**Additional file 3**

**Table S1**. Sequencing results of selected variants from the biopanning of the T7ZfQQR expressing the Zfm2 library on a panel of substrates. The frequency logo was generated using all extracted amino acid residue sequences, where first, second, third and fourth column is the frequency obtained for randomized positions Q56, S58, N59 and K62, respectively.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sample number** | **Nucleotide sequence used for selection of second module** | **Number of reads passing filters** | **Frequency logo of amino acid residues from randomized region** |
| 1 | GGG | 44832 | https://lh3.googleusercontent.com/h8PXjNdGzXQiLB90BOX6ZaCjPS1lVwlt2ju89SAp8KGdQRetqHRpy0oFzDsv8KShmhMSxEiG6zSrrNxwzGe67Ycc0FCP7pdCYFOJhNsPvLTZ_8Thom9vkgGdmWJqLU_4jh_02tUv |
| 2 | GGA | 21417 | https://lh5.googleusercontent.com/Ms94NB6PqFTPk85mtvSb_57-x6cLxJmO54hoF-GKofi-AJPFGiimREBAbhIUaaLbMZQ9A7BfqLuhJg-ZjlUpCH-fMuaTKxYEii8Bj87v9nMYgb_oNJoMdQYUV02WYfLBg3l4n1q2 |
| 3 | GGT | 163949 | https://lh5.googleusercontent.com/Zn87zr3zK7YM2xULFBbJSTlsRsIUcLE4RalG8EjFMtwu-WZeXX-1U3a9raK359Wz58-LwimYLqVDp7i6cAQa-4pd26PcOV_YkA1fAuCtsp3DNNhR_8V6f8eNk5udMGXTxcsgGCm8 |
| 4 | GGC | 111596 | https://lh6.googleusercontent.com/W4wsp_jNnjRn1fd4408HW4P4ggynBzyVgs5BHI_XEoCSgx-26s0bXnIUV437AD5J8HkdSwYdGbsb2ikFqECZs1kdEAB0UqP4BVW0wp-CCFUyDSd3ckn8wALF2_MoH3mO814DOc0V |
| 5 | GAG | 149445 | https://lh4.googleusercontent.com/Ez3DGzQmnEhnLQmYCOk4gOZnATx-dRlFX4QlQxsD1s-Vt42l51QabqQlDg6JfHB6IpF_RX2qH7FLBasjWpRWSBULeQS0d39VlLkcid-I11nVVI1Ke7O_eTbDnHxa4HRJFcegrDfF |
| 6 | GTG | 87366 | https://lh3.googleusercontent.com/Y_VUdNp3xjR5_3XM3URYqIK-FU7KGC_EVVxqobUS7aw2Y62TWnm3scSPabSJADqNpJY-n9z2gDkEyNJUwY3nvb1Jo7VTvYps87fEASqoifUTso_s1vnLGdusLgUYO5gUsUPrPrPs |
| 7 | GCG | 63335 | https://lh6.googleusercontent.com/49p0LZ5fMyN2ehzNMzPhQSS8yxpdCNGveMDBxr-rH2eVLqjmGiEkOZG7V1t9Rffn_M4yjmhRp8pV-WjGZk099uBQBoMyfQ_kHW8jrNhBULj6tvWSzMu9AZU2u5Hd2LV8mdf1jA-l |
| 8 | GAA | 60898 | https://lh4.googleusercontent.com/fzTNJ1zhyvaFNlQll-AS2z1zxP6FoGgE7S7B8G-i0aUuel_5kNOrq6G_lMBkcG6QIWwcXKgwedIeQyuVEajh5_GRcqJvcQY-9cBWMz5WmLXVbjfWlArprIfRZIk2PzRD-Z7gLW6l |
| 9 | GTA | 53852 | https://lh4.googleusercontent.com/FunvKc_wlwoxbD1_jNjeBHx9uWIAvlL3vDwiKW6TExS-8OmKPqnsw2oiGrOaEeHp93shIVX_REmcU5Z6HOBVXm0tp9E8N6z7a_KFTSfU5PPXCM9iqZU-bXX-Lq9hXsqqSK1uSv2F |
| 10 | GCA | 50223 | https://lh5.googleusercontent.com/vftTEnQAGZJ0RBRP3nHqO_7E2ga09K5J7fnSwzh0Z7OyiNTbdX5KpWMnI7z_WcKrBFW3GjsfpNH_0Q8K2aTtHN3vEsWhYaoir7lyiLkIFmt5pl7zyhuIDN-WG22e2n9KYdJXOLW_ |
| 11 | GAT | 53432 | https://lh3.googleusercontent.com/O948SwYCKTC7gPVU5zjnBm-bWzxGuQo9OefI9e5xZTusM9C3772vwqHK6sNJ5h9VHeK6zNagRaf_5vsuFV_FbmmUdyKxYtjfD22VPLFm7fUDzoEry5sMYaJMQ5O2laWFsSgPw2Jb |
| 12 | GTT | 50531 | https://lh4.googleusercontent.com/7o0x8KlV-kXXg3QUVunVL-v8N9IFKLg_99UXGFQkMMf43MU8AqjO_TyezRrZqOITMB4YvNHSLunljW8By5pZpgZ47n_-dwaS26nnfOfk25Pjo4d8WyJX0a_X4PiFeaTtJOG3cenL |
| 13 | GCT | 54399 |  |
| 14 | GAC | 22991 |  |
| 15 | GTC | 44998 |  |
| 16 | GCC | 25933 |  |
| 17 | AAG | 39927 |  |
| 18 | AAA | 56184 |  |
| 19 | AAT | 68649 |  |
| 20 | AAC | 41015 |  |
| 21 | AGG | 46739 |  |
| 22 | ATG | 72661 |  |
| 23 | ACG | 38390 |  |
| 24 | AGA | 34345 |  |
| 25 | ATA | 52693 |  |
| 26 | ACA | 63623 |  |
| 27 | AGT | 47149 |  |
| 28 | ATT | 68079 |  |
| 29 | ACT | 59210 |  |
| 30 | AGC | 78613 |  |
| 31 | ATC | 55096 |  |
| 32 | ACC | 45545 |  |
| 33 | TTG | 41691 |  |
| 34 | TTA | 43464 |  |
| 35 | TTT | 69425 |  |
| 36 | TTC | 46208 |  |
| 37 | TGG | 152133 |  |
| 38 | TAG | 55822 |  |
| 39 | TCG | 40287 |  |
| 40 | TGA | 100686 |  |
| 41 | TAA | 87310 |  |
| 42 | TCA | 38444 |  |
| 43 | TGT | 100767 |  |
| 44 | TAT | 74930 |  |
| 45 | TCT | 59807 |  |
| 46 | TGC | 76625 |  |
| 47 | TAC | 63637 |  |
| 48 | TCC | 32919 |  |
| 49 | CCG | 105741 |  |
| 50 | CCA | 144997 |  |
| 51 | CCT | 89487 |  |
| 52 | CCC | 62703 |  |
| 53 | CGG | 65713 |  |
| 54 | CAG | 46101 |  |
| 55 | CTG | 233659 |  |
| 56 | CGA | 159945 |  |
| 57 | CAA | 146246 |  |
| 58 | CTA | 63761 |  |
| 59 | CGT | 52372 |  |
| 60 | CAT | 119039 |  |
| 61 | CTT | 42718 |  |
| 62 | CGC | 44639 |  |
| 63 | CAC | 39525 |  |
| 64 | CTC | 44527 |  |

**Table S2**. Sequencing results of selected variants from the biopanning of the T7ZfQQR expressing the L5 and L6 library. The frequency logo was generated using all extracted amino acid residue sequences, where consecutive columns represent the frequencies obtained for randomized positions in the linkers.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sample name** | **Nucleotide sequence used for selection** | **Number of reads passing filters** | **Frequency logo of amino acid residues from randomized region** |
| L5 | GGGGAAGAA | 41490 | See Figure 2B |
| L6 | GGGGAAGAA | 27850 | See Figure 2C |