**Organic and Biomolecular Chemistry**

# Supplementary Material

## The Design, Synthesis, and Characterization of the First Cavitand-Based *De Novo* Hetero-Template Assembled Synthetic Proteins (Hetero-TASPs).

Heidi E. K. Huttunen-Hennelly, John C. Sherman*

*Correspondence to John C. Sherman, Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, BC, V6T 1Z1, Canada, Fax: (+1) 604)822-2847
**email:** John C. Sherman (sherman@chem.ubc.ca)

*Synthesis of the Hetero-TASPs via the Protecting Group Approach:*

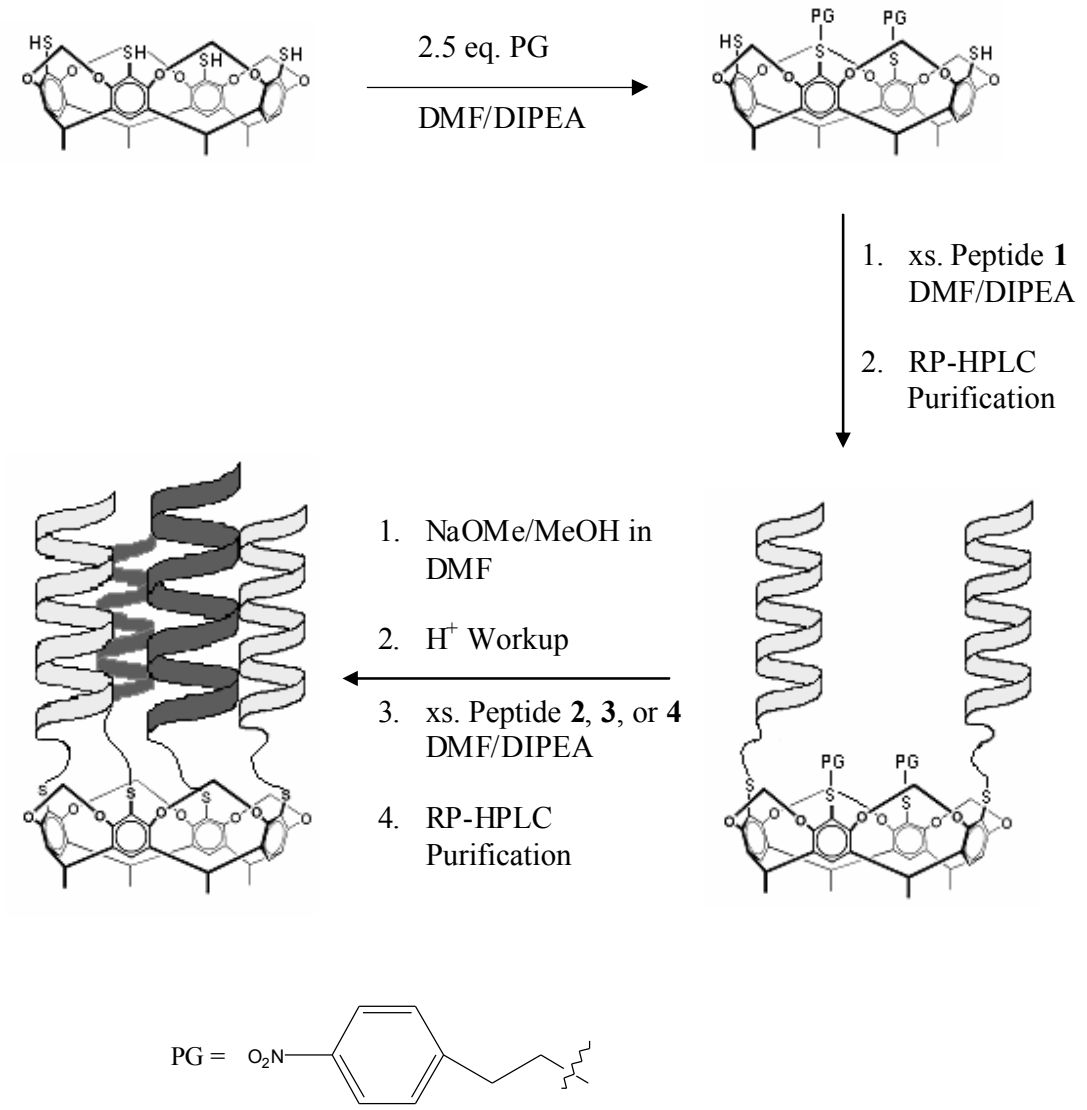


**Figure 19.** Schematic Representation Outlining the General Synthesis for an a,c Hetero- TASP where PG is the 2-(4-nitrophenyl)ethyl Group.

The general syntheses of hetero-TASPs **9-16** involved the protection of two thiol moieties on the cavitand template, followed by peptide attachment at the remaining two thiols,

selective deprotection of the protecting groups, and subsequent different peptide sequence attachment. A solution of the arylthiol cavitand (5 mg, 6.9 μmol, 1 equiv.) with 2-(4-nitrophenyl)ethyl bromide (npe) (4 mg, 17.4 μmol, 2.5 equiv.) in degassed DMF (5 mL) was stirring under $N_2$ to which DIPEA (100 μL) was then added in excess. The reaction was left to stir for 4 hours. The crude reaction mixture was then evaporated *in vacuo* and not purified further. A crude reaction mixture of protected cavitand (5 mg) and excess lg3 peptide (30 mg) was stirring in degassed DMF under $N_2$. DIPEA (100 μL) was then added in excess and the reaction was left to stir for 4 hours. The reaction mixture was then evaporated *in vacuo*, dissolved in $H_2O$, filtered, and purified by reversed-phase HPLC to isolate the mono, di (a,b and a,c), tri, and tetrasubstituted caviteins with npe protecting groups at the remaining thiol moieties.

The isolated a,b and a,c disubstituted caviteins (2 mg, 0.4 μmol, 1 equiv.) were then individually subjected to a solution of excess MeONa/MeOH (0.03 mL) in degassed DMF under $N_2$ to cleave the npe protecting groups. After 2 hours the reaction mixture was acidified with 0.1 M HCl until acidic (acidity tested with litmus paper). The mixture was then evaporated *in vacuo*, dissolved in $H_2O$, filtered and purified by RP-HPLC to afford pure a,b and a,c lg3 disubstituted caviteins, respectively.

The isolated disubstituted a,b and a,c hetero-TASP intermediates were then individually subjected to an excess of appropriate peptide to yield the a,b and a,c disubstituted hetero-TASPs. Note the LG3/AG3 family included the synthesis of 3LG3•1AG3 and 1LG3•3AG3 hetero-TASPs in addition to the disubstituted counterparts. The crude products were then purified by RP-HPLC and lyophilized to afford the desired hetero-TASPs.

*Identification of the a,b and a,c hetero-TASPs:*

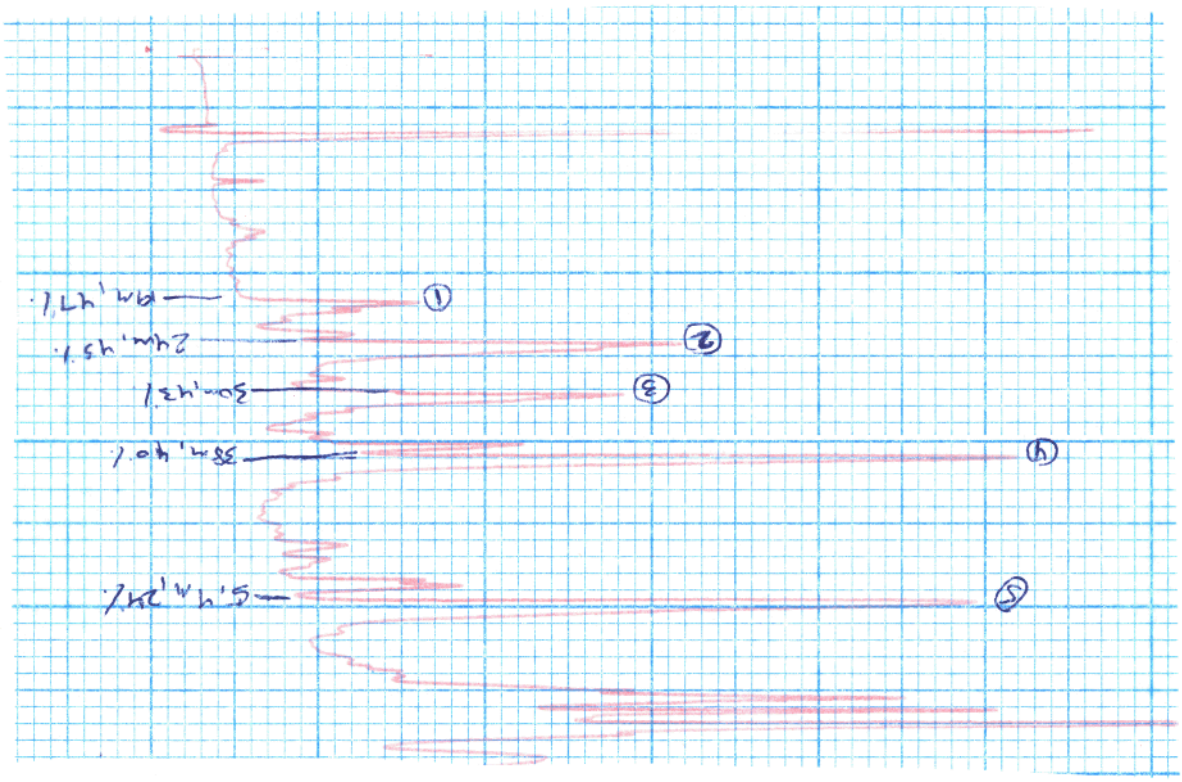

**Figure 20.** The Purification of the Crude Reaction Mixture of the lg3 Substituted Hetero-TASPs by Preparatory Reversed-Phase HPLC Using a Gradient of 20-70 % Acetonitrile (with 0.05 % TFA) in Water (with 0.1 % TFA) Over 60 Minutes.

Peak (1) = LG3/4pep

Peak (2) = LG3/3pep

Peak (3) = LG3/2pep_ac

Peak (4) = LG3/2pep_ab

Peak (5) = LG3/1pep

From the HPLC spectrum it was deduced that peaks 3 and 4 correspond to the a,c and a,b disubstituted cavitein variants, respectively. Peak 4 is about two times as large as peak 3

which supports the expected 2:1 product ratio of the a,b:a,c disubstituted cavitein variants. Furthermore, in the $^{1}$H NMR spectrum for the 2LG3•2LG2C_ab cavitein in Figure 15, the cavitand signal ($H_{out}$) at ~6.1 ppm is split into three signals in a 1:2:1 ratio. A 1:2:1 splitting of the $H_{out}$ cavitand signal supports the a,b symmetry. In this way, the HPLC and $^{1}$H NMR spectra provide evidence for the supposed identities of the a,c and a,b disubstituted products.

*ANS Binding Data for the hetero-TASPs:*

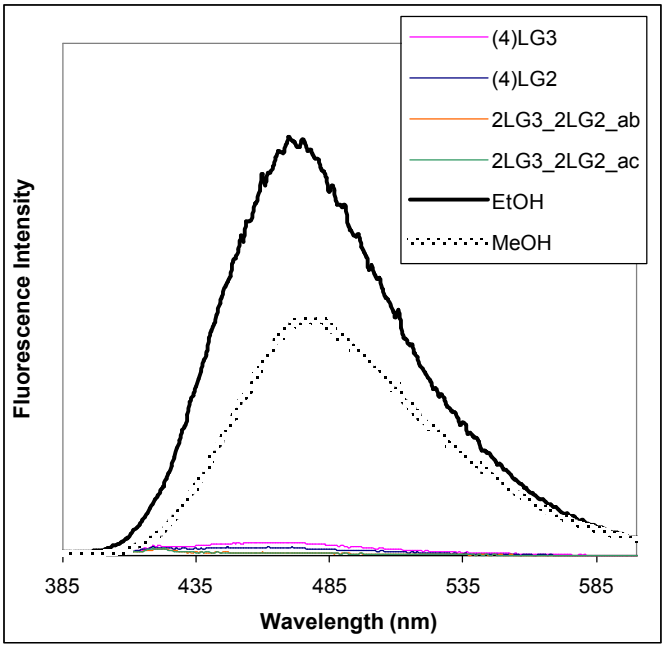


**Figure 21.** Fluorescence Emission Spectra of 2 μM ANS in the Presence of 95 % Ethanol, 100 % Methanol, 50 μM LG3/LG2 Substituted Caviteins at 20 °C in pH 7.0 50 mM Phosphate Buffer.

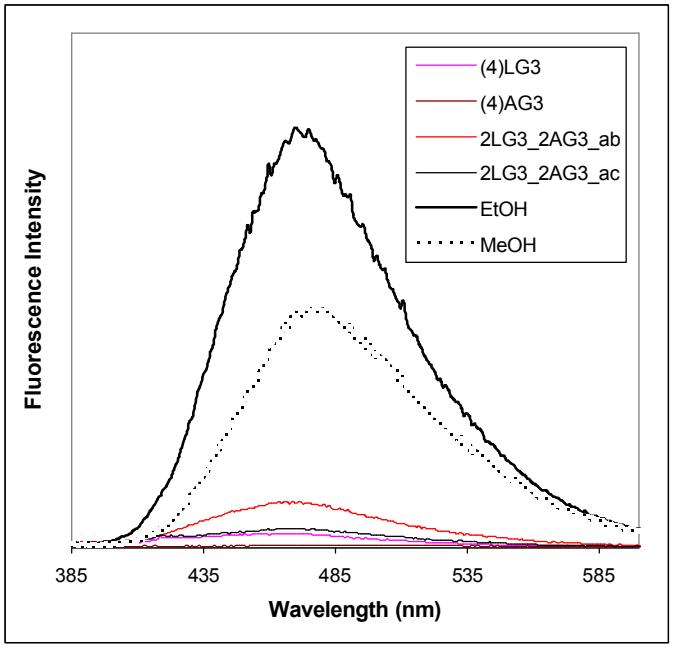


**Figure 22.** Fluorescence Emission Spectra of 2 μM ANS in the Presence of 95 % Ethanol, 100 % Methanol, 50 μM LG3/AG3 Substituted Caviteins at 20 °C in pH 7.0 50 mM Phosphate Buffer.

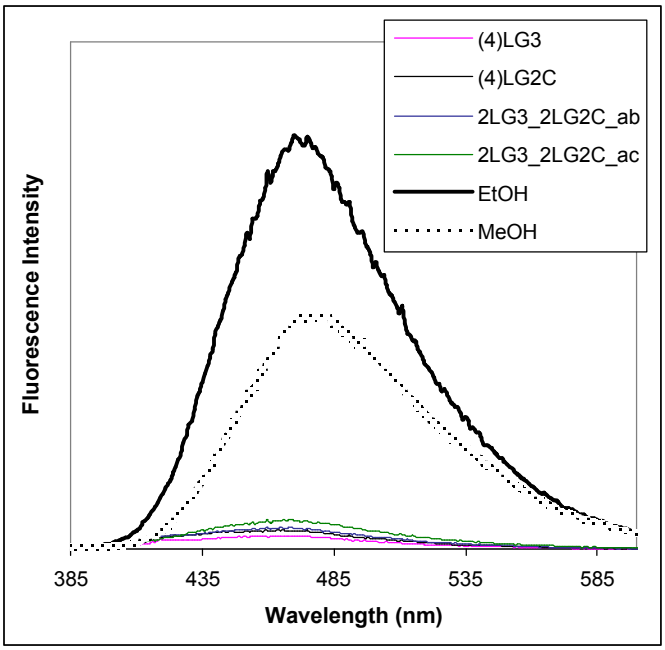


**Figure 23.** Fluorescence Emission Spectra of 2 μM ANS in the Presence of 95 % Ethanol, 100 % Methanol, 50 μM LG3/LG2C Substituted Caviteins at 20 °C in pH 7.0 50 mM Phosphate Buffer.

*Synthesis of Peptide lg2c:*

A C-terminal cysteine residue was added on the automated peptide synthesizer containing an S-trityl side chain protecting group. After peptide lg2c was removed from the automated peptide synthesizer the free N-terminus was acetylated through a manual treatment of lg2c (~300 mg resin, ~100 mg peptide **4**, ~0.6 mmol) with excess acetic anhydride (3.5 mL) in 1 mL NMP for 1 hour at room temperature. The peptide resin was then filtered through a medium frit filter with DCM. The lg2c peptide was then cleaved off the resin and of protecting groups using a mixture of TFA (95 %), $H_2O$ (2.5 %) and 1,2 ethanedithiol (2.5 %). The last step involved activating the free C-terminus of lg2c by adding lg2c (20 mg, 10 μmol) in 3 mL ethanol to a rapidly stirring solution of 2,2'-dipyridyl disulfide (12 mg, 55 μmol) in 2 mL ethanol. The reaction was stirred at room temperature for 1 hour. The ethanol was reduced to 1 mL *in vacuo*, and the solution was pipetted onto ice-cold diethyl ether. The resulting solid was re-dissolved in water (1.5 mL) containing 0.1 % TFA, and filtered using a 0.45 μm nylon filter. Subsequent purification of lg2c by RP-HPLC and lyophilization afforded peptide **4** as a white solid (50 mg, 19 %). All of the peptides were characterized for purity by the inspection of a single peak by analytical reversed-phase HPLC (>95 % pure), and using MALDI-MS.

*Sedimentation Equilibrium NONLIN Fits:*

The initial fit of the data using NONLIN was completed as follows: (1) let delta y, baseline offset and reduced molecular weight vary, (2) fix delta y, let baseline offset and reduced molecular weight vary (note: a series of iterations followed until a reduced molecular weight value was converged upon) (3) evaluate error limits. The experimental molecular weight can then be calculated from:

$$\sigma = (M(1 - \bar{v}\rho)\omega^2)/RT$$

where $\sigma$ is the reduced molecular weight.

In order to verify that the monomer fit was indeed accurate, the data were fit to the equation below describing possible oligomeric species:

$$A_r = A_{(monomer,\ ro)} \exp\left[ ((1 - \bar{v}\rho)\omega^2 / 2RT)\ (M(r^2 - r_o^2))\right]$$
$$+ A_{(monomer,\ ro)}^{n2} K_{a,2} \exp\left[ ((1 - \bar{v}\rho)\omega^2 / 2RT)\ n_2(M(r^2 - r_o^2))\right]$$
$$+ A_{(monomer,\ ro)}^{n3} K_{a,3} \exp\left[ ((1 - \bar{v}\rho)\omega^2 / 2RT)\ n_3(M(r^2 - r_o^2))\right]$$
$$+ A_{(monomer,\ ro)}^{n4} K_{a,4} \exp\left[ ((1 - \bar{v}\rho)\omega^2 / 2RT)\ n_4(M(r^2 - r_o^2))\right] + E$$

where $A_{(monomer,\ ro)}$ is the absorbance of the monomer at the reference radius $r_o$, $n_2$ is the stoichiometry for species 2, $K_{a,2}$ is the association constant for the monomer –*n*– mer equilibrium of species 2, $n_3$ is the stoichiometry for species 3, $K_{a,3}$ is the association constant for the monomer –*n*– mer equilibrium of species 3, $n_4$ is the stoichiometry for species 4, $K_{a,4}$ is the association constant for the monomer –*n*– mer equilibrium of species 4.

The second fit included a monomer-dimer equilibrium and was completed as follows: (1) let delta y, baseline offset and $K_{a,2}$ vary, and fix the reduced molecular weight for a monomer (if unknown let vary), (2) fix delta y, let baseline offset and $K_{a,2}$ vary, and fix the reduced molecular weight for a monomer (if unknown let vary) (note: a series of iterations followed until a value for $K_{a,2}$ was converged upon) (3) evaluate error limits. For thoroughness, the data were fit to the equations describing a monomer-trimer and a monomer-tetramer equilibrium respectively, to eliminate other possible association states.