

UNIVERSIDAD DE BARCELONA
Facultad de Farmacia
Departamento de Bioquímica y Biología Molecular

**REDESIGN OF CARNITINE ACETYLTRANSFERASE
SPECIFICITY BY PROTEIN ENGINEERING**

ANTONIO FELIPE GARCIA CORDENTE

2006

APPENDIX

1. PRIMER SEQUENCES

In all the cases the mutated nucleotides are underlined.

PRIMER	SEQUENCE (5' → 3')
CrAT-ATG.for	GAT GTT AGC TTT TGC TGC CAG
CrAT-2100.rev	CTT GTT CAG CCT CTG GGC TCA GC
<i>Hind</i> III-CrAT.for	TCG ATA AGC TTA TAA AAT GTT AGC CTT TGC TGC CAG AAC
<i>Eco</i> RI-CrAT.rev	CGG AAT TCC GCC AAA GTG GGC TTG GCT GTG
CrAT A106M.for	GAG TGG TGG CTC AAG ACA <u>ATG</u> TAT CTC CAG TTC CGC CAG
CrAT A106M.rev	CTG GCG GAA CTG GAG ATA <u>CAT</u> TGT CTT GAG CCA CCA CTC
CrAT S119N. for	GTG GTC ATC TAC TCC <u>AAT</u> CCA GGA GTG CTG CTG
CrAT S119N. rev	CAG CAG CAC TCC TGG <u>ATT</u> GGA GTA GAT GAC CAC
CrAT F138R. for	CAG GGT CAG CTC CGG <u>CGT</u> GCT GCC AAA CTC ATC
CrAT F138R. rev	GAT GAG TTT GGC AGC <u>ACG</u> CCG GAG CTG ACC CTG
CrAT G249A. for	CAA CAA GGA GCC TGT <u>TGC</u> CAT CCT AAC CTC CAA C
CrAT G249A. rev	GTT GGA GGT TAG GAT <u>GGC</u> AAC AGG CTC CTT GTT G
CrAT H343A.for	GTG GGA TGG TTT ATG <u>AAG</u> CTG CAG CTG CAG AAG GG
CrAT H343A.rev	CCC TTC TGC AGC TGC <u>AGC</u> TTC ATA AAC CAT CCC AC
CrAT E347A.for	GAA CAT GCA GCT GCA <u>GCA</u> GGG CCC CCC ATT GTC
CrAT E347A.rev	GAC AAT GGG GGG CCC <u>TGC</u> TGC AGC TGC ATG TTC
CrAT D356A.for	CAT TGT CGC TCT TGT <u>GGC</u> CCA TGT CAT GGA G
CrAT D356A.rev	CTC CAT GAC ATG <u>GGC</u> CAC AAG AGC GAC AAT G
CrAT A515G. for	GCC ATC CGA GGG GAG <u>GGC</u> TTT GAC CGG CAC TTG
CrAT A515G. rev	CAA GTG CCG GTC AAA <u>GCC</u> CTC CCC TCG GAT GGC
CrAT R518N.for	GGG GAG GCC TTT GAC <u>AAC</u> CAC TTG CTG GGT CTG
CrAT R518N.rev	CAG ACC CAG CAA GTG <u>GTT</u> GTC AAA GGC CTC CCC
CrAT M564A.for	CAA GAC AGA CTG TGT <u>CGC</u> GTC CTT CGG ACC TGT G
CrAT M564A.rev	CAC AGG TCC GAA GGA <u>CGC</u> GAC ACA GTC TGT CTT G
CrAT M564G.for	CAA GAC AGA CTG TGT <u>CGG</u> GTC CTT CGG ACC TGT G
CrAT M564G.rev	CAC AGG TCC GAA GGA <u>CCC</u> GAC ACA GTC TGT CTT G
CrAT T465V/T467N.for	GTT TCA CCT GGG CCG <u>CGT</u> TGA CAA CAT CCG CTC CGC CTC C
CrAT T465V/T467N.rev	GGA GGC GGA GCG GAT <u>GTT</u> GTC <u>AAC</u> GCG GCC CAG GTG AAA C
CrAT GGG.for	CAA GAC AGA CTG TGG <u>CGG</u> <u>GGG</u> CTT CGG ACC TGT G
CrAT GGG.rev	CAC AGG TCC GAA <u>GCC</u> <u>CCC</u> <u>GCC</u> ACA GTC TGT CTT G

Appendix

PRIMER	SEQUENCE (5' → 3')
CAT2-OE-For	GACTGAATTCATGAGGATCTGTCATTCGAGAACTCTCTC
CAT2-OE-Rev	GACTCTCGAGTCATAACTTTGCTTTTCGTTTATTCTCATTTTC
CAT2-ΔN.For	GACTGAATTCATGCATTCGGCCATTGTCAATTACTC
CAT2-ΔC.Rev	GACTCTCGAGTCATTTTCGTTTATTCTCATTTTCCAAG
CAT2-Y606G.For	CGGAATATTTTGACGGTGGTGGTTGGTCCCAAG
CAT2-Y606G.Rev	CTTGGGACCAACCACCACCGTCAAAATATTCCG
COT D343A.for	GAA CAT TGC TCA CTA TGT TGC TGA GAA GCT CCT AGA G
COT D343A.rev	CTC TAG GAG CTT CTC AGC AAC ATA GTG AGC AAT GTT C
COT I551A.for	CTG GTT GGT TAC TTA CGA GCT CAG GGA GTC GTG GTT C
COT I551A.rev	GAA CCA CGA CTC CCT GAG CTC GTA AGT AAC CAA CCA G
COT G553M.for	GTT ACT TAC GAA TTC AGA TGG TCG TGG TTC CCA TG
COT G553M.rev	CAT GGG AAC CAC GAC CAT CTG AAT TCG TAA GTA AC

2. SEQUENCE OF THE RAT CARNITINE OCTANOYL TRANSFERASE

Sequence of the rat COT (Choi, 1995), with accession number in GenBank Data Libraries U26033.

DEFINITION: *Rattus norvegicus* carnitine octanoyltransferase mRNA, complete cds.

ACCESSION NUMBER: U26033

SOURCE: *Rattus norvegicus* (Norway rat)

CODIFYING SEQUENCE: 52...1890 (612 amino acid residues).

```
/translation="MENQLAKSIEERTFQYQDSLPLPVPSLEESLKKYLESVKPFAN
EDEYKKTEEIVQKFQDGVGKTLHQKLLERAKGKRNLWLEEWLNVAYLDVRIPSQLNVN
FVGPSPHFEHYWPAREGTQLERGSILLWHNLNYWQLLRREKLPVHKSGNTPLDMNQFR
MLFSTCKVPGITRDSIMNYFKTESEGHCPHTIAVLCRGRAFVFDVLHDGCLITPPELL
RQLTYIYQKCWNEPVGPSIAALTSEERTRWAKAREYLIIGLDPENLTLLEKIQSSLFVY
SIEDTSPHATPENFSQVFEMLLGGDPAVRWGDKSYNLI SFANGIFGCS CDHAPYDAML
MVNIAHYVDEKLLETEGRWKGSEKVRDIPLPEELAF TVDEKILNDVYQAKAQHLKAAS
DLQIAASTFTSFGKKLTKKEALHPDTFIQLALQLAYYRLHGRPGCCYETAMTRYFYHG
RTETVRSCTVEAVRWCQSMQDPSASLLERQQKMLDAFAKHNKMMRDCSHGKGFDRHLL
GLLLIAKEEGLPVPELFDPLFSRSGGGNFVLSTSLVGYLRIQGVVPMVHNGYGFF
YHIRDDRFVVTCSWRSCL E TDAEKL VEMI FHAFHDMI HLMNTAHL "
```

```

1  gagtgagag agccaagccg ggtgcaggag ttttcttact gtgactatac catggaaaat
61  caattggcta agtcaattga agaacgaaca ttccagtacc aggactctct tccgccttg
121 cccgttcctt cgcttgaaga atcactgaag aagtaccttg agtcagtga gccatttga
181 aatgaagacg aatacaagaa aactgaagaa atagttcaaa agtttcaaga tggagttggc
241 aagacattgc atcagaagt acttgaaagg gctaaaggaa aaagaaactg gctggaagag
301 tgggtggctca atgtcgctca cttggatgtg cgtattccat cacaactgaa cgtgaacttt
361 gtgggtccgt ctccccactt tgaacactac tggcctgcaa gggaaggcac tcagttgaa
421 agaggaagca tactactgtg gcacaacttg aactactggc agctgctaag aagagaaaaa
481 ttgcctgtac ataaatctgg aaatactcct ctagacatga accaattccg gatgctgttt
541 tctacctgca aggttccggg aatcactaga gattcgatta tgaattatth taagactgag
601 agcgaggggc attgtccgac ccacattgcc gtgctgtgtc gaggcagagc gtttgcctc
661 gatgtcctcc atgacggttg tttgatcacc ccaccagaac ttctcagaca actgacatac
721 atctaccaga aatgctggaa tgaacctgtt gggcccagta tagcggcatt aaccagtga
781 gagcgaactc ggtgggcgaa ggcaagagaa tatctgattg gtcttgatcc agagaactg
841 actttattag aaaaaattca atccagttta tttgtgtatt ccatagaaga caccagtcca
901 catgcaacc cagaaaatth ttctcaggtc tttgaaatgc ttcttggtgg agatccagca
961 gtgcgctggg gtgacaagtc ctataatctg atttcctttg ctaacggaat atttggctgt
1021 agctgtgatc atgctcctta tgatgcaatg cttatggtga acattgctca ctatgttgat
1081 gagaagctcc tagagacgga agggagatgg aagggttcag aaaaagtcag ggatataccg
1141 ttgccagagg agctggcttt cactgtggat gagaagatac tgaatgacgt ctaccaagcc
1201 aaagcccaac acctcaaagc agcatctgat ttacagatag cagcatctac cttcacatct
1261 tttggcaaaa agctcactaa gaaggaggcc cttcaccctg acacctttat tcagctcgt
1321 cttcagctcg cctactacag acttcatgga cgccccggtt gctgctatga aacagctatg
1381 acaagatact tttaccatgg ccgaacagag actgtgcat cttgtacagt ggaggccgtc

```

Appendix

1441 aggtggtgcc agtccatgca ggatccttct gccagtctcc ttgaacgtca gcaaaagatg
1501 ttagacgctt ttgcaaagca taacaagatg atgagagatt gttcccatgg aaaaggattt
1561 gaccgtcacc ttttaggcct tttgctcata gcaaaagagg aaggcctccc tgttccagaa
1621 ctgtttgagg atccactttt ctccagaagt ggaggagggtg ggaattttgt gctgtcaaca
1681 agtctgggtg gttactttacg aattcagggg gtcgtgggtc ccatggtaca taatggatac
1741 ggctttttct accacatcag agatgacagg tttgtgggtga catgttcac ctaggaggtca
1801 tgtcttgaga ctgatgcaga aaagttagtg gagatgattt ttcattgctt ccacgatatg
1861 atacatctga tgaacacggc tcatctttag agactcagag acatacaggt cacagaaact
1921 gggtagcggag aatgggatgg tgatacgaca tggaagggaat gttgacttaa aggaaacctg
1981 ttaatgcagg gattagagag ggatgcactc tagatttatt ctaccttaa gccttctgtt
2041 gcaacagcaa tgcaaaactca gacatagtg atagaactat gcaatgtttt aagcctcaac
2101 aatgcacatc tgtatatttt aacaatacaa atcctactct aatgttaaaa tatttttgtt
2161 ggcacatgtg taggttgcaa gtctctctgtg aacataatta tagagtattt ctcaagcact
2221 ttaatacttt ctaatggcca gagggataaa aacccatggg tagatgctaa tttccctgac
2281 atcagtgcct tctacatcca gcacaggagt acaagcctat gagatttcat gggaaaacca
2341 ctattgttca atattgatct aaaatagctc ctttgaacag acaaaagtat caagttgtat
2401 tagaaaagaa tattagcaaa actcattatg atatgttgta attaattttg tgaatataaa
2461 atcaaaacac ttccatttaa atctacttgg tagagttagt ggctttaaag ggttaaatgt
2521 cgagtatgat tctcagaact ttataattat ttcccactgt tattcaaaat gttagcatat
2581 agacattctc ccattgtaat tcagtgttta tattctcaaa gaataaagca tccagaatcc
2641 ttgtaatttc tcatttattt tcaataaaaa tgattcctga t

3. ALIGNMENT OF REPRESENTATIVE SEQUENCES OF CARNITINE/CHOLINE ACYLTRANSFERASES

Amino acid sequence of 22 representative enzymes that catalyse short acyl-CoAs as substrates: CrAT (CACP) from yeast, *C. tropicalis*, human, mouse and pigeon; and ChAT (CLAT) from mouse, rat, pig, human, fruit fly, and *C. elegans*; and enzymes which have medium- and long-chain acyl-CoAs as substrates: L-CPT I (CPT1) from rat, mouse and human; M-CPT I (CPTM) from rat and mouse; CPT II (CPT2) from rat, mouse and human; and COT (OCTC) from human, rat and cattle; were obtained from the SwissProt data bank and aligned using ClustalW. Residues have been coloured by conservation. The position of the putative amino acids involved in acyl-CoA specificity is indicated with an asterisk.

CPT2_RAT	1MMPRLLLFRAMP,RCPSVLVGPASRP	24
CPT2_MOUSE	1MMPRLLLLRWLP,RCPSVLVGPASRP	24
CPT2_HUMAN	1MVPRLLLLRAMP, RGPVAVGPGASRP	24
CACP_YEAST	1MRICHSRITLNLKDLPISTRAMHSAIVNYS, TQKA	25
CACP_CANTR	1MNFNFKLSQQVLKN, ...STKS	17
CACP_HUMAN	1AFAARTVVKP	10
CACP_MOUSE	1MLFAARTVVKP	12
CACP_COLLII	1MDRKQKQAEKARP	13
CLAT_MOUSE	1*	1
CLAT_RAT	1*	1
CLAT_PTIG	1*	1
CLAT_HUMAN	1	MGLRTAKKRGLGGGGKWKREEGGGTRGRREVPRACFLQSGGRGDPGDVGGPAGNPGCSPHPRAATRPPPLPAHTPAHTPEWCGAASAEAAEPRRAGPHLCIPAPGLTKTP	110
CLAT_DROME	1VAGNEASTSAGSGPESAAALFSLRFSIGSGPNSPQRVVSNLRGLFTHR	50
CLAT_CAEEEL	1MEKEK	5
CPTM_RAT	1MAEAHQAVAFQFTVTPDGVDFRLSREALRHIYLSGINSJJKRIRIKNGILRG	53
CPTM_HUMAN	1MAEAHQAVAFQFTVTPDGVDFRLSREALKHVYLSGINSJJKRIRIKNGILRG	53
CPT1_RAT	1MAEAHQAVAFQFTVTPDGVDFRLSREALKQICLSGLHSJKKFFIRFKNGIITG	53
CPT1_MOUSE	1*FQFTVTPDGVDFRLSREALKQICLSGLHSJKKFFIRFKNGIITG	44
CPT1_HUMAN	1MAEAHQAVAFQFTVTPDGVDFRLSREALRQIYLSGLHSJKKFFIRFKNGIITG	53
CPT2_RAT	25	LSAVS.....GPDYDYLQH.....SIVPTMHYQD.....	47
CPT2_MOUSE	25	LSAVS.....GPAEYLQH.....SIVPTMHYQD.....	47
CPT2_HUMAN	25	LSAGS.....GPGQYLQR.....SIVPTMHYQD.....	47
CACP_YEAST	36	QFPVETNNGEHY, WAE.....KPNKIFYQNK, RPNFQGITAKQQ.....	72
CACP_CANTR	18	IMPILK.....KP, FSTS.....HAKGDLFKYQS.....	40
CACP_HUMAN	11	LGFLKRF.....SLMKASS.....RFKAHQD.....	31
CACP_MOUSE	13	LGLLKPS.....SLMKVSG.....RFKAHQD.....	33
CACP_COLLII	14	YGLLKRA.....ALGKIPG.....RFQHQE.....	34
CLAT_MOUSE	2	ILEKVP.....PKMPVQA.....SS, CEVVL.....	21
CLAT_RAT	2	ILEKAP.....QKMPVKA.....SS, WEEL.....	20
CLAT_PTIG	2	ILEKTP.....PKMAAKS.....PSS, EEPP.....	21
CLAT_HUMAN	111	ILEKVP.....RKMAAKT.....PSS, EES.....	129
CLAT_DROME	51	LSNITTSDTGKWDLSL.....IPKKWJST.....AESVDFGFP.....	85
CLAT_CAEEEL	6	VDE.....LPP.....NDNMYET.....	18
OCTC_HUMAN	1MENQLAKS.....TEERTFYQD.....	18
OCTC_RAT	1MENQLAKS.....IEERTFYQD.....	18
OCTC_BOVIN	1MENQLAKS.....TEERTFYQD.....	18
CPTM_RAT	54	VYPGSRITSWLVVVMATVGSNYCKVDISMGLVHCIQRCLPTRYGSYGTPQTETLLSMVIFSTGWWATGIFLFRQTLKLLSYHGWMFEMHS, KTSBATKIWAICVRLLSR	162
CPTM_HUMAN	54	VYPGSRITSWLVVIMATVGSFNCVDISLGLVSCIQRCLPQCGPYQTPQTRALLSMAIFSTGWWATGIFFRQTLKLLCYHGWMFEMHG, KTSNLTRIWAMCIRLLSSR	162
CPT1_RAT	54	VFPANSSWLIVVGVISSMHAKVDPSPSLGMIAKISRTLDTTG, RMSSTKNIVSGVLFGTGLWVAIIMTRYSKLVLLSYHGWMFAEHG, KMSRSTKIWAMVKKVLSGR	160
CPT1_MOUSE	45	VFPANSSWLIVVGVISSMHTKVDPSPSLGMIAKINRDLDTTG, RMSSTKNIVSGVLFGTGLWVAIIMTRYSKLVLLSYHGWMFAEHG, KMSRSTKIWAMVKKVLSGR	151
CPT1_HUMAN	54	VFPANSSWLIVVGVIMTTHYAKIDPSLGIIAKINRDLTETAN, CMSSTKNIVSGVLFGTGLWVAIIMTRYSKLVLLSYHGWMFTEHG, KMSRATKIWAMVKKVLSGR	160

Appendix

Table with columns for species (e.g., CPT2_RAT, CPT2_MOUSE), protein domains (e.g., SLRPLPIPKLEDTMKRYLNAQKPLLD), and positions (48-161). Includes alignment markers like * and *.

Table with columns for species (e.g., CPT2_RAT, CPT2_MOUSE), protein domains (e.g., KSEYNDQ, LTRATNLTVSAVRFKLTQAGL), and positions (142-237). Includes alignment markers like * and *.

Table with columns for species (e.g., CPT2_RAT, CPT2_MOUSE), protein domains (e.g., ARHLVLRKGFHYVFDVLDQDGN, IWNPLEIQAHKYLIS), and positions (238-329). Includes alignment markers like * and *.

Table with columns for species (e.g., CPT2_RAT, CPT2_MOUSE), protein domains (e.g., FFMKDLI, HLSHTMLHG), and positions (330-410). Includes alignment markers like * and *.

Appendix

CPT2_RAT 411 SVETLSNLSGALKAGTAAKEKFDTTVKTSIDSIQFORGGKFLKKKQLSPDAVAQLAFQMAFLROYG...QTVATYESCSTAFAKHGRTEIC.PASIFTKRCSEAFV 517
CPT2_MOUSE 411 SWQKLSFKLSSALKAGTAAKEKFDATMKTIDAIQFORGGKFLKKKQLSPDAVAQLAFQMAFLROYG...QTVATYESCSTAFAKHGRTEIC.PASIFTKRCSEAFV 517
CPT2_HUMAN 411 TVQKLFNFDLTKAGTAAKEKFDATMKTIDAIQFORGGKFLKKKQLSPDAVAQLAFQMAFLROYG...QTVATYESCSTAFAKHGRTEIC.PASIFTKRCSEAFV 517

CPT2_RAT 518 R.....DPSKHSVG.....ELQHMMAECSKYHGQLTKEANAGGDFDRHLYALRYLATARGNLN...PE...LYLDPAYQRMN.....HNILSTSTLN 593
CPT2_MOUSE 518 R.....EPSKHSVG.....ELQHMMAECSKYHGQLTKEANAGGDFDRHLYALRYLAARGVLT...PE...LYQDPAYQRIN.....HNILSTSTLS 593
CPT2_HUMAN 518 R.....EPSRHSAG.....ELQHMMAECSKYHGQLTKEANAGGDFDRHLYALRYLAARGVLT...PE...LYLDPAYQRMN.....HNILSTSTLS 593

CPT2_RAT 594 SPAVS...LGGAPVVPDGFGIAYAVH,DDWIGCNVSSYSGR...NAREFLHCQKCLEDFDALE.....GKAIKT..... 658
CPT2_MOUSE 594 SPAVS...LGGAPVVPDGFGIAYAVH,DDWIGCNVSSYSGR...NAREFLHCQKCLEDFDALE.....GKAIKT..... 658
CPT2_HUMAN 594 SPAVN...LGGAPVVPDGFGIAYAVH,DDWIGCNVSSYSGR...NAREFLHCQKCLEDFDALE.....GKAIKT..... 658