DISSERTATION
submitted to the
Combined Faculties of the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Biochemikerin Jana Sticht
born in Berlin

Oral examination:

HIV-1 capsid assembly and its inhibition

Referees: Prof. Dr. Hans-Georg Kräusslich
PD Dr. Stephan Urban

meinem Vater

Summary

The structural polyprotein Gag alone drives HIV (human immunodeficiency virus) assembly leading to the release of spherical particles by budding from the plasma membrane. Maturation occurs upon Gag processing, which induces the formation of the infectious virus with its characteristic conical core. Virus assembly and maturation represent targets for anti-viral substances, since these processes are driven by multiple weak protein-protein interactions. A small molecule binding specifically to one of the relevant interfaces is likely to inhibit virus assembly.

In order to select a potential peptide inhibitor of HIV assembly, I screened a library of phages that present random peptide sequences on their surface for binding to Gag-derived proteins. A 12mer peptide (Capsid Assembly Inhibitor, *CAI*) was selected that binds specifically to the C-terminal domain (C-CA) of the HIV-1 capsid protein (CA). *CAI* was shown to inhibit efficiently assembly of immature-like as well as mature-like particles *in vitro*. The complex of *CAI* with the dimerisation defective protein C-$CA_{W184A/M185A}$ was analysed by nuclear magnetic resonance (NMR) spectroscopy. Backbone resonance assignments were obtained for the free as well as the peptide-bound protein. The binding site for *CAI* was mapped to amino acids 169–191 of CA using chemical shift perturbation analysis, which was also used to derive the dissociation constant of the complex ($K_D = 15 \pm 7\,\mu M$). The postulated binding site includes the CA dimer interface, but *CAI* binding does not abolish dimerisation of C-CA. This indicates that *CAI* exerts its inhibitory effect by blocking an assembly interface different from the dimer interface. A preliminary crystal structure of the complex was obtained by our collaborators and strongly supported the above results. The peptide presented here is the first described HIV inhibitor that targets assembly of immature as well as mature particles.

In addition to the characterisation of the new assembly inhibitor, I investigated two aspects of assembly and maturation of HIV-1. pH-induced changes occurring in helix 3 and helix 6 of the N-terminal domain of CA had been correlated to changes occurring *in vivo* upon maturation [45]. In order to characterise thoroughly these conformational changes, I analysed monoclonal antibodies for their affinity to various Gag-derived proteins at pH 6 and 8. The results suggested that pH-induced changes in helix 3 correlate to changes occurring upon Gag cleavage. It had been shown that Gag processing induces the formation of a β-hairpin structure at the N-terminus of CA [41]. Although this conformational change had been suggested to be a major determinant of virus maturation [91], the observed changes in N-CA were independent of hairpin formation; the data presented here suggested the presence of an immature conformation of Gag that is not reflected by protein structures available to date. Additionally, the presence of structurally disordered regions at the C-terminus of CA was shown to influence the efficiency and shape of CA *in vitro* assembly.

Zusammenfassung

Zusammenbau und Freisetzung des Humanen Immundefizienzvirus (HIV) werden von den viralen Strukturproteinen Gag vermittelt. Die Reifung von HIV vom nichtinfektiösen, unreifen Partikel zum infektiösen Virus erfordert die Prozessierung des Vorläuferproteins Gag durch die virale Protease. Dies führt zur Ausbildung des konischen Kapsids, das essentiell für die Infektiösität der freigesetzten Partikel ist. Virus-Zusammenbau und Reifung basieren auf zahlreichen schwachen Protein-Protein-Interaktionen und bieten somit geeignete Angriffspunkte für antivirale Substanzen. Ein kleines Molekül, das an eine dieser Interaktionsflächen bindet, könnte Virus-Zusammenbau oder Reifung inhibieren.

Mit dem Ziel, einen potentiellen HIV Inhibitor zu finden, habe ich mit Hilfe einer Phagenbibliothek ein 12mer Peptid (*Capsid Assembly Inhibitor*, *CAI*) selektiert, das spezifisch an die C-terminale Domäne (C-CA) des HIV-1 Capsid Proteins (CA) bindet und *in vitro* sehr effizient den Zusammenbau von sphärischen sowie tubulären Partikeln hemmt. Eine monomere Variante des Proteins C-CA ($\text{C-CA}_{\text{W184A/M185A}}$) wurde im Komplex mit *CAI* mit Kernresonanz Spektroskopie (NMR) untersucht. Sowohl für das freie als auch für das komplexierte Protein konnte eine Zuordnung der Signale des Protein-Rückgrats erhalten werden. Anhand einer Analyse der Veränderungen der chemischen Verschiebungen durch *CAI* wurde gezeigt, daß die Dissoziationskonstante des Komplexes $15 \pm 7\,\mu$M beträgt und CA-Aminosäuren 169–191 die Bindungsstelle für das Peptid bilden. Obwohl diese Region die CA-Dimerisierungsdomäne beinhaltet, verhinderte *CAI*-Bindung nicht die Dimerisierung von C-CA. Demnach scheint das Peptid eine noch unbekannte Interaktion zu stören. Eine vorläufige Kristallstruktur des Komplexes, die von unseren Kooperationspartnern erhalten wurde, bestätigte die oben beschriebenden Ergebnisse. Mit dem Peptid *CAI* wird hier der erste HIV-Inhibitor beschrieben, der den Zusammenbau unreifer und reifer Partikel verhindert.

Neben der Charakterisierung dieses Inhibitors habe ich zwei Aspekte von HIV-Zusammenbau und Reifung untersucht. Es war vermutet worden, daß pH-induzierte Veränderungen in Helix 3 und Helix 6 der N-terminalen Domäne von CA (N-CA) Konformationsänderungen widerspiegeln, die *in vivo* durch Gag-Prozessierung induziert werden [45]. Mit dem Ziel, diese Konformationsänderungen zu charakterisieren, habe ich die Affinität verschiedener monoklonaler Antikörper für diverse Gag-abgeleitete Proteine bei pH 6 und pH 8 untersucht. Die pH induzierten Konformationsänderungen in Helix 3 schienen den durch Gag-Prozessierung induzierten Vorgängen zu entsprechen, sie waren jedoch unabhängig von den mit Virusreifung assoziierten Konformationsänderungen am N-Terminus von CA [91]. Die Ergebnisse legten nahe, daß keine der verfügbaren Proteinstrukturen die vollständig unreife Konformation des Gag Proteins widerspiegelt. Weiterhin konnte ich zeigen, daß strukturell ungeordnete Regionen am C-Terminus von CA *in vitro* einen Einfluß auf Form und Effizienz der Partikelbildung haben.

Contents

1. Introduction

1.1. The human immunodeficiency virus (HIV)

The human immunodeficiency virus (HIV) is the causative agent of the acquired immunodeficiency syndrome (AIDS), which is currently one of the most prevalent causes of death world wide. HIV belongs to the family of retroviruses, which are named after their characteristic ability to transcribe their RNA genome into double stranded DNA using their own enzyme reverse transcriptase (RT).

Retroviruses can be differentiated into two groups: Simple retroviruses comprise three open reading frames (ORF) coding for the structural proteins (*gag*), the viral enzymes (*pol*) and the viral envelope proteins (*env*). Complex retroviruses such as HIV additionally code for regulatory proteins [25] that optimise the viral replication strategy. Within the retroviruses, HIV belongs to the genus of the lentiviruses [24].

In the following sections the genome organisation and the replication cycle of HIV-1 will be briefly introduced. Since the subject of my PhD thesis included the investigation of several aspects of HIV assembly and maturation, this part of the virus life cycle will be described in more detail focusing on the interactions of the structural proteins driving these processes.

1.1.1. Genomic organisation of HIV-1

The HIV-1 genome consists of two identical linear, single-stranded (+) RNA molecules of about 9 kb length. The genome structure of the provirus is depicted schematically in Fig. 1.1. Shown are the three major ORFs (*gag*, *pol* and *env*, common to all retroviruses and the six additional genes coding for the regulatory proteins of HIV-1 (Vif, Vpr, Vpu, Nef, Tat, Rev). The *gag*-ORF codes for the 55 kD structural precursor protein (Gag) with its domains matrix (MA), capsid (CA), nucleocapsid (NC) and p6 and the two spacer peptides SP1 and SP2. With a frequency of about 5% ribosomal frameshifting leads to the expression of the Gag-Pol precursor protein coding additionally for the

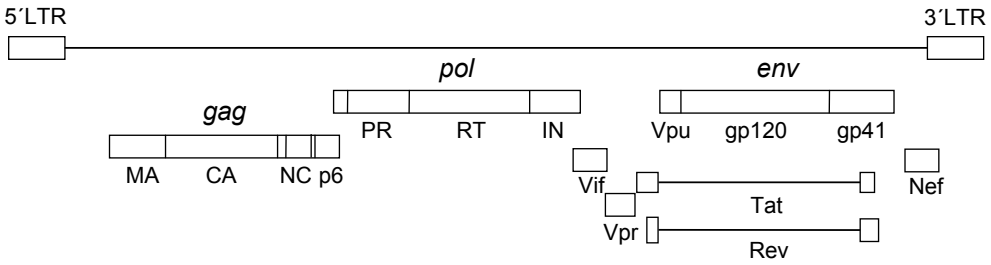

Figure 1.1.: Schematic representation of the HIV-1 proviral DNA. Names of encoded proteins are written below the corresponding gene segments.

viral enzymes protease (PR), reverse transcriptase (RT), and integrase (IN). The *env*-ORF encodes the glycoprotein precursor gp160 with its domains gp120 (surface protein, SU) and gp41 (transmembrane protein, TM). The coding region of the proviral DNA is flanked by long terminal repeats (LTRs), which are essential for reverse transcription and integration. They include the promotor, the primer binding site (PBS) and the binding site for the viral transcriptional activator protein Tat as well as the RNA packaging signal. This so-called ψ-sequence mediates specific incorporation of the viral RNA gennome into nascent particles by interaction with the NC-domain of Gag.

1.1.2. HIV-1 replication cycle

The replication cycle of HIV-1 is depicted schematically in Fig. 1.2. The entry of HIV-1 into cells is mediated by the viral glycoproteins. The protein gp120 specifically binds to CD4 receptor molecules present on cells of the immune system, mainly on T-helper lymphocytes and on macrophages. Subsequent interaction with a coreceptor induces fusion of viral and cellular membranes. This leads to the release of the viral core into the cytoplasm, where it disassembles. In the cytosol the viral RT transcribes the RNA genome into double stranded DNA. The preintegration complex (PIC), which besides the genome probably contains MA, RT, IN and Vpr, is transported to the nucleus [24], where the DNA is integrated into the host cell genome as provirus.

After integration the virus exploits the cellular machinery for gene expression. Under the control of the 5'-LTR promotor, transcription is initiated but is not very efficient. Transcripts are spliced completely before they are exported to the cytoplasm, leading to the successive translation of the regulatory proteins Tat, Nef and Rev. Tat is imported into the nucleus and enhances transcription by recruiting the transcriptional elongation complex. The Rev protein then mediates the export of unspliced and singly spliced

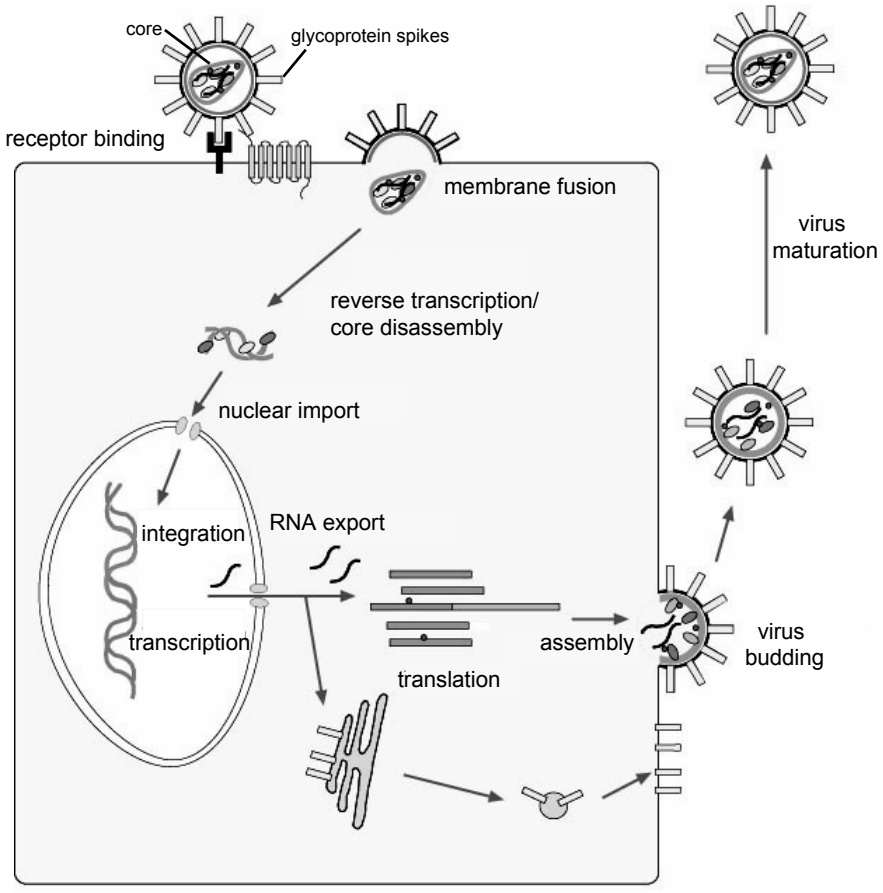

Figure 1.2.: Schematic representation (courtesy of Barbara Müller) of the HIV-1 replication cycle as described in the text.

mRNAs, leading to the efficient translation of all viral proteins and to the availability of genomic RNA in the cytoplasm for packaging into nascent viral particles.

The N-terminally myristylated Gag and GagPol precursor proteins as well as the genomic RNA reach the plasma membrane via a currently not understood pathway, while Env glycoproteins are delivered there by the secretory pathway. All viral components presumably meet at lipid raft microdomains of the plasma membrane [22], where Gag proteins fulfil their functions in mediating assembly [24]: The M-domain (membrane binding domain) of Gag mediates membrane targeting and binding. Gag-Gag interactions (including the CA dimer interface) and Gag-RNA interactions drive immature particle assembly by leading to the oligomerisation of Gag. The viral RNA genome is specifically incorporated into nascent particles by interaction with the NC domain. The incorporation of viral glycoproteins into the viral envelope presumably

is mediated by interaction with the MA domain at the site of budding [34]. Finally, the L-domain (Late-domain), a P(T/S)AP motif present in the p6 domain, recruits the endosomal sorting machinery to the plasma membrane [75]. This leads to the release of initially immature virus particles, displaying a spherical layer of Gag below the membrane (Fig. 1.3). Virus maturation is induced by processing of Gag by the viral protease. Highly regulated sequential cleavage of Gag induces the formation of the infectious virus, with MA remaining as spherical layer below the membrane and CA forming the characteristic conical core, which encases the NC-RNA complex (Fig. 1.3). The appearance of the immature and mature virus particles in the electron microscope is exemplified by pictures of thin sections in Fig. 1.3.

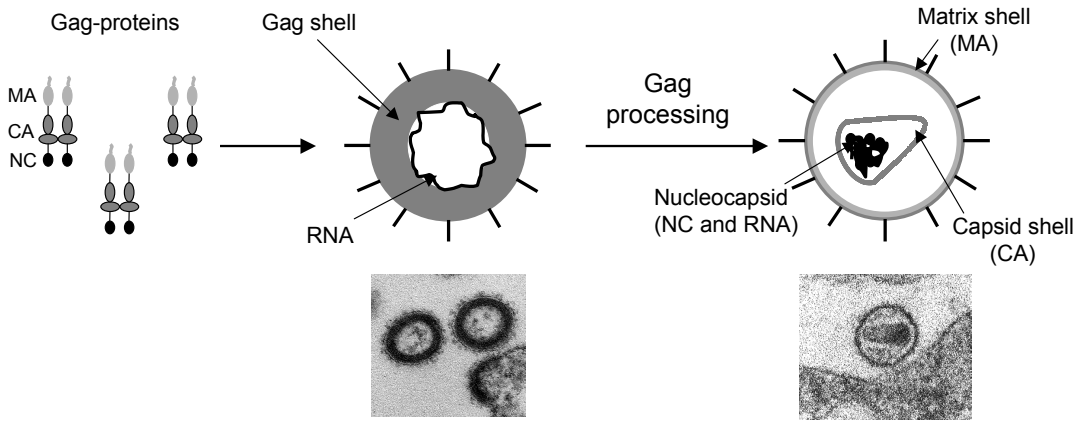

Figure 1.3.: Schematic representation of assembly and maturation of HIV-1. Examples of EM images (Hanswalter Zentgraf) of thin sections of infected cells are shown for illustration. The spherical layer of Gag characteristic for immature virus particles as well as the conical core of the mature virus is clearly visible.

The processes of assembly and maturation of HIV-1 are driven by multiple protein-protein interactions and conformational changes, which will be the focus of the following sections.

1.2. Assembly and maturation of HIV-1

Gag polyproteins have been shown to be essential and sufficient for the assembly of virus-like particles [40]. Driven by multiple weak interactions they initiate a budding structure at the plasma membrane and are released as stable, immature virus particles. These particles were shown to display a spherical layer of 5,000 Gag proteins on average

[11] below the membrane, radially arranged with MA at the membrane and NC in the particle interior [35, 96] (Fig. 1.3). Virus particles are not perfectly round and their size ranges from 119-207 nm [11]. Although the location of assembly is not the same for all retroviruses (e.g. the Mason-Pfizer monkey virus (M-PMV) assembles in the cytoplasm), in all cases initially immature virus particles are released.

Virus maturation is induced by processing of Gag. Conformational rearrangements lead to the infectious particle with a spherical layer of MA below the membrane and in the case of HIV-1 the characteristic conical core that encases the NC-RNA complex (Fig. 1.3). However, core morphology varies between retroviruses: Eccentric spherical (e.g. mouse mammary tumour virus, MMTV) or central spherical (e.g. Rous sarcoma virus, RSV) and cylindrical (e.g. Mason-Pfizer monkey virus, M-PMV) cores are observed as well [24]. The conical capsid of HIV-1 is probably built of only 1,000-1,500 CA molecules [65, 13]. This is less than half of the total amount of CA available in the virus particle after cleavage [11]. Thus, the maturation process does not occur by the condensation of the spherical CA shell into a densely packed conical core as had been assumed earlier [95]. It rather includes disassembly of the spherical shell and the formation of a more loosely packed core particle by a second assembly step [11]. The presence of assembled as well as unassembled CA proteins within the mature particle was also suggested by Lanman et al. [61]. Thus, the loosely packed core might be a property essential for disassembly after cell entry. Such a core destabilising effect had initially been suggested for the cellular protein CypA, which is incorporated into nascent viral particles by interaction with CA. However, this hypothesis had been disproved [94]. It is currently assumed that CypA, after cell entry, counteracts a recently determined cellular factor that restricts HIV-1 infection in certain cell-types (Trim5α) [90]. Also incorporated into nascent viral particles is the viral regulatory protein Vif, which had been shown to interfere with another cellular restriction factor (APOBEC3G) [83].

1.2.1. *In vitro* assembly assays

In vitro systems have been developed over the last years, in order to study retrovirus assembly and maturation. Since it had been shown that the Gag protein alone is both necessary and sufficient for assembly of virus-like particles [40], these systems are solely based on purified recombinant Gag-derived proteins. Assembly of purified recombinant Gag-derived proteins can be induced at high protein concentrations by changing ionic strength and pH. For proteins containing the NC domain, assembly efficiency can be significantly increased in the presence of nucleic acid [24]. This finding supported the

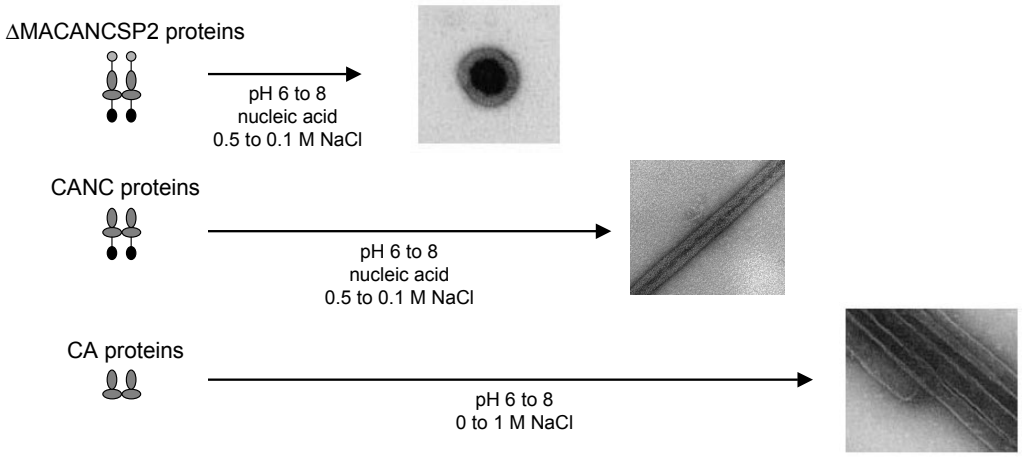

Figure 1.4.: *In vitro* assembly of spherical, immature-like and tubular, mature-like particles is illustrated. The system used here is based on dialysis of purified recombinant Gag-derived proteins. A change in pH and ionic strength induces the formation of spherical particles of ΔMACANCSP2 or of tubular particles of CA or CANC.

hypothesis that RNA forms a scaffold supporting Gag assembly. Depending on the Gag-derived protein used, spherical, immature-like as well as tubular, mature-like particles can be formed, which can be analysed by electron microscopy (EM) as indicated for HIV-1 proteins in Fig. 1.4.

Purified recombinant CA or CANC proteins derived from HIV-1 or RSV Gag have been shown to assemble into tubular particles *in vitro* [16, 44]. Cryo-EM analysis of tubular particles assembled of HIV-1 CA (Fig. 1.4) revealed that the ultrastructure of *in vitro* particles is remarkably similar to that of the conical core of the mature virus [44, 13]. Tubular particles assembled of HIV-1 CANC (Fig. 1.4) were shown to display a lattice packing in Cryo-EM, that is similar to CA tubes and the mature conical capsid (John Briggs, personal communication). In contrast to that, purified recombinant HIV-1 Gag proteins as well as N-terminal elongations of CA only assembled into very small, heterogenous spherical particles [43, 91, 15]. The first *in vitro* particles shown to be very similar to immature viruses were assembled from a HIV-1 Gag-derived protein with a deletion of p6 and MA amino acids 16-99 (ΔMACANCSP2) [45] (Fig. 1.4). In the context of a proviral clone this mutation had been shown to alter the budding site of the virus but not its morphology [31]. The ultrastructure of the immature-like particles assembled *in vitro* was shown to be remarkably similar to that of immature viruses [45, 96, 11]. Finally, *in vitro* assembly of HIV-1 Gag lacking p6 into spherical particles resembling immature viruses was shown to be possible in the presence of

inositol phosphates [14].

In addition to a change in particle shape upon N-terminal elongation of CA, the deletion of the SP1 domain within HIV-1 ΔMACANCSP2 also led to assembly of tubular particles. In the case of RSV, a similar switch was observed [17]. Thus, the *in vitro* system is suitable for the analysis of shape determinants within Gag.

Finally, it had even been shown that depending on pH differently shaped particles can be assembled of the same protein [45, 30]. The protein ΔMACANCSP2 was shown to form spherical, immature-like particles at pH 8 but tubular mature-like particles at pH 6 [45]. Two monoclonal antibodies directed against the N-terminal domain of CA bound specifically to the protein at pH 8 but not at pH 6, while they detected their peptide epitopes independent of pH. These results are illustrated in Fig. 1.5 since they will be discussed in detail in section 3.3. Based on these findings, it was proposed that the two antibodies differentiate between two different conformations of ΔMACANCSP2. At pH 8 the protein is assumed to obtain an immature conformation having the ability to assemble into spherical particles. At pH 6 slight conformational changes would occur as indicated by the loss of antibody binding, leading to assembly of tubular particles similar to those assembled of the mature CA protein. Thus, it was assumed that the pH shift triggers changes within ΔMACANCSP2 that mimic those triggered *in vivo* by Gag processing.

Apart from being a good tool for the analysis of Gag assembly domains, the *in vitro* system can also be used to test potential assembly inhibitors. However, this is limited

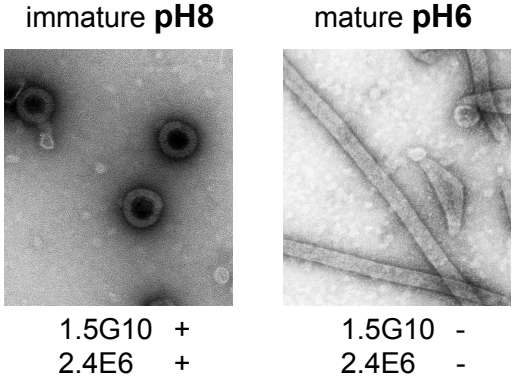

Figure 1.5.: EM images of particles resulting from *in vitro* assembly reactions of ΔMACANCSP2. At pH 8 spherical, immature-like (left) and at pH 6 tubular, mature-like particles are formed. For two monoclonal antibodies (1.5G10 and 2.4E6) a loss in antibody affinity was observed at pH 6 (-) in contrast to pH 8 (+). Courtesy of Ingolf Gross [45].

to the analysis of few selected compounds, since the method used here is based on EM analysis, which is inherently unsuitable for any sort of screening approach. The analysis of assembly by turbidity measurements [62] or by dynamic light scattering [30] would be better suitable for that, but the readout of both of these assays cannot differentiate between protein aggregation and particle assembly. Discerning the nature of the assemblies also requires EM analysis.

1.2.2. The mature HIV-1 CA protein

The Gag protein alone mediates assembly of virus-like particles. The minimal part of Gag sufficient for particle formation was shown to contain a myristylated CA protein, the SP1 spacer peptide and a leucine zipper [3] substituting for the NC-RNA scaffold. Thus, it can be inferred that the CA domain of Gag is essentially required for immature particle assembly. In the mature virus the CA protein alone forms the conical core, which is essential for virus infectivity. Since CA is thus the major determinant of HIV assembly, the protein will be described in detail here.

CA consists of two subdomains, the N-terminal domain (N-CA, amino acids 1-145) and the C-terminal domain (C-CA, amino acids 151-231), linked by five flexible residues, which might allow for different domain-domain orientations during the assembly and maturation. The structures of the single domains have been solved, but no structure of the entire CA protein is available to date. All retroviral CA proteins contain an N-terminal and a C-terminal domain. The CA protein structures solved so far are all very similar despite limited sequence similarity [18]. Deletion analyses have shown that C-CA is essential for virus particle formation [68]. N-CA was dispensable for assembly of immature particles but essential for the formation of the mature core [3, 9]. This indicates that the CA protein mediates both assembly of immature as well as of mature particles.

As depicted in Fig. 1.6, N-CA consists of seven α-helices, a β-hairpin structure and an exposed flexible loop [41]. The N-terminal β-hairpin structure is stabilised by a salt bridge between Pro^{1} and Asp^{51} and this structural element is highly conserved among several retroviruses. By solving the crystal structure of the complex of N-CA with Cyclophilin A [36], its binding site could be mapped to the flexible loop (amino acids 85-99) of N-CA.

The structure of C-CA consists of four α-helices [37, 100] (Fig. 1.6). The protein crystallises as dimer and forms a dimer in solution with a K_{D} of about 18 μM [37]. The

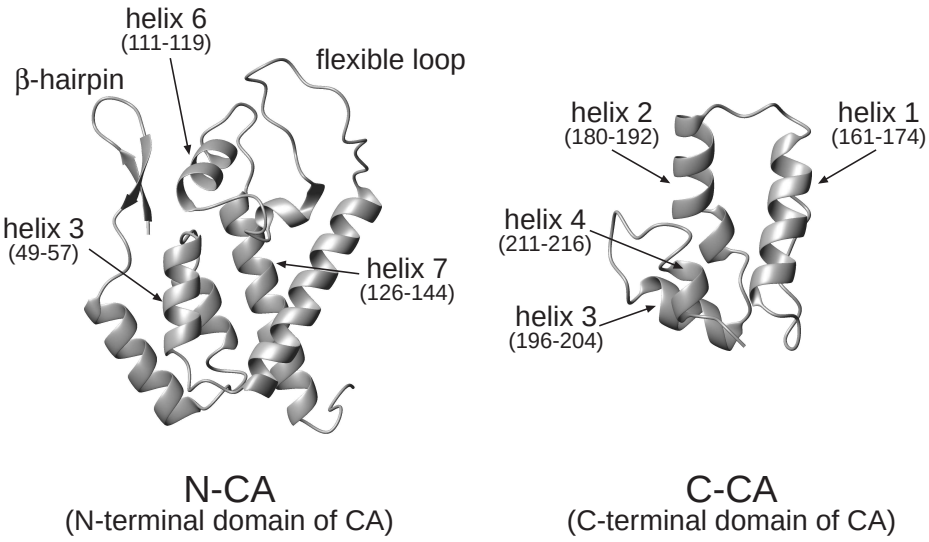

Figure 1.6.: Ribbon diagram of the structure of the N-terminal domain of CA (pdb ID: 1GWP [41]) and of the C-terminal domain of CA (pdb ID: 1A8O [37]). The helices and structural elements relevant for this work are indicated, numbers correspond to CA residues. The structures were generated with MOLMOL [58].

centre of the dimer interface is formed by amino acids W184 and M185, and it is stabilised by several additional residues mainly located within helix 1 and 2 of C-CA [37, 5]. Although this dimer interface was shown to be very important for particle assembly [37], not all retroviral CA proteins readily dimerise. For example the HTLV-I and RSV CA proteins have been shown to be monomeric in solution. In the case of HTLV-I, disulfide formation had been proposed as alternative mechanism to stabilise the CA dimer upon assembly [57]. For RSV it had been suggested that RNA interaction promotes dimerisation of the protein [67]. Apart from the dimer interface the major homology region (MHR, amino acids 153-172), which is the Gag sequence most conserved within all retroviruses, is also located in C-CA.

1.2.3. Assembly interfaces within the immature HIV-1 particle

While the structures of all separate HIV-1 Gag domains have been solved [70, 72, 73, 49, 41, 36, 37, 100, 26], that of full-length Gag is not available to date. However, Gag proteins seem to consist of independently folded domains flexibly linked together. Thus, it is likely that the overall structure of the single domains is largely retained upon cleavage, as confirmed by structures of Gag-derived proteins comprising MA and N-CA

[89] or C-CA and NC [78]. However, local conformational changes induced by cleavage might lead to the formation of interfaces different from those present in Gag. Such Gag-Gag interfaces presently can only be inferred from mutational analyses. After rough indications on the importance of Gag domains had been obtained from insertion and deletion analyses [21, 81, 19], side-chain mutations yielded a detailed idea of important regions: The integrity of the C-CA dimer interface was shown to be influential but not essential for immature particle assembly as was analysed by mutation of interface residues W184 and M185 in the context of an infectious clone [37, 92]. Mutation of amino acids 212-216 of CA also impaired virus release [53]. An assembly interface specifically important for immature particle assembly might be located within helices 4-6 of N-CA, since mutations in that region reduced virus release without effecting CA *in vitro* assembly [92]. Mutations at the base of C-CA (K158, D197, P224) abolished virus release concomitant with Gag accumulating in sheets below the plasma membrane [92]. This suggests a defect in the curvature of the Gag shell needed for the formation of a virus bud. The same phenotype had been observed for SP1 deletion [59] and point mutants [2], indicating that both regions might be involved in curvature of the Gag shell.

The MA domain within Gag is important for virus assembly in that it targets Gag proteins to the membrane. By solving the structure of MA [70, 72, 73, 49] it became evident, that a large basic patch within the first 31 MA amino acids mediates membrane binding together with the N-terminal myristate. The fact that the mature monomeric protein crystallises as trimer suggested that MA might provide a trimerisation interface for immature particle assembly. However, this had not been proven to date. The NC protein as domain of Gag is responsible for the specific packaging of the viral genome. Two CCHC-zinc fingers, which are highly conserved among all retroviruses, as well as flanking basic residues interact with the ψ-sequence in the 5'-LTR of the viral genome. Originally, NC-NC interactions had been postulated to be essential for virus assembly. However, it is still a matter of debate whether NC-NC interactions are important or whether RNA binding mediates Gag oligomerisation [24].

The question of which of the described interfaces stabilise the Gag lattice in the immature particle could be answered by structural analysis of the whole particle. However, since HIV-virions are not icosahedral and inconsistent in size and shape [35, 95] no high resolution structure or model of the immature virus particle is available to date.

1.2.4. Conformational changes accompanying maturation

Conformational transformations triggered by protein processing or by pH changes occur during maturation of many viruses. In the case of HIV-1, the radial Gag lattice described above is rearranged after highly regulated and sequential [93] cleavage by the viral protease. The first cleavage between SP1 and NC leads to the condensation of the nucleocapsid. Cleavage between MA and CA then releases CASP1 from the membrane, but only the last processing step between CA and SP1 induces the formation of the mature conical core.

The MA-CA cleavage was shown to have two major effects: MA as subdomain of Gag seems to have a higher membrane binding affinity than the cleaved MA protein [82, 104]. This was attributed to the masking of the myristate in the mature protein [85], which presumably is triggered by the removal of self-associating Gag domains upon processing [88]. On the site of CA, the release from MA induces the formation of a β-hairpin structure, as depicted in Fig. 1.6. Since this structure is stabilised by a salt bridge from the N-terminal amino-group of Pro^{1} and the side-chain carboxy group of Asp^{51}, it is evident that it cannot form in the uncleaved protein. Accordingly, the structure of $MACA_{151}$ showed a flexible and unstructured linker between MA and CA [89]. Thus, hairpin formation was postulated to present a conformational switch between immature and mature proteins and the mutation of Asp^{51} indeed led to altered core formation and inhibited assembly of tubular particles *in vitro* [91]. However, the deletion of the first 13 amino acids of CA did not abolish assembly of tubular particles *in vitro* [44], indicating that hairpin formation is not the major determinant of mature particle assembly, although this structural element is conserved among all retroviruses analysed. In addition to conformational changes directly at the site of processing, helix 3 and helix 6 of N-CA (Fig. 1.6) were suggested to undergo conformational changes characteristic of virus maturation [45], as described in detail above in section 1.2.1.

The final cleavage of CA-SP1 was shown to be important for virus maturation since mutation of the cleavage-site led to non-infectious particles with aberrant cores. This cleavage had been postulated to be slowed down by the presence of a putative α-helix spanning the CA-SP1 junction, which was also implied to represent a Gag specific assembly interface [2]. A similar motif was suggested to be present in MLV CA [23]. However, in the case of HIV-1 the presence of the helix could not have been confirmed by structural analysis [100, 78].

1.2.5. Assembly interfaces within the mature HIV-1 core

Conical core formation is the result of the conformational changes induced upon cleavage. Thanks to the development of the *in vitro* assembly system tubular particles could be formed and analysed as representatives for conical cores. Modelling based on cryo-EM reconstructions of *in vitro* assembled tubular particles suggested that the conical core is formed of hexameric arrangements of CA, with a few pentameric defects leading to closed ends of the particle [65]. The hexameric rings were suggested to be based on interactions of helix 1 and 2 of N-CA. Recently the crystal structure of hexameric murine leukemia virus (MLV) N-CA has been solved. MLV N-CA is structurally very similar to HIV N-CA and the structure suggested that hexamers are stabilised by interactions of helix 1, 2 and 3 [76]. Mutational analysis of surface residues of CA supported the importance of helix 1-3 for assembly [92, 38]. Hexamers were suggested to be connected to six neighbouring CA proteins via the dimer interface in the C-terminal domain of CA (C-CA) [65].

Apart from the well described dimer interface, Lanman et al. recently provided evidence for an additional N-CA/C-CA intersubunit interface in the mature virus [60, 61]. Within particles assembled *in vitro* out of CA amino acids 55-68 within N-CA were shown to be involved in a protein-protein interaction that was not seen in the soluble CA protein. Crosslinking of this region to C-CA suggested the presence of an intersubunit interface [60] as had been postulated earlier [62]. A similar analysis performed on mature as well as immature virus-like particles indicated that this interaction only occurs in mature particles [61]. Evidence for a N-CA/C-CA interaction had been obtained in the case of RSV as well [10].

1.2.6. HIV-1 assembly inhibitors

Since the mid-90ies, the progression to AIDS can be delayed or halted using a combination of at least three anti-viral drugs targeting either HIV reverse transcriptase (RT) or protease (PR) (highly active anti-retroviral therapy, HAART). Since recently, additional drugs are used that block HIV entry by targeting the fusion mechanism of Env proteins. The emergence of drug resistant mutants make the development of new anti-viral drugs directed against other steps in the virus life cycle a subject of highest priority in AIDS research. Virus assembly and maturation of infectious particles represent additional targets for anti-viral substances, since these processes are driven by multiple specific protein-protein interactions. Compounds interfering with some of

these weak interfaces are likely to abolish virus structure and infectivity. However, as described above, especially the interactions driving assembly of immature virus particles and thus driving virus release are not understood in detail, which impeded the development of such drugs to date.

The attempt to develop assembly inhibitors targeting Gag-Gag interfaces was mainly pursued with limited success using CA fragments [79, 50, 39]. The only two small-molecule compounds described to interfere with Gag target the maturation of the virus. The compound *CAP-1* abolishes conical core formation by binding to N-CA but does not inhibit virus release [87]. The compound PA-457 prevents cleavage of CA-SP1 [64, 103, 102] and thus also interferes with core formation. However, it is evident that a compound that targets assembly of immature particles will affect the virus life cycle in an earlier stage. Ideally, a compound should be developed that targets assembly of both immature and mature particles and thus would interfere with two steps of the viral life cycle.

1.3. The objective of this work

As mentioned above, a major aim of HIV research is the development of new inhibitors. Assembly and maturation of HIV represent suitable targets for anti-viral substances, but drug development have been impeded by the lack of knowledge concerning the interactions driving assembly and maturation.

One goal of my PhD thesis was the selection of peptides that bind tightly and specifically to Gag-derived proteins using a Phage Display approach. The selected peptides were to be analysed for assembly inhibition in the *in vitro* system. The complex of Gag-derived proteins with peptides inhibiting assembly *in vitro* was to be characterised biochemically and structurally in order to derive a mechanism of inhibition.

Furthermore, in order to elucidate the mechanisms driving assembly and maturation, the influence of Gag domains on *in vitro* assembly was to be investigated using various CA derived proteins. Additionally, the conformational changes within N-CA that had been suggested to be characteristic for virus maturation [45] were to be analysed in detail.

2. Materials and methods

2.1. Molecular Biology

2.1.1. Buffers

LB-medium:	50 g peptone, 25 g yeast extract, 50 g NaCl ad 5 l H_2O → autoclave
LB-agar:	1 l LB-medium, 15 g agar, 100 mg ampicillin
TE:	10 mM Tris pH 8.0, 1 mM EDTA pH 8.0
TENS:	10 ml TE-buffer, 250 μl 10% SDS, 100 μl 10 M NaOH
1× Loading buffer:	10% glycerol, 0.25% bromphenol blue, 0.25% xylene cyanol
1× TAE:	40 mM Tris, 5 mM acetic acid, 1 mM EDTA

2.1.2. Bacteria

For plasmid preparations and cloning the *E. coli* strain XL1-Blue (Stratagene) with the genotype hsdR17, endA1, supE44, thi-1, recA1, gyrA96, relA1, lac[F', proAB, $lac^qZ\Delta M15$, $Tn10(Tet^R)$] was used, since these bacteria show particularly few recombination events. For protein expression either the *E. coli* strain BL21(DE3) (Stratagene) with the genotype B, F^-, dcm, ompT, $hsdS(r_{B^-}m_{B^-})$, gal, (DE3) or the BL21(DE3)CodonPlus-RIL (Stratagene) was used. The latter overcomes limited expression due to *E. coli* codon usage, because it supplies rare tRNAs on a chloramphenicol resistance plasmid. Both strains code for the T7 RNA polymerase under the control of a *lac* promotor and operator and express the lac repressor which binds to the operator and inhibits T7 RNA polymerase expression, until it is induced by the addition of the lactose analogue IPTG.

For phage amplification the *E. coli* strain ER2738 (New England Biolabs) with the genotype lac^q Δ(lacZ)M15 $proA^+B^+$ zzf:: $Tn10(Tet^R)$/fhuA2 supE thi Δ(lac-proAB) Δ(hsdMS-mcrB)5 ($r_k^-m_k^-McrB^-$) was used. This is a robust F^+ strain with a rapid growth rate and it is particularly well suited for M13 phage amplification. The F-factor also confers tetracyclin resistence to allow for selection.

2.1.3. Plasmids

The used protein expression plasmids are all based on the vector pET11c (Novagen). The vector map is shown in the appendix A.1. The plasmid contains a T7 operator and promotor, which is recognised by the T7 RNA polymerase but not by the host *E. coli* RNA polymerase. Thus, host strain bacteria such as BL21(DE3) are needed that express the T7 RNA polymerase tightly regulated. The pET11c plasmid also codes for the lac repressor to keep the level of basal expression of the target gene very low. Upon induction with IPTG host bacteria express the T7 RNA polymerase, which then transcribes the target gene. An ampicillin resistence gene is also encoded by the pET11c plasmid to allow for selection. All target genes used here were expressed without tag and were inserted between the NdeI and BamHI restriction sites in the multiple cloning site. The insert coding for ΔMACANCSP2 was derived from the HIV-1 BH10 isolate [80], all other proteins were derived from the pNL4-3 isolate [4]. CA variants were cloned by PCR or PCR mutagenesis using a modified plasmid (pET11c-CA(Cass), obtained from Mira Grättinger). This plasmid codes for the HIV-1 CA_{NL4-3} protein, but the DNA sequence contains silent mutations introducing additional restriction sites.

Infectious plasmids coding for mutant CA proteins (CA207A, E212A, Q219A and P224A) were obtained from Wesley I. Sundquist, University of Utah, Salt Lake City, USA. The capsid coding region was cloned into the pET11c expression plasmid by PCR.

2.1.4. Polymerase chain reaction (PCR)

PCR allows specific and exponential amplification of DNA sequences by repetitive melting of double stranded DNA, hybridisation of oligonucleotides complementary to the 3'-end of each DNA strand and synthesis of new strands by a thermostable polymerase in the presence of Mg^{2+} and dNTPs. Temperature and length of PCR reaction steps was adjusted as recommended by the manufacturer for Pfu Polymerase (Promega), an enzyme with relatively low processivity but a reliable proof reading function. The oligonucleotides used here had a length of 20–35 nt (s. appendix A.2), a melting temperature of 58–70°C as determined by T_M=4(G+C)+2(A+T) and if necessary contained restriction sites for suitable enzymes. PCR mutagenesis was performed using overlapping oligonucleotides including the mutation. The amplified DNA fragments were purified by gel electrophoresis: Samples were mixed with loading buffer and separated in TAE buffer using 1–1.5% (w/v) agarose gels including 2 μg/ml ethidium bromide.

A 1 kb ladder marker was used as standard. Bands, visualised by UV fluorescence of intercalated ethidium bromide, were cut out and the DNA was extracted with the NucleoSpin® Extract Kit (Macherey-Nagel) follwoing the instructions of the manufacturer (the agarose is melted at 50°C in the presence of chaotropic salts, DNA is then bound to a small ion exchange column and eluted with TE buffer).

2.1.5. Plasmid cloning

Purified PCR fragments and 5–10 μg pET11c plasmid were cut with suitable restriction enzymes (here usually NdeI and BamHI, MBI Fermentas, buffers and volume according to the instructions). To avoid religation, linearised plasmids were dephosphorylated at the 5'-end with 5 U calf intestinal alkaline phosphatase (CIAP, MBI Fermentas) for 30 min at 37°C. Vector fragments were purified by gel electrophoresis as described above and digested PCR products were purified with the NucleoSpin® Extract Kit (Macherey-Nagel) according to the instructions.

For ligation about 200 ng linearised plasmid was incubated over night with a threefold molar excess of PCR product in a volume of 20 μl with 1 U T4-DNA Ligase (MBI Fermentas) at 16°C. As negative control a ligation was performed without PCR product. Reaction mixtures were used to transform bacteria.

2.1.6. Transformation of bacteria and plasmid preparation

50 μl competent bacteria were thawed on ice, mixed with 20 μl ligation reaction or with 1 μl pure plasmid DNA and left on ice for 20 min. To allow plasmid DNA to enter cells, bacteria were incubated for 90 s at 42°C and then cooled on ice. After addition of 800 μl LB-medium, bacteria were incubated for 30–45 min at 37°C to allow for the synthesis of β-lactamase, which confers ampicillin (AMP) resistence. Bacteria were pelleted, the supernatant was decanted and the resuspended pellet was spread on LB-AMP-Agar plates. Plates were left over night at 37°C. As negative control the ligation without insert was transformed as well.

For minipreparation, colonies were picked from plates and transferred to 3 ml LB-medium (100μg/ml AMP) and incubated over night. Plasmid DNA was then extracted by alcaline lysis: 1.5 ml overnight culture was centrifuged for 10 s at 14,000 rpm. The supernatant was decanted and the pellet was resuspended in the remaining medium. Addition of 250 μl TENS-buffer was followed by 5 s mixing, additon of 150 μl 4 M

NaAc pH 4.5 and again 5 s mixing, then centrifugation for 2 min at 14,000 rpm at room temperature. The supernatant was transferred in a new reaction vessel and DNA was isolated by ethanol precipitation: The supernatant was mixed with 1/10 of its volume 3 M NaAc pH 5.2 and the threefold volume iced absolute ethanol. After incubation on ice for 20 min precipitated DNA was pelleted by centrifugation for 30 min at 4°C and 13,000 rpm. The pellet was washed with 70% ethanol, dried and resuspended in 50 μl TE buffer with 0.02 mg/ml RNAseA. To test purified plasmids for the right inserts, 4 μl miniprep DNA was digested with suitable restriction enzymes and the cleavage pattern was analysed on agarose gels.

Larger amounts of plasmid DNA were purified from 100 ml overnight culture using the NucleoBond® PC100 Kit (Macherey-Nagel) according to the manufacturer's instructions.

2.1.7. Concentration and purity of DNA

The absorption of DNA samples was measured with a DU640 spectrophotometer (Beckmann Coulter) at 260 and 280 nm against TE buffer. DNA bases have an absorption maximum at 260 nm and for dsDNA an optical density (OD_{260}) of 1.0 corresponds to a concentration of about 50 μg/ml. Proteins have an absorption maximum at 280 nm so that a DNA sample is considered to be pure when the ratio of OD_{260}/OD_{280} is 1.8–2.

2.1.8. Sequencing of plasmid DNA

For the sequencing reaction 50 fmol plasmid-DNA and 1.6 pmol of a forward or reverse oligonucleotide was used. The molar mass of a plasmid was calculated according to the following fomula $\left[\frac{2}{3}\,\text{kb}\right]\,\mu\text{g} \sim 1\,\text{pmol}$. Plasmid DNA was mixed with water to a final volume of 20 μl and heated for 2 min at 96°C. The mixture was cooled on ice and oligonucleotide and 6 μl Master Mix (Beckman Coulter) were added. The PCR reaction was performed as follows:

	96°C	50°C	60°C	4°C
30 cycles	20 s	20 s	4 min	-
1 cylce	-	-	-	∞

The reaction was stopped by addition of 5 μl stop-solution (2 μl 3 M NaAc pH 5.2 (Sigma), 2 μl 100 mM EDTA pH 8.0 (Sigma) and 1 μl glycogen (Beckman). DNA was precipitated by addition of 60 μl 95% ethanol (−20°C) for 10 min at −20°C. After centrifugation for 15 min at 15,000 rpm and 4°C the pellet was washed with 70% ethanol (−20°C) and dried in a speed vac. Pellets were resuspended in 40 μl smple loading solution (SLS) (Beckman Coulter), overlayed with mineral oil and loaded on a CEQ-2000 sequencer according to the manufacturer's protocol (Beckman Coulter).

2.2. Protein biochemistry

2.2.1. Buffers

10× PBS:	80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4 ad 1 l H_2O
1× SDS buffer:	30.3 g Tris, 144 g glycine, 10 g SDS ad 10 l H_2O
Fixative:	45% methanol (v/v), 10% acetic acid (v/v)
Coomassie G:	0.0025% Coomassie G (w/v), 10% acetic acid (v/v)
Blot transfer buffer:	40 mM Tris, 20% methanol, 39 mM glycine, 0.0375% SDS → pH 9.2
3× Sample buffer:	62.5 mM Tris pH 6.8, 10% glycerol, 2% SDS 5% β-mercaptoethanol, 0.005% bromphenol blue (w/v)
30% AA/Bis (200:1):	500 ml 30% acrylamide, 0.75 g bisacrylamide
5× M9 stock:	30 g Na_2HPO_4, 15 g KH_2PO_4, 2.5 g NaCl → pH 7.0, ad 1 l, autoclave
M9 medium:	100 ml M9 stock, 2 ml 50% glucose, 0.5 ml 0.1 M $CaCl_2$ 0.5 ml 1 M $MgSO_4$, 0.5 ml 100 mg/ml AMP, 0.5 ml 0.4 g/ml thiamin, 0.5 g $^{15}NH_4Cl$, (1 g ^{13}C-glucose), ad 500 ml H_2O

2.2.2. Expression of recombinant proteins in *E. coli*

To test for protein expression, plasmids were transformed in *E. coli* BL21(DE3) bacteria as described. A colony from an LB-agar plate was transferred to 20 ml LB medium with 100 μg/ml ampicillin and incubated at 37°C to an optical density ($OD_{600\,nm}$) of 0.6. Protein expression was induced with 0.5 mM IPTG and cultivation continued for another 1–4 h. Samples before induction and after protein expression were analysed on denaturing polyacrylamide gels.

For large scale purification in 1–5 l LB medium, either colonies were picked from LB-AMP plates and amplified in a small volume over night prior to expression in the liter scale or freshly transformed bacteria were directly transferred to 100 ml LB-AMP medium for over night growth. Expression of the CA protein as well as all CA variants was allowed for a maximum of 2 h in order to ensure assembly competence, all the other proteins were expressed for 4 h.

^{15}N and / or ^{13}C isotope labelling of proteins was achieved by growing the transformed bacteria in 20 ml LB-AMP medium over night. Bacteria were pelleted, resuspended in PBS and transferred in large volumes of M9 minimal medium supplemented with 1 g/l ^{15}N-NH_4Cl (CK Gas Products Limtd.) or 1 g/l ^{15}N-NH_4Cl and 2 g/l ^{13}C-Glucose (Cambridge Isotope Laboratories, Inc.). Protein expression was induced at an OD_{600} of 0.6 with 0.5 mM IPTG and bacteria were allowed to grow for an additional 5 h.

After expression bacteria were pelleted and stored at −20°C.

2.2.3. Bacterial lysis and ammonium sulfate precipitation

Bacteria were lysed by 3× sonification (Branson sonifier 250) for 2 min at 50% intensity while cooling on ice. Cellular debris was removed by centrifugation for 10–20 min at 14,000 rpm. The pellet was resuspended in half of the original volume lysis buffer and sonified and centrifuged again. Both supernatants were combined and the overexpressed protein was precipitated by addition of 25%-50% ammonium sulfate, which was added successively over a time of 15 min to the protein solution under constant stirring. After additional stirring for at least 301 min, the precipitated protein was pelleted by centrifugation for 15 min at 14,000 rpm. Pellets were resuspended and stirred for another 30 min. A final centrifugation step of 5 min at 14,000 rpm was performed to pellet unsoluble material. The supernatant was used for further purification.

2.2.4. Protein purification in general

Three different chromatographic methods were used here: Ion exchange chromatography to separate proteins based on there surface charge, affinity chromatography to specifically purifiy proteins with characteristic functions and gel filtration to separate proteins by their size. If not mentionend differently all chromatographic purifications were run on a BioCAD Sprint Perfusion chromatography system (Applera Dtld.

GmbH). Buffers used for POROS columns had to be filtered and buffers used for gel filtration had to be degassed in addition.

Proteins were concentrated using Vivaspin concentrators (Vivascience) and dialysis was performed in SpectraPor dialysis tubing using the suitable molecular weight cutoff for the respective proteins. Purified proteins were stored at −80°C.

Molecular weight, pI value and extinction coefficients as determined by the computer program PROTPARAM are shown in the appendix A.3.

2.2.4.1. Anion exchange chromatography

For all proteins purified here, anion exchange chromatography was used initially to remove nucleic acids and other highly negative charged impurities. A batch adsorption protocol with diethylaminoethyl (DEAE)-cellulose (Whatman DE52) was used for CA and C-terminal truncated variants as described [44, 42] and a POROS HQ20 column with a volume of 11 ml and a flow rate of 10 ml/min was used for all other proteins. For this purification step, pH and salt concentration were chosen in order to avoid binding of the target protein to the matrix.

2.2.4.2. Cation exchange and heparin affinity chromatography

As second purification step either cation exchange (POROS HS, column volume of 1.6 ml, flow rate 10 ml/min, sulfopropyl matrix, Perseptive Biosystems) or heparin affinity (POROS HE, column volume of 1.6 ml, flow rate 10 ml/min, Perseptive Biosystems) chromatography was used.

Conditions for cation exchange chromatography were chosen so that the protein of interest bound to the column, meaning that the overall surface charge of the protein had to be positive and the salt concentration low enough. The theoretical pI as calculated by the program PROTPARAM was used as starting point to find the optimal buffer conditions for ion exchange chromatography. However, in some cases the real pI can differ significantly if charged residues are buried in the folded protein. Salt concentrations have to be kept low for weakly charged proteins. After washing away unbound proteins, the bound fraction was eluted with increasing salt concentration.

By using affinity chromatography usually a higher grade of purity can be achieved, because column binding is much more specific. Nucleic acid binding proteins such

as the HIV-1 NC protein also bind to heparin. Thus, the proteins CANC, C-CANC and NC were purified by heparin affinity chromatography instead of cation exchange chromatography. Bound protein was also eluted from the column by increasing the salt concentration.

2.2.4.3. Gel filtration

After the second purification step proteins were pure to a grade of near homogeneity. Those proteins used for structural analysis were further purified by gel filtration. This method is based on porous column material, so that a larger volume is available for small proteins. Thus, at a constant flow rate large proteins elute faster from the column than small proteins. In case of non-globular proteins the retention time is also dependent on their shape. Proteins larger than the pore size elute with the void volume from the column and cannot be separated. Here, Superdex 75 (Amersham Biosciences) columns were used with a separation range of 3 to 70 kD.

2.2.5. Purification of ΔMACANCSP2, CANC and CA

ΔMACANCSP2 was expressed and purified as described [45, 42]. For *in vitro* assembly of ΔMACANCSP2 in 0.1 kD cutoff dialysis tubing, gel filtration (Superdex 75 HR10/30, Amersham Biosciences) was performed with a flow rate of 0.5 ml/min in 30 mM MES pH 6.0, 500 mM NaCl, 1 mM EDTA, 2 mM DTT, in order to separate the protein from higher order aggregates. CANC was expressed and purified as described [44], except that heparin affinity chromatography was performed in HEPA buffer (20 mM Tris pH 7.5, 7% glycerol, 10 μM $ZnCl_2$, 2 mM DTT) and the protein was stored in 50 mM Tris pH 7.5, 500 mM NaCl, 10% glycerol, 5 μM $ZnCl_2$, 2 mM DTT [42]. CA as well as truncated or mutated CA variants and CASP1 were purified as described for CA [44, 42], with the exception that protein expression was only allowed for a maximum of 2 h to ensure assembly competence and the lysis buffer used here contained 1 M NaCl. The pH of cation exchange buffers was adjusted for CA variants to account for a different pI value, if necessary. In case of the protein CA220, the pH of the buffer was not adjusted, instead the protein was collected pure enough in the flow-through of the cation exchange column.

2.2.6. Purification of C-CANC and NC

After cell lysis by sonification in 50 mM Tris pH 7.5, 1 M NaCl, 20 μM $ZnCl_2$, 10% glycerol, 5 mM DTT, C-CANC was precipitated with 25% ammonium sulfate. The pellet was resuspended in HQ-buffer (30 mM Tris pH 8.2, 500 mM NaCl, 20 μM $ZnCl_2$, 5 mM DTT) and C-CANC was purified by anion exchange chromatography (POROS HQ20). The flow-through was diluted with three volumes 1× HE buffer (2× HE buffer: 60 mM HEPES pH 7.4, 40 μM $ZnCl_2$, 20% glycerol, 5 mM DTT) and loaded onto a heparin affinity column (POROS HE, Applied Biosystems). Bound protein was eluted with step segments of 300 mM, 500 mM, 600 mM and 1 M NaCl in 1× HE buffer. Peak fractions were pooled, concentrated and purified to a high grade by gel filtration (Superdex75, HR10/30, Amersham Biosciences) in 50 mM Tris pH 8, 100 mM NaCl, 2 mM DTT. The purified protein was stored in 25 mM MES pH 6.5, 50 μM, 250 mM NaCl, 10% glycerol, 5 mM DTT.

The protein NC was obtained by cleavage of C-CANC with the HIV protease (PR). C-CANC was dialysed against 2× protease buffer (50 mM MES pH 6.5, 100 μM $ZnCl_2$, 500 mM NaCl, 20% glycerol, 2 mM DTT). PR was added in a molar ratio of C-CANC:PR of 1000:1, the mixture was diluted to a final 1× protease buffer concentration and digestion was allowed for 3 h at room temperature. The sample was diluted 1:1 in 2× HEPA buffer (100 mM Tris pH 7.4, 40 μM $ZnCl_2$, 20% glycerol, 5 mM DTT) and C-CA was separated from NC by heparin affinity chromatography. NC protein that bound to the column was eluted with a gradient from 0.1 to 0.8 M NaCl. Peak fractions were pooled and dialysed against 50 mM Tris pH 7.4, 20 μM $ZnCl_2$, 10% glycerol, 100 mM NaCl, 2 mM DTT. The flow-through containing C-CA was concentrated and further purified by gel filtration in 50 mM Tris pH 8, 100 mM NaCl, 2 mM DTT (Superdex75, HR10/30, Amersham Biosciences).

2.2.7. Purification of C-CA and C-CA$_{W184A/M185A}$

For expression of ^{15}N and $^{15}N/^{13}C$ labelled protein, protein expressing bacteria were grown in M9 minimal medium as described above. After cell lysis by sonification in 50 mM Tris pH 7.5, 10% glycerol, 1 M NaCl, 5 mM DTT, proteins were precipitated with 50% ammonium sulfate. The pellet was resuspended in 30 mM Tris pH 8.0, 500 mM NaCl, 1 mM EDTA, 1 mM DTT and the protein was purified by anion exchange chromatography (POROS HQ20). The flow-through was dialysed over night against 30 mM MES pH 5.5, 1 mM EDTA, 2 mM DTT, to adjust the pH to 5.5 and to reduce the salt

concentration. Protein was then loaded onto a cation exchange column (POROS HS) and bound protein was eluted with a gradient from 0 to 0.5 M NaCl. Peak fractions were pooled, concentrated by Vivaspin centrifugation (Vivascience) and further purified by gel filtration. For C-CA a Superdex 75 HR 10/30 was used. In the case of C-$CA_{W184A/M185A}$ a Superdex75 HR 26/60 run by a Pharmacia FPLC system (Stephan Urban, Heidelberg) with a flow rate of 2 ml/min was used.

The purified proteins were dialysed against 100 mM ammonium acetate pH 7.0, 5 mM DTT. Sample concentrations were 1 mM for 3D-heteronuclear and triple-resonance experiments and 130 μM for chemical shift perturbation in 100 mM ammonium acetate pH 7.0 and 5% D_2O, 0.02% sodium azide. Samples contained 2-10 mM TCEP in order to ensure that proteins are in the reduced state. The correct molecular weight of the proteins for NMR analysis was verified by mass spectrometry.

2.2.8. Additional proteins and peptides

The proteins GagΔp6, Gag$\Delta p6_{H12A}$, ΔMACANC, $\Delta MACANC_{H12A}$, $MACA_{151}$ and $MA_{129}CA_{146}$ were obtained as pure ammonium sulfate pellets from our collaborator Wesley I. Sundquist (University of Utah, Salt Lake City, USA).

Peptides were obtained as lyophilised TCA salts (PSL, Heidelberg) and either directly resuspended in buffer or dialysed prior to usage in 0.1 kD cutoff tubing (SpectraPor). The peptide concentration was estimated from A_{280} measurements using extinction coefficents calculated as described below. The molecular weight, the theoretical pI and the extinction coefficient for the used proteins is given in the appendix A.3.

2.2.9. Determination of protein concentration

Protein concentrations were estimated from A_{280} measurements using extinction coefficents (ε_{280}) calculated based on the sequence of the protein by the computer program PROTPARAM. The calculated extinction coefficient is only correct if all absorbing amino acids are accessible and not influenced by their surroundings. Although this is not the case in a folded protein, measurements under native conditions were reasonably accurate for the used proteins.

For peptides ε_{280} was calculated based on the contributions of absorbing amino acids present in the sequence with $\varepsilon_{280,\mathrm{Trp}} = 5500\,\mathrm{M^{-1}cm^{-1}}$, $\varepsilon_{280,\mathrm{Tyr}} = 1490\,\mathrm{M^{-1}cm^{-1}}$.

Extinction coefficients for all used proteins and peptides are shown in the appendix A.3.

2.2.10. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Denaturing polyacrylamide gel electrophoresis allows the separation of proteins according to their mass. SDS induces protein unfolding and coats proteins with negative charges, which allows for a separation of proteins independent of their inherent charge. In order to obtain a good resolution in a range of 15–45 kD for most applications 17.5% acrylamide gels with a ratio of acrylamide/bisacrylamide (AA/Bis) of 200:1 were used. For proteins smaller than 15 kD 17.5% acrylamide gels with a ratio of 29:1 (AA/Bis) were used. The following protocol yields two $10 \times 6\,\mathrm{cm}^2$ gels with a thickness of 0.75 mm:

	17.5% seperating gel		stacking gel	
	200:1	29:1	200:1	29:1
H_2O	1.5 ml	2.92 ml	3.4 ml	3.645 ml
1.5 M Tris/HCl pH 8.8	2.5 ml	2.5 ml	-	-
1 M Tris/HCl pH 6.8	-	-	630 µl	630 µl
10% SDS	100 µl	100 µl	50 µl	50 µl
AA / Bis	5.8 ml (30%)	4.375 ml (40%)	830 µl(30%)	625 µl (40%)
10% APS	100 µl	100 µl	50 µl	50 µl
TEMED	4 µl	4 µl	5 µl	5 µl

Prior to loading, samples were cooked for 5 min in sample buffer. The electrophoresis was performed in SDS buffer (Tris-glycine) at 28 mA per $10 \times 6\,\mathrm{cm}^2$. Gels were then either used for western blot analysis (section 2.2.11) or stained with Coomassie G (Coomassie Brilliant Blue G250, Serva): Gels were first incubated for 20 min in fixative then for at least 20 min in Coomassie G solution. Unspecific staining was removed with water.

2.2.11. Western blot analysis

After gel electrophoresis proteins can be transferred to nitrocellulose membranes for specific detection with antibodies. The gel and pieces of Whatman paper and nitrocellulose of the same size were soaked in blot transfer buffer. Whatman paper, the

nitrocellulose membrane, the gel and more Whatman paper were piled up on the anode. The transfer was performed for 1 h at 0.8 mA/cm^2. To visualise marker proteins the membrane was stained with PonceauS and afterwards washed with PBS for 10 min. To block unspecific binding sites, the membrane was incubated in 5% milkpowder in PBS for 1 h. After washing for 10 min with PBS/0.05% Tween membranes were incubated over night with antibodies (polyclonal rabbit anti-CA antiserum (RαCA) or mouse monoclonal antibodies) diluted in 1% milkpowder in PBS. After washing 3× for 10 min with PBS/0.05% Tween and 3× for 10 min with PBS, membranes were incubated for 75 min with secondary horse radish peroxidase (HRP) conjugated antibodies (goat anti-rabbit (GαR) or goat anti-mouse (GαM), Dianova). Washing steps were repeated and reactive bands were visualised by ECL detection: As substrate a 1:1 mixture of SuperSignal® West Pico Luminol/Enhancer solution and Stable Peroxide solution (PIERCE) was used. The resulting chemiluminescence of the substrate reaction was detected by autoradiography.

2.2.12. Monoclonal antibodies (mAbs)

Seven monoclonal antibodies (mABs) were produced by Klaus Wiegers, HPI Hamburg, by hybridoma-technology. Mice were immunised either with the recombinant protein ΔMACA (in the case of antibodies 1.5G10, 2.3D12, 2.3F12, 2.4E6, 2.4G3), baculovirus expressed Gag (3.1B5) or MA (1.1E10). One antibody (AK183) had been obtained from the NIH, USA. An αPolio mAb was used as control. 2.3D12 and 2.3F12 are IgM antibodies, the remaining are IgGs.

The concentration of mAbs was determined by absorption at 280 nm using the following relations: 1 mg/ml corresponds to 1.43 $OD_{280\,nm}$ for IgG mAbs and to 1.18 $OD_{280\,nm}$ for IgM mAbs (Current protocols in immunology, NIH).

2.2.13. Peptide scan technology

The synthesis of overlapping 15mer peptides covering the sequence of the target protein on activated membranes by chemistry was performed by Martin Blüthner, Insitute of Immunology and Molecular Genetics, Karlsruhe. For hydration, membranes were soaked in methanol for 2 min and then washed 3× for 10 min in PBS. A Ponceau staining was performed to visualise peptide spots. After washing with PBS to remove the stain, membranes were blocked in 5% milk powder in PBS for 1 h. After washing for

10 min in PBS/0.05% Tween, membranes were incubated over night with the primary antibody diluted in 1% milk powder in PBS (mAbs final concentration of 12.5 μg/ml). After washing 3× for 10 min with PBS/0.05% Tween and 3× for 10 min with PBS, membranes were incubated for 75 min with the secondary antibody diluted in 1% milk powder in PBS (for detection of bound mAb HRP-conjugated goat anti-mouse (GαM) antibody (Dianova) diluted 1:10,000). After washing as before with PBS/0.05% Tween and PBS, reactive spots were detected by ECL reaction as for western blot analysis.

2.2.14. ELISA assay to detect pH dependent conformational changes

ELISA plates (96well, Nunc Maxisorp) were coated over night with a polyclonal rabbit anti-CA (RαCA) antiserum (Prof. Kräusslich) at room temperature. In parallel, 200 μl 50 nM protein solutions were prepared in 100 mM NaCl, 0.05% Tween, 100 μg/ml BSA and 30 mM MES pH 6.0 or 30 mM Tris pH 8.0. 50 μg/ml antibody was diluted 1:10 to 1:10^4 in protein solutions containing Gag, ΔMACANCSP2 and $MACA_{151}$. 500 μg/ml antibody was diluted 1:10 to 1:10^4 in solutions containing $MA_{129}CA$ and CA derived proteins. These mixtures were also incubated over night at room temperature.

The coated plate was washed 3× with water. Plates were blocked with 10% FCS in PBS for 1 h at 37°C and then washed 3× with 100 mM NaCl, 0.05% Tween and 30 mM MES pH 6.0 or 30 mM Tris pH 8.0, respectively. Antibody-protein mixtures were incubated on plates for 3 h at room temperature and then washed as before. Plates were incubated with HRP-conjugated GαM antibody (Dianova, diluted 1:2000 in 100 mM NaCl, 0.05% Tween, 10% FCS and 30 mM MES pH 6.0 or 30 mM Tris pH 8.0, respectively) for 1 h at 37°C to detect antibody-protein complexes. After washing the plates as before, the ELISA was developed by addition of 100 μl per well of a mixture of 10 ml 0.1 M sodium acetate with 2 μl H_2O_2 and 100 μl 3,3',5,5' tetramethylbenzidine (TMB) as HRP substrate. Upon oxidation the absorption maximum of TMB shifts from 285 nm to 652 nm (blue colour). The reaction is stopped by addition of 0.5 M H_2SO_4 leading to a yellow colour that can be measured at 450 nm and allows for a higher sensitivity of the quantification than the blue colour.

2.3. *In vitro* assembly and electron microscopy

To assemble spherical particles, 40 μl of 10–25 μM ΔMACANCSP2 in 30 mM MES pH 6.0, 500 mM NaCl, 1 mM EDTA, 2 mM DTT was dialysed for 3–4 h in 10–14 kD cutoff dialysis tubing (SpectraPor) in the presence of 5% (w/w) 73mer oligonucleotide (the sequence is given in the appendix A.2) against 50 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 2 mM DTT. Assembled particles were pelleted by centrifugation for 5 min at 15,000 rpm and resuspended in dialysis buffer in a quarter of the original sample volume.

For EM analysis, 300 mesh copper grids (G300 HEXC3, SCI) were coated on glass slides with a thin film of Formvar (1595 Serva, 1% in chloroform). In order to achieve an optimal surface for adsobtion of hydrophobic as well as hydrophilic substances, the grids were coated with a layer of carbon after drying and were glow-discharged for 15 s prior to usage. 5μl of the resuspended pellets was then adsorbed onto prepared copper grids. After 5 min incubation, surplus solution sticking to the grids was carefully removed with Whatman paper (size 1). Grids were then incubated on a 5 μl drop of 2% uranyl acetate (MERCK), to visualise assembled particles by negative staining. After 5 min surplus solution was removed again and grids were dried on the air. Assembled particles were analysed by transmission electron microscopy using a Philips 400T, Philips BioTWIN, Morgagni or Zeiss EM10 microscope (EMBL Heidelberg).

Tubular particles were assembled either from CANC or from CA. 40 μl of 15μM CANC (stored in 50 mM Tris pH 7.5, 5μM $ZnCl_2$, 500 mM NaCl, 10% glycerol, 2 mM DTT) was assembled as described above. 40 μl of 100–130μM CA proteins (as well as the variants CASP1, CA220, CA207, CA177) stored in 30 mM MES pH 6.0, 500 mM NaCl, 1 mM EDTA, 2 mM DTT was dialysed against 50 mM Tris pH 8.0, 1 M NaCl, 1 mM EDTA, 2 mM DTT. Assembled particles were pelleted and treated for EM analysis as described above.

For assembly inhibition studies with monoclonal antibodies (mAb) 10μM ΔMACANCSP2 was preincubated for 30 min at 4°C with 5 μM of each mAb in a volume of 40 μl. To allow for antibody binding a pH of 7.5 was used. Nucleic acid was added after the preincubation and assembly reactions were performed as described above, except that they were started at pH 7.5.

To test peptides for inhibition of *in vitro* assembly, the system had to be adjusted for the usage of 0.1 kD dialysis tubing (SpectraPor). In the case of ΔMACANCSP2 the protein had to be purified by gel filtration (section 2.2.5) and for all proteins

dialysis had to be performed over night to achieve high efficiency of assembly. 25μM ΔMACANCSP2 or 15 μM CANC was preincubated without or with a 0.5–5 fold molar excess of peptide (dissolved in PBS) for 30 min at 4°C in a volume of 40 μl. After addition of nucleic acid assembly reactions were performed over night in 0.1 kD cutoff tubing and treated for EM analysis as described above.

To test peptides for inhibition of CA assembly additional adjustments were necessary. CA assembly is triggered by a pH change from 6 to 8 and an increase in the NaCl concentration from 0 to 1 M. The latter was found to be not compatible with dialysis in 0.1 kD cutoff tubing, since membranes were drained after 15 min dialysis. Adjusting the NaCl concentration to 0.5 M prior to dialysis allowed for particle assembly with decreased efficiency. In addition, in order to use a 10-fold molar excess of peptide over protein, peptides had to be dissolved in DMSO because of the high protein concentration needed for assembly. This led to a final concentration of 1.75% DMSO within the tubing, which also had a detrimental effect on the efficiency of the reaction as well as on the shape of the assembled particles. 117μM CA was preincubated with a 10-fold molar excess of peptide in a volume of 40 μl for 30 min at 4°C Apart from the changes mentioned here, assembly of CA was performed over night in 0.1 kD cutoff tubing and was treated for EM analysis as described above.

2.4. Phage display

The phage display technology was used here to screen a library of random peptides for ligands binding tightly and specifically to a target protein. The used libraries are based on the M13 phage. This filamentous phage with a length of 1 μm and a diameter of 9 nm has a circular ssDNA genome packaged within 2700 copies of the pVIII coat protein. Its head is built of five copies of the pIII coat protein, which by interaction with the F-pilus of bacteria confers its infection. Phages inject their single stranded DNA via the F-pilus into the bacterium, where it is converted into a dsDNA, the so-called replicative form, which is transmitted vertically in the bacterial life cycle. The release of progeny phages is a non-lytic process.

Generally, phage libraries as used for phage display present peptides at the N-terminus of either the pIII or pVIII coat protein. Higher stringency of the selection procedure is conferred by pIII-fusion proteins due to the lower copy number per phage. The libraries used here (Ph.D.-12™or Ph.D.-C7C™, New England Biolabs) consisted of M13 based phages presenting random linear 12mer peptides (2.7×10^9 sequences) or cysteine

constrained 7mer peptides (1.2×10^9 sequences) fused to the N-terminus of their pIII coat protein. The low complexity of the 12mer library (theoretically possible are 4×10^{15} sequences) is compensated by the fact, that a single 12mer sequence might cover several shorter epitopes. The C7C library has a high complexity (theoretically possible are 1.3×10^9 sequences) and the advantage, that some peptide folds might be stabilised by the formation of a loop structure. Folds made impossible by the constraint might still be included in the 12mer library. During generation of the random library the statistically equal incorporation of all amino acids was ensured by a genetic code reduced to 32 codons.

These libraries were used to select peptide ligands binding to HIV-1 CA or C-CANC and to map the epitope of antibody AK183. For the biopanning procedure the purified target proteins CA and C-CANC as well as antibody AK183 were immobilised on magnetic beads. Magnetic beads (tosyl acitvated Dynabeads® M-280, Dynal Biotech) were treated as recommended by the manufacturer using 12 μg protein or 5 μg anti mouse antibody per 10^7 beads. Beads coated with protein were ready to use after immobilisation, 20 μl beads coated with anti-mouse antibody were incubated for 2 h at room temperature on a rotator with 1 ml of 2 μg/ml AK183 or anti-Polio antibody in TBSG prior to use. Coated beads were stored at 4°C for the duration of the biopanning reaction.

2.4.1. Buffers

TBS:	50 mM Tris pH 7.5, 150 mM NaCl
TBSG:	50 mM Tris pH 7.5, 150 mM NaCl, 0.25% gelatin
TBSGT:	50 mM Tris pH 7.5, 150 mM NaCl, 0.25% gelatin, 0.1% Tween
Neutralisation buffer:	1 M Tris pH 9.1
Elution buffer:	0.2 M glycine/HCl pH 2.2, 1 $\frac{\text{mg}}{\text{ml}}$ BSA
PEG/NaCl:	25% PEG 6000, 2.5 M NaCl
Iodide buffer:	10 mM Tris pH 8.0, 1 mM EDTA, 4 M NaI → store dark

2.4.2. Biopanning reaction

To select for specific and tight binding phages a combination of positive and negative selections was performed. First, 20 μl Dynabeads coated with CA, C-CANC or AK183, were washed six times with 1 ml TBSGT and then incubated on a rotator over night at 4°C with 10 μl phage library in 1 ml TBSG. This corresponds to $\sim$ 55 copies of each

sequence present in either of the two libraries. After washing the beads 10× with 1 ml TBSGT, bound phages were eluted from the beads by a 10 min incubation at room temperature on a rotator with 600 µl 0.2 M glycine pH 2.2 with 1 mg/ml BSA. The eluates were neutralised with 90 µl 1 M Tris pH 9.1 and subsequently incubated over night with dynabeads coated with the negative targets BSA, NC or α Polio antibody. 25 ml of a 1 : 100 dilution of an overnight culture of *E. coli* ER2738 bacteria was then infected with 900 µl of the supernatant of the negative selection. After 4.5 h incubation at 37°C und 225 rpm, bacteria were pelleted for 10 min at 10,000 rpm and 4°C. 8 ml of PEG/NaCl was added to the supernatant and left over night at 4°C. Precipitated phages were pelleted by centrifugation for 30 min at 10,000 rpm and 4°C. Pellets were resuspended in 1 ml TBS. Bacterial debris was removed by a 2 min centrifugation at 15,000 rpm. Phages were again precipitated by addition of 333 µl PEG/NaCl and incubated for 1 h on ice. After centrifugation for 15 min at 15,000 rpm and 4°C phage pellets were resuspended in 200 µl TBS, 0.02% sodium azide. 150 µl of this pool of selected and amplified phages was then used for an additional round of positive and negative selection, followed by amplification and a third positive selection.

2.4.3. Determination of phage titers

The phage libraries are based on M13mp19 phages carrying the *lacZα* gene under the control of a lac promotor. Thus, phage plaques appear blue on IPTG/XGal plates, whereas wild type M13 phages yield white plaques. 10 µl of serial dilutions (10^{-4} to 10^{-7}) of the selected phage pool in TBS was mixed with 200 µl of an overnight culture of *E.coil* ER2738. After 5 min incubation at room temperature the infected cultures were mixed with 3 ml warm top agar and poured onto IPTG/XGal plates. After incubation over night at 37°C plaque forming units (pfu) were counted.

2.4.4. Test amplification of single phages for ELISA assay

Single phage plaques from IPTG/XGal plates with a total of about 100–150 plaques were transferred to 200 µl LB-medium in a 96 well plate. Phages were amplified over night at 30°C und 175 rpm. As negative control 2 µl of a M13 wild-type phage stock solution was amplified as well. In parallel, Nunc Maxisorp ELISA plates were coated over night with 2 µg/ml CA or BSA, C-CANC or NC and AK183 or α Polio antibody at room temperature. M13 phage was coated as positive control. After washing the ELISA plates 4× with 200 µl PBS/0.05% Tween, they were blocked with 100 µl 5%

milk powder in PBS/0.05% Tween for 4 h at room-temperature. To remove bacteria, 96well plates with the amplified phages were centrifuged for 7 min at 2500 rpm. 25 μl of the supernatant was transferred to an ELISA well and plates were incubated over night at 4°C. After washing, bound phages were detected with 100 μl HRP-conjugated anti-M13 monoclonal antibody (Amersham Biosciences) diluted 1 : 5000 in 5% milk powder in PBS/0.05% Tween for 2 h at room temperature. After washing the ELISA was developed as described above, section 2.2.14. Phages that displayed a fivefold better binding than the wild type M13 phage and background values for the reaction with the negative target were amplified for further usage and for DNA preparation and sequencing.

2.4.5. Phage purification, ssDNA preparation, sequencing

From phages positive in ELISA tests, 100 μl of the small scale amplification was used to infect 2 ml of a 1 : 100 dilution of an over night culture of *E.coli* ER2738. Infected bacteria were cultivated for 4.5 h under fast shaking (225 rpm) at 37°C. Subsequently, bacterial debris was removed by centrifugation for 5 min at 3500 rpm and 4°C. For purification, 1.5 ml of the supernatant was incubated over night at 4°C with 500 μl PEG/NaCl, for single stranded phage DNA (ssDNA) preparation 500 μl supernatant and 200 μl PEG/NaCl were used. Precipitated phages were pelleted for 15 min at 15,000 rpm. For storage pellets were resuspended in 200 μl TBS/0.02% sodium azide and glycerol was added to a final concentration of 50%. These stock solutions were stored at −20°C.

Phages precipitated for ssDNA preparation were pelleted the same way. Pellets were resuspended in 100 μl iodide buffer and mixed with 250 μl pure ethanol. After 10 min incubation at room temperature ssDNA was precipitated by centrifugation at 15,000 rpm, washed with 70% ehtanol and resuspended in 20 μl H_2O. The region of interest was sequenced with the 96gIII primer (New England Biolabs) supplied with the phage library mixing 3 μl ssDNA (163 μg/ml) with 5 μl 96gIII oligonucleotide (100 pmol/μl), 6 μl H_2O and 6 μl Mastermix (Beckman). The sequencing reaction was performed as described in section 2.1.8.

2.4.6. Phage and peptide ELISA assays

CA protein (2 μg/ml) was coated on 96well Nunc Maxisorp ELISA plates. Plates were washed with PBS/0.05% Tween and blocked with 5% milk powder in PBS/0.05% Tween. Purified phage presenting *CAI* (phage-*CAI*, 1.2 to 4.4×10^6 pfu/μl) was mixed with peptides at final concentrations of 50, 25, 5, 2.5 and 0 μg/ml. Plates were incubated over night with 50 μl of the mixture at 4°C. After washing, bound phages were detected with 100 μl HRP-conjugated αM13 monoclonal antibody (Amersham Biosciences), diluted 1 : 5,000 in 5% milk powder in PBS/0.05% Tween for 2 h at room temperature. After washing, the ELISA was developed following standard procedures. To test for competition of various untitered phages, the volume of phage solution corresponding to 1.2×10^7 pfu/μl phage-*CAI* was mixed with 51.9μM *CAI*.

In order to test phage-CAI binding to CA derived proteins, 96well Nunc Maxisorp ELISA plates were coated with 2 μg/ml of these proteins. Plates were washed and blocked as described above. Phage-*CAI* (2.4×10^6 pfu/μl) was added to the plate and incubated over night at 4°C. After washing, bound phages were detected as described above.

2.5. Cell culture experiments

293T and Hela cells were cultured in Dulbecco's modified Eagle medium containing 10% FCS and antibiotics. Cells were transfected with FuGENE 6 (Roche) according to the instructions of the manufacturer using 2 μg plasmid DNA per six-well. The plasmid used for these transfections was pNL4-3ΔPBSRT- (obtained from Eva Gottwein), a plasmid leading to the CMV-driven synthesis of full-length HIV-1 with a deletion of the primer binding site and an active-site mutation in the reverse transcriptase. This leads to release of non-infectious HIV-1 particles. The transfections were performed either in the absence or in the presence of 40 μg/ml peptide. Alternatively cells were transfected with the Chariot™ protein transfection kit (Active Motif) prior to plasmid transfection. Chariot transfections were performed according to the manufacturer's instructions with 250 or 500 ng *CAI* or *CAI*ctrl or with 0.5 or 1 μg *CAI*-C-FITC. A GFP expression plasmid was used to control for DNA transfection.

After 17–20 h the virus containing cell-culture supernatant was harvested and filtered through a 0.45 μm filter. Virus particles were pelleted through 20% (w/w) sucrose cushions by ultracentrifugation for 90 min at 39,000 rpm in a SW41 rotor. Pellets were

resuspended in 50 μl PBS. Cells were washed and lysed in 200 μl SDS-PAGE sample loading buffer. Lysed cells and particle preparations were analysed by western blotting as described in section 2.2.11 using a 1:2000 dilution of anti-CA polyclonal antiserum (RαCA) as first antibody.

Hela cells were microinjected in CO_2 independent medium with either plasmid DNA alone (10 ng/μl), a mixture of 10 ng/μl plasmid DNA and 1 mg/ml *CAI* or 1 mg/ml *CAI*-C-FITC. The plasmid used here was pNL4-3ΔenvGFP [48]. This plasmid has a deletion within the *env*-ORF and an inserted GFP gene within the *nef*-ORF. This leads to GFP expression and to the synthesis of non-infectious particles. Particle release in the presence and in the absence of *CAI* was quantified by RT-PCR. In order to remove plasmid DNA that leaked into the supernatant during microinjection DNAse digestion was performed prior to PCR analysis. The RT-PCR for the detection of virus particles per ml supernatant was performed by the virological diagnostics group using the COBAS® TaqMan HIV-1 Test (Roche).

2.6. Nuclear magnetic resonance (NMR) spectroscopy

In the last decades solution state NMR spectroscopy has evolved into a powerful method for the analysis of protein structures at the atomic level that is highly complementary to x-ray crystallography. Apart from structural analysis in the solution state, NMR is a versatile tool for the analysis of intramolecular dynamics and protein folding as well as of protein-ligand interactions. It allows the determination of equilibrium binding constants, localisation of binding sites and in favourable cases it allows inferences on binding kinetics. The inherently low sensitivity and high spectral complexity of NMR limits the molecular weight range of proteins that are suitable for full structure analysis to about 30–40 kD. However, recent advances in spectrometers and experimentation allow the analysis of much larger proteins.

The exact description of the principles of NMR requires a quantum mechanical analysis, which is beyond the scope of this introduction. Rather, this is supposed to briefly describe the NMR parameters and types of experiments used here. A detailed description of the applied methods can be found in current text books [20].

Here, NMR spectroscopy was to characterise a protein-ligand complex in terms of the ligand binding site and the affinity of the interaction as well as the oligomerisation state

of the complex. NMR is a suitable method to study protein-ligand interactions, since it is able to detect and localise with atomic resolution subtle changes in the electronic environment, which are induced for example upon ligand binding.

NMR is based on the fact that protons (^{1}H) present in proteins, or nitrogen (^{15}N) and carbon (^{13}C) nuclei, that can be incorporated into proteins by isotope labelling, are magnetically active (they have a spin). If these nuclei are placed in a magnetic field, their local electronic environment influences their exact resonance frequency, which is the frequency needed to induce the transition of the spin of a nucleus to a higher energy state. These differences in resonance frequency from a standard value are called chemical shift. The chemical shift of nitrogens and amide protons is very sensitive to changes in the local environment that are for example due to ligand binding. The $^{13}C\alpha$ shift of an amino acid depends in a characteristic manner on the local secondary structure.

Having induced the spin transition of nuclei to a higher energy state by applying a radio frequency pulse matching the respective resonance frequency, the return of the nuclei to the equilibrium state is recorded. The time constants that govern this return report the rotational motions of a macromolecule and can therefore be used to obtain quantitative information on protein dynamics with atomic resolution.

In order to achieve the atomic resolution, chemical shifts of individual nuclei within ideally all amino acids of a protein have to be identified (assignment). This assignment becomes possible by measuring correlations between nuclei. These are obtained by using either scalar (through-bond) or dipolar (through-space) couplings between the nuclei. In a homonuclear experiment for example protons within the spin system of each amino acid are correlated (COSY and TOCSY experiments). Heteronuclear single quantum coherences (HSQC) are typically used to correlate protons with nitrogens or carbons as in ^{1}H-^{15}N-HSQC and triple resonance experiments. The dipolar coupling (through-space) is used in NOESY experiments and its characteristic dependence on the internuclear distance yields essential informations for structure determination.

Multi-dimensional NMR experiments consist of a sequence of radio frequency pulses and delays designed to achieve a specific correlation between two or more nuclei. While the precise pulse sequence may be rather complicated, these experiments are all based on the elements described above.

2.6.1. ^{1}H-^{15}N-heteronulear single quantum coherence (HSQC) spectroscopy

The ^{1}H-^{15}N-HSQC spectrum uses the scalar (through-bond) coupling between the amide proton and the isotope labelled nitrogen nucleus to correlate the chemical shifts of these nuclei. Hence, in principle, this spectrum gives rise to a single resonance for each amide and nitrogen nucleus present in the protein. At the conditions used here, every residue except proline appears once and some side-chain NH groups are observed as well. Thus, these spectra provide a characteristic fingerprint of a protein. The dispersion of the signals is a good indicator of the folding state of the protein. Excess number of resonances indicate conformational heterogeneity while fewer than expected resonances indicate either fast solvent exchange (usually highly flexible and exposed residues) or additional line-broadening due to for example oligomerisation.

2.6.2. Total correlation spectroscopy (TOCSY)

A ^{1}H-^{1}H-TOCSY pulse sequence transfers magnetisation through a proton spin system via scalar coupling, giving rise to a peak pattern characteristic for certain classes of amino acids. A ^{1}H-^{1}H-TOCSY and a ^{1}H-^{15}N-HSQC spectrum can be combined into a three dimensional (3D) spectrum. Since peaks in the ^{1}H-^{1}H plane are now spread in the ^{15}N chemical shift dimension, the spectral overlap is significantly reduced and the spectrum is simplified.

2.6.3. Nuclear Overhauser effect spectroscopy (NOESY)

A ^{1}H-^{1}H-NOESY experiments uses of the dipolar coupling (through-space) to transfer magnetisation between protons resulting in additional peaks (nuclear Overhauser effects, NOEs) depending on the interproton distance. In proteins the range of observable distances is typically limited to about 5 Å. The combination of a ^{1}H-^{1}H-NOESY with a ^{1}H-^{15}N-HSQC spectrum yields the same 3D spectrum as for the TOCSY experiment, but for a given ^{1}H-^{15}N resonance, addtional peaks are observed resulting from protons outside the spin system within a distance of 5 Å.

Using the above described combinations of HSQC, TOCSY and NOESY spectra the resonances of adjacent or nearby amino acids can be assigned to specific protons in the protein. Thus, these spectra contain information about the primary structure of the

protein and the NOESY spectrum additionally contains informations on the distance between different nuclei.

2.6.4. Triple resonance experiments

If a ^{13}C-^{15}N double labelled protein is available, triple resonance experiments can be performed to simplify the sequence resonance assignment. Here, HNCA, HNCACB and CBCA(CO)NH spectra were performed and the magnetisation transfer occurring in these experiments is depicted schematically in Fig. 2.1. In the HNCA experiment magnetisation is transferred from the amide proton to the nitrogen, then to the Cα nucleus and back to the amid proton for detection. This transfer occurs via scalar coupling. The coupling constant for the transfer from N to Cα within the same amino acid is very similar to that to the Cα of the previous amino acid. Thus, magnetisation is transferred to both Cα atoms, which leads to two Cα peaks at the resonance frequency of a given ^{15}N-H^{N} peak. In principle, this method allows the sequential assignment of the backbone ^{13}Cα, ^{15}N and H^{N} resonances.

The principle of the HNCACB spectrum is the same as for the HNCA spectrum, but magnetisation is additionally transferred to the Cβ, leading to altogether four peaks at a given ^{15}N-H^{N} resonance frequency, the Cα and Cβ for the same as well as the preceding amino acid. The additonal Cβ peak greatly helps in backbone assignment in

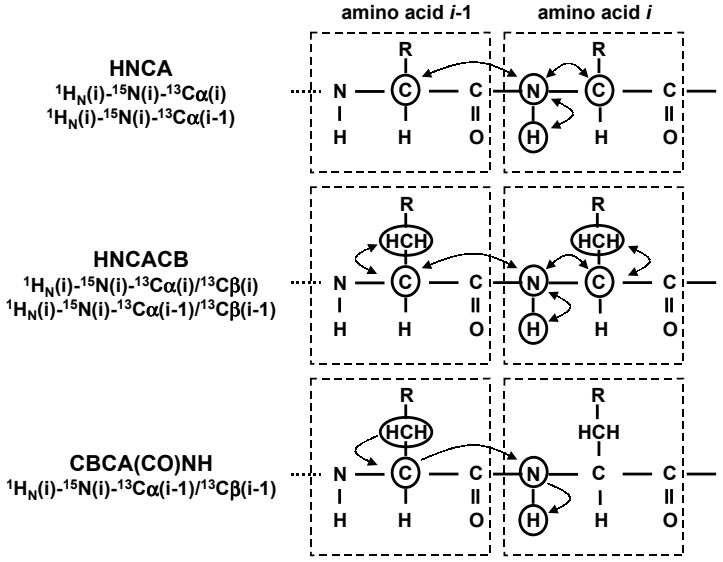

Figure 2.1.: Magnetisation transfers of HNCA, HNCACB and CBCA(CO)NH experiments.

cases of degenerate Cα peaks in HNCA spectra. Delineation of the same and preceeding residues can be achieved by using the CBCA(CO)NH experiment. Here, magnetisation is transferred from the Hβ and Hα protons to the Cβ and Cα nuclei. The chemical shifts of the Cβ and Cα nuclei are recorded, magnetization is then transferred via the carbonyl carbon (CO) to the nitrogen, where the ^{15}N chemical shift is recorded and finally to the amide proton (H^N) for detection. Hence, this experiment correlates resonances of the Cβ and Cα nuclei of the previous amino acid with the ^{15}N-H^N resonances of the residue itself.

All five spectra mentioned here complement each other and can lead to a reliable assignment of the backbone NH as well as the Cα and Cβ carbon resonances.

2.6.5. Protein secondary structure analysis using ^{13}C shifts

The above described triple resonance spectra also include information about the secondary structure of the protein. Cα chemical shifts (δCα) are characteristic for the local secondary structure [86]. Random coil (rc) δCα are derived from peptides measured in urea. Subtraction of these rc-δCα from the observed δCα values of a structured protein yields the secondary chemical shift. This value correlates with the secondary structure of a protein. Large positive secondary shifts typically greater than 2 ppm are characteristic for α-helical regions and negative shifts are observed for extended structures such as β-sheets.

2.6.6. Protein-ligand interaction studies

Chemical shifts of backbone NH nuclei are very sensitive to their local environment. Thus, ^{1}H-^{15}N-HSQC spectra are particularly well suited for the analysis of protein-ligand interactions. Upon titration of a protein with a ligand, chemical shifts of peaks in HSQC-spectra change, when the corresponding amino acid is in proximity to the bound ligand. These changes contain a wealth of information on the binding site as well as on the kinetics of the interaction. Thus, with the complete backbone assignment at hand, the ligand binding site can be mapped onto the primary sequence of a protein. In case of structural rearrangements, however, chemical shift changes may reflect both ligand binding and structural changes and the pattern becomes more complex. Chemical shift perturbations ($\Delta\delta$) observed for proton and nitrogen atoms of an amide bond can be summarised in a single weighted value using the following formula. The scaling factor

for the ^{15}N nucleus is used to approximately account for the overall differences in chemical shift perturbation between the ^{1}H and ^{15}N nuclei. Plotting the chemical shift perturbation for each assigned amino acid allows the localisation of the binding site.

$$\Delta\delta_{\mathrm{H,N}} = \sqrt{(\Delta\delta_{\mathrm{H}})^2 + \frac{1}{6}(\Delta\delta_{\mathrm{N}})^2} \tag{2.1}$$

The kinetics of the binding process affect the appearance of the NMR spectra upon titration with a ligand. Assuming the absence of intermediate states, the chemical shift difference ($\Delta\delta$) between the free and bound form of the observed nucleus can be correlated to the off-rate of the complex. For example a $\Delta\delta$ of 1 ppm in the proton dimension measured at 600 MHz corresponds to a time scale of 600 Hz. Broadly three time scales may be differentiated: Slow exchange is observed if the off-rate of the ligand is much slower than $\Delta\delta$. Upon ligand titration resonances of the free state gradually disappear while resonances of the bound state simultaneously appear. In this case, peak intensities correspond to the population of free and bound form. The other extreme is the fast exchange regime (the off-rate is much faster than $\Delta\delta$) leading to peaks gradually moving from the resonance of the free to that of the bound state. Here, the resonance position reflects the average of bound and free population. Inbetween these two extremes intermediate exchange is observed. An off-rate similar to $\Delta\delta$ leads to broadening of the resonances to the extend that they cannot be detected.

In the case of fast exchange the binding constant (K_{D}) of the interaction can be determined by fitting the K_{D} to the fractional shift ($\Delta/\Delta_{\mathrm{max}}$) according to the following formula, where Δ is the difference of the measured shift at any (maximal, Δ_{max}) ligand concentration and the shift of the free protein, and [L] and [P] are the ligand (CAI) and protein concentrations, respectively:

$$\frac{\Delta}{\Delta_{\mathrm{max}}} = \frac{(K_{\mathrm{D}} + [L] + [P]) - \sqrt{(K_{\mathrm{D}} + [L] + [P])^2 - 4[P][L]}}{2[P]} \tag{2.2}$$

2.6.7. Protein dynamics

The dynamics of a nucleus in a macromolecule can be inferred from the time constants that govern the return of its NMR signal to the equilibrium state. In a simple case, this process is exponential and can be described by empirical T_1 (longitudinal or spin-lattice relaxation) and T_2 (transverse relaxation) time constants.

T_1 and T_2 relaxation is induced by random fluctuations of the local magnetic field. In macromolecules the fluctuations are dominated by the random rotation of the molecules and can be characterised by a correlation time τ_c. For a well structured protein the random rotations of the backbone NHs can be approximated by the overall rotational correlation time of the protein (τ_m) plus additional fast internal fluctuations. These are characterised by an internal correlation time and an order parameter S^2 [66]. A S^2 of 1 means that there is no internal motion, while S^2=0 may indicate full rotational freedom. T_1 and T_2 values are dependent on τ_m and for a globular protein τ_m increases with the molecular weight. Accordingly, the average τ_m and thus the molecular weight and the oligomerisation state of a protein can be determined from the analysis of T_1 and T_2 values.

It is convenient to represent T_1 and T_2 values graphically, by plotting T_1 values for every amino acid against the corresponding T_2 values. These experimental values can be compared to the theory by adding T_1 and T_2 values as a function of correlation time and order parameter according to the above mentioned model free approach by Lipari and Szabo [66]. If the rotation of a ^{15}N nucleus is governed by overall tumbling of the molecule, the corresponding peak should be located within an order parameter of 1 to 0.8. If peaks of most residues are located within this region the protein can be assumed to tumble as a whole and T_1 and T_2 values of these residues are suitable for the determination of τ_m. Residues giving rise to peaks in a region of small S^2 can be assumed to be very flexible and rotate on their own time scale. Peaks shifted to smaller T_2 values outside the Lipari-Szabo model are affected by other slow relaxation processes such as oligomerisation or ligand binding.

2.6.8. NMR experimental setup

NMR spectra of ^{15}N labelled C-CA and ^{15}N/^{13}C C-CA$_{W184A/M185A}$ were collected at 25°C on a 600 MHz four-channel Varian INOVA NMR spectrometer equipped with a room temperature 5 mm triple-resonance z-gradient probe (in collaboration with Dr. Jörn Werner, University of Southampton). Resonance assignments of the backbone ^{15}N and HN nuclei of C-CA$_{W184A/M185A}$ in the free form and in complex with a fourfold molar excess of *CAI* were obtained by a combination of ^{15}N-edited and triple resonance spectra.

In addition to the 2D ^{1}H-^{15}N HSQC spectra [8, 6] 3D ^{1}H-^{15}N NOESY-HSQC [56] and ^{1}H-^{15}N TOCSY-HSQC [29] spectra were recorded with mixing times of 120 ms and

50 ms (DIPSI-3, 9000 kHz), respectively. ^{13}C decoupling was achieved with adiabatic inversion pulses [84]. In the HNCA [55] experiments the ^{13}C carrier was placed in the middle of the Cα region at 57 ppm, while in the CBCACONH [46] and HNCACB [98] sequences the carrier was at ^{13}C 40 ppm. ^{13}Cα and ^{13}C' pulses were generated as off-resonance amplitude and phase-modulated pulses adjusting the RF field strength of the ^{13}Cα/β and ^{13}Cα pulses to give minimal excitation in the ^{13}C' region, and *vice versa*. In the triple resonance experiments the ^{1}H carrier was placed at the H_2O frequency throughout and protons were decoupled using a 7196 Hz WALTZ-16 field. During acquisition ^{15}N was decoupled by using a 1295 Hz GARP1_7 sequence. All experiments were carried out with Gaussian water-flipback pulses [47] and gradient-enhancement was used for the ^{15}N-^{1}H coherences [54]. States-TPPI [69] was applied in the indirectly detected dimensions of all experiments to give quadrature detection. Acquisition times are listed in the Table 2.1.

Table 2.1.: Acquisition times of experiments used for assignment.

	^{1}H [ms]	^{15}N [ms]	^{13}C [ms]	HN [ms]
$CCA_{W184A/M185A}$				
^{1}H-^{15}N HSQC	-	43.5	-	127.8
3D ^{1}H-^{15}N NOESY-HSQC	20	16,3	-	51.2
3D ^{1}H-^{15}N TOCSY-HSQC	16	16.3	-	51.2
HNCA	-	24.5	14.1	51.2
HNCACB	-	24.5	7.9	51.2
CBCACONH	-	24.5	5.3	51.2
$CCA_{W184A/M185A}/CAI$				
^{1}H-^{15}N HSQC	-	65.3	-	51.2
HNCA	-	24.5	14.1	51.2
HNCACB	-	24.5	7.9	51.2

Protein samples were ensured to be in the reduced form by analysis of cysteine Cα and Cβ chemical shifts (literature values for the reduced cysteine 58.63 (Cα) and 28.34 (Cβ) and for the disulfide 55.63 (Cα) and 41.42 (Cβ)). For experiments with peptides, the lyophhilised TCA salts were either resuspended directly in protein solution or were dialysed at high concentrations against the same buffer prior to mixing with protein solution. Generally, protein concentrations were kept constant during titrations while ligand concentrations were increased. The localisation of the peptide binding site was obtained by recording a series of eight 2D ^{1}H-^{15}N-HSQC spectra with acquisition times of 128 (^{1}H) and 65 (^{15}N) ms at a concentration ratio of peptide:protein of 0:1, 0.25:1,

0.5:1, 1:1, 1.5:1, 2:1, 3:1, 4:1, 7.3:1. Rotational correlation times of ^{15}N C-CA were determined at 390 μM protein concentration without or with an equimolar amount of *CAI* from the ratio of the average ^{15}N-T_1 and ^{15}N-T_2 relaxation time constants. The T_1 and T_2 values for non-overlapping peaks were obtained by exponential fits to the measured intensity decays for all non-overlapping peaks of a series of 2D ^{1}H-^{15}N-T_1,T_2 measurements with delay times of 0.02, 0.08, 0.16, 0.32, 0.6, 1.0, 1.4 and 0.01, 0.03, 0.05, 0.07, 0.11, 0.15 s for T_1 and T_2, respectively.

NMR data were processed with NMRPipe [27], applying mild resolution enhancement and linear prediction in the heteronuclear dimension prior to fourier transformation and were analysed with NMRView [52].

2.7. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a method that allows the direct determination of several thermodynamic parameters that characterise a biomolecular interaction. In addition to the binding affinity (K_D) of an interaction, changes in enthalpy (ΔH) and entropy (ΔS) and the stoichiometry of the interaction (n) can be obtained. The method is based on the measurement of heats of interaction evolving upon titration of a protein solution with a highly concentrated ligand solution. Protein and ligand should be thoroughly dialysed against the same buffer and degassed immediately before starting the titration to minimise heats associated with solvent dilution, changes in buffer conditions or air bubbles.

A highly concentrated ligand solution is added step-wise from a syringe to a sample cell containing the protein solution under constant stirring. At each injection the heating energy (in μcal/s) that is needed to keep the temperature of the sample cell the same as that of a reference cell is measured. This accounts for heat released or absorbed by the system upon ligand addition and this is directly proportional to enthalpy changes (ΔH) in the system. Since these changes do not exclusively reflect the binding reaction, raw data have to be corrected for heats due to dilution of the ligand into the comparatively large volume of protein solution. The heats of dilution may be measured in a separate experiment or may be retrieved from heat release after the saturation of an interaction. Delineation of enthalpy changes due to the binding reaction from concomitant changes such as conformational changes, oligomerisation equilibria, protonation and dislufide formation require several experiments and in some cases may not be obtainable.

Ideally concentrations of protein and ligand are adjusted such that for the first steps of the titration all added ligand is bound by the protein, and that binding sites are only saturated later in the experiment. Finally, excess peptide is added that does not interact with the protein and gives rise only to heats of dilution. The measured heat release (ΔQ_i) per injection can be determined by integration over the peak area after correction for heats of dilution. By fitting the isotherm [97, 63] that relates ΔQ_i and the reaction enthalpy ΔH dependent on protein, ligand and complex concentrations at each (i-th) injection considering corrections for concentration change during the experiment, ΔH, K_{D} and the stoichiometry (n) of the interaction may be derived.

2.7.1. ITC experimental setup

A VP-ITC calorimeter (MicroCal Inc., Northampton, MA, USA) was used. Protein and peptide samples were dialysed against 100 mM ammonium acetate pH 7, 5 mM DTT and degassed prior to titration and all experiments were performed at 25°C. A solution of 2 mM *CAI* was titrated into the sample cell containing 1.39 ml of a 130 μM C-CA$_{\mathrm{W184A/M185A}}$ solution. 46 injections of 2 μl followed by 29 injections of 8 μl were performed with a spacing time of 240 s and a stirring speed of 310 rotations per second. The heat released was quantified by integration over the area of each titration peak. For data analysis and curve fitting the Origin 5.0 software package supplied by MicroCal was used. Correction for heat of dilution was achieved by subtraction of the heat realased after saturation of the complex. The reverse experiment was performed by titrating a 2 mM C-CA$_{\mathrm{W184A/M185A}}$ solution into 130 μM *CAI* using the same experimental parameters.

2.8. Computer programs

The extinction coefficient, theoretical pI value and molecular weight of proteins was calculated using the program PROTPARAM of the swissprot database (http://www.expasy.org/tools/protparam.html). The homology of peptide sequences was analysed using the ClustalW multiple alignment software (http://www.ebi.ac.uk/clustalw/index.html). Theoretical helicity of peptides was determined using AGADIR (http://www.embl-heidelberg.de/services). Protein structures were visualised using MOLMOL [58]. To write this thesis, the type setting program LaTeX was used.

3. Results

3.1. *In vitro* assembly of HIV-1 Gag-derived proteins

During assembly of HIV-1, Gag proteins accumulate at the plasma membrane and initiate a budding profile. The virus is released as an immature, non-infectious particle containing a spherical shell of Gag polyproteins underneath the viral membrane. The formation of the mature, infectious virus is initiated by processing of Gag proteins by the viral protease. Five sequential cleavages of Gag lead to the separation of the Matrix (MA), Capsid (CA), Nucleocapsid (NC), and p6 proteins, thereby also releasing the two spacer peptides SP1 and SP2. Complete maturation is only possible after the final processing step, which separates CA from the C-terminally adjacent 14 amino acid spacer peptide 1 (SP1). This process leads to the formation of the characteristic conical core, which is essential for viral infectivity.

In order to study assembly and maturation of retroviruses, *in vitro* systems have been developed over the last years [16, 44, 43, 15, 45, 14]. As had been established in our lab, assembly of recombinantly expressed and purified Gag-derived proteins can be triggered by changes in ionic strength and pH during protein dialysis. Assembled particles can then be analysed by electron microscopy (EM). Depending on the protein variant, spherical, immature-like as well as tubular, mature-like particles can be formed. The HIV-1 Gag-derived protein ΔMACANCSP2, which lacks p6 and amino acids 16–99 of MA, was shown to form spherical particles *in vitro* that are very similar to immature viruses [45, 96]. Tubular particles that resemble the mature conical core can be assembled *in vitro* of HIV-1 CA [44, 13] and the C-terminally elongated protein CANC [44]. The proteins ΔMACANCSP2, CANC and CA as well as variants used here for *in vitro* assembly are depicted schematically in Fig. 3.1 A.

Here, the previously established *in vitro* system was used as tool to test peptides and antibodies for assembly inhibition as will be described in the following sections. The above mentioned proteins ΔMACANCSP2, CA and CANC were expressed in

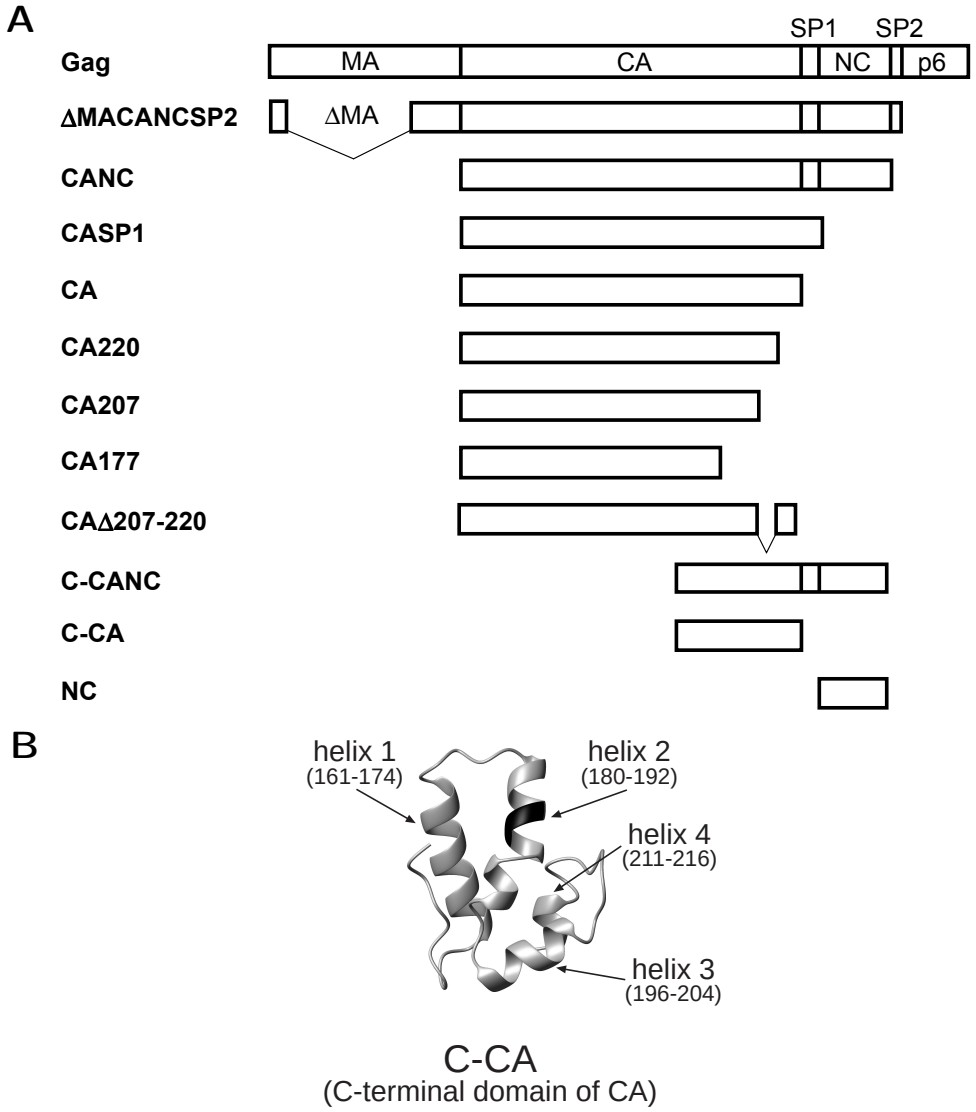

Figure 3.1.: A) Schematic representation of Gag-derived proteins used in this and the next section. B) Ribbon diagram of the structure of the C-terminal domain of CA (pdb ID: 1A8O [37]). CA truncations were made according to the borders of the helices as indicated. The dimer interface amino acids W184 and M185 are coloured black. The structure was generated with MOLMOL [58] and the orientation of the structure was chosen to best visualise the dimer interface.

E. coli and purified as described [44, 45] to near homogeneity. As an example, the SDS-PAGE analysis of the purification procedure of the CA protein is shown in Fig. 3.2. CA expression was induced in *E. coli* for not more than 2 h (lanes 1 and 2). After cell lysis by sonification (lane 3), protein remaining in solution after ultracentrifugation (lane 4) was precipitated with ammonium sulfate (lane 5–7).

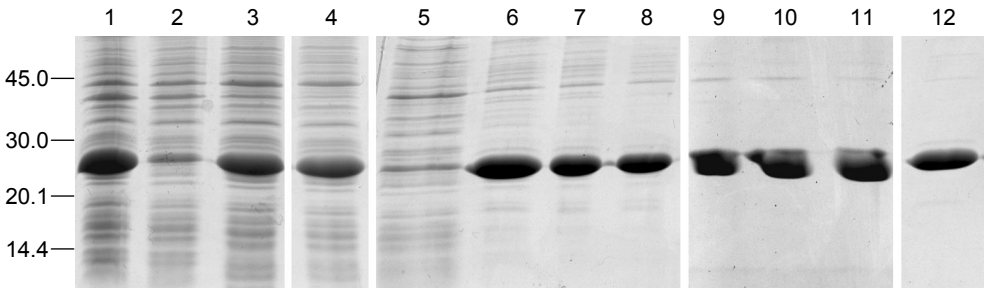

Figure 3.2.: Coomassie stained SDS-PAGE representing the purification of recombinantly expressed CA protein. Shown are lysates of induced (lane 1) and non-induced (lane 2) bacteria, the supernatant after sonification (lane 3) and subsequent ultracentrifugation (lane 4), supernatant (lane 5) and pellet (lane 6) as well as the resuspended pellet (lane 7) after 30% ammonium sulfate precipitation, the supernatant after anion exchange batch procedure (lane 8), peak fractions from cation exchange chromatography (lanes 9–11) and the pure protein after a final 60% ammonium sulfate precipitation step (lane 12). A molecular weight marker (in kD) is depicted on the left.

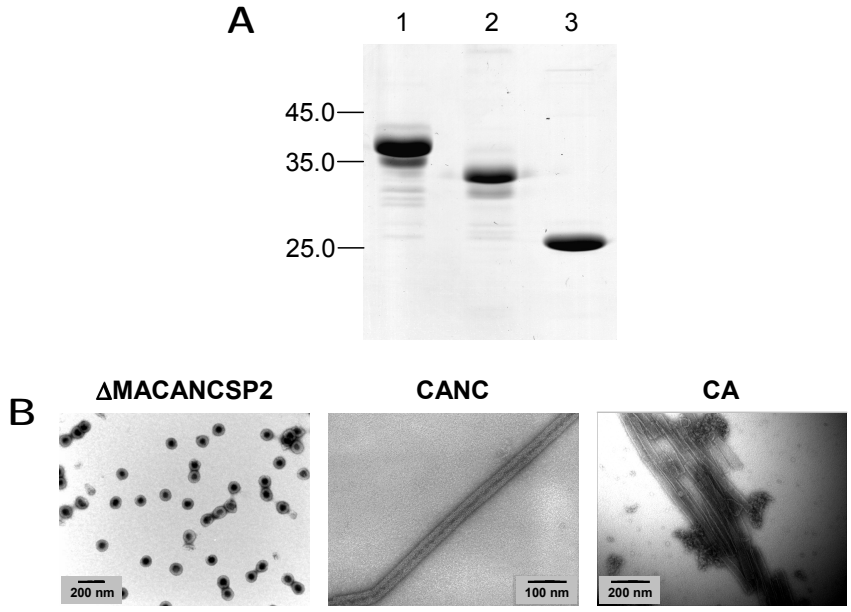

Figure 3.3.: A) Coomassie stained SDS-PAGE of 3μg of purified recombinant proteins ΔMACANCSP2, CANC and CA. A molecular weight marker (in kD) is depicted on the left. B) Electron micrographs of negatively stained particles resulting from *in vitro* assembly reactions: 25 μM ΔMACANCSP2 (46,000-fold magnification), 15 μM CANC (80,000-fold magnification), 117 μM CA (65,000-fold magnification).

In order to remove nucleic acids, an anion exchange batch procedure was performed (lane 8). The protein was further purified by cation exchange chromatography and peak fractions (lanes 9–11) were pooled and precipitated with ammonium sulfate (lane 12).

All three purified proteins that I used for *in vitro* assembly reactions throughout my PhD thesis are shown on a Coomassie stained polyacrylamide gel in Fig. 3.3 A. For assembly of immature-like, spherical particles the protein ΔMACANCSP2 was dialysed as described in the methods 2.3 [45]. An electron micrograph of negatively stained *in vitro* assembled spherical particles is shown in the left panel of Fig. 3.3 B. For the analysis of tubular, mature-like particles CANC was assembled using the same conditions as for ΔMACANCSP2 and assembly of CA was induced as described in the methods 2.3 [44]. Electron micrographs of tubular particles assembled of CANC and CA are shown in Fig. 3.3 B.

3.1.1. Analysis of the influence of the CA-SP1 region on assembly

Deletion of SP1 had been shown to abolish the formation of spherical immature particles *in vivo* [59] and *in vitro* [45]. Additionally, the mature conical core can only form after SP1 had been released from CA [93]. However, SP1 had been shown to be structurally disordered [100, 78] so that its function cannot be derived from structural analysis. Having established the *in vitro* assembly system, the influence of the CA-SP1 region on assembly was characterised.

The CASP1 protein, the full-length CA protein (231 amino acids) as well as C-terminally truncated variants (CA220, CA207, CA177) were expressed and purified as described above [44]. All proteins are depicted schematically in Fig. 3.1 A. CA220 lacks the last eleven flexible amino acids, CA207 lacks helix 4 of the C-terminal domain of CA (C-CA) and in CA177 only the first helix of C-CA is left. As can be derived from Fig. 3.1 B, the truncations of the CA protein were performed in order to theoretically preserve the structural integrity of the remaining helices.

In vitro assembly of all proteins was performed and the shape of the formed particles was analysed by electron microscopy. Pictures of negatively stained particles are shown in Fig. 3.4 A. In comparison to tubes formed by CA, CASP1 particles were slightly thicker, shorter and rounder, but not generally more conical, as had been described before [44]. In contrast to that, tubular particles formed by CA220 were slightly thinner,

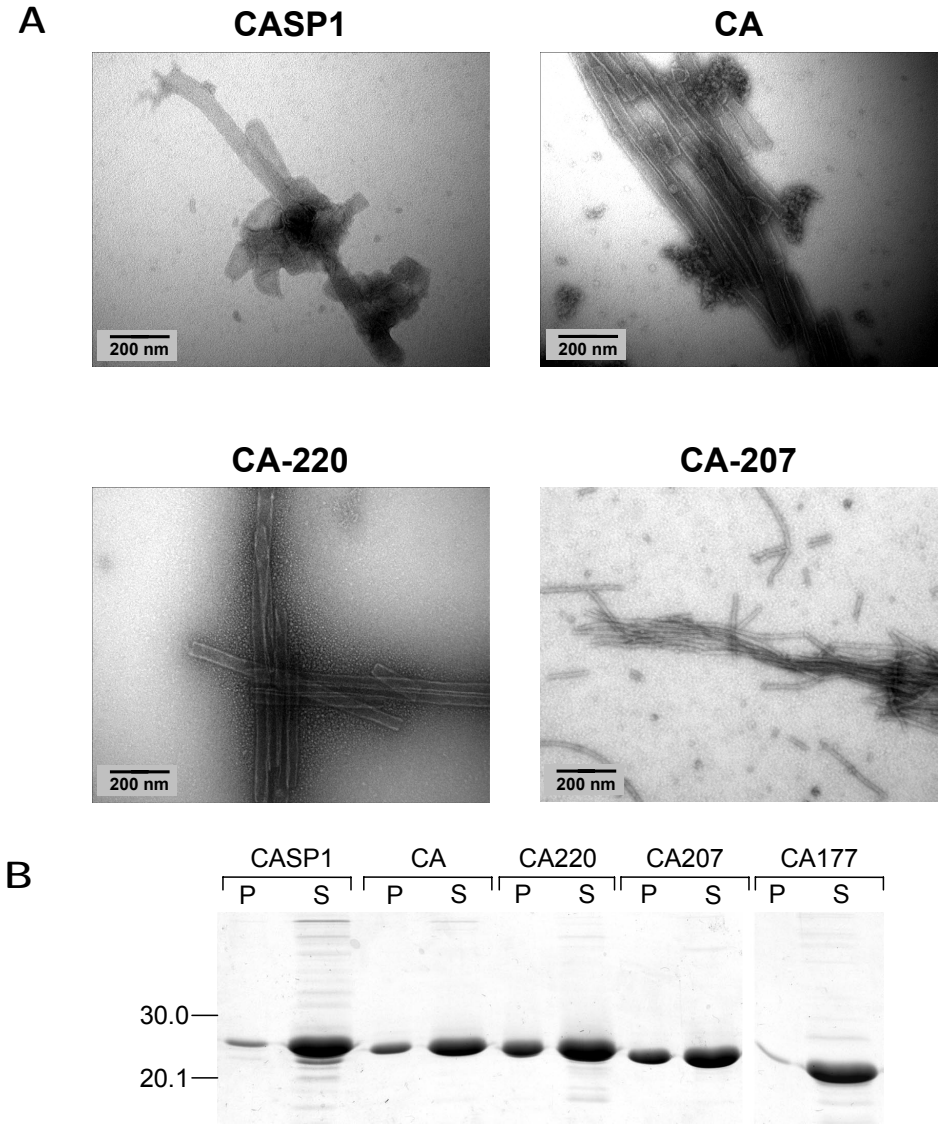

Figure 3.4.: A) Electron micrographs of negatively stained particles resulting from *in vitro* assembly reactions of CASP1, CA, CA220 and CA207 at 60,000-fold magnification. *In vitro* assembly was performed at 3 mg/ml concentration and particles were pelleted prior to adsorption on grids. B) Coomassie stained SDS-PAGE of pellet (P) and supernatant (S) fractions of *in vitro* assembly reactions of CASP1, CA, CA-220, CA207 and CA177 after centrifugation. A molecular weight marker (in kD) is depicted on the left.

but their shape was significantly more homogenous than that of CA tubes. Particles assembled of CA207 displayed only a diameter of about 10 nm. CA tubes were suggested to be built of helical arrangements of CA hexamers [65]. A tubular structure of only 10 nm diameter would not accommodate such a structural organisation. Thus, CA207 particles can be assumed to be helices rather than tubes. This suggested that either an important assembly interface is lost in CA207 or the structure of the protein is significantly altered upon deletion of helix 4 of C-CA. As expected for a protein lacking the dimer interface (amino acids 184/185), CA177 did not form any particles (data not shown). These results indicated that removal of flexible amino acids such as SP1 and the C-terminus of CA increased the regularity of *in vitro* assembled particles.

The efficiency of particle formation was difficult to judge by EM analysis, since tubular particles have the tendency to aggregate and therefore are distributed irregularly over the EM grid. However, judging from an overview over the whole grid the assembly efficiency increased for shorter proteins, but was abolished for CA177. In order to compare the assembly efficiency of the different proteins, aliquots of pellet and supernatant fractions were analysed by SDS-PAGE after centrifugation. A picture of the Coomassie stained gel is shown in Fig. 3.4 B. Comparing pellet (P) and supernatant (S) fractions for all five analysed proteins, most of the protein was found in the supernatant. This indicated that only a fraction of the protein assembled into particles. The pellet fraction can be assumed to correspond to the amount of assembled particles. In the case of CA177, practically no protein was found in the pellet consistent with the fact that no particles were observed by EM. For CASP1 only a small portion of the protein was found in the pellet fraction while increasingly more protein is found in the pellet for CA, CA220 and CA207.

Thus, removal of the last three helices of CA including the dimer interface abolished assembly, removal of helix 4 of C-CA abolished the ability to form tubular particles and removal of SP1 and the C-terminal amino acids of CA increased the efficiency of CA assembly *in vitro*.

3.2. A novel peptide inhibitor of HIV-1 assembly

HIV assembly and maturation present suitable targets for the development of antiviral substances, since these processes are driven by multiple weak protein-protein interactions. A molecule binding specifically to an interface that is relevant for assembly

is likely to be a good inhibitor of HIV replication and could thus serve as a basis for anti-retroviral drug design.

In order to discover a potential inhibitor targeting immature as well as mature interfaces, I performed a phage display screen to select for peptides that bind tightly and specifically to Gag-derived proteins. The selected peptides were tested for assembly inhibition *in vitro*.

3.2.1. Phage display: target proteins, biopanning, selected phages

Phage display is a method that allows to screen a library of peptides for their interaction with a target protein. Here a library of M13 derived phages presenting random 12mer peptides was screened for tight and specific binding to Gag-derived proteins. The rationale for this approach was the assumption that a peptide that interacts with Gag might have the potential to inhibit assembly of HIV.

The phage display biopanning reaction was performed as presented schematically in Fig. 3.5: An immobilised target protein was incubated with the phage library. Unbound phages were washed away, bound phages were eluted and amplified. In order to counterselect for unspecific binding, the phage pool was then incubated with an immobilised negative target and the unbound fraction was used as input for the next round of selection.

This biopanning reaction was performed on two different Gag-derived target proteins. The full-length CA protein (Fig. 3.1 A) was chosen, since it is the major determinant of particle assembly. The C-CANC protein (Fig. 3.1 A) was used because the region spanning the CA-SP1 junction was previously shown to be essential for assembly of immature particles [45]. For the C-CANC reaction the NC protein was chosen for negative selection to avoid a preferential selection of peptides targeting the nucleic acid binding domain. In the reaction with CA, BSA was used to counterselect for phages binding unspecifically to proteins in general. This strategy was used to obtain peptides that would bind specifically and with high affinity to the mature HIV-1 capsid protein (CA) and to a supposedly immature CA conformation in the context of C-CANC. All used proteins are depicted schematically in (Fig. 3.1 A).

The CA protein was purified as described in section 3.1. C-CANC was expressed in *E. coli* and purified by a combination of anion exchange and heparin affinity chromato-

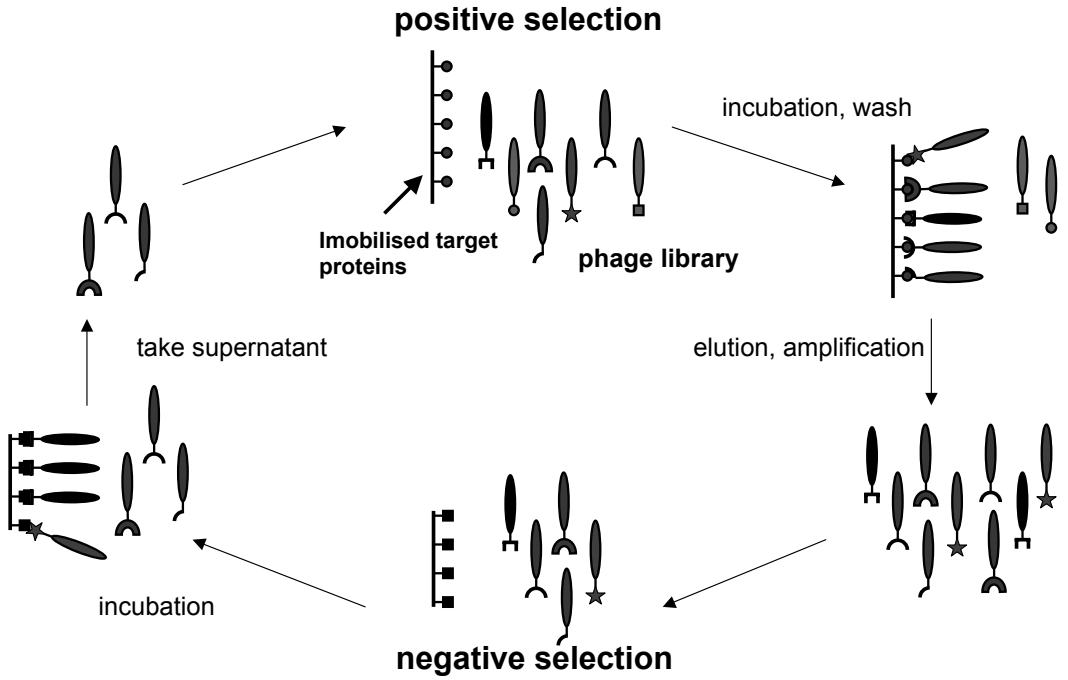

Figure 3.5.: Schematic representation of the phage display biopanning reaction that was used here. A detailed description is found in the text.

graphy as described in the methods (section 2.2.6). The NC protein was obtained by cleavage of C-CANC with the viral protease followed by heparin affinity chromatography as described in the methods (section 2.2.6) and in [42]. The Coomassie stained SDS-PAGE of the pure proteins used for the biopanning reaction is shown in Fig. 3.6.

The recombinant purified proteins as well as BSA were coated on magnetic beads (Dynabeads) for the biopanning reaction with a library of M13 derived phages presenting random 12mer peptides at the N-terminus of their pIII coat protein. After three positive and two negative selections phage pools with a titer of 1.2×10^7 pfu/μl (CA selection) and 3.2×10^6 pfu/μl (C-CANC selection) were obtained. The same approach was performed using a library presenting cyclic 7mer peptides, but titers of the selected phage pool were very low and none of the selected phages displayed specific binding. Thus, all of them were excluded from further analysis (data not shown).

Single phage clones from the pool selected with C-CANC or CA as target proteins were analysed for specific binding in ELISA experiments. Amplified single phages were incubated on multiwell plates coated with CA or BSA in the case of the CA selection and with C-CANC or NC in the case of the C-CANC selection. Phage binding was

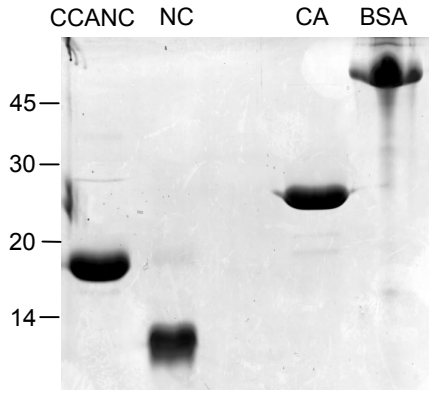

Figure 3.6.: Coomassie stained SDS-PAGE of 5μg of each purified recombinant protein used as target (C-CANC, CA) and negative target (NC, BSA) for the phage display selection. A molecular weight marker (in kD) is depicted on the left.

detected with a HRP-conjugated anti-M13 monoclonal antibody. In the case of the CA selection 36 of 41 analysed phages bound to CA but not to BSA (data not shown). The peptide coding region of the DNA of 29 of the specific phages was sequenced and found to code for only 2 different peptides as summarised in Table 3.1. In the case of the C-CANC selection the wild-type M13 phage as well as all selected phages were found to bind to the NC protein non-specifically. This seems to be an inherent property of all M13 phages, but does not necessarily mean that there is no CA specific binding. Thus, we repeated the test for specific binding using multiwell plates coated with CA and BSA. Out of 64 analysed phages 57 were found to specifically interact with CA (data not shown). Of those, 46 were sequenced and were found to code for 17 different peptides as summarised in Table 3.1.

ClustalW multiple alignment analysis allowed classification of the selected peptide sequences into four distinct groups (Table 3.1). Group 1 comprises all sequences selected with CA and most sequences selected with C-CANC as target protein. All group 1 peptides displayed significant sequence homology as indicated by shaded amino acids in Table 3.1. Intriguingly one sequence was selected with C-CANC as well as with CA as target protein (Table 3.1, line four). Additionally, phages presenting group 1 sequences displayed highest affinities for CA. Thus, group 1 peptides were chosen for further analysis. Phages presenting peptides classified in group 2 and 3 displayed slightly reduced affinity for CA and were therefore initially precluded from further analysis. Phages presenting group 4 peptides displayed low affinities for CA but showed unusually high affinities for C-CANC as compared to NC. Therefore these phages might

Table 3.1.: Upper part: Peptide sequences selected for each target protein, classification in groups based on sequence homology. Conserved amino acids are underlayed in gray. Numbers indicate, how often sequences were selected with C-CANC or CA as target proteins. Lower part: Synthetic peptides used for further analyses. Peptides 1 (Capsid Assembly Inhibitor, *CAI*), 2 and 3 were synthesised based on selected sequences. *CAI*ctrl has the same amino acid composition as *CAI* but a scrambled sequence. *CAI* and *CAI*ctrl were synthesised with an additional cysteine and labelled with the fluorophore FITC or with biotin.

	sequence	× selected (C-CANC)	× selected (CA)
group 1	ITFEDLLDYYGP	15	-
	ISWSELDAFMQM	6	-
	ISWMDLTAYYRG	-	2
	VSYSELTSYYMR	6	27
	VKYHDLQTFFDP	1	-
	VTYAQLQAYFPD	1	-
	LEFSDLEDFFRA	1	-
	LNFSDLNNYFLL	1	-
group 2	YNEPWWLTPSMF	3	-
	LDYPWWLSMNNI	1	-
group 3	STTWQDFFKTFG	5	-
	SYTQWDNAPGTR	1	-
group 4	IADRPRAWIGSP	1	-
	AMKTHTAIAPRA	1	-
others	HPQMHATPYQTT	1	-
	SPSNLYEQLLHW	1	-
	LPMIDIYRTAEL	1	-

synthetic peptides:

CAI (Pep1)	ITFEDLLDYYGP	s. above line 1
Pep2	VSYSELTSYYMR	s. above line 4
Pep3	IADRPRAWIGSP	s. above line 13
CAI-C	ITFEDLLDYYGPC	
CAI-C-FITC	ITFEDLLDYYGPC-FITC	
CAI-C-biotin	ITFEDLLDYYGPC-biotin	
*CAI*ctrl	IYDPTLYGLEFD	
*CAI*ctrl-FITC	IYDPTLYGLEFD-C-FITC	
*CAI*ctrl-biotin	IYDPTLYGLEFD-C-biotin	

interact with SP1 and were stored for later analysis. The remaining peptides that could not be classified into any of the groups seemed to bind unspecifically in repeated ELISA tests and therefore might be artefacts.

The following analysis will focus on group 1 peptides because of their obvious homology and good affinity for CA.

3.2.2. Synthetic peptides compete with phages for CA binding

Three of the selected peptides (Table 3.1, lower part) were synthesised and analysed for their ability to compete with peptide-presenting phages for binding to CA coated on ELISA plates. The peptide most abundant in the pool selected with C-CANC (Pep1) was chosen for further analysis and named *CAI* (Capsid Assembly Inhibitor). The peptide most abundant in the pool selected with CA (Pep2) and an additional non-homologous peptide classified into group 4 (Pep3) were analysed as well.

Multiwell plates were coated with CA and incubated with purified phages presenting *CAI* (phage-*CAI*) in the presence of increasing peptide concentrations. Bound phages were detected with a HRP-conjugated anti-M13 monoclonal antibody. The data were normalised against phage-binding to CA in the absence of peptides (100%). The fraction of bound phage-*CAI* in the presence of peptide is plotted against the peptide concentration in Fig. 3.7. In the presence of increasing concentrations of both group 1 peptides (*CAI* and Pep2) phage binding was reduced with similar efficiency to less than 20%, whereas Pep3 had no effect (Fig. 3.7 A). The fact that *CAI* and Pep2 both inhibited binding of phage-*CAI* indicated the interchangeability of group 1 peptides. This hypothesis was tested by analysing, whether *CAI* competes with phages presenting any group 1 peptide for CA binding in ELISA assays. Indeed, binding of all phages was reduced in the presence of *CAI* (data not shown), which confirms the cross-reactivity of group 1 peptides. Based on these results and on sequence homology it can be inferred that all group 1 peptides present variations on a common peptide motif and bind to the same region in the C-terminal domain of CA (C-CA).

Pep2 and *CAI* both inhibited phage binding with similar efficiency, but due to the limited solubility of Pep2 *CAI* was chosen for further analysis. In order to prove specificity of binding a peptide with the same amino acid composition as *CAI* but a scrambled sequence (*CAI*ctrl) was used as control in ELISA experiments. An additional cysteine was added to the C-terminus of *CAI* and *CAI*ctrl and these peptides (*CAI*-C and *CAI*ctrl-C) were labelled with biotin or with the fluorophore FITC. All peptide

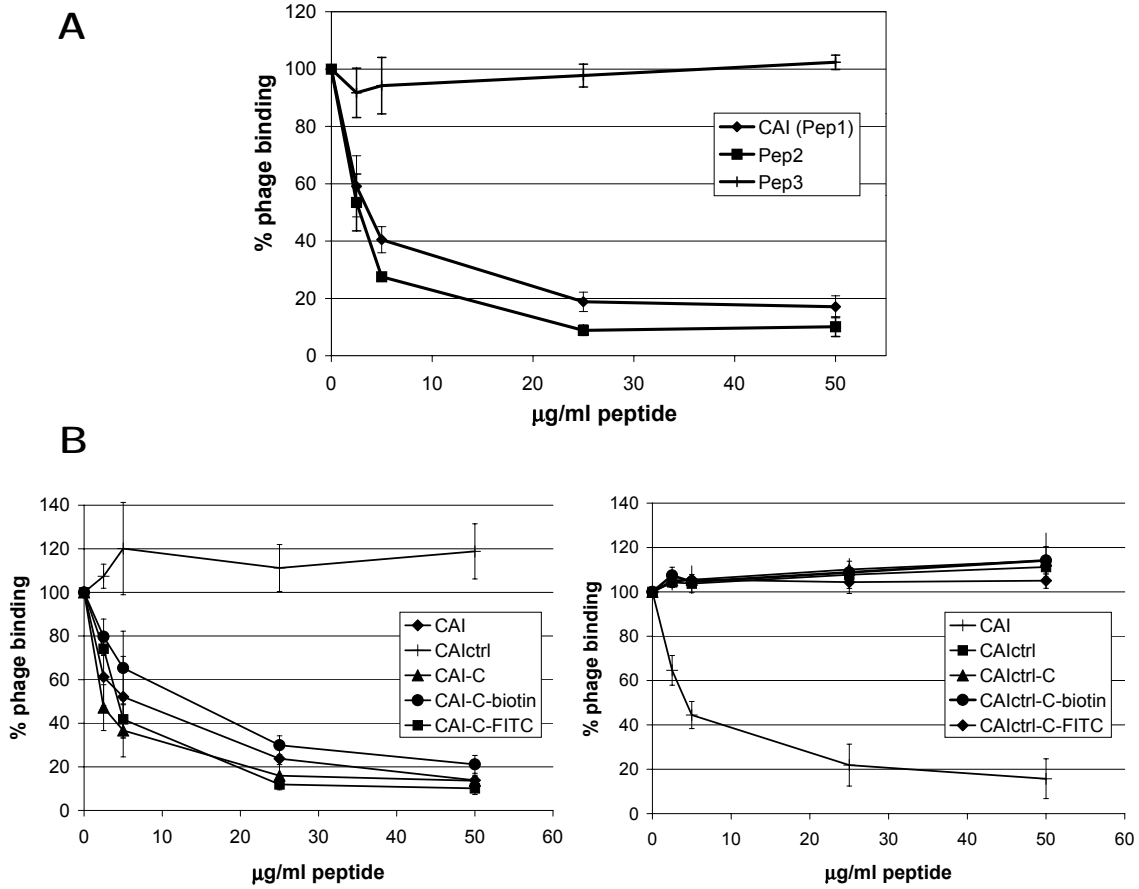

Figure 3.7.: ELISA assay to detect inhibition of phage-*CAI* binding to CA by synthetic peptides. ELISA plates were coated with CA and incubated with phage-*CAI* (1.2×10^6 pfu/μl) in the presence of increasing concentrations of peptides. A) Phage-*CAI* binding to CA (100%) is shown for increasing concentrations of *CAI* (Pep1), Pep2 and Pep3. B) Phage-*CAI* binding to CA (100%) is shown for increasing concentrations of *CAI*, its labelled derivatives and the control peptide *CAI*ctrl (left), and for *CAI*, *CAI*ctrl and its labelled derivatives (right).

derivatives are listed in Table 3.1 and they were all analysed in ELISA assays for their functionality. As seen in Fig. 3.7 B, all derivatives of *CAI* inhibited phage binding with similar efficiency (left panel), whereas the control peptide *CAI*ctrl as well as its derivatives had no effect (right panel). In contrast to that, biotin labelling of the N-terminus of *CAI*, completely abolished the functionality of the peptide (data not shown).

Since the function of *CAI* is lost when its sequence is scrambled, the binding to CA can be postulated to be a property inherent to the primary and possibly secondary structure of *CAI*. This emphasises the specificity of the interaction. The addition of a cysteine and a tag at the C-terminus does not abolish the functionality of the peptide.

3.2.3. *CAI* inhibits *in vitro* assembly of immature-like and mature-like particles

Since assembly of HIV is mediated by multiple weak Gag-Gag interactions, a small molecule that binds to CA might function as an assembly inhibitor. The ability of *CAI* to inhibit assembly *in vitro* was investigated by its influence on the formation of spherical or tubular particles from Gag-derived proteins.

In vitro assembly of spherical particles was performed using ΔMACANCSP2 as described in [45] and in the methods (section 2.3). Cryo-EM analysis had previously revealed that the ultrastructure of these *in vitro* assembled particles is remarkably similar to that of immature virus particles [45, 11]. Thus, *in vitro* assembly can be assumed to reliably reflect *in vivo* processes. ΔMACANCSP2 *in vitro* assembly was induced by changes in pH and salt concentration during protein dialysis in 0.1 kD cutoff tubing. Reactions were performed either without or with a fivefold molar excess of *CAI*, *CAI*ctrl, Pep2 or Pep3. Electron micrographs of negatively stained particles resulting from assembly reactions in the absence and in the presence of the peptides are shown in Fig. 3.8 A. Formation of spherical particles as seen for ΔMACANCSP2 alone on the top left panel is clearly abolished in the presence of a fivefold molar excess of *CAI* (top middle) or Pep2 (top right), whereas a fivefold molar excess of *CAI*ctrl or Pep3 (lower panel) had no effect on assembly of spherical particles. The labelled derivatives of *CAI* inhibited assembly with invariable efficiency (data not shown). The formation of aberrant structures or aggregates seemed to be enhanced in the presence of *CAI*, Pep2 or *CAI*-C as seen in Fig. 3.8, but these structures were generally observed in ΔMACANCSP2 assembly reactions upon dialysis in 0.1 kD cutoff tubing. Due to the high contrast they are hardly visible in the left panel of Fig. 3.8 A. In the absence of spherical particles the contrast is reduced and the aberrant structures become better visible as on pictures where assembly was inhibited. This might indicate that in the presence of the inhibiting peptides some sort of multimerisation (or aggregation) occurs but not in the ordered form of spherical particles.

In order to quantify the inhibitory potential of *CAI*, a titration was performed: A 0.5- to 5-fold molar excess of *CAI*-C was added to the *in vitro* assembly reaction and the number of spherical particles per unit area was counted at a 28,000-fold magnification. A representative titration is shown in Fig. 3.8 B. At equimolar concentrations of ΔMACANCSP2 and *CAI*-C the number of spherical particles was reduced by 80%. An EM picture of the titration endpoint confirms that no spherical particles are formed in the presence of a fivefold molar excess of *CAI*-C. Preliminary data indicated that

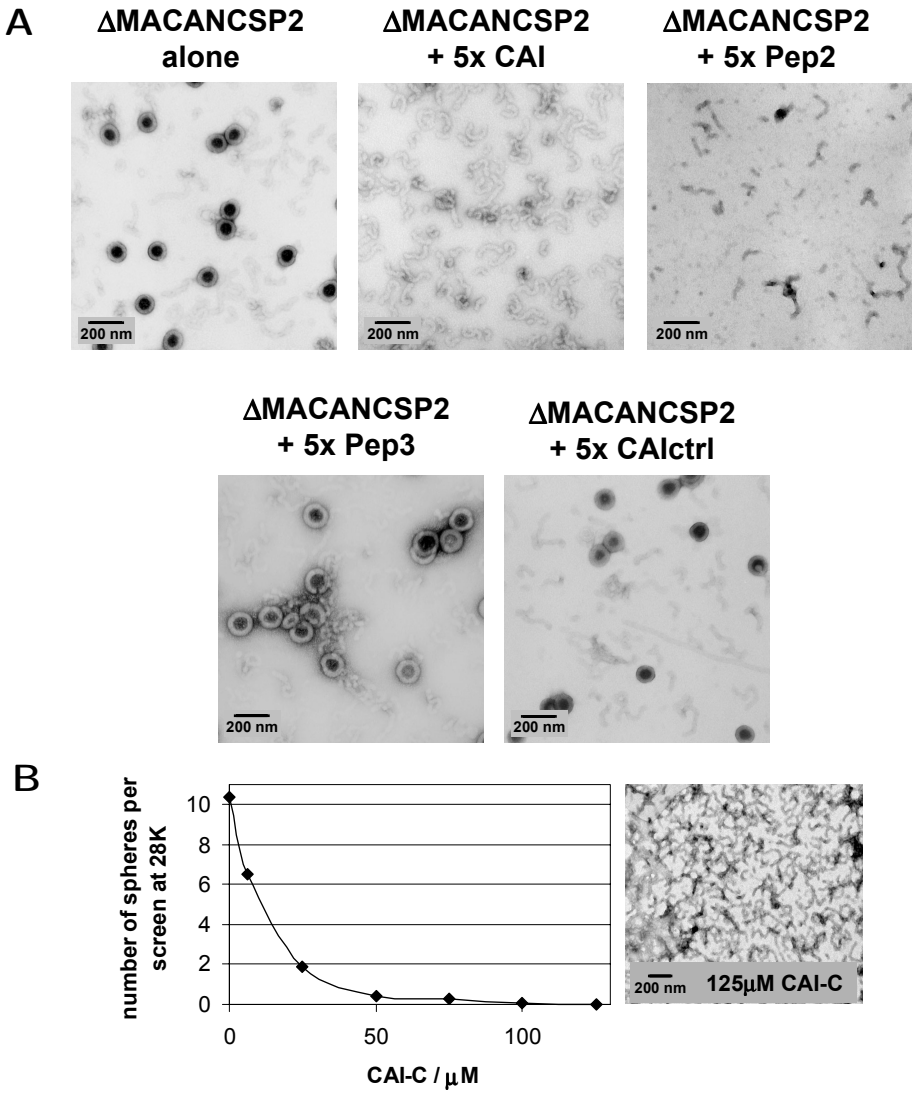

Figure 3.8.: A) Negatively stained EM images of particles resulting from *in vitro* assembly reactions of 25 μM ΔMACANCSP2 at 46,000-fold magnification in the absence or in the presence of *CAI*, Pep2, Pep3 or *CAI*ctrl. B) Titration curve of ΔMACANCSP2 assembly inhibition. The number (average of 20 counted screens) of spherical particles counted on the EM screen at a 28,000-fold magnification is plotted against the concentration of *CAI*-C present in the assembly reaction (corresponding to a peptide/protein ratio of 0.5 to 5). An electron micrograph of the endpoint of the titration is shown on the right.

the inhibition of the *in vitro* assembly reaction might be irreversible: After the usual dialysis in 0.1 kD cutoff tubing, assembly reactions were further dialysed in 12-14 kD cutoff tubing, in order to induce particle formation by removal of *CAI*. Even after 8 h dialysis still no spherical particles were found (data not shown). Taken together, these

results indicated that *CAI* efficiently inhibits assembly of immature-like particles.

Mature-like, tubular particles can be assembled from either CANC or CA *in vitro* [44]. For assembly of CANC, the same conditions as for ΔMACANCSP2 were used. Assembly reactions were performed without or with a 0.5- to 5-fold molar excess of *CAI* or a fivefold molar excess of *CAI*ctrl. Electron micrographs of negatively stained particles resulting from assembly reactions with increasing concentrations of *CAI* are shown in Fig. 3.9 A. Relatively low magnification was used to visualise decreasing par-

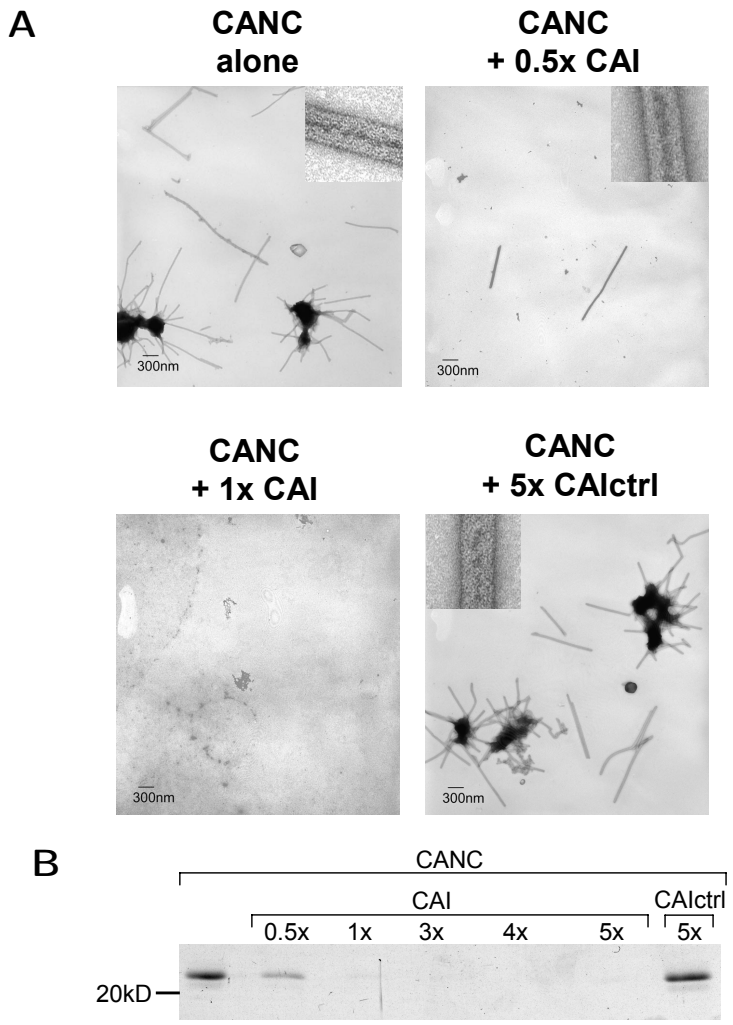

Figure 3.9.: A) EM images of negatively stained particles resulting from *in vitro* assembly reactions of 15 μM CANC at 13,000-fold magnification without or with *CAI* or *CAI*ctrl. Integrity of tubular particles is shown by 100,000-fold magnification insets. B) Coomassie stained SDS-PAGE of pelleted particles in the presence of increasing concentrations of *CAI* or *CAI*ctrl.

ticle numbers, while the integrity of tubular particles is shown in high magnification insets. Assembly of CANC into tubular particles was inhibited efficiently at equimolar concentrations of *CAI*, whereas the addition of a fivefold molar excess of *CAI*ctrl had no effect. The quantification of the inhibition of tubular particle assembly was complicated by the propensity of tubes to aggregate. Accordingly, counting of single particles was not possible in this case. Instead, particles were pelleted by centrifugation and the amount of protein present in resuspended pellets was analysed by SDS-PAGE. The Coomassie stained gel is shown in Fig. 3.9 B, indicating that assembly of CANC into tubular particles was abolished completely at equimolar concentrations of *CAI*. Thus, the formation of mature-like particles from CANC is inhibited with similar efficiency as the formation of immature-like particles.

The mature CA protein assembles into tubular particles independent of nucleic acid but only at high salt and protein concentrations [44]. This complicates the analysis of peptide mediated assembly inhibition because very high peptide concentrations are needed, invariably resulting in 1.75% DMSO in the reaction. This had deleterious effects on efficiency and quality of CA assembly. Still, assembly of CA was found to be abolished by addition of *CAI* as seen in Fig. 3.10, even though the quality of EM data was significantly reduced. However, no quantitative conclusions should be drawn.

Taken together, *CAI* inhibits assembly of spherical, immature-like as well as of tubular, mature-like particles with similar efficiency *in vitro*.

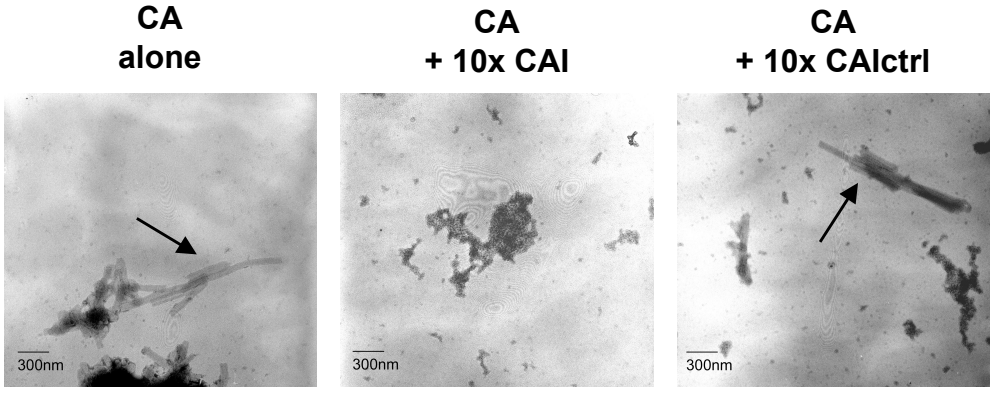

Figure 3.10.: EM images of negatively stained particles resulting from *in vitro* assembly reactions of 117 μM CA at 25,000-fold magnification without or with a 10-fold molar excess of *CAI* or *CAI*ctrl in the presence of 1.75% DMSO.

3.2.4. Preliminary analysis of *CAI* effects in cell culture

In order to determine whether *CAI* inhibits HIV release in cell culture, 293T cells were transfected with a plasmid (pNL4-3ΔPBSRT-) leading to the synthesis of non-infectious HIV particles. In parallel, cells were transfected with a mixture of plasmid and *CAI*. After 20 h particles were prepared from the cell culture supernatant. Western blot analysis was performed on particle preparations and lysed cells in order to detect differences in particle release or in the protein pattern inside cells. No differences were observed between cells transfected in the presence or in the absence of peptide (data not shown). The likeliest explanation for these negative results is that the peptide was not taken up by the cells.

In order to facilitate cellular uptake the Chariot™ protein transfection kit (Active Motif) was used. Prior to plasmid (pNL4-3ΔPBSRT-) transfection, 293T cells were either not treated at all or transfected with the Chariot reagent in the absence or in the presence of two concentrations of *CAI* or *CAI*ctrl according to the manufacturer's instructions. After 17 h virus particles were prepared from the supernatant and western blot analysis was performed on particle preparations and lysed cells using a polyclonal anti-CA antiserum (RαCA). The autoradiograph is depicted in Fig. 3.11. Again no difference was detectable for cells transfected with the plasmid alone, treated with the Chariot reagent without peptide or transfected with peptides. In order to evaluate the

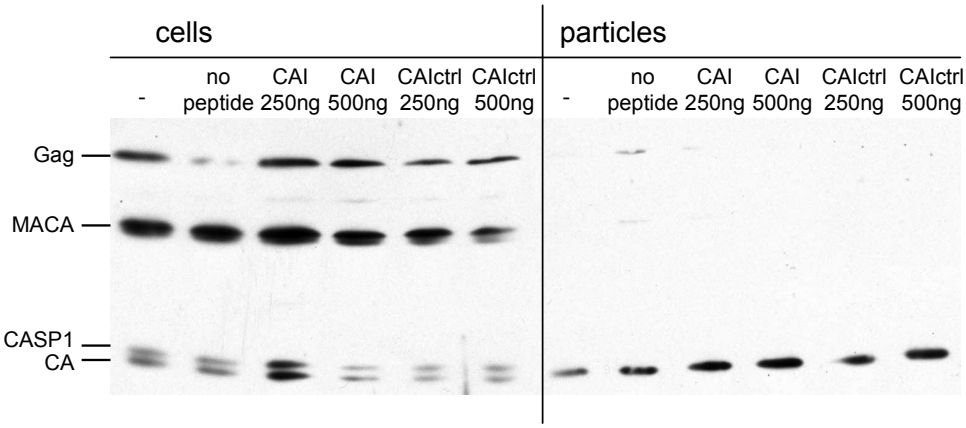

Figure 3.11.: Western blot analysis of transfected 293T cells and released virus particles. Cells were transfected with *CAI* and *CAIctrl* as indicated prior to transfection with pNL4-3ΔPBSRT-. Lysed cells and virus particle preparations were analysed for Gag-derived proteins by western blotting using a polyclonal anti-CA antiserum. Gag-derived proteins are identified on the left.

peptide uptake, the Chariot kit was used to transfect 293T cells with the fluorescently labelled peptide *CAI*-C-FITC. Only fluorescent aggregates were observed outside cells, but no intracellular fluorescence was detectable (data not shown). This indicated that no peptide uptake was achieved using the Chariot system.

In an alternative approach we used microinjection of Hela cells. *CAI*-C-FITC alone, a plasmid that leads to the synthesis of non-infectious HIV particles or a mixture of plasmid and *CAI* was injected. RT-PCR was chosen as readout for particle release, since the number of released particles was expected to be too low for any other approach. The number of RNA molecules released per microinjected cell was unexpectedly high and did not significantly differ between samples from cells injected with peptide (18,400 virions per ml and injected cell) and without peptide (24,500 virions per ml and injected cell). In addition, cells injected with *CAI*-C-FITC showed strong fluorescence immediately after injection, but no fluorescence was detected 24 h post injection, indicating that the peptide might already be degraded at the time HIV assembly occurs.

Taken together, no conclusion on the effect of *CAI* on HIV assembly in cell culture can be drawn from these experiments. It cannot be ruled out that the results are due to a lack of peptide uptake or stability or to technical problems.

3.2.5. Biochemical mapping of regions within CA important for the interaction with *CAI*

With the aim of mapping the peptide binding site within CA, the affinity of phage-*CAI* for CA derivatives was analysed in ELISA assays. C-terminally truncated (CA220, CA207) as well as point mutated CA variants (CA_{P207A}, CA_{E212A}, $CA_{Q219A,\ P224A}$, $CA_{I201T/E212A/E213G}$) and the C-terminal domain of CA (C-CA) as well as its dimerisation defective variant (C-$CA_{W184A/M185A}$) were chosen for this analysis. The affinity of phage-*CAI* for these proteins was expected to be reduced in case mutations or truncations affected the binding site of *CAI*. Protein purification protocols are given in the methods and all proteins used here are depicted schematically in Fig. 3.1 A. Point mutations can be located in the structure of the protein as shown in Fig. 3.1 B.

ELISA assays were performed as described above. Phage-*CAI* binding to the various proteins relative to binding to CA is plotted in Fig. 3.12. The deletion of the last 11 flexible amino acids of CA (221–231) had no effect on phage binding as indicated by results for the protein CA220. CA proteins lacking helix 4 of C-CA (CA207 and CAΔ207-220) showed almost complete loss of phage binding. Point mutation of surface

exposed amino acids within that region (P207A, E212A, Q219A, I201T/E212A/E213G, P224A) had no effect. Thus, either the deletion of helix 4 of C-CA induces structural rearrangements as suggested in section 3.1.1 that prevent *CAI* binding to another site of the protein, or amino acids different from those mutated here are involved in the interaction.

The affinity of phage-*CAI* for C-CA was as high as for full-length CA, but binding of the dimerisation defective protein C-CA$_{W184A/M185A}$ was significantly reduced (Fig. 3.12). This indicated that either the side chains of the dimer interface amino acids W184 and M185 are part of the *CAI* binding site or defective dimerisation abolishes binding of *CAI* to another site of the protein.

Taken together, the integrity of the dimer interface and of helix 4 of C-CA are important for the interaction of CA with *CAI*.

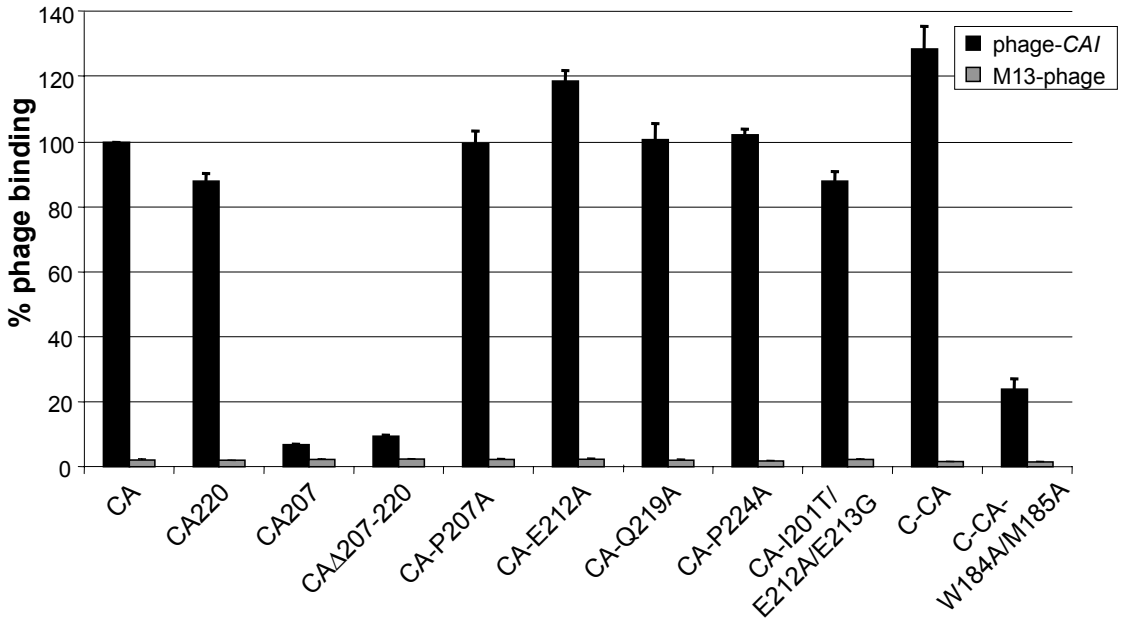

Figure 3.12.: ELISA assay to detect binding of phage-*CAI* to various CA derived proteins. ELISA plates were coated with CA derived proteins and incubated with phage-*CAI* (2.4×10^6 pfu/μl). Used were C-terminally truncated proteins (CA220, CA207), a deletion mutant (CAΔ207-220), CA proteins containing point mutations (P207A, E212A, Q219A, I201T/E212A/E213G, P224A), the C-terminal domain alone (C-CA) as well as its dimerisation defective variant (C-CA$_{W184A/M185A}$).

3.2.6. Binding constant and stoichiometry of the interaction as determined by isothermal titration calorimetry (ITC)

In the attempt to obtain thermodynamic parameters, such as the binding stoichiometry (n) and the dissociation constant (K_D) of the interaction, isothermal titration calorimetry (ITC) was performed. Since *CAI* binds to the independently folded C-terminal domain of CA as good as to the full-length CA protein (Fig. 3.12), the C-terminal domain alone was chosen for ITC analysis. Additionally, the dimerisation defective protein C-$CA_{W184A/M185A}$ was used in order to keep the thermodynamic complexity of the system low by avoiding effects arising from the monomer-dimer equilibrium of C-CA. C-$CA_{W184A/M185A}$ appears to be monomeric in solution up to about 2 mM concentration as deduced from NMR experiments (section 3.2.7).

Complex formation of *CAI* with C-$CA_{W184A/M185A}$ was clearly detectable by ITC analysis, as indicated by the heat released upon peptide addition. This is in accordance with the finding that the affinity of *CAI* for the mutant protein was reduced but not abolished in ELISA experiments (Fig. 3.12). Raw data as well as the calculated heat release plotted against the molar ratio of ligand and protein are shown in Fig. 3.13 A for the titration of C-$CA_{W184A/M185A}$ with *CAI*. The sudden increase in heat release seen in the raw data was due to an increase in the injection volume for all consecutive injections in order to achieve saturation. The reverse titration of *CAI* with C-$CA_{W184A/M185A}$ is shown in Fig. 3.13 B.

The measued heat release (plotted in the lower part of Fig. 3.13) is related to the reaction enthalpy dependent on protein and ligand concentrations in a differential binding curve. This curve may be fitted to the measured heat release to obtain thermodynamic data. The form of the curve depends on the binding model. The simplest case describes a single binding site. Apart from that, the curve can also be fitted assuming two equal binding sites or even two co-operative binding sites. In our case, best fits were obtained for the one-site model, but were still not satisfactory. This is partly due to the fact that the relatively high affinity of the interaction precluded to reach a plateau in the beginning of the reaction under the chosen conditions. This means that the protein/ligand ratio, where every added ligand molecule binds, is not well defined in our case. This inevitably influences the curve fits and thus the determination of the thermodynamic parameters. Also, no final saturation seemed to be reached at the end of the titration, which might indicate the presence of a second process that cannot be further specified. Additionally, relatively large heats of dilution can be observed indicating for example dissociation of aggregates upon dilution. Thermodynamic data given in Fig. 3.13 were

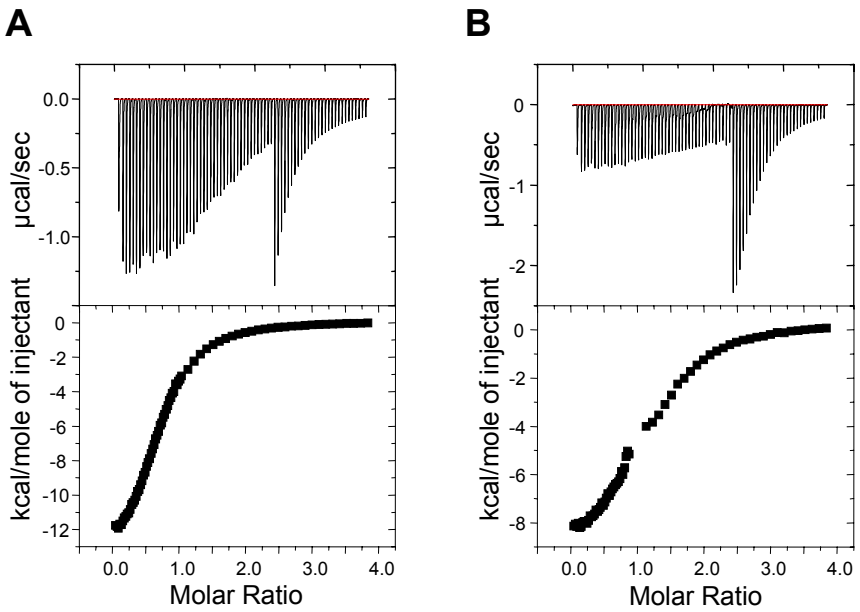

Figure 3.13.: Isothermal titration calorimetry of C-$CA_{W184A/M185A}$ with *CAI* in 100 mM ammonium acetate pH 7.0, 5 mM DTT. A) 130 μM C-$CA_{W184A/M185A}$ titrated with 2 mM *CAI*. The best fit ($\chi^2 = 6410$) was obtained for a one-site model giving: $K_D = 14.3 \pm 0.3\,\mu$M, $n = 0.74 \pm 0.00$, $\Delta H = -13.99 \pm 0.06$ kcal/mol, $\Delta S = -24.75$ cal/mol K. B) The reverse titration using 130 μM *CAI* titrated with 2 mM C-$CA_{W184A/M185A}$. The best fit ($\chi^2 = 28069$) was obtained for a one-site model giving: $K_D = 19.5 \pm 0.1\,\mu$M, $n = 1.16 \pm 0.01$, $\Delta H = -9.4 \pm 0.1$kcal/mol, $\Delta S = -10.0$ cal/mol K.

inferred from the best fit obtained with the one-site model. These data have to be treated with care and it can only be inferred that the interaction presumably displays 1:1 stoichiometry with a K_D in the range of 10 to 30 μM.

3.2.7. Nuclear magnetic resonance (NMR) spectroscopy of the *CAI*-protein complex

Since the above experiments were not sufficient to derive a binding site for *CAI* or a stoichiometry or affinity of the interaction, the interaction of C-CA and C-$CA_{W184A/M185A}$ with *CAI* was analysed in more detail using NMR spectroscopy. Since *CAI* binds to the independently folded C-terminal domain of CA as good as to the full-length CA protein (Fig. 3.12), the protein C-CA was tested for its suitability for NMR analysis. 2D ^{1}H-^{15}N-HSQC spectra reflect the correlation between nitrogens attached to a proton and thus give rise to a single peak for every backbone NH corresponding

to a single amino acid. This leads to a pattern that is very characteristic for the respective protein. Thus, these spectra allow to investigate whether a sample is suitable for detailed NMR analysis. A Coomassie stained gel of a SDS-PAGE analysis showing the pure C-CA protein is depicted together with a 2D ^{1}H-^{15}N-HSQC spectrum in Fig. 3.14 A. The spectrum was of relatively poor quality. Much less than the expected number of resonances and general line broadening was observed. This was described previously and was attributed to C-CA dimerisation [78].

Problems associated with C-CA dimerisation were circumvented by using the dimerisation defective double mutant C-CA$_{\mathrm{W184A/M185A}}$. Single or double dimer interface mutations had been used in several previous analyses [37, 62, 71, 92, 60, 38] and the integrity of these residues was shown to be important but not essential for immature

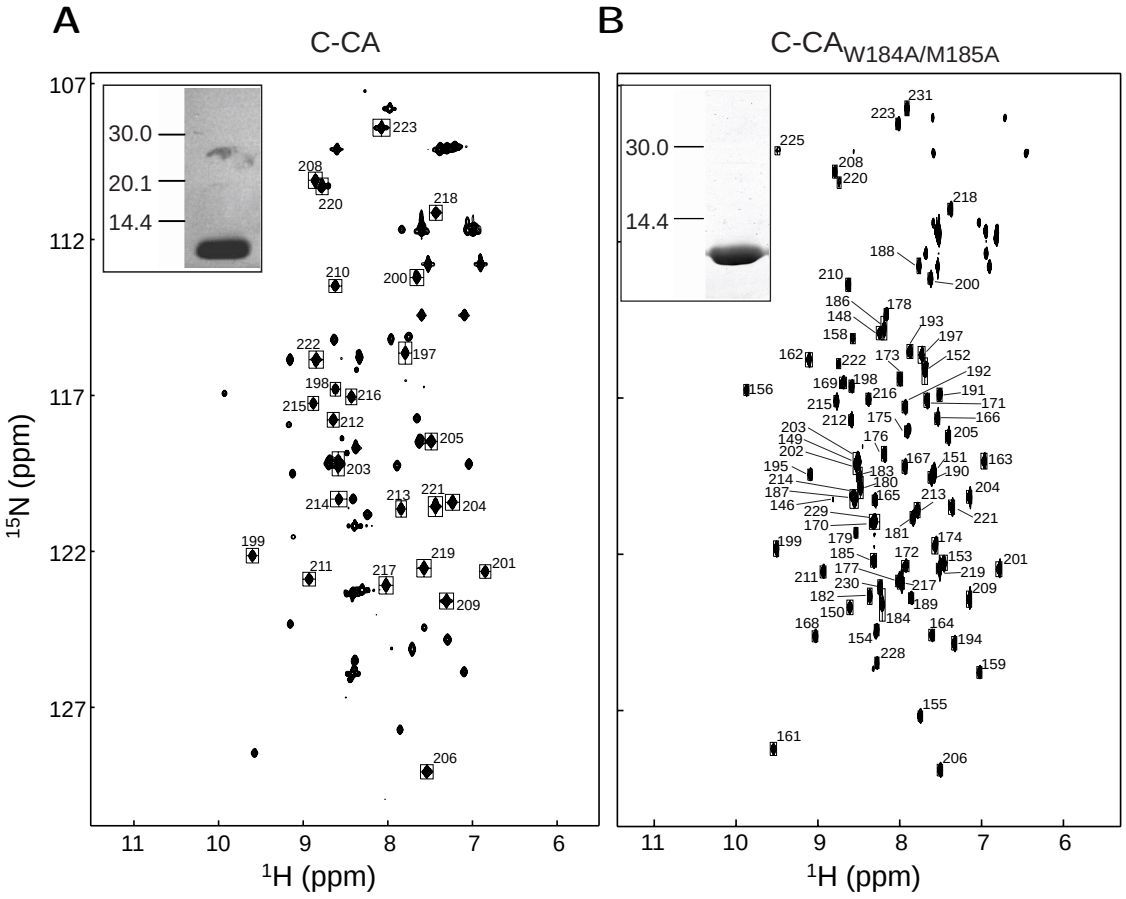

Figure 3.14.: ^{1}H-^{15}N-HSQC spectra for C-CA (A) and C-CA$_{\mathrm{W184A/M185A}}$ (B). Assigned ^{15}N-H resonances are labelled with the corresponding CA residue number. A Coomassie stained SDS-PAGE of 10μg of each purified protein is shown in the insets with a molecular weight marker (in kD) on the left.

particle assembly [37], which indicated that the overall functionality of the protein seems to be retained. In Fig. 3.14 B an HSQC spectrum together with an SDS-PAGE analysis of the pure protein is shown. The quality of the NMR spectrum improved significantly as may be seen by comparison of panels A and B of Fig. 3.14. Only two peaks less than expected were observed and the line shapes were as expected for a 10 kD monomeric protein. Spectra did not change significantly within the investigated concentration range of 0.13 to 2 mM protein.

3.2.7.1. Sequence resonance assignment and structural integrity of C-CA$_{W184A/M185A}$

To date a complete backbone resonance assignment for C-CA has not been reported. Partial assignments mainly for helix 3 and 4 of C-CA in the context of the protein C-CANC are available [78]. In accordance with that, only partial backbone assignments were obtained for C-CA based on NOESY and TOCSY experiments, as indicated in Fig. 3.14 A. However, I obtained sequence resonance assignments (appendix A.4) for all but two amino acids (H226, K227 part of the flexible C-terminus) of the dimerisation defective protein C-CA$_{W184A/M185A}$, using standard triple-resonance experiments on $^{13}C/^{15}N$ double labelled protein. The assigned HSQC resonances in both spectra are numbered with the corresponding CA residue in Fig. 3.14. As an example for the sequential assignment, selective strips taken from a HNCACB spectrum of C-CA$_{W184A/M185A}$ are depicted in Fig. 3.15. For illustration, the sequential connectivities are marked by lines for Cα (black) and Cβ (red) resonances.

The structural integrity of the mutant protein was investigated by analysis of secondary $^{13}C\alpha$ chemical shifts. Cα and Cβ chemical shift values have been correlated empirically to secondary structure elements as described in the methods 2.6.5 [86]. For each assigned Cα of C-CA$_{W184A/M185A}$ the chemical shift deviation from the random coil value is plotted in Fig. 3.16. For comparison, the helical elements (helix 1 to helix 4) of C-CA are indicated as derived from the published crystal structure. Large positive chemical shift deviations are indicative of α-helix formation, whereas negative deviations are obtained for extended structures such as loop regions or β-sheet structures. As is clearly seen by comparison of clusters of positive secondary shifts with the helical elements of C-CA, all helices are conserved except for the beginning of helix 2 (framed), which seems to be destabilised. This region of reduced propensity of helix formation includes the dimer interface mutation.

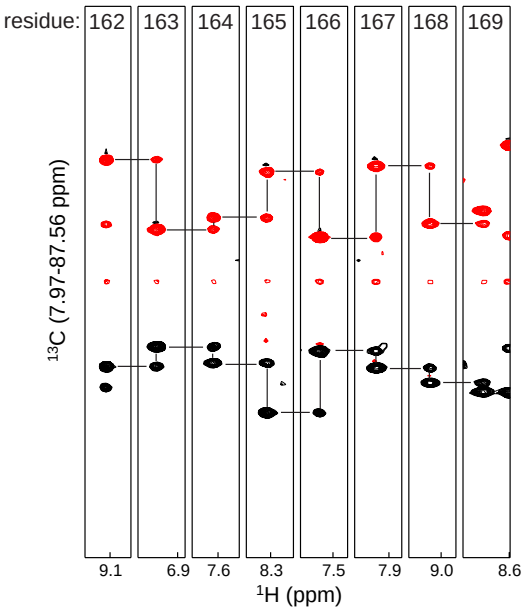

Figure 3.15.: Representative strips taken from a HNCACB spectrum of C-CA$_{W184A/M185A}$. Shown are strips extracted at the amide chemical shifts of CA residues 162–169, as indicated. Sequential connectivities are indicated by lines.

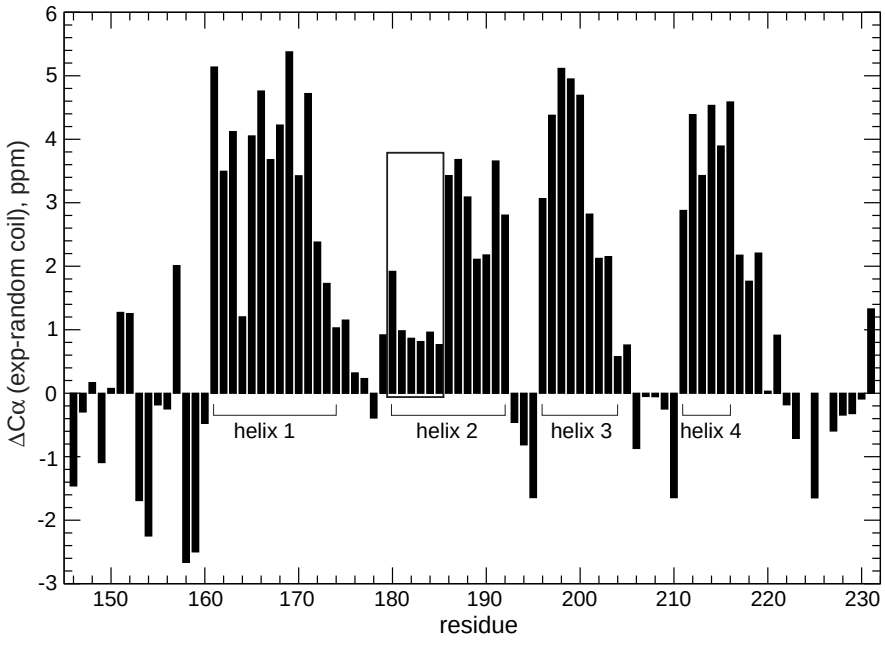

Figure 3.16.: Cα chemical shift deviations from random coil (peptides in urea) values. Chemical shift deviations are plotted for each residue of C-CA$_{W184A/M185A}$. The location of helices 1 to 4 of C-CA are indicated. The beginning of helix 2 of C-CA is framed to emphasise the observed effect.

Thus, it can be inferred that the removal of the hydrophobic dimer interface side chains (W184, M185) destabilises helix 2, which might contribute to the reduction in dimerisation of the mutant protein. With the exception of the dimer interface region the overall secondary structure of the protein seems to be retained.

3.2.7.2. Determination of the *CAI* binding site by chemical shift perturbation

NMR chemical shift perturbation was used to determine the binding site of *CAI*. For the analysis of protein-ligand interactions 2D ^{1}H-^{15}N-HSQC spectra are particularly well suited, since NH resonances are very sensitive to changes in their local environment. Thus, ligand binding induces chemical shift perturbation at the binding site. With the knowledge of NH backbone assignments, the binding site can then be mapped to the protein sequence. Since spectra obtained with the dimerisation defective protein C-CA$_{\mathrm{W184A/M185A}}$ were of much better quality and backbone assignments could be readily obtained, this protein was used to analyse the interaction with *CAI*, although the affinity of this interaction was shown to be reduced as compared to wild-type C-CA (Fig. 3.12). Complex formation was clearly detectable by NMR chemical shift perturbation analysis and this is in accordance with the ITC data. The overlay of a series of 2D ^{1}H-^{15}N-HSQC spectra obtained for the C-CA$_{\mathrm{W184A/M185A}}$ protein without and with increasing concentrations of *CAI* is shown in Fig. 3.17 A. While most peaks were unaffected (examples are encircled), significant chemical shift perturbation was detectable for a subset of peaks, as indicated by arrows in Fig. 3.17 A. These resonances displayed large shifts indicative of large environmental changes as expected for the binding of a peptide with three aromatic residues. The fact that resonances corresponding to the free protein gradually disappeared and resonances corresponding to the complex concomitantly appeared undergoing slight line broadening is indicative of a slow to intermediate exchange regime on the NMR time scale.

A small number of resonances showing slight chemical shift perturbation were in the fast exchange regime. One of these residues is labelled with a star in Fig. 3.17 A. These resonances are suitable for the calculation of the K_{D} value of the interaction since the fractional shift (chemical shift difference between free protein and protein with a given ligand concentration divided by the maximal observed difference) is a function of the K_{D} value. Thus, the resulting equation given in the methods 2.6.6 can be fitted to experimentally obtained fractional shifts which leads to an average K_{D} of $15.0 \pm 7.2\,\mu$M. This value is in good agreement with the slow to intermediate exchange regime observed

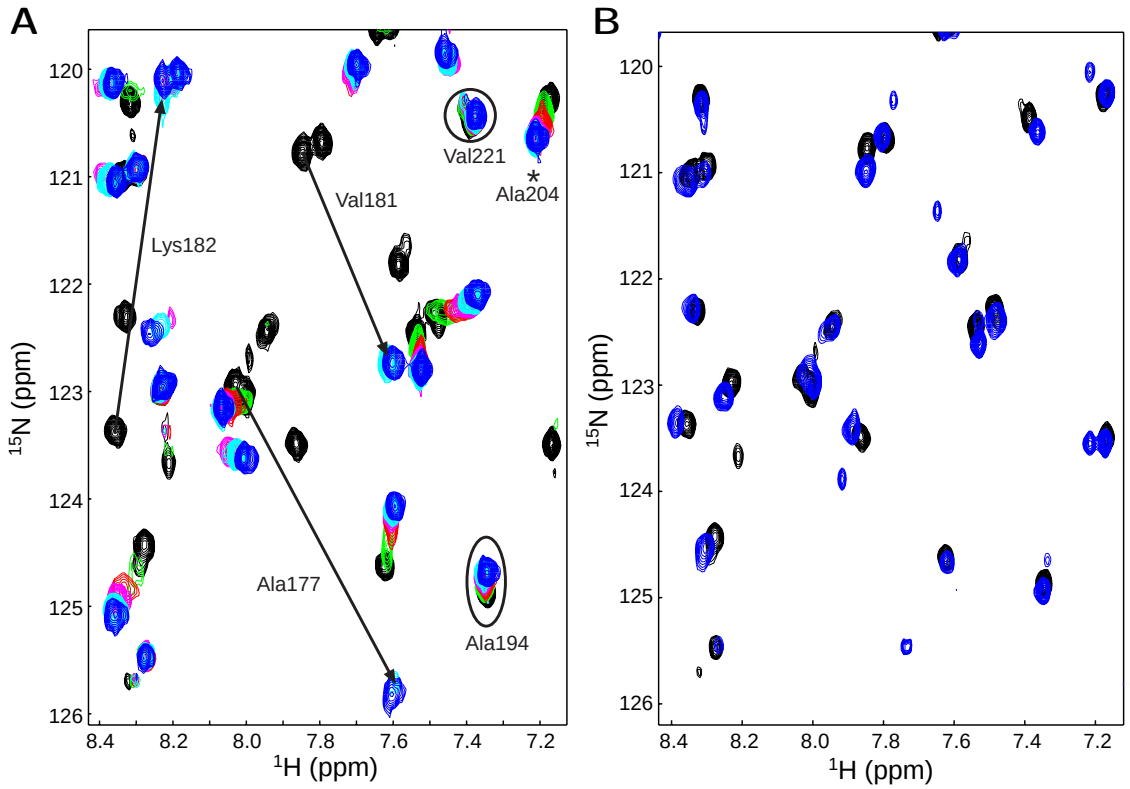

Figure 3.17.: A) Overlay of 2D-^{1}H-^{15}N-HSQC spectra obtained for the titration of free C-CA$_{W184A/M185A}$ (130 μM, black) with increasing concentrations of *CAI* (peptide:protein ratio as follows: 0.5:1 (green), 1:1 (red), 2:1 (magenta), 4:1 (cyan), 7.3:1 (blue)). Examples of peaks shifting upon *CAI* titration are indicated by arrows, while unaffected peaks are encircled. For clarity only a subset of resonances is annotated with residue assignments. B) Overlay of 2D ^{1}H-^{15}N-HSQC spectra obtained for 130 μM C-CA$_{W184A/M185A}$ without peptide (black) and with a fourfold molar excess of *CAI*ctrl (blue).

for resonances displaying large chemical shift perturbations and it matches the range for the K_D value obtained by ITC.

A separate 2D ^{1}H-^{15}N-HSQC experiment was performed with the control peptide *CAI*ctrl. The fact that no significant chemical shift perturbation was observed indicated that C-CA$_{W184A/M185A}$ does not interact with the control peptide. This can be seen in Fig. 3.17 B.

In addition to the assignment of the free protein (Fig. 3.14 B), backbone resonance assignments (appendix A.4) based on standard triple resonance experiments were also obtained for the peptide-bound form of C-CA$_{W184A/M185A}$. Assignments obtained with protein samples at 1 mM concentration and a fourfold molar excess of *CAI* were used

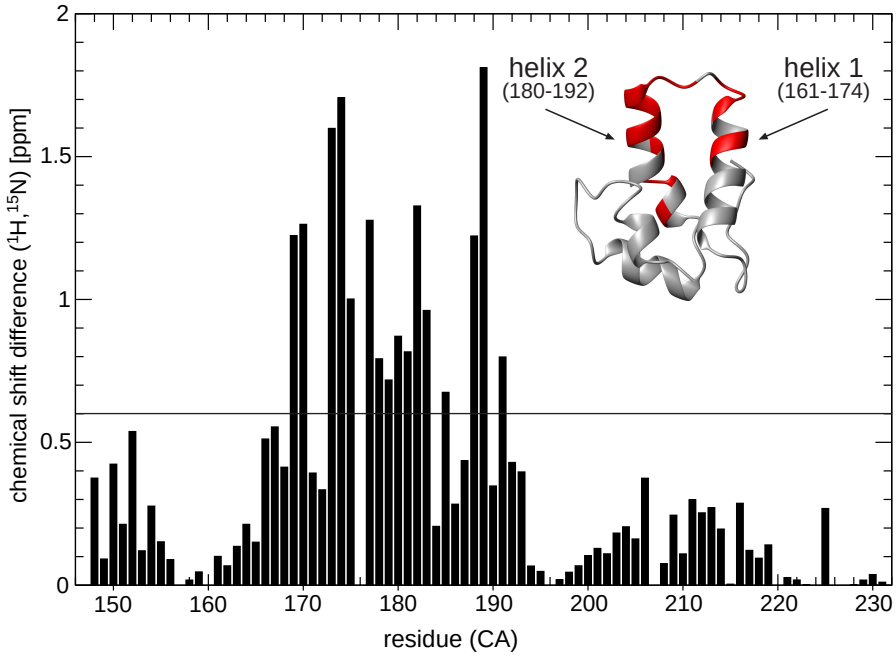

Figure 3.18.: ^{1}H-^{15}N chemical shift perturbations plotted for each unambiguously assigned amino acid of C-CA$_{W184A/M185A}$ (CA residue numbering). The inset represents a ribbon diagram of the structure of the C-terminal domain of CA (pdb ID: 1A8O [37]) with amino acids showing chemical shift perturbations larger than 0.6 ppm coloured in red. The structure was generated with MOLMOL [58] and the orientation of the structure was chosen to best visualise the affected residues.

to assign the resonances observed upon chemical shift perturbation analysis (appendix A.4) as indicated for some peaks in Fig. 3.17 A. ^{1}H-^{15}N chemical shift perturbations induced by *CAI* are were calculated (methods 2.6.6) and plotted for each assigned CA residue of C-CA$_{W184A/M185A}$ in Fig. 3.18 A. The biggest shifts occurred within amino acids 169–191 suggesting that this region contains the binding site for *CAI*. Based on the chemical shift analysis (Fig. 3.16), the x-ray structure of C-CA [37, 100, 7] can be assumed to be a reasonably good model for the structure of C-CA$_{W184A/M185A}$. In order to visualise the binding site, amino acids showing chemical shift perturbations larger than 0.6 ppm (marked by a line in Fig. 3.18) are coloured red in the crystal structure of C-CA, as shown in the inset in Fig. 3.18. The orientation of the C-CA structure was chosen in order to best visualise the affected residues. In this orientation the dimer interface is hidden in the back of the protein indicating that the peptide might bind to the side opposite to the dimer interface. Taken together, the top part of helix 1, the inter-helical linker and helix 2 are strongly affected by peptide binding.

The structural integrity of the peptide-bound protein was analysed by secondary Cα chemical shift analysis. Chemical shift deviations for each residue are plotted in Fig. 3.19. Again helical elements (helix 1 to helix 4) of C-CA are indicated and were compared to clusters of positive chemical shift deviations. The structural integrity of all helices seems to be retained, and, interestingly, even helix 2 (framed) seems to be restored as compared to the mutant protein without peptide (Fig. 3.16).

Thus, the affinity of *CAI* for the mutant protein might be reduced either because the side chains of the mutated residues are part of the binding site, or because binding energy is required for the formation and stabilisation of helix 2. This also indicates that chemical shift perturbation observed in the area of the restored helix might partly be due to the conformational rearrangement.

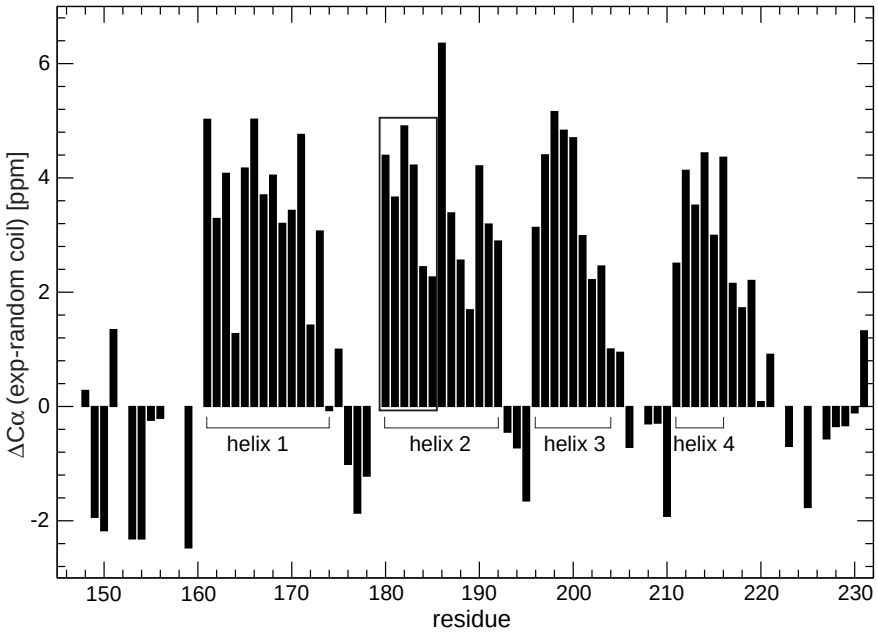

Figure 3.19.: Cα chemical shift deviations from random coil (peptides in urea) values. Chemical shift deviations are plotted for each residue of C-CA$_{\mathrm{W184A/M185A}}$ in the presence of a fourfold molar excess of *CAI*. The location of helices 1 to 4 of C-CA are indicated and the beginning of helix 2 is framed to emphasise the observed effect.

Taken together, these data identify a binding site (CA amino acids 169–191) and a dissociation constant ($K_{\mathrm{D}} = 15.0 \pm 7.2\,\mu$M) for the interaction of *CAI* with C-CA$_{\mathrm{W184A/M185A}}$. Additionally, the destabilised helix 2 in the mutant protein seems to be restored upon *CAI* binding. The fact that no significant chemical shift perturbation was observed for amino acids 207–220 indicates, that this region is not an essential part

of the binding site. The peptide binding site comprises the mutated dimer interface residues W184A and M185A and the beginning of helix 2 seems to be stabilised upon peptide binding. Thus, the peptide might exert its inhibitory effect either by interfering with the CA dimer interface or with a yet unknown interface within the C-terminal domain of CA.

3.2.7.3. *CAI* does not inhibit dimerisation of C-CA

In order to determine whether peptide binding abolishes CA dimerisation, the overall rotational correlation times (τ_m) of free and peptide-bound C-CA were determined based on ^{15}N-T_1 and ^{15}N-T_2 relaxation measurements. In case the *CAI* binding disrupts the C-CA dimer the correlation time should be reduced as compared to free (dimeric) C-CA. If not, the correlation time is expected to increase because of the additional molecular weight of the bound peptide. In the case of free C-CA, τ_m was determined to be 9.0 ± 0.2 ns, which is a typical value for molecules of about 20 kD (a C-CA dimer). This value increased to 10.2 ± 0.2 ns upon *CAI* binding, which is consistent with the increased molecular weight expected for a protein-dimer in complex with the peptide (about 23 kD). For a monomeric peptide-protein complex the expected correlation time would be closer to 5 ns. Thus, C-CA dimerisation is not abolished upon *CAI* binding indicating that peptide binding does not block the dimer interface.

A graphical representation of T_1 and T_2 values for each analysed backbone ^{15}N is depicted in Fig. 3.20. Experimental values obtained for the free protein are shown in black circles, values for the protein bound form are shown in red circles. Added to this plot are curves of calculated T_1 and T_2 values as a function of the rotational correlation time and order parameter S^2 according to the Lipari-Szabo model [66] (methods 2.6.7). Most values cluster in a small region of the plot within an S^2 of 1.0 to 0.8, indicating that the movement of these residues is dominated by the tumbling of the protein as a whole. Residues corresponding to smaller S^2 values are highly flexible and can be assumed to rotate independently of the globular part of the protein. These peaks for example correspond to residues located within the flexible C-terminus of C-CA (amino acids 220–231). Residues shifted to small T_2 values, outside of the model, are affected by slow motion processes, such as peptide binding or oligomerisation. Thus, the plot indicates that all analysed residues except for those at the flexible C-terminus tumble as a single unit in solution. The overall increase in T_1 and decrease in T_2 values in the presence of the peptide indicates a slight increase in molecular weight, as expected for the peptide binding to the protein dimer.

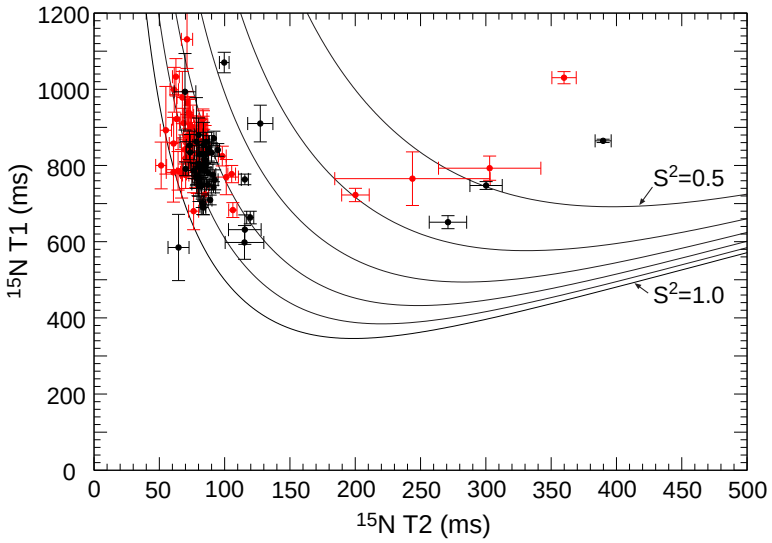

Figure 3.20.: ^{15}N-T_1 and ^{15}N-T_2 experimental values obtained for C-CA in the absence (black circles) and in the presence of equimolar amounts of *CAI* (red circles). Added to this plot are curves of calculated T_1 and T_2 values as a function of the rotational correlation time and the order parameter S^2 according to the Lipari-Szabo model [66]. The protein concentration was 390 μM in 100 mM ammonium acetate pH 7.0, 5 mM DTT.

These data establish that *CAI* exerts its inhibitory effect on HIV assembly by interfering with an assembly interface different from the C-CA dimer interface.

3.2.8. Crystallography of the complex of C-CA with *CAI*

The crystallisation of the complex of *CAI* with C-CA was initiated in collaboration with Felix Rey's group at the CNRS, Gif-sur-Yvette. All the following experiments on crystallisation and crystallography of the complex were performed by François Ternois. Co-crystals of C-CA with *CAI* were obtained and despite their fishbone-like appearance (Fig. 3.21 A) they were shown to diffract to 1.8 Å. A preliminary automated reconstruction of the backbone structure of the complex was obtained by molecular replacement as depicted in the MOLMOL representation in Fig. 3.21 A. The peptide is coloured blue and the backbone shifts observed within C-CA$_{W184A/M185A}$ by NMR spectroscopy are coloured red this structure of complexed C-CA. The dimer interface residues W184 and M185 are hidden in the back of the structure.

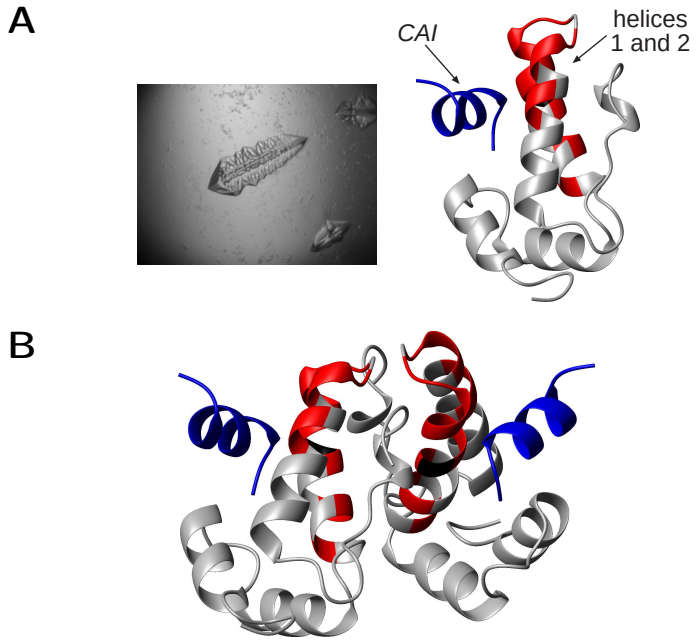

Figure 3.21.: A) A representative co-crystal obtained for C-CA with *CAI* (courtesy of François Ternois, crystallisation conditions were 32% PEG 4000, 10 mM $MgCl_2$, 100 mM sodium acetate pH 4.6, C-CA:*CAI*=1:2). A ribbon diagram of the preliminary structure of the C-terminal domain of CA with *CAI* (blue) is shown on the right. Residues undergoing chemical shift perturbations in NMR experiments are coloured red. B) The dimeric form of the complex is shown. Structures were generated with MOLMOL [58].

This preliminary crystal structure suggests the following: No major structural rearrangements of the protein occur upon peptide binding. The *CAI* binding site is located opposite to the dimer interface indicating that dimerisation of the protein is not affected by binding and the interaction is mainly mediated by the N-terminus of the peptide. The model of a dimer of the C-CA/*CAI* complex shown in Fig. 3.21 B. Taken together, this preliminary structure is in very good agreement with NMR and biochemical data described above.

3.3. Monoclonal antibodies as tool to analyse virus assembly

The above described *in vitro* system allowed to study assembly of immature-like, spherical and mature-like, tubular particles independently. However, it did not elucidate the conformational changes within the CA protein that transform the immature spherical

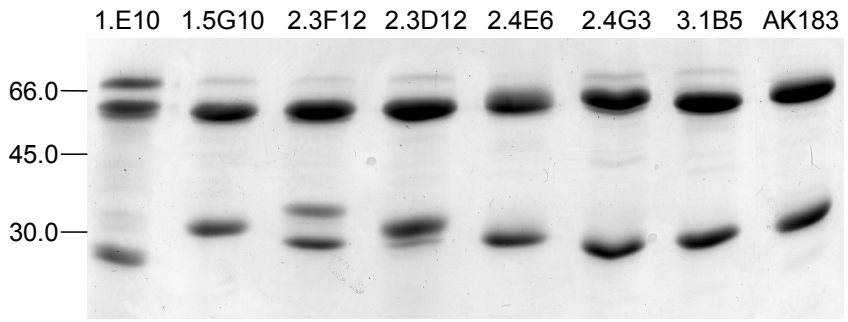

Figure 3.22.: Coomassie stained reducing SDS-PAGE of 10 μg of each used mAbs. A molecular weight marker (in kD) is depicted on the left.

shell into the mature conical core upon cleavage *in vivo*. It was postulated by Gross et al. that this conformational switch, which is triggered by Gag processing *in vivo* might be mimicked by a change in pH [45]. This hypothesis was based on the finding that under *in vitro* assembly conditions the immature protein ΔMACANCSP2 formed immature-like particles at pH 8, but the same protein formed mature-like particles at pH 6 as illustrated in Fig. 1.5 in the introduction. This indicated that pH might trigger the same conformational rearrangements that *in vivo* accompany maturation. A change in pH from 8 to 6 went along with a loss in binding of the monoclonal antibodies (mAbs) 1.5G10 and 2.4E6 to ΔMACANCSP2 (Fig. 1.5), although they recognised their peptide epitopes independent of pH. Thus, these antibodies were hypothesised to differentiate between two possible conformations of the CA domain: one that allows assembly of mature particles and one that drives assembly of immature particles [45].

In order to test this hypothesis, I performed a detailed analysis of the observed conformational changes. A panel of seven mouse mAbs targeting different regions within CA was used (1.5G10, 2.3D12, 2.3F12, 2.4E6, 2.4G3, 3.1B5, AK183), including the antibodies mentioned above. A mAb directed against MA (1.1E10) and a anti-Polio (αPolio) mAb were used as controls. Antibody integrity was confirmed by SDS-PAGE analysis under reducing conditions. As seen in Fig. 3.22, antibody solutions were pure and heavy and light chains were present in comparable amounts for all antibodies.

The following aspects will be described in detail below: The antibody epitopes were mapped and mAbs were tested for inhibition of *in vitro* assembly of spherical particles. A detailed analysis of pH dependent binding to various Gag-derived proteins was performed in order to test the above hypothesis and to delineate the functional relevance and the mechanism of conformational changes.

3.3.1. Epitope mapping of a set of monoclonal antibodies

3.3.1.1. Epitope mapping by peptide scan technology

The peptide scan technology, a method for linear epitope analysis, was used to map the epitopes of the available mAbs. Membranes with spots of overlapping peptides covering the sequence of the respective Gag-derived target protein were stained with Ponceau to visualise all peptide spots. After incubation with mAbs, reactive spots comprising the epitope sequence were detected with HRP-conjugated antibodies and visualised by autoradiography.

Ponceau stained Gag membranes and autoradiographs showing the reactive peptide spots after incubation with antibodies 2.4E6 and 3.1B5 are presented in Fig. 3.23 as examples. The variation in intensity of Ponceau staining is due to the peptides' amino acid composition. An overlay of the autoradiograph and the stained membrane allowed the determination of the number of the reactive spot and thus the corresponding amino acid sequence within CA as indicated in Fig. 3.23. Results obtained from these experiments for all mABs are summarised in Table 3.2. Epitopes of some mAbs had been mapped previously using an ELISA based approach (Klaus Wiegers, personal communication and [45]). These results correlated very well with the data obtained by peptide scan technology (Table 3.2). Two previously unknown epitopes could be determined additionally, and results were confirmed using membranes covering the CA sequence with an alternative set of overlapping peptides (Table 3.2): The epitope of antibody 3.1B5 is located in helix 6 of N-CA (amino acids 109–118) and AK183 binds to amino acids 156–166 of the C-terminal domain of CA (C-CA). These residues are part of the major homology region (MHR, CA amino acids 153–172), a highly conserved sequence among several retroviruses. All determined epitopes are highlighted in the structures of N-CA and C-CA in Fig. 3.24. Antibody 2.4G3 did not react with any membrane,

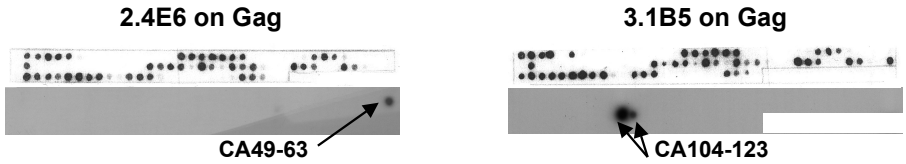

Figure 3.23.: Antibody epitope mapping by peptide scan technology. Membranes with spots of 15mer overlapping oligo peptides covering Gag sequences were incubated with antibody 2.4E6 or 3.1B5, respectively. Ponceau stained membranes and the autoradiographs visualising the reactive peptide spots are shown. CA amino acid sequences forming the epitope are indicated.

Table 3.2.: List of antibody epitopes as determined by peptide scanning, western blot analysis and phage display as well as results obtained by K. Wiegers. Numbers indicate CA and MA residues forming the epitope.

antibody	mapping results	epitope
1.1E10	**121MA-3CA (peptide scan)**	MA 122–132
	MA 122–132 (K. Wiegers)	
1.5G10	**CA 108–122 (peptide scan)**	CA 109–120
	CA 109–120 (K. Wiegers)	
2.3D12	**CA 129–143 (peptide scan)**	CA 129–143
	CA 121–144 (K. Wiegers)	
2.3F12	**CA 180–194 (peptide scan)**	CA 182–192
	CA 182–192 (K. Wiegers)	
2.4E6	**CA 49–63 (peptide scan)**	CA 49–62
	CA 49–62 (K. Wiegers)	
2.4G3	**CA 194–207 (western blot)**	
3.1B5	**CA 109–119 (peptide scan I)**	CA 109–118
	CA 104–118 (peptide scan II)	
AK183	**CA 156–170 (peptide scan I)**	CA 156–164
	CA 153–166 (peptide scan II)	
	CA 156–158, 160, 164 (phage display)	

indicating that a conformational epitope might be recognised by this antibody. Only antibody 1.1E10 was found to react strain specifically binding to HIV-1 BH10 but not to NL4-3 derived MA proteins.

3.3.1.2. Epitope mapping by western blot analysis

In order to determine the epitope of antibody 2.4G3, which could not be identified by peptide scanning, western blot analysis of purified C-terminally truncated CA proteins was performed. The full-length CA protein (amino acids 1–231) as well as variants truncated after helix 1 (1–177), helix 2 (1–194), helix 3 (1–207) or helix 4 (1–220) of the C-terminal domain were purified and separated by SDS-PAGE. Autoradiographs of the subsequent western blot analysis performed with antibody 2.4G3 and polyclonal RαCA antiserum are shown in Fig. 3.25. As seen in the first two lanes, antibody 2.4G3 did not react with CA proteins lacking amino acids 194–207, while the general detectability of all truncated proteins was confirmed by reaction with the polyclonal antiserum, as

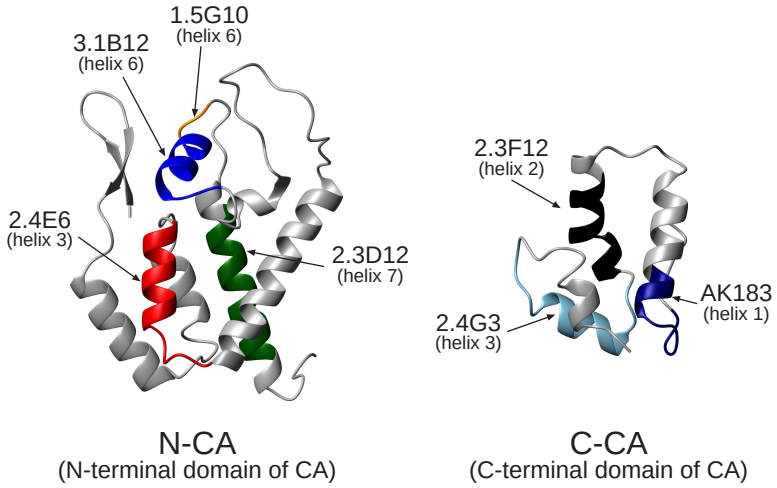

Figure 3.24.: Antibody epitopes highlighted in ribbon diagram of the structure of N-CA (pdb ID: 1GWP [41]) and C-CA (pdb ID: 1A8O [37]). Within N-CA the epitope of 2.4E6 (red, amino acids 49–62) overlaps with helix 3, epitopes of 3.1B5 (blue, 109–118) and 1.5G10 (orange, 109–120) with helix 6 and 2.3D12 (green, 129–143) overlaps with helix 7. Within C-CA the epitope of AK183 (dark blue, amino acids 156–164) overlaps with helix 1, 2.3F12 (black, amino acids 182–192) with helix 2 and 2.4G3 (light blue, amino acids 194–207) with helix 3. Structures were generated with MOLMOL [58].

depicted in the right part of Fig. 3.25. This suggests that part of the epitope of 2.4G3 might be located within amino acids 194–207.

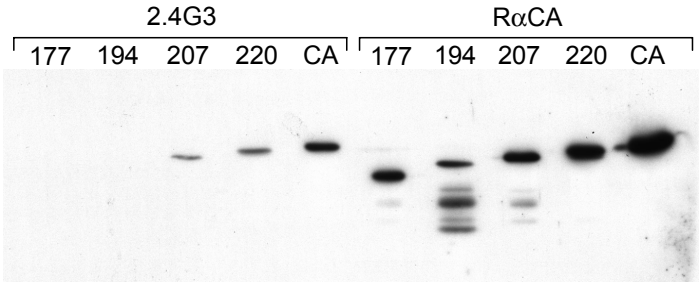

Figure 3.25.: Detection of C-terminally truncated CA proteins by antibody 2.4G3 in western blot analysis. 55 ng of the full-length CA protein as well as C-terminal truncated variants (CA177, CA194, CA207, CA220) were separated by SDS-PAGE. Western blot analysis was performed with antibody 2.4G3 and a polyclonal RαCA antiserum in parallel to ensure that truncated proteins are detectable in principal.

3.3.1.3. Epitope mapping by phage display

In an alternative approach, the phage display technology was tested as method for antibody epitope mapping. Antibody AK183 and the αPolio mAb were coated on Dynabeads to function as positive and negative target proteins for the biopanning reaction. Two different libraries based on M13 phages presenting either random linear 12mer or cyclic C7C peptide sequences fused to the N-terminus of the pIII coat protein were used for the biopanning reaction. After three rounds of positive and two negative selections a phage pool with a titer of 3.8×10^7 pfu/μl for the 12mer library and 7.7×10^6 pfu/μl for the C7C library was obtained. Of each pool 88 single phages were amplified and analysed for their ability to bind to AK183 or to the αPolio mAb coated on ELISA plates. Bound phages were detected with a αM13 HRP conjugated antibody. The wildtype M13 phage was used as negative control and did neither bind to AK183 nor to the α Polio antibody. In the case of the 12mer library, 77 phage clones bound specifically to AK183 but not to the Polio antibody. The 17 phages displaying the highest affinities were amplified for DNA preparation and sequencing. A conserved sequence KGPxxYP was found in all selected 12mer peptides as highlighted in Table 3.3. In the case of the C7C library, 71 phage cloned specifically bound to AK183 but not to the Polio antibody. The seven phages showing the highest affinity were sequenced and were found to all code for the same peptide, including the motif PYxPK, which seems to resemble the conserved motif found in the selected 12mer peptides (Table 3.3).

Based on these results it was assumed that the sequence KGPxxYP represents the epitope of antibody AK183. However, this sequence does not occur linearly within CA.

Table 3.3.: Shown are sequences selected with the C7C and 12mer phage libraries. Numbers indicate how often sequences were selected. The amino acid motif assumed to be the epitope is underlayed in gray.

sequence	× selected
IKGPYYYPFMEL	7
KGPALYPGLIFS	4
AKGPADYPHWHQ	2
HKGPMCTPSYCP	2
TWSKGPALYPSL	1
KGPKDYPLSLAH	1
CHPYWPKWC	7

Therefore a computer program (Michael Humbert, Georg-Speyer-Haus, Frankfurt, unpublished) was used, that screens the surface of a protein structure for a putative conformational epitope. Running the program with all selected 12mer sequences on the structure of C-CA (pdb file 1AUM, [37]), several possible solutions were found dependent on which peptide sequence was entered. Results included CA amino acids 156–169, 196–200 or 203–209. Amino acids 156–158 (KGP) and 164, 160 (YP) most likely are the correct epitope, since this sequence was found in all selected 12mer peptides and correlated with results obtained by peptide scanning. This finding indicates that within the sequence determined by peptide scanning (CA amino acids 156–166, coloured dark blue in Fig. 3.24) amino acids 156–158, 160, 164 are those most strongly involved in AK183 antibody binding.

3.3.2. Inhibition of *in vitro* assembly of spherical particles by monoclonal antibodies

Binding of a molecule to one of the CA interfaces involved in assembly might block an essential interaction and thus interfere with virus replication. Due to their inherent specificity and affinity mAbs targeted against HIV CA appeared to be suitable to serve as basis for the development of an assembly inhibitor. Therefore all nine mAbs were tested for inhibition of *in vitro* assembly of ΔMACANCSP2 using the above described system. The *in vitro* assembly reaction usually is started at pH 6, but antibodies 1.5G10 and 2.4E6 had been found not to bind to ΔMACANCSP2 at a pH lower than 7.3 [45]. Thus, antibodies were preincubated with a twofold molar excess of protein at pH 7.5 prior to starting the reaction at the same pH. ΔMACANCSP2 was shown to efficiently assemble under these modified conditions as seen on the electron micrograph in the left panel of Fig. 3.26. Under these conditions all analysed antibodies (except for the αPolio antibody) abolished *in vitro* assembly of ΔMACANCSP2. This is exemplified in the right panel of Fig. 3.26, showing an electron micrograph of the assembly reaction in the presence of antibody 2.4G3. As is clearly seen, no spherical particles were detectable. Since all antibodies tested abolished assembly with similar efficiency, this assay most likely cannot discriminate between specific blockage of an assembly interface and inhibition due to sterical hindrance. However, even if the mAbs seem to be unsuitable for the investigation of assembly inhibition, their pH dependent affinity has the potential to elucidate the mechanism of assembly and maturation as will be discussed below.

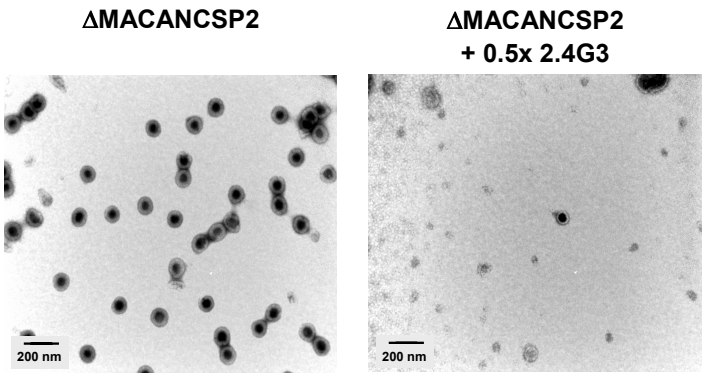

Figure 3.26.: Electron micrographs of negatively stained particles resulting from ΔMACANCSP2 assembly reactions at 46,000-fold magnification. 10 μM protein was used, either without or with 5 μM antibody 2.4G3.

3.3.3. Monoclonal antibodies as tool to detect conformational changes

As described above Gross et al. postulated that pH-induced conformational changes detected by monoclonal antibodies 1.5G10 and 2.4E6 might correlate with conformational changes induced *in vivo* upon maturation [45]. The binding of all available mAbs to ΔMACANCSP2 was analysed by sandwich ELISA experiments at pH 6 and pH 8, to determine if any other antibodies display pH dependent binding. ELISA plates were coated with a polyclonal rabbit anti-CA antiserum and then incubated with a mixture of ΔMACANCSP2 with decreasing concentrations of mAbs. Bound protein-mAb complexes were then detected with HRP-conjugated polyclonal anti-mouse antiserum. All incubation and washing steps were performed at pH 6 or pH 8, respectively. In Fig. 3.27, the absorption at 450 nm is plotted as a measure of antibody affinity over the antibody dilution. As seen by comparison of the right (pH 8) with the left (pH 6) panel, the affinity of all antibodies is somewhat reduced at pH 6, but antibodies 2.3D12 and 2.3F12 in addition to 1.5G10 and 2.4E6 showed a significant reduction in binding at pH 6.

Taken together, four mAbs are available that display pH-dependent affinities. However, the IgM antibodies 2.3D12 and 2.3F12 were excluded from further analysis to restrict the system to IgG antibodies for the sake of simplicity. In order to investigate whether the state of maturation of Gag correlates to antibody affinities, a detailed analysis was performed using only the pH sensitive antibodies 1.5G10 and 2.4E6 together with the pH-independent antibody 3.1B5, all of them binding to N-CA (see also Fig. 3.24).

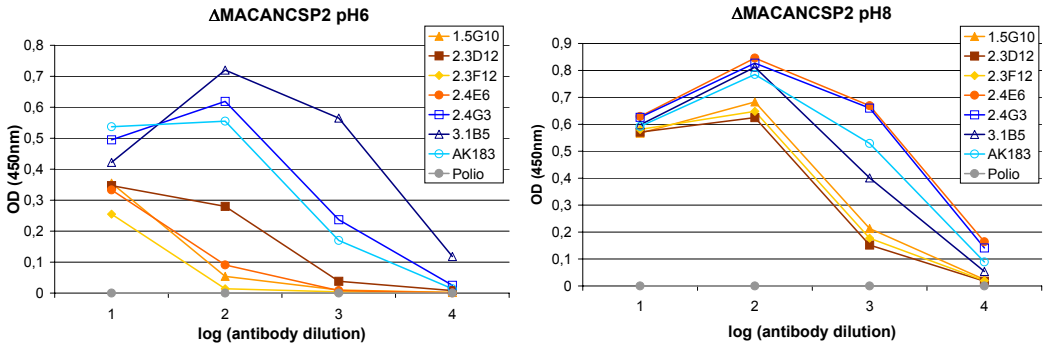

Figure 3.27.: Binding of mAbs to ΔMACANCSP2 was analysed by sandwich ELISA experiments at pH 6 and pH 8 as described in the text. Plotted is the absorption at 450 nm as a measure of antibody affinity over the antibody dilution. Antibodies displaying reduced binding at pH 6 are coloured in tones of red, the remaining antibodies are coloured in tones of blue.

These antibodies were used for sandwich ELISA assays at pH 6 and pH 8 together with a set of C- and N-terminally truncated Gag proteins. These proteins were chosen in order to determine if other regions of Gag might have an influence on antibody binding to N-CA and if cleavage at the N- or the C-terminus of CA affects antibody affinities. The chosen proteins are depicted schematically in Fig. 3.28.

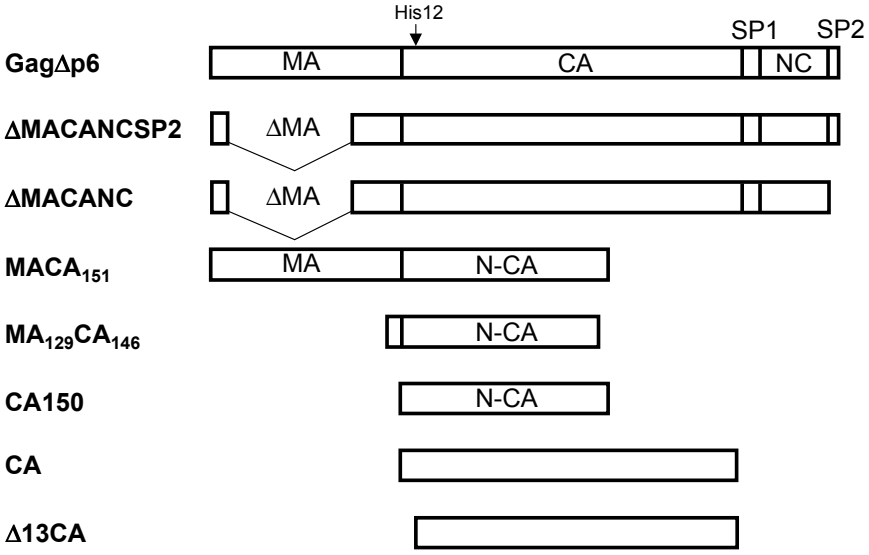

Figure 3.28.: Schematic representation of Gag-derived proteins used to analyse pH dependent conformational changes.

3.3.3.1. The C-terminal part of Gag has no influence on antibody binding

ELISA assays were performed with GagΔp6, ΔMACANC as well as $MACA_{151}$ (Fig. 3.28), in order to investigate the influence of Gag domains other than N-CA on binding of antibodies 1.5G10, 2.4E6 and 3.1B5. These proteins as well as $MA_{129}CA_{146}$ were obtained from our collaborator Wesley I. Sundquist (University of Utah, Salt Lake City, USA). In Fig. 3.29 antibody affinities are plotted over the antibody dilution as described before. The comparison of right (pH 8) and left (pH 6) panels shows that

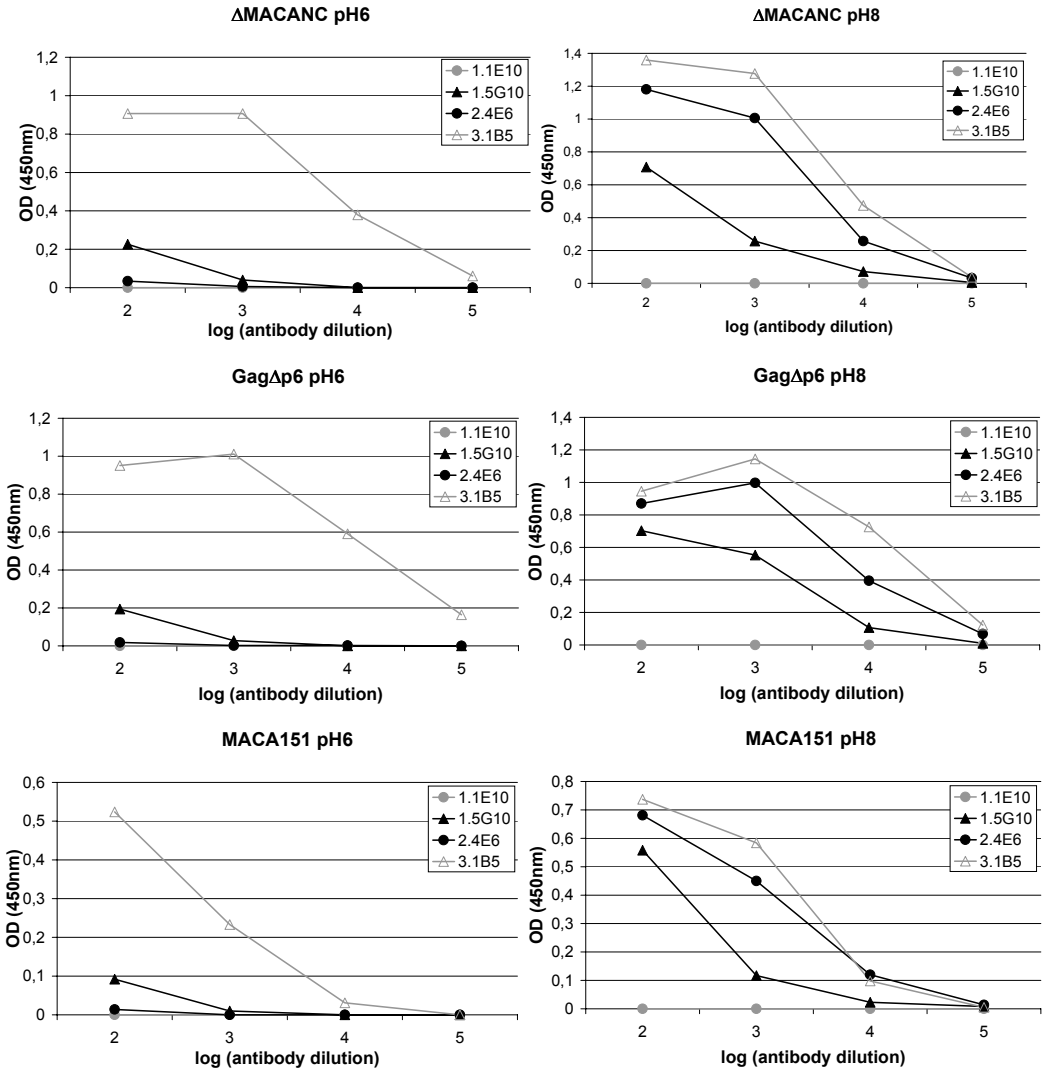

Figure 3.29.: Binding of mAbs to ΔMACANC, GagΔp6 and $MACA_{151}$ was analysed by sandwich ELISA experiments at pH 6 (left) and pH 8 (right). Plotted is the absorption at 450 nm as a measure of antibody affinity over the antibody dilution.

affinities of 1.5G10 and 2.4E6 for all three proteins at pH 6 were strongly reduced as compared to pH 8. The affinity of antibody 3.1B5 remained unchanged. This indicates, that neither the deletion of MA amino acids 16–99 nor the truncation of the whole C-terminal part of Gag including C-CA, SP1, NC and SP2 had an influence on pH dependence of 1.5G10 and 2.4E6 binding.

3.3.3.2. The CA protein shows a different pattern of antibody binding

According to the hypothesis by Gross et al., the conformation of the N-terminally unprocessed proteins (GagΔp6, ΔMACANC and $MACA_{151}$) at pH 6 reflects a mature conformation, indicating that the mature CA protein should show the same or a similar pattern of antibody binding at pH 6. In order to test this hypothesis ELISA experiments were performed with the mature CA protein as well. As seen in Fig. 3.30 A, antibody binding to CA was not dependent on pH and antibody 1.5G10 displayed high affinity for CA whereas antibodies 2.4E6 and 3.1B5 did not bind. This pattern was partly unexpected and different from that observed for the proteins analysed before as seen by comparison of this figure with Fig. 3.29: Antibody 2.4E6 did not bind to CA, nor to ΔMACANC at pH 6. This part is consistent with the hypothesis that pH-induced changes in helix 3 are also detectable upon cleavage at the N-terminus of CA. The situation seems to be more complicated for the epitopes located in helix 6: Antibody 3.1B5 bound independent of pH to N-terminally unprocessed proteins but not to CA. Antibody 1.5G10 bound to N-terminally unprocessed proteins only at pH 8 but displayed pH-independent binding to CA.

An influence of C-CA on antibody affinities in the context of the mature CA protein was ruled out by the analysis of N-CA alone (CA150). As is evident from the comparison of panel A and B of Fig. 3.30, there is no difference in the pattern of antibody affinities between the full-length CA protein and CA150.

3.3.3.3. Conformational changes are not dependent on β-hairpin formation induced upon cleavage

The differences seen between N-terminally immature proteins and mature CA indicated, that cleavage at the N-terminus of CA induces changes in the same regions as induced by a pH reduction. The underlying mechanism seemed likely to be explained by the well described β-hairpin formation induced upon maturation *in vivo*: The NMR structure of the N-terminally unprocessed protein $MACA_{151}$ showed a flexible and

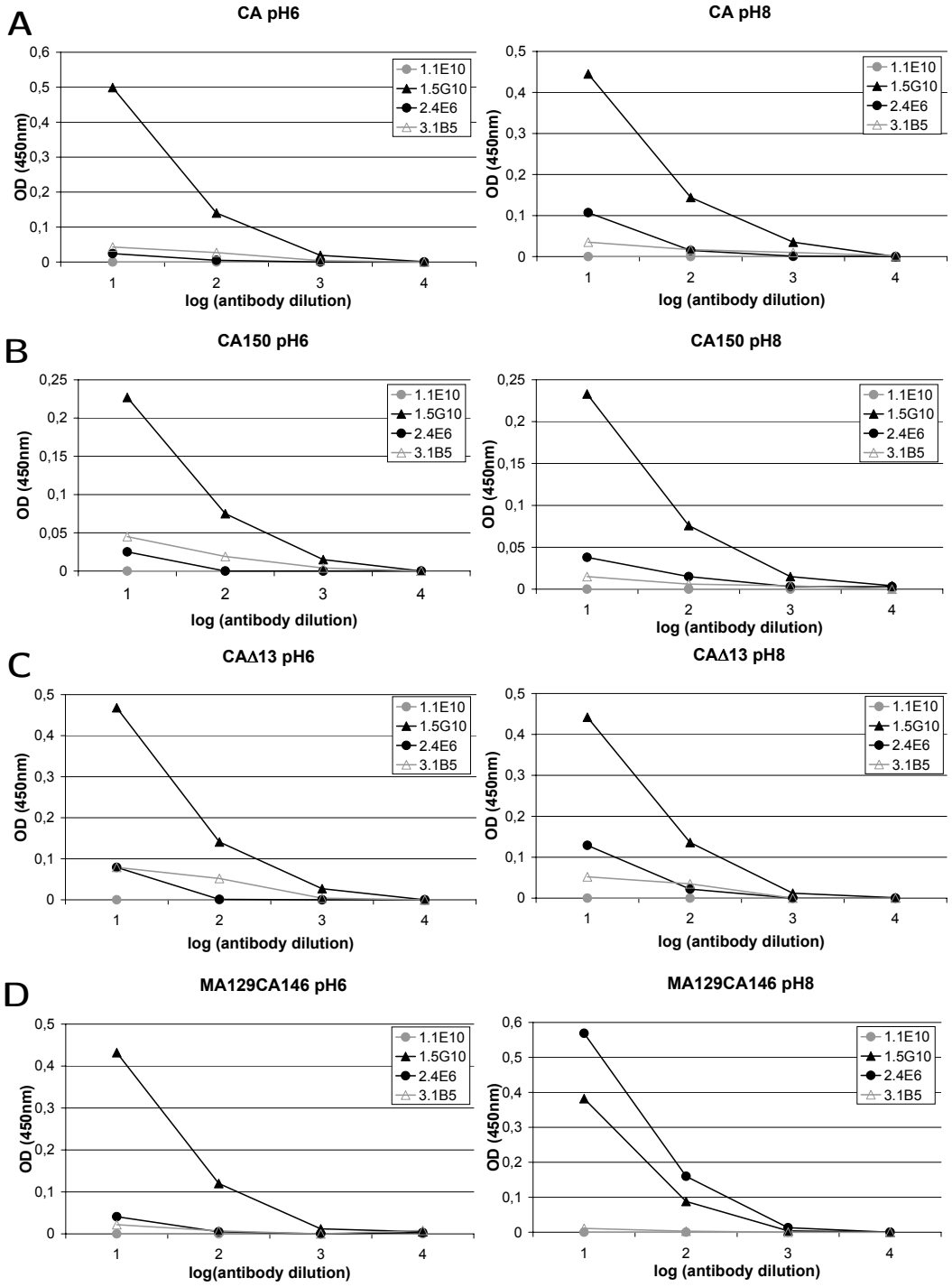

Figure 3.30.: Binding of mAbs to CA (A), CA150 (B), Δ13CA (C) and MA_{129} CA_{146} (D) was analysed by sandwich ELISA experiments at pH 6 and pH 8. Plotted is the absorption at 450 nm as a measure of antibody affinity over the antibody dilution.

unstructured linker between MA and CA [89]. However, upon cleavage a β-hairpin stabilised by a salt bridge between Pro1 and Asp51 forms at the N-terminus of the mature CA protein [41]. This hairpin cannot form in the unprocessed protein and was therefore postulated to present a conformational switch [91]. In addition to that, NMR data obtained for the protein $MA_{129}CA_{146}$ indicated the possibility of an intermediate hairpin structure, stabilised by a salt bridge between His12 and Asp51 (Wesley I. Sundquist, personal communication). Since His12 was also shown to deprotonate at pH 7 (Wesley I. Sundquist, personal communication), it was a likely candidate to trigger the pH-induced conformational changes observed here. Taken together, these findings led to the following hypothesis: The linker region is unstructured at pH 8 but adopts an intermediate His12-Asp51 stabilised conformation at pH 6. Upon maturation, the Pro1-Asp51 stabilised hairpin would form. These changes were hypothesised to correlate with the observed different patterns of antibody affinities.

In order to test this hypothesis, a CA protein lacking the first 13 amino acids and therefore being deficient in hairpin formation (CAΔ13) was analysed for mAb binding using ELISA assays as described before. If the hairpin were involved in the observed conformational changes, this protein should show a pattern different from CA. However, as seen by comparison of panels A and C in Fig. 3.30, there was no difference between CA and CAΔ13, indicating that hairpin-formation does not influence the antibody affinities. Also, the mutant protein GagΔp6-H12A was analysed for antibody binding by ELISA. This protein should have lost the pH dependence of antibody binding, if His12 functioned as pH sensor. As seen by comparison of GagΔp6-H12A (Fig. 3.31) with results for GagΔp6 shown in Fig. 3.29, there was no difference in the antibody binding pattern of both proteins.

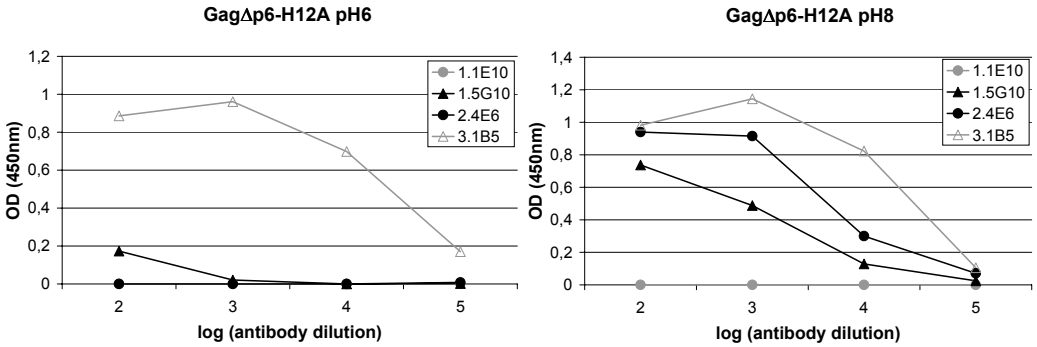

Figure 3.31.: Binding of mAbs to GagΔp6-H12A was analysed by sandwich ELISA experiments at pH 6 and pH 8. Plotted is the absorption at 450 nm as a measure of antibody affinity over the antibody dilution.

These results indicated that neither the β-hairpin formation nor amino acid His12 influences the observed conformational changes and the antibody affinities. However, the protein $MA_{129}CA_{146}$, which showed the His12-Asp51 salt bridge in NMR studies, indeed displayed an intermediate pattern (Fig. 3.30, D): At pH 6 the fully mature conformation was observed as for CA (comparison of all left site graphs in Fig. 3.30), but antibody 2.4E6 displayed high affinities at pH 8 as was observed for immature proteins (comparison of the right graph of panel D with the right graph of panel A, Fig. 3.30 and with the top right graph in Fig. 3.29).

In summary, pH dependent changes observed in helix 3 were also detected upon cleavage between MA and CA. Changes in helix 6 were shown to be influenced by the presence of MA residues as well as by pH. None of the observed changes were influenced by β-hairpin formation and the amino acid His12 is not the pH sensor that triggers these changes.

4. Discussion

4.1. The influence of the CA-SP1 region on *in vitro* assembly

In vivo HIV-1 Gag proteins accumulate at the plasma membrane to form a budding structure. Stable, immature virus particles are then released from the cell, displaying a spherical layer of about 5,000 Gag proteins [11] below the membrane. The formation of the mature conical core is induced by the processing of Gag. Upon infection the conical core is released into the cell. Possibly, the core remains stable during reverse transcription but has to disassemble at some point to release the viral genome to be transported into the nucleus for integration. This indicates, that in contrast to the stable, immature virus particle, the mature conical core has to be a metastable structure. The hypothesis that core stability has to be precisely balanced was confirmed by mutational analyses that identified a region within the N-terminal domain of CA that influences core stability [32, 28]. Additionally, Gag-RNA interactions were shown to significantly stabilise immature particles [77]. Recently, the protein lattice forming the immature particle was shown to be more densely packed than the lattice forming the mature core, also suggesting a difference in stability [11]. In accordance with a looser lattice packing, the mature core seems to be formed by only less than half of the CA proteins (1,000 to 1,500 molecules [65, 13]) available in the virus particle after cleavage.

Here, I analysed the influence of the C-terminus of CA as well as of SP1 on shape and efficiency of tubular particle formation, since mutations within this region were shown to influence virus assembly [92, 1, 74]. The result of the analysis of the proteins CASP1, CA and CA220 indicated that removal of the structurally disordered [100] last 11 residues of CA as well as SP1 significantly enhanced the efficiency of the assembly reaction and the regularity of the formed particles. The protein CA220 that lacks the flexible regions formed more regular particles with higher efficiency than CA. Thus, it can be hypothesised that too much flexibility as in CASP1 prevents core formation,

but a bit of flexibility might be retained in CA in order to achieve the formation of metastable particles.

The increased regularity of particles formed from CA220 is of interest for another question as well. So far, the only model available for the structure of the mature conical core is based on low resolution image reconstructions of Cryo-EM pictures of tubular particles assembled from CA [65]. The limitations of the proposed model were demonstrated by recent studies suggesting the formation of alternative interfaces [60, 76]. A higher resolution might be achieved using the more regular CA220 particles for cryo-EM reconstruction. In collaboration with the group of Stephen Fuller (Wellcome Trust Centre for Human Genetics, Oxford, UK) the Cryo-EM reconstruction of the tubular particles assembled of CA220 will be pursued, in order to obtain a more reliable structural model that might allow inferences on assembly interfaces.

After having achieved an increase in regularity and efficiency by truncating the flexible C-terminus, the additional removal of helix 4 of C-CA significantly changed the shape of assembled particles. Thin helical rather than tubular structures were formed. This indicates either that profound structural rearrangements are induced upon truncation or that helix 4 is important for assembly as had previously been suggested by mutational analysis [53]. In contradiction to that, a more recent analysis suggested that this region is not involved in any CA assembly interface [60]. Thus, it is postulated that conformational changes within the protein lead to the altered assembly phenotype.

4.2. A peptide inhibitor of HIV-1 Assembly

The phage display technology was successfully used to select a 12mer peptide that interacts specifically with the C-terminal domain of CA (C-CA) and efficiently inhibits HIV assembly *in vitro*. Originally, the phage display technology was developed for the selection of high-affinity ligands of receptors or antibodies. In contrast to these molecules, proteins that are not designed for interactions of nanomolar affinity are unlikely to have highly active surfaces. Additionally, it was observed that only a limited number of sites on a protein surface is available for interaction with phage-display-selected peptides [51]. Therefore it is not surprising that most of the phages selected with C-CANC bound to the same region within C-CA, although it certainly was not expected to select essentially the same sequences with CA as target protein. This implies that the surface of N-CA, at least in the context of CA, does not provide high affinity binding sites for 12mer peptides that were present in the used library. This

library was of relatively low complexity, comprising 2.7×10^9 sequences in contrast to the theoretically possible 4×10^{15} sequences. Moreover, sequences might have been selected prevalently that do not disturb or may even enhance phage infectivity, which is mediated by the pIII coat protein.

Within C-CA one major binding site seems to exist as indicated by the prevalence in selection of highly homologous peptides using two different target proteins. The observed sequence similarities between these peptides allow the identification of residues that might be important for the interaction with C-CA. The N-terminus was more conserved within group 1 peptides and generally seemed to be more important for the interaction with CA than the C-terminus, which could be modified without loosing the peptide function. Hydrophobic amino acids as well as an acidic pI value are features that are best conserved within group 1 peptide sequences. A helical wheel representation of the *CAI* sequence is depicted in Fig. 4.1. This suggests that the peptide forms an amphipathic helix with a hydrophobic site that is very well conserved in group 1 peptides. The intrinsic helicity as calculated by the computer program AGADIR (http://www.embl-heidelberg.de/services) was negligible for all peptides, but might be induced upon CA binding.

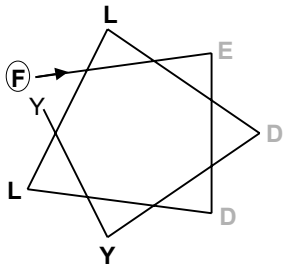

Figure 4.1.: Helical wheel presentation of part of the sequence of *CAI*: Amino acids FEDLLDYY correspond to the heptad repeat positions *abcdefg*. The first amino acid (*a*) of the first heptad repeat is encircled. Hydrophobic residues are coloured black, negatively charged residues are coloured grey. The hydrophobic site as well as the negative charge in position *c* is very well conserved within group 1 peptides.

4.2.1. *CAI* binding site within CA

NMR analysis was performed using the dimerisation defective protein C-$CA_{W184A/M185A}$. Backbone chemical shift perturbation measurements indicated that the region of amino acids 169–191 is involved in the interaction of *CAI* with C-$CA_{W184A/M185A}$. This region includes eleven residues shown to be involved in the CA

dimer interface [5]. However, *CAI* binding does not abolish dimerisation of C-CA. This result, together with the pattern of chemical shift perturbations, indicated that *CAI* binds to the side of helix 1 and 2 that is opposite to the dimer interface. This also suggests that dimer interface side chains 184 and 185 are not part of the peptide binding site. However, the affinity of *CAI* was shown to be significantly reduced for C-CA upon mutation of W184 and M185. This reduction in affinity presumably does not directly reflect the mutation of the side chains. Instead, the beginning of helix 2 (amino acids 180–185) seems to be conformationally labile in the mutant protein, but was shown to be stabilized by *CAI* binding. Thus, binding energy is likely to be expended on the stabilisation of the helix leading to a reduced overall affinity of *CAI* for C-CA$_{W184A/M185A}$. This conformational stabilisation might contribute to the observed chemical shift perturbations within amino acids 169–191. However, the peptide contains three aromatic residues, which are likely to induce large shifts in the binding site. The fact that these residues are conserved within all group 1 peptides indicates that they are involved in interactions with C-CA and thus should contribute most to the observed chemical shift perturbations. Taken together, the *CAI* binding site can be postulated to be located at the side of the protein opposite to the dimer interface including amino acids 169–191. This region comprises the C-terminus of helix 1, the interhelical linker and helix 2.

Although the deletion of amino acids 207–220 almost abolished phage-*CAI* binding to CA in ELISA experiments, this region is postulated not to form an essential part of the binding site because of the absence of significant chemical shift perturbations in NMR experiments. Thus, deletion of helix 4 of C-CA might induce structural rearrangements that influence *CAI* binding to helix 1 and 2. This hypothesis is supported by the finding that the protein lacking helix 4 of C-CA has lost the ability for tubular particle formation *in vitro*, as described in section 3.1. An alternative explanation to reconcile results from ELISA and NMR experiments would be the following: As can be derived from the preliminary crystal structure of the complex (Fig. 3.21), *CAI* binds to helix 1 and 2 but is located in proximity to helix 4 as well. It is conceivable that single side chains might stabilise the complex. The influence of side chains 207, 212, 213 and 219 was ruled out already by mutational analysis. However, the preliminary crystal structure indicates that Leu 211 is part of the groove that forms the peptide binding site and might thus stabilise the interaction. Such a stabilising side-chain interaction would be hypothesized to induce no significant backbone shift perturbations, but might account for resonances displaying subtle shifts in the fast exchange regime as observed in that region. The involvement of amino acids 207-220 in the interaction will be ruled

out by the refined crystal structure of the complex of *CAI* with C-CA.

The above described inferences based on NMR and biochemical analyses are strongly supported by the preliminary crystal structure of the complex of *CAI* with C-CA that has been obtained recently by our collaborators. The crystallographic interface matches very well the pattern of backbone chemical shift perturbations observed by NMR, indicating that the peptide binds to the groove formed by helix 1 and helix 2 at the side of the protein opposite to the dimer interface. Dimerisation of that complex does not seem to be impaired. The structure also indicates that the N-terminus of the peptide is strongly involved in the interaction while the C-terminus does not form contacts with C-CA, which is consistent with the finding that the C-terminus, but not the N-terminus, can be labelled without loosing functionality. Thus, it can be postulated that the peptide blocks an assembly interface that involves helix 1 and 2 of C-CA and is different form the dimer interface.

4.2.2. The mechanism of assembly inhibition

Assembly of immature- as well as mature-like particles was shown to be efficiently inhibited by *CAI*. Since the structural composition of *in vitro* assembled particles is very similar to that of immature viruses and the mature conical core, respectively [44, 45, 13, 11]. Hence, an interface important for virus assembly as well as maturation must be targeted by *CAI*.

As mentioned above the peptide binds on the side opposite to the dimer interface indicating that its inhibitory effect is exerted by blocking an assembly interface different from the dimer interface. Few indications are found in the literature that the peptide binding site region might represent an assembly interface. An insertion of five amino acids at CA residue 170 was shown to significantly abolish virus release [81]. Point mutations of CA amino acids that might be involved in the interaction as deduced from the preliminary crystal structure have not been included in the mutational analysis of CA surface residues [92] since they are buried in a cleft formed by helix 1 and 2.

Recently a novel N-CA/C-CA interaction has been described, which was shown to be present in CA *in vitro* assemblies but not in the unassembled protein. The interface involves amino acids 55–68 of N-CA. In the C-terminal domain a region comprising the C-terminus of helix 1 through the N-terminus of helix 2 was suggested to be part of the interface [60]. This region would perfectly match the region affected by *CAI* binding. However, an interface including amino acids 55–68 was detected only in

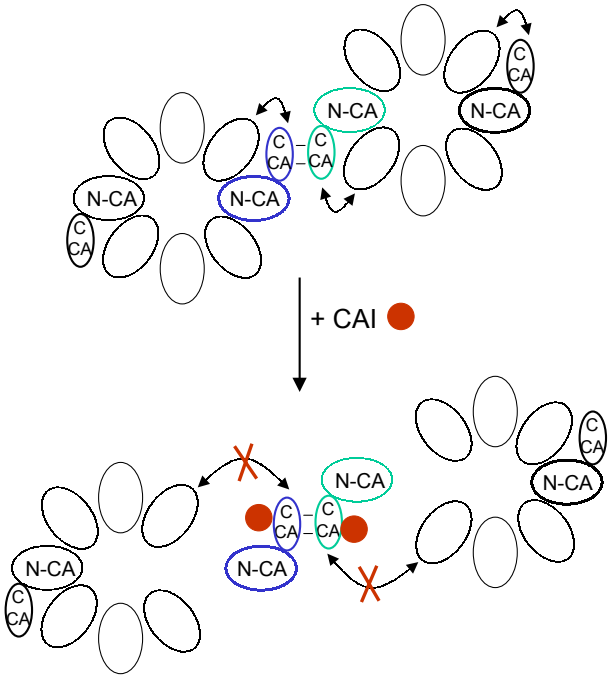

Figure 4.2.: Potential inhibition of the N-CA/C-CA interaction based on the model presented by Lanman et al. [60]. The hexameric arrangements of CA N-terminal domains (N-CA) are depicted. For clarity only few C-terminal domains (C-CA) are shown that link two adjacent hexamers (indicated by double lines). The N-CA/C-CA interaction is indicated by arrows. Upon addition of peptide, this interaction might be blocked thereby abolishing hexamer formation.

mature but not in immature virus-like particles [61] indicating that this interaction is characteristic of mature particle assembly. This was also supported by the inhibitor *CAP-1*, which targets the N-CA region involved in the interaction and specifically inhibits cone formation but not virus release [87]. Since *CAI* appears to block an interface equally important for immature and mature particle assembly, this N-CA/C-CA interaction appears to be unlikely to present the major inhibitory mechanism of *CAI*. However, it is conceivable that the peptide binding site is involved in slightly different interfaces, each characteristic for immature or mature particle assembly. Thus, a schematic representation of a model for the inhibition of this N-CA/C-CA interaction in the mature particle is depicted in Fig. 4.2. This is based on the model by Lanman et al. [60]. The hexameric arrangements of CA N-terminal domains (N-CA) are depicted, with C-terminal domains (C-CA) linking two adjacent hexamers (indicated by double lines). The N-CA/C-CA interaction is indicated by arrows. Upon addition of peptide, this interaction might be blocked and the hexamers, which are the building blocks of

mature HIV particles could not form properly. However, this model is on somewhat shaky foundations since the region of C-CA involved in the N-CA/C-CA interaction is still to be proven and the model would only explain the inhibitory effect of *CAI* on mature particle assembly.

Apart from the model described above, *CAI* might block a yet undetermined assembly interface. Alternatively, it might exert its inhibitory effect by prohibiting certain orientations of N-CA to C-CA that are important for assembly or by locking the domains in an unfavourable orientation. Generally, an assembly inhibitor could either function by reduction of the amount of protein available for assembly or by inactivation of assembly intermediates by blocking or misdirecting the assembly process. It is evident that the first mechanism requires nearly stoichiometric concentrations of inhibitor and target protein whereas in the second case a single inhibitor molecule might inactivate a large number of target proteins because they might be trapped in an assembly intermediate or assemble into aberrant particles. Such a scenario had been suggested for a small molecule inhibitor of hepatitis B virus assembly [105]. In the case of *CAI*, the high efficiency of assembly inhibtion *in vitro* as well as the increased formation of aberrant structures detectable by EM in the presence of *CAI* might indicate that the peptide indeed misdirects assembly. However, it was not analysed if the aberrant structures are indeed misdirected or trapped assembly products or if they are just protein aggregates of undefined nature. Additionally, first hints were obtained that assembly inhibtion by *CAI* might be irreversible. Both aspects, assembly misdirection or trapping of intermediates as well as irreversible inhibition, would significantly increase the validity of *CAI* for drug development, if they held true.

A potent inhibitor should display sufficiently strong binding affinities to disrupt the interactions driving assembly. ITC measurements indicated that the dissociation constant (K_D) for the interaction of *CAI* with the dimerisation defective protein C-$CA_{W184A/M185A}$ is in the range of 10–30 μM. This was confirmed by the determination of the K_D based on NMR chemical shift perturbations yielding a value of $15.0 \pm 7.2\,\mu$M. Since the affinity of *CAI* was shown to be significantly better for the wild-type C-CA or CA protein, it can be assumed that the interaction leading to inhibition of assembly displays a K_D of significantly less than $15.0 \pm 7.2\,\mu$M. Thus, the affinity of the peptide-protein interaction is likely to be stronger than the affinity of the CA dimer ($K_D = 18\,\mu$M) [37]. Since multiple interactions, presumably of even weaker affinity than dimerisation, seem to drive HIV assembly, a K_D in the low μM range is likely to be good enough for efficient inhibition. This is consistent with the high efficiency of inhibition that was observed in the *in vitro* system. Additionally, the molecule *CAP-1*, which

was shown to inhibit HIV-1 maturation by interacting with the N-terminal domain of CA displayed only a very weak affinity with a K_D of about 800 μM. Nevertheless, this compound efficiently inhibited maturation.

Another property of the *CAI*/CA interaction advantageous for drug development is the fact that the region within CA that is targeted by *CAI* is highly conserved among all HIV-1 subtypes listed in the Los Alamos database. This indicates that a drug based on *CAI* might have a broad application range and more importantly resistance mutations might not evovle as fast.

As described in the introduction, so far all approaches to develop HIV assembly inhibitors have failed to target the formation of immature particles [87, 64]. In contrast to that, *CAI* inhibits both, assembly of immature-like as well as mature-like particles *in vitro*. Thus, *CAI* is likely to combine the inhibition of two steps of the virus life cycle in one single molecule. It was suggested recently that HIV maturation consists of disassembly of the spherical CA shell followed by de novo assembly of the conical core particle [13, 12]. Accordingly, *CAI* can be assumed to first reduce virus release by abolishing Gag-Gag interactions and subsequently target escaping particles. Nascent virions are likely to incorporate *CAI*, since the affinity of the interaction is in the same range as described for the interaction with CypA [101], which is incorporated into virus particles as well [33]. Additionally, the above mentioned small molecule inhibitor of HIV maturation (*CAP-1*) displays 10-fold weaker affinity for N-CA [87] and still must be incorporated into virus particles in order to exert its inhibitory effects. Within the released particles *CAI* should then inhibit the formation of the mature conical core and inactivate residual particles.

Taken together, the peptide inhibitor (*CAI*) presented here is the first HIV inhibtor targeting immature particle assembly and may represent the first inhibitor that could block virus replication in two consecutive steps resulting in increased efficacy. *CAI* meets all the requirements to form the basis for anti-viral drug development, since it binds specifically and with reasonable affinity to a highly conserved region within HIV-1 CA. In order to replace the peptide by a small molecule that should be more suitable as a drug and for cell culture analyses, the established multiwell ELISA assays based on phage-*CAI* will be used for high-throughput screening of compound libraries. The high-resolution crystal structure of the C-CA/*CAI* complex, which will be available soon, will allow for the rational design of peptidomimetics.

4.3. Conformational changes accompanying virus maturation

A panel of monoclonal antibodies (mAbs) directed against HIV-1 CA was used here to elucidate structural changes occurring upon virus maturation. This section will describe the characterisation of the mAbs with respect to epitope mapping and their potential to inhibit assembly *in vitro*. This will be followed by a detailed discussion of conformational changes accompanying virus maturation as detected by the mAbs.

4.3.1. Epitope mapping and assembly inhibition

The epitopes of a set of mAbs were mapped using the peptide scan technology. In the case of five antibodies, previously determined epitopes were confirmed by our results. For two antibodies (3.1B5, AK183) previous mapping attempts using an ELISA approach had failed: No reactivity had been detectable with CA derived peptides coated on ELISA plates. Using the peptide scan technology, these epitopes were successfully defined, which is most likely explained by an increased accessibility of peptides synthesised on membranes as compared to peptides adsorbed on ELISA plates.

One antibody (2.4G3) is likely to recognise a conformational epitope, since the binding site could not be determined by either of these methods. However, antibody 2.4G3 still binds to CA under western blot conditions, and binding was only abolished for CA proteins lacking amino acids 194–207. Either this region is part of the epitope or protein truncation induces structural rearrangements that prevent antibody binding to another site. To unambigously define the epitope of 2.4G3 an appoach should be used that allows the determination of conformational epitopes. With the aim of establishing such a method, a phage display screening was performed using the previously mapped antibody AK183 for evaluation. The aim was to select those peptides from random libraries that comprise the epitope sequence. An inherent problem of such an appoach is the fact that most sequences selected do not linearly occur in the target protein but are mimotopes or even improved epitopes. Simple linear sequence similarity searches will fail in these cases. The only solution to this problem is the use of computer programs to search the surface of the structure of the target protein for putative epitopes. In our case, the structure of the target protein C-CA as well as a a suitable computer program (Michael Humbert, Georg-Speyer-Haus, Frankfurt/Main, unpublished) were available. However, without having determined the epitope region by peptide scanning before,

it would not have been possible to unambiguously use the output of the computer program to define the epitope.

Taken together, the peptide scan technology turned out to be a very powerful technique for the mapping of linear epitopes, whereas phage display as tool to map linear and conformational epitopes will most probably only work with the target protein structure at hand and a rough knowledge about the location of the epitope.

Additionally, antibodies were tested for their potential to inhibit assembly *in vitro*. All used antibodies that bind to ΔMACANCSP2, regardless of their epitope location, efficiently inhibited assembly *in vitro*. This indicated an indirect, sterical effect on assembly. Sterical hindrance is very likely to occur upon binding of a 150 kD IgG mAb to the 42 kD ΔMACANCSP2 protein. Such an indirect effect had also been suggested for a mAb that blocks receptor binding of the Adeno associated virus (AAV) [99]. A possibility to achieve more specific effects would be the usage of Fab fragments. However, Fab fragments are still half the size of Gag and thus likely to retain the problem of sterical hindrance. Alternatively, in cases antibody epitopes are located in regions involved in assembly, structural analysis could lead to the rational design of an inhibitor.

4.3.2. Helix 3 and Helix 6 of N-CA are involved in conformational changes during Gag maturation

In order to test the hypothesis that conformational changes that are triggered by Gag processing *in vivo* might be mimicked by a change in pH *in vitro* [45], a detailed analysis was performed. Three mAbs were analysed for detecting pH-induced conformational changes in Gag-derived proteins. Antibody 2.4E6 binds to helix 3 of N-CA, 1.5G10 and 3.1B5 bind to helix 6 of N-CA (Fig. 3.24). The results of these experiments will be discussed below, mainly focusing on the following points: the correlation between cleavage of Gag and pH-induced conformational changes, the underlying mechanism triggering these changes and their relevance for virus maturation.

4.3.2.1. Conformational changes can be triggered by pH and MA-CA cleavage

Several Gag-derived proteins were analysed with three mAbs, in order to investigate whether the observed pH dependent antibody affinities for the protein ΔMACANCSP2 are intrinsic to N-CA or are influenced by other Gag domains. The pattern of antibody

affinities was not influenced by the deletion of the whole C-terminal part of Gag (C-CA, SP1, NC, SP2), indicating that the observed conformational changes do not correlate with cleavage between CA and SP1.

In contrast to that, the presence of MA indeed had an effect on the conformation of N-CA. The comparison of the mature CA protein with proteins including either full-length MA, MA with a deletion of amino acids 16-99 (ΔMA) or the last four amino acids of MA revealed that several intermediate conformations can be obtained. This indicated the need to consider the conformational states of helix 3 and helix 6 separately. Antibody 3.1B5 bound to helix 6 independent of pH, but only in the presence of MA or ΔMA, indicating that this antibody differentiates between states of MA-CA processing. In the presence of full-length MA or ΔMA antibody 1.5G10 binding was lost at pH 6, whereas the antibody bound independent of pH to the remaining proteins. Thus, MA has an influence on the accessibility or conformation of helix 6, and in case of antibody 1.5G10 this effect seems to be superimposed by a pH effect that is also dependent on MA. This led to the hypothesis that helix 6 is mainly effected by the presence of MA. It was shown that several mutations in helices 4 to 6 induce a defect in Gag assembly but did not influence CA assembly *in vitro* [92]. This suggests that only in the unprocessed protein, helix 6 adopts a conformation that is important for particle assembly.

The binding properties of antibody 2.4E6 seem to present a link between pH and cleavage induced changes, because it neither binds to the mature CA protein nor to N-terminally unprocessed proteins at pH 6, but it binds well to all unprocessed proteins at pH 8. In the case of this antibody, the hypothesis that immature proteins adopt a mature conformation at pH 6 is supported, and indicates that the pH trigger induces similar changes in helix 3 as induced by processing of Gag at the N-terminus of CA.

In summary, different conformations of N-CA were observed that can be characterised by different affinities of the three mAbs. 2.4E6 differentiated between mature and immature conformations induced by pH as well as by cleavage. 3.1B5 differentiated between processing states of MA-CA and 1.5G10 displayed a combination of effects. These results indicated the complexity of the system and emphasised that although only two helices of the protein are seen by our antibodies, the three antibodies displayed different effects. Additionally, the epitopes of antibodies 1.5G10 and 3.1B12 largely overlap and still both antibodies display different binding properties. To verify the observed results and to gain knowledge about the remaining regions of N-CA it will be necessary extend the panel of antibodies used for this analysis.

4.3.2.2. Conformational changes are not dependent on β-hairpin formation

As described above, the removal of MA residues upon cleavage seems to induce changes in helix 3 that can be correlated to pH-induced conformational changes. How the same changes could be triggered by pH as well as by cleavage remains an open question. The β-hairpin formation, induced upon cleavage between MA and CA, seemed to be the most obvious candidate: The NMR structure of the N-terminally unprocessed protein $MACA_{151}$ showed a flexible and unstructured linker between MA and CA [89]. However, upon cleavage a β-hairpin stabilised by a salt bridge between Pro1 and Asp51 forms at the N-terminus of the mature CA protein [41]. Since this hairpin cannot form in the unprocessed protein, it was postulated to function as switch between immature and mature conformation [91]. In addition to that, NMR data obtained for the protein $MA_{129}CA_{146}$ indicated the possibility of an intermediate hairpin structure, stabilised by a salt bridge between His12 and Asp51 (Wesley I. Sundquist, personal communication). Since His12 was shown to deprotonate at pH 7, this titrable amino acid seemed likely to represent the link between pH and cleavage induced conformational changes, leading to the following hypothesis to explain the observed patterns of antibody affinities: The MA-CA linker region is unstructured at pH 8 but adopts an intermediate His12-Asp51 stabilised conformation at pH 6. Upon maturation, the Pro1-Asp51 stabilised hairpin would form. In order to test this hypothesis a haiprin deficient CA variant (Δ13CA) as well as a His12Ala mutant were analysed for antibody binding: No difference was detectable between CA and Δ13CA, suggesting that hairpin formation is not involved in the conformational changes detected in helix 3 and 6. This is consistent with the finding that the protein Δ13CA retained the ability to assemble into tubular particles *in vitro* [43]. Taken together, hairpin formation does not seem to be the conformational switch that triggers the transformation of the whole protein into the mature form and it does not affect the conformational changes seen for helix 3 and 6. Additionally, immature Gag-derived proteins with the His12Ala mutation still displayed pH dependent binding, indicating that His12 does not function as pH sensor.

The pH sensor that triggers the changes in helix 3 has to be located within N-CA, but seems to be functional only in the unprocessed protein, since antibody 2.4E6 did not display pH sensitivity in the mature CA protein, but in the presence of at least four (non-titrable) amino acids of MA. Histidines are the best candidates to translate a pH reduction from 8 to 6 into conformational changes, since their theoretical pK_a-value is 6.04. This value would only have to be shifted to about pH 7, as was found to be the case for His12 (Wesley I. Sundquist, personal communication), to trigger the

conformational changes observed here. After having ruled out the influence of His12, four histidines within N-CA remain as candidates for the pH sensor. Two of those are located close to helix 3 (His62 in front of helix 4 and His84 behind helix 4). However, since no additional evidence supports the influence of these histidines and since in principle pK_a values of glutamate and aspartate could be shifted to pH 7 as well, it seemed difficult to chose candidates for site directed mutagenesis.

4.3.2.3. A model to consolidate the observed conformational changes

In order to elucidate the relevance of the observed conformational changes, the structures of mature and immature Gag-derived proteins were compared. While the structure of N-CA had been studied in detail [41, 36], not much is known to date about Gag. Recently, the NMR structure of the N-terminally unprocessed protein $MACA_{151}$ [89] was solved and claimed to represent the immature conformation of Gag. This structure was solved at pH 5, and our results indicate that the protein adopts a mature-like conformation under these conditions. Thus, it can be proposed that the structure published for $MACA_{151}$ does not reflect the immature conformation of the protein. This is supported by the comparison of the structure of $MACA_{151}$ at pH 5 [89] with the structure of mature N-CA [41]: A 2 Å shift in helix 6 but no differences in helix 3 are detectable upon comparison of these two structures [89], as indicated schematically in Fig. 4.3. This is in agreement with the antibody pattern observed for a mature-like conformation of $MACA_{151}$ at pH 6 as compared to that observed for CA (Fig. 4.3). Thus, it can be concluded that the immature conformation of Gag-derived proteins competent to assemble into spherical particles is not reflected by the structure of $MACA_{151}$. Based on our results a structure of the immature conformation (Fig. 4.3, left) would be expected to show shifts in helix 3 with respect to the structure of both the mature CA protein (Fig. 4.3, right) as well as to that of the mature-like conformation of $MACA_{151}$ (Fig. 4.3, bottom).

A structure of the proposed immature conformation of Gag might reveal either a new assembly interface important for the formation of immature particles or might elucidate the burial of an interface characteristic only for the mature protein. The HIV-1 CA protein assembles into tubular particles *in vitro* and forms conical cores *in vivo*. Both were shown to be built of helical arrangements of CA hexamers. Low resolution image reconstructions of *in vitro* assembled particles suggested interactions within helix 1 and 2 of N-CA as the basis for the hexameric arrangement [65]. While HIV-1 CA hexamers have not been detected to date, a crystal structure of hexameric murine leukemia virus

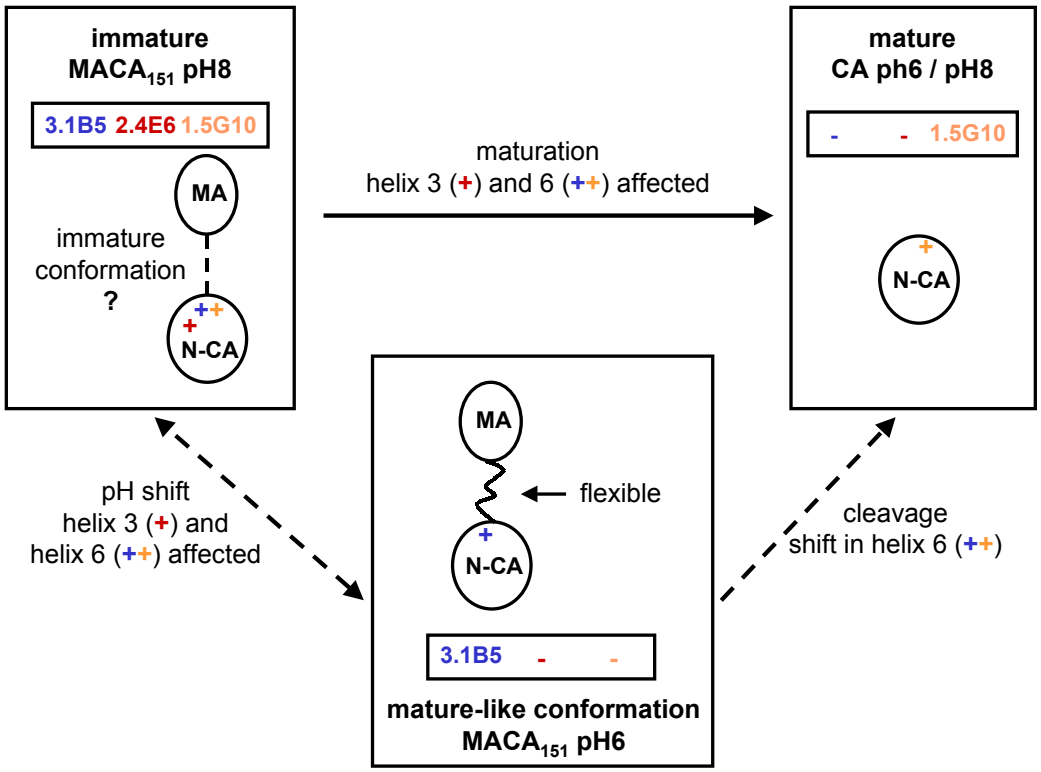

Figure 4.3.: Model to illustrate observed pH dependent changes in helix 3 and 6 of N-CA. Antibody epitopes are coloured, written antibody names and a + indicate binding to the respective protein. The immature conformation (pH 8) can be transformed by pH shift to the mature-like conformation (pH 6). Upon cleavage of the immature form, N-CA obtains the mature conformation, differing from the mature-like form in helix 6 and from the immature form in helix 3 and 6.

(MLV) N-CA was published recently. This structure reveals an interaction between helices 1, 2 and 3 as the basis for the hexameric arrangement [76]. Since the structures of both MLV and HIV N-CA are very similar and general mechanisms driving virus assembly and maturation (such as β hairpin formation, CA protein dimerisation, the relevance of a spacer peptide at the C-terminus of CA) are conserved among several retroviruses, a revision of the model presented for HIV might reveal the same interaction. In this case, cleavage of MA-CA as well as the pH trigger might induce a conformation of helix 3 capable of the formation of hexameric arrangements, concomitantly loosing recognition by antibody 2.4E6. Thus, a structure of the immature conformation might reveal a change in helix 3 that prevents hexamer formation.

We will further pursue the structural analysis of unprocessed Gag-derived proteins by co-crystallisation of $MACA_{151}$ with Fab fragments of 2.4E6 in collaboration with

Wesley I. Sundquist (University of Utah, Salt Lake City, USA). Since antibody 2.4E6 only binds to the immature conformation of the protein, co-crystals should represent the complex of interest. However, it has to be kept in mind that it might not be possible to force a Gag-derived protein into the same conformation that Gag adopts in the virus particle.

The hypothesis of the existence of several relevant CA conformations has been raised before. The model presented for the structure of the mature conical core suggested the presence of more than one conformation of the mature CA protein [65]. A conical core built of helical arrangements of hexamers can only form a closed structure upon the integration of pentameric defects. Based on that it was hypothesised that CA might form a flat, hexameric mature conformation and a slightly more curved, pentameric conformation. This hypothesis of two conformations of the CA proteins was supported by the finding that the shape of CA particles assembled *in vitro* differs at pH 6 and pH 8 (Wesley I.Sundquist, personal communication and [30]). However, with the currently available set of mAbs, mature CA conformations cannot be differentiated. This suggests that these conformations do not show changes in antibody binding to helix 3 or 6 of N-CA but of course does not exclude their existence. More mAbs have been obtained and will be characterised in order to provide a panel of mAbs, whose epitopes cover the whole CA protein. Such a panel of mAbs should allow for the differentiation between two conformations of CA and will lead to further characterisation. This should verify the observed conformational changes in helix 3 and helix 6 of N-CA.

A. Appendix

A.1. The protein expression plasmid pET11c

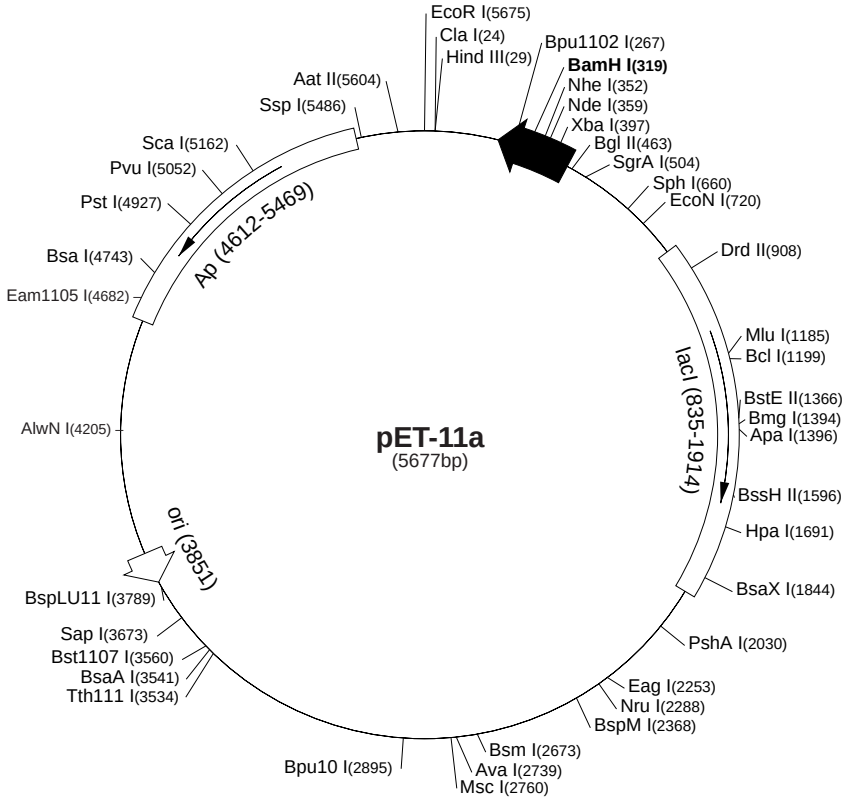

Figure A.1.: Map of the plasmid pET11a (Novagen). The map is the same for pET11c, except that its size is 5675 bp, 2 bp have to be subtracted from each site beyond the *Bam*HI restriction site.

A.2. Oligonucleotides

List of all oligonucleotides used for the amplification of different parts of the HIV-1 *gag* ORF. The presence of a restriction site, a start or stop codon as well as the first encoded amino acid is indicated on the right.

Oligo	sequence	purpose
CA1fwd	5'-GTTCCAGCATATGCCTATAGTGCAGAACCTC-3'	NdeI ATG CA1
CCANCfwd:	5'-GCCGCATATGAGCCCTACCAGCATTCTGG-3'	NdeI ATG CA146
CA177rev	5'-CGGCGGGATCCTTAAGCTTGCTCGGCTCTTAGAG-3'	CA177 stop BamHI
CA194rev	-5'-CGGCGGGATCCTTACGCATTTTGGACCAACAAGG-3'	CA194 stop BamHI
CA207rev	5'-CGGCGGGATCCTTATGGTCCCAATGCTTTTAAAATAG-3'	CA207 stop BamHI
CA220rev	5'-CGGCGGGATCCTTATCCCTGACATGCTGTCATCATTTC-3'	CA220 stop BamHI
CArev	5'-CGGCGGGATCCTTACAAAACTCGAGCTTTATGCC CG-3'	CA231 stop BamHI
SP1rev	5'-CGGCGGGATCCTTACATTATGGTAGCTGGATTTGTTAC-3'	SP1-14 stop BamHI
NCrev:	5'-TGACTCTCTGTCCGATTAATTCCTAGGGCGGC-3'	NC55 stop BamHI
CAΔ207-220rev	5'-CGGCGGGATCCTTACAAAACTCGAGCTTTATGCCCGG GTCCCCCCACTCCCAATGCTTTTAAAATAGTCTTAC-3'	CA199-206 / 221-231 stop BamHI
WMtoAAfwd	5'-GAGGTAAAAAATGCGGCGACAGAAACCTTG-3'	W184A/M185A
WMtoAArev	5'-CAAGGTTTCTGTCGCCGCATTTTTTACCTC-3'	W184A/M185A
73mer	5'-GGCTAGAAGGATCCATATGGGTGCGAGAGCGTCAGTA TTAAGCGGGGGAGAATTAGATCGACCTATAGTGCAG-3'	in vitro assembly
T7fwd	5'-TAATACGACTCACTAT-3'	pET11c sequencing
pET11crev	5'-GCTAGTTATTGCTCAGCG-3'	pET11c sequencing
96gIII	5'-CCCTCATAGTTAGCGTAACG-3'	phage sequencing

A.3. Biophysical parameters of the proteins and peptides used here

Molecular weight, theoretical pI value and molar extinction coefficient at 280 nm as determined by the computer program PROTPARAM (http://www.expasy.org/tools/protparam.html).

protein	M/kD	theoret. pI	ε/$M^{-1}cm^{-1}$
CA	25.6	6.36	22690
CA220	24.5	5.77	33570
CA207	23.2	6.25	33690
CA177	19.9	6.24	27880
CAΔ207-220	24.2	7.49	33570
CA-P207A	25.6	6.36	33690
CA-E212A	25.5	6.71	33690
CA-Q219A	25.5	6.36	33690
CA-W184A/M185A	25.4	6.36	28000
Δ13CA	24.1	6.25	33690
CASP1	27.0	6.11	33690
CCA	9.5	7.75	8370
CCA-W184A/M185A	9.5	7.78	2680
CCANC	17.4	9.46	14420
CANC	33.4	9.09	39740
ΔMACANCSP2	40.4	9.25	47990
NC	6.3	9.93	6050
peptide			
CAI/*CAI*ctrl	1.44	3.4	2400
CAI-C/*CAI*ctrl-C	1.55	n.d.	2400
CAI-C-FITC/*CAI*ctrl-C-FITC	1.94	n.d.	n.d.
CAI-C-biotin/*CAI*ctrl-C-biotin	2.09	n.d.	n.d.
Pep2	1.50	6.0	3600
Pep3	1.34	9.6	5560

A.4. NMR resonance assignments

Backbone resonance assignments were obtained using standard triple resonance experiments. Spectra of free C-CA$_{W184A/M185A}$ were recorded at 1.3 mM protein concentration, spectra of the complex were recorded using 1 mM C-CA$_{W184A/M185A}$ and 4 mM *CAI*. The endpoint of the chemical shift perturbation titration (130 μM protein with to 950 μM *CAI*) was assigned based on the assignments obtained for the free and bound protein at mM concentrations.

Free C-CA$_{W184A/M185A}$

CA residue	H^N [ppm]	N^H [ppm]	Cα [ppm]	Cβ [ppm]
Ser 146	8.82	120.25	57.21	64.77
Pro 147	-	-	63.41	32.29
Thr 148	8.25	114.92	62.18	69.82
Ser 149	8.51	118.88	57.58	64.50
Ile 150	8.61	123.70	61.70	38.60
Leu 151	7.59	119.44	56.74	41.64
Asp 152	7.70	116.14	54.25	41.92
Ile 153	7.47	122.27	59.93	35.43
Arg 154	8.28	124.42	54.23	32.89
Gln 155	7.75	127.18	56.04	27.49
Gly 156	9.88	116.74	45.14	-
Pro 157	-	-	65.71	32.59
Lys 158	8.57	115.10	54.05	32.59
Glu 159	7.03	125.78	53.59	32.18
Pro 160	-	-	63.23	32.68
Phe 161	9.54	128.23	63.23	39.46
Arg 162	9.11	115.78	59.98	30.14
Asp 163	6.97	119.02	57.11	40.15
Tyr 164	7.61	124.59	59.48	38.43
Val 165	8.31	120.27	66.66	31.86
Asp 166	7.54	117.66	57.75	41.30
Arg 167	7.94	119.20	60.16	31.00
Phe 168	9.04	124.63	62.31	39.33
Tyr 169	8.69	116.52	63.66	37.50
Lys 170	8.34	120.99	60.14	32.59

CA residue	H^N [ppm]	N^H [ppm]	Cα [ppm]	Cβ [ppm]
Thr 171	7.67	117.07	66.73	68.35
Leu 172	7.93	122.38	57.85	42.43
Arg 173	8.00	116.39	58.21	30.15
Ala 174	7.57	121.73	53.85	18.88
Glu 175	7.91	118.07	57.24	30.14
Gln 176	8.19	118.78	56.54	28.51
Ala 177	8.02	122.86	53.05	19.46
Ser 178	8.17	114.31	58.28	64.18
Gln 179	8.54	121.31	57.14	29.19
Glu 180	8.48	119.86	58.01	29.89
Val 181	7.84	120.82	63.59	32.43
Lys 182	8.37	123.34	57.58	32.64
Asn 183	8.49	119.45	54.14	38.80
Trp 184	8.21	123.62	53.78	19.14
Met 185	8.32	122.19	53.59	18.59
Thr 186	8.19	114.77	65.44	69.30
Glu 187	8.59	120.14	59.77	29.44
Thr 188	7.77	112.77	65.10	69.20
Leu 189	7.86	123.40	57.58	42.40
Leu 190	7.62	119.54	57.65	42.10
Val 191	7.51	116.90	66.27	32.14
Gln 192	7.94	117.29	59.03	29.16
Asn 193	7.88	115.51	52.87	38.60
Ala 194	7.33	124.85	52.01	19.14
Asn 195	9.10	119.45	51.69	35.62
Pro 196	-	-	66.76	32.35
Asp 197	7.73	115.62	57.37	41.14
Cys 198	8.59	116.62	63.75	28.01
Lys 199	9.51	121.82	61.66	32.59
Thr 200	7.63	113.18	66.70	68.98
Ile 201	6.79	122.48	64.44	38.56
Leu 202	8.52	119.01	57.59	42.27
Lys 203	8.53	119.10	58.86	31.95
Ala 204	7.15	120.16	53.40	18.53
Leu 205	7.41	118.24	56.23	42.65
Gly 206	7.51	128.91	44.52	-
Pro 207	-	-	63.65	32.02
Gly 208	8.79	109.76	45.34	-

CA residue	H^N [ppm]	N^H [ppm]	Cα [ppm]	Cβ [ppm]
Ala 209	7.15	123.44	52.57	20.00
Thr 210	8.63	113.38	60.37	71.56
Leu 211	8.94	122.57	58.35	41.46
Glu 212	8.60	117.71	60.48	29.22
Glu 213	7.79	120.59	59.52	29.70
Met 214	8.56	120.23	60.30	33.27
Met 215	8.78	117.10	59.66	32.99
Thr 216	8.38	117.03	66.60	68.73
Ala 217	7.99	122.93	54.99	19.26
Cys 218	7.38	110.96	60.39	28.56
Gln 219	7.52	122.46	58.43	28.61
Gly 220	8.74	110.11	45.42	-
Val 221	7.37	120.49	63.52	31.87
Gly 222	8.75	115.90	45.21	-
Gly 223	8.02	108.23	44.68	-
Pro 224	-	-	-	28.79
Gly 225	9.50	109.09	43.74	-
His 226	-	-	-	-
Lys 227	-	-	56.12	33.14
Ala 228	8.28	125.48	52.48	19.49
Arg 229	8.31	120.96	56.16	31.11
Val 230	8.24	123.04	62.52	32.91
Leu 231	7.91	107.73	56.80	43.59

C-CA$_{W184A/M185A}$ (1 mM) in complex with *CAI* (4 mM)

CA residue	H^N [ppm]	N^H [ppm]	Cα [ppm]	Cβ [ppm]
Ser 146	-	-	-	-
Pro 147	-	-	-	32.19
Thr 148	8.26	115.73	62.29	69.95
Ser 149	8.48	118.58	56.73	64.70
Ile 150	9.15	126.28	59.45	39.19
Leu 151	7.53	119.45	56.82	41.36
Asp 152	7.60	114.68	54.26	42.12
Ile 153	7.33	121.88	59.30	34.52
Arg 154	8.39	125.33	54.16	33.03
Gln 155	7.87	127.46	55.98	27.60

CA residue	H^N [ppm]	N^H [ppm]	Cα [ppm]	Cβ [ppm]
Gly 156	9.93	116.82	45.18	-
Pro 157	-	-	-	-
Lys 158	-	-	-	-
Glu 159	7.07	125.64	53.62	32.18
Pro 160	-	-	-	32.82
Phe 161	9.56	128.43	63.12	39.78
Arg 162	9.13	115.55	59.77	30.17
Asp 163	7.04	119.28	57.07	40.27
Tyr 164	7.60	123.98	59.56	38.39
Val 165	8.19	119.95	66.78	30.63
Asp 166	7.51	118.97	58.02	41.77
Arg 167	7.74	117.80	60.18	30.81
Phe 168	9.00	125.50	62.14	39.81
Tyr 169	9.41	119.09	61.48	38.56
Lys 170	8.51	117.76	60.14	32.52
Thr 171	7.87	116.04	66.77	68.18
Leu 172	8.41	122.76	56.89	41.69
Arg 173	9.00	119.51	59.55	29.61
Ala 174	6.77	117.95	52.75	19.40
Glu 175	7.09	119.74	57.09	30.53
Gln 176	-	-	-	27.65
Ala 177	7.64	125.98	50.95	22.05
Ser 178	8.55	116.00	57.45	64.84
Gln 179	-	-	-	-
Glu 180	9.09	117.83	60.49	29.07
Val 181	7.59	122.44	66.27	30.87
Lys 182	8.26	119.65	61.62	31.78
Asn 183	9.09	121.04	-	37.94
Trp 184	7.96	123.73	55.27	18.62
Met 185	9.10	120.93	55.09	18.30
Thr 186	8.11	115.83	68.37	69.61
Glu 187	8.30	121.16	59.48	30.84
Thr 188	7.90	110.06	64.57	70.33
Leu 189	8.41	118.45	57.16	42.84
Leu 190	7.22	118.98	59.68	41.66
Val 191	8.29	116.23	65.80	31.75
Gln 192	7.45	116.78	59.12	29.12
Asn 193	8.19	114.75	52.88	37.90

CA residue	H^N [ppm]	N^H [ppm]	Cα [ppm]	Cβ [ppm]
Ala 194	7.34	124.56	52.10	18.81
Asn 195	9.16	119.36	51.68	35.58
Pro 196	-	-	-	32.46
Asp 197	7.76	115.55	57.39	41.27
Cys 198	8.64	116.62	63.79	27.91
Lys 199	9.59	121.90	61.54	32.52
Thr 200	7.65	113.35	66.71	68.99
Ile 201	6.86	122.69	64.61	38.70
Leu 202	8.65	119.00	57.69	42.40
Lys 203	8.60	119.53	59.17	32.01
Ala 204	7.23	120.64	53.83	18.36
Leu 205	7.54	117.89	56.42	42.91
Gly 206	7.34	127.95	44.67	-
Pro 207	-	-	-	-
Gly 208	-	-	-	-
Ala 209	6.98	122.99	52.53	20.42
Thr 210	8.59	113.14	60.09	71.22
Leu 211	9.25	122.38	57.98	41.23
Glu 212	8.80	118.11	60.22	29.45
Glu 213	7.71	119.89	59.61	29.70
Met 214	8.38	120.09	60.21	33.32
Met 215	8.79	117.04	58.77	32.27
Thr 216	8.40	116.28	66.37	68.78
Ala 217	8.08	123.07	54.97	19.41
Cys 218	7.46	110.69	60.36	28.48
Gln 219	7.53	122.87	58.43	28.55
Gly 220	-	-	-	-
Val 221	7.36	120.42	-	31.84
Gly 222	-	-	-	-
Gly 223	8.03	108.12	44.69	-
Pro 224	-	-	-	28.95
Gly 225	9.48	109.77	-	-
His 226	-	-	-	-
Lys 227	-	-	-	33.19
Ala 228	8.26	125.29	52.47	31.03
Arg 229	8.29	120.85	56.14	-
Val 230	8.23	122.81	62.50	32.91
Leu 231	7.91	107.68	56.79	43.63

C-CA$_{W184A/M185A}$ (130 μM) in complex with *CAI* (950 μM)

CA residue	H^N [ppm]	N^H [ppm]
Ser 146	-	-
Pro 147	-	-
Thr 148	8.29	115.84
Ser 149	8.52	118.66
Ile 150	8.66	124.73
Leu 151	7.46	119.86
Asp 152	7.58	114.85
Ile 153	7.37	122.10
Arg 154	8.36	125.08
Gln 155	7.83	127.49
Gly 156	9.92	116.93
Pro 157	-	-
Lys 158	8.57	115.06
Glu 159	7.07	125.73
Pro 160	-	-
Phe 161	9.54	128.47
Arg 162	9.11	115.61
Asp 163	7.04	119.31
Tyr 164	7.60	124.06
Val 165	8.19	120.05
Asp 166	7.51	118.91
Arg 167	7.74	117.93
Phe 168	8.96	125.62
Tyr 169	9.38	119.00
Lys 170	8.54	117.94
Thr 171	7.84	116.21
Leu 172	8.26	122.46
Arg 173	9.02	119.41
Ala 174	6.79	118.01
Glu 175	7.07	119.41
Gln 176	-	-
Ala 177	7.60	125.82
Ser 178	8.55	116.00
Gln 179	9.11	122.36
Glu 180	8.98	118.12
Val 181	7.60	122.73

CA residue	H^N [ppm]	N^H [ppm]
Lys 182	8.23	120.11
Asn 183	9.13	121.21
Trp 184	8.01	123.62
Met 185	8.74	120.89
Thr 186	8.09	115.41
Glu 187	8.36	121.05
Thr 188	7.88	109.79
Leu 189	8.40	119.17
Leu 190	7.28	119.43
Val 191	8.28	116.34
Gln 192	7.51	117.13
Asn 193	8.21	114.97
Ala 194	7.34	124.69
Asn 195	9.14	119.44
Pro 196	-	-
Asp 197	7.75	115.62
Cys 198	8.63	116.68
Lys 199	9.56	121.93
Thr 200	7.64	113.43
Ile 201	6.85	122.76
Leu 202	8.62	119.04
Lys 203	8.57	119.53
Ala 204	7.21	120.64
Leu 205	7.51	117.94
Gly 206	7.35	128.08
Pro 207	-	-
Gly 208	8.77	109.58
Ala 209	6.98	123.02
Thr 210	8.56	113.17
Leu 211	9.23	122.42
Glu 212	8.77	118.14
Glu 213	7.70	119.96
Met 214	8.37	120.13
Met 215	8.78	117.11
Thr 216	8.39	116.32
Ala 217	8.07	123.15
Cys 218	7.46	110.82
Gln 219	7.52	122.80

CA residue	H^N [ppm]	N^H [ppm]
Gly 220	-	-
Val 221	7.38	120.43
Gly 222	8.74	115.94
Gly 223	8.02	108.23
Pro 224	-	-
Gly 225	9.43	109.73
His 226	-	-
Lys 227	-	-
Ala 228	8.28	125.47
Arg 229	8.30	120.92
Val 230	8.23	122.96
Leu 231	7.90	107.73

A.5. Abbreviations

α	anti
A_{280}	absoprtion at 280 nm
AA	acrylamide
abs	absolute
AIDS	acquired immunodeficiency syndrome
AMP	ampicillin
APS	ammonium persulfate
BIS	N,N'-methylene bisacrylamide
bp	base pair
BSA	bovine serum albumin
CA	capsid protein
C-CA	C-terminal domain of the capsid protein
CIAP	Calf intestinal alkaline phosphatase
C-terminus	carboxy terminus of a protein
D	Dalton (g/mol)
δ	chemical shift
DEAE	diethylaminoethyl
DMSO	dimethyl sulfoxide
DTT	1,4-dithiothreitol
ε	extinction coefficient
ELISA	enzyme-linked immunosorbent assay
E. coli	*Escherichia coli*

EM	electron microscopy
EtOH	ethanol
FCS	fetal calf serum
FID	free induction decay
Fig.	Figure
FT	Fourier transform
Gag	group specific antigen, viral structural protein
H	enthalpy
HAART	highly active anti retroviral therapy
HE	heparin
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
HIV	human immunodeficiency virus
HQ	quaternary polyethylene imine
HRP	horse radish peroxidase
HS	sulfopropyl
HSQC	heteronuclear single quantum coherence
HTLV	Human T-cell leukemia virus
IPTG	isopropyl-β-D-thio-galactoside
ITC	isothermal titration calorimetry
K_D	dissociation constant
LB	Luria broth
LTR	long terminal repeat
MA	matrix
MES	2-morpholinoethanesulfonic acid
MLV	Murine leukemia virus
M-PMV	Mason-Pfizer monkey virus
mRNA	messenger RNA
n	binding stoichiometry
NaAc	sodium acetate
NC	nucleocapsid
N-CA	N-terminal domain of the capsid protein
NMR	nuclear magnetic resonance spectroscopy
NOESY	nuclear Overhauser effect spectroscopy
nt	nucleotide
N-terminus	amino terminus of a protein
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline

PBS	primer binding site
PCR	polymerase chain reaction
PEG	polyethylene glycol
pI	isoelectric point
PIC	preintegration complex
ppm	parts per million
PR	protease
Q	heat
rpm	rotations per minute
RSV	Rous sarcoma virus
RT	reverse transcriptase
S	entropy
SDS	sodium dodecylsulphate
SP1	spacer peptide 1
τ	correlation time
t	temperature
T_1,T_2	relaxation times
TCA	Trichloroacetic acid
TCEP	Tris(2-carboxyethyl)phosphine
TE	Tris/EDTA
TMB	3,3',5,5' tetramethylbenzidine
Tris	Tris/HCl, tris(hydroxmethylaminomethane hydrochloride)
TOCSY	total correlation spectroscopy
wt	wild type
XGal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Bibliography

[1] **Abdurahman S., Höglund S., Goodbar-Larsson L., Vahlne A.** (2004). Selected amino acid substitutions in the C-terminal region of human immunodeficiency virus type 1 capsid protein affect virus assembly and release. J. Gen. Virol. **85**:2903–13.

[2] **Accola M.A., Höglund S., Göttlinger H.G.** (1998). A putative alpha-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. J. Virol. **72**:2072–2078.

[3] **Accola M.A., Strack B., Göttlinger H.G.** (2000). Efficient particle production by minimal gag constructs which retain the carboxy-terminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain. J. Virol. **74**:5395–5402.

[4] **Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A., Martin M.A.** (1986). Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. **59(2)**:284–291.

[5] **del Álamo M., Neira J.L., Mateu M.G.** (2003). Thermodynamic dissection of a low affinity protein-protein interface involved in human immunodeficiency virus assembly. J. Biol. Chem. **278(30)**:27923–27929.

[6] **Bax A., Ikura M., Kay L.E., D. A. Torchia a.R.T.** (1990). Comparison of different modes of 2-dimensional reverse-correlation NMR for the study of proteins. J. Magn. Reson. **86**:304–318.

[7] **Berthet-Colominas C., Monaco S., Novelli A., Siba G., Mallet F., Cusack S.** (1999). Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. EMBO J. **18**:1124–1136.

[8] **Bodenhausen G., Ruben D.J.** (1980). Natural abundance nitrogen-15 nmr by enhanced heteronuclear spectroscopy. Chem. Phys. Lett. **69**:185–189.

[9] **Borsetti A., Ohagen A., Göttlinger H.G.** (1998). The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. J. Virol. **72**:9313–9317.

[10] **Bowzard J.B., Wills J.W., Craven R.** (2001). Second-site suppressors of Rous sarcoma virus CA mutations: Evidence for interdomain interactions. J. Virol. **75**:6850–6856.

[11] **Briggs J.A.G., Simon M.N., Gross I., Krausslich H.G., Fuller S.D., Vogt V.M., Johnson M.C.** (2004). The stoichiometry of Gag protein in HIV-1. Nat. Struct. Mol. Biol. **11(7)**:672–675.

[12] **Briggs J.A.G., Watson B.E., Gowen B.E., Fuller S.D.** (2004). Cryoelectron microscopy of mouse mammary tumor virus. J. Virol. **78(5)**:2606–2608.

[13] **Briggs J.A.G., Wilk T., Welker R., Kräusslich H.G., Fuller S.D.** (2003). Structural organization of authentic, mature HIV-1 virions and cores. EMBO J. **22(7)**:1707–1715.

[14] **Campbell S., Fisher R.J., Towler E.M., Fox S., Issaq H.J., Wolfe T., Phillips L.R., Rein A.** (2001). Modulation of HIV-like particle assembly in vitro by inositol phosphates. Proc. Natl. Acad. Sci. USA **98**:10875–10879.

[15] **Campbell S., Rein A.** (1999). In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J. Virol. **73**:2270–2279.

[16] **Campbell S., Vogt V.M.** (1995). Self-assembly in vitro of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J. Virol. **69**:6487–6497.

[17] **Campbell S., Vogt V.M.** (1997). In vitro assembly of virus-like particles with rous sarcoma virus Gag deletion mutants: Identification of the p10 domain as a morphological determinant in the formation of spherical particles. J. Virol. **71**:4425–4435.

[18] **Campos-Olivas R., Newman J.L., Summers M.F.** (2000). Solution structure and dynamics of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. J. Mol. Biol. **296(2)**:633–649.

[19] **Carrière C., Gay B., Chazal N., Morin N., Boulanger P.** (1995). Sequence requirements for encapsidation of deletion mutants and chimeras of human immunodeficiency virus type 1 Gag precursor into retrovirus-like particles. J. Virol. **69(4)**:2366–2377.

[20] **Cavanagh J., Fairbrother W.J., Palmer III A.G., Skelton N.J.** (1996). Protein NMR Spectroscopy: Principles and Practice. Academic Press, San Diego.

[21] **Chazal N., Carrière C., Gay B., Boulanger P.** (1994). Phenotypic characterization of insertion mutants of the human immunodeficiency virus type 1 Gag precursor expressed in recombinant baculovirus-infected cells. J. Virol. **68(1)**:111–122.

[22] **Chazal N., Gerlier D.** (2003). Virus entry, assembly, budding, and membrane rafts. Microbiol. Mol. Biol. Rev. **67**:226–237.

[23] **Cheslock S.R., Poon D.T.K., Fu W., Rhodes T.D., Henderson L.E., Nagashima K., McGrath C.F., Hu W.S.** (2003). Charged assembly helix motif in murine leukemia virus capsid: an important region for virus assembly and particle size determination. J. Virol. **77**:7058–7066.

[24] **Coffin J.M., Hughes S.H., Varmus H.E.** (1999). Retroviruses. Cold Spring Harbor Laboratory Press, New York.

[25] **Cullen B.R.** (1991). Human immunodeficiency virus as a prototypic complex retrovirus. J. Virol. **65**:1053–1056.

[26] **DeGuzman R.N., Wu Z.R., Stalling C.C., Pappalardo L., Borer P.N., Summers M.F.** (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. Science **279**:384–388.

[27] **Delaglio F., Grzesiek S., Vuister G.W., Zhu G., Pfeifer J., Bax A.** (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR **6**:277–293.

[28] **Douglas C.C., Thomas D., Lanman J., Prevelige Jr P.** (2004). Investigation of N-terminal domain charged residues on the assembly and stability of HIV-1 CA. Biochemistry **43**:10435–10441.

[29] **Driscoll P.C., Marion G.M.C.D., Wingfieldand P.T., Gronenborn A.M.** (1990). Complete resonance assignment for the polypeptide backbone of interleukin-1-beta using 3-dimensional heteronuclear NMR-spectroscopy. Biochemistry **29**:3542–3556.

[30] **Ehrlich L.S., Liu T., Scarlata S., Chu B., Carter C.A.** (2001). HIV-1 capsid protein forms spherical (immature-like) and tubular (mature-like) particles in vitro: Structure switching by pH-induced conformational changes. Biophys. J. **81**:586–594.

[31] **Fäcke M., Janetzko A., Shoeman R.L., Kräusslich H.G.** (1993). A large deletion in the matrix domain of the human immunodeficiency virus gag gene redirects virus particle assembly from the plasma membrane to the endoplasmatic reticulum. J. Virol. **67**:4972–4980.

[32] **Forshey B.M., von Schwedler U., Sundquist W.I., Aiken C.** (2002). Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. J. Virol. **76**:5667–5677.

[33] **Franke E.K., Yuan H.E., Luban J.** (1994). Specific incorporation of cyclophilin A into HIV-1 virions. Nature **372**:359–362.

[34] **Freed E.O.** (1998). HIV-1 Gag proteins: Diverse functions in the virus life cycle. Virology **251**:1–15.

[35] **Fuller S.D., Wilk T., Gowen B.E., Kräusslich H.G., Vogt V.M.** (1997). Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. Curr. Biol. **7**:729–738.

[36] **Gamble T.R., Vajdos F.F., Yoo S., Worthylake D.K., Houseweart M., Sundquist W.I., Hill C.P.** (1996). Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell **87**:1285–1294.

[37] **Gamble T.R., Yoo S., Vajdos F.F., von Schwedler U.K., Worthylake D.K., Wang H., McCutcheon J.P., Sundquist W.I., Hill C.P.** (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. Science **278**:849–852.

[38] **Ganser-Pornillos B.K., von Schwedler U.K., Stray K.M., Aiken C., Sundquist W.I.** (2004). Assembly properties of the human immunodeficiency virus type 1 CA protein. J. Virol. **78(5)**:2545–2552.

[39] **Garzón M.T., Lidon-Moya M.C., Barrera F.N., Prieto A., Gomez J., Mateu M.G., Neira J.L.** (2004). The dimerization domain of the HIV-1 capsid protein binds a capsid protein-derived peptide: a biophysical characterization. Protein Sci. **13(6)**:1512–1523.

[40] **Gheysen D., Jacobs E., de Foresta F., Thiriart C., Francotte M., Thinnes D., De Wilde M.** (1989). Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. Cell **59**:103–112.

[41] **Gitti R.K., Lee B.M., Walker J., Summers M.F., Yoo S., Sundquist W.I.** (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. Science **273**:231–235.

[42] **Grättinger M.** (2000). Charakterisierung der *in vitro* Assemblierung und Nukleinsäureinteraktion von HIV Strukturproteinen. Dissertation, Universität Hamburg.

[43] **Gross I., Hohenberg H., Huckhagel C., Kräusslich H.** (1998). N-terminal

extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. J. Virol. **72**:4798–4810.

[44] **Gross I., Hohenberg H., Kräusslich H.** (1997). In vitro assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. Eur. J. Biochem. **249**:592–600.

[45] **Gross I., Hohenberg H., Wilk T., Wiegers K., Grättinger M., Müller B., Fuller S., Kräusslich H.G.** (2000). A conformational switch controlling HIV-1 morphogenesis. EMBO J. **19**:103–113.

[46] **Grzesiek S., Bax A.** (1992). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. J. Magn. Reson. **99**:201–207.

[47] **Grzesiek S., Bax A.** (1992). Measurement of amide proton-exchange rates and noes with water in c-13/n-15-enriched calcineurin-b. J. Biomol. NMR **3**:627–638.

[48] **He J., Chen Y., Farzan M., Choe H., Ohagen A., Gartner S., Busciglio J., Yang X., Hofmann W., Newman W., Mackay C.R., Sodroski J., Gabuzda D.** (1997). CCR3 and CCR5 are co-receptors of HIV-1 infection of microglia. Nature **385**:645–649.

[49] **Hill C.P., Worthylake D., Bancroft D.P., Christensen A.M., Sundquist W.I.** (1996). Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: Implications for membrane association and assembly. Proc. Natl. Acad. Sci. USA **93**:3099–3104.

[50] **Höglund S., Su J., Reneby S.S., Vegvari A., Hjerten S., Sintorn I.M., Foster H., Wu Y.P., Nystrom I., Vahlne A.** (2002). Tripeptide interference with human immunodeficiency virus type 1 morphogenesis. Antimicrob. Agents Chemother. **46(11)**:3597–3605.

[51] **Hyde-DeRuyscher R., Paige L.A., Christensen D.J., Hyde-DeRuyscher N., Lim A., Fredericks Z.L., Kranz J., Gallant P., Zhang J., Rocklage S.M., Fowlkes D.M., Wendler P.A., Hamilton P.T.** (2000). Detection of small-molecule enzyme inhibitors with peptides isolated from phage-displayed combinatorial peptide libraries. Chem. Biol. **7(1)**:17–25.

[52] **Johnson B.A., Blevins R.A.** (1994). NMRView: A computer program for the visualization and analysis of NMR data. J. Biomol. NMR **4**:603–614.

[53] **Kattenbeck B., von Poblotzki A., Rohrhofer A., Wolf H., Modrow S.** (1997). Inhibition of human immunodeficiency virus type 1 particle formation by alterations of defined amino acids within the C terminus of the capsid protein. J. Gen. Virol. **78**:2489–2496.

[54] **Kay L.E.** (1995). Pulsed-field gradient multidimensional NMR methods for the study of protein structure and dynamics. Prog. Biophys. Mol. Biol. **63**:277–299.

[55] **Kay L.E., Ikura M., Tschudin R., Bax A.** (1989). 3-dimensional triple-resonance NMR-spectroscopy of isotopically enriched proteins. J. Magn. Reson. **89**:496–51.

[56] **Kay L.E., Marion D., Bax A.** (1989). Practical aspects of 3d heteronuclear NMR of proteins. J. Magn. Reson. **84**:72–84.

[57] **Khorasanizadeh S., Campos-Olivas R., Summers M.F.** (1999). Solution structure of the capsid protein from the human T-cell leukemia virus type-I. J. Mol. Biol. **291(2)**:491–505.

[58] **Koradi R., Billeter M., Wüthrich K.** (1996). MOLMOL: A program for display and analysis of macromolecular structure. J. Mol. Graphics **14**:51–55.

[59] **Krausslich H.G., Facke M., Heuser A.M., Konvalinka J., Zentgraf H.** (1995). The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. J. Virol. **69(6)**:3407–3419.

[60] **Lanman J., Lam T.T., Barnes S., Sakalian M., Emmett M.R., Marshall A.G., Prevelige Jr P.E.** (2003). Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. J. Mol. Biol. **325**:759–772.

[61] **Lanman J., Lam T.T., Emmett M.R., Marshall A.G., Sakalian M., Prevelige Jr P.E.** (2004). Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. Nat. Struct. Mol. Biol. **11(7)**:676–677.

[62] **Lanman J., Sexton J., Sakalian M., Prevelige P.** (2002). Kinetic analysis of the role of intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly in vitro. J. Virol. **76**:6900–6908.

[63] **Leavitt S., Freire E.** (2001). Direct measurement of protein binding energetics by isothermal titration calorimetry. Curr. Opin. Struct. Biol. **11**:560–566.

[64] **Li F., Goila-Gaur R., Salzwedel K., Kilgore N.R., Reddick M., Matallana C., Castillo A., Zoumplis D., Martin D.E., Orenstein J.M., Allaway G.P., Freed E.O., Wild C.T.** (2003). PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. Proc. Natl. Acad. Sci. U S A **100(23)**:13555–13560.

[65] **Li S., Hill C., Sundquist W., Finch J.** (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. Nature **407**:409–413.

[66] **Lipari G., Szabo A.** (1989). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecule. 1. theory and range of validity. J. Am. Chem. Soc. **104**:4546–4559.

[67] **Ma Y.M., Vogt V.M.** (2004). Nucleic acid binding-induced Gag dimerization in the assembly of Rous sarcoma virus particles in vitro. J. Virol. **78(1)**:52–60.

[68] **Mammano F., Ohagen A., Höglund S., Göttlinger H.G.** (1994). Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. J. Virol. **68**:4927–4936.

[69] **Marion D., Ikura M., Tschudin R., Bax A.** (1989). Rapid recording of 2D NMR-spectra without phase cycling - application to the study of hydrogen-exchange in proteins. J. Magn. Reson. **85**:393–399.

[70] **Massiah M.A., Starich M.R., Paschall C., Summers M.F., Christensen A.M., Sundquist W.I.** (1994). Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein. J. Mol. Biol. **244**:198–223.

[71] **Mateu M.G.** (2002). Conformational stability of dimeric and monomeric forms of the C-terminal domain of human immunodeficiency virus-1 capsid protein. J. Mol. Biol. **318(2)**:519–531.

[72] **Matthews S., Barlow P., Boyd J., Barton G., Russell R., Mills H., Cunningham M., Meyers N., Burns N., Clark N.** (1994). Structural similarity between the p17 matrix protein of HIV-1 and interferon-γ. Nature **370**:666–668.

[73] **Matthews S., Barlow P., Clark N., Kingsman S., Kingsman A., Campbell I.** (1995). Refined solution structure of p17, the HIV matrix protein. Biochem. Soc. Trans. **23**:725–729.

[74] **Melamed D., Mark-Danieli M., Kenan-Eichler M., Kraus O., Castiel A., Laham N., Pupko T., Glaser F., Ben-Tal N., Bacharach E.** (2004). The conserved carboxy terminus of the capsid domain of the human immunodeficiency virus type 1 Gag protein is important for virion assembly and release. J. Virol. **78**:9675–9688.

[75] **Morita E., Sundquist W.I.** (2004). Retrovirus budding. Annu. Rev. Cell Dev. Biol. **20**:395–425.

[76] **Mortuza G.B., Haire L.F., Stevens A., Smerdon A.J., Stoye J.P., Taylor I.A.** (2004). High-resolution structure of a retroviral capsid hexameric amino-terminal domain. Nature **431**:481–485.

[77] **Muriaux D., Mirro J., Harvin D., Rein A.** (2001). RNA is a structural element in retrovirus particles. Proc. Natl. Acad. Sci. USA **98**:5246–5251.

[78] **Newman J.L., Butcher E.W., Patel D.T., Mikhaylenko Y., Summers M.F.** (2004). Flexibility in the P2 domain of the HIV-1 Gag polyprotein. Protein Sci. **13(8)**:2101–2107.

[79] **Niedrig M., Gelderblom H.R., Pauli G., März J., Bickhard H., Wolf H., Modrow S.** (1994). Inhibition of infectious human immunodeficiency virus type 1 particle formation by Gag protein-derived peptides. J. Gen. Virol. **75**:1469–74.

[80] **Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.S., Gallo R.C., Wong-Staal F.** (1987). Complete nucleotide sequences of functional clones of the AIDS virus. AIDS Res. Hum. Retroviruses **3**:57–69.

[81] **Reicin A.S., Paik S., Berkowitz R.D., Luban J., Lowy I., Goff S.P.** (1995). Linker insertion mutations in the human immunodeficiency virus type 1 gag gene: effects on virion particle assembly, release, and infectivity. J. Virol. **69(2)**:642–650.

[82] **Scarlata S., Ehrlich L.S., Carter C.A.** (1998). Membrane-induced alterations in HIV-1 Gag and matrix protein-protein interactions. J. Mol. Biol. **277**:161–169.

[83] **Sheehy A.M., Gaddis N.C., Choj J.D., Malim M.H.** (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature **418**:646–650.

[84] **Silver M.S., Joseph R.I., Hoult D.I.** (1984). Highly selective pi/2 and pi-pulse generation. J. Magn. Reson. **59**:347–351.

[85] **Spearman P., Horton R., Ratner L., Kuli-Zade I.** (1997). Membrane binding of human immunodeficiency virus type 1 matrix protein in vivo supports a conformational myristyl switch mechanism. J. Virol. **71**:6582–6592.

[86] **Spera S., Bax A.** (1991). Empirical correlation between protein backbone conformation and C-alpha and C-beta C-13 nuclear-magnetic-resonance chemical-shifts. J. Am. Chem. Soc. **113**:5490–5492.

[87] **Tang C., Loeliger E., Kinde I., Kyere S., Mayo K., Barklis E., Sun Y., Huang M., Summers M.F.** (2003). Antiviral inhibition of the HIV-1 capsid protein. J. Mol. Biol. **327(5)**:1013–1020.

[88] **Tang C., Loeliger E., Luncsford P., Kinde I., Beckett D., Summers M.F.** (2004). Entropic switch regulates myristate exposure in the HIV-1 matrix protein. pnas **101**:517–522.

[89] **Tang C., Ndassa Y., Summers M.F.** (2002). Structure of the N-terminal 283-residue fragment of the immature HIV-1 gag polyprotein. Nature Struct. Biol. **9**:537–543.

[90] **Towers G.J., Hatziioannou T., Cowan S., Goff S.P., Luban J., Bieniasz P.D.** (2003). Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. Nat. Med. **9**:1138–1143.

[91] **von Schwedler U.K., Stemmler T.L., Klishko V.Y., Li S., Albertine K.H., Davis D.R., Sundquist W.I.** (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. EMBO J. **17(6)**:1555–1568.

[92] **von Schwedler U.K., Stray K.M., Garrus J.E., Sundquist W.I.** (2003). Functional surfaces of the human immunodeficiency virus type 1 capsid protein. J. Virol. **77(9)**:5439–5450.

[93] **Wiegers K., Rutter G., Kottler H., Tessmer U., Hohenberg H., Kräusslich H.G.** (1998). Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. J. Virol. **72**:2846–2854.

[94] **Wiegers K., Rutter G., Schubert U., Grättinger M., Kräusslich H.G.** (1999). Cyclophilin A incorporation is not required for human immunodeficiency virus type 1 particle maturation and does not destabilize the mature capsid. Virology **257**:261–274.

[95] **Wilk T., Fuller S.D.** (1999). Towards the structure of the human immunodeficiency virus: Divide and conquer. Curr. Opin. Struct. Biol. **9**:231–243.

[96] **Wilk T., Gross I., Gowen B.E., Rutten T., de Haas F., Welker R., Kräusslich H.G., Boulanger P., Fuller S.D.** (2001). The organization of immature HIV-1. J. Virol. **75**:759–771.

[97] **Wiseman T., Williston S., Brandts J.F., Lin L.N.** (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal. Biochem. **179**:131–135.

[98] **Wittekind M., Mueller L.** (1989). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide- proton and nitrogen resonances with the alpha-carbon and beta-carbon resonances in proteins. J. Magn. Reson., Ser. B **101**:201–205.

[99] **Wobus C.E., Hügle-Dörr B., Girod A., Petersen G., Hallek M., Kleinschmidt J.A.** (2000). Monoclonal antibodies against the adeno-associated-virus type 2 (AAV-2) capsid: Epitope mapping and identification of capsid domains involved in AAV-2–cell interaction and neutralization of AAV-2 infection. J. Virol. **74(19)**:9281–9293.

[100] **Worthylake D.K., Wang H., Yoo S., Sundquist W.I., Hill C.P.** (1999). Structures of the HIV-1 capsid protein dimerization domain at 2.6 åresolution. Acta Crystallogr. D. Biol. Crystallogr. **55**:85–92.

[101] **Yoo S., Myszka D.G., Yeh C., McMurray M., Hill C.P., Sundquist W.I.** (1997). Molecular recognition in the HIV-1 Capsid/Cyclophilin A complex. J. Mol. Biol. **269(5)**:780–795.

[102] **Zhou J., Chen C.H., Aiken C.** (2004). The sequence of the CA-SP1 junction accounts for the differential sensitivity of HIV-1 and SIV to the small molecule maturation inhibitor 3-*O*-(3',3'-dimethylsuccinyl)-betulinic acid. Retrovirology **15(1)**.

[103] **Zhou J., Yuan X., Dismuke D., Forshey B.M., Lundquist C., Lee K.H., Aiken C., Chen C.H.** (2004). Small-molecule inhibition of human immunodeficiency virus type 1 replication by specific targeting of the final step of virion maturation. J. Virol. **78(2)**:922–929.

[104] **Zhou W., Resh M.D.** (1996). Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. J. Virol. **70**:8540–8548.

[105] **Zlotnick A., Ceres P., Singh S., Johnson J.M.** (2002). A small molecule inhibits and misdirects assembly of hepatitis B virus capsids. J. Virol. **76(10)**:4848–4854.

Publications

Sticht J., Humbert M., Findlow S., Bodem J., Müller B., Dietrich U., Werner J., Kräusslich H.-G. (2004). A peptide inhibitor of HIV-1 assembly (submitted for publication)

Ternois F., Sticht J., Werner J., Müller B., Kräusslich H.-G., Rey F. (2005). Crystal structure of the C-terminal domain of the HIV-1 capsid protein in complex with an assembly inhibitor (in preparation)

Sticht, J., Wiegers K., Blüthner M., Kräusslich H.-G. (2003). Inhibition of HIV-1 assembly by monoclonal antibodies. Poster presentation at the Jahrestagung der Gesellschaft für Virologie in March 2003, Berlin, Germany.

Sticht, J., Gross I., Grättinger M., Müller B., Wiegers K., Hohenberg H., Rutter G., Wilk T., Fuller S., Kräusslich H.-G. (2003). In vitro assembly of human immunodeficiency virus. Workshop talk at the 152nd Meeting of the Society for General Microbiology in April 2002 at Edinburgh, UK

Sticht, J., Müller B., Bodem J., Wiegers K., Blüthner M., Wang H., Anderson D., Sundquist W. I., Kräusslich H.-G. (2003). PH dependent conformational changes in HIV-1 Gag derived proteins linked to viral maturation. Poster presentation at the Fourth international retroviral NC symposium in September 2003, Strasbourg.

Sticht J., Werner J., Humbert M., Dietrich U., Müller B., Bodem J.,Kräusslich H.-G. (2004) A peptide inhibitor of HIV-1 assembly. Poster presentation at the FASEB summer research conference in July 2004, Vermont, USA.

Sticht J., Humbert M., Findlow S., Bodem J., Müller B., Dietrich U., Werner J., Kräusslich H.-G. (2004). A peptide inhibitor of HIV-1 assembly. Poster presentation at the Retrovirus assembly meeting in October 2004, Prague, Czech Republic.

Acknowledgements

First and foremost thanks to my supervisor Hans-Georg Kräusslich for constant support, long discussions and for providing the general direction of the project combined with the trust in me treading my own paths. Thanks to Stephan Urban for being available as referee and for the interest and help in my project. I would also like to thank Barbara Müller for help and discussions and Jochen Bodem for the most critical but still helpful attitude towards my project and for his computational support in any Linux related question. And of course thanks to all the others in the lab, especially to Eva and Stefan, with whom I really enjoyed sharing the time in the lab.

Many thanks to Jörn Werner for opening the door to NMR spectroscopy to me and firing my enthusiasm for it. Thanks also for countless discussions and the real interest in my project, which just made it such great fun to work together. This collaboration allowed to solve many problems related to the project and it turned my trips to England to a very enjoyable time, which is also thanks to Anna.

Thanks to Gareth Griffiths for the possibility to use the electron microscopes at EMBL and to Anja Habermann and Ken Goldie for constant EM support in technical questions. I thank Michael Humbert for introducing me to the phage display technology and for all subsequent discussions. Thanks to Martin Blüthner for more phage support and the peptide scan membranes. I also thank François Ternois, Felix Rey, Wes Sundquist, Steve Fuller and John Briggs for great collaboration and fruitful discussions.

Finally, many thanks to Matthias for support at all times, including phases of frustration and overwork, and for the effort he put in helping with the computational refinement of this thesis. And thanks to all those who never stopped trying to drag me away from work for climbing, mountaineering or just for running up and down the hills surrounding Heidelberg.