A Versatile Biosynthetic Approach to Amide Bond Formation

Helena K. Philpott, Pamela J. Thomas, David Tew, Doug E. Fuerst, Sarah L. Lovelock,*

Contents

	Page
1. Gene Sequences	S1
2. General protocols for the chemical synthesis of amide standards	S13
3. General protocol for the expression and purification of CoA ligases	S18
4. Pyrophosphate detection assay for the evaluation of CoA ligase substrate specificity	S18
5. Analytical scale CoA ligase reactions	S19
6. General protocol for the expression of the N-acyltransferase panel	S20
7. General protocol for N-acyltransferase assays	S20
7.1 Preliminary assay: NAT Activity towards a panel of amines (**2a-2d**)	S21
7.2 Assay 2: NAT Activity towards a panel of CoA esters	S23
7.3 Assay 3: NAT Activity towards a panel of amines (**2c-2i**)	S24
8. Evaluation of L-azetidine carboxylic acid N-acyltransferases	S24
9. Synthesis of the Losmapimod amide intermediate using whole cell biocatalysts	S25

1. Gene Sequences

CoA ligases

All CoA ligase genes were synthesized at Genscript and the DNA sequences were codon optimised for *E.coli* expression. The genes were cloned into pET28 using NdeI and NotI restriction sites so that the expressed protein contains an N-terminal 6-His tag:

phl from *Penicillium chrysogenum:* (Uniprot O74725)
MGSSHHHHHHSSGLVPRGSHMVFLPPKESGQLDPIPDNIPISEFMLNERYGRVRHASSRDPYTCGITGKSYSSKEV
ANRVDSLARSLSKEFGWAPNEGSEWDKTLAVFALNTIDSLPLFWAVHRLGGVLTPANASYSAAELTHQLLDSKAKA
LVTCVPLLSISLEAAAKAGLPKNRIYLLDVPEQLLGGVKPPAGYKSVSELTQAGKSLPPVDELRWSAGEGARRTAFVC
YSSGTSGLPKGVMISHRNVIANTLQIKAFEQNYRDGGGTKPASTEVALGLLPQSHIYALVVIGHAGAYRGDQTIVLP
KFELKSYLNAIQQYKISALFLVPPIIIHMLGTQDVCSKYDLSSVTSLFTGAAPLGMETAADFLKLYPNILIRQGYGLTETC
TVVSSTHPHDIWLGSSGALLPGVEARIVTPENKEITTYDSPGELVVRSPSVVLGYLNNEKATAETFVDGWMRTGDE
AVIRRSPKGIEHVFIVDRIKELIKVKGLQVAPAELEAHILAHPDVSDCAVIAIPDDRAGEVPKAIVVKSASAGSDESVSQ
ALVKYVEDHKARHKWLKGGIRFVDAIPKSPSGKILRRLIRDQEKEARRKAGSKI

PhCL from *Petunia hybrida:* (Uniprot I3PB37)
MGSSHHHHHHSSGLVPRGSHMPMETETNQGDLIFRSKLPDIYIPKHLPLHSYCFENISEFSSRPCLINGANNHIYTYA
DVELTSRKVAAGLNKLGIQQKDTIMILLPNSPEFVFAFMGASYLGAISTMANPLFTPAEVVKQAKASNAKLIITQACF
VNKVKDYAFDNNLNVICIDSAPEGCIHFSELTQADEHDIPDVKIQSDDVVALPYSSGTTGLPKGVMLTHKGLVTSVA
QQVDGENANLYMHSEDVLMCVLPLFHIYSLNSVLLCGLRVGAAILIMQKFDIVQFCELIEKYKVTIGPFVPPIVLAIAK
SPVVDNYDLSSVRTVMSGAAPLGKELEDAVRIKFPNAKLGQGYGMTEAGPVLAMCLAFAKEPFDIKSGACGTVVR
NAEMKIVDPDTGCSLPRNQPGEICIRGDQIMKGYLNDPAATTRTIDKEGWLHTGDIGYIDNDDELFIVDRLKELIKY
KGFQVAPAELEALLLNHPNISDAAVVPMKDEQAGEVPVAFVVRSNGSDITEDEVKDFVSKQVIFYKRIKRVFFVETV
PKSPSGKILRKDLRARLAAGVPN

CBL from *Alcaligenes sp.* ALP83: (Uniprot Q8GN86)
MGSSHHHHHHSSGLVPRGSHMQTVNEMLRRAATRAPDHCALAVPARGLRLTHAELRARVEAVAARLHADGLRP
QQRVAVVAPNSADVVIAILALHRLGAVPALLNPRLKSAELAELIKRGEMTAAVIAVGRQVADAIFQSGSGARIIFLGD
LVRDGEPYSYGPPIEDPQREPAQPAFIFYTSGTTGLPKAAIIPQRAAESRVLFMSTQVGLRHGRHNVVLGLMPLYHV
VGFFAVLVAALALDGTYVVVEEFRPVDALQLVQQEQVTSLFATPTHLDALAAAAAHAGSSLKLDSLRHVTFAGAT
MPDAVLETVHQHLPGEKVNIYGTTEAMNSLYMRQPKTGTEMAPGFFSEVRIVRIGGGVDEIVANGEEGELIVAAS
DSAFVGYLNQPQATAEKLQDGWYRTSDVAVWTPEGTVRILGRVDDMIISGGENIHPSEIERVLGTAPGVTEVVVIG
LADQRWGQSVTACVVPRLGETLSADALDTFCRSSELADFKRPKRYFILDQLPKNALNKVLRRQLVQQVSS

ipfF from *Sphingomonas* Ibu-2: (Uniprot A1E027)
MGSSHHHHHHSSGLVPRGSHMLARDLVKRCARNYPTKTAYLCGERSRSWREMDQRSDRFGVALQQLGHRPGE
AVAILTQESIEVYEHFFACMKIAAPRVGLNTGYVWPEMLHVLKDSEVKFLLLDTRCRHLLAERLGELKALGITLIGYG
AGHGLERDYESLLATAEGEPHWPALAPDDILFVSYTSGTTGVPKGVMLTQEGGVNCILHSLISFGFGPDDVWYMP
AASAWVVVILNAFGLGNGMTTVIPDGGYQLQAYLRDIERFRVTVGMLVPTMLQRAIVEIQTNPVYDLSSLRMVVY
GSSPATPKLIRDARATFKGIKLLQAYAMTEATGGWISYLTDADHEHALREEIELLKSVGRIGIHYDCSIRDESGQPVPI
GQSGEIWLRGNTMMKGYRNLPEATAEAMPDGWLRTNDIGRLDERGYLYLLDRQKFLIITGAVNVFPTTVEAILVE
HPAVEEVAVVGVPHPEWGEAVVAVVVRKPSHRDVTVQALIDFCHGKLSRPETPKHVVFVDELPKTSNAKLKKGEL
KKWLSGGAVPLPWQLEVA

stlB from *Photorhabdus luminescens*: (Uniprot Q7N526)
MGSSHHHHHHSSGLVPRGSHMEKVWLKHYPADVPAEIDPDRYASLVEMFENAVAHYADQPAFINMGEIMTFRK
LEERSRAFAAYLQNGLGLSKGDRVALMMPNLLQYPVALFGVLRAGMIVVNVNPLYTPRELEHQLNDSGTSAIVIVS
NFAHTLEKIVFNTKVKHVILTRMGDQLSRPKGTLVDFAVKYIKRLVPKYNLPDAISFRCAIQKGYRMQYVKPEINGN
DLAFLQYTGGTTGVAKGAMLTHRNILANLEQAKAVYSPLLRVGQELIVTALPLYHIFALMVNCLLLIYLGGCNLLITN
PRDITGTAKELGRYPFTSVTGVNTLFNAWLNNEEFKKLDFSTLRLVVGGGMPVQKAVAEKWAKVTGTNLLEGYGL
TECSPLVSCNPYNSKHYTGSIGFPVSSTEIKLVDDDGNEVEMGQQGELWIRGPQVMAGYWNRPDATEEVLKDG
WVATGDIANVNEQGSIHIVDRKKDMILVSGFNVYPNEVEDVVSAHPKVLESAAIGVSSESSGETVKVFVVRIDPGLT
EDELKTHCRRYLTGYKVPKIIEFRDELPKSNVGKILRRELRDEEEKVRNVA

RpCL from *Rhodopseudomonas palustris*: (Uniprot O07454)
MGSSHHHHHHSSGLVPRGSHMEFDAVLLPPRRAASIAAGLWHDRTINDDLDACVAHCPDKVALTAVRLDGGAV
RRFSYRELATLADRVAVGLNRLGVGRGDVVAMQLPNWWQFTVLYLACSRIGAVLNPLMPIFRERELSFMLKHGD
AKVLVVPKSFRGFDHEAMARSLQPDLPALRTIVVVDGGGADDFDTLLTTPEWEKQPDAAAILQGSRPSPDDITQLI
YTSGTTGEPKGVMHSANTLMANIVPYAQRLALRESDVILMASPMAHQTGFMYGLMMPIMLRASAVLQDIWEPT
KAAELIRTERVTFTMASTPFLTDLTRVVKESGEPVPSLKTFLCAGAPIPGPLVEQAQAGLGAKIVSAWGMTENGAVT
LIKLDDDDKLASTTDGCPLPGVEVKVIDGDGKTLPPNQIGRLVVRSCSNFGGYLKRPHWNGTDADGWFDTGDLAY
MTADGYIRISGRSKDVIIRGGENIPVVEVEALLYKHPAVAQVAIVAYPDERLGERACAVVVPKTGASIDFAAMVEFLK
AQKLALQYIPERLVVRDAMPATPSGKIQKFRLREMLQHNDL

LcCL from *Leptothrix cholodnii*: (Uniprot B1Y4L8)
MGSSHHHHHHSSGLVPRGSHMNFDTVLLPPRRAASVAAGHWFDRTINDDLDACVAACPDKVALTAVQVESGEV
RRFTYRELAAMADRVAVGLSRLGVGRNDVVAMQLPNGWQFTVVYLACSRIGAVVNPLMHIFRERELTFMLGHG
EAKVLIVPKTFRGFDHERMVDTIRPDLPKLQQVVVVGGSGANSFEALLCGPAWENEPDAHEVLTRSRPGPDDVTQ
LIYTSGTTGEPKGVMHTANTVMANIIPYAERLHLGSDDVVLMASPMAHQTGFMYGLMMPIMLRASAVLQDLW
DARRAVELIRSEGATFTMASTPFLSDLAKTVAETGTSVPTLRTFLCAGAPIPGALVEQARKVLGTKIVSAWGMTENG
AVTLIKLDDDDQRAFTTDGCPLPGVELKVVDADGAELPAGQAGKLLVRAASNFGGYLHRPQWNGTDADGWFDT

GDLARIDAQGYIRISGRSKDVIIRGGENIPVVEVEALLYRHPAVAQVAIVAYPDERLGERACAFITTKPGQSLDFAGM
VEFLKAQKLAIQYIPERLVVRDALPSTPSGKLQKFKLREMVRDGSI

ArCL from *Acinetobacter radioresistens*: (Uniprot C6RPQ0)
MGSSHHHHHHSSGLVPRGSHMDFDAVLIPARKEAMLKQGYWLNQTILDFLRSAVEKNPDKTALVSVKVENQTEQ
TFSYQQLWNMTNKIALGLKQLGIEKNDVVSCQLPNWWEFTLLYLACSRIGAVLNPLMPIFRERELEFMLKRGESKV
FVVPKTFRNFNHEQLANQLQNKLDSLKHVVVVNGEGENNFDHLLLNHSLEQDASAVAELDNMESGPDDITQLIFT
SGTTGEPKGVMHTANTLFSNIVPYAERLHLTENDVVLMASPMAHQTGFMYGLMMPIQLNTKVVLQDVWDVAK
AVDLIHQHQVNFTMASTPFLNDLSNTVAEQHDKVDSLKIFLCAGAPIPGPLVQKARETLGVKVISAWGMTECGAV
TLTRPEDKDERSFNTDGIALPGVEIKIVDKKGQSKTVNEAGRLMIRSCSSFGGYLKRPDLNDTNVEGWFDTGDIAYQ
DEQGYIRICGRKKDVIIRGGENIPVAEIESLLYKHPNIAVVALVAYADERLGERACAIIKLKDPAQLLSLNELVDFLKTH
NLAIQYIPERLEIWEEIPMTPSGKIQKFKLRELLQ

MsCL from *Mycobacterium smegmatis*: (Uniprot A0QZQ6)
MGSSHHHHHHSSGLVPRGSHMTAPTERRLANVLNGQYTPIDDDTAANWRAAGWWENRSIRSLLADAAQAHPD
RIALVGRRADGGRVARTYQEFDRNANQVASVLASLGVRPDDAVVVMLPNWVEYPEFLFGINELGAIYAGIPVAYG
DQQAAAILRRSRARVLVIPRRWRGNNILEQSRRLRDQIPTLQQVIVLDDDGTDLRDGESLWSDHAHVAARQFPPP
DPGQICYLGFTSGTTGEPKGAMHSHNTLIYSARRQAEHIGTEAFGEPVVNLVASPMGHHTGYVWGGVFTVMLAG
TAVHVDRWDPTWGAQVVREEGVTTFFGAPTFLQDIIRTELAGDPACPLRCMVVAGAPVPRNLPAQAAEALGAYV
APAWGMTECSILTSCTPDEPDAILRTDGSVFAGSEVKIVDDTGAAVAAGVVGDLLMRGPGVVYGYYDRPDATRD
AYLPGLWFKTGDRADVDENGWLRLRGRSKDIIIRGGENIPVTDVESAIFDHPDVLNAAVIGLPDERLGERVCAVLVT
KSGCPELTVDTLGEYLLGQGLSKHYLPEKVVHLDELPMTPSGKIQKFKLREQYS

N-acyltransferase panel
Genes were synthesized at Genscript and the DNA sequences were codon optimised for *E. coli* expression. The genes were cloned into pCDF-Duet vectors using BamHI and HindIII restriction sites so that the expressed protein contains an N-terminal 6-His tag.

01StAT from *Salmonella typhimurium*: (Uniprot Q00267)
MGSSHHHHHHSQDPMTSFLHAYFTRLHCQPLGVPTVEALRTLHLAHNCAIPFENLDVLLPREIQLDETALEEKLLYA
RRGGYCFELNGLFERALRDIGFNVRSLLGRVILSHPASLPPRTHRLLLVDVEDEQWIADVGFGGQTLTAPLRLQAEIA
QQTPHGEYRLMQEGSTWILQFRHHEHWQSMYCFDLGVQQQSDHVMGNFWSAHWPQSHFRHHLLMCRHLP
DGGKLTLTNFHFTRYHQGHAVEQVNVPDVPSLYQLLQQQFGLGVNDVKHGFTEAELAAVMAAFDTHPEAGK

02MsAT from *Mycobacterium smegmatis*: (Uniprot O86309)
MGSSHHHHHHSQDPMAMDLGGYLTRIGLDGRPRPDLGTLHAIVAAHNRSIPFENLDPLLGIPVADLSAEALFAKL
VDRRRGGYCYEHNGLLGYVLEELGFEVERLSGRVVWMRADDAPLPAQTHNVLSVAVPGADGRYLVDVGFGGQT
LTSPIRLEAGPVQQTRHEPYRLTRHGDDHTLAAQVRGEWQPLYTFTTEPRPRIDLEVGSWYVSTHPGSHFVTGLTV
AVVTDDARYNLRGRNLAVHRSGATEHIRFDSAAQVLDAIVNRFGIDLGDLAGRDVQARVAEVLDT

03BaAT from *Bacillus anthracis*: (Uniprot Q81R98)
MGSSHHHHHHSQDPMMTNLQKEFFKRLKIPAKEITFNDLDEILLNMGMILPYENLDIMAGTIKNISKNNLVEKLLI
QKRGGLCYELNSLLYYFLMDCGFQVYKVAGTVYDLYDNKWKPDDGHVIIILHHNKKDYVIDAGFASHLPLHPVPFS
GEVISSQTGEYRIRKRTTQKGTHILEMRKGANGESTNFLQSEPSDEWKIGYAFTLDPIDEQKVNNIQKVIVEHKESPF
NKGAITCKLTNYGHISLTNKNYTETFKGTKNKRPIESKDYARILRESFGITQVKYVGKTLERG

04RlAT from *Rhizobium loti:* (Uniprot Q98D42)
MGSSHHHHHHSQDPMNDAPPFDLDAYLARIGYTGPRNASLDTLKALHFAHPQAIPFENIDPFLGRPVRLDLAALQ
DKIVLGGRGGYCFEHNLLFMHALKALGFEVGGLAARVLWGQSEDAITARSHMLLRVELDGRTYIADVGFGGLTLTA

PLLLEPGREQKTPHEPFRIVEADDHFRLQAAIGGDWRSLYRFDLQPQYEVDYSVTNYFLSTSPTSHFLSSVIAARAAP
DRRYALRGNRLSIHHLGGRTEQTEIATAADLADTLQGLLGIIIPDRTAFEAKVRETKIVETNA

05PaAT from *Pseudomonas aeruginosa*: (Uniprot Q9HUY3)
MGSSHHHHHHSQDPMTPLTPEQTHAYLHHIGIDDPGPPSLANLDRLIDAHLRRVAFENLDVLLDRPIEIDADKVFA
KVVEGSRGGYCFELNSLFARLLLALGYELELLVARVRWGLPDDAPLTQQSHLMLRLYLAEGEFLVDVGFGSANPPRA
LPLPGDEADAGQVHCVRLVDPHAGLYESAVRGRSGWLPLYRFDLRPQLWIDYIPRNWYTSTHPHSVFRQGLKAAI
TEGDLRLTLADGLFGQRAGNGETLQRQLRDVEELLDILQTRFRLRLDPASEVPALARRLAGLISA

06MtAT from *Mycobacterium tuberculosis*: (Uniprot I6XHK1)
MGSSHHHHHHSQDPMALDLTAYFDRINYRGATDPTLDVLQDLVTVHSRTIPFENLDPLLGVPVDDLSPQALADKL
VLRRRGGYCFEHNGLMGYVLAELGYRVRRFAARVVWKLAPDAPLPPQTHTLLGVTFPGSGGCYLVDVGFGGQTP
TSPLRLETGAVQPTTHEPYRLEDRVDGFVLQAMVRDTWQTLYEFTTQTRPQIDLKVASWYASTHPASKFVTGLTA
AVITDDARWNLSGRDLAVHRAGGTEKIRLADAAAVVDTLSERFGINVADIGERGALETRIDELLARQPGADAP

07NfAT from *Nocardia farcinica*: (Uniprot Q5YYQ3)
MGSSHHHHHHSQDPMSKPDDPAYHWNGAELDLDAYLARIGFAGERAPTLATLRELVYRHTTAIPFENLEAVLGRP
VRLDLATLQDKLVHSRRGGYCYENAGLFAAALERLGFGVTGHTGRVTMGAGGLRPATHALLRVTTADDDRVWM
CDVGFGRGPLRPYELRPQPDEFTLGDWRFRLERRTGELGTDLWVLHQFGRDGWVDRYTFTTAPQYRIDFEVGNH
FVSTSPRSPFTTRPFLQRFHSDRHHVLDGLTLITERPDGSADIRALTPGELPEVINELFDIELPGPDLDALTTGSWLERV
AAGTP

08MaAT from *Mycobacterium abscessus*: (Uniprot B1ME52)
MGSSHHHHHHSQDPMWNGDELQLDEYLAFIGFDGDRSPTLETLRRLQRGHVLNIKWENLDAVLHKHVALDIPAV
QAKLLRSPRGGYCYEHVALFGAVLQRLGFDFYGIQGRVQMGATTIRPATHGMLVVRLAAEQWLCDVGFGTSPLA
PIRLVDEAVVADESWTYRLRRGEVTPGADGWTLSEAAGDGSEPGWLSRHTFVLEPQYPIDYRAASYFVASSPHSPF
STRAFVQQISPDHAYILDHRELHEIQPGVGRKTRQLTPAEVLATLREIFGIELGADDSTLLLERLAEQ

09MmAT from *Mycobacterium marinum*: (Uniprot B2HIZ6)
MGSSHHHHHHSQDPMALDLTGYLDRINYRGATDPTLDVLRDLVSAHTGAIAFENLDPLMGVPVDDLSAEALADK
LVDRRRGGYCYEHNGLIGYVLAELGYRVRRLAGRVVWLAPPDAPTPAQTHTVLAVTFPGCQGPYLVDVGFGGMT
PTAPLRLETGTVQQTALEPYRLDDRGDGLVLQAMVRDEWQALYEFSTLTRPQVDLRVGSWFVSTHPTSHFVTGL
MAATVADDARWNLMGRNLAIHRRGGTEKILLEDAAAVVDTLGDRFGINVADVGERGRLEARIDKVCFGAENR

10HvAT from *Hordeum vulgare*: (Uniprot A9ZPJ7)
MGSSHHHHHHSQDPMKITVHSSKAVKPEYGACGVAPGCTADVVPLTVLDKANFDTYISVIYAFHPPAPPNAVLEA
GLGRALVDYREWAGRLGVDANGDRAILLNDAGARFVEATADVALDSVMPLKPTSEVLSLHPSGDDGPEELMLIQV
TRFACGSLVVGFTAQHLVSDGRATSNFFLAWSQATRGVAVDPVPVHDRASFFHPREPLHVEYEHRGVEFKPYEKA
HDVVCGADGDEDEVVVNKVHFSREFISKLKAQASAGAPRPCSTLQCVVAHLWRSMTMARGLDGGETTSVAIAVD
GRARMSPQVPDGYTGNVILWARPTTTAGELVDRPVKHAVELISREVARINDGYFKSFIDFANSGAVEKERLVATAD
AADMVLSPNIEVDSWLRIPFYDMDFGGGRPFFFMPSYLPVEGLLILLPSFLGDGSVDAYVPLFSRDMNTFKNCCYSL
D

11AtAT from *Arabidopsis thaliana*: (Uniprot Q9FNP9)
MGSSHHHHHHSQDPMALKVIKISRVSPATASVDPLIVPLSFFDLQWLKLNPTEQVFFYKLTESSSSRDVFYSSILPKLE
RSLSLILTHFRLFTGHLKWDSQDPKPHLVVLSGDTLSLTVAETDADFSRISGRGLRPELELRPLIPELPIYSDSGAVVSL
QVTLFPKQGFCIGTTAHHVVLDGKTAEKFNKAWAHTCKHGTIPKILPTVLDRSVVNVPAGLEQKMLELLPYLTEDD
KENGRTLKLPPVKEINAKDNVLRITIEISPENIEKLKERAKKESTRAELHLSTFVVTFAHVWTCMVKARSGDPNRPVRF
MYAADFRNRLEPPVPVTYFGTCVLAMDFYKYKAKEFMGEDGFVNTVEILSDSVKRLASQGVESTWKVYEEGTKTM
KWGTQLLVVNGSNQIGMYETDFGWGRPIHTETMSIYKNDEFSMSKRRDGIGGVEIGISLKKLEMDTFLSLFYKWIG
N

12HvAT from *Hordeum vulgare*: (Uniprot A9ZPJ6)
MGSSHHHHHHSQDPMKITVHSSKAVKPEYGACGLAPGCTADVVPLTVLDKANFDTYISVIYAFHAPAPPNAVLEA
GLGRALVDYREWAGRLGVDASGGRAILLNDAGARFVEATADVALDSVMPLKPTSEVLSLHPSGDDGPEELMLIQV
TRFACGSLVVGFTTQHIVSDGRSTGNFFVAWSQATRGAAIDPVPVHDRASFFHPREPLHVEYEHRGVEFKPCEKAH
DVVCGADGDEDEVVVNKVHFSREFISKLKAHASAGAPRPCSTLQCVVAHLWRSMTMARGLDGGETTSVAIAVDG
RARMSPQVPDGYTGNVILWARPTTTAGELVTRPVKHAVELISREVARINDGYFKSFIDFANSGAVEKERLVATADA
ADMVLSPNIEVDSWLRIPFYDMDFGGGRPFFFMPSYLPVEGLLILLPSFLGDGSVDAYVPLFSRDMNTFKNCCYSL
D

13AtAT from *Arabidopsis thaliana*: (Uniprot Q8GYW8)
MGSSHHHHHHSQDPMANQRKPILPLLLEKKPVELVKPSKHTHCETLSLSTLDNDPFNEVMYATIYVFKANGKNLD
DPVSLLRKALSELLVHYYPLSGKLMRSESNGKLQLVYLGEGVPFEVATSTLDLSSLNYIENLDDQVALRLVPEIEIDYES
NVCYHPLALQVTKFACGGFTIGTALTHAVCDGYGVAQIIHALTELAAGKTEPSVKSVWQRERLVGKIDNKPGKVPG
SHIDGFLATSAYLPTTDVVTETINIRAGDIKRLKDSMMKECEYLKESFTTYEVLSSYIWKLRSRALKLNPDGITVLGVAV
GIRHVLDPPLPKGYYGNAYIDVYVELTVRELEESSISNIANRVKKAKKTAYEKGYIEEELKNTERLMRDDSMFEGVSD
GLFFLTDWRNIGWFGSMDFGWNEPVNLRPLTQRESTVHVGMILKPSKSDPSMEGGVKVIMKLPRDAMVEFKRE
MATMKKLYFGDTN

14AtAT from *Arabidopsis thaliana*: (Uniprot O80467)
MGSSHHHHHHSQDPMPIHIGSSIPLMVEKMLTEMVKPSKHIPQQTLNLSTLDNDPYNEVIYKACYVFKAKNVADD
DNRPEALLREALSDLLGYYYPLSGSLKRQESDRKLQLSCGGDGGGVPFTVATANVELSSLKNLENIDSDTALNFLPVL
HVDIDGYRPFALQVTKFECGGFILGMAMSHAMCDGYGEGHIMCALTDLAGGKKKPMVTPIWERERLVGKPEDD
QPPFVPGDDTAASPYLPTDDWVTEKITIRADSIRRLKEATLKEYDFSNETITTFEVIGAYLWKSRVKALNLDRDGVTV
LGLSVGIRNVVDPPLPDGYYGNAYIDMYVPLTAREVEEFTISDIVKLIKEAKRNAHDKDYLQEELANTEKIIKMNLTIK
GKKDGLFCLTDWRNIGIFGSMDFGWDEPVNIVPVVPSETARTVNMFMRPSRLESDMVGGVQIVVTLPRIAMVKF
KEEMEALE

15SpAT from *Schizosaccharomyces pombe*: (Uniprot O59735)
MGSSHHHHHHSQDPMGNNTSRRSQSQKFKNIPSISAGSFSTMPVQHRGRRRRSKSIVGRAAQNAVLRSKEKTWI
VPLILLTLLVGWYFVNPNGYIKYGIFLSYPIPGTNPAQYGKGRLDIAFCLFYALFFTFCREFIMQEIIARIGRHFNIRAPA
KLRRFEEQAYTCLYFTVMGSWGLYVMKQTPMWFFNTDAFWEEYPHFYHVGSFKAFYLIEAAYWIQQALVLILQLE
KPRKDFKELVVHHIITLLLIGLSYYFHFTWIGLAVFITMDTSDIWLALSKCLNYVNTVIVYPIFVIFVFVWIYMRHYLNF
KIMWAVWGTMRTINSFDLDWAAEQYKCWISRDVTLILLTALQLVNIYWLILILRIGYRAFTTNDTHDERSEDEDEEV
SDEKSSAKKND

16ScAT from *Saccharomyces cerevisiae*: (Uniprot A6ZSP9)
MGSSHHHHHHSQDPMTSATDKSIDRLVVNAKTRRRNSSVGKIDLGDTVPGFAAMPESAASKNEAKKRMKALTG
DSKKDSDLLWKVWFSYREMNYRHSWLTPFFILVCVYSAYFLSGNRTESNPLHMFVAISYQVDGTDSYAKGIKDLSF
VFFYMIFFTFLREFLMDVVIRPFTVYLNVTSEHRQKRMLEQMYAIFYCGVSGPFGLYIMYHSDLWLFKTKPMYRTYP
DITNPFLFKIFYLGQAAFWAQQACVLVLQLEKPRKDYKELVFHHIVTLLLIWSSYVFHFTKMGLAIYITMDVSDFFLSL
SKTLNYLNSVFTPFVFGLFVFFWIYLRHVVNIRILWSVLTEFRHEGNYVLNFATQQYKCWISLPIVFVLIAALQLVNLY
WLFLILRILYRLIWQGIQKDERSDSDSDESAENEESKEKCE

17EcAT from *Escherichia coli*: (Uniprot P0A951)
MGSSHHHHHHSQDPMPSAHSVKLRPLEREDLRYVHQLDNNASVMRYWFEEPYEAFVELSDLYDKHIHDQSERRF
VVECDGEKAGLVELVEINHVHRRAEFQIIISPEYQGKGLATRAAKLAMDYGFTVLNLYKLYLIVDKENEKAIHIYRKLG
FSVEGELMHEFFINGQYRNAIRMCIFQHQYLAEHKTPGQTLLKPTAQ

18BsAT from *Bacillus subtilis*: (Uniprot P21340)
MGSSHHHHHHSQDPMSVKMKKCSREDLQTLQQLSIETFNDTFKEQNSPENMKAYLESAFNTEQLEKELSNMSSQ
FFFIYFDHEIAGYVKVNIDDAQSEEMGAESLEIERIYIKNSFQKHGLGKHLLNKAIEIALERNKKNIWLGVWEKNENAI
AFYKKMGFVQTGAHSFYMGDEEQTDLIMAKTLI

19CpAT from *Cryptosporidium parvum*: (Uniprot Q5CPU3)
MGSSHHHHHHSQDPMISSFEVRKATIDDYFELRNLICDVTRCTETLSREQAEERFRYNTYHPYCLVDTENGRIVGYA
GFYIIPHLGRKNDSRIEHVIISKEYRNRGLGRLLCKQIIEDAKNKFNCGRIDLTVESHIAKKLYSSLEFEKVNTEVMRNSF
LDLTPKSD

20HsAT from *Homo sapiens*: (Uniprot Q14032)
MGSSHHHHHHSQDPMIQLTATPVSALVDEPVHIRATGLIPFQMVSFQASLEDENGDMFYSQAHYRANEFGEVDL
NHASSLGGDYMGVHPMGLFWSLKPEKLLTRLLKRDVMNRPFQVQVKLYDLELIVNNKVASAPKASLTLERWYVA
PGVTRIKVREGRLRGALFLPPGEGLFPGVIDLFGGLGGLLEFRASLLASRGFASLALAYHNYEDLPRKPEVTDLEYFEE
AANFLLRHPKVFGSGVGVVSVCQGVQIGLSMAIYLKQVTATVLINGTNFPFGIPQVYHGQIHQPLPHSAQLISTNAL
GLLELYRTFETTQVGASQYLFPIEEAQGQFLFIVGEGDKTINSKAHAEQAIGQLKRHGKNNWTLLSYPGAGHLIEPPY
SPLCCASTTHDLRLHWGGEVIPHAAAQEHAWKEIQRFLRKHLIPDVTSQL

21BcAT from *Beutenbergia cavernae*: (Uniprot C5BVE1)
MGSSHHHHHHSQDPMELTAEPTHPTQDARLRLVVTGADPGEPVTVEAVEAGRSASATFVADAGGRVDLQAQA
PVDGTYRGVDAMGLFWSMDGPPVAADPLAPLSVRLTATTQVGNRAVLVVDRLRRPPGVTRTEIRDRGVVGTLFT
PPGRGPRPGVVLLGGAEGGMHERDAALLAGHGFAALALAYFGLPGLPTGLVDVPLEYFGTAIELLRGMPDVGPAV
GVIGASRGGEAALLVGATFPEVAAVVSTVGSGLVTQGIPPHPSLLQILTDARASWTFEGRPLPYLPHTLTPEFVSDVE
RGAPVSLAGTYRPDLADPALVTEATIPAERVRGAVLFVTAGRDQGWPCLELSRYAMDRMNDYEHAHYADAGHAI
APPPYTSSVEAPSPGPGVTFAAGSADPQANAAARADAWRRTLAFLAEHLGAATPSPVSADALEGHR

22HsAT from *Homo sapiens*: (Uniprot Q9UHE5)
MGSSHHHHHHSQDPMAPCHIRKYQESDRQWVVGLLSRGMAEHAPATFRQLLKLPRTLILLLGGPLALLLVSGSWL
LALVFSISLFPALWFLAKKPWTEYVDMTLCTDMSDITKSYLSERGSCFWVAESEEKVVGMVGALPVDDPTLREKRL
QLFHLFVDSEHRRQGIAKALVRTVLQFARDQGYSEVILDTGTIQLSAMALYQSMGFKKTGQSFFCVWARLVALHTV
HFIYHLPSSKVGSL

23DmAT from *Drosophila melanogaster*: (Uniprot Q94521)
MGSSHHHHHHSQDPMEVQKLPDQSLISSMMLDSRCGLNDLYPIARLTQKMEDALTVSGKPAACPVDQDCPYTIE
LIQPEDGEAVIAMLKTFFFKDEPLNTFLDLGECKELEKYSLKPLPDNCSYKAVNKKGEIIGVFLNGLMRRPSPDDVPEK
AADSCEHPKFKKILSLMDHVEEQFNIFDVYPDEELILDGKILSVDTNYRGLGIAGRLTERAYEYMRENGINVYHVLCS
SHYSARVMEKLGFHEVFRMQFADYKPQGEVVFKPAAPHVGIQVMAKEVGPAKAAQTKL

24LlAT from *Lactococcus lactis*: (Uniprot G6FA50)
MGSSHHHHHHSQDPMNEITYTQIMPNDIDDVIRIERATFSENEALSVESMIERINLIPDSFIAARNSEGTVVGYVSG
PVTEGRYLDDESFERTEANPKTGGFQKIISLTVDPDYQGLGIATNLLLLLEKEAKSKKRLGISLTCHDYLVPYYEKHGFT
NEGLSESKFGGETWYNMVMEF

25DmAT from *Drosophila melanogaster*: (Uniprot Q95SX8)
MGSSHHHHHHSQDPMAQFTLYNKHSAPPSSESTRVDCEHVPLCSINDVQLRFLVPDDLTEVRQLCQEWFPIDYPL
SWYEDITSSTRFFALAAVYNLAIIGLIVAEIKPYRNVNKEVIANMSDSDELYTRLSGFPMQDKGILPDSMGRSADVGY
ILSLGVHRSHRRNGIGSLLLDALMNHLTTAERHSVKAIFLHTLTTNQPAIFFYEKRRFTLHSFLPYYYNIRGKGKDGFTY
VNYINGGHPPWTLLDHIKHYASMVRHTSSLCAWLAGRVQQVVRWFYHKLLTRFNFIE

26AtAT from *Arabidopsis thaliana*: (Uniprot O80438)
MGSSHHHHHHSQDPMEKEMEDKEEFDEGEIEYTSYAGEHHLPLIMSLVDQELSEPYSIFTYRYFVYLWPQLCFLAF
HKGKCVGTIVCKMGDHRQTFRGYIAMLVVIKPYRGRGIASELVTRAIKAMMESGCEEVTLEAEVSNKGALALYGRL
GFIRAKRLYHYYLNGMDAFRLKLLFPKPRVPQIPSQVQTQQEYETFPRPRVP

27BsAT from *Bacillus subtilis*: (Uniprot O34981)
MGSSHHHHHHSQDPMKMMDANEIISFIQNSTKSTPVKVYVKGELEGINFGESAKAFINGNTGVVFGEWSEIQTAI
EENQSKIEDYVVENDRRNSAIPMLDLKNIKARIEPGAIIRDQVEIGDNAVIMMGASINIGSVIGEGTMIDMNVVLGG
RATVGKNCHIGAGSVLAGVIEPPSAKPVVIEDDVVIGANAVVLEGVTVGKGAVVAAGAIVVNDVEPYTVVAGTPAK
KIKDIDEKTKGKTEIKQELRQL

28BaAT from *Bacillus anthracis*: (Uniprot Q81MQ2)
MGSSHHHHHHSQDPMKMMDANEIISFIQKSEKKTPVKVYIKGDLKEVTFPETVQAFVNKKSGVLFGEWSEIKTILD
ENSKYIVDYVVENDRRNSAIPMLDLKGIKARIEPGAIIRDHVEIGDNAVIMMNATINIGAVIGEGSMIDMNAVLGG
RATVGKNCHVGAGAVLAGVIEPPSAKPVIVEDDVVIGANVVVLEGVTVGKGAVVAAGAVVTEDVPPYTVVAGTPA
RVIKEIDEKTKAKTEIKQELRQLNPEK

29EfAT from *Enterococcus faecalis*: (Uniprot Q836H8)
MGSSHHHHHHSQDPMDAYEIIQYIGDAKKQTLVKVTLKGQLKEVTFPETIKVFNNCKTGTLFGDWADVKPFLEAN
KEKIEDYVVENDARNSAIPFLDLKDINARIEPGALIREKVEIGDQAVIMMGAILNIGAVVGAGTMIDMGAVLGGRAT
VGKHCHIGAGTVLAGVIEPPSAAPVVIENEVVIGANAVVLEGVRVGEGAVVAAGAVVVEDVPAHTVVAGVPAKVI
KQIDDKTKSKTEILEELRKL

30SaAT from *Staphylococcus aureus*: (Uniprot Q7A2S0)
MGSSHHHHHHSQDPMVQHLTAEEIIQYISDAKKSTPIKVYLNGNFEGITYPESFKVFGSEQSKVIFCEADDWKPFYE
AYGSQFEDIEIEMDRRNSAIPLKDLTNTNARIEPGAFIREQAIIEDGAVVMMGATINIGAVVGEGTMIDMNATLGG
RATTGKNVHVGAGAVLAGVIEPPSASPVIIEDDVLIGANAVILEGVRVGKGAIVAAGAIVTQDVPAGAVVAGTPAK
VIKQASEVQDTKKEIVAALRKLND

31NtAT from *Nicotiana tabacum*: (Uniprot P80969)
MGSSHHHHHHSQDPMATTNNKNLTITEKVYVRVRLANEADISHIYKLFYQIHEYHNYTHLYKATESSLCDLLFKANP
NPLFYGPSVLLLEVSPTPFENTKKDEKFKPVLKTFDLRATVEDKEAEEFKSKSCGDEKEDVFIAGYAFFYANYSCFYDK
AGIYFESLYFRESYRKLGMGGLLFGTVASIAANNGFASVEGIVAVWNKKSYDFYVNMGVEIFDEFRYGKLVGDALQ
KYADKEKV

32NtAT from *Nicotiana tabacum*: (Uniprot Q9SMB8)
MGSSHHHHHHSQDPMATTNNKNLTITEKVYVRVRLANEADISHIYKLFYQIHEYHNYTHLYKATESSLCDLLFKANP
NPLFYGPSVLLLEVSPTPFENTKKDEKFKPVLKTFDLRATVEDKEAEEFKSKSCGDEKEDVFIAGYAFFYANYSCFYDK
AGIYFESLYFRESYRKLGMGSLLFGTVASIAANNGFASVEGIVAVWNKKSYDFYVNMGVEIFDEFRYGKLVGDALQK
YADKEKA

33MtAT from *Mycobacterium tuberculosis*: (Uniprot P9WP21)
MGSSHHHHHHSQDPMSTVTGAAGIGLATLAADGSVLDTWFPAPELTESGTSATSRLAVSDVPVELAALIGRDDDR
RTETIAVRTVIGSLDDVAADPYDAYLRLHLLSHRLVAPHGLNAGGLFGVLTNVVWTNHGPCAIDGFEAVRARLRRR
GPVTVYGVDKFPRMVDYVVPTGVRIADADRVRLGAHLAPGTTVMHEGFVNYNAGTLGASMVEGRISAGVVVGD
GSDVGGGASIMGTLSGGGTHVISIGKRCLLGANSGLGISLGDDCVVEAGLYVTAGTRVTMPDSNSVKARELSGSSN
LLFRRNSVSGAVEVLARDGQGIALNEDLHAN

34PaAT from *Pseudomonas aeruginosa*: (Uniprot G3XD76)
MGSSHHHHHHSQDPMSQSLFSLAFGVGTQNRQEAWLEVFYALPLLKPSSEIVAAVAPILGYAAGNQALTFTSQQ
AYQLADALKGIDAAQSALLSRLAESQKPLVATLLAEDAAPSSTAEAYLKLHLLSHRLVKPHAVNLSGIFPLLPNVAWT

NIGAVDLAELAELQLEARLKGKLLEVFSVDKFPKMTDYVVPAGVRIADTARVRLGAYIGEGTTVMHEGFVNFNAGT
EGPGMIEGRVSAGVFVGKGSDLGGGCSTMGTLSGGGNIVISVGEGCLIGANAGIGIPLGDRNIVEAGLYITAGTKV
ALLDEQNALVKVVKARDLAGQPDLLFRRNSQNGAVECKTNKTAIELNEALHAHN

35MaAT from *Methylomicrobium alcaliphilum*: (Uniprot Q4JQJ5)
MGSSHHHHHHSQDPMLPDKTALPIITLSQPTAEVGAQVHRLISKCPPLDPNSMYCNLLQSSHFSETAVAAKIGDEL
VGFVSGYRIPQRPDTLFVWQVAVGEKARGQGLATRMLKAILARPVNQDINRIETTITPNNKASWALFEGLAKKLDT
QIGSAVMFDKTRHFADQHETEMLVKVGPFKAVQA

36PaAT from *Pseudomonas aeruginosa*: (Uniprot A6VCX3)
MGSSHHHHHHSQDPMSASIRDAGVADLPGILAIYNDAVGNTTAIWNETPVDLANRQAWFDARARQGYPILVAS
DAAGEVLGYASYGDWRPFEGFRGTVEHSVYVRDDQRGKGLGVQLLQALIERARAQGLHVMVAAIESGNAASIGL
HRRLGFEISGQMPQVGQKFGRWLDLTFMQLNLDPTRSAP

37HsAT from *Homo sapiens*: (Uniprot P30419)
MGSSHHHHHHSQDPMADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGGLSPANDTGAKKK
KKKQKKKKEKGSETDSAQDQPVKMNSLPAERIQEIQKAIELFSVGQGPAKTMEEASKRSYQFWDTQPVPKLGEVV
NTHGPVEPDKDNIRQEPYTLPQGFTWDALDLGDRGVLKELYTLLNENYVEDDDNMFRFDYSPEFLLWALRPPGW
LPQWHCGVRVVSSRKLVGFISAIPANIHIYDTEKKMVEINFLCVHKKLRSKRVAPVLIREITRRVHLEGIFQAVYTAGV
VLPKPVGTCRYWHRSLNPRKLIEVKFSHLSRNMTMQRTMKLYRLPETPKTAGLRPMETKDIPVVHQLLTRYLKQFH
LTPVMSQEEVEHWFYPQENIIDTFVVENANGEVTDFLSFYTLPSTIMNHPTHKSLKAAYSFYNVHTQTPLLDLMSD
ALVLAKMKGFDVFNALDLMENKTFLEKLKFGIGDGNLQYYLYNWKCPSMGAEKVGLVLQ

38ShAT from *Streptomyces hygroscopicus*: (Uniprot P16426)
MGSSHHHHHHSQDPMSPERRPADIRRATEADMPAVCTIVNHYIETSTVNFRTEPQEPQEWTDDLVRLRERYPWL
VAEVDGEVAGIAYAGPWKARNAYDWTAESTVYVSPRHQRTGLGSTLYTHLLKSLEAQGFKSVVAVIGLPNDPSVR
MHEALGYAPRGMLRAAGFKHGNWHDVGFWQLDFSLPVPPRPVLPVTEI

39SeAT from *Salmonella enteritidis*: (Uniprot Q9R381)
MGSSHHHHHHSQDPMDIRQMNKTHLEHWRGLRKQLWPGHPDDAHLADGEEILQADHLASFIAMADGVAIGF
ADASIRHDYVNGCDSSPVVFLEGIFVLPSFRQRGVAKQLIAAVQRWGTNKGCREMASDTSPENTISQKVHQALGF
EETERVIFYRKRC

40BtAT from *Bos taurus*: (Uniprot Q2KIR7)
MGSSHHHHHHSQDPMFLLQGAQMLQMLEKSLRKSLPMSLKVYGTVMHMNHGNPFNLKALVDKWPDFQTVVI
RPQEQDMKDDLDHYTNTYHVYSEDLKNCQEFLDLPEVINWKQHLQIQSTQSSLNEVIQNLAATKSFKVKRSKNILY
MASETIKELTPSLLDVKNLPVGDGKPKAIDPEMFKLSSVDPSHAAVVNRFWLFGGNERSLRFIERCIQSFPNFCLLGT
EGTPVSWSLMDQTGEMRMAGTLPEYRAQGLVTHAIYQQAQCLLKRGFPVYSHVDPKNQIMQKMSQSLNHVP
MPSDWNQWNCEPL

41GmAT from *Gibberella moniliformis*: (Uniprot D8FSU9)
MGSSHHHHHHSQDPMADRIRYSKSQLEKYYDRIAFPASDRRYDISNLSSEDQRSYLDTLTKQQILTVPFENLTLHYS
WHRTVDVNADHLYDKIVNERRGGYCMENNTFFNTVLLSLGYHTYMVGSRVFNPDADRFGGTSHCLSLVIIDGKTL
AVDVGFGGRNPTEPLEVEHERVHTGSSGFQMRLRYDAIAQNVSNQKLWIYEYRSRDGAEWVPQWCFMDFEVLP
EDIRVFNLSPSKSPSSFFTFKVVSVQFTSEKEDYSDGSARDLNNVGGDVDGAFVIDGELFKYRKGGETKWERTFKSE
DERLAALRKYYGVEVTKENERAIGGTAGAISYRRTGS

42GmAT from *Gibberella moniliformis*: (Uniprot B7SP66)
MGSSHHHHHHSQDPMARLEDPTALTQLPDESARVRYTSSELQDYFETLKFPQRFLDLGNSVLKDPSLARTKENGLP
LLQAITRYHTCNVPFENLVLHYDPHKIVTLDPAELYTKIVTRRRGGRCMENNIFLGTALRSLGYEVRNCGGRVSRAM
SPYPEVRKNQSATYDGWNHMLLLVFLGDEWYGVDVGMGSMGPNLPFPLQDGFESLSIAPREIRIQKRSISETHAT

GPSHATKMWCYDVCYNPAESKKTWTPVYCFTETEFLPQDYEVMSWFTSTNPRSFFTRYITCTKMIMDEDKEVIIG
NLTLFKDTVRETIGSDRKVVKKFETEEERIKGLVEIFDVNLTEEEKNSLPQEKRLA

43GmAT from *Gibberella moniliformis*: (Uniprot D8FSU8)
MGSSHHHHHHSQDPMSAIYSEAQVAGFLKHLQIPQEFYVGNEAILDHAFLKVLHQHMIATVPYDNLTLHYSSHRN
ITLEPQALYQKIVGDGRGRGGYCMESNLFFCYMLRALGFQVYPVGVRVRLRNNGIPFGGYPGWVHIVNIVTLPDN
SRWVMDASFGGDGPTQPMPLVEGAEWHNMGTQTARLIKDFIPGQTELTSGRRLWIYQCRNSPDLPWTSFYAFS
HSVEWLPADFEITNCYTGTSPRSFQTSTVLIVKFLLRESKTSPTGEEIYGKRMLINDVVKENPGGKTKVLKELKTEDER
VEALKEYFDIDLTTEEREAIKGFQTEIKSE

44TcAT from *Taxus canadensis*: (Uniprot Q8LL69)
MGSSHHHHHHSQDPMEKAGSTDFHVKKFDPVMVAPSLPSPKATVQLSVVDSLTICRGIFNTLLVFNAPDNISADP
VKIIREALSKVLVYYFPLAGRLRSKEIGELEVECTGDGALFVEAMVEDTISVLRDLDDLNPSFQQLVFWHPLDTAIEDL
HLVIVQVTRFTCGGIAVGVTLPHSVCDGRGAAQFVTALAEMARGEVKPSLEPIWNRELLNPEDPLHLQLNQFDSIC
PPPMLEELGQASFVINVDTIEYMKQCVMEECNEFCSSFEVVAALVWIARTKALQIPHTENVKLLFAMDLRKLFNPPL
PNGYYGNAIGTAYAMDNVQDLLNGSLLRAIMIIKKAKADLKDNYSRSRVVTNPYSLDVNKKSDNILALSDWRRLGF
YEADFGWGGPLNVSSLQRLENGLPMFSTFLYLLPAKNKSDGIKLLLSCMPPTTLKSFKIVMEAMIEKYVSKV

45SpAT from *Schizosaccharomyces pombe*: (Uniprot Q09927)
MGSSHHHHHHSQDPMKDPNTIPPWRCTDFNAWCIAVDKSTNVKNKEELLSTLTYFINYEIEMGQTYPIDIKMTRN
EAEDFFFKFCTVICVPVESETSPAPDLATASIDWKTSLLGAFYIKPNYPGRCSHICNGGFLVSPSHRSKGIGRNLANAY
LYFAPRIGFKSSVFNLVFATNIKSIRLWERLNFTRAGIIKDAGRLKGHEGYVDAYIYQYHFPSLEDALK

46LmAT from *Leishmania major*: (Uniprot Q4Q5S8)
MGSSHHHHHHSQDPMSRNPSNSDAAHAFWSTQPVPQTEDETEKIVFAGPMDEPKTVADIPEEPYPIASTFEWW
TPNMEAADDIHAIYELLRDNYVEDDDSMFRFNYSEEFLQWALCPPNYIPDWHVAVRRKADKKLLAFIAGVPVTLR
MGTPKYMKVKAQEKGEGEEAAKYDEPRHICEINFLCVHKQLREKRLAPILIKEATRRVNRTNVWQAVYTAGVLLPT
PYASGQYFHRSLNPEKLVEIRFSGIPAQYQKFQNPMAMLKRNYQLPSAPKNSGLREMKPSDVPQVRRILMNYLDS
FDVGPVFSDAEISHYLLPRDGVVFTYVVENDKKVTDFFSFYRIPSTVIGNSNYNLLNAAYVHYYAATSIPLHQLILDLLI
VAHSRGFDVCNMVEILDNRSFVEQLKFGAGDGHLRYYFYNWAYPKIKPSQVALVML

47AtAT from *Arabidopsis thaliana*: (Uniprot Q9ZV05)
MGSSHHHHHHSQDPMAPPTAAPEPNTVPETSPTGHRMFSRIRLATPTDVPFIHKLIHQMAVFERLTHLFVATESG
LASTLFNSRPFQAVTVFLLEISPSPFPTTHDASSPDFTPFLETHKVDLPIEDPDREKFLPDKLNDVVVAGFVLFFPNYPS
FLAKQGFYIEDIFMREPYRRKGFGKLLLTAVAKQAVKLGVGRVEWIVIDWNVNAINFYEQMGAQVFKEWRLCRLT
GDALQAIDKLNI

48CaAT from *Capsicum annuum*: (Uniprot Q5D8C0)
MGSSHHHHHHSQDPMASAISETITTNGPSENNNLTITGKIHTRVRLATKSDLHHIYQLFYQIHAYHNFTHLYKATES
SLGDLLFKENPLPLFYGPSVLLLEVSPTPFTQPKNNKDEGFKPVLTTFNLKFPVVEGQVEEFQSKYDDGNDKRDVFIA
GYAFFYANYSCFYDKPGFYFESLYFRESYRKLGMGRLLFGTVASIAANNGFVSVEGIVAVWNKKSYDFYIDMGVEIF
DEFRYGKLHGENLQKYADKQKNEGGNC

49StAT from *Solanum tuberosum*: (Uniprot Q8H9D9)
MGSSHHHHHHSQDPMAPAPQQPTPSETIITDASSENNNVTITGKIYTRVRLATKSDLSHIYQLFYQIHVYHNFTHLY
KATESSLEGLLFKENPLPLFYGPSVLLLEVSPTPFNEPKNTTDEGFNPVLTTFDLKFPVVEGQVEEFRSKYDDKSDAYIA
GYAFFYANYSCFNDKPGFYFESLYFRESYRKLGMGKLLFGTVSSIAADNGFVSVDGIVAVWNKKSYDFYINMGVEIF
DEFRYGKLHGENLQKYADKGKIEEETC

50EcAT from *Escherichia coli*: (Uniprot P0AE37)
MGSSHHHHHHSQDPMMVIRPVERSDVSALMQLASKTGGGLTSLPANEATLSARIERAIKTWQGELPKSEQGYVF
VLEDSETGTVAGICAIEVAVGLNDPWYNYRVGTLVHASKELNVYNALPTLFLSNDHTGSSELCTLFLDPDWRKEGN
GYLLSKSRFMFMAAFRDKFNDKVVAEMRGVIDEHGYSPFWQSLGKRFFSMDFSRADFLCGTGQKAFIAELMPKH
PIYTHFLSQEAQDVIGQVHPQTAPARAVLEKEGFRYRNYIDIFDGGPTLECDIDRVRAIRKSRLVEVAEGQPAQGDF
PACLVANENYHHFRVVLVRTDPATERLILTAAQLDALKCHAGDRVRLVRLCAEEKTA

51EcAT from *Escherichia coli*: (Uniprot P76112)
MGSSHHHHHHSQDPMSIRFARKADCAAIAEIYNHAVLYTAAIWNDQTVDADNRIAWFEARTLAGYPVLVSEENG
VVTGYASFGDWRSFDGFRHTVEHSVYVHPDHQGKGLGRKLLSRLIDEARDCGKHVMVAGIESQNQASLHLHQSL
GFVVTAQMPQVGTKFGRWLDLTFMQLQLDERTEPDAIG

52PlAT from *Pseudoalteromonas luteoviolacea*: (Uniprot V4HKD0)
MGSSHHHHHHSQDPMSEKLDSYKLTQDHQSWTSKPATLEEWQIVNEWALAEKWDLGLGDTECFFNIDDQGFYL
GYVNGEPVASVSVVNYSDEYAYAGFYLVAPGARGKGYGLRLSYDAFHHCDKRSVGLDGMPEQEENYKKGGFVTH
YETSRLVGVHNQQVDAPQGVQNISADNIEAVIEFDAQITGYSRAALLQNWFSGEGRHGFLIDSGDGVLGVVGIRR
STDGYRLGPLYAENQAVSEKLFAMAMAQVPLGTQVTIDAPTLDLGFINTLKALGYEEIFHTFRMYRGTEPQGEKHKI
QAIASLELG

53PhAT from *Photobacterium halotolerans*: (Uniprot A0A0F5VBE7)
MGSSHHHHHHSQDPMDIVKNKTETETSNSVWRVESANLAEWQAVIAWAKDEGWDMGIGDANAFFEVDDQG
FFIGYLGKEPVAAMSVVNYSSRFSFLGHYLVSPAYRGQGYGLKLCQSAFYHGGERCMGLDGMPAQVDNYAKWGF
AGDRYNLRMVGQVQQDYHCPDHIALVQAADIDALIAYDQSCTGIDRSALLKHWFIGETRYGFICRSGQEVTGIIGV
RQSQEGYRVGPLFANSPDLVEPLFIAALSAVPQGGKVTLDVPEKADRTLISLAETHGFESIFQTLRMYRGEPPKEQEH
KVWCIASLELG

54SlAT from *Solanum lycopersicum*: (Uniprot Q8RXB6)
MGSSHHHHHHSQDPMASSLSETITTDASSENNNVTITGKIYTRLRLATKSDLSHIYQLFYQIHAYHNNTHLYKATESS
LANLLFKENPLPLFYGPSVLLLEVSPTPFNEPTNEGFKPVLTTFDLKFPVVEGQVEEFRSKYDDKSDVYIAGYAFFYVN
YSCFSDKPGFYFESLYFRESYRKLGMGSLLFGTVASIAANNGFVSVEGIVAVWNKKSYDFYVNMGVEIFDEFRYGKL
HGENLQKYANDKEKNDGGN

55SlAT from *Solanum lycopersicum*: (Uniprot Q8RXB8)
MGSSHHHHHHSQDPMAPALEQAITSDASSDVTITGKIYTRVRLATKSDLSHIYRLFYQIHEYHNYTHLYKATESSLAN
LLFKENPLPLFYGPSVLLLEVSPTPFDEPKNTTDEGFKPVLTTFDLKFPVVEGEVEEFRSKYDDKSDVYIAGYAFFYANY
SCFYDKPGFYFESLYFRESYRKLGMGSLLFGTVASIAANNGFVSVEGIVAVWNKKSYDFYVNMGVEIFDEFRYGKLH
GENLQKYAHN

56SlAT from *Solanum lycopersicum*: (Uniprot Q8RXB7)
MGSSHHHHHHSQDPMAPTSQQPTPSPSLSDSLTTDASSDVTITGKIYTRVRLATKSDLSHIYKLFYQIHEYHNFTHLY
KATESSLEGLLFKENPLPLFYGPSVLLLEVSPTPFNEPTNQAFKPVLTTFDLKFPVVEGQVEEFRSKHDDKSDAYIAGY
AFFYANYSCFNDKPGFYFESLYFRESYRKLGMGKLLFGTVSSIAANNGFVSVDGIIAVWNKKSYDFYINMGVEIFDEF
RYGKLHGENLQKYAHNKGKTEEETC

57SpAT from *Streptococcus pasteurianus*: (Uniprot F5X5E8)
MGSSHHHHHHSQDPMRKLVKKTIKINLKSESLLADDKFDIMINQLKPNQIIDVHMTLLDYYNINAPMRIGSETPWY
AWAKYQADDNGVVSVNNCISLEGDYVGKIDMGLFFSCRPLKSKKNHLLEDVRDIPVYDSFYVKIEIFSKSKKIAEKIFK
RYYQLPEISHETIISSRYQARLFYPSNAQNLPAIIVFSGSDGRIEKAQNIAQLLASRGFVTLAVAYFGLRGISRYLDRIPLE
IVKENLDYLANLEIVDENRIGLYGRSKGAELSLVAASLFSQIRCLVVNSPSCAILEGLKGRINSKHSSWTYQNKELPYTK
FSIREFFKEKFLGQKFSDYELSSLIKAKEIGANVLMIGEKSDEIWNTEFSIKLLEKQLQKRKNGLVKSILLNNGGHMLTI
AYQPNQRYTKILKENLLSDSVISWKATIDFFKNYL

58AxAT from *Achromobacter xylosoxidans*: (Uniprot E3HPZ9)
MGSSHHHHHHSQDPMAEQPSITVTPERALIDVAREIRVSGFPAYAFITITASMTMCGAPWRSQAVFVAGHDGSL
DLARDCPVSGSYAEPSAMGIVWSMACEDIASVVFPPDRTEPLAIHIEASDGRHTATATLVQEFQADGVTHRAVRE
QVGGMTLSGELYTPAGPGPHPLVIYMNGSSGGVNAPRAALFASRGYQCLALAIFNYEGRPKYLNDMPLEYFQHVL
RWARAELAPRDGFVALSGISRGGETSLLVASHYPELVSAVVAYVPSPVMHGVVSAGAPGTGRDAQVWTKAGEPL
PHLWQDNASANWDAAYASQPPYRQTHAFLSATRDSAAFERARIPVENYPGPVLLISASDDGFWPSTAYSEIVMRR
RQEHGLPTFHHVCMGAGHHVHYPCLPATLISKPHAMSGLLLDAGGTPSANAAGNEGSYLAVLEFLGRVGGVAW

59XbAT from Xenorhabdus bovienii: (Uniprot A0A077PX88)
MGSSHHHHHHSQDPMFVSAFIPQYCWKDILDMSEKSNKKCVFKRVTANEWRSTIFHWANTEQWNMVESDISR
FFNVDPHGFFITYRNGMPISSMSMVNYSDTYTHGGHYIVPPEFRNHVCAGKIWKSSIAYAEHRTISCDGNSHLSSLY
EKRGFVPHYRTFRMSGVITQKVIQPAGVLPITKINIDGVIHYDAECTGVYRGRLLADWFWGEGRYGFCTHSDTGVT
GIIGIRKSTDGYRLGPFHADNADVLETLFRAALAQVPTGEAVTFDAPETDNNQFIPLAKEYGLRELFYTFRMYKGNVI
PKGRLEKLRAIASLELG

60MbAT from *Candidatus Collierbacteria bacterium*: (Uniprot A0A0G1MLC1)
MGSSHHHHHHSQDPMKYSIRHINQKELETAIEWAATEGWNPGLYDADAFDAMDPQGYFMGYVDNEPVASISA
VSYPGQFGFLGFYIVKPEYRGKGYGWKLWQEAMKYLNSHNVGLDGVVAQQENYKKSGFTRAYANIRYEGIGSNS
VKKSKNIVPLSKVSFDELLKYDREVFPADRSSFLKIWFNQPKSLALGYIKDGKLLGFGMVRPCRVGYKVGPLFADNEK
IAEELFQQFRVFVRKEKIYLDVPEVNKKAVFLAKKYKMVPMFETARMYTKEMPQIDLHKIFGVTTFELG

61SrAT from *Streptomyces rimosus*: (Uniprot L8EWB7)
MGSSHHHHHHSQDPMTSPGEIPRQPVPTEAAAPSGPAPGTATATLTAGPDAPAGFTVRPATLDEWREVAAWAA
AESWNPGRADVDCFHPTDPDGFFVGHLDGQLVSSVSVVNYSAAYAFLGYYLVHPDFRGRGLGLATWRAAVAHA
GDRTIGLDAVPAQEATYRRSGFTTAYRTTRYRGRPVRPGTPAVPVVPVTADRLDEIADFDRQCFPAERRAFLERWL
TAAGHHAYAQIRDGRVAGYGVLRPAQDGYRLGPLFADTSRGAEAVFDALTSHLGPEDEVHIDIPDPHHSARSLALS
RGLEPDSHTMRMYTGPVPPARAEHGYAVTSLELG

62SrAT from *Streptomyces roseochromogenus*: (Uniprot V6JJA0)
MGSSHHHHHHSQDPMTPAAPHDLVVTQATLADWPVIAGWAAAEGWNPGLSDGPAFFAQDADGFFLGRIDGE
PVSAISVVNYGSDYAFLGCYLVRPDLRGHGHGLTTWKTALAHAGDRTVGLDGVVAQQDNYRQSGFELAYRTIRFT
GAAPRAEVPVGVRPAGPADLAAITAYDSACYPADRPRFLAEWLTGPGHRAFVRHDGARLTGYGVLRPGRDTLRIG
PLFADTADDAHALFAALTTEAAGSEVAIDVPEPHAAGVALAEEAGFQASFETARMYTGPVRAYAQHRVFGVTTLEL
G

63ScAT from *Saccharomyces cerevisiae*: (Uniprot Q08649)
MGSSHHHHHHSQDPMSHDGKEEPGIAKKINSVDDIIIKCQCWVQKNDEERLAEILSINTRKAPPKFYVHYVNYNKR
LDEWITTDRINLDKEVLYPKLKATDEDNKKQKKKKATNTSETPQDSLQDGVDGFSRENTDVMDLDNLNVQGIKDE
NISHEDEIKKLRTSGSMTQNPHEVARVRNLNRIIMGKYEIEPWYFSPYPIELTDEDFIYIDDFTLQYFGSKKQYERYRK
KCTLRHPPGNEIYRDDYVSFFEIDGRKQRTWCRNLCLLSKLFLDHKTLYYDVDPFLFYCMTRRDELGHHLVGYFSKEK
ESADGYNVACILTLPQYQRMGYGKLLIEFSYELSKKENKVGSPEKPLSDLGLLSYRAYWSDTLITLLVEHQKEITIDEISS
MTSMTTTDILHTAKTLNILRYYKGQHIIFLNEDILDRYNRLKAKKRRTIDPNRLIWKPPVFTASQLRFAW

64ScAT from *Saccharomyces cerevisiae*: (Uniprot Q03330)
MGSSHHHHHHSQDPMVTKHQIEEDHLDGATTDPEVKRVKLENNVEEIQPEQAETNKQEGTDKENKGKFEKETER
IGGSEVVTDVEKGIVKFEFDGVEYTFKERPSVVEENEGKIEFRVVNNDNTKENMMVLTGLKNIFQKQLPKMPKEYIA
RLVYDRSHLSMAVIRKPLTVVGGITYRPFDKREFAEIVFCAISSTEQVRGYGAHLMNHLKDYVRNTSNIKYFLTYADN
YAIGYFKKQGFTKEITLDKSIWMGYIKDYEGGTLMQCSMLPRIRYLDAGKILLLQEAALRRKIRTISKSHIVRPGLEQFK
DLNNIKPIDPMTIPGLKEAGWTPEMDALAQRPKRGPHDAAIQNILTELQNHAAAWPFLQPVNKEEVPDYYDFIKE
PMDLSTMEIKLESNKYQKMEDFIYDARLVFNNCRMYNGENTSYYKYANRLEKFFNNKVKEIPEYSHLID

65PsAT from *Pseudoalteromonas sp*: (Uniprot F8J3H2)
MGSSHHHHHHSQDPMSEKLDSYKLMQEHQTWTSKPASLEEWQIVNEWAIAEKWDLGLGDTERFFNIDEEGFYL
GYVNDEPVASVSVVNYTDEYAYAGFYLVAPGARGKGYGLRLSYDAFRHCDKRSVGLDGMPEQEENYKKGGFVTH
YETSRLVGIHNQQVDAPDGVQNITADNIDEVIKFDEKITGYPRAALLKDWFSGEGRHGFVINSGDGVIGVVGIRRST
DGYRLGPLYSENQAVCDKLFAMALAQVPQGTQVTIDAPTLDLGFINGLKKMGFEEIFHTFRMYRGKEPQGEKHKI
QAIASLELG

66BcAT from *Bacillus cereus*: (Uniprot Q81AS3)
MGSSHHHHHHSQDPMTDFQKQFFARLHIEEKDTVSFEDLSNIMYAMAQTVPFENLNILEKNFKEISKENLKEKILV
NNRGGLCYELNPTMYYFLKDSGFDVHLVSGTVYNAANSIWAVDSGHIATVLTHHNELYLIEVGFGSYLPLAPVPFLG
EVIHSATGDYRIRKEMTEKGNYILEMRKNNEFLDQSAADDWTLGYAFYIEEVDEEKANTAQKIIVEHEGSPFNKVPLI
VKLTEDGHASLTKDSLTVAKNGKKTKETVTDMQYTNLLHSKFGITL

67CaAT from *Capsicum annuum*: (Uniprot Q9ATJ3)
MGSSHHHHHHSQDPMASAPQPPTLSEKTTNLSPENDNVTITGKIYTRVRLATKSDLHHAYQLFYQIHAYHNQFHL
FKATESSLSDLFFKENPLPLFYGPTLLLLEVSPTAFTEPKNNKDEGFKPVFTALDLKFPVVEGQVEEFRSKYDDGTDKR
DVFIAGYAYFFASYSLFGNDKPGIHFDSLYFRESYRKLGMGKLLFGTVASIAANNGFAAVEGIVAVWNKKSYDFYVS
MGVEMHDDFRFGKLDGENLQKYADKEKNGAGSC

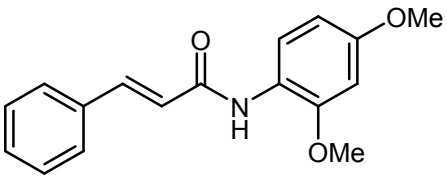
2. General protocols for the chemical synthesis of amide standards

General

All compounds were purchased from Sigma-Aldrich, TCI chemicals, Apollo Scientific and Alfa Aesar and used without further purification. NMR data was collected using a Bruker Avance 400 Ultrashield instrument and analyzed using ACD/SpecManager version 12.5. Coupling constants (*J*) are given in Hz. LC-MS analysis was conducted on a Waters Aquity UPLC with a Waters ZQ mass detector with alternate scan positive and negative electrospray.

Method 1: Amine coupling reactions with acid chlorides

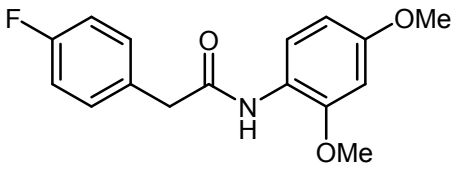
To an ice-cold solution of amine (1.0 mmol, 1.0-1.2 eq) in dichloromethane (8mL), triethylamine (376 µl) and acid chloride (1.0 eq) were added dropwise (cinnamoyl chloride was dissolved in a minimal amount of dichloromethane) and the reaction was stirred at room temperature for 1 h. The reaction was washed with saturated $NaHCO_3$ (2 x 10 ml) followed by 1M HCl (1 x 10 ml). The organic phase was dried with $MgSO_4$, filtered and concentrated under reduced pressure to afford the product which was used directly unless otherwise stated.



N-(2,4-dimethoxyphenyl)cinnamamide

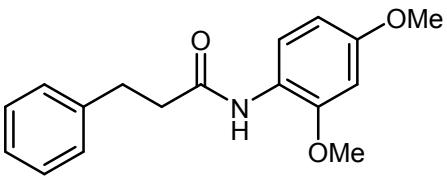
The crude product required further purification by silica chromatography: Heptane/ethyl acetate 90:10 followed by 70:30 and 50:50

^{1}H NMR ($CDCl_3$): δ 7.74 (d, J=15.5, 1H), 7.61 (br s, 1H), 7.50 (m, 3H), 7.36 (m, 3H), 6.98 (br d, J=7.8, 1H), 6.81 (d, J=8.6, 1H), 6.56 (d, J=15.5, 1H), 3.87 (s, 3H), 3.85 (s, 3H), m/z: 284.1 $[M]^+$, 10 mg, 4 % yield as a yellow solid



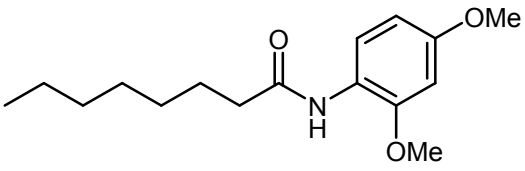
N-(2,4-dimethoxyphenyl)-2-(4-fluorophenyl)acetamide

^{1}H NMR ($CDCl_3$): δ 7.29 (m, 2H), 7.16 (br s, 1H), 7.07 (m, 3H), 6.75 (m, 2H), 3.83 (s, 6H), 3.64 (s, 3H), m/z: 290.2 $[M]^+$, 142 mg, 55 % yield as a lilac solid



N-(2,4-dimethoxyphenyl)-3-phenylpropanamide

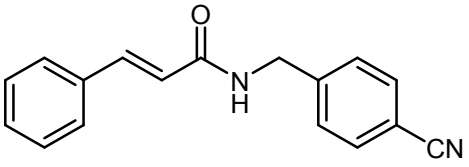
^{1}H NMR ($CDCl_3$): δ 7.26 (m, 6H), 6.98 (br s, 1H), 6.76 (m, 2H) 3.86 (s, 3H), 3.84 (s, 3H), 3.05 (t, J=7.6, 2H), 2.64 (t, J=7.6, 2H), m/z: 286.2 $[M]^+$,180 mg, 70 % yield as a lilac solid



N-(2,4-dimethoxyphenyl)octanamide

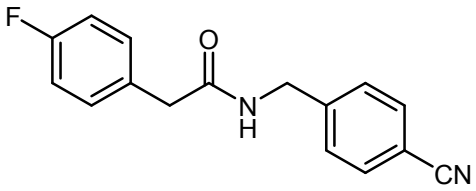
^{1}H NMR ($CDCl_3$): δ 7.39 (d, J=2.2, 1H), 7.14 (br s, 1H), 6.84 (dd, J=8.4, 2.5, 1H), 6.79 (d, J=8.4, 1H), 3.87 (s, 3H), 3.02 (s, 3H), 2.33 (t, J=7.6, 2H), 1.72 (quin, J=7.4, 2H), 1.34 (m, 8H), 0.88 (t, J=6.9, 3H), m/z: 280.2 $[M]^+$,

190 mg, 76 % yield as a lilac solid



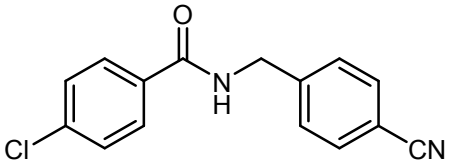
N-(4-cyanobenzyl)cinnamamide

^{1}H NMR ($CDCl_3$): δ 7.69 (d, J=15.5, 1H), 7.62 (m, 2H), 7.50 (m, 2H), 7.43 (m, 2H), 7.37 (m, 3H), 6.45 (d, J=15.5, 1H), 6.14 (br t, J=6.1, 1H), 4.63 (d, J=6.1, 2H), m/z: 263.2 $[M]^+$, 135 mg, 57% % yield as a white solid



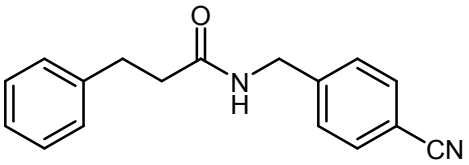
N-(4-cyanobenzyl)-2-(4-fluorophenyl)acetamide

^{1}H NMR ($CDCl_3$): δ 7.59 (m, 2H), 7.26 (m, 4H), 7.05 (m, 2H), 5.82 (br s, 1H), 4.45 (d, J=6.1, 2H), 3.60 (s, 2H), m/z: 267.2 $[M]^+$, 158 mg, 65% % yield as a white solid



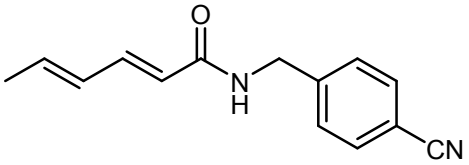
4-chloro-N-(4-cyanobenzyl)benzamide

^{1}H NMR ($CDCl_3$): δ 7.74 (m, 2H), 7.64 (m, 2H), 7.44 (m, 4H), 6.51 (br m, 1H), 4.70 (d, J=6.0, 2H), m/z: 269.3 $[M]^-$, 138 mg, 57 % yield as a white solid



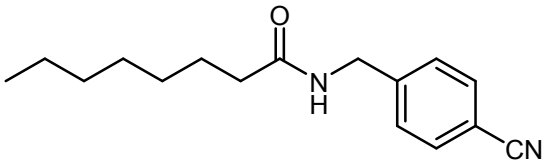
N-(4-cyanobenzyl)-3-phenylpropanamide

^{1}H NMR ($CDCl_3$): δ 7.57 (m, 2H), 7.25 (m, 7H), 5.75 (br m, 1H), 4.47 (d, J=5.9, 2H), 3.04 (t, J=7.4, 2H), 2.60 (t, J=7.4, 2H), m/z: 265.2 $[M]^+$, 187 mg, 79 % yield as a white solid



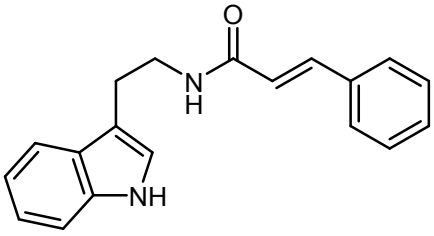
(2E,4E)-N-(4-cyanobenzyl)hexa-2,4-dienamide

^{1}H NMR ($CDCl_3$): δ 7.60 (m, 2H), 7.40 (m, 2H), 7.24 (m, 1H), 6.14 (m, 2H), 5.93 (br m, 1H), 5.78 (d, J=15.0, 1H), 4.57 (dd, J= 11.8, 5.3, 2H), 1.85 (d, J=5.3, 3H), m/z: 227.1 $[M]^+$, 172 mg, 84 % yield as a white solid



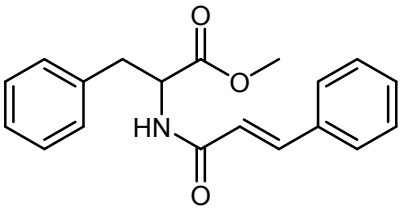
N-(4-cyanobenzyl)octanamide

^{1}H NMR ($CDCl_3$): δ 7.63 (m, 2H), 7.40 (m, 2H), 6.00 (br m, 1H), 4.51 (d, J=6.4, 2H), 2.26 (t, J=7.4, 2H), 1.68 (m, 2H), 1.32 (m, 8H), 0.90 (m, 3H), m/z: 259.2 $[M]^+$, 150 mg, 65 % yield as a white solid



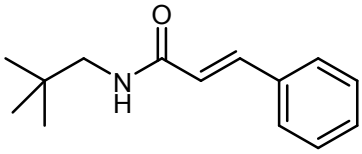
N-(2-(1H-indol-3-yl)ethyl)cinnamamide

^{1}H NMR ($CDCl_3$): δ 8.11 (br s, 1H), 7.64 (d, J=7.9, 1H), 7.60 (d, J=15.6, 1H), 7.46 (m, 2H), 7.39 (m, 1H), 7.34 (m, 3H), 7.22 (ddd, J=8.4, 7.1, 1.2, 1H), 7.14 (ddd, J=8.0, 7.1, 1.0, 1H), 7.07 (d, J=2.2, 1H), 6.28 (d, J=15.6, 1H), 5.69 (br m, 1H), 3.75 (q, J=6.6, 2H), 3.06 (t, J=6.6, 2H), m/z: 291.1 $[M]^+$, 201 mg, 77 % yield as a pale orange solid



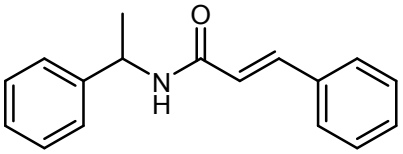
Methyl cinnamoylphenylalaninate

^{1}H NMR ($CDCl_3$): δ 7.63 (d, J=15.6, 1H), 7.50 (m, 2H), 7.36 (m, 3H), 7.28 (m, 2H), 7.12 (m, 2H), 6.39 (d, J=15.6, 1H), 6.09 (br d, J=7.5, 1H), 5.04 (dt, J=7.7, 5.6), 3.75 (s, 3H), 3.24 (dd, J=13.9, 5.4, 1H), 3.18 (dd, J=13.5, 5.3, 1H), m/z: 310.0 $[M]^+$, 183 mg, 66 % yield as a white solid



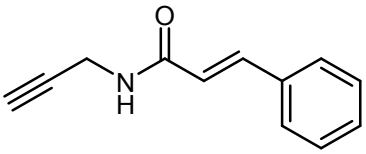
N-neopentylcinnamamide

^{1}H NMR ($CDCl_3$): δ 7.64 (d, J=15.5, 1H), 7.51 (m, 2H), 7.36 (m, 3H), 6.42 (d, J=15.6, 1H), 5.65 (br m, 1H), 3.22 (d, J=6.4, 2H) 0.96 (s, 9H) m/z: 218.2 $[M]^+$, 137 mg, 70 % yield as a white solid



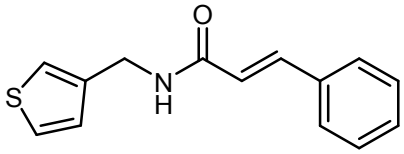
N-(1-phenylethyl)cinnamamide

^{1}H NMR ($CDCl_3$): δ 7.63 (d, J=15.5), 7.48 (m, 2H), 7.36 (m, 7H), 7.27 (m, 1H), 6.39 (d, J=15.6, 1H), 5.85 (br d, J=7.1, 1H), 5.28 (m, 1H), 1.57 (d, J=6.9, 3H), m/z: 252.2 $[M]^+$, 109 mg, 48 % yield as a white solid



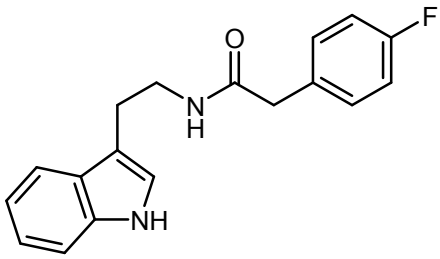
N-(prop-2-yn-1-yl)cinnamamide

^{1}H NMR ($CDCl_3$): 7.66 (d, J=15.6, 1H), 7.50 (m, 2H), 7.36 (m, 3H), 6.40 (d, J=15.6, 1H),5.88 (br s, 1H), 4.20 (dd, J=5.3, 2.6, 2H), 2.27 (t, J=2.5, 1H), m/z: 186.1 $[M]^+$, 126 mg, 76 % yield as a white solid



N-(thiophen-3-ylmethyl)cinnamamide

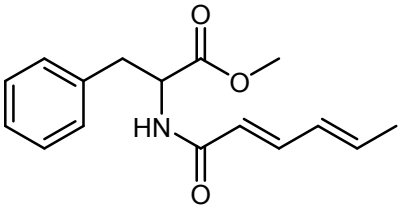
^{1}H NMR ($CDCl_3$): δ 7.66 (d, J=15.6, 1H), 7.48 (m, 2H), 7.35 (m, 3H), 7.30 (dd, J=4.9, 3.0, 1H) 7.19 (m, 1H), 7.06 (dd, J= 4.9, 1.2, 1H), 6.42 (d, J=15.6, 1H), 6.04 (br m, 1H), 4.57 (d, J=5.7, 2H), m/z: 244.1 $[M]^+$, 129 mg, 59 % yield as a white solid



N-(2-(1H-indol-3-yl)ethyl)-2-(4-fluorophenyl)acetamide

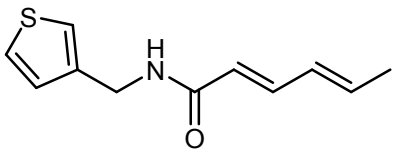
The product required further purification by silica chromatography: Heptane/ethyl acetate 90:10 followed by heptane/ethylacetate 70:30

^{1}H NMR ($CDCl_3$): δ 7.98 (br s, 1H), 7.53 (d, J=7.9, 1H), 7.37 (d, J=8.1, 1H), 7.21 (t, J=7.4, 1H), 7.08 (m, 3H), 6.93 (m, 2H), 6.83 (d, J=1.9, 1H), 5.38 (br s, 1H), 3.55 (q, J=6.4, 2H), 3.45 (s, 2H), 2.91 (t, J=6.6, 2H), m/z: 297.1 $[M]^+$, 33 mg, 12 % yield as an off-white solid



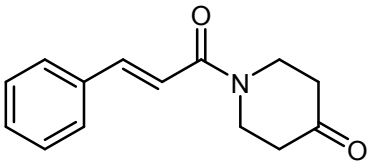
Methyl ((2E,4E)-hexa-2,4-dienoyl)phenylalaninate

^{1}H NMR ($CDCl_3$): δ 7.27 (m, 4H), 7.09 (m, 2H), 6.13 (m, 2H), 5.89 (br d, J= 7.3, 1H), 5.73 (d, J=15.1, 1H), 4.98 (m, 1H), 3.73 (s, 3H), 3.17 (t, J=6.8, 2H), 1.84 (d, J=5.5, 3H), m/z: 274.1 $[M]^+$, 173 mg, 70 % yield as a viscous yellow oil



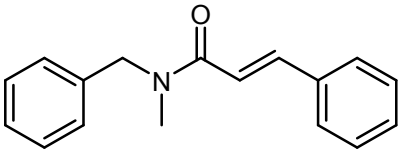
(2E,4E)-N-(thiophen-3-ylmethyl)hexa-2,4-dienamide

^{1}H NMR ($CDCl_3$): δ 7.29 (dd, J=4.9, 3.1, 1H), 7.21 (m, 1H), 7.16 (m, 1H), 7.04 (dd, J=4.9, 1.2, 1H), 6.12 (m, 2H), 5.73 (d, J=15.1, 1H), 5.69 (br s, 1H), 4.53 (d, J=5.7, 2H), 1.84 (d, J=5.7, 3H), m/z: 208.0 $[M]^+$, 102 mg, 55 % yield as a brown solid



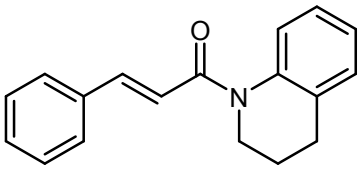
1-cinnamoylpiperidin-4-one

^{1}H NMR ($CDCl_3$): δ 7.75 (d, J=15.4, 1H), 7.54 (m, 2H), 7.40 (m, 3H), 6.94 (d, J=15.4, 1H), 3.98 (br m, 4H), 2.55 (t, J=6.3, 4H), m/z: 231.0 $[M]^+$, 98 mg, 48 % yield as a white solid



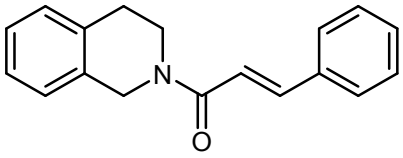
N-benzyl-N-methylcinnamamide

^{1}H NMR ($CDCl_3$): δ 7.76 (dd, J=15.4, 5.6, 1H), 7.55 (br d, J=6.2, 1H), 7.46 (m, 1H), 7.30 (m, 8H), 6.90 (m, 1H), 4.71 (br d, J=12.7, 2H), 3.08 (d, J=9.0, 3H), m/z: 252.0 $[M]^+$, 120 mg, 53 % yield as a colourless oil



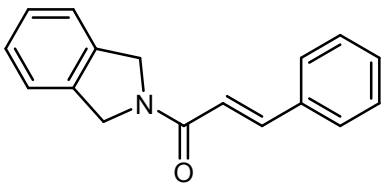
(E)-1-(3,4-dihydroquinolin-1(2H)-yl)-3-phenylprop-2-en-1-one

^{1}H NMR ($CDCl_3$): δ 7.74 (d, J=15.6, 1H), 7.44 (m, 2H), 7.34 (m, 3H), 7.17 (m, 4H), 6.84 (d, J=15.6, 1H), 3.92 (t, J=6.7, 2H), 2.76 (t, J=6.5, 2H), 2.01 (quin, J=6.6, 2H), m/z: 264.1 $[M]^+$, 132 mg, 56 % yield as a yellow solid



(E)-1-(3,4-dihydroisoquinolin-2(1H)-yl)-3-phenylprop-2-en-1-one

^{1}H NMR ($CDCl_3$): δ 7.71 (d, J=15.4, 1H), 7.55 (dd, J=7.7, 1.5, 2H), 7.37 (m, 3H), 7.17 (m, 4H), 6.98 (d, J=15.4, 1H), 4.84 (s, 2H), 3.91 (m, 2H), 2.95 (m, 2H), m/z: 264.0 $[M]^+$, 132 mg, 56 % yield as a white solid

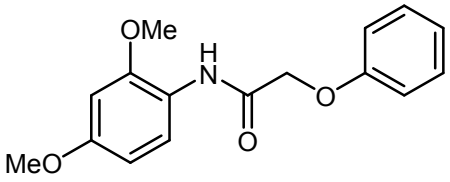


(E)-1-(isoindolin-2-yl)-3-phenylprop-2-en-1-one

^{1}H NMR ($CDCl_3$): δ 7.80 (d, J=15.4, 1H), 7.59 (m, 2H), 7.40 (m, 3H), 7.32 (m, 4H), 6.85 (d, J=15.4, 1H), 5.04 (s, 2H), 4.94 (s, 2H), m/z: 250.0 $[M]^+$, 118 mg, 53 % yield as a brown solid

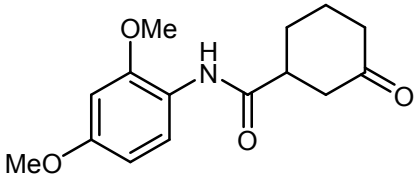
Method 2: Amine and Acid coupling reactions using EDC coupling reagent

1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide HCl (240 mg, 1.25 eq), DMAP (24.5 mg, 0.2 eq) and triethylamine (175 ul) were added to a suspension of acid (1 eq) in DCM (8 ml), followed by addition of the amine (1 eq). The reactions were stirred at room temperature overnight. Reactions were washed with saturated $NaHCO_3$ (2 x 10 ml), followed by 1M HCl (1 x 10 ml) and brine (1 x 10 mL). The organic phase was dried ($MgSO_4$), filtered and concentrated under reduced pressure.



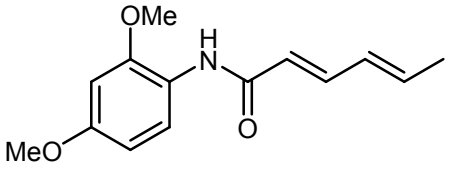
N-(2,4-dimethoxyphenyl)-2-phenoxyacetamide

^{1}H NMR ($CDCl_3$): δ 8.19 (br s, 1H), 7.36 (m, 3H), 7.07 (t, J=7.4, 1H), 6.99 (m, 3H), 6.83 (d, J=8.6, 1H), 4.61 (s, 2H), 3.90 (s, 3H), 3.87 (s, 3H), m/z: 288.1 $[M]^+$, 95 mg, 33 % yield as a brown solid



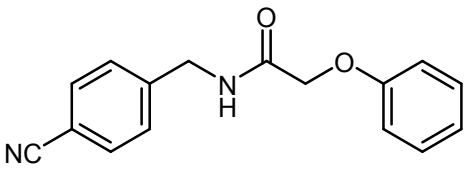
N-(2,4-dimethoxyphenyl)-3-oxocyclohexane-1-carboxamide

^{1}H NMR ($CDCl_3$): δ 7.38 (d, J=2.2, 1H), 7.17 (br s, 1H), 6.85 (dd, J=9.1, 3.6, 1H), 6.80 (d, J=9.1, 1H), 3.87 (s, 3H), 3.86 (s, 3H), 2.71 (m, 2H), 2.55 (m, 1H), 2.40 (m, 2H), 2.14 (m, 2H), 2.00 (m, 1H), 1.75 (m, 1H), m/z: 278.2 $[M]^+$, 147 mg, 53 % yield as a brown solid



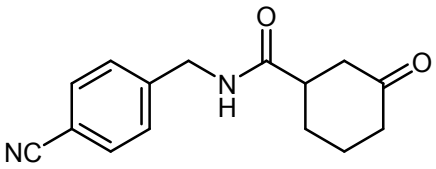
(2E,4E)-N-(2,4-dimethoxyphenyl)hexa-2,4-dienamide

^{1}H NMR ($CDCl_3$): δ 7.49 (s, 1H), 7.31 (dd, J=14.8, 10.0, 1H), 7.17 (s, 1H), 6.88 (dd, J=8.3, 1.3, 1H), 6.80 (d, J=8.5, 1H), 6.16 (m, 2H), 5.87 (d, J=14.8, 1H), 3.89 (s, 3H), 3.86 (s, 3H), 1.86 (d, J=5.8, 3H), m/z: 248.0 $[M]^+$, 98 mg, 40 % yield as a brown solid



N-(4-cyanobenzyl)-2-phenoxyacetamide

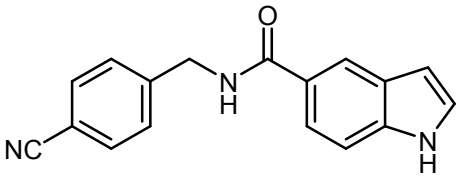
^{1}H NMR ($CDCl_3$): δ 7.61 (m, 2H), 7.34 (m, 4H), 7.05 (m, 1H), 6.92 (m, 2H), 4.60 (d, J=6.3, 2H), 4.58 (s, 2H), m/z: 267.2 $[M]^+$, 95 mg, 36 % yield as a white solid



N-(4-cyanobenzyl)-3-oxocyclohexane-1-carboxamide

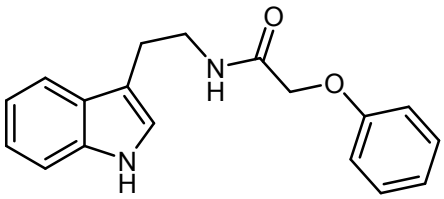
The product required further purification by silica chromatography: Heptane/ethyl acetate 90:10 followed by heptane/ethylacetate 70:30

^{1}H NMR ($CDCl_3$): δ 7.62 (d, J=8.2, 2H), 7.36 (d, J=8.4, 2H), 6.00 (br, m, 1H), 4.50 (t, J=6.2, 2H), 2.66 (m, 2H), 2.48 (m, 1H), 2.37 (m, 2H), 2.13 (m, 1H), 1.97 (m, 2H), 1.71 (m, 1H), m/z: 257.2 $[M]^+$ 80 mg, 31 % yield as a white solid



N-(4-cyanobenzyl)-1H-indole-5-carboxamide

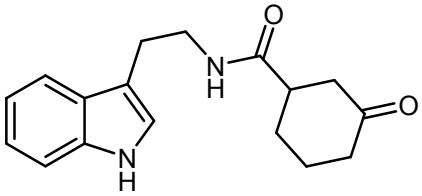
The product precipitated during the course of the reaction, the solid was filtered under vacuum and washed with dichloromethane to yield the product which was used without further purification. ^{1}H NMR (DMSO-d6): δ 11.41 (br s, 1H), 9.04 (t, J=6.0, 1H), 8.19 (m, 1H), 7.80 (m, 2H), 7.68 (dd, J=8.5, 1.7, 1H), 7.52 (d, J=8.5, 2H), 7.43 (m, 2H), 6.53 (m, 1H), 4.56 (d, J=5.6, 2H), m/z: 276.1 $[M]^+$, 215 mg, 78 % yield as a light orange solid



N-(2-(indolin-3-yl)ethyl)-2-phenoxyacetamide

The product required purification by silica chromatography: Heptane/ethyl acetate 90:10 followed by heptane/ethyl acetate 70:30,

^{1}H NMR ($CDCl_3$): δ 7.97 (br s, 1H), 7.61 (d, J=7.9, 1H), 7.38 (d, J=8.1, 1H), 7.29 (m, 2H), 7.22 (m, 1H), 7.12 (m, 1H), 7.01 (t=7.3, 1H), 6.90 (d, J=2.2, 1H), 6.81 (d, J=8.0, 2H), 6.68 (br m, 1H), 4.47 (s, 2H), 3.68 (q, J=6.6, 2H), 3.01 (t, J= 6.8, 2H), m/z: 295.1 $[M]^+$, 14 mg, 5 % yield as an off white solid



N-(2-(indolin-3-yl)ethyl)-3-oxocyclohexane-1-carboxamide

^{1}H NMR ($CDCl_3$): δ 8.08 (br s, 1H), 7.60 (d, J=7.5, 1H), 7.38 (d, J=8.2, 1H), 7.22 (m, 1H), 7.13 (m, 1H), 7.03 (d, J=2.2, 1H), 5.50 (br m, 1H), 3.61 (dquin, J=12.8, 6.6, 2H), 2.98 (t, J=6.6, 2H), 2.61 (m, 1H), 2.36 (m, 4H), 2.07 (m, 1H), 1.87 (m, 2H), 1.65 (m, 1H), m/z: 285.1 $[M]^+$, 80 mg, 28 % yield as a yellow oil

3. General protocol for the expression and purification of CoA ligases

LB media (5 ml) containing kanamycin (50 μg/ml) was inoculated with a single colony of *E. coli* BL21(DE3) transformed with pET28-CL, and incubated overnight at 37 °C, 200 rpm. For expression, LB media (500 ml) containing 1 % glycerol, 0.05 % antifoam and kanamycin 50 μg/ml was inoculated with the overnight seed culture (5 ml). The culture was grown to an OD_{600} 0.6-1.0 (2.5 hours) and protein expression was induced with the addition of IPTG to a final concentration of 0.3 mM. Cultures were incubated for a further 18-24 hours at 18°C. Cells were harvested by centrifugation (20 mins, 4000 rpm, 4°C) and either purified directly or stored at -20°C until required. The general protocol was followed for all proteins with the following exception; genes PhCL and stlB were expressed using Rosetta2(DE3) cells and chloramphenicol (34 μg/ml) in addition to kanamycin (50 μg/ml).

For purification, cell pellets were resuspended in 35 ml lysis buffer (50 mM Tris, 150 mM NaCl, 10 % glycerol, pH 7.4-8.0) and lysed by sonication (13 mm probe, 5 minutes, 9 sec on, 9.0 sec off, 40 % amplitude). The lysates were cleared by centrifugation (39,191 *g*, 4°C, 45 minutes) and bound to pre-equilibrated Ni-NTA resin (2-3 ml). The matrix was washed with wash buffer (50 mM Tris, 150 mM NaCl, 10 % glycerol, 20-25 mM imidazole, pH 7.4-8.0) and the protein was eluted in elution buffer (50 mM Tris, 150 mM NaCl, 10 % glycerol, pH 7.4-8.0) containing increasing concentrations of imidazole (50-400 mM). Purified protein was identified by SDS-PAGE analysis and pure fractions were desalted using a PD10 column (GE Healthcare). Protein concentration was determined using a Bradford assay according to the manufacturers protocol (Bio-Rad).

4. Pyrophosphate detection assay for the evaluation of CoA ligase substrate specificity

The activities of nine CoA ligases were evaluated towards acid substrates **1a-o** using an EnzChek pyrophosphate assay (Invitrogen) according to the manufacturer's protocol. Briefly, assay solution was prepared as follows: 50 μL reaction buffer (x20), 200 μL MESG (2-amino-6-mercapto-7-methylpurine ribonucleoside, 1 mM solution), 10 μL PNP (purine nucleoside phosphorylase, 100 U/ml), 10 μL inorganic pyrophosphate (30 U/ml) and 430 μL water. CoA ligase (5 μL, 0.5 mg/ml) and

assay solution (25 µL) were mixed in a 384 well plate, to initiate the reaction substrate solution (10 µL) containing 1.6 mM acid, 1.6 mM ATP and 1.6 mM CoA in 50 mM Tris buffer pH 7.5 was added and the change in absorbance at 360 nm was measured over time. Reactions were performed in triplicate and corrected against a blank containing no CoA ligase. The specific activity was calculated using a standard curve generated from pyrophosphate standards.

5. Analytical scale CoA ligase reactions

CoA ligase activity was confirmed by UPLC for selected reactions using purified enzyme. Reactions were prepared as follows: 750 µl acid substrate (20 mM, in tris buffer 50 mM, pH 8.0), 300 µl CoA (50 mM), 300 µl ATP (50 mM), 300 µl $MgCl_2$ (50 mM), 600 µL Tris buffer pH 8.0 (50 mM), 3.25 ml water. The final reaction mix was pH adjusted to 8.0, and the reaction was initiated with the addition of CoA ligase (500 µl, 0.5 mg/ml). Reactions were incubated at 21°C, 150 rpm for 24 hours. A 500 µl sample was removed from each reaction and 50 µl HCl (5M) was added to quench the reaction. The reaction was vortexed and the precipitated protein was cleared by centrifugation (13000 rpm, 1 minute).

UPLC analysis was conducted on an Agilent 1290 infinity II, Zorbax SB-C18 column (50 mm x 3.0 mm, i.d. 1.8 µm packing diameter). The solvents employed were A = water + 0.1 % formic acid and B = Acetonitrile + 0.1 % formic acid and a gradient method was used according to table S1. Flow rate: 1.5 ml/min, Temperature: 60 °C, Injection volume: 2 µL, UV detection: 220 nm.

Table S1: UPLC gradient method

Time (min)	% A	% B
0.00	100	0
2.50	5	95
2.70	5	95
2.71	100	0
4.00	100	0

Figure S1: UPLC trace of the reaction between cinnamic acid (**1a**) and CoA catalyzed by the enzyme PhCL.

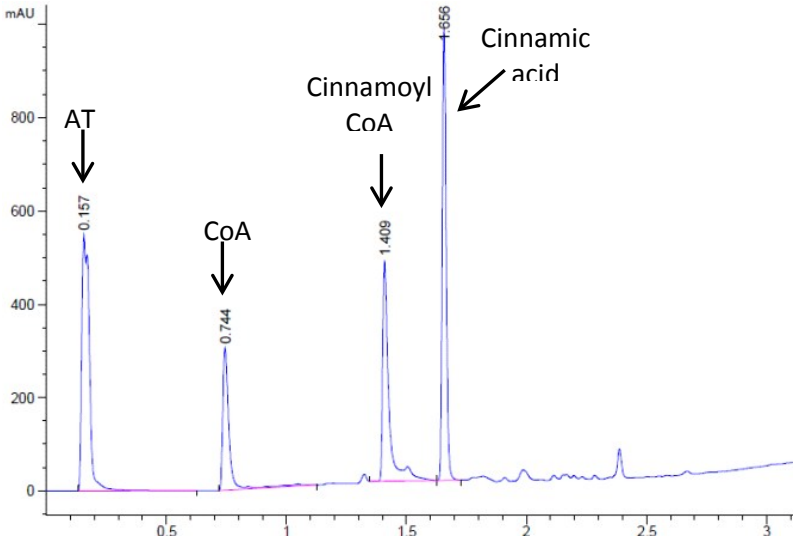


Table S2: Conversion of selected acid substrates **1a, c-g, l, k & m** to the corresponding CoA esters

Acid	CL	Conversion[1]
1a	PhCL	44.6 %
1c	phl	8.0 %
1d	phl	82.3 %
1e	ipfF	86.8 %
1f	ipfF	69.2 %
1g	CBL	73.0 %
1i	CBL	52.8%
1k	PhCL	32.3%
1m	PhCL	17.6 %

[1] Conversion was calculated from the product peak area relative to the remaining acid peak area

6. General protocol for the expression of the N-acyltransferase panel

All transfer and aliquotting steps were performed using a Beckman Biomek FX liquid handling robot.

Glycerol stocks of BL21(DE3) transformed with pCDF-Duet-NAT constructs were used to inoculate 170 µl TB media containing 50 µg/ml streptomycin in each well of a 96-well microtitre plate, the overnight cultures were sealed with a gas permeable seal and incubated overnight at 37 °C, 200 rpm, 85% humidity. 10 µl of the overnight culture was used to inoculate 390 µl TB media containing 50 µg/ml streptomycin + 0.4 % glycerol in each well of a 2 ml 96-deep well plate. The plates were sealed and incubated at 30 °C, 250 rpm, 85% humidity until they reached an OD_{600} of 0.6-0.8 (2 hours). For induction, 40 µl IPTG (10 mM solution) was added to each well and the cultures were grown for an additional 20 hours at 20 °C, 250 rpm, 85% humidity. The cells were harvested by centrifugation (4000 rpm, 10 minutes, 4 °C) and the supernatant was poured off, cell pellets were stored at -80 °C until required.

Protein expression was confirmed by SDS-PAGE analysis, under the expression conditions described the following NATs did not yield soluble protein: 03BaAT, 04RlAT, 06MtAT, 08MaAT, 09MmAT, 13AtAT, 15SpST, 16ScAT, 17EcAT, 20HsAT 21BsAT, 22HsAT, 25DmAT, 26AtAT, 27AtAT, 40BtAT, 44TcAT, 46LmAT, 50EcAT, 57SpAT, 59XbAT, 60MbAT

7. General protocol for N-acyltransferase assays

All transfer and aliquotting steps were performed using a Beckman Biomek FX liquid handling robot.

Cell Lysis

Plates were thawed on the bench for 1 hour, 200 µL lysis buffer (50 mM phosphate buffer pH 7.5 containing 1 mg/ml lysozyme, 0.5 mg/ml polymixin B, 0.1 U/ml Benzonase) was added to each well and the cell pellets were resuspended with shaking (850 rpm) for 2 hours at room temperature. Cell debris was cleared by centrifugation (4000 rpm, 4 °C, 10 minutes). 100 µl of the cleared cell lysate was transferred to a fresh 96-well microtitre plate, the lysates were used directly in the assays.

Reactions

Reactions contained 50% lysate loading and a final concentration of 2.5 mM amine and 2.5 mM Acid or CoA ester, and were incubated at 25 °C, 800 rpm, for 18 hours. To quench, 100 µl acetonitrile was

added to each well and the plate was incubated on a plate shaker for 1 hour at 850 rpm. Precipitated protein was removed by filtration using Acroprep Filter plates (96-well, 350 µl, 0.2 µm, GHP) centrifuged at 4000 rpm for 5 minutes. The samples were collected in a 96-well microtitre plate and injected onto the UPLC-MS for analysis.

UPLC-MS method

UPLC analysis was conducted on an Acquity UPLC, CSH C18 column (50mm x 2.1mm
i.d. 1.7µm packing diameter) at 40 °C. The solvents employed were: A = 10 mM Ammonium Bicarbonate in water adjusted to pH 10 with Ammonia solution, B = Acetonitrile, and a gradient method was used according to table S3. The UV detection was a summed signal from wavelength of 210nm to 400nm. Injection volume: 2 µl. Flow rate: 1 ml/min.

Table S3: UPLC gradient method

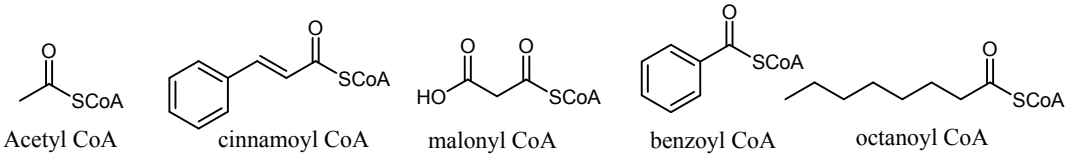
Time (min)	% A	% B
0.00	97	3
0.05	97	3
1.50	5	95
1.90	5	95
2.00	97	3

MS detection was performed using a Waters ZQ, alternate-scan positive and negative electrospray, Scan Range: 100 to 1000 AMU, Scan Time: 0.27 seconds, Inter scan Delay: 0.10 seconds

7.1 Preliminary assay: NAT Activity towards a panel of amines (2a-2d)

Cleared cell lysates were screened using either 1) commercial CoA esters or 2) Cinnamoyl CoA produced *in situ* from cinnamic acid and CoA ligase, PhCL. Commercial CoA esters (figure S2) were purchased from Sigma: Acetyl coenzyme A sodium salt CAS# 102029-73-2, Malonyl coenzyme A lithium salt CAS# 108347-84-8, Benzoyl coenzyme A lithium salt CAS# 102185-37-5, octanoyl Coenzyme A lithium salt CAS# 324518-20-9. Acid and amine solutions (20 mM) were prepared in 50 mM Tris buffer and pH adjusted to 8.0

Figure S2: Naturally occurring CoA esters used for the initial substrate profiling of NATs



Acetyl CoA cinnamoyl CoA malonyl CoA benzoyl CoA octanoyl CoA

Assay mix 1: Commercially available CoA ester substrates (volume/well)
5 µl CoA ester (50 mM in water)
32.5 µL buffer (50 mM Tris, pH 8.0)

Assay mix 2: Cinnamoyl CoA produced in situ from **1a** *(Volume/well)*
5 µl CoA (50 mM in water, adjusted to pH 5 with NaOH)
6 µL ATP (50 mM in water)
12.5 µL Cinnamic acid (20 mM solution in 50 mM Tris pH 8.0)
10 µL purified PhCL (0.5 mg/ml)
4 µL buffer (50 mM Tris, pH 8.0)

Amine solution (12.5 µl) and CoA ester assay mix (37.5 µl) were added to each well of a 96-well microtitre plate. Lastly, using the Biomek FX liquid handling robot, N-acyltransferase lysate (50 µl) was added and reactions were incubated at 25 °C, 800 rpm, for 18 hours. The final reaction contained either 2.5 mM CoA ester, 2.5 mM Amine in Tris pH 8.0 with 50% acyltransferase lysate loading or 2.5 mM Acid, 2.5 mM CoA, 5 µg CoA ligase, 3 mM ATP, 2.5 mM Amine in Tris pH 8.0 with 50 % acyltransferase lysate loading. Conversion was calculated as a percentage relative to the amine (Conversion = product peak area/(product peak area + amine peak area)*100)

Table S4: Activity of N-acyltransferases towards amines **2a-2d** using naturally occurring CoA ester partners

	CoAester	Percentage Conversion (%)			
		Amine **2a**	Amine **2b**	Amine **2c**	Amine **2d**
01StAT	acetyl	51.1	nd[1]	0	nd[1]
02MsAT	acetyl	18.9	nd[1]	0	nd[1]
05PaAT	acetyl	92.5	nd[1]	19.4	nd[1]
07NfAT	acetyl	89.8	nd[1]	67.9	nd[1]
10HvAT	cinnamoyl	0	4.5	0	0
11AtAT	cinnamoyl	0	31.8	0	0
12HvAT	cinnamoyl	0	4	0	0
14AtAT	cinnamoyl	0	0	0	0
18BsAT	acetyl	0	nd[1]	4.4	nd[1]
19CpAT	acetyl	3.3	nd[1]	0	nd[1]
23DmAT	acetyl	21.2	nd[1]	100	nd[1]
24LlAT	cinnamoyl	0	0	0	0
28BaAT	malonyl[2]	0	0	0	0
29EfAT	malonyl[2]	0	0	0	0
30SaAT	malonyl[2]	0	0	0	0
31NtAT	cinnamoyl	0	26.9	11.9	18.2
32NtAT	cinnamoyl	0	30.2	8.5	18.5
33MtAT	malonyl[2]	0	0	0	0
34PaAT	malonyl[2]	0	0	0	0
35MaAT	acetyl	3.3	nd[1]	0	nd[1]
36PaAT	acetyl	3.5	nd[1]	0	nd[1]
37HsAT	octanoyl	0	0	0	0
38ShAT	acetyl	3.5	nd[1]	0	nd[1]
39SeAT	acetyl	0	nd[1]	0	nd[1]
41GmAT	acetyl	90	nd[1]	24.8	nd[1]
42GmAT	acetyl	57	nd[1]	23.8	nd[1]
43GmAT	acetyl	24.7	nd[1]	0	nd[1]
45SpAT	acetyl	3.1	nd[1]	0	nd[1]
47AtAT	acetyl	0	nd[1]	0	nd[1]
48CaAT	cinnamoyl	8	31.6	24.7	49.1
49StAT	cinnamoyl	8.8	33.7	21.6	49.5

51EcAT	acetyl	0	nd[1]	0	nd[1]
52PlAT	octanoyl	29.1	23.5	0	59.1
53PhAT	octanoyl	49	21.3	0	48.9
54SlAT	cinnamoyl	0	24.8	5.2	81.6
55SlAT	cinnamoyl	6.4	30.5	25.7	48.4
56SlAT	cinnamoyl	0	36.2	12.5	43.6
58AxAT	octanoyl	0	0	0	0
61SrAT	octanoyl	0	61.1	0	6.3
62SrAT	octanoyl	0	0	0	0
63ScAT	acetyl	0	nd[1]	0	nd[1]
64ScAT	acetyl	4.4	nd[1]	0	nd[1]
65PsAT	octanoyl	14	13.6	0	52.1
66BcAT	acetyl	86.1	nd[1]	82.1	nd[1]
67CaAT	cinnamoyl	0	31.7	24.9	33

[1] N-acetyl phenylalanine methyl ester and N-acetyl 4-(aminomethyl)benzonitrile product peaks and the corresponding amine substrate peak were not resolved using this analytical method and the conversion was not determined.
[2] No product peaks were observed using malonyl CoA as an acyl donor, however in the absence of a product standard the product retention time was not known.

7.2 Assay 2: NAT Activity towards a panel of CoA esters

The activity of the panel of N-acyltransferases were screened towards CoA esters which were produced *in situ* from acids **1a, 1c-e, 1g, 1k-m &1o** using either substrate **2a** or **2b** as the amine coupling partner. 31 N-acyltransferases which had demonstrated activity in the preliminary assay were evaluated.

CoA ester assay mix (Volume/well)
0.05 µl CoA (50 mM in water, adjusted to pH 5 with NaOH)
6 µL ATP (50 mM in water)
12.5 µL Acid solution (20 mM, in 50 mM Tris pH 8.0)
10 µL CoA Ligase (0.5 mg/ml)
6.45 µL buffer (50 mM Tris, pH 8.0)

CoA ester assay mix (37.5 µl) and amine solution (20 mM, 12.5 µl) were added to each well of a 96-well microtitre plate. Using the Biomek FX liquid handling robot, N-acyltransferase lysate (50 µl) was added and reactions were incubated at 25 °C, 800 rpm, for 18 hours. The final reaction mix contained 2.5 mM Acid, 2.5 mM Amine, 0.025 mM CoA, 3 mM ATP, 5 µg CoA ligase in Tris pH 8.0 with 50 % N-acyltransferase lysate loading. The percentage conversion was calculated relative to standard curves of the chemically synthesised amides and amine starting materials.

NATs 23DmAT, 28BaAT, 36PaAT and 65PsAT (table S5) were active towards only a single CoA ester substrate tested and were omitted from the table in the main paper.

Table S5: Activity of NATs 23DmAT, 28BaAT, 36PaAT and 65PsAT towards CoA esters produced in situ from acids **1a, 1c-e, 1g, 1k-m & 1o**

NAT	Amine	Percentage conversion (%)								
		1a +PhCL	**1c** +ipfF	**1d** +phl	**1e** +ipfF	**1g** +CBL	**1k** +PhCL	**1l** +ipfF	**1m** +PhCL	**1o** +RpCL
23DmAT	**2a**	0	0	0	0	0	0	0	0	2
28BaAT	**2b**	0	0	0	0	0	0	55.8	0	0
36PaAT	**2a**	0	0	7.4	0	0	0	0	0	0
65PsAT	**2b**	0	0	0	0	0	0	40.4	0	0

7.3 Assay 3: NAT Activity towards a panel of amines (2c-2i)

A panel of 21 N-acyltransferases were screened towards amines **2c-2i**. A CoA ester substrate was selected for each N-acyltransferase which was produced *in situ* using purified CoA ligases.

CoA ester assay mix (Volume/well)
0.05 µl CoA (50 mM in water, adjusted to pH ~5 with NaOH)
6 µL ATP (50 mM in water)
12.5 µL Acid solution (20 mM, made up in 50 mM Tris pH 8.0)
10 µL CoA Ligase (0.5 mg/ml) prepared in N36158-4
6.45 µL buffer (50 mM Tris, pH 8.0)

CoA ester assay mix (37.5 µl) and amine solution (20 mM, 12.5 µl) were added to each well of a 96-well microtitre plate. Using the Biomek FX liquid handling robot, N-acyltransferase lysate (50 µl) was added and reactions were incubated at 25 °C, 800 rpm, for 18 hours. The final reaction mix contained 2.5 mM Acid, 2.5 mM Amine, 0.025 mM CoA, 3 mM ATP, 5 µg CoA ligase in Tris pH 8.0 with 50 % N-acyltransferase lysate loading. The percentage conversion was calculated relative to standard curves of the chemically synthesised amides and amine starting materials. Amines (R)-**2d** and (R)-**2e** were also screened but were not substrates for any of the NATs.

23DmAT was active towards only a single amine tested (table S6) and was omitted from the table in the main paper.

Table S6: Activity of 23DmAT towards amines **2c-2i**

NAT	Acid	CL	Converion (%)						
			2c	**(S)-2d**	**(S)-2e**	**2f**	**2g**	**2h**	**2i**
23DmAT	**1o**	RpCL	14.3	0	0	0	0	0	0

The activity of acyltransferase 11AtAT was evaluated towards amines **2k**-**2o** according to the same protocol, using cinnamic acid **1a** and CoA ligase PhCL.

8. Evaluation of L-azetidine carboxylic acid N-acyltransferases

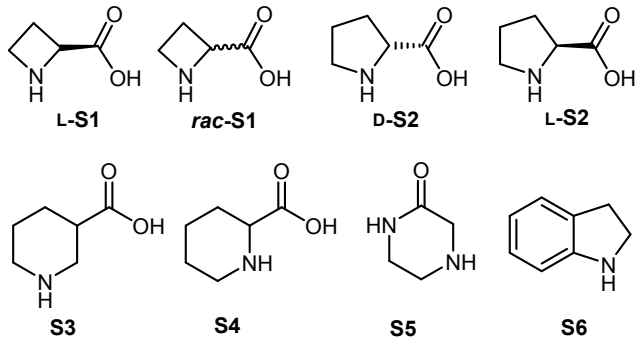
Mpr1 from *Saccharomyces cerevisiae*
MDAESIEWKLTANLRNGPTFFQPLADSIEPLQFKLIGSDTVATAFPVFDTKYIPDSLINYLFKLFNLEIESGKTYPQLHS
LTKQGFLNYWFHSFAVVVLQTDEKFIQDNQDWNSVLLGTFYIKPNYAPRCSHNCNAGFLVNGAHRGQKVGYRLA
QVYLNWAPLLGYKYSIFNLVFVTNQASWKIWDKLNFQRIGLVPHAGILNGFSEPVDAIIYGKDLTKIEPEFLSME

KfAT from *Kribbella flavida*
MSPEALQVRDAEDADWPAILPFFREIVSAGETYAYDPELTDEQARSLWMTPSGAPQSRTTVAVDADGTVLGSAN
MYPNRPGPGAHVASASFMVAAAARGRGVGRALCQDMIDWAGREGFRAIQFNAVVETNTVAVKLWQSLGFRVI
GTVPEAFHHPTHGYVGLHVMHRPL

Two L-azetidine carboxylic acid N-acyltransferases (MpR1 and KfAT) were expressed and purified following the protocol described for CoA ligases (section 3). The activity of the purified enzymes towards a small panel of secondary amines (**S1-6**) and acetyl CoA was evaluated using a spectrophotometric Ellman's assay. Ellman's reagent (5,5-dithio-bis-(2-nitrobenzoic acid) reacts quantitatively with the free sulfhydryl group of CoA to yield 2-nitro-5-thiobenzoate, which can be detected at 412 nm.

Purified N-Acyltransferases (0.25mg/ml) was added to amine substrate (0.5 mM), acetylCoA (0.5mM) and Ellman's reagent (0.5mM) in Tris Buffer (50 mM, pH 8.0) and the increase in absorbance at 412 nm was measured over time on a M2 Spectramax (Molecular Devices). MpR1 and KfAT displayed activity towards the known substrate azetidine-2-carboxylic acid (**L-S1** and ***rac*-S1**), however substrates **S2-6** were not accepted.

Figure S3: Panel of secondary amines **S1-6**



9. Synthesis of the Losmapimod amide intermediate using whole cell biocatalysts

For co-expression of CoA ligase CBL and N-acyltransferase CASHT, the CBL gene was subcloned into the pCDF-Duet-CASHT construct using NdeI and NotI restriction sites. A single colony of E. coli BL21(DE3) transformed with the pCDF-duet-CASHT-CBL construct was picked and used to inoculate an overnight culture of LB media (5 ml) containing streptomycin (50 µg/ml), which was incubated at 37 °C, 200 rpm. For expression, LB media (500 ml) containing 1 % glycerol, 0.05 % antifoam and streptomycin (50 µg/ml) was inoculated with the overnight seed culture (5 ml). The culture was grown to an OD_{600} 0.6-1.0 (2.5 hours) and protein expression was induced with the addition of IPTG to a final concentration of 0.3 mM. Cultures were incubated for a further 18 hours at 28°C. Cells were pelleted by centrifugation (4000 rpm, 20 minutes), the supernatant was removed and the cell pellet was resuspended in 80 mL of fresh LB media which was used directly in the amide formation reaction.

6-chloronicotinic acid (93.6 mg, 0.6 mmol) and neopentylamine (70 µl, 0.6 mmol) were added to the resuspended cells (60 mL) in a 250 mL Erlenmyer flask and the reaction was incubated at 21 °C for 16 hours with shaking 200 rpm. To analyse the final conversion, a 200 µl sample was taken from the

reaction and 200 µl acetonitrile was added. The sample was vortexed and the cell debris was cleared by centrifugation (1 min at 13,000 rpm). The sample was filtered using a Mini-UniPrep Syringeless Filter (0.45 µm, PVDF, Fisher) and analysed by UPLC-MS.

UPLC analysis was conducted on an Agilent 1290 infinity II, Zorbax SB-C18 column (50 mm x 3.0 mm, i.d. 1.8 µm packing diameter). The solvents employed were A = water + 0.1 % formic acid and B = Acetonitrile + 0.1 % formic acid and a gradient method was used according to table S1. Flow rate: 1.0 ml/min, Temperature: 60 °C, Injection volume: 0.5 µL, UV detection: 220 nm.

Table S7: UPLC gradient method

Time (min)	**% A**	**% B**
0.00	97	3
1.50	5	95
1.90	5	95
2.00	97	3

The UV detection was a summed signal from wavelength of 210nm to 350nm. MS detection was performed using a Waters ZQ, Ionisation mode: Alternate-scan positive and negative electrospray, Scan Range: 100 to 1000 AMU, Scan Time: 0.27 seconds, Inter scan Delay: 0.10 seconds. The product eluted at 0.94 minutes and the acid at 0.69 minutes.

The conversion (83%) was calculated relative to the remaining acid. Next, the cells were pelleted by centrifugation (4000 rpm, 10 minutes) and the product was extracted from the supernatant using ethyl acetate (2 x 50 ml). The organic phases were combined, dried over $MgSO_4$ and concentrated under reduced pressure to give the product as a white solid (99 mg, 74 % yield).

^{1}H NMR ($CDCl_3$): δ 8.7 (d, J=2.09 Hz, 1H), 8.1 (dd, J=8.31, 2.52 Hz, 1H), 7.4 (d, J=8.31 Hz, 1H), 6. (br. s, 1H), 3.3 (d, J=6.40 Hz, 2H), 1.0 (s, 9H)