A Framework for Studying Crucial Steps in Proteasome Core Particle Assembly

By

Elizabeth Grotemeyer

B.S., California State University Dominguez Hills, 2015

Submitted to the graduate degree program in Molecular Biosciences and the Graduate Faculty
of the University of Kansas in partial fulfillment of the requirements
for the degree of Master of Arts.

_____

Chair: Eric Deeds

_____

Christopher Fischer

_____

Roberto de Guzman

Date Defended: 20 July 2017

The thesis committee for Elizabeth Grotemeyer certifies that this is
the approved version of the following thesis:

A Framework for Studying Crucial Steps in Proteasome Core Particle Assembly

_____

Co-Chair: Eric Deeds

_____

Co-Chair: Christopher Fischer

Date Approved: 20 July 2017

Abstract

The ability of proteins to repeatedly and reliably self-assemble in the cell is a critical element of the maintenance of life. Despite this, the mechanisms that underlie these events are poorly understood. The proteasome core particle from *Rhodococcus erythropolis* is an excellent model system for understanding assembly processes. This system has a two step assembly pathway where individual subunits first assemble into half proteasomes. Then, two half proteasomes dimerize to produce a full proteasome core particle. The beta subunit of this complex is synthesized in an inactive with an N-terminal propeptide that is cleaved after assembly is complete, rendering the CP enzymatically active. Evidence suggests that the propeptides plays a crucial role in both steps of the assembly process. To date, however, it has been impossible to fully characterize the role of the propeptide in assembly because this protein is typically produced as a heterogeneous mixture with a variety of N-terminally truncations in the propeptides itself. Here, we used Ligation Independent Cloning to produce a beta variant, which we call D3, that is homogeneous for the full-length propeptide. We also used Native PAGE to begin to characterize the kinetics of half proteasome dimerization. We found that there is a temperature-dependent effect on the dimerization process and that the presence of the full propeptide dramatically slowed assembly when compared to the heterogeneous beta. Using these methods, we can now study the thermodynamics and kinetics of this system much more rigorously than has been possible to date.

Contents

List of Figures

Introduction

It is well established in biology that proteins carry out most of the processes necessary for the maintenance of life. However, it is often not the case that individual proteins carry out these tasks. Rather, it is large macromolecular complexes that are generally responsible for these processes.[1] These large molecular machines are composed of many different proteins that must assemble correctly and efficiently in the cell. The ability of proteins to consistently and reliably self-assemble into the macromolecular complexes necessary for carrying out cellular processes is a fascinating and poorly understood phenomenon in biology. To better understand the mechanisms that drive assembly, we use the proteasome from the bacterial species *Rhodococcus erythropolis* as a model system.

The proteasome is a large multiprotein nonlysosomal protease complex that that serves the critical cellular function of regulating protein quality and amount by irreversibly degrading proteins.[2] The proteasome is found in all domains of life, though its presence in bacteria is limited to the actinomycetes which include *Rhodococcus erythropolis* and *Mycobacterium tuberculosis*[3]. The catalytically active core particle of the proteasome is comprised of 14 α- and 14 β-type subunits which self-assemble in solution to form a barrel structure composed of four stacked heptameric rings in the order $\alpha_7\beta_7\beta_7\alpha_7$ that is 15nm long and 11nm in diameter.[4,5] The inside of the barrel is separated into three chambers separated by narrow constrictions which only unfolded proteins can pass through.[6,7] This overall structure is conserved across all domains of life, however, the subunit composition of proteasomes varies greatly between them.[8–11] In eukaryotes, there are 7 distinct alpha genes and 7 distinct beta genes while in archaea and bacteria there is only one of each.[12] The only known exception to this is *R. erythropolis* which

has two copies of each subunit[3], though the biological relevance of this is still unclear since it is possible to produce catalytically active proteasomes using only one version of each subunit.[13,14]

The proteasome is a member of the N-terminal nucleophile family. The catalytically active N-terminal threonine residues on the beta subunits are contained within the barrel allowing the cell to regulate the activity of the proteasome by sequestering the proteolytic residues away from the cytoplasm. Regulation of substrates that are to be degraded is controlled by the activity of the 19S regulatory particles present on either end of the barrel opening. These regulatory particles recognize cellular proteins that have been marked for degradation and unfold them so that they can be passed through the proteasome core particle where they are cleaved into small peptides by the proteolytic activity of the active sites.[15–17]

In bacterial and archaeal species, proteasome subunits will self-assemble in vitro into active core particles in a two-step process. First, monomeric subunits come together to form half proteasomes comprised of a single α-ring and a single β-ring ($\alpha_7\beta_7$). Then, two half proteasomes dimerize to produce a full, active proteasome core particle.[14] These processes occur on very different timescales in vitro: monomer assembly into half proteasomes occurs much faster than half proteasome dimerization. It has been shown repeatedly that half proteasome assembly is nearly complete within a minute of mixing at 37°C while dimerization to full proteasomes requires approximately 3 hours to go to completion under the same conditions.[14]

The beta subunits are synthesized with N-terminal propeptides that vary dramatically in length between species.[12] In *R. erythropolis* the propeptide is 65 amino acids long. The propeptides renders the β subunits inactive, which protects the cell from uncontrolled protease activity during assembly. Once the barrel has formed and the catalytic activity has been sequestered, the propeptides are autocatalytically cleaved resulting in active proteasome core

particles.[18] When the beta subunits are expressed separately for in vitro assembly studies, the protein manifests as a series of bands on an SDS PAGE gel rather than as a single band as is expected for pure proteins. Sequencing and mass spectrometric analysis have revealed that these additional bands are N-terminally truncated versions of the protein.[13] These truncations are restricted to the propeptide and range from -18 to -50 amino acids.[19]

In some species, including *R. erythropolis*, the propeptide has also been shown to be a critical intramolecular chaperone for assembly. Crystal structures in which the propeptide is not cleaved demonstrate that the C-terminal region of the propeptide plays an important role in the formation of the half proteasome by providing a large portion of the interface between the beta subunit and its neighboring alpha subunits[20]. It is important to note that in this crystal structure, there is a significant portion of the N-terminal region of the propeptide that is unresolved. It is also worth noting that there are no crystal structures of the half proteasome in its native conformation.

In assembly studies in which the propeptide was completely removed, half proteasomes were shown to dimerize to form complete core particles much more slowly than in wild type assembly studies. When the propeptide was added back exogenously as a short peptide, core particle formation was nearly instantaneous.[14] These studies suggest a critical, but not essential, role for the propeptide in both steps of assembly.

Though it is clear that the propeptide participates in the interactions that result in the formation of the proteasome core particle, it is not yet clear what that role is. In order to fully characterize the role the propeptide plays in the assembly process, we must first develop the tools necessary to address two key problems. First, we must develop a beta construct that produces a truly homogeneous population of beta with a well-defined propeptide length and sequence. Then

we must also develop assays that allow us to quantitatively measure the rates of the various steps in assembly. In addition, these assays must allow us to separate initial half proteasome assembly from half proteasome dimerization so that the rates of these processes can be determined separately.

Development of the D3 Beta Construct

Previous constructs for the production of the beta subunit have all failed to produce a homogenous population of the protein. Instead, these constructs (including the pET22B beta construct (Fig 1) used in our lab) have resulted in the presence of a series of bands that are beta proteins that have been truncated to various degrees from the N-terminus of the protein.[19] It is thought that these truncations are the result of *E. coli* proteases during the expression process, though it has never been conclusively demonstrated that this is the cause.[13]

To address this problem, we developed a method for producing the beta subunit of the *R. erythropolis* proteasome that does not have these truncations. We accomplished this by using Ligation Independent Cloning (LIC) to produce the

Figure 1: pET22B and D3 construct designs
The pET22B construct includes an uncleaved C-terminal His6 tag. The D3 construct contains a TEV cleavable N-terminal MBP which protects the N-terminus of the protein from cellular proteases during growth and purification.

beta subunit with an N-terminal TEV cleavable MBP. This approach allowed us to protect the N-terminus of the beta protein until it had been purified away from other cellular proteases.

The beta sequence was inserted into the pET His6 MBP N10 TEV LIC cloning vector (2C-T), a gift from Scott Gradia (Addgene plasmid # 29706). The correct insertion of the beta gene was confirmed by DNA sequencing. This construct was then cloned into BL21-DE3 cells containing the PLysS vector for expression. The protein was initially isolated from clarified cell lysate using metal affinity chromatography. The purified protein was dialyzed at 4°C overnight with TEV protease to cleave the beta from the MBP. After TEV cleavage, the protein underwent another round of Co-Affinity chromatography to remove the MBP. The final purification step was anion exchange chromatography which resulted in highly pure protein. Despite the purity of
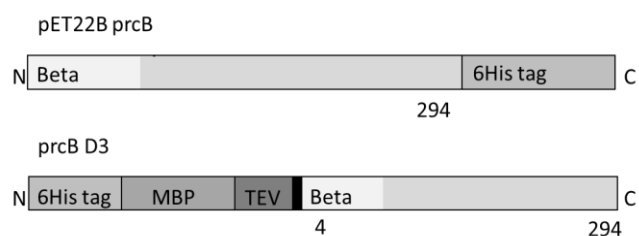
the protein, this construct did not produce proteasome core particles that were catalytically active even after 24 hours of incubation with the alpha subunit at 37°C based on the standard enzymatic activity assay for proteasomes.[13]

Since the beta propeptide in *R. erythropolis* is very long, we hypothesized that the additional three amino acids remaining on the N-terminus of the propeptide after TEV cleavage were responsible for the lack of activity in our construct. Thus, we produced a variant in which we removed the first three amino acids from the beta propeptides; we
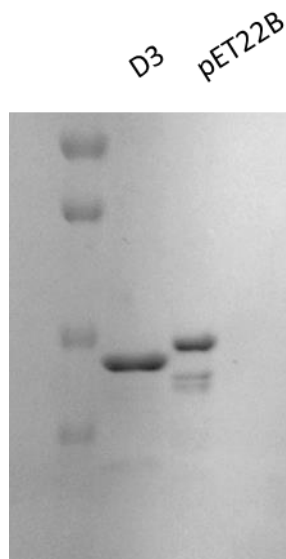

Figure 2: SDS PAGE of the D3 and pET22B beta proteins
The D3 construct produces a single population of beta protein while the pET22B construct produces a heterogeneous mixture of beta proteins. The difference in size of the major bands is the result of an uncleaved C-terminal 6his tag in the pET22B construct.

named this construct D3 (Fig 1). This resulted in a beta subunit that had a final length that was equivalent to the full-length protein with a sequence variation in the first three amino acids. Following the same cloning and purification approaches describe above, we were able to produce homogeneous beta, confirmed by SDS PAGE (Fig 2) and intact mass spectrometry. In initial assembly and activity assays, this mutant assembled well and displayed enzymatic activity similar to that of previous constructs.

By protecting the N-terminus of the beta subunit from cleavage during expression, we are now able to produce homogeneous populations of the beta subunit, which makes quantitative exploration of the role of the propeptide in the assembly process possible by allowing for the production of deliberate truncations or mutations aimed at identifying critical regions or residues in the propeptide. Mass spectrometric studies of the beta truncations have indicated that there are 12 variants in the heterogeneous beta that are incorporated into proteasome core particles with

various levels of success.[19] Using the LIC approach, we have been able to produce plasmids for the expression of all of these variants. While the studies described below focus on the D3 protein, by expressing homogeneous populations of these truncation variants we will be able to distinguish the roles that various regions of the propeptide are playing in the assembly process in future work.

Initial Characterization of the Dimerization Rate of Half Proteasomes

Before we are able to clarify the role of the propeptide in assembly, we must first develop an assay that will allow us to distinguish between the two steps of assembly: half proteasome formation and half proteasome dimerization. To do this, we used native PAGE analysis of assembly timecourses, which gives information about the concentration of the half proteasome and the full proteasome at each time point.

For the assembly reactions, 8 µM alpha and 8 µM beta subunits in HNE (20 mM HEPES, 20 mM NaCl, 1 mM EDTA) were mixed and flash frozen in 20 µL aliquots and stored at -80°C. At the start of each time point, samples were removed from the freezer in sets of 4 replicates, allowed to thaw for 1 minute at 4°C, centrifuged for 15 seconds, and then transferred to a water bath at 4°C, 10°C, 17°C, 22°C, 30°C, or 37°C. At the end of the time course series, loading buffer was added to each sample and 20 uL of the reaction mix were then loaded onto a Novex Tris glycine 4-20% native gel along with BSA standards of 0.6, 1, and 1.5 µM. The gels were then stained in SYPRO Ruby protein stain per the manufacturer's instructions and imaged using a densitometer.

Representative gels from each set of experiments are shown in Figure 3. Band intensities of the half and full proteasomes, as well of the BSA standards, were quantified using ImageJ software. The concentration of half and full proteasomes were determined from these intensities using a standard curve generated from the BSA. By running the BSA standard on each gel, we were able to correct for differences in staining intensity on a gel-to-gel basis, allowing for accurate normalization of data from multiple gels (Fig 4).

This data indicates several things about the assembly process. First, there is a clear difference in the assembly rates of each of the constructs over this temperature range. For example, in the pET22B construct, assembly is near complete after 30 minutes at 30°C while the D3 construct is has only partially assembled into full proteasomes. Additionally, it is interesting to note that at all temperatures and for both constructs, assembly of the half proteasome is nearly complete at 0 min. This highlights the extreme separation of timescales for these two processes.

These experiments also highlight the temperature dependence of these reactions. At higher temperatures, both constructs are capable of approaching complete assembly into full proteasomes, however at 4°C assembly is dramatically slowed. Furthermore, at low temperatures it appears that the core particle reaches saturation at a much lower concentration than at higher temperatures (Fig 2), though the reason for this is still unclear.
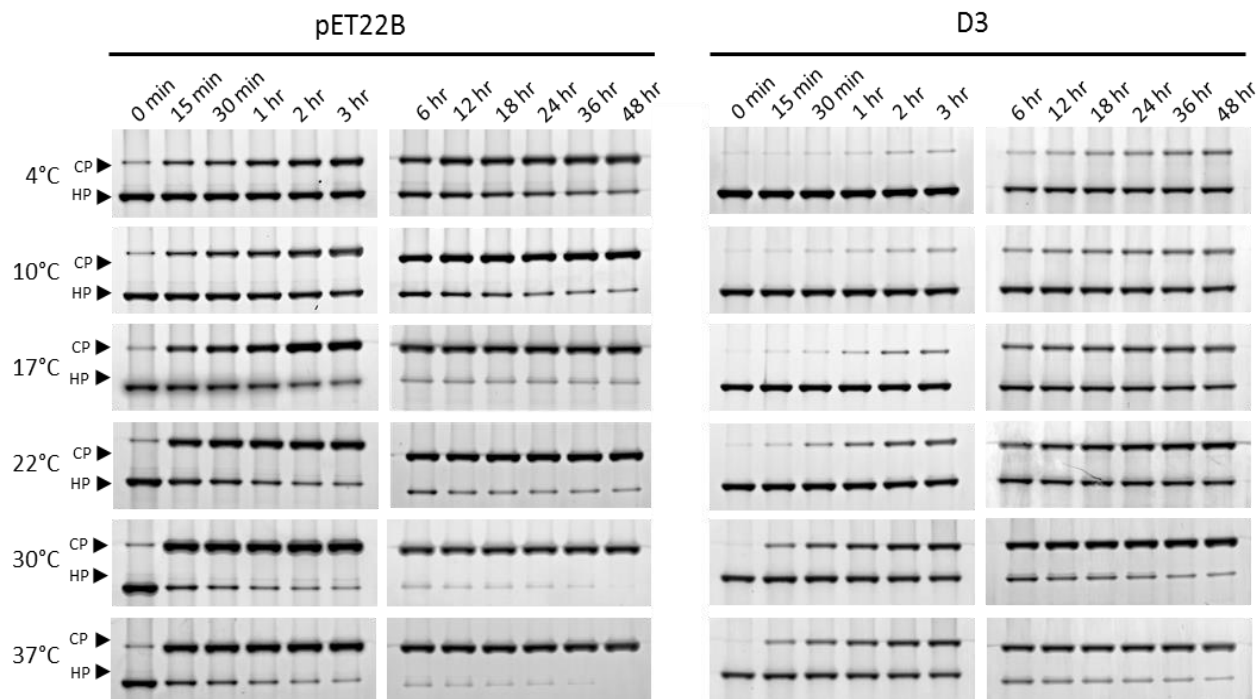

Figure 3: Native PAGE of proteasome assembly assays
8 µM of each subunit were allowed to assemble at each temperature over the course of 48 hours. The assembly reaction was then analyzed via Native PAGE stained with SYPRO Ruby.
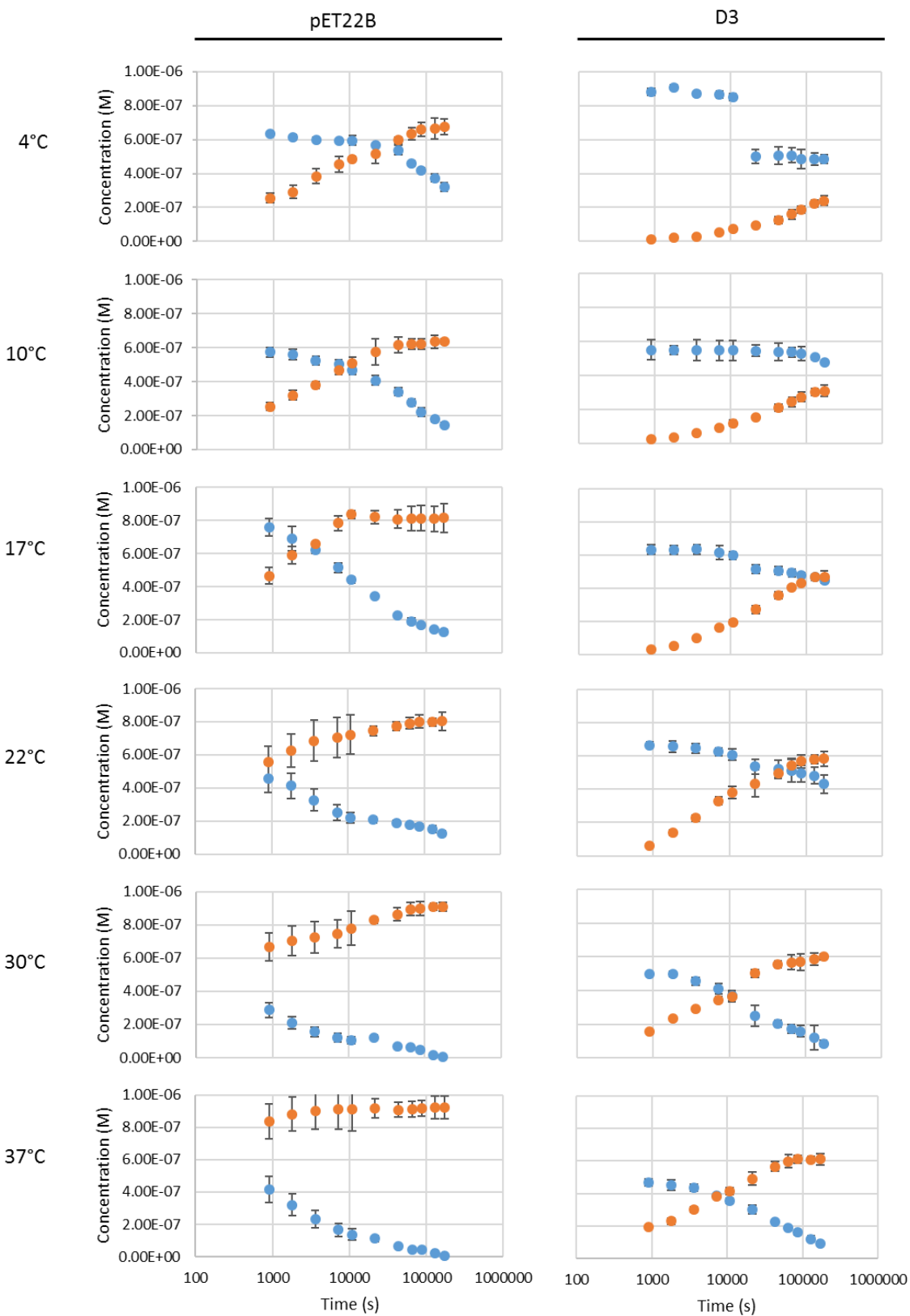
Figure 4: Concentrations of half and full proteasomes
The concentrations of half (blue) and full (orange) proteasomes were determined from quantitative analysis of native gel bands
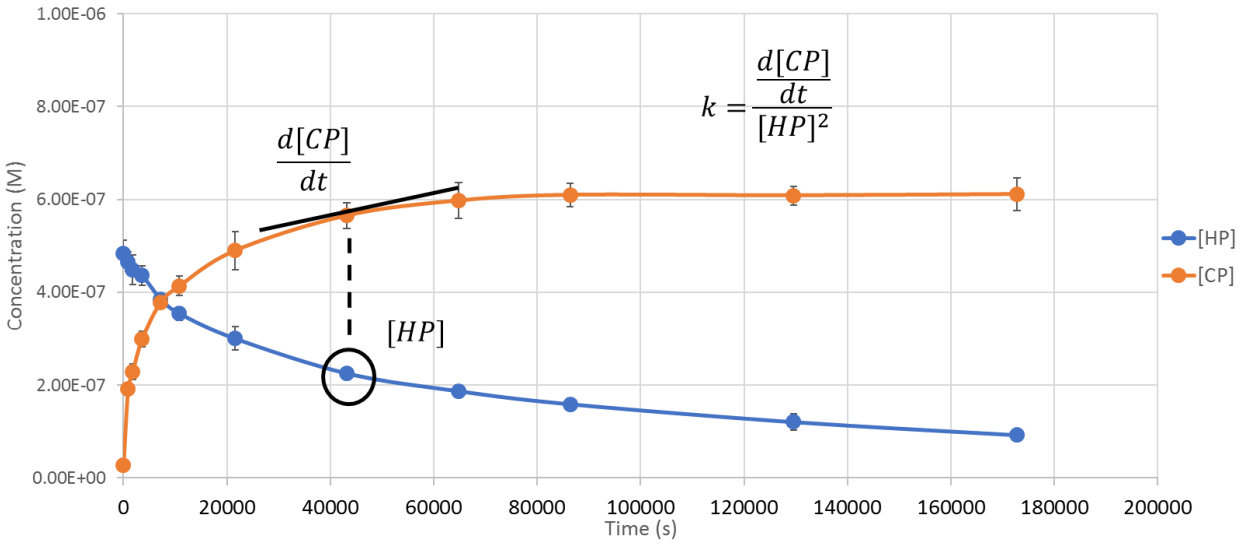
Figure 5: Determination of Rate Constants

The concentration of core particle over time was fit using a spline function using the statistical software R, which generates a differentiable curve fit to each data set. Using these, and the concentration of half proteasomes identified in Figure 4, we could then determine the rate constant for each construct at each temperature according to Equation 1:

$$rate\ of\ dimerization = \frac{d[CP]}{dt} = k[HP]^2 \qquad [1]$$

This process is outlined in Figure 5. Once rate constants had been determined, it was then possible to create an Arrhenius plot for both constructs (Fig 6).

In this plot it is clear that the temperature dependence is linear for both constructs, indicating that the reaction kinetics are governed by standard thermodynamics in this temperature range. The slopes of these lines are not significantly different from one another (t-test, $p = 0.05$). With this data we can begin to identify the contributions of the energetic components by examining the Arrhenius equation (Eq. 2) and its linear form (Eq. 3):

$$k = \omega_0 e^{\frac{-\Delta G^{\ddagger}}{RT}} \qquad [2]$$

$$[3]$$

$$\ln k = \frac{-\Delta G^{\ddagger}}{RT} + \ln \omega_0.$$

Separating $\Delta G^{\ddagger}$ into its components according to Eq 4 gives Eq 5:

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} + T\Delta S^{\ddagger} \qquad [4]$$

$$\ln k = -\left(\frac{\Delta H^{\ddagger}}{RT} + \frac{\Delta S^{\ddagger}}{R}\right) + \ln \omega_0. \qquad [5]$$

This can be rewritten in the form $y = mx + b$ as in Eq 6:

$$\ln k = \left(\ln \omega_0 - \frac{\Delta S^{\ddagger}}{R}\right) - \left(\frac{\Delta H^{\ddagger}}{R}\right)\frac{1}{T}. \qquad [6]$$
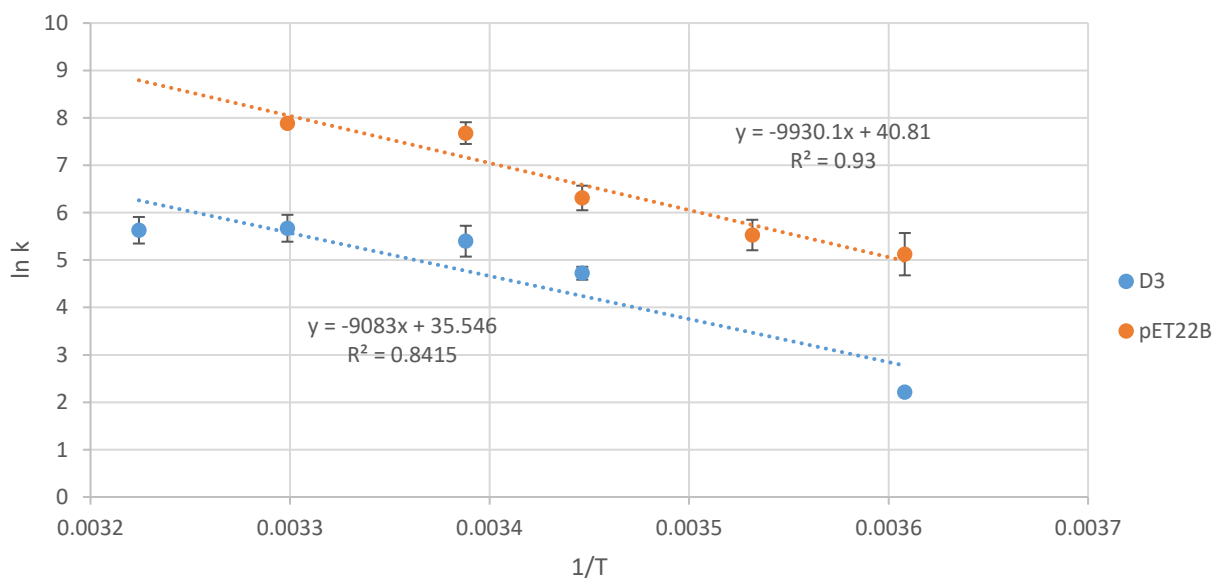

Figure 6 Arrhenius Plot for the pET22B and D3 Constructs

Since the slopes of the lines in Figure 5 are statistically indistinguishable, the enthalpic contributions to the activation free energy of the constructs are the same. This means that the energetic differences that result in the differences in assembly rates are dictated by the terms that define the intercept, which we will call $\widehat{\beta_0}$. The term ω_0 is the frequency factor, which is likely

to be constant for these different constructs, which leaves the entropic contribution as the primary source of the difference (Eq 7):

$$\widehat{\beta_0}(D3) - \widehat{\beta_0}(pET22B) = \frac{\Delta S^{\ddagger}(pET22B)}{R} - \frac{\Delta S^{\ddagger}(D3)}{R} = 5.471$$

This means that the difference in the entropic component of the free energy is approximately 11 cal/mol·K. Note that the main difference between the two constructs is the presence of a heterogeneous mix of propeptides lengths in pET22B vs. a homogeneous population in D3 (Fig. 2). This supports a model in which a major determinant of the free energy barrier to dimerization is the fact that the propeptides must move "out of the way" in order for dimerization to occur, since the entropic cost of that conformational change would likely be lower for subunits with shorter propeptides lengths.

While there are many interesting observations to be drawn from this data, it is currently not possible to make definitive conclusions about the role of the propeptide on the assembly process. This data does demonstrate, however, that our approach can be used to identify differences in the assembly rates of different constructs, which will be critical for furth [7] lies into the specific role that the propeptide is playing in this process.

Discussion

The primary purpose of this work was to develop the experimental and analytical framework necessary for determining the roles that the propeptide plays in proteasome core particle assembly. The method that we describe here that allows for the production of a homogeneous population of beta subunits has and will continue to be a critical development for the understanding of the assembly of the propeptide. With this development we are now able to begin to address questions about the role that the propeptide is playing in the assembly process in a way that was not previously possible.

It was also necessary to develop a framework that allows us to explore the distinct steps in the assembly processes individually. By using native PAGE, we can track changes in both the half proteasome and full proteasome concentrations individually over time. This allows us to separately quantify the rates of the processes that produce the half proteasome from those responsible for dimerization.

Initial studies using these approaches have yielded promising preliminary data demonstrating the validity of this approach. The presence of the full-length propeptide generated using the D3 construct slows the dimerization process dramatically when compared to a heterogeneous mix produced using pET22B. It is also quite evident that there is a dramatic temperature dependence for this dimerization process. Our approach allows for in-depth study of the thermodynamic properties of this system, which has not been possible to date.

Though it is clear that this framework can be used to complete this analysis, there are some issues that will need to be addressed in subsequent experiments. In particular, for fast reactions, there is a lot of data that is missing given the time points chosen here. As an example, in the pET22b construct at 37°C, the assembly process is already near completion after 15

minutes, which is only the second time point in this analysis. This means that much of the data regarding the initial dimerization rate is lost to us. In future work, it will be necessary to increase the number of data points collected in the earliest parts of the time course.

Future experiments should also focus on using homogeneous, N-terminally truncated versions of the beta protein, which, in conjunction with the data here for the full-length beta, will be able to begin to address the questions of the role of the propeptide in the dimerization process of the proteasome assembly pathway.

References

1. Alberts, B. The Cell as a Collection of Protein Machines : Preparing the Next Generation of Molecular Biologists. **92,** 291–294 (1998).
2. Hochstrasser, M. Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. *Curr. Opin. Cell Biol.* **7,** 215–223 (1995).
3. De Mot, R., Nagy, I. & Baumeister, W. A self-compartmentalizing protease in Rhodococcus: The 20S proteasome. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **74,** 83–87 (1998).
4. Baumeister, W. *et al.* Electron microscopy and image analysis of the multicatalytic proteinase. *FEBS Lett.* **241,** 239–245 (1988).
5. Zwickl, P., Kleinz, J. & Baumeister, W. Critical elements in proteasome assembly. *Nat. Struct. Biol.* **1,** 765–770 (1994).
6. Aoyama, K., Zühl, F., Tamura, T. & Baumeister, W. 2-D Crystallization of the Rhodococcus 20S Proteasome. *J. Struct. Biol.* **116,** 438–442 (1996).
7. Aoyama, K. 2D-crystallization of Rhodococcus 20S proteasome at the liquid-liquid interface. *J. Cryst. Growth* **168,** 198–203 (1996).
8. Kunjappu, M. J. & Hochstrasser, M. Assembly of the 20S proteasome. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843,** 2–12 (2014).
9. Dahlmann, B. *et al.* The multicatalytic proteinase (prosome) is ubiquitous from eukaryotes to archaebacteria. *FEBS Lett.* **251,** 125–131 (1989).
10. Tamura, T. *et al.* The first characterization of a eubacterial proteasome: the 20S complex of Rhodococcus. *Curr. Biol.* **5,** 766–774 (1995).
11. Jastrab, J. B. & Darwin, K. H. Bacterial Proteasomes. *Annu. Rev. Microbiol.* **69,** 109–127 (2015).
12. Marques, A. J., Palanimurugan, R., Matias, A. C., Ramos, P. C. & Dohmen, R. J. Catalytic Mechanism and Assembly of the Proteasome Catalytic Mechanism and Assembly of the Proteasome. *Chem. Rev.* **109,** 1509–1536 (2009).
13. Zühl, F. *et al.* Subunit topology of the Rhodococcus proteasome. *FEBS Lett.* **400,** 83–90 (1997).
14. Zühl, F., Seemüller, E., Golbik, R. & Baumeister, W. Dissecting the assembly pathway of the 20S proteasome. *FEBS Lett.* **418,** 189–194 (1997).
15. Baumeister, W., Walz, J., Zühl, F. & Seemüller, E. The proteasome: Paradigm of a self-compartmentalizizing protease. *Cell* **92,** 367–380 (1998).
16. Zwickl, P., Baumeister, W. & Steven, A. Dis-assembly lines: The proteasome and related ATPase-assisted proteases. *Curr. Opin. Struct. Biol.* **10,** 242–250 (2000).
17. Groll, M. & Huber, R. Substrate access and processing by the 20S proteasome core particle. *Int. J. Biochem. Cell Biol.* **35,** 606–616 (2003).
18. Chen, P. & Hochstrasser, M. Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell* **86,** 961–972 (1996).
19. Sharon, M., Witt, S., Glasmacher, E., Baumeister, W. & Robinson, C. V. Mass spectrometry reveals the missing links in the assembly pathway of the bacterial 20 S proteasome. *J. Biol. Chem.* **282,** 18448–18457 (2007).
20. Kwon, Y. Do, Nagy, I., Adams, P. D., Baumeister, W. & Jap, B. K. Crystal structures of the Rhodococcus proteasome with and without its pro-peptides: Implications for the role of the pro-peptide in proteasome assembly. *J. Mol. Biol.* **335,** 233–245 (2004).