

High level expression and facile purification of recombinant silk-elastin-like polymers in auto induction shake flask cultures

Raul Machado, João Azevedo-Silva, Cristina Correia, Tony Collins, Francisco Javier Arias, Jose Carlos Rodríguez-Cabello, Margarida Casal

Additional file 1 – Nucleotide and amino acid sequences of SELP constructions.

SELP-109-A**(S₁₀E₉)₇***Nucleotide Sequence*

GTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCT
CCGGTGCAGGCGCGCTGGTTCTGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGC
GGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTG
GGTGTTCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGG
TGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTC
TGGTGCAGGCGCGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTAGGTGCAGGCGCG
GGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTG
GTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCG
GCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAG
GTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCG
GTGCGGGCGCTGTTCTGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGG
TAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGT
GTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTG
GCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGG
TGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTAGGTGCAGGCGCGGGT
TCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTT
CTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGC
GGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGT
GCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGT
GCGGGCGCTGTTCTGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTA
GCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTG
TCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGG
GTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGT
CAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTAGGTGCAGGCGCGGGTTC
CGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCT
GGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGG
TGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCA
GGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCG
GGCGCTGGTTCTGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCG
GCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCC
GGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTT
CCTGCCGTGGGTGTTCCGGCAGTGGGC

Amino Acid Sequence

VGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGAGAGSGA
GAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
VGVGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGAGAGS
GAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
PAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGAGA
GSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
VPAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGA
GAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
AVGVPAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGS
GAGAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVG
VPAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGAGA
GSGAGAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPA
VGVPVAVGVPAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGA
GAGSGAGAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGV
PAVGVPVAVGVPAVG

SELP-59-A

(S₅E₉)₉

Nucleotide Sequence

GTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCT
CCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCC
GGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCCCTGCCGTGGGTGTT
CCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGC
GCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGG
CTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCC
TGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGC
GCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAG
TGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGC
GGTGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGC
TCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGACAGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGG
GCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGT
GGGTGTTCCGGCGGTGGGCGTTCCCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGC
GCTGGTGCAGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCG
TACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGG
CGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCCCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGC
GCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGC
GCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGT
TCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCCCTGCCGTGGGTGTTCCGGCAGTGGGC
GTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCT
CCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCC
GGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCCCTGCCGTGGGTGTT
CCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGC
GCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGG
CTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCC
TGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGC
GCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCGGGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAG
TGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGC
GGTGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGC

Amino Acid Sequence

VGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
AVGVPAVGVPVAVGVPAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGS
GVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVGAGAGSG
AGAGSGAGAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
VPAVGVPVAVGVPAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVPAVGVP
AVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVGAGAGSGAGAGSGAG
AGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPA
VGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVG
VPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
GAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVPAVGVPVAVGVPAVGVPVAVGVPAVGVP
VGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSG
VPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVGVPVAVGVPAVG

(S₅E₂₀)₅

[illegible][illegible]

SELP-1020-A

(S₁₀E₂₀)₄

Nucleotide Sequence

GTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCT
CCGGTGCAGGCGCGCTGGTTCTGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGC
GGGTAGCGGCGCTGGTGCAGGCTCCGGTGCAGGCGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTG
GGTGTTCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGG
TGGGCGTTCCGGCGGTGGGTGTTCCGGCTGTGGGCGTTCCAGCCGTGGGTGTTCCGGCAGTGGGCGTACCGGC
CGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCG
GCTGTGGGTGTTCCGGCGGTGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTT
CCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCAGGCGCTGGTTCC
TGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCA
GGCTCCGGTGCAGGCGCGCTGGTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCG
TTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCCGGCGGTGGG
TGTTCCGGCTGTGGGCGTTCCAGCCGTGGGTGTTCCGGCAGTGGGCGTACCGGCCGTGGGCGTTCCAGCAGTGG
GTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGT
GGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCT
GGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCAGGCGCTGGTTCTGGCGTAGGTGCAGGCGCG
GGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCAGGCGCTG
GTTCTGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCG
GCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCCGGCGGTGGGTGTTCCGGCTGTGGGCGTTCC
CAGCCGTGGGTGTTCCGGCAGTGGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTT
CCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCTGCCGTGGGTG
TTCCGGCAGTGGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGCTCTGGTGCAGGCGCGGGTAGCGG
CGCTGGTGCAGGCTCCGGTGCAGGCGCTGGTTCTGGCGTAGGTGCAGGCGCGGGTTCCGGCGCTGGTGCCGGC
TCTGGTGCAGGCGCGGGTAGCGGCGCTGGTGCAGGCTCCGGTGCAGGCGCTGGTTCTGGCGTACCGGCCGTGG
GCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGGCGGTGGGCGTTCCGGCTGT
GGGTGTTCCGGCGGTGGGCGTTCCGGCGGTGGGTGTTCCGGCTGTGGGCGTTCCAGCCGTGGGTGTTCCGGCA
GTGGGCGTACCGGCCGTGGGCGTTCCAGCAGTGGGTGTTCCGGCTGTGGGCGTTCCGGCGGTGGGTGTTCCGG
CGGTGGGCGTTCCGGCTGTGGGTGTTCCGGCGGTGGGCGTTCTGCCGTGGGTGTTCCGGCAGTGGGC

Amino Acid Sequence

VGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGAGAGSGA
GAGSGAGAGSGVPAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVP
VGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVP
VGVGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGAGAGS
GAGAGSGAGAGSGVPAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVG
PAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVP
PAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVP
GSGAGAGSGAGAGSGVPAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAV
GVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAV
GVPVAVGVGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGVVGAGAGSGAGAGSGA
GAGSGAGAGSGAGAGSGVPAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVP
AVGVPAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVPVAVGVP
AVGVPAVG