Supporting Information

Computational De Novo Design and Characterization of a Protein that Selectively Binds a Highly Hyperpolarizable Abiological Chromophore

H. Christopher Fry,[‡] Andreas Lehmann,[‡¶] Louise E. Sinks,[‡] Inge Asselberghs,[⊥] Andrey Tronin,[‡] Venkata Krishnan,[‡] J. Kent Blasie,[‡] Koen Clays,[⊥] William F. DeGrado,[▽*] Jeffrey G. Saven,[‡*] and Michael J. Therien[§*]

[‡]Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104-6323; [⊥]Department of Chemistry, University of Leuven, B-3001 Leuven, Belgium; [▽]Department of Pharmaceutical Chemistry, Investigator, Cardiovascular Research Institute, University of California at San Francisco, San Francisco, CA 94158-9001; [§]Department of Chemistry, Duke University, Durham, NC 27708-0354

Chart S1. Sequence alignment between the monomeric single chain protein ***PA$_{SC}$*** and ***SCRPZ-1***. The highlighted regions represent segments of helical overlap that were used to guide the alignment.

Figure S1. Backbone alignment of models of ***PA$_{sc}$*** (blue) and the template tertiary structure (red) of ***SCRPZ-1*** and ***SCRPZ-2***.

Figure S2. Model structures illustrating acidic and basic residues of ***SCRPZ-1*** (top), ***SCRPZ-2*** (middle) and ***PA$_{SC}$*** (Bottom) rendered in a spacefilling format from four different viewpoints. Lysine and arginine residues (blue) and aspartic and glutamic acid (red) and glutamine (purple) are highlighted to detail the differences between ***SCRPZ-1*** (Top) and ***SCRPZ-2*** (Bottom).

Gene Sequences for *SCRPZ 1-3*

Synthesis of [RuPZn](PF$_6$)$_2$

Scheme S1. Synthesis of the ***RuPZn*** chromophore.

Chart S1. Sequence alignment between the monomeric single chain protein ***PA$_{SC}$*** and ***SCRPZ-1***. The highlighted regions represent segments of helical overlap that were used to guide the alignment.

```
                          N15--- --N6    N C     C6--- -----C15
PA(SC)     1     SPEEAMQEAQ QTAREAEQAM QKHRQAYDKG DQQKALQTAK EFQQAMQKHK
SCRPZ      1         ELEKLR QTGEQILQIA KQVNEIMLKG DDDSLEQLLK LAYELIQQHT
                                 N15------N6     NC      C6------
PA(SC)     51    QYMNP---QA ISESVQKTAR YFEQAMQKHR QAYDKGDQQK ALQTAKEAQQ
SCRPZ      47    QLAYNRQEAA DTE-IMKQGQ QILEIAQQVN EVLLKGDKDS LEQLIKLAYQ
                 -C15
PA(SC)     98    AMQKHSQALR G
SCRPZ      96    LIQQLQELFE KKN------- --
```

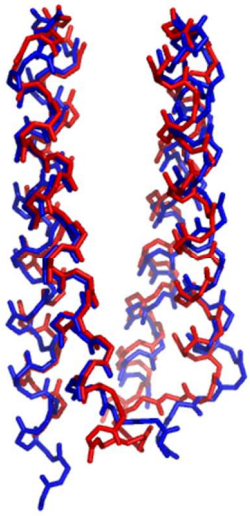


Figure S1. Backbone alignment of models of ***PA$_{sc}$*** (blue) and the template tertiary structure (red) of ***SCRPZ-1*** and ***SCRPZ-2***.

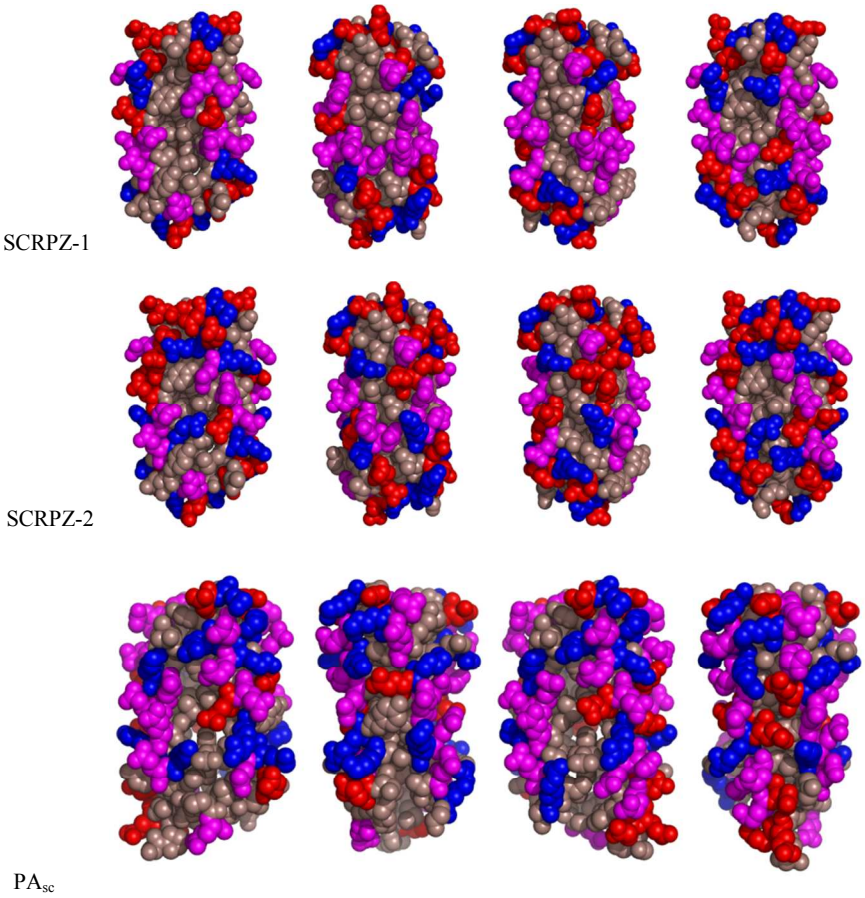


Figure S2. Model structures illustrating acidic and basic residues of ***SCRPZ-1*** (top), ***SCRPZ-2*** (middle) and ***PA$_{SC}$*** (Bottom) rendered in a spacefilling format from four different viewpoints. Lysine and arginine residues (blue) and aspartic and glutamic acid (red) and glutamine (purple) are highlighted to detail the differences between ***SCRPZ-1*** (Top) and ***SCRPZ-2*** (Bottom).

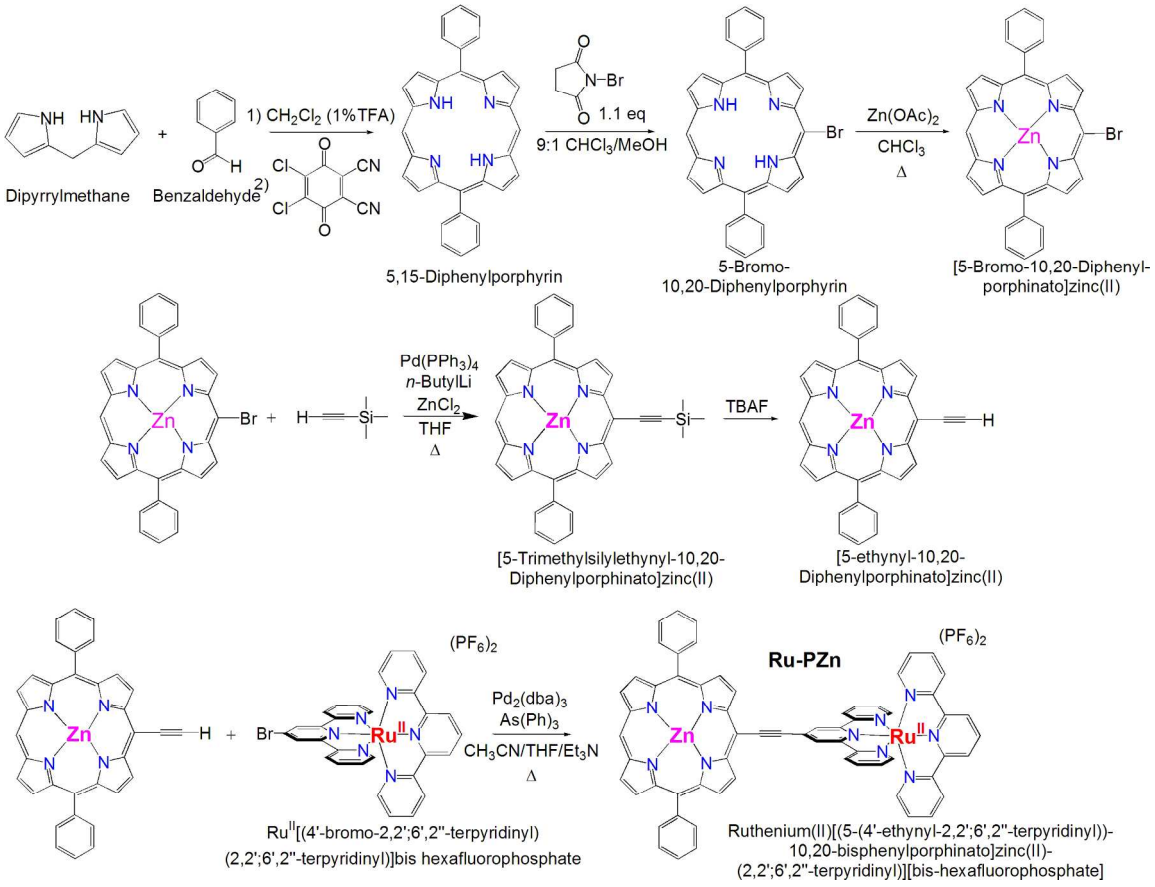
Gene Sequences for *SCRPZ 1-3.*

***SCRPZ*-1:**
5'*CATATG*GAACTGGAAAAACTGCGTCAGACTGGTGAACAGATCCTGCAAATTGCGA
AACAGGTTAACGAGATTATGCTGAAGGGTGATGACGACTCTCTGGAGCAGCTGATT
AAACTGGCGTACGAACTGATCCAACAACACACCCAGCTGGCGTACAATCGTCAGGA
AGCTGCGGATACGGAGATCATGAAACAGGGTCAGCAGATTCTGGAAATTGCCCAGC
AGGTGAACGAAGTGCTGCTGAAAGGCGACAAAGACTCTCTGGAACAGCTGTTAAAG
CTGGCGTATCAGCTGATTCAGCAACTGCAGGAGCTGTTCGAAAAGAAAAACTAA*AA
GCTT* 3'

SCRPZ-2:
5'*CATATG*GAACTGGAAAAGCTGCGTCAAACTGGTGATGAGATCCTGCAAATTGCCCA
GCGCGTGCGTGAGATTATGGACAAAGGTGACGACGATAGCCTGGAGCAACTGTTGG
AGGAAGCCTATGAGCTGATCCAAAAGCACCGCCAACTGGCCTACAACCGTCAGGAG
GCCGCGGACACCGAAATTATGAAACAAGGTGACCAGATTCTGCAAATCGCACAGCG
TGTGCGCGAGGTTCTGGACAAGGGCGACAAGGACTCCCTGGAGCAGTTGATTGAGG
AGGCGTACCAGCTGATTCAAAAGTTGCGCGAGCTGTTCGAGAAAAAGAACT*AAGCTT*
3'
SCRPZ-3:
5'*CATATG*GAACTGGAAAAACTGCGTCAGACTGGTGAACAGATCCTGCAAATTGCGA
AACAGGTTAACGAGATTATGCTGAAGGGTGATGACGACTCTCTGGAGCAGCTGATT
AAACTGGCGTACGAACTGATCCAACAACACACCCAGCTGGCGTACAATCGTCAGGA
AGCTGCGGATACGGAGATCATGAAACAGGGTCAGCAGATTCTGGAAATTGCCCAGC
AGGTGAACGAAGTGCTGCTGAAAGGCGACAAAGACTCTCTGGAACAGCTGTTAAAG
CTGGCGTATCAGCTGATTCAGCAACTGCAGGAGCTGTTCGAAAAGAAAAACTAA*AA
GCTT* 3'

Synthesis of $[RuPZn](PF_6)_2$

The synthesis of the compound followed previous literature methods[13] but used the unsubstituted 5,15-diphenylporphyrin. For clarity, the reaction scheme is provided in the supporting information. ^{1}H-NMR (d_6-DMSO, 500 MHz) 10.2 (s, 1), 10.01 (d, 2), 9.35 (d, 2), 9.30 (d, 2), 9.04 (d, 2), 8.89 (d, 2), 8.77 (d, 2), 8.73 (d, 2), 8.51 (d, 2), 8.45 (t, 1), 8.25 (d, 4), 7.97 (t, 2), 7.93 (t, 2), 7.85 (m, 6), 7.56 (d, 2), 7.41 (d, 2), 7.22 (t, 4). UV/Visible (THF): λ_{max} = 428, 506, 637 nm MALDI-TOF MS: *m/z* = 1258. 46 (M^+, -PF_6) (calculated *m/z* for $C_{64}H_{40}F_6N_{10}PRuZn$: 1259.14) and 1114.40 ($M^+$, -2$PF_6$) (calculated *m/z* for $C_{64}H_{40}N_{10}RuZn$: 1114.18).



Scheme S1. Synthesis of the ***RuPZn*** chromophore.