Semisynthetic protein nanoreactor for single-molecule chemistry

Joongoo Lee and Hagan Bayley†

Department of Chemistry, University of Oxford, Oxford OX1 3TA, United Kingdom

Edited by Cynthia J. Burrows, University of Utah, Salt Lake City, UT, and approved September 16, 2015 (received for review June 2, 2015)

The semisynthesis of single-molecule chemistry has attracted interest because it provides insights into fundamental chemistry. The staphylococcal α-hemolysin (αHL) pore has been engineered for a variety of applications in biotechnology. For example, the pore can function as a detector for stochastic sensing (1–3) and DNA sequencing (4, 5), and act as a “nanoreactor” for single-molecule covalent chemistry (6). In the nanoreactor approach, individual bond-making and bond-breaking steps can be observed with submillisecond time resolution by monitoring the ionic current carried by the pore under an applied potential. Single-molecule chemistry has attracted interest because it provides insights into fundamental chemical processes that cannot be observed at the ensemble scale (6–8). However, chemistry carried out within the αHL nanoreactor has so far been confined to the reactions of thiolates (i.e., deprotonated Cys side chains) (7–12) and derivatives of the side chains of Cys residues (13, 14). The incorporation of unnatural amino acids into the αHL pore would substantially advance single-molecule chemistry, through the incorporation of a large variety of individual side chains as well as multiple substitutions. We have produced αHL polypeptides with unnatural alkyld aminocarboxylic acid chains (15) in vitro by using chemically aminoacylated tRNAs. However, the method is demanding and often gives poor results, e.g., when more than one amino acid is introduced. In vivo methods that use engineered synthetase/tRNA pairs also enable the incorporation of unnatural amino acids (16, 17). Two different amino acids (17, 18), but so far not more, can be incorporated. However, not all amino acids are successfully handled, and a new one often requires additional synthetase development. Further, various amino acid side chains are destroyed in vivo, often by reduction (18).

Another means to incorporate unnatural amino acids is through native chemical ligation (NCL) (19–21), wherein the amino acid of interest is introduced in a peptide obtained by chemical synthesis (22, 23). The advantage of NCL is that virtually any amino acid or collection of amino acids can be incorporated into the peptide chain, provided the side chain is stable toward the conditions of solid-phase peptide synthesis (SPPS) and NCL, or can be protected and deprotected. An extension of NCL that has seen widespread success is expressed protein ligation (24–26), a method for semisynthesis in which large segments of the polypeptide chain are bacterially expressed. However, relatively little work has been reported on the semisynthesis of membrane proteins (27–31). This deficiency can be attributed to the hydrophobic nature of membrane proteins, which are poorly soluble in either folded or unfolded form (32). For example, a polypeptide segment encompassing the transmembrane region of the αHL pore precipitates in neutral buffer. The use of a denaturing environment during NCL can solve the solubility problem (20). However, it raises the question of whether the semisynthetic protein can be refolded into the native form. Here, we incorporate an unnatural amino acid with a terminal alkyne side chain into the αHL polypeptide and demonstrate the functional activity of heptameric pores made from it. Further, we investigate click chemistry of the alkyne at the single-molecule level by using the semisynthetic pore as a nanoreactor.

Results

Preparation of Polypeptide Fragments. We designed polypeptide fragments that would provide for the efficient semisynthesis of a full-length αHL monomer by a three-way ligation strategy. In our design (Fig. 1A), a central segment [SP (synthetic peptide): SP[Thz-Met-127-Gly-128-N-acyl-benzimidazolone(Nbz)] was obtained by chemical synthesis. An N-terminal fragment (NTF: NTF[Ala-1-Met-13]-thioester) and a C-terminal fragment (CTF: CTF[Cys<sup>127</sup>-Asn<sup>205</sup>-D<sub>2</sub>H<sub>2</sub>]) were obtained after expression in <i>Escherichia coli</i>. Recombinant NTF[Ala-1-Met-13]-thioester was produced by thiolysis of a fusion protein, NTF-GyrA intein-CBD (NTF[Ala-1-Met-13]-<i>Mycobacterium xenopi</i> DNA gyrase A (Mxe GyrA) intein-chitin binding domain (CBD)) (SI Appendix, Fig. S1) derived from the plasmid pTXB3. The DNA encoding residues 1–113 of αHL were PCR-amplified from a plasmid bearing a full-length αHL gene and cloned into pTXB3. After expression at 37 °C, high levels of the fusion protein were found within inclusion bodies in the <i>E. coli</i> host. Following cell lysis and centrifugation, most of the pellet was dissolved in 8 M urea, 100 mM sodium phosphate, pH 7, for purification.

Significance

The modulation of ionic current flowing through an individual protein pore provides information at the single-molecule level about chemical reactions occurring within the pore. However, chemistry investigated in this way has been largely confined to the reactions of thiolates, presented by the side chains of cysteine residues. The introduction of unnatural amino acids would provide a large variety of reactive side chains with which additional single-molecule chemistry could be investigated. Here we have produced semisynthetic protein pores containing a terminal alkyne and used the pores as a nanoreactor to study Cu(I)-catalyzed azide-alkyne cycloaddition. A long-lived intermediate (4.5 s) in the reaction was directly observed.

Author contributions: J.L. and H.B. designed research; J.L. performed research; J.L. analyzed data; and J.L. and H.B. wrote the paper.

The authors declare no conflict of interest.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1510565112/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1510565112

PNAS Early Edition | 1 of 6
Lee and Bayley

Full-Length αHL Monomer Is Obtained by Native Chemical Ligation. We then proceeded to assemble a full-length αHL polypeptide bearing the propargyl group (Fig. L4) with two sequential NCL reactions each producing a native peptide bond: (i) ligation of the synthetic peptide to CTF (SI Appendix, Figs. S8 and S9), which yielded SP-CTF in 63% yield, and (ii) ligation of NTF to the first ligation product, which yielded the full-length synthetic monomer (SM) of αHL in 16% yield (SI Appendix, Fig. S10). Therefore, the overall yield of SM was 10%. The potential deletion product NTF-CTF was absent (Fig. 1B and SI Appendix, Fig. S10). We purified SM by gel filtration under denaturing conditions (8 M urea) followed by ion-exchange chromatography (SI Appendix, Fig. S10 C and D), and characterized it by LC-MS, which yielded a single peak containing material of the theoretical mass (Fig. 1B). The purified SM was concentrated under denaturing conditions (8 M urea) to prevent premature assembly into a heptameric pore.

Semisynthetic αHL Monomers Correctly Refold to Heptameric Pores. We next tested whether SM could be correctly folded into active monomers. WT αHL monomers prepared by in vitro transcription and translation (IVTT) and denatured in 8 M urea completely regained hemolytic activity after the urea concentration was reduced to 1 M (36). The urea concentration of the SM solution was reduced to ∼60 mM by repeated dilution and concentration in a centrifugal filter (Amicon, MWCO 3k). The SM was then examined for (i) hemolytic activity toward rabbit red blood cells (rRBCs); (ii) heptamer formation on liposomes and red blood cell membranes; and (iii) conformational integrity as reflected by limited proteolysis. The specific hemolytic activity of SM [HC30 = 39 ng mL−1 (n = 3); HC50, the concentration of protein giving 50% lysis at 100 min, Fig. 2A] was very similar to (SI Appendix, Fig. S11) that of recombinant WT αHL monomer [HC30 = 35 ± 12 ng mL−1 (n = 3), previously measured (37) as 31 ng mL−1]. To visualize the formation of αHL heptamers, we incubated SM for 1 h at 37 °C in the presence of liposomes (1 mg mL−1, diphytanoyl phosphatidylcholine, DPhPC) (38) and upon SDS-polyacrylamide gel electrophoresis, we observed a new band that migrated with a mobility corresponding to the heptameric state (Fig. 1C). In addition, we incubated SM in different ratios with radiolabeled WT αHL could be solubilized in denaturing buffer containing 8 M urea. To allow binding to chitin beads, the concentration of the urea was reduced to 1 M. Unbound proteins were washed off the column and the bound fusion protein was treated with 2-mercaptopetethane sulfonate under gravity flow, yielding the NTF[Ala1-Met113]4-thioester in high purity as determined by SDS/PAGE (SI Appendix, Fig. S1B) and LC-MS (liquid chromatography–mass spectrometry) (SI Appendix, Fig. S2).

We prepared another plasmid (pT7-SC1-CDH, SI Appendix, Fig. S3) for the production of CTF[Cys127-Asn128]-D3H4, under control of the T7 promoter. In this plasmid, the codons for residues 1–126 in the pT7-SC1 plasmid, which contains a full-length αHL gene, are replaced with codons for Met-Cys. CTF was again obtained from inclusion bodies, and then purified in 6 M guanidine hydrochloride by use of the His tag at the C terminus. The N-terminal Met was found to be clefted and the unmasked Cys residue had undergone condensation with pyruvic acid, an abundant metabolite (33). We regenerated free Cys at the N terminus by cleaving the thiazolidine (Thz) with methoxylamine (MeONH2) at pH 4 (SI Appendix, Fig. S4B). The purified CTF (SI Appendix, Fig. S4C) was characterized by LC-MS (SI Appendix, Fig. S5).

We then proceeded to assemble a full-length αHL polypeptide (Fig. 1A) and the synthetic monomer (SM) containing an alkyn group was prepared by NCL, purified, and characterized by LC-MS: [M+H]+ = 34,994 (obs), 34,994 Da (calcd). (C) SDS/PAGE analysis of WT and synthetic αHL. Lane 1: molecular markers; lane 2: radiolabeled αHL monomer produced by IVTT; lane 3: radiolabeled WT; pores assembled in the presence of DPhPC liposomes (1 mg mL−1); lane 4: purified synthetic αHL, monomer (SM); lane 5: purified SM after treatment with DPhPC liposomes. SM, pores are formed, which comigrate with the WT, pore. An autoradiogram is superimposed on the Coomassie Blue stained gel.

The central synthetic peptide (SP: Thz116Thr117LeuPrp118TyrGly119-Prp) with a side chain containing a functional group (Fig. 1A) was incorporated into the peptide. Two sequential NCLs splice the three segments into a full-length αHL polypeptide. (b) The synthetic αHL monomer (SM) containing an alkyn group was prepared by NCL, purified, and characterized by LC-MS: [M+H]+ = 34,994 (obs), 34,994 Da (calcd). The central peptide, SP, was produced by SPPS. An unnatural amino acid residue (in this case 1-proparglyglycine, Prp) with a side chain containing a functional group of interest (yellow ball) was incorporated into the peptide. Two sequential NCLs splice the three segments into a full-length αHL polypeptide. (b) The synthetic αHL monomer (SM) containing an alkyn group was prepared by NCL, purified, and characterized by LC-MS: [M+H]+ = 34,994 (obs), 34,994 Da (calcd). The central peptide, SP, was produced by SPPS. An unnatural amino acid residue (in this case 1-proparglyglycine, Prp) with a side chain containing a functional group of interest (yellow ball) was incorporated into the peptide. Two sequential NCLs splice the three segments into a full-length αHL polypeptide.
produced by IVTT in the presence of rabbit red blood cell membranes (rRBCm). SM oligomerized to form heteroheptamers with different stoichiometries (WT7αSM1, n = 0–7) (Fig. 3A). Together, these data show that SM forms functional heptameric pores. A limited proteolysis experiment was carried out with proteinase K, to determine the conformational state of the αHL polypeptide in the monomer and the heptamer (39–41). We found that the monomer was cleaved into two fragments of 20 and 15 kDa, indicating that it is correctly folded (Fig. 2B, lane 4). When the heptamer formed on liposomes was similarly treated, a large fraction of the polypeptide chains was protected from the protease (Fig. 2B, lane 5), indicating that the heptamer was correctly assembled (SI Appendix, Fig. S12).

The semisynthetic pore reveals an intermediate in the copper(I)-catalyzed azide-alkyne cycloaddition reaction. We then used the protein nanoreactor approach to investigate click chemistry of the αHL subunit neither affects the electrical properties of the αHL pore nor its ability to bind the βCD adaptor. The copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC), a click reaction (44), has been widely studied in materials science (45), bioconjugation (46), and drug discovery (47). We first examined the unfunctionalized WT7 pore (SI Appendix, Fig. S15) in the presence of the reagents used for the click reaction, namely the Cu(I) catalyst and azide substrates (SI Appendix, Fig. S14, APB400: 0.4 kDa azide-PEG-biotin; APB700: 0.7 kDa azide-PEG-biotin). Cu(I) (20 mM) and APB400 (0.5–2.0 mM), neither separately nor together, affected the WT7 pore (SI Appendix, Fig. S15A–G), whereas APB700, either by itself (SI Appendix, Fig. S15 H–K) or with Cu(I) (SI Appendix, Fig. S15J), produced reversible events, which increased in frequency with APB700 concentration. These events most likely arise from entry of APB700 into the lumen of the pore, without covalent attachment, for periods sufficient to be detected (48). By examining the concentration dependence of the blockades (APB700, 0.5 to 2 mM) in the absence of Cu(I), we obtained $k_{on} = 1.3 \pm 0.2 \times 10^5 \text{ M}^{-1} \text{s}^{-1}$, $k_{off} = 8.7 \pm 0.3 \times 10^3 \text{ s}^{-1}$; $K_D = k_{off}/k_{on} = 6.7 \pm 1.0 \times 10^{-2} \text{ M}$ (SI Appendix, Fig. S17), similar to the values obtained by Movileanu et al. for a 940-Da PEG (48). The mean lifetime of the blockades was short (~100 μs), resulting in a broad distribution of amplitudes under our data acquisition conditions. No permanent current decrease was observed with the WT7 pore during 3 h of monitoring after the addition of Cu(I), with either of the azides present in both compartments.

We then carried out chemistry with single WT7αSM1 pores. As in the case of the WT7 pore, Cu(I) added to both the cis and trans compartments at the same time did not generate significant changes in the current (SI Appendix, Fig. S16 A and B), suggesting that if a Cu acetylide (Fig. 4 A, I) is formed it is short-lived. When APB700 was added to the trans compartment in the absence of Cu(I), short reversible current blockades characteristic of entry and exit from the pore were seen, which were similar to those observed with the WT7 pore: $k_{on} = 1.6 \pm 0.2 \times 10^5 \text{ M}^{-1} \text{s}^{-1}$; $k_{off} = 5.6 \pm 1.1 \times 10^3 \text{ s}^{-1}$; $K_D = k_{off}/k_{on} = 3.5 \pm 0.8 \times 10^{-2} \text{ M}$ (SI Appendix, Fig. S16 C–I). In the case of APB400, the events were presumably too short to be registered at the data acquisition rate used (48). By contrast with WT7, the addition of Cu(I) to both compartments in the presence of 2 mM APB400 or APB700 (trans) eventually led to an irreversible partial blockade of the WT7αSM1 pore, suggesting that Cu(I) triggered a reaction between the substrate and the alkynyl within the pore.
Interestingly, the blockade of the pore proceeded in two steps (Fig. 4 B–D and SI Appendix, Table S1). An intermediate state ($S_i$) was formed first, characterized by a small current decrease ($\delta I_{b1} = I_{S_i} - I_{S_0}$, where $S_0$ is the unreacted state of the pore), which was subsequently converted to a second state ($S_n$) associated with a larger current decrease ($\delta I_{b2} = I_{S_n} - I_{S_i}$). The percent decreases in the open pore current for APB400 (SI Appendix, Fig. S18A) were $\%\delta I_{b1, \text{APB400}} = 6.5 \pm 3.4$ and $\%\delta I_{b2, \text{APB400}} = 18.9 \pm 3.2\%$ ($n = 11$, where $n$ represents the number of experiments). For the larger APB700, the percent decreases (SI Appendix, Fig. S18B) were $\%\delta I_{b1, \text{APB700}} = 9.5 \pm 3.6$ and $\%\delta I_{b2, \text{APB700}} = 31.1 \pm 3.7\%$ ($n = 10$). The first step (to form $S_i$) was irreversible, and the final step yielded a permanent partly blocked state, $S_n$. The $S_i$ state is relatively noisy (Fig. 4 C and D), which suggests that the intermediate undergoes a reversible transformation, such as a conformational change or protonation and deprotonation [for APB400: $I_{\text{max}}(S_0) = 12.7 \pm 1.6\%$, $I_{\text{max}}(S_i) = 16.3 \pm 2.9\%$, $I_{\text{min}}(S_i) = 14.1 \pm 2.3\%$ ($n = 11$); for APB700: $I_{\text{max}}(S_0) = 17.4 \pm 3.2\%$, $I_{\text{min}}(S_0) = 22.1 \pm 2.8\%$, $I_{\text{min}}(S_i) = 14.7 \pm 3.2\%$ ($n = 10$), filtering: 5 kHz, sampling: 25 kHz, SI Appendix, Table S1]. The mean lifetime of $S_i$ (SI Appendix, Figs. S19 and S20) was $\tau_i = 4.5 \pm 0.6$ s ($k_i = 0.22 \pm 0.03$ s$^{-1}$) [$n = 21$, for APB400 ($n = 11$); for APB700 ($n = 10$)] and the mean overall reaction rate constant was $0.24 \pm 0.01$ M$^{-1}$ s$^{-1}$ (SI Appendix, Fig. S21). It is possible that the intermediate is one of the three sequential structures (II–IV, Fig. 4A) proposed by Worrell et al. (49), and based on its lifetime it is most likely the Cu triazoleide (IV). However, the mechanism of the CuAAC reaction remains contentious, and our assignment is strictly hypothetical. More important is the fact that we observe a long-lived intermediate that must be incorporated into any viable reaction mechanism.

Discussion

Previous work has demonstrated the ability of single αHL pores to display the time course of bond-making and bond-breaking reactions in individual molecules. Here, we have incorporated an unnatural amino acid bearing an alkyne group into the αHL pore to expand the range of covalent chemistry that can be examined by the nanoreactor approach.

By native chemical ligation, we synthesized a full-length αHL monomer including a central synthetic peptide incorporating Prp. The alkyne side chain of the amino acid was predicted to project into the lumen of the transmembrane β-barrel in the heptameric αHL pore. Our biochemical assays and single-channel current recordings demonstrated that the semisynthetic αHL monomer can be renatured efficiently and included in correctly assembled αHL heptamers. We used pores containing a single alkyne subunit to monitor CuAAC, a click reaction, at the single-molecule level and observed a long-lived (4.5 ± 0.6 s, $n = 21$) intermediate. We also obtained the overall rate constant ($k = 0.24 \pm 0.01$ M$^{-1}$ s$^{-1}$) of the click reaction within the pore, which is surprisingly low (50). To compare the rate constant with the reaction in bulk solution, we carried out the CuAAC reaction (SI Appendix, Fig. S23A) with a water-soluble synthetic tripeptide containing a terminal alkyne (ArgPrpArg, RXR, SI Appendix, Fig. S224) and Fmoc-L-β-azidalanine (Aza) (SI Appendix, Fig. S22B). The reaction was monitored by measuring the absorbance of the product (SI Appendix, Fig. S23B) at 260 nm by HPLC. The rate constant in bulk solution was $2.1 \pm 0.4$ M$^{-1}$ s$^{-1}$ (SI Appendix, Fig. S24), compared with earlier values for similar reactions in bulk solution, which are in the range of 10–200 M$^{-1}$ s$^{-1}$ (50). The ability of copper ions to form a complex with Bis-Tris propane (51) in the reaction buffer might contribute to the slower reaction rate. In previous work, we have found that reactions of small molecules inside the αHL pore exhibit rates within a factor of 10 of the rates in solution (6). In the present case, the low rate constant must reflect the ability of the bulky reagents to partition into the pore (52) and their ability to become correctly oriented close to the internal wall.

We suggest that the nanoreactor system, which we have extended significantly here, will be applicable to many reactions that occur on a millisecond time scale in aqueous conditions over
buffer and S
\[ S = \text{trans buffer (adjusted to pH 4.0)} \]

The buffer was 1 M KCl, 20 mM Bis-Tris propane, 25 M α to S
The authors thank H. Choi for assistance with the state. The triazolide (Click reaction in a WT

The N-terminal Thz group was subsequently cleaved by treatment with 0.2 M carboxyethyl)phosphine (TCEP), and 20 mM MPAA]. After overnight reaction (pH 6.8) containing 6 M guanidine hydrochloride (Gu.HCl), 40 mM Tris(2-

The mixture was desalted on a P6 column (Bio-Rad) and analyzed by LC-MS. The purified product was desalted with a P6 column and chromatography. The purified product was desalted with a P6 column and characterized by SDS/PAGE and LC-MS: ([M+H] \[ = 34,994 \] (obs), 34,994 (calcd)). For full details, see SI Appendix.

**Materials and Methods**

**Native Chemical Ligation.** The two sequential native chemical ligation reactions were carried out under previously established conditions (56). CTF (0.5 mM), from which pyruvate had been removed, was mixed with SP-Nbz (5 mM) under Ar in 0.5 mL of degassed NCL buffer [200 mM NaH

The mixture was desalted on a P6 column (Bio-Rad) and analyzed by LC-MS. The N-terminal Thz group was subsequently cleaved by treatment with 0.2 M MeONH_2.HCl in 200 mM NaH_2 PO_4 buffer (adjusted to pH 4.0) containing 6 M Gu.HCl, 40 mM TCEP for 4 h at room temperature. For the next round of ligation, the buffer was replaced with NCL buffer containing 200 mM MPAA by repeated dilution and concentration with a centrifugal filter (Amicon, MWCO 3k). NTF–thioester (0.1 mM) was mixed with the first ligation product (0.05 mM) and the reaction was allowed to proceed overnight. The final ligation product was purified by gel filtration followed by ion-exchange chromatography. The purified product was desalted with a P6 column and characterized by SDS/PAGE and LC-MS: ([M+H]^+ = 34,994 (obs), 34,994 (calcd)). For full details, see SI Appendix.

**Ionic Current Recordings and Single-Molecule Chemistry.** Ionic current measurements were performed with a planar bilayer apparatus at room temperature. Data were analyzed with pClamp (Version 10.1, Molecular Devices). OriginPro 8.5 was used for plotting IV curves and statistical analyses. For single-molecule click chemistry, substrate APB400 or APB700 [100 mM in 1 M KCl, 20 mM Bis-Tris propane (pH 8.5), 25 μM TCEP] was added in 2.5-μL portions to the trans compartment to give final concentrations of 0.5 to 2.0 μM. CuBr_2 was dissolved in the same buffer and mixed with 2 equivalents of Na ascorbate to reduce Cu(I) to Cu(II) before addition to both the cis and trans compartments at 20 mM final concentration. For full details, see SI Appendix.

**ACKNOWLEDGMENTS.** The authors thank H. Choi for assistance with the molecular dynamic simulations and Dr. S. Cheley, Dr. O. Daltrop, Dr. M. Boudreau, and Dr. A. Boersma for earlier work on this project and helpful discussions. This work was supported by a European Research Council Advanced Grant. J.L. was supported in part by a Korean Government Scholarship.