

Supplemental Information

A Re-wired Green Fluorescent Protein: Folding and Function in a Non-sequential, Non-circular GFP Permutant

Philippa J. Reeder¹, Yao-Ming Huang², Jonathan S. Dordick^{2,3*}, and Christopher Bystroff^{2*},

¹Department of Chemical and Biological Engineering, University of Colorado, Boulder, 80309, ²Department of Biology and ³Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180

*To whom correspondence should be addressed (bystrc@rpi.edu; dordick@rpi.edu)

Table 1S. Sequences of constructs. New loops are highlighted with bold text.

GFP-OPT

MGSSHHHHHHSSGLVPRGSHM

KGEELFTGVVPILEVELDGDVNGHKFSVRGEGEGDATIGKLTCLKFICTTGKLPVPWPTLVTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRATVKFVGDTLVNR IELKGTDFKEDGNILGHKLEYNFNSHNVIITADKQKNGIKANFTVRHNVEDGQSVQLADHY QQNTPIGDGPVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITHGMDELYK

rGFP1

MGSSHHHHHHSSGLVPRGSIHM

GHKLEYNFNSHNVIITADKQKNGIKANFTVRHNIEDGQSVQLADHYQQNT**FEGDTLV**QERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGN**IMPGDGP**WPTLVTTLYGVQCFSRYPDHMKRHDFFKSA**EKG**VLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITHGMDELYK**GGTGG**SMASKGEELFTGVVPILEVELDGDVNGHKFSVRGEGEGDATIGKLTCLKFICTTGKL

rGFP2

MGSSHHHHHHSSGLVPRGSIHM

GGASGSKGEELFTGVVPILEVELDGDVNGHKFSVSGEEGEGDATYKLTCLKFICTTGKLPVPWPTLVTTLYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKIDFKEDGNILGHKLEYNYNSHNVI**IDGG**VYLSTQSALSKDPNEKRDHMLLEFVTAAGIT**GGDGP**KLV**PDS**GIKANFKIRHNIEDGQSVQLADHYQQNTPIGDAGGS

rGFP3

MGSSHHHHHHSSGLVPRGSIHM

QKNGIKANFTVRHNIEDGSVQLADHYQQNTPIGDG**SGTGS**GKGEELFTGVVPILEVELDGD
VNGHKFSVRGEGEGDATIGKLTLLKFICTTGKLPVPWPTLVTTTLTYGVQCFSRYPDHMKRH
DFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKL
EYNFNHSHNVYITAD**GGSGGT**GVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH

rGFP3 CP map. CP name is indicated by the numbers in brackets. Crossed out sequences are removed in the CP construct closest to it.

MGSSHHHHHHSSGLVPRGSIHM

QKNGIKANFTVRHNI ~~ED~~ (9-8) GSVQLADHYQQNTPIGDG **SGTGS** (1-9)
KGEELFTGVVPILEVELDGDVN (2-1) GHKFSVRGEGEGDAT (3-2)
IGKLTLLKFICTTGK (A-3) LPVPWPTLVTTTLTYGVQCFSRYPDHMKRH
DFFKSAM (4-A) ~~P~~ EGYVQERTISFKD (5-4) DGKYKTRAVVKFEGD (6-5)
TLVNRIELKGTDFK (L-6) EDGNILGHKLE ~~Y~~ (7-L) NFNHSHNVYITAI
~~DGGSGGT~~ (10-7) GVLLPDNHYLSTQTVLSKDPN (11-10)
EKRDHMLLEFVTAAGITH **GGTGS**

GGTGS is the loop introduced for circular permutation

rGFP3 CP 9-8

MGSSHHHHHHSSGLVPRGSIHM

GSVQLADHYQQNTPIGDGSGTGSKGGEELFTGVVPILEVELDGDVNGHKFSVRGEGEGDAT
IGKLTLLKFICTTGKLPVPWPTLVTTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTI
SFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNHSHNVYITAI
DGGSGGTGVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTV
RHNI

rGFP3 CP 1-9

MGSSHHHHHHSSGLVPRGSIHM

KGEELFTGVVPILEVELDGDVNGHKFSVRGEGEGDATIGKLTLLKFICTTGKLPVPWPTLV
TLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVN
RIELKGTDFKEDGNILGHKLEYNFNHSHNVYITAI
DGGSGGTGVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTVRHNI
EDGSVQLADHYQQNTPIGDG

rGFP3 CP 2-1

MGSSHHHHHHSSGLVPRGSIHM

GHKFSVRGEGEGDATIGKLTLLKFICTTGKLPVPWPTLVTTTLTYGVQCFSRYPDHMKRH
DFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKL
EYNFNHSHNVYITAI
DGGSGGTGVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTVRHNI
EDGSVQLADHYQQNTPIGDGSGTGSKGGEELFTGVVPILEVELDGDVN

rGFP3 CP 3-2

MGSSHHHHHHSSGLVPRGSIHM

IGKLTLLKFICTTGKLPVPWPTLVTTLLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTI
SFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITAIDGG
SGGTGVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTV
RHNIEDGSVQLADHYQQNTPIGDGSGTGSGKGEELFTGVVPILEVELDGDVNGHKFSVRGE
GEGDAT

rGFP3 CP A-3

MGSSHHHHHHSSGLVPRGSIHM

LPVPWPTLVTTLLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTISFKDDGKYKTRAVV
KFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITAIDGGSGGTGVLLPDNHYL
STQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTVRHNIEDGSVQLADH
YQQNTPIGDGSGTGSGKGEELFTGVVPILEVELDGDVNGHKFSVRGEGEGDATIGKLTLLK
FICTTGK

rGFP3 CP 4-A

MGSSHHHHHHSSGLVPRGSIHM

EGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHN
VYITAIDGGSGGTGVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QK
NGIKANFTVRHNIEDGSVQLADHYQQNTPIGDGSGTGSGKGEELFTGVVPILEVELDGDVN
GHKFSVRGEGEGDATIGKLTLLKFICTTGKLPVPWPTLVTTLLTYGVQCFSRYPDHMKRHDF
FKSAM

rGFP3 CP 5-4

MGSSHHHHHHSSGLVPRGSIHM

DGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITAIDGGSGGT
GVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTVRHNI
EDGSVQLADHYQQNTPIGDGSGTGSGKGEELFTGVVPILEVELDGDVNGHKFSVRGEGEGD
ATIGKLTLLKFICTTGKLPVPWPTLVTTLLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQER
TISFKD

rGFP3 CP 6-5

MGSSHHHHHHSSGLVPRGSIHM

TLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITAIDGGSGGTGVLLPDNHYLSTQTV
LSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTVRHNIEDGSVQLADHYQQNT
PIGDGSGTGSGKGEELFTGVVPILEVELDGDVNGHKFSVRGEGEGDATIGKLTLLKFICTTG
KLPVPWPTLVTTLLTYGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTISFKDDGKYKTRAV
VKFEGD

rGFP3 CP L-6

MGSSHHHHHHSSGLVPRGSIHM

EDGNILGHKLEYNFNSHNVYITAIDGGSGGTGVLLPDNHYLSTQTVLSKDPNEKRDHML
LEFVTAAGITH**GGTGS**QKNGIKANFTVRHNIEDGSVQLADHYQQNTPIGDGSGTGSGKGE
ELFTGVVPILEVELDGDVNGHKFSVRGEGEGDATIGKLTLLKFICTTGKLPVPWPTLVTTLL
YGVQCFSRYPDHMKRHDFFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIEL

KGTFD

rGFP3 CP 7-L

MGSSHHHHHHSSGLVPRGSIHM

NFNShNVYITAI DGGSGGTGVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH
GGTGSQKNGIKANFTVRHNI EDGSVQLADHYQQNTPIGDGSGTGSGKGEELFTGVVPI LV
ELDGDVNGHKFSVRGEGEGDATIGKLT LKFICTTGKLPVPWPTLVTTLT YGVQCFSRYPD
HMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNI
LGHKLE

rGFP3 CP 10-7

MGSSHHHHHHSSGLVPRGSIHM

GVLLPDNHYLSTQTVLSKDPNEKRDHMLLEFVTAAGITH**GGTGS**QKNGIKANFTVRHNI
EDGSVQLADHYQQNTPIGDGSGTGSGKGEELFTGVVPI LV EL DGDVNGHKFSVRGEGEGD
ATIGKLT LKFICTTGKLPVPWPTLVTTLT YGVQCFSRYPDHMKRHDFFKSAMPEGYVQER
TISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITAI

rGFP3 CP 11-10

MGSSHHHHHHSSGLVPRGSIHM

EKRDMVLEFVTAAGITHGGTGSQKNGIKANFTVRHNI EDGSVQLADHYQQNTPIGDG
GTGSGKGEELFTGVVPI LV EL DGDVNGHKFSVRGEGEGDATIGKLT LKFICTTGKLPVP
PTLVTTLT YGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDGKYKTRAVVKFEGDT
LVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITAI D**GGSGGT**GVLLPDNHYLSTQTVL
SKDPN