Characterization of the substrate specificity of squalene-hopene cyclases (SHCs)

Untersuchungen zur Substratspezifität von Squalen-Hopen Zyklasen (SHCs)

- supporting information -

Von der Fakultät 3: Chemie der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von Miriam Seitz

aus Rastatt

Hauptberichter: Prof. Dr. Bernhard Hauer

Mitberichter: Prof. Dr. Bernd Plietker

Vorsitzende: Prof. Dr. Cosima Stubenrauch

Tag der mündlichen Prüfung: 6. Februar 2013

Institut für Technische Biochemie der Universität Stuttgart

2012

Table of content

I Sequences

1 Overview

Name	Original host	GI number	NCBI accession number	sequences	available at ITB?	ITB number
*Aaci*SHC	*Alicyclobacillus acidoterrestris*	927384	CAA61950.1	see 2.1	yes	ITB286
*Aac*SHC	*Alicyclobacillus acidocaldarius*	2851526	P33247.4	see 2.2	yes	ITB285
*Ace*SHC	*Acidothermus cellolyticus*	117928904	YP_873455.1	see 2.3	yes	ITB287
*Apa*SHC1	*Acetobacter pasteurianus*	258541105	YP_003187836.1	see 2.4	yes	ITB312
*Apa*SHC2	*Acetobacter pasteurianus*	258541296	YP_003186729.1	see 2.5	no	-
*Cac*SHC	*Catenulispora acidiphila*	256395787	YP_003117351.1	see 2.6	yes	ITB288
*Gth*SHC	*Geobacillus thermodentrificans*	138895534	YP_001125987.1	see 2.7	yes	ITB162
*Mca*SHC	*Methylococcus capsulatus*	53804820	YP_113312.1	see 2.8	yes	ITB164
*Pca*SHC1	*Pelobacter carbinolicus*	77544139	ABA87701.1	see 2.9	no	-
*Pca*SHC2	*Pelobacter carbinolicus*	77544053	ABA87615.1	see 2.10	yes	ITB313
*Rpa*SHC1	*Rhodopseudomonas palustris*	115526460	YP_783371.1	see 2.11	yes	ITB314
*Rpa*SHC2	*Rhodopseudomonas palustris*	90421528	YP_531598.1	see 2.12	no	-
*Sco*SHC	*Streptomyces coelicolor SHC*	21225057	NP_630836.1	see 2.13	yes	ITB315
*Sfu*SHC1	*Syntrophobacter fumaroxidans*	116698484	ABK17672.1	see 2.14	yes	ITB316
*Sfu*SHC2	*Syntrophobacter fumaroxidans*	116699226	ABK18414.1	see 2.15	no	-
*Ssc*SHC	*Streptomyces scabiei*	260645368	CBG68454.1	see 2.16	yes	ITB289
*Ssv*SHC	*Streptomyces sviceus*	197784692	YP_002207454.1	see 2.17	yes	ITB290

*Sth*SHC	*Spherobacter thermophilus*	269838031	YP_003320259.1	see 2.18	yes	ITB291
*Svi*SHC	*Saccharomonospora viridis*	257056311	YP_003134143.1	see 2.19	yes	ITB292
*Tel*SHC	*Thermesynechococcus elongatus*	22299852	P_683099.1	see 2.20	yes	ITB171
*Ttu*SHC	*Teredinibacter turnerae*	254787171	YP_003074600.1	see 2.21	yes	ITB320
*Zmo*SHC1	*Zymomonas mobilis*	56552444	YP_163283.1	see 2.22	yes	ITB104
*Zmo*SHC2	*Zymomonas mobilis*	6466213	AAF12829.1	see 2.23	yes	ITB283
*Zmo*SHC1_F486Y	*Zymomonas mobilis*_F486Y	-	-	see 10.2.2 and CD-ROM	yes	ITB294
*Aac*SHC_Y420C	*Zymomonas mobilis*_Y420C	-	-	see 10.2.2 and CD-ROM	yes	ITB304
*Zmo*SHC1_Loop	*Zymomonas mobilis*_Loop	-	-	see 10.2.2 and CD-ROM	yes	ITB322

2 Wild type SHCs

2.1 *Aaci*SHC

2.1.1 Amino acid sequence

MTKQLLDTPMVQATLEAGVAHLLRRQAPDGYWWAPLLSNVCMEAEYVLLCHCLGKKNPEREA
QIRKYIISQRREDGTWSIYPGGPSDLNATVEAYVALKYLGEPASDPQMVQAKEFIQNEGGIE
STRVFTRLWLAMVGQYPWDKLPVIPPEIMHLPKSVPLNIYDFASWARATIVTLSYRHESPTC
DATSGLCKGSGIVRGEGPPKRRSAKGGDSGFFVALDKFLKAYNKWPIQPGRKSGEQKALEWI
LAHQEADGCWGGIQPPWFYALLALKCLNMTDHPAFVKGFEGLEAYGVHTSDGGWMFQASISP
IWDTGLTVLALRSAGLPPDHPALIKAGEWLVSKQILKDGDWKVRRRKAKPGGWAFEFHCENY
PDVDDTAMVVLALNGIQLPDEGKRRDALTRGFRWLREMQSSNGGWGAYDVDNTRQLTKSDSI
FATSGEVIDPPSEDVTAHVLECFGSFGYDEAWKVIRKAVEYLKAQQRPDGSWFGRWGVNYVY
GIGAVVPGLKAVGVDMREPWVQKSLDWLVEHQNEDGGWGEDCRSYDDPRLAGQGVSTPSQTA
WALMALIAGGRVESDAVLRGVTYLHDTQRADGGWDEEVYTGTGFPGDFYLAYTMYRDILPVW
ALGRYQEAMQRIRG

2.1.2 DNA sequence (*E. coli* codon optimized)

ATGACGAAACAACTGCTGGACACCCCGATGGTACAAGCGACCCTGGAAGCTGGCGTGGCTCA
TCTGCTGCGTCGCCAAGCACCGGATGGCTACTGGTGGGCACCGCTGCTGTCGAATGTCTGCA
TGGAGGCAGAATACGTGCTGCTGTGCCACTGCTTGGGCAAGAAAAACCCGGAACGTGAGGCG
CAAATCCGTAAATACATCATTTCCCAGCGTCGCGAGGATGGTACTTGGAGCATTTATCCGGG
TGGTCCTTCCGACCTGAATGCCACCGTGGAGGCATACGTGGCGCTGAAATATCTGGGCGAGC
CGGCATCTGATCCGCAAATGGTTCAGGCAAAGGAGTTTATTCAGAACGAAGGTGGTATCGAA
AGCACCCGCGTTTTCACCCGTCTGTGGCTGGCTATGGTTGGCCAGTACCCGTGGGACAAGCT
GCCGGTTATCCCGCCAGAGATTATGCACCTGCCGAAAAGCGTCCCGTTGAACATCTATGACT
TCGCGAGCTGGGCGCGTGCCACCATTGTGACGCTGTCTTATCGTCACGAGTCCCCGACCTGC
GACGCCACCAGCGGCTTGTGTAAAGGTAGCGGTATCGTTCGTGGTGAGGGTCCGCCGAAACG
TCGTAGCGCAAAGGGTGGCGACAGCGGCTTTTTCGTTGCTCTGGACAAGTTCTTGAAGGCAT
ACAACAAATGGCCAATCCAGCCGGGTCGCAAGAGCGGCGAGCAAAAGGCCCTGGAATGGATC
TTGGCGCACCAGGAAGCTGACGGCTGTTGGGGTGGTATCCAGCCACCATGGTTTTACGCGCT
GCTGGCGTTGAAATGCTTGAACATGACCGATCATCCGGCGTTTGTGAAAGGCTTCGAAGGTC
TGGAGGCTTATGGTGTCCACACGAGCGACGGCGGTTGGATGTTCCAAGCGAGCATCTCCCCG
ATTTGGGACACCGGCCTGACGGTCCTGGCGCTGCGTAGCGCGGGTTTGCCGCCTGACCATCC
GGCCCTGATTAAAGCAGGCGAGTGGCTGGTATCTAAACAGATCTTGAAGGATGGTGACTGGA
AAGTGCGTCGCCGTAAAGCCAAGCCGGGTGGCTGGGCGTTTGAGTTTCACTGTGAGAATTAC
CCGGACGTTGATGACACCGCCATGGTCGTTCTGGCTCTGAATGGTATTCAGCTGCCGGATGA
GGGCAAGCGCCGTGATGCGCTGACCCGTGGTTTCCGTTGGCTGCGCGAAATGCAAAGCAGCA
ACGGTGGTTGGGGCGCATACGACGTGGATAATACGCGTCAACTGACTAAGAGCGATAGCATT
TTTGCGACCTCTGGTGAAGTGATTGACCCGCCGTCCGAAGATGTGACCGCCCATGTTCTGGA
GTGTTTTGGTAGCTTCGGCTATGACGAGGCGTGGAAGGTGATTCGCAAGGCGGTCGAGTATC
TGAAGGCGCAGCAGCGTCCGGATGGTAGCTGGTTCGGCCGCTGGGGCGTCAACTATGTTTAC
GGCATCGGTGCCGTCGTTCCAGGTCTGAAAGCAGTGGGTGTTGATATGCGTGAGCCGTGGGT
TCAGAAGAGCCTGGACTGGCTGGTTGAGCATCAGAATGAGGACGGCGGCTGGGGTGAAGATT
GCCGTTCTTACGACGATCCGCGCCTGGCAGGCCAGGGTGTGAGCACTCCGAGCCAGACCGCG
TGGGCACTGATGGCGCTGATTGCGGGTGGTCGTGTCGAAAGCGACGCCGTGCTGCGCGGTGT

GACGTATTTGCACGATACGCAACGTGCGGATGGCGGTTGGGACGAAGAGGTCTATACGGGCA
CGGGTTTTCCGGGTGATTTCTATCTGGCGTACACCATGTACCGCGATATTCTGCCGGTCTGG
GCGCTGGGTCGTTACCAAGAAGCGATGCAACGTATCCGCGGT

2.2 *Aac*SHC

2.2.1 Amino acid sequence

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRME
KIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIE
SSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPV
FPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWL
LERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISP
VWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYY
PDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPF
CDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYG
TGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLA
LGRYKQAIERR

2.2.2 DNA sequence (*E. coli* codon optimized)

ATGGCGGAACAGCTGGTGGAAGCGCCGGCGTATGCGCGCACCCTGGATCGCGCGGTGGAATA
TCTGCTGAGCTGCCAGAAAGATGAAGGCTATTGGTGGGGCCCGCTGCTGAGCAACGTGACCA
TGGAAGCGGAATATGTGCTGCTGTGCCATATTCTGGATCGCGTGGATCGCGATCGCATGGAA
AAAATTCGCCGCTATCTGCTGCATGAACAGCGCGAAGATGGCACCTGGGCGCTGTATCCGGG
CGGCCCGCCGGATCTGGATACCACCATTGAAGCGTATGTGGCGCTGAAATATATTGGCATGA
GCCGCGATGAAGAACCGATGCAGAAAGCGCTGCGCTTTATTCAGAGCCAGGGCGGCATTGAA
AGCAGCCGCGTGTTTACCCGCATGTGGCTGGCGCTGGTGGGCGAATATCCGTGGGAAAAAGT
GCCGATGGTGCCGCCGGAAATTATGTTTCTGGGCAAACGCATGCCGCTGAACATTTATGAAT
TTGGCAGCTGGGCGCGCGCGACCGTGGTGGCGCTGAGCATTGTGATGAGCCGCCAGCCGGTG
TTTCCGCTGCCGGAACGCGCGCGCGTGCCGGAACTGTATGAAACCGATGTGCCGCCGCGCCG
CCGCGGCGCGAAAGGCGGCGGcGGCTGGATTTTTGATGCGCTGGATCGCGCGCTGCATGGCT
ATCAGAAACTGAGCGTGCATCCGTTTCGCCGCGCGGCGGAAATTCGCGCGCTGGATTGGCTG
CTGGAACGCCAGGCGGGCGATGGCAGCTGGGGCGGCATTCAGCCGCCGTGGTTTTATGCGCT
GATTGCGCTGAAAATTCTGGATATGACCCAGCATCCGGCGTTTATTAAAGGCTGGGAAGGCC
TGGAACTGTATGGCGTGGAACTGGATTATGGCGGCTGGATGTTTCAGGCGAGCATTAGCCCG
GTGTGGGATACCGGCCTgGCGGTGCTGGCGCTGCGCGCGGCGGGCCTGCCGGCGGATCATGA
TCGCCTGGTGAAAGCGGGCGAATGGCTGCTGGATCGCCAGATTACCGTGCCGGGCGATTGGG
CGGTGAAACGCCCGAACCTGAAACCGGGCGGCTTTGCGTTTCAGTTTGATAACGTGTATTAT
CCGGATGTGGATGATACCGCGGTggtGGTGtGGGCGCTGAACACCCTGCGCCTGCCGGATGA
ACGCCGCCGCCGCGATGCGATGACCAAAGGCTTTCGCTGGATTGTGGGCATGCAGAGCAGCA
ACGGCGGCTGGGGCGCGTATGATGTGGATAACACCAGCGATCTGCCGAACCATATTCCGTTT
TGCGATTTTGGCGAAGTGACCGATCCGCCGAGCGAAGATGTGACCGCGCATGTGCTGGAATG
CTTTGGCAGCTTTGGCTATGATGATGCGTGGAAAGTGATTCGCCGCGCGGTGGAATATCTGA
AACGCGAACAGAAACCGGATGGCAGCTGGTTTGGCCGCTGGGGCGTGAACTATCTGTATGGC
ACCGGCGCGGTGGTGAGCGCGCTGAAAGCGGTGGGCATTGATACCCGCGAACCGTATATTCA
GAAAGCGCTGGATTGGGTGGAACAGCATCAGAACCCGGATGGCGGCTGGGGCGAAGATTGCC
GCAGCTATGAAGATCCGGCGTATGCGGGCAAAGGCGCGAGCACCCCGAGCCAGACCGCGTGG
GCGCTGATGGCGCTGATTGCGGGCGGCCGCGCGGAAAGCGAAGCGGCGCGCCGCGGCGTGCA

GTATCTGGTGGAAACCCAGCGCCCGGATGGCGGCTGGGATGAACCGTATTATACCGGCACCG
GGTTCCCGGGCGATTTTTATCTGGGCTATACCATGTATCGCCATGTGTTTCCGACCCTGGCG
CTGGGCCGCTATAAACAGGCGATTGAACGCCGCTAA

2.3 *Ace*SHC

2.3.1 Amino acid sequence

MTQASVREDAKAALDRAVDYLLSLQDEKGFWKGELETNVTIEAEDLLLREFLGIRTPDITAE
TARWIRAKQRSDGTWATFYDGPPDLSTSVEAYVALKLAGDDPAAPHMEKAAAYIRGAGGVER
TRVFTRLWLALFGLWPWDDLPTLPPEMIFLPSWFPLNIYDWGCWARQTVVPLTIVSALRPVR
PIPLSIDEIRTGAPPPPRDPAWTIRGFFQRLDDLLRGYRRVADHGPARLFRRLAMRRAAEWI
IARQEADGSWGGIQPPWVYSLIALHLLGYPLDHPVLRRGLDGLNGFTIREETADGAVRRLEA
CQSPVWDTALAVTALRDAGLPADHPRVQAAARWLVGEEVRVAGDWAVRRPGLPPGGWAFEFA
NDNYPDTDDTAEVVLALRRVRLEDADQQALEAAVRRATTWVIGMQSTDGGWGAFDADNTREL
VLRLPFCDFGAVIDPPSADVTAHIVEMLAALGMRDHPATVAGVRWLLAHQEPDGSWFGRWGA
NHIYGTGAVVPALIAAGVSPDTPPIRRAIRWLEEHQNPDGGWGEDLRSYTDPALWVGRGVST
ASQTAWALLALLAAGEEASPAVDRGVRWLVTTQQPDGGWDEPHYTGTGFPGDFYINYHLYRL
VFPISALGRYVNR

2.3.2 DNA sequence (*E. coli* codon optimized)

ATGACCCAAGCAAGCGTACGCGAGGATGCAAAAGCGGCCCTGGACCGTGCGGTTGATTACCT
GCTGAGCCTGCAAGATGAGAAAGGTTTCTGGAAGGGTGAGTTGGAAACCAATGTGACGATTG
AAGCGGAGGACCTGCTGCTGCGTGAGTTCTTGGGTATCCGTACCCCGGATATCACCGCCGAA
ACCGCACGCTGGATTCGTGCGAAACAGCGTTCTGATGGCACGTGGGCGACCTTCTACGACGG
TCCGCCAGACCTGAGCACCTCGGTGGAAGCGTACGTTGCGTTGAAGTTGGCGGGTGATGACC
CGGCAGCTCCGCACATGGAAAAGGCGGCTGCCTATATCCGCGGTGCGGGTGGCGTCGAGCGT
ACCCGCGTCTTTACCCGCCTGTGGCTGGCTCTGTTCGGCCTGTGGCCTTGGGATGACCTGCC
GACTCTGCCTCCGGAGATGATTTTCCTGCCGAGCTGGTTTCCACTGAATATCTACGACTGGG
GCTGTTGGGCGCGTCAGACGGTGGTCCCGCTGACCATCGTCAGCGCACTGCGTCCGGTGCGT
CCGATTCCGTTGAGCATCGACGAGATCCGTACGGGCGCTCCGCCACCACCGCGTGATCCGGC
ATGGACCATCCGTGGCTTTTTCCAGCGTCTGGACGATCTGCTGCGCGGCTATCGTCGTGTGG
CCGATCATGGTCCGGCACGTCTGTTTCGCCGTCTGGCTATGCGTCGTGCAGCAGAATGGATT
ATTGCGCGCCAAGAGGCGGACGGCTCCTGGGGTGGTATTCAGCCGCCGTGGGTGTACAGCCT
GATCGCGCTGCACCTGCTGGGCTACCCGTTGGATCACCCGGTGCTGCGCCGTGGTCTGGACG
GTCTGAATGGTTTTACCATTCGCGAGGAAACGGCCGACGGTGCGGTCCGTCGTCTGGAGGCA
TGCCAAAGCCCGGTTTGGGATACGGCGCTGGCCGTGACGGCCCTGCGTGATGCGGGCTTGCC
TGCCGACCACCCGCGTGTTCAGGCCGCAGCCCGCTGGCTGGTCGGCGAGGAAGTCCGCGTGG
CTGGCGATTGGGCGGTGCGTCGTCCGGGTTTGCCGCCTGGTGGCTGGGCGTTTGAATTTGCT
AACGATAACTATCCGGACACCGATGATACCGCGGAGGTCGTTCTGGCATTGCGCCGTGTCCG
CTTGGAGGACGCAGACCAACAGGCACTGGAGGCAGCTGTGCGTCGCGCCACGACCTGGGTTA
TCGGTATGCAAAGCACTGATGGTGGTTGGGGCGCATTTGACGCGGACAACACTCGCGAACTG
GTTCTGCGTCTGCCGTTCTGCGATTTCGGTGCCGTCATTGACCCGCCGAGCGCCGACGTTAC
GGCGCACATTGTTGAAATGCTGGCAGCGCTGGGCATGCGTGACCATCCGGCTACGGTTGCGG
GTGTTCGCTGGCTGTTGGCGCATCAGGAGCCGGACGGCAGCTGGTTCGGTCGCTGGGGTGCT
AATCATATCTACGGTACCGGCGCAGTTGTCCCGGCTCTGATTGCAGCCGGTGTTAGCCCGGA
CACGCCACCGATTCGCCGTGCGATTCGTTGGTTGGAGGAGCATCAGAACCCTGATGGCGGCT
GGGGTGAAGATCTGCGTAGCTACACCGACCCGGCGTTGTGGGTCGGTCGTGGTGTGTCCACC
GCGAGCCAAACGGCGTGGGCGCTGCTGGCGCTGTTGGCAGCCGGCGAAGAAGCATCCCCGGC

AGTCGACCGTGGCGTGCGTTGGCTGGTGACCACCCAGCAACCGGACGGCGGTTGGGATGAGC
CGCACTACACCGGTACGGGTTTCCCGGGTGACTTCTATATCAACTACCACCTGTATCGCCTG
GTTTTTCCGATCTCTGCACTGGGTCGCTATGTGAATCGTTAA

2.4 *Apa*SHC1

2.4.1 Amino acid sequence

MNMASRFSLKKILRSGSDTQGTNVNTLIQSGTSDIVRQKPAPQEPADLSALKAMGNSLTHTL
SSACEWLMKQQKPDGHWVGSVGSNASMEAEWCLALWFLGLEDHPLRPRLGKALLEMQRPDGS
WGTYYGAGSGDINATVESYAALRSLGYAEDDPAVSKAAAWIISKGGLKNVRVFTRYWLALIG
EWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMMPLAILSARRPSRPLRPQDRLDALFP
GGRANFDYELPTKEGRDVIADFFRLADKGLHWLQSSFLKRAPSREAAIKYVLEWIIWHQDAD
GGWGGIQPPWVYGLMALHGEGYQFHHPVMAKALDALNDPGWRHDKGDASWIQATNSPVWDTM
LSLMALHDANAEERFTPEMDKALDWLLSRQVRVKGDWSVKLPNTEPGGWAFEYANDRYPDTD
DTAVALIAIASCRNRPEWQAKGVEEAIGRGVRWLVAMQSSCGGWGAFDKDNNKSILAKIPFC
DFGEALDPPSVDVTAHVLEAFGLLGLPRDLPCIQRGLAYIRKEQDPTGPWFGRWGVNYLYGT
GAVLPALAALGEDMTQPYISKACDWLINCQQENGGWGESCASYMEVSSIGHGATTPSQTAWA
LMGLIAANRPQDYEAIAKGCRYLIDLQEEDGSWNEEEFTGTGFPGYGVGQTIKLDDPAISKR
LMQGAELSRAFMLRYDLYRQLFPIIALSRASRLIKLGN

2.4.2 DNA sequence (*E. coli* codon optimized)

ATGAATATGGCCAGCCGCTTTAGCCTGAAAAAAATTCTGCGTAGCGGTAGCGATACCCAGGG
CACCAATGTTAATACCCTGATTCAGAGCGGCACCAGCGATATTGTTCGTCAGAAACCGGCAC
CGCAGGAACCGGCAGATCTGAGCGCACTGAAAGCAATGGGTAATAGCCTGACCCATACCCTG
AGCAGCGCATGTGAATGGCTGATGAAACAGCAGAAACCGGATGGTCATTGGGTTGGTAGCGT
GGGTAGCAATGCAAGCATGGAAGCAGAATGGTGTCTGGCACTGTGGTTTCTGGGTCTGGAAG
ATCATCCGCTGCGTCCTCGTCTGGGTAAAGCACTGCTGGAAATGCAGCGTCCGGATGGTAGC
TGGGGCACCTATTATGGTGCAGGTAGCGGTGATATTAATGCAACCGTTGAAAGCTATGCAGC
ACTGCGTAGCCTGGGTTATGCAGAAGATGATCCGGCAGTTAGCAAAGCAGCAGCATGGATTA
TTAGCAAAGGTGGTCTGAAAAATGTGCGTGTTTTTACCCGTTATTGGCTGGCACTGATTGGT
GAATGGCCGTGGGAAAAAACCCCGAATCTGCCTCCGGAAATTATTTGGTTTCCGGATAATTT
TGTGTTTAGCATTTATAATTTTGCCCAGTGGGCACGTGCAACCATGATGCCGCTGGCAATTC
TGAGCGCACGTCGTCCGAGCCGTCCGCTGCGTCCGCAGGATCGTCTGGATGCACTGTTTCCG
GGTGGTCGTGCAAATTTTGATTATGAACTGCCGACCAAAGAAGGTCGCGACGTTATTGCAGA
TTTTTTTCGTCTGGCCGATAAAGGTCTGCATTGGCTGCAGAGCAGCTTTCTGAAACGTGCAC
CGAGCCGTGAAGCAGCAATTAAATATGTTCTGGAATGGATTATTTGGCATCAGGATGCAGAT
GGTGGTTGGGGTGGTATTCAGCCTCCGTGGGTTTATGGTCTGATGGCACTGCATGGTGAAGG
TTATCAGTTTCATCATCCGGTTATGGCAAAAGCACTGGATGCCCTGAATGATCCGGGTTGGC
GTCATGATAAAGGTGATGCAAGCTGGATTCAGGCAACCAATAGTCCGGTTTGGGATACCATG
CTGAGCCTGATGGCCCTGCATGATGCAAATGCAGAAGAACGTTTTACACCGGAAATGGATAA
AGCCCTGGATTGGCTGCTGAGCCGTCAGGTTCGTGTTAAAGGTGATTGGAGCGTTAAACTGC
CGAATACCGAACCGGGTGGTTGGGCATTTGAATATGCAAATGATCGTTATCCTGATACCGAT
GATACCGCAGTTGCACTGATTGCAATTGCAAGCTGTCGTAATCGTCCGGAATGGCAGGCAAA
AGGTGTTGAAGAAGCAATTGGTCGCGGAGTTCGTTGGCTGGTTGCAATGCAGAGCAGTTGTG
GTGGCTGGGGTGCATTTGATAAAGATAATAATAAAAGCATTCTGGCCAAAATTCCGTTTTGC
GATTTTGGTGAAGCACTGGACCCTCCGAGCGTTGATGTTACCGCACATGTTCTGGAAGCATT
TGGTCTGCTGGGTCTGCCTCGTGATCTGCCGTGTATTCAGCGTGGTCTGGCATATATTCGTA
AAGAACAGGACCCGACCGGTCCGTGGTTTGGTCGTTGGGGTGTTAATTATCTGTATGGCACC

GGTGCAGTTCTGCCTGCACTGGCAGCACTGGGTGAAGATATGACCCAGCCGTATATTAGCAA
AGCATGCGATTGGCTGATTAATTGTCAGCAGGAAAATGGCGGTTGGGGAGAAAGCTGTGCAA
GCTATATGGAAGTTAGCAGCATTGGTCATGGTGCAACCACCCCGAGCCAGACCGCATGGGCA
CTGATGGGTCTGATTGCAGCAAATCGTCCGCAGGATTATGAAGCAATTGCAAAAGGTTGCCG
CTATCTGATTGATCTGCAGGAAGAGGACGGCAGCTGGAACGAAGAAGAATTCACCGGCACCG
GTTTTCCGGGTTATGGTGTTGGTCAGACCATTAAACTGGATGATCCGGCTATTAGTAAACGT
CTGATGCAGGGTGCAGAACTGAGCCGTGCATTTATGCTGCGTTATGATCTGTATCGTCAGCT
GTTTCCGATTATTGCCCTGAGCCGTGCCAGCCGTCTGATTAAACTGGGTAATTAATAA

2.5 *Apa*SHC2

2.5.1 Amino acid sequence

MAADGSALSESRLSSEALDRAVLSAHTALSQAQQDDGHWVYELEADATIPAEYILLEHFMDR
IDDALEQKIAIYLRRIQSEEHGGWPLYHNGKFDLSATVKAYFALKAVGDDINAPHMQRAREA
ILDHGGAERSNVFTRSQLALFGEVPWRATPVMPVELMLLPAKAFFSVWNMSYWSRTVIAPLL
VLAALRPVAANPRQVHVRELFVTPPEKVQDWIRGPYRSAWGYVFKGLDSVLRPVVPFIPEKT
HKKAIQAALDFIEPRLNGKDGLGAIYPAMANVVMMYRAMGVPDEDPRAKTAWEAVQALIVEK
DDEAYCQPCVSPIWDTGLSGHAMIEAASGPNGIAPEKTVAELKKASAWLRSKQILNVKGDWA
VRNPNLAPGGWAFQYGNDYYPDVDDTAVVGMLLHREGDPTNAEAIERARTWIVGMQSTDGGW
GAFDIDNNKDVLNHIPFADHGALLDPPTADVTARCISFLAQLRNPEDEPVIQRGLEYLRKEQ
EKDGSWFGRWGTNYIYGTWSALCALNAAGVSHDDPAVVKAVEWLRSVQRADGGWGEGCESYE
GGPHGTYGESLPSQTAWAVLGLMAAGRRDDPAVTRGIAWLADQQDANGEWHEDPYNAVGFPK
VFYLRYHGYKQFFPLMALARYRNLESSNTRRVSFGF

2.5.2 DNA sequence

ATGGCCGCCGATGGGAGTGCTCTTTCCGAATCACGCCTTTCTTCAGAGGCTCTGGATCGTGC
GGTCCTTAGTGCGCATACGGCGCTCAGTCAGGCCCAGCAAGATGATGGACATTGGGTTTATG
AACTGGAAGCCGATGCCACCATTCCTGCTGAATATATCCTGCTCGAACATTTTATGGACAGG
ATTGATGATGCGCTGGAGCAGAAAATTGCCATCTACCTGCGCCGCATCCAAAGCGAAGAACA
CGGCGGCTGGCCCCTTTACCACAATGGCAAGTTTGACCTTTCAGCCACTGTAAAAGCCTATT
TCGCACTGAAAGCTGTGGGGGATGATATTAACGCCCCCCATATGCAACGTGCACGAGAAGCC
ATTCTGGATCATGGCGGCGCAGAACGCTCAAATGTATTCACACGCTCCCAGCTTGCTCTGTT
TGGGGAAGTGCCATGGCGTGCAACCCCGGTTATGCCGGTAGAGTTGATGCTTCTGCCTGCCA
AGGCATTCTTTTCCGTATGGAATATGTCTTACTGGTCTCGCACCGTTATTGCACCGCTTCTG
GTGCTGGCAGCCCTGCGCCCTGTGGCGGCAAACCCGCGGCAAGTTCATGTCCGCGAGCTGTT
TGTAACGCCACCAGAAAAAGTGCAGGACTGGATCCGCGGTCCTTATCGCTCTGCATGGGGGT
ATGTTTTTAAAGGGCTGGATAGCGTTTTACGGCCGGTCGTGCCGTTTATTCCCGAAAAAACA
CATAAAAAGGCTATTCAAGCCGCCCTTGATTTTATCGAGCCTCGCCTGAACGGCAAAGATGG
ATTGGGGGCTATTTACCCCGCCATGGCCAATGTGGTGATGATGTATCGGGCCATGGGCGTGC
CGGATGAAGACCCACGTGCAAAAACGGCATGGGAAGCCGTGCAGGCCCTCATCGTTGAAAAA
GACGACGAAGCTTACTGTCAGCCCTGCGTTTCCCCCATTTGGGACACCGGACTTTCTGGCCA
TGCCATGATTGAGGCAGCCTCCGGTCCCAATGGAATCGCACCAGAAAAAACTGTTGCTGAGC
TGAAAAAAGCCTCTGCATGGCTCCGCAGCAAGCAGATCCTGAACGTGAAGGGAGATTGGGCT
GTTCGTAACCCCAATCTGGCTCCCGGTGGGTGGGCTTTCCAATACGGAAACGACTATTACCC
GGATGTGGATGATACAGCCGTAGTAGGCATGTTGCTGCACCGTGAAGGCGACCCCACAAATG
CTGAAGCCATTGAGCGCGCACGCACATGGATTGTGGGCATGCAAAGCACAGATGGTGGCTGG
GGTGCTTTTGATATCGACAACAACAAGGATGTGCTCAACCACATTCCCTTTGCCGATCACGG
CGCCTTACTAGACCCGCCTACTGCCGATGTTACCGCCCGCTGCATCTCCTTTCTGGCCCAAT

TGCGGAACCCGGAAGATGAACCCGTTATTCAACGCGGGCTGGAATATCTACGCAAAGAGCAG
GAAAAAGATGGCTCCTGGTTTGGGCGCTGGGGCACAAACTACATTTACGGCACATGGTCTGC
CCTGTGCGCCCTGAATGCTGCTGGCGTTTCCCACGATGACCCTGCCGTGGTGAAAGCTGTGG
AATGGCTACGCTCCGTTCAGCGCGCAGATGGTGGCTGGGGTGAAGGTTGCGAATCTTATGAA
GGTGGCCCGCACGGCACATATGGCGAAAGCCTGCCCTCGCAAACTGCATGGGCTGTGCTAGG
GCTGATGGCCGCAGGGCGGCGGGATGATCCAGCCGTAACACGCGGTATTGCATGGTTGGCAG
ACCAGCAGGATGCGAACGGGGAATGGCATGAAGACCCCTATAATGCTGTTGGCTTCCCCAAA
GTGTTTTACCTGCGTTACCACGGCTATAAGCAGTTCTTCCCGCTTATGGCATTAGCACGCTA
CCGCAACCTTGAAAGCAGCAATACGCGCCGCGTTTCCTTTGGGTTCTAA

2.6 *Cac*SHC

2.6.1 Amino acid sequence

MTDVIDKAVAATGPADPSQGAAATLQAAADHLLGLQDDAGWWKGELETNVTMDAEDLLLRQF
LGIRTEEVTREAGDWIRSQQRADGTWANFFDGPADLSTTIEAYTALRMAGDAKDAEHMRAAR
TYILDSGGIEASRVFTRIWLALFGEWQWSDLPVMPPELIYLPKWFPLNVYDWACWARQTVVP
LTIVNALRPVRPLGFDLKELRTGRRAPAQRGLFSTLDRALHVYERKPLRSVRDAALRRSADW
IIARQEADGSWGGIQPPWVYSLMALNLLGYGVDHPVMRKGIEGLDRFTIRDERGRRLEACQS
PVWDTVLAMTALRDAELPENHPALVKAADWVLGEEITNPGDWSVRRPRVAPGGWAFEFDNDG
YPDVDDTAEVVLALNRVAHPDAPAAIRRGVDWLEGMACKDGGYGAFDADNTRTLALKLPFCD
FGAVIDPPTADVTAHTLEAYAALGLANSRASQRALEWLVKAQERDGSWFGRWGANHVYGTGA
VVPAMVAVGVDPEDEMIRRAVRWLEEHQNDDGGWGEDLRSYRDKSWIGRGVSTASQTAWALL
ALLAAGEERGTAVEQGVRFLIRTQRADGTWDEDHYTGTGFPGDFYLNYHLYRLVFPISALGR
YVRAVGAAGDGGDAGHAGHAGTVS

2.6.2 DNA sequence (*E. coli* codon optimized)

ATGACGGACGTAATCGATAAAGCAGTAGCAGCGACCGGCCCAGCGGACCCGAGCCAAGGTGC
GGCAGCGACCCTGCAAGCGGCTGCGGACCACCTGCTGGGCTTGCAGGACGACGCAGGTTGGT
GGAAAGGCGAACTGGAAACTAACGTGACGATGGACGCAGAAGATCTGCTGCTGCGCCAGTTC
CTGGGTATTCGTACCGAAGAGGTCACGCGCGAGGCTGGTGATTGGATTCGCAGCCAACAGCG
TGCTGATGGCACCTGGGCGAACTTCTTTGATGGTCCAGCGGATCTGAGCACCACCATCGAGG
CCTATACTGCACTGCGCATGGCGGGTGACGCGAAAGACGCCGAACACATGCGTGCTGCACGT
ACTTACATCCTGGACAGCGGCGGTATCGAGGCAAGCCGCGTCTTTACCCGTATTTGGCTGGC
TCTGTTTGGCGAGTGGCAGTGGAGCGATCTGCCGGTTATGCCACCGGAACTGATCTACTTGC
CGAAATGGTTTCCGCTGAACGTTTATGACTGGGCGTGTTGGGCCCGTCAAACCGTTGTTCCG
CTGACTATTGTCAATGCGCTGCGTCCGGTGCGTCCGTTGGGTTTCGATCTGAAGGAACTGCG
CACCGGTCGTCGTGCTCCAGCTCAGCGCGGTTTGTTTAGCACGTTGGACCGTGCATTGCACG
TCTATGAACGTAAGCCGCTGCGCTCGGTCCGTGATGCGGCACTGCGTCGCAGCGCGGATTGG
ATCATTGCGCGCCAGGAAGCGGACGGTTCCTGGGGCGGTATCCAACCGCCGTGGGTATACAG
CTTGATGGCGCTGAACTTGCTGGGCTATGGTGTGGATCATCCGGTCATGCGCAAGGGTATCG
AAGGCCTGGACCGTTTTACCATCCGCGACGAGCGTGGTCGTCGCCTGGAGGCGTGCCAGAGC
CCGGTTTGGGATACGGTGCTGGCGATGACCGCATTGCGCGACGCGGAACTGCCGGAGAACCA
TCCGGCACTGGTTAAGGCAGCCGATTGGGTGCTGGGTGAGGAGATTACGAATCCGGGCGACT
GGAGCGTCCGCCGTCCGCGCGTGGCACCGGGTGGTTGGGCCTTCGAGTTCGACAACGACGGT
TATCCGGATGTGGATGACACGGCAGAGGTTGTTTTGGCTCTGAACCGTGTTGCGCATCCGGA
CGCGCCTGCGGCAATCCGTCGTGGCGTGGATTGGTTGGAAGGTATGGCCTGCAAAGACGGCG
GCTACGGTGCGTTTGATGCTGACAATACCCGTACCCTGGCCCTGAAACTGCCGTTCTGTGAT
TTCGGTGCCGTCATTGACCCGCCGACCGCGGACGTGACGGCCCACACCCTGGAAGCATACGC

TGCACTGGGTCTGGCGAATTCTCGTGCGTCGCAACGCGCACTGGAGTGGCTGGTAAAGGCGC
AAGAGCGTGATGGTTCCTGGTTCGGTCGTTGGGGTGCGAATCACGTGTACGGCACGGGTGCG
GTCGTGCCGGCGATGGTTGCAGTCGGTGTTGACCCTGAGGATGAGATGATTCGTCGCGCAGT
CCGCTGGCTGGAAGAGCATCAGAATGACGATGGTGGCTGGGGCGAGGACCTGCGCTCTTACC
GTGATAAGAGCTGGATCGGCCGTGGCGTTAGCACCGCGTCCCAGACCGCCTGGGCGTTGCTG
GCACTGCTGGCTGCGGGCGAGGAGCGCGGCACGGCTGTGGAACAGGGTGTCCGTTTTCTGAT
TCGTACGCAACGTGCGGACGGTACCTGGGATGAAGATCACTATACTGGTACCGGTTTTCCGG
GTGATTTCTATCTGAATTACCACCTGTACCGTCTGGTGTTCCCGATTAGCGCCCTGGGTCGT
TACGTGCGTGCCGTTGGTGCCGCAGGTGACGGTGGCGATGCTGGTCATGCGGGCCACGCGGG
CACCGTGAGCTAA

2.7 *Gth*SHC

2.7.1 Amino acid sequence

MAGERSALITALKRSQAADGSWRFPFETGISTDAYMIILLRTLDINDEPLIQALVERIESRQ
EANGAWKLFADEGDGNVTATVEAYYALLYSGYRQPTDRHMQKAKRRILDMGGLDRVHLFTKV
MLALTGQYPWPGRFPLPLEFFLLPPSFPLNMYDLSVYGRANMIPLLIAADSRYSRKTDKSPD
LSDLFASRGDWGMPESRSLLTYVKRSLIGLPAQLHQAAKQRAVRYLFEHIEPDGTLYSYFSS
TFLFIFALLALGYRNDDPRIRQAVRGLRSLRTTIDGHVHLQYTTASVWNTALASYTLQEAGV
PMTDRAIEKANRYLLSRQNVRYGDWAVHNPYSTPGGWGFSDVNTMNPDVDDTTAALRAIRQA
AAKETAFRHAWDRANQWLFSMQNDDGGFAAFEKNVSSRFWRYLPIEGAEFLLMDPSTADLTG
RTLEYFGTFAGLTKDQRAVSRAVDWLLSHQERNGSWYGRWGICYIYGTWAAITGLTAVGVPA
HHPALQKAVRWLLSIQNDDGGWGESCKSDGAKTYVPLGDSTPVHTAWALDALVAAAERPTLE
MKAGFRALFRLLHHPDWTASYPVGQGMAGAFYIHYHSYRYIFPLLALAHYEQKFGPLDD

2.7.2 DNA sequence (*E. coli* codon optimized)

ATGGCAGGCGAACGTAGCGCACTGATTACCGCACTGAAACGTAGCCAGGCAGCAGATGGTAG
CTGGCGTTTTCCGTTTGAAACCGGTATTAGCACCGATGCCTATATGATTATTCTGCTGCGTA
CCCTGGATATTAATGATGAACCGCTGATTCAGGCACTGGTTGAACGTATTGAAAGCCGTCAA
GAGGCAAATGGTGCATGGAAACTGTTTGCAGATGAAGGTGATGGTAATGTTACCGCAACCGT
TGAAGCATATTACGCACTGCTGTATAGCGGCTATCGTCAGCCGACCGATCGTCACATGCAGA
AAGCAAAACGTCGTATTCTGGATATGGGTGGTCTGGATCGTGTTCACCTGTTTACCAAAGTT
ATGCTGGCACTGACAGGTCAGTATCCGTGGCCTGGTCGTTTTCCGCTGCCGCTGGAATTTTT
TCTGCTGCCTCCGAGCTTTCCGCTGAATATGTATGATCTGAGCGTTTATGGTCGTGCAAATA
TGATTCCGCTGCTGATTGCAGCAGATAGCCGTTATAGCCGTAAAACCGATAAAAGTCCGGAT
CTGAGCGACCTGTTTGCAAGCCGTGGTGATTGGGGTATGCCGGAAAGCCGTAGCCTGCTGAC
CTATGTTAAACGTAGCCTGATTGGTCTGCCTGCACAGCTGCATCAGGCAGCAAAACAGCGTG
CAGTTCGTTACCTGTTTGAACATATTGAACCGGATGGCACCCTGTATAGCTATTTTAGCAGC
ACCTTTCTGTTTATTTTTGCACTGCTGGCACTGGGTTATCGTAATGATGATCCGCGTATTCG
TCAGGCAGTTCGTGGTCTGCGTAGCCTGCGTACCACCATTGATGGTCATGTTCATCTGCAGT
ATACCACCGCAAGCGTTTGGAATACCGCACTGGCAAGCTATACCCTGCAAGAGGCAGGCGTT
CCGATGACCGATCGTGCAATTGAAAAAGCAAATCGTTATCTGCTGAGCCGTCAGAATGTTCG
TTATGGTGATTGGGCAGTTCATAATCCGTATAGCACACCGGGTGGTTGGGGTTTTAGTGATG
TTAATACCATGAATCCGGATGTTGATGATACCACCGCAGCACTGCGTGCAATTCGCCAGGCA
GCCGCAAAAGAAACCGCATTTCGTCATGCATGGGATCGTGCAAATCAGTGGCTGTTTAGCAT
GCAGAATGATGATGGTGGTTTTGCAGCCTTTGAAAAAAATGTTAGCAGCCGTTTTTGGCGTT
ATCTGCCGATTGAAGGTGCAGAATTTCTGCTGATGGACCCGAGCACCGCAGATCTGACCGGT
CGTACCCTGGAATACTTTGGCACCTTTGCAGGTCTGACCAAAGATCAGCGTGCCGTTAGCCG

TGCAGTTGATTGGCTGCTGAGCCATCAAGAACGTAATGGTAGCTGGTATGGTCGTTGGGGTA
TTTGTTATATTTATGGCACCTGGGCAGCAATTACCGGTCTGACCGCAGTTGGTGTTCCGGCA
CATCATCCGGCACTGCAGAAAGCAGTTCGTTGGCTGCTGTCAATTCAGAACGATGATGGCGG
TTGGGGTGAAAGCTGTAAAAGTGATGGTGCAAAAACCTATGTTCCGCTGGGTGATAGTACAC
CGGTTCATACCGCATGGGCACTGGATGCACTGGTTGCAGCAGCAGAACGTCCGACCCTGGAA
ATGAAAGCAGGTTTTCGTGCACTGTTTCGTCTGCTGCATCATCCGGATTGGACCGCAAGCTA
TCCGGTTGGTCAGGGTATGGCAGGCGCATTTTATATTCATTATCACAGCTACCGCTACATTT
TTCCGCTGCTGGCCCTGGCACATTATGAACAGAAATTTGGTCCGCTGGATGATTAA

2.8 *Mca*SHC

2.8.1 Amino acid sequence

MLREATAISNLEPPLTASYVESPLDAAIRQAKDRLLSLQHLEGYWVFELEADCTIPAEYILM
MHFMDEIDAALQAKIANYLRSHQSADGSYPLFRGGAGDISCTVKVYYALKLAGDSIDAPHMK
KAREWILAQGGAARSNVFTRIMLAMFEQIPWRGIPFIPVEIMLLPKWFPFHLDKVSYWSRTV
MVPLFILCSHKVTARNPSRIHVRELFTVDPQKERHYFDHVKTPLGKAILALERFGRMLEPLI
PKAVRKKATQKAFDWFTARLNGVDGLGAIFPAMVNAYEALDFLGVPPDDERRRLARESIDRL
LVFQGDSVYCQPCVSPIWDTALTSLTLQEVARHTADLRLDAALSKGLKWLASKQIDKDAPGD
WRVNRAGLEGGGWAFQFGNDYYPDVDDSAVVAHALLGSEDPSFDDNLRRAANWIAGMQSRNG
GFGAFDADNTYYYLNSIPFADHGALLDPPTADVSARCAMFLARWVNRQPELRPVLERTIDYL
RREQEADGSWFGRWGTNYIYGTWSVLLAYEAAGVPNDDPSVRRAVAWLKSIQREDGGWGEDN
FSYHDPSYRGRFHTSTAFQTGFALIALMAAGEAGSPEVQAGVDYLLRQQRPDGFWNDECFTA
PGFPRVFYLKYHGYDKFFPLWALARYRNERYALA

2.8.2 DNA sequence (*E. coli* codon optimized)

ATGCTGCGTGAAGCAACCGCAATTAGCAATCTGGAACCGCCTCTGACCGCAAGCTATGTTGA
AAGTCCGCTGGATGCAGCAATTCGTCAGGCAAAAGATCGTCTGCTGAGCCTGCAGCATCTGG
AAGGTTATTGGGTTTTTGAACTGGAAGCCGATTGTACCATTCCGGCAGAATATATCCTGATG
ATGCACTTTATGGACGAAATTGATGCAGCACTGCAGGCAAAAATTGCAAATTATCTGCGTAG
CCATCAGAGCGCAGATGGTAGCTATCCGCTGTTTCGTGGTGGTGCCGGTGATATTAGCTGTA
CCGTTAAAGTTTACTACGCACTGAAACTGGCAGGCGATAGCATTGATGCACCGCACATGAAA
AAAGCACGTGAATGGATTCTGGCACAGGGTGGTGCAGCACGTAGCAATGTTTTTACCCGTAT
TATGCTGGCAATGTTTGAGCAGATTCCGTGGCGTGGTATTCCGTTTATTCCGGTTGAAATTA
TGCTGCTGCCGAAATGGTTTCCGTTTCATCTGGATAAAGTGAGCTATTGGAGCCGTACCGTT
ATGGTTCCGCTGTTTATTCTGTGTAGCCATAAAGTTACCGCACGTAATCCGAGCCGTATTCA
TGTTCGTGAACTGTTTACCGTTGATCCGCAGAAAGAACGCCATTATTTTGATCATGTGAAAA
CACCGCTGGGTAAAGCAATTCTGGCACTGGAACGTTTTGGTCGTATGCTGGAACCGCTGATT
CCGAAAGCAGTTCGTAAAAAAGCAACCCAGAAAGCCTTTGATTGGTTTACCGCACGTCTGAA
TGGTGTTGATGGTCTGGGTGCAATTTTTCCGGCAATGGTTAATGCCTATGAAGCACTGGATT
TTCTGGGTGTTCCGCCTGATGATGAACGTCGTCGTCTGGCACGTGAAAGCATTGATCGCCTG
CTGGTTTTTCAGGGTGATAGCGTTTATTGTCAGCCGTGTGTTAGCCCGATTTGGGATACCGC
ACTGACCAGCCTGACCCTGCAAGAAGTTGCACGTCATACCGCAGATCTGCGTCTGGATGCAG
CCCTGAGCAAAGGTCTGAAATGGCTGGCAAGCAAACAAATTGATAAAGATGCACCGGGTGAT
TGGCGTGTTAATCGTGCAGGTCTGGAAGGTGGTGGTTGGGCATTTCAGTTTGGCAATGATTA
TTATCCGGATGTTGATGATAGCGCAGTTGTTGCACATGCACTGCTGGGTAGCGAAGATCCGA
GCTTTGATGATAATCTGCGTCGTGCAGCAAATTGGATTGCAGGTATGCAGAGCCGTAATGGT
GGTTTTGGTGCATTTGATGCCGATAACACCTATTATTACCTGAACAGCATTCCGTTTGCAGA
TCATGGTGCACTGCTGGACCCTCCGACCGCAGATGTTAGCGCACGTTGTGCAATGTTTCTGG

CACGTTGGGTTAATCGTCAGCCGGAACTGCGTCCGGTTCTGGAACGTACCATTGATTATCTG
CGTCGCGAACAAGAAGCAGACGGTAGCTGGTTTGGTCGTTGGGGCACCAATTATATTTATGG
CACCTGGTCAGTTCTGCTGGCGTATGAAGCAGCCGGTGTTCCGAATGATGATCCGAGCGTTC
GTCGTGCAGTTGCATGGCTGAAAAGCATTCAGCGTGAAGATGGTGGCTGGGGTGAAGATAAT
TTTAGCTATCATGATCCGAGCTATCGTGGTCGTTTTCATACCAGCACCGCATTTCAGACCGG
TTTTGCACTGATTGCCCTGATGGCAGCCGGTGAAGCAGGTAGTCCGGAAGTTCAGGCAGGCG
TGGATTATCTGCTGCGTCAGCAGCGTCCGGATGGTTTTTGGAATGATGAATGTTTTACCGCA
CCGGGTTTTCCGCGTGTTTTTTATCTGAAATATCATGGCTATGATAAATTTTTCCCGCTGTG
GGCACTGGCACGTTATCGTAATGAACGTTATGCACTGGCCTAA

2.9 *Pca*SHC1

2.9.1 Amino acid sequence

MDKIKMKNINQPKFRVFRGGQKAATPCPGTTNERRGALDRGRLSASLKHSREWLLSLQADAG
NWVFALEADTTIASEYVMLQRFLGRPLAPELQQRLANYLLSRQLPDGGWPLYAEDGFANIST
TVKAYLALKLLGYPTHCDPLVRARQIVLALGGAEKCNVFTRIALALFGQIPWRTTPAMPVEI
MLLPRWFYFHLSKISYWARTVVVPLLILYAKRPVCRLEPWEGIPELFVTPPDKLGYLDVCKP
GQWRKNVFIWVDRLTRKMVRCVPRRLHNLALRAAETWTREHMQGAGGIGAIFPAMANAVMAL
RTLGCSPDDADYQRGLKALDDLLIDRCDVPPREDTPVSPCWCTGTSAAPMLDPSPAGSHAQG
GDQGICQPCASPIWDTGLALTALLEGGLDARHPAVDRAVRWLLDQQVDVKGDWAQRVPNLEA
GGWAFQFENALYPDLDDTSKVLMSLIRAGAMDNPGYRQELSRAINWVIGMQNSDGGWGAFDV
DNNYLYLNDIPFADHGALLDPSTADVTGRCIEMLAMAGFGRDFLPIARGVDFLRREQEDFGG
WYGRWGVNYIYGTWSALSGLIHAGEDLQAPYIRQAVGWLESVQNPDGGWGETCYSYDDPALA
GRGVSTASQTAWALLGLMAAGEVDNLAVRRGIQYLVEEQNRAGGWDERHFTGTGFPRVFYLR
YHGYSQYFPLWALGLYERLSSGNPSRQQMVRRAGPAGLHLPVLDRRKKLRRKRKA

2.9.2 DNA sequence

ATGGATAAAATCAAAATGAAAAACATAAACCAGCCCAAATTTAGAGTTTTTCGTGGTGGACA
AAAAGCCGCTACGCCTTGTCCGGGGACGACTAACGAGAGACGCGGTGCCCTGGATCGCGGTC
GTCTGTCGGCTTCCCTCAAGCATTCTCGCGAATGGCTGCTTTCGCTGCAGGCCGACGCCGGT
AATTGGGTTTTTGCGTTGGAAGCGGATACCACCATCGCTTCCGAATATGTGATGCTGCAACG
TTTTCTCGGTCGCCCTCTTGCGCCTGAATTGCAACAACGTCTGGCCAATTATTTACTCAGTC
GTCAATTGCCCGACGGTGGCTGGCCGTTATATGCGGAAGACGGGTTTGCCAATATCAGCACC
ACCGTCAAGGCTTACCTGGCGCTTAAGCTGTTGGGTTACCCGACCCACTGCGACCCCCTGGT
GCGGGCGCGGCAAATCGTTTTGGCCCTCGGCGGTGCCGAAAAATGCAATGTGTTCACGCGCA
TCGCGCTGGCGCTGTTCGGGCAGATTCCCTGGCGCACGACTCCGGCCATGCCGGTTGAAATC
ATGCTTTTGCCGCGCTGGTTCTATTTTCATTTAAGTAAGATTTCCTATTGGGCTCGTACCGT
GGTGGTGCCGTTGCTGATTCTGTACGCCAAACGCCCGGTCTGCCGTCTGGAGCCCTGGGAAG
GGATCCCTGAGCTGTTTGTCACGCCGCCGGATAAACTCGGTTACCTCGATGTCTGTAAACCC
GGTCAGTGGCGTAAAAATGTCTTTATCTGGGTGGATCGCCTGACCCGCAAAATGGTGCGCTG
TGTCCCCCGGCGTCTGCACAACCTGGCGCTGAGGGCTGCAGAGACATGGACACGGGAGCATA
TGCAGGGCGCCGGAGGTATCGGGGCTATTTTCCCGGCCATGGCCAATGCCGTCATGGCGCTG
CGGACTCTGGGCTGCTCGCCGGATGATGCCGATTATCAGCGCGGCCTCAAGGCTCTCGACGA
TCTGCTGATTGACCGTTGTGACGTTCCTCCCCGGGAGGATACGCCGGTTTCGCCGTGCTGGT
GCACAGGCACCTCAGCCGCTCCGATGCTCGATCCCAGCCCTGCCGGCAGCCATGCGCAGGGT
GGCGATCAGGGTATCTGTCAACCTTGTGCGTCGCCGATCTGGGATACGGGACTTGCCCTTAC
GGCGCTGCTTGAAGGGGGGCTTGATGCCAGGCATCCGGCGGTCGATCGTGCGGTTCGCTGGC
TGCTGGATCAGCAGGTCGATGTCAAGGGCGACTGGGCGCAGCGGGTGCCGAACCTCGAAGCG

GGCGGTTGGGCATTTCAGTTCGAAAACGCTCTGTATCCCGATCTGGACGATACCAGCAAGGT
GCTGATGTCCCTGATACGCGCCGGTGCGATGGATAACCCGGGCTATCGACAGGAGCTGTCGC
GGGCTATCAATTGGGTTATCGGCATGCAGAACAGCGATGGAGGATGGGGTGCCTTCGACGTT
GACAATAATTACCTTTATTTAAATGATATCCCTTTCGCCGATCATGGGGCGTTGCTCGATCC
CAGTACTGCGGATGTGACGGGGCGGTGCATCGAAATGCTCGCTATGGCAGGTTTCGGCCGGG
ATTTTTTGCCCATTGCCAGGGGGGTGGATTTCCTGCGTCGTGAGCAGGAGGACTTCGGCGGT
TGGTATGGTCGCTGGGGCGTGAACTATATTTATGGGACCTGGTCGGCCCTGTCCGGGTTGAT
CCACGCCGGCGAGGATTTGCAGGCTCCTTATATCCGGCAGGCGGTGGGCTGGCTTGAATCGG
TACAGAACCCGGATGGTGGATGGGGCGAAACCTGTTATTCCTATGACGATCCCGCCCTGGCC
GGACGTGGCGTCAGCACCGCCTCGCAGACAGCCTGGGCGCTCCTGGGGTTGATGGCCGCCGG
CGAGGTGGACAACCTGGCAGTGCGGCGCGGGATTCAGTATTTGGTGGAGGAACAAAACCGAG
CCGGGGGCTGGGATGAACGCCATTTCACCGGAACCGGTTTCCCTCGGGTCTTTTATTTGCGT
TACCACGGGTACAGTCAGTACTTTCCCCTCTGGGCCCTCGGTCTGTACGAACGGCTCAGCTC
CGGGAACCCGAGCAGGCAGCAGATGGTACGGCGGGCGGGGCCTGCCGGTTTGCATCTGCCGG
TTCTCGACCGGCGCAAAAAACTACGTCGCAAGCGCAAAGCGTAA

2.10*Pca*SHC2

2.10.1Amino acid sequence

MNVIRQLNSGVNAAKSLDDGIESAIEWLAENQDKEGFWVGMLESNSCIEAEWILAMHLLGVK
DDPKYDKVVQAILNEQREDGSWAVYYDAPAGDINATVEAYAALRTAGFGAGDERLIKARNWI
FSHGGLKNVRVFTRYWLALIGEWPWDETPALAPEIIYLPAWCPLNIYDFACWARATLVPLSV
LSVRRPVKPLPAESRLDELFPEGRENADYSLPESEKGLAERFFLVVDWFLKKYNRLPMQFGR
EKAIRLCLEWIVRHQDYDGGWGGIQPPLIYSLIALNTEGYGINHPVISKGLDAFNPPWAYEK
NGGVYLQCSESPVWDTLFTMLALFESGCSFDDTPMMRPALDWILSKQITSWGDWQVKVRGVR
PGGWAFERANTAYPDVDDTALALVVLAEARRHVKDSAAVDAALERAEEWILGLQCRNGGWAA
FDRDNNSAIVTKIPFCDFGEVLDPPSVDVTAHVVEALAALGRDRHDPVVARALKYIRSEQEP
GGSWFGRWGVNHIYGTCAVLPALAAIGEDMRAPYVLRAADWLVRHQNDDGGWGESCASYMDD
SQCGQGSSTASQTGWALMALVAMSSHDYDEAIRRGLDYLLSHQKSGTWDEPQYTGTGFPGYG
VGERTNLKEAGATLDQGCELARGFMINYNMYRHYFPLIAMARARRHLGLAANPRHQDSRSSV
EVAPEALRGRACG

2.10.2DNA sequence (*E. coli* codon optimized)

ATGAATGTGATTCGCCAGCTGAATAGCGGTGTTAATGCAGCAAAAAGCCTGGATGATGGTAT
TGAAAGCGCAATTGAATGGCTGGCAGAAAATCAGGATAAAGAAGGTTTTTGGGTTGGTATGC
TGGAAAGCAATAGCTGTATTGAAGCAGAGTGGATTCTGGCAATGCATCTGCTGGGTGTTAAA
GATGATCCGAAATATGATAAAGTGGTGCAGGCCATTCTGAATGAACAGCGTGAAGATGGTAG
CTGGGCAGTTTATTATGATGCACCGGCAGGAGATATTAACGCAACCGTTGAAGCCTATGCAG
CACTGCGTACCGCAGGTTTTGGTGCCGGTGATGAACGTCTGATTAAAGCACGTAATTGGATT
TTTAGCCATGGTGGTCTGAAAAATGTGCGTGTTTTTACCCGTTATTGGCTGGCACTGATTGG
TGAATGGCCGTGGGATGAAACACCGGCACTGGCACCGGAAATTATTTATCTGCCTGCATGGT
GTCCGCTGAATATTTATGATTTTGCATGTTGGGCACGTGCAACCCTGGTTCCGCTGAGCGTT
CTGAGCGTGCGTCGTCCGGTTAAACCGCTGCCTGCAGAAAGCCGTCTGGATGAACTGTTTCC
GGAAGGTCGTGAAAATGCAGATTATAGCCTGCCGGAAAGCGAAAAAGGTCTGGCAGAACGTT
TTTTTCTGGTTGTTGATTGGTTTCTGAAAAAATATAATCGCCTGCCGATGCAGTTTGGTCGT
GAAAAAGCAATTCGTCTGTGTCTGGAATGGATTGTTCGTCACCAGGATTACGATGGTGGTTG
GGGTGGTATTCAGCCTCCGCTGATTTATAGCCTGATTGCACTGAATACCGAAGGCTATGGTA
TTAATCATCCGGTTATTAGCAAAGGTCTGGATGCATTTAATCCTCCGTGGGCCTATGAAAAA

AATGGTGGTGTTTATCTGCAGTGTAGCGAAAGTCCGGTTTGGGATACCCTGTTTACCATGCT
GGCACTGTTTGAAAGCGGTTGTAGCTTTGATGATACCCCGATGATGCGTCCGGCACTGGATT
GGATTCTGAGCAAACAAATCACCAGCTGGGGTGATTGGCAGGTTAAAGTTCGTGGTGTTCGT
CCGGGTGGTTGGGCATTTGAACGTGCAAATACCGCATATCCGGATGTTGATGATACCGCACT
GGCACTGGTTGTTCTGGCCGAAGCACGTCGTCATGTTAAAGATAGCGCAGCAGTTGATGCAG
CACTGGAACGTGCAGAAGAATGGATTCTGGGTCTGCAGTGTCGTAATGGTGGCTGGGCAGCA
TTTGATCGTGATAATAATAGCGCCATTGTGACCAAAATTCCGTTTTGTGATTTTGGTGAAGT
TCTGGACCCTCCGAGCGTTGATGTGACCGCACATGTTGTTGAAGCACTGGCAGCACTGGGTC
GTGATCGTCATGATCCGGTTGTTGCACGTGCACTGAAATATATTCGTAGCGAACAGGAACCG
GGAGGTAGCTGGTTTGGTCGTTGGGGTGTTAATCATATTTATGGCACCTGTGCAGTTCTGCC
TGCACTGGCTGCAATTGGTGAAGATATGCGTGCACCGTATGTTCTGCGTGCAGCAGATTGGC
TGGTGCGTCATCAGAATGATGATGGTGGATGGGGTGAAAGCTGTGCAAGCTATATGGATGAT
AGCCAGTGTGGTCAGGGTAGCAGCACCGCAAGCCAGACCGGTTGGGCACTGATGGCCCTGGT
TGCAATGAGCAGCCATGATTATGATGAAGCCATTCGTCGTGGTCTGGATTATCTGCTGAGCC
ATCAGAAAAGCGGCACCTGGGATGAACCGCAGTATACCGGCACCGGTTTTCCGGGTTATGGT
GTTGGTGAACGTACCAATCTGAAAGAAGCAGGCGCAACACTGGATCAGGGTTGTGAACTGGC
ACGTGGTTTTATGATTAATTATAATATGTATCGCCATTATTTTCCGCTGATTGCAATGGCAC
GTGCCCGTCGTCATCTGGGTCTGGCAGCAAATCCGCGTCATCAGGATAGCCGTAGCAGCGTT
GAAGTTGCACCGGAAGCACTGCGTGGTCGTGCATGTGGTTAATAAGGA

2.11*Rpa*SHC1

2.11.1Amino acid sequence

MDSILAPRADAPRNIDGALRESVQQAADWLVANQKPDGHWVGRAETNATMEAQWCLALWFLG
LEDHPLRVRLGRALLDTQRPDGAWHVFYGAPNGDINATVEAYAALRSLGHRDDEEPLRKARD
WILSKGGLANIRVFTRYWLALIGEWPWEKTPNILPEVIWLPTWFPFSIYNFAQWARATLMPI
AVLSAHRPSRPLAPQDRLDALFPQGRDSFNYDLPARLGAGVWDVIFRKIDTILHRLQDWGAR
RGPHGIMRRGAIDHVLQWIIRHQDYDGSWGGIQPPWIYGLMALHTEGYAMTHPVMAKALDAL
NEPGWRIDIGDATFIQATNSPVWDTMLSLLAFDDAGLGERYPEQVERAVRWVLKRQVLVPGD
WSVKLPDVKPGGWAFEYANNFYPDTDDTSVALMALAPFRHDPKWQAEGIEDAIQRGIDWLVA
MQCKEGGWGAFDKDNDKKILAKIPFCDFGEALDPPSADVTAHIIEAFAKVGLDRNHPSIVRA
LDYLKREQEPEGPWFGRWGVNYVYGTGAVLPALAAIGEDMRQPYIARACDWLIARQQANGGW
GESCVSYMDAKQAGEGTATASQTAWALMALIAADRPQDRDAIERGCLYLTETQRDGTWQEVH
YTGTGFPGYGVGQTIKLNDPLLSKRLMQGPELSRSFMLRYDLYRHYFPMMAIGRVLRQRGDR
SGH

2.11.2DNA sequence

ATGGATTCTATTCTGGCACCGCGGGCCGACGCGCCGCGCAATATCGACGGGGCGTTGCGGGA
GAGCGTGCAGCAGGCGGCCGACTGGCTGGTCGCCAACCAGAAGCCGGACGGCCACTGGGTCG
GGCGCGCCGAGACCAACGCCACCATGGAGGCGCAATGGTGCCTGGCGCTGTGGTTCCTCGGC
CTCGAGGATCATCCGCTGCGGGTTCGGCTCGGCCGCGCGCTGCTCGATACCCAGCGCCCCGA
CGGCGCCTGGCACGTGTTTTACGGCGCGCCGAACGGCGACATCAACGCCACGGTCGAGGCCT
ATGCGGCGCTGCGTTCGCTCGGCCATCGCGACGATGAAGAGCCGCTGCGCAAGGCGCGCGAC
TGGATTCTGTCGAAGGGCGGCCTCGCCAACATCCGCGTCTTCACCCGCTACTGGTTGGCGCT
GATCGGCGAGTGGCCGTGGGAGAAGACGCCGAACATTCTGCCCGAAGTGATCTGGCTGCCGA
CCTGGTTTCCGTTCTCGATCTATAATTTCGCGCAATGGGCCCGCGCCACGCTGATGCCGATC
GCGGTGCTGTCGGCGCATCGGCCGAGCCGGCCGTTGGCGCCGCAAGACCGGCTCGACGCGCT
GTTTCCGCAAGGCCGCGACAGCTTCAACTACGATCTGCCGGCGCGGTTAGGCGCCGGGGTGT

GGGATGTCATCTTCCGCAAGATCGACACCATTCTGCATCGCCTGCAGGACTGGGGCGCCAGA
CGCGGCCCGCACGGCATCATGCGCCGCGGCGCGATCGATCACGTGCTGCAATGGATCATCCG
TCATCAGGACTATGACGGCAGCTGGGGCGGCATCCAGCCGCCCTGGATCTACGGGTTGATGG
CGCTGCATACCGAGGGCTACGCCATGACCCATCCGGTGATGGCGAAAGCGCTCGACGCGCTG
AACGAACCCGGCTGGCGCATTGACATCGGCGACGCCACCTTCATCCAGGCCACCAATTCGCC
GGTGTGGGACACCATGCTGTCGCTGCTGGCGTTCGACGACGCCGGCCTCGGCGAACGCTACC
CTGAGCAGGTCGAGCGCGCGGTGCGCTGGGTGCTGAAGCGCCAGGTGCTCGTGCCCGGCGAT
TGGTCGGTGAAGCTGCCCGACGTCAAGCCGGGCGGCTGGGCGTTCGAATACGCCAACAATTT
CTATCCCGACACCGACGATACCTCGGTGGCGCTGATGGCGCTGGCGCCGTTCCGGCACGATC
CGAAATGGCAGGCCGAAGGCATCGAGGATGCGATCCAGCGCGGCATCGACTGGCTGGTGGCG
ATGCAGTGCAAGGAAGGCGGCTGGGGCGCCTTCGACAAGGACAACGACAAGAAGATTCTGGC
CAAGATTCCGTTCTGCGATTTCGGCGAGGCGCTCGACCCGCCGTCGGCCGACGTCACCGCGC
ATATCATCGAGGCCTTCGCCAAGGTCGGGCTCGACCGCAACCATCCCTCGATCGTTCGCGCG
CTGGATTATCTGAAGCGCGAGCAGGAGCCGGAGGGCCCGTGGTTCGGCCGCTGGGGCGTCAA
CTACGTCTACGGCACCGGCGCGGTGCTGCCGGCCTTGGCCGCGATCGGCGAGGACATGCGCC
AGCCCTATATCGCGCGCGCCTGCGACTGGCTGATCGCGCGGCAGCAGGCCAATGGCGGCTGG
GGCGAAAGCTGCGTCTCCTACATGGACGCCAAGCAGGCCGGCGAAGGCACCGCCACCGCCTC
GCAGACCGCGTGGGCGCTGATGGCGCTGATCGCCGCCGACCGGCCGCAGGACCGCGACGCGA
TCGAGCGCGGCTGCCTGTATCTGACCGAGACCCAGCGCGACGGCACCTGGCAGGAAGTGCAC
TACACCGGCACCGGCTTTCCCGGCTACGGCGTCGGCCAGACCATCAAGCTGAACGATCCGTT
GTTGTCGAAGCGGCTGATGCAGGGACCAGAACTGTCGCGCTCCTTCATGCTGCGCTACGACC
TCTACCGCCACTATTTTCCGATGATGGCGATCGGGCGGGTGCTGCGGCAGCGCGGTGATCGG
TCAGGGCATTGA

2.12*Rpa*SHC2

2.12.1Amino acid sequence

MESGNNKQPAAAIGALDASIESATNALLGYRQPDGHWVFELEADCTIPAEYVLLRHYLGEPV
DAALEAKIANYLRRVQGAHGGWPLVHDGGFDMSASVKGYFALKMIGDDIDAPHMAKAREAIR
SRGGAIHSNVFTRFLLSMFGITTWRSVPVLPVEIMLLPMWSPFHLNKISYWARTTIVPLMVL
AALKPRAVNRLDIGLDELFLQDPKSIKMPAKAPHQSWALFKLFAGIDAVLRTIEPLFPKRLR
DHAIKLAVDFVEERLNGEDGLGAIYPPMANTVMMYKVLGFPEDHPPRAITRRGIDKLLVIGE
DEAYCQPCVSPVWDTALTCHALLEVGGEAAVPPAKRGMDWLLPKQVLDLKGDWAVKRPNLRP
GGWAFQYNNAHYPDLDDTAVVVMAMDRSRRATGSREYDEAIARAREWIEGMQSDDGGWAAFD
VNNLEYYLNNIPFSDHGAMLDPPTEDVTARCVSMLSQLGETAASSKAVADGVEYLRRTQLPD
GSWYGRWGLNYIYGTWSVLCALNAAGVDHQDPVIRKAVTWLASVQNPDGGWGEGAESYRLNY
TRYEQAPTTASQTSWALLGLMAAGEVDSPVVARGVEYLKSTQTGKGLWDEQRYTATGFPRVF
YLRYHGYAKFFPLWALARYRNLRSTNSKVVGVGM

2.12.2DNA sequence

ATGGAGTCCGGGAACAACAAGCAGCCCGCGGCGGCAATCGGCGCTCTCGATGCGAGCATCGA
GAGCGCGACCAACGCCTTGCTGGGCTATCGGCAGCCCGACGGGCACTGGGTGTTCGAACTTG
AGGCCGACTGCACCATTCCTGCGGAATACGTGCTGCTGCGGCATTACCTCGGCGAGCCGGTC
GACGCCGCGTTGGAGGCCAAGATCGCCAACTATCTGCGCCGCGTGCAGGGCGCCCATGGCGG
CTGGCCGCTGGTGCACGACGGCGGCTTCGACATGAGCGCCAGCGTCAAGGGCTACTTCGCGC
TGAAGATGATCGGTGACGACATCGACGCGCCGCACATGGCGAAGGCGCGCGAGGCGATCCGC
TCGCGCGGCGGCGCGATCCACAGCAACGTGTTCACCCGCTTCCTGCTGTCGATGTTCGGCAT
CACCACCTGGCGCAGCGTGCCGGTGCTGCCGGTCGAGATCATGCTGCTGCCGATGTGGTCGC

CGTTCCATCTCAACAAGATCTCCTATTGGGCGCGCACCACCATCGTGCCGCTGATGGTGCTG
GCGGCCTTGAAGCCGCGCGCGGTCAACCGGCTCGACATCGGACTCGACGAACTGTTCTTGCA
GGATCCGAAGTCGATCAAGATGCCGGCCAAGGCGCCGCATCAGAGCTGGGCGCTGTTCAAGC
TGTTCGCCGGCATCGATGCGGTGTTGCGCACGATCGAGCCGTTGTTCCCGAAGCGGCTGCGC
GATCATGCGATCAAGCTCGCGGTGGATTTCGTCGAGGAGCGGCTGAACGGCGAGGACGGGCT
CGGCGCGATCTATCCGCCGATGGCCAACACCGTGATGATGTACAAGGTGCTGGGCTTTCCCG
AGGATCATCCGCCGCGCGCGATCACCCGGCGCGGCATCGACAAGCTGTTGGTGATCGGCGAG
GACGAAGCCTATTGCCAGCCTTGCGTGTCGCCGGTGTGGGACACCGCGCTGACCTGCCACGC
GCTGCTCGAAGTCGGCGGCGAGGCGGCGGTGCCGCCGGCCAAGCGCGGTATGGACTGGCTGC
TGCCCAAGCAGGTGCTCGACCTCAAGGGCGACTGGGCGGTGAAGCGGCCGAACCTGCGGCCC
GGCGGCTGGGCGTTCCAGTACAACAACGCGCACTATCCAGACCTCGACGACACCGCGGTGGT
GGTGATGGCGATGGACCGCTCGCGCCGCGCCACCGGCAGCCGCGAATATGACGAGGCGATCG
CCCGGGCCCGGGAGTGGATCGAGGGCATGCAGTCCGACGACGGCGGCTGGGCGGCGTTCGAC
GTCAACAATCTGGAATATTACCTCAACAACATCCCGTTCTCCGACCACGGCGCGATGCTCGA
CCCGCCGACCGAGGACGTCACCGCGCGCTGTGTTTCGATGCTGTCACAGCTCGGCGAGACCG
CGGCGAGCAGCAAGGCGGTCGCCGACGGCGTCGAATATCTGCGCAGGACTCAGCTGCCGGAC
GGCTCCTGGTACGGCCGCTGGGGGCTGAATTACATCTACGGCACCTGGTCGGTGCTGTGCGC
GCTGAACGCCGCCGGGGTCGATCATCAGGATCCGGTGATTCGCAAGGCGGTGACCTGGCTGG
CTTCGGTCCAGAACCCCGACGGCGGTTGGGGCGAGGGTGCCGAGAGCTACCGGCTGAATTAC
ACGCGATACGAGCAGGCGCCGACCACCGCCTCGCAGACCTCATGGGCTTTGCTCGGCCTGAT
GGCGGCCGGTGAGGTGGATTCCCCCGTAGTTGCCCGCGGCGTGGAGTACCTAAAAAGCACAC
AGACCGGAAAAGGGCTCTGGGACGAGCAGCGATACACCGCGACGGGCTTTCCGCGGGTGTTT
TATTTGCGTTATCATGGCTATGCGAAGTTCTTTCCGCTGTGGGCGCTGGCGCGGTATCGAAA
CCTGAGGAGCACCAACAGTAAGGTGGTAGGGGTCGGGATGTGA

2.13*Sco*SHC

2.13.1Amino acid sequence

MTATTDGSTGASLRPLAASASDTDITIPAAAAGVPEAAARATRRATDFLLAKQDAEGWWKGD
LETNVTMDAEDLLLRQFLGIQDEETTRAAALFIRGEQREDGTWATFYGGPGELSTTIEAYVA
LRLAGDSPEAPHMARAAEWIRSRGGIASARVFTRIWLALFGWWKWDDLPELPPELIYFPTWV
PLNIYDFGCWARQTIVPLTIVSAKRPVRPAPFPLDELHTDPARPNPPRPLAPVASWDGAFQR
IDKALHAYRKVAPRRLRRAAMNSAARWIIERQENDGCWGGIQPPAVYSVIALYLLGYDLEHP
VMRAGLESLDRFAVWREDGARMIEACQSPVWDTCLATIALADAGVPEDHPQLVKASDWMLGE
QIVRPGDWSVKRPGLPPGGWAFEFHNDNYPDIDDTAEVVLALRRVRHHDPERVEKAIGRGVR
WNLGMQSKNGAWGAFDVDNTSAFPNRLPFCDFGEVIDPPSADVTAHVVEMLAVEGLAHDPRT
RRGIQWLLDAQETDGSWFGRWGVNYVYGTGSVIPALTAAGLPTSHPAIRRAVRWLESVQNED
GGWGEDLRSYRYVREWSGRGASTASQTGWALMALLAAGERDSKAVERGVAWLAATQREDGSW
DEPYFTGTGFPWDFSINYNLYRQVFPLTALGRYVHGEPFAKKPRAADAPAEAAPAEVKGS

2.13.2DNA sequence

ATGACAGCGACGACCGACGGAAGCACCGGAGCCTCCCTGCGGCCCCTGGCAGCCTCGGCCAG
CGACACCGACATCACGATCCCCGCCGCGGCGGCCGGGGTACCCGAAGCCGCCGCCCGCGCCA
CCCGGCGTGCCACCGACTTCCTGCTCGCCAAGCAGGACGCCGAGGGCTGGTGGAAGGGCGAC
CTCGAGACGAACGTCACGATGGACGCCGAGGACCTGCTCCTGCGTCAGTTCCTGGGCATCCA
GGACGAGGAGACCACCCGCGCCGCCGCGCTGTTCATCCGCGGCGAGCAGCGCGAGGACGGCA
CCTGGGCCACCTTCTACGGCGGCCCCGGCGAACTGTCCACGACCATCGAGGCCTACGTCGCC
CTCCGCCTGGCCGGCGACTCACCCGAGGCGCCCCACATGGCGCGGGCCGCGGAGTGGATCAG

GTCCCGCGGCGGCATCGCCTCCGCCCGGGTCTTCACCCGGATCTGGCTGGCCCTGTTCGGCT
GGTGGAAGTGGGACGACCTGCCCGAACTCCCGCCGGAGCTGATCTACTTCCCCACCTGGGTC
CCGCTCAACATCTACGACTTCGGCTGCTGGGCCCGGCAGACCATCGTGCCGCTCACCATCGT
CTCCGCGAAGCGGCCGGTGCGTCCCGCGCCGTTCCCGCTGGACGAACTGCACACCGACCCGG
CCCGCCCCAACCCGCCACGCCCCCTGGCACCCGTGGCCAGCTGGGACGGCGCCTTCCAGCGC
ATCGACAAGGCCCTGCACGCCTACCGCAAGGTCGCCCCGCGCCGGCTGCGCCGGGCCGCGAT
GAACAGCGCCGCCCGCTGGATCATCGAGCGGCAGGAGAACGACGGCTGCTGGGGCGGCATCC
AGCCGCCTGCGGTCTACTCGGTCATCGCCCTCTACCTGCTCGGCTACGACCTCGAACACCCC
GTGATGCGCGCGGGACTGGAGTCGCTGGACCGTTTCGCCGTCTGGCGCGAGGACGGCGCCCG
GATGATCGAGGCCTGCCAGTCCCCGGTGTGGGACACCTGCCTGGCCACCATCGCGCTGGCCG
ACGCGGGCGTGCCCGAGGACCACCCGCAGCTGGTGAAGGCCTCGGACTGGATGCTCGGCGAA
CAGATCGTGCGCCCCGGCGACTGGTCGGTGAAGCGCCCCGGACTCCCGCCCGGCGGCTGGGC
GTTCGAGTTCCACAACGACAACTACCCCGACATCGACGACACCGCCGAGGTGGTCCTCGCCC
TGCGCCGGGTCAGGCACCACGACCCGGAACGGGTGGAGAAGGCGATCGGGCGCGGGGTGCGC
TGGAACCTCGGCATGCAGTCGAAGAACGGCGCCTGGGGCGCCTTCGACGTCGACAACACCAG
CGCCTTCCCCAACCGGCTGCCGTTCTGCGACTTCGGCGAGGTCATCGACCCGCCGTCCGCGG
ACGTCACCGCGCACGTCGTCGAGATGCTCGCCGTCGAGGGCCTCGCCCACGACCCGCGCACC
CGCCGCGGCATCCAGTGGCTGCTCGACGCCCAGGAGACGGACGGTTCGTGGTTCGGCCGCTG
GGGCGTCAACTACGTCTACGGCACCGGTTCCGTGATCCCCGCGCTGACCGCGGCCGGACTGC
CCACCTCGCACCCGGCCATCCGCCGGGCGGTGCGCTGGCTGGAGTCCGTCCAGAACGAGGAC
GGCGGCTGGGGCGAGGACCTGCGCTCCTACCGCTACGTCCGGGAGTGGAGCGGCCGGGGCGC
CTCGACCGCCTCGCAGACCGGCTGGGCGCTGATGGCCCTGCTGGCGGCAGGGGAGCGGGACT
CCAAAGCCGTGGAGCGCGGCGTCGCATGGCTCGCGGCCACCCAGCGGGAGGACGGCTCCTGG
GACGAGCCCTACTTCACGGGCACCGGCTTCCCGTGGGACTTCTCCATCAACTACAACCTCTA
CCGCCAGGTCTTCCCGCTCACCGCTCTCGGCCGGTACGTCCACGGGGAGCCCTTCGCCAAGA
AGCCCCGCGCGGCCGACGCCCCCGCCGAAGCCGCCCCGGCCGAGGTGAAGGGCAGCTGA

2.14*Sfu*SHC1

2.14.1Amino acid sequence

MRRLDTFPPEIPTGSRDKPPSGEEHSCSTPAEPLRSRLDEGILRAVDWLVCDQHPDGFWAGM
LQSNSCMEAEWVLAMHFLGIDDDPKYDGVIRAILGEQRADGSWGVFHKAPNGDINTTVECYA
ALRASGLAPESAPLSSAREWILAGGGLANIRNFTKYWLALIGEWPWEGTPTIPPELIFFPPR
MPLNIYHFASWARSTIVPLSILSARRPVRPLPEDRRLDELFPQGRSAFDFRLPRKDGWLSWE
GFFHVCDRILRLYARTRRAPFRETAIRVCLEWIIRRQETDGAWSGIQPPWIYALLALHAEGY
GLDHPILRAGLRAFDSHWSYERDGGIYLQASESPVWDTVLSLRALADCGEERKASVSIASAL
EWLLNRQISVPGDWAVRVPSVPCGGWAFQRANSFYPDVDDTAVAIEVLARLRPFTANQSAVD
RAIRSARDWVLAMQCSNGGWAAFDRDNDFKLVTKIPFCDFGELLDPPSVDVTAHVIEALAAL
GWDMTSREIEAAVSFIRREQEAEGSWFGRWGVNHIYGTATVLPALRAIGEDMSSAYVLRAAD
WLASRQNADGGWGETPASYMDDSLRGVGESTASQTAWAIMGLVAVGSGAHDDTVRRGIDFLL
FAQHGGTWEEPQYTGTGFPGYSVGERIRLRDMGASLKQGTELQRAFMINYNLYRHYFPLMAL
GRARYHLQLRRSAREGGNGETTPNGSAL

2.14.2DNA sequence (*E. coli* codon optimized)

ATGCGTCGCCTGGATACCTTTCCTCCGGAAATTCCGACCGGTAGCCGTGATAAACCTCCGAG
CGGTGAAGAACATAGCTGTAGCACACCGGCAGAACCGCTGCGTAGCCGTCTGGATGAAGGTA
TTCTGCGTGCAGTTGATTGGCTGGTTTGTGATCAGCATCCGGATGGTTTTTGGGCAGGTATG
CTGCAGAGCAATAGCTGTATGGAAGCAGAATGGGTTCTGGCAATGCATTTTCTGGGTATTGA

TGATGATCCGAAATATGATGGTGTGATTCGTGCAATTCTGGGTGAACAGCGTGCAGATGGTA
GCTGGGGTGTTTTTCATAAAGCACCGAATGGCGATATTAATACCACCGTTGAATGCTATGCA
GCACTGCGTGCAAGCGGTCTGGCACCGGAAAGCGCACCGCTGAGCAGCGCACGTGAATGGAT
TCTGGCAGGCGGTGGTCTGGCAAATATTCGTAATTTTACCAAATATTGGCTGGCCCTGATTG
GTGAATGGCCGTGGGAAGGCACCCCGACCATTCCTCCGGAACTGATTTTTTTTCCTCCGCGT
ATGCCGCTGAATATTTATCATTTTGCAAGCTGGGCACGTAGCACCATTGTTCCGCTGAGCAT
TCTGAGCGCACGTCGTCCGGTTCGTCCGCTGCCGGAAGATCGTCGTCTGGATGAACTGTTTC
CGCAGGGTCGTAGCGCATTTGATTTTCGTCTGCCTCGTAAAGATGGTTGGCTGAGCTGGGAA
GGTTTTTTTCATGTTTGTGATCGTATTCTGCGTCTGTATGCACGTACCCGTCGTGCACCGTT
TCGTGAAACCGCAATTCGTGTTTGTCTGGAATGGATTATTCGTCGTCAGGAAACCGACGGAG
CCTGGTCAGGTATTCAGCCTCCGTGGATTTATGCACTGCTGGCACTGCATGCCGAAGGTTAT
GGTCTGGATCATCCGATTCTGCGTGCCGGTCTGCGTGCCTTTGATAGCCATTGGAGCTATGA
ACGTGATGGTGGTATTTATCTGCAGGCAAGCGAAAGTCCGGTTTGGGATACCGTTCTGAGCC
TGCGTGCACTGGCAGATTGTGGTGAAGAACGTAAAGCAAGCGTTAGCATTGCAAGCGCACTG
GAATGGCTGCTGAATCGTCAGATTAGCGTTCCGGGTGATTGGGCAGTTCGTGTTCCGAGCGT
TCCGTGTGGTGGTTGGGCATTTCAGCGTGCAAATAGCTTTTATCCGGATGTTGATGATACCG
CAGTTGCAATTGAAGTTCTGGCACGTCTGCGTCCGTTTACCGCAAATCAGAGCGCAGTTGAT
CGTGCAATTCGTAGCGCACGTGATTGGGTGCTGGCCATGCAGTGTAGCAATGGTGGCTGGGC
AGCATTTGATCGTGATAATGATTTTAAACTGGTGACCAAAATTCCGTTTTGCGATTTTGGTG
AACTGCTGGACCCTCCGAGCGTTGATGTTACCGCACATGTTATTGAAGCACTGGCAGCACTG
GGTTGGGATATGACCAGCCGTGAAATTGAAGCAGCAGTTAGCTTTATTCGTCGTGAACAGGA
AGCAGAAGGTAGCTGGTTTGGTCGTTGGGGTGTTAATCATATTTATGGCACCGCAACCGTTC
TGCCTGCACTGCGTGCCATTGGTGAAGATATGAGCAGCGCCTATGTTCTGCGTGCGGCAGAT
TGGCTGGCAAGCCGTCAGAATGCAGATGGCGGTTGGGGTGAAACACCGGCAAGCTATATGGA
TGATAGCCTGCGTGGTGTTGGTGAAAGCACCGCAAGCCAGACCGCATGGGCAATTATGGGTC
TGGTTGCAGTTGGTAGCGGTGCACATGATGATACCGTTCGTCGTGGTATTGATTTTCTGCTG
TTTGCACAGCATGGTGGCACCTGGGAAGAACCGCAGTATACCGGCACCGGTTTTCCGGGTTA
TAGCGTTGGTGAACGTATTCGTCTGCGTGATATGGGTGCAAGCCTGAAACAGGGCACCGAAC
TGCAGCGTGCATTTATGATTAATTATAATCTGTATCGCCATTATTTTCCGCTGATGGCACTG
GGTCGTGCACGTTATCATCTGCAGCTGCGTCGTAGTGCACGTGAAGGTGGTAATGGTGAAAC
CACCCCGAATGGTAGCGCACTG

2.15*Sfu*SHC2

2.15.1Amino acid sequence

MNPIRGKRGSAADFLEEEYQWENLADHGESGRTPGGGHPAALKEYEAGSATEHTGHHCVHHL
GVRNSWLRKIEKAIDNACGQLFKTQYEDGYWWSELESNVTITSEYIMLLYLLEVSRPEQQKS
MVKYLLNQQRPDGSWGLYYGDGGNLSTTIEAYFALKLAGEHCESEPMRRAREFILSKGGIES
ARVFTKIWLALFSQYDWDKVPSMPVELVLLPSSLYFNIYEFSSWARGTVVPLSIVMSIRPRC
PLPAKCSIKELYVPGSKHKNFASCTHKLFFLFDRIAKAFERRPVPSLRNKAVQAAETWVLDH
QEDSGDWGGIQPPMVYSVLALYYLGYPLDHEVIVKGIKALDAFCMEDEEGTRMQSCVSPVWD
TALTVLSMLDAGVAAEHPGLEKAGRWLLENQVLTGGDWQIKNDSLPGGWAFEFYNTRYPDVD
DSAVVLSTLNRFNAERVEGLEFAKCRGMEWCLSMQSSNGGWAAFDKDNTLEILNRIPFADQE
AMVDYPTADVTGRVLEAMGYLGYDGSHPRARKAIQFLKKRQERDGCWWGRWGVNYIYGTWSV
LKGLISIGEDPRAAYIRAAVRWVKDHQNSDGGWGETCESYENPELRGQGPSTPSQTAWALMS
LIACGEMKSQEASRGIQYLLRTQKRDGTWEELHFTGTGFPKHFYIRYHNYRNCFPLMALGQY
LRALER

2.15.2DNA sequence

ATGAATCCAATCAGGGGCAAGAGAGGAAGCGCGGCGGATTTCCTCGAAGAAGAGTATCAGTG
GGAGAATCTTGCTGACCATGGCGAATCGGGGCGCACTCCCGGAGGCGGTCATCCGGCCGCGT
TGAAGGAGTACGAGGCCGGGAGCGCAACGGAGCACACCGGTCATCACTGCGTTCATCATCTG
GGGGTGCGGAATTCATGGTTGCGAAAGATCGAGAAGGCCATCGACAATGCGTGCGGTCAGCT
TTTCAAGACTCAATATGAAGACGGGTACTGGTGGTCGGAACTGGAATCGAACGTCACGATCA
CCAGCGAGTACATCATGCTGCTCTACCTTCTGGAGGTGAGCAGGCCCGAGCAGCAGAAAAGC
ATGGTGAAATACCTGCTCAATCAACAGCGGCCCGACGGTTCGTGGGGATTGTACTACGGAGA
CGGCGGGAATTTGAGCACCACGATCGAGGCCTATTTCGCGCTCAAGCTTGCGGGTGAGCACT
GCGAGTCGGAGCCGATGAGGAGGGCCCGCGAATTTATTCTGTCCAAGGGCGGCATCGAGTCG
GCGCGGGTATTCACGAAGATCTGGCTGGCGCTTTTTTCTCAGTACGACTGGGACAAGGTGCC
GTCCATGCCCGTCGAGCTGGTGTTGCTCCCAAGCAGTCTGTATTTCAATATTTATGAGTTTT
CGAGCTGGGCCAGAGGCACGGTGGTTCCGTTGTCCATTGTGATGTCCATCCGGCCGCGTTGT
CCGTTGCCCGCAAAGTGTTCGATCAAGGAGCTCTACGTCCCGGGCAGCAAGCACAAGAATTT
CGCATCGTGCACGCACAAGCTGTTTTTTCTTTTCGACCGTATTGCGAAGGCGTTTGAGCGGC
GCCCGGTTCCTTCCTTGCGGAACAAGGCGGTGCAGGCGGCGGAGACCTGGGTTTTGGATCAC
CAGGAGGACAGCGGAGATTGGGGCGGGATACAGCCGCCGATGGTCTACTCTGTCCTGGCGCT
GTACTACCTGGGGTACCCGCTGGATCACGAGGTCATCGTCAAGGGAATAAAGGCGCTTGACG
CCTTCTGCATGGAAGACGAGGAGGGAACGCGGATGCAGTCCTGTGTTTCTCCCGTCTGGGAC
ACGGCCCTCACCGTTCTGTCCATGCTCGACGCAGGCGTCGCTGCGGAACACCCCGGCCTGGA
AAAAGCGGGAAGGTGGCTTTTGGAGAACCAGGTTCTGACGGGTGGAGACTGGCAGATCAAGA
ACGATAGTCTTCCGGGCGGATGGGCGTTCGAATTCTACAACACCCGCTATCCCGACGTGGAT
GATTCCGCGGTTGTGCTGAGCACTCTGAACCGCTTCAATGCGGAGCGGGTCGAAGGGCTGGA
ATTCGCCAAGTGCAGGGGCATGGAATGGTGCCTCAGCATGCAGAGCTCCAATGGAGGATGGG
CCGCCTTCGACAAGGACAATACTCTCGAGATCCTCAATCGCATTCCTTTTGCCGACCAGGAA
GCGATGGTTGATTACCCCACCGCCGATGTTACAGGCCGGGTGCTCGAAGCCATGGGATATCT
CGGATACGACGGTTCACACCCGCGGGCGCGAAAGGCAATCCAATTCCTGAAGAAGCGCCAGG
AACGCGACGGTTGCTGGTGGGGACGCTGGGGCGTCAACTACATCTACGGCACTTGGTCCGTC
CTCAAGGGGCTGATATCCATCGGCGAGGACCCCAGGGCGGCTTACATCAGGGCGGCTGTGCG
CTGGGTGAAGGATCACCAGAATTCGGACGGCGGGTGGGGAGAGACCTGTGAGAGTTACGAGA
ACCCTGAACTGCGCGGTCAGGGGCCGAGCACTCCCTCCCAGACCGCCTGGGCCCTGATGTCG
CTGATCGCCTGCGGGGAAATGAAATCTCAGGAAGCCAGCCGCGGGATTCAGTATCTGCTCAG
AACGCAGAAACGGGACGGCACTTGGGAGGAACTCCATTTTACGGGAACGGGCTTTCCCAAGC
ACTTCTACATCCGCTACCACAATTATCGGAATTGTTTCCCCCTGATGGCTCTGGGGCAGTAT
CTGCGGGCTCTTGAGCGGTAA

2.16*Ssc*SHC

2.16.1Amino acid sequence

MTATTDGSTGALPPRAPSASDTDHGTPVAAGVQEAALHAVGRATDFLLSRQDAQGWWKGDLE
TNVTMDAEDLLLRQFLGIRDDATTRAAALFIRGEQRPDGTWATFYGGPPDLSATVEAYVALR
LAGDDPAAPHMAKASAWIRARGGIAAARVFTRIWLALFGWWKWDDLPEMPPEIVYFPTWMPL
NIYDFGCWARQTIVPLTVVSAKRPVRPAPFPLDELHTDPGRPNPPRPLDRLGSWEGAFQRLD
RALHGYHKVALKRLRRAAMNRAARWIVERQENDGCWGGIQPPAVYSVIALHLLGYDLGHPVM
RAGLESLDRFAVWREDGARMIEACQSPVWDTCLATIALADAGLPPDHPQLVKAADWMLGEEI
VRPGDWSVKRPQLPPGGWAFEFHNDNYPDIDDTAEVVLALRRVRHPDPERVERAVRRGVRWT
LGMQSGNGAWAAFDADNTSPFPNRLPFCDFGEVIDPPSADVTAHVVEMLAAEGLSHDPRTRR
GIEWLLAEQEPGGAWFGRWGVNYVYGTGSVVPALVTAGLPAAHPAIRRAVAWLETVQNDDGG

WGEDLRSYPDPAEWGGKGASTASQTAWALLALLAAGERDGKATERGVAWLARTQREDGSWDE
PYFTGTGFPWDFSINYHLYRQVFPLTALGRYVHGEPAVLKPGTR

2.16.2DNA sequence (*E. coli* codon optimized)

ATGACCGCAACTACCGACGGCTCGACCGGCGCACTGCCACCACGTGCCCCGAGCGCGAGCGA
TACCGACCACGGTACGCCGGTTGCAGCAGGTGTGCAGGAAGCAGCATTGCACGCGGTTGGTC
GTGCAACCGACTTCCTGCTGAGCCGCCAAGACGCCCAAGGCTGGTGGAAAGGTGACCTGGAA
ACCAACGTTACTATGGACGCAGAGGACCTGCTGCTGCGCCAGTTCCTGGGCATCCGTGATGA
CGCTACGACCCGTGCGGCTGCATTGTTTATCCGTGGTGAACAGCGTCCGGATGGCACGTGGG
CAACCTTTTATGGCGGTCCGCCGGATCTGAGCGCAACCGTGGAGGCCTATGTCGCGCTGCGT
TTGGCGGGTGACGATCCGGCTGCTCCTCACATGGCGAAAGCTAGCGCTTGGATTCGCGCACG
CGGTGGTATCGCTGCCGCTCGCGTGTTTACCCGTATCTGGTTGGCGCTGTTCGGTTGGTGGA
AATGGGACGACCTGCCGGAGATGCCTCCGGAAATTGTTTACTTTCCGACTTGGATGCCGCTG
AATATCTATGACTTTGGCTGTTGGGCACGCCAGACTATTGTCCCGCTGACGGTTGTGTCTGC
GAAACGCCCGGTCCGTCCTGCCCCGTTCCCGCTGGACGAACTGCACACCGATCCGGGTCGTC
CGAACCCACCGCGTCCGCTGGATCGTTTGGGTTCGTGGGAAGGTGCGTTCCAACGTCTGGAT
CGCGCACTGCATGGTTACCATAAGGTCGCTCTGAAACGTCTGCGTCGTGCGGCCATGAATCG
TGCGGCACGTTGGATTGTTGAGCGCCAGGAGAACGACGGCTGCTGGGGTGGCATCCAGCCGC
CAGCGGTTTACAGCGTGATTGCACTGCACCTGCTGGGTTATGATCTGGGCCACCCGGTTATG
CGTGCGGGTTTGGAGTCCCTGGACCGCTTCGCCGTGTGGCGTGAAGATGGTGCGCGTATGAT
TGAGGCGTGCCAGAGCCCGGTGTGGGACACCTGTCTGGCCACGATCGCTCTGGCGGACGCCG
GTTTGCCGCCAGACCACCCGCAACTGGTGAAGGCTGCGGATTGGATGCTGGGTGAAGAGATC
GTCCGTCCGGGCGATTGGAGCGTGAAACGTCCGCAGCTGCCACCGGGTGGTTGGGCCTTCGA
ATTCCACAACGATAACTACCCGGATATTGATGATACCGCGGAAGTCGTGCTGGCCTTGCGCC
GTGTTCGTCATCCAGACCCGGAACGTGTCGAGCGTGCGGTGCGTCGTGGTGTGCGCTGGACC
CTGGGTATGCAAAGCGGCAACGGTGCTTGGGCAGCGTTCGACGCTGACAATACCTCTCCGTT
TCCGAACCGTCTGCCGTTCTGCGATTTTGGCGAGGTCATTGATCCGCCATCCGCAGATGTCA
CCGCCCATGTCGTTGAGATGTTGGCGGCAGAAGGCCTGAGCCACGATCCGCGTACTCGTCGC
GGCATTGAGTGGCTGCTGGCGGAGCAAGAACCTGGTGGTGCGTGGTTCGGCCGCTGGGGTGT
GAATTATGTGTACGGTACCGGTTCCGTTGTTCCGGCACTGGTGACCGCTGGCCTGCCAGCAG
CGCATCCGGCCATTCGTCGCGCCGTGGCTTGGTTGGAAACGGTCCAAAATGACGACGGCGGT
TGGGGCGAGGATCTGCGCAGCTACCCTGACCCGGCAGAATGGGGTGGTAAGGGTGCCAGCAC
CGCGAGCCAGACGGCTTGGGCACTGCTGGCCCTGCTGGCCGCAGGCGAGCGCGATGGCAAGG
CAACGGAGCGCGGTGTCGCGTGGCTGGCGCGTACGCAACGCGAGGACGGCAGCTGGGACGAG
CCGTACTTTACCGGTACGGGTTTTCCTTGGGATTTCAGCATCAATTATCACCTGTACCGCCA
GGTCTTTCCGCTGACCGCGCTGGGCCGTTACGTTCATGGTGAGCCGGCAGTTTTGAAGCCGG
GCACGCGTTAA

2.17*Ssv*SHC

2.17.1Amino acid sequence

MHEGEAMTATTDGSTGALPPRAAAASETHLDTPVAAGIQEAAVRAVQRATEHLLARQDAEGW
WKGDLETNVTMDAEDLLLRQFLGIRDESTTRAAAKFIRGEQREDGTWAGFYGGPGELSTTVE
AYVALRLDGDAPDAPHMAKASAWIRAQGGIAAARVFTRIWLALFGWWKWEDLPELPPELIYF
PKWAPLNIYDFGCWARQTIVPLTIVSAKRPVRPAPFPLDELHADPADPNPAKPLAPVASWDG
AFQRLDKAMHQLRKVAPRRLRRAAMNSAARWIIERQENDGCWGGIQPPAVYSVIALHLLGYD
LQHPVMRAGLESLDRFAIWREDGSRMIEACQSPVWDTCLATIALVDAGVPADHPQLVKAADW
MLGEEIVRPGDWSVKRPQLPPGGWAFEFHNDNYPDIDDTAEVVLALRRVRHHDPDRVENAIG

RGVRWNLGMQSKNGAWGAFDVDNTSPFPNRLPFCDFGEVIDPPSADVTAHVVEMLAVEGLSH
DPRTRRGIEWLLAEQEPDGSWFGRWGVNYIYGTGSVVPALTAAGLPASHPAIRRAVAWLEKV
QNDDGGWGEDLRSYKYVKEWSGRGASTASQTAWALMALLAAGERDSKAVERGVEWLASTQRA
DGSWDEPYFTGTGFPWDFSINYHLYRQVFPLTALGRYVHGEPFSRTEAL

2.17.2DNA sequence (*E. coli* codon optimized)

ATGCACGAGGGCGAAGCGATGACGGCAACCACGGACGGCAGCACCGGCGCTTTGCCGCCTCG
CGCAGCAGCGGCGAGCGAAACCCACCTGGACACCCCGGTCGCAGCAGGTATTCAGGAAGCAG
CGGTTCGTGCGGTTCAGCGCGCGACCGAACACTTGCTGGCACGTCAGGACGCGGAAGGTTGG
TGGAAGGGCGACCTGGAAACGAATGTGACGATGGATGCGGAAGATCTGCTGCTGCGTCAGTT
CCTGGGTATTCGCGACGAGTCCACCACGCGTGCGGCAGCGAAGTTCATCCGTGGTGAACAGC
GTGAAGATGGTACGTGGGCTGGCTTCTACGGCGGTCCGGGTGAACTGTCCACCACTGTGGAG
GCGTACGTTGCCCTGCGCCTGGACGGTGATGCCCCGGACGCACCGCACATGGCCAAGGCAAG
CGCGTGGATTCGCGCTCAAGGTGGCATCGCGGCAGCACGCGTCTTTACCCGCATTTGGCTGG
CGCTGTTTGGCTGGTGGAAATGGGAGGATCTGCCGGAACTGCCACCGGAGCTGATCTACTTC
CCAAAATGGGCACCGTTGAACATCTATGATTTCGGTTGTTGGGCTCGCCAAACGATCGTGCC
GCTGACCATCGTCAGCGCAAAACGCCCGGTTCGCCCAGCGCCGTTTCCGTTGGACGAGCTGC
ACGCGGACCCAGCAGACCCGAATCCGGCAAAACCGCTGGCACCGGTGGCAAGCTGGGACGGT
GCCTTCCAGCGTCTGGACAAGGCAATGCACCAACTGCGCAAAGTAGCTCCGCGTCGTCTGCG
TCGTGCCGCCATGAATTCCGCTGCGCGTTGGATTATCGAACGCCAAGAGAATGACGGCTGCT
GGGGCGGTATCCAGCCTCCGGCCGTTTACTCCGTTATTGCGCTGCATCTGCTGGGCTATGAC
TTGCAACATCCGGTGATGCGTGCGGGCTTGGAGAGCCTGGATCGTTTCGCGATTTGGCGCGA
AGATGGCAGCCGTATGATTGAGGCGTGTCAGTCTCCGGTCTGGGATACGTGCCTGGCCACGA
TCGCGTTGGTTGACGCGGGTGTGCCGGCAGACCATCCGCAGCTGGTCAAAGCGGCAGACTGG
ATGCTGGGCGAGGAGATCGTCCGTCCGGGCGATTGGTCTGTCAAACGTCCGCAACTGCCGCC
TGGTGGCTGGGCGTTTGAGTTTCATAATGATAACTACCCGGATATCGATGACACCGCTGAGG
TTGTTTTGGCCCTGCGTCGCGTCCGCCACCATGATCCGGACCGTGTCGAGAATGCAATTGGT
CGCGGCGTTCGCTGGAACCTGGGCATGCAATCGAAGAATGGTGCATGGGGTGCGTTCGACGT
AGACAACACGAGCCCGTTCCCGAACCGTCTGCCGTTTTGCGATTTCGGTGAGGTGATTGATC
CGCCGAGCGCGGACGTCACCGCACACGTCGTCGAGATGCTGGCGGTGGAAGGTCTGAGCCAT
GACCCACGCACCCGTCGCGGCATTGAGTGGCTGCTGGCGGAGCAAGAGCCGGACGGTAGCTG
GTTTGGTCGTTGGGGTGTTAACTATATCTATGGTACGGGTTCCGTGGTGCCGGCTCTGACTG
CCGCTGGCCTGCCGGCGAGCCACCCAGCCATCCGCCGTGCAGTGGCATGGCTGGAGAAGGTT
CAGAACGATGACGGTGGTTGGGGTGAGGACCTGCGTAGCTACAAATACGTTAAGGAGTGGAG
CGGCCGTGGCGCTAGCACCGCCTCTCAAACCGCTTGGGCGCTGATGGCGCTGCTGGCAGCGG
GCGAGCGTGATAGCAAGGCGGTCGAGCGTGGTGTGGAATGGCTGGCGAGCACCCAACGCGCC
GATGGCAGCTGGGATGAGCCATACTTTACCGGCACCGGTTTTCCGTGGGATTTCAGCATTAA
CTACCACTTGTATCGTCAGGTGTTCCCGCTGACGGCGTTGGGCCGTTATGTTCACGGTGAAC
CGTTTAGCCGTACTGAAGCACTGTAA

2.18*Sth*SHC

2.18.1Amino acid sequence

MDPALSRAVDWLLEHQDPAGWWCGEFETNVTITAEHILLLRFLGLDPSPLRDAVTRYLLGQQ
REDGSWALYYEGPADLSTSIEAYAALKVLGLDPTSEPMRRALQVIHDLGGVAQARVFTRIWL
AMFGQYPWDGVPSMPPELIWLPPSAPFNLYDFACWARATITPLLIILARRPVRPLGCDLGEL
VLPGSEHLLTRVPGSGPFWWGDKVLKRYDHLVRHPGRDRACQRIVEWIIARQEADGSWGGIQ

SAWVMSLIALHLEGLPLDHPVMRAGLAGFDRVALEDERGWRLQASTSPVWDTAWAVLALRRA
GLPREHPRLALAVDWLLQEQIPGGGDWQVRTGTIPGGGWAFEFDNDHYPDIDDTAVVVLALL
EAGHEDRVRNAVERAARWILAMRSTDGGWGAFDRDNAREVIHRLPIADFGTLIDPPSEDVTA
HVLEMLARLSFPSTDPVVARGLEFLQQTQRPDGAWFGRWGVNYIYGTWCAVSALTAFADTDA
TARAMVPRAVAWLLDRQNADGGWGETCGSYEDPNLAGVGRSTPSQTAWAVLALQAAGLGQHP
ACRRGLDFLRERQVGGTWEEREHTGTGFPGDFFINYHLYRHVFPTMALAGAATGMDSPR

2.18.2DNA sequence (*E. coli* codon optimized)

ATGGACCCTGCACTCTCACGCGCCGTCGATTGGCTGCTGGAACACCAGGACCCAGCGGGTTG
GTGGTGCGGTGAATTTGAAACGAACGTGACCATTACCGCAGAACATATTCTGCTGCTGCGTT
TTCTGGGCCTGGACCCAAGCCCGCTGCGTGATGCGGTGACCCGCTACCTGCTGGGCCAGCAG
CGCGAGGATGGTAGCTGGGCACTGTACTACGAGGGTCCGGCTGACCTGTCTACTAGCATCGA
AGCGTACGCCGCGTTGAAAGTGCTGGGTCTGGACCCGACCAGCGAGCCGATGCGTCGTGCGT
TGCAAGTTATTCATGACCTGGGTGGTGTTGCCCAGGCCCGTGTGTTCACCCGCATTTGGCTG
GCAATGTTTGGCCAATATCCGTGGGACGGTGTACCGAGCATGCCGCCTGAACTGATCTGGTT
GCCGCCGAGCGCACCGTTCAACCTGTACGACTTCGCGTGTTGGGCACGTGCGACCATTACTC
CGCTGCTGATTATCCTGGCGCGTCGTCCAGTCCGTCCGCTGGGCTGTGACTTGGGTGAGCTG
GTTCTGCCGGGTAGCGAGCACCTGTTGACCCGTGTGCCGGGTTCCGGTCCTTTCTGGTGGGG
CGATAAGGTTCTGAAACGTTATGATCATCTGGTCCGTCACCCGGGTCGTGACCGTGCATGTC
AGCGCATTGTGGAGTGGATCATCGCGCGTCAAGAAGCCGACGGTTCTTGGGGTGGCATCCAA
AGCGCTTGGGTTATGAGCCTGATTGCCCTGCACCTGGAAGGTCTGCCGTTGGATCACCCGGT
CATGCGTGCCGGCCTGGCCGGTTTCGACCGTGTCGCGCTGGAGGACGAGCGCGGCTGGCGCC
TGCAAGCTAGCACGAGCCCTGTGTGGGATACCGCGTGGGCAGTCCTGGCGCTGCGCCGTGCG
GGTCTGCCGCGTGAGCACCCGCGCTTGGCCCTGGCGGTCGACTGGCTGTTGCAAGAACAGAT
CCCGGGTGGTGGCGACTGGCAGGTCCGTACGGGTACCATTCCTGGTGGCGGCTGGGCATTCG
AGTTCGATAATGACCATTATCCGGACATCGACGATACCGCGGTTGTTGTTCTGGCGTTGCTG
GAGGCGGGTCATGAGGATCGCGTCCGCAACGCGGTTGAGCGTGCAGCTCGCTGGATTCTGGC
AATGCGCTCGACCGATGGCGGTTGGGGTGCCTTCGACCGCGATAATGCCCGCGAGGTCATCC
ACCGCCTGCCGATCGCCGATTTTGGTACGCTGATCGATCCGCCGTCCGAGGATGTGACCGCG
CACGTCCTGGAGATGCTGGCTCGTCTGTCTTTCCCGAGCACGGACCCGGTTGTGGCGCGTGG
TCTGGAATTCTTGCAGCAAACGCAGCGTCCGGATGGTGCCTGGTTCGGCCGTTGGGGCGTGA
ACTACATTTATGGCACCTGGTGCGCGGTGTCCGCGCTGACGGCATTTGCCGATACGGACGCT
ACGGCTCGCGCGATGGTGCCGCGTGCGGTTGCGTGGCTGTTGGACCGTCAGAATGCAGATGG
TGGCTGGGGTGAAACCTGCGGCAGCTACGAGGACCCGAACCTGGCGGGTGTTGGCCGCAGCA
CCCCGAGCCAAACTGCATGGGCGGTTCTGGCTCTGCAAGCCGCAGGCCTGGGTCAGCACCCG
GCGTGCCGTCGTGGCCTGGATTTTCTGCGCGAGCGTCAAGTGGGCGGTACGTGGGAAGAACG
TGAACATACCGGCACCGGTTTTCCGGGCGATTTCTTTATCAATTATCATTTGTATCGTCACG
TGTTTCCGACCATGGCGTTGGCAGGTGCGGCAACGGGCATGGACAGCCCGCGTTAA

2.19*Svi*SHC

2.19.1Amino acid sequence

MTDVLTRELSPNSTRDRVRSCVSSARQYLLSLQHEEGWWKGELDTNVTMEAEDLLLRQFLGI
SDEQVTQETARWIRSCQREDGTWATFHGGPPDLSTTVEAYVALRLAGDAMDAAHLRKAREYI
LDSGGIESTRVFTRIWLALFGEWPWSRLPVLPPEMMLLPDWFPLNIYDWASWARQTVVPLTI
VGSLRPTRDLGFSVRELRTGIQRRDLESPLSWAGVFHGLDSVLHRLEKLPLKPLRKVALARA
EQWILDRQESDGGWGGIQPPWVYSILALHLRGYPLDHPVLRKALDGLDGFTIRHRTENGWIR
KLEACQSPVWDTALAMTALLDSGTPPNDPALVRAADWILRQEIRVSGDWRVRRPALEPSGWA

FEFANDHYPDTDDTAEVVLGLQRVRHPEPHRVNAAVERATAWLVGMQSSDGGWGAFDADNTR
TLCEKLPFCDFGAVIDPPSADVTAHIVEMLAARGMADSESARRGVRWLLEHQEVDGSWFGRW
GANHVYGTGAVVPALVACGISPQHEAVRAAVQWLVAHQNADGGWGEDLRSYVDRTWVGRGTS
TPSQTAWALLALLAAGERGEVVRRGVEWLMAAQRPDGGWDEPQYTGTGFPGDFYISYHMYRI
VFPLTALGRYLGRGGDVGTG

2.19.2DNA sequence (*E. coli* codon optimized)

ATGACTGACGTACTGACCCGCGAACTGAGCCCGAACAGCACCCGTGACCGCGTTCGTAGCTG
TGTGAGCTCCGCGCGTCAGTACCTGCTGTCTCTGCAACATGAAGAGGGTTGGTGGAAAGGTG
AGCTGGATACCAATGTGACCATGGAGGCGGAGGATCTGCTGCTGCGTCAGTTCCTGGGCATC
AGCGACGAGCAGGTTACCCAGGAAACGGCTCGTTGGATTCGTTCGTGCCAACGTGAGGATGG
CACCTGGGCAACGTTTCACGGTGGTCCGCCGGACCTGAGCACGACCGTCGAAGCCTATGTTG
CGTTGCGTCTGGCCGGTGATGCAATGGACGCGGCACATCTGCGTAAAGCGCGTGAATACATT
CTGGACAGCGGTGGTATCGAAAGCACGCGTGTTTTCACCCGCATTTGGCTGGCGTTGTTCGG
TGAATGGCCGTGGAGCCGTCTGCCGGTCCTGCCACCGGAGATGATGCTGTTGCCGGATTGGT
TTCCGTTGAACATCTACGACTGGGCGTCCTGGGCGCGTCAAACGGTCGTTCCTCTGACGATT
GTAGGCAGCCTGCGTCCGACCCGTGATCTGGGCTTCTCCGTGCGTGAGCTGCGTACGGGTAT
TCAGCGCCGTGACCTGGAAAGCCCTCTGAGCTGGGCCGGTGTCTTTCACGGTCTGGACTCCG
TGCTGCATCGTCTGGAGAAACTGCCGCTGAAGCCGCTGCGTAAGGTTGCCCTGGCCCGTGCA
GAGCAATGGATTCTGGACCGCCAGGAGTCTGACGGTGGTTGGGGTGGTATCCAACCGCCGTG
GGTTTACTCTATCTTGGCACTGCACCTGCGTGGTTATCCGTTGGACCATCCGGTTCTGCGCA
AAGCGCTGGATGGCCTGGACGGTTTCACCATCCGTCATCGTACTGAAAACGGCTGGATTCGC
AAGCTGGAAGCGTGCCAGAGCCCGGTGTGGGACACCGCGCTGGCGATGACCGCACTGTTGGA
TAGCGGTACCCCGCCAAACGACCCGGCGTTGGTGCGCGCTGCGGATTGGATCTTGCGTCAAG
AAATCCGTGTCAGCGGCGATTGGCGCGTCCGTCGCCCTGCGCTGGAACCGAGCGGTTGGGCC
TTCGAGTTTGCCAATGACCACTACCCGGATACCGATGACACGGCAGAAGTTGTGCTGGGTCT
GCAACGCGTTCGTCACCCGGAGCCGCACCGTGTGAACGCTGCCGTCGAACGTGCAACGGCGT
GGCTGGTCGGCATGCAGTCGAGCGATGGTGGCTGGGGCGCTTTTGATGCAGACAATACCCGC
ACCCTGTGTGAAAAGCTGCCGTTCTGTGACTTCGGCGCAGTTATTGACCCGCCGAGCGCGGA
TGTCACCGCGCACATTGTTGAGATGCTGGCAGCCCGTGGCATGGCGGACTCTGAGAGCGCGC
GTCGCGGCGTCCGTTGGCTGCTGGAGCACCAAGAGGTCGATGGTAGCTGGTTTGGCCGTTGG
GGTGCGAATCACGTCTACGGTACTGGTGCGGTTGTGCCAGCGTTGGTCGCGTGCGGTATTTC
GCCGCAGCATGAGGCTGTGCGTGCAGCTGTGCAGTGGCTGGTGGCACACCAAAATGCGGACG
GCGGCTGGGGTGAGGACCTGCGCAGCTACGTTGATCGCACCTGGGTGGGTCGCGGTACGAGC
ACGCCGAGCCAGACCGCATGGGCCCTGCTGGCGCTGCTGGCGGCTGGCGAACGCGGCGAGGT
GGTTCGTCGCGGTGTGGAGTGGCTGATGGCCGCACAGCGCCCAGATGGCGGTTGGGACGAGC
CGCAATATACGGGTACCGGCTTTCCGGGCGATTTCTATATCAGCTACCACATGTATCGCATC
GTTTTTCCGCTGACGGCACTGGGTCGTTATCTGGGCCGTGGTGGTGATGTGGGTACCGGTTA
A

2.20*Tel*SHC

2.20.1Amino acid sequence

MPTSLATAIDPKQLQQAIRASQDFLFSQQYAEGYWWAELESNVTMTAEVILLHKIWGTEQRL
PLAKAEQYLRNHQRDHGGWELFYGDGGDLSTSVEAYMGLRLLGVPETDPALVKARQFILARG
GISKTRIFTKLHLALIGCYDWRGIPSLPPWIMLLPEGSPFTIYEMSSWARSSTVPLLIVMDR
KPVYGMDPPITLDELYSEGRANVVWELPRQGDWRDVFIGLDRVFKLFETLNIHPLREQGLKA
AEEWVLERQEASGDWGGIIPAMLNSLLALRALDYAVDDPIVQRGMAAVDRFAIETETEYRVQ

PCVSPVWDTALVMRAMVDSGVAPDHPALVKAGEWLLSKQILDYGDWHIKNKKGRPGGWAFEF
ENRFYPDVDDTAVVVMALHAVTLPNENLKRRAIERAVAWIASMQCRPGGWAAFDVDNDQDWL
NGIPYGDLKAMIDPNTADVTARVLEMVGRCQLAFDRVALDRALAYLRNEQEPEGCWFGRWGV
NYLYGTSGVLTALSLVAPRYDRWRIRRAAEWLMQCQNADGGWGETCWSYHDPSLKGKGDSTA
SQTAWAIIGLLAAGDATGDYATEAIERGIAYLLETQRPDGTWHEDYFTGTGFPCHFYLKYHY
YQQHFPLTALGRYARWRNLLAT

2.20.2DNA sequence (*E. coli* codon optimized)

ATGCCGACCAGCCTGGCAACCGCAATTGATCCGAAACAGCTGCAGCAGGCAATTCGTGCAAG
CCAGGATTTTCTGTTTAGCCAGCAGTATGCCGAAGGTTATTGGTGGGCAGAACTGGAAAGCA
ATGTTACCATGACCGCAGAAGTTATTCTGCTGCATAAAATTTGGGGCACCGAACAGCGTCTG
CCGCTGGCAAAAGCAGAACAGTATCTGCGTAATCATCAGCGTGATCATGGTGGTTGGGAACT
GTTTTATGGTGATGGTGGTGATCTGAGCACCAGCGTTGAAGCATATATGGGTCTGCGTCTGC
TGGGTGTTCCGGAAACCGATCCGGCACTGGTTAAAGCACGTCAGTTTATTCTGGCACGTGGT
GGTATTAGCAAAACCCGTATTTTTACCAAACTGCATCTGGCACTGATTGGTTGTTATGATTG
GCGTGGTATTCCGAGCCTGCCTCCGTGGATTATGCTGCTGCCGGAAGGTAGCCCGTTTACCA
TTTATGAAATGAGCAGCTGGGCACGTAGCAGCACCGTTCCGCTGCTGATTGTTATGGATCGT
AAACCGGTTTATGGTATGGACCCTCCGATTACCCTGGATGAACTGTATAGCGAAGGTCGTGC
AAATGTTGTTTGGGAACTGCCTCGTCAGGGTGATTGGCGTGATGTTTTTATTGGTCTGGATC
GTGTGTTTAAACTGTTTGAAACCCTGAATATTCATCCGCTGCGTGAACAGGGTCTGAAAGCA
GCAGAAGAATGGGTTCTGGAACGTCAAGAAGCATCAGGCGATTGGGGTGGTATTATTCCGGC
AATGCTGAATAGCCTGCTGGCACTGCGTGCACTGGATTATGCAGTTGATGATCCGATTGTTC
AGCGTGGTATGGCAGCAGTTGATCGTTTTGCAATTGAAACCGAAACCGAATATCGTGTTCAG
CCGTGTGTTAGTCCGGTTTGGGATACCGCACTGGTTATGCGTGCAATGGTTGATAGCGGTGT
TGCACCGGATCATCCGGCTCTGGTGAAAGCCGGTGAATGGCTGCTGAGCAAACAAATTCTGG
ATTATGGCGATTGGCATATTAAAAATAAAAAAGGTCGTCCGGGTGGTTGGGCATTTGAATTT
GAAAATCGCTTTTATCCGGATGTGGATGATACCGCAGTTGTTGTTATGGCCCTGCATGCAGT
TACCCTGCCGAATGAAAATCTGAAACGTCGTGCAATTGAACGTGCAGTTGCATGGATTGCAA
GCATGCAGTGCCGTCCTGGTGGCTGGGCAGCATTTGATGTTGATAATGATCAGGATTGGCTG
AATGGTATTCCGTATGGTGATCTGAAAGCAATGATTGATCCGAATACCGCAGATGTTACCGC
ACGTGTTCTGGAAATGGTTGGTCGTTGTCAGCTGGCATTTGATCGTGTTGCACTGGATCGTG
CACTGGCATATCTGCGCAATGAACAAGAACCGGAAGGTTGTTGGTTTGGTCGTTGGGGTGTT
AATTATCTGTATGGCACCAGCGGTGTTCTGACCGCACTGAGCCTGGTTGCACCGCGTTATGA
TCGTTGGCGTATTCGTCGTGCAGCAGAATGGCTGATGCAGTGTCAGAATGCAGACGGTGGCT
GGGGTGAAACCTGTTGGAGCTATCATGATCCGAGCCTGAAAGGTAAAGGTGATAGCACCGCA
AGCCAGACCGCATGGGCAATTATTGGTCTGCTGGCAGCCGGTGATGCAACCGGTGATTATGC
AACCGAAGCCATTGAACGTGGTATTGCATATCTGCTGGAAACCCAGCGTCCGGATGGCACCT
GGCATGAAGATTATTTTACCGGCACCGGTTTTCCGTGCCATTTTTATCTGAAATACCACTAT
TATCAGCAGCATTTTCCGCTGACCGCTCTGGGTCGTTATGCCCGTTGGCGTAATCTGCTGGC
AACCTAA

2.21*Ttu*SHC

2.21.1Amino acid sequence

MEIQDEVDLLEPQESLTASADSAVDRALFWLLDAQYEDGYWAGILESNACMEAEWLLCFHVL
GIANHPMSRGLVQGLLQRQRADGSWDVYYGARAGDINTTVEVYAALRCQGYAADHPDIKRAR
DWIQLQGGVKQVRVFTRFWLALIGEWPWEETPNLPPEILFFPRWFPFNIYHFAAWARATLVP
LCILSARRMVVPLNKKSCLQELFPEDRSAVVALGKKAGAWSTFFYHADRALKKYQRTFKRPP

GRQQAIKMCLEWILRRQDADGAWGGIQPPWIYSLMALKAEGYPVTHPVMAKGLAALDAHWSY
ERPGGARFVQACESPVWDTLLSSFALLDCGFSCTSSSELRKAVDWILDQQVLLPGDWQQKLP
TVSPGGWAFERANVHYPDVDDTAVALIVLAKVRPDYPDTARVNLAIERGLNWLFAMQCRNGG
WGAFDKDNDKDLLTKIPFSDFGETIDPASVDVTAHVLEALGLLGYRTTHPAVAKALEFIRSE
QENDGCWFGRWGVNYIYGTAAVLPALASLNMNMNQEFIRRAANWILGKQNNDGGWGESCASY
MDDTQRGRGPSTASQTAWAMMSLLAVDGGTYAESLLRAEAYLKTTQTPEGTWDEPYYTGTGF
PGYGIGRREIKRQRSLQQHAELSRGFMINYNLYRHYFPLMALGRLAALRGA

2.21.2DNA sequence (*E. coli* codon optimized)

ATGGAAATTCAGGATGAAGTGGATCTGCTGGAACCGCAGGAAAGCCTGACCGCAAGCGCAGA
TAGCGCAGTTGATCGTGCACTGTTTTGGCTGCTGGATGCACAGTATGAAGATGGTTATTGGG
CAGGTATTCTGGAAAGCAATGCATGTATGGAAGCAGAATGGCTGCTGTGTTTTCATGTTCTG
GGTATTGCAAATCATCCGATGAGCCGTGGTCTGGTTCAGGGTCTGCTGCAGCGTCAGCGTGC
AGATGGTAGCTGGGATGTTTATTATGGTGCACGTGCCGGTGATATTAACACCACCGTTGAAG
TTTATGCAGCACTGCGTTGTCAGGGTTATGCAGCAGATCATCCGGATATTAAACGTGCACGT
GATTGGATTCAGCTGCAGGGTGGTGTTAAACAGGTTCGTGTTTTTACCCGTTTTTGGCTGGC
ACTGATTGGTGAATGGCCGTGGGAAGAAACCCCGAATCTGCCTCCGGAAATTCTGTTTTTTC
CGCGTTGGTTTCCGTTTAATATTTATCATTTTGCAGCATGGGCACGTGCAACCCTGGTTCCG
CTGTGTATTCTGAGCGCACGTCGTATGGTTGTTCCGCTGAATAAAAAAAGCTGTCTGCAGGA
ACTGTTTCCGGAAGATCGTTCTGCAGTTGTTGCACTGGGTAAAAAAGCCGGTGCATGGTCAA
CCTTTTTTTATCATGCAGATCGTGCCCTGAAAAAATATCAGCGTACCTTTAAACGTCCGCCT
GGTCGTCAGCAGGCAATTAAAATGTGTCTGGAATGGATTCTGCGTCGTCAGGATGCAGATGG
TGCATGGGGTGGTATTCAGCCTCCGTGGATTTATAGCCTGATGGCACTGAAAGCAGAAGGTT
ATCCGGTTACCCATCCGGTTATGGCAAAAGGTCTGGCAGCACTGGATGCACATTGGAGCTAT
GAACGTCCGGGTGGTGCACGTTTTGTTCAGGCATGTGAAAGTCCGGTTTGGGATACCCTGCT
GAGCAGCTTTGCACTGCTGGATTGTGGTTTTAGCTGTACCAGCAGCAGCGAACTGCGTAAAG
CAGTTGATTGGATTCTGGATCAGCAAGTTCTGCTGCCTGGTGATTGGCAGCAGAAACTGCCG
ACCGTTAGTCCGGGTGGTTGGGCATTTGAACGTGCAAATGTTCATTATCCGGATGTTGATGA
TACCGCAGTTGCACTGATTGTTCTGGCAAAAGTTCGTCCGGATTATCCTGATACCGCACGTG
TTAATCTGGCAATTGAACGTGGTCTGAATTGGCTGTTTGCAATGCAGTGTCGTAATGGTGGT
TGGGGTGCATTTGATAAAGATAATGATAAAGATCTGCTGACCAAAATTCCGTTTAGCGATTT
TGGCGAAACCATTGATCCGGCAAGCGTTGATGTTACCGCACATGTTCTGGAAGCACTGGGTC
TGCTGGGTTATCGTACCACCCATCCGGCAGTTGCAAAAGCACTGGAATTTATTCGTAGCGAA
CAGGAAAATGACGGTTGTTGGTTTGGTCGTTGGGGTGTGAATTATATTTATGGCACCGCAGC
AGTTCTGCCTGCACTGGCAAGCCTGAATATGAATATGAATCAGGAATTCATTCGTCGCGCAG
CAAATTGGATTCTGGGTAAACAGAATAATGATGGTGGCTGGGGTGAAAGCTGTGCAAGCTAT
ATGGATGATACCCAGCGTGGTCGTGGTCCGAGCACCGCAAGCCAGACCGCATGGGCAATGAT
GAGCCTGCTGGCAGTTGATGGTGGCACCTATGCAGAAAGCCTGCTGCGTGCAGAAGCATATC
TGAAAACCACCCAGACACCGGAAGGCACCTGGGATGAACCGTATTACACCGGCACCGGTTTT
CCGGGTTATGGTATTGGTCGTCGTGAAATTAAACGTCAGCGTAGCCTGCAGCAGCATGCAGA
ACTGAGCCGTGGTTTTATGATTAATTATAATCTGTATCGCCATTATTTTCCGCTGATGGCCC
TGGGTCGTCTGGCAGCTCTGCGTGGTGCA

2.22*Zmo*SHC1

2.22.1Amino acid sequence

MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIFKTMGNS
LNNTLVSACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDHPLRPRLGNALLEMQ

REDGSWGVYFGAGNGDINATVEAYAALRSLGYSADNPVLKKAAAWIAEKGGLKNIRVFTRYW
LALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMVPIAILSARRPSRPLRPQDRL
DELFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNLLKRNSLREAAIRHVLEWIIR
HQDADGGWGGIQPPWVYGLMALHGEGYQLYHPVMAKALSALDDPGWRHDRGESSWIQATNSP
VWDTMLALMALKDAKAEDRFTPEMDKAADWLLARQVKVKGDWSIKLPDVEPGGWAFEYANDR
YPDTDDTAVALIALSSYRDKEEWQKKGVEDAITRGVNWLIAMQSECGGWGAFDKDNNRSILS
KIPFCDFGESIDPPSVDVTAHVLEAFGTLGLSRDMPVIQKAIDYVRSEQEAEGAWFGRWGVN
YIYGTGAVLPALAAIGEDMTQPYITKACDWLVAHQQEDGGWGESCSSYMEIDSIGKGPTTPS
QTAWALMGLIAANRPEDYEAIAKGCHYLIDRQEQDGSWKEEEFTGTGFPGYGVGQTIKLDDP
ALSKRLLQGAELSRAFMLRYDFYRQFFPIMALSRAERLIDLNN

2.22.2DNA sequence

ATGGGTATTGACAGAATGAATAGCTTAAGTCGCTTGTTAATGAAGAAGATTTTCGGGGCTGA
AAAAACCTCGTATAAACCGGCTTCCGATACCATAATCGGAACGGATACCCTGAAAAGACCGA
ACCGGCGGCCTGAACCGACGGCAAAAGTCGACAAAACGATATTCAAGACTATGGGGAATAGT
CTGAATAATACCCTTGTTTCAGCCTGTGACTGGTTGATCGGACAACAAAAGCCCGATGGTCA
TTGGGTCGGTGCCGTGGAATCCAATGCTTCGATGGAAGCAGAATGGTGTCTGGCCTTGTGGT
TTTTGGGTCTGGAAGATCATCCGCTTCGTCCAAGATTGGGCAATGCTCTTTTGGAAATGCAG
CGGGAAGATGGCTCTTGGGGAGTCTATTTCGGCGCTGGAAATGGCGATATCAATGCCACGGT
TGAAGCCTATGCGGCCTTGCGGTCTTTGGGGTATTCTGCCGATAATCCTGTTTTGAAAAAAG
CGGCAGCATGGATTGCTGAAAAAGGCGGATTAAAAAATATCCGTGTCTTTACCCGTTATTGG
CTGGCGTTGATCGGGGAATGGCCTTGGGAAAAGACCCCTAACCTTCCCCCTGAAATTATCTG
GTTCCCTGATAATTTTGTCTTTTCGATTTATAATTTTGCCCAATGGGCGCGGGCAACCATGG
TGCCGATTGCTATTCTGTCCGCGAGACGACCAAGCCGCCCGCTGCGCCCTCAAGACCGATTG
GATGAACTGTTTCCAGAAGGCCGCGCTCGCTTTGATTATGAATTGCCGAAAAAAGAAGGCAT
CGATCTTTGGTCGCAATTTTTCCGAACCACTGACCGTGGATTACATTGGGTTCAGTCCAATC
TGTTAAAGCGCAATAGCTTGCGTGAAGCCGCTATCCGTCATGTTTTGGAATGGATTATCCGG
CATCAGGATGCCGATGGCGGTTGGGGTGGAATTCAGCCACCTTGGGTCTATGGTTTGATGGC
GTTACATGGTGAAGGCTATCAGCTTTATCATCCGGTGATGGCCAAGGCTTTGTCGGCTTTGG
ATGATCCCGGTTGGCGACATGACAGAGGCGAGTCTTCTTGGATACAGGCCACCAATAGTCCG
GTATGGGATACAATGTTGGCCTTGATGGCGTTAAAAGACGCCAAGGCCGAGGATCGTTTTAC
GCCGGAAATGGATAAGGCCGCCGATTGGCTTTTGGCTCGACAGGTCAAAGTCAAAGGCGATT
GGTCAATCAAACTGCCCGATGTTGAACCCGGTGGATGGGCATTTGAATATGCCAATGATCGC
TATCCCGATACCGATGATACCGCCGTCGCTTTGATCGCCCTTTCCTCTTATCGTGATAAGGA
GGAGTGGCAAAAGAAAGGCGTTGAGGACGCCATTACCCGTGGGGTTAATTGGTTGATCGCCA
TGCAAAGCGAATGTGGCGGTTGGGGAGCCTTTGATAAGGATAATAACAGAAGTATCCTTTCC
AAAATTCCTTTTTGTGATTTCGGAGAATCTATTGATCCGCCTTCAGTCGATGTAACGGCGCA
TGTTTTAGAGGCCTTTGGCACCTTGGGACTGTCCCGCGATATGCCGGTCATCCAAAAAGCGA
TCGACTATGTCCGTTCCGAACAGGAAGCCGAAGGCGCGTGGTTTGGTCGTTGGGGCGTTAAT
TATATCTATGGCACCGGTGCGGTTCTGCCTGCTTTGGCGGCGATCGGTGAAGATATGACCCA
GCCTTACATCACCAAGGCTTGCGATTGGCTGGTCGCACATCAGCAGGAAGACGGCGGTTGGG
GCGAAAGCTGCTCTTCCTATATGGAGATTGATTCCATTGGGAAGGGCCCAACCACGCCGTCC
CAGACTGCTTGGGCTTTGATGGGGTTGATCGCGGCCAATCGTCCCGAAGATTATGAAGCCAT
TGCCAAGGGATGCCATTATCTGATTGATCGCCAAGAGCAGGATGGTAGCTGGAAAGAAGAAG
AATTCACCGGCACCGGATTCCCCGGTTATGGCGTGGGTCAGACGATCAAGTTGGATGATCCG
GCTTTATCGAAACGATTGCTTCAAGGCGCTGAACTGTCACGGGCGTTTATGCTGCGTTATGA
TTTTTATCGGCAATTCTTCCCGATTATGGCGTTAAGTCGGGCAGAGAGACTGATTGATTTGA
ATAATTGA

2.23*Zmo*SHC2

2.23.1Amino acid sequence

MTVSTSSAFHHSPLSDDVEPIIQKATRALLEKQQQDGHWVFELEADATIPAEYILLKHYLGE
PEDLEIEAKIGRYLRRIQGEHGGWSLFYGGDLDLSATVKAYFALKMIGDSPDAPHMLRARNE
ILARGGAMRANVFTRIQLALFGAMSWEHVPQMPVELMLMPEWFPVHINKMAYWARTVLVPLL
VLQALKPVARNRRGILVDELFVPDVLPTLQESGDPIWRRFFSALDKVLHKVEPYWPKNMRAK
AIHSCVHFVTERLNGEDGLGAIYPAIANSVMMYDALGYPENHPERAIARRAVEKLMVLDGTE
DQGDKEVYCQPCLSPIWDTALVAHAMLEVGGDEAEKSAISALSWLKPQQILDVKGDWAWRRP
DLRPGGWAFQYRNDYYPDVDDTAVVTMAMDRAAKLSDLHDDFEESKARAMEWTIGMQSDNGG
WGAFDANNSYTYLNNIPFADHGALLDPPTVDVSARCVSMMAQAGISITDPKMKAAVDYLLKE
QEEDGSWFGRWGVNYIYGTWSALCALNVAALPHDHLAVQKAVAWLKTIQNEDGGWGENCDSY
ALDYSGYEPMDSTASQTAWALLGLMAVGEANSEAVTKGINWLAQNQDEEGLWKEDYYSGGGF
PRVFYLRYHGYSKYFPLWALARYRNLKKANQPIVHYGM

2.23.2DNA sequence

ATGACTGTATCGACTTCCTCGGCTTTTCATCATAGCCCGTTGTCTGATGATGTTGAGCCGAT
TATCCAAAAGGCCACCCGTGCCTTGCTTGAGAAGCAGCAGCAGGATGGCCATTGGGTTTTTG
AATTGGAAGCCGATGCAACCATTCCCGCTGAATACATCCTGTTAAAGCATTATTTGGGTGAA
CCCGAAGATTTAGAAATAGAGGCCAAGATAGGTCGCTATTTGCGTCGTATTCAGGGCGAGCA
TGGCGGATGGTCTTTGTTTTATGGTGGTGATCTTGATTTGAGCGCCACGGTCAAAGCCTATT
TTGCCTTGAAAATGATCGGAGATTCTCCTGATGCGCCTCATATGCTTCGAGCCAGAAATGAA
ATTTTGGCACGGGGTGGGGCGATGCGTGCCAATGTCTTTACACGTATTCAATTAGCTCTGTT
CGGGGCAATGTCATGGGAGCATGTCCCTCAAATGCCCGTAGAGTTGATGTTGATGCCGGAAT
GGTTTCCGGTTCACATCAATAAAATGGCCTATTGGGCAAGAACCGTTTTAGTCCCGTTATTG
GTTTTACAGGCGTTAAAGCCTGTCGCCCGTAATCGGCGCGGTATCTTGGTTGATGAATTATT
TGTGCCGGATGTTTTACCGACCCTTCAGGAAAGCGGTGACCCTATATGGCGTCGTTTTTTTT
CGGCACTTGATAAGGTATTGCATAAAGTAGAACCTTATTGGCCGAAAAATATGCGCGCGAAG
GCTATTCATAGCTGTGTCCATTTTGTGACCGAGCGTTTGAATGGTGAAGACGGGTTGGGTGC
TATTTATCCGGCGATTGCCAATAGCGTCATGATGTATGATGCCTTGGGATATCCCGAAAACC
ATCCAGAAAGAGCCATTGCCCGTCGGGCTGTCGAAAAATTGATGGTGTTAGATGGCACGGAA
GATCAGGGTGATAAAGAAGTCTACTGTCAGCCTTGTTTATCCCCGATTTGGGATACCGCTTT
GGTTGCCCATGCCATGTTGGAAGTCGGAGGCGATGAGGCTGAAAAATCGGCTATTTCTGCCT
TGAGCTGGTTAAAGCCGCAACAAATTTTGGATGTAAAGGGCGATTGGGCATGGCGGCGGCCT
GATCTCAGACCCGGGGGATGGGCCTTTCAATATAGAAATGACTATTATCCCGATGTCGATGA
TACGGCTGTTGTGACTATGGCGATGGATCGAGCCGCAAAATTGTCGGATCTTCACGATGATT
TTGAGGAATCTAAAGCGCGTGCCATGGAATGGACCATTGGGATGCAAAGCGATAATGGCGGT
TGGGGCGCTTTCGATGCCAATAACAGCTATACTTATCTGAATAATATTCCCTTTGCTGATCA
TGGCGCGTTACTTGATCCGCCAACGGTCGATGTCTCGGCACGCTGCGTTTCAATGATGGCGC
AAGCCGGTATCTCGATTACAGATCCCAAAATGAAAGCGGCAGTTGATTATCTTCTGAAAGAG
CAAGAAGAGGATGGTAGCTGGTTCGGGCGTTGGGGTGTCAATTACATATATGGCACATGGTC
GGCCTTATGTGCATTGAATGTGGCCGCTTTACCCCATGATCATTTAGCTGTTCAGAAAGCTG
TGGCTTGGCTGAAAACTATTCAAAATGAAGATGGTGGTTGGGGTGAAAATTGCGATAGCTAT
GCCCTTGATTATAGCGGATACGAGCCGATGGATTCGACGGCTTCCCAAACAGCATGGGCTTT
ATTGGGCTTGATGGCTGTTGGGGAAGCTAATTCCGAGGCCGTGACAAAGGGTATAAACTGGT
TGGCACAAAATCAGGATGAAGAAGGATTGTGGAAAGAAGATTATTATAGTGGCGGTGGTTTT
CCCCGTGTTTTTTATCTTCGGTATCACGGTTATTCCAAATATTTTCCTCTTTGGGCTTTAGC
GCGCTATCGCAATTTGAAAAAAGCCAATCAGCCGATTGTTCATTATGGGATGTAA

3 Mutant SHCs

3.1 *Zmo*SHC1_F486Y

3.1.1 Amino acid sequence

MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIFKTMGNS
LNNTLVSACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDHPLRPRLGNALLEMQ
REDGSWGVYFGAGNGDINATVEAYAALRSLGYSADNPVLKKAAAWIAEKGGLKNIRVFTRYW
LALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMVPIAILSARRPSRPLRPQDRL
DELFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNLLKRNSLREAAIRHVLEWIIR
HQDADGGWGGIQPPWVYGLMALHGEGYQLYHPVMAKALSALDDPGWRHDRGESSWIQATNSP
VWDTMLALMALKDAKAEDRFTPEMDKAADWLLARQVKVKGDWSIKLPDVEPGGWAFEYANDR
YPDTDDTAVALIALSSYRDKEEWQKKGVEDAITRGVNWLIAMQSECGGWGAYDKDNNRSILS
KIPFCDFGESIDPPSVDVTAHVLEAFGTLGLSRDMPVIQKAIDYVRSEQEAEGAWFGRWGVN
YIYGTGAVLPALAAIGEDMTQPYITKACDWLVAHQQEDGGWGESCSSYMEIDSIGKGPTTPS
QTAWALMGLIAANRPEDYEAIAKGCHYLIDRQEQDGSWKEEEFTGTGFPGYGVGQTIKLDDP
ALSKRLLQGAELSRAFMLRYDFYRQFFPIMALSRAERLIDLNN

3.1.2 DNA sequence

ATGGGTATTGACAGAATGAATAGCTTAAGTCGCTTGTTAATGAAGAAGATTTTCGGGGCTGA
AAAAACCTCGTATAAACCGGCTTCCGATACCATAATCGGAACGGATACCCTGAAAAGACCGA
ACCGGCGGCCTGAACCGACGGCAAAAGTCGACAAAACGATATTCAAGACTATGGGGAATAGT
CTGAATAATACCCTTGTTTCAGCCTGTGACTGGTTGATCGGACAACAAAAGCCCGATGGTCA
TTGGGTCGGTGCCGTGGAATCCAATGCTTCGATGGAAGCAGAATGGTGTCTGGCCTTGTGGT
TTTTGGGTCTGGAAGATCATCCGCTTCGTCCAAGATTGGGCAATGCTCTTTTGGAAATGCAG
CGGGAAGATGGCTCTTGGGGAGTCTATTTCGGCGCTGGAAATGGCGATATCAATGCCACGGT
TGAAGCCTATGCGGCCTTGCGGTCTTTGGGGTATTCTGCCGATAATCCTGTTTTGAAAAAAG
CGGCAGCATGGATTGCTGAAAAAGGCGGATTAAAAAATATCCGTGTCTTTACCCGTTATTGG
CTGGCGTTGATCGGGGAATGGCCTTGGGAAAAGACCCCTAACCTTCCCCCTGAAATTATCTG
GTTCCCTGATAATTTTGTCTTTTCGATTTATAATTTTGCCCAATGGGCGCGGGCAACCATGG
TGCCGATTGCTATTCTGTCCGCGAGACGACCAAGCCGCCCGCTGCGCCCTCAAGACCGATTG
GATGAACTGTTTCCAGAAGGCCGCGCTCGCTTTGATTATGAATTGCCGAAAAAAGAAGGCAT
CGATCTTTGGTCGCAATTTTTCCGAACCACTGACCGTGGATTACATTGGGTTCAGTCCAATC
TGTTAAAGCGCAATAGCTTGCGTGAAGCCGCTATCCGTCATGTTTTGGAATGGATTATCCGG
CATCAGGATGCCGATGGCGGTTGGGGTGGAATTCAGCCACCTTGGGTCTATGGTTTGATGGC
GTTACATGGTGAAGGCTATCAGCTTTATCATCCGGTGATGGCCAAGGCTTTGTCGGCTTTGG
ATGATCCCGGTTGGCGACATGACAGAGGCGAGTCTTCTTGGATACAGGCCACCAATAGTCCG
GTATGGGATACAATGTTGGCCTTGATGGCGTTAAAAGACGCCAAGGCCGAGGATCGTTTTAC
GCCGGAAATGGATAAGGCCGCCGATTGGCTTTTGGCTCGACAGGTCAAAGTCAAAGGCGATT
GGTCAATCAAACTGCCCGATGTTGAACCCGGTGGATGGGCATTTGAATATGCCAATGATCGC
TATCCCGATACCGATGATACCGCCGTCGCTTTGATCGCCCTTTCCTCTTATCGTGATAAGGA
GGAGTGGCAAAAGAAAGGCGTTGAGGACGCCATTACCCGTGGGGTTAATTGGTTGATCGCCA
TGCAAAGCGAATGTGGCGGTTGGGGAGCCTATGATAAGGATAATAACAGAAGTATCCTTTCC
AAAATTCCTTTTTGTGATTTCGGAGAATCTATTGATCCGCCTTCAGTCGATGTAACGGCGCA
TGTTTTAGAGGCCTTTGGCACCTTGGGACTGTCCCGCGATATGCCGGTCATCCAAAAAGCGA

TCGACTATGTCCGTTCCGAACAGGAAGCCGAAGGCGCGTGGTTTGGTCGTTGGGGCGTTAAT
TATATCTATGGCACCGGTGCGGTTCTGCCTGCTTTGGCGGCGATCGGTGAAGATATGACCCA
GCCTTACATCACCAAGGCTTGCGATTGGCTGGTCGCACATCAGCAGGAAGACGGCGGTTGGG
GCGAAAGCTGCTCTTCCTATATGGAGATTGATTCCATTGGGAAGGGCCCAACCACGCCGTCC
CAGACTGCTTGGGCTTTGATGGGGTTGATCGCGGCCAATCGTCCCGAAGATTATGAAGCCAT
TGCCAAGGGATGCCATTATCTGATTGATCGCCAAGAGCAGGATGGTAGCTGGAAAGAAGAAG
AATTCACCGGCACCGGATTCCCCGGTTATGGCGTGGGTCAGACGATCAAGTTGGATGATCCG
GCTTTATCGAAACGATTGCTTCAAGGCGCTGAACTGTCACGGGCGTTTATGCTGCGTTATGA
TTTTTATCGGCAATTCTTCCCGATTATGGCGTTAAGTCGGGCAGAGAGACTGATTGATTTGA
ATAATTGA

3.2 *Aac*SHC_Y420C

3.2.1 Amino acid sequence

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRME
KIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIE
SSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPV
FPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWL
LERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISP
VWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYY
PDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGACDVDNTSDLPNHIPF
CDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYG
TGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLA
LGRYKQAIERR

3.2.2 DNA sequence

ATGGCGGAACAGCTGGTGGAAGCGCCGGCGTATGCGCGCACCCTGGATCGCGCGGTGGAATA
TCTGCTGAGCTGCCAGAAAGATGAAGGCTATTGGTGGGGCCCGCTGCTGAGCAACGTGACCA
TGGAAGCGGAATATGTGCTGCTGTGCCATATTCTGGATCGCGTGGATCGCGATCGCATGGAA
AAAATTCGCCGCTATCTGCTGCATGAACAGCGCGAAGATGGCACCTGGGCGCTGTATCCGGG
CGGCCCGCCGGATCTGGATACCACCATTGAAGCGTATGTGGCGCTGAAATATATTGGCATGA
GCCGCGATGAAGAACCGATGCAGAAAGCGCTGCGCTTTATTCAGAGCCAGGGCGGCATTGAA
AGCAGCCGCGTGTTTACCCGCATGTGGCTGGCGCTGGTGGGCGAATATCCGTGGGAAAAAGT
GCCGATGGTGCCGCCGGAAATTATGTTTCTGGGCAAACGCATGCCGCTGAACATTTATGAAT
TTGGCAGCTGGGCGCGCGACCGTGGTGGCGCTGAGCATTGTGATGAGCCGCCAGCCGGTG
TTTCCGCTGCCGGAACGCGCGCGCGTGCCGGAACTGTATGAAACCGATGTGCCGCCGCGCCG
CCGCGGCGCGAAAGGCGGCGGcGGCTGGATTTTTGATGCGCTGGATCGCGCGCTGCATGGCT
ATCAGAAACTGAGCGTGCATCCGTTTCGCCGCGCGGCGGAAATTCGCGCGCTGGATTGGCTG
CTGGAACGCCAGGCGGGCGATGGCAGCTGGGGCGGCATTCAGCCGCCGTGGTTTTATGCGCT
GATTGCGCTGAAAATTCTGGATATGACCCAGCATCCGGCGTTTATTAAAGGCTGGGAAGGCC
TGGAACTGTATGGCGTGGAACTGGATTATGGCGGCTGGATGTTTCAGGCGAGCATTAGCCCG
GTGTGGGATACCGGCCTgGCGGTGCTGGCGCTGCGCGCGGCGGGCCTGCCGGCGGATCATGA
TCGCCTGGTGAAAGCGGGCGAATGGCTGCTGGATCGCCAGATTACCGTGCCGGGCGATTGGG
CGGTGAAACGCCCGAACCTGAAACCGGGCGGCTTTGCGTTTCAGTTTGATAACGTGTATTAT
CCGGATGTGGATGATACCGCGGTggtGGTGtGGGCGCTGAACACCCTGCGCCTGCCGGATGA
ACGCCGCCGCCGCGATGCGATGACCAAAGGCTTTCGCTGGATTGTGGGCATGCAGAGCAGCA
ACGGCGGCTGGGGCGCCTGTGATGTGGATAACACCAGCGATCTGCCGAACCATATTCCGTTT

TGCGATTTTGGCGAAGTGACCGATCCGCCGAGCGAAGATGTGACCGCGCATGTGCTGGAATG
CTTTGGCAGCTTTGGCTATGATGATGCGTGGAAAGTGATTCGCCGCGCGGTGGAATATCTGA
AACGCGAACAGAAACCGGATGGCAGCTGGTTTGGCCGCTGGGGCGTGAACTATCTGTATGGC
ACCGGCGCGGTGGTGAGCGCGCTGAAAGCGGTGGGCATTGATACCCGCGAACCGTATATTCA
GAAAGCGCTGGATTGGGTGGAACAGCATCAGAACCCGGATGGCGGCTGGGGCGAAGATTGCC
GCAGCTATGAAGATCCGGCGTATGCGGGCAAAGGCGCGAGCACCCCGAGCCAGACCGCGTGG
GCGCTGATGGCGCTGATTGCGGGCGGCCGCGCGGAAAGCGAAGCGGCGCGCCGCGGCGTGCA
GTATCTGGTGGAAACCCAGCGCCCGGATGGCGGCTGGGATGAACCGTATTATACCGGCACCG
GGTTCCCGGGCGATTTTTATCTGGGCTATACCATGTATCGCCATGTGTTTCCGACCCTGGCG
CTGGGCCGCTATAAACAGGCGATTGAACGCCGCTAA

3.3 *Zmo*SHC1_Loop

3.3.1 Amino acid sequence

MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIFKTMGNS
LNNTLVSACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDHPLRPRLGNALLEMQ
REDGSWGVYFGAGNGDINATVEAYAALRSLGYSADNPVLKKAAAWIAEKGGLKNIRVFTRYW
LALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMVPIAILSARRPSRPLRPQDRL
DELFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNLLKRNSLREAAIRHVLEWIIR
HQDADGGWGGIQPPWVYGLMALHGEGYQLYHPVMAKALSALDDPGWRHDRGESSWIQATNSP
VWDTMLALMALKDAKAEDRFTPEMDKAADWLLARQVKVKGDWSIKLPDVEPGGWAFEYANDR
YPDTDDTAVALIALSSYRDKEEWQKKGVEDAITRGVNWLIAMQSECGGWGAFDKDNNRSILS
KIPFCDFGESIDPPSVDVTAHVLEAFGTLGLSRDMPVIQKAIDYVRSEQEAEGAWFGRWGVN
YIYGTGAVLPALAAIGEDMTQPYITKACDWLVAHQQEDGGWGESCSSYMEIDSIGKGPTTPS
QTAWALMGLIAANRPEDYEAIAKGCHYLIDRQEQDGSWKEEEFTGTGFPRAFMLRYDFYRQF
FPIMALSRAERLIDLNN

3.3.2 DNA sequence

ATGGGTATTGACAGAATGAATAGCTTAAGTCGCTTGTTAATGAAGAAGATTTTCGGGGCTGA
AAAAACCTCGTATAAACCGGCTTCCGATACCATAATCGGAACGGATACCCTGAAAAGACCGA
ACCGGCGGCCTGAACCGACGGCAAAAGTCGACAAAACGATATTCAAGACTATGGGGAATAGT
CTGAATAATACCCTTGTTTCAGCCTGTGACTGGTTGATCGGACAACAAAAGCCCGATGGTCA
TTGGGTCGGTGCCGTGGAATCCAATGCTTCGATGGAAGCAGAATGGTGTCTGGCCTTGTGGT
TTTTGGGTCTGGAAGATCATCCGCTTCGTCCAAGATTGGGCAATGCTCTTTTGGAAATGCAG
CGGGAAGATGGCTCTTGGGGAGTCTATTTCGGCGCTGGAAATGGCGATATCAATGCCACGGT
TGAAGCCTATGCGGCCTTGCGGTCTTTGGGGTATTCTGCCGATAATCCTGTTTTGAAAAAAG
CGGCAGCATGGATTGCTGAAAAAGGCGGATTAAAAAATATCCGTGTCTTTACCCGTTATTGG
CTGGCGTTGATCGGGGAATGGCCTTGGGAAAAGACCCCTAACCTTCCCCCTGAAATTATCTG
GTTCCCTGATAATTTTGTCTTTTCGATTTATAATTTTGCCCAATGGGCGCGGGCAACCATGG
TGCCGATTGCTATTCTGTCCGCGAGACGACCAAGCCGCCCGCTGCGCCCTCAAGACCGATTG
GATGAACTGTTTCCAGAAGGCCGCGCTCGCTTTGATTATGAATTGCCGAAAAAAGAAGGCAT
CGATCTTTGGTCGCAATTTTTCCGAACCACTGACCGTGGATTACATTGGGTTCAGTCCAATC
TGTTAAAGCGCAATAGCTTGCGTGAAGCCGCTATCCGTCATGTTTTGGAATGGATTATCCGG
CATCAGGATGCCGATGGCGGTTGGGGTGGAATTCAGCCACCTTGGGTCTATGGTTTGATGGC
GTTACATGGTGAAGGCTATCAGCTTTATCATCCGGTGATGGCCAAGGCTTTGTCGGCTTTGG
ATGATCCCGGTTGGCGACATGACAGAGGCGAGTCTTCTTGGATACAGGCCACCAATAGTCCG
GTATGGGATACAATGTTGGCCTTGATGGCGTTAAAAGACGCCAAGGCCGAGGATCGTTTTAC
GCCGGAAATGGATAAGGCCGCCGATTGGCTTTTGGCTCGACAGGTCAAAGTCAAAGGCGATT

GGTCAATCAAACTGCCCGATGTTGAACCCGGTGGATGGGCATTTGAATATGCCAATGATCGC
TATCCCGATACCGATGATACCGCCGTCGCTTTGATCGCCCTTTCCTCTTATCGTGATAAGGA
GGAGTGGCAAAAGAAAGGCGTTGAGGACGCCATTACCCGTGGGGTTAATTGGTTGATCGCCA
TGCAAAGCGAATGTGGCGGTTGGGGAGCCTTTGATAAGGATAATAACAGAAGTATCCTTTCC
AAAATTCCTTTTTGTGATTTCGGAGAATCTATTGATCCGCCTTCAGTCGATGTAACGGCGCA
TGTTTTAGAGGCCTTTGGCACCTTGGGACTGTCCCGCGATATGCCGGTCATCCAAAAAGCGA
TCGACTATGTCCGTTCCGAACAGGAAGCCGAAGGCGCGTGGTTTGGTCGTTGGGGCGTTAAT
TATATCTATGGCACCGGTGCGGTTCTGCCTGCTTTGGCGGCGATCGGTGAAGATATGACCCA
GCCTTACATCACCAAGGCTTGCGATTGGCTGGTCGCACATCAGCAGGAAGACGGCGGTTGGG
GCGAAAGCTGCTCTTCCTATATGGAGATTGATTCCATTGGGAAGGGCCCAACCACGCCGTCC
CAGACTGCTTGGGCTTTGATGGGGTTGATCGCGGCCAATCGTCCCGAAGATTATGAAGCCAT
TGCCAAGGGATGCCATTATCTGATTGATCGCCAAGAGCAGGATGGTAGCTGGAAAGAAGAAG
AATTCACCGGCACCGGATTCCCCCGGGCGTTTATGCTGCGTTATGATTTTTATCGGCAATTC
TTCCCGATTATGGCGTTAAGTCGGGCAGAGAGACTGATTGATTTGAATAATTGA

II GC chromatograms and NMR, IR and MS spectra

4 CG-FID and GC-MS chromatograms and spectra

4.1 Squalene, hopene, hopanol

4.1.1 GC-FID squalene conversion

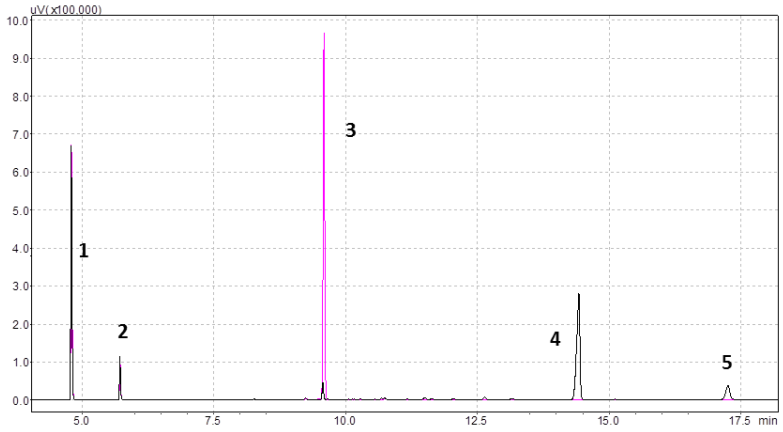


Fig. 4.1: GC-FID chromatograms of the conversion of squalene with *Aac*SHC (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): internal standard 2 (ISTD2 1-dodecanol); (3): substrate squalene; (4): product hopene; (5): product hopanol

4.1.2 GC-MS squalene conversion

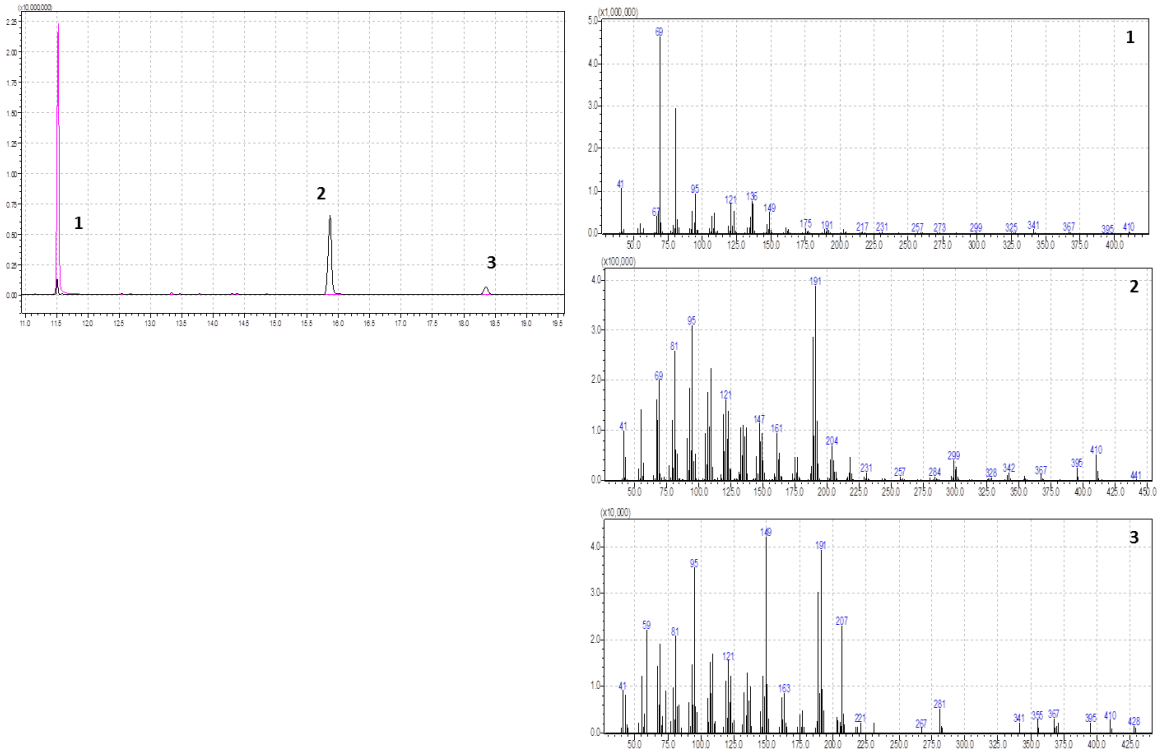


Fig. 4.2: GC-MS chromatograms of the conversion of squalene with *Aac*SHC (black) and negative control (pink), (1): substrate squalene; (2): product hopene; (3): product hopanol.

4.2 Homofarnesol, ambroxan

4.2.1 GC-FID homofarnesol conversion

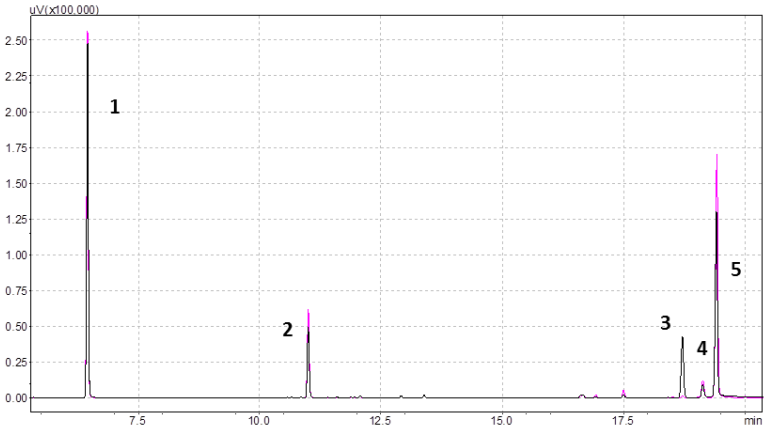


Fig. 4.3: GC-FID chromatograms of the conversion of homofarnesol with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): internal standard 2 (ISTD2 1-dodecanol); (3): product ambroxan; (4): (*E*,*Z*)-homofarnesol; (5): substrate (*E*,*E*)-homofarnesol.

4.2.2 GC-MS homofarnesol conversion

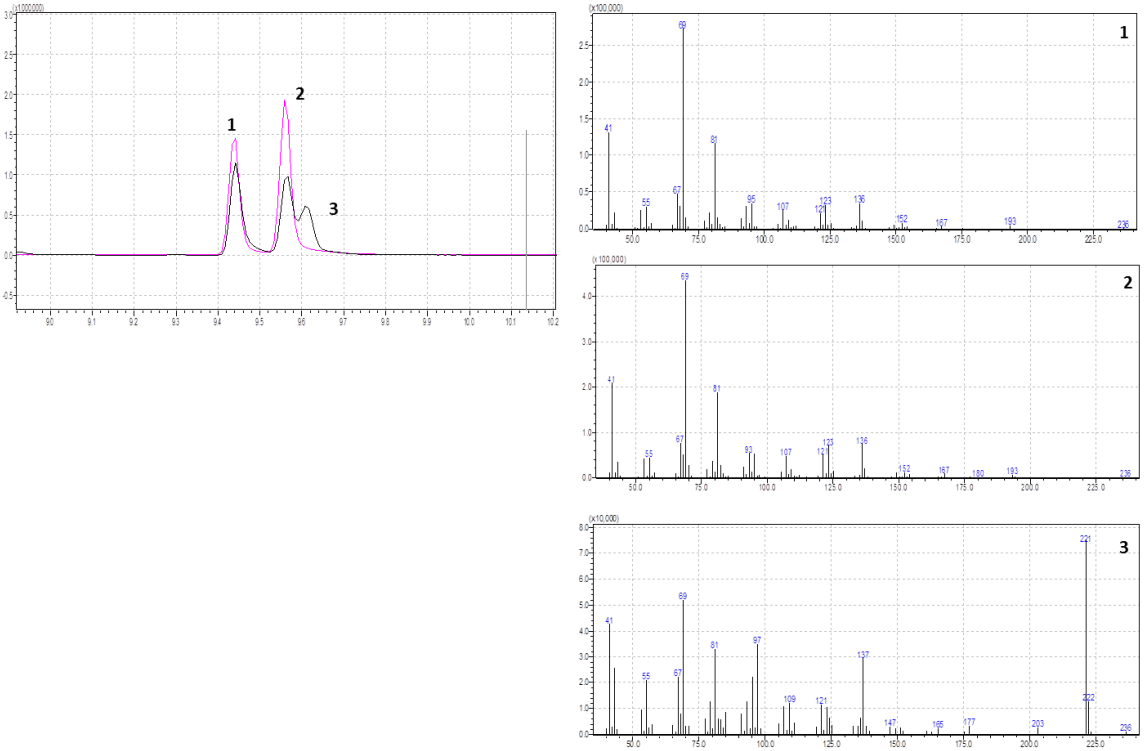


Fig. 4.4: GC-MS chromatograms and spectra of the conversion of homofarnesol with *Zmo*SHC1 (black) and negative control (pink), (1): (*E*,*Z*)-homofarnesol; (2): substrate (*E*,*E*)-homofarnesol, (3): product ambroxan.

4.3 Citronellal, isopulegol

4.3.1 GC-FID citronellal conversion

Substrate and product could not be separated under the given conditions.

4.3.2 GC-MS citronellal conversion

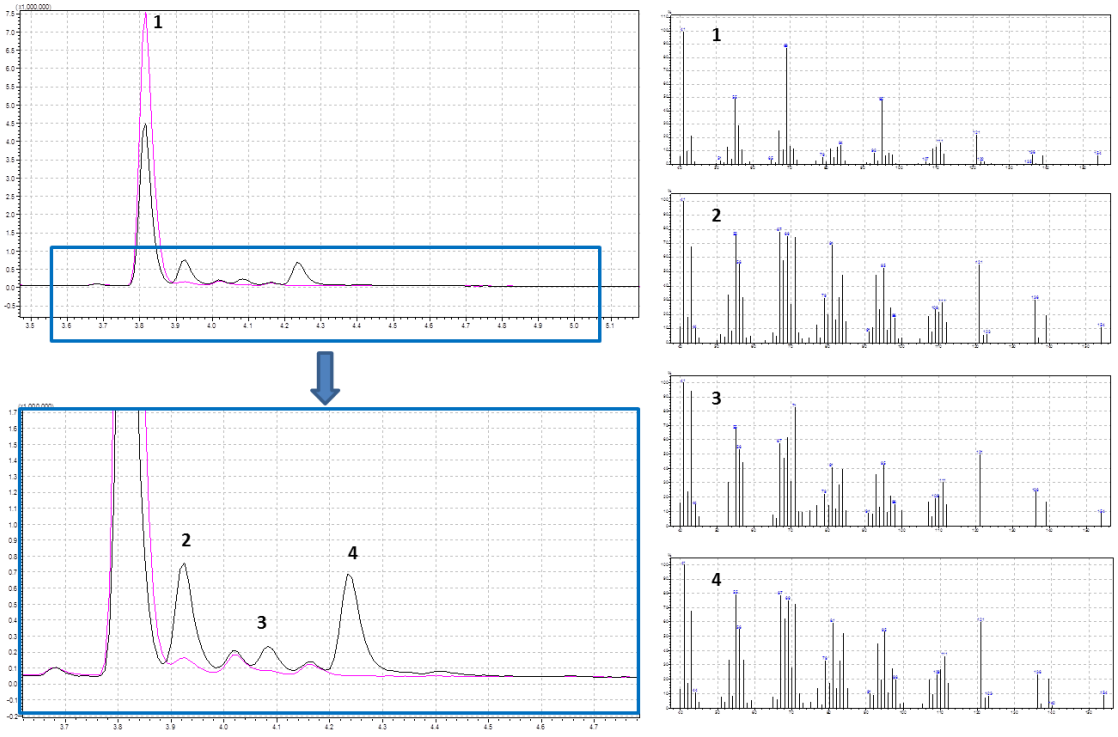


Fig. 4.5: GC-MS chromatograms and spectra of the conversion of (*S*)-citronellal with *Zmo*SHC1 (black) and negative control (pink), (1): citronellal; (2): isopulegol, (3): *neo*-isopulegol, (4): *iso*-isopulegol.

4.4 Homofarnesoic acid, sclareolide

4.4.1 GC-FID homofarnesoic acid conversion

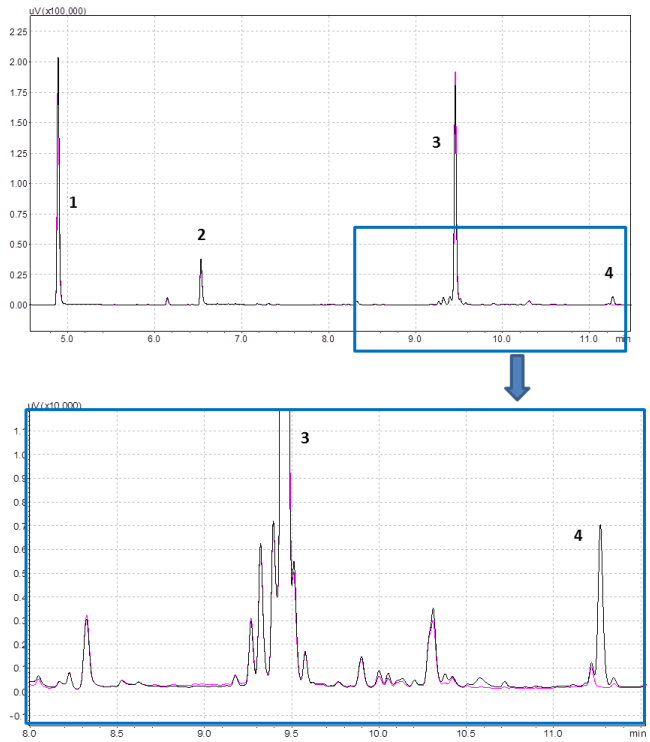


Fig. 4.6: GC-FID chromatograms of the conversion of homofarnesoic acid with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): internal standard 2 (ISTD2 1-dodecanol); (3): substrate homofarnesoic acid; (6): product sclareolide. The extract was derivatized with TMSH prior to GC analysis.

4.4.2 GC-MS homofarnesoic acid conversion

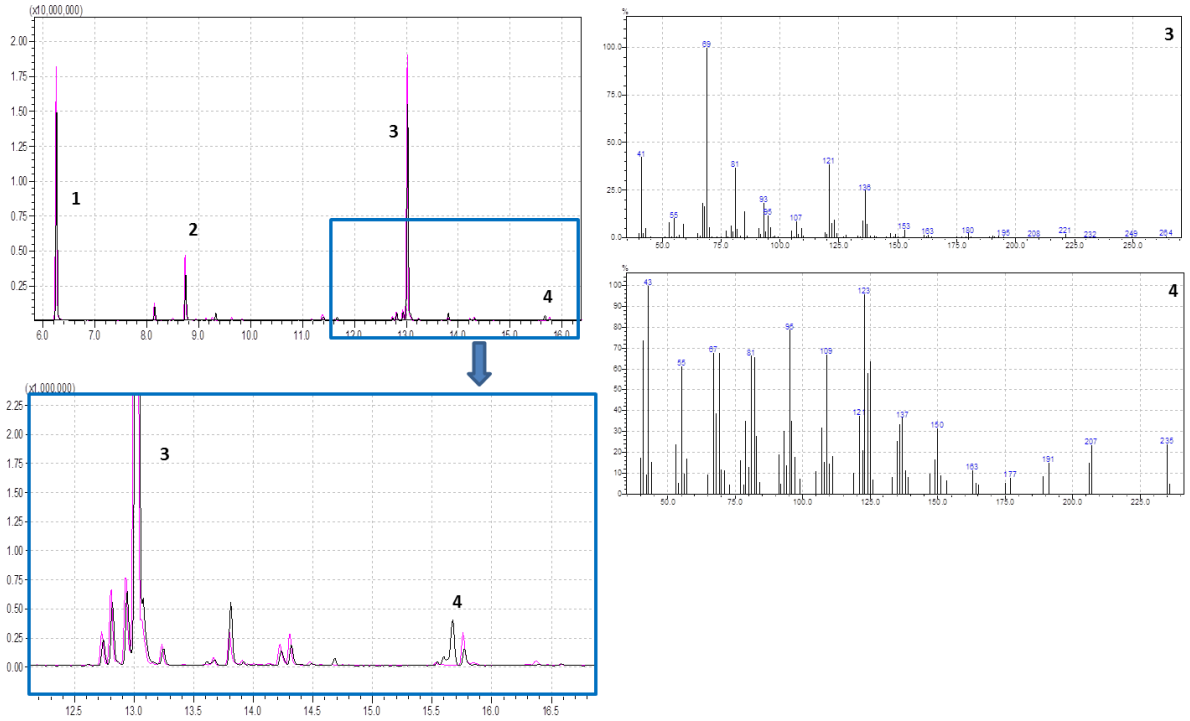


Fig. 4.7: GC-MS chromatograms and spectra of the conversion of homofarnesoic acid with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): internal standard 2 (ISTD2 1-dodecanol); (3): substrate homofarnesoic acid; (6): product sclareolide. The extract was derivatized with TMSH prior to GC analysis.

4.5 Farnesylacetone, sclareoloxide

4.5.1 GC-FID farnesylacetone conversion

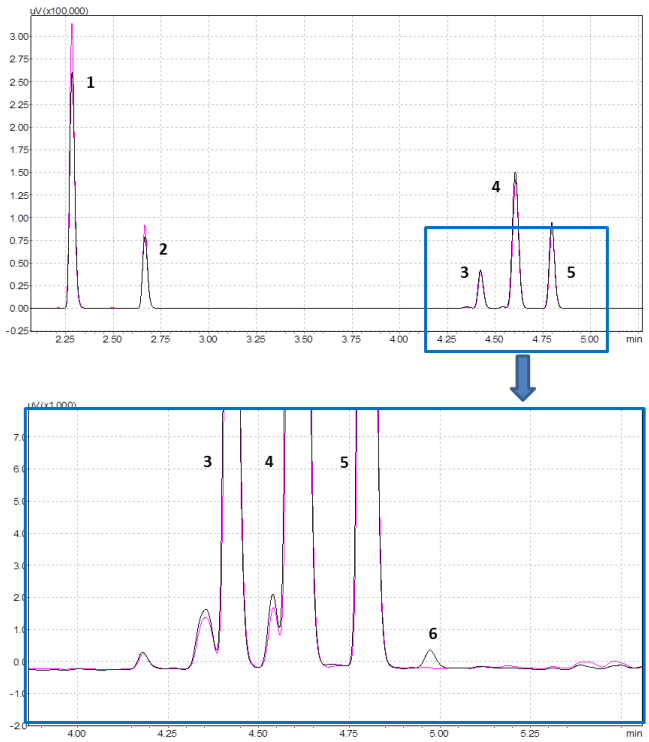


Fig. 4.8: GC-FID chromatograms of the conversion of farnesylacetone with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): internal standard 2 (ISTD2 1-dodecanol); (3), (4), (5): substrate farnesylacetone; (6): sclareoloxide.

4.6 GC-MS farnesylacetone conversion

Substrate and product could not be separated under the given conditions.

4.7 Geranylacetone, hexahydrochromene

4.7.1 GC-FID geranylacetone conversion

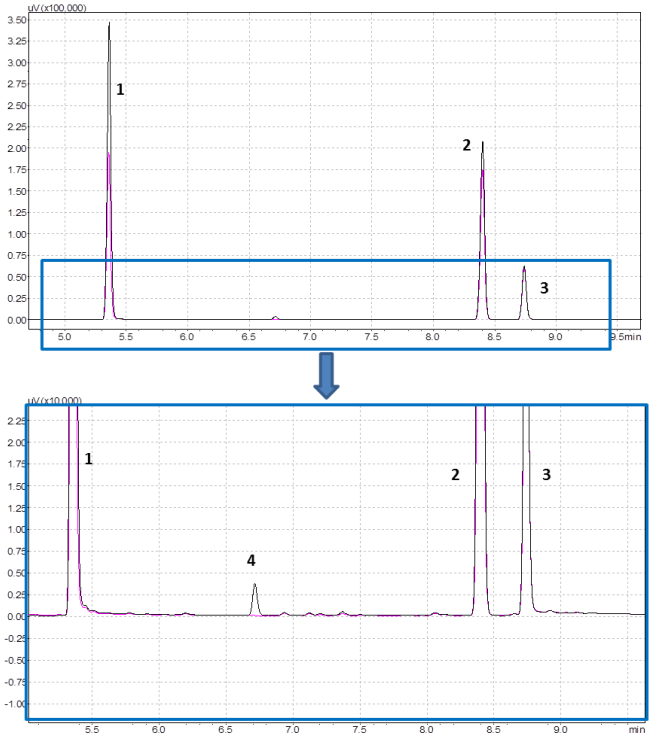


Fig. 4.9: GC-FID chromatograms of the conversion of geranylacetone with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): substrate geranylacetone; (3): internal standard 2 (ISTD2 1-dodecanol); (4): product hexahydrochromene.

4.7.2 GC-MS geranylacetone conversion

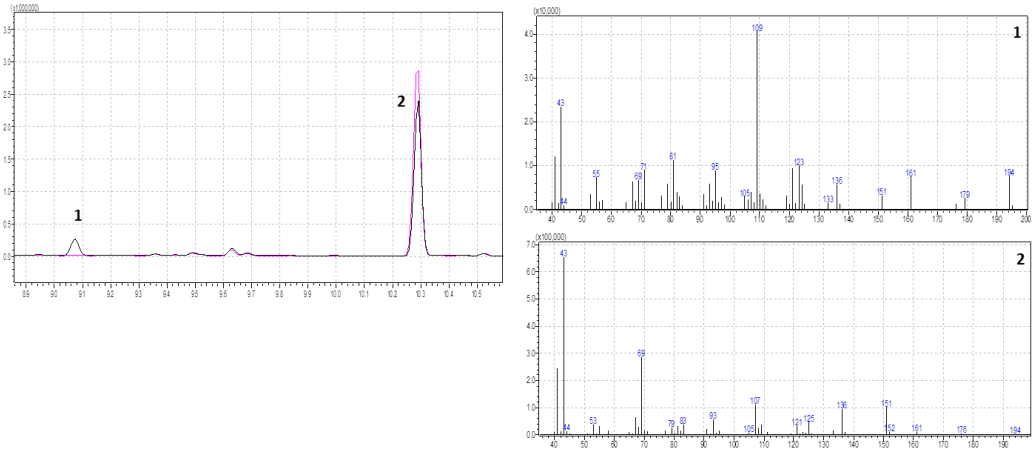


Fig. 4.10: GC-MS chromatograms and spectra of the conversion of geranylacetone with *Zmo*SHC1 (black) and negative control (pink), (1): product hexahydrochromene, (2) substrate geranylacetone.

4.8 Nerolidol, caparrapioxide

4.8.1 GC-FID nerolidol conversion

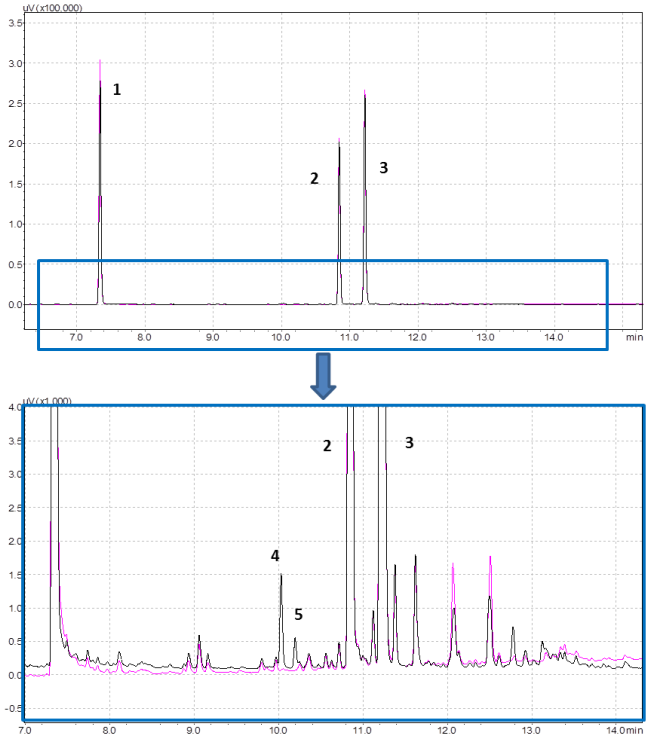


Fig. 4.11: GC-FID chromatograms of the conversion of nerolidol with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): substrate *cis*-nerolidol; (3): substrate *trans*-nerolidol; (4): product (-)-caparrapioxide; (5): product (-)-8-*epi*-caparrapioxide.

4.8.2 GC-MS nerolidol conversion

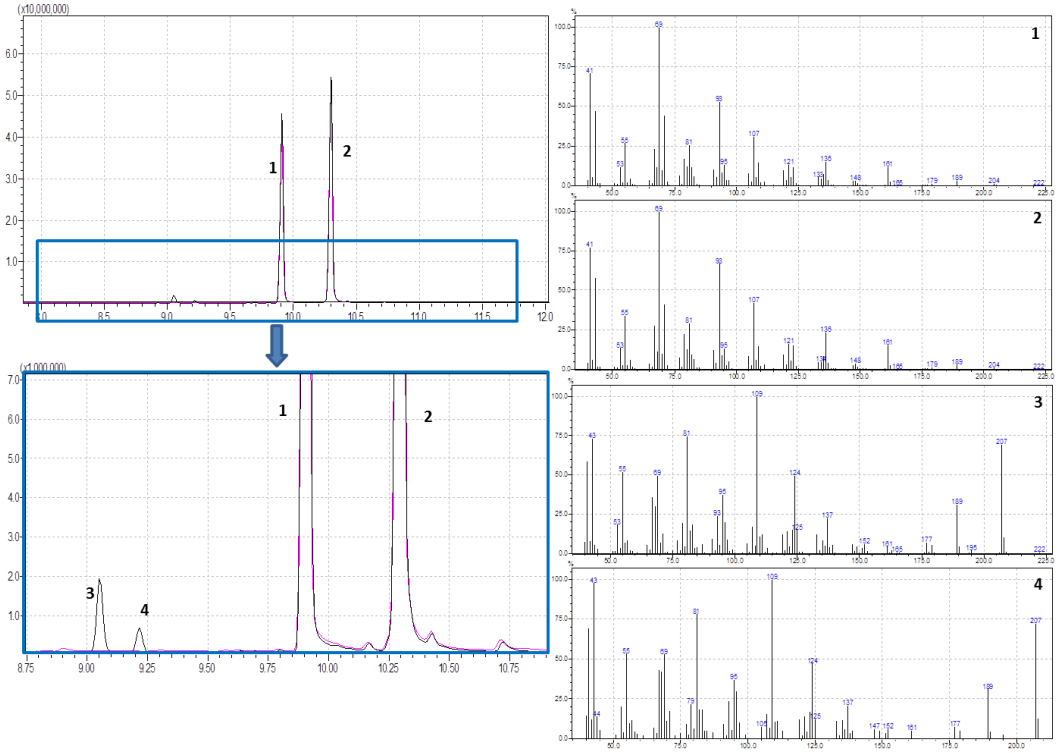


Fig. 4.12: GC-MS chromatograms and spectra of the conversion of nerolidol with *Zmo*SHC1 (black) and negative control (pink), (1): substrate *cis*-nerolidol; (2): substrate *trans*-nerolidol; (3): product (-)-caparrapioxide; (4): product (-)-8-*epi*-caparrapioxide.

4.9 Other substrates tested

4.9.1 GC-FID bishomofarnesol conversion

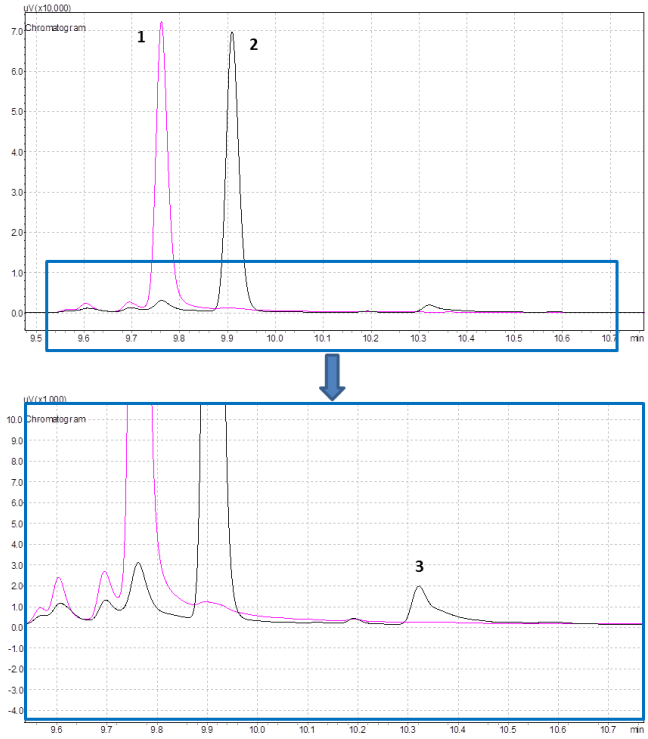


Fig. 4.13: GC-FID chromatograms of the conversion of bishomofarnesol with *Zmo*SHC1 (black) and negative control (pink). The substrate bishomofarnesol (1) was converted into two products (2 and 3).

4.9.2 GC-MS bishomofarnesol conversion

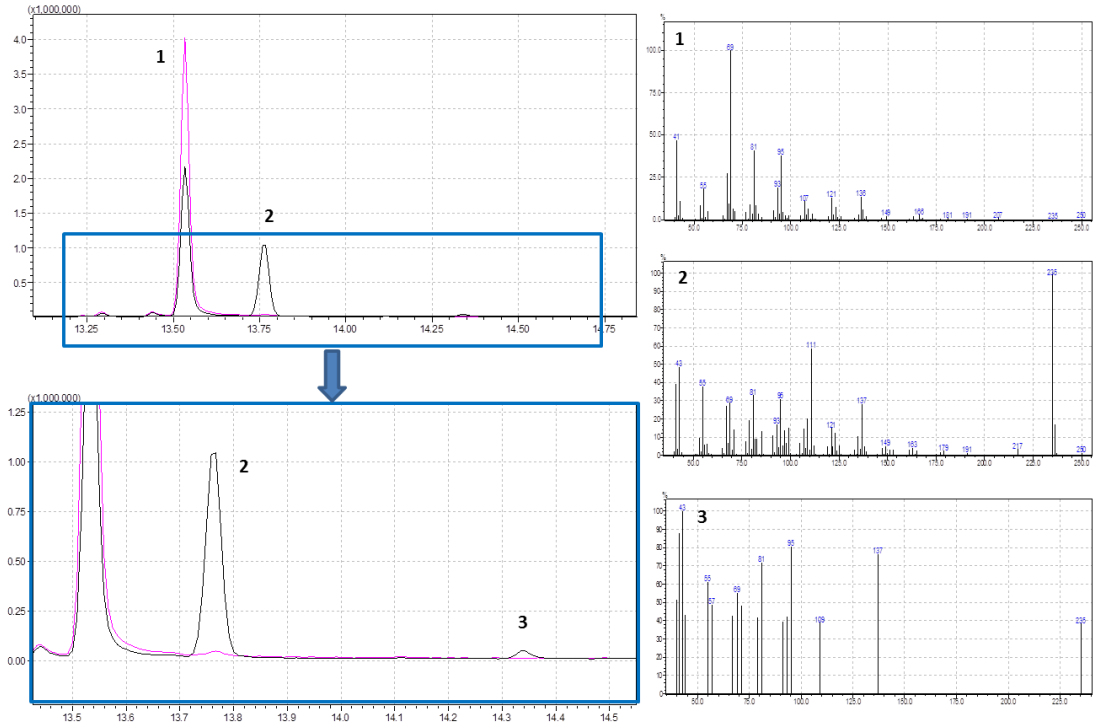


Fig. 4.14: GC-MS chromatogram and spectra of the conversion of bishomofarnesol with *Zmo*SHC1 (black) and negative control (pink), the GC-MS spectra of the substrate (1), the major product (2) and the minor product (3) are shown on the right side.

4.9.3 GC-FID geraniol conversion

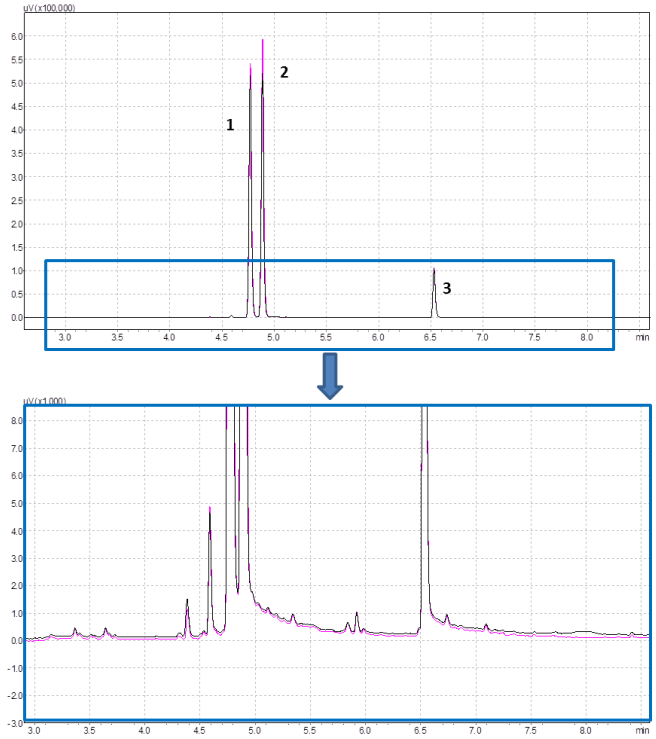


Fig. 4.15: GC-FID chromatograms of the conversion of geraniol with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): substrate geraniol; (3): internal standard 2 (ISTD2 1-dodecanol). No product peak was detected.

4.9.4 GC-FID linalool conversion

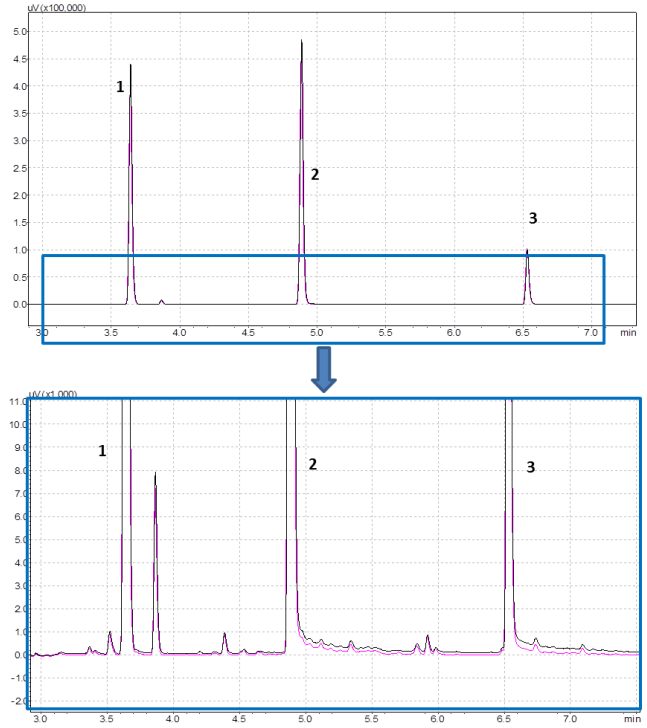


Fig. 4.16: GC-FID chromatograms of the conversion of linalool with *Zmo*SHC1 (black) and negative control (pink), (1): substrate linalool; (2): internal standard 1 (ISTD1 1-decanol); (3): internal standard 2 (ISTD2 1-dodecanol). No product peak was detected.

4.9.5 GC-FID pseudoionone conversion

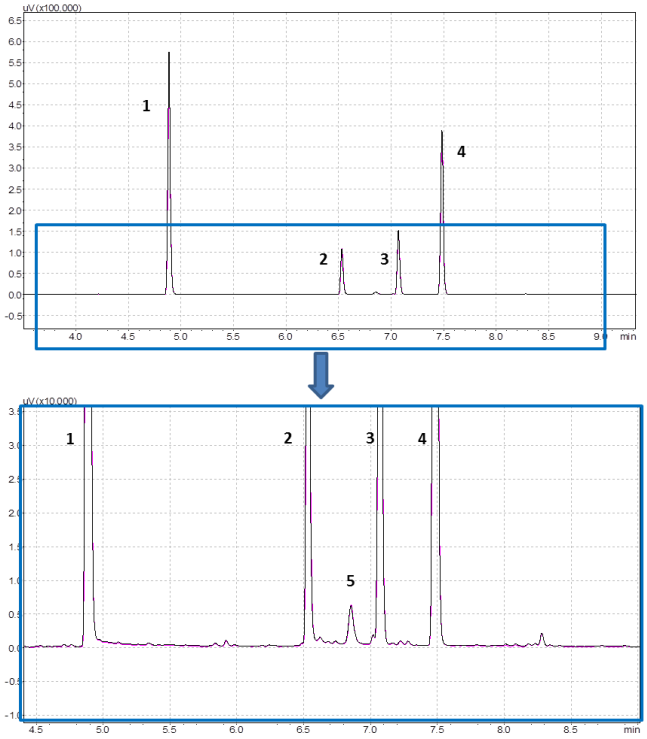


Fig. 4.17: GC-FID chromatograms of the conversion of pseudoionone with *Zmo*SHC1 (black) and negative control (pink), (1): internal standard 1 (ISTD1 1-decanol); (2): internal standard 2 (ISTD2 1-dodecanol); (3), (4), (5): substrate pseudoionone. No product peak was detected.

4.9.6 GC-FID geranic acid conversion

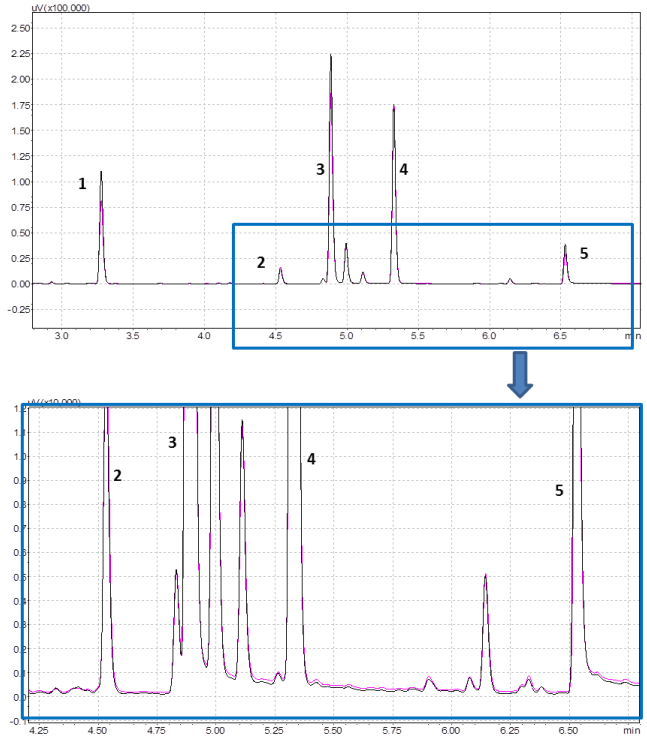


Fig. 4.18: GC-FID chromatograms of the conversion of geranic acid with *Zmo*SHC1 (black) and negative control (pink), (1), (2), (4): substrate geranic acid; (3): internal standard 1 (ISTD1 1-decanol); (5): internal standard 2 (ISTD2 1-dodecanol). No product peak was detected. The extract was derivatized with TMSH prior to GC analysis.

4.9.7 GC-FID bishomofarnesoic acid conversion

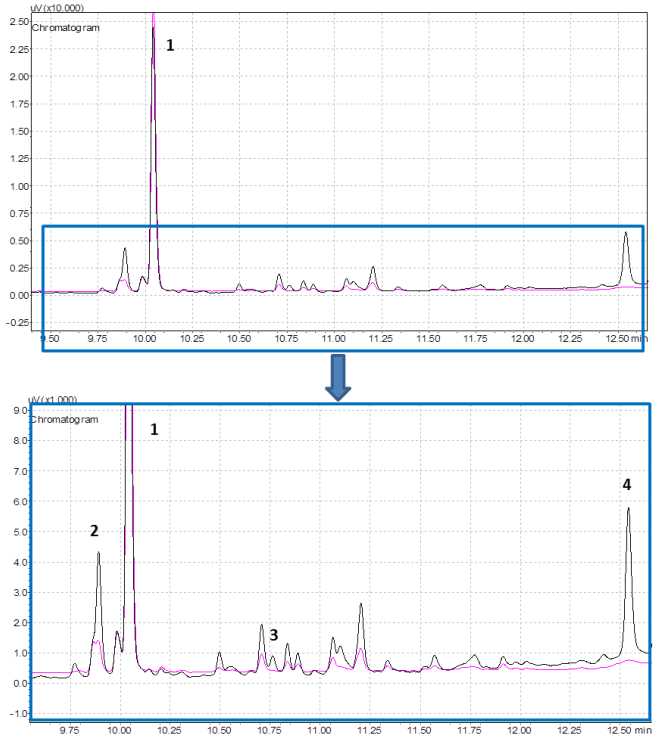


Fig. 4.19: GC-FID chromatogram of the conversion of bishomofarnesoic acid with *Zmo*SHC1 (black) and negative control (pink), (1): substrate bishomofarnesoic acid; (2), (3), (4): products.

4.9.8 GC-MS bishomofarnesoic acid conversion

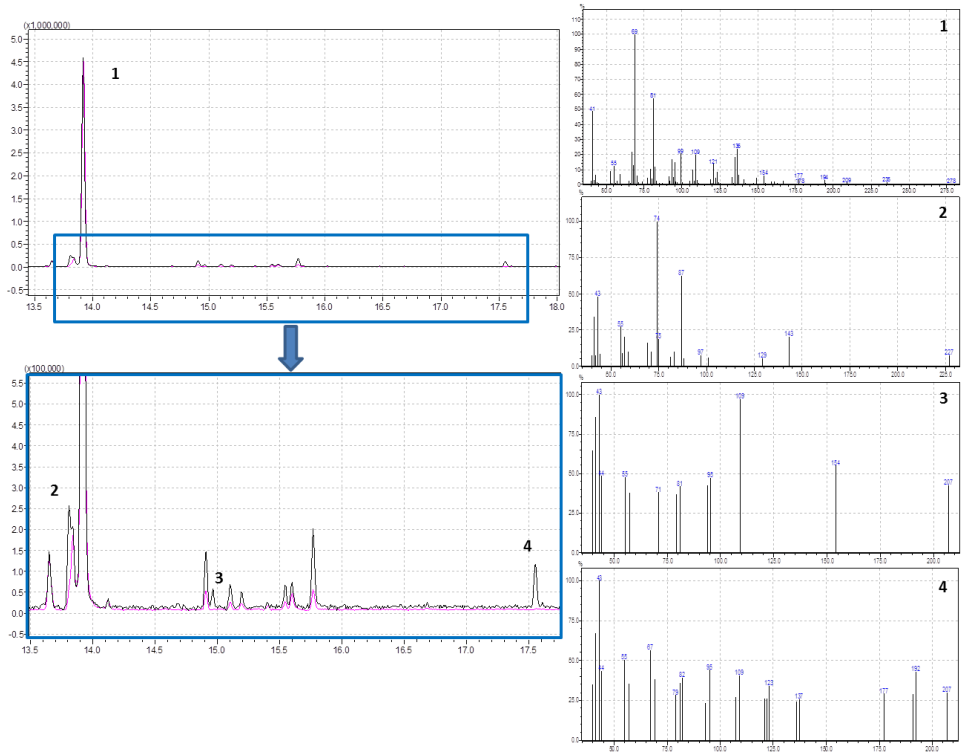


Fig. 4.20: GC-MS chromatogram and spectra of the conversion of bishomofarnesoic acid with *Zmo*SHC1 (black) and negative control (pink), the GC-MS spectra of the substrate (1), and the products (2), (3) and (4) are shown on the right side.

4.9.9 GC-MS bishomofarnesal conversion

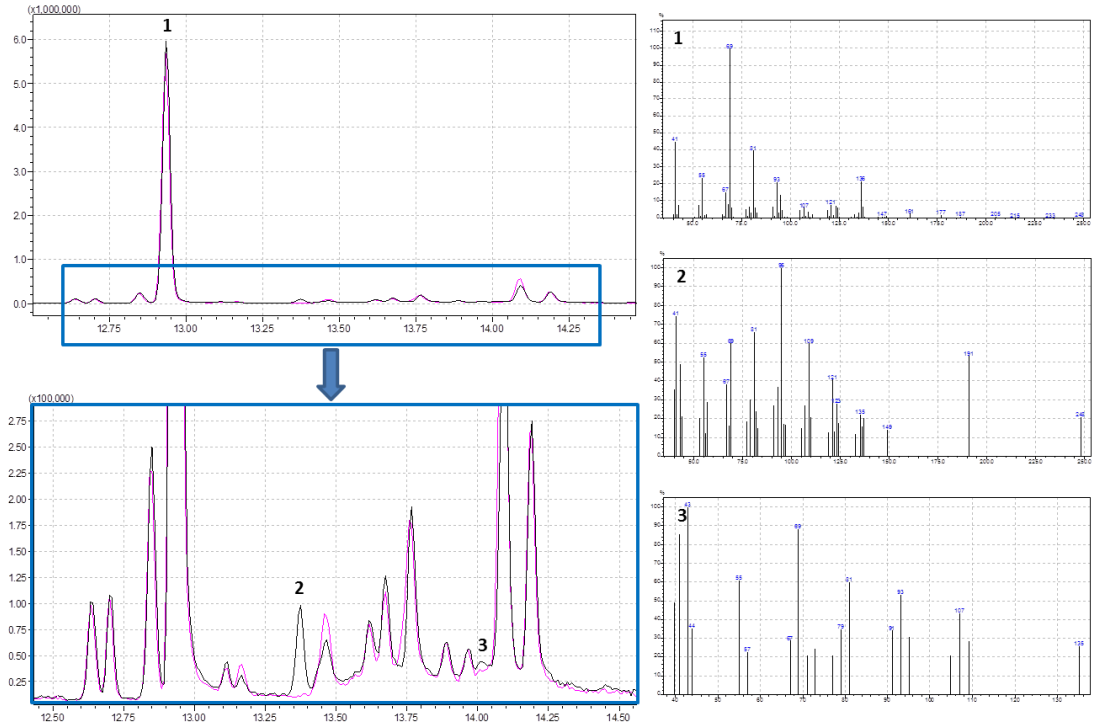
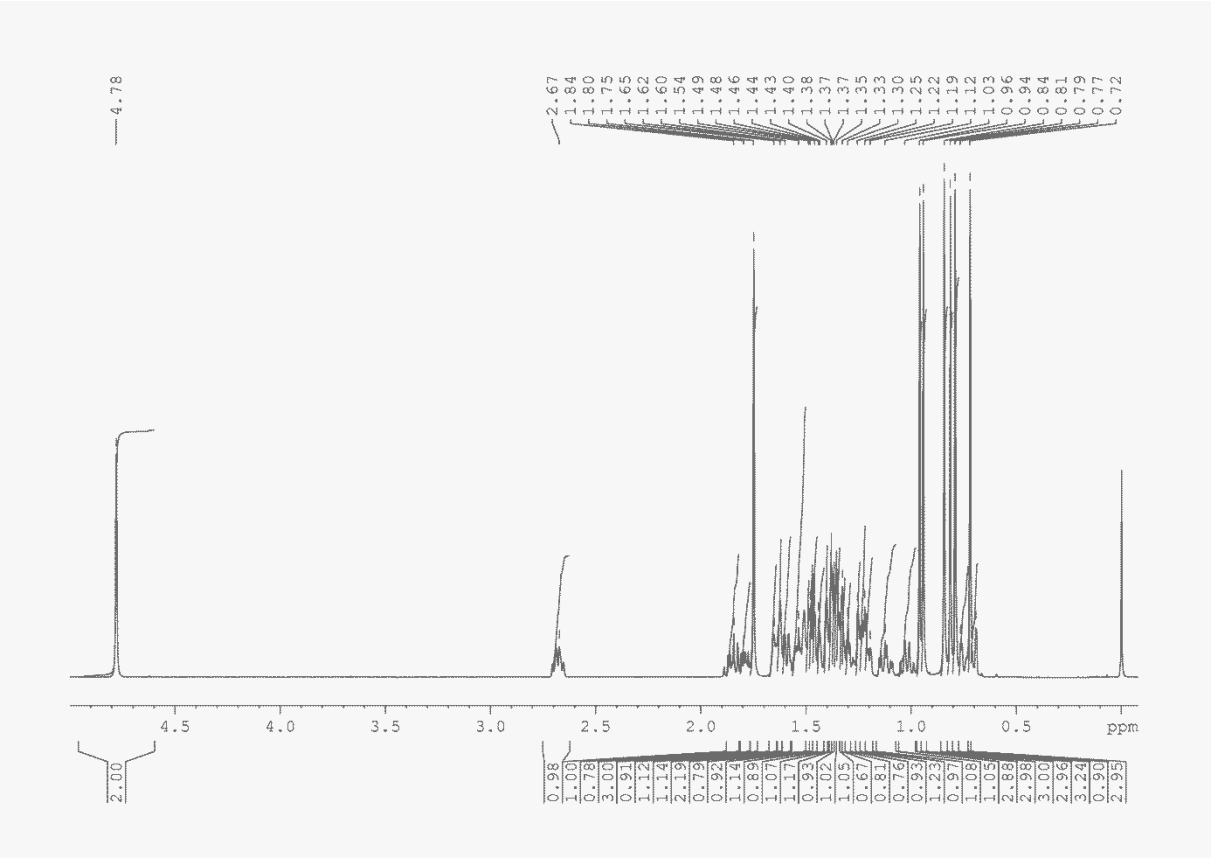


Fig. 4.21: GC-MS chromatograms of the conversion of bishomofarnesal with *Zmo*SHC1 (black) and negative control (pink), the GC-MS spectra of the substrate (1), the major product (2) and the minor product (3) are shown on the right side.

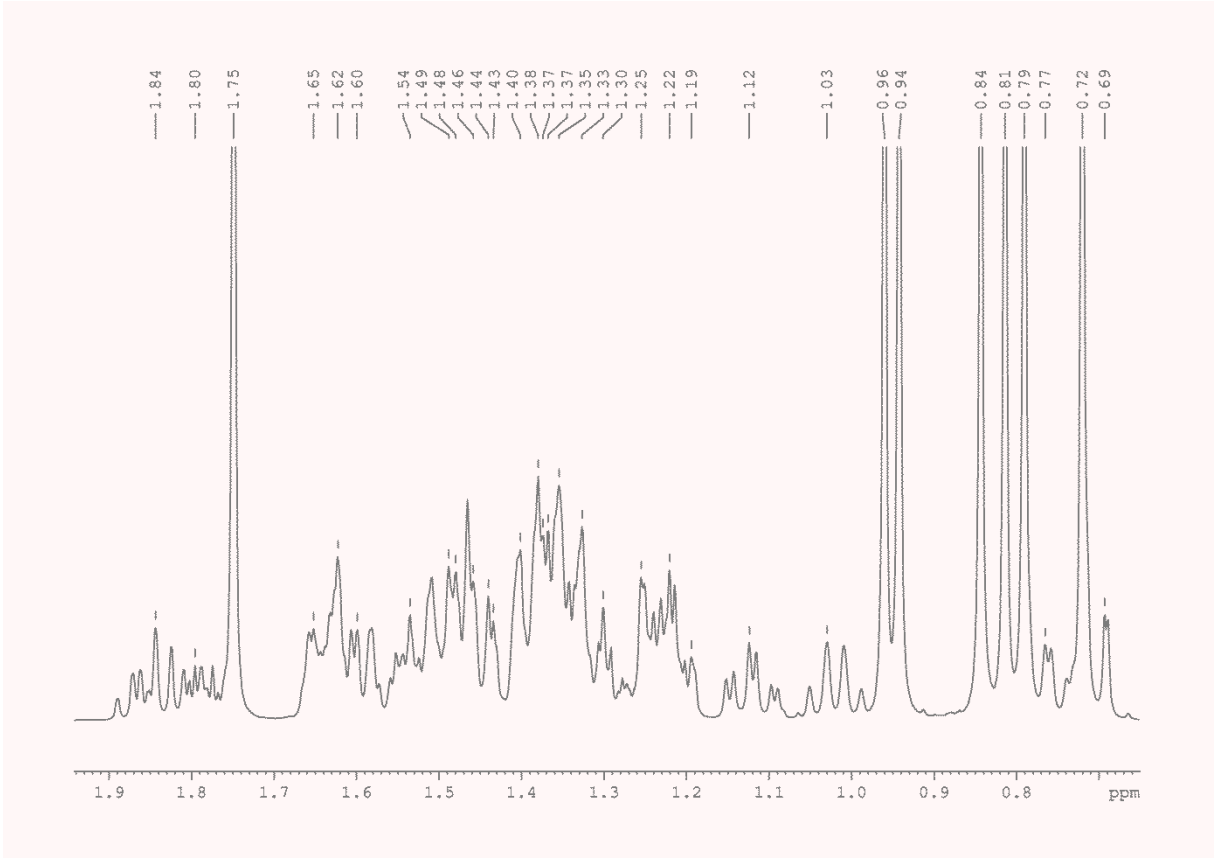
5 NMR spectra

5.1 NMR spectra hopene

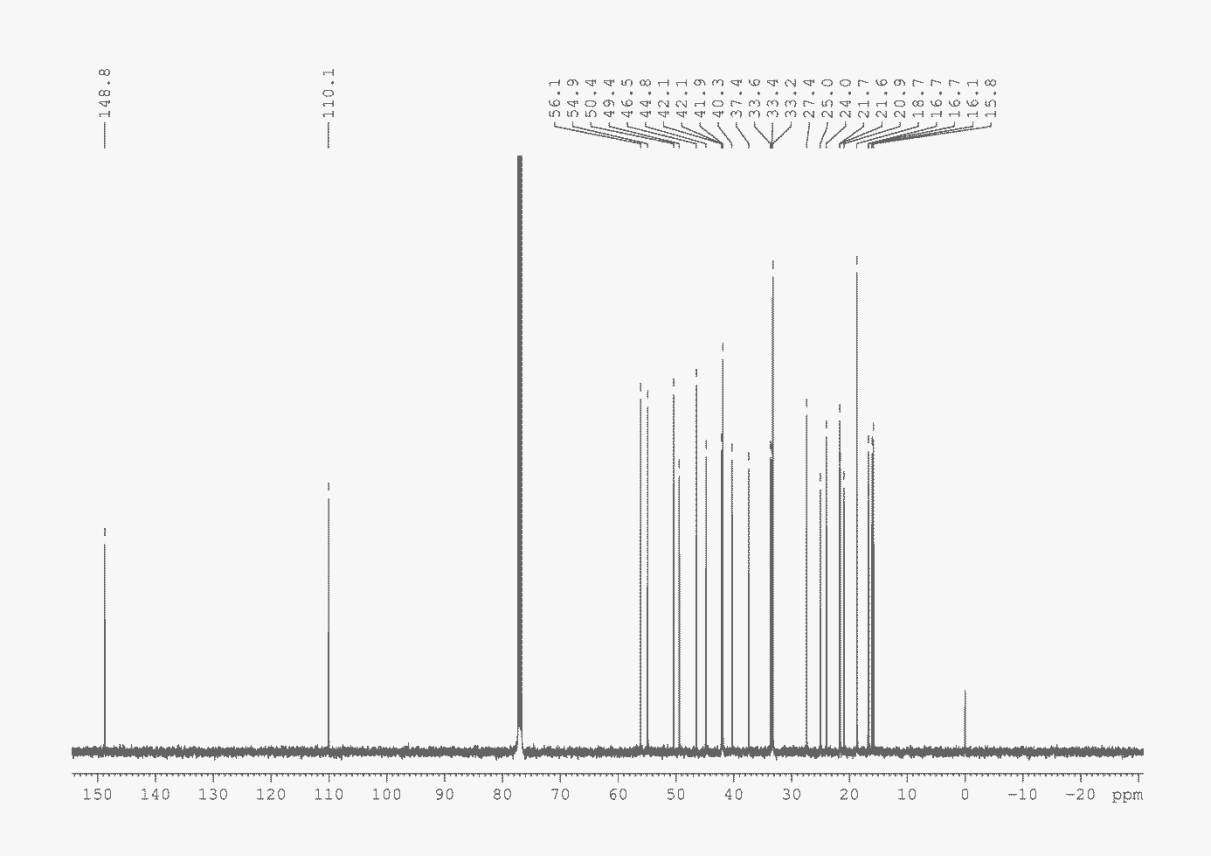
5.1.1 ^{1}H NMR Spectrum hopene



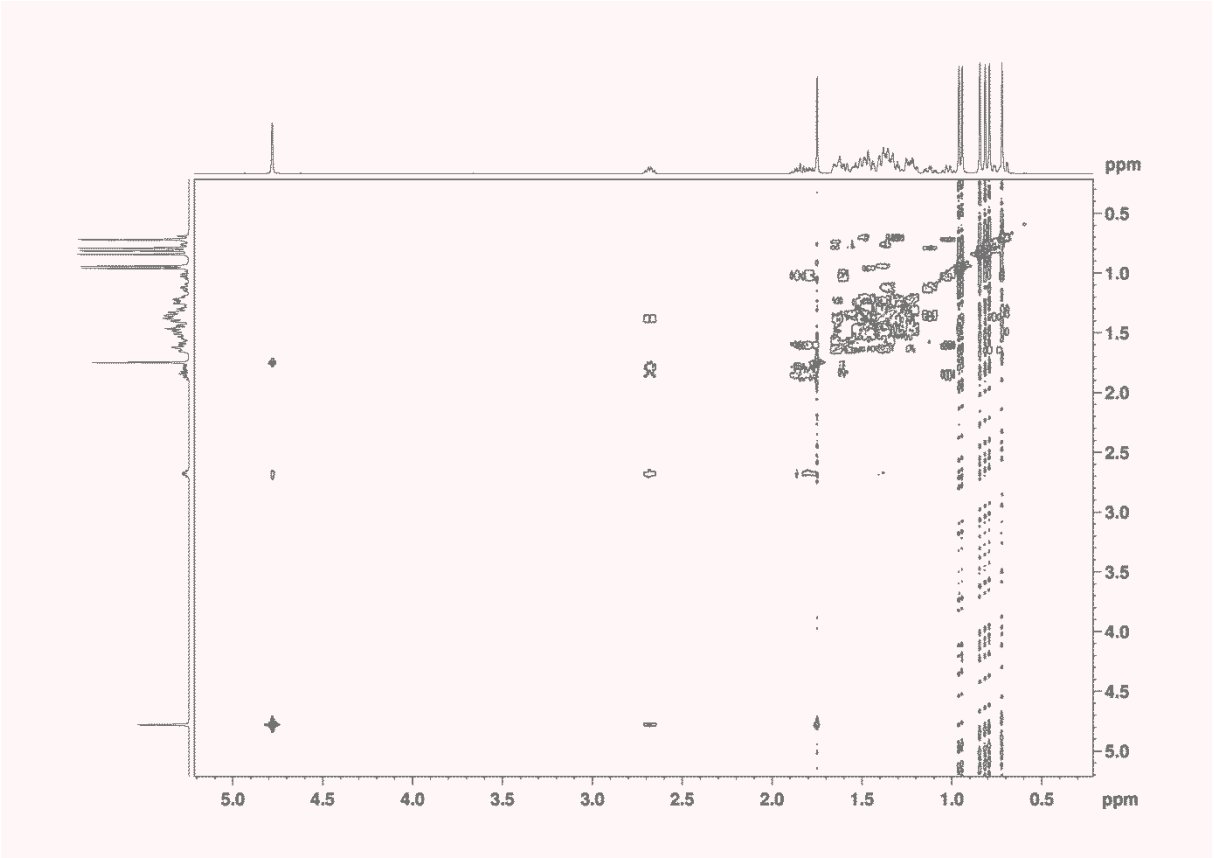
5.1.2 ^{1}H NMR spectrum hopene zoomed



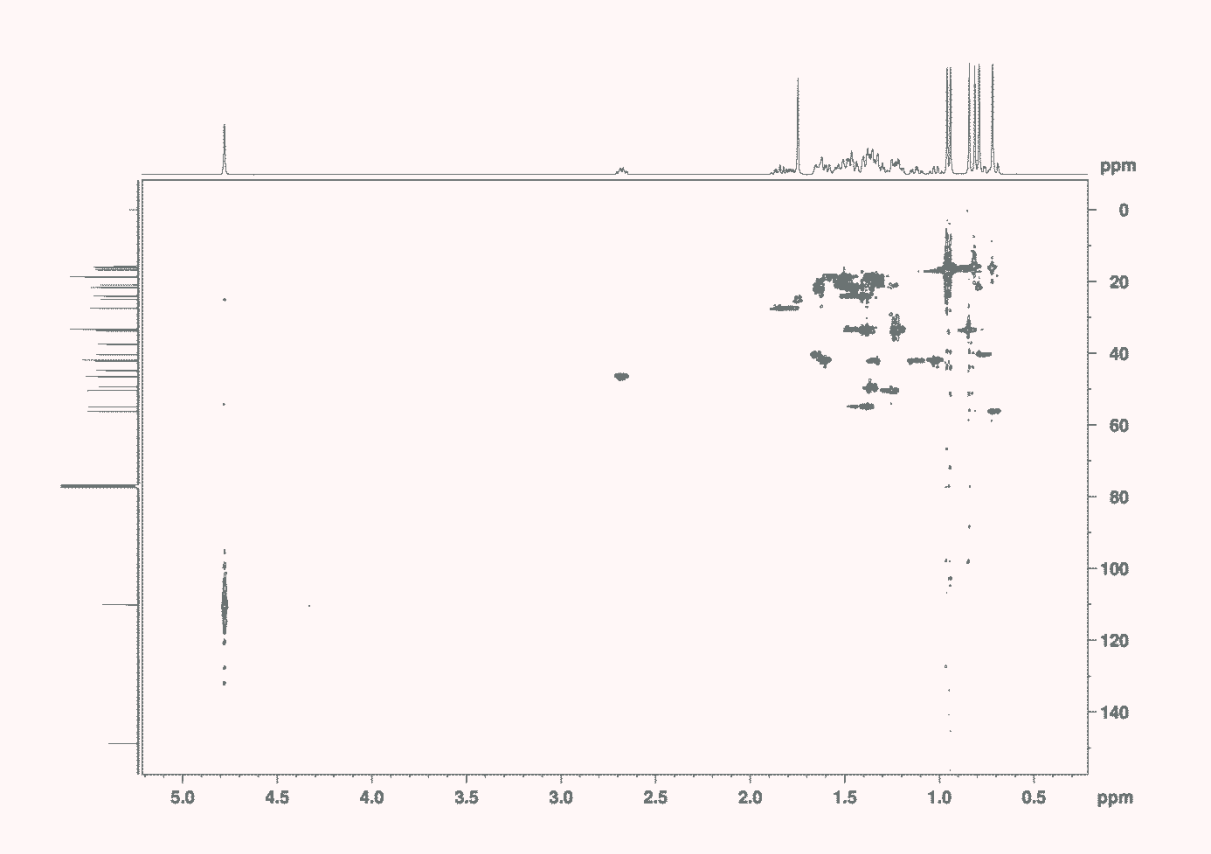
5.1.3 ^{13}C NMR spectrum hopene



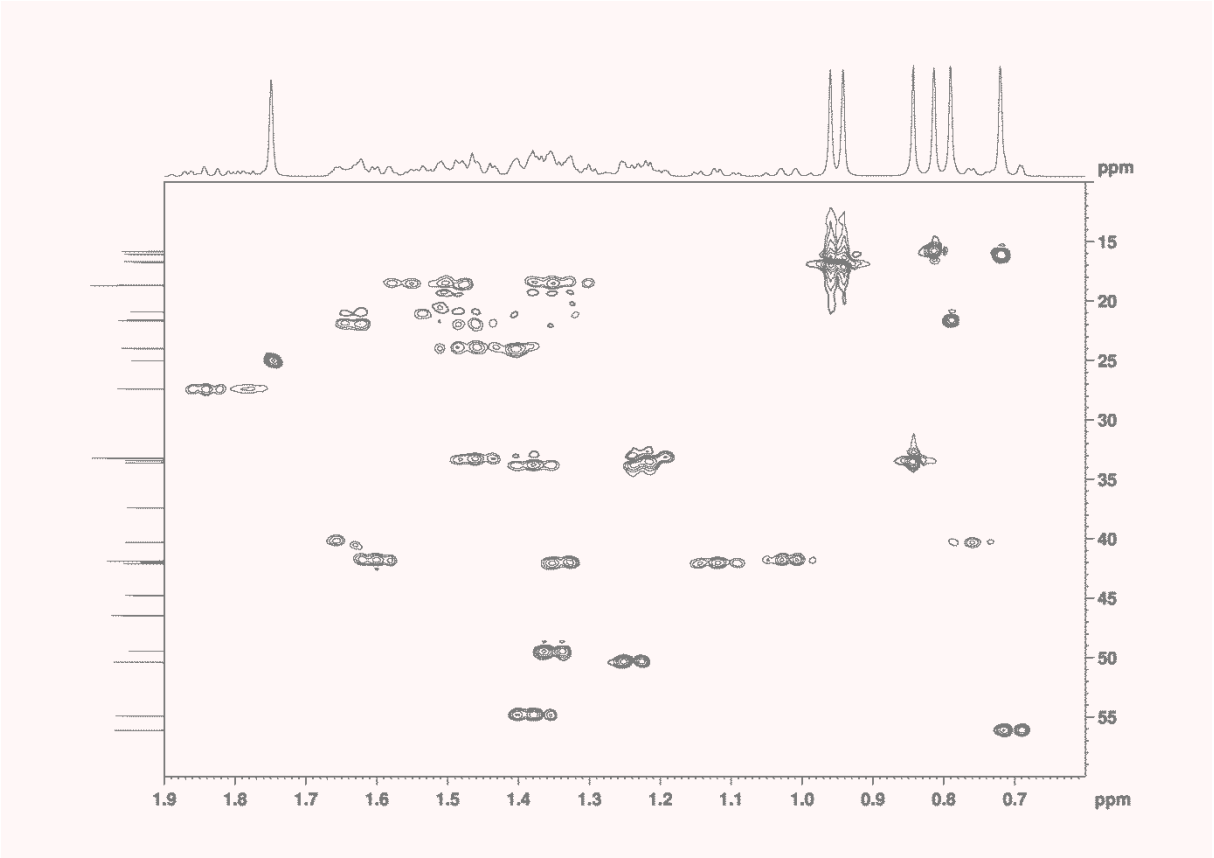
5.1.4 COSY NMR spectrum hopene



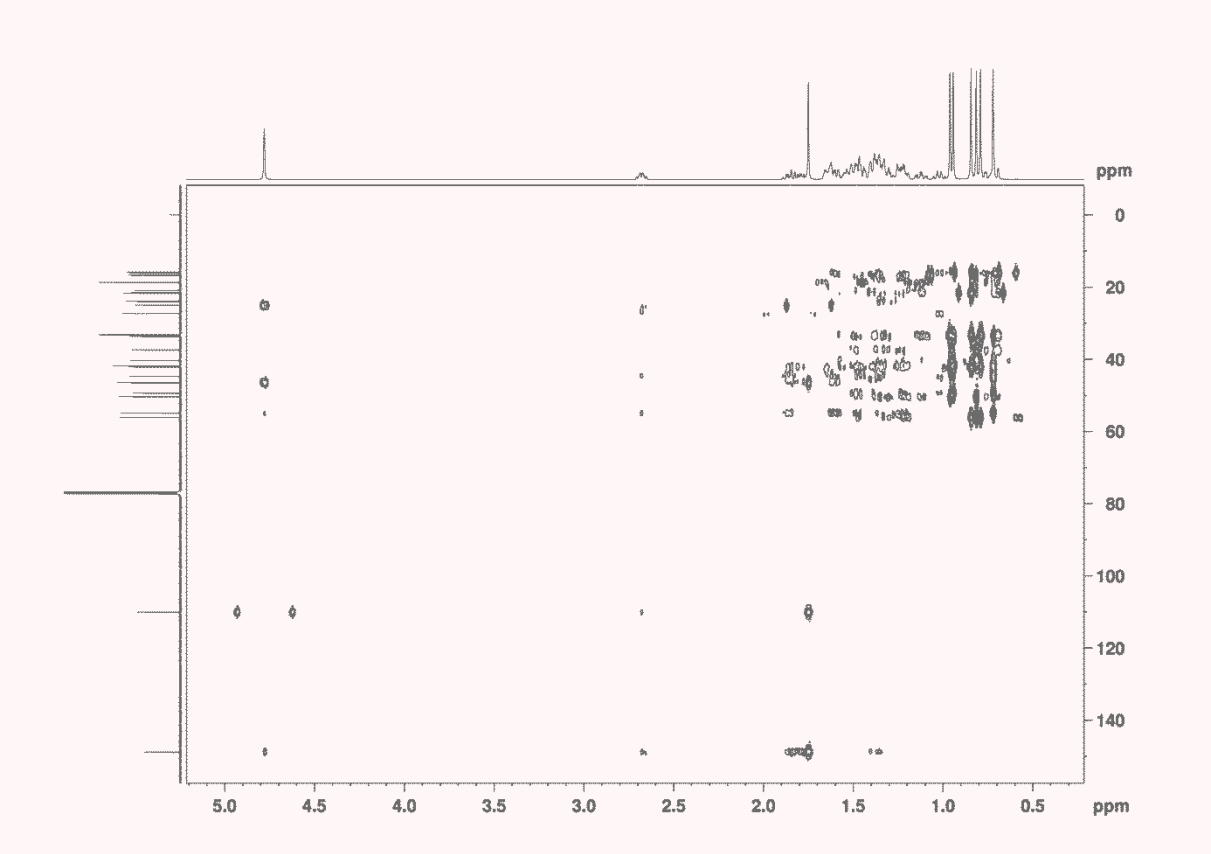
5.1.5 HSQC NMR spectrum hopene



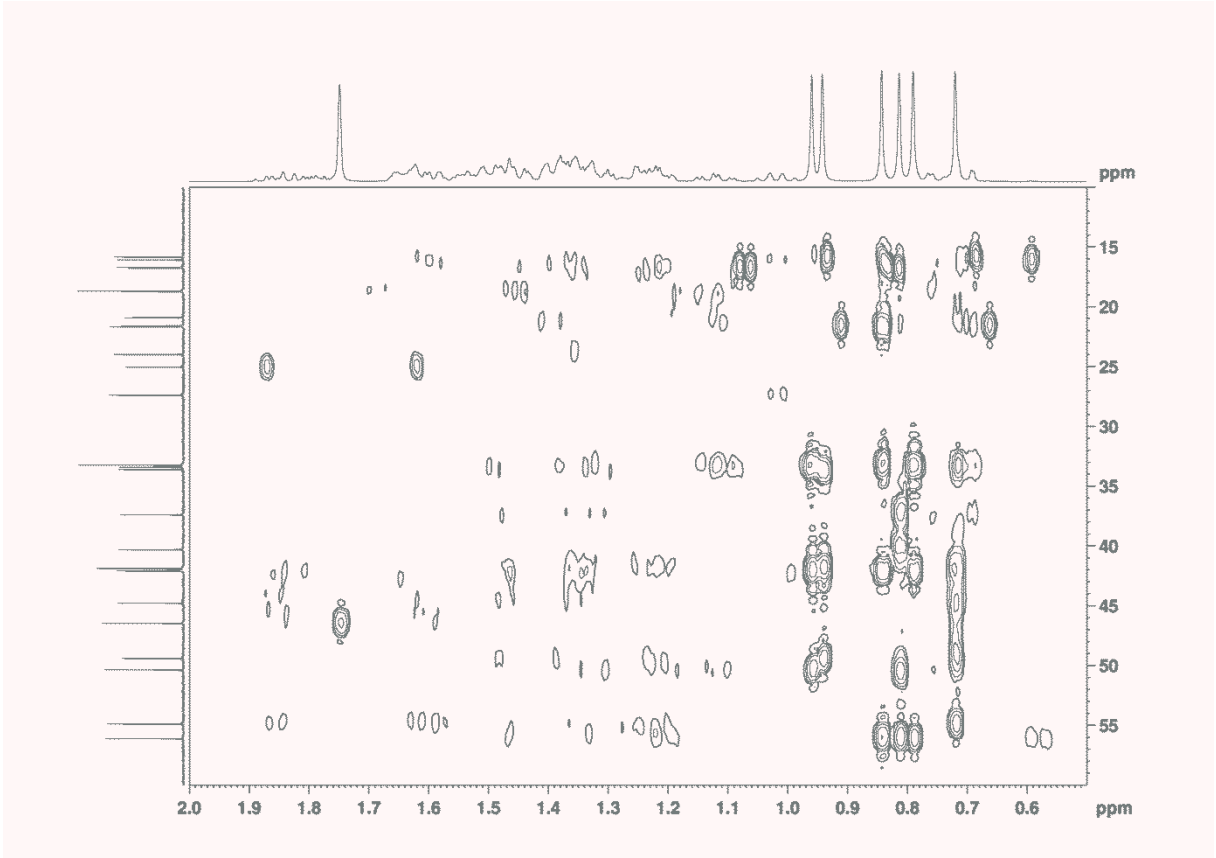
5.1.6 HSQC NMR spectrum hopene zoomed



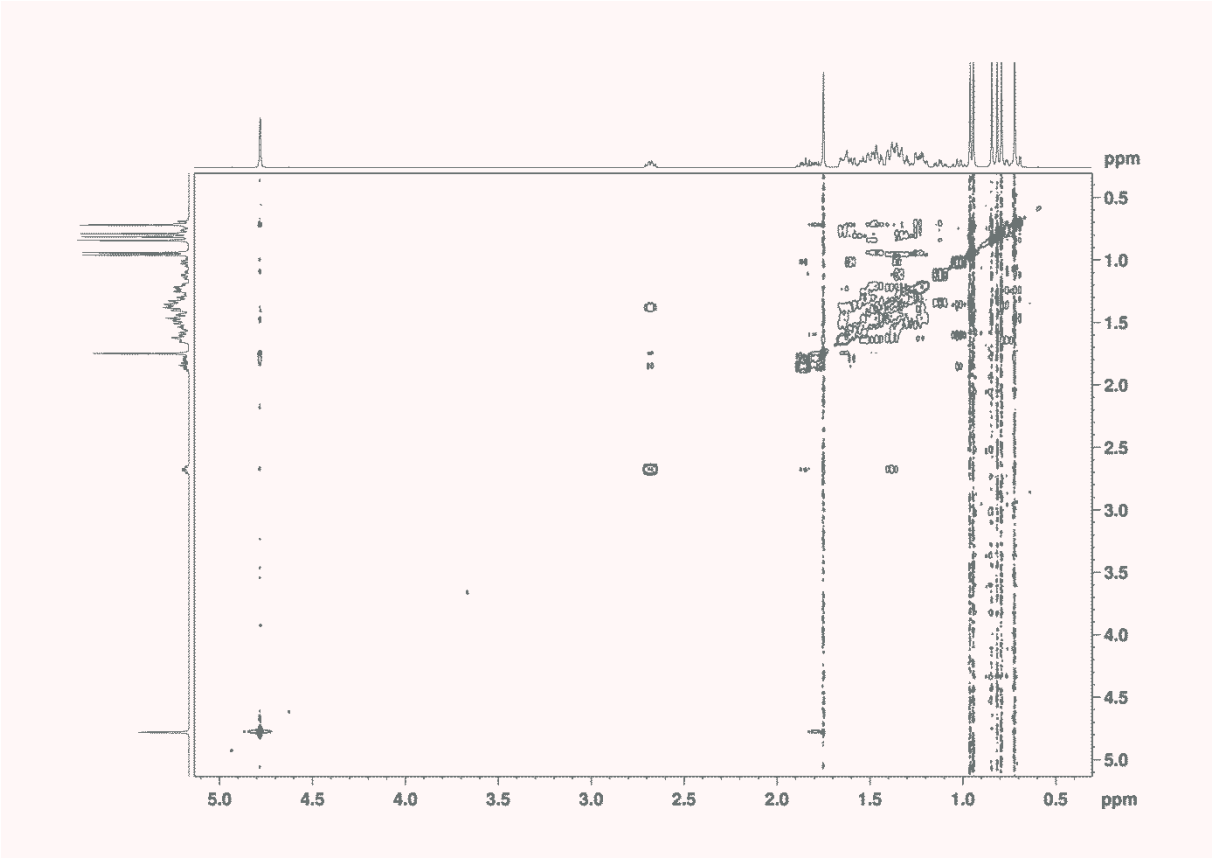
5.1.7 HMBC NMR spectrum hopene



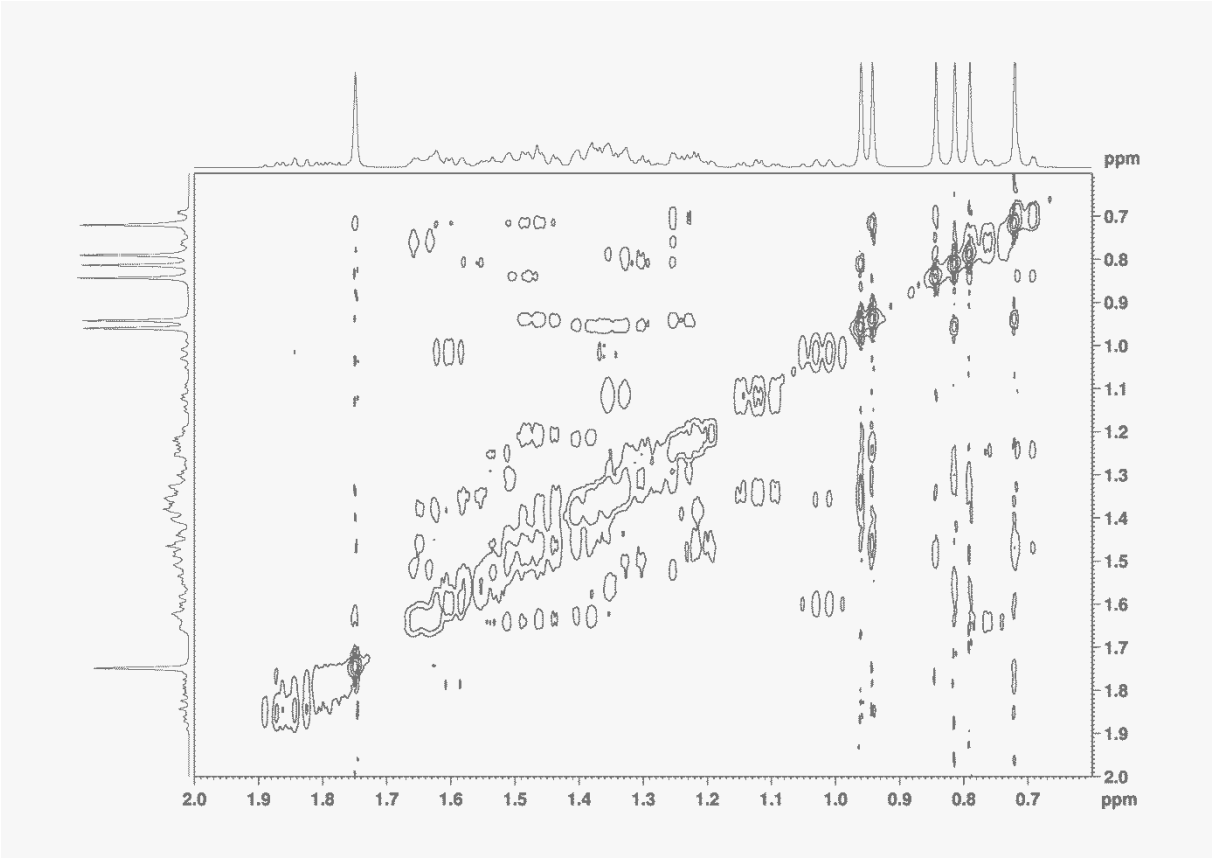
5.1.8 HMBC NMR spectrum hopene zoomed



5.1.9 NOESY NMR spectrum hopene

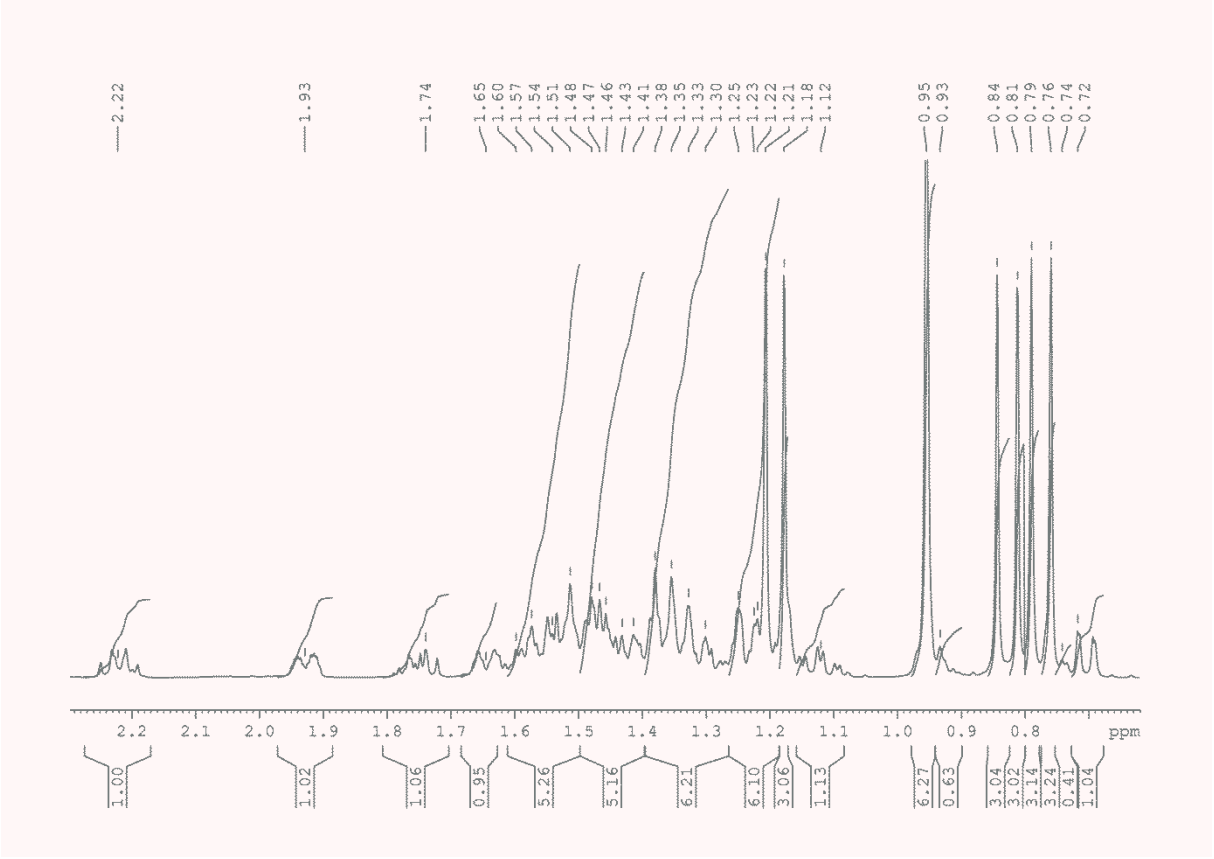


5.1.10NOESY NMR spectrum hopene zoomed

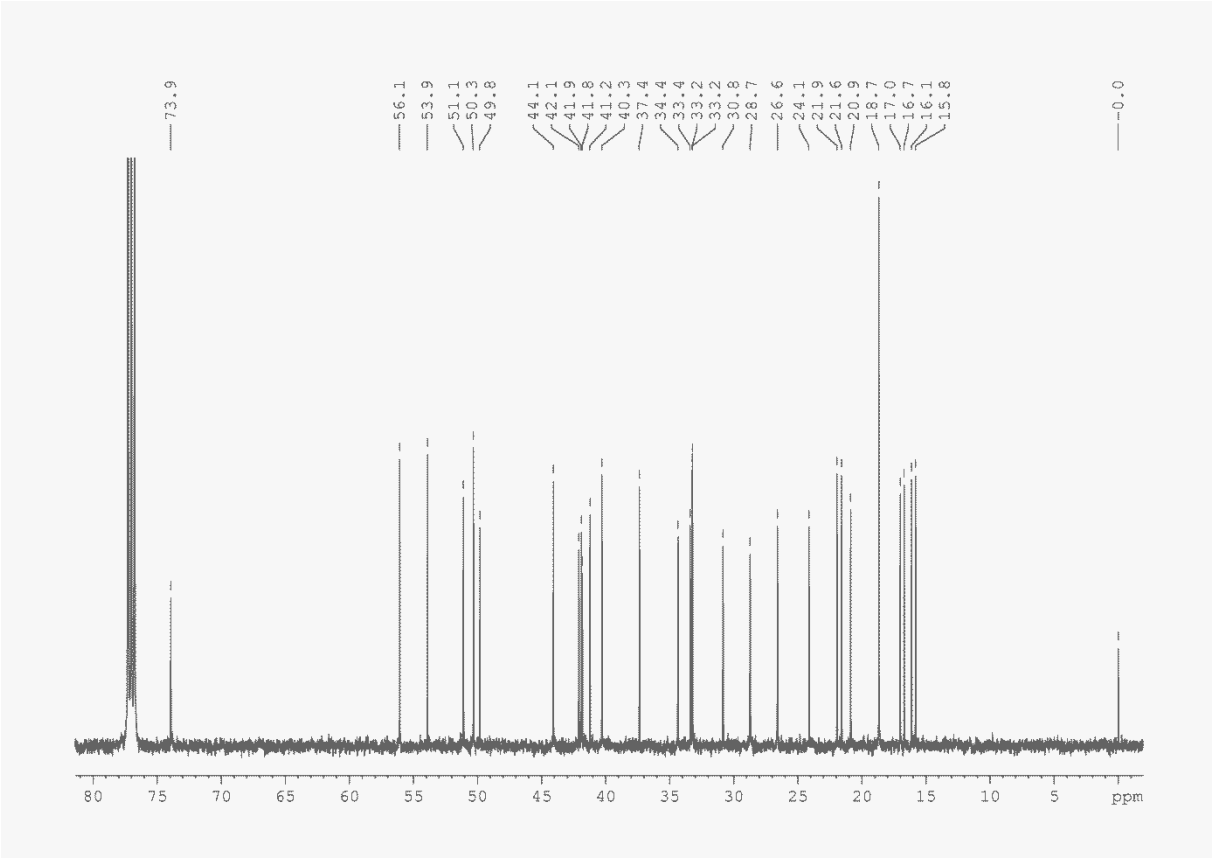


5.2 NMR spectra hopanol

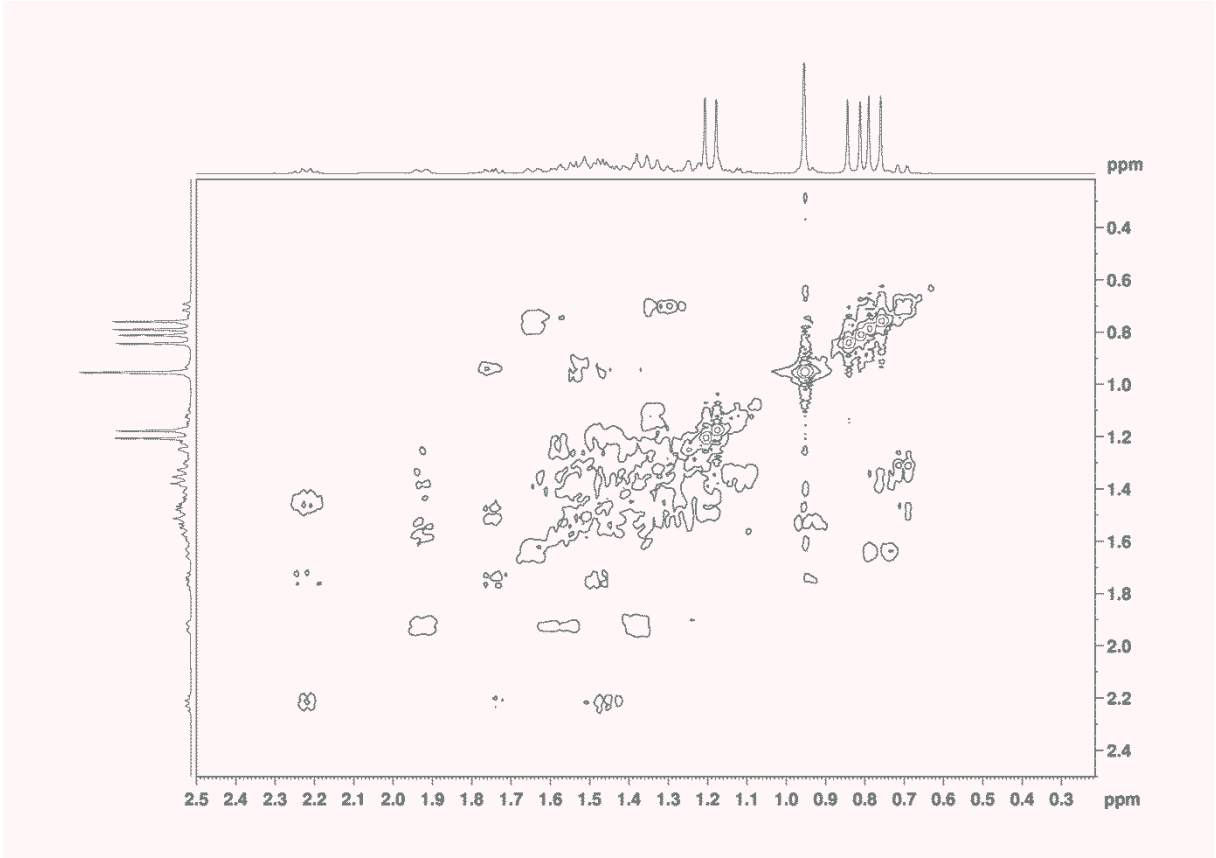
5.2.1 ^{1}H NMR spectrum hopanol



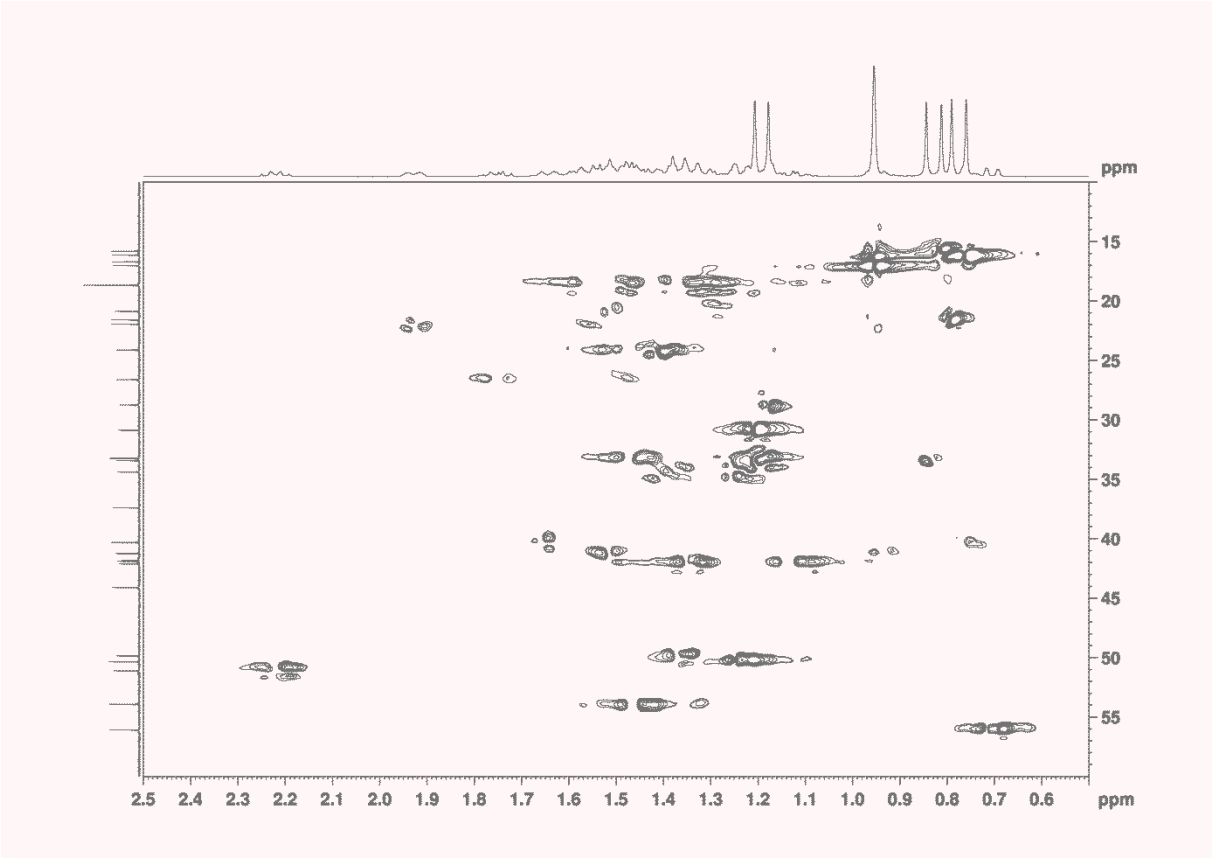
5.2.2 ^{13}C NMR spectrum hopanol



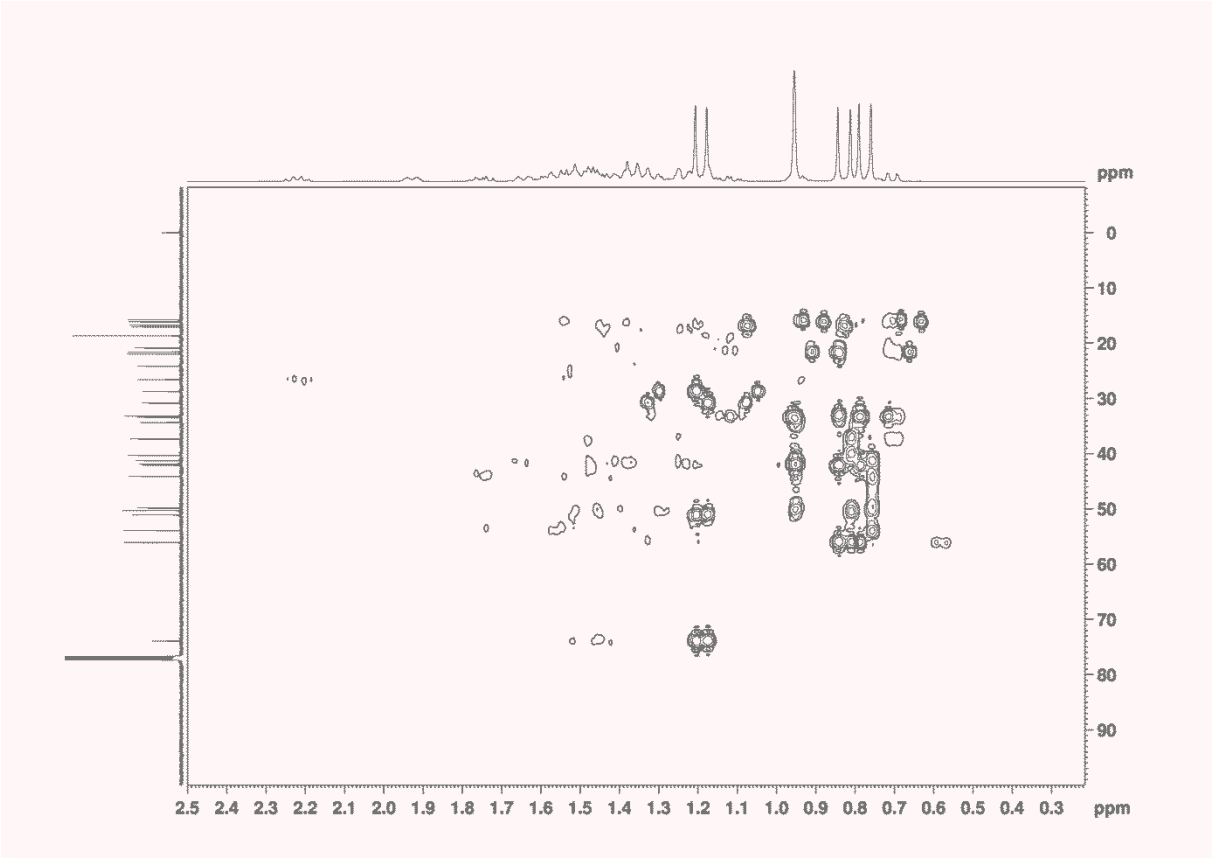
5.2.3 COSY NMR spectrum hopanol



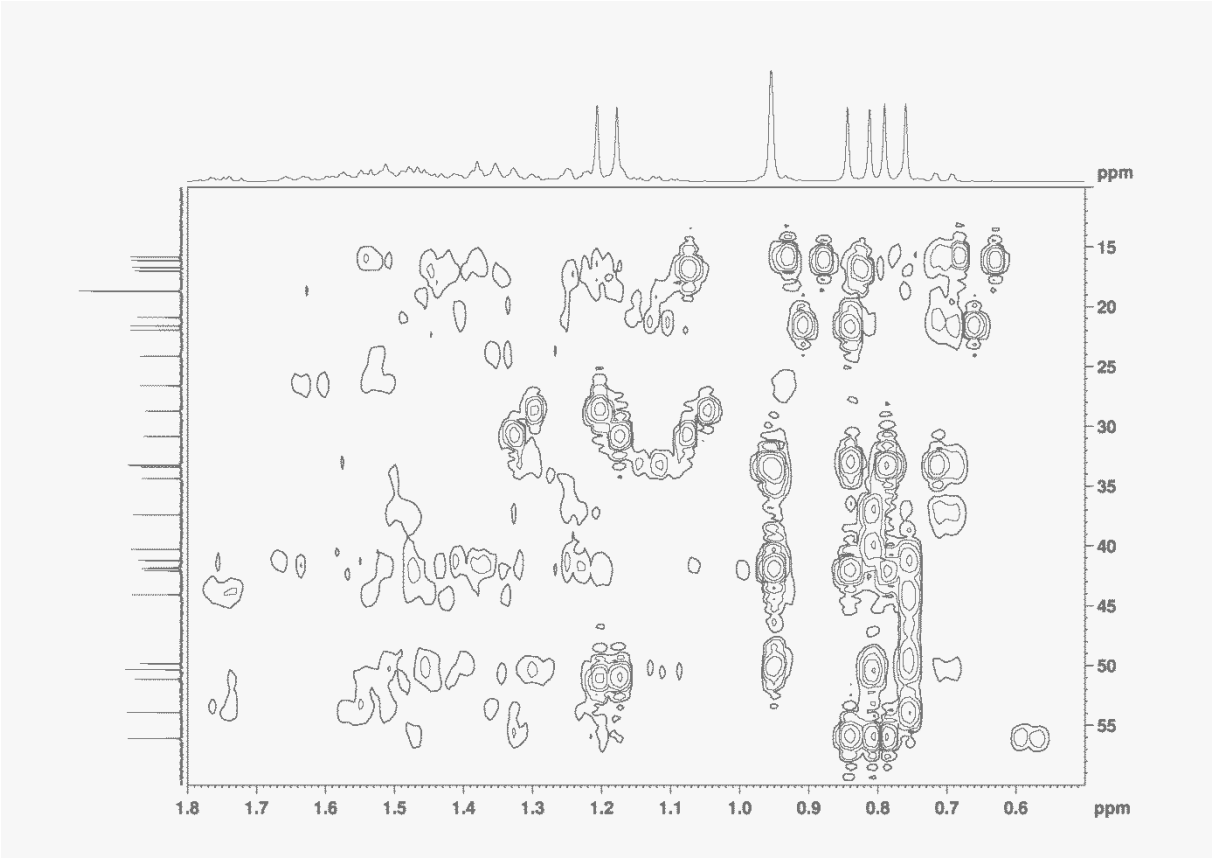
5.2.4 HSQC NMR spectrum hopanol



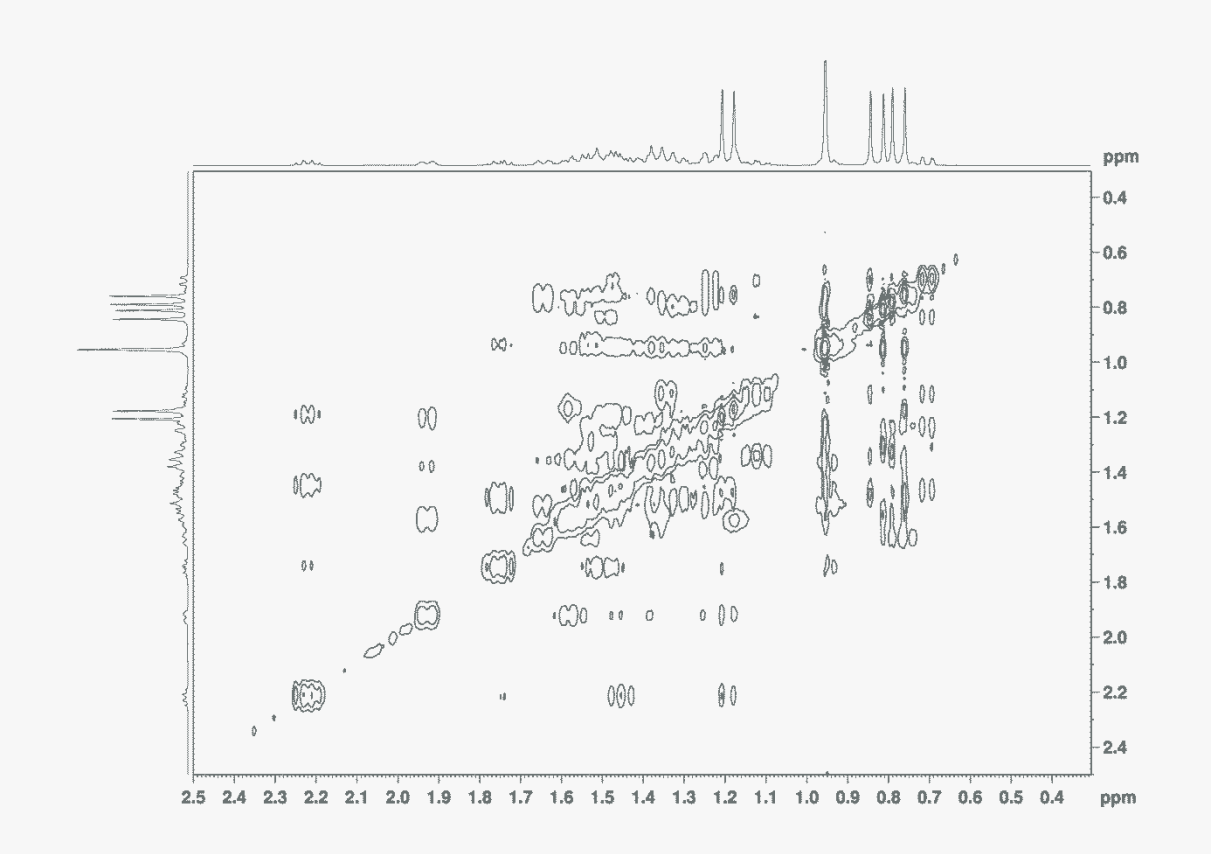
5.2.5 HMBC NMR spectrum hopanol



5.2.6 HMBC NMR spectrum hopanol zoomed

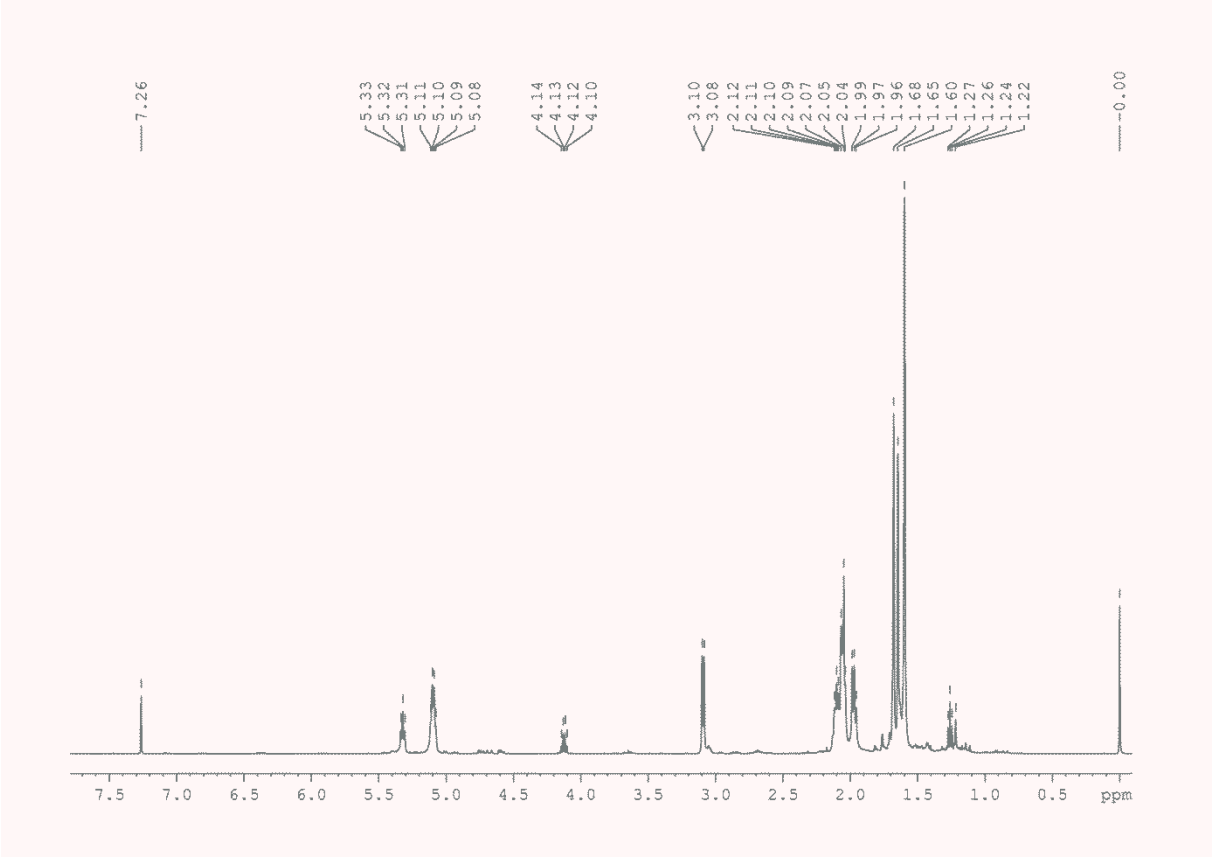


5.2.7 NOESY NMR spectrum hopanol

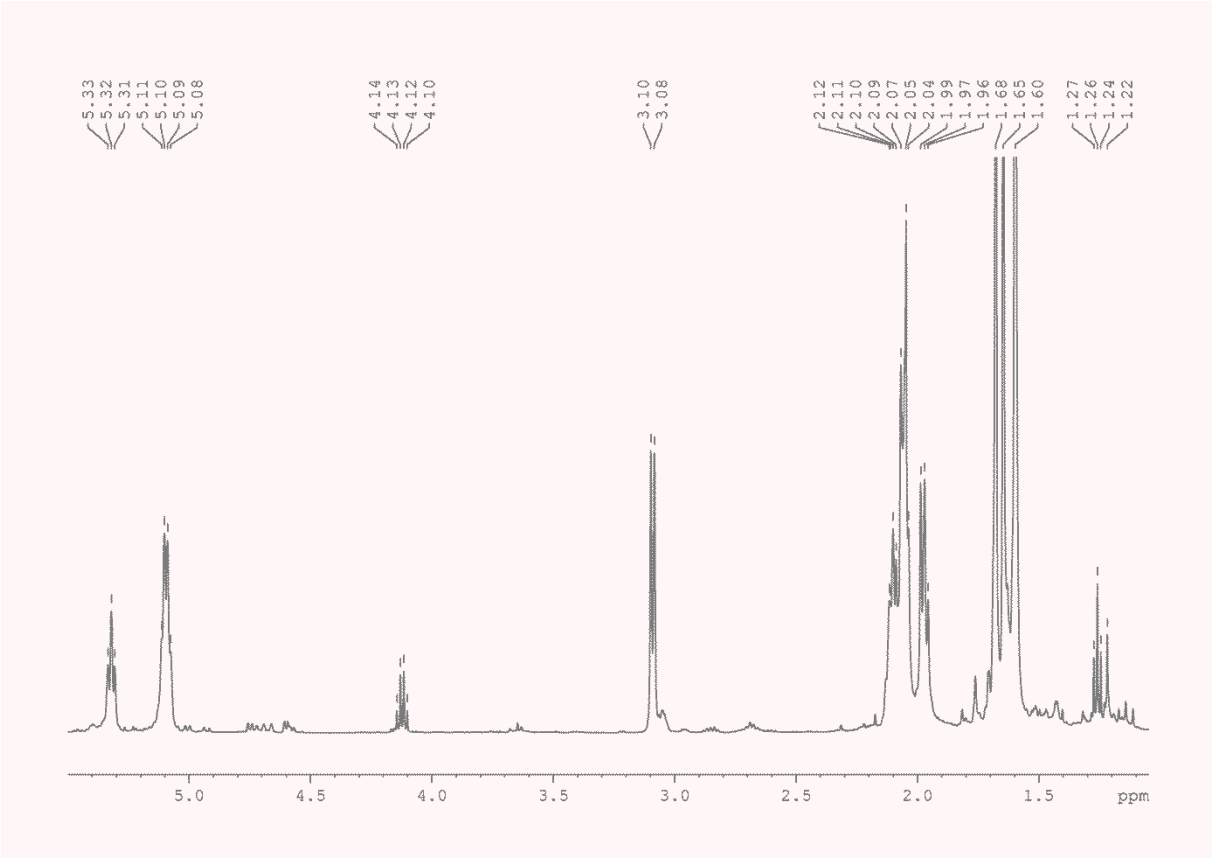


5.3 NMR spectra homofarnesoic acid

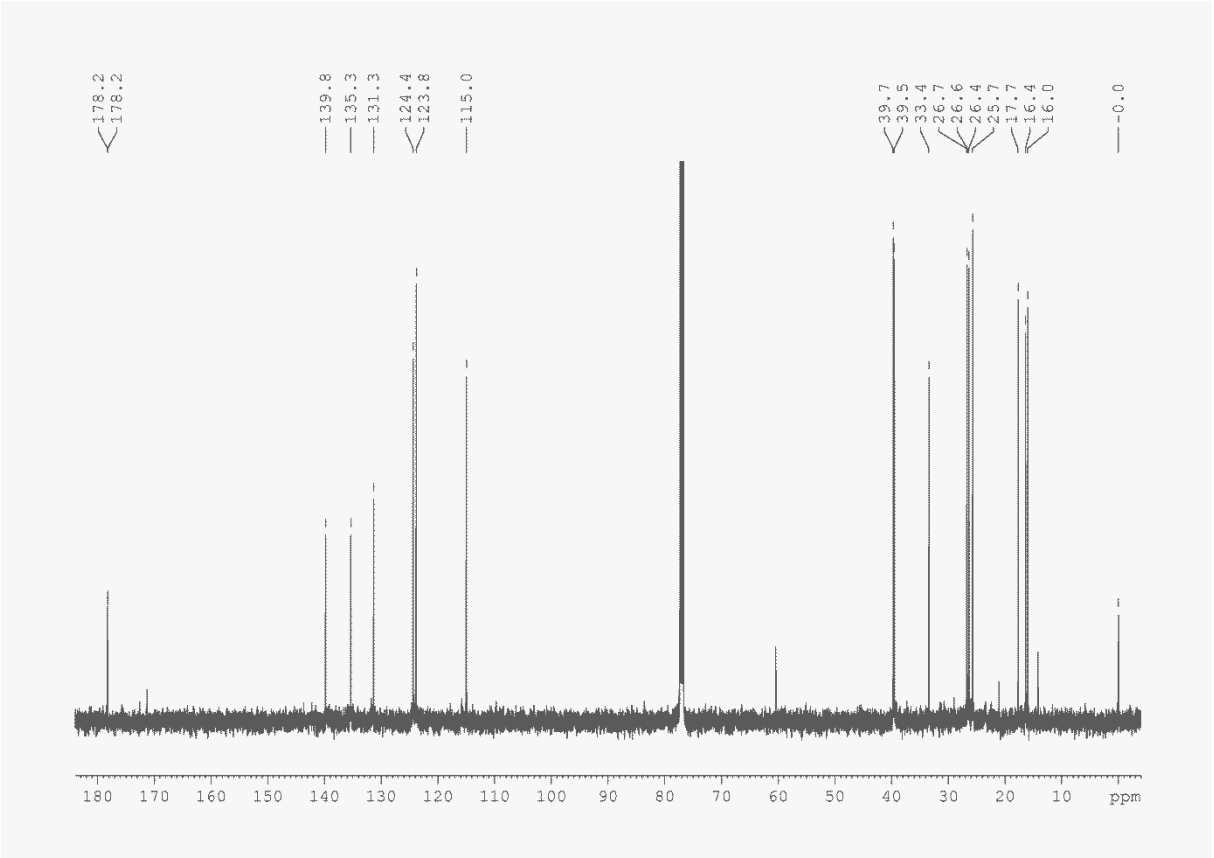
5.3.1 ^{1}H NMR spectrum homofarnesoic acid



5.3.2 ^{1}H NMR spectrum homofarnesoic acid zoomed

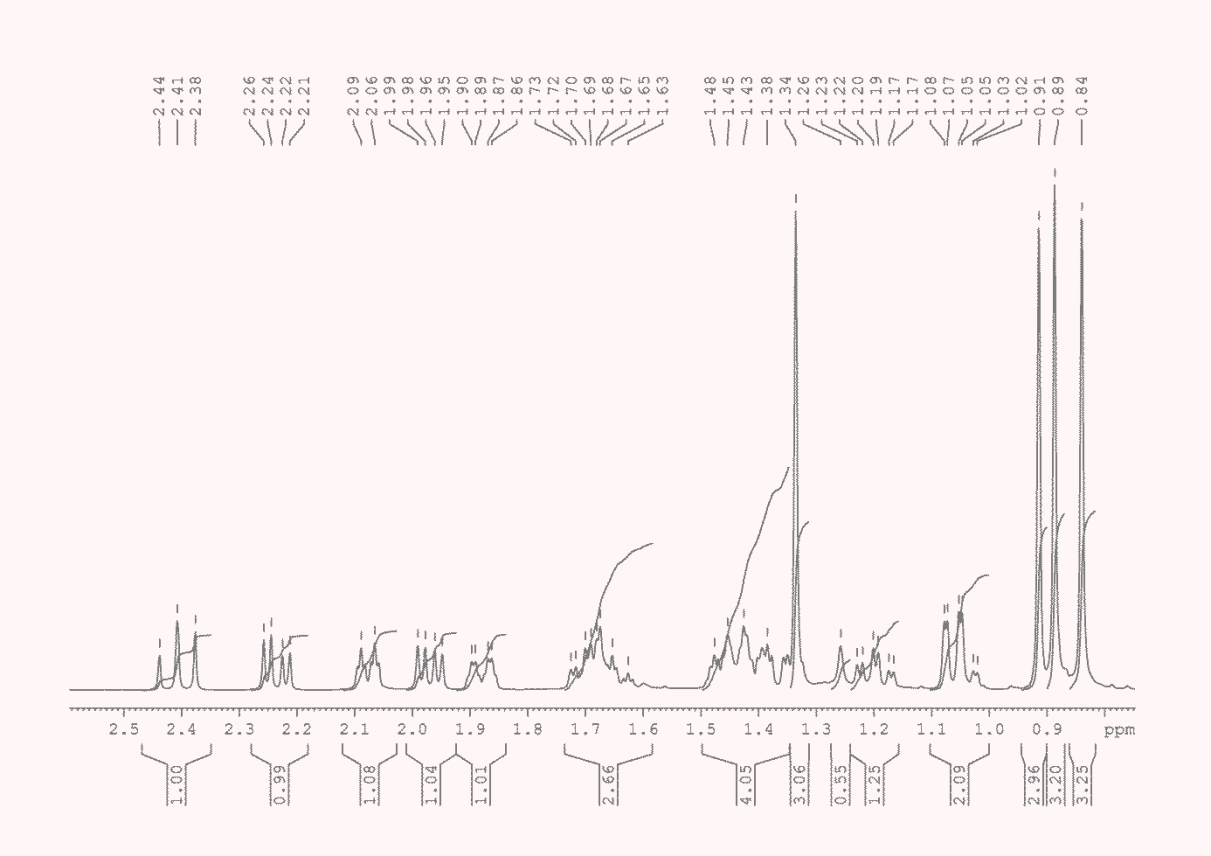


5.3.3 ^{13}C NMR spectrum homofarnesoic acid

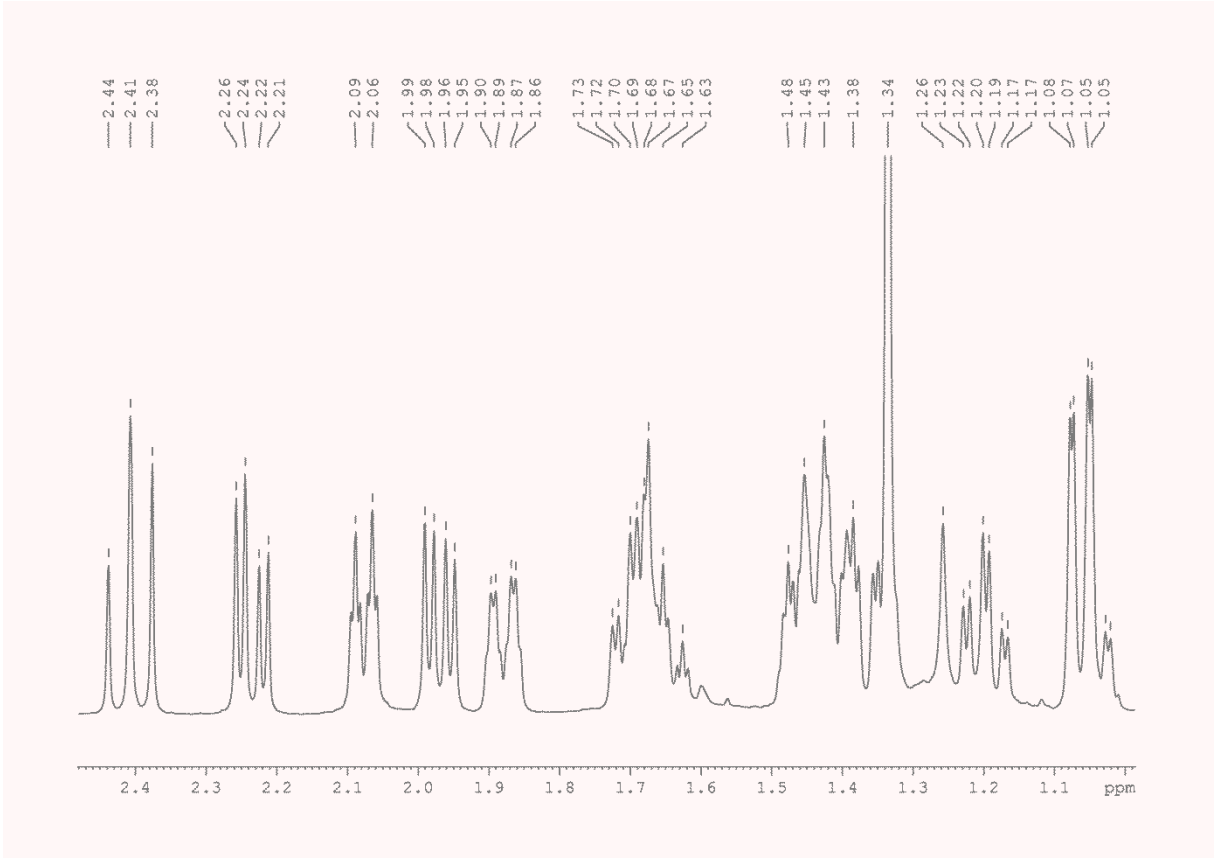


5.4 NMR spectra sclareolide

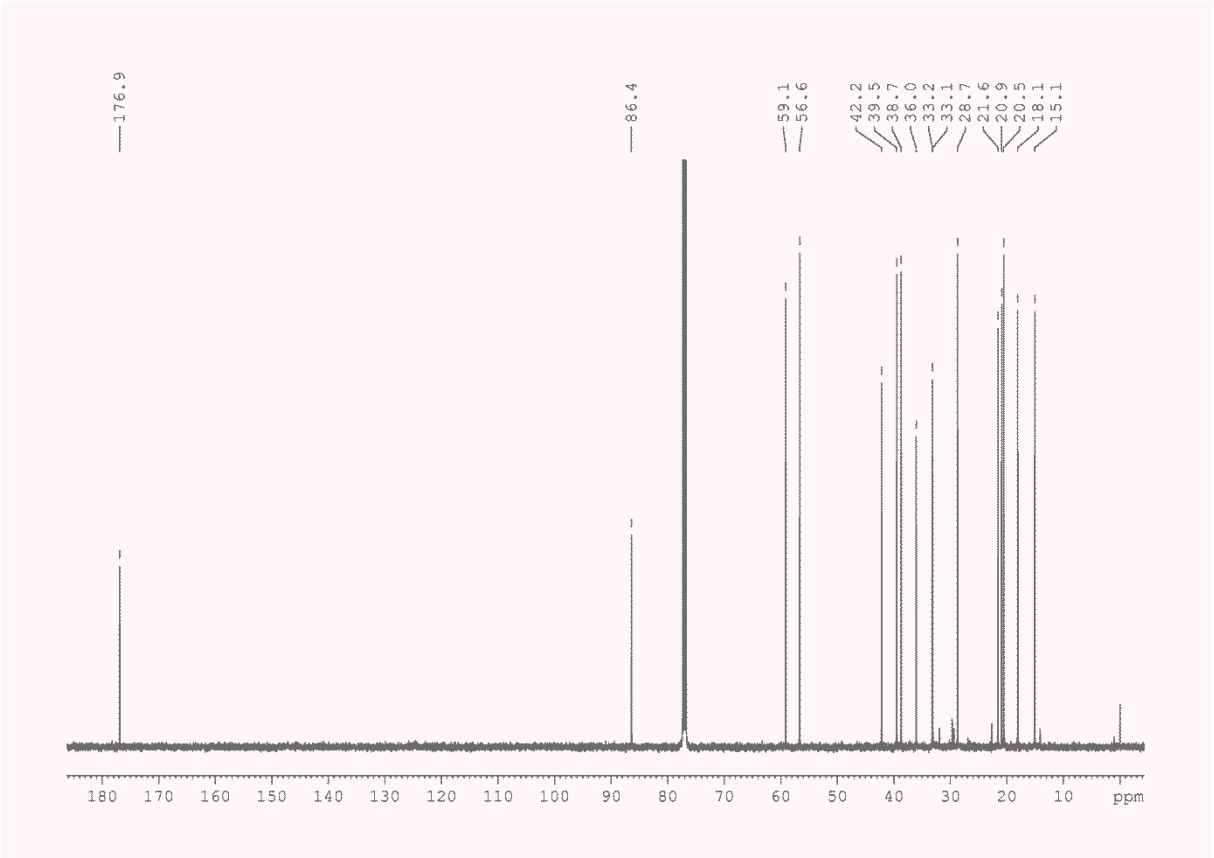
5.4.1 ^{1}H NMR spectrum sclareolide



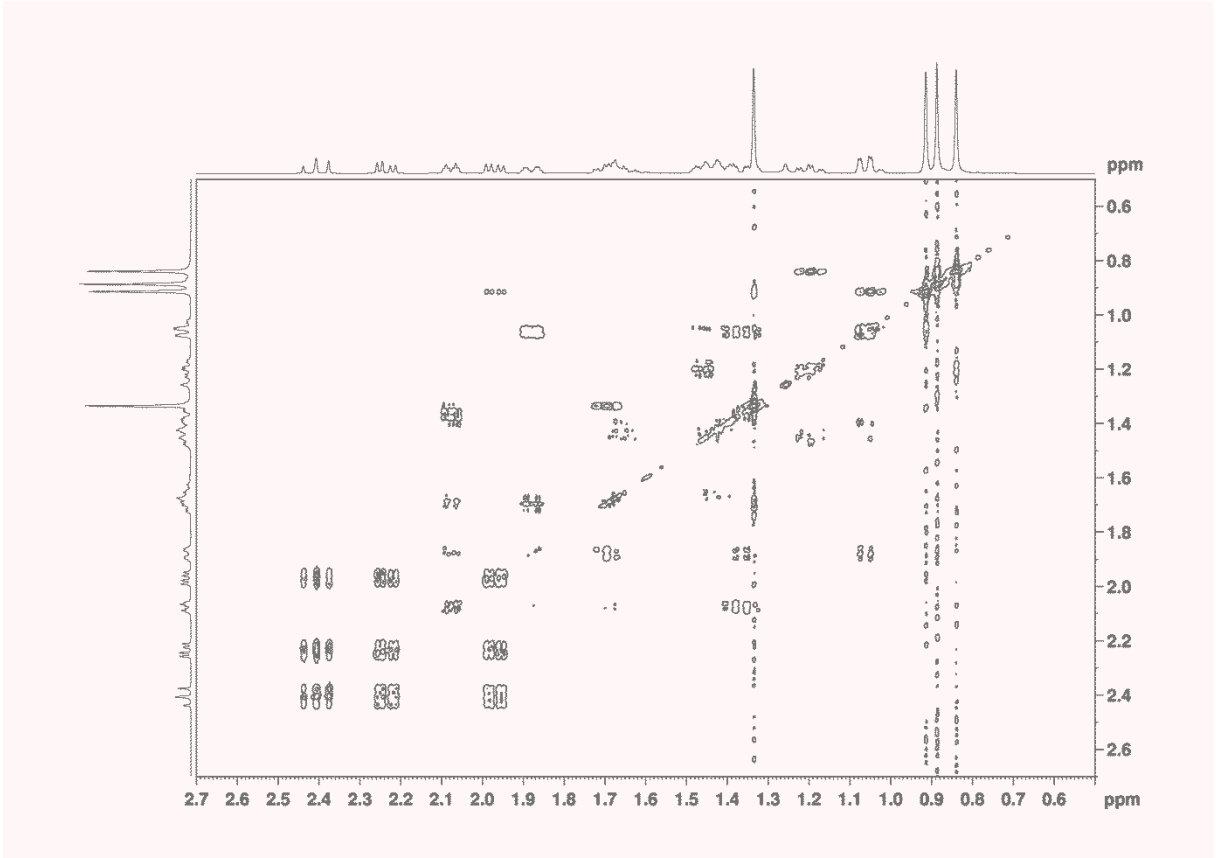
5.4.2 ^{1}H NMR spectrum sclareolide zoomed



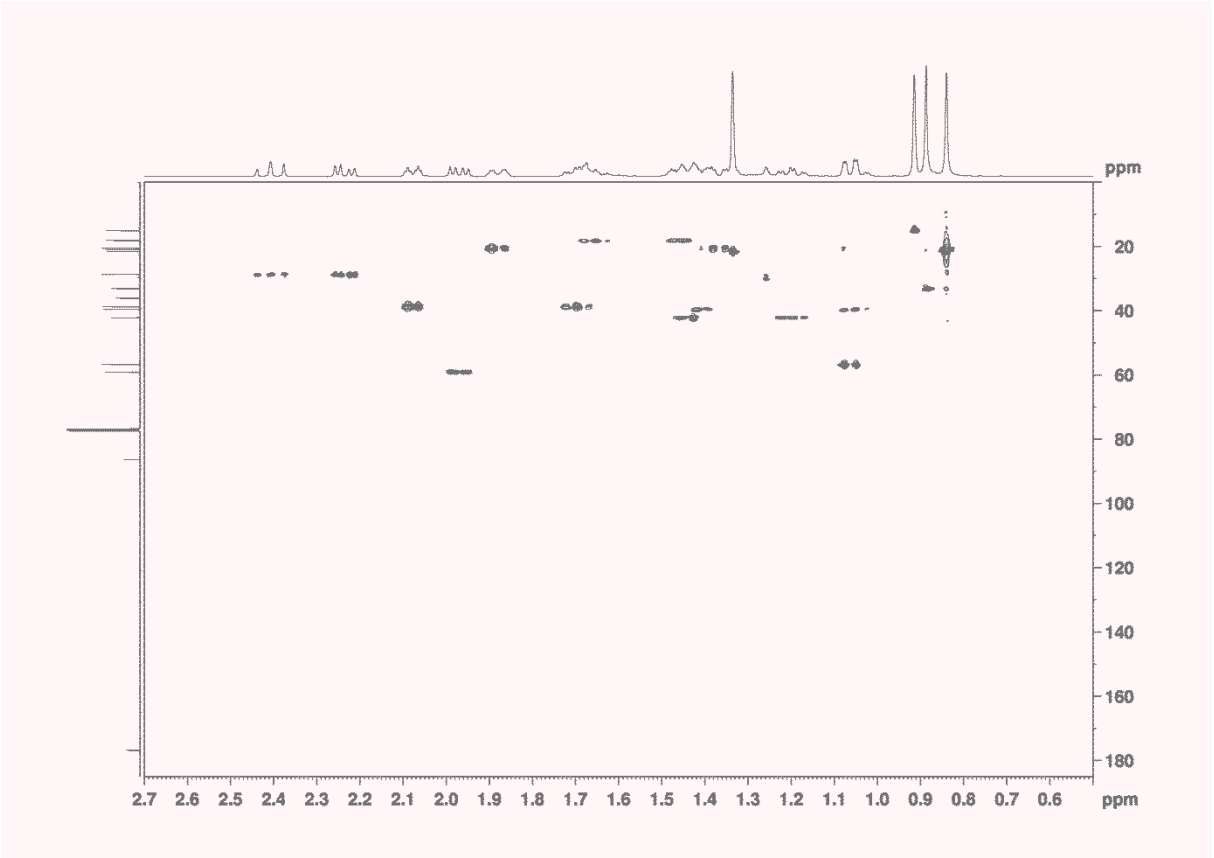
5.4.3 ^{13}C NMR spectrum sclareolide



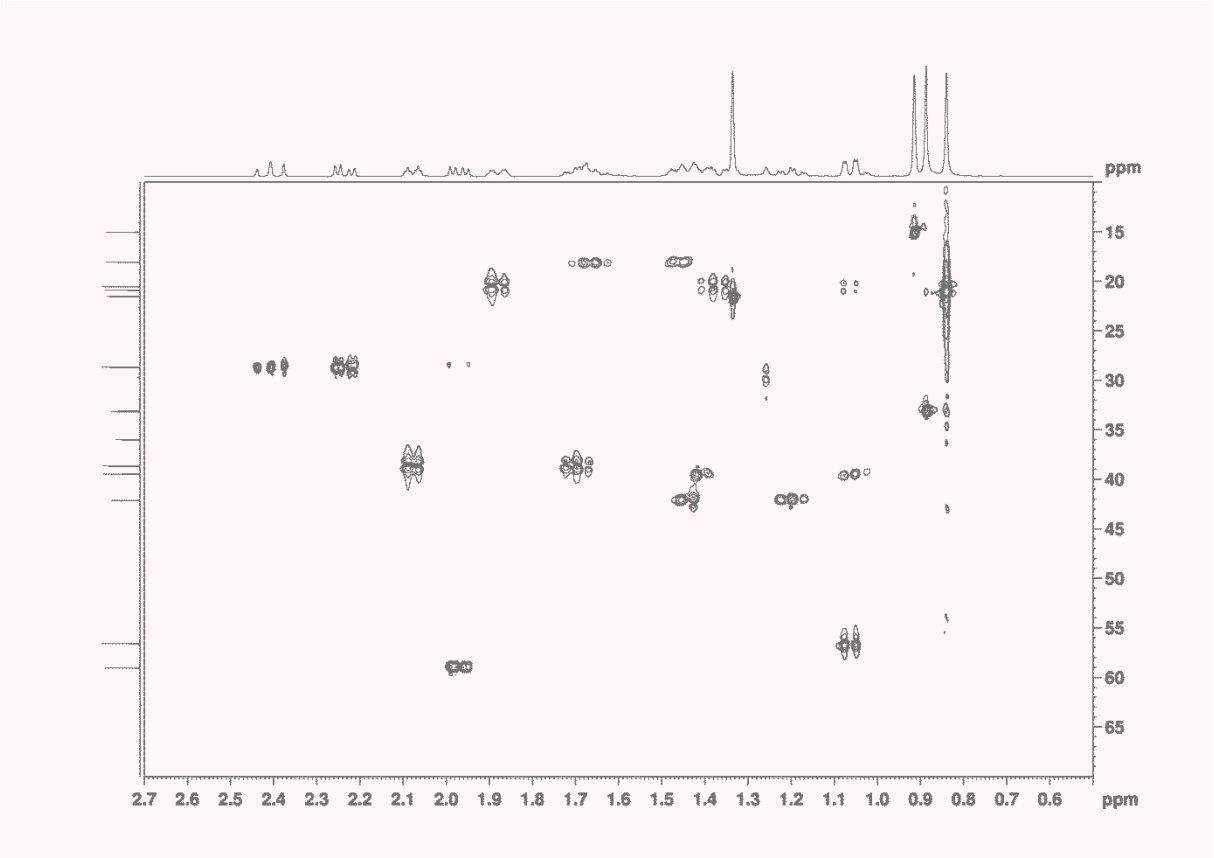
5.4.4 COSY NMR spectrum sclareolide



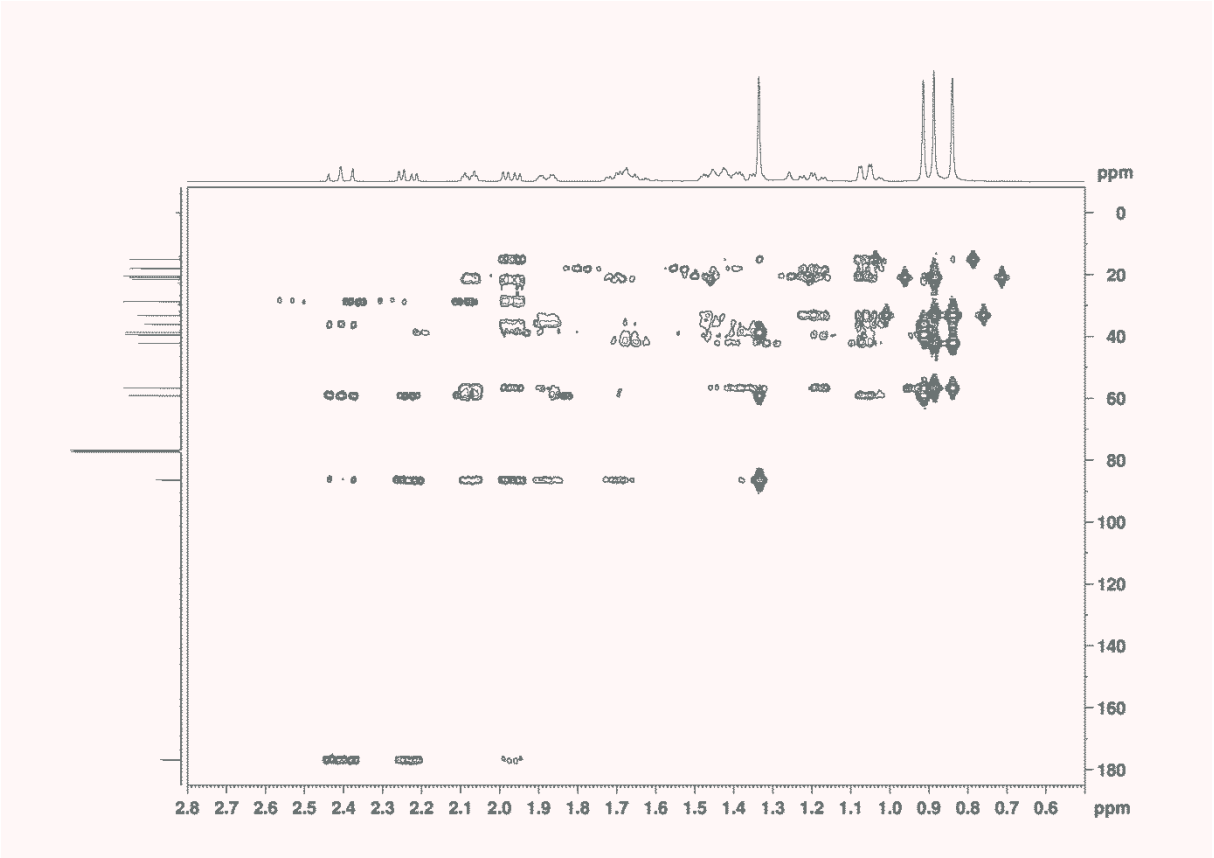
5.4.5 HSQC NMR spectrum sclareolide



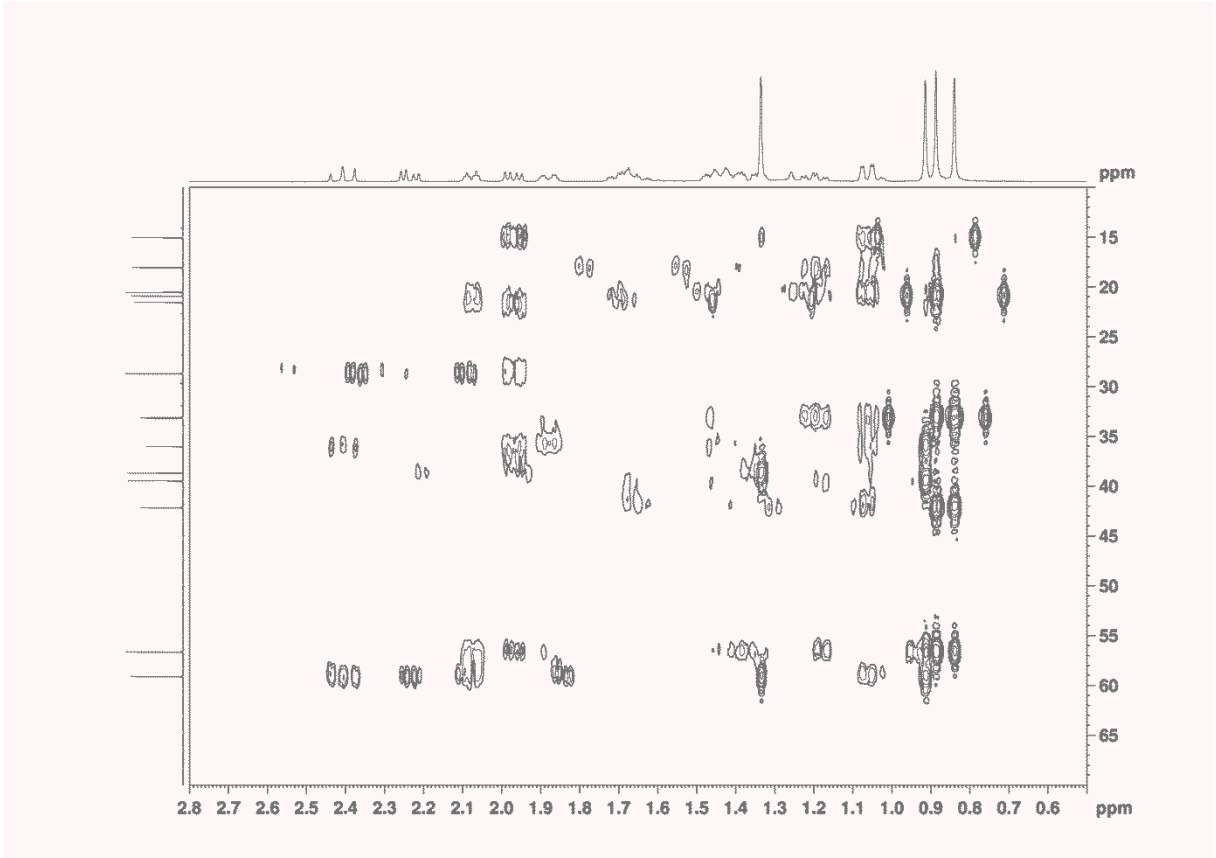
5.4.6 HSQC NMR spectrum sclareolide zoomed



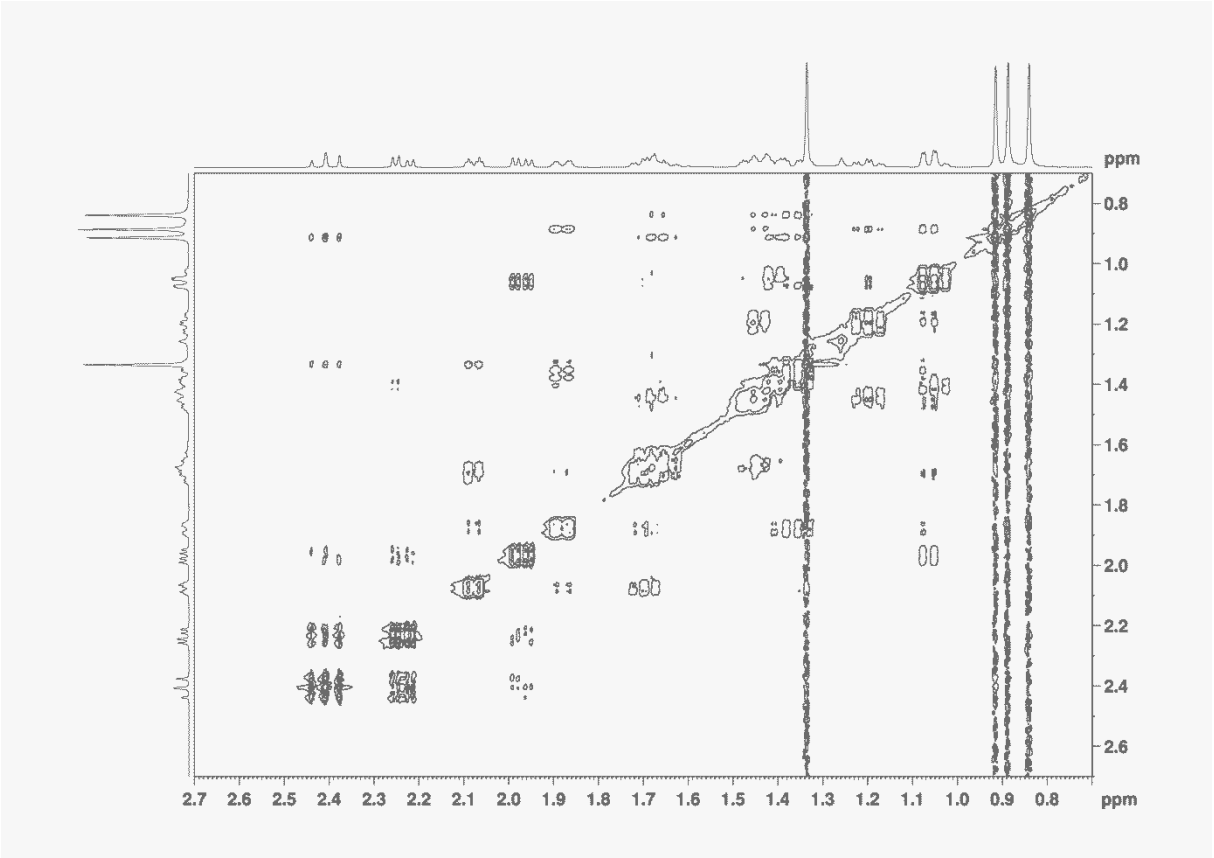
5.4.7 HMBC NMR spectrum sclareolide



5.4.8 HMBC NMR spectrum sclareolide zoomed

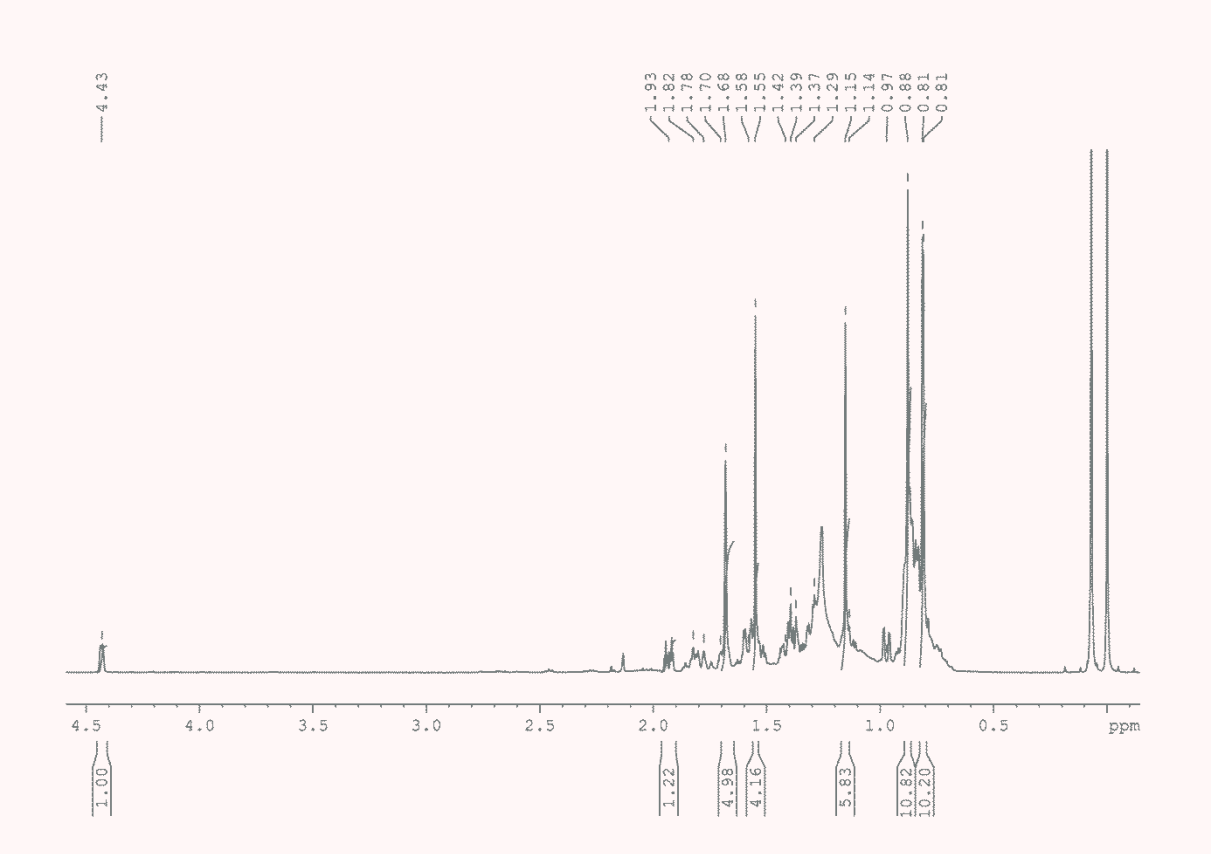


5.4.9 NOESY NMR spectrum sclareolide

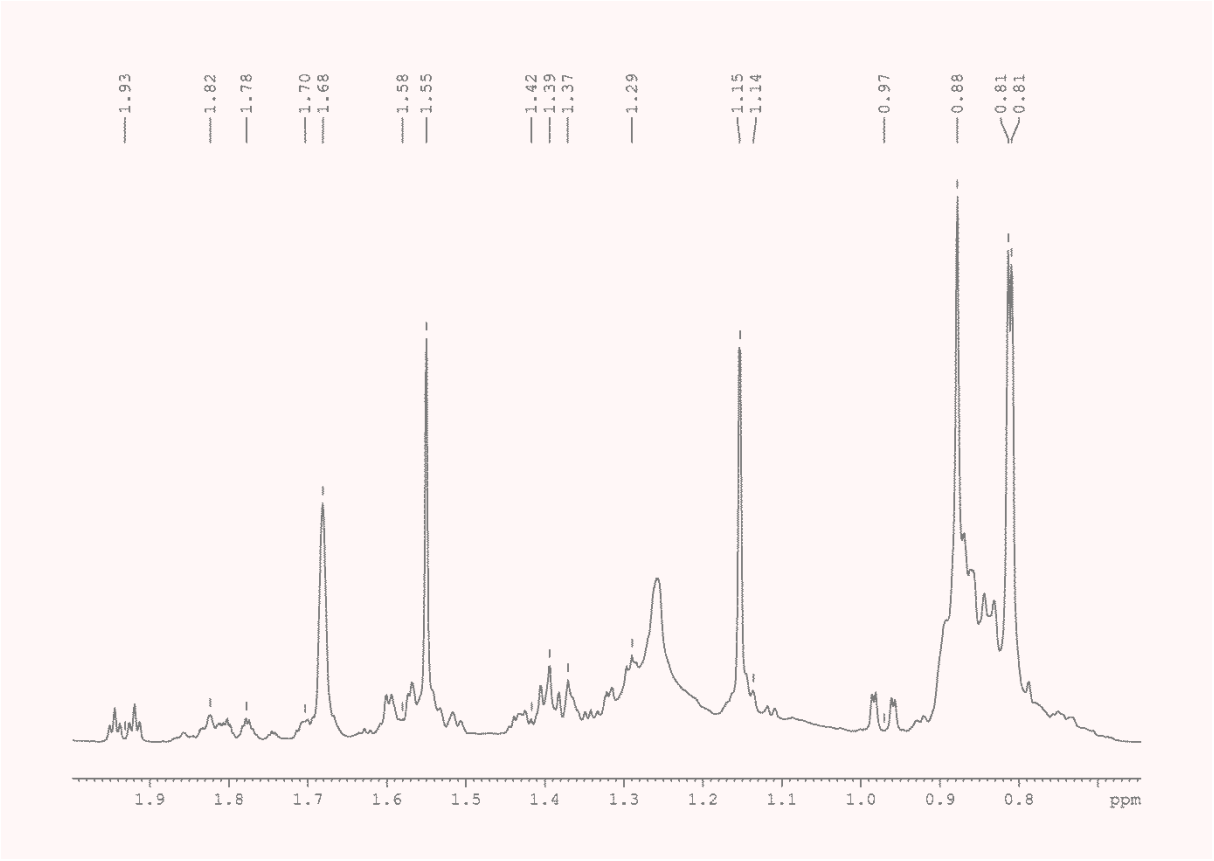


5.5 NMR spectra sclareoloxide

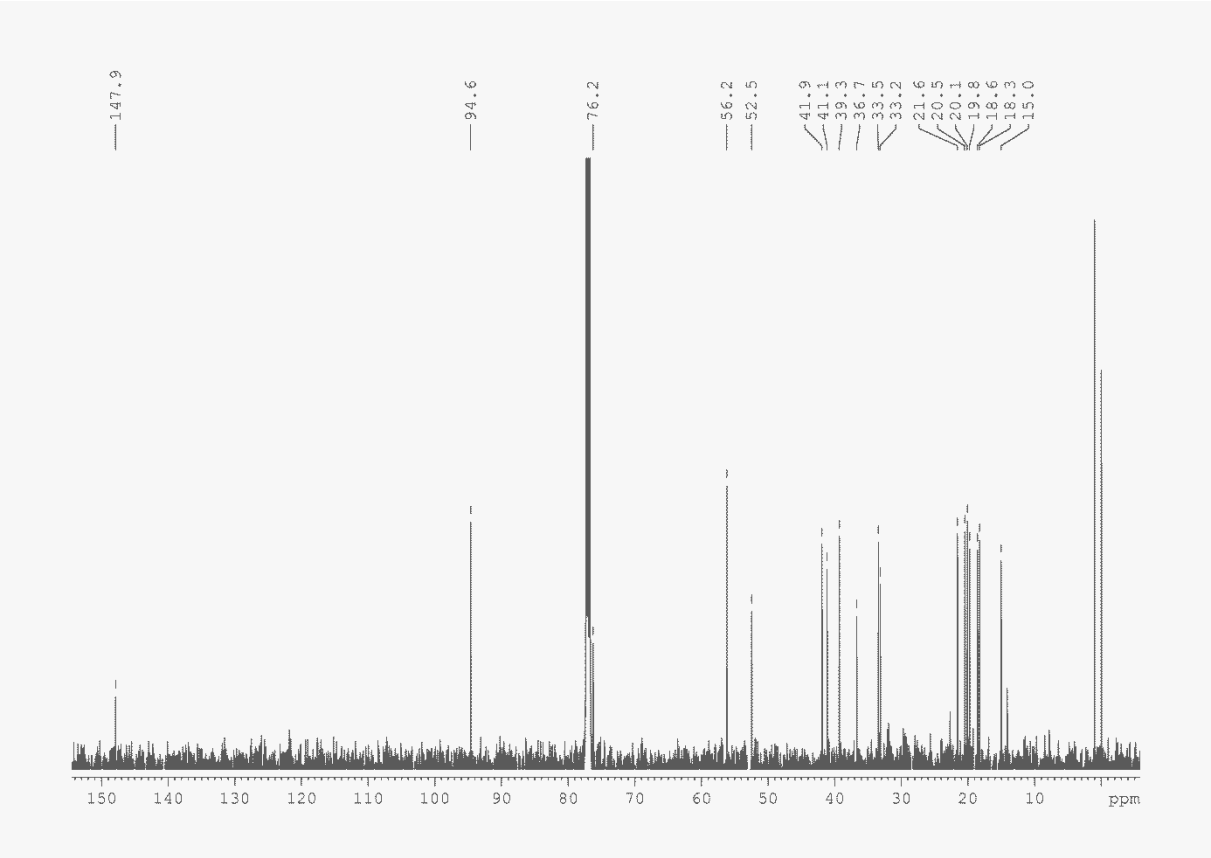
5.5.1 ^{1}H NMR spectrum sclareoloxide



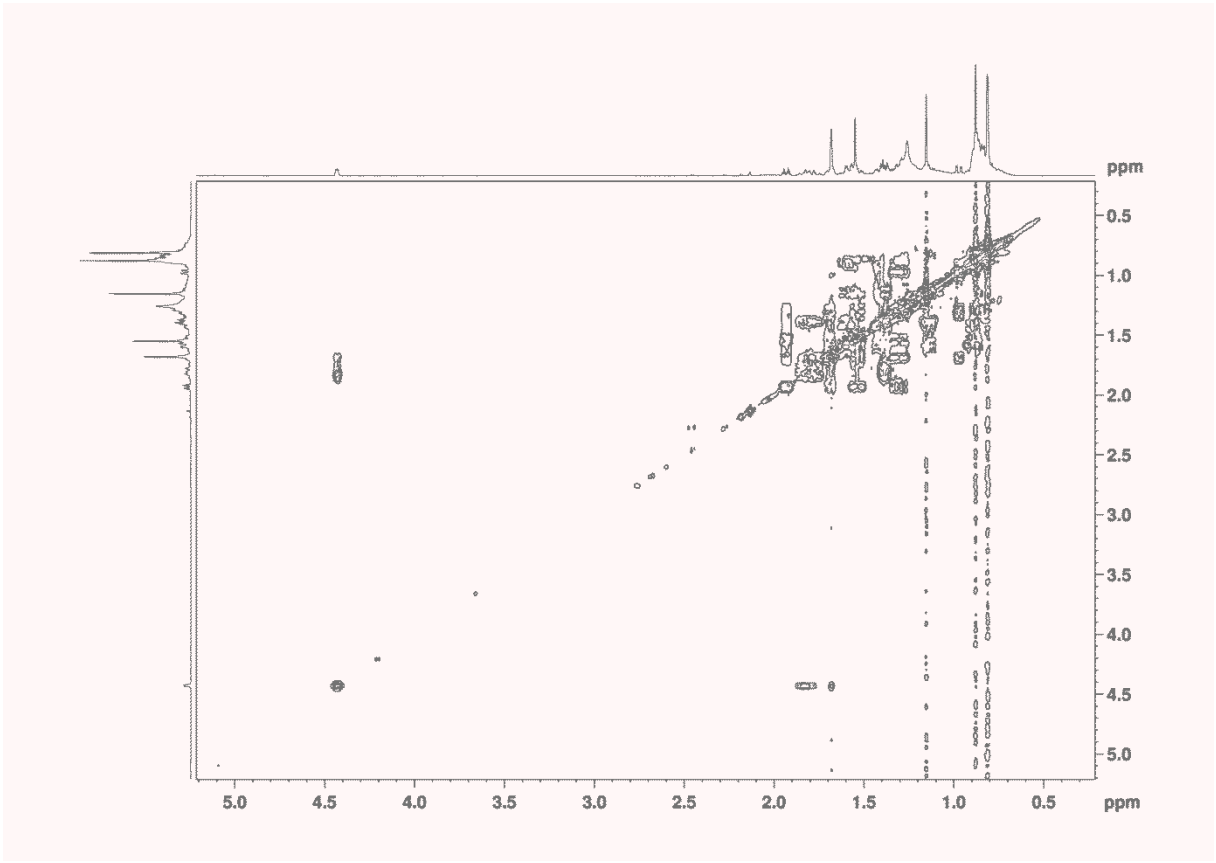
5.5.2 ^{1}H NMR spectrum sclareoloxide zoomed



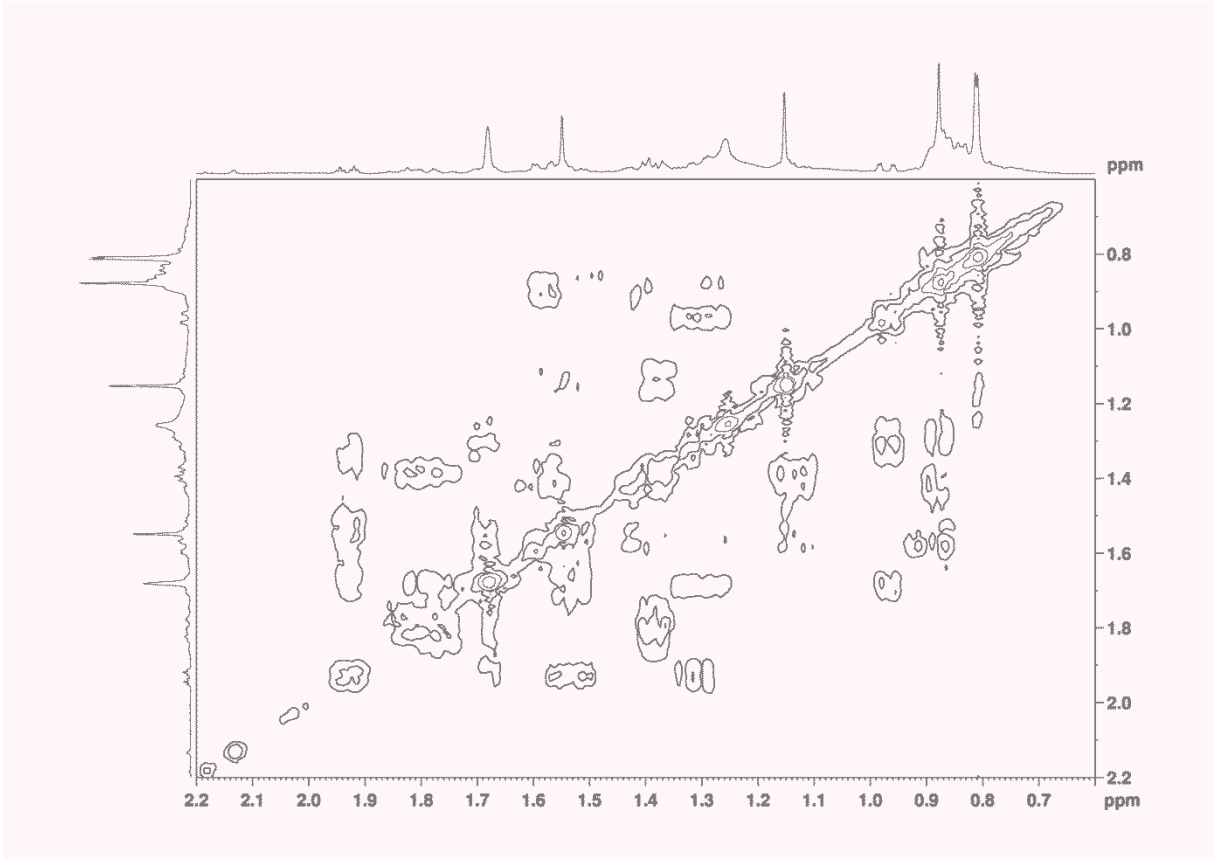
5.5.3 ^{13}C NMR spectrum sclareoloxide



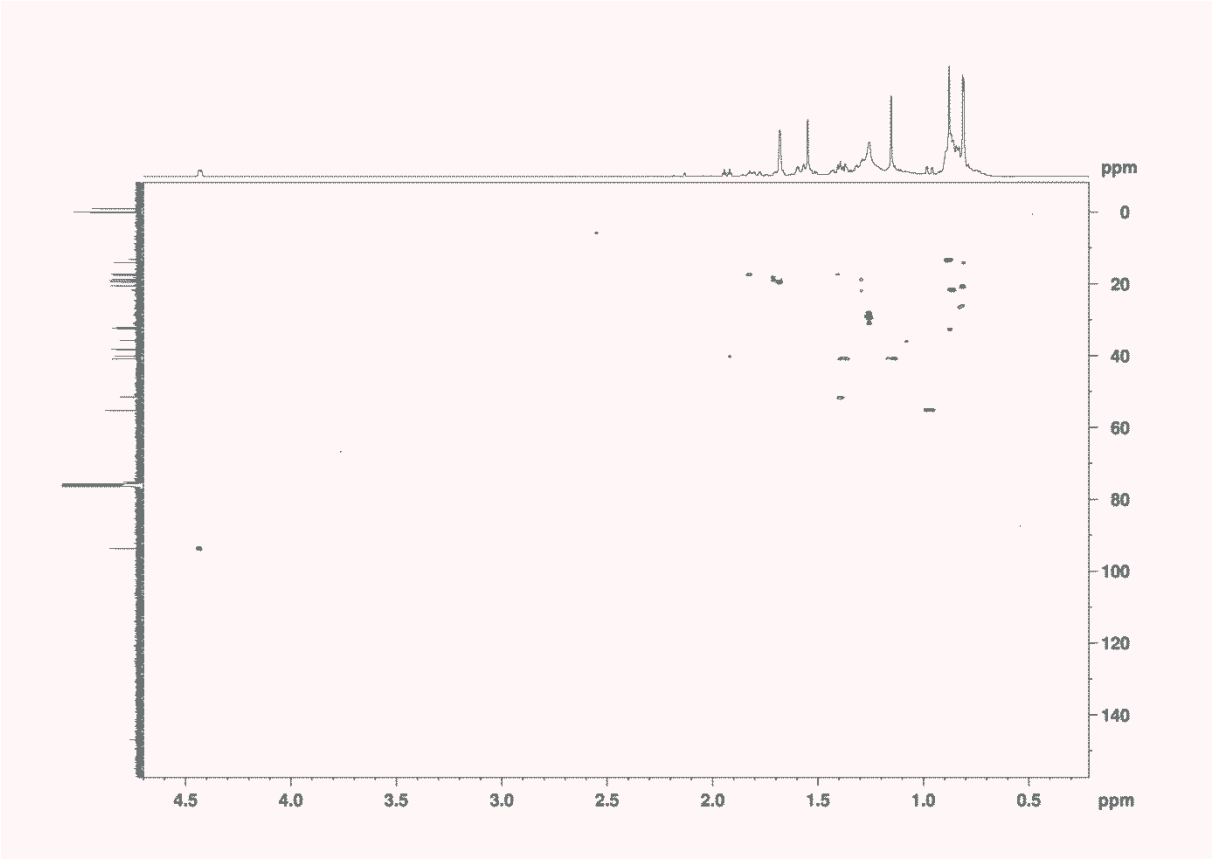
5.5.4 COSY NMR spectrum sclareoloxide



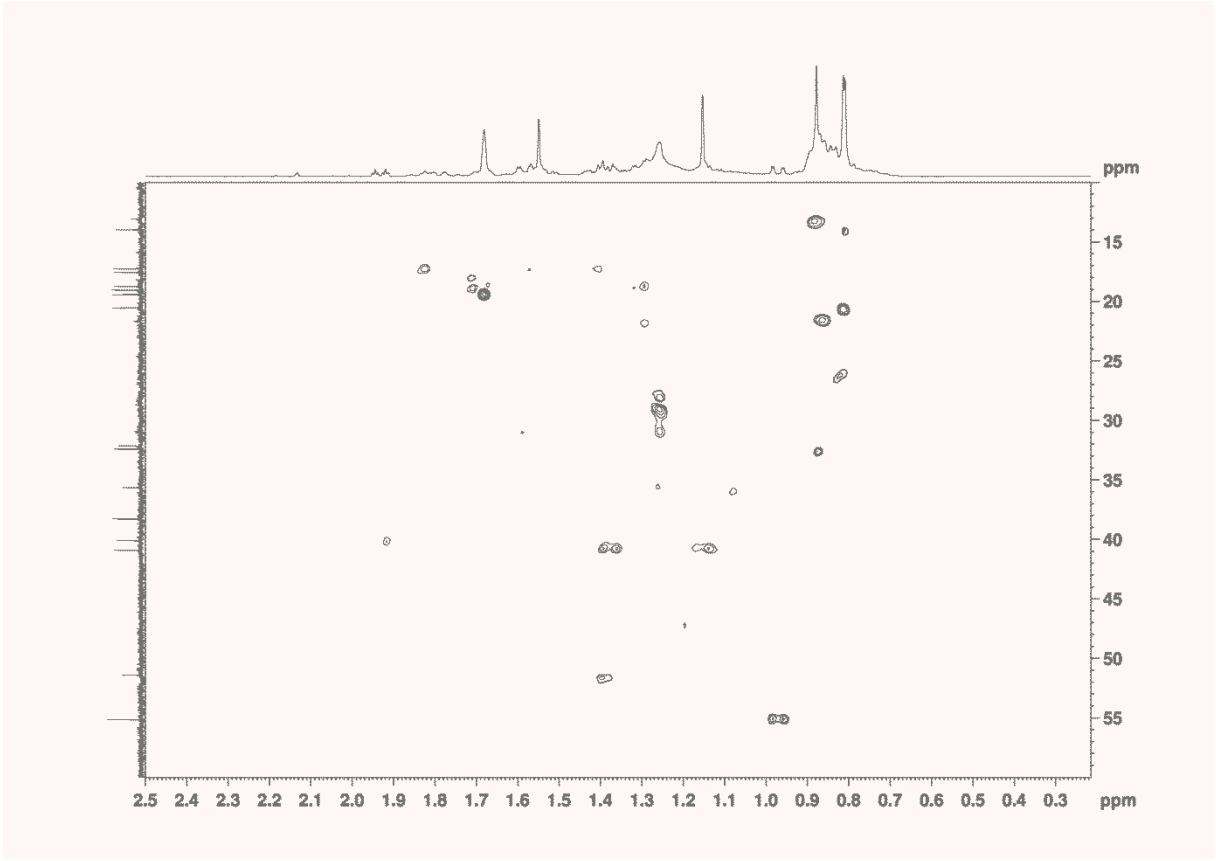
5.5.5 COSY NMR spectrum sclareoloxide zoomed



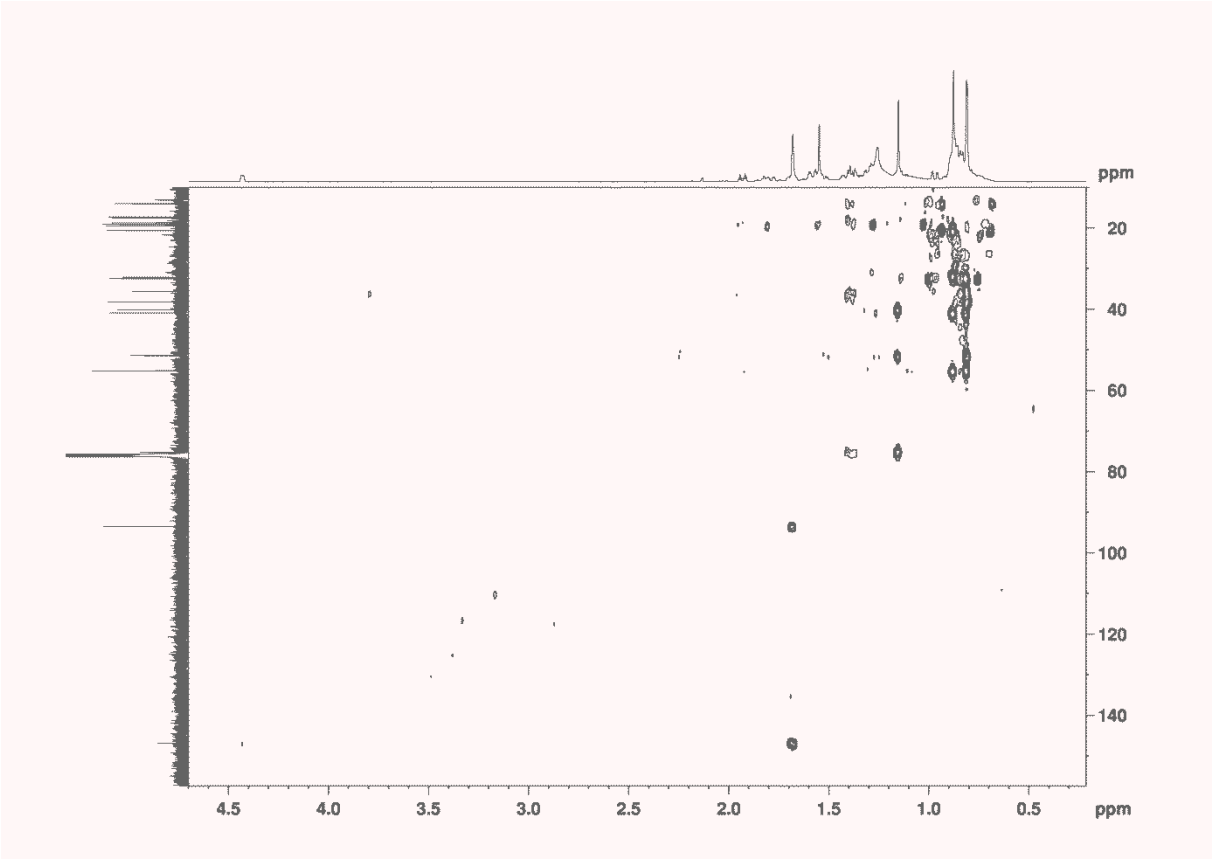
5.5.6 HSQC NMR spectrum sclareoloxide



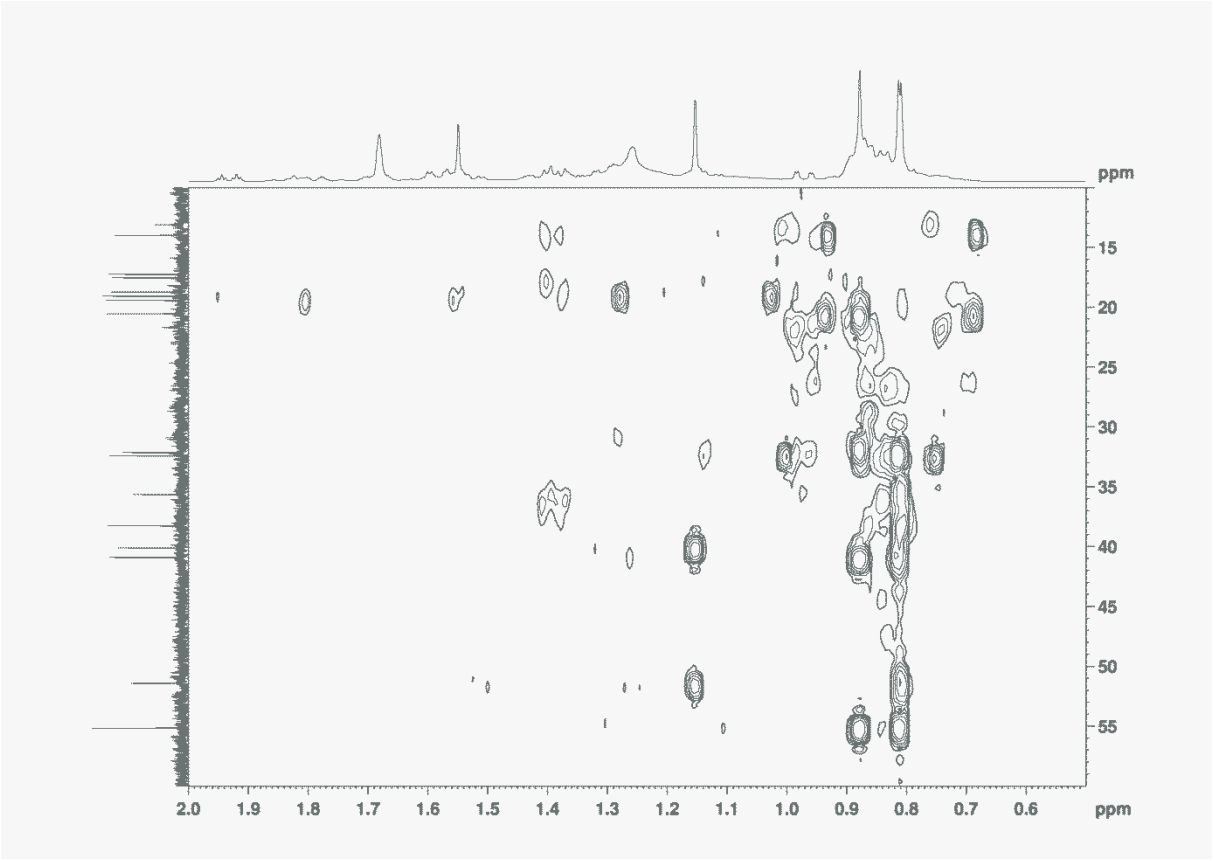
5.5.7 HSQC NMR spectrum sclareoloxide zoomed



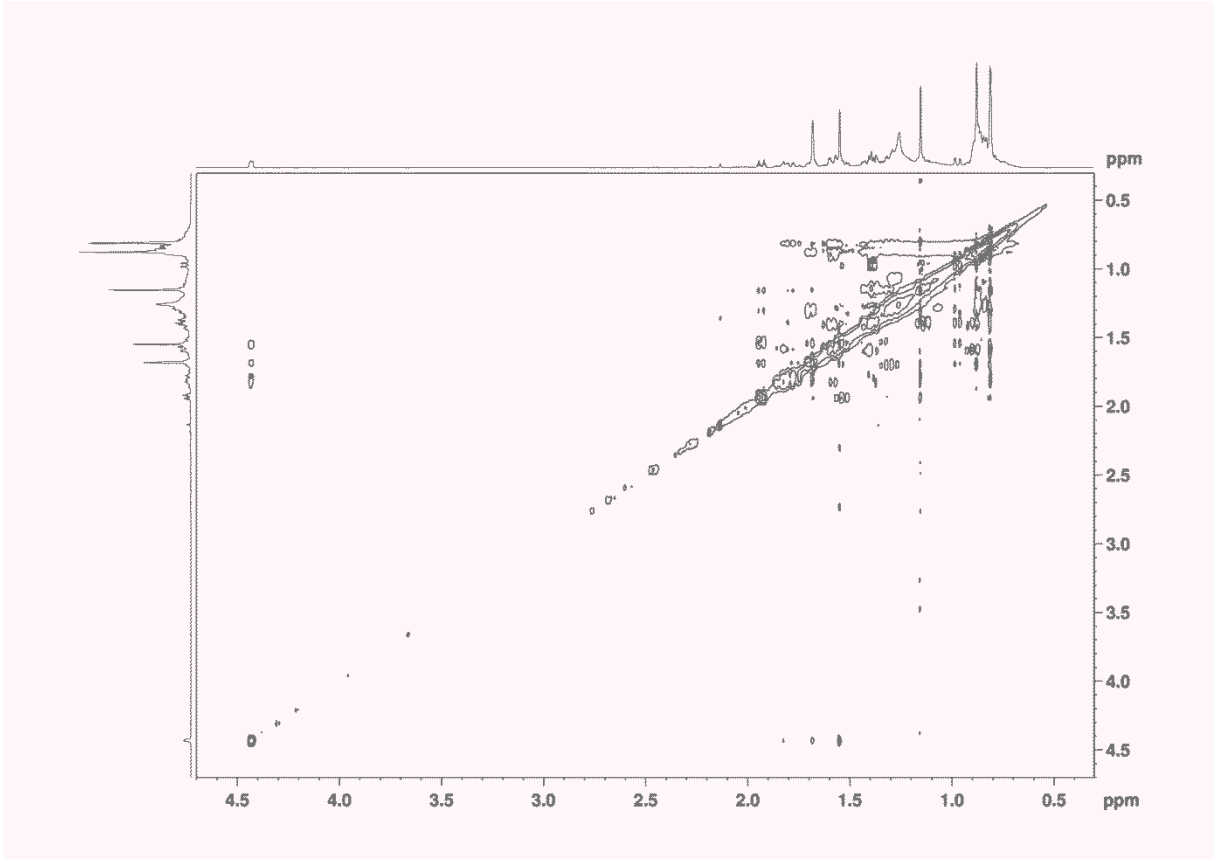
5.5.8 HMBC NMR spectrum sclareoloxide



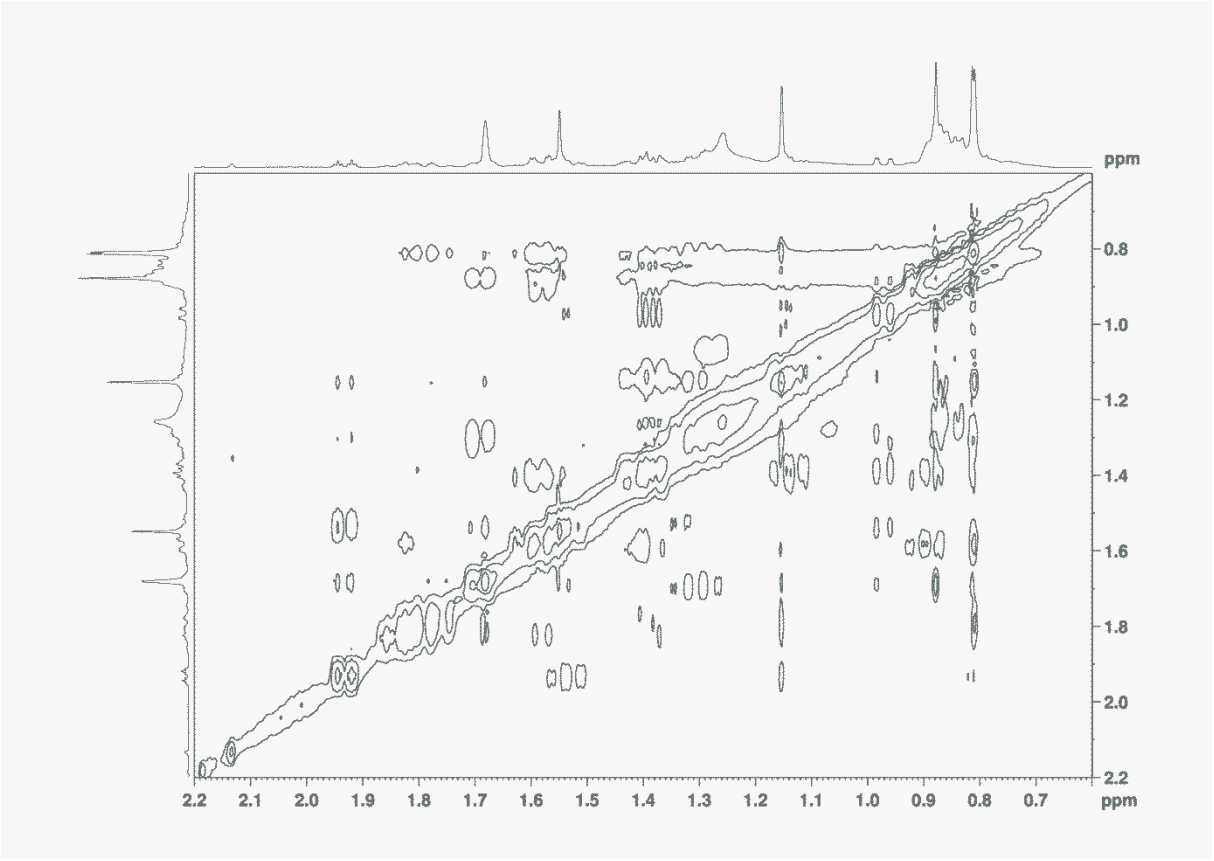
5.5.9 HMBC NMR spectrum sclareoloxide zoomed



5.5.10NOESY NMR spectrum sclareoloxide

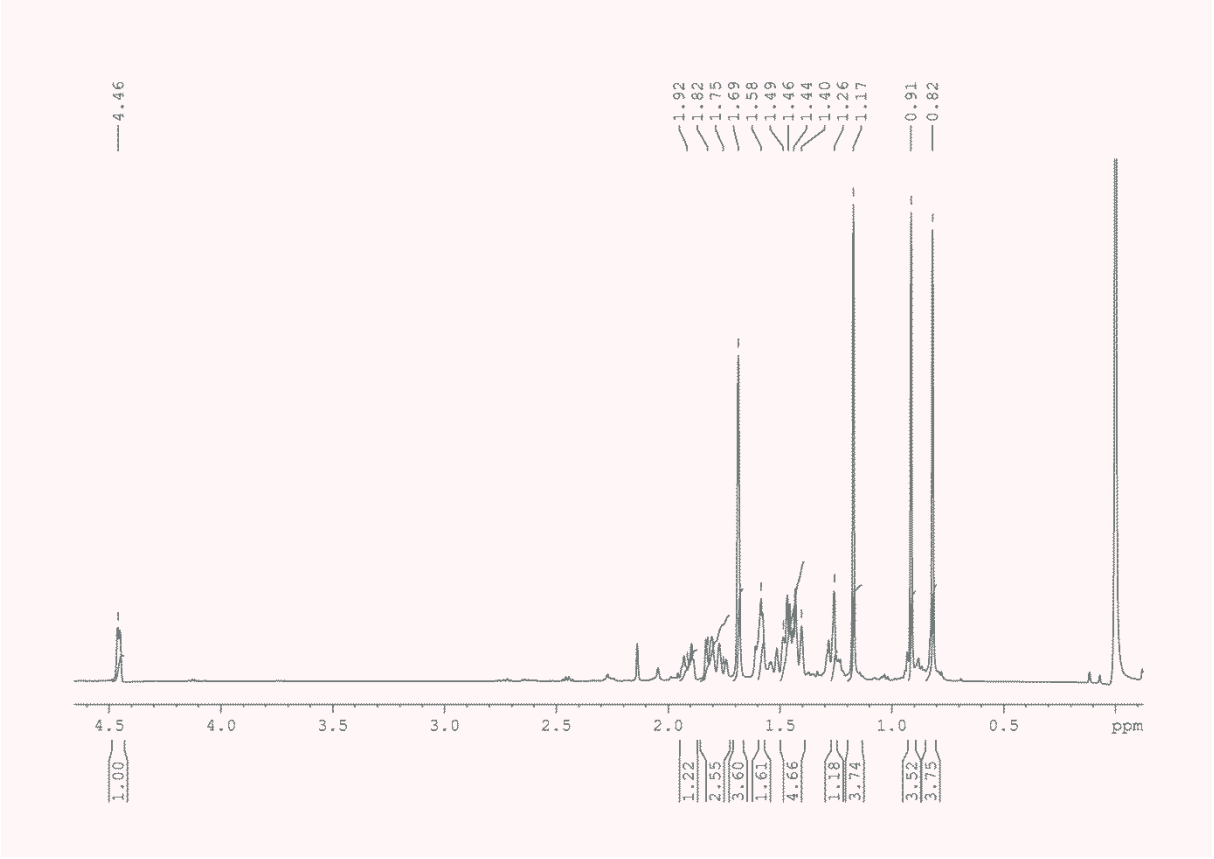


5.5.11NOESY NMR spectrum sclareoloxide zoomed

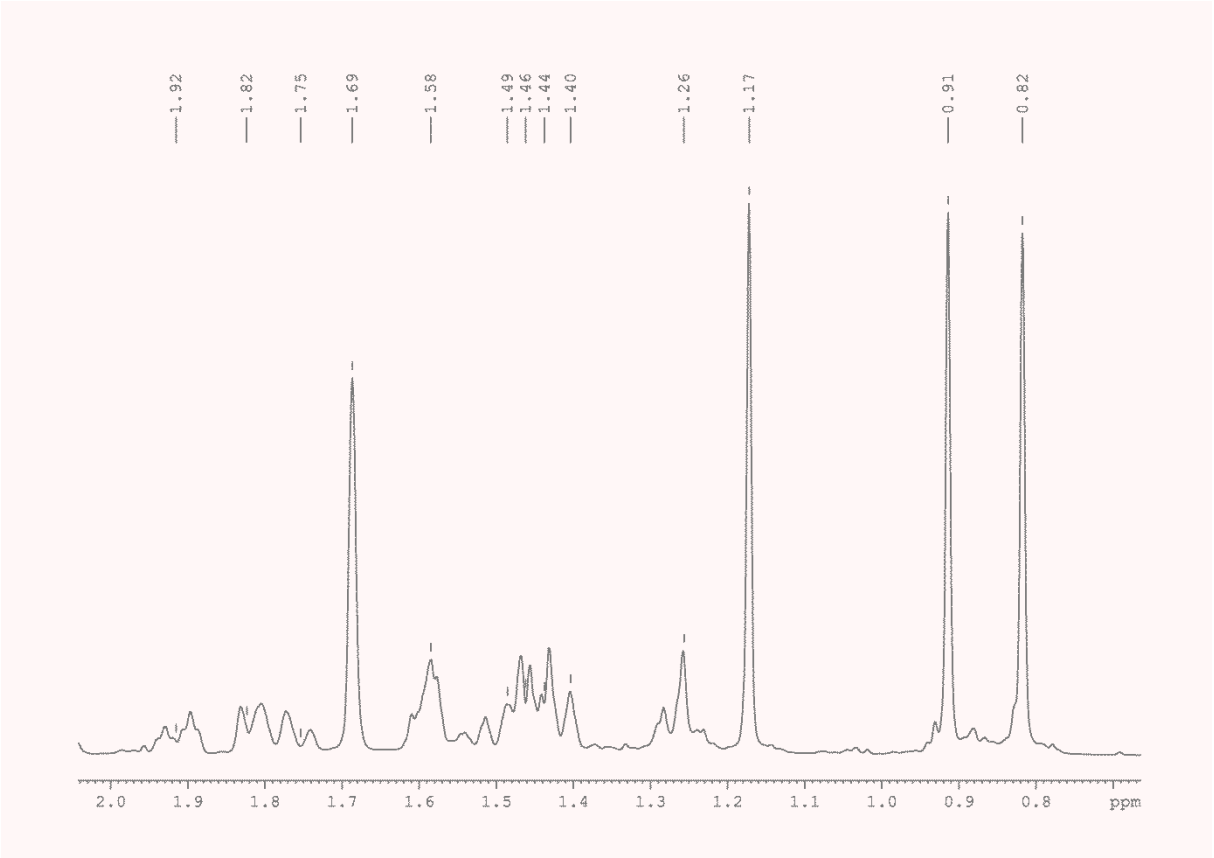


5.6 NMR spectra hexahydrochromene

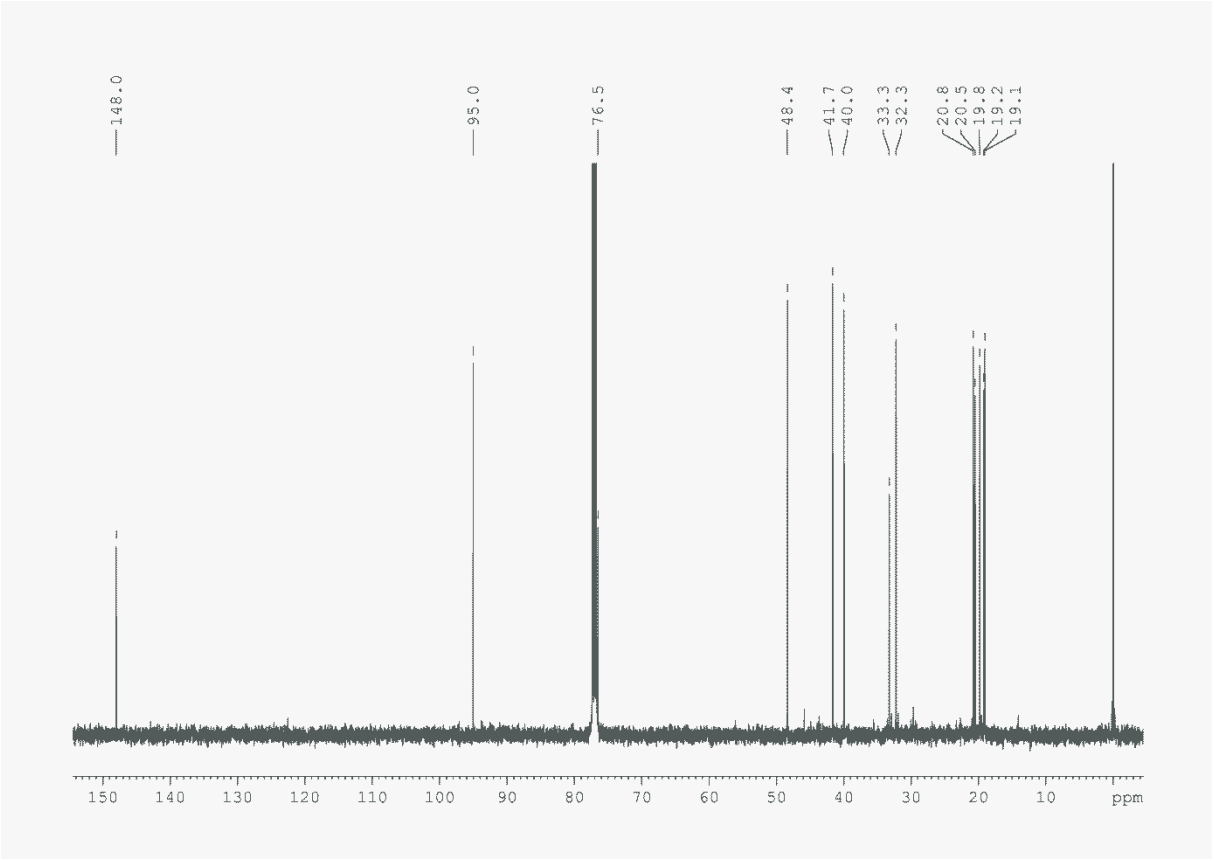
5.6.1 ^{1}H NMR spectrum hexahydrochromene



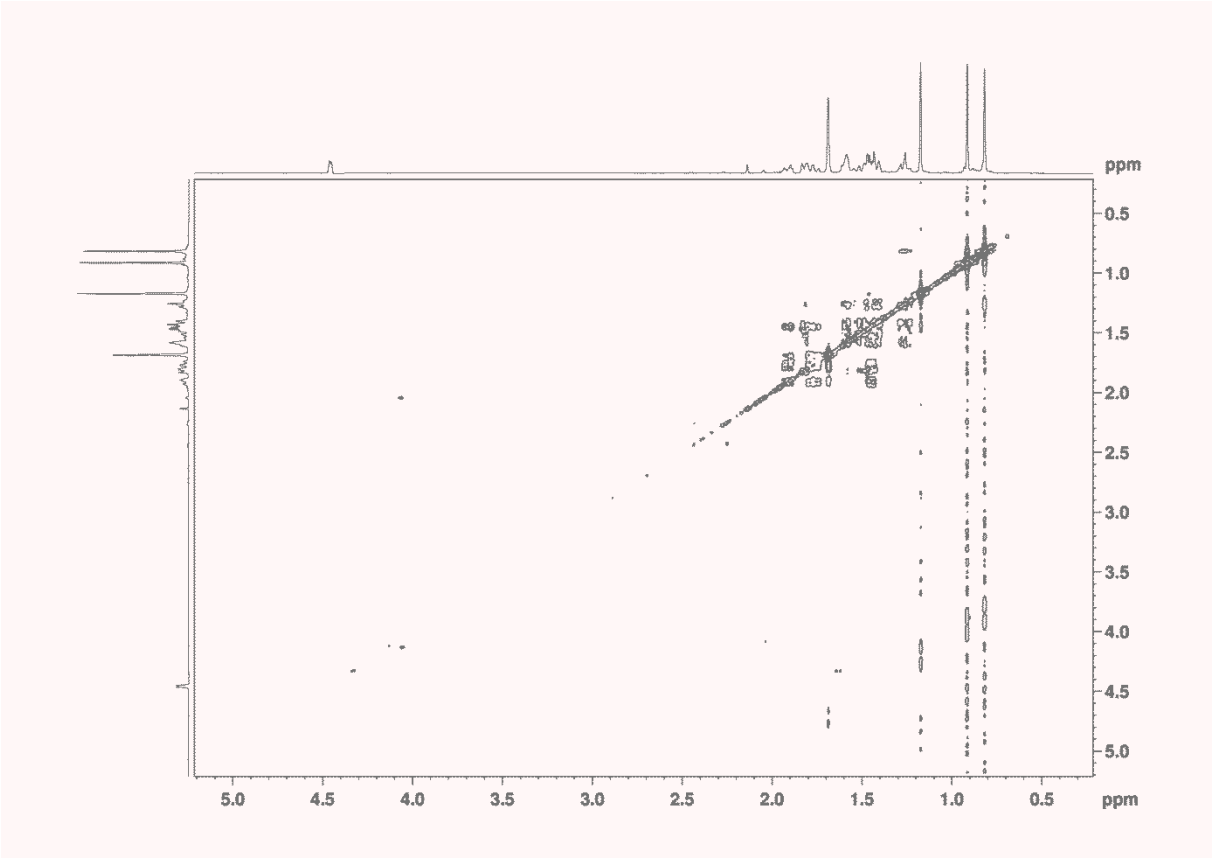
5.6.2 ^{1}H NMR spectrum hexahydrochromene zoomed



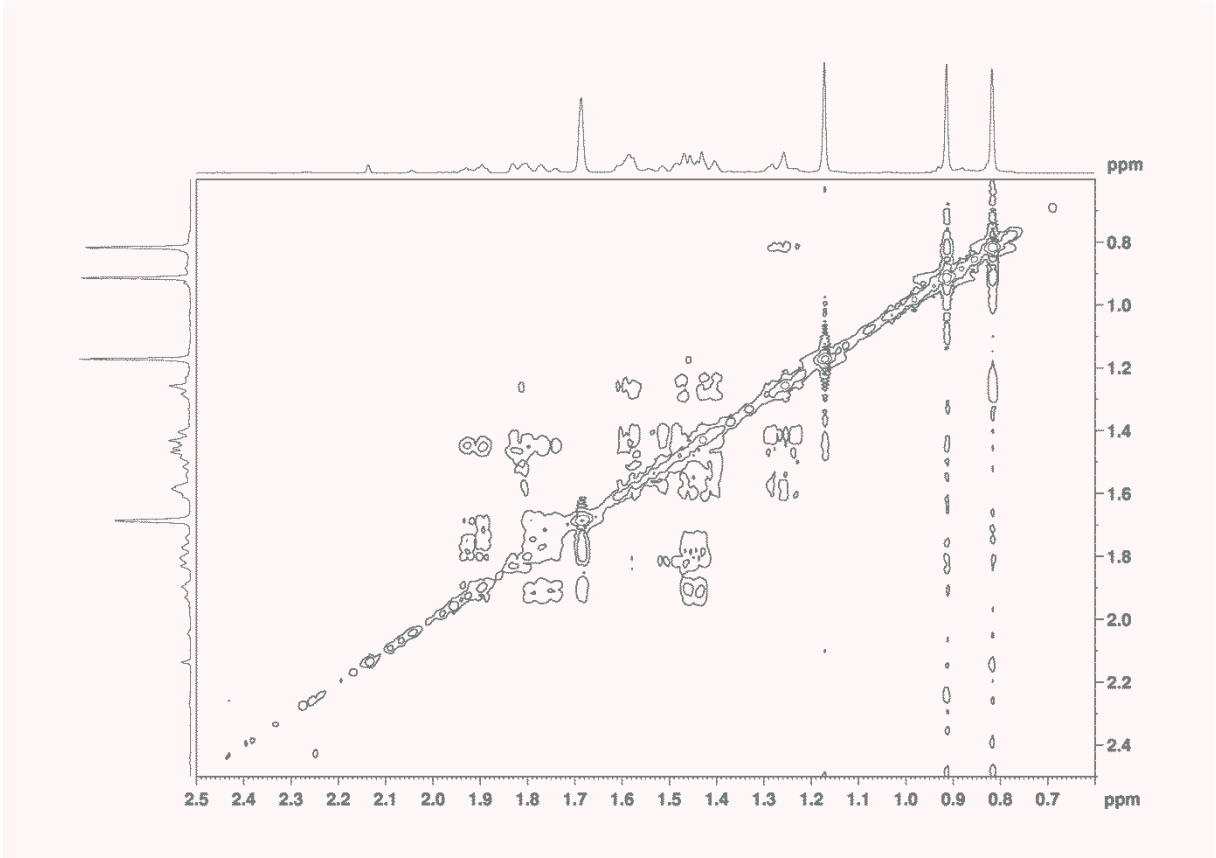
5.6.3 ^{13}C NMR spectrum hexahydrochromene



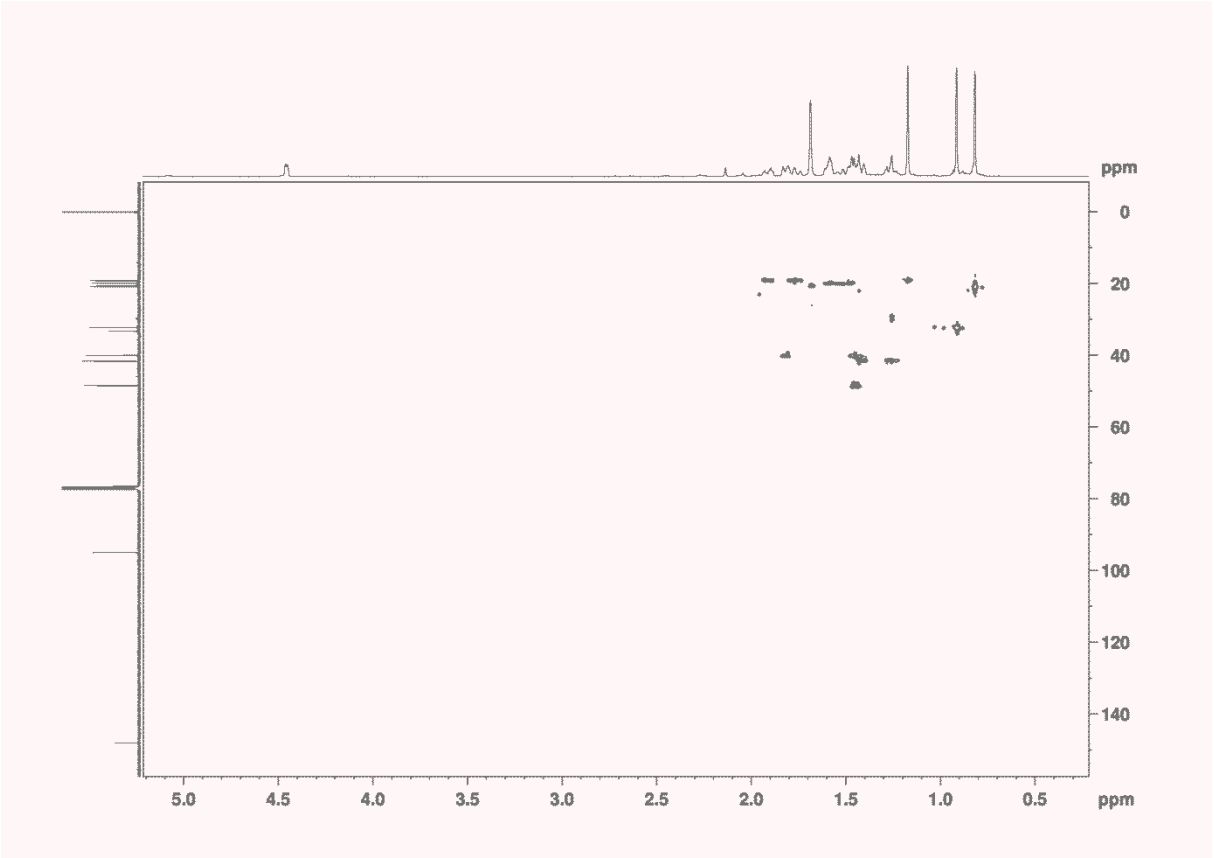
5.6.4 COSY NMR spectrum hexahydrochromene



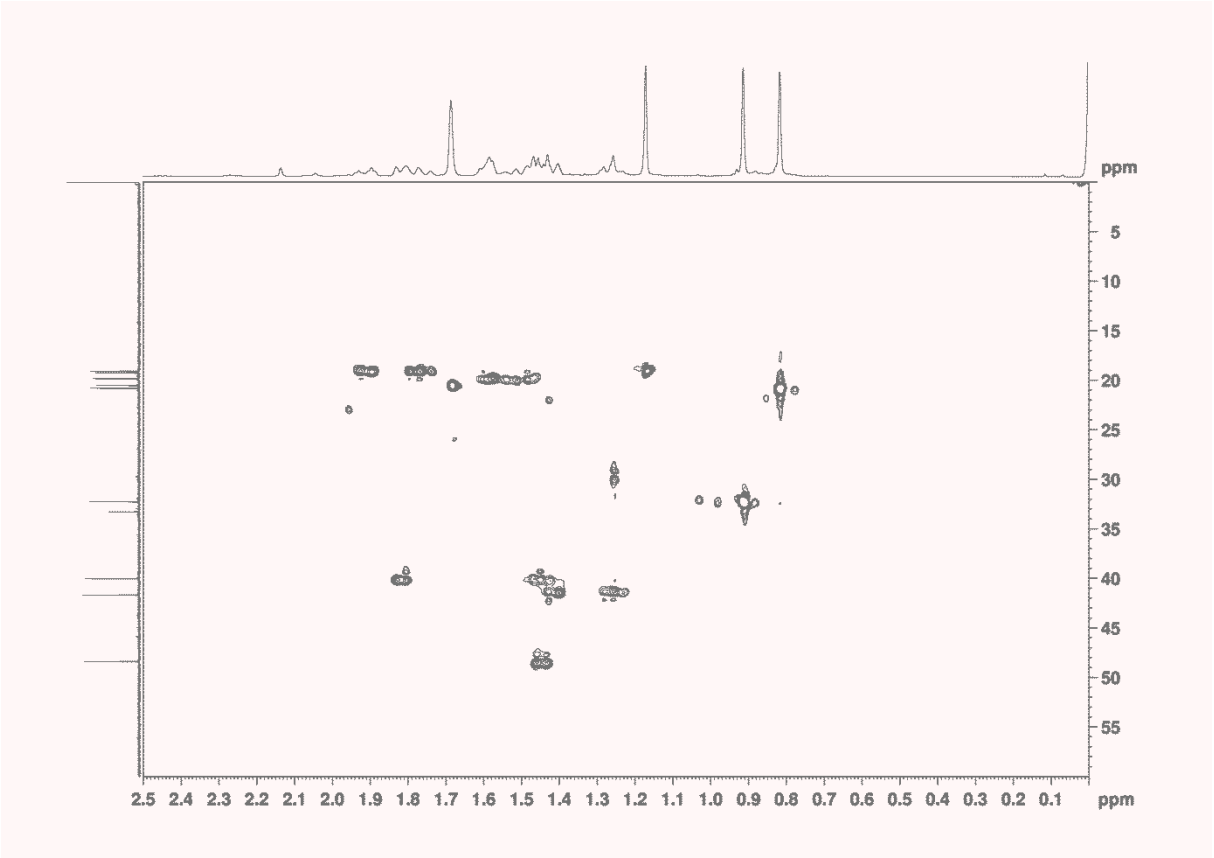
5.6.5 COSY NMR spectrum hexahydrochromene zoomed



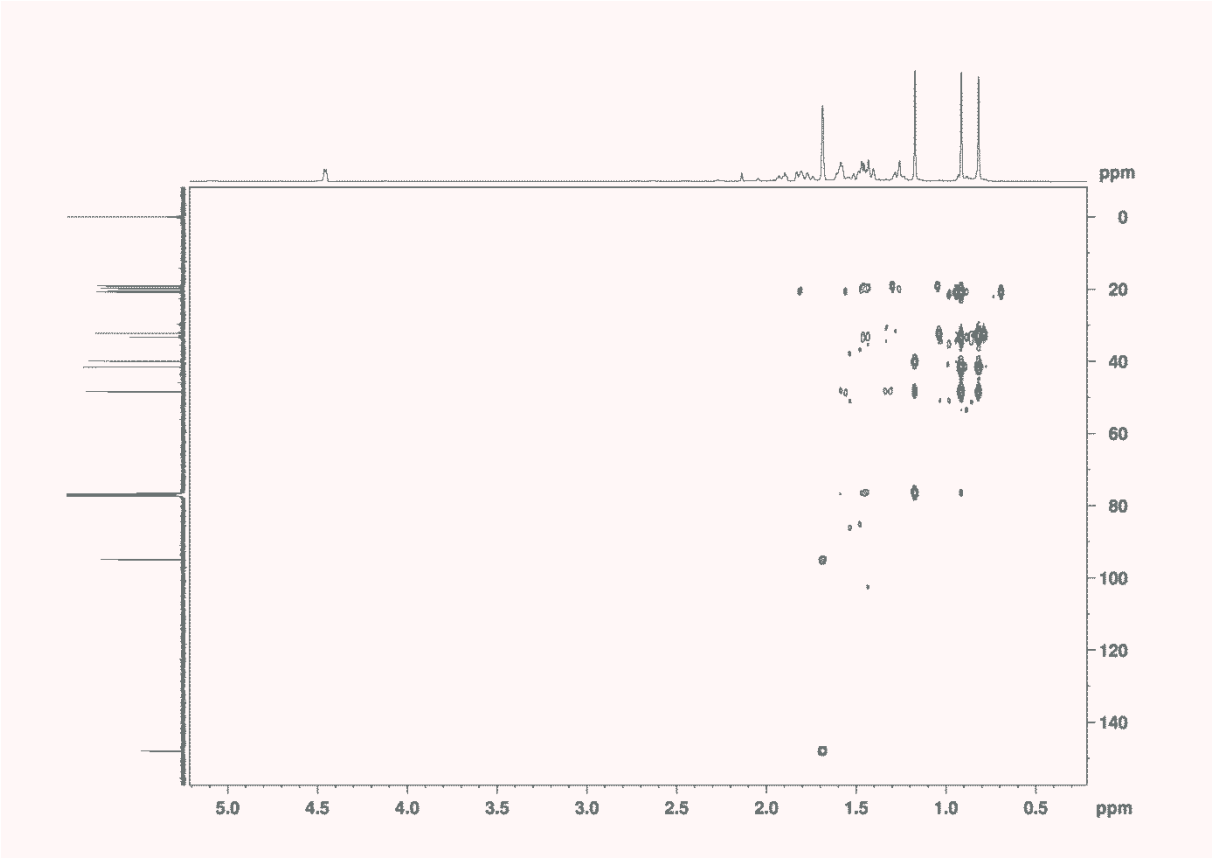
5.6.6 HSQC NMR spectrum hexahydrochromene



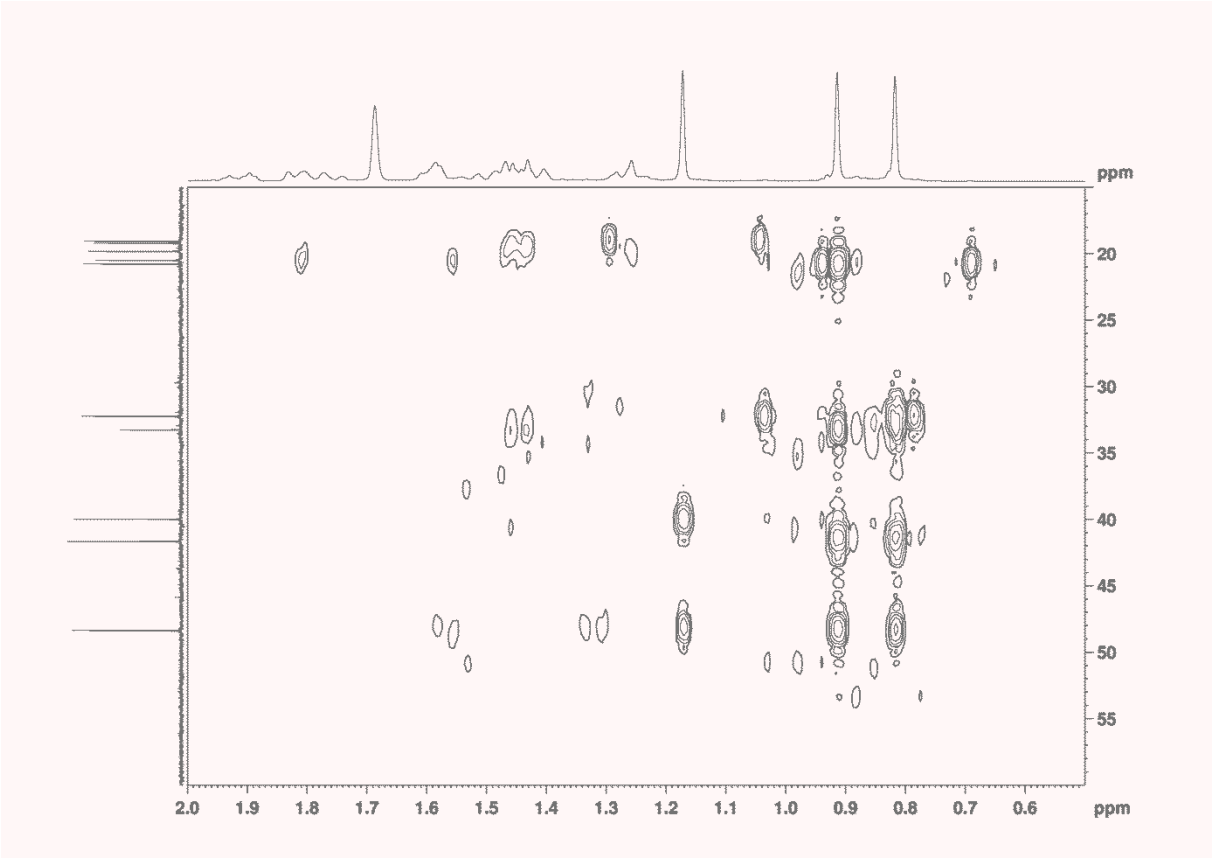
5.6.7 HSQC NMR spectrum hexahydrochromene zoomed



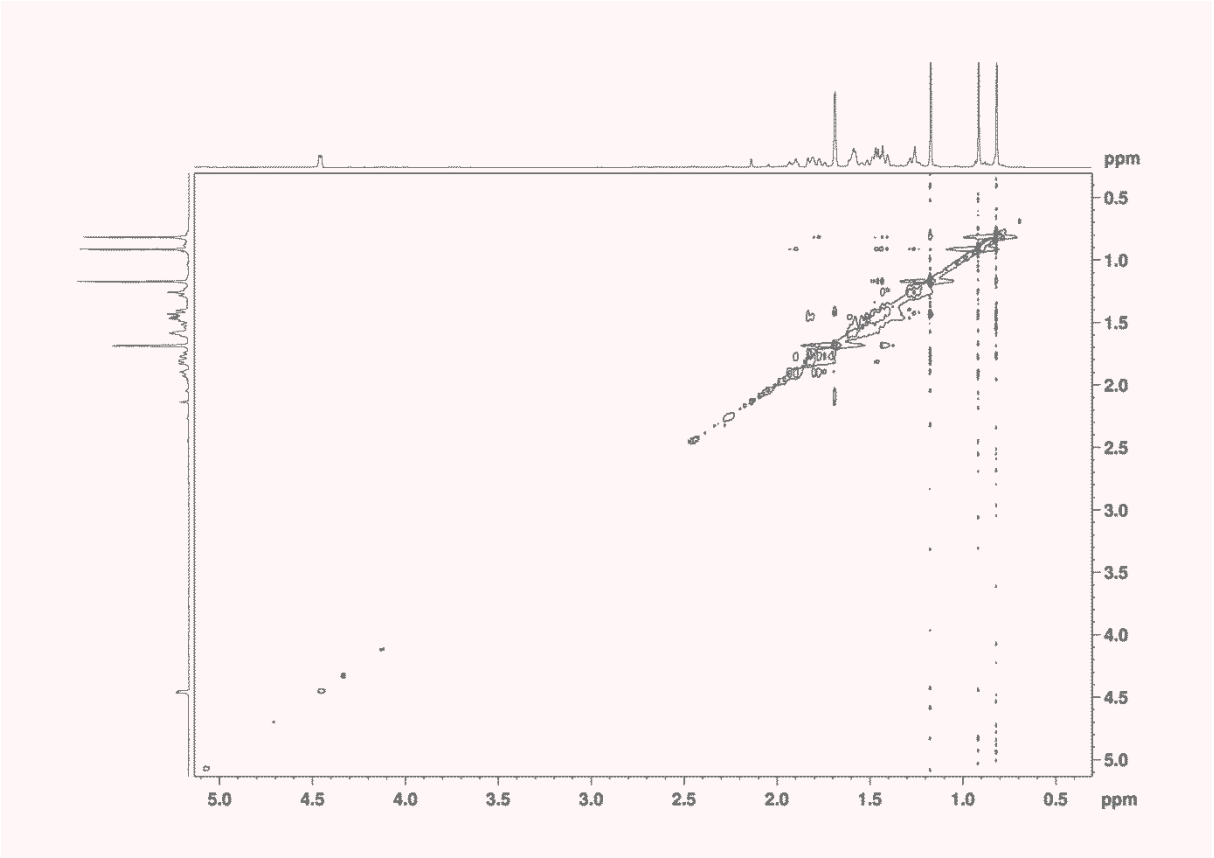
5.6.8 HMBC NMR spectrum hexahydrochromene



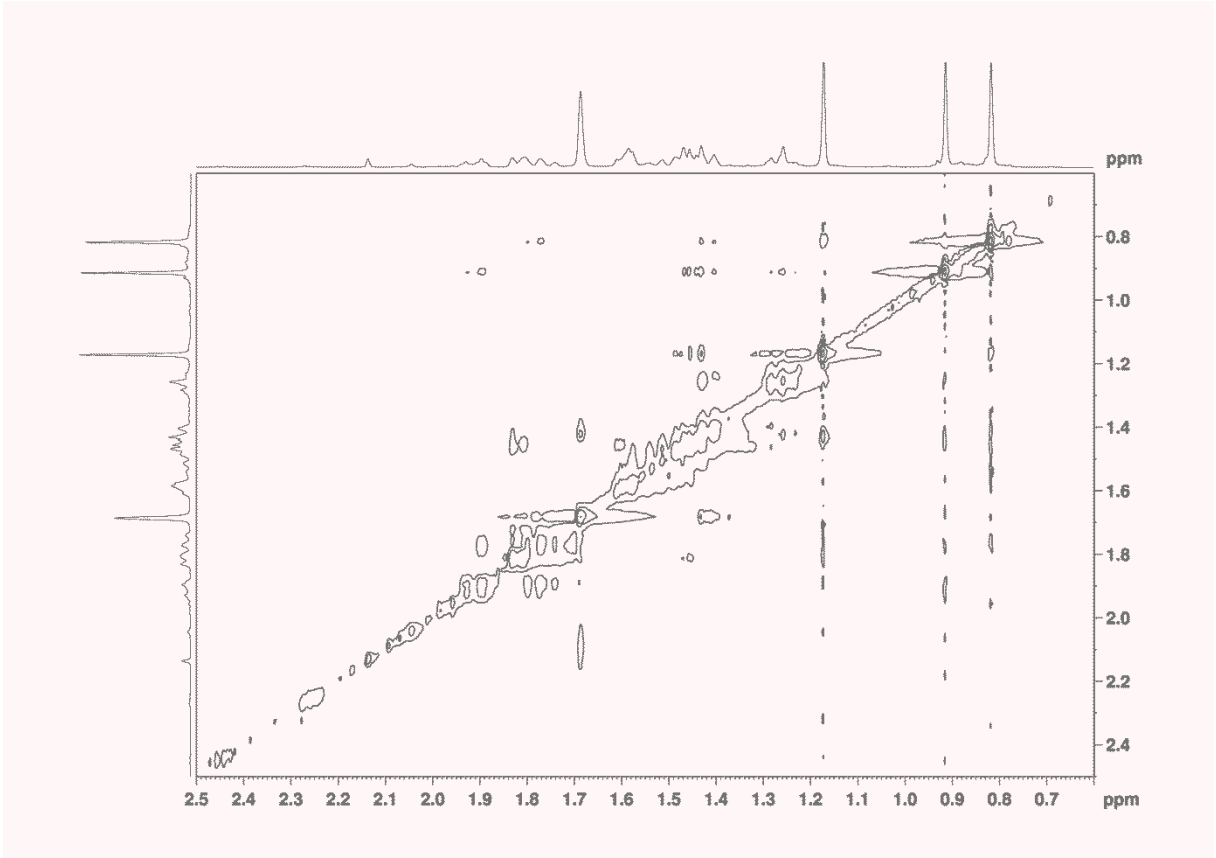
5.6.9 HMBC NMR spectrum hexahydrochromene zoomed



5.6.10NOESY NMR spectrum hexahydrochromene

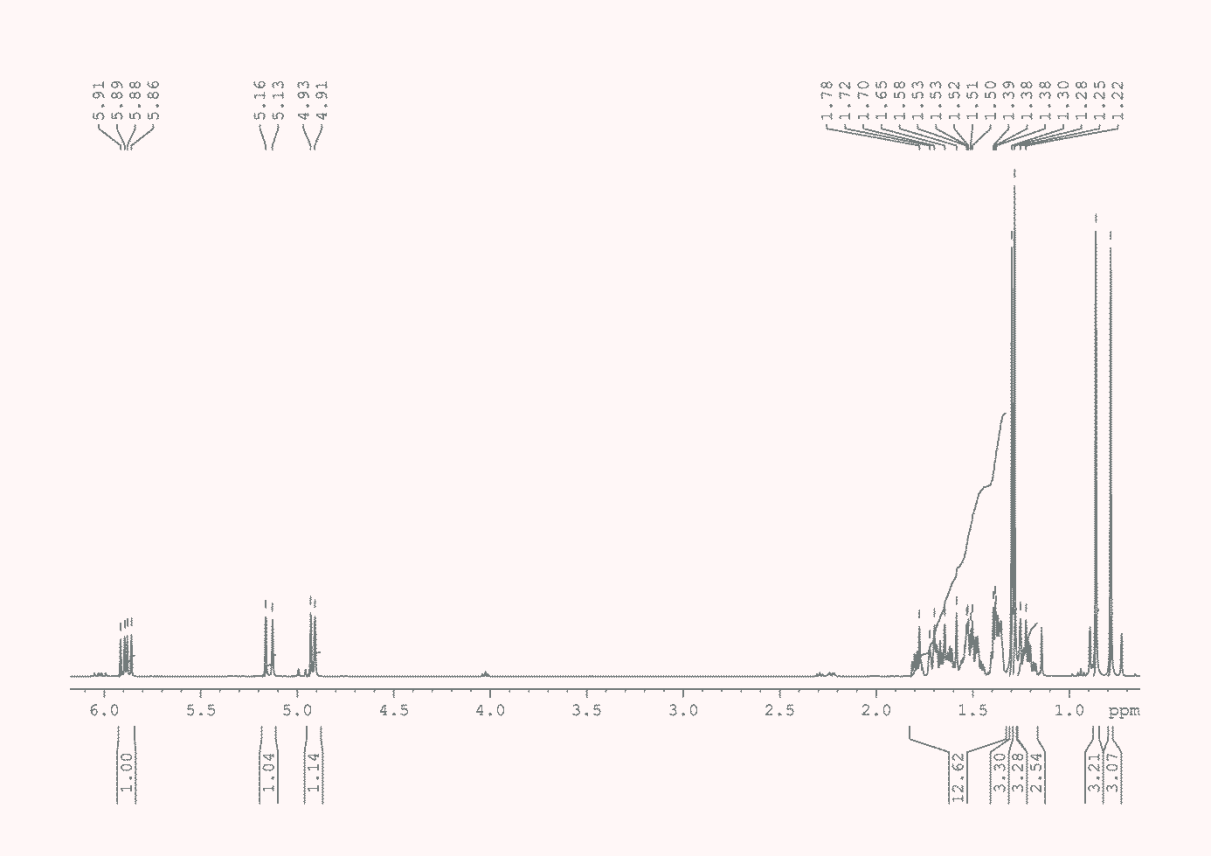


5.6.11NOESY NMR spectrum hexahydrochromene zoomed

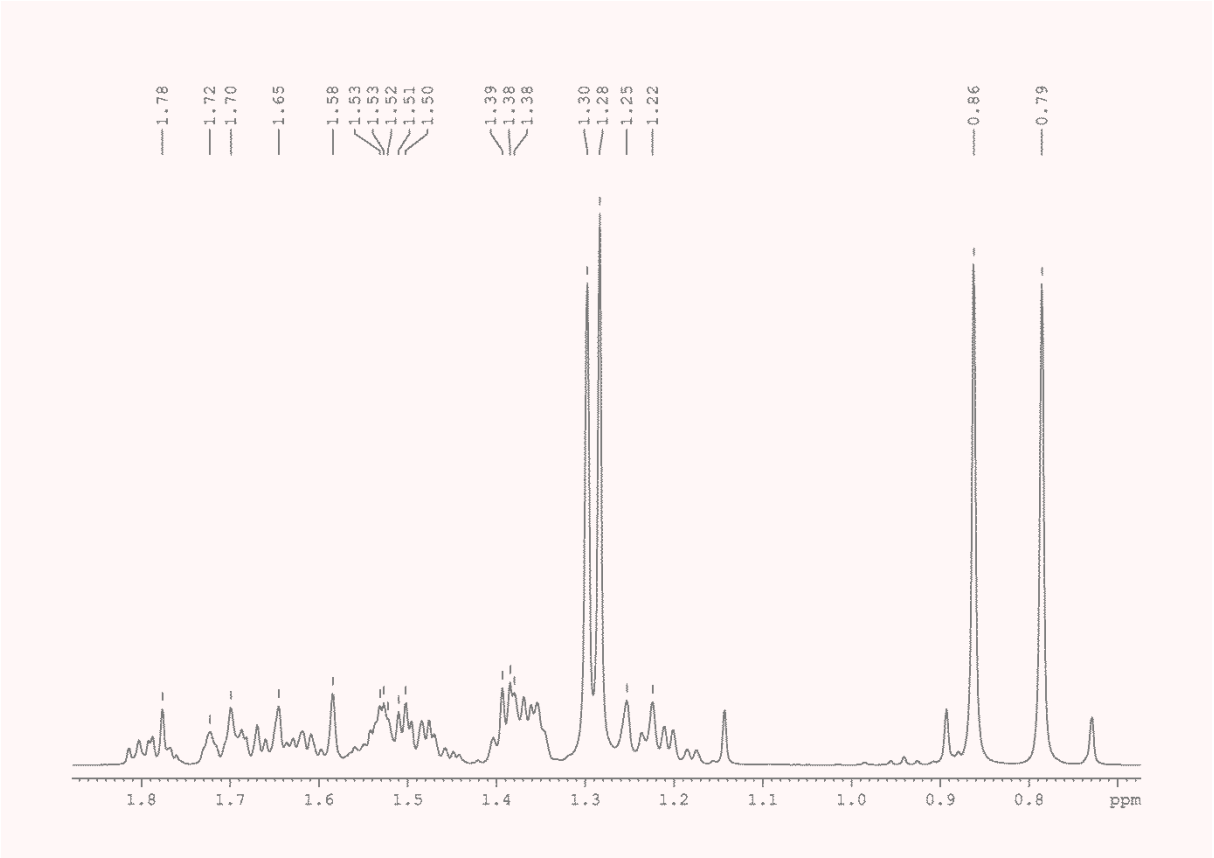


5.7 NMR spectra (-)-caparrapioxide

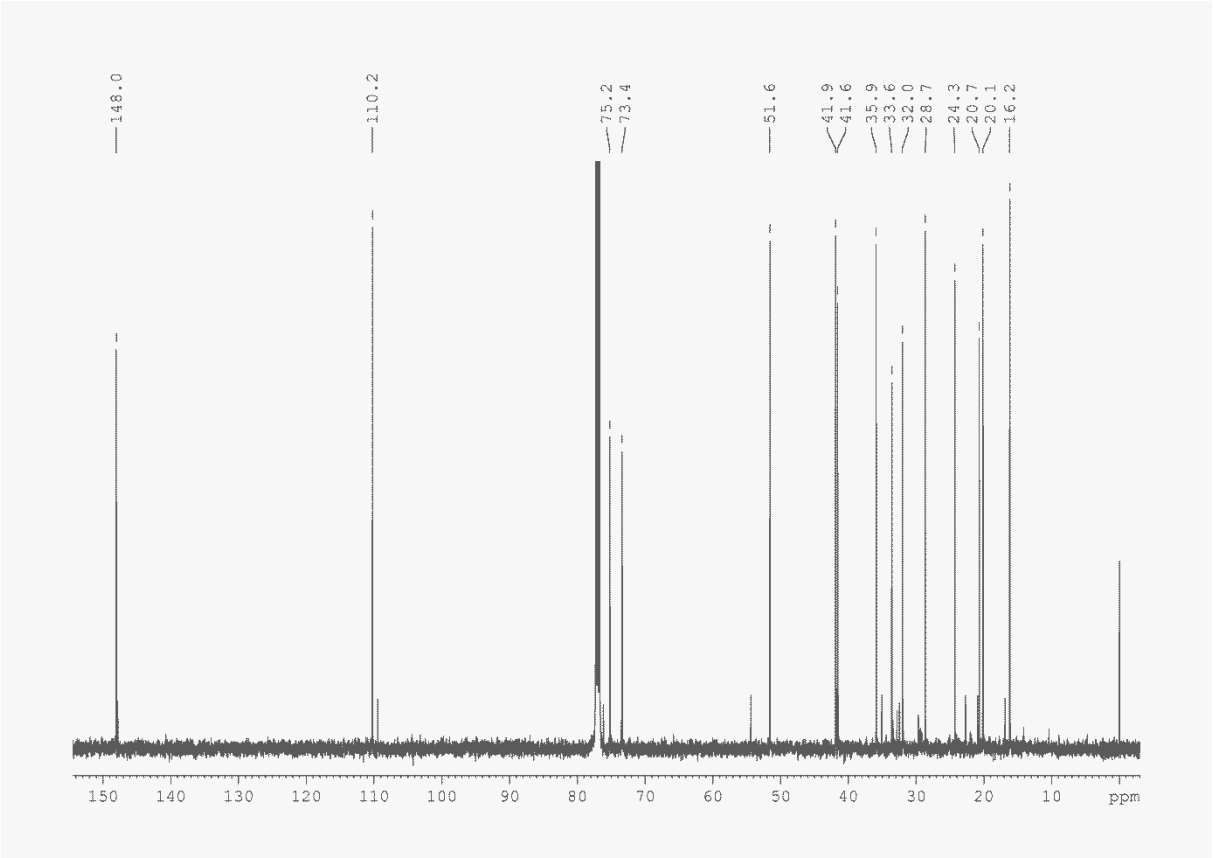
5.7.1 ^{1}H NMR spectrum (-)-caparrapioxide



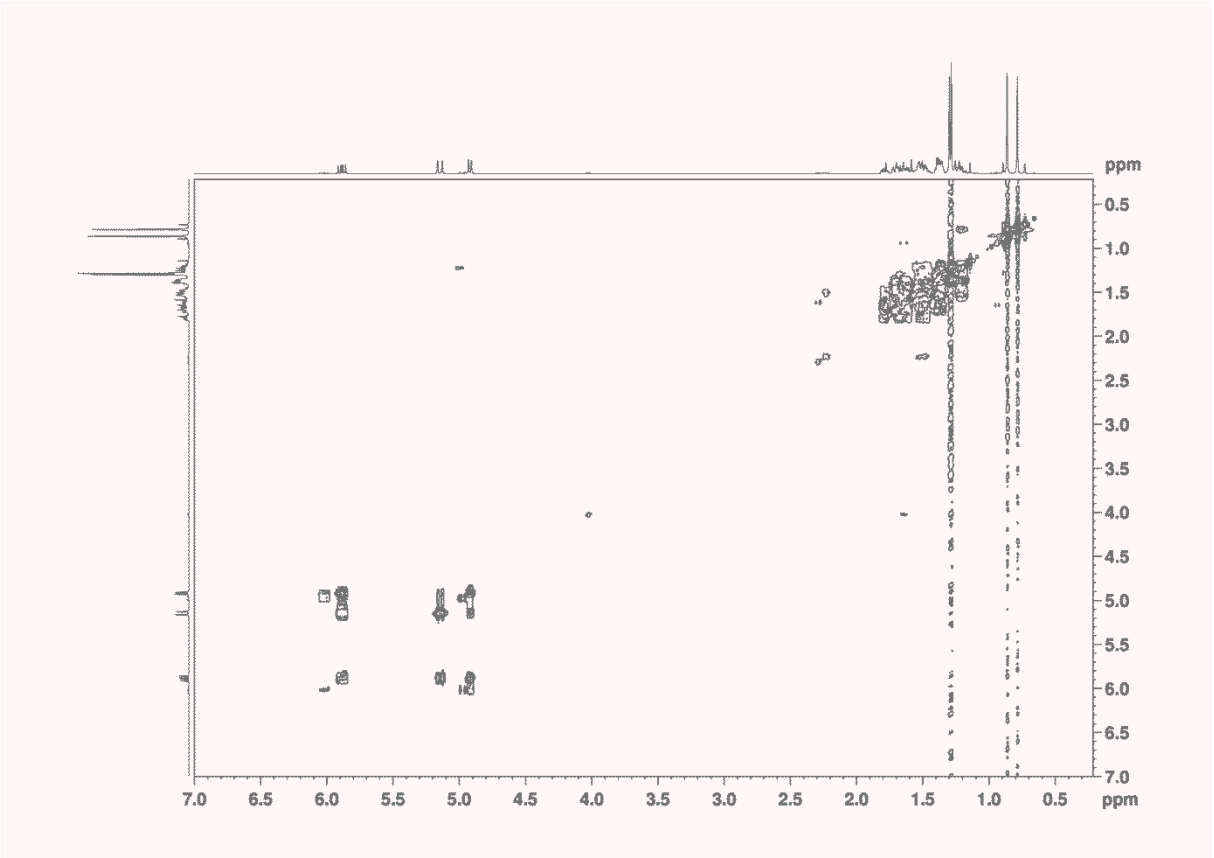
5.7.2 ^{1}H NMR spectrum (-)-caparrapioxide zoomed



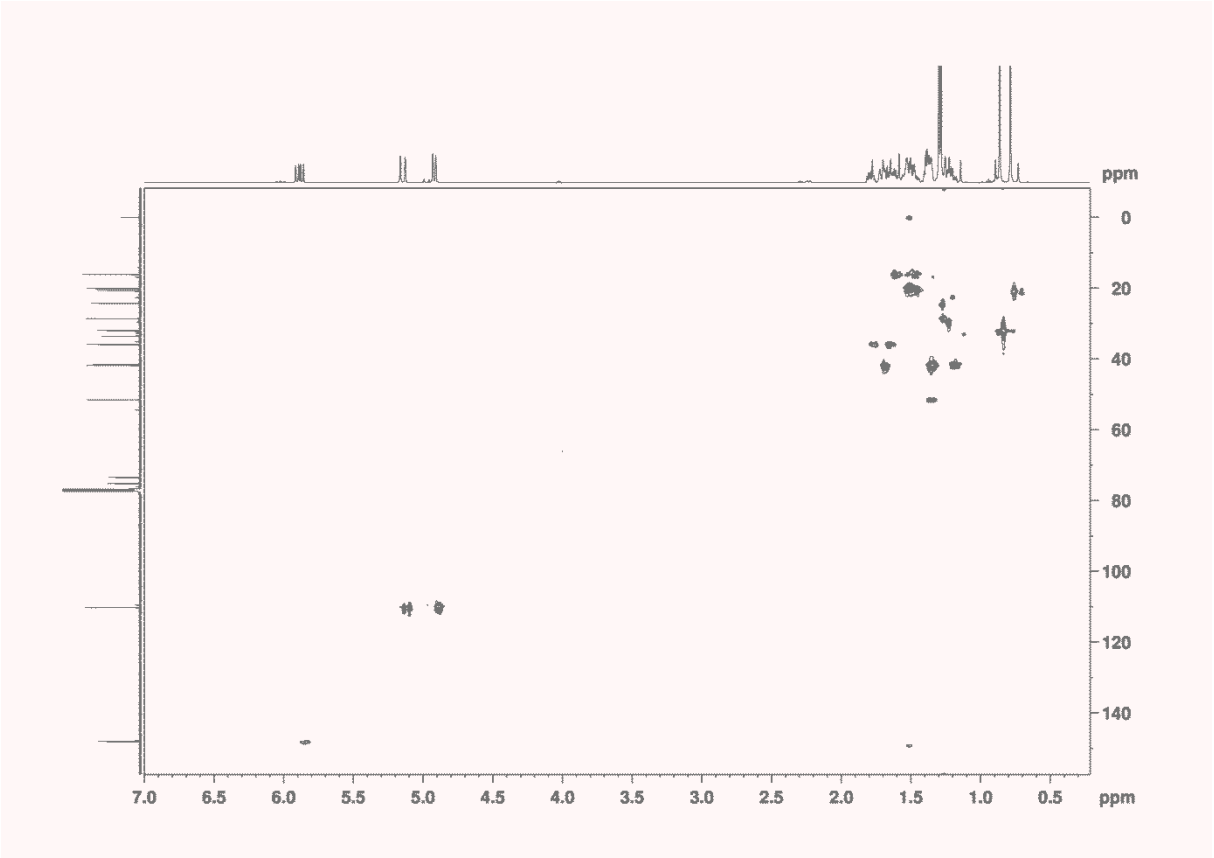
5.7.3 ^{13}C NMR spectrum (-)-caparrapioxide



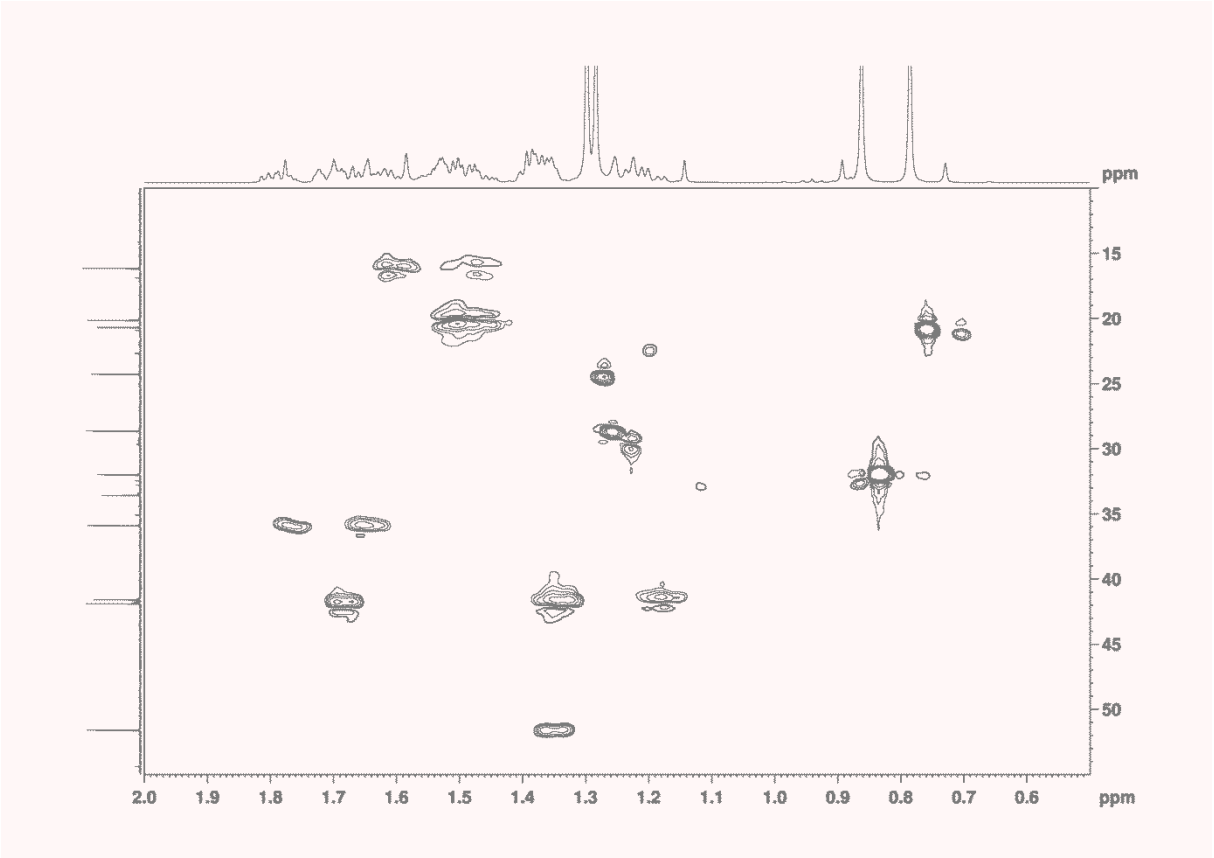
5.7.4 COSY NMR spectrum (-)-caparrapioxide



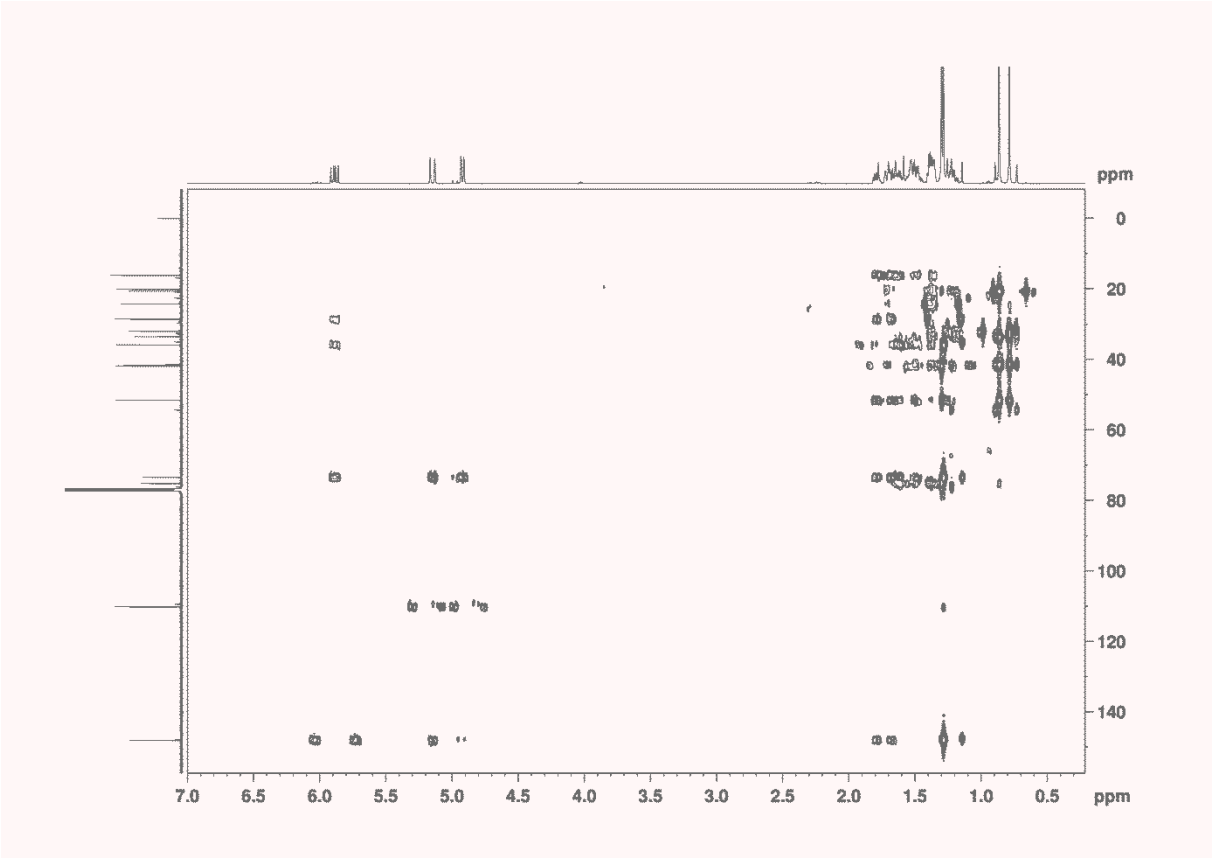
5.7.5 HSQC NMR spectrum (-)-caparrapioxide



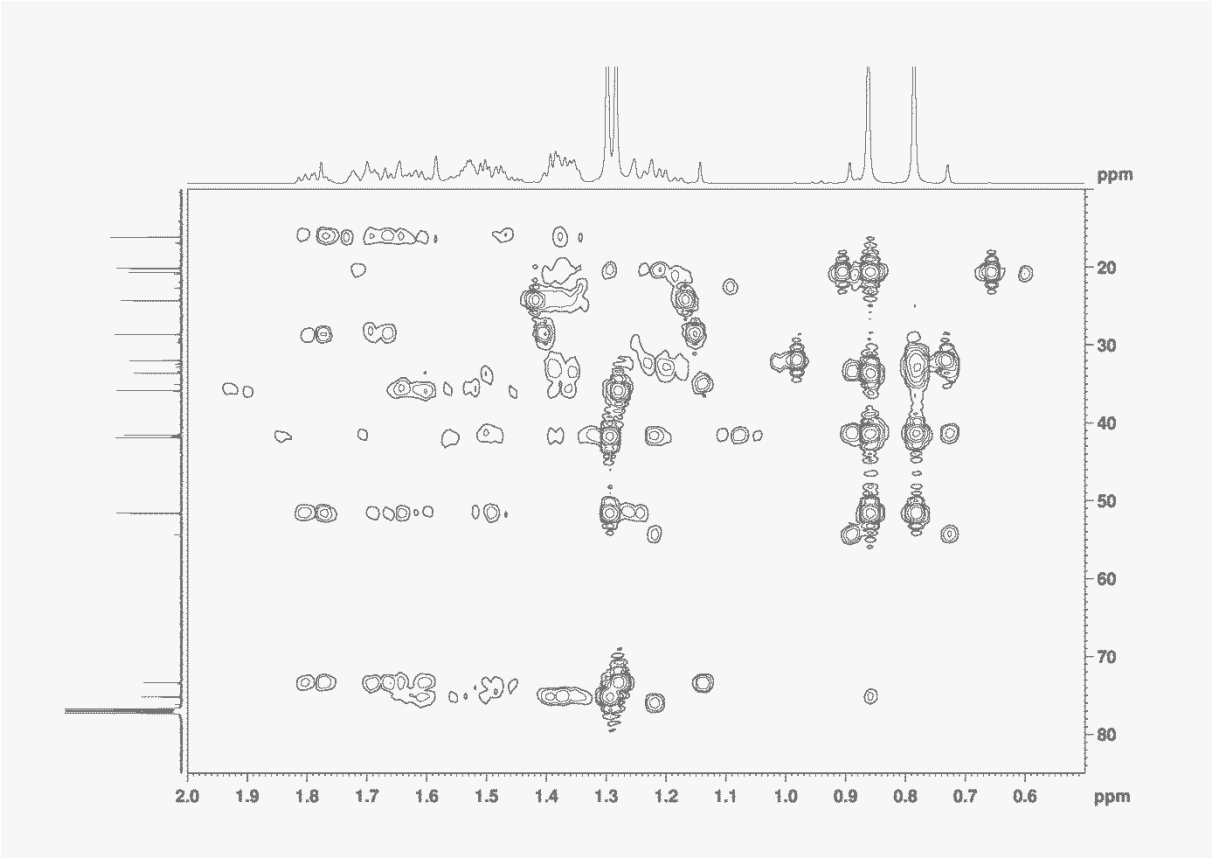
5.7.6 HSQC NMR spectrum (-)-caparrapioxide zoomed



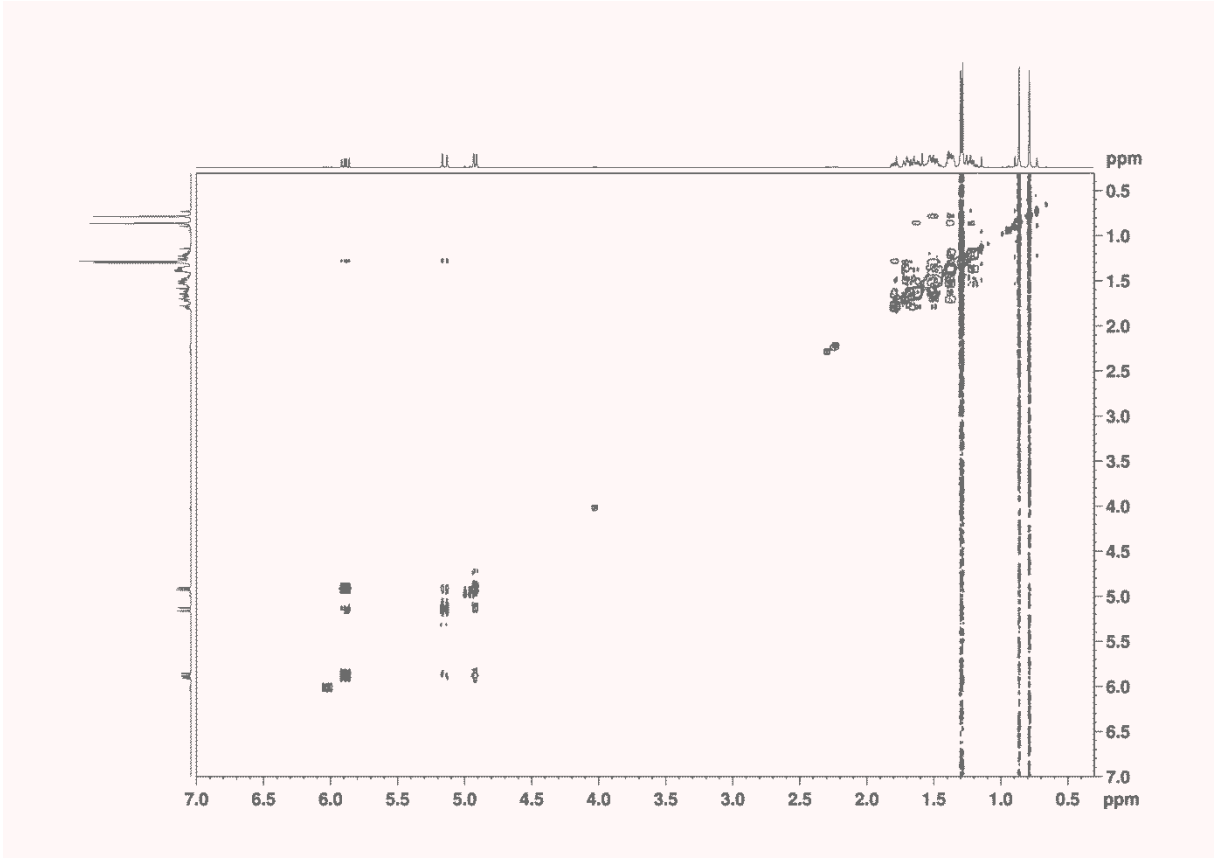
5.7.7 HMBC NMR spectrum (-)-caparrapioxide



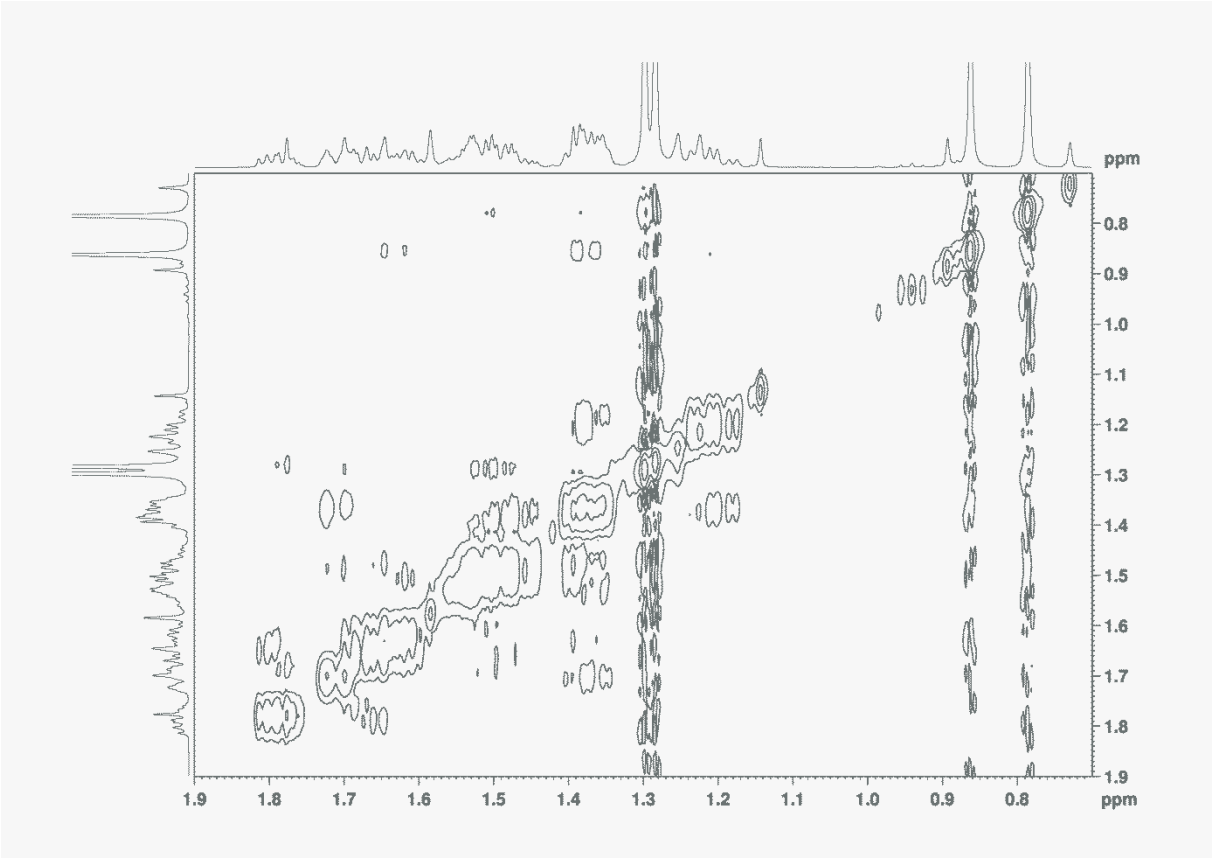
5.7.8 HMBC NMR spectrum (-)-caparrapioxide zoomed



5.7.9 NOESY NMR spectrum (-)-caparrapioxide

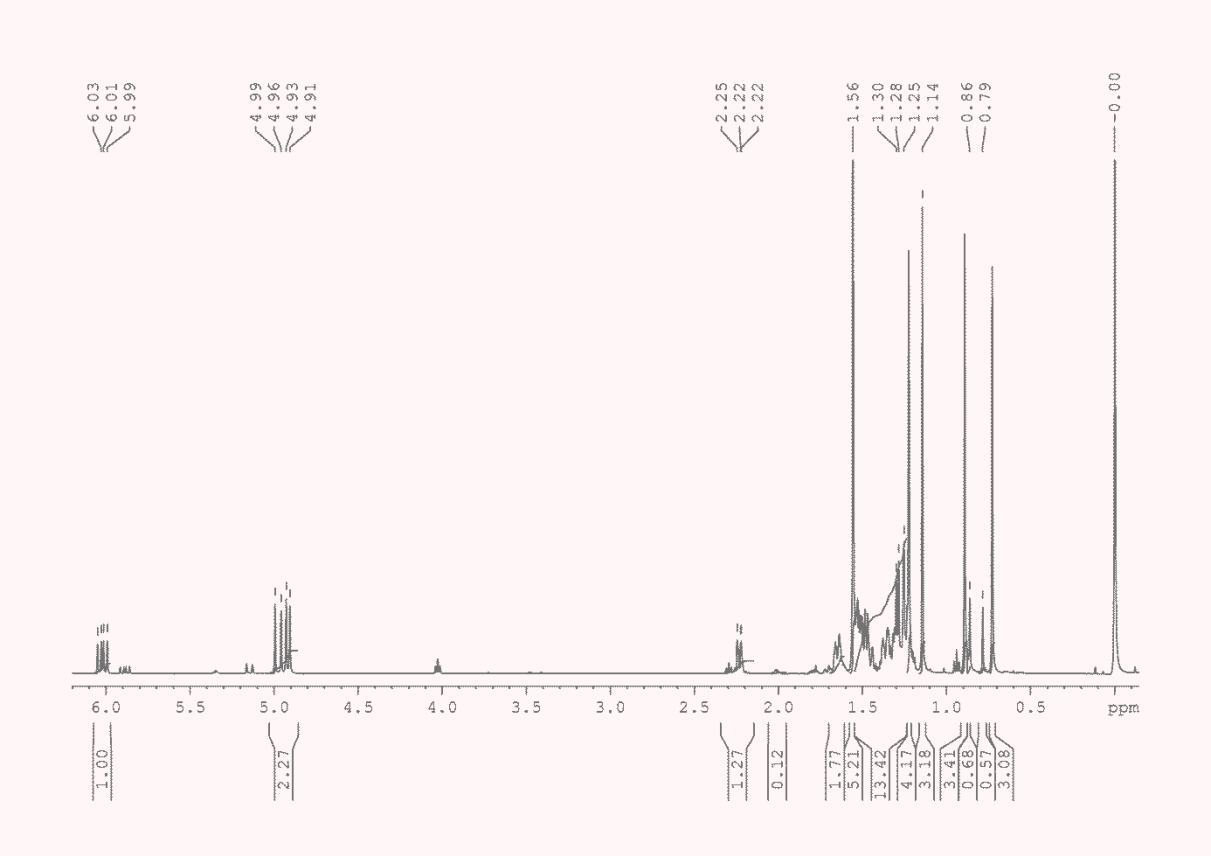


5.7.10NOESY NMR spectrum (-)-caparrapioxide zoomed

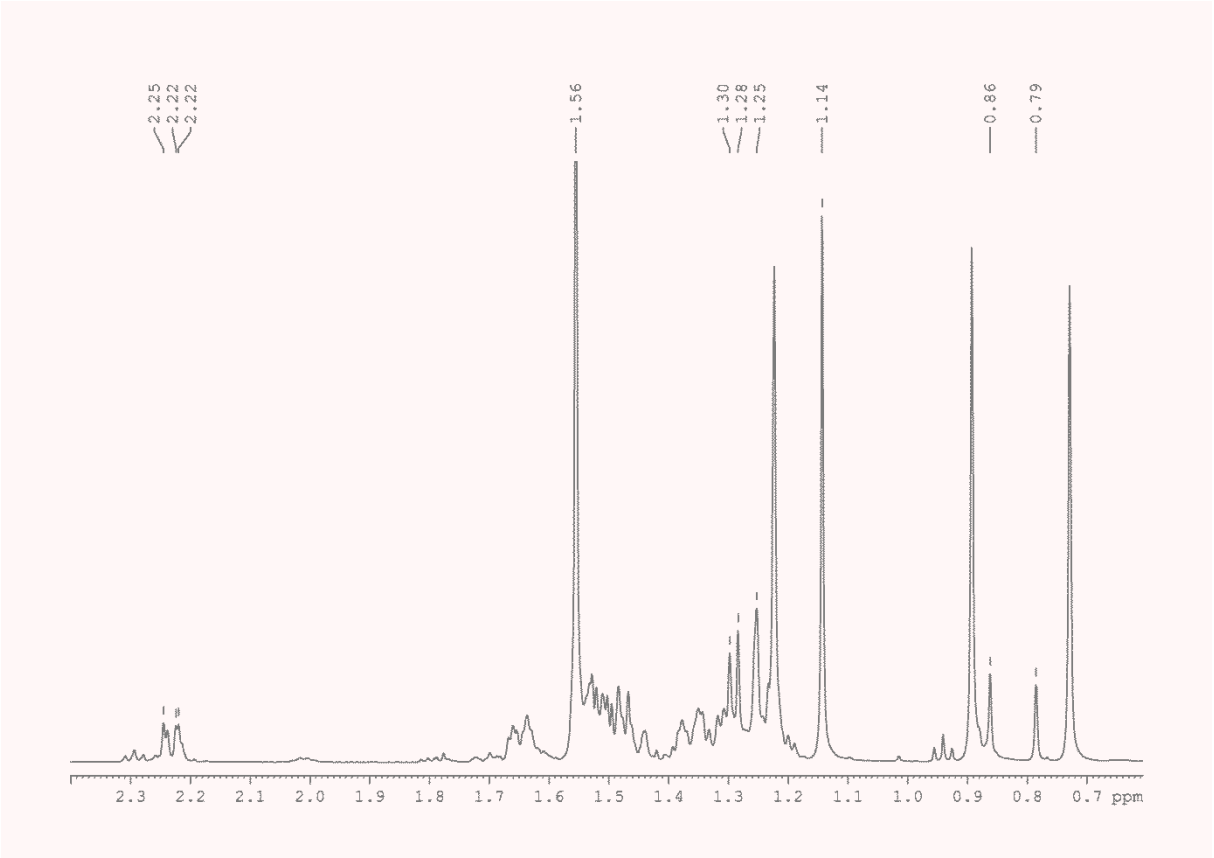


5.8 NMR spectra (-)-8-epi-caparrapioxide

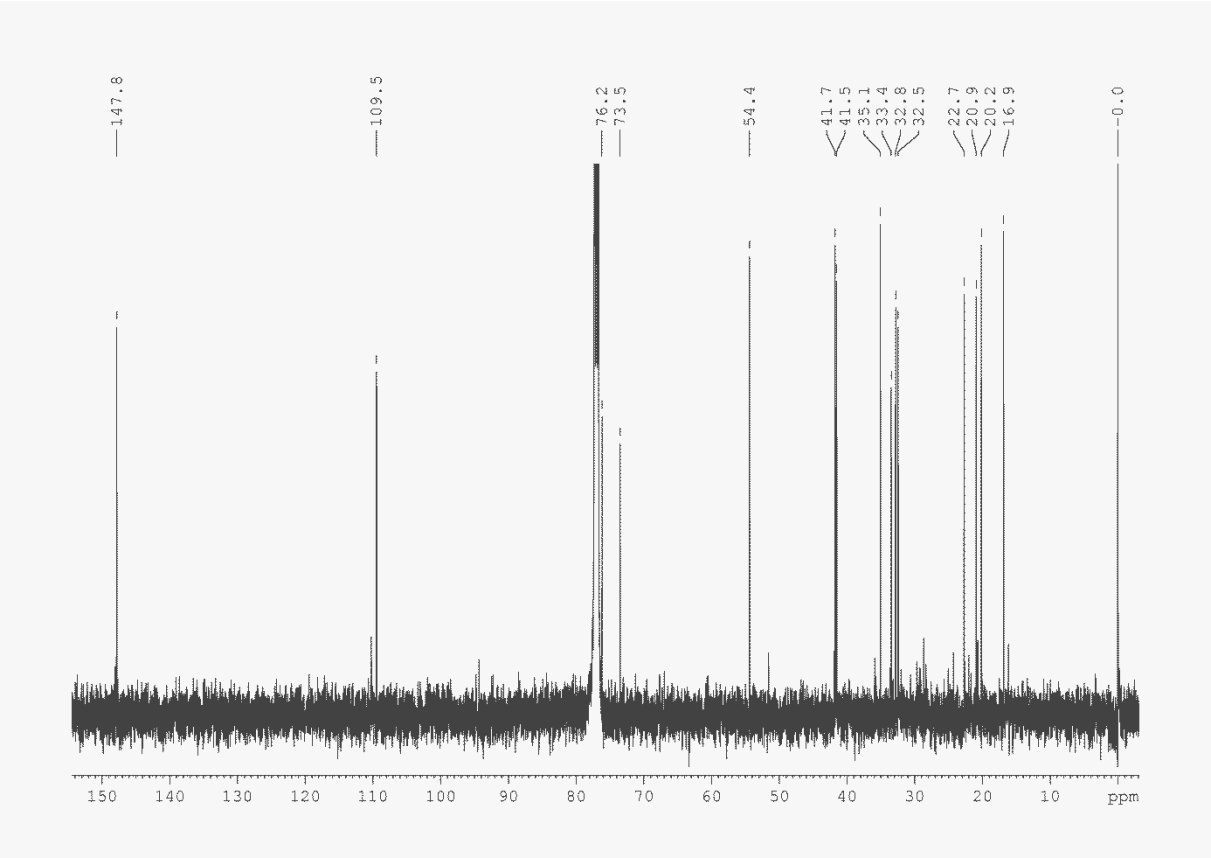
5.8.1 ^{1}H NMR spectrum (-)—8-*epi*-caparrapioxide



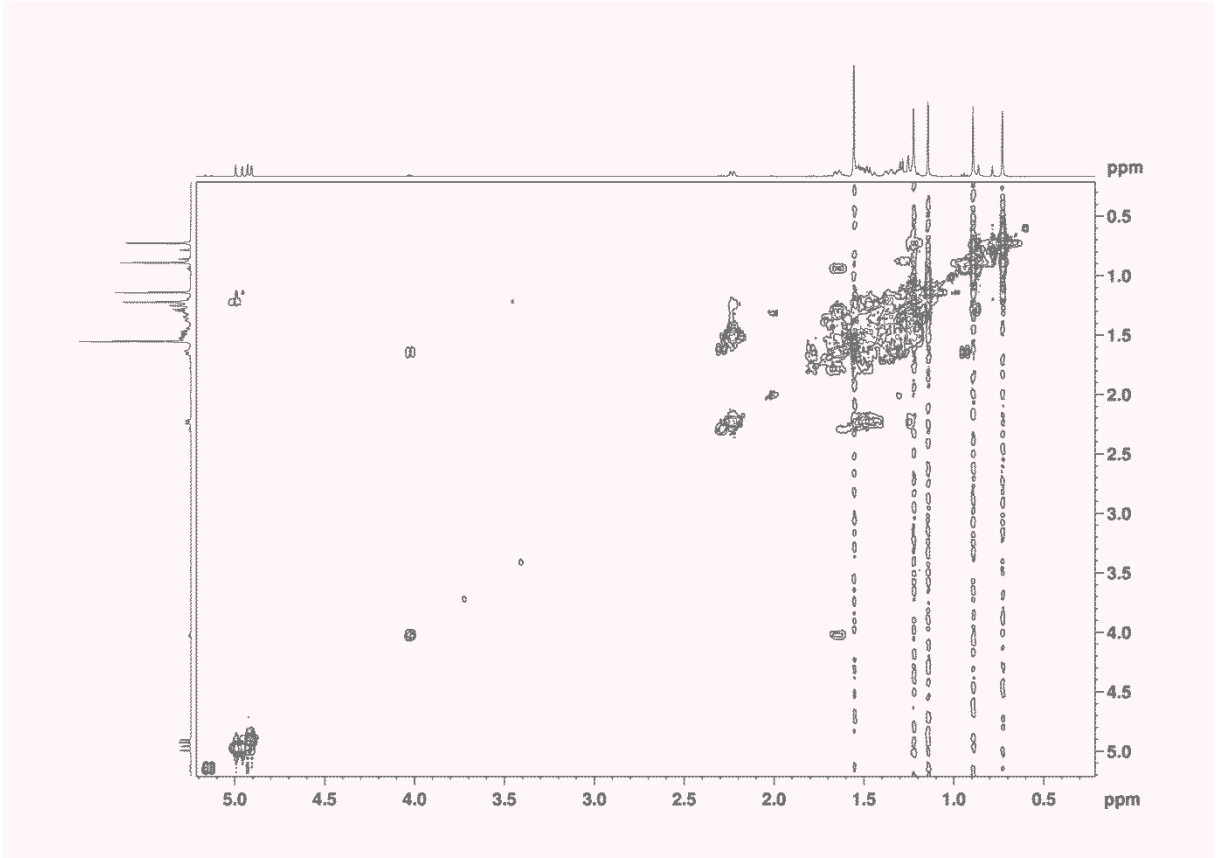
5.8.2 ^{1}H NMR spectrum (-)—8-*epi*-caparrapioxide zoomed



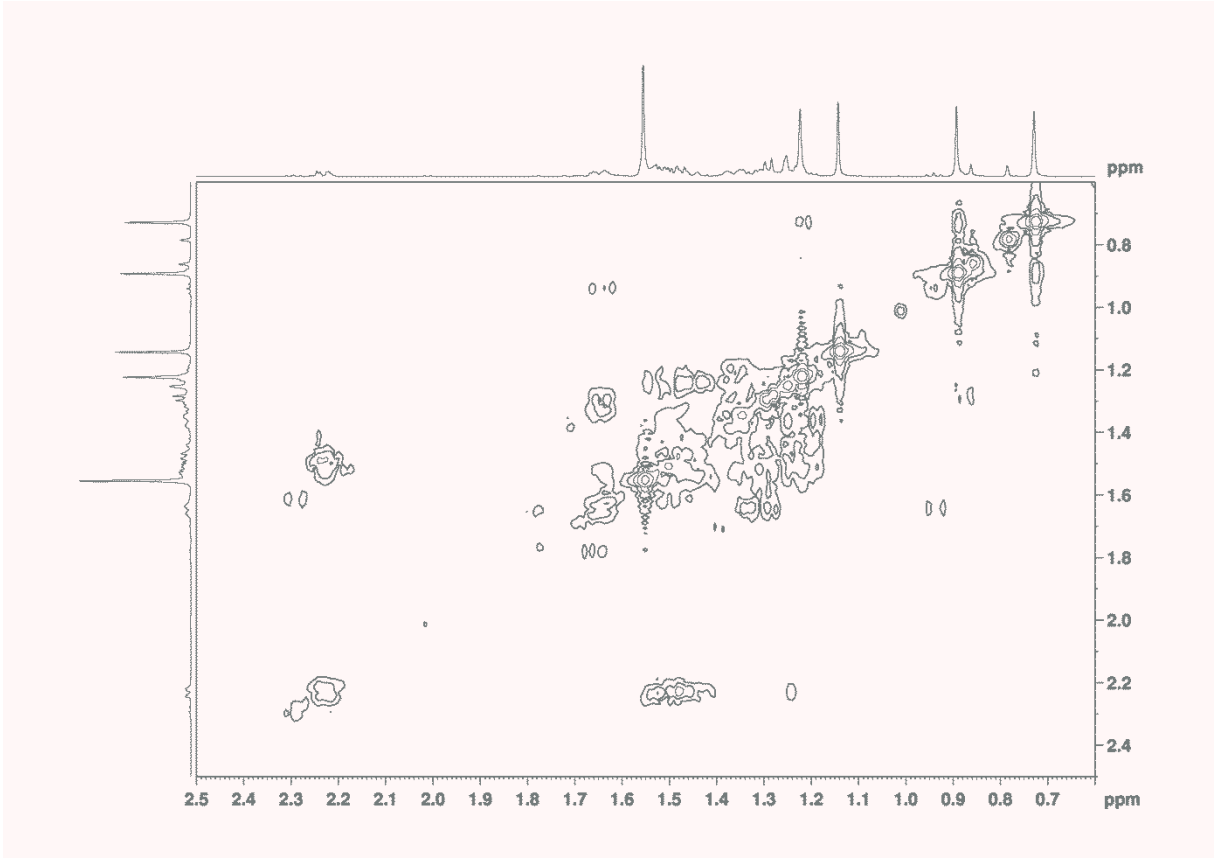
5.8.3 ^{13}C NMR spectrum (-)—8-*epi*-caparrapioxide



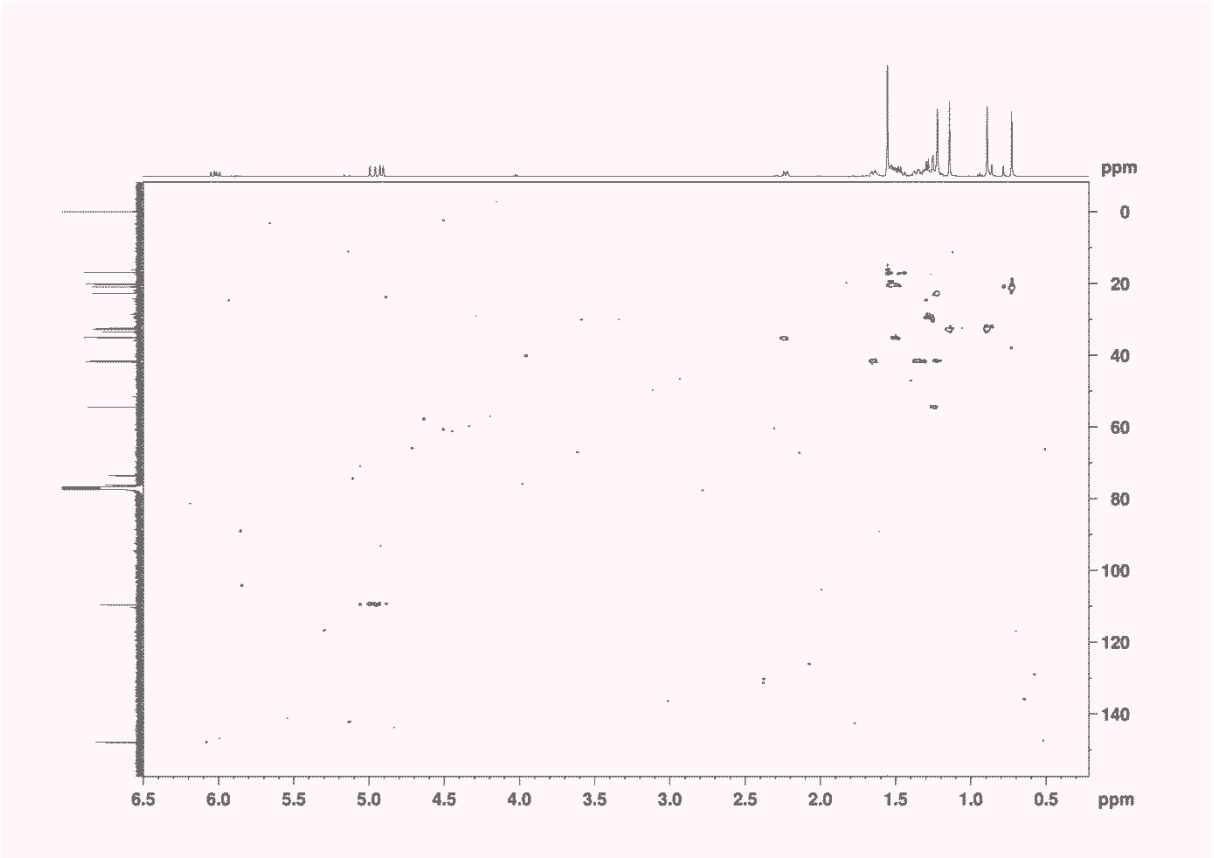
5.8.4 COSY NMR spectrum (-)—8-*epi*-caparrapioxide



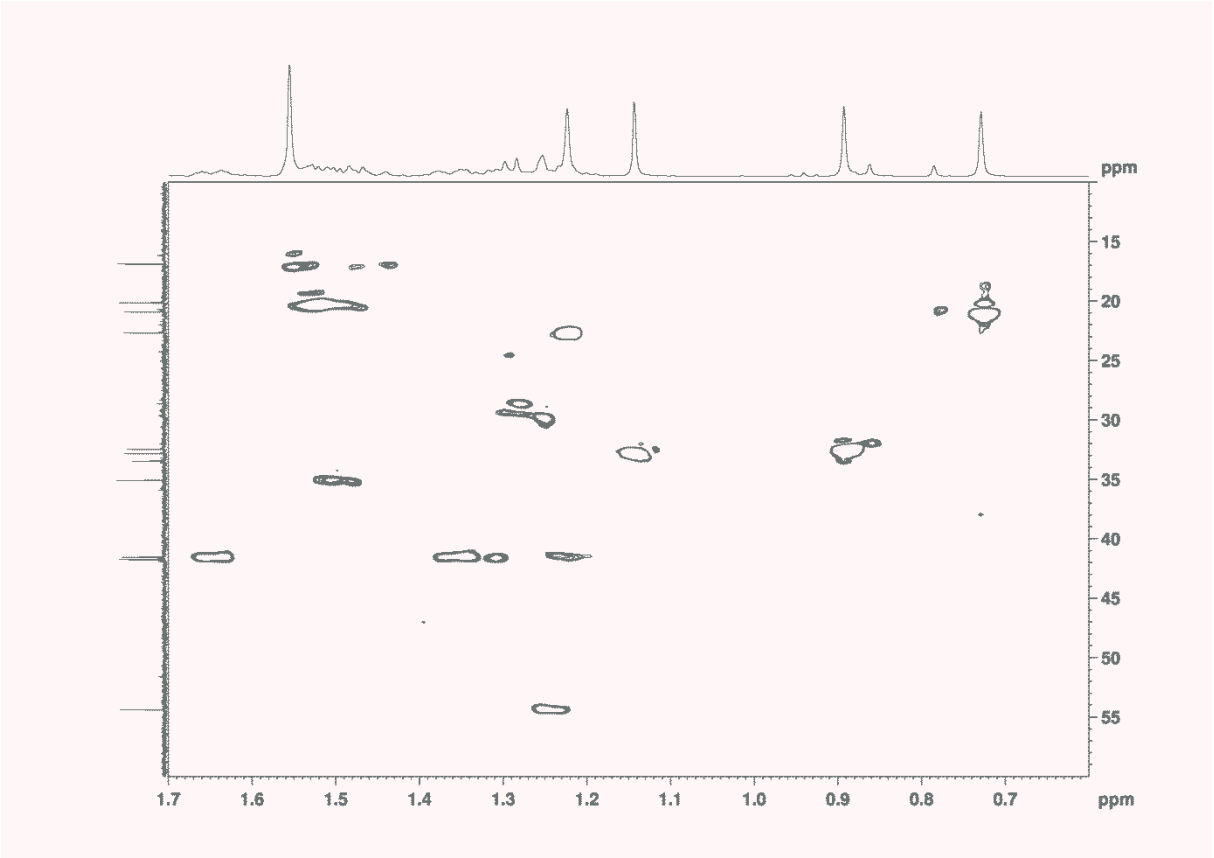
5.8.5 COSY NMR spectrum (-)—8-*epi*-caparrapioxide zoomed



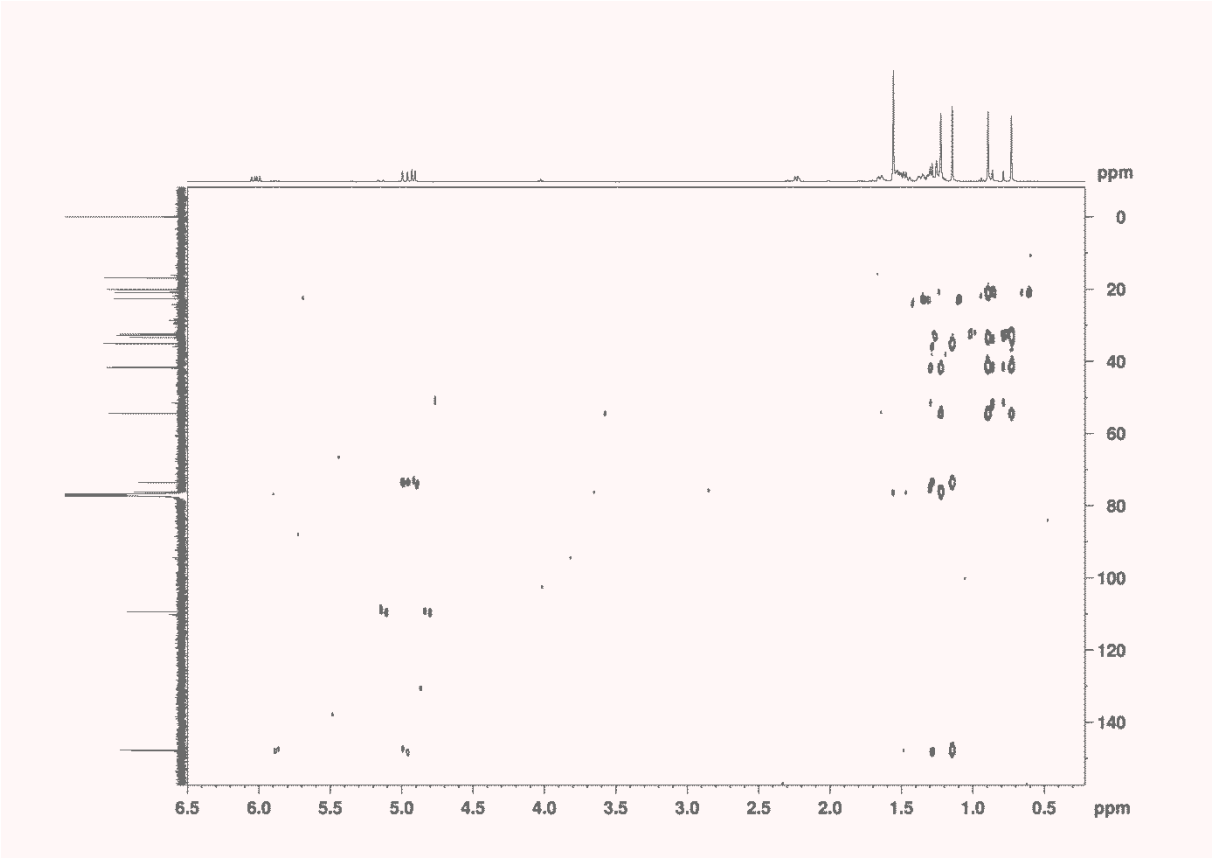
5.8.6 HSQC NMR spectrum (-)—8-*epi*-caparrapioxide



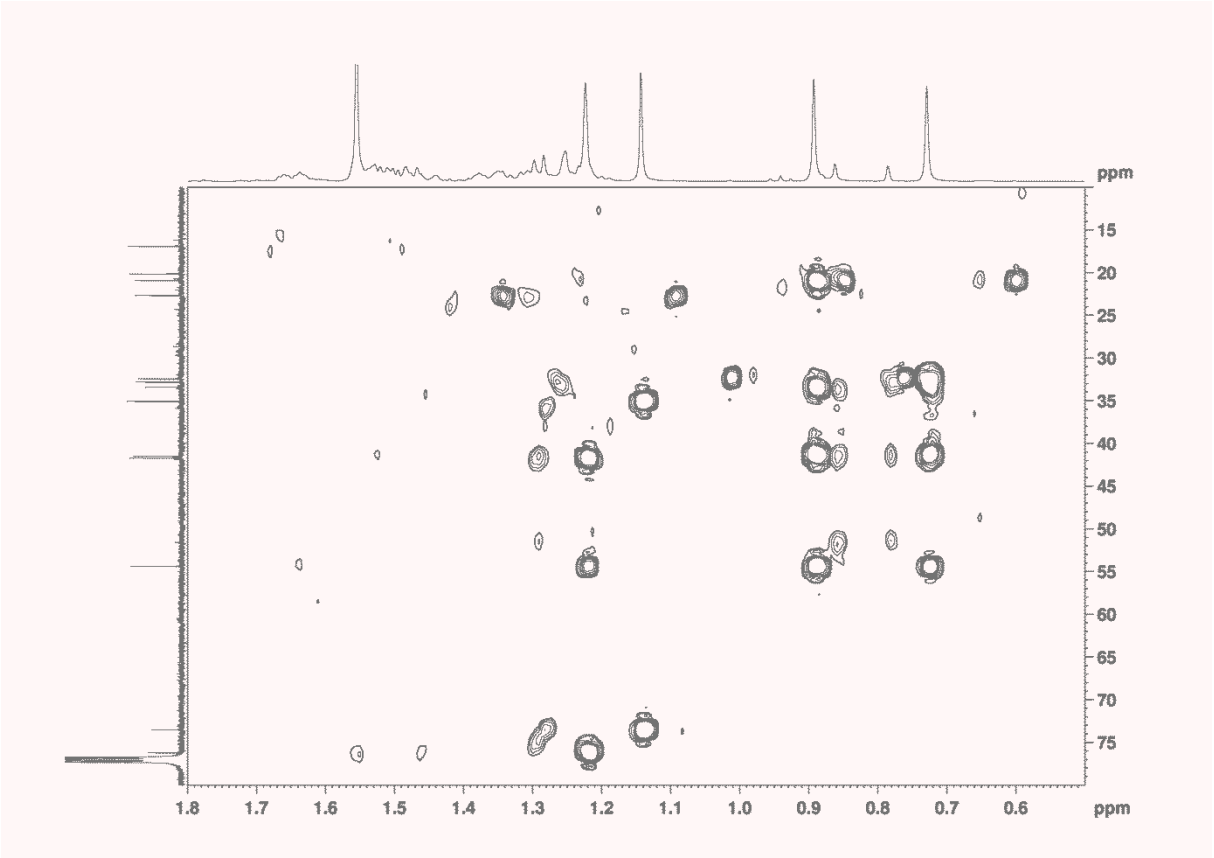
5.8.7 HSQC NMR spectrum (-)—8-*epi*-caparrapioxide zoomed



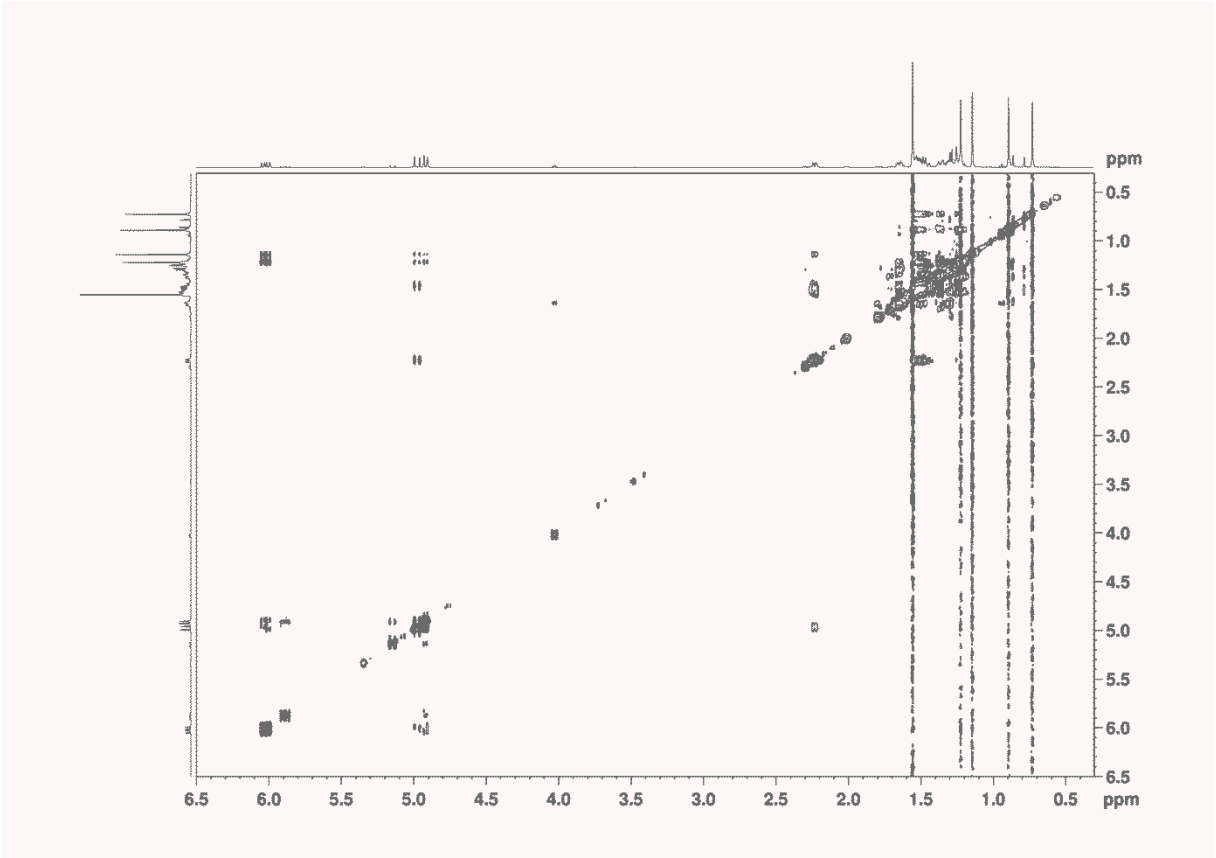
5.8.8 HMBC NMR spectrum (-)—8-*epi*-caparrapioxide



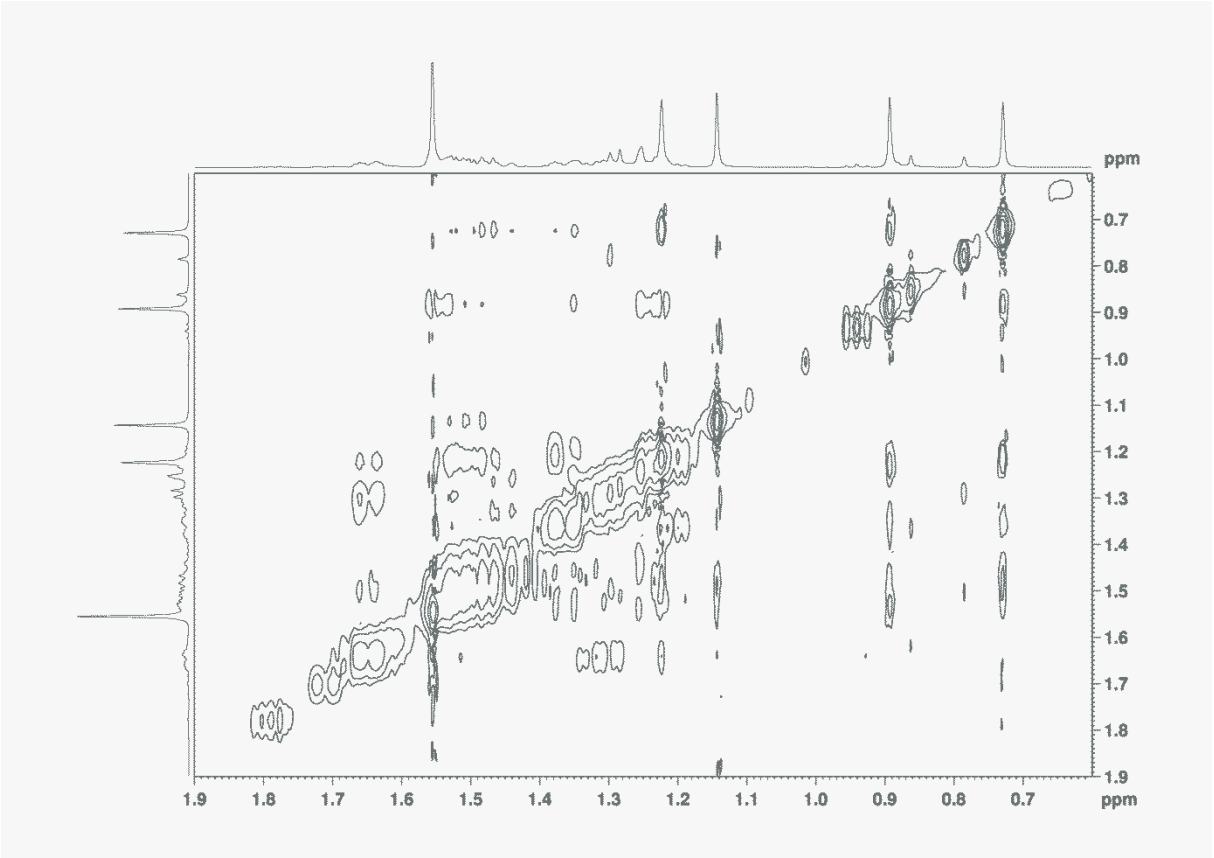
5.8.9 HMBC NMR spectrum (-)—8-*epi*-caparrapioxide zoomed



5.8.10NOESY NMR spectrum (-)—8-*epi*-caparrapioxide



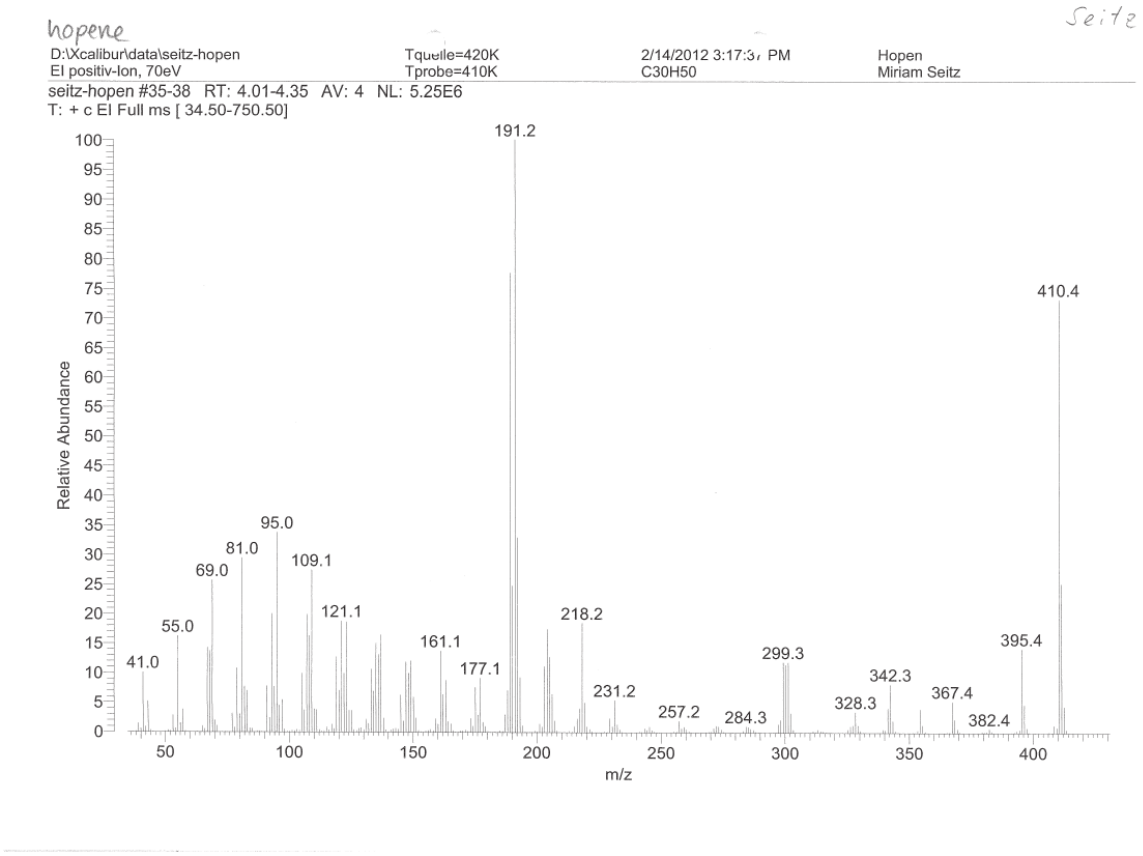
5.8.11NOESY NMR spectrum (-)—8-*epi*-caparrapioxide zoomed



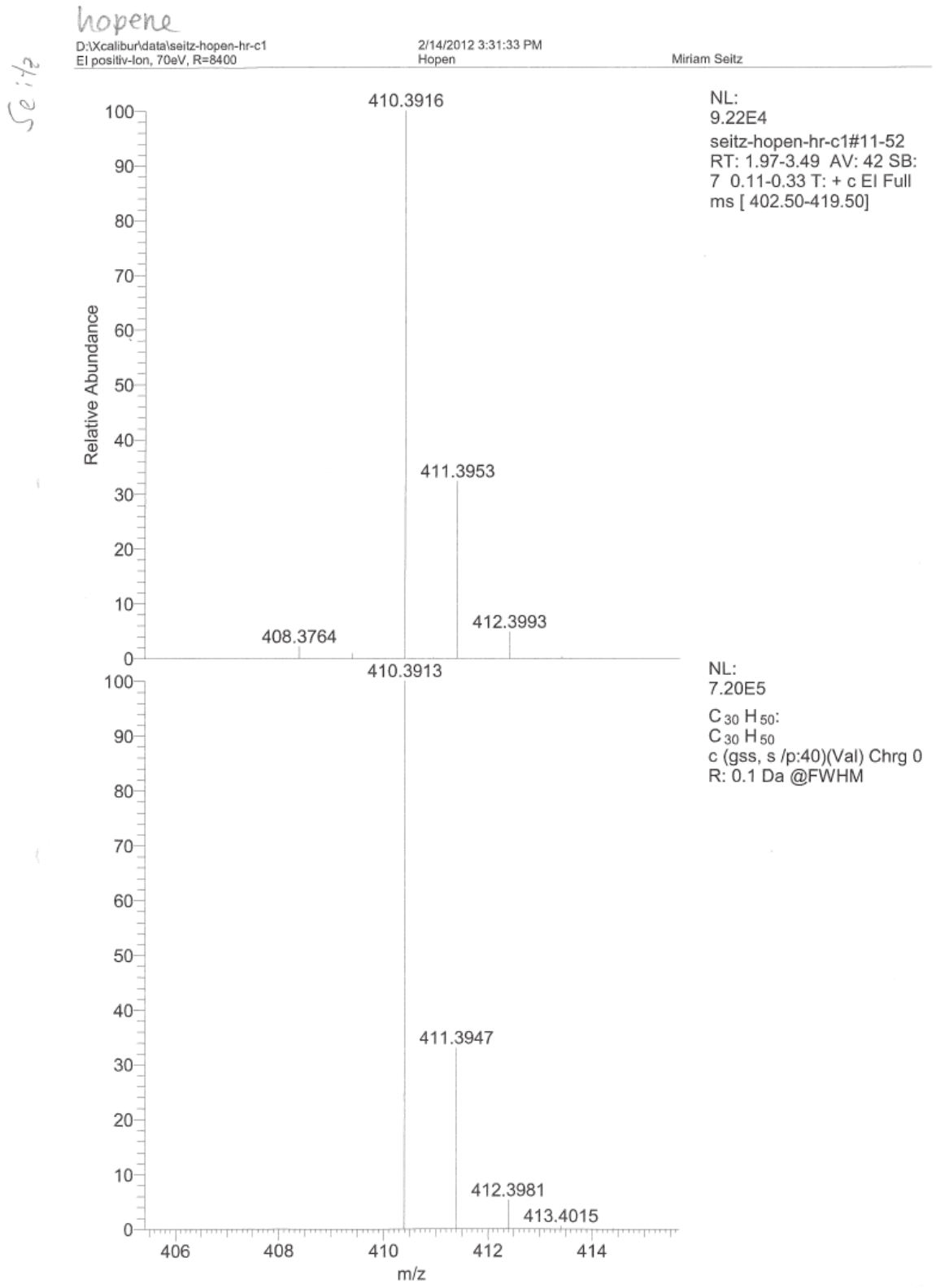
6 HREIMS spectra

6.1 HREIMS spectra hopene

6.1.1 Full HREIMS spectrum hopene

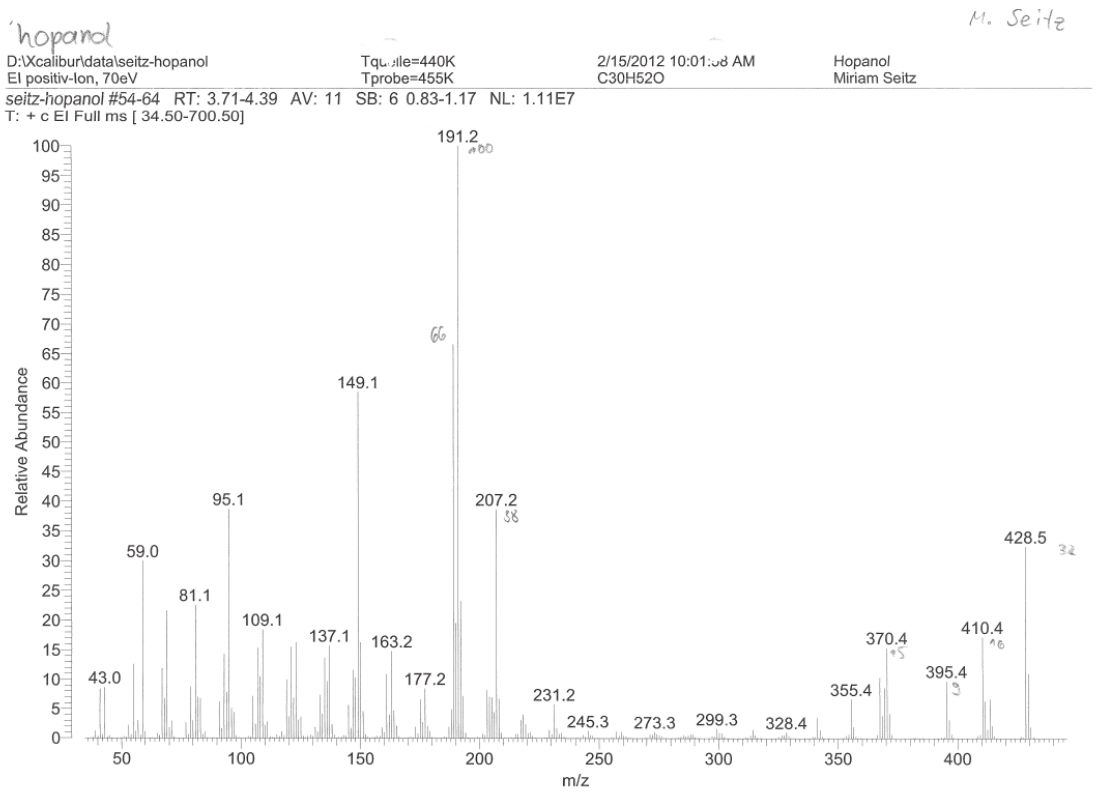


6.1.2 Molecule HREIMS spectrum hopene

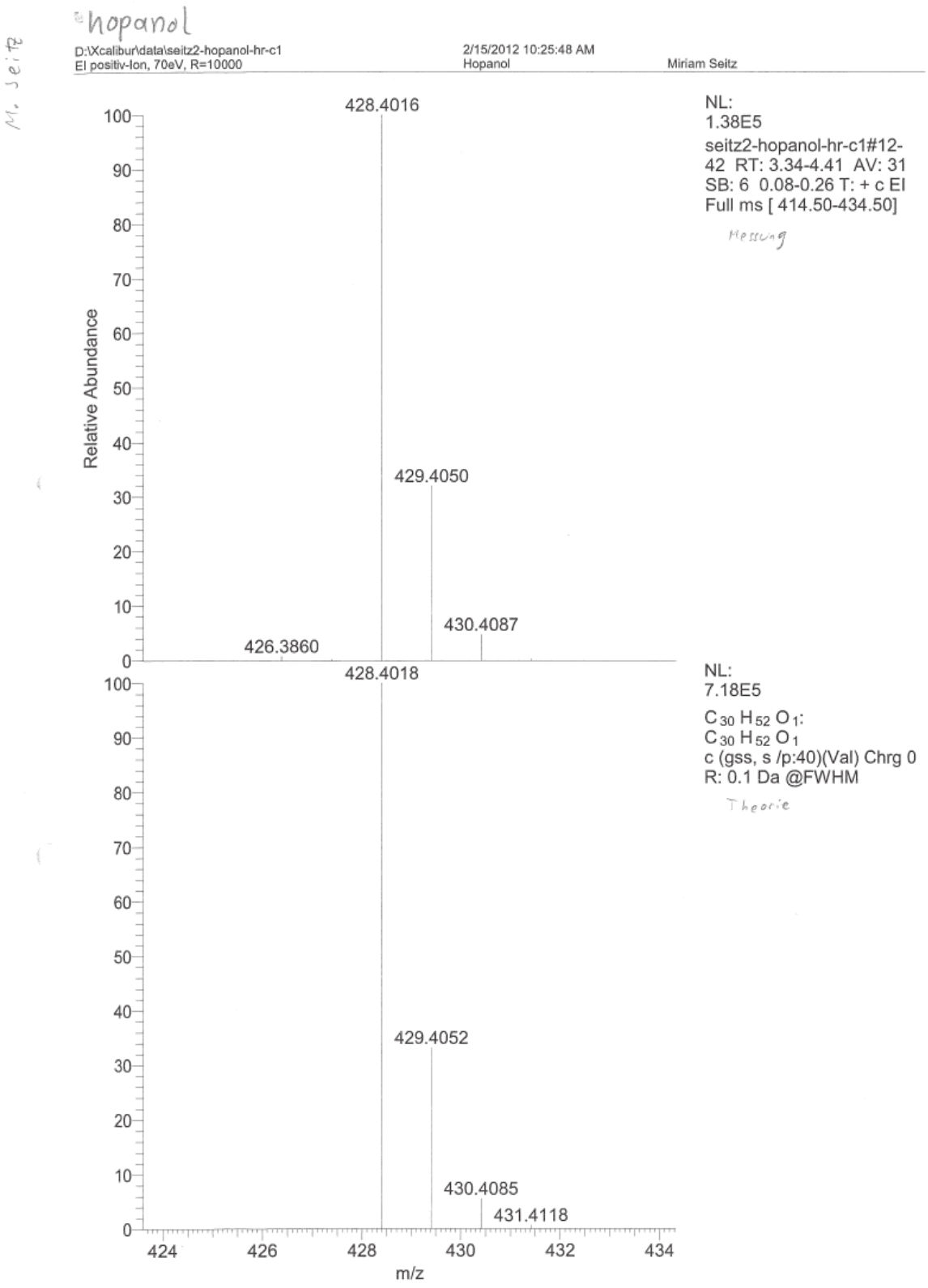


6.2 HREIMS spectra hopanol

6.2.1 Full HREIMS spectrum hopanol

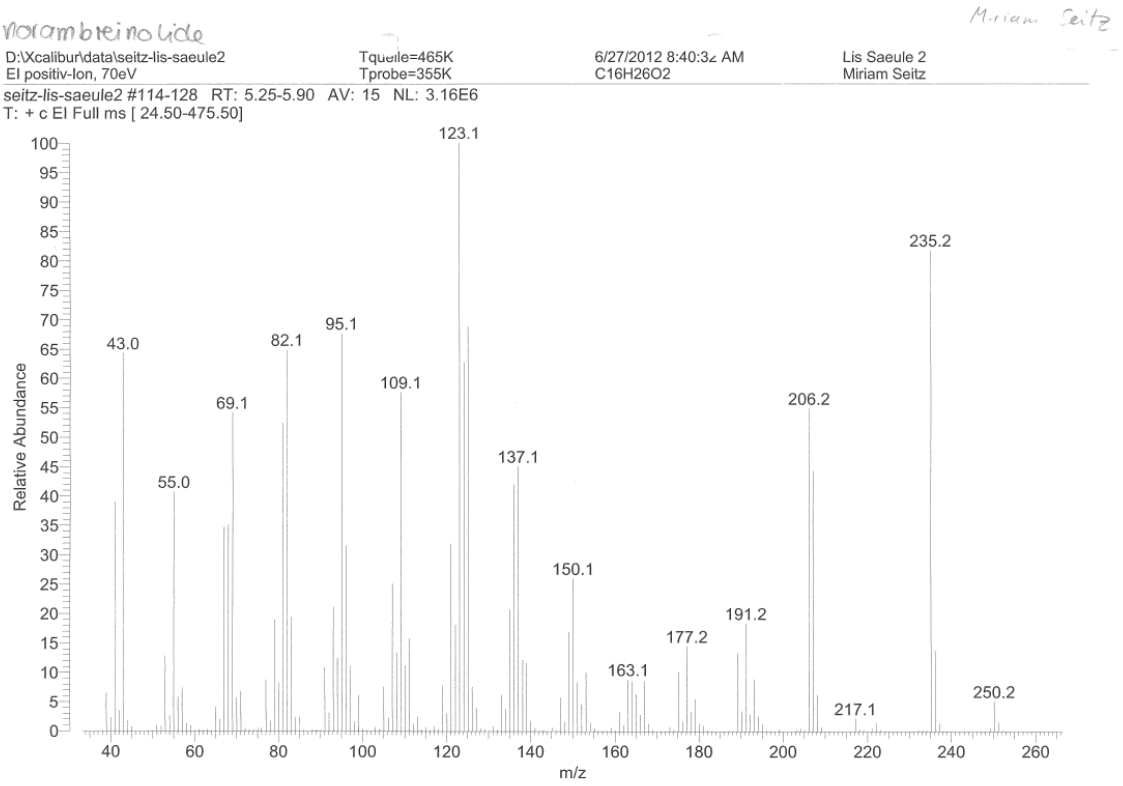


6.2.2 Molecule HREIMS spectrum hopanol

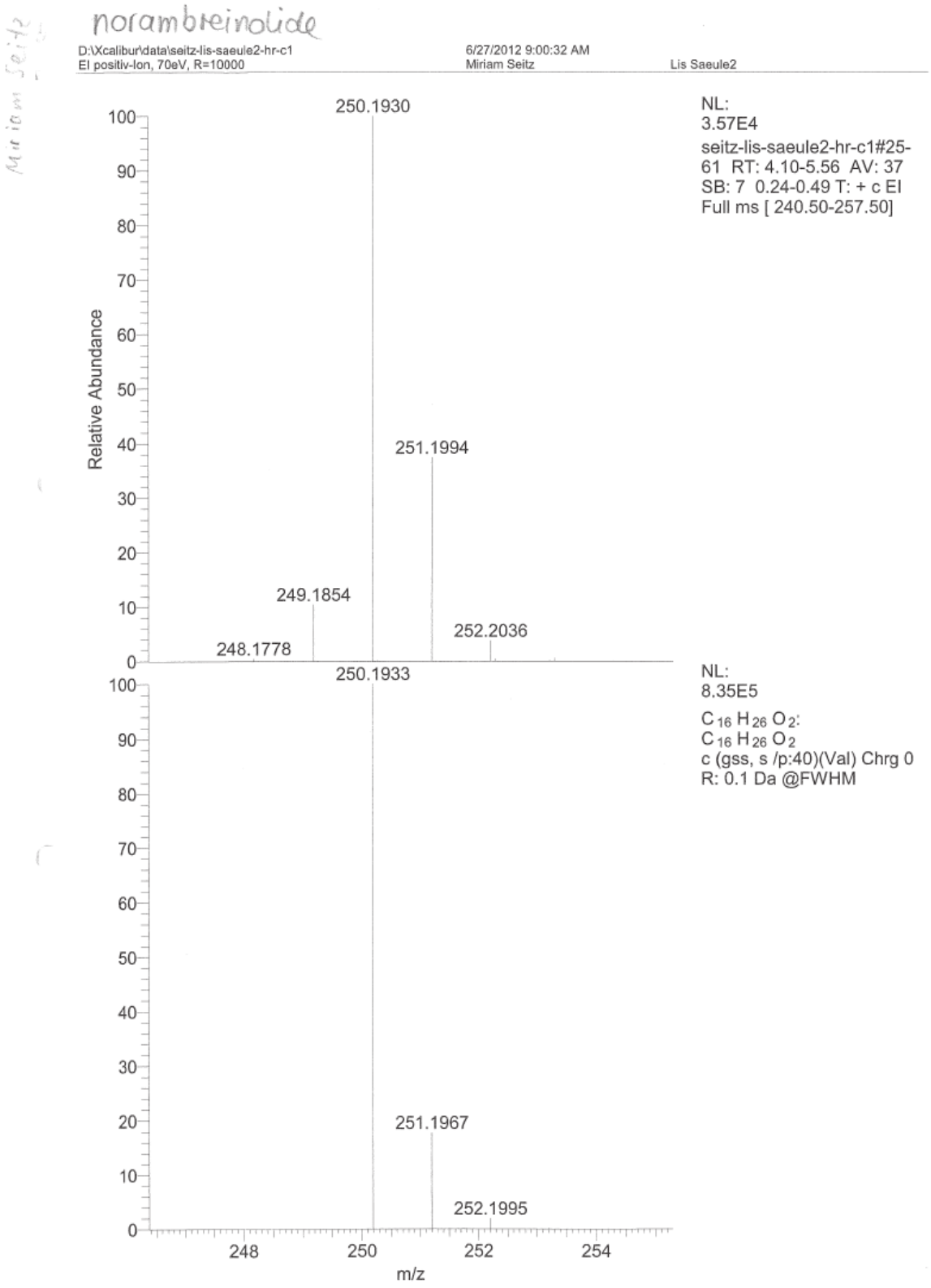


6.3 HREIMS spectra sclareolide

6.3.1 Full HREIMS spectrum sclareolide

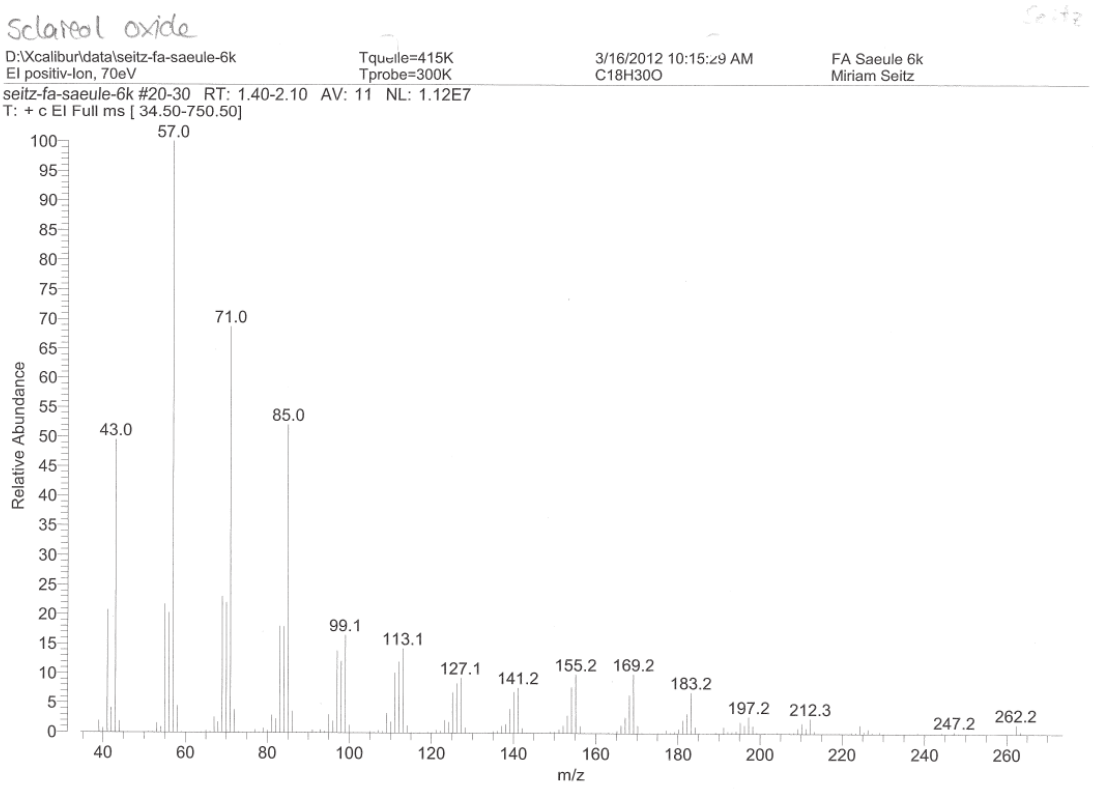


6.3.2 Molecule spectrum HREIMS sclareolide

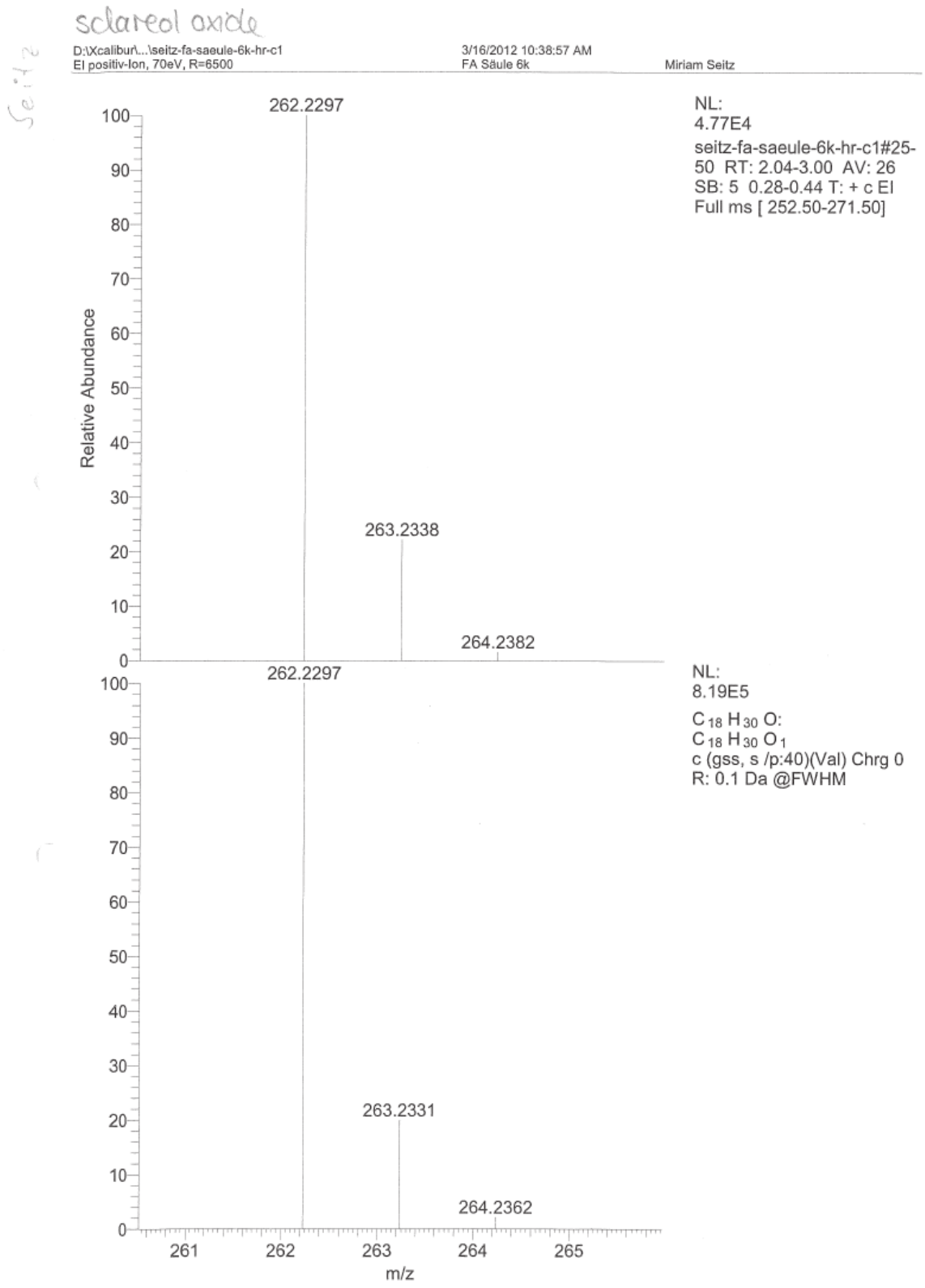


6.4 HREIMS spectra sclareoloxide

6.4.1 Full HREIMS spectrum sclareoloxide

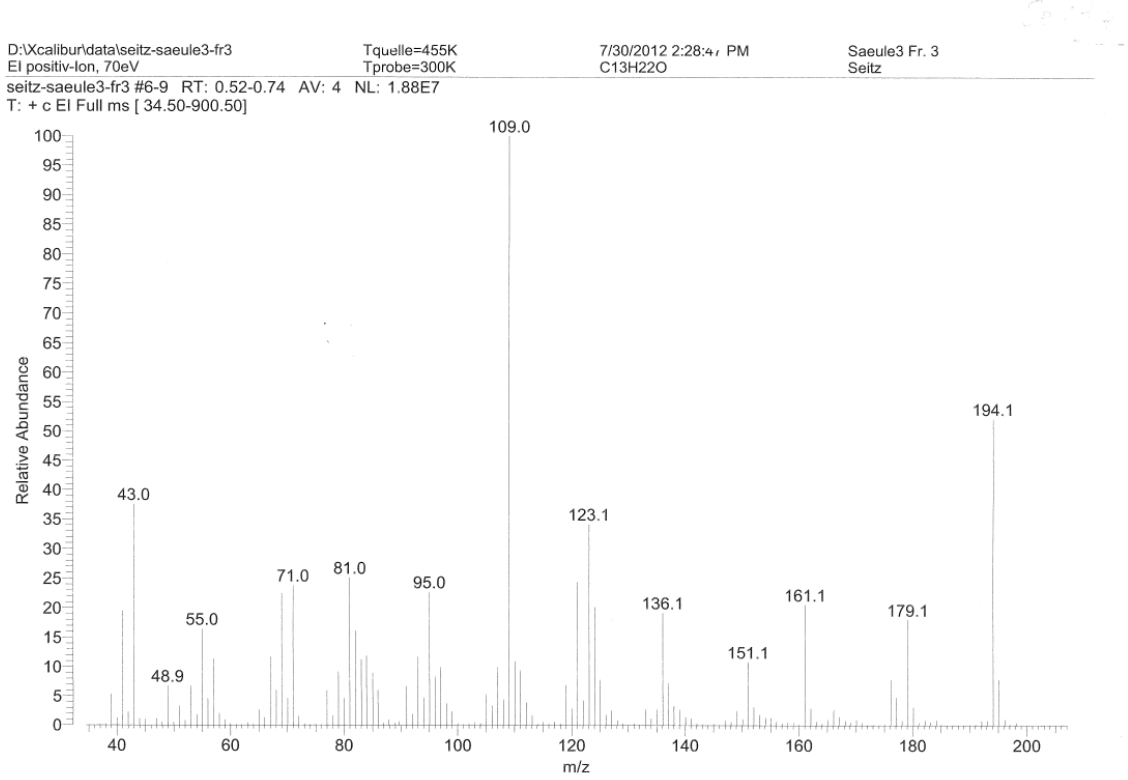


6.4.2 Molecule HREIMS spectrum sclareoloxide

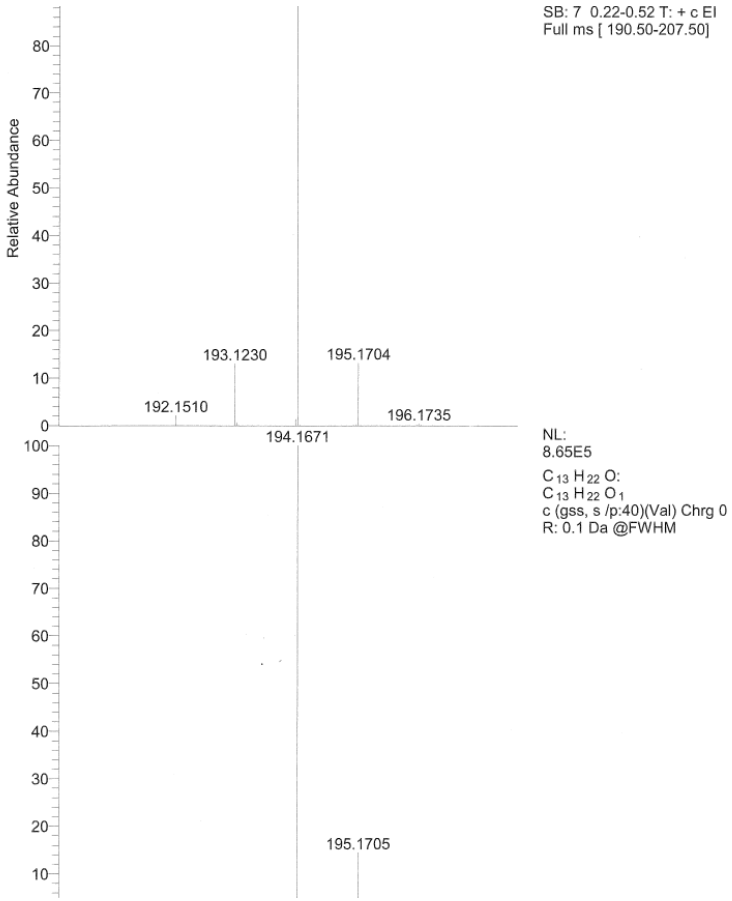


6.5 HREIMS spectra hexahydrochromene

6.5.1 Full HREIMS spectrum hexahydrochromene

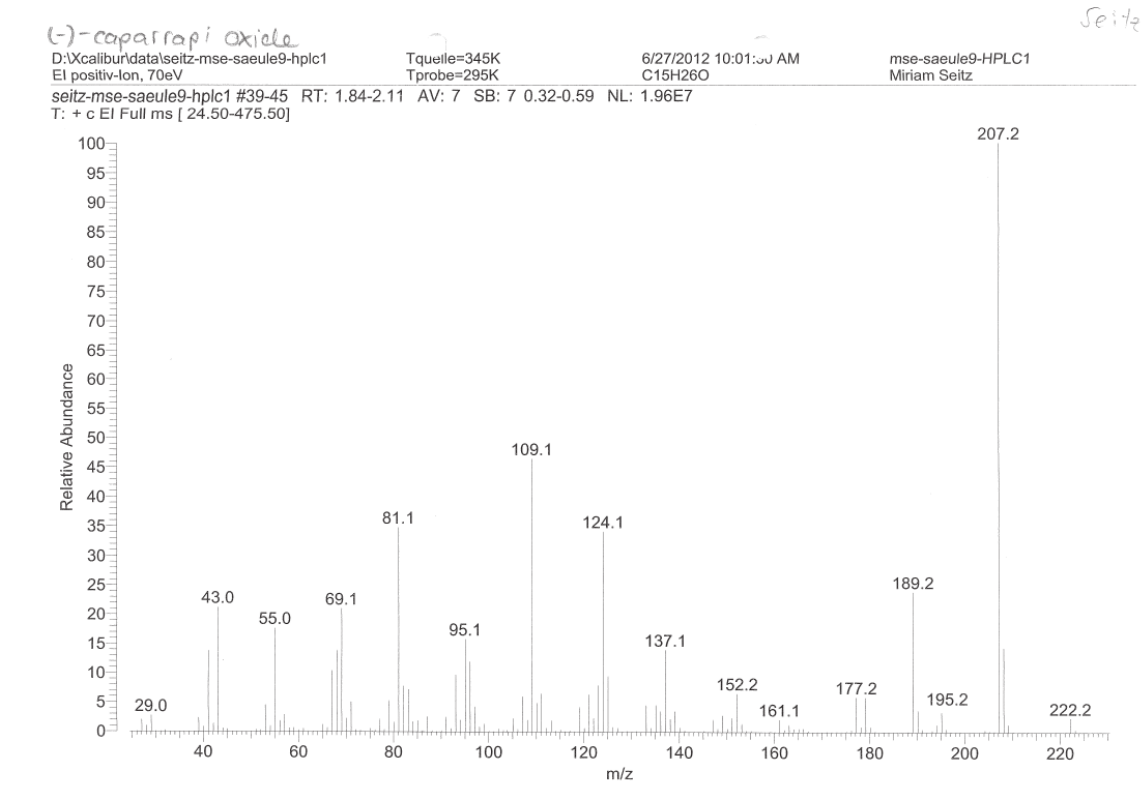


6.5.2 Molecule HREIMS spectrum hexahydrochromene

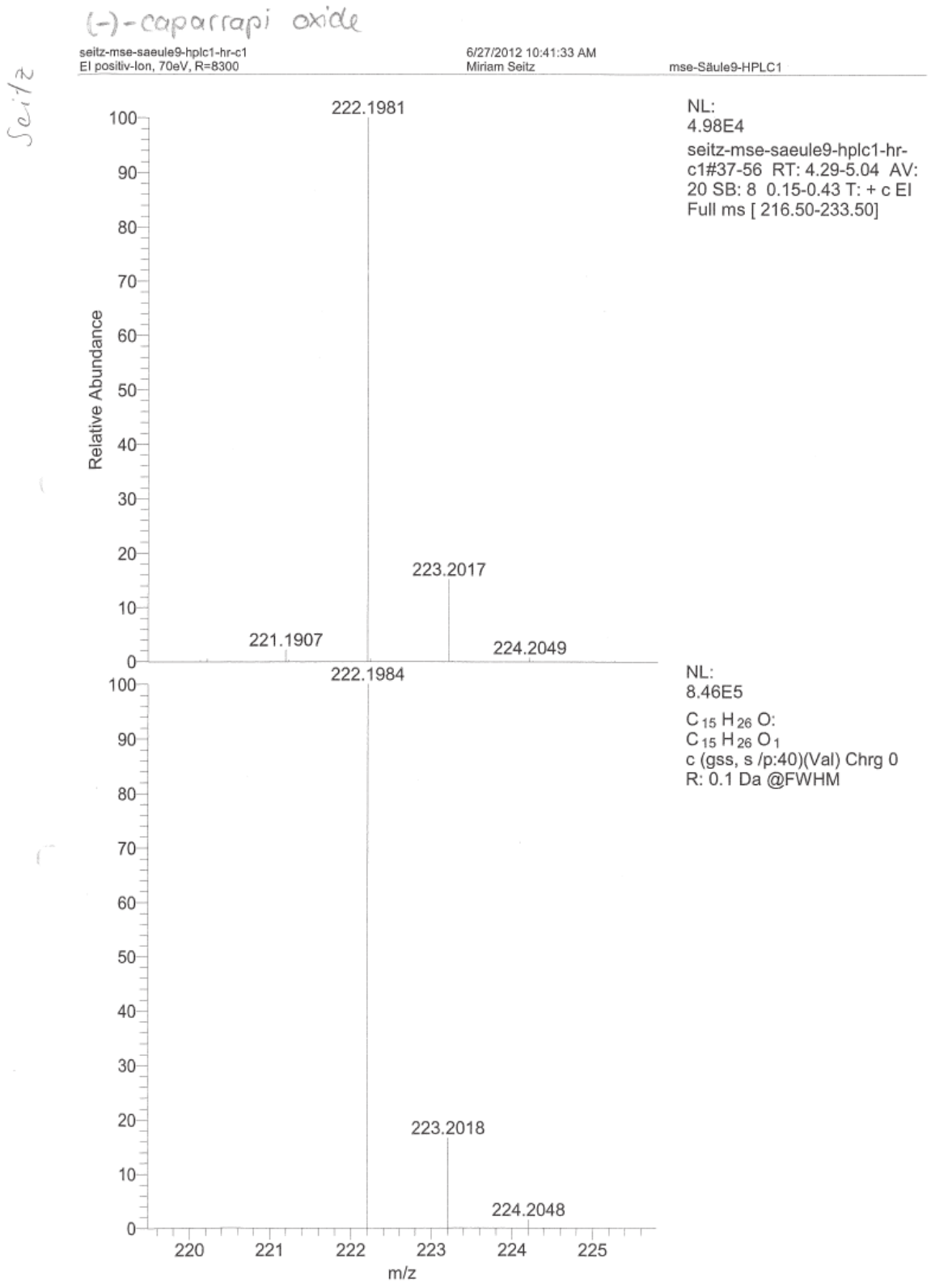


6.6 HREIMS spectra (-)-caparrapioxide

6.6.1 Full HREIMS spectrum (-)-caparrapioxide

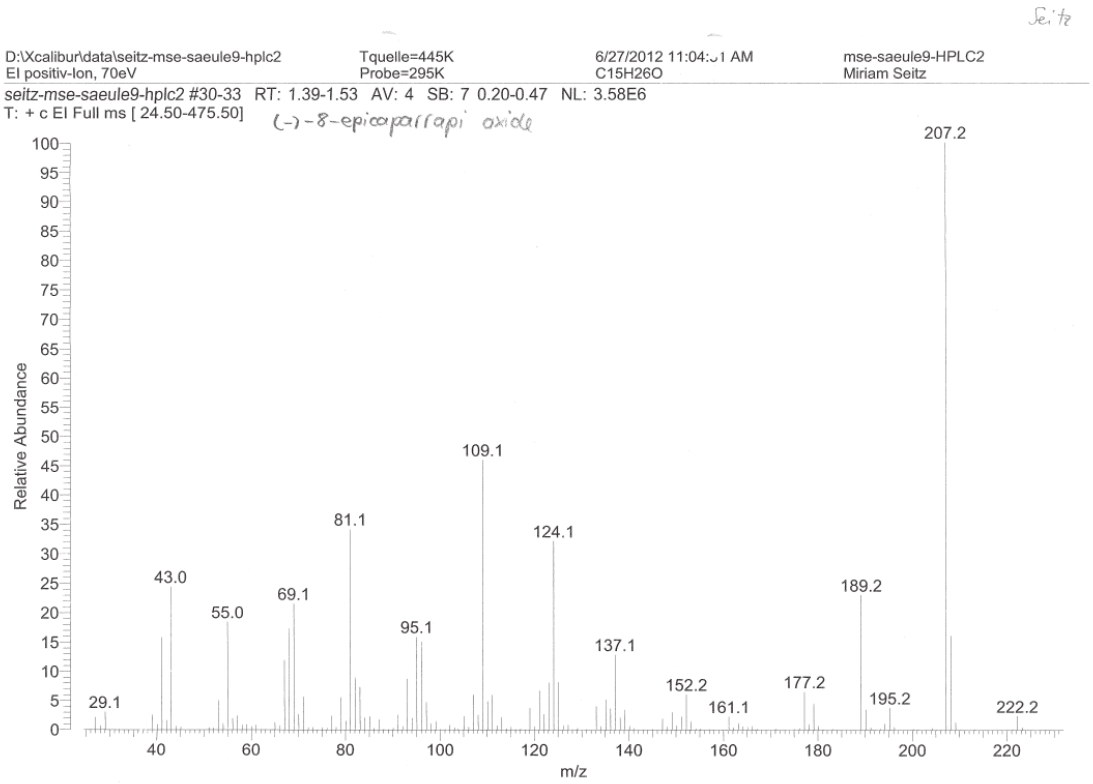


6.6.2 Molecule spectrum HREIMS (-)-caparrapioxide

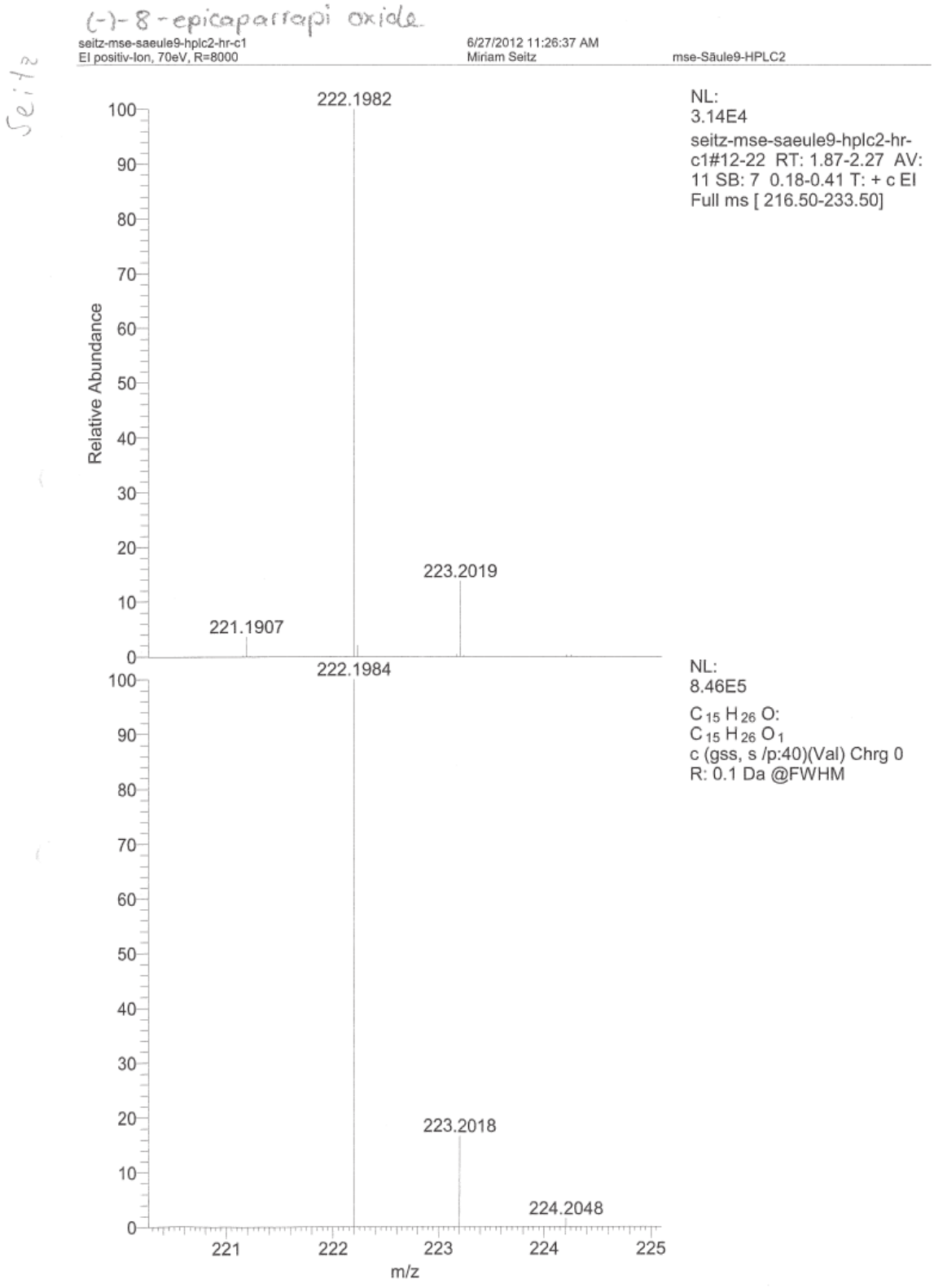


6.7 HREIMS spectra (-)-8-epi-caparrapioxide

6.7.1 Full HREIMS spectrum (-)—8-*epi*-caparrapioxide

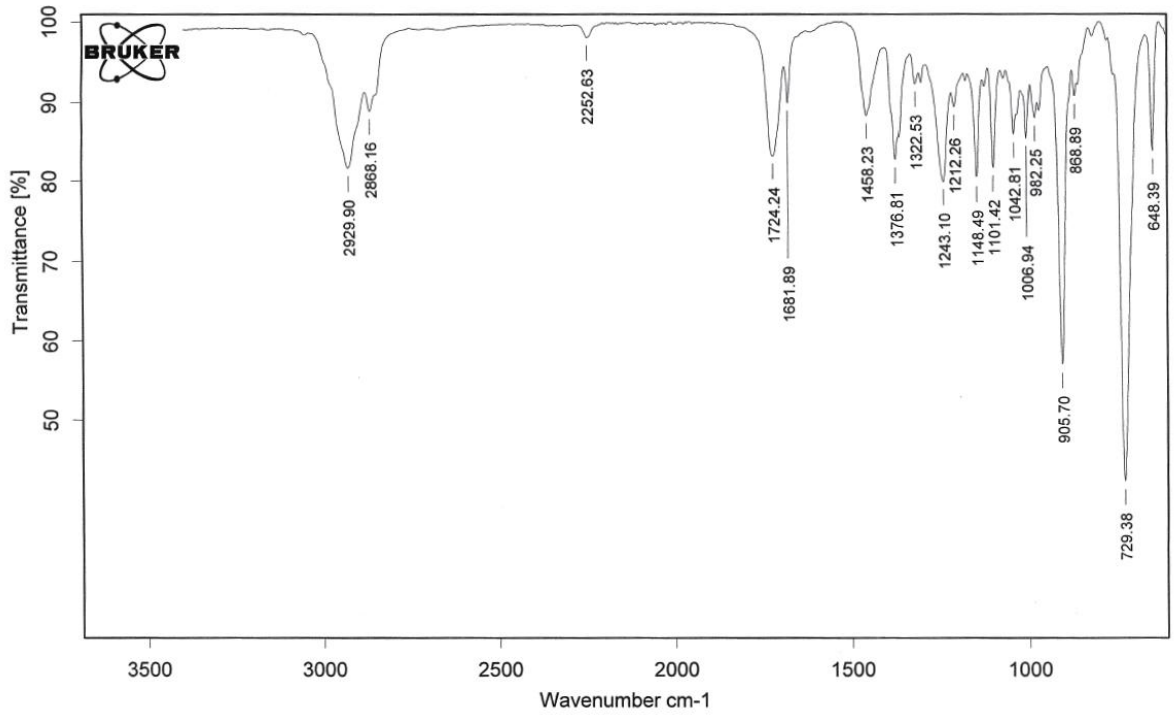


6.7.2 Molecule HREIMS spectrum (-)—8-*epi*-caparrapioxide

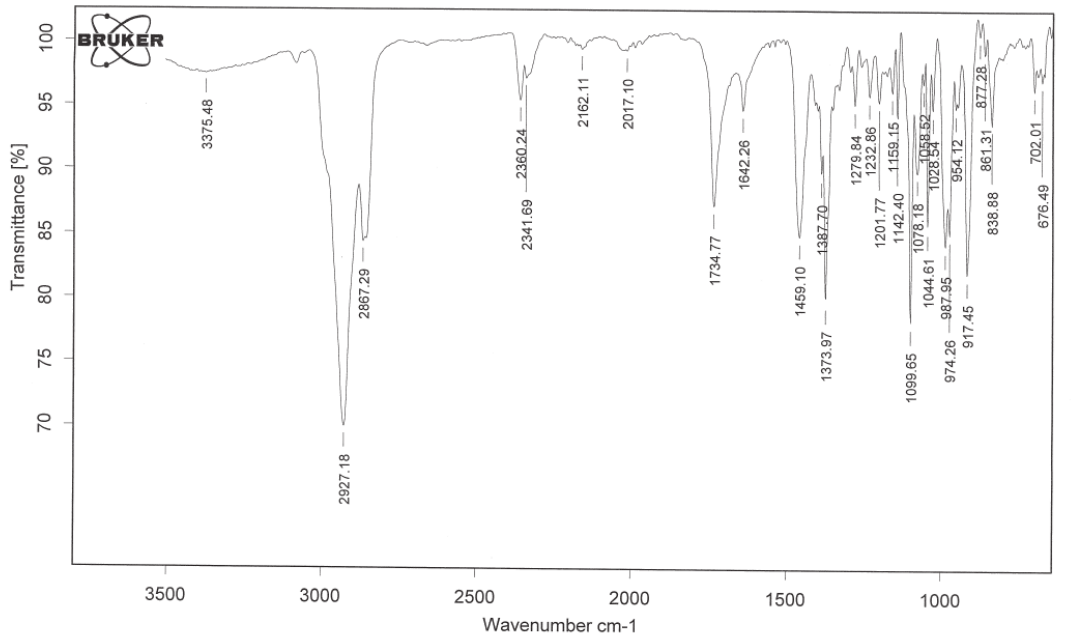


7 IR spectra

7.1 IR spectrum hexahydrochromene



7.2 IR spectrum (-)-caparrapioxide

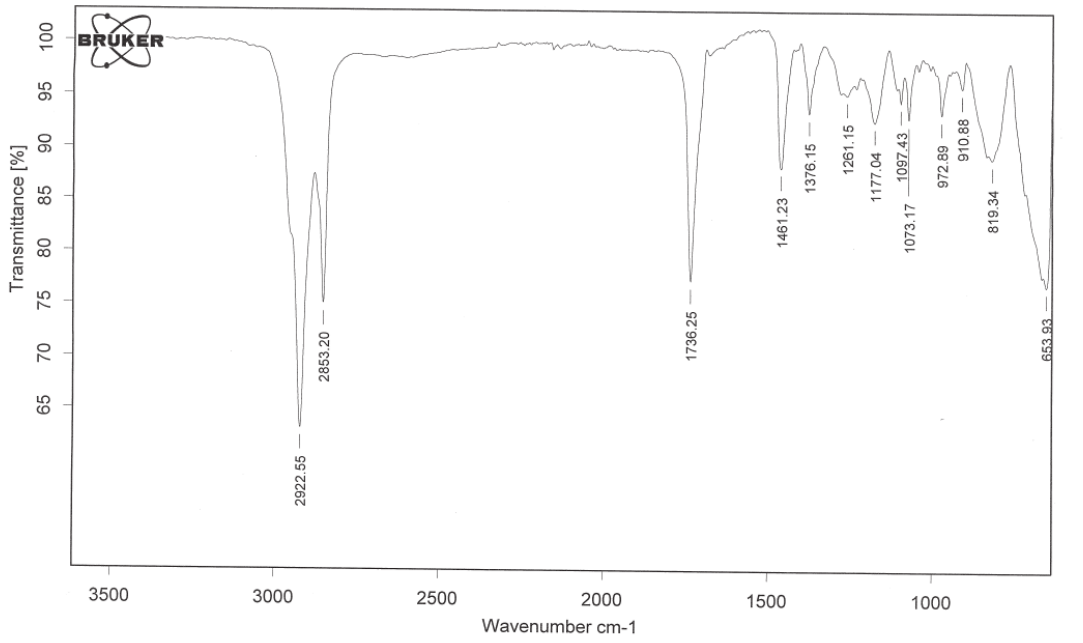


(-)-caparrapi oxide

D:\IR-DATEN\Wegner\mse Säule 9 HPLC1.0 Seitz/mse Säule 9 HPLC1 28/06/2012

Seite 1 von 1

7.3 IR spectrum (-)-8-epi-caparrapioxide



Seite 1 von 1