the binding of calmodulin, the regulatory segment is displaced from the substrate-binding pocket, resulting in an active enzyme. Phosphorylation of the regulatory segment (at Thr287 in CaMKIIβ) further disrupts binding to the kinase domain, leading to an active enzyme even in the absence of calcium/calmodulin. The enzyme is thus able to store a “memory” of previous activation signals, a property thought to be essential for its role in LTP.

Calcium/calmodulin-dependent protein kinases (CaMKs) were initially discovered by Paul Greengard (Schulman and Greengard, 1978) and historically named based on the order of elution from a fractionation column loaded with brain extract (Yamauchi and Fujisawa, 1980). There are multiple CaMKs and the name is often misleading, as not all CaMKs are regulated by calcium/calmodulin ($Ca^{2+}$/CaM). Such an example is CaMKIII, now renamed to eukaryotic elongation factor 2 (eEF2) kinase and classified

as an atypical kinase. Similarly, not all kinases regulated by $Ca^{2+}$/CaM are officially classified as CaMKs. Of the originally reported CaMKs, CaMKI and CaMKIV are closely related, while CaMKII is evolutionarily closer to the phosphorylase kinase (PhK). CaMKs can be further divided into two families, CaMK1 and CaMK2. The CaMK1 family contains 12 members, including the multifunctional kinases CaMKI and CaMKIV, which are also involved in neuronal plasticity and gene regulation. The CaMK2 family contains the four CaMKII genes (*Camk2α, β, γ, δ*), two PhK variants targeted at specific substrates and the pseudokinase CASK, which lacks enzymatic activity (Bayer and Schulman, 2019).

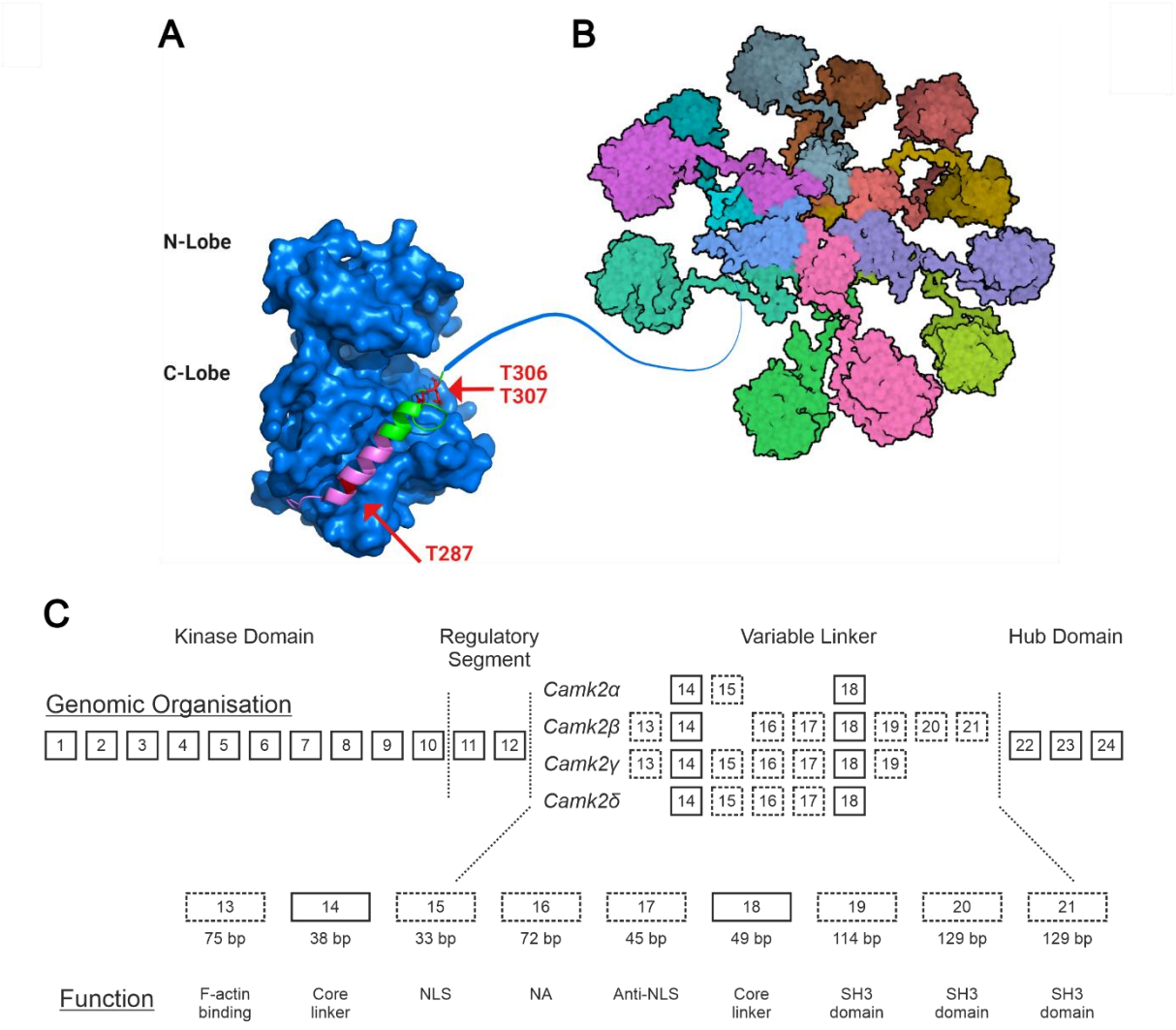



Figure 2: CaMKII structure and genomic organization.

(A) Molecular structure of the CaMKII monomer displayed as an interaction surface (PDB: 2VN9). The regulatory segment is shown as a helix. Magenta indicates the autoinhibitory segment, green indicates the calmodulin binding motif. The activating autophosphorylation (T287) and the two inhibitory autophosphorylations (T306/T307) are indicated in red. (B) Oligomeric assembly of CaMKII into a dodecameric holoenzyme. (PDB: 2UGY). (C) Genomic organisation of the four CaMKII genes. Boxes represent individual exons, solid lines indicating constitutive exons, dashed lines indicating

alternatively spliced exons. For the variable linker domain, gene names indicate which exons are present in that particular gene. For all other domains, the intron/exon structure does not differ between the genes. The association with CaMKII protein domains is displayed above the intron/exon structure. The exons of the variable linker segment are shown enlarged underneath. Published functions of individual exons are indicated (Sloutsky et al., 2020; Tombes et al., 2003).

## Genomic organization and evolutionary conservation of CaMKII genes

Whereas simple organisms, such as *Caenorhabditis elegans* or *Drosophila melanogaster,* harbour a single ancestral CaMKII gene, duplication resulted in four genes in mammals, termed α, β, γ and δ (Tombes *et al.*, 2003) (Figure 2C). These genes and their various splicing isoforms are expressed in a tissue-specific manner, with α and β being the predominant isoforms in neuronal cells. Together, they are estimated to contribute up to 2% of the total protein mass in the hippocampus and 1% of total brain protein mass in neurons (Erondu and Kennedy, 1985). Notably, conservation in CaMKII dates back to the evolutionary stage when the first synapse is thought to have formed (Ryan and Grant, 2009), and all essential features of the kinase are well conserved between metazoans. Alternative splicing of the four genes leads to the expression of over 70 distinct isoforms in mammals (Sloutsky *et al.*, 2020; Tombes *et al.*, 2003). Other species have acquired additional gene copies, such as the teleost fish *Danio rerio,* in which *Camk2β*, *γ* and *δ* have been duplicated. Interestingly, this redundancy is mirrored by the corresponding activation factor, calmodulin. Three gene copies exist in mammals, and six in zebrafish, which all encode the exact same protein (Rothschild et al., 2020). While many CaMKII variants, including individual splice isoforms, are known to have specific functions, this redundancy could also be a safety mechanism. A disruptive mutation in one of the genes can under certain conditions be rescued by another CaMKII gene. The genetic conservation of the paralogs is therefore at least partially reflective of their ability to compensate for each other's absence (Elgersma et al., 2004; Rothschild *et al.*, 2020).

Genetic variation in CaMKII has mostly been restricted to the variable linker segment and almost all mammalian splice variants are derived from the alternative splicing of one of the nine exons encoding this variable segment. A notable exception is the catalytically inactive α-kinase anchoring protein (αKAP), which is transcribed from an internal promotor within the *Camk2α* gene. It contains a transmembrane domain and targets the CaMKII holoenzymes to the sarcoplasmic reticulum in cardiac and skeletal muscle cells (Bayer et al., 1998; Bayer et al., 1996).

The first ten exons of CaMKII encode for the kinase domain, including the highly conserved ribose anchor, ATP binding pocket and activation loop. They are highly similar between the four CaMKII genes and have been conserved during evolution. The kinase domain is followed by the regulatory domain, which features an

autoinhibitory segment, the $Ca^{2+}$/CaM binding site, and several functionally relevant posttranslational modification sites. Of these, the activating autophosphorylation site T287 (based on the CaMKIIβ amino acid numeration, shifted to T286 in CaMKIIα), and the inhibiting sites T306/307 stand out. This regulatory domain is encoded by exons 11 and 12 and similarly conserved between genes and species.

The adjacent flexible linker domain consists of nine exons (13-21), of which individual CaMKII genes contain between three (*Camk2α*) and eight (*Camk2β*) (Sloutsky *et al.*, 2020; Tombes *et al.*, 2003). These exons were initially numbered I-X (Tombes and Krystal, 1997), but have more recently been given a number based on the position within the entire spectrum of CaMKII genes (Sloutsky *et al.*, 2020). Exon 13 is specific to the *Camk2β* and *γ* genes and modulates the affinity for calmodulin (Brocke et al., 1999). It is known to strongly increase the interaction with filamentous actin (F-actin) (Khan et al., 2019; O'Leary et al., 2006) which is modulated by several important autophosphorylation sites. Exon 14 is considered constitutive and functions as the first part of a “core linker”. It has, however, recently been reported to be excluded in the so-called zero-linker isoform of *Camk2α* (Sloutsky *et al.*, 2020). Exon 15 is not present in *Camk2β* and contains a highly conserved 11-residue nuclear targeting sequence (Srinivasan et al., 1994). Exon 16 is specific to *Camk2β, γ* and *δ* and the least conserved among the three genes. In *Camk2γ*, it additionally includes an alternative donor site enabling the first part of the exon to be included independently. No definitive function of this exon is known. Exon 17 is similarly absent in *Camk2α* and theoretically counters the nuclear targeting sequence of exon 15 (Takeuchi et al., 2000). It is however also present in *Camk2β*, which lacks this localization sequence. Exon 18 is highly conserved and forms the second part of the “core linker”. It is therefore constitutively included in all isoforms except the zero-linker variant of *Camk2α*. This “core linker” supposedly tethers the different parts of the variable linker domain together and also provides a minimal level of flexibility for the kinase domain, linked to the central hub. Exons 19-21 are exclusive to *Camk2β* and under strict tissue-specific control, with the highest expression in endocrine tissue (Breen and Ashcroft, 1997; Urquidi and Ashcroft, 1995). They all contain SH3 binding sites and have presumably been generated by exon duplication. A homolog of exon 19 is also present in *Camk2γ*, but additionally harbours an internal acceptor site and has been classified as an unrelated exon in some publications (Tombes *et al.*, 2003). Following the variable linker domain is the association or hub domain, encoded by three exons. Three hydrophilic

domains within these exons are thought to be responsible for oligomerization of the enzyme (Kolb et al., 1998). These exons have been conserved during evolution and are largely invariant between CaMKII genes, enabling the formation of hetero-oligomers.

Of the two CaMKII genes predominantly expressed in neurons, *Camk2α* has three reported alternative splicing isoforms. Two are derived from the exclusion or inclusion of exon 15, while the zero-linker variants lacks the entire flexible linker domain (Sloutsky *et al.*, 2020). There are eleven known *Camk2β* isoforms generated by alternative splicing, of which up to eight have been detected in a single tissue (Sloutsky *et al.*, 2020; Tombes *et al.*, 2003). The historically termed full-length (FL) isoform, which actually lacks exons 19 to 21, is expressed throughout the brain but enriched in cerebellum (Beaman-Hall et al., 1992; Miller and Kennedy, 1985). The longest possible isoform, also containing exons 19 to 21, is termed $\beta_M$ and expressed in endocrine and muscle tissue. Another isoform, termed $\beta_2$, contains exons 20 and 21 and shows a similar tissue specific expression pattern. Both isoforms presumably target the holoenzyme to the plasma membrane (Breen and Ashcroft, 1997; Urquidi and Ashcroft, 1995). A developmental regulation has been reported for exon 13 in rodents, and isoforms lacking this exon are predominantly expressed in the embryonic mouse brain. This splicing pattern changes towards the FL isoform within two weeks after birth (Brocke *et al.*, 1999; Brocke et al., 1995). Exon 17 alternative splicing has been reported, but no strict tissue- or developmental specific regulation of this exon is known. Strikingly, whereas most isoforms of *Camk2β* have been identified in multiple species, including humans and various rodents, exon 16 alternative splicing has been exclusively reported for human tissues (Rochlitz et al., 2000; Sloutsky *et al.*, 2020). However, as of yet, no conclusions regarding a species-specific regulation of this exon have been made.

## Structure and oligomeric assembly of CaMKII

Each CaMKII monomer consists of a highly conserved N-terminal kinase domain that displays the typical conserved protein kinase core (Figure 2A). A small N-terminal lobe featuring β-sheets is connected via a flexible linker to a larger C-terminal lobe that is predominantly α-helical. The activation loop is part of the C-terminal lobe, while the ATP binding pocket is located in the cleft between the two regions. While most kinases, including CaMKI and CaMKIV, require phosphorylation and subsequent structural rearrangement of their activation loop in order to adopt an active state, CaMKII does not contain a phosphorylation site in this segment. Instead, CaMKII is in a theoretically constitutively active conformation, but can be inhibited by a regulatory segment. In the inactive state, an autoinhibitory segment forms a helix that binds tightly to the substrate binding site, preventing access of substrates and also mildly affecting nucleotide binding (Chao *et al.*, 2011; Rellos *et al.*, 2010; Rosenberg et al., 2005). The $Ca^{2+}$/CaM binding motif partially overlaps with the autoinhibitory segment and binding of the activator calmodulin thus relieves the autoinhibitory block. Upon activation, autophosphorylation of Thr287 (Thr286 in CaMKIIα), which is located between two kinase domains of the same holoenzyme, further disrupts the binding of the regulatory segment to the kinase domain (Chao *et al.*, 2011; Hanson et al., 1994). This leads to an active enzyme, even in the absence of an external activation signal. Phosphorylation of this residue drastically increases the affinity for calmodulin, an effect known as CaM-trapping (Meyer *et al.*, 1992). The calcium/calmodulin-binding motif features two autophosphorylation sites, Thr306 and Thr307. Phosphorylation prevents the binding of $Ca^{2+}$/CaM but does not impact the autonomy generated by a prior phosphorylation of Thr287 (Chao *et al.*, 2011; Rellos *et al.*, 2010). The main effect of the “inhibiting” phosphorylation is therefore not to inhibition the enzyme *per se*, but to prevent it from reaching its maximal activity (Bayer and Schulman, 2019). Sequential phosphorylation of these three residues has a great impact on the activity of the enzyme (Hashimoto et al., 1987). The triple-phosphorylated state, T287/306/307, is therefore not inactive, but capped at the level of autonomous activity.

The regulatory segment is followed by the variable linker domain. The linker composition differs from gene to gene and is heavily affected by alternative splicing. Most isoforms contain a “core linker” which allows for a basic level of flexibility between the central hub and the kinase domains, protruding into the periphery. The actual length of the linker differs considerably, from 30 residues in the shortest CaMKIIα

isoform (except for the zero-linker variant), to 218 residues in the longest isoform of CaMKIIβ. The linker can serve multiple functions, depending on the exact exon composition. In all isoforms, it connects the kinase and regulatory domains to the association or hub domain.

Multiple CaMKII subunits associate via mostly hydrophobic interactions to form a holoenzyme consisting of two rings (Figure 2B). CaMKIIα forms predominantly dodecamers, consisting of two hexameric rings (Chao *et al.*, 2011). However, a considerable subpopulation also forms tetradecamers in solution (Sloutsky *et al.*, 2020). A slightly shifted distribution has been reported for CaMKIIβ using negative-stain electron microscopy, which also shows a small population of hexadecamers (Buonarati et al., 2021). Different oligomeric states have been reported for individual hub assemblies and structural homologs found in bacteria and algae can adopt 16 to 20-subunit assemblies (McSpadden et al., 2019). These are supported by extensive hydrogen-bonding networks, which are absent in mammalian CaMKII. As the hub domain is conserved among CaMKII genes, subunits from different genes and isoforms can interact and form heterooligomeric assemblies. This is thought to enable fine-tuning of the holoenzyme characteristics to the exact needs of the different cellular contexts. The oligomeric assembly has also been proposed to enable subunit exchange between holoenzymes (Bhattacharyya et al., 2016; Rosenberg et al., 2006; Stratton et al., 2014), potentially enabling the repair of damaged parts or adapting the holoenzyme stoichiometry in response to changing cellular conditions. It also has been proposed to enable mixing of active and inactive holoenzymes in the aftermath of an activation signal, to facilitate the spread of CaMKII activity. However, as phosphorylation and thereby activation of additional subunits still requires binding of calmodulin to make the activating phosphorylation site accessible, its influence remains to be established. The oligomeric assembly can also increase the apparent affinity for binding partners through avidity effects. It further enables multivalent interactions and crosslinking of cellular components via simultaneous binding of multiple interaction partners (Colbran and Brown, 2004; Coultrap and Bayer, 2012; Hell, 2014). A prominent example is the ability of CaMKIIβ to bundle actin filaments (Okamoto et al., 2007).

## Influence of the variable linker domain of CaMKII

The frequency-dependent response to calcium/calmodulin ($Ca^{2+}$/CaM) spikes does not require a multimeric holoenzyme. However, the oligomeric assembly of the kinase is highly conserved and can already be found in ancient versions of the protein, such as in *C. elegans* (Reiner et al., 1999). Although the frequency-dependent response to $Ca^{2+}$/CaM is a general feature of the enzyme, it has been observed that the length of the variable linker region, as determined by alternative splicing, alters the exact extent of the autonomous activity that is generated at a particular frequency (Bayer et al., 2002). Whereas the regulatory segment is sandwiched between the kinase and the hub domain, the linker segment is located at the surface of the holoenzyme. This enables the incorporation of additional residues present in alternative splicing isoforms with longer linker segments without disturbing the overall architecture of the holoenzyme. Based on a crystal structure of CaMKIIα (Chao *et al.*, 2011), an early model suggested that the inactive holoenzyme exists in at least two distinct conformations, an extended, activatable, and a compact, non-activatable state. In this model, the length of the linker determines the equilibrium between the two conformations, modulating the auto-inhibition of the enzyme to set a tuneable threshold for the response to calcium spikes. However, different studies (Buonarati *et al.*, 2021; Myers et al., 2017) that analysed the structural heterogeneity of the inactive CaMKIIα or β holoenzymes by negative stain electron microscopy found this compact conformation to be only sparsely populated (<3%). Instead, the studies reported the presence of kinase domain dimers, consistent with a previous crystal structure of isolated kinase domain dimers from *C. elegans* (Rosenberg *et al.*, 2005). In this latter structure, the kinase domains dimerise via a coiled-coil interaction of their regulatory segment, potentially preventing activation by calcium/calmodulin in the context of the holoenzyme. The exact mechanism of how the linker length affects the kinetics of the enzyme therefore remains enigmatic. Further complicating, not all alternative splicing events seem to change the calmodulin-dependency of the enzyme (Sloutsky *et al.*, 2020). Instead, the effect seems to at least partially depend on the identity of the kinase and hub domains and therefore differ between CaMKII genes, despite their high level of conservation.

A recent study reported an influence of the variable linker on the acquisition of the activating and inhibiting autophosphorylation (Bhattacharyya et al., 2020). The difference was found between isoforms with longer linkers (full-length CaMKIIβ and γ)

and isoforms with shorter linkers (CaMKIIα and δ). Whereas CaMKIIα shows robust levels of the activating autophosphorylation, it shows only low levels of the inhibiting autophosphorylation, which are rapidly lost due to phosphatase activity. CaMKIIβ displays higher levels of the inhibitory autophosphorylation and diminished levels of the activating phosphorylation. Interestingly, this study used the actual full-length isoform of CaMKIIβ, containing exons 19 to 21. CaMKIIγ similarly contains a homolog of exon 19, raising the possibility that these exons might be relevant for the observed differences in autophosphorylation pattern. A shorter-linker isoform of CaMKIIβ, which displayed autophosphorylation pattern similar to CaMKIIα, lacks these additional exons. As the isoforms containing exons 19 to 21 are presumably not expressed in neurons (Tombes *et al.*, 2003), a functional relevance of this mechanism for modulating synaptic plasticity remains to be established.

An interesting parallel can be drawn between CaMKII and the AKAP-PKA system. The A-kinase anchoring protein (AKAP) binds to the protein kinase A (PKA) and tethers the enzyme in proximity to substrates. An intrinsically disordered flexible linker within PKA imparts the molecular plasticity of the complex and the linker length determines the spatial range over which PKA can phosphorylate its substrates (Smith et al., 2013). Interestingly, the linker length of CaMKII has been shown to affect substrate specificity in flies *in vitro* (GuptaRoy et al., 2000) and a much stronger effect could be envisioned in an intact cellular environment.

## Generating $Ca^{2+}$-independent autonomous activity in CaMKII

Autonomous activity of CaMKII has been proposed as a form of molecular memory and is primarily mediated by autophosphorylation. However, other means of generating autonomy exist and include binding to specific proteins and various other posttranslational modifications (Bayer and Schulman, 2019). All stimulating modifications require an initial $Ca^{2+}$/CaM-stimulus to displace the regulatory segment and make many of the targeted sites accessible. Most of these modifications act by preventing the autoinhibitory segment from re-associating with the kinase domain and thus re-inhibiting the enzyme. Autonomous activity is significantly weaker compared to $Ca^{2+}$/CaM-stimulated activity. An exception are so called high-autonomous targets, which are specifically inhibited at high calcium levels by binding of calmodulin to the CaMKII target sequence (Woolfrey et al., 2018) and possibly by additional, as of yet unknown, mechanisms (Coultrap and Bayer, 2014; Coultrap et al., 2014). The activating autophosphorylation T287 (nomenclature based on the CaMKIIβ amino acid numeration) generates the highest level of autonomy, reaching about 20% of the $Ca^{2+}$/CaM-stimulated activity. This autonomy is, however, not suited for long-term information storage, as it is relatively quickly reversed by phosphatases (Bayer and Schulman, 2019; Bhattacharyya *et al.*, 2020). It primarily seems to act as a frequency decoder and produces graded CaMKII activity based on the frequency and number of stimulating signals (Bayer *et al.*, 2002; Chang et al., 2017; De Koninck and Schulman, 1998). A longer-lasting autonomy is generated by CaMKII binding to the NMDAR subunit GluN2B (Bayer et al., 2001; Bayer et al., 2006). This can be induced by $Ca^{2+}$/CaM-stimulation or T287 autophosphorylation alone. The stimulus is no longer required upon binding and the enzyme becomes more resistant to phosphatases.

Other modifications such as nitrosylation, oxidation, and GlcNAcylation, were initially identified in disease states. Nitrosylation generates autonomy and has been reported for CaMKIIα, β and δ (Coultrap and Bayer, 2014; Coultrap *et al.*, 2014; Erickson et al., 2015). The mechanism differs slightly between CaMKII homologues. While CaMKIIα requires modification of both C280 and C289, CaMKIIβ and δ require only modification of C290 and lack residue C281. Instead, an inhibiting nitrosylation of C273 has been reported for CaMKIIδ (Erickson *et al.*, 2015). A direct effect of GlcNAcylation has not been shown, but generation of an autonomous state has been suggested due to structural similarities to other activated states (Erickson et al., 2013). Oxidation can induce autoactivity in all CaMKII isoforms and, depending on the gene, might

additionally be modulated by nitrosylation of neighbouring residues (Coultrap and Bayer, 2014; Coultrap *et al.*, 2014; Erickson et al., 2008).

## Canonical function of CaMKII in synaptic plasticity

CaMKII is a multifunctional kinase and involved in numerous pathways, depending on the exact cellular environment. CaMKIIα and β fulfil their canonical functions in neurons, where they mediate key steps of synaptic plasticity and are essential for long-term potentiation (LTP) (Hell, 2014; Malenka and Bear, 2004).

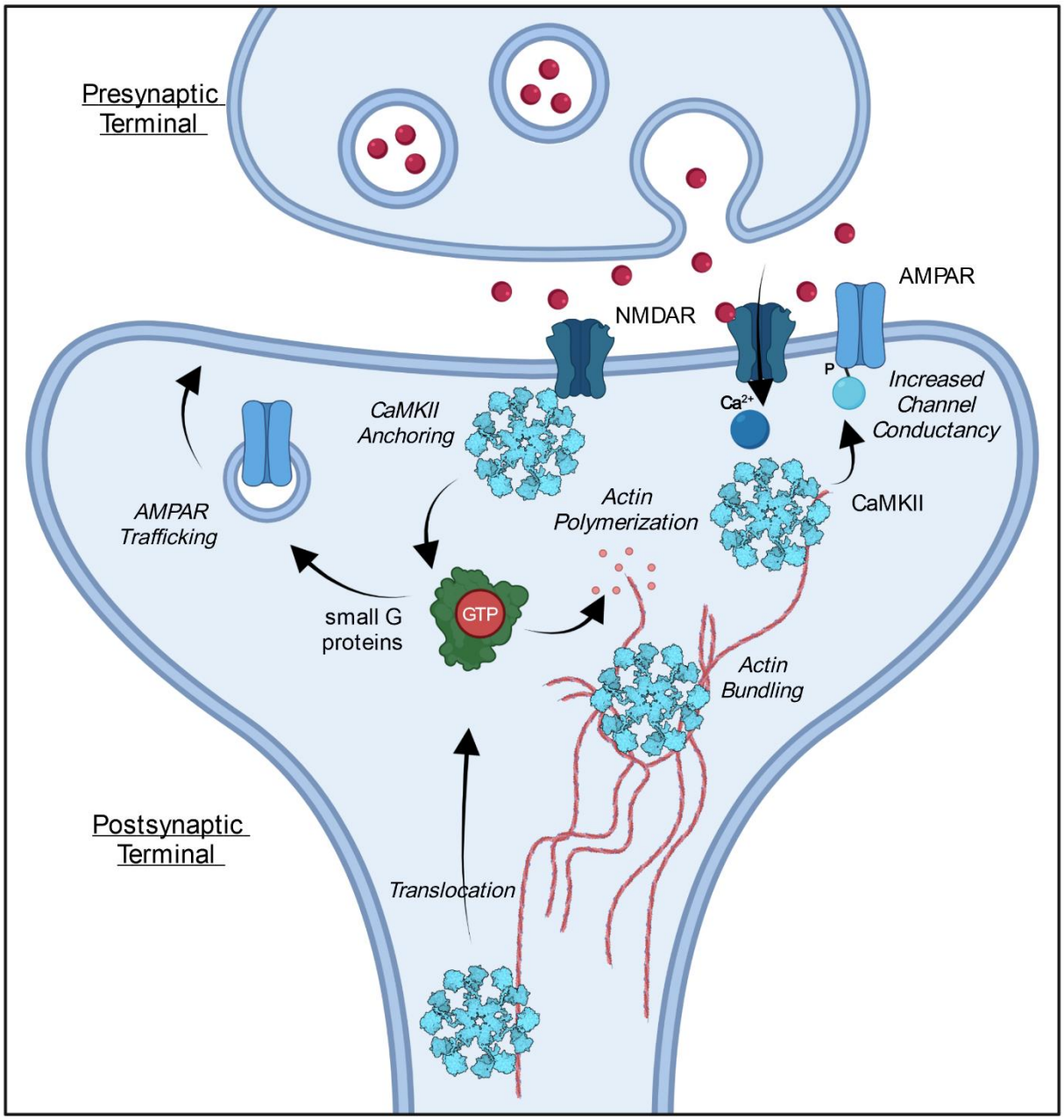



Figure 3: Mechanism of CaMKII-mediated long term potentiation.

Schematic representation of a chemical synapse and the involvement of CaMKII in generating synaptic plasticity. Calcium influx into the postsynapse due to synaptic activity activates CaMKII via calmodulin. Active CaMKII translocates from the synaptic spine to the postsynaptic density (PSD). CaMKII can directly phosphorylate AMPA receptors to increase their single channel conductance. CaMKII also regulates small GTPases such as Ras and Rac which affects actin polymerization and AMPA receptor trafficking. Activated CaMKII can bind to the NMDAR subunit GluN2B, anchoring

CaMKII near the site of calcium influx. Once dephosphorylated, CaMKII can bind to F-actin and various actin-associated proteins and thereby bundle actin filaments.

A synapse is an asymmetric cellular junction that allows a neuron to pass electrical or chemical signals to another neuron or other type of target cell (Südhof, 2018) (Figure 3). In humans, most synapses form during pre- and postnatal development and half of all synapses are subsequently pruned, predominantly in the first two decades of life. In a chemical synapse, a presynaptic terminal releases neurotransmitter molecules, using a canonical, vesicle trafficking-based machinery that ejects the neurotransmitters into the synaptic cleft. A specialized structure on the postsynaptic neuron, called the postsynaptic density (PSD), senses the signals through diverse receptors and triggers electrical and chemical signalling pathways in the signal-receiving cell. The nature of the presynaptic specialization - a cluster of synaptic vesicles, some of which are docked at a thickened part of the presynaptic plasma membrane - imposes strong temporal and spatial restrictions on neurotransmitter release. Synapses differ vastly in their properties, such as the type of neurotransmitter they release, their receptor composition, and the short-term and long-term changes induced by their activity. While presynaptic terminals are uniquely neuronal, postsynaptic specializations are predominantly neuronal in the brain, forming mostly on protoplasmic projections called dendrites. Outside of the central nervous system, they can form in virtually any cell type.

The probably best-characterized synapses are glutamatergic synapses, which are excitatory, mediate fast synaptic transmissions in the central nervous system using glutamate as a neurotransmitter, and feature a prominent postsynaptic density (Sheng and Hoogenraad, 2007). In excitatory synapses, the PSD is housed on small actin-rich protrusions called dendritic spines, whereas inhibitory synapses are located on dendritic shafts. Spines are small structures of 0.5-2 µm and usually feature a bulbous head that is connected to the dendrite via a thin spine neck. Spines can be classified as “mushroom”, “thin”, or “stubby” and their structure correlates with synaptic strength. The thin spine neck restricts diffusion and spines thus serve as microcompartments, enabling differential regulation of neighbouring synapses. The PSD typically lies at the distal top of the spine head. It is a very large membrane-associated protein complex, specialized for synaptic signalling and plasticity. The PSD can differ considerably in size, ranging from 200-800 nm wide and 30-50 nm thick (Carlin et al., 1980; Sheng

and Hoogenraad, 2007) and an average PSD of ~360 nm length has been estimated to have a molecular weight of over 1 gigadalton (Chen et al., 2005). The abundance of individual proteins within the PSD differs considerably, and is specific to the synapse type. Strikingly, CaMKIIα (28 pmol/20 µg PSD protein) and CaMKIIβ (4.7 pmol/20 µg) are the most abundant proteins in this compartment (Cheng et al., 2006). Together, they make up around 9% of the total mass of the PSD and are significantly more abundant than other core components, like SynGAP (2.1 pmol/20 µg), a Ras GTPase activating protein (RasGAP). The most abundant exclusively structural protein is PSD-95 (2.3 pmol/20 µg), suggesting that CaMKII and SynGAP also fulfil structural roles. The large NMDA-receptor (NMDAR) protein complex acts as an interaction and organization hub and is regarded as a core substructure of the PSD, even though it is considerably less abundant than PSD-95 or CaMKII (0.24 pmol/ 20 µg for the NR1 subunit). NMDARs are anchored and directly bound by PSD-95 and interact with CaMKII via the cytoplasmic tail of their GluN2B subunit (Bayer *et al.*, 2001). The concentration of NMDARs is relatively uniform and about ~20 receptors are estimated to be present per average PSD (Sheng and Hoogenraad, 2007). Another class of neurotransmitter receptors, AMPA-receptors (AMPAR) have an average concentration of ~15 receptors per PSD. These receptors show a much larger variability between individual synapses and their concentration can change in response to synaptic stimulation (Araki et al., 2015; Herring and Nicoll, 2016). CaMKII, which also shows a high level of variability in term of abundance, is concentrated at the cytoplasmic face of the PSD. In electron microscopy images of intact PSDs, CaMKII forms towers that protrude into the spine (Petersen et al., 2003). So far, these have not been structurally resolved.

During synaptic transmission, ligand-gated ion channels, such as AMPARs, open and allow the influx of sodium ions, potentially leading to the depolarization of the postsynaptic membrane. NMDARs initially stay closed, as they are blocked by magnesium ions and require prior depolarization of the plasma membrane. They therefore act as coincidence detectors, requiring both a change in membrane potential, as well as binding of the neurotransmitter, in order to transduce signals. NMDARs allow the influx of calcium ions and thus the activation of calcium-dependent kinases, such as CaMKII. The enzyme can directly phosphorylate AMPARs to increase their single-channel conductance (Derkach et al., 1999; Kristensen et al., 2011). CaMKII also regulates various proteins involved in AMPAR trafficking, such as small GTPases

(Araki *et al.*, 2015; Herring and Nicoll, 2016; Zhu et al., 2002), and AMPAR-associated proteins (TARPs) (Opazo et al., 2010; Park et al., 2016). This leads to an increase of the concentration of AMAPRs at the postsynaptic membrane (Herring and Nicoll, 2016). CaMKII binds to the NMDAR subunit GluN2B, an interaction that initially depends on stimulation by $Ca^{2+}$/CaM, but once established is independent and more resistant than CaMKII autophosphorylation. It ensures recruitment of CaMKII to strategically ideal locations near the source of calcium influx (Leonard et al., 1999), in addition to generating autoactivity itself (Bayer *et al.*, 2001). This interaction is important for maintaining LTP, whereas the activating autophosphorylation is initially required, but largely reversed shortly after the stimulus (Chang *et al.*, 2017; Rossbach et al., 2009). In addition to its role in synapses, CaMKII has recently been found to directly phosphorylate the voltage-gated sodium channel Nav1.6. Mainly expressed at the axon initial segment, this channel is critical for the initiation and propagation of action potentials (Zybura et al., 2020).

## CaMKII is involved in structural rearrangement of the synapse

CaMKII is a highly unusual kinase, as it is far more abundant than most of its substrates. Even though CaMKII is a multifunctional kinase and targets many different proteins, it is by far the most abundant component of the PSD, even exceeding structural proteins (Cheng *et al.*, 2006). Exceeding other proteins by an order of magnitude, the principles of normal signalling cascades do not hold true for this enzyme. In related kinases, nucleotide exchange seems to be the rate-limiting step (1 event per ~50 ms), compared to the relatively fast phosphorylation reaction (1 event per ~2 ms) (Grant and Adams, 1996; Zhou and Adams, 1997). Thus, phosphorylation by CaMKII might happen on a timescale similar to electrophysical responses (Bayer and Schulman, 2019). It also means that most CaMKII molecules will only phosphorylate a single substrate during an individual stimulus, thus deviating from the principle of signal amplification.

CaMKII forms large multimeric assemblies, which can act as interaction hubs for many proteins (Hell, 2014) and thereby additionally fulfil structural roles. It directly interacts with the F-actin cytoskeleton, mainly via CaMKIIβ (Bayer *et al.*, 2006; Fink et al., 2003; Shen et al., 1998), promoting spine and synapse formation via F-actin bundling (Fink *et al.*, 2003; Sugawara et al., 2017). CaMKII adopts different actin-binding modes, most easily rationalized by multiple subunit contacts between the CaMKII holoenzyme and the F-actin cytoskeleton (Khan *et al.*, 2019). The interaction differs strongly between CaMKII genes and is affected by alternative splicing of exon 13, but not completely lost in the CaMKIIα variant lacking this exon (Bayer *et al.*, 2006; Hoffman et al., 2013). Thus, multiple actin binding sites may exist on CaMKII. The enzyme also binds to α-actinin via its calmodulin-binding motif (Jalan-Sakrikar et al., 2012). CaMKII generally stabilizes F-actin bundles and augments spine size in a kinase-independent manner (Lin and Redmond, 2008; Okamoto *et al.*, 2007). The influence of CaMKIIβ in specific forms of LTP was further found to only depend on its actin-binding properties, rather than its catalytic activity (Borgesius et al., 2011). Interaction with actin results in CaMKII being mostly concentrated in the interior of the spine. Synaptic activity leads to the dissociation of CaMKII from the actin cytoskeleton and its translocation to the PSD (Shen and Meyer, 1999; Shen et al., 2000). This is mediated by the autophosphorylation of specific residues (Lin and Redmond, 2008), located in and around exon 13. F-actin binding therefore requires CaMKIIβ to be in a non-phosphorylated state (Zalcman et al., 2018). As actin is the main structural component

of dendritic spines, CaMKII has a strong influence on the structural aspect of LTP (sLTP). Synaptic potentiation induces an increase in spine size, while long-term depression (LTD) results in the shrinking of spines (Bosch and Hayashi, 2012). Lack of CaMKIIβ specifically results in longer dendrites (Puram et al., 2011). CaMKII has also been shown to bind to the Rac-GEF Tiam1 in stimulated spines, generating CaMKII autoactivity. Phosphorylation of Tiam1 maintains Rac activity, which promotes stable actin polymerization and is required for LTP (Saneyoshi et al., 2019). In the presynapse, CaMKIIβ regulates axon growth by affecting F-actin length (Xi et al., 2019). While CaMKIIα predominantly regulates synaptic strength by regulating neurotransmitter receptor numbers and physical properties, CaMKIIβ regulates dendritic morphology, neurite extension, synapse number and spine size (Fink *et al.*, 2003). This does not exclude a structural role for CaMKIIα, which is more abundant in many synapses and known to increase in concentration, both in response to activity-dependent translocation, as well as activity-dependent *de novo* translation (D'Amours et al., 2011; Mayford et al., 1996; Mori et al., 2000).

## Mouse models of CaMKII reveal role in neurological disease states

As a multifunctional kinase, CaMKII is involved in numerous steps of generating synaptic plasticity. Although it has been proven or suspected to be involved in various neurological diseases, the exact mechanisms are often unclear (Robison, 2014; Sałaciak et al., 2021). The impact of a misregulation of CaMKII has been tested in various mouse models, many of them showing distinct learning-deficient phenotypes (Elgersma *et al.*, 2004; Zalcman *et al.*, 2018). The majority of studies has focussed on CaMKIIα, as it is the most abundant isoform in the mammalian brain. Homozygous and heterozygous *Camk2α* knockout (KO) mice display learning impairments in several hippocampus-dependent tasks and diverse behavioural phenotypes, including altered locomotor activity and anxiety-related behaviour (Silva et al., 1992; Yamasaki et al., 2008). Several knockin (KI) mouse strains modulate the activity of CaMKIIα. Disabling of the activating autophosphorylation (T286A) impairs NMDAR-dependent LTP and results in various learning and memory deficits (Giese et al., 1998; Irvine et al., 2005). Introduction of a kinase-dead mutant (K42R) inhibits both functional as well as structural LTP and mice display severe impairments in hippocampus-based memory tasks. Interestingly, the activity-dependent translocation of CaMKII remained unaffected in these animals (Yamagata et al., 2009). A KO model of *Camk2β* displays altered anxiety-related behaviour, cognitive impairments and diverse motor deficits, like ataxia and impaired grip strength (Bachstetter et al., 2014). Knockin overexpression of *Camk2β* diminished, but not completely impaired LTP. This manipulation also had a more subtle effect on learning and memory, with animals only showing deficits for long-term memory tasks (Cho et al., 2007). The impact on neuronal plasticity thus depends on the exact manipulation of CaMKII. KO severely impairs most memory tasks, whereas overexpression only affects memory consolidation over longer periods. A KO model of *Camk2δ* has recently been reported, but not tested for learning and memory impairment (Mohamed et al., 2019). CaMKIIδ is the main isoform in the axons of developing neurons in rodents (Omar Faison et al., 2002) and functions in learning and memory have been suggested (Zalcman *et al.*, 2018). CaMKIIγ has been proposed as a cytonuclear signalling protein, shuttling $Ca^{2+}$/calmodulin to the nucleus to induce gene expression. A KO model for *Camk2γ* shows impaired activity-dependent expression of core immediate-early genes (BDNF, c-Fos, Arc) and displays spatial learning impairments. Interestingly, a point mutation that impairs calmodulin

shuttling by CaMKIIγ is associated with intellectual disability in humans (Cohen et al., 2018).

Together, these models suggest a relevance of neuronal CaMKII for maintaining cognitive abilities. Not surprisingly, mutations in *Camk2α* and *β* (Küry et al., 2017) or *Camk2γ* (de Ligt et al., 2012; Quervain and Papassotiropoulos, 2006) have been linked to intellectual disability in humans. CaMKII disfunction might furthermore underlie numerous neuropsychiatric illnesses, including addiction, schizophrenia, epilepsy and depression, as well as several neurodevelopmental and neurodegenerative disorders (Robison, 2014; Sałaciak *et al.*, 2021). Many of these processes, such as addiction and depression, are thought of as maladaptations of learning and reward circuitries. Expression levels of CaMKII are often altered in depression (Li et al., 2013; Robison, 2014), but the exact effect depends on the model system, treatment, and induced CaMKII modifications (Sałaciak *et al.*, 2021). In general, an increase in CaMKIIβ and, conversely, a decrease in CaMKIIα are linked with depression-like phenotypes in animals. CaMKII is also involved in neurodegenerative diseases like Alzheimer's disease (Ghosh and Giese, 2015), Parkinson's disease (Picconi et al., 2004) and Angelman syndrome (van Woerden et al., 2007). Together, the involvement of CaMKII in numerous neurological disease states seems obvious, yet exact mechanism are often missing. As CaMKII is a multifunctional kinase and frequently acts upstream of signalling and remodelling processes fundamental to synaptic plasticity, the precise cause or impact of CaMKII alterations are difficult to decipher.

# Aims

How species identity is determined by cellular gene expression programs is a fundamental question in biology. Although it touches upon our very existence, it presently lacks a convincing answer. This project has tested the intriguing hypothesis that species-specific alternative splicing gives rise to species-specific protein isoforms that mediate species-specific functions. This thesis was based on initial findings obtained during my Master's thesis in the laboratory of Prof. Florian Heyd, entitled "Species-Specific Alternative Splicing" (2017). In this, I was able to show the species-specific alternative splicing of *Camk2β* exon 16 and was able to identify the regulatory *cis*-acting element, located in the upstream intron. The aim of this thesis was to fully understand the exact mechanism of the observed species-specific differences in *Camk2β* alternative splicing and elucidate their functional consequences on a transcriptomic, proteomic, cellular and organismal level. The following specific aims were investigated:

1. Revealing the governing principles behind the species-specific splicing event. Decipher the exact regulatory mechanism, including the precise location and nature of the *cis*-acting element. If applicable, finding the corresponding *trans*-acting factor. Having established the mechanism, correlating it with genomic differences between species to elucidate the evolution of this splicing event.

2. Identify kinetic differences between various CaMKIIβ isoforms. Test if these variants differ regarding their cooperativity, activity or calmodulin-dependency. Evaluate if the length, or the composition of the linker is associated with kinetic differences.

3. Identify differences in the substrate spectra between various CaMKIIβ isoforms. Test if there are isoform- or group-specific phosphorylation targets. Analyse if these correlate with the composition or length of the variable linker domain, or the subcellular localization of the enzyme. Evaluate if the substrate spectra indicate functional differences between CaMKIIβ isoforms, such as differential targeting of proteins involved in synaptic plasticity.

4. Using a mouse model with a humanized *Camk2β* splicing pattern, test the electrophysical changes in synapses with an altered ratio of CaMKIIβ isoforms. How does the presence of species-specific splice isoforms affect long-term potentiation in standard glutaminergic neurons?

5. Investigate the impact of species-specific alternative splicing of *Camk2β* in the context of an entire organism, utilizing the new mouse model. Identify behavioural phenotypes that are linked to the altered splicing pattern of *Camk2β* and connect them to global gene expression and alternative splicing changes. How does the balance of splice isoforms affect the behaviour of an organism?

# Material and Methods

## Identification of endogenous *Camk2* alternative splicing isoforms

RT-PCR was performed with total RNA from human cerebellum (Clonetech, Cat# 636535), mouse cerebellum and frog brain tissue. Human cerebellum RNA contained material pooled from three male Asians, aged 21-29 (information provided by supplier). Total cerebellum RNA from mouse (Mus musculus) and total brain tissue RNA from frog (Xenopus laevis) was extracted via Trizol (see below). Human cerebellum RNA was adjusted to 125 ng/µl, mouse and frog RNA to 500 ng/µl. Specificity for *Camk2β* was inferred by a gene-specific RT-primer, annealing to the less conserved exon 25 (numbering based on all CaMKII genes, see Figure 2C). Primers for the splice-sensitive radioactive PCR annealed to exons 9 and 23. PCR products were sequenced using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific).

## RT-PCR

2 µl RNA, 1 µl RT-primer (concentration set to 1 ng/µl), 1 µl 5x hybridisation buffer (1.5 M NaCl, 50 mM Tris pH 7.5, 10 mM EDTA) and 1 µl $H_20$ were mixed in a standard PCR-tube. The annealing mix was heated to 90°C and slowly cooled to 43°C (90-70°C: 20 sec/°C, 69-51°C: 30 sec/°C, 50-43°C: 40 sec/°C). For each RT-PCR sample, 19.75 µl of 1.25x reaction buffer (12.5 mM DTT, 12.5 mM Tris pH 8.0, 1.5 mM MgCl2, 1.25 mM dNTPs) and 0.25 µl MMuLV reverse transcriptase (NEB) were premixed and heated to 43°C, before adding to the annealing mix. For –RT controls, 20 µl of 1.25x reaction buffer without the enzyme were added. The reaction mix was incubated at 43°C for 30 min, heated to 95°C for 5 min and cooled to 4°C.

## Radioactive labelling of PCR primer

10 µl PNK buffer (70 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM DTT), 1 µl T4PNK (10 U/µl, MOLOX), 2 µl forward primer (set to 100 ng/µl), 2.5 µl $^{32}$P-γ-ATP (Hartmann Analytic, SRP-501) and 84.5 µl $H_20$ were mixed and incubated at 37°C for 45-60 min. After incubation, 23 µl 3 M NaOAc (pH 5), 1 µl glycogen, 76 µl $H_20$ and 200 µl phenol-chloroform-isoamyl alcohol-mix (Roti-Aqua P/C/I, Roth) were added. The sample was mixed and centrifuged for 3 min in a standard microcentrifuge at 10000 rpm. The upper phase was transferred into a new tube and 500 µl pre-cooled EtOH was added. The sample was mixed and centrifuged for 10 min. The pellet was washed with 70% EtOH. After washing, the pellet was air-dried for 5 min and resuspended in 80 µl $H_20$.

Labelling of the marker (pBR322 DNA-MspI Digest, NEB) was done similarly, using 2 µl marker instead of forward primer. The marker pellet was resuspended in 40 µl $H_20$ and 40 µl formamide loading dye (95% (v/v) formamide, 21 mM EDTA, 0.02 w/v bromphenol blue, 0.02 w/v xylene cyanole).

## Radioactive PCR

Each reaction contained 5 µl RT-PCR mix, 1.5 µl 10x Taq-buffer (0.5 M KCl, 0.1 M Tris pH 8.3, 15 mM $MgCl_2$, 0.01% gelatine), 1 µl unlabelled forward primer (set to 2.5 ng/µl), 1 µl reverse primer (5 ng/µl), 1 µl radioactively labelled forward primer (set to approximately 2.5 ng/µl, see description above) and 0.5 µl Taq-polymerase (self-made) in a total of 20 µl. dNTPs were already contained in the RT-PCR mix. PCR reactions were performed as following: initial heating to 95°C for 1 min, 30 sec at 95°C, 30 sec at 60-65°C, 1 min 30 sec at 72°C, final elongation for 7 min at 72°C, cooling to 4°C. Steps 1-3 were repeated for 25-32 cycles. Elongation time at 72°C was adjusted for each individual PCR reaction, according to the expected maximal product size. 20 µl formamide loading dye was added to each sample and 24 µl of the mix was transferred into a new reaction tube. The samples were heated to 95°C for 5 min and analysed on a denaturing polyacrylamide gel (7 M urea, 5% polyacrylamide in 0.5x TBE). The gel was run in 0.5x TBE (44.5 mM Tris-Borate, 1 mM EDTA). The gel was transferred onto a filter paper and dried. PCR products were detected via a photostimulable phosphor (PSP) plate.

## Minigene Design

Minigenes were designed using the pcDNA3.1+ vector backbone. The minigenes contained the following sequences: Exon 11 with 300 bp of the downstream intron, Exon 16 with 100 bp of the upstream intron, the full intron in between exon 16 and 17, exon 17 with 300 bp of the downstream intron and exon 23 with 100 bp of the upstream exon (see Figure S2A).

## Minigene Splicing Assay

Alternative splicing of the minigenes was analysed in N2A, SH-SY5Y, HEK and HeLa cells in biological triplicates. HEK and HeLa cells were cultivated in DMEM High Glucose medium (Biowest) with 10% fetal bovine serum (FBS) and penicillin/streptomycin (Biowest). N2A cells were cultivated in a 1:1 mix of Opti-MEM and DMEM medium (Opti-MEM with GlutaMAX, Gibco and DMEM with GlutaMAX,

Gibco). SH-SY5Y cells were cultivated in DMEM High Glucose medium with 10% FBS, penicillin/ streptomycin and additional L-glutamine (1% v/v of 200 mM). Cells were seeded in 12-well plates with a concentration of $1*10^5$ cells/well (HEK, SH-SY5Y, N2A) or $1.5*10^5$ cells/well (HeLa). After 24h, the cells were transfected with 1 µg plasmid and 2 µl Roti-Fect (Carl Roth GmBH) transfection reagent per well. Cells were harvested 48 hours after transfection and RNA was extracted using RNA Tri-Liquid (BioSell) reagent. All reactions, if not otherwise specified, were performed on ice. All centrifugation steps were performed in a cooled microcentrifuge with >12,000 rpm. 500 µl RNA Tri-Liquid (BioSell) was added, and resuspended cells transferred into new reaction tubes. After the addition of 100 µl chloroform, the samples were vigorously mixed for at least 10 seconds and incubated at room temperature for 10 min. The samples were centrifuged for 10 min. The upper phase was transferred into a new reaction tube and 300 µl isopropanol was added. The samples were mixed and centrifuged for 10 min. The precipitated RNA pellet was washed with 500 µl pre-cooled 70% EtOH and again centrifuged for 5 min. The RNA pellets were air dried for 5 min at RT and dissolved in 16 µl $H_20$. DNase I digestion was performed to minimize contamination with plasmid DNA. 2 µl of 10x DNase 1 buffer (100 mM Tris-HCl pH 7.5, 25 mM $MgCl_2$, 5 mM $CaCl_2$) and 2 µl of DNase I (1 U/µl, Epicentre) were added and the samples incubated for 20 min at 37°C. 180 µl $H_20$ and 200 µl phenol-chloroform-isoamyl alcohol-mix (Roti-Aqua P/C/I, Roth) were added. The samples were thoroughly mixed and centrifuged for 5 min. The upper phase was transferred into a new reaction tube and the RNA was precipitated by adding 20 µl 3 M NaOAc and 500 µl 100% EtOH. After centrifugation for 10 min, the resulting RNA pellet was washed with 500 µl pre-cooled 70% EtOH. Next, the RNA pellet was air dried for around 5 min at RT and dissolved in 12 µl $H_20$. The concentration was measured via spectroscopy and adjusted to 500 ng/µl. Alternative splicing was analysed by radioactive RT-PCR as described above, with a vector specific primer pair. Each minigene was tested in at least three independent experiments with biological triplicates.

## siRNA Knockdown

Potential *trans*-acting factors were predicted with DeepClip (Grønning et al., 2020). This deep-learning based model enables comparison of two sequences and both the human and mouse sequence of the identified functionally relevant segment (overlap

between exchange segments 4 and 8) were compared. The results were filtered as to identify RNA binding proteins that would preferentially bind to one of the two species. Additional *trans*-acting factors were predicted using RBP map (Paz et al., 2014) and the list extended to include homologous proteins and potentially non-binding control targets. Pre-trained models for the following datasets were used for the prediction: (Ray et al., 2013) for hnRNPL , QKI, PTBP1, hnRNPLL and hnRNPH2; (Galarneau and Richard, 2005) for QKI; (García-Blanco et al., 1989) for PTBP1 and (Rossbach *et al.*, 2009) for hnRNPL. HEK or N2A cells were seeded at a density of $5*10^4$ cells/well in 12-well plates. After 24 h, cells were transfected with 2 μl of a 20 μM siRNA using Roti-Fect transfection reagent. After 24 h, cells were transfected with the minigenes (see above). After 24 h, cells were harvested and RNA extracted via RNA Tri-Liquid (BioSell). Alternative splicing was analysed via radioactive RT-PCR as described above. Quantitative PCR (qPCR) was performed to validate knockdown efficiency. The housekeeping gene gapdh was used as an internal standard. The experiment was performed twice in biological triplicates.

## Prediction of branch point sequences

The SVM-BPfinder (Corvelo et al., 2010) tool was used to predict branch point sequences. If not otherwise specified, human was selected as target organism to predict branch point strength. Predicted branch points were further validated using RNABP (Nazari et al., 2019).

## Expression and purification of selected CaMKIIβ isoforms

Selected CaMKIIβ isoforms were expressed in High Five insect cells via the baculovirus system. All purification steps were performed at 4°C. Cell pellets were resuspended in CaMKII lysis buffer (10 mM Tris/HCl pH7.5, 500 mM NaCl, 1 mM EDTA, 1 mM EGTA, 5% Glycerol, 1 mM DTT) supplemented with protease inhibitors (cOmplete, Roche) and lysed by sonication. Insoluble particles were separated by centrifugation at 21,500 rpm for 1 h. The soluble fraction was incubated with Strep-Tactin Sepharose beads (IBA Lifesciences) for 1 h and washed with CaMKII lysis buffer. Bound protein was eluted with CaMKII SEC buffer (50 mM PIPES pH 7.5, 500 mM NaCl, 1 mM EGTA, 10% glycerol, 1 mM DTT) containing 2.5 mM desthiobiotin (IBA Lifesciences). Eluted protein was concentrated and run on a Superose 6 10/300 GL size exclusion column with CaMKII SEC buffer. Fractions were pooled according to SDS-PAGE and chromatogram, concentrated to approx. 1 mg/ml and flash frozen

in single-use aliquots in liquid nitrogen. Before use, aliquots were thawed on ice, gently mixed by pipetting and centrifuged at 20,000 rcf for 5 min. Exactly equal concentrations were determined by repeated SDS-PAGE, coomassie staining and quantification with ImageQuant TL (Cytiva).

## Expression and purification of CaMKII kinase domains

The kinase domains of all 4 human CaMKII genes (α: 1-315, β: 1-316, γ: 1-316, δ: 1-316) were expressed in BL21 RIL cells, in TB or autoinduction buffer. The expression construct featured an N-terminal StrepII-His-TEV tag. All purification steps were performed at 4°C. Cell pellets were resuspended in CaMKII Kinase Domain Lysis Buffer (50 mM PIPES pH 7.5, 300 mM NaCl), supplemented with protease inhibitors (cOmplete, Roche) and lysed by sonication. Insoluble particles were separated by centrifugation at 21,500 rpm for 1 h. The soluble fraction was loaded on StrepTactinXT beads (IBA Lifesciences) by gravity flow and washed with 10 CV CaMKII Kinase Domain Lysis Buffer. Bound protein was eluted in CaMKII Kinase Domain Lysis Buffer supplemented with 50 mM biotin (IBA Lifesciences). Fractions were pooled according to SDS-PAGE, supplemented with 10% glycerol and flash frozen in liquid nitrogen.

## Expression and purification of CaMKII substrate Syntide 2-GST

The sequence for Syntide 2 (PLARTLSVAGLPGKK) was expressed as a GST fusion protein, with a TEV cleavable N-terminal His-tag. A short linker (GGGGSGGGGS) was inserted between the Syntide 2 sequence and the C-terminal GST-tag. The fusion protein was expressed in BL21-RIL cells using auto-induction medium. All purification steps were performed at 4°C. Cell pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 20 mM imidazole, 1 mM DTT) containing protease inhibitors (cOmplete, Roche) and lysed by sonication. Insoluble particles were separated by centrifugation at 21,500 rpm for 1 h. The soluble fraction was loaded on a HisTrap Crude column (GE) and eluted with a linear gradient of elution buffer (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 500 mM imidazole, 1 mM DTT). Target fractions were pooled, supplied with TEV protease (self-made) and dialysed against lysis buffer overnight. Digested samples were re-run on a HisTrap Crude column. The flow through was collected, concentrated and run on a High Load Superdex 75 26/60 size exclusion column, equilibrated with SEC buffer (20 mM PIPES pH 7.5, 50 mM NaCl). Target fractions were pooled, concentrated to 22 mg/ml and flash frozen in liquid nitrogen.

## Expression and purification of human full-length tau (tau 441)

Human full-length tau (tau 441) was expressed as a fusion protein with an N-terminal His- and a C-terminal StrepII-tag. The protein was expressed in BL21 RIL cells in TB medium. Bacteria were grown at 37°C until an optical density of 0.6-0.8. Protein expression was induced with 1 mM IPTG for 3 h at 37°C. Cell pellets were resuspended in PBS buffer supplemented with 5 mM imidazole and protease inhibitors (cOmplete, Roche). Cells were lysed by sonication and incubated at 80°C in a water bath for 10 min with sporadic manual agitation. The lysate was cooled on ice for 10 min and supplemented with fresh protease inhibitors and 2 mM DTT. The lysate was cleared by centrifugation at 21,500 rpm for 30 min. The supernatant was loaded onto a HisTrap FF Crude 5 ml column (Cytiva) equilibrated with PBS supplemented with 5 mM imidazole and 1 mM DTT. The column was washed until baseline and the protein eluted with a linear gradient from 5-500 mM imidazole. Fractions were pooled based on the chromatogram and SDS-PAGE. Pooled fractions were loaded on a StrepTrap 5 ml column (Cytiva), equilibrated with PBS + 1 mM DTT. The column was washed until baseline and the protein was eluted with PBS containing 1 mM DTT and 2.5 mM desthiobiotin (IBA Lifesciences). Fractions were pooled based on the chromatogram and SDS-PAGE. The pooled sample was concentrated using a molecular weight cut-off of 3 kDa and run on a Superdex S200 26/60 (GE), equilibrated in PBS supplemented with 1 mM DTT. Fractions were pooled based on the chromatogram and SDS-PAGE, concentrated to ~ 15 mg/ml and flash frozen in liquid nitrogen.

## *In vitro* kinase assay

The protocol was adapted from (Coultrap and Bayer, 2012). CaMKII activity was measured by $^{32}P$ incorporation into the substrate Syntide-2-GST or tau 441 (human). The model substrate Syntide-2 (Hashimoto and Soderling, 1987) was linked to GST to increase its molecular weight, facilitate purification and enable deficient separation on an SDS-PAGE. Reactions were performed in cap-less 0.2 ml PCR-stripes. Purified CaMKIIβ was diluted to 10 nM in a mix containing 50 mM PIPES pH 7.2, 0.1% BSA, 2 mM $CaCl_2$, 10 mM $MgCl_2$, 50 µM Syntide-2-GST or 10 µM tau (human tau 441). The reaction was started by adding 1 nM to 4 µM calmodulin (Calbiochem) and 100 µM ATP (~1 Ci $mmol^{-1}$ [$\gamma^{32}P$]-ATP). Reagents were pre-incubated at 30°C for 5 min. Reactions were carried out in a final volume of 30 µl for 2 min at 30°C. Reactions were terminated by adding 10 µl SDS sample buffer. Samples were run on a 12.5% SDS-

PAGE, dried and analysed via a photostimulable phosphor plate. Gels were quantified using ImageQuant 5.2 or ImageQuant TL (Cytiva). Results were plotted using Graph Pad Prism 6 and fit to a Hill equation. For the standard IVK assay, the experiment was repeated two times in triplicates. To compare the maximal activity at optimal calmodulin concentrations, $V/Vmax_{FL}$ values for calmodulin concentrations from 100-1000 nM were pooled and plotted using Graph Pad Prism 6. Normal distribution and equality of variances was tested via Shapiro-Wilk test, Q-Q-Plots and F-test. Based on the results, a Student's t-test or Welch's t-test was performed. Resulting p-values were adjusted for multiple comparison using Holm's method. Statistical analysis was performed in R and RStudio. For the temperature gradient, reaction mixtures containing 100 nM calmodulin were incubated at the designated temperature for 15 min and the reaction was started by adding ATP. This pre-incubation of CaMKII with calmodulin in the absence of ATP was essential to detect the temperature dependency. Reactions were carried out for 3 min at the designated temperature and terminated by addition of 10 µl SDS sample buffer. For the autoactivity assay, the activation of CaMKII with varying concentrations of calmodulin was performed in the absence of the substrate protein. After a 2-minute incubation, the activation was quenched by addition of 5.3 mM EGTA. The substrate protein was added together with 3.3 mM $MgCl_2$ to enable the phosphorylation reaction. The sample was incubated for 3 min and the reaction terminated with SDS sample buffer. The analysis was performed as described above. In addition to the normal reaction buffer, an alternative high-salt buffer (50 mM PIPES pH 7.2, 0.1% BSA, 2 mM $CaCl_2$, 10 mM $MgCl_2$, 300 mM NaCl) was used.

## Cryogenic electron microscopy

Selected CaMKIIβ isoforms were expressed and purified as described above. Samples were concentrated to 1 mg/ml (initial data collection and statistical analysis) or 5 mg/ml (final reconstruction of the CaMKIIβΔ16,17 hub) in size exclusion buffer. For the final reconstruction, the buffer was supplemented with 0.002% (w/v) lauryl maltose neopentyl glycol (LMNG, Anatrace) to overcome preferred orientation. A total of 3.8 µl was applied onto glow-discharged 300 mesh holey gold UltrAuFoil R1.2/1.3 grids (Quantifoil Micro Tools GmbH) and vitrified using a Vitrobot Mark IV (Thermo Fisher Scientific) set to 10°C and 100% humidity by plunging into liquid ethane after 4s of blotting. Initial dataset of the FL and Δ16,17 isoform were collected in a FEI Talos Arctica transmission electron microscope (Thermo Fisher Scientific) operated at 200

kV and equipped with a Falcon 3EC direct electron detector operated in linear mode. 2D classification as described below was used to assess the statistical distribution of the different oligomeric states for the FL and Δ16,17 isoforms.

For the final reconstruction of the CaMKIIβΔ16,17 hub, data was collected on a FEI Titan Krios G3i transmission electron microscope (Thermo Fisher Scientific) operated at 300 kV equipped with a Falcon 3EC at a nominal magnification of 96,000×, corresponding to a calibrated pixel size of 0.832 Å. Objective astigmatism and coma were corrected with AutoCTF (Thermo Fisher Scientific) under the final imaging conditions. An electron flux of 0.7 e− per pixel per second on the detector was selected, corresponding to an exposure rate of 1 $e^{-}$ $Å^{-2}$ $s^{-1}$ on the sample. Images were taken at a nominal defocus of between −0.6 and −2 µm, accumulating a total electron exposure of 40 $e^{-}$ $Å^{-2}$ during a 40 s exposure, fractionated into 33 images. For automated data acquisition, EPU 2.8.1 (Thermo Fisher Scientific) was utilized with aberration-free image shift (AFIS) enabled, allowing 6 µm image-beam-shift acquisition. Raw movies were aligned and dose-weighted with patch-motion correction implemented in cryoSPARC version 2.9 (Punjani et al., 2017). Initial CTF estimation was achieved using Patch CTF. Particle picking was initially done with the Blob picker using a gaussian disc of 120 – 180 Å diameter on a subset of the micrographs. Particle images were extracted with a box size of 280 px, Fourier-cropped to 70 px (pixel size 3.328 Å) and subjected to 2D classification. The best appearing classes were used for ab initio refinement into 3 classes without imposing symmetry. One of the reconstructions clearly represented the typical 12-mer CamKII hub domain and was further refined by NU refinement. To improve particle picking and enrich the dataset for rare side views, low-pass filtered projections of the reconstruction were used for template-based picking of 668,392 particles from a total of 1,185 micrographs.

After iterative 2D classification cycles, 301,046 particles were selected for further processing. The 2D class averages showed obvious differences in the architecture of the hub domain, further referred to open and closed 12-mer as well as a 14-mer. A subset of 245,495 particles was subjected to local motion correction with re-extraction using a box size of 320 px (0.832 Å/px). Iterative 2D classification cycles were applied to carefully dissect all the existing conformations in the sample which turned out not to separate all states properly. Finally, 3D variability analysis (3DVA) was applied to isolate an open 12-mer (30,621), closed symmetric 12-mer (139,526), closed

asymmetric (27,310) and 14-mer hub structure (38,726) which were refined individually by NU refinement. A small subset of 3,430 particles were assigned to a 16-mer hub that could only be found by 2D classification and was absent during 3D classification attempts. See supplementary table 1 for cryo-EM data collection and refinement statistics.

For model building, initial models were generated by fitting in the structure of the CaMKIIα hub (pdb: 5IG3). This model was adjusted to the cryoEM density manually using Coot (version 0.8.9.1) (Casañal et al., 2020) and refined against the cryoEM map using the real-space refinement function in Phenix (version 1.19.1) (Liebschner et al., 2019). Structure figures were prepared using PyMOL (version 2.4.0, Schrödinger) or UCSF ChimeraX (Pettersen et al., 2021). Secondary structure elements were assigned using the dss function in PyMOL, or dssp in ChimeraX , and manually adjusted in the case of the open 12-mer structure, based on the assigned for the symmetric 12-mer structure.

## Analog-sensitive kinase assay – cloning and western blot

Analog-sensitive (AS) kinase mutants of CaMKIIβ were based on a previously published AS mutant for CaMKIIα (Wang et al., 2003). Sequences of the different CaMKIIβ variants were PCR amplified, fused to an N-terminal TEV-cleavable Twin-Strep-tag and inserted into the pcDNA3.1 vector backbone. The previously published F89G mutation was inserted via QuickChange Mutagenesis PCR (Agilent). A kinase dead version (K43R) was generated in a similar fashion. For initial testing, the different variants were expressed in HEK cells as described above and total lysate was generated by resuspending the cell pellet in SDS sample buffer. Semi-dry western blotting was performed using standard methods and a PVDF-blotting membrane (Immobilon-P, Merck). The blot was blocked in 2% BSA (w/v) in LS-TBST and subsequently developed with an anti-CaMKIIβ antibody (sc-376828, Santa Cruz) 1:500 in 2% BSA (w/v) in LS-TBST overnight at 4°C. An HRP-linked anti-mouse antibody (Cell Signalling) was used as a secondary antibody and incubated for 1 h at RT. The western blot was developed with a standard ECL reagent (Pierce ECL Western Blotting-Substrate, ThermoFisher).

## Analog-sensitive kinase assay – pulldown and *in vitro* kinase assays

HEK cells were cultured as described above and seeded at a concentration of $0.2*10^5$ cells/ml and 12 ml in 10 cm dishes or 30 ml in 15 cm dishes. The CaMKIIβΔ13,16 analog-sensitive mutant was transfected 24 h after seeding as described above, using 12 µg for 10 cm dishes and 36 µg for 15 cm dishes. 24 h after transfection, cells were harvested with trypsin, transferred to 1.5 ml reaction tubes and washed with PBS before being flash frozen in liquid nitrogen and stored at -80°C. Cell pellets corresponding to 3x 10 cm dishes and 2x 15 cm dishes were thawed on ice and resuspended in CaMKII lysis buffer (10 mM Tris/HCl pH7.5, 500 mM NaCl, 1 mM EDTA, 1 mM EGTA, 5% Glycerol, 1 mM DTT) supplemented with protease inhibitors (cOmplete EDTA-free, Roche). Cells were lysed by sonication on ice at 40% amplitude, 0.5 cycle and six rounds of 5 s. Lysates were cleared by centrifugation at 20,000 rcf, 4°C for 30 min. The supernatant was transferred to a new reaction tube and mixed with 50 µl pre-equilibrated StrepTactinXT beads (IBA) and supplemented with biotin-blocking solution (IBA). Samples were incubated for 1 h at 4°C with slow rotation. Beads were sedimented by centrifugation at 500 rcf, 4°C for 5 min. Beads were washed three times in CaMKII SEC buffer (50 mM PIPES pH 7.5, 500 mM NaCl, 1 mM EGTA, 10% glycerol, 1 mM DTT) and bound protein eluted in six steps with CaMKII SEC buffer supplemented with 50 mM biotin (IBA). The eluate was dispersed into single-use aliquots, flash frozen in liquid nitrogen and stored at -80°C. To compare the AS mutant to the wt kinase, a standard *in vitro* kinase assay (IVK) was performed as described above, using a limited range of calmodulin concentrations and roughly estimating the concentration via UV-absorption at 280 nm. To test the inhibition by various ATP analogs, the standard IVK assay was modified and set to a single calmodulin concentration of 100 nM. The reaction mixture contained varying concentrations of non-radioactive ATP (0-1 mM) or 0.5 mM of one of the following non-radioactive ATP analogs: $N^6$-methyl-ATP, $N^6$-etheno-ATP, $N^6$-phenyl-ATP, $N^6$-benzyl-ATP (Jena Bioscience). To test the utilization of different ATP analogs, an *in vitro* kinase was performed with two different reaction buffers, the standard IVK reaction buffer (50mM PIPES pH 7.2, 0.1% BSA, 2 mM CaCl2, 10mM MgCl2) and an alternative high salt reaction buffer (40 mM PIPES pH 7.5, 500 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$). All reactions contained 200 nM calmodulin, 31 µM Syntide-2-GST and 200 ng analog-sensitive kinase mutant. The reaction was started by addition of $N^6$-benzyl-ATPγS, $N^6$-phenyl-ATPγS (Jena Bioscience) or $H_2O$ for control reactions. The samples were

incubated for 30 min at 30°C and subsequently alkylated with 50 mM PNBM (*p*-nitrobenzyl mesylate, Agilent) at a final concentration of 2.5 mM for 1 h at RT. The reaction was terminated by addition of SDS sample buffer (containing DTT) and the samples analysed via standard SDS-PAGE and semi-dry western blotting. The blot was developed using an anti-thiophosphate ester antibody (ab92570, Abcam) and an HRP-linked anti-rabbit antibody (Cell Signalling).

## Analog-sensitive kinase assay – optimization of labelling conditions

To find the optimal labelling conditions, CaMKIIβ analog-sensitive mutants were expressed in N2a cells. Cells were seeded at a density of $0.2*10^6$ cells/ml and 12 ml in 10 cm dishes. 24 h after seeding, cells were transfected as described above. 24 h after transfection, cells were washed with PBS and harvested using a cell scraper. Reactions were performed in AS lysis buffer (20 mM PIPES pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM EGTA, 200 nM calmodulin), using varying concentrations of ATP, GTP and Tween-20 (Sigma) or nOG (octyl-β-D-glucopyranoside, Sigma) as permeabilization agent. Alternatively, cells were lysed by addition of 0.5% (v/v) NP40, 0.5% (v/v) Cell Triton X-100 and sonication at 40% amplitude, 0.5 cycle and 2 rounds of 6s. Some reactions additionally contained the recombinant substrate Syntide-2-GST at 40 µM. Samples were incubated for 30 min at 30°C and quenched by addition of EDTA/EGTA to a final concentration of 10 mM each. 50 mM PNBM was added to a final concentration of 2.5 mM and incubated for 1-2 h at RT. SDS-PAGE and western blot with a thiophosphate-specific primary antibody was performed as described above.

## Analog-sensitive kinase assay – cell permeabilization test

N2a cells were cultures as described above and seeded at a concentration of $0.4*10^6$ cells/ml in 12-well plates (1 ml/well) and 24-well plates (0.5 ml/well). After 24 h, cells were washed with PBS and incubated with the detergents sodium dodecyl sulfate (SDS, Sigma), lauryl maltose neopentyl glycol (LMNG, Anatrace) or n-octyl-β-D-glucopyranoside (nOG, Sigma) dissolved in PBS at varying concentrations for 10-30 min at 37°C. Permeabilized cells were stained with low concentrations of trypan blue contained in the permeabilization buffer. Additionally, existence of protein-permeable pores was confirmed by addition of an anti-β-actin antibody (sc-47778, 1:200 in permeabilization buffer), as well as an anti-hnRNPL antibody (sc-32317, 1:2000 in permeabilization buffer). Both antibodies were detected via a fluorescence-labelled

secondary antibody (anti-mouse-Alexa Fluor 568, ThermoFisher, 1:2000 in permeabilization buffer). Cells were analysed on an epifluorescence microscope. Three biological replicates were tested per condition.

## Analog-sensitive kinase assay – *in vivo* labelling and thiophosphate enrichment

The thiophosphate enrichment strategy was based on (Michowski et al., 2020), with modifications. The analog-sensitive kinase mutants were PCR amplified with primers omitting the Twin-Strep-tag and cloned back into the pcDNA3.1 expression plasmid, to avoid interference from the affinity tag. N2a cells were cultures as described above and seeded into 15 cm dishes at a concentration of $0.1*10^6$ cells/ml and 30ml/dish. Cells were incubated for 24 h and transfected with 37.5 µg DNA and 75 µl Rotifect (Carl Roth) per 15 cm dish, as described above. Cells were grown for 48 h, washed with 20 ml PBS and subsequently 20 ml AS lysis buffer (20 mM PIPES pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM EGTA). The liquid was removed and the dish was carefully washed with 1.2 ml AS lysis buffer, supplemented with protease inhibitors (cOmplete, Roche), phosphatase inhibitors (PhosSTOP, Roche) and 0.5 mM TCEP. The liquid was removed thoroughly, the cells detached with a cell scraper, transferred to a reaction tube and kept on ice until all samples had been harvested. From then on, samples were processed in parallel in Protein LoBind tubes (ThermoFisher). Each 15 cm dish resulted in approximately 1.2 ml cell suspension, which was split into two 600 µl aliquots. The remaining cells were discarded. Each aliquot was supplemented with 75 µl detergent mix (3.6% nOG, 36 mM $CaCl_2$) and briefly mixed. The labelling reaction was started by addition of 225 µl reaction mix (200 nM calmodulin, 0.1 mM $N^6$-benzyl-ATPγS, 0.2 mM ATP, 3 mM GTP, PhosSTOP phosphatase inhibitors in AS lysis buffer) and incubated for 30 min at 30°C with sporadic manual agitation. For the untransfected control, calmodulin was omitted. The reaction was terminated by addition of EDTA/EGTA to a final concentration of 10 mM each. Labelled samples were briefly sonicated to create a homogeneous suspension and concentrations were determined by Pierce 660 nM assay (ThermoFisher). Samples were flash frozen in liquid nitrogen and stored at -80°C.

Samples were thawed and lysate corresponding to 6 mg protein was transferred into a 15 ml tube for protein precipitation. All samples were equalized in volume with AS lysis buffer and supplemented with 5 volumes of ice-cold methanol/chloroform mix

(ration 4:1), followed by 3 volumes of ice-cold $H_2O$. The samples were thoroughly mixed, incubated 10 min on ice and centrifuged for 20 min at 2000 rcf. The resulting pellet, located at the interface, was washed in 5 volumes ice-cold methanol and centrifuged for 20 min at 2000 rcf. The supernatant was removed and the pellet dried at RT. The dried pellet was resuspended in 800 µl denaturation buffer (100 mM $NH_4HCO_3$, 2 mM EDTA, 10 mM TCEP adjusted to pH 7-8, 8 M urea), adjusted to 6 M urea with $H_2O$ and incubated at 55°C for 1 h with agitation at 300 rpm. The sample was slowly cooled to RT for 10 min and diluted to 2 M urea with 50 mM $NH_4HCO_3$ in $H_2O$. TCEP (pH adjusted to 7-8) was added to a final concentration of 10 mM. Trypsin (Trypsin, TPCK treated from bovine pancreas, Sigma) was added at a ratio of 1:20 (w/w, based on starting material) and the samples were digested overnight at 37°C. The following morning, 10 M NaOH was added to a final concentration of 0.08 mM and the digestion continued for 3 h. The digest was acidified with 2.5% trifluoroacetic acid (TFA) to a final concentration of 0.1% and a pH of ~ 2.5. If required, more TFA was added to lower the pH. The digest was centrifuged for 3 min at 1400 rcf and the supernatant aliquoted to a new tube. SepPak Plus cartridges (Waters) were equilibrated by sequential washing with 10 ml 0.1% TFA/50% acetonitrile (in $H_2O$) and 10 ml 0.1% TFA (in $H_2O$). The sample was loaded by passing it through the cartridge 5 times. The cartridge was washed with 10 ml 0.1% TFA (in $H_2O$). Bound peptides were eluted with 4 ml 80% acetonitrile/0.1% acetic acid and dried overnight in a vacuum centrifuge. SulfoLink beads (ThermoFisher) were transferred to a Protein LoBind tube and washed with 200 mM HEPES pH 7.0. Beads were incubated with 200 mM HEPES pH 7.0, 25 µg/ml BSA for 10 min at RT in the dark. Beads were sequentially washed with 200 mM HEPES pH 7.0 and 2 times with 4 M urea, 0.1 M Tris pH 8.8, 10 mM TCEP (pH of stock solution ~ 2.5, this lowers the total pH to ~ 8.0). The dried peptides were resuspended in 4 M urea, 0.1 M Tris pH 8.8, 10 mM TCEP and acidified to pH 5 with 5% (v/v) formic acid. The peptide solution was added to the equilibrated beads and rotated overnight at RT in the dark. The next day, the beads were centrifuged at 2000 rcf for 3 min and the supernatant discarded. The beads were washed sequentially with 4 M urea in 20 mM HEPES pH 7.0, $H_2O$, 5 M NaCl, 50% acetonitrile in $H_2O$ and 5% (v/v) formic acid. Unused binding sites were blocked by incubation with a fresh solution of 10 mM DTT for 10 min in the dark. Bound peptides were eluted in three steps with a solution of 2 mg/ml Oxone (Potassium peroxymonosulfate, Sigma) in $H_2O$. Eluates were pooled and desalted using SepPak

Plus cartridges as described above. Samples were dried in a vacuum centrifuge and stored at -80°C.

To remove remaining contaminants, peptides were further purified with Styrene Divinyl Benzene (SDB) StageTips. StageTips were prepared by inserting the material into standard 200 µl pipet tips and washing sequentially with methanol, 80% acetonitrile in 0.1% formic acid and two steps of 0.1% formic acid in $H_2O$. The resuspended samples were acidified with 10% formic acid to a final concentration of 1%. The samples were loaded and passed through the StageTips, followed by sequential washing with 0.1% formic acid in $H_2O$ and two rounds of 80% acetonitrile in 0.1% formic acid. Bound peptides were eluted with 5% $NH_4OH$ in 60% acetonitrile, split into two equal aliquots and dried in a vacuum centrifuge. Peptides were measured on an Orbitrap Q Exactive HF (Thermo Scientific) or an Orbitrap Exploris 480 (Thermo Scientific). Resulting spectra were analysed using MaxQuant (Version 1.6.5.0) against the UniProt mouse reference proteome. Subsequent analysis was done in python (version 3.8.5, Anaconda distribution) using the packages pandas, numpy, matplotlib, seaborn, upsetplot, scipty, sklearn, matplotlib_venn. Contaminants and reverse peptide hits were removed and the analysis restricted to phosphorylated peptides with a localization probability ≥ 0.75. The overlap between the two datasets was calculated using the unique phosphosite (protein/gene name + identity of phosphorylated residue) as an index. The intensity values of both datasets were normalized before merging, using the min-max normalization: $x_{norm.} = \frac{x - \min(x)}{\max(x) - \min(x)}$. Min(x) and max(x) were set to the respective minimal or maximal value of the individual datasets. When pooling replicates, an average intensity value was calculated. If only one replicate featured an intensity value for the respective target, this value was kept. Correlation matrices were calculated using a Pearson correlation coefficient. For Boolean data, the Jaccard similarity index was used. To display isoform- and group-specific targets via UpSetPlots, targets identified by MS2 spectra were used. This included additional targets for which no MS1 peak (and thus no intensity value) could be extracted.

## Confocal microscopy

The selected CaMKII isoforms were cloned into a eukaryotic expression vector with an N-terminal tandem-tomato (tdTomato) tag, followed by a short linker sequence. Expression was controlled by a UbC promotor. N2a cells were seeded on poly-L-lysine coated glass slides placed in 24-well plates at a concentration of $0.4*10^5$ cells/well.

Plasmid were transfected the next day. After 24 h, cells were fixed for 20 min at RT with 3.7% PFA in PBS and permeabilized with 0.1% Trition X-100 in PBS for 5 min at RT. The actin cytoskeleton was stained with Phalloidin-iFluor™ 647 Conjugate (AAT Bioquest) according to the manufacturer's recommendations. Glass slides were embedded on an object plate with ProLong™ Diamond Antifade Mountant with DAPI (ThermoFisher) and dried overnight at RT. Images were collected on a confocal laser scanning microscope (Leica) with a 63x oil immersion objective.

## Generation of the mouse model

The mouse model was generated in the Transgenics Facility at the Max Delbrück Center for Molecular Medicine Berlin (MDC) under the supervision of Dr. Ralf Kühn. The model was based on the *Camk2β* minigenes. CRISPR/Cas9 was used to remove the 100 bp initially found to harbour the *cis*-acting element in the endogenous mouse *Camk2β* gene. A synthetic gene was used as a repair template to insert the human ortholog of the excised sequence into the endogenous mouse gene (humanized strain). A deletion strain was generated in which the repair process failed and only the mouse sequence was deleted.

## RNA Seq analysis

Mouse model

Total RNA was extracted from mouse cerebellum tissue as described above (minigene splicing assay). 4 male wt, 2 male and 2 female heterozygous and 4 male homozygous mice of the humanized strain were selected for RNA-Seq. For library preparation, DNAse I-digested RNA samples were filtered using the polyA+ selection method at BGI Genomics and sequenced using DNBSeq PE150 sequencing. This yielded ~50-60 million paired-end 150 nt reads. Reads were aligned to the GRCm38 genome using the STAR aligner (v.2.7.9a) (Dobin et al., 2013), yielding on average ~75% uniquely mapped reads. Files were indexed using SAMtools (Danecek et al., 2021) and the splicing pattern analysed using rMATS (v3.1.0) (Shen et al., 2014). Downstream analyses and data visualization were performed using standard python code (v3.8.5). Data was visualized and sashimi plots generated via IGV (Robinson et al., 2011). Gene expression patterns were analysed using Salmon (v1.8.0) (Patro et al., 2017) and DESeq2 (Love et al., 2014). Volcano plots were generated using GraphPad Prism 5-6.

Various mammals

Publicly available RNA-Seq data was analyzed for various mammals. For human, chimpanzee (*Pan trogodytes*), bonobo (*Pan paniscus*) and rhesus macaque (*Macaca mulatta*) data from cerebellum white tissue and cerebellum grey tissue from multiple individuals was selected. For gibbon (*Hylobates lar*), data from different brain regions from a single individual were selected. For gorilla (*Gorilla gorilla*) and orangutan (*Pongo pygmaeus*), data from cerebellum and total brain tissue was selected. For pig (*Sus scrofa*), data from cerebellum tissue was selected. Reads were aligned to the respective genome (human: GRCh38; chimpanzee: panTro6; bonobo: panPan1.1; gorilla: gorGor6; orangutan: ponAbe3 (*Pongo abelii*); Gibbon: nomLeu3 (*Nomascus leucogenys*); rhesus macaque: rheMac10; pig: SusScr11 using the STAR aligner (v.2.7.9a) (Dobin *et al.*, 2013). Subsequent analysis was performed as described above. To calculate %Skipped values for *Camk2β* exon 16, the sum of all individual exon 16 skipping events was calculated. For final visualization, cerebellum grey and white matter files (where available) were merged to create combined cerebellum files. A list of all used publicly available RNA-Seq data, including species, tissue, read length and used reference genome can be found in supplementary table S2.

## Electrophysical characterization

All experiments regarding the electrophysical characterization were entirely performed in the research group of Prof. Dietmar Schmitz (Charité, NeuroCure) under the supervision of Dr. Alexander Stumpf. Hippocampal slices were prepared from adult C57/BL6J and transgenic (deletion and humanized) mice. Animals were anesthetized with isoflurane and decapitated. The brain was quickly removed and chilled in ice-cold sucrose-based artificial cerebrospinal fluid (sACSF) containing (in mM): NaCl 87, $NaHCO_3$ 26, glucose 10, sucrose 50, KCl 2.5, $NaH_2PO4$ 1.25, $CaCl_2$ 0.5 and $MgCl_2$ 3, saturated with 95% (vol/vol) $O_2$/5% (vol/vol) $CO_2$, pH 7.4. Horizontal slices (300 µm) were cut and stored submerged in sACSF for 30 min at 35 °C and subsequently stored in ACSF containing (in mM): NaCl 119, $NaHCO_3$ 26, glucose 10, KCl 2.5, $NaH_2PO4$ 1, $CaCl_2$ 2.5 and $MgCl_2$ 1.3 saturated with 95% (vol/vol) $O_2$/5% (vol/vol) $CO_2$, pH 7.4, at RT. Experiments were started 1 to 6 h after the preparation.

Recordings were performed in a submerged recording chamber (Warner instruments RC-27L), filled with ACSF with solution exchange speed set to 3-5 ml/min at RT (22-24°C). Stimulation electrodes were placed in the stratium radiatum of CA1 (near CA3)

to stimulate Schaffer collaterals. Recording electrodes were placed in the str. radiatum of the CA1 field. Stimulation was applied every 10 s. In order to analyze the input-output relationship, stimulation intensities were adjusted to different FV amplitudes (0.05 mV increments, 0.05mV – 0.4 mV) and correlated with the corresponding field excitatory postsynaptic potential (fEPSP). Paired pulse ratios (PPR) were determined by dividing the amplitude of the second fEPSP (50 ms inter-stimulus interval) with the amplitude of the first (average of ten repetitions). Long term potentiation (LTP): Basal stimulation was applied every 10 s in order to monitor stability of the responses at least for 10 min before LTP was induced by one single high frequency stimulation train (100 pulses, 100 Hz). Magnitude of LTP was determined by normalizing the average of the initial fEPSP slopes 25-30 min and 55 -60 min after LTP induction to average baseline fEPSP slope. Data collection and quantification was performed blindly.

## Behavioural studies

All animal experiments were performed with C57BL/6 mice in accordance with institutional and governmental recommendations and laws (permission T 0311/13). Mice were kept under 12-hour light-dark conditions. For all behavioural experiments, littermate wt control animals were used. The nest building assay was performed as described in (Deacon, 2006). For the open field test, mice were acclimatised to a transparent plastic handling tunnel in their home cage for one week. Animals were handled for 1-3 minutes per animal for three consecutive days before the experiment was performed. On the fourth day, animals were allowed to freely explore an open, rectangular arena (40 x 40 cm) made out of white plastic for 10 min. Sessions were videotaped and analysed using optimouse (Ben-Shaul, 2017) in MATLAB 2021 (The MathWorks Inc.). For the deletion strain experiments, 2 sessions per animal were analysed and the open field arena additionally contained two small plastic objects. Subsequent data analysis was done in R (R Core Team, 2021) and RStudio (RStudio Team, 2021). Data was tested for normal distribution via Shapiro-Wilk test and Q-Q Plots, normality of variances was confirmed via F-test or Levene-Test. Based on these, means were compared via Student's t-test, Welch's t-test or Wilcoxon signed-rank test. P-values were adjusted for multiple comparisons via Holm's method.

## Figures

Figures 1, 2 and 3 were created with biorender.com (2022). Figure 3A was adapted and modified from the template "mRNA Splicing Types", retrieved from

https://app.biorender.com/biorender-templates. Figures 4 and 15 contain individual icons from biorender.com.

# Results

## Alternative splicing of CaMKII is species-specific

Alternative splicing of CaMKII has long been established and multiple studies have reported developmental stage- and tissue-specific splicing events (Tombes *et al.*, 2003). Differences between species are known but have so far been restricted to species that are evolutionary distant from humans, such as *Drosophila melanogaster* (Griffith and Greenspan, 1993) or *Caenorhabditis elegans* (Reiner *et al.*, 1999). For these species, the architecture and number of CaMKII genes differs considerably from vertebrate genomes, often featuring a single ancestral CaMKII gene. As opposed to that, CaMKII has largely been conserved during vertebrate evolution and all mammals feature the same four genes (Figure 4A). These genes show a conserved architecture and differences mostly relate to the presence or absence of certain exons in the variable linker domain. Alternative splicing of CaMKII in different vertebrates has been reported, but not systematically compared (Tombes *et al.*, 2003) (Cook et al., 2018; Rochlitz *et al.*, 2000; Sloutsky *et al.*, 2020). To address this, we performed radioactive RT-PCR with gene-specific primers on total cerebellum RNA from human and mouse (Figure 4B). Species-specific differences in the alternative splicing pattern can be seen for three of the four CaMKII genes (*Camk2β, γ*, and *δ*), such as a shorter isoform of *Camk2β* that appears to be exclusively present in the human sample. *Camk2γ* shows the highest diversity of splice isoforms in the cerebellum, despite featuring only five alternative exons compared to the six alternative exons of the *Camk2β* gene. However, three of these alternative exons in *Camk2β* (exons 19 to 21) are only included in endocrine tissue (Bayer *et al.*, 1998; Urquidi and Ashcroft, 1995). The identity of CaMKII splice isoforms was also confirmed by partial sequencing of the detected splice isoforms. Importantly, although the primers are located within exons represented in all CaMKII paralogues and conserved between all species, the relative positions had to be shifted in some of the genes. One can therefore not directly compare the transcript sizes and thus linker length between the CaMKII genes based on the PCR product sizes. Similarly, the primer position for *Camk2α* had to be shifted between human and mouse, leading to a perceived difference of their splicing pattern. However, for both species the main bands visible on the gel correspond to the same *Camk2α* splice isoforms, as was confirmed by sequencing.

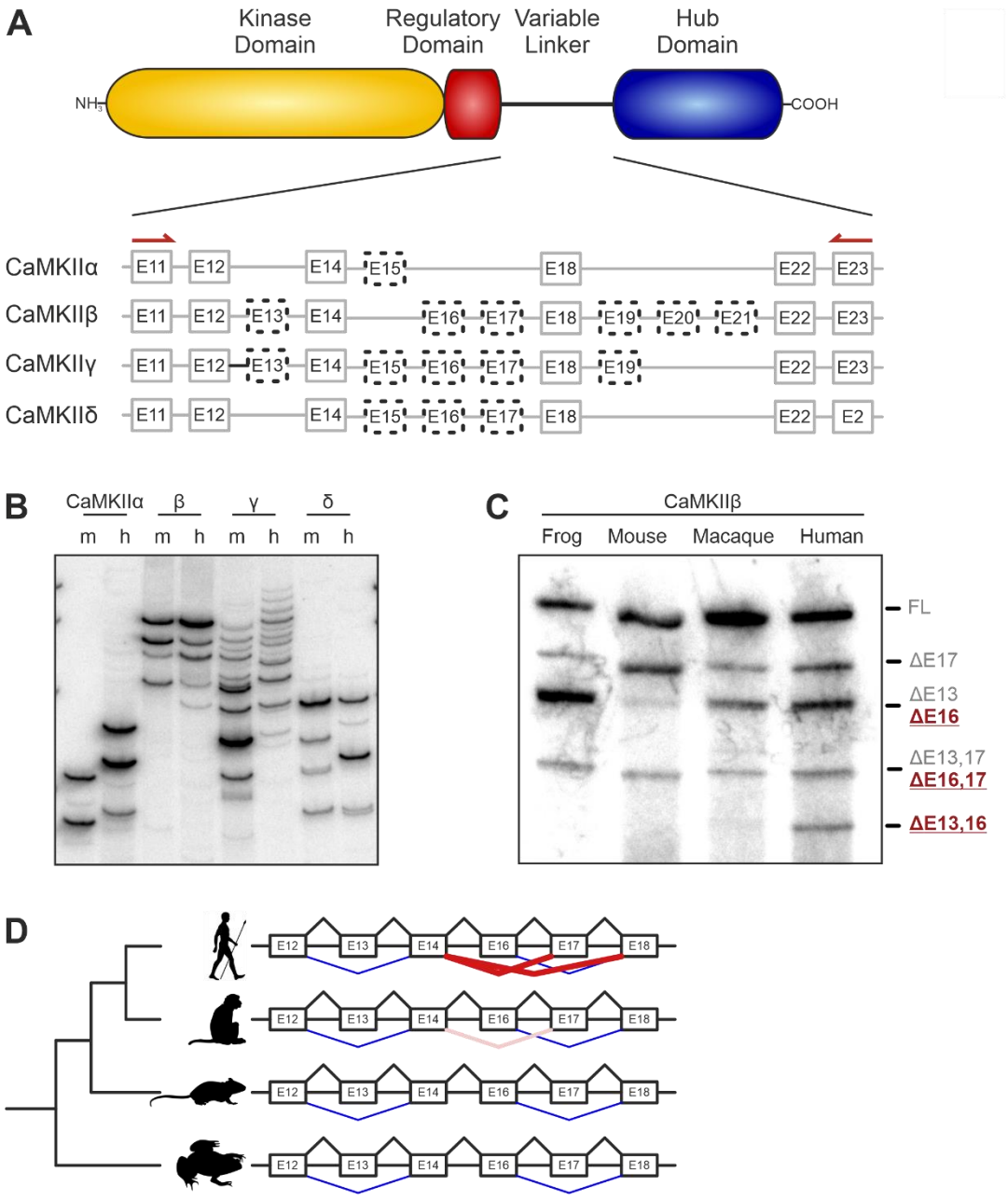



Figure 4: Exon 16 of *Camk2β* is spliced in a primate-specific manner.

(A) Schematic representation of the domain architecture of CaMKII and the intron-exon structure of the variable linker region of the four mammalian CaMKII genes. Numbered boxes represent exons, connecting lines represent introns. Boxes with dashed lines represent known alternatively spliced exons. (B, C) Endogenous CaMKII splice isoforms were identified by radioactive isoform-specific RT-PCR with frog (*Xenopus laevis*), and primate (*Macaca mulatta*) total brain RNA and mouse (Mus musculus) and human cerebellum RNA. Isoforms were separated on a denaturing polyacrylamide gel. Isoforms are indicated on the right and named according to the exons that are skipped. As exons 19-21 are missing in neuronal tissue, they were excluded from the naming scheme. (D) Schematic representation of the identified species-specific splicing events for *Camk2β*. Coloured lines indicate alternative splicing events detected by RT-PCR.

We also extended our analysis to include rhesus macaque (*Macaca mulatta*), which mostly resembles the human or mouse splicing pattern (Figure S1A). More precisely, the macaque sample shows the same two known splice isoforms for *Camk2α*, although the shorter transcript seems to be predominant. While it looks essentially identical to the human splicing pattern for *Camk2γ*, it is different from both species for CaMKIIδ. For *Camk2β* on the other hand, the rhesus sample seems identical to the mouse splicing pattern. *Camk2β*, together with *Camk2α*, is one of the brain-specific CaMKII genes and has long been established as a key regulator of synaptic plasticity (Bayer and Schulman, 2019). The observed splicing pattern appears human-specific, and distinct enough for a detailed analysis. We further investigated this, also including the African clawed frog (*Xenopus laevis*) (Figure 4C). All visible splice isoforms were identified by Sanger sequencing and revealed the species-specific alternative splicing of *Camk2β* exon 16 (previously also named exon IV/V (Tombes *et al.*, 2003)), whose inclusion or exclusion leads to three species-specific splice isoforms. The shortest of these, lacking exons 13 and 16 (termed Δ13,16) can easily be identified as primate-specific on the polyacrylamide gel. It also appears to be a more frequent outcome in humans, as only a faint band is visible for rhesus macaque. Partial sequencing of the splice isoforms from *Camk2γ* and *δ* suggests that exon 16 is not spliced in a species-specific manner for these genes, as exon 16 exclusion isoforms can also be identified in mouse samples for both genes. Exon 16 is furthermore the least conserved exon in the linker segment, differs in size between the CaMKII genes and, in *Camk2γ*, features an additional splice donor site (Tombes *et al.*, 2003). Importantly, exons 19 to 21 were not present in any of the detected *Camk2β* isoforms, for any of the investigated species. Therefore, the full-length (FL) isoform refers to the longest isoform detected in the cerebellum. Together, this establishes the species-specific alternative splicing of *Camk2β, γ* and *δ* and reveals a novel primate-specific regulation of *Camk2β* exon 16 (Figure 4D).

## Minigene splicing assays indicate the presence of a *cis*-acting element

Based on these findings, minigenes from human, mouse (*Mus musculus*, C57BL/6 strain) and frog (*Xenopus laevis*) were designed. Minigenes are eukaryotic expression constructs that contain a genomic segment from a gene of interest, including alternatively spliced exons and flanking genomic regions (Cooper, 2005) (Figure S2A). In contrast to normal expression constructs, they also contain intronic sequences and therefore undergo splicing. As endogenous genes can be exceedingly long (the human *Camk2β* gene spans ~ 108 kb), minigenes enable the selection of a specific region for a detailed, yet simplified splicing analysis. In its most simple version, a minigene contains two constitutive exons, one with the transcription start site, the other with the terminator sequence and poly(A) signal. Inserted in between these two exons is a sequence containing one or several alternatively spliced exons, including the flanking or intermediate introns. As mammalian introns can be too long for plasmid-based cloning and expression systems, they are often shortened in a way that maintains the intron/exon boundaries.

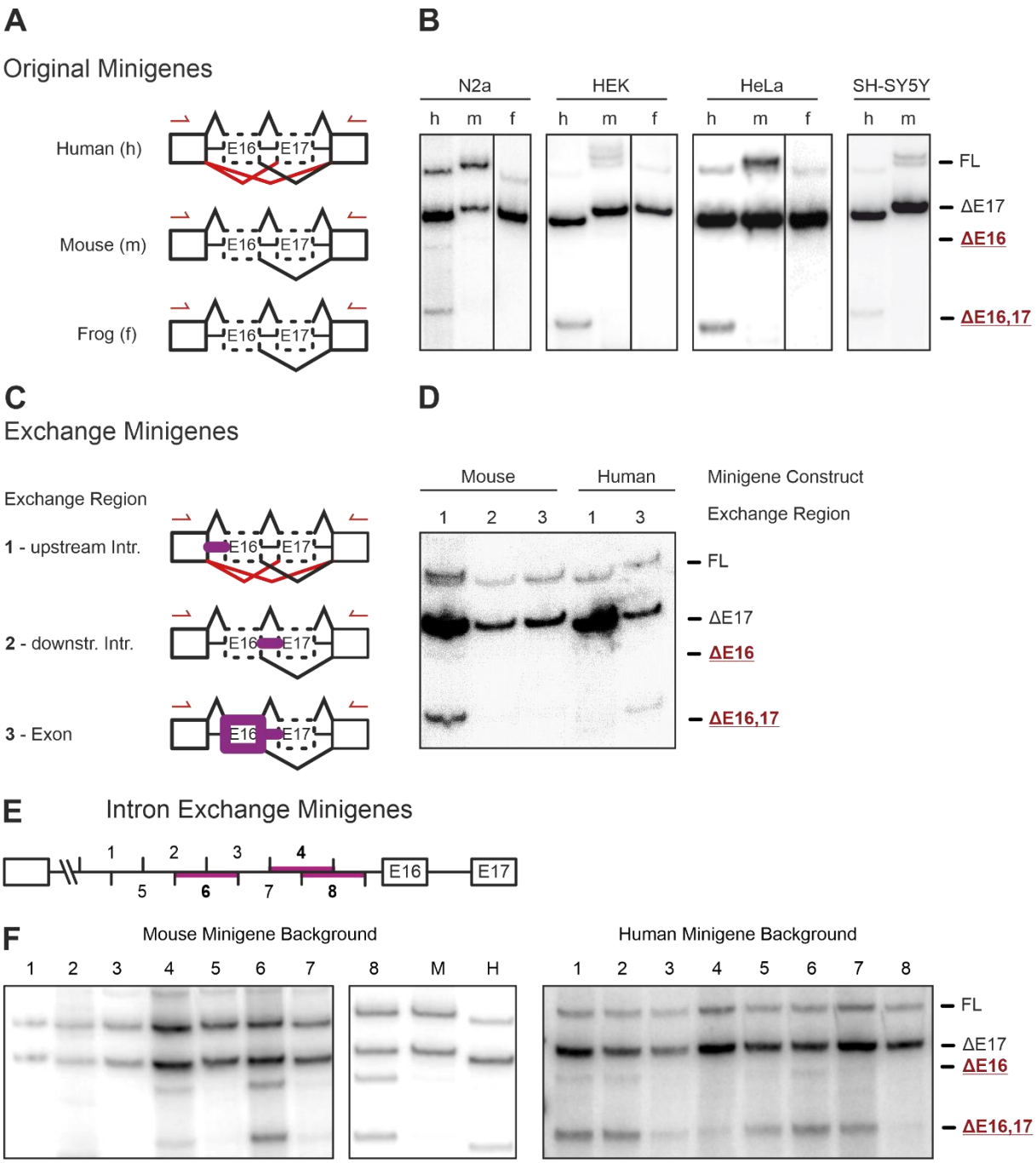



Figure 5: Minigenes reveal a *cis*-acting element in the upstream intron.

(A) Schematic representation of the minigene constructs used in B. Red lines indicate human-specific splicing events. Arrows indicate positions of primer used for RT-PCR. (B) The human (h), mouse (m) and frog (f) (*Xenopus laevis*) sequences of exons 16 and 17, including the adjacent introns, were cloned in between two constitutive exons and transfected into N2A (mouse), HEK, HeLa and SH-SY5Y cells. Resulting splice isoforms were identified by radioactive RT-PCR. Isoforms are indicated on the right and named according to the exons skipped. (C) Schematic representation of the minigene constructs used in D. Red lines indicate human-specific splicing events. Purple lines highlight segments of the minigene that were exchanged between the human and mouse construct. (D) Human and mouse exchange minigenes were transfected into HEK cells and resulting splice isoforms identified by radioactive RT-PCR. (E) Schematic representation of the intron containing the identified functionally relevant *cis*-acting element. Numbers indicate 20 bp segments that were exchanged between the human and mouse construct. Purple lines highlight segments of functional relevance identified in F. (F) Human and mouse exchange minigenes were transfected into N2a cells and resulting splice isoforms identified by radioactive RT-PCR. Experiments E and F were performed by Nicole Dimos.

The *Camk2β* minigenes feature two constitutive exons, in between which the alternative exons 16 and 17 were inserted (Figure 5A, S2A). Both the central intron, and the proximal regions of the flanking introns were included. In order to maintain the intron/exon boundaries of the constitutive exons, the proximal region of their flanking introns was likewise inserted. The minigenes were transfected into various cell lines and the splicing pattern analysed by radioactive RT-PCR with a vector specific primer pair (Figure 5B). The splicing pattern of the minigenes recapitulates the observed endogenous *Camk2β* splicing pattern. Specifically, all minigenes show bands corresponding to the full-length (FL) and Δ17 isoforms, whereas only the human minigene shows additional bands for the Δ16 and Δ16,17 isoforms. Transfection of the minigenes into various human and mouse cell lines reveals that the observed splicing pattern is independent of the cell line and thus of the *trans*-acting environment. This strongly indicates a *cis*-regulated mechanism, in which differences on the sequence level determine the observed species-specific splicing patterns. The actual PSI (percent spliced in) value might differ between cell lines, minigenes and sometimes even experiments and reflects the artificial nature of the minigene construct. The species-specific aspect is nevertheless fully captured as the alternative splicing of exon 16 is strictly limited to the human minigene.

To pinpoint the location of the *cis*-acting element, a second set of minigenes was designed (Figure 5C). In these, intronic or exonic sequences that were thought to possibly harbour the regulatory element of interest were systematically exchanged between the human and mouse minigene. Subsequent splicing analyses locate the *cis*-acting element to the intron upstream of exon 16 (Figure 5D). Insertion of this sequence into the mouse minigene is sufficient to induce the human splicing pattern. Conversely, transfer of the mouse sequence into the human minigene abolishes exon 16 exclusion. Transfer of any other sequence does not lead to an observable change of exon 16 splicing. Together, this confirms the primate-specific regulation of *Camk2β* exon 16 and shows that the mechanism is *cis*-regulated, with the regulatory element located in the upstream intron.

Having identified the approximate position of the *cis*-regulatory element, we set out to determine its exact location and sequence. As described above, the *Camk2β* minigenes contain only a part of the intron upstream of the alternative exon 16 (Figure S2A). These 100 base pairs (bp) were further subdivided into eight overlapping

segments of 20 bp (Figure 5E). The splice site itself, including the first 15 bp upstream of it, is identical between human and mice. The eight segments were exchanged between the human and mouse minigenes, and the resulting splicing pattern analysed after expression in human and mouse cell lines (Figure 5F, Figure S2B). No difference between the tested human and mouse cell lines can be observed, further supporting the *cis*-regulated nature of the splicing event. RT-PCR identified three segments of functional importance, two of which overlap by 10 bp. These two segments (segment 4 and 8) act bidirectionally: transfer of the mouse sequence into the human minigene is sufficient to abolish the human-specific exclusion of exon 16, whereas transfer of the human sequence into the mouse context induces de-novo exclusion of exon 16. The third identified segment (segment 6) only works in one direction: transfer from human to mouse induces exon 16 exclusion, whereas the corresponding mouse sequence inserted into the human minigene does not change the splicing pattern. Interestingly, the overlap of segments 4 and 8 is relatively well-conserved between the two species, further narrowing down the possible choice of functionally relevant nucleotides (Figure S2C). Together, this reveals two sequences in the intron upstream of *Camk2β* exon 16, that regulate its alternative splicing.

## Identification of potential *trans*-acting factors

The presence of a *cis*-acting element suggests the existence of a corresponding *trans*-acting factor as a binding partner. In general, *cis*-acting elements act as recognition motifs for *trans*-acting proteins, which themselves are either part of the spliceosome or recruit components of it (Ule and Blencowe, 2019). The identified *cis*-acting element regulates the primate-specific exclusion of *Camk2β* exon 16. It can therefore be classified as an intronic splicing silencer binding to the human sequence, or an intronic splicing enhancer binding to the mouse sequence. To identify candidate *trans*-acting factors, we first screened publicly available CLIP datasets of various splicing regulators but failed to find any significant hits. This might, in part, be due to *Camk2β* being predominantly expressed in neuronal tissue and thus not being present in most publicly available CLIP datasets. As an alternative, we tried to predict potential *trans*-acting factors based on the previously identified *cis*-regulatory sequences. We used the tools DeepCLIP (Grønning *et al.*, 2020) and RBPmap (Paz *et al.*, 2014) and focused our analysis on the overlap between segments 4 and 8. DeepCLIP consists of a context-aware neuronal network that predicts protein binding to nucleic acid sequences and generates context-dependent binding motifs.The program enables the comparison of two sequences, and we restricted the search to proteins that were predicted to bind specifically to either the human or mouse sequence (Figure S3A). Nine candidates, including control targets, were selected and validated in an siRNA knockdown, combined with the established minigene splicing assay (Figure 6A, B). None of the knockdown targets showed a reproducible effect on exon 16 exclusion. Nova1 seems to induce low levels of exon 16 skipping in the mouse minigene, though this could not be replicated (Figure S3B, C). The triple knockdown of hnRNPH1/H2/F, which all recognize similar sequences, seems to have a weak effect on the human minigene. However, this knockdown leads to a slight increase in exon 16 skipping and does not fit the observed minigene splicing patterns, which suggest a splicing silencer in the human sequence. The observed effects might therefore rather be caused by general perturbations of the splicing machinery.

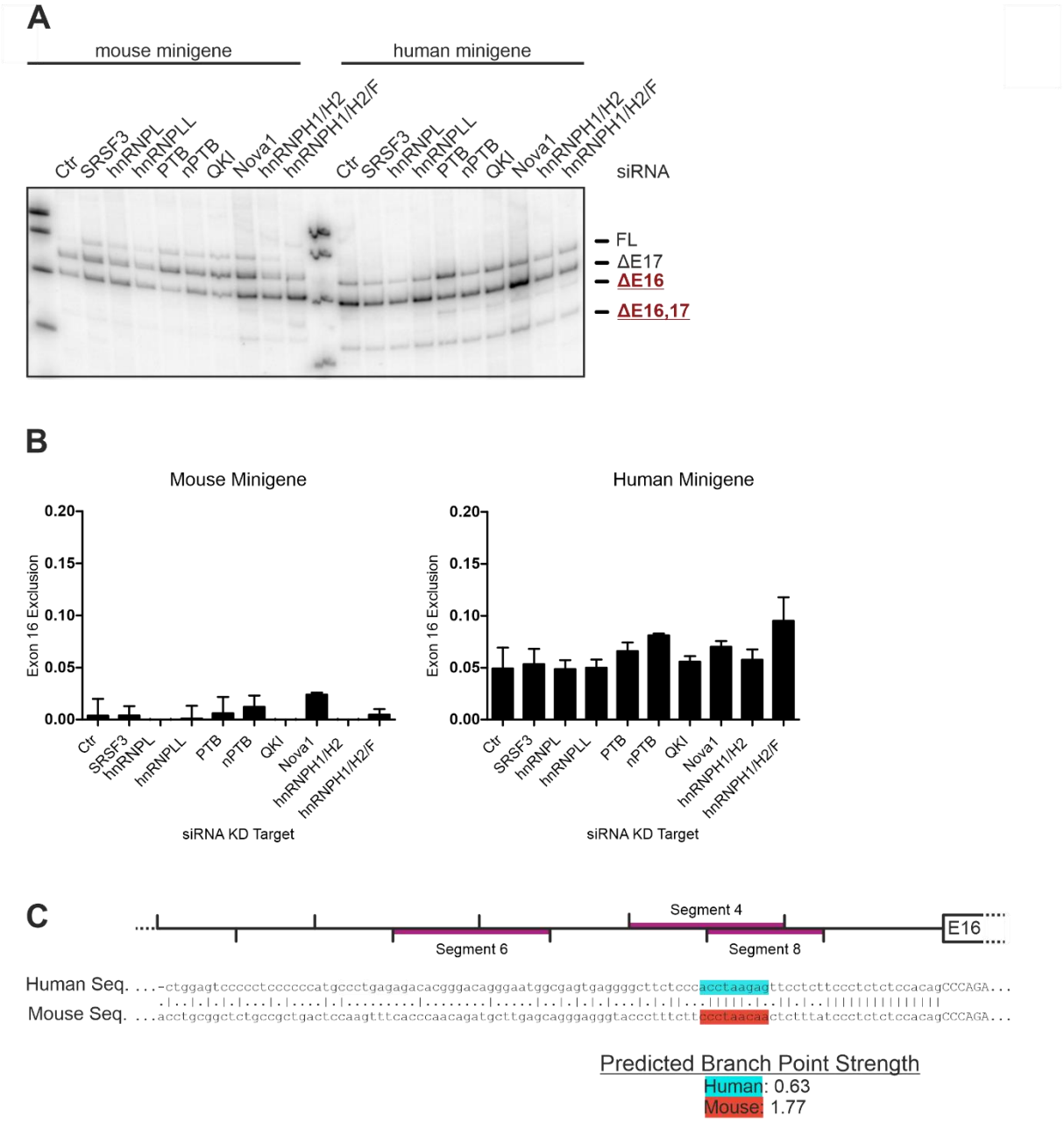



Figure 6: Identification of a functionally relevant branch point motif.

(A) siRNA KD of potential *trans*-acting factors regulating the species-specific AS of *Camk2β*. Indicated *trans*-acting factors were downregulated by siRNA-mediated KD in N2a cells. Human and mouse minigenes of *Camk2β* were *trans*fected and resulting splice isoforms identified by radioactive RT-PCR. Isoforms are indicated on the right and named according to the exons skipped. (B) Quantification of A, error bars indicate standard deviation (n=3). (C) Sequence alignment between human and mouse of the intron harbouring the identified *cis*-acting element. Purple lines highlight segments of functional relevance. Highlighted sequences indicate locations of predicted branch point sequences (SVM-BP) (Corvelo *et al.*, 2010), dark colours indicate stronger predicted BP strength.

Following these results, we turned our attention to other sequence elements present in the regulatory regions, and analysed the predicted splice site and branch point (BP) strengths. As expected from the strong conservation of the splice site-proximal

nucleotides, the predicted strength of the 3' splice site does not differ between the two species (human – ASSP: 10.15, SpliceRover: 0.9953; mouse – ASSP: 10.15, SpliceRover: 0.9991 (Wang and Marín, 2006; Zuallaert et al., 2018)). However, prediction of potential BP sequences (Corvelo *et al.*, 2010; Nazari *et al.*, 2019) revealed that both species harbour the most salient BP sequences in the overlap of segments 4 and 8 (Figure 6C). Strikingly, while this sequence resembles a near-optimal BP that lies within the AG dinucleotide exclusion zone (AGEZ) in the mouse intron, the corresponding human sequence scores much lower and lies slightly outside of the AGEZ. Including the other two species for which we have analysed the endogenous splicing pattern (rhesus macaque and African clawed frog), the predicted BP strength rates as such: mouse > frog > rhesus > human (Table 1) and correlates well with the observed splicing pattern. The predicted frog BP sequence scores lower than the corresponding mouse sequence, but an alternative BP is found in very close proximity. When added, the frog BP scores reach that of the mouse sequence. The predicted rhesus sequence scores higher than the corresponding human sequence, but considerably lower than the mouse sequence. Notably, in the RT-PCR, the rhesus macaque sample also shows a faint band for the Δ13,16 exclusion isoform for endogenous *Camk2β* (Figure 1C). Taken together, these results strongly suggest that the *cis*-regulatory element is not a conventional, splicing-factor bound enhancer or repressor motif, but the splicing differences are instead mediated by the evolution of the BP sequence. Interestingly, due to the challenges of experimentally determining BPs, the functional relevance of this motif in regulating alternative splicing has only recently been studied (Mercer et al., 2015).

Table 1. Predicted branch point scores for the intron upstream of *Camk2β* exon 16.

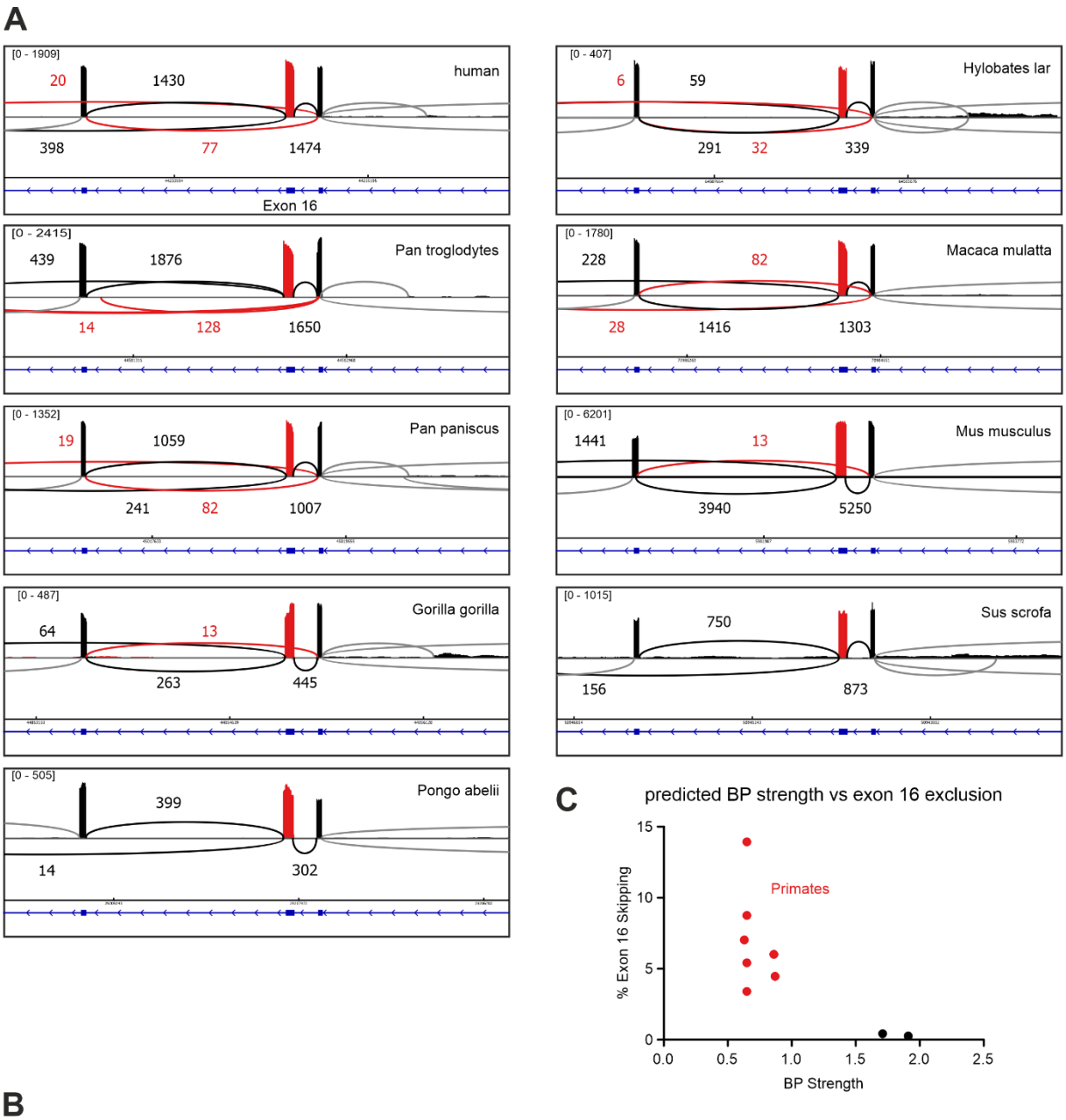
| Species | AGEZ | Distance | Sequence | Branch Point Score |
|---|---|---|---|---|
| Mouse | 44 | 26 | ccctaacaa | 1,77 |
| Frog | 18 | 20 | aactaagtc | 1,11 |
| Frog | 18 | 24 | ctttaacta | 0,74 |
| Rhesus | 24 | 26 | gcctaaggg | 0,83 |
| Human | 22 | 26 | acctaagag | 0,63 |

Table 1. Branch Point (BP) scores were calculates using SVM-BP (Corvelo *et al.*, 2010). AGEZ: AG dinucleotide exclusion zone, distance: distance to 3' splice site, sequence: sequence of identified BP, branch point score: predicted BP score (scaled vector model).

To confirm these findings, publicly available RNA-Seq data from different mammals was analysed, including species from which DNA or RNA samples were not obtainable during the course of this study. The focus was set on primates, in order to better understand the evolution of the *cis*-acting element. RNA-Seq data from cerebellar tissue or, where cerebellum data was not available, total brain tissue from different species was mapped to the corresponding genome (Figure 7A). Exon 16 and exon 16,17 exclusion isoforms could be confirmed in humans, even though they only amount to ~5-7% of all *Camk2β* transcripts. Exclusion of exon 16 was also observed in mouse tissue, but at a much lower frequency of ~0.4%. Even less exon 16 skipping could be observed in the more distant pig (*Sus scrofa*). All analysed primates show significant exon 16 skipping except for the orangutan (*Pongo abelii*). This might be due to low sequencing depth or quality, and exon 16 skipping could indeed be observed in testis data for this species (Figure S4A). There is some variation regarding the exact PSI levels among primates, but it does not strictly correlate with the evolutionary relationships. Alignment of the BP sequence shows a clear similarity between all primates, which is distinct from mouse and pig (Figure 7B). The latter two show a significantly higher BP score, which explains the observed splicing differences. The core of the BP motif seems conserved among primates, with only two nucleotides showing some variation. These variations correlate with the evolutionary relationship and result in slightly different predicted BP strengths. However, these differences do not correlate with the exon 16 exclusion levels observed in the RNA-Seq data (Figure 7C). Even though the exact PSI values differ between human, chimpanzee (*Pan troglodytes*), bonobo (*Pan paniscus*) and gorilla (*Gorilla gorilla*), their BP sequence is identical. A recent study on the expression of CaMKII in human hippocampus found exon 16-exclusion isoforms of *Camk2β* to range from ~4% to 16% between donors (Sloutsky *et al.*, 2020). This could suggest some variation in the exact interpretation of these regulatory elements and hints at additional regulatory layers that are specific to individuals or developmental stage.

We extended our analysis regarding the conservation of the BP sequence to include additional species (Figure S4B). All primates show a weak BP with a conserved consensus sequence. This also includes the order *Dermoptera*, the flying lemurs, which are the closest relatives of primates. All other species feature a strong BP motif, that shows a medium degree of sequence conservation among most mammals. The sequences diverge with increased evolutionary distance, but the high BP strength is

maintained. Notable exceptions like the *Anolis carolinensis* lizard feature intron sequences that do not return any valid, predicted BPs in close proximity to the splice site, suggesting fundamental differences in the interpretation of these core elements. Taken together, the RNA-Seq data confirms the radioactive PCRs and minigene splicing assays, revealing species-specific differences in the canonical branch point sequence. Comparing different mammals suggests that this feature has emerged during primate evolution.




B

| Species | Intron | Branch Point | Exon | Branch Point Score |
|---|---|---|---|---|
| Sus scrofa | ccctagtctctcccctaatggcttctcttcctctctctgcag | | CCCCAGACGAATAGCACCA | 1.93 |
| Mus musculus | acccttttcttcccctaacaactctttatccctctctccacag | | CCCCAGACAAACAGCACCA | 1.77 |
| Macacca mulatta | ggcttctcccgcctaagggttcctcttccctctctccacag | | CCCCAGACGAATAGCACCA | 0.83 |
| Hylobates moloch | ggcttctcccgcctaagggttcctcttccctctctccacag | | CCCCAGACGAATAGCACCA | 0.83 |
| Pongo abelii | ggcttctcccacctaagggttcctcttccctctctccacag | | CCCCAGACGAATAGCACCA | 0.54 |
| Gorilla gorilla | ggcttctcccacctaagagttcctcttccctctctccacag | | CCCCAGACGAATAGCACCA | 0.63 |
| Homo sapiens | ggcttctcccacctaagagttcctcttccctctctccacag | | CCCCAGACGAATAGCACCA | 0.63 |
| Pan troglodytes | ggcttctcccacctaagagttcctcttccctctctccacag | | CCCCAGACGAATAGCACCA | 0.63 |
| Pan paniscus | ggcttctcccacctaagagttcctcttccctctctccacag | | CCCCAGACGAATAGCACCA | 0.63 |

Figure 7: RNA-Seq analysis reveals evolution of *Camk2β* exon 16 alternative splicing.

(A) Sashimi plot showing the alternative splicing of *Camk2β* exon 16 in RNA-Seq data from human, chimpanzee (*Pan troglodytes*), bonobo (*Pan paniscus*), gorilla (*Gorilla gorilla*), orangutan (*Pongo abelii*), gibbon (*Hylobates lar*), rhesus macaque (*Macaca mulatta*), mouse (*Mus musculus*) and wild pig (*Sus scrofa*). RNA-Seq data from cerebellum was used for all species, except orangutan, for which RNA-Seq data from total brain tissue was used. Red colour indicates exon 16 and exon 16 exclusion reads. Numbers indicate number of reads per splice junction, with the minimum set to 3 junction reads. Shown in blue is the intron/ exon-structure of the displayed region. (B) Alignment of the

identified functionally relevant branch point (BP) sequence. Nucleotides that belong to the predicted BP sequence are colour coded. Lower case letters indicate intronic sequence, upper case letters indicate exonic sequence. The BP score was predicted using SVM-BPfinder (Corvelo *et al.*, 2010) with the human BP model. BP score refers to the scaled vector model. (C) The predicted BP strength was plotted against the experimentally determined exon 16 exclusion levels. BP Strength was predicted using SVM-BPfinder (Corvelo *et al.*, 2010) with the human model for the human sequence, the *Pan troglodytes* model for chimpanzee, bonobo, gorilla and gibbon, the rhesus model for rhesus macaque, the mouse model for mouse and the *Bos taurus* model for pig. Exon 16 exclusion levels were calculated from RNA Seq data using rMATS (Shen *et al.*, 2014).

## *In vitro* kinase assay reveals minor kinetic differences

Having established the genomic causes and the transcriptomic consequences of the species-specific alternative splicing of *Camk2β*, we set out to determine its effect on protein level. We selected two species-specific isoforms (Δ16,17 and Δ13,16) as well as two control isoforms (FL and Δ13) for expression and purification from insect cells (Figure 8A). Again, the full-length (FL) isoform refers to the longest detected isoform in cerebellum and lacks exons 19 to 21 (Figure 1A,C). These four isoforms were tested in a radioactive *in vitro* kinase assay with the model substate Syntide 2 (Hashimoto and Soderling, 1987), linked to GST (Figure 8B, C). Activity was monitored as a function of calmodulin concentration, to test the cooperativity of the enzyme. Consistent with a recent publication (Sloutsky *et al.*, 2020), we did not observe major differences in the $EC_{50}$ value or the Hill coefficient (Table 2) between the four CaMKIIβ variants. Instead, we observe small but significant differences in the maximal activity ($V_{max}$) reached at optimal calmodulin concentrations (Figure 8D). At concentrations of 100-1000 nM calmodulin, both species-specific protein isoforms reach a slightly higher maximal activity, compared the FL and Δ13 isoforms. The same effect was also observed using human full-length tau protein (tau 441) as an alternative protein substrate (Figure S5A, B).

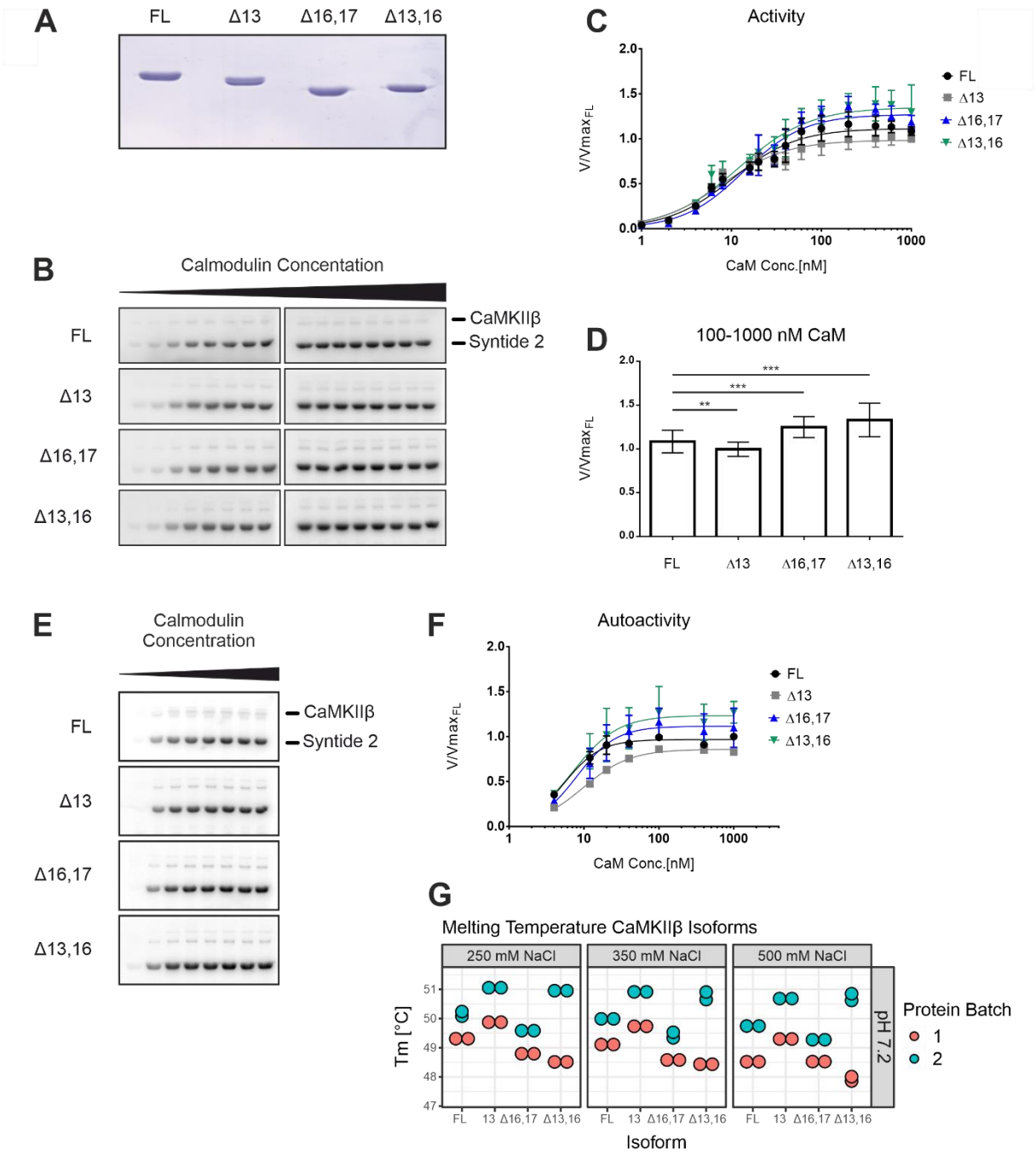



Figure 8: *In vitro* kinase assay reveals minor kinetic differences.

(A) SDS-PAGE of purified CaMKIIβ isoforms. Proteins were expressed in insect cells and purified via Strep-affinity and size exclusion chromatography. Protein concentration was determined via UV-absorption at 280 nm and precisely levelled by repeated SDS-PAGE, Coomassie-staining and subsequent quantification. (B) *In vitro* kinase assay with different CaMKIIβ isoforms. CaMKII activity against a protein substrate (Syntide 2, fused to GST) was measured as a function of calmodulin concentration. Direct phosphorylation of the substrate by CaMKIIβ was measured via $^{32}P$ incorporation. Samples were separated on an SDS-PAGE and detected using autoradiography. (C, D) Quantification of B, normalized to the maximum activity of the FL isoform (n = 6). Error bars indicate standard deviation. (D) Samples at approximate maximal activity were combined. Error bars indicate standard deviation. *p<0.05, **p<0.01, ***p<0.001 calculated by Student's or Welch's t-test. Adjusted for multiple comparisons using Holm's method. (E) *In vitro* kinase assay with different CaMKIIβ isoforms to test the autoactivity generated at increasing calmodulin concentrations. CaMKII was first activated with calmodulin in the presence of ATP. EGTA was added to chelate calcium and quench the binding of calmodulin. Addition of a protein substrate (Syntide 2, fused to GST) enabled detection of generated autoactivity via $^{32}P$ incorporation. Samples were separated on an SDS-PAGE and

detected using autoradiography. (F) Quantification of E, normalized to the maximum activity of the FL isoform (n=3). (G) Melting temperature of purified CaMKIIβ isoforms under different buffer conditions, as determined by a thermofluor assay. Tm: melting temperature in °C. Colour indicates protein batch.

One of the key properties of CaMKII is the ability to adopt different activation states, based on its own phosphorylation pattern (Bayer and Schulman, 2019). Upon stimulation, the enzyme quickly *trans*-autophosphorylates on T287 and adopts an auto-active state, that persists even in the absence of calcium/calmodulin. Recent studies suggest that the rate at which certain activating and inhibiting phosphorylation are acquired differs between CaMKII genes and might also be influenced by the length and composition of the variable linker segment (Bhattacharyya *et al.*, 2020). We thus tested the autoactivity of the selected CaMKIIβ isoforms, generated at varying calmodulin concentrations (Figure 8E). This should directly reflect the activation and hence phosphorylation state of the enzyme. CaMKII was first stimulated in the absence of the substrate protein, after which calcium was quenched by supplementing EGTA. Addition of the substrate Syntide-2-GST allowed for specific measurement of the autoactivity previously generated. Similar differences regarding $V_{max}$ could be observed, with the species-specific protein isoforms reaching slightly higher maximal activities (Figure 8F). We also performed the assay in the presence of an increased salt concentration, as CaMKII proved to require high salt concentrations during purification (Figure S5D, E). Interestingly, under these conditions the $V_{max}$ differences essentially disappear, suggesting that they depend on the exact reaction conditions. Together, these experiments confirm that the tested CaMKIIβ splice isoforms do not differ in their $EC_{50}$ values or Hill coefficients. Instead, a slight difference in maximal activity at optimal calcium/calmodulin concentrations can be observed.

Table 2. Kinetic parameters of purified CaMKIIβ isoforms.

| | Substrate: Syntide 2 | | | | Substrate: Tau-441 | |
|---|---|---|---|---|---|---|
| | Isoform | | | | Isoform | |
| | FL | Δ13 | Δ16,17 | Δ13,16 | FL | Δ16,17 |
| $V_{max}$ | 1.11 ± 0.02 | 0.99 ± 0.02 | 1.27 ± 0.03 | 1.35 ± 0.03 | 1.04 ± 0.02 | 1.25 ± 0.03 |
| h | 1.22 ± 0.11 | 1.20 ± 0.09 | 1.29 ± 0.11 | 1.11 ± 0.11 | 1.53 ± 0.19 | 1.58 ± 0.17 |
| $EC_{50}$ | 10.30 ± 0.82 | 7.80 ± 0.56 | 13.65 ± 0.97 | 11.98 ± 1.16 | 12.2 ± 0.90 | 18.8 ± 1.30 |

Table 2. Kinetic parameters as determined by in vitro kinase assay and subsequent fitting of a Hill equation (Figure 8B,C; Figure S5A,B). h: Hill Coefficient.

To test whether these characteristics are caused by fluctuations in protein stability, a thermofluor assay was performed using two independently expressed and purified protein batches (Figure S6A, B). Although the absolute melting temperatures differ between the two batches, the relative difference between the CaMKIIβ isoforms remains similar, except for isoform Δ13,16 (Figure 8G, S6A). Notably, these small variations (0.5-1°C melting temperature) in protein stability do not correlate with the observed activity differences in the *in vitro* kinase assays. All CaMKIIβ isoforms show a similar dependency on pH and salt, with the protein being stabilized at higher pH and lower salt concentrations (Figure S6B, C). Interestingly, this contrasts with the behaviour of CaMKIIβ during protein purification, which requires high NaCl concentrations. In summary, these variations in protein stability are comparably small and fail to explain the observed activity differences of the enzyme variants.

## CaMKII is a temperature-dependent kinase

We also tested the hypothesis that CaMKII might be a temperature-sensitive kinase. A recent study from our research group showed that protein kinases can be temperature-dependent, exhibiting higher enzymatic activity at lower temperatures (Haltenhof et al., 2020). Such temperature fluctuations are linked to the circadian rhythm and dynamically change over the course of the day. We adapted our *in vitro* kinase assay and performed it under saturated calcium/calmodulin concentrations at varying temperatures (Figure 9A, B). CaMKIIβ shows a strong, near-linear decrease of activity at increasing temperatures, with a 2-fold activity change within a physiologically relevant temperature range for humans (36-38°C). No difference between the CaMKIIβ protein isoforms can be observed, indicating that the variable linker segment does not regulate the temperature sensitivity of the enzyme. As no temperature sensitivity had previously been reported for any CaMKII gene, we next tested whether the observed effect is specific to the β homologue. As no difference could be observed between the CaMKIIβ protein isoforms, we restricted our analysis to the kinase domains. We expressed and purified the kinase domains of all four human CaMKII genes (α, β, γ, δ) and tested them in our *in vitro* kinase assay (Figure 9C, D). When analysing a similar temperature range as before, the kinase domains also show a near-linear temperature dependency, comparable to full-length CaMKIIβ (Figure 9D). When the temperature range is extended, differences become visible at lower and higher temperatures (Figure 9E). The δ isoform seems to be more stable, retaining about half of its maximal activity at 40°C. Together, this establishes CaMKII as a temperature-sensitive kinase that exhibits increased kinetic activity at lower temperatures. It provides a potential link between body temperature, circadian rhythm and synaptic plasticity. This feature is an inherent property of the kinase domain itself and conserved among all four human CaMKII genes.

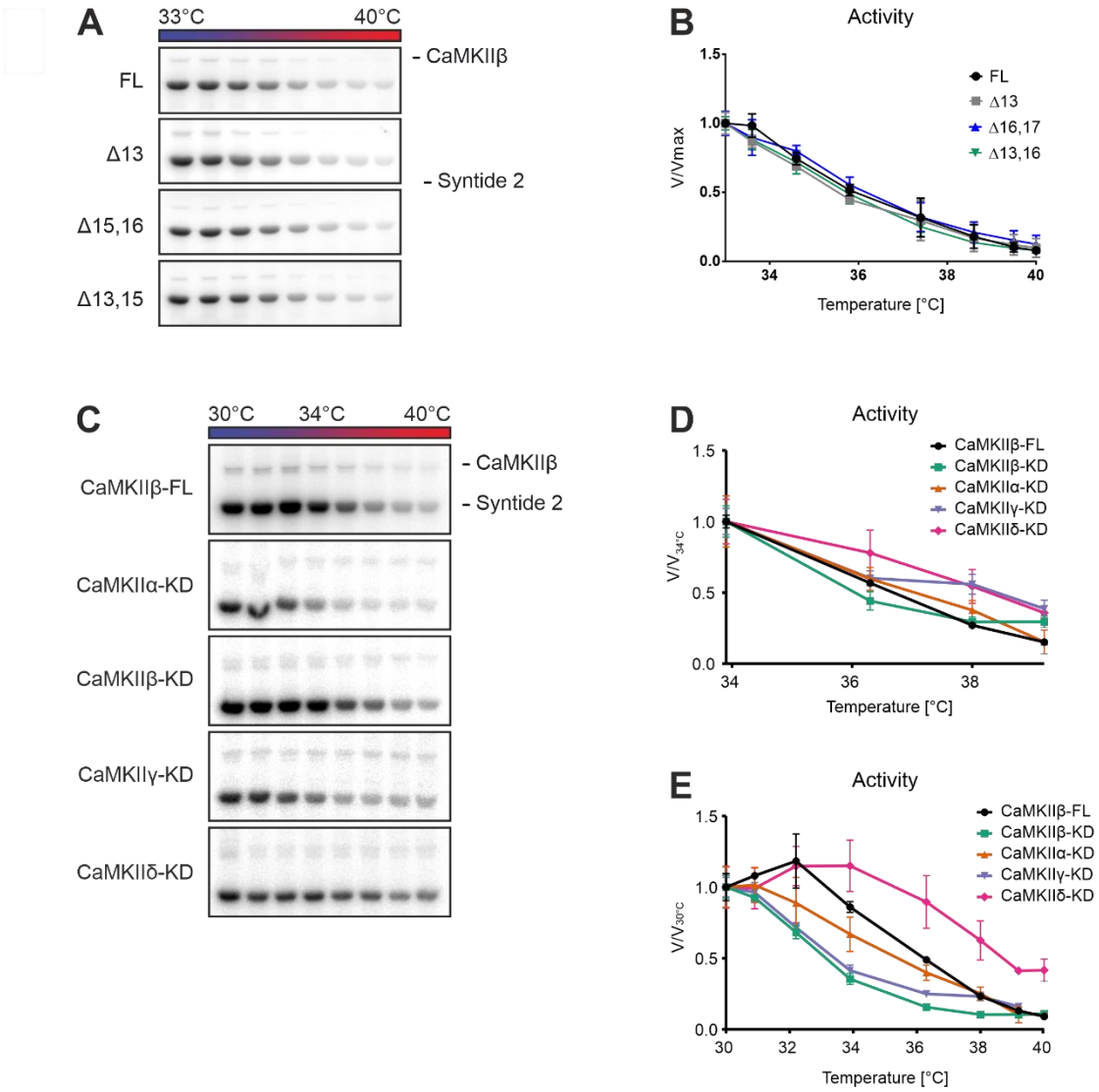



Figure 9: CaMKII is a temperature-dependent kinase.

(A, C) *In vitro* kinase assay with different CaMKIIβ isoforms (A) or different CaMKII kinase domains (C). CaMKIIβ isoforms were expressed in insect cells and purified via Strep-affinity and size exclusion chromatography. CaMKIIα, β, γ, δ-kinase domains were expressed in bacteria and purified via Strep-affinity chromatography. CaMKII activity against a protein substrate (Syntide 2, fused to GST) was measured at different reaction temperatures. Direct phosphorylation of the substrate by CaMKII was measured via $^{32}P$ incorporation. Samples were separated on an SDS-PAGE and detected using autoradiography. (B, D E) Quantification of A (B) and C (D, E), normalized to the activity at the lowest temperature (n = 3). Error bars indicate standard deviation.

## cryoEM reveals diversity of oligomeric states

Selected, purified CaMKIIβ isoforms (Figure 8A) were analysed using cryogenic electron microscopy (cryoEM). Datasets from the FL and Δ16,17 isoforms were collected and 2D classification revealed an unexpected diversity of oligomeric states (Figure 10A). Only the central oligomerization hub of each holoenzyme is visible, as the peripheral kinase domains are flexible and thus disappear in the 2D classes and final 3D reconstructions. Whereas for CaMKIIα, only 12-mer and 14-mer assemblies have been reported (Myers *et al.*, 2017; Sloutsky *et al.*, 2020), we also see a 16-mer and an open 12-mer state in our data (Figure 10A). The equilibrium between these states does not significantly differ between the two CaMKII isoforms, except for the 16-mer state, which was not detected in the dataset of the FL variant. However, as this dataset was significantly smaller and the 16-mer state is the least populated (0.4%), it cannot be excluded that this state might have gone unnoticed for the FL variant. Most holoenzymes consist of 12 subunits (dodecameric), followed by 14 subunit holoenzymes (tetradecameric). These results are similar to a negative-stain analysis of purified FL CaMKIIβ (also lacking exons 19 to 21), where 93% of detected holoenzymes adopted a dodecameric state (Buonarati *et al.*, 2021). However, they found significantly less 14-mers compared to our data, and did not detect the open 12-mer state.

Use of the detergent LMNG helped to overcome the problem of preferred orientations, evident by the detection of side views (Figure 10A, S8B). It enabled the structural elucidation of all states (Figure S7). The 12-mer further separated into a symmetric and an asymmetric state, which are mostly identical (Figure 10B, S7B). When aligned, both structures have a root mean square deviation (RMSD) of 1.15 Å. If outliers are removed, this number is decreased to 0.40 Å. The structures mainly differ in loop regions and at the N-terminus, where the density is less well-resolved. In the asymmetric model, one subunit features an extended N-terminus, with 3 additionally resolved residues (Figure 10C). When compared to an existing structure of the CaMKIIα hub (McSpadden *et al.*, 2019), our model for the symmetric 12-mer shows high structural similarity (Figure S8C). The overlayed structures have a RMSD of 1.53 Å, which can be expected for the high level of structural conservation between the α and β variant. We are the first to report a model for an open dodecameric structure and existence of such a state has been proposed to facilitate subunit exchange between

holoenzymes (Bhattacharyya *et al.*, 2016; Stratton *et al.*, 2014). In contrast to the proposed model of a spiral form, our open conformation shows an in-plane movement of the individual subunits relative to each other (Figure 10D,E). A gap between the subunits is visible when the map is contoured at 9 RMSD, which partially disappears at a contour level of 6 RMSD (Figure S9A). The refined model of the open 12-mer shows variability in inter-subunit distances, with a significantly increased gap between two neighbouring subunits. There is also variability within individual subunits, although this is more evenly distributed across the holoenzyme (Figure S9B). All distances are similar to the closed (symmetric) 12-mer, suggesting that the open conformation might be an intermediate state and not a fully open holoenzyme (Figure S9C). The increased gap partially disappears when contouring the map at 6 RMSD, due to a curvature in the N-terminal helix of one of the contributing subunits (Figure S9D). This helix appears straight in most other subunits (Figure S9E). Together, our results reveal a high diversity of oligomeric states for CaMKIIβ, which does not differ between the splice isoforms. Similar to previous studies, we identify the dodecamer as the main state, but additionally find a high number of tetradecameric holoenzymes. A novel open-dodecameric structure could be resolved, which was previously postulated to facilitate in subunit exchange between individual holoenzymes.

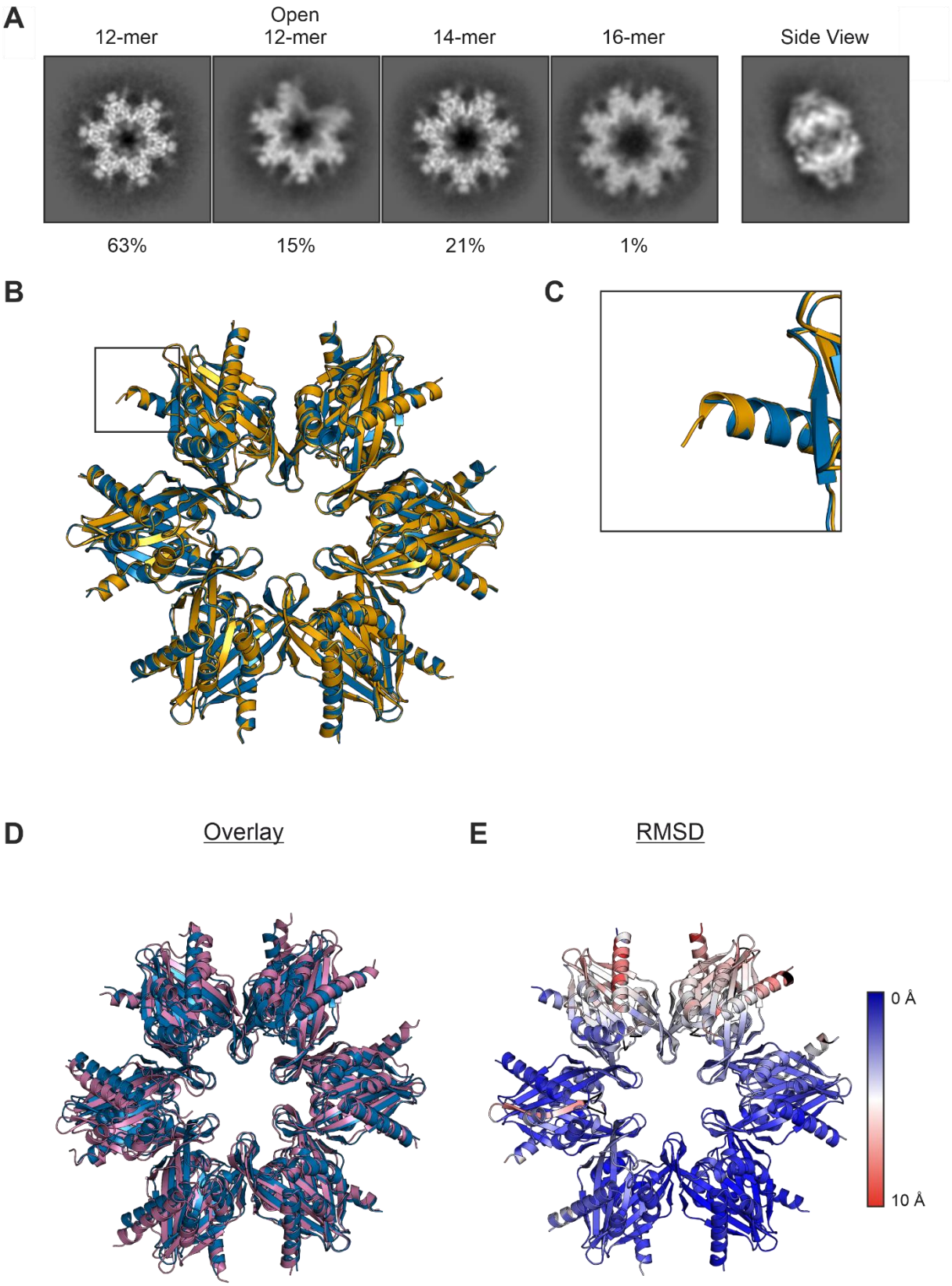



Figure 10: cryoEM analysis of CaMKIIβ.

(A) 2D class averages of the CaMKIIβΔ16,17 hub. CaMKIIβ was expressed in insect cells and purified via Strep-affinity and size exclusion chromatography and subjected to single particle cryoEM. 2D classes were generated using cryoSPARC (Punjani *et al.*, 2017). The numbers represent the ratio between the individual states, as determined by the number of particles in the respective 2D classes. (B) Final structure of the closed 12-mer (dodecamer) of the CaMKIIβ hub. Displayed is an overlay of

the symmetric (blue) and asymmetric (yellow) model. The black box highlights the segment enlarged in C (C) Magnified view showing the main structural difference between the two closed 12-mer states. (D) Overlay of the closed, symmetric 12-mer state (blue) and the open 12-mer state (magenta). (E) Structure of the closed symmetric 12-mer state, coloured based on the root mean square deviation (RMSD) between the Cα atoms of the closed symmetric 12-mer and open 12-mer state. Scale bare indicates colouring scheme, with black symbolising residues not used for alignment.

## CaMKIIβ isoforms have different substrate spectra

### Characterization of an analog-sensitive kinase mutant of CaMKIIβ

Having found only minor kinetic variations between the CaMKIIβ isoforms, we set out to test whether they show differences in their substrate spectra. As CaMKII is a multifunctional kinase, involved in numerous signalling pathways, the list of known and putative phosphorylation targets is long. Testing individual substrates *in vitro* is thus impractical and has previously only been realized for fly CaMKII (GuptaRoy *et al.*, 2000). Such an approach also fails to account for target specificity generated by distinct subcellular localization or secondary protein-protein interactions. We instead adopted the analog-sensitive kinase system (Lopez et al., 2014), which allows for the direct labelling of kinase substrates, followed by specific enrichment and subsequent detection by mass spectrometry. This method is unbiased and does not require prior knowledge of potential kinase targets.

An analog-sensitive mutant had previously been described for CaMKIIα (Wang *et al.*, 2003). Due to the high sequence similarity of the kinase domain between the α and β isoforms, we hypothesized that the homologous mutation should also be functional in CaMKIIβ. The mutation (phenylalanine 89 to glycine) affects the ATP binding pocket, enlarging it and enabling the enzyme to utilize ATP analogs with bulky side chains on their $N^6$-atom (Lopez *et al.*, 2014). When expressed in HEK cells, clear bands for all four selected CaMKIIβ variants could be seen on a Western blot (Figure 11A). All isoforms show the typical triple-band pattern caused by the Twin-Strep-tag, which can also be observed for the uncleaved wild type protein (data not shown). The Δ13,16 isoform was selected for *in vitro* characterization and purified via Strep-affinity chromatography (Figure 11B). The analog-sensitive mutant was tested in an *in vitro* kinase assay with the model substrate Syntide-2-GST and shows a calmodulin dependency similar to the wild type protein (Figure 11C,D). Whereas the exact $EC_{50}$ value and Hill coefficient cannot be reliably determined due to the low concentration, the activity profile shows the expected sigmoidal shape and reaches a plateau at around 100 nM CaM. We next tested whether the enzyme could be competitively inhibited by an excess of non-radioactive unmodified ATP, or ATP analogs (Figure 11E,F). The wild type protein can be efficiently inhibited by an excess of non-radioactive ATP, whereas the mutant retains most of its activity. This is somewhat unexpected, as the non-labelled and radioactively labelled ATP are chemically

identical and should be used with equal efficiency by both enzymes. The inhibition with selected ATP analogs shows a clear difference between wild type and mutant CaMKIIβ. Whereas the wild type retains almost maximal activity, the mutant shows a strong inhibition with two of the tested ATP analogs. These two, $N^6$-benzyl- and $N^6$-phenyl-ATP, are structurally similar and distinct from the much smaller $N^6$-methyl-ATP or much larger $N^6$-etheno-ATP. Importantly, the ATP analogs were used at a concentration where the wild type kinase is almost completely inhibited by the same concentration of unmodified ATP, proving that the wild type-kinase does not efficiently utilize the tested ATP analogs. Together, these experiments show that the analog-sensitive kinase mutant of CaMKIIβ can be successfully overexpressed and purified from eukaryotic cells. Its kinetic properties are similar to those of the wild type kinase, except for the specific inhibition by $N^6$-modified ATP analogs.

The enrichment of peptides labelled in the analog-sensitive kinase assay and their detection via mass spectrometry requires the use of ATPγS, which contains a γ-terminal thiophosphate group. The mutant and wild type CaMKII were again tested in an *in vitro* kinase assay against the model substrate Syntide-2-GST, where radioactive ATP was replaced with $N^6$-benzyl-ATPγS or $N^6$-phenyl-ATPγS. This enables direct visualization of the substrate phosphorylation by Western blotting with a thiophosphate ester-specific antibody (Figure 11G). Whereas the wild type-kinase shows no signal for either ATP analog, the analog-sensitive mutant shows a strong signal for the protein substrate and a weaker signal for CaMKII autophosphorylation. Both ATP analogs are functional, as are both tested reaction buffers. No signal can be observed for the high salt buffer for $N^6$-Benzyl-ATPγS, which is most likely due to a blotting error. The signal depends on the addition of the alkylation agent PNBM (*p*-nitrobenzyl mesylate), as the antibody is specific to the thiophosphate *ester* epitope. Together, this shows the efficient utilization of $N^6$-benzyl/phenyl-ATPγS by the analog-sensitive kinase mutant of CaMKIIβ, under varying reaction conditions.

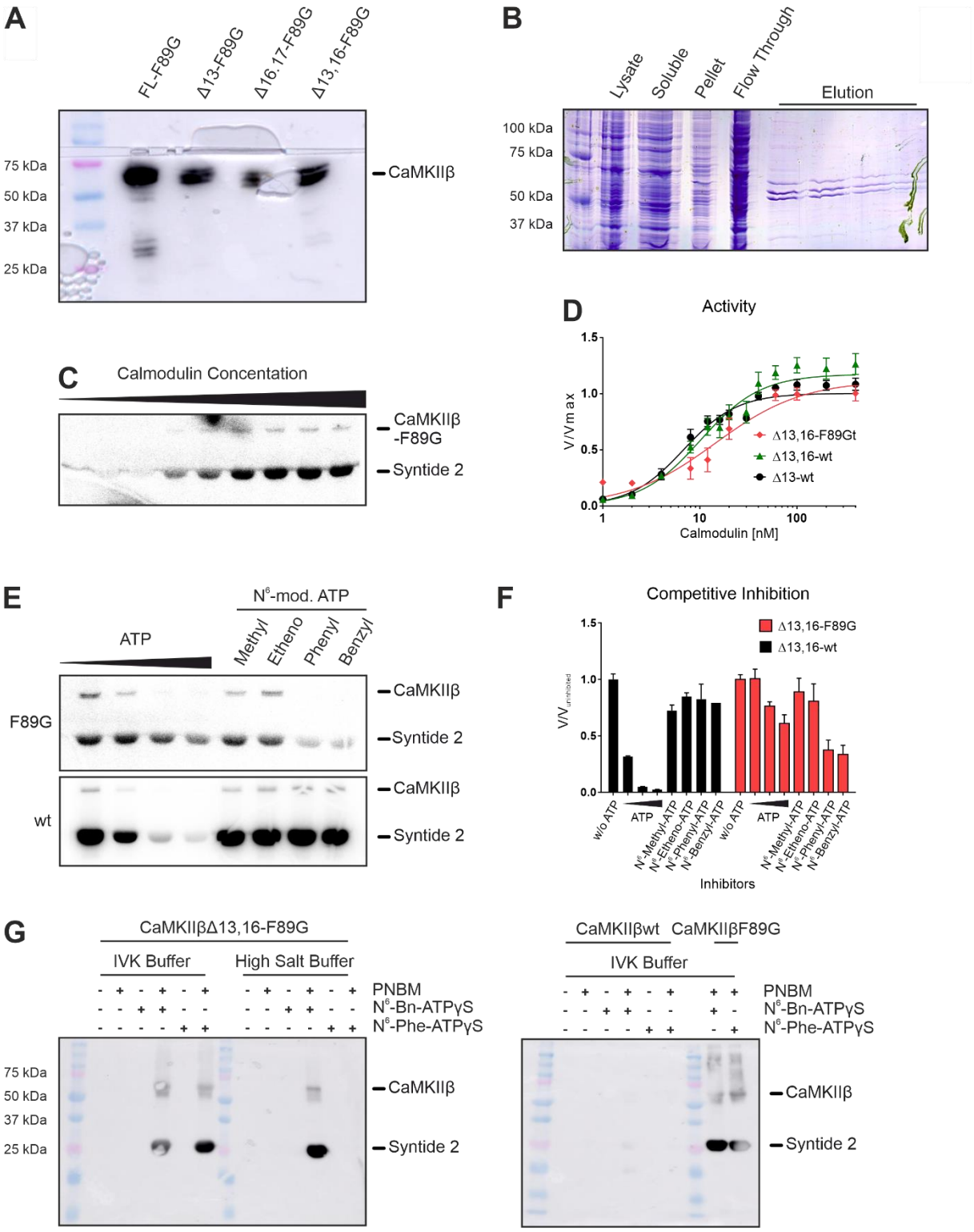



Figure 11: Functional characterization of a CaMKIIβ analog-sensitive kinase mutant.
(A) Western blot showing the expression of analog-sensitive kinase mutants (F89G) of selected CaMKIIIβ isoforms in HEK cells. (B) SDS-PAGE monitoring the purification of the AS mutant Twin-Strep-CaMKIIIβΔ13,16-F89G. The protein was expressed in HEK cells and purified via Strep-affinity chromatography. As the tag was not cleaved, the purified protein shows a triple-band pattern (C) *In vitro* kinase assay with the purified AS mutant of CaMKIIβ. CaMKII activity against a protein substrate (Syntide 2, fused to GST) was measured as a function of calmodulin concentration. Direct phosphorylation of the substrate by CaMKIIβ was measured via $^{32}P$ incorporation. Samples were separated on an SDS-PAGE and detected using autoradiography. (D) Quantification of C, normalized to the maximum activity at 400 nM calmodulin. Data for the wt variant of CaMKIIβΔ13 and

CaMKIIβΔ13,16 are taken from Figure 4B,C. Error bars indicate standard deviation. (E) Inhibition of the AS mutant of CaMKIIβ by various $N^6$-modified ATP analogs. An *in vitro* kinase assay was performed at an optimal calmodulin concentration and supplemented with different non-radioactive ATP analogs. Enzymatic activity was measured as described for C. (F) Quantification of E, normalized to the un-inhibited signal. (G) Western blot showing the active usage of $N^6$-modified ATPγS analogs by the AS mutant of CaMKIIβ. An *in vitro* kinase assay was performed in the presence of $N^6$-Bn-ATPγS ($N^6$-benzyl-ATPγS) or $N^6$-Phe-ATPγS ($N^6$-phenyl-ATPγS). Thiophosphorylated protein substrate (Syntide 2, linked to GST) was alkylated with PNBM and detected via standard western blotting using a thiophosphate ester-specific antibody. Two reaction buffers were tested, the standard reaction buffer for the radioactive in vitro kinase assay, as well as a high salt reaction buffer.

## Optimization of labelling conditions *in vivo*

The goal was to perform the assay in cells, as the subcellular localization is known to differ between CaMKII genes (Shen *et al.*, 1998) and splice isoforms (O'Leary *et al.*, 2006) and might impact the substrate specificity of the enzyme. Previous attempts to perform the analog-sensitive kinase assay *in vivo* have used the saponin detergent digitonin (Banko et al., 2011; Barilari et al., 2017; Maas et al., 2014; Michowski *et al.*, 2020) to permeabilize the plasma membrane and deliver the $N^6$-modified ATPγS analogs. Digitonin has the disadvantage of drastically destabilizing ATP (Chaudhary et al., 2002), leading to the generation of free thiophosphate. Incorporated into unmodified ATPγS by metabolic enzymes, it can be used by endogenous kinases and strongly increases the background noise. It also rapidly depletes the supplied pool of $N^6$-modified ATPγS and thereby the required co-substrate for the kinase of interest. We therefore tested alternative lysis and permeabilization methods, such as treatment with the detergent Tween-20. The large pores formed by this detergent enable diffusion of extracellularly supplied calmodulin and thus efficient activation of the enzyme. Overexpression of the analog-sensitive mutant leads to a strong phosphorylation of the substrate Syntide-2-GST (Figure 12A). From the tested methods, permeabilization by Tween-20 led to the most efficient phosphorylation of the substrate. Both ATP analogs show a similar efficiency under low ATP/ GTP conditions (I), whereas $N^6$-benzyl-ATPγS is more efficiently utilized under high ATP/ GTP conditions (II). Detection of the recombinant protein substrate Syntide-2-GST indicates pores large enough to accommodate the uptake of certain proteins, as this substrate was added extracellularly. Together, these experiments show that the permeabilization with Tween-20 can deliver both ATP analogs and extracellularly added proteins into eukaryotic cells, to efficiently label CaMKIIβ substrates *in vivo*. Thus, permeabilization with Tween-20 was selected for subsequent assays.

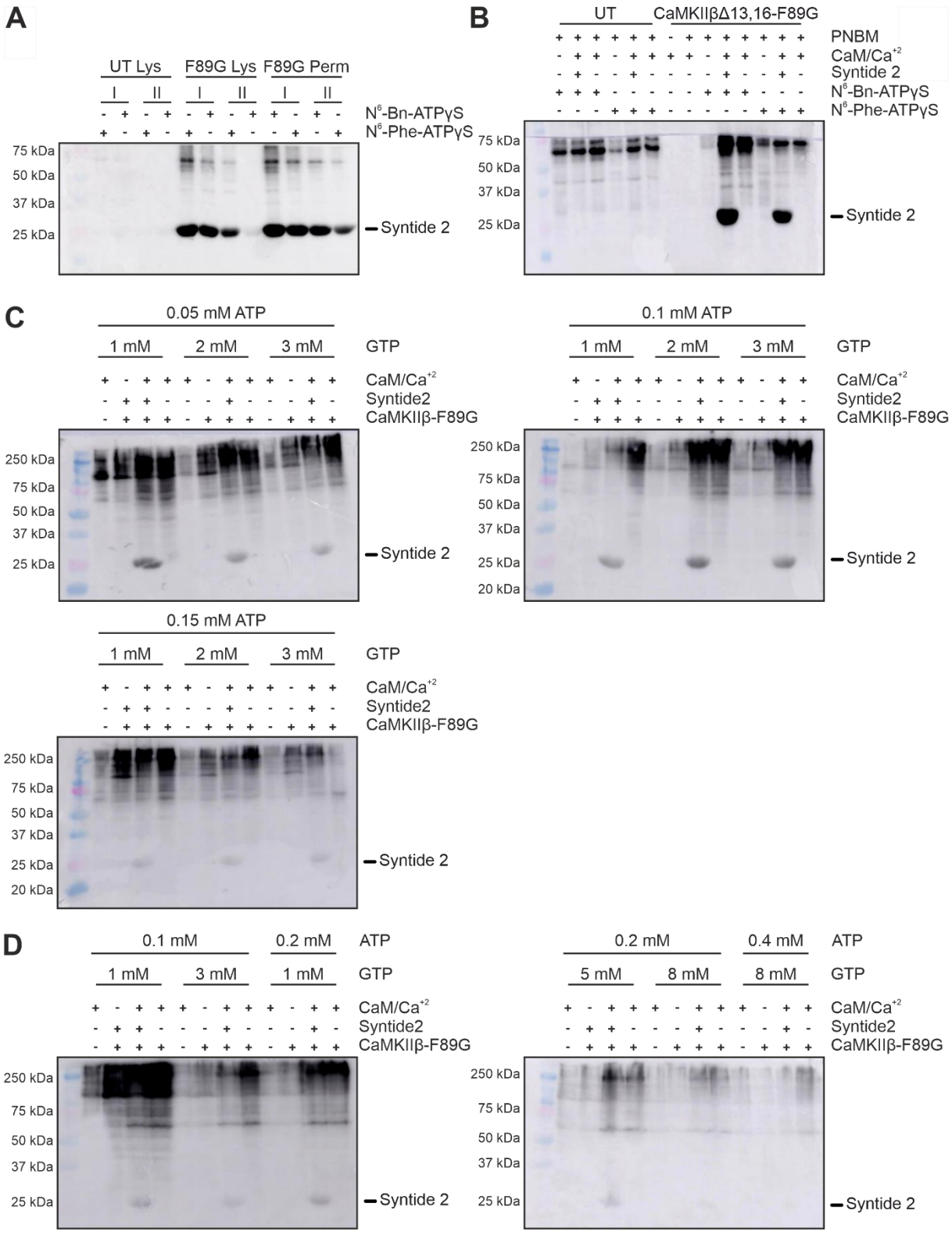



Figure 12: Optimization of *in vivo* labelling conditions for the analog-sensitive kinase assay.

(A) Comparison of cell lysis methods. N2a cells overexpressing Twin-Strep-CaMKIIIβΔ13,16-F89G were collected and either permeabilized with 0.025% (v/v) Tween-20 (Perm) or lysed with 0.5% (v/v) NP-40, 0.5% (v/v) Triton X-100 and sonication (Lys). Reactions were performed with either one of two $N^6$-modified ATPγS analogs (benzyl, phenyl), alkylated with PNBM, run on an SDS-PAGE and analysed via western blot with a thiophosphate ester-specific antibody. Two reaction conditions were tested, I: 0.1 mM ATP, 2 mM GTP; II: 0.2 mM ATP, 5 mM GTP. Upper band: cellular proteins, lower

band: external substrate (Syntide 2, linked to GST). UT: untransfected control, F89G: analog-sensitive kinase mutant overexpressed. (B) Further comparison of $N^6$-benzyl- vs $N^6$-phenyl-ATPγS, testing the specificity of the labelling reaction in the presence/ absence of stimulating conditions ($Ca^{2+}$/CaM) or the external substrate Syntide 2. Reactions were performed in Tween-20 permeabilized N2a cells overexpressing Twin-Strep-CaMKIIIβΔ13,16-F89G. (C, D) Systematic optimization of *in vivo* labelling conditions for Twin-Strep-CaMKIIIβΔ13,16-F89G with $N^6$-Benzyl-ATPγS in permeabilized N2a cells. Varying concentrations of ATP and GTP were tested as indicated. Reactions were performed in the presence/ absence of stimulating conditions (CaM/$Ca^{2+}$) or an external substrate (Syntide 2, linked to GST).

When including untransfected and unstimulated controls, a relatively strong signal can be observed around the 70 kDa mark for both conditions in a Western blot detected with the thiophosphate ester-specific antibody (Figure 12B). As no signal for the protein substrate Syntide-2-GST can be observed in these controls, it is unlikely that this background signal originates from endogenous CaMKII present in N2a cells. Activating conditions are essential for transfected cells to observe an increased phosphorylation, both for the supplied protein substrate and cellular proteins. Again, a stronger signal can be observed for $N^6$-benzyl-ATPγS. The controls lacking an $N^6$-modified ATPγS analog or the alkylation agent PNBM show that the specificity of the thiophosphate ester antibody is maintained in the context of a more complex protein sample. The background signal in the samples from untransfected cells is expected, as endogenous kinases can also utilize ATP analogs when provided in high concentrations. Together with the previous results, this indicates $N^6$-benyzl-ATPγS as an ideal analog for further assays, but also highlights the need to further optimize the precise labelling conditions.

To screen different reaction conditions, the concentration of $N^6$-benzyl-ATPγS was kept constant, while the concentrations of ATP and GTP were varied (Lopez *et al.*, 2014) (Figure 12C). The overall signal intensity decreases with both increasing ATP and GTP concentrations. Importantly, the difference between untransfected or unstimulated cells and transfected and stimulated cells becomes larger at increasing ATP/ GTP concentrations. Increased ATP and GTP concentrations mostly prevent the background noise generated by the endogenous kinases using the supplied $N^6$-benzyl-ATPγS. At very high concentrations, these differences largely disappear, as most likely all kinases get efficiently inhibited, including the overexpressed analog-sensitive mutant. The optimal conditions, chosen based on a large difference between control and stimulated samples, and a strong overall signal intensity, are at concentrations revolving around 0.1 mM ATP and 2-3 mM GTP. This was confirmed by further extending the range of tested concentrations (Figure 12D). Almost no signal can be

observed for GTP concentrations of 5 mM or above and ATP concentrations of 0.4 mM. Together, these experiments reveal the optimal *in vivo* labelling conditions for Tween-20 permeabilized cells to be around 0.1 mM ATP and 3 mM GTP.

As shown above, the detergent Tween-20 enables efficient cell permeabilization and substrate labelling. Unfortunately, it proved to be incompatible with downstream mass spectrometry (MS) analysis, as Tween-20 contains multiple repeat units of polyethylene glycol (PEG) and leads to a high degree of contamination in the MS samples. To find a replacement for Tween-20, three presumably MS-compatible detergents were tested for their ability to permeabilize N2a cells (Figure 13A). Sodium dodecyl sulfate (SDS), n-octyl-β-D-glucopyranoside (nOG) and lauryl maltose neopentyl glycol (LMNG) were employed and the degree of permeabilization was monitored via Trypan Blue staining. Only nOG was able to efficiently permeabilize the cells, whereas SDS and LMNG did not result in sufficient permeabilization at concentrations where cell integrity and surface attachment were maintained. nOG was therefore tested more extensively, with concentrations ranging from 0.1-0.5 % (w/v) at two different time points (Figure 13B). Efficient cell permeabilization can be observed from 0.3 % (w/v) nOG onwards, with a further increase at 0.4 % (w/v). Both ten and thirty min incubation times show efficient cell permeabilization, while overall cell integrity is maintained. Even at high detergent concentrations, the cells remained attached to the surface. We confirmed these results by performing immunofluorescent staining of nOG-permeabilized cells, and could observe good antibody permeation into the cytoplasm, with low nuclear signal. Together, this revealed nOG as an efficient replacement for Tween-20, with concentrations between 0.3 – 0.4 % (w/v) optimal for inducing protein-permeable pores.

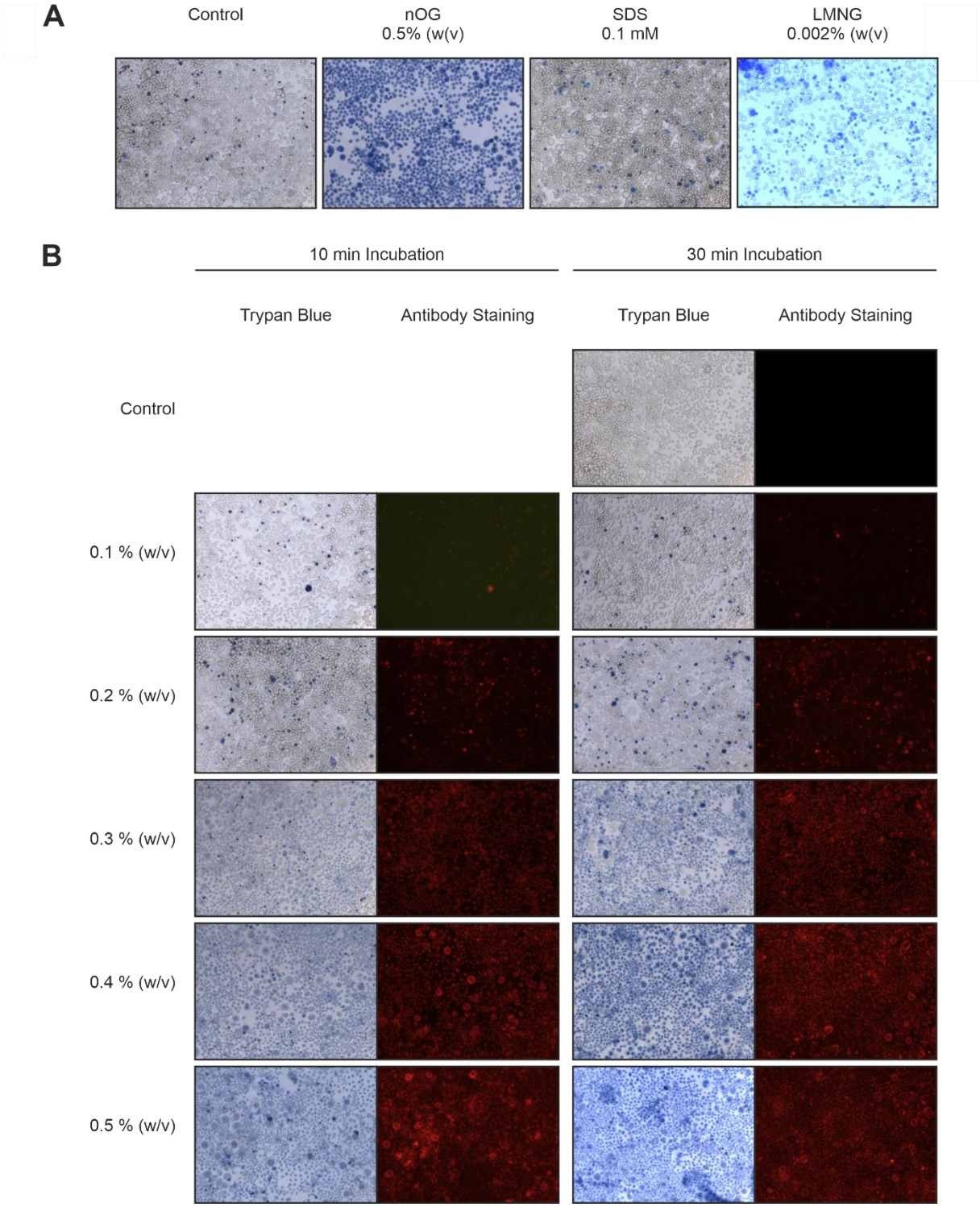



Figure 13: nOG can efficiently permeabilize N2a cells, suitable for *in vivo* labelling.

(A) Effect of different detergents on the permeabilization of N2a cells. N2a cells were incubated with PBS supplemented with either 0.5% (w/v) nOG (n-octyl-β-D-glucopyranoside), 0.1 mM SDS (sodium lauryl sulphate) or 0.002% (w/v) LMNG (lauryl maltose neopentyl glycol) for 30 min. Successful cell permeabilization was confirmed via trypan blue staining. (B) Systematic comparison of different concentrations of nOG for permeabilization of N2a cells. Cells were incubated with PBS supplemented with different nOG concentrations for 10 or 30 min. Successful cell permeabilization was confirmed via Trypan Blue staining. The presence of protein-permeable pores was confirmed by staining with an anti-GAPDH antibody. The integrity of the nuclear membrane was confirmed by staining with an anti-HNRNPL antibody. Both primary antibodies were detected via the same secondary antibody.

To compare labelling efficiencies between nOG and Tween-20, the established *in vitro* kinase assay was repeated using the previously optimized labelling conditions (Figure 14A). A clear difference between control samples and stimulated, analog-sensitive mutant overexpressing samples can be observed. However, the signal for stimulated untransfected cells remains relatively high, underscoring the importance of background noise control in subsequent analyses. Both detergents lead to similar results, although Tween-20 shows a slightly increased labelling efficiency, most evident for the protein substrate Syntide-2-GST. To avoid potential interference, the Twin-Strep-tag was removed, and the assay repeated with all four selected CaMKIIβ isoforms (Figure 14B). Again, a clear difference between the control samples and isoform samples can be observed. As an additional control, a kinase-dead CaMKIIβ variant (K43R, mutation of lysine 43 to arginine) was included. A stronger signal for CaMKIIβ autophosphorylation can be observed for the FL and Δ13 isoform, even though this, to some extent, depends on the blotting conditions.

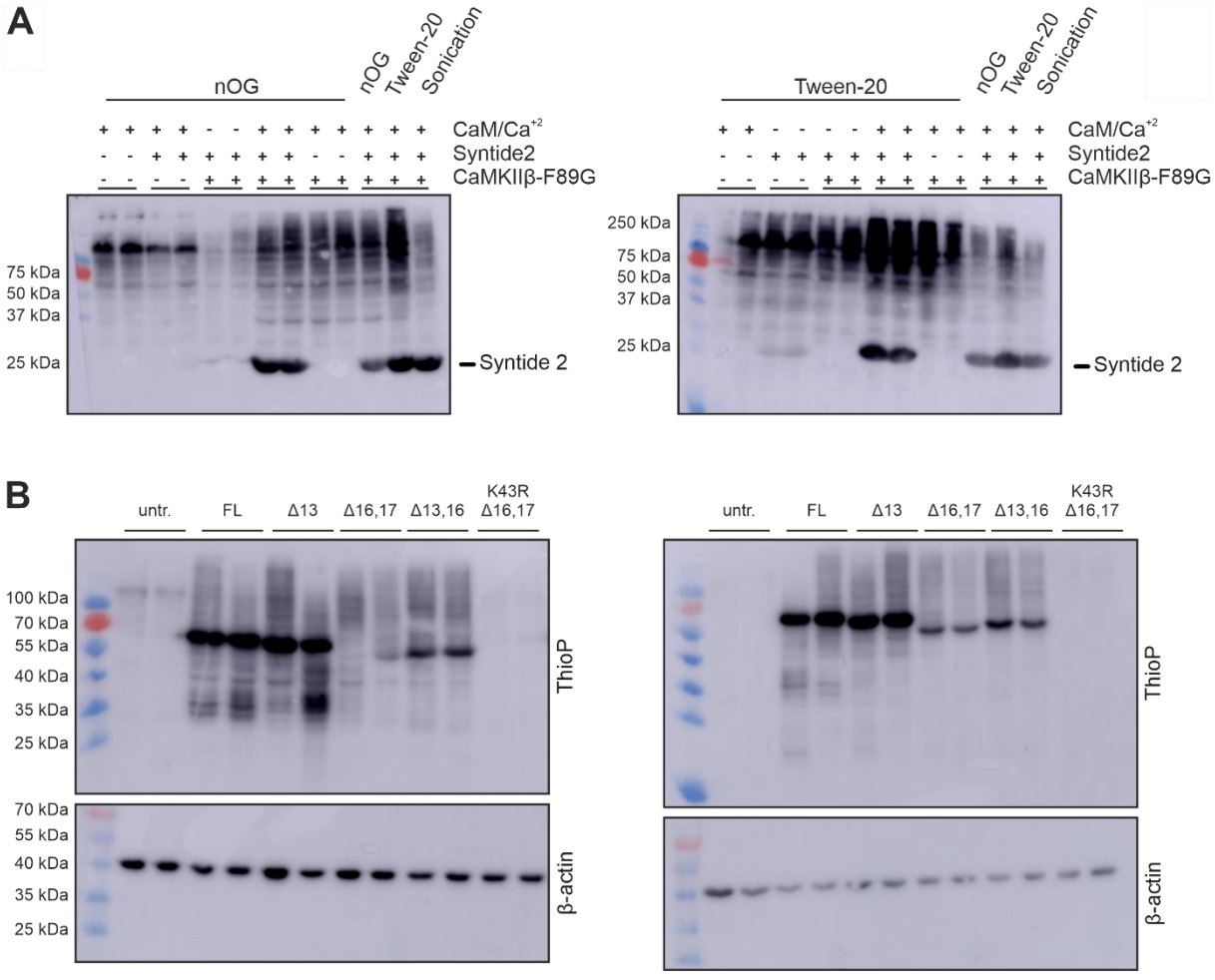



Figure 14: *In vivo* labelling of specific CaMKIIβ substrates in permeabilized cells.

(A) Comparison of nOG (n-octyl-β-D-glucopyranoside) and Tween-20 as permeabilization agents under optimized conditions. N2a cells overexpressing Twin-Strep-CaMKIIIβΔ13,16-F89G were collected and permeabilized with nOG or Tween-20 as indicated, or lysed by brief sonication. Reactions were performed with $N^6$-benzyl-ATPγS in the presence/ absence of stimulating conditions (CaM/$Ca^{2+}$) or an external substrate (Syntide 2, linked to GST). Samples were alkylated, run on an SDS-PAGE and analysed via western blot with a thiophosphate ester-specific antibody. (B) Final *in vivo* labelling reaction with selected, untagged CaMKIIIβ-F89G isoforms overexpressed in N2a cells and permeabilized with nOG. Untransfected cells and a kinase dead mutant (K43R) were used as negative controls. Successful labelling was confirmed by SDS-PAGE followed by western blot with a thiophosphate ester-specific antibody (ThioP). Equal cell concentrations were confirmed with a β-actin antibody (β-actin). Left and right blots show to same samples to illustrate how the signal varies between blotting conditions.

## Identification of isoform- and group-specific substrate spectra

The two control samples and four CaMKIIβ isoform samples were then used for thiophospho-enrichment and subsequent detection via mass spectrometry. The purification is based on the covalent capture of thiol-containing peptides, such as thiophosphorylated peptides, to iodoacetyl-beads. The exact procedure can be followed in the section “Analog-sensitive kinase assay – in vivo labelling and thiophosphate enrichment” in the “Material and Methods” chapter. Two datasets have been collected, as the samples were measured on two independent instruments. Each sample consists of two independent biological replicates and the analysis was restricted to phosphorylated peptides. As a large amount of starting material was used (4 mg cell lysate), a significant number of phosphopeptides were detected in the control samples (UT, untransfected and K43R, kinase-dead), emphasizing the need to control for background noise.

The two datasets show a large overlap regarding the total number of identified phosphorylation sites (Figure S10A) and a correlation matrix was generated based on their intensity values (Figure S10B). Both datasets were normalized to the same intensity range. Most samples correlate well between and within the two datasets and only individual replicates deviate significantly. This matrix still includes the control samples, and therefore background noise, which might explain the somewhat irregular correlation pattern. If all phosphorylation sites found in at least one control sample are removed, a clear correlation pattern becomes evident (Figure 15A). All biological and technical replicates from a specific sample correlate well with each other, suggesting that both datasets are highly similar and can be merged for subsequent analysis. A correlation is also revealed to exist between sample pairs generated by particular CaMKIIβ isoforms: the FL and Δ16,17 samples show a high correlation coefficient, as do the Δ13 and Δ13,16 samples. Together, this indicates an efficient enrichment for

specific CaMKIIβ substrates and suggests that both collected datasets can be combined for subsequent analysis.

Importantly, missing data is interpreted as the phosphorylation site not being present in the substrate spectra of that sample. This is different from standard whole proteome experiments, where most proteins and peptides are expected to be present in all analysed samples. Missing data, meaning the absence of a specific peptide, is therefore omitted from these analysis. However, due to the specific enrichment procedure for phosphopeptides, and the potentially different substrate spectra of individual CaMKIIβ isoforms, it is highly possible that not all phosphorylation sites are detected in all samples. Interpreting missing data as unknown would ignore these isoform-specific differences. When both datasets are merged and biological replicates are pooled (Figure S10C, D), this distinction becomes evident. If the analysis is restricted to pairwise shared targets (Figure S10D), all samples except the K43R control correlate exceptionally well. There is also a high correlation between the UT control and the CaMKIIβ samples, as the correlation is only calculated based on the background noise shared between all samples. K43R shows a lower correlation coefficient with the CaMKIIβ samples, except for the Δ16,17 isoform. The analog-sensitive mutant was derived from this variant and even though the mutation directly affects nucleotide binding, it retains some of its autophosphorylation-activity (Tullis et al., 2020). Together, this shows that the merged dataset is highly similar to the individual datasets and underscores the need to incorporate isoform-exclusive phosphorylation targets.

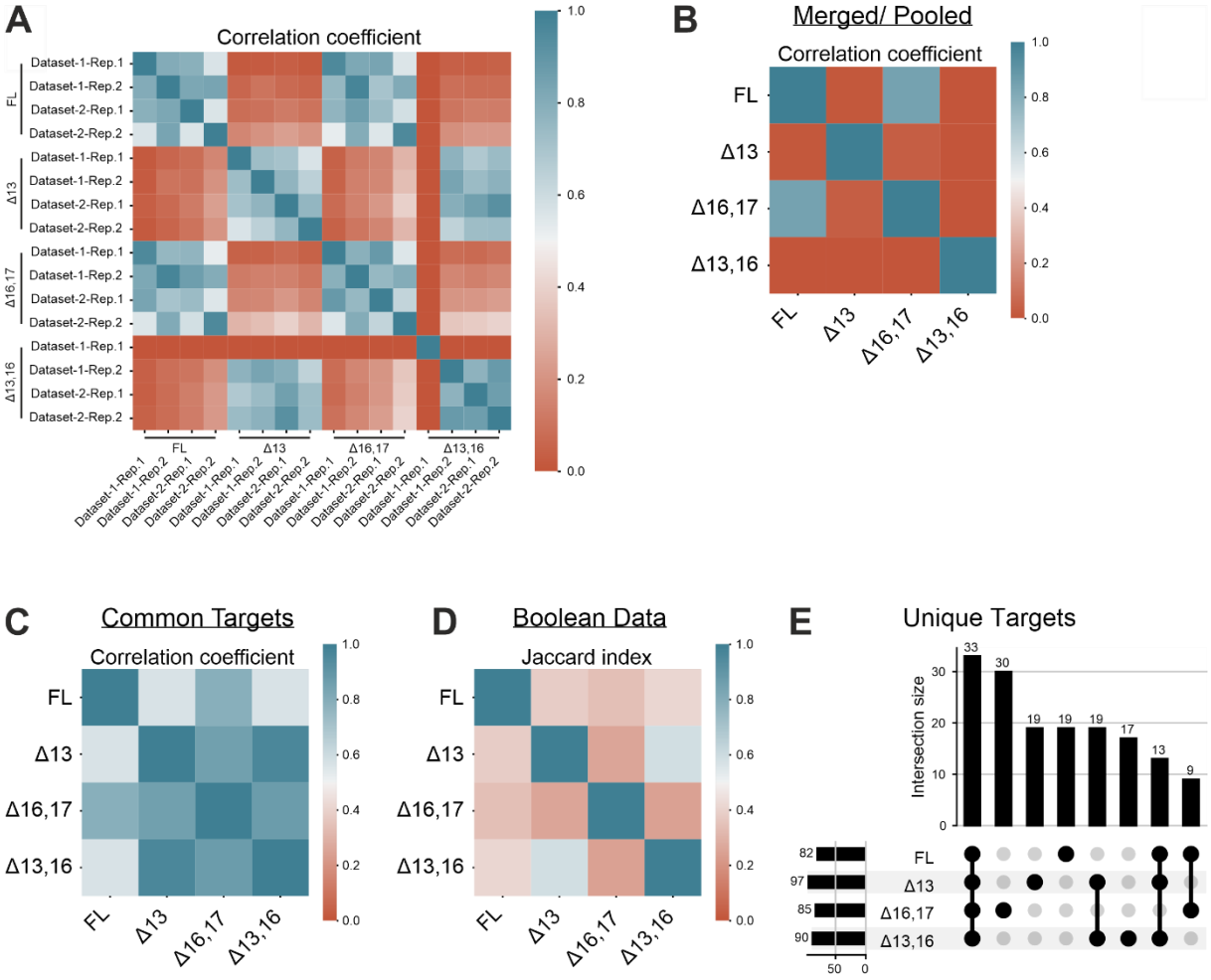



Figure 15: CaMKIIβ isoforms differ in their substrate spectra.

(A-C) Correlation matrices (Pearson) based on the intensity values of individual phosphorylation sites. Phosphorylation sites detected in the control samples were removed. (A) Comparison of the individual samples of the unpooled and unmerged datasets, ordered according to isoform. (B) Comparison of the merged and pooled dataset. (C) Comparison of the merged and pooled dataset, restricted to common targets (shared among all CaMKII samples) (D) Similarity matrix (Jaccard index) of the merged and pooled dataset. Phosphorylation sites detected in the control samples were removed. Intensity values were transformed into Boolean data (detected/ not detected). (E) Intersection plot showing the isoform and group-exclusive phosphorylation sites. Analysis was restricted to CaMKII-specific targets. Numbers on the left indicate the total number of phosphorylation sites detected in a sample. Numbers on the top indicate the intersection size between samples, meaning the number of phosphorylation sites that are unique to this group of samples. Black dots and connecting lines indicate the exact group of samples for which the intersection size is displayed.

If the merged dataset is filtered to remove the control samples, a clear correlation pattern between the CaMKIIβ isoforms becomes evident (Figure 15B). A strong correlation between the FL and Δ16,17 isoforms can be observed, whereas no correlation is detectable between the Δ13 and Δ13,16 isoform samples. This is due to a single Δ13,16 replicate whose phosphorylation levels deviate significantly from all other samples (Figure 15A). If removed, the correlation between the Δ13 and Δ13,16 isoforms is restored (Figure S10E). A similar pattern can be observed when the

analysis is restricted to targets shared among pairs of samples (Figure S10F). To circumvent the above mentioned problem of missing data, the analysis was restricted to targets shared among all pooled CaMKIIβ samples (Figure 15C). A similar correlation pattern can be observed, reinforcing the previous interpretation of isoform- and group-specific substrate spectra. Even if shared targets are analysed, the observed differences between the CaMKIIβ isoforms persist. In addition to exclusive phosphorylation targets, the substrate spectra therefore additionally differ in the abundance of shared targets. To confirm this, the dataset was transformed into a Boolean format, with phosphorylation sites being assigned a binary value of either present or absent, and used to calculate a similarity matrix (Figure 15D). The pattern observed in the previous correlation matrices mostly disappears, except for a slightly higher similarity between the Δ13 and Δ13,16 isoforms. When analysing isoform- and group-exclusive targets, the largest intersection is for targets shared between all CaMKIIβ samples (Figure 15E). Interestingly, this is followed by the intersection between the Δ13 and Δ13,16 isoforms, indicating that this group has a relatively large number of group-exclusive phosphorylation sites and confirming the observations from the similarity matrix. The FL and Δ16,17 pair can also be found, but with a lower number of group-exclusive hits. A comparably large number of phosphorylation sites are found in all CaMKIIβ isoforms except Δ16,17. However, this variant also features the highest number of isoform-exclusive phosphorylation targets, suggesting a particular functional relevance of exon 17 in conferring target specificity. The strongest effect on the substrate specificity is observed for exon 13, which is known to mediate the interaction with the F-actin cytoskeleton (Bayer *et al.*, 2006). This is confirmed using fluorescently tagged CaMKIIβ, overexpressed in N2a cells (Figure S10G). Full-length CaMKIIβ shows an enrichment in neurites and close to the plasma membrane, suggesting an interaction with F-actin. This is significantly reduced for the isoforms lacking exon 13, and to a lesser extent in the Δ16,17 isoform. Together, these experiments show isoform-specific differences in the substrate spectra between multiple CaMKIIβ variants. Although certain substrates appear to be isoform-specific, there are strong correlations between the samples generated by some of the isoforms, depending mostly on the presence or absence of exon 13.

Although isoform- and group-exclusive targets exist, many of the phosphorylation sites are shared between multiple samples. The main difference between the substrate spectra lies in the relative abundance of the various phosphorylation sites, which

represents the probability that a certain substrate will be phosphorylated. As the CaMKIIβ variants have identical kinase domains, each isoform should theoretically phosphorylate all possible targets. However, it seems that the variable linker segment, which differs between the isoforms, changes the probability of particular substrates to be phosphorylated. This can be explained by effects of the linker length and composition on either secondary protein-protein interactions or on the subcellular localization of the kinase isoforms.

When comparing individual phosphorylation targets, the clearest distinction can be seen for CaMKII autophosphorylation (Table 3). Eleven different autophosphorylation sites were detected, many of them showing the same correlation pattern observed before. Exon 13 exclusion isoforms naturally loose the phosphosites within exon 13, and additionally lack the S315 site located in exon 12. This site is also involved in F-actin binding, and upon phosphorylation disrupts the interaction with the cytoskeleton. T306 and T307 are the inhibitory phosphorylation sites, preventing association of calmodulin and thereby the full activation of the enzyme. Whereas T306 is completely absent in exon 13 exclusion isoforms, T307 is significantly downregulated. The corresponding activating autophosphorylation, T287, was detected for all isoforms, but seems upregulated in exon 16 inclusion isoforms. A similar effect can be seen for CaMKIIδ, where the signal for T287 decreases with a shortening of the linker. Interestingly, the distribution of these activating and inhibitory phosphorylation does not correlate with the kinetic differences observed *in vitro* (Figure 8B, C). Although the exon 16 exclusion variants Δ16,17 and Δ13,16 show a slightly higher activity *in vitro*, they show reduced levels of T287 autophosphorylation in the analog-sensitive kinase assay. The phosphorylation of T306/307 on the other hand does not correlate in any way with the inclusion of exon 16, but instead depends on the presence of exon 13. Other autophosphorylation sites without a known function also seem group-specific (S71 for exon 13-exclusion isoforms) or group-preferential (S534 downregulated in exon 13-exclusion isoforms). S276, which can be activating or inhibiting depending on the exact conditions (Berchtold et al., 2021), is specifically absent in the Δ16,17 isoform. Taken together, we observe clear differences regarding CaMKII autophosphorylation between the isoforms, with the main effect being the expected decrease of F-actin interaction sites in exon 13 exclusion isoforms. We furthermore see a clear difference regarding the activating and inhibitory phosphorylation, which show a preference for exon 16-inclusion and exon 13 inclusion isoforms, respectively.

Table 3. CaMKIIβ autophosphorylation sites are isoform-specific.

| | Log2 (Enrichment) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Unique phospho-site** | **Exon 13-incl.** | **Exon 16-incl.** | **FL/Ctr** | **13/Ctr** | **16,17/ Ctr** | **13,16/ Ctr** | **Location** | **Function** |
| **Camk2b-S315** | unique | 0,75 | unique | absent | unique | absent | Exon 12 | F-actin binding |
| **Camk2b-T320** | unique | 0,86 | unique | absent | unique | absent | Exon 13 | F-actin binding |
| **Camk2b-T321** | unique | 1,00 | unique | absent | unique | absent | Exon 13 | F-actin binding |
| **Camk2a-T306** | unique | 1,62 | unique | absent | unique | absent | Exon 12 | Inact. |
| **Camk2a-T307** | 4,26 | 1,12 | unique | unique | unique | unique | Exon 12 | Inact. |
| **Camk2b-S534** | 0,23 | 1,71 | 1,17 | 0,95 | -0,51 | -0,81 | Hub Domain | NA |
| **Camk2b-S280** | -0,08 | 0,84 | 5,68 | 5,58 | 4,58 | 4,98 | Exon 11 | GlcNAc site |
| **Camk2b-T287** | -0,09 | 0,84 | 6,57 | 6,48 | 5,47 | 5,88 | Exon 11 | Act. |
| **Camk2d-S280** | -0,20 | 1,03 | 2,74 | 2,47 | 1,58 | absent | Exon 11 | GlcNAc site |
| **Camk2d-T287** | -0,20 | 1,03 | 2,74 | 2,47 | 1,58 | absent | Exon 11 | Act. |
| **Camk2b-T254** | -0,84 | -0,31 | unique | unique | unique | unique | Kinase Domain | NA |
| **Camk2b-S276** | -1,77 | -0,75 | unique | unique | absent | unique | Exon 11 | act./inact. |
| **Camk2b-S71** | absent | 0,57 | absent | unique | absent | unique | Kinase Domain | NA |

Table 3. The list of detected phosphorylation sites was restricted to gene names CaMK2a, CaMK2b, CaMK2d. Exon 13-incl./ Exon 16-incl.: Log2 ratio between the average intensity values for exon 13 or 16 inclusion variants vs. exon 13 or 16 skipping variants. Isoform/Ctr: Log2 ratio between the average intensity values for this isoform vs. controls. Unique: this phosphorylation target was not detected in the corresponding control samples (13/16-exclusion variants for 13-inclusion/16-inclusion; UT and K43R for isoform/Ctr. Absent: this phosphorylation site was not detected in the corresponding sample.

The strong signal for CaMKII autophosphorylation, including for two sites located within exon 13 itself (T320/321), could potentially confound the previous correlation analysis. These sites can physically not be present in exon 13 exclusion isoforms and therefore

need to be removed when comparing the substrate spectra (Figure S11A). After this transformation, the observed correlation pattern is maintained, including for the merged and pooled datasets (Figure S11B). If all CaMKII autophosphorylation targets are removed, the only remaining correlation is that between the Δ13 and Δ13,16 isoforms, if the Δ13,16 outlier is omitted (Figure S11C). The similarity matrix, based on the transformation into Boolean data, is also not affected by removing CaMKII autophosphorylation targets (Figure S11D). This is in line with the observed high number of targets exclusively found in exon 13 exclusion isoforms (Figure 15E). Together, this indicates that the observed similarities in substrate spectra between the FL and Δ16,17 isoforms are mainly due to similarities regarding CaMKII autophosphorylation. In contrast, exon 13 exclusion isoforms show additional group-specific non-CaMKII targets.

When comparing general phosphorylation sites, differences between isoforms become less clear. A classic GO term enrichment analysis only results in a significant enrichment of the term GO:0032880 (regulation of protein localization) for the Δ13 isoform and of GO:0005515 (protein binding) for the Δ16,17 isoform. Interestingly, if one instead manually compares targets associated with individual GO terms of interest (e.g., “synapse”, “postsynapse organisation”, “actin cytoskeleton”) between the substrate spectra of different CaMKIIβ isoforms, we do not see a decrease of actin cytoskeleton associated targets in the exon 13 exclusion variants. The only notable clustering can be observed for targets associated with “myelin sheath” (GO:0043209) as a cellular component, which seems specifically enriched in the Δ16,17 isoform (Table 4 and Supplementary Table S3). In summary, although the general principles of isoform-specific substrate spectra seem evident, the exact functional consequences based on individual phosphorylation targets remains elusive.

Table 4. Myelin sheath associated targets are specific to the Δ16,17 isoform.

| | Log2 (Enrichment) | | | | | |
|---|---|---|---|---|---|---|
| **Unique phosphosite** | **Exon 13-incl.** | **Exon 16-incl.** | **FL/Ctr** | **13/Ctr** | **16,17/Ctr** | **13,16/Ctr** |
| **Plec-T3795** | -0,45 | 0,49 | unique | unique | unique | unique |
| **Pkm-T25** | unique | absent | absent | absent | unique | absent |
| **Tubb4b-S126** | unique | absent | absent | absent | unique | absent |
| **Eef1a2-S396** | unique | absent | absent | absent | unique | absent |
| **Slc25a4-T254** | unique | absent | absent | absent | unique | absent |
| **Gdi2-S302** | unique | absent | absent | absent | unique | absent |
| **Plcb1-S564** | 0,26 | absent | absent | absent | unique | unique |
| **Actb-S33** | absent | unique | absent | unique | absent | absent |

Table 4. Identified phosphorylation targets that associate with the GO term GO:0043209 (myelin sheath). The table is restricted to phosphorylation targets that were not detected in a control sample (UT, K43R). Exon 13-incl./ Exon 16-incl.: Log2 ratio between the average intensity values for exon 13 or 16 inclusion variants vs. exon 13 or 16 skipping variants. Isoform/Ctr: Log2 ratio between the average intensity values for this isoform vs. controls. Unique: this phosphorylation target was not detected in the corresponding control samples (13/16-exclusion variants for 13-inclusion/16-inclusion; UT and K43R for isoform/Ctr. Absent: this phosphorylation site was not detected in the corresponding sample.

## A mouse model with humanized *Camk2β* splicing

To study the consequences of *Camk2β* alternative splicing *in vivo*, we generated a mouse model with a humanized *Camk2β* splicing pattern. Based on the results obtained with our minigenes, we used CRISPR/Cas9 and introduced the identified human intronic regulatory sequence into the mouse genome. We generated two mouse strains, one containing the human intronic regulatory sequence, termed “humanized strain”, and one in which the mouse sequence had simply been deleted, termed “deletion strain” (Figure 16A). The intron-exon structure is retained, as only a part of the intron was exchanged or deleted, leaving both splice sites intact. Indeed, this results in a human-like *Camk2β* splicing pattern in the brain of mutant mice, most evident by the presence of the primate-specific splice isoform Δ13,16 (Figure 16B). Sanger sequencing also confirmed the presence of the other previously identified primate-specific exon 16-exclusion isoforms. Interestingly, the mouse model shows an additional band for a Δ13,16,17 isoform, that was previously not detected in our human sample. Strikingly, both mouse model strains show an identical splicing pattern. This confirms the minigene splicing assays and the postulated model of species-specific differences in branch point strength. The mechanism is *cis*-regulated, with the mouse sequence harbouring a functionally relevant, strong branch point motif. The sequence of the human intron is inconsequential, and its insertion into the mouse endogenous context has the same effect as just deleting of the corresponding mouse sequence.

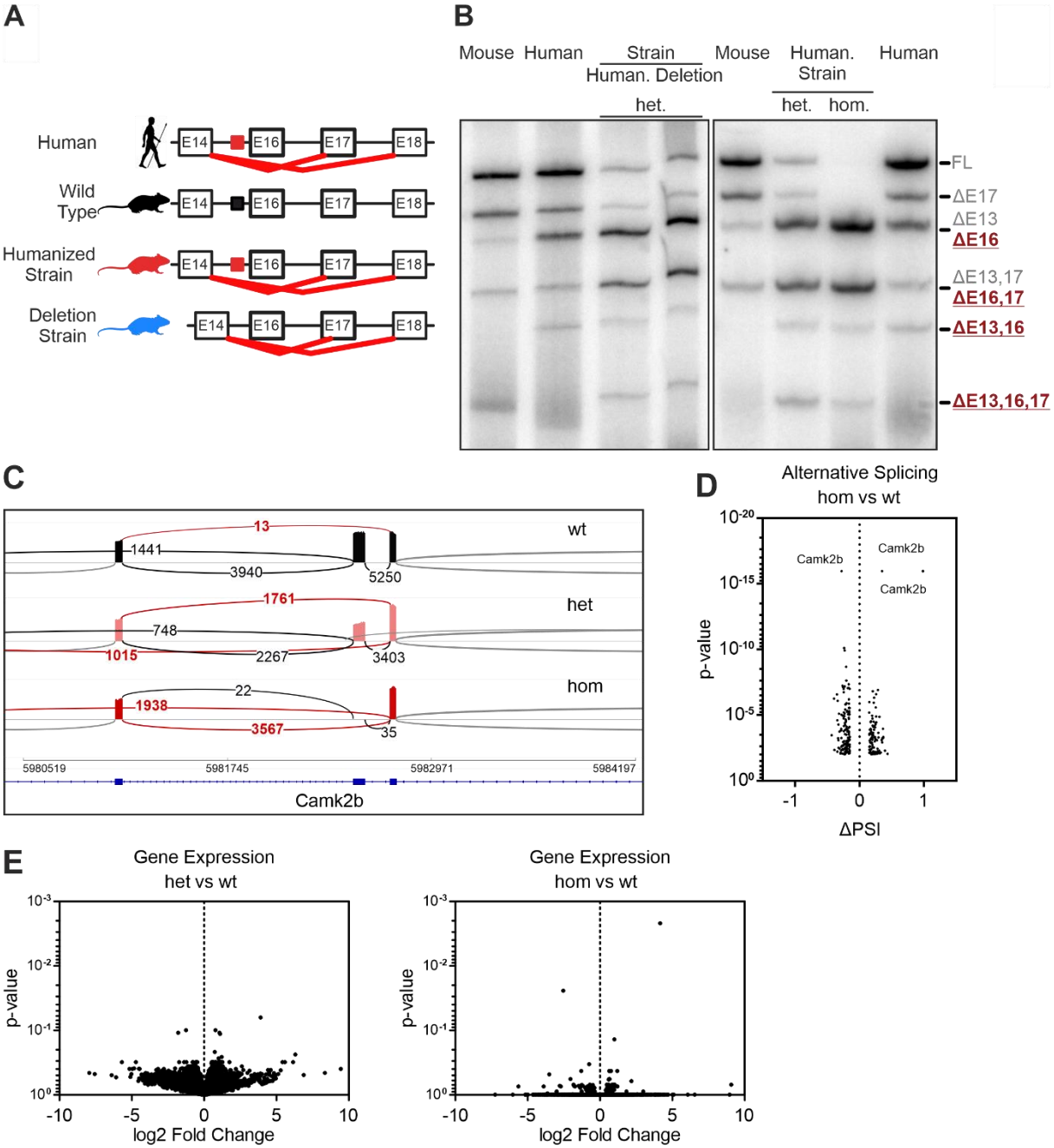



Figure 16: Generation of a mouse model with humanized *Camk2β* splicing.

(A) Schematic representation of the intron-exon structure of the variable linker region of the *Camk2β* gene and comparison of the identified alternative splicing isoforms in human, wild type mice, and the novel mouse model with a humanized *Camk2β* splicing pattern (humanized strain and deletion strain). Red lines indicate identified species-specific splicing events. Coloured boxes indicate location of the identified *cis*-regulatory element in human (red) and mice (black). (B) Endogenous *Camk2β* splice isoforms were identified by radioactive isoform-specific RT-PCR with mouse (*Mus musculus*) and human cerebellum RNA. Isoforms were separated on a denaturing polyacrylamide gel. Isoforms are indicated on the right and named according to the skipped exons. Human.: humanized strain, Deletion: deletion strain, wt: wild type animals, het: heterozygous animals, hom: homozygous animals. (C) Sashimi plot showing the alternative splicing of *Camk2β* exon 16 in RNA-Seq data from wild type, heterozygous and homozygous mice of the humanized strain. Each graph summarizes RNA-Seq data of 4 biological replicates.(D) Volcano plot mapping the differences in percentage spliced in (PSI) of cassette exons of homozygous vs. heterozygous animals against their respective p-values. Individual splicing events affecting *Camk2β* exon 16 are labelled. (E) Volcano plot mapping gene expression changes in the mouse model for both heterozygous and homozygous animals of the humanized strain.

There appear to be strong splicing differences between heterozygous and homozygous animals. Whereas the heterozygous animals show a very human-like *Camk2β* splicing pattern, the homozygous animals lack the FL and Δ17 isoforms. Instead, the bands for the Δ16 and Δ16,17 isoforms display increased intensity, indicating that exon 16 is efficiently removed in these animals. To confirm this, we performed RNA-Seq on cerebellum samples from the humanized strain. Indeed, we can see a nearly complete inclusion of exon 16 in the wild type mouse, which is reduced to around 50 % in heterozygous animals and essentially absent in homozygous animals (Figure 16C). We also checked whether alternative splicing was affected on a global level in our mouse model. Strikingly, exon 16 of *Camk2β* comes up as the most affected splicing event, far ahead of any other exon (Figure 16D). Most other splicing events that were found to be altered in our mouse model affect genes with low expression levels, suggesting that they represent natural and stochastic fluctuation in splicing, or imprecisions of the analysis pipeline. We selected nine of the highest confidence targets for validation by RT-PCR (Figure S12). None of the targets identified by RNA-Seq could be validated, confirming the specificity of our mouse model and suggesting that alternative splicing is not affected on a global level. We could also not detect any significant differences in gene expression levels between wild type and heterozygous animals, and only two genes that differ between wild type and homozygous animals (Figure 16E). These two genes, Gm12184 and Zbtb37, have no known connection to synaptic plasticity or learning and memory. This suggests that under resting conditions, gene expression is not significantly altered in our mouse model of humanized *Camk2β* alternative splicing. In summary, our mouse model confirms the proposed mechanism of a functionally relevant branch point sequence upstream of exon 16 that is potent in the mouse *Camk2β* gene, and absent in primates. Deletion of this sequence results in a human-like *Camk2β* splicing pattern in the brain of mutant mice. RNA-Seq analysis confirms the specificity of our mouse model and shows that alternative splicing and gene expression are not altered on a global level.

## Mice with humanized *Camk2β* splicing pattern show reduced levels of LTP

Having confirmed the validity of our mouse model, we next set out to determine the effects of an altered *Camk2β* altered splicing pattern on synaptic plasticity. We performed an electrophysiological characterization of CA1-CA3 neurons in acute hippocampal slices of both mouse strains. Basal synaptic transmission, as well as short-term plasticity, measured as paired pulse ratio, do not appear affected in the mouse model (Figure 17A, B). However, high frequency induced long-term potentiation (LTP) was significantly impaired in both mouse strains, both after 30 and 60 minutes post stimulation (Figure 17C-E). Again, short-term potentiation measured as the immediate response after the tetanic pulse (post-tetanic potentiation) was not affected. LTP was induced with a single high frequency tetanic pulse and stimulation with multiple pulses led to similar results (Figure S13 A-C). Together, these experiments show that in our mouse model with humanized *Camk2β* alternative splicing, basal synaptic transmission as well as short-term plasticity does not appear affected, whereas LTP is severely impaired in both mouse strains.

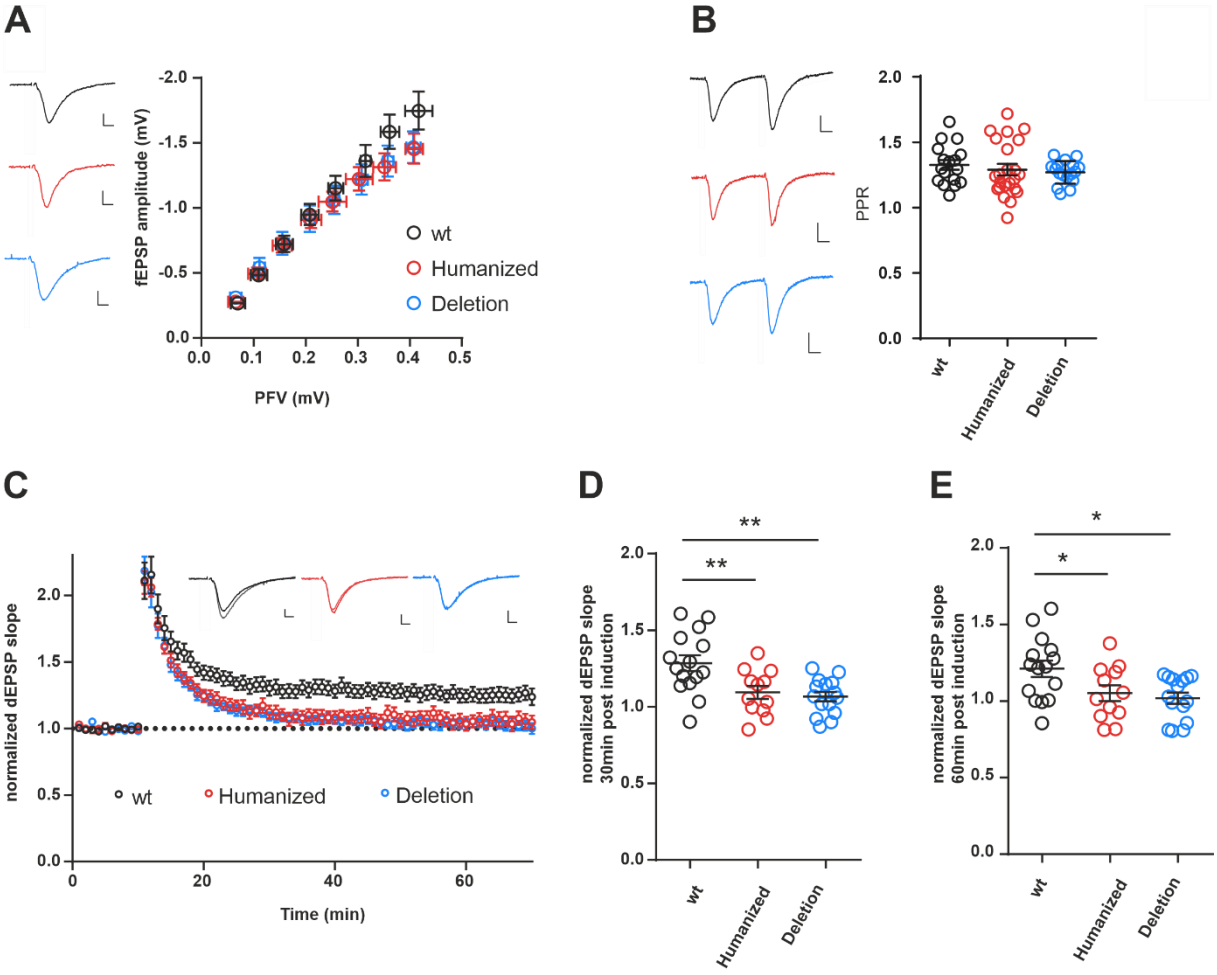



Figure 17: Long-term potentiation is severely impaired in the mouse model.

(A) Input/ output characterization: relationship between amplitudes of presynaptic fiber volley (PFV) and field excitatory postsynaptic potential (fEPSP). Scale bars: 0.2 mV/ 5ms. (B) Short term plasticity (paired pulse ratio (PPR) with 50 ms inter-stimulus interval). Scale bars: 0.2 mV / 10ms. (C) Time course of LTP induction in CA3-CA1 synapses in acute hippocampal slices. LTP was induced after 10 min with a single train of 100 Hz, 1 s. Example traces show average of baseline and potentiated field excitatory postsynaptic potentials (fEPSP) 60 min after LTP induction. Scale bar: 0.2 mV/ 5 ms. wt (wild type): 15 slices, 6 mice, humanized (humanized strain, homozygote): 12 slices 6 mice; deletion (deletion strain, homozygote):15 slices, 6 mice. (D, E) Bar plots depicting the field EPSP slope after 30 min (D) or 60 min (E) after LTP induction. *p<0.05, **p<0.01, calculated by ANOVA followed by Dunnet's test. These experiments were performed by Dr. Alexander Stumpf in the research group of Prof. Schmitz (Charité, Universitätsmedizin Berlin).

## Mice with humanized *Camk2β* splicing pattern show diverging behavioural traits

Having established the electrophysiological consequences in our mouse model, we set out to identify potential resulting behavioural phenotypes. We first performed a nest building assay with the humanized strain, in which the animals are individually housed overnight and supplied with cotton pads as nest building material. Homozygous animals of the humanized strain built better nests, measured as the amount of nestlet material remaining (Figure 18A) or nest score (Figure S14A). Heterozygous animals scored lower than their wild type littermates. Even though a clear trend for both groups can be observed in comparison to wild type animals, it only reaches statistical significance in the direct comparison of heterozygous and homozygous mice. We could not detect a significant difference between male and female animals regarding nest building behaviour (Figure S14B, C). Even though a definitive interpretation of nest building behaviour is difficult, the test can partially be interpreted as a measure of locomotor activity and general motor function. As CaMKIIβ is known to be expressed in muscle tissue, where it presumably helps to target the holoenzyme to the sarcoplasmic reticulum (Bayer *et al.*, 1998; Urquidi and Ashcroft, 1995), an effect on motor function is highly plausible.

To investigate this, we performed an open field test with the humanized strain, in which the animals are allowed to freely explore an open arena. We could not detect significant differences between the genotypes regarding locomotor activity (Figure 18B, S14D), but instead found increased overall activity levels in female animals (Figure S14E, F). When analysing the behaviour of male and female mice separately, a trend of increased locomotor activity can be seen in male homozygous animals (Figure S14G, H). We did not observe significant differences between the genotypes regarding anxiety levels, measured as time spent in the central part of the arena (Figure 18C). Although the median remains similar across all genotypes, we observe an increased variance in anxiety levels for heterozygous and homozygous animals. Specifically, more mice spent an extended time in the centre, which is considered a proxy for reduced anxiety levels. No significant difference was observed between the anxiety levels of male and female mice (Figure S14I, J). These observations are also evident in the defecation levels, an alternative measure of anxiety in mice (Figure S14K, L).

To confirm these results, we repeated the open field test with the deletion strain of our mouse model, in which we observed similar phenotypes regarding locomotor activity (Figure S15A-C). Specifically, when analysed separately, male homozygous animals showed significantly increased levels of locomotor activity (Figure S15C). Interestingly, we observe a clear decrease in anxiety levels, with homozygous animals of the deletion strain showing increased time spent in the central part of the arena (Figure S15D), as well as decreased defecation levels (Figure S15E). This might be explained by slight differences in the experimental setup (see Material and methods). Together, these results show distinct phenotypes for our mouse model of humanized *Camk2β* alternative splicing, namely changes in nest building behaviour and altered locomotor activity. As we have not modified exonic coding regions but an intronic splicing regulatory sequence, any phenotype in these mice is the result of the humanized *Camk2β* splicing pattern and thus directly connects species-specific alternative splicing with differences in behaviour.

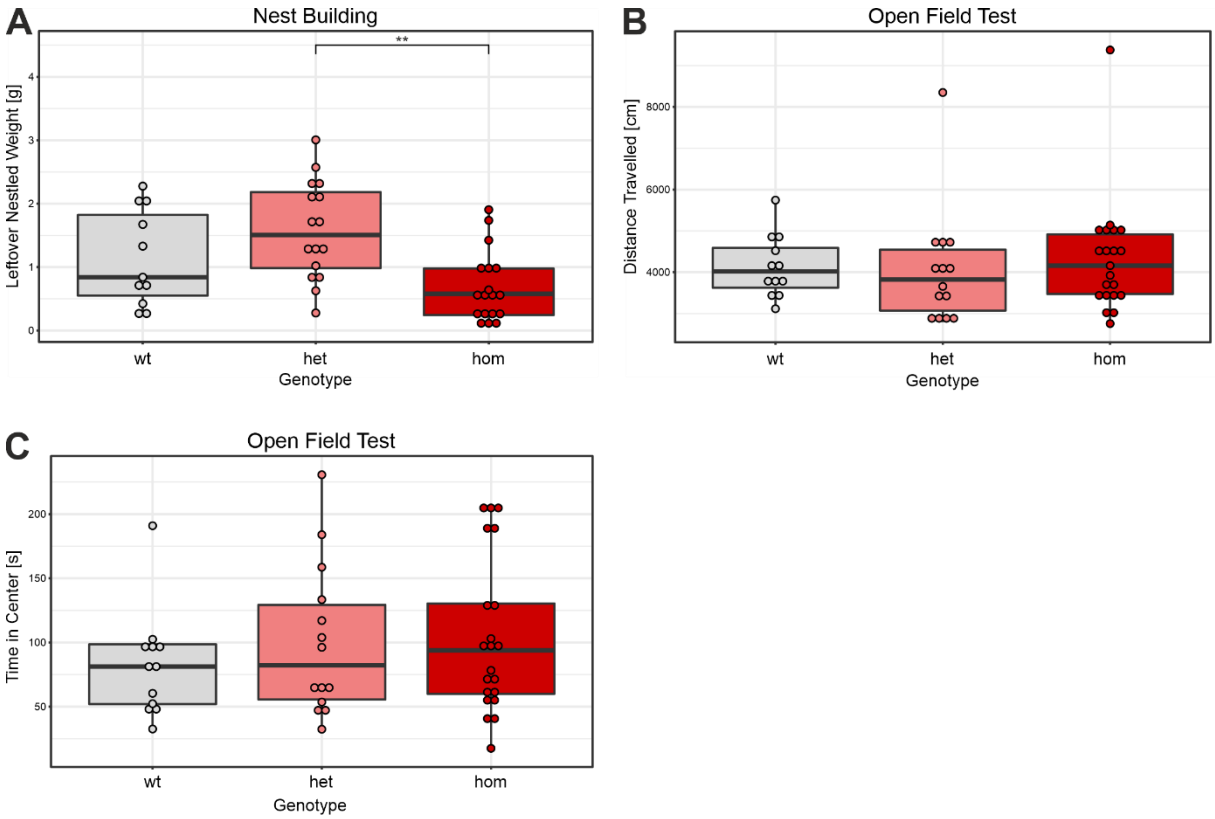



Figure 18: Behavioural characterization of the mouse model.

(A) Boxplot of the nest building test, displaying leftover nestlet weight. Mice were singly housed and supplied with cotton pads of 3.0 g overnight. Unused material was collected and weighted. *p<0.05, **p<0.01, ***p<0.001 calculated by Wilcoxon signed-rank test. Adjusted for multiple comparisons using Holm‘s method. (B) Locomotor activity as measured in the open field test. Mice were allowed to freely explore an open arena (40 x 40 cm) for 10 min. Locomotor activity is displayed as distance travelled in cm over the entire session. (C) Anxiety levels as measured in the open field test. Sessions were videotaped and the time spend in the central part of the arena (20 x 20 cm) determined. wt: wild type mice (n=11 (nest building), n=12 (open field)), het: heterozygous mice (n=16 (nest building), n=14 (open field)), hom: homozygous mice (n=18 (nest building), n=21 (open field)) of the humanized strain.

# Discussion

In this thesis, I report the species-specific alternative splicing of three of the four CaMKII genes (Figure 4). For *Camk2β*, I decipher the exact mechanism consisting of an altered branch point (BP) sequence, regulating the primate-specific exclusion of exon 16 (Figure 5). Tissue- and developmental stage-specific alternative splicing has long been established for CaMKII (Sloutsky and Stratton, 2021; Tombes *et al.*, 2003). Formerly reported differences between species focused mostly on the number and exact composition of CaMKII genes. The primate-specific exon 16-exclusion isoforms that I report here have been published before for human samples (Rochlitz *et al.*, 2000; Sloutsky *et al.*, 2020) but not for mouse (Cook *et al.*, 2018). Yet, previous studies failed to recognize the species-specific characteristics therein. Interestingly, the mechanism we discovered is not guided by the standard combination of a *cis*-acting splicing enhancer or silencer motif that is bound by an auxiliary *trans*-acting protein factor. Instead, it affects the constitutively present BP sequence. Importantly, the intron itself is not conserved, yet all tested species feature one or more potential BPs at nearly identical distances to the 3' splice site. The calculated differences in BP strength are sufficient to explain the observed differences between primates and other species, including other mammals (Figure 7, Table 1). It does not fully explain the variance between primates, which might reflect minor differences in the canonical splicing machinery and thus the interpretation of conserved regulatory elements. It is also possible that additional, as of yet unknown, regulatory mechanism exist, which could explain the large variance between individuals of the same species (Sloutsky *et al.*, 2020). The general importance of the BP sequence and strength in regulating alternative splicing is not fully understood, in part due to experimental difficulties in validating BP sequences (Mercer *et al.*, 2015).

## Difficulties in determining isoform-specific kinetic characteristics

Alternative splicing of CaMKII leads to over 70 different variants in humans (Sloutsky et al., 2020; Tombes et al., 2003) many of them lacking a known isoform-specific function. Some redundancy between the four CaMKII genes has been reported, and knockdown of a single gene can result in expression, localization or phosphorylation changes of other CaMKII isoforms (Elgersma et al., 2002; Yamasaki *et al.*, 2008). Almost all known CaMKII splice isoforms differ exclusively regarding the length and composition of their flexible linker domains. An initial model proposed that this linker

modulates an equilibrium between various inactive states of the holoenzyme, thus affecting its kinetics. However, the exact nature of these states is still under debate (Buonarati *et al.*, 2021; Chao *et al.*, 2011; Myers *et al.*, 2017; Sloutsky *et al.*, 2020). In this study, we confirm recently published results that for CaMKIIβ, the linker segment does not affect the kinetics of the enzyme regarding $EC_{50}$ and Hill coefficient (Sloutsky *et al.*, 2020). Instead, we find differences regarding the maximal activity ($V_{max}$) reached at optimal calmodulin concentrations (Figure 8, Table 2). These characteristics might have gone unnoticed before, as most studies have focused on CaMKIIα and utilized a spectroscopic assay in which these differences are difficult to identify. A potential mechanistic explanation could be found in the behaviour of the holoenzyme under varying experimental conditions. A recent publication proposed stability differences between CaMKIIα holoenzymes (Torres-Ocampo et al., 2020), and we could show that the observed differences, at least regarding the autoactivity, depend on the exact buffer composition (Figure S5). When testing the stability of the enzyme, all isoforms show a similar dependency on salt concentration and pH (Figures 8G, S6). Reproducible isoform-specific differences in protein stability exist but are relatively small and do not correlate with the observed activity differences. This lack of major differences in kinetic properties and protein stability is supported by structural data, where two CaMKIIβ isoforms with different length linkers show a similar equilibrium between various oligomeric states (Figure 10A, Figure S7).

We could also detect minor differences regarding the autoactivity generated at various calmodulin concentrations, which is influenced by the autophosphorylation state of the holoenzyme. A recent publication showed that the speed at which these phosphorylations are acquired differs between CaMKII genes and isoforms (Bhattacharyya *et al.*, 2020). Isoforms with longer linker, such as full-length CaMKIIβ, were biased towards inhibitory autophosphorylation. However, these differences are not able to explain the observed activity differences, as the inhibitory autophosphorylation do not themself affect the activity state of the enzyme. They only prevent re-association of calmodulin and thereby full activation of CaMKII. In addition to the inhibitory autophosphorylation, the rate at which the activating autophosphorylation is acquired was also shown to differ according to linker length. This could indeed explain the observed activity differences, as this phosphorylation strongly induces autoactivity and further enhances calmodulin binding (Meyer *et al.*,

1992). Contradictory, this phosphorylation is enriched in the longer-linker isoforms in our mass spectrometry data (Table 3).

It remains to be established whether these small kinetic differences could have a significant effect in the context of a functioning synapse. CaMKII, at least in neuronal tissue, is a very unusual kinase as it is significantly more abundant than most of its substrates (Cheng *et al.*, 2006; Erondu and Kennedy, 1985). Together with the fact that nucleotide exchange is most likely the rate limiting step (Bayer and Schulman, 2019; Grant and Adams, 1996; Zhou and Adams, 1997), every CaMKII subunit only catalyses a single reaction during an individual calcium spike. It is therefore likely that our current kinetic assays are insufficient to reflect the native environment in neurons, such as the transient nature of calcium spikes, or the large excess of kinase. Most kinetic assays, including ours and recent studies (Sloutsky *et al.*, 2020) use steady-state kinetics, consisting of a permanent activation with calmodulin. Early experiments already showed that these might not fully capture isoform-specific differences, such as differential responses to short activation spikes (Bayer *et al.*, 2002). It therefore appears that kinetic differences between CaMKII isoforms come in many forms, including cooperativity differences between CaMKIIα and β (Sloutsky *et al.*, 2020), differences in maximal activity at optimal calmodulin concentrations (this study), differential rates of acquiring various autophosphorylations (Bhattacharyya *et al.*, 2020) and isoform-specific responses to transient calcium-spikes (Bayer *et al.*, 2002). None of these studies fully reflect the native environment in synapses and it remains questionable which characteristics remain in an intact neuron.

## An unexpected temperature-sensitivity of CaMKII

Another novel kinetic aspect of CaMKII is the observed temperature-sensitivity of the enzyme (Figure 9). Even though CaMKII has been studied for nearly 40 years (Bayer and Schulman, 2019), no such dependency has been reported. Previous studies might have missed this aspect, as it requires a deviation from the standard protocol. Namely, CaMKII has to be pre-incubated with calmodulin and failure to do so eliminates the temperature-dependency (see material and methods). This seems to be a conserved property of the kinase domain, as we could not observe a significant difference between CaMKII genes and splice isoforms within a physiologically relevant temperature range. All tested variants show a near-linear temperature-dependency with an approximate 2-fold activity change (Figure 9D). When the temperature range

is extended, gene-specific differences become visible but should not play a major role for most cellular processes (Figure 9E). This 2-fold activity change is indeed a much larger effect than the previously described $V_{max}$ differences between various CaMKIIβ isoforms. It provides a direct mechanistic link between synaptic plasticity and body temperature, as influenced by circadian rhythm, aging, physical activity and disease state. We recently reported another class of temperature-dependent kinases, linking circadian rhythm to changes in alternative splicing (Haltenhof *et al.*, 2020). Interestingly, whereas for the previously identified temperature-dependent CLKs the activation loop of the kinase domain plays a central role, this feature is not conserved in CaMKII. Most kinases are activated by phosphorylation of their activation loop, which is positioned near the active site (Nolen et al., 2004). In CaMKII, the phosphorylation site is replaced by a conserved glycine and the activation loop is instead stabilized by a hydrogen bond network (Rellos *et al.*, 2010), locking the enzyme in an activation-competent conformation. The inactivation is mediated by a regulatory segment immediately adjacent to the kinase domain, featuring multiple functionally relevant phosphorylation sites. In the absence of calmodulin or activating phosphorylations, it keeps the enzyme in an autoinhibited state (Bayer and Schulman, 2019; Chao *et al.*, 2011). It therefore remains to be established whether the activation loop plays a similar role in mediating the temperature-dependency for CaMKII. Other temperature-sensing systems exist and often consist of receptors or ion channels (Bautista et al., 2007; Gong et al., 2019; Wetsel, 2011). However, CaMKII takes a special role, being highly abundant in neurons and not primarily a temperature-sensor. Instead, it is directly involved in mediating synaptic plasticity as a key signaling molecule and structural component. Due to the abundance of CaMKII and the observed strong temperature-dependency, we expect even the relatively small physiologically relevant temperature changes to have a significant impact on brain function and synaptic plasticity. It is also noteworthy that even though the most studied functions for this enzyme are found in neuronal and cardiac tissue, CaMKIIγ shows a more ubiquitous expression compared to CaMKIIα or β and an isoform of CaMKII can be found in most human cells. This also includes tissues such as peripheral sensory neurons in skin (Carlton, 2002; Uhlén et al., 2015), for which much stronger temperature fluctuations are common. Interestingly, reduced CaMKII activity has been linked to a switch from sensory to pain pathways in specialized sensory neurons termed low-threshold mechanoreceptors (Yu et al., 2015). Cold is known to effectively reduce pain sensations and the herein reported

increase of CaMKII activity at lower temperatures could provide an additional simple mechanistic explanation.

## Subcellular localization affects substrate specificity

Apart from its enzymatic function, CaMKII has a well-established structural role. It is known to bind to and bundle actin filaments, a function that differs between CaMKII genes and is also influenced by alternative splicing (O'Leary *et al.*, 2006). Gene specific differences have been shown for CaMKIIβ and γ. The shorter γ-variant results in a layered structure, in contrast to the rod-like actin bundles formed by CaMKIIβ (Hoffman *et al.*, 2013). Contrary, another study found no differences in actin-bundling between CaMKII variants with a 10-fold difference in linker length (Khan *et al.*, 2019). This study was however restricted to CaMKIIβ variants from different species (rat and human) and CaMKII from *C. elegans*. It therefore remains likely that the exact composition and length of the variable linker affects F-actin bundling, both by modulating functionally relevant interaction and phosphorylation sites, and by setting a distance-threshold for the bundling effect. One could potentially envision CaMKII as an interaction hub, with the unstructured and flexible linker domain acting as a multivalent binding site for various proteins. Holoenzymes containing multiple isoforms would then enable mixing of different interaction sites and fine-tuning of the exact distance between interaction partners. This would also affect subcellular localization of CaMKII itself and could potentially result in differential access to phosphorylation targets. For fly CaMKII, it could already be shown *in vitro* that different isoforms have different substrate spectra (GuptaRoy *et al.*, 2000), an effect that should be further potentiated in an intact cellular environment.

We performed an unbiased screen for CaMKII targets in permeabilized N2a cells (a mouse neuroblastoma cell line), to better recapitulate the *in vivo* situation. We find clear differences in the substrate spectra, including isoform-specific and group-specific characteristics (Figure 15). As expected, these strongly correlate with known localization differences. Namely, the presence of exon 13 has the strongest effect on the substrate spectra and is known as the main interaction site with the F-actin cytoskeleton. We find significant differences regarding the autophosphorylation state of CaMKII itself (Table 3). However, these do not strictly correlate with the length of the linker, nor the *in vitro* activity assays. The activating autophosphorylation is enriched in the longer-linker isoforms, whereas the inhibitory phosphorylations are

enriched in exon 13-exclusion isoforms. The similarity between the substrate spectra of the FL and Δ16,17 isoforms is largely based on CaMKII autophosphorylation, whereas the similarity between the Δ13 and Δ13,16 isoforms is also based on additional targets (Figure S11). In general, a key limitation of the technique is the limited sampling of a large space of potential CaMKII targets. As a multifunctional kinase, CaMKII is known to phosphorylate a great variety of protein substrates and any two random samples out of this large space are expected to show only a limited overlap. This might explain the absence of major enrichments in specific GO terms. However, manual inspection reveals some isoform-specific characteristics in targets associated with specific terms, such as “myelin sheath” (Table 4). We do not observe the presence or absence of specific shared targets for the exon 16-exclusion isoforms. Instead of suggesting that this primate-specifically regulated exon does not affect substrate specificity, it is more likely that the current assay lacks sensitivity and specificity to determine more subtle effects. Mainly, even though a mouse neuroblastoma cell line was used, N2a cells are still far from recapitulating actual synaptic connections. The absence of an enrichment for proteins involved in long-term plasticity, even for the combined or shared lists of CaMKII targets, strongly suggests that a more complex system is needed to fully mimic the native environment.

## A mouse model with humanized *Camk2β* splicing

Regardless of the exact mechanism, our mouse model clearly proves that modulating alternative splicing of *Camk2β* has drastic effects on the organism. This is, to our knowledge, the first mouse model in which a species-specific *cis*-regulatory element has been modulated to impact alternative splicing. Existing mouse models of CaMKII all feature rather drastic changes, like partial or complete gene knockout (KO) or global knockin (KI) of phospho-mutants (Elgersma *et al.*, 2004). Often, these models show compensatory effects, like altered localization or phosphorylation states of other CaMKII isoforms (Elgersma *et al.*, 2002; Yamasaki *et al.*, 2008). In our model with a humanized *Camk2β* splicing pattern, we see a clear and drastic impairment of LTP, indicating that the precise balance of CaMKIIβ isoforms is required for normal brain function and synaptic plasticity (Figure 16). This strongly suggests a functional relevance for different splice isoforms, even if the kinetics do not show robust differences. The observed effect on long-term potentiation (LTP) is not due to a general impairment of synaptic function, as short-term plasticity and basal synaptic

transmission appear unaffected. A similar trend can be observed in the standard CaMKII KO/KI mouse models, where basal synaptic transmission is often similarly unaffected (Elgersma *et al.*, 2004).

Both gene expression and alternative splicing are not globally affected in our mouse model. As discussed, CaMKIIβ is involved in structural LTP (sLTP), meaning the compositional and structural rearrangement of the synapse. As gene expression is not affected in the resting state, these effects seem to be predominately mediated by restructuring of pre-existing components of the synapse and postsynaptic density (PSD). The mass spectrometry data clearly shows that the autophosphorylation state differs between CaMKII isoforms, which is known to affect its interaction capabilities. Autophosphorylation regulates the interaction with the F-actin cytoskeleton (Lin and Redmond, 2008) and other components of the PSD like NMDA receptors (Bayer *et al.*, 2006; Elgersma *et al.*, 2002). One could envision a model in which CaMKII isoforms mediate localization differences to and within specialized compartments like the PSD. Due to the high abundance of the enzyme and the multivalency of its interaction sites in multimeric holoenzymes, CaMKII itself strongly affects the cytoskeleton and structure of the synapse. It also determines which substrates are in immediate close vicinity and can therefore be phosphorylated during a single transient calcium spike. Depending on the isoform, CaMKII acquires different autophosphorylations upon activation, which regulate the level and persistence of the generated autoactivity. These autophosphorylations might also depend on the exact cellular environment, as suggested by differences between our mass spectrometry results in permeabilized N2a cells and overexpressed and purified holoenzymes *in vitro* (Bhattacharyya *et al.*, 2020). The phosphorylation state of CaMKII also affects the localization changes following synaptic activity. CaMKII dissociates from F-actin, enabling structural rearrangement triggered by direct phosphorylation or indirect regulation of various GEFs and GAPs by CaMKII. With the exact timing depending on the persistence of certain acquired autophosphorylations, CaMKII reassociates with F-actin and locks the restructured synapse in this new state. The isoform-specific characteristics in the impact of different CaMKII variants seem to be of transient nature, as they do not lead to long-term differences in gene expression. The impact on LTP therefore seems to be entirely based on the modification and restructuring of pre-existing components of the synapse and therefore not lead to global changes in gene expression, at least in the

context of larger brain regions. Together, these results prove that the correct balance of CaMKII isoforms is essential for normal synaptic plasticity and brain function.

The phenotypes observed in our behavioural tests, namely the effects on locomotor activity and anxiety, are reminiscent of a previous CaMKIIβ knockout model (Bachstetter *et al.*, 2014). Interestingly, these animals showed reduced anxiety levels and diverse deficits in various locomotor tasks but did not show altered activity in an open field test. CaMKIIβ is expressed in muscle and endocrine tissue, were it presumably helps to target the holoenzyme to the sarcoplasmic reticulum (Bayer *et al.*, 1998; Urquidi and Ashcroft, 1995). Deficits in CaMKIIβ are therefore expected to globally affect motor function. The observed phenotypic differences between heterozygous and homozygous animals fit the hypothesis that the precise balance of splice isoforms is important. Both feature an altered splicing pattern but differ in the exact ratio of expressed splice isoforms. Based on the hypothesis that the correct balance of CaMKII variants is required for normal brain function, one would indeed expect different phenotypes depending on the exact distribution of splice variants. It is therefore this balance which affects behaviour and dictates the exact phenotype, and not just the presence or absence of specific variants.

Together, our mouse validates the hypothesis of functionally relevant species-specific differences in a *cis*-regulatory element, namely differences in branch point strength, which modulate *Camk2β* alternative splicing. It proves that the exact balance of CaMKIIβ variants is essential for normal brain function and perturbations of this equilibrium has profound repercussions on synaptic plasticity and behaviour.

# References

Ajith, S., Gazzara, M.R., Cole, B.S., Shankarling, G., Martinez, N.M., Mallory, M.J., and Lynch, K.W. (2016). Position-dependent activity of CELF2 in the regulation of splicing and implications for signal-responsive regulation in T cells. RNA Biology *13*, 569-581. 10.1080/15476286.2016.1176663.

Alt, F.W., Bothwell, A.L., Knapp, M., Siden, E., Mather, E., Koshland, M., and Baltimore, D. (1980). Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends. Cell *20*, 293-301. 10.1016/0092-8674(80)90615-7.

Araki, Y., Zeng, M., Zhang, M., and Huganir, Richard L. (2015). Rapid Dispersion of SynGAP from Synaptic Spines Triggers AMPA Receptor Insertion and Spine Enlargement during LTP. Neuron *85*, 173-189. https://doi.org/10.1016/j.neuron.2014.12.023.

Bachstetter, A.D., Webster, S.J., Tu, T., Goulding, D.S., Haiech, J., Watterson, D.M., and Van Eldik, L.J. (2014). Generation and behavior characterization of CaMKIIβ knockout mice. PLoS One *9*, e105191. 10.1371/journal.pone.0105191.

Banko, M.R., Allen, J.J., Schaffer, B.E., Wilker, E.W., Tsou, P., White, J.L., Villén, J., Wang, B., Kim, S.R., Sakamoto, K., et al. (2011). Chemical genetic screen for AMPKα2 substrates uncovers a network of proteins involved in mitosis. Mol Cell *44*, 878-892. 10.1016/j.molcel.2011.11.005.

Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., and Frey, B.J. (2010). Deciphering the splicing code. Nature *465*, 53-59. 10.1038/nature09000.

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., et al. (2012). The evolutionary landscape of alternative splicing in vertebrate species. Science *338*, 1587-1593. 10.1126/science.1230612.

Barilari, M., Bonfils, G., Treins, C., Koka, V., De Villeneuve, D., Fabrega, S., and Pende, M. (2017). ZRF1 is a novel S6 kinase substrate that drives the senescence programme. The EMBO Journal *36*, 736-750. https://doi.org/10.15252/embj.201694966.

Bautista, D.M., Siemens, J., Glazer, J.M., Tsuruda, P.R., Basbaum, A.I., Stucky, C.L., Jordt, S.-E., and Julius, D. (2007). The menthol receptor TRPM8 is the principal detector of environmental cold. Nature *448*, 204-208. 10.1038/nature05910.

Bayer, K.U., De Koninck, P., Leonard, A.S., Hell, J.W., and Schulman, H. (2001). Interaction with the NMDA receptor locks CaMKII in an active conformation. Nature *411*, 801-805. 10.1038/35081080.

Bayer, K.U., De Koninck, P., and Schulman, H. (2002). Alternative splicing modulates the frequency-dependent response of CaMKII to Ca(2+) oscillations. Embo j *21*, 3590-3597. 10.1093/emboj/cdf360.

Bayer, K.U., Harbers, K., and Schulman, H. (1998). alphaKAP is an anchoring protein for a novel CaM kinase II isoform in skeletal muscle. Embo j *17*, 5598-5605. 10.1093/emboj/17.19.5598.

Bayer, K.U., LeBel, É., McDonald, G.L., O'Leary, H., Schulman, H., and De Koninck, P. (2006). Transition from Reversible to Persistent Binding of CaMKII to Postsynaptic Sites and NR2B. The Journal of Neuroscience *26*, 1164-1174. 10.1523/jneurosci.3116-05.2006.

Bayer, K.U., Löhler, J., and Harbers, K. (1996). An alternative, nonkinase product of the brain-specifically expressed Ca2+/calmodulin-dependent kinase II alpha isoform gene in skeletal muscle. Mol Cell Biol *16*, 29-36. 10.1128/mcb.16.1.29.

Bayer, K.U., and Schulman, H. (2019). CaM Kinase: Still Inspiring at 40. Neuron *103*, 380-394. https://doi.org/10.1016/j.neuron.2019.05.033.

Beaman-Hall, C.M., Hozza, M.J., and Vallano, M.L. (1992). Detection of mRNAs encoding distinct isoenzymes of type II calcium/calmodulin-dependent protein kinase using the polymerase chain reaction. J Neurochem *58*, 1259-1267. 10.1111/j.1471-4159.1992.tb11337.x.

Ben-Shaul, Y. (2017). OptiMouse: a comprehensive open source program for reliable detection and analysis of mouse body and nose positions. BMC Biology *15*, 41. 10.1186/s12915-017-0377-3.

Berchtold, M.W., Munk, M., Kulej, K., Porth, I., Lorentzen, L., Panina, S., Zacharias, T., Larsen, M.R., and la Cour, J.M. (2021). The heart arrhythmia-linked D130G calmodulin mutation causes premature inhibitory autophosphorylation of CaMKII. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1868*, 119119. https://doi.org/10.1016/j.bbamcr.2021.119119.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proceedings of the National Academy of Sciences *74*, 3171-3175. doi:10.1073/pnas.74.8.3171.

Bhattacharyya, M., Lee, Y.K., Muratcioglu, S., Qiu, B., Nyayapati, P., Schulman, H., Groves, J.T., and Kuriyan, J. (2020). Flexible linkers in CaMKII control the balance between activating and inhibitory autophosphorylation. Elife *9*. 10.7554/eLife.53670.

Bhattacharyya, M., Stratton, M.M., Going, C.C., McSpadden, E.D., Huang, Y., Susa, A.C., Elleman, A., Cao, Y.M., Pappireddi, N., Burkhardt, P., et al. (2016). Molecular mechanism of activation-triggered subunit exchange in Ca(2+)/calmodulin-dependent protein kinase II. Elife *5*. 10.7554/eLife.13405.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem *72*, 291-336. 10.1146/annurev.biochem.72.121801.161720.

Borgesius, N.Z., van Woerden, G.M., Buitendijk, G.H.S., Keijzer, N., Jaarsma, D., Hoogenraad, C.C., and Elgersma, Y. (2011). βCaMKII Plays a Nonenzymatic Role in Hippocampal Synaptic Plasticity and Learning by Targeting αCaMKII to Synapses. The Journal of Neuroscience *31*, 10141-10148. 10.1523/jneurosci.5105-10.2011.

Bosch, M., and Hayashi, Y. (2012). Structural plasticity of dendritic spines. Curr Opin Neurobiol *22*, 383-388. 10.1016/j.conb.2011.09.002.

Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, Brenton R., and Blencowe, Benjamin J. (2013). Dynamic Integration of Splicing within Gene Regulatory Pathways. Cell *152*, 1252-1269. https://doi.org/10.1016/j.cell.2013.02.034.

Breen, M.A., and Ashcroft, S.J.H. (1997). A truncated isoform of Ca2+/calmodulin-dependent protein kinase II expressed in human islets of Langerhans may result from trans-splicing. FEBS Letters *409*, 375-379. https://doi.org/10.1016/S0014-5793(97)00555-3.

Brocke, L., Chiang, L.W., Wagner, P.D., and Schulman, H. (1999). Functional implications of the subunit composition of neuronal CaM kinase II. J Biol Chem *274*, 22713-22722. 10.1074/jbc.274.32.22713.

Brocke, L., Srinivasan, M., and Schulman, H. (1995). Developmental and regional expression of multifunctional Ca2+/calmodulin-dependent protein kinase isoforms in rat brain. J Neurosci *15*, 6797-6808. 10.1523/jneurosci.15-10-06797.1995.

Brooks, A.N., Yang, L., Duff, M.O., Hansen, K.D., Park, J.W., Dudoit, S., Brenner, S.E., and Graveley, B.R. (2011). Conservation of an RNA regulatory map between Drosophila and mammals. Genome Research *21*, 193-202. 10.1101/gr.108662.110.

Buljan, M., Chalancon, G., Eustermann, S., Wagner, Gunter P., Fuxreiter, M., Bateman, A., and Babu, M.M. (2012). Tissue-Specific Splicing of Disordered Segments that Embed Binding Motifs Rewires Protein Interaction Networks. Molecular Cell *46*, 871-883. https://doi.org/10.1016/j.molcel.2012.05.039.

Buonarati, O.R., Miller, A.P., Coultrap, S.J., Bayer, K.U., and Reichow, S.L. (2021). Conserved and divergent features of neuronal CaMKII holoenzyme structure, function, and high-order assembly. Cell Reports *37*, 110168. https://doi.org/10.1016/j.celrep.2021.110168.

Burnette, J.M., Miyamoto-Sato, E., Schaub, M.A., Conklin, J., and Lopez, A.J. (2005). Subdivision of large introns in Drosophila by recursive splicing at nonexonic elements. Genetics *170*, 661-674. 10.1534/genetics.104.039701.

Carelli, F.N., Liechti, A., Halbert, J., Warnefors, M., and Kaessmann, H. (2018). Repurposing of promoters and enhancers during mammalian evolution. Nat Commun *9*, 4066. 10.1038/s41467-018-06544-z.

Carlin, R.K., Grab, D.J., Cohen, R.S., and Siekevitz, P. (1980). Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. Journal of Cell Biology *86*, 831-845. 10.1083/jcb.86.3.831.

Carlton, S.M. (2002). Localization of CaMKIIα in rat primary sensory neurons: increase in inflammation. Brain Research *947*, 252-259. https://doi.org/10.1016/S0006-8993(02)02932-3.

Casañal, A., Lohkamp, B., and Emsley, P. (2020). Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. Protein Science *29*, 1055-1064. https://doi.org/10.1002/pro.3791.

Chang, J.-Y., Parra-Bueno, P., Laviv, T., Szatmari, E.M., Lee, S.-J.R., and Yasuda, R. (2017). CaMKII Autophosphorylation Is Necessary for Optimal Integration of Ca2+ Signals during LTP Induction, but Not Maintenance. Neuron *94*, 800-808.e804. https://doi.org/10.1016/j.neuron.2017.04.041.

Chao, L.H., Stratton, M.M., Lee, I.H., Rosenberg, O.S., Levitz, J., Mandell, D.J., Kortemme, T., Groves, J.T., Schulman, H., and Kuriyan, J. (2011). A mechanism for tunable autoinhibition in the structure of a human Ca2+/calmodulin- dependent kinase II holoenzyme. Cell *146*, 732-745. 10.1016/j.cell.2011.07.038.

Chaudhary, A., Brugge, J.S., and Cooper, J.A. (2002). Direct phosphorylation of focal adhesion kinase by c-Src: evidence using a modified nucleotide pocket kinase and ATP analog. Biochem Biophys Res Commun *294*, 293-300. 10.1016/s0006-291x(02)00475-8.

Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nature Reviews Molecular Cell Biology *10*, 741-754. 10.1038/nrm2777.

Chen, X., Vinade, L., Leapman, R.D., Petersen, J.D., Nakagawa, T., Phillips, T.M., Sheng, M., and Reese, T.S. (2005). Mass of the postsynaptic density and enumeration of three key molecules. Proceedings of the National Academy of Sciences *102*, 11551-11556. doi:10.1073/pnas.0505359102.

Cheng, D., Hoogenraad, C.C., Rush, J., Ramm, E., Schlager, M.A., Duong, D.M., Xu, P., Wijayawardana, S.R., Hanfelt, J., Nakagawa, T., et al. (2006). Relative and Absolute Quantification of Postsynaptic Density Proteome Isolated from Rat Forebrain and Cerebellum*S. Molecular & Cellular Proteomics *5*, 1158-1170. https://doi.org/10.1074/mcp.D500009-MCP200.

Cho, M.H., Cao, X., Wang, D., and Tsien, J.Z. (2007). Dentate gyrus-specific manipulation of β-Ca/calmodulin-dependent kinase II disrupts memory consolidation. Proceedings of the National Academy of Sciences *104*, 16317-16322. doi:10.1073/pnas.0703344104.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell *12*, 1-8. 10.1016/0092-8674(77)90180-5.

Cohen, S.M., Suutari, B., He, X., Wang, Y., Sanchez, S., Tirko, N.N., Mandelberg, N.J., Mullins, C., Zhou, G., Wang, S., et al. (2018). Calmodulin shuttling mediates cytonuclear signaling to trigger experience-dependent transcription and memory. Nature Communications *9*, 2451. 10.1038/s41467-018-04705-8.

Colbran, R.J., and Brown, A.M. (2004). Calcium/calmodulin-dependent protein kinase II and synaptic plasticity. Current Opinion in Neurobiology *14*, 318-327. https://doi.org/10.1016/j.conb.2004.05.008.

Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. Genes Dev *11*, 2755-2766. 10.1101/gad.11.21.2755.

Cook, S.G., Bourke, A.M., O'Leary, H., Zaegel, V., Lasda, E., Mize-Berge, J., Quillinan, N., Tucker, C.L., Coultrap, S.J., Herson, P.S., and Bayer, K.U. (2018). Analysis of the CaMKIIα and β splice-variant distribution among brain regions reveals isoform-specific differences in holoenzyme formation. Sci Rep *8*, 5448. 10.1038/s41598-018-23779-4.

Cooper, T.A. (2005). Use of minigene systems to dissect alternative splicing elements. Methods *37*, 331-340. 10.1016/j.ymeth.2005.07.015.

Corvelo, A., Hallegger, M., Smith, C.W., and Eyras, E. (2010). Genome-wide association between branch point properties and alternative splicing. PLoS Comput Biol *6*, e1001016. 10.1371/journal.pcbi.1001016.

Coultrap, S.J., and Bayer, K.U. (2012). Ca2+/Calmodulin-Dependent Protein Kinase II (CaMKII). In Protein Kinase Technologies, H. Mukai, ed. (Humana Press), pp. 49-72. 10.1007/978-1-61779-824-5_4.

Coultrap, S.J., and Bayer, K.U. (2014). Nitric oxide induces Ca2+-independent activity of the Ca2+/calmodulin-dependent protein kinase II (CaMKII). J Biol Chem *289*, 19458-19465. 10.1074/jbc.M114.558254.

Coultrap, S.J., Zaegel, V., and Bayer, K.U. (2014). CaMKII isoforms differ in their specific requirements for regulation by nitric oxide. FEBS Letters *588*, 4672-4676. https://doi.org/10.1016/j.febslet.2014.10.039.

Csuros, M., Rogozin, I.B., and Koonin, E.V. (2011). A Detailed History of Intron-rich Eukaryotic Ancestors Inferred from a Global Survey of 100 Complete Genomes. PLOS Computational Biology *7*, e1002150. 10.1371/journal.pcbi.1002150.

D'Amours, G., Bureau, G., Boily, M.-J., and Cyr, M. (2011). Differential gene expression profiling in the mouse brain during motor skill learning: Focus on the striatum structure. Behavioural Brain Research *221*, 108-117. https://doi.org/10.1016/j.bbr.2011.02.030.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*. 10.1093/gigascience/giab008.

De Conti, L., Baralle, M., and Buratti, E. (2013). Exon and intron definition in pre-mRNA splicing. WIREs RNA *4*, 49-60. https://doi.org/10.1002/wrna.1140.

De Koninck, P., and Schulman, H. (1998). Sensitivity of CaM kinase II to the frequency of Ca2+ oscillations. Science *279*, 227-230. 10.1126/science.279.5348.227.

de Ligt, J., Willemsen, M.H., van Bon, B.W.M., Kleefstra, T., Yntema, H.G., Kroes, T., Vulto-van Silfhout, A.T., Koolen, D.A., de Vries, P., Gilissen, C., et al. (2012). Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. New England Journal of Medicine *367*, 1921-1929. 10.1056/NEJMoa1206524.

Deacon, R.M.J. (2006). Assessing nest building in mice. Nature Protocols *1*, 1117-1119. 10.1038/nprot.2006.170.

Derkach, V., Barria, A., and Soderling, T.R. (1999). Ca2+/calmodulin-kinase II enhances channel conductance of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionate type glutamate receptors. Proceedings of the National Academy of Sciences *96*, 3269-3274. doi:10.1073/pnas.96.6.3269.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England) *29*, 15-21. 10.1093/bioinformatics/bts635.

Duff, M.O., Olson, S., Wei, X., Garrett, S.C., Osman, A., Bolisetty, M., Plocik, A., Celniker, S.E., and Graveley, B.R. (2015). Genome-wide identification of zero nucleotide recursive splicing in Drosophila. Nature *521*, 376-379. 10.1038/nature14475.

Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R., and Hood, L. (1980). Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. Cell *20*, 313-319. 10.1016/0092-8674(80)90617-0.

Elgersma, Y., Fedorov, N.B., Ikonen, S., Choi, E.S., Elgersma, M., Carvalho, O.M., Giese, K.P., and Silva, A.J. (2002). Inhibitory autophosphorylation of CaMKII controls PSD association, plasticity, and learning. Neuron *36*, 493-505. 10.1016/s0896-6273(02)01007-3.

Elgersma, Y., Sweatt, J.D., and Giese, K.P. (2004). Mouse genetic approaches to investigating calcium/calmodulin-dependent protein kinase II function in plasticity and cognition. J Neurosci *24*, 8410-8415. 10.1523/jneurosci.3622-04.2004.

Ellis, Jonathan D., Barrios-Rodiles, M., Çolak, R., Irimia, M., Kim, T., Calarco, John A., Wang, X., Pan, Q., O'Hanlon, D., Kim, Philip M., et al. (2012). Tissue-Specific Alternative Splicing Remodels Protein-Protein Interaction Networks. Molecular Cell *46*, 884-892. https://doi.org/10.1016/j.molcel.2012.05.037.

Erickson, J.R., Joiner, M.-l.A., Guan, X., Kutschke, W., Yang, J., Oddis, C.V., Bartlett, R.K., Lowe, J.S., O'Donnell, S.E., Aykin-Burns, N., et al. (2008). A Dynamic Pathway for Calcium-Independent Activation of CaMKII by Methionine Oxidation. Cell *133*, 462-474. https://doi.org/10.1016/j.cell.2008.02.048.

Erickson, J.R., Nichols, C.B., Uchinoumi, H., Stein, M.L., Bossuyt, J., and Bers, D.M. (2015). S-Nitrosylation Induces Both Autonomous Activation and Inhibition of Calcium/Calmodulin-dependent Protein Kinase II δ*. Journal of Biological Chemistry *290*, 25646-25656. https://doi.org/10.1074/jbc.M115.650234.

Erickson, J.R., Pereira, L., Wang, L., Han, G., Ferguson, A., Dao, K., Copeland, R.J., Despa, F., Hart, G.W., Ripplinger, C.M., and Bers, D.M. (2013). Diabetic hyperglycaemia activates CaMKII and arrhythmias by O-linked glycosylation. Nature *502*, 372-376. 10.1038/nature12537.

Erkelenz, S., Mueller, W.F., Evans, M.S., Busch, A., Schöneweis, K., Hertel, K.J., and Schaal, H. (2013). Position-dependent splicing activation and repression by SR and

hnRNP proteins rely on common mechanisms. Rna *19*, 96-102. 10.1261/rna.037044.112.

Erondu, N.E., and Kennedy, M.B. (1985). Regional distribution of type II Ca2+/calmodulin-dependent protein kinase in rat brain. J Neurosci *5*, 3270-3277. 10.1523/jneurosci.05-12-03270.1985.

Ewing, B., and Green, P. (2000). Analysis of expressed sequence tags indicates 35,000 human genes. Nature Genetics *25*, 232-234. 10.1038/76115.

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2009). The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. Mol Cell *36*, 593-608. 10.1016/j.molcel.2009.09.040.

Fink, C.C., Bayer, K.-U., Myers, J.W., Ferrell, J.E., Schulman, H., and Meyer, T. (2003). Selective Regulation of Neurite Extension and Synapse Formation by the β but not the α Isoform of CaMKII. Neuron *39*, 283-297. https://doi.org/10.1016/S0896-6273(03)00428-8.

Fu, X.-D., and Ares, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. Nature Reviews Genetics *15*, 689-701. 10.1038/nrg3778.

Galarneau, A., and Richard, S. (2005). Target RNA motif and target mRNAs of the Quaking STAR protein. Nature Structural & Molecular Biology *12*, 691-698. 10.1038/nsmb963.

Gallo, C.M., Labadorf, A.T., Ho, A., and Beffert, U. (2022). Single molecule, long-read Apoer2 sequencing identifies conserved and species-specific splicing patterns. Genomics *114*, 110318. 10.1016/j.ygeno.2022.110318.

García-Blanco, M.A., Jamison, S.F., and Sharp, P.A. (1989). Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. Genes Dev *3*, 1874-1886. 10.1101/gad.3.12a.1874.

Ghosh, A., and Giese, K.P. (2015). Calcium/calmodulin-dependent kinase II and Alzheimer's disease. Molecular Brain *8*, 78. 10.1186/s13041-015-0166-2.

Giese, K.P., Fedorov, N.B., Filipkowski, R.K., and Silva, A.J. (1998). Autophosphorylation at Thr<sup>286</sup> of the α Calcium-Calmodulin Kinase II in LTP and Learning. Science *279*, 870-873. doi:10.1126/science.279.5352.870.

Gong, J., Liu, J., Ronan, E.A., He, F., Cai, W., Fatima, M., Zhang, W., Lee, H., Li, Z., Kim, G.-H., et al. (2019). A Cold-Sensing Receptor Encoded by a Glutamate Receptor Gene. Cell *178*, 1375-1386.e1311. https://doi.org/10.1016/j.cell.2019.07.034.

Gracheva, E.O., Cordero-Morales, J.F., González-Carcacía, J.A., Ingolia, N.T., Manno, C., Aranguren, C.I., Weissman, J.S., and Julius, D. (2011). Ganglion-specific splicing of TRPV1 underlies infrared sensation in vampire bats. Nature *476*, 88-91. 10.1038/nature10245.

Grant, B.D., and Adams, J.A. (1996). Pre-Steady-State Kinetic Analysis of cAMP-Dependent Protein Kinase Using Rapid Quench Flow Techniques. Biochemistry *35*, 2022-2029. 10.1021/bi952144+.

Grau-Bové, X., Ruiz-Trillo, I., and Irimia, M. (2018). Origin of exon skipping-rich transcriptomes in animals driven by evolution of gene architecture. Genome Biology *19*, 135. 10.1186/s13059-018-1499-9.

Graveley, B.R. (2008). The haplo-spliceo-transcriptome: common variations in alternative splicing in the human population. Trends Genet *24*, 5-7. 10.1016/j.tig.2007.10.004.

Griffith, L.C., and Greenspan, R.J. (1993). The diversity of calcium/calmodulin-dependent protein kinase II isoforms in Drosophila is generated by alternative splicing of a single gene. J Neurochem *61*, 1534-1537. 10.1111/j.1471-4159.1993.tb13650.x.

Grønning, Alexander Gulliver B., Doktor, T.K., Larsen, Simon J., Petersen, Ulrika Simone S., Holm, L.L., Bruun, Gitte H., Hansen, M.B., Hartung, A.-M., Baumbach, J., and Andresen, B.S. (2020). DeepCLIP: predicting the effect of mutations on protein–RNA binding with deep learning. Nucleic Acids Research *48*, 7099-7118. 10.1093/nar/gkaa530.

Gueroussov, S., Gonatopoulos-Pournatzis, T., Irimia, M., Raj, B., Lin, Z.Y., Gingras, A.C., and Blencowe, B.J. (2015). An alternative splicing event amplifies evolutionary differences between vertebrates. Science *349*, 868-873. 10.1126/science.aaa8381.

GuptaRoy, B., Marwaha, N., Pla, M., Wang, Z., Nelson, H.B., Beckingham, K., and Griffith, L.C. (2000). Alternative splicing of Drosophila calcium/calmodulin-dependent protein kinase II regulates substrate specificity and activation. Brain Res Mol Brain Res *80*, 26-34. 10.1016/s0169-328x(00)00115-7.

Haltenhof, T., Kotte, A., De Bortoli, F., Schiefer, S., Meinke, S., Emmerichs, A.K., Petermann, K.K., Timmermann, B., Imhof, P., Franz, A., et al. (2020). A Conserved Kinase-Based Body-Temperature Sensor Globally Controls Alternative Splicing and Gene Expression. Mol Cell *78*, 57-69.e54. 10.1016/j.molcel.2020.01.028.

Hanson, P.I., Meyer, T., Stryer, L., and Schulman, H. (1994). Dual role of calmodulin in autophosphorylation of multifunctional CaM kinase may underlie decoding of calcium signals. Neuron *12*, 943-956. 10.1016/0896-6273(94)90306-9.

Hashimoto, Y., and Soderling, T.R. (1987). Calcium . calmodulin-dependent protein kinase II and calcium . phospholipid-dependent protein kinase activities in rat tissues assayed with a synthetic peptide. Arch Biochem Biophys *252*, 418-425. 10.1016/0003-9861(87)90048-8.

Hatton, A.R., Subramaniam, V., and Lopez, A.J. (1998). Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon-exon junctions. Mol Cell *2*, 787-796. 10.1016/s1097-2765(00)80293-2.

Hell, J.W. (2014). CaMKII: claiming center stage in postsynaptic function and organization. Neuron *81*, 249-265. 10.1016/j.neuron.2013.12.024.

Herring, B.E., and Nicoll, R.A. (2016). Long-Term Potentiation: From CaMKII to AMPA Receptor Trafficking. Annual Review of Physiology *78*, 351-365. 10.1146/annurev-physiol-021014-071753.

Hillier, L.W., Reinke, V., Green, P., Hirst, M., Marra, M.A., and Waterston, R.H. (2009). Massively parallel sequencing of the polyadenylated transcriptome of C. elegans. Genome Res *19*, 657-666. 10.1101/gr.088112.108.

Hoffman, L., Farley, M.M., and Waxham, M.N. (2013). Calcium-calmodulin-dependent protein kinase II isoforms differentially impact the dynamics and structure of the actin cytoskeleton. Biochemistry *52*, 1198-1207. 10.1021/bi3016586.

Hudmon, A., and Schulman, H. (2002). Neuronal CA2+/calmodulin-dependent protein kinase II: the role of structure and autoregulation in cellular function. Annu Rev Biochem *71*, 473-510. 10.1146/annurev.biochem.71.110601.135410.

Irvine, E.E., Vernon, J., and Giese, K.P. (2005). αCaMKII autophosphorylation contributes to rapid learning but is not necessary for memory. Nature Neuroscience *8*, 411-412. 10.1038/nn1431.

Jalan-Sakrikar, N., Bartlett, R.K., Baucum, A.J., 2nd, and Colbran, R.J. (2012). Substrate-selective and calcium-independent activation of CaMKII by α-actinin. J Biol Chem *287*, 15275-15283. 10.1074/jbc.M112.351817.

Kaiser, T.S., Poehn, B., Szkiba, D., Preussner, M., Sedlazeck, F.J., Zrim, A., Neumann, T., Nguyen, L.-T., Betancourt, A.J., Hummel, T., et al. (2016). The genomic basis of circadian and circalunar timing adaptations in a midge. Nature *540*, 69-73. 10.1038/nature20151.

Kannan, S., Halter, G., Renner, T., and Waters, E.R. (2018). Patterns of alternative splicing vary between species during heat stress. AoB Plants *10*, ply013. 10.1093/aobpla/ply013.

Kastner, B., Will, C.L., Stark, H., and Lührmann, R. (2019). Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes. Cold Spring Harb Perspect Biol *11*. 10.1101/cshperspect.a032417.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet *11*, 345-355. 10.1038/nrg2776.

Khan, S., Downing, K.H., and Molloy, J.E. (2019). Architectural Dynamics of CaMKII-Actin Networks. Biophysical Journal *116*, 104-119. 10.1016/j.bpj.2018.11.006.

Kim, E., Magen, A., and Ast, G. (2006). Different levels of alternative splicing among eukaryotes. Nucleic Acids Research *35*, 125-131. 10.1093/nar/gkl924.

Kolb, S.J., Hudmon, A., Ginsberg, T.R., and Waxham, M.N. (1998). Identification of domains essential for the assembly of calcium/calmodulin-dependent protein kinase II holoenzymes. J Biol Chem *273*, 31555-31564. 10.1074/jbc.273.47.31555.

Kristensen, A.S., Jenkins, M.A., Banke, T.G., Schousboe, A., Makino, Y., Johnson, R.C., Huganir, R., and Traynelis, S.F. (2011). Mechanism of Ca2+/calmodulin-dependent kinase II regulation of AMPA receptor gating. Nature Neuroscience *14*, 727-735. 10.1038/nn.2804.

Küry, S., van Woerden, G.M., Besnard, T., Proietti Onori, M., Latypova, X., Towne, M.C., Cho, M.T., Prescott, T.E., Ploeg, M.A., Sanders, S., et al. (2017). De Novo Mutations in Protein Kinase Genes CAMK2A and CAMK2B Cause Intellectual Disability. The American Journal of Human Genetics *101*, 768-788. https://doi.org/10.1016/j.ajhg.2017.10.003.

Lee, Y., and Rio, D.C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu Rev Biochem *84*, 291-323. 10.1146/annurev-biochem-060614-034316.

Leonard, A.S., Lim, I.A., Hemsworth, D.E., Horne, M.C., and Hell, J.W. (1999). Calcium/calmodulin-dependent protein kinase II is associated with the N-methyl-D-aspartate receptor. Proceedings of the National Academy of Sciences *96*, 3239-3244. doi:10.1073/pnas.96.6.3239.

Lev-Maor, G., Sorek, R., Shomron, N., and Ast, G. (2003). The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. Science *300*, 1288-1291. 10.1126/science.1082588.

Li, K., Zhou, T., Liao, L., Yang, Z., Wong, C., Henn, F., Malinow, R., Yates, J.R., and Hu, H. (2013). βCaMKII in Lateral Habenula Mediates Core Symptoms of Depression. Science *341*, 1016-1020. 10.1126/science.1240729.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol *75*, 861-877. 10.1107/s2059798319011471.

Lin, S., Lin, Y., Nery, J.R., Urich, M.A., Breschi, A., Davis, C.A., Dobin, A., Zaleski, C., Beer, M.A., Chapman, W.C., et al. (2014). Comparison of the transcriptional landscapes between human and mouse tissues. Proceedings of the National Academy of Sciences *111*, 17224-17229. 10.1073/pnas.1413624111.

Lin, Y.-C., and Redmond, L. (2008). CaMKIIb binding to stable F-actin *in vivo* regulates F-actin filament stability. Proceedings of the National Academy of Sciences *105*, 15791-15796. doi:10.1073/pnas.0804399105.

Liu, M., and Grigoriev, A. (2004). Protein domains correlate strongly with exons in multiple eukaryotic genomes--evidence of exon shuffling? Trends Genet *20*, 399-403. 10.1016/j.tig.2004.06.013.

Lopez, M.S., Kliegman, J.I., and Shokat, K.M. (2014). Chapter Eight - The Logic and Design of Analog-Sensitive Kinases and Their Small Molecule Inhibitors. In Methods in Enzymology, K.M. Shokat, ed. (Academic Press), pp. 189-213. https://doi.org/10.1016/B978-0-12-397918-6.00008-2.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550. 10.1186/s13059-014-0550-8.

Maas, N.L., Singh, N., and Diehl, J.A. (2014). Generation and characterization of an analog-sensitive PERK allele. Cancer Biol Ther *15*, 1106-1111. 10.4161/cbt.29274.

Malenka, R.C., and Bear, M.F. (2004). LTP and LTD: an embarrassment of riches. Neuron *44*, 5-21. 10.1016/j.neuron.2004.09.012.

Mayford, M., Baranes, D., Podsypanina, K., and Kandel, E.R. (1996). The 3'-untranslated region of CaMKIIa is a cis-acting signal for the localization and translation of mRNA in dendrites. Proceedings of the National Academy of Sciences *93*, 13250-13255. doi:10.1073/pnas.93.23.13250.

McSpadden, E.D., Xia, Z., Chi, C.C., Susa, A.C., Shah, N.H., Gee, C.L., Williams, E.R., and Kuriyan, J. (2019). Variation in assembly stoichiometry in non-metazoan homologs of the hub domain of Ca(2+) /calmodulin-dependent protein kinase II. Protein Sci *28*, 1071-1082. 10.1002/pro.3614.

Mercer, T.R., Clark, M.B., Andersen, S.B., Brunck, M.E., Haerty, W., Crawford, J., Taft, R.J., Nielsen, L.K., Dinger, M.E., and Mattick, J.S. (2015). Genome-wide discovery of human splicing branchpoints. Genome Res *25*, 290-303. 10.1101/gr.182899.114.

Merkin, J., Russell, C., Chen, P., and Burge, C.B. (2012). Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. Science *338*, 1593-1599. 10.1126/science.1228186.

Meyer, T., Hanson, P.I., Stryer, L., and Schulman, H. (1992). Calmodulin trapping by calcium-calmodulin-dependent protein kinase. Science *256*, 1199-1202. 10.1126/science.256.5060.1199.

Michowski, W., Chick, J.M., Chu, C., Kolodziejczyk, A., Wang, Y., Suski, J.M., Abraham, B., Anders, L., Day, D., Dunkl, L.M., et al. (2020). Cdk1 Controls Global Epigenetic Landscape in Embryonic Stem Cells. Mol Cell *78*, 459-476.e413. 10.1016/j.molcel.2020.03.010.

Miller, S.G., and Kennedy, M.B. (1985). Distinct forebrain and cerebellar isozymes of type II Ca2+/calmodulin-dependent protein kinase associate differently with the postsynaptic density fraction. J Biol Chem *260*, 9039-9046.

Modrek, B., and Lee, C.J. (2003). Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. Nature Genetics *34*, 177-180. 10.1038/ng1159.

Mohamed, B.A., Elkenani, M., Jakubiczka-Smorag, J., Buchholz, E., Koszewa, S., Lbik, D., Schnelle, M., Hasenfuss, G., and Toischer, K. (2019). Genetic deletion of calcium/calmodulin-dependent protein kinase type II delta does not mitigate adverse myocardial remodeling in volume-overloaded hearts. Scientific Reports *9*, 9889. 10.1038/s41598-019-46332-3.

Mori, Y., Imaizumi, K., Katayama, T., Yoneda, T., and Tohyama, M. (2000). Two cis-acting elements in the 3′ untranslated region of α-CaMKII regulate its dendritic targeting. Nature Neuroscience *3*, 1079-1084. 10.1038/80591.

Myers, J.B., Zaegel, V., Coultrap, S.J., Miller, A.P., Bayer, K.U., and Reichow, S.L. (2017). The CaMKII holoenzyme structure in activation-competent conformations. Nat Commun *8*, 15742. 10.1038/ncomms15742.

Nazari, I., Tayara, H., and Chong, K.T. (2019). Branch Point Selection in RNA Splicing Using Deep Learning. IEEE Access *7*, 1800-1807. 10.1109/ACCESS.2018.2886569.

Nolen, B., Taylor, S., and Ghosh, G. (2004). Regulation of Protein Kinases: Controlling Activity through Activation Segment Conformation. Molecular Cell *15*, 661-675. https://doi.org/10.1016/j.molcel.2004.08.024.

O'Leary, H., Lasda, E., and Bayer, K.U. (2006). CaMKIIbeta association with the actin cytoskeleton is regulated by alternative splicing. Mol Biol Cell *17*, 4656-4665. 10.1091/mbc.e06-03-0252.

Okamoto, K.-I., Narayanan, R., Lee, S.H., Murata, K., and Hayashi, Y. (2007). The role of CaMKII as an F-actin-bundling protein crucial for maintenance of dendritic spine structure. Proceedings of the National Academy of Sciences *104*, 6418-6423. doi:10.1073/pnas.0701656104.

Omar Faison, M., Perozzi, E.F., Caran, N., Stewart, J.K., and Tombes, R.M. (2002). Axonal localization of delta Ca2+/calmodulin-dependent protein kinase II in developing P19 neurons. International Journal of Developmental Neuroscience *20*, 585-592. https://doi.org/10.1016/S0736-5748(02)00107-7.

Opazo, P., Labrecque, S., Tigaret, C.M., Frouin, A., Wiseman, P.W., De Koninck, P., and Choquet, D. (2010). CaMKII triggers the diffusional trapping of surface AMPARs through phosphorylation of stargazin. Neuron *67*, 239-252. 10.1016/j.neuron.2010.06.007.

Pai, A.A., Paggi, J.M., Yan, P., Adelman, K., and Burge, C.B. (2018). Numerous recursive sites contribute to accuracy of splicing in long introns in flies. PLOS Genetics *14*, e1007588. 10.1371/journal.pgen.1007588.

Pan, Q., Shai, O., Misquitta, C., Zhang, W., Saltzman, A.L., Mohammad, N., Babak, T., Siu, H., Hughes, T.R., Morris, Q.D., et al. (2004). Revealing Global Regulatory Features of Mammalian Alternative Splicing Using a Quantitative Microarray Platform. Molecular Cell *16*, 929-941. https://doi.org/10.1016/j.molcel.2004.12.004.

Park, J., Chávez, Andrés E., Mineur, Yann S., Morimoto-Tomita, M., Lutzu, S., Kim, Kwang S., Picciotto, Marina R., Castillo, Pablo E., and Tomita, S. (2016). CaMKII Phosphorylation of TARPγ-8 Is a Mediator of LTP and Learning and Memory. Neuron *92*, 75-83. https://doi.org/10.1016/j.neuron.2016.09.002.

Parra, M., Booth, B.W., Weiszmann, R., Yee, B., Yeo, G.W., Brown, J.B., Celniker, S.E., and Conboy, J.G. (2018). An important class of intron retention events in human erythroblasts is regulated by cryptic exons proposed to function as splicing decoys. RNA *24*, 1255-1265. 10.1261/rna.066951.118.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods *14*, 417-419. 10.1038/nmeth.4197.

Paz, I., Kosti, I., Ares, M., Jr., Cline, M., and Mandel-Gutfreund, Y. (2014). RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res *42*, W361-367. 10.1093/nar/gku406.

Petersen, J.D., Chen, X., Vinade, L., Dosemeci, A., Lisman, J.E., and Reese, T.S. (2003). Distribution of Postsynaptic Density (PSD)-95 and $Ca^{2+}$/Calmodulin-Dependent Protein Kinase II at the PSD. The Journal of Neuroscience *23*, 11270-11278. 10.1523/jneurosci.23-35-11270.2003.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci *30*, 70-82. 10.1002/pro.3943.

Picconi, B., Gardoni, F., Centonze, D., Mauceri, D., Cenci, M.A., Bernardi, G., Calabresi, P., and Di Luca, M. (2004). Abnormal Ca2+-calmodulin-dependent protein kinase II function mediates synaptic and motor deficits in experimental parkinsonism. J Neurosci *24*, 5283-5291. 10.1523/jneurosci.1224-04.2004.

Plaschka, C., Newman, A.J., and Nagai, K. (2019). Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast. Cold Spring Harb Perspect Biol *11*. 10.1101/cshperspect.a032391.

Preußner, M., Goldammer, G., Neumann, A., Haltenhof, T., Rautenstrauch, P., Müller-McNicoll, M., and Heyd, F. (2017). Body Temperature Cycles Control Rhythmic Alternative Splicing in Mammals. Mol Cell *67*, 433-446.e434. 10.1016/j.molcel.2017.06.006.

Preußner, M., Wilhelmi, I., Schultz, A.S., Finkernagel, F., Michel, M., Möröy, T., and Heyd, F. (2014). Rhythmic U2af26 alternative splicing controls PERIOD1 stability and the circadian clock in mice. Mol Cell *54*, 651-662. 10.1016/j.molcel.2014.04.015.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature Methods *14*, 290-296. 10.1038/nmeth.4169.

Puram, S.V., Kim, A.H., Ikeuchi, Y., Wilson-Grady, J.T., Merdes, A., Gygi, S.P., and Bonni, A. (2011). A CaMKIIβ signaling pathway at the centrosome regulates dendrite patterning in the brain. Nat Neurosci *14*, 973-983. 10.1038/nn.2857.

Quervain, D.J.-F.d., and Papassotiropoulos, A. (2006). Identification of a genetic cluster influencing memory performance and hippocampal activity in humans. Proceedings of the National Academy of Sciences *103*, 4270-4274. doi:10.1073/pnas.0510212103.

R Core Team (2021). R: A Language and Environment for Statistical Computing.

Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., et al. (2013). A compendium of RNA-binding motifs for decoding gene regulation. Nature *499*, 172-177. 10.1038/nature12311.

Reiner, D.J., Newton, E.M., Tian, H., and Thomas, J.H. (1999). Diverse behavioural defects caused by mutations in Caenorhabditis elegans unc-43 CaM kinase II. Nature *402*, 199-203. 10.1038/46072.

Rellos, P., Pike, A.C.W., Niesen, F.H., Salah, E., Lee, W.H., von Delft, F., and Knapp, S. (2010). Structure of the CaMKIIδ/Calmodulin Complex Reveals the Molecular Mechanism of CaMKII Kinase Activation. PLOS Biology *8*, e1000426. 10.1371/journal.pbio.1000426.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature biotechnology *29*, 24-26. 10.1038/nbt.1754.

Robison, A.J. (2014). Emerging role of CaMKII in neuropsychiatric disease. Trends in Neurosciences *37*, 653-662. https://doi.org/10.1016/j.tins.2014.07.001.

Rochlitz, H., Voigt, A., Lankat-Buttgereit, B., Göke, B., Heimberg, H., Nauck, M.A., Schiemann, U., Schatz, H., and Pfeiffer, A.F. (2000). Cloning and quantitative determination of the human Ca2+/calmodulin-dependent protein kinase II (CaMK II) isoforms in human beta cells. Diabetologia *43*, 465-473. 10.1007/s001250051330.

Rodriguez, J., Ren, G., Day, C.R., Zhao, K., Chow, C.C., and Larson, D.R. (2019). Intrinsic Dynamics of a Human Gene Reveal the Basis of Expression Heterogeneity. Cell *176*, 213-226.e218. https://doi.org/10.1016/j.cell.2018.11.026.

Rosenberg, O.S., Deindl, S., Comolli, L.R., Hoelz, A., Downing, K.H., Nairn, A.C., and Kuriyan, J. (2006). Oligomerization states of the association domain and the holoenyzme of Ca2+/CaM kinase II. The FEBS Journal *273*, 682-694. https://doi.org/10.1111/j.1742-4658.2005.05088.x.

Rosenberg, O.S., Deindl, S., Sung, R.J., Nairn, A.C., and Kuriyan, J. (2005). Structure of the autoinhibited kinase domain of CaMKII and SAXS analysis of the holoenzyme. Cell *123*, 849-860. 10.1016/j.cell.2005.10.029.

Rossbach, O., Hung, L.-H., Schreiner, S., Grishina, I., Heiner, M., Hui, J., and Bindereif, A. (2009). Auto- and Cross-Regulation of the hnRNP L Proteins by Alternative Splicing. Molecular and Cellular Biology *29*, 1442-1451. doi:10.1128/MCB.01689-08.

Rothschild, S.C., Ingram, S.R., Lu, F.I., Thisse, B., Thisse, C., Parkerson, J.A., and Tombes, R.M. (2020). Genetic compensation of γ CaMKII, an evolutionarily conserved gene. Gene *742*, 144567. 10.1016/j.gene.2020.144567.

RStudio Team (2021). RStudio: Integrated Development Environment for R.

Ryan, T.J., and Grant, S.G.N. (2009). The origin and evolution of synapses. Nature Reviews Neuroscience *10*, 701-712. 10.1038/nrn2717.

Sałaciak, K., Koszałka, A., Żmudzka, E., and Pytka, K. (2021). The Calcium/Calmodulin-Dependent Kinases II and IV as Therapeutic Targets in Neurodegenerative and Neuropsychiatric Disorders. Int J Mol Sci *22*. 10.3390/ijms22094307.

Saneyoshi, T., Matsuno, H., Suzuki, A., Murakoshi, H., Hedrick, N.G., Agnello, E., O'Connell, R., Stratton, M.M., Yasuda, R., and Hayashi, Y. (2019). Reciprocal Activation within a Kinase-Effector Complex Underlying Persistence of Structural LTP. Neuron *102*, 1199-1210.e1196. https://doi.org/10.1016/j.neuron.2019.04.012.

Schulman, H., and Greengard, P. (1978). Stimulation of brain membrane protein phosphorylation by calcium and an endogenous heat-stable protein. Nature *271*, 478-479. 10.1038/271478a0.

Shatkin, A.J. (1976). Capping of eucaryotic mRNAs. Cell *9*, 645-653. https://doi.org/10.1016/0092-8674(76)90128-8.

Shen, K., and Meyer, T. (1999). Dynamic control of CaMKII translocation and localization in hippocampal neurons by NMDA receptor stimulation. Science *284*, 162-166. 10.1126/science.284.5411.162.

Shen, K., Teruel, M.N., Connor, J.H., Shenolikar, S., and Meyer, T. (2000). Molecular memory by reversible translocation of calcium/calmodulin-dependent protein kinase II. Nature Neuroscience *3*, 881-886. 10.1038/78783.

Shen, K., Teruel, M.N., Subramanian, K., and Meyer, T. (1998). CaMKIIbeta functions as an F-actin targeting module that localizes CaMKIIalpha/beta heterooligomers to dendritic spines. Neuron *21*, 593-606. 10.1016/s0896-6273(00)80569-3.

Shen, S., Park, J.W., Lu, Z.-x., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proceedings of the National Academy of Sciences *111*, E5593-E5601. 10.1073/pnas.1419161111.

Sheng, M., and Hoogenraad, C.C. (2007). The Postsynaptic Architecture of Excitatory Synapses: A More Quantitative View. Annual Review of Biochemistry *76*, 823-847. 10.1146/annurev.biochem.76.060805.160029.

Shi, Y., Su, Z., Yang, H., Wang, W., Jin, G., He, G., Siddique, A.N., Zhang, L., Zhu, A., Xue, R., and Zhang, C. (2019). Alternative splicing coupled to nonsense-mediated mRNA decay contributes to the high-altitude adaptation of maca (Lepidium meyenii). Gene *694*, 7-18. 10.1016/j.gene.2018.12.082.

Sibley, C.R., Blazquez, L., and Ule, J. (2016). Lessons from non-canonical splicing. Nature Reviews Genetics *17*, 407-421. 10.1038/nrg.2016.46.

Sibley, C.R., Emmett, W., Blazquez, L., Faro, A., Haberman, N., Briese, M., Trabzuni, D., Ryten, M., Weale, M.E., Hardy, J., et al. (2015). Recursive splicing in long vertebrate genes. Nature *521*, 371-375. 10.1038/nature14466.

Silva, A.J., Paylor, R., Wehner, J.M., and Tonegawa, S. (1992). Impaired Spatial Learning in α-Calcium-Calmodulin Kinase II Mutant Mice. Science *257*, 206-211. doi:10.1126/science.1321493.

Sloutsky, R., Dziedzic, N., Dunn, M.J., Bates, R.M., Torres-Ocampo, A.P., Boopathy, S., Page, B., Weeks, J.G., Chao, L.H., and Stratton, M.M. (2020). Heterogeneity in human hippocampal CaMKII transcripts reveals allosteric hub-dependent regulation. Sci Signal *13*. 10.1126/scisignal.aaz0240.

Sloutsky, R., and Stratton, M.M. (2021). Functional implications of CaMKII alternative splicing. Eur J Neurosci *54*, 6780-6794. 10.1111/ejn.14761.

Smith, F.D., Reichow, S.L., Esseltine, J.L., Shi, D., Langeberg, L.K., Scott, J.D., and Gonen, T. (2013). Intrinsic disorder within an AKAP-protein kinase A complex guides local substrate phosphorylation. Elife *2*, e01319. 10.7554/eLife.01319.

Sohail, M., and Xie, J. (2015). Evolutionary Emergence of a Novel Splice Variant with an Opposite Effect on the Cell Cycle. Molecular and Cellular Biology *35*, 2203-2214. doi:10.1128/MCB.00190-15.

Srinivasan, M., Edman, C.F., and Schulman, H. (1994). Alternative splicing introduces a nuclear localization signal that targets multifunctional CaM kinase to the nucleus. J Cell Biol *126*, 839-852. 10.1083/jcb.126.4.839.

Stolc, V., Gauhar, Z., Mason, C., Halasz, G., van Batenburg, M.F., Rifkin, S.A., Hua, S., Herreman, T., Tongprasit, W., Barbano, P.E., et al. (2004). A gene expression map for the euchromatic genome of Drosophila melanogaster. Science *306*, 655-660. 10.1126/science.1101312.

Stratton, M., Lee, I.H., Bhattacharyya, M., Christensen, S.M., Chao, L.H., Schulman, H., Groves, J.T., and Kuriyan, J. (2014). Activation-triggered subunit exchange between CaMKII holoenzymes facilitates the spread of kinase activity. Elife *3*, e01610. 10.7554/eLife.01610.

Südhof, T.C. (2018). Towards an Understanding of Synapse Formation. Neuron *100*, 276-293. https://doi.org/10.1016/j.neuron.2018.09.040.

Sugawara, T., Hisatsune, C., Miyamoto, H., Ogawa, N., and Mikoshiba, K. (2017). Regulation of spinogenesis in mature Purkinje cells via mGluR/PKC-mediated phosphorylation of CaMKIIβ. Proc Natl Acad Sci U S A *114*, E5256-e5265. 10.1073/pnas.1617270114.

Takeuchi, Y., Yamamoto, H., Fukunaga, K., Miyakawa, T., and Miyamoto, E. (2000). Identification of the isoforms of Ca(2+)/Calmodulin-dependent protein kinase II in rat astrocytes and their subcellular localization. J Neurochem *74*, 2557-2567. 10.1046/j.1471-4159.2000.0742557.x.

Terai, Y., Morikawa, N., Kawakami, K., and Okada, N. (2003). The complexity of alternative splicing of hagoromo mRNAs is increased in an explosively speciated lineage in East African cichlids. Proceedings of the National Academy of Sciences *100*, 12798-12803. doi:10.1073/pnas.2132833100.

Tian, B., and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. Nature Reviews Molecular Cell Biology *18*, 18-30. 10.1038/nrm.2016.116.

Tombes, R.M., Faison, M.O., and Turbeville, J.M. (2003). Organization and evolution of multifunctional Ca(2+)/CaM-dependent protein kinase genes. Gene *322*, 17-31. 10.1016/j.gene.2003.08.023.

Tombes, R.M., and Krystal, G.W. (1997). Identification of novel human tumor cell-specific CaMK-II variants. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1355*, 281-292. https://doi.org/10.1016/S0167-4889(96)00141-3.

Torres-Ocampo, A.P., Özden, C., Hommer, A., Gardella, A., Lapinskas, E., Samkutty, A., Esposito, E., Garman, S.C., and Stratton, M.M. (2020). Characterization of CaMKIIα holoenzyme stability. Protein Science *29*, 1524-1534. https://doi.org/10.1002/pro.3869.

Tullis, J.E., Rumian, N.L., Brown, C.N., and Bayer, K.U. (2020). The CaMKII K42M and K42R mutations are equivalent in suppressing kinase activity and targeting. PLOS ONE *15*, e0236478. 10.1371/journal.pone.0236478.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419. 10.1126/science.1260419.

Ule, J., and Blencowe, B.J. (2019). Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. Mol Cell *76*, 329-345. 10.1016/j.molcel.2019.09.017.

Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). An RNA map predicting Nova-dependent splicing regulation. Nature *444*, 580-586. 10.1038/nature05304.

Urquidi, V., and Ashcroft, S.J. (1995). A novel pancreatic beta-cell isoform of calcium/calmodulin-dependent protein kinase II (beta 3 isoform) contains a proline-rich tandem repeat in the association domain. FEBS Lett *358*, 23-26. 10.1016/0014-5793(94)01381-a.

van Woerden, G.M., Harris, K.D., Hojjati, M.R., Gustin, R.M., Qiu, S., de Avila Freire, R., Jiang, Y.H., Elgersma, Y., and Weeber, E.J. (2007). Rescue of neurological deficits in a mouse model for Angelman syndrome by reduction of alphaCaMKII inhibitory phosphorylation. Nat Neurosci *10*, 280-282. 10.1038/nn1845.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701-718. 10.1016/j.cell.2009.02.009.

Wahle, E., and Keller, W. (1992). THE BIOCHEMISTRY OF 3′-END CLEAVAGE AND POLYADENYLATION OF MESSENGER RNA PRECURSORS. Annual Review of Biochemistry *61*, 419-438. 10.1146/annurev.bi.61.070192.002223.

Wan, Y., Anastasakis, D.G., Rodriguez, J., Palangat, M., Gudla, P., Zaki, G., Tandon, M., Pegoraro, G., Chow, C.C., Hafner, M., and Larson, D.R. (2021). Dynamic imaging

of nascent RNA reveals general principles of transcription dynamics and stochastic splice site selection. Cell *184*, 2878-2895.e2820. 10.1016/j.cell.2021.04.012.

Wang, H., Shimizu, E., Tang, Y.P., Cho, M., Kyin, M., Zuo, W., Robinson, D.A., Alaimo, P.J., Zhang, C., Morimoto, H., et al. (2003). Inducible protein knockout reveals temporal requirement of CaMKII reactivation for memory consolidation in the brain. Proc Natl Acad Sci U S A *100*, 4287-4292. 10.1073/pnas.0636870100.

Wang, J., Lu, Z.X., Tokheim, C.J., Miller, S.E., and Xing, Y. (2015). Species-specific exon loss in human transcriptomes. Mol Biol Evol *32*, 481-494. 10.1093/molbev/msu317.

Wang, M., and Marín, A. (2006). Characterization and prediction of alternative splice sites. Gene *366*, 219-227. 10.1016/j.gene.2005.07.015.

Wang, Z., Kayikci, M., Briese, M., Zarnack, K., Luscombe, N.M., Rot, G., Zupan, B., Curk, T., and Ule, J. (2010). iCLIP Predicts the Dual Splicing Effects of TIA-RNA Interactions. PLOS Biology *8*, e1000530. 10.1371/journal.pbio.1000530.

Wetsel, W.C. (2011). Sensing hot and cold with TRP channels. International Journal of Hyperthermia *27*, 388-398. 10.3109/02656736.2011.554337.

Will, C.L., and Lührmann, R. (2011). Spliceosome structure and function. Cold Spring Harb Perspect Biol *3*. 10.1101/cshperspect.a003707.

Witten, J.T., and Ule, J. (2011). Understanding splicing regulation through RNA splicing maps. Trends in Genetics *27*, 89-97. https://doi.org/10.1016/j.tig.2010.12.001.

Woolfrey, K.M., O'Leary, H., Goodell, D.J., Robertson, H.R., Horne, E.A., Coultrap, S.J., Dell'Acqua, M.L., and Bayer, K.U. (2018). CaMKII regulates the depalmitoylation and synaptic removal of the scaffold protein AKAP79/150 to mediate structural long-term depression. Journal of Biological Chemistry *293*, 1551-1567. https://doi.org/10.1074/jbc.M117.813808.

Xi, F., Xu, R.J., Xu, J.H., Ma, J.J., Wang, W.H., Wang, F., Ma, Y.X., Qi, S.B., Zhang, H.C., Zhang, H.N., et al. (2019). Calcium/calmodulin-dependent protein kinase II regulates mammalian axon growth by affecting F-actin length in growth cone. J Cell Physiol *234*, 23053-23065. 10.1002/jcp.28867.

Yamagata, Y., Kobayashi, S., Umeda, T., Inoue, A., Sakagami, H., Fukaya, M., Watanabe, M., Hatanaka, N., Totsuka, M., Yagi, T., et al. (2009). Kinase-Dead Knock-In Mouse Reveals an Essential Role of Kinase Activity of $Ca^{2+}$/Calmodulin-Dependent Protein Kinase IIα in Dendritic Spine Enlargement, Long-Term Potentiation, and Learning. The Journal of Neuroscience *29*, 7607-7618. 10.1523/jneurosci.0707-09.2009.

Yamasaki, N., Maekawa, M., Kobayashi, K., Kajii, Y., Maeda, J., Soma, M., Takao, K., Tanda, K., Ohira, K., Toyama, K., et al. (2008). Alpha-CaMKII deficiency causes immature dentate gyrus, a novel candidate endophenotype of psychiatric disorders. Molecular Brain *1*, 6. 10.1186/1756-6606-1-6.

Yamauchi, T., and Fujisawa, H. (1980). Evidence for three distinct forms of calmodulin-dependent protein kinases from rat brain. FEBS Letters *116*, 141-144. https://doi.org/10.1016/0014-5793(80)80628-4.

Yanaizu, M., Washizu, C., Nukina, N., Satoh, J.I., and Kino, Y. (2020). CELF2 regulates the species-specific alternative splicing of TREM2. Sci Rep *10*, 17995. 10.1038/s41598-020-75057-x.

Yang, X., Coulombe-Huntington, J., Kang, S., Sheynkman, Gloria M., Hao, T., Richardson, A., Sun, S., Yang, F., Shen, Yun A., Murray, Ryan R., et al. (2016). Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing. Cell *164*, 805-817. https://doi.org/10.1016/j.cell.2016.01.029.

Yu, H., Pan, B., Weyer, A., Wu, H.-E., Meng, J., Fischer, G., Vilceanu, D., Light, A.R., Stucky, C., Rice, F.L., et al. (2015). CaMKII Controls Whether Touch Is Painful. The Journal of Neuroscience *35*, 14086-14102. 10.1523/jneurosci.1969-15.2015.

Zalcman, G., Federman, N., and Romano, A. (2018). CaMKII Isoforms in Learning and Memory: Localization and Function. Frontiers in Molecular Neuroscience *11*. 10.3389/fnmol.2018.00445.

Zhang, X.H., and Chasin, L.A. (2006). Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons. Proc Natl Acad Sci U S A *103*, 13427-13432. 10.1073/pnas.0603042103.

Zhou, J., and Adams, J.A. (1997). Participation of ADP Dissociation in the Rate-Determining Step in cAMP-Dependent Protein Kinase. Biochemistry *36*, 15733-15738. 10.1021/bi971438n.

Zhu, J.J., Qin, Y., Zhao, M., Van Aelst, L., and Malinow, R. (2002). Ras and Rap Control AMPA Receptor Trafficking during Synaptic Plasticity. Cell *110*, 443-455. https://doi.org/10.1016/S0092-8674(02)00897-8.

Zuallaert, J., Godin, F., Kim, M., Soete, A., Saeys, Y., and De Neve, W. (2018). SpliceRover: interpretable convolutional neural networks for improved splice site prediction. Bioinformatics *34*, 4180-4188. 10.1093/bioinformatics/bty497.

Zybura, A.S., Baucum, A.J., Rush, A.M., Cummins, T.R., and Hudmon, A. (2020). CaMKII enhances voltage-gated sodium channel Nav1.6 activity and neuronal excitability. Journal of Biological Chemistry *295*, 11845-11865. https://doi.org/10.1074/jbc.RA120.014062.

# Contributions

The following experiments were done in part or entirely by other researchers. Nicole Dimos performed part of minigene exchange experiments (Figure 2E,F, Figure S2B), including cloning, minigene splicing assay and evaluation. Marco Preußner assisted with the evaluation of the RNA-Seq data, both of the CaMKIIβ mouse model and the publicly available mammalian datasets. Ji Yanlong (AG Urlaub, MPI for Multidisciplinary Sciences) and Benno Kuropka (FU Berlin, BioSupraMol) measured each a part of the analog-sensitive kinase samples and assisted with the data analysis. Alexander Stumpf (AG Schmitz, Charité, NeuroCure) performed the electrophysical characterization of the mouse model. Some of the results regarding the endogenous splicing of *Camk2β*, as well as the initial minigene experiments (Figure 5A-D) were already partially reported in my master thesis titled “Species-Specific Alternative Splicing” (2017). Tarek Hilal (FU Berlin, BioSupraMol) assisted with the cryo-electron microscopy analysis, and performed the sample preparation, data collection and generation of 3D reconstructions.

# Supplement

Supplementary Table S1. cryoEM data collection and refinement statistics.

| Complex | 12-mer sym. | 12-mer asym. | open 12-mer | 14-mer | 16-mer |
|---|---|---|---|---|---|
| **Data collection** | | | | | |
| Microscope | FEI Titan Krios G3i | | | | |
| Voltage [keV] | 300 | | | | |
| Camera | Falcon 3EC | | | | |
| Magnification | 96,000 | | | | |
| Pixel size [Å/px] | 0.832 | | | | |
| Total electron exposure [$e^-$ /Å$^2$ ] | 40 | | | | |
| Exposure rate [$e^-$ /px/s] | 1 | | | | |
| No. of frames collected during exposure | 33 | | | | |
| Defocus range [μm] | 0.6 - 2 | | | | |
| Automation software | EPU 2.8.1 | | | | |
| Micrographs collected | 1185 | | | | |
| Micrographs used | 1184 | | | | |
| **Data processing** | | | | | |
| Total extracted particles | 668,392 | | | | |
| Refined particles | 301,046 | | | | |
| Final particles | 139,526 | 27,310 | 30,621 | 38,726 | 3,430 |
| Point-group or symmetry implied | D6 | C1 | C1 | C1 | D8 |
| Global resolution [Å] | 2.55 | 3.20 | 3.26 | 3.17 | 3.50 |
| FSC0.143 unmasked/masked [Å] | 3.3/2.6 | 7.3/3.2 | 7.1/3.3 | 6.8/3.2 | 7.3/3.5 |
| Map sharpening B factor [Å$^2$ ] / (B factor range) | -93 | -65 | -77 | -74 | -81 |
| Map sharpening method | global B-factor | | | | |
| **Model refinement** | | | | | |
| CC mask | 0.89 | 0.84 | 0.68 | NA | NA |
| CC volume | 0.89 | 0.84 | 0.68 | | |
| Model | | | | | |
| Non-H atoms | 12726 | 12743 | 12664 | | |
| Protein residues | 1584 | 1587 | 1577 | | |
| RMSD | | | | | |
| Bond lengths [Å] | 0.006 | 0.006 | 0.003 | | |
| Bond angles [°] | 0.637 | 0.769 | 0.739 | | |
| Ramachandran plot | | | | | |
| Favored [%] | 98.33 | 97.89 | 93.3 | | |
| Allowed [%] | 1.67 | 2.11 | 6.57 | | |
| Outliers [%] | 0 | 0 | 0.13 | | |
| Model quality | | | | | |
| Clash score | 3.33 | 11.69 | 17.39 | | |
| Rotamer outliers [%] | 0.52 | 0.89 | 0.6 | | |
| Overall score | 1.12 | 1.61 | 2.18 | | |

Supplementary Table S1 cryoEM data collection and refinement statistics. RMSD: root mean square deviation. Refinement statistics, including model quality (MolProbity) were assessed in Phenix (Liebschner *et al.*, 2019).

Supplementary Table S2. Publicly available RNA-Seq datasets used in this study.

| SRA | Species | Tissue | Read length | paired/single | Reference Genome |
|---|---|---|---|---|---|
| SRR8750487 | *Human* | Cerebellar White Matter | 150 bp | paired-end | GRCh38 |
| SRR8750488 | *Human* | Cerebellar White Matter | 150 bp | paired-end | GRCh38 |
| SRR8750489 | *Human* | Cerebellar White Matter | 150 bp | paired-end | GRCh38 |
| SRR8750490 | *Human* | Cerebellar White Matter | 150 bp | paired-end | GRCh38 |
| SRR8750491 | *Human* | Cerebellar Grey Matter | 150 bp | paired-end | GRCh38 |
| SRR8750492 | *Human* | Cerebellar Grey Matter | 150 bp | paired-end | GRCh38 |
| SRR8750493 | *Human* | Cerebellar Grey Matter | 150 bp | paired-end | GRCh38 |
| SRR8750647 | *Pan troglodytes* | Cerebellar White Matter | 150 bp | paired-end | panTro6 |
| SRR8750679 | *Pan troglodytes* | Cerebellar White Matter | 150 bp | paired-end | panTro6 |
| SRR8750711 | *Pan troglodytes* | Cerebellar White Matter | 150 bp | paired-end | panTro6 |
| SRR8750648 | *Pan troglodytes* | Cerebellar Grey Matter | 150 bp | paired-end | panTro6 |
| SRR8750680 | *Pan troglodytes* | Cerebellar Grey Matter | 150 bp | paired-end | panTro6 |
| SRR8750712 | *Pan troglodytes* | Cerebellar Grey Matter | 150 bp | paired-end | panTro6 |
| SRR8750448 | *Pan paniscus* | Cerebellar White Matter | 150 bp | paired-end | panPan1.1 |
| SRR8750449 | *Pan paniscus* | Cerebellar White Matter | 150 bp | paired-end | panPan1.1 |
| SRR8750450 | *Pan paniscus* | Cerebellar White Matter | 150 bp | paired-end | panPan1.1 |
| SRR8750451 | *Pan paniscus* | Cerebellar Grey Matter | 150 bp | paired-end | panPan1.1 |
| SRR8750452 | *Pan paniscus* | Cerebellar Grey Matter | 150 bp | paired-end | panPan1.1 |
| SRR8750595 | *Pan paniscus* | Cerebellar Grey Matter | 150 bp | paired-end | panPan1.1 |
| SRR5804509 | *Gorilla gorilla* | Cerebellum | 101 bp | paired-end | gorGor6 |

| | | | | | |
|---|---|---|---|---|---|
| SRR5804501 | *Gorilla gorilla* | Cerebellum | 101 bp | paired-end | gorGor6 |
| SRR306801 | *Gorilla gorilla* | Brain | 101 bp | paired-end | gorGor6 |
| SRR10393301 | *Pongo pygmaeus abelii* | Testis | 150 bp | paired-end | ponAbe3 |
| SRR10393303 | *Pongo pygmaeus abelii* | Testis | 150 bp | paired-end | ponAbe3 |
| SRR10393302 | *Pongo pygmaeus abelii* | Testis | 150 bp | paired-end | ponAbe3 |
| SRR10393304 | *Pongo pygmaeus abelii* | Testis | 150 bp | paired-end | ponAbe3 |
| DRR128395 | *Pongo pygmaeus* | Skin | 125 bp | paired-end | ponAbe3 |
| DRR128394 | *Pongo pygmaeus* | Skin | 100 bp | paired-end | ponAbe3 |
| DRR128393 | *Pongo pygmaeus* | Skin | 100 bp | paired-end | ponAbe3 |
| SRR306792 | *Pongo pygmaeus* | Brain | 150 bp | paired-end | ponAbe3 |
| SRR5804517 | *Hylobates lar* | Cerebellum | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |
| SRR5804510 | *Hylobates lar* | Dorsolateral Prefrontal Cortex | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |
| SRR5804511 | *Hylobates lar* | Ventrolateral Prefrontal Cortex | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |
| SRR5804512 | *Hylobates lar* | Premotor Cortex | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |
| SRR5804513 | *Hylobates lar* | Primary Visual Cortex | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |
| SRR5804514 | *Hylobates lar* | Anterior Cingulate Cortex | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |
| SRR5804515 | *Hylobates lar* | Striatum | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |

| | | | | | |
|---|---|---|---|---|---|
| SRR5804516 | *Hylobates lar* | Hippocampus | 100 bp | paired-end | nomLeu3 (*Nomascus leucogenys*) |
| SRR8750552 | *Macaca mulatta* | Cerebellar White Matter | 150 bp | paired-end | rheMac10 |
| SRR8750553 | *Macaca mulatta* | Cerebellar White Matter | 150 bp | paired-end | rheMac10 |
| SRR8750554 | *Macaca mulatta* | Cerebellar White Matter | 150 bp | paired-end | rheMac10 |
| SRR8750549 | *Macaca mulatta* | Cerebellar Grey Matter | 150 bp | paired-end | rheMac10 |
| SRR8750550 | *Macaca mulatta* | Cerebellar Grey Matter | 150 bp | paired-end | rheMac10 |
| SRR8750551 | *Macaca mulatta* | Cerebellar Grey Matter | 150 bp | paired-end | rheMac10 |
| SRR11939284 | *Sus scrofa* | Cerebellum | 150 bp | paired-end | SusScr11 |
| SRR11939285 | *Sus scrofa* | Cerebellum | 150 bp | paired-end | SusScr11 |
| SRR11939286 | *Sus scrofa* | Cerebellum | 150 bp | paired-end | SusScr11 |

Supplementary Table S2. Publicly available RNA-Seq datasets used in this study. SRA: Identifier for the Sequence Read Archive (SRA), Reference Genome: genome that was used to map the RNA-Seq reads.

Supplementary Table S3. Myelin sheath associated targets are specific to the Δ16,17 isoform.

| Unique phosphosite | Log2 (Enrichment) | | | | | |
|---|---|---|---|---|---|---|
| | 13-inclusion | 16-inclusion | FL/Ctr | 13/Ctr | 16,17/Ctr | 13,16/Ctr |
| **Pkm-T25** | unique | absent | absent | absent | unique | absent |
| **Tubb4b-S126** | unique | absent | absent | absent | unique | absent |
| **Eef1a2-S396** | unique | absent | absent | absent | unique | absent |
| **Plec-T3795** | -0,45 | 0,49 | unique | unique | unique | unique |
| **Slc25a4-T254** | unique | absent | absent | absent | unique | absent |
| **Gdi2-S302** | unique | absent | absent | absent | unique | absent |
| **Plcb1-S564** | 0,26 | absent | absent | absent | unique | unique |
| **Pkm-S362** | 3,21 | -3,13 | 1,98 | 0,59 | 4,58 | absent |
| **Mdh2-S276** | 0,95 | -0,99 | 2,41 | 1,88 | 3,16 | absent |
| **Eef1a1-S396** | 1,14 | -0,16 | 1,30 | 1,17 | 2,13 | -0,21 |
| **Eef1a1-S414** | 1,32 | -0,53 | 0,61 | 0,40 | 1,78 | -0,58 |
| **Aldoa-T235** | 1,17 | -0,91 | 0,75 | -0,21 | 1,76 | 0,47 |
| **Actg1-S14** | 1,06 | -0,65 | 0,11 | 0,22 | 1,49 | -0,51 |
| **Eef1a2-S414** | 0,93 | -1,56 | -0,48 | -0,40 | 1,12 | absent |
| **Dpysl2-Y431** | 1,04 | -0,79 | -0,45 | -0,41 | 1,00 | -0,80 |
| **Pkm-T328** | 2,02 | -0,16 | 0,04 | absent | 0,96 | -1,46 |
| **Msn-Y291** | 0,77 | 0,19 | 0,08 | 0,39 | 0,80 | -1,61 |
| **Ppia-T68** | 0,90 | -0,22 | -0,46 | -0,09 | 0,69 | -1,65 |
| **Hsp90aa1-Y529** | -0,10 | 0,17 | 0,32 | 0,35 | 0,17 | absent |
| **Actb-Y218** | 0,23 | -0,70 | -0,62 | -0,75 | 0,16 | -0,13 |
| **Sptan1-S2423** | 1,08 | -1,13 | -1,79 | -1,50 | 0,14 | -1,72 |
| **Cct2-S413** | 1,11 | -0,28 | -0,58 | -1,02 | 0,13 | -1,64 |
| **Stip1-Y476** | 0,28 | 0,75 | 0,55 | 0,36 | -0,09 | -0,53 |
| **Eno1-S63** | -0,72 | -1,31 | -1,26 | -0,81 | -0,18 | 0,65 |
| **Cltc-Y1206** | 0,65 | -0,65 | absent | -1,14 | -0,49 | absent |
| **Actb-S14** | 0,18 | 0,01 | -0,52 | -0,68 | -0,51 | -0,70 |
| **Atp1a1-S668** | -0,09 | 0,23 | -0,27 | -0,23 | -0,55 | -0,41 |
| **Plec-S3772** | 0,95 | 1,01 | 0,98 | 0,32 | absent | -0,32 |
| **Cct3-Y274** | -0,51 | -0,51 | -1,05 | absent | absent | -0,53 |
| **Septin8-S313** | -0,17 | unique | -2,06 | -1,89 | absent | absent |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Tuba1b-T292** | 1,95 | unique | 1,60 | -0,35 | absent | absent |
| **Actb-S33** | absent | unique | absent | unique | absent | absent |

Supplementary Table S3. Identified phosphorylation targets that associate with the GO term GO:0043209 (myelin sheath). The table contains all identified phosphorylation targets, including targets also detected in control samples (UT, K43R). Exon 13-incl./ Exon 16-incl.: Log2 ratio between the average intensity values for exon 13 or 16 inclusion variants vs exon 13 or 16 skipping variants. Isoform/Ctr: Log2 ratio between the average intensity values for this isoform vs controls. Unique: this phosphorylation target was not detected in the corresponding control samples (13/16-exclusion variants for 13-inclusion/16-inclusion; UT and K43R for isoform/Ctr. Absent: this phosphorylation site was not detected in the corresponding sample.

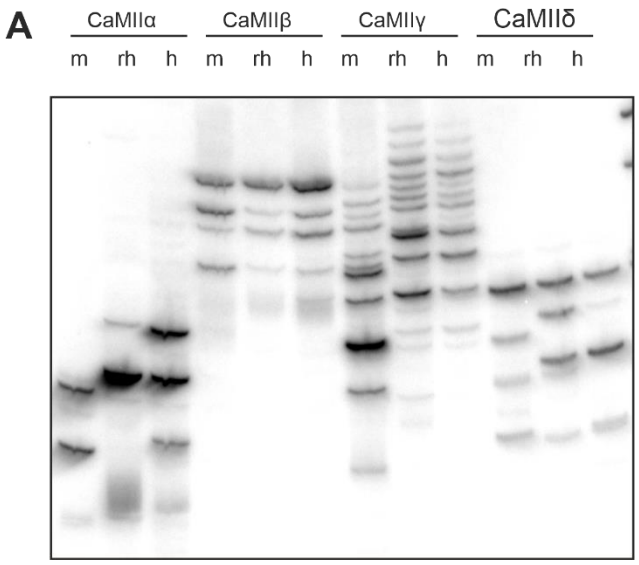



Figure S1: Species-specific AS of CaMKII, additional rhesus samples for all 4 genes.

(A) Endogenous CaMKII splice isoforms were identified by radioactive isoform-specific RT-PCR with primate (*Macaca mulatta*) total brain RNA and mouse (*Mus musculus*) and human cerebellum RNA. Isoforms were separated on a denaturing polyacrylamide gel and detected via autoradiography. m: mouse, rh: rhesus, h:human.

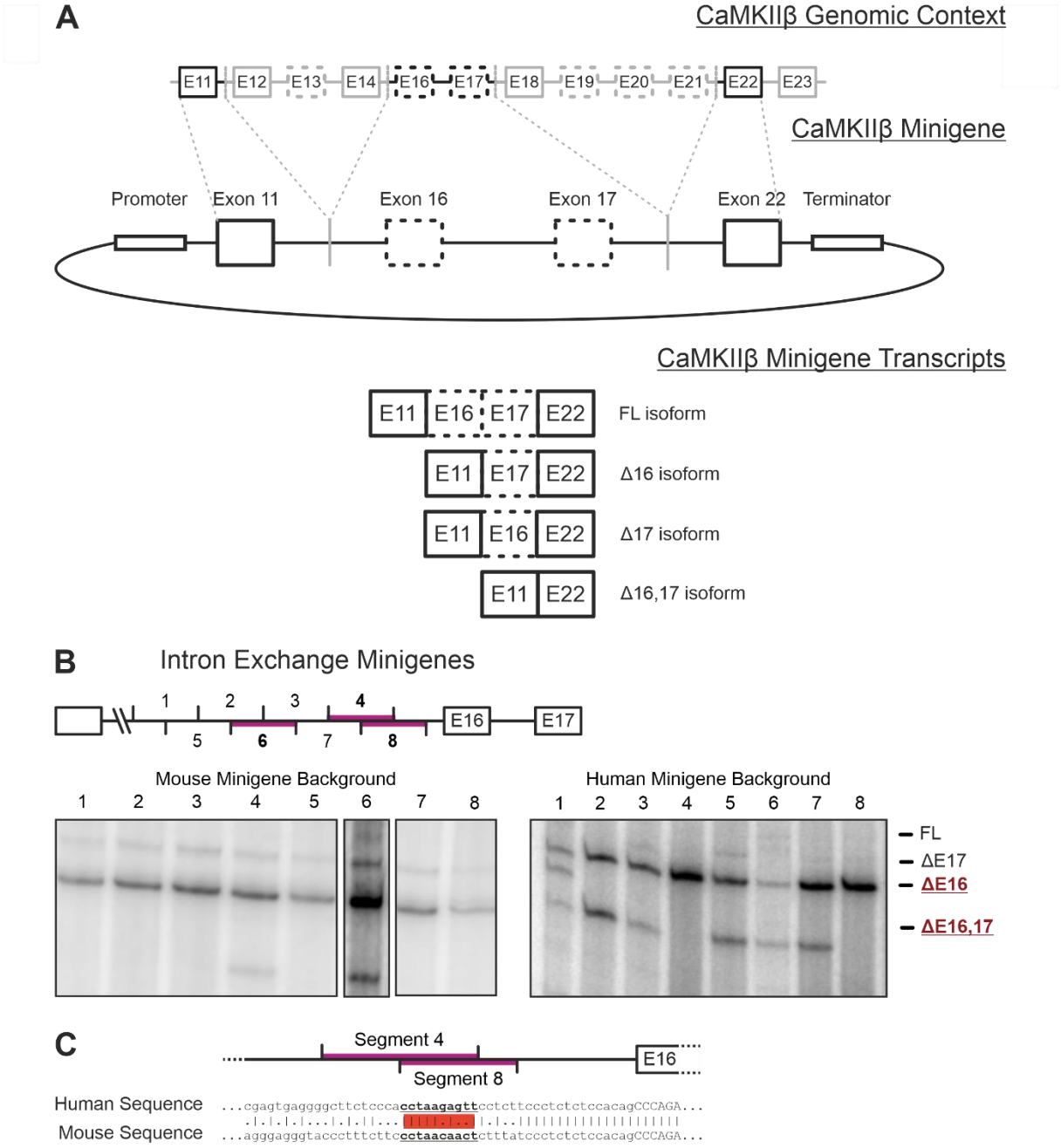



Figure S2: Minigene design and minigene splicing assays.

(A) Schematic representation of the design for the CaMKIIβ minigenes. The top represents the genomic background, boxed represent exons, lines represent introns. Dashed boxes represent alternatively spliced exons. The middle represents the minigene construct, containing two alternative exons flanked by two constitutive exons. Efficient transcription is ensured by promoter and terminator sequences. Dashed lines indicate which exonic and intronic regions from the genomic background were used in the minigene construct. The bottom displays all four possible alternatively spliced transcripts. (B) Top: Schematic representation of the intron containing the identified functionally relevant *cis*-acting element. Numbers indicate 20 bp segments that were exchanged between the human and mouse construct. Purple lines highlight segments of functional relevance. Bottom: Human and mouse exchange minigenes were transfected into HEK cells and resulting splice isoforms identified by radioactive RT-PCR. (C) Alignment of the human and mouse sequence of the functionally relevant segments 4 and 8. The overlap between both segments is highlighted in red. Experiment B was performed by Nicole Dimos.

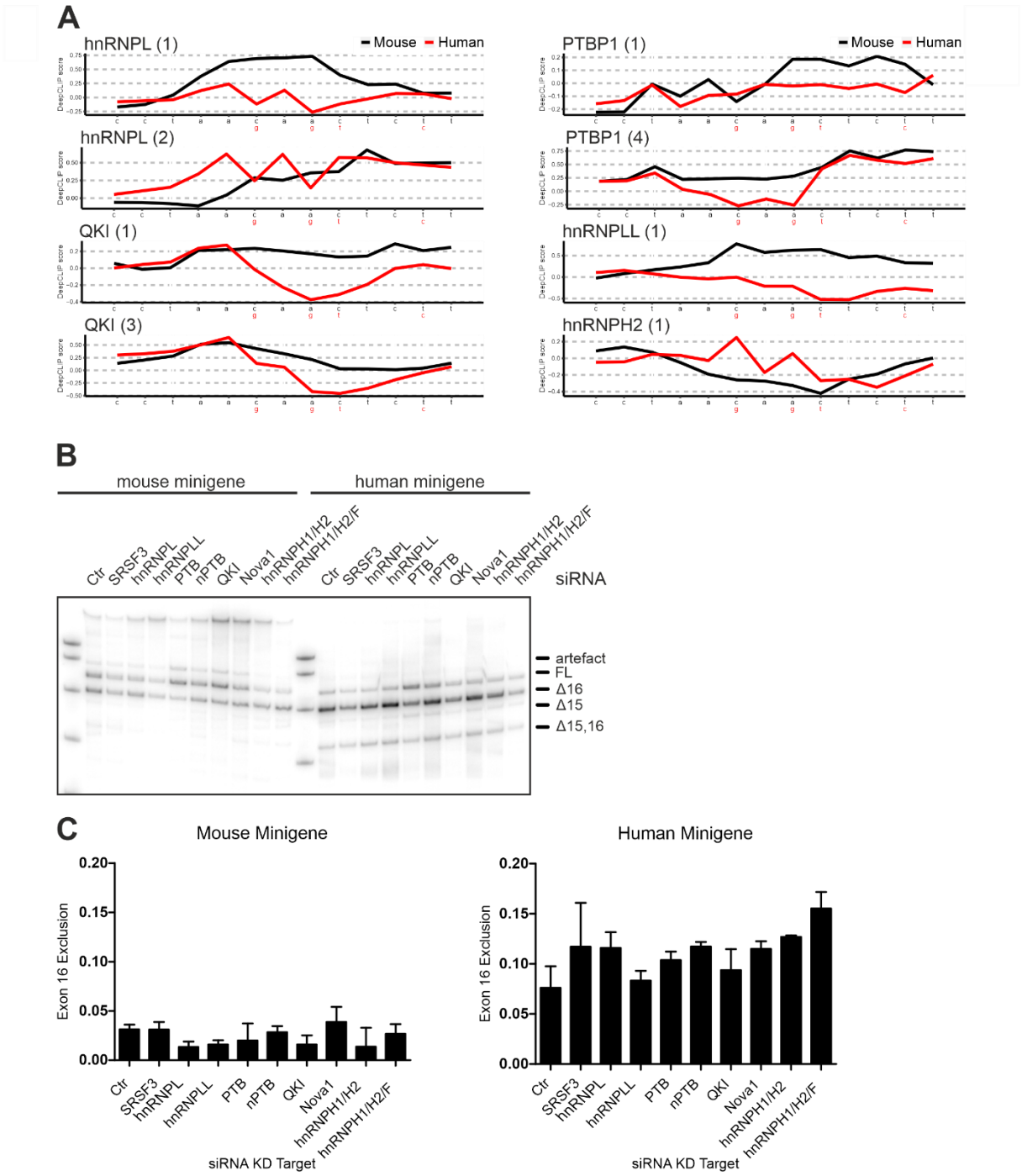



Figure S3: siRNA knockdown fails to identify a corresponding *trans*-acting factor.
(A) DeepClip (Grønning *et al.*, 2020) predictions for potential *trans*-acting factors binding to the identified segments of interest in the intron upstream of exon 16 of the CaMKIIβ gene. The human and mouse sequence of the identified functionally relevant segment were compared and results filtered as to identify RNA binding proteins that would preferentially bind to one of the two species. Numbers indicate the pre-trained model that was used for the prediction: 1: (Ray et al., 2013), 2: (Rossbach et al., 2009), 3: (Galarneau and Richard, 2005), 4: (García-Blanco et al., 1989). (B) siRNA KD of potential *trans*-acting factors regulating the species-specific AS of CaMKIIβ. Indicated *trans*-acting factors were downregulated by siRNA-mediated KD in N2a cells. Human and mouse minigenes of CaMKIIβ were transfected and resulting splice isoforms identified by radioactive RT-PCR. Isoforms are indicated on the right and named according to the exons skipped. (C) Quantification of B, error bars indicate standard deviation (n=3).

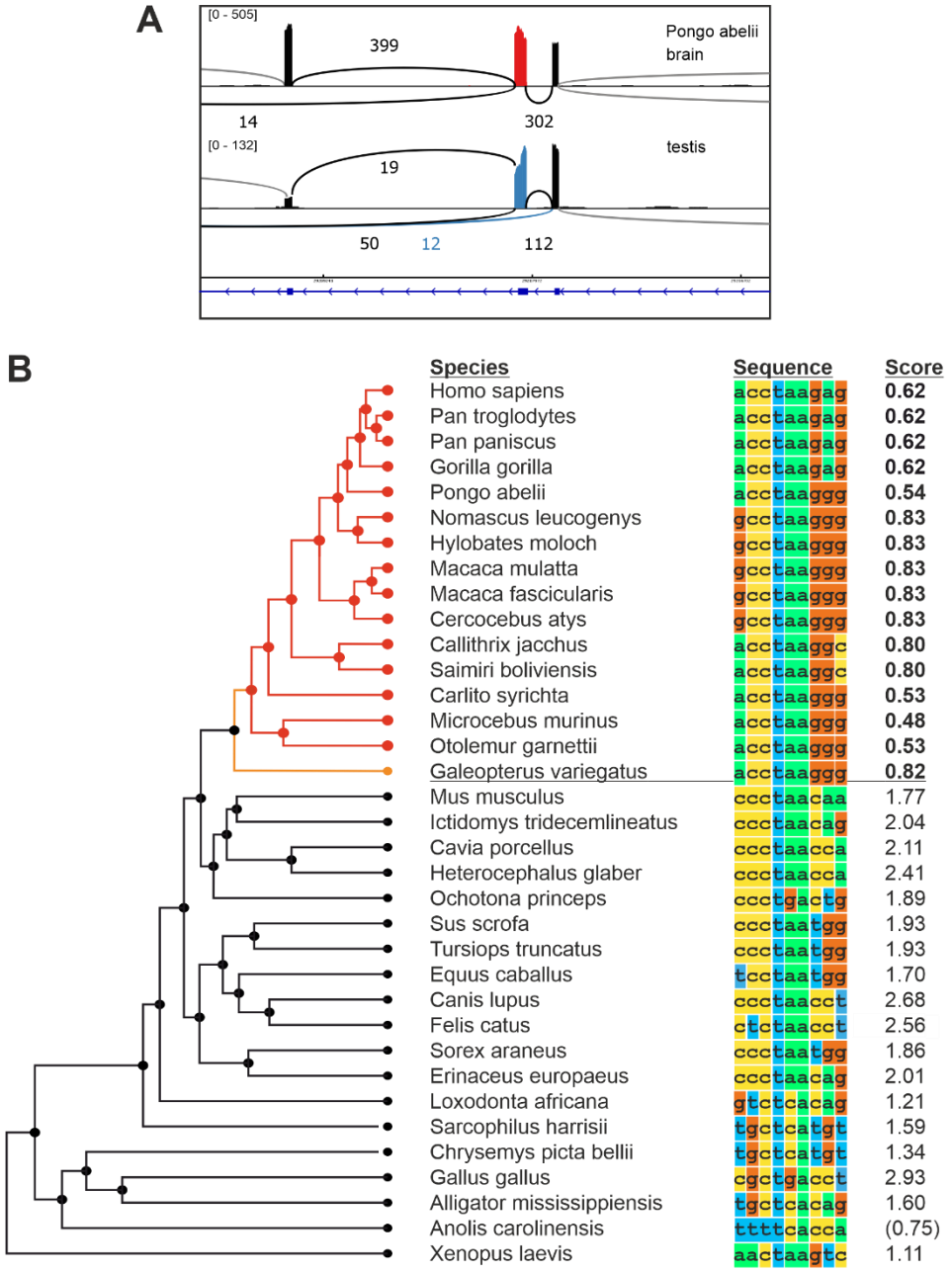



Figure S4: RNA-Seq Analysis reveals evolution of CaMKIIβ exon 16 alternative splicing.

(A) Sashimi plot showing the alternative splicing of CaMKIIβ exon 16 in RNA-Seq data from orangutan (*Pongo abelii*). RNA-Seq data from total brain tissue and testis was used. Red/blue colour indicates exon 16 and exon 16 exclusion reads. Numbers indicate number of reads per splice junction. Shown in blue is the intron/ exon-structure of the displayed region. (B) Alignment of the highest scoring BP sequence in various species. Only BPs within 100 nt of the 3' splice site were considered. Left is a phylogenetic tree depicting evolutionary relationships. Red: Primate order, orange: Dermoptera order. Score revers to the BP score (scaled vector model) calculated via SVM-BPfinder (Corvelo *et al.*, 2010). Brackets indicate species for which no acceptable BP was found within 100 nt of the 3' splice site.

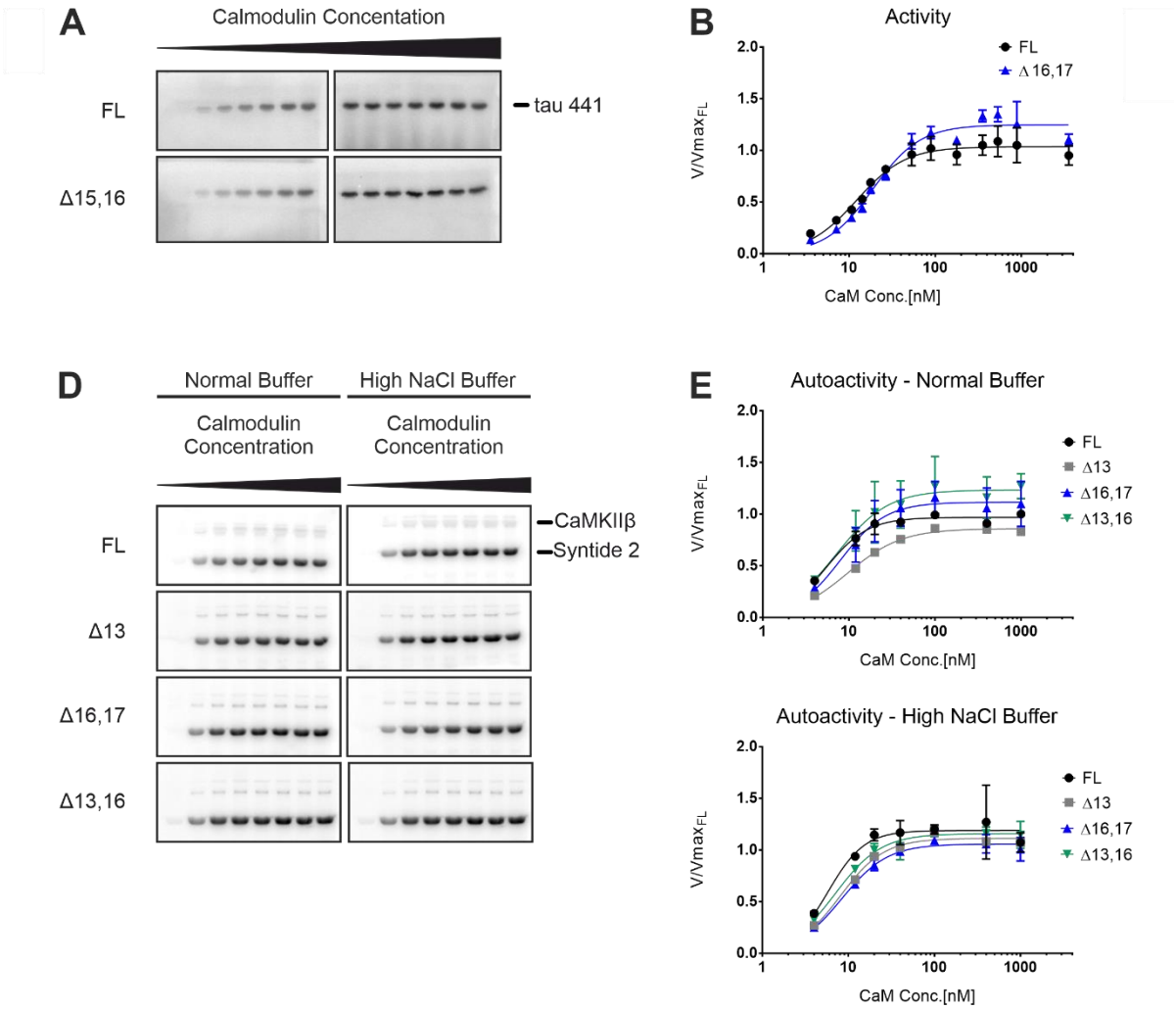



Figure S5: The kinetic differences are substrate independent and depend on the reaction conditions.

(A) *In vitro* kinase assay with different CaMKIIβ isoforms. CaMKII activity against a protein substrate (human full-length tau-441, with N-terminal polyhistidine- and C-terminal StrepII-tag) was measured as a function of calmodulin concentration. Direct phosphorylation of the substrate by CaMKIIβ was measured via $^{32}P$ incorporation. Samples were separated on an SDS-PAGE and detected using autoradiography. (B) Quantification of D, normalized to the maximum activity of the FL isoform (n = 3). (C) Table with $V_{max}$, $EC_{50}$ values and Hill coefficients (h) of the different CaMKIIβ isoforms for two different substrate proteins. Data was fitted to a Hill equation (allosteric sigmoidal nonlinear fit) in GraphPad Prism 6. (D) *In vitro* kinase assay with different CaMKIIβ isoforms to test the autoactivity generated at increasing calmodulin concentrations. CaMKII was first activated with calmodulin in the presence of ATP. EGTA was added to chelate calcium and quench the binding of calmodulin. Addition of a protein substrate (Syntide 2, fused to GST) enabled detected of generated autoactivity via $^{32}P$ incorporation. Samples were separated on an SDS-PAGE and detected using autoradiography. Left: no-salt buffer, right: high-salt buffer. (E) Quantification of D, normalized to the maximum activity of the FL isoform (n=3)

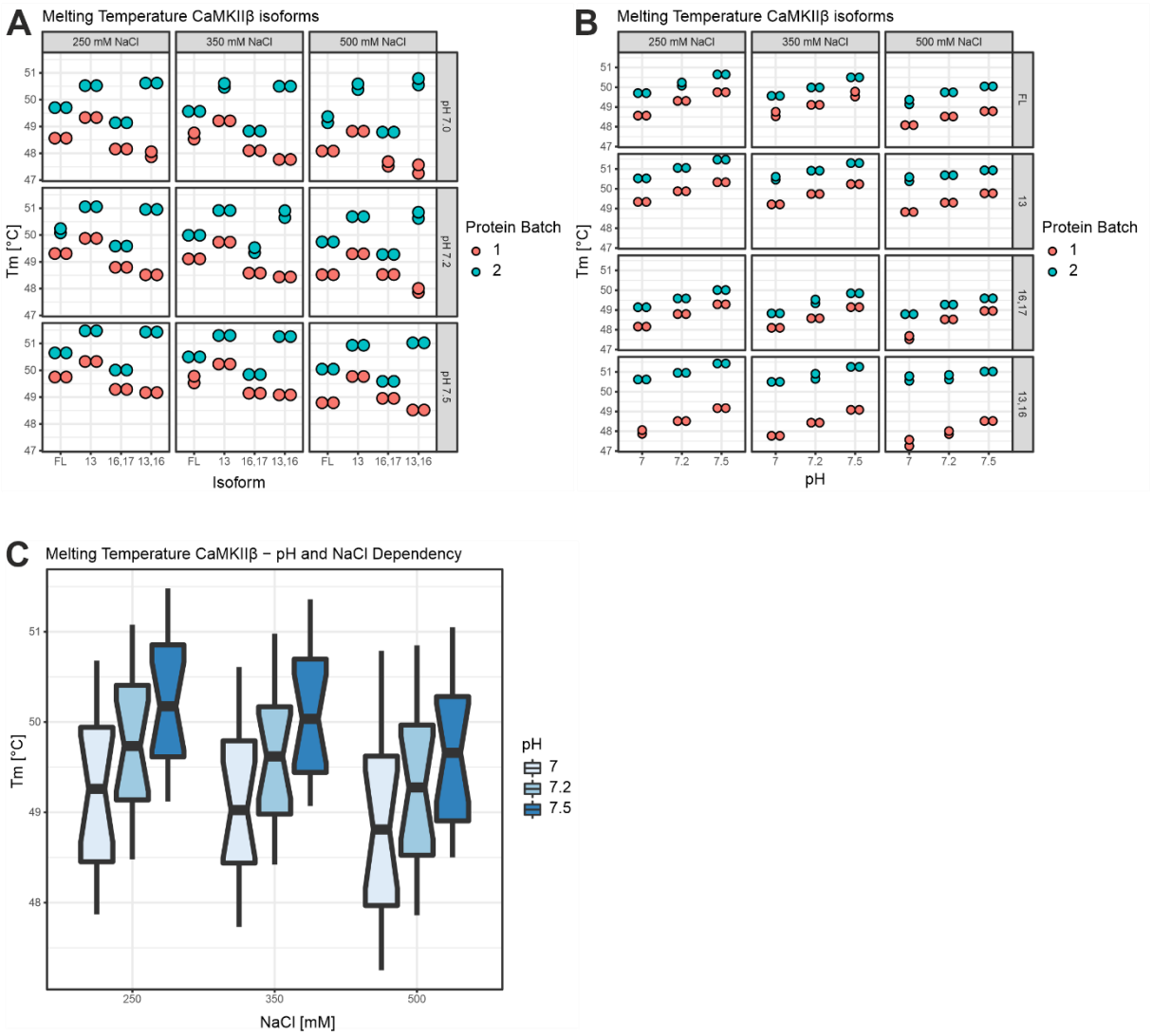



Figure S6: CaMKIIβ isoforms show differences in protein stability.

(A,B) Thermofluor assay with purified CaMKIIβ isoforms performed under varying pH and NaCl concentrations. A: samples are classified according to NaCl concentration and pH, x-axis represents different isoforms. B: samples are classified according to NaCl concentration and isoform, x-axis represents different pH conditions. Colour code indicates protein batch used. (C) Pooled samples (all isoforms, both protein batches), melting temperature (Tm) plotted against NaCl concentration and pH.

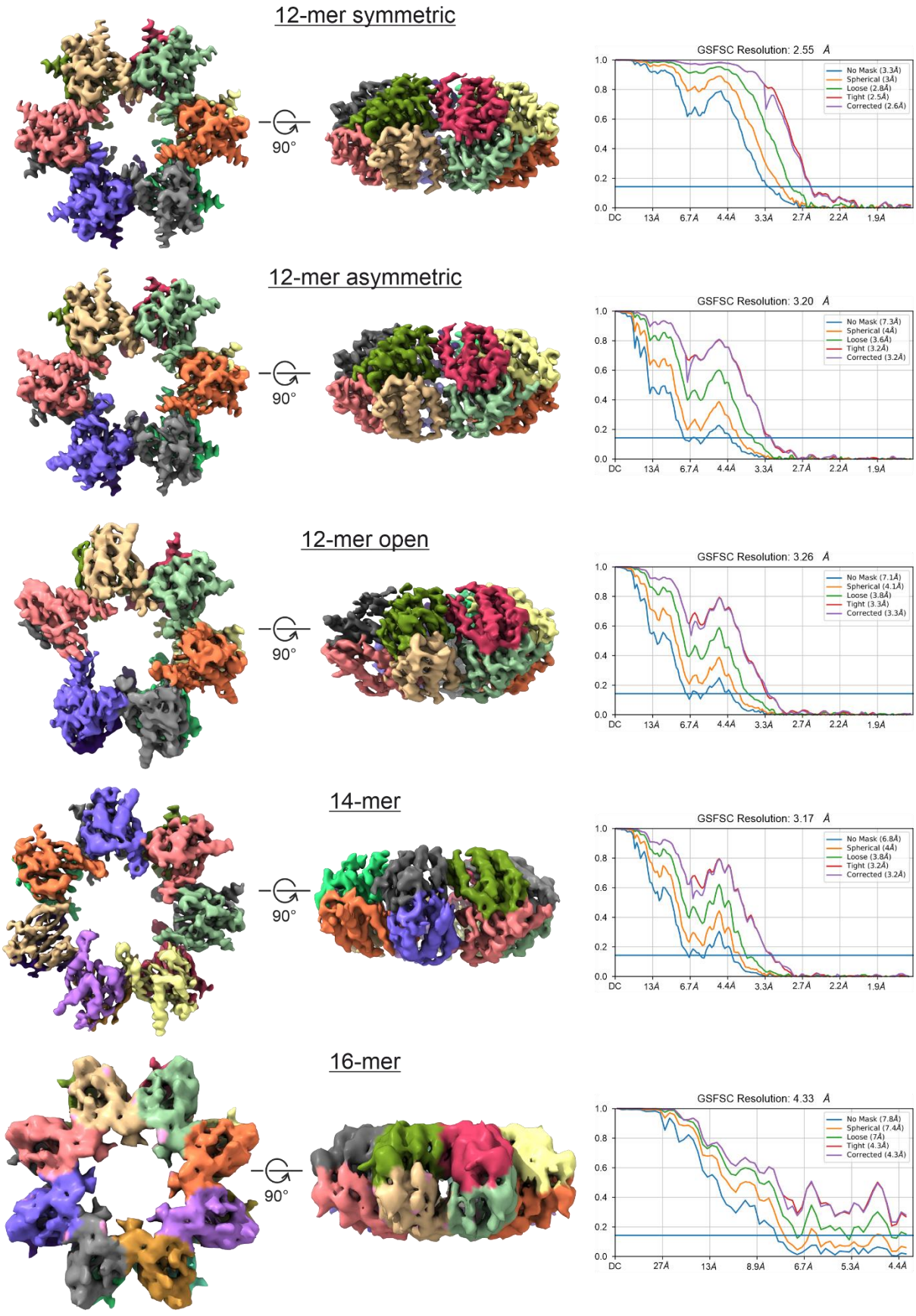



Figure S7: cryoEM reconstructions of the CaMKIIβ hub.

Reconstructions of the different oligomeric and conformational states of the CaMKIIβΔ16,17 central hub. Maps are coloured according to subunit. All maps were contoured at 9 RMSD, except the 16-mer, which was contoured at 6 RMSD. Right: Gold standard Fourier shell correlation of the respective maps using different masks.

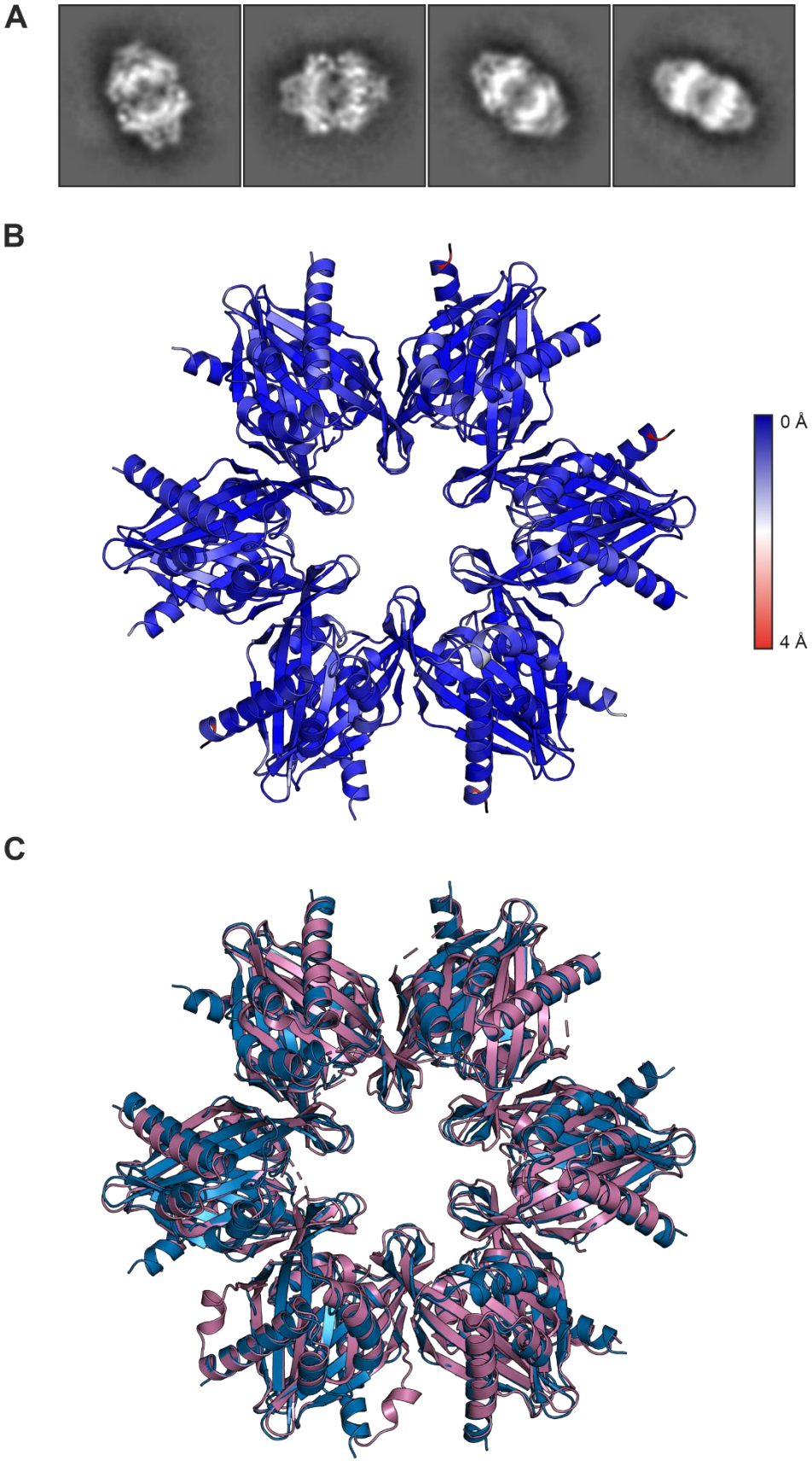
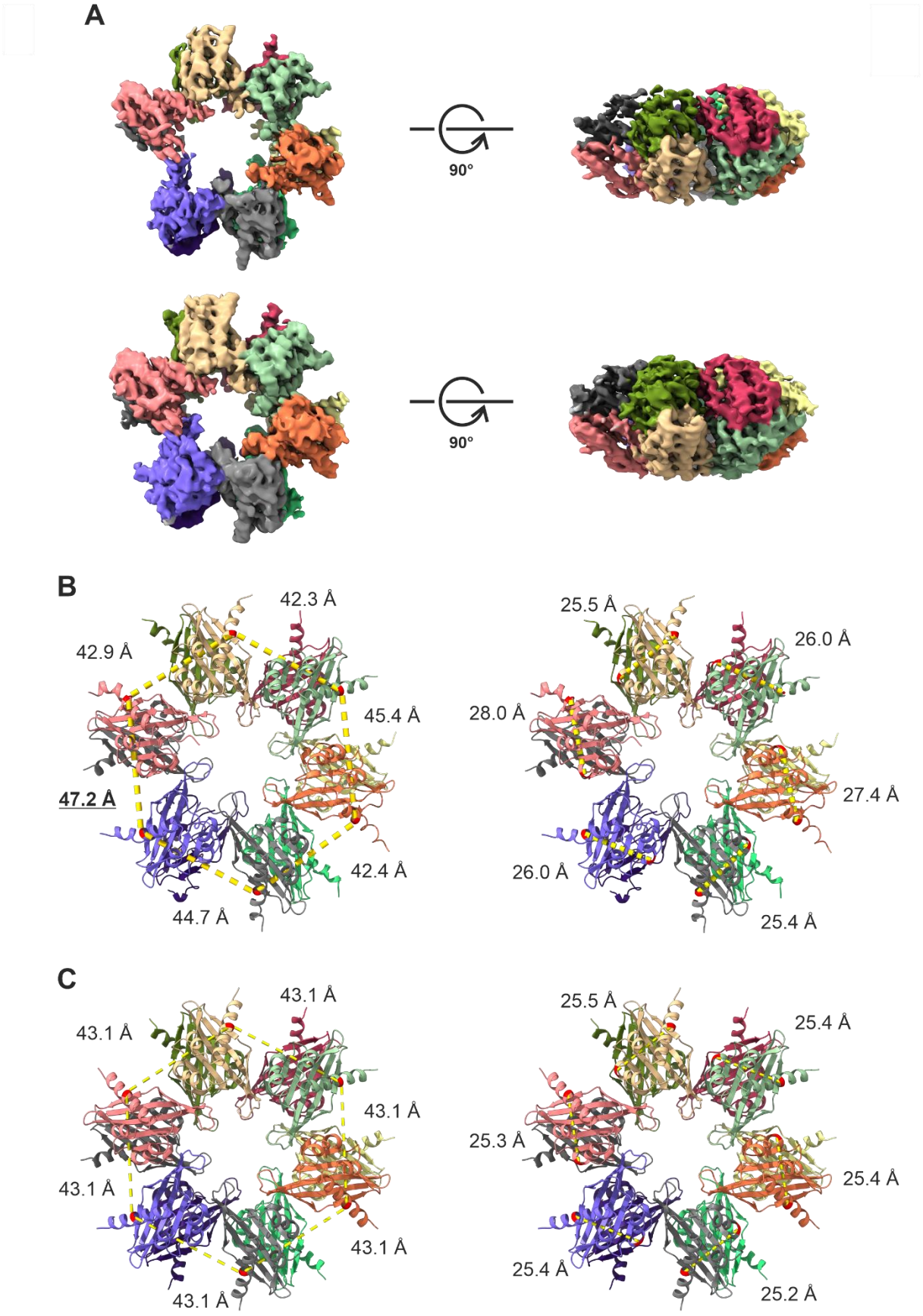

A
B
0 Å
4 Å
C

Figure S8: cryoEM analysis of the 12-mer.

(A) 2D classes of side views of the CaMKIIβΔ16,17 hub. CaMKII was expressed in insect cells and purified via Strep-affinity and size exclusion chromatography. The purified protein was frozen vitrified and data collected on a FEI Titan Krios electron microscope. 2D classes were generated using cryoSPARC (Punjani *et al.*, 2017). (B) Final structure of the 12-mer (dodecamer) of the CaMKIIβ hub. Displayed is the symmetric model, coloured based on the root mean square deviation (RMSD) between the Cα atoms of the symmetric and asymmetric model. (C) Overlay of the symmetric 12-mer (blue) with the published hub of CaMKIIα (magenta, pdb: 5IG3) (McSpadden *et al.*, 2019).


A
90°
90°
B
42.3 Å
42.9 Å
45.4 Å
47.2 Å
42.4 Å
44.7 Å
25.5 Å
26.0 Å
28.0 Å
27.4 Å
26.0 Å
25.4 Å
C
43.1 Å
43.1 Å
43.1 Å
43.1 Å
43.1 Å
43.1 Å
25.5 Å
25.4 Å
25.3 Å
25.4 Å
25.4 Å
25.2 Å

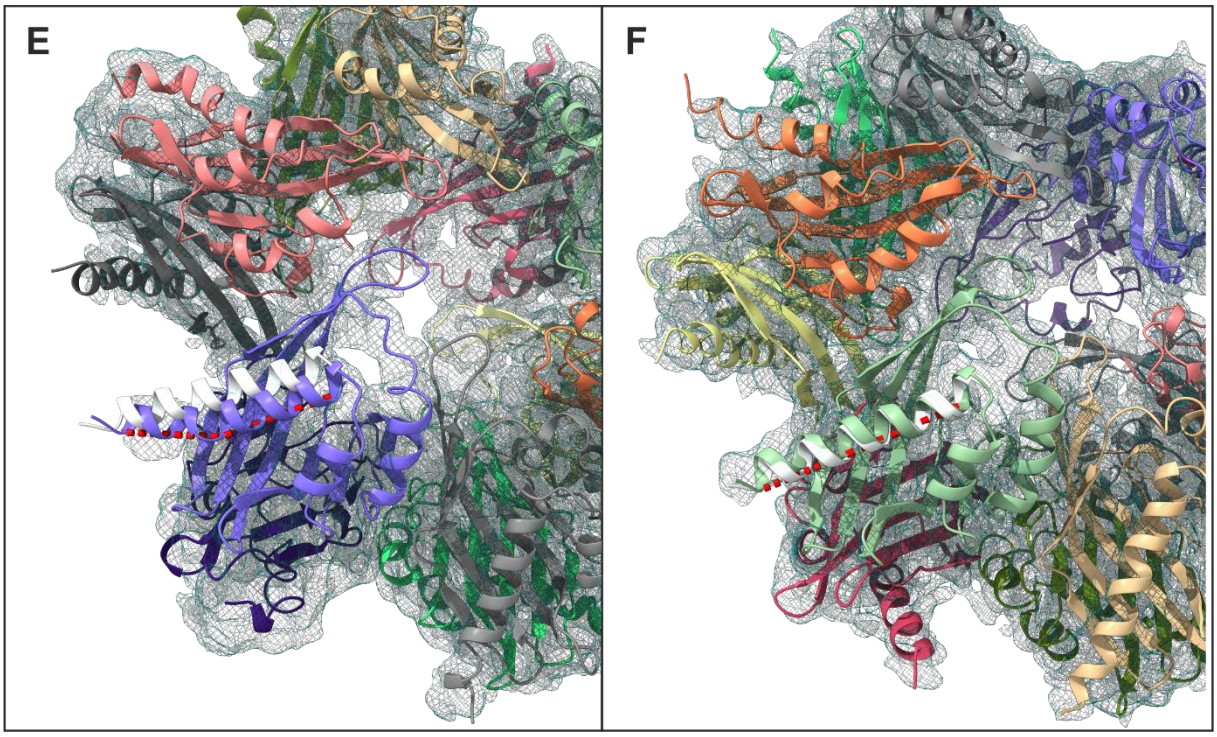



Figure S9: cryoEM analysis of the open 12-mer.

(A) Reconstructions of the open 12-mer state of the CaMKIIβΔ16,17 central hub. Maps are coloured according to subunit. Top: map contoured at 9 RMSD, bottom: map contoured at 6 RMSD. (B,C) Inter- and intra-subunit distances of the refined model for the open and closed 12-mer states. Structures are coloured according to chain. Selected residues for distance measurements are highlighted in red, distances are symbolized by dashed red lines. Numbers indicate distances in Å. Left: inter-subunit distances between the Cα-Atoms of lysine 401. Right: intra-subunit distances between the Cα-Atoms of lysine 401 and aspartic acid 442. B: open 12-mer, C: closed symmetric 12-mer. (D,E) Refined model for the open 12-mer state. Structures are coloured according to chain. A red line highlights the N-terminal helix that is curved in one of the subunits neighbouring the gap in the open 12-mer (D) but not other subunits (E). The respective helix of the closed symmetric 12-mer is shown in white. The electron density of the open 12-mer contoured at 6 RMSD is shown as light blue mesh.

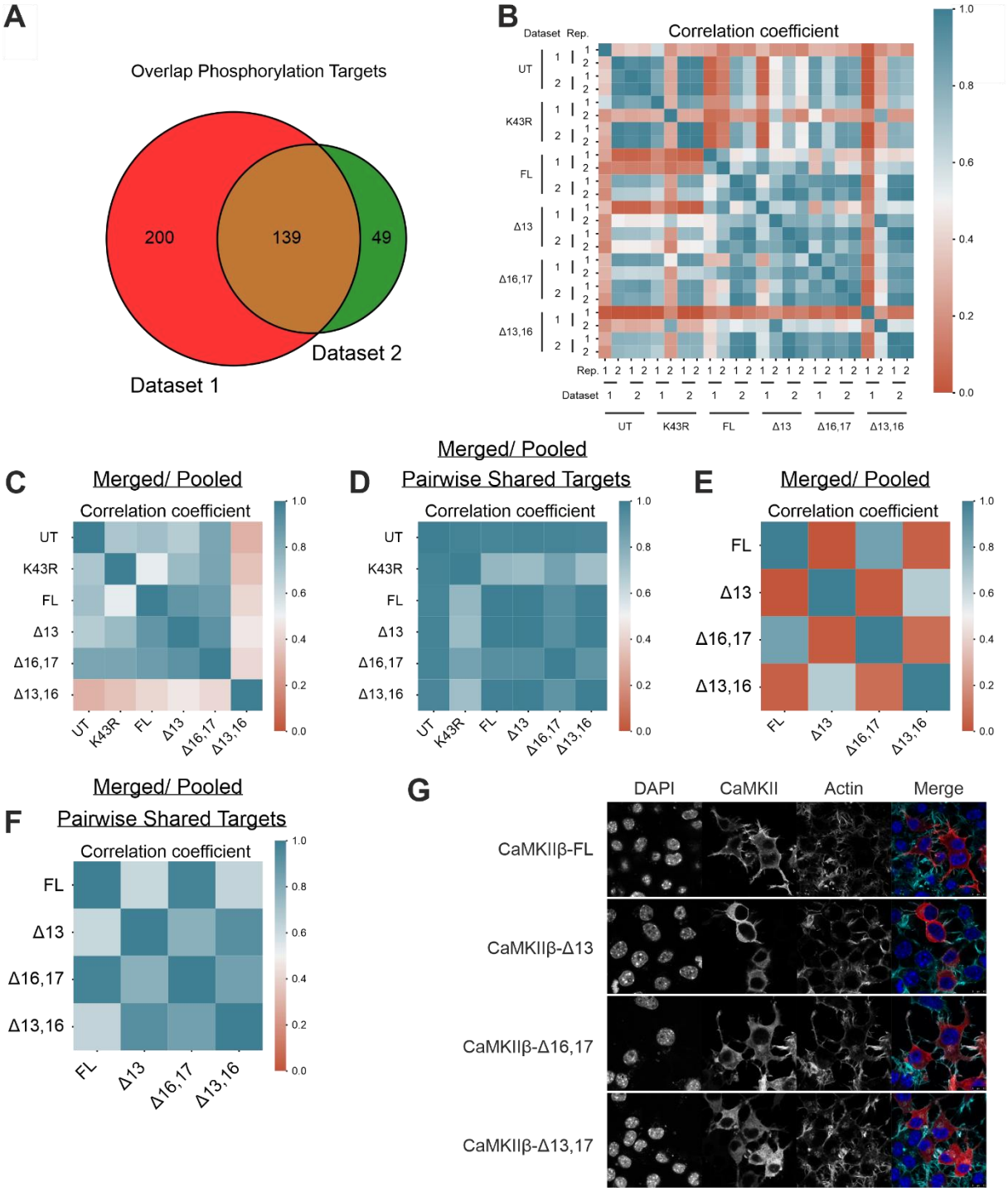



Figure S10: CaMKIIβ isoforms differ in their substrate spectra.

(A) Venn diagram showing the overlap between the two datasets, based on unique phosphorylation sites. (B-F) Correlation matrices (Pearson) based on the intensity values of individual phosphorylation sites. (B) Comparison of the individual samples of the unpooled and unmerged datasets, ordered according to isoform. Rep.: technical replicates (C) Comparison of the merged and pooled dataset. (D) Comparison of the merged and pooled dataset, analysis was restricted to pairwise shared targets. (E) Comparison of the merged and pooled dataset. Phosphorylation sites detected in the control samples were removed. The outlier (Dataset 1, isoform Δ13,16, replicate 1) was removed from the analysis. (F) Comparison of the merged and pooled dataset. Phosphorylation sites detected in the control samples were removed. Analysis was restricted to pairwise shared targets. (G) Subcellular localization of CaMKIIβ isoforms, expressed in N2a cells with an N-terminal tdTomato-tag. Nuclei were stained with DAPI, the actin filament with phalloidin-647. Individual channels are displayed as black and white, merged shown in colour. Blue: DAPI, red: CaMKII, cyan: actin.

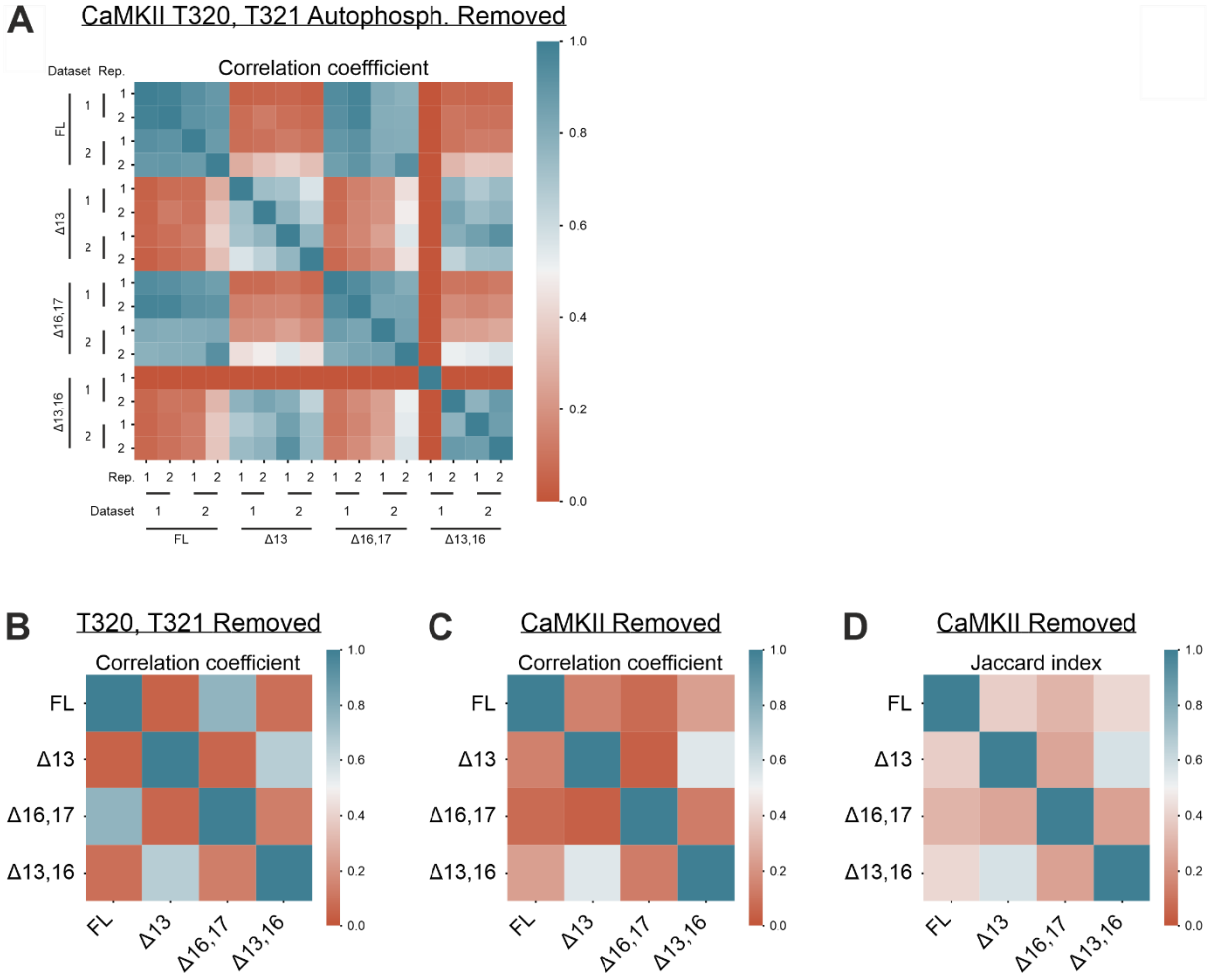
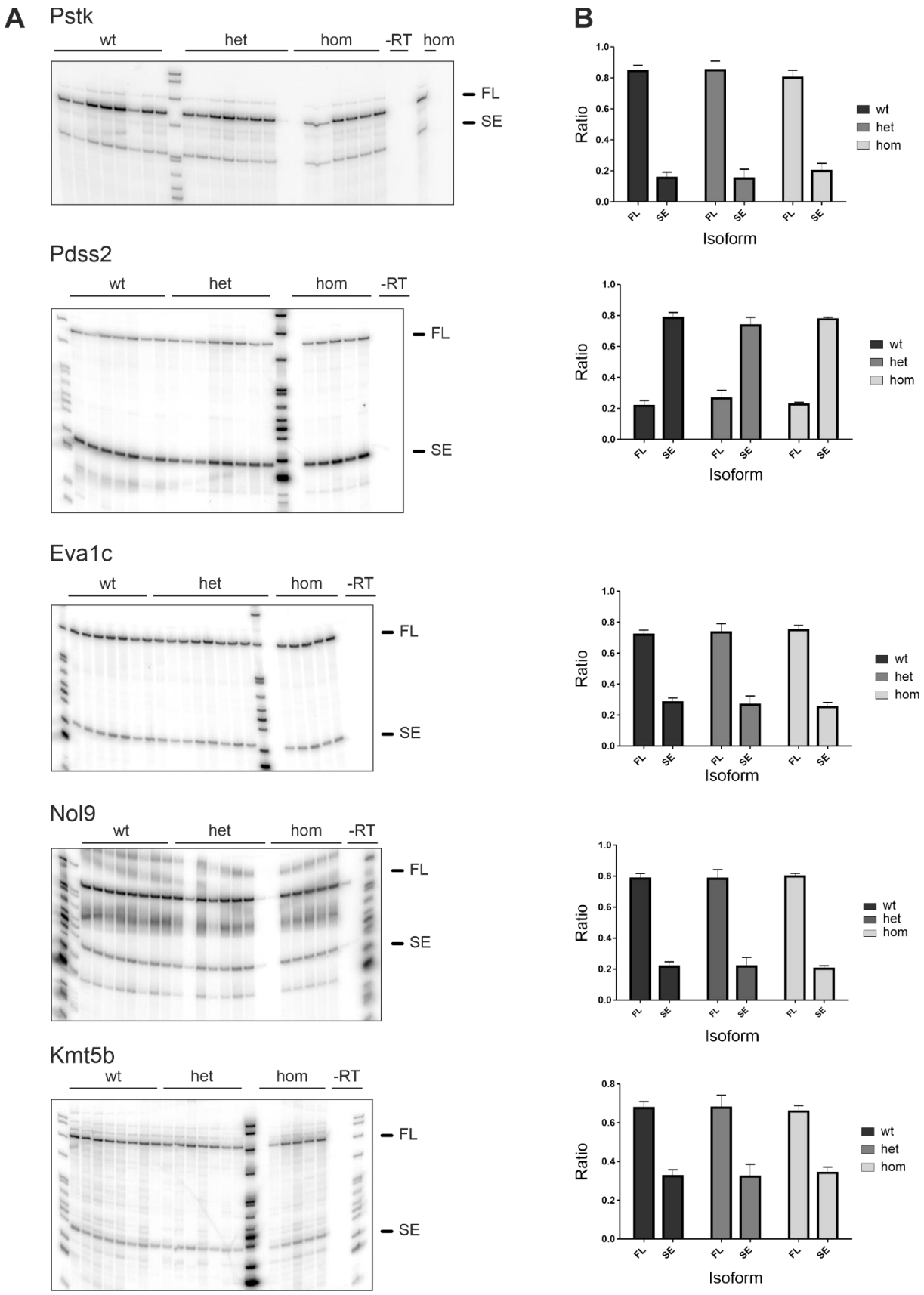



Figure S11: CaMKIIβ autophosphorylation differs between isoforms.

(A-C) Correlation matrices (Pearson) based on the intensity values of individual phosphorylation sites. (A) Comparison of the individual samples of the unpooled and unmerged datasets, ordered according to isoform. The CaMKII autophosphorylation sites T320 and T321 were removed from the analysis. Rep.: technical replicates (B) Similar to B, but with the pooled and merged dataset. (C) Similar to C, but all CaMKII autophosphorylation sites were removed from the analysis (D) Similarity matrix (Jaccard index) of the merged and pooled dataset. Phosphorylation sites detected in the control samples, and all CaMKII autophosphorylation sites were removed from the analysis. Intensity values were transformed into Boolean data (detected/ not detected).


A
Pstk
wt
het
hom
-RT
hom
FL
SE
Pdss2
wt
het
hom
-RT
FL
SE
Eva1c
wt
het
hom
-RT
FL
SE
Nol9
wt
het
hom
-RT
FL
SE
Kmt5b
wt
het
hom
-RT
FL
SE
B
Ratio
Isoform
FL
SE
wt
het
hom

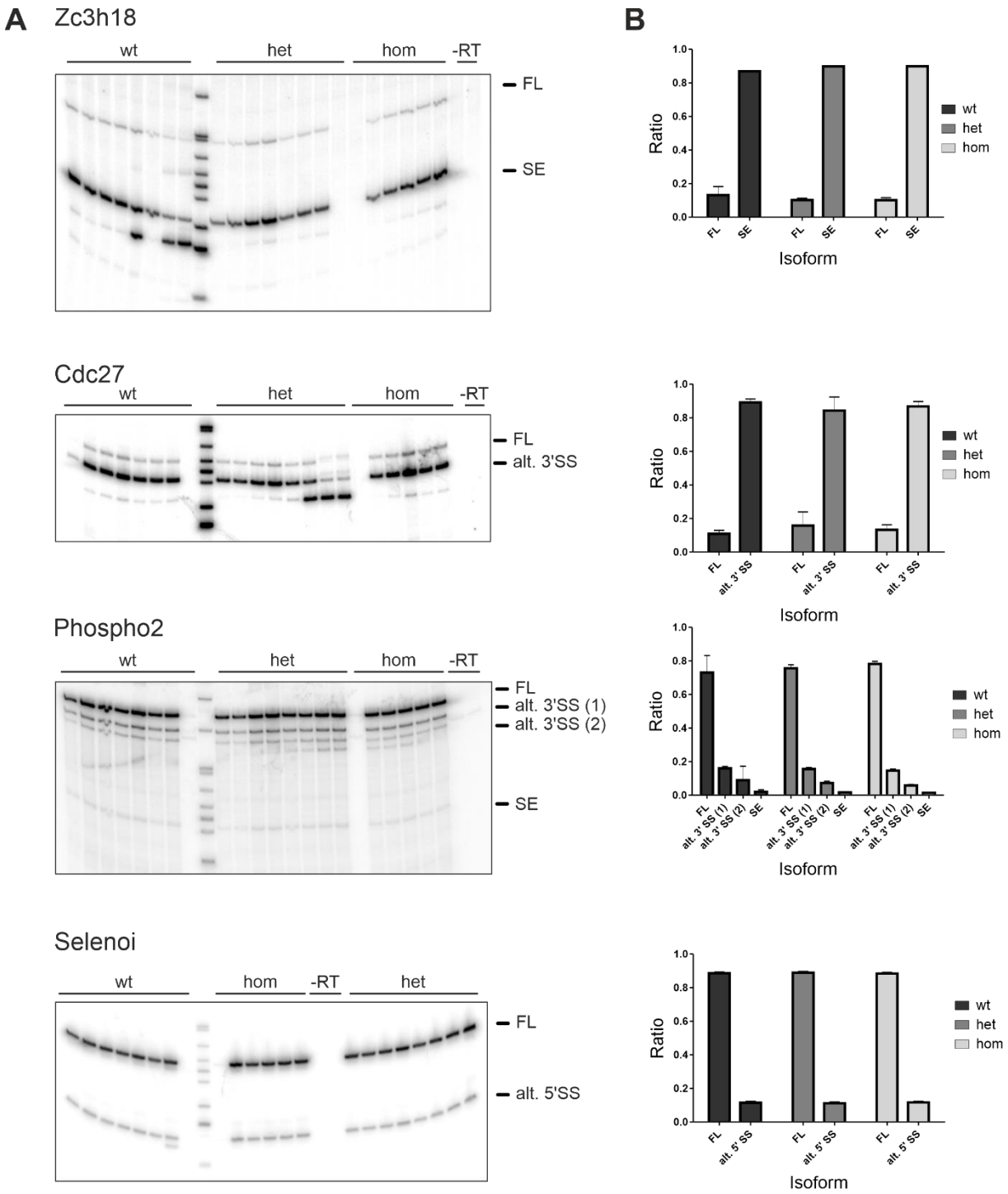



Figure S12: Alternative splicing is not globally affected in the mouse model.

(A) Validation of RNA-Seq targets found to be differentially spliced in the mouse model. 9 high-scoring targets were selected and validated via radioactive RT-PCR with cerebellum RNA from 8 wild type (wt) mice, 7 heterozygous (het) mice and 6 homozygous (hom) mice. (B) Quantifications of A. FL: full-length isoform, SE: skipped exon isoform, alt.3`SS: alternative 3`splice site, alt.5`SS: alternative 5' splice site.

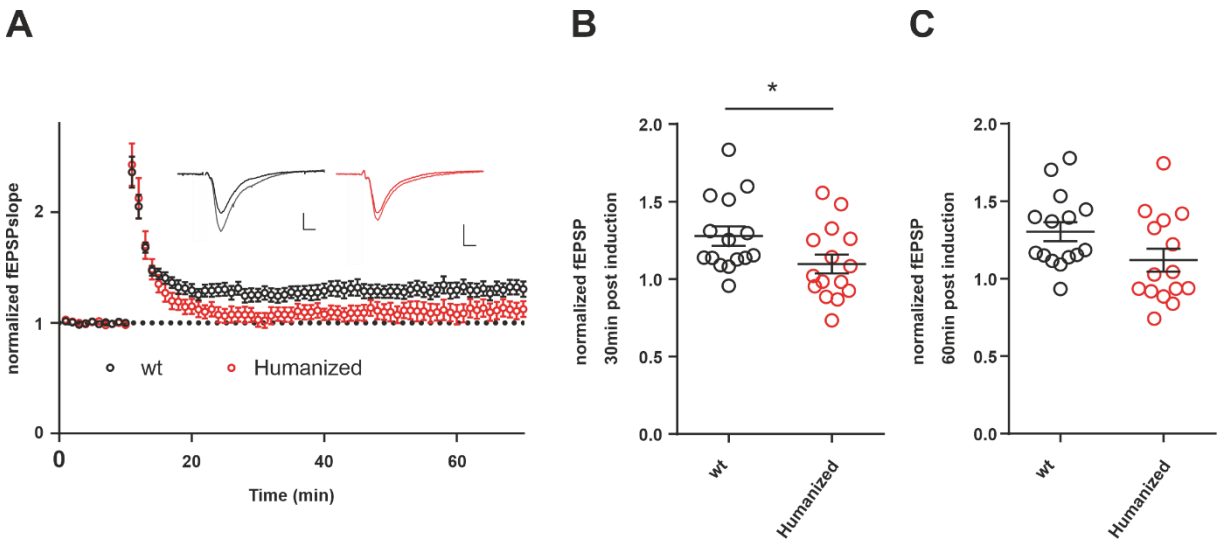
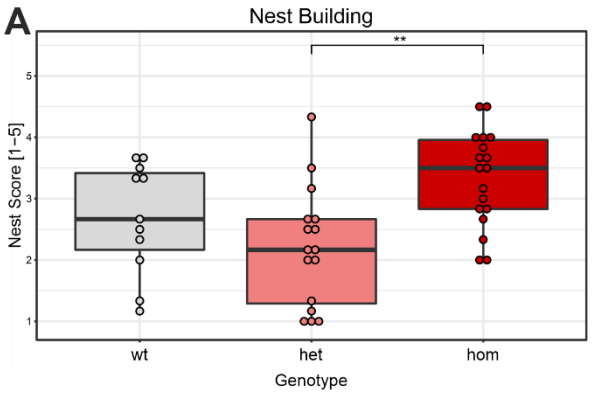
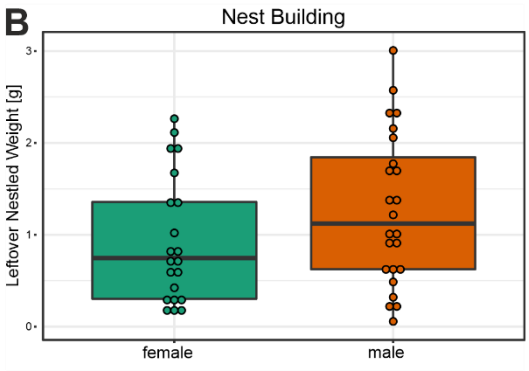
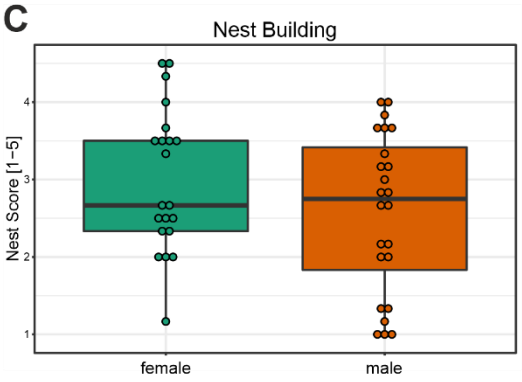


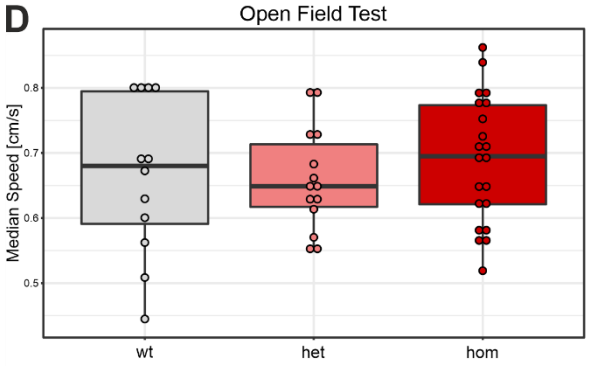

Figure S13: Long term potentiation is severely impaired in the mouse model.

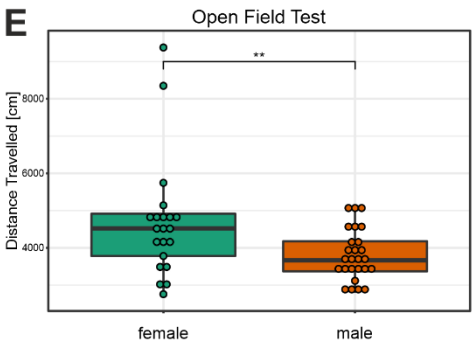
(A) Time course of LTP induction in CA3-CA1 synapses in acute hippocampal slices. LTP was induced after 10 min with a four trains of 100 Hz, 1s. Example traces show average of baseline and potentiated field excitatory postsynaptic potentials (fEPSP) 60 min after LTP induction. Scale bar: 0.2 mV / 5 ms. wt (wild type): 15 slices, 6 mice, humanized (humanized strain, homozygote): 12 slices 6 mice. (B, C) Bar plots depicting the field EPSP slope after 30 min (B) or 60 min (C) after LTP induction. *p<0.05, **p<0.01, calculated by ANOVA followed by Dunnet's test.

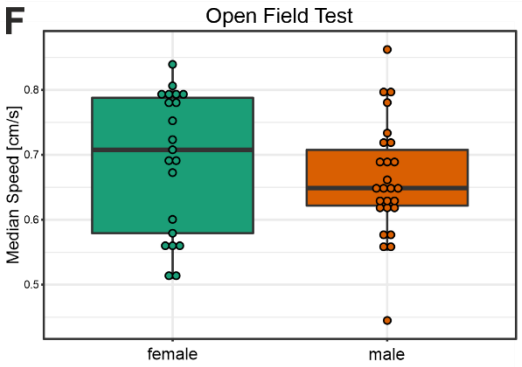

A
Nest Building
**
Nest Score [1–5]
wt
het
hom
Genotype

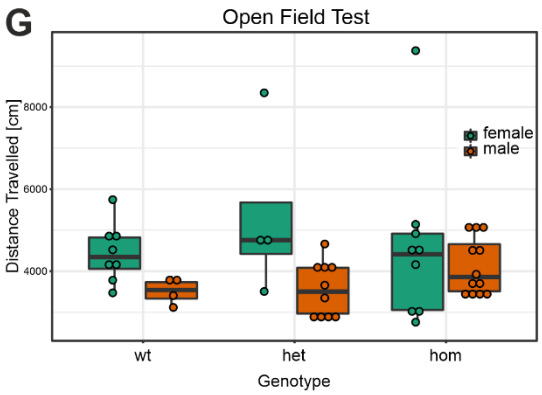

B
Nest Building
Leftover Nestled Weight [g]
female
male

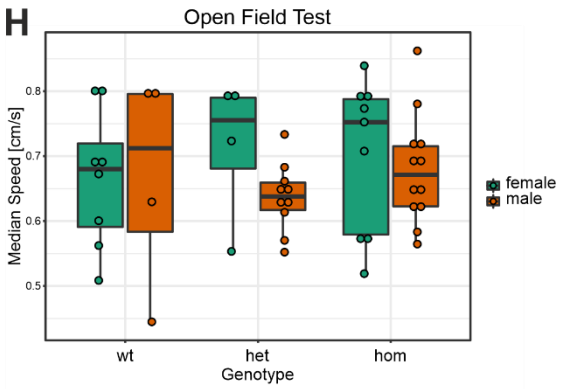

C
Nest Building
Nest Score [1–5]
female
male


D
Open Field Test
Median Speed [cm/s]
wt
het
hom


E
Open Field Test
**
Distance Travelled [cm]
female
male


F
Open Field Test
Median Speed [cm/s]
female
male


G
Open Field Test
Distance Travelled [cm]
female
male
wt
het
hom
Genotype


H
Open Field Test
Median Speed [cm/s]
female
male
wt
het
hom
Genotype

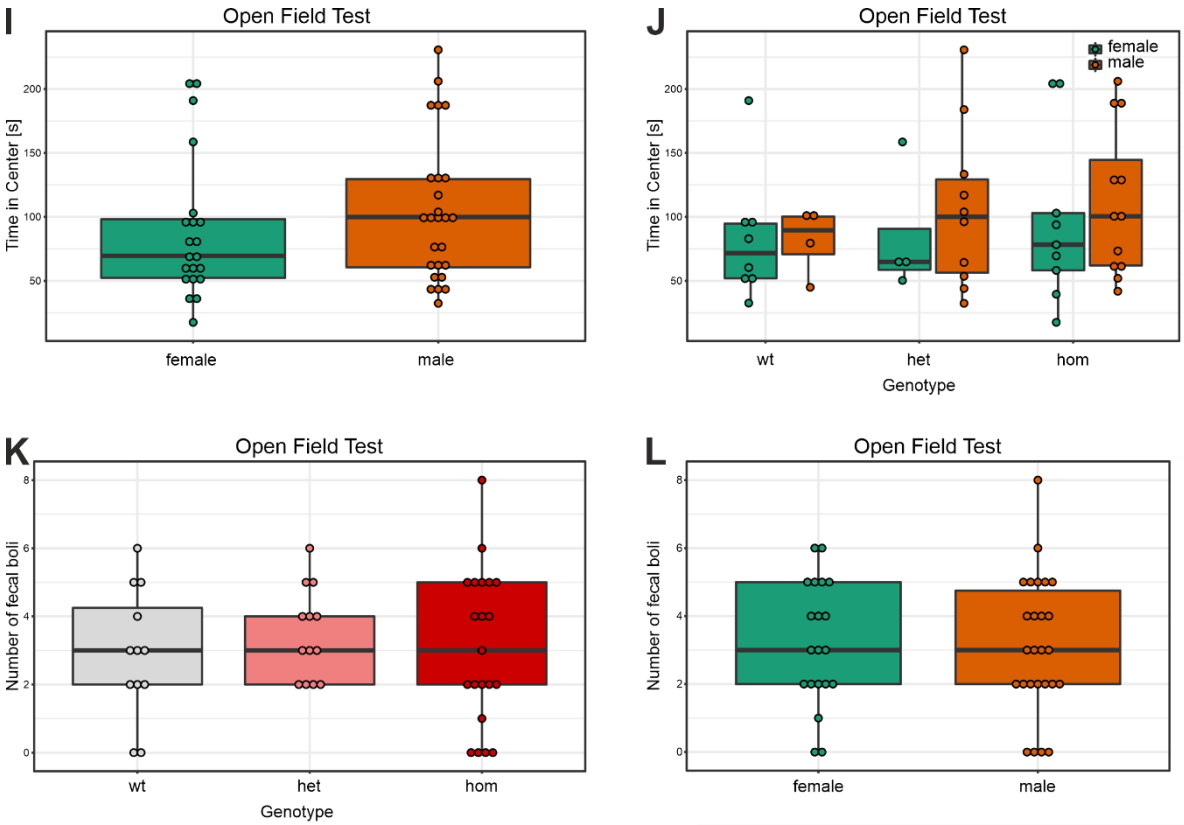



Figure S14: Behavioural characterization of the mouse model.

(A, B, C) Results of the nest building test, displayed as nest score (A, C) or leftover nestlet weight (B). Mice were singly housed and supplied with cotton pads of 3.0 g overnight. Nests were scored according to a previously described scale of 1-5 (Deacon, 2006), with 5 being a perfect nest. Unused material was collected and weighted. (D-L) Results of the open field test. Mice were allowed to explore an open arena (40 x 40 cm) freely for 10 min. (D-L) Locomotor activity measured as distance travelled during the session or median speed. (I-L) Anxiety levels measured as time spend in the center (20 x 20 cm) of the arena, or defecation during the session. *p<0.05, **p<0.01, ***p<0.001 calculated by Student‘s t-test (A) or Wilcoxon signed-rank test (D). Adjusted for multiple comparisons using Holm‘s method. wt: wild type mice (nest building: 8 female, 3 male; open field: 8 female, 4 male), het: heterozygous mice (nest building: 6 female, 10 male; open field: 4 female, 10 male) hom: homozygous mice (nest building: 7 female, 11 male; open field: 9 female, 12 male) of the humanized strain.

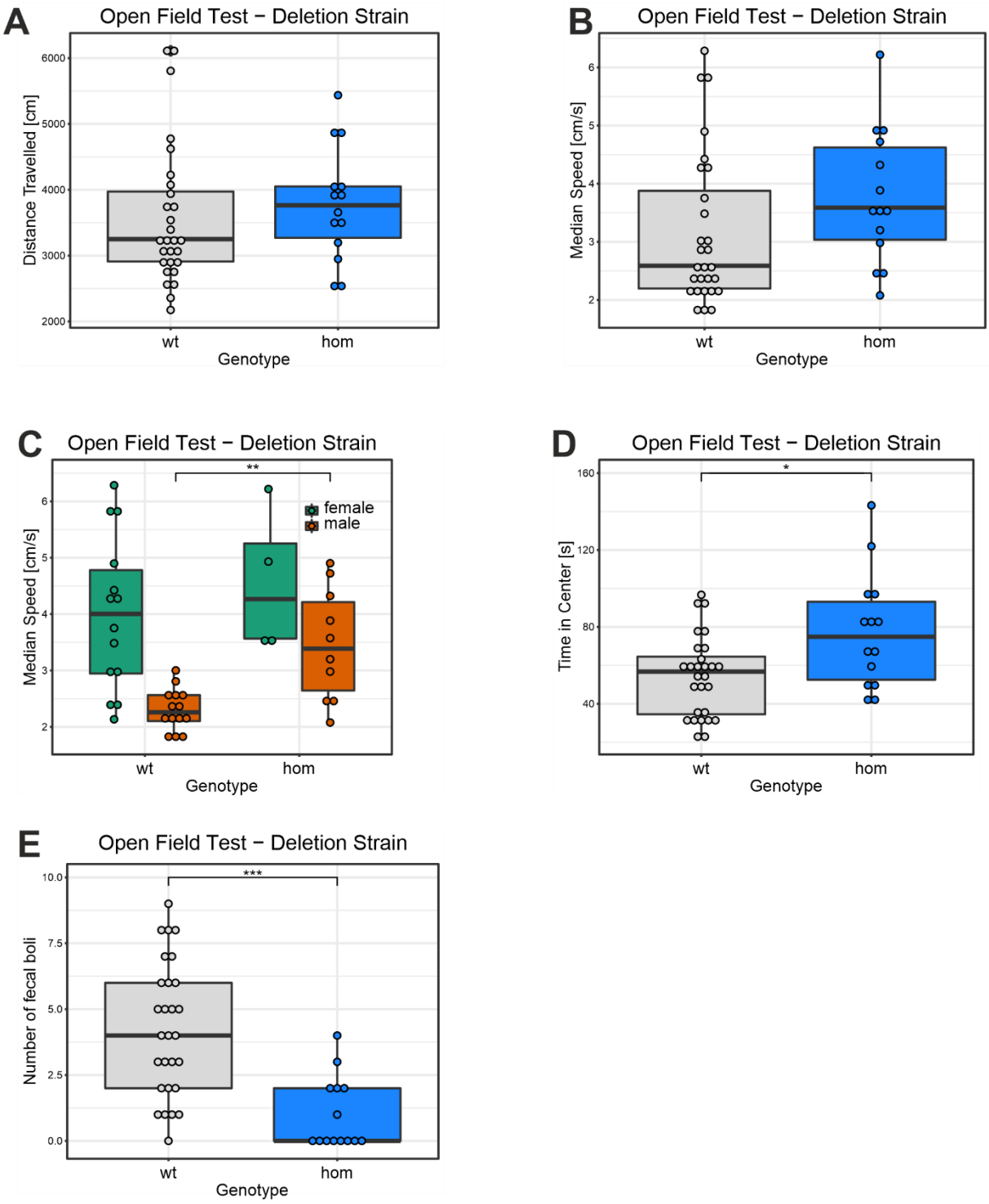



Figure S15: Behavioural characterization of the mouse model, deletion strain.

Results of the open field test for the deletion strain. Mice were allowed to explore an open arena (40 x 40 cm) freely for 10 min. (A-C) Locomotor activity measured as distance travelled or median speed. (D,E) Anxiety levels measured as time spend in the center (20 x 20 cm) of the arena (D), or defecation during the session (E). *p<0.05, **p<0.01, ***p<0.001 calculated by Student's t-test (D) or Wilcoxon signed-rank test (C, D). Adjusted for multiple comparisons using Holm's method (C). wt: wild type mice (n=14, 7 female, 7 male), hom: homozygous mice (n=14, 4 female, 10 male) of the deletion strain.