# Molecular insights into the regulatory interactions of Dystrophia Myotonica Protein Kinase

## Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

**Pilar Garcia**
aus Spanien

Basel, 2006

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Olga Mayans
Dr. Jörg Stetefeld
Prof. Markus Rüegg

Basel, den 5. 7, 2006

Prof. Dr. Hans-Jakob Wirz
Dekan

# Declaration

I declare that I wrote this thesis, **Molecular insights into the regulatory interactions of Dystrophia Myotonica Protein Kinase**, with the help indicated and only handed it into the Faculty of Science of the University of Basel and to no other faculty and no other university.

# Abstract

Dystrophia Myotonica Protein Kinase (DMPK) is the defining member of a family of complex, multidomain kinases of major biomedical relevance. These kinases are characterized by a highly conserved catalytic domain and a coiled-coil motif (CC) involved in the regulation of their activity. DMPK has been related to a progressive neuromuscular disorder known as Myotonic Dystrophy (DM), the most prevalent muscular dystrophy in adults. DM is a multisystemic disease characterized by myotonia and progressive skeletal muscle weakness. Affected patients suffer from DMPK insufficiency due to retention of mRNA transcripts in the nucleus. However, it is not currently understood how low levels of this kinase affect cellular function and its involvement in disease remains elusive. The sequestration of the CUG-binding protein (CUGBP), a putative DMPK substrate by a trinucleotide repeat extension in the 3' non-coding region of the DMPK gene has been suggested to play a role in the disease. DMPK is speculated to be involved in the modulation of the plasma membrane depolarization and reorganization of the actin cytoskeleton during tissue development, possibly, acting as a downstream effector of the actin cytoskeleton-linked GTPase Rac1. In order to gain a better understanding of the complex mechanism of regulation of this kinase, its oligomeric state and intrasteric regulation were investigated. Self-assembly is crucial for the regulation of DMPK and related kinases. Their CC domains are thought to form dimeric arrangements and, thus, to mediate dimerization in this family of kinases. In addition, a role as intrasteric regulators has been attributed to these moieties. In the current work, the role of the CC domain of DMPK in kinase assembly and possible Rac1 docking has been analyzed using structural, biophysical and biochemical approaches. Contrary to expectations, the self-assembly of DMPK is not dictated by the association properties of its CC domain, instead, it appears driven by sequence segments flanking both N- and C-termini of the catalytic kinase fraction, leading to the formation of head-to-head dimers. Our findings support a shared pattern of assembly across DMPK, ROCKs and MRCK members of this family.

# List of Figures

# List of Abbreviations

| | |
|---|---|
| **aa** | Amino acid |
| **ACK** | Activated Cdc42-associated tyrosine kinase |
| **APS** | Ammonium persulphate |
| **AU** | Analytical ultracentrifugation |
| **β-ME** | beta-mercaptoethanol |
| **BCA** | Bicinchoninic acid |
| **BN-PAGE** | Blue-Native Polyacrylamide Gel-Electrophoresis |
| **bp** | Base pair |
| **BSA** | Bovine serum albumin |
| **CC** | Coiled-coil |
| **cDNA** | Complementary DNA |
| **CUG-BP** | CUG-binding protein |
| **CELF** | CUG-BP and ETR-3-like factor |
| **CH** | Citron homology domain |
| **Clc-1** | Skeletal muscle chloride channel |
| **CR** | Cysteine-rich domain |
| **CRIB** | Cdc42/Rac interactive binding site |
| **cTNT** | Cardiac troponin T |
| **CV** | Column volume |
| **DHPR** | Dihydropyridine receptor |
| **DM** | Myotonic Dystrophy |
| **DM1** | Myotonic Dystrophy Type 1 |
| **DM2** | Myotonic Dystrophy Type 2 |
| **DMPK** | Dystrophia Myotonica Protein Kinase |
| **DMRFK** | DMPK-related family of kinases |
| **DMWD** | Dystrophia Myotonica-containing WD repeat motif |
| **DNA** | Deoxyribonucleic acid |
| **DTT** | Dithiothreitol |
| **EDTA** | Ethylenediaminetetraacetic acid |
| **ELAV** | Embryonically lethal abnormal vision |

| | |
|---|---|
| **ETR-3** | ELAV-type RNA binding protein 3 |
| **FCGRT** | Fc fragment of IgG receptor |
| **FRAXA** | Fragile X syndrome |
| **FRAXE** | Fragile XE syndrome |
| **GAP** | GTP activating protein |
| **GBD** | GTPase-binding domain |
| **GDI** | Guanine nucleotide dissociation inhibitor |
| **GEP** | Guanine nucleotide exchange factor |
| **GIPR** | Gastric inhibitory polypeptide receptor |
| **GTPγS** | Guanosine-5'-0-3-thiotriphosphate |
| **$His_6$-tag** | Hexahistidine tag |
| **hnRNP** | Heterogeneous nuclear ribonucleoprotein |
| **IPTG** | Isopropil-β-D-thiogalactopyranoside |
| **IR** | Insulin receptor |
| **IS** | Inhibitory switch |
| **LB** | Luria-Bertani broth medium |
| **LMW** | Low molecular weight |
| **LR** | Leucine-rich |
| **MALDI-TOF** | Matrix assisted laser desorption-Time of flight |
| **MBNL** | Muscleblind-like protein |
| **MD** | Muscular dystrophy |
| **MKBP** | DMPK-binding protein |
| **MM** | Molecular mass |
| **MRCK** | Cdc42-binding kinase |
| **mRNA** | Messenger ribonucleic acid |
| **MYPT1** | Myosin phosphatase targeting subunit 1 |
| **OD** | Optical density |
| **PEG** | Polyethylene glycol |
| **PCR** | Polymerase chain reaction |
| **PDZ** | Post-synaptic density-95/discs large/zonula occludens-1 domain |
| **PH** | Pleckstrin homology domain |

| | |
|---|---|
| **PMF** | Mass fingerprinting |
| **PMSF** | Phenylmethylsulphonyl fluoride |
| **PRD** | Proline-rich domain |
| **PROMM** | Proximal Myotonic Myopathy |
| **RBD** | Rho-binding domain |
| **REM** | Rho effector motif |
| **ROCK** | Rho-associated kinase |
| **RT** | Room temperature |
| **RRM** | RNA recognition motif |
| **RSHL1** | Radial spokehead-like 1 |
| **SCA** | Spinocerebellar ataxia |
| **SDS-PAGE** | Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis |
| **SIX5** | Sine oculis homeobox homolog 5 |
| **TEV** | Tobacco etch virus |
| **UTR** | Untranslated region |
| **VR** | Variable region |
| **WASP** | Wiskott-Aldrich Syndrome Protein |
| **ZNF9** | Zinc finger protein 9 |

# Acknowledgements

I would like to express my most sincere gratitute to my supervisor Prof. Olga Mayans for giving me the opportunity to work in her group and for providing me with a very interesting project. I highly appreciate her guidance, support and comprehension both scientifically and emotionally throughout my PhD study, especially during the last year. Thanks for everything.

I thank Dr. Marco Marino for never loosing his patience and always being available for valuable advice. He is a great colleague to work with and I wish him all the best in his new role as a teacher and as a father of little Luca.

I also would like to thank Dr. Jörg Stetefeld and Prof. Markus Rüegg for kindly agreeing to be my second examiner and chairman, respectively.

I would like to thank the people in my group Darko, Elli, Michael and Zöhre. It has been a pleasure to work with them. I wish Michael and Rainer all the best for their exams and future carriers.

My special regards to Zöhre and Elli for being good friends during all my PhD life. I will miss all the nice coffee breaks. A thank you also goes to Zöhre for helping me with cloning.

Thanks also to Mrs. Ute Gruetter for always being so kind and helpful with the administrative work. I thank Roland Buerki and Margrit Jenny for the support in computer installations.

Thanks to Dr Paul Jenö and Ariel Lustig for their expert analyses on mass fingerprinting and analytical ultracentrifugation.

I am also greatful to Carmen Chan, Dinesh Palanivelu, Fabio Casagrande, Frida, Joachim Koeser and Sara Paulillo for their excellent company and all the laughs we shared. Frida and Sara I wish you both the best luck with your exams. My special regards to Ainhoa Mielgo. Her continous care and support throughout these years are highly appreciated. I wish you all the best in San Diego and I hope to see you soon.

Many thanks to my friends in Spain Esther, Marta, Mercedes, Pilar, Rosana, Rocio Enriquez, Rocio Torquemada, Silvia, Cristina, my cousin Maria and many others for

your unconditional support despite the distance and for being always there. I am looking forwards to see you all.

Thanks to the Geamed people, Alicia, Enrique, Levan, Rafael, Rosa and the others for their companionship and support. Special thanks go to Fernando Chacon for his understanding and encouragement in difficult moments.

To my dearest family, thank you all so much, my wonderful parents and brothers Carlos, Javier y Pedro for your endless support, encouragement and love.

Finally but most importantly, my deepest thank-you to my beloved brother Javier, thanks for the marvelous seventeen years, full of happiness and many memorable moments that you have shared with us. Thanks for being an example of strength and determination. To him I dedicate this work.

Te adora, tu hermana Pilar

# Contents

# Chapter 1

## Introduction

# 1.1 Myotonic Dystrophy

## 1.1.1 Introduction

Myotonic Dystrophy (DM), also known as dystrophia myotonica or Steinert's disease, was first described one century ago (Batten 1909; Steinert 1909). It is the most common form of muscular dystrophy in adults, with an incidence of 1:8000 (Harper 1989). This is an autosomal dominant neuromuscular disorder characterized by an extremely variable clinical phenotype. Onset is usually in early adulthood and the life span of affected individuals is typically six decades. The causative genetic agent for DM was identified in 1992 as a CTG repeat expansion in chromosome 19q13.3 within the 3' untranslated region (UTR) of a gene coding for a Ser/Thr kinase, Dystrophia Myotonica Protein Kinase (DMPK) (Figure 1.1) (Brook et al. 1992; Fu et al. 1992; Mahadevan et al. 1992).

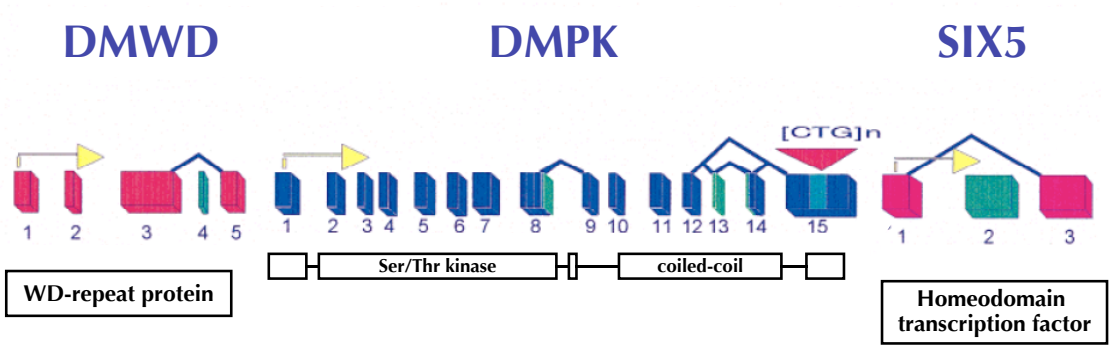



**Figure 1.1: Schematic representation of the DM locus.** The DM locus is strongly conserved between mouse and man and contains three genes, DMWD, DMPK and SIX5, located closely together in a gene-dense region. DMWD gene is split over five exons and encodes a WD-repeat protein, DMPK gene is split over 15 exons and encodes a Ser/Thr protein kinase, and SIX5 gene is split over three exons and encodes a homeodomain transcription factor. The unstable $(CTG)_n$ repeat is located in the 3' proximal exon (exon 15) of the human DMPK gene.Transcription initiation sites are indicated with arrows and alternative splice events with curved lines. DMWD and SIX5 are coloured in red, DMPK in blue and alternative exons are in green (Groenen and Wieringa 1998).

## 1.1.2 Clinical Features

DM symptoms are characterized by involuntary persistence of muscle contraction (myotonia) and progressive weakness and wasting of muscle (dystrophy). However, this is a multisystemic disorder with a consistent constelation of associated disturbances including cardiac conduction defects[1], smooth muscle[2] involvement, mental changes, hypersomnia, ocular cataract[3], digestive problems and diabetes mellitus type 2[4]. Also common are testicular atrophy and premature frontal balding in men. In several patients, tumors have been also found but a direct link to DM is yet to be proven (Harper 1989; Brewster 1998).

A congenital form of DM exhibits the most severe phenotype and faces a neonatal mortality rate of 25 %. Features include severe mental retardation, hypotonia, as well as facial diplegia[5] and compromised respiratory function (Harper 1989).

## 1.1.3 Myotonic Dystrophy Type 2

In 1995, several reports described families with dominantly inherited multisystemic disorders that were genetically distinct from the classical DM (DM1) (Ricker et al. 1994). This new form of DM was designated Myotonic Dystrophy Type 2 (DM2) or Proximal Myotonic Myopathy (PROMM). The manifestations of DM1 and DM2 are similar. The muscle histopathology in DM2 resembles that seen in DM1. The non-muscle manifestations of DM1 are also observed in DM2, including cardiac arrhythmias, frontal balding, insulin resistance, fatigue and cataracts.

[1] A cardiac conduction disorder is a dysfunction in the heart's electrical conduction system, which can make the heart beat too fast, too slow, or at an uneven rate.

[2] Smooth muscle is a type of non-striated muscle, found within the "walls" of hollow organs such as blood vessels, the bladder, the uterus, and the gastrointestinal tract.

[3] Cataract is any opacity of the lens or capsule of the eye, causing impairment of vision or blindness.

[4] Diabetes mellitus type 2 (formerly called non-insulin-dependent diabetes) is a long-term metabolic disorder that is primarily characterized by insulin resistance, relative insulin deficiency, and hyperglycemia.

[5] Facial diplegia is a paralysis of both sides of the face. DM patients present a characteristic facial diplegia with an inverted “V” shaped upper lip.

Although clinically similar to DM1, DM2 patients tend to experience less severe symptoms of cognitive impairment, ptosis[6], muscle wasting, dysphagia[7], and gastrointestinal or respiratory insufficiency. DM2 symptoms are commonly attributed to arthritis, fibromyalgia[8] and overuse of muscles. Some of the brain manifestations of DM1, such as hypersomnolence or mental retardation, have not been observed in DM2 patients. The mutation in patients with DM2, was recently characterized as and expanded CCTG repeat (Liquori et al. 2001) located in a non-coding region of the ZNF9 (zinc finger protein 9) gene on chromosome 3q21.3.

### 1.1.4 Triplet-repeat Disorders

Exactly how the expansions in the DMPK gene result in DM is still uncertain. So far, there are 14 diseases known to be caused by triplet repeats (Sobczak et al. 2003). Triplet-repeat disorders are generally classified into two groups, according to the location of the expansion. In the first group the expanded repeats are exonic, whereas in the second group the repeats are localized in non-coding sequences. The non-coding triplet-repeat disorders typically have large and variable repeat expansions that result in multiple tissue dysfunction or degeneration, and exhibit **anticipation** (genetic instability leading to longer tract lengths with each generation, which is generally accompanied by increasing severity). They include Fragile X syndrome (FRAXA), Fragile XE syndrome (FRAXE), Friedreich ataxia, spinocerebellar ataxia SCA8, SCA12 and DM1. Tetranucleotide (CCUG)n repeats involved DM2 occur in an intron.

[6] Ptosis is the paralysis of the muscles of the eyelid. It is commonly called droopy eyelids.
[7] Dysphagia is the difficulty in swallowing.
[8] Fibromyalgia is a common form of generalized muscular pain and fatigue.

## 1.2 Proposed Mechanisms of DM pathogenesis

The mechanism by which CTG expansion leads to disease in DM is not currently understood. Several pathological models, non-mutually excluding, have been proposed at the DNA, RNA and protein levels.

### 1.2.1 Haploinsufficiency of DMPK

Initially, the location of the mutation at the 3' end of a kinase gene suggested that alterations in DMPK expression might be the obvious cause of the disease. Most of the early expression studies were consistent with this hypothesis, indicating that DMPK mRNA and protein levels were reduced in patient muscle and cell culture (Fu et al. 1992; Hofmann-Radvanyi et al. 1993; Novelli et al. 1993). However, DMPK knockout mice generated to test this hypothesis did not have the typical multisystemic features of the disease. Initial reports showed that, surprisingly, the heterozygote animals presented none of the symptoms characteristic of the DM phenotype, whereas, homozygous mice, did eventually showed a very mild, late-onset myopathy (Reddy et al. 1996). Thus, the characteristic plethora of symptoms associated with the DM phenotype did not develop despite complete loss-of-function of DMPK. More recently, the same question was approached (Jansen et al. 1996) using an elaborate experimental design: in addition to knocking out the endogenous mouse DM gene, the normal human DM gene was simultaneously overexpressed. These knock-out/overexpressor mice with multiple copies of the human DM gene also did not develop the characteristic human DM phenotype. Nevertheless, mice with >20 copies of the human DM gene showed signs of cardiac conduction abnormalities. When considered together, these results suggested that DMPK may contribute to the cardiac features of DM1, but haploinsufficiency of DMPK does not cause the multisystemic clinical features of DM1 (Jansen et al. 1996).

## 1.2.2 CTG-Expansion Effects on Chromosome Structure

A second proposed mechanism is that the expanded repeat affects the expression of multiple genes in the region. Support for this hypothesis comes from the observation that the CTG expansion is a strong nucleosome-binding site that could potentially alter chromatin structure and have regional effects on the expression of DMPK and its neighboring genes (Wang et al. 1994; Otten and Tapscott 1995). This mapped six further genes: *DMWD* and *SIX5* near the vicinity of *DMPK,* and *SYMPLEKIN*, *20-D7*, human gastric inhibitory polypeptide receptor gene (*GIPR*) and Radial spokehead-like gene *RSHL1* to within 200 kb of *DMPK* and its associated repeat. All six additional genes could play a contributory role in the DM1 phenotype. The expansion overlaps not only the 3' end of *DMPK,* but also the 5' promoter region of the neighboring gene *SIX5*. ***SIX5*** has a strong resemblance to the fruit fly gene *sine oculis* (needed for eye development) and to a family of mouse genes that regulate distal limb muscle development. Because cataracts and distal muscle wasting are common in DM1, haploinsufficiency of SIX5 was suggested as a possible contributor to DM1 pathogenesis (Shaw et al. 1993; Boucher et al. 1995; Jansen et al. 1995). The most compelling support for the involvement of neighboring genes in DM pathogenesis came from the fact that *Six5* knockout mice developed cataracts (Klesert et al. 1997). ***DMWD*** gene, immediately upstream of *DMPK,* is expressed in the testis and therefore is proposed to play a role in male infertility (Junghans et al. 2001). ***SYMPLEKIN***, has been suggested to function in the assembly of the polyadenylation machinery (Takagaki and Manley 2000). Obvious implications therefore arise with respect to this gene and the multisystemic nature of DM1. The other three transcripts in this region may be involved in DM-associated testis problems and diabetes. ***RSHL1,*** which lies immediately upstream of *DMWD*, and ***20-D7***, located downstream of *SIX5* have been detected in adult testis (Eriksson et al. 2001), thus, altered expression of these genes may result in testicular atrophy in males. Finally ***GIPR*** is involved in insulin secretion and so could be responsible for the increased incidence of diabetes observed in DM1 patients (Alwazzan et al. 1998).

## 1.2.3 The RNA gain-of-function hypothesis

Several independent lines of evidence indicate that the predominant mechanism for DM pathogenesis is a gain-of-function of RNA transcribed from the expanded alleles. First, no point mutations or deletions within the *DM1* or *DM2* protein products cause DM, indicating that the repeats are determinative for these diseases rather than a loss of function associated with the *DM1* or *DM2* loci. Second, the fact that two different loci containing similar expanded repeats cause strikingly similar diseases strongly suggests that DM1 and DM2 share a common pathogenic mechanism that is independent of a loss of function for the affected locus. Third, RNAs containing long tracks of CUG or CCUG repeats are transcribed from the expanded *DMPK* and *ZNF9* alleles, and both repeat-containing RNAs accumulate in discrete nuclear foci (Taneja et al. 1995; Davis et al. 1997; Liquori et al. 2001) (see Figure 1.2).

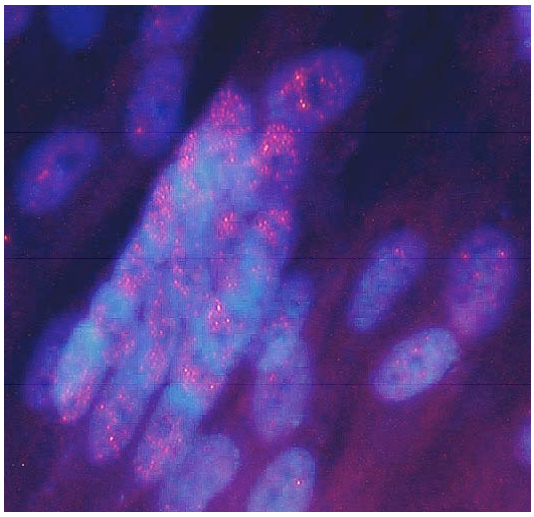


**Figure 1.2: Mutant DMPK mRNA forms foci in the nucleus of DM1 myoblasts.** Fluorescence *in situ* hybridization performed on DM1 myoblasts with 3000 CTG repeats using a Cy3-labeled $(CAG)_5$ PNA[9] probe and DAPI[10] staining (Langlois et al, unpublished data). Reproduced from www.theses.ulaval.ca/2003/21404/21404.pdf.

[9] Peptide nucleic acid (PNA) oligomers are DNA/RNA analogs in which the natural sugar–phosphate backbone is replaced by a synthetic peptide backbone. PNA were labeled with the fluorescence dye Cyanine 3 (Cy3).
[10] DAPI (41,6-Diamidino-2-Phenylindole Dihydrochloride) is a highly fluorescent cationic dye, which specifically binds to A-T-rich DNA. This fluorescent character permits the use of DAPI for fluorescent microscopy and analytical biochemistry.

Fourth, transgenic mice expressing 250 CUG repeats in the 3'-UTR of the human skeletal α-actin gene reproduced myotonia and the histopathological features observed in DM1 muscle (Mankodi et al. 2000), demonstrating that expression of CUG repeats independent of the *DM1* locus is sufficient to induce major features of the disease.

According to the RNA gain-of-function hypothesis, DMPK expanded-repeats exert a toxic gain of function through their inappropriate association with CUG-binding proteins, which are required for the maduration, stability and translation of specific mRNAs (Timchenko et al. 1996). The first RNA-binding protein found to be sequestered by the expansion repeats was CUG-binding protein (CUG-BP), a conserved heterogeneous ribonucleoprotein, which associates with pre-mRNAs during transcription and is believed to play a role in their processing (Timchenko et al. 1996). CUG-BP was first identified in band shift assays using a $(CUG)_8$ repeat probe (Timchenko et al. 1996). However, recent data questions the role of this protein in the pathology of DM1. As shown above (Figure 1.2), *DMPK* transcripts accumulate as nuclear foci in DM1 cell lines, whereas the localization pattern of CUG-BP in the same lines was unaltered. Furthermore, this protein has been shown to localize at the base of the RNA hairpin structure and not along the CUG repeat (Michalowski et al. 1999). On the other hand, muscleblind-like protein (MBNL), a second class of CUG binding protein, has a strong affinity for double-stranded CUG-repeat RNA *in vitro* and colocalizes with the nuclear foci containing CUG and CCUG repeat RNA in DM cells (Begemann et al. 1997; Michalowski et al. 1999; Miller et al. 2000; Fardaei et al. 2001). This protein is a homolog of a protein required for development of muscle and photoreceptor cells in *Drosophila* (Begemann et al. 1997) and contain two $Cys_3His$-type zinc finger motifs.

Although the physical evidence links MBNL and not CUG-BP with the nuclear foci of CUG-repeat RNA, functional analyses indicate that increased activity of CUG-BP is responsible for the aberrant regulation of cardiac troponin T (cTNT), insulin receptor (IR), and muscle-specific chloride channel (Clc-1) alternative splicing observed in DM1, contributing to cardiac conduction problems, insulin resistance and myotonia in DM (Philips et al. 1998; Savkur et al. 2001; Charlet et al. 2002). Further details on CUG-BP and its role in DM are given in section 1.7.

Interestingly, spinocerebellar ataxia SCA8 is caused by a CTG expansion in the 3' UTR of chromosome 13q21, a genetic flaw of the same nature than that found in DM1. The striking similarities between the *SCA8*, *DM1* and *DM2* mutations suggest that a toxic RNA mechanism at the cellular level may cause SCA8. Supporting this hypothesis, progressive retinal degeneration is seen in *Drosophila melanogaster* on expression of both expanded and normal human *SCA8* RNA (Mutsuddi et al. 2004). *SCA8* transcripts are almost exclusively expressed in the brain, which is consistent with the central nervous system involvement of that disease (Janzen 1999).

## 1.2.4 Additive Model of DM1 Pathogenesis

Subsequently, an additive model was proposed in which each of the above mechanisms contributes to DM1 pathogenesis, with some aspects of the disease caused by haploinsufficiency of DMPK, SIX5, and other neighboring genes and other clinical features resulting from effects of the CUG expansion in RNA (see Figure 1.3).

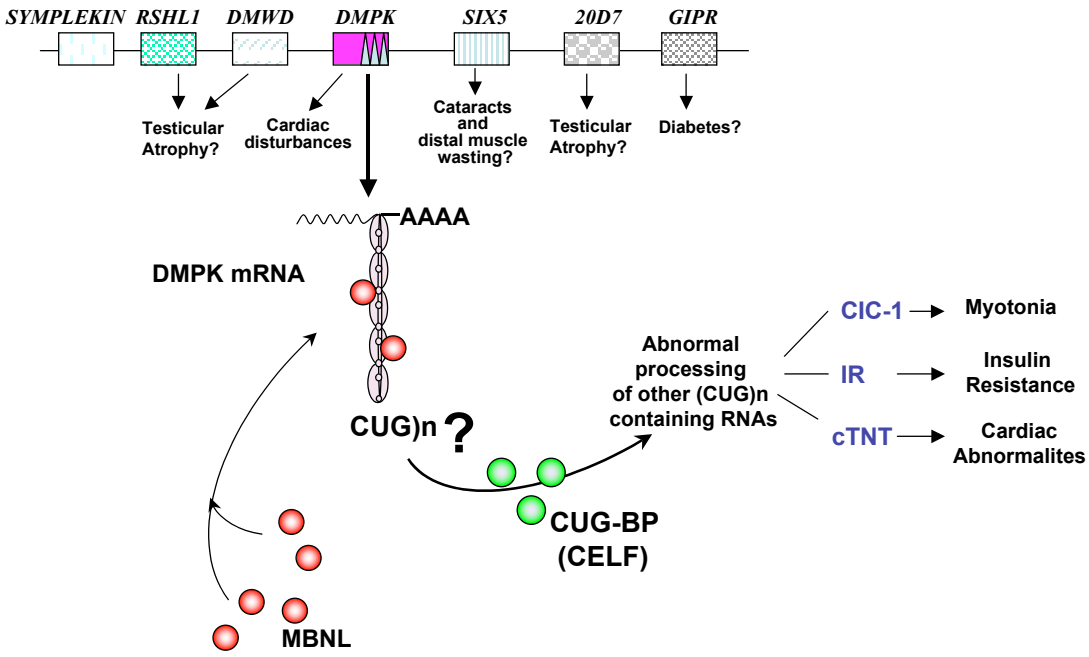



**Figure 1.3: Model of pathogenesis in DM1**. CUG expanded repeats affect the expression of multiple genes in the region by altering chromatin structure. Additionally, triplet repeats form imperfect hairpins that are bound by the double-stranded RNA-binding protein MBNL (red). Nuclear accumulation of the repeat-containing RNA elevates expression of CUG-BP (green) by an unknown mechanism. The consecuences of MBNL sequestration remain to be determined. Increased CUG-BP activity alters regulated splicing of three CUG-BP targets, ultimately resulting in the characteristic symptoms of DM.

## 1.3 Therapeutic strategies for DM

Currently, there is no cure for DM and treatment is aimed at controlling symptoms to maximize the quality of life of affected patients. In general, patients are given supportive care, such as occupational and physical therapy to maximize their ability to function in daily life. Stretching limbs to avoid tightened tendons and muscles is particularly important (when tightness of tendons develops surgery can be performed). It is very important to monitor the function of the respiratory muscles on regular basis, as some patients may require assistance with their breathing. Similarly, the heart may need to be checked by a cardiologist on a regular basis to detect early signs of cardiac dysfunction. Patients with cardiac and respiratory problems can be treated using a pacemaker and ventilator, respectively. Cataract surgery in the form of lens replacement may be done to improve the vision of affected patients.

More severe cases might demand treatment with corticosteroids to slow muscle degeneration and anticonvulsants such as quinine, phenytoin and others to control seizures and some muscle activity. Medication can also be prescribed for some muscular dystrophy (MD)-related heart problems.

In addition, people with MD are given age-appropriate dietary therapy to help them follow a healthy meal plan while avoiding obesity. Obesity is especially harmful to patients with muscular dystrophy because it places additional strain on their already weak muscles.

Two experimental procedures may hold some promise for treating MD. One of these is called myoblast transfer. In this procedure, millions of immature muscle cells are injected into a patient's damaged muscle. The goal is to provide the person's body with normal, healthy cells that may be able to function in the place of damaged ones. Thus far, there seems to be no evidence that this procedure is successful, although research of this topic continues. The second procedure is gene therapy. In this case, a person with MD is injected with artificially produced genes that are correct copies of the faulty genes in their body. The expectation is that the correct genes will restore the function of genes responsible for the disorder. However, gene therapy is a difficult task complicated with many side effects. Nevertheless, it may become the most likely way of curing DM.

# 1.4 Dystrophia Myotonica Protein Kinase (DMPK)

## 1.4.1 Characteristics of DMPK

DMPK was first identified and cloned as a protein kinase in the quest to establish the molecular basis of disease in DM, now more than a decade ago (Brook et al. 1992; Fu et al. 1992; Mahadevan et al. 1992). The *DMPK* gene is composed of 15 exons, which via three different alternative splicing events lead to six major DMPK isoforms A to F (Groenen et al. 2000), conserved between human and mice, and one less abundant isoform designated G, only present in humans (Tiscornia and Mahadevan 2000).

All isoforms share an N-terminal leucine-rich (**LR**) region, a Ser/Thr protein kinase domain (**K**), and a coiled-coil region (**CC**) (Leung et al. 1998; Madaule et al. 1998; Tan et al. 2001), but vary in the presence or absence of a five-amino acid **VSGGG** motif and in the nature of a C-terminal variable region (**VR**) (Groenen et al. 2000), see Figure 1.4.

The **LR** region was initially suggested to act as a signal peptide (Groenen et al. 2000). However, recent availability of the crystal structure from ROCKI (Jacobs et al. 2006), as well as our own results on DMPK, show that this region participates in protein dimerization. Likewise, the **CC** domain is also thought to mediate oligomerization (Cohen and Parry 1990). Recently, the **VSGGG** motif was suggested to modulate DMPK autophosphorylation activity and probably folding configuration based on indirect evidence inferred from the characteristic protein gel migration behaviour and *in vitro* kinase assays (Wansink et al. 2003). Four different types of C-terminal **VR** exist on DMPK isoforms: two long hydrophobic (tail 1) and relatively hydrophilic (tail 2) C-termini, a two-residue C-terminus (tail 3) and a fourth type of tail, of which the N-terminal half is identical to tail 1 (tail 4) (Tiscornia and Mahadevan 2000; Wansink et al. 2003). DMPK isoforms with long tails are predominant in the heart, skeletal muscle and brain, and DMPKs wih a 2-amino-acid tail in smooth muscle (Jansen et al. 1992; Groenen et al. 2000). Besides this, different alternative tails have been reported to anchor DMPK isoforms in distinct intracellular localization; long isoforms containing either tail 1 (DMPK A and B) or tail 2 (DMPK C and D) are membrane associated and anchor to the endoplasmatic reticulum and to the mitochondrial outer membrane, respectively, whereas, isoforms containing either

short tail 3 (DMPK E and F) or tail 4 (DMPK G) adopt a cytosolic localization (Wansink et al. 2003; van Herpen et al. 2005).

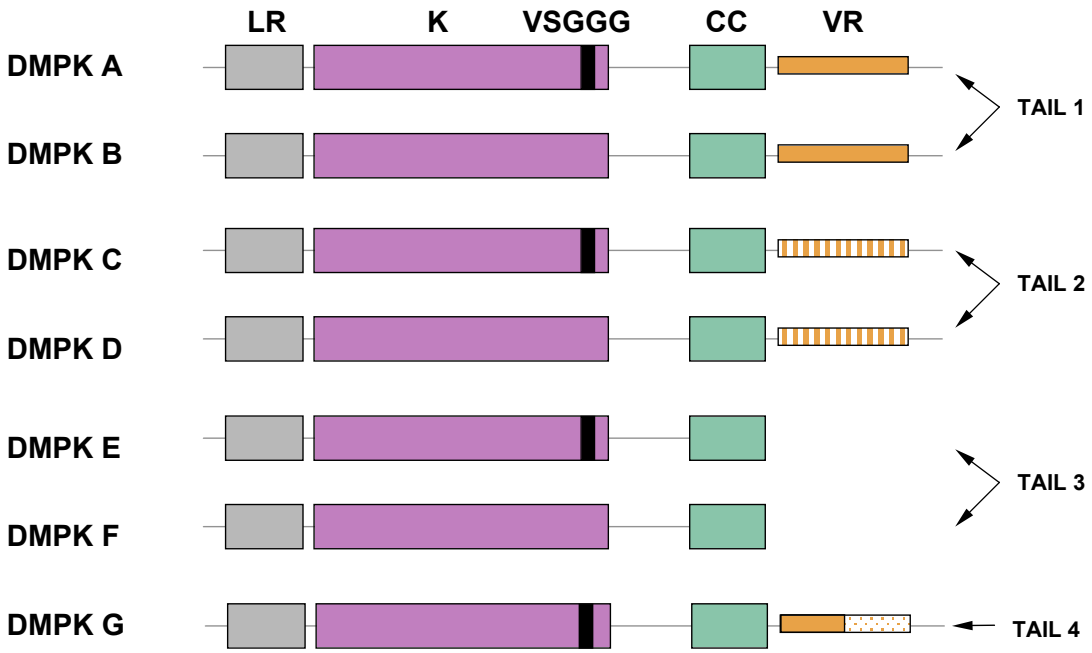



**Figure 1.4: Domain organization of DMPK isoforms**. Major DMPK isoforms A to F have an N-terminal leucine-rich (LR) region, a Ser/Thr kinase domain (K), and a coiled-coil region (CC). Differences between isoforms originate from alternative splicing, conserved between human and mice: (i) a VSGGG-sequence can be present (isoforms A, C and E) or absent (isoforms B, D and F) and (ii) three different C-terminal variable regions (VR) occur. Less abundant spliceoform DMPK G, only present in humans, carries a fourth type of VR, similar to tail 1. The figure was adapted from Wansink et al. 2003.

## 1.4.2 DMPK localization

DMPK protein is expressed in high levels in skeletal and cardiac muscle, but it is also present in brain cortex, cerebellum, eye lens, liver, testis, ovaries and intestine (Jansen et al. 1996; Balasubramanyam et al. 1998). DMPK was localized at neuromuscular junctions (van der Ven et al. 1993), myotendinous junctions (Whiting et al. 1995), and terminal cisternae of the sarcoplasmic reticulum in the skeletal muscle (Dunne et al. 1996; Salvatori et al. 1997; Shimokawa et al. 1997) and intercalated discs in the cardiac muscle (Ueda et al. 1998).

### 1.4.3 DMPK-related family of kinases (DMRFK)

DMPK was described as a member of the large AGC group of Ser/Thr kinases (Hanks and Hunter 1995), but it has been recently revealed as the canonical member of the "DMPK-related family of kinases" (DMRFK) (Zhao et al. 1997), a family of multidomain kinases which is expected to regulate novel signalling pathways in the tissues involved. Members of this family include kinases such as Rho-associated kinases I and II (ROCKI/II, also known as β and α, respectively) (Leung et al. 1995; Ishizaki et al. 1996), *Caenorhabditis elegans* LET-502 (a *C. elegans* homolog of ROCK) (Wissmann et al. 1997), *Neurospora* Cot1 (Yarden et al. 1992), *Drosophila* Warts (Justice 1995), *Drosophila* Genghis Khan (GEK) (a *Drosophila* homologue of MRCK) (Luo et al. 1997), murine Citron Rho-interacting kinase (Di Cunto et al. 1998) and Cdc42-binding kinases (MRCKs) (Leung et al. 1998; Ng et al. 2004) (see Figure 1.5). Members of this family show strong conservation in their kinase domain and a CC domain, expected to regulate activity by oligomerization.

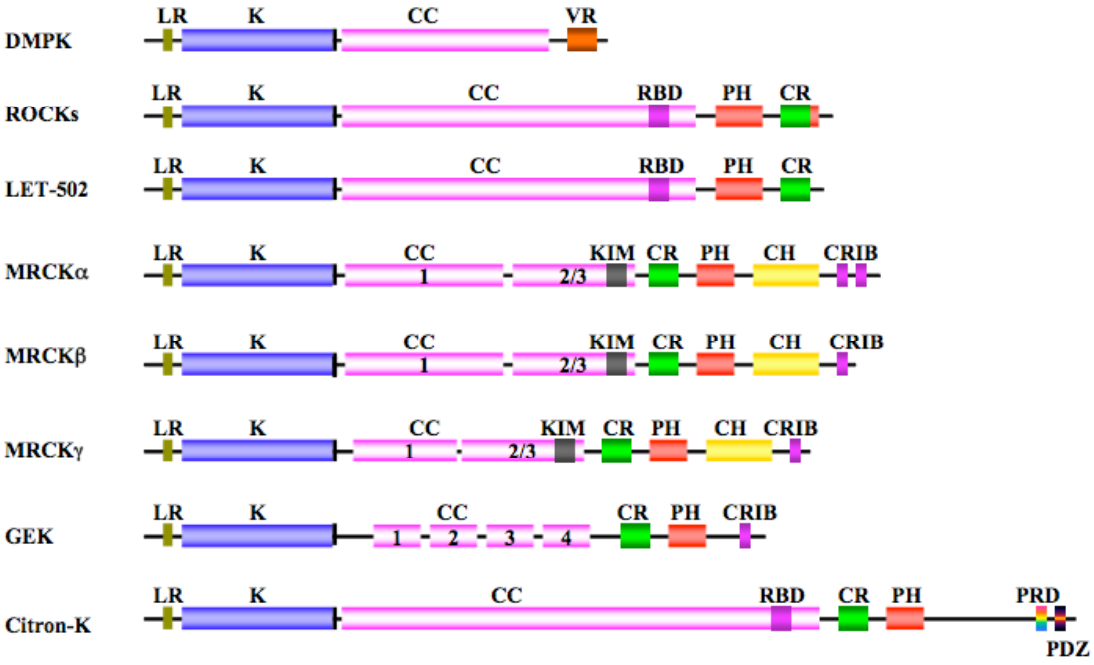



**Figure 1.5: Domain organization in DMPK-related family of kinases.** Domain composition of DMPK, ROCKs, LET-502, MRCKα, MRCKβ, MRCKγ, GEK and Citron kinase. LR: leucine-rich region; K: kinase; CC: coiled-coil; VR: variable region; RBD: Rho-binding domain; KIM: kinase inhibitory motif; PH: pleckstrin homology domain; CR: cysteine-rich domain, CH: citron homology domain; CRIB: Cdc42/Rac interactive binding site; PRD: proline-rich domain; PDZ: post-synaptic density-95/discs large/zonula occludens-1 domain.

## 1.4.4 Function of DMPK and related kinases

Although the cellular function of DMPK has not been established so far, recent findings indicate that it might participate in a variety of cellular processes including *i)* regulation of the electrical excitability of the plasma membrane, *ii)* myoblast differentiation, *iii)* cytoskeletal movement or intracellular transport dynamics and *iv)* tumors formation.

DMPK has been suggested to be involved in the **regulation of the electrical excitability of the plasma membrane**. Evidence reveals that lack of DMPK protein may be associated with typical DM1 symptoms like myopathy and heart conduction defects, perhaps via effects on sodium and calcium ion homeostasis (Jansen et al. 1996; Reddy et al. 1996; Benders et al. 1997; Mounsey et al. 2000b).

In addition to this, a role of DMPK in **myoblast differentiation** has also been suggested. Over-expression studies of DMPK demostrate that the kinase is expressed in various myoblast cell lines and slightly up-regulated during skeletal myogenesis *in vitro*. Importantly, a 4-10-fold DMPK over-expression markedly inhibits myogenesis, which seems to correlate with the marked delay in muscle terminal differentiation observed in severe cases of DM (Sabourin et al. 1997).

Besides that, DMPK has also been suggested to participate in **cytoskeletal movement or intracellular transport dynamics** (Jin et al. 2000; Schulz et al. 2003). Given the sequence homology between DMPK and other members of the DMFPK family, it is expected that certain functionality is shared. Most of these kinases function as effectors of small Rho GTPases (Rho, Rac and Cdc42 subclasses), which are known to regulate cytoskeletal dynamics. Accordingly, several members of the family are functionally related to cytoskeletal reorganization events, as illustrated by the involvement of Gek in actin polymerization (Luo et al. 1997) and the role of Citron kinase, LET-502 and ROCKII in regulating cleavage furrow contraction during cytokinesis (Madaule et al. 1998; Kosako et al. 2000; Piekny and Mains 2002).

An effect of DMPK in synaptic plasticity through an effect on actin cytoskeleton could be relevant to the cognitive dysfunction associated with DM (Schulz et al. 2003).

In addition, a Ser/Thr kinase from fission yeast orb6 related to DMPK (Verde et al. 1998) has been proven to maintain cell polarity during interphase and to induce actin reorganization during morphological transitions, thus, coordinating cell

morphogenesis with cell cycle. Orb6 shares functional similarities with *Drosophila* warts. Mutations in the latter gene cause an overgrowth phenotype and loss of warts results in cell-autonomous formation of epithelial tumors. This suggests that this is a tumor repressor gene and, therefore, is essential for the control of cellular morphogenesis as well as proliferation. This observation could be highly relevant to DM, considering that sometimes this disorder presents associated calcifying epitheliomas, neurofibromas and parathyroid adenomas.

## 1.4.5 Substrates of DMPK

Various potential substrates have been proposed for DMPK by *in vitro* studies including the β-subunit of the dihydropyridine receptor (DHPR) calcium channel (Timchenko et al. 1995), phospholemman (Mounsey et al. 2000a), phospholamban (Kaliman et al. 2005), the myosin phosphatase targeting subunit 1 (MYPT1) (Muranyi et al. 2001) and CUG-BP (Timchenko et al. 1996).

**DHPR** is a voltage-dependent calcium channel, which functions in skeletal muscle essentially as a voltage sensor, triggering intracellular calcium release for excitation-contraction coupling (Rios and Brum 1987). **Phospholamban** is a regulator of the calcium pump in cardiac muscle cells. **Phospholemman** is a membrane-bound protein involved in ion transport. Modification of muscle calcium and sodium ion channels homeostasis could lead to an alteration of muscle excitability, as seen in DM (Benders et al. 1997; Mounsey et al. 2000b). Nevertheless, an *in vivo* demostration of the phosphorylation of these substrates by DMPK remains to be established, and a link with clinical manifestations of DM is unclear.

**MYPT1** has also been identified as a substate for DMPK. DMPK controls the status of myosin regulatory light chain posphorylation, either directly or indirectly via regulation of myosin phosphatase activity, thereby affecting stress fiber formation, smooth muscle contraction or cytokinesis. However, the effects of DMPK-mediated phosphorylation on actomyosin dynamics have not yet been studied in detail (Muranyi et al. 2001).

**CUG-BP** has also been suggested as a DMPK substrate. It exists in a hyperphosphorylated cytoplasmic (CUG-BP1) and hypophosphorylated nuclear (CUG-BP2) forms and concentration of the latter isoform is increased in nuclei in DM

disease. Data from two biological systems with reduced levels of DMPK, homocygous DM patients and DMPK knockout mice, show that DMPK regulates both phosphorylation and intracellular localization of CUG-BP (Timchenko et al. 1996; Roberts et al. 1997). Besides this, DMPK mRNAs are accumulating in nuclei of DM patients (Taneja et al. 1995). Because CUG-BP was suggested to interact with DMPK mRNA (Timchenko et al. 1996) and hypophosphorylated form is accumulating in nuclei, it is speculated that nuclear retention of DMPK transcripts is, at least in part, mediated by CUG-BP protein. Therefore, a feedback mechanism has been proposed, whereby DMPK phosphorylates CUGBP, which, in turn, facilitates the transport or processing of the mRNA for DMPK and other CUG-BP dependent mRNAs. According to this autoregulatory loop, it is likely that the expansion of the triplet repeats in DM cells impairs the ability of DMPK to phosphorylate CUG-BP. Decreased phosphorylation of CUG-BP results in decrease of DMPK mRNA expression and possibly other mRNAs resulting in the multiple organ involvement that is observed in DM. (Timchenko et al. 1996; Roberts et al. 1997).

### 1.4.6 Regulation of DMPK activity

**Self-assembly** is a main feature of the DMPK-related family of kinases. Given that CCs are established oligomerization motifs, these were initially thought to be essential for molecular assembly in these kinases. CC domains from DMPK-related kinases were thought to form dimeric assemblies and, thereby, to mediate dimerization. To date, only two closely related CC motifs from this family have been characterized. These correspond to the C-terminal Rho-binding domains (RBDs) of ROCKII-CC and ROCKI-CC that share 52% sequence identity. For these, crystal structures have been reported in free form (Shimizu et al. 2003) and complexed to RhoA (Dvorsky et al. 2004), respectively, confirming that both coils form parallel dimers (Figure 1.6).

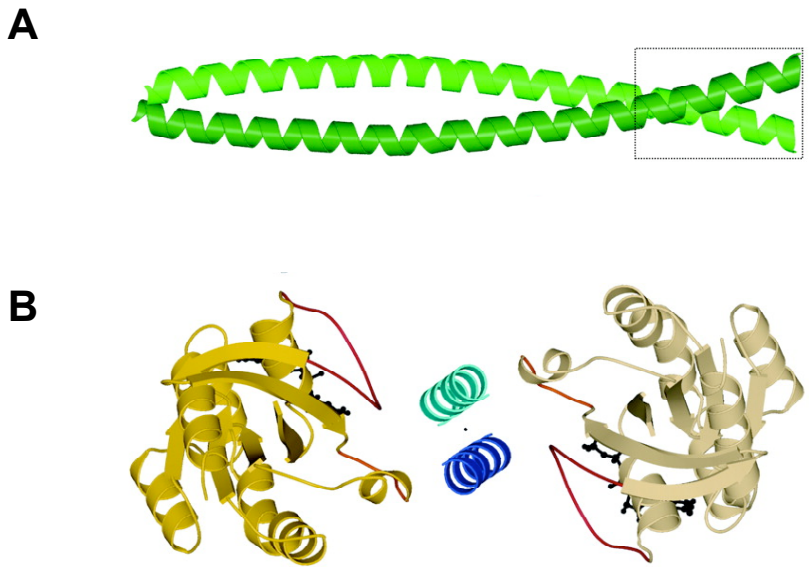



**Figure 1.6: Structures of the RBD from ROCKII in free form and ROCKI complexed with RhoA**. **(A)** Structure of ROCKII-RBD. A ribbon representation of the parallel CC structure is shown with the two helices in green. The *dashed-lined box* points at the Rho binding motif (Shimizu et al. 2003). **(B)** Top view of the complex between two RhoA molecules (gold and beige) and the Rho-interacting motif of the RBD molecules (blue and cyan) shown in ribbon representation. The bound GTPγS[11] molecule (black) is shown as a ball-and-stick model. Switch regions[12] I and II of both RhoA molecules are highlighted in red and orange, respectively (Dvorsky et al. 2004).

However, recent studies on ROCKs have shown that CC moieties are not required for their oligomerization and other domains are responsible for self-association in these kinases, namely the N-terminal LR region and a tail C-terminal to the kinase domain. (Doran et al. 2004; Jacobs et al. 2006). The structure from the dimeric N-terminal fraction (i.e., lacking CC segments) from ROCKI bound to the ATP-inhibitor Y-27632[13] (reported after the present work) shows how head-to-head dimers interact via two dimerization domains: a small, N-terminal domain comprising the α-helical LR region as well as a region, C-terminal to the kinase domain, which wraps around the catalytic core to reach the subunit interface (Jacobs et al. 2006), see Figure 1.7. The latter agrees with a biochemical study on MRCK, which also reported self-association in the absence of CC fractions (Tan et al. 2001).

[11] GTPγS (guanosine-5'-0-3-thiotriphosphate) is a non-hydrolyzable GTP analogue.
[12] Switch I and switch II regions, often referred to as the effector region, are involved in effector interactions.
[13] Y-27632 is a highly potent, cell permeable, selective ATP competitive inhibitor of ROCK1 and ROCK2.

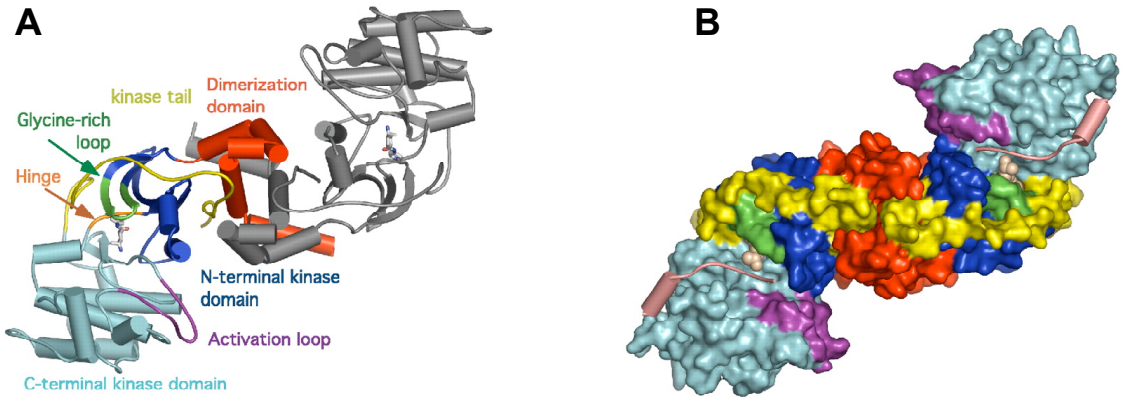



**Figure 1.7: Crystal structure of the ROCK/Y-27632 complex**. **(A)** The ROCK protein dimer is drawn with β-sheets as arrows and α-helices as cylinders. One monomer is drawn in gray, and the others are colored by protein region. The N-terminal dimerization domain is shown in red. The N-terminal kinase domain (dark blue) is shown with the glycine-rich loop drawn in green. The hinge connecting the two domains is colored orange. The C-terminal kinase domain is shown in light blue with the activation loop in purple and the kinase tail in yellow. Y-27632 is shown in the active site near the glycine-rich loop and the hinge. **(B)** Surface representation of the dimer, where both monomers are colored by region. Y-27632 is shown in the active site as spheres. The figure was reproduced from Jacobs et al. 2006.

Often, self-association is aimed to ensure efficient **autophosphorylation** in trans by increasing the local concentration of the catalytic domains. This phenomenon is now generally known to be an activation mechanism common to many protein kinases including nonreceptor and receptor tyrosine kinases as well as Ser/Thr kinases. Phosphorylation of key residues at two sites within the kinase domain: a conserved threonine residue in the **activation loop** and a Ser/Thr residue in a **hydrophobic phosphorylation motif** (consensus sequence FXX(F/Y)(S/T)(F/Y)), (Wansink et al. 2003) near the C-terminus, provide the means of activation of many protein kinases (Palaty et al. 1997; Chan et al. 1999; Millward et al. 1999).

Studies on MRCK and ROCK found that residues S204, T240 and T403 of MRCK (Tan et al. 2001) within the activation loop and the hydrophobic motif, respectively, as well as T398 within the hydrophobic motif of ROCK (Tan et al. 2001; Zhao and Manser 2005), are involved in modulation of the kinase activity. Although no consensus phosphorylation site has been proposed for DMPK, and autophosphorylation is considered to be mainly conformationally driven (Wansink et

al. 2003), however, it might be that DMPK undergoes a similar regulatory phosphorylation scenario. Nevertheless, the **VSGGG motif** present at the C-terminus of the kinase domain of DMPK, unique among DMPK-related kinases, has been suggested to modulate both autophosphorylation activity and probably folding in this kinase (Wansink et al. 2003).

In addition to its multimerization property, CC motifs of DMPK-related kinases have also been found to act as **intrasteric regulators**. In ROCKs and MRCKs, the CC regions interact with their respective kinase domains, functioning as intrasteric inhibitors of kinase activity (Amano et al. 1999; Tan et al. 2001). In ROCKs, the putative negative regulatory region involves both RBD at the C-terminal portion of its large CC motif and the pleckstrin homology (PH) domain. Constructs lacking the C-terminal portion of ROCK kinase appeared to be constitutively active, whilst those without the kinase region or various C-terminal fractions of Rho-kinase act as dominant negative forms. So far, the only model for regulation of the catalytic activity in ROCK that has been proposed, suggests that the RBD and the PH domain interact with the catalytic domain to keep the enzyme in the resting state, inactive (Amano et al. 1999). Accordingly, binding of the Rho GTPase to the RBD disrupts the inhibition by the RBD and the PH domain in response to extracellular signals and results in an open conformation that facilitates N-terminus mediated oligomerization, auto-phosphorylation and subsequent activation of the kinase. Furthermore, It has been reported that also some lipids, such as arachidonic acid, can activate ROCK *in vitro* (Feng et al. 1999), by binding to the negative autoregulatory region, most likely to the PH domain, and subsequent release from the catalytic site.

In contrast to ROCK, the regulation of MRCK does not appear to involve the PH domain, but a small region within the extensive MRCK CC region, termed the kinase inhibitory motif (KIM) and located at the end of the distal CC2/CC3 domains, which acts as a negative autoregulatory domain by interacting to and inhibiting the catalytic domain (Tan et al. 2001; Ng et al. 2004). Co-expression of MRCKγ KIM domain is sufficient to inhibit the activity of either MRCKα or β. A proposed model of activation suggests that upon phorbol 12-myristate 13-acetate (PMA) binding to the CR domain, the kinase inhibition by KIM would be disrupted, thus facilitating N-terminus mediated dimerization, auto-transphosphorylation and following activation.

Opposite to ROCK and MRCK, DMPK-CC has been proposed to act as an enhancer of its kinase activity (Zhang and Epstein 2003). In addition, the presence of an autoinhibitory domain at the extreme C-terminus of the full-length kinase (residues 617-629) has been suggested (Bush et al. 2000). Studies with synthetic peptide substrates and inhibitors suggested this region as a pseudosubstrate autoinhibitory domain (Bush et al. 2000). For an illustration of the models of activation proposed for the DMPK-related kinases ROCK and MRCK, see Figure 1.8.

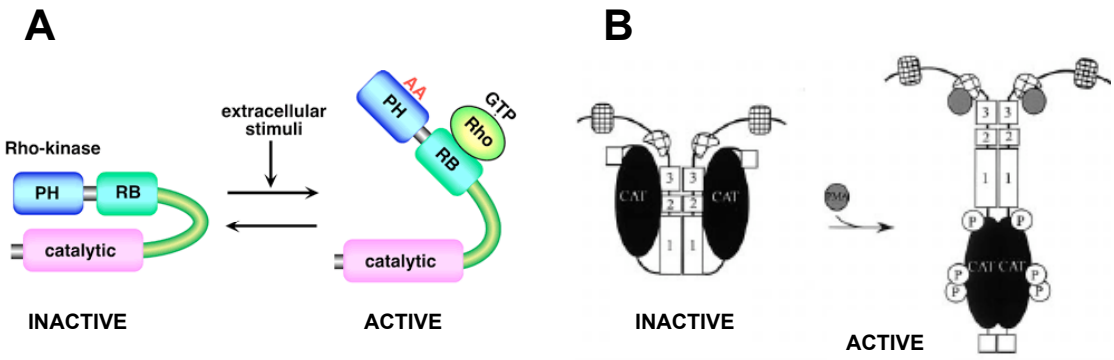



**Figure 1.8: Models for the regulation of DMPK-related kinases ROCK and MRCK.**
**(A) Model for the regulation of ROCK**. In ROCK, It is likely that the RB and PH domains interact with the catalytic domain to inactivate the enzyme in the resting state, and the active form of Rho interacts with the RB domain, alters the conformation of Rho-kinase and thereby cancels the inhibition by the RB and PH domains in response to extracellular signals. It has been reported that arachidonic acid cancels the inhibitory effect of RB/PH region, and activates Rho-kinase *in vitro*. **(C) Model for regulation of the catalytic activity of MRCK.** The intramolecular interaction between the kinase inhibitory motif (KIM) located at the end of the distal CC2/CC3 and the kinase domain keeps the kinase in a closed, inactive, dimeric structure. Disruption of this interaction (e.g., PMA[14] binding to the cysteine rich domain CRD or co-expression with a mutant kinase domain) resulted in an open structure that facilitates N terminus-mediated dimerization, autophosphorylation, and subsequent kinase activation. These figures were reproduced from Amano et al. 1999 and Tan et al. 2001.

[14] PMA: phorbol 12-myristate 13-acetate is the most commonly used phorbol ester.

In this family of kinases, a complex, multistep regulation process is expected, comprising both, phosphorylation by up-stream kinases and interactions with regulatory molecules. For DMPK, it has been reported that DMPK-binding protein (MKBP) enhances its catalytic activity (Suzuki et al. 1998). MKBP is a novel member of the small heat shock protein family that specifically associates with DMPK. It activates DMPK kinase activity *in vitro* and protects it from heat-induced inactivation. The expression of MKBP in skeletal muscles from DM1 patients is selectively upregulated, suggesting that a protein interaction between DMPK and MKBP might be involved in the pathogenesis of DM1 (Suzuki et al. 1998).

In addition, several DMPK-related kinases have been shown to interact with members of the Rho family of small GTPases, which includes Rho, Rac1 and Cdc42. To date, no endogenous activators of DMPK are known. However, a recent study based on pull-down assays using bacterially expressed Rho GTPases and DMPK (Shimizu et al. 2000) suggests that the actin cytoskeleton-linked **GTPase Rac1 binds to DMPK**, and co-expression of Rac1 and DMPK activates its transphosphorylation activity in a GTP-sensitive manner. DMPK seems to also bind **Raf-1 kinase,** the Ras-activated molecule of the MAP kinase pathway, which phosphorylates and activates DMPK (Shimizu et al. 2000).

# 1.5 Coiled-coils (CCs)

α-Helical CCs, first proposed by Crick in 1953, represent what is probably the most widespread assembly motif found in proteins. CCs are comprised of α-helices that wind around each other, usually forming a left-handed superhelix (Crick 1953; Chothia et al. 1981; Burkhard et al. 2001). The sequences of left-handed CCs are characterized by heptad repeats, $(\mathbf{a\ b\ c\ d\ e\ f\ g})_n$ , where positions **a** and **d** are characteristically occupied by hydrophobic residues like Leu, Ile, or Val, and constitute the structural core of the motif. (Hodges 1992; Adamson et al. 1993; Hodges 1996). Whereas positions **e** and **g** are often charged and form networks of intra- and inter- helical salt bridges crucial to fold stability and interchain recognition (O'Shea et al. 1991; Ellenberger et al. 1992; Glover and Harrison 1995) (Figure 1.9).

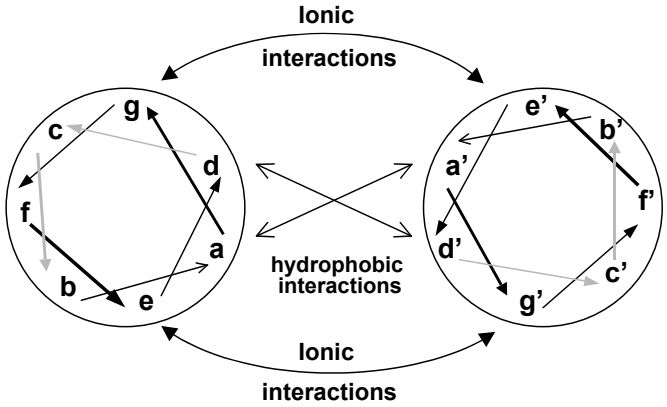



**Figure 1.9: Helical wheel diagram of a parallel dimeric CC.** The residues are labeled **a–g** in one helix and **a'–g'** in the other. The hydrophilic interactions (**g–e`**/**g`–e**) and the hydrophobic core (**a/a'** and **d/d'**) are shown.

The CC geometry can be readily parameterized (Crick 1953). Any helical structure is mathematically defined by three parameters, radius, pitch, and multiplicity. The CC radius refers to the radius of the superhelix, the CC pitch is the distance along the axis of the superhelix that corresponds to an exact 360° turn, and the multiplicity of a superhelix is defined by the number of α-helices forming it (Strelkov and Burkhard 2002).

Typically, CCs are very specific oligomers with high thermodynamic stability, which enables them to function as oligomerization domains in many proteins (Burkhard et al. 2001). However, CCs can also be implicated in refolding processes and molecular recognition events (Lupas and Gruber 2005).

CCs can form dimers, trimers, tetramers and pentamers depending on their amino acid sequence and the enviroment. Furthermore, CCs can display 11-residues and 15-residues repeats, alternative to the common 7-residues pattern, and may present discontinuities along the α-helices (Figure 1.10).

Although the majority of two-stranded CCs have parallel strands, a number of CCs with antiparallel strands have been observed. Such CCs may be intrachain, where the CC is formed by two helices joined by a turn, or they can be interchain interactions between separate polypeptide chains.

Few factors have been found, so far, to be crucial for determining the number of strands in a CC. For example, changing a two-stranded helix to a three- or four-stranded one, by changing the sets of buried hydrophobic residues at the core **a** and **d** positions (Harbury et al. 1993), this indicates that the shape of the hydrophobic side chain is an important determinant. Moreover, residues at positions **e** and **g** have also been found to play an important role in the stoichiometry of a CC. A study by Zeng and colleagues changed the oligomerization state of GCN4 by alteration of the g/e pairings (Zeng et al. 1997). Particularly, by varying specific g/e residues, they were able to increase the oligomerization state, presumably by widening the hydrophobic interface available.

To date, numerous structures are available for proteins containing the heptad repeat characteristic of trimeric CCs, including α-fibrous proteins (Conway and Parry 1991; Woolfson and Alber 1995) and other proteins like, for example, a soluble trimeric fragment of influenza hemagglutinin (Bullough et al. 1994) and a trimeric fragment of C-type mannose-binding proteins (Weis and Drickamer 1994). However, there is a relative lack of reports on the stability and folding of four-chain and five-chain CCs compared to the number of reports on dimeric CCs. Nevertheless, some tetrameric (DeGrado et al. 1989; Peters et al. 1995) and pentameric (Kammerer 1997) CCs have been characterized.

In addition, the heptad repeat can have discontinuities, known as *stutter*, *stammer* and *shift*, which correspond to a four- three- and one-residue insert, respectively, into a

sequence with an otherwise continuos heptad repeat (Brown et al. 1996).

In addition to the most common CCs with heptad repeats, structures with longer repeats are possible. In 1951, Pauling and co-workers mentioned the possibility of other periodicities, including 11-residue repeats (three α-helical turns, 11/3) and 15-residue repeats (15/4) (Pauling et al. 1951). Accordingly, 11-residue repeats and 15-residue repeats were found recently in the tetrabrachion protein from *Staphylothermus marinus* (Peters et al. 1996) and the tetramerization domain (TD) from human vasodilator-stimulated phosphoprotein (VASP) (Kuhnel et al. 2004), respectively, which revealed a right-handed supercoiling.

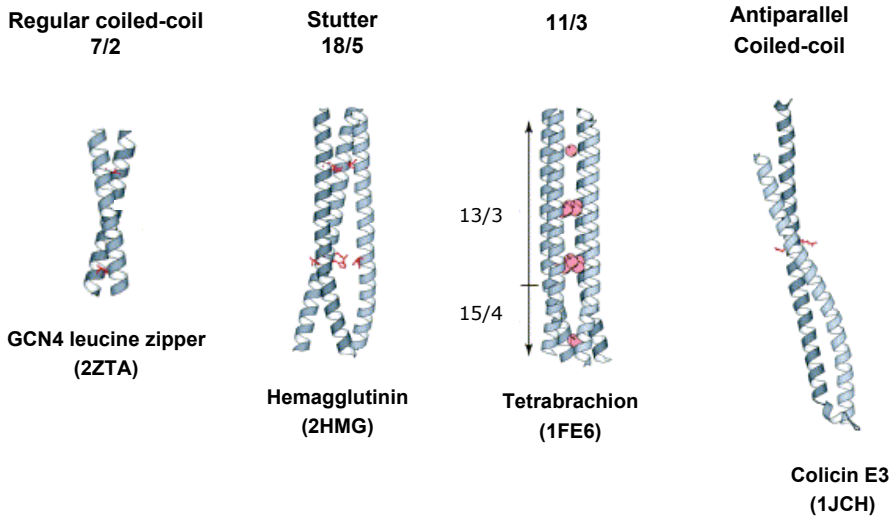



**Figure 1.10: Periodicities of CC proteins.** Gallery of parallel and antiparallel CC structures with heptad and non-heptad repeats. From the left to the right: General control protein GCN4 shows a regular dimeric CC, influenza hemagglutinin reveals a trimeric CC plus a stutter, tetrabrachion shows a parallel right-handed tetrameric with the periodically occuring water molecules (spheres) in the cavities of the hydrophobic core, and colicin E3 shows an antiparallel CC (adapted from Gruber and Lupas 2003).

# 1.6 Rac1

## 1.6.1 Introduction

The actin cytoskeleton mediates a variety of essential biological functions in all eukaryotic cells. In addition to providing a structural framework around which cell shape and polarity are defined, its dynamic properties provide the driving force for cells to move and to divide. Understanding the biochemical mechanisms that control the organization of actin is thus a major goal of contemporary cell biology, with implications for health and disease. Members of the Rho family of small GTPases have emerged as key regulators of the actin cytoskeleton. A recent study suggests that the small Rho GTPase Rac1 interacts with DMPK and co-expression of Rac and DMPK triggers its activity in a GTP-sensitive manner (Shimizu et al. 2000). Protein binding assays were performed by Shimizu and co-workers to study a possible interaction between DMPK and the three small Rho GTPases RhoA, Rac1 and Cdc42 using pull-downs and bacterially expressed GTPases and DMPK samples. DMPK was shown to bind exclusively to Rac1, but not to RhoA or Cdc42. These results indicated that Rac1 can physically interact with DMPK. Furthermore, co-transfection of DMPK with GTP-bound Rac1 activated DMPK phosphorylation of histone H1 by almost 3-fold, consistent with a potential regulatory interaction. This interaction of DMPK with the actin cytoskeleton-linked Rho GTPase Rac1 may play a critical role in the diverse pleiotropy of DM.

Rac GTPase has three mammalian isoforms, Rac1, Rac2 and Rac3. Rac1 is expressed in most cell tissues (Moll et al. 1991), Rac2 is specifically expressed in hematopietic cells (Gu et al. 2002) whereas Rac3 is localized mostly to the adult brain and expression in the developing nervous system (Haataja et al. 1997; Albertinazzi et al. 1998).

Rac1 tertiary structure consists of one central β-sheet made up of six strands, five parallel and one anti-parallel, six α-helices and two short $3_{10}$ helices (Figure 1.11).

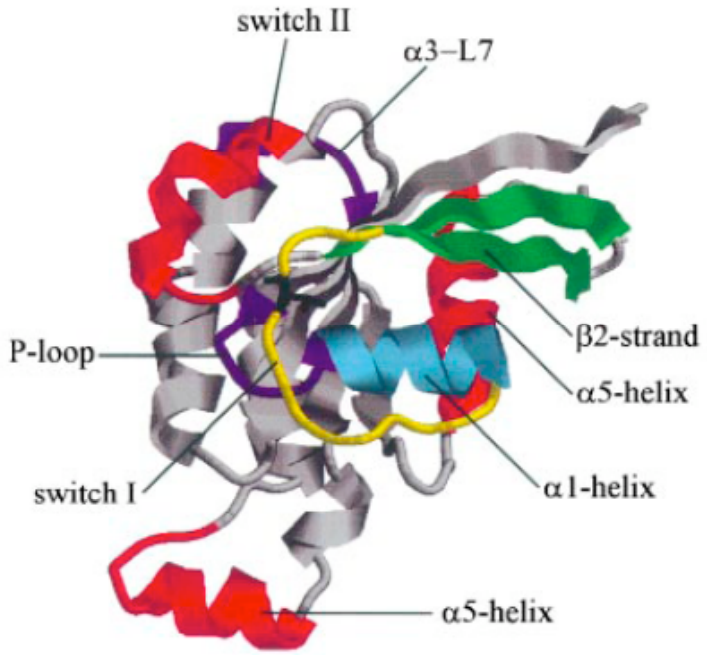



**Figure 1.11: Structure of the GTP-coordinated human Rac1 molecule.** Ribbon diagram of the GTP-Rac1 molecule. Secondary elements are labeled (reproduced from Muller et al. 1999).

## 1.6.2 Rho GTPases

### 1.6.2.1 Characteristics of the Rho GTPases

***The Rho GTPase cycle***

Rho GTPases are proteins with a typical molecular mass (MM) of 20-25 kDa that act as molecular switches utilizing a simple biochemical strategy to regulate complex cellular processes. They cycle between two conformational states: one bound to GTP (active state) and the other bound to GDP (inactive state), hydrolyzing GTP to GDP. In the “on” (GTP) state, GTPases recognize target proteins and generate a response until GTP hydrolysis returns the switch to the “off" state (Takai et al. 2001). For a graphic illustration of the GTPase cycle see Figure 1.12.

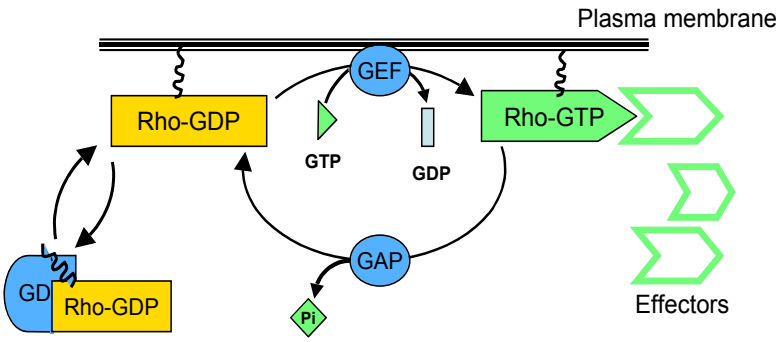



**Figure 1.12: The GDP/GTP cycle of Rho GTPases.** Small GTPases are in an active conformation when they have GTP bound, and are inactivated when the GTP is hydrolyzed to GDP. When the GTPase is in the active conformation it is able to interact with effector proteins, which transmit the signal downstream of the GTPase. The rate of the intrinsic hydrolysis can be enhanced by the interaction with GTPase activating proteins (GAPs), which promote the inactivation of the GTPase. Guanine nucleotide exchange factors (GEFs) accelerate the exchange of GDP for GTP, thereby activating the GTPase Rho guanine nucleotide dissociation inhibitors (Rho GDIs), which keeps the Rho GTPase in the GDP-bound, inactive state (Etienne-Manneville and Hall 2002).

### *Key functional domains of Rho GTPases*

Rho GTPases typically consists of the ***GTPase-domain*** and additional short sequences in the C-terminus and N-terminus (Takai et al. 2001; Wennerberg and Der 2004). The ***Rho insert motif*** is an insertion of about 13 amino acid residues in the GTPase-domain of Rho GTPases, not present in GTPase-domains of other small GTPases (Vetter and Wittinghofer 2001) The differences in the conformation of GDP- and GTP-bound small GTPases are basically manifested in two regions within the GTPase-domain, the so-called ***Switch I*** and ***Switch II regions,*** often referred to as the ***effector region***. In the GTP-bound Rho GTPase, the conformation of the switches are stabilized. When GTP is hydrolyzed to GDP, the conformation relaxes (Vetter and Wittinghofer 2001; Hakoshima et al. 2003), see Figure 1.13. Amino acid residues in the switch regions mediate the interaction between the GTPase and its effector proteins, which bind only to the GTP-bound conformation (Vetter and Wittinghofer 2001). The presence of $Mg^{2+}$ ions in the catalytic cleft of the GTPase-domain is crucial for both guanine nucleotide binding and catalytic function (Vetter and Wittinghofer 2001; Hakoshima et al. 2003).

Most of the small GTPases require lipid modifications to be fully active, and this occurs at the conserved ***CAAX-motif*** in the C-terminus of the protein. The most common lipid modification of Rho GTPases is isoprenylation of the cysteine residue by addition of farnesyl or geranylgeranyl isoprenoid, but palmitoylation also occur on some Rho GTPases. The lipid modifications promote the localization of Rho GTPases to membranes, which is crucial for their function (Wennerberg and Der 2004; Winter-Vann and Casey 2005).

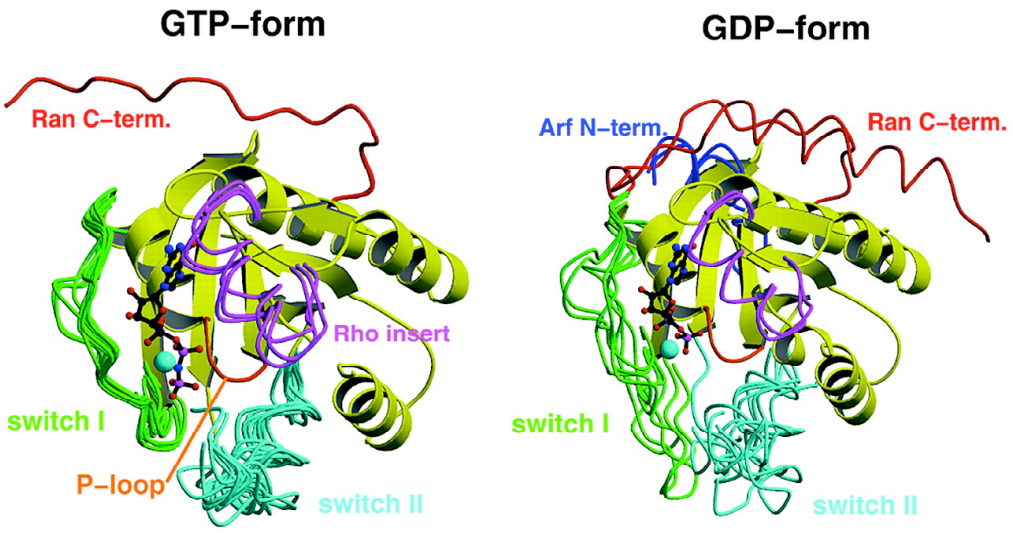
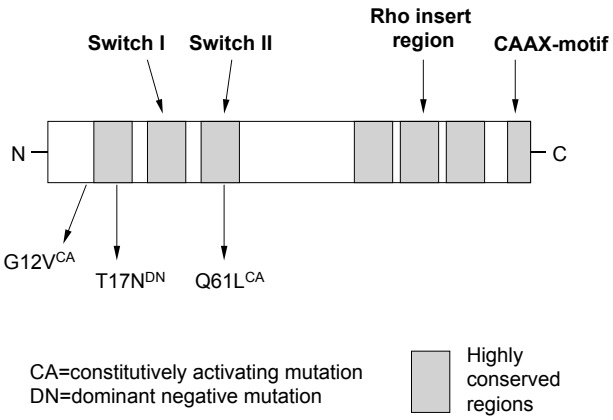



**Figure 1.13**: **Superposition of selected Ras-related proteins in the GDP and the GTP bound form.** Whereas in the GTP bound form, all switches are stabilized in a similar conformation, they are much more flexible in the GDP bound form and show divergent conformations. Extra elements in the structures of Ras-related proteins are highlighted (reproduced from Vetter and Wittinghofer, 2001).

***Rho GTPase mutants as research tools***

Over-expression of mutant Rho GTPases has been an extensively used tool to investigate their role in signaling pathways and regulation of cellular processes. These mutations were first identified and evaluated in Ras, and then, based on sequence homology the corresponding mutants of Rho GTPases were produced, see Figure 1.14 (Feig 1999).




***Key residues:***

**Gly 12**: Naturally occurring mutation to valine (12V) destroys the intrinsic GTPase activity and results in a constitutively active protein. This mutation has been found in certain tumors.

**Gln 61**: Mutation to leucine (61L) creates an activate protein similar to 12V.

**Thr 17 to Asn (T17N)**: Mutation that causes the Ras molecule to bind irreversibly to GEFs and remain inactive. This mutation is dominant negative, presumably because the Ras proteins bind and sequester the GEFs, preventing them from activating native molecules.

**Figure 1.14**: **Schematic diagram of the structure of a typical Rho GTPase.** The key funcional domains of a typical Rho GTPase; the insert region, the switch I and II regions and the CAAX-motif, as well as mutations used in studies, are depicted in the figure.

### 1.6.2.2 Rho GTPase Subfamilies

The Rho GTPase family is defined by the presence of the Rho-like GTPase domain, and 22 human genes have been found to encode Rho GTPases. A phylogenetic tree where the Rho GTPases are organized according to sequence homology in their GTPase domain is shown in Figure 1.15.

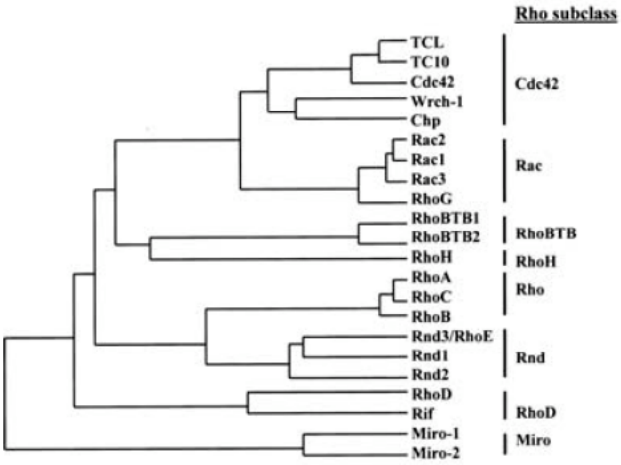



**Figure 1.15: Phylogenetic tree representation of Rho GTPases**. The mammalian Rho GTPases can be subdivided, according to sequence homology in their GTPase domain, into eight different classes consisting of the following members: **Rho** (RhoA, RhoB, RhoC), **Rac** (Rac1, Rac2, Rac3, RhoG), **Cdc42** (Cdc42, TC10, TCL, Wrch-1, Chp), **RhoBTB** (RhoBTB1, RhoBTB2), **RhoH, Rnd** (RhoE/Rnd3, Rnd1, Rnd2), **RhoD** (RhoD, Rif) and **Miro** (Miro-1, Miro-2). These classes can be classified into 2 further groups: "typical" or "atypical" categories. The typical Rho GTPases include most of the members, except Chp, Wrch-1, RhoBTB1-2 and Miro-1 and 2 small GTPases, which constitute the atypical category of small Rho GTPases (Aspenstrom et al. 2004).

The amino acid sequences of Rho GTPases show approximately 30% similarity to Ras and within the Rho GTPase family the homology is 40-95 % (Takai et al. 2001; Furstoss et al. 2002; Burridge and Wennerberg 2004; Wennerberg and Der 2004). RhoA, Rac1 and Cdc42 are the best studied Rho GTPases.

#### 1.6.2.3 Functional features of Rho GTPases

Rho GTPases are key regulators of actin cytoskeleton reorganization. Most of the functional information on Rho-family proteins has come from studies of RhoA, Rac1 and Cdc42. Rho promotes stress fibres formation and focal adhesion assembly; Rac1 promotes actin-rich surface protrusions (lamellipodia) formation and membrane ruffling; and Cdc42 was revealed to induce the formation of actin microspikes and actin-rich finger-like membrane extensions (filopodia). Consequently, Rho GTPases have been implicated in the regulation of cell polarity, cell movement, cell shape, and cell-cell and cell-matrix interactions, as well as in regulation of endocytosis and exocytosis (Ridley 2001). Beside this, interest in Rho GTPases has increased since they have also been found to participate in the regulation of many other signal transduction pathways, in addition to those linked to the cytoskeleton. They are involved in the regulation of gene transcription, cell cycle progression, microtubule dynamics, vesicular transport pathways and a variety of enzymatic activities ranging from a NADPH oxidase in phagocytes to a glucan synthase in yeast (Etienne-Manneville and Hall 2002).

The multiple effectors identified for RhoA, Rac1, Cdc42 and other family members reflect the complex and diverse functional properties of these proteins.

#### 1.6.2.4 Rho GTPases effector proteins

Given the involvement of Rho GTPase in such wide variety of important cellular processes, a great deal of effort has been put into identifying their cellular targets o effector proteins. To date about 30 potential effectors for Rho, Rac and Cdc42 have been identified, primarily using affinity chromatography pull-down assays and the yeast two-hybrid system.

**GTPase binding motifs**

Sequence elements involved in the interaction between Rho family GTPases and their effectors are of considerable interest since their definition should permit the identification of further potential effector proteins. The first such motif to be identified was the **CRIB** (Cdc42/Rac interactive binding) motif, which specifies interaction with Rac1 and Cdc42, but not Rho (Burbelo et al. 1995). A second motif, **REM** (Rho effector motif), is found in the N-terminal Rho-binding domains of the Rho effectors (Watanabe et al. 1996) (Figure 1.16).

***CRIB motif***

The CRIB motif is necessary for the binding to Cdc42 and Rac GTPases, however, it is not sufficient to give high-affinity binding (Rudolph et al. 1998; Thompson et al. 1998). A less conserved inhibitory switch (IS) domain responsible for maintaining the proteins in a basal (autoinhibited) state is located C-terminally of the CRIB-motif (Hoffman and Cerione 2000; Kim et al. 2000; Lei et al. 2000). CRIB domains can adopt related but distinct folds depending on context. Although CRIB domains are largely unstructured in the free state, the IS domain forms an N-terminal β-hairpin that immediately follows the conserved CRIB motif and a central bundle of three α-helices in the autoinhibited state. The interaction between CRIB domains and their respective GTPase proteins leads to the formation of a high-affinity complex in which unstructured regions of both the effector and the GTPase become rigid. Recently reported NMR structures of Cdc42 bound to **a**ctivated **C**dc42-associated tyrosine **k**inase (ACK) (Mott et al. 1999) and **W**iskott-**A**ldrich **S**yndrome **P**roteins (WASP) (Rudolph et al. 1998) show that the CRIB motifs have no discernable structure in solution, but when bound to Cdc42 they form tight intermolecular β-sheet across the GTPase switch regions. Binding of Cdc42 or Rac to the CRIB domain causes a dramatic conformational change, refolding part of the IS domain and unfolding the rest (Rudolph et al. 1998; Mott et al. 1999; Hoffman and Cerione 2000; Kim et al. 2000; Lei et al. 2000; Morreale et al. 2000).

***REM motif***

The REM repeat, which is also called Rho effector or HR1 domain, was first described as a three times repeated homology region of the N-terminal non-catalytic part of protein kinase PRK1 (PKN) (Palmer et al. 1995). The first two of these repeats were later shown to bind the small G protein Rho (Shibata et al. 1996; Flynn et al. 1998) known to activate PKN in its GTP-bound form. Similar Rho-binding domains also occur in a number of other protein kinases and in the Rho-binding proteins rhophilin and rhotekin. Recently, the structure of the N-terminal REM repeat of PKN complexed with RhoA has been determined by X-ray crystallography (Maesaki et al. 1999). It forms an antiparallel CC fold termed an ACC finger. No sequence motifs common to other RhoA effectors have been identified, although the defined Rho-binding regions of Citron, Kinectin, and the ROCKs are all found in regions of extended α-helical CC structure (Alberts et al. 1998).

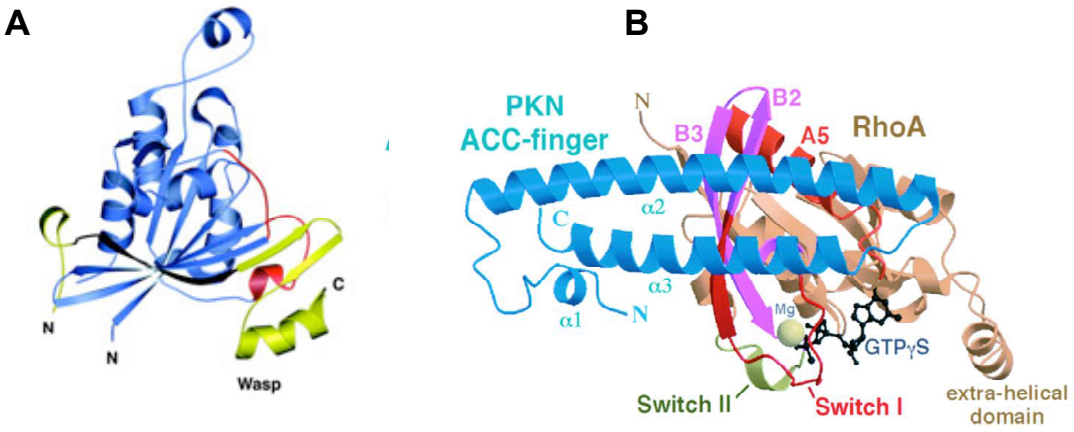



**Figure 1.16: Crystal structures of CRIB and REM motifs. (A)** Crystal structure of the complex between activated Cdc42 and an extended N-terminal GTPase-binding domain (GBD) peptide that includes the CRIB motif from WASP. Ribbon diagram with Cdc42 colored blue and WASP yellow. The Switch I and II regions of Cdc42 are colored in red and the CRIB motif of WASP is colored in black (Abdul-Manan et al. 1999). **(B)** Crystal structure of the complex between RhoA and the effector domain of PKN. Interaction RhoA-PKN occurs via the second REM motif from the PKN effector domain. Ribbon representation of the PKN effector domain (blue) bound to RhoA (brown). The bound GTPγS molecule (black) and the magnesium ion (yellow) are shown in ball-and-stick models (Maesaki et al. 1999).

## 1.6.3 Activation of effectors by Rho GTPases

When considering how a GTPase might regulate the activity of a target protein, and how these proteins subsequently transduce signals to the cell, the whole effector protein must be considered. The most common mechanism of effector activation by Rho GTPases appears to be the disruption of intramolecular autoinhibitory interactions, to expose functional domains within the effector protein (see Figure 1.17).

According to this model, inhibitory domains have been described for several Rho GTPase effector proteins. Among Rac/Cdc42 targets, PAK1-3 (Rac effectors) (Bagrodia and Cerione 1999; Tu and Wigler 1999) and the DMPK-related MRCK kinase (Cdc42 effector) (Tan et al. 2001; Ng et al. 2004) contain an intramolecular regulatory domain that inhibits kinase activity (see section 1.4.6).

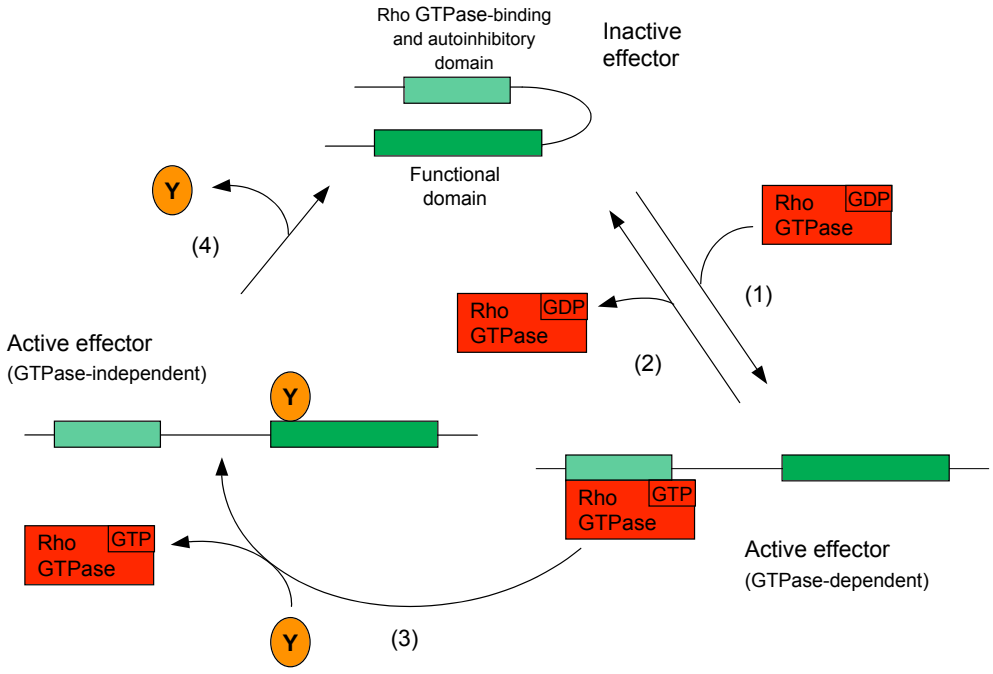



**Figure 1.17: General model of activation of Rho GTPase effectors. (1)** Binding of Rho GTPase to effector relieves an autoinhibitory intramolecular interaction (this is clearly the case for the kinase PAK, WASP, and maybe also for other effectors that contain autoinhibitory domains; see the text). **(2)** The effector remains active until GTP hydrolysis takes place. **(3)** Alternatively, a modification of the effector (Y, orange ellipses) (e.g. autophosphorylation, phosphorylation by a separate kinase or binding to a different activating protein) may maintain activity even after dissociation of the GTPase. **(4)** Inactivation of the effector occurs through removal of modification Y (e.g. dephosphorylation or removal of a bound activating protein), allowing the effector to re-enter its inactive conformation (Bishop and Hall 2000).

WASP, N-WASP and protein IQGAP[14], three related Cdc42 targets, also appear to be regulated by an intramolecular interaction (Kim et al. 2000). Rho effectors include ROCK and PKN kinases. In ROCKα, a region of the autoinhibitory domain (amino acids 941-1388), located at the C-terminal part of the CC domain, comprises the GTPase-binding and PH domains, and when mutated so as to disrupt binding to Rho, will inhibit endogenous ROCK (Amano et al. 1999) (see section 1.4.6).

Activation of DMPK-related kinases via small Rho GTPases involves different binding sequences. In ROCK and Citron kinases a RBD overlaps the extensive CC region. The recent crystal structure of a region of ROCKI (residues 947-1015) indicates the homodimer binds GTP-RhoA as a CC, but employs only 10-13 residues present at the extreme C-terminal end of the extended parallel CC region to bind a predominantly hydrophobic patch assembled by the switch regions of RhoA (Dvorsky et al. 2004). The specificity of ROCK interaction with RhoA (versus Rac1 and Cdc42) appears to involve residues F39, Y66, and L69 of RhoA.

On the other hand, MRCK possess a CRIB motif. Uniquely among CRIB containing effectors MRCKα can exist as a product with two PDBs, while the other two isoforms, MRCKβ and MRCKγ, are restricted to a single such CRIB motif (Tan et al. 2003). Interstingly CRIB1 binds preferentially to GTP-Cdc42, but addition of CRIB2 increases MRCK interaction with GTP-Rac.

[14] IQGAPs are scaffolding proteins that have been shown to be components of the actin cytoskeleton.

# 1.7 CUG-BP protein

## 1.7.1 Introduction

CUG-BP is a human RNA-binding protein that was first characterized as a novel human heterogeneous nuclear ribonucleoprotein (hnRNP). HnRNPs bind to nascent RNA polymerase II transcripts, and have been suggested to play important roles in nuclear pre-mRNA splicing and polyadenylation as well as the nucleocytoplasmic export and cytoplasmic stability of mRNAs. CUG-BP was purified based on its sequence-specific binding to RNA containing CUG repeats (Timchenko et al. 1996) and has been proposed to play a role in the pathogenesis of the trinucleotide expansion disease DM1, in which nuclear accumulation of DMPK transcripts containing expanded triplet repeats may sequester CUG-BP, thus creating an RNA gain-of-function mutation (see section 1.2.3).

CUG-BP is the defining member of the CELF (CUG-BP and ETR-3-like factor) family of RNA-binding proteins, which also includes an embryonically lethal abnormal vision (ELAV)-type RNA binding protein 3 (ETR-3), CELF3, CELF4, CELF5 and CELF6 proteins (Lu et al. 1999; Ladd et al. 2001; Ladd et al. 2004). CELF proteins have a similar domains arrangement and act as regulators of pre-mRNA alternative splicing.

## 1.7.2 Domain structure

All six proteins have the same domain structure including two adjacent RNA recognition motifs (RRM1 and RRM2) near the N-terminus and one RRM (RRM3) near the C-terminus, which are separated by a 160-230 residue "divergent domain" of unknown function (see Figure 1.18A). The CELF RRMs show a high degree of sequence identity among family members, however, there is little sequence identity in the divergent domain. The six CELF proteins can be divided into two groups based on sequence identity (see Figure 1.18B). One group contains CUG-BP and ETR-3, which are 78% identical. The second group contains CELF3, CELF4, CELF5 and CELF6, which exhibit at most 44% identity to CUG-BP and 62-66% identity to each other (Ladd et al. 2001).

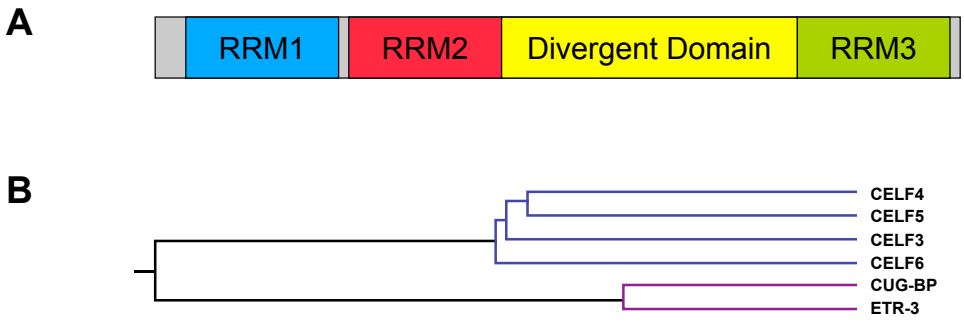



**Figure 1.18: CELF family members are closely related**. **(A)** All of the CELF proteins possess the same domain structure: three RRMs and a divergent domain of unknown function between RRM2 and RRM3. **(B)** Evolutionary relationship of human CELF proteins. Nomenclature is designed to consider CUG-BP and ETR-3 as CELF1 and CELF2, respectively. Figure adapted from Ladd et al. 2001.

The RNA recognition motif, RRM, also known as ribonucleoprotein domain (RNP) is one of the most abundant and best-characterized protein domains in eukaryotes and the most common nucleic acid binding protein motif (Birney et al. 1993; Burd and Dreyfuss 1994). The structure of RRM mainly has two identical binding domains, RNP1 and RNP2 that bind on opposite sides of a RNA loop. Each RRM has a $\alpha/\beta$-fold, which consists of a compact globule made of a four-stranded antiparallel β-sheet that forms the RNA-binding surface backed by two $\alpha$-helices. Thus, specific interactions are found in the space between the two RRM domains that serve as a region of recognition. Though many of the protein–RNA interactions are through amino acids located on the β-sheet of the RRM core structure, the specificity for the RNA ligands usually requires residues outside the conserved structural elements. For example, loops connecting the β-strands or regions outside the conserved RRM domain can play prominent roles in RNA recognition (Scherly et al. 1990; Ding et al. 1999). In addition, many RNA-binding proteins have multiple RRMs. To date, six structures of proteins with two tandem RRMs have been determined (Scherly et al. 1990; Shamoo et al. 1997; Xu et al. 1997; Crowder et al. 1999; Deo et al. 1999; Allain et al. 2000; Conte et al. 2000; Wang and Tanaka Hall 2001). For a graphic illustration of an RRM motif, see the crystal structure of the complex between the two RRMs of

the Sex-lethal (Sxl) protein and a 12-nucleotide single-stranded RNA derived from the *tra* polypyrimidine tract (Handa et al. 1999) (Figure 1.19).

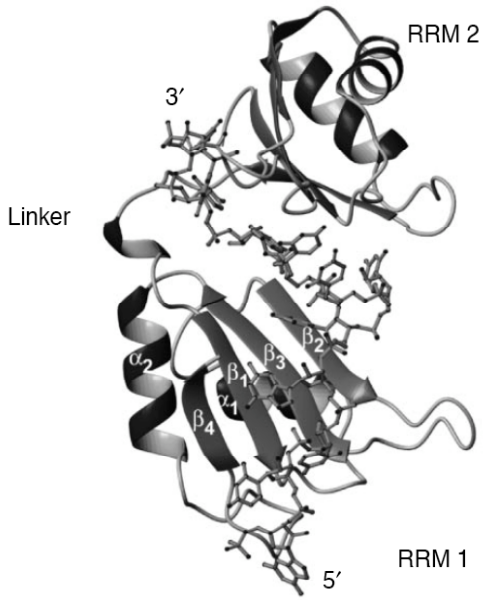



**Figure 1.19: Crystal structure of the complex between the two RRMs of the Sex-lethal (Sxl) protein and the *tra* mRNA precursor.** Overall view of the complex between *Drosophila* Sxl protein, a developmental regulator of alternative splicing, bound to a 12-nucleotide single-stranded RNA derived from the *tra* polypyrimidine tract (Perez-Canadillas and Varani 2001).

## 1.7.3 Function

CELF proteins are well-characterized alternative splicing regulators of a subset of pre-mRNAs, including human and chicken cardiac troponin T (cTNT), human insulin receptor (IR) and human muscle-specific chloride channel (Clc-1) (Cooper 1998; Savkur et al. 2001; Charlet et al. 2002).

Apart from the difference in their sequences, the two CELF protein groups also differ in their ability for RNA-binding and splicing function (Ladd et al. 2004). To define regions required for RNA-binding and splicing regulation, a functional analysis was performed on two CELF protein family members, CELF4 and ETR-3, which represent the two different CELF family subgroups.

For CELF4, it had been demonstrated that either RRM1 or RRM2 are necessary and sufficient for RNA binding and RRM2 plus 66 amino acids within the divergent

domain were sufficient to function as a splicing activator for cTNT exon 5 (Singh et al. 2004).

Whereas deletion analysis within ETR-3 suggest that the residues required for activation by ETR-3 involve non-overlapping N- and C-terminal regions containing either RRM1 and RRM2 or RRM3 plus segments of the adjacent divergent domain (Han and Cooper 2005).

## 1.7.4 Involvement of CUG-BP protein in DM1

Five pre-mRNAs have been identified to be aberrantly spliced in DM1 tissues and mouse models: tau (Sergeant et al. 2001), myotubularin-related protein 1 (MTMR1) (Buj-Bello et al. 2002), IR (Savkur et al. 2001), cTNT (Philips et al. 1998) and Clc-1 (Charlet et al. 2002; Mankodi et al. 2002). While the role of MTMR1 is still elusive in DM1, disruption of appropriate splicing of cTNT, IR and CLc-1 mRNAs can be associated with symptoms of the disease. It has not yet been demonstrated that the splicing defect of tau is associated with particular symptoms of DM1. However, tau is a microtubule-associated protein implicated in axon and neurite development in the brain (Shahani and Brandt 2002). Alterations in the balance of tau isoforms may well be responsible for the abnormal personality traits of myotonic dystrophy patients and could explain the severe involvement of the CNS in congenital cases.

Human cTNT exists as two splicing isoforms with either inclusion or exclusion of exon 5. The isoform that includes exon 5 is normally only expressed during embryonic development in the heart and skeletal muscle, while in adults exon 5 is excluded (Cooper and Ordahl 1984). Homozygous and heterozygous DM1 patients show alternate splicing patterns with a significant increase in cTNT mRNAs with exon 5 inclusion (Philips et al. 1998).

The insulin receptor is composed of two extracellular $\alpha$ and two intracellular $\beta$ subunits. Alternative splicing of exon 11 generates two isoforms of the $\alpha$ subunit: isoform A (IRA) lacking exon 11 and isoform B (IRB) that includes exon 11 (Figure 1.20A) (Seino and Bell 1989). Expression of the two isoforms is tissue specific depending on glucose requirement of each particular tissue since both isoforms do not have the same signalling capabilities (Moller et al. 1989; Kellerer et al. 1992). The B isoform, that is expressed dominantly in skeletal muscle, adipose tissue and the liver,

has been reported to signal more efficiently in response to insulin binding despite having a 2-fold lower affinity for insulin (Moller et al. 1989; McClain 1991; Mosthaf et al. 1991; Yamaguchi et al. 1993). In myotonic dystrophy, the IR is aberrantly spliced thus promoting the IRA isoform in skeletal muscle (Savkur et al. 2001).

Myotonia is one of the major cardinal features of DM. The inability to relax muscles after contractions is due to hyperpolarization of the muscle fibers after ion entry (Lehmann-Horn and Jurkat-Rott 1999). Several channelopathies such as mutations in the muscle-specific sodium (SCN4A) and chloride (CLCN1) channel genes are known to cause myotonia in humans and animals (Lehmann-Horn and Jurkat-Rott, 1999). The Clc-1 channel is the product of the CLCN1 gene and is the major skeletal muscle chloride channel in humans (Lehmann-Horn and Jurkat-Rott 1999).

Since defects in both sodium and chloride conductance have been documented in DM1, Charlet and co-workers investigated whether SCN4A and CLCN1 gene products are misregulated in the disease (Franke et al. 1990; Koty et al. 1996; Charlet et al. 2002). Their study showed, that similar to cTNT and the IR, the Clc-1 mRNA was aberrantly spliced resulting in 3 defective isoforms containing a premature termination codon (Fig. 1.20B) (Charlet et al. 2002; Mankodi et al. 2002). Using a DM mouse model it was determined that loss of Clc-1 function was sufficient to induce myotonia (Mankodi et al. 2002).

A shared mechanism links the processing of IR, cTNT and Clc-1: they are all natural targets for the alternative splicing-regulator CUG-BP (Ladd et al. 2001). CUG-BP has been shown to bind U/G-rich motifs in introns adjacent to splice sites in all three of these pre-mRNAs (Charlet et al. 2002; Philips et al. 1998; Savkur et al. 2001). Furthermore, over-expression of CUG-BP in normal cells induces the same defective alternative splicing as that seen in DM1 patients (Faustino and Cooper 2003). The exact mechanism by which CUG-BP accumulates in the nucleus of DM1 cells is unknown.

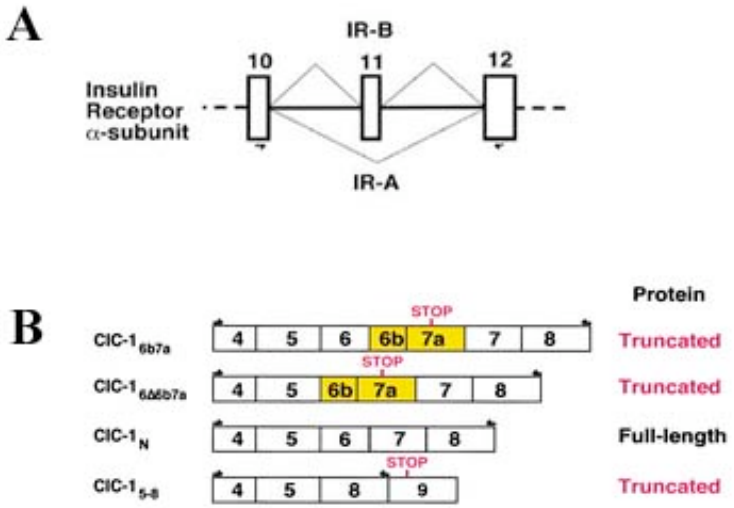



**Figure 1.20: Diagram of the alternative splicing of IR and Clc-1. (A)** Insulin receptor (IR) a-subunit alternative splicing. IR-A is produced with inclusion of exon 11, while the IR-B isoform is produced by exclusion of exon 11. IR-A is the high-signalling isoform normally found in insulin-responsive tissue like skeletal muscle and the liver. **(B)** Alternative splicing of the chloride channel-1 (Clc-1) mRNA. Aberrant splicing generates 3 isoforms coding for inactive proteins with premature stop codons. Illustration reproduced from: (A) (Savkur et al. 2001) (B) (Charlet et al. 2002).

The current explanation suggests that double strand RNA binding proteins (dsRNA-BP) attach themselves to the CUG expansion hairpin loop thereby forming the characteristic foci and consequently inducing an elevation in CUG-BP steady-state levels in DM1 cells (Charlet et al. 2002; Faustino and Cooper 2003; Miller et al. 2000; (Timchenko et al. 2001). In addition, CUG-BP has also been proposed as a substrate of DMPK. The haploinsufficiency or complete lack of DMPK might favours the hypophosphorylated form of CUG-BP that is shown to accumulate in the nucleus (Philips et al. 1998; Roberts et al. 1997).

# Experimental Work

# Chapter 2

## Crystallization and preliminary X-ray analysis of the coiled-coil domain of dystrophia myotonica kinase

Garcia, P., Marino, M. and Mayans, O. (2004) Crystallization and preliminary X-ray analysis of the coiled-coil domain of dystrophia myotonica kinase. *Acta Crystallogr D Biol Crystallogr*, 60, 2336-233

Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Crystallization and preliminary X-ray analysis of the coiled-coil domain of dystrophia myotonica kinase

**Pilar Garcia, Marco Marino and Olga Mayans***

Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Correspondence e-mail: olga.mayans@unibas.ch

Received 6 August 2004
Accepted 22 October 2004

The coiled-coil domain of dystrophia myotonica protein kinase (DMPK) has been cloned, overexpressed, purified and crystallized. Two crystal forms have been obtained that belong to space groups $P3$ and $P2_12_12_1$ and diffract to 2.4 and 1.6 Å resolution, respectively. Experimental phases were obtained by MAD from an SeMet derivative. The location of selenium sites used molecular-replacement phases obtained from search models lacking sequence similarity with the coiled-coil under study. Both crystal forms contain three polypeptide chains in the asymmetric unit.

## 1. Introduction

Regulation by oligomerization is a key feature of the dystrophia myotonica protein kinase (DMPK) related family of Ser/Thr kinases. This group of complex multi-domain kinases includes important members such as human $p160^{ROCK}$ (ROCKI), Rho-kinase (ROCKII), citron Rho-interacting kinase and Cdc42-binding kinase (MRCK) among others (Riento & Ridley, 2003). They are functionally related to reorganization events in the cytoskeleton during cellular processes such as cytokinesis, neurite outgrowth and smooth-muscle contraction. These kinases are characterized by the high conservation of their catalytic domains and by the unusual presence of one or more coiled-coil domains (CC) that are central to regulation of their activity. Based on sequence data, the CC domains from members of this family are predicted to form dimeric arrangements. However, it is remarkable that little or no overall sequence homology can be detected across these domains and that different biochemical roles have been attributed to them. At their C-termini, the CC moieties of ROCKI, ROCKII and citron kinase (Di Cunto *et al.*, 1998; Riento & Ridley, 2003) house binding motifs for small Rho GTPases that are known to regulate cytoskeletal dynamics. The CC domains of MRCK and DMPK are not known to mediate protein interactions, but have instead been proposed to act as an inhibitor and an enhancer of kinase activity, respectively (Tan *et al.*, 2001; Zhang & Epstein, 2003). Furthermore, it is generally accepted that these CC sections mediate dimerization in these kinases (Zhang & Epstein, 2003). However, it has been shown that other domains within these complex systems can also lead to oligomerization (Chen *et al.*, 2002; Tan *et al.*, 2001). In summary, the functionality of CC domains from this kinase family remains unclear in the absence of structural data.

To date, the structures of only two closely related CC fragments from this kinase family have been reported. These correspond to the C-termini of ROCKII-CC and ROCKI-CC, both of which are dimeric and span the Rho-binding motif. Their structures have been elucidated in the free form (PDB code 1uix; Shimizu *et al.*, 2003) and in complex with RhoA (PDB code 1s1c; Dvorsky *et al.*, 2004), respectively. In order to gain further understanding of the contribution of CC moieties to kinase regulation in this family, we have undertaken the study of the CC domain from DMPK (DMPK-CC). DMPK is speculated to be related to a progressive neuromuscular disorder known as myotonic dystrophy (DM), which is the most common form of muscular dystrophy in adults (Larkin & Fardaei, 2001). Its involvement in disease, however, remains polemic. Understanding the activation mechanism of DMPK will contribute to further clarifying its functionality.

## 2. Construction of the plasmid

Based on structure-prediction results from *MULTI-COIL* (Wolf *et al.*, 1997), DMPK-CC was identified to span residues 461–537. An expression construct comprising these residues was derived from human DMPK cDNA (GenBank L08835) using primers 5′-CA-TG**CCATGG**CAGAGGCTGAGGCCGAG-3′ and 5′-GG**GGTACC**TTACCCCGTGACAGC TGTGG-3′ carrying *Nco*I and *Kpn*I restriction sites (in bold). The PCR product was purified (Wizard, Promega) and cloned into the similarly digested expression vector pET-M11 (EMBL). This modified version of pET-24d (Novagen) includes a $His_6$ tag plus a TEV (tobacco etch virus) protease-cleavage site

N-terminal to the expression construct. The plasmid was constructed in *Escherichia coli* strain DH5$\alpha$.

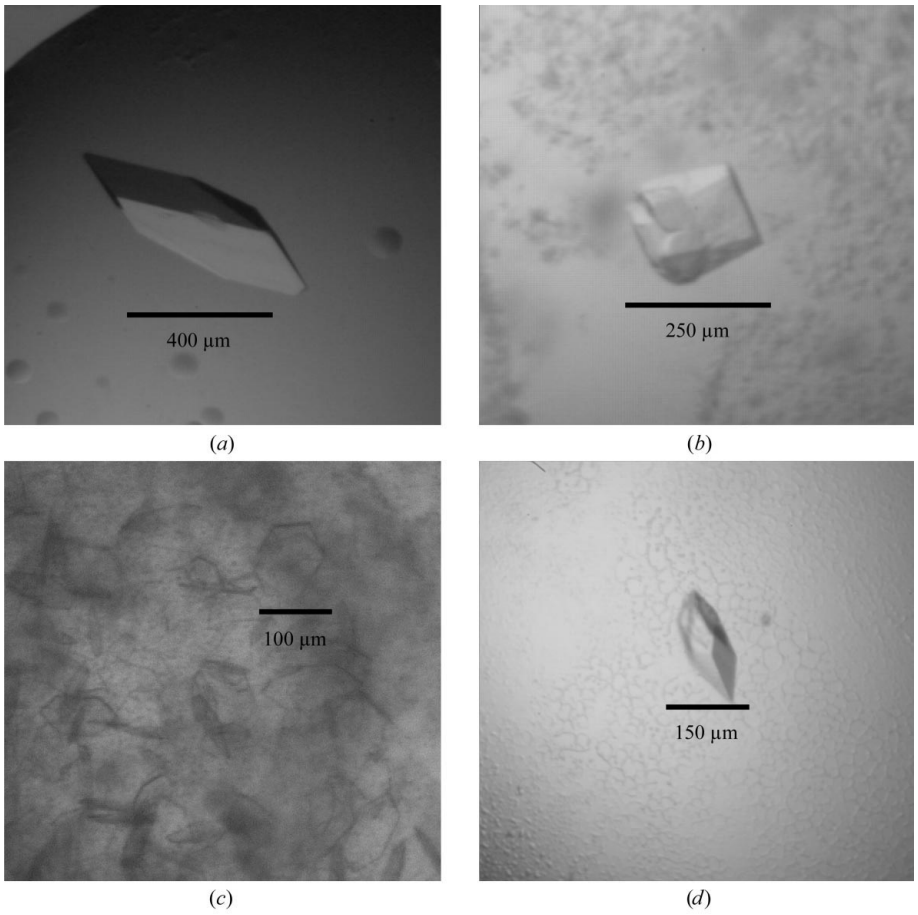
## 3. Protein overexpression and purification

DMPK-CC was expressed in *E. coli* BL21(DE3) grown in LB medium supplemented with 25 µg ml$^{-1}$ kanamycin and 34 µg ml$^{-1}$ chloramphenicol at 310 K. Induction took place at an $OD_{600\,nm}$ of 0.6 by the addition of 1 m*M* isopropyl-$\beta$-D-thiogalactopyranoside (IPTG) and expression took place overnight at 298 K. Cells were harvested by centrifugation at 5000*g* for 30 min at 277 K. The bacterial pellet was resuspended in lysis buffer (50 m*M* Tris pH 7.5, 50 m*M* NaCl) and sonicated in the presence of 1 mg ml$^{-1}$ lysozyme. The homogenate was clarified by centrifugation at 20 000*g* for 1 h and 277 K. For affinity purification, the supernatant was applied onto a HiTrap $Ni^{2+}$-Chelating HP column (Amersham Biosciences) equilibrated in lysis buffer containing 40 m*M* imidazole. Protein was eluted at 200 m*M* imidazole. Tag removal took place by incubation with TEV protease (0.4 mg ml$^{-1}$) for 2 h at room temperature during dialysis against 50 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl. Subtractive purification of protease ($His_6$-tagged at the C-terminus) and non-digested protein was carried out on a subsequent chelating step. The cleaved protein was then concentrated to 5–10 mg ml$^{-1}$ according to a Bradford assay (Bradford, 1976) and used in crystallization trials. Higher purity DMPK-CC could be obtained by further size-exclusion chromatography using a Hi-Load 16/60 Superdex 75 prep-grade column (Amersham Biosciences) equilibrated in 50 m*M* Tris pH 7.5 and 200 m*M* NaCl. However, this additional procedure did not result in detectable variation of crystallization results and was discontinued.

**Table 1**
X-ray data and phasing statistics.

Data processing and reduction used the *MOSFLM*/*SCALA* software (Kabsch, 1988; Leslie, 1992). Phase calculation and refinement was performed in *SHARP* (de La Fortelle and Bricogne, 1997) using data to 3.0 Å resolution. Phasing statistics are given for acentric reflections.

| | *P*3 | $P2_12_12_1$ | | | |
|---|---|---|---|---|---|
| | Native | Native | SeMet peak | SeMet remote | SeMet edge |
| Unit-cell parameters (Å) | $a = b = 38.05$, $c = 114.37$ | $a = 39.09$, $b = 46.17$, $c = 143.48$ | | | |
| No. chains in AU | 3 | 3 | | | |
| $V_M$ (Å$^3$ Da$^{-1}$) | 1.3 | 2.4 | | | |
| Solvent content (%) | 32 | 48 | | | |
| Wavelength (Å) | 0.886 | 0.98 | 0.9793 | 0.9392 | 0.9795 |
| Resolution (Å) | 20–2.4 (2.53–2.40) | 20–1.6 (1.69–1.6) | 20–2.3 (2.42–2.3) | 20–2.5 (2.64–2.5) | 20–2.9 (3.06–2.9) |
| Unique reflections | 5395 (728) | 33876 (3951) | 11724 (1440) | 9565 (1379) | 6105 (873) |
| Completeness (%) | 74.1 (68.2) | 96.3 (79.2) | 97.2 (83.9) | 99.5 (99.5) | 99.2 (100.0) |
| $R_{sym}(I)$ | 0.123 (0.228) | 0.059 (0.330) | 0.079 (0.204) | 0.091 (0.316) | 0.104 (0.304) |
| Redundancy | 2.0 (1.3) | 4.4 (2.3) | 12.7 (8.7) | 6.8 (7.0) | 6.7 (7.0) |
| $\langle I/\sigma(I)\rangle$ | 6.5 (1.4) | 15.0 (2.5) | 25.6 (8.2) | 17.8 (6.2) | 16.5 (6.1) |
| Phasing power (iso/ano) | | | 0.15/1.83 | −/0.96 | 0.58/0.55 |
| $R_{Cullis}$ | | | 1.0/0.64 | −/0.84 | 0.91/0.93 |
| FOM (before/after solvent treatment) | | | | 0.45/0.87 | |




(*a*) (*b*) (*c*) (*d*)

**Figure 1**
Crystal habits. (*a*) Crystals grown under condition (i) supplemented with 3% 2-propanol. (*b*) Crystals grown under condition (i) supplemented with 5% dioxane. (*c*) Thin crystal plates grown from heavy precipitate under condition (ii). (*d*) Best diffracting crystals obtained from condition (iii).

## 4. Crystallization

Crystallization was carried out by the hanging-drop vapour-diffusion method at 291 K. Each drop consisted of equal volumes (1.5 µl) of protein solution and reservoir solution. The components of the protein solution are described above and the protein concentration was 10 mg ml$^{-1}$. Initial screening used a wide range of commercially available crystallization kits. Crystals were obtained from Structure Screen 1 (Molecular Dimensions) under two sets of conditions: (i) 2.0 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 8.5 and (ii) 100 m*M* sodium acetate pH 4.6, 8% PEG 4000. Crystals obtained under condition (i) grew as stacks of multiple plates. Single crystals grew from optimized media containing 2.5 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 7.5 and either 3% 2-propanol or 5% dioxane (Figs. 1*a* and 1*b*). Despite the good visual appearance of these crystals, their poor diffraction quality made

them unsuitable for structure elucidation. As a consequence, this crystal form was abandoned. Crystals grown under condition (ii) appeared one month after drop mixing. They grew from heavy precipitate as extremely thin hexagonal plates of small dimensions (Fig. 1*c*). Efforts to increase the crystal size, including seeding protocols, yielded no improvement.

Finally, a customized screen resulted in an additional crystal type grown from 100 m*M* Tris–HCl pH 7.5, 2.0 *M* sodium/potassium tartrate (condition iii). Crystals exhibited a polyhedral habit and grew to final dimensions within two to three weeks of drop mixing (Fig. 1*d*). The diffraction properties of these showed them to be suitable for structure elucidation.

## 5. Data collection and NCS analysis

Native diffraction data have been recorded from crystals grown under conditions (ii) and (iii). For condition (ii), data were collected from a crystal extracted from the drop shown in Fig. 1(*c*). Diffraction was recorded to 2.4 Å resolution at beamline XS0A6 (SLS, Villingen) using a MAR CCD 165 detector and under cryoconditions in 27.5%(*v*/*v*) glycerol. Reflections were recorded in a single sweep of 70°, using 1.0° non-overlapping rotation steps. Crystals belong to space group *P*3. Unit-cell dimensions and data statistics are given in Table 1. The diffraction from these thin plate crystals suffered strongly from radiation damage, impeding the collection of complete data of high quality. Analysis of the crystal lattice using the self-rotation function in *AMoRe* (Navaza, 2001) revealed NCS relations at Eulerian angles $\alpha = 60$ and 180°, $\beta = \gamma = 0°$. This is consistent with the presence of three $\alpha$-chains in the asymmetric unit aligned with the crystallographic *c* axis, which is in agreement with the fact that the dimension of this axis closely resembles the theoretical length calculated for this coiled-coil.

X-ray data optimal for structure elucidation were recorded to 1.6 Å resolution from crystals grown under condition (iii). Data collection took place at beamline ID14-2 (ESRF, Grenoble) on an ADSC Quantum 4 CCD detector and under cryoconditions using 20%(*v*/*v*) glycerol as a cryoprotectant. Reflections were recorded over a total rotation range of 105°, using 0.70° non-overlapping steps and two sweeps at high and low exposure doses (data statistics are given in Table 1). Crystals belong to space group $P2_12_12_1$ as confidently estimated from the pattern of systematic absences along the crystallographic axes. The calculated $V_M$ value (Matthews, 1968) (Table 1) suggested the presence of more than one chain in the asymmetric unit of this crystal form. However, inspection of self-rotation functions and native Patterson maps did not reveal NCS relations.

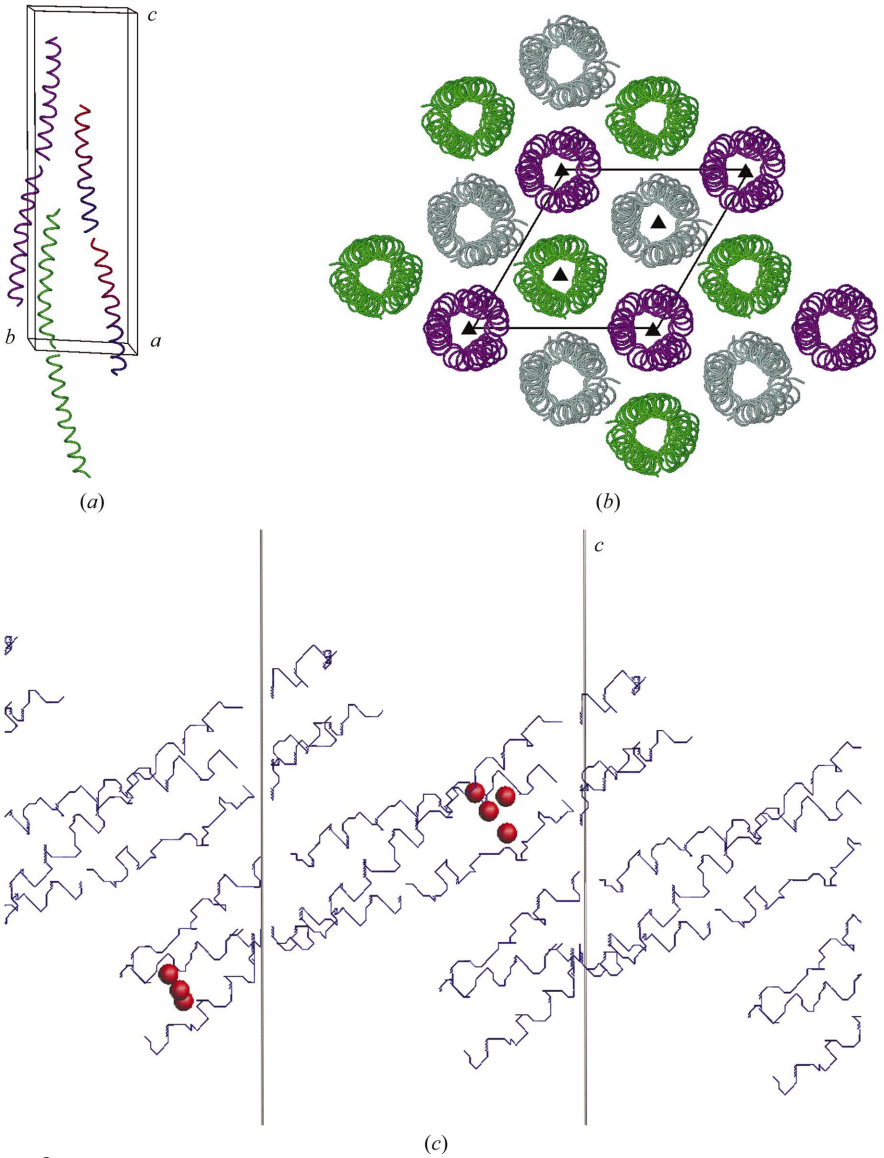



**Figure 2**
(*a*) Crystal lattices in the *P*3 and $P2_12_12_1$ space groups. The asymmetric unit of the *P*3 cell contains three $\alpha$-helical chains aligned with the *c* crystallographic axis. Each chain originates from the fusion of two consecutive short helices as positioned by MR. To indicate this, one of the chains has its two halves coloured in a blue-to-red gradient, where blue indicates the N-termini and red the C-termini. (*b*) View of the *P*3 lattice perpendicular to the crystallographic threefold axis. Each NCS chain generates a full coiled-coil by applying crystal symmetry (reflected by different colouring). Crystallographic threefolds are indicated. (*c*) Experimental skeletons (after solvent treatment in *SHARP*) obtained by MAD phasing in $P2_12_12_1$ symmetry. SeMet sites located using MR phases and employed in phase calculation are shown as full circles. Five sites are unique, while one additional site has two related peaks, possibly owing to a double Met conformation.

## 6. Phasing

Phasing was initially attempted by molecular replacement (MR) in *AMoRe* (Navaza, 2001). Although DMPK-CC shares no sequence similarity with other coiled-coils for which structures are available, the structural similarity within this fold was expected to be sufficiently high to yield successful results in MR. This has been the case for other folds with a highly conserved architecture, such as the parallel $\beta$-helix, where correct solutions were obtained from search models with sequence identity values as low as 17% (Mayans *et al.*, 1997). In the current study, coiled-coil structures as well as their constituent single chains have been assayed for PDB entries 1avy, 1bb1, 1gzl, 1ij0 (trimeric models), 1ic2 (dimeric model),

1fe6 and 1gcl (tetrameric models). Only CC fragments without discontinuities were considered. All models were used as poly-alanine variants to account for the lack of sequence similarity. Because of the elongated shape of the models, a wide range of Patterson integration radii were tested. In general, short chains performed better than long ones, probably owing to long-range deviation along the coiled-coil axis between the search and target structures.

In space group *P*3, the best search model consisted of a single $\alpha$-helical chain from 1ij0 (a trimeric variant of GCN4 leucine zipper 31 amino acids long). This yielded two best MR solutions (correlation coefficient 60.1% and *R* factor 52.2% after rigid-body fitting in *AMoRe* in the resolution range 12–3.5 Å), which corresponded to two different halves of the same $\alpha$-helical chain of DMPK-CC (Fig. 2*a*). Both halves were fused to create an optimized search model 62 residues long. A new MR search resulted in three solutions corresponding to three NCS copies (correlation coefficient 68.2%, *R* factor 47.1%; Fig. 2*a*). Each chain generated a full trimeric coiled-coil by applying crystal symmetry. The lattice thus calculated (Fig. 2*b*) was in good agreement with NCS data for this crystal form.

In space group $P2_12_12_1$, trimeric 1ij0 was the best search model. Two solutions were obtained (correlation coefficient 53.3%, *R* factor 54.2% after rigid-body fitting in *AMoRe* in the resolution range 12–3.5 Å), each corresponding to half of DMPK-CC with a slight overlap along the helical axis. An equivalent MR solution was obtained when using the trimeric DMPK-CC model obtained by MR in the *P*3 space group as a search model.

Despite good crystal packing, the resultant MR phases were of insufficient quality to allow model building and refinement. Thus, experimental phasing was pursued by MAD on an SeMet derivative crystallized in space group $P2_12_12_1$. Diffraction was recorded at peak, inflection and high-energy remote points at ID14-2 (ESRF, Grenoble) (Table 1). Six selenium sites out of a total of nine (it should be noted that the three N-terminal Met residues are derived from the plasmid and are expected to be structurally disordered) have been located by anomalous difference Fourier ($>5\sigma$) using phases calculated from MR solutions (Fig. 2*c*). The sites within the core of the MR model are in good agreement with anomalous difference Patterson maps and are at the expected separation distance along the coil axis. Visual inspection of high-resolution electron-density maps calculated from the experimental phases using these selenium positions showed the MR solutions to be essentially correct and confirmed the trimeric state of DMPK-CC in these crystals. However, they also revealed an inversion of the $\alpha$-helix termini in the MR model, which was probably a consequence of the lack of definition in the helix when removing side chains and at the resolution of the MR studies. This explains the low quality of these initial phases. Structure elucidation will now proceed making use of experimental phases. The fact that both crystal forms in this study, which were obtained at high and low salt conditions as well as at pH 4.6 and pH 7.5, contain a trimer in their asymmetric units leads us to speculate that the association state of this coiled-coil must be stable. However, further biophysical studies must be undertaken to establish the oligomeric state of this coiled-coil in solution and to resolve whether the currently observed arrangement might be the result of the crystallization media or of packing forces.

Very special thanks go to Professor Benjamin Perryman for facilitating the initial DNA coding for DMPK. We would like to acknowledge the support to this project from the Swiss National Foundation (grant 3100A0-100852) and the Association Française contre les Myopathies (grant No. 8421). Also, a most special mention should go to the staff of beamlines ID14-2 at ESRF and XS06A at SLS for providing exceptional assistance during data collection.

## References

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Chen, X. Q., Tan, I., Ng, C. H., Hall, C., Lim, L. & Leung, T. (2002). *J. Biol. Chem.* **277**, 12680–12688.
Di Cunto, F., Calautti, E., Hsiao, J., Ong, L., Topley, G., Turco, E. & Dotto, G. P. (1998). *J. Biol. Chem.* **273**, 29706–29711.
Dvorsky, R., Blumenstein, L., Vetter, I. R. & Ahmadian, M. R. (2004). *J. Biol. Chem.* **279**, 7098–7104.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Larkin, K. & Fardaei, M. (2001). *Brain Res. Bull.* **56**, 389–395.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EAMCB Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., Pickersgill, R. & Jenkins, J. (1997). *Structure*, **5**, 677–689.
Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.
Riento, K. & Ridley, A. J. (2003). *Nature Rev. Mol. Cell Biol.* **4**, 446–456.
Shimizu, T., Ihara, K., Maesaki, R., Amano, M., Kaibuchi, K. & Hakoshima, T. (2003). *J. Biol. Chem.* **278**, 46046–46051.
Tan, I., Seow, K. T., Lim, L. & Leung, T. (2001). *Mol. Cell. Biol.* **21**, 2767–2778.
Wolf, E., Kim, P. S. & Berger, B. (1997). *Protein Sci.* **6**, 1179–1189.
Zhang, R. & Epstein, H. F. (2003). *FEBS Lett.* **546**, 281–287.

# Chapter 3

## Molecular insights into the self-assembly mechanism of Dystrophia Myotonica Kinase

Garcia, P., Ucurum, Z., Bucher, R., Svergun, D.I., Huber, T., Lustig, A., Konarev, P.V., Marino, M., and Mayans, O. 2006. Molecular insights into the self-assembly mechanism of dystrophia myotonica kinase. *Faseb J* **20**(8): 1142-1151.

# Molecular insights into the self-assembly mechanism of dystrophia myotonica kinase

**Pilar Garcia,* Zöhre Ucurum,* Rainer Bucher,* Dmitri I. Svergun,† Thomas Huber,‡ Ariel Lustig,* Petr V. Konarev,† Marco Marino,* Olga Mayans*,[1]**

*Division of Structural Biology, Biozentrum, University of Basel, Basel, Switzerland; †European Molecular Biology Laboratory, Hamburg Outstation, c/o DESY, Hamburg, Germany and Institute of Crystallography, Russian Academy of Sciences, Moscow, Russia; and ‡Biochemisches Institut, Universität Zürich, Zürich, Switzerland

ABSTRACT **Self-assembly via coiled-coil domains (CC) is crucial for the regulation of the dystrophia myotonica kinase (DMPK) -related family of kinases. These CC domains are thought to form dimeric arrangements and thus to mediate dimerization in these enzymes. Using size exclusion chromatography combined with multiangle static light scattering, we analyzed the oligomeric state of DMPK as well as that of a truncated variant lacking the CC fraction. Remarkably, both forms were found to assemble into robust dimers. In contrast, the CC domain in isolation yielded trimeric assemblies, indicating that the oligomerization properties of CC domains from this kinase family are more diversified than anticipated. The crystal structure of this CC has been elucidated to 1.6 Å resolution and its properties in solution established using sedimentation equilibrium and thermal denaturation. These data show that, contrary to expectations, the self-assembly of DMPK is not dictated by the association properties of its CC domain. Instead, it appears to be driven by sequence segments flanking both N and C termini of the catalytic kinase fraction, as suggested by models of head-to-head dimers based on small angle X-ray scattering data. Our findings support a shared pattern of assembly across DMPK, ROCKs, and MRCK members of this family.—Garcia, P., Ucurum, Z., Bucher, R., Svergun, D. I., Huber, T., Lustig, A., Konarev, P. V., Marino, M., Mayans, O. Molecular insights into the self-assembly mechanism of dystrophia myotonica kinase. *FASEB J.* 20, 1142–1151 (2006)**

*Key Words: kinase · coiled-coil · protein assembly · protein structure · X-ray crystallography · multiangle static light scattering · small angle X-ray scattering*

Self-assembly is a main feature of the dystrophia myotonica kinase (DMPK) -related family of kinases. These complex multidomain kinases share a highly conserved catalytic domain and a coiled-coil fraction (CC) involved in the regulation of their activity. To this group belong DMPK (1, 2), Rho-associated kinases I and II (ROCKI/ II, also known as β and α, respectively) (3, 4), *C. elegans* LET-502 (5), *Drosophila* Genghis Khan kinase (Gek) (6), murine citron kinase (7), and rat Cdc42 binding kinase (MRCK) (8), among others. Most of these kinases act as effectors of small Rho GTPases (Rho, Rac, and Cdc42 subclasses) known to regulate cytoskeletal dynamics. Accordingly, several members of the family are functionally linked to cytoskeletal reorganization events, as illustrated by the involvement of Gek in actin polymerization (6) and the role of citron kinase, LET-502, and ROCKII in regulating the cleavage furrow contraction during cytokinesis (9–11).

DMPK is regarded as the defining member of this kinase family. It is most closely related to MRCK and ROCKs, with which it shares sequence similarity across the N-terminal catalytic fraction but not across CC moieties (**Fig. 1**). By analogy to other family members, DMPK is suspected to regulate cytoskeletal events (12, 13) and may act as a downstream effector in the regulatory pathways of RhoA (12) and Rac1 (14). Other reports, however, defend that DMPK intervenes in the modulation of the electrical excitability of the plasma membrane (15, 16). Taken together, functional data suggest that DMPK might be related to muscle stiffness and hyperexcitability, heart conduction defects, and cognitive deficit in a progressive neuromuscular disorder known as myotonic dystrophy (DM1), which is the most common form of muscular dystrophy in adults (17). DM1 patients suffer from DMPK insufficiency due to the retention of mRNA transcripts in the nucleus (18, 19). However, it is not currently understood how low levels of this kinase affect cellular function, and thus its involvement in disease remains polemic.

Oligomerization via CC segments is thought to be central to the regulation of DMPK-related kinases. CC motifs consist of two to five amphipathic α-helices that wind around each other to form, typically, a left-handed supercoil. They mediate self-assembly events in proteins but can also be implicated in refolding processes and molecular recognition (for a recent review, see ref 20). CC sequences are characterized by a repeat

[1] Correspondence: Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstr. 70, Basel CH-4056, Switzerland. E-mail: olga.mayans@unibas.ch

doi: 10.1096/fj.05-5262com



0892-6638/06/0020-1142 © FASEB

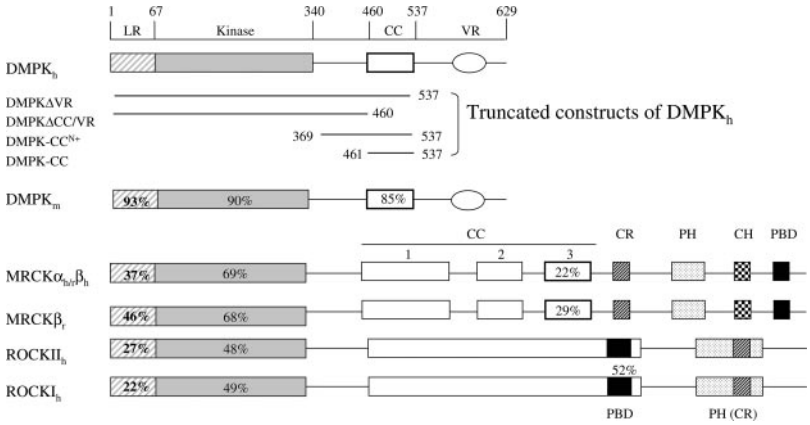



**Figure 1.** Domain organization in DMPK, MRCK, and ROCK. Domain composition of human DMPK (Q09013–9), mouse DMPK (P54265), human MRCKα (Q5VT26), rat MRCKα (AF021935), human MRCKβ (Q81WQ7), mouse ROCKII (V58513), and mouse ROCKI (V58512). LR: leucine-rich region; CC: coiled-coil; VR: region variable across DMPK spliceoforms. CR: cysteine-rich domain, PH: pleckstrin homology domain; CH: citron homology domain; PBD: p21 GTPase binding domain. Sequence identity to human DMPK is given as percentage where applicable. DMPK constructs used in the current study are shown.

pattern of seven amino acids denoted *a* to *g*. Positions *a* and *d* classically harbor hydrophobic residues, which constitute the structural core of the motif. Positions *e* and *g* often host charged groups that form networks of intra- and interhelical salt bridges crucial to fold stability and interchain recognition. CC domains from DMPK-related kinases are believed to form dimeric assemblies and thereby to mediate dimerization in these enzymes. To date, only two closely related CC fragments from this family have been characterized. These correspond to the C-terminal, Rho binding motifs of ROCKII-CC and ROCKI-CC that share 52% sequence identity (Fig. 1). For these, crystal structures have been reported in free form (21) and complexed to RhoA (22), respectively, confirming that both coils form parallel dimers. In a similar way, only a few analytical studies have reported on the oligomerization state of full-length kinases. Currently, data are available only for DMPK (23) and ROCKs (24, 25), which have been shown to adopt dimeric states. Despite the similarity of these kinases, different mechanisms have been proposed for their dimerization. While data on DMPK defended a self-assembly driven by its CC motif (23), the reports on ROCKs prove that CC moieties are not required for their oligomerization. The latter agrees with a biochemical study of MRCK, which also reported self-association in the absence of CC fractions (26). The associative interactions in ROCKs are the best understood so far due to the recent availability of the crystal structure from the dimeric N-terminal fraction (i.e., lacking CC segments) from ROCKI (25). This shows how head-to-head dimers interact via a small, N-terminal domain comprising the α-helical LR region (Fig. 1) as well as a tail, C-terminal to the kinase domain, which wraps around the catalytic core to reach the subunit interface. Given that dimerization segments show a certain homology across DMPK, ROCKs, and MRCK (Fig. 1), it would be reasonable to expect that related mechanism of assembly could be supported by these kinases, even if conflicting data have been reported (23).

Moreover, different biochemical roles have been attributed to the CC moieties of these kinases. The CC domains of ROCKs and citron kinase act as intrasteric inhibitors of their respective kinase domains and house at their C termini binding motifs for small Rho GTPases, where binding of the latter releases the inhibition (7, 27). CC segments of MRCK (CC2/CC3) also act as intrasteric inhibitors but do not host a GTPase binding region (26), and physiological release factors are yet to be identified. On the contrary, the CC fraction of DMPK (DMPK-CC) has been proposed to act as an enhancer of its kinase activity (23), although it is unknown whether this involves transphosphorylation events as those observed in tyrosine kinases. In summary, both the functionality of CC domains from this family and the extent to which they govern molecular assembly remain unclear. Given that the CC from DMPK appears to differ from those of ROCKs and MRCK both in its regulatory properties and in its role during assembly (23), we have investigated the architectural details of DMPK to assess whether a common association principle exists in this family of kinases.

## MATERIALS AND METHODS

### Cloning

Expression constructs were derived from human DMPK cDNA (L08835) spanning residue ranges as shown in Fig 1. DMPKΔVR was cloned into the *Eco*RI and *Sal*I sites of the insect cell expression vector pFastBac1 (Invitrogen, San Diego, CA, USA). A $His_6$-tag and a TEV (tobacco etch virus) protease cleavage site were inserted N-terminal to the target construct. DMPKΔCC/VR was cloned into pET-15b (Novagen, Madison, WI, USA) using primers 5′-GGAATTCCATATGTCAGCCGAGGTGCG-3′ and 5′-CGGGATCCTTACCCCGTGACAGCTGTGG-3′ with *Nde*I and *Bam*HI restriction sites, respectively. This vector adds a $His_6$-tag and a thrombin cleavage site N-terminal to the target protein. Cloning of DMPK-CC has been previously reported (28). DMPK-$CC^{N+}$ was cloned into the pETM-11 vector, a modified version of pET-24d (Novagen) that includes a $His_6$-tag plus a TEV protease cleavage site N-terminal to the expression construct. This used primers 5′-CGTTCGAAGACTACATGTTGGTGGAGGACGGGCTC-3′ and 5′-GGGGTACCTTACCCCGTGACAGCTGTGG-3′ carrying BbsI (*Nco*I-compatible) and *Kpn*I restriction sites.

### Protein production

Sf9 cells were transfected with the bacmic cDNA construct of DMPKΔVR according to the manufacturer's instructions

(Life Technologies Bac-to-Bac system). The transfected cells were incubated for 3 days at 27°C and the supernatant, containing recombinant virus for protein expression, was collected. Amplified virus (10 ml) was added to $1 \times 10^6$ Sf9 cells and incubated at 27°C for 3 days. Lysis was performed by incubation on ice for 30 min in the presence of DNase and protease inhibitors (Complete EDTA-free; Roche, Nutley, NJ, USA) in 50 mM Tris pH 8.0, 300 mM NaCl, 5 mM β-mercaptoethanol, 2% glycerol. A clarified supernatant was obtained by centrifugation at 20,000 rpm and 4°C for 1 h. The protein product was purified by affinity chromatography on a His-Trap HP column (GE Healthcare) equilibrated in lysis media. The bound protein was eluted by supplementing the medium with 200 mM imidazole. Subsequent purification used size exclusion chromatography on a Superdex 200 Tricorn 10/30 GL column (GE Healthcare) in 50 mM Tris pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 2% glycerol.

DMPKΔCC/VR was expressed in *E. coli* Rosetta (DE3) (Novagen) grown in LB medium supplemented with 34 μg/ml of chloramphenicol and 100 μg/ml ampicillin. Induction used 1 mM IPTG and incubation was overnight at 25°C. Cells were harvested by centrifugation at 8000 *g* at 4°C for 15 min. Lysis was by French pressing in the presence of DNase and protease inhibitors in 50 mM Tris pH 8.0, 50 mM NaCl, 5 mM β-mercaptoethanol. The clarified supernatant, obtained by centrifugation at 20,000 rpm (4°C), was purified by affinity chromatography and subsequent gel filtration (as described for DMPKΔVR) in lysis media. Expression and purification of DMPK-CC were as reported (28). Production of DMPK-$CC^{N+}$ used protocols as for DMPK-CC, except that all media were supplemented with 5 mM β-mercaptoethanol. All measurements used fresh preparations and sample identification was by peptide-digest mass spectrometry.

### Structure determination

Crystallization, data collection, data statistics, and experimental phasing of DMPK-CC have been described (28). Briefly, crystals grew from 100 mM Tris-HCl pH 7.5, 2.0 M $Na^+/K^+$ tartrate and 5–10 mg/ml protein solutions. X-ray diffraction data collected at beamline ID14–2 (ESRF, Grenoble) revealed the crystals to belong to space group $P2_12_12_1$ with unit cell dimensions of a = 39.1 Å, b = 46.2 Å, c = 143.5 Å, and three polypeptide chains per asymmetric unit. Experimental phasing used 3λ MAD on an Se-Met derivative. Phase refinement was in SHARP (29) and automatic model improvement in ARP/warp v6.1 (30). Manual model building used the software O (31) and refinement was in REFMAC (32). For cross-validation, crystallographic data were divided into a working and a test set using FREERFLAG (33). Restraints for non-crystallographic symmetry were applied during the initial phase of refinement and released at later stages. Geometrical parameters of intermediate and final models were assessed using PROCHECK (34). The final model lacks several N- and C-terminal residues that could not be resolved in electron density maps, namely 461–465/530–537 in chain A, 461–463/528–537 in chain B and 461–465/527–537 in chain C. This amounts to ~20% of disorder in the structure and explains the slightly elevated R-factor values, which are somewhat above those characteristic of this resolution. Coordinates have been deposited at the Protein Data Bank with accession number 1WT6.

### Analytical ultracentrifugation

Sedimentation velocity and equilibrium data were recorded using an Optima XL-A analytical ultracentrifuge (Beckman Instruments, Fullerton, CA, USA) equipped with 4 and 12-mm Epon double-sector cells in an An-60 Ti rotor and absorption optics. Runs were performed at 20°C in *1)* 50 mM Tris pH 7.5, 50 mM NaCl and *2)* 100 mM sodium phosphate pH 7.0 using detection wavelengths of 277 nm and 234 nm. Sedimentation velocity experiments were carried out at 54,000 rpm and 0.65 mg/ml sample concentration. Sedimentation equilibrium measurements were performed at 19,000, 21,000, 24,000, and 26,000 rpm, and protein concentrations of 0.16–3.3 mg/ml for DMPK-CC and 0.09–1.4 mg/ml for DMPK-$CC^{N+}$, where optical density values remained in a linear range. Average molecular masses were evaluated using SEGAL (G. Machaidze and A. Lustig, unpublished software), which adjusts the baseline absorbance to obtain the best linear fit of ln(absorbance) vs. radial distance square. A protein partial specific volume of 0.73 ml/g was used, while solution density was taken as 1.003 g/ml and viscosity as 1.02 centipoise.

### Circular dichroism spectroscopy

Thermal denaturation of DMPK-CC was monitored at 222 nM on a 62A DS circular dichroism spectropolarimeter (AVIV) equipped with a temperature-controlled quartz cell of 2 mm optical path over a temperature range of 5–90°C. The scan speed was 0.5°C/min with a spectral bandwidth of 2.0 nm.

### Binding assays

The binding of DMPK-CC to DMPKΔCC/VR was assayed in 50 mM Tris-HCl pH 7.5, 100 mM NaCl by *1)* affinity pull-down assays using $His_6$-tagged DMPKΔCC/VR immobilized in a His-Trap HP column (GE Healthcare) vs. a flow of untagged DMPK-CC; *2)* simultaneous coexpression of both constructs followed by attempts of copurification via metal affinity plus size exclusion chromatography (protocols as above).

### Size exclusion chromatography combined with multiangle light scattering (SEC-MALS)

The oligomeric state of DMPKΔVR and DMPKΔCC/VR in solution was determined via SECMALS measurements performed on an ÅKTA explorer 10 system (GE Healthcare) connected to a triangle light scattering detector and a differential refractometer (miniDAWN Tristar and Optilab, respectively; Wyatt Technology, Santa Barbara, CA, USA). A Superdex 200 HR 10/30 column (GE Healthcare) was used in 50 mM Tris-HCl pH 8.0, 50 mM NaCl at a flow rate of 0.5 ml/min. Sample volumes of 50 μl (DMPKΔVR) and 100 μl (DMPKΔCC/VR) were injected at a concentration of 2 mg/ml. A specific refractive index increment (dn/dc) value of 0.185 ml/g was used. The data were recorded and processed using ASTRA software (Wyatt Technology). To determine the detector delay volumes and normalization coefficients for the MALS detector, a BSA sample (Sigma A-8531) was used as reference. Neither despiking nor band broadening correction was applied.

### Small angle X-ray scattering (SAXS)

Synchrotron radiation SAXS data were collected on the X33 camera (35) of EMBL (DESY, Hamburg). Solutions of DMPKΔCC/VR were measured at protein concentrations of 1.4, 3.8, 5.7, 8.5, and 9.3 mg/ml in 50 mM TrisHCl pH 8.0, 50 mM NaCl, 2.5 mM β-mercaptoethanol at 4°C. Measurements were recorded using a MAR345 IP detector at a sample detector distance of 2.7 m and wavelength λ = 1.5 Å in the range of momentum transfer $0.011 < s < 0.45$ $Å^{-1}$. Exposure time for a single measurement was 5 min. To check for

TABLE 1. *X-ray data statistics and model parameters*

| | |
|---|---|
| X-ray data | |
| Space group | $P2_12_12_1$ |
| Resolution (Å) | 20–1.6 (1.69–1.6) |
| Unique reflections | 33876 (3951) |
| Completeness (%) | 96.3 (79.2) |
| $R_{sym}(I)$ | 0.059 (0.330) |
| Multiplicity | 4.4 (2.3) |
| $<I/\sigma(I)>$ | 15.0 (2.5) |
| Model refinement | |
| Resolution (Å) | 20–1.6 |
| Reflections working/free | 32044/1712 |
| R-factor/R-free (%) | 22.7/27.0 |
| Chains per asymmetric unit | 3 |
| No. protein/solvent atoms | 1619/294 |
| Average B ($Å^2$) | 26.05 |
| Rmsd bond lengths (Å)/ angles (°) | 0.019/1.717 |
| Residues in Ramachandran core / disallowed (%) | 99.5/0 |

radiation damage, two consecutive 3 min exposures at the highest protein concentration were compared and no changes were observed. The data were normalized, the scattering of the buffer subtracted, and the difference curves extrapolated to zero solute concentration using the program package PRIMUS (36). The maximum particle dimension $D_{max}$, forward scattering I (0) and the radius of gyration $R_g$ were evaluated using GNOM (37). The molecular mass (MM) of the solute was calculated by comparing its forward scattering with that of a reference solution of BSA (MM=66 kDa). The volume of the hydrated particle was computed using the Porod invariant equation (38). The low resolution shape of the molecule was computed *ab initio* using the program DAMMIN (39), which uses an assembly of densely packed beads to represent the protein molecule. Simulated annealing was used to build a compact, interconnected configuration of beads inside a sphere with the diameter $D_{max}$ that fits the experimental data $I_{exp}(s)$ to minimize the discrepancy:

$$\chi^2 = \frac{1}{N-1}\sum_j \left[\frac{I_{exp}(s_j) - cI_{calc}(s_j)}{\sigma(s_j)}\right]^2$$

where N is the number of experimental points, $c$ is a scaling factor and $I_{calc}(s)$, and $\sigma(s_j)$ are the calculated intensity and the experimental error at the momentum transfer $s_j$, respectively.

The scattering from the crystallographic model of ROCKI (protein fraction of PDB entry 2ETR; 25) was calculated using CRYSOL (40). The model of dimeric DMPKΔCC/VR was constructed based on the structure of ROCKI using the program BUNCH (41). The dimerization domains consisting of the 70 N-terminal and 20 C-terminal residues of ROCKI remained fixed during the calculations, so that the subunit interface stayed unaltered. The kinase domain was allowed to move as a rigid body while the missing C-terminal tail in DMPKΔCC/VR was represented as an interconnected chain of dummy residues (for domain definition, see Fig 6*a*) (42). A simulated annealing protocol was used to position the kinase domain and find the configuration of the C-terminal fragment that could fit the experimental scattering while obeying the P2 symmetry of the dimer.

## RESULTS

### Crystal structure of DMPK-CC

The structure of DMPK-CC (residues 461–537; Fig. 1) has been elucidated by X-ray crystallography up to 1.6 Å resolution. X-ray data statistics and model refinement parameters are given in **Table 1**. Crystals contain three polypeptide chains in the asymmetric unit, which form a consecutive coil of 86.5 Å length, spanning residues 466–527 and comprising eight complete heptad repeats. They build a parallel, left-handed trimeric motif (**Fig 2*a***).

The architecture of DMPK-CC is restrained at its terminal sections while a central region exhibits moderate structural fluctuations as indicated by crystallographic temperature factors and interchain conformational agreement (Fig 2*a, b*). The relative structural slackness at the middle point of the domain correlates with a widening of the coil at that position (Fig 2*c*), which is likely to be caused by the presence of a sequence motif, 493-IRTD-496, with low propensity for CC formation (estimated using MultiCoil; 43). Also noticeable is a residue located upstream, R483, which contributes a charged group in core position *a* with a reduced Crick-phase (–2.8° instead of the commonly

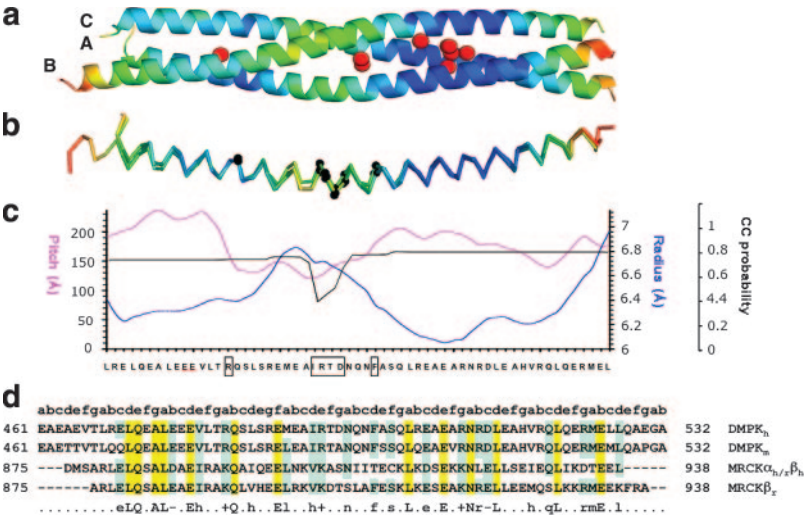



**Figure 2.** Crystal structure of DMPK-CC. *a)* Ribbon representation colored in a blue-to-red gradient according to crystallographic B-factors (8 to 55 $Å^2$), where blue designates low values. Solvent atoms in the interior of the CC motif are displayed as red spheres; *b)* Superimposition of chains A, B, and C colored as in panel *a*. Cα atoms of residues influencing coil parameters are shown as black spheres; *c)* coil parameter values for pitch, radius and residue tendency to coil formation. Residues leading to coil alterations are boxed; *d)* sequence alignment of DMPK-CC and MRCK-CC3. Identity is shown in yellow and similarity in green.

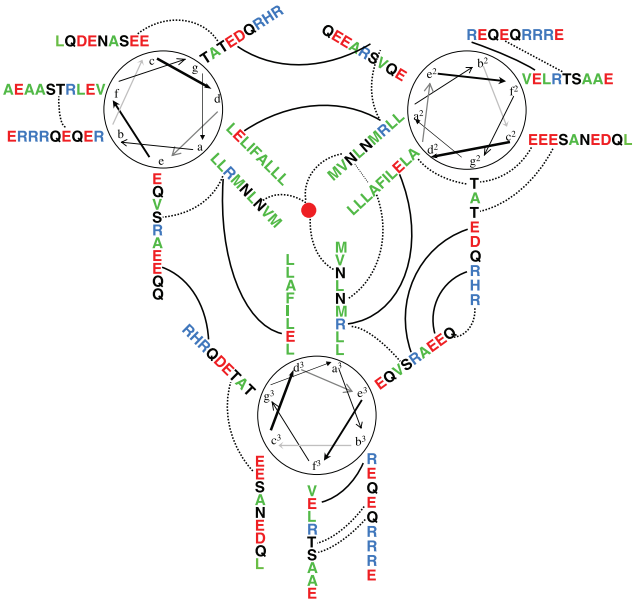



**Figure 3.** Intra- and interchain interactions in DMPK-CC. Green represents hydrophobic groups, red is negatively charged, blue is positively charged, and black, all other residues. The central red sphere represents an internal solvent atom located at the 3-fold axis and coordinating the three side chains of N511. Interactions are shown as connecting lines, where salt bridges are in black.

encountered value of −40°, as calculated using Twister; 44). This leads to overtwisting of the subsequent coil portion (Fig 2*c*) and an associated suboptimal hydrophobic core packing in that region. Thus, central coil widening and pitch parameters appear not to simply reflect crystal lattice restraints but to be an intrinsic structural feature of this CC domain.

Residue arrangements around the 3-fold axis of DMPK-CC are mostly symmetrical, an exception are the side chains of F500, an aromatic residue in position *d*. The presence of aromatic groups in core positions is extremely rare in dimeric CCs, but it is observed in trimeric motifs (45). For example, the trimeric coil of lung surfactant protein D (46) contains F225 and Y228 in *a* and *d* positions, respectively, which exhibit similar asymmetrical arrangements of their side chains. This is likely to reflect the difficulty of packing bulky groups in the limited interior of a trimeric coil. Remarkably, DMPK-CC also houses several solvent molecules in its core. In particular, its C-terminal section comprises an internal cavity filled by solvent that stabilizes a polar residue, N511, in *a* core position. A similar arrangement has been found in other trimeric CCs, such as bacteriophage T4 fibritin (47) and a domain of a coat protein from MoMuLVirus (48). Such buried polar groups in CCs are thought to contribute to the specificity of the motif and to maintain chain registry (49). Other relevant interactions in DMPK-CC (**Fig. 3**) include 6 interchain and 2 intrachain salt bridges. Interchain bridges help defining the association state of CCs, while intrachain bridges are important for helix stability (50). The interchain bridges of DMPK-CC belong both to the classical types *eg'*RE and *eg'*ER and to unusual core contacts of type *ad'*RE (notation as in ref 40; where ' designate residues in a subsequent heptad repeat). Intrachain salt bridges in DMPK-CC are of type *bf*RE. It is striking that the sequence of DMPK-CC is rich in charged residues (16E, 9R, vs. 2D, 0K per chain), although less than half of these are involved in interactions (8E, 4R, 0D). Thus, availability of reactive groups in the surface of this domain is high (predominantly of a negatively charged character and toward the $NH_2$ terminus of the domain), so that it could support electrostatic binding to other molecular components.

## Characterization of DMPK-CC in solution

In agreement with crystallographic data, the association state of DMPK-CC in solution is trimeric. One single oligomeric species could be detected using native-PAGE, size exclusion chromatography (**Fig 4*a***), and a sedimentation velocity study, which produced a single migrating boundary corresponding to a $S_{20W}$ of 2,1. Thus, the sample population in solution appeared to be highly homogeneous, with no dynamic states of assem-

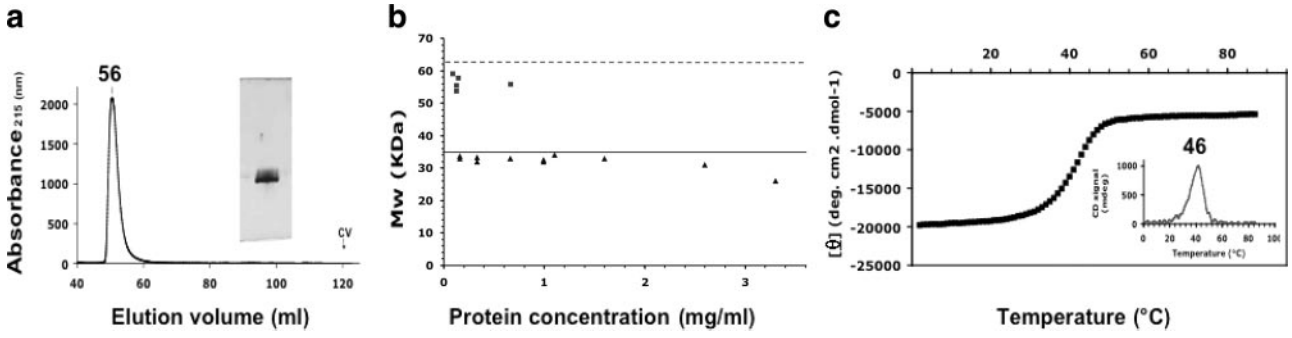



**Figure 4.** Biophysical characterization of DMPK-CC in solution. *a)* Size exclusion chromatogram in 50 mM Tris-HCl pH 7.5, 100 mM NaCl using a Superdex 75 Hi Load 16/60 column. A native-PAGE is shown in the inset; *b)* molecular mass calculated from sedimentation equilibrium measurements for DMPK-CC (triangles) and DMPK-$CC^{N+}$ (squares). Buffer conditions as in panel *a*. Theoretical mass values for trimeric formations are indicated by horizontal bars, where a continuous line corresponds to DMPK-CC and a dashed line to DMPK $CC^{N+}$. *c)* Thermal denaturation curve monitored by circular dichroism. First derivative is shown.

bly being observed. Sedimentation equilibrium analysis of DMPK-CC yielded an average molecular mass (MM) of ~33 kDa over the entire protein concentration range assayed (0.16–3.3 mg/ml), consistent with the single presence of trimers in the solutions (theoretical MM for a trimer of this construct 34.9 kDa; Fig 4*b*). Since the sample remained fully assembled at the lowest concentration assayed, the Kd of the interaction must be lower than 4 μM.

These studies carried out at moderate ionic strength (50–100 mM NaCl) and limited protein concentration are in good agreement with crystallographic data obtained from high salt solutions (2.0 M $Na^+/K^+$ tartrate) or polyethylene glycol media and concentrated protein samples (5–10 mg/ml), both at pH 4.6 and 7.5 (28). This suggests that the trimeric formations of DMPK-CC are stable in a broad range of media conditions. To gain further insight into the stability of this domain, it was investigated by thermal denaturation monitored by circular dichroism. A monophasic melting curve was recorded characteristic of a cooperative helix-coil transition of the three chains as a single domain, with a $T_{1/2m}$ of 46°C (Fig 4*c*). This measurement was carried out in 100 mM sodium phosphate buffer pH 7.0, 100 mM NaCl at a protein concentration of 0.6 mg/ml. To account for specific buffer effects, measurements were repeated in 50 mM bis-Tris propane pH 7.5, 100 mM NaCl with equivalent results. This indicates a strong association of the component chains and confirms the robustness of the assembly. Therefore, it can be concluded that DMPK-CC has a strong tendency to self-associate into well-defined trimeric formations.

### Oligomeric state of DMPK variants containing and lacking the CC domain

The assembly level of a DMPK construct comprising the LR sequence, the kinase domain, and the CC fraction, DMPKΔVR (Fig. 1), was investigated using multiangle light scattering coupled to size exclusion chromatography (SEC-MALS). This construct closely resembles an expressed DMPK spliceoform and is a suitable representative of this kinase in the cell. The study yielded MMs of 122′7 kDa and 124′5 kDa for protein concentrations estimated using UV-$A_{280}$ and refractive index, respectively. These values are in excellent agreement with the calculated MM of 120′2 kDa for a dimer of this sample (**Fig. 5**). This result is consistent with previous sedimentation equilibrium data on a comparable construct of this kinase (23).

To investigate the contribution of DMPK-CC to the dimerization of this kinase, a truncated variant lacking the CC fraction, DMPKΔCC/VR (Fig. 1), was examined. Data from size exclusion chromatography and native PAGE indicated the presence of a single species of DMPKΔCC/VR in solution, and SEC-MALS measurements confirmed that this corresponded to dimeric formations (Fig. 5). The experimental MM estimated for this sample was 101′4 kDa and 107′4 kDa using

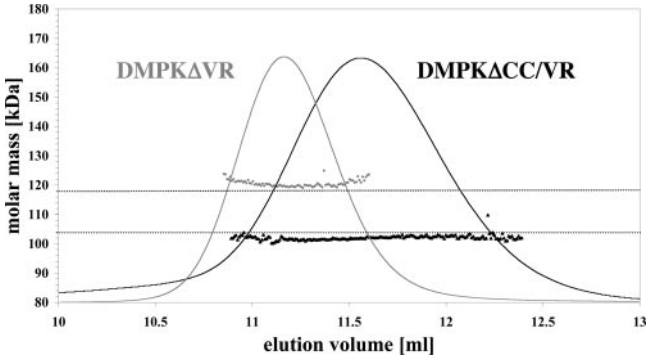



**Figure 5.** SEC-MALS characterization of DMPKΔVR and DMPKΔCC/VR. Molar mass (dots) and normalized UV signals are shown. The theoretical molar masses of the dimers of DMPKΔVR and DMPKΔVR/CC (120.0 kDa and 106.6 kDa, respectively) are indicated with horizontal lines.

UV-$A_{280}$ and refractive index to estimate the protein concentration, respectively. This is in close agreement with the theoretical MM of 106′6 kDa for a dimer of this construct. Furthermore, samples remained monodisperse and data were consistent through injected protein concentrations of 0.3–2 mg/ml, indicating that results are concentration independent (data not shown). The assembly state of DMPKΔVR coincides with that of DMPKΔCC/VR but not DMPK-CC, showing that the association level of this kinase is not dictated by the oligomerization properties of its CC domain and that segments associated with the kinase domain bear this function instead.

### Analysis of DMPKΔCC/VR dimers by small angle X-ray scattering

The kinase domain of DMPK is flanked by an N-terminal region of ~65 residue length (LR) and a C-terminal tail containing ~120-residues that links it to the CC domain (Fig. 1). Recent crystallographic data on ROCKI (25) show that, in that kinase, sequences both N- and C-terminal to the kinase domain pack together to form a dimerization domain. It has been shown that the single presence of the α-helical LR region is not sufficient to secure assembly in this case, so that the removal of the kinase C-terminal tail results in ~80% monomeric species (24). Given that ROCKI and DMPK share homology in the mentioned regions (22% and 36% seq. id. in the LR domain and kinase tail, respectively; **Fig 6*a***), we investigated whether DMPK might follow the same architectural principles of ROCKI by analyzing the overall conformation of DMPKΔCC/VR dimers using SAXS.

The experimental SAXS pattern of DMPKΔCC/VR is displayed in Fig 6*b*. The estimated MM, 95 ± 10 kDa, confirms that the protein is dimeric under the SAXS experimental conditions. This is further corroborated by the volume of the hydrated particle, 210 ± 10 × $10^3$Å (3), noting that, for globular proteins, the hydrated volume in $Å^3$ should be approximately twice the MM in daltons. The radius of gyration $R_g$ and the maximum particle size $D_{max}$ were 39.9 ± 0.6 Å and

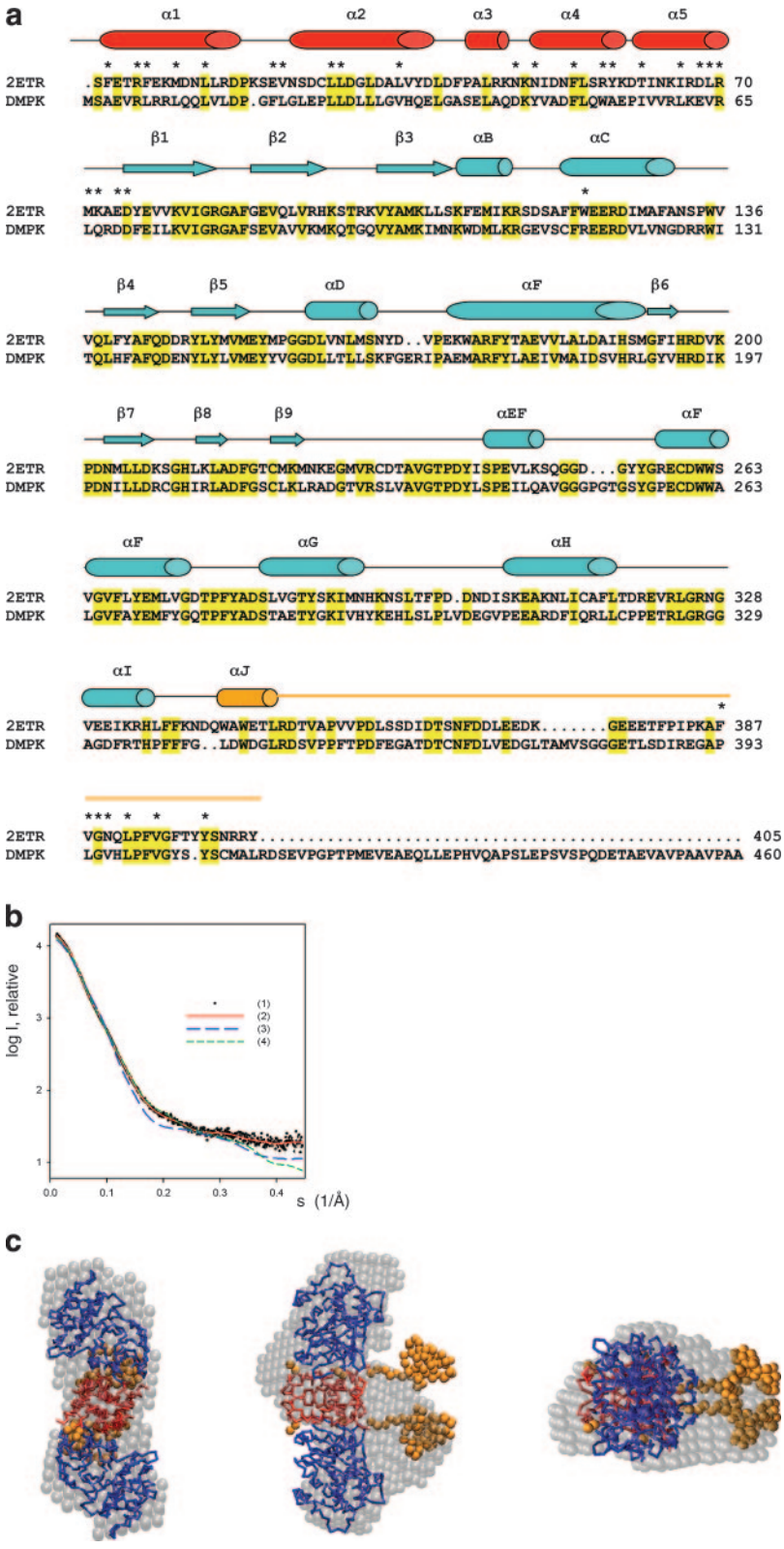



**Figure 6.** SAXS analysis of DMPKΔCC/VR dimers (*a*). Sequence similarity between DMPKΔCC/VR and the crystallographic model of ROCKI (PDB code 2ETR). Sequence identity is highlighted in yellow. Secondary structure elements are as in ref 25. Domain definition is as follows: red corresponds to the LR domain, blue to the kinase domain, and golden to the C-terminal tail. ROCKI residues involved in dimerization are marked with an asterisk. *b)* Experimental scattering of DMPKΔCC/VR (1) and calculated patterns from the *ab initio* low-resolution bead model (2), the ROCKI crystallographic model (3) and the reconstructed DMPKΔCC/VR model with an added C-terminal fraction (4); *c) ab initio* model of DMPKΔCC/VR (gray semitransparent beads) superimposed on the model constructed by adding C-terminal particles (orange beads) to the protomers of ROCKI (Cα-trace). Domain coloring as in *a*. Three orthogonal views are shown.

130 ± 10 Å, respectively, suggesting an elongated structure. This agrees with the low-resolution shape of DMPKΔCC/VR reconstructed *ab initio*, which fits neatly the experimental data with a discrepancy $\chi = 0.95$ (Fig 6*b*). A comparison of this with the crystal structure of ROCKI indicates that DMPKΔCC/VR has a similar overall conformation, also appearing to form head-to-head dimers (Fig 6*c*). The low-resolution model of DMPKΔCC/VR, however, displayed extra electron density that was not satisfied by the ROCKI structure, as can be expected from the fact that each protomer of DMPKΔCC/VR has 65 additional residues at its C terminus (Fig 6*a*). The missing residues account for 14% of the total scattering mass of this sample. They

play a significant role in the scattering of this molecule, possibly due to their contribution to an anisometric molecular shape (see below), so that the patterns computed from ROCKI yielded a poor fit to the experimental data ($\chi$=6.6; Fig 6*b*). To construct a model of DMPKΔCC/VR, the missing C-terminal portion was generated for each monomer using beads and maintaining a 2-fold symmetry as described in Materials and Methods. Although during the calculation the kinase domain was allowed to drift and rotate as a rigid body, its position was practically unchanged with respect to ROCKI. Models of DMPKΔCC/VR so generated showed that the C-terminal portions of each protomer formed rather compact formations, suggesting the existence of a protein domain at this position. A typical model of DMPKΔCC/VR in solution provides a good fit to the experimental data ($\chi$=1.75; Fig 6*b*) and agrees well with the molecular shape calculated *ab initio* as revealed by a comparative superimposition (Fig 6*c*). A deviation between models, however, shows that the volumes occupied by the C-terminal fraction are represented somewhat differently by the two modeling methods. It should be noted that the *ab initio* model was obtained without imposing any prior knowledge or symmetry restrictions, whereas the reconstructed model applies a 2-fold symmetry that may not be entirely adhered to by a potentially flexible C-terminal tail. Such flexibility could underlie localized deviations between otherwise consistent models and scattering patterns.

CC domains can be found in proteins in combination with other motifs that influence their association. The most prominent example is that of the trimeric CC of fibritin, which is followed by a C-terminal extension of ~30 residue length that folds into a β-hairpin (51). This intertwines with equivalent β-hairpins from other subunits, being essential to achieve CC trimerization in that case. Given that SAXS models of DMPKΔCC/VR suggested the presence of a structural formation immediately prior to the CC fraction, we assayed whether this might act as an association motif that could modulate CC stoichiometry. For this, we constructed an N-terminally extended version of DMPK-CC, DMPK-CC$^{N+}$ (Fig. 1), and analyzed its oligomeric state by sedimentation equilibrium (Fig 4*b*). Results indicated that DMPK-CC$^{N+}$ remained trimeric (estimated MM 56.1 kDa compared to a theoretical MM of 63.5 kDa for a trimer of this construct), suggesting that possible structural elements of the kinase-CC linker sequence do not influence the oligomerization of this coil.

### DMPK-CC is not involved in domain interactions

The CC fractions of ROCKs and MRCK have been shown to interact directly with their respective kinase domains, thereby acting as intrasteric inhibitors of kinase activity (26, 27). In ROCKs, the interaction involves the Rho binding site at the C terminus of its long CC domain, whereas in MRCK the binding has been mapped to a fragment spanning CC2/CC3, whose deletion resulted in about a 3-fold increase in activity. Given that DMPK-CC and MRCK-CC3 share certain sequence similarity (up to 31% identity according to isoforms) (Fig 2*d*), we investigated the possible interaction of DMPK-CC with other domains in this kinase using pull-down assays as well as coexpression/copurification experiments on DMPK-CC and DMPKΔCC/VR. Binding between these species could not be detected by either procedure, which yielded independent original species. Hence, DMPK-CC appears to neither interact with the kinase domain nor form a hetero-coil against the α-helical LR sequence. This indicates that a CC association/dissociation regulatory model such as that proposed for MRCK and ROCKs is unlikely to be applicable to DMPK and that the kinase and CC fractions component of DMPK are likely to be arranged as architecturally distinct fractions, resulting in a globular portion and a filamentous shaft.

## DISCUSSION

The CC domains of DMPK-related kinases are essential to the regulation of their catalytic activities. However, the molecular events mediated by these domains are yet to be accurately understood. CC sequences in this family exhibit large variability in length and composition, sharing little or no similarity. Only minimal structural data exist that aid to clarify the respective properties and functional roles of CC domains within this family. Here we report the crystal structure of DMPK-CC, which is unrelated to the ROCKs-CCs previously described (21, 22). Contrary to the dimeric character of the latter fragments, DMPK-CC forms a trimeric motif that is characterized by rigid termini and a central widening of its coil. It is currently unknown whether this feature might reflect functional properties of this domain. As proven by sedimentation equilibrium data, the assembly level of DMPK-CC in solution coincides with that in the crystal. No other oligomeric state has been detected in the current study under a broad range of protein concentration and ionic strength conditions. Furthermore, thermal denaturation studies have shown this assembly to be stable. Thus, our data prove that the oligomeric states of CC domains from this kinase family are more diversified than previously expected and that more coils must be investigated before general conclusions on their self-association properties can be drawn.

Experimental results on DMPK-CC contrast markedly with sequence-based predictions, which suggests this coil forms dimeric assemblies. For example, the software MultiCoil (43) estimates 82% probability of dimer vs. 10% trimer formation. The structural determinants of CC association remain poorly understood, and thus such predictions should be considered with caution. Only a few systematic analyses have addressed the basis of CC association experimentally, focusing on either the role of hydrophobic core residues (52, 53) or salt bridges (e.g., 54, 55). Consequently, cases are described where experimental data disagree with predictions (for

example, see ref 46). To understand the molecular basis of DMPK-CC trimerization, we analyzed its interchain contacts. The only factor that might be proposed at this time to favor formation of trimers over dimers is the presence of an aromatic residue, F500, at core position *a*, since this is rarely observed in dimeric coils (45). The contribution of the several salt bridges in DMPK-CC to its oligomerization cannot be evaluated based on current knowledge.

Given that CCs are established oligomerization motifs, these were initially thought to be indispensable for the molecular assembly of DMPK-related kinases. However, recent studies of ROCKs have shown that other domains mediate self-association in these kinases, namely, the N-terminal LR region and a tail C-terminal to the kinase domain (24, 25). Similarly, in the current study, a truncated DMPK variant lacking the CC moiety formed dimeric assemblies according to SEC-MALS data. A low-resolution structural analysis of these dimers using SAXS indicated that DMPK and ROCKI share similar dimerization principles involving head-to-head protomer interactions, even though a certain number of residues involved in the dimerization of ROCKI are not conserved in DMPK (Fig 6*a*). Our data on the truncated DMPK variant differ from those of a previous study (23), where a similar DMPK construct was argued to be monomeric according to sucrose density gradient sedimentation and pull-down assays. The dimeric form in our study (residues 1–460) contains 29 more residues at its C terminus than the construct previously reported. The truncation in that case might have disrupted a possible structural element C-terminal to the kinase domain, which has been insinuated by current SAXS data. More studies are required to identify the exact structural details of this segment of DMPK and to establish whether its integrity is essential for the assembly of this kinase. Contrary to that former study, we conclude that the CC domain of DMPK is not required for its self-association and that a common architectural design is likely to be shared by the closest members of this family: DMPK, MRCK, and ROCKs.

The fact that an interaction between DMPK-CC and other domains in DMPK could not be identified in the current study suggests it is unlikely that the assembly properties of this CC are altered by a possible embedding within the kinase core. Neither is the influence of flanking structural elements expected, given that an N-terminally extended CC construct retained the same oligomerization concentration. Thus, we speculate that the stoichiometry of this coil within the context of full DMPK must be dominated by the prior association of the N-terminal catalytic fraction into dimers. Prediction data confirm that the DMPK-CC sequence is geometrically compatible with a dimeric association, and it is known that CC oligomerization can be dynamically altered by local factors (20).

Oligomerization is crucial for the activation of numerous kinases. Often self-assembly is directed to ensure efficient autophosphorylation in *trans* by increasing the local concentration of catalytic domains (56). Given that DMPK-CC does not determine the self-association of this kinase and that, contrary to MRCK and ROCKs, it does not act as an intrasteric inhibitor but as an activator, further investigations are needed to clarify whether its function might relate instead to the achievement of an active conformation or even substrate docking. Moreover, given the heterogeneity in CC sequences and their diversified interaction patterns across DMPK-related kinases, studies are required to reveal the exact role of these moieties and to establish the principles of their regulatory mechanisms. FJ

Thanks go to Prof. B. Perryman for facilitating the initial DNA coding for DMPK. A special mention goes to R. Ravelli at ID14–2 (ESRF) for exceptional assistance during data collection. Thanks go to S. Strelkov for critical reading of this manuscript. Most sincere gratitude goes to Prof. A. Plückthun for generous support with SEC-MALS. We acknowledge the support of Swiss National Foundation (grant 3100A0–100852) and to P.G. of Association Française contre les Myopathies (grant 9812).

## REFERENCES

1. Fu, Y. H., Pizzuti, A., Fenwick, R. G., Jr., King, J., Rajnarayan, S., Dunne, P. W., Dubel, J., Nasser, G. A., Ashizawa, T., de Jong, P., et al. (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**, 1256–1258
2. Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H., Hunter, K., Stanton, V. P., Thirion, J. P., Hudson, T., et al. (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808
3. Nakagawa, O., Fujisawa, K., Ishizaki, T., Saito, Y., Nakao, K., and Narumiya, S. (1996) ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein serine/threonine kinase in mice. *FEBS Lett.* **392,** 189–193
4. Ishizaki, T., Maekawa, M., Fujisawa, K., Okawa, K., Iwamatsu, A., Fujita, A., Watanabe, N., Saito, Y., Kakizuka, A., Morii, N., and Narumiya, S. (1996) The small GTP-binding protein Rho binds to and activates a 160 kDa Ser/Thr protein kinase homologous to myotonic dystrophy kinase. *EMBO J.* **15,** 1885–1893
5. Wissmann, A., Ingles, J., McGhee, J. D., and Mains, P. E. (1997) Caenorhabditis elegans LET-502 is related to Rho-binding kinases and human myotonic dystrophy kinase and interacts genetically with a homolog of the regulatory subunit of smooth muscle myosin phosphatase to affect cell shape. *Genes Dev.* **11,** 409–422
6. Luo, L., Lee, T., Tsai, L., Tang, G., Jan, L. Y., and Jan, Y. N. (1997) Genghis Khan (Gek) as a putative effector for Drosophila Cdc42 and regulator of actin polymerization. *Proc. Natl. Acad. Sci. USA* **94,** 12963–12968
7. Di Cunto, F., Calautti, E., Hsiao, J., Ong, L., Topley, G., Turco, E., and Dotto, G. P. (1998) Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase encompassing the Rho-Rac-binding protein Citron. *J. Biol. Chem.* **273,** 29706–29711
8. Leung, T., Chen, X. Q., Tan, I., Manser, E., and Lim, L. (1998) Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector in promoting cytoskeletal reorganization. *Mol. Cell. Biol.* **18,** 130–140
9. Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuoka, T., Bito, H., Ishizaki, T., and Narumiya, S. (1998) Role of citron kinase as a target of the small GTPase Rho in cytokinesis. *Nature* **394,** 491–494
10. Piekny, A. J., and Mains, P. E. (2002) Rho-binding kinase (LET-502) and myosin phosphatase (MEL-11) regulate cytokinesis in the early *Caenorhabditis elegans* embryo. *J. Cell Sci.* **115,** 2271–2282
11. Kosako, H., Yoshida, T., Matsumura, F., Ishizaki, T., Narumiya, S., and Inagaki, M. (2000) Rho-kinase/ROCK is involved in cytokinesis through the phosphorylation of myosin light chain

and not ezrin/radixin/moesin proteins at the cleavage furrow. *Oncogene* **19,** 6059–6064
12. Jin, S., Shimizu, M., Balasubramanyam, A., and Epstein, H. F. (2000) Myotonic dystrophy protein kinase (DMPK) induces actin cytoskeletal reorganization and apoptotic-like blebbing in lens cells. *Cell Motil. Cytoskeleton* **45,** 133–148
13. Schulz, P. E., McIntosh, A. D., Kasten, M. R., Wieringa, B., and Epstein, H. F. (2003) A role for myotonic dystrophy protein kinase in synaptic plasticity. *J. Neurophysiol.* **89,** 1177–1186
14. Shimizu, M., Wang, W., Walch, E. T., Dunne, P. W., and Epstein, H. F. (2000) Rac-1 and Raf-1 kinases, components of distinct signaling pathways, activate myotonic dystrophy protein kinase. *FEBS Lett.* **475,** 273–277
15. Mounsey, J. P., Mistry, D. J., Ai, C. W., Reddy, S., and Moorman, J. R. (2000) Skeletal muscle sodium channel gating in mice deficient in myotonic dystrophy protein kinase. *Hum. Mol. Genet.* **9,** 2313–2320
16. Lee, H. C., Patel, M. K., Mistry, D. J., Wang, Q., Reddy, S., Moorman, J. R., and Mounsey, J. P. (2003) Abnormal Na channel gating in murine cardiac myocytes deficient in myotonic dystrophy protein kinase. *Physiol. Genomics* **12,** 147–157
17. Harper, P. S. (1989) *Myotonic Dsytrophy,* Saunders, London/Philadelphia
18. Davis, B. M., McCurrach, M. E., Taneja, K. L., Singer, R. H., and Housman, D. E. (1997) Expansion of a CUG trinucleotide repeat in the 3′ untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc. Natl. Acad. Sci. USA* **94,** 7388–7393
19. Jiang, H., Mankodi, A., Swanson, M. S., Moxley, R. T., and Thornton, C. A. (2004) Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* **13,** 3079–3088
20. Lupas, A. N., and Gruber, M. (2005) The structure of alpha-helical coiled coils. *Adv. Protein Chem.* **70,** 37–78
21. Shimizu, T., Ihara, K., Maesaki, R., Amano, M., Kaibuchi, K., and Hakoshima, T. (2003) Parallel coiled-coil association of the RhoA-binding domain in Rho-kinase. *J. Biol. Chem.* **278,** 46046–46051
22. Dvorsky, R., Blumenstein, L., Vetter, I. R., and Ahmadian, M. R. (2004) Structural insights into the interaction of ROCKI with the switch regions of RhoA. *J. Biol. Chem.* **279,** 7098–7104
23. Zhang, R., and Epstein, H. F. (2003) Homodimerization through coiled-coil regions enhances activity of the myotonic dystrophy protein kinase. *FEBS Lett.* **546,** 281–287
24. Doran, J. D., Liu, X., Taslimi, P., Saadat, A., and Fox, T. (2004) New insights into the structure-function relationships of Rho-associated kinase: a thermodynamic and hydrodynamic study of the dimer-to-monomer transition and its kinetic implications. *Biochem. J.* **384,** 255–262
25. Jacobs, M. D., Hayakawa, K., Swenson, L., Bellon, S. F., Fleming, M., Taslimi, P., and Doran, J. (2005) The structure of dimeric rock 1 reveals the mechanism for ligand selectivity. *J. Biol. Chem.* Oct 24; [Epub ahead of print]
26. Tan, I., Seow, K. T., Lim, L., and Leung, T. (2001) Intermolecular and intramolecular interactions regulate catalytic activity of myotonic dystrophy kinase-related Cdc42-binding kinase alpha. *Mol. Cell. Biol.* **21,** 2767–2778
27. Amano, M., Chihara, K., Nakamura, N., Kaneko, T., Matsuura, Y., and Kaibuchi, K. (1999) The COOH terminus of Rho-kinase negatively regulates rho-kinase activity. *J. Biol. Chem.* **274**, 32418–32424
28. Garcia, P., Marino, M., and Mayans, O. (2004) Crystallization and preliminary X-ray analysis of the coiled-coil domain of dystrophia myotonica kinase. *Acta Crystallogr. D* **60,** 2336–2339
29. La Fortelle, E., and Bricogne, G. (1997) *Methods Enzymol.* **276,** 472–494
30. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6,** 458–463
31. Jones, T.A, Zou, J.Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119
32. Murshudov, G. N., Vagin A.A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53,** 240–255
33. Collaborative Computational Project, N. 4 (1994) *Acta Crystallogr. D* **50,** 760–763
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291
35. Koch, M. H. J., and Bordas, J. (1983) X-ray diffraction and scattering on disordered systems using synchrotron radiation. *Nucl. Instrum. Methods* **208,** 461–469
36. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUSā Windows-PC based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36,** 1277–1282
37. Svergun, D. I. (1992) Determination of the regularization parameter in indirect transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25,** 495–503
38. Porod, G. (1982) Small angle X-ray scattering. (Glatter, O., and Kratky, O., eds) Academic Press, London
39. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76,** 2879–2886
40. Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) CRYSOLā program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28,** 768–773
41. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modelling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89,** 1237–1250
42. Petoukhov, M. V., Eady, N.A., Brown, K. A., and Svergun, D. I. (2002) Addition of missing loops and domains to protein models by x-ray solution scattering. *Biophys. J.* **83,** 3113–3125
43. Wolf, E., Kim, P. S., and Berger, B. (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci.* **6,** 1179–1189
44. Strelkov, S. V., and Burkhard, P. (2002) Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. *J. Struct. Biol.* **137,** 54–64
45. Walshaw, J., and Woolfson, D. N. (2001) Socket: a program for identifying and analysing coiled-coil motifs within protein structures. *J. Mol. Biol.* **307,** 1427–1450
46. Hakansson, K., Lim, N. K., Hoppe, H. J., and Reid, K. B. (1999) Crystal structure of the trimeric alpha-helical coiled-coil and the three lectin domains of human lung surfactant protein D. *Structure Fold Des.* **7,** 255–264
47. Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V., and Rossmann, M. G. (1997) Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. *Structure* **5,** 789–798
48. Fass, D., Harrison, S. C., and Kim, P. S. (1996) Retrovirus envelope domain at 1.7 Å resolution. *Nat. Struct. Biol.* **3,** 465–469
49. Lumb, K. J., and Kim, P. S. (1998) A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry* **34,** 8642–8648
50. Burkhard, P., Ivaninskii, S., and Lustig, A. (2002) Improving coiled-coil stability by optimizing ionic interactions. *J. Mol. Biol.* **318,** 901–910
51. Tao Y, Strelkov, S.V., Mesyanzhinov, V.V., and Rossmann, M. G. (1997) Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. *Structure* **5,** 789–798
52. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262,** 1401–1407
53. Gonzalez, L., Jr., Woolfson, D. N., and Alber, T. (1996) Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nat. Struct. Biol.* **3,** 1011–1018
54. Kammerer, R. A., Jaravine, V. A., Frank, S., Schulthess, T., Landwehr, R., Lustig, A., Garcia-Echeverria, C., Alexandrescu, A. T., Engel, J., and Steinmetz, M. O. (2001) An intrahelical salt bridge within the trigger site stabilizes the GCN4 leucine zipper. *J. Biol. Chem.* **276,** 13685–13688
55. Meier, M., Lustig, A., Aebi, U., and Burkhard, P. (2002) Removing an interhelical salt bridge abolishes coiled-coil formation in a de novo designed peptide. *J. Struct. Biol.* **137,** 65–72
56. Hubbard, S. R., and Till, J. H. (2000) Protein tyrosine kinase structure and function. *Annu. Rev. Biochem.* **69,** 373–398

*Received for publication December 3, 2005.*
*Accepted for publication February 3, 2006.*

# Annex A

## Studies on the oligomeric state of DMPK-CC$^{N+}$

**(Extended version of the DMPK-CC)**

# A.1 Introduction

DMPK-CC has been found to constitute a trimeric motif both in solution and in the crystalline state. On the contrary, the truncated DMPKΔCC/VR is a dimer according to SEC-MALS and SAXS. Since the full-length DMPK has been shown to be dimeric, the association of the whole system must be driven by a dimerization motif other than the CC domain.

Interestingly, CC domains can be found in proteins in combination with other motifs that influence their association. The most notorious example is that of the trimeric CC of fibritin, which is followed by a C-terminal extension of ~30 residues length, termed "foldon", that folds into a β-hairpin and serves as a registration motif for the triple CC domain of fibritin (Tao et al. 1997). This small foldon domain has a β-propeller conformation and can fold and trimerize autonomously (Frank et al. 2001). It has been reported that fibritins with a deleted or mutated foldon domain fail to fold correctly, whereas mutants with an altered CC motif but an intact foldon domain trimerize successfully (Letarov et al. 1999).

SAXS models of DMPKΔCC/VR suggested the presence of a yet unidentified domain in this kinase located immediately prior to its CC domain. In order to study whether a CC-associated domain might modulate the oligomerization of DMPK-CC, an N-terminally extended version of the latter, hereby termed DMPK-CC$^{N+}$ was created and its oligomeric state analyzed by sedimentation equilibrium.

The constructs employed in the current and the following studies are shown in Figure A.1.

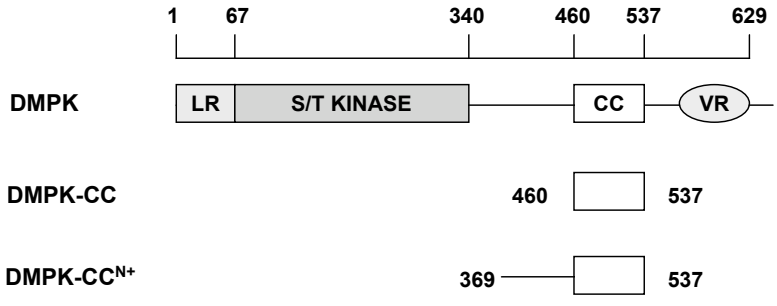



**Figure A.1: Schematic domain representation of human DMPK constructs used in this study.**
LR: leucine-rich region; CC: coiled-coil; VR: variable region are shown.

# A.2 Experimental Work

## A.2.1 DMPK-CC$^{N+}$ native

### A.2.1.1 Cloning

DMPK-CC$^{N+}$ construct comprises the CC domain and the residues within the linker region N-terminal to the CC. The C-terminal boundary of the kinase domain was determined by sequence alignment with homologous kinases, and a caspase cleavage site was found immediately after the kinase domain in the linker region at position 368. Therefore, in the expectation that the DMPK-CC$^{N+}$ construct would be stable in solution, its N-terminus was designed to start at this caspase cleavage site. Thus, DMPK-CC$^{N+}$ comprises residues 369-537, including the region that, according to modelling based on SAXS data, might allocate a putative structural element C-terminal to the kinase domain. An expression construct encoding DMPK-CC$^{N+}$ was derived from human DMPK complementary DNA (cDNA) (see Figure A.2).

The nucleotide sequence of the forward primer (N-terminus) was 5'-CGTTCGAAGACTACATGTTGGTGGAGGACGGGCTC- 3' with *Bbs*I restriction site, which creates a compatible overhang for *Nco*I in the vector pETM-11 (EMBL, Figure A.3). The reverse primer (C-terminus) was 5'-GGGGTACCTTACCCCGTGACAGCTGTGG- 3', which includes a *Kpn*I restriction site and was already employed to syntesize the DMPK-CC construct.

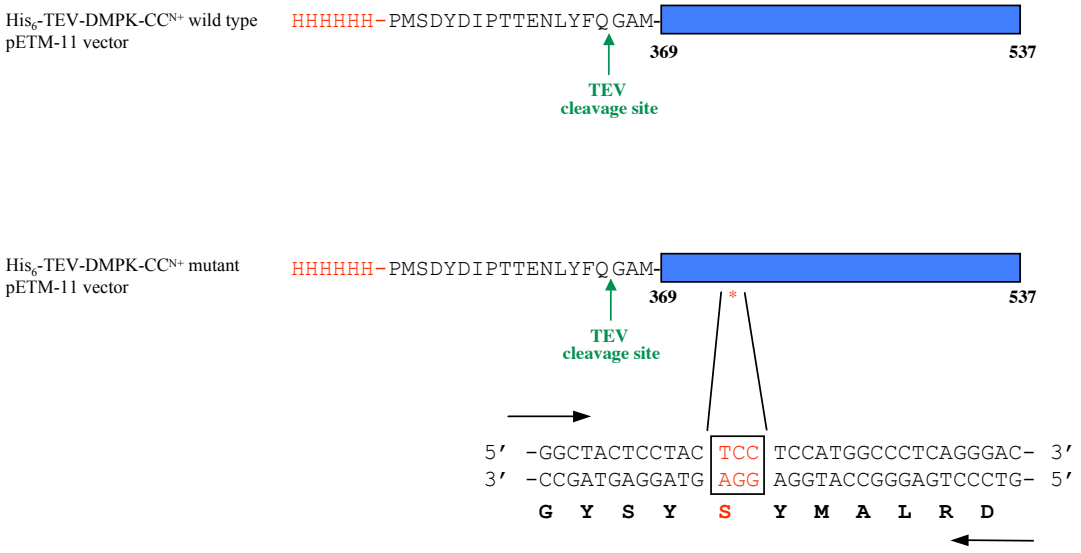



**Figure A.2: Design of DMPK-CC$^{N+}$ constructs.** Two constructs of DMPK-CC$^{N+}$ were produced. The first one corresponds to the wild-type, which contains the $His_6$-tag and a TEV protease cleavage site. The second one is a variant form of the former, where the cysteine at position 407 has been point mutated into a serine.

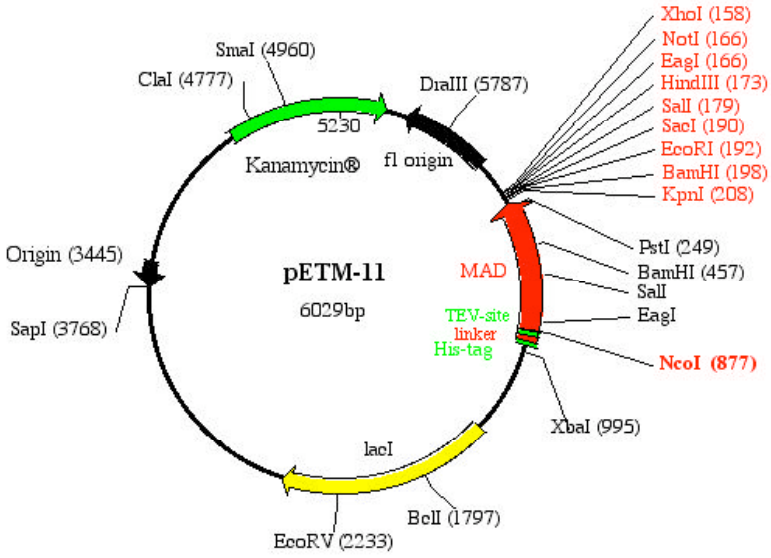



**Figure A.3: pETM-11 vector map.** The pETM-11 expression vector (Günter Stier; EMBL plasmid collection, Heidelberg, Germany) is a modified version of the pET24d with an upstream sequence coding for a hexahistidine-tag ($His_6$-tag) plus a TEV protease cleavage site followed by the non-essential gene of a protein called MAD. For construct design, the MAD insert was excised and replaced with the gene encoding the desired protein.

Polymerase chain reaction (PCR) was carried out with cDNA (150 ng) as a template, direct and reverse primers, and DNA polymerase (Expand Long Template PCR System, Roche) on a Tpersonal Biometra thermocycler. The PCR process consisted of the following steps: **(1)** *Initiation.* The mixture was heated at 95 °C for 5 minutes to ensure the melting of the DNA strands as well as the primers. **(2)** *Melting*. The double-stranded DNA is heated at 95 °C for 30 seconds in order to separate the strands. **(3)** *Annealing*. After separating the DNA strands, the temperature is lowered to 55 °C and incubated for 30 seconds, so the primers can attach themselves to the single DNA strands. **(4)** *Elongation*. Finally, the DNA-Polymerase has to copy the DNA strands. It starts at the annealed primer and works its way along the DNA strand. Elongation occurs by incubating at 72 °C for 1 minute. **(5)** Steps 2-4 are repeated 30 times. The PCR product was verified on an agarose gel as illustrated in Figure A.4.

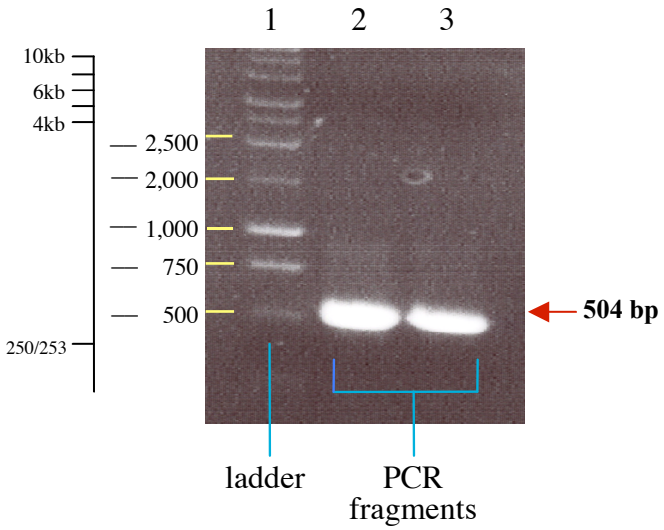



**Figure A.4: Verification of the PCR product on agarose gel**. Lane 1, 1kb DNA ladder (Promega). Lanes 2-3, agarose-gel purified PCR products with an expected MM of 504 base pairs (bp).

Subsequently, a PCR fragment of approximately 500 bp was gel-purified by using QIAquick Gel Extraction Kit (Qiagen) from an approximately pure sample and cloned cohesive-ended into the expression vector pETM-11. This modified version of pET-24d (Novagen) includes a $His_6$-tag plus a TEV (tobacco etch virus) protease-cleavage site at the N-terminus to the expression construct.

Ligation employed the "Fast Ligation" method (Promega). Conventional ligation requires a long incubation period (typically overnight) at temperatures ranging from

12-22 °C. The Fast Ligation provides an alternative method designed for the ligation of cohesive-ended DNA inserts into plasmid vectors in 5 minutes at room temperature (RT) with higher efficiency than traditional ligation procedures. This system, based on T4 DNA ligase, is designed to eliminate any further purification prior to transformation of ligated DNA.

The ligation mix (both complementarily digested PCR product and expression vector) was incubated with Fast Ligation mixture for 10 minutes at RT followed by heat-shock transformation into XLI-Blue cells at 42 °C for 90 seconds. The plasmid was constructed *in E. coli* strain DH5α.

DNA manipulations and bacterial transformations were carried out using standard protocols (Sambrook et al. 1989). The correct sequence of the clone was checked by DNA sequencing (Microsynth AG, Balgach, Switzerland).

### A.2.1.2 Over-expression

The plasmid coding for DMPK-CC$^{N+}$ was transformed in *E. coli* Rosetta (DE3) strain (Novagen). The cells were grown in Luria-Bertani broth (LB medium), supplemented with 34 μg/ml chloramphenicol and 25 μg/ml kanamycin at 37 °C. At an optical density (OD) of 0.6 at 600 nm, expression was induced by addition of 1mM isopropil-β-D-thiogalactopyranoside (IPTG) and cells were incubated overnight at 25 °C until an OD of 1.2 was reached. The cells were harvested by centrifugation at 18,000 rpm, for 30 minutes at 4 °C, washed with LB medium and stored at –80 °C.

### A.2.1.3 Purification

For product purification, the cells were resuspended in lysis buffer composed of 50 mM Tris pH 7.5, 100 mM NaCl and 5 mM β-mercaptoethanol (β-ME) supplemented with 40 mM imidazole, 1-5 μg/ml of DNAse, Phenylmethylsulphonylfluoride (PMSF), a spatula-tip of lysozyme and a standard protease inhibitor cocktail without the chelating agent ethylenediaminetetraacetic acid EDTA (Roche). Further cell lysis was performed by sonication in ice (Branson Digital Sonifier) using a number of short pulses (3-5 seconds) with pauses (7-10 seconds) to re-establish a low temperature. The lysate was harvested by centrifugation at 18,000 rpm for 30 minutes at 4 °C and the clarified supernatant was then purified by affinity chromatography on a His-Trap

HP chelating column (GE Healthcare) pre-equilibrated in lysis buffer. To discard the contaminant proteins the column was washed with lysis buffer followed by a gradient of salt washes, where lysis buffer had been progressively supplemented with 150, 200 and 250 mM NaCl. Each step consisted of eight column volumes, always in the presence of 40 mM imidazole. The bound protein was eluted by increasing the imidazole concentration to 200 mM as shown in Figure A.5A. The elution product was then dialysed (membrane 12-14,000 KDa cut off, Spectra/Por) in 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 5 mM β-ME overnight at 4 °C. Subsequent purification used size exclusion chromatography on a Superdex 200 Tricorn 10/300 GL column (GE Healthcare), equilibrated in 50 mM Tris pH 7.5, 100 mM NaCl and 5 mM β-ME. In the elution profile, two peaks were observed, which corresponded to native DMPK-CC$^{N+}$ and to lysozyme contaminant (according to mass spectroscopy), respectively (Figure A.5B). DMPK-CC$^{N+}$ eluted at 10.66 ml with an apparent MM comprised between 440-232 kDa, much higher that expected for a dimer (46.8 kDa) or a trimer (70.2 kDa), probably due to the anisotropic shape of the protein. This protocol yielded ~10 mg of pure protein per liter of culture. The sample was stored at 4 °C for further studies. Fractions corresponding to native DMPK-CC$^{N+}$ were pooled and subjected to $His_6$-tag cleavage. Tag removal used in-house produced TEV protease (0.7 mg/ml) (see TEV-protease production protocol in Annex C) in 50 mM Tris pH 8, 150 mM NaCl and 2 mM dithiothreitol (DTT) overnight at 4 °C. The digested protein was purified on a second subtractive metal-affinity step employing a His-Trap HP chelating 1 ml column (GE Healthcare). Tag-cleaved protein was concentrated to 5 mg/ml with a 10 kDa cut off concentrator (Millipore) and stored at 4 °C. The protein obtained was pure according to SDS-PAGE (Figure A.6) and mass spectroscopy. All purification procedures were executed at RT.

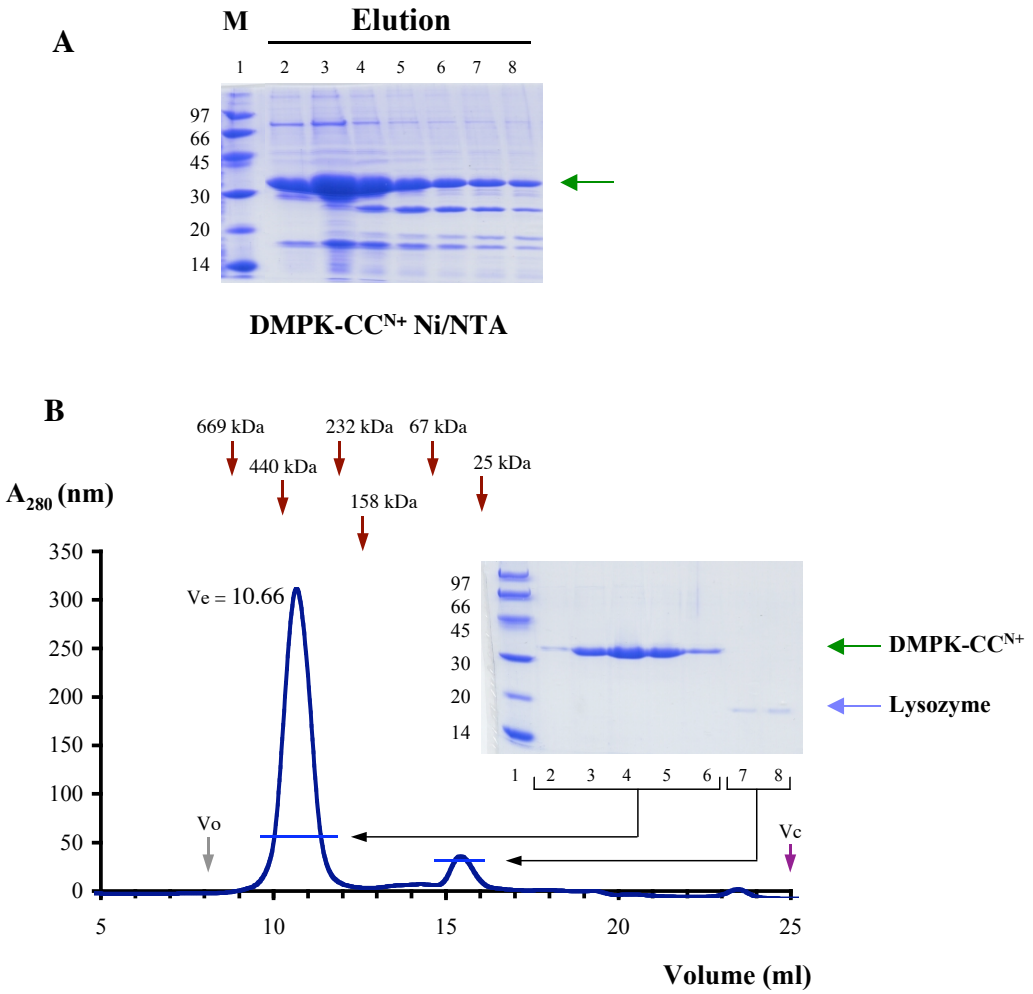



**Figure A.5: Purification of DMPK-CC$^{N+}$ wild-type. (A)** SDS-PAGE analysis of DMPK-CC$^{N+}$ native form (arrow) after a first purification using Ni-NTA chromatography. Lane 1, marker. Lanes 2-8, elution fractions. **(B)** Elution profile of DMPK-CC$^{N+}$ in a second purification step using size exclusion chromatography on a Superdex 200 Tricorn 10/300 GL column (GE Healthcare). Two peaks were observed, with the major one being DMPK-CC$^{N+}$ and the second peak lysozyme. The inset shows a SDS-PAGE gel of both peak fractions. Accordingly, fractions from DMPK-CC$^{N+}$ were pooled and used for further experiments. Positions of size standards thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), albumin (67 kDa) and chymotrypsinogenA (25 kDa) are indicated.

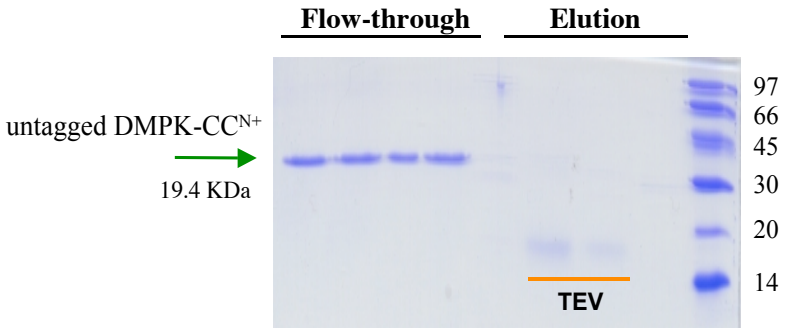



**Figure A.6: SDS-PAGE of a subtractive purification of untagged DMPK-CC$^{N+}$ protein.** After digestion with TEV protease, the purified untagged DMPK-CC$^{N+}$ protein, with an expected MM of 19.4 kDa, was recovered in the flow-through (arrow) in a subtractive $Ni^{2+}$-NTA affinity purification step.

### A.2.1.4 Protein determination

The protein concentration was determined by Bradford assay (Bradford 1976) with standard Bovine serum albumin (BSA) as standard (Pierce). The Bradford Protein Assay is based on the observation that the absorbance maximum for an acidic solution of Coomassie Brilliant Blue G-250 shifts from 465 to 595 nm when binding to protein occurs. The assay is useful since the extinction coefficient of a dye-albumin complex solution is constant a 10-fold concentration range. The assay can be performed in the presence of reducing agents such as DTT and β-ME.

The main disadvantage is that identical amounts of different standard proteins can cause considerable differences in the resulting absorption coefficients. In addition to this, the assay is non-linear over wide ranges, the dye binds strongly to quartz cuvettes and some proteins precipitate in the dye reagent. The assay should be used over narrow concentration ranges and plastic (or glass) cuvettes should be used.

In the current study, a calibration curve was constructed using standard BSA solution (2.0 mg/ml) in the range of 0.2 -1 mg/ml. Protein samples with concentrations ranging from 0.2-1 mg/ml were made up to 10 µl with distilled $H_2O$ in 1.5 ml eppendorf microcentrifuge tubes. 1 ml of Bradford reagent was then added. The solution was allowed to incubate at RT for 5 minutes before absorbances were read at 595 nm wavelength using absorption spectroscopy (Hewlett Packard Model 8453 UV-VIS system, Switzerland). Protein concentration estimations were calculated by interpolation from the calibration curve.

### A.2.1.5 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Sodium-Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis, or SDS-PAGE, is a technique used in biochemistry and molecular biology to separate proteins according to their size (length of polypeptide chain or MM). The solution of proteins to be analyzed is first mixed with SDS, an anionic detergent, which denatures secondary and non–disulfide–linked tertiary structures, and applies a negative charge to each protein in proportion to its mass.

Electrophoresis was carried out using 15% SDS-PAGE gels, according to the method of Laemmli (Laemmli 1970). Protein bands were visualized with Coomassie Brilliant Blue staining and the MM determination was performed referring to low molecular weight (LMW) SDS (GE Healthcare) as standards.

### A.2.1.6 Oxidation analysis

Proteins that contain thiol groups are susceptible to oxidation by free radicals, electrophilic molecules, and nitric oxide donors. Oxidation leads to the formation of disulfide bridges either within, or between polypeptidic chains. Particularly, cysteine is a highly vulnerable oxidative target because of the reactivity of the thiol group. Cysteines readily dimerize oxidatively to cystine. The tendency of proteins with free thiol groups to form artifactual disulfide-linked aggregates during and after purification is well known and can often be detected by comparing apparent size using reducing and non-reducing SDS-PAGE. Thus, in order to decipher whether DMPK-CC$^{N+}$ suffered from cysteine-oxidation upon storage, the protein was purified both in the presence and in the absence of reducing agent (β-ME).

Interestingly, one unique band was observed under reducing conditions, while two bands appeared when ß-ME was not employed (Figure A.7). Inter-chain oxidation in this case must be mediated by C407, the only cysteine residue present in the sequence of this construct.

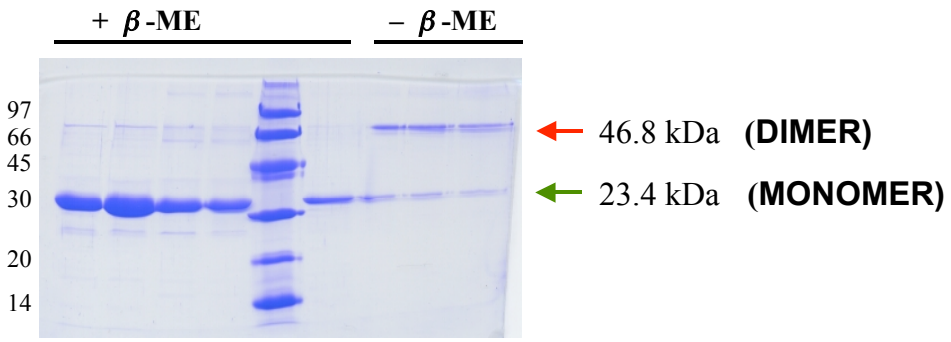



**Figure A.7: Oxidation test of DMPK-CC$^{N+}$.** To investigate the oxidation of DMPK-CC$^{N+}$, the protein was expressed and further purified by $Ni^{2+}$-NTA affinity chromatography, both in the presence and in the absence of the reducing agent β-ME. The presence of a second band at high MM under non-reducing conditions is indicated with a red arrow. Expected MMs for both monomeric and dimeric forms of DMPK-CC$^{N+}$ are also indicated.

### A.2.1.7 Crystallization

Several crystallization trials were attempted for both His-tagged and cleaved DMPK-CC$^{N+}$ native proteins using the hanging-drop vapour-diffusion method (Suzuki et al. 1993) using 24-well VDX crystallization plates (Hampton Research) at RT. Each drop contained 2 µl of protein solution (5 mg/ml and 1 mg/ml of His-tagged or cleaved DMPK-CC$^{N+}$, respectively) and an equal amount of crystallization reservoir solution, equilibrated against 1 ml reservoir solution. Initial screening was performed using Sparse matrix screens (kits I and II from Molecular Dimensions) based on the procedure of Jancarik & Kim (Jancarik 1991). In parallel, crystallization conditions that yielded crystals for DMPK-CC were assayed. No crystals were obtained for DMPK-CC$^{N+}$ native under any of the protein concentrations and crystallization conditions tested.

## A.2.2 DMPK-CC$^{N+}$ mutant

### A.2.2.1 Cloning

In order to prevent protein aggregation mediated by cysteine oxidation, *in vitro* site-directed mutagenesis was carried out (QuikChange Site Directed Mutagenesis Kit, Stratagene). The cysteine at position 407 was point mutated into a serine, creating a mutant hereon called pETM-11:DMPK-CC$^{N+}$C407S (Figure A.2). Serine was chosen because its volume and composition are very similar to those of cysteine, but instead of a reactive sulfhydryl it contains a hydroxyl group. The pETM-11:DMPK-CC$^{N+}$ wild-type plasmid was used as a template DNA and the two mutagenic primers utilized to amplified the mutated DMPK-CC$^{N+}$ are as follows:

```
5′ -GGCTACTCCTACTCCTCCATGGCCCTCAGGGAC- 3′
3′ -CCGATGAGGATGAGGAGGTACCGGGAGTCCCTG- 5′
```

Positive clones pETM-11:DMPK-CC$^{N+}$C407S were verified by DNA sequencing (Microsynth AG, Balgach, Switzerland). Primers were purchased from Microsynth and used after dissolving in distilled water.

### A.2.2.2 Over-expression and purification

Expression of DMPK-CC$^{N+}$C407S was as for the wild-type (see section A.2.1.2). Purification was performed in a single-step purification using a $Ni^{2+}$-NTA affinity chromatography column. The protein obtained with this procedure was pure according to SDS-PAGE and was used in subsequent crystallization trials. $His_6$-tag removal was as for the wild-type (see section A.2.1.3). Pure tag-cleaved protein was then concentrated to 20 mg/ml at 4 °C and was also used in subsequent crystallization trials. Protein procedures were executed at RT and protein concentration was measured employing the BCA method (Pierce) (see section A.2.2.3). As for the wild-type form, this protocol yielded ~10 mg of cleaved pure protein per liter of culture.

Despite size exclusion chromatography was not required for purification of DMPK-CC$^{N+}$C407S, comparison of the elution profile of DMPK-CC$^{N+}$C407S with that of the wild-type over a size exclusion column showed that the mutant eluted from the

column with a retention time similar to the wild-type protein. This suggested that the quaternary structure of the wild-type protein is unaltered in the mutant.

### A.2.2.3 Protein determination

Protein concentration was determined by bicinchoninic acid assay (also known as BCA assay). The BCA assay is a biochemical procedure for detecting the presence of protein in a solution that changes colour from green to purple in proportion to protein concentration in a given sample, which can then be measured using colorimetric techniques.

There is no perfect assay method for determining protein concentration. Each of the many methods has numerous advantages and disadvantages. Accordingly, the BCA assay is a popular choice for the assay of protein concentration due to its high sensitivity, compatibility with detergents and with a wide range of common buffer components and its overall simplicity. Although the BCA assay requires more time than the Bradford technique, BCA has some advantages over the former protein determination method besides those mentioned above; it has less variability, the color complex is stable and it is applicable over a broad range of protein concentration.

DMPK-CC$^{N+}$C407S protein concentration was estimated using the BCA assay with BSA as protein standard, following the manufacturer's protocols (Pierce). This procedure employs two reagents: A) a stock BCA solution containing 1 g bicinchoninic acid (BCA), 2 g sodium carbonate, 0.16 g sodium tartrate, 0.4 g NaOH, and 0.95 g sodium bicarbonate per 100 ml of aqueous solution and B) 0.4 gm cupric sulfate (5 x hydrated) in 10 ml distilled water. The BCA assay relies on two reactions. First, the $Cu^{2+}$ ions, BCA and the peptide bonds or protein residues such as cysteine, cystine, tryptophan and tyrosine, form a complex, which reduces $Cu^{2+}$ to $Cu^{+1}$. The amount of $Cu^{2+}$ reduction is proportional to the amount of protein present in the solution. Then, the $Cu^{+1}$ ion and BCA form a complex that has a purple-blue color and strong absorption at 562 nm. The protein concentration can be quantified by measuring the absorption spectra and comparing with protein solutions with known concentrations.

The BCA assay was carried out as follows: protein sample with concentrations ranging from 0.05 – 0.25 mg/ml were made up to 50 µl 5 mM Hepes pH 7.2 buffer

supplemented with 1% SDS, in 1.5 ml eppendorf microcentrifuge tubes. 1 ml of working reagent, composed of 1part of reagent B and 50 parts of reagent A, was then added. The solution was incubated at 60 °C for 30 minutes, and absorbances were measured at 562 nm wavelength using absorption spectroscopy (Hewlett Packard Model 8453 UV-VIS system, Switzerland). Protein concentrations were estimated by interpolation from the calibration curve.

### A.2.2.4 Characterization of DMPK-CC$^{N+}$ mutant in solution

#### A.2.2.4.1 *Blue-Native Polyacrylamide Gel-Electrophoresis (BN-PAGE)*

Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE) is a charge shift method, in which proteins run on PAGE in the absence of SDS and are separated according to their mass-charge ratio.

To assess the different oligomerization states possible for DMPK-CC$^{N+}$C407S, a native gel analysis was performed, followed by Coomassie staining for protein. The separation was performed on a gradient gel (10-16%), in the presence of blue cathode buffer (15 mM Bis-Tris-HCl pH 7.0, 50 mM Tricine and 0.005 % Coomassie Brilliant Blue G-250) and anode buffer (50 Bis-Tris-HCl pH 7.0). Special standard markers (ferritin 440 KDa, catalase 332 KDa, aldolase 158 kDa, albumin 66 kDa and ovalbumine with 43 kDa) were used for identification. The experiment was conducted overnight at 15 mA or 100 V. BN-PAGE revealed a single band with an apparent molecular mass (MM) of ~140-150 kDa, as illustrated in Figure A.9. Remarkably, this value was much higher than expected for a dimer or a trimer. This was in accordance with results obtained from size exclusion chromatography. Therefore, it can be concluded that DMPK-CC$^{N+}$C407S exists as a single species, similar to DMPK-CC$^{N+}$, however, MM cannot be assessed probably due to the anisometric dimensions of the molecule.

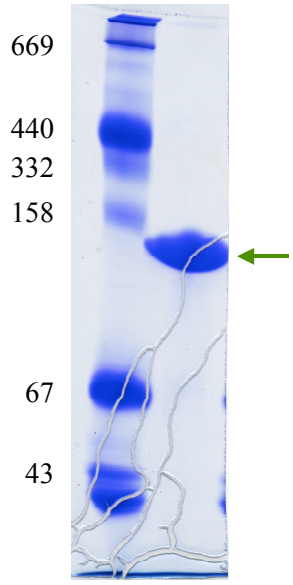



**Figure A.9: Blue Native-PAGE analysis (10-16%) of purified DMPK-CC$^{N+}$ sample.** The presence of a single species is indicated (arrow). MM markers from top: thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), bovine serum albumine (67 kDa) and ovalbumine (43 kDa).

### A.2.2.4.2 *Analytical Ultracentrifugation*

Theodor Svedberg invented the analytical ultracentrifuge in 1923, and won the Nobel Prize in Chemistry in 1926 for his research on colloids and proteins using the ultracentrifuge. Analytical ultracentrifugation (AU) is a versatile and rigorous technique that provides quantitative information about the sizes and shapes of macromolecules in solution. Two types of analytical centrifugation experiments exist, which are very complementary: sedimentation velocity and sedimentation equilibrium.

***Sedimentation velocity*** measures how fast the macromolecules move in response to centrifugal force. The sedimentation boundary movement gives information about mass and shape of macromolecules.

Sedimentation velocity experiments are performed at high-speed rotation (30,000 - 50,000 rpm) and this type of experiment can determine the following: sample homogeneity (or heterogeneity) in protein mass and conformation, sedimentation coefficients of solute, presence of aggregates and their size distribution, overall shape/assymetry of macromolecules, MM and stoichiometry.

***Sedimentation equilibrium*** is based on the balancing of centrifugal force and diffusion of the particle in order to form a solute gradient, which remains constant for an indefinite period of time. This experiment is usually performed at a lower rotor speed and takes several days to complete. This technique allows determination of the following: MM, binding constants, oligomeric state, and the non-ideality of the solution.

To assess the oligomerization state and homogeneity of untagged DMPK-CC$^{N+}$C407S in solution, analytical ultracentrifugation experiments were carried out. Sedimentation velocity and equilibrium data were recorded using an Optima XL-A analytical ultracentrifuge (Beckman Instruments) equipped with 4 and 12-mm Epon double-sector cells in an An-60 Ti rotor and absorption optics. Runs were performed at 20 $^{o}$C in 50 mM Tris pH 7.5, 50 mM NaCl using a detection wavelength of 230 nm. Sedimentation velocity experiments were carried out at 50,000 rpm and 0.67 mg/ml sample concentration for checking monodispersity (Figure A.10). Sedimentation equilibrium measurements were performed at different speeds (13,000, 15,000 and 24,000 rpm) and protein concentrations of 0.09-1.4 mg/ml, where OD values remained in a linear range, to analyze the self-association state of DMPK-CC$^{N+}$C407S (Figure A.11). Average MMs were evaluated using SEGAL (G. Machaidze & A. Lustig, unpublished software), which adjusts the baseline absorbance to obtain the best linear fit of ln (absorbance) versus radial distance square. A partial specific volume of 0.73 ml/g and a solution density of 1.005 g/ml were used; for the conversion of S' to S20w, the solution viscosity was taken as 1.31 centipoise.

In agreement with the results obtained by BN-PAGE, one single oligomeric species could be detected using the sedimentation velocity technique, which produced a single migrating boundary corresponding to a sedimentation coefficient $S_{20W}$ of 1.9. Thus, the sample population in solution appeared to be highly homogeneous, with no dynamic states of assembly being observed. Sedimentation equilibrium analysis of untagged DMPK-CC$^{N+}$C407S yielded an average MM of ~56 kDa over the entire protein concentration range assayed (0.09-1.4 mg/ml), compared to a theoretical MM of 58.2 kDa for a trimer of this construct (Figure A.12).

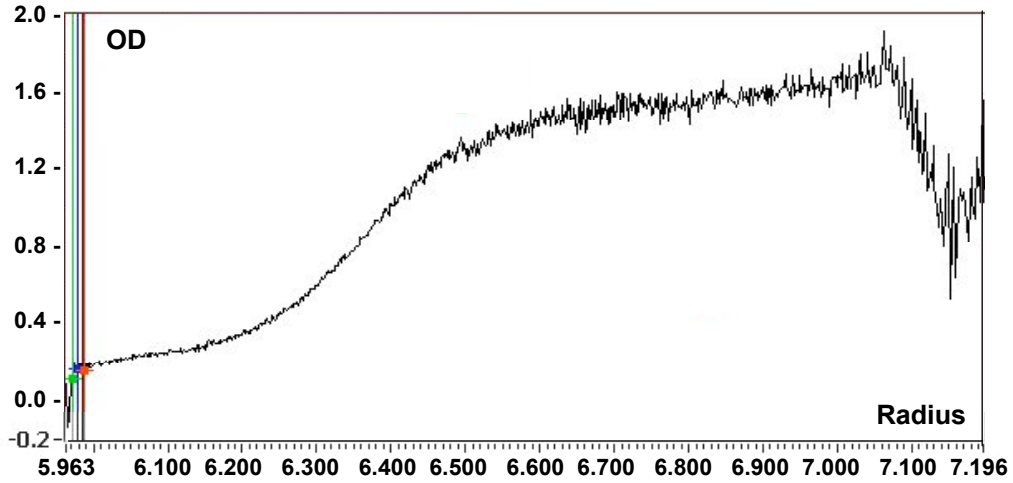



**Figure A.10: Sedimentation velocity profile of DMPK-CC$^{N+}$.** Sedimentation velocity revealed a monodisperse species ($S_{20,w}$=1.9 S). Measurements were performed at 20 °C in 50 mM Tris pH 7.5, 50 mM NaCl with a loading concentration of 0.67 mg/ml using a detection wavelength of 232nm. Runs were carried out at a rotor speed of 50,000 rpm for 190 minutes. The direction of the sedimentation is from left to right.

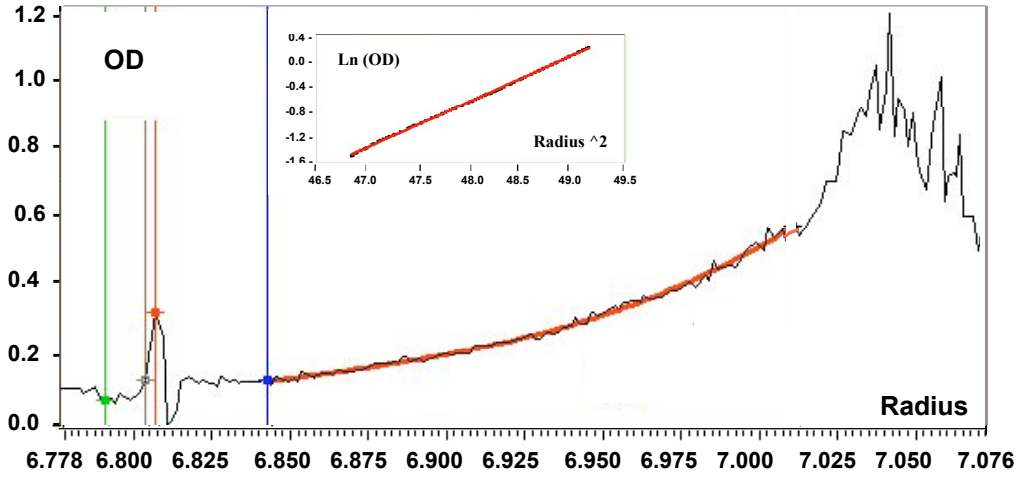



**Figure A.11: Sedimentation equilibrium ultracentrifugation of DMPK-CC$^{N+}$.** Experiments were performed at a rotor speed of 15,000 rpm in buffer 50 mM Tris pH 7.5, 50 mM NaCl at 20 °C with loading concentrations of 0.13 mg/ml, monitoring absorbance at 231 nm. The natural logarithm of the absorbance at 280 nm is plotted against the square of the radial position as shown in the inlet.

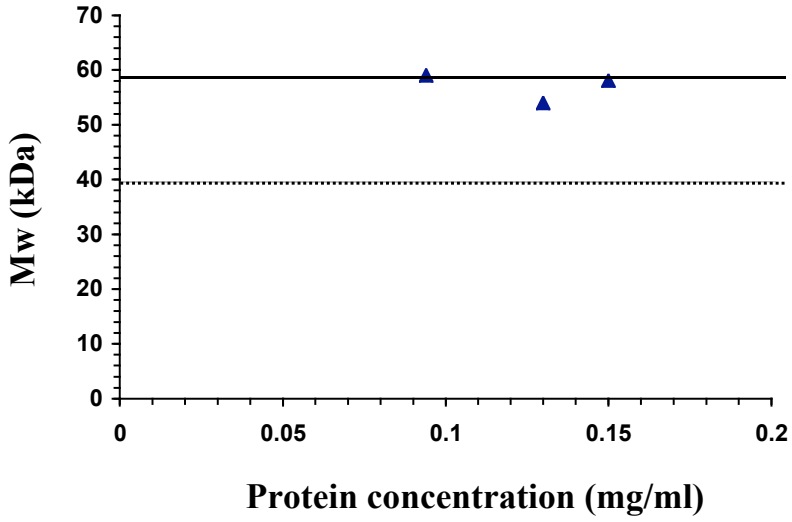



**Figure A.12: Sedimentation equilibrium experiments for DMPK-CC$^{N+}$.** MM calculated from sedimentation equilibrium experiments on untagged DMPK-CC$^{N+}$ (triangles). Experiments were performed in buffer 50 mM Tris pH 7.5, 50 mM NaCl at 20 °C and rotor speeds of 13,000 rpm with loading concentrations ranging from 0.1 to 0.27 mg/ml. Runs were carried out using a detection wavelength of 230 nm Theoretical mass values for dimeric and trimeric formations of DMPK-CC$^{N+}$ are indicated by an horizontal bar, where a continuous line corresponds to the trimer and a dashed line to the dimer.

## A.2.3 Crystallization

Several crystallization trials were undertaken for tag-cleaved DMPK-CC$^{N+}$C407S protein utilizing the hanging-drop vapour-diffusion method (Suzuki et al. 1993) using 24-well VDX plates (Hampton Research) at RT. Each drop contained equal volumes (2 µl) of protein solution and crystallization reservoir solution, equilibrated against 1 ml reservoir solution. Protein concentration varied from 3 to 20 mg/ml. A wide variety of crystallization conditions were tested using several commercially available sparse matrix screens. Sparse matrix screens employed include kits I and II from Hampton Research and Molecular Dimensions, Wizard and Cryo screens from Emerald Biostructures as well as systematic polyethylene glycol (PEG)/Ion screens, which screens a complete profile of anions and cations in the presence of PEG 3350 with pH 4.5 - 9.2 (Hampton Research). Besides this, The Quik Screen (Hampton Research) was used to test 0.8 - 1.8 M Sodium/Potassium Phosphate versus pH 5 - 8.2 in a grid format. In addition, Initial crystallization was also attempted using crystallization conditions as those of DMPK-CC. No crystal growth was observed for DMPK-CC$^{N+}$C407S under any of the protein concentrations and crystallization conditions tested.

## A.3 Discussion

In order to study whether the presence of a CC-associated domain in DMPK might modulate the oligomerization of its CC domain, as suggested by reported SAXS experiments on DMPKΔCC/VR, an N-terminally extended version of DMPK-CC, named DMPK-CC$^{N+}$ was created and its oligomeric state analyzed by sedimentation equilibrium. Crystallization was attempted in a wide range of crystallization conditions for DMPK-CC$^{N+}$ but no crystals were obtained under any of the conditions tested. A possible explanation may be that the protein suffered from aggregation due to oxidation of the sample, degradation or even additional residues N-terminal to the CC might have also created a tail disorder, thus, preventing protein from crystallization.

The presence of oxidation in the sample was assessed by expressing DMPK-CC$^{N+}$ both in the presence and in the absence of a reducing agent. As a result, one unique band was observed when the protein was expressed under reducing conditions, while two bands appeared when β-ME was not employed, indicating oxidation in the sample. In order to prevent oxidation, the C407, the only present in the DMPK-CC$^{N+}$ sequence was point mutated into a serine, creating the DMPK-CC$^{N+}$ mutant termed DMPK-CC$^{N+}$C407S.

Analysis of DMPK-CC$^{N+}$C407S using native PAGE and sedimentation velocity techniques yielded one single oligomeric species, indicating the sample population in solution to be highly homogeneous. Furthermore, sedimentation equilibrium analysis of DMPK-CC$^{N+}$ mutant yielded and average MM of 56 kDa over the entire protein concentration range assayed (0.09-1.4 mg/ml), compared to the theoretical MM of 58.2 KDa for a trimer of this construct. Contrary to expectations, results showed that DMPK-CC$^{N+}$ mutant, in accordance with DMPK-CC, remained trimeric in solution and no additional oligomerization motif was identified. However, it might be possible that such motif existed but needed the presence of additional residues or even the whole kinase to be structurally or functionally integral. Crystallization of mutant DMPK-CC$^{N+}$C407S yielded no crystal form.

The nature of this region of DMPK will be revealed by structural data on full-length DMPK. Nevertheless, crystallization of the full-lengh DMPK construct has been attemped by Zöhre Ucurum in our laboratory, so far without success.

Moreover, it is known that CC oligomerization can be altered by enviromental conditions, such as pH and temperature. For example the TlpA protein of *Salmonella* encodes a repressor protein that makes use of the CC motif to sense temperature changes. At temperatures below 37 °C, TlpA forms a dimer that can bind DNA and act as a repressor. Whereas temperatures above 37 °C cause the CC dimerization domain to unfold. The random coiled monomers are unable to remain bound to DNA, lifting the repression of the promoter (Hurme et al. 1996; Hurme et al. 1997). An example for pH regulation of a CC dimer is the apoptosis-associated protein Par-4. The C-terminal region of Par-4, required for apoptotic activity, self-association and association with all presently known interaction partners, contains a CC domain (Diaz-Meco et al. 1996; Sells et al. 1997). The protein forms a CC dimer only at an acidic pH and low temperature. Through site-directed mutagenesis predicting the effects of unfavorable charged residues at higher pH, Par-4 was altered from a pH-dependent, labile dimer into a pH-independent, stabile dimer (Dutta et al. 2003).

Oligomerization is crucial for the activation of numerous kinases. Often, self-assembly is aimed to ensure efficient autophosphorylation in trans by increasing the local concentration of catalytic domains (Hubbard and Till 2000). Given that DMPK-CC does not mediate the self-association of this kinase and that, contrary to MRCK and ROCKs, it does not act as an intrasteric inhibitor but as an activator, further investigations are needed to clarify whether its function might relate instead to the achievement of an active conformation or even substrate docking. Supporting this idea, recent studies on a DMPK variant carrying a mutated CC domain, showed that it was still capable of autophosphorylation and transphosphorylation of peptides, but the rate of its kinase activity was significantly lowered (van Herpen et al. 2006). A possible explanation for this could be that distortion of the CC structure affects conformational flexibility in the kinase domain itself and that this feature is needed for efficient intramolecular autophosphorylation (van Herpen et al. 2006). Moreover, binding of the DMPK to MYPT2, a paralogue of the DMPK natural substrate MYPT1, was strongly reduced, proposing that the DMPK-CC motif may be important in strengthening the binding between DMPK and MYPT2, and bringing protein sequences, crucial for cross-link formation, in close proximity (van Herpen et al. 2006).

In conclusion, the identification of the additional domain has not taken place, but if

such domain existed, it might not be sufficient to modulate the oligomerization properties of DMPK-CC.

# Annex B

## Analysis of the interactions between DMPK-CC and other DMPK domains

# B.1 Introduction

The functionality of CC domains from the DMPK-related family of kinases remains unclear. Apart from their multimerization properties, different biochemical roles have been attributed to these moieties. CC motifs of ROCKs and Citron kinase have been found to act as intrasteric inhibitors of their respective kinase domains and harbour at their C-termini binding motifs for small Rho GTPases, where binding of the latter relieves the inhibition (Di Cunto et al. 1998; Amano et al. 1999). Whereas CC segments of MRCK (CC2/CC3) also act as intrasteric inhibitors but do not host a GTPase binding region (Tan et al. 2001) and physiological release factors are yet to be identified (see Figure 1.5 in section 1.4.3).

Full-length ROCK, in the absence of effector molecules, exists in an autoinhibited state. The C-terminal region of ROCK, encompassing CC and PH domains, has been shown to partially inhibit the kinase catalytic activity by binding directly to the kinase domain (Amano et al. 1999). When GTP-bound RhoA interacts with to the Rho-binding region of the CC domain, the contacts between the catalytic kinase domain and the C-terminal region are disrupted, relieving the inhibition (Ishizaki et al. 1996). Cleavage of the C-terminal inhibitory domain from the catalytic domain by caspase-3 during apoptosis also activates the enzyme (Coleman et al. 2001; Sebbagh et al. 2001). An analogous form of truncated Citron kinase has also been observed *in vivo* (Di Cunto et al. 98). It has been reported that arachidonic acid cancels the inhibitory effect of RB/PH region, and activates ROCK *in vitro* (Feng et al. 1999).

Unlike ROCK, the regulation of MRCK activity does not appear to involve the PH domain; instead, a small region within the extensive MRCK CC region, located at the end of the distal CC2/CC3 domains, termed KIM (KI motif), acts as a negative autoregulatory domain by complexing to and inhibiting the catalytic domain (Tan et al. 2001; Ng et al. 2004). Activation of the kinase was observed upon binding of phorbol ester to the neighbouring CR domain, presumably by releasing the constraint of the inhibitory effect on the catalytic activity. Co-expression of MRCKγ KIM is sufficient to inhibit the activity of either MRCKα or MRCKγ (Ng et al. 2004).

On the contrary, the CC domain of DMPK (DMPK-CC) has been proposed to act as an enhancer of its kinase activity (Zhang and Epstein 2003), although it is still unknown whether this involves transphosphorylation events as those occuring in

tyrosine kinases. Recent studies using a DMPK form where its CC was mutated into a segment with low coil-forming probability, showed that autophosphorylation and the transphosphorylation activity towards a peptide substrate, KKRNRRLTVA (Wansink et al. 2003), was 2-3-fold reduced when compared with the corresponding activities of wild-type DMPK, suggesting that the DMPK-CC region must have a facilitating role in the determination of DMPK activity (van Herpen et al. 2006).

The trimeric assembly of DMPK-CC differs from that of the full-length kinase, which is dimeric. Dynamic alterations in the oligomeric state of DMPK-CC might be explained by embedding of the CC motif within the kinase core. Therefore, given that DMPK-CC shares certain sequence similarity with the CC domains CC2/CC3 of MRCK (up to 30 % identity according to isoforms), a similar regulatory model to that proposed for MRCK might be applicable for DMPK (Figure B.1). Therefore, to assess whether the DMPK-CC is an autoregulatory region, the ability of the CC domain to interact with other domains within the kinase was investigated. Binding studies of DMPK-CC and a truncated DMPK variant lacking the CC domain, DMPKΔCC/VR, were performed.

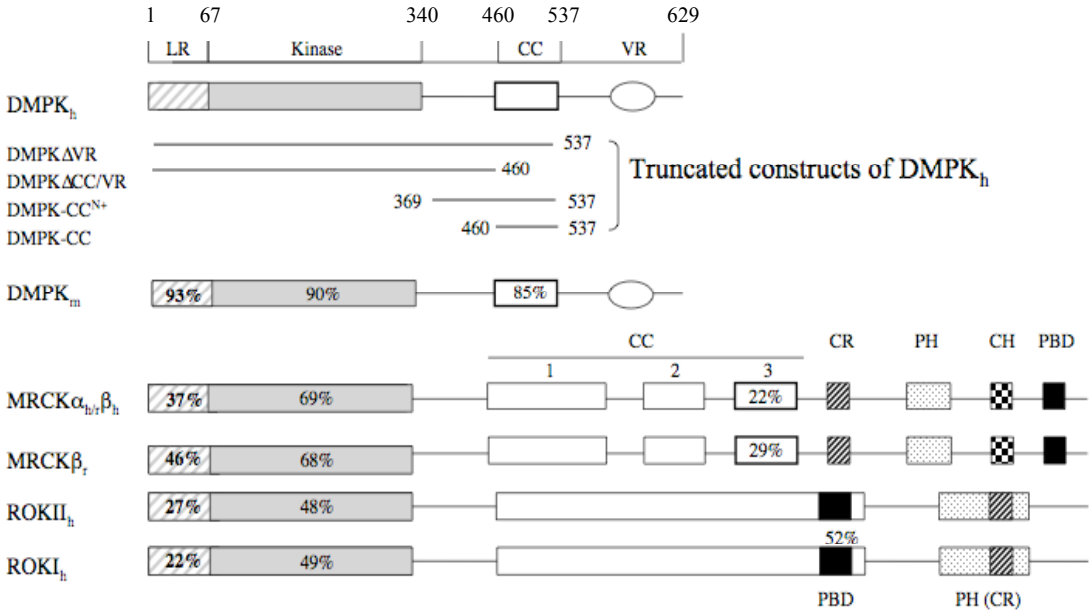



**Figure B.1: Domain organization in DMPK, MRCK and ROCKs.** Domain composition of human DMPK (Q09013-9), mouse DMPK (P54265), human MRCKα (Q5VT26), rat MRCKα (AF021935), human MRCKβ (Q81WQ7), mouse ROCKII (V58513), mouse ROCKI (V58512). LR indicates a leucine-rich region, CC the coiled-coil, VR a region variable across spliceoforms, PDB is the GTPase binding domain. Sequence identity to human DMPK is shown as percentage and on a per domain basis. Identity between the PBD domains of ROCKI and II is also shown.

# B.2 Experimental Work

## B.2.1 Cloning and protein production

Cloning and production of DMPK-CC and DMPKΔCC/VR was as previously reported (Chapters 2 and 3, respectively). The DMPKΔCC/VR construct was provided by Rainer Burcher (Biozentrum, University of Basel).

## B.2.2 Determination of protein concentration

The concentration of DMPK-CC was determined by the Bradford protein assay as reported for the DMPK-CC$^{N+}$ wild-type form (see section A.2.1.4).

DMPKΔCC/VR protein concentration was determined with absortion spectroscopy (Hewlett Packard 8453, HP845X UV-Visible system, Switzerland) at 280 nm. The theoretical extinction coefficient $\varepsilon$ at 280 nm for DMPKΔCC/VR was found to be 56,260 $M^{-1}$ $cm^{-1}$ based on the method by Gill and von Hippel (Gill and von Hippel 1989).

## B.2.3 *In vitro* Binding Studies

Binding studies of DMPK-CC to DMPKΔCC/VR were as follows:

***A) $Ni^{2+}$-NTA Pull-down assays using $His_6$-tagged DMPKΔCC/VR immobilized on a His-Trap HP column versus a flow of untagged DMPK-CC***

Both DMPK-CC and DMPKΔCC/VR were expressed separately. Expression and purification procedures have been described (Chapters 2 and 3, respectively). DMPKΔCC/VR was loaded and immobilized on a His-Trap HP column pre-equilibrated in 50 mM Tris-HCl pH 7.5, 100 mM NaCl. The $His_6$-tag of DMPK-CC was removed using TEV protease (see Annex C) and the cleaved protein was applied to the column. The mixture was incubated for 4 hours at 30 °C in the column and the bound protein eluted from the resin with 200 mM imidazole. A temperature of 30 °C was chosen for this study since most of the reported work on expression and purification of DMPK has been carried out at this temperature (Bush et al. 2000). Untagged DMPK-CC eluted entirely in the flow-through, whereas His-DMPKΔCC/VR eluted in the presence of 200 mM imidazol. No complex DMPK-

CC:DMPKΔCC/VR was observed within the elution fraction. Lysozyme was observed in the elution fractions, identified according to mass spectroscopy. (Figure B.2).

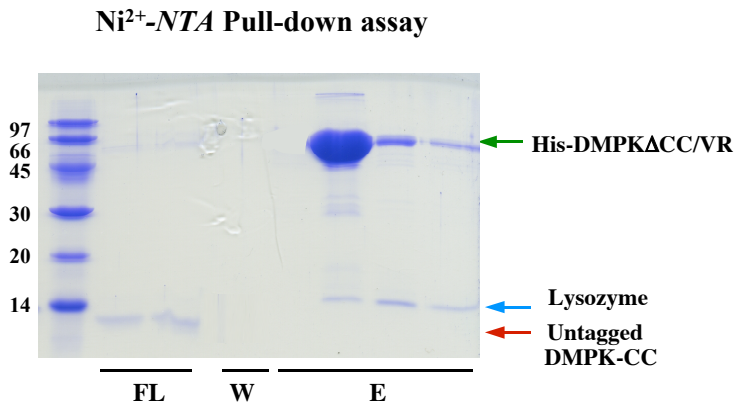



**Figure B.2: Pull-down assay for His-tagged DMPKΔCC/VR against a flow of untagged DMPK-CC.** DMPK-CC does not bind to DMPKΔCC/VR and elutes completely in the flow-through (FL) fractions, while DMPKΔCC/VR and the lysozyme contaminant co-elute (E) at 200 mM imidazole. Washes (W) with 40 mM imidazole are shown (SDS-PAGE analysis).

To confirm results obtained by pull-down assays, and make sure that the lack of interaction was not due to compromised accessibility of the binding site by the in-column immobilization, we examined the interaction by incubating sample mixtures in solution and subsequently analysing them by size exclusion chromatography.

***B) Co-purification attempts of DMPK-CC and DMPKΔCC/VR using size exclusion chromatography after incubation of the mixture***

Both purified proteins were incubated together at a 1:1 molar ratio under two sets of conditions: ***i)*** at 30 °C for 40 minutes and ***ii)*** at RT overnight.

Three different runs were then performed on a size exclusion chromatography column (Superdex 200 16/60, GE Healthcare) for the samples incubated under conditions ***i***; DMPK-CC and DMPKΔCC/VR proteins, as well as the mixture containing the putative complex DMPK-CC:DMPKΔCC/VR were loaded and purified separately on a gel filtration (Superdex 200 16/60, GE Healthcare) column pre-equilibrated in buffer containing 50 mM Tris-HCl pH 7.5, 100 mM NaCl. The combined elution profiles are

shown in Figure B.3. DMPK-CC eluted in a single peak at 73.4 ml. DMPKΔCC/VR eluted in a major peak at 62.3 ml, although the elution profile exhibits "shoulders", indicating the presence of a number of discrete complexes of various sizes possibly due to the presence of contaminants. Finally, purification of the mixture over the gel filtration column revealed several peaks identical to those obtained when the samples were applied separately. This indicated that DMPK-CC and DMPKΔCC/VR did not form any complex under these binding conditions. Similar elution profile was obtained when the samples were incubated under conditions ***ii***.

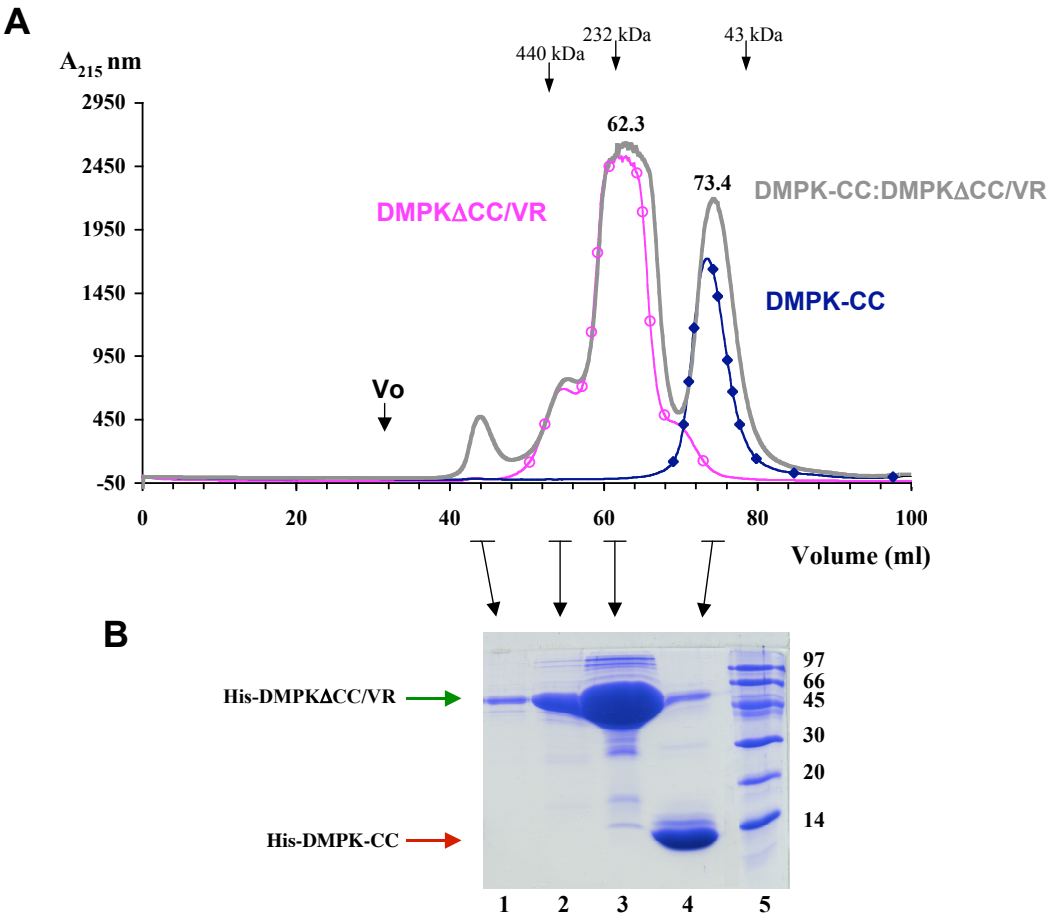



**Figure B.3: Purification of DMPK-CC, DMPKΔCC/VR and DMPK-CC:DMPKΔCC/VR mixture by size exclusion chromatography (Superdex 200 16/60 Healthcare). (A)** The elution profiles of DMPK-CC (◆), DMPKΔCC/VR (o) and the mixture DMPK-CC:DMPKΔCC/VR (—) show that no binding occurred between DMPK-CC and DMPKΔCC/VR when both proteins were in incubated together at a 1:1 molar ratio for 40 minutes at 30 °C under the conditions of this study. Positions of size standards ovalbumin (43 kDa), catalase (232 kDa) and ferritin (440 kDa) are indicated. **(B)** The SDS-PAGE analysis shows the elution peaks of DMPK-CC:ΔMPKΔCC/VR, with lanes 1, 2 and 3 corresponding to DMPKΔCC/VR and lane 4 to DMPK-CC. Lane 4 also shows the presence of DMPKΔCC/VR.

The lack of binding in samples expressed separately could be due to partial misfolding of the binding motifs, which might require of the partner to achieve a native bound state. To account for possible structural dissorders or aberrations at this respect, we analyzed interaction of the samples when present together during their folding step in the bacterial cell by co-expression studies and subsequent co-purification by $Ni^{2+}$-NTA affinity and size exclusion chromatography.

### *C) Co-expression of both constructs followed by co-purification by metal-affinity and size exclusion chromatography*

Plasmids carrying constructs for both DMPK-CC and DMPKΔCC/VR were co-transformed in *E. coli* Rosetta (DE3) strain (Novagen). Culture, lysis and purification conditions were as for separate constructs (Chapters 2 and 3, respectively). The cells were grown in LB medium supplemented with 34 μg/ml chloramphenicol, 25 μg/ml kanamycin and 100 μg/ml ampicillin at 37 °C. Co-expression was induced at an OD of 0.6 by addition of 1mM IPTG and cells were incubated overnight at 30 °C. Co-purification of DMPK-CC and DMPKΔCC/VR was first attempted using a metal-affinity step in 50 mM Tris pH 7.5, 100 mM NaCl and 5 mM β-ME. Elution fractions were subjected to further size exclusion chromatography on a Superdex 200 Tricorn 10/300 GL column (GE Healthcare). The profile of the size exclusion chromatogram of the co-expressed sample was very similar to that obtained when the samples where expressed and purified independently. No complex was observed (Figure B.4).

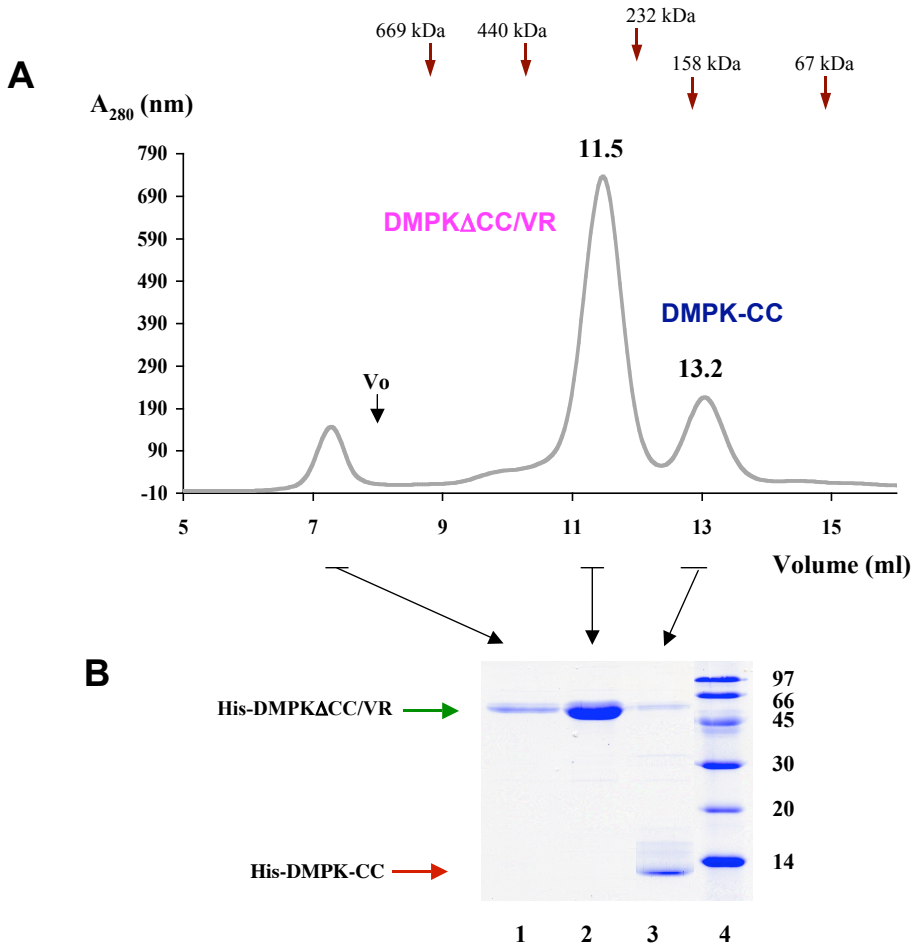



**Figure B.4: Purification of co-expressed DMPK-CC and DMPKΔCC/VR by size exclusion chromatography. (A)** Elution profile of the mixture of DMPK-CC and DMPKΔCC/VR on a Superdex 200 Tricorn 10/300 GL column (GE Healthcare) shows that the samples elute in two different peaks corresponding to DMPKΔCC/VR and DMPK-CC. Positions of size standards thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), Aldolase (158 kDa) and Albumin (67 kDa) are indicated. **(B)** SDS-PAGE analysis lanes 1 and 2 correspond to DMPKΔCC/VR and lane 3 to DMPK-CC.

## B.3 Discussion

To assess whether DMPK-CC is an autoregulatory region, as it occurs in ROCK and MRCK kinases, and if its oligomeric properties might change by interacting within the catalytic domain, we have looked into the possible interaction of DMPK-CC and the latter. Binding studies of DMPK-CC and a truncated DMPK variant lacking the CC domain, DMPKΔCC/VR, were performed by pull-down assays, co-purification trials, as well as co-expression/co-purification experiments. Binding between these species could not be detected by any of the procedures, which yielded independent original species. Hence, DMPK-CC appears neither to interact with the kinase domain nor to form a hetero-coil against the α-helical LR sequence, also present in the DMPKΔCC/VR construct assayed. This indicates that a CC association/dissociation regulatory model such as that proposed for MRCK and ROCKs might not to be applicable to DMPK. In the current study, complex formation was always assessed by methods based on column fractionation and no in-solution method was attempted (such as isothermal titration calorimetry or surface plasmon resonance). Thus, it should not be excluded that the shier forces of the resin could have prevented complex formation. However, our results are in agreement with recent studies on DMPK (Zhang and Epstein 2003; van Herpen et al. 2006) where co-immunoprecipitation experiments on various N- and C-terminal DMPK truncated mutants showed that the CC domain does not interact with the DMPK variant containing only the kinase domain.

Based on these results, it is unlikely that the assembly properties of this CC are altered by a possible embedding within the kinase core. These results also indicate that architecturally, DMPK must consist of a globular fraction and structurally independent shaft. Further details on DMPK architecture are expected from structural studies on full-length forms of DMPK, which are being carried out in our laboratory by Zöhre Ucurum.

Given the heterogeneity in CC sequences and their diversified interaction patterns across DMPK-related kinases, studies are required to reveal the exact role of these moieties and to establish the principles of their regulatory mechanism.

# Annex C

## TEV protease

# C.1 Introduction

TEV protease is the common name for the 27 kDa catalytic domain of the Nuclear Inclusion a (NIa) protein encoded by the tobacco etch virus (TEV). Because its sequence specificity is far more stringent than that of factor Xa, thrombin, or enterokinase, TEV protease is a very useful reagent for cleaving fusion proteins. It is also relatively easy to over-produce and purify large quantities of the enzyme. TEV protease recognizes a linear epitope of the general form E-X-X-Y-X-Q-G/S, with cleavage occurring between Q and G or Q and S. However, the most commonly used sequence is E-N-L-Y-F-Q-G.

The TEV protease employed in this study was produced in-house and consists of the 27 kDa catalytic domain with an N-terminal $His_6$-tag. The seven amino acid consensus sequence is E-N-L-Y-F-Q-G. Cleavage occurs between the conserved Q and G residues.

# C.2 TEV protease production

Protein expression was carried in *E. coli* BL21 (DE3) pLysS strain using pET-24d (Novagen) plasmid, which carries a C-terminal $His_6$-tag and a kanamycin resistance marker. The cells were grown in 500 ml of LB medium, supplemented with 34 μg/ml chloramphenicol and 25 μg/ml kanamycin at 37 °C. Expression was induced by addition of 1 mM IPTG at an OD of 0.6 at 600 nm, and cells were incubated overnight at 20 °C. The cells were harvested by centrifugation at 18,000 rpm, 30 minutes at 4 °C, washed with LB medium and stored at –80 °C until required. Cell lysis, centrifugation and chromatography stepts were carried out at 4 °C.

For purification, the cell paste was defrosted and resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl and 1 mM β-ME supplemented with 10 mM imidazole, 1-5 μg/ml of DNAse, PMSF, and standard protease inhibitor cocktail without EDTA (Roche).

Further cell disruption was achieved by sonication with cooling (Branson Digital Sonifier) using a number of short pulses (3-5 seconds) with pauses (7-10 seconds) for 3 minutes time. Debris was removed by centrifugation at 20,000 rpm for 45 minutes and imidazole was added to the supernatant to a final concentration of 10 mM. Then,

the supernatant fraction was applied to a $Ni^{2+}$-NTA affinity column (GE Healthcare) pre-equilibrated in binding buffer (50 mM Tris pH 8.0, 300 mM NaCl and 1 mM β-ME supplemented with 10 mM imidazole and 20% glycerol). To remove contaminants, the column was washed with binding buffer followed by a gradient of imidazole washes, where binding buffer had been progressively supplemented with 30 and 80 mM imidazole. Each step was composed of ten column volumes. Purified protein was further eluted with binding buffer containing 300 mM imidazole.
The elution product was dialysed (membrane 12-14,000 KDa cut off, Spectra/Por) in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT and 20% glycerol overnight at 4 °C.

## C.3 TEV protease reaction conditions

The reaction buffer used for TEV protease was 50 mM Tris-HCl pH 8.0, 150 mM NaCl and 1mM DTT. TEV protease is maximally active at 34 °C, however, cleavage can be performed at temperatures as low as 4 °C. The duration of the cleavage reaction was overnight at 4 °C or 1 hour at RT (TEV protease is only 3-fold less active at 4 °C than at RT). After digestion of the fusion protein substrate with TEV protease, the cleaved protein was recoved in the flow-through after a subtractive affinity purification step, while the uncleaved sample and the His-tagged TEV protease were absorbed to the affinity resin.

# Chapter 4

## *In vitro* studies on the interaction between DMPK-CC and Rac1 GTPase

# 4.1 Introduction

By analogy to other family members, DMPK is expected to modulate cytoskeletal events (Jin et al. 2000; Schulz et al. 2003) and to possibly act as down-stream effector in the regulatory pathways of small Rho GTPases. Shimizu and co-workers (Shimizu et al. 2000) suggested that the cytoskeleton-linked Rho GTPase Rac1 interacts with DMPK and co-expression of Rac and DMPK triggers its activity in a GTP-dependent manner. Protein binding studies were performed to examine a possible interaction between DMPK and the three small GST-tagged Rho GTPases RhoA, Rac1 and Cdc42 by GST pull-down assays (Shimizu et al. 2000). Equal amounts of bacterially expressed fusion proteins RhoA, Rac1 and Cdc42 were incubated with bacterially expressed FLAG-tagged DMPK and then glutathione-sepharose beads. Bound proteins were eluted from beads by boiling in SDS sample buffer and separated by SDS-PAGE. The presence of DMPK was further detected on Western blots of the solubilized Rac1 complexes, but not of the other GTPases by anti-FLAG. These results were further confirmed by an independent experiment using Rac1 expressed in mammalian cells, indicating that Rac1 physically interacts with DMPK. Furthermore, co-transfection of DMPK with constitutively activated Rac1 mutated variants, activated DMPK transphosphorylation of histone H1 by almost 3-fold, consistent with a potential regulatory interaction.

Since DMPK close relatives ROCK and Citron kinase, have been found to contain a GTPase binding domain within their CC motif, it is logical to speculate that the interaction between DMPK and Rac1 might take place via its CC segment.

In order to investigate a possible interaction between DMPK-CC and the small GTPase Rac1, *in vitro* binding studies of DMPK-CC and Rac1 have been performed. For this, a Rac1 mutant protein (Rac1+), containing a glutamine to leucine substitution at amino acid 61 (Q61L mutation) was employed. This particular mutation was originally described in Ras, where it inhibits GAP-stimulated GTPase activity and therefore, unlike wild-type Rac1, remains in the GTP-bound state during the course of the binding study. Nevertheless, neither CRIB nor REM motifs have been described for DMPK.

# 4.2 Experimental Work

## 4.2.1 Rac1 and Rac1+ production

Bacterial expression constructs pGEX-4T3 encoding both human wild-type Rac1 and Rac1+, constitutively active Q61L mutant, were kindly provided by Daniel W. Baird at Cornell University, Ithaca, New York.
Plasmid pGEX-4T3 encodes a thrombin cleavage site between the GST affinity tag and the N-terminal target protein. The nucleotide sequences of the two clones were confirmed by DNA sequencing (Microsynth AG, Balgach, Switzerland).

Plasmids coding for Rac1 and Rac1+ were used to transform *E. coli* Rosetta (DE3) strain (Novagen). Over-expression of the GST-Rac1 and Rac1+ fusion proteins was achieved by growing the cell cultures in LB medium supplemented with 34 μg/ml chloramphenicol and 100 μg/ml ampicillin at 37 °C. Expression was induced by addition of 1 mM IPTG at an $OD_{600}$ of 0.6 and cells were incubated overnight at 20 °C. Cells were pelleted by centrifugation at 18,000 rpm for 30 minutes and stored frozen at -80 °C. The cell pellets were then resuspended in lysis buffer containing 50 mM Tris 7.5, 100 mM NaCl, supplemented with 1-5 μg/ml of DNAse, PMSF, a spatula-tip of lysozyme and standard protease inhibitor cocktail without EDTA (Roche). Further cell lysis was performed by sonication in ice (Digital Sonifier, Branson) using a number of short pulses (3-5 seconds) with pauses (7-10 seconds). The cell lysate was clarified by centrifigation at 18,000 rpm for 30 minutes at 4 °C. The supernatant was loaded onto a 1 ml GSTrap FF column (GE Healthcare), which had been pre-equilibrated in lysis buffer. The unbound protein was washed out and the bound protein eluted by supplementing the buffer with reduced gluthathione 8.0. All the buffers were filter-sterilized before use and contained freshly prepared 5 mM β-ME. The eluted fractions were pooled and dialysed against 20 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT. The overall yield was estimated to be 5-10 mg of pure Rac1 and 5-10 mg of Rac1+ per liter of *E. coli* culture. Purified Rac1 and Rac1+ were subsequently used for *in vitro* binding studies to the CC domain of the DMPK protein kinase. Protein concentrations were determined with absorption spectroscopy at 280 nm. The theoretical Extinction coefficient ε of Rac1 and Rac1+

was calculated to be 62,900 $M^{-1}$ $cm^{-1}$ according to Gill and von Hippel (Gill and von Hippel 1989).

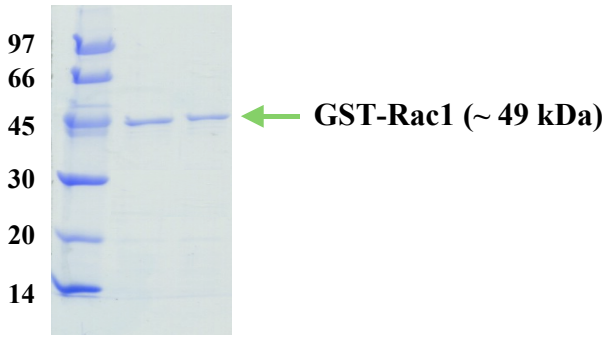



**Figure 4.1: SDS-PAGE analysis of wild-type GST-Rac1 purification by GST-affinity chromatography.** The recombinant GST-Rac1 protein is ~ 49 kDa, consisting of the 21.5 kDa Rac1 protein plus a 27 kDa N-terminal GST tag.

## 4.2.4 *In vitro* Binding Studies

Binding studies between DMPK-CC and both Rac1 and Rac1+ were attempted as it follows:

***A) Affinity pull-down assays using GST-tagged Rac1 immobilized on a GSTrap FF column versus a flow of His-tagged DMPK-CC in the presence of a GTP analogue (GTPγS), followed by a $Ni^{2+}$-NTA affinity chromatography step***

*In vitro* binding studies protocols were carried out with a modification of *Barz et al.* (Barz et al. 2000). A GSTrap FF column was equilibrated in binding buffer (20 mM Tris pH 7.5 / 50 mM NaCl / 1 mM EDTA / 1 mM DTT). Then, GST-tagged Rac1 was loaded onto the column at a flow rate of 0.15 ml/min, followed by washes with 7 column volumes (CV) nucleotide depletion buffer (NDB) containing 20 mM Tris pH 7.5 / 50 mM NaCl / 1 mM EDTA / 1 mM DTT / 5% glycerol / 0.05% Triton X-100 to deplete wild-type Rac1 of bound GDP and GTP, and incubated for 20 minutes at RT. To load the non-hydrolyzable GTPγS, washing was continued with 8 CV GTP analogue loading buffer (20 mM Tris pH 7.5 / 50 mM NaCl / 1 mM DTT / 5% glycerol / 0.05% Triton X-100 /10 mM $MgCl^{2+}$ / 20 µM GTP analogue) and incubated

for 30 minutes at RT. The His-tagged DMPK-CC was then added to GST-Rac1 and the mix was incubated for 1 hour at 4 °C. Unbound DMPK-CC was removed with washing buffer (20 mM Tris pH 7.5 / 50 mM NaCl 1 mM DTT) and elution was accomplished by addition of 15 CV elution buffer (20 mM Tris pH 7.5 / 50 mM NaCl 1 mM DTT 10 mM glutathione). Where indicated, DMPK-CC was used instead of Rac1 as a negative control. Rac1 was observed to elute mainly in fractions 1 and 2 (1.5 ml each). Therefore, these fractions were pooled and further purified by affinity chromatography using a nickel resin according to the manufacturer's instructions in order to co-isolate DMPK-CC and Rac1 proteins. Where indicated, Rac1 was used instead of DMPK-CC as a negative control (Figure 4.2).

During these studies, no interaction between DMPK-CC and Rac1 was detected in the presence of GTPγS. DMPK-CC was revealed mainly in the flow-through during the GST pull-down assay as depicted in the SDS-PAGE analysis A. However, Rac1 appears in the elution fraction of the His-pull down (B), probably due to binding to lysozyme.

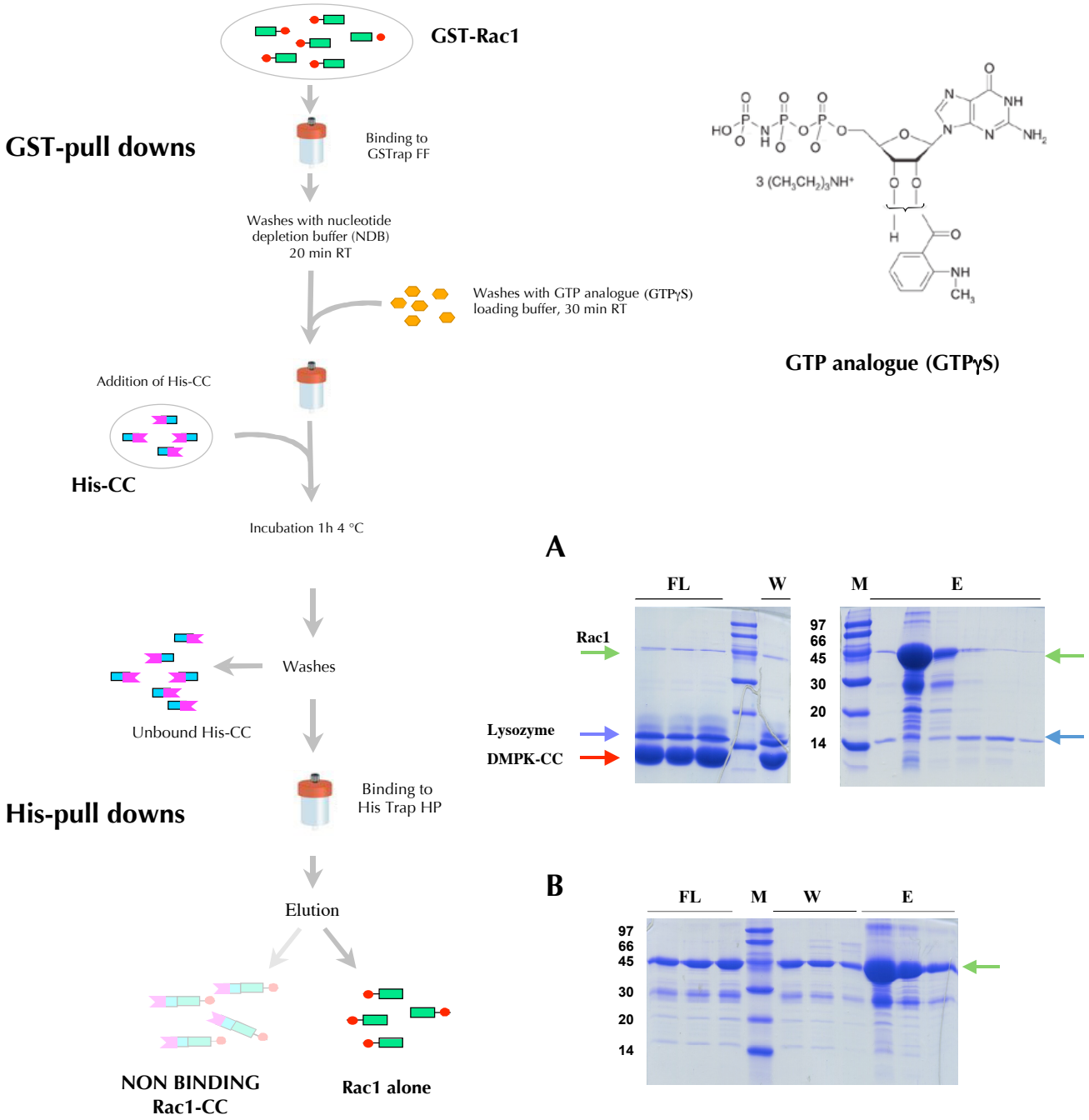



**Figure 4.2: Overview of the binding assay flow scheme of GST-tagged Rac1 and His-tagged DMPK-CC using GST-pull downs followed by His-pull downs.** GST-tagged Rac1 was loaded onto GSTrap FF column pre-equilibrated in binding buffer, followed by washes with NDB to deplete Rac1 of GDP and GTP, and incubated for 20 minutes at RT. To load the non-hydrolyzable GTP analogue, (GTPγS), the column was subsequently washed with GTP analogue loading buffer and incubated for 30 minutes at RT. The His-tagged DMPK-CC was then added to GST-Rac1 and the mix was incubated for 1 hour at 4 °C. Unbound DMPK-CC was removed with washing buffer and elution was by addition of 10 mM glutathione. Elution fractions were pooled together and further loaded onto a His-Trap HP column. Purification was according to the manufacturer's instructions. The inlets show the non-hydrolyzable GTP analogue, guanosine-5'-0-3-thiotriphosphate (GTPγS) and the SDS-PAGE analyses of GST-pull downs **(A)** and His-pull downs **(B)**.

***B) Affinity pull-down assays using His-tagged DMPK-CC captured on a $Ni^{2+}$-NTA column versus a flow of GST-tagged Rac1 and Rac1+***

Expression and purification of DMPK-CC, Rac1 and Rac1+ were as previously reported (Chapter 2 and sections 4.2.2-3). First, to obtain the active conformation of Rac1, the purified GTPase was incubated with 1:1 molar ratio of $MgCl^{2+}$ and a 2-fold molar excess of GTPγS for 30 minutes at RT. During the incubation process, the sample precipitated or aggregated. Therefore, the incubation temperature was lowered to 4 °C, with no better results. Finally, Rac1 was incubated with 5 mM $MgCl^{2+}$ and a 2-fold molar excess of GTPγS for 15 minutes at 30 °C in the presence of 10 mM EDTA to facilitate nucleotide exchange, as described in Knaus *et al*. (Knaus et al. 1992). Subsequently increasing the $MgCl^{2+}$ concentration (to 50-70 mM) stopped the nucleotide exchange. The sample did not precipitate nor aggregate under these conditions.

Co-purification of DMPK-CC and both Rac1 and Rac1+ was attempted by affinity pull-down assays using His-tagged DMPK-CC immobilized on a His-Trap HP column (GE Healthcare) versus a flow of either activated GST-tagged Rac1 or constitutively active Rac1. His-pull down assays results revealed no binding of DMPK-CC and Rac1 proteins (Figure 4.3).

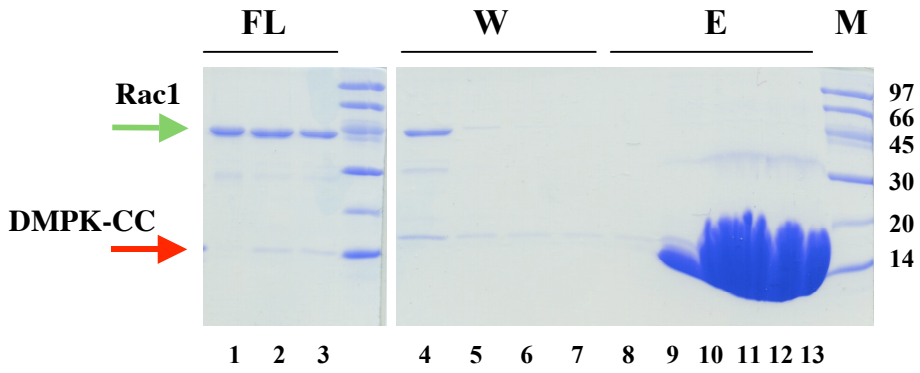



**Figure 4.3: SDS-PAGE analysis of His-pull-down assays using His-tagged DMPK-CC captured on a His-Trap HP column versus a flow of GST-tagged activated Rac1. (A)** Lane 1-3, flow-through of the purified Rac1 loaded onto the $Ni^{2+}$-NTA affinity column with immobilized His-tagged DMPK-CC. Lanes 4-7, washes. Lanes 8-13, elution with elution buffer containing 200 mM imidazole. DMPK-CC and Rac1 were indicated with a red and a green arrow, respectively.

Besides this, co-expression of DMPK-CC and Rac1+ was also attempted. However, no co-expression was observed, resulting in the loss of Rac1+. The inability of both plasmids to coexist might be explained by incompatibility of the plasmids.

## 4.3 Discussion

In order to study whether the binding of Rac1 to DMPK could be mapped to its CC domain, *in vitro* binding assays of His-tagged DMPK-CC with either activated GST-tagged Rac1 or constitutively active Rac1 mutant have been tested. Binding studies included His- and GST-pull-downs.

Binding studies revealed no detectable specific interaction of DMPK-CC with either wild-type Rac1 or the constitutively active Rac1 mutant in any of the assays performed. Previous studies (Shimizu et al. 2000) employed Western blotting techniques, which are much more sensitive than SDS-PAGE. Moreover, they used the full-length DMPK kinase for the binding assays, which suggests that Rac1 interaction to DMPK and subsequent activation, presumably involve other regions apart from the CC motif, as it occurs for MRCK. Nonetheless, binding studies between full-length DMPK and the three small GTPases Rho, Rac1 and Cdc42, using His- and GST-pull-downs as well as co-expression followed by co-purification using native-PAGE were performed in our group by Zöhre Ucurum. However, no interaction was observed between DMPK and any of the GTPases.

Current results apparently disagree with reported studies, which defend that Rac1 can physically interact with DMPK, thereby triggering its trans-phosphorylation activity in a GTP-sensitive manner (Shimizu et al. 2000).

It should be borne in mind that the presence of the complex was always assessed on column runs and any non-column based method has been employed. Thus, it should not be excluded that the shier forces of the resin could have avoided complex formation.

In conclusion, we have shown that DMPK is unlikely to interact with Rac1 through its CC region under the conditions tested. A possible explanation for this might be that the GTPase-kinase interaction needs of any region in DMPK out of the CC domain to

be effective. Besides, the level of assembly of DMPK-CC in isolation might no the same than in the context of the full-kinase.

Therefore, to analyse the binding between DMPK-CC and Rac1, binding studies should employ untagged proteins. Likewise, other purification methods than those column-based such as western blot should also be tried. Furthermore, in order to gain a better understanding of the interactions of DMPK and Rac1 and to investigate whether the small GTPase is responsible for the subsequent activation of the kinase, further experiments should be performed within the context of the full DMPK kinase taking into account tags and plasmids employed.

# Chapter 5

## Production of CUG-BP, a novel alternative splicing-regulator involved in DM

# 5.1 Introduction

According to the RNA gain-of-function hypothesis, nuclear accumulation of the DMPK expanded-repeats cause DM disease via recruitment of the splicing regulator, CUG-BP, which subsequently causes misregulated splicing of a subset of pre-mRNAs that are natural CUG-BP targets. Given the involvement of this novel splicing-regulator CUG-BP in DM1 pathogenesis and its expected role as DMPK substrate, CUG-BP was incorporated in the current study, both in its full-length form and as a subfragment encoding CUG-BP-$RRM_{1-2}$ (sequence coding for RRM1 and RRM2 domains of CUG-BP). The primary goal of this work was to establish their recombinant over-expression for further biophysical studies both in the absence and presence of DMPK.

# 5.2 Experimental Work

## 5.2.1 Cloning

Expression constructs were derived from human CUG-BP cDNA (residues 1-482) in the pET15b vector (Figure 7.1) provided by Maurice Swanson, University of Florida. Fragments 1-482 (full-length) and 1-186 (sequence encoding RRM1 and RRM2 domains of CUG-BP) were digested from pET15b and subcloned into the pETM-11 expression vector.

The sequence coding for the full-length CUG-BP was amplified by PCR with Expand Long Template PCR System (Roche) using the forward primer 5'-CGTTCGGTCTCACATGAACGGCACCCTGGACCAC-3' with *Bsa*I restriction site, which creates a compatible overhang for *Nco*I in the vector and the reverse primer 5'-CGCGGATCCTTAGTAGGGCTTGCTGTCATTC-3' which includes a *BamH*I restriction site. The sequence coding for RRM1-2 domains of CUG-BP was amplified with Pfu polymerase (Stratagene) using the same sense primer than the full-length CUG-BP and the antisense primer 5'-CGCGGATCCTTAAGCAAATTTTACCACCATG-3' which includes a *BamH*I restriction site. PCR was carried out as reported for DMPK-CC$^{N+}$ (see section A.2.2.1), slightly modified for the full-length CUG-BP cDNA fragment, where the cycle profile was: 5 minutes at 95 °C (*Initiation*), then 45 seconds at 95 °C (*Melting*), 45 seconds at 60 °C (*Annealing*) and 2 minutes at 72 °C (*Elongation*) for a total of 30 cycles, followed by 30 minutes at 72 °C. The PCR products were then digested with *Nco*I and *BamH*I enzymes (sites are underlined in the primers) and ligated into the pETM-11 vector with T4 DNA ligase as previously described for DMPK-CC$^{N+}$ (see section A.2.2.1).

DNA manipulations and bacterial transformations were carried out using standard protocols (Sambrook et al. 1989). Transformants were assayed for correct incorporation of the insert by PCR screening and their sequence verified by DNA sequencing (Microsynth AG, Balgach, Switzerland).

The resulting clones were termed CUG-BP (full-length) and CUG-BP-RRM$_{1-2}$ (1-186), respectively.

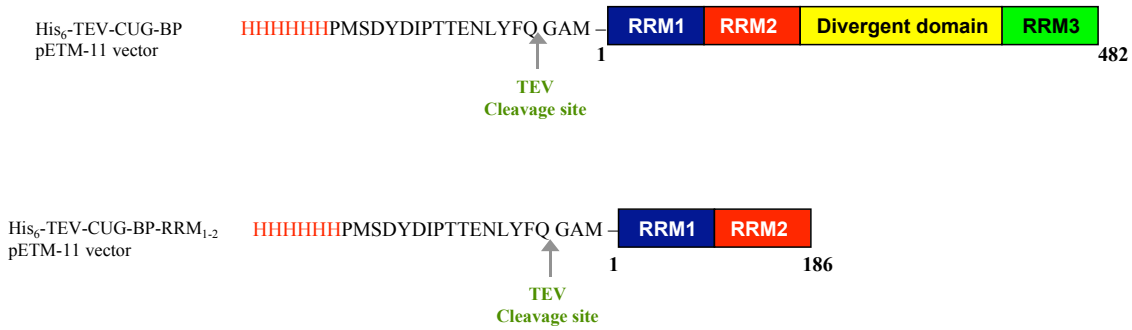



**Figure 5.1: Design of CUG-BP constructs**. Two constructs of CUG-BP were produced. The first one corresponds to the full-length CUG-BP and the second one is a shorter variant, which encodes for the RNA-binding motifs 1 and 2 ($RRM_{1-2}$) of CUG-BP. Both constructs contain the $His_6$-tag, shown in red, and a TEV protease cleavage site with its recognition sequence marked in green.

## 5.2.2 Over-expression

CUG-BP and CUG-BP-$RRM_{1-2}$ were expressed in *E. coli* Rosetta (DE3) (Novagen) grown in LB medium supplemented with 34 µg/ml chloramphenicol and 25 µg/ml kanamycin at 37 °C. Expression was induced by addition of 1mM IPTG at an $OD_{600}$ of 0.6 and cells were incubated overnight at 25 °C. Cells were harvested, washed with LB medium and stored at –80 °C.

## 5.2.3 Purification of full-length CUG-BP

Cells were resuspended in lysis buffer composed of 50 mM Tris pH 7.0, 150 mM NaCl and 5 mM β-ME supplemented with 10 mM imidazole, 1-5 µg/ml of DNAse, PMSF, a spatula-tip of lysozyme and standard protease inhibitor cocktail without EDTA (Roche). Cell lysis was performed by sonication in ice (Branson Digital Sonifier). The lysate was harvested by centrifugation at 18,000 rpm for 30 minutes at 4 °C and the clarified supernatant was subjected to a single-step purification using $Ni^{2+}$-NTA affinity chromatography pre-equilibrated in lysis buffer. The bound protein was eluted with 200 mM imidazole.

SDS-PAGE of purified CUG-BP showed protein degradation (Figure 5.2). Two strong bands, along with several minor bands were identified as CUG-BP by peptide mass fingerprinting (PMF) technique. The sizes of the two major bands were ~55 and 27 kDa. The higher band corresponded to the expected MM of full-length CUG-BP (theoretical MM of 54.7 kDa). This result suggests that full-length CUG-BP could have been cleaved during expression generating this unexpected truncated CUG-BP form.

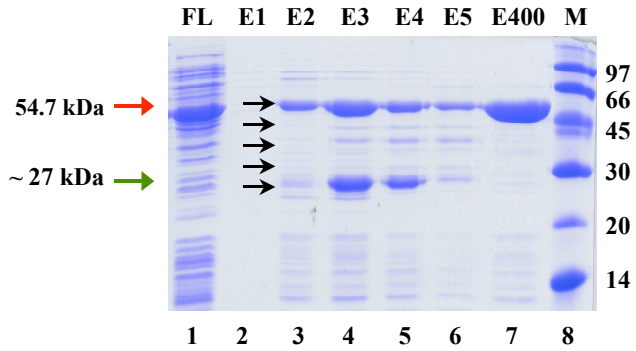



**Figure 5.2: SDS-PAGE analysis of CUG-BP full-length purification on a $Ni^{2+}$-NTA affinity column.** Lane 1, flow-through (FL). Lanes 2-6, elution fractions using 200 mM imidazol. Lane 7, elution with 400 mM imidazol. All fractions contain 1.5 ml. Lane 8, marker. CUG-BP full-length and the 27 kDa band are indicated with red and green arrows, respectively. Bands identified as CUG-BP by mass spectroscopy are shown with black arrows.

## 5.2.4 Peptide Mass fingerprinting

In order to analyze the nature of the multiple bands appeared during the full-length CUG-BP purification procedure, and to check whether the sample suffered from degradation, a peptide mass fingerprinting experiment was performed.

PMF is an analytical technique developed by John Yates and colleagues (Griffin et al. 1995) used to identify proteins by matching their constituent fragment masses (peptide masses) to the theoretical peptide masses generated from a protein or DNA database.

In first, Protein bands of the SDS-gel corresponding to the MMs of the target protein were cut and subjected to trypsic digestion. Subsequently, the absolute masses of the digested fragments were accurately measured using matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometric analysis. MALDI-TOF mass spectra were acquired on a Bruker Reflex III instrument (Bruker

Daltonik). Peptides were analysed either in linear or in reflector mode by using a-cyano-4-hydroxycinnamic acid (1 mg/ml in 80% acetonitrile / 0.1% TFA) matrix. The matrix molecules are required for the desorption of the peptide molecules. Samples were prepared by mixing 1 ml peptide solution with 1 ml matrix solution. Then, 300 nl of the mix were pipetted onto a MALDI target. Matrix and peptide molecules co-crystallized on the MALDI. The instrument was calibrated with angiotensin II, substrate P, bombesin, and ACTH-18-39. For each proteolysis intermediate, the MMs of the ejected species were measured and searched in the sequence library of the theoretically trypsined CUG-binding proteins using the programme MASCOT. These fragments were used to map the boundary of the proteolysis intermediate protein. The PMF was carried out by the group of Dr. Paul Jenö, Biozentrum, Basel.

As a result, according to PMF experiment, all bands tested were identified as CUG-BP. Therefore, CUG-BP protein suffered from degradation. In addition, since the major band observed of ~27 kDa has also been identified as CUG-BP, it is likely that it corresponds to a truncated form of the full-length CUG-BP protein as a consecuence of cleavage of the latter occurred during the expression.

## 5.2.5 Purification of CUG-BP-RRM$_{1-2}$

For CUG-BP-RRM$_{1-2}$ purification, lysis was achieved as for the full-length CUG-BP construct in a lysis buffer composed of 50 mM Tris pH 8.5, 150 mM NaCl and 5 mM β-ME. Clarified supernatant was applied to a $Ni^{2+}$-NTA column (GE Healthcare) pre-equilibrated in lysis buffer. The bound protein was eluted with 200 mM imidazole. SDS-PAGE analysis of the $Ni^{2+}$-NTA fractions showed one major band corresponding to CUG-BP-RRM$_{1-2}$ (theoretical MM of 24.1 kDa) along with some contaminant proteins. Elution fractions were then pooled and dialysed (membrane 12-14,000 KDa cut-off, Spectra/Por) in 50 mM Tris-HCl pH 8.5, 150 mM NaCl and 5 mM β-ME overnight at 4 °C. The $His_6$-tag was removed by incubation with TEV protease and separation of the cleaved protein was performed on a second metal-affinity step as previously reported (see section A.2.2.3). The last purification step used size exclusion chromatography on a column (Superdex 200 Tricorn 10/300 GL, GE Healthcare), equilibrated in 50 mM Tris pH 8.5, 150 mM NaCl and 5 mM β-ME

(Figure 5.3). The fractions corresponding to pure CUG-BP-RRM$_{1-2}$ were then pooled and concentrated to approximately 20 mg/ml (10,000 Da MW cut-off, Millipore) for the subsequent crystallization trials. This protocol yielded approximately 3 mg per liter of culture of pure protein. All purification procedures were executed at RT and protein concentration was measured utilizing both Bradford assay (Bradford 1976) and BCA method (Pierce).

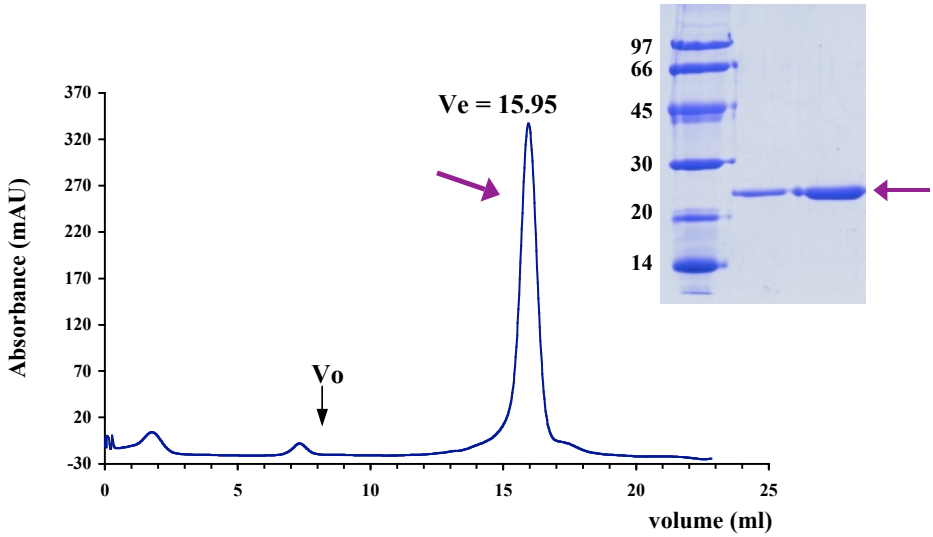



**Figure 5.3: Purification of CUG-BP-RMM1-2 by size exclusion chromatography**. Elution profile of CUG-BP-RMM$_{1-2}$ on a Superdex 200 Tricorn 10/300 GL column (arrow). Elution peak is labeled with its peak elution volume. The inlet shows an SDS-gel of the elution peak fractions, with highly purified CUG-BP-RMM$_{1-2}$ (arrow).

## 5.2.6 Crystallization

Crystallization trials for non-tagged CUG-BP-RRM$_{1-2}$ were carried out at RT using the hanging-drop vapour-diffusion technique on 24-well VDX plates (Hampton Research) (Suzuki et al. 1993). Each drop contained equal volumes (2 μl) of protein solution and crystallization reservoir solution, equilibrated against 1 ml reservoir solution.

Protein concentration varied from 4 to 15 mg/ml. A large range of crystallization conditions were tested using numerous commercially available sparse matrix screens including kits I and II from Hampton Research and Molecular Dimensions, Wizard and Cryo screens from Emerald Biostructures as well as polyethylene glycol (PEG) screens with pHs varying from 4.5 to 9.2 (Hampton Research). Besides this, The

Quik Screen (Hampton Research) was used to Screen 0.8 - 1.8 M Sodium/Potassium Phosphate versus pH 5 - 8.2 in a grid format.
No crystal growth was observed for CUG-BP-$RRM_{1-2}$ under any of the protein concentrations and crystallization conditions tested.

## 5.3 Discussion

In order to characterize CUG-BP, we have cloned two constructs encoding CUG-BP (full-length CUG-BP) and CUG-BP-$RRM_{1-2}$ (sequence coding for RRM1 and RRM2 domains of CUG-BP). Purification of the short construct yielded single pure protein, apparently stable, whereas, isolation of the long construct has not been possible due to progressive protein degradation. His-tagged proteins were expressed in *E.coli* and purification employed $Ni^{2+}$-NTA affinity and size exclusion chromatography. SDS-PAGE analysis of purified CUG-BP-$RRM_{1-2}$ showed the presence of a single band at the expected MM (24.1 kDa). Pure protein was concentrated and used for crystallization trials. However, no crystal growth was observed. In the case of the full-length CUG-BP, SDS-PAGE analysis during protein purification showed protein degradation.
So far, little information is available on CUG-BP production procedures. Two main research groups have lead studies on this novel alternative splicing-regulator. The first group employed extracts from mammalian HeLa cells, fibroblast and myotubes as protein sources and further purification was by denaturation/elution techniques. The second group utilized bacterially expressed His-tagged CUG-BP and purification was performed by metal affinity chromatography. Interestingly, bound protein eluted in the presence of high concentrations of EDTA. Since we have attempted expression of stable full-length CUG-BP so far without success, protein degradation might be explained by improper folding caused by the affinity tags or the absence of needed post-tranlational modifications. Nevertheless, any of the previous group referred to the stability of the protein, yields nor crystallizations trials attempted. Therefore, in order to purify stable protein, different expression systems, as well as untagged protein should be used. Crystallization was attempted for CUG-BP-$RRM_{1-2}$ in a wide

range of crystallization conditions, but no crystals growth was observed.

The characterization of CUG-BP is of great interest since it is involved in the pathogenesis of DM and it could serve as a new therapeutic target for eventual treatment of DM patients.

# Chapter 6

## Crystallization of the N-terminal Z1Z2 of human titin

Marino, M., Zou, P., Svergun, D.I., Garcia, P., Edlich, C., Simon, B., Wilmanns, M., Muhle-Goll, C. & Mayans, O. (submitted). The Ig-doublet Z1Z2 – a model system for the hybrid analysis of conformational dynamics in poly-Ig tandems from titin.

The crystallization in this chapter is my responsibility.

# The Ig-doublet Z1Z2 – a model system for the hybrid analysis of conformational dynamics in Ig tandems from titin

Marco Marino[1], Peijian Zou[2], Dmitri Svergun[2,3], Pilar Garcia[1], Christian Edlich[4], Bernd Simon[4], Matthias Wilmanns[2], Claudia Muhle-Goll[4,5]*, Olga Mayans[1]*

*[1]Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH4056 Basel, Switzerland; [2]EMBL Outstation, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany; [3]Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333 Moscow, Russia; [4]EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany; [5]Max-Planck-Institut for Medical Research, Jahnstrasse 29, 69120 Heidelberg, Germany.*

**Running title:** Conformational dynamics of Z1Z2

* Authors for correspondenceTel: +41 61 267 2083 / Fax: +41 61 267 2109 /
Olga.Mayans@unibas.ch Tel: +49 6221 387 397 / Fax: +49 6221 387 306 /
Muhle@embl-heidelberg.de

**Keywords**: Muscle sarcomere, titin, poly-Ig tandems, X-ray crystallography, NMR residual dipolar couplings, SAXS, filamentous protein structure and dynamics

**Summary**

**Titin is a gigantic elastic filament that determines sarcomere ultrastructure and stretch response in vertebrate muscle. It folds into numerous Ig and FnIII domains connected in tandem. Data on interdomain arrangements and dynamics are essential to understand the architecture and function of this filament. We have performed a mechanistic analysis of the conformational dynamics of a two Ig fragment from the N-terminus of titin, Z1Z2, using in combination X-ray crystallography, SAXS, NMR relaxation data and residual dipolar couplings. Z1Z2 adopts a defined semi-extended conformation in solution, where close-hinge arrangements represent low probability states. Although specific interdomain contacts are not encountered, the linker acquires semi-rigid properties via a weak hydrophobic component, which appears to allow "viscous" modular oscillations but prevents free motions. Thus, Z1Z2 constitutes an adaptable modular system with restricted dynamics. Linker adaptability allows Z1Z2 to induced-fit its binding partner telethonin, where a pre-existent conformation prevents an entropically unfavorable complexation and, possibly, contribute to selective partner recruitment onto the repetitive surface of the filament. The molecular interconversion of Z1Z2 conformations is analyzed.**

Filamentous proteins are characterized by intrinsic flexibility and a poorly defined long-range order that defies the elucidation of its atomic principles. Crystallographic approaches suffer from troubled crystal growth and crystalline conformations can be of doubtful biological relevance, while NMR methods based on NOE-derived distance restraints are compromised by the scarcity of close contacts across modular components, which prevents an accurate description of the overall conformation. These difficulties have obstructed the analysis of a crucial modular protein, titin, which via formation of giant filaments controls the ultrastructure and passive mechanics of vertebrate muscle. Titin is central to muscle elasticity, stretch-stress biomechanical pathways and myofibrillogenesis (for a review see Granzier and Labeit, 2004). It folds into ~300 Ig and

FnIII modules assembled in tandems that span over 1 $\mu$m length from the Z-disc, at the beginning of the sarcomere, to the central M-line. To date, atomic models have only been reported for a few components of titin (listed in Marino et al., 2005) primarily as isolated domains that do not clarify further the architecture of this filament. An initial insight into domain arrangements in the tandems has been obtained through non-atomic methods such as small angle X-ray scattering (SAXS), NMR relaxation studies and electron microscopy. These have addressed I91-I92 (formerly I27-I28) and I91-I94 from the distal I-band of titin (Improta et al., 1998); the Ig-doublet Z1Z2 at the N-terminus (Zou et al., 2003) and I65-I70 from the skeletal I-band (Marino et al., 2005). Complementary, interactions between consecutive Ig in I-band fragments I10-I11 and I91-I96 (Politou et al., 1996; Scott et al., 2002) have been examined using thermodynamics of unfolding, cross-linking and inter-Ig complexation. Taken together, these data suggest that Ig modules of titin follow a "beads-on-a-string" model of organization with limited interdomain motion but weakly interacting neighbouring modules. However, in the absence of atomic models of tandem domain fragments of titin, the molecular principles underlying the conformation and dynamics of this filament remain unclear.

Here, we analyze the conformational dynamics of the N-terminal doublet Z1Z2. This is integral to the sarcomeric Z-disc, where it interacts with the small protein telethonin (Mues et al., 1998). Overexpression of Z1Z2 and telethonin disrupts the association of titin filaments and the ultrastructure of the Z-disc (Gregorio et al., 1998), highlighting the central skeletal role of these components. The crystal structure of Z1Z2 complexed to telethonin has been recently reported (Zou et al., 2006), this being the only domain pair of titin for which a model is available. The complex consists of a pseudosymmetric arrangement where two antiparallel Z1Z2 sandwich one telethonin molecule. The latter forms an elongated _-sheet structure that interacts with its two Z1Z2 partners via main-chain hydrogen bonds, thereby mediating titin assembly and anchoring at the Z-disc. We investigate here the atomic characterization of isolated Z1Z2 applying in combination X-ray crystallography, SAXS, NMR $^{15}$N relaxation studies and residual dipolar couplings (RDC). Our findings provide an understanding of the structural dynamics of this doublet and its possible significance for the complexation of telethonin. Thus, Z1Z2 constitutes the first Ig-tandem of titin for which an architectural, dynamic and binding characterization becomes completed. This might serve as a model system to understand further the molecular mechanisms of the large poly-Ig arrays of titin.

## RESULTS

### Crystal structures of two conformational states

The crystal structure of Z1Z2 has been elucidated at 2.0 Å resolution. Data and model statistics are given in Table 1. Z1 and Z2 belong to the I(intermediate)- set of Ig folds (Harpaz and Chothia, 1994). They share 41% seq. id. and high structural similarity (rmsd ~1.02 Å calculated for C_ atoms using TOP; Lu, 2000). The crystal type used in this study contains three Z1Z2 copies per asymmetric unit, corresponding to two different conformations (Fig 1a). One copy exhibits an extended domain arrangement where the two Ig are in almost coaxial orientation. The two others display a compact, V-shaped conformation in which the two Ig nearly pack laterally. Values for hinge opening and domain torsions in these models are given in Table 2. The two compact models are basically identical (rmsd 0.67 Å for C_ atoms from the entire molecule). These data reveal that domains in Z1Z2 can undergo long-range, complex motions.

In the crystal, every Z1Z2 copy binds a metal ion (Fig 1a). The latter is thought to be cadmium, which was present in large quantity in the crystallization medium. This was supported by anomalous diffraction data recorded at _=1.54 Å, which yielded peaks of 16.5_, 15.4_ and 14.0_ height for metal sites calculated using apo-Z1Z2 phases. In its open conformation, Z1Z2 coordinates the metal via residues from Z1 (E26 and H28) and Z2 (E165) of another molecule. In the closed forms, metal binding involves both Z1 (E26, H28) and Z2 (E155) from a same molecule. Here, metal coordination constitutes the most prominent interaction across domains where, otherwise, only scarce, non-conserved contacts are observed, namely hydrogen bonds Z1:Q13–Z2:D159 (2.53 Å) and Z1:Q10–Z2:E155 (2.45 Å) respectively for each closed model. The open conformation is void of any specific Ig-Ig contacts, but it includes a weak hydrophobic component at the linker region derived from the clustering of the aliphatic portion of E19 and L18 in Z1, the aliphatic chain of E100 at the linker and I130 in Z2 (Fig 1b).

The arrangement of Ig modules in Z1Z2 depends on the conformation of a three-residue linker with sequence "AET" (residues 99-101). Both in open and closed models, linker groups are well defined in electron density maps (Fig 1c) and their crystallographic thermal factors do not deviate significantly from the protein average. A comparative analysis of main chain torsion angles across models shows that residue A99 constitutes the effective hinge in these structures. A99 changes its conformation from extended to helical by a rotation around its dihedral angle psi (Fig 1d). Thus, the long-range domain motions in these structures are torsional in nature and derive primarily from a rotation around one single main chain bond, that additionally acts here as mechanism of hinge closure.

**Domain orientation as derived from RDC data**

Crystal structures may suffer from distortions due to packing forces or exhibit selected conformations that, even if sparsely populated in solution, might favour specific lattice contacts. Thus, we analyzed the relative orientation of domains in Z1Z2 using NMR data derived from RDC in a partial alignment medium (Tjandra et al., 1997; Bax et al., 2001). RDC define an external coordinate system for the molecule of interest termed the "alignment tensor", which is characterized by the axial component Da and the rhombicity R. For a modular protein, the orientation of each domain can be determined separately within this coordinate frame and the relative interdomain orientation deduced (Losonczi et al., 1999).

Partial alignment of Z1Z2 was induced using compressed, positively charged polyacrylamide gels (Ulmer et al., 2003; Cierpicki and Bushweller, 2004). Chemical shift analysis (Edlich and Muhle-Goll, 2003) showed that the secondary structure elements of the molecule in solution are identical to those in the crystal. Only a single set of chemical shifts was observed for all residues, including linker residues, suggesting that in solution Z1Z2 either adopts a preferential conformation or alternative forms are dynamically averaged. The linewidth of linker and core residues was equivalent. The backbone torsion angles of the linker groups derived from the chemical shifts indicated an extended conformation for all three residues, but did not fit any of the two crystallographic

arrangements. Similarly, RDC values calculated from each of the crystal models, or a linear combination of these, did not fit the experimental data. The goodness-of-fit factor Q remained above 0.6 in all cases, whereas values of 0.25-0.40 are typical for fits of crystal structures at 1.8-2.5 Å resolution (Cornilescu et al., 1998). Individual domains also yielded poor fits, especially Z2 (Q factors for RDC of secondary structure elements from Z1 and Z2 were 0.38 and 0.60, respectively) (Fig 2a & b). Furthermore, values for the Da component of the alignment tensor obtained by singular value decomposition (SVD) were underestimated by 27% for Z2 compared to the value obtained from the powder pattern. This indicated a high degree of structural "noise" in the protein models compared to the solution structure (Zweckstetter and Bax, 2002). While the error in the crystallographic coordinates was calculated as 0.27 Å based on a Luzatti plot (Luzatti, 1952), the structural variability of the individual Ig as estimated from the rmsd in atom positions across models was ~0.56 ± 0.10 Å for Z1 (residues 4-98) and ~0.39 ± 0.14 Å for Z2 (residues 102-194) (TOP software; Lu, 2000). This showed that the local geometry within domains varies slightly, possibly justifying the poor agreement between calculated and experimental RDC.

Because of the mismatch in RDC values, the Da and R components of the alignment tensor could not be determined accurately. Thus, we tested the dependence of the calculated domain orientation on these parameters by refining the individual domains against experimental RDC data under an extensive range of values. Da and R quantities were varied stepwise and sweeping the range of values between those determined by SVD and those from the powder pattern. Only RDC data for residues in secondary structure elements were used in the calculation since these were less affected by structural variability. Only well isolated RDC peaks with good intensity were considered (Fig 2a & b). Data were not detectably correlated (Fig 2d). During these calculations, which employed simulated annealing, the crystal structures were kept virtually unchanged by means of geometrical restraints. The influence of force constants for geometrical and experimental RDC terms was assessed, revealing that NH bonds underwent a reorientation of $5^{o}$-$10^{o}$ depending on the force applied. These changes are within the NH-bond variability calculated from the superposition of individual crystal structures which amounted to as much as $30^{o}$ in loops and $14^{o}$ in secondary structure elements. This work proved the calculated orientation of both HN vectors and alignment tensor to be robust, so that domain arrangements obtained by the diverse protocols were

in agreement by 10°. Final orientations were computed by refinement of full structures, where translational drifts between the two Ig were restrained by confining the linker using two further RDC and additional torsion angle restraints derived from the chemical shifts. As for individual Ig, calculated orientations remained within a _10° range (Fig 2c). From all trials performed, a representative family of conformers (n=200) was obtained (Fig 3c), where the alignment tensor calculated from full structures corresponded exactly to the average obtained from individual domains.

The conformers so calculated divided into four clusters (Fig 3c). This segregation arises from the four-fold degeneracy of the RDC values, which are even functions of _ and _ polar angles of the respective HN-bonds and the alignment tensor (see methods, eq. 1). In order to unambiguously determine domain orientations by RDC, measurements in different media are needed that, by comparative analysis, can resolve this degeneracy and allow the selection of a given cluster. In the current case, additional RDC data could not be recorded because Z1Z2 interacted with all other media assayed (5% [w/v] hexaethylene glycol monododecyl-ether [C12E6]/n-hexanol, phages, DHPC/DMPC). None of the four clusters of conformers calculated from polyacrylamide gels coincided with available crystal structures, so that also these could not point to a representative cluster. Finally, we carried out cluster selection by comparison to SAXS data that report on the average conformation of a molecule in solution.

**Average conformation in solution determined by SAXS**

The overall parameters of Z1Z2 derived from SAXS are similar to those reported (Zou et al., 2003) (Fig 3a). The experimental radius of gyration, $R_g$= 3.1±0.05 nm, and maximum size, $D_{max}$= 11.0±1.0 nm, indicate an anisometric structure. This is further confirmed by the profile of the distance distribution function (Fig 3a), typical of elongated particles (Feigin and Svergun, 1987). A low-resolution model of Z1Z2 reconstructed *ab initio* fitted optimally the experimental data ($\chi$=0.74; Fig 3a) and displayed two separated domains in a semi-extended conformation (Zou et al., 2003) (Fig 3b).

In order to validate crystallographic and RDC models their fitting to SAXS data was analyzed (Fig 3a). The scattering calculated from the compact crystal structure failed to fit the experimental scattering ($\chi$=3.63). The extended crystal form yielded a significantly better fit ($\chi$=1.22), but the computed pattern still deviated noticeably from the data. Calculations on the 200 RDC conformers showed that the 50 models with lowest discrepancy values ($\chi$=0.80 – 0.93) constituted a defined cluster (Fig 3c). The best-fitting RDC conformer within this cluster ($\chi$=0.80) was neatly superimposable with the *ab initio* SAXS model computed independently (Fig 3b). The full RDC cluster yielded a fit with $\chi$=0.90, while the best possible mixture of crystallographic models (85% extended plus 15% compact) provided _=1.00, better than those of each model separately but still worse than any of the RDC conformers in the chosen cluster. This suggested that crystal models are worse representatives of the average conformation of Z1Z2 in solution than RDC conformers, whose domain arrangement (Table 2) is more bent than the extended crystal structure (Fig 4a). No improved fit was achieved by assaying populations of mixed crystal and RDC models and best mixtures never contained more than 34% of the compact crystal model, indicating that, if present in solution, compact conformations amount to only a minor fraction of the population. Thus, it can be concluded that the selected RDC cluster is a good representative of Z1Z2 in solution.

**Absence of fast interdomain motions**

Modular flexibility around a hinge region can result in fast interconverting conformations and give rise to uncertainty in domain arrangements calculated by RDC. To assess whether fast domain motions take place in Z1Z2 and to address motion on a per residue basis, we measured $^{15}$N relaxation data (R1, R2) and heteronuclear NOE values at 800 MHz. The latter were evenly distributed around an average of 0.76, at exception of the two N-terminal residues (Fig 3d). The C'D loop region of Z1, which includes a two-residue insertion respect to Z2, displays reduced values of ~0.5, indicating that this is less rigid than the core of the domain. Of the residues in and around the linker region, only T101 shows a reduced NOE value of 0.56. Neither the R2 nor R1 values of the linker residues deviated noticeably from an average value. Correspondingly, the Lipari-Szabo

order parameter $S^2$ (that describes the angular motion for each residue) of linker residue T101 was only slightly diminished from the average value of 0.8-0.9 to a value of 0.75. These data indicate that ultra-fast interconversion of Z1Z2 conformations in solution is unlikely.

The size and shape of the molecular diffusion tensor can also be determined from relaxation data. The diffusion tensor calculated for Z1Z2 was clearly anisotropic whereas a flexible domain arrangement would have yielded more isotropic values. Unfortunately, the quality of our data did not allow us to determine motional parameters on a per residue basis or to calculate domain arrangements from the individually determined diffusion tensors. Instead, we used relaxation values to estimate the mobility of Z1Z2 through comparison to other proteins.

This approach has been used previously in the study of Fn9-Fn10 from fibronectin (Spitzfaden et al., 1997). There, the $^{15}N$ relaxation behaviour of a native construct was compared to that of engineered variants with increasing number of glycine linker residues that progressively abolished the native contacts between domains. In that work, both T1 and T2 values varied significantly with linker length, reflecting an increase in modular flexibility and, hence, the comparative rigidity of the native protein. For Z1Z2, average R1 and R2 values of 1/1000ms and 1/60 ms, respectively, are consistent with those from a slightly elongated linker construct (2 added residues) of the FnIII doublet. The values are also fully consistent with those measured for an I-band pair from titin, I91-I92 (formerly I27-I28) (Improta et al., 1998). In Z1Z2, the overall correlation time $\tau_c$ of 14.2 ns (TENSOR2; Dosset et al., 2000) is below that expected for a rigid domain arrangement (~20 ns) but distinctly different from the value of a mobile two-domain construct (~9 ns) (Spitzfaden et al., 1997). Based on this comparison, Z1Z2 appears to tumble as a fairly rigid body in solution in the nanosecond timescale. Its linker exhibits certain rigidity, even that this is not as pronounced as that of Fn9-Fn10.

**DISCUSSION**

Here we present a comprehensive, mechanistic analysis of the conformational dynamics of an Ig-doublet from titin. This is based on the comparison of crystallographic models, RDC conformers, $^{15}$N NMR relaxation values and low resolution SAXS data. Crystal structures reveal the atomic details of the molecule, particularly of linker groups, but in this study the presence of mixed conformers within one same asymmetric unit hints at the influence of external forces on domain arrangement. RDC data allow determining unbiased interdomain orientations in solution - albeit at a four-fold degeneracy - but require the availability of a structural model for ready interpretability. To resolve the RDC degeneracy, independent measurements are needed that confirm the calculated domain orientation. In this study, this was achieved using SAXS data, which on their own only allow a global description of the average molecular shape. To assess whether such averaged conformations in solution result from possible dynamic-averaging due to molecular flexibility or correspond to conformationally monodisperse samples, NMR relaxation methods can be applied that measure motions in a polypeptide chain on a per residue basis in the subnanosecond time scale. The study reported in this manuscript combines all these methodologies to conclude on the conformational properties of Z1Z2.

Our findings suggest that Z1Z2 adopts a preferential conformation in solution, even if the length of its three-residue linker (_10 Å) and the lack of Ig-Ig contacts could well allow for complex modular motions. Structural models calculated independently using RDC and SAXS agree excellently in revealing Z1Z2 in a semi-extended conformation (Table 2; Fig 3b). Structural noise (Zweckstetter and Bax, 2002) and/or dynamical averaging (Bertini et al., 2004) can account for a reduced consistency between crystal structures and RDC values as seen here. The conformation calculated for Z1Z2, however, is unlikely to merely correspond to an average of fast interconverting forms in solution, since $^{15}$N relaxation data did not reveal fast motions within the linker region (Fig 3d) and values are consistent with at least a semi-rigid arrangement.

The fact that little structural restrictions oppose interdomain re-arrangements in Z1Z2 suggests that this might undergo moderate or even long-range motions either derived from relatively slow dynamics (below the NMR time scale) or adaptation to a local

environment, for example an induced-fit to a binding partner or a crystalline lattice. The several Z1Z2 conformations identified to date (Fig 4a) confirm that these events do, in fact, take place. The extended crystal structure of Z1Z2 identified in this study is closely related to the preferential state in solution but it has a wider hinge opening (Table 2), likely to reflect an accommodation to lattice interactions. The compact crystal model, however, appears to represent a lowly populated state in solution. The fact that this is enriched in the crystal indicates that it is favoured by the crystallization medium and/or selectively incorporated into the lattice. Considering that the conformational energy of Z1Z2 must account for hinge geometry as well as Ig-Ig interactions, it is feasible that cadmium binding, the only prominent interdomain contact in this case, could have driven and locked the molecule into this state. This metal, abundant in the crystallization medium, is unlikely to be physiologically relevant. This is supported by the fact that metal ions were not found in the crystal structure of Z1Z2 complexed to telethonin (Zou et al., 2006).

Z1Z2 bound to telethonin is fully extended (Zou et al., 2006) (Table 2, Fig 4a), probably resulting from an induced-fit of Z1Z2 permitted by its adaptable linker. The interaction is highly selective, even that it merely relies on main chain groups that mediate the formation of an intermolecular β-sheet and, as shown by mutagenesis (Zou et al., 2006), it is sequence independent. In titin, numerous Ig are encountered along the filament, where pairs with similar linker length and close structural similarity to Z1 and Z2 are abundant, particularly in the skeletal I-band section (Fig 4c) (Marino et al., 2005). Partner recognition under these conditions could be expected to involve primarily steric criteria. Hence, we speculate that the pre-existent conformation of Z1Z2, where interacting surfaces are in a productive spatial arrangement, must contribute to an effective and selective binding, whereas free inter-Ig motions would render such complexation sterically and entropically unfavourable. This hypothesis is supported by recent dynamic simulations (Lee et al., 2006) in which crystallographic coordinates of Z1Z2 in its complexed state were freed from the telethonin partner and equilibrated on its own in solution. Results showed how the linker region changed its geometry from fully extended to slightly bent, so that the angle between domains decreased Conformational dynamics of Z1Z2
to 137° at the end of the simulation, which is in agreement with the experimental conformation of Z1Z2 in solution calculated in the current study (Table 2).

The comparative analysis of Z1Z2 conformations reveals the molecular mechanisms of linker adaptability and, thus, of interdomain motions. In the conformations so far identified, variations in hinge opening are associated with domain torsions, as revealed by the direction of the inter-rotation axes (Fig 4a). Structures of free Z1Z2 show that an acute domain closure can be achieved by rotation around a single main chain bond. On the other hand, interconversion of the several extended arrangements involves only slight rotations of the linker groups. Interestingly, both in the acute hinge closure and the extended domain variations observed in this study, residue A99 constitutes the central mechanical point of torsion (Fig 4b), probably because of the absence of a hindering side chain. Even that moderate domain re-arrangements across extended conformations involve only minor alterations of the linker, NMR relaxation data show that modular dynamics is restricted in Z1Z2. This is unusual, given the absence of Ig-Ig contacts and the relatively long linker. Data available on other poly -Ig and -Fn proteins indicate that, commonly, modules within tandems stabilize their respective orientations either by close domain-domain interactions, such as in fibronectin (Leahy et al., 1996), or by stiffening the conformation of its linker sequences, as cadherin which uses calcium binding (Boggon et al., 2002). The linker of Z1Z2 appears to be semi-rigid, deformable, but not freely flexible. This might be due to its minor hydrophobic component (Fig 1b), which could be speculated to act as a "greasy" domain interface allowing viscous domain oscillations but preventing free motions. Extreme domain rearrangements such as that of the compacted Z1Z2 alter the disposition of hydrophobic groups significantly and, in the absence of compensatory inter-domain forces, might be unfavourable.

Long poly-Ig tandems predominate in the I-band of titin (Fig 4c). Their elastic properties are central to the passive mechanics of the sarcomere (references within Granzier and Labeit, 2004). Understanding the mechanical function of these tandems intrinsically requires knowledge of their conformation and dynamics. Z1 and Z2 share high structural resemblance to modules in the prominent skeletal I-band fraction (Fig 4c), which are characterized by conserved sequence motifs at their N-terminal loop cluster (for a description of Ig and linker conservation at the I-band see Marino et al., 2005). Furthermore, the linker length of Z1Z2 coincides with that of the regularly interspersed super-repeats pairs (Fig 4c). NMR relaxation data on Z1Z2 are consistent with those reported for I91-I92 (formerly I27-I28) from the distal I-band (Improta et al., 1998), suggesting that I-band modules and Z1Z2 exhibit comparable dynamics even that I91 and I92 do not retain conservation of its N-terminal loop cluster. As in Z1Z2, I91-I92 appeared to show a small degree of bending respect to a fully linear conformation, with

an estimated hinge opening value of ~155° according to SAXS data (Improta et al., 1998). Furthermore, calculations on a six-Ig fragment from the skeletal I-band, I65-I70, based on electron microscopy and SAXS data and following a freely-joint chain polymer model suggested a certain stiffness of the chain and an average opening angle between any two-Ig in this fragment of ~143° (Marino et al., 2005), in agreement with observations in the current study. This indicates that the conformational dynamics of Z1Z2 might be closely related to that of I-band Ig-tandems. Consequently, modules within the I-band tandems of titin might resemble Z1Z2 in constituting adaptable, modular formations with extended domain arrangements, characterized by a restricted and slow dynamics - probably reflecting a certain viscosity due to small hydrophobic interdomain surfaces. As in Z1Z2, it could be expected that at least the longer linkers (3-4 residues) of I-band pairs might be able to adopt a preferential conformation at equilibrium in solution but could undergo long-range motions, likely to involve domain torsions, when subjected to directed stretch-contraction forces in the sarcomere. Additional data on poly-Ig arrays of titin will have to become available to confirm whether this might constitute a generic conformational principle of the structure and mechanics of this filament.

## METHODS

### Crystal structure determination

The cloning, expression and purification of Z1Z2 from human titin (residues 1-196; EMBL X90568) have been recently reported (Zou et al., 2003). Crystallization used the hanging-drop method. First, thin needles grew from 100 mM MES pH 6.5, 50 mM $CdSO_4$ and 1 M sodium acetate at a protein concentration of 30 mg/ml. The needles were used for microseeding of solutions containing 100 mM MES pH 5.5, 200 mM $CdSO_4$, 400 mM sodium acetate and 1520 mg/ml protein. Crystals grew as clusters of thin plates (thickness ~15 $\mu$m). Single crystals were excised from the clusters.

For X-ray data collection, crystals were frozen at 100K in mother liquor solutions

supplemented with 30% [v/v] glycerol as cryoprotectant. Although initial diffraction was measurable up to 1.5 Å resolution, radiation damage limited effective recordings to 2.0 Å and required frequent recentering to irradiate fresh portions of the crystal. Data processing used MOSFLM/SCALA (Leslie, 1992; Evans, 1993). Phasing was by molecular replacement using a poly-alanine model derived from the single Ig domain of telokin (1FHG) (approx. 40% seq. id. to Z1 or Z2 and accounting for 1/6 of the asymmetric unit) as search model in AmoRe (Navaza, 2001). Five Ig domains could be located, resulting in a correlation coefficient of 30.5% and R-factor of 49.9%. The sixth domain was revealed in electron density maps after initial model refinement had taken place. Manual model building used O (Jones et al., 1991), while refinement employed overall anisotropic B-factor scaling, bulk solvent correction and restrained atomic B-factor minimization in CNS (Brunger et al., 1998). Non-crystallographic symmetry restraints were applied across individual domains during the early stages of refinement and progressively released. Experimental data were divided into a working and a test set using FREERFLAG (Collaborative Computational Project N. 4, 1994). The solvent was built with ARP/warp (Perrakis et al., 2001) and validated visually in electron density maps. The final model includes all protein residues for each Z1Z2 copy in the asymmetric unit.

Model coordinates and structure factor amplitudes have been deposited at the Protein Data Bank with accession code 2A38.

**NMR spectroscopy**

Spectra were acquired in 20 mM sodium phosphate buffer pH 6.8, 100 mM NaCl at a protein concentration of 0.4 mM and 22°C. Spin-state-selective $^{15}$N-TROSY experiments without $^{1}$H decoupling were recorded on a $^{15}$N-labeled sample to extract HN–N dipolar coupling at 22°C on either a Bruker DMX600 or DMX500 spectrometer, both equipped with a cryoprobe. A laterally compressed 7% polyacrylamide gel consisting of 50% positively charged copolymer (cross-linking ratio 20:1): (3-acrylamidopropyl)-

trimethylammonium (Ulmer et al., 2003; Cierpicki and Bushweller, 2004) with a compression ratio of 17/23 was used as alignment medium.

*Relaxation and Dynamics*

Standard sets of $^{15}$N R1, R2, and ($^{1}$H-$^{15}$N)NOE relaxation experiments were measured at 800 MHz. The program TENSOR2 (Dosset et al., 2000) was used for data analysis.

*NMR structure evaluation and calculation*

RDC values can be calculated from the polar angles of the respective HN-bonds in the alignment tensor frame, following:

*RDC=Da*((3cos_ _1)*1.5**R*sin_ cos2_)(eq. 1), where *Da* and *R* specify the axial component and the rhombicity of the tensor.

Values for Da and R were obtained either by SVD (Losonczi et al., 1999) or optimized from a systematic grid search. A modification of the program Aria1.2/CNS was used for structure calculation (Linge et al., 2001). This employs a non-crystallographic symmetry term to harmonically restrain the coordinates of the individual domains to the respective crystal structures. An artificial tetra-atomic molecule representing the axes of the alignment tensor was allowed to float during the torsion angle dynamics period of the refinement protocol (Clore et al., 1998). This allows reorientation of the domains to match the experimental RDC during simulated annealing while a minor local reorientation of the HN vectors is permitted. Hydrogens were added to the crystal structures with CNS (Brunger et al., 1998). The final calculation of the domain orientation used common Da and R values of -7 and 0.5 obtained from the powder pattern, but results were essentially unchanged if other and/or two separate values were employed. Torsion backbone angles calculated from chemical shifts (Cornilescu et al., 1999) were added to restrain the conformation of the linker. 500 conformers were computed in each calculation trial, of which the best 200 were refined in an explicit water shell.

**SAXS**

SAXS data were collected at X33, EMBL (DESY, Hamburg) using a linear gas detector and at a wavelength _=1.5 Å. Solutions of Z1Z2 were measured at concentrations of 2.0, 5.0 and 10.0 mg/ml. Data acquisition and processing was as reported (Zou et al., 2003). The low resolution model of Z1Z2 was constructed *ab initio* using the program GASBOR (Svergun et al., 2001). The scattering from crystal structures and RDC models was calculated using CRYSOL (Svergun et al., 1995). The volume fractions of mixed populations of conformers were computed with OLIGOMER (Konarev et al., 2003).

**Acknowledgements**

Very special thanks go to C. Schulze and T. Tomizaki for exceptional assistance during X-ray data collection. We are grateful to S. Grzesziek, L. Duma and D. Häussinger (Biozentrum, Basel) for helpful discussion and critical reading of this manuscript. C.M.-G. is an Emmy Noether-fellow of the DFG (Mu-1606-1).

**REFERENCES**

Bax, A., Kontaxis, G., and Tjandra, N. (2001). Dipolar couplings in macromolecular structure determination Methods in Enzymology 339, 127-174.

Bertini, I., Del Bianco, C., Gelis, I., Katsaros, N., Luchinat, C., Parigi, G., Peana, M., Provenzani, A., Zoroddu, M.A. (2004). Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. Proc Natl Acad Sci U S A. 101, 6841-6.

Boggon, T.J., Murray, J., Chappuis-Flament, S., Wong, E., Gumbiner, B.M. & Shapiro, L.C. (2002). Cadherin ectodomain structure and implications for cell adhesion mechanisms. Science 296, 1308-1313.

Brunger, A.T., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D54, 905-921.

Cierpicki, T., and Bushweller, J.H. (2004). Charged gels as orienting media for measurement of residual dipolar couplings in soluble and integral membrane proteins. J Am Chem Soc 126,16259-16266.

Clore, G.M., Gronenborn, A.M., and Tjandra, N. (1998). Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. J Magn Reson 131, 159-162.

Collaborative Computational Project, N. 4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D50, 760-3.

Cornilescu, G., Delaglio F., and Bax A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology J Biomol NMR 13, 289–302.

Cornilescu, G., Marquardt, J.L., Ottiger, M., and Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J Am Chem Soc 120, 6836-6837.

Dosset, P, Hus J.C., Blackledge M, and Marion D. (2000). Efficient analysis of

macromolecular rotational diffusion from heteronuclear relaxation data. J Biomol NMR 16, 23-28.

Edlich, C., and Muhle-Goll, C. (2003). 1H, 13C, and 15N backbone assignment of the first two Ig domains Z1Z2 of the giant muscle protein Titin. J Biomol NMR 27, 283-284.

Evans, P.R. (1993). Data reduction. Proceedings of CCP4 Study Weekend on Data Collection and Processing, 114-122.

Feigin, L.A. and Svergun D.I. (1987). Structure analysis by small-angle x-ray and neutron scattering. New York, Plenum Press.

Granzier, H.L., and Labeit, S. (2004). The giant protein titin: a major player in myocardial mechanics, signaling, and disease. Circ Res 94, 284-295.

Gregorio, C.C. et al. (1998). The NH2 terminus of titin spans the Z-disc: its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity. J Cell Biol 143, 10131027.

Harpaz, Y., and Chothia, C. (1994). Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. J Mol Biol 238, 528-539.

Hayward, S., and Lee, R.A. (2002). Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. J Mol Graph Model 21, 181-183.

Improta, S., Krueger, J.K., Gautel, M., Atkinson, R.A., Lefevre, J.F., Moulton, S., Trewhella, J., and Pastore, A. (1998). The assembly of immunoglobulin-like modules in titin: implications for muscle elasticity. J Mol Biol 284, 761-777.

Jones, T.A, Zou, J.Y., Cowan S.W and Kjeldgaard M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst. A47, 110-119.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis J Appl Crystallogr 36, 1277-1282.

Leahy, D.J., Aukhil, I., and Erickson, H.P. (1996). 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and

synergy region. Cell 84, 155-64.

Lee, E.H., Gao, M., Pinotsis, N., Wilmanns, M. and Schulten, K. (2006). Mechanical strength of the titin Z1Z2-telethonin complex. Structure *14*, 497-509.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Prot Crystallogr.

Linge J.P, O'Donoghue S.I., and Nilges M. (2001). Automated assignment of ambiguous nuclear overhauser effects with ARIA. Methods Enzymol 339, 71–90.

Losonczi, J.A, Andrec, M., Fischer, M.W., and Prestegard, J.H. (1999). Order matrix analysis of residual dipolar couplings using singular value decomposition. J Magn Reson 138, 334342.

Lu, G. (2000). TOP: a new method for protein structure comparisons and similarity searches. J Appl Crystallogr 33, 176-183.

Luzatti, V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. Acta Crystallogr 5, 802-810.

Marino, M., Svergun, D.I., Kreplak, L., Konarev, P.V., Maco, B., Labeit, D., and Mayans, O. (2005). Poly-Ig tandems from I-band titin share extended domain arrangements irrespective of the distinct features of their modular constituents. J Muscle Res Cell Motil, 1-11.

Mues, A., van der Ven, P.F., Young, P., Furst, D.O., and Gautel, M. (1998). Two immunoglobulin-like domains of the Z-disc portion of titin interact in a conformation dependent way with telethonin. FEBS Lett 428, 111-114.

Navaza, J. (2001). Implementation of molecular replacement in AMoRe. Acta Crystallogr D57, 1367-1372.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. Acta Crystallogr D57, 1445-1450.

Politou, A.S, Gautel, M., Improta, S., Vangelista, L., Pastore, A. (1996). The elastic I-band region of titin is assembled in a "modular" fashion by weakly interacting Ig-like domains. J Mol Biol 255, 604-616.

Scott, K.A., Steward, A., Fowler, S.B., & Clarke, J. (2002). Titin; a multidomain protein that behaves as the sum of its parts. J Mol Biol 315, 819-829.

Spitzfaden, C., Grant, R.P., Mardon, H.J., and Campbell, I.D. (1997). Module-module

interactions in the cell binding region of fibronectin: stability, flexibility and specificity. J Mol Biol 265, 565-579.

Svergun, D.I., Barberato, C., and Koch, M.H.J. (1995). CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates J Appl Crystallogr 28, 768-773.

Svergun, D.I., Petoukhov, M.V., and Koch, M.H. (2001). Determination of domain structure of proteins from X-ray solution scattering. Biophys J. 80, 2946-53.

Tjandra, N., Omichinski, J.G., Gronenborn, A.M, Clore, G.M., and Bax, A. (1997). Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution. Nat Struct Biol 4, 732-738.

Ulmer, T.S., Ramirez, B.E., Delaglio, F., and Bax, A. (2003). Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. J Am Chem Soc 125, 9179-9191.

Zou, P., Gautel, M., Geerlof, A., Wilmanns, M., Koch, M.H., and Svergun, D.I. (2003). Solution scattering suggests cross-linking function of telethonin in the complex with titin. J Biol Chem 278, 2636-44.

Zou, P., Pinotsis, N., Lange, S., Song, Y.H., Popov, A., Mavridis, I., Mayans, O., Gautel, M., and Wilmanns, M. (2006). Palindromic assembly of the giant muscle protein titin in the sarcomeric Z-disk. Nature 439, 229-233.

Zweckstetter, M. and Bax, A. (2002). Evaluation of uncertainty in alignment tensors obtained from dipolar couplings. J Biomol NMR 23, 127-137.

## FIGURES

**Figure 1: Crystal structure of Z1Z2**

**a.** Superimposition of open- and close-hinge crystallographic conformations. To enhance visualization of the domain torsion, _-sheets of Z1 are coloured differently. Metal ions and coordinating protein residues are displayed. The rotation axis is shown, where the magnitude of the rotation is ~155° (determined using DYNDOM; Hayward and Lee, 2002); **b.** Hydrophobic component of the linker region of open Z1Z2, where non-polar groups are coloured green; **c**. $(2F_{obs}-F_{calc})\_{calc}$ electron density maps contoured at 1_ for the open (above) and close (below) hinge conformations; **d.** Ramachandran plot for linker residues in open- (black) and closed-(red) conformations;

**Figure 2**: **RDC analysis**

Comparison of experimental and calculated RDC values for **a.** Z1 and **b.** Z2 domains. Experimental data are shown using a continuous line and small circles. The dotted line represents calculated average values for the individual Ig of all three crystallographic copies (Da=-7; R=0.5). Error bars indicate the corresponding lowest and highest values. **c**. Comparison of domain orientations in models best satisfying RDC data resulting from diverse refinement protocols. Backbone atoms are shown. **d**. Globe of NH vector distribution (dotted lines). The alignment frame is depicted as solid lines, where z-, x- and y- axes are denoted by a black, red and green sphere, respectively.

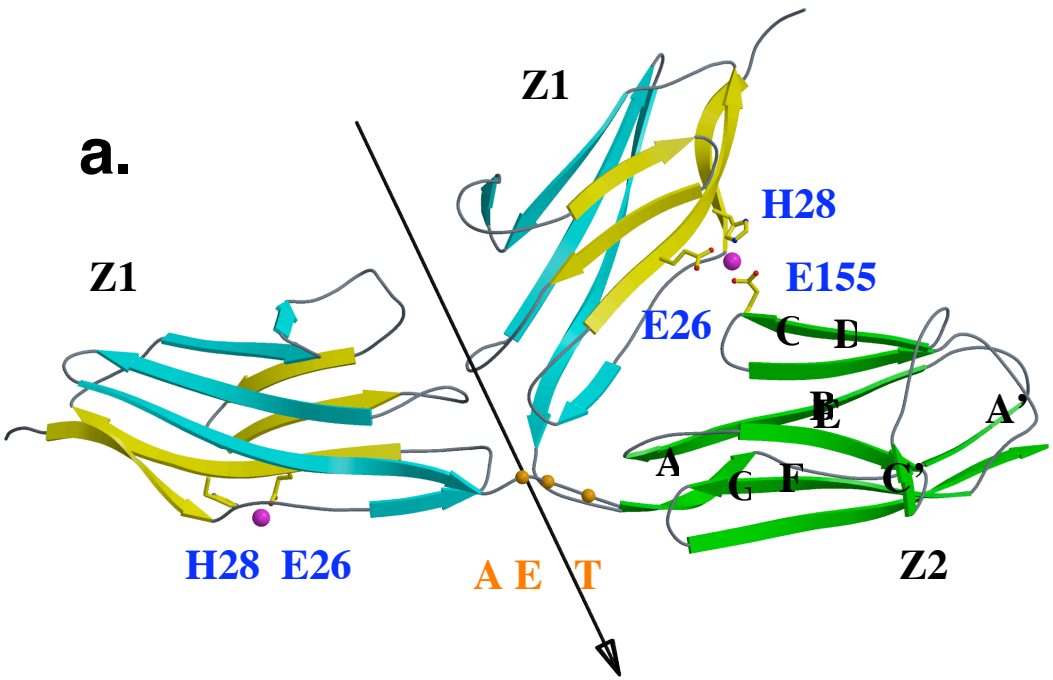
**Figure 3: Conformational analysis of Z1Z2 in solution**

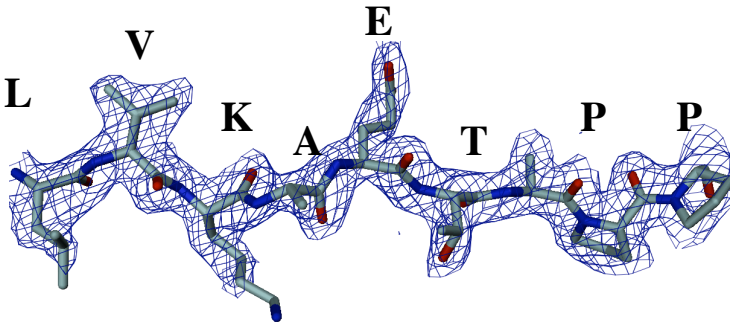
**a.** Scattering patterns from Z1Z2 sample and its models: (1) experimental data; scattering computed from (2) ab initio sphere model, (3) extended crystal model, (4) compact crystal model, and (5) best SAXS-fitting RDC model. Insert: distance distribution of Z1Z2; **b**. ab initio model of Z1Z2 (golden spheres) superimposed on the best SAXS-fitting RDC model (wireframe). The views shown correspond to a -

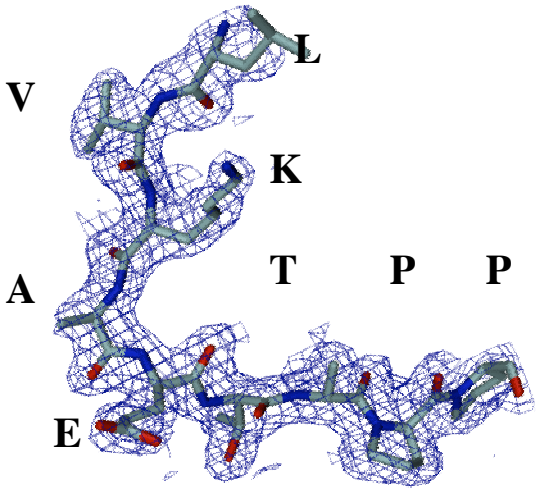
90° rotation around the vertical axis; **c**. Family of RDC conformers (n=200) superimposed on Z1 composed of four defined conformational clusters. Conformers are colour coded according to their fitting discrepancy to SAXS data as follows: red indicates the best fitting cluster with $\chi$ values of 0.80-0.93, green corresponds to $\chi$=0.96-1.19; blue to $\chi$=1.3-1.39; cyan to $\chi$=1.20-1.50 and black to $\chi$>2.22; c. Heteronuclear NOE values per residue. Secondary structure elements are shown in cyan. Linker groups are in red.

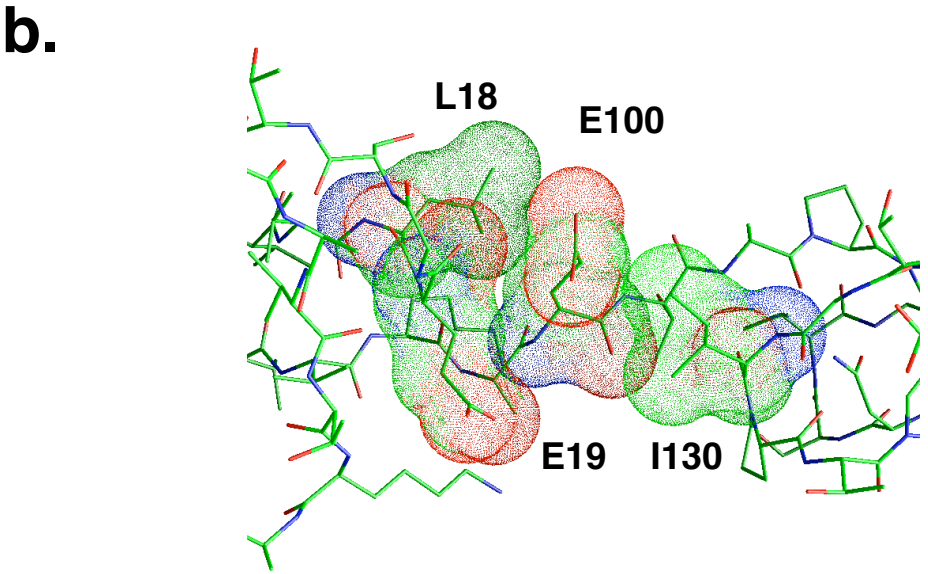
**Figure 4: Inter-domain motions in Z1Z2 and relation to Ig-tandems of the I-band.**
**a.** Stereo representation of models of Z1Z2 superimposed on Z2 (green). Z1 is coloured coded so that dark brown represents the telethonin-bound form (9) (1YA5, where the main area of interaction to telethonin is coloured purple), red shows the crystallographic open form, orange the best SAXS-fitting RDC model and pink the closed crystal model. Domain arrangements are described in Table 2. Inter-rotational axes are displayed, where the axis in black relates telethonin-bound and crystal open Z1Z2; the axis in grey relates the open crystal structure to the best SAXS-fitting RDC conformer; and the axis in blue relates open and closed crystal models; **b**. Detailed view of the linker region of crystal structures, where rotational axes are coloured as in **a.**; **c.** Poly-Ig composition of the I-band of titin, where Ig domains and linker segments are represented by filled boxes. For simplicity, N2A, N2B and PEVK elements are excluded. Constitutively and differentially expressed I-band fractions are indicated, the latter being prominent in skeletal muscle. Ig domains coloured orange are of the same type that Z1 and Z2, sharing a conserved N-terminal loop cluster (as described in Marino et al., 2005). Linkers depicted in light blue consist of very short sequences (0-1 residue), while longer linkers of comparable length to that of Z1Z2 are shown in dark blue (3-4 residues). Super-repeats are indicated by a capped-bar. Modules for which an atomic structure has been elucidated are marked by a thick bar. Fragments boxed correspond to I91-I92, whose domain arrangement has been analyzed by SAXS and its relaxation times by NMR, and I65-I70, whose conformational parameters has been studied by SAXS and electron microscopy.

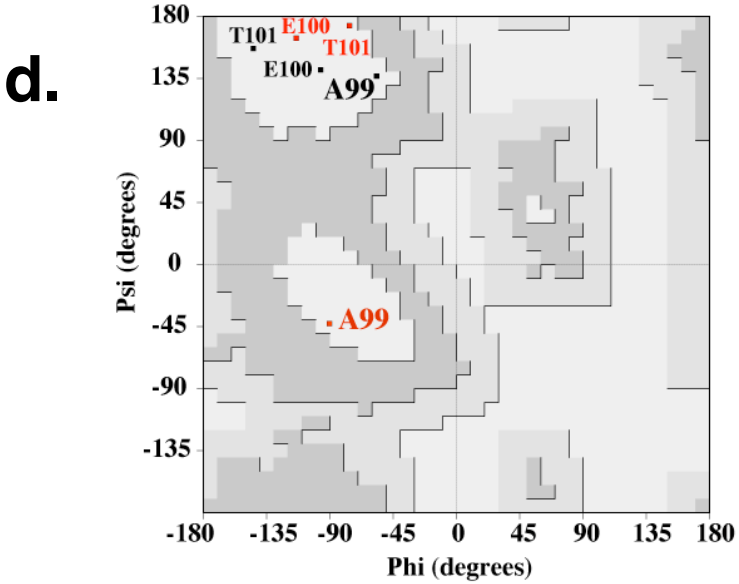
# Fig 1


a.
Z1
Z1
H28
E155
E26
C
D
E
A'
A
G
F
C'
H28
E26
A E T
Z2

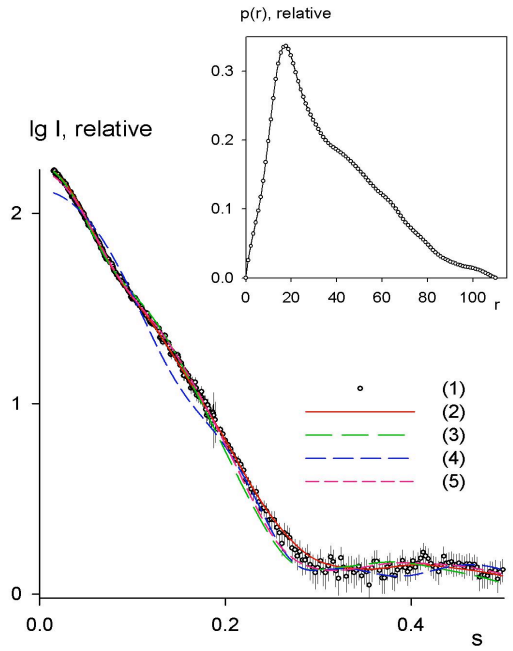

c.
V
E
L
K
A
T
P
P


L
V
K
T
P
P
A
E

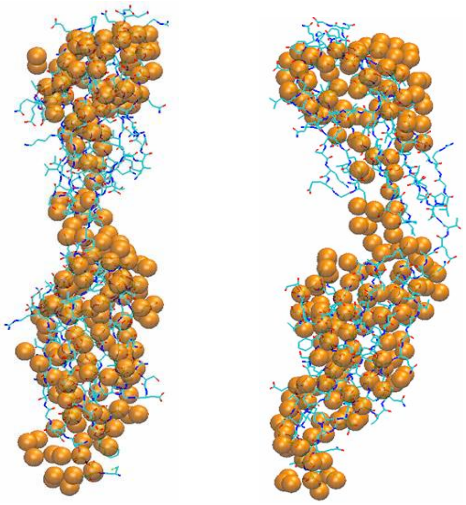

b.
L18
E100
E19
I130


d.
T101
E100
T101
E100
A99
A99
180
135
90
45
0
-45
-90
-135
-180
-135
-90
-45
0
45
90
135
180
Psi (degrees)
Phi (degrees)

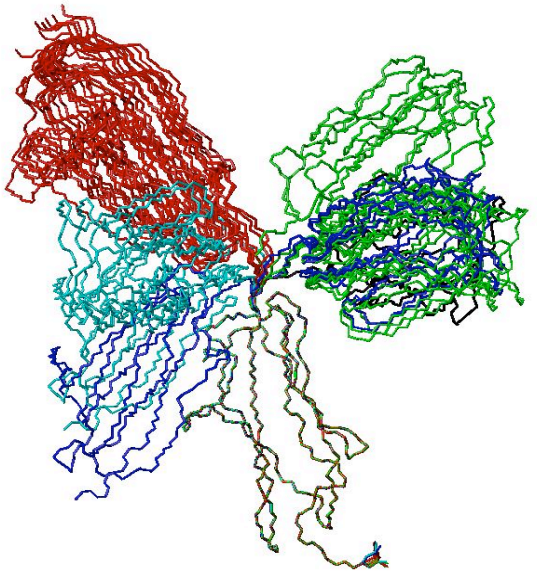
**a.**


p(r), relative
lg I, relative
s
r
(1)
(2)
(3)
(4)
(5)

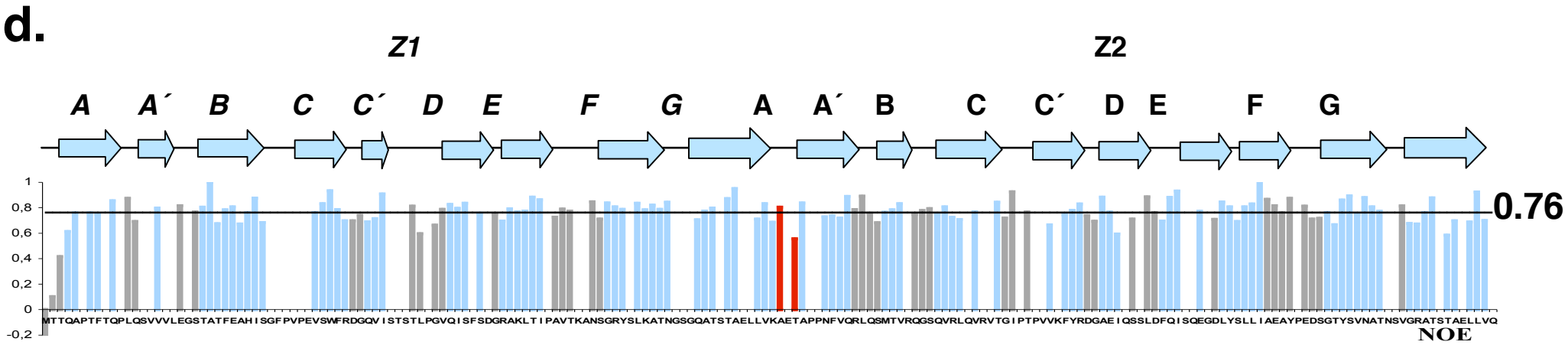
**b.**



**c.**

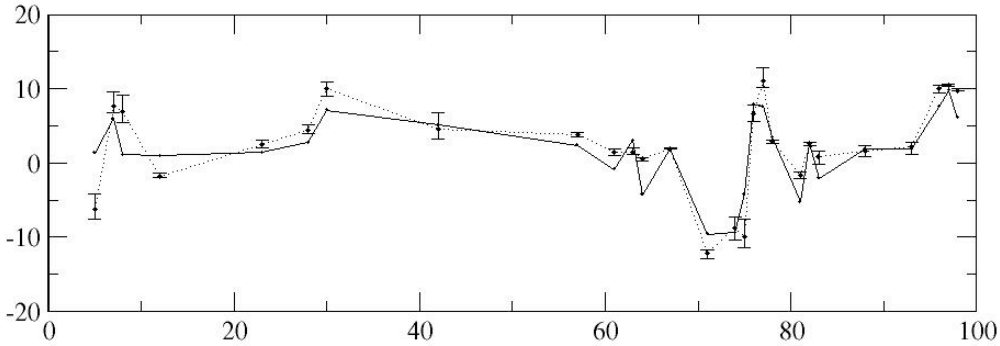


**d.**

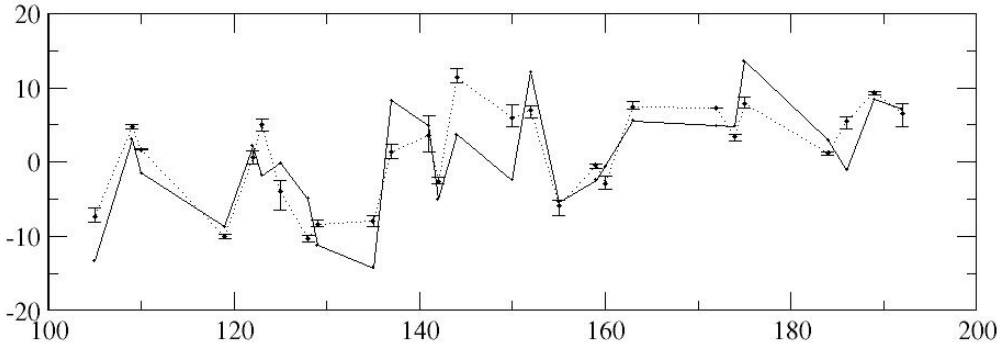

Z1
Z2
A A´ B C C´ D E F G A A´ B C C´ D E F G
0.76
MTTQAPTFTQPLQSVVVLEGSTATFEAHISGFPVPEVSWFRDGQVISTSTLPGVQISFSDGRAKLTIPAVTKANSGRYSLKATNGSGQATSTAELLVKAETAPPNFVQRLQSMTVRQGSQVRLQVRVTGIPTPVVKFYRDGAEIQSSLDFQISQEGDLYSLLIAEAYPEDSGTYSVNATNSVGRATSTAELLVQ
NOE

(Fig 2)

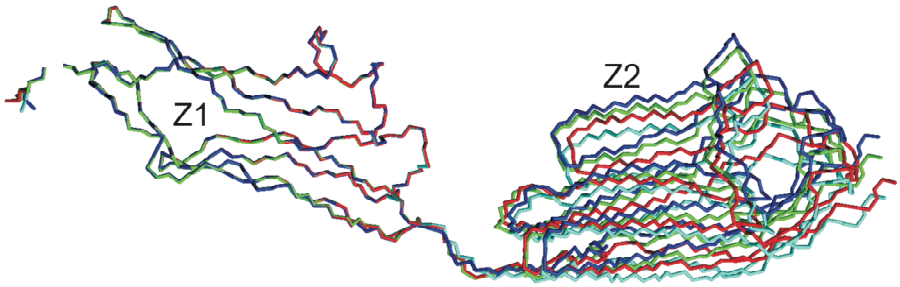
a.


20
10
0
-10
-20
0
20
40
60
80
100

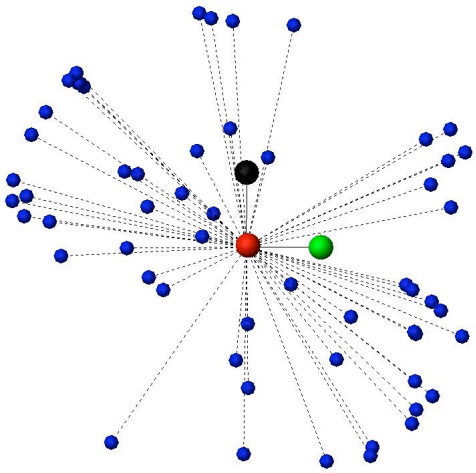
b.


20
10
0
-10
-20
100
120
140
160
180
200

c.


Z1
Z2

d.



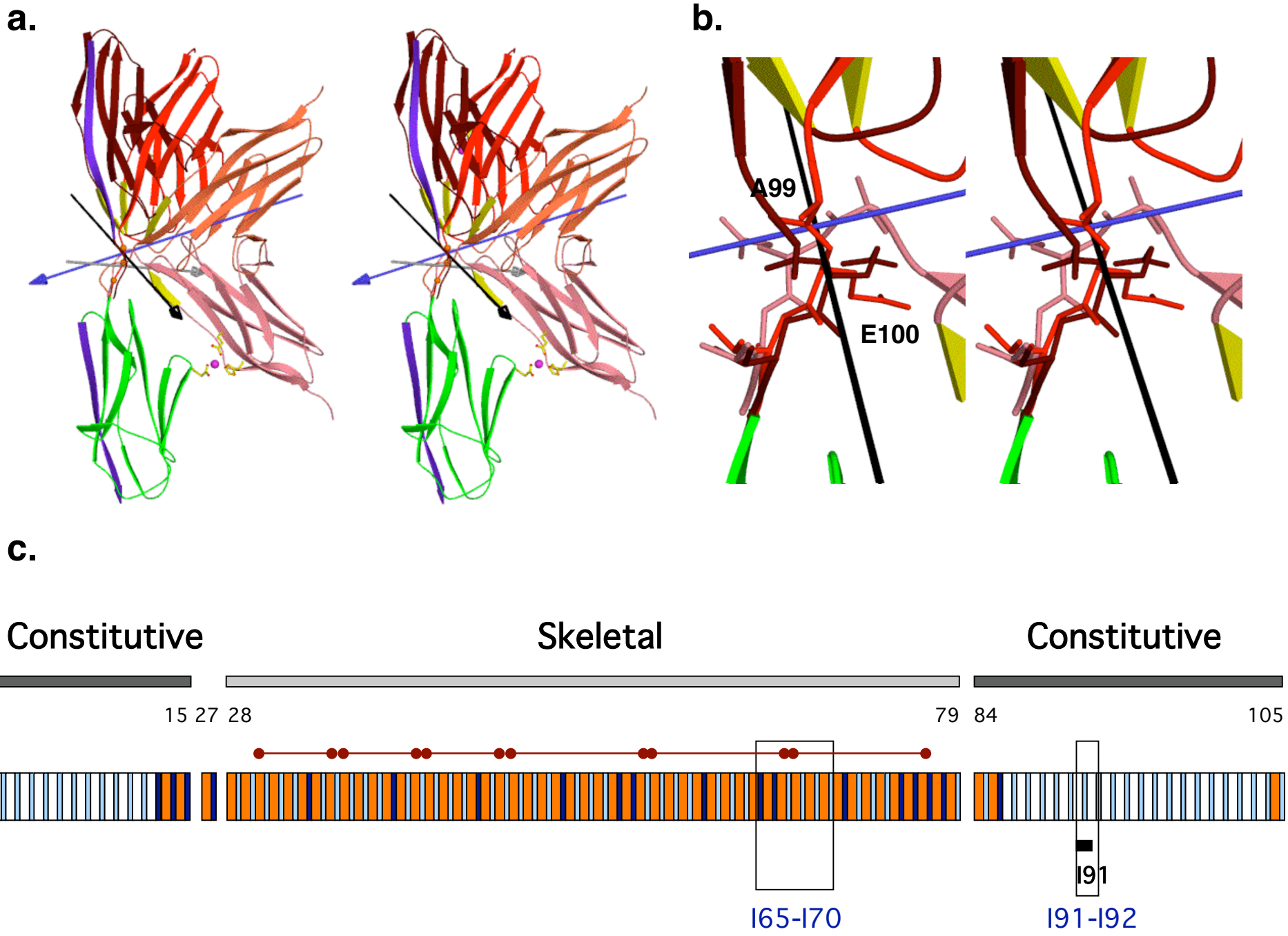

a.
b.
A99
E100
c.
Constitutive
Skeletal
Constitutive
1
15 27
28
79
84
105
I1
I65-I70
I91
I91-I92

**Table 1: X-ray data statistics and model refinement parameters**

| | |
|---|---|
| Space group | P1 |
| Unit Cell (a, b, c) | 55.40 Å, 56.29 Å, 74.41 Å |
| ($\alpha$, $\beta$, $\gamma$) | 71.96°, 88.49°, 75.41° |
| Beamline | X06SA (SLS) |
| Wavelength | 0.920 Å |
| Resolution (Å) | 20 – 2.0 (2.05 – 2.0) |
| Unique reflections | 51011 (2876) |
| Completeness (%) | 91.4 (69.4) |
| $R_{sym}$(I) (%) | 9.2 (34.9) |
| $\langle I \rangle / \sigma(I)$ | 8.9 (2.8) |
| Redundancy | 1.8 (1.4) |
| | |
| R-factor / $R_{free}$(%)[&] | 20.91 /25.46 |
| No. of Protein residues | 582 |
| No. of Solvent / Metal atoms | 457 / 3 |
| Ramachandran plot | |
| Favored / disallowed regions | 93.2% / 0% |
| Rmsd bonds (Å) / angles ($^o$) | 0.007 / 1.45 |

[&] The R-free set comprised 1296 reflections corresponding to 2.3% of the total data.

**Table 2: Domain arrangements in Z1Z2 models**

| | Open form | Compact form | RDC cluster | best SAXS-fitting RDC | Z1Z2:telethonin |
|---|---|---|---|---|---|
| Hinge opening angle (°) | 146.3 | 36.7 / 37.6 | 131.2 ± 10.3 | 136 | 167 |
| Domain torsion (°) | +75 | -99 | +89.5 ± 10.2 | +85 | +48 |

# Summary

The primary objective of this study was to investigate the complex mechanism of regulation of DMPK. CCs are thought to form dimeric arrangements in the DMPK-related family of kinases and thus to mediate dimerization in these enzymes. We analyzed the oligomeric state of the CC domain of DMPK using analytical ultracentrifugation and established that, in isolation, this motif forms trimers as unique species. Its crystallographic structure was elucidated to analyze inter-chain interactions and establish the molecular principles of this unexpected level of association. Using size exclusion chromatography combined with multi-angle static light scattering, we analyzed the oligomeric state of DMPK as well as that of a truncated variant lacking the CC region. Interestingly, both forms were found to associate into robust dimers. In contrast, DMPK-CC in isolation is a trimeric motif, indicating that the self-assembly of DMPK is not dictated by the association properties of its CC domain. Instead, it appears to be driven by sequence segments flanking both N and C termini of the catalytic kinase fraction, as suggested by models of head-to-head dimers based on small angle X-ray scattering data. In order to decipher the presence of a possible structural element C-terminal to the kinase domain, native PAGE and sedimentation velocity analyses of an N-terminally extended DMPK-CC were performed, however such motif was not identified. In addition, the absence of interaction between DMPK-CC and other domains in DMPK, suggested that a CC association/dissociation regulatory model such as that proposed for MRCK and ROCKs might not be applicable for DMPK. Thus, we speculate that the oligomeric state of this CC within the context of full DMPK must be dominated by the prior association of the N-terminal catalytic fraction into dimers. In addition, DMPK is unlikely to interact with Rac1 through the CC region under the condition tested. A possible explanation might be that the GTPase-kinase interaction needs of any region in DMPK out of the CC domain to be effective or even, the level of assembly of DMPK-CC in isolation might not be the same than in the context of the full-kinase. Given the heterogeneity in CC sequences and their diversified interaction patterns across MDPK-related family of kinases, studies are required to reveal the exact role of these moieties and to establish the principles of their regulatory mechanisms.

# Appendix

## Curriculum Vitae

# PILAR GARCIA HERMOSA

**Personal details**

Address: Biozentrum, Uni Basel, Klingelbergstr. 50-70, CH 4056, Basel.
Phone: +41 61 267 2082
Email: Pilar.Garcia@stud.unibas.ch
Nationality: Spanish
Date of birth: March 24, 1978

**Education**

**PhD Biophysics**, Prof O. Mayans Group, Division of Structural Biology.
2004–Present, Biozentrum, University of Basel, Switzerland.

*Thesis:* Molecular insights into the regulatory interactions of Dystrophia Myotonica Protein Kinase.

**Research Assistant** in Molecular Structural Biology.
7/2002–9/2003, Division of Structural Biology.
Biozentrum, University of Basel, Switzerland.

**EU Erasmus exchange postgraduate student.**

10/2001–6/2002 Division of Structural Biology.
Biozentrum, University of Basel, Switzerland.

**Bachelor degree in Biological Sciences.**

1996-2001 Universidad Complutense de Madrid, Spain.
Including 2-year major in "Biosanitaria".

Part-time course containing lectures and practical training on Physiotherapy.
01/1996-06/1996, A.N.D.E, (Spanish Foundation for Mental Disabilities) in Madrid, Spain.

Part-time course on techniques for evaluation and management of ecosystems
01/1998-06/1998, I.M.E.F.E. (Institute for Employment and Business Training) in Madrid, Spain.

Monthly training on Immunology, Cellular Immunology and Immunochemistry
03/1998, 03/1999, 02/2000 "Hospital 12 de Octubre", Madrid, Spain.

**Skills employed**

**Biochemistry:** protein purification, biochemical and biophysical characterization.
**Molecular biology:** cloning and protein expression.
**Structural biology:** crystallization, data collection (in-house and at SLS and ESRF synchrotrons), X-ray crystallography, structure elucidation and analyses.
**Computing skills** Working knowledge on the crystallographic packages (CCP4 Suite, O), molecular display (DINO) and bioinformatics softwares.

**Publications**

**Garcia, P**., Marino, M. and Mayans, O. (2004) Crystallization and preliminary X-ray analysis of the coiled-coil domain of dystrophia myotonica kinase. *Acta Crystallogr D Biol Crystallogr*, 60, 2336-2339

**Garcia, P.,** Ucurum, Z., Bucher, R., Svergun, D.I., Huber, T., Lustig, A., Konarev, P.V., Marino, M., and Mayans, O. 2006. Molecular insights into the self-assembly mechanism of dystrophia myotonica kinase. *Faseb J* 20(8): 1142-1151.

Marino, M., Zou, P., Svergun, D.I., **Garcia, P.,** Edlich, C., Simon, B., Wilmanns, M., Muhle-Goll, C. & Mayans, O. (submitted). The Ig-doublet Z1Z2 – a model system for the hybrid analysis of conformational dynamics in poly-Ig tandems from titin.

**Conference talks**

"17$^{th}$ Regio meeting" (on protein crystallography) Falkau, Black Forest, Germany, 2003.

**Poster presentations**

European Crystallography Meeting 22, Hungary, 2004.

**Attended Workshops and Meetings**

" A day on High-Throughput Techniques in Structural Biology",EMBL, Heidelberg Laboratory, Germany, 2002.

"16$^{th}$ Regio meeting" (on protein crystallography), Vosges, France, 2002.

**Teaching**

Instructing and assessing final year undergraduates in small group practicals.

**Attendant of intensive German courses**

Switzerland, 2001-Present.

**Languages**

Spanish, fluent English and conversational German.

**References**

PhD supervisor Prof O. Mayans Olga.Mayans@unibas.ch

# References

Abdul-Manan, N., Aghazadeh, B., Liu, G.A., Majumdar, A., Ouerfelli, O., Siminovitch, K.A., and Rosen, M.K. 1999. Structure of Cdc42 in complex with the GTPase-binding domain of the 'Wiskott-Aldrich syndrome' protein. *Nature* **399**(6734): 379-383.

Adamson, J.G., Zhou, N.E., and Hodges, R.S. 1993. Structure, function and application of the coiled-coil protein folding motif. *Curr Opin Biotechnol* **4**(4): 428-437.

Albertinazzi, C., Gilardelli, D., Paris, S., Longhi, R., and de Curtis, I. 1998. Overexpression of a neural-specific rho family GTPase, cRac1B, selectively induces enhanced neuritogenesis and neurite branching in primary neurons. *J Cell Biol* **142**(3): 815-825.

Alberts, A.S., Bouquin, N., Johnston, L.H., and Treisman, R. 1998. Analysis of RhoA-binding proteins reveals an interaction domain conserved in heterotrimeric G protein beta subunits and the yeast response regulator protein Skn7. *J Biol Chem* **273**(15): 8616-8622.

Allain, F.H., Bouvet, P., Dieckmann, T., and Feigon, J. 2000. Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin. *Embo J* **19**(24): 6870-6881.

Alwazzan, M., Hamshere, M.G., Lennon, G.G., and Brook, J.D. 1998. Six transcripts map within 200 kilobases of the myotonic dystrophy expanded repeat. *Mamm Genome* **9**(6): 485-487.

Amano, M., Chihara, K., Nakamura, N., Kaneko, T., Matsuura, Y., and Kaibuchi, K. 1999. The COOH terminus of Rho-kinase negatively regulates rho-kinase activity. *J Biol Chem* **274**(45): 32418-32424.

Aspenstrom, P., Fransson, A., and Saras, J. 2004. Rho GTPases have diverse effects on the organization of the actin filament system. *Biochem J* **377**(Pt 2): 327-337.

Balasubramanyam, A., Iyer, D., Stringer, J.L., Beaulieu, C., Potvin, A., Neumeyer, A.M., Avruch, J., and Epstein, H.F. 1998. Developmental changes in expression of myotonic dystrophy protein kinase in the rat central nervous system. *J Comp Neurol* **394**(3): 309-325.

Barz, C., Abahji, T.N., Trulzsch, K., and Heesemann, J. 2000. The Yersinia Ser/Thr protein kinase YpkA/YopO directly interacts with the small GTPases RhoA and Rac-1. *FEBS Lett* **482**(1-2): 139-143.

Batten, F.G., H. 1909. Myotonia atrophica. *Brain* **32**: 187-205.

Begemann, G., Paricio, N., Artero, R., Kiss, I., Perez-Alonso, M., and Mlodzik, M. 1997. muscleblind, a gene required for photoreceptor differentiation in Drosophila, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development* **124**(21): 4321-4331.

Benders, A.A., Groenen, P.J., Oerlemans, F.T., Veerkamp, J.H., and Wieringa, B. 1997. Myotonic dystrophy protein kinase is involved in the modulation of the Ca2+ homeostasis in skeletal muscle cells. *J Clin Invest* **100**(6): 1440-1447.

Birney, E., Kumar, S., and Krainer, A.R. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res* **21**(25): 5803-5816.

Bishop, A.L. and Hall, A. 2000. Rho GTPases and their effector proteins. *Biochem J* **348 Pt 2**: 241-255.

Boucher, C.A., King, S.K., Carey, N., Krahe, R., Winchester, C.L., Rahman, S., Creavin, T., Meghji, P., Bailey, M.E., Chartier, F.L., and et al. 1995. A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)n repeat. *Hum Mol Genet* **4**(10): 1919-1925.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Brewster, B.G.P.W.B. 1998. Myotonic dystrophy: clinical and molecular aspects. *In Emery AEH (ed) Neuromuscular disorders: clinical and molecular genetics John Wiley & Sons, Nova York*

Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T., and et al. 1992. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**(4): 799-808.

Brown, J.H., Cohen, C., and Parry, D.A. 1996. Heptad breaks in alpha-helical coiled coils: stutters and stammers. *Proteins* **26**(2): 134-145.

Buj-Bello, A., Furling, D., Tronchere, H., Laporte, J., Lerouge, T., Butler-Browne, G.S., and Mandel, J.L. 2002. Muscle-specific alternative splicing of myotubularin-related 1 gene is impaired in DM1 muscle cells. *Hum Mol Genet* **11**(19): 2297-2307.

Bullough, P.A., Hughson, F.M., Skehel, J.J., and Wiley, D.C. 1994. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**(6492): 37-43.

Burbelo, P.D., Drechsel, D., and Hall, A. 1995. A conserved binding motif defines numerous candidate target proteins for both Cdc42 and Rac GTPases. *J Biol Chem* **270**(49): 29071-29074.

Burd, C.G. and Dreyfuss, G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**(5172): 615-621.

Burkhard, P., Stetefeld, J., and Strelkov, S.V. 2001. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol* **11**(2): 82-88.

Burridge, K. and Wennerberg, K. 2004. Rho and Rac take center stage. *Cell* **116**(2): 167-179.

Bush, E.W., Helmke, S.M., Birnbaum, R.A., and Perryman, M.B. 2000. Myotonic dystrophy protein kinase domains mediate localization, oligomerization, novel catalytic activity, and autoinhibition. *Biochemistry* **39**(29): 8480-8490.

Chan, T.O., Rittenhouse, S.E., and Tsichlis, P.N. 1999. AKT/PKB and other D3 phosphoinositide-regulated kinases: kinase activation by phosphoinositide-dependent phosphorylation. *Annu Rev Biochem* **68**: 965-1014.

Charlet, B.N., Savkur, R.S., Singh, G., Philips, A.V., Grice, E.A., and Cooper, T.A. 2002. Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol Cell* **10**(1): 45-53.

Chothia, C., Levitt, M., and Richardson, D. 1981. Helix to helix packing in proteins. *J Mol Biol* **145**(1): 215-250.

Cohen, C. and Parry, D.A. 1990. Alpha-helical coiled coils and bundles: how to design an alpha-helical protein. *Proteins* **7**(1): 1-15.

Coleman, M.L., Sahai, E.A., Yeo, M., Bosch, M., Dewar, A., and Olson, M.F. 2001. Membrane blebbing during apoptosis results from caspase-mediated activation of ROCK I. *Nat Cell Biol* **3**(4): 339-345.

Conte, M.R., Grune, T., Ghuman, J., Kelly, G., Ladas, A., Matthews, S., and Curry, S. 2000. Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. *Embo J* **19**(12): 3132-3141.

Conway, J.F. and Parry, D.A. 1991. Three-stranded alpha-fibrous proteins: the heptad repeat and its implications for structure. *Int J Biol Macromol* **13**(1): 14-16.

Cooper, T.A. 1998. Muscle-specific splicing of a heterologous exon mediated by a single muscle-specific splicing enhancer from the cardiac troponin T gene. *Mol Cell Biol* **18**(8): 4519-4525.

Cooper, T.A. and Ordahl, C.P. 1984. A single troponin T gene regulated by different programs in cardiac and skeletal muscle development. *Science* **226**(4677): 979-982.

Crick, F.H.C. 1953. The Fourier transform of a coiled-coil. *Acta Crystallographica* **6**: 685-689.

Crowder, S.M., Kanaar, R., Rio, D.C., and Alber, T. 1999. Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal. *Proc Natl Acad Sci U S A* **96**(9): 4892-4897.

Davis, B.M., McCurrach, M.E., Taneja, K.L., Singer, R.H., and Housman, D.E. 1997. Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc Natl Acad Sci U S A* **94**(14): 7388-7393.

DeGrado, W.F., Wasserman, Z.R., and Lear, J.D. 1989. Protein design, a minimalist approach. *Science* **243**(4891): 622-628.

Deo, R.C., Bonanno, J.B., Sonenberg, N., and Burley, S.K. 1999. Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98**(6): 835-845.

Di Cunto, F., Calautti, E., Hsiao, J., Ong, L., Topley, G., Turco, E., and Dotto, G.P. 1998. Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase encompassing the Rho-Rac-binding protein Citron. *J Biol Chem* **273**(45): 29706-29711.

Diaz-Meco, M.T., Municio, M.M., Frutos, S., Sanchez, P., Lozano, J., Sanz, L., and Moscat, J. 1996. The product of par-4, a gene induced during apoptosis, interacts selectively with the atypical isoforms of protein kinase C. *Cell* **86**(5): 777-786.

Ding, J., Hayashi, M.K., Zhang, Y., Manche, L., Krainer, A.R., and Xu, R.M. 1999. Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* **13**(9): 1102-1115.

Doran, J.D., Liu, X., Taslimi, P., Saadat, A., and Fox, T. 2004. New insights into the structure-function relationships of Rho-associated kinase: a thermodynamic and hydrodynamic study of the dimer-to-monomer transition and its kinetic implications. *Biochem J* **384**(Pt 2): 255-262.

Dunne, P.W., Ma, L., Casey, D.L., Harati, Y., and Epstein, H.F. 1996. Localization of myotonic dystrophy protein kinase in skeletal muscle and its alteration with disease. *Cell Motil Cytoskeleton* **33**(1): 52-63.

Dutta, K., Engler, F.A., Cotton, L., Alexandrov, A., Bedi, G.S., Colquhoun, J., and Pascal, S.M. 2003. Stabilization of a pH-sensitive apoptosis-linked coiled coil through single point mutations. *Protein Sci* **12**(2): 257-265.

Dvorsky, R., Blumenstein, L., Vetter, I.R., and Ahmadian, M.R. 2004. Structural insights into the interaction of ROCKI with the switch regions of RhoA. *J Biol Chem* **279**(8): 7098-7104.

Ellenberger, T.E., Brandl, C.J., Struhl, K., and Harrison, S.C. 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell* **71**(7): 1223-1237.

Eriksson, M., Ansved, T., Anvret, M., and Carey, N. 2001. A mammalian radial spokehead-like gene, RSHL1, at the myotonic dystrophy-1 locus. *Biochem Biophys Res Commun* **281**(4): 835-841.

Etienne-Manneville, S. and Hall, A. 2002. Rho GTPases in cell biology. *Nature* **420**(6916): 629-635.

Fardaei, M., Larkin, K., Brook, J.D., and Hamshere, M.G. 2001. In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res* **29**(13): 2766-2771.

Faustino, N.A. and Cooper, T.A. 2003. Pre-mRNA splicing and human disease. *Genes Dev* **17**(4): 419-437.

Feig, L.A. 1999. Tools of the trade: use of dominant-inhibitory mutants of Ras-family GTPases. *Nat Cell Biol* **1**(2): E25-27.

Feng, J., Ito, M., Kureishi, Y., Ichikawa, K., Amano, M., Isaka, N., Okawa, K., Iwamatsu, A., Kaibuchi, K., Hartshorne, D.J., and Nakano, T. 1999. Rho-associated kinase of chicken gizzard smooth muscle. *J Biol Chem* **274**(6): 3744-3752.

Flynn, P., Mellor, H., Palmer, R., Panayotou, G., and Parker, P.J. 1998. Multiple interactions of PRK1 with RhoA. Functional assignment of the Hr1 repeat motif. *J Biol Chem* **273**(5): 2698-2705.

Frank, S., Kammerer, R.A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bachinger, H.P., and Engel, J. 2001. Stabilization of short collagen-like triple helices by protein engineering. *J Mol Biol* **308**(5): 1081-1089.

Franke, C., Hatt, H., Iaizzo, P.A., and Lehmann-Horn, F. 1990. Characteristics of Na+ channels and Cl- conductance in resealed muscle fibre segments from patients with myotonic dystrophy. *J Physiol* **425**: 391-405.

Fu, Y.H., Pizzuti, A., Fenwick, R.G., Jr., King, J., Rajnarayan, S., Dunne, P.W., Dubel, J., Nasser, G.A., Ashizawa, T., de Jong, P., and et al. 1992. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**(5049): 1256-1258.

Furstoss, O., Dorey, K., Simon, V., Barila, D., Superti-Furga, G., and Roche, S. 2002. c-Abl is an effector of Src for growth factor-induced c-myc expression and DNA synthesis. *Embo J* **21**(4): 514-524.

Gill, S.C. and von Hippel, P.H. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **182**(2): 319-326.

Glover, J.N. and Harrison, S.C. 1995. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature* **373**(6511): 257-261.

Griffin, P.R., MacCoss, M.J., Eng, J.K., Blevins, R.A., Aaronson, J.S., and Yates, J.R., 3rd. 1995. Direct database searching with MALDI-PSD spectra of peptides. *Rapid Commun Mass Spectrom* **9**(15): 1546-1551.

Groenen, P. and Wieringa, B. 1998. Expanding complexity in myotonic dystrophy. *Bioessays* **20**(11): 901-912.

Groenen, P.J., Wansink, D.G., Coerwinkel, M., van den Broek, W., Jansen, G., and Wieringa, B. 2000. Constitutive and regulated modes of splicing produce six major myotonic dystrophy protein kinase (DMPK) isoforms with distinct properties. *Hum Mol Genet* **9**(4): 605-616.

Gu, Y., Byrne, M.C., Paranavitana, N.C., Aronow, B., Siefring, J.E., D'Souza, M., Horton, H.F., Quilliam, L.A., and Williams, D.A. 2002. Rac2, a hematopoiesis-specific Rho GTPase, specifically regulates mast cell protease gene expression in bone marrow-derived mast cells. *Mol Cell Biol* **22**(21): 7645-7657.

Haataja, L., Groffen, J., and Heisterkamp, N. 1997. Characterization of RAC3, a novel member of the Rho family. *J Biol Chem* **272**(33): 20384-20388.

Hakoshima, T., Shimizu, T., and Maesaki, R. 2003. Structural basis of the Rho GTPase signaling. *J Biochem (Tokyo)* **134**(3): 327-331.

Han, J. and Cooper, T.A. 2005. Identification of CELF splicing activation and repression domains in vivo. *Nucleic Acids Res* **33**(9): 2769-2780.

Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. 1999. Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **398**(6728): 579-585.

Hanks, S.K. and Hunter, T. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *Faseb J* **9**(8): 576-596.

Harbury, P.B., Zhang, T., Kim, P.S., and Alber, T. 1993. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**(5138): 1401-1407.

Harper, P.S. 1989. Myotonic Dystrophy. Second edition.: 384 pp.

Hodges, R.S. 1992. Unzipping the secrets of coiled-coils. *Curr Biol* **2**(3): 122-124.

-. 1996. Boehringer Mannheim award lecture 1995. La conference Boehringer Mannheim 1995. De novo design of alpha-helical proteins: basic research to medical applications. *Biochem Cell Biol* **74**(2): 133-154.

Hoffman, G.R. and Cerione, R.A. 2000. Flipping the switch: the structural basis for signaling through the CRIB motif. *Cell* **102**(4): 403-406.

Hofmann-Radvanyi, H., Lavedan, C., Rabes, J.P., Savoy, D., Duros, C., Johnson, K., and Junien, C. 1993. Myotonic dystrophy: absence of CTG enlarged transcript in congenital forms, and low expression of the normal allele. *Hum Mol Genet* **2**(8): 1263-1266.

Hubbard, S.R. and Till, J.H. 2000. Protein tyrosine kinase structure and function. *Annu Rev Biochem* **69**: 373-398.

Hurme, R., Berndt, K.D., Namork, E., and Rhen, M. 1996. DNA binding exerted by a bacterial gene regulator with an extensive coiled-coil domain. *J Biol Chem* **271**(21): 12626-12631.

Hurme, R., Berndt, K.D., Normark, S.J., and Rhen, M. 1997. A proteinaceous gene regulatory thermometer in Salmonella. *Cell* **90**(1): 55-64.

Ishizaki, T., Maekawa, M., Fujisawa, K., Okawa, K., Iwamatsu, A., Fujita, A., Watanabe, N., Saito, Y., Kakizuka, A., Morii, N., and Narumiya, S. 1996. The small GTP-binding protein Rho binds to and activates a 160 kDa Ser/Thr protein kinase homologous to myotonic dystrophy kinase. *Embo J* **15**(8): 1885-1893.

Jacobs, M., Hayakawa, K., Swenson, L., Bellon, S., Fleming, M., Taslimi, P., and Doran, J. 2006. The structure of dimeric ROCK I reveals the mechanism for ligand selectivity. *J Biol Chem* **281**(1): 260-268.

Jancarik, J.a.K., S. H. 1991. Sparse matrix sampling: a screening method for crystallization of proteins. *J Appl Cryst* **24**: 409-411.

Jansen, G., Bachner, D., Coerwinkel, M., Wormskamp, N., Hameister, H., and Wieringa, B. 1995. Structural organization and developmental expression pattern of the mouse WD-repeat gene DMR-N9 immediately upstream of the myotonic dystrophy locus. *Hum Mol Genet* **4**(5): 843-852.

Jansen, G., Groenen, P.J., Bachner, D., Jap, P.H., Coerwinkel, M., Oerlemans, F., van den Broek, W., Gohlsch, B., Pette, D., Plomp, J.J., Molenaar, P.C., Nederhoff, M.G., van Echteld, C.J., Dekker, M., Berns, A., Hameister, H., and Wieringa, B. 1996. Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nat Genet* **13**(3): 316-324.

Jansen, G., Mahadevan, M., Amemiya, C., Wormskamp, N., Segers, B., Hendriks, W., O'Hoy, K., Baird, S., Sabourin, L., Lennon, G., and et al. 1992. Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. *Nat Genet* **1**(4): 261-266.

Janzen, M.L.M., K.A. Benzow, J.W. Day, M.D. Koob and L.P.W. Ranum. 1999. Limited expression of SCA8 is consistent with cerebellar pathogenesis and toxic gain of function RNA model. *Am J Hum Genet* **65**: A267.

Jin, S., Shimizu, M., Balasubramanyam, A., and Epstein, H.F. 2000. Myotonic dystrophy protein kinase (DMPK) induces actin cytoskeletal reorganization and apoptotic-like blebbing in lens cells. *Cell Motil Cytoskeleton* **45**(2): 133-148.

Junghans, R.P., Ebralidze, A., and Tiwari, B. 2001. Does (CUG)n repeat in DMPK mRNA 'paint' chromosome 19 to suppress distant genes to create the diverse phenotype of myotonic dystrophy? A new hypothesis of long-range cis autosomal inactivation. *Neurogenetics* **3**(2): 59-67.

Justice, R.W. 1995. The Drosophila tumor suppressor gene warts encodes a homolog of human myotonic dystrophy kinase and is required for the control of cell shape and proliferation. *Genes and Development* **9**(5): 534-546.

Kaliman, P., Catalucci, D., Lam, J.T., Kondo, R., Gutierrez, J.C., Reddy, S., Palacin, M., Zorzano, A., Chien, K.R., and Ruiz-Lozano, P. 2005. Myotonic dystrophy protein kinase phosphorylates phospholamban and regulates calcium uptake in cardiomyocyte sarcoplasmic reticulum. *J Biol Chem* **280**(9): 8016-8021.

Kammerer, R.A. 1997. Alpha-helical coiled-coil oligomerization domains in extracellular proteins. *Matrix Biol* **15**(8-9): 555-565; discussion 567-558.

Kellerer, M., Lammers, R., Ermel, B., Tippmer, S., Vogt, B., Obermaier-Kusser, B., Ullrich, A., and Haring, H.U. 1992. Distinct alpha-subunit structures of human insulin receptor A and B variants determine differences in tyrosine kinase activities. *Biochemistry* **31**(19): 4588-4596.

Kim, A.S., Kakalis, L.T., Abdul-Manan, N., Liu, G.A., and Rosen, M.K. 2000. Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. *Nature* **404**(6774): 151-158.

Klesert, T.R., Otten, A.D., Bird, T.D., and Tapscott, S.J. 1997. Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat Genet* **16**(4): 402-406.

Knaus, U.G., Heyworth, P.G., Kinsella, B.T., Curnutte, J.T., and Bokoch, G.M. 1992. Purification and characterization of Rac 2. A cytosolic GTP-binding protein that regulates human neutrophil NADPH oxidase. *J Biol Chem* **267**(33): 23575-23582.

Kosako, H., Yoshida, T., Matsumura, F., Ishizaki, T., Narumiya, S., and Inagaki, M. 2000. Rho-kinase/ROCK is involved in cytokinesis through the phosphorylation of myosin light chain and not ezrin/radixin/moesin proteins at the cleavage furrow. *Oncogene* **19**(52): 6059-6064.

Koty, P.P., Pegoraro, E., Hobson, G., Marks, H.G., Turel, A., Flagler, D., Cadaldini, M., Angelini, C., and Hoffman, E.P. 1996. Myotonia and the muscle chloride channel: dominant mutations show variable penetrance and founder effect. *Neurology* **47**(4): 963-968.

Kuhnel, K., Jarchau, T., Wolf, E., Schlichting, I., Walter, U., Wittinghofer, A., and Strelkov, S.V. 2004. The VASP tetramerization domain is a right-handed coiled coil based on a 15-residue repeat. *Proc Natl Acad Sci U S A* **101**(49): 17027-17032.

Ladd, A.N., Charlet, N., and Cooper, T.A. 2001. The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol Cell Biol* **21**(4): 1285-1296.

Ladd, A.N., Nguyen, N.H., Malhotra, K., and Cooper, T.A. 2004. CELF6, a member of the CELF family of RNA-binding proteins, regulates muscle-specific splicing enhancer-dependent alternative splicing. *J Biol Chem* **279**(17): 17756-17764.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259): 680-685.

Lehmann-Horn, F. and Jurkat-Rott, K. 1999. Voltage-gated ion channels and hereditary disease. *Physiol Rev* **79**(4): 1317-1372.

Lei, M., Lu, W., Meng, W., Parrini, M.C., Eck, M.J., Mayer, B.J., and Harrison, S.C. 2000. Structure of PAK1 in an autoinhibited conformation reveals a multistage activation switch. *Cell* **102**(3): 387-397.

Leung, T., Chen, X.Q., Tan, I., Manser, E., and Lim, L. 1998. Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector in promoting cytoskeletal reorganization. *Mol Cell Biol* **18**(1): 130-140.

Leung, T., Manser, E., Tan, L., and Lim, L. 1995. A novel serine/threonine kinase binding the Ras-related RhoA GTPase which translocates the kinase to peripheral membranes. *J Biol Chem* **270**(49): 29051-29054.

Liquori, C.L., Ricker, K., Moseley, M.L., Jacobsen, J.F., Kress, W., Naylor, S.L., Day, J.W., and Ranum, L.P. 2001. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**(5531): 864-867.

Lu, X., Timchenko, N.A., and Timchenko, L.T. 1999. Cardiac elav-type RNA-binding protein (ETR-3) binds to RNA CUG repeats expanded in myotonic dystrophy. *Hum Mol Genet* **8**(1): 53-60.

Luo, L., Lee, T., Tsai, L., Tang, G., Jan, L.Y., and Jan, Y.N. 1997. Genghis Khan (Gek) as a putative effector for Drosophila Cdc42 and regulator of actin polymerization. *Proc Natl Acad Sci U S A* **94**(24): 12963-12968.

Lupas, A.N. and Gruber, M. 2005. The structure of alpha-helical coiled coils. *Adv Protein Chem* **70**: 37-78.

Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuoka, T., Bito, H., Ishizaki, T., and Narumiya, S. 1998. Role of citron kinase as a target of the small GTPase Rho in cytokinesis. *Nature* **394**(6692): 491-494.

Maesaki, R., Ihara, K., Shimizu, T., Kuroda, S., Kaibuchi, K., and Hakoshima, T. 1999. The structural basis of Rho effector recognition revealed by the crystal structure of human RhoA complexed with the effector domain of PKN/PRK1. *Mol Cell* **4**(5): 793-803.

Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., and et al. 1992. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255**(5049): 1253-1255.

Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C.A. 2000. Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* **289**(5485): 1769-1773.

Mankodi, A., Takahashi, M.P., Jiang, H., Beck, C.L., Bowers, W.J., Moxley, R.T., Cannon, S.C., and Thornton, C.A. 2002. Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell* **10**(1): 35-44.

McClain, D.A. 1991. Different ligand affinities of the two human insulin receptor splice variants are reflected in parallel changes in sensitivity for insulin action. *Mol Endocrinol* **5**(5): 734-739.

Michalowski, S., Miller, J.W., Urbinati, C.R., Paliouras, M., Swanson, M.S., and Griffith, J. 1999. Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Res* **27**(17): 3534-3542.

Miller, J.W., Urbinati, C.R., Teng-Umnuay, P., Stenberg, M.G., Byrne, B.J., Thornton, C.A., and Swanson, M.S. 2000. Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *Embo J* **19**(17): 4439-4448.

Millward, T.A., Hess, D., and Hemmings, B.A. 1999. Ndr protein kinase is regulated by phosphorylation on two conserved sequence motifs. *J Biol Chem* **274**(48): 33847-33850.

Moller, D.E., Yokota, A., Caro, J.F., and Flier, J.S. 1989. Tissue-specific expression of two alternatively spliced insulin receptor mRNAs in man. *Mol Endocrinol* **3**(8): 1263-1269.

Morreale, A., Venkatesan, M., Mott, H.R., Owen, D., Nietlispach, D., Lowe, P.N., and Laue, E.D. 2000. Structure of Cdc42 bound to the GTPase binding domain of PAK. *Nat Struct Biol* **7**(5): 384-388.

Mosthaf, L., Vogt, B., Haring, H.U., and Ullrich, A. 1991. Altered expression of insulin receptor types A and B in the skeletal muscle of non-insulin-dependent diabetes mellitus patients. *Proc Natl Acad Sci U S A* **88**(11): 4728-4730.

Mott, H.R., Owen, D., Nietlispach, D., Lowe, P.N., Manser, E., Lim, L., and Laue, E.D. 1999. Structure of the small G protein Cdc42 bound to the GTPase-binding domain of ACK. *Nature* **399**(6734): 384-388.

Mounsey, J.P., John, J.E., 3rd, Helmke, S.M., Bush, E.W., Gilbert, J., Roses, A.D., Perryman, M.B., Jones, L.R., and Moorman, J.R. 2000a. Phospholemman is a substrate for myotonic dystrophy protein kinase. *J Biol Chem* **275**(30): 23362-23367.

Mounsey, J.P., Mistry, D.J., Ai, C.W., Reddy, S., and Moorman, J.R. 2000b. Skeletal muscle sodium channel gating in mice deficient in myotonic dystrophy protein kinase. *Hum Mol Genet* **9**(15): 2313-2320.

Muller, S., von Eichel-Streiber, C., and Moos, M. 1999. Impact of amino acids 22-27 of Rho-subfamily GTPases on glucosylation by the large clostridial cytotoxins TcsL-1522, TcdB-1470 and TcdB-8864. *Eur J Biochem* **266**(3): 1073-1080.

Muranyi, A., Zhang, R., Liu, F., Hirano, K., Ito, M., Epstein, H.F., and Hartshorne, D.J. 2001. Myotonic dystrophy protein kinase phosphorylates the myosin phosphatase targeting subunit and inhibits myosin phosphatase activity. *FEBS Lett* **493**(2-3): 80-84.

Mutsuddi, M., Marshall, C.M., Benzow, K.A., Koob, M.D., and Rebay, I. 2004. The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in Drosophila. *Curr Biol* **14**(4): 302-308.

Ng, Y., Tan, I., Lim, L., and Leung, T. 2004. Expression of the human myotonic dystrophy kinase-related Cdc42-binding kinase gamma is regulated by promoter DNA methylation and Sp1 binding. *J Biol Chem* **279**(33): 34156-34164.

Novelli, G., Gennarelli, M., Zelano, G., Pizzuti, A., Fattorini, C., Caskey, C.T., and Dallapiccola, B. 1993. Failure in detecting mRNA transcripts from the mutated allele in myotonic dystrophy muscle. *Biochem Mol Biol Int* **29**(2): 291-297.

O'Shea, E.K., Klemm, J.D., Kim, P.S., and Alber, T. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**(5031): 539-544.

Otten, A.D. and Tapscott, S.J. 1995. Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc Natl Acad Sci U S A* **92**(12): 5465-5469.

Palaty, C.K., Kalmar, G., Tai, G., Oh, S., Amankawa, L., Affolter, M., Aebersold, R., and Pelech, S.L. 1997. Identification of the autophosphorylation sites of the Xenopus laevis Pim-1 proto-oncogene-encoded protein kinase. *J Biol Chem* **272**(16): 10514-10521.

Palmer, R.H., Ridden, J., and Parker, P.J. 1995. Cloning and expression patterns of two members of a novel protein-kinase-C-related kinase family. *Eur J Biochem* **227**(1-2): 344-351.

Pauling, L., Corey, R.B., and Branson, H.R. 1951. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci U S A* **37**(4): 205-211.

Perez-Canadillas, J.M. and Varani, G. 2001. Recent advances in RNA-protein recognition. *Curr Opin Struct Biol* **11**(1): 53-58.

Peters, J., Baumeister, W., and Lupas, A. 1996. Hyperthermostable surface layer protein tetrabrachion from the archaebacterium Staphylothermus marinus: evidence for the presence of a right-handed coiled coil derived from the primary structure. *J Mol Biol* **257**(5): 1031-1041.

Peters, J., Nitsch, M., Kuhlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., Engelhardt, H., Pfander, J.P., Muller, S., Goldie, K., and et al. 1995. Tetrabrachion: a filamentous archaebacterial surface protein assembly of unusual structure and extreme stability. *J Mol Biol* **245**(4): 385-401.

Philips, A.V., Timchenko, L.T., and Cooper, T.A. 1998. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* **280**(5364): 737-741.

Piekny, A.J. and Mains, P.E. 2002. Rho-binding kinase (LET-502) and myosin phosphatase (MEL-11) regulate cytokinesis in the early Caenorhabditis elegans embryo. *J Cell Sci* **115**(Pt 11): 2271-2282.

Reddy, S., Smith, D.B., Rich, M.M., Leferovich, J.M., Reilly, P., Davis, B.M., Tran, K., Rayburn, H., Bronson, R., Cros, D., Balice-Gordon, R.J., and Housman, D. 1996. Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat Genet* **13**(3): 325-335.

Ricker, K., Koch, M.C., Lehmann-Horn, F., Pongratz, D., Otto, M., Heine, R., and Moxley, R.T., 3rd. 1994. Proximal myotonic myopathy: a new dominant disorder with myotonia, muscle weakness, and cataracts. *Neurology* **44**(8): 1448-1452.

Ridley, A.J. 2001. Rho family proteins: coordinating cell responses. *Trends Cell Biol* **11**(12): 471-477.

Rios, E. and Brum, G. 1987. Involvement of dihydropyridine receptors in excitation-contraction coupling in skeletal muscle. *Nature* **325**(6106): 717-720.

Roberts, R., Timchenko, N.A., Miller, J.W., Reddy, S., Caskey, C.T., Swanson, M.S., and Timchenko, L.T. 1997. Altered phosphorylation and intracellular distribution of a (CUG)n triplet repeat RNA-binding protein in patients with myotonic dystrophy and in myotonin protein kinase knockout mice. *Proc Natl Acad Sci U S A* **94**(24): 13221-13226.

Rudolph, M.G., Bayer, P., Abo, A., Kuhlmann, J., Vetter, I.R., and Wittinghofer, A. 1998. The Cdc42/Rac interactive binding region motif of the Wiskott Aldrich syndrome protein (WASP) is necessary but not sufficient for tight binding to Cdc42 and structure formation. *J Biol Chem* **273**(29): 18067-18076.

Sabourin, L.A., Tamai, K., Narang, M.A., and Korneluk, R.G. 1997. Overexpression of 3'-untranslated region of the myotonic dystrophy kinase cDNA inhibits myoblast differentiation in vitro. *J Biol Chem* **272**(47): 29626-29635.

Salvatori, S., Biral, D., Furlan, S., and Marin, O. 1997. Evidence for localization of the myotonic dystrophy protein kinase to the terminal cisternae of the sarcoplasmic reticulum. *J Muscle Res Cell Motil* **18**(4): 429-440.

Savkur, R.S., Philips, A.V., and Cooper, T.A. 2001. Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet* **29**(1): 40-47.

Scherly, D., Boelens, W., Dathan, N.A., van Venrooij, W.J., and Mattaj, I.W. 1990. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B'' and their cognate RNAs. *Nature* **345**(6275): 502-506.

Schulz, P.E., McIntosh, A.D., Kasten, M.R., Wieringa, B., and Epstein, H.F. 2003. A role for myotonic dystrophy protein kinase in synaptic plasticity. *J Neurophysiol* **89**(3): 1177-1186.

Sebbagh, M., Renvoize, C., Hamelin, J., Riche, N., Bertoglio, J., and Breard, J. 2001. Caspase-3-mediated cleavage of ROCK I induces MLC phosphorylation and apoptotic membrane blebbing. *Nat Cell Biol* **3**(4): 346-352.

Seino, S. and Bell, G.I. 1989. Alternative splicing of human insulin receptor messenger RNA. *Biochem Biophys Res Commun* **159**(1): 312-316.

Sells, S.F., Han, S.S., Muthukkumar, S., Maddiwar, N., Johnstone, R., Boghaert, E., Gillis, D., Liu, G., Nair, P., Monnig, S., Collini, P., Mattson, M.P., Sukhatme, V.P., Zimmer, S.G., Wood, D.P., Jr., McRoberts, J.W., Shi, Y., and Rangnekar, V.M. 1997. Expression and function of the leucine zipper protein Par-4 in apoptosis. *Mol Cell Biol* **17**(7): 3823-3832.

Sergeant, N., Sablonniere, B., Schraen-Maschke, S., Ghestem, A., Maurage, C.A., Wattez, A., Vermersch, P., and Delacourte, A. 2001. Dysregulation of human brain microtubule-associated tau mRNA maturation in myotonic dystrophy type 1. *Hum Mol Genet* **10**(19): 2143-2155.

Shahani, N. and Brandt, R. 2002. Functions and malfunctions of the tau proteins. *Cell Mol Life Sci* **59**(10): 1668-1680.

Shamoo, Y., Krueger, U., Rice, L.M., Williams, K.R., and Steitz, T.A. 1997. Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 A resolution. *Nat Struct Biol* **4**(3): 215-222.

Shaw, D.J., McCurrach, M., Rundle, S.A., Harley, H.G., Crow, S.R., Sohn, R., Thirion, J.P., Hamshere, M.G., Buckler, A.J., Harper, P.S., and et al. 1993. Genomic organization and transcriptional units at the myotonic dystrophy locus. *Genomics* **18**(3): 673-679.

Shibata, H., Mukai, H., Inagaki, Y., Homma, Y., Kimura, K., Kaibuchi, K., Narumiya, S., and Ono, Y. 1996. Characterization of the interaction between RhoA and the amino-terminal region of PKN. *FEBS Lett* **385**(3): 221-224.

Shimizu, M., Wang, W., Walch, E.T., Dunne, P.W., and Epstein, H.F. 2000. Rac-1 and Raf-1 kinases, components of distinct signaling pathways, activate myotonic dystrophy protein kinase. *FEBS Lett* **475**(3): 273-277.

Shimizu, T., Ihara, K., Maesaki, R., Amano, M., Kaibuchi, K., and Hakoshima, T. 2003. Parallel coiled-coil association of the RhoA-binding domain in Rho-kinase. *J Biol Chem* **278**(46): 46046-46051.

Shimokawa, M., Ishiura, S., Kameda, N., Yamamoto, M., Sasagawa, N., Saitoh, N., Sorimachi, H., Ueda, H., Ohno, S., Suzuki, K., and Kobayashi, T. 1997. Novel isoform of myotonin protein kinase: gene product of myotonic dystrophy is localized in the sarcoplasmic reticulum of skeletal muscle. *Am J Pathol* **150**(4): 1285-1295.

Singh, G., Charlet, B.N., Han, J., and Cooper, T.A. 2004. ETR-3 and CELF4 protein domains required for RNA binding and splicing activity in vivo. *Nucleic Acids Res* **32**(3): 1232-1241.

Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W.J. 2003. RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res* **31**(19): 5469-5482.

Steinert, H. 1909. Myopathologische Beitrage 1: Uber das klinischeund anatomische Bild des Muskelchwunds der Myotoniker. *Dtsch Z Nervenheilkd* **37**: 58-104.

Strelkov, S.V. and Burkhard, P. 2002. Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. *J Struct Biol* **137**(1-2): 54-64.

Suzuki, A., Kurasawa, T., Tashiro, C., Hasegawa, K., Yamane, T., and Ashida, T. 1993. Crystallization and preliminary X-ray studies on the trypsin inhibitor I-2 from wheat germ and its complex with trypsin. *Acta Crystallogr D Biol Crystallogr* **49**(Pt 6): 594-596.

Suzuki, A., Sugiyama, Y., Hayashi, Y., Nyu-i, N., Yoshida, M., Nonaka, I., Ishiura, S., Arahata, K., and Ohno, S. 1998. MKBP, a novel member of the small heat

shock protein family, binds and activates the myotonic dystrophy protein kinase. *J Cell Biol* **140**(5): 1113-1124.

Takagaki, Y. and Manley, J.L. 2000. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol Cell Biol* **20**(5): 1515-1525.

Takai, Y., Sasaki, T., and Matozaki, T. 2001. Small GTP-binding proteins. *Physiol Rev* **81**(1): 153-208.

Tan, I., Cheong, A., Lim, L., and Leung, T. 2003. Genomic organization of human myotonic dystrophy kinase-related Cdc42-binding kinase alpha reveals multiple alternative splicing and functional diversity. *Gene* **304**: 107-115.

Tan, I., Seow, K.T., Lim, L., and Leung, T. 2001. Intermolecular and intramolecular interactions regulate catalytic activity of myotonic dystrophy kinase-related Cdc42-binding kinase alpha. *Mol Cell Biol* **21**(8): 2767-2778.

Taneja, K.L., McCurrach, M., Schalling, M., Housman, D., and Singer, R.H. 1995. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* **128**(6): 995-1002.

Tao, Y., Strelkov, S.V., Mesyanzhinov, V.V., and Rossmann, M.G. 1997. Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. *Structure* **5**(6): 789-798.

Thompson, G., Owen, D., Chalk, P.A., and Lowe, P.N. 1998. Delineation of the Cdc42/Rac-binding domain of p21-activated kinase. *Biochemistry* **37**(21): 7885-7891.

Timchenko, L., Nastainczyk, W., Schneider, T., Patel, B., Hofmann, F., and Caskey, C.T. 1995. Full-length myotonin protein kinase (72 kDa) displays serine kinase activity. *Proc Natl Acad Sci U S A* **92**(12): 5366-5370.

Timchenko, L.T., Miller, J.W., Timchenko, N.A., DeVore, D.R., Datar, K.V., Lin, L., Roberts, R., Caskey, C.T., and Swanson, M.S. 1996. Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res* **24**(22): 4407-4414.

Timchenko, N.A., Cai, Z.J., Welm, A.L., Reddy, S., Ashizawa, T., and Timchenko, L.T. 2001. RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. *J Biol Chem* **276**(11): 7820-7826.

Tiscornia, G. and Mahadevan, M.S. 2000. Myotonic dystrophy: the role of the CUG triplet repeats in splicing of a novel DMPK exon and altered cytoplasmic DMPK mRNA isoform ratios. *Mol Cell* **5**(6): 959-967.

Ueda, H., Kameda, N., Baba, T., Terada, N., Shimokawa, M., Yamamoto, M., Ishiura, S., Kobayashi, T., and Ohno, S. 1998. Immunolocalization of myotonic dystrophy protein kinase in corbular and junctional sarcoplasmic reticulum of human cardiac muscle. *Histochem J* **30**(4): 245-251.

van der Ven, P.F., Jansen, G., van Kuppevelt, T.H., Perryman, M.B., Lupa, M., Dunne, P.W., ter Laak, H.J., Jap, P.H., Veerkamp, J.H., Epstein, H.F., and et al. 1993. Myotonic dystrophy kinase is a component of neuromuscular junctions. *Hum Mol Genet* **2**(11): 1889-1894.

van Herpen, R.E., Oude Ophuis, R.J., Wijers, M., Bennink, M.B., van de Loo, F.A., Fransen, J., Wieringa, B., and Wansink, D.G. 2005. Divergent mitochondrial and endoplasmic reticulum association of DMPK splice isoforms depends on unique sequence arrangements in tail anchors. *Mol Cell Biol* **25**(4): 1402-1414.

van Herpen, R.E., Tjeertes, J.V., Mulders, S.A., Oude Ophuis, R.J., Wieringa, B., and Wansink, D.G. 2006. Coiled-coil interactions modulate multimerization, mitochondrial binding and kinase activity of myotonic dystrophy protein kinase splice isoforms. *Febs J* **273**(6): 1124-1136.

Verde, F., Wiley, D.J., and Nurse, P. 1998. Fission yeast orb6, a ser/thr protein kinase related to mammalian rho kinase and myotonic dystrophy kinase, is required for maintenance of cell polarity and coordinates cell morphogenesis with the cell cycle. *Proc Natl Acad Sci U S A* **95**(13): 7526-7531.

Vetter, I.R. and Wittinghofer, A. 2001. The guanine nucleotide-binding switch in three dimensions. *Science* **294**(5545): 1299-1304.

Wang, X. and Tanaka Hall, T.M. 2001. Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat Struct Biol* **8**(2): 141-145.

Wang, Y.H., Amirhaeri, S., Kang, S., Wells, R.D., and Griffith, J.D. 1994. Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* **265**(5172): 669-671.

Wansink, D.G., van Herpen, R.E., Coerwinkel-Driessen, M.M., Groenen, P.J., Hemmings, B.A., and Wieringa, B. 2003. Alternative splicing controls myotonic dystrophy protein kinase structure, enzymatic activity, and subcellular localization. *Mol Cell Biol* **23**(16): 5489-5501.

Watanabe, G., Saito, Y., Madaule, P., Ishizaki, T., Fujisawa, K., Morii, N., Mukai, H., Ono, Y., Kakizuka, A., and Narumiya, S. 1996. Protein kinase N (PKN) and PKN-related protein rhophilin as targets of small GTPase Rho. *Science* **271**(5249): 645-648.

Weis, W.I. and Drickamer, K. 1994. Trimeric structure of a C-type mannose-binding protein. *Structure* **2**(12): 1227-1240.

Wennerberg, K. and Der, C.J. 2004. Rho-family GTPases: it's not only Rac and Rho (and I like it). *J Cell Sci* **117**(Pt 8): 1301-1312.

Whiting, E.J., Waring, J.D., Tamai, K., Somerville, M.J., Hincke, M., Staines, W.A., Ikeda, J.E., and Korneluk, R.G. 1995. Characterization of myotonic dystrophy kinase (DMK) protein in human and rodent muscle and central nervous tissue. *Hum Mol Genet* **4**(6): 1063-1072.

Winter-Vann, A.M. and Casey, P.J. 2005. Post-prenylation-processing enzymes as new targets in oncogenesis. *Nat Rev Cancer* **5**(5): 405-412.

Wissmann, A., Ingles, J., McGhee, J.D., and Mains, P.E. 1997. Caenorhabditis elegans LET-502 is related to Rho-binding kinases and human myotonic dystrophy kinase and interacts genetically with a homolog of the regulatory subunit of smooth muscle myosin phosphatase to affect cell shape. *Genes Dev* **11**(4): 409-422.

Woolfson, D.N. and Alber, T. 1995. Predicting oligomerization states of coiled coils. *Protein Sci* **4**(8): 1596-1607.

Xu, R.M., Jokhan, L., Cheng, X., Mayeda, A., and Krainer, A.R. 1997. Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure* **5**(4): 559-570.

Yamaguchi, Y., Flier, J.S., Benecke, H., Ransil, B.J., and Moller, D.E. 1993. Ligand-binding properties of the two isoforms of the human insulin receptor. *Endocrinology* **132**(3): 1132-1138.

Yarden, O., Plamann, M., Ebbole, D.J., and Yanofsky, C. 1992. cot-1, a gene required for hyphal elongation in Neurospora crassa, encodes a protein kinase. *Embo J* **11**(6): 2159-2166.

Zeng, X., Zhu, H., Lashuel, H.A., and Hu, J.C. 1997. Oligomerization properties of GCN4 leucine zipper e and g position mutants. *Protein Sci* **6**(10): 2218-2226.

Zhang, R. and Epstein, H.F. 2003. Homodimerization through coiled-coil regions enhances activity of the myotonic dystrophy protein kinase. *FEBS Lett* **546**(2-3): 281-287.

Zhao, Y., Loyer, P., Li, H., Valentine, V., Kidd, V., and Kraft, A.S. 1997. Cloning and chromosomal location of a novel member of the myotonic dystrophy family of protein kinases. *J Biol Chem* **272**(15): 10013-10020.

Zhao, Z.S. and Manser, E. 2005. PAK and other Rho-associated kinases--effectors with surprisingly diverse mechanisms of regulation. *Biochem J* **386**(Pt 2): 201-214.