Biophysical Journal, Volume 99
Supporting Material

pK_a values for the unfolded state under native conditions explain the pH dependent stability of PGB1

Stina Lindman, Mikael C. Bauer, Mikael Lund, Carl Diehl, Frans A. A. Mulder, Mikael Akke, Sara Linse¶

Cloning, expression and purification of ^{13}C and ^{15}N labeled PGB1 fragments
The PGB1-QDD fragments were cloned into the Intein Plasmid system pTXB1 (New England Biolabs, Beverly, MA) according to the supplied protocol. The protein or peptide of choice is fused N-terminally to an intein tag followed by a chitin binding domain (CBD). In the presence of thiol reagents, such as DTT, the intein undergoes specific self-cleavage which releases the target protein from the chitin-bound intein tag (1). The IMPACT system offers a convenient purification protocol, and also provides a means of avoiding degradation during production of small peptides.

The N-terminal fragment (denoted N41) covers residues 1–40 of PGB1-QDD and an additional Gly that was added to increase the yield. The identity of the C-terminal residue affects the yield, where Asp—which is residue number 40 in PGB1—typically results in poor yields (IMPACT kit, instruction manual, New England Biolabs). The C-terminal fragment (denoted C16) comprises residues 41–56 of the wt PGB1 sequence, which is the same as in PGB1-QDD. Plasmids containing either N41 or C16 were transformed into *E. coli* ER2566 cells using electroporation. Single colonies were picked to inoculate 50 mL LB containing 50 μg/mL ampicillin for overnight cultures. 5 mL of overnight culture was transferred to 500 mL minimal medium containing 13.7 mM $^{15}NH_4Cl$, 12.5 mM ^{13}C-glucose, 0.042 M Na_2HPO_4, 0.22 M KH_2PO_4, 0.00855 M NaCl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1 μg/mL Vitamin B1, 18 μM $FeCl_3$, 50 μg/mL ampicillin. Production of the N41-intein-CBD or C16-intein-CBD fusion proteins was induced with 100 μg/mL IPTG at OD_{600} = 0.6. Six hours after induction the cells were harvested and the pellets were collected and stored at –80 ^oC until purification.
Pellets were thawed, and sonicated in 1 mM EDTA, 10 mM Tris pH 7.5. NaCl was added to the supernatant to a final concentration of 0.4 M and the solution was shaken with chitin matrix (equilibrated with 0.4 M NaCl, 1 mM EDTA, 10 mM Tris pH 7.5) for one hour at 4 ^oC. The suspension was poured into a column and washed with several column volumes of 0.4 M NaCl, 10 mM Tris pH 8.0. Cleavage of the intein tag was started by flushing the column with one column volume of 50 mM DTT, 0.4 M NaCl, 10 mM Tris pH 8.0. The column was incubated at 4 ^oC overnight and the peptide fragments were eluted by applying 50 mM DTT, 0.15 M NaCl, 10 mM Tris pH 8.0 in several fractions. Elution was performed once a day for four days to elute as much peptide as possible. From the design of the IMPACT system one would expect that only the peptide is eluted from the column. However, the affinity between the chitin binding domain and chitin seems to be too weak and we observe that both the peptide and cleaved intein tag are eluted. Fractions were analyzed using agarose gel-electrophoresis and peptide containing

fractions were pooled and lyophilized. Lyophilized fractions were dissolved and separated on a 3.4 x 180 cm Sephadex G50 superfine gel filtration column (GE Healthcare, Uppsala, Sweden) using 50 mM ammonium acetate, pH 6 as running buffer. Fractions containing peptide were lyophilized and desalted separately on Sephadex G25 or Sephadex G10 superfine gel filtration columns (GE Healthcare) in water. All fractions were analyzed using MALDI mass spectrometry to confirm peptides of correct size.

Circular dichroism (CD) spectroscopy

To investigate the structure of the fragments CD spectra were recorded at pH 7, 5 and 3 in water at a low salt concentration (no salt added except for small amounts of HCl or NaOH to adjust the pH). CD spectra were recorded using a JASCO J-815 CD spectrometer (Jasco Corporation., Tokyo, Japan) with a JASCO PTC- 423S/15 Peltier type thermostated cell holder. CD spectra were recorded from 250 to 190 nm at 25°C in a 0.1 mm cuvette at a protein concentration of 600 µM with an average of 3 scans (scan rate 20 nm/min, response 8 s, bandwidth 1 nm, resolution 1 nm). Spectra are shown in Fig. S1 and indicates mainly unfolded fragments. Fragment of residues 1-41 shows some evidence of a minor fraction of secondary structure at pH 3.

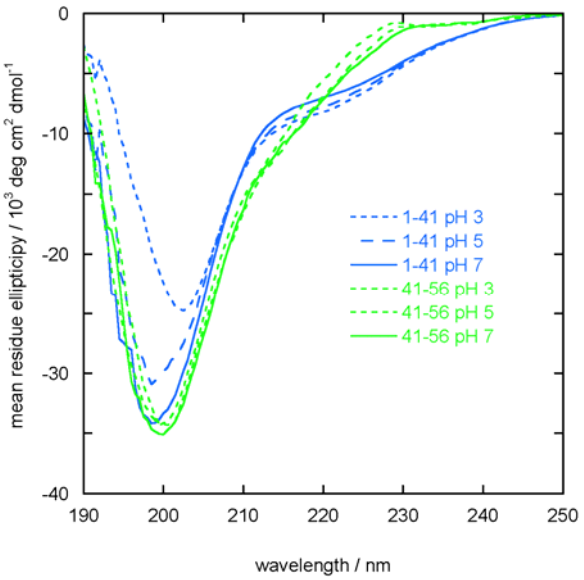

Fig S1. Far UV CD spectra of fragments N41 (blue) and C16 (green) at pH 3 (short dashed line), pH 5 (long dashed line) and pH 7 (solid line).

Gaussian chain model to correct for cleavage

To evaluate the difference in pK_a values between unfolded intact protein and unfolded fragments a Gaussian chain model (2) was used. This model describes the protein as an ideal polymer where the distances (r) between charged groups follow a Gaussian distribution p(r):

$$p(r) = 4\pi r^2 (3/2\pi d^2)^{3/2} \exp(-3r^2/2d^2) \qquad \text{(S1)}$$

where d is the RMS distance calculated as,

$$d=bl^{1/2}+s \quad \text{(S2)}$$

where the effective distance (b) between two consecutive residues is set to 7.5 Å, l is the number of peptide bonds separating the two charged groups, and s (set to 5 Å) is a shift that accounts for that the distance of interest is between two side chains (2, 3). The Debye-Hückel theory calculates the electrostatic interaction (kJ/mol) between residue i and j:

$$U_{ij} = \frac{N_A q_i q_j}{4\pi\varepsilon_0\varepsilon}\exp(-\kappa r)/r \quad \text{(S3)}$$

where q is the charge, ε is the dielectric constant of the solvent (set to 78.5), κ the inverse of the Debye screening length (set to 1/30.4 Å) and r is the distance between the charges (in Å). The mean electrostatic interaction energy between residue i and j is found as,

$$W_{ij} = \frac{N_A q_i q_j}{4\pi\varepsilon_0\varepsilon}\int_0^\infty drp(r)\exp(-\kappa r)/r \quad \text{(S4)}$$

The interaction energy is calculated using numerical integration and the average charge of a residue is found through Monte-Carlo sampling using the following Boltzmann factor,

$$\exp\left[-\Delta W_{tot}/kT \pm \ln 10(pH - pK_a^{mod})\right] \quad \text{(S5)}$$

W_{tot} is the interaction energy for the sum over all pairs of charges and ΔW_{tot} refers to the change in interaction energy when randomly changing the ionization of a residue. pK_a^{mod} is the model pK_a value where values of 4.0, 4.4 and 3.8 were used for Asp, Glu and the C-terminus, respectively (4). At each pH the average charge of all negative residues and C-termini in the intact protein and in N41 and C16 was obtained. Titration curves were obtained by performing the simulations at pH-values from 1–8 and Eq. S6 was fit to data using Kaleidagraph to obtain pK_a values.

$$q^i = \frac{Q}{1+10^{n_H Q(pH-pK_a^i)}} \quad \text{(S6)}$$

Experimental pK_a values of the fragments were corrected for cleavage and introduction of new N- and C-termini. From MC simulations, using the Gaussian chain model, the pK_a values of unfolded, intact chain and unfolded fragments could be calculated. The corrected values were calculated as:

$$pK_{a,corr}=pK_{a,exp}+(pK_{a,intact}-pK_{a,fragment}) \quad \text{(S7)}$$

References, Supporting Material

1. Chong, S. R., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M. Q. (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, *Gene. 192*, 271-281.
2. Zhou, H. X. (2002) A Gaussian-chain model for treating residual charge-charge interactions in the unfolded state of proteins, *Proc. Natl. Acad. Sci. USA. 99*, 3569-3574.
3. Zhou, H. X. (2001) Loops in proteins can be modeled as worm-like chains, *J. Phys. Chem. B. 105*, 6763-6766.
4. Nozaki, Y., and Tanford, C. (1967) Examination of titration behavior, *Methods Enzymol. 11*, 715-734.