

Biophysical and biochemical studies of HIV-1 capsid interactions

Juliana Piacentini

Hollidaysburg, Pennsylvania

B.A. Biochemistry & Molecular Biology, Franklin & Marshall College, 2017

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Molecular Physiology & Biological Physics

University of Virginia

March 2024

Abstract

Human immunodeficiency virus type 1 (HIV-1) is a virus of global public health concern that weakens a host's immune system by attacking their immune cells, increasing susceptibility to other diseases and often leading to the development of acquired immune deficiency syndrome (AIDS). The HIV-1 capsid is a dynamic molecular machine essential to productive infection as it functions in numerous events in the viral life cycle. The mature capsid consists of ~1500 copies of the monomeric capsid (CA) protein arranged within an extended lattice that contains about 250 hexamers and exactly 12 pentamers to form a fullerene cone shell. Structural stability of the capsid lattice is crucial for proper completion of early infection events. Many binding partners of the capsid modulate lattice stability with either promotive or restrictive effects. In this work, we used a clean biochemical system to elucidate important details about the HIV-1 capsid and its binding partners, including host cell factor polyglutamine binding protein 1 (PQBP1). This intrinsically disordered protein bridges viral capsid recognition with the recruitment of cGAS (cyclic GMP-AMP synthase), a pattern recognition receptor that stimulates downstream induction of interferon upon detection of cytosolic double-stranded DNA. We provide structural evidence of PQBP1 binding to the capsid with the primary interaction occurring between the electronegative N-terminus of PQBP1 and the electropositive arginine-18 ring found within CA hexamers, though additional regions of PQBP1 may contribute to capsid binding. Furthermore, we demonstrate that PQBP1 binding promotes capsid lattice destabilization, likely as a result of PQBP1 self-association driven by C-terminal dimerization. Through structural determination of cross-linked CA pentamers bound to HIV-1 antagonist GS-CA1, we also demonstrate structural plasticity of the CA pentamer as evidenced by its hexamer-like conformation in complex with this ligand. All together, this dissertation reveals new and important details about the HIV-1 capsid and two of its many binding partners, allowing for a better overall understanding of capsid's structural dynamics throughout infection.

Acknowledgements

It's me. Hi. Even as I write this, I am still processing the fact that I am on the precipice of completing my doctoral degree, a milestone beyond my wildest dreams. One thing I know for certain, however, is that I could never have done it, and indeed didn't do it, without the endless support of so many incredible people.

First, to my advisors, Dr. Owen Pornillos and Dr. Barbie Ganser-Pornillos – thank you for taking a chance on me and for always seeing the potential in me. I had no concept of what graduate school would look like when I came to UVA, but I couldn't have imagined a more perfect fit for me than the Pornillos lab. You two are incredible scientists and people, and I will always be proud to call myself your student. I also want to thank all the members of the Pornillos lab, past and present. I wish you all success in your future endeavors and will fondly think back on the memories we shared. I especially want to thank Kaneil Zadrozny, who always looked after us in more ways than just in our science. You were the rock of our lab, and I'll always remember us bonding over our moon tattoos in the first few weeks of my rotation.

I am grateful to have developed a fantastic network of mentors outside of my thesis lab, as well. Thank you to my thesis committee members – Dr. Anna Cliffe, Dr. Anne Kenworthy, Dr. Robert Nakamoto, and Dr. Stefanie Redemann – for your time and invaluable contributions to my training. Thank you to Dr. Lukas Tamm for hosting me in his lab for the final months of my degree and to Dr. Janet Cross for constantly offering me support during this period. I also want to thank the faculty and administration within Biophysics, the Molecular Physiology & Biological Physics department, and the Molecular Electron Microscopy Core for their impact on my graduate school experience. Biophysics has grown so much in my time at UVA, and I can't wait to see what new accomplishments members of the community will achieve.

I wouldn't have ever made it to graduate school in the first place without the guidance of my undergrad PIs, Dr. Christine Phillips-Piro and Dr. Scott Brewer. It was in your labs that I began my journey with research and discovered a passion for science I always knew I had but never fully understood. I will never forget the conversation I had with you, Dr. Brewer, the day before I was supposed to take the MCAT (and hadn't studied, so I was just going to wing it...) in which you so patiently and kindly described your grad school experience and planted the seed in my mind that I could do something similar after college. And Dr. Piro, you held my hand through the tumultuous year that was my 2017. You fostered my personal and professional growth, supporting me through heartbreak, holding me accountable when I needed it, and guiding me through the graduate school application process. I will never have the words to adequately express how grateful I am for all you have done for me in that year and all the years since, and I feel so honored to think of you as a colleague & friend.

In coming to UVA for graduate school, I have been incredibly fortunate to meet some of the best people I have ever known and make friends I know will last a lifetime. Thank you to my BIMS besties for the chats, laughs, tears, movie nights, game nights, silent discos, and granting me the title of Cat Auntie to all my cute little furry nieces & nephews. I'm not sure I will miss graduate school, but I will certainly miss the grad school experiences we shared. Thank you to my Biophysics cohort, members of Women In Medical Sciences (WIMS), and the UVA grad student friends I made outside of science for the bonding moments, both in work and in leisure. Thank you to communities like my church, Wesley Memorial UMC, and my band, the Albemarle County Concert Band, that truly made Charlottesville my home for the past six years. I have treasured my time here, and a piece of my heart will always live in Cville.

As proud as I feel of the academic work I have done and the life I have created in Charlottesville, the single greatest gift that graduate school has given me is my life partner. Mitchell, you are everything to me. You have supported me through some of the hardest times of my life and given me belief in myself when I had none. Even in times when my world felt like it was falling to pieces, I always knew that I did one thing right in finding you. What a blessing that we got to go on this journey together, growing alongside and with one another throughout these years in graduate school. I love you, I am endlessly thankful for you, and I am so excited to see what future we will continue building together as Dr. & Dr. (and husband & wife).

Everything I have accomplished is because of my family, and so it is to them that I dedicate this achievement. To you, Mom, for constantly reminding me who I am and what I am capable of ever since I was a little girl. To you, Dad, so you can call me “Dr. Piacentini” and know your daughter earned that title. To you, Nick, so I can be a sister worthy of such a hard-working, altruistic brother. To you, Grammie, so we can continue to wonder at God’s Creation together. To you, Nonna, e per la nostra famiglia italiana. To those who have passed from this life to the next – Granddaddy, Nonno Enzo, Pappy, Nonno Patrick – to honor your legacy. To all of you, so that I can make you proud. In the hard days when I felt like giving up, I kept going because of you, and so this achievement is for us.

Finally, thank you to my chosen family for rooting me on in my graduate school journey: Aunties Heather & Jess and everyone back in Central PA, my best friend Shaina and all my lifelong pals from dear old F&M, the McGinty & Candelori families who care for me as their own, and all others that have made a significant impact on my life. You have walked with me through this entire endeavor and have shown me that meaningful relationships transcend any distance, length of time, and even a global pandemic.

Long story short... I survived!

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	vi
List of Figures	ix
List of Tables	x
Chapter 1: Introduction	1
1.1 Human Immunodeficiency Virus.....	1
1.2 The HIV-1 capsid: Master regulator of post-entry events.....	4
1.3 Structural properties of the HIV-1 capsid	5
1.4 Experimental recapitulation of HIV-1 capsid assembly states.....	8
1.5 Dissertation overview	9
Chapter 2: Molecular determinants of PQBP1 binding to the HIV-1 capsid lattice.	10
2.1 Abstract	10
2.2 Introduction	10
2.3 Results	12
2.3.1 PQBP1 Nt directly binds the HIV-1 CA hexamer	12
2.3.2 CryoEM structure of the HIV-1 CA hexamer bound to PQBP1 Nt ..	17
2.3.3 A tri-aspartate motif is required for PQBP1 Nt binding to the HIV-1 CA hexamer.....	22
2.3.4 Binding of PQBP1 to the HIV-1 capsid lattice.....	25
2.4 Discussion.....	27
2.5 Methods.....	29
2.5.1 Protein production	29
2.5.2 Nanodifferential scanning fluorimetry	30
2.5.3 Pull-down assays.....	30
2.5.4 ISG54 induction assay.....	31
2.5.5 CryoEM structure determination.....	31
2.5.6 3D variability analysis	33

2.5.7 Coordinate modeling	33
2.5.8 Data deposition.....	33
2.6 Acknowledgements	33
Chapter 3: PQBP1 Ct is required to induce HIV-1 capsid lattice destabilization	35
3.1 Introduction	35
3.2 Results	36
3.2.1 Cryo-tomography analysis of PQBP1 binding to the HIV-1 capsid lattice	36
3.2.2 PQBP1 binding induces capsid lattice destabilization.....	37
3.2.3 PQBP1 self-association is mediated by Nt and Ct	40
3.2.4 PQBP1 Ct dimerization is required for cGAS signaling.....	42
3.3 Discussion.....	44
3.4 Methods.....	46
3.4.1 Protein production	46
3.4.2 Electron microscopy	47
3.4.3 Pull-down assays.....	48
3.4.4 Mass photometry	48
3.4.5 Proximal ligation assay.....	49
3.5 Acknowledgements	49
Chapter 4: The hexamer-like cross-linked CA pentamer binds ultra-potent capsid inhibitor GS-CA1	50
4.1 Introduction	50
4.2 Results	51
4.2.1 GS-CA1 binds to crosslinked HIV-1 CA pentamer.....	51
4.2.2 Structure of disulfide-stabilized CA pentamer in complex with GS-CA1	52
4.2.3 Structural comparison of HIV-1 CA hexamers and pentamers	55
4.3 Discussion	58
4.4 Methods.....	58
4.4.1 Protein production	58
4.4.2 Nanodifferential scanning fluorimetry	59

4.4.3 Crystallization	59
4.4.4 Structural determination	59
4.4.5 Structure comparisons.....	60
4.5 Acknowledgements	60
Chapter 5: Conclusion.....	61
5.1 Summary of results	61
5.2 Impacts and future directions	63
Chapter 6: References	64

List of Figures

Figure 1.1. Early events of the HIV-1 life cycle	3
Figure 1.2. Structural components of the mature HIV-1 capsid	6
Figure 1.3. Structures of small molecules binding to different sites on CA hexamers.....	7
Figure 2.1. PQBP1 is an intrinsically disordered protein (IDP)	11
Figure 2.2. Initial characterization of PQBP1 binding to soluble disulfide-stabilized HIV-1 CA A14C/E45C/W184A/M185A hexamers.....	13
Figure 2.3. The N-terminal (Nt) region of PQBP1 directly binds soluble disulfide-stabilized HIV-1 CA hexamer	14
Figure 2.4. PQBP1 Nt binding to the hexamer is mediated by charge interactions	16
Figure 2.5. CryoEM structure determination of soluble disulfide-stabilized HIV-1 CA hexamer in complex with PQBP1 Nt.....	18
Figure 2.6. CryoEM structure of CA hexamer bound to Nt.....	19
Figure 2.7. 3D variability analysis of the consensus cryoEM map	20
Figure 2.8. High-affinity hexamer binding requires a tri-aspartate motif in Nt	23
Figure 2.9. Thermostabilization of HIV-1 CA hexamer by PQBP1 Nt peptide fragments	24
Figure 2.10. PQBP1 binding to the HIV-1 capsid lattice in vitro	26
Figure 3.1. PQBP1 FL decorates cross-linked CA tubes	36
Figure 3.2. PQBP1 binding results in capsid lattice destabilization	38
Figure 3.3. Binding of PQBP1 affects CA tube morphology	39
Figure 3.4. PQBP1 Ct self-association is unmasked by Nt.....	41
Figure 3.5. C-terminal PQBP1 dimerization facilitates HIV-1-induced cGAS signaling..	43
Figure 4.1. Analysis of GS-CA1 binding to cross-linked CA pentamer	52
Figure 4.2. GS-CA1 structure bound to CA pentamer.....	55
Figure 4.3. Structural comparison of HIV-1 CA capsomers.....	56

List of Tables

Table 2.1. Binding parameters for PQBP1 peptides and disulfide-stabilized HIV-1 CA hexamer	15
Table 2.2. CryoEM data collection, structure refinement, and coordinate modeling statistics	21
Table 4.1. Structure determination statistics	53
Table 4.2. HIV-1 CA structures used for comparisons in Figure 4.3	57

Chapter 1: Introduction

1.1 Human Immunodeficiency Virus

Human immunodeficiency virus (HIV) attacks the body's immune system by infecting CD4⁺ T-cells,¹ dendritic cells,² and macrophages,³ increasing susceptibility to other diseases and, when untreated, leads to the development of acquired immune deficiency syndrome (AIDS).⁴ Although significant advances to combat the virus have been made over the last three decades, HIV remains a substantial global public health concern: according to a report published by the World Health Organization at the end of 2022, 39 million people were identified to be living with HIV, and 1.3 million were newly infected in that year alone.⁵ Preventative measures such as safe sex practices and pre-exposure prophylaxis (PrEP) or post-exposure prophylaxis (PEP) medications are the most effective methods to reduce HIV transmission.⁶ Should a person become infected with HIV, the main source of treatment is the use of antiretroviral therapy (ART), which suppresses viral replication and thus decrease the clinical burden of infection.⁷ Consistent use of ART can robustly reduce viral loads in HIV-positive individuals such that they no longer transmit the virus.⁸ However, even though ARTs can mitigate disease progression, it cannot cure HIV infection. In addition to causing adverse side effects being susceptible to drug-resistance mutations, ART is not readily and consistently accessible to all infected individuals, particularly those in resource-poor communities.^{5,9} This poses a serious problem since these therapies must be continuously taken to prevent viral rebound and subsequent disease progression. As such, the need persists for more permanent, manageable, and accessible treatment options.

HIV is a single-stranded, positive-sense RNA lentivirus of two types: type 1 (HIV-1) and type 2 (HIV-2). HIV-1 accounts for more than 95% of all infections and is the cause

of the global HIV pandemic, whereas HIV-2 is more geographically limited as it is endemic to West African regions.¹⁰ HIV-2 is more than 55% genetically different from HIV-1 and demonstrates lower transmission rates and slower progression than HIV-1.^{10–13} HIV-1 will be the primary focus within this dissertation, but it should not go unsaid that equivalent knowledge about HIV-2 is imperative for a comprehensive understanding of HIV and its impact on global public health.

The early stages of HIV-1 infection (**Figure 1.1**; reviewed in ¹⁵) begin with cellular entry of a mature viral particle, which occurs through receptor-mediated membrane fusion (reviewed in ¹⁴): the viral envelope glycoprotein Env binds to its CD4 receptor and a co-receptor, either C-C chemokine receptor 5 (CCR5) or C-X-C chemokine receptor 4 (CXCR4), on the host cell membrane. A hydrophobic peptide from one of Env's subunits then inserts into the membrane and refolds to bring the viral and host cell membranes into close proximity. The membranes fuse, and the viral core is released into the cytoplasm.¹⁴ This core particle consists of the viral genome, the nucleocapsid (NC) protein that coats the genome, replicative enzymes reverse transcriptase (RT) and integrase (IN), and the surrounding viral capsid. Cytoplasmic entry triggers activation of RT within the capsid, which begins reverse transcribing the viral RNA to form the reverse transcription complex (RTC).¹⁵ As reverse transcription continues, the viral core traffics to the nucleus along the host cell's microtubule network,^{16–19} docking at the nuclear pore complex (NPC) before undergoing nuclear import.^{20–37} Once inside the nucleus, reverse transcription will finish, resulting in formation of the pre-integration complex (PIC). Viral capsid uncoating releases the PIC,^{38–40} in which the 5' and 3' ends of the fully reverse-transcribed, double-stranded viral DNA is primed by IN for insertion into transcriptionally active sites of the host cell genome.^{26,28,41–43} Proper completion of these so-called early events relies on the assembled HIV-1 capsid,⁴⁴ which is essential for viral infectivity.^{21,45,46}

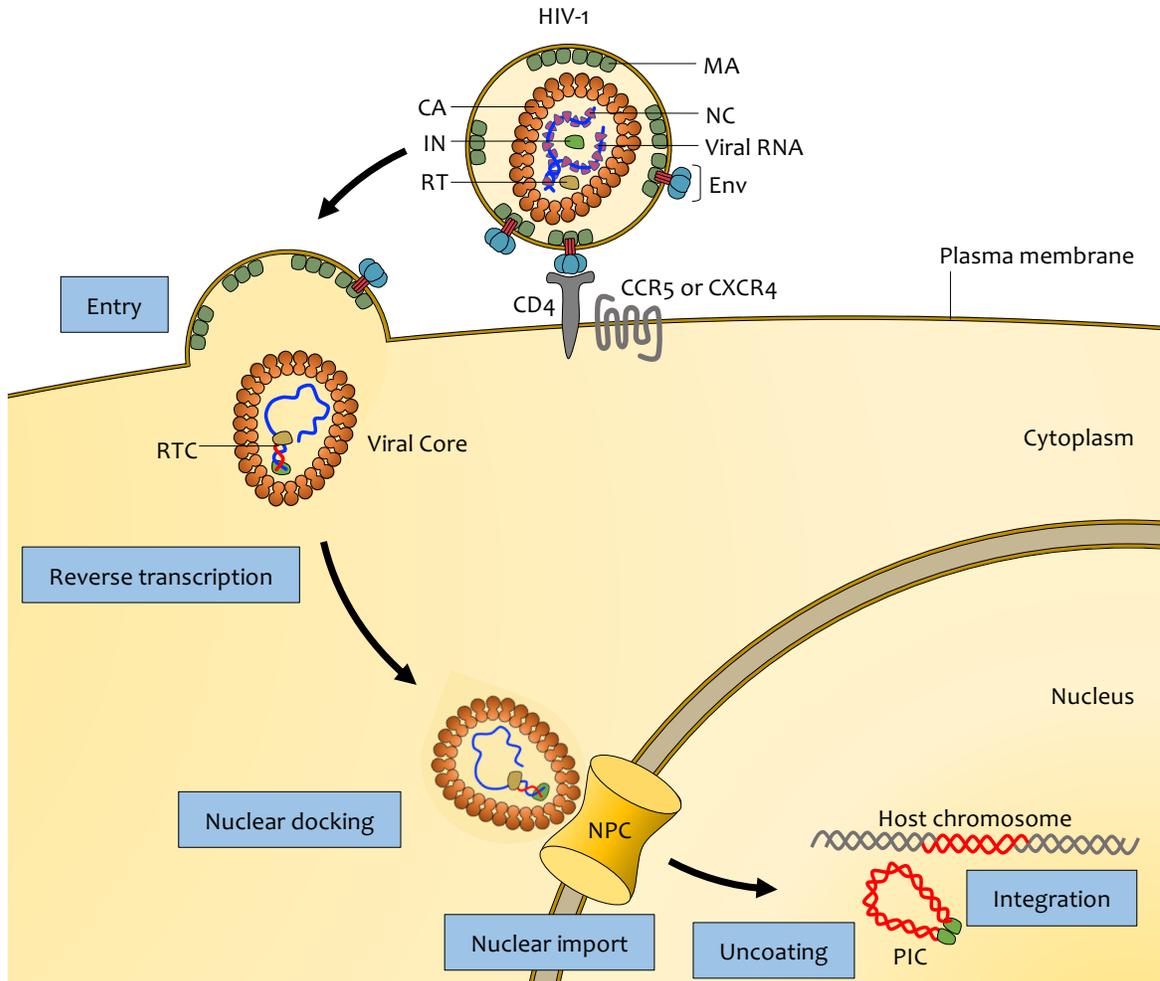


Figure 1.1. Early events of the HIV-1 life cycle. HIV-1 infection is initiated upon virion binding to a CD4 receptor and co-receptor (either CCR5 or CXCR4) on the plasma membrane of the host cell. Upon membrane fusion, the viral core is released into the cytoplasm. Reverse transcription begins, and the core is trafficked to the nucleus. Once the viral core arrives at the nucleus, it docks at the NPC and undergoes nuclear import. In the nucleus, the capsid uncoats, reverse transcription completes, and the provirus is integrated into host cell chromosomes. MA = matrix protein, NC = nucleocapsid protein.¹⁵

1.2 The HIV-1 capsid: Master regulator of post-entry events

As noted above, the HIV-1 capsid has many functions in early infection events. Upon viral particle release into the cell, the capsid provides the molecular platform for reverse transcription,^{47,48} protects the reverse-transcribing viral genome from innate immune sensors,^{2,49,50} and interacts with the microtubule network to traffic the viral core to the nucleus.^{16–19} The capsid is the determining factor of HIV-1's ability to infect non-dividing cells, facilitating nuclear import of the virus by directly docking at the NPC via interactions with nucleoporins (NUPs) and other import factors as the core is translocated into the host cell nucleus.^{15,20,21,23–30,32,33,37,51–60} The assembled capsid also functions in post-nuclear entry steps; it is thought to remain associated with the RTC until RT has finished and interacts with factors that target the virus to sites of DNA integration.^{32,61–63} Only when the capsid uncoats can integration of viral DNA into the host cell genome occur.^{41,64}

The HIV-1 capsid, specifically its structural stability, regulates successful completion of the events in the early viral life cycle. For example, premature capsid destabilization causes decreased RT activity, lack of viral nuclear import, and viral destruction as a consequence of early host cell detection and immune sensing.^{45,65–68} Conversely, overstabilization of the capsid stalls RT, blocks nuclear import, and prevents integration by inhibiting viral uncoating.^{41,64} Additionally, many crucial interactions between the capsid and its binding partners during infection steps occur at interfaces that do not exist outside the context of a fully-assembled capsid,⁶⁹ which will be discussed more in the following section. Thus, HIV-1 capsids with improper structural stability do not allow for productive viral infection.

1.3 Structural properties of the HIV-1 capsid

The viral capsid is composed of around 1500 copies of the capsid protein (CA),⁷⁰ a 231-residue, highly helical protein with an N-terminal domain (NTD) and C-terminal domain (CTD) separated by a flexible linker (**Figure 1.2A**).^{71,72} Using extensive intermolecular interactions, the CA protein self-assembles into protein shell (the capsid) containing around 250 hexamers and exactly 12 pentamers. The CA hexamer and pentamer each have two concentric rings: an inner ring of NTDs and an outer ring of CTDs (**Figure 1.2B**).^{69,72-76} The inner ring interactions are mediated by N-terminal helices 1, 2 and 3 of each protomer, which interact with one another to form a rigid helical core of either 18 or 15 helices in the hexamer or pentamer, respectively.^{72,75,77,78} Additionally, a crucial NTD-CTD interface forms between the NTDs and CTDs of neighboring protomers,^{72,75-77} an important feature that is unique to the mature assembled capsid.^{79,80} The outer CTD ring is responsible for interacting with CTDs of other hexamers and pentamers to form the extended capsid lattice.^{72,75,77-79} The flexibility of CTDs is imperative for these interactions as it permits proper lattice curvature and closing to form the asymmetric fullerene cone shell (**Figure 1.2C**).^{72,74,77,78,81}

The HIV-1 capsid is a binding target for various factors that can assist or restrict infection, often by modulating capsid stability. Many of these binding partners interact with the capsid at one of two different sites found in CA oligomers: the arginine 18 (R18) ring, and the NTD-CTD interface. The former is an electropositive pore made of five or six arginine residues that form the binding site for molecules containing acidic motifs. These include inositol-6-phosphate (IP6; **Figure 1.3A-C**)⁸², a host cell metabolite required for infection that can stabilize the capsid lattice throughout the viral life cycle,⁸³ deoxynucleotide triphosphates (dNTPs), which are transported into the capsid during reverse-transcription,⁸⁴ fasciculation and elongation protein zeta 1 (FEZ1), a kinesin-1

cofactor that promotes cellular trafficking of the virus,¹⁷ and tripartite motif (TRIM) proteins, known restriction factors of HIV-1 that are thought to attenuate infection by prematurely uncoating the capsid upon binding.⁸⁵ The other binding site is the NTD-CTD interface formed between CA monomers within pentamers and hexamers. This hydrophobic pocket binds molecules such as NUP153, which facilitates nuclear docking and import, and cleavage and polyadenylation specific factor 6 (CPSF6), which functions in nuclear import and in targeting viral DNA to transcriptionally-active chromatin. The NTD-CTD interface is also the target site of small molecule inhibitors, including PF74 (**Figure 1.3D-F**) and lenacapavir, which can both outcompete viral infection binding partners and modulate capsid stability as a means of restricting infection.^{30,55,86–88}

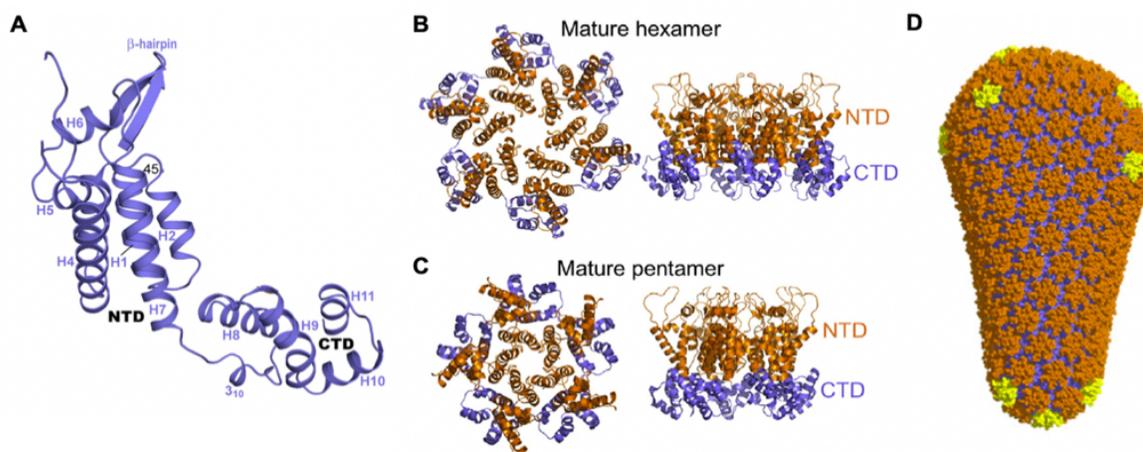


Figure 1.2. Structural components of the mature HIV-1 capsid. (A) Side view of a CA monomer. Secondary structural elements are labeled. Adapted from ⁷⁷. (B and C) Structures of the mature hexamer and pentamer, respectively. NTDs are shown in orange, and CTDs are shown in blue. (D) Structure of the mature HIV-1 capsid, which is a fullerene cone shell formed by the extended lattices of ~250 hexamers (orange and blue) and 12 pentamers (yellow and blue). B, C, and D adapted from ⁷⁵.

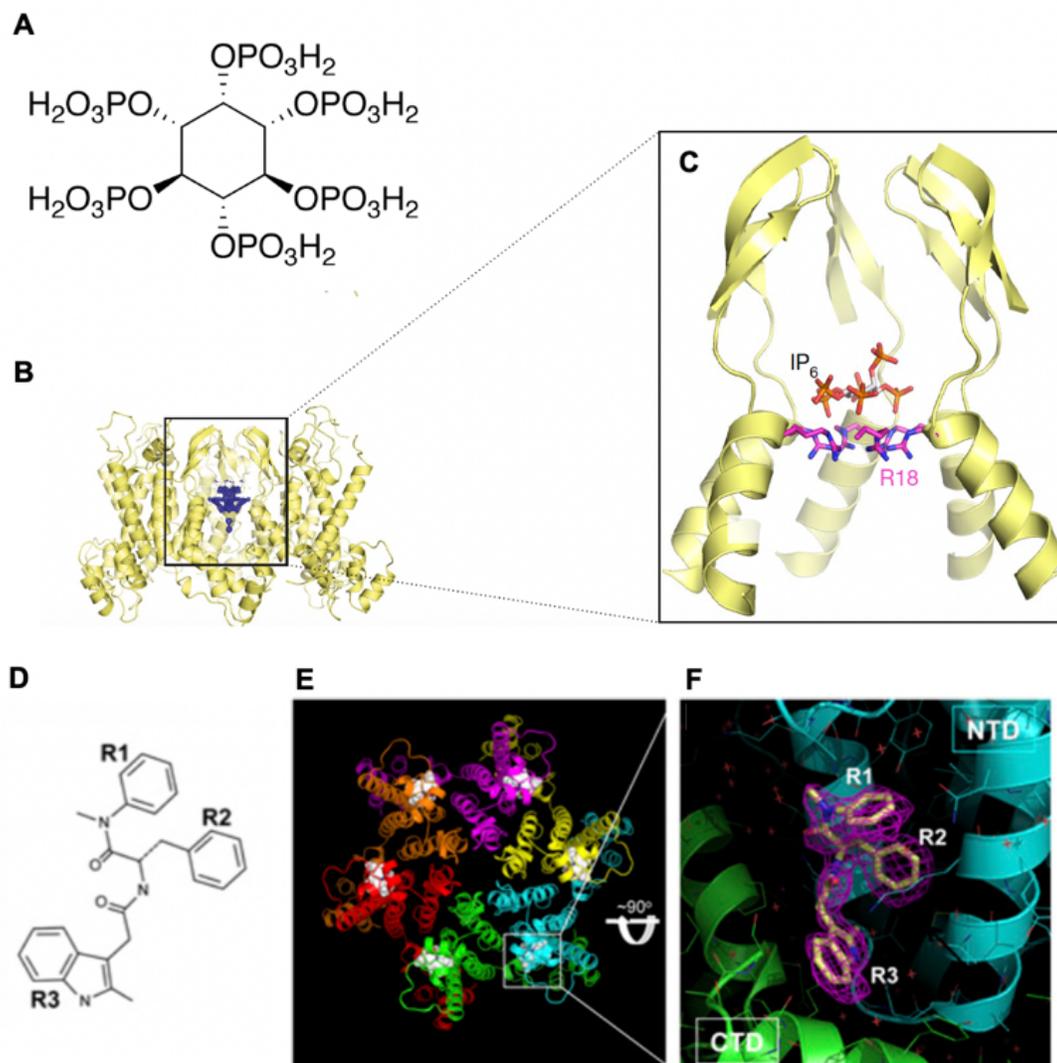


Figure 1.3. Structures of small molecules binding to different sites on CA hexamers. (A) Chemical structure of inositol-6-phosphate (IP6). (B) Side view ribbon diagram of a CA hexamer (yellow) with unbiased mF_o-DF_c difference density contoured at 2.2σ (blue mesh), indicating the binding site of IP6. (C) Illustration of the CA hexamer's electropositive pore formed by six R18 residues (magenta) binding to an IP6 molecule (orange). (D) Chemical structure of PF-350074 (PF74) with all three aromatic rings labeled (R1, R2, R3). (E) Top view of the CA hexamer with six PF74 molecules bound. Each CA monomer is represented by a different color, and PF74 is represented in white. (F) Unbiased mF_o-DF_c difference density contoured at 3σ (magenta mesh) at the NTD-CTD (cyan and green, respectively) interface, indicating the binding site of PF74 (modeled in yellow). Figures adapted from ⁸² (A), ⁸³ (B and C), and ³⁰ (D-F).

1.4 Experimental recapitulation of HIV-1 capsid assembly states

Although HIV-1 capsid assembly is well-characterized, the process of uncoating is much less understood and remains a topic of debate.^{15,89} Previous studies have suggested that the capsid uncoats in the cytoplasm either immediately after entry or gradually as the core is trafficked to the nucleus, though this would likely lead to host cell immune sensing.^{20,90–94} Other publications have proposed that the viral capsid can arrive at the nucleus fully intact but will undergo uncoating before import,^{21–23,54,95,96} citing the ~40 nm diameter permeability barrier of the NPC as too small to allow the ~60 nm capsid to fit through.^{75,97,98} More recent data have indicated that a fully- or nearly-intact capsid can enter the nucleus before it uncoats,^{22,41,64,99–101} suggesting that structural rearrangements of the capsid and/or the NPC in addition to viral capsid plasticity are likely key to achieving import.^{102–106} Nevertheless, the biochemical and biophysical mechanisms of viral uncoating have yet to be determined.

To evaluate HIV-1 capsid uncoating, and how capsid stability can be affected by binding factors, it is imperative to understand the function of each level of capsid assembly. Though previous studies utilizing cell-based, genetic, and/or molecular biology techniques have been informative, many biochemical and biophysical details about each assembly state and their functions remain unclear. In order to assess these properties, we established a series of protein reagents that recapitulate each distinct assembly species *in vitro*. In addition to purifying monomeric CA, both soluble CA oligomers (i.e., hexamers and pentamers) as well as extended CA lattices in the form of tubes can be obtained.^{77,107} Even more recently developed protocols allow for the isolation of entire replication-competent viral cores.⁴⁷ Employing these highly specific constructs allows us to systematically probe distinct characteristics of each level of assembly, revealing new

details that are fundamental to understanding the interactions of these proteins not just with one another but also with external binding partners.

1.5 Dissertation overview

This dissertation will discuss novel biophysical and biochemical details about the interactions between the HIV-1 capsid and its binding partners as well as the consequences of these interactions.

Chapters 2 and 3 focus on polyglutamine binding protein 1 (PQBP1), a host protein recently described to be a binding partner of the HIV-1 capsid as part of innate immune sensing through the cGAS/STING pathway. Chapter 2 shows the first structural data of binding and elucidates the primary mode of PQBP1 interaction with the capsid. Chapter 3 provides evidence for how PQBP1 self-association can promote both stable capsid binding and downstream signaling upon viral infection. These results give insight into how HIV-1 activates the innate immune response upon host cell entry.

In Chapter 4, the first structure of the cross-linked CA pentamer bound to GS-CA1, a small molecule inhibitor from which lenacapavir is derived, is presented. Structural analyses demonstrate the hexamer-like characteristics of this pentamer, which may be a biologically relevant conformation required for proper capsid function.

Collectively, this work reveals important details about the HIV-1 capsid, its binding partners, and the downstream consequences of binding. Understanding these capsid dynamics is fundamental to the study of post-cell entry viral events and will contribute to future therapeutic efforts focused on mitigating HIV-1 infection at large.

Chapter 2: Molecular determinants of PQBP1 binding to the HIV-1 capsid lattice

This chapter was reformatted from The Journal of Molecular Biology:

Piacentini, J., Allen, D.S., Ganser-Pornillos, B.K., Chanda, S.K., Yoh, S.M., and Pornillos, O. Molecular determinants of PQBP1 binding to the HIV-1 capsid lattice. *J Mol Biol.* **436**(4), 168409 (2024).

2.1 Abstract

Human immunodeficiency virus type 1 (HIV-1) stimulates innate immune responses upon infection, including cyclic GMP-AMP synthase (cGAS) signaling that results in type I interferon production. HIV-1-induced activation of cGAS requires the host cell factor polyglutamine binding protein 1 (PQBP1), an intrinsically disordered protein that bridges capsid recognition and cGAS recruitment. However, the molecular details of PQBP1 interactions with the HIV-1 capsid and their functional implications remain poorly understood. Here, we show that PQBP1 binds to HIV-1 capsids through charge complementing contacts between acidic residues in the N-terminal region of PQBP1 and an arginine ring in the central channel of the HIV-1 CA hexamer that makes up the viral capsid. These studies reveal the molecular details of PQBP1's primary interaction with the HIV-1 capsid and suggest that additional elements are likely to contribute to stable capsid binding.

2.2 Introduction

Cyclic GMP-AMP synthase (cGAS) is a pattern recognition receptor (PRR) that induces an interferon response upon recognition of cytosolic double-stranded DNA

(dsDNA).^{108–110} The cGAS pathway functions as a first-line defense against DNA and RNA viruses, including HIV-1.^{2,49,111,112} Reverse-transcribed HIV-1 DNA is one of the pathogen-associated molecular patterns (PAMPs) that can be recognized by cGAS, but these nascent transcripts are normally sequestered and protected within the viral capsid. Activation of cGAS during HIV-1 infection requires polyglutamine binding protein 1 (PQBP1), which binds to the capsid and recruits cGAS to the viral transcripts.^{113,114}

PQBP1 is an intrinsically disordered protein (IDP) of 265 amino-acid residues (**Figure 2.1A**) that is found both in the cell nucleus and cytoplasm.^{115–117} Being an IDP, PQBP1 lacks a well-ordered globular fold, apart from a small WW domain (**Figure 2.1B**);^{115,117–119} this property makes it challenging to define the protein's structure-function correlates. Previous studies have shown that PQBP1 recognizes the HIV-1 capsid through its N-terminal region, whereas cGAS recruitment and activation requires the WW domain as well as elements in the C-terminal region (**Figure 2.1A**).^{113,114}

However, many details of these separate activities and how they mechanistically integrate to enable HIV-dependent cGAS signaling remain unknown.

In this study, we focus on understanding the molecular mechanism by which PQBP1 binds and recognizes the HIV-1 capsid. Recent studies have revealed that PQBP1 binding to the capsid requires Arg18 in HIV-1 CA, the viral protein that makes up the capsid

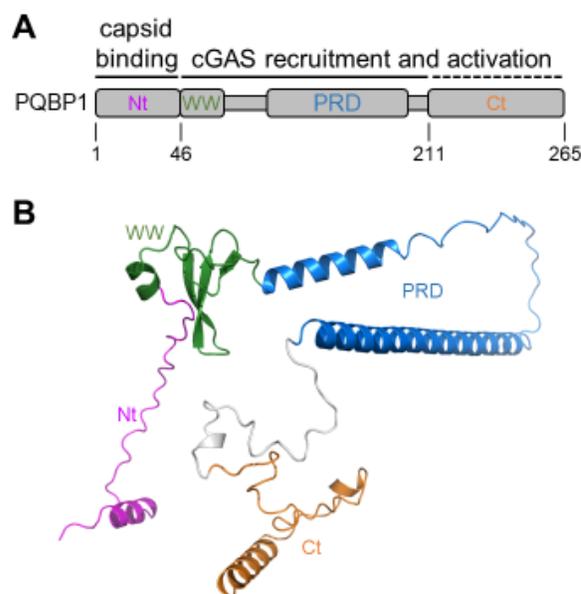


Figure 2.1. PQBP1 is an intrinsically disordered protein (IDP). (A) Schematic of the protein primary sequence. Functionally-defined regions for capsid binding and cGAS recruitment and activation are indicated. (B) Ribbons representation of predicted PQBP1 structure by AlphaFold.¹¹⁹ Functional domains are shown in different colors and labeled.

subunits. Six Arg18 residues form a ring or annulus within the central channel of the CA hexamer, which is one of the oligomeric building blocks of the capsid. The Arg18 ring is a well-characterized binding site for acidic ligands, including nucleotide triphosphates,⁸⁴ the capsid-stabilizing factor inositol hexakisphosphate (IP₆),^{83,120} and the transport factor FEZ1, which contains poly-glutamate motifs.¹⁷ The N-terminal (Nt) region of PQBP1 also contains acidic motifs,¹¹⁴ leading to the inference that, analogous to the above ligands, PQBP1 also binds the Arg18 ring via charge complementation. However, direct demonstration of this interaction has not been reported and the corresponding amino-acid residue determinants have not been thoroughly defined. Likewise, whether and how other regions of PQBP1 contribute to productive capsid recognition have not been elucidated. To address these gaps in knowledge, we used purified PQBP1 proteins and disulfide-stabilized HIV-1 CA assemblies to reconstitute PQBP1-capsid complexes in vitro for biochemical and structural analyses. Our results uncover the molecular details of the minimal binding interaction between PQBP1 Nt and the HIV-1 CA hexamer. Furthermore, we also report evidence that PQBP1 regions outside of Nt contribute to stable capsid binding. We propose that, similar to other well-characterized factors that target the HIV-1 capsid, PQBP1 likely requires avidity for functional capsid recognition.

2.3 Results

2.3.1 PQBP1 Nt directly binds the HIV-1 CA hexamer

We purified recombinant, full-length PQBP1 (**Figure 2.2A**) and characterized its binding to soluble, disulfide-crosslinked HIV-1 CA hexamers. As described before, the crosslinked CA hexamer is held together by designed disulfides (A14C/E45C), which, in combination with mutations (W184A/M185A) that prevent assembly of the extended

capsid lattice, allow preparation of discrete hexamers in solution.¹⁸ These hexamers have been used in many studies of HIV-1 capsid-binding factors. To measure binding, we utilized nanodifferential scanning fluorimetry (nanoDSF), a technique used to measure the thermostability of folded proteins. In this experiment, protein unfolding is indicated by the shift in peak Trp fluorescence as these residues transition from hydrophobic (protein interior) to hydrophilic (exposed to solvent) environments. Since the CA hexamers contain Trp residues only in the N-terminal domain (NTD), their nanoDSF profiles report on the unfolding of the central NTD ring that contains the Arg18 ring. The hexamer displayed a single unfolding transition, with an apparent melting temperature (T_m) of 50.7 ± 0.5 °C (**Figure 2.2B**, dark blue curves). On the other hand, purified PQBP1, which contains three Trp residues, displayed a non-canonical nanoDSF profile indicative of a largely unfolded state in solution (**Figure 2.2B**, orange curves). A sample containing both proteins revealed a shift in CA's apparent T_m by around 2°C (**Figure 2.2B**, blue curves). This degree of thermostabilization in nanoDSF and DSF assays is generally accepted to indicate a direct binding interaction.¹²¹

To begin mapping the binding determinants in PQBP1, we tested various fragments for their ability to increase thermostability of the CA hexamer. Because PQBP1

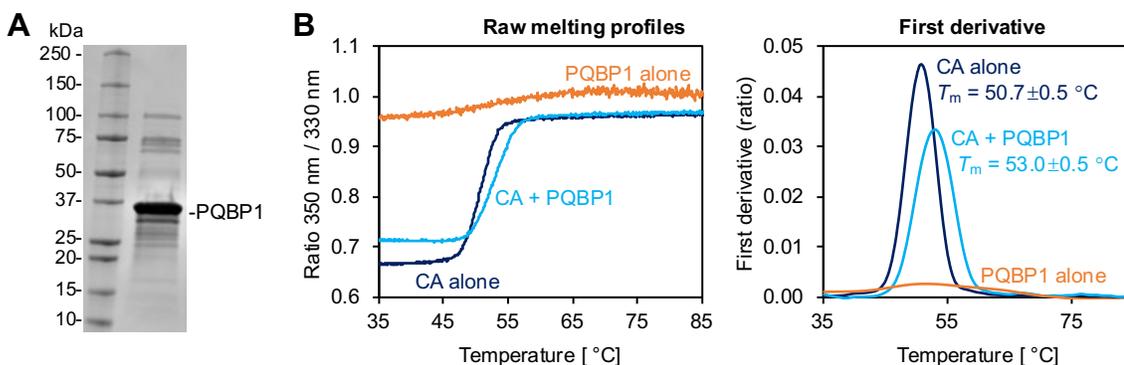


Figure 2.2. Initial characterization of PQBP1 binding to soluble disulfide-stabilized HIV-1 CA A14C/E45C/W184A/M185A hexamers. (A) SDS-PAGE profile of full-length, C-terminally His-tagged recombinant PQBP1. (B) NanoDSF melting profiles (left panel) and first derivative curves (right panel) of the indicated protein samples. Results are representative of two experiments.

is an IDP, we used functional annotations (e.g., **Figure 2.1A**) rather than structurally defined domains for fragment design (**Figure 2.3A**). Results show that the PQBP1 fragment spanning Nt (residues 1-46) shifted the apparent T_m of the hexamer (**Figure 2.3B**, compare magenta and black curves). In contrast, the WW domain, polar-rich domain (PRD) and C-terminal (Ct) fragments had no effect (**Figure 2.3B**, compare green, blue and orange curves to black). Dose-response curve fitting showed that Nt binds the hexamer with an apparent dissociation constant (K_d) of around 500 nM (**Figure 2.3C** and **Table 2.1**), indicating a relatively high-affinity complex in vitro.

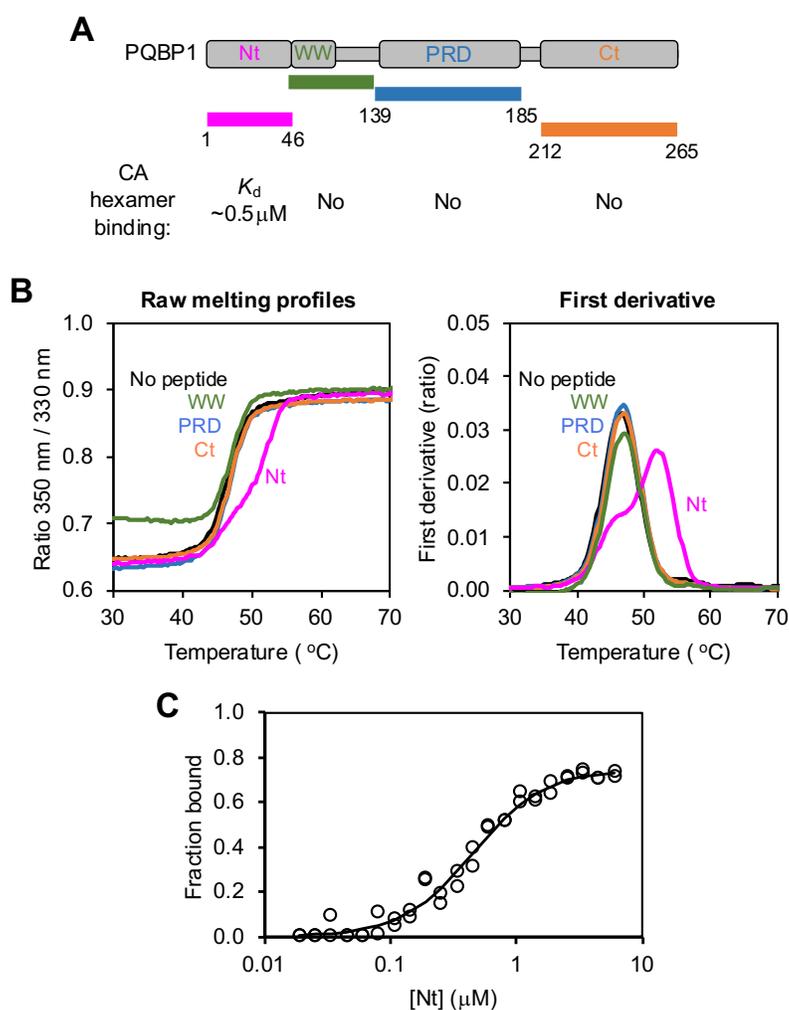


Figure 2.3. The N-terminal (Nt) region of PQBP1 directly binds soluble disulfide-stabilized HIV-1 CA hexamer. (A) Schematic of synthetic PQBP1 peptides spanning Nt, WW, PRD (polar-rich domain), and Ct (C-terminal region). **(B)** NanoDSF raw melting profiles (left panel) and first derivative curves (right panel) of HIV-1 CA hexamer in the presence of the indicated PQBP1 constructs. Results are representative of at least two experiments. **(C)** Nt binding data were fit to a simple binding model. Results are representative of 4 independent experiments, each with 2-3 replicates.

Table 2.1. Binding parameters for PQBP1 peptides and disulfide-stabilized HIV-1 CA hexamer

PQBP1 peptide	Apparent K_d (nM)	Hill coefficient	CA sp. activity
Nt	503±66	1.6±0.1	0.8
15-5	85.0±1.0	1.8±0.4	0.8
15-6	54.0±0.4	3.0±0.2	0.8

Note: K_d and Hill coefficient were treated as floating parameters during fitting, but not CA specific activity. Values are presented as mean±S.D. of four (Nt) or three (15-5 and 15-6) independent determinations.

As noted previously,¹¹⁴ Nt contains clusters of acidic residues, and in particular two tripeptide motifs (₂₅EEE₂₇ and ₃₄DDD₃₆). To determine if these motifs contribute to the interaction, we tested an Nt construct harboring GGG substitutions for each of the motifs (**Figure 2.4A,C**). The mutant Nt did not bind the CA hexamer.

The Arg18 ring of the HIV-1 CA hexamer has also been implicated in PQBP1 binding,¹¹⁴ so we performed the same experiments with an R18A mutant of the hexamer (**Figure 2.4B,C**). None of the PQBP1 fragments, including Nt, induced thermostabilization in this case, confirming the importance of Arg18 in PQBP1 binding to the CA hexamer.

Finally, we used cryoEM to compare 2D class averages of the non-mutated CA (Arg18) and R18A mutant hexamers. We observed extra density within the central channel of the non-mutated hexamer but not the R18A hexamer (red arrows in **Figure 2.4D**). These results are consistent with the interpretation that, similar to other acidic ligands,^{17,83,84,120} PQBP1 Nt binds to the ring of six Arg18 residues located within the hexamer channel.

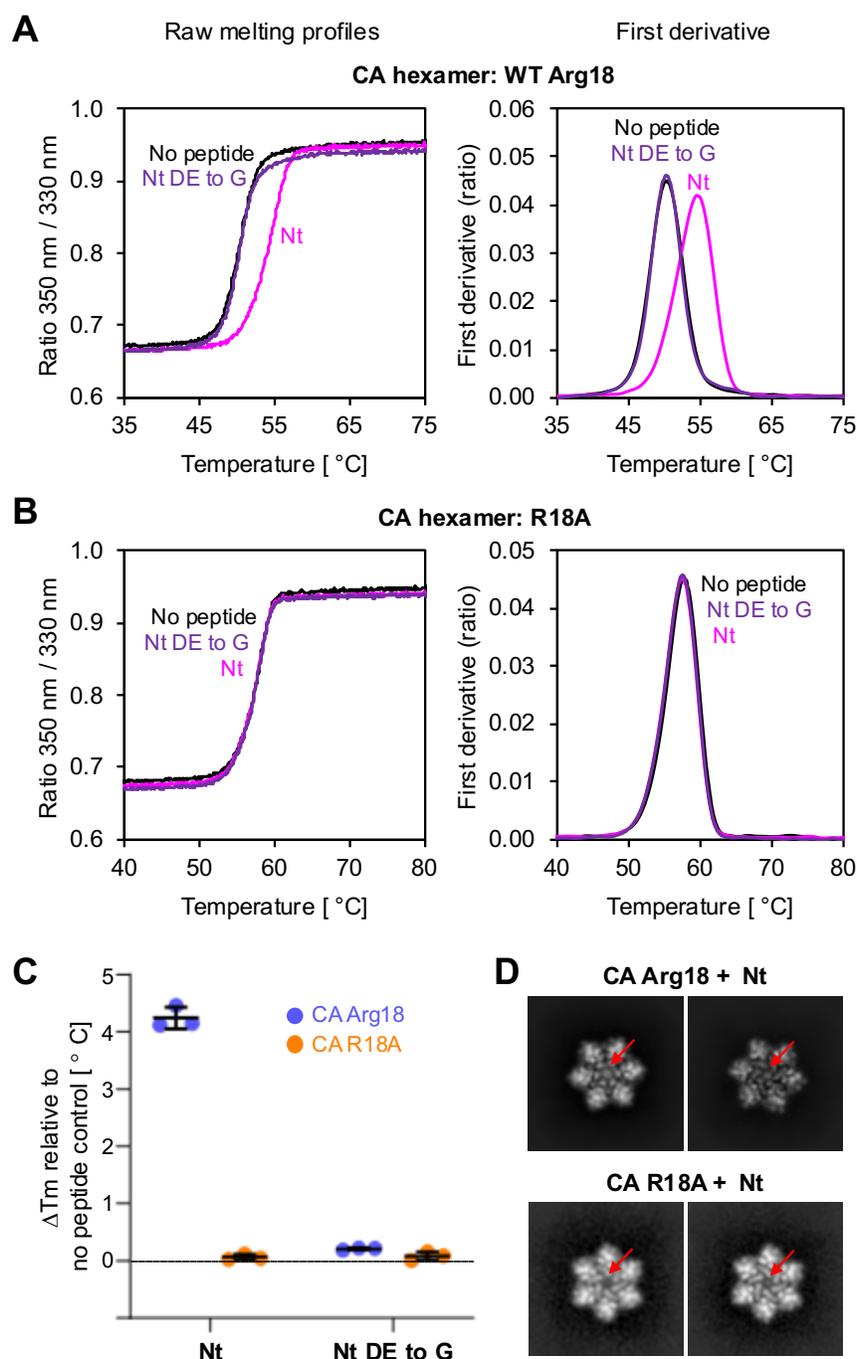


Figure 2.4. PQBP1 Nt binding to the hexamer is mediated by charge interactions. (A) and (B) NanoDSF raw melting profiles (left panels) and first derivative curves (right panels) of HIV-1 CA hexamer (0.83 μ M) containing non-mutated Arg18 (A) or R18A mutation (B), after incubation with 5 μ M Nt peptide (red), or Nt DE-to-G mutant (purple), or without peptide (black). Data are representative of at least 3 technical replicates. **(C)** Summary of thermostabilization data, presented as mean \pm S.D. of three independent experiments. **(D)** Representative 2D class averages of the indicated CA hexamers after incubation with 3-fold excess Nt peptide. Red arrows indicate the central channel.

2.3.2 CryoEM structure of the HIV-1 CA hexamer bound to PQBP1 Nt

To directly visualize binding of PQBP1 Nt to the HIV-1 CA hexamer, we vitrified samples in cryoEM grids, collected cryoEM image data, and solved the structure by using single-particle averaging techniques (**Figure 2.5A** and **Table 2.2**). The hexamers had a marked preferred orientation on the grids, and so we combined particle sets picked from images collected at 0° and 30° tilts. An initial consensus map of the CA hexamer bound to Nt was refined to a nominal resolution of 3.3 Å (according to the gold-standard 0.143 cut-off) without imposing symmetry (C1) (**Figure 2.5B** and **Table 2.2**). The CA residues are very well-defined, indicating that the map is of high quality (**Figure 2.5C**). The extra density within the central channel that we attribute to Nt appears as an elongated blob that has proteinaceous features but is less well-defined than CA densities (**Figure 2.6A,B**). The peptide density is not 6-fold symmetric. Our interpretation is that even though the ordered peptide segment is only 3% that of the CA subunits, it is sufficiently well-ordered to contribute to the alignment. However, the peptide remains the average of multiple bound orientations, indicating that the 6-fold symmetric features of the CA ring still dominates particle alignment.

In an attempt to improve the Nt densities, we tested a variety of post-alignment processing and 3D classification strategies. None had sufficient power to resolve a unique peptide orientation, but 3D variability analysis (3DVA)¹²² (see Methods and **Figure 2.7** for details) highlighted some key features. Firstly, the consensus map indicates a single peptide chain bound within the central channel of the hexamer. We could observe evenly spaced ridges in the peptide density (dashed lines in **Figure 2.6B**), which we interpret to be consistent with a helical configuration. Although requiring further confirmation, this feature would be consistent with the typical IDP behavior of binding-induced folding into

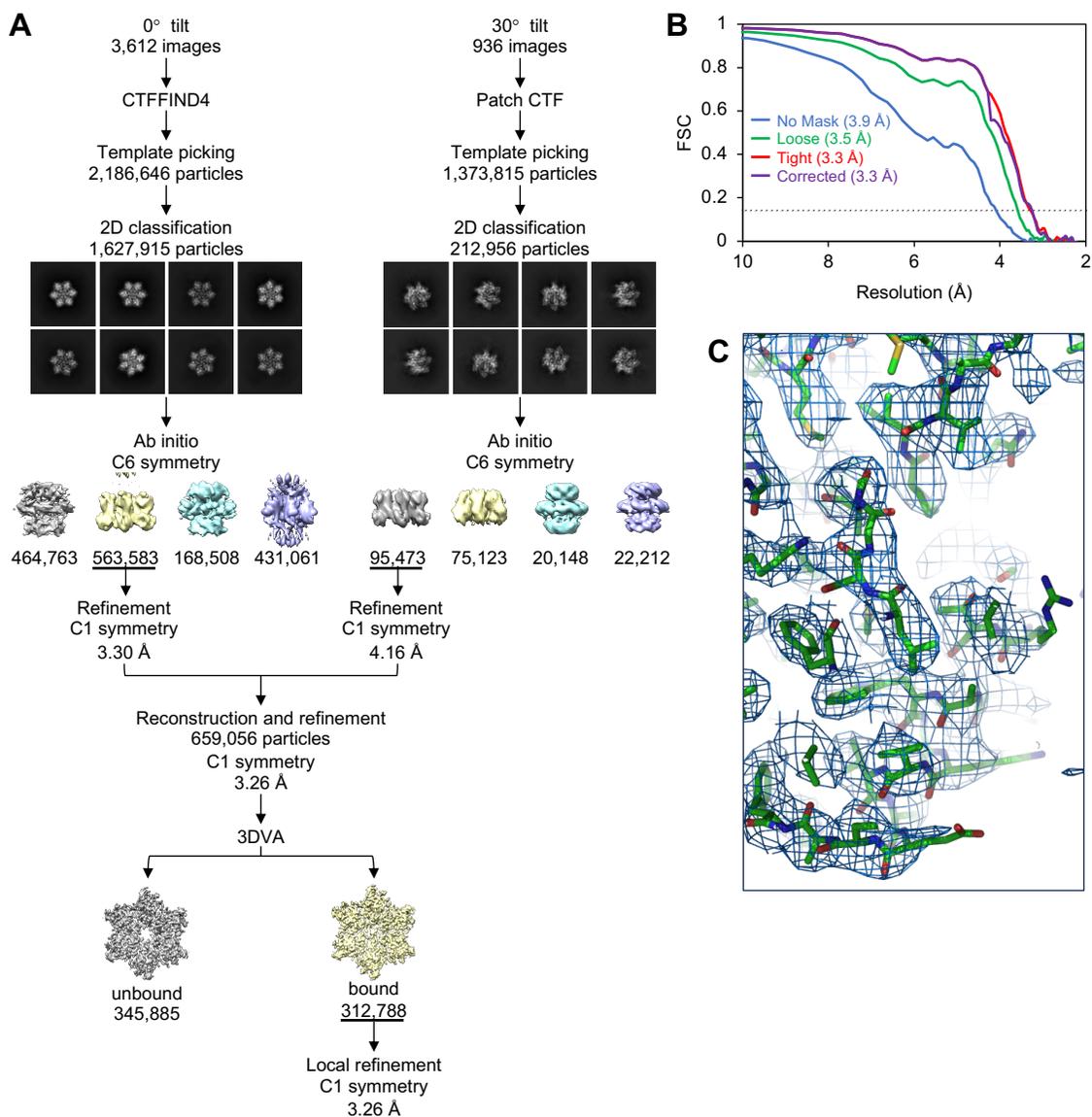


Figure 2.5. CryoEM structure determination of soluble disulfide-stabilized HIV-1 CA hexamer in complex with PQBP1 Nt. (A) Single particle averaging workflow. **(B)** Fourier shell correlation curves for the final local refinement of Nt-bound CA hexamer. **(C)** Density for the CA portion of the consensus map (blue mesh), with coordinate model (green sticks).

canonical secondary structural motifs when complexed with their partners.¹²³ In principle, the density can also accommodate two chains in elongated or strand configuration. However, 2D class averages (face views only) confirmed that the hexamers predominantly contain a single density feature in the central channel, and only a very small proportion (<1%) contains two peptide densities (**Figure 2.7D**). Secondly, although the peptide

density in the consensus map is centered within the hexamer, the peptide is actually positioned asymmetrically. This key point is illustrated by modes of variation, derived from 3DVA, that run along vectors perpendicular to the central axis of the hexamer channel (**Figure 2.6C**). Maps reconstructed from the extreme positions of these vectors (**Figure 2.7A-C** and **Table 2.2**) show that the long axis of the peptide does not coincide with the channel center (**Figure 2.6D**). Thirdly, the maps clearly indicate direct contacts to the CA Arg18 sidechains. Notably, the 3DVA subset maps show contacts to only one or two, and not all six, Arg18 residues (**Figure 2.6D**, red asterisks).

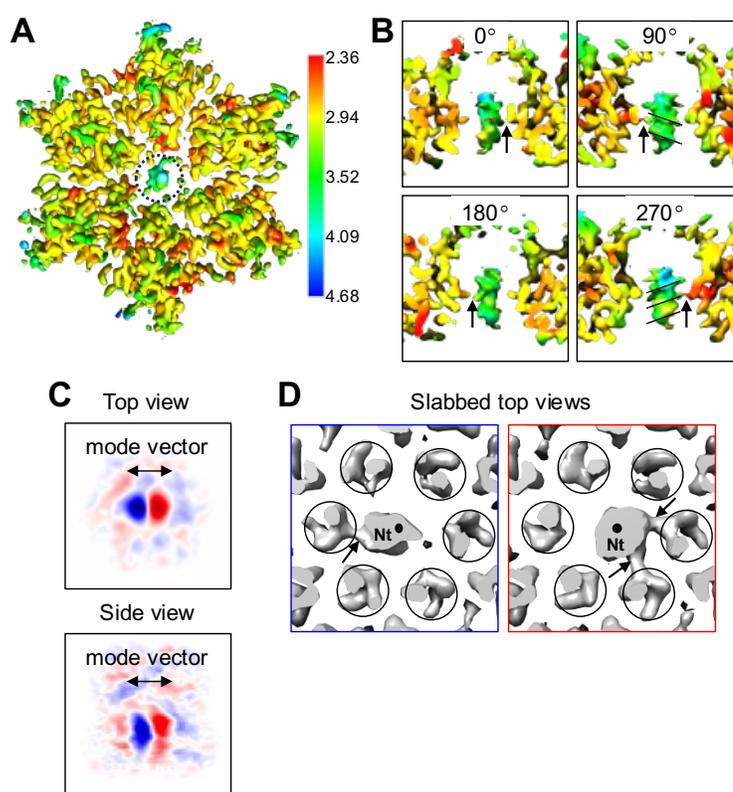


Figure 2.6. CryoEM structure of the CA hexamer bound to Nt. (A) Top view of the consensus map refined with C1 symmetry, colored according to local density. The peptide in the central channel is encircled. (B) Orthogonal side views of the peptide density. Arrows indicate apparent contacts between the peptide and Arg18. Dashed lines indicate evenly spaced ridges that we interpret to be consistent with a predominant helical configuration. (C) Orthogonal views of central slices through the major 3D variability component for the consensus map, as determined by 3DVA analysis. Positive (red) and negative (blue) values indicate different positions of the peptide relative to the central hexamer channel. (D) Reconstructions of particle subsets, without further alignment, revealing two peptide orientations determined by 3DVA. Circles indicate densities for CA helix 1 surrounding the central channel. Black dots indicate the central axis of the hexamer channel. Arrows indicate clearly defined contacts between the peptide and Arg18.

In summary, our cryoEM structure empirically establishes that Nt indeed binds within the CA hexamer channel and directly contacts Arg18. Unfortunately, resolving the bound peptide structure within its 6-fold symmetric binding site is pushing the limits of the single-particle averaging technique, and will have to await further studies.

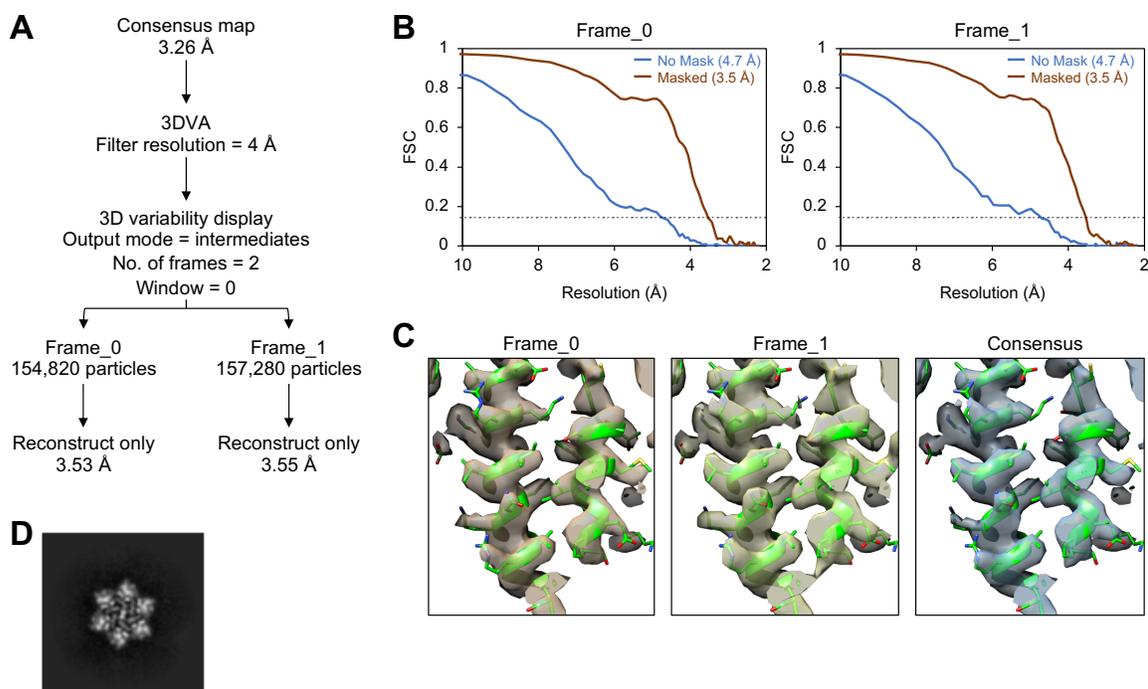


Figure 2.7. 3D variability analysis of the consensus cryoEM map. (A) 3DVA workflow. **(B)** FSC curves for indicated reconstructions, which were calculated using the particle positions and orientations derived from the consensus map refinement, i.e., without any additional alignment. **(C)** Representative densities from the CA portions of the maps, with the consensus map (right) shown for reference. **(D)** 2D class average of a minor population (~200 particles) suggesting two peptide densities inside the hexamer channel. This could represent hexamers with either two bound peptide molecules or a single peptide in hairpin configuration (our preferred interpretation). A third formal possibility is that each individual hexamer contains only a single bound peptide, and yet CA-dominated alignment just so happened to produce an average with two strong density features in the channel; we surmise that this is unlikely.

Table 2.2. CryoEM data collection, structure refinement, and coordinate modeling statistics

Data collection			
Magnification	81,000x		
Voltage	300 kV		
Electron dose	50 e/Å ²		
Defocus range	0.5 to 2.5 μm		
Pixel size	1.08 Å		
Image processing and map calculation			
	Consensus	3DVA Frame 0	3DVA Frame 1
Symmetry imposed	C1	C1	C1
No. of particles	312,788	154,820	157,280
Map resolution / FSC threshold	3.3 Å / 0.143	3.5 Å / 0.143	3.5 Å / 0.143
Map resolution range* (Å)	2.4-31.4 Å	2.3-9.7 Å	2.3-10.0 Å
Coordinate modeling and refinement			
Initial model	PDB 3H47		
Model resolution / threshold	3.2 Å / 0.143		
Map sharpening B-factor	109 Å ²		
Mean real-space cross-correlation (CC)	0.81		
No. of non-hydrogen atoms	9,678		
Protein residues	1,260		
Mean B-factor	74.26 Å ²		
RMSD bond lengths	0.002 Å		
RMSD bond angles	0.443°		
MolProbity score / clash score	1.61 / 4.89		
Poor rotamers	2.71%		
Ramachandran favored / outliers / z-score	98.9% / 0% / 2.36		

2.3.3 A tri-aspartate motif is required for PQBP1 Nt binding to the HIV-1 CA hexamer

Assuming an α -helix configuration, the bound Nt density would be consistent with at most 15 amino-acid residues, i.e., one-third of Nt. To more precisely determine the portion of Nt that binds the hexamer, we used nanoDSF to test 15-mer peptides with 10-residue overlaps (**Figure 2.8A**). Results indicate that 15-mer peptides containing either the EEE or DDD motifs displayed measurable binding (**Figure 2.8A** and **Figure 2.9B**), but two fragments containing the DDD motif bound with the highest affinities (85 nM for 15-5 and 54 nM for 15-6) (**Figure 2.8B** and **Table 2.1**). Experiments with 10-mer and 20-mer fragments produced similar results (**Figure 2.9A,B**). To confirm the importance of Asp over Glu, we generated variants of peptide 15-6 (EIIAE DYDDD PVDYE), which bound the hexamer with highest affinity. As expected from results with Nt, substitution of the DDD motif for AAA abolished binding (**Figure 2.9C**). Interestingly, a peptide containing only Asp residues (DIIAD DYDDD PVDYD) afforded higher thermostabilization of the CA hexamer compared to a peptide containing only Glu residues (EIIAE EYEEE PVEYE) (**Figure 2.9C**). These results suggest that, in addition to the charge complementing interactions with Arg18, steric requirements help to define binding specificity of peptide ligands to the CA hexamer, at least in vitro.

To confirm the importance of the above acidic motifs in Nt for PQBP1-mediated sensing of HIV-1, we assayed for HIV-dependent ISG54 induction in the monocyte-like THP-1 cell line differentiated with PMA (phorbol 12-myristate-13-acetate), an established model system for human macrophages (**Figure 2.8C**). Cells stably expressing the indicated siRNA-resistant PQBP1-eYFP constructs were transfected with siRNA against endogenous PQBP1, then infected with HIV-1, followed by RT-qPCR analysis for ISG54

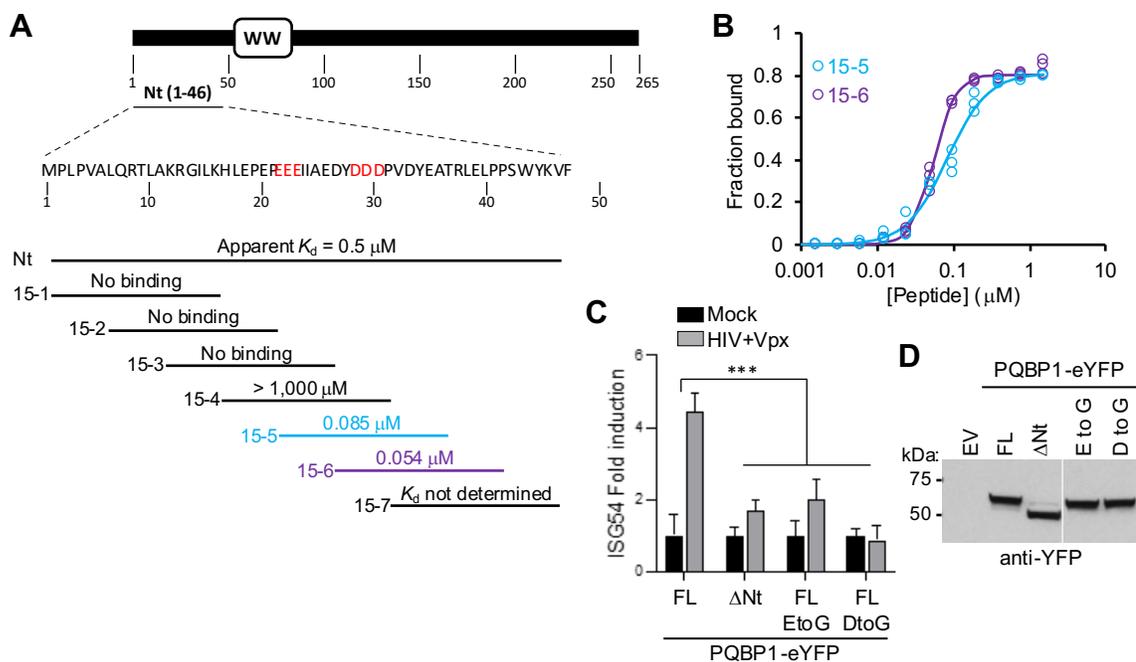


Figure 2.8. High-affinity hexamer binding requires a tri-aspartate motif in Nt. (A) Sequence of Nt and positions of 15-mer peptides. (B) Binding data for peptides 15-5 and 15-6. Data are representative of three independent experiments, each in three replicates. (C) HIV-dependent ISG54 induction by the indicated PQBP1-eYFP constructs in differentiated THP-1 cells, shown as mean \pm S.D. Experiments were done in biological triplicate; shown is representative of two independent experiments. ***, $p < 0.001$ (two-way ANOVA). (D) Anti-YFP immunoblot of THP-1 whole cell lysates to detect expression levels of indicated PQBP1-eYFP constructs. EV, empty vector.

expression (see Methods for details). As previously shown,¹¹⁴ control experiments with cells expressing wild type, full-length (FL) PQBP1-eYFP induced >4-fold higher ISG54 mRNA expression compared to mock-infected cells (positive control), whereas the Nt-deletion mutant (ΔNt) had attenuated (<2-fold) response (negative control). We then tested E-to-G and D-to-G full-length PQBP1-eYFP mutant constructs, in which either the ₂₅EEE₂₇ or ₃₄DDD₃₆ motif was substituted for GGG. Both mutants induced ISG54 levels comparable to ΔNt . Western blotting confirmed that all PQBP1-eYFP constructs expressed at comparable levels (Figure 2.8D). These data confirm the importance of the acidic motifs in Nt for PQBP1-mediated innate immune responses to HIV-1. Taken together, our aggregate data also indicate that the pronounced difference between DDD and EEE in

context of hexamer binding in vitro is attenuated in context of PQBP1's overall anti-viral activity in cells. Our interpretation is that functional binding in cells requires both acidic motifs (see also Discussion).

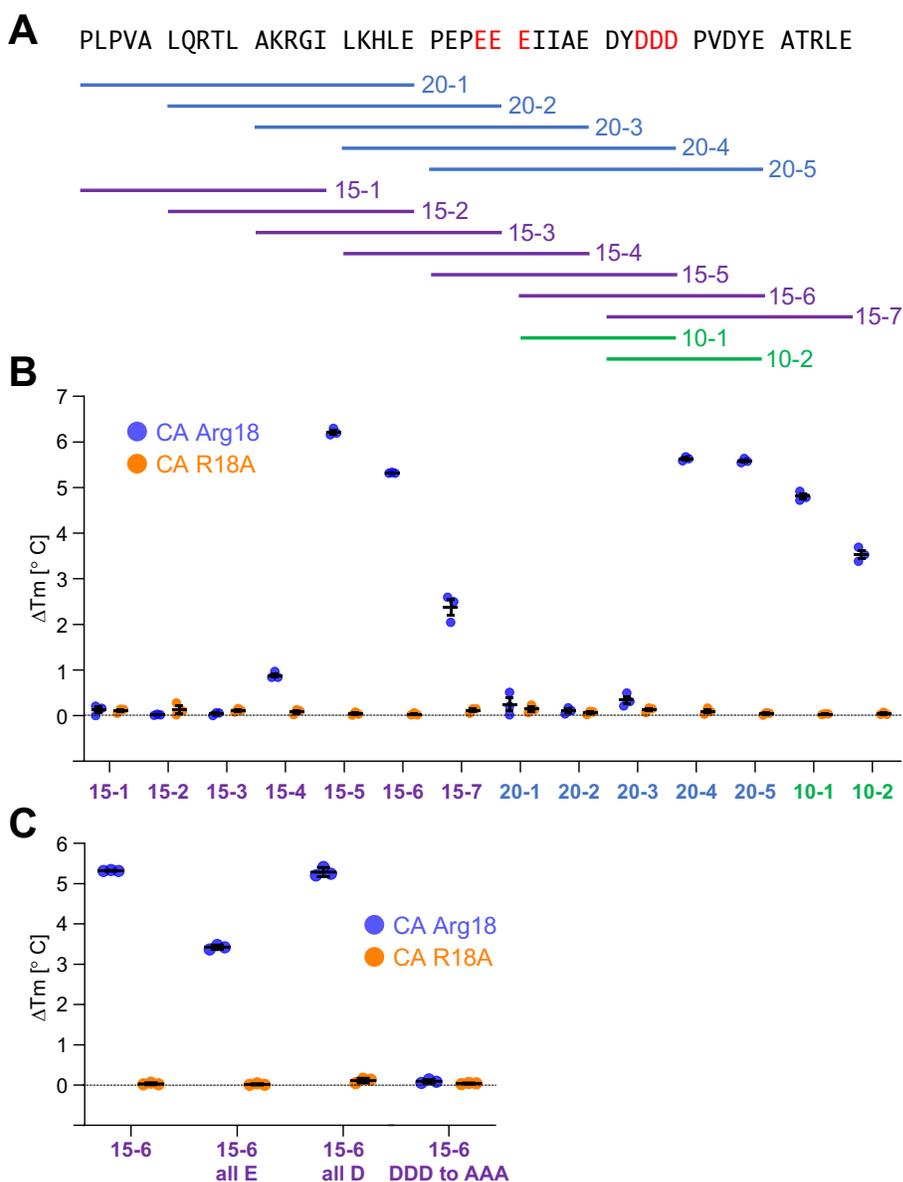


Figure 2.9. Thermostabilization of HIV-1 CA hexamer by PQBP1 Nt peptide fragments. (A) Primary sequences of 20-mer (blue), 15-mer (purple), and 10-mer (green) peptides. All were obtained by chemical synthesis, with N-terminal acetylation (except for peptides starting with proline) and C-terminal amidation. (B) Thermostabilization of the indicated hexamers by 20-mer, 15-mer, and 10-mer peptides, presented as mean \pm S.D. of three independent experiments. (C) Thermostabilization of the indicated hexamers by variants of the 15-6 peptide, presented as mean \pm S.D. of three independent experiments.

2.3.4 Binding of PQBP1 to the HIV-1 capsid lattice

In cells, the functional target of PQBP1 is the assembled capsid, and indeed PQBP1 is shown to colocalize with incoming HIV-1 cores.¹¹⁴ We therefore extended our binding studies by using full-length recombinant PQBP1 protein and HIV-1 capsid-like assemblies in vitro. Although IP₆-induced HIV-1 CA fullerene cones are more accurate mimics of actual viral capsids,¹²⁴ our results with the hexamer indicated that PQBP1 binding will be inhibited by the mM levels of IP₆ in these assemblies. Experiments were thus performed with A14C/E45C HIV-1 CA tubes (lacking the W184A/M185A mutations of the soluble, crosslinked CA hexamer) because these tubes can be reconstituted in almost any buffer solution (i.e., without IP₆) once the disulfide crosslinks form.¹⁰⁷ The disulfide-crosslinked tubes are established to be functional mimics of the HIV-1 capsid and have been used in many published pull-down binding studies. We incubated CA tubes with full-length recombinant PQBP1 (with PQBP1 at 6-fold excess over CA hexamers), and then assessed binding by visualizing the protein fraction associated with pelleted CA tubes after centrifugation. As shown in **Figure 2.10A** (top row), substantially greater amounts of PQBP1 pelleted with CA tubes than without. Densitometry quantification indicated that 17±6% (N=9) of input PQBP1 pelleted with the CA tubes (**Figure 2.10B**), which is the theoretical saturation point assuming 1 binding site per CA hexamer. However, this number does not take into account contributions from non-specific binding. To estimate non-specific binding, we used the Nt-deletion and DE-to-G PQBP1 mutants, which pelleted with CA tubes at lower levels than wild-type PQBP1 (**Figure 2.10A**, middle two rows; quantified in **Figure 2.10B**). Thus, assuming that wild-type PQBP1 has the same level of non-specific binding to the CA tubes as the mutants, our combined pull-down results indicate that the PQBP1 molecules occupy around half of the available sites in the CA tubes, which in our view is reasonably efficient binding in

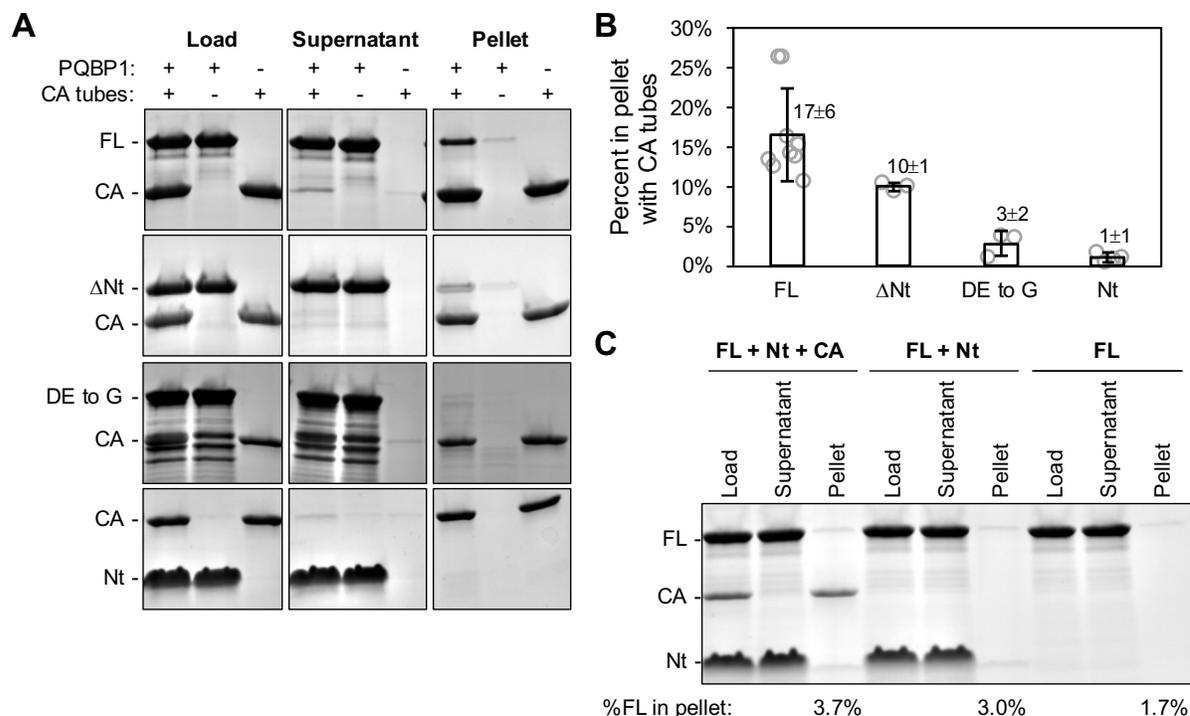


Figure 2.10. PQBP1 binding to the HIV-1 capsid lattice in vitro. (A) Indicated combinations of PQBP1 constructs (50 μ M) and disulfide-stabilized CA tubes (50 μ M in terms of CA subunits, which is equivalent to 8.3 μ M of hexamer) were incubated for 1 h at room temperature, then centrifuged. Load, Supernatant, and Pellet fractions were analyzed using SDS-PAGE and Coomassie staining. Representative of 9 (FL, full-length PQBP1), 3 (Δ Nt mutant), 3 (DE-to-G mutant), or 4 (Nt peptide) experiments. (B) Band intensities for pelleted PQBP1 in the presence of CA tubes after subtraction of matching negative controls in the absence of CA tubes, reported as mean \pm S.D. (C) CA tubes were incubated with both FL PQBP1 and Nt (50 μ M each), prior to centrifugation. Representative of 2 independent experiments.

vitro.

Interestingly, the Nt peptide did not bind CA tubes efficiently in this experiment, with only \sim 1% of input found in the pellet (Figure 2.10A, bottom row; quantified in Figure 2.10B). This result was unexpected, given Nt's apparent nM-level affinity for soluble CA hexamers in the nanoDSF experiments (Figure 2.3C). We surmised that Nt may dissociate from the CA tubes during centrifugation or that the central channel of the CA hexamer may be more resistant to peptide insertion in context of the assembled capsid lattice. In an attempt to distinguish between these possibilities, we performed a competition experiment in which equal amounts of Nt and full-length PQBP1 protein were

incubated with the CA tubes prior to centrifugation. Results show diminution of pellet-associated full-length PQBP1 to background levels (**Figure 2.10B**), which indicates that Nt is capable of binding the tubes. More important, these data also suggest that other regions of PQBP1, outside of Nt, are likely contributing to stable association of the full-length protein with the CA tubes. This result is reminiscent of TRIM5 α : the principal binding interaction between CA and the TRIM5 α capsid-binding module (the SPRY domain) is insufficient for stable binding to the assembled capsid lattice and requires contributions from other regions of the protein (reviewed by ref. 85).

2.4 Discussion

Our finding that Nt binds the central channel of the CA hexamer provides a formal biochemical proof that PQBP1 is a capsid-binding protein, because the hexamer only exists in context of an assembled capsid. Our results also indicate that binding involves charge complementation between acidic residues in PQBP1 and the basic Arg18 ring in the central channel of the HIV-1 CA hexamer. Other host cell factors, including nucleotide triphosphates, IP₆, and FEZ1,^{17,83,84,120} have been previously shown to bind to the Arg18 site via charge interactions. The addition of PQBP1 to the list highlights the importance of the Arg18 ring in HIV-1 capsid multifunctionality.

While our biochemical and structural data show that a tri-aspartate (DDD) motif in context of a small peptide represents the “minimal” PQBP1 binding element for the HIV-1 CA hexamer, the surrounding sequences are also clearly important. Specifically, a 46-mer peptide (Nt) binds with 10-fold weaker affinity to the CA hexamer than a 15-mer peptide. We surmise that this difference might be explained by the need to thread the peptide through the hexamer channel. Since the DDD motif is located in the middle of both Nt and

the 15-6 peptide, binding would be more inefficient with the former because more residues will need to be threaded, increasing the probability of dissociation before the DDD motif can make contact with the Arg18 ring. This model can explain why multiple acidic motifs are found in both PQBP1 and FEZ1, i.e., to maintain productive contacts. Furthermore, our cryoEM structure suggests that the bound portion of Nt might fold into a helix. Although pending confirmation, binding-induced folding of PQBP1 would contribute additional stabilization to the interaction. We envision that PQBP1 acts in analogy to a grapnel or grappling hook, and that the Nt peptide segment that we observe in our cryoEM map represents a secure anchor point for PQBP1. Our observation of a minor subpopulation of hexamers containing two peptide densities can be interpreted as a hairpin configuration of Nt in an intermediate stage of insertion into the hexamer channel.

In cells, small molecule ligands for the Arg18 ring are thought to be at considerable excess over protein factors (e.g., IP_6 is estimated at 20 μM or more, and nucleotide triphosphates are collectively at around 10 mM).^{125–127} These numbers indicate that productive protein binding to the HIV-1 capsid is not just determined by simple thermodynamic parameters. Indeed, many protein factors are now shown to stably bind to the HIV-1 capsid through avidity, and an important emerging concept is that the capsid acts as a multivalent recruitment platform for both pro-viral and anti-viral proteins.^{85,128–136} Along these lines, we find that the PQBP1 Nt peptide does not sustainably bind to HIV-1 CA tubes to withstand centrifugation, whereas full-length PQBP1 does. We propose, in line with other capsid-binding factors, that other regions of PQBP1 outside of Nt also promote stable binding through avidity. Further studies are now required to delineate these regions in PQBP1 and determine how they promote capsid recognition and eventual cGAS activation.

2.5 Methods

2.5.1 Protein production

Full-length PQBP1 protein was sub-cloned into pET16b (Novagen) with a C-terminal His-tag and Gly-Gly linker. Transformed BL21(DE3) *E. coli* cells were induced mid-log phase with 1 mM IPTG for 4.5 hours. Cells were pelleted, resuspended in buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 1 mM TCEP) supplemented with a protease inhibitor tablet (Roche), and then lysed by a combination of lysozyme treatment and sonication. The soluble fraction was filtered into a gravity column packed with Ni-NTA beads (Qiagen) and incubated for 30 minutes at 4°C. Following a wash of the beads, protein was eluted with resuspension buffer containing increasing concentrations of imidazole (50-300 mM). Appropriate fractions as determined by SDS-PAGE were pooled and dialyzed overnight at 4°C into 20 mM Tris, pH 8.0, 100 mM NaCl, 1mM TCEP. The sample was then loaded onto a HiTrap Q FF column (GE Healthcare) and eluted with a linear NaCl gradient. Protein fractions identified via SDS-PAGE were pooled for overnight dialysis at 4°C into the final protein buffer of 20 mM Tris, pH 8.0, 100 mM NaCl, and either 1 mM TCEP or 5 mM β -mercaptoethanol. Samples were concentrated to 2-7 mg/mL and flash-frozen in liquid nitrogen for storage at -80°C. The Nt-deletion (Δ Nt) and DE-to-G PQBP1 mutant proteins were prepared in the same way. We found that the DE-to-G mutant was somehow prone to proteolysis, and thus the preparation contained significant amounts of lower-MW fragments. We judged this mutant to be suitable for use in pull-down experiments because the CA band was still distinguishable in Coomassie-stained SDS-PAGE gels (**Figure 2.10A**).

Peptides corresponding to PQBP1 fragments (**Figure 2.9A**) were obtained from a commercial source (Celltein) at >95% purity (HPLC purification) and verified by mass

spectrometry. All peptides were N-terminally acetylated and C-terminally amidated, except for Nt and those starting with proline, which were modified only at the C-terminus. Freeze-dried peptides were reconstituted in water or PQBP1 buffer to generate 30-50 mM stock solutions prior to use.

Disulfide-crosslinked A14C/E45C/W184A/M185A CA hexamers with or without the R18A mutation, and disulfide-crosslinked A14C/E45C CA tubes were prepared as previously described.⁷⁷

2.5.2 Nanodifferential scanning fluorimetry

CA hexamers were briefly incubated (5 min) with PQBP1 constructs, loaded onto capillaries, and melting profiles (both raw curves and first derivative curves) were obtained using a Tycho NT.6 (NanoTemper), except for data shown in **Figure 2.3B**, which were obtained with a Prometheus (NanoTemper). Dose-response curves were obtained by serially diluting Nt peptide into 0.83 μM of hexamer (equivalent to 5 μM of CA subunits), or 15-mer peptide into 0.083 μM of hexamer. Each first derivative profile was fit as a sum of two Gaussian curves (corresponding to bound and unbound species) using Excel (Microsoft), as described.⁶⁵ Bound fractions were calculated from the area under each Gaussian curve. The dose-response data were then fit to a simple cooperative model, using the equation $y = y_{\text{max}} * [\text{Nt}]^n / (K_d^n + [\text{Nt}]^n)$, where K_d is the apparent dissociation constant, n is the apparent Hill coefficient, and y_{max} is the specific activity of CA.

2.5.3 Pull-down assays

CA tubes (50 μM in terms of CA subunits, which is equivalent to 8.3 μM in terms of CA hexamers) were incubated without or with 50 μM of indicated PQBP1 constructs for 1 h at room temperature. Pellet and supernatant fractions were separated in a microcentrifuge at maximum speed ($\sim 21,000$ g) for 30 min at room temperature. Fractions

were analyzed by using SDS-PAGE with Coomassie staining; pre-cast Mini-Protean 15-well 4-20% gradient gels (Biorad) and a standard protocol were used for uniform staining. Band quantification was performed with Fiji.¹³⁷

2.5.4 ISG54 induction assay

THP-1 cells stably expressing both wild-type and mutant PQBP1-eYFP fusion proteins (with silent mutations conferring resistance to siRNAs targeted against endogenous PQBP1) were generated as described previously.¹¹³ siRNAs were introduced into PMA-differentiated THP-1 cells using Stemfect RNA transfection kits. Typically, 5 pmol of siRNAs and 0.17 μ L of Stemfect were used per 2.5×10^4 cells. Forty-eight hours after siRNA transfection, cells were infected with VSV-G pseudotyped HIV-1 virus in the presence of VLP-VPX for 16 hours, followed by RT-qPCR analysis for ISG54 expression.^{111,113} The following siRNAs were used: siPQBP1 (5'-AAGCTCAGAAGCAGTAATGCA-3'), and non-targeting control (5'-AATCGATCATAGGACGAACGC-3'). qPCR primers were described previously.¹¹³ 5 ng of p24 or 2.5 RT units of HIV-1 were used per 25,000 cells in 100 μ L of media.

2.5.5 CryoEM structure determination

CA hexamers (50 μ M) were incubated with 3-fold excess of Nt peptide at room temperature for 1 h. Samples (4 μ L) were diluted, and then applied to glow-discharged 1.2/1.3 c-flat 300-mesh copper grids. The grids were briefly blotted manually, and then plunge-frozen in liquid ethane using a home-built device. CryoEM data were collected at the University of Virginia Molecular Electron Microscopy Core, using a Krios (ThermoFisher) operating at 300 kV and equipped with an energy filter and K3 direct detector (Gatan). Images were collected using EPU (ThermoFisher) at a pixel size of 1.08 \AA in counting mode, with a total dose of 50 electrons/ \AA^2 over 40 frames, and target

defocus of -0.5 to -2.5 μm . Two data sets were collected, one at 0° tilt (3,612 images) and a second at 30° tilt (936 images).

All image processing and map calculations were performed in cryoSPARC v.3.3.1–3.¹³⁸ Raw movies were corrected for beam-induced motion using MotionCor2.¹³⁹ CTF estimation was performed with CTFFIND4¹⁴⁰ for the 0° tilt data and Patch CTF for the 30° tilt data, both as implemented in cryoSPARC. Initially, the two data sets were processed independently (**Figure 2.5A**). After template-based picking, particles for ab initio calculation were selected through two or three rounds of reference-free 2D classification. Ab initio maps ($n=4$) were calculated using default parameters and C6 symmetry applied. One class from each data set was selected then refined in a single round, with C1 symmetry. At this point, the two particle sets (659,056 total) were combined by reconstructing a single map with C1 symmetry. After local refinement, the particle set was separated into bound and unbound hexamers, using a 3D Variability Analysis (3DVA) run in cryoSPARC (as described below). The Nt-bound hexamer (312,788 particles) was first reconstructed without further alignment, and then subjected to one round of local refinement in C1 symmetry, yielding a final consensus map at nominal resolution of 3.3 Å (0.143 cutoff) (**Figure 2.5B**).

The consensus map was further processed through another round of 3DVA, which defined alternative positions of the Nt peptide. Each new particle set was reconstructed without further alignment, resulting in two ligand-bound maps (154,820 and 157,280 particles), each at nominal resolution of 3.5 Å (at the 0.143 cut-off) (**Figure 2.7B,C**). Note that odd and even particle assignments that were set after ab initio calculations were kept throughout this analysis, thus maintaining gold-standard conditions. We also confirmed that similar proportions of particles from tilted and untilted images were maintained. In principle, 3DVA could have been performed with each new map to define additional

subsets, but ever-decreasing particle numbers became limiting.

2.5.6 3D variability analysis

Variability modes were calculated inside a mask that included the central part of the hexamer and excluded all CA density except for the β -hairpin and helix 1, with filter resolution set to 4 Å. Non-overlapping particle subsets were subsequently defined using a 3D Variability Display run in “intermediates” output mode, with number of frames set at 2 and width at 0.

2.5.7 Coordinate modeling

The consensus map was refined for one additional round with C6 symmetry applied. PDB 3H47,⁷⁷ which consists of a single chain of CA, was docked into this map, then refined in real space using Phenix,¹⁴¹ with a single round of morphing and 3H47 as reference. The other five CA subunits were then fit by copying the first chain, and then docked into the C1 consensus map. After manual inspection and minor rebuilding in Coot,¹⁴² real-space refinement was performed with non-crystallographic symmetry restraints. The Nt density was left unmodeled.

2.5.8 Data deposition

The consensus map has been deposited at the EMDB (EMD-41711) and the coordinate model of the CA hexamer has been deposited at the PDB (8TY6).

2.6 Acknowledgements

The University of Virginia Molecular Electron Microscopy Core facility (RRID: SCR_019031) was built in part with NIH grant G20-RR31199. Purchase of the Krios was funded in part by S10-RR025067. We especially thank K. Dryden and M. Purdy for

assistance in cryoEM data collection, and M. Purdy for help with optimizing server and software set up. This study was funded by NIH U54-AI170856 (O.P. and B.K.G.-P.), and by NIH R01-AI127302, a CFAR Development Grant, and CHRP Basic Bio Pilot Grant BB19-SBMR-0 (S.M.Y and S.K.C.). J.P. was supported by NIH T32-GM080186.

Chapter 3: PQBP1 Ct is required to induce HIV-1 capsid lattice destabilization

3.1 Introduction

As described in the previous chapter, PQBP1 Nt binds to the electropositive Arg18 pore of HIV-1 CA hexamer found within the capsid lattice.^{114,143} Although this interaction represents the minimal requirements for binding, the complete interaction network between full-length PQBP1 and the HIV-1 capsid remains undefined, as do the consequences of binding. This interaction network is likely determined by the dynamic structural nature of PQBP1 since it is an IDP. Though many biophysical characteristics of this protein continue to be elusive, PQBP1 is thought to have the potential to self-associate, including by undergoing dimerization at its C-terminus.¹¹⁵ This highly conserved, semi-ordered region has been shown to be important for both endogenous functions of PQBP1 and the stimulation of cGAS signaling upon HIV-1 infection.^{113,117} However, the details of how PQBP1 self-association may affect the HIV-1 capsid after binding in order to bridge capsid recognition and cGAS recruitment & downstream signaling are still unknown.

This chapter describes results from an ongoing project focused on elucidating the biochemical consequences of PQBP1 binding to the HIV-1 capsid and how PQBP1 self-association contributes to these effects. Cryo-electron tomograms of cross-linked CA tubes incubated with PQBP1 show tube decoration and confirm a direct binding interaction. Pull-down assays performed after extended incubation periods demonstrate an increase in soluble CA signal, but only in the presence of full-length (FL) PQBP1. CA solubilization indicates capsid lattice disruption induced by PQBP1 binding and is consistent with negative stain electron microscopy (EM) images showing altered CA tube

morphology. Mass photometry data reveal the propensity for PQBP1 to self-associate, in particular by dimerizing at its Ct. Finally, experiments performed with a PQBP1 construct in which Ct was replaced with a leucine zipper recapitulate self-dimerization and cGAS signaling stimulation observed with PQBP1 FL. Taken together, these data suggest that Ct dimerization contributes to PQBP1 induced destabilization of the HIV-1 capsid lattice, thus facilitating cGAS signaling during HIV-1 infection.

3.2 Results

3.2.1 Cryo-tomography analysis of PQBP1 binding to the HIV-1 capsid lattice

To examine the structure of FL PQBP1 bound to the extended HIV-1 capsid lattice, pelleted samples of A45C/E45C cross-linked CA tubes were incubated for an hour in the presence of PQBP1 FL and then prepared for cryo-electron tomography (cryo-ET). Twenty tomograms were collected and await further processing for structural determination. Although additional control experiments should be performed, preliminary images from these data show potential decoration of cross-linked CA tubes by PQBP1 (**Figure 3.1**).

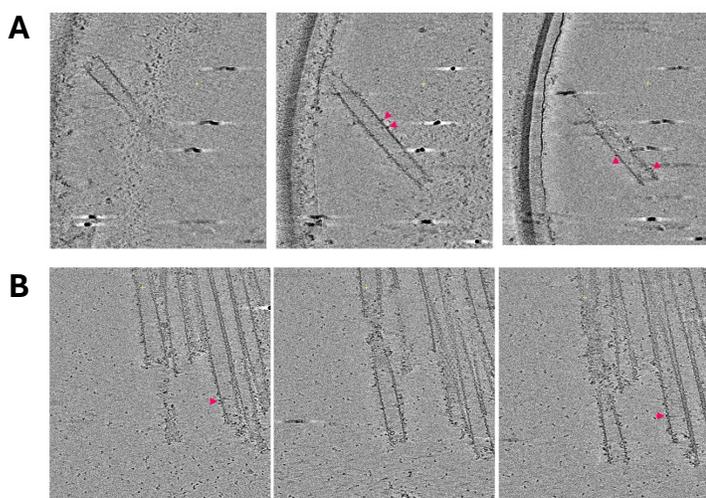


Figure 3.1. PQBP1 FL decorates cross-linked CA tubes. **A** and **B** are two representative tomograms representative. Left-to-right images show the same CA tube(s) in the presence of equimolar PQBP1 FL across different field views along the z-axis. Possible PQBP1 decoration is indicated by the pink arrows.

3.2.2 PQBP1 binding induces capsid lattice destabilization

In order to assess the consequences of PQBP1 binding, pull-down experiments were performed after both one hour and overnight incubation between PQBP1 FL and cross-linked CA tubes. As previously shown, PQBP1 FL presence in the pellet was observed after one hour, indicating that the protein is binding to and being pulled down by the cross-linked CA tubes.¹⁴³ However, there was also a modest increase in soluble CA signal when PQBP1 was present, indicating that destabilization of the capsid lattice may occur upon PQBP1 binding. Indeed, after overnight incubation, minimal PQBP1 FL was observed in the pellet, and there was a substantial increase of soluble CA that corresponded to a decrease of pelleted CA (**Figure 3.2A**). The latter observation is particularly striking; cross-linked CA tubes normally pellet completely as they have almost 100% assembly efficiency and are extremely stable once assembled (**Figure 3.2D**).¹⁰⁷ Since the tubes are stabilized by engineered disulfide bonds, control experiments were performed to test the effect of reducing agent. Though there is ~2.2 mM β me in the final pulldown sample buffer, the tubes have demonstrated resistance to β me concentrations of up to 4 mM (**Figure 3.2D**), indicating the reducing agent is not causing tube destabilization. As an additional control, PQBP1 binding and destabilization effects were abolished by addition of 40 μ M IP6 (**Figure 3.2E**), which competitively binds at the same site on CA hexamers as PQBP1.⁸³ These results suggest that FL PQBP1 binding leads to capsid lattice destabilization.

Pull-downs were repeated with PQBP1 constructs containing a C-terminal deletion (Δ Ct; residues 1-211) or N-terminal deletion (Δ Nt; residues 47-265). PQBP1 Δ Ct pelleted with the cross-linked tubes after an hour, but there was no increase in soluble CA signal after overnight incubation (**Figure 3.2B**), indicating that Ct is required to observe the destabilization phenotype. Pull-downs performed with PQBP1 Δ Nt demonstrated non-

specific binding (as determined previously¹⁴³) and no increase in soluble CA signal after overnight incubation (**Figure 3.2C**).

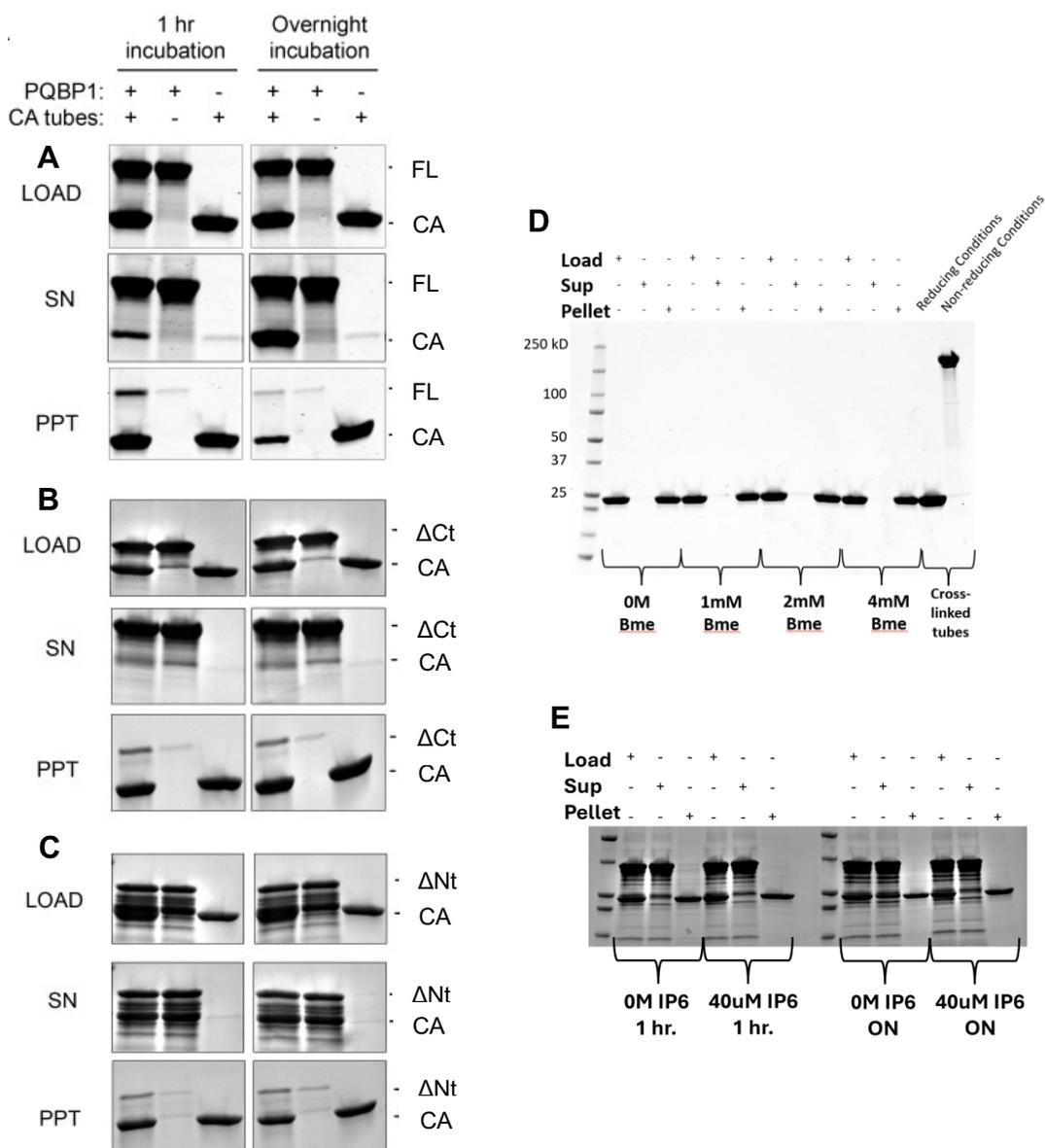


Figure 3.2. PQBP1 binding results in capsid lattice destabilization. (A-C) Pull-down results between PQBP1 FL (**A**), Δ Ct (**B**), or Δ Nt (**C**) and cross-linked CA tubes after indicated incubation times, as analyzed by SDS-PAGE and uniform Coomassie staining. (**D**) Pull-down control results of cross-linked CA tubes incubated overnight (ON) in the presence of β me. Assembly confirmation of the cross-linked CA tubes is determined by comparing protein sizes in reducing (high [β me], sample boiled) vs. non-reducing (no β me, sample not boiled) conditions as seen in the right-most lanes, reproducibly demonstrating the nearly perfect efficiency of assembly.¹⁰⁷ (**E**) Pull-down results of PQBP1 FL and CA tubes samples with 40uM IP6 added prior to incubation for one hour (left) or ON (right). All presented data are representative of at least three independent experiments for each set of conditions.

Pelleted pull-down samples of CA tubes after one hour incubation with or without PQBP1 FL and/or 40 μ M IP6 were visualized using negative stain EM. Images show that while CA tubes alone are mostly straight and elongated (**Figure 3.3A**), CA tubes in the presence of PQBP1 appear distorted, e.g. wavy, pinched, wobbly (**Figure 3.3B**). In general, CA tubes incubated with PQBP1 were less numerous and appeared more fragile than pelleted CA tubes alone. However, the addition of IP6 restored CA tube fidelity, even in the presence of PQBP1 (**Figure 3.3C**). These results are in agreement with the observations made from pull-down data and visually confirm that PQBP1 FL can affect capsid lattice stability when bound.

Taken together, these data suggest that PQBP1 induces HIV-1 capsid lattice destabilization as a consequence of binding, but only when PQBP1 retains its C-terminus.

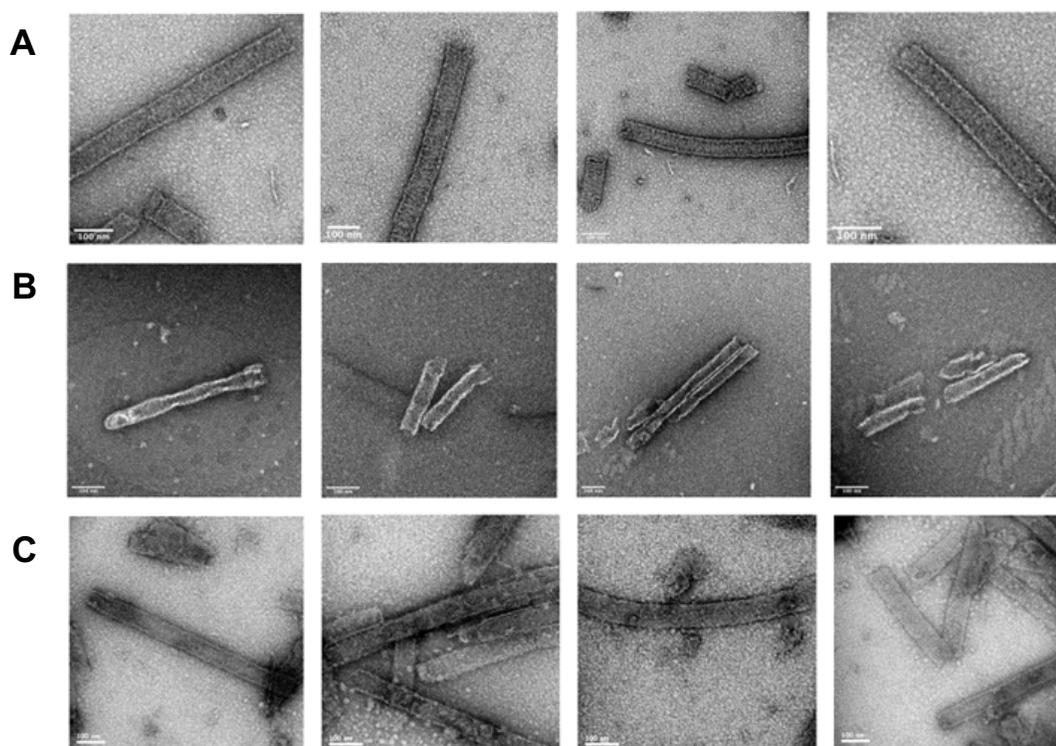


Figure 3.3. Binding of PQBP1 affects CA tube morphology. Negative stain EM images of pelleted cross-linked tubes CA alone (**A**) or in the presence of PQBP1 FL without (**B**) or with (**C**) the addition of 40 μ M IP6 after one hour of incubation. Images are representative of at least two independent experiments for each condition, and all scale bars show 100nm.

3.2.3 PQBP1 self-association is mediated by Nt and Ct

The result that Ct is necessary to observe capsid lattice destabilization implicates Ct in destabilizing the HIV-1 capsid, though its role in this process was unclear. Previous data have shown that while Nt is crucial for PQBP1 binding to the capsid, it is not the only region of the protein contributing to these interactions,¹⁴³ thus Ct may also be involved in PQBP1 binding. Additionally, incubation of Nt with PQBP1 FL in the absence of CA tubes causes the full PQBP1 protein to pellet, indicating the formation of higher-order assemblies.¹⁴³ This observation, as well as the prediction of Ct to be a potential site of self-interaction,¹¹⁵ suggest that self-association of PQBP1 is critical to understanding the behavior of PQBP1, especially in the context of HIV-1 capsid binding.

To investigate PQBP1 self-association capabilities, additional pull-downs were performed on 50 μ M FL PQBP1 samples mixed with various amounts of Nt in the absence of CA tubes. Addition of 20-fold excess Nt caused sample turbidity indicative of higher-order self-association. Analysis of the supernatant and pellet fractions after centrifugation revealed that the pellet consisted primarily of full-length PQBP1 (**Figure 3.4A**). Interestingly, the pellet contained only trace amounts of Nt (last lane in **Figure 3.4A**). Interpreted through the standard nucleation-propagation model of higher-order assembly, the simplest explanation of this result is that Nt binding to full-length PQBP1 nucleates the process, which is then propagated by additional full-length PQBP1 molecules.

To examine this further, we used mass photometry (MP) to analyze self-association of PQBP1 at lower protein concentrations. Full-length PQBP1 was largely monomeric at 30 and 50 nM, with only small traces of apparent dimer and trimer species, if at all (**Figure 3.4B**). Addition of 30 nM or 1 μ M Nt to 30 nM of full-length PQBP1 induced greater amounts of PQBP1 dimer in a concentration-dependent manner (**Figure 3.4C**), confirming our findings with the pull-down experiments that Nt induces PQBP1 self-association.

Since PQBP1 Ct is the proposed self-association site, experiments focused on this region of the protein were also performed. Analysis of Δ Nt MP data showed this construct as primarily dimeric with minimal monomer species at both 30 and 100 nM (**Figure 3.4D**). Furthermore, Ct on its own also exists as a dimer, with the caveat that the monomer peak is below the limit of detection for this method (**Figure 3.4E**). These collective results suggest that PQBP1 dimerizes at its C-terminus.

Overall, these results indicate that PQBP1 harbors a self-association property – likely Ct dimerization – that is hidden but is unmasked by deletion of Nt or by addition of Nt peptide in trans.

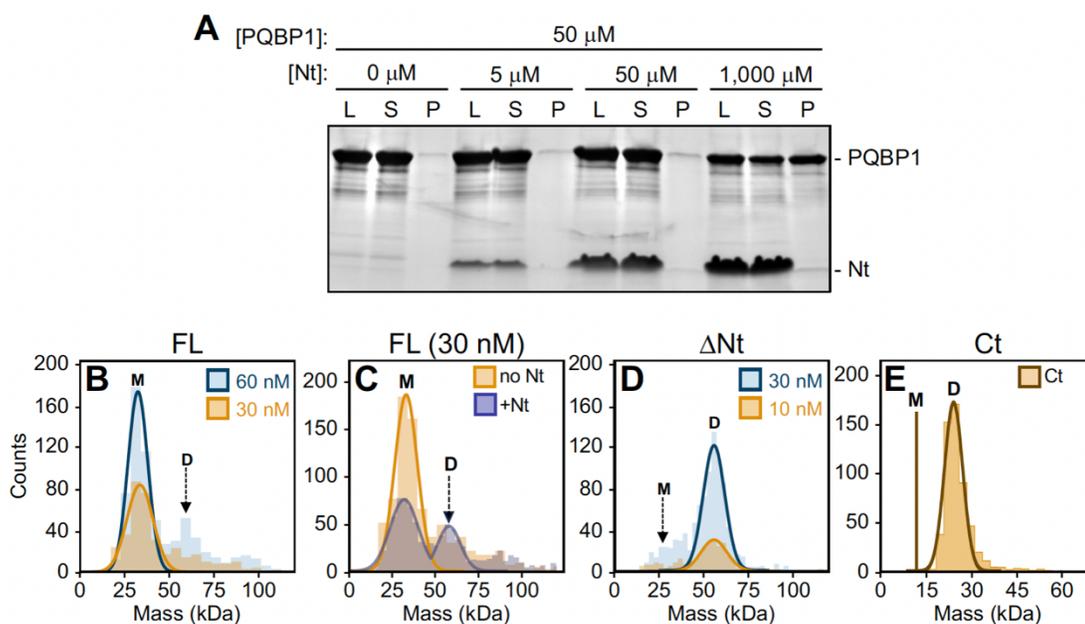


Figure 3.4. PQBP1 Ct self-association is unmasked by Nt. (A) PQBP1 FL (50 μ M) was incubated with indicated amounts of Nt peptide for 1 h at room temperature, then centrifuged. Load (L), Supernatant (S), and Pellet (P) fractions were analyzed using SDS-PAGE and Coomassie staining. Representative of three independent experiments. (B-E) Mass photometry profiles of PQBP1 FL (B), FL incubated with or without Nt peptide (C), Δ Nt (D), and Ct (E). Bars are histogram counts, and curves are Gaussian fits. M, monomer; D, dimer. Profiles are representative of at least 2 independent measurements. Data kindly provided by D. Allen and S. Yoh.

3.2.4 PQBP1 Ct dimerization is required for cGAS signaling

In order to determine the role of Ct dimerization within the context of PQBP1's function in bridging capsid recognition and cGAS signaling, a plasmid construct was designed and commercially synthesized such that PQBP1 Ct was replaced by the leucine zipper (LZ) sequence from Cre binding protein (CREB), which is known to form a stable homo-dimer (**Figure 3.5A**).¹⁴⁴ The construct was purified using a protocol adapted from previous methods (**Figure 3.5B-D**),¹⁴³ and MP experiments were performed with 30 nM quantities of both PQBP1-LZ and PQBP1 Δ Ct. The profile of Δ Ct looks similar to that of FL in that it is primarily monomeric with little to no multimers (**Figure 3.5E**; teal). Importantly, PQBP1 LZ resembles the MP profile of Δ Nt as it is mainly dimeric with trace amounts of both monomer and other higher-order species (**Figure 3.5E**; blue), which validates dimerization of the leucine zipper region of the protein. Interestingly, PQBP1 LZ forms a dimer even without the addition of Nt peptide or the deletion of Nt, indicating that the LZ substitution promotes a more stable dimeric state, as intended.

To test the functional role of PQBP1 Ct dimerization in cGAS signaling, a proximal ligation assay (PLA) between FLAG-tagged cGAS and CA was performed (as previously described¹¹⁴) in HIV-1 infected THP-1 cells stably expressing PQBP1 FL, Δ Ct, and LZ. Colocalization of cGAS and CA was observed in the presence of PQBP1 FL but not in the absence of PQBP1 even during infection, recapitulating previous results that PQBP1 is required to induce cGAS signaling upon HIV-1 infection.¹¹³ However, there was no significant cGAS-CA colocalization with PQBP1 Δ Ct during infection, indicating the importance of Ct for downstream recruitment of cGAS. PQBP1 LZ rescued cGAS signaling activity as evidenced by the restored colocalization of cGAS and CA, indicating that self-association of PQBP1 at its C-terminus is necessary for the stimulation of cGAS. These data, paired with the MP results, suggest that PQBP1 can dimerize at its C-terminus and

that this dimerization is crucial to the induction of cGAS signaling.

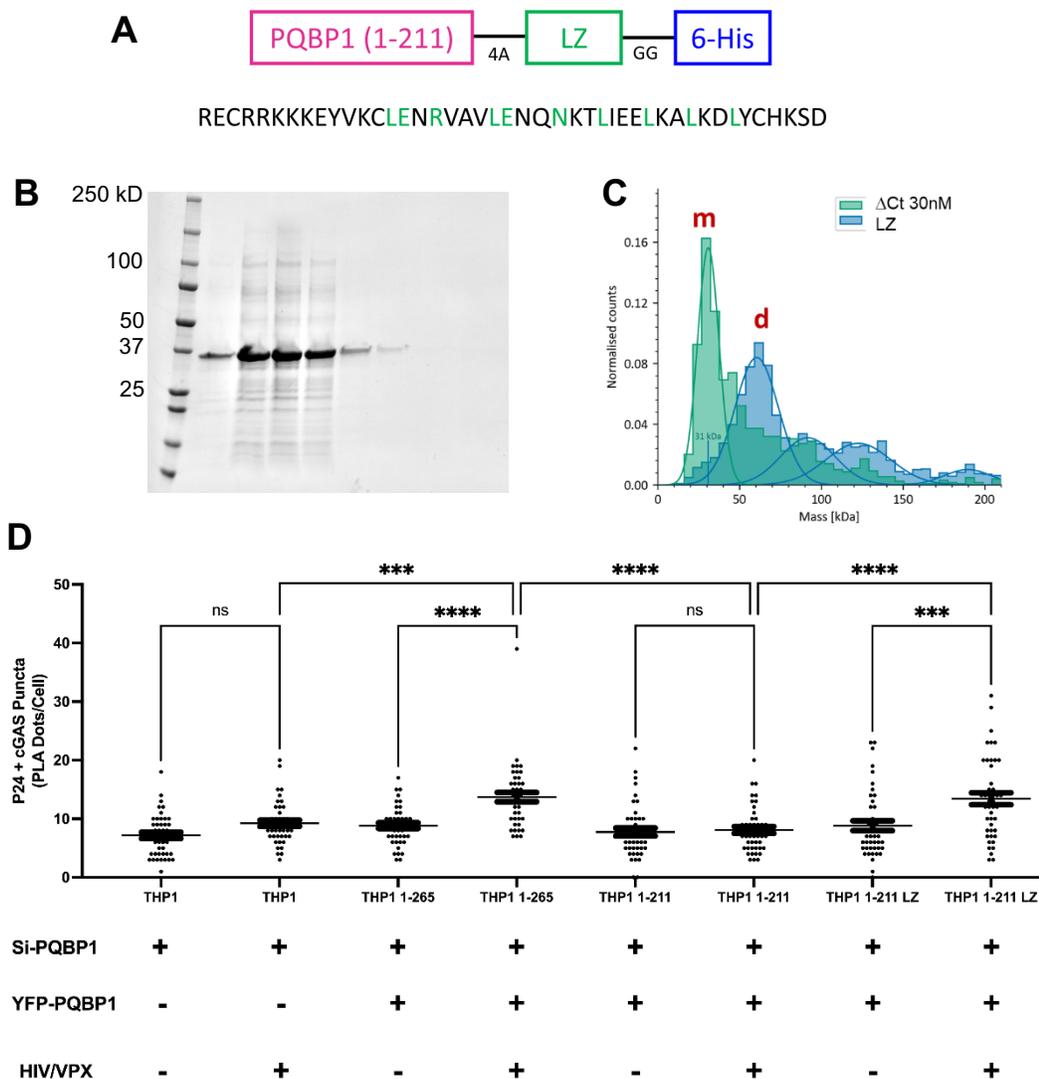


Figure 3.5. C-terminal PQBP1 dimerization facilitates HIV-1-induced cGAS signaling. (A) Schematic of the synthesized PQBP1 LZ construct within a pET30a vector (top) and the amino acid sequence of the leucine zipper (bottom). “4A” corresponds to a four-alanine linker, “GG” corresponds to a glycine-glycine linker, and “6-His” corresponds to a 6-histidine tag. Leucine zipper residues that contribute to the homodimeric interaction are colored in green.¹⁴⁴ (B) SDS-PAGE gel of fractions containing PQBP1 LZ (~31kD) after ion exchange chromatography demonstrate the successful characterization of this construct. (C) Mass photometry profiles of 30nM PQBP1 ΔCt (teal) or LZ (blue). Bars are histogram counts, and curves are Gaussian fits. M, monomer; D, dimer. Profiles are representative of at least 2 independent measurements. (D) PLA dots per THP-1 cell infected with HIV-1 in the presence of VLP-Vpx and the indicated PQBP1 constructs. Cells were paraformaldehyde-fixed for analysis 2 hours post-infection with N=45 cells per group, and dots were quantified using ImageJ. One-way ANOVA, ****p<, ***p<, and “ns” is no significance. Panels C and D kindly provided by D. Allen and S. Yoh.

3.3 Discussion

The data presented in this chapter confirm previous results that full-length PQBP1 protein can bind to the extended HIV-1 capsid lattice beyond just the minimal requirements of PQBP1 Nt and the R18 pore of CA hexamers. Interestingly, our results also demonstrate that prolonged incubations of CA tubes in the presence of PQBP1 leads to lattice destabilization as a consequence of PQBP1 binding. Our negative stain EM images reinforce the notion that PQBP1 itself is causing lattice disruption upon binding since CA tubes become warped after PQBP1 incubation, but only when IP6 is not present to outcompete PQBP1 as a capsid binding partner. Furthermore, the destabilization phenotype is dependent upon the Ct of PQBP1.

Our results also reveal that PQBP1 has a propensity for self-association that is unmasked by manipulating Nt. At lower (nM) PQBP1 concentrations, we find that excess Nt peptide in trans or Nt deletion promotes PQBP1 dimerization. Additionally, at higher (μ M) concentrations, we find that Nt induces formation of higher-order (pelletable) species of PQBP1. Since PQBP1 Δ Nt, FL (in certain conditions), and likely Ct itself exist as dimers in solution while Δ Ct is primarily monomeric, the simplest explanation is that PQBP1 dimerization occurs at the C-terminus. This is in agreement with the predicted structural properties of Ct as it possesses a semi-ordered helical region with the potential to self-associate.¹¹⁵ Based on these data, a simple model for PQBP1 self-association in vitro is: 1) In context of full-length PQBP1, Nt inhibits self-association, perhaps by binding to or otherwise sequestering Ct in cis. 2) Deletion of Nt or addition of Nt peptide in trans releases or exposes the Ct, promoting PQBP1 self-dimerization. 3) Self-association at low protein concentrations is limited to dimerization, with higher-order species forming at higher protein concentrations. An important corollary of this model is that the previously-described insertion of Nt into the CA hexamer channel could be expected to have the

same effect as Nt deletion and promote PQBP1 self-association. Combined with CA multivalency, the net effect would be to drive self-association of PQBP1 molecules on the capsid surface. The formation of these higher-order PQBP1 assemblies may provide mechanical stress sufficient to break apart the capsid lattice, resulting in capsid destabilization.

Additionally, by using an engineered PQBP1 construct that is confirmed to undergo spontaneous dimerization via a C-terminal leucine zipper, we showed that cGAS signaling upon HIV-1 infection requires dimerization of PQBP1. These results suggest that the self-association of PQBP1, not just binding of the protein itself, facilitates cGAS signaling during infection. Formation of higher-order PQBP1 assemblies may increase local cGAS concentrations required to reach the signaling threshold.¹⁰⁸ Furthermore, since capsid lattice destabilization is a consequence of PQBP1 self-association, cGAS recruited to capsid-bound PQBP1 would have access to the viral dsDNA that becomes exposed to the cytoplasm to stimulate downstream signaling. These results support a model that PQBP1 does not just bridge capsid recognition and cGAS recruitment but also serves as a signalosome, or a higher-order signaling platform as seen across biology.¹⁴⁵

To complete this project, there are several experiments that will be performed. Cryo-electron tomograms of PQBP1 FL binding to the extended capsid lattice will be processed for structural determination, and additional data will be collected as needed. For quality control purposes, pull-downs between CA tubes and PQBP1 in the presence and absence of IP6 will be repeated, and negative stain EM images of CA tubes in the conditions seen in **Figure 3.3** will be quantified. To biochemically confirm that PQBP1-induced destabilization of CA tubes depends upon Ct dimerization, pull-down assays will be performed as described with the PQBP1 LZ construct.

As a general note, exploring a different model system other than cross-linked or

IP6-stabilized tubes may be necessary for these and other experiments. Even though cross-linked tubes are useful in some circumstances, their stability can translate to rigidity that does not always recapitulate the flexibility of the capsid lattice and may conceal behaviors that would otherwise be observed. An alternative to assembling tubes through cross-linking is assembling them either in high salt conditions or with the addition of IP6, but such assemblies cannot be faithfully applied to these studies since both high salt and IP6 outcompete PQBP1 for binding to the capsid lattice.^{17,50} While other extended capsid assembly intermediates have been characterized in the literature,¹⁴⁶ a promising capsid reagent for future experiments is purified viral cores.⁴⁷

Regardless, the data collected thus far preliminary give rise to a possible explanation for how PQBP1 is able to bridge capsid recognition and cGAS signaling during HIV-1 infection: upon binding to the HIV-1 capsid, PQBP1 undergoes Ct dimerization and additional formation of higher-order assemblies, ultimately disrupting the capsid lattice while simultaneously serving as a cGAS recruitment platform to facilitate downstream signaling. Further studies are now required to more precisely delineate how PQBP1 Ct dimerization leads to higher-order self-association and how these properties promote both capsid recognition and eventual cGAS activation.

3.4 Methods

3.4.1 Protein production

For preparation of PQBP1-LZ, a synthesized plasmid containing a gene encoding the first 211 residues of PQBP1, a four-alanine linker, the CREB leucine zipper sequence,¹⁴⁴ a Gly-Gly linker, and a C-terminal 6-His tag within a pET30a vector was obtained (GenScript). Sequence confirmation and protein expression levels of this

construct upon 1mM IPTG induction of BL21 (DE3) *E. coli* cells at mid-log phase were confirmed before performing large-scale overexpression for 4.5 hours. Cells were pelleted, resuspended in buffer (50mM Tris-pH 8.0, 500mM NaCl, 1mM TCEP) supplemented with a protease inhibitor tablet (Roche), and then lysed by a combination of lysozyme treatment and sonication. The soluble fraction was filtered into a gravity column packed with Ni-NTA beads (Qiagen) and incubated for 30 min. at 4°C. Following a wash of the beads, protein was eluted with resuspension buffer containing increasing concentrations of imidazole (50-300mM). Appropriate fractions as determined by SDS-PAGE (**Figure 3.5B**) were pooled and dialyzed overnight at 4°C into 20mM Tris-pH 8.0, 500mM NaCl, 1mM TCEP. The sample was then run through a HiTrap Q FF column (GE Healthcare). Protein fractions identified via SDS-PAGE (**Figure 3.5C**) were pooled for overnight dialysis into 20mM Tris-pH 8.0, 500mM NaCl, 5mM β me and concentrated before injection onto a Superdex 75 Increase 10/300 size exclusion column (Cytiva). Protein-containing fractions identified by SDS-PAGE (**Figure 3.5D**) were pooled and concentrated to ~1mg/mL before flash-freezing in liquid nitrogen for storage at -80°C. Prior to this protein being used for experiments, it was thawed and sequentially dialyzed for 9 h. first into 20 mM Tris-pH 8.0, 250mM NaCl, 5mM β me and then into the final buffer of 20 mM Tris-pH 8.0, 100mM NaCl, 5mM β me. Since we found that this protein was prone to precipitation when lowering the salt concentration from 500mM, it was spun at 20,000xg for 10 min. before being flash-frozen for storage or used for experiments. The above protocol was adapted from previous methods, which is how all other PQBP1 proteins were prepared.¹⁴³ Disulfide-crosslinked A14C/E45C CA tubes were prepared as previously described.⁷⁷

3.4.2 Electron microscopy

To prepare samples, CA tubes (50 μ M in terms of CA subunits, which is equivalent to 8.3 μ M in terms of CA hexamers) were incubated without or with 50 μ M of indicated

PQBP1 constructs at room temperature for an hour. Pellet and supernatant fractions were separated in a microcentrifuge at maximum speed (~21,000xg) for 30 min. at 4°C, and the pellet was resuspended in 20mM Tris-pH 8.0, 100mM NaCl.

For negative stain EM, samples were applied to carbon film grids and stained with 2% uranyl acetate. Data were collected using a Spirit (Tecnai) operating at 300kV.

For cryo-ET, samples were diluted and applied to glow-discharged 1.2/1.3c-flat 300-mesh copper grids. The grids were briefly blotted manually and then plunge-frozen in liquid ethane using a home-built device. CryoET data were collected at the University of Virginia Molecular Electron Microscopy Core using a Krios (ThermoFisher) operating at 300kV and equipped with an energy filter and K3 direct detector (Gatan).

3.4.3 Pull-down assays

CA tubes were incubated with or without equimolar quantities of indicated PQBP1 constructs, either at room temperature for an hour or overnight at 4°C. Pellet and supernatant fractions were separated via centrifugation and were analyzed by using SDS-PAGE with Coomassie staining. Pre-cast Mini-Protean 15-well 4-20% gradient gels (Biorad) and a standard protocol were used for uniform staining. This protocol was adapted from a previously used method.¹⁴³

3.4.4 Mass photometry

Mass Photometry measurements were determined using the Refeyn Two^{MP} system (Refeyn Ltd, Oxford, UK). Recombinant PQBP1 proteins were diluted in binding buffer (20mM HEPES, 100mM KCL, 2.5mM MgCl₂, pH 8) at indicated concentrations (10-200nM) at room temperature. High Precision No. 1.5H glass coverslips were cleaned and 20 ul of the diluted sample were loaded onto silicone gaskets chambers (CultureWellTM reusable gasket, Grace Bio-Labs). Samples were read for 60

seconds in regular view using Acquire^{MP} software (Refeyn Ltd, Oxford, UK). Binding buffer was used to focus and determine background noise. Data was analyzed using Discover^{MP} software (Refeyn Ltd, Oxford, UK) and linear mass calibration standards were generated from BSA or Gel Filtration Standard (Bio-Rad). Representative histograms were fitted with Gaussian Curves. Data are representative of at least two independent experiments.

3.4.5 Proximal ligation assay

THP-1 cells stably expressing both wild-type and mutant PQBP1-eYFP fusion proteins, as well as possessing silent mutations conferring resistance to siRNAs targeted against endogenous PQBP1, were generated as described previously.^{113,114} These cells were fixed after 2 hours post-infection in 4% PFA, in PBS for 20 min, permeabilized with 0.2% Triton X-100 in PBS, and subjected to the PLA assay according to the manufacturer's protocol for the Duolink in situ detection reagents red kit (DUO92008, Sigma-Aldrich). Antibody incubations were used at 1;400 dilution: anti-rabbit HIV1 p24 (Abcam-ab32352) and mouse anti-FLAG® M2 antibody (Sigma). PLA dots were detected using Nikon A1R HD confocal with 60X objective. Analysis was done using ImageJ. This protocol was adapted from previous methods.¹¹⁴

3.5 Acknowledgements

We would like to thank the Chanda laboratory at The Scripps Institute, especially D. Allen and S. Yoh, for their contributions to this project (**Figures 3.4B-E, 3.5C&D**; Mass Photometry methods). We also thank the University of Virginia Molecular Electron Microscopy Core for their assistance in sample setup and EM data collection.

Chapter 4: The hexamer-like cross-linked CA pentamer binds ultra-potent capsid inhibitor GS-CA1

4.1 Introduction

The HIV-1 capsid is a fullerene cone, made of variable numbers of viral CA protein hexamers and twelve CA pentamers.⁷⁴ The pentamers occupy sites of high curvature, called declinations, exactly 12 of which are required to form a completely closed capsid shell. The HIV-1 CA protein does not normally form discrete hexamers and pentamers in solution, but we have previously developed a disulfide crosslinking strategy to generate these reagents for biochemical studies and structure analysis through X-ray crystallography.^{76,77,107} We found two sets of disulfides – N21C/A22C and P17C/T19C – which, combined with mutations (W184A and M185A) that prevent stable assembly of capsid-like particles (CLPs), allowed purification and crystallization of discrete, soluble HIV-1 CA pentamers.^{76,107} The disulfide-stabilized CA pentamers are distinct from the pentamers found in actual HIV-1 capsids.¹⁴⁷ Nevertheless, the crosslinked pentamers can bind to small molecule inhibitors that target the HIV-1 capsid, specifically the compounds PF-3450074 (PF74) and GS-CA1.^{30,148} The crosslinked pentamers without the W184A/M185A mutations assemble into 30-nm spherical CLPs that are hyperstable *in vitro*,¹⁰⁷ and thus have attracted interest as nanocages with potential utility in medicine and materials science.^{149,150} Particularly relevant here, the N21C/A22C spheres seem especially amenable for use in high-throughput screening for capsid-binding small molecules.¹⁵⁰

GS-CA1 is the parental compound of the newest anti-HIV drug,^{148,151} which was approved for clinical use in the United States in 2022 with the generic name lenacapavir (LEN). The natural target of GS-CA1 is the CA hexamer within the assembled HIV-1

capsid. X-ray crystal structures of LEN in complex with the CA hexamer indicate that the compound binds to the so-called NTD-CTD interface, which consists of the N-terminal domain (NTD) and C-terminal domain (CTD) of adjacent CA subunits.^{151,152} GS-CA1 makes close contacts with both the NTD and CTD, and binds with very high affinity ($K_d = 380$ pM) to the hexamer.¹⁴⁸ We have shown before that the NTD-CTD interface of disulfide-stabilized HIV-1 CA pentamers is very similar to that of the hexamer.⁷⁶ Indeed, GS-CA1 also binds to the N21C/A22C pentamer with very high affinity ($K_d = 270$ pM).¹⁴⁸ In this study, we determined the X-ray crystal structure of the N21C/A22C pentamer in complex with GS-CA1, and confirm that drug binding occurs in the same manner as the CA hexamer. Our work provides further support for the idea that crosslinked N21C/A22C assemblies are useful biochemical surrogates for identifying and characterizing capsid-binding ligands.

4.2 Results

4.2.1 GS-CA1 binds to the cross-linked HIV-1 CA pentamer

We first tested for binding of GS-CA1 to the N21C/A22C-crosslinked CA pentamer, by using nanodifferential scanning fluorimetry (nanoDSF). In this assay, tryptophan fluorescence was used to monitor protein thermal melting in the presence and absence of equimolar amounts of drug (5 μ M protein and 5 μ M GS-CA1, which represents a 5-fold excess of drug over pentamer). Disulfide-crosslinked CA hexamer⁷⁷ and unassembled wild type (WT) CA monomer were used as comparative controls. As expected from previous studies,¹⁴⁸ GS-CA1 significantly thermostabilized, by around 12 °C, both the cross-linked pentamer and hexamer, but not the unassembled CA protein (**Figure 4.1A**). First derivative melting curves showed that all the pentamers in the sample bound to the drug

(**Figure 4.1B**, compare blue and black curves). Interestingly, the bound species was polydisperse. This result could indicate lack of saturation of the five available binding sites per pentamer, or that denaturation of the disulfide-stabilized pentamer is a multi-step process. In either case, the nanoDSF experiments confirmed that the cross-linked CA pentamer binds GS-CA1 at sufficient stoichiometry, and thus samples suitable for crystal structure determination could be produced.

4.2.2 Structure of the disulfide-stabilized CA pentamer in complex with GS-CA1

The pentamer-drug complex crystallized in space group P21212, with two pentamers in the asymmetric unit (**Figure 4.1C**, **Table 4.1**). The previously described unbound pentamer was in P21, with unrelated crystal packing.⁷⁶ The quality of the electron density map varied for different parts of the protein-drug complexes. Densities for all ten NTDs were generally of high quality (e.g., **Figure 4.1D**), but four CTDs were of

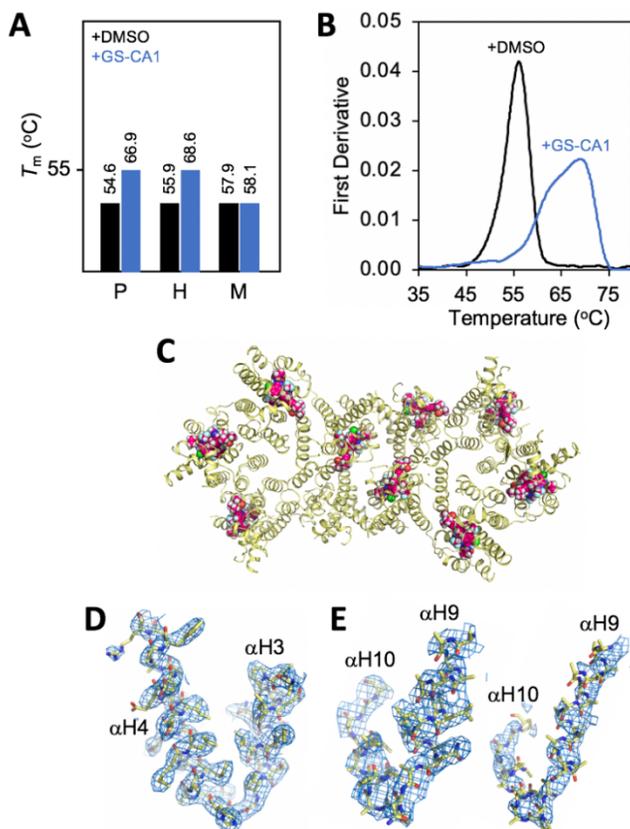


Figure 4.1. Analysis of GS-CA1 binding to cross-linked CA pentamer. **(A)** Apparent melting temperatures (T_m) of the indicated CA constructs in the presence of GS-CA1 (blue) or DMSO control (black). Error bars represent the standard deviation of triplicate measurements. P, crosslinked pentamer; H, crosslinked hexamer; M, unassembled CA monomer. **(B)** Melting profiles of crosslinked pentamer, shown as first derivative curves. Representative of three independent measurements. **(C)** View of two pentamers in the P21212 asymmetric unit. CA protein is rendered as ribbons (yellow). Nine bound drug molecules are rendered as spheres (pink). **(D)** Electron density for the NTD (blue mesh, $2mF_o-DF_c$, 1.2σ), illustrated for helices 3 and 4 in chain F, which is representative of all 10 protein chains. **(E)** Variable electron density for the CTD (blue mesh, $2mF_o-DF_c$, 1σ), illustrated for helices 9 and 10 in chain B (left) and chain C (right).

Table 4.1. Structure determination statistics

Data Collection	
Beamline	APS 22-ID
Wavelength, Å	1.000
Data processing program	HKL2000
Space group	P21212
<i>a</i> , <i>b</i> , <i>c</i> , Å	137.49, 149.46, 157.11
α , β , γ , °	90, 90, 90
Resolution range, Å	50-3.46 (3.58-3.46)
R_{meas}	0.11 (>1)
$R_{\text{p.i.m.}}$	0.04 (0.50)
$CC_{1/2}$	1.00 (0.70)
Mean $\ \sigma(I) \ $	19.7 (1.0)
Completeness, %	95.8 (90.0)
Average multiplicity	9.9 (6.5)
Wilson <i>B</i> factor, Å ²	
Phasing	
Phasing program	MOLREP
Method	Molecular replacement
Search model	PDB entry 3p05
Model Refinement	
Refinement program	Phenix
Resolution range, Å	40.94-3.46 (3.58-3.46)
No. of unique reflections	24,887 (641)
No. of reflections for R_{free}	1,273 (37)
R_{work}	0.23 (0.25)
R_{free}	0.29 (0.26)
No. of non-H atoms	
Protein	14,550
GS-CA1	900
Average <i>B</i> factor, Å ²	
Protein	43.26
GS-CA1	32.47
R.m.s.d., bond lengths, Å	0.003
R.m.s.d., bond angles, °	0.570
Validation and Deposition	
Ramachandran favored, %	96.21
Ramachandran outliers, %	0
Ramachandran Z-score	
Rotamer outliers, %	0.93
<i>MolProbability</i> clashscore	5.13
PDB code	

Values in parentheses are for the highest resolution shell.

much poorer quality, particularly in helices 9 and 10 (e.g., **Figure 4.1E**); as described before for disulfide-stabilized HIV-1 CA hexamers, this is due to both the inherent flexibility of the CTD and amplified flexibility caused by the W184A and M185A mutations in helix 9.⁷⁷

Density for bound GS-CA1 was observed at each of the ten crystallographically independent NTD-CTD interfaces. Again, there was variation in the drug electron densities, which did not track with density quality for the CTD and more likely reflects differing occupancies. Importantly, ligand positions were also within the best-defined regions of the map, as illustrated in **Figure 4.2A**. For nine of the ten positions, drug densities were sufficiently well-resolved to justify modeling. All nine modeled copies of GS-CA1 adopted the same configuration relative to, and made the same interactions with, each of their binding sites.

Comparison with published crystal structures of the CA hexamer in complex with LEN^{151,152} revealed the same binding pose (**Figure 4.2B**), and that key interactions are preserved: a difluorinated phenyl ring is buried in a hydrophobic pocket lined by Leu56, Met66, Leu69, and Ile73 in the NTD (**Figure 4.2C**); drug nitrogen atoms make hydrogen bonds with Asn57 (**Figure 4.2D**); and a substituted pyrazole ring stacks against Arg173 in the CTD (**Figure 4.2E**). There are three minor chemical differences between GS-CA1 and LEN (circled in **Figure 4.2B** and **4.2F**): a cyclopropyl ring in GS-CA1 is a methyl group in LEN, and two CHF₂ (difluoromethyl) groups in GS-CA1 are CF₃ (trifluoromethyl) in LEN. Of these, only the CHF₂/CF₃ group attached to the pyrazole ring is likely to contribute significantly to interactions with the protein, being buried in a hydrophobic subpocket lined by Tyr169 and Leu172 in the CTD (**Figure 4.2F**). LEN is more metabolically stable than GS-CA1, with improved anti-viral potency.^{148,151,153} Our structural observations are

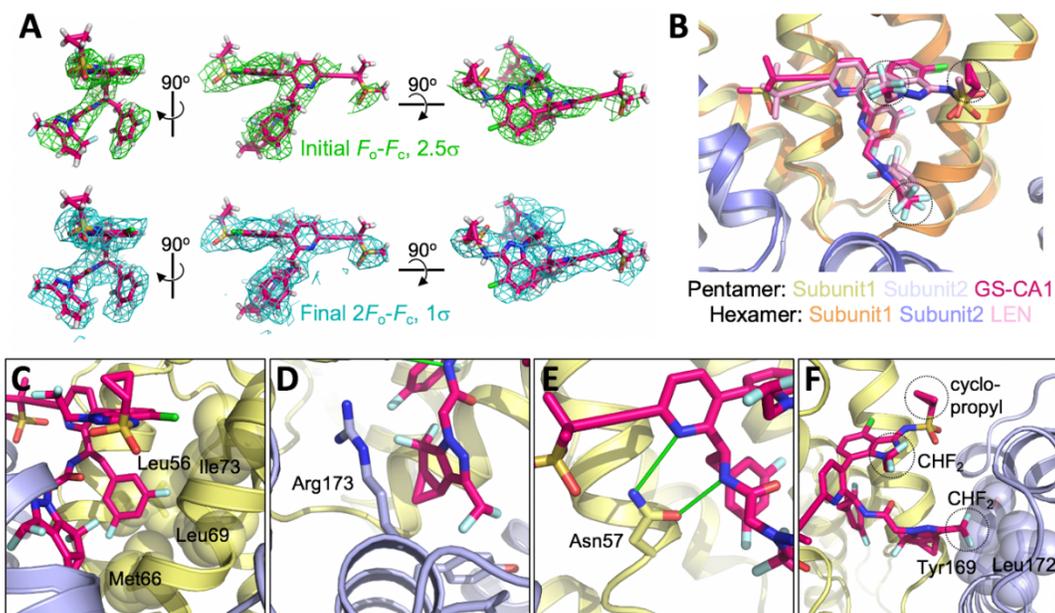


Figure 4.2. GS-CA1 structure bound to CA pentamer. (A) Apparent melting temperatures (T_m) of the indicated CA constructs in the presence of GS-CA1 (blue) or DMSO control (black). Error bars represent the standard deviation of triplicate measurements. P, crosslinked pentamer; H, crosslinked hexamer; M, unassembled CA monomer. (B) Melting profiles of crosslinked pentamer, shown as first derivative curves. Representative of three independent measurements. (C) View of two pentamers in the P21212 asymmetric unit. CA protein is rendered as ribbons (yellow). Nine bound drug molecules are rendered as spheres (pink). (D) Electron density for the NTD (blue mesh, $2mF_o - DF_c$, 1.2σ), illustrated for helices 3 and 4 in chain F, which is representative of all 10 protein chains. (E) Variable electron density for the CTD (blue mesh, $2mF_o - DF_c$, 1σ), illustrated for helices 9 and 10 in chain B (left) and chain C (right).

consistent with findings from biochemical studies that the cross-linked CA pentamer and hexamer bind to GS-CA1 and related ligands with similar affinities.^{30,148}

4.2.3 Structural comparison of HIV-1 CA hexamers and pentamers

Next, we compared our cross-linked CA pentamer to other X-ray and cryoEM structures of the HIV-1 CA hexamer and pentamer. While some of this analysis has been done previously at low resolution,¹⁴⁷ we perform here a more systematic comparison and include recent high-resolution cryoEM structures.^{124,154,155} Using PDB 3p05 (unbound crosslinked pentamer) as reference, we superimposed the two assembly interfaces that hold together the capsomers from structures listed in **Table 4.2**. The NTD-NTD interface (**Figure 4.3A**) consists of helix 2 of one subunit packing against helices 1 and 3 of the

adjacent subunit, whereas the NTD-CTD interface (**Figure 4.3B**) consists of helix 8 from the CTD packing against helices 3 and 4 of the interacting NTD. **Figure 4.3C** is a scatter plot of the RMSDs for crystallographically independent interfaces; similar structures cluster together in this plot. **Figure 4.3D** plots the average of RMSDs for each type of capsomer, indicating degree of similarity to the reference structure. Results show that: (1) GS-CA1 binding does not significantly perturb the structure of the cross-linked pentamer, as evidenced by clustering of these structures in the scatter plot (**Figure 4.3C**) and by average RMSDs $<1 \text{ \AA}$ (**Figure 4.3D**). (2) The cross-linked pentamer is similar to both cross-linked and native hexamers, regardless of drug-binding status, again evidenced by

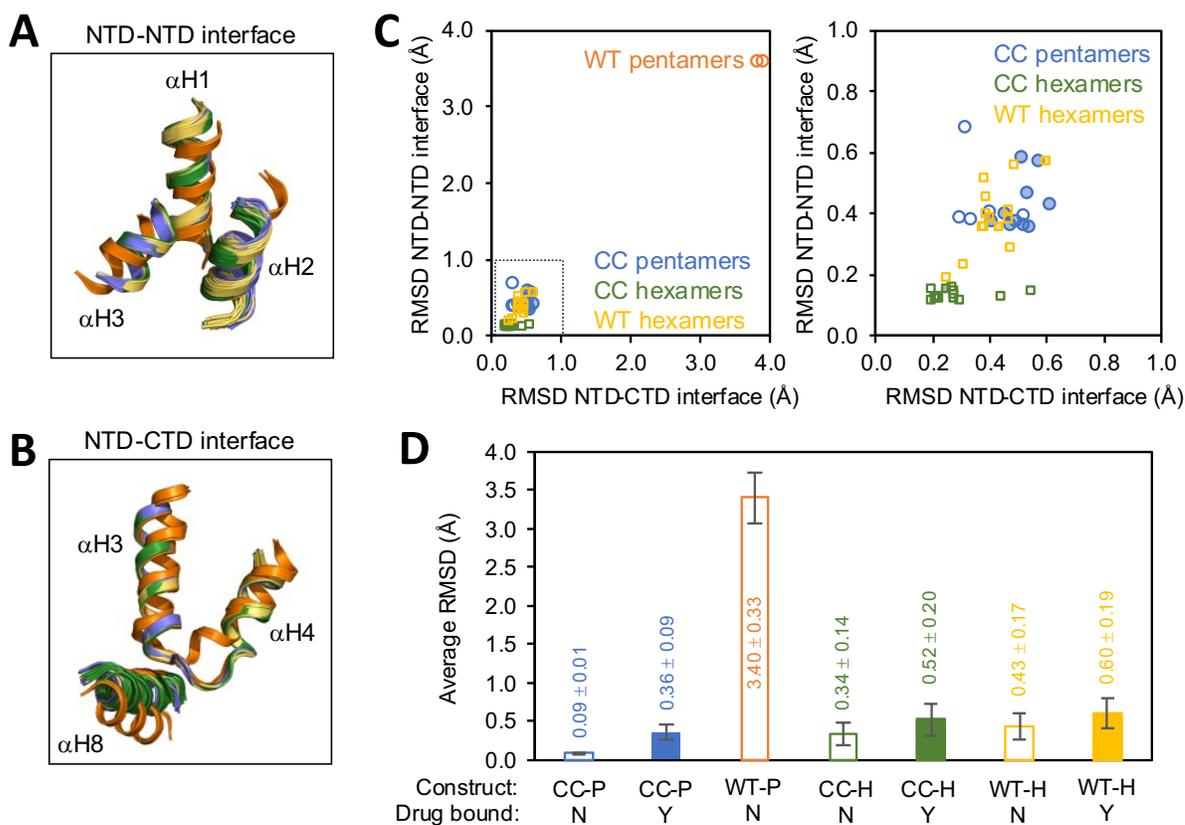


Figure 4.3. Structural comparison of HIV-1 CA capsomers. Crosslinked pentamers are colored in blue, WT pentamers in orange, crosslinked hexamers in green, and WT hexamers in yellow. For each structure, equivalent C α residues were superimposed with PDB 3h47 as reference. **(A)** Superposition of NTD-NTD interfaces in helices 1, 2, and 3. **(B)** Superposition of NTD-CTD interfaces in helices 3, 4, and 8. **(C)** Scatter plot of the RMSDs derived from the superpositions (left panel). Each symbol indicates a crystallographically independent structure extracted from PDB models listed in **Table 4.2**. A zoomed-in view of the graphical region within the dotted lines is shown in the right panel. **(D)** Average of the RMSDs for each group of structures. Error bars indicate standard deviations.

Table 4.2. HIV-1 CA structures used for comparisons in Figure 4.3

Capsomer	Bound drug	PDB	Method	No. of NTD-NTD interfaces	No. of NTD-CTD interfaces	Reference
Crosslinked pentamer	None	3p05	X-ray	5	5	Pornillos <i>et al.</i> (2011) <i>Nature</i> 469 , 424
	GS-CA1	TBD	X-ray	10	10	This study
WT pentamer	None	7urn	CryoEM	1	1	Schirra <i>et al.</i> (2023) <i>Nat Struct Mol Biol</i> 30 , 383
	None	8ckw	CryoEM	1	1	Stacey <i>et al.</i> (2023) <i>Proc Natl Acad Sci USA</i> 120 , e2220557120
Crosslinked hexamer	None	3h47	X-ray	1	1	Pornillos <i>et al.</i> (2009) <i>Cell</i> 137 , 1282
	None	3h4e	X-ray	12	12	Pornillos <i>et al.</i> (2009) <i>Cell</i> 137 , 1282
	PF74	4qnb	X-ray	1	1	Bhattacharya <i>et al.</i> (2014) <i>Proc Natl Acad Sci USA</i> 111 , 18625
	PF74	4u0e	X-ray	1	1	Price <i>et al.</i> (2014) <i>PLoS Pathog</i> 10 , e1004459
	BI-2	4u0f	X-ray	1	1	Price <i>et al.</i> (2014) <i>PLoS Pathog</i> 10 , e1004459
	LEN	6v2f	X-ray	6	6	Link <i>et al.</i> (2020) <i>Nature</i> 584 , 614
	LEN	6vkv	X-ray	3	3	Bester <i>et al.</i> (2020) <i>Science</i> 370 , 360
WT hexamer	None	4xfx	X-ray	1	1	Gres <i>et al.</i> (2015) <i>Science</i> 349 , 99
	None	4xfy	X-ray	1	1	Gres <i>et al.</i> (2015) <i>Science</i> 349 , 99
	None	7urn	CryoEM	2	2	Schirra <i>et al.</i> (2023) <i>Nat Struct Mol Biol</i> 30 , 383
	None	8ckv	CryoEM	1	1	Stacey <i>et al.</i> (2023) <i>Proc Natl Acad Sci USA</i> 120 , e2220557120
	None	8ckx	CryoEM	6	6	Stacey <i>et al.</i> (2023) <i>Proc Natl Acad Sci USA</i> 120 , e2220557120
	PF74	4xfz	X-ray	1	1	Gres <i>et al.</i> (2015) <i>Science</i> 349 , 99
	LEN	8g6o	CryoEM	2	2	Highland <i>et al.</i> (2023) <i>Proc Natl Acad Sci USA</i> 120 , e2220545120

clustering (**Figure 4.3C**) and by low RMSDs (**Figure 4.3D**). (3) The cross-linked pentamer, and both the crosslinked and WT hexamers, have very different NTD-NTD and NTD-CTD interfaces from the native pentamer. The WT pentamers form cluster widely separated from all the others (**Figure 4.3C**) and have RMSDs ~ 10 Å (**Figure 4.3D**). These observations provide further evidence that the cross-linked pentamer and WT hexamer are functionally equivalent, whereas both are functionally distinct from the WT pentamer.

4.3 Discussion

Although interactions of the HIV-1 CA hexamer with many of its known binding partners have been well-characterized, the CA pentamer's ability to interact with the same ligands and the consequences of these interactions remain poorly understood. The key conclusion from this study is that the disulfide-crosslinked HIV-1 CA pentamer recapitulates hexamer-like functional characteristics. Specifically, the crosslinked pentamer binds ligand, such as GS-CA1, in the same manner as the hexamer. Due to their favorable characteristics compared to hexamer-containing and hexamer-only assemblies, the 30-nm spheres made of crosslinked CA pentamers may be attractive functional surrogates of the hexamer.¹⁵⁰

Although distinct from the native pentamer, we suggest that similarities with the hexamer warrant further studies, to determine whether the alternative configuration adopted by the crosslinked pentamer represents a biologically functional state of HIV-1 CA. It is possible that specific molecular triggers, such as a high-affinity ligand or host factor binding, might allow conformational switching of the native pentamer into this configuration. For example, we speculate that such a switch might be relevant to the "capsid remodeling" process that is proposed to occur during the nuclear import step of HIV-1 replication.^{102,104}

4.4 Methods

4.4.1 Protein production

HIV-1 CA proteins were expressed and purified as described previously.¹⁵⁶ Soluble pentamers were generated as described.¹⁰⁷ In brief, purified CA protein harboring the

N21C, A22C, W184A, and M185A mutations was sequentially dialyzed (at 4 °C and for 8-12 h each step), first into assembly buffer (50 mM Tris-pH 8.0, 1 M NaCl, 100 mM β -mercaptoethanol), then into assembly buffer without mercaptoethanol to allow the disulfide bonds to form, and finally into crystallization buffer (20 mM Tris, pH 8.0).

4.4.2 Nanodifferential scanning fluorimetry

Protein samples (0.4 mM stock) and drug (10 mM stock in DMSO) were diluted to 5 μ M concentrations in 20 mM Tris, pH 8.0, 100 mM NaCl, and then incubated for 5 mins. Melting profiles were measured by using a Tycho NT.6 (NanoTemper). First derivative curves and apparent melting temperatures (T_m) were determined from raw curves using the manufacturer's software.

4.4.3 Crystallization

Protein-drug samples were prepared by incubating 0.4 mM of crosslinked pentamers and 1 mM GS-CA1 overnight at room temperature. GS-CA1 was directly diluted from a 50 mM stock in DMSO, which resulted in significant drug precipitation. The protein-drug mixture was crystallized in sitting drops at 20 °C. Crystals formed within 48 h in 0.1 M HEPES, pH 6.8, 21 (v/v) % PEG 8000, and 10% (v/v) ethylene glycol at a 1:1 sample to precipitant ratio.

4.4.4 Structure determination

Crystals were cryoprotected with 25% (v/v) ethylene glycol in mother liquor and flash-frozen in liquid nitrogen. Diffraction data were collected on beamline 22-ID at the Advanced Photon Source. Indexing, integration, and scaling were performed with HKL-2000.¹⁵⁷ The crystals belonged to space group P21212, with two pentamers in the asymmetric unit. The structure was solved by molecular replacement with MOLREP,¹⁵⁸

using as search model PDB entry 3p05.⁷⁶ Iterative refinement and model building were performed using Phenix¹⁵⁹ and Coot.¹⁴¹ Initial model building was performed with 10-fold non-crystallographic symmetry (NCS) maps, calculated separately for the NTD and CTD. The drug was built and refined only after complete refinement of the protein. NCS restraints were used throughout the refinement process. Structure validation was performed using MolProbity¹⁶⁰ as implemented in Phenix. Figures were prepared using PyMOL (Schrödinger Scientific).

4.4.5 Structure comparisons

Pair-wise alignments of published structures (Supplementary Table 1) of crosslinked pentamers, wild type (WT) pentamers, crosslinked hexamers, and WT hexamers – with or without bound drug – were performed with PyMOL, using chains E and F of PDB 3p05 as reference. RMSDs were calculated for corresponding C α atoms. The following residues were used for alignment: NTD-NTD interface, helix 1 (a.a. 17-29), helix 3 (a.a. 49-56), and helix 2 (a.a. 34-44); NTD-CTD interface, helix 3 (a.a. 49-56), helix 4 (a.a. 63-73), and helix 8 (a.a. 161-174).

4.5 Acknowledgements

We thank Yueping Wan and Kaneil Zadrozny for assistance with protein purifications and crystallization, and Gilead Sciences for providing initial samples of GS-CA1.

Chapter 5: Conclusion

5.1 Summary of results

The HIV-1 capsid is a dynamic molecular machine essential to post-cell entry viral events. This work systematically employed a series of biochemical reagents that recapitulate various levels of capsid assembly to assess interactions with important binding factors and their related consequences. The results reveal novel biochemical and biophysical details fundamental to understanding capsid dynamics and their larger role in early HIV-1 infection.

We show the first structural evidence of the HIV-1 capsid binding to PQBP1, a host cell protein recently identified to bridge capsid recognition and activation of the cGAS innate immune sensing pathway upon viral entry. Our thermostability and structural data demonstrate that the primary mode of this interaction is through charge complementation between the electronegative N-terminus of PQBP1 and the arginine 18 pore found within the CA hexamer. Pulldown assays confirm binding of full-length PQBP1 to the extended capsid lattice and also suggest that additional regions of PQBP1 may be contributing to capsid binding. Furthermore, pulldown assays performed with CA tubes after extended periods of FL PQBP1 incubation demonstrate increases in levels of soluble CA, which can be interpreted as capsid lattice destabilization as a consequence of PQBP1 binding. Negative stain EM images of CA tubes from these experiments appear damaged, reinforcing pulldown results. However, these effects no longer occur when the C-terminus of PQBP1 is not present.

We also investigate PQBP1's propensity to self-associate as an intrinsically disordered protein. Pulldown assays of FL PQBP1 incubated with increasing amounts of Nt show that self-association of the protein can be unmasked by the N-terminal region of

PQBP1. Mass photometry data confirm these results and demonstrate that PQBP1 undergoes dimerization, likely at its C-terminus. In-cell experiments performed with a characterized PQBP1 construct that C-terminally dimerizes via a leucine zipper indicate that C-terminal dimerization of PQBP1 is required to induce cGAS signaling upon viral infection.

The collective data suggest a potential mechanism for how PQBP1 can bridge capsid recognition with cGAS recruitment to the virus to induce downstream signaling: after viral entry, PQBP1 binds to the capsid primarily through contacts between its N-terminus and the R18 pore within CA hexamers, which may be reinforced by additional interactions via other regions of the protein. Nt binding to the capsid frees PQBP1 Ct to dimerize, and this self-association induces conformational changes and/or mechanical stress sufficient to promote capsid lattice destabilization. PQBP1 self-association can also serve as a platform for cGAS recruitment, enabling multiple cGAS molecules to bind at sites of capsid lattice breakage and bringing them into close proximity of the reverse-transcribing viral genome. These local concentrations of activated cGAS would then reach the signaling threshold necessary to induce downstream events.

Finally, we solved the structure of small molecule inhibitor GS-CA1 in complex with the cross-linked HIV-1 CA pentamer. Though binding interactions between CA hexamers and small molecule inhibitors including but not limited to GS-CA1 have been previously elucidated, we provide the first evidence of GS-CA1 binding to a CA pentamer. Structural and thermostability data reveal the hexamer-like characteristics of the cross-linked CA pentamer, which adopts a conformation alternative to the canonical CA pentamer. These results contribute to the mounting body of evidence supporting CA pentamer structural plasticity, a characteristic of the HIV-1 capsid that may have important functional relevance.

5.2 Impacts and future directions

Additional studies focused on PQBP1 should aim to elucidate a more complete molecular understanding of how capsid recognition occurs; while we describe the primary interactions between PQBP1 and the capsid and note that other interactions contribute to stable binding, these regions and the resulting complex(es) that form remain to be fully characterized. Our pulldown assays, which distinguish genetically separate functionalities of binding and destabilization based on length of incubation (short times show binding, long times show destabilization), should be particularly useful for these studies. We also successfully used functional replacement of distinct PQBP1 regions to assess self-association properties, specifically the dimerization propensity of Ct. Both of these approaches will be informative in further work not just with PQBP1 but also in determining how cGAS and further higher-order complex formation generates the proposed anti-viral signaling platform threshold. Once these biochemical details have been uncovered, they should be validated *in vivo*.

The alternative CA pentamer should also continue to be investigated for functional relevance during infection. Unpublished findings from our group demonstrate that this conformation can be obtained multiple ways *in vitro* through cross-linking or point mutations (e.g. R18L, K25A), giving a toehold for in-cell experiments that may reveal important functional information.

In conclusion, the data presented in this dissertation establish the fundamental molecular bases of important interactions of the HIV-1 capsid and set the foundation for future mechanistic studies, including potential development of therapeutics that target the HIV-1 capsid to mitigate viral infection.

Chapter 6: References

1. Zhang, Z.-Q. *et al.* Sexual Transmission and Propagation of SIV and HIV in Resting and Activated CD4⁺ T Cells. *Science* **286**, 1353–1357 (1999).
2. Lahaye, X. *et al.* The Capsids of HIV-1 and HIV-2 Determine Immune Detection of the Viral cDNA by the Innate Sensor cGAS in Dendritic Cells. *Immunity* **39**, 1132–1142 (2013).
3. Weinberg *et al.* Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *J. Exp. Med.* **174**, 1477–1482 (1991).
4. Centers for Disease Control and Prevention. *About HIV/AIDS*. <https://www.cdc.gov/hiv/basics/whatishiv.html#is-there-a-cure> (2019).
5. World Health Organization. HIV statistics, globally and by WHO region, 2023. <https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics> (2023).
6. Centers for Disease Control and Prevention. Prevention | HIV Basics | HIV/AIDS | CDC. <https://www.cdc.gov/hiv/basics/prevention.html> (2023).
7. Antiretroviral Therapy Cohort Collaboration. Survival of HIV-positive patients starting antiretroviral therapy between 1996 and 2013: a collaborative analysis of cohort studies. *Lancet HIV* **4**, e349–e356 (2017).
8. Rodger, A. J. *et al.* Sexual Activity Without Condoms and Risk of HIV Transmission in Serodifferent Couples When the HIV-Positive Partner Is Using Suppressive Antiretroviral Therapy. *JAMA* **316**, 171–181 (2016).
9. National Institutes of Health. Adverse Effects of Antiretroviral Agents. <https://clinicalinfo.hiv.gov/en/guidelines/hiv-clinical-guidelines-adult-and-adolescent-arv/adverse-effects-antiretroviral-agents> (2023).
10. HIV-2 Infection | NIH. <https://clinicalinfo.hiv.gov/en/guidelines/hiv-clinical-guidelines-adult-and-adolescent-arv/hiv-2-infection> (2019).
11. Marlink, R. *et al.* Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science* **265**, 1587–1590 (1994).

12. Matheron, S. *et al.* Factors associated with clinical progression in HIV-2 infected-patients: the French ANRS cohort. *AIDS Lond. Engl.* **17**, 2593–2601 (2003).
13. Motomura, K., Chen, J. & Hu, W.-S. Genetic Recombination between Human Immunodeficiency Virus Type 1 (HIV-1) and HIV-2, Two Distinct Human Lentiviruses. *J. Virol.* **82**, 1923–1933 (2008).
14. Wilen, C. B., Tilton, J. C. & Doms, R. W. HIV: Cell Binding and Entry. *Cold Spring Harb. Perspect. Med.* **2**, (2012).
15. Campbell, E. M. & Hope, T. J. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat. Rev. Microbiol.* **13**, 471–483 (2015).
16. Naghavi, M. H. HIV-1 capsid exploitation of the host microtubule cytoskeleton during early infection. *Retrovirology* **18**, 19 (2021).
17. Huang, P.-T. *et al.* FEZ1 Is Recruited to a Conserved Cofactor Site on Capsid to Promote HIV-1 Trafficking. *Cell Rep.* **28**, 2373-2385.e7 (2019).
18. Malikov, V. *et al.* HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus. *Nat. Commun.* **6**, 6660 (2015).
19. Malikov, V. & Naghavi, M. H. FEZ1 Plays Dual Roles in Early HIV-1 Infection by Independently Regulating Capsid Transport and Host Interferon-Stimulated Gene Expression. *J. Virol.* **0**, e00499-23 (2023).
20. Matreyek, K. A., Yücel, S. S., Li, X. & Engelman, A. Nucleoporin NUP153 Phenylalanine-Glycine Motifs Engage a Common Binding Pocket within the HIV-1 Capsid Protein to Mediate Lentiviral Infectivity. *PLoS Pathog.* **9**, e1003693 (2013).
21. Francis, A. C. & Melikyan, G. B. Single HIV-1 imaging reveals progression of infection through CA-dependent steps of docking at the nuclear pore, uncoating and nuclear transport. *Cell Host Microbe* **23**, 536-548.e6 (2018).
22. Burdick, R. C. *et al.* Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog.* **13**, e1006570 (2017).
23. Schaller, T. *et al.* HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathog.* **7**, e1002439 (2011).

24. Kane, M. *et al.* Nuclear pore heterogeneity influences HIV-1 infection and the antiviral activity of MX2. *eLife* **7**, e35738 (2018).
25. Ocwieja, K. E. *et al.* HIV Integration Targeting: A Pathway Involving Transportin-3 and the Nuclear Pore Protein RanBP2. *PLoS Pathog.* **7**, e1001313 (2011).
26. Dharan, A. *et al.* KIF5B and Nup358 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. *PLoS Pathog.* **12**, e1005700 (2016).
27. Di Nunzio, F. *et al.* Human Nucleoporins Promote HIV-1 Docking at the Nuclear Pore, Nuclear Import and Integration. *PLoS ONE* **7**, e46037 (2012).
28. Buffone, C. *et al.* Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells. *J. Virol.* **92**, e00648-18 (2018).
29. Matreyek, K. A. & Engelman, A. The Requirement for Nucleoporin NUP153 during Human Immunodeficiency Virus Type 1 Infection Is Determined by the Viral Capsid. *J. Virol.* **85**, 7818–7827 (2011).
30. Bhattacharya, A. *et al.* Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc. Natl. Acad. Sci.* **111**, 18625–18630 (2014).
31. Bejarano, D. A. *et al.* HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *eLife* **8**, e41800 (2019).
32. Achuthan, V. *et al.* Capsid-CPSF6 interaction licenses nuclear HIV-1 trafficking to sites of viral DNA integration. *Cell Host Microbe* **24**, 392-404.e8 (2018).
33. Zhou, L. *et al.* Transportin 3 Promotes a Nuclear Maturation Step Required for Efficient HIV-1 Integration. *PLoS Pathog.* **7**, e1002194 (2011).
34. Lee, K. *et al.* Flexible Use of Nuclear Import Pathways by HIV-1. *Cell Host Microbe* **7**, 221–233 (2010).
35. Bichel, K. *et al.* HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. *Retrovirology* **10**, 81 (2013).
36. Chen, N.-Y. *et al.* HIV-1 capsid is involved in post-nuclear entry steps. *Retrovirology* **13**, 28 (2016).

37. Chin, C. R. *et al.* Direct Visualization of HIV-1 Replication Intermediates shows that Viral Capsid and CPSF6 Modulate HIV-1 Intra-nuclear Invasion and Integration. *Cell Rep.* **13**, 1717–1731 (2015).
38. Francis, A. C., Marin, M., Prellberg, M. J., Palermino-Rowland, K. & Melikyan, G. B. HIV-1 Uncoating and Nuclear Import Precede the Completion of Reverse Transcription in Cell Lines and in Primary Macrophages. *Viruses* **12**, 1234 (2020).
39. Rankovic, S., Deshpande, A., Harel, S., Aiken, C. & Rousso, I. HIV-1 uncoating occurs via a series of rapid biomechanical changes in the core related to individual stages of reverse transcription. *J. Virol.* (2021) doi:10.1128/JVI.00166-21.
40. Müller, T. G. *et al.* HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *eLife* **10**, e64776 (2021).
41. Burdick, R. C. *et al.* HIV-1 uncoats in the nucleus near sites of integration. *Proc. Natl. Acad. Sci.* **117**, 5486–5493 (2020).
42. Marini, B. *et al.* Nuclear architecture dictates HIV-1 integration site selection. *Nature* **521**, 227–231 (2015).
43. Sowd, G. A. *et al.* A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1054–E1063 (2016).
44. Fassati, A. Multiple roles of the capsid protein in the early steps of HIV-1 infection. *Virus Res.* **170**, 15–24 (2012).
45. Forshey, B. M., von Schwedler, U., Sundquist, W. I. & Aiken, C. Formation of a Human Immunodeficiency Virus Type 1 Core of Optimal Stability Is Crucial for Viral Replication. *J. Virol.* **76**, 5667–5677 (2002).
46. von Schwedler, U. K., Stray, K. M., Garrus, J. E. & Sundquist, W. I. Functional Surfaces of the Human Immunodeficiency Virus Type 1 Capsid Protein. *J. Virol.* **77**, 5439–5450 (2003).
47. Christensen, D. E., Ganser-Pornillos, B. K., Johnson, J. S., Pornillos, O. & Sundquist, W. I. Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. *Science* **370**, (2020).

48. Jennings, J. *et al.* The HIV-1 capsid serves as a nanoscale reaction vessel for reverse transcription. *bioRxiv* 2023.11.08.566350 (2023)
doi:10.1101/2023.11.08.566350.
49. Altfeld, M. & Gale Jr, M. Innate immunity against HIV-1 infection. *Nat. Immunol.* **16**, 554–562 (2015).
50. Papa, G. *et al.* IP6-stabilised HIV capsids evade cGAS/STING-mediated host immune sensing. *EMBO Rep.* **n/a**, e56275 (2023).
51. Yamashita, M. & Emerman, M. Capsid Is a Dominant Determinant of Retrovirus Infectivity in Nondividing Cells. *J. Virol.* **78**, 5670–5678 (2004).
52. Wu, J., Matunis, M. J., Kraemer, D., Blobel, G. & Coutavas, E. Nup358, a Cytoplasmically Exposed Nucleoporin with Peptide Repeats, Ran-GTP Binding Sites, Zinc Fingers, a Cyclophilin A Homologous Domain, and a Leucine-rich Region. *J. Biol. Chem.* **270**, 14209–14213 (1995).
53. Krull, S., Thyberg, J., Björkroth, B., Rackwitz, H.-R. & Cordes, V. C. Nucleoporins as Components of the Nuclear Pore Complex Core Structure and Tpr as the Architectural Element of the Nuclear Basket. *Mol. Biol. Cell* **15**, 4261–4277 (2004).
54. Fernandez, J. *et al.* Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. *Nat. Microbiol.* **4**, 1840–1850 (2019).
55. Balasubramaniam, M. *et al.* PF74 Inhibits HIV-1 Integration by Altering the Composition of the Preintegration Complex. *J. Virol.* **93**, (2019).
56. Peng, K. *et al.* Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. *eLife* **3**, e04114 (2014).
57. Price, A. J. *et al.* Host Cofactors and Pharmacologic Ligands Share an Essential Interface in HIV-1 Capsid That Is Lost upon Disassembly. *PLoS Pathog.* **10**, e1004459 (2014).
58. Li, S. *et al.* Defining the HIV Capsid Binding Site of Nucleoporin 153. *mSphere* **0**, e00310-22 (2022).
59. Shen, Q. *et al.* The capsid lattice engages a bipartite NUP153 motif to mediate nuclear entry of HIV-1 cores. *Proc. Natl. Acad. Sci.* **120**, e2202815120 (2023).

60. Xue, G. *et al.* The HIV-1 capsid core is an opportunistic nuclear import receptor. *Nat. Commun.* **14**, 3782 (2023).
61. Bedwell, G. J. & Engelman, A. N. Factors that mold the nuclear landscape of HIV-1 integration. *Nucleic Acids Res.* (2020) doi:10.1093/nar/gkaa1207.
62. Achuthan, V., Perreira, J. M., Ahn, J. J., Brass, A. L. & Engelman, A. N. Capsid-CPSF6 interaction: Master regulator of nuclear HIV-1 positioning and integration. *J. Life Sci. Westlake Village Calif* **1**, 39–45 (2019).
63. Engelman, A. N. HIV Capsid and Integration Targeting. *Viruses* **13**, 125 (2021).
64. Dharan, A., Bachmann, N., Talley, S., Zwickelmaier, V. & Campbell, E. M. Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. *Nat. Microbiol.* **5**, 1088–1095 (2020).
65. Eschbach, J. E. *et al.* Capsid lattice destabilization leads to premature loss of the viral genome and integrase enzyme during HIV-1 infection. *J. Virol.* (2020) doi:10.1128/JVI.00984-20.
66. Roa, A. *et al.* RING Domain Mutations Uncouple TRIM5 α Restriction of HIV-1 from Inhibition of Reverse Transcription and Acceleration of Uncoating. *J. Virol.* **86**, 1717–1727 (2012).
67. Stremlau, M. *et al.* The cytoplasmic body component TRIM5 α restricts HIV-1 infection in Old World monkeys. *Nature* **427**, 848–853 (2004).
68. Stremlau, M. *et al.* Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5 α restriction factor. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5514–5519 (2006).
69. Grime, J. M. A. *et al.* Coarse-grained simulation reveals key features of HIV-1 capsid self-assembly. *Nat. Commun.* **7**, 11568 (2016).
70. Briggs, J. A. G., Wilk, T., Welker, R., Kräusslich, H.-G. & Fuller, S. D. Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.* **22**, 1707–1715 (2003).
71. Shin, R., Tzou, Y.-M. & Krishna, N. R. Structure of a monomeric mutant of the HIV-1 capsid protein. *Biochemistry* **50**, 9457–9467 (2011).

72. Gres, A. T. *et al.* X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **349**, 99–103 (2015).
73. Pornillos, O. & Ganser-Pornillos, B. K. Maturation of retroviruses. *Curr. Opin. Virol.* **36**, 47–55 (2019).
74. Ganser, B. K. Assembly and Analysis of Conical Models for the HIV-1 Core. *Science* **283**, 80–83 (1999).
75. Ganser-Pornillos, B. K., Cheng, A. & Yeager, M. Structure of Full-Length HIV-1 CA: A Model for the Mature Capsid Lattice. *Cell* **131**, 70–79 (2007).
76. Pornillos, O., Ganser-Pornillos, B. K. & Yeager, M. Atomic-level modelling of the HIV capsid. *Nature* **469**, 424–427 (2011).
77. Pornillos, O. *et al.* X-Ray Structures of the Hexameric Building Block of the HIV Capsid. *Cell* **137**, 1282–1292 (2009).
78. Li, S., Hill, C. P., Sundquist, W. I. & Finch, J. T. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409–413 (2000).
79. Lanman, J., Sexton, J., Sakalian, M. & Prevelige, P. E. Kinetic Analysis of the Role of Intersubunit Interactions in Human Immunodeficiency Virus Type 1 Capsid Protein Assembly In Vitro. *J. Virol.* **76**, 6900–6908 (2002).
80. Lanman, J. *et al.* Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. *Nat. Struct. Mol. Biol.* **11**, 676–677 (2004).
81. Grime, J. M. A. & Voth, G. A. Early Stages of the HIV-1 Capsid Protein Lattice Formation. *Biophys. J.* **103**, 1774–1783 (2012).
82. Toronto Research Chemicals. Phytic Acid. <https://www.trc-canada.com/product-detail/?CatNum=P398700> (2018).
83. Dick, R. A. *et al.* Inositol phosphates are assembly co-factors for HIV-1. *Nature* **560**, 509–512 (2018).
84. Jacques, D. A. *et al.* HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349–353 (2016).

85. Ganser-Pornillos, B. K. & Pornillos, O. Restriction of HIV-1 and other retroviruses by TRIM5. *Nat. Rev. Microbiol.* **17**, 546–556 (2019).
86. Dostálková, A. *et al.* PF74 and Its Novel Derivatives Stabilize Hexameric Lattice of HIV-1 Mature-Like Particles. *Molecules* **25**, 1895 (2020).
87. Nka, A. D. *et al.* Evaluation of HIV-1 capsid genetic variability and lenacapavir (GS-6207) drug resistance-associated mutations according to viral clades among drug-naive individuals. *J. Antimicrob. Chemother.* **dkac388** (2022)
doi:10.1093/jac/dkac388.
88. Selyutina, A. *et al.* GS-CA1 and lenacapavir stabilize the HIV-1 core and modulate the core interaction with cellular factors. *iScience* **25**, 103593 (2021).
89. Ambrose, Z. & Aiken, C. HIV-1 Uncoating: Connection to Nuclear Entry and Regulation by Host Proteins. *Virology* **0**, 371–379 (2014).
90. Yang, Y., Fricke, T. & Diaz-Griffero, F. Inhibition of Reverse Transcriptase Activity Increases Stability of the HIV-1 Core. *J. Virol.* **87**, 683–687 (2013).
91. Arfi, V. *et al.* Characterization of the Behavior of Functional Viral Genomes during the Early Steps of Human Immunodeficiency Virus Type 1 Infection. *J. Virol.* **83**, 7524–7535 (2009).
92. Francis, A. C., Marin, M., Shi, J., Aiken, C. & Melikyan, G. B. Time-Resolved Imaging of Single HIV-1 Uncoating In Vitro and in Living Cells. *PLoS Pathog.* **12**, e1005709 (2016).
93. Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C. & Rousso, I. Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J. Virol.* **91**, e00289-17 (2017).
94. Yang, Y., Luban, J. & Diaz-Griffero, F. The Fate of HIV-1 Capsid: A Biochemical Assay for HIV-1 Uncoating. *Methods Mol. Biol.* **1087**, 29–36 (2014).
95. Arhel, N. J. *et al.* HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J.* **26**, 3025–3037 (2007).
96. Bönisch, I. Z. *et al.* Capsid-Labelled HIV To Investigate the Role of Capsid during Nuclear Import and Integration. *J. Virol.* **94**, e01024 (2020).

97. Lin, D. H. *et al.* Architecture of the symmetric core of the nuclear pore. *Science* **352**, aaf1015 (2016).
98. Panté, N. & Kann, M. Nuclear Pore Complex Is Able to Transport Macromolecules with Diameters of ~39 nm. *Mol. Biol. Cell* **13**, 425–434 (2002).
99. Zila, V. *et al.* Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell* (2021) doi:10.1016/j.cell.2021.01.025.
100. Li, C., Burdick, R. C., Nagashima, K., Hu, W.-S. & Pathak, V. K. HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proc. Natl. Acad. Sci.* **118**, (2021).
101. Schifferdecker, S. *et al.* Direct Capsid Labeling of Infectious HIV-1 by Genetic Code Expansion Allows Detection of Largely Complete Nuclear Capsids and Suggests Nuclear Entry of HIV-1 Complexes via Common Routes. *mBio* **0**, e01959-22 (2022).
102. Francis, A. C. *et al.* Localization and functions of native and eGFP-tagged capsid proteins in HIV-1 particles. *PLOS Pathog.* **18**, e1010754 (2022).
103. Monette, A., Panté, N. & Moulard, A. J. HIV-1 remodels the nuclear pore complex. *J. Cell Biol.* **193**, 619–631 (2011).
104. Guedán, A. *et al.* HIV-1 requires capsid remodelling at the nuclear pore for nuclear entry and integration. *PLOS Pathog.* **17**, e1009484 (2021).
105. Deshpande, A. *et al.* Elasticity of the HIV-1 Core Facilitates Nuclear Entry and Infection. *BioRxiv Prepr. Serv. Biol.* 2023.09.29.560083 (2023) doi:10.1101/2023.09.29.560083.
106. Li, X. *et al.* HIV-1 viral cores enter the nucleus collectively through the nuclear endocytosis-like pathway. *Sci. China Life Sci.* **64**, 66–76 (2021).
107. Pornillos, O., Ganser-Pornillos, B. K., Banumathi, S., Hua, Y. & Yeager, M. Disulfide Bond Stabilization of the Hexameric Capsomer of Human Immunodeficiency Virus. *J. Mol. Biol.* **401**, 985–995 (2010).
108. Mosallanejad, K. & Kagan, J. C. Control of innate immunity by the cGAS-STING pathway. *Immunol. Cell Biol.* **100**, 409–423 (2022).

109. Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP Synthase is a Cytosolic DNA Sensor that Activates the Type-I Interferon Pathway. *Science* **339**, 10.1126/science.1232458 (2013).
110. Wu, J. *et al.* Cyclic-GMP-AMP Is An Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA. *Science* **339**, 10.1126/science.1229963 (2013).
111. Manel, N. *et al.* A cryptic sensor for HIV-1 activates antiviral innate immunity in dendritic cells. *Nature* **467**, 214–217 (2010).
112. Gao, D. *et al.* Cyclic GMP-AMP Synthase Is an Innate Immune Sensor of HIV and Other Retroviruses. *Science* **341**, 903–906 (2013).
113. Yoh, S. M. *et al.* PQBP1 is a Proximal Sensor of the cGAS-dependent Innate Response to HIV-1. *Cell* **161**, 1293–1305 (2015).
114. Yoh, S. M. *et al.* Recognition of HIV-1 capsid by PQBP1 licenses an innate immune sensing of nascent HIV-1 DNA. *Mol. Cell* (2022) doi:10.1016/j.molcel.2022.06.010.
115. Rees, M. *et al.* Solution Model of the Intrinsically Disordered Polyglutamine Tract-Binding Protein-1. *Biophys. J.* **102**, 1608–1616 (2012).
116. Mizuguchi, M. & Okazawa, H. [Structural study of polyglutamine tract-binding protein 1]. *Yakugaku Zasshi* **133**, 519–526 (2013).
117. Tanaka, H. & Okazawa, H. PQBP1: The Key to Intellectual Disability, Neurodegenerative Diseases, and Innate Immunity. *Int. J. Mol. Sci.* **23**, 6227 (2022).
118. Sudol, M., McDonald, C. B. & Farooq, A. Molecular insights into the WW domain of the Golabi-Ito-Hall syndrome protein PQBP1. *FEBS Lett.* **586**, 2795–2799 (2012).
119. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
120. Mallery, D. L. *et al.* IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *eLife* **7**, e35335 (2018).
121. Gao, K., Oerlemans, R. & Groves, M. R. Theory and applications of differential scanning fluorimetry in early-stage drug discovery. *Biophys. Rev.* **12**, 85–104 (2020).

122. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
123. Malagrino, F. *et al.* Unveiling induced folding of intrinsically disordered proteins - Protein engineering, frustration and emerging themes. *Curr. Opin. Struct. Biol.* **72**, 153–160 (2022).
124. Schirra, R. T. *et al.* A molecular switch modulates assembly and host factor binding of the HIV-1 capsid. *Nat. Struct. Mol. Biol.* **30**, 383–390 (2023).
125. Letcher, A. J., Schell, M. J. & Irvine, R. F. Do mammals make all their own inositol hexakisphosphate? *Biochem. J.* **416**, 263–270 (2008).
126. Diamond, T. L. *et al.* Macrophage Tropism of HIV-1 Depends on Efficient Cellular dNTP Utilization by Reverse Transcriptase*. *J. Biol. Chem.* **279**, 51545–51553 (2004).
127. Kennedy, E. M. *et al.* Ribonucleoside Triphosphates as Substrate of Human Immunodeficiency Virus Type 1 Reverse Transcriptase in Human Macrophages*. *J. Biol. Chem.* **285**, 39380–39391 (2010).
128. Yap, M. W., Mortuza, G. B., Taylor, I. A. & Stoye, J. P. The design of artificial retroviral restriction factors. *Virology* **365**, 302–314 (2007).
129. Javanbakht, H. *et al.* The ability of multimerized cyclophilin A to restrict retrovirus infection. *Virology* **367**, 19–29 (2007).
130. Fribourgh, J. L. *et al.* Structural Insight into HIV-1 Restriction by MxB. *Cell Host Microbe* **16**, 627–638 (2014).
131. Goldstone, D. C. *et al.* Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. *Proc. Natl. Acad. Sci.* **111**, 9609–9614 (2014).
132. Liu, C. *et al.* Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nat. Commun.* **7**, 10714 (2016).
133. Wagner, J. M. *et al.* Mechanism of B-box 2 domain-mediated higher-order assembly of the retroviral restriction factor TRIM5 α . *eLife* **5**, e16309 (2016).

134. Wagner, J. M. *et al.* General Model for Retroviral Capsid Pattern Recognition by TRIM5 Proteins. *J. Virol.* **92**, e01563-17 (2018).
135. Ning, J. *et al.* Truncated CPSF6 Forms Higher-Order Complexes That Bind and Disrupt HIV-1 Capsid. *J. Virol.* **92**, e00368-18 (2018).
136. Wei, G. *et al.* Prion-like low complexity regions enable avid virus-host interactions during HIV-1 infection. *Nat. Commun.* **13**, 5879 (2022).
137. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
138. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
139. Zheng, S. Q. *et al.* MotionCor2 - anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331 (2017).
140. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
141. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486 (2010).
142. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861 (2019).
143. Piacentini, J. *et al.* Molecular Determinants of PQBP1 Binding to the HIV-1 Capsid Lattice. *J. Mol. Biol.* **436**, 168409 (2024).
144. Hoeffler, J. P., Meyer, T. E., Yun, Y., Jameson, J. L. & Habener, J. F. Cyclic AMP-Responsive DNA-Binding Protein: Structure Based on a Cloned Placental cDNA. *Science* **242**, 1430–1433 (1988).
145. Wu, H. Higher-Order Assemblies in a New Paradigm of Signal Transduction. *Cell* **153**, 287–292 (2013).
146. Summers, B. J. *et al.* Modular HIV-1 Capsid Assemblies Reveal Diverse Host-Capsid Recognition Mechanisms. *Cell Host Microbe* **26**, 203-216.e6 (2019).

147. Mattei, S., Glass, B., Hagen, W. J. H., Kräusslich, H.-G. & Briggs, J. A. G. The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science* **354**, 1434–1437 (2016).
148. Yant, S. R. *et al.* A highly potent long-acting small-molecule HIV-1 capsid inhibitor with efficacy in a humanized mouse model. *Nat. Med.* **25**, 1377–1384 (2019).
149. Zhang, Z. *et al.* T = 4 Icosahedral HIV-1 Capsid As an Immunogenic Vector for HIV-1 V3 Loop Epitope Display. *Viruses* **10**, 667 (2018).
150. Lau, D. *et al.* Self-Assembly of Fluorescent HIV Capsid Spheres for Detection of Capsid Binders. *Langmuir* **36**, 3624–3632 (2020).
151. Link, J. O. *et al.* Clinical targeting of HIV capsid protein with a long-acting small molecule. *Nature* **584**, 614–618 (2020).
152. Bester, S. M. *et al.* Structural and mechanistic bases for a potent HIV-1 capsid inhibitor. *Science* **370**, 360–364 (2020).
153. Singh, K. *et al.* GS-CA Compounds: First-In-Class HIV-1 Capsid Inhibitors Covering Multiple Grounds. *Front. Microbiol.* **10**, 1227 (2019).
154. Highland, C. M., Tan, A., Ricaña, C. L., Briggs, J. A. G. & Dick, R. A. Structural insights into HIV-1 polyanion-dependent capsid lattice formation revealed by single particle cryo-EM. *Proc. Natl. Acad. Sci.* **120**, e2220545120 (2023).
155. Stacey, J. C. V. *et al.* Two structural switches in HIV-1 capsid regulate capsid curvature and host factor binding. *Proc. Natl. Acad. Sci.* **120**, e2220557120 (2023).
156. Ganser-Pornillos, B. K., von Schwedler, U. K., Stray, K. M., Aiken, C. & Sundquist, W. I. Assembly Properties of the Human Immunodeficiency Virus Type 1 CA Protein. *J. Virol.* **78**, 2545–2552 (2004).
157. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
158. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).

159. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
160. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).