

## SUPPORTING INFORMATION

YVEFGLGAGAPGEPGNPGSPGNQGQPGNKGSPGNPGQPGNEGQPGQPGQNGQPGEPGSNGPQGS  
QGNPGKNGQPGSPGSQGS PGNQGS PGQPGNPGQPGEQ GKPGNQG PAGE PGNPGSPGNQGQPGNK  
GSPGNPGQPGNEGQPGQPGQNGQPGEPGSNGPQGSQGNPGKNGQPGSPGSQGS PGNQGS PGQPG  
NPGQPGEQ GKPGNQG PAGE (GAGAGAGH)<sub>n</sub> GAGAPGEPGNPGSPGNQGQPGNKGSPGNPGQPGNE  
GQPGQPGQNGQPGEPGSNGPQGSQGNPGKNGQPGSPGSQGS PGNQGS PGQPGNPGQPGEQ GKPG  
NQG PAGE PGNPGSPGNQGQPGNKGSPGNPGQPGNEGQPGQPGQNGQPGEPGSNGPQGSQGNPGK  
NGQPGSPGSQGS PGNQGS PGQPGNPGQPGEQ GKPGNQG PAGE GA

Figure S1: Amino acid sequence of protein polymers (n = 8, 16, 24 or 48) used in this study.

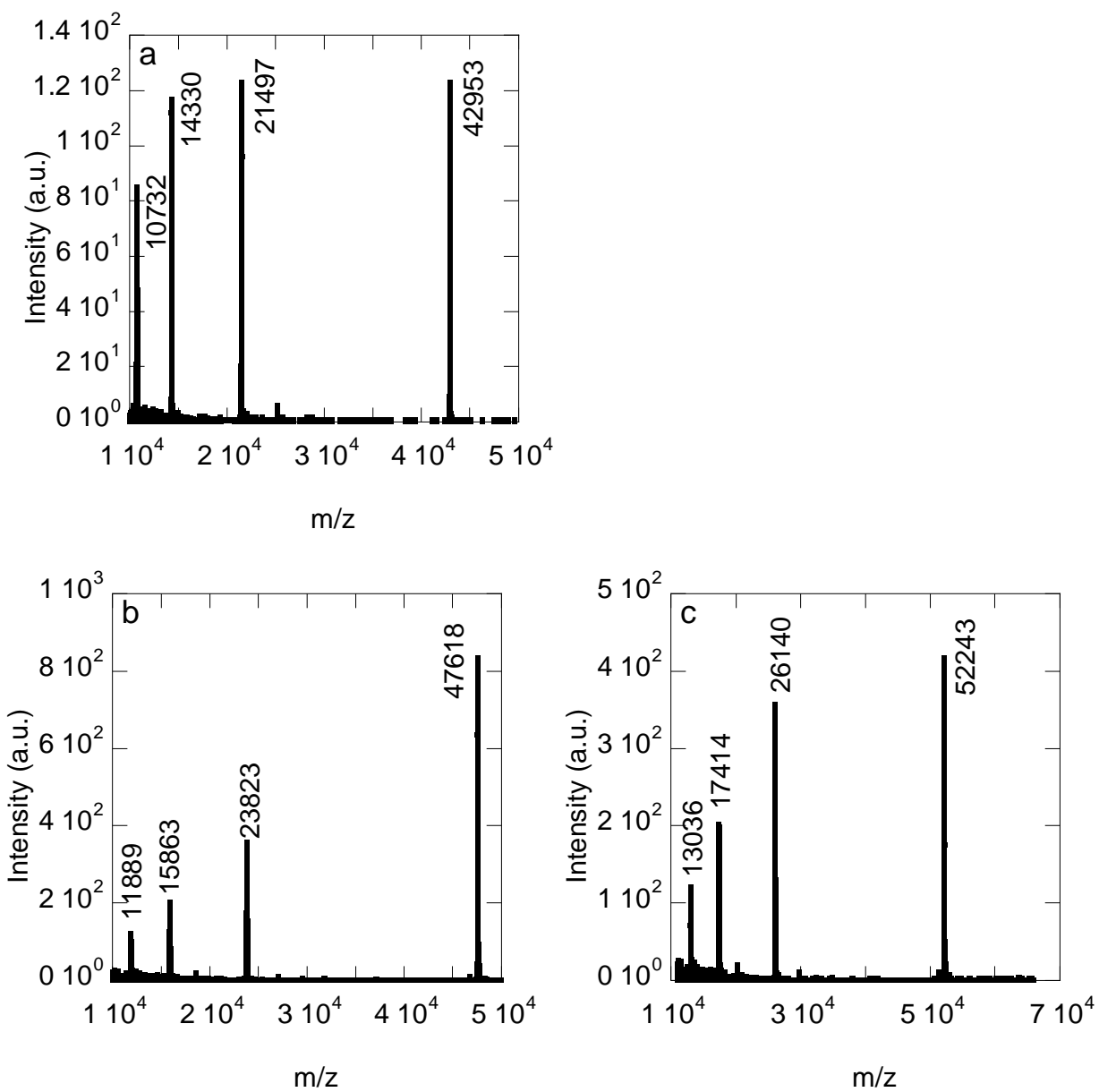


Figure S2: MALDI-TOF spectra of  $C_2S^H_8C_2$  (a),  $C_2S^H_{16}C_2$  (b) and  $C_2S^H_{24}C_2$  (c). Each spectrum shows four peaks corresponding to the  $(M + H)^+$ ,  $(M + 2H)^{2+}$ ,  $(M + 3H)^{3+}$  and  $(M + 4H)^{4+}$ .