ELSEVIER

Biophysical Chemistry 96 (2002) 163–171

Biophysical Chemistry

www.elsevier.com/locate/bpc

Intein-mediated cyclization of a soluble and a membrane protein in vivo: function and stability☆

Christian Siebold, Bernhard Erni*

Departement für Chemie und Biochemie, Universität Bern, Freiestrasse 3, CH-3012, Bern, Switzerland

Received 8 February 2001; received in revised form 25 September 2001; accepted 28 September 2001

Abstract

Cyclized subunits of the *E. coli* glucose transporter were produced in vivo by intein mediated *trans*-splicing. IIA^{Glc} is a β-sandwich protein, $IICB^{Glc}$ spans the membrane eight times. Genes encoding the circularly permuted precursors $U_{C\Delta}$-IIA^{Glc}-$U_{N\Delta}$ and $U_{C\Delta}$-$IICB^{Glc}$-$U_{N\Delta}$ were assembled from DNA fragments encoding the 3′ and 5′ segments of the *recA* intein of *M. tuberculosis* and *crr* and *ptsG* of *E. coli*, respectively. A 20-residues long, Ala-Pro rich linker peptide and/or a histidine tag were used to join the native N- and C-termini in the cyclized proteins. The cyclized proteins complemented growth of glucose auxotrophic strains. Purified, cyclized IIA^{Glc} and $IICB^{Glc}$ had 100 and 25%, respectively, of wild-type glucose phosphotransferase activity. They had an increased electrophoretic mobility, which decreased upon linearization of the proteins with chymotrypsin. Cyclized IIA^{Glc} displayed increased stability against temperature and GuHCl-induced unfolding (75 vs. 70 °C; 1.52 vs. 1.05 M).

Keywords: Circular protein; Glucose transporter; Intein; Phosphotransferase system; Protein splicing; Protein stability

1. Introduction

Protein splicing is a postranslational process, by which an internal sequence (intein) of a precursor protein is removed by cleavage of two peptide bonds and the two flanking sequences (N-extein and C-extein) are ligated through a native peptide bond. The splice reaction involves four nucleophilic substitution reactions (N-S, S-S, N-N and S-N acyl shifts) and is catalyzed by the intein. The N- and C-exteins are the substrates of this reaction, but by themselves do not have splicing activity. Some inteins in addition have sequence-specific endonuclease activity, mediating the replicative transposition of the intein encoding DNA into an intein free site (intein homing). The protein splicing activity is confined to the N- and C-terminal flanking region of the intein, as inferred from sequence comparison of inteins with and without homing endonuclease activity. In addition to intramolecular splicing, intermolecular splicing between exteins on a 'split inteins' encoded by

☆ Dedicated to Rainer Jaenicke, the biophysicist of rigorous standards, the always generous colleague, the homo universalis with a fine sense for the arts and humanities.

*Corresponding author. Tel.: +41-31-631-4346; fax: +41-31-631-4887.

E-mail address: erni@ibc.unibe.ch (B. Erni).

0301-4622/02/$ - see front matter
PII: S0301-4622(02)00012-1

two separate genes has also been observed [1]. Approximately 50% of all mature spliced exteins are DNA processing enzymes (for comprehensive reviews see [2,3].

Gene reconstruction experiments revealed that a large portion of the central part of inteins can be deleted without compromising extein splicing, and that inteins can be split and expressed as two polypeptide chains, which upon non-covalent association regain splicing activity. This allows for an intein mediated *trans*-splicing of two independently translated polypeptides [1,4–9]. Three inteins have been used for *trans*-splicing in vivo and in vitro, the RecA intein of *Mycobacterium tuberculosis*, the *Pyrococcus* Psp Pol-I intein [6–8], and the DnaE split intein from *Synechocystis* sp. PCC6803 [1,10]. The maximum reported in vitro splicing efficiencies were 75% with Psp Pol-I and 60% with the RecA intein. The in vivo splicing efficiency of the Rec A intein was at most, 50% [8].

Because the central portion of the intein is not necessary and the flanking regions do not need to be linked covalently for splicing activity, inteins can be used to prepare cyclized proteins [10–12]. A circularly permuted linear precursor comprising, in this order, the C-terminal half of an intein, the target protein to be cyclized and the N-terminal half of the intein, can be converted into a cyclized product by intramolecular *trans*-splicing. Cyclization of a protein could, in principle, have three desirable effects: (i) increased stability, because of reduced conformational entropy of the unfolded state; (ii) increased rate of folding, due to a reduced number of possible folding pathways; and (iii) resistance to amino- and carboxypeptidases.

Here, we describe the use of the RecA intein to cyclize the subunits of the *E. coli* glucose transporter in vivo. This transporter acts by a mechanism that couples translocation to phosphorylation of the substrate (for a review see [13]). It consists of two subunits, IIA^{Glc} and $IICB^{Glc}$, which sequentially transfer phosphorylgroups from the phosphorylcarrier protein HPr to glucose (Fig. 1). His-90 of IIA^{Glc} and Cys-421 of $IICB^{Glc}$ are transiently phosphorylated in this process. IIA^{Glc} (18 kD) is a sandwich of two β-sheets, each consisting of six antiparallel β-strands and two

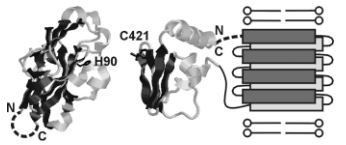

Fig. 1. Model of the IIA^{Glc} and $IICB^{Glc}$ subunits of the glucose transporter. $IICB^{Glc}$ consists of the cytoplasmic IIB^{Glc} domain (PDB code 1IBA) and the membrane spanning IIC^{Glc} domain. The active site Cys-421 of $IICB^{Glc}$ and the active site His-90 of IIA^{Glc} (PDB code 2F3G) are shown in *stick* representation. The sites of cyclization are shown as *broken lines* between the native N- and C-termini.

very short α-helices [14–16]. $IICB^{Glc}$ (53 kD) consists of two domains. The C domain spans the membrane eight times. The B domain consists of a four-stranded antiparallel β-sheet, covered on one face by three helices [17]. The N- and C-termini of $IICB^{Glc}$ are on the cytoplasmic face of the membrane and can be connected by a linker peptide, as demonstrated by functional expression of circularly permuted variants of $IICB^{Glc}$ [18,19].

2. Materials and methods

2.1. Plasmid construction

The full-length *recA* intein (codons 1–790) was PCR amplified from the *M. tuberculosis* genomic DNA (gift of Dr C. Aebi, University Hospital, Bern) and cloned into vector pBluescript. The expression vectors were constructed by sequential ligation of three PCR amplified fragments, with plasmid pMSEH2 encoding the Ptac promoter, T7 ribosome binding site, a polylinker, β-lactamase and the $LacI^{Q}$ repressor [20]. The DNA for $U_{C\Delta}$ (codon 591–692 [8]) was PCR amplified with primers P5(NdeI) and P3(SalI), digested with *Nde*I, and ligated with pMSEH2, opened with *NdeI* and *SmaI* to afford pCint. The DNA for $U_{N\Delta}$ (codon 251–344 [8]) was PCR amplified with primers P5(XbaI) and P3(HindIII), digested with *Hin*dIII, and ligated with pCint, opened with *Bam*HI (blunted) and *Hin*dIII to afford pCintN. The gene *ptsG* encoding $IICB^{Glc}$ with a C-terminal His-tag was PCR amplified with P5(SalI) and P3(XbaI) primers, and pTSGH11 [21] as template. Target DNA and vector pCintN, were digested

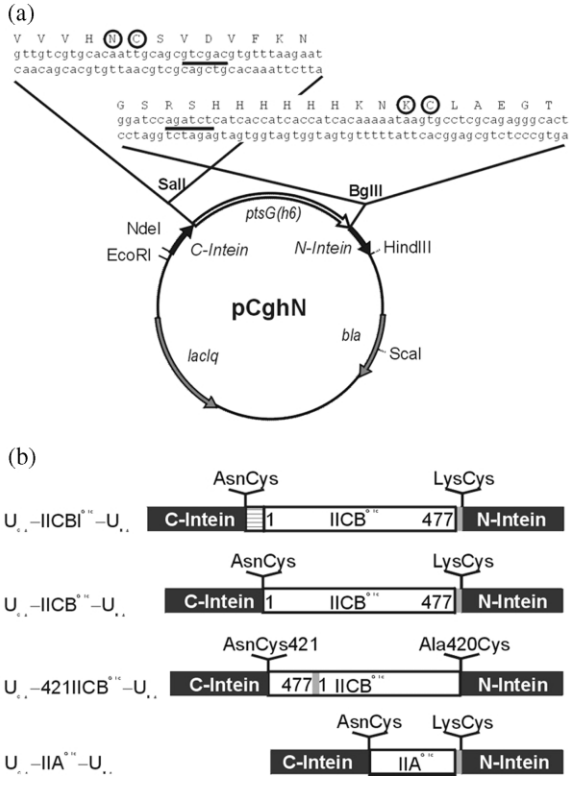

Fig. 2. Plasmid map, linker sequences and structure of precursor proteins. (a) Plasmid pCghN encodes the precursor proteins $U_{C\Delta}$-IICB^{Glc}-$U_{N\Delta}$ under the control of the Ptac promoter. The nucleotide and amino acid sequences of the $U_{C\Delta}$-IICB^{Glc} and IICB^{Glc}-$U_{N\Delta}$ splice sites are indicated. The residues from the native RecA intein–extein junction are *circled*. The restriction sites *Sal*I and *Bgl*II are *underlined*. (b) Structure of the expressed fusion proteins. Intein sequences, *filled*; Sequences of proteins to be cyclized, *open*; His-tag at the splice junction, *gray*; Ala-Pro linker, *striped*. Residue numbers refer to the IICB^{Glc} wild-type sequence (Swiss-Prot accession number P05053; IIA^{Glc} P08837).

with *Sal*I and *Xba*I, and then ligated to afford pCghN (Fig. 2a). The vectors encoding the precursors $U_{C\Delta}$-IICBl^{Glc}-$U_{N\Delta}$ and $U_{C\Delta}$-IIA^{Glc}-$U_{N\Delta}$ (Fig. 2b) were constructed by exchange of the *Sal*I/*Bgl*II fragment of pCghN with a PCR fragment, encoding the desired target sequence. The following plasmids served as templates for target DNA amplification with appropriate P5(SalI) and P3(BglII) primer pairs: pTSHIC9 for IIA^{Glc} [22], pTSGxG for IICBl^{Glc}. pTSGxG encodes two IICB^{Glc} fused in frame via an Ala-Pro linker (unpublished). The vector encoding the precursor $U_{C\Delta}$-421IICB^{Glc}-$U_{N\Delta}$ was constructed as follows: the *Sal*I site of pCintN was changed into a *Mun*I site to create the correct Asn/Cys421 splice junction. The circularly permuted *ptsG* gene (codons 421–420) was assembled from two PCR fragments. Fragment A (codons 421–477 plus His-tag) was amplified with primers P5(MunI) and P3(overhang), fragment B (codons 1–420) was PCR amplified with P5(overhang) and P3(XbaI). Fragments A and B were amplified by overlap extension PCR with primers P5(MunI) and P3(XbaI), and the full-length circularly permuted *ptsG* gene was ligated into the modified pCintN opened with *Mun*I and *Xba*I.

E. coli strain WA2127Δcrr was constructed by P1 transduction of the Δcrr/Kan^R cassette from *E. coli* TP2862 [23] into WA2127 (*manXYZ*) [24].

2.2. Expression and purification of cyclized proteins

E. coli ZSC112LΔG was transformed with plasmids, encoding the $U_{C\Delta}$-IICB^{Glc}-$U_{N\Delta}$ and $U_{C\Delta}$-IIA^{Glc}-$U_{N\Delta}$ precursor proteins. Cells were grown at 37 °C and induced with 100 μM IPTG, when the culture (3 l) had reached $OD_{600}=0.8$ and harvested by centrifugation after 5 h. The cell sediment was resuspended in lysis buffer (25 ml, 50 mM Tris–HCl, pH 8.0, 10 mM β-mercaptoethanol, 500 mM NaCl) broken by two passages in a French pressure cell and fractionated by differential centrifugation into cell debris plus inclusion bodies (10 min, $3000\times g$), membrane fraction (60 min, $150\,000\times g$), and cytoplasmic fraction (supernatant). Membranes containing cyclized IICB^{Glc} were resuspended in buffer (6 ml, 20 mM Tris–glycine, pH 9.1, 10 mM β-mercaptoethanol, 500 mM NaCl) and solubilized with *n*-dodecylmaltoside (60 mM final concentration). The extract was freed of insoluble material by centrifugation (60 min, $150\,000\times g$), mixed with 6 ml of Ni-NTA resin and allowed to adsorb for 30 min. The resin was transferred to a column and washed with 75 ml of two buffers (50 mM NaP_i, pH 8.0, 10 mM β-mercaptoethanol, 500 mM NaCl, 0.4 mM *n*-dodecylmaltoside) containing 0 and 40 mM imidazole. The protein was eluted

with 200 mM imidazole in the same buffer. Cyclized IIA^{Glc} was adsorbed to 6 ml Ni-NTA resin and processed as described as above, but in detergent free buffers. Cyclized IIA^{Glc} was eluted with 200 mM imidazole. The purified proteins were dialyzed against buffer (20 mM NaP_i, pH 8.0, 1 mM EDTA, 1 mM DTT, 150 mM NaCl) with and without 0.4 mM *n*-dodecylmaltoside, respectively, and stored at 4 °C. Insoluble, $U_{C\Delta}$-IIA^{Glc}-$U_{N\Delta}$ precursor was extracted from the cell debris with 4 M urea in buffer A (20 mM NaP_i, pH 8.0, 1 mM EDTA, 1 mM DTT, 150 mM NaCl), freed of insoluble material by centrifugation (10 min, $3000 \times g$) and dialyzed against several changes of the same buffer.

2.3. Assays for phosphotransferase activity and transport

To test for in vivo glucose transport, *E. coli* ZC112ΔG (*ptsG manZ glk*) [25] and WA2127Δcrr (*crr manXYZ*) were transformed to express wild-type and cyclized forms of $IICB^{Glc}$ and IIA^{Glc}, respectively, and streaked on McConkey agar base (Difco) supplemented with 100 μg/ml ampicillin and 0.4% glucose. Sugar–phosphotransferase activity of purified wild-type and cyclized proteins was measured as described [26], in the presence of saturating amounts of the *E. coli* proteins enzyme I, HPr, and either wild-type IIA^{Glc} or wild-type $IICB^{Glc}$.

2.4. Protein characterization

Precursor and processed proteins were identified by Western-blotting with anti-IIA^{Glc} serum, monoclonal anti-$IICB^{Glc}$ antibodies and goat anti-rabbit (rabbit anti-mouse) antibody–alkaline phosphatase conjugates [27]. Purified wild-type and cyclized proteins (5 μg/in 20 μl buffer A without EDTA) were treated with limiting concentrations of chymotrypsin for 30 min at 37 °C and then analyzed on a 20% polyacrylamide gel. The masses of the cyclized proteins were identified by electro-spray ionization mass spectrometry. The stabilities of wild-type and cyclized IIA^{Glc} were determined by guanidinium HCl- and temperature-induced equilibrium unfolding. Samples were dissolved at 2 mg/ml in 20 mM NaP_i pH 7.4, in high purity GuHCl (Sigma) at the indicated concentrations and incubated for 2 h at 22 °C. For temperature-induced unfolding and refolding, the temperature was changed at a rate of 1 °C/min. The progress curves were recorded by circular dichroism spectroscopy at 222 nm in a Jasco spectropolarimeter (J715-A) using a 0.5-mm light-path cell. Data were normalized to the differences between the linear changes observed below and above the transition region. Normalized data are reported as fraction of folded protein. The smoothed curves for GuHCl-induced unfolding were obtained by non-linear least square fit [28].

3. Results

3.1. In vivo cyclization and function

Two linear precursors, $U_{C\Delta}$-$IICB^{Glc}$-$U_{N\Delta}$ and $U_{C\Delta}$-IIA^{Glc}-$U_{N\Delta}$, were designed such that the N- and C-termini in the cyclized product were bridged with a hexahistidine tag (Figs. 1 and 2). In a third precursor ($U_{C\Delta}$-$IICBl^{Glc}$-$U_{N\Delta}$), the bridge was extended with a 20 residues long Ala-Pro rich linker peptide. The $U_{C\Delta}$-II splice junction consisted of the Asn/Cys pair, which is invariant in all inteins, the II-$U_{N\Delta}$ junction consisted of the Lys/Cys pair, which is found in the *M. tuberculosis* RecA intein. A fourth precursor of $U_{C\Delta}$-$IICB^{Glc}$-$U_{N\Delta}$ was designed differently. The native N- and C-termini of $IICB^{Glc}$ were fused and the splice junctions were introduced at the active site (Cys-421) in the IIB^{Glc}-domain. In this variant, the $U_{C\Delta}$-IIB^{Glc} splice junction consisted of Asn/Cys-421 and the II-$U_{N\Delta}$ junction of Ala-420/Cys. *E. coli* strain ZSC112ΔG (*ptsG*$^-$ *manZ*$^-$*crr*$^+$) was transformed with plasmids encoding the $IICB^{Glc}$ precursors and WA2127Δcrr (*crr*$^-$ *manXYZ*$^-$*ptsG*$^+$) with the plasmid for the IIA^{Glc} precursor. Transformants were plated on McConkey glucose indicator plates to assay for glucose fermentation, which is contingent on IIA^{Glc} and $IICB^{Glc}$ mediated uptake. Transformants expressing $IICB^{Glc}$ and IIA^{Glc} with splice junctions at the native N- and C-termini formed red colonies, suggesting that the precursors were either active by themselves or processed to a functional form.

Cells expressing the precursor with the splice junction in the active site Cys-421 of the IIB^{Glc} domain formed yellow colonies, indicating that it was not processed correctly.

3.2. Purification and characterization of cyclized proteins

The active IIA^{Glc} and $IICBl^{Glc}$ variants, with a linker peptide between the native C- and N-termini (Fig. 2b), were purified. Protein expression was induced with IPTG and cells were harvested after 5 h of induction. The cell lysate was separated into three fractions: inclusion bodies and cell debris, membranes, and cytoplasmic proteins. The proteins in each fraction were identified on immunblots, with monoclonal anti-$IICB^{Glc}$ and polyclonal anti-IIA^{Glc} antibodies. The $U_{C\Delta}$-IIA^{Glc}-$U_{N\Delta}$ precursor was strongly expressed, but by and large, was in inclusion bodies (Fig. 3a, lane 3). The cytoplasmic fraction contained processed IIA^{Glc}, which could be purified by metal chelate affinity chromatography and gel filtration (Fig. 3a, lanes 4 and 5). The $U_{C\Delta}$-$IICBl^{Glc}$-$U_{N\Delta}$ precursor was expressed in much smaller amounts, as expected of an inner membrane protein. Both precursor and processed forms were in the membrane fraction (Fig. 3b, lane 2). The processed form could be solubilized completely in *n*-dodecylmaltoside; the precursor, by and large, was insoluble in the non-ionic detergent (Fig. 3b, lanes 3–4). Processed $IICB^{Glc}$ and a small amount of precursor were purified by metal chelate affinity chromatography. Only a smear of antigenic material could be detected in extracts of the variant, with the splice junction in the active site Cys-421 (results not shown).

Processed IIA^{Glc} had a molecular mass of 20075.2 ± 5.2 g/mol, which agrees well with 20074.2 g/mol expected of cyclized IIA^{Glc} (Table 1). Processed $IICB^{Glc}$ had a molecular mass of 54007.0 ± 14.0 g/mol, which is 355 g/mol greater than the expected mass of 53652.0 g/mol. The reason for this difference is not known. The DNA sequence encoding the precursor was as expected.

The processed IIA^{Glc} and $IICB^{Glc}$ had increased electrophoretic mobilities relative to the linear wild-type forms, as expected of cyclized proteins unable to assume a fully unfolded linear conformations (Fig. 4). When the purified proteins were cleaved with a limiting concentration of chymotrypsin, the electrophoretic mobilities first decreased to the rate of the linear wild-type proteins and then increased again as proteolytic degradation progressed (Fig. 4). These results indicate that the IIA^{Glc} and $IICB^{Glc}$ precursors were proc-

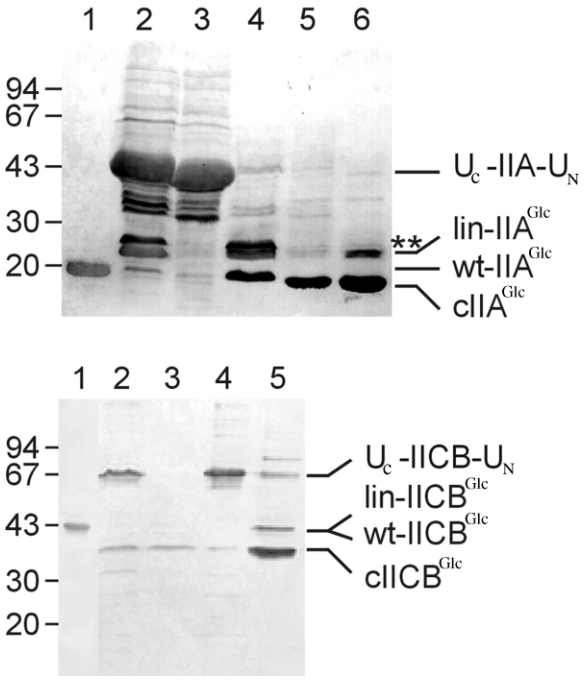

Fig. 3. Immunblot analysis of in vivo protein cyclization. (a) IIA^{Glc}: *Lane 1*, purified wild-type IIA^{Glc} (reference). *Lane 2*, cell lysate. *Lane 3*, inclusion bodies, containing the precursor and products of abortive processing or proteolysis. *Lane 4*, soluble, cytoplasmic fraction, containing cyclized IIA^{Glc}. *Lane 5*, cyclic IIA^{Glc}, after Ni^{2+} affinity chromatography. *Lane 6*, cyclic IIA^{Glc}, after gel filtration chromatography. A small amount of the cyclic protein is already linearized by unspecific endogenous protease activity. The immunblot is stained with anti-IIA^{Glc} antiserum. The proteins in *lanes 2 and 4* indicated with *stars* are cross-reacting species and unrelated to IIA^{Glc}. (b) $IICB^{Glc}$: *Lane 1*, purified wild-type $IICB^{Glc}$ (reference). *Lane 2*, membrane fraction, containing precursor and cyclized $IICB^{Glc}$. *Lane 3*, detergent extract, soluble fraction containing cyclized $IICB^{Glc}$. *Lane 4*, detergent insoluble material, containing precursor. *Lane 5*, cyclized $IICB^{Glc}$, after Ni^{2+} affinity chromatography. The fraction is contaminated by a small amount of linearized and precursor proteins. The immunblot is stained with a mixture of monoclonal IIB^{Glc}-specific antibodies.

Table 1
Properties of cyclized IIA^{Glc} and $IICB^{Glc}$

Protein	Yield (mg/l)	In vivo glucose uptake	In vitro specific activity (%)	Mr ESI-MS/ expected	N-terminus after chymotrypsin
IIA^{Glc} (w.t.)	20–25	+	100	n.d./18251	KSLV, SLVS, VSDD
$ccIIA^{Glc}$	2	+	100	20075 ± 5/ 20074	DKLK
$IICB^{Glc}$(w.t.)	2–3	+	100	n.d.	n.d.
$ccIICB^{Glc}$	0.15–0.2	+	25–30	n.d.	n.d.
$ccIICBl^{Glc}$	0.15–0.2	+	25–30	54007 ± 14/ 53652	n.d.

essed to cyclized forms by intein-mediated *trans*-splicing. Linearization of cyclized IIA^{Glc} by chymotrypsin occurred after Phe-4 (numbering of wild-type sequence) near the splice junction. The N-terminus of of the linear wild-type IIA^{Glc} was heterogeneous, as previously described [29], with starts at Lys-8, Ser-9 and Val-11 of the original sequence. Cyclized IIA^{Glc} could not be microsequenced, as expected if the N-terminus is blocked by cyclization. Similarly, cyclized $IICB^{Glc}$ could not be sequenced, whereas wild-type $IICB^{Glc}$ had a free N-terminus (results not shown).

The specific glucose phosphotransferase activities of wild-type and cyclized IIA^{Glc} were identical (Fig. 5a). The activity of purified cyclized $IICB^{Glc}$ was 25% of the wild-type control (Fig. 5b). This is slightly less than the activities of two circularly permuted forms of $IICB^{Glc}$, which had 40 and 70% of wild-type activity. These variants have their native N- and C-termini connected by an Ala-Pro rich linker, and have new N- and C-termini in the third cytoplasmic loop and in the linker between the IIC^{Glc} and IIB^{Glc} domain, respectively [18,19].

The stability of linear and cyclized IIA^{Glc} were compared by temperature-induced unfolding and refolding, and by GuHCl-induced unfolding (Fig. 6), which was monitored by circular dichroism spectroscopy at 222 nm. The midpoint of transition for GuHCl-induced unfolding was shifted from 1.1 to 1.5 M for cyclized IIA^{Glc} (Fig. 6a). The temperature at the midpoint of unfolding was 75 °C for cyclized IIA^{Glc} and 70 °C for linear IIA^{Glc} (Fig. 6 b,c). The midpoint temperatures of refolding were 71 °C for cyclized IIA^{Glc} and 66 °C for wild-type IIA^{Glc}. Phosphotransferase activity of both protein variants was recovered completely after cooling (results not shown), indicating that the proteins assumed the catalytically active fold and did not undergo chemical modifications during the heating. Taken together, these observations indicated that cyclization of IIA^{Glc} results in a stabilization of the protein.

4. Discussion

The N- and C-termini of proteins are often close together and surface exposed [30], and it has been demonstrated by circular permutation of amino acid sequences that termini can be joined by short loop regions, without changing the overall conformation and activity of the protein [18,31–33]. If circularly permuted proteins are active, cyclized proteins should be so, too. In addition, they are expected to have lower entropy in the unfolded state and should therefore be more stable. Lacking free termini, they are resistant to exopeptidases. Partial immobilization of unstructured ends by cross-linking could facilitate protein crystallization.

Here, we demonstrate that: a truncated RecA intein [8] is capable of in vivo-cyclizing the soluble and membrane bound subunits of the glucose transporter in *E. coli*, the cyclic subunits can be purified by metal chelate affinity chromatography and are catalytically active. Cyclization of soluble IIA^{Glc} results in a 5 °C stabilization against reversible temperature-induced unfolding.

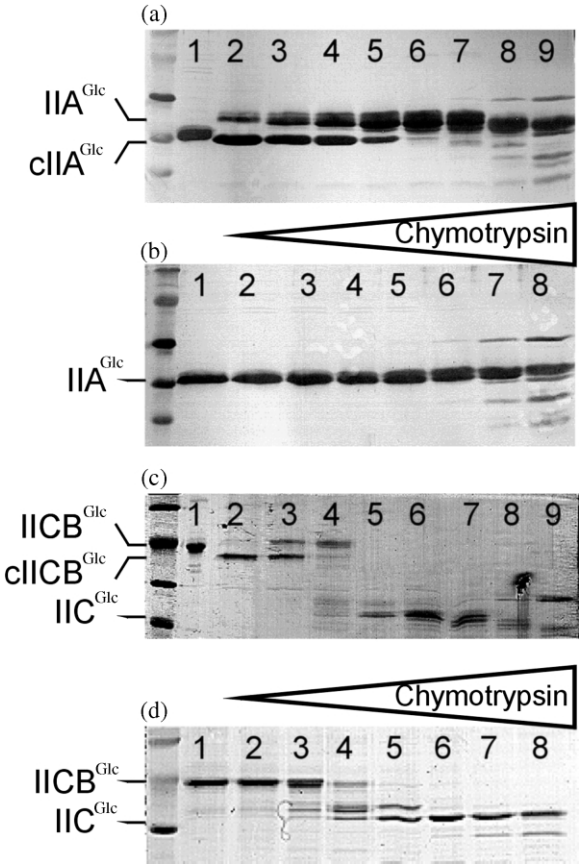

Fig. 4. Linearization of cyclized IIA^{Glc} and $IICB^{Glc}$ with chymotrypsin. (a) cyclized IIA^{Glc}; (b) wild-type IIA^{Glc} (control); (c) cyclized $IICB^{Glc}$; (d) wild-type $IICB^{Glc}$ (control). Note the decrease of electrophoretic mobility upon linearization of the cyclic forms. *Lanes 1*: purified wild-type IIA^{Glc} and $IICB^{Glc}$ (references). *Lanes 2*: no chymotrypsin. (a, b) *Lanes 3–8*: with chymotrypsin (2, 4, 8, 16, 32, 125, 500 μg/ml) (c, d) *Lanes 3–8*: with chymotrypsin (2, 20, 125 ng/ml, 1, 5, 27, 167 μg/ml).

This value is in line with the 5 °C stabilization against heat precipitation, which was measured with in vitro cyclized β-lactamase (Mw 30 kD) [11] and the increased thermostability observed with cyclized dihydrofolate reductase [10]. We do not know how this degree of stabilization compares with the effect of a more physiological modification of IIA^{Glc}, such as phosphorylation or complex formation with the phosphorylcarrier protein HPr. However, it is noteworthy that phosphorylation of IIA^{Chb}, which is structurally different but functionally analogous to IIA^{Glc}, decreased the Tm for thermal denaturation from 54 to 40 °C [34].

Of the two precursors of $IICB^{Glc}$, only the one with the splice sites at the native N- and C-termini was inserted into the membrane and cyclized, whereas the second, with the splice site in the invariant active site Cys, was rapidly degraded. Although inteins were found to interrupt conserved sequences in DNA polymerases [35], the interruption of a hairpin loop in a β-sheet is not tolerated, suggesting that there are limits to intein-mediated target cyclization.

The efficiency of in vivo cyclization is far below the values of 50 to 75% reported for in vitro splicing reactions. 2 mg/l of pure IIA^{Glc} are of the same order as the 5 mg/l reported for cyclized dihydrofolate reductase [10], but this is little in relation to the amount of precursor and partially processed products, which are present in the inclusion bodies. Attempts to cyclize IIA^{Glc} from the urea solubilized precursor, according to the methods described for in vitro splicing [36], were unsuccessful. The precursor precipitated upon removal of urea and not even partial splice reactions were observed. Judged by Western-blotting, only small amounts of partially processed intermediates (thioester, lariat) are formed. The splice reaction, once initiated, appears to proceed to completion.

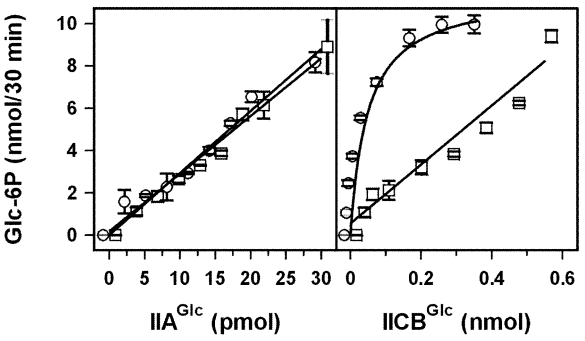

Fig. 5. Glucose phosphotransferase activity. (a) cyclized IIA^{Glc} (*squares*), wild-type IIA^{Glc} (*circles*). (b) cyclized $IICB^{Glc}$ (*squares*), wild-type $IICB^{Glc}$ (*circles*). The reaction volume was 0.1 ml per sample. Incubation was for 30 min at 37 °C. Shown, are means and S.D. of two independent assays.

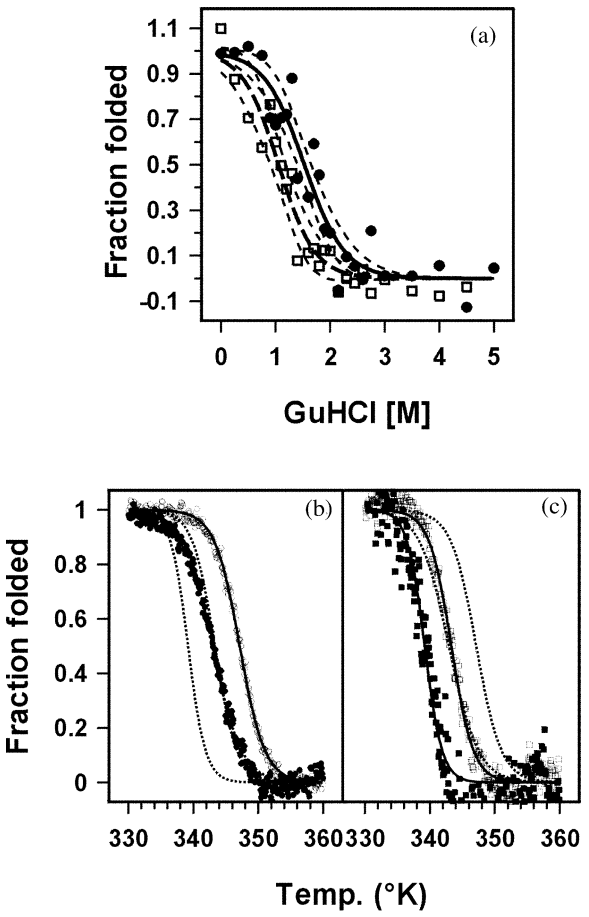

Fig. 6. Guanidinium HCl- (a) and temperature-induced (b, c) unfolding of cyclized and wild-type IIA^{Glc}. (a) cyclized IIA^{Glc} (*closed circles*); wild-type, linear IIA^{Glc} (*open squares*). The smoothed curves show the mean and the 95% confidence band (*thin broken lines*). (b) cyclized IIA^{Glc}. Unfolding transition (*open circles*), refolding transition (*closed circles*). (c) wild-type IIA^{Glc}. Unfolding transition (*open squares*), refolding transition (*closed squares*). To facilitate comparison, the smoothed curves drawn through the data points of c are also plotted in b and vice versa in c (*dotted lines*).

The objective of future experiments is to improve the efficiency of cyclization. Results reported for DnaE mediated in vivo cyclization [10] suggest that this naturally split intein is more suitable, because its two halves have an intrinsically strong affinity. The association between the two RecA intein fragments might be strengthened by fusion of leucine zippers to the N- and C-termini of the precursor.

Acknowledgments

We would like to thank J. Schaller and U. Kämpfer (Bern) for sequence- and mass spectroscopic protein analyses, Dr C. Aebi (Bern) for the gift of *M. tuberculosis* genomic DNA, and P.W. Postma (Amsterdam) for the IIA^{Glc} antiserum. This research was supported by the Swiss National Science Foundation (Grant 31-45838.95), ARPI-DA AG Münchenstein, and Ciba-Geigy Jubiläumsstiftung, Basel.

References

[1] H. Wu, Z. Hu, X.Q. Liu, Protein *trans*-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803, Proc. Natl. Acad. Sci. USA 95 (1998) 9226–9231.

[2] H. Paulus, Protein splicing and related forms of protein autoprocessing, Annu. Rev. Biochem. 69 (2000) 447–496.

[3] X.Q. Liu, Protein splicing intein: genetic mobility, origin and evolution, Annu. Rev. Genet. 34 (2000) 61–76.

[4] T.C.J. Evans, J. Benner, M.Q. Xu, The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins, J. Biol. Chem. 274 (1999) 18359–18363.

[5] B.M. Lew, K.V. Mills, H. Paulus, Characteristics of protein splicing in *trans* mediated by a semisynthetic split intein, Biopolymers 51 (1999) 355–362.

[6] K.V. Mills, B.M. Lew, S.Q. Jiang, H. Paulus, Protein splicing in *trans* by purified N- and C-terminal fragments of the *Mycobacterium tuberculosis* RecA intein, Proc. Natl. Acad. Sci. USA 95 (1998) 3543–3548.

[7] M.W. Southworth, E. Adam, D. Panne, R. Byer, R. Kautz, F.B. Perler, Control of protein splicing by intein fragment reassembl, EMBO J. 17 (1998) 918–926.

[8] K. Shingledecker, S.Q. Jiang, H. Paulus, Molecular dissection of the *Mycobacterium tuberculosis* RecA intein: design of a minimal intein and of a *trans*-splicing system involving two intein fragments, Gene 207 (1998) 187–195.

[9] H. Wu, M.Q. Xu, X.Q. Liu, Protein *trans*-splicing and functional mini-inteins of a cyanobacterial dnaB intein, Biochim. Biophys. Acta 1387 (1998) 422–432.

[10] C.P. Scott, E. Abel-Santos, M. Wall, D.C. Wahnon, S.J. Benkovic, Production of cyclic peptides and proteins in vivo, Proc. Natl. Acad. Sci. USA 96 (1999) 13638–13643.

[11] H. Iwai, A. Pluckthun, lactamase: stability enhancement

by cyclizing the backbone, FEBS Lett. 459 (1999) 166–172.
[12] T.C. Evans, D. Martin, R. Kolly, D. Panne, L. Sun, I. Ghosh, et al., Protein *trans*-splicing and cyclization by a naturally split intein from the dnaE gene of *Synechocystis* species PCC6803, J. Biol. Chem. 275 (2000) 9091–9094.
[13] G.T. Robillard, J. Broos, Structure/function studies on the bacterial carbohydrate transporters, enzymes II, of the phosphoenolpyruvate-dependent phosphotransferase system, Biochim. Biophys. Acta 1422 (1999) 73–104.
[14] J.H. Hurley, H.R. Faber, D. Worthylake, et al., Structure of the Regulatory Complex of *Escherichia coli* III(Glc) with Glycerol Kinase, Science 259 (1993) 673–677.
[15] D.I. Liao, G. Kapadia, P. Reddy, M.H.J. Saier, J. Reizer, O. Herzberg, Structure of the IIA domain of the glucose permease of *Bacillus subtilis* at 2.2-A resolution, Biochemistry 30 (1991) 9583–9594.
[16] D. Worthylake, N.D. Meadow, S. Roseman, D.I. Liao, O. Herzberg, S.J. Remington, Three-dimensional structure of the *Escherichia coli* phosphocarrier protein IIIglc, Proc. Natl. Acad. Sci. USA 88 (1991) 10382–10386.
[17] M. Eberstadt, S.G. Grdadolnik, G. Gemmecker, H. Kessler, A. Buhr, B. Erni, Solution structure of the IIBGlc domain of the glucose transporter of *Escherichia coli*, Biochemistry 35 (1996) 11286–11292.
[18] R. Beutler, F. Ruggiero, B. Erni, Folding and activity of circularly permuted forms of a polytopic membrane protein, Proc. Natl. Acad. Sci. USA 97 (2000) 1477–1482.
[19] R. Gutknecht, M. Manni, Q.C. Mao, B. Erni, The glucose transporter of *Escherichia coli* with circularly permuted domains is active in vivo and in vitr, J. Biol. Chem. 273 (1998) 25745–25750.
[20] R. Beutler, M. Kaufmann, F. Ruggiero, B. Erni, The glucose transporter of the *Escherichia coli* phosphotransferase system: linker insertion mutants and split variants, Biochemistry 39 (2000) 3745–3750.
[21] R. Lanz, B. Erni, The glucose transporter of the *Escherichia coli* phosphotransferase system-Mutant analysis of the invariant arginines, histidines and domain linker, J. Biol. Chem. 273 (1998) 12239–12243.
[22] Q. Mao, T. Schunk, K. Flükiger, B. Erni, Functional reconstitution of the purified mannose phosphotransferase system of *Escherichia coli* into phospholipid vesicles, J. Biol. Chem. 270 (1995) 5258–5265.
[23] S. Levy, G.Q. Zeng, A. Danchin, Cyclic AMP synthesis in *Escherichia coli* strains bearing known deletions in the pts phosphotransferase operon, Gene 86 (1990) 27–33.
[24] J. Elliott, W. Arber, *Escherichia coli* mutants which block phage lambda DNA injection coincide with ptsM which determines a component of a sugar transport system, Mol. Gen. Genet. 161 (1978) 1–8.
[25] A. Buhr, K. Flükiger, B. Erni, The glucose transporter of *Escherichia coli*. Overexpression, purification and characterization of functional domains, J. Biol. Chem. 269 (1994) 23437–23443.
[26] B. Erni, H. Trachsel, P.W. Postma, J.P. Rosenbusch, Bacterial phosphotransferase system. Solubilization and purification of the glucose-specific enzyme II from membranes of *Salmonella typhimurium*, J. Biol. Chem. 257 (1982) 13726–13730.
[27] M. Meins, B. Zanolari, J.P. Rosenbusch, B. Erni, Glucose permease of *Escherichia coli*. Purification of the IIGlc subunit and functional characterization of its oligomeric forms, J. Biol. Chem. 263 (1988) 12986–12993.
[28] Z. Markovic-Housley, B. Stolz, R. Lanz, B. Erni, Effects of tryptophan to phenylalanine substitutions on the structure, stability and enzyme activity of the IIAB(Man) subunit of the mannose transporter of *Escherichia coli*, Protein Sci. 8 (1999) 1530–1535.
[29] N.D. Meadow, P. Coyle, A. Komoryia, C.B. Anfinsen, S. Roseman, Limited proteolysis of IIIGlc, a regulatory protein of the phosphoenolpyruvate:glycose phosphotransferase system, by membrane-associated enzymes from *Salmonella typhimurium* and *Escherichia coli*, J. Biol. Chem. 261 (1986) 13504–13509.
[30] J.M. Thornton, B.L. Sibanda, Amino and carboxy-terminal regions in globular proteins, J. Mol. Biol. 167 (1983) 443–460.
[31] K. Luger, U. Hommel, M. Herold, J. Hofsteenge, K. Kirschner, Correct folding of circularly permuted variants of a beta alpha barrel enzyme in vivo, Science 243 (1989) 206–210.
[32] R. Koebnik, L. Krämer, Membrane assembly of circularly permuted variants of the *E. coli* outer membrane protein OmpA, J. Mol. Biol. 250 (1995) 617–626.
[33] J. Hennecke, P. Sebbel, R. Glockshuber, Random circular permutation of DsbA reveals segments that are essential for protein folding and stability, J. Mol. Biol. 286 (1999) 1197–1215.
[34] N.O. Keyhani, O. Boudker, S. Roseman, Isolation and characterization of IIAChb, a soluble protein of the enzyme II complex required for the transport/phosphorylation of *N*,*N*′-Diacetylchitobiose in *Escherichia coli*, J. Biol. Chem. 275 (2000) 33091–33101.
[35] F.B. Perler, D.G. Comb, W.E. Jack, et al., Intervening sequences in an Archaea DNA polymerase gene, Proc. Natl. Acad. Sci. USA 89 (1992) 5577–5581.
[36] S. Chong, Y. Shao, H. Paulus, J. Benner, F.B. Perler, M.Q. Xu, Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage and establishment of an in vitro splicing system, J. Biol. Chem. 271 (1996) 22159–22168.