SUPPORTING INFORMATION

Loading of malonyl-CoA onto tandem acetyl carrier protein domains of polyunsaturated fatty acid synthases

Omar Santín and Gabriel Moncalián

Material included: Supplementary Figures S1 to S9

mmKSAT CpKSAT SbKSAT SchKSAT SoceKSAT SoceKSAT consensus>70	1 10 MAKKNTTSIK MAKKQQKSSC SK MAARL ME	HAKDV.LS	29 SDDQQLNSR NEEQAFNSR .KDKRLNKR QEQKGG	30 LQECPIAII LQECPIAVV LLENDYVAIV EMDTRIAII IAVV EVVIA !a!v	40 GMASUFADAK GMASIFADAK GMSAIFANSR GMSAILPCGT GVSAIFPGSL GMSGKLPESE Gm. i	50 NLDQFNDNIV NLDVCDNIF YLNKFOLISE TVRESCETIR DAHGFCRDILS NLEEFNNLIC	ε ο S VDAILDVP S VDAIRDVP K IDAIRDVP G IDCLSDLP G IDCLSDLP G IDLITDVP G VDMVT.AD . vD.i.d.p	7 Q SD RWN TD D HY SD RWA ID D YY D THWR AE D YYI ED RVD YYAYFI ST HWL VE D YYI DR RWK AG L YG JWYY	SOOKAADKTYCK SADKKAADKTYCK SOPKAADKTYCK OADKSKADKSYCK OPVKTKDKIYCK OPVFXFVDKIYCK OPDPSAPDKTYAK
mmKSAT CpKSAT SbKSAT SchKSAT SoceKSAT SscrKSAT consensus>70	90 RGGFIPEL.D RGGFLPEY.D RGGFIPEY.D RGGFIPEY.D RGAFLKDV.P RMGKLKDLSF Rggf#	100 DPMEFGLP DPMEFGLP DPMEFGLP DAREFGLN DDASFFGVH F#efG	110 PNILELTDI PNILELTDI PNILELTDI PNILELTDI SKQANTMDP eD.	120 AQLLSTVA AQLLSTVA SQLLSTVA NQTISTLA ISTLAT QLRMLTEVT .q.l.L.v.	130 RDVLSDAGIG RDVLNDAGIG KEVLADANLP KEALQDAGID KEVLEDAAQG YEAIVDGGIN .e.l.Da	140 SDYDHDKIC DGSGYDRDKVC TDYDRDRIC ALGKERKNIC JFESMSREN .PASLRGTSTC	150 ITL <u>C</u> VGGGO ITL <u>C</u> VGGGO ITL <u>C</u> VGGGO IVL <u>C</u> VTSAO IVL <u>C</u> VTSAO IVWV <u>C</u> VSSSD I.1G!#	160 KQISPLT <mark>SRLQ</mark> KQISPLT <mark>SRLQ</mark> KISQSLNSRLQ KISQSLNSRLQ ELLASMVSRLQ 	170 GPVLEKVLKASG GPVLEKVLKASG YVVVKVFKSSG YVVVEKVLRKMG RPVWAKALRDLG
mmKSAT CpKSAT SbKSAT SchKSAT SoceKSAT SocrKSAT consensus>70	180 IDEDRAMII VDAEDRAMII LSDEDRAMII MPEEDVKVAV YPEDEVKRAC ASEAI ed	190 DKFKKAYIG KKFQDQYIH EXFKANFPE DKIAGNYVP SRDFETLVG . kyi.	200 WEENSFPGM WEENSFPGS WRLDSFPGF WRLDSFPGF YSMIGC WSfpG.	219 LGNVIAGRI LGNVIAGRI LGNVTAGRI LGNVTAGRI QRAMMANRI Ignv.agR.	220 ANRFDFGGTN ANRFDFGGMN ANRFDFGGMN TNTFNLDGMN ANRLDLGGTN SFFFDFKGPS .n.f#G.n	230 CVVDACAGSI CVVDACAGSI CVVDACASSI CVVDACASSI CVVDACASSI CVTDACASSI CVVDAACASSI	240 AAUKMAISD AAIKLAISD AAIRMALTE IAVKVAIDE SAMSMAINE LALQSAYOA . A A. e	250 LEYRSEVMIS LEHRSEVMI LYGCSDMM7 LYGCCDMW7 ALGCSDLVIJ RGGEC <u>SAAV</u> Ld!	260 GVCCDNSPFMY GCVCCDNSPFMY GCVCTDNSPSMY GACCDNSIGMY GCDTMNDAFMY GCLNVLLKPNSS Ggn.my
mmKSAT CpKSAT SbKSAT SchKSAT SccKSAT SscrKSAT consensus>70	270 MSESKTPAFT MSESKTPAFT MADSKTPAFT MADSKTPVFS MCESKTPALS LOEMKLGMLS \$.Fsktp	280 TNDDIRPFD TNETIQFFD TDPSVRAYD KSGDCRPFS ODGTCRSFD	290 DDSKGMLVG NDSKGMMIG IDSKGMMIG EKTKGMLIG DKADGTLLG AEGTGYCRA G.l.g	300 SGIGMMAFK SGIGMMAFKI SGIGMVALKI SGIAMVALKI SGIAMVALKI SGIAMVALKI SGIAMVALKI SGIAMVALKI	310 RLEDAERDGD RLEDAERDGD RYADAVRDGD RYADAVRDGD KKSLAR rda.rdgd	320 KIYSVLKGIGT KIYAVLKGIGT RIYAVIKGVG2 EIHAVIRGCAS RVYAVIRGIGS RVYATILNAGT .!yavg.g.	330 SSDGR.FKS SDGR.FKS SDGR.AAG SSDGR.SKS NTDGSKEQG SSDG	340 IYADR PDG QAH IYADR PDG QAH IYADR PBG QAH IYT PT ISG QEH YYADV PEG QAH YF PS GD VQE Q IY P. dg Q.	350 (ALKRAYEDAGFA (ALKRAYDDAGFA (ALKRAYDDAGFA (ALRRAYDDAGFA (ALRRAYNRACVO (ALRRTYAAAGYG (DIFSLYAPAGPD al.r.Y.Ag.)
mmKSAT Cp_KSAT SbKSAT SchKSAT SoceKSAT SscrKSAT consensus>70	360 PETCGLIEGH PRSCGLIEAH PHSIGLVEAH PATVTLVEGH PETVELMEAH PESLE <u>VIEA</u> H Plie.H	379 GTGTKAGDAJ GTGTKAGDAJ GTGTAAGDVJ GTGTVGDRJ GTGTKAGDAJ GTGTKVGDP GTGT.GD.	380 EFAGLTKH EFSGLVKH EFSGLVKH ELTALRNL EFEGLRAM EFEGLRAM Eglf	390 GAASDEH SAN.NEQH AQG.NDTH DKAYGEGN DKAYGEGN CATRH	0 400 KQY IALGSVK: KQH IALGSVK: NQH IALGSVK: RQK IALGSVK: RQK IALGSVK: RQK IALGSVK: REP LLIGSVK: #) 410 OIGHTKSAAG OIGHTKAAAG SUGHTKSTAG SIGHLKAVAG OIGHTKAAAG SIGHLKAVAG SIGHLKAAAG	42(SAGMIKALI SAGMIKAVLI TAGVIKAALI LAGMIKVIMI AAGLIKAIMI VALIKVILI . AgmiK \$a) 430 HHKILPATI HHKVLPATI HHKVLPATI HHKVLPATI HHKVLPATI HHKVLPPTI SEHGVWAPNI L.Hkvlp.ti	HIDKPSEAE HIDQPNTSL NVSQPNFXL NVDNPPNLYDNT KVDKPNPKL HYDTPNFL HYTTPNF
4 4 mmKSAT Cp_KSAT SbKSAT SchKSAT SoceKSAT SoceKSAT consensus>70	40 45 DTKNSPLYLN ATENSPMYLN NTESSPFYLN FINESSLYIN DTEKTAFYLN PALQD.G <u>RLQ</u> .iyl#	0 40 SETRPWMARE TETRPWMARE TETRPWFQRJ TMNRPWFQRJ TMNRPWFQRJ TQARPWFRP, VVDRPLPIR.	50 EDGI <mark>PRRAGI</mark> EDGL <mark>PRRAGI</mark> ADAT <mark>PRRAGI</mark> O GDHPRRAGI GDH PRASI GGNVGI P rrag!	SFGFGGG SFGFGGT SFGFGGT SFGFGGT SFGFGGS NSFGFGGS SFGFGG	480 YEMYLEEYY YEMYLEEYY YELVLEEYY YELVLEEYY YELVLEEYY YUYILOP NSI YEH.VL#e	490 GH. DSAYRI AA. QGYYRI PEHSRDEQYRQ EH. TTAYRI FPA. PKAWRV RPA. PPPAQH	500 NSVSQTVLT: NSVPQTILV7 RSVPQTILF NKRPQPVLM RALPAELFUI AALPRLLQA: P	510 EANDQQ <mark>GI</mark> VAE TAKDEHALITÇ AAANKAALLSE MAATPAALQSI ISADTPAALAI SGRTLE <mark>AV</mark> QTI	520 LINNWRTKL LLQQWQDKL LLKALSQSV CEAQLKEFEAAI RARALAKE LLEQGLRH
mmKSAT Cp_KSAT SbKSAT SchKSAT SoceKSAT SoceKSAT consensus>70	530 AVDADH.QGF SSVADA.QPY NTNANKSSAA KENETVKNTA AE.VPE SRDLAFVG d	VFNELVT IFNALVT SLNAIAQ YIKCVKFGEQ ILRFLAR MLNEIAA	540 TWPLKTPS DCALQTPP QYPLRTLP 2FKFPGSIPP ESVLSFD VSPVAMPF	550 SVNQARLGF AENLARCGF ASTDARLGF ASTDARLGF ASTPARLGL YRGY <mark>AVLG</mark> GF ArlG.	560 VARNANEAIAM VARNANEAIAM VAKNANEAIXM VAKDARDACS CATDEADLRKH SAGS	570 IDTALKQFNA IAGALKQFAS LNQSIAHLET LRAICAQFAK LEQVAAHLEA	580 NADKMTWSV NVGSDSWSV NA.IEVWQL DVTKEAWRL RP.EQALSA QEVQQV	590 T. GVYYRQAO T. GIYYRKSF S. GISYRSHF REGVSFRAKO LVHCAS	600 IIDATGKVVA IIKTKGKVVA ILVAKNESKKVA IIATNGAVA GEAPGRVA V
mmKSAT Cp_KSAT SbKSAT SchKSAT SoceKSAT SoceKSAT consensus>70	610 LFSGCSSYV LFSGCSSYV LFACCSSYV LFSGCAOYL LFSGCSSYV LFPGCSSYV LFPGCSSYV LFPGCSSYV LFSGRAOWQ lf.GqG.Qy.	620 NMGRELTCN MGRELACN NMGRELACN HMFSEVAMN GMGADALMT GMGLSL.NRI .Mg.e	630 PSMMHSAA PSVMQAAQI PPEMRQQVM PQFRQSIA DPARAMDA DRFRDSILF	640 MDKEFSAA MDTEFSNA LSDKVFAHH MDAAQSKVI AAGV.AI LSDQALKPL .d	GLGQL GLGQL GQTPL AGSDKDFERV ADAPL GL.R	650 SAVTFPIPVYT PTTYPIPVFS SNILYPIPAFD QVLYPRKPYE HEVVFPRPYS 	660 DAERKLQEE(DDARKAQDE ADAVKAQEA) REPE.QDHKI DEDRAAQEA DLLLSTDEA #e.	670 DERLTQHAOPF VIRTCHAOPS ALTNTLFAOSS USLTAYSOPS LERTRAOPS VLDDIVSSPVS 1tq.	680 TGSLSVGLFKTF IGTLSVGLYKTF IGAISMAQYSLL TLACALGAFEIF IGATSLAHLALL LTSIQIALIDLL
69 mmKSAT Cp_KSAT SbKSAT SchKSAT SoceKSAT SscrKSAT consensus>70	Q 700 KQAGFKADFA TNAGFKADFT TQAGFAPDMV KEAGFTPDFA AALGVRAEAF TSLGLQPDGI G#	Q 71 AGHSFGELTA AGHSFGELTA AGHSFGELSA AGHSLGEFAA AGHSLGEVAC AGHS.GEa	Q 7 LWAADVLSE LWAAGVLND LCAAGVISN LYAAGCVDR LHAAGCVDR HAAGCLDR GYADGCLTQ 1. Aag	2 Q SDYMMLARS SDYMMLARS EDYVELAFA DELFELVCR ADLLRVARR EEAVLSSYW	739 RGQAMAAPEQ RGQAMAAPKD RGHAMAQVPS RARIMGGKDA RGELMGAA RGYCIKEA RGm	DTASKAADTGT	40 DFDAGKMAA TVDTGTMIA OTDLGTMFA ATPKGCMAA SSERGAMIA NVLPGAMAA 	750 VGDPK VGAPD ILKQKNDIDA IGPNA JPRGID GLSWE I	760 VAVIIDTLDDV VAADIKDIKDI LNSCLAQFDGV NIKVQAANV VRALGLGGV CKQRCPPGI
mmKSAT Cp_KSAT SbKSAT SchKSAT SoceKSAT SscrKSAT consensus>70	779 SIANFNSNNQI SIANYNSNNQI KIANYNAPTQI WLGNSNSPSQ VIANHNGPKQI VPACHNSKDI an.Nq	780 VVAGTTEQV VVAGTTSQI LVIAGGTEQT IVISGSVAGI VVLSGSVAAI VTISGPQAAM .vi.G	790 AVAVTTLGN. AIAIEELKG QLAAKAISE QAESARLQK EAAEERLKG SEFLQQLKR.	800 AGFKVVPLP LGFKAIALP EGFRVVPLA AGIQARRLD DVFVKEVR .gl.	810 V.SAFHTPL V.SAFHTPL V.SGAFHTPL C.ESAFHSPQ V.AAFHSPL TGGIAFHSYF AFH.p.	820 VRHACKFFAKA VGHACKFFSDA VGHACKFFSDA VGHASSAFKDV VAEASAFLREF MESIAFTLRC a	830 VDSAKF.KA IDNAKF.NK IDKAKF.SA ISKVSF.RT LAGIAV.EA LRKVILDPK	840 VV. PVSNO VV. PVSNO SV. ALYANO KAETKLFSNV AA. PVSN RS. KRWLST P	GLVHSSK TAKAHSNK TQLHPSD SGIYPT. EIEPYAGG SIPEAQWQGSL
mmKSAT Cp KSAT SbKSAT SchKSAT SoceKSAT SscrKSAT consensus>70	850 PNDIKKN.LK AADIKKS.LK GKAIKAE.FK DAREM.LT GDAARDR.LA ART <mark>FS</mark> AEYSV	860 NHMLESVHEN QHMLQSVHEN QHMLQSVHEN QHMLSSVHEN NNLVSPVLE mV.F.	879 NEEIDNIYAD SEQLQAMYDA JTQVRNMYDA JEEIERMYAD 2EALQHVPAH .ee.d	880 GGRVFIÐF GGRVFVÐF GARIFVÐF GVRIFVÐV GVRIFVÐV G. RVVVÐI g.rvf!E.	890 ENVUTKLV ENVUTKLV ENVUKLV EVUSKLV ESVUSKLV ESVUSKLV APHALLOAVL FL.VL.1V	900 NIITEKSD.V NIIKDKDD.V NTIGEHLNEL ETIKDDPS.DIISGPPH.R RSIESSCT.I n.L	910 TATAVNANP TATAVNANP CLVSMNPNP VTVSVNPASC .AVALDRKG IPL.MKKDH	920 QPADVQMRQA KSADMONRQA GDSDSQIRLA TDSDIQIRLA HG.VTSILEA DN.LEFFLSN 1.a	930 ALQMAVIGVAID AVQMSVIGLELM AVQLAVVGVALT AVQLVVAGVNLQ VARLAVAGVFD VGRLHAGVSV \$.v.Gv.1d
mmKSAT Cp_KSAT SbKSAT SchKSAT SoceKSAT SscrKSAT	940 NIDPYDAVKR DIDPYSAVKR EIDPYCAITS GFDKWDAPDA ARVLWEGFAA PNGLFPPVEF	950 P.LVAP P.LSAP Q.EIAE TRMQAIE PSDP.RALP PAP	96 CASPMLMB CMSPLAMB CASPLAMB CMSPLAMB CKRRTTLB CKRRLALC CGTP.L	Q S Q S A A S Y V S E C L T G A S Y V S E C L T A T N H I S A C L S A A T Y V S I Q I N G S N Y G K E 	970 2KTK 2KTK 2KTK 2YPP 2KTK.WDHSQAW	980 . KAFADALTD . KAFDDALND . AKMAKSLAT . KVRDAAMD . QGGAKALPA NDVPSAAD <mark>F</mark> PS	GWTVKQAKA. GWTIKQAHT. GSVTSQVQY. GRCVTYLKG. ASPPRAAQK. GSSCSS <mark>V</mark> AVY	90 10 .VPAVVSQPQ .TPVAVPAPQ .VDRIVET .AAPLIK P KFDVSPES	00 1010 VIEKIVEVEKIV VVEKIV VVEKEV PDHYLV

Figure S1. Multiple sequence alignment of KS-AT domains. Alignment of representative KS-AT protein regions from Pfa and FAS systems. Sequences correspond to PfaA (mmKSAT) from *Moritella marina* (Acc. no. Q9RA21), PfaA (Cp_KSAT) from *Colwellia psychrerythraea* (Acc. no. Q47ZG8), PfaA (SbKSAT) from Shewanella baltica (Acc. no. A9KUH8), PFA1 (SchKSAT) from *Schizochytrium* ATCC 20888 (Acc. no. AAK72879), pfa2 (SoceKSAT) from *Sorangium cellulosum* (Acc. no. A9EPF7) and mFAS (SscrKSAT) from Sus scrofa (Acc. no. A5YV76). Identical residues are shown in white on a red background, while similar residues are shown in red. The unstructured loop that connects both domains is marked with a green line. The positions of the catalytic residues of both domains have been marked with black (KS) and blue (AT) arrows, respectively.



Figure S2. Multiple sequence alignment of representative tandem ACPs individual domains of PfaA-like proteins. The alignment of the tandem ACP individual domains from PfaA-like proteins is shown. Sequences correspond to PfaA (mmPfaA1, mmPfaA2, mmPfaA3, mmPfaA4, mmPfaA5) from *Moritella marina* (Acc. no. Q9RA21), PfaA (cpPfaA1, cpPfaA2, cpPfaA3, cpPfaA4, cpPfaA5, cpPfaA6) from *Colwellia psychrerythraea* (Acc. no. Q47ZG8), PfaA (sbPfaA1, sbPfaA2, sbPfaA3, sbPfaA4, sbPfaA5) from *Shewanella baltica* (Acc. no. A9KUH8), pfa2 (abPfa1, abPfa2, abPfa3) from *Aetherobacter fasciculatus* (Acc. no. A0A076Q2L6), pfa2 (scPfa1, scPfa2, scPfa3, scPfa4, scPfa5) from *Sorangium cellulosum* (Acc. no. A9EPF7) and PFA1 (schPfa1, schPfa2, schPfa3, schPfa4, schPfa5, schPfa6, schPfa7, schPfa8, schPfa9) from *Schizochytrium* ATCC 20888 (Acc. no. A4K72879). The position of the selected residues for the alanine screening of mmPfaA1 is marked with black arrows and the serine active site with a red arrow. Identical residues are shown in white on a red background, while similar residues are shown in red. The type of organism is indicated by color bars. Position of α -helices is represented with grey boxes at the bottom of the alignment.



Figure S3 Multiple sequence alignment of KR'-KR domains. Alignment of representative KR'-KR protein regions from Pfa and PKS system. Sequences correspond to PfaA (mmKR) from *Moritella marina* (Acc. no. Q9RA21), PfaA (CpKR) from *Colwellia psychrerythraea* (Acc. no. Q47ZG8), PfaA (SbKR) from Shewanella baltica (Acc. no. A9KUH8), PFA1 (SchKR) from *Schizochytrium* ATCC 20888 (Acc. no. AAK72879), Macrolactin PKS (MacrKR) from *Bacillus amyloliquefaciens* (Acc. no. Q1RS63) and PikaII module (PikAIIKR) from *Streptomyces venezuelae* (Acc. no. Q9ZGI4). Identical residues are shown in white on a red background, while similar residues are shown in red. The unstructured loop that connect both KR domains was marked with a green line. The positions of the catalytic residues of the KR domains have been marked with black arrows.



Figure S4. Multiple sequence alignment showing the swapped KR K/N residues wihin PfaA homologous proteins. The alignment of the KR active site region responsible for catalysis in omega-3 synthases in comparison with PKS and FAS systems is shown. Sequences correspond to PfaA (mmKR) from *Moritella marina* (Acc. no. Q9RA21), PfaA (CpKR) from *Colwellia psychrerythraea* (Acc. no. Q47ZG8), PfaA (SbKR) from *Shewanella baltica* (Acc. no. A9KUH8), PFA1 (SchKR) from *Schizochytrium* ATCC 20888 (Acc. no. AAK72879), Pfa2 (SoceKR) from *Sorangium cellulosum* (Acc. no. A9EPF7), SpnC PKS (SacKR) from *Saccharopolyspora spinosa* (Acc. no. Q9ALM4), DEBS (DEBSKR) from *Streptomyces hygroscopicus* (Acc. no. AQW50878), mammalian FAS (mFASKR) from *Sus scrofa* (Acc. no. A5YV76), Actinorhodin PKS (actKR) from *Streptomyces coelicolor* (Acc. no. P16544), Granaticin PKS (GranKR) from *Streptomyces sp. ERV7* (Acc. no. AAO65349). Identical residues are shown in white on a red background, while similar residues are shown in red. The position of the catalytic motifs is marked with black stars, and K/N swapping is represented with two arrows connecting both residues.



Figure S5. Multiple sequence alignment of DH domains. Alignment of representative DH protein regions from Pfa and PKS systems. Aligned sequences correspond to PfaA (mmDH) from *Moritella marina* (Acc. no. Q9RA21), PfaA (CpDH) from *Colwellia psychrerythraea* (Acc. no. Q47ZG8), PfaA (SbDH) from *Shewanella baltica* (Acc. no. A9KUH8), PFA1 (SchDH) from *Schizochytrium* ATCC 20888 (Acc. no. AAK72879) and a PKS type I (StrepDH) from *Streptomyces sp.* MUSC 93 (Acc. no. OIJ85280). Identical residues are shown in white on a red background, while similar residues are shown in red. The positions of the catalytic residues of the DH domain have been marked with black arrows.



Figure S6. Alanine substitution screening for mmACP1. (A) Wild type mmACP1, the apo ACP form and the four mutants incubated with radiolabeled [14C]-malonyl-CoA analyzed by radio-SDS-PAGE. Lane 1: holo-mmACP1+Malonyl-CoA; 2: apo-mmACP1+Malonyl-CoA 3: holo-mmACP1(T1272A)+Malonyl-CoA; 4: holo-mmACP1(Y1274A)+Malonyl-CoA; 5: holo-mmACP1(R1322A)+Malonyl-CoA; 6: holo-mmACP1(T1323A)+Malonyl-CoA. (B) Quantification of the protein bands intensities of mmACP1 wild type and the mutants analyzed in A by densitometric scanning, normalized with molarity of each sample. (C) and (D) correspond to the same protein samples but incubated with mmKS-AT. Black arrows indicate the theoretical mmACP1 molecular weight.



holoACP

Figure S7. High Definition Mass Spectrometry. High definition mass spectra of apo mmACP1 (A) and holo mmACP1 (B). The amount of desalted protein used for both experiments was 40 μ g. MS spectra were manually acquired in the m/z range 500-1700. Default deconvolution parameters were used.



Figure S8. (A) Predicted structure of mmPfaE from *M. marina*. mmPfaE structure was modelled by Phyre2 using *Bacillus subtillis* Spf (1QR0) as template. Predicted catalytic residues D124, E126 and E170 are shown. The ACP-PfaE interaction zone is indicated with a yellow circle. (B) Alignment of mmPfaE with homologous proteins. The active site residues responsible for catalysis are highlighted with red dots and the residues responsible for ACP-PfaE interaction are marked in yellow.



Figure S9. Predicted structure of mmPfaB from *M. marina***.** A linker (grey) connects the pseudo keto synthase N-terminal domain (blue) and the acyl transferase domain (green).