THE UNIVERSITY OF CHICAGO

STRUCTURAL INSIGHTS INTO INTERACTIONS BETWEEN THE VAULT CAGE AND COMPONENTS OF AN ADP-RIBOSYLTRANSFERASE SYSTEM

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOCHEMISTRY AND MOLECULAR BIOPHYSICS

BY

JANE ELIZABETH LODWICK

CHICAGO, ILLINOIS

DECEMBER 2023

© Copyright 2023 by Jane Elizabeth Lodwick

All Rights Reserved

For my parents, Sheila & David

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my thesis advisor, Minglei Zhao. His obligations during my time in the lab came with enormous pressure. Despite his daunting remit to simultaneously teach, build a research program, and bring the BMB Department into the cryo-EM age, however, he never failed to treat his trainees with kindness. I am grateful for the compassion and understanding he showed me while I developed a project trajectory, and I have benefitted immensely from his wealth of structural biology expertise and personal investment in my collaboration projects. Finally, I am very lucky that he provided me with the opportunity to work on such an exciting and beautiful molecular system.

I appreciate the scientific and emotional support that Zhao group members offered me during my time in the lab. Thanks to Yuan Xie for providing me with plasmids and viruses that accelerated the pace of my research, to Man Pan for bringing considerable knowledge and skill to our native vault collaboration, and to Yuanyuan Yu for establishing the lab's coimmunoprecipitation-based proteomics workflow, providing a valuable template for other lab members to follow.

My warm thanks to Chang Liu for her ongoing support and understanding as the other early graduate student in the lab. Her ability to consistently model positivity and optimism, even during the worst phases of the COVID-19 pandemic, inspired me to work through some of the most difficult periods of my PhD. I had the privilege of teaching Karolina Roganowicz and Kasandra Medrano, both of whom brought fresh energy and curiosity to the vault project. I thank Karolina, in particular, for her year of tireless work in the lab and her patience with me as I learned how to teach.

I am grateful for the opportunity to have worked with Bowei Yang, who strove to offer help and support to his labmates, even as he was experiencing profound personal difficulty and pain. He offered me encouragement and affirmation that I will strive to live up to going forward.

Many thanks to Bixia Zhang and Erick Torres Guasp for creating such a vibrant and welcoming lab environment. I am grateful to Bixia for her perceptive insights as a scientist, labmate, and friend and for going out of her way to encourage and help me. I am thankful to Erick for bringing his enthusiasm for research, as well as his humor, exceptional kindness, and Puerto Rican sesame candies, to the lab.

I have benefitted from the assistance of some incredible collaborators, including Satchal Erramilli and Amanda DiGiulio, without whom my cell biology work would not have been possible. Thanks to Rong Shen for brightening my days with his friendly greetings and for his MD simulation data, which has both validated and broadened the horizons of my work so far.

I am exceedingly grateful to the staff at the University of Chicago Advanced Electron Microscopy Facility, especially Jotham Austin and Tera Lavoie, who provided training and technical support to users at all hours while overseeing the growth of the core into a world-class facility. Thanks to Hyun Lee and her post-doctoral fellows at the University of Illinois-Chicago Biophysics Core for their hard work, patience, and willingness to troubleshoot in order to generate essential biochemical data for my project. I am also thankful to Allen Huff at the University of Chicago Proteomics Platform Facility and Lasanthi Jayathilaka at the University of Illinois-Chicago Mass Spectrometry Core for their mass spectrometry data generation and analysis.

Many thanks to the BMB administrative office for keeping me on track to finish my degree, supporting me through the F31 grant application process, and facilitating talks and

activities for the BMB community. I am especially thankful to Lisa Anderson for providing me with invaluable advice and support during my first year and to Shani Charles for guiding me through the logistics of my last.

Thank you to the members of my thesis committee, Erin Adams, Engin Özkan, and my chair, Joseph Piccirilli, for providing me with encouragement and technical expertise that proved critical to the project. I am grateful for their investment in both my scientific and personal success. The supportive atmosphere they cultivated during my committee meetings made me feel comfortable enough to honestly convey whatever challenges my project was presenting and to try out new hypotheses and ideas without fear of judgement.

I am incredibly lucky to have had wonderful scientific mentors going back to junior high school, when Mrs. Moinihan allowed me to participate in a collaborative environmental research project and guided me through my first science fairs. Mrs. Santo and Mrs. Beigi continued to foster my passion for science in high school, ultimately giving me the confidence to pursue a collegiate chemistry degree. Many thanks to my undergraduate advisor, Don Elmore, for introducing me to biochemistry research, encouraging my aspirations to attain a PhD, and expressing confidence in my abilities as a scientist when I was most unsure of them. Thank you also to Julia Klaips for her support during this time.

I am grateful to Ronen Marmorstein for providing me with additional research and lab management experience in an R1 lab, which was critical to my preparation for graduate school. Thank you to the postdoc who mentored me there, John Domsic, for teaching me how to work in a protein biochemistry lab and how to train students with humor and humility. I am incredibly thankful for the friendships I forged with other Marmorstein lab members. Mike Grasso, Gleb Bazilevsky, Yadilette Rivera-Colon, Robert Magin, and Emily Mattes made my time as a lab tech

some of the most socially and scientifically rewarding of my academic life, and they continue to give me invaluable support and advice to this day.

Thank you to the friends I have made at the University of Chicago. I want to acknowledge the women of my BMB cohort: Soo Ji Kim, Cole Ladd, Caraline Sepich-Poore, and Yujiao Wu. I have consistently been inspired by the sense of perspective Soo Ji brings to her work and life, and I have benefitted immensely from her advice and gift for listening. I am so grateful for the support I have received from Cole during our time in the program, as well as the opportunity to bond over music, memes, and our favorite drag queens. I have both learned from our outward differences and felt a deep sense of recognition from our shared emotional understanding. I am thankful to Chris Katanski for his reassurance and feedback when I was planning key phases of my project and for continuing to encourage me as I was planning my defense. Thank you to Phil Ross for doing so much to mitigate my stress and increase my confidence in the final stretch of my graduate career. His frequent check-ins and our conversations about the future were critical to keeping me focused and optimistic while I worked towards graduating. Thank you to all of my close friends for watching my wonderful kitty, Harold, through the years.

I have benefitted so much from the active involvement of extended family in my life. My grandparents, Lillian & Paul Lodwick and Herbert & Mary Alice Ramerman, as well as my aunts and uncles, especially Eileen Ramerman, have shown me more support than I could have asked for. Knowing that they loved and were proud of me, at every stage of my life, was incredibly motivating as I worked towards this goal.

My sister, Mary, has been an incredible role model for succeeding in a challenging and specialized career path. I am equally inspired by her matchless sensitivity and caring, with which

she has shown me extraordinary support. It has been wonderful to spend so much time with her in Chicago over the past couple of years, but, wherever we are, I know that I will always be able to rely on and treasure seeing her.

Finally, I am inexpressibly grateful to my parents, Sheila and David, for working so hard to provide me with every opportunity for happiness and success. They completely devoted themselves to my sister and I when we were growing up: cheering us on at our extracurricular activities, working to ensure that we could have the best educations possible, and responding to our little triumphs and failures with compassion, encouragement, and unconditional love. I am thankful that they passed their passion for learning along to us and that they encouraged us to pursue work that we found interesting and meaningful. They nurtured my interest in science in any way they could, guiding me through experiments at home, connecting me with educators who could help me further, and always expressing their confidence in my abilities. They also helped me in innumerable material ways, by assisting me in my moves to research institutions across the country, sending me fortifying care packages, and welcoming me home for breaks over the holidays. I always felt that I could challenge myself and chase my loftiest goals because I would have their steadfast support, and I owe them everything.

ABSTRACT

Among the most intriguing mysteries in contemporary biology is the functional mechanism of the vault complex. Renowned for its immense size and arched structure resembling that of "cathedral vaults," it is found across tissue types, cellular compartments, and eukaryotic species, with a patchy pattern of conservation that stretches from slime molds to humans. Although work since vault's initial discovery in 1986 has failed to define its precise function, studies of the vault's primary component – the major vault protein (MVP) – have provided some clues. Research observing MVP knockout mice and condition-dependent changes in MVP expression has demonstrated vault's ability to promote cell survival following exposure to stress. Structures of the giant oligomeric MVP "cage," which comprises most of vault's mass, have spurred hypotheses that it acts as a molecular "cargo" transport module and/or signaling scaffold. This supposition has been bolstered by reports of vault's ability to rapidly relocalize to different cellular compartments following cell perturbation, as well as observations of heterogeneous density within its Fabergé egg-like cage. Lacking any enzymatic activity of its own, the vault cage most likely acts as a relatively inert vessel. We hypothesized that regulation by additional molecules is critical to MVP's ability to accept transiently associated cargo, move between organelles, and otherwise mediate cell signaling. Some obvious candidates to fill these regulatory roles are vault's understudied "minor" components: the enzyme poly (ADP-ribose) polymerase 4 (PARP4), the NOD-like protein telomerase component 1 (TEP1), and short non-coding transcripts known as vault RNAs (vtRNAs), all of which reside within the MVP cage. My research has therefore sought to explore the roles of vault's minor components and expand our understanding of vault as a multispecies complex. Due to its established enzymatic activity, I decided to begin by focusing on PARP4, an ADP-ribosyltransferase that consumes the metabolite

NAD^+ to deposit post-translational modifications (PTMs) onto its target substrates. In Chapter 2, I report our novel complex structures, which resolve the interface between MVP and PARP4, as well as a surprising interaction between MVP and NAD^+. Accounts of my initial efforts to study the cellular function of vault-bound PARP4 are described in Chapter 3. I summarize my thesis work and its import in Chapter 4. Additional data supporting the findings listed in Chapter 2 and preliminary work to elucidate the structure of the regions of PARP4 that do not interact with vault can be found in Appendices A and B, respectively.

CHAPTER 1
INTRODUCTION

1.1 Structural Overview of the Vault Cage

The "vault cage" is a massive, cavernous shell composed entirely of homo-oligomeric copies of major vault protein (MVP). This cage houses the minor vault components: the proteins telomerase component 1 (TEP1) and poly (ADP-ribose) polymerase 4 (PARP4), as well as a short non-coding RNA dubbed vault RNA (vtRNA). Together with MVP, these macromolecules constitute the "vault complex." The basic structural features of the cage have been known since vault's discovery. After serendipitously isolating it from a gradient purification of large cytosolic complexes, the researchers who discovered vault visualized it using negative stain electron microscopy. Though somewhat distorted by the uranyl acetate stain that coated it, the intact vault cage's unique, dense, ovular morphology (Figure 1.1b) was immediately apparent. Vault's size (about 40 x 70 nm at its widest points) and the fact that a large proportion of its mass was attributable to a single protein species, initially led the researchers to believe that it was a coated vesicle. After a failed attempt to detect phospholipid within the cage, however, they confirmed that the vault complex was solely composed of protein and vtRNA [7]. Vault's relative compositional homogeneity was, for a free-standing, cytosolic complex of its size, unprecedented. This, as well as its unique structure, immediately marked vault as a singular cellular entity. It would take several more years to establish that MVP was the sole component of vault's cage-like exterior shell and that each cage was comprised of two symmetrical halves. This, in turn, led to the discovery that the cage had the capacity to open at the medial interface between the halves. That finding prompted speculation that vault might engulf "cargo" molecules

to transport them between cells and organelles [8] and spurred efforts to capture the structural details of the cage at high resolution.

While early work established that numerous copies of MVP existed in each cage half, the precise stoichiometry was not settled until the crystal structure of the rat homolog was solved in 2008 [9]. (Because the rat and human MVP protein share over 90% sequence identity, the following conclusions can be broadly applied to both species.) In addition to establishing that the vault cage was a dimer of MVP 39mers, the structure's high resolution (3.5 Å) allowed for speculation about the dynamics of the cage's two halves. The study also introduced a corporeal nomenclature (stretching from MVP's C-terminus at the vault ends to the N-terminus in the cage's center) to refer to different regions of the cage (Figure 1.1b).

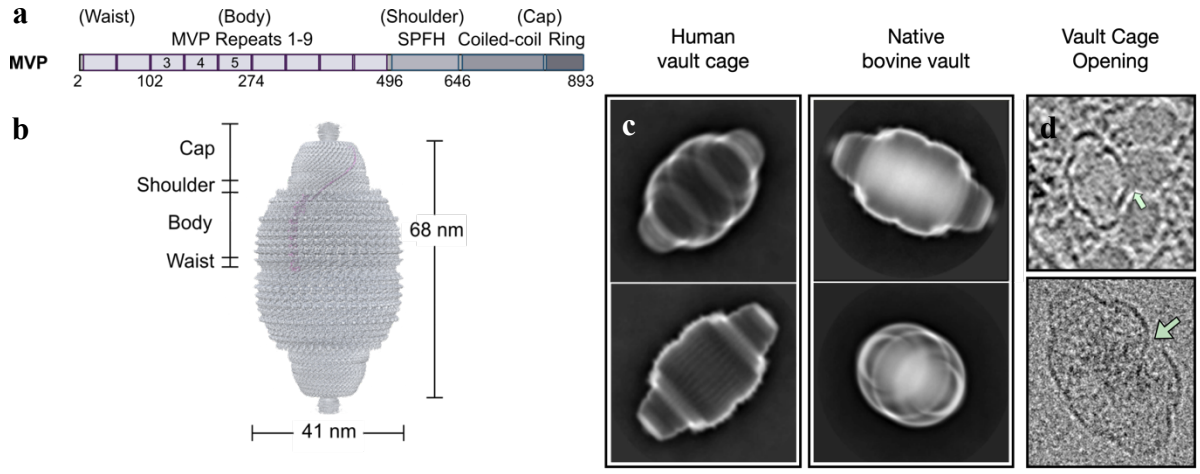


Figure 1.1 Vault cage nomenclature and evidence of molecular uptake

a, Domain diagram of the major vault protein. **b**, Labeled regions of the vault cage. **c**, 2D class averages of the MVP cage alone (left), and of the native vault complex purified from bull liver tissue (right), highlighting the difference in the internal occupancy of native vaults. **d**, Cryo-electron tomography image of vault in a human cell slice, opening at the waist to interact with (and possibly take up the components) of a cytosolic granule adapted from [5] (upper). Single particle cryo-electron microscopy image of a vault particle purified from vault liver with internal molecular cargo and an opening at its waist (lower).

Each vault "cap," comprising the tapered ends of the cage, consists of MVP's tightly packed coiled-coil domains, which engage in lateral hydrophobic interactions with neighboring monomers. Atomic force microscopy (AFM) experiments have confirmed that the energetically favorable packing of these coiled-coils promotes the stability of each vault half. This stability permits domains closer to the center of the vault to exhibit more dynamic behavior without wrenching the particle apart [10]. The "shoulder" region, just downstream of the cap, also likely contributes to MVP multimerization via its Stomatin/Prohibitin/Flotillin/HflKC-like (SPFH) domain. SPFH family members evolved convergently, appear throughout the tree of life, and readily oligomerize, often while localized to lipid microdomains [11] or other sterol-rich membranes in the cell [12]. Other than a unique beta-loop-beta motif (Figure 2.5c) that appears in the middle of its SPFH domain (which will be referred to here as the "keyhole loop"), MVP's shoulder region shares broad structural homology with other SPFH proteins.

Similar to MVP, the bacterial HflK/C proteins (which serve to sequester membrane proteins and inhibit their downstream activity), exhibit radially symmetrical coiled-coils upstream of their SPFH domains [13]. These striking structural similarities can provide some potential clues about MVP's evolutionary history. However, in the space where HflK/C proteins encode for transmembrane domains to anchor them around their membrane-bound sequestration targets, MVP instead contains soluble beta-rich domains. These nine eponymous repeat domains constitute the cage "body" and engage in relatively weak polar, intermolecular interactions with adjacent MVP molecules, endowing this region of the vault cage with greater flexibility than the cap region. Adjacent copies of MVP in this region have also been observed to "self-heal" transient fractures following pressure exerted by AFM tips. Interestingly, these AFM mechanical stress studies revealed that the vault cage has a much lower threshold for breakage but a higher

capacity for self-repair than similar protein capsid structures, indicating that it may be optimized to tolerate frequent conformational disruption [10].

The "waist," or interface between the vault cage halves, is the component of the cage most suggestive of vault's potential role as a macromolecular ferry that internalizes its cargo. The first vault cage crystal structure suggested that conserved residues at this site engaged in two salt bridge interactions per mirrored MVP dimer, for a total of 78 around the circumference of the cage. Several subsequent studies (including ours) have refuted the presence of these particular salt bridges in the vault solution structure. However, the integrity of the interaction between the halves has been shown to be susceptible to disruption from protonation in low pH environments [14], suggesting that electrostatic interactions may play an important role in cage stability. I will discuss our findings regarding inter-vault half interactions in Chapter 2.

Disruptions between vault halves can sometimes provoke the particle's complete cleavage [15]. Accordingly, previous studies indicate that vault particles can exchange halves *in vivo* and *in vitro* [16], offering one potential mechanism for the uptake of cargo. Conversely, some cryo-electron tomography (cryo-ET) images suggest that separation between the halves can remain more localized [5], permitting the entry of cargo into the cage without breaking it apart completely (Figure 1.1c). This mechanism seems somewhat more consistent with the observed dynamics of the vault cage body. The exceptional structure of the vault cage has served as a powerful hypothesis generator for vault function. Unsurprisingly, its capacious surface and vast lumenal cavity has encouraged numerous proposals that it may act to sequester, transport, or increase the local concentration of other molecules in the cell, suppositions further supported by the discovery of heterogeneous density within natively purified vault molecules (Figure 1.1c). Importantly, many researchers have also drawn insights about vault's potential function by

mining the sequence and structural similarities that vault components share with other proteins throughout the tree of life. I will discuss these potential functions in more detail in Chapter 1.2.1.

1.2 Functional Overview of the Vault Complex

A few years after vault's discovery, upregulated expression of MVP was found to be correlated with chemotherapeutic drug resistance in non-small cell lung cancer [17], heralding a potential breakthrough in vault biology. Thus began decades of research attempting to uncover the mechanistic underpinnings of vault's role in chemoresistance. As it stands today, researchers still have not uncovered a causal mechanism for this relationship. This frustrated effort to elucidate vault function is emblematic of much of the vault research conducted since. The literature is largely populated by papers in which MVP has been identified as a player in some kind of cellular stress response via a genetic screen or immunoprecipitation assay, then follow-ups in which researchers struggle to validate or explain their initial observation [17-20]. As such, we still have not arrived at a comprehensive understanding of vault function after many years of study. Throughout this period, methods in vault biology primarily centered around observing or manipulating MVP expression and assessing the effects of that expression in cells and organisms. Although this approach has not been sufficient to fully elucidate vault's mechanism, it has uncovered much of what we do know about vault's general cellular role to this point. As such, I will review hypotheses about vault function derived from the differential tissue expression of its component proteins and their phylogenetic distribution, as well as research evaluating the impact of MVP's depletion at both the cellular and organismal levels below.

1.2.1 Vault Tissue Specificity and Taxonomy

RNA sequencing and immunohistochemistry data indicate that MVP is constitutively expressed across tissue types, but that it is upregulated in tissues frequently exposed to foreign molecules, including pathogens. These tissues include the lung and gut epithelia as well as "scavenging" cells in the lining of other organs, such as glandular cells in bladder tissues and microglia in the brain [21]. Similarly, it is abundant in white blood cells as well as other immune tissues like the spleen and appendix, which filter pathogens from blood and regulate gastrointestinal bacteria, respectively. While PARP4 and TEP1 are expressed at lower levels overall (dampening the differences in their tissue specificity), the general pattern of their expression is similar to MVP's. Namely, both still show higher than average expression in the lung, gut, bladder, lymphoid tissues, and white blood cells, bolstering the supposition that vault is important for the clearance of pathogens and xenobiotics (Figure 1.2a).

Vault's presence in these tissues has informed interpretations of phylogenetic analyses examining the conservation of its components. In addition to considering sequence similarities between potential vault component orthologs, one ancestral reconstruction of MVP evaluated whether those conserved sequences would be sufficient to form the vault cage structure. By those criteria, it indicated that vaults form in species throughout eukaryota and most likely existed in the last eukaryotic common ancestor [4], while also showing that vault has been lost in a seemingly arbitrary combination of groups, including fungi, insects, and roundworms (Figure 1.2b). Additionally, whether MVP is present in plants is controversial; homologs have been identified in very few species (and are missing from model organisms like *A. thaliana*) and those purported to exist have been suspected to be possible contaminants within the plant samples. One explanation the authors of this study proposed for vault's loss (or likely loss) in the above species

involves differences in the organisms' propensity for engulfing bacteria: while plants and fungi do not typically engage in bacterial consumption, insects (which do so frequently) have developed other mechanisms for protecting themselves from pathogens. This hypothesis arose partly from MVP's differential tissue expression and partly from the results of a study (that I will describe further in Chapter 1.1.2) demonstrating vault's role in epithelial internalization of pathogenic bacteria.

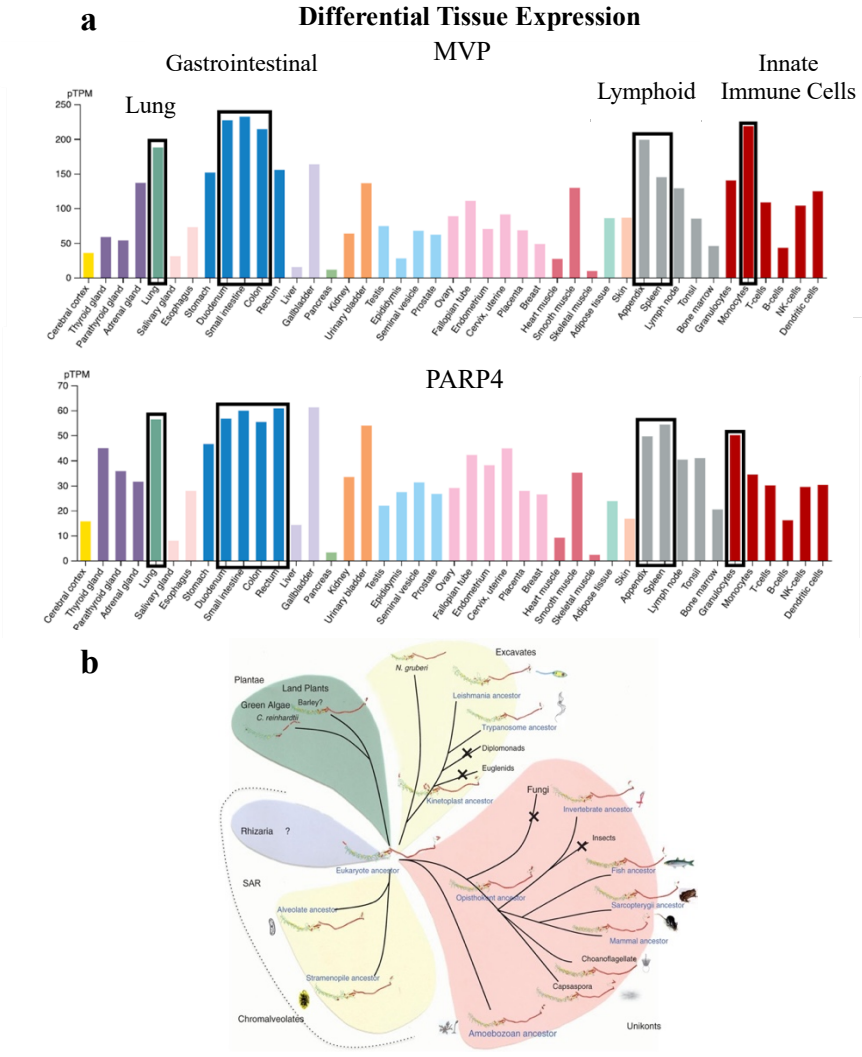


Figure 1.2 MVP Expression in human tissues and throughout eukaryota

a, Figure adapted from the Human Proteome Database [1] showing differential tissue expression patterns of MVP (upper) and PARP4 (lower) as measured by RNA-seq. **b**, Figure adapted from [4], showing MVP expression throughout eukaryotes.

Vault's abundance in tissues with particularly prevalent bacterial exposure has invited some interesting speculation that MVP genes may have originated in eukaryotes via horizontal gene transfer from bacterial species. Comparisons of MVP to bacterial proteins frequently focus on similarities between the structure of MVP and those of some multimeric bacterial proteins. In addition to MVP's general structural resemblance to bacterial HflK/C proteins in its cap and shoulder regions, one earlier study noted that MVP's repeat domains exhibited striking resemblance to repeat domains in bacterial toxic anion resistance (TelA) proteins. In addition to highly similar (predicted) beta-rich structures, the repeat domains shared >30% sequence identity with one another. TelA proteins engage in thiol:redox reductions to neutralize toxic anions bacteria encounter in the soil. As such, immunofluorescence studies in cells were conducted to observe the effect of toxic anions on MVP behavior, and authors found that MVP readily formed puncta near the cell membrane in response [22].

More recently, MVP homologs have been identified in additional bacterial species, as first compiled by Sokolskyi et al. Although the work in question has only been submitted to bioarχiv, it did correctly highlight the existence of cyanobacterial homologs (acquired by whole genome shotgun sequencing) of MVP [23]. These sequences share >50% sequence identity, as well as strong predicted structural homology to MVP, though their predicted structures seem to lack the proper placement of residues for interactions between vault halves and (therefore) formation of a full cage. Interestingly, the identified bacterial sequence homologs exhibit nucleotide-binding function, which is mediated by their shoulder domains. In Chapter 2, I will describe a nucleotide binding pocket we discovered in the course of our cryogenic electron microscopy (cryo-EM) analyses of the MVP-PARP4 complex. In aggregate, these analyses

suggest that vault's ancestral function may have been to respond to pathogenic and xenobiotic entities that invade the cell and that similar functions may persist in mammals to this day.

1.2.2 in vivo *MVP Functional Studies*

The first mammalian MVP knockout (KO) phenotype was not observed until 2008. Knocking out MVP in mice (and thereby preventing vault formation) was shown to reduce mouse survival following bacterial infection relative to their wild type (WT) counterparts. Authors proposed that this could be attributed to vault's role ferrying protein components that stabilized lipid microdomains to the cell membrane, which in turn facilitated bacterial internalization by epithelial cells. Although the MVP-expressing epithelial cells did exhibit greater capacity to take up bacteria, the authors of the study failed to demonstrate any direct interactions between vault and the proteins that they claimed were its cargo [24]. By contrast, another study showing diminished survival of mice following viral infection, claimed that vault delivered pro-inflammatory transcription factors into the nucleus to promote cytokine secretion [25]. (I will note, however, that our lab was unable to recapitulate their findings showing that MVP migrates to the nucleus following stimulation with interferon-gamma in the same cell line).

Nonetheless, other cell-based studied investigating vault's immune function have also suggested that it mediates cell signaling pathways involved in inflammation [26-28]. Notably, these studies also indicated that vault operates (at least in part) by localizing to the cell membrane. In one, authors once again detected MVP in lipid rafts, this time following cell treatment with a bacterially-derived ligand [26]. In two others, researchers identified interactions between MVP and cell surface receptors involved in the detection of pathogenic molecules that alert the cell to the presence of an infectious agent [27, 28]. In composite, these studies indicate

that vault's ability to rapidly localize to a particular cellular compartment (often the cell membrane) is important for the execution of its function. They also demonstrate that increased MVP expression has been correlated with both enhanced and diminished cellular inflammation depending on cell type and perturbation, demonstrating vault's ability to respond differentially to dysregulated inflammatory signals in different environments.

Vault's role in mediating inflammation has also been proposed to operate in systems outside of the immune response. The third major study examining the impact of MVP KO in mice identified a role for vault in metabolic regulation. The authors generated two mouse lines: one global mouse knockout (MVP^{KO}) and one in which MVP was specifically knocked out in the macrophages of the mice (MVP^{MacKO}). The study noted that MVP^{KO} and MVP^{MacKO} mice fed a high-fat diet were more susceptible to obesity and dysregulated glycolipid and insulin metabolism than WT animals fed the same food [29]. Authors noted upregulation of both fatty acid synthases and pro-inflammatory cytokines and traced those events back to MVP's regulation of the NF-kB pathway. This study explicitly linked metabolic disease to vault's role in mediating inflammation, suggesting that vault may be able to sense and modulate its activity in response to the presence of metabolic stress. In addition to MVP's putative capacity to bind metabolites, the presence of vault's nicotinamide adenine dinucleotide (NAD^+)-dependent enzyme PARP4 also presents an intriguing link between vault activity and metabolic regulation. PARP4 is part of a fascinating superfamily of enzymes that are highly responsive to cellular changes in nutrient availability and to markers of infection and inflammation, suggesting that PARP4 may play an important role in mammalian vault function. I will discuss this enzyme family, as well as its dependence and effect on cellular NAD^+ stores, in greater detail below.

1.3 Poly (ADP-ribose) Polymerase (PARP) Enzymes and NAD^+ Metabolism

1.3.1 The PARP Superfamily of Proteins

Humans express 17 PARP proteins, all of which encode for an ADP-ribosyltransferase (ART) domain that shares structural homology with the ancestral "diphtheria toxin" ADP-ribosyltransferase [6]. (Consequently, each PARP may be referred to by its more biochemically accurate moniker "ART – Diphtheria toxin-like" or "ARTD" [30].) In all but one case (in the inactive PARP13), the PARP's ART domain allows the enzyme to deposit an ADP-ribose (ADPr) post-translational modification (PTM) onto its target substrate, consuming the small molecule NAD^+ in the process. Specifically, it binds and positions NAD^+ in such a way that a nucleophile on the PARP's target can attack the N-glycosidic bond between NAD^+'s nicotinamide and ADP-ribose (ADPr) moieties. The nicotinamide then leaves, and the ADPr group is transferred to the target [2]. Other domains in the PARP enzyme act to recruit these targets for modification. Paradigmatically, these "targeting" domains either promote direct binding to the enzyme's modification target, indirect binding to molecules proximal to the target, and/or localization to a compartment in which the target is enriched [6]. Once within striking distance of the PARP, the target can be modified in a number of ways. Originally, PARPs were thought to primarily modify acidic protein residues, while only occasionally targeting cysteine, asparagine, and basic amino acids [31]. More recently, however, PARP-derived modifications of all of the polar and charged amino acids have been documented [32, 33], with serine emerging as the most prevalent mammalian ADPr recipient [34]. In addition, several PARPs have even been shown to ADP-ribosylate (ADPrylate) the ends of DNA *in vitro* [35], as well as RNA, both *in vitro* [36] and *in vivo* [37]. Within the PARP family, four enzymes exhibit the ability to deposit modifications with long, often branched chains of ADPr known as poly (ADP-ribose) (PAR), while the remainder

modify substrates with single ADP-ribose groups, known as mono (ADP-ribose) (MAR). As a result, this latter class of enzymes have recently begun to be referred to as "mono (ADP-ribosyl) transferases," or "MARTs."

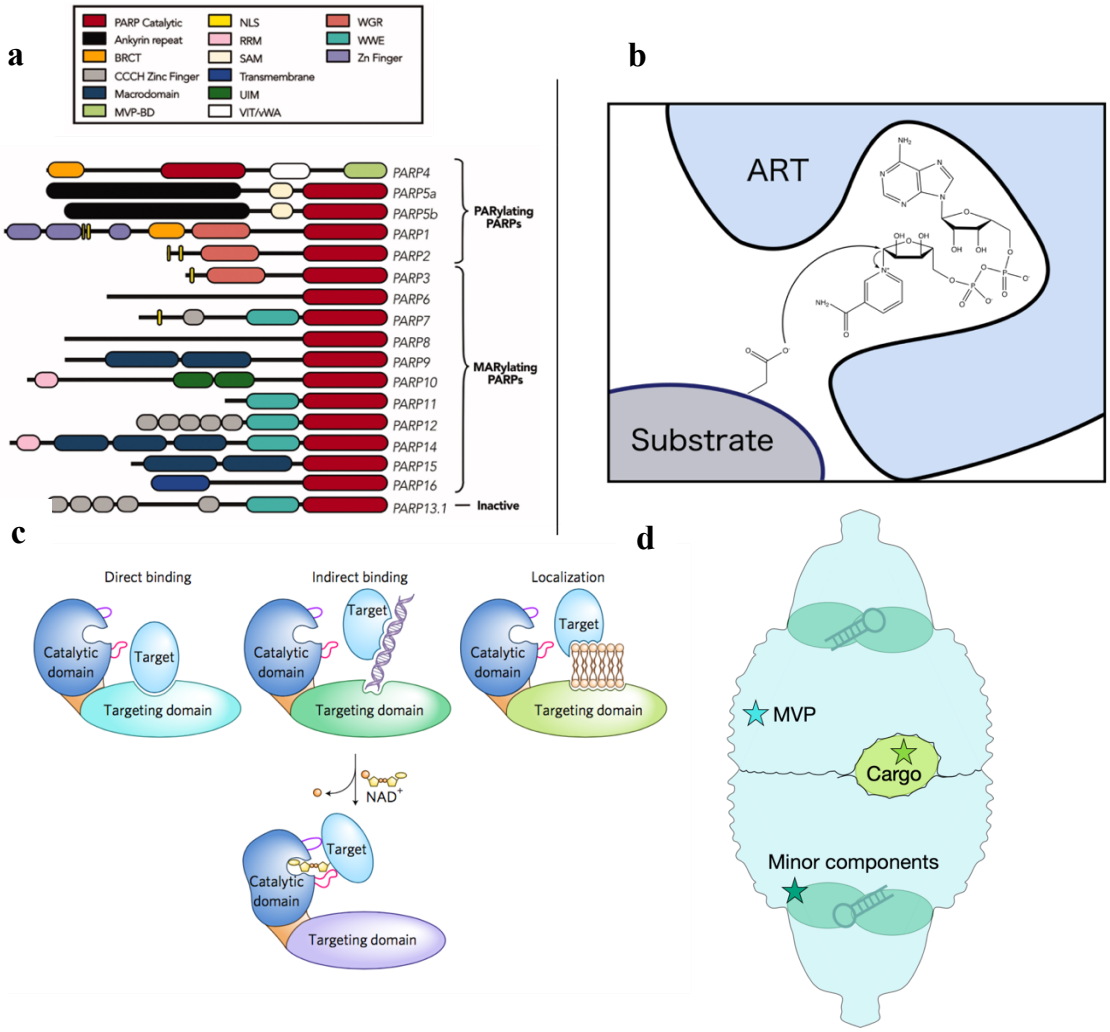


Figure 1.3 PARP family enzyme activity

a, Figure adapted from [2] showing domain diagrams of members of the PARP family. PARP4's MINT domain is colored in light green at the top. **b**, ADPrylation mechanism of the PARP family ART domain. **c**, Figure adapted from [6] showing models for target recruitment to PARP enzymes. **d**, Potential targets of vault-bound PARP4's enzymatic activity according to the models specified in c. PARP4 may function by ADPrylating its direct interaction partner (MVP), molecules proximal to MVP (other internal minor components), or molecules that localize to the vault lumen (transient cargo).

ADPr modifications are deposited in the service of a number of molecular functions, including disrupting macromolecular complexes [38], inducing conformational changes in modified proteins or their interacting partners [39], stabilizing certain molecules [37] or marking others for degradation [40], and regulating recruitment of molecules to certain cellular sites [41-43]. ADPr modifications often engage in "crosstalk" with other PTMs to mediate a variety of signaling pathways [34, 41, 44]. Unsurprisingly, given the breadth of potential PARP targets and the functional versatility of their PTMs, these enzymes have been implicated in a diverse collection of biological processes. Though best known for their role in DNA repair, PARPs have also been shown to regulate transcription [45], translation [46], mitosis, cell migration, cytoskeletal growth [47], the stress response [48], inflammation [41], and programmed cell death [49]. Along similar lines, PARPs have long loomed large as an inhibition target for cancer, initially due to their DNA repair activity [50]. More recently, additional PARP activities that promote cancer cell proliferation, including maintaining cancer cell proteostasis [46] and promoting rapid cytoskeletal assembly and turnover [51], have been recognized as well. In light of recent events, the regulatory roles of MARTs in immunity, particularly the antiviral response, have generated particular interest [52]. Intriguingly, many viruses express ADP-ribosylhydrolases ("deMARylating") enzymes that are critical to their success [53-55]. In addition to dysregulating the specific function of the macromolecules they deMARylate, these viral enzymes also induce the expression of MARTs and force the cell to deplete its stores of NAD^+ [56]. This clever viral mechanism underscores the importance of maintaining adequate concentrations of NAD^+ throughout the cell, as well as maintaining the delicate balance that exists between compartments. Because the K_m of most NAD^+-consuming enzymes lies above the

metabolite's equilibrium concentration, NAD^+ production and availability is a key regulator of PARP activity [57].

1.3.2 NAD^+ Biosynthesis and Regulation

In addition to its role as an enzymatic substrate for PARP enzymes, NAD^+ is a critical coenzyme for proteins involved in cellular energy production. The NAD redox couple – the reduced (NADH) and oxidized (NAD^+) forms – donate and accept electrons to and from metabolic enzymes and metabolites over the course of glycolysis and oxidative phosphorylation (OXPHOS). These are the processes by which the cell generates adenosine triphosphate (ATP), a proxy for cellular energy which powers a vast number of essential enzymatic reactions in the cell. Between its prodigious consumption by enzymes and its vital role in maintaining energetic homeostasis, NAD^+'s biosynthesis is tightly regulated. It occurs through three main routes: the Kynurenine, Preiss-Handler (P-H), and salvage pathways [3]. The two former pathways transform dietary metabolites (tryptophan and nicotinic acid (NA), respectively) into NAD^+. Tryptophan's relatively low abundance in the cell places the onus of most of the NAD^+ production onto the P-H and salvage pathways. Consequently, adequate expression of the key enzymes in those pathways is extremely important to the health of the cell.

Nicotinamide phosphoribosyltransferase (NAMPT) is the rate-limiting enzyme in the salvage pathway, which recycles the byproducts of NAD^+-consuming reactions to resynthesize NAD^+. This highly efficient process is responsible for producing most of the cell's NAD^+. Both the P-H and salvage pathways also utilize nicotinamide mononucleotide adenylyl transferase (NMNAT). Three NMNAT isoforms exist in the cell, and each is associated with a particular

compartment. The nucleus houses NMNAT1, the cytosol and Golgi apparatus share NMNAT2, and the mitochondria (due to the necessity of maintaining high [NAD^+] for OXPHOS) has its

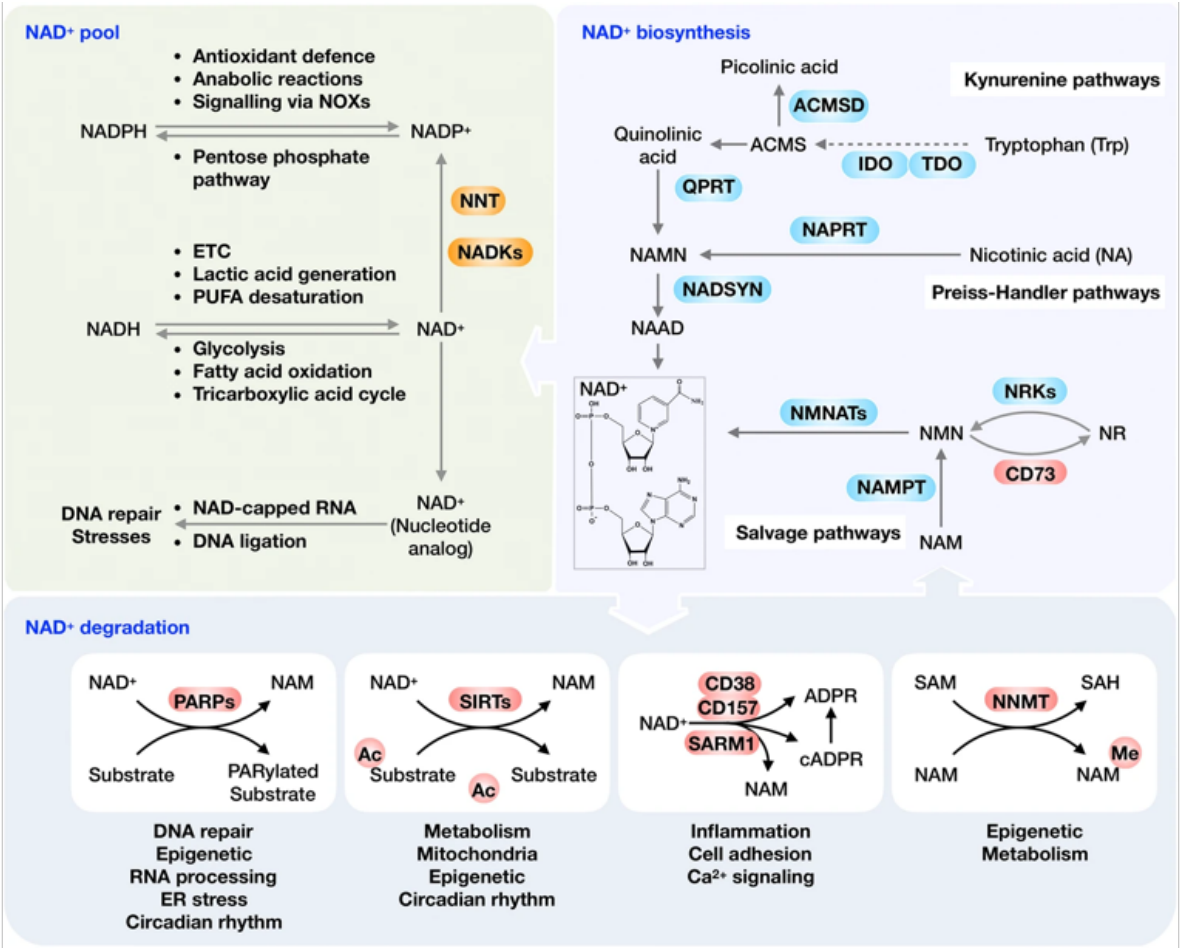


Figure 1.4 NAD^+ biosynthesis and depletion pathways

Figure adapted from [3], showing common cellular processes that deplete NAD^+ (upper left panel), the three NAD^+ biosynthesis pathways (upper right panel), and the enzyme families that consume NAD^+ to deposit PTMSs and generate cyclic ADPr (cADPR) for signaling, as well as those that preclude use of the NAD^+ precursor, NAM, for NAD^+ biosynthesis by methylating it (lower panel).

own dedicated isoform, NMNAT3. The microenvironment and metabolic status of the cell can impact differential expression of these enzymes and subsequently alter the net concentration and compartmental distribution of NAD^+ [58]. A number of triggers can induce NAMPT and NMNAT expression, including the presence of DNA damage agents [59], fasting [60], viral

infection [56], and the introduction of certain nutrients [58]. In addition, NAD^+ biosynthesis enzymes are frequently upregulated in cancer cells [61], and the NAD^+ concentration in immune cells is highly sensitive to molecular signals of infection and inflammation. Inflammatory astrocytes have demonstrated the capacity to induce NAD^+-biosynthesis enzymes and more than double NAD^+ levels over a span of six hours following the introduction of pathogen-associated molecular patterns (PAMPs) [62]. The abundant NAD^+ supports the OXPHOS-dependent polarization of the astrocytes into pro-inflammatory cells. Conversely, the detection of PAMPs in other cell types can rapidly diminish NAD^+ pools, instead of augmenting them [63]. By being attuned to these changes in NAD^+'s cellular concentration, NAD^+-dependent molecules can operate according to cell need. In Chapter 2, I will discuss my cryo-EM structure of NAD^+ bound to novel ligand binding pockets in the MVP cage. More details on the functional implications of the NAD^+-dependent PARP4 enzyme will be discussed in section 1.4.2.

1.3.3 ADP-Ribose- and NAD^+-Binding Proteins

Although some nucleotide-binding proteins possess fine-grained specificity for certain molecules (i.e., the ability to distinguish between PAR and MAR or NAD^+ and NADH), there is mounting evidence that many of them interact with multiple classes of nucleotides. PAR-binding proteins, in particular, tend to overlap significantly with RNA- and DNA-binding proteins. While, in some cases, this promiscuity is enabled by separate domains encoded by the protein, it is frequently the result of multipurpose modules such as RNA-recognition motifs, Tri-RGG domains, various helicase domains, OB folds, and the well-known PAR binders: WWE- and macrodomains [64]. Unsurprisingly, recent gene ontology analyses have implicated a number of PAR-binding proteins in DNA repair, but also in chromatin organization and RNA metabolism.

Others (especially those that preferentially interact with MAR and short PAR chains) are associated with pathways that regulate glycolipid metabolism [65, 66], inviting the possibility that they may also interact with small molecule metabolites directly. Relatedly, NUDIX enzymes specifically recognize the phosphodiester linkage of PAR chains (often prior to cleaving them) and can also interact with NAD^+ and other nucleotide small molecules [67]. Similarly, the NAD^+-binding modules in several proteins, including GAPDH and BRCT-containing proteins, have since been shown to interact with RNA [68] and DNA [69], respectively.

As discussed in Chapter 1.3.1, ADPr readers contribute to a wide range of biological processes. Outside of anchored ADPr modifications, proteins that associate with free ADPr in the form of cyclic ADPr and free PAR polymers, also play important cellular roles. Cyclic ADPr (cADPR) is a second messenger best known for its ability to regulate calcium signaling by interacting with relevant ion channels [70]. Parthanatos, a form of regulated cell death that occurs in the wake of PARP1 overactivation (itself the result of severe DNA damage), depends on the migration of PAR polymers to the cytosol. There, they trigger the release of apoptosis-inducing factor (AIF) from mitochondria [49] to the nucleus, likely by inducing a conformational change in AIF that degrades its affinity for the mitochondrial membrane [71]. AIF then forms a complex with several other proteins that induce cell death through mass chromatinolysis [72]. In addition, PAR has long been known to "seed" stress granules by recruiting critical masses of multivalent binders, often with RNA-interaction domains. In each case, although the half-life of cellular PAR is short (on the order of 20 minutes), its abundance in the cytosol has driven researchers to hypothesize that it may be operating as a free polymer, with added stability if it exhibits a large size and branched structure [73].

A recent pathway enrichment analysis of NAD^+-binding proteins indicated that the category primarily comprises metabolic redox enzymes (with dehydrogenases and reductases being heavily represented) and/or NAD^+-consuming enzymes (as discussed in 1.3.1). The overwhelming majority of proteins in the consensus dataset contained Rossman folds or Ankyrin repeats (the two best known NAD^+-binding domains) or various conserved NAD^+-binding domains characteristic to particular enzyme families. Few NAD^+-binding proteins exist outside of this primary group of redox and post-translational modification enzymes which directly consume the nucleotide. Those that do, however, tend to be solute carriers (SLCs), transcription factors (TFs), and proteins that do not exclusively bind NAD^+ [74].

NAD^+'s inability to efficiently diffuse through most cellular membranes necessitates that it be taken up into the cell and various organelles via specialized transporters. The NAD^+-binding transporters in the pathway enrichment dataset set tended to belong to the SLC family of proteins. The transcription regulators in the dataset were observed to undergo conformational changes following interactions with NAD^+, which served to regulate their interactions with target DNA [75, 76]. The BRCT domain in DNA Ligase IV has been shown to bind NAD^+ in order to transfer it to a neighboring adenylation domain in the enzyme when ATP is unavailable [69]. Though unusual, these cases point to comparatively uncommon functions for NAD^+-binding proteins, including transport, conformation regulation, and the storage and transfer of NAD^+ to other molecules and domains. In Chapter 2, I will discuss the potential ramifications of an apparent NAD^+-binding site within the MVP protein.

1.4 The Minor Vault Components

After their initial discovery of the vault cage, researchers classified three co-purified macromolecular species as “minor vault components.” In addition to vtRNA, researchers characterized two abundant protein species that were co-purified with the vault cage. The proteins, originally referred to as “p240” and “p193” were later identified as telomerase component 1 (TEP1) and poly (ADP-ribose) polymerase 4 (PARP4), respectively [77]. They were deemed specific, constitutive members of the complex that resided within the cage. They are not required for the construction or stability of the vault cage, suggesting that they hold other regulatory roles in vault complex activity. Full-length structures have not been solved for either protein, but their predicted domain architectures offer some clues to their potential function.

1.4.1 Telomerase Component 1 (TEP1) and vault RNA (vtRNA)

TEP1 is predicted to consist of a suite of domains common to NOD-like receptors (which self-associate in proinflammatory complexes and also contain nucleotide-binding sites), an RNA-binding domain (RBD), and a von Willebrand Factor A (VWA) domain (a module that frequently acts to mediate protein-protein interactions) [78]. Early electron microscopy experiments attributed some poorly resolved density in the vault cap to TEP1, but its localization within the cage is still being debated [79]. As its name suggests, TEP1 has also been observed to associate with the telomerase complex and bind telomeric RNA, but its role in telomerase activity is not clear [80]. Similarly, it binds vtRNA (and is essential for the RNA’s recruitment to the vault cage) [81], but neither the role of TEP1 nor vtRNA in vault’s function is understood. Interestingly, the predicted structure of TEP1’s RBD is homologous to a promiscuous small RNA-binding protein in the cytosol, whose association to its primary RNA binding partner (Y

RNA) mediates its interactions with other RNA species [82]. Analogously, TEP1 (with possible regulation by vtRNA) could conceivably recruit a repertoire of RNA species to vault, where they would be protected from degradation by the vault cage. Interestingly, vaults isolated from both rat and human tissues [83] have been co-purified with different RNA species. Outside of vault, vtRNA has been observed to regulate a host of processes including accelerated cell proliferation, anti-apoptotic signaling [84], and anti-autophagic activity [85]. Notably, it has also been observed to activate the innate immune response via RIG-I like receptor (RLR) binding following viral infection. Interestingly, its ability to bind RLRs rests on a triphosphate modification that alerts the cell to the infection-dependent down-regulation of short RNA processing enzymes [86]. Although only about 5% of vtRNAs are vault-bound, vault's potential to protect, sequester, or transport reserves of these RNAs may be a part of its function as well.

1.4.2 Poly (ADP-ribose) Polymerase 4 (PARP4)

The final minor component, PARP4, was determined to have enzymatic activity soon after vault was discovered [87], making it the fourth member of the PARP superfamily of proteins. Upstream of its ART domain, PARP4 is predicted to share several domains with other members of the PARP family. The first, BRCA1-C terminus (BRCT), is also found in the DNA repair enzyme PARP1. There, it has been proposed to support PARP1 function both through its recruitment of proteins with BRCT domains to sites of DNA damage [88] and its ability to associate with DNA via its phosphate-binding motif [89]. This motif has also been shown to bind ADPr modifications [90] as well as NAD^+. The second domain PARP4 has in common with other PARPs is its conserved "HD" (helical domain), which acts as an auto-inhibitory domain in

PARP1, limiting off-target ADPrylation [91]. In PARP4, however, it is critical to the stability of the ART domain, which is not soluble without it (our data).

PARP4's domain organization begins to diverge from that of other PARPs at its ART domain. While the ARTs in most PARP proteins are situated at their C-termini, PARP4's sits in the middle of its sequence (Figure 1.3a). This placement is the hallmark of a clade of PARPs found only in organisms that express major vault protein homologs (though there are some MVP-expressing organisms that do not express proteins from the PARP4 clade) [92]. PARPs from this clade are also notable for lacking many of the domains (ankyrins, macrodomains, etc.) used to classify other PARPs into groups [2], underscoring PARP4's unique status within the family. PARP4's ART active site is, itself, notable for the presence of a small threonine in its NAD^+ binding pocket, where the remaining PARPs have a comparatively bulky lysine, tyrosine, leucine, or isoleucine. This has allowed for the design of a PARP4-specific inhibitor that exploits this unique steric arrangement [93]. Beyond its ART domain, PARP4 contains vault-inter-alpha trypsin inhibitor (VIT) and VWA domains, both of which are known for mediating protein-protein interactions (often at the cell surface), inviting the possibility that PARP4 may use those domains to recruit or stabilize potential cargo within the vault cage. Additionally, the VWA domains shared between PARP4 and TEP1 may facilitate their binding within the vault.

PARP4 can also localize to the nucleus and the mitotic spindle in the absence of MVP [87]. It is therefore possible that some of PARP4's domains support cellular roles unrelated to its vault function or even that the enzyme's recruitment to the vault cage is meant to inhibit these activities, with the cage acting as a sequestration depot. Even though investigation into PARP4's cellular role has been limited, however, it has been shown to share several functions with MVP. One study in which the proteins were knocked down individually showed that depletion of each

protein contributed to dysregulation of the PI3K/AKT pathway by diminishing insulin-induced AKT phosphorylation [94]. In addition to showing functional cooperation between PARP4 and MVP, this study is striking for demonstrating vault's role in mediating a process that unifies metabolic regulation (cellular influx of insulin) and tumor development (overactive PI3K/AKT). The broad similarity of the tissue-specific expression profile between PARP4 and MVP also supports the notion that they work to regulate similar processes. Most convincing, at the end of PARP4's C-terminus (which also comprises a long, disordered loop region) there is a relatively short sequence required for PARP4's interaction with the vault cage, known as the MVP-INTeraction or MINT domain (also known as the MVP-BD in some papers) [95].

PARP4's MINT domain comprises a small, α-helix-rich fold that does not belong to an established domain family. However, a FoldSeek [96] search of similar structures to MINT's in the AlphaFold database, returned several hits to small helical domains at the C-termini of "VWA-containing protein 5B1," an extracellular protein of unclear function. Although the structures of the domains in question do not match MINT perfectly (and lack sequence homology with it), the broad similarities between the VWA-containing proteins and the region of PARP4 downstream of its catalytic domain are striking. Interestingly, data from our lab show that the proteins coimmunoprecipitated with MVP from the flat worm species *S. mediterranea* do not include a PARP4 homolog, but one of the most highly enriched proteins in the dataset also possesses a VWA domain. It is, therefore, intriguing to speculate about a possible path by which early vault complexes took up VWA-containing proteins (potentially for their ability to stabilize protein cargo within the cage), which later acquired PARP activity in more recently evolved species. In Chapter 2, I will describe our cryo-EM structure of the MVP-PARP4 interface and what it

reveals about the structural elements and individual residues essential for MVP's association with the MINT domain.

Our knowledge of the PARP family of proteins and the ways in which they are regulated has influenced a great deal of our lab's conception of vault-bound PARP4's potential role. Due to its observed capacity to auto-MARylate [47], PARP4 has been classified as a MART, though there has been some question as to whether it has the capacity to PARylate substrates as well. Strikingly, its active domain contains a "catalytic triad" motif only otherwise conserved in the four PARylating PARPs and which, based on available structural evidence, appears to be a driving force for their ability to stabilize growing ADPr chains [97]. Imagining a conditional difference in PARP4's ability to deposit MAR or PAR is plausible when one considers that interactions between PARPs and auxiliary proteins have been shown to influence the catalytic properties of those PARPs. These factors have, in various circumstances, conferred activity to an otherwise inactive PARP [98], induced PARPs to deposit shorter chains [99], and biased PARP activity towards modifying a particular type of amino acid [98, 100]. It is therefore possible that, in the context of the vault complex, PARP4 interacts with other factors that drive vault-specific modification behavior, including a conditional ability to PARylate or MARylate substrates.

Given the ability of PARP targeting domains to dictate which molecules the PARP modifies, one may infer that PARP4's MINT domain determines at least some of its targets. Under the direct/indirect/localization model (Figure 1.3c), PARP4 could conceivably ADPrylate MVP, other minor vault components, or cargo engulfed in the cage. Potential consequences of such modifications could include: influencing the dynamics of the vault cage or other complex components, regulating the stability of macromolecules in the vault, altering the repertoire of macromolecules recruited to the complex, and/or modifying cargo or transiently bound signaling

molecules to direct signaling pathways. In addition to using cell biology and mass spectrometry to begin to answer these questions, we will employ cryo-EM to begin to understand the impact of vault's ADP-ribosylation module in its dynamics and function as a multispecies complex.

Taken together, prior studies of MVP function and vault cage morphology depict a nanoparticle whose structure renders it amenable to, and even optimized for, dynamic interactions with transient binding partners, potentially engulfing them in the vault body. Physiological factors that may induce or exploit vault cage dynamics, however, have not been identified in these studies. In the cellular context, it seems likely that additional factors are required to regulate vault cage assembly, dynamics, localization, and condition-dependent interactions with various macromolecules. Though I will not address vault cage assembly here, my work will attempt to determine whether the role of one of the particle's minor components influences the latter three factors. Specifically, I will ask how PARP4's ability to deposit post-translational modifications (PTMs), its numerous protein-interaction domains (Figure 1.3a), and localization within the cage may influence vault dynamics or cargo recruitment. In addition, I will begin to try to uncover whether PARP4's ability to modify or interact with vault-associated molecules influences the activity of the particle. Furthermore, following a surprising result, I will attempt to characterize and seek to determine the functional implications of MVP's interaction with PARP4 ADP-ribosylation substrate, NAD^+. To do this, I will use a combination of structural biology, biochemistry (outlined in Chapter 2), and cell biology approaches (discussed in greater detail in Chapters 2 and 3).

CHAPTER 2

STRUCTURAL BASIS FOR THE RECRUITMENT OF PARP4 AND ITS SUBSTRATE NAD^+ TO THE VAULT CAGE

2.1 Foreword

The data and text in this chapter are to be reformatted into a manuscript for publication. Several authors contributed to the data generation and analysis, as well as written descriptions featured in the Methods section (2.6) of this chapter. Their contributions are delineated in 2.1.2.

2.1.1 Authors & Affiliations

Jane Lodwick[a], Rong Shen[a], Satchal Erramilli[a], Lasanthi Jayathilaka[b], Allen Huff[c], Anthony A. Kossiakoff[a], Minglei Zhao[a],

[a]Department of Biochemistry & Molecular Biology, The University of Chicago

[b]Mass Spectrometry Core, The University of Illinois-Chicago

[c]University of Chicago Proteomics Platform Facility The University of Chicago

2.1.2 Author Contributions

Study Design & Structural Data Analysis: J.L. and M.Z.

Protein Purification, Structural Data Production, Cell Culture, Biochemistry, & Co-IP: J.L.

Fab Selection & Evaluation: S.E. and A.A.K.

Molecular Dynamics Simulation Data Generation, Analysis, & Methods Description: R.S.

LC-MS Post-translational Modification (PTM) Detection Workflow & Methods Description: L.J.

Proteomics Sample Workflow & Analysis: A.H.

2.2 Abstract

Vault is a massive ribonucleoprotein complex present throughout eukaryotes. It contains 78 copies of the major vault protein (MVP), which oligomerize into an ovular cage, enveloping three constitutive minor vault components, as well as transiently bound "cargo" molecules. Vault mediates key biological functions including metabolic regulation [29, 94], innate immunity [24, 25], and the cell's response to disparate forms of stress [94, 101, 102], but its mode of action remains unclear. In this study, we shed light on vault's functional mechanism by using cryo-EM to resolve a structure of human MVP in complex with the understudied minor vault component, poly (ADP)-ribose polymerase 4 (PARP4). PARP4 is an ADP-ribosyltransferase that relies on the available pool of nicotinamide adenine dinucleotide (NAD^+) to deposit ADP-ribose moieties onto macromolecular targets [6]. Our sub-3.5 Å structures are the first to characterize the MVP-PARP4 binding interface to near-atomic resolution, and they also reveal a surprising interaction between NAD^+ and MVP. In addition to its direct interaction with MVP, mass spectrometry data indicate that NAD^+ enables PARP4-dependent ADP-ribosylation of MVP at a site near the medial interface of the cage. This modification could serve as a dynamic platform for the association of cargo molecules taken up into the vault lumen. Along these lines, we show proteomics data suggesting that vaults purified from WT cells interact with a different subset of molecules than those purified from PARP4-depleted cells. Together, our findings begin to illuminate vault cage regulation, and bring us closer to an understanding of vault's functional mechanism.

2.3 Introduction

Vault has been shown to promote eukaryotic cell survival in response to numerous threats, including xenobiotic [22], genotoxic [101], ER [94], and hyperosmotic stress [102]. Most notably, it plays important roles in mammalian metabolism and immunity. Although it is not essential, organisms unable to produce vaults show greater susceptibility to death from infectious disease [2, 3] and the development of some metabolic disorders [29]. Vault's singular structure, an immense, cavernous protein cage (comprised of 78 copies of major vault protein, or MVP), has inspired numerous hypotheses about its function. The cage derives most of its stability from strong, lateral hydrophobic packing between the coiled-coil domains of adjacent copies of MVP in its "cap" region.

Beneath this cap lies a Stomatin/Prohibitin/Flotillin/HflK-C (SPFH) domain, belonging to a family of structural homologs known for their ability to multimerize and associate with components of lipid rafts. While its structure is nearly identical to that of other proteins in the family, MVP's SPFH module exhibits a unique, "keyhole"-like (Figure 2.5c) beta-loop-beta structure that extends into the vault lumen. Finally, a series of eponymous beta sandwich-like repeat domains extend towards a medial interface between the two symmetrical vault cage halves (Figure 2.2a). Researchers have speculated that vault's vast surface area and largely hollow interior may serve as a signaling scaffold, reaction crucible, or molecular transport module for transiently bound "cargo" molecules. To this end, several studies have linked changes in MVP expression to signaling regulation and the movement of certain macromolecules to different compartments in the cell and even into the extracellular space [27, 83, 94, 103]. However, vault's molecular-level mechanism is not well understood.

Although vault research has overwhelmingly focused on the MVP cage, the particle contains three additional "minor" components: the enzyme poly (ADP-ribose) polymerase 4 (PARP4), the NLR-like protein telomerase component 1 (TEP1), and short non-coding transcripts known as vault RNAs (vtRNAs). These understudied components, previously estimated to occupy about an eighth of the total mass of the vault particle, are unnecessary for the formation and stability of the cage [9]. Nevertheless, understanding their function is likely key to unraveling the mystery of vertebrate vault function and mechanism. In particular, PARP4, which has known catalytic activity, may be contributing to vault function in a variety of ways.

PARP4 is a member of the PARP superfamily of enzymes, which consume NAD^+ to deposit ADP-ribose (ADPr) modifications onto macromolecular substrates. The function of PARP proteins is driven in large part by their "targeting" domains which determine their subcellular localization and, consequently, the molecules they modify [6]. PARP4 is stably recruited to the interior of the vault cage via its MVP-Interaction or "MINT" domain (Figure 2.2a). We therefore posited that PARP4's association to MVP could result in the modification of MVP, other minor vault components, or cargo molecules. Such modifications could potentially regulate the dynamics or subcellular localization of the vault cage, the repertoire of cargo molecules recruited to the vault, or the modification status of those cargo molecules, thereby impacting downstream signaling. In addition, PARP4's unusual combination of protein-protein interaction domains, including BRCT, VWA, and VIT, (Figure 2.2a) could encourage its direct interactions to vault cargo within the cage. These features motivated us to determine the structural basis for PARP4's recruitment to MVP, the impact of adding PARP4's substrate NAD^+ to the complex, and the effect of PARP4's presence on the composition of the vault interactome.

2.4 Results

2.4.1 Structure of the Human MVP-PARP4 Complex

Full-length human MVP and PARP4 constructs were expressed in insect cells through baculoviral infection. An initial attempt to separately express and then co-purify the complex proteins from co-lysate proved unsuccessful (Figure 2S.1). Although the vault cage was successfully isolated from the co-lysate purification, it did not contain PARP4, indicating that additional factors are likely required for PARP4's encapsulation into the vault. Subsequently, we co-infected insect cells with both species' baculoviruses, adding 5-6 times the volume of PARP4 virus as that of MVP. This allowed us to saturate PARP4's binding sites (in spite of MVP's very robust expression) and consequently maximize the structural information we could derive about the MVP-PARP4 interface. Following exhaustive ultracentrifugation, gradient purification, and gel filtration into buffer optimized for ADP-ribosylation (ADPrylation) assays, the MVP-PARP4 complex was applied to carbon coated grids.

The resulting structure of human MVP is broadly similar to previous rat homolog structures, with some notable differences. The first is human vault's knob-like mass of poorly resolved peptide density above the vault cap. This is the product of 30 additional residues at the human MVP C-terminus that are absent in the rat sequence. A cryo-EM map of bovine vault (whose MVP sequence also contains the additional C-terminal residues) contains a similar knob-like structure (data not shown). Although we have not probed its function experimentally, the knob's presence near two openings at the end of the vault suggest that it could act as some kind of filter for molecules that might otherwise freely enter the cage. Alternatively, it could interact with and anchor vault to cellular structures, in line with previous reports that human vault interacts with cytoskeletal elements through its cap region [104].

	apo-hMVP	hMVP-NAD^+	hMVP-PARP4-NAD^+
Microscope	Krios (University of Chicago)	Krios (University of Chicago)	Krios (University of Chicago)
Magnification	64,000x	64,000x	64,000x
Voltage (kV)	300	300	300
Spherical aberration (mm)	2.7	2.7	2.7
Detector	K3	K3	K3
Camera mode	Super resolution counting	Super resolution counting	Super resolution counting
Total exposure (e^- /$Å^2$)	45	45	40
Defocus range (μm)	-1.0 to -2.5	-1.0 to -2.5	-1.0 to -2.5
Pixel size (Å)	0.672 (1.344 physical)	0.672 (1.344 physical)	0.672 (1.344 physical)
Mode of data collection	Image shift	Image shift	Image shift
Software for data collection	EPU	EPU	EPU
Number of micrographs	6,968	4,952	9,126
Symmetry imposed	D39	D39	D39
Box size (pixel)	800	800	800
Initial particle images (no.)	205,194	318,989	758,987
Particle images for 3D (no.)	130, 552	163,428	525,525
Final particle images (no.)	98,674	108,991	439,225
Map resolution, unmasked (Å)	3.9	3.8	3.9
Map resolution, masked (Å)	3.3	3.1	3.2

Table 2.1 Statistics for data collections

Our structure also shed some light on the stability and specific properties of interactions between the symmetrical halves of the human vault. Contrary to the interactions shown in the model of the original rat vault crystal structure, we did not observe salt bridges between MVP monomers directly opposite one another (despite the fact that the relevant residues were conserved in both species). This finding was in line with subsequent vault solution structures, suggesting that the original crystallography conditions may have artificially compressed the cage. Copies of MVP that meet at the half share a greasy interface to which each monomer contributes two isoleucines (I) and a phenylalanine (F) (Figure 2.1c). Interestingly, F6 in the human sequence takes the place of an alanine in the rat MVP, potentially bolstering the strength of the hydrophobic packing in human vault. It is clear from the cryo-EM density (gray mesh, Figure 2.1) that MVP monomers come into closest proximity with opposing MVP molecules at diagonal junctions.

While the vault's ability to dynamically "breathe" in this region makes it impossible to unequivocally assign rotamer positions here, it is clear that these diagonal junctions include more polar and charged residues. In particular, the side chains of D39, R9, and Q21 represent a loose triad of residues with the potential to forge or break electrostatic interactions with one another depending on cage dynamics. In addition, N22 and N40 also reside within range to potentially interact with one another (Figure 2.1d). In aggregate, the presence of polar and charged residues at the vault equator likely contribute to the overall stability of the cage and could potentially influence the dynamics of vault cage opening through their potential to be disrupted by changes in pH, salt concentration, or other polar or charged molecules that come into contact with the vault waist.

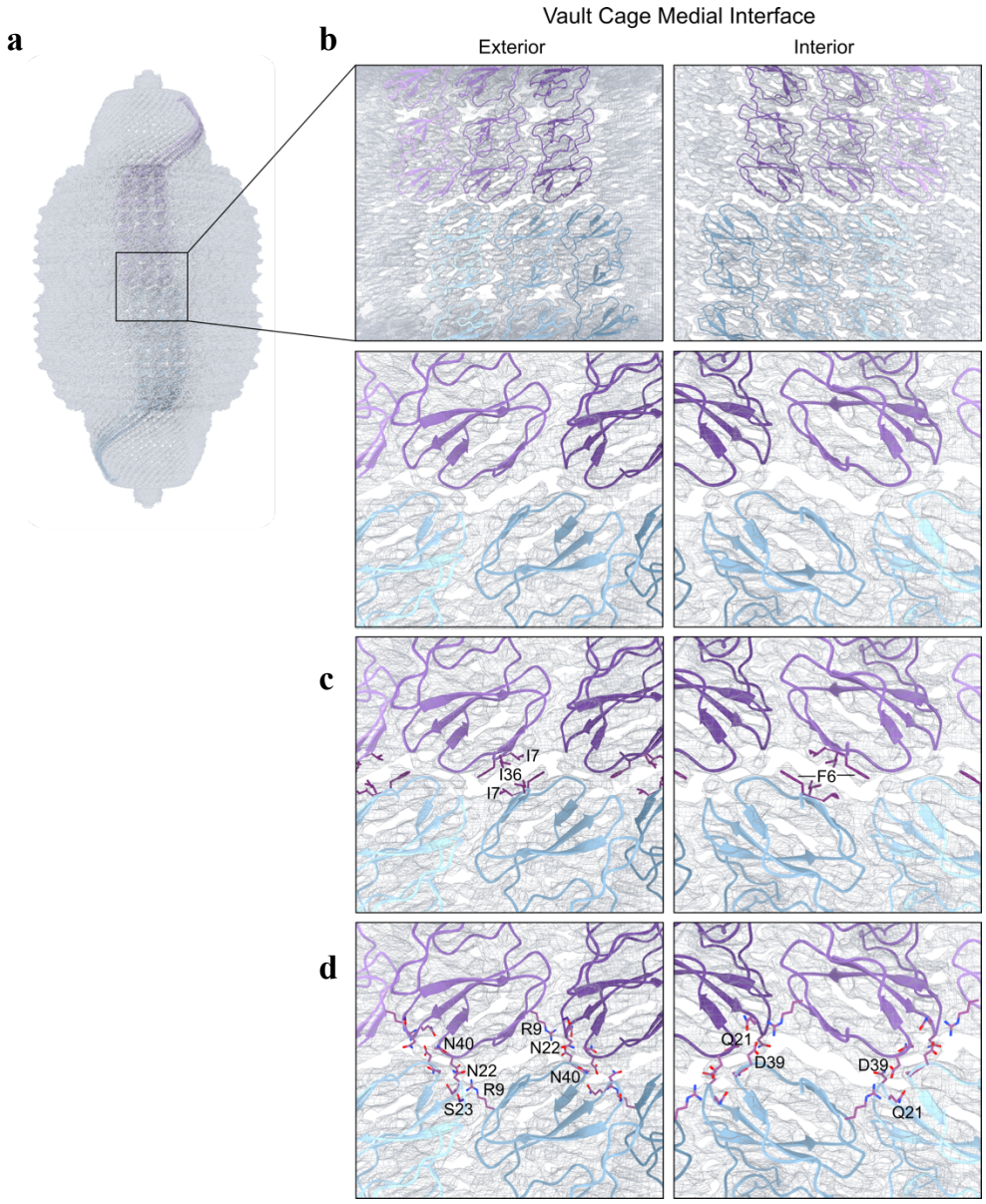


Figure 2.1 The human vault cage medial interface

a, cryo-EM map of the human vault cage, with MVP trimer models in the upper half of the vault particle shown in purple and those in the lower half colored cyan. **b,** Views of the vault medial interface, with models fit to our cryo-EM map (gray mesh). **c,** Hydrophobic residues form the interface between two directly opposed MVP monomers. The side chains of I7 and I36 are labeled in the panel on the left, and F6 is labeled in the right-hand panel. **d,** Polar and charged side chains between MVP monomers positioned diagonally across from one another at the medial interface. N22 and N40, which sit within range to hydrogen bond with one another, are labeled on the left, along with R9. Q21 and D39, both of which are highly conserved and may interact with (the also highly conserved) R9, are labeled on the right.

The most striking feature in our structures, however, was the presence of internal PARP4 density in our complex. It was clearly visible in the 3D reconstruction (Figure 2.2c), 2D classes, and could even be seen in individual particles (Figure 2S.2). Several different symmetries were imposed over the course of single particle reconstruction to achieve the best possible PARP4

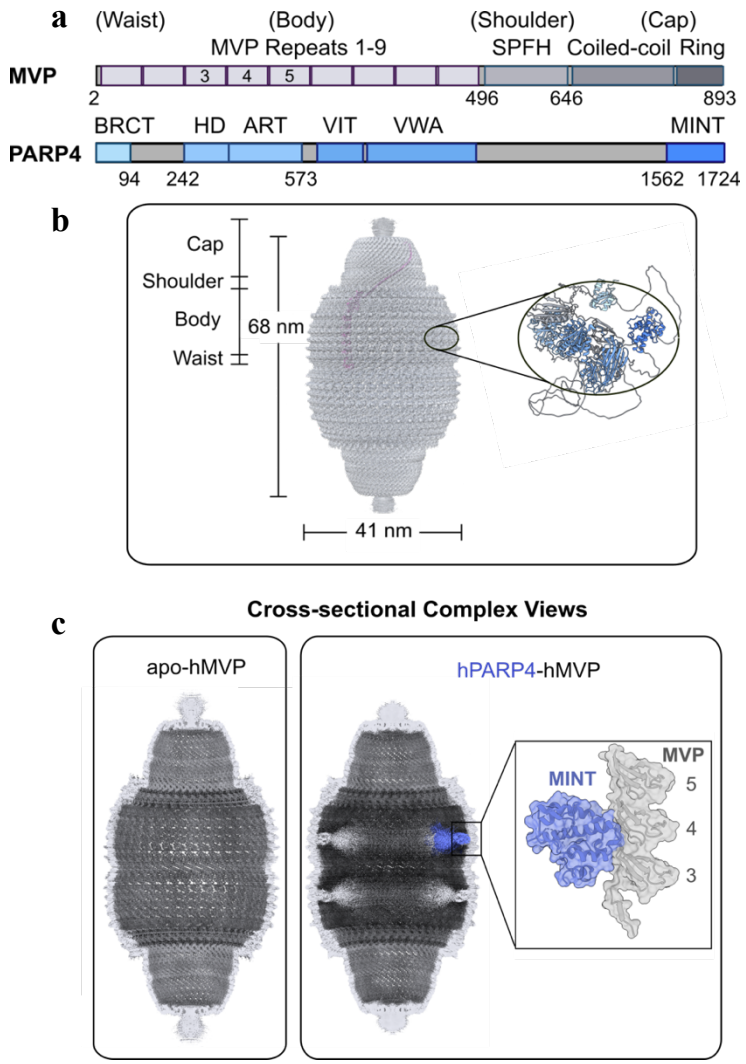


Figure 2.2 Cryo-EM maps reveal PARP4 occupancy in vault's interior

a, Domain diagrams of Major Vault Protein (MVP) and Poly (ADP-ribose) Polymerase 4 (PARP4). **b**, Dimensions of the vault cage, with one MVP monomer highlighted in pink (left) and the AlphaFold structural model of PARP4 in blue (right), with the inset showing its approximate localization in the vault cage. **c**, Cross-sectional views of the apo-MVP (left) and MVP-PARP4 cage cryo-EM maps, with density corresponding to PARP4 colored in blue. Surface representation of the MINT and MVP3-5 domains that mediate binding between PARP4 and MVP, respectively (inset).

resolution. Because the enzyme's N-terminal domains sit in the center of the vault cage and are interspersed with long disordered regions, they were too dynamic to be averaged to high resolution. After we imposed D39 symmetry, however, we were able to resolve the complex structure directly at the PARP4-MVP interface to near-atomic resolution (Figure 2S.2c). This allowed us to uncover key details of the PARP4-MVP binding interface (Figure 2.3).

The elements of the MINT domain critical to its interaction with MVP appear between its α-helices-4 and -5 and α-helices -7 and -8 (Figure 2.3b). (This stands in contrast to previous predictive computational models indicating that the interaction is driven by residues at the very C-terminal end of MINT in α-helix 10 [105].) Surface analyses of the interface reveal that it falls in a region of substantial electrostatic complementarity between MVP and PARP4. Specifically, an acidic stretch that runs along MVP's third and fourth repeat domains aligns with a basic region between α-helices-4 and -5 of the MINT domain (Figure 2.3a). Accordingly, PARP4's K1618 and MVP's D214 residues forge a salt bridge, while several nearby residues engage in additional specific electrostatic interactions. Notably, PARP4 S1614 and MVP R169 engage in hydrogen bonds with backbone amide groups of the opposing protein, as do PARP4 R1689 and MVP Q164. The side chains of PARP4 Q1682 and MVP Q172 are in range to hydrogen bond as well (Figure 2.3b). Because the majority of the vault cage interior exhibits some degree of negative electrostatic potential (Figure 2.5a), the structural arrangement of the complex components is also critical for the interaction. Several specific structural elements also appear to encourage the interaction, a phenomenon best exemplified by the fit of MVP's Q164 side chain into a small, complementary pocket in the PARP4 density (Figure 2.3c). Having identified these residues from our structure, we sought to verify their import in complex formation by imposing alanine mutations at those positions.

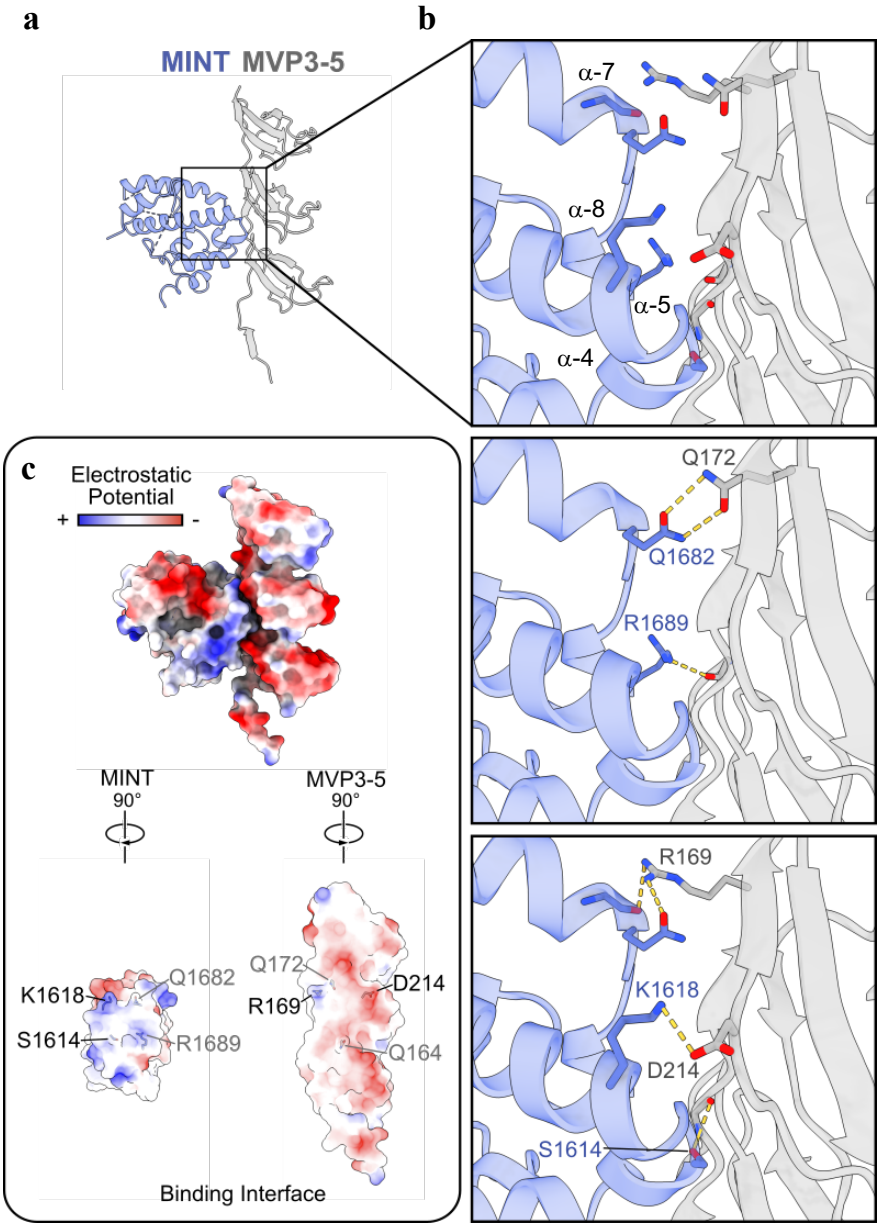


Figure 2.3 Electrostatic interactions drive MVP-PARP4 complex formation

a, Atomic model of the binding interface between PARP4's MINT domain (blue) and MVP's 3-5 repeat domains (gray) built into the cryo-EM map of the FL protein complex. **b**, All side chains that could potentially participate in electrostatic interactions between PARP4 and MVP (upper inset). Two subsets of electrostatic interactions that appear to drive complex formation (central and lower insets). **c**, Surface representation of the MINT-MVP3-5 complex model colored by relative electrostatic potential (upper). The exposed surface areas of MINT (left) and MVP3-5 (right) where they contact one another (in "open book view"). Surface is colored according to electrostatic potential and key residues are labeled.

2.4.2 Minimal Mutations Disrupt MVP-PARP4 Binding

To more efficiently test the complex formation capacity of our mutants, we designed PARP4 and MVP constructs containing only the necessary interaction domains (MINT and MVP3-5, respectively), which could be readily expressed in and purified from *E. coli*. After we confirmed that these domains could form a stable complex that remained intact over a size exclusion column, we began to impose potentially disruptive mutations to each one. Based on their electrostatic and steric properties, we began by mutating the PARP4 residues S1614 and K1618 as well as the MVP residues R169 and D214, to alanines (A) (Figure 2.3b, lower panel and 2.3c, black text). We dubbed the respective constructs the "SKA" and "RDA" mutants and used gel filtration to assay each construct's ability to interact with the corresponding species. Prior to injecting them over a gel filtration column, we incubated approximately equimolar concentrations of WT MINT and WT MVP3-5. The resulting peak eluted at a lower volume than the elution volumes of each individual species. We confirmed that they co-eluted by SDS-PAGE gel (Figure 2.4a). When we pooled fractions from the complex peak and re-injected the complex over the same column, we observed a peak that eluted at the same volume, demonstrating the stability of the complex.

We repeated this process three times, first incubating SKA MINT with RDA MVP, then SKA MINT with WT MVP3-5, and finally WT MINT with RDA MVP. We injected each mixture over the same gel filtration column, once again comparing the elution volumes of the individual components to the elution volumes of the mixtures. In each case, we saw no evidence of complex formation (Figure 2.4b-d), indicating that the PARP4-MVP interaction relies on specific residues. In addition to disrupting specific electrostatic interactions, the alanine mutations also likely buried charge (in the mutant-WT construct mixtures) and imposed steric clashes.

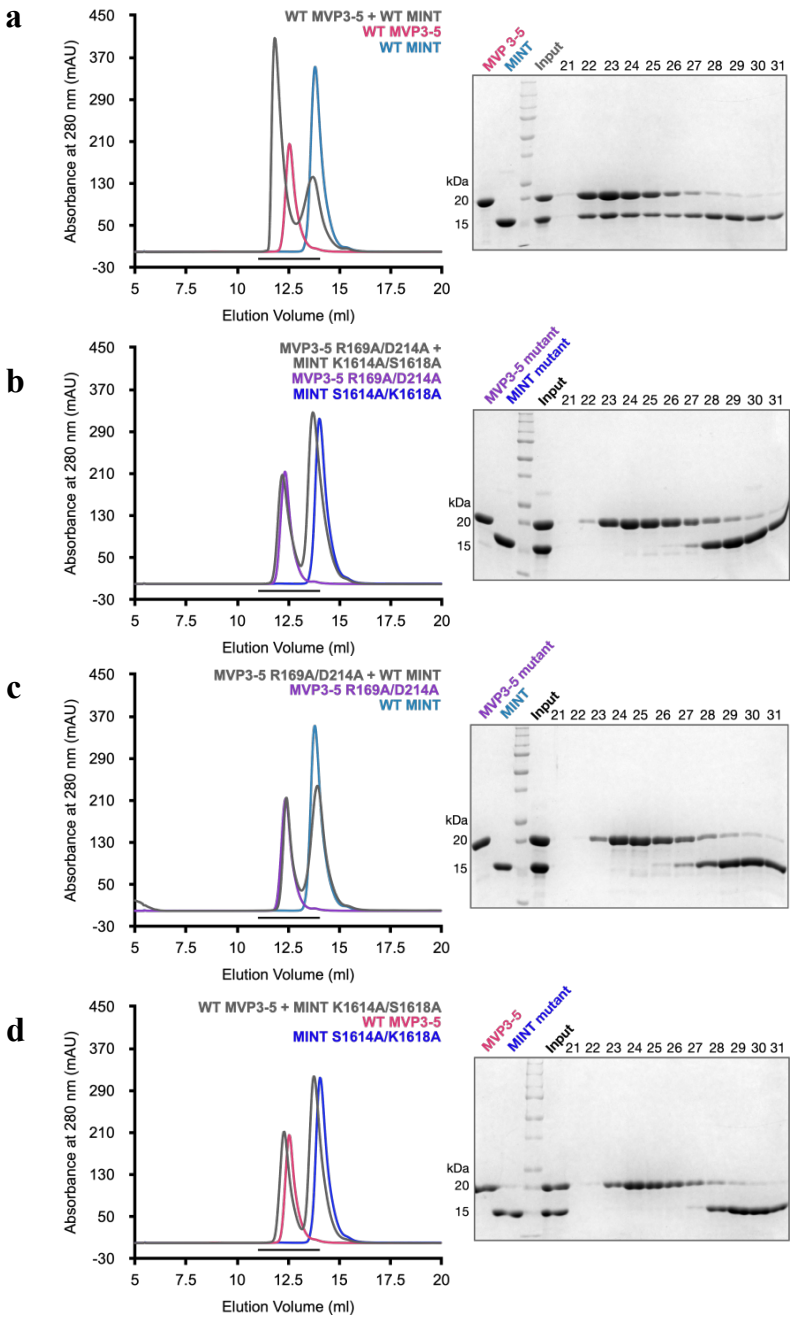


Figure 2.4 Minimal mutations abolish MINT - MVP3-5 binding

a, WT MINT and MVP3-5 constructs co-elute from a Superdex 75 gel filtration column following co-incubation. The gray chromatogram shows complex elution, and chromatograms corresponding to the individual injections of MVP3-5 and MINT are pink and cyan, respectively (left). The black bar denotes the area over which fraction samples were collected, incubated with 2X Laemmli buffer, run over a reducing SDS-PAGE gel, and visualized with Coomassie stain (right). Analogous results are shown for each combination of constructs, with both mutants shown in **b,** and a combination of one WT and one mutant construct **c-d**.

2.4.3 Structure of the Human MVP-PARP4-NAD$^+$ Complex

As a complement to our structure of MVP bound to PARP4 alone, we sought to collect a cryo-EM dataset of the complex with added NAD$^+$. In doing so, we hoped to determine whether active PARP4 would adopt an alternate conformation or whether its activity would influence the conformation or stability of the vault cage. To this end, we incubated the MVP-PARP4 complex with NAD$^+$ for several hours, dialyzed excess metabolite out overnight, and froze grids with the complex solution for data collection. Following single particle reconstruction, we observed a broad conformational change within the PARP4 N-terminus. Although it still lacked high (or even medium) resolution details, density corresponding to PARP4's innermost domains clearly shifted in the vault lumen, moving so that that they were adjacent to the MVP SPFH domain (Figure 2.9a). There did not appear to be a direct interaction between PARP4 and MVP. However, there is a loop (which we were not able to build into the map due to its flexibility) predicted to sit just beneath MVP's SPFH domain. This part of the protein could conceivably move in range to interact with the N-terminal domains of PARP4, potentially stabilizing its alternate conformation. Alternatively, the conformational shift could occur following an interaction between PARP4 and NAD$^+$, the details of which we were not able to resolve.

Interestingly, we noticed some additional, previously unobserved density within the SPFH domain of the PARP4-MVP-NAD$^+$ complex. We soon found that we could dock NAD$^+$ coordinates within it. In a structure of MVP-NAD$^+$ (lacking PARP4), the density was still present, indicating that MVP specifically binds NAD$^+$ and that PARP4 is not necessary for the interaction. The NAD$^+$ binding site sits within the vault cage interior, nestled behind the "keyhole loop" of MVP's SPFH domain, which extends into the vault lumen and creates a small pocket between itself and the remainder of the domain (Figure 2.6c – pink loop in inset). This

introduces a striking structural feature into the interior of the vault cage. Whereas the majority of vault's interior surface is relatively smooth and acidic, NAD^+'s binding site is a recessed pocket lined by basic residues (Figure 2.5b).

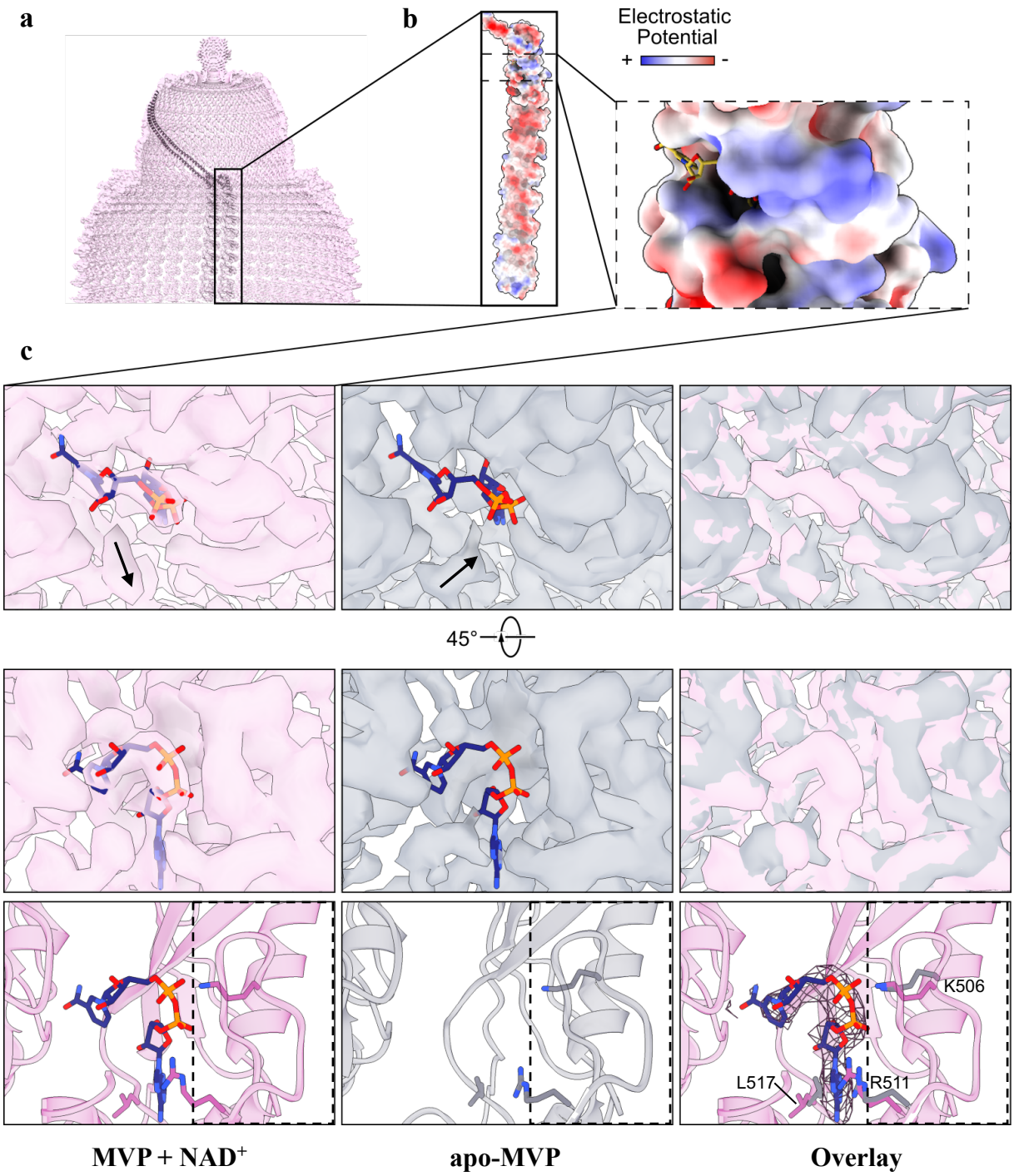


Figure 2.5 NAD^+ density behind the MVP keyhole loop

a, Interior of the vault half and inset view of one MVP monomer following electrostatic surface analysis. **b**, Surface representation of the monomer model (center). Inset view of NAD^+ (yellow sticks) fit within the surface representation model (right). **c,** Top view of density in the MVP-PARP4-NAD^+ map (pink), with the NAD^+ atomic model docked into it (left) and in the identical position in the apo-MVP map (center), where the corresponding density is missing. Map overlay (right). Directionality of an adjacent loop is shown by accompanying arrows (upper row). Side view of the binding pocket in the same maps (middle row). Atomic model of the MVP keyhole loop (dashed black box), with the sidechains of NAD^+-binding residues that undergo rotamer shifts when NAD^+ is present shown in relation to bound NAD^+, the absence of NAD^+ (center), and overlaid (right - with NAD^+ density represented as black mesh – left).

In order to better understand this interaction, we looked to both our structural model and estimated pairwise interaction energies between specific MVP side chains and NAD^+ moieties calculated by molecular dynamics (MD) simulations. Both suggested that positively charged side chains in the binding pocket are critical to neutralizing the negative charge of NAD^+'s diphosphate group. MVP R504, K506, and R511 engage in salt bridges with these phosphate groups, with K506 and R511 adopting different rotamer conformations in the presence and absence of NAD^+ (Figure 2.5c). The pairwise interaction energies between these residues and the diphosphate group are extremely favorable (Figure 2.6b). The basic residue R437 also plays an important role in NAD^+ binding, likely through a combination of phosphate charge neutralization and pi-cation stacking against NAD^+'s niacin group. Additional residues that drive binding include D521 and W545, which engage NAD^+'s ribose sugars in hydrogen bonds. The bulky, hydrophobic residues L517, F522, and F523 serve to delineate the boundaries of the pocket.

The vault cage contains as many theoretical NAD^+-binding pockets as it does copies of MVP, and the majority of highly favorable contacts occur between NAD^+ and the MVP monomer ("MVP2" – Figure 2.6c) whose keyhole loop it sits behind (Figure 2.6c – pink loop).

However, residues on the copy of MVP that sits immediately to the left ("MVP1" – Figure 2.6c) also support NAD^+'s association with the vault cage. Two proline residues on MVP1 introduce additional steric constraints around the binding pocket to further stabilize bound NAD^+. The amide group of NAD^+'s nicotinamide engages in hydrogen bonds with backbone amide groups of D566 from MVP1, while its K537 residue can also contribute to the neutralization of phosphate charge. While several of the above side chains undergo rotamer shifts, the metabolite's presence does not seem to induce any larger changes in the SPFH domain. However, some density extending from a flexible loop below, seems to be displaced further out into the vault lumen in the presence of NAD^+ (Figure 2.5c). Unfortunately, the fact that it's been largely averaged out limits our ability to interpret the impact of this shift, but it does present the possibility that bound NAD^+ may induce some degree of conformational change within the vault cage.

Initially, it was not entirely clear whether the observed density represented bound NAD^+ or covalently attached ADP-ribose. This was due to the similarity of the metabolites' structures and the fact that the density for NAD^+'s flexible nicotinamide moiety was not as clearly defined as the rest of the molecule. However, the geometry of our docked NAD^+ coordinates could not accommodate an interaction between a potentially modifiable residue (D521) and the relevant hydroxyl group on NAD^+'s distal ribose; the distance between them exceeded the length of a C-O covalent bond. Furthermore, after repeating our MVP-NAD^+ incubation and submitting the sample for post-translational modification analysis via LC-MS/MS, we were unable to identify an ADPr modification on D521. The MS results did, however, indicate that there was an ADPrylated residue elsewhere on the cage.

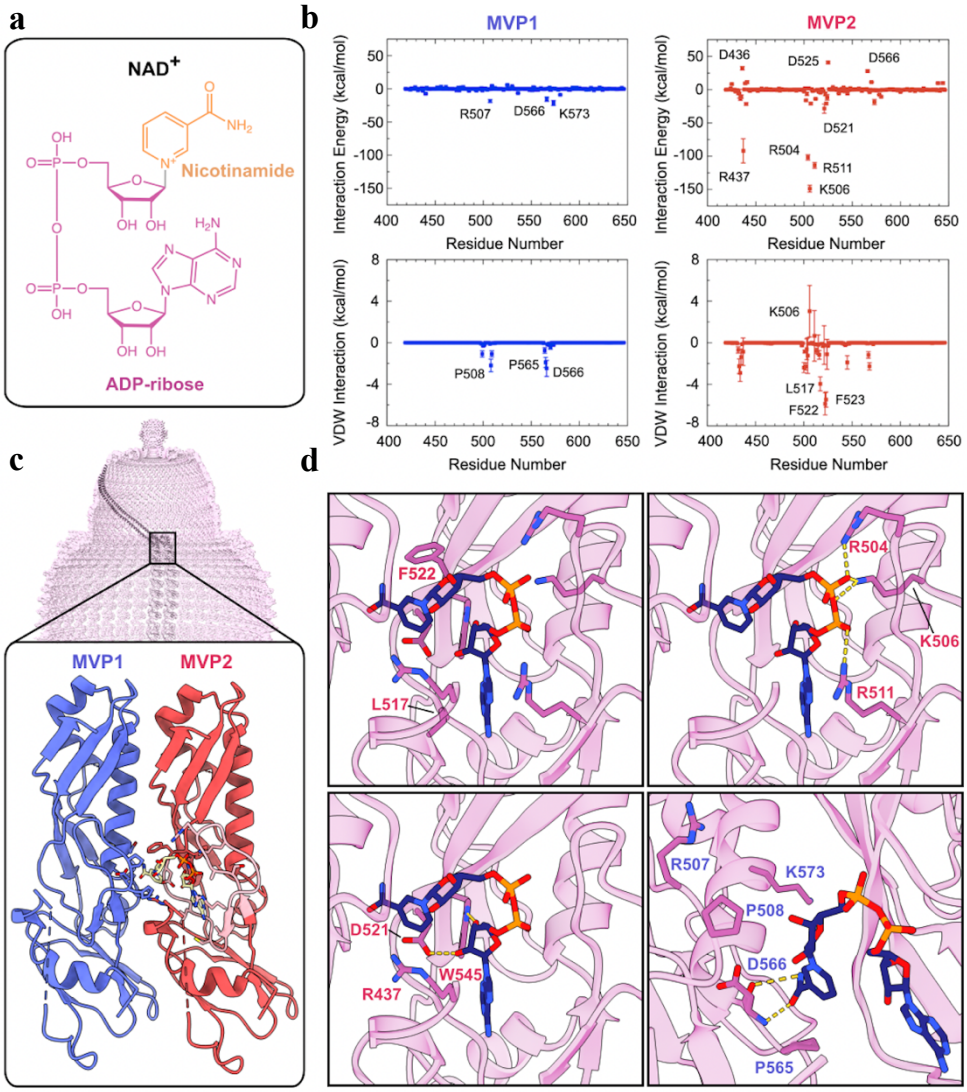


Figure 2.6 MVP residues critical to NAD⁺ binding

a, Chemical structure of NAD^+. ADP-ribose and nicotinamide moieties are shown in pink and gold, respectively. **b,** Pairwise interaction energies between individual MVP residues and NAD^+ moieties calculated by MD simulations. Results for residues corresponding to MVP1 are shown in blue, and those for MVP2 are shown in red. **Data provided by Dr. Rong Shen.** **b,** Interior of the vault half (top) with inset view of two SPFH domain monomers (below). The loop containing essential NAD^+ binding residues is colored in light pink within the atomic model of SPFH from "MVP2," with bound NAD^+ shown in yellow. **d,** Close-up views of essential residues for MVP's interaction with NAD^+, where MVP2 residues are labeled in red, and MVP1 residues are labeled in blue. Bulky, hydrophobic residues line the NAD^+-binding pocket (top left). Basic residues forge electrostatic interactions with the ADP-ribose diphosphate group (top right). Polar and charged residues interact with the ribose and niacin groups of NAD^+ (lower left). Residues from MVP1 provide additional steric constraints around the NAD^+ binding pocket (lower right).

2.4.4 MVP is ADP-ribosylated in the Presence of PARP4

In order to identify ADPr modifications using mass spectrometry, both the MVP-only and the MVP-PARP4 complex were supplemented with NAD^+ as described above. Afterwards, the proteins were visualized using a denaturing SDS-PAGE gel, and the bands corresponding to each protein were extracted and submitted for in-gel digestion and LC-MS/MS analysis. Because PARP4 had previously been characterized as a mono (ADP-ribosyl) transferase (MART) enzyme [97] (one that appends a single ADPr to its modification targets), we looked for mass shift signatures corresponding to mono-ADP-ribose (MAR). We did not detect any modifications associated with the apo-MVP cage. However, MVP from our MVP-PARP4 complex did appear to exhibit the characteristic mass shift of ADPrylation on a serine (S23) near the vault cage equator (Figure 2S.5), suggesting that it undergoes PARP4-dependent MARylation. While we did not observe clear density for this modification in our ternary complex map, the flexibility of an ADPr group unconstrained by a binding pocket (like the one that confines bound NAD^+) could easily have been averaged out over the course of single particle reconstruction, especially if only a few residues on the cage were modified.

The exterior placement of this apparent modification is surprising given PARP4's association to the vault cage interior. Due to the established flexibility of the vault cage (Figure 2.7b) and its ability to exchange halves with other vaults, however, it is possible that PARP4 could (at least temporarily) access this region of the cage to modify it. The stability and overall structure of the vault cage did not change substantially following NAD^+ addition in the presence of PARP4 (Figure A.2). This negated an initial hypothesis we held that PARP4 might regulate the opening of the cage by depositing ADPr groups on residues that engage in inter-half salt bridges. Although the modified residue sits on the edge of an MVP monomer, its rotamer position would

not seem to encourage inter-vault half disruption if modified by an ADPr group. However, it could potentially serve as a binding platform for cargo molecules about to be engulfed within the cage lumen. This possibility motivated us to determine whether PARP4's presence could alter the repertoire of vault-interacting proteins.

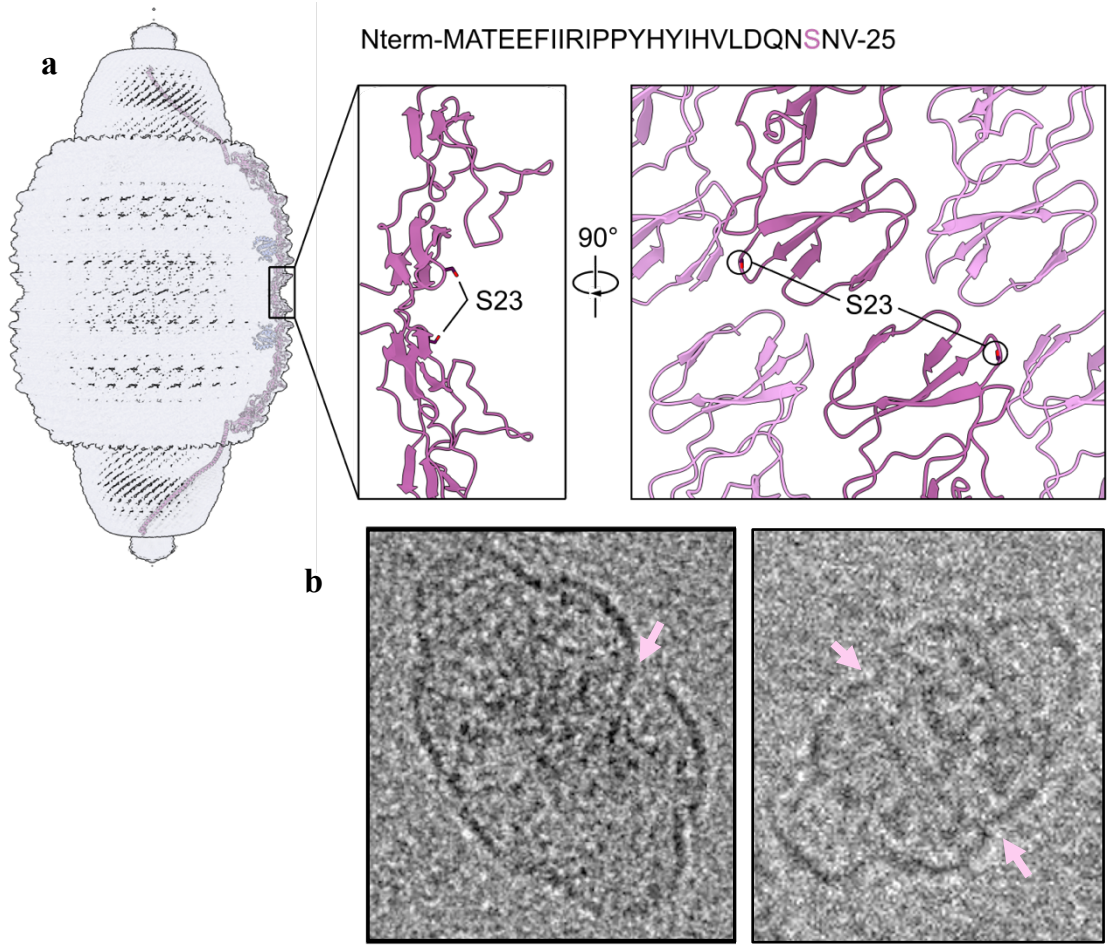
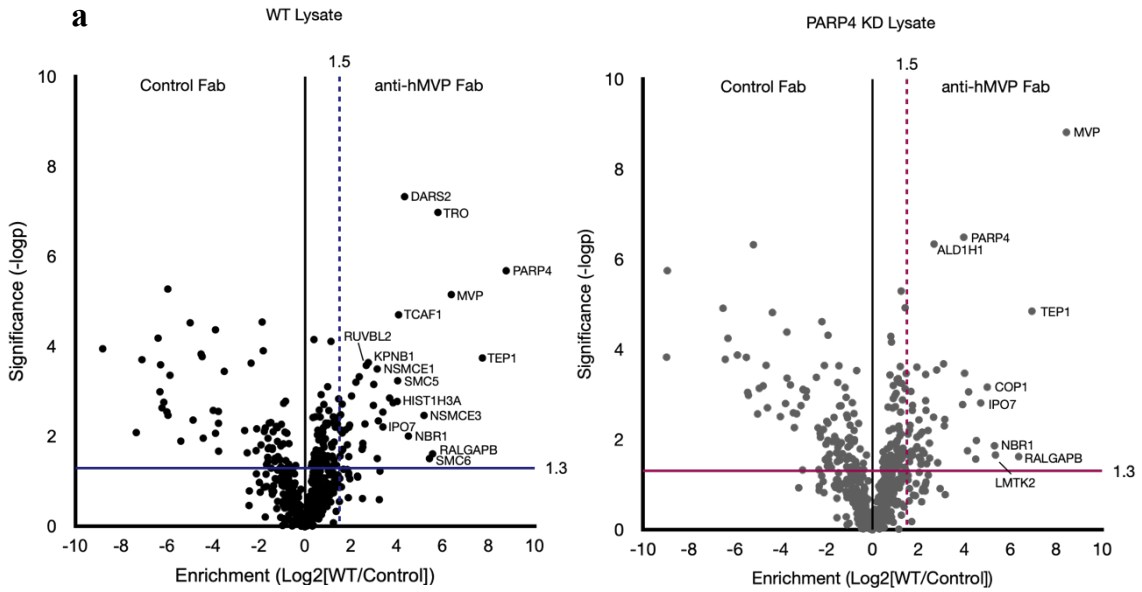


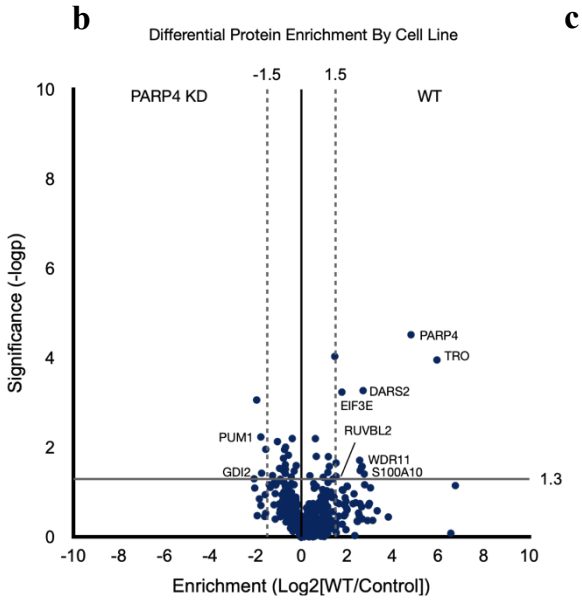
Figure 2.7 Serine 23 of MVP appears to be ADP-ribosylated in the presence of PARP4

a, Cartoon representation of the vault cage, with MVP and bound MINT shown in pink and blue, respectively. The location of serine 23 (S23) is highlighted in pink in MVP's N-terminal sequence (top) and shown on the MVP cage (inset), both from the side (left) and in an exterior view of the cage (right). **b,** cryo-EM images of vault particles from native bovine vault (left) and recombinant MVP-PARP4 purifications (right) in which the flexible MVP cage moves inward towards the particle's lumen, creating a potential opportunity for internal PARP4 to modify MVP. Points where the termini of MVP apparently move into the cage are indicated by light pink arrows.

2.4.5 Vault Coimmunoprecipitation (Co-IP) from WT and PARP4-Depleted Cells

In order to assess PARP4's influence on the vault interactome, we engineered a PARP4-depleted A549 cell line using CRISPR-Cas9. Although PARP4 expression was not completely abolished, it was substantially downregulated (Figure A.1). We chose to use A549 cells (derived from human lung epithelia) due to their high constitutive expression of vault proteins and vault's documented, medically relevant activity in lung tissue [24-26]. Having previously retained the services of the Kossiakoff lab to generate anti-hMVP fragment antigen-binding proteins (Fabs) (Figure 2S.4), we constructed Avi-tag versions that could be biotinylated and subsequently isolated from heterogeneous cell lysate solutions with magnetic streptavidin beads. To control for the detection of abundant background proteins and those that might interact with the anti-hMVP Fab's immunoglobulin backbone, we compared the set of proteins pulled down by our anti-hMVP fab with a set that associated to an anti-GFP control Fab. Three anti-hMVP and control IP biological replicates were conducted per cell line (WT versus PARP4 KD) for a total of twelve samples. Fresh protease and PARG inhibitors were added to lysis buffer immediately prior to the experiment to avoid protein degradation and the loss of potential ADPr modifications, respectively. To avoid false positives, we limited the Fab-bead/lysate incubation to two hours and subjected the beads to a very stringent series of washes. After eluting the immunoprecipitated proteins with acetic acid, we submitted the samples for LC-MS/MS analysis to identify vault binding partners.





c

	PARP4 KD		WT	
	Gene	Log2 (Exp/Ctrl)	Gene	Log2 (Exp/Ctrl)
1	GAPDH	2.3	PARP4	8.8
2	PKM	3.1	TRO	6.7
3	DDX3X	3.0	RACK1	3.1
4	ENO1	1.9	PPIA	2.6
5	RPS9	2.5	TCAF1	4.2
6	PUM1	4.9	DARS2	4.3
7	ZNF45	5.0	WDR11	5.8
8	ZNF268	3.5	MMTAG2	1.6
9	LDHB	1.5	RPL27A	1.5
10	SFRP1	1.9	LUC7L3	1.6
11	GPI	1.9	AKR1B1	3.4
12	DDX56	1.4	AKR1C2	2.5
13	MRPS21	2.7	RUVBL2	2.7
14	PUM2	2.5	PRDX1	1.6
15	CASP14	1.6	ACAD11	1.9
16	UGDH	2.1	S100A10	2.2
17	GDI2	4.5	HNRNPDL	1.6
18	RPL39	2.4	TXN	2.5
19	MED18	2.1	HIST1H3X	4.0
20	GSTP1	1.7	HTRA1	1.5
21	NUMB	2.1	SLC25A1	1.7
22	MATCAP2	2.1	EIF1B	1.6
23	SPEG	1.7	MICAL2	3.0
24	TRIM7	2.2		
25	LMTK2	5.4		

Figure 2.8 Differential vault binding activity in WT and PARP4-depleted cell lines

a, Volcano plots showing the relative enrichment of proteins pulled down by either an anti-hMVP experimental Fab or anti-GFP control Fab from WT (left) and PARP4-depleted cell lysate (right). Proteins to the upper right of the significance and enrichment cutoffs (colored lines) are considered "hits." Top hits from each condition are labeled. **b,** Volcano plot showing a direct comparison of protein abundances between each cell line. **c,** Table of proteins that are either significantly enriched or uniquely identified as hits in one cell line versus the other. Log2(Exp/Ctrl) values refer to the enrichment of each protein in anti-hMVP versus control Fab IP samples.

To identify hits, we looked for proteins with differential abundance values in the anti-hMVP Fab IP versus control Fab IP samples of $\log_2(\text{Fold Change}) \geq 1.5$. To assess the significance of these hits, we subjected normalized protein abundance readings from each replicate to a 2-tailed, homoscedastic t-test and only kept those hits whose resultant -log(p-value) was > 1.3 (indicating a p-value of < 0.05). Using this framework, we identified significantly enriched vault component proteins, as well as COP1 (aka RFWD2), a ubiquitin ligase previously observed to interact with vault [101]. We also observed about 70X PARP4 protein abundance in vault elution samples purified from the WT versus KD cells. Now confident in our approach, we identified a total of 42 significant hits from the WT lysate and 44 from the PARP4-depleted sample (Figure 2.8c). As this numerical mismatch would suggest, each dataset contained vault binding partners unique to that cell line (Figure 2.8c). In addition, several proteins present in both datasets were significantly more abundant in one condition than the other (Figure 2.8b-c). While the biological meaning of individual hits has not yet been elucidated, these data offer an initial indication that PARP4's presence in the vault may work to recruit specific cargo to the particle.

2.5 Discussion

Our structure of the MVP-PARP4 complex is the first to reveal the interaction between MVP and a native protiein binding partner at near-atomic resolution. It also resolves a longstanding question in the field concerning the precise nature of the MVP-PARP4 interaction. To date, vault researchers have worked from an incorrect model produced by an early generation of protein structure prediction software [106]. Although the AlphaFold model of MINT, itself, was a vast improvement over the previous model, AlphaFold still failed to predict the correct

interaction interface with MVP, possibly due to the unique properties of the folds involved. Despite its apparent simplicity, the MINT domain has no close analogs in the RCSB, and, to date, is only predicted to share substantial structural similarities with models of domains in the FoldSeek database (notably those present in "VWA-containing protein 5B1"). Our structure of this tricky interface will allow for precisely targeted assays probing the function of vault-bound PARP4 in the future.

Our discovery of MVP's novel NAD^+-binding pocket introduces a number of new questions to the field. Provided that NAD^+ is MVP's preferred ligand, its binding could herald myriad functions. The overwhelming majority of NAD^+-binding proteins are redox enzymes or enzymes that consume NAD^+ [74], neither of which are structurally similar to MVP's NAD^+-interaction domain. MVP's retention of the NAD^+ molecule in its binding pocket and its apparent inability to deposit any NAD^+-derived PTMs further indicate that it is not acting as an NAD^+ hydrolase. Outside of these two main categories, most NAD^+ interaction proteins either bind the molecule to stabilize a particular conformational state [75, 76], dissociate from a binding partner [107], or play some role in NAD^+'s movement throughout the cell. Along these lines, NAD^+'s displacement of the loop below MVP's shoulder (Figure 2.5) could conceivably impact MVP function. For instance, it could potentially allow this region of MVP to interact with additional molecules (such as PARP4), possibly in response to changes in cellular NAD^+ concentration. NAD^+'s occupancy of the pocket could also preclude some as yet undiscovered substrate from binding there instead, introducing another layer of vault cage regulation.

Alternatively, NAD^+ could be thought of as an additional form of cargo. While MVP does not possess a standard channel or transporter structure, it could conceivably traffic NAD^+ between cellular compartments or even into the extracellular space where the NAD^+

concentration is very low [108]. Mounting evidence suggests that vault can traverse the extracellular space to deliver cargo between cells [83]. It stands to reason that it could retain some NAD^+ for consumption by its resident PARP enzyme or other bound NAD^+-utilizing molecules in this relatively depleted space. While one could imagine a scenario in which MVP's NAD^+- and PARP4-binding functions co-evolved to provide a repository of substrate for the enzyme, MVP multiple sequence alignments suggest that this is not the case; the NAD^+- and PARP4 interaction residues are not co-conserved (Figure 2.9c). Although several NAD^+-binding residues are conserved in invertebrate species, the PARP4-binding residues (or at least those dictating the interaction in human vaults) are not present in those homologs. However, another distinct possibility remains: that NAD^+ is not the preferred ligand for the MVP binding pocket.

NAD^+'s broad structural similarity to other nucleotides and MVP's specific affinity for its phosphate and ribose moieties indicate that other small molecules could be bound by the pocket. Our MD simulations indicate that if the nucleobase itself were swapped out for a guanine, the resulting nucleotide would still be capable of binding MVP as long as the ligand retained polar ribose and diphosphate groups (Figure 2S.7). While GDP's relatively low abundance in the cell makes it an unlikely candidate as the site's primary ligand, the simulations nonetheless illustrate the versatility of the binding site. This kind of versatility is characteristic to many NAD^+-binding proteins. NAD^+ transporters can accommodate multiple nucleotide species, with some acting to transport multiple related metabolites. Proteins in the NAD^+-binding NUDIX family are known for their ability to interact with oxidized, mutagenic nucleotides that they "clean" from the cell by hydrolyzing them [74]. By analogy, MVP could either currently act, or at some point in its evolutionary history have acted, to detect oxidized nucleotides or related xenobiotics from the cell via this binding module.

Intriguingly, NAD^+'s positioning in the pocket (with the adenine situated furthest into the binding site) indicates that the site could also potentially bind an ADPr modification. This is consistent with previous work that identified MVP as a PAR-binding protein. The 20-residue peptide those authors identified as the driver of the interaction (Figure 2.6c – the pink loop on MVP1) maps perfectly onto our NAD^+-binding site [109], indicating that the terminal ADPr moiety of a PAR chain could occupy the pocket in a similar way. By the same logic, a MAR modification could theoretically occupy the space (though the keyhole loop would likely have to move substantially to accommodate the remainder of the modified molecule). If such a structural change were not possible, this would call into question whether PARP4 substrates (known to be MARylated) could be bound at the site. However, substrates of PARylating enzymes, or even free PAR chains themselves could potentially be anchored there. Although MVP has not, to this point, been shown to directly sequester chemotherapeutic drugs, an ability to interact with molecular hallmarks of DNA damage (including PAR chains or oxidized nucleotides) could contribute to its association with chemotherapeutic resistance.

MVP's ability to interact with PAR could also portend an additional binding function. Unsurprisingly, there is considerable overlap between proteins that bind PAR and those that bind single-stranded nucleic acids, introducing yet another possible function for MVP's nucleotide binding pocket. Although there is strong evidence to suggest that TEP1 is necessary for vtRNA's inclusion within the particle, one could imagine a role for MVP's nucleotide binding pocket in the recruitment of additional RNA species shown to associate to vault [83]. To this point, we have only biochemically verified the interaction between the MVP cage and NAD^+ (using isothermal titration calorimetry) (Figure 2S.6), but we have additional plans to investigate the

binding site's versatility by measuring its affinity for ADP, ADPr, and other dinucleotides as well.

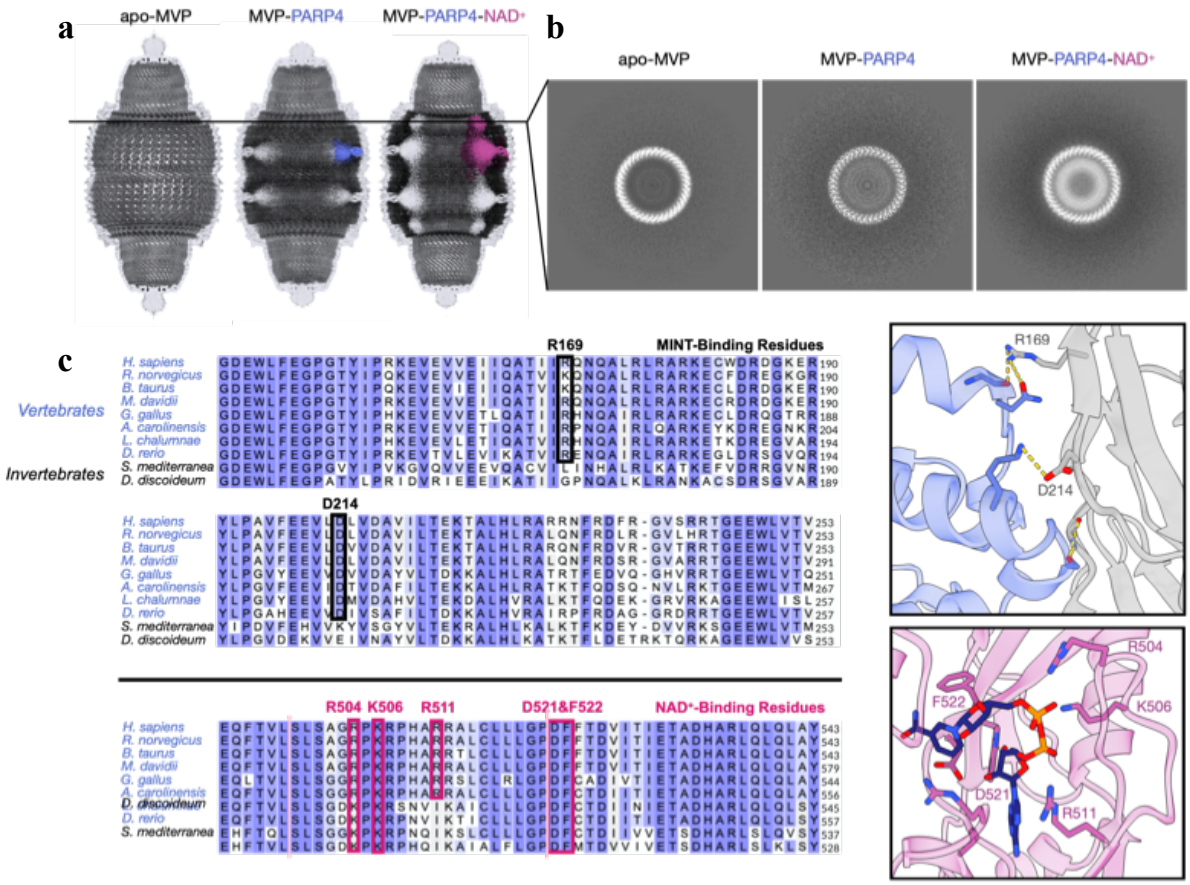


Figure 2.9 Bound NAD⁺ changes the conformation of vault-incorporated PARP4, but residues necessary for vault's interactions with NAD⁺- and PARP4 are not co-conserved

a, Cross-sectional views of the apo-hMVP, MVP-PARP4, and MVP-PARP4-NAD$^+$ cryo-EM maps. **b**, A 2D "slice" of cryo-EM density taken from the same plane in each map (indicated by the line passing through them). **c,** Multiple sequence alignment (MSA) of the PARP4 and NAD$^+$-binding sites of MVP expressed in several vertebrate and invertebrate species. Residues responsible for human MVP's interaction with PARP4 and NAD$^+$ are outlined in black and pink, respectively. Atomic models featuring the side chains of the outlined residues are shown on the right.

An apparent PARP4-dependent ADPrylation site near vault's medial interface could contribute to a dynamic PTM code that would attract (or repel) different molecules to the MVP cage depending on which PTMs decorated it at the time. ADPrylation is known to target sites otherwise occupied by other PTMs (especially serine phosphorylation [34]) under specific cellular conditions. Analogously, the interchange of PTMs at a specific vault cage site depending on cell state could alter the repertoire of vault's interacting cargo molecules depending on which modification was currently favored. The modification's presence right next to the medial interface could play a role in positioning bound cargo near a site where it could be internalized into the cage. This could, in part, begin to explain vault's ability to participate in such diverse cellular processes with its wide range of interacting partners. It also would have clear implications for elucidating PARP4's regulatory function within the vault cage. Along these lines, our proteomics data show that vaults purified from WT and PARP4-depleted A549 cells do associate with distinct species and abundances of various macromolecules.

2.6 Methods

2.6.1 Plasmid Construction and Baculovirus Generation

Human MVP and PARP4 genes were synthesized by GenScript and subcloned into the pVL1393 baculovirus transfer vector using Gibson Assembly. Ligation products were transformed into Top10 competent cells and isolated using the Qiagen miniprep kit. pVL1393-hMVP and -hPARP4 plasmids were incubated with BestBac2.0 linearized baculoviral DNA (Expression Systems, 91-002) and transfected into adherent Sf9 insect cells using Cellfectin II (Invitrogen 10362-100). After 7 days, the cells were pelleted by centrifugation and P0 baculovirus was harvested from the supernatant fraction. P0 virus was amplified in Sf9

suspension cells for an additional week to generate P1 baculovirus stocks. P1 viruses were, in turn, amplified in Sf9 cells to generate P2 stocks. P1 and P2 stocks were used to infect Hi5 cells) for protein expression. Constructs for *E. coli* expression of WT and mutant MVP 3-5 and MINT domains were cloned into the pET47b(+) vector, which introduced N-terminal His tag and HRV cut sites. Alanine mutations were imposed by round-the-horn PCR and confirmed by Sanger sequencing.

2.6.2 Expression and Purification of Vault Cage Complexes

Hi5 cells were grown to one liter at a density of about 2 x 10^6 cells / ml and infected with 0.8 ml hMVP P_1 baculovirus. hMVP-hPARP4 complex samples were infected with an additional 4-5 ml hPARP4 P_2 baculovirus. Infected cells were rotated on an orbital shaker at 120 rpm in a room with an ambient temperature of 27°C. After 65 hours, they were harvested by centrifugation at 900 *g* for 15 minutes. Cell pellets were washed with 20 ml of phosphate buffered saline (PBS) per 0.5 L of cells, transferred to 50 ml Falcon tubes, and centrifuged for an additional 15 minutes at 900 *g*. Cells were decanted, and aliquots were flash frozen in liquid nitrogen prior to long term storage at -80°C. For purification, 0.5 L cell aliquots were thawed in room temperature water for 10 minutes, then resuspended in 120 ml Buffer A (50 mM Tris pH 7.4, 75 mM NaCl, 1.5 mM $MgCl_2$, & 1 mM DTT), with 1% Triton X-100 and 1 mM phenylmethylsulfonyl fluoride (PMSF). Pellets were disrupted with 30-40 strokes of a Type A Dounce homogenizer. Lysate was transferred to centrifuge tubes and vortexed twice during a 20-minute incubation on ice. Lysate was cleared by centrifugation at 20,000 *g* for 15 minutes, and the supernatant was harvested for additional purification steps. High-molecular weight complexes were pelleted by centrifugation at 150,000 *g* for one hour. Pellets were resuspended

(with the assistance of a 1-ml dounce) in six milliliters of Buffer A with 7% Ficoll and sucrose added.

The resuspension was centrifuged for 45 minutes at 43,000 *g* in order to pellet microsomal contaminants. The supernatant fraction was diluted in 21 ml of Buffer A and pelleted by ultracentrifugation for 3 hours at 200,000 *g*. Pellets were resuspended in 0.8 ml of Buffer A and incubated with 500 µg RNase A as well as 50 U of RNase T1 for 15 minutes. Denatured ribosomal proteins were pelleted by centrifugation at 30,000 *g* for 20 minutes. The supernatant was then transferred to a new 1.5 ml tube for an additional 15-minute spin. The remaining supernatant was taken up to 1 ml with Buffer A and gently pipetted over a 20%-60% sucrose gradient. Gradients were centrifuged at 78,000 *g* for 16 hours. Fractions from the 45% and upper half of the 50% gradient fractions were pooled and dialyzed in ADP-ribosylation buffer (50 mM HEPES pH 8, 5 mM $MgCl_2$, 5 mM $CaCl_2$, 0.25 mM DTT) overnight. Dialysate was concentrated to 500 µl and injected over a Superose 6 increase 10/300 GL column (Cytiva). Fractions were aliquoted and flash frozen with liquid nitrogen prior to storage at -80°C.

2.6.3 Expression and Purification of MVP3-5 and MINT Truncation Constructs

pET47b-MVP3-5 and -MINT WT and mutant constructs were transformed into Rosetta Singles BL21(DE3) *E. coli* competent cells. 50 ml Luria Broth (LB) growth media were inoculated with a single colony and grown overnight at 37°C. 8 ml of each starter culture were added to 990 ml LB and grown at 37°C until the OD600 reached 0.6-0.75. Protein production was induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and cells were kept at 18°C and rotated at 220 rpm overnight.

For purification, cells were pelleted by centrifugation at 6,000 rpm for 20 minutes and each pellet was resuspended in 50 ml Ni-NTA buffer (20 mM Na_2HPO_4/NaH_2PO_4 pH 7.4, 300 mM NaCl, 1 mM PMSF) with 10 mM of imidazole. Resuspended cell solutions were subjected to sonication for 3 minutes at the 75% amplitude setting, with a pulse sequence of 3s on/5s off. Lysate was cleared by centrifugation at 16,500 rpm for 45 minutes, and the supernatant was flowed over 6 milliliters of washed Ni-NTA resin in a gravity column. Resin beads were washed with 150-200 milliliters of Ni-NTA buffer with 20 mM Imidazole. Proteins were eluted from the column following a 5-minute incubation in 15-25 milliliters of Ni-NTA buffer with 250 mM Imidazole. Protein eluate was dialyzed into two liters of SEC buffer (50 mM Tris pH 8, 150 mM NaCl, and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP)-HCl) and subjected to His-tag cleavage with PreScission Protease overnight. Dialyzed protein was concentrated, filtered, and injected over a Superdex 75 increase 10/300 GL column (Cytiva). Peak fractions were pooled and used immediately for gel filtration peak shift assays.

2.6.4 Cryo-electron Microscopy Sample Preparation

Vault cage peak fractions were concentrated and incubated at 4°C for 3.5-4 hours in the presence or absence of 1 mM NAD^+. Samples were dialyzed overnight to remove excess metabolite in the NAD^+-added solutions. Vault proteins were then concentrated and subjected to 5 minutes of centrifugation at 13,000 rpm to pellet aggregates. Quantifoil (R1.2/1.3, 200 mesh) copper grids with 2 nanometers of continuous carbon film (Electron Microscopy Sciences) were glow discharged using a 15 second, 5 watt program on the Solarus 950 Plasma Cleaner System (Gatan). Three microliters of sample (1.5-1.8 mg ml^{-1}) were applied to the carbon-coated grid surface. The climate chamber of the Vitrobot Mark IV (ThermoFisher) was set to 8°C and 100%

humidity. Grids were blotted for 1-2 seconds at blot force 0 with standard Vitrobot filter paper (Ted Pella, 47000-100), then plunge-frozen into liquid ethane.

2.6.5 Data Collection

Data were collected at the Advanced Electron Microscopy Facility at the University of Chicago on a 300 kV Titan Krios transmission electron microscope (ThermoFisher Scientific), equipped with a K3 direct detector camera (Gatan). Movies were recorded at a nominal magnification of 64,000x (translating to a pixel size of 0.672 Å at the detector), in super-resolution counting mode by image shift. EPU software was set to automated acquisition mode and collected (depending on the dataset) between 2745 and 9126 image stacks, each with 40 frames, subject to a total exposure of 40-45 $e^-/Å^2$. The defocus range was set to -1.0 to -2.5 μm. See **Table 2.1** for additional details.

2.6.6 Image Processing

Stack images were exported to cryoSPARC live (v3.3.1), where they underwent motion correction and CTF determination. Particles were picked using 2D class averages generated from initial templates, then exported to cryoSPARC for additional processing. Contaminants, broken vaults, and poorly-aligned classes were eliminated following initial 2D classification. Depending on the dataset, between 100,000 and 318,989 particles were used to generate *ab initio* classes that were typically of low quality but useful as templates to sort particles during heterogenous refinement. For the first homogeneous refinement, a low-pass filtered MVP cage map from an earlier dataset served as the initial model, and D39 symmetry imposed to accommodate the alignment of the 39 MVP monomers in each half of the vault cage. The refined map and previous

low-quality *ab initio* volumes were used as inputs for a heterogeneous refinement job to further sort vault cage particles. Particles aligned to the intact MVP cage map were subjected to non-uniform refinement. The resulting map was typically used as the input for an additional round of either homogeneous or non-uniform refinement, in which the program was instructed to optimize the per-particle defocus and per-exposure-group CTF parameter optimization, as well as beam tilt, anisotropic magnification, and Ewald's sphere correction. See Figures 2S.2 and 2S.3 for more details.

2.6.7 Isothermal Titration Calorimetry (ITC)

MVP cage protein was purified mostly as described in 2.6.2, with the exception that the buffer was DTT-free (since it tends to oxidize during injection and distort the ITC baseline). Sample was concentrated to 320 µl and subjected to a 10-minute spin at 13,000 rpm, at which point the supernatant was transferred to a new tube. The MVP monomer concentration was calculated to be about 80 µM protein absorbance measurement at 280 nm. A 5 mM solution of NAD^+ (ChemImpex 00229) was prepared in the exact ADP-ribosylation buffer used for MVP SEC. ITC binding experiments were conducted using a MicroCal PEAQ-ITC instrument (Malvern Panalytical). Following an initial 0.4 µl ligand injection, 18 injections of 2 µl NAD^+ solution were titrated into buffer (as a control) or MVP at a rate of 2 s μl^{-1} at 150 s time intervals. Stirring speed and experimental temperature were kept at 750 rpm and 25 °C, respectively.

2.6.8 Gel Filtration Peak Shift Assays

Wildtype and mutant MINT (100 µM) and MVP 3-5 (90 µM) constructs were individually injected onto a Superdex 75 increase 10/300 GL column (Cytiva) to establish the

peak height and elution volume of each construct. To assess complex formation, wild-type MINT (100 μM) was incubated with wild-type MVP 3-5 (90 μM) overnight at 4°C. The mixture was then injected onto the Superdex 75 increase column. This process was repeated with a mixture of the mutant MINT and MVP 3-5 constructs, as well as with combinations of each WT and mutant construct. 2X Laemmli loading dye was added to peak fraction samples, which were then boiled, and loaded onto 15% SDS-PAGE gels. Bands were resolved with Coomassie stain.

2.6.9 Simulation System and Molecular Dynamics Simulation

The all-atom simulation system was built using a dimer of the SPFH domain (residues: 419-646) from the atomic cryo-EM structure of the vault cage with an NAD^+ molecule in the binding pocket. Two MVP SPFH domain monomers and the NAD^+ molecule between them were solvated in a water box using the VMD program. Titratable residues of the protein were assigned to their default protonation state at neutral pH. Potassium and chloride ions were added into the water box to make the simulation system electronically neutral and reach a final concentration of 100 mM KCl. The resulting system contains ~55,600 atoms with orthorhombic box dimensions of ~70 x 80 x 95 $Å^3$.

The molecular dynamics simulation was performed using the program NAMD. The CHARMM36 force field with torsional backbone corrections was used for protein, ion, NAD, and the (transferable intermolecular potential 3P) TIP3P model for water. The system temperature and pressure were maintained at 300 K and 1 atm using the Langevin dynamics and the Nose−Hoover Langevin piston method, respectively. The long-range electrostatic interaction was calculated using the particle mesh Ewald (PME) method, and the van der Waals interaction was gradually switched off at 10−12 Å. A time step of 2 fs was used.

The simulation system was initially minimized for 5,000 steps, and then equilibrated for 50 ns with 3 and 7 restraints being applied to the protein and the NAD molecule, respectively, to restrain the conformation, position and orientation of the protein, and the conformation and relative position/orientation of NAD with respect to the protein. The restraints were applied using the Colvars module, with the collective variables being generated using the Binding Free-Energy Estimator 2 (BFEE2) software. The last 10 ns simulation trajectory (1,000 snapshots) was used to calculate the average pairwise interaction between NAD and each protein residue using NAMD.

2.6.10 Mass Spectrometry for ADPr Modification Detection

MVP and MVP-PARP4 particles were incubated with NAD^+ as described above, after which they were diluted two-fold with Laemmli buffer and loaded into a 12% SDS-PAGE gel. Gel bands were chopped to 1-2 mm^3 pieces and destained using three 20 min incubations/changes of 50% acetonitrile, 25 mM ammonium bicarbonate (ABC). Once all stain was removed, gel slices were reduced with 10mM DTT in 25mM ABC and alkylated with 55 mM 2-Iodoacetamide (IAA). Gel pieces were washed with 25 mM ABC, dehydrated in 100% acetonitrile, and then air-dried for 10 minutes or to complete dryness. Gel pieces were re-hydrated in 50 mM ABC containing 5 µg/mL Trypsin (Promega), then incubated at 37 °C overnight. For extraction of peptides from the gel pieces, 100 µl 50% acetonitrile, 0.1% formic acid (FA) was added, and samples were incubated while shaking at room temperature for 20 min. Samples were centrifuged at 20,000 *g* for 5 minutes and the supernatant was transferred to a clean tube. This step was repeated two more times. The eluates were combined and dried. Upon reconstituting with 20 µl of 5% acetonitrile in 0.1% FA in water the samples were desalted using

a C18 zip-tip. After washing with 5% acetonitrile in 0.1% FA twice, the peptides were eluted with 50% ACN, 0.1% FA twice. Then the eluates were dried using a vacuum centrifugation system (Eppendorf). The peptides were suspended in 10 uL 5% acetonitrile, 0.1% FA buffer for LC-MS.

5 µl of digested peptides were analyzed using Q Exactive HF mass spectrometer coupled with an UltiMate 3000 RSLC nanosystem with a Nanospray Frex Ion Source (Thermo Fisher Scientific). Digested peptides were loaded into a Waters nanoEase M/Z C18 (100 Å, 5 µM, 180 µM x 20 mm) trap column and then a 75 µm x 150 mm Waters BEH C18 (130 Å, 1.7 µM, 75 µM x 15 cm) and separated at a flow rate of 300 nL/min. Solvent A was 0.1% FA in water and solvent B was 0.1% FA, 80% ACN in water. The solvent gradient of LC was 5% B in 0-3 minutes, 10% B in 3.2 minutes, 10-35% B in 29 minutes, 35-95% B in 30 minutes, wash 95% in 35 minutes, followed by 5% B equilibration until 45 minutes.

Full MS scans were acquired in the Q-Exactive mass spectrometer over 374-1400 m/z range with a resolution of 120,000 (at 200 m/z) from 5 minutes to 45 minutes. The AGC target value was 3.00E+06 for the full scan. The 15 most intense peaks with charge states 2, 3, 4, 5 were fragmented in the HCD collision cell with a normalized collision energy of 28%, these peaks were then excluded for the 30s within a mass window of 1.2 m/z. A tandem mass spectrum was acquired in the mass analyzer with a resolution of 60,000. The AGC target value was 1.00E+05. The ion selection threshold was 1.00E+04 counts, and the maximum allowed ion injection time was 50 ms for full scans and 50 ms for fragment ion scans.

Spectra were searched against the Uniprot human database using Mascot daemon (2.0.3.0, updated on 08/11/20) with the following parameters: parent mass tolerance of 10 ppm, constant medication on cysteine alkylation, variable modification on methionine oxidation,

deamidation of asparagine and glutamine and ADP-ribosylation on glutamate, aspartate, cysteine, arginine, lysine, histidine, serine and tyrosine. Search results were entered into Scaffold Q+S software (v5.1.2, Proteome Software, Portland, OR) for compilation, normalization, and comparison of spectral counts, etc. The filtering criteria of protein identification were a 1% false discovery rate (FDR) of protein and peptide with 1 minimum peptide count.

2.6.11 A549 Cell Culture and PARP4 KD Cell Line Generation

A549 cells were cultured in DMEM (MilliporeSigma) supplemented with 10% fetal bovine calf serum (Cytiva), 100 U ml^{-1} penicillin, and 100 µg ml^{-1} streptomycin in a humidified incubator (with 95% air, 5% CO_2) at 37 °C. Mutant cells were generated via CRISPR–Cas9 transfection. A guide RNA targeting exon 1 of PARP4 (sense primer 5′-CACCGCTGGGTTTGCAATATGAACG-3′ and anti-sense primer 5'-AAACCGTTCATATTGCAAACCCAGC-3'), was annealed into the pSpCas9(BB)-2A-GFP plasmid (Addgene PX458, gift from the Keenan Lab). The guide RNA-Cas9 plasmid was transfected into A549 cells using Lipofectamine 3000 (ThermoFisher) and a cell line-specific protocol provided by ThermoFisher [110]. Cells were viewed under a fluorescence microscope to verify that transfection had occurred, then trypsinized and pooled for fluorescence assisted cell sorting (FACS) at the University of Chicago Flow Cytometry Core. Each cell that exhibited low-level fluorescence was selected and transferred to an individual well of a 96-well plate. Plates were transferred to a 37 °C incubator and kept there overnight, after which the cells were supplemented with fresh media. Single clone cell lines were expanded to larger plates as needed.

The genomic DNA of the cell lines that persisted was isolated and genotyped to identify mutants. Western blot assays were performed to verify protein-level depletion (Figure A.1).

2.6.12 anti-MVP Fragment Antigen Binding Protein (Fab) Selection and Purification

Human MVP was expressed and purified as described above. The purity of the human vault cage was established by SDS-PAGE gel and visualization via negative stain (Figure 2S.4b). MVP was purified in amine-free buffer (150 mM NaCl, 50 mM HEPES pH 7.4, 1.5 mM $MgCl_2$, 0.1 mM TCEP-HCl) prior to submission to the Kossiakoff lab for Fab selection using a phage display library. MVP was biotinylated and immobilized on streptavidin beads, at which point the beads were washed with selection buffer and subjected to four rounds of phage display panning at progressively lower concentrations of MVP (200-10 nM). The nineteen Fab candidates that emerged from the final round of clones were subjected to a phage ELISA assay (Figure 2S.4c), and the twelve that displayed the greatest affinity and specificity for MVP were subcloned into the RH2.2 expression vector for subsequent purification. High affinity Fabs were observed to bind the MVP cage cap region by cryo-EM imaging (Figure 2S.4d).

Our anti-GFP control and anti-hMVP fabs were subcloned into an RH2.2 vector featuring an Avi-tag and purified for our coimmunoprecipitation experiments. Briefly, RH2.2-Avi-Fab plasmids were transformed into BL21-BirA *Escherichia coli* competent cells, which were spread over ampicillin/chloramphenicol-resistant agar plates. A single colony was used to inoculate a 50 ml starting culture, which was grown for 6-8 hours. 12 ml of the culture were transferred into 1 L of autoinduction media – autoclaved Luria Broth with 1X NPS (25 mM $(NH_4)_2SO_4$ 50 mM KH_2PO_4, 50 mM Na_2HPO_4), 1X 5052 (1 g glycerol, 0.1 g glucose, and 0.4 g lactose), and 1 mM

$MgSO_4$ – with 50 μM of biotin added. Cells were grown in autoinduction media overnight at 30°C, then pelleted at 5000 rpm for 15 minutes.

Pellets were resuspended in 30 ml lysis buffer (20 mM Tris pH 7.5, 150 mM NaCl) and the resulting solution was sonicated at 50% amplitude, with a 1s on/1s off cycle for 4 minutes. Sonicant was incubated in a 63°C water bath for 30 minutes to precipitate contaminant proteins and Fab degradation proteins, and the lysate was cleared by a 40-minute spin at 38,000 *g*. The supernatant was filtered and transferred to a 50 ml superloop (GE Healthcare), to be loaded onto a 5 ml HiTrap Protein L column (Cytiva). Fabs were eluted with 0.1 M acetic acid. Fab-containing fractions were immediately consolidated and loaded onto a 5 ml Mono S 4.6/100 PE cation exchange chromatography column (GE Healthcare). Fabs were eluted via salt gradient and immediately dialyzed into lysis buffer overnight.

2.6.13 Vault Immunoprecipitation and Mass Spectrometry

WT and PARP4 KD A549 cells were seeded (at 7.5×10^5 cells per 15-cm plate) three days prior to the experiment. Two 15-cm plates were seeded per condition replicate (for a total of 24 plates). On the day of the experiment, cells were washed with 8 and then 4 ml ice cold PBS, then subjected to lysis with 1.7 ml NETN buffer (150 mM NaCl, 50 mM Tris pH 7.5, 5 mM EDTA, 0.5% NP-40, 1 mM DTT, 1 mM PMSF, 1 complete protease inhibitor tablet (Roche) per 20 ml buffer, and 250 nM ADP-HPD PARG inhibitor) per plate. Cells were incubated with NETN buffer at 4°C for 6 minutes, at which point the lysate was collected with a cell scraper and transferred to a 2 ml centrifugation tube. Tubes were vortexed for 30 s and kept on ice for an additional 15 minutes. Cell debris and nuclei were pelleted for 8 minutes at 13,000 rpm, after which the lysate supernatant was transferred to a fresh 2 ml tube and placed on ice. Following a

Bradford assay to quantify total lysate protein material per replicate, volumes containing 6 mg lysate protein were used as starting material for each coimmunoprecipitation replicate.

During harvest, a total of 55 mg per replicate of biotinylated anti-hMVP or anti-GFP (control) Fabs were incubated with magnetic streptavidin beads at 4°C. After being washed with NETN buffer three times, the Fab-bead mixtures were added to harvested lysate and incubated in an overhead rotator at 4°C for two hours. Samples were then subjected to a stringent series of eight washes by NETN buffers with varying NaCl concentrations. The wash cycle started with the highest (2X at 500 mM) NaCl concentration and progressed to lower (3X at 300 mM), then lowest (3X at 150 mM) concentration buffers. All washes were conducted on an overhead rotator at 4°C. Finally, immunoprecipitated proteins were briefly washed with 900 µl ddH2O to eliminate buffer contaminants, then eluted from the streptavidin beads with 82 µl 0.1 M acetic acid, pH 4.5, incubated at 37°C. Elutions were immediately stored at -80°C for subsequent MS analysis. Samples were reduced, alkylated, and trypsin-digested overnight. Peptides were eluted and desalted with C18 stagetip, then sent for MS analysis. Data was searched with Proteome Discoverer 3.0, normalized with Perseus 2.0.10.0, and analyzed using Microsoft Excel and R Studio.

2.7 Supplementary Figures

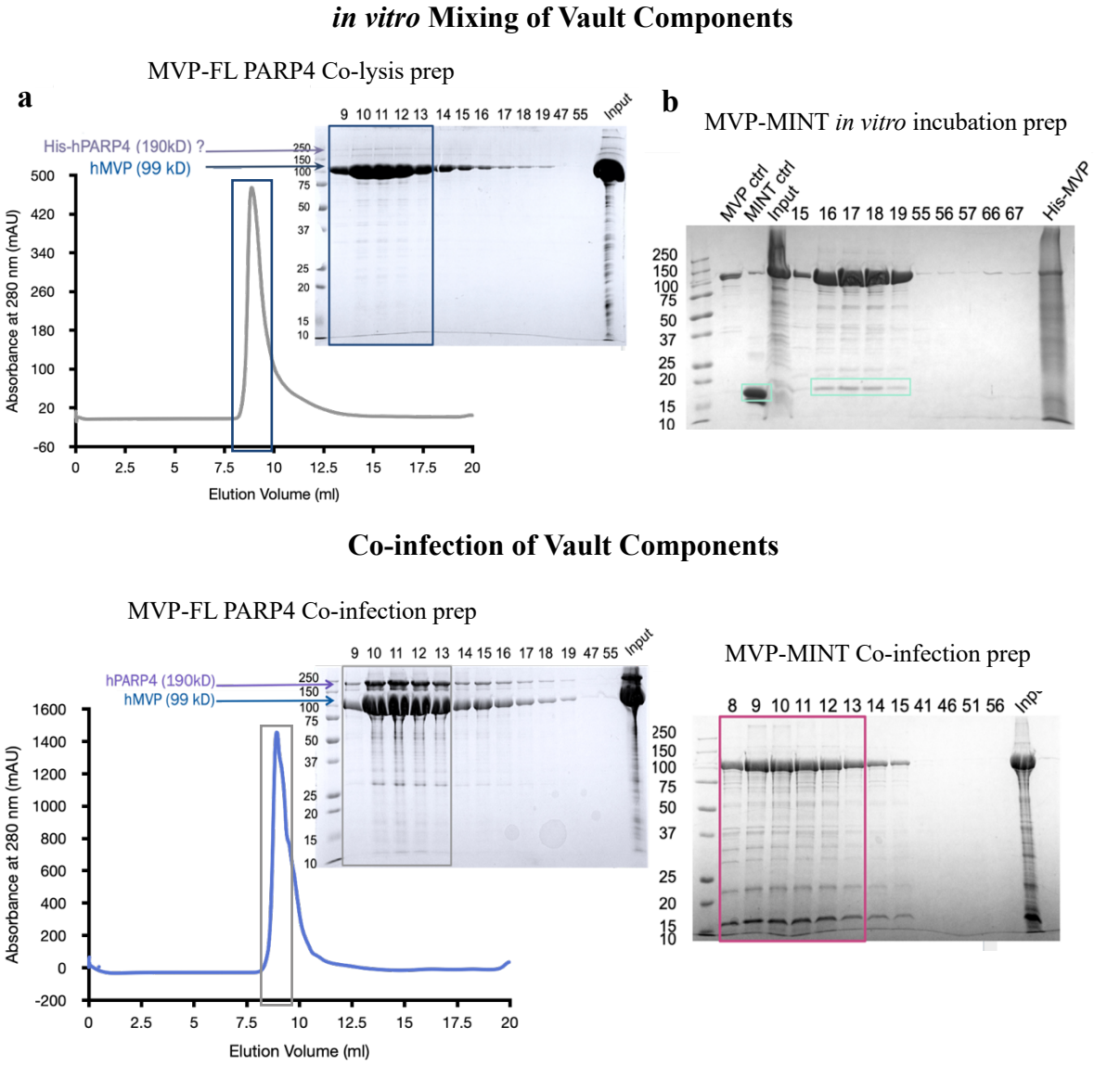


Figure 2S.1 PARP4 and MINT are not efficiently incorporated into the vault cage in vitro

a, SEC chromatogram of MVP-PARP4 complex purified via insect cell co-lysis (upper) and co-infection (lower) preps. SDS-PAGE gels indicate that PARP4 can only be efficiently internalized into the MVP cage when they are co-expressed in insect cells (following co-infection of their respective baculoviruses). **b,** Gels comparing MINT domain uptake into the MVP cage following co-incubation of MVP and MINT *in vitro* versus purification from co-infected insect cells. Some MINT (boxed in light green) can be internalized by the MVP cage *in vitro,* but uptake is more efficient if the complex is purified from insect cells co-infected with MINT and MVP viruses (complex boxed in pink).

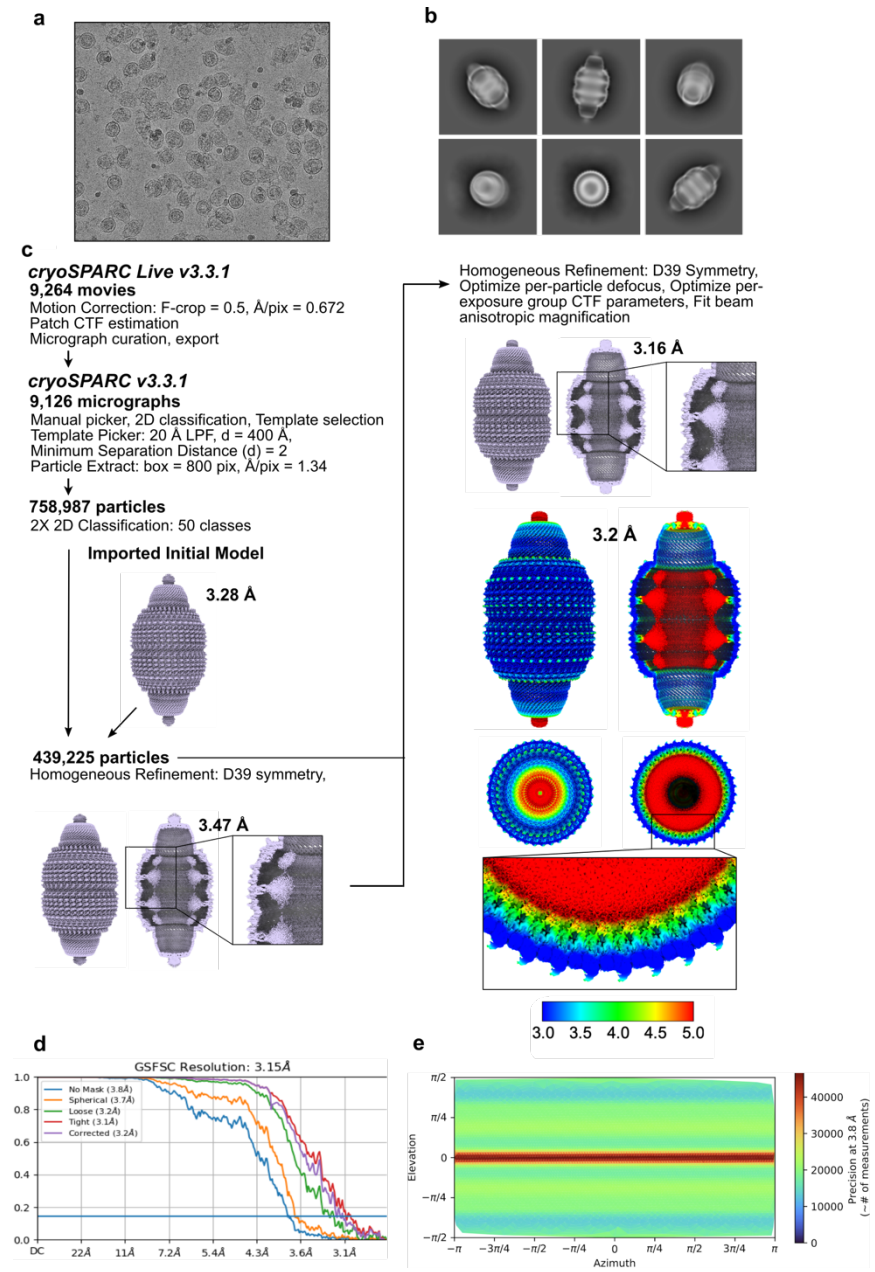


Figure 2S.2 Data processing workflow for the MVP-PARP4-NAD$^+$ structure

a, A representative motion-corrected electron micrograph from the hMVP-PARP4-NAD$^+$ dataset (n = 9,126), collected on a Titan Krios microscope. **b**, A subset of initial 2D class averages, following template particle picking. **c,** The image processing workflow for hMVP-PARP4-NAD$^+$. Initial processing was completed using cryoSPARC LIVE v3.3.1. Selected particles were exported to cryoSPARC v3.3.1. Exported particles were subjected to several rounds of homogeneous refinement, beginning with an initial model resolved from a previous dataset. The final map is shown with local resolution values calculated in cryoSPARC v4. **d**, Fourier Shell Correlation (FSC) curves as calculated in cryoSPARC. **e**, Euler angle distributions.

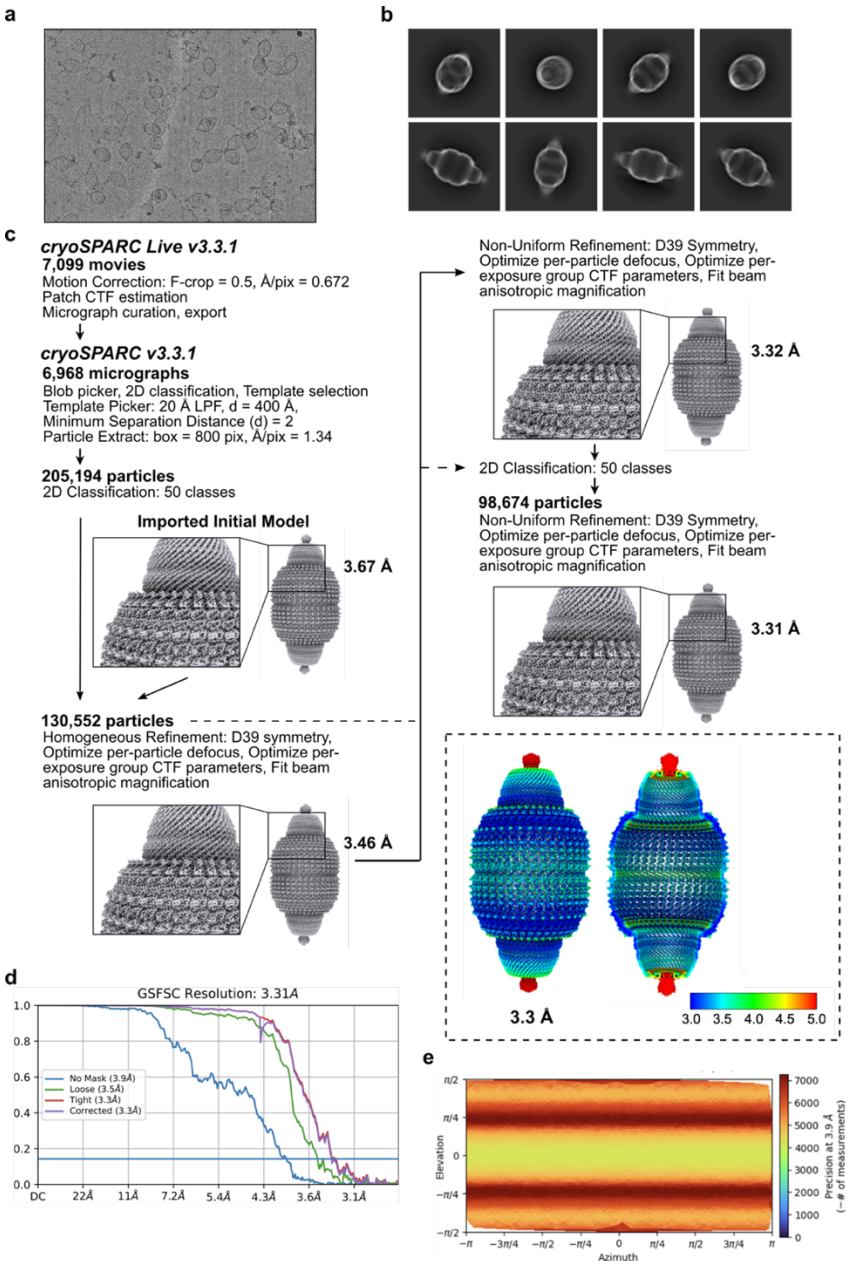
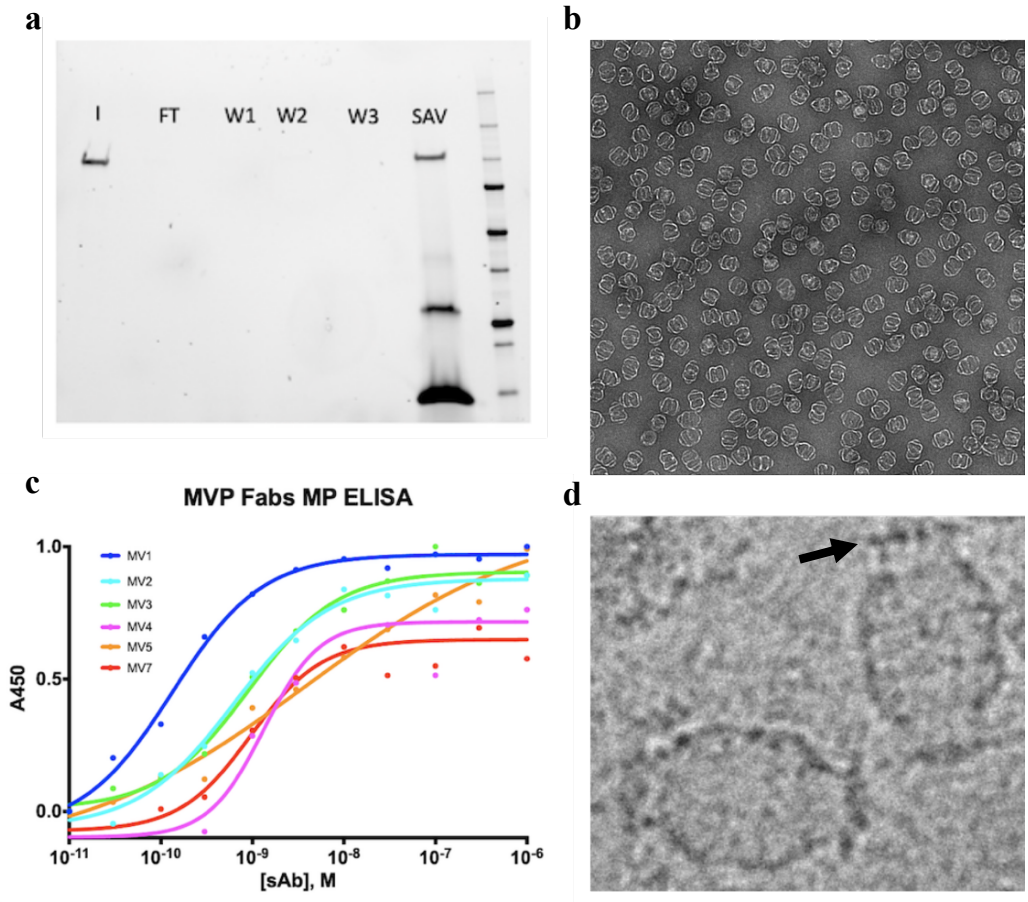


Figure 2S.3 Data processing workflow for the apo-hMVP structure

a, A representative motion-corrected electron micrograph from the apo-hMVP dataset (n = 6,968), collected on a Titan Krios microscope. **b**, A subset of initial, reference-free 2D class averages, following "blob" particle picking. **c,** The image processing workflow for apo-hMVP. Initial processing was completed using cryoSPARC LIVE v3.3.1. Selected particles were exported to cryoSPARC v3.3.1. Exported particles were subjected to several rounds of homogeneous and non-uniform refinement, beginning with an initial model resolved from a previous dataset. The final map is shown with local resolution values calculated in cryoSPARC v4. **d**, Fourier Shell Correlation (FSC) curves as calculated in cryoSPARC. **e**, Euler angle distributions.



e

Fab	MV1	MV2	MV3	MV4	MV5	MV7
IC50 (nM)	0.133	0.627	0.908	1.33	3.8102	0.963

Figure 2S.4 Fab selection for the hMVP cage

a, Gel depicting purified hMVP (far left, "I" lane) and pulldown demonstrating successful binding to streptavidin beads (far right, "SAV" lane) for immobilization during phage display. **b,** Negative stain electron microscopy image of purified vault cage sample submitted for Fab selection. **c,** Phage ELISA assay showing fluorescence associated with each Fab's affinity for hMVP. Fabs were named "MV" (and later, "VM") and numbered according to relative affinity. **d,** cryo-EM images depicting Fabs (shown by arrow) bound to the caps of some MVP-MINT particles, visualized by the Talos 200 kV microscope. **e,** Table of Fab-hMVP binding IC50 values to compare their affinities. **a & c provided by Dr. Satchal Erramilli**

a

Unmodified peptide

B	B Ions	B+2H	B-NH3	B-H2O	AA		Y Ions	Y+2H	Y-NH3	Y-H2O	Y
1	114.1					I	2,152.1	1,076.6	2,135.1	2,134.1	18
2	211.1					P	2,039.0	1,020.0	2,022.0	2,021.0	17
3	308.2					P	1,942.0	971.5	1,924.9	1,923.9	16
4	471.3			453.2		Y	1,844.9	923.0	1,827.9	1,826.9	15
5	608.3	304.7		590.3		H	1,681.8	841.4	1,664.8	1,663.8	14
6	771.4	386.2		753.4		Y	1,544.8	772.9	1,527.8	1,526.8	13
7	884.5	442.7		866.5		I	1,381.7	691.4	1,364.7	1,363.7	12
8	1,021.5	511.3		1,003.5		H	1,268.6	634.8	1,251.6	1,250.6	11
9	1,120.6	560.8		1,102.6		V	1,131.6	566.3	1,114.5	1,113.6	10
10	1,233.7	617.3		1,215.7		L	1,032.5	516.8	1,015.5	1,014.5	9
11	1,348.7	674.9		1,330.7		D	919.4	460.2	902.4	901.4	8
12	1,476.8	738.9	1,459.7	1,458.8		Q	804.4	402.7	787.4	786.4	7
13	1,590.8	795.9	1,573.8	1,572.8		N	676.3	338.7	659.3	658.3	6
14	1,677.8	839.4	1,660.8	1,659.8		S	562.3		545.3	544.3	5
15	1,791.9	896.4	1,774.9	1,773.9		N	475.3		458.2	457.3	4
16	1,890.9	946.0	1,873.9	1,872.9		V	361.2		344.2	343.2	3
17	1,978.0	989.5	1,961.0	1,960.0		S	262.2		245.1	244.1	2
18	2,152.1	1,076.6	2,135.1	2,134.1		R	175.1		158.1		1

B	B Ions	B+2H	B-NH3	B-H2O	AA		Y Ions	Y+2H	Y-NH3	Y-H2O	Y
1	114.1	57.5				I	2,152.1	1,076.6	2,135.1	2,134.1	18
2	211.1	106.1				P	2,039.0	1,020.0	2,022.0	2,021.0	17
3	308.2	154.6				P	1,942.0	971.5	1,924.9	1,923.9	16
4	471.3	236.1		453.2		Y	1,844.9	923.0	1,827.9	1,826.9	15
5	608.3	304.7		590.3		H	1,681.8	841.4	1,664.8	1,663.8	14
6	771.4	386.2		753.4		Y	1,544.8	772.9	1,527.8	1,526.8	13
7	884.5	442.7		866.5		I	1,381.7	691.4	1,364.7	1,363.7	12
8	1,021.5	511.3		1,003.5		H	1,268.6	634.8	1,251.6	1,250.6	11
9	1,120.6	560.8		1,102.6		V	1,131.6	566.3	1,114.5	1,113.6	10
10	1,233.7	617.3		1,215.7		L	1,032.5	516.8	1,015.5	1,014.5	9
11	1,348.7	674.9		1,330.7		D	919.4	460.2	902.4	901.4	8
12	1,476.8	738.9	1,459.7	1,458.8		Q	804.4	402.7	787.4	786.4	7
13	1,590.8	795.9	1,573.8	1,572.8		N	676.3	338.7	659.3	658.3	6
14	1,677.8	839.4	1,660.8	1,659.8		S	562.3	281.7	545.3	544.3	5
15	1,791.9	896.4	1,774.9	1,773.9		N	475.3	238.1	458.2	457.3	4
16	1,890.9	946.0	1,873.9	1,872.9		V	361.2	181.1	344.2	343.2	3
17	1,978.0	989.5	1,961.0	1,960.0		S	262.2	131.6	245.1	244.1	2
18	2,152.1	1,076.6	2,135.1	2,134.1		R	175.1	88.1	158.1		1

Peptide with apparent ADPr mass shift signature

B	B Ions	B+2H	B-NH3	B-H2O	AA		Y Ions	Y+2H	Y-NH3	Y-H2O	Y
1	114.1	57.5				I	2,694.1	1,347.6	2,677.1	2,676.1	18
2	211.1	106.1				P	2,581.1	1,291.0	2,564.0	2,563.0	17
3	308.2	154.6				P	2,484.0	1,242.5	2,467.0	2,466.0	16
4	471.3	236.1		453.2		Y	2,386.9	1,194.0	2,369.9	2,368.9	15
5	608.3	304.7		590.3		H	2,223.9	1,112.4	2,206.9	2,205.9	14
6	771.4	386.2		753.4		Y	2,086.8	1,043.9	2,069.8	2,068.8	13
7	884.5	442.7		866.5		I	1,923.8	962.4	1,906.7	1,905.8	12
8	1,021.5	511.3		1,003.5		H	1,810.7	905.8	1,793.7	1,792.7	11
9	1,120.6	560.8		1,102.6		V	1,673.6	837.3	1,656.6	1,655.6	10
10	1,233.7	617.3		1,215.7		L	1,574.6	787.8	1,557.5	1,556.5	9
11	1,348.7	674.9		1,330.7		D	1,461.5	731.2	1,444.4	1,443.5	8
12	1,477.7	739.4	1,460.7	1,459.7		Q+1	1,346.4	673.7	1,329.4	1,328.4	7
13	1,591.8	796.4	1,574.8	1,573.8		N	1,217.4	609.2	1,200.4	1,199.4	6
14	2,219.9	1,110.4	2,202.9	2,201.9		S+541	1,103.4	552.2	1,086.3	1,085.3	5
15	2,333.9	1,167.5	2,316.9	2,315.9		N	475.3	238.1	458.2	457.3	4
16	2,433.0	1,217.0	2,416.0	2,415.0		V	361.2	181.1	344.2	343.2	3
17	2,520.0	1,260.5	2,503.0	2,502.0		S	262.2	131.6	245.1	244.1	2
18	2,694.1	1,347.6	2,677.1	2,676.1		R	175.1	88.1	158.1		1

b

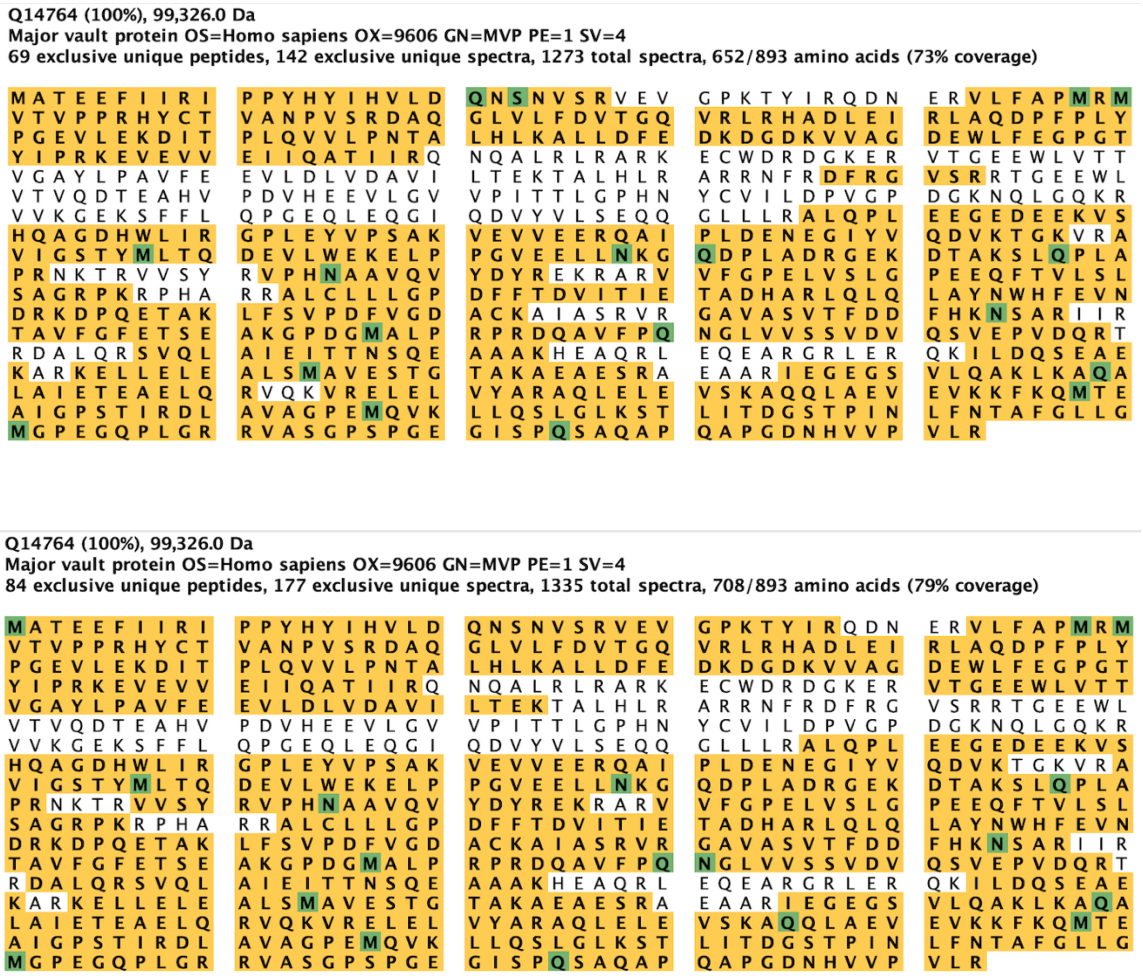


Figure 2S.5 MS fragmentation tables of unmodified vs. modified MVP peptides

a, Fragmentation tables showing the mass/charge (m/z) values of the peptide spanning the hMVP sequence I10 to R27. The B ion column shows the additive m/z values for amino acids when measured from the N-to C-terminus of the peptide in question, while the Y ions represent the sum of m/z values measured in the opposite direction. Highlighted values represent fragments detected by initial MS, while the remainder were detected by MS2. Top table shows m/z values of an unmodified peptide in which all fragments could be assigned with 100% probability. The central and lower tables show peptide fragments detected by MS2 that appear to be either unmodified (middle) or exhibiting an apparent ADP-ribose modification on S23 (lower). **b,** Color-coded sequences mapping the sequence coverage (yellow residues) and modified residues (green) observed in MVP alone (upper) and MVP in the presence of PARP4 (lower). Coverage of the peptide spanning I10 to R27 was achieved for each condition.

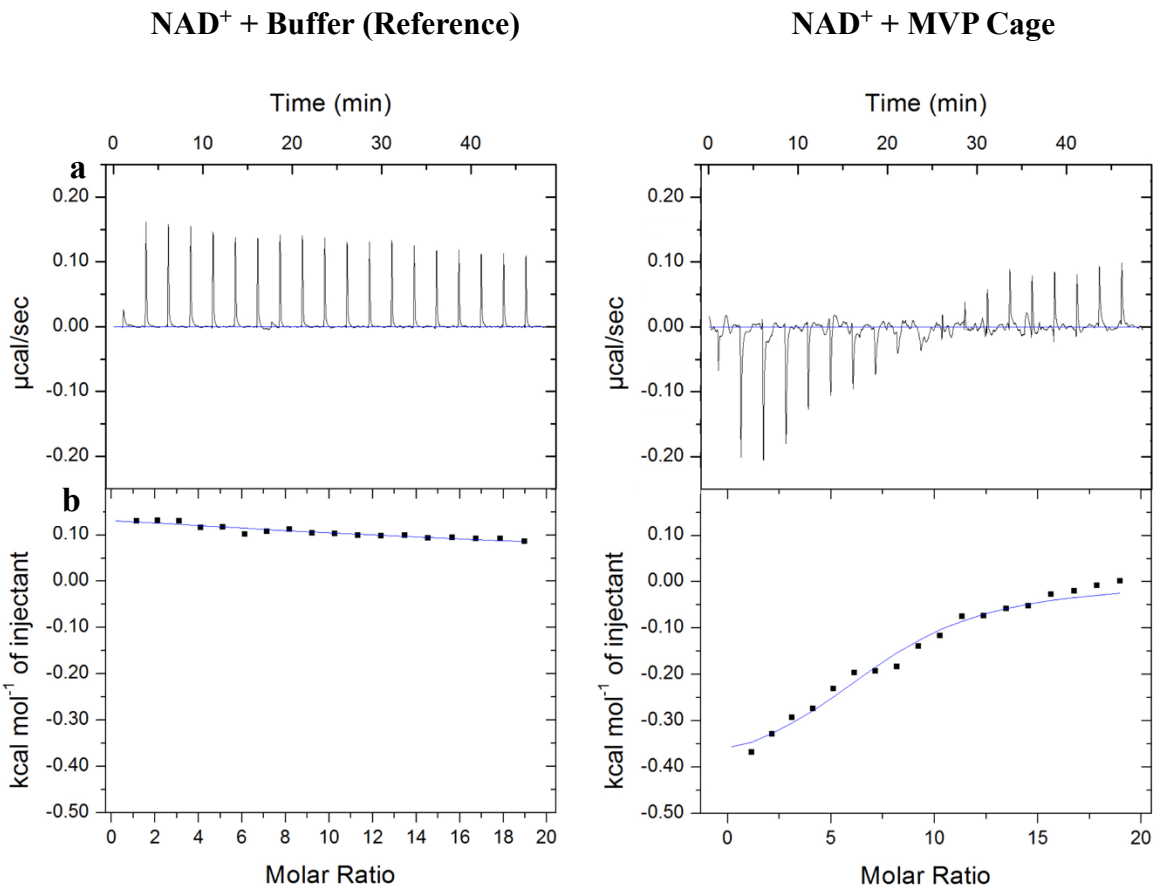


Figure 2S.6 NAD$^+$ specifically binds the hMVP cage

a, 2 µl titrations of 5 mM NAD$^+$ were injected 18 times (following an initial 0.4 µl injection) into a cell containing MVP purification buffer, inducing a dose-independent endothermic response (left). Identical concentrations and volumes of NAD$^+$ were injected into a solution of the MVP cage complex (with MVP monomer concentration of 50 µM), inducing a dose-dependent exothermic response (right). **b,** Line fits to normalized ITC data. First injection values were disregarded, and the MVP + NAD$^+$ binding isotherm is reference-subtracted, with adjacent averaging smoothing applied.

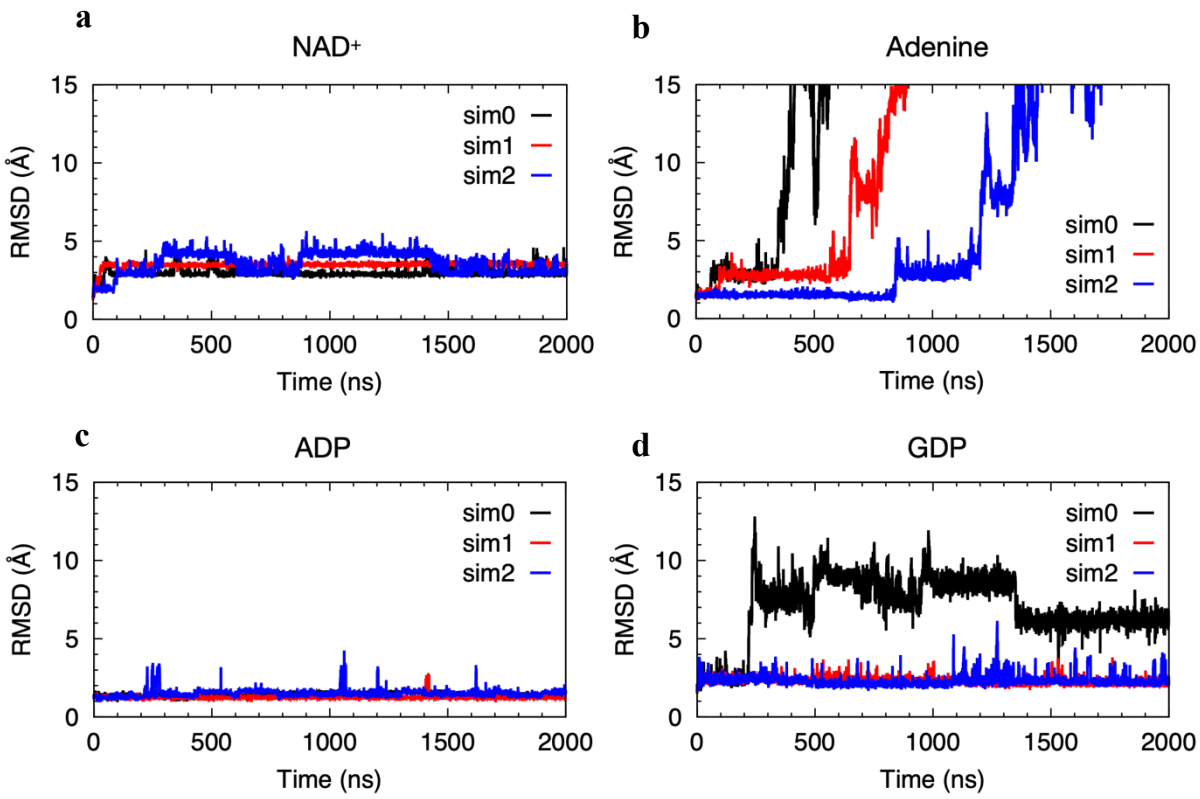


Figure 2S.7 Nucleotide diphosphate molecules stably occupy the NAD$^+$-binding pocket

a, 2 μs RMSD simulation of NAD$^+$ and our cryo-EM-derived model of the MVP binding pocket. The simulation was repeated three times. Small RMSD fluctuations (≤5 Å) indicate base-flipping, while larger deviations suggest that the metabolite has left the binding pocket. **b-d**, Identical treatments of adenine, ADP, and GDP suggest that a purine base with ribose and diphosphate moieties is sufficient for stable occupancy of MVP's NAD$^+$-binding pocket.

CHAPTER 3

TOWARDS IDENTIFYING THE MODIFICATION SUBSTRATES OF VAULT-BOUND PARP4

3.1 Abstract

Poly (ADP-ribose) polymerases comprise a superfamily of transferases that deposit post-translational modifications (PTMs) onto macromolecular targets. They consume NAD^+ to transfer its ADP-ribose (ADPr) moiety (either single ADPr molecules – MAR – or longer polymers – PAR) onto polar and charged units of polypeptides and nucleic acids. PARPs are best known for their roles in DNA repair, but, of the seventeen PARPs expressed by humans (sixteen of which exhibit catalytic activity), only three are primarily associated with the DNA damage response. The remainder are found throughout the cell and possess a wide range of functions, which are best understood by identifying the targets of each PARP. Recent progress in reagent development and chemical biology strategies for identifying and isolating MAR- and PAR-modified molecules has greatly accelerated the discovery of native PARP substrates over the past ten years [45, 111]. To date, *in vivo* targets of almost every PARP have been identified. The fourth member of the family, a member of the enigmatic vault complex, is the exception in this respect. It has been observed to modify itself and the primary component of the enigmatic vault complex to which it belongs *in vitro*. However, PARP4's presence among a diverse set of molecules (both within and outside of the vault), suggests it likely targets a larger repertoire of molecules. Here, we will outline our initial efforts towards using chemical biology to determine the native substrates of PARP4.

3.2 Introduction

ADPrylation mediated by the three most abundant nuclear PARPs (PARPs 1-3) has long been studied in the context of its role in DNA base excision repair and, consequently, cancer. In recent years, however, the biology and medical relevance of the remaining 13 active PARPs has generated increasing interest. This curiosity has only intensified as their roles in antiviral activity and various cancer phenotypes have begun to be uncovered. Studying the impact of their modifications in these processes, however, has been a challenge due to the outsized activity of PARPs -1 and -2. PARP1, in particular, has been estimated to produce ~90% of PAR formed in the cell [112], and recent studies demonstrate that, under certain conditions [113, 114], it can switch to primarily depositing MAR onto its targets. As a result, even assays employing reagents that have been engineered to specifically identify MAR modifications can be swamped by PARP1 signal, obscuring the enzymatic impact of the remaining PARP enzymes.

In order to specifically ascribe modified targets to the correct PARP, a number of protein-ligand engineering ("bump-in-hole") approaches have been developed, exploiting the fundamental similarities between PARP enzyme active sites as well as the relative chemical complexity of the ADPr molecule [115]. The catalytic domain structures of eleven human PARPs have been deposited into the protein data bank. Common to each of them is a series of anti-parallel beta strands that comprise the base of the NAD^+-binding pocket, the placement of two conserved residues common to the enzymes' "catalytic triad" (in all but the structure of non-catalytic PARP13), and several bulky residues in conserved positions that outline the PARPs' flexible donor or "D loops" which sit above the bound NAD^+ molecule.

The PARP catalytic triad consists of three residues that interact with specific NAD^+ moieties. The first two residues, a histidine (H) and a tyrosine (Y), are universally conserved among active PARPs, and interact with the "proximal" ribose (adjacent to the adenine in NAD^+) and nicotinamide groups of NAD^+, respectively, while the third residue varies across the PARP family members. The five PARPs with established capacity to deposit PAR chains possess a glutamate in the third position. This residue promotes chain elongation by interacting with the ADPr group that has already been deposited onto the target; namely, it forges a hydrogen bond with the ADPr group's distal ribose (adjacent to the nicotinamide in NAD^+) and subsequently serves as a general base to promote nucleophilic attack of the 2' hydroxyl group on the proximal ribose, creating the conditions for a PAR polymer to form [116]. The remaining PARPs contain a hydrophobic residue (leucine or isoleucine) in the third position. This preserves similar steric constraints within the pocket but precludes chain elongation. (Interestingly, PARP4 possesses the PARylating H-Y-E catalytic motif despite only being observed to have MARylating activity to date.) Outside of the catalytic triad, each PARP's "D-loop" serves to stabilize the entire NAD^+ "donor" prior to its deposition of ADPr onto the substrate, or "acceptor." While the precise sequence and structure of PARP donor loops vary considerably, they all sit above the bound NAD^+ and tend to be flexible. Unsurprisingly, mutations to either of the above PARP features can significantly impact their catalytic activity, even ablating it entirely. Even mutations to some residues adjacent to these features can impact activity, making it imperative that ligand engineering approaches are planned and tested to prevent killing the enzyme activity.

One of the first of these approaches appended an alkylated butylthiol group onto the 8 position of the NAD^+ adenine. Its relative bulk, compared to a standard NAD^+ molecule, facilitated complementarity with an "analog-sensitive" mutant PARP. The asPARP molecule

featured an alanine in place of a bulky "gating" leucine residue adjacent to the D-loop (Figure 3.1c), whose presence would preclude binding of the 8-Butylthio(3-yne)T-NAD^+ molecule in the WT enzyme. Therefore, the introduction of a specific asPARP species along with the NAD^+ analog enables specific deposition of the analog ADPr (which retains the 8-Bu(3-yne) group) onto the asPARP's substrates. The inclusion of the alkyne allows for copper (I)-mediated click reactions conjugating the analog ADPr modification to molecules (including fluorophores or pulldown tags) with an azide group. The conservation of the particular gating residue (L453A in PARP4 – Figure 3.1c) renders the technique generalizable across the PARP superfamily.

Thus far, it has been used to identify the substrates of PARP7 substrates in ovarian cancer cells [51], as well as a comprehensive list of the targets of PARPs -1 to -3 (demonstrating their roles outside of DNA repair) in unstressed cells [45]. The technique's primary limitation stems from NAD^+'s inability (and, therefore, that of its analogs) to efficiently pass through cell membranes. Two strategies have been devised to mitigate this. In the first, recombinant asPARP proteins are introduced to cell lysate. Another approach has been to engineer analog-sensitive nuclear PARPs (PARPs -1 to -3), introduce them to the cells in lentiviral vectors, and introduce the NAD^+ analog to the purified nuclei (whose membranes are not as resistant to passive NAD+ entry). Our system requires an additional approach as PARP4 is a primarily cytosolic protein, which, in our hands, cannot be taken up into the vault particle when added to cell lysate (Figure 2S.1). While it would still undoubtedly be useful to determine the substrates of PARP4 outside of its capacity as a member of the vault complex, a vault-focused study would require that PARP4 be engulfed in the cage.

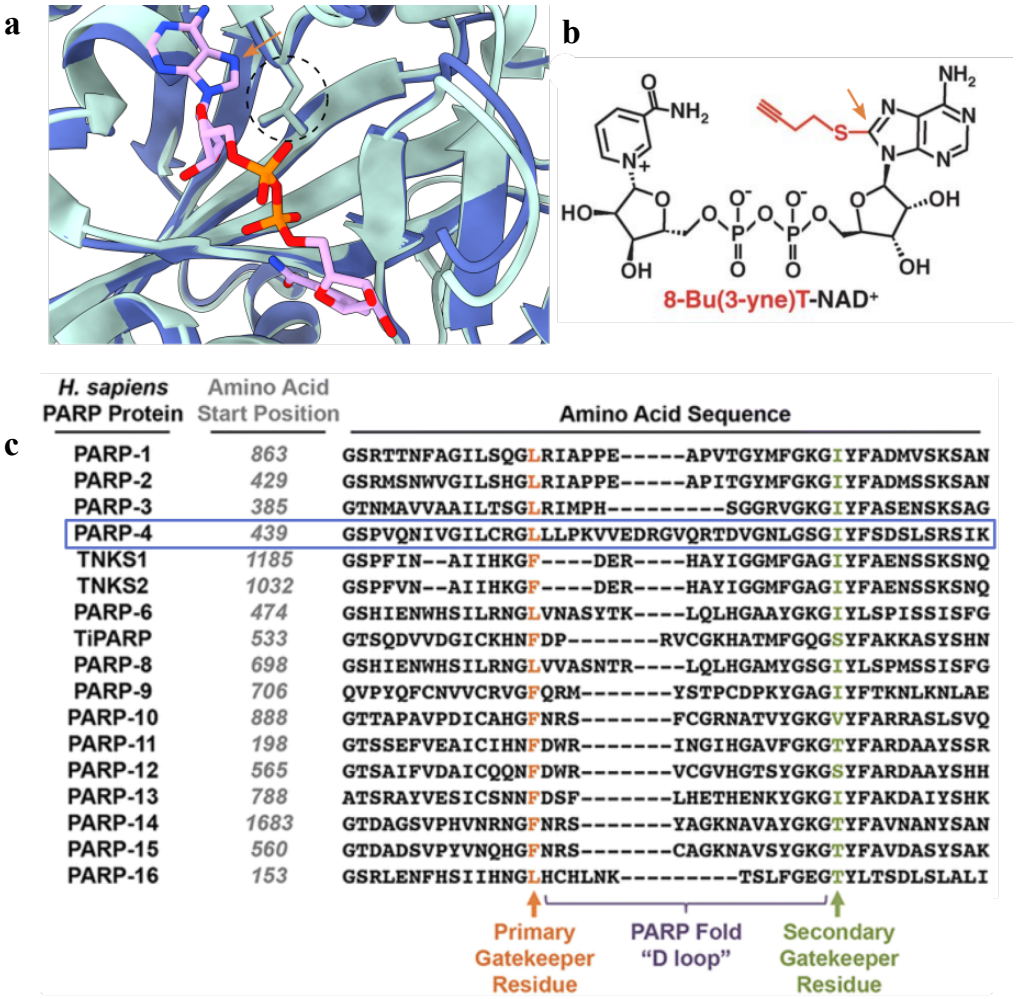


Figure 3.1 Structural basis for asPARP4 construct design

a, Overlay of a PARP1 active site structure (green, accession no. 6bhv) bound to a non-hydrolyzable NAD^+ analog (pink) and the AlphaFold predicted model of the PARP4 catalytic domain (blue). PARP1's L877 "gatekeeper" residue (dashed circle) sits in the exact same position as PARP4's L453 residue in a predicted AlphaFold model of PARP4. **b,** Chemical structure of 8-Bu-NAD^+. **c,** MSA highlighting the bulky "gatekeeper" residues surrounding each PARP enzyme's "D loop." PARP4's sequence is boxed in blue, and its "gatekeepter" target residue, L453, residue is colored in orange.

Therefore, we have conceived two strategies to promote PARP4 vault incorporation in our asPARP4 study. In the first, a plasmid encoding for asPARP4 under the control of a CMV promoter would be transfected into cells. Either the whole-cell lysate, or a fraction containing isolated vault complexes could then be incubated with the 8-Bu-NAD^+, after which the sample could be trypsinized, with its modified peptides clicked to azide-conjugated agarose resin. Alternatively, either the full vault complex or the recombinant MVP-asPARP4 complex could be recombinantly expressed and introduced to the cell lysate. However, it would not necessarily incorporate standard vault cargo, potentially skewing the results. We have started with the first option, showing that we can both successfully express, and incorporate asPARP4 into vault particles in the cell when we transiently overexpress asPARP4 and MVP in HEK293 cells. Planned controls include an asPARP4 construct with targeted alanine mutations imposed on key PARP4-MVP interaction residues so that we can capture PARP4 targets that are not associated with vault. Versions of each construct without the L453A mutation will also be delivered into cells to detect potential false positives.

3.3 Preliminary Results

Before attempting the asPARP4 substrate pulldown in cells, we needed to establish that we could construct an analog-sensitive mutant. Sequence and (predicted) structure alignments suggested that that L453 was the PARP4 residue that best corresponded to the "gatekeeper" residues in PARPs -1 and -2. Because we had previously expressed but failed to purify the full-length PARP4 construct alone from insect cells (Figure B.1), the *in vitro* analog-sensitive activity test was conducted with the PARP4 HD-ART domain. After imposing an L453A mutation, we purified the WT and asPARP4 HD-ART constructs. Both behaved similarly, eluting from a size

exclusion column at the same volume and appearing equally resistant to degradation on SDS-PAGE gels (Figure 3S).

To test their activity and substrate selectivity, we first performed a Western-blot based assay, in which the WT and mutant constructs were incubated with standard NAD^+ and 8-Butylthio(3-yne)T-NAD^+, respectively. Automodification of each construct was detected through use of anti-pan-ADP-ribose reagent. (The ADPr-binding component of the reagent is a macrodomain whose structures suggest its interaction with ADPr would not be occluded by an additional moiety at the 8-position of the adenine.) As expected, WT HD-ART only automodified itself when supplemented with standard NAD+, while the L453A mutant was only active in the presence of the analog (Figure 3.2c). After verifying its activity in the presence of analog, we attempted to subject our analog-modified asHD-ART to an azide-alkyne cycloaddition reaction with an azide-rhodamine fluorophore. We detected clear in-gel fluorescent signal at the molecular weight corresponding to our asHD-ART construct, demonstrating that the L453A PARP4 HD-ART mutant can bind and consume 8-Butylthio(3-yne)T-NAD^+, rendering it a good candidate for an asPARP4 pulldown (Figure 3.2c).

We initially intended to transiently overexpress asPARP4 into A549 PARP4 KD cells to maximize its occupancy in the vault cage, but our attempts to do so were unsuccessful due to the rapid and near-complete death of the cells (versus limited cell death in vector-only transfection controls) over several < 20 hr timeframes. After several efforts to optimize, we attempted expression with (more transfection-robust) HEK293 cells. Because vault protein expression is naturally much lower in HEK293 cells than in A549's, we decided to deliver both asPARP4 and MVP into the cells to maximize the potential yield of vault-associated PARP4 targets.

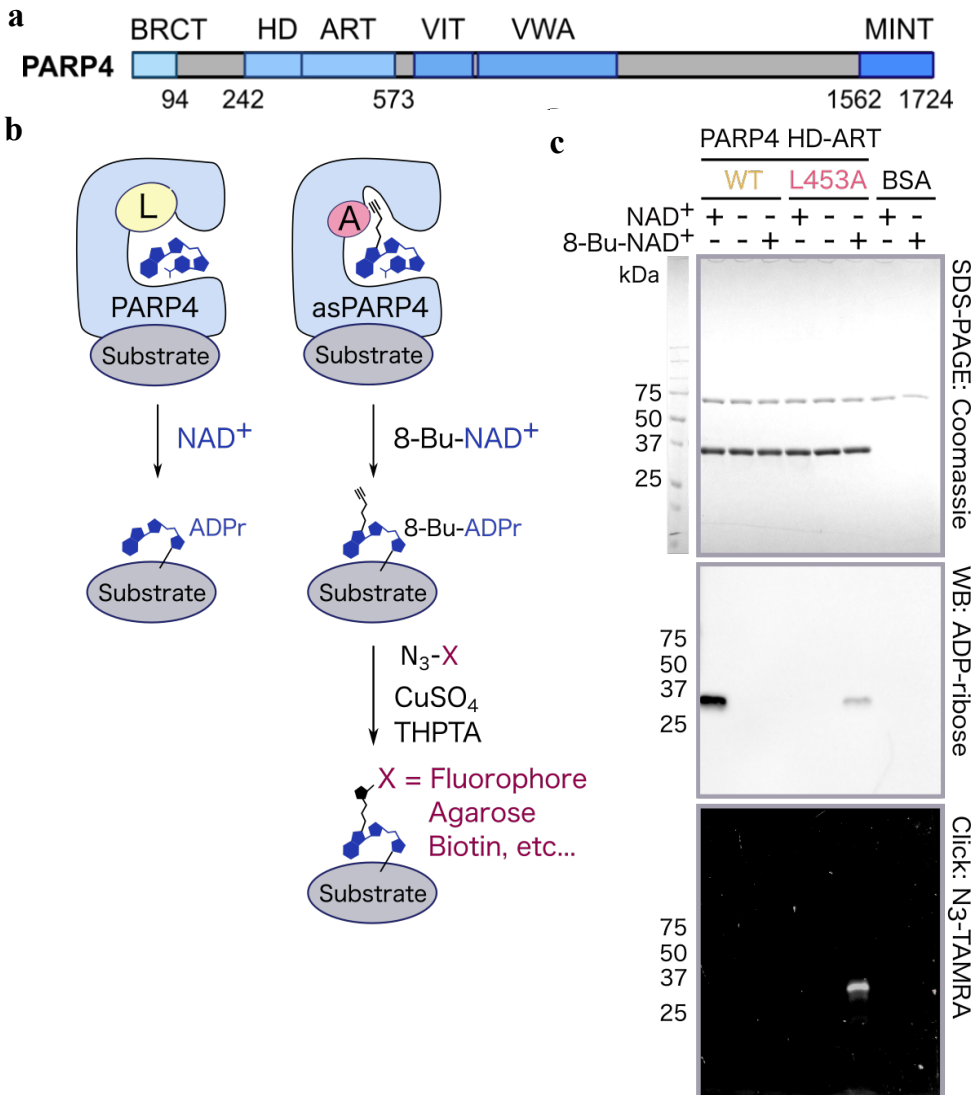


Figure 3.2 Schematic and results of the asPARP4 catalytic domain activity assay

a, Domain diagram of PARP4. **b,** Cartoon depiction of the WT (left) and analog-sensitive (right) PARP4 catalytic domain interacting with unfunctionalized NAD^+ and 8-Bu-NAD^+, respectively. Using click chemistry, various molecules can be conjugated to the ADPr modification for subsequent visualization or isolation of the modified substrate(s). **c,** Proof-of-principle assay demonstrating that the L543A PARP4 catalytic domain mutant can modify substrates with 8-Bu-ADPr, and that a fluorophore can be clicked to those modifications and visualized in an SDS-PAGE gel. SDS-PAGE gel showing protein input for each assay condition (upper panel). ADPr and 8-Bu-ADPr modifications were detected with anti-pan-ADPr reagent (MABE1061) following Western blot transfer (middle), and rhodamine clicked to asPARP4 was detected by in-gel fluorescence excited by the rhodamine green light wavelength settings on the ChemiDoc Imaging System.

The HEK293 cells proved much more tolerant of the vector. When we performed our vault IP (as described in 2.6.13) with asPARP4-transfected HEK293 cells, we could only detect clear PARP4 signal in the cell lysate (input) and IP elution fractions of cells that had been transfected with asPARP4 DNA (Figure 3.3b). Interestingly, we did not detect significant quantities of PARP4 in the elution fraction when we overexpressed MVP without transfecting the asPARP4 vector as well. It is difficult to determine whether this is because endogenous PARP4 is unable to efficiently incorporate into the transiently expressed vault cage or if there is simply too little endogenous PARP4 to observe. Similarly, we did not detect meaningful quantities of endogenous MVP in the elution from the cell samples to which we did not add the MVP expression vector, so we cannot say whether asPARP4 would be efficiently encapsulated into the endogenous vault cage. Regardless, these results have convinced us to co-transfect vectors expressing asPARP4 and MVP to boost any signal we might observe from asPARP4 activity.

We concede that the lack of an additional tag to verify that this upregulation is purely due to our transfected protein does cloud the interpretation of the results. Though we have plans to repeat the experiment with Flag-tagged asPARP4 to confirm our initial findings, we can confidently say that HEK293 cells can tolerate pCMV-asPARP4 transfection and say, with reasonable confidence, that the cells demonstrate an increase in subsequent asPARP4 expression and encapsulation into the vault cage.

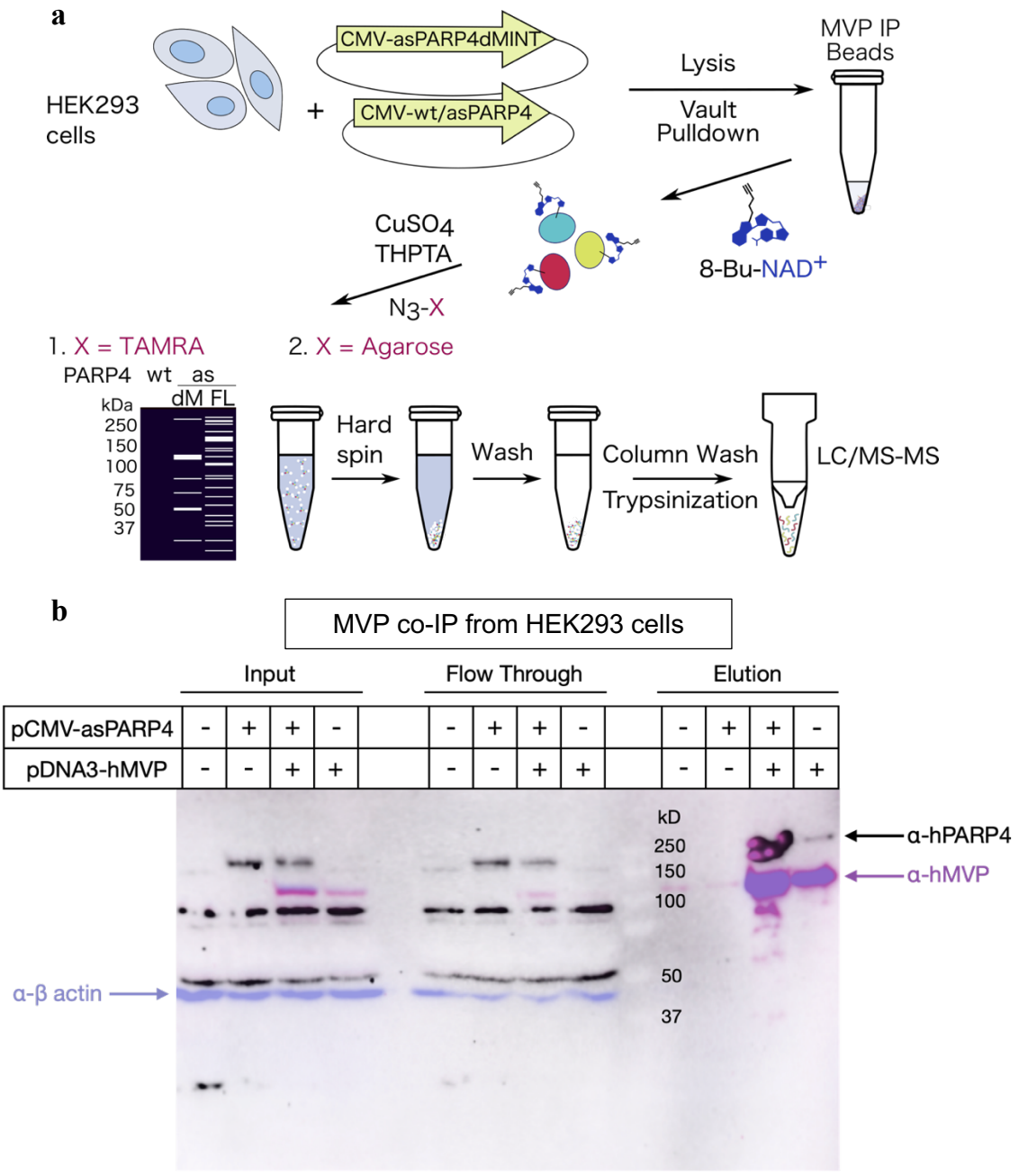


Figure 3.3* in vivo *asPARP4 substrate pulldown

a, Schematic showing planned workflow for the *in vivo* transfection, purification, NAD^+ analog treatment, validation, and click-chemistry pulldown of asPARP4 and an MVP-binding incompetent mutant. **b,** Western blot demonstrating PARP4 detection in HEK293 cells and vault IP samples (as detected by exposure to anti-PARP4 antibody) that were transfected with pCMV-asPARP4 and pcDNA3-MVP. PARP4 signal is shown in black, while MVP and β-actin signal are shown in purple and light blue, respectively.

3.4 Discussion

We have thus far demonstrated the broad applicability of 8-Bu(3-yne)T-NAD^+ click chemistry by verifying that it works even with a PARP enzyme that contains an unusual D-loop (Figure 3.1c). We have also established a viable system for transient PARP4 overexpression in human cells. These efforts will hopefully situate us to discover vault-associated PARP4 modification targets. Such a finding could potentially explain, at least in part, vault's role in signaling pathway mediation. Alternatively, even if PARP4 is not modifying vault cargo and is instead only ADPrylating itself and other vault components, determining those specific sites in the context of a natively derived vault particle could prove incredibly useful to understanding vault regulation. Even information about PARP4's substrates outside the context of the vault particle would provide the field with important insights into this enigmatic PARP protein.

3.5 Methods

3.5.1 Expression and Purification of WT and asPARP4 Catalytic Fragments

PARP4's catalytic fragment (encompassing its HD and ART domains) was subcloned into the pET47b(+) vector. The L453A mutation was introduced via round-the-horn PCR using primers 5'-GTGTCGAGGGGCGCTTTTACCCAAAGTAGTGGAAGATCGTG-3' and 5'-GGTAAAAGCGCCCCTCGACACAAGATTCCCACGATG-3' and confirmed by Sanger sequencing. pET47b-HD-ART and -asHD-ART were transformed into BL21(DE3) *E. coli* competent cells. 50 ml of Luria Broth (LB) growth media were inoculated with a single colony and grown for 6-8 hours at 37°C, until the OD600 reached 0.6-0.75. 8 ml of each starter culture were added to 990 ml of autoinduction media (autoclaved Luria Broth with 1X NPS (25 mM

$(NH_4)_2SO_4$ 50 mM KH_2PO_4, 50 mM Na_2HPO_4), 1X 5052 (1 g glycerol, 0.1 g glucose, and 0.4 g lactose), and 1 mM $MgSO_4$) and grown at 25°C and rotated at 220 rpm overnight.

Cells were pelleted by centrifugation at 6,000 rpm for 20 minutes and each pellet was resuspended in 80 ml Ni-NTA buffer (20 mM Na_2HPO_4/NaH_2PO_4 pH 7.4, 300 mM NaCl, 1 mM PMSF) with 10 mM of imidazole. Resuspended cell solutions were subjected to sonication for 3 minutes at the 75% amplitude setting, with a pulse sequence of 3s on/5s off. Lysate was cleared by centrifugation at 16,500 rpm for 30 minutes, and the supernatant was flowed over 6 milliliters of washed Ni-NTA resin (ThermoFisher) in a gravity column. Resin beads were washed with 150-200 ml 20 mM Imidazole in Ni-NTA buffer. Proteins were eluted from the column following a 5-minute incubation in 15-25 ml 250 mM imidazole in Ni-NTA buffer. Protein eluate was dialyzed into two liters of anion exchange Buffer A (25 mM Tris pH 8, 30 mM NaCl, and 0.5 mM TCEP-HCl) and subjected to His-tag cleavage with PreScission Protease overnight. Protein was concentrated to 10 ml, injected over a 5 ml MonoQ PE 1.6/100 column (Cytiva), and eluted via a salt gradient, with increasing concentrations of ion exchange Buffer B (25 mM Tris pH 8, 500 mM NaCl, and 1 mM DTT). Fractions containing the catalytic fragment were pooled and concentrated to 500 µl, then injected over a Superdex 200 increase 10/300 GL column (Cytiva) and eluted with SEC buffer (25 mM Tris pH 7.4, 150 mM NaCl, 0.2 mM TCEP-HCl, 2 mM MgSO4). Peak fractions were combined, concentrated, and stored at -80°C.

3.5.2 Western Blot WT and asPARP4 Activity Assay

The WT and asHD-ART domains were dialyzed into ADPrylation buffer (50 mM HEPES pH 8, 5 mM $MgCl_2$, 5 mM $CaCl_2$, 1 mM DTT) overnight. PARP enzymes, an activity assay "Master Mix," and either standard or analog NAD^+ were added to a reaction mixture, so that the

solution contained final concentrations of 100 ng/ul BSA, 0.01% NP-40, and 100 ng/ul sheared salmon sperm DNA (sssDNA), and of 500 μM β-NAD^+. The proteins were rocked at room temperature for 1 hour, at which point 2X Laemmli loading buffer was added to each tube and and boiled at 100 °C for 5 minutes. Two 12% acrylamide gels were run at 200 V for 45 minutes, after which one was incubated in Coomassie stain for 20 minutes, and the other was used for Western blot transfer. Proteins were transferred onto polyvinylidene difluoride (PVDF) membrane at 100 V for one hour. The membrane was blocked using a 5% milk solution in 1X TBST buffer (with 0.1% tween), then incubated in 5 ml of a 1:1000 dilution of α-pan-ADP-ribose (MABE1016) in 5% milk overnight. The next day, the blot was rinsed with TBST and incubated in a 1:4000 dilution of goat α-rabbit IgG HRP antibody (Jackson 111-035-045) at room temperature for one hour prior to visualization with the SuperSignal™ West Pico PLUS Chemiluminescent Substrate (ThermoFisher). The blot was exposed using the Chemi Blot setting on the ChemiDoc Imaging System (Bio-Rad).

3.5.3 Azide-Alkyne Cycloaddition and In-gel Visualization of Conjugated Fluorophore

An ADPrylation assay was performed exactly as described in 3.4.2 up until the samples were prepared for analysis by SDS-PAGE gel, at which point two thirds of the ADPrylated (or control) sample was diluted with 2X Laemmli loading buffer and boiled as described above. The remaining third of sample was added to 4 reaction volumes of methanol, 1 volume chloroform, and 3 volumes water. The sample was centrifuged at 1,000 *g* for 10 minutes, after which the protein layer began to precipitate and settle beneath the water/methanol layer, which was decanted. The sample was washed with an additional 4 reaction volumes of methanol, which were again decanted, and the protein pellet was air-dried in the Eppendorf tube. Once the pellet

was gummy, 5 reaction volumes of urea solubilization buffer (8 M urea, 4% CHAPS detergent, 100 mM HEPES, and 1 M NaCl, 0.2 mM PSMF) were distributed over the pellet and briefly incubated on ice. Samples were diluted with 4.5 reaction volumes of water in preparation for the cycloaddition reaction.

$CuSO_4$ and THPTA solutions were combined at a 1:5 molar ratio just prior to their addition to the protein mixture. 1 µl 10 mM N_3-TAMRA fluorophore, 2 µl 50:250 mM $CuSO_4$:THPTA, 1 µl 500 mM aminoguanidine-HCl, and 1 µl 500 mM NaAscorbate were added stepwise to the protein solution, which was then protected with foil and rocked for 2 hours at room temperature. The click reaction sample underwent an additional methanol: chloroform precipitation prior to being air-dried (under foil) once more. 1X Laemmli loading dye was added to each of the samples, which were heated at 65 °C for 15 minutes, then boiled at 100 °C for five. Gel samples were cleared of precipitate with a 5-minute, 13,000 rpm spin, then loaded into a 12% acrylamide gel, which was run at 20 mA for 65 minutes. The gel was rinsed with 30 ml of 10% methanol for ten minutes, twice, then 30 ml of water for 10 minutes, twice. Samples and gels were left under foil throughout. The gel was imaged directly on the ChemiDoc Imaging System (Bio-Rad) using a rhodamine green light setting with an excitation wavelength of 532 nm, and emission wavelength of 605 nm.

3.6 Supplementary Figure

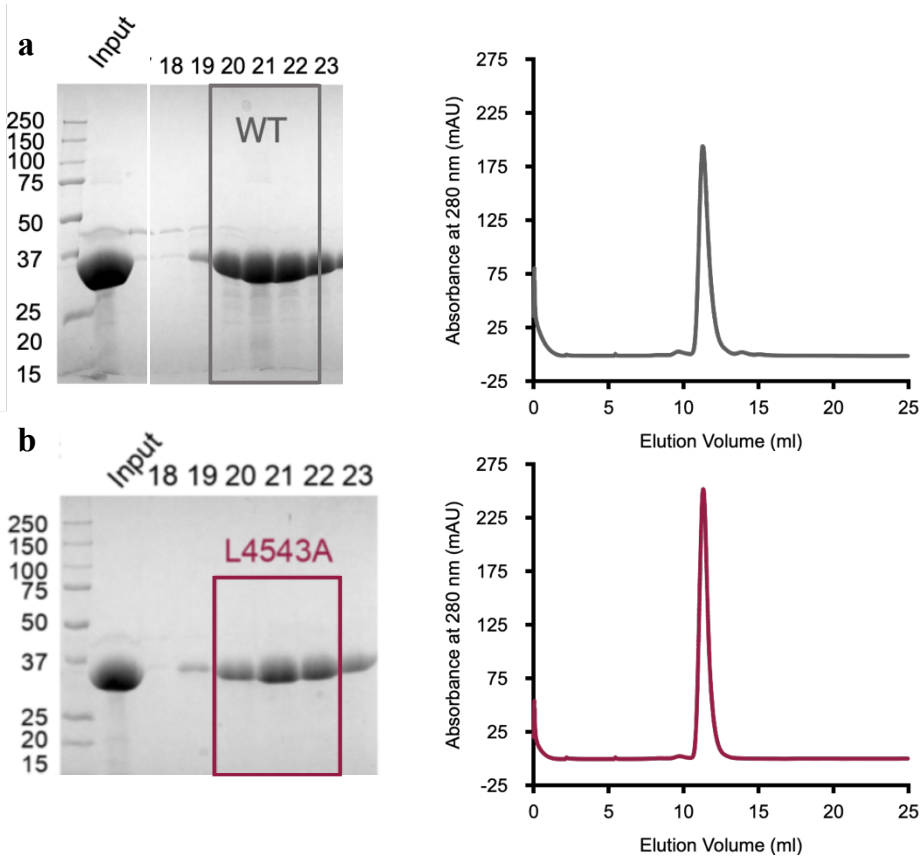


Figure 3S WT and asPARP4 HD-ART SEC Purification

a, SDS-PAGE gel (left) loaded with fractions of the WT HD-ART domain eluted from a Superdex 75 increase column after being purified as described in 3.5.1. SEC trace showing elution of protein at around 11 ml (right). **b,** SDS-PAGE gel and SEC trace show similar behavior of L453A mutant.

CHAPTER 4

CONCLUDING REMARKS

My thesis work sought to illuminate vault's function, mechanism, and regulation by other molecules. In Chapter 2, I used structural biology, biochemistry, and MD simulations to characterize the association between the MVP cage and its resident enzyme, PARP4, as well as its novel interaction with PARP4's enzymatic substrate, NAD^+. We showed that MVP-PARP4 complex formation was driven by broad shape complementarity and regions of attractive electrostatic potential between MINT and MVP's repeat "body" domains. We also highlighted steric features and electrostatic interactions forged by specific residue side chains in our structure. Mutagenesis and binding assays using SEC confirmed the significance of those residues to MVP-PARP4 complex formation. This work resolves a longstanding question in the field concerning the precise nature of the MVP-PARP4 interaction, which was (possibly due to the unusual structures of their respective binding domains) previously modeled incorrectly by prediction software [106]. Our work will allow for precisely targeted assays probing the function of PARP4-containing vaults in the future. Although the majority of the PARP4 structure could not be resolved from our complex maps, I describe our initial efforts to determine the structure of its ordered domains in Appendix B (Figure B.2.2). Refining these efforts will hopefully shed light on how PARP4's somewhat unusual activity is regulated, and how it may target modification substrates within the vault.

Our structures also reveal a novel NAD^+-binding fold behind the "keyhole loop" of MVP's shoulder region. We use the structural model derived from our cryo-EM map, as well as MD simulations, to identify critical NAD^+-interacting residues, showing that the metabolite is primarily stabilized via basic residues that line a small binding pocket and neutralize the charge

of its diphosphate group. Additional stability is achieved through hydrogen bonds between NAD^+'s ribose sugars and polar MVP residues, as well as some stacking interactions with NAD^+'s nicotinamide group. Isothermal titration calorimetry data showing a dose-dependent exothermic response to the titration of NAD^+ into MVP also provide evidence for the interaction. Because little precedent exists for NAD^+ interactions with protein domains that do not belong to a limited set of enzymes, we still have many questions about the function of this interaction. For one, we do not know if our *in vitro* observations correspond to a specific cellular function for NAD^+ binding or if they are instead suggestive of a more functionally relevant interaction between MVP and a different nucleotide or class of nucleotides. However, limited examples of proteins within the NAD^+ interactome suggest that NAD^+ binding partners may interact with the molecule to stabilize a particular conformational state or transfer the metabolite to a particular compartment or complex. We aim to address this question by assaying MVP's ability to interact with similar nucleotide-based structures. From there, we can begin to conceive *in vivo* experiments probing the role of this MVP-nucleotide interaction on a stronger footing.

Chapter 2 also features some initial mass spectrometry (MS) data regarding PARP4's function within the vault particle. MS PTM detection experiments following *in vitro* ADPrylation assays indicate that a specific serine near MVP's medial interface is apparently ADPrylated in the presence of PARP4. Such a modification in the cell could potentially regulate which macromolecules interact with, or even enter, the MVP cage depending on the cell environment. Previous work on ADPr modifications suggests that they function not only by recruiting readers to various sites, but also by displacing other (especially) phosphate PTMs from a set of common target residues. A paradigm in which vault is able to differentially recruit cargo corresponding to cell state, with the help of a dynamic PTM code could explain (at least in part)

its diverse cellular functions. In an effort to begin to understand PARP4's role in such a paradigm (or, alternatively, its role in recruiting cargo through direct binding), we used a proteomics approach to assess differences in vault cargo recruitment in a WT versus PARP4 KD cell line. While there was substantial overlap between their enriched interaction partners ("hits"), immunoprecipitated vaults pulled down several hits that were unique to a specific cell line, as well as hits common to both conditions that showed cell-line dependent enrichment. Upcoming work will attempt to resolve the vault-associated functions of these hits and PARP4's role in regulating them.

In Chapter 3, I discussed some of my other initial efforts to expand our understanding of vault-bound PARP4's role in the cell. We have begun to develop an analog-sensitive PARP4 platform to "click" PARP4 modification targets to azide-conjugated molecules that can be readily pulled out of lysate. In this way, we hope to identify PARP4's targets in both the presence and absence of vault to map its potential regulatory role in signaling pathways mediated by vault. Taken together, the work has offered the field new approaches to probe vault function and regulation. It has generated new lines of inquiry that will hopefully bring us closer to understanding the bases of vault's multifaceted cellular function.

APPENDIX A

APPENDIX TO CHAPTER 2

In this section, I've appended quality control experiments done with the Fabs we used for the coimmunoprecipitation assays described in Chapter 2 as well as results of an experiment done to directly compare the stability of the vault cage when the MVP-PARP4 complex is incubated with and without NAD^+ over a 30 min period.

A.1 Fab and Cell Line Quality Control Tests

A.1.1 Avi-Fab Biotinylation Test

We conducted an initial test to ensure that the avi-tagged anti-hMVP Fabs we constructed were efficiently biotinylated following purification from a BirA-expressing cell line supplemented with biotin. Biotinylated Fabs were incubated with streptavidin beads for 30 min at room temperature and washed 3X. Washed beads were combined with 2X Laemmli loading buffer and loaded onto a stain-free gel alongside the input, flow through, and final wash samples from the pulldown. An avi-tagged anti-MBP Fab was used as a positive control. The appearance of bands representing our Fabs in the streptavidin bead fraction indicated that the experiment was successful and that we could use the Fabs for co-IP with magnetic streptavidin beads.

A.1.2 Vault IP from MVP- and PARP4-depleted Cell Lines

After verifying, by Sanger sequencing, that we had successfully mutated the target loci in our engineered MVP- and PARP4-depleted cell lines, we verified that we had downregulated their expression at the protein level (and importantly in the case of PARP4, its inclusion into the

vault). We conducted a Western blot assay following the MVP IP, with the Fab used in our co-IP experiments described above (Figure 2.4.8). We blotted for PARP4, then MVP (due to MVP's relative abundance over PARP4 in a native purification) to verify that our cell lines were suitable for assays investigating the loss of vault component proteins in the cell.

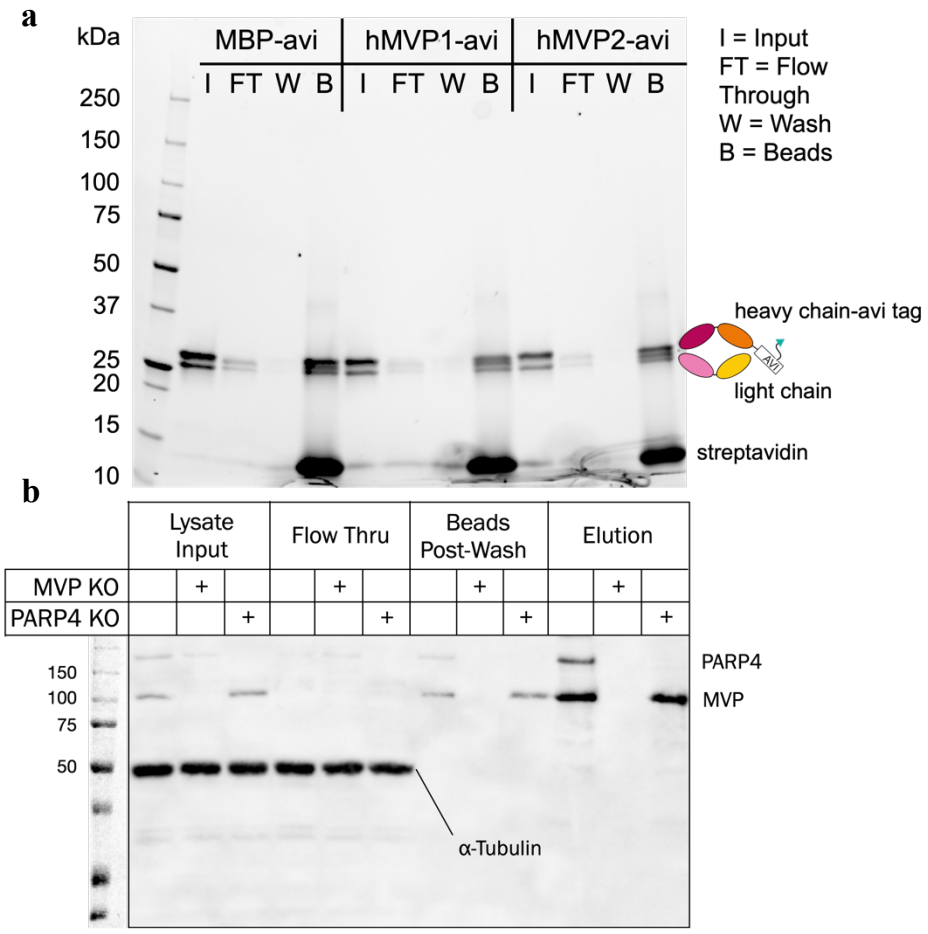


Figure A.1 Fab tests for coimmunoprecipitation assays

a, SDS-PAGE gel showing results of biotinylated Fab test IP with streptavidin beads. Bands running near 25 kD represent the Fab chains, and their presence in the "bead" lanes with streptavidin (band near 14 kD) indicates that they were successfully pulled down and suitable for co-IP experiments. **b,** Test IP of vaults from engineered cell lines to deplete MVP and PARP4. Vaults were pulled down from A549 cell lysate with the VM1 Fab and subjected to Western blot with anti-MVP and anti-PARP4 antibodies. Significantly more MVP and PARP4 were detected in the WT than respective engineered cell lines. Alpha-tubulin is shown in input and flow thru lanes as a loading control.

A.2 Vault Cage Stability Test Comparing MVP-PARP4 Complexes +/- NAD^+

In an early hypothesis about PARP4's role in vault regulation, we speculated that it might deposit ADPr modifications onto polar or charged residues near the medial interface to disrupt the MVP intermolecular interactions, thereby initiating cage opening. We thought this might be a mechanism to regulate cargo internalization. Once we had successfully purified the MVP-PARP4 complex, we observed the effects of adding PARP4's NAD^+ substrate to the complex mixture for 30 min prior to freezing on a grid with MVP-PARP4 particles that had been incubated in buffer. We hoped that, within this relatively short timeframe, we would be able to observe a number of conformational states (or degrees of aggregation or particle collapse) if this was, in fact, a legitimate phenomenon. While we did not subject the data to any quantitative analysis, the apparently equal proportion of intact vaults in each micrograph (Figure A.2), contested our hypothesis that they would disrupt vault cage stability to an extent that we did not pursue additional attempts to test it. These results did, however, did leave the door open to pursue high-resolution vault cage structures containing PARP4 and NAD^+ later on.

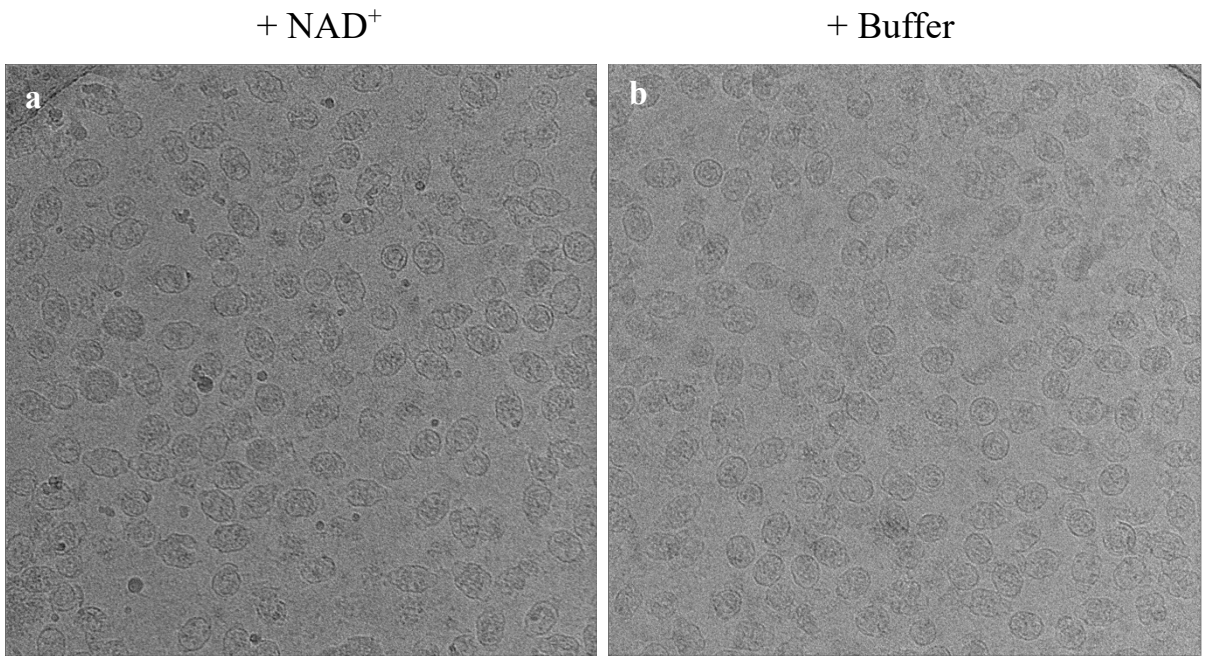


Figure A.2 NAD+ treatment did not destabilize the vault cages containing PARP4

a, MVP-PARP4 complexes were incubated with 500 μM NAD^+ for 30 min in ADPrylation buffer prior to being applied to grids and plunge-frozen into liquid ethane for subsequent analysis using the Talos 200 kV cryo-electron microscope. **b,** An identical sample of MVP-PARP4 molecules were supplemented with an equivalent volume of buffer, frozen onto cryo-EM grids, and compared to the + NAD^+ sample. Both images were taken at 73 kx magnification.

APPENDIX B

ATTEMPTS TO FURTHER CHARACTERIZE PARP4 STRUCTURE

Here, I show my efforts to purify and structurally characterize the N-terminal region of hPARP4 that extends into the vault lumen and which we could not resolve in our full-length cryo-EM structure. I also show data referenced in Chapter 3 regarding attempts to purify full length His-hPARP4 following expression in Hi5 insect cells.

B.1 Purification Attempt of Full-length (FL) His-PARP4 Expressed in Insect Cells

Because we recognized that PARP4's flexibility within the comparatively rigid MVP cage would make structural resolution difficult, we wanted to employ some additional strategies to resolve its structure and potentially fit the resultant model into a cryo-EM map. Because PARP4 is a very large protein, we decided it would be worthwhile to try purifying the FL construct for cryo-EM analysis. A group had previously purified the FL construct by overexpressing the construct in mammalian cells with a GFP tag [97], so we decided to attempt it with our insect cell system. Szymon Kordon, the rotation student at the time, put considerable work into the effort and made several optimization attempts. After those proved unsuccessful as well, I decided to pursue a more hospitable construct.

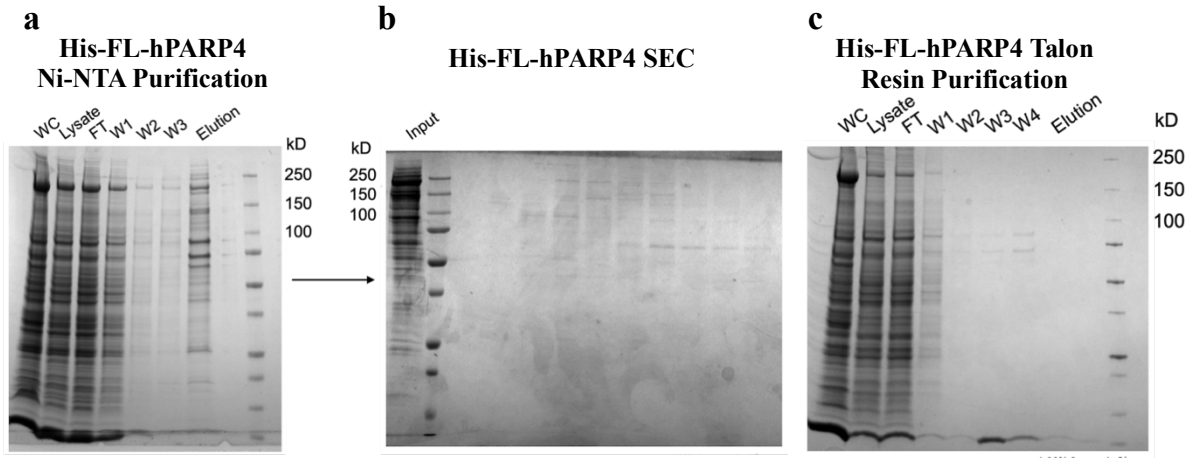


Figure B.1 His-FL-hPARP4 purification attempts

a, SDS-PAGE gel showing His-FL-hPARP4 Ni-NTA affinity purification samples. His-FL-hPARP4 was overexpressed in insect cells (WC), but it was lost over the course of the prep until contaminants could be seen in about equal quantities to PARP4 in the elution. **b,** SDS-PAGE gel showing His-FL-hPARP4 fractions from a Superdex 200 increase gel filtration column. **c,** His-FL-hPARP4 Talon affinity purification attempt. Using Talon resin resulted in an even lower yield.

Data from Szymon Kordon's Zhao Lab rotation presentation

B.2 Purification and Initial Structural Analysis of the hPARP4 1-1243 Construct

After my initial failure to purify (FL) hPARP4, I designed a construct that was predicted (pre-AlphaFold) to contain regions of ordered secondary structure (occasionally broken up by loops between known domains) from its N-terminus to its 1243rd residue. Between residue 1243 and the MINT domain, PARP4 contains a proline-serine rich, predicted-to-be disordered region that (even if FL PARP4 purification from our insect cell system was possible) would not have been amenable to structural determination. Having, with considerable effort from Szymon Kordon – a rotation student at the time – previously attempted to purify the FL construct and noticing that much of it had been lost in the flow through during the affinity purification, I constructed a GST-tagged version out of concern that an N-terminal His-tag might be occluded

during the affinity prep. After optimizing an on-column cleavage protocol to more effectively purify cut protein (which was cleaved very inefficiently), I tested the construct's activity, which was detectable by anti-pan-ADP-ribose Western blot, but markedly lower than the HD-ART domain alone. An interesting future direction may be to further investigate whether this is due, perhaps, to its diminished stability compared to the more compact catalytic domain, or if any of its other domains exercise an autoinhibitory effect. I also used this assay as an opportunity to compare the construct's modification efficiency in the NaCl-free ADPrylation buffer recommended in [45] and in the same buffer with normal (150 mM) salt and did note slightly lower efficiency in the NaCl-containing buffer (Figure B.2.1c).

I also initiated some grid optimization and eventually collected a mid-size dataset (in 150 mM NaCl SEC buffer, rather than the salt-free reaction buffer) (Figure B.2.2). Although the particles exhibited good contrast (and decent 2D classes), a combination of excessive particle distribution on most micrographs (generating a lot of noise around individual particles), large variation in ice thickness within each hole, and preferred orientation of the particle, meant that the resolution only reached about 8 Å. In addition, although the grids were frozen immediately after protein purification, the presence of a persistent degradation product associated with this construct (Figure B.2.1b) may also have thwarted structural analysis by muddling the homogeneity of the various classes. Since my initial collection, I have attempted some shorter constructs to address the degradation issue (though none have expressed as well) and done some screening with the original construct in DDM (a detergent that tends to work well for structural determination of soluble proteins) but have not optimized the grids to a point that they have merited additional data collection. A structure of PARP4's ordered region would potentially answer a number of interesting mechanistic questions about the black sheep of the PARP family.

It could shed light on the aforementioned possibility that PARP4 is regulated by an autoinhibitory mechanism, resolve the properties of its active site that (as far as we know) may prevent it from depositing PAR chains on its own despite containing the necessary catalytic triad, and even give us some insight into how it might interact with potential binding partners within the vault cage.

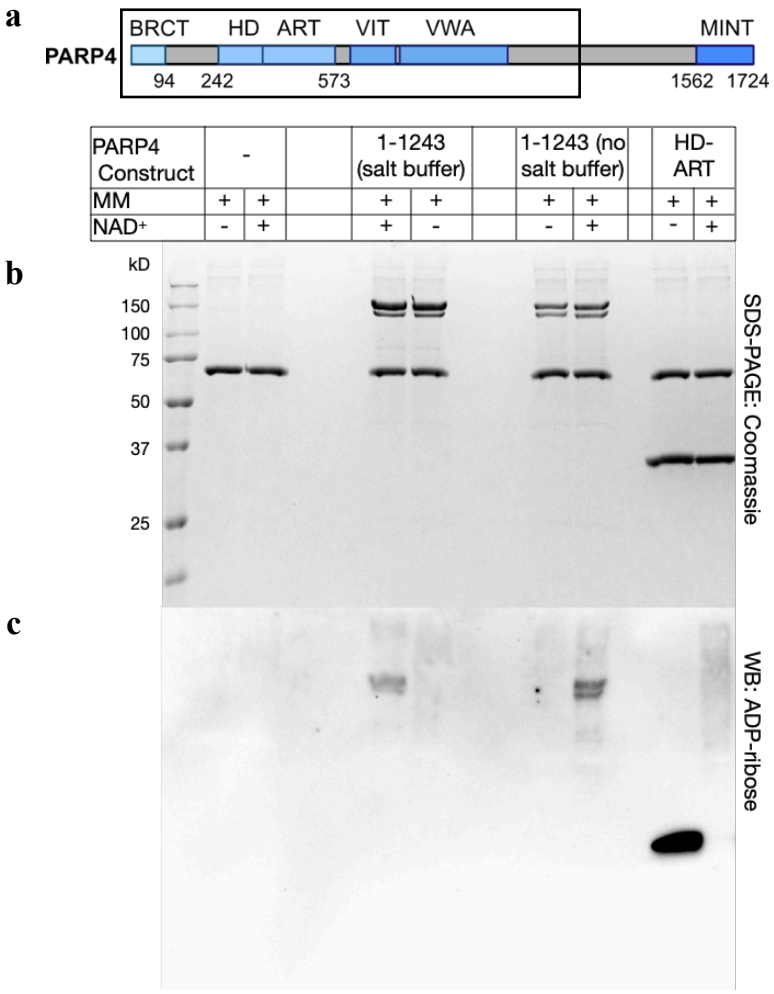


Figure B.2.1 Catalytic activity test for hPARP4 1-1243 construct

a, PARP4 domain diagram, with box around residues 1-1243. **b,** SDS-PAGE gel showing inputs for Western Blot based activity assay. Negative control (BSA master mix with added NAD^+ and no PARP protein) shown on the left and positive control (PARP4 catalytic domain) on the right. hPARP4 1-1243 construct incubated with ADP-ribosylation buffers containing 150 mM NaCl and salt-free ADP-ribosylation buffer shown in the center left and center right lanes, respectively. **c,** Western blot showing detected automodification (black bands) of each PARP4 construct using anti-pan-ADP-ribose reagent.

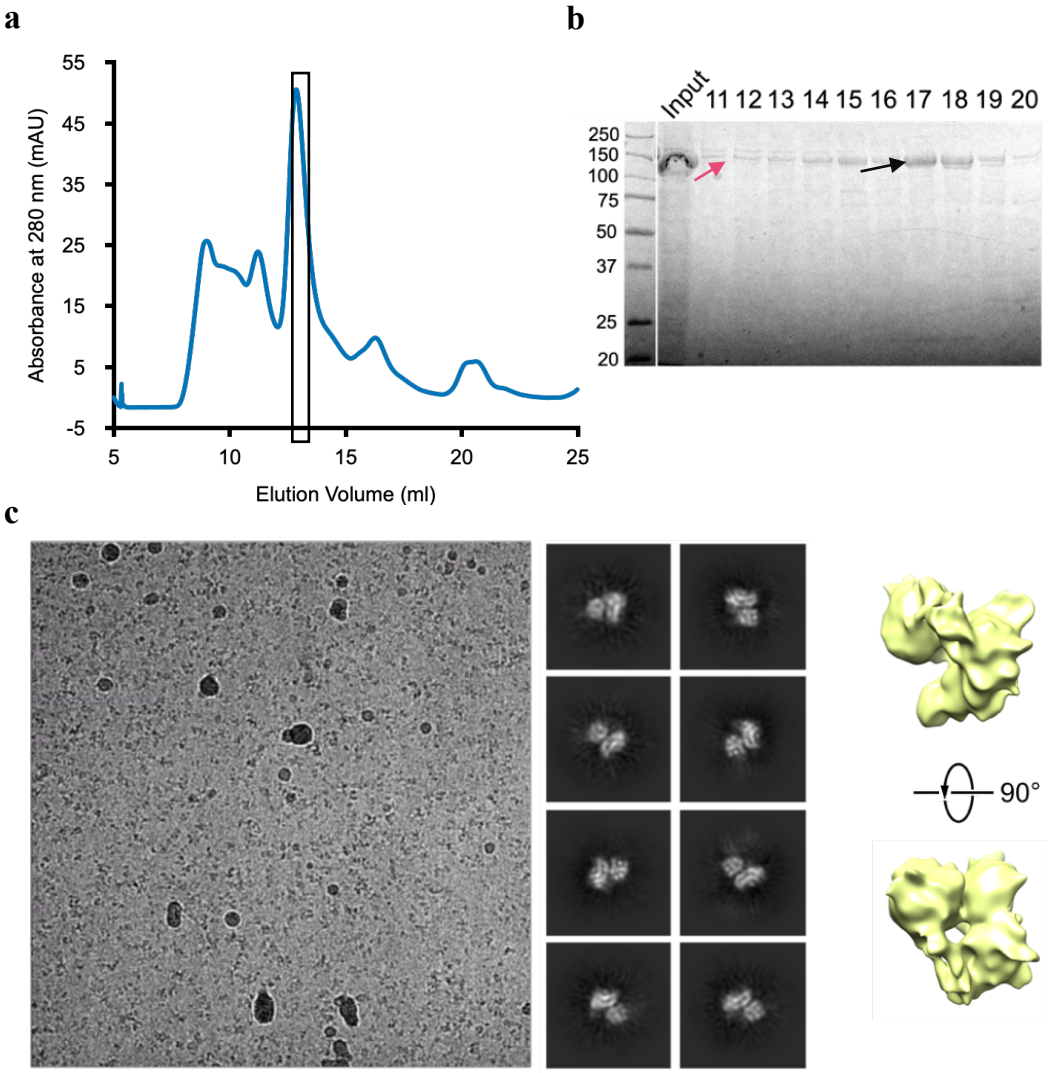


Figure B.2.2 Purification and structural studies of PARP4's ordered N-terminal domains

a, SEC chromatogram of hPARP4 1-1243 elution from a Superdex 200 increase column. Absorbance values at 280 nm are shown in blue, with the portion of the peak contained in fraction 17 boxed in black. **b**, Coomassie stained 12% SDS-PAGE gel of SEC fraction samples. The pink arrow highlights residual uncleaved GST-hPARP4 1-1243, while the black arrow shows the fully cleaved construct. Sample in fraction 17 was used for cryo-EM grid preparation. **c,** Sample micrograph, 2D classes, and low-resolution 3D reconstruction of the hPARP4 1-1243 construct from an initial dataset. Data were processed in the RELION software suite, and 3D volumes were visualized in ChimeraX.

REFERENCES

1. Uhlén, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

2. Sanderson, D.J. and M.S. Cohen, *Mechanisms governing PARP expression, localization, and activity in cells.* Critical Reviews in Biochemistry and Molecular Biology, 2020. **55**(6): p. 541-554.

3. Xie, N., et al., *NAD+ metabolism: pathophysiologic mechanisms and therapeutic potential.* Signal Transduction and Targeted Therapy, 2020. **5**(1): p. 227.

4. Daly, T.K., A.J. Sutherland-Smith, and D. Penny, *In Silico Resurrection of the Major Vault Protein Suggests It Is Ancestral in Modern Eukaryotes.* Genome Biology and Evolution, 2013. **5**(8): p. 1567-1583.

5. Woodward, C.L., L.M. Mendonça, and G.J. Jensen, *Direct visualization of vaults within intact cells by electron cryo-tomography.* Cellular and Molecular Life Sciences, 2015. **72**(17): p. 3401-3409.

6. Cohen, M.S. and P. Chang, *Insights into the biogenesis, function, and regulation of ADP-ribosylation.* Nature Chemical Biology, 2018. **14**(3): p. 236-243.

7. Kedersha, N.L. and L.H. Rome, *Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA.* Journal of Cell Biology, 1986. **103**(3): p. 699-709.

8. Kedersha, N.L., et al., *Vaults. III. Vault ribonucleoprotein particles open into flower-like structures with octagonal symmetry.* J Cell Biol, 1991. **112**(2): p. 225-35.

9. Tanaka, H., et al., *The Structure of Rat Liver Vault at 3.5 Angstrom Resolution.* Science, 2009. **323**(5912): p. 384-388.

10. Llauró Portell, A., et al., *Mechanical Stability and Reversible Fracture of Vault Particles*. Vol. 106. 2014. 687-95.

11. Liu, J., et al., *The Stomatin/Prohibitin/Flotillin/HflK/C Domain of Flotillin-1 Contains Distinct Sequences That Direct Plasma Membrane Localization and Protein Interactions in 3T3-L1 Adipocytes.* Journal of Biological Chemistry, 2005. **280**(16): p. 16125-16134.

12. Browman, D.T., M.B. Hoegg, and S.M. Robbins, *The SPFH domain-containing proteins: more than lipid raft markers.* Trends in Cell Biology, 2007. **17**(8): p. 394-402.

13. Ma, C., et al., *Structural insights into the membrane microdomain organization by SPFH family proteins.* Cell Research, 2022. **32**(2): p. 176-189.

14. Llauró Portell, A., et al., *Decrease in pH destabilizes individual vault nanocages by weakening the inter-protein lateral interaction*. Vol. 6. 2016. 34143.

15. Guerra, P., et al., *Symmetry disruption commits vault particles to disassembly.* Science Advances, 2022. **8**(6): p. eabj7795.

16. Yang, J., et al., *Vaults Are Dynamically Unconstrained Cytoplasmic Nanoparticles Capable of Half Vault Exchange.* ACS Nano, 2010. **4**(12): p. 7229-7240.

17. Mossink, M.H., et al., *Disruption of the Murine Major Vault Protein (MVP/LRP) Gene Does Not Induce Hypersensitivity to Cytostatics.* Cancer Research, 2002. **62**(24): p. 7298-7304.

18. Mossink Marieke, H., et al., *Unimpaired dendritic cell functions in MVP/LRP knockout mice.* Immunology, 2003. **110**(1): p. 58-65.

19. van Zon, A., et al., *Efflux Kinetics and Intracellular Distribution of Daunorubicin Are Not Affected by Major Vault Protein/Lung Resistance-Related Protein (Vault) Expression.* Cancer Research, 2004. **64**(14): p. 4887-4892.

20. Huffman, K.E. and D.R. Corey, *Major vault protein does not play a role in chemoresistance or drug localization in a non-small cell lung cancer cell line.* Biochemistry, 2005. **44**(7): p. 2253-61.

21. Chugani, D., N. Kedersha, and L. Rome, *Vault immunofluorescence in the brain: new insights regarding the origin of microglia.* Journal of Neuroscience, 1991. **11**(1): p. 256-268.

22. Suprenant, K.A., et al., *The major vault protein is related to the toxic anion resistance protein(TelA) family.* Journal of Experimental Biology, 2007. **210**(6): p. 946-955.

23. Sokolskyi, T., *Bacterial Major Vault Protein homologs shed new light on origins of the enigmatic organelle.* bioRxiv, 2019: p. 872010.

24. Kowalski, M.P., et al., *Host Resistance to Lung Infection Mediated by Major Vault Protein in Epithelial Cells.* Science, 2007. **317**(5834): p. 130-132.

25. Peng, N., et al., *Inducible Major Vault Protein Plays a Pivotal Role in Double-Stranded RNA– or Virus-Induced Proinflammatory Response.* The Journal of Immunology, 2016. **196**(6): p. 2753-2766.

26. Rayo, J., et al., *Immunoediting role for major vault protein in apoptotic signaling induced by bacterial N-acyl homoserine lactones.* Proceedings of the National Academy of Sciences, 2021. **118**(12): p. e2012529118.

27. Liu, S., et al., *Human hepatitis B virus surface and e antigens inhibit major vault protein signaling in interferon induction pathways.* Journal of Hepatology, 2015. **62**(5): p. 1015-1023.

28. Ben, J., et al., *Major Vault Protein Regulates Class A Scavenger Receptor-mediated Tumor Necrosis Factor-α Synthesis and Apoptosis in Macrophages.* Journal of Biological Chemistry, 2013. **288**(27): p. 20076-20084.

29. Ben, J., et al., *Major vault protein suppresses obesity and atherosclerosis through inhibiting IKK–NF-κB signaling mediated inflammation.* Nature Communications, 2019. **10**(1): p. 1801.

30. Lüscher, B., et al., *ADP-ribosyltransferases, an update on function and nomenclature.* The FEBS Journal, 2022. **289**(23): p. 7399-7410.

31. Zhang, Y., et al., *Site-specific characterization of the Asp- and Glu-ADP-ribosylated proteome.* Nature Methods, 2013. **10**(10): p. 981-984.

32. Buch-Larsen, S.C., et al., *Mapping Physiological ADP-Ribosylation Using Activated Ion Electron Transfer Dissociation.* Cell Reports, 2020. **32**(12): p. 108176.

33. van der Heden van Noort, G.J., *Chemical Tools to Study Protein ADP-Ribosylation.* ACS Omega, 2020. **5**(4): p. 1743-1751.

34. Larsen, S.C., et al., *Systems-wide Analysis of Serine ADP-Ribosylation Reveals Widespread Occurrence and Site-Specific Overlap with Phosphorylation.* Cell Reports, 2018. **24**(9): p. 2493-2505.e4.

35. Munnur, D. and I. Ahel, *Reversible mono-ADP-ribosylation of DNA breaks.* The FEBS Journal, 2017. **284**(23): p. 4002-4016.

36. Munnur, D., et al., *Reversible ADP-ribosylation of RNA.* Nucleic Acids Research, 2019. **47**(11): p. 5658-5669.

37. Weixler, L., K.L.H. Feijs, and R. Zaja, *ADP-ribosylation of RNA in mammalian cells is mediated by TRPT1 and multiple PARPs.* Nucleic Acids Research, 2022. **50**(16): p. 9426-9441.

38. Verheugd, P., et al., *Regulation of NF-κB signalling by the mono-ADP-ribosyltransferase ARTD10.* Nature Communications, 2013. **4**(1): p. 1683.

39. Lehmann, L.C., et al., *Mechanistic Insights into Autoinhibition of the Oncogenic Chromatin Remodeler ALC1.* Mol Cell, 2017. **68**(5): p. 847-859.e7.

40. Li, L., et al., *PARP12 suppresses Zika virus infection through PARP-dependent degradation of NS1 and NS3 viral proteins.* Science Signaling, 2018. **11**(535): p. eaas9332.

41. Boehi, F., P. Manetsch, and M.O. Hottiger, *Interplay between ADP-ribosyltransferases and essential cell signaling pathways controls cellular responses.* Cell Discovery, 2021. **7**(1): p. 104.

42. Catara, G., et al., *PARP1-produced poly-ADP-ribose causes the PARP12 translocation to stress granules and impairment of Golgi complex functions.* Scientific Reports, 2017. **7**(1): p. 14035.

43. D'Amours, D., et al., *Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions.* Biochem J, 1999. **342 (Pt 2)**(Pt 2): p. 249-68.

44. Zhao, Y., et al., *Crosstalk between ubiquitin and other post-translational modifications on chromatin during double-strand break repair.* Trends in Cell Biology, 2014. **24**(7): p. 426-434.

45. Gibson, B.A., et al., *Chemical genetic discovery of PARP targets reveals a role for PARP-1 in transcription elongation.* Science, 2016. **353**(6294): p. 45-50.

46. Challa, S., et al., *Ribosome ADP-ribosylation inhibits translation and maintains proteostasis in cancers.* Cell, 2021. **184**(17): p. 4531-4546.e26.

47. Vyas, S., et al., *A systematic analysis of the PARP protein family identifies new functions critical for cell physiology.* Nature Communications, 2013. **4**(1): p. 2240.

48. Jin, X., et al., *Functional Roles of Poly(ADP-Ribose) in Stress Granule Formation and Dynamics.* Frontiers in Cell and Developmental Biology, 2021. **9**.

49. David, K.K., et al., *Parthanatos, a messenger of death.* Front Biosci (Landmark Ed), 2009. **14**(3): p. 1116-28.

50. Rose, M., et al., *PARP Inhibitors: Clinical Relevance, Mechanisms of Action and Tumor Resistance.* Frontiers in Cell and Developmental Biology, 2020. **8**.

51. Palavalli Parsons, L.H., et al., *Identification of PARP-7 substrates reveals a role for MARylation in microtubule control in ovarian cancer cells.* eLife, 2021. **10**: p. e60481.

52. Lüscher, B., et al., *Intracellular mono-ADP-ribosyltransferases at the host–virus interphase.* Cellular and Molecular Life Sciences, 2022. **79**(6): p. 288.

53. Alhammad, Y.M.O., et al., *The SARS-CoV-2 Conserved Macrodomain Is a Mono-ADP-Ribosylhydrolase.* Journal of Virology, 2021. **95**(3): p. e01969-20.

54. Voth, L.S., et al., *Unique Mutations in the Murine Hepatitis Virus Macrodomain Differentially Attenuate Virus Replication, Indicating Multiple Roles for the Macrodomain in Coronavirus Replication.* Journal of Virology, 2021. **95**(15): p. e00766-21.

55. McPherson, R.L., et al., *ADP-ribosylhydrolase activity of Chikungunya virus macrodomain is critical for virus replication and virulence.* Proceedings of the National Academy of Sciences, 2017. **114**(7): p. 1666-1671.

56. Heer, C.D., et al., *Coronavirus infection and PARP expression dysregulate the NAD metabolome: An actionable component of innate immunity.* Journal of Biological Chemistry, 2020. **295**(52): p. 17986-17996.

57. Hopp, A.-K. and M.O. Hottiger, *Uncovering the Invisible: Mono-ADP-ribosylation Moved into the Spotlight.* Cells, 2021. **10**(3): p. 680.

58. Ryu, K.W., et al., *Metabolic regulation of transcription through compartmentalized NAD+ biosynthesis.* Science, 2018. **360**(6389): p. eaan5780.

59. Pan, L.-Z., et al., *The NAD+ synthesizing enzyme nicotinamide mononucleotide adenylyltransferase 2 (NMNAT-2) is a p53 downstream target.* Cell Cycle, 2014. **13**(6): p. 1041-1048.

60. Yang, H., et al., *Nutrient-Sensitive Mitochondrial NAD+ Levels Dictate Cell Survival.* Cell, 2007. **130**(6): p. 1095-1107.

61. Heske, C.M., *Beyond Energy Metabolism: Exploiting the Additional Roles of NAMPT for Cancer Therapy.* Front Oncol, 2019. **9**: p. 1514.

62. Meyer, T., et al., *NAD+metabolism drives astrocyte proinflammatory reprogramming in central nervous system autoimmunity.* Proceedings of the National Academy of Sciences, 2022. **119**(35): p. e2211310119.

63. Cameron, A.M., et al., *Inflammatory macrophage dependence on NAD+ salvage is a consequence of reactive oxygen species–mediated DNA damage.* Nature Immunology, 2019. **20**(4): p. 420-432.

64. Dasovich, M., et al., *Identifying Poly(ADP-ribose)-Binding Proteins with Photoaffinity-Based Proteomics.* Journal of the American Chemical Society, 2021. **143**(8): p. 3037-3042.

65. Cho, S.H., et al., *Glycolytic rate and lymphomagenesis depend on PARP14, an ADP ribosyltransferase of the B aggressive lymphoma (BAL) family.* Proceedings of the National Academy of Sciences, 2011. **108**(38): p. 15972-15977.

66. Iwata, H., et al., *PARP9 and PARP14 cross-regulate macrophage activation via STAT1 ADP-ribosylation.* Nature communications, 2016. **7**(1): p. 12849.

67. Kulikova, V.A. and A.A. Nikiforov, *Role of NUDIX Hydrolases in NAD and ADP-Ribose Metabolism in Mammals.* Biochemistry (Mosc), 2020. **85**(8): p. 883-894.

68. Nagy, E., et al., *Identification of the NAD+-Binding Fold of Glyceraldehyde-3-Phosphate Dehydrogenase as a Novel RNA-Binding Domain.* Biochemical and Biophysical Research Communications, 2000. **275**(2): p. 253-260.

69. Chen, S.-H. and X. Yu, *Human DNA ligase IV is able to use NAD+ as an alternative adenylation donor for DNA ends ligation.* Nucleic Acids Research, 2018. **47**(3): p. 1321-1334.

70. Ribeiro, J.M., et al., *Specific cyclic ADP-ribose phosphohydrolase obtained by mutagenic engineering of Mn2+-dependent ADP-ribose/CDP-alcohol diphosphatase.* Scientific Reports, 2018. **8**(1): p. 1036.

71. Wang, Y., et al., *Poly(ADP-Ribose) (PAR) Binding to Apoptosis-Inducing Factor Is Critical for PAR Polymerase-1-Dependent Cell Death (Parthanatos).* Science Signaling, 2011. **4**(167): p. ra20-ra20.

72. Artus, C., et al., *AIF promotes chromatinolysis and caspase-independent programmed necrosis by interacting with histone H2AX.* Embo j, 2010. **29**(9): p. 1585-99.

73. Reber, J.M. and A. Mangerich, *Why structure and chain length matter: on the biological significance underlying the structural heterogeneity of poly(ADP-ribose).* Nucleic Acids Research, 2021. **49**(15): p. 8432-8448.

74. Duarte-Pereira, S., et al., *Study of NAD-interacting proteins highlights the extent of NAD regulatory roles in the cell and its potential as a therapeutic target.* Journal of Integrative Bioinformatics, 2023(20220049).

75. Kumar, V., et al., *Transcription corepressor CtBP is an NAD(+)-regulated dehydrogenase.* Mol Cell, 2002. **10**(4): p. 857-69.

76. Rutter, J., et al., *Regulation of Clock and NPAS2 DNA Binding by the Redox State of NAD Cofactors.* Science, 2001. **293**(5529): p. 510-514.

77. Kong, L.B., et al., *Structure of the vault, a ubiquitous celular component.* Structure, 1999. **7**(4): p. 371-9.

78. Bateman, A. and V. Kickhoefer, *The TROVE module: A common element in Telomerase, Ro and Vault ribonucleoproteins.* BMC Bioinformatics, 2003. **4**(1): p. 49.

79. Mikyas, Y., et al., *Cryoelectron Microscopy Imaging of Recombinant and Tissue Derived Vaults: Localization of the MVP N Termini and VPARP.* Journal of Molecular Biology, 2004. **344**(1): p. 91-105.

80. Liu, Y., et al., *Telomerase-associated protein TEP1 is not essential for telomerase activity or telomere length maintenance in vivo.* Mol Cell Biol, 2000. **20**(21): p. 8178-84.

81. Kickhoefer, V.A., et al., *The Telomerase/Vault-Associated Protein Tep1 Is Required for Vault RNA Stability and Its Association with the Vault Particle.* The Journal of Cell Biology, 2001. **152**(1): p. 157-164.

82. Liu, Y., et al., *Vault Poly(ADP-Ribose) Polymerase Is Associated with Mammalian Telomerase and Is Dispensable for Telomerase Function and Vault Structure In Vivo.* Molecular and Cellular Biology, 2004. **24**(12): p. 5314-5323.

83. Teng, Y., et al., *MVP-mediated exosomal sorting of miR-193a promotes colon cancer progression.* Nat Commun, 2017. **8**: p. 14448.

84. Amort, M., et al., *Expression of the vault RNA protects cells from undergoing apoptosis.* 2015. **6**: p. 7030.

85. Horos, R., et al., *The Small Non-coding Vault RNA1-1 Acts as a Riboregulator of Autophagy.* Cell, 2019. **176**(5): p. 1054-1067.e12.

86. Zhao, Y., et al., *RIG-I like receptor sensing of host RNAs facilitates the cell-intrinsic immune response to KSHV infection.* Nature Communications, 2018. **9**(1): p. 4841.

87. Kickhoefer, V.A., et al., *The 193-Kd Vault Protein, Vparp, Is a Novel Poly(Adp-Ribose) Polymerase.* The Journal of Cell Biology, 1999. **146**(5): p. 917-928.

88. Loeffler, P.A., et al. *Structural studies of the PARP-1 BRCT domain*. BMC structural biology, 2011. **11**, 37 DOI: 10.1186/1472-6807-11-37.

89. Rudolph, J., et al., *The BRCT domain of PARP1 binds intact DNA and mediates intrastrand transfer.* Molecular Cell, 2021. **81**(24): p. 4994-5006.e5.

90. Li, M., et al., *The FHA and BRCT domains recognize ADP-ribosylation during DNA damage response.* Genes & Development, 2013. **27**(16): p. 1752-1768.

91. Langelier, M.-F., et al., *NAD+ analog reveals PARP-1 substrate-blocking mechanism and allosteric communication from catalytic center to DNA-binding domains.* Nature Communications, 2018. **9**(1): p. 844.

92. Citarelli, M., S. Teotia, and R.S. Lamb, *Evolutionary history of the poly(ADP-ribose) polymerase gene family in eukaryotes.* BMC Evolutionary Biology, 2010. **10**(1): p. 308.

93. Kirby, I.T., A. Person, and M. Cohen, *Rational design of selective inhibitors of PARP4.* RSC Medicinal Chemistry, 2021. **12**(11): p. 1950-1957.

94. Zhang, W., et al., *Feedback regulation on PTEN/AKT pathway by the ER stress kinase PERK mediated by interaction with the Vault complex.* Cellular Signalling, 2015. **27**(3): p. 436-442.

95. Barkauskaite, E., G. Jankevicius, and I. Ahel, *Structures and Mechanisms of Enzymes Employed in the Synthesis and Degradation of PARP-Dependent Protein ADP-Ribosylation.* Molecular Cell. **58**(6): p. 935-946.

96. van Kempen, M., et al., *Fast and accurate protein structure search with Foldseek.* Nature Biotechnology, 2023.

97. Vyas, S., et al., *Family-wide analysis of poly(ADP-ribose) polymerase activity.* Nature Communications, 2014. **5**(1): p. 4426.

98. Yang, C.-S., et al., *Ubiquitin Modification by the E3 Ligase/ADP-Ribosyltransferase Dtx3L/Parp9.* Molecular Cell, 2017. **66**(4): p. 503-516.e5.

99. Bonfiglio, J.J., et al., *An HPF1/PARP1-Based Chemical Biology Strategy for Exploring ADP-Ribosylation.* Cell, 2020. **183**(4): p. 1086-1102.e23.

100. Suskiewicz, M.J., et al., *HPF1 completes the PARP active site for DNA damage-induced ADP-ribosylation.* Nature, 2020. **579**(7800): p. 598-602.

101. Yi, C., et al., *Major Vault Protein, in Concert with Constitutively Photomorphogenic 1, Negatively Regulates c-Jun–Mediated Activator Protein 1 Transcription in Mammalian Cells.* Cancer Research, 2005. **65**(13): p. 5835-5840.

102. Ikeda, R., et al., *Hyperosmotic stress up-regulates the expression of major vault protein in SW620 human colon cancer cells.* Experimental Cell Research, 2008. **314**(16): p. 3017-3026.

103. Wakatsuki, S., et al., *Small noncoding vault RNA modulates synapse formation by amplifying MAPK signaling.* J Cell Biol, 2021. **220**(2).

104. Eichenmüller, B., et al., *Vaults bind directly to microtubules via their caps and not their barrels.* Cell Motil Cytoskeleton, 2003. **56**(4): p. 225-36.

105. Yu, K., et al., *Understanding the Interaction between Major Vault Protein and INT for Application in Modulated Molecular Release*. 2013.

106. Yu, K., et al., *Modulation of the Vault Protein-Protein Interaction for Tuning of Molecular Release.* Scientific Reports, 2017. **7**(1): p. 14816.

107. Li, J., et al., *A conserved NAD+ binding pocket that regulates protein-protein interactions during aging.* Science, 2017. **355**(6331): p. 1312-1317.

108. Audrito, V., et al., *The Extracellular NADome Modulates Immune Responses.* Frontiers in Immunology, 2021. **12**.

109. Gagné, J.-P., et al., *Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes.* Nucleic Acids Research, 2008. **36**(22): p. 6959-6976.

110. *Transfecting Plasmid DNA Into A549 Cells Using Lipofectamine 3000 Reagent*. 2020; Available from: https://www.thermofisher.com/us/en/home/references/protocols/cell-culture/transfection-protocol/a549-cells-protocol.html#order.

111. Carter-O'Connell, I., et al., *Combining Chemical Genetics with Proximity-Dependent Labeling Reveals Cellular Targets of Poly(ADP-ribose) Polymerase 14 (PARP14).* ACS Chemical Biology, 2018. **13**(10): p. 2841-2848.

112. Sukhanova, M.V., et al., *A Single-Molecule Atomic Force Microscopy Study of PARP1 and PARP2 Recognition of Base Excision Repair DNA Intermediates.* Journal of Molecular Biology, 2019. **431**(15): p. 2655-2673.

113. Bonfiglio, J.J., et al., *Serine ADP-Ribosylation Depends on HPF1.* Molecular Cell, 2017. **65**(5): p. 932-940.e6.

114. Dasovich, M. and A.K. Leung, *Molecular tools unveil distinct waves of ADP-ribosylation during DNA repair.* Cell Reports Methods, 2023. **3**(5).

115. Carter-O'Connell, I., et al., *Engineering the Substrate Specificity of ADP-Ribosyltransferases for Identifying Direct Protein Targets.* Journal of the American Chemical Society, 2014. **136**(14): p. 5201-5204.

116. Alemasova, E.E. and O.I. Lavrik, *Poly(ADP-ribosyl)ation by PARP1: reaction mechanism and regulatory proteins.* Nucleic Acids Res, 2019. **47**(8): p. 3811-3827.