Electronic Supplementary Information for

Evolution-guided engineering of nonribosomal peptide synthetase adenylation domains

Max Crüsemann,[a] Christoph Kohlhaas,[a] and Jörn Piel*[a,b]

LEGENDS

Figure S1: Alignment of the A domains of HrmO1-P3 and CdaPSI-A5/H6. Black bars indicate the selected fusion sites.

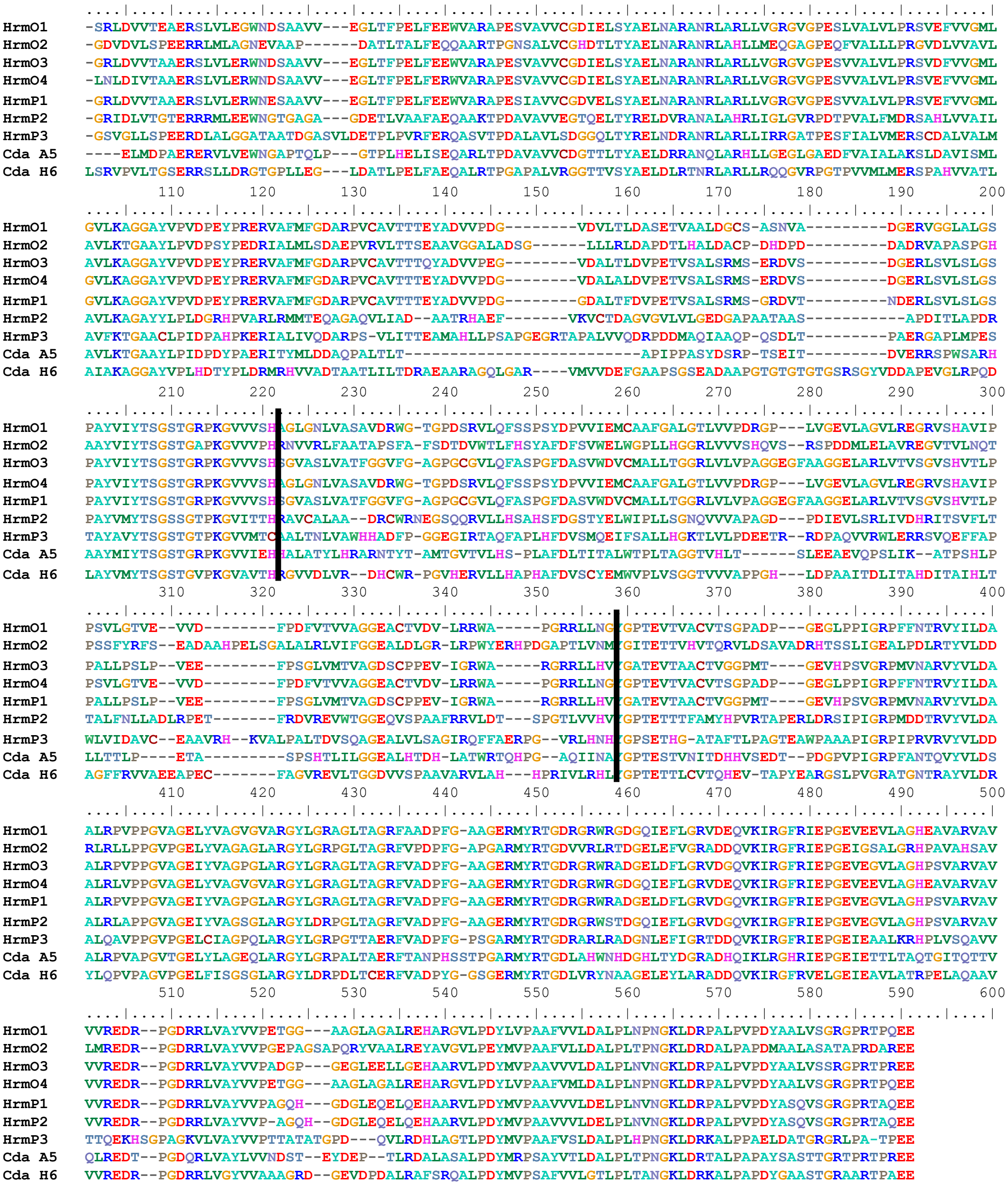
Figure S2: SDS-PAGE analysis (15%, Tris/Tricine) of the expressed and Ni-purified A domain/MbtH complexes HrmO1$_A$, HrmO3$_A$, HrmO4$_A$, HrmP1$_A$, HrmP2$_A$, O2f$_A$, O4f$_A$, P2f$_A$, A5f$_A$, H6f$_A$, coeluted with HrmR. A5f$_x$ shows a coelution of A5f$_A$ with CdaX. M=Marker.

Figure S3: SDS-PAGE analysis (15%, Tris/Tricine) of the FPLC-purified A domain/HrmR complexes HrmO1$_A$, HrmO2$_A$, HrmO3$_A$, O2f$_A$, HrmP1$_A$, HrmP3$_A$. M=Marker.

Figure S4: Results of the $\gamma^{18}O_4$-ATP assay for **A**, HrmO1$_A$; **B**, HrmO2$_A$; **C**, HrmO3$_A$; **D**, HrmO4$_A$; **E**, HrmP1$_A$; **F**, HrmP2$_A$ (black); **G**, HrmP3$_A$. The preferred substrate is equal 100%. X-axis: % relative activation.

Table S1: Sequences of the primers used in this study. Inserted restriction sites are underlined, complementary parts to *hrmO3* are in italics.

Table S2: Obtained C- and TM-scores generated by I-TASSER for HrmO3$_A$ and the engineered A domains. The quality parameters of the created models are in the same range as HrmO3$_A$, indicating sufficient quality of the models.



```
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HrmO1     -SRLDVVTEAERSLVLEGWNDSAAVV---EGLTFPELFEEWVARAPESVAVVCGDIELSYAELNARANRLARLLVGRGVGPESLVALVLPRSVEFVVGML
HrmO2     -GDVDVLSPEERRLMLAGNEVAAP------DATLTALFEQQAARTPGNSALVCGHDTLTYAELNARANRLAHLLMEQGAGPEQFVALLLPRGVDLVVAVL
HrmO3     -GRLDVVTAAERSLVLERWNDSAAVV---EGLTFPELFEEWVARAPESVAVVCGDIELSYAELNARANRLARLLVGRGVGPESVVALVLPRSVDFVVGML
HrmO4     -LNLDIVTAAERSLVLERWNDSAAVV---EGLTFPELFERWVARAPESVAVVCGDIELSYAELNARANRLARLLVGRGVGPESVVALVLPRSVEFVVGML
HrmP1     -GRLDVVTAAERSLVLERWNESAAVV---EGLTFPELFEEWVARAPESIAVVCGDVELSYAELNARANRLARLLVGRGVGPESVVALVLPRSVEFVVGML
HrmP2     -GRIDLVTGTERRRMLEEWNGTGAGA---GDETLVAAFAEQAAKTPDAVAVVEGTQELTYRELDVRANALAHRLIGLGVRPDTPVALFMDRSAHLVVAIL
HrmP3     -GSVGLLSPEERDLALGGATAATDGASVLDETPLPVRFERQASVTPDALAVLSDGGQLTYRELNDRANRLARLLIRRGATPESFIALVMERSCDALVALM
Cda A5    ----ELMDPAERERVLVEWNGAPTQLP---GTPLHELISEQARLTPDAVAVVCDGTTLTYAELDRRANQLARHLLGEGLGAEDFVAIALAKSLDAVISML
Cda H6    LSRVPVLTGSERRSLLDRGTGPLLEG---LDATLPELFAEQALRTPGAPALVRGGTTVSYAELDLRTNRLARLLRQQGVRPGTPVVMLMERSPAHVVATL
                 110       120       130       140       150       160       170       180       190       200
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HrmO1     GVLKAGGAYVPVDPEYPRERVAFMFGDARPVCAVTTTEYADVVPDG--------VDVLTLDASETVAALDGCS-ASNVA---------DGERVGGLALGS
HrmO2     AVLKTGAAYLPVDPSYPEDRIALMLSDAEPVRVLTTSEAAVGGALADSG------LLLRLDAPDTLHALDACP-DHDPD---------DADRVAPASPGH
HrmO3     AVLKAGGAYVPVDPEYPRERVAFMFGDARPVCAVTTTQYADVVPEG--------VDALTLDVPETVSALSRMS-ERDVS---------DGERLSVLSLGS
HrmO4     GVLKAGGAYVPVDPEYPRERVAFMFGDARPVCAVTTTEYADVVPDG--------VDALALDVPETVSALSRMS-ERDVS---------DGERLSVLSLGS
HrmP1     GVLKAGGAYVPVDPEYPRERVAFMFGDARPVCAVTTTEYADVVPDG--------GDALTFDVPETVSALSRMS-GRDVT---------NDERLSVLSLGS
HrmP2     AVLKAGAYYLPLDGRHPVARLRMMTEQAGAQVLIAD--AATRHAEF-------VKVCTDAGVGVLVLGEDGAPAATAAS-----------APDITLAPDR
HrmP3     AVFKTGAACLPIDPAHPKERIALIVQDARPS-VLITTEAMAHLLPSAPGEGRTAPALVVQDRPDDMAQIAAQP-QSDLT---------PAERGAPLMPES
Cda A5    AVLKTGAAYLPIDPDYPAERITYMLDDAQPALTLT-------------------------APIPPASYDSRP-TSEIT---------DVERRSPWSARH
Cda H6    AIAKAGGAYVPLHDTYPLDRMRHVVADTAATLILTDRAEAARAGQLGAR----VMVVDEFGAAPSGSEADAAPGTGTGTGTGSRSGYVDDAPEVGLRPQD
                 210       220       230       240       250       260       270       280       290       300
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HrmO1     PAYVIYTSGSTGRPKGVVVSHAGLGNLVASAVDRWG-TGPDSRVLQFSSPSYDPVVIEMCAAFGALGTLVVPDRGP---LVGEVLAGVLREGRVSHAVIP
HrmO2     AAYVIYTSGSTGAPKGVVVPHRNVVRLFAATAPSFA-FSDTDVWTLFHSYAFDFSVWELWGPLLHGGRLVVVSHQVS--RSPDDMLELAVREGVTVLNQT
HrmO3     PAYVIYTSGSTGRPKGVVVSHSGVASLVATFGGVFG-AGPGCGVLQFASPGFDASVWDVCMALLTGGRLVLVPAGGEGFAAGGELARLVTVSGVSHVTLP
HrmO4     PAYVIYTSGSTGRPKGVVVSHAGLGNLVASAVDRWG-TGPDSRVLQFSSPSYDPVVIEMCAAFGALGTLVVPDRGP---LVGEVLAGVLREGRVSHAVIP
HrmP1     PAYVIYTSGSTGRPKGVVVSHSGVASLVATFGGVFG-AGPGCGVLQFASPGFDASVWDVCMALLTGGRLVLVPAGGEGFAAGGELARLVTVSGVSHVTLP
HrmP2     PAYVMYTSGSSGTPKGVITTHRAVCALAA--DRCWRNEGSQQRVLLHSAHSFDGSTYELWIPLLSGNQVVVAPAGD---PDIEVLSRLIVDHRITSVFLT
HrmP3     TAYAVYTSGSTGTPKGVVMTCAALTNLVAWHHADFP-GGEGIRTAQFAPLHFDVSMQEIFSALLHGKTLVLPDEETR--RDPAQVVRWLERRSVQEFFAP
Cda A5    AAYMIYTSGSTGRPKGVVIEHHALATYLHRARNTYT-AMTGVTVLHS-PLAFDLTITALWTPLTAGGTVHLT------SLEEAEVQPSLIK--ATPSHLP
Cda H6    LAYVMYTSGSTGVPKGVAVTHRGVVDLVR--DHCWR-PGVHERVLLHAPHAFDVSCYEMWVPLVSGGTVVVAPPGH---LDPAAITDLITAHDITAIHLT
                 310       320       330       340       350       360       370       380       390       400
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HrmO1     PSVLGTVE--VVD--------FPDFVTVVAGGEACTVDV-LRRWA-----PGRRLLNGYGPTEVTVACVTSGPADP---GEGLPPIGRPFFNTRVYILDA
HrmO2     PSSFYRFS--EADAAHPELSGALALRLVIFGGEALDLGR-LRPWYERHPDGAPTLVNMYGITETTVHVTQRVLDSAVADRHTSSLIGEALPDLRTYVLDD
HrmO3     PALLPSLP--VEE--------FPSGLVMTVAGDSCPPEV-IGRWA-----RGRRLLHVYGATEVTAACTVGGPMT----GEVHPSVGRPMVNARVYVLDA
HrmO4     PSVLGTVE--VVD--------FPDFVTVVAGGEACTVDV-LRRWA-----PGRRLLNGYGPTEVTVACVTSGPADP---GEGLPPIGRPFFNTRVYILDA
HrmP1     PALLPSLP--VEE--------FPSGLVMTVAGDSCPPEV-IGRWA-----RGRRLLHVYGATEVTAACTVGGPMT----GEVHPSVGRPMVNARVYVLDA
HrmP2     TALFNLLADLRPET-------FRDVREVWTGGEQVSPAAFRRVLDT---SPGTLVVHVYGPTETTTFAMYHPVRTAPERLDRSIPIGRPMDDTRVYVLDA
HrmP3     WLVIDAVC--EAAVRH--KVALPALTDVSQAGEALVLSAGIRQFFAERPG--VRLHNHYGPSETHG-ATAFTLPAGTEAWPAAAPIGRPIPRVRVYVLDD
Cda A5    LLTTLP----ETA---------SPSHTLILGGEALHTDH-LATWRTQHPG--AQIINAYGPTESTVNITDHHVSEDT--PDGPVPIGRPFANTQVYVLDS
Cda H6    AGFFRVVAEEAPEC-------FAGVREVLTGGDVVSPAAVARVLAH---HPRIVLRHLYGPTETTLCVTQHEV-TAPYEARGSLPVGRATGNTRAYVLDR
                 410       420       430       440       450       460       470       480       490       500
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HrmO1     ALRPVPPGVAGELYVAGVGVARGYLGRAGLTAGRFAADPFG-AAGERMYRTGDRGRWRGDGQIEFLGRVDEQVKIRGFRIEPGEVEEVLAGHEAVARVAV
HrmO2     RLRLLPPGVPGELYVAGAGLARGYLGRPGLTAGRFVPDPFG-APGARMYRTGDVVRLRTDGELEFVGRADDQVKIRGFRIEPGEIGSALGRHPAVAHSAV
HrmO3     ALRPVPPGVAGEIYVAGPGLARGYLGRAGLTAGRFVADPFG-AAGERMYRTGDRGRWRADGELDFLGRVDGQVKIRGFRIEPGEVEGVLAGHPSVARVAV
HrmO4     ALRLVPPGVAGELYVAGVGVARGYLGRAGLTAGRFVADPFG-AAGERMYRTGDRGRWRGDGQIEFLGRVDEQVKIRGFRIEPGEVEEVLAGHEAVARVAV
HrmP1     ALRPVPPGVAGEIYVAGPGLARGYLGRAGLTAGRFVADPFG-AAGERMYRTGDRGRWRADGELDFLGRVDGQVKIRGFRIEPGEVEGVLAGHPSVARVAV
HrmP2     ALRLAPPGVAGEIYVAGSGLARGYLDRPGLTAGRFVADPFG-AAGERMYRTGDRGRWSTDGQIEFLGRVDGQVKIRGFRIEPGEVEGVLAGHPSVARVAV
HrmP3     ALQAVPPGVPGELCIAGPQLARGYLGRPGTTAERFVADPFG-PSGARMYRTGDRARLRADGNLEFIGRTDDQVKIRGFRIEPGEIEAALKRHPLVSQAVV
Cda A5    ALRPVAPGVTGELYLAGEQLARGYLGRPALTAERFTANPHSSTPGARMYRTGDLAHWNHDGHLTYDGRADHQIKLRGHRIEPGEIETTLTAQTGITQTTV
Cda H6    YLQPVPAGVPGELFISGSGLARGYLDRPDLTCERFVADPYG-GSGERMYRTGDLVRYNAAGELEYLARADDQVKIRGFRVELGEIEAVLATRPELAQAAV
                 510       520       530       540       550       560       570       580       590       600
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HrmO1     VVREDR--PGDRRLVAYVVPETGG---AAGLAGALREHARGVLPDYLVPAAFVVLDALPLNPNGKLDRPALPVPDYAALVSGRGPRTPQEE
HrmO2     LMREDR--PGDRRLVAYVVPGEPAGSAPQRYVAALREYAVGVLPEYMVPAAFVLLDALPLTPNGKLDRDALPAPDMAALASATAPRDAREE
HrmO3     VVREDR--PGDRRLVAYVVPADGP---GEGLEELLGEHAARVLPDYMVPAAVVVLDALPLNVNGKLDRPALPVPDYAALVSSRGPRTPREE
HrmO4     VVREDR--PGDRRLVAYVVPETGG---AAGLAGALREHARGVLPDYLVPAAFVMLDALPLNPNGKLDRPALPVPDYAALVSGRGPRTPQEE
HrmP1     VVREDR--PGDRRLVAYVVPAGQH---GDGLEQELQEHAARVLPDYMVPAAVVVLDELPLNVNGKLDRPALPVPDYASQVSGRGPRTAQEE
HrmP2     VVREDR--PGDRRLVAYVVP-AGQH--GDGLEQELQEHAARVLPDYMVPAAVVVLDELPLNVNGKLDRPALPVPDYASQVSGRGPRTAQEE
HrmP3     TTQEKHSGPAGKVLVAYVVPTTATATGPD---QVLRDHLAGTLPDYMVPAAFVSLDALPLHPNGKLDRKALPPAELDATGRGRLPA-TPEE
Cda A5    QLREDT--PGDQRLVAYLVVNDST--EYDEP--TLRDALASALPDYMRPSAYVTLDALPLTPNGKLDRTALPAPAYSASTTGRTPRTPREE
Cda H6    VVREDR--PGDRRLVGYVVAAAGRD--GEVDPDALRAFSRQALPDYMVPSAFVVLGTLPLTANGKLDRKALPAPDYGAASTGRAARTPAEE
```

Figure S1: Alignment of the A domains of HrmO1-P3 and CdaPSI-A5/H6. Black bars indicate the selected fusion sites.

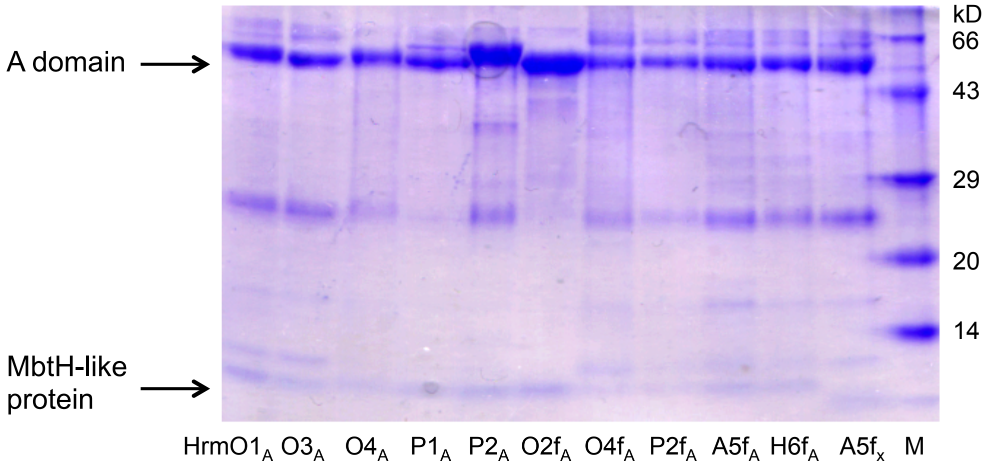


Figure S2: SDS-PAGE analysis (15%, Tris/Tricine) of the expressed and Ni-NTA affinity-purified A domain/MbtH complexes $HrmO1_A$, $HrmO3_A$, $HrmO4_A$, $HrmP1_A$, $HrmP2_A$, $O2f_A$, $O4f_A$, $P2f_A$, $A5f_A$, $H6f_A$, coeluted with HrmR. $A5f_x$ shows a coelution of $A5f_A$ with CdaX. M=Marker

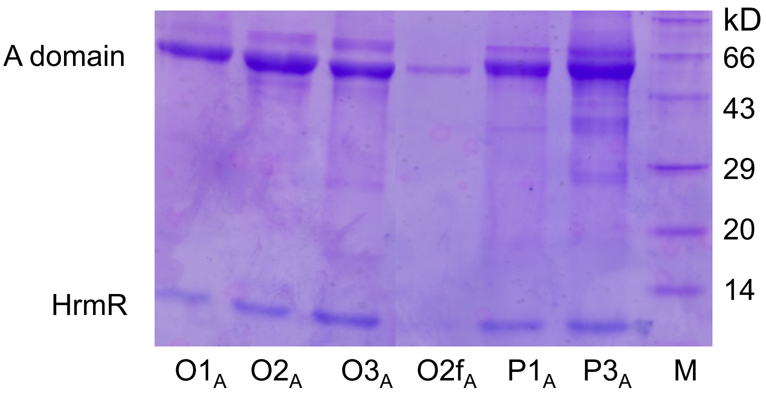


Figure S3: SDS-PAGE analysis (15%, Tris/Tricine) of the FPLC-purified A domain/HrmR complexes HrmO1$_A$, HrmO2$_A$, HrmO3$_A$, O2f$_A$, HrmP1$_A$, HrmP3$_A$. M=Marker

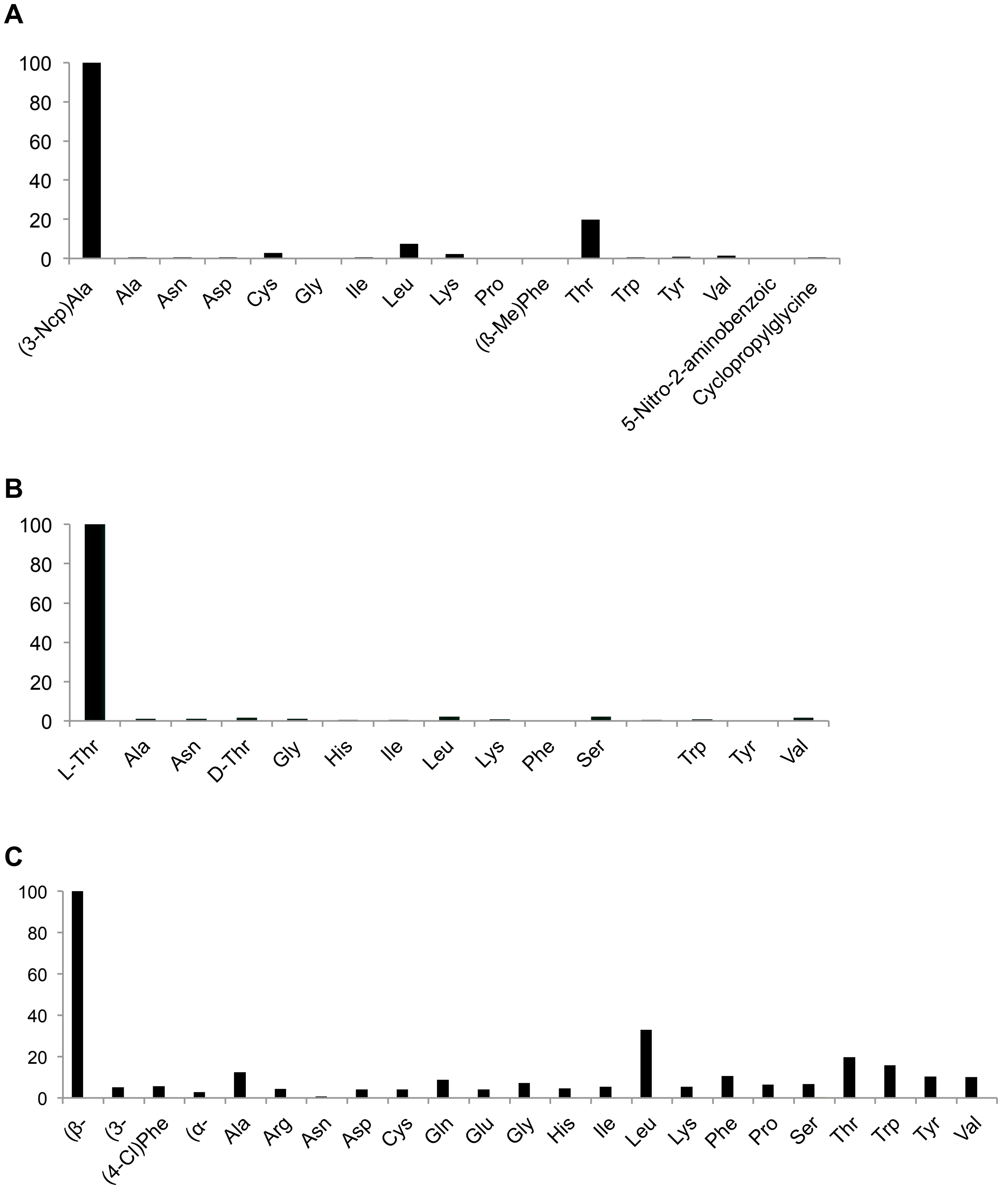


Figure S4: Results of the $\gamma^{18}O_4$-ATP assay for **A**, HrmO1$_A$; **B**, HrmO2$_A$; **C**, HrmO3$_A$; **D**, HrmO4$_A$; **E**, HrmP1$_A$; **F**, HrmP2$_A$ (black); **G**, HrmP3$_A$. The preferred substrate is equal 100%. Y-axis: % relative activation.

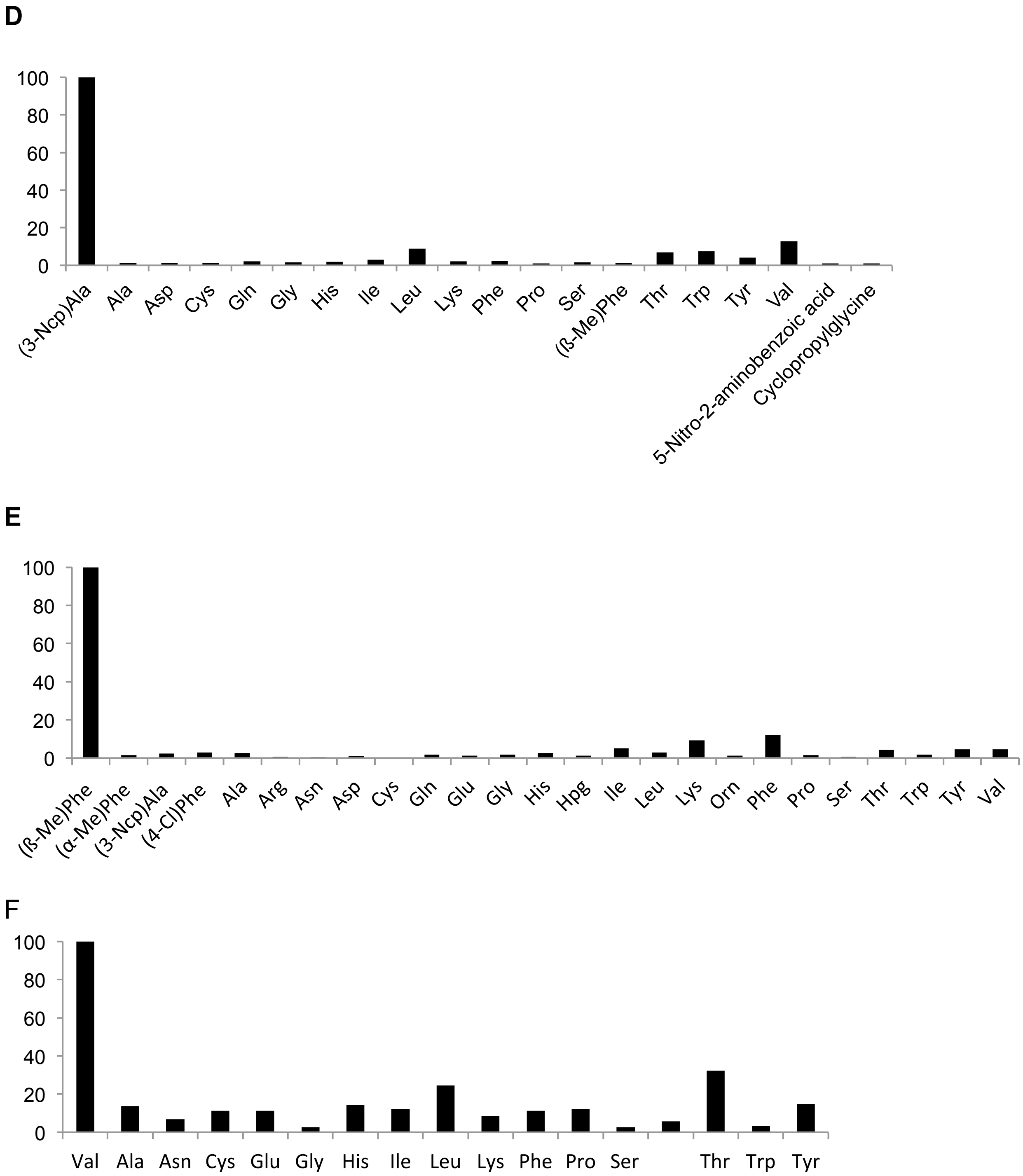


Figure S4, continued.

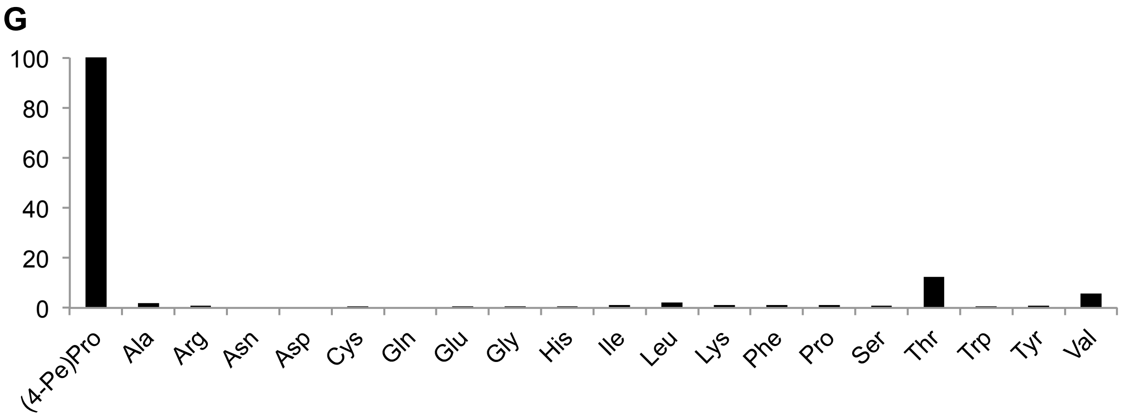


Figure S4, continued.

Table S1: Sequences of the primers used in this study. Inserted restriction sites are underlined, complementary parts to *hrmO3* are in italics. Pairs O1F/O1R to P3F/P3R were used to amplify DNA regions encoding A domains $O1_A$ to $P3_A$.

O1F: AAAGAATTCAGCCGGCTGGACGTGGTTA

O1R: AAAAAGCTTTCACTCCTCCTGCGGCGTACG

O2F: AAAGAATTCGGCGACGTCGACGTCCTG

O2R: AAAAAGCTTTCATTCCTCGCGGGCGTCGCG

O3F: AAAGAATTCGGCCGGCTGGACGTCGTC

O3R: AAAAAGCTTTCACTCCTCCCGCGGTGTGCG

O4F: AAAGAATTCCTGAATCTTGACATCGTCACTG

O4R: AAAAAGCTTTCACTCCTCTGCGGCGTACG

P1F: AAAGAATTCGGCCGGCTGGACGTCGT

P1R: AAAAAGCTTTCACTCCTCCTGCGCGGTACGCG

P2F: AAACAATTGGGCCGCATCGACCTGGTC

P2R: AAAAAGCTTTCACTCCTCCTGCGCCGGTACG

P3F: AAAGAATTCGGTTCCGTCGGGCTGTTGT

P3R: AAAAAGCTTTCACGCCTCCTCGGGGTGG

cdaXF: AAAAAACATATGAGTAATCCCTTCGACGAC

cdaXR: AAA AAACTCGAGTCCGGTCAGTTGCCGGTG

hrmRF: AAAAAACATATGCGAGATCGAGAAAAGAATGGA

hrmFR: AAAAAAGAGCTCTCATGCCGCGCTCCCTTC

O3_nestedF: CTGGACGTCGTCACGGCG

O3_nestedR: CTCCTCCCGCGGTGTGCG

O3_leftF: AAAGAATTCGCTTTCGCGGCCGATCCG

O3_leftR: ATGCGAAACCACCACACCT

O3_rightF: TACGGGGCGACGGAGGTG

O3_rightR: AAAAAGCTTACACCGACCCGCTCCAGC

O4_insertF: *AAGGTGTGGTGGTTTCGCAT*GCGGGTCTGGGGAATTTG

O4_insertR: *CGTCACCTCCGTCGCCCCGTA*CCCGTTCAGCAACCGCCG

O2_insertF: *AAGGGTGTGGTGGTTTCGCAT*CGGAACGTCGTGCGCCTCT

O2_insertR: *CGTCACCTCCGTCGCCCCGTA*CATGTTGACCAGGGTCGGA

P2_insertF: AAGGGTGTGGTGGTTTCGCATAGGGCCGTCTGCGCGTTG

P2_insertR: *CGTCACCTCCGTCGCCCCGTA*GACGTGCACCACCAGCGT

Asp_insertF: AAGGGTGTGGTGGTTTCGCATCACGCCCTGGCCACCTAC

Asp_insertR: *CGTCACCTCCGTCGCCCCGTA*GGCGTTGATGATCTGGGCG

HPG_insertF: *AAGGGTGTGGTGGTTTCGCAT*CGCGGGGTCGTGGACCTG

HPG_insertR: *CGTCACCTCCGTCGCCCCGTA*CAGGTGACGCAGTACGATCC

Table S2: Obtained C- and TM-scores generated by I-TASSER for HrmO3$_A$ and the engineered A domains. The quality parameters of the created models are in the same range as HrmO3$_A$, indicating sufficient quality of the models.

Protein	**C-score**	**TM-Score**
HrmO3$_A$	0,96	0,84
O2f$_A$	0,81	0,82
O4f$_A$	0,54	0,79
P2f$_A$	0,51	0,78
A5f$_A$	0,77	0,82
H6f$_A$	0,81	0,82