Electronic Supplementary Material (ESI) for Molecular BioSystems

Supplementary Information

Design of protein congeners containing β-cyclopropylalanine

Carlos G. Acevedo-Rocha,*[a,b] Anna-Skrollan Geiermann, [c] Nediljko Budisa, [d] and Lars Merkel[d]

Materials and Methods

1. Chemicals

All standard chemicals were obtained from Sigma (Steinheim, Germany) or Merck KGaA (Darmstadt, Germany) unless otherwise indicated. The L-isomers from methionine (Met) and β-cyclopropylalanine (Cpa) were purchased from Sigma. The radioactively labelled Pyrophosphate (^{32}P) and Met (^{35}S) were bought from Perkin Elmer (Boston MA, USA).

2. Plasmids

a) ***Thermoanaerobacter thermohydrosulfuricus* lipase (TTL)**. The thermostable TTL was cloned as previously described.[1] Briefly, the lipase gene was cloned between the *Eco*RI and *Pst*I of plasmid pQE80L (Qiagen, Hilden, Germany), leaving the open reading frame to stop after the C-terminal His-Tag and a short Gly-Ser spacer sequence.

b) **Human annexin V (anxA5)**. The gene of the human annexin V was cloned within the sites *Nco*I and *Hind*III of the plasmid pRSET-5d under the control of the T7 promoter, as previously described.[2] The plasmid encoding HAV was used as a template for site-directed mutagenesis using appropriate primers via QuikChange Kit (Strategene, La Jolla, CA, USA) to generate the mutants HAV(A2G) and HAV(A2R).

c) **Enhanced green fluorescent protein (eGFP)**. The gene encoding eGFP (747 bp) was PCR-amplified with suitable primers and cloned into the ampicillin-resistance pQE60L (Qiagen) plasmid under the control of the strong inducible T5 promoter between the *Nco*I and *Hind*III sites as previously described.[3] This plasmid contains a C-terminal His-Tag but needs the co-transformation

of the kanamycin-resistant vector pREP4 (Qiagen, Hildesheim, Germany) encoding the repressor gene *lacI*^{q}. The plasmids encoding eGFP were subjected to site-directed mutagenesis using suitable primers to generate the mutants eGFP(G2A) and eGFP(G2R).

d) **Barstar (b*)**. The gene encoding *B. amyloliquefaciens* b* (270 bp) was originally cloned between the *Eco*RI and *Hind*III sites of vector pKK223-3 (Pharmacia, now GE Healthcare, Munich, Germany). Thereafter, using the same restriction sites b* was cloned into pQE80L eliminating the N-terminal His-Tag as previously described.[4] The plasmids with b* were used as template to generate the corresponding mutant b*(K23M/E47M/K79M) via QuikChange with suitable primers.

e) ***Escherichia coli* methionine amino peptidase (*Ec*MetAP)**. The *Ec*MetAP plasmid was a gift from Christian Klein (Heidelberg University).

f) ***E. coli* methionyl-tRNA synthetase (*Ec*MetRS)**. The gene encoding the *Ec*MetRS (2034 bp) was amplified by PCR from genomic DNA of *E. coli* BL21(DE3) using the primers: 5'-agctagctgaattcattgaaggccgtactcaagtcgcgaagaaa-3' / 5'-agctagctgaattctcattatttcac-ctgatgacccggttt-3' and cloned between the *Eco*RI sites (underlined) of plasmid pET-28a (Novagen, Merck Chemicals Ltd., Nottingham, UK) following standard procedures. This plasmid contains an N-terminal His-Tag, the T7 promoter and is resistant to kanamycin. A T7-Tag and Fxa cleavage site was added between the His-Tag and before starting the *Ec*MetRS coding region.

3. Protein sequences

The Met residues to be replaced by Cpa are indicated in **bold**, His-Tags are underlined, T7-Tag is double underlined, a Fxa protease cleavage site is underlined in bold and the stop signal as *.

a) TTL:

MQKAVEITYN GKTLRG**M**M**H**L PDDVKGKVP**M** VI**M**FHGFTGN KVESHFIFVK **M**SRALEKVGI
GSVRFDFYGS GESDGDFSE**M** TFSSELEDAR QILKFVKEQP TTDPERIGLL GLS**M**GGAIAG
IVAREYKDEI KALVLWAPAF N**M**PELI**M**NES VKQYGAI**M**EQ LGFVDIGGHK LSKDFVEDIS
KLNIFELSKG YDKKVLIVHG TNDEAVEYKV SDRILKEVYG DNATRVTIEN ADHTFKSLEW
EKKAIEESVE FFKKELLKGG SHHHHHH*

b) AnxA5:

MAQVLRGTVT DFPGFDERAD AETLRKA**M**KG LGTDEESILT LLTSRSNAQR QEISAAFKTL
FGRDLLDDLK SELTGKFEKL IVAL**M**KPSRL YDAYELKHAL KGAGTNEKVL TEIIASRTPE
ELRAIKQVYE EEYGSSLEDD VVGDTSGYYQ R**M**LVVLLQAN RDPDAGIDEA QVEQDAQALF
QAGELKWGTD EEKFITIFGT RSVSHLRKVF DKY**M**TISGFQ IEETIDRETS GNLEQLLLAV
VKSIRSIPAY LAETLYYA**M**K GAGTDDHTLI RV**M**VSRSEID LFNIRKEFRK NFATSLYS**M**I
KGDTSGDYKK ALLLLCGEDD *

c) AnxA5(A2G):

MGQVLRGTVT DFPGFDERAD AETLRKA**M**KG LGTDEESILT LLTSRSNAQR QEISAAFKTL
FGRDLLDDLK SELTGKFEKL IVAL**M**KPSRL YDAYELKHAL KGAGTNEKVL TEIIASRTPE
ELRAIKQVYE EEYGSSLEDD VVGDTSGYYQ R**M**LVVLLQAN RDPDAGIDEA QVEQDAQALF
QAGELKWGTD EEKFITIFGT RSVSHLRKVF DKY**M**TISGFQ IEETIDRETS GNLEQLLLAV
VKSIRSIPAY LAETLYYA**M**K GAGTDDHTLI RV**M**VSRSEID LFNIRKEFRK NFATSLYS**M**I
KGDTSGDYKK ALLLLCGEDD *

d) AnxA5(A2R):

MRQVLRGTVT DFPGFDERAD AETLRKA**M**KG LGTDEESILT LLTSRSNAQR QEISAAFKTL
FGRDLLDDLK SELTGKFEKL IVAL**M**KPSRL YDAYELKHAL KGAGTNEKVL TEIIASRTPE
ELRAIKQVYE EEYGSSLEDD VVGDTSGYYQ R**M**LVVLLQAN RDPDAGIDEA QVEQDAQALF
QAGELKWGTD EEKFITIFGT RSVSHLRKVF DKY**M**TISGFQ IEETIDRETS GNLEQLLLAV
VKSIRSIPAY LAETLYYA**M**K GAGTDDHTLI RV**M**VSRSEID LFNIRKEFRK NFATSLYS**M**I
KGDTSGDYKK ALLLLCGEDD *

e) eGFP:

MGSSVSKGEE LFTGVVPILV ELDGDVNGHK FSVSGEGEGD ATYGKLTLKF ICTTGKLPVP
WPTLVTTLTY GVQCFSRYPD H**M**KQHDFFKS A**M**PEGYVQER TIFFKDDGNY KTRAEVKFEG
DTLVNRIELK GIDFKEDGNI LGHKLEYNYN SHNVYI**M**ADK QKNGIKVNFK IRHNIEDGSV
QLADHYQQNT PIGDGPVLLP DNHYLSTQSA LSKDPNEKRD H**M**VLLEFVTA AGITLG**M**DEL
YKHHHHHH*

d) eGFP(G2A):

MASSVSKGEE LFTGVVPILV ELDGDVNGHK FSVSGEGEGD ATYGKLTLKF ICTTGKLPVP
WPTLVTTLTY GVQCFSRYPD H**M**KQHDFFKS A**M**PEGYVQER TIFFKDDGNY KTRAEVKFEG
DTLVNRIELK GIDFKEDGNI LGHKLEYNYN SHNVYI**M**ADK QKNGIKVNFK IRHNIEDGSV
QLADHYQQNT PIGDGPVLLP DNHYLSTQSA LSKDPNEKRD H**M**VLLEFVTA AGITLG**M**DEL
YKHHHHHH*

e) eGFP(G2R):

MRSSVSKGEE LFTGVVPILV ELDGDVNGHK FSVSGEGEGD ATYGKLTLKF ICTTGKLPVP
WPTLVTTLTY GVQCFSRYPD H**M**KQHDFFKS A**M**PEGYVQER TIFFKDDGNY KTRAEVKFEG
DTLVNRIELK GIDFKEDGNI LGHKLEYNYN SHNVYI**M**ADK QKNGIKVNFK IRHNIEDGSV
QLADHYQQNT PIGDGPVLLP DNHYLSTQSA LSKDPNEKRD H**M**VLLEFVTA AGITLG**M**DEL
YKHHHHHH*

f) b*$_{1AUG}$:

MKKAVINGEQ IRSISDLHQT LKKELALPEY YGENLDALWD CLTGWVEYPL VLEWRQFEQS
KQLTENGAES VLQVFREAKA EGCDITIILS *

g) b*$_{4AUG}$ (K23M/E47M/K79M):

MKKAVINGEQ IRSISDLHQT LK**M**ELALPEY YGENLDALWD CLTGWV**M**YPL VLEWRQFEQS
KQLTENGAES VLQVFREA**M**A EGCDITIILS *

h) *Ec*MetAP:

MAISIKTPED IEKMRVAGRL AAEVLEMIEP YVKPGVSTGE LDRICNDYIV NEQHAVSACL
GYHGYPKSVC ISINEVVCHG IPDDAKLLKD GDIVNIDVTV IKDGFHGDTS KMFIVGKPTI
MGERLCRITQ ESLYLALRMV KPGINLREIG AAIQKFVEAE GFSVVREYCG HGIGRGFHEE
PQVLHYDSRE TNVVLKPGMT FTIEPMVNAG KKEIRTMKDG WTVKTKDRSL SAQYEHTIVV
TDNGCEILTL RKDDTIPAII SHDE*

i) *Ec*MetRS:

MGSSHHHHHH SSGLVPRGSH MASMTGGQQM GRGSEFIEGR TQVAKKILVT CALPYANGSI
HLGHMLEHIQ ADVWVRYQRM RGHEVNFICA DDAHGTPIML KAQQLGITPE QMIGEMSQEH
QTDFAGFNIS YDNYHSTHSE ENRQLSELIY SRLKENGFIK NRTISQLYDP EKGMFLPDRF
VKGTCPKCKS PDQYGDNCEV CGATYSPTEL IEPKSVVSGA TPVMRDSEHF FFDLPSFSEM
LQAWTRSGAL QEQVANKMQE WFESGLQQWD ISRDAPYFGF EIPNAPGKYF YVWLDAPIGY
MGSFKNLCDK RGDSVSFDEY WKKDSTAELY HFIGKDIVYF HSLFWPAMLE GSNFRKPSNL
FVHGYVTVNG AKMSKSRGTF IKASTWLNHF DADSLRYYYT AKLSSRIDDI DLNLEDFVQR
VNADIVNKVV NLASRNAGFI NKRFDGVLAS ELADPQLYKT FTDAAEVIGE AWESREFGKA
VREIMALADL ANRYVDEQAP WVVAKQEGRD ADLQAICSMG INLFRVLMTY LKPVLPKLTE
RAEAFLNTEL TWDGIQQPLL GHKVNPFKAL YNRIDMRQVE ALVEASKEEV KAAAAPVTGP
LADDPIQETI TFDDFAKVDL RVALIENAEF VEGSDKLLRL TLDLGGEKRN VFSGIRSAYP
DPQALIGRHT IMVANLAPRK MRFGISEGMV MAAGPGGKDI FLLSPDAGAK PGHQVK*

4. Strains

To incorporate Cpa into proteins, the Met auxotrophic strain CAG18491 (LAM-rph-1 metEo 3079:Tn10), obtained from the *E. coli* Genetic Resources at Yale (CGSC# 7464), was employed. For overexpressing the *Ec*MetRS and *Ec*MetAP, *E. coli* BL21(DE3) cells (Stratagene) were used.

5. Protein expression

a) **Congeners:** The Met auxotrophic *E. coli* strain CAG18491 was transformed by electroporation with TTL, anxA5, eGFP, b* and mutants thereof. To express the recombinant proteins, the cells were firstly grown in New Minimal Medium[5] with 100 mg/mL ampicillin and a limiting concentration of 30 mM Met at 37 °C and continuous shaking at 220 rpm. Upon Met depletion, as indicated by growth arrest at an OD_{600} of 0.6-0.8, 0.5 mM Cpa were added to the medium, followed by the addition of 1 mM IPTG after 10-15 min to induce target protein expression. After 4-6 hrs, the cells were harvest by low-speed centrifugation and re-suspended in 50 mM NaH_2PO_4/300 mM NaCl/10 mM imidazole/pH 8.0 for the His-tagged proteins (C-terminus: anxA5, eGFP and N-terminus: TTL) and in 50 mM Tris-HCl/100 mM NaCl/pH 8.0 in the case of tag-free b*. All proteins were expressed in soluble form, except b* that was obtained in inclusion bodies.

b) **Non-congeners**: Electrocompetent *E. coli* BL21(DE3) cells were transformed with the plasmid encoding *Ec*MetRS and plated on LB media containing 36 µg/mL kanamycin. Single colonies were picked and grown in 5 mL of LB culture with the same antibiotic concentration at 37°C. When cells reached an OD_{600} of 0.6-0.8, protein expression was induced by the addition of 1 mM IPTG. The culture was vigorous shaked for 4-6 h at 30 °C.

6. Protein purification

a) **His-Tagged proteins**: In the case of TTL, anxA5, eGFP and *Ec*MetRS, standard Ni-NTA purification strategy was applied as reported elsewhere.[3] Briefly, 1 mg/mL DNase (Roche, Mannheim, Germany), RNase and lysozyme were added to the re-suspended cells, followed by sonication and high-speed centrifugation. The lysate was loaded onto a 1 mL HiTrap chelating HP column (GE Healthcare), washed with sodium phosphate buffer containing 20 mm imidazole, followed by high-salt washing buffer (50 mM NaH_2PO_4/2 M NaCl/20 mM imidazole/pH 8.0). The proteins were eluted with an imidazole gradient (20-500 mM) in sodium phosphate buffer. The

fractions with the desired proteins, as judged by SDS-PAGE, where pooled, dialyzed against 50 mM Tris-HCl pH 8.0 and concentrated by ultra-filtration (Vivaspin 20 MWCO 10,000; Sartorius AG, Goettingen, Germany).

b) **Tag-less proteins**: b* mutants were processed as above but the cell pellet containing inclusion bodies was dissolved in 50 mM Tris-HCl/7.5 M urea/pH 8.0. The suspension was then separated by high-speed centrifugation and the supernatant was dialyzed (3,500 Da MWCO) against 50 mM Tris-HCl/100 mM NaCl/pH 8.0 thrice. The lysate was cleared from solid particles and loaded onto a 5 ml HiTrap Q Sepharose column (GE Healthcare), washed with 50 mM Tris-HCl/100 mM NaCl/pH 8.0, and barstar was then eluted with a 100-1000 mM NaCl gradient in 50 mM Tris-HCl/pH 8.0. The fractions were analyzed by SDS-PAGE and the desired congeners were pooled, dialyzed against 50 mM Tris-HCl/pH 8.0 and concentrated by ultra-filtration (Vivaspin 10 MWCO 5,000; Sartorius AG).

Using BSA for calibration, all samples were assayed for protein concentration by the Bradford assay.[6] To estimate purity, 10-12% and 16% SDS-PAGE was respectively carried out for His and non-tagged proteins and congeners, yielding the expected bands with 70-95% purity (See Figure S1).

7. Analytical characterization by ESI-MS

After successful expression and purification of proteins/mutants and congeners thereof, 20 µl of sample (approx. 1 mg/mL) were separated on a Waters RP C4 column (100 mm large x 2.1 mm wide, 30 nm pore size; 3.5 mm particle size; Waters, Eschborn, Germany) by applying within 20 min a gradient elution (flow rate of 250 µl/min) from 80 % A to 90% B (A: 0.05 % (v/v) TFA in water, B: 0.05 % (v/v) TFA in acetonitrile). The masses of the eluted protein fractions were analyzed on a MicroTOF ESI-MS (Bruker Daltonics, Bremen, Germany). The peak intensities of the fully and partially substituted eGFP and barstar congeners were deduced, summed up from the corresponding mass spectra and indicated in percentages in order to give a semiquantitative result.

8. *N*-terminal Met excision assay

The model peptides were synthesized according standard protocols using the Fmoc-strategy. In this study, the following peptides were used: H-Met-Gly-Ser-Ser-Leu-Phe-NH_2, H-Cpa-Gly-Ser-Ser-Leu-Phe-NH_2, H-Cpa-Gly-Gln-Val-Lys-Tyr-NH_2 and H-Cpa-Ala-Gly-Gln-Val-Lys-Tyr-NH_2. A 72 mM solution of peptide in H_2O containing 21.4 % DMSO was prepared. Afterwards, 25 µl of the 72 mM solution were added to the reaction mixture containing 325 µl H_2O, 50 µl 2 mM $CoCl_2$, and 50 µl Na_2HPO_4/NaH_2PO_4 buffer solution pH 7.5. The reaction was started by the addition of 50 µl *Ec*MetAP (0.13 mg/ml) and agitated at 37 °C / 500 rpm. After 10 min, 30 min, 60 min, 90 min, 120 min and 300 min, a sample of 50 µl was taken from the reaction mixture and added to 10 µl of 1 M HCl to stop the enzymatic activity. The reaction products were subsequently analyzed by RP-HPLC. Analytical RP-HPLC was performed on a Waters GmbH (Eschborn) HPLC system equipped with a controller 600, pumps 626 with inline-degasser AF, autosampler 717plus, data module and dual λ absorbance detector 2487. Peptides were analyzed on an ET 125/4 Nucleosil 100-5 C8 column (Macherey & Nagel, Düren, Germany) using a gradient from 95 % A (0.1 % TFA in water) to 90 % B (0.08 % TFA in acetonitrile) within 13 min at a flow rate of 1.5 ml/min.

9. tRNA aminoacylation competition assay

Partially inspired by the competition assay reported by Szostak and colleagues[7, 8], we developed a similar assay but using radioactively-labelled Met. In this case, our method allows determining the tRNA

aminoacylation preference of the *Ec*MetRS for Met analogues when compared to Met. In the reaction containing a minimal amount of labelled ^{35}S-Met (5 nM), bigger amounts of unlabelled Met are added until no radioactivity can be detected. The 'cold' Met will compete against the labelled one for the active site of the cognate *Ec*MetRS. This indicates an optimal methionylation level that can be compared to the scenario presented by any Met analogue, for example, Cpa. In this way, the aminoacylation efficiency of both Met- and Cpa-tRNA^{fMet} by the *Ec*MetRS was compared using initiator tRNA^{fMet}, which thereafter was precipitated onto a filter and the amount of radioactivity was measured. The reaction was performed as previously described[7] but with the following modifications: The reaction mix (30 µl) was incubated at 37 °C in 50 mM HEPES/pH 7.4, 10 mM $MgCl_2$, 50 mM KCl, 3 mM BME, 5 mM ATP, 1 mg/mL BSA, 5 nM [^{35}S]-Met (3 cpm/µmol), tRNA^{fMet} (10 µM), and *Ec*MetRS (5 µM). After incubation for 20 min, 20 µl of the reaction mixture were transferred onto a Whatman GF/F paper (GE Healthcare) and the tRNA^{fMet} was precipitated by the addition of 1 ml of trichloroacetic acid or TCA (10%). The tRNAs were step-wise washed (1 min each) with 1 ml of TCA (5%), 1 ml of ethanol (abs.) and 1 ml of ethanol:ether (solution 1:1). Finally, the filters were dried for 5 min and mixed with 7 ml scintillation solution (Carl Roth) to determine the amount of radioactivity in a counter (Perkin Elmer).

10. Literature

1. M. G. Hoesl, C. G. Acevedo-Rocha, S. Nehring, M. Royter, C. Wolschner, B. Wiltschi, N. Budisa and G. Antranikian, *ChemCatChem*, 2011, **3**, 213-221.
2. N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann and R. Huber, *Eur. J. Biochem.*, 1995, **230**, 788-796.
3. T. Steiner, P. Hess, J. H. Bae, B. Wiltschi, L. Moroder and N. Budisa, *PLoS One*, 2008, **3**, e1680.
4. S. Lepthien, L. Merkel and N. Budisa, *Angew Chem Int Ed Engl*, 2010, **49**, 5446-5450.
5. N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann and R. Huber, *European journal of biochemistry / FEBS*, 1995, **230**, 788-796.
6. M. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248-254.
7. M. C. Hartman, K. Josephson and J. W. Szostak, *Proc. Natl. Acad. Sci. U S A*, 2006, **103**, 4356-4361.
8. M. C. T. Hartman, K. Josephson, C.-W. Lin and J. W. Szostak, *PLoS ONE*, 2007, **2**, e972.