Selective Incorporation of Nitrile-Based Infrared Probes into Proteins via Cysteine Alkylation

Hyunil Jo,[‡] Robert M. Culik,[‡] Ivan V. Korendovych,[‡] William F. DeGrado,[†,‡,*] and Feng Gai[†,*]

[†]Department of Chemistry and [‡]Department of Biochemistry & Biophysics, University of Pennsylvania, Philadelphia, PA 19104

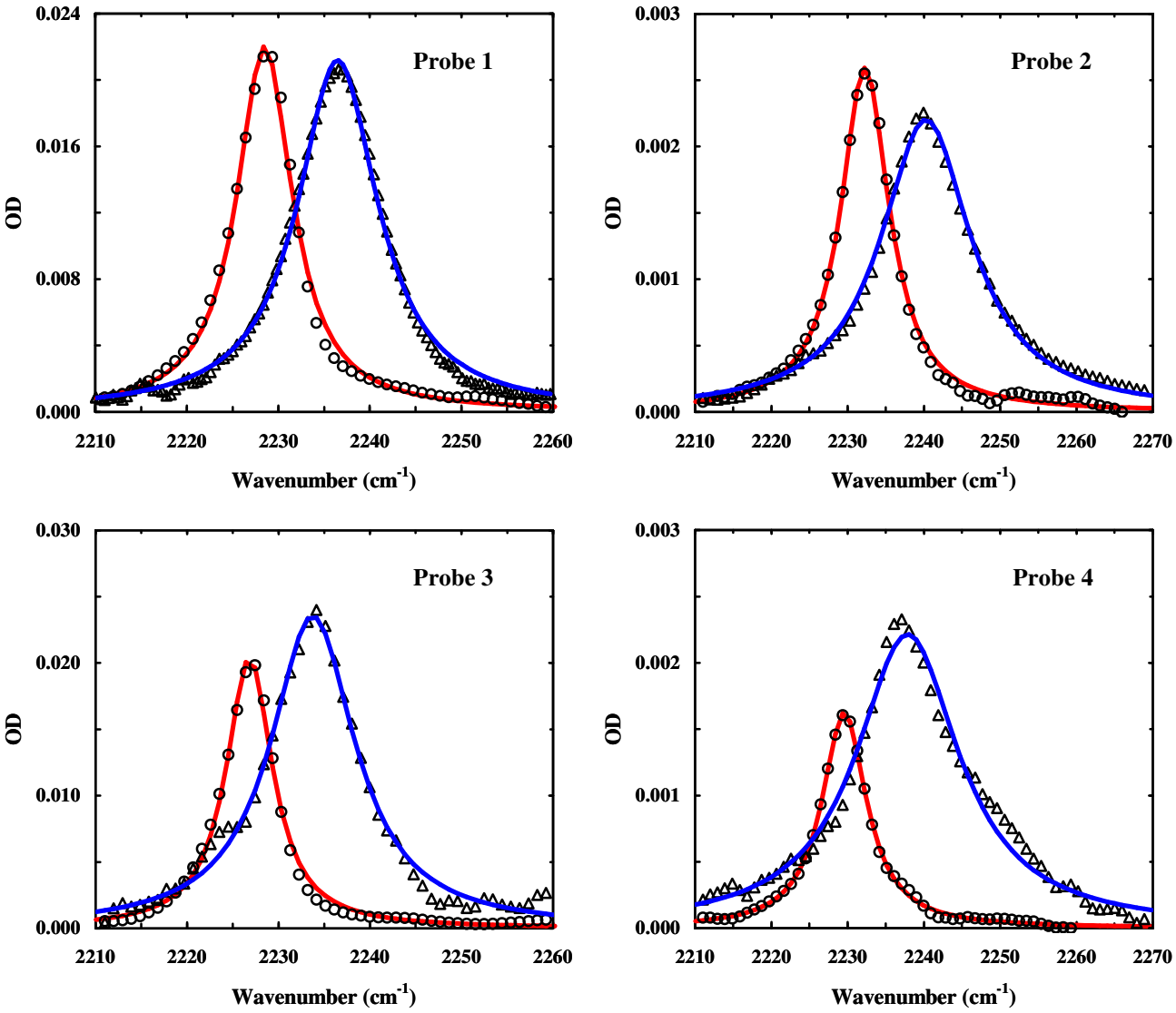

Figure S1. The C≡N stretching bands of the four cysteine derivatives, as indicated, obtained in water (open triangles) and THF (open circles). Lines are respective fits of these data to a Lorentzian function and the fitting parameters are given in Table 1. The sample concentrations were in the range of 5–10 mM, estimated by weight.

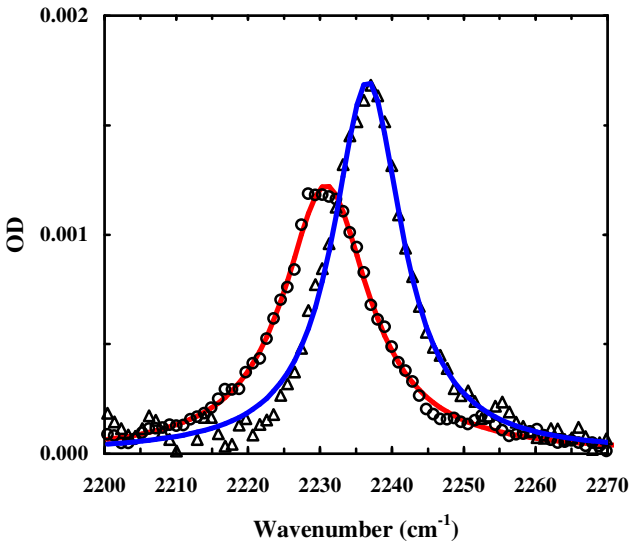

Figure S2. The C≡N stretching bands of A8C-CN obtained in the absence (open triangles) and presence (open circles) of CaM (50 mM HEPES buffer, pH 7.4, 30 mM $CaCl_2$). The concentrations of A8C-CN and CaM were estimated to be about 1-2 mM. Lines are respective fits of these data to a Lorentzian function with the following parameters: for A8C-CN ν = 2236.6 cm^{-1} and $\Delta\nu$ = 11.7 cm^{-1}, for A8C-CN/CaM ν = 2230.8 cm^{-1} and $\Delta\nu$ = 14.4 cm^{-1}.

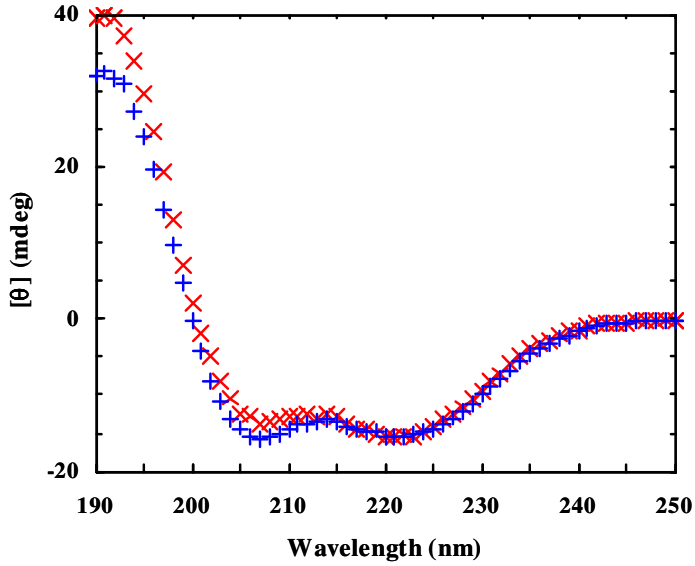

Figure S3. Far-UV CD spectra of the nitrile-labeled (+) and unlabeled (×) human calmodulin-like proteins. In both cases, the concentration of the protein was approximately 18 μM (in 50 mM HEPES buffer, pH 7.4).

Materials and Methods

Cysteine derivatives: The four model nitrile probes (**1**, **2**, **3**, and **4**) were synthesized from N-acetyl cysteine using previously reported procedures (see reference 30 for detail). All the prepared compounds show satisfactory NMR (^{1}H and ^{13}C) and ESI-MS data, which will be reported elsewhere. The IR samples of these probes were prepared by directly dissolving lyophilized solids in either Millipore water or 99.5+% spectroscopic grade tetrahydrofuran (Acros Organics) and the corresponding concentrations were estimated by weight.

Mastoparan-X mutants: The two mutants of mastoparan-X (sequence: INWKGIAAMAKKLL), W3C and A8C, were synthesized on an Argonaut Quest 210 synthesizer (Argonaut Technologies) at 0.092 mmol scales using Fmoc-CLEAR^{TM} Amide resin (substitution level – 0.46 mmol/g). Activation of Fmoc-amino acids (5-fold excess) was achieved with 0.95 equiv (relative to the amino acid) excess of HBTU in the presence of 10 equiv of DIPEA in DMF (HPLC grade, Aldrich). Side chain deprotection and simultaneous cleavage from the resin were carried out using a mixture of TFA/thioanisole/ethanedithiol/anisole (90:5:3:2, v/v) at room temperature, for 3 hours. After filtration, most of the solvent was then evaporated using a stream of N_2. The crude peptide was obtained by ether precipitation, which was then purified to homogeneity by reverse-phase chromatography (Agilent 1100 Series) with a C18 preparative column (Vydac) using a constant 10% buffer A (1% TFA in Millipore water), and a linear gradient of buffer B (100% acetonitrile) and buffer C (100% Millipore water). The mass of all peptides was verified using a Bruker Ultraflex III mass spectrometer.

Purified MpX mutants (W3C and A8C) were then used to perform the cysteine alkylation reaction with 4-bromomethyl benzonitrile under the conditions specified in Scheme 1. Purification was performed as described above for the cysteine derivatives. The IR samples were prepared by directly dissolving lyophilized peptide solids into 50 mM HEPES buffer (pH 7). For the peptide binding experiments, the buffer also contains 30 mM $CaCl_2$ and 2 mM CaM. The final concentration of the A8C-CN was approximately 2 mM, estimated by Trp absorbance at 280 nm using an extinction coefficient of 5,500 $M^{-1}cm^{-1}$. The final concentration of the W3C-CN was estimated to be between 1-2 mM based on the absorbance at 214 nm using an extinction coefficient of 15,521 $M^{-1}cm^{-1}$.

Calmodulin variants: Human CaM[1] was used in the MpX binding experiments, which has the following sequence:

ADQL TEEQI AEFKE AFSLF DKDGD GTITT KELGT VMRSL GQNPT EAELQ DMINE VDADG NGTID FPEFL TMMAR KMKDT DSEEE IREAF RVFDK DGNGY ISAAE LRHVM TNLGE KLTDE EVDEM IREAD IDGDG QVNYE EFVQM MTAK

A His-tagged version of human calmodulin-like protein CALM3, which contains a single cysteine residue, was used to test the cysteine alkylation reaction in proteins. The CALM3 gene was obtained from Invitrogen (cloned into pEXP5-NT/TOPO vector) and has the following sequence:

SGSHH HHHHG SSGEN LYFQS LMADQ LTEEQ VTEFK EAFSL FDKDG DGCIT TRELG TVMRS LGQNP TEAEL RDMMS EIDRD GNGTV DFPEF LGMMA RKMKD TDNEE EIREA FRVFD KDGNG FVSAA ELRHV MTRLG EKLSD EEVDE MIRAA DTDGD GQVNY EEFVR VLVSK

Both CaM and CALM3 were expressed in E. coli BL21(DE3) pLysS cells (Novagen) at 37˚ and purified on an ACTA FPLC system (Amersham Biosystems) over a Ni-NTA column (Qiagen) in the case of CALM3, and twice over a phenyl sepharose HiTrap column in the case of CaM. The purity of the proteins was assessed by SDS-PAGE, MALDI and ESI-MS.

Cysteine alkylation of CALM3: 3 mL of a buffered (50 mM HEPES, pH 7.0, 100 mM NaCl) solution of CALM3 (100 μM) was first diluted with a NH_4HCO_3 buffer (15 mL, 100 mM, pH = 8.4). Then, an aliquot of an aqueous TCEP solution (0.12 mL, 4.2 mM) was added to the above solution and stirred for 1 hour at room temperature, which was followed by addition of an aliquot of 4-bromomethyl benzonitrile DMF solution (0.25 mL, 2.5 mM). These reaction conditions are incredibly mild – it should be noted that the DMF present in the final reaction solution is less than 2% v/v. The reaction mixture was then stirred overnight at room temperature. Reaction progress was monitored by LC-MS. The small-molecule reagents were removed using a desalting column 10DG (BioRad). The solution obtained was then dialyzed twice into 50 mM HEPES buffer (pH 7.4) using a 10 kD dialysis cassette (Thermo Scientific).

Circular Dichroism Measurements: CD spectra were collected on an AVIV 410 spectropolarimeter using a 1 mm quartz cell.

FTIR Measurements: FTIR spectra were collected on a Nicolet Magna-IR 860 spectrometer using 1 cm^{-1} resolution and a CaF_2 sample cell that was divided into two compartments with a Teflon spacer (52 μm). For all reported spectra, a background was subtracted.

Reference

1. The gene of CaM is a generous gift of Professor A. Joshua Wand [Urbauer, J. L.; Short, J. H.; Dow, L. K.; Wand, A. J. (1995) Structural analysis of a novel interaction by calmodulin – high-affinity binding of a peptide in the absence of calcium. *Biochemistry 34,* 8099-8109]