Facile Method of Constructing Polyproteins for Single-Molecule Force Spectroscopy Studies

Peng Zheng, Yi Cao, † and Hongbin Li *
Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

ABSTRACT: Constructing polyproteins consisting of identical tandem repeats of proteins provides an unambiguous method of investigating the mechanical properties of proteins at the single-molecule level using force spectroscopy techniques. Here we report a maleimide–thiol coupling-based facile method of constructing polyproteins for single-molecule force spectroscopy studies on the mechanical properties of proteins. This method allows for the construction of polyproteins in an efficient fashion under room temperature. The resultant thioether bonds are resistant to reduction and make it possible to carry out single-molecule force spectroscopy studies under various redox conditions. This novel method complements existing polyprotein engineering methods and can be easily applied to a wide variety of proteins.

INTRODUCTION

Protein mechanics plays an important role in a wide variety of biological processes across different length scales.1–3 Many proteins are subject to the mechanical stretching force under their biological conditions and play various structural and mechanical roles. Elastomeric proteins are one representative class of such mechanical proteins, and many of them are tandem modular proteins made of multiple individually folded domains.4,5 They function as molecular springs under their biological settings to provide tissues with elasticity, extensibility, and strength3,6 and also play important regulatory roles in a wide variety of biological processes.2 The development of single-molecule force spectroscopy techniques, in particular, the atomic force microscopy (AFM)-based force spectroscopy techniques, over the last two decades has made it possible to measure the mechanical properties of elastomeric proteins at the single-molecule level and has provided tremendous insight into the mechanical design of elastomeric proteins.4,7–11 Moreover, single-molecule AFM has evolved into a general method to characterize the mechanical unfolding/folding dynamics of proteins along a well-defined reaction coordinate at the single-molecule level.7,10,12

In single-molecule AFM studies, it is critical to unambiguously distinguish single-molecule stretching events of the protein of interest from nonspecific interactions between the AFM tip and sample as well as from non-single-molecule events.13,14 Constructing polyproteins made of multiple identical tandem repeats of proteins of interest provides unambiguous fingerprints for identifying single-molecule events and has thus become a gold standard in single-molecule AFM studies of protein mechanics. Because polyproteins are made of identical tandem repeats of proteins, the resultant force—extension curves will exhibit characteristic sawtooth-like force—extension curves with identical contour length increments between consecutive sawtooth peaks. Such repetitive patterns allow one to identify single-molecule stretching events readily and unambiguously from a myriad of nonspecific interactions and complex multimolecule stretching events.13,14

The construction of polyproteins is typically accomplished using a recombinant DNA technology-based DNA concatamerization method.15 This method provides precise control of the molecular structure and composition of the resultant polyprotein and has been the most widely used method.13–17 However, it requires stepwise repetitive cloning and is laborious and expensive. To facilitate the faster construction of polyproteins, alternative methods have been pursued. A cysteine engineering-based method was developed for this purpose. This method was first successfully demonstrated in a solid-state synthesis of polyproteins based on the specific crystallographic arrangement of proteins in their crystals18 and later was generalized to proteins in solution.19,20 In this method, a pair of cysteines were introduced into the protein of interest at chosen locations. The subsequent oxidation of cysteines results in the polymerization of the protein via the formation of intermolecular disulfide bonds. This method provides an efficient alternative to the recombinant DNA technology-based polyprotein engineering approach and also allows the stretching of polyproteins in any defined pulling direction.19 This method has found many applications in a variety of proteins.21,22 However, the oxidation of cysteines is a slow process, and the disulfide linkage in the resultant polypeptides limits the use of this method of studying redox-dependent mechanical
properties of proteins.\textsuperscript{20} Here, we report a facile method based on maleimide–thiol coupling chemistry to construct polyproteins to complement these existing polyprotein engineering methods.

\section*{MATERIALS AND METHODS}

\textbf{Protein Engineering.} The model proteins we used in this study are GB1, the B1 IgG binding domain of protein G from Streptococcus, and rubredoxin from \textit{C. pasteurianum} (cpRD). The gene coding cys-GB1-cys was amplified by using the polymerase chain reaction (PCR) and subcloned into the expression vector pQE80L (Qigen). The gene encoding cys-cpRD-GB1-cys was constructed using standard molecular biology techniques and cloned into expression vector pQE80L between \textit{BamH}I and \textit{Kpn}I sites. The sequences of both genes were verified by direct DNA sequencing. Both constructs were overexpressed in \textit{E. coli} strain DH5\textsubscript{α} and purified using Co\textsubscript{2}⁺-affinity chromatography. The purified protein was digested with thrombin to obtain a 100 kDa protein chimera (GB1-cpRD) and purified using a TALON His-Tag purification resin (Clontech). The purified protein was digested with thrombin to obtain a 100 kDa protein chimera and purified using a TALON His-Tag purification resin (Clontech). The purified protein was digested with thrombin to obtain a 100 kDa protein chimera and purified using a TALON His-Tag purification resin (Clontech).

\textbf{Construction of Polyproteins through a Maleimide-Thiol Coupling Chemistry.} BM(PEO)$_3$ can link two GB1s together by reacting sequentially with the thiol groups on the protein via maleimide groups on both ends to form stable thioether bonds and further extension leading to the formation of polyprotein (GB1)$_n$.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Construction of polyprotein using maleimide—thiol coupling chemistry. (A) The maleimide group can specifically react with the sulphydryl group, forming a covalent thioether bond that cannot be cleaved under reducing conditions. (B) Construction of the polyprotein (GB1)$_n$ using the maleimide—thiol coupling chemistry. BM(PEO)$_3$ can link two GB1s together by reacting sequentially with the sulphydryl groups on the protein via maleimide groups on both ends to form stable thioether bonds and further extension leading to the formation of polyprotein (GB1)$_n$.}
\end{figure}

\section*{RESULTS AND DISCUSSION}

\textbf{General Principle.} Maleimide—thiol coupling-based bioconjugation methods have been used extensively in biochemistry because of the rapid and specific reaction of maleimide with sulphydryl groups under mild experimental conditions.\textsuperscript{25,26} Here we employ this maleimide–thiol coupling chemistry to develop an efficient method of constructing polyproteins for single-molecule AFM studies. The basic experimental design is schematically

solution of BM(PEO)$_3$ (1,8-bis-maleimido-(PEO)$_3$, Molecular Biosciences) in Tris buffer at pH 7.4. The solution was incubated at room temperature for \~{}2 h, and then the resultant polyprotein solution was directly used in AFM experiments.

\textbf{Characterization of Polyproteins Constructed Using Maleimide—Thiol Coupling Chemistry.} Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed to examine the molecular weight of the resultant polyproteins. The polymerization efficiency of GB1 was also analyzed using a gel filtration chromatography technique with a Superdex 75 10/300GL column in an ATKA FPLC system (GE Healthcare). The buffer contains 50 mM sodium phosphate plus 100 mM sodium chloride at pH 6.5, and the flow rate was 0.5 mL/min.

\textbf{Single-Molecule AFM Experiments.} Single-molecule AFM experiments were carried out on a custom-built AFM as described.\textsuperscript{24,25} Each Si$_3$N$_4$ cantilever (MLCT, Bruker Corp.) was calibrated in solution using the equipartition theorem before each experiment to obtain the spring constant (typically around 40 pN/nm). All experiments were done in Tris buffer at pH 7.4 at room temperature.

In a typical experiment, 2 μL of the polyprotein sample (at a concentration of \~{}2 mg/mL) was added to a clean glass coverslip covered with \~{}50 μL of Tris buffer. The protein was allowed to absorb for \~{}5 min before starting the AFM measurements. During the experiment, the cantilever was brought into contact with the substrate at a contact force of \~{}1 nN to pick up proteins. The pulling speed was 400 nm/s.

For experiments on the Fe(II)-cpRD-GB1 polyprotein, 10 μL of 200 mM dithioreitol (DTT) was added to the solution after the (Fe(II)-cpRD-GB1)$_n$ protein was absorbed on the glass coverslip. AFM experiments started after incubation for \~{}20 min. The same amount of DTT solution was added every hour thereafter to ensure a reduced environment.
depicted in Figure 1. First, two cysteine residues will be introduced into the protein of interest via site-directed mutagenesis at well-defined, solvent-exposed positions to provide accessible sulfhydryl groups. Then, the protein will react with bifunctional maleimide compound BM(PEO)$_3$ in a molar ratio of 1:1, leading to the polymerization of the protein of interest via the formation of intermolecular thioether bonds. In bifunctional maleimide compound BM(PEO)$_3$, PEO serves as a flexible linker with increased solubility. (PEO)$_3$ has a contour length of $\sim$1.5 nm and can effectively minimize domain-domain interactions because of the formation of thioether bonds.26,27

**Synthesis and Characterization of the Polyprotein (GB1)$_n$ Constructed via Maleimide–Thiol Coupling Chemistry.** To validate the design of this maleimide–thiol coupling-based polyprotein engineering method, we first constructed a model polyprotein of GB1, which is an excellent model of an artificial elastomeric protein domain.28,29 Polyprotein (GB1)$_n$ has been constructed using a recombinant DNA technology-based strategy, and its mechanical properties have been characterized in great detail in our previous single-molecule AFM studies.28,29 Thus, constructing the GB1 polyprotein will provide a good test case for the proposed maleimide–thiol coupling approach.

By reacting Cys-GB1-Cys, which was pretreated with $\beta$-mercaptoethanol at 4°C overnight to prevent the formation of intermolecular disulfide bonds, with BM(PEO)$_3$ in a molar ratio of 1:1 in PBS buffer at room temperature for 2 h, we found that polyprotein (GB1)$_n$ readily formed. From the standard reducing SDS-PAGE picture (Figure 2A), it is evident that the vast majority of monomeric GB1 has been converted to a multimer of GB1, and the formation of multimeric GB1 (up to 7) can be readily identified. At higher molecular weight, the protein band becomes smeared, indicating the formation of a mixture of higher-order multimers of GB1. From the intensity of the bands, it is possible to estimate the abundance of different multimers formed during this polymerization reaction.

The reaction of Cys-GB1-Cys and BM(PEO)$_3$ is similar to 2 + 2 condensation polymerization. The molecular weight of the resultant polymer depends on the stoichiometry of the two functional groups.30 To obtain higher multimers of GB1, it is important to ensure that the two functional groups are present in a 1:1 molar ratio. However, because of the air oxidation of cysteine residues into intermolecular disulfide bonds, the molar ratio of GB1 versus BM(PEO)$_3$ may not be necessarily equal to that of sulfhydryl groups versus maleimide groups. Thus, fine-tuning the molar ratio of GB1 and BM(PEO)$_3$ is necessary. As shown in Figure 2B, because of the incomplete reduction of disulfide bonds, the apparent 1:1 molar ratio of GB1 versus BM(PEO)$_3$ led to a large fraction of monomeric GB1 remaining in the reaction mixture, as monitored by analytical size exclusion chromatography (Figure 2B, curve 1). By adjusting the apparent molar ratio of GB1/BM(PEO)$_3$ to 1:0.8, the fraction of remaining monomeric GB1 decreased dramatically after polymerization and the yield of multimeric GB1 increased significantly (Figure 2B, curve 2).

We also found that the reaction temperature is an important parameter affecting the formation of multimeric GB1. It can be seen that incubating the reaction mixture at 4°C overnight led to the formation of a very small amount of oligomeric GB1, and a large amount of GB1 remained monomeric (Figure 2B, curve 3). However, reaction at room temperature for 2 h led to the efficient polymerization of GB1.

Thus, reacting Cys-GB1-Cys with BM(PEO)$_3$, at room temperature provides an efficient alternative method of constructing polyproteins for single-molecule AFM experiments.

**Single-Molecule AFM Experiments Revealing That (GB1)$_n$ Shows the Same Mechanical Properties as (GB1)$_h$ Constructed on the Basis of the Recombinant DNA Approach.** The formation of polyprotein (GB1)$_n$ is further confirmed by single-molecule AFM experiments. Stretching polyprotein (GB1)$_n$ results in characteristic sawtooth-like force–extension curves, where each individual sawtooth peak corresponds to the mechanical unfolding of individual GB1 domains in the polyprotein chain (Figure 3A). The force–extension curve contained as many as...
17 unfolding events, suggesting that the polyprotein contained 17 or more GB1 domains (where the degree of polymerization is 17 or higher). Fitting force—extension curves using the wormlike chain (WLC) model of polymer elasticity revealed that the contour length increment $\Delta L_c$ of GB1 is $\sim 18$ nm, which is identical to that of recombinant-produced polyprotein (GB1)$_n$. Dotted lines are fits to the WLC model at a pulling speed of 400 nm/s. (B) Histogram of the contour length increment of (GB1)$_n$, showing 18.0 ± 0.7 nm (average ± standard deviation). (C) Histogram of the unfolding force showing an average force of 182 ± 42 pN ($n = 2628$). (D) Five consecutive force—extension curves of polyprotein (GB1)$_n$ from a repeated stretching and relaxation experiment that lasts 100 cycles.

**Figure 3.** Mechanical unfolding and refolding of polyprotein (GB1)$_n$ constructed using maleimide—thiol coupling chemistry has an identical mechanical signature to that of (GB1)$_8$ constructed using the recombinant DNA technique. (A) Typical force—extension curves of the unfolding of (GB1)$_n$ show a characteristic contour length increment $\Delta L_c$ of $\sim 18$ nm, which is identical to the result of recombinant-produced polyprotein (GB1)$_n$. Dotted lines are fits to the WLC model at a pulling speed of 400 nm/s. (B) Histogram of the contour length increment of (GB1)$_n$, showing 18.0 ± 0.7 nm (average ± standard deviation). (C) Histogram of the unfolding force showing an average force of 182 ± 42 pN ($n = 2628$). (D) Five consecutive force—extension curves of polyprotein (GB1)$_n$ from a repeated stretching and relaxation experiment that lasts 100 cycles.

and relaxed more than 100 times. These results clearly indicated that polyprotein (GB1)$_n$ constructed using maleimide—thiol coupling chemistry is the same as (GB1)$_8$, suggesting that this method is suitable for constructing polyproteins for single-molecule force spectroscopy studies.

**Construction of the Pure Form Polymetalloprotein Rubredoxin.** Many proteins require cofactors to carry out their biological functions. Metalloproteins are one representative class of such proteins. Interactions between proteins and their cofactors can facilitate the folding of proteins and increase their thermodynamic stability. However, the existence of holo and apo forms of proteins leads to a mixture of proteins in different conformations. In addition, nonendogenous ligands can compete with endogenous ones, creating additional complexity for the composition of the protein. For example, metalloprotein rubredoxin can accommodate different metal ions in its tetra-cysteine metal-binding center. When overexpressed in *E. coli*, two different forms, wt Fe-rubredoxin and Zn-substituted rubredoxin, coexpress naturally, leading to a mixture of the two different metal-containing rubredoxins. These issues make it difficult, if not impossible, to use recombinant DNA methodology to construct

**Figure 4.** (A) UV—vis spectra of cpRD-GB1 before (top panel) and after (bottom panel) reacting with BM(PEO)$_3$. For clarity, the two spectra are offset from each other. The two proteins show identical spectral features, indicating that maleimide groups do not react with Fe-bound cysteine residues in rubredoxin. (B) Mechanical unfolding of (cpRD-GB1)$_n$ under normal and reduced environments shows that the polyprotein is stable under reduced conditions. Typical force—extension curves of Fe(III)-cpRD-GB1 (curves 1 and 2) and Fe(II)-cpRD-GB1 (curves 3 and 4) are shown. They both show characteristic contour length increments of $\Delta L_c \approx 18$ nm, which is from the unfolding of fingerprint GB1, and $\Delta L_c \approx 13$ nm, which is from the rupture of the Fe–S$_4$ center in cpRD, and a much lower unfolding force is observed in Fe(II)-cpRD-GB1 (in green). These results also reveal the robustness of the thioether linkage in the polyprotein under reducing conditions.
a polyprotein in which all individual domains are in the same form for single-molecule AFM studies. Instead, the maleimide—thiol coupling method provides the possibility to overcome this technical hurdle. Figure 4 shows one example of using the maleimide—thiol coupling method to construct a polyprotein of metalloprotein rubredoxin from *Clostridium pasteurianum* (cpRD) for single-molecule AFM experiments.

Rubredoxin is the simplest iron–sulfur protein containing one Fe(III) bound by four cysteines and has been used in our laboratories as a model system for investigating the mechanical unfolding/folding dynamics of metalloproteins. To construct pure Fe(III)-cpRD for single-molecule AFM experiments, we first constructed Cys-cpRD-GB1-Cys chimera as a building block for the construction of polyprotein (cpRD-GB1). We used an anion-exchange chromatography technique to separate the Zn(II) and Fe(III) forms of cpRD and obtained Fe(III)-cpRD-GB1. Then, purified Cys-cpRD-GB1-Cys was used to react with BM(PEO) at room temperature for 2 h to produce polyprotein (cpRD-GB1) that contains Fe(III)-cpRD only. The formation of the polyprotein was verified by SDS-PAGE (Figure 2A). In addition, the maleimide—thiol coupling reaction did not alter the characteristic UV–vis spectrum of Fe(III)-cpRD (Figure 4A), which originates from the absorbance by the Fe–S₄ center, suggesting that the polymerization process does not affect the Fe–S₄ center, and BM(PEO) reacted specifically with the cysteine residues at the N and C termini and the Fe(III)-coordinating cysteines did not react with maleimide groups.

Thus, we can now use single-molecule AFM to investigate the mechanical unfolding process of Fe(III)-cpRD in a well-controlled manner using polyprotein (cpRD-GB1) where GB1 domains serve as single molecular fingerprints for identifying single-molecule stretching events as well as an internal force caliber. Stretching polyprotein (Fe(III)-cpRD-GB1) results in characteristic sawtooth-like force–extension curves, where the sawtooth peaks correspond to the mechanical unfolding of GB1 and Fe(III)-cpRD domains (Figure 4C, curves 1 and 2). Unfolding events of GB1 are characterized by Δ₁ ≈ 18 nm (in black), and unfolding events of Δ₂ ≈ 13 nm are due to the mechanical unfolding of Fe(III)-cpRD domains (in red), providing the characteristic signatures of the unfolding of cp-rubredoxin.

**CONCLUSIONS**

We reported a novel, facile chemical coupling strategy to prepare polyproteins for single-molecule force spectroscopy studies. This maleimide—thiol coupling-based strategy offers a new alternative method that not only complements the existing strategies but also offers some unique advantages that will facilitate some special single-molecule AFM experiments that are otherwise difficult to study using current polyprotein construction strategies. We anticipate that the continuous development of a polyprotein engineering approach will greatly facilitate protein mechanics studies and help to develop the single-molecule AFM technique into a general biophysical tool for protein characterization.

**AUTHOR INFORMATION**

Corresponding Author
*E-mail: hongbin@chem.ubc.ca.*

Present Addresses
1Department of Physics, Nanjing University, Nanjing, Jiangsu Province, PR China.

**ACKNOWLEDGMENT**

This work is supported by the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chairs program, and the Canada Foundation for Innovation.