Coil 2 of intermediate filaments: its complete structure and impact of desminopathy-related mutations.

Inauguraldissertation

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

von

Stefan Nicolet
aus Tramelan BE

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

<div style="text-align: right">

Prof. Dr. Sergei V. Strelkov

Prof. Dr. Harald Herrmann

Prof. Dr. Ueli Aebi

</div>

Basel, den 27. April 2010

<div style="text-align: right">

Prof. Dr. Eberhard Parlow

Dekan

</div>

Abbreviation table

Å	ångstroms (10^{-10} m)
AUC	analytical ultracentrifugation
CD	circular dichroism
Da	Dalton, atomic unit 1.6 × 10^{-24} g
DNA	desoxyribonucleic acid
DTT	dithiothreitol
EM	electron microscopy
GF	gel filtration
IF(s)	intermediate filament(s)
MAD	multiwavelength anomalous dispersion
PEG	polyethylene glycol
RNA	ribonucleic acid
SAXS	small angle X-ray scattering
TEV	tobacco Etch virus
ULF(s)	unit-length filament(s)
wt	wild type

Acknowledgments

Table of contents

1. Introduction

1.1 The cytoskeleton

Unlike insects and plant cells, which are protected respectively by a hard chitin exoskeleton or by a cellulose cell wall, vertebrate cells are only surrounded by a fine lipid bilayer. To maintain their shapes and organelles organization, vertebrate cells lean on the cell cytoskeleton which consists of three different filament-forming polymer networks. The thinnest filaments of about 6nm diameter called microfilaments are made up of two twisted helical chains of globular G-actin (Holmes *et al.*, 1990)[1]. The largest filaments called microtubules have a diameter of 25nm and form a hollow tube of 13 to 15 protofilaments consisting of globular α and β tubulin heterodimers. The filaments we will focus on have a diameter of 7 to 16 nm and are consequently called intermediate filaments (IFs). IFs have unique properties compared to microtubules and actin microfilaments. First, the elementary IFs building blocks are elongated α-helical proteins. Such long monomeric α-helices are unstable in aqueous solutions and associate into multimers to preserve, as globular proteins, a hydrophobic core and expose to their surface hydrophilic amino-acids (Crick, 1953)[2]. IFs dimers can already be observed by electron microscopy (Aebi *et al.*, 1988). This fibrous rope-like nature confers them more *flexibility* and *resistance* to stress than microfilaments and microtubules (Janmey *et al.*, 1991)[3]. Second, in contrast to microtubules and microfilaments, assembled IFs are *apolar* since elementary dimers adopt both "up and down" directions relative to the filament axis. Finally, IFs are highly *polymorphic* as their number of subunits varies, even within the same filament protein type. Altogether, these IFs properties mirror their biological functions. Their high mechanical resistance prevents cell disruption upon stress whereas their flexibility procures to cells and tissues a sufficient malleability to suite with dynamic processes like, for instance, leukocyte extravasation or muscular contraction. As apolar structures, IFs have never been shown to drive directional motor processes as myosin with actin or kinesin and dynein with microtubules. Lastly, IFs polymorphism may explain the diverse cell or tissue specific requirements for cytoplasmic organization or resistance to stress.

1.2 Intermediate filaments

Intermediate filaments are tissue-specific proteins encoded by a multi-gene family of 70 genes (Hesse *et al.*, 2001)[4]. At least 72 pathologies are associated to IFs mutations. Such a large amount of information is accessible online through the human IF database (Szeverenyi *et al.*, 2008)[5].

A detailed sequence comparison allows their classification into six sequence homology classes (SHC): acidic keratins (SHC I), basic keratins (SHC II), desmin-/vimentin-type proteins (SHC III), neurofilaments (SHC IV), nuclear lamins (SHC V) and eye-lens filaments (SHC VI) (Fuchs *et al.*, 1994, Hess *et al.*, 1998)[6,7]. This thesis only focuses on SHC III IFs.

The primary sequence analysis reveals several features of IFs. First, one can observe an underlying heptad repeat sequence of the form $(a$-b-c-d-e-f-$g)_n$ where the *a* and *d* positions are predominantly occupied by apolar residues like leucine, valine and isoleucine. Such repeat is characteristic for twined α-helices into two-stranded left-handed coiled-coils. This ~310 amino acids (~350 for nuclear lamins and invertebrates) long α-helical rod domain is flanked by variable in length and non helical N-terminal "head" and C-terminal "tail" domains. In addition, the heptad repeat discontinues at three positions called linkers L1, L12 and L2. The linker L1 divides the coil 1 into the 1A and 1B subdomains and the linker L2 splits the coil 2 into 2A and 2B. The linker L12 is found in-between coil 1 and coil 2 (Figure 1.2.1).
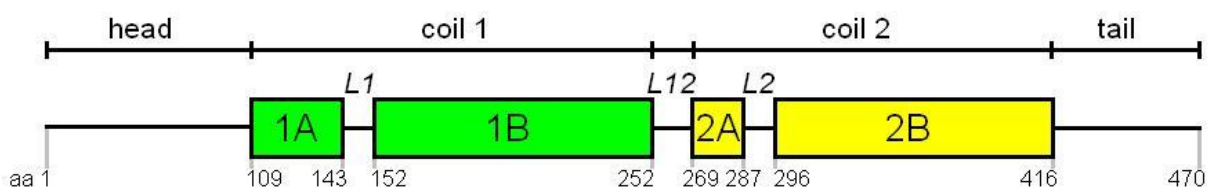
Figure 1.2.1 The domain organization of vimentin (SHC III): the rod domain, which consists of two coils separated by the linker L12, is flanked by the non-helical N-terminal head and the C-terminal tail domains. The coil 1 is divided into the two sub-domains 1A and 1B by the linker L1. Similarly, the coil 2 has two sub-domains 2A and 2B interrupted by the linker L2.

Remarkably, the number of residues in heptad-containing regions is absolutely conserved in cytoplasmic IF of vertebrates. Links, except L2, are more variable (Table 1.2.1).

Type	Head	1A	L1	1B	L12	2A	L2	2B	Tail
Ia	variable	35	11	101	16	19	8	116	variable
Ib	"	35	11-14	101	16	19	8	116	"
IIa	"	35	10	101	17	19	8	116	"
IIb	"	35	11-14	101	17	19	8	116	"
III	"	35	8	101	16,18	19	8	116	55
IV-L	91	35	10	101	17	19	8	116	143
IV-M	102	35	10	101	17	19	8	116	variable
IV-H	95	35	10	101	22	19	8	116	"
IV-INX	94*	35	10	101	15	19	8	116	97*
V	"	41	0	148	19	19	8	116	variable

Table 1.2.1 The number of residues in segments of the different SHCs is absolutely conserved in segments 1A, 1B, 2A and 2B and L2 with the exception of nuclear lamins (SHC V). The SHC Ia and IIa are the hard keratins, Ib and IIb the epidermal keratins and the neurofilaments (SHC IV) are divided in low (NF-L), middle (NF-M) and high (NF-H) according to their polypeptide chain mass. *Only one sequence is available for IV-INX (α-internexin). The number of residues for coil 2B has been adjusted to Table adapted from Parry and Steinert, 1995[8].

Finally, two sequence motifs are found at opposite end of the rod domain. The first one corresponds to the first three heptads of the coil 1A and has eight absolutely conserved residues. The second motif is located at the end of coil 2B and harbor the IF 'consensus motif' YRKLLEGEE. These motifs have been shown to be critical for filament assembly and point mutations within produce severe phenotypes (Strelkov *et al., 2002*)[9].

1.3 Atomic structures of intermediate filaments

The history of IFs structure is linked to famous biology discoveries of the 20 century. The correct interpretation of hard keratin X-ray diffraction pattern by William Astbury in 1930 required Linus Pauling's description of α-helices in 1951[10] and their organization into coiled coils by Francis Crick in 1952[2,11]. It is only in 1981 that the first coiled coil single crystal structures, i.e. the influenza virus heamagglutinin glycoprotein (Wilson *et al.*, 1981)[12] and the catabolite gene activator protein (McKay *et al.*, 1981, 1982)[13,14] were published. The interest in coiled coil grew up in the 90s when several transcription factors were shown to harbor the periodic signature of coiled coils (Landschulz *et al.*, 1988; O'Shea *et al.*, 1991)[15,16]. In 1998, the analysis of the long 18 heptads structure of the oligomerization domain of cortexilin I (Burkhard *et al.*, 1998) [17] reveals that the proper assembly of α-helices in coiled coil requires a 13-residues pattern or 'trigger sites' (Kammerer *et al.*, 1998)[18]. Analogous 'trigger

sites' have been identified in IFs (Wu *et al.*, 2000)[19] and form a remarkable network of intra- and interhelical salt bridges.

First knowledge about IFs atomic structure was essentially inherited from other coiled coil structures, as attempts to crystallize entire or large segments of IFs failed (Herrmann *et al.*, 2004)[20]. The reason behind is that complete IFs or coiled coil fragments of >100 residues become flexible and have a limited rotational diffusion in solution. Thus, such long fragments cannot align properly to form a crystal lattice. To overcome this problem, a 'divide-and-conquer' strategy consisting of crystallizing shorter rod fragments <100 residues was established. This approach is however only valid for IF proteins since the coiled coil folding does not require a precise surrounding as for globular proteins (Strelkov *et al.*, 2001)[21].

The first crystallized IF fragment was a fusion between the yeast GCN4 coiled coil leucine-zipper domain and the vimentin carboxy-terminal consensus motif, residues 385-412. This chimera construct was designed to ensure a proper coil formation. It confirms the presence of the hydrophobic seam in *a* and *d* positions and the presence of an interhelical salt bridge between the Glu396(*g*) and Arg401(*e*) of the other chain stabilizing the end of coil 2. Intrahelical salt bridges are found between Lys390(*a*) and Asp394(*d*). Interestingly, the helices bent away from the coiled-coil axis after residue Leu405, what defines the exact end of the coil 2 (Herrmann *et al.*, 2000)[22].

This chimera construct was confirmed by the crystallization of an overlapping vimentin fragment ('Cys2') from residues 328 to 411. This structure exhibits a discontinuity of the coiled coil geometry, called 'stutter', where α-helices shortly run in parallel. A stutter (residues 351-354) can be considered as a heptad repeat plus four residues or as a heptad minus three residues (Strelkov *et al.*, 2002)[9].

The third known fragment represents the vimentin A coil 1A from residues 102-138 (Strelkov *et al.*, 2002)[9]. This fragment behaves as a monomer both in solution and within the crystal. Surprisingly, its α-helix adopts the same bending as in a coiled coil. One can thus conclude that the α-helical bending only depends of the primary amino acid sequence. The coiled coil dimer model was created by superimposing the coil 1A to the GCN4-vimentin structure. The modeled dimer reveals a strong extra-helical salt bridge between Lys120(*g*) and Glu125(*e*). Recently, the same fragment with a

single stabilizing mutation (Y117L) crystallized as a coiled coil dimer(Meier *et al.*, 2009)[23].

1.4 Coiled coil analysis
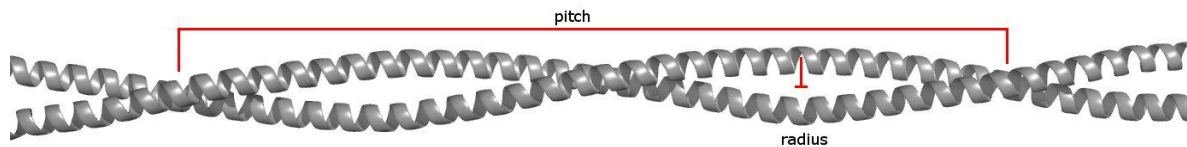
Figure 1.4.1 The two major parameters two describe a coiled coil are the coiled coil pitch and the coiled coil radius. The pitch corresponds to the distance needed for a coiled coil to do a full rotation. The coiled coil radius is the distance between the helix axis and the coiled coil axis.

There are few important parameters to describe a coiled coil. First, the distance required for a coiled coil to do a complete turn is called the *pitch*. The second required parameter is the coiled coil radius, which corresponds to the distance between one helix axis and the coiled coil axis (Strelkov *et al.*, 2002)[24]. Alternatively, Crick used a vector representation to describe coiled coils (Crick, 1953)[25]. In addition each α-helices present in coiled coils are described by their *pitch*, *radius* and *α-helical rise*. The α-helical rise represents the distance along the helix axis per residue. Within this thesis, the program Twister has been used to calculate coiled-coil parameters (Strelkov *et al.*, 2002)[24]

The classical coiled coil, as present in the IF rod domain, has a heptad pattern repeat of residues with ~3.5 residues per turn and is *left-handed* to compensate the small difference of the parallel α-helix 3.6 residues per turn. Alternatively, some coiled coils have *hendecads* (11-residues) periodicities and can be treated as a heptad (7 residues) plus stutter (4 residues). Such α-helices have ~3.7 residues per turn and a resulting *right-handed* tetramer coiled-coil has been observed in a fragment of the tetrabrachion protein from *Staphylothermus marinus* (Stetefeld *et al.*, 2000)[26]. Finally, a *pentadecad* periodicity was found in the NMR structure of the tetramerization domain of the Mnt repressor of Salmonella. These coiled coils are right-handed with an α-helical number of residues per turn of 3.8 (Nooren *et al.*, 1999)[27].

5

1.5 Intermediate filaments assembly

There are three assembly groups; the keratins (SHC I and II) forms obligate heterodimers while the desmin-/vimentin-type (SHC III) and neurofilaments (SHC IV) form homopolymers *in vitro.* However, *in vivo*, filaments of SHC III and IV are frequently observed as copolymers. Finally, lamins (SHV V) are not cytoplasmic proteins but constitute the nuclear lamina attached to the inner nuclear membrane.

Denatured vimentin filaments already forms dimers at an urea concentration of 6M. At 4.5M urea, vimentin further assemble into tetramers. Upon dialysis into buffers close to physiological conditions, vimentin forms long and uniform filaments. In low ionic strength buffer (5mM Tris-HCl, pH 8.4 or 2mM sodium phosphate, pH 7.5) only tetramers species are observed. Increasing the ionic strength of the solution by adding salts readily initiates filament formation. Thus, the assembly process can be monitored at different time points by arresting filament polymerization with glutaraldehyde (Figure 1.5.1).

Within 2s after salt addition, rod-like structures of ~60nm length and ~16nm diameter are observed (Figure 1.5.1A). Such structures, called *unit-length-filaments* (ULFs), result from antiparallel half-staggered tetramers. ULFs then longitudinally anneal into long filaments (Figure 1.5.1B). Concomitantly with a further extension, filaments undergo a structural rearrangement of their subunits to produce mature filaments of ~11nm diameter (Figure 1.5.1C) (Strelkov *et al.*, 2003)[28].
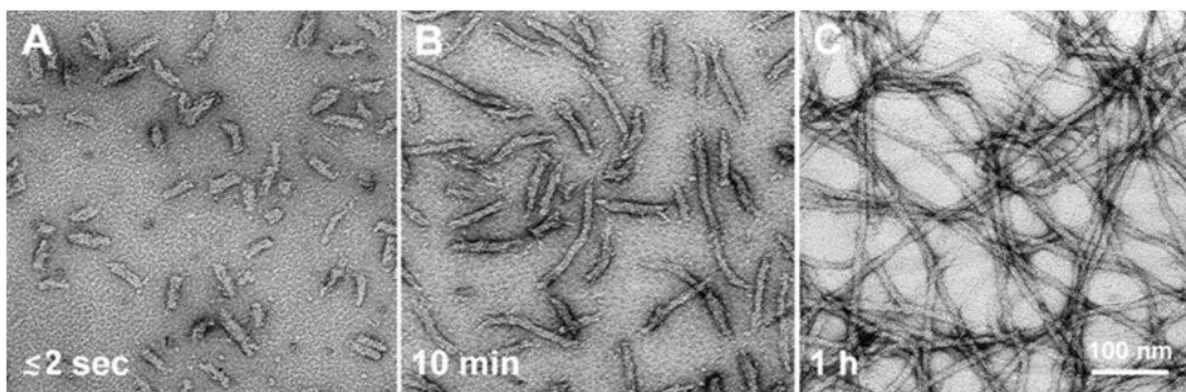
Figure 1.5.1 Electron microscopy time-course pictures of recombinant vimentin assembly. (A) Within 2s after salt addition, unit-length filament structures (ULFs) of about 60nm of length for 16nm diameter are observed. (B) ULFs then longitudinally assemble into immature filaments. (C) Concomitantly with filament elongation, filaments undergo an internal rearrangement or "radial compaction process" to yield mature filaments of about 11nm diameter.

Within this thesis I focus only on SHC III IFs. It should however be mentioned here that the assembly processes considerably vary between IF types.

1.6 Desmin and desmin-related myopathies

Desmin (old name skeletin, MW ~52kDa) is a muscle-specific protein which forms a cytoplasmic scaffold to interconnect myofibers and links the contractile apparatus to the sarcolemma (Morita *et al*, 2005)[29], the mitochondria (Tokuyasu *et al.*, 1983)[30,31] and the nucleus (Stromer *et al.*, 1988)[32]. Desmin is also found in cell-cell adhesion complexes as desmosomes in cardiac muscle (Clark *et al.*, 2002)[33] and costamers in skeletal muscles (Lazarides, 1980)[34] (Figure 1.5.1).

Desmin knock-out mice are still viable and reproduce normally. However, these mice exhibit a loss of lateral alignment of myofibrils, disrupted anchorage of myofibrils to the sarcolemma and disturbed mitochondrial function. Muscle pathologies as degeneration, necrosis and calcification were more pronounced with increased muscle usage (Clark *et al.*, 2002)[33].
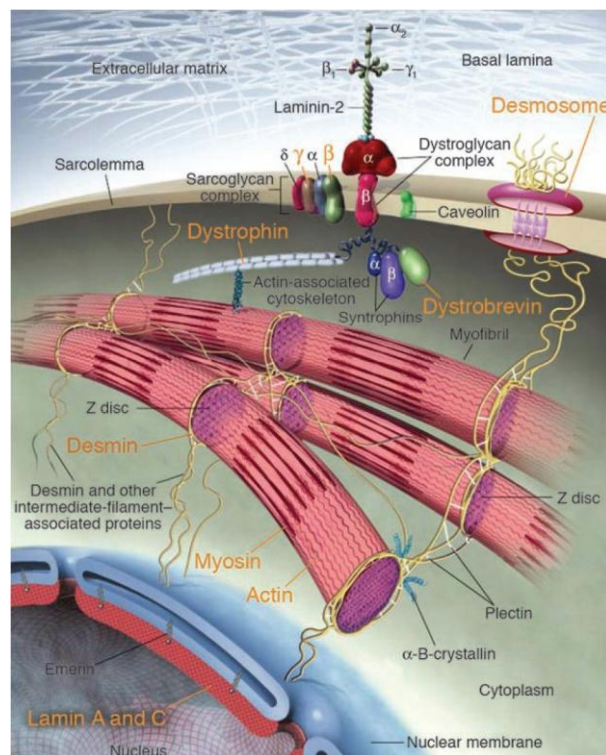
Figure 1.5.1 Schematic view of desmin localization within cardiac muscle cell. Desmin laterally interlink the myofibril Z-discs and connects them to the sarcolemma, nuclear membrane and desmosomes (Morita *et al*, 2005)[29].

Desmin-related myopathies or desminopathies (DRMs; OMIM 601419) in humans are characterized by a progressive skeletal myopathy, cardiomyopathy and respiratory insufficiencies, leading to death by cardiac arrest. Histologically, DRMs are classified in two distinct groups; in the first group desmin deposits are punctual cytoplasmic inclusions or "spherical bodies" whereas the second group consists of disseminated accumulation of granulofilamentous material (Goebel *et al.*, 2000)[35]. DRMs have been associated to missense or deletion mutations of the desmin gene (Goldfarb *et al.*, 1998)[36]. Desminopathies were often misdiagnosed; one can reasonably expect that their number will still increase over time. Interestingly, most of desmin mutations resides in the coil 2B (Figure 1.5.2) and are to proline, a residue considered as a "helix-breaker" (Chou and Fasman, 1974)[37].
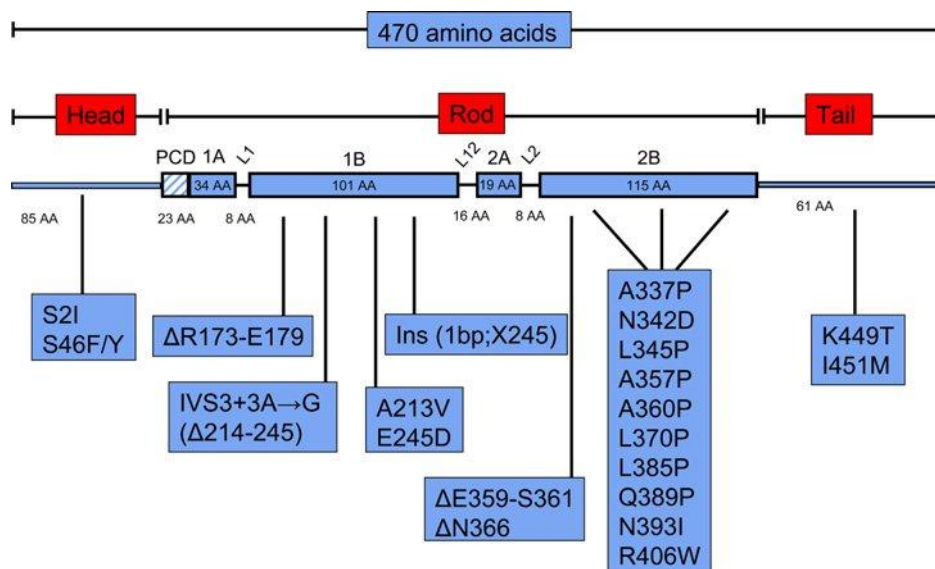
Figure 1.5.2 Schematic view of desmin structure and localization of known desmin mutations (Bar *et al.,* 2004)[38].

Both ΔN366 and ΔE359-S361 desmin deletions were modeled according to the coil 2B vimentin structure (Kaminska *et al.*, 2004)[39]. The ΔN359-S369 deletion generates an additional stutter after the natural occurring one (residues F356-E359) and further unwinds the coiled-coil. In contrast, the ΔN366 deletion close to the stutter converts it to a stammer structure which corresponds to a four residue deletion in the heptad repeat motif. Stammers are predicted to overwind coiled coil structures (Brown *et al.* 1996; Strelkov and Burkhard, 2002)[24,40]. These coiled coil geometry distortions modify the angular positions of head and tail domains altering the filament assembly (Figure 1.5.3).
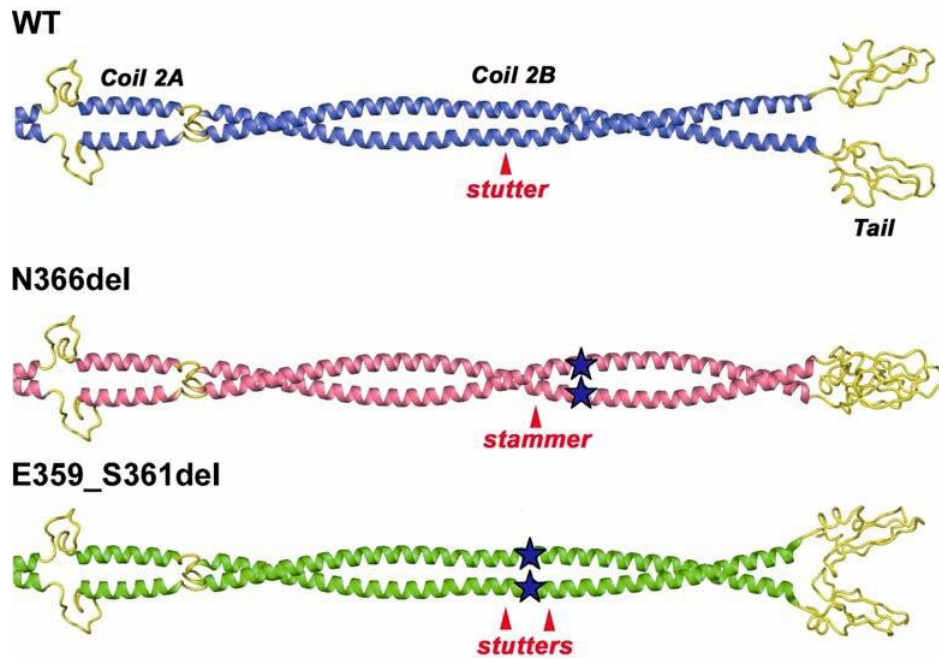
Figure 1.5.3 Coiled coil model of desmin partly based on the vimentin coil 2B structure. The naturally occurring stutter in wt desmin locally unwinds the coiled coil. The ΔN366 deletion produces a stammer which corresponds to a coil overwinding whereas the ΔE359-S361 deletion generates a second stutter which further unwind the coil 2B. The tail angles are drastically modified compared to the wt (picture from Kaminska *et al.***, 2004)[39].**

2.   Aim of this thesis

2.1   Atomic structure of the N-terminal part of coil 2

One goal of this thesis was to obtain further vimentin X-ray atomic structures. Thus far, only two atomic structures of the vimentin rod domain are available, i.e. the coil 1A and coil 2B. Here, we were interested to obtain a crystal structure around the L2 region, which length of eight residues is remarkably conserved between IFs (cf. table 1.2.1).

Early predictions showed that this region maintains a high α-helical content. Moreover, modeling attempts to extend the classical heptad periodicity from coil 2A into the linker 2 were compromised since the apolar stripe formed by residues in *a* and *d* positions in coil 2B would have been externalized. Consequently, it was suggested that the linker 2 was forming particular non-helical structures separating classical coiled coil 2A and 2B (cf. Figure 1.2.1) (Steinert *et al.*, 1994)[41].
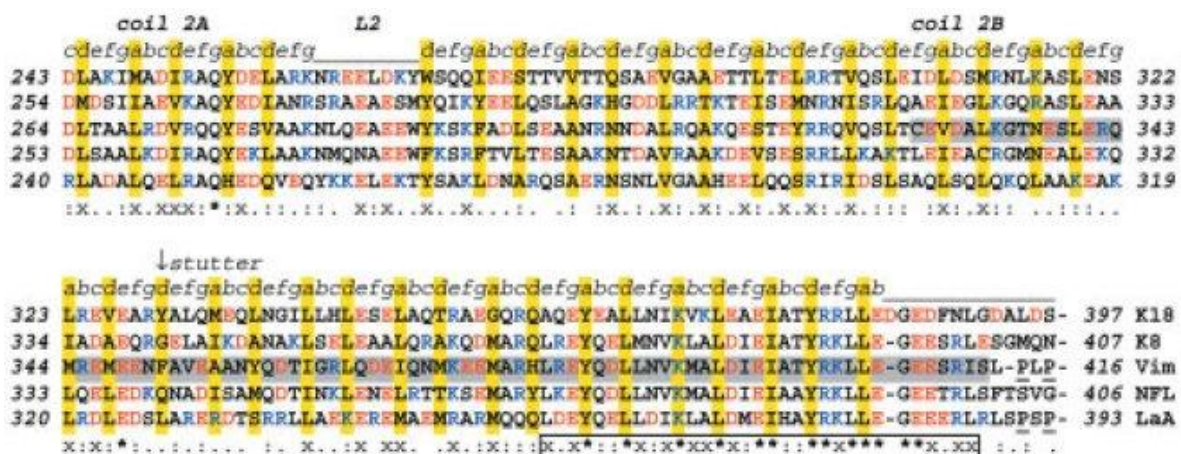
Figure 2.1 Coil 2 alignments of representatives of the IF SHC, the cytokeratins 18(I) and 8(II), vimentin(III), the low neurofilament and lamin A. The heptad repeat is marked as (*abcdefg)* and core positions (*a* and *d*) are high-lighted in yellow. This alignment shows a clear separation between coil 2A and coil 2B. (from Strelkov *et al.*, 2003)[28].

However, it has been recently hypothesized that the sequence of coil 2A and linker 2 corresponds to a hendecad residue periodicity (Parry *et al.*, 2006)[42]. These repeats were already reported by Pauling as a coil possibility and can be treated has a heptad repeat plus a stutter. The existence of coils with hendecad-residue repeat was demonstrated later by crystallography in the RH4 artificial protein (Harbury *et al.*, 1998)[43] and in the tetrabrachion protein (Stetefeld *et al.*, 2000)[26]. Both structures are parallel right-handed coiled coil tetramer. Therefore, it was suggested that the 25 first

residues of coil 2 form a long-period right-handed coil followed by a shorter-pitch classical left-handed coil.

2.2 Impact of myopathic mutations on desmin structure

With desmin, we wanted to understand how some myopathic point mutations located in coil 2B alter the desmin structure. For this purpose, we wanted to obtain crystal structures of a desmin fragment homologous to the previously crystallized fragments of coil 2B in vimentin (cys2) and lamin A (lam1) containing seven myopathic mutations. In addition, some of these fragments were analyzed by gel-filtration, analytical ultracentrifugation and circular dichroism.

Four of these mutations are to proline, a residue often considered as a helix-breaker since it cannot form a hydrogen bond with the residue -4 of the helix. However, helices containing proline after the fourth position exist and have a pronounced kink (cf. Figure 2.2.1) (Woolfson *et al.*, 1990)[44].
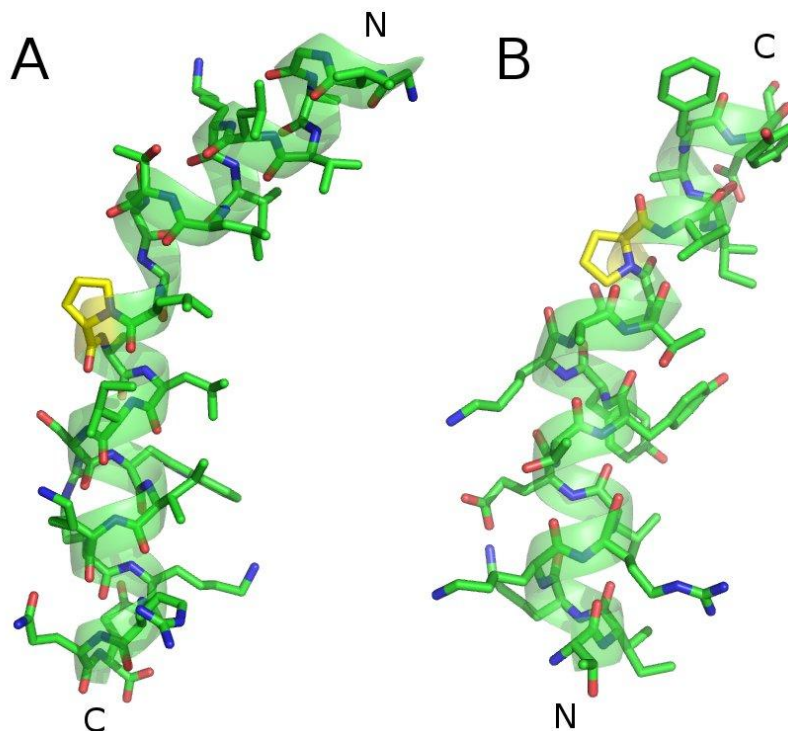
Figure 2.2.1 Illustration of kinked alpha-helices by a proline residue. In A) the mellitin chain A residues 1 to 26 (PDB entry 1MLT) and in B) the adenylate kinase residues 99 to 113 (PDB entry 3ADK).

Surprisingly however, some desmin mutants as A360P assemble to seemingly normal filaments *in vitro* (Bar *et al.*, 2005)[45]. It could thus be imagined that some of these proline mutations, within the much broader context of the desmin rod domain, have only a limited local impact on desmin structure. As an ultimate goal, this study may give some insights into the impact of proline in coiled coils.

In addition to the proline mutations, we were interested in one deletion supposed to produce a stammer (Δ366). To date, stammers are only speculative structures and have never been experimentally observed.

Lastly, the mutation E401S may disrupt a potential salt bridge with R406. However, this salt bridge was shown to only a partial occupancy in available crystal structures of 2B and may consequently not be necessary for the dimer stabilization (Strelkov *et al.*, 2002)[9].

2.3   Small angle X-ray scattering analysis of desmin wt and mutants

With SAXS we will compare and analyze assembly complexes of full-length recombinant desmin myopathic mutants versus desmin wild-type in various buffers. For this, it was shown that the desmin filament assembly can be triggered *in vitro* from a low ionic strength buffer by increasing either the ionic strength or by decreasing the pH (Stromer, 1987)[46]. Therefore, we will first dialyze our 8M urea denatured desmin samples to a low ionic strength buffer called *tetramer buffer*. Then, we will initiate desmin polymerization by adding to these samples different NaCl concentration or *assembly buffer* which both increases the ionic strength and lower the pH. For each sample we will probe their viscosity or gel-formation by pipetting. Only the intensity curves of sample solutions sufficiently fluid to be handled in the SAXS synchrotron machinery will be then be collected.

The resulting desmin intensity curves will be sorted according to their correlation R-factor within a correct angular range window corresponding to the middle part of the curve. This window will be chosen to contain the tertiary structure information of our desmin samples. With this, we will be able to compare and classify desmin mutant assembly complexes versus desmin wt.

Then, a specific treatment of these intensity curves for rod-shaped molecules will be used to calculate the averaged cross-section diameter of desmin assemblies. The cross-section diameter will help us to determine how mutant desmin complexes deviate from wt assembly.

Finally, our SAXS results will be compared to available data in the literature, i.e. desmin pathologies, analytical ultracentrifugation and electron microscopy data.

3.   Material and methods

3.1   Molecular biology techniques

3.1.1 Template vectors

Project 'Atomic structure of the N-terminal part of coil 2

A plasmid containing the full-length human vimentin cDNA (EMBL, Z19544) was kindly provided to us by H. Herrmann (DKFZ, Heidelberg, Germany).

Project 'Impact of myopathic mutations on desmin structure'

We received six pPET-T expression plasmids containing fragments of desmin wild-type and mutants from H. Bär (DKFZ, Heidelberg, Germany). The human desmin protein fragments (residues 332 to 416) are expressed with an N-terminal His-tagged laminin (residues 1799 to 1832) moiety delineated from the desmin fragment by a thrombin cleavage site (LVPR↓ GS) (Figure 3.2.1.1). The pPET-T vector contains an ampicillin resistance gene.
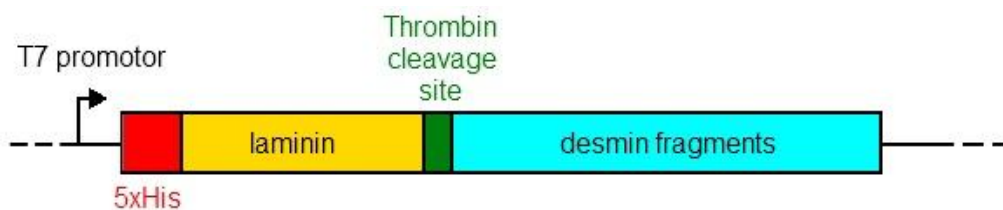
Figure 3.2.1.1 Schematic representation of the pPET-T vector. Desmin fragments are expressed together with an N-terminal His-tagged laminin. Desmin fragments are obtained by cleaving the chimera protein with thrombin.

These desmin plasmids contained the following myopathic point mutations: L345P, A360P, L370P, L385P and E401S.

3.1.2 Primers

All primers were ordered to Microsynth AG, Switzerland.

Project 'Atomic structure of the N-terminal part of coil 2

Six different vimentin DNA fragments (D1-D7) were amplified by PCR from the full-length vimentin plasmid. Forward primers have a *BamH1* restriction site whereas reverse primers contain an *EcoR1* restriction site (Table 3.2.2.1). The melting

temperature Tm was calculated according to the following formula which works for small oligonuclotides (about <20 bp):

$$Tm\ [°C] = 4 \times number\ of\ G·C + 2 \times number\ of\ A·T$$

fragment (sequence)	primers (5' to 3')	Tm [°C]
D1 forward (209-283)	GTGG↓GATCCGATGTTGACAATGCGTCT	52
D1 reverse	TATG↓AATTCTTAGTTCTTGGCAGCCACACT	56
D3 forward (261-335)	ATTTG↓GGATCCAAGCCTGACCTCAC	54
D3 reverse	TATG↓AATTCTTATCCTTTAAGGGCATCCAC	54
D4 forward (261-326)	ATTTG↓GGATCCAAGCCTGACCTCAC	54
D4 reverse	TATG↓AATTCTTAGGTGAGGGACTGCAC	50
D5 forward (251-330)	ATTTG↓GGATCCGAACAGCATGTCCAAATC	52
D5 reverse	TATG↓AATTCTTAGGTGAGGGACTGCAC	50
D6 forward (261-350)	ATTTG↓GGATCCAAGCCTGACCTCAC	54
D6 reverse	TATG↓AATTCTTAGTTCTCTTCCATTTCACG	52
D7 forward (216-310)	ATTTG↓GGATCCGCACGTCTTGACCTTGAA	54
D7 reverse	TATG↓AATTCTTAGCGCAGGGCGTCATT	50

Table 3.2.2.1 Forward and reverse primers used for PCR amplification of vimentin fragments (D1-D7) around the coil 2A region. The corresponding amino acid positions in vimentin are indicated in bracket. All primers are indicated in the 5' to 3' sense. The down arrow (↓) indicates the cleavage position by BamH1 for forward primers and EcoR1 for reverse primers.

Project 'Impact of myopathic mutations on desmin structure'

The table 3.2.2.2 describes oligonucleotides used to remove the N-terminal cystein of initial desmin fragments and to remove Asp366 (cf. stammer chapter 1.6) using the Stratagene Quickchange method. This formula is recommended by Stratagene for long oligonucleotides (>25 bp) :

$$Tm\ [°C] = 81.5 + 0.41(\%GC) - 675/N\ with\ N = primer\ length\ in\ base\ pairs$$

desmin deletions	oligonucleotides (5' to 3')	Tm [°C]
ΔC333 (forward)	CGTGGATCCACCGAGATTGACGCCC	80.7
ΔC333 (reverse)	GGGCGTCAATCTCGGTGGATCCACG	80.7
ΔN366 (forward)	GGCTACCAGGACATTGCACGCCTGG	80.7
ΔN366 (reverse)	CCAGGCGTGCAATGTCCTGGTAGCC	80.7

Table 3.2.2.2 Oligonucleotides used to delete the N-terminal desmin cystein 333 and the asparagine 366 by the Quickchange method.

Furthermore, several shorter desmin fragments were amplified by PCR. The initial plasmids from H. Bär were used as templates. The table 3.2.2.3 describes the primers used for 'shifting' desmin sequences. Forward primers have *BamH1* and

reverse primers *EcoR1* restriction sites. A combinatorial approach was used to generate amplicons. The melting temperature was calculated as in table 3.2.2.1.

'shifts'	primers (5' to 3')	Tm [°C]
N+00 (332)	CCACCCGG↓GATCCACCTGCGAGATTG	60
N+11 (343)	CCACCCGG↓GATCCGACTCCCTGATGAGGCAGATGC	60
C-00 (416)	CCACCCGG↓AATTCTTAGATCCGGCTCTC	60
C-05 (411)	CCACCCGG↓AATTCTTATCCCTCCAGCAGCTTCCG	62

Table 3.2.2.3 Primers used for 'shifting' desmin fragment sequences. Sequence shifts are indicated by N+number of amino acids and C-number of amino acids. The numbers in brackets correspond to the amino acid positions in desmin. N+00 and C-00 primers are equivalent to the template plasmid sequence.

Lastly, we wanted to analyze these mutations in a lamin context. The lamin A wt pPEP-T plasmid was provided to us by L. Kapinos (Biozentrum, Basel). The expressed lamin A fragment (lam1) is homologuous to desmin. The table 3.2.2.4 presents primers which were used to introduce point and deletion mutations by the Stratagene Quickchange method. For convenience, point and deletion mutations are named according to the desmin ones. The melting temperature was calculated as in table 3.2.2.2.

mutations	oligonucleotides (5' to 3')	Tm [°C]
L345P (forward)	CAGCTGGCAGCCCCGGAGGCGAAGCTTC	86.7
L345P (reverse)	GAAGCTTCGCCTCCGGGGCTGCCAGCTG	86.7
A360P (forward)	GCCGGCTGGTGTCCGGCTCACGGGCC	88.7
A360P (reverse)	GGCCCGTGAGCCGGACACCAGCCGGC	88.7
L370P (forward)	CTGCTGGCGGAACCGGAGCGGGAGATG	82.6
L370P (reverse)	CATCTCCCGCTCCGGTTCCGCCAGCAG	82.6
L377P (forward)	GAGATGGCCGAGCCGCGGGCAAGGATG	85.3
L377P (reverse)	CATCCTTGCCCGCGGCTCGGCCATCTC	85.3
L385P (forward)	GATGCAGCAGCAGCCGGACGAGTACCAG	83.7
L385P (reverse)	CTGGTACTCGTCCGGCTGCTGCTGCATC	83.7
ΔN366 (forward)	CACCAGCCGGCGGCTGGCGGAAAAGG	85.5
ΔN366 (reverse)	CCTTTTCCGCCAGCCGCCGGCTGGTG	85.5
E401S (forward)	GCCCTGGACATGTCGATCCACGCCTAC	82.3
E401S (reverse)	GTAGGCGTGGATCGACATGTCCAGGGC	82.3

Table 3.2.2.4 Primers to mutate or delete the lamin A sequence. Mutation residues and numbers correspond to the desmin ones. The lamin mutant 'L377P' was introduced after a mistake in primer design.

3.1.3 Insert preparation by PCR

Phusion polymerase and buffers are from New England Biolabs. The table 3.2.3.1 describes the PCR mixes used during this thesis.

Components	Final concentrations
32.5 µl ddH$_2$O	-
10 µl of 5x Phusion HF buffer	-
2.5 µl dNTPs from 2.5mM stock	100 µM each
1.25 µl Forward primer (20 µM stock)	600 µM
1.25 µl Reverse primer (20 µM stock)	600 µM
1 µl DNA template	-
1 µl DMSO	2%
0.5 µl Phusion DNA polymerase*	0.02 U/µl

Table 3.2.3.1 Components and final concentrations of PCR mix. The Phusion polymerase was introduced in the PCR mix after 'hot-start'.

DNA was amplified in the thermocycler using the parameters of table 3.2.3.2. The annealing temperature was chosen according to the calculated Tm. When PCR reactions were done in parallel, the lowest Tm was used. The Phusion polymerase was added to the PCR mix after 3 min at 95°C ('hot-start').

cycle	T°C (time)	description
1	95 (3min)	hot-start
2	95 (30s)	denaturation
3	~55* (30s)	annealing
4	72 (1min/kbp)	elongation
5		**35 x** back to step 2
6	72 (10min)	
7	4 (pause)	

Table 3.2.3.1 Thermocycler cycles description. The Phusion polymerase was added to the PCR mix after 3 min. The annealing temperature was chosen according to the calculated Tm.

The Qiagen QIAquick PCR Purification Kit was used to clean-up the PCR reaction. As a small modification, the DNA was eluted from the 'spin-column' with 30 µl ddH$_2$O instead of 50 µl of the kit buffer.

After cleaning, the DNA size was controlled by migrating 3 µl of the sample (10%) on 1 x TAE agarose gel (50X TAE stock solution: 2M Tris, 57 ml glacial acetic acid, 100 ml of 0.5M EDTA solution pH 8, water added to 1l)

3.1.4 Mutations and deletions with the QuickChange method

This protocol was adapted from the Stratagen QuickChange manual. The main idea is to amplify the template plasmid by PCR using primers which contain the point mutation(s) or deletion(s) of interest. The 'artificial' plasmids, generated by PCR, do not contain methylated cytosines. Thus, the methyl-cytosine specific restriction enzyme *Dpn1*, which is added to the PCR mix after reaction, cleaves template plasmids but let newly generated plasmids intact. Competent cells cannot be transformed with linear DNA. Consequently, after transformation(see chapter 3.2.8) and plating, the resulting colonies only contain the plasmid harboring the point mutation(s) or deletion(s).

The table 3.2.4.1 describes the PCR mix used for whole plasmid amplification.

Components	Final concentrations
26 µl ddH$_2$O	-
10 µl of 5x Phusion HF buffer	-
10 µl dNTPs from 1 mM stock	200 µM each
1 µl Forward primer (20 µM stock)	400 µM
1 µl Reverse primer (20 µM stock)	400 µM
1 µl DNA template (20 ng/µl)	400 pg
1 µl Phusion DNA polymerase	0.04 U/µl

Table 3.2.4.1 PCR mix for QuickChange PCR reaction.

The thermocycler was programmed as mentioned in table 3.2.4.2.

cycle	T°C (time)	description
1	95 (3 min)	hot-start
2	95 (30s)	denaturation
3	55 (30s)	annealing
4	72 (2 min/kbp)	elongation
5		**16 x** back to step 2
6	72 (10 min)	
7	4 (pause)	

Table 3.2.4.2 Thermocycler program for the QuickChange PCR reaction. The elongation time is overestimated to ensure complete reactions. Here the Phusion polymerase was already introduced into the PCR mix before the 'hot-start'.

After the PCR reaction, 1µl of *Dpn1* is added to the PCR mix and incubated for 3 hours at 37°C. The PCR mix is then heated at 65°C for 15 min to inactivate enzymes before cell transformation.

3.2.6 Digestion

All the vectors and inserts used in this thesis were digested with *BamH1* and *EcoR1*. These enzymes were provided by New England Biolabs (NEB). The mix is presented in table 3.2.6.1.

components	insert	plasmid
ddH$_2$O	10 µl	23 µl
NEB buffer for *EcoR1*	5 µl	5 µl
BSA (diluted 1:10)	5 µl	5 µl
Insert / plasmid	28 µl	5 µl
BamH1	1 µl	1 µl
EcoR1	1 µl	1 µl

Table 3.2.6.1 Digestion of inserts and plasmids for cloning using BamH1 and EcoR1 restriction enzymes.

The double digestions were performed overnight at 37°C. Inserts and plasmid were then purified on a 1.5% TAE agarose gel. Gel slices were cut out and purified using the Qiagen QIAquick Gel Extraction Kit. The DNA samples were eluted in 30 µl ddH$_2$O instead of the recommended kit buffer.
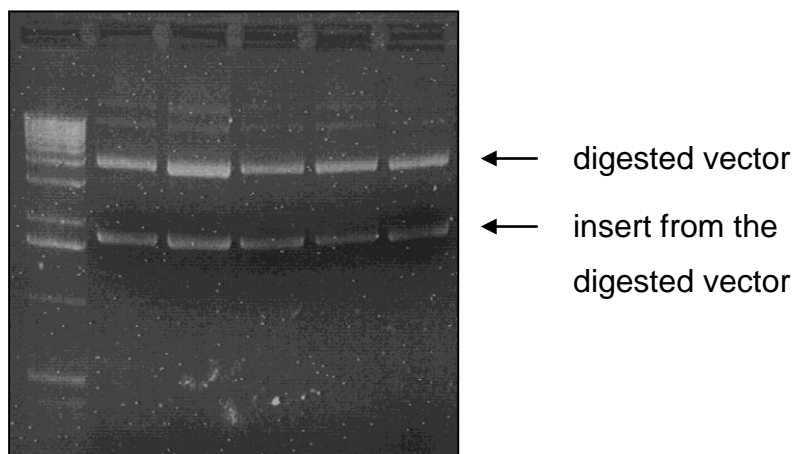
Figure 3.2.6.1 Example of a digested pPET-T vector which was migrated on a 1% TAE agarose gel.

3.2.7 Ligation

3 µl of digested insert and plasmid were migrated together on a 1.5% agarose gel to estimate the ligation ration. The usual insert-to-plasmid ratio evaluated by eyes was

of 3:1. The table 3.2.7.1 summarizes the ligation with the T4 DNA ligase provided from New England Biolabs. All ligations were done overnight at 16°C.

components	volume
ddH$_2$O	4 µl
Insert	3 µl
Vector	1 µl
10X T4 DNA ligase buffer	1 µl
T4 DNA ligase	1 µl

Table 3.2.7.1 T4 DNA ligation of vector and insert as done in this thesis.

3.2.8 *E.coli* transformation

Two type of *E.coli* cells were used during this thesis. The DH5-α competent cells were transformed after ligation to amplify DNA and BL-21(DE3) competent cells were used to produce high amounts of protein fragments. Both cell types were transformed using heat-shock procedure. Briefly, cells are incubated on ice for 20 min together with 1 µl of plasmid. Cells are then 'heat-shocked' for 1 min at 42°C and transferred again on ice for 2 min. The lysogenic broth media (LB) is added to the sample which is then incubated for 1 hour at 37°C. Cells are then platted on LB agar plates containing the 0.1 µg/ml of ampicillin.

3.2.9 DNA purification

Single colonies from plates were grown in 4 ml LB medium and purified using the Qiagen Spin Miniprep Kit. DNA was eluted with 30 µl ddH$_2$O instead of the recommended buffer. All plasmids in this study were sent to Microsynth AG, Switzerland for sequencing.

3.3 Fragments purification

Both desmin and vimentin protein fragments were purified during this thesis. This protocol applies for 'Atomic structure of the N-terminal part of vimentin' and 'Impact of myopathic mutations on desmin structure' projects.

E.coli BL-21 cells transformed with pPEP-T (thrombin cleavage site) or pPEP-TEV (Tobacco Etch Virus protease cleavage site) expression plasmids were grown at 37°C in 1-2 l of LB media supplied with ampicillin to 0.1 µg/ml. At $OD_{600} = 0.7$, the protein fragments expression was induced for 4 hours with 1mM isopropyl β-D-1 thiogalacto-pyranoside (IPTG). Cells were then harvested by centrifugation at 5'000 rpm on a Sigma 6K 15 centrifuge. The cell pellet was then washed in 20 mM Tris pH 7.5 and resuspended in 35 ml elution buffer (10 mM Tris, 100mM NaH_2PO_4, 8M urea pH 8.0). Cells were disrupted by sonication (40% amplitude, 3s/3s on/off, 5 min) and centrifuged for 40 min at 16'000 rpm (Sigma 6K 15) to remove cellular dust. The supernatant was then incubated for 1 hour with Sigma His-Select Nickel Affinity resin pre-equilibrated with elution buffer pH 8.0.

Variant 1 for desmin protein fragments

Desmin protein fragments were eluted by washing the column with elution buffers of lower pH's, respectively of pH 6.3, 5.9 and 4.5. Pure desmin fractions were then dialyzed in 'thrombin buffer' (20 mM Tris pH 8.4, 150mM NaCl, 2.5mM $CaCl_2$) and cleaved overnight at 4°C with 0.75U human thrombin (Sigma T6884) per mg of protein. The digestion was checked on SDS-PAGE and the sample dialyzed back in elution buffer. A second Ni-chelating column was used to remove undigested fragment and His-tagged laminin. Samples were dialyzed back in a gel-filtration buffer (2 mM Tris, 0.1M NaCl, 1mM β-mercaptoethanol, pH 8.4) and a final 'polishing' gel-filtration was performed in a semi-analytic Superdex 75pg 16/60 GE Healthcare column.

Variant 2 for vimentin protein fragments

The vimentin protein fragments were eluted directly with 1M imidazole. Fractions were then analyzed on SDS-page gels. Vimentin protein fragments were dialyzed in 'TEV buffer' (50mM Tris pH 8.0, 100mM NaCl, 2mM DTT) and cleaved overnight at 4°C with 1mg of recombinant TEV protease per 40mg of protein. TEV and thrombin cleavages were then verified on SDS-PAGE. The protein sample was then dialyzed again in elution buffer and passed a second time through a His-Select Nickel Affinity resin to remove undigested fragments and His-tagged laminin fragments. The collected fragments were then dialyzed in gel-filtration buffer (2 mM Tris, 0.1M NaCl,

1mM β-mercaptoethanol, pH 8.4). A final 'polishing' gel-filtration was performed in a semi-analytic Superdex 75pg 16/60 GE Healthcare column.

Protein sample were then concentrated with a centrifugal ultrafiltration device (Amicon, 5kDa cut-off).

3.4 Analytical ultracentrifugation

Some desmin samples were analyzed by analytical ultracentrifugation by Ariel Lustig (Biozentrum, Basel). The vimentin D3 fragment was analyzed by Norbert Mücke (DKFZ, Heidelberg, Germany).

3.5 Circular dichroism

Some desmin samples were analyzed by circular dichroism (CD). Shortly, in a CD experiment, the light is passed through a polarizer and decomposed into its left- and right- circular components. If the polarized light crosses a solution containing polar chromophores, typically protein helices, sheets and random coils, then light left- and right absorption is different. Typical CD spectra recorded for proteins are shown in figure 3.5.1. The instrument was a special Jasco photospectrometer that records the difference in absorption $\Delta A = A_L - A_R$. Using the Beer-Lambert equation for $\Delta A \ll 1$, we get $\Delta A = (\varepsilon_L - \varepsilon_R) \times C \times l$ with the molar concentration C and the light path l. CD spectra are often given in molar elipsity $[\theta] = \alpha \times (\varepsilon_L - \varepsilon_R)$ with $\alpha = 3298 \ m^2 \ mol^{-1}$. Before the experiment, the desmin samples concentration was precisely adjusted to 125 µg/ml. Characteristic minima for α-helices are at 208 and 222 nm. Spectra were then deconvoluted against known protein spectra to estimate the secondary structures content using the program CDNN version 2.1 from Gerald Bohm, Martin-Luther-Universität Halle-Wittenberg, Germany.
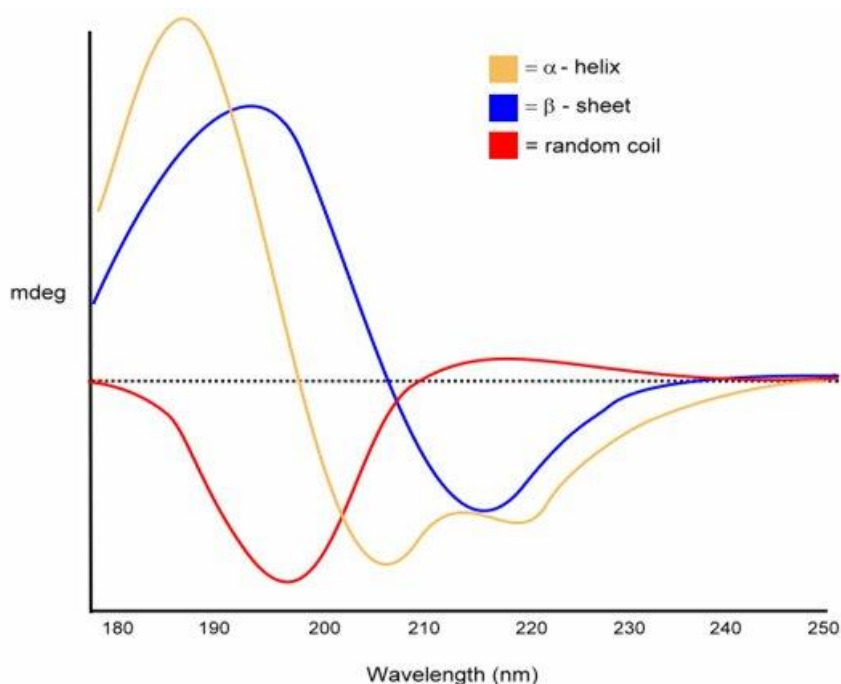
Table 3.5.1 Theorical CD spectra of known protein secondary structures. The α-helices have two characteristic minima at 208 and 222 nm.

3.6 X-ray crystallography

3.6.1 Crystallization

The crystallization conditions were searched based on the sparse-matrix screening technique. The commercial kits used during this study are listed in table 3.6.1.1. Initial protein samples were either kept in gel-filtration buffer or dialyzed into a low ionic strength buffer (10 mM Tris, 1 mM EDTA, 5mM DTT, pH 8.4) to maximize the crystallization mother liquor effect.

crystallization kit (number of conditions)	company
Crystal screen 1 and 2 (96)	Hampton research
Crystal screen cryo 1 and 2 (96)	Hampton research
Index (96)	Hampton research
Additive screen (96)	Hampton research
JCSG+ (96)	Qiagen
AmSO4 Suite (96)	Qiagen
Wizard I and II (96)	Emerald BioSystems

Table 3.6.1.1 Crystallization kits used during this thesis. The Hampton crystal screen 1 and 2 have been 'replaced' by the Index kit. Both the AmSO4 Suite and Additive screen are 'optimization' kits.

Protein samples were concentrated to at least 4 mg/ml and mixed in a 1 µl:1 µl ratio with the crystallization condition using the hanging drop technique. The protein screens were randomly placed at 22°C or 4°C and inspected visually under microscope for at least 1 month. Initial potential conditions showing crystals or crystal-like structures were optimized by refining the kit 'hit', e.g. by modifying the precipitant , salt concentrations, pH's and/or the protein concentration.

3.6.2 Crystal mounting and data acquisition

Crystals were mounted in Hampton cryo-loops and transferred, if necessary, to adequate cryo-solutions preventing formation of ice crystals. Crystals were flash-frozen in liquid nitrogen and kept in until synchrotron data acquisition. As an advantage, the radiation damage is reduced with frozen crystals.

The data acquisition was done at synchrotrons Soleil (France), SLS (Switzerland) and DESY (Germany). There, the mounted crystals were kept frozen under nitrogen flow (100K) during data collection.

The wavelengths for data collection of single wavelength anomalous dispersion (SAD) or multiple wavelength anomalous dispersion (MAD) were chosen according to a prior fluorescence scan. For this, the crystal is exposed to a low dose X-ray radiation and the fluorescence intensity per energy (or wavelength) is recorded. In a SAD experiment, only the peak dataset is used, whereas in a MAD experiment, both peak and inflexion datasets are used.

A typical fluorescence scan as collected for the vimentin crystal D3 is shown in figure 3.6.2.1.
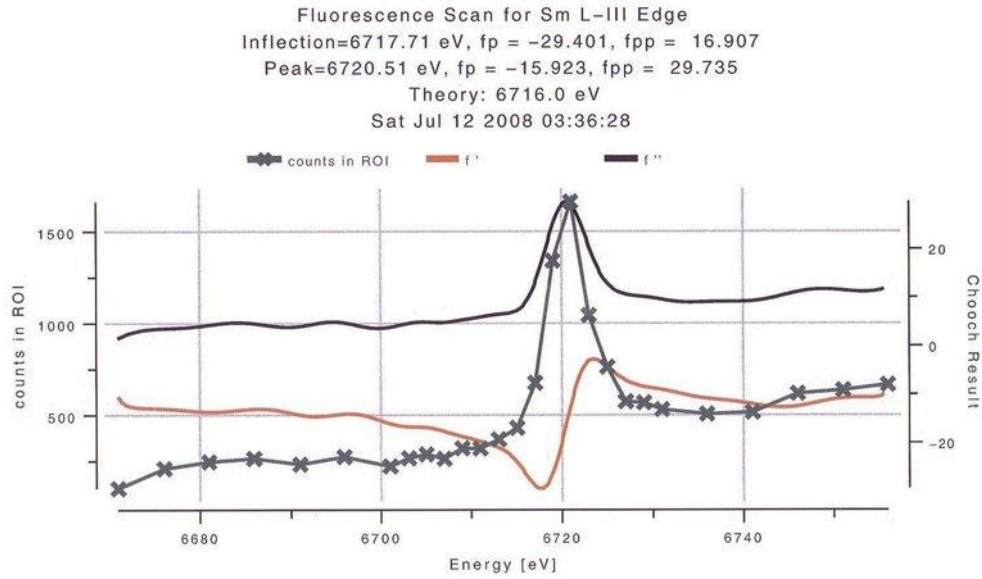
Figure 3.6.2.1 The recorded fluorescence scan for the vimentin D3 fragment. The experimental measured fluorescence intensities are marked by a cross. Inflexion and peak wavelengths are determined from the maxima of the first derivative (f') and the minima of the second derivative (f") curves respectively.

3.6.3 Data processing

Data were processed with the program *iMosflm*[47]. Space-groups having the smallest unit cell and the highest symmetry are always selected. Screw-axes (e.g. $P2_X$) are chosen by looking at systematic absences in reflexions. Additional rotational axis (e.g. $P2_12_x2_y$) can only be determined by processing (or re-scaling) frames and inspecting the R_{sym} value:

$$ Rsym = \frac{\sum_h \sum_i |I_{i,h} - \langle I_h \rangle|}{\sum_h \sum_i I_{i,h}} $$

where $I_{i,h}$ is the *i*th observation of I_h and $\langle I_h \rangle$ is the mean of of N_h measurements.

The additional rotational axis is kept if the R_{sym} value does not change after reprocessing [nb: the *Rmerge* term is sometimes wrongly used instead of the *Rsym* value. *Rmerge* is used to compare different crystal datasets]. Processed data were then scaled using *Scala*[47]. The main purpose of scaling is to correct the intensity decay due to radiation damage.

3.6.4 Phasing and model building

Multi-wavelength anomalous dispersion phasing (MAD) datasets were collected according to the fluorescence peak (Figure 3.6.2.1). The datasets were processed and scaled without merging F^+ and F^- Bijvoet pairs. The electron density map was calculated by the program *autoSharp*[48] and the model was built with *Buccaneer* and *Coot*[47].

3.6.5 Refinement

A particular attention was given to the refinement. For technical reasons, the refinement statistics using the program *Refmac*[47] were worse than with the program *Phenix*[49]. Moreover, statistics were better by using the vimentin D3 peak dataset (R_{work}=0.29, R_{free}=0.33) instead of the native dataset (R_{work}=0.34, R_{free}=0.39). It might be that the heavy atom stabilized the crystal packing. In addition, external restraint options of *Phenix* were used to define α-helical geometry and help to correctly place amino acids at the C-termini were the electron density map quality is poorer. The final model has an excellent geometry (see Table 4.1.1), with a *Molprobity* (Davis et al., 2007)[50] score of 1.7 and 98.9% of residues in favoured regions of the Ramachandran plot with no outliers. The refined structure includes residues 264 to 334. The four N-terminal residues GGSKP (including two exogenous glycines and the first three residues of the vimentin fragment) as well as the C-terminal residue Gly335 are disordered. The final coordinates and experimental structure factors have been deposited to the PDB with accession code 3KLT.

3.6.6 Structure analysis and modeling

The coiled coil parameters (see chapter 1.4) of the vimentin D3 fragment were calculated using the program *Twister* (Strelkov *et al.*, 2002)[24]. This program does not only output coil parameters, but also returns the α-helices and coiled coil axis.

These axes were used to move the N-terminal part (chains A and B) of the vimentin D3 structure forming an unusual tetramer to a physiological dimer using simple geometry. The distance between parallel helices was reduced to 9Å which is twice the average radius of the left-handed coiled coil in the C-terminal part. The model was then idealized with *Refmac*[47].

3.7 Small angle X-ray scattering

3.7.1 Desmin samples source and purification

The desmin mutants samples A213V, E245D, A337P, L345P, R350P, A357P, ΔE359-S361, ΔN366, L370P, L385P, Q389P, D399Y, R406W, N342D, E413K and I451M were provided to us by Harald Bär, DKFZ, Heidelberg, Germany. Brielfy, desmin wt and mutant sequences were cloned into a pDS5 expression vector as previously (Herrmann *et al.*, 1999)[51] and expressed in a JM109 *E. coli* strain. Cells were lyzed with 1% NP-40 and washed with 0.5% Triton X-100 buffers. Inclusion bodies enriched with desmin were solubilized in urea-containing buffer depleted of detergent and loaded on DEAE and CM Sepharose columns (for a detailed protocol see Herrmann *et al.*, 2004)[52]. Fractions were then analyzed on a Coomassie-stained SDS-PAGE (Figure 3.7.1.1 A)

Alternative desmin wild type and mutant samples L345P, A360P, L370P and R406W were provided to us from Lidiya Kurochkina, Intstitute of Bioorganic Chemistry, Moscow, Russia. Here, desmin samples were cloned into a Novagen pET-23b(+) vector and expressed in BL21(DE3) *E. coli* cells. After cell lysis by sonication and centrifugation, the resulting pellet containing desmin and cellular debris was dissolved in 50mM Tris-HCl pH 7.5, 1mM EDTA, 1mM DTT, 8M urea buffer and loaded on a 10ml hand-made Q-Sepharose column. The protein was eluted with a gradient of the same buffer to 0.3M KCl. The fractions purity was probed on Coomassie-stained SDS-PAGE (Figure 3.7.1.1 B)

All desmin samples were provided to us in 8M urea denaturing conditions and frozen.
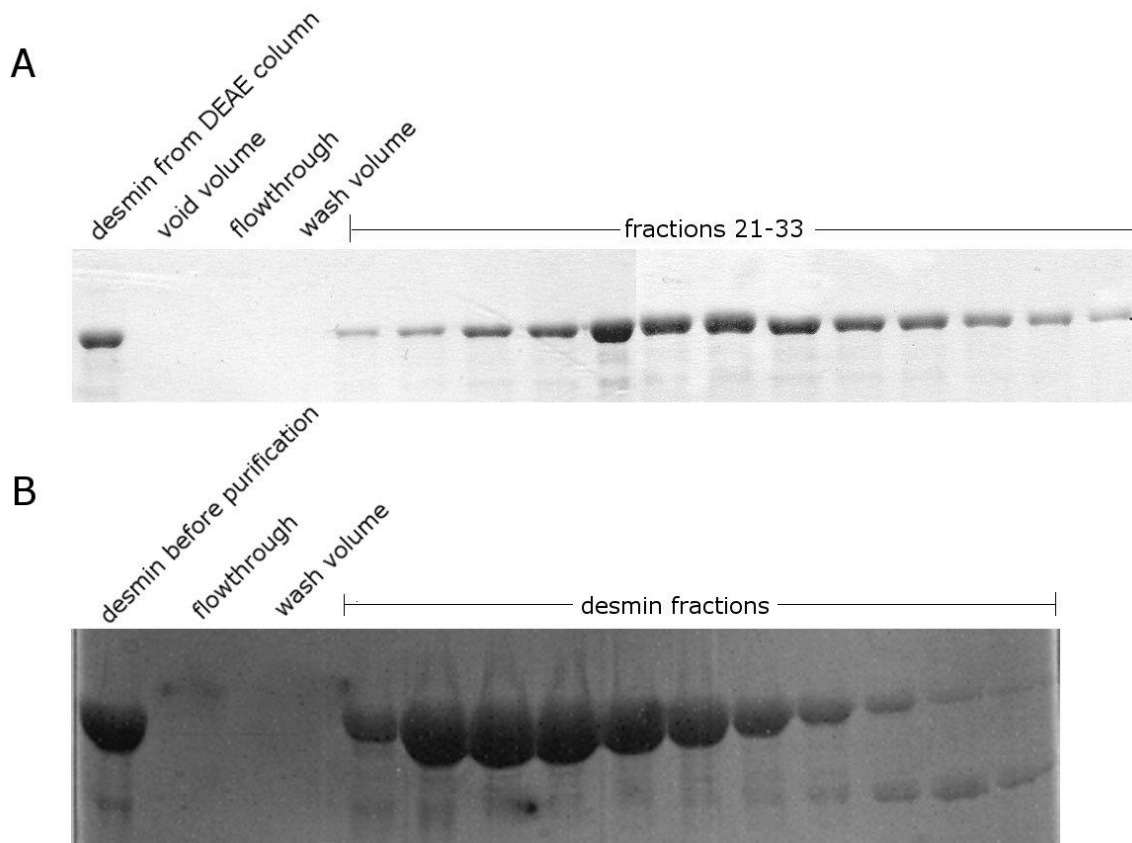
Figure 3.7.1.1 Coomassie-stained SDS-PAGE A) Desmin wt fractions from the CM-Sepharose column elution as done by Bär *et al.* and B) alternative desmin elution from a Q-Sepharose column as done by L. Kurochkina.

3.7.2 Samples preparation

Denatured desmin samples were dialyzed in regenerated cellulose membranes of 3.5kDa cut-off (Spectra/Por) at 4°C in 1l of *tetramer buffer* (5mM Tris pH 8.4, 0.1mM EGTA, 1mM EDTA, 1mM DTT) supplied with 8M urea. Using a peristaltic pump, 4l of *tetramer buffer* were added gradually. The remaining urea (about 1.6M), was then removed by two consecutive dialysis in 4l *tetramer buffers* (about $1/100^2$ v/v dialyze ratio). Samples were concentrated from 0.2 to 5.0 mg/ml in centrifugal ultrafiltration devices (Amicon) of 5 kDa cut-off. The sample concentration was determined by OD at 280 nm.

Desmin polymerization was initiated by adding NaCl to the final salt concentrations of 15, 20, 30 and 75 mM in 50 μl aliquots. Alternatively, *10X assembly buffer* (250mM Tris, 500mM NaCl, pH 7.0) was added to the 50 μl aliquots to a final salt concentrations of 10, 20 and 50mM NaCl. The pH of desmin samples respectively

shifted to pH 8.24, 8.20 and 8.0. The desmin polymerization lasted for at least one hour prior to synchrotron measurements.

Before data acquisition, desmin samples were centrifuged for 10 min at 5000 x rcf to remove insoluble protein aggregates and dust. We could then see either a phase separation, with a lowest phase being very viscous, or one single phase. This single phase was probed by pipetting. In some cases it was found to be fluid or very viscous.

3.7.3 Data collection

SAXS measurements were done at the dedicated SAXS beamline X33 of the Deutsche Elektronen-Synchrotron (DESY), Hamburg, Germany and at the beamline SWING of the Soleil synchrotron, St-Aubin, France.

Desmin samples and buffers were manually injected into the SAXS device (Figure 3.7.3.1) and pumped into the glass capillary exposed to the X-ray beam. The thinness (0.1 mm) of the tubing system prevents data acquisition of gel-forming or too viscous samples. The temperature of the SAXS devices was kept constant at 20°C.
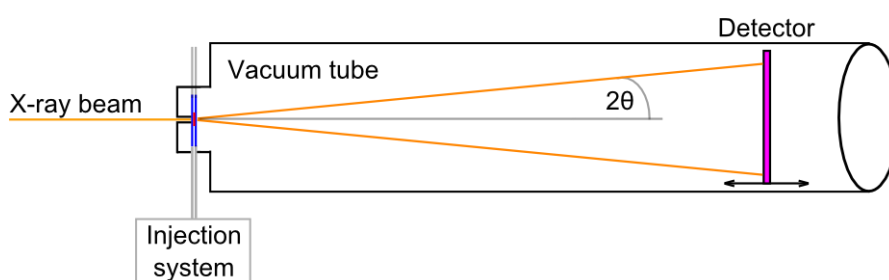
Figure 3.7.3.1 schematic view of a small-angle scattering device. The sample is loaded in the injection system and pumped into the capillary (in blue) exposed to the X-ray beam. The whole device, including the detector (in magenta), is under vacuum to prevent air-scattering. The detector can be moved to have a suitable scattering angle (2θ).

The device records several frames while the sample flows through the glass capillary. These frames were automatically processed at synchrotron places and we obtained text files containing angles versus the averaged intensity. The intensity from diluted monodisperse samples is measured using the following formula: $I(s) = \langle I(s) \rangle_\Omega = 4\pi \int_{r=0}^{\infty} p(r) \frac{\sin(sr)}{sr} dr$ with s being the scattering angle, $\langle I(s) \rangle_\Omega$ the spherical average of the intensity and p(r) the electron density for a sphere radius of radius r. For all measurements, the scattering angle θ was kept below 5° to ensure the relationship $\sin(\theta) \cong \theta$.

3.7.4 Data processing

First, the sample intensity curves were subtracted to their corresponding buffer intensity curves using the program *Primus* (Konarev *et al.*, 2003)[53]. This operation is allowed as X-ray intensity is additive. Thus, the resulting intensity curves only represent the protein scattering. Finally, all intensity curves were adjusted to have the same SAXS angular convention defined as Q [nm^{-1}] = 4π sin(θ) / λ with λ being the X-ray wavelength.

3.7.5 Data classification

A rational method using the correlation R-factor was used to classify our desmin intensity curves. The R-factor between two intensity curves I_1(s) and I_2(s) with Q being the scattering angle of desmin samples was calculated for all datasets. The formula for the R-factor is

$$R = \frac{\sum_{i=1}^{N}\left(I_1(Q_i) - \overline{I_1}\right)\left(I_2(Q_i) - \overline{I_2}\right)}{(n-1)s_{I_1}s_{I_2}}$$

with $I_1(Q_i)$, $I_2(Q_i)$ the datasets intensities at the angle Q_i, $\overline{I_1}, \overline{I_2}$ the mean intensity of each datasets and s_{I_1}, s_{I_2} the standard deviation of each datasets defined as follow:

$$s_{I_{1,2}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_{1,2}(Q_i) - \overline{I_{1,2}})^2}$$

The angular window for the scattering angle Q was chosen to be between 1 and 2.5 [nm^{-1}]. This window contains the tertiary structure information of our samples. Further, this window is nearly insensitive to the intensity deviation due to desmin sample aggregation observed at low angles (Q < 0.7) or to very low signal to noise ratio (Q > 3). Samples were then classified according to their R-factor difference (in percent).

The R-factor calculation was done using a small software called 'Rfactor_simple' written by Anna Sokolova and dedicated for SAXS intensities comparison.

3.7.6 Distance distribution function calculation

The distance distribution function p(r) for monodisperse solutions or the size distribution function D(r) for polydisperse functions can be calculated with the software *Gnom*. The distance distribution function is defined as follow:

$$p(r) = \frac{1}{2\pi^2} \int_{s=0}^{D_{max}} I(s) \frac{\sin(sr)}{sr} ds$$

Concretely, the distance distribution function represents all possible vectors that can be traced within the protein. The D_{max} value corresponds to the longest vector, ie. the protein maximum diameter (Figure 3.7.6.1).
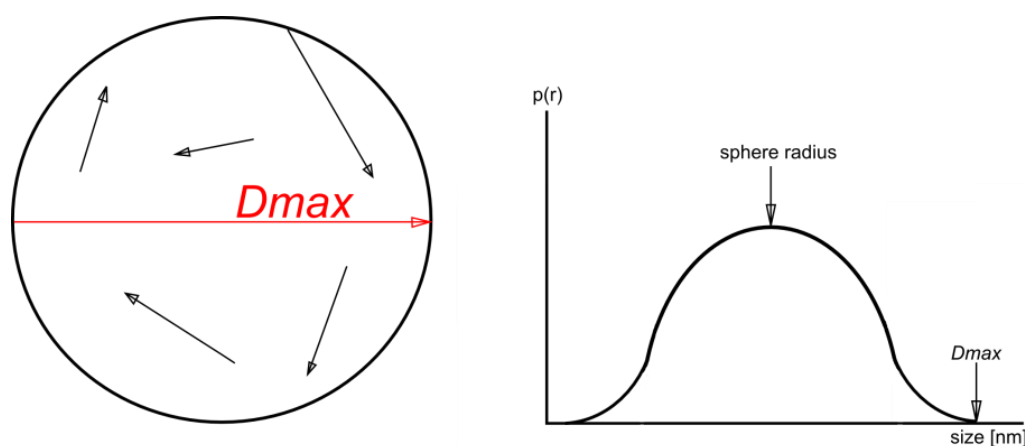
Figure 3.7.6.1 Distance distribution function of a sphere. In (A) the software calculates all vectors within the sphere. The longest vector, *Dmax*, corresponds to the sphere diameter. (B) The resulting Gaussian-like function is centered on the sphere radius. Dmax is found when p(r) = 0. For ellipsoid-shaped particles, the function maximum is shifted to the averaged radius.

Within this thesis, I will deal with a particular case of SAXS processing for rod-like particles. Here, the software *Gnom* was used to calculate the distance distribution of the cross-section P_{cr} (Figure 3.7.6.2). Technically, the cross-section P_{cr} is obtained from the *Kratky-Porod* approximation $I(Q) \propto \frac{\pi}{Q\,L} e^{\frac{-R^2 Q^2}{4}}$ with *L* being the rod-particle length and *R* its radius for *L* being much bigger than *R*.
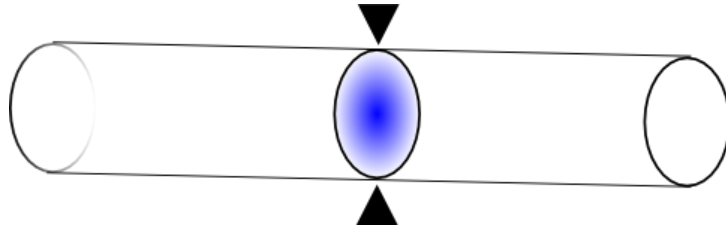
Figure 3.7.6.2 For elongated rod-like particles, the software _Gnom_ calculates an average of the cross-section (black arrows) distance distribution P_{cr} along the whole particle.

To obtain the distance distribution functions of the cross-section, several _Dmax_ values ranging from 10 to 25 nm were tried in _Gnom_. Further, the angular range Q for processing data was usually comprised between 1 and 3 nm^{-1}. This angle range excludes the low scattering angle deviation due to protein-protein interactions and the low signal to noise of high scattering angles (in our case for Q > 3 nm^{-1}). The resulting curves were then smoothed by finding a reasonable α factor.

4 Results

4.1 Atomic structure of vimentin coil 2 domain

4.1.1 Fragments design

Seven (D1–D7) vimentin fragments (Figure 4.1.1.1) covering the predicted linkers L12 and L2 were cloned, expressed in E. coli and purified. While the success of protein crystallization cannot be predicted, working with multiple overlapping fragments can highly improve the chances for obtaining suitable crystals (see e.g. Strelkov *et al.*, 2001)[21]. The secondary structure prediction program Jpred (Cole *et al.*, 2008)[54] predicts the stretch including residues 200–360 to be α-helical, except within the linker L12 region 212 (residues 248–262). The fragments D3, D5, and D6 were designed to overlap with the previously crystallized cys2 fragment. Fragments D1, D2, and D7 included parts of coil 1B and the whole length of the predicted linker L12, in the hope to support the presumably flexible linker by α-helical segments on both sides.
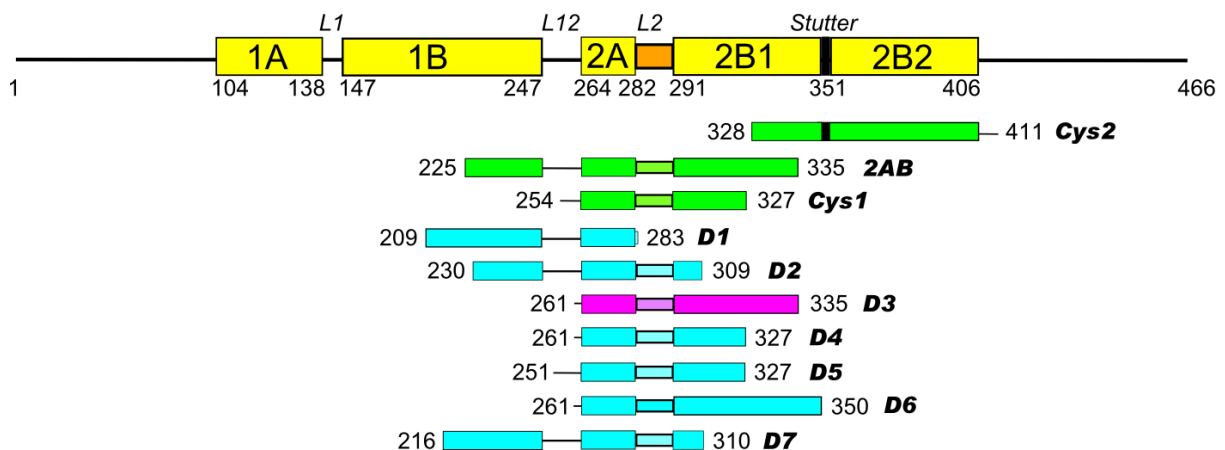
Figure 4.1.1.1 Schematic view of human vimentin chain and the used fragments. Above, the widely accepted predicted structure of an IF chain including four coiled-coil segments 1A, 1B, 2A, and 2B is shown (see e.g. Strelkov *et al.*, 2003)[28]. Below, the fragment cys2 with the known crystal structure (Strelkov *et al.*, 2002)[9] as well as the other designed fragments are indicated.

4.1.2 Overall structure of the D3 fragment

After extensive crystallization screening, only crystals for the fragment D3 including residues 261–335 (Figure 4.1.1.1) could be obtained. The refined D3 crystals diffracted X-rays to 2.4 Å resolution (see Table 4.1.2.1). After MAD phasing, the obtained electron density map (Figure 4.1.2.1A) could be readily used for model tracing. It shows a purely α-helical structure. Four protein chains were found within the crystallographic asymmetric unit. There are two parallel in-register dimers formed by chains A and B, and also chains C and D. These dimers feature a standard left-handed coiled coil in their C-terminal parts (residues 303–334). Most interestingly, the dimers overlap in their N-terminal parts (residues 264–302) to yield a symmetrical tetrameric assembly which features a nearly straight α-helical bundle (Figure 4.1.2.1B). The nearest neighboring helices (A and C, A and D and so forth, see Figure 4.1.2.1B) of the bundle are in antiparallel orientation, so that the residues Asn283 of all four chains are aligned.
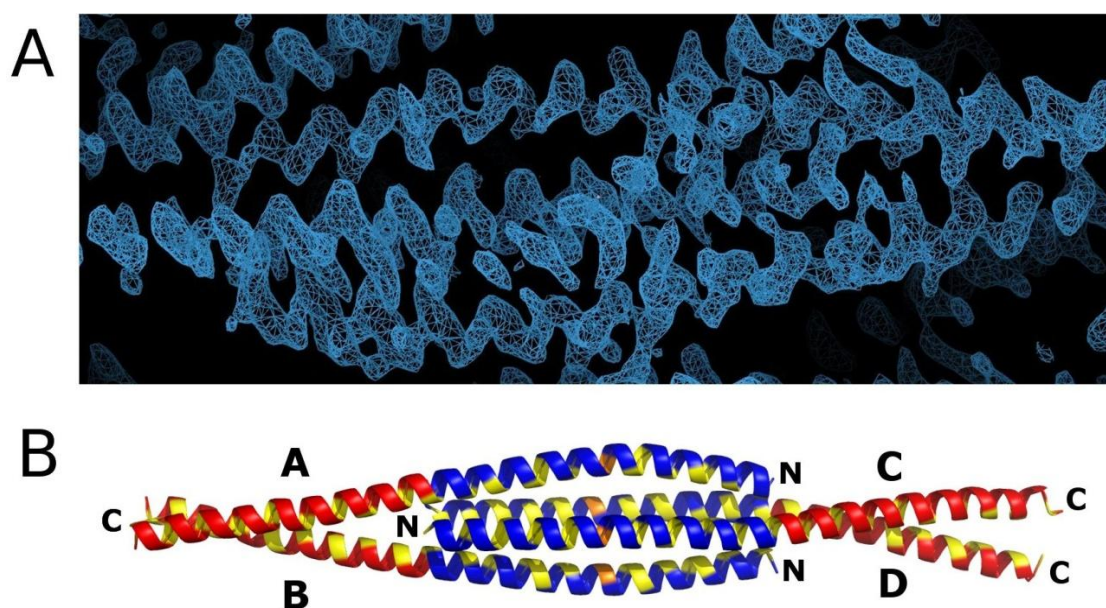
Figure 4.1.2.1 Crystal structure of the D3 fragment. (A) Experimental MAD-phased electron density map after solvent flattening. (B) Ribbon diagram of the structure. The N-terminal overlap part forming the tetramer is shown in blue. C-terminal parts forming coiled coils are in red. The residues in core positions of the heptad and those of the hendecad repeats are in yellow (compare Figure 4.1.3.1B). Residues Asn283 are shown in orange.

Crystallographic statistics.

	Native	Sm derivative (peak)
Wavelength (Å)	1.072	1.8449
Space group	$P2_1$	
Unit cell dimensions (Å)	63.8, 30.4, 82.6	63.1, 30.1, 82.8
Unit cell angles (°)	$\alpha = \gamma = 90$, $\beta = 101.7$	$\alpha = \gamma = 90$, $\beta = 102.7$
Matthews coef (Å^3/Da)/ Solv fract (%)	2.3/47	
Resolution	37.6–2.4 (2.5–2.4)	44.4–2.6 (2.74–2.6)
R_{sym}* (%)	5.7 (53.3)	3.9 (43.3)
I/σ	13.5 (2.5)	15.2 (2.8)
Unique reflexions	13,067	9176
Completeness (%)	99.8 (100)	95.5 (91.8)
Ano. completeness (%)		94.9 (90.2)
Multiplicity	4.1	
Anomalous multiplicity	2.1	
Mosaicity (°)	0.81	0.67
Wilson B-factor (Å^2)	61.3	77.5
Model refinement		
R_{work}/R_{free}**		0.291/0.331
Number of atoms		
Total		2399
Protein		2290
Ligand/ion		102
Water mol.		7
Average B-factor (Å^2)		
Overall		78.3
Protein		78.7
Ligand/ion		77.4
Water mol.		59.2
Rms deviations		
Bond lengths (Å)		0.002
Bond angles (°)		0.369

Data in parentheses are for the highest resolution bin.

* $R_{sym} = \Sigma_h \Sigma_i |I_{i,h} - \langle I_h \rangle| / \Sigma_h \Sigma_i I_{i,h}$, where $I_{i,h}$ is the ith observation of I_h and $\langle I_h \rangle$ is the mean of N_h measurements.

** $R = \Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$, with the 'free' value calculated with 6% of reflections.

Table 4.1.1.2 Crystallography statistics of the vimentin D3 fragment.

At this point, we used analytical ultracentrifugation to investigate whether the observed crystallographic tetramer is the solution form. Equilibrium runs with the D3

fragment (0.25–0.9 mg/ml) in either 5 mM Tris–HCl buffer, 1mM DTT (pH 8.4) or 5mM Tris–HCl buffer, 160 mM NaCl, 1mM DTT (pH 7.5) revealed average molecular masses of 8.7 and 8.4 kDa, respectively. This is in a perfect agreement with the calculated mass of the monomeric fragment (8.6 kDa). Hence we conclude that up to at least 1 mg/ml concentration there is no noticeable oligomerization of the D3 fragment in solution. During crystallization however, concentrated (5 mg/ml) protein solution was exposed to high concentrations of PEG in the presence of Ca2+ ions, which apparently drives the formation of the tetramers and their subsequent crystallization. It should be noted in particular that in the resulting crystal structure five stretches of PEG molecules were found to be bound inside the relatively 'loose' tetramer, making multiple hydrogen bonds with the protein (to be discussed below).

4.1.3 C-terminal part (residues 303-334) is a left-handed coiled coil

The geometry of the D3 structure was analyzed using the program Twister (Strelkov and Burkhard, 2002)[24] which in particular computes the local coiled-coil radius and pitch as a function of residue number (Figure 4.1.3.1A). On input, a pair of parallel chains (i.e. either dimer AB or CD) was taken, and the results of dimers AB and CD were averaged. The C-terminal part (residues 303–334) was found to have a typical left-handed coiled-coil geometry with an average pitch of about 140 Å and an average radius of 4.5 Å. In comparison, the GCN4 dimer has a pitch of 148 Å and a radius of 4.9 Å (Harbury *et al.*, 1993)[55]. Furthermore, the program Twister can assign the heptad positions in standard left-handed coiled coils. Thereby the first position of the heptad (*a*) is recognized by a small positive value (~20°) of the so-called Crick angle which defines the side-chain position with respect to the coiled-coil axis (for details, see Strelkov and Burkhard, 2002)[24]. For the C-terminal part of the structure, the heptad assignment found by this structural criterion (Figure 4.1.3.1B) coincides with the assignment previously made on the basis of sequence analyses (see e.g. Fig. 1B in Strelkov *et al.*, 2003)[28].
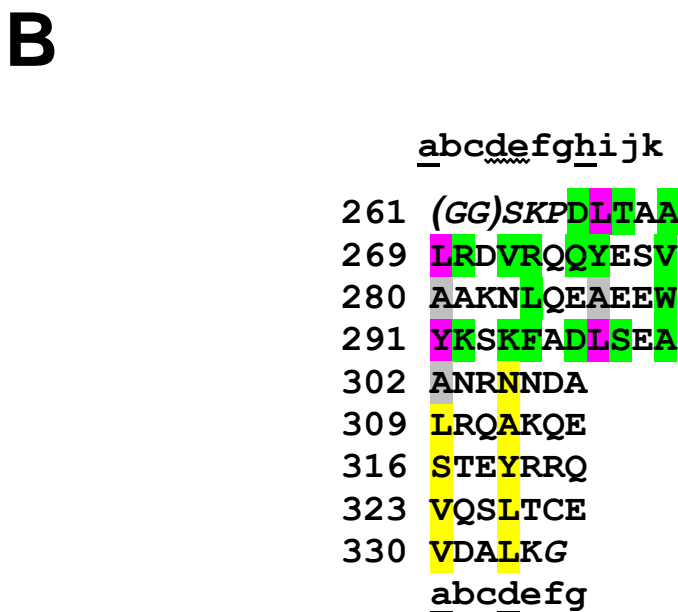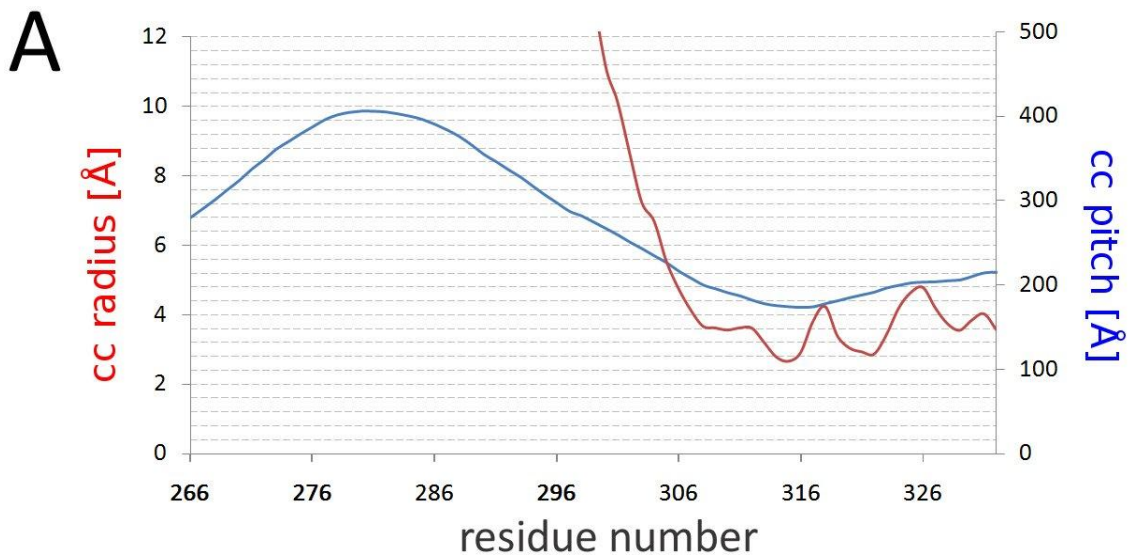
Figure 4.1.3.1 Geometry and repeat patterns within the D3 fragment. (A) The coiled-coil radius (red) and pitch (blue) as a function of residue number. (B) Amino-acid sequence formatted to show the hendecad repeat pattern (abcdefghijk) in the N-terminal part up to residue 302 and heptad repeats (*abcdefg*) in the C-terminal part. The residues highlighted in magenta contribute to the hydrophobic core of the N-terminal tetrameric part and are involved in the interaction of both the 'nearest neighbor' (antiparallel) and diagonally located (parallel) α-helices (see text for explanations). The residues highlighted in green are only involved in the 'nearest neighbor' interactions, as determined by the PISA server (Krissinel and Henrick, 2007)[56]. The residues in *a* and *h* positions highlighted in gray point towards the axis of the tetramer but however are too far away to form a dense hydrophobic core. The residues forming the core of the C-terminal coiled coil are highlighted in yellow. The actual vimentin sequence starting with residue 261 is preceded by two exogenous glycines. Disordered residues at N- and C-termini are shown in italics.

4.1.4 N-terminal part (residues 264-302) is a parallel α-helical bundle

Within the N-terminal parallel bundle, the program Twister reveals a left-handed geometry for residues 264–277 and downstream of residue 298 but a right-handed one for residues 278–297. However, the coiled-coil pitch value exceeds 700 Å everywhere up to the residue 297 indicating that the supercoiling is negligible. However, the coiled-coil radius, which is equal to half the distance between the parallel helices (A–B or C–D), varies considerably. It starts at a value of about 7 Å near the N-terminus, increases to a maximum of 9.9 Å at residue 281 and decreases again to 6.1 Å at residue 302 (Figure 4.1.3.1A). This implies, on one hand, that the helices within the N-terminal stretch are curved (just like helices forming a standard coiled coil (Strelkov *et al.*, 2002)[9] despite the absence of the supercoiling). On the other hand, the tetramer bundle reveals a pronounced 'bulging' in the middle (residue 283), with α-helices separating much more than in a typical four-stranded coiled coil (such as e.g. the tetrameric GCN4 mutant which has a coiled-coil radius of 7.6 Å (Harbury *et al.*, 1993)[55]. The possible reasons for this 'bulging' lie in the packing of the tetramer, to be discussed below.

Furthermore, residues 269, 280, and 291 have small positive values for the Crick angles that are reminiscent of residues in the *a* positions in a heptad repeat. In the parallel bundle, these residues mark the start points of three consecutive hendecad repeats *abcdefghijk*, where *a*, *d*, *e*, and *h* are the core positions. This assignment largely coincides with the sequence-based prediction by Parry (2006)[42]. As should be expected for a hendecad-based structure (see Figure 2D in Kuhnel *et al.,* 2004)[57], all residues in the *a* and *h* positions point roughly in the direction of the bundle axis. However, the direct contact between the pairs of chains belonging to the same dimer (i.e. A–B and C–D) is limited. Indeed, it is only the residues Leu265 (position *h*), Leu269(*a*), Tyr291(*a*), and Leu298(*h*) that are directly packing against each other, while the large distance between the helices in the middle of the tetramer precludes a direct contact of residues Tyr276(*h*), Ala280(*a*), and Ala287(*h*) of these parallel chains. At the same time, the major factor stabilizing the tetramer appears to be the hydrophobic contacts between the nearest neighbor antiparallel helices, such as between helices A and C, A and D, etc. (see Figure 4.1.2.1A). Notably, nearly every residue in *a*, *b*, *d*, *e*, *g*, *h*, and *k* positions is involved in this contact, as highlighted in Figure 4.1.3.1B. This behavior is well correlated with the predominantly hydrophobic

character of residues in these positions. In contrast, the residues in positions *c*, *f*, *i*, and *j* are mostly polar (see Figure 4.1.3.1B).

As already mentioned, the middle part of the tetramer centered on residue 283 features a very loose packing. In particular, the core residues Ala280(*a*), Asn283(*d*), and Ala287(*h*) are not engaged in any interhelical contacts. This is linked to the fact that, with the coiled-coil radius up to 9.9 Å in this region, the distance between the axes of the adjacent antiparallel helices and the parallel helices belonging to the same dimer can be as large as 14.0 and 19.8 Å, respectively. Interestingly, the voids between the helices in this region were found to be occupied by five partially ordered polyethylene glycol molecules (Figure 4.1.4.1A). Four of these are located in a quasi-symmetrical way (one per protein chain), wrapped around the side chain of Asn283, and reside in a hydrophobic pocket formed by residues Trp290, Tyr291, Tyr276, Val279, and Ala287. A reason for the described 'bulging out' in the middle part of the tetramer may lie in the presence of two symmetrical aromatic clusters (Figure 4.1.4.1A). Such cluster is formed by the Tyr276 residues of two 'upstream' chains and the Trp290 and Tyr291 residues of two 'downstream' chains. These residues are located roughly in the same plane perpendicular to the tetramer axis (Figure 4.1.4.1B). Since accommodating these bulky residues requires considerable volume, the helices are pushed further to a larger distance from each other than those near either end of the tetrameric bundle. Apparently, due to the intrinsic rigidity of an α-helix this further causes a yet larger separation of the chains near residue 283.

A

39
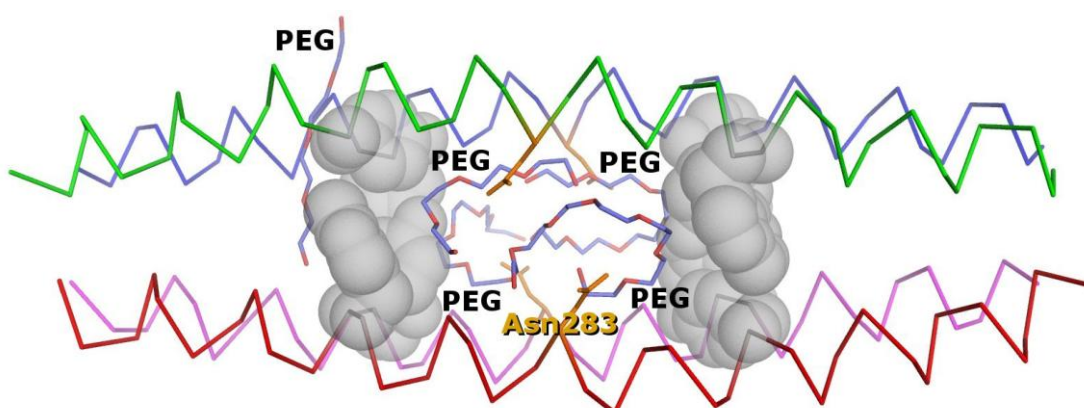
B
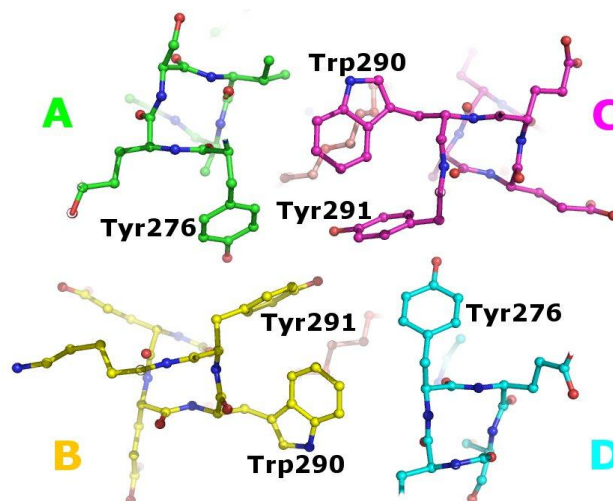
Figure 4.1.4.1 Structural features of tetrameric α-helical bundle formed by residues 264–302. (A) Structure diagram showing the five bound PEG molecules. Side chains of Asn283 are in orange. Two aromatic clusters formed by the side chains of Tyr276, Trp290, and Tyr291 are shown in gray. (B) The aromatic cluster of Tyr276, Trp290, and Tyr291 viewed along the tetramer axis.

4.1.5 Native dimer model

As discussed above, within the α-helical bundle part the distance between the two helices of the same D3 dimer is considerably higher than expected for the full-length native dimer structure. Consequently, the native situation was modeled by bringing the N-terminal parts of the helices closer. The distance between the α-helical axes was chosen to be 9 Å which is twice the average radius of the left-handed coiled coil in the C-terminal part. The resulting model was energy-minimized whereby the side chains conformations were adjusted to avoid steric clashes (Figure 4.1.5.1). In agreement with the theoretical predictions (Kuhnel *et al.*, 2004; Peters *et al.*, 1996)[57,58], the helices within the constructed dimer interact mainly via hydrophobic packing of the residues in *a*, *d*, *e*, and *h* positions of the hendecad repeat (compare Fig. 3B). Starting from the N-terminus, the residues Leu265(*h*), Leu269(*a*), Val272(*d*), Arg273(*e*), and Tyr276(*h*) (all representing frequent side-chain types in the core positions of coiled coils (Walshaw and Woolfson, 2001)[59] appear to provide for a reasonable hydrophobic core). However, the downstream core residues Ala280(*a*), Asn283(*d*), and Ala287(*h*) are all suboptimal for a good hydrophobic stabilization. The next core residue Tyr291(*a*) is preceded by a bulky Trp290(*k*). Whereas in the

40

crystallographic tetramer this tryptophan is inserted into the hydrophobic core (Figure 4.1.4.1B), in the native dimer it is highly exposed (Figure 4.1.4.5).
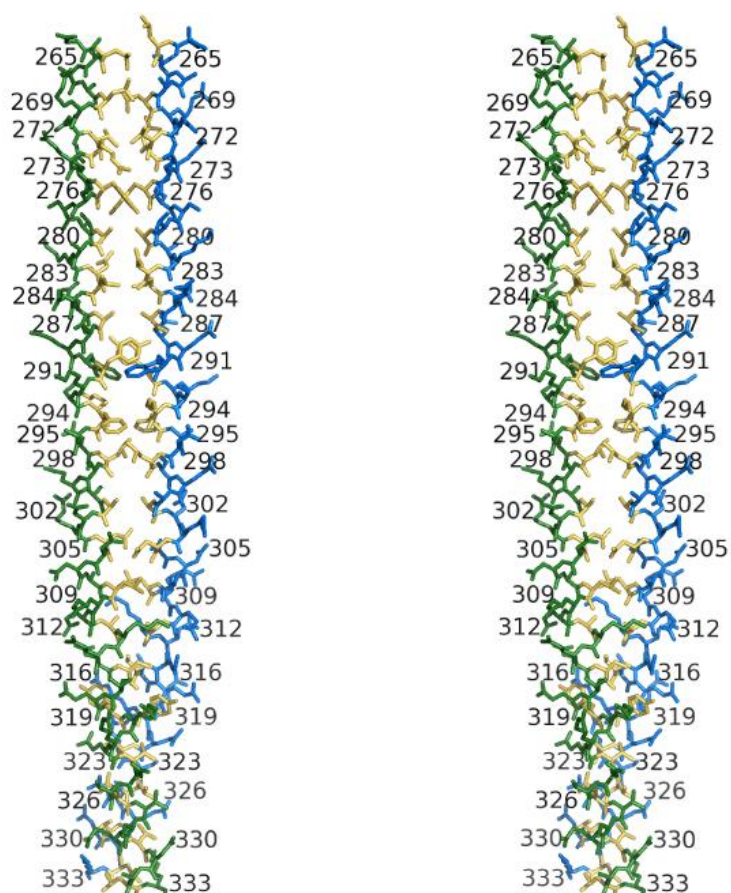
Figure 4.1.4.5 Stereo image of the native D3 dimer model. Residues in hydrophobic core positions are in yellow.

4.2 Impact of myopathic mutations on desmin structure

4.2.1 Fragment design

The initial expression pPEP-T vectors of desmin wt and myopathic mutants L345P, A360P, L370P, L385P and E401S fragments were provided to us by Harald Bär, DKFZ, Heidelberg, Germany. In addition, a short protein fragment of the deletion mutant ΔN366 (stammer) was synthesized chemically. This deletion was later introduced in our desmin wt fragment pPEP-T vector. To maximize our chances of crystallization, all our desmin fragments (residues 332-416) are homologous to the previously crystallized vimentin *cys2* fragment (residues 328-411) (pdb entry 1gk4, ref) and lamin A *lam1* fragment (residues 305-387) (pdb entry 1x8y, ref). The purification of desmin and *lam1* fragments involves a thrombin cleavage (LVPR/GS) which lets exogenous GS residues at their N-termini. In contrast, such extra-residues are absent of the *cys2* fragment obtained by chemical cleavage of the full-length recombinant desmin at two cystein positions (Figure 4.2.1.1).

```
desmin   GSTCEIDALKGTNDSLMRQMRELEDRFASEASGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEGEESRI-
cys2     ---CEVDALKGTNESLERQMREMEENFAVEAANYQDTIGRLQDEIQNMKEEMARHLREYQDLLNVKMALDIEIATYRKLLEGEESRIS
lam1     -GSCQLSQLQKQLAAKEAKLRDLEDSLARERDTSRRLLAEKEREMAEMRARMQQQLDEYQELLDIKLALDMEIHAYRKLLEGEEERL-
            *::. *:     :    ::*::*: :* *     :  :.. : *: .:: .* ::* ***:**::*:***:** :*********.*:
```

Figure 4.2.1.1 Multiple alignments of the desmin wt fragment with vimentin cys2 and lamin A lam1. Conserved residues (*) are in red and similar residues (:) in green. Desmin fragments have 75% identity with *cys2* and 41% with *lam1*. The alignment was produced with NPS@.

We choose mutants in coil 2B which belong to different assembly groups as observed by Bär *et al.* (2005, 2006)[45,59]. Further, these mutants cover different positions of the heptad repeat *abcdefg* (Table 4.2.1.1). Last but not least, four of these mutations are replacements to a proline residue expected to cause a pronounced local distortion of the helical structure due to the sterical properties of this amino acid (cf Figure 2.2.1).

desmin mutant	heptad position	assembly group
L345P	*d*	IV
A360P	*a*	I
ΔN366	*g*	-
L370P	*d*	IV
L385P	*e*	II
E401S	*g*	-

Table 4.2.1.1 Mutants studied in this thesis with their heptad position and corresponding assembly group according to Bär *et al.* (2005, 2006)[45,59]. Mutants of the assembly group I form seemingly normal filaments. Mutants of the group II have a regular unit-length filament (ULF) formation but reduced longitudinal annealing properties. Mutants of group III (not studied here) form heavy aggregates and mutants of group IV form short-living assembly precursors that rapidly disintegrate.

The A360P mutant (assembly group I) forms filaments quite similar to the wild-type ones both *in vitro* and *in vivo*. However, the measurements using STEM reveal that the A360P filaments have a somewhat larger width (14.5±1.3 nm) than the wild-type filaments (12.6±1.3 nm) (Bär *et al.*, 2006)[59]. These filaments also appear more flexible and sometimes kinked and branched. The L385P mutant (assembly group II) is capable of forming ULFs which, however, fail to elongate and radially compact properly. Correspondingly, after the initiation of assembly the viscosity of the L385P solution does not increase significantly past the value typical for isolated ULFs. Finally, the L345P and L370P mutants (assembly group IV) are capable of forming ULF-like structures which, however, aggregate rapidly. The deletion ΔN366 is predicted to produce a 'stammer' structure (cf Figure 1.5.3). However, this mutation has never been observed in any coiled coil structure (Kaminska *et al.*, 2004; Burkhard *et al.*, 2001)[39,45,60]. Finally, the mutation E401S may disrupt a predicted interhelical salt bridge between residues E401 and R406.

4.2.2 Purification, characterization and stability of initial desmin fragments

The desmin wt fragment and mutants were successfully purified in Ni-chelating columns and each yield ~30 mg/ml of pure protein per litre of culture. To the exception of the L385P mutant which was less soluble, all of them were concentrated to 7-10 mg/ml.

We also investigate the oligomeric state of our desmin samples. For this purpose, analytical ultracentrifugation experiments were carried out by Ariel Lustig, Biozentrum, Basel. Interestingly, both the wild-type and L345P mutant fragments are dimers in solution, whereas the A360P and E401S mutant fragment tend to be monomers (Table 4.2.2.1).

	Calculated mass (kDa)	AUC average (kDa)	Type
Desmin wild-type	20.3	18.1	dimer
Desmin L345P	20.3	23.6	dimer
Desmin A360P	20.3	12.5	monomer + dimer
Desmin E401S	19.9	13.3	monomer + dimer
Desmin stammer (synthetic)	12.7	7.3	monomer

Table 4.2.2.1 Analytical ultracentrifugation results of desmin wild-type and mutants L345P, A360P and E401S; the wild-type and L345P desmin fragment are dimers whereas the mutant fragments A360P and E401S are mostly monomeric. All samples were measured at three different concentrations and two speeds (28'000 and 34'000 rpm). The desmin stammer sample has been measured at three different speeds (26'000, 32'000 and 42'000 rpm).

All our protein fragments were screened towards crystallization using the hanging-drop technique in commercially available kits. Since no crystal hits were found, we investigated the stability of the desmin wild-type fragment. For this, our wild-type sample at 7 mg/ml kept during one month at 4°C was eluted in a semi-analytic chromatography column (Superdex 75pg 16/60) separating proteins by size from 3 to 70 kDa. The major part of this sample eluted in the void volume (i.e. >70kDa) (Figure 4.2.2.1A), which means that it forms aggregates or polymers of at least eight monomers (10.2 kDa / monomer). Further, the void volume peak is not a consequence of our high sample concentration, since the same sample diluted to 1 mg/ml applied to the gel-filtration column yields the same void volume peak to dimer peak ratio (Figure 4.2.2.1B). However, the dimer peak gets more prominent when the protein was denatured in 8M urea and renatured in our protein buffer containing fresh DTT prior gel-filtration (Figure 4.2.2.1C). This suggests that DTT oxidizes over time allowing the N-terminal cystein (residue 333) residues to form cross-bridges yielding to aggregates.
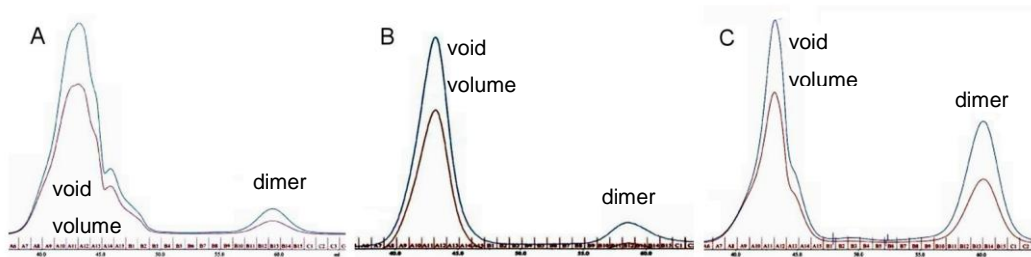
Figure 4.2.2.1 Gel-fitration elution profiles of the desmin wild-type fragment at 7 mg/ml containing the N-terminal cystein after one month of storage at 4°C (A). The same sample diluted to 1mg/ml (B) has the same void volume peak to dimer ratio. In the same sample (7 mg/ml) denatured in 8M urea and renatured in a buffer containing fresh DTT, the dimer peak gets more prominent (C)

Consequently, we envisaged to produce another set of pPEP-T vectors with desmin wild-type lacking the N-terminal cystein.

4.2.3 Desmin wild-type and mutants protein fragments lacking the N-terminal cystein

The seven desmin fragments lacking cystein (Figure 4.2.3.1) were expressing at the same level than of the cystein-containing desmin fragments, yielding also ~30 mg/ml protein per liter of culture. These fragments were gel-filtered and screened towards crystallization.

Figure 4.2.3.1 Desmin fragments (residues 332-416) lacking the N-terminal Cys333.

We showed by circular dichroism spectra (cf. Chapter 3.5) that all fragments have a predominance of α-helical structure (Figure 4.2.3.1A). The differences in the spectra are caused by scaling due to inaccuracies in protein concentration determination. The cystein deletion does not alter the α-helical content. The gel-filtration shows that the sample 'aggregation' (void volume) is reduced (Figure 4.2.3.1B). However, despite the fact that the quality of desmin samples was largely improved, the crystallization screen did not yield crystal. Therefore, we opt for another crystallization strategy by modifying the fragment sequences.
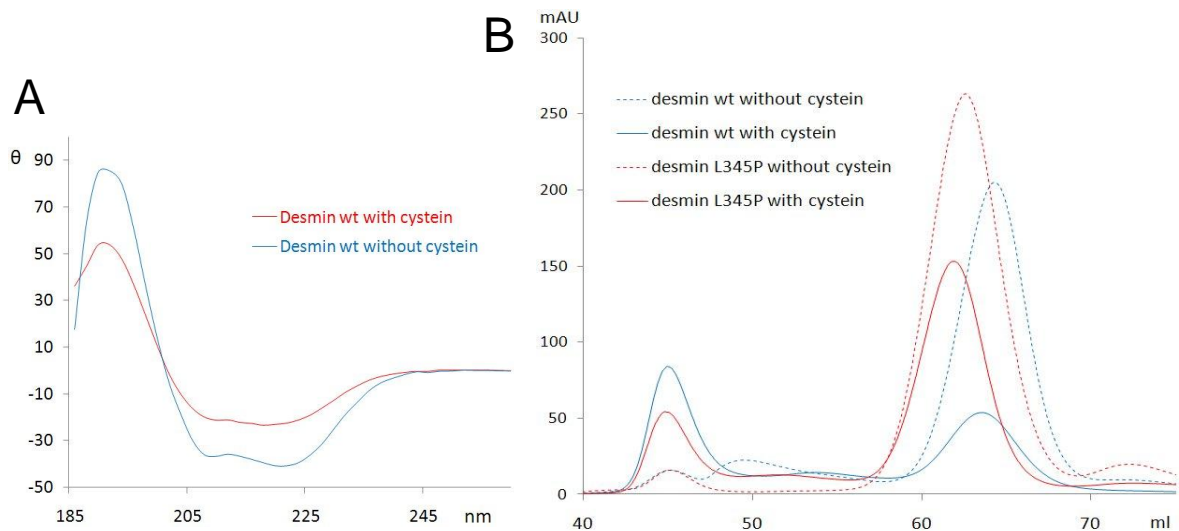
Figure 4.2.3.1 (A) Circular dichroism spectra of the desmin wild-type protein fragment (residues 332-416) and of the desmin protein fragment lacking the N-terminal cystein Cys333. Minima at 208 and 222 nm are characteristic of α-helical samples. (B) Gel-filtration elution profiles of desmin wt (in blue) and mutant L345P (in red) with and without cystein (continuous and dashed line respectively). The void volume peaks are nearly absent in cystein-depleted fragments.

4.2.4 Desmin wild-type and mutant sequence 'shifts'

The pPEP-T constructs with sequence shifts were generated by combining forward primers starting eleven residues after the N-terminus (at residue 343) and reverse primers lacking five residues at the C-terminus (at residue 411) compared to our desmin original plasmids (residues 332-416). The three possible derived wild-type sequences were cloned as well as six randomly chosen mutants (Figure 4.2.4.1).

```
GSTCEIDALKGTNDSLMRQMRELEDRFASEASGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEG----    Des Wt 332-411
GS-----------DSLMRQMRELEDRFASEASGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEGEESRI    Des Wt 343-416
GS-----------DSLMRQMRELEDRFASEASGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEG----    Des Wt 343-416

GSTCEIDALKGTNDSPMRQMRELEDRFASEASGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEG----    Des L345P 332-411
GS-----------DSLMRQMRELEDRFASEPSGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEGEESRI    Des A360P 343-416
GSTCEIDALKGTNDSLMRQMRELEDRFASEPSGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEG----    Des A360P 332-411
GS-----------DSLMRQMRELEDRFASEASGYQDNIARPEEEIRHLKDEMARHLREYQDLLNVKMALDVEIATYRKLLEGEESRI    Des L370P 343-416
GS-----------DSLMRQMRELEDRFASEASGYQDNIARLEEEIRHLKDEMARHPREYQDLLNVKMALDVEIATYRKLLEG----    Des L385P 343-411
GSTCEIDALKGTNDSLMRQMRELEDRFASEASGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKMALDVSIATYRKLLEG----    Des E401S 332-411
```

Figure 4.2.4.1 Desmin sequences shifts with residues highlighted in yellow representing the heptad repeat positions *a*. The residues in red are the mutations.

These proteins were expressing at levels comparable to the wild-type desmin fragment. The circular dichroism measurement of two shifted samples (Figure 4.2.4.2A), i.e. desmin wild-type fragments 332-411 and 343-411, showed that the α-helicity was more pronounced than of the original fragment sequence (residues 332-416). The geometry conservation is confirmed by gel-filtration since elution peaks are

almost proportional to the samples molecular weights (Figure 4.2.6B). Unfortunately none of these shifted wt and mutant fragments yield crystal.
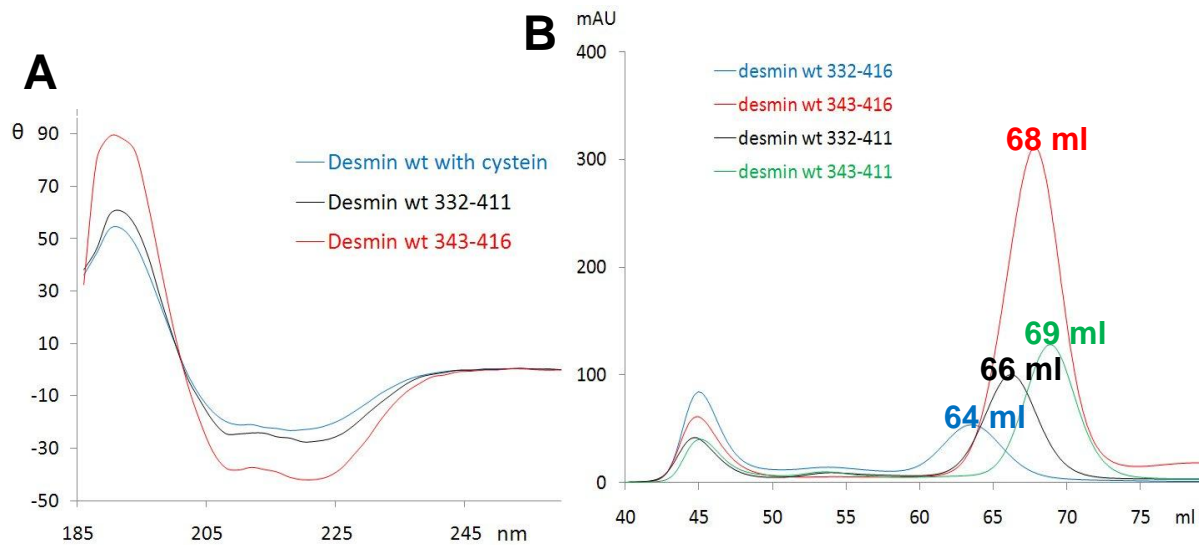
Figure 4.2.4.2 A) Circular dichroism spectra of the desmin wild-type protein fragment including cystein (residues 332-416, in blue) and of the desmin wild-type fragments residues 332-411 (black) and 343-416 (red). Minima at 208 and 222 nm, characteristic of α-helical samples, are more pronounced for the 'shifted' fragments. B) Gel-filtration profiles of desmin wt fragments. The elution volume peaks are almost proportional to the fragment masses, showing that the overall secondary structure of these fragments is conserved: desmin wt 332-416 MW = 20 kDa, desmin wt 343-416 MW = 18 kDa, desmin wt 332-411 MW = 19 kDa, desmin wt 343-411, MW = 17 kDa (dimer masses).

4.2.5 Point mutation in a lamin A context

We then decided to insert desmin-related mutations in the homologous *lam1* fragment (PDB entry 1x8y Strelkov *et al.*, 2004)[61]. We choose the *lam1* protein fragment since it crystallizes readily in several different conditions and has a high symmetry space group (P 6_5 2 2). For convenience, these lamin fragments are followed by the number which corresponds to the desmin mutation position. Anyway, one can rapidly calculate the lamin A residue number by subtracting 28 to the desmin residue number (Figure 4.2.5.1). The lamin 377 was produced after a mistake in primers design.

```
GSCQLSQLQKQLAAKEAKLRDLEDSLARERDTSRRLLAEKEREMAEMRARMQQQLDEYQELLDIKLALDMEIHAYRKLLEGEEERL    Lamin Wt
GSCQLSQLQKQLAAPEAKLRDLEDSLARERDTSRRLLAEKEREMAEMRARMQQQLDEYQELLDIKLALDMEIHAYRKLLEGEEERL    Lamin 345
GSCQLSQLQKQLAAKEAKLRDLEDSLAREPDTSRRLLAEKEREMAEMRARMQQQLDEYQELLDIKLALDMEIHAYRKLLEGEEERL    Lamin 360
GSCQLSQLQKQLAAKEAKLRDLEDSLARERDTSRRLLAEPEREMAEMRARMQQQLDEYQELLDIKLALDMEIHAYRKLLEGEEERL    Lamin 370 *
GSCQLSQLQKQLAAKEAKLRDLEDSLARERDTSRRLLAEKEREMAEPRARMQQQLDEYQELLDIKLALDMEIHAYRKLLEGEEERL    Lamin 377 *
GSCQLSQLQKQLAAKEAKLRDLEDSLARERDTSRRLLAEKEREMAEMRARMQQQPDEYQELLDIKLALDMEIHAYRKLLEGEEERL    Lamin 385
GSCQLSQLQKQLAAKEAKLRDLEDSLARERDTSRR-LAEKEREMAEMRARMQQQLDEYQELLDIKLALDMEIHAYRKLLEGEEERL    Lamin 366
GSCQLSQLQKQLAAKEAKLRDLEDSLARERDTSRRLLAEKEREMAEMRARMQQQLDEYQELLDIKLALDMSIHAYRKLLEGEEERL    Lamin 401
```

Figure 4.2.5.1 Desmin mutation inserted into the lamin A *lam1* sequence. The desmin mutations number were kept for convenience. The relationship is lamin A residue number = desmin residue number – 28.

As a control, the *lam1* fragment was reproduced and crystallized readily in many different crystallization conditions (Figure 4.2.5.2), indicating that our purification and crystallization procedures are correct.

100 µm
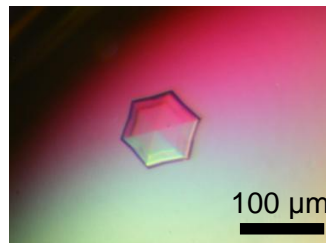
Figure 4.2.5.2 The lamin A (*lam1*) reproduced crystal. This crystal grew out of the initial screen without any further crystallization optimizations. The six-fold symmetry is evident from the crystal shape.

CD spectra of the *lam1* and lamin 401 fragments do not reveal particular features (Figure 4.2.5.3A). These lamin samples have an α-helical content similar that of homologous desmin samples.

The gel-filtration of lamin fragments, containing cysteins, has a comparable void volume to dimer peak ratio as observed for desmin fragments. An interesting feature of lamin fragments 345, 360, 370 and 385 is the presence of a third peak at ~78 ml elution volume which corresponds to monomers. Even, the lamin 370 is exclusively monomeric. Here gel-filtration (Figure 4.2.5.3B) supports analytical ultracentrifugation result showing that the lamin wt fragment is a dimer in solution (Strelkov, 2004)[61]. We also noticed by analytical centrifugation that desmin L345P and A360P fragments (Table 4.2.2.1) are of a mixture of dimers and monomers in solution. Again, gel-filtration shows the presence of monomeric peak for both the homologous lamin 345 and 360 fragments. However, a great care should be taken by comparing desmin and lamin fragments. For instance, no monomer peak is found for the lamin 401 which corresponding desmin E401S analytical ultracentrifugation also showed to be a mixture of dimer and monomer.
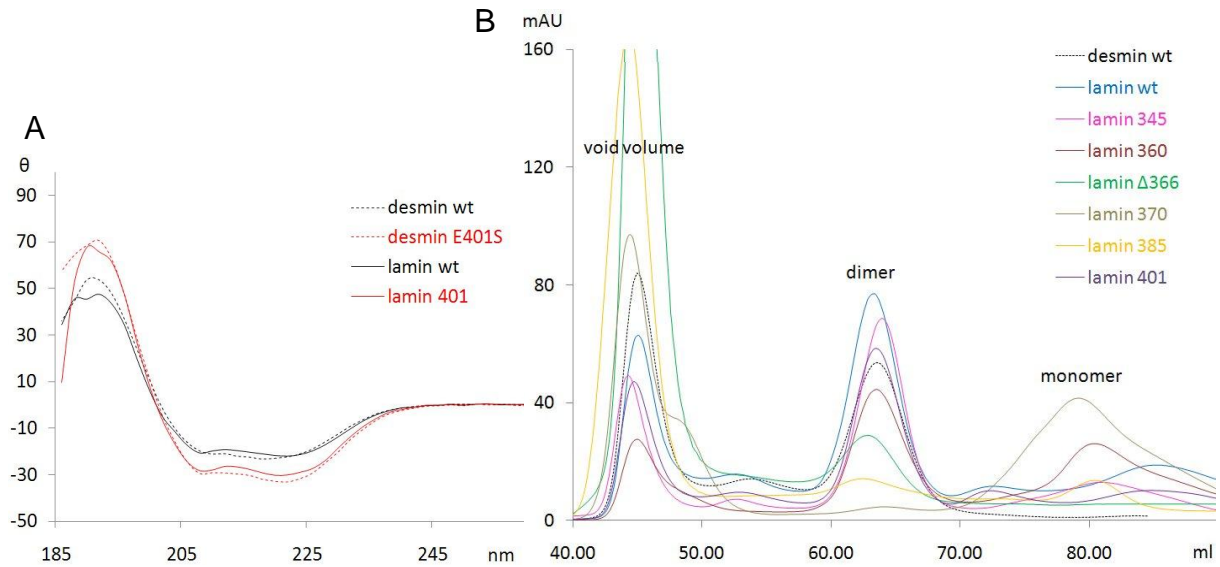
Figure 4.2.5.3 A) CD spectra of desmin and lamin wt and mutant 401. B) Gel-filtrations of lamin fragments. The void-volume peaks (45ml) size of lamins are comparable, if not bigger, to desmin. In contrast to desmin samples, the lamin fragments wt, 345, 360, 370 and 385 have an additional peak around 80ml corresponding to monomers. For both lamin 345 and 360, the presence of a monomeric peak can be correlated with analytical ultracentrifugation showing a mixture of monomers and dimers. However, the lamin 370 is exclusively monomeric whereas analytical ultracentrifugation profiles of its homologous desmin L370P showed only dimers.

Unfortunately, the crystallization of lamin fragments containing desmin mutations was unsuccessful.

4.3 Small angle X-ray scattering analysis of desmin wt and mutants

4.3.1 Buffer conditions and samples

All desmin samples were first dialyzed in tetramer buffer. After concentration, we added either NaCl or assembly buffer to our samples. The table 4.3.1.1 describes the different buffer conditions we used in this SAXS experiment.

Name	Composition	pH	I* [mM]	Name	Composition	pH	I* [mM]
Tetramer buffer (TB)	5mM Tris-HCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT	8.4	5	10X Assembly buffer (AB)	250mM Tris-HCl, 500mM NaCl	7.0	
TB + 15 mM NaCl	5mM Tris-HCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT, 15 mM NaCl	8.4	20	TB + 10mM** AB	10mM Tris-HCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT, 10 mM NaCl	8.24***	20
TB + 30 mM NaCl	5mM Tris-HCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT, 30 mM NaCl	8.4	35	TB + 20mM** AB	15mM Tris-HCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT, 20 mM NaCl	8.2***	35
TB + 75 mM NaCl	5mM Tris-HCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT, 75 mM NaCl	8.4	80	TB + 50mM** AB	30mM Tris-HCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT, 50 mM NaCl	8.0***	80

* estimate of the ionic strength based on the molarity of Tris buffer and added salt

** NaCl molarity resulting after addition of an appropriate volume of 10X AB

*** pH as measured after addition of an appropriate volume of 10X AB

Table 4.3.1.1 Buffers used for SAXS measurements. The ionic strength has been calculated using the following formula: I [mol l^{-1}] = 0.5 x Σ (C_i z_i^2) with C_i being the molar concentration in mol x l^{-1} and z_i the ion charge. The Tris charge has been chosen to be +1.

The table 4.3.1.2 (next page) shows all collected samples, their buffer conditions and concentrations as measured by OD at 280 nm.

Table groups: **Tetramer** | **I = 20 mM** (15mM salt, 10mM AB) | **I = 35 mM** (30mM salt, 20mM AB) | **I = 80 mM** (75mM salt, 50mM AB)

Sample	Tetramer	g/l	15mM salt	g/l	10mM AB	g/l	30mM salt	g/l	20mM AB	g/l	75mM salt	g/l	50mM AB	g/l
Vimentin Sokolova	1H	3.5	2H**	2.3	-	-	-	-	-	-	-	-	3H	?
Vimentin Hamburg	4H	1.7	-	1.0	-	-	-	-	5H	1.7	-	-	-	-
Desmin Lydia 28.5	6H 7S 8S 9S 10H	0.5 1 2.5 1.7	11H**	-	-	-	-	-	-	-	12H	1.0	-	-
Desmin Lydia 2.6	13H	1.0	-	-	14H	1.0	gel	-	15H	1.0	-	-	-	-
L345P Lydia	16H	1.5	17H 18H	0.6 1.5	gel	1.5	-	1.5	gel	0.6 1.5	gel	0.6 1.5	-	-
L345P Harald	19S	1.4	-	-	gel	1.4	-	-	-	-	-	-	-	-
A360P Lydia	20H	1.0	21H	1.0	24H	1.0	22H	1.0	25H	1.0	23H	1.0	-	-
L370P Lydia	26S	0.6	-	-	gel	0.9	gel	0.7	gel	0.7	-	-	-	-
L370P Harald	27S	1.6	-	-	-	-	-	-	-	-	-	-	28S #	1.6
L385P Lydia	-	-	gel	0.7	gel	0.7	-	-	gel	0.7	-	-	-	-
L385P Harald	29S	1.4	-	-	-	-	-	-	gel	0.8	-	-	gel	1.4
N342D Harald	30H 31S	1.0 2.1	gel	1.0	-	-	-	-	gel	1.0	-	-	gel	2.1
A337P Harald	32S	1.2	-	-	gel	0.5	-	-	gel	0.5	-	-	33S *	1.2
R406W Lydia	34S	2.7	-	-	36H	1.0	35H	1.0	37H	1.0	-	-	gel	1.0
R406W Harald	38S	1.1	-	-	39S	0.4	-	-	40S #	0.4	-	-	gel	1.8
Q389P Harald	41H $	0.2	-	-	-	-	-	-	-	-	-	-	-	-
I451M Harald	42S	0.5	-	-	43H	0.4	-	-	44H	0.4	-	-	gel	0.5
E245D Harald	45S	2.7	46H %	1.2	-	-	-	-	47H	1.2	-	-	gel	2.7
E413K Harald	48S	2.1	49H	0.7	51H	0.7	50H	0.7	52H	0.7	-	-	gel	0.7
R350P Harald	53S	0.9	54H #	0.8	-	-	gel	0.8	-	-	-	-	gel	0.9
A357P Harald	55S	1.8	-	-	gel	1.1	-	-	56H #	1.1	-	-	gel	1.8
D399Y Harald	57S	0.6	60H	0.5	58H #	0.2	-	-	gel	0.2	-	-	gel	0.6
Δ359-361 Harald	59S	2.2	-	-	-	-	gel	0.5	gel	0.5	-	-	gel	2.2
A213V Harald	61S	2.4	-	-	-	-	-	-	gel	0.7	-	-	gel	2.4
Δ366 Harald	62H 63S	0.7 2.1	64H	0.7	gel	0.7	65H #	0.7	-	-	gel	0.7	gel	2.1

\# two phases: the liquid one was measured
* X-Ray induced polymerization
% very visquous
$ gel formation in tetramer buffer during concentration
** 20mM NaCl addition rather than 15mM

Table 4.3.1.2 SAXS measured samples. For each sample, the sample number is given in bold, followed by a letter indicating the beamline used (H for Hamburg, S for Soleil) and the concentration of the sample. Refer to the table 4.3.1.1 for a complete description of buffers.

4.3.2 Desmin wt and mutants viscosity

We added NaCl and assembly buffer in aliquots of our dialysed samples in tetramer and let them polymerize for at least one hour. Before data acquisition these samples were centrifuged. Some samples formed two visible phases. The viscosity of samples forming a single phase or of the supernatant, when phase separation occurred, was probed by pipetting. This *qualitative* indication of desmin samples viscosity already demonstrates that several desmin mutants behave differently that of desmin wt under the same buffer conditions (see Table 4.3.2.1).

	TB	+15 mM NaCl	+20 mM AB	+30 mM NaCl	+50 mM AB	+75 mM NaCl
A360P	green	green	green	green	green	green
Vimentin	green	green	green	green	green	green
Desmin	green	green	green	green	green	green
R406W	green	green	green	green	white	red
E413K	green	green	green	green	white	red
I451M	green	green	green	green	white	red
E245D	green	green	green	orange	orange	red
dN366	orange	orange	orange	orange	orange	red
dE359-S361	orange	orange	white	white	red	red
D399Y	orange	orange	white	white	red	red
R350P	green	orange	white	red	red	red
L345P	green	orange	white	red	red	red
A213V	orange	white	white	red	red	red
L385P	green	red	red	red	red	red
L370P	green	red	red	red	red	red
A357P	orange	red	red	red	red	red
A337P	orange	red	red	red	red	red
N342D	orange	red	red	red	red	red
Q389P	red	red	red	red	red	red

Table 4.3.2.1 Desmin wt and mutants qualitative solubility assay in tetramer buffer (TB) or with TB supplied with NaCl or assembly buffer (AB). The green cases are samples which stayed liquid. In orange, samples only formed a single but very viscous phase after centrifugation. In red the samples did form a gel. In that case, only the upper phase after centrifugation, when fluid enough to be pipetted, was measured. No measurements were done for samples in white cases.

Samples solubility in tetramer buffer

Samples in tetramer buffer can be classified into three solubility classes. First, the desmin wt and mutants E245D, L345P, R350P, A360P, L370P, L385P, R406W, E413K and I451M remain soluble after dialysis and at high concentrations (>1mg/ml).

52

As for comparison, the vimentin sample also belonged to this class. Second, the samples A213V, A337P, N342D, A357P, ΔE359-S361, ΔN366 and D399Y were still soluble after dialysis but tend to form gel at concentrations higher than 1 mg/ml. Finally, the mutant Q389P already formed a gel into the dialysis bag at a concentration of 0.2 mg/ml. Consequently, all samples in tetramer buffer except Q389P were sufficiently fluid to flow into the small 0.1 mm capillaries of the SAXS machine (cf. Figure 3.7.3.1).

Samples solubility with salt or assembly buffer addition

The viscosity of desmin samples supplied with of 15, 30 and 75 mM NaCl was found to be equivalent to the addition of 10, 20 and 50 mM assembly buffer respectively. Our ionic strength calculation shows that the Tris buffer present in assembly buffer respectively compensates the excess of NaCl.

Here also, mutants can be classified in three groups. First, the mutants E245D, A360P, R406W, E413K and I451M stayed as fluid as desmin and vimentin wt up to 30 mM NaCl addition. Second, the mutants L345P, R350P, ΔE359-S361, ΔN366 and D399Y had a single viscous phase up to 30 mM NaCl addition. However, with 30 mM NaCl addition, these mutants turned into gel. Third, the mutants A337P, A357P, N342D, L370P and L385P already formed a gel with 15 mM NaCl addition. Finally, the mutant A213V cannot be classified since its viscosity was not probed with 15 mM NaCl addition.

4.3.3 Desmin wt assembly complexes as monitored by SAXS

Before starting to analyze our desmin mutants against desmin wt, we will first show how the desmin wt SAXS intensity curves and distance distribution functions vary in tetramer buffer and increased ionic strength buffers.

Several measurements of desmin wt were done in tetramer buffer. According to the R-factor tests done with intensity curves between the angular Q range 1 to 2.5 [nm^{-1}], only one of our five desmin wt curve moderately deviates (Rf > 5%) from other measurements (Table 4.3.3.1). Interestingly, this desmin wt in tetramer buffer 'outsider' is not from a different purification batch and thus such deviation might be

related to few impurities or air bubbles within the sample or SAXS machinery. Therefore, I decided to skip this sample since it is more likely to compromise further analysis and comparison with desmin mutants.

conc. [mg/ml]	0.5	1.0	2.5	5.0	1.7
0.5	-				
1.0	4.9%	-			
2.5	4.4%	5.0%	-		
5.0	5.0%	6.2%	1.8%	-	
1.7	4.2%	5.1%	2.3%	2.8%	-
1.0*	4.2%	5.5%	1.9%	2.3%	2.2%

Table 4.3.3.1 Desmin wt at different concentrations in tetramer buffer and R-factor differences as calculated between the angular Q ranges from 1 to 2.5 nm^{-1}. The differences below 5% are in green, whereas differences of >5% are in yellow. The dataset marked with a * at 1 mg/ml correspond to another purification preparation.

The fact we measured desmin wt in tetramer buffer with different concentrations (0.5 mg/ml to 5 mg/ml) (Figure 4.3.2.1) permits us to determine if desmin intensity curves deviates at low angles. Such deviations are either due to aggregation or close proximity between molecules present in the sample. Indeed, at low angles the intensity curve for the molecules attracting themselves will go up whereas intensity curves for the molecules repulsing each others will go down. These effects are typically seen at the lowest scattering angles (Q below 0.5 nm^{-1}).

All our R-factor comparisons were done in a Q scattering angle range from 1 to 2.5 nm^{-1} independent of such effects.
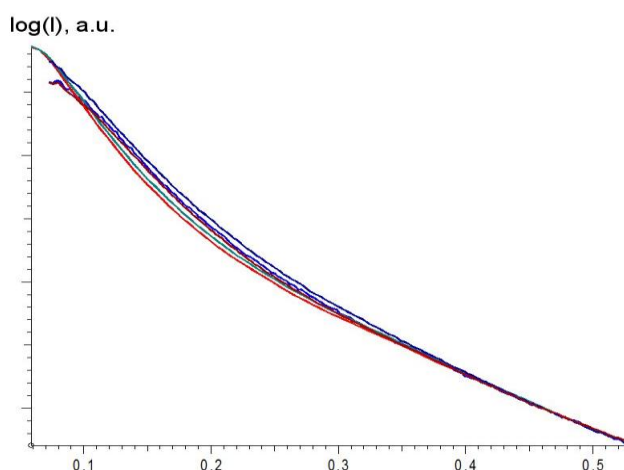
Figure 4.3.2.1 Scattering curves of all desmin wt samples at the lowest angles. No aggregation is observed.

The SAXS curves of desmin wt in tetramer buffer and with increased ionic strength (Figure 4.3.2.2A) reveal that the cross-section diameter (Dmax) of desmin wt in tetramer buffer is of ~12 nm and increased to ~20 nm with increased ionic strength buffers (Figure 4.3.2.2B). The distance distribution function in the tetramer buffer shows only one maximum at ~1.5-2 nm whereas upon salt addition a second smaller peak at ~7 nm emerges. While there is a distinct difference between the desmin wt curve in tetramer buffer and after 10 mM AB addition, any further addition of AB or salt does not lead to significant changes in the SAXS signal and, correspondingly, the distance distribution function curve. It should be concluded that, with 10-15 mM salt addition, desmin wt already reaches its maximal cross-section diameter. This is an unexpected result compared to the vimentin SAXS experiment (Sokolova *et al.*)[62]. First, the vimentin curves measured in tetramer buffer showed a diameter of ~6 nm (see Sokolova *et al.*[62], table 1 and figure 1A and B) that is half the diameter we observed for desmin. Second, the desmin cross-section diameter (~12 nm) in tetramer buffer is similar to vimentin supplied with 20 mM NaCl. Third, the maximal cross-section diameter of desmin (~20 nm) is somewhat larger than vimentin (Dmax ~16 nm). Therefore, we can conclude that desmin form bigger species than vimentin in tetramer buffer and requires less salt to reach its maximal cross-section diameter. However, a detailed direct comparison of wt desmin and vimentin needs to be carried out in future SAXS experiments to verify these conclusions.
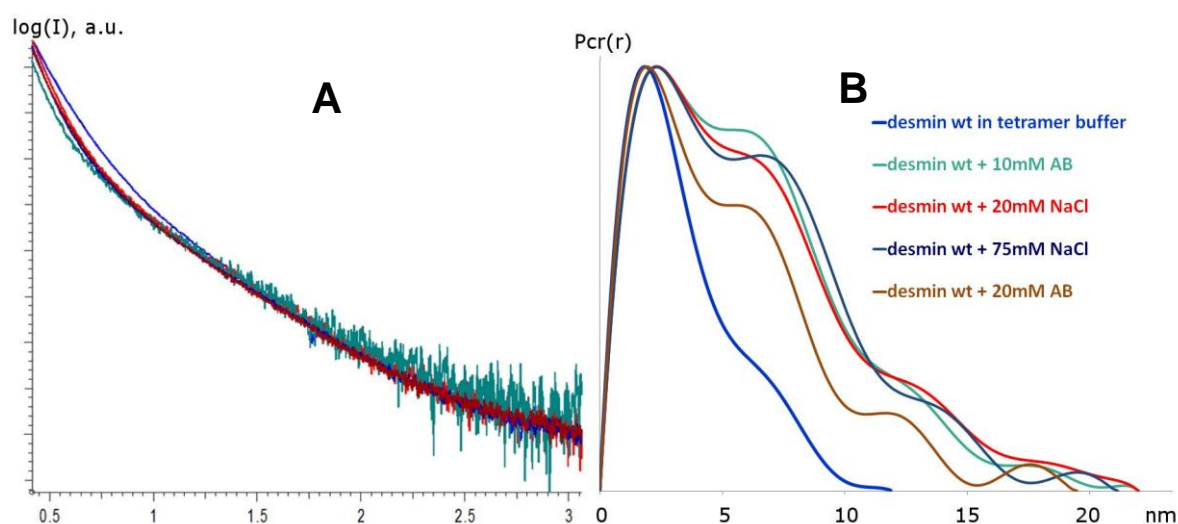
Figure 4.3.3.2 Desmin wt in tetramer buffer is show in blue, desmin wt in 10mM AB in green, with 20 mM NaCl in red, with 75 mM NaCl in dark blue and with 20 mM AB in brown. The intensity curves (A) show a neat slope difference between desmin wt in tetramer buffer (blue)

and desmin wt supplied with salts or assembly buffer. The distance distribution functions (B) indicate that the cross-section diameter of desmin wt in tetramer buffer is of ~12 nm whereas it increases to 19-22 nm for desmin wt supplied with salts or assembly buffer.

4.3.4 Desmin mutants in tetramer buffer

The analysis of desmin wt serves as a template to determine if desmin mutants resemble or not to desmin wt complexes. For this, a complete R-factor analysis was done to sort our samples in different classes. The R-factor difference between mutants and the averaged five desmin wt samples is shown in Table 4.3.4.1. We consider here mutants to be similar if their R-factors are < 5%, moderately different if their R-factor is comprised between 5% and 10% and different for Rf > 10%. The 5% difference in R-factor corresponds to the observed difference *within* desmin wt datasets in tetramer buffer.

sample	%Rf vs. desWt	Dmax [nm]
Desmin wt	up to 5%	12.0
L345P	3.4%	11.8
R406W	3.7%	9.5
R350P	3.9%	11.0
A360P	3.9%	13.0
ΔN366	4.3%	12.5
A337P	4.7%	12.3
N342D	4.9%	12.1
E413K	6.5%	12.0
A357P	6.8%	12.0
ΔE359-S361	7.1%	11.0
L385P	8.2%	12.0
E245D	8.4%	12.0
A213V	8.5%	11.0
D399Y	11.5%	11.0
I451M	11.9%	13.0
L370P	17.3%	22.0

Table 4.3.4.1 Desmin mutants in tetramer buffer sorted according to their averaged R-factor difference with desmin wt in tetramer buffer. Desmin mutant samples with a R-factor difference of <5%, in green, are considered to be equivalent to desmin wt. Desmin mutants having a R-factor comprised between 5 and 12%, in yellow, are considered as moderately different. Desmin mutants with a R-factor difference of more than 12%, in red, strongly deviate of desmin wt. Except L370P, all these mutants have a cross-section diameter close to desmin wt (average = 11.7 nm, excluding wt and L370P).

Except for L370P, all mutants in tetramer buffer were found to have the same cross-section diameter (Dmax of ~10 nm) as desmin wt. The shape of the distance distribution function has been investigated in more detail. For the first group of

mutants (i.e. L345P, R406W, R350P, A360P, ΔN366, A337P and N342D), the distance distribution functions closely resemble the one of desmin wt and have only *one* maximum at ~2 nm (Figure 4.3.4.1A). However the distance distribution function of mutants forming the second group (i.e. E413K, A357P, ΔE359-S361, L385P, E245D, A213V, D399Y and I451M) exhibits *two* maxima at ~2 nm and ~3.7 nm (Figure 4.3.4.1B). This peak duplicity indicates either a mixture of two different species or a less compact core structure within the filament precursor core than desmin wt. Finally, the mutant L370P strongly deviates and forms our last group (Figure 4.3.4.1C). Interestingly, the mutant L370P reaches the same cross-section diameter as desmin wt in higher ionic strength buffers (~22 nm).
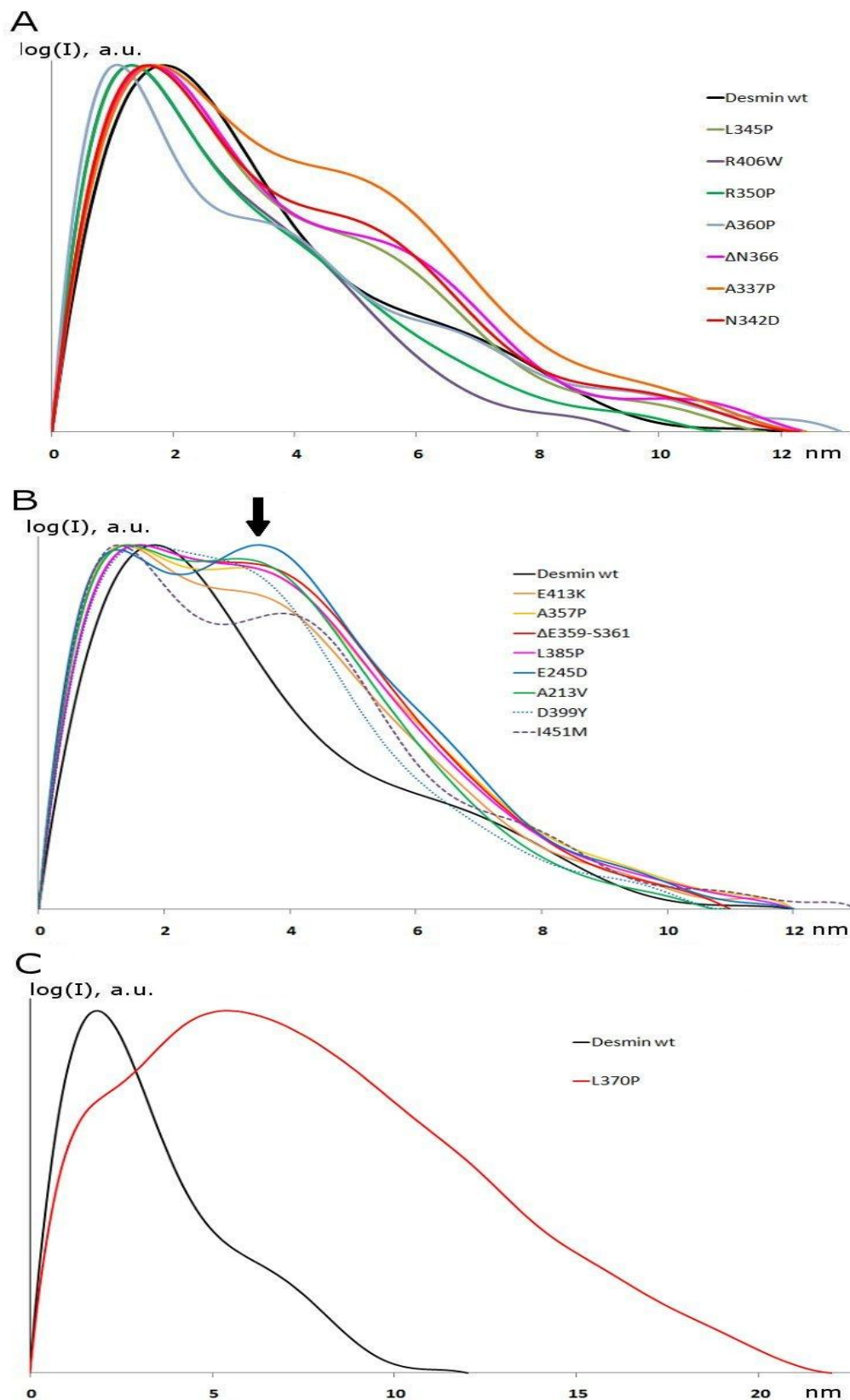
Figure 4.3.4.1 Distance distribution functions of desmin mutants and desmin wt (black curves). (A) The first group of mutants share similar distance distributions than of the wt. These curves have only one maximum at ~2 nm. (B) The second group of mutants have distance distribution functions exhibiting a second maxima at ~4 nm. (C) The mutant L370P has a cross-section diameter of ~22 nm which is twice the size than of desmin wt in tetramer buffer.

4.3.5 Desmin mutants after salt or assembly buffer addition

We have seen that almost all our mutants in tetramer buffer, except L370P, look similar to desmin wt.

However, as discussed above (section 4.3.2), only some of the studied mutants did not reveal gel formation upon addition of 15 mM NaCl. These are the mutants E245D, A360P, R406W, E413K and I451M.

The SAXS data recorded after salt or assembly buffer addition reveal (Figure 4.3.5.1) that all of the above-mentioned mutants but R406W assemble in roughly the same way as the wt. For all these mutants in increased ionic strength buffers, a second maximum at ~7 nm appears, whereas the cross-section diameter (Dmax) increases to ~15 nm. Interestingly, mutants E245D and I451M show disappearance of the distance distribution peak at ~2 nm which is present in all other curves.

The mutant R406W has revealed a slower assembly than the wt (Figure 4.3.5.1), since the 7 nm peak in the distance distribution function was much weaker even with 30 mM NaCl addition.

As discussed in section 4.3.2, for many mutants, addition of salt or assembly buffer resulted in gel formation. Often (see Table 4.3.2.1) the sample could be separated into two phases: the lower gel phase and the upper soluble phase. In the case of mutants L345P, ΔS359-E361, ΔN366, R350P, A357P and D399Y the upper phase could still be measured (Table 4.3.1.2 and Figure 4.3.5.1). However it should be stressed that the resulting SAXS curve does not represent the complete sample but only the still soluble part. As seen in figure 4.3.5.1, the distance distribution functions of all these samples show profiles that are only somewhat 'thicker' than that for the wt desmin in tetramer buffer, with a single maximum at ~2 nm. The second peak at ~6 nm is much weaker for most of these mutants.

It should be therefore concluded that the upper fluid phase of these samples contains a predominance of an oligomeric species similar to that present in the tetramer buffer (most likely a tetramer, see the discussion section below).

Finally, the mutants A213V, A337P, N342D, L370P and L385P could not be measured since they already formed a gel in the tetramer buffer supplied with 15 mM NaCl. For two of these mutants, A337P and N342D, this confirms a previous EM and viscometry study showing that they tend to stick and aggregate (Bar *et al.*, 2005)[45].
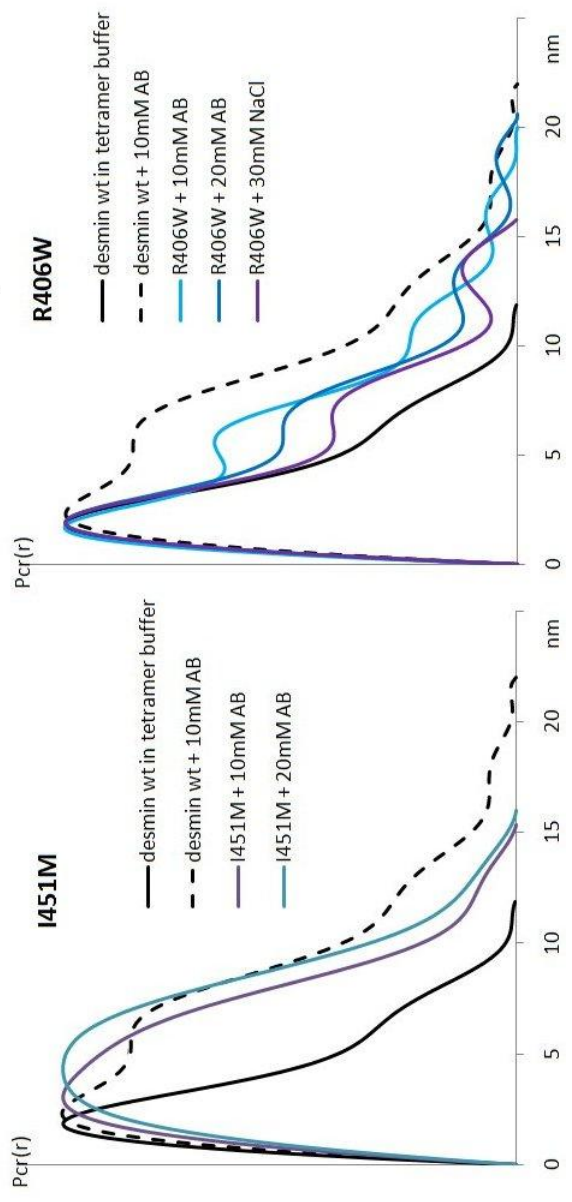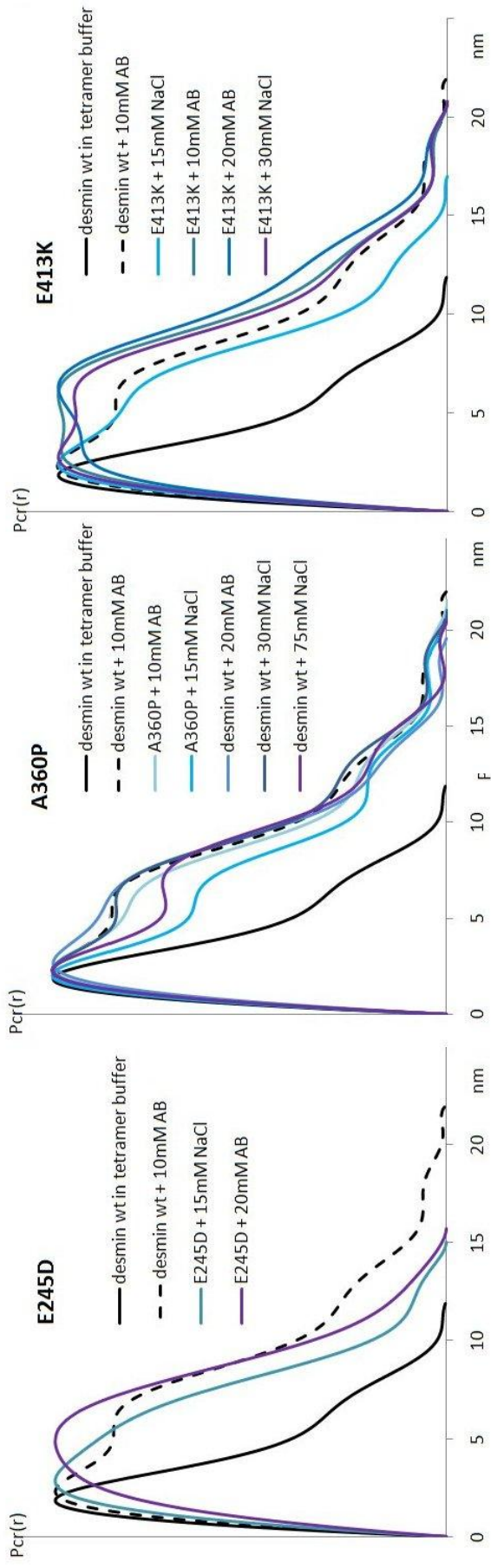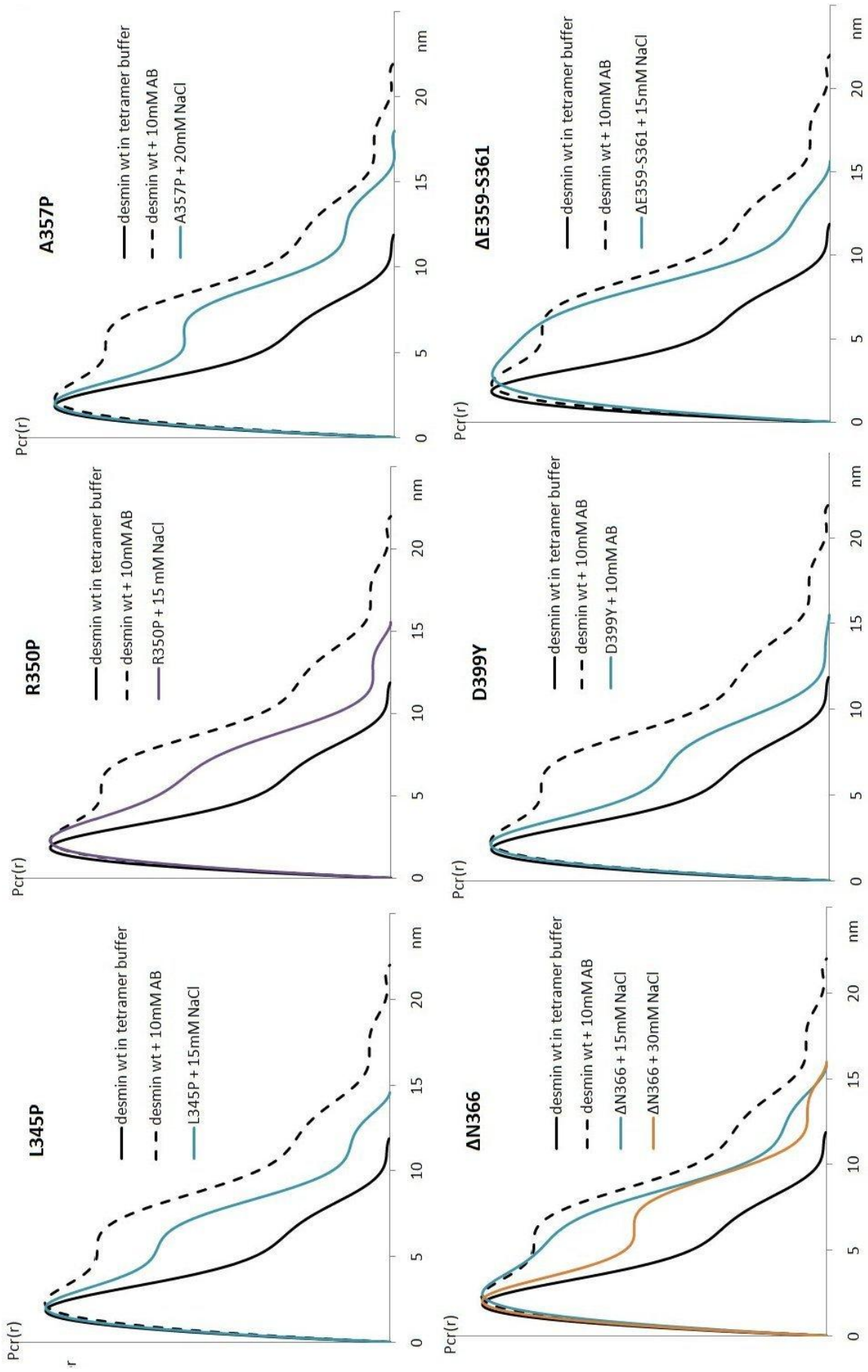
Figure 4.3.5.1 Distance distribution functions of desmin mutants supplied with salts or assembly buffer.

4.3.6 Conclusions of the SAXS study

By taking into account all obtained data, we can summarize the conclusions in a table (Table 4.3.6.1). We would classify the studied mutants into two large groups.

The first group includes the mutants that show the overall assembly behavior in tetramer buffer and upon ionic strength increase similar to the wt. Most of these mutants were also studied by Bär *et al.*[45,59]. In particular, most mutants revealed a sedimentation coefficient in tetramer buffer which is close to that of the wt (5.2S). Correspondingly, these mutants are assumed to be present as tetramers in the tetramer buffer. Moreover, using EM Bär *et al.* observed normal filaments assembly after addition of an equal amount of assembly buffer (45 mM Tris∘HCl pH 7.0, 100 mM NaCl) to the tetramer buffer. Correspondingly, Bär *et al.* have classified these mutants as belonging to the 'Class I'.

In particular, the mutants A360P and E245D were found by EM to belong to the class I. Our SAXS measurement confirms that these two mutants assemble as the wt. Further, our SAXS data show that the mutant R406W assemble less than the wt (Figure 4.3.5.1). This correlates to the fact that it shows abnormalities by EM and was classified in the assembly class II (disturbed longitudinal annealing and radial compaction). Moreover, the mutant R406W distance distribution function shows several maxima (~2, 6, 11, 14... nm) that might be characteristic for a less dense core due to a radial compaction defect. Further, its cross-section diameter (~19 nm) stays close to the desmin wt (~22 nm). However, we cannot conclude if ULFs anneal or not since our SAXS datasets resolution is below the ULF size. Finally, it would be interesting to verify by EM if the mutants I451M and E413K, which were not studied by Bär *et al.* also assemble as desmin wt.

The second big group that we have observed includes mutants that assemble more readily even after minimal increase in the ionic strength. Many, but not all of these mutants belonging to this group were showing signs of gel formation already in the tetramer buffer at concentrations above 1 mg/ml. The gel formation was distinct as soon as some salt or assembly buffer was added. In some cases, the resulting

sample could be separated into the gel fraction and the supernatant. We could measure the supernatant by SAXS, but it should be stressed that the obtained curves did not represent the complete sample.

For instance, the supernatant of the mutant L345P and R350P contains species which have a similar cross-section diameter (~ 15 nm) that of desmin wt in tetramer buffer (~ 12 nm). These mutants were shown by Bär *et al.* to form ULFs which rapidly degrade into small fibrous and roundish ~50 nm aggregates (class IV). In our case, these aggregates were most likely removed by the centrifugal step and our SAXS curves show that these two mutants do not form filaments. Further, the distance distribution functions of these two mutants confirm that they only partly assemble, since they have only a weak second peak at ~7 nm compared to the desmin wt.

The mutants A337P, N342D and A357P of our second group of mutants were shown by Bär *et al.* to have an enhanced adhesiveness and aggregate formation (class III). Indeed, all these mutants formed a gel already in tetramer buffer. With salt addition, the solution turns so viscous for mutants A337P and N342D that no SAXS data could be measured. The supernatant of the mutant A357P was measured and showed a distance distribution function similar than that of the wt. This suggests that this mutant still form filaments as observed by Bär *et al.*

Finally, the mutant D399Y was shown by EM to form seemingly normal filaments. In our case, this mutant turned into a gel in both tetramer buffer and higher ionic strength buffers. Further, the SAXS curve of the supernatant of this mutant in tetramer buffer and 10 mM assembly buffer has a very weak second peak at ~7 nm, suggesting to us that it does only partly assemble. However, we could not exclude that complete but sticky filaments were not removed by the centrifugal step before SAXS acquisition.

Interestingly, the sedimentation coefficients in tetramer buffer of mutants D399Y (~12S) and Q389P (~10S) were shown to be higher than desmin wt and most mutants (5.2S). The mutant Q389P has precipitated in our hands already in tetramer buffer, suggesting that it rapidly form higher species or aggregates.

Lastly, the mutant L370P was soluble in tetramer buffer but formed abnormally large oligomers (Dmax = 22 nm). This observation contradicts the data by Bär *et al.* that

measured normal sedimentation coefficients in tetramer buffer. At the same time, the further assembly that they have observed was also not normal; this filament was shown to have a conserved ULF formation that break down into small aggregates (class IV).

Table 4.3.6.1

Group I: mutants similar to desmin wt

	Tetramer buffer (TB)		+ 10 mM AB or 15 mM NaCl		AUC sed. coef in TB	Class Bär *et al.*
	Gel form[1]	SAXS Rf [2] in %	Gel form[1]	SAXS Rf [3] in %		
WT	no	up to 5	no	-	5.2	I
E245D	no	8.4	no	- / 3.5	5.7	I
A360P	no	3.9	no	3.4 / 4.4	5.2	I
E413K	no	6.5	no	5.8 / 3.4	n/a	n/a
I451M	no	11.9	no	6.8 / -	n/a	n/a
R406W*	no	3.7	no	4.9 / -	5.4	II

[1] as in Table 4.3.2.1

[2] R-factor versus desmin wt in tetramer buffer as in Table 4.3.4.1

[3] R-factor versus desmin wt in the same condition (1st number with 10 mM AB, 2nd with 15 mM NaCl)

* R406W revealed decreased assembly upon buffer addition

Group II: mutants similar to desmin wt in tetramer buffer but increased assembly upon ionic strength increase

	Tetramer buffer (TB)		+ 10 mM AB or 15 mM NaCl		AUC sed. coef in TB	Class Bär *et al.*
	Gel form[1]	SAXS Rf[2] in %	Gel form[1]	SAXS Rf[3] in %		
L345P	no	3.4	yes	- / 8.1	4.9	IV
R350P	no	3.9	yes	- / 9.7	4.8	IV
L385P	no	8.2	yes	n/a	5.3	II
A213V	yes	8.5	yes	n/a	5.0	I
A337P	yes	4.7	yes	n/a	5.3	III
N342D	yes	4.9	yes	n/a	5.0	III
A357P[#]	yes	6.8	yes	18.0	5.2	III
ΔE359-S361	yes	7.1	Yes	- / 5.8	n/a	n/a
ΔN366	yes	4.3	Yes	- / 4.6	n/a	n/a
D399Y	yes	11.5	yes	17.9 / -	11.9/ 12.7	I

Outliers: L370P forms large oligomers already in tetramer buffer and Q389P precipitates

	Tetramer buffer (TB)		+ 10 mM AB or 15 mM NaCl		AUC sed. coef in TB	Class Bär *et al*
	Gel form[1]	SAXS Rf[2] in %	Gel form[1]	SAXS Rf[3] in %		
L370P	no	17.3	yes	n/a	5.2	IV
Q389P	precip	n/a	n/a	n/a	9.5/10.6	I

[1] as in Table 4.3.2.1

[2] R-factor versus desmin wt in tetramer buffer as in Table 4.3.4.1

[3] R-factor versus desmin wt in the same condition (1st number with 10 mM AB, 2nd with 15 mM NaCl

[#] comparison in tetramer buffer + 20 mM assembly buffer

5   Discussion

5.1   Atomic structure of vimentin coil 2 domain

5.1.1 Hendecad-based parallel α-helical bundle near the N-term of coil 2

The crystal structure described here represents the regions of the vimentin molecule formerly designated as coil 2A, linker L2 and the beginning of coil 2B. Importantly, we show that this region forms a single, contiguous α-helix. This is in agreement with the secondary structure predictions made for the amino-acid sequence, as noted earlier (Hess *et al.*, 2006; Parry, 2006)[42,63]. Our findings call for a revision of the widely accepted separation of coil 2 into segments 2A and 2B by the proposed non-helical linker L2. We show that the only substructure that could be found within coil 2 is linked to the differences in the hydrophobic repeat period and therefore the degree of supercoiling. The N-terminal part starting with Pro263 and ending with Ala302 is not twisted and should therefore be designated as a parallel α-helical bundle rather than a coiled coil. The rest of coil 2 is a standard coiled coil with a fairly regular left-handed geometry, with the exception of a short unwound region near the stutter at position 351. Along with the ubiquitous heptad repeat pattern, the hendecad periodicity in the distribution of hydrophobic side chains has also been recognized as a possible α-helical oligomerization motif (Peters *et al.*, 1996)[58]. While theoretical calculations for such a hendecad periodicity originally suggested a right-handed coiled-coil structure (Peters *et al.*, 1996)[58], the few crystal structures that have become available recently all indicate essentially a parallel geometry. For instance, analysis of the tetrabrachion RHCC structure (Stetefeld *et al.*, 2000)[26] using the program Twister yields an overall coiled-coil pitch of ~4200 Å for the part featuring hendecad repeats (residues 19–52), meaning that the supercoiling is indeed negligible.

Unexpectedly, the crystal structure of the D3 fragment did not reveal an in-register dimer for its entire length, an arrangement that would be expected for this region within the native full-length IF dimer. This is likely due to the differences in the properties of the isolated relatively short fragment and the full-length protein. Indeed, we have shown that the D3 fragment stays monomeric in solution, just like the two earlier vimentin fragments including the L2 region, namely 2AB (residues 225–335) and cys1 (residues 254–327) (Strelkov et al., 2001)[21]. This finding nicely correlates with the analysis of the hydrophobic core of the 'reconstructed' native dimer which

suggests that the coiled coil in this region is only marginally stable. A similar property was previously observed for another relatively short vimentin fragment corresponding to the 1A segment, which forms a monomeric α-helix in solution (Strelkov *et al.*, 2002)[9]. However, after introducing a single stabilizing Y117L mutation into this fragment, a stable dimeric coiled coil is observed (Meier et al., 2009)[23].

Evidence for a continuous α-helical structure near the N-terminus of vimentin coil 2 has previously been obtained from the SDSL–EPR experiments (Hess *et al.*, 2006)[63]. These studies, performed with assembled vimentin filaments, suggest a highly ordered structure for residues 281–304. This stretch is situated within the α-helical bundle according to our crystal data. Furthermore, EPR spectra point to the residues of the two parallel chains that are positioned especially close to each other (as revealed by a high d_1/d ratio; see Table 1 in Hess *et al.*, 2006) [63]. Notably, the pattern of these residues exhibits a distinct hendecad period, even though the authors do not discuss this. Moreover, the highest d_1/d ratios are observed for residues in core positions, in perfect agreement with our assignment (Figure 4.1.3.1B). Finally, Hess et al. (2006) suggest that the regular left-handed coiled coil starts with the residue 302, which again fully agrees with our crystallographic data.

5.1.2 Molecular structure and properties of IF coil 2

There is some variability in the length of coil 1 across different IF types, which is caused, in particular, by a highly variable linker L1 in different IFs and a longer 1B segment in nuclear lamins (Strelkov *et al.*, 2003)[28]. However, coil 2 is generally more conserved in sequence across different IFs than coil 1, and its length (which is the total length of segments formerly known as coil 2A, linker L2 and coil 2B) is absolutely conserved (Parry and Steinert, 1999)[64]. Sequence alignment of coil 2 of vimentin including residues 264–405 (Figure 5.1.2.1A) with the same region of other IF types points to a considerable sequence conservation especially near its N-terminus i.e. within the hendecad repeat region (as analyzed in detail by Parry (2006)[42] and within the C-terminal region known as the IF consensus sequence. The overall sequence similarity between coil 2 of vimentin and every other IF type is indeed considerable, amounting to 80% for NFL protein and to about 60% for both keratin types and nuclear lamins (Figure 5.1.2.1B).

Correspondingly, the parallel bundle assembly exemplified by our vimentin D3 fragment is clearly a conserved structural feature of all IF proteins. While the initiation of the α-helical structure of vimentin coil 2 is likely to be facilitated by the N-terminal capping by the Pro263 residue, most other IF types have a different amino acid in this position (Figure 5.1.2.1A). Importantly, our D3 dimer structure (Figure 4.1.4.5) can be fused with the previously determined structure of the cys2 fragment (Strelkov *et al.*, 2002)[9], since both fragments include residues 328–335. Upon superimposing the two structures by this region, the root mean-square deviation of their Cα positions is 1.3 Å indicating a good match. As a result, we can obtain a complete atomic model of vimentin coil 2 (Figure 5.1.2.1C). The length of the 141-residue long coil 2 as measured for this model is 21 nm. Notably, this value that should be conserved for all IF types exactly matches the ~21-nm longitudinal beading periodicity that was consistently found in the EM micrographs of various assembled IFs after glycerol spraying and rotary metal shadowing (Milam and Erickson, 1982; Heins *et al.*, 1993)[65,66]. This is an interesting correlation, although the observed periodicity in EM pictures is likely to depend on a number of additional parameters, such as the overlap length in staggered interdimer contacts (i.e. A11 and A22 modes) and positions of terminal domains within the mature IF architecture (Parry and Steinert, 1999)[64].

Last but not least, the observation of a contiguous α-helical structure for the entire length of coil 2 has an important implication with regard to the mechanical properties of the IF dimer. In particular, coil 2 should be considered as a single comparably rigid 'rod-like' structure based on two intertwined α-helices. Indeed, the persistence length of a regular two-stranded coiled coil has been estimated as 25 ± 15 nm (Schwaiger *et al.,* 2002)[67] which is somewhat longer than the contour length of coil 2. This stiffness of coil 2 is rather opposite to the properties of the preceding linker L12 and the downstream tail domain which are expected to be quite flexible (Parry, 2006)[42].

A

```
           -ghijkabcdefghijkabcdefghijkabcdefghijkabcdefgabcdefgabcdefgabcdefgabcdefg
Vim   263  PDLTAALRDVRQQYESVAAKNLQEAEEWYKSKFADLSEAANRNNDALRQAKQESTEYRRQVQSLTCEVDALKGT  336
NFL   252  PDLSAALKDIRAQYEKLAAKNMQNAEEWFKSRFTVLTESAAKNTDAVRAAKDEVSESRRLLKAKTLEIEACRGM  325

K8    254  LDMDSIIAEVKAQYEDIANRSRAEAESMYQIKYEELQSLAGKHGDDLRRTKTEISEMNRNISRLQAEIEGLKGQ  327
K18   243  QDLAKIMADIRAQYDELARKNREELDKYWSQQIEESTTVVTTQSAEVGAAETTLTELRRTVQSLEIDLDSMRNL  316

LaA   239  YRLADALQELRAQHEDQVEQYKKELEKTYSAKLDNARQSAERNSNLVGAAHEELQQSRIRIDSLSAQLSQLQKQ  312
LaB1  240  SKLAQALHEMREQHDAQVRLYKEELEQTYHAKLENARLSSEMNTSTVNSAREELMESRMRIESLSSQLSNLQKE  313

           :x. :x xxx:*:x..: :.. x.x  :  :.:...  .: .:.. x..x:.x:.x.x: x :x. :x..:..
```

```
           abcdefgabcdefghijkabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgab
337  NESLERQMREMEENFAVEAANYQDTIGRLQDEIQNMKEEMARHLREYQDLLNVKMALDIEIATYRKLLE  405  Vim
326  NEALEKQLQELEDKQNADISAMQDTINKLENELRTTKSEMARYLKEYQDLLNVKMALDIEIAAYRKLLE  394  NFL

328  RASLEAAIADAEQRGELAIKDANAKLSELEAALQRAKQDMARQLREYQELMNVKLALDIEIATYRKLLE  396  K8
317  KASLENSLREVEARYALQMEQLNGILLHLESELAQTRAEGQRQAQEYEALLNIKVKLEAEIATYRRLLE  385  K18

313  LAAKEAKLRDLEDSLARERDTSRRLLAEKEREMAEMRARMQQQLDEYQELLDIKLALDMEIHAYRKLLE  381  LaA
314  SRACLERIQELEDLLAKEKDNSRRMLTDKEREMAEIRDQMQQQLNDYEQLLDVKLALDMEISAYRKLLE  382  LaB1

       ..:x   x.x:*:. :.:...  . .x .:x xx. .x :x.::x.x*:.*x:x*xx*x:**:.**x***
```

B

|	Vim	NFL	K8	K18	LaA	LaB1	
Vim	X	80	63	58	60	60	
NFL	57	X	60	61	60	57	
K8	41	38	X	63	52	49	Sequence similarity
K18	33	33	31	X	56	56	
LaA	32	32	32	33	X	83	
LaB1	29	30	25	33	62	X	

Sequence identity

C

Figure 5.1.2.1 Structure of IF coil 2. (A) Alignment of the coil 2 sequences for different IF types including human vimentin (type III protein), neurofilament L protein (type IV), cytokeratin 8 (type I), cytokeratin 18 (type II), and lamins A and B1 (type V). New assignment of the hendecad positions in the N-terminal region is highlighted in cyan. Assignment of the heptad positions is in yellow. Stutter is in green. The line below the sequences indicates the sequence conservation score calculated as in Strelkov _et al._, 2002[9]. (B) Amino-acid sequence identity (below the diagonal) and similarity (above the diagonal) for coil 2 of different IF types. The values shown are in percent rounded up to an integer value. For sequence similarity calculation, Blosum62 matrix was used (Henikoff and Henikoff, 1992)[68]. (C) Atomic model of the full vimentin coil 2 shown as ribbon diagram and semitransparent molecular surface. The N-terminal parallel bundle is colored blue, the rest of the D3 fragment is in orange and the previously determined C-terminal cys2 fragment is in yellow. The stutter is colored green.

5.2   Impact of myopathic mutations on desmin structure

Despite the fact we get strong expressions and large quantities (up to ~10 mg per purification batches) of pure and relatively well soluble desmin wt and mutant fragments, our crystallization trials were unsuccessful. Also, the desmin fragments stability over time was greatly increased by removing the N-terminal cystein, which was shown to oxidize over time. These cystein-less fragments make us confident that the nature of our protein fragments during our crystallization process, which was followed for all our samples up to ~3 month (dried drops), stayed the same.

Together with crystallization trials, our protein fragments were partly characterized by circular dichroism, analytical ultracentrifugation and semi-analytical gel-filtrations. Clearly, all our CD spectra showed that these fragments were α-helical. However, with respect to analytical ultracentrifugation and gel-filtration, many of our desmin samples contained a relatively high amount of monomers. In the crystallographer's logic, one would expect IF fragments dimers to crystallize better than monomers, since our relatively short fragments may nearly adopt a globular shape favoring crystal contacts. However, this perspective has to be revisited, since several IF fragment structures do crystallize, even though their solution state was of a monomer. This perspective is reinforced in this thesis by our N-terminal coil 2 structure which was measured as a monomer in solution.

We also made up the assumption that a strong homology or identity with previously crystallized IF fragments (as shown in Figure 4.2.1.1) favor the crystallization of our desmin and lamin fragments. Obviously, such predictions for coiled coils were too naïve and keeps opened more perspectives for fundamental research within this field. One reason behind might be that coiled coils do not necessary staggered among their length into the crystal, but rather crystallize into various spacegroups suggesting that only few crystal contacts are sufficient for crystallization. Also, the D3 structure shown previously reveals that PEG molecules can support a relatively loose helix packing by forming several intermolecular hydrogen bridges.

No crystals were obtained out of the homologous lamin A fragments carrying homologous desmin mutation despite the wild-type fragment readily crystallized in ~10% of our conditions. Therefore, it is likely that the mutations we introduced deeply disturb the α-helical structure of our fragments. Further, our desmin wt fragment

which shares 63% homology (40% identity) to the lamin A fragment did never crystallized. Thus, it would be interesting to mutate several desmin amino-acids into their lamin corresponding ones to get the desmin fragment crystallized. However, according to our mutated lamin experience, I believe that these mutants would not crystallize, even into a desmin context.

Alternatively, all the desmin fragments 'shifts' we produced were smaller than our desmin template. If we consider that the proline mutations accommodate within a coiled coil, then we may eventually consider producing longer flanking sequences to the mutations. With our vimentin D3 structure, we know now that the coiled coil starts already at desmin residue 307 (first heptad *a* position).

Another consideration is the presence of a mixture of *cis-* and *trans-* proline in our expressed mutants, the *trans* form being the energetically favored one. Here, we cannot exclude that such a conformational mixture prevented crystallization. One idea would be to incubate our desmin mutant fragments with *cis-* or *trans-* prolyl isomerases prior to crystallization.

Finally, one might try to get nuclear magnetic resonance (NMR) structures of IF fragments. The NMR technique was already used in 1999 to solve ~35kDa proteins (Yu, 1999)[69], what is large enough to solve our IF fragments which are usually ~20kDa (as dimers).

5.3   Small angle X-ray scattering analysis of desmin wt and mutants

In addition to our efforts to get atomic structures of desmin mutants, we compared with SAXS our desmin mutant assembly complexes versus the desmin wt ones.

In tetramer buffer, we were able to sort our mutants into three classes. The first class of mutants has a similar distance distribution functions than desmin wt, with only one maxima and about the same cross-section diameter as desmin wt. The second class has distance distribution functions showing a second maximum, suggesting the presence of other assembly species. Alternatively, a less dense core due to a 'bulging' within an oligomeric assembly could explain the presence of this second maximum. However, their cross-section diameters (Dmax) stay comparable to that of desmin wt.

Finally, our mutant L370P is an outlier which forms our latest class. It has about twice the size of desmin wt in tetramer buffer. This cross-section size matches the one we observed for our biggest observed desmin wt assembly species (~20 nm) which is comparable to those as measured by EM (~16 nm) for the ULFs. With SAXS measurements, one has to take into account the presence of a hydration shell with ordered water molecules. This hydration shell accounts for few additional nm, but cannot explain such a large discrepancy.

We then added salts or assembly buffer to our samples. Our first observation was that when we triggered desmin assembly in higher ionic strength buffers, about half of our samples turned into gel. In tetramer buffer, however, the only gel-forming mutant was Q389P. This mutant, together with D399Y, was shown to have a high sedimentation coefficient but was not shown to form a gel (Class I, Bar *et al.*, 2006)[59].

With our measurements in increased ionic strength buffer, we were then able to define two major assembly classes (cf. table 4.3.6.1). The first class of mutants is composed of mutants that behave as desmin wt. All of them stay fluid, either in tetramer buffer or with increased ionic strength buffers.

For two of these mutants, i.e. E245D and A360P, our SAXS experiment confirms the EM observation that they form seemingly normal filaments (Bär *et al.*)[59]. Further, the mutant R406W, which was shown to have a disturbed longitudinal annealing and radial compaction defect (Class II), assembles in our experiment more slowly than wt (Figure 4.3.5.1). However, with our SAXS resolution, we cannot assess if ULFs anneal or not. Finally the desmin mutants E413K and I451M were never measured previously, but our SAXS data suggest that they assemble as desmin wt.

Our second group consists of mutants that are similar to desmin wt in tetramer buffer but have an increased assembly upon ionic strength increase. Three of these mutants, i.e A337P, N342D and A357P were shown by EM to have enhanced filament adhesiveness and to form aggregates (Class III). Correspondingly, these mutants turned into gel in our experiment. We could only measure the supernatant of the sample A357P which revealed to have a similar distance distribution function as desmin wt, suggesting that indeed this mutant is able to form filament intermediates like tetramer and ULFs.

In addition, the mutants L345P and R350P, which were shown to have a conserved ULF formation but deterioration of assembly and break-down into small aggregates (Class II), also belong into our second class. These two mutants turned into a gel only after salt addition. Most likely, these aggregates were removed by our centrifugal step prior to the SAXS measurement. Here, the disappearance of a second peak in their distance distribution function compared to the desmin wt correlates the EM observation of a partial assembly.

Lastly, we had a third group composed of two outliers. First, the mutant Q389P completely precipitate in tetramer buffer, despite it was shown by EM to form seemingly normal filaments (Class I). Second, the mutant L370P was shown by SAXS to already assemble in large species. This mutant was reported to have a conserved ULF formation which degrades into small aggregates (Class IV). Bär *et al.* reports that this mutant often retained the outline of earlier precursors, and we cannot exclude here that we measured one of them.

However, we also have some discrepancies between the EM studies and our SAXS data. The mutants A213V, D399Y and Q389P were shown by EM to behave as desmin wt. The most striking difference concerns the mutant Q389P which, as mentioned, precipitate. Interestingly, this mutant and the mutant D399Y had a high sedimentation coefficient in tetramer buffer, suggesting that this two mutants deviates already at the early stages of assembly. Bär *et al.* also reported that the mutant A213V formed a gel that clogged the viscosimeter at concentrations higher than 0.1 mg/ml. Here, the difference we observed might be related to the difference in concentration of our sample (2.4 mg/ml), which may have favored precipitation and gel formation.

As another limitation, we cannot determine with SAXS if our solutions were mono- or polydisperse. Due to the complex nature of intermediate filaments, we may consider that most of our measurements are of mixtures, and may thus deviate of EM or AUC analysis. To overcome this difficulty, it might be interesting in the future to measure desmin assembly in 'real-time' by adding salts directly to the sample under X-rays. Such way would also help us to catch desmin filament intermediates which we have missed in our experiment. Also, we should take some EM pictures to determine if our experimental procedure did not lead to badly assembled filaments.

In particular, it has to be mentioned that while most of the mutants were provided to us by H. Bär / H. Herrmann (Heidelberg), the desmin wt and mutants A360P, L370P and R406W were provided to us by L. Kurochkina (Moscow). Both groups used different expression plasmids and purification protocols as discussed in material and methods (section 3.7.1). Potentially, minor differences may lead to different assembly species and a precise EM comparison between these samples should be carried on.

Last but not least, there were also some differences in the way we prepared our samples. First, Bär *et al.* dialyzed their samples in a high molecular weight regenerated cellulose dialysis tubing (50 kDa cut-off) which excludes impurities but retains desmin monomers (53 kDa). In our case, desmin samples were dialyzed in 3.5 kDa cut-off membranes, yielding to less pure samples. However, we dialyzed our samples more slowly than Bär *et al.*, as 4l of tetramer buffer were gradually added overnight, using a peristaltic pump, to 1l of 8M urea tetramer buffer. We then dialyzed twice more our samples (at 1.6M urea) in 4l tetramer buffer to remove almost all urea. Second, Bär *et al.* initiate filament assembly by adding an equal amount of assembly buffer to the sample in tetramer buffer yielding to the exact final pH of 7.5. In our case, various amounts of assembly buffer were added, lowering our sample pH's to only pH 8.0. Finally, we added NaCl directly to our samples, which is a procedure which was not used by Bär *et al.* Such differences might also be investigated in the future, since they may modify the assembly of desmin.

6. References

1. Holmes, K. C., Popp, D., Gebhard, W. and Kabsch, W. Atomic model of the actin filament. *Nature* 347, 44-49 (1990).

2. Crick, F. H. C. The Packing of alpha-Helices: Simple Coiled-Coils. *Acta Cryst.* B, 1880-1884 (1953).

3. Janmey, P. A., Euteneuer, U., Traub, P. and Schliwa, M. Viscoelastic properties of vimentin compared with other filamentous biopolymer networks. *J Cell Biol* 113, 155-160 (1991).

4. Hesse, M., Magin, T. M. and Weber, K. Genes for intermediate filament proteins and the draft sequence of the human genome: novel keratin genes and a surprisingly high number of pseudogenes related to keratin genes 8 and 18. *J Cell Sci* 114, 2569-2575 (2001).

5. Szeverenyi, I. et al. The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases. *Hum Mutat* 29, 351-360 (2008).

6. Fuchs, E. & Weber, K. Intermediate filaments: structure, dynamics, function, and disease. *Annu Rev Biochem* 63, 345-382 (1994).

7. Hess, J. F., Casselman, J. T., Kong, A. P. and FitzGerald, P. G. Primary sequence, secondary structure, gene structure, and assembly properties suggests that the lens-specific cytoskeletal protein filensin represents a novel class of intermediate filament protein. *Exp Eye Res* 66, 625-644 (1998).

8. Parry, D. A. D. & Steinert, P. M. *Intermediate filament structure* (Springer-Verlag, Heidelberg, 1995).

9. Strelkov, S. V. et al. Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. *EMBO J* 21, 1255-1266 (2002).

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

11. Crick, F. H. C. Is alpha-keratin a coiled coil? *Nature* 170, 882-883 (1952).

12. Wilson, I. A., Skehel, J. J. and Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature* 289, 366-373 (1981).

13. McKay, D. B. & Steitz, T. A. Structure of catabolite gene activator protein at 2.9 A resolution suggests binding to left-handed B-DNA. *Nature* 290, 744-749 (1981).

14. McKay, D. B., Weber, I. T. and Steitz, T. A. Structure of catabolite gene activator protein at 2.9-A resolution. Incorporation of amino acid sequence and interactions with cyclic AMP. *J Biol Chem* 257, 9518-9524 (1982).

15. Landschulz, W. H., Johnson, P. F. and McKnight, S. L. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240, 1759-1764 (1988).

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

17. Burkhard, P. et al. Crystallization and preliminary X-Ray diffraction analysis of the 190-A-long coiled-coil dimerization domain of the actin-bundling protein cortexillin I from dictyostelium discoideum. *J Struct Biol* 122, 293-296 (1998).

18. Kammerer, R. A. et al. An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc Natl Acad Sci U S A* 95, 13419-13424 (1998).

19. Wu, K. C. et al. Coiled-coil trigger motifs in the 1B and 2B rod domain segments are required for the stability of keratin intermediate filaments. *Mol Biol Cell* 11, 3539-3558 (2000).

20. Herrmann, H. & Aebi, U. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. *Annu Rev Biochem* 73, 749-789 (2004).

21. Strelkov, S. V. et al. Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. *J Mol Biol* 306, 773-781 (2001).

22. Herrmann, H. et al. The intermediate filament protein consensus motif of helix 2B: its atomic structure and contribution to assembly. *J Mol Biol* 298, 817-832 (2000).

23. Meier, M. et al. Vimentin Coil 1A-A Molecular Switch Involved in the Initiation of Filament Elongation. *J Mol Biol* (2009).

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

25. Crick, F. H. C. The Fourier Transform of a Coiled-Coil. *Acta Cryst* 6, 685-689 (1953).

26. Stetefeld, J. et al. Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat Struct Biol* 7, 772-776 (2000).

27. Nooren IM, K. R., Sauer RT, Boelens R. The tetramerization domain of the Mnt repressor consists of two right-handed coiled coils. *Nat Struct Biol* 6, 755-759 (1999).

28.	Strelkov, S. V., Herrmann, H. and Aebi, U. Molecular architecture of intermediate filaments. *Bioessays* 25, 243-251 (2003).

29.	Morita, H., Seidman, J. and Seidman, C. E. Genetic causes of human heart failure. *J Clin Invest* 115, 518-526 (2005).

30.	Tokuyasu, K. T., Dutton, A. H. and Singer, S. J. Immunoelectron microscopic studies of desmin (skeletin) localization and intermediate filament organization in chicken cardiac muscle. *J Cell Biol* 96, 1736-1742 (1983).

31.	Tokuyasu, K. T. Visualization of longitudinally-oriented intermediate filaments in frozen sections of chicken cardiac muscle by a new staining method. *J Cell Biol* 97, 562-565 (1983).

32.	Stromer, M. H. & Bendayan, M. Arrangement of desmin intermediate filaments in smooth muscle cells as shown by high-resolution immunocytochemistry. *Cell Motil Cytoskeleton* 11, 117-125 (1988).

33.	Clark, K. A., McElhinny, A. S., Beckerle, M. C. and Gregorio, C. C. Striated muscle cytoarchitecture: an intricate web of form and function. *Annu Rev Cell Dev Biol* 18, 637-706 (2002).

34.	Lazarides, E. Intermediate filaments as mechanical integrators of cellular space. *Nature* 283, 249-256 (1980).

35.	Goebel, H. H. & Warlo, I. A. Progress in desmin-related myopathies. *J Child Neurol* 15, 565-572 (2000).

36.	Goldfarb, L. G. et al. Missense mutations in desmin associated with familial cardiac and skeletal myopathy. *Nat Genet* 19, 402-403 (1998).

37.	Chou, P. Y. & Fasman, G. D. Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. *Biochemistry* 13, 211-222 (1974).

38.	Bar, H., Strelkov, S. V., Sjoberg, G., Aebi, U. and Herrmann, H. The biology of desmin filaments: how do mutations affect their structure, assembly, and organisation? *J Struct Biol* 148, 137-152 (2004).

39.	Kaminska, A. et al. Small deletions disturb desmin architecture leading to breakdown of muscle cells and development of skeletal or cardioskeletal myopathy. *Hum Genet* 114, 306-313 (2004).

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

41.	Steinert, P. M., North, A. C. and Parry, D. A. Structural features of keratin intermediate filaments. *J Invest Dermatol* 103, 19S-24S (1994).

42.	Parry, D. A. D. Hendecad repeat in segment 2A and linker L2 of intermediate filament chains implies the possibility of a right-handed coiled-coil structure. *J Struct Biol* 155, 370-374 (2006).

43.   Harbury, P. B., Plecs, J. J., Tidor, B., Alber, T. and Kim, P. S. High-resolution protein design with backbone freedom. *Science* 282, 1462-1467 (1998).

44.   Woolfson, D. N. & Williams, D. H. The influence of proline residues on alpha-helical structure. *FEBS Lett* 277, 185-188 (1990).

45.   Bar, H. et al. Severe muscle disease-causing desmin mutations interfere with in vitro filament assembly at distinct stages. *Proc Natl Acad Sci U S A* 102, 15099-15104 (2005).

46.   Stromer, M. H., Ritter, M. A., Pang, Y. Y. and Robson, R. M. Effect of cations and temperature on kinetics of desmin assembly. *Biochem J* 246, 75-81 (1987).

47.   The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763 (1994).

48.   Vonrhein, C., Blanc, E., Roversi, P. and Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol Biol* 364, 215-230 (2007).

49.   Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58, 1948-1954 (2002).

50.   Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* 35, 375-383 (2007).

51.   Herrmann, H., Haner, M., Brettel, M., Ku, N. O. and Aebi, U. Characterization of distinct early assembly units of different intermediate filament proteins. *J Mol Biol* 286, 1403-1420 (1999).

52.   Herrmann, H., Kreplak, L. and Aebi, U. Isolation, characterization, and in vitro assembly of intermediate filaments. *Methods Cell Biol* 78, 3-24 (2004).

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

54.   Cole, C., Barber, J. D. and Barton, G. J. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* 36, W197-201 (2008).

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

56.   Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372, 774-797 (2007).

57.   Kuhnel, K. et al. The VASP tetramerization domain is a right-handed coiled coil based on a 15-residue repeat. *Proc Natl Acad Sci U S A* 101, 17027-17032 (2004).

58.   Peters, J., Baumeister, W. and Lupas, A. 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, 1031-1041 (1996).

59.   Bar, H. et al. Impact of disease mutations on the desmin filament assembly process. *J Mol Biol* 360, 1031-1042 (2006).

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

61.   Strelkov, S. V., Schumacher, J., Burkhard, P., Aebi, U. and Herrmann, H. Crystal structure of the human lamin A coil 2B dimer: implications for the head-to-tail association of nuclear lamins. *J Mol Biol* 343, 1067-1080 (2004).

62.   Sokolova, A. V. et al. Monitoring intermediate filament assembly by small-angle x-ray scattering reveals the molecular architecture of assembly intermediates. *Proc Natl Acad Sci U S A* 103, 16206-16211 (2006).

63.   Hess, J. F., Budamagunta, M. S., Shipman, R. L., FitzGerald, P. G. and Voss, J. C. Characterization of the linker 2 region in human vimentin using site-directed spin labeling and electron paramagnetic resonance. *Biochemistry* 45, 11737-11743 (2006).

64.   Parry, D. A. D. & Steinert, P. M. Intermediate filaments: molecular architecture, assembly, dynamics and polymorphism. *Q Rev Biophys* 32, 99-187 (1999).

65.   Milam, L. & Erickson, H. P. Visualization of a 21-nm axial periodicity in shadowed keratin filaments and neurofilaments. *J Cell Biol* 94, 592-596 (1982).

66.   Heins, S. et al. The rod domain of NF-L determines neurofilament architecture, whereas the end domains specify filament assembly and network formation. *J Cell Biol* 123, 1517-1533 (1993).

67.   Schwaiger, I., Sattler, C., Hostetter, D. R. and Rief, M. The myosin coiled-coil is a truly elastic protein structure. *Nat Mater* 1, 232-235 (2002).

68.   Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* 89, 10915-10919 (1992).

69.   Yu, H. Extending the size limit of protein nuclear magnetic resonance. *Proc Natl Acad Sci U S A* 96, 332-334 (1999).

Curriculum Vitae

Stefan Nicolet

Landskronstrasse 62
4056 Basel, Switzerland

Phone: +41 061 321 08 84
Mobile: +41 078 776 08 84
E-Mail: s.nicolet@unibas.ch
          nicolet_stefan@gmx.ch

Born 18.09.1979 in Biel/Bienne, Switzerland
Married

Citizenship: Swiss

STUDIES

1999-2005	Diploma in biochemistry, University of Fribourg, Swizerland Thesis title: "The *Trypanosoma brucei* RNAse Z homologue & Computational analysis of the methionyl-tRNA formyltransferase in trypanosomatids and apicomplexans" Supervisor Prof. A. Schneider
2005-2009	PhD in structural biology, M.E. Müller Institute for structural biology, University of Basel, Switzerland, and Department of Pharmaceutical Sciences, Catholic University of Leuven, Belgium. Thesis title: "Coil 2 of intermediate filaments: its complete structure and impact of desminopathy-related mutations." Supervisors: Prof S.V. Strelkov and Prof. U. Aebi (After spending the first year in Basel I followed my promoter Prof. S.V Strelkov to the Catholic University of Leuven)
2009	Working within the group of Prof. Guy Cornelis, Biozentrum, University of Basel
2010.04.18	Doctorexam

SKILLS

Biochemistry and molecular biology

- Molecular cloning, Northern and Western blotting
- Work with P^{32} labeled nucleotides and H^3 labeled lipids
- Immunofluorescence
- *Trypanosoma brucei brucei* cell culture
- Recombinant protein expression and purification
- Fast protein liquid chromatography (Akta system)

Crystallography and other biophysical methods

- Protein crystallization, crystal soaking, crystal mounting
- Circular dichroism spectroscopy
- X-ray diffraction data collection, SAD/MAD and molecular replacement phasing, refinement
- Small-angle X-ray scattering

Additional experience

- Helped in supervising one technician and two Master students
- Helped in Installation of a new laboratory in Leuven, Belgium
- Set up of Linux computer network for crystallographic computing
- Windows/Linux regular user, C/C++, C# programming and shell scripts

LANGUAGES

Mother tongue: French

Foreign languages: English, German, Portuguese

CONFERENCE PRESENTATIONS

S. Nicolet, H. Bär, H. Herrmann, U. Aebi, S.V. Strelkov "Towards understanding the impact of disease-related mutations on desmin structure", Poster, 4th European Workshop on Intermediate Filaments, Bonn, Germany, 2005

S. Nicolet, L. Kapinos, H. Baer, N. Muecke, J. Schumacher, H. Hermann, U. Aebi, S.V. Strelkov "Toward atomic details of intermediate filaments", Poster, 23rd European Crystallographic Meeting (ECM23), Leuven, Belgium, 2006

S. Nicolet, U. Aebi, S.V. Strelkov "The 2A-L2-2B structure of vimentin", Gordon research conference on intermediate filaments, Oxford, England, 2008

S. Nicolet, H. Herrmann, L.P. Kurochkina, H. Baer, U. Aebi, S.V. Strelkov, "Assembly of myopathy-related desmin mutants as monitored using small-angle X-ray scattering", Poster, 7th Swiss Meeting on Muscle Research, Macolin/Magglingen, Switzerland, 2008

WORKSHOPS

"Data Collection and Phasing Workshop", ESRF Grenoble, France, 2007

Erice 40th course in crystallography "From Molecules to Medicines: Integrating Crystallography in Drug Discovery", Erice, Italy, 2007

PUBLICATIONS

Nicolet, S., Herrmann, H., Aebi, U. & Strelkov, S. V. 2010, **'Atomic structure of vimentin coil 2**', *J Struct Biol*, vol. 170, no. 2, pp. 369-376.

Christ, F., Voet, A., Marchand, A., Nicolet, S., Desimmie, B. A., Marchand, D., Bardiot, D., Van der, V. e. N., Van Remoortel, B., Strelkov, S. V., De Maeyer, M., Chaltin, P. & Debyser, Z. 2010, **'Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication**', *Nat Chem Biol*, vol. 6, no. 6, pp. 442-448.