(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 April 2010 (29.04.2010)

PCT

(10) International Publication Number
WO 2010/046713 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/GB2009/051435

(22) International Filing Date:
23 October 2009 (23.10.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0819563.8 24 October 2008 (24.10.2008) GB

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ISIS INNOVATION** [GB/GB]; Ewert House, Ewert Place, Summertown, Oxford, Oxfordshire OX2 7SG (GB). **SCHOFIELD, Christopher, Joseph** [GB/GB]; University of Oxford, Department of Chemistry, Chemistry Research Laboratory, Mansfield Road, Oxford Oxfordshire OX1 3TA (GB). **HAMED, Refaat, B.** [EG/GB]; University of Oxford, Department of Chemistry, Chemistry Research Laboratory, Mansfield Road, Oxford Oxfordshire OX1 3TA (GB). **BATCHELAR, Edward, Timothy** [GB/GB]; University of Oxford, Department of Chemistry, Chemistry Research Laboratory, Mansfield Road, Oxford Oxfordshire OX1 3TA (GB). **DUCHO, Christian** [DE/DE]; Valentinsbreite 19, 37077 Goettinger (DE).

(74) Agent: **BARKER BRETTELL LLP**; 138 Hagley Road, Edgbaston, Birmingham B16 9PW (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

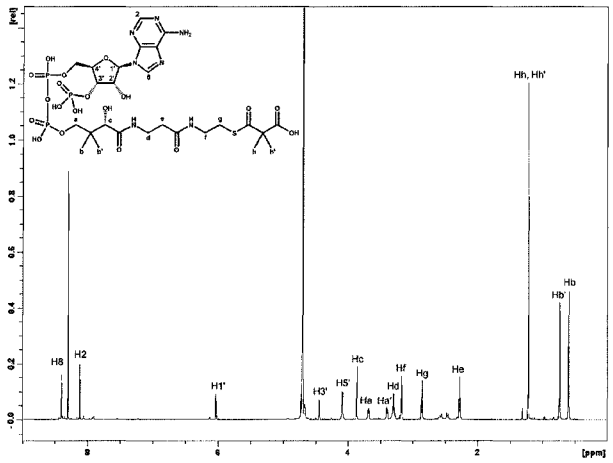
**Published**:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: METHODS FOR PREPARING HETEROCYCLIC RINGS

Figure 1: $^1$H-NMR spectrum of purified dimethylmalonyl-CoA.




(57) **Abstract**: A process for preparing an enantiomerically enriched compound containing a substituted heterocyclic ring, said process comprising a carbon-carbon bond formation reaction in the presence of a crotonase superfamily protein or a homolog or variant thereof.

WO 2010/046713 A2

# METHODS FOR PREPARING HETEROCYCLIC RINGS

**[001]** The invention relates to a process for preparing compounds containing a substituted heterocyclic ring, to compounds obtainable by said process and to the use of said compounds as intermediates in the synthesis or semi-synthesis of compounds of medicinal interest, including antibiotics.

**[002]** Asymmetric C-C bond formation remains a difficult task for synthetic organic chemists. *N*-Containing heterocycles are components of many clinically important pharmaceuticals and agrochemicals. Pyrrolidine, piperidine and azepine alkaloids have been isolated from numerous natural sources. The literature reveals thousands of references to these ring systems in clinical and preclinical research. The development of new methods, particularly those suitable for the production scale preparation of chiral derivatives, for preparation of these *N*-heterocycles is therefore of considerable interest. In the case of saturated heterocycles, and in particular ring functionalised compounds, their utilisation in medicinal chemistry is often limited by the availability of commercially viable compounds. Construction of these heterocyclic rings in an enantioenriched fashion has been a subject of considerable synthetic attention.

**[003]** Thus, according to a first aspect of the invention, there is provided a process for preparing an enantiomerically enriched compound containing a substituted heterocyclic ring, said process comprising a carbon-carbon bond formation reaction in the presence of a crotonase superfamily protein or a homolog or variant thereof.

**[004]** References herein to "enantiomerically enriched" refer to the increase in concentration of one particular enantiomer with respect to the other enantiomer. In one embodiment, the enrichment is of a stereoisomer. For example, enrichment will occur once a greater than 50:50 mixture of stereoisomers is obtained.

**[005]** In one embodiment, enrichment is of a *trans*-carboxymethylproline moiety. In a further embodiment, the enrichment is of a diastereoisomer. In one embodiment, the stereoisomer created by C-C bond formation is enriched as the (*S*)-stereoisomer. In a further embodiment, the stereocentre created by C-C bond formation is *trans* with respect to the C-2 carboxylate group.

**[006]** In one embodiment, the racemic mixture is enriched to a level greater than any one of the following percentages: 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% or 95%.

**[007]** References herein to "crotonase superfamily protein" include references to any crotonase superfamily protein containing carbon-carbon bond formation activity. The skilled person will appreciate that the crotonase superfamily (CS) is a family of enzymes with a

repeated ββα motif, which catalyse a variety of diverse reactions in which a common structural feature involved in catalysis (in most cases) is an oxyanion hole present in their active sites (R. B. Hamed *et al*, Cellular and Molecular Life Sciences, 2008, 65, 2507-2527).

**[008]** In one embodiment, the crotonase superfamily protein comprises one of the following: 3,5-Dihydroxyphenylglyoxylate synthase (DpgC), Transcarboxylase 12S (TC 12S), *Anabaena* β-Diketone hydrolase (ABDH), 6-Oxocamphor hydrolase (6-OCH), 4-Chlorobenzoyl-CoA dehalogenase (4-CBD), Methylmalonyl-CoA decarboxylase (MMCD), Glutaconyl-CoA de-carboxylase-α subunit (Gcdα), $ECH_2$ Decarboxylase domain of CurF (CurF), Naphthoate synthase (MenB), Adenine, Uracil binding ECH homologue (AUH), Enoyl-CoA hydratase (ECH), Dienyl-DoC isomerase (DCI), Hydroxylcinnamoyl-CoA hydratase-ligase (HCHL), $\Delta^3,\Delta^2$-Enoyl-CoA isomerase (ECI), Acetyl-CoA carboxylase carboxyltransferase subunit from yeast (ACC CT), Carboxymethylproline synthase (CarB), The proteolytic subunit of caseinolytic protease (ClpP), Photosystem II D1 CTPase (D1-CTPase), Interphotoreceptor retinoid-binding protein (IRBP) or Tricorn protease (Tricorn).

**[009]** In a further embodiment, the crotonase superfamily protein comprises a carboxymethylproline synthase enzyme. In a further embodiment, the carboxymethylproline synthase enzyme comprises CarB or ThnE or a homolog or variant thereof. In a yet further embodiment, the carboxymethylproline synthase enzyme comprises CarB or a homolog or variant thereof. In an alternative embodiment, the carboxymethylproline synthase enzyme comprises ThnE or a homolog or variant thereof.

**[010]** While the essential steps in the biosynthesis of the simplest carbapenem antibiotic, (5*R*)-carbapenem-3-carboxylate, are catalysed by three enzymes (CarB, CarA and CarC) in *Pectobacterium carotovorum*, higher number of enzymes is involved in the biosynthesis of the more complex carbapenem thienamycin in *Streptomyces cattleya*. The early stages in both cases are catalysed by two members of the crotonase superfamily (CS) enzymes, namely CarB and ThnE, converting malonyl-CoA and L-glutamate semialdehyde/5-hydroxyproline/pyroline-5-carboxylate (L-GHP) into *trans*-carboxymethylproline (*t*-CMP). CarB and ThnE are unusual among the CS because in addition to catalysing decarboxylation of malonyl-CoA it has been surprisingly found that CarB and ThnE also catalyse diastereoselective C-C bond formation in addition to thioester hydrolysis.

**[011]** References herein to "protein", "polypeptide" and "peptide" means a compound composed of at least five constituent amino acids connected by peptide bonds. The constituent amino acids may be from the group of the amino acids encoded by the genetic code and they may be natural amino acids which are not encoded by the genetic code, as well

as synthetic amino acids. Natural amino acids which are not encoded by the genetic code are e.g. hydroxyproline, γ-carboxyglutamate, ornithine, phosphoserine, D-alanine and D-glutamine. Synthetic amino acids comprise amino acids manufactured by chemical synthesis, e.g. D-isomers of the amino acids encoded by the genetic code such as D-alanine and D-leucine, Aib (α-aminoisobutyric acid), Abu (α-aminobutyric acid), Tle (*tert*-butylglycine), β-alanine, 3-aminomethyl benzoic acid and anthranilic acid.

**[012]** In one embodiment, the homolog of the crotonase superfamily has an amino acid sequence having at least 80% identity to a crotonase superfamily member (e.g. CarB or ThnE). In one embodiment, the homolog of the crotonase superfamily has an amino acid sequence having at least 85%, such at least 90%, for instance at least 95%, such as for instance at least 99% identity to a crotonase superfamily member (e.g. CarB or ThnE).

**[013]** The term "identity" as known in the art, refers to a relationship between the sequences of two or more peptides, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between peptides, as determined by the number of matches between strings of two or more amino acid residues. "Identity" measures the percentage of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program (i.e., "algorithms"). Identity of related peptides can be readily calculated by known methods. Such methods include, but are not limited to, those described in Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M. Stockton Press, New York, 1991; and Carillo et al., SIAM J. Applied Math. 48, 1073 (1988).

**[014]** Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity are described in publicly available computer programs. Preferred computer program methods to determine identity between two sequences include the GCG program package, including GAP (Devereux et al., Nucl. Acid. Res. 12, 387 (1984); Genetics Computer Group, University of Wisconsin, Madison, Wis.), BLASTP, BLASTN, and FASTA (Altschul et al., J. Mol. Biol. 215, 403-410 (1990)). The BLASTX program is publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul et al.

NCB/NLM/NIH Bethesda, Md. 20894; Altschul et al., supra). The well known Smith Waterman algorithm may also be used to determine identity.

**[015]** For example, using the computer algorithm GAP (Genetics Computer Group, University of Wisconsin, Madison, Wis.), two peptides for which the percentage sequence identity is to be determined are aligned for optimal matching of their respective amino acids (the "matched span", as determined by the algorithm). A gap opening penalty (which is calculated as 3 times the average diagonal; the "average diagonal" is the average of the diagonal of the comparison matrix being used; the "diagonal" is the score or number assigned to each perfect amino acid match by the particular comparison matrix) and a gap extension penalty (which is usually {fraction (1/10)} times the gap opening penalty), as well as a comparison matrix such as PAM 250 or BLOSUM 62 are used in conjunction with the algorithm. A standard comparison matrix (see Dayhoff et al., Atlas of Protein Sequence and Structure, vol. 5, supp.3 (1978) for the PAM 250 comparison matrix; Henikoff et al., Proc. Natl. Acad. Sci USA 89, 10915-10919 (1992) for the BLOSUM 62 comparison matrix) is also used by the algorithm.

**[016]** Preferred parameters for a peptide sequence comparison include the following:

**[017]** Algorithm: Needleman et al., J. Mol. Biol. 48, 443-453 (1970); Comparison matrix: BLOSUM 62 from Henikoff et al., PNAS USA 89, 10915-10919 (1992); Gap Penalty: 12, Gap Length Penalty: 4, Threshold of Similarity: 0.

**[018]** The GAP program is useful with the above parameters. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps) using the GAP algorithm.

**[019]** In one embodiment, the homolog of the crotonase superfamily has an amino acid sequence, which sequence is at least 80% similar to a crotonase superfamily member (e.g. CarB or ThnE). In one embodiment, the homolog of the crotonase superfamily has an amino acid sequence, which sequence is at least 85%, such as at least 90%, for instance at least 95%, such as for instance at least 99% similar to a crotonase superfamily member (e.g. CarB or ThnE).

**[020]** The term "similarity" is a concept related to identity, but in contrast to "identity", refers to a sequence relationship that includes both identical matches and conservative substitution matches. If two polypeptide sequences have, for example, (fraction (10/20)) identical amino acids, and the remainder are all non-conservative substitutions, then the percentage identity and similarity would both be 50%. If, in the same example, there are 5 more positions where there are conservative substitutions, then the percentage identity remains 50%, but the percentage similarity would be 75% ((fraction (15/20))). Therefore, in

cases where there are conservative substitutions, the degree of similarity between two polypeptides will be higher than the percentage identity between those two polypeptides.

**[021]** Conservative modifications of a peptide comprising a given amino acid sequence (and the corresponding modifications to the encoding nucleic acids) will produce peptides having functional and chemical characteristics similar to those of a peptide comprising the given amino acid sequence. In contrast, substantial modifications in the functional and/or chemical characteristics of such peptide as compared to an original peptide may be accomplished by selecting substitutions in the amino acid sequence that differ significantly in their effect on maintaining (a) the structure of the molecular backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain.

**[022]** For example, a "conservative amino acid substitution" may involve a substitution of a native amino acid residue with a nonnative residue such that there is little or no effect on the polarity or charge of the amino acid residue at that position. Furthermore, any native residue in the polypeptide may also be substituted with alanine, as has been previously described for "alanine scanning mutagenesis" (see, for example, MacLennan et al., Acta Physiol. Scand. Suppl. 643, 55-67 (1998); Sasaki et al., Adv. Biophys. 35, 1-24 (1998), which discuss alanine scanning mutagenesis).

**[023]** Desired amino acid substitutions (whether conservative or non-conservative) may be determined by those skilled in the art at the time such substitutions are desired. For example, amino acid substitutions can be used to identify important residues of the peptides according to the invention, or to increase or decrease the affinity of the peptides described herein for the receptor in addition to the already described mutations.

**[024]** Naturally occurring residues may be divided into classes based on common side chain properties:

1) hydrophobic: norleucine, Met, Ala, Val, Leu, Ile;
2) neutral hydrophilic: Cys, Ser, Thr, Asn, Gln;
3) acidic: Asp, Glu;
4) basic: His, Lys, Arg;
5) residues that influence chain orientation: Gly, Pro; and
6) aromatic: Trp, Tyr, Phe.

**[025]** In making such changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics, these are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine

(-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

**[026]** The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is understood in the art. Kyte et al., J. Mol. Biol. 157, 105-131 (1982). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those that are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

**[027]** It is also understood in the art that the substitution of like amino acids may be made effectively on the basis of hydrophilicity, particularly where the biologically functionally equivalent protein or peptide thereby created is intended for use in immunological embodiments, as in the present case. The greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, i.e., with a biological property of the protein.

**[028]** The following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0±1); glutamate (+3.0±1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5±1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those that are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. One may also identify epitopes from primary amino acid sequences on the basis of hydrophilicity. These regions are also referred to as "epitopic core regions".

**[029]** References herein to “variant” refer to a modified member of the crotonase superfamily wherein one or more amino acid residues of the peptide have been substituted by other amino acid residues and/or wherein one or more amino acid residues have been deleted from the peptide and or wherein one or more amino acid residues have been added to the peptide. Such addition or deletion of amino acid residues can take place at the *N*-terminal of the peptide and/or at the C-terminal of the peptide. All amino acids for which the optical isomer is not stated are to be understood to mean the L-isomer. A simple system is used to describe modified members of the crotonase superfamily. For example, CarBM108V

designates a variant of CarB obtained by substituting the naturally occurring amino acid residue Methionine (M) in position 108 with Valine (V).

**[030]** In one embodiment, the crotonase superfamily protein variant is a variant comprising one or more amino acid substitutions, e.g. two amino acid substitutions.

**[031]** In one embodiment, when enrichment is of a *trans*-carboxymethylproline moiety, the crotonase superfamily protein or homolog or variant thereof is wild-type CarB. In an alternative embodiment, the crotonase superfamily protein or homolog or variant thereof is other than wild-type CarB.

**[032]** In one embodiment, the crotonase superfamily protein variant is a variant of CarB having an amino acid substitution between residues 60 and 235, such as between 70 and 120, such as between 79 and 120, e.g. between 105 and 115. In a further embodiment, the crotonase superfamily protein variant is a variant of CarB having an amino acid substitution between residues 75 and 85 (e.g. 79), between 105 and 115 (e.g. 108 or 111) or between 225 and 230 (e.g. 229).

**[033]** In one embodiment, the crotonase superfamily protein variant is a variant of CarB having an amino acid substitution at position 108. Without being bound by theory, residue 108 in wild-type CarB is believed to form an oxyanion hole together with G62 which is believed to stabilise the enolate intermediate of malonyl-CoA. In a further embodiment, the variant of CarB at position 108 is selected from any one of the following: M108A, M108V, M108L or M108I. In an alternative embodiment, the crotonase superfamily protein or homolog or variant thereof is other than M108A.

**[034]** In an alternative embodiment, the crotonase superfamily protein variant is a variant of CarB having an amino acid substitution at position 111. Without being bound by theory, residue 111 in wild-type CarB is believed to be important in binding the L-GSA substrate of CarB (or analogue thereof). In a further embodiment, the variant of CarB at position 111 is selected from Q111N.

**[035]** In one embodiment, the crotonase superfamily protein variant is a variant of CarB having an amino acid substitution at position 79. Without being bound by theory, residue 79 in wild-type CarB is believed to be one of the residues of the hydrophobic face of the active site. In a further embodiment, the variant of CarB at position 79 is selected from W79F or W79A. In a further embodiment, the variant of CarB is W79F.

**[036]** In one embodiment, the crotonase superfamily protein variant is a variant of CarB having an amino acid substitution at position 229. In a further embodiment, the variant of CarB at position 229 is selected from H229A.

**[037]** In one embodiment, the crotonase superfamily protein variant is a variant of CarB comprising one or more of the amino acid substitutions hereinbefore defined. In a further embodiment, the crotonase superfamily protein CarB variant comprises W79F and M108A.

**[038]** In one embodiment, the crotonase superfamily protein or homolog or variant thereof is wild-type ThnE. In an alternative embodiment, the crotonase superfamily protein or homolog or variant thereof is other than wild-type ThnE.

**[039]** In one embodiment, the crotonase superfamily protein variant is a variant of ThnE having an amino acid substitution between residues 100 and 290, such as between 110 and 285, such as between 115 and 280, e.g. between 120 and 275. In a further embodiment, the crotonase superfamily protein variant is a variant of ThnE having an amino acid substitution between residues 120 and 130 (e.g. 124), between 150 and 160 (e.g. 153) or between 270 and 280 (e.g. 274).

**[040]** In one embodiment, the crotonase superfamily protein variant is a variant of ThnE having an amino acid substitution at position 124. In a further embodiment, the variant of ThnE at position 124 is selected from: W124F.

**[041]** In an alternative embodiment, the crotonase superfamily protein variant is a variant of ThnE having an amino acid substitution at position 153. In a further embodiment, the variant of ThnE at position 153 is selected from: V153A, V153M, V153I, or V153L.

**[042]** In one embodiment, the crotonase superfamily protein variant is a variant of ThnE having an amino acid substitution at position 274. In a further embodiment, the variant of ThnE at position 274 is selected from: H274A.

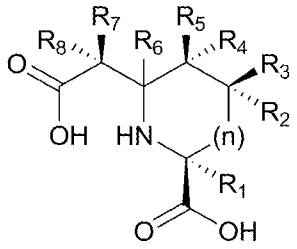
**[043]** In one embodiment, the crotonase superfamily protein variant is a variant of ThnE comprising one or more of the amino acid substitutions hereinbefore defined. In a further embodiment, the crotonase superfamily protein ThnE variant comprises V153M and W124F.

**[044]** In one embodiment, the crotonase superfamily variant is selected from: a variant of CarB wherein the variant is selected from one substituted at M108 (e.g. substitution by A, V, I or L), substituted at Q111 (e.g. substitution by N), substituted at W79 (e.g. substitution by F or A), substituted at H229 (e.g. substitution by A) and substituted at both W79 and M108 (e.g. W79FM108A); and a variant of ThnE wherein the variant is selected from one substituted at V153 (e.g. substitution by A, M, I or L), substituted at W124 (e.g. substitution by F), substituted at H274 (e.g. substitution by A), and substituted at both V153 and W124 (e.g. V153MW124F). Accordingly, the crotonase superfamily variant may be CarB M108A, M108V, M108I, M108L, Q111N, W79F, W79A, H229A, or W79FM108A; or ThnE V153A, V153M, V153I, V153L, W124F, H274A, or V153MW124F.

**[045]** References herein to a "substituted heterocyclic ring" include references to aromatic, non-aromatic, unsaturated, partially saturated and fully saturated heterocyclic rings. In general, such groups may be monocyclic or bicyclic and may contain, for example, 3 to 12 ring members, more usually 5 to 10 ring members. Examples of monocyclic groups are groups containing 3, 4, 5, 6, 7, and 8 ring members, more usually 3 to 7, and preferably 5 or 6 ring members.

**[046]** In one embodiment, the heterocyclic ring is a monocyclic ring. In a further embodiment, the heterocyclic ring contains at least one nitrogen atom. In a yet further embodiment, the heterocyclic ring contains a single nitrogen atom. In a yet further embodiment, the heterocyclic ring contains a carbon containing substituent linked to a carbon atom on the heterocyclic ring via a carbon atom on the substituent. In a further embodiment, the carbon containing substituent is present on the C-4 position (such as 4-methyl or 4- -methyl). In a further embodiment, the heterocyclic ring comprises a 5, 6 or 7 membered ring. In a yet further embodiment, the heterocyclic ring comprises a 6 or 7 membered ring. In an alternative embodiment, the heterocyclic ring comprises a 5 or 6 membered ring. In a yet further embodiment, the heterocyclic ring comprises a 5 membered ring.

**[047]** In a yet further embodiment, the heterocyclic ring comprises a compound of formula (I):




(I)

wherein n represents an integer from 0 to 3; and

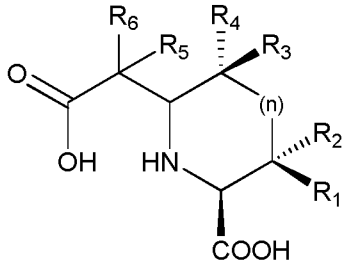
$R_1$, $R_2$, $R_3$, $R_4$, $R_5$, $R_6$, $R_7$ and $R_8$ independently represent hydrogen or optionally substituted $C_{1-6}$ alkyl.

**[048]** References herein to optional substituents for $C_{1-6}$ alkyl include one or more halogen, amino, hydroxyl or cyano groups or the like.

**[049]** In one embodiment, n represents an integer from 1 to 3. In a further embodiment, n represents 1 or 2. In a yet further embodiment, n represents 1.

**[050]** In one embodiment, $R_1$, $R_2$, $R_3$, $R_4$, $R_5$, $R_6$, $R_7$ and $R_8$ independently represent hydrogen or optionally substituted $C_{1-3}$ alkyl. In a further embodiment, $R_1$, $R_2$, $R_3$, $R_4$, $R_5$, $R_6$, $R_7$ and $R_8$ independently represent hydrogen, methyl or ethyl.

**[051]** In a yet further embodiment, the heterocyclic ring comprises a compound of formula (A):




(A)

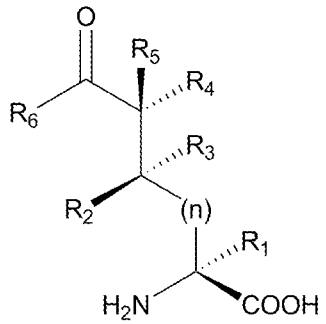
wherein n represents an integer from 0 to 3; and

$R_1$, $R_2$, $R_3$, $R_4$, $R_5$ and $R_6$ independently represent hydrogen or optionally substituted $C_{1-6}$ alkyl.

**[052]** References herein to optional substituents for $C_{1-6}$ alkyl include one or more halogen, amino, hydroxyl or cyano groups or the like. In one embodiment, n represents an integer from 1 to 3. In a further embodiment, n represents 1 or 2. In a yet further embodiment, n represents 1. In one embodiment, $R_1$, $R_2$, $R_3$, $R_4$, $R_5$ and $R_6$ independently represent hydrogen or optionally substituted $C_{1-3}$ alkyl. In a further embodiment, $R_1$, $R_2$, $R_3$, $R_4$, $R_5$ and $R_6$ independently represent hydrogen, methyl or ethyl. In one embodiment, the heterocyclic ring comprises a compound selected from Examples 1-31. In a further embodiment, the heterocyclic ring comprises a compound selected from Examples 1-24 or 1-22.

**[053]** In one embodiment, the process comprises reaction of an amino acid aldehyde compound, for example a semialdehyde compound, in the presence of a malonyl-CoA compound or derivatives thereof.

**[054]** In one embodiment, the aminoacid aldehyde compound is a semialdehyde compound of formula (II):

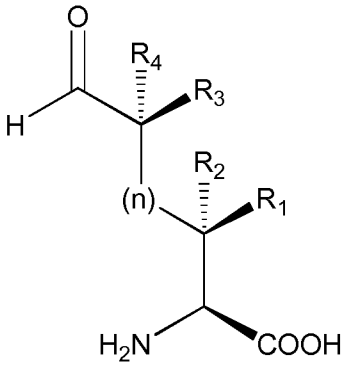



(II)

wherein n represents an integer from 0 to 3; and

$R_1$, $R_2$, $R_3$, $R_4$, , $R_5$ and $R_6$ independently represent hydrogen or optionally substituted $C_{1-6}$ alkyl. Preferred values for n and for $R_1$, $R_2$, $R_3$, $R_4$, , $R_5$ and $R_6$ may be as discussed above.

**[055]** In one embodiment, the aminoacid semialdehyde compound is a compound of formula (B):

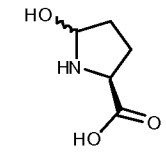


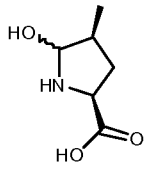

(B)

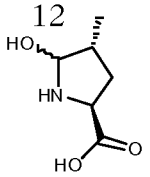
wherein n represents an integer from 0 to 3; and

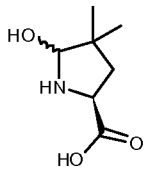
$R_1$, $R_2$, $R_3$ and $R_4$ independently represent hydrogen or optionally substituted $C_{1-6}$ alkyl. Preferred values for n and for $R_1$, $R_2$, $R_3$ and $R_4$ may be as discussed above.

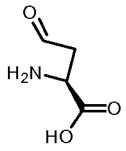
**[056]** In a further embodiment, the amino acid semialdehyde compound or derivative thereof is selected from any one of:

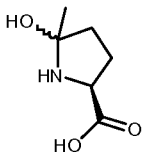

L-glutamate semialdehyde (GSA)

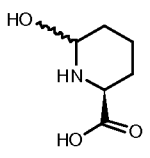

(4*S*)-methylGSA

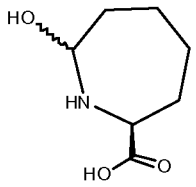

(4*R*)-methylGSA

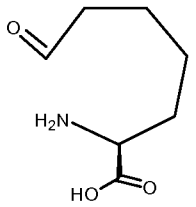

4,4-dimethylGSA


L-aspartate semialdehyde

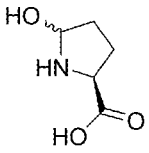

5-methylGSA

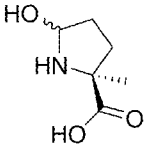

L-aminoadipate semialdehyde

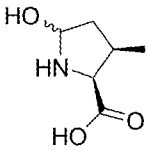

(*S*)-2-amino-7-oxoheptanoic acid

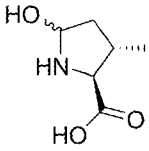

(*S*)-2-amino-7-oxoheptanoic acid

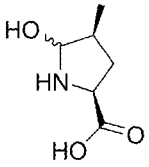
**Aminoacid semialdehydes used in this study.**

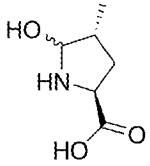

L-GSA

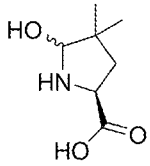

2-methyl-L-GSA

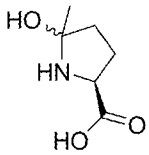

(3*R*)-methyl-L-GSA

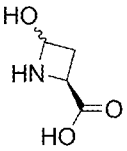

(3*S*)-methyl-L-GSA

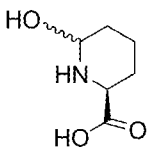

(4*S*)-methyl-L-GSA

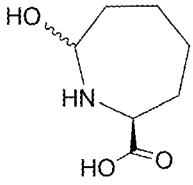

(4*R*)-methyl-L-GSA

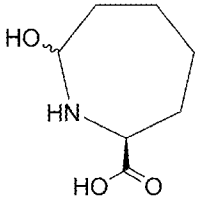

4,4-dimethyl-L-GSA


5-methyl-L-GSA


L-aspartate semialdehyde


L-aminoadipate semialdehyde

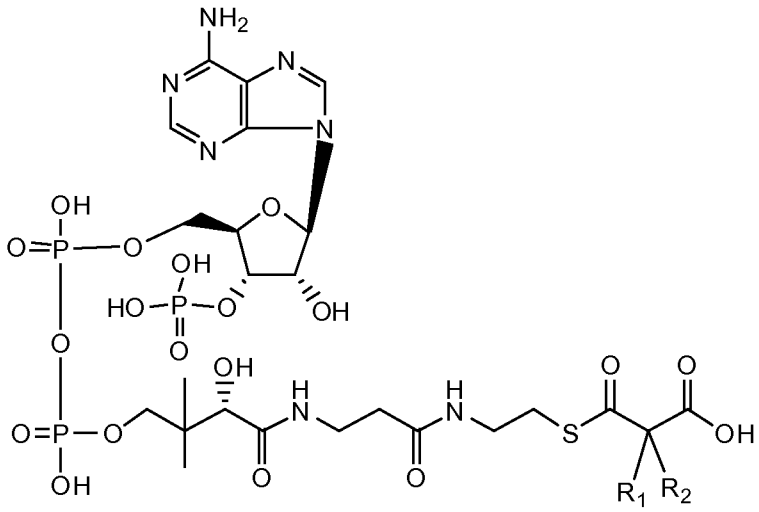

(*S*)-2-amino-7-oxoheptanoic acid


(*S*)-7-hydroxyazepane-2-carboxylic acid

**Aminoacid semialdehydes used in this study**

It will be appreciated that the above mentioned amino acid aldehydes will generally exist in equilibrium with tautomeric derivatives of said compounds. For example, L-GSA exists in equilibrium with L-5-hydroxyproline (5HP) and L-pyrroline-5-carboxylate (P5C).

**[057]** In one embodiment, the malonyl-CoA compound or derivative thereof is selected from any one of: malonyl coenzyme A, methylmalonyl coenzyme A, ethylmalonyl coenzyme A, isopropylmalonyl coenzyme A and dimethylmalonyl coenzyme A:



**$R_1$ = H $R_2$ = H Malonyl Coenzyme A**

**$R_1$ = Me $R_2$ = H Methylmalonyl Coenzyme A**

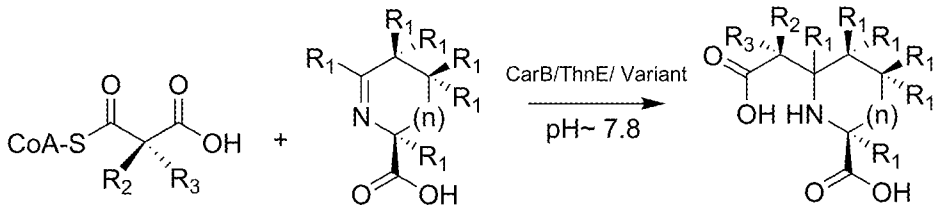
**$R_1$ = Et $R_2$ = H Ethylmalonyl Coenzyme A**

**$R_1$ = Me $R_2$ = Me Dimethylmalonyl Coenzyme A**

**$R_1$ = i-Pr $R_2$ = H Isopropylmalonyl Coenzyme A**

**[058]** Optimisation of a fermentation procedure for the *in vivo* production of the above mentioned *N*-heterocycles utilising a crotonase superfamily protein or homolog or variant thereof may represent a breakthrough for large scale enantioenriched production of these compounds.

**[059]** In one embodiment, the process may be as set out in the following reaction scheme, where the crotonase superfamily protein or homolog or variant thereof is CarB or ThnE or a homolog or variant thereof:




$R_1$= H or $CH_3$
$R_2$= H, $CH_3$, $CH_2CH_3$ or *i*-pr
$R_3$= H or $CH_3$
n = 0 to 3

**[060]** In one embodiment, n is 1 or 2. Examples of CarB or ThnE variants that can be used are discussed above.

**[061]** According to a second aspect of the invention, there is provided a process for preparing a substituted heterocyclic ring, said process comprising a carbon-carbon bond formation reaction in the presence of a crotonase superfamily protein or a homolog or variant thereof.

**[062]** According to a further aspect of the invention, there is provided a process for enhancing the substrate specificity and/or substrate acceptance of a crotonase superfamily protein which comprises the step of preparing a variant of said protein as hereinbefore defined.

**[063]** According to a further aspect of the invention, there is provided a compound containing a substituted heterocyclic ring obtainable by a process as hereinbefore defined.

**[064]** It will be appreciated that the crotonase superfamily protein variants disclosed herein and the compounds containing heterocyclic rings also constitute novel aspects of the invention. Thus, according to a further aspect of the invention, there is provided: a CarB variant selected from any one or more of the following: W79F, W79A, M108A, M108V, M108L, M108I, Q111N, and/or H229A. In one embodiment, the CarB variant is W79F. In an alternative embodiment, the CarB variant is a W79F/M108A double variant. According to a further aspect of the invention, there is provided a ThnE variant selected from any one of the following: V153A, V153M, V153I, V153L, W124F, or H274A. In an alternative embodiment, the ThnE variant is a V153M/W124F double variant.

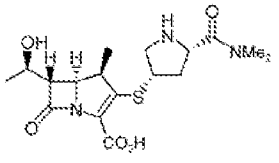
**[065]** According to a further aspect of the invention, there is provided a compound of formula (I) as hereinbefore defined.

**[066]** In one embodiment, the compound of formula (I) is a compound of Examples 1-31. In a further embodiment, the compound of formula (I) is a compound of Examples 1-24 or 1-22.

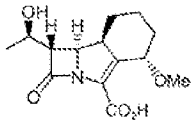
**[067]** It will also be appreciated that the wild-type crotonase superfamily proteins disclosed herein have not previously been disclosed in the synthesis of compounds containing heterocyclic rings with 6,6'-dialkyl substituents. Therefore, according to a further aspect of the invention there is provided a wild-type crotonase superfamily protein as defined hereinbefore for use in the synthesis of compounds containing heterocyclic rings with 6,6'-dialkyl substituents. In one embodiment, the wild-type crotonase superfamily protein is CarB. In an alternative embodiment, the wild-type crotonase superfamily protein is ThnE.

**[068]** According to a further aspect of the invention, there is provided a compound containing a substituted heterocyclic ring as hereinbefore defined for use in the synthesis of a medicament. In one embodiment, said medicament is an antibiotic. In a further embodiment, said antibiotic is a substituted proline. In an alternative embodiment, said antibiotic is a cephem (e.g. carbacephem). In an alternative embodiment, said antibiotic is a carbapenem (e.g. (5*R*)-carbapenem or thienamycin). In a further embodiment, said antibiotic is thienamycin.

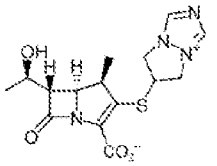
**[069]** Further examples of carbapenem antibiotics that can be synthesised are:



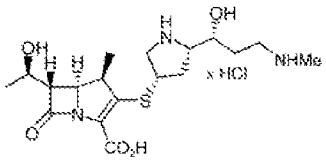

meropenem (55)



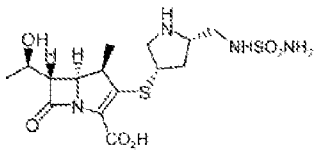

sanfetrinem (56)



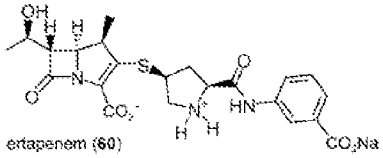

biapenem (57)




lenapenem (58)

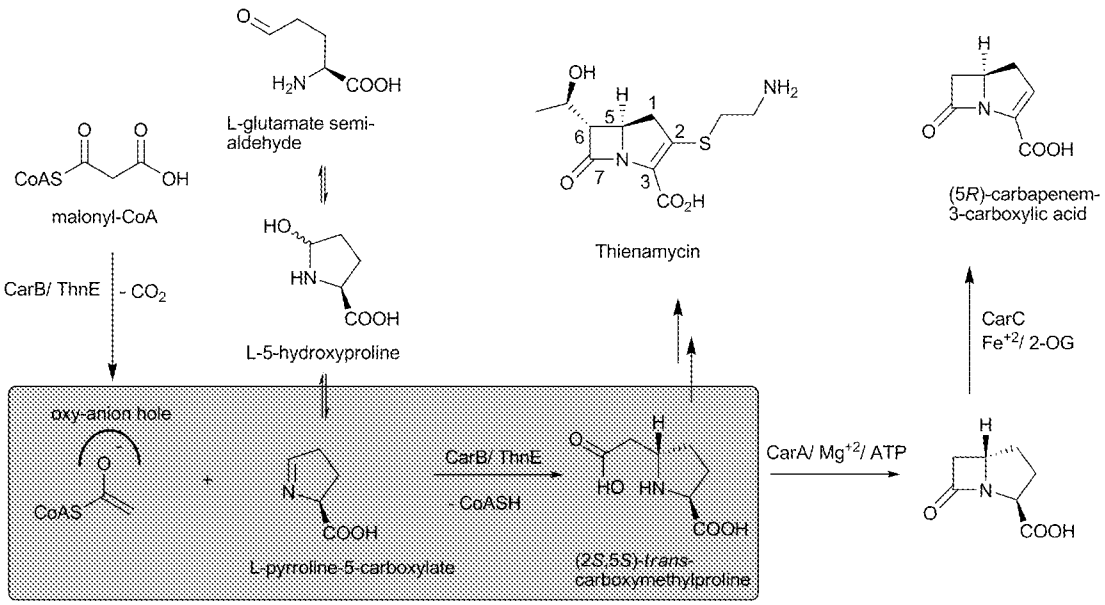



doripenem (59)



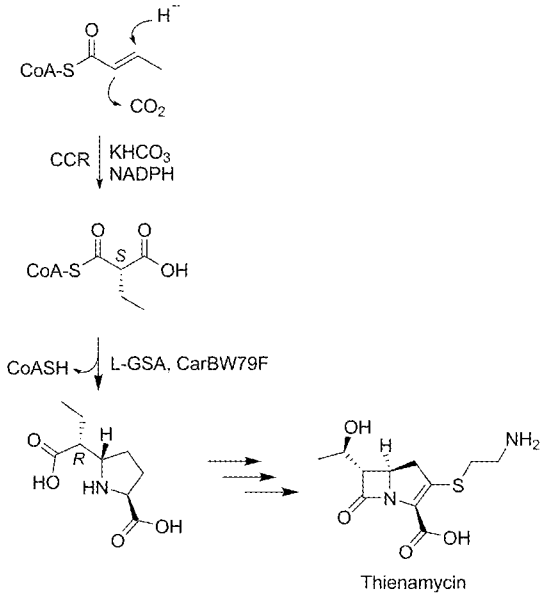

ertapenem (60)

**[070]** The following scheme shows the role of the crotonase superfamily enzymes CarB and ThnE in the biosynthesis of (i) (5*R*)-carbapenem-3-carboxylic acid in *Pectobacterium carotovorum* and (ii) thienamycin in *Streptomyces cattleya*.




**[071]** According to a further aspect of the invention, there is provided a crotonase superfamily protein variant for use in the synthesis of thienamycin. In one embodiment, said variant is a CarB or a ThnE variant. In a further embodiment, said variant is a CarB variant. In a yet further embodiment, said CarB variant is CarBW79F or CarBW79A. It has surprisingly been found that the use of CarBW79F or CarBW79A could advantageously reduce the number of steps required to synthesise thienamycin and therefore constitutes an efficient and cost effective alternative to thienamycin synthesis.

**[072]** For example, thienemycin may be made according to the following reaction scheme in which CarB W79F is used together with Crotonyl-CoA carboxylase reductase (CCR) (See T. J. Erb *et al*, Proceedings of the National Academy of Sciences, 2007, 104, 10631) to provide 6(*R*)-*t*-CMP. CarBW79A could be used as an alternative to CarBW79F.

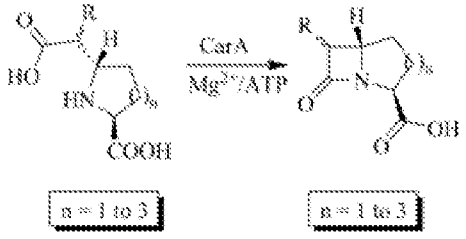



CCR: Crotonyl-CoA carboxylase reductase

**[073]** In another embodiment of the invention there is provided a method for preparing a bicyclic β-lactam by reacting a compound containing a substituted heterocyclic ring as described above with carbapenam synthetase (CarA). The reaction preferably takes place in the presence of suitable cofactors, for example ATP and $Mg^{2+}$.

**[074]** The compound containing a substituted heterocyclic ring may be of formula (I). The compound containing a substituted heterocyclic ring may have been prepared by a process as described above. In one embodiment, there is, therefore, provided a method for preparing a bicyclic β-lactam by (i) preparing a compound containing a substituted heterocyclic ring according to a process as described above and (ii) reacting this compound with carbapenam synthetase (CarA). The CarA mediated reaction preferably takes place in the presence of suitable cofactors, for example ATP and $Mg^{2+}$.

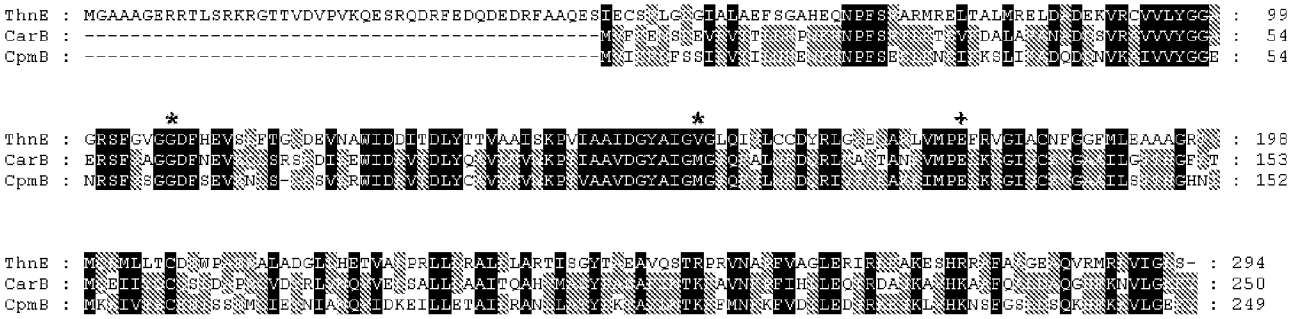
**[075]** In one embodiment, the method is in accordance with the following scheme.



Where R is hydrogen or optionally substituted $C_{1-6}$ alkyl, e.g. optionally substituted methyl, ethyl or propyl.

**[076]** References herein to optional substituents for $C_{1-6}$ alkyl include one or more halogen, amino, hydroxyl or cyano groups or the like.

**[077]** Sequence alignment for ThnE (*S. cattleya*)(Seq. I.D. No. 1), CarB (*P. carotovorum*)(Seq. I.D. No.2) and CpmB (*Photorhabdus luminescens*)(Seq. I.D.No3). Note the high degree of similarity between the three enzymes apart from the first 46 amino acid *N*-



terminal residues in ThnE. The proposed oxyanion hole forming residues are marked with * and the catalytically important (at least in CarB) glutamate residue with +. The figure was generated using Clustal W [J. D. Thompson et al, Nucleic Acids Res. 1994, 22, 4673-4680] and Genedoc[ K. B. Nicholas et al, EMBNEW. NEWS 1997, 4, 14].

The invention will now be illustrated by reference to the following non-limiting examples:

## Materials and Methods

### CarB purification and assay

CarB and CarB variants were purified following the reported method (M. C. Sleeman *et al*, J. of Biol. Chem. 2004, 279, 6730-6736) except that glycerol was omitted from all buffers and that enzyme was buffer exchanged into 50mM 2-amino-2-hydroxymethyl-propane-1,3-diol hydrochloride (TRIS.HCl) pH 7.5 before storage at -80 °C. CarB incubations were performed by sequential addition of the following: 600mM TRIS.HCl pH 9.0 (35 µL), 10mM CoA derivative (8 µL), 50mM GSA/P5C in 10% formic acid (5 µL) and 2mM CarB (2 µL), then incubation at 37 °C for 10 mins. An equal volume of methanol was then added and the mixture cooled on ice for 10 mins before centrifugation at 12,500 x *g* for 10 mins. The supernatant was decanted and analysed by Liquid Chromatography/ Time Of Flight Mass spectrometry (LC/TOFMS). Control assays were performed as before but with substitution of 50mM TRIS.HCl pH 7.5 for CarB.

**Small scale assay analyses**

Products from small scale assays were analysed on a Waters LCT Classic with 2790 sample/ solvent manager with a Primesep 100 column using a gradient from 5 % aqueous MeCN + 0.1 % formic acid to 100 % MeCN + 0.1 % formic acid.

**Large scale enzymatic product isolation and characterisation**

Products for NMR analysis were produced by scale-up of assay conditions (10x), quenching with MeOH (500 µL), centrifugation (13,000 x *g*) and freeze-drying of the supernatant. The resultant residue was re-suspended in 15 % aqueous methanol (200 µL) and purified using a mixed mode Waters Spherisorb column (250 mm x 10 mm, 5 µ) pre-equilibrated in 5 % aqueous MeOH before a gradient was run to 5-25 % aqueous MeOH (according to the polarity of the product) with 0.1 % formic acid. Elution was monitored using a Waters ZMD mass spectrometer, 2700 sample manager and 600 controller. Fractions with masses corresponding to anticipated products were collected (5-10 mL) and lyophilised. The resultant residue was re-suspended in $D_2O$ (500 µL), transferred to an Eppendorf vial and freeze-dried. The final residue was re-suspended in $D_2O$ (6 - 12 µL), transferred into a 1mm NMR tube using a hand centrifuge, and analysed by NMR using a Bruker DQX500 machine fitted with a 1mm TXI microprobe or a Bruker AVIII 700 with $^1H$ inverse cryoprobe. Deuterium chloride was added to some samples and the temperature was elevated to 335 K when required.

**Synthesis and purification of malonyl-CoA derivatives**

Malonyl-CoA derivatives *i.e.* ethyl-, isopropyl- and dimethylmalonyl-CoA were prepared from coenzyme A and ethyl-, isopropyl- and dimethylmalonic acid, respectively, by the method of Taoka *et al.*(Taoka *et al*, J. Biol. Chem., 1994, 269, 31630-31634) Dimethyl- and isopropylmalonyl-CoA were purified by HPLC on a Waters liquid chromatography system, comprising a Waters 2996 photodiode array UV detector monitoring absorbance at 254 and 263 nm, Waters 717plus autosampler, Waters 600E pump and a 250-mm C18 semi-prep column (10 mm internal diameter, Phenomenex, C18 Luna). Solvent reservoir A contained 100 mM ammonium formate pH 5.0 and reservoir B contained methanol. The column was pre-equilibrated in 5 % B at 2 mL/min. After 7 min, a gradient was run to 95 % B over 15 min. These conditions were maintained for 17 min before returning to 5% B over 1 min and the column re-equilibrated for 10 min. Dimethyl- and isopropylmalonyl-CoA eluted at 26 and 30 min, respectively. See figures for $^1H$-NMR assignments.

**Dimethylmalonyl-CoA $^1$H NMR** ($D_2O$, referenced to the HOD peak at 4.7 ppm): δ 0.60 (s, 3H, b), 0.75 (s, 3H, b'), 1.20 (s, 6H, h and h'), 2.30 (t, 2H, e), 2.85 (t, 2H, g), 3.15 (t, 2H, f), 3.30 (m, 2H, d), 3.40 (m, 1H, a), 3.70 (m, 1H, a), 3.85 (s, 1H, c), 4.10 (bs, 2H, 5'), 4.45 (bs, 1H, 3'), H2' and H4' were obscured by the HOD signal, 6.05 (d, 1H, 1'), 8.10 (s, 1H, H2), 8.40 (s, 1H, H8) (results shown in Figure 1).

**Isopropylmalonyl-CoA $^1$H NMR** ($D_2O$, referenced to the HOD peak at 4.7 ppm): δ 0.59 (s, 3H, b), 0.72 (d, 3H, j), 0.74 (s, 3H, b'), 0.77 (d, 3H, k), 2.13 (m, 1H, i), 2.28 (m, 2H, e), 2.88 (m, 2H, g), 3.11 (s,1H, h), 3.18 (m, 2H, F), 3.30 (m, 2H, d), 3.40 (m, 1H, a), 3.69 (m, 1H, a), 3.87 (s, 1H, c), 4.09 (bs, 2H, 5'), 4.51 (bs, 1H, 4'), 6.03 (d, 1H, 1'), 8.12 (s, 1H, H2), 8.41 (s, 1H, H8). H2' and H4' were obscured by the HOD signal (results shown in Figure 2).

**Ethylmalonyl-CoA $^1$H NMR** ($D_2O$, referenced to the HOD peak at 4.7 ppm): δ 0.66 (s, 3H, b), 0.80 (m, 6H, b' and j), 1.72 (t, 2H, i), 2.34 (t, 2H, e), 2.94 (m, 2H, g), 3.25 (m, 2H, F), 3.36 (m, 2H, d), 3.47 (m, 1H, a), 3.75 (m, 1H, a), 3.93 (s, 1H, c), 4.16 (bs, 2H, 5'), 4.51 (bs, 1H, 3'), 6.10 (d, 1H, 1'), 8.18 (s, 1H, H2), 8.46 (s, 1H, H8). H2' and H4' were obscured by the HOD signal. $H_h$ was exchanged with $D_2O$ (results shown in Figure 3).

**Example A: Chemo-enzymatic introduction of a β-methyl substituent at C-4 of carboxymethylproline**

Thienamycin is one of the most potent naturally occurring broad spectrum β-lactam antibiotics. The chemical instability of thienamycin in concentrated solution and in the solid state has held back its development as a clinical drug candidate. Synthetic structural modification of thienamycin produced by capping the nucleophilicity of the C-2 amino group e.g. imipenem (*N*-formimidoyl thienamycin) resulted in a product with improved stability and with antibacterial properties significantly superior to thienamycin but like other naturally occurring carbapenems is readily metabolised by renal dehydropeptidase-I (DHP-I), thus requiring the co-administration of DHP-I inhibitor e.g. cilastatin. The introduction of 1 β-methyl thienemycin derivatives resulted in antibiotics biologically more active than thienemycin and more significantly are highly resistant to enzymatic hydrolysis by DHP-I. The α-methyl isomers are rather resistant to DHP-I hydrolysis but its antibacterial activities are known to be very much decreased. The present study therefore chemo-enzymatically introduced a β-methyl substituent at C-4 of carboxymethylproline. This was achieved by synthesising and testing the two epimers of 4-methylGSA as substrates for

carboxymethylprolines (CarB and ThnE or variants thereof). Firstly, protected forms of C-4 mono- (the two epimers) and di-methylated L-GSA derivatives were synthesised from pyroglutamic acid. Following CarB/ThnE incubation in the presence of malonyl-CoA, 4,4-dimethyl L-GSA was converted to a single product assigned as 4,4-dimethyl-*t*-CMP on the basis of MS and $^{1}H$ NMR analyses (with stereochemical assignment by NOE analyses). However, when either of the diastereomericlly pure (> 95 %) epimers of protected 4-methyl-glutamate semialdehyde were deprotected and incubated with CarB and malonyl-CoA (under standard conditions), they each gave a ca. 1 : 1 mixture of epimeric products **4** and **5**. Evidence that the C-4 epimerisation occurred during the acid mediated deprotection came from deprotection of protected form of glutamate semialdehyde in $DCOOD/D_2O$ (D = $^{2}H$) which led to the incorporation of two deuterium atoms at C-4 of the GSA backbone. Subsequent incubation with CarB and malonyl-CoA led to the production of 4,4-$[^{2}H]_2$-*t*-CMP. When either of the two epimers of 4-methyl-GSA was incubated with ThnE/CarBM108V/CarBM108Ile under the same conditions mentioned before with CarB, the major product was compound 4 with diasteromeric bias more than 90%. Taken together, these results reveal that CarB can accept substrates substituted at either of the epimeric C-4 positions and CarB variants/homologs can stereospecifically introduce a methyl group having β-configuration which is desirable therapeutically. Further optimisation of reaction conditions (pH, temperature, time) can lead to enhanced enrichment of a particular stereoisomer.

**Example B: Spectroscopic identification of products of carboxymethylproline synthases (CarB, ThnE or variants) with alternate substrates (see Table A):**

The following general considerations apply:

1. In all cases, the LC/MS data (negative ion electrospray ionisation, unless otherwise stated) supported the formation of the product as shown by observation of the molecular ion and the ion arising from decarboxylation of the product.

2. The formation of a ring structure (five, six or seven) was assigned in part from the $^{1}H$-NMR chemical shift of the bridgehead proton (H-5, H-6 or H-7 in case of five, six or seven-membered rings, respectively).

3. For all the compounds reported, the stereochemical assignments of the bridgehead carbons as having the (S)-stereochemistry was in part based on the nOe data which showed no correlation between H-2 and the bridgehead proton. The nOe data between other protons supported this assignment. This assignment assumes that the (S)-stereochemistry at C-2 is maintained during the reaction. This has been shown to be the case for CarB catalysed

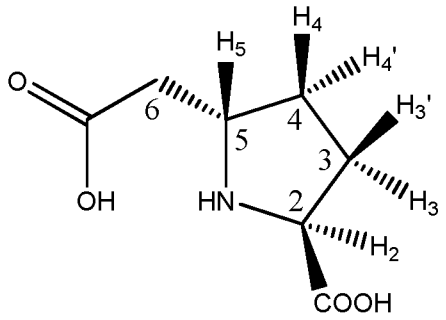
conversion of L-glutamate semialdehyde to (2*S*,5*S*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid (M. C. Sleeman and C. J. Schofield, *Journal of Biological Chemistry*, 2004, 279, 6730-6736).

4. Some of the products differ in their stereochemistry at C-4 and/or C-6 (in case of the 5-membered ring structures) or C-7 (in case of the 6-membered ring structures). Assignment of the stereochemistry at C-4 was from the nOe data between C-4 protons and either the α-proton (H-2) or the bridgehead (C-5) proton. Assignment of the stereochemistry at C-6/C-7 was more complex (and hence is less secure) due to the rotation about the C-5/C-6 (bridgehead) to C-6/C-7 C-C bond. Careful analysis of nOe and coupling constant data, coupled to other NMR experiments (e.g. decoupling of selected protons and TOCSY experiments) were used in these assignments.

**Structure elucidation of products of CarB/ThnE/Variants following incubation**

**Example 1**

**(2*S*,5*S*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid, 1**




(2*S*,5*S*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid

| **Proton no.** | δH compound **1** |
|---|---|
| H-2 | 4.25 (br.t, J = 8.3 Hz) |
| H-5 | 4.0, (m) |
| H-6 | 2.84 (2H, m) |
| H-3 | 2.42 (m) |
| H-4 | 2.25 (m) |
| H-3' | 2.03 (m) |
| H-4' | 1.77 (m) |

**Table 1**

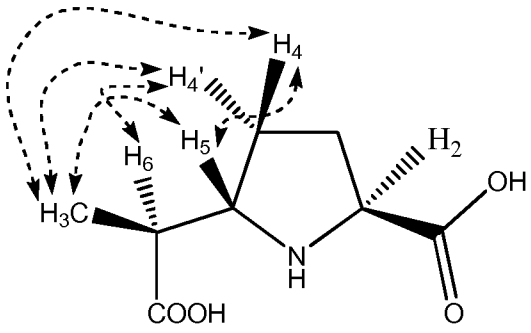
**$^1$H-NMR data of compound 1 isolated from CarB catalysed reaction (500MHz, $D_2O$).**

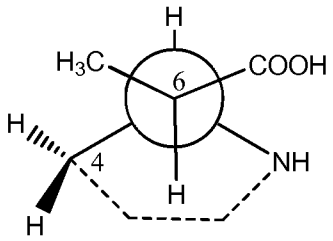
For **1** and all other assigned compounds, the "centre of gravity" of the chemical shift for multiplets is reported.

Authentic (2*S*,5*S*)- and (2*S*,5*R*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid were prepared by synthesis.{M. C. Sleeman *et al*, J. of Biol. Chem. 2004, 279, 6730-6736} The data for (2*S*,5*S*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid prepared by CarB, CarB variants and ThnE were very similar to authentic (2*S*,5*S*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid (*t*-CMP). For **1**, m/z (negative ion electrospray ionisation) 172 [M-H$^-$], 128 [M-COOH$^-$].

**Example 2**

**(2*S*,5*S*)-5-((*S*)-1-carboxyethyl)-2-methylpyrrolidine-2-carboxylic acid, 2**




$J_{5,6}$= 9.8 Hz
(2*S*, 5*S*, 6*S*)

(2*S*,5*S*)-5-((*S*)-1-carboxyethyl)-2-methylpyrrolidine-2-carboxylic acid

NB: For all assigned compounds, the dashed arrows accompanying chemical structures represent the observed nOes.

| **Proton no.** | δH compound **2** | nOe | δH compound **3** | nOe |
|---|---|---|---|---|
| H-2 | 4.23 (br.t, J = 8.50 Hz) | | 4.20 (br.t, J = 8.55 Hz) | |
| H-5 | 3.82 (m) | 4, 6Me, 6 (w) | 3.85 (m) | 4, 6Me, 6 |
| H-6 | 2.80 (dq, J = 9.8, | 4', 6 Me | 2.90 (dq, J = 6.8, | 4' (w), 6 |

| | | | | |
|---|---|---|---|---|
| | 7.16 Hz) | | 7.00 Hz) | Me |
| H-3 | 2.42 (m) | | 2.41 (m) | |
| H-4 | 2.25 (m) | 6Me | 2.22 (m) | 6Me (w), 4', 3' |
| H-3' | 2.01 (m) | | 2.00 (m) | |
| H-4' | 1.78 (m) | 6Me, 6 | 1.78 (m) | 6Me, 6 (w) |
| 6-Me | 1.23 (d, J = 7.16 Hz) | | 1.24 (d, J = 7.00 Hz) | |

**Table 2: $^1$H-NMR and 2D NOESY data for compounds 2 and 3 (500MHz, $D_2O$).**

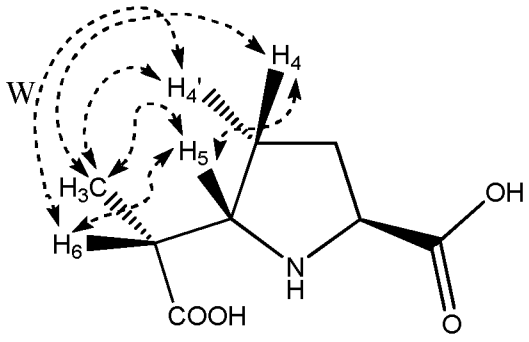
The nOe data refers to the observed nOe between the protons in column 1 and those indicated in columns 3 and 5.

For compounds **2** and **3**: m/z (negative ion electrospray ionisation) 186 [M-H$^-$], 142 [M-COOH$^-$]. The stereochemistry of C-6 of compound **2** was assigned as (*S*)- based on the following observations:

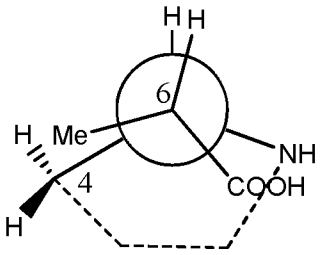
- The $J_{5,6}$ value of 9.8 Hz (predicted $\phi \sim 170^{o}$) together with a weak nOe observed between H-5 and H-6 indicating an *anti* conformation for these two protons.
- The observation of an nOe between H-6 and H-4', together with the observation of an nOe between the C-6 methyl group and both C-4 protons.

**Example 3**

**(2*S*,5*S*)-5-((*R*)-1-carboxyethyl)pyrrolidine-2-carboxylic acid, 3**




(2*S*,5*S*)-5-((*R*)-1-carboxyethyl)pyrrolidine-2-carboxylic acid




$J_{5,6}$= 6.8 Hz,
(2*S*, 5*S*, 6*R*)

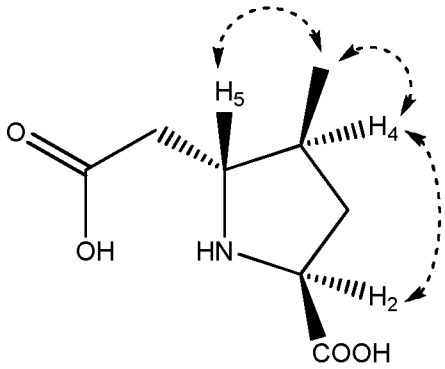
The stereochemistry of C-6 of compound **3** was identified as (*R*)- based on the following observations:

- The $J_{5,6}$ value of 6.8 Hz (predicted $\phi \sim 35^{o}$) together with the strong nOe observed between H-5 and H-6, indicating a *syn* arrangement between these two protons.
- The observed nOe between the C-6 methyl group to both protons at C-4 (H-4' > H-4), together with the observation of a weak nOe between H-6 and H-4'.

**Examples 4 and 5**

**(2*S*,4*S*,5*S*)-5-(carboxymethyl)-4-methylpyrrolidine-2-carboxylic acid, 4**

**(2*S*,4R,5*S*)-5-(carboxymethyl)-4-methylpyrrolidine-2-carboxylic acid, 5**




(2*S*,4*S*,5*S*)-5-(carboxymethyl)-4-methylpyrrolidine-2-carboxylic acid

| **Proton no.** | δH compound **4** | nOe | δH compound **5** |
|---|---|---|---|
| H-2 | 4.21 (br.t, J = 9.20 Hz) | 3, 4(w) | 4.27 (br.t, J = 8.64 Hz) |
| H-5 | 3.51 (m) | 4Me, 6 > 6', 3' | 4.06 (m) |
| H-6 | 2.93 (dd, J = 3.5, 17.90 Hz) | | 2.75 (m) |
| H-6′ | 2.73 (dd, J = 9.4, 17.90 Hz) | | |
| H-3 | 2.56 (m) | 3',4 | 2.56 (m) |
| H-4 | 2.13 (m) | 4Me, 6 < 6' | 2.21 (m) |
| H-3′ | 1.67 (m) | | 2.13 (m) |
| 4-Me | 1.03 (d, J = 6.50 Hz) | | 0.94 (d, J = 7.18 Hz) |

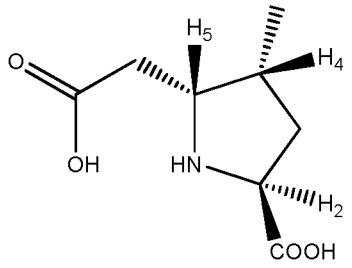
**Table 3: $^{1}$H-NMR and 2D NOESY data for compounds 4 and 5 (500MHz, $D_2O$).**

The nOe data refers to the observed nOe between the protons in column 1 and those indicated in columns 3 and 5.

For compounds **4** and **5**: m/z (negative ion electrospray ionisation) 186 [M-H$^-$], 142 [M-COOH$^-$]. The stereochemistry of C-4 of compound **4** was assigned as (*S*)- based on the following observations:

- The strong nOe observed between H-5 and the C-4 methyl group and the absence of an nOe between H-5 and H-4.
- The weak nOe between H-2 and H-4.

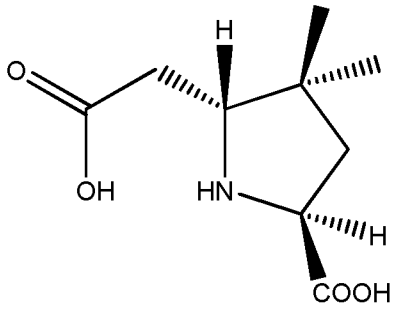
The stereochemistry at C-4 of the other diasteromer (compound **5**, which was obtained as a 1:1 mixture in the case of CarB assays) was therefore assigned (*R*)-.




(2*S*,4R,5*S*)-5-(carboxymethyl)-4-methylpyrrolidine-2-carboxylic acid

**Example 6**

**(2*S*,5*S*)-5-(carboxymethyl)-4,4-dimethylpyrrolidine-2-carboxylic acid, 6**




(2*S*,5*S*)-5-(carboxymethyl)-4,4-dimethylpyrrolidine-2-carboxylic acid

| **Proton no.** | δH Compound **6** |
|---|---|
| H-2 | 4.25 (br.t, J = 9.3 Hz) |
| H-5 | 3.63 (dd, J= 3.5, 10.4 Hz) |
| H-6 | 2.78 (dd, J= 3.5, 17.6 Hz) |
| H-6′ | 2.63 (dd, J= 10.4, 17.6 Hz) |

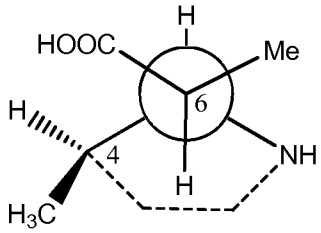
| | |
|---|---|
| H-3 | 2.28 (dd, J = 8.5, 13.2 Hz) |
| H-3′ | 1.86 (dd, J = 10.2, 13.2 Hz) |
| 4Me | 1.07 (s) |
| 4′Me | 0.96 (s) |

**Table 4: $^1$H-NMR data for compound 6 (500MHz, $D_2O$).**

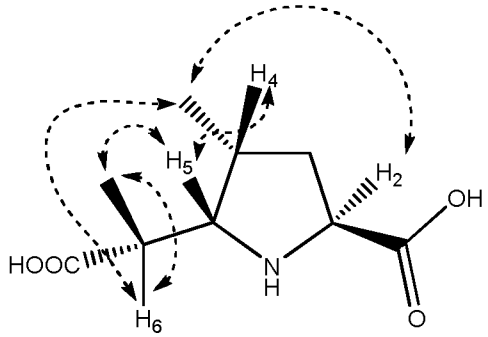
For compounds **6**: m/z (negative ion electrospray ionisation) 200 [M-H$^-$], 156 [M-COOH$^-$]. The $^1$H-NMR data of compound **6** exhibited a similar pattern of chemical shifts as compounds **1-5**.

**Example 7**

**(2*S*,4*R*,5*S*)-5-((*R*)-1-carboxyethyl)-4-methylpyrrolidine-2-carboxylic acid, 7**




$J_{5,6}$= 10.0 Hz
$J_{4,5}$=5.8 Hz
(2*S*, 4*R*, 5*S*, 6*R*)




(2*S*,4*R*,5*S*)-5-((*R*)-1-carboxyethyl)-4-methylpyrrolidine-2-carboxylic acid

| Proton no. | δH Compound **7** | nOe | δH Compound **8** | nOe | δH Compound **9** | nOe |
|---|---|---|---|---|---|---|
| H-2 | 4.24 (br.t, J = 9.20 Hz) | 3, 4Me | 4.18 (br.t, J = 9.0 Hz) | 3, 4 | 4.14 (br.t, J = 9.10 Hz) | |
| H-5 | 3.81 (m) | 4, 3', 6Me, 6 (w) | 3.49 (m) | 4Me, 6Me, 6 | 3.39 (m) | 4Me, 6Me, 4 (w), 6 |
| H-6 | 2.80 (dd, J = 10.0,7.0 Hz) | 4Me , 6 Me | 2.92 (dd, J = 4.8,6.8 Hz) | H-4, 4Me , 6 Me | 2.87 (dd, J = 4.8,6.8 Hz) | 4Me (w), 6 Me |
| H-3 | 2.63 | | 2.53 (m) | | 2.55 (m) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| H-4 | 2.14 (m) | | 2.22 (m) | 4Me, 6Me, 6 | 2.12 (m) | |
| H-3′ | | | 1.68 (m) | | 1.68 (m) | |
| 6-Me | 1.30 (d, J = 7.0 Hz) | | 1.22 (d, 6.8 J = Hz) | 4Me | 1.33 (d, J = 6.80 Hz) | |
| 4-Me | 0.91 (d, J = 7.3 Hz) | | 0.99 (d, J = 7.0 Hz) | | 1.07 (d, J = 6.90 Hz) | |

**Table 5: $^1$H-NMR and 2D NOESY data for compounds 7, 8 and 9 (500MHz, $D_2O$).**

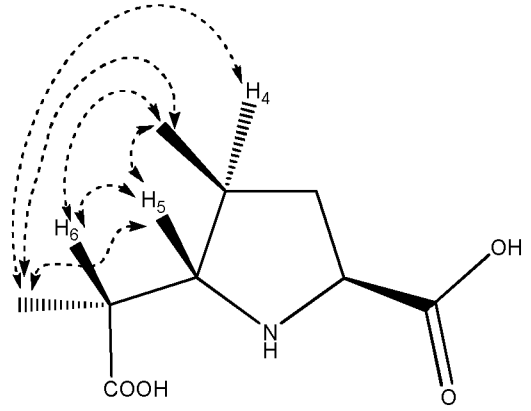
The nOe data refers to the observed nOe between the protons in column 1 and those indicated in columns 3, 5 and 7.

For compounds **7**, **8** and **9**: m/z (negative ion electrospray ionisation) 200 [M-H$^-$], 156 [M-COOH$^-$]. The stereochemistry of C-4 of compound **7** was assigned as *R* based on the observation of a strong nOe between H-2 and the C-4 methyl group, together with the absence of any nOe between C-4 methyl group and H-5. The stereochemistry of C-6 was assigned as *R* on the basis of the following observations:

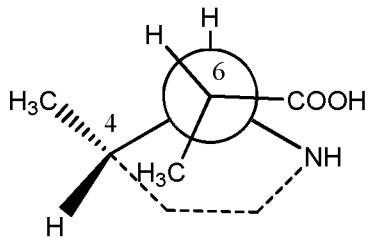
- A $J_{5,6}$ value of 10 Hz (predicted $\phi \sim 170^{o}$) in addition to a weak nOe observed between H-5 and H-6 indicating an *anti* arrangement for these two protons.
- A strong nOe observed between H-6 and C-4 methyl group, as well as the absence of any nOe between the methyl group on C-6 to either H-4 or the methyl group at C-4.

**Example 8**

**(2*S*,4*S*,5*S*)-5-((*R*)-1-carboxyethyl)-4-methylpyrrolidine-2-carboxylic acid, 8**




(2*S*,4*S*,5*S*)-5-((*R*)-1-carboxyethyl)-4-methylpyrrolidine-2-carboxylic acid




$J_{5,6}$= 4.8 Hz
$J_{4,5}$=10.0 Hz
(2*S*, 4*S*, 5*S*, 6*R*)

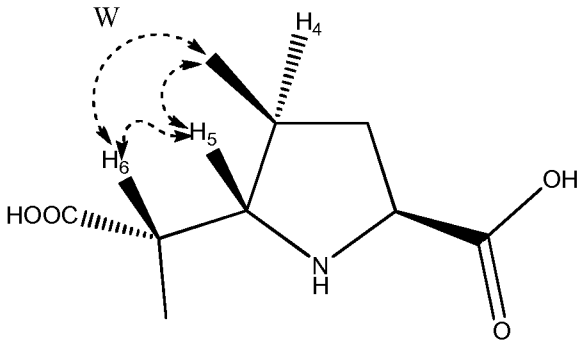
The stereochemistry of C-4 of compound **8** was assigned as (*S*)- based on the observation of a strong nOe between H-5 and the C-4 methyl group together with the observation of a weak

nOe between H-5 and H-4. The stereochemistry of the C-6 was assigned as (*S*)- based on the following observations:

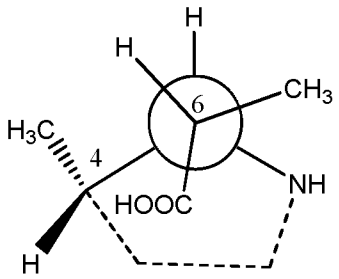
- A $J_{5,6}$ value of 4.8 Hz (predicted $\phi \sim 40^{o}$) together with a strong nOe between H-5 and H-6 indicating a *syn* relationship between these two protons.
- A strong nOe between H-6 to both H-4 and the methyl group on C-4 as well as the observation of an nOe between the methyl group on C-6 to both H-4 and the C-4 methyl group.

**Example 9**

**(2*S*,4*S*,5*S*)-5-((*S*)-1-carboxyethyl)-4-methylpyrrolidine-2-carboxylic acid, 9**




(2*S*,4*S*,5*S*)-5-((*S*)-1-carboxyethyl)-4-methylpyrrolidine-2-carboxylic acid



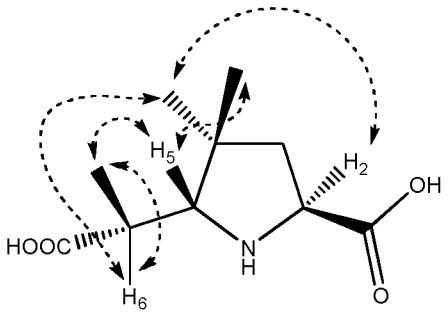

$J_{5,6}$= 4.8Hz
$J_{4,5}$= 9.8Hz
(2*S*, 4*S*, 5*S*, 6*S*)

The stereochemistry of C-4 of compound **9** was assigned as (*S*)- based on the observation of a strong nOe between H-5 and the C-4 methyl group together with the observation of a weak nOe between H-5 and H-4. The stereochemistry of the C-6 was assigned as (*R*)- based on the following observations:

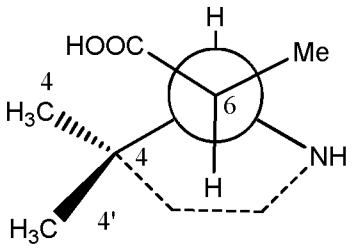
- The $J_{5,6}$ value of 4.8 Hz (predicted $\phi \sim 50^{o}$), together with the strong nOe observed between H-5 and H-6 indicating a *syn* relationship between these two protons.
- The observation of a weak nOe between H-6 and the methyl group on C-4, together with the absence of any observed nOe between H-6 and H-4, and the observation of no nOe between C-6 methyl group to either H-4 or C-4 methyl group.

**Example 10**

**(2*S*,5*S*)-5-((*R*)-1-carboxyethyl)-4,4-dimethylpyrrolidine-2-carboxylic acid, 10**




(2*S*,5*S*)-5-((*R*)-1-carboxyethyl)-4,4-dimethylpyrrolidine-2-carboxylic acid




$J_{5,6}$= 9.8 Hz
(2*S*, 5*S*, 6*R*)

| Proton no. | δH compound 10 | nOe |
|---|---|---|
| H-2 | 4.13 (br.t, J = 9.3 Hz) | 3, 4'Me |
| H-5 | 3.53 (d, J = 9.8 Hz) | 4Me, 3', 6Me, 6 (w) |
| H-6 | 2.75 (dq, J = 6.8, 9.8 Hz) | 4'Me > 4Me , 6 Me |
| H-3 | 2.23 (dd, J = 8.16, 13.2 Hz) | 3', 4'Me |
| H-3' | 1.90 (dd, J = 10.4, 13.2 Hz) | |
| 4Me | 1.04 (s) | |
| 4'Me | 1.02 (s) | |
| 6Me | 1.30 (d, J = 6.8 Hz) | |

**Table 6: $^1$H- NMR and 2D nOe data for compound 10 (500MHz, $D_2O$).**

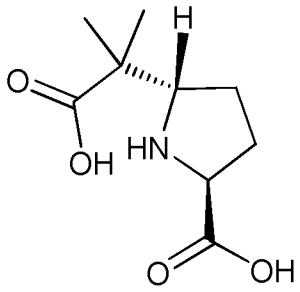
The nOe data refers to the observed nOe between the protons in column 1 and those indicated in column 3.

For compound **10**: m/z (negative ion electrospray ionisation) 214 [M-H$^-$], 170 [M-COOH$^-$]. The stereochemistry of the C-6 was assigned as (*R*)- on the basis of the following observations:

- The $J_{5,6}$ value of 9.8 Hz (predicted $\phi \sim 170^{\circ}$) together with a weak nOe observed between H-5 and H-6, indicating an *anti* relationship between the two protons.
- The strong nOe observed between H-6 and one of the methyl groups on C-4 (4'Me ), together with the absence of an nOe observed between C-6 methyl group to either of the two C-4 methyl groups (supported by 1D NOESY data involving the C-6 methyl group).

**Example 11**

**(2*S*,5*S*)-5-(2-carboxypropan-2-yl)pyrrolidine-2-carboxylic acid, 11**




(2*S*,5*S*)-5-(2-carboxypropan-2-yl)-pyrrolidine-2-carboxylic acid

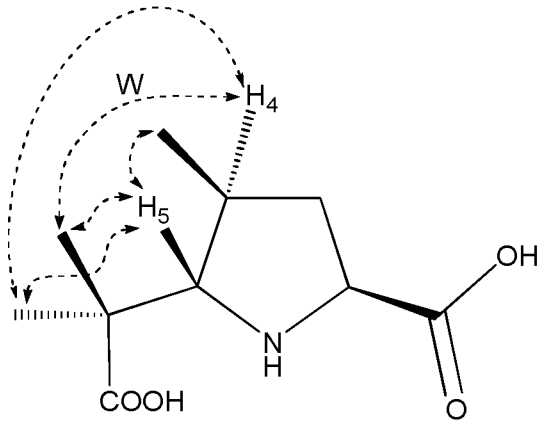
| Proton no. | δH compound **11** |
|---|---|
| H-2 | 4.15 (br. t, J = Hz) |
| H-5 | 3.77 (dd, J = 7.75 Hz) |
| H-3 | 2.39 (m) |
| H-4 | 2.20 (m) |
| H-3' | 1.97 (m) |
| H-4' | 1.80 (m) |
| 6-Me | 1.29 (s) |
| 6'-Me | 1.19 (s) |

**Table 7: $^1$H-NMR data for compound 11 (700MHz, D2O).**

For compound **11**: m/z (negative ion electrospray ionisation) 200 [M-H$^-$], 156 [M-COOH$^-$].

**Example 12**

**(2*S*,4*S*,5*S*)-5-(2-carboxypropan-2-yl)-4-methylpyrrolidine-2-carboxylic acid, 12**




(2*S*,4*S*,5*S*)-5-(2-carboxypropan-2-yl)-4-methylpyrrolidine-2-carboxylic acid

| Proton | δH compound **12** | nOe |
|---|---|---|
| H-2 | 4.00 (br. t, J = Hz) | |
| H-5 | 3.23 (dd, J = 7.3,10.6 Hz) | 4-Me, 6-Me, 6'Me |
| H-3 | 2.37 (m) | |
| H-4 | 2.07 (m) | 6-Me (W), 6'Me |
| H-3' | 1.72 (m) | |
| 4-Me | 0.98 (d, J = 6.75 Hz) | |
| 6-Me | 1.14 (s) | |
| 6'-Me | 1.06 (s) | |

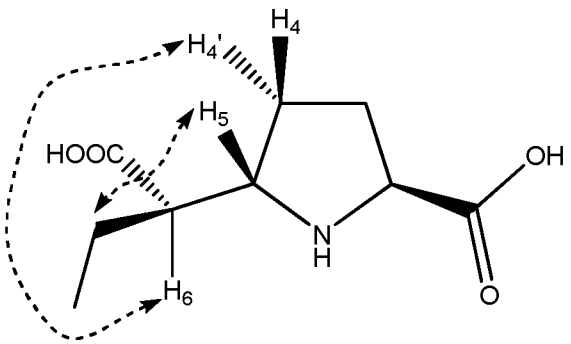
**Table 8: $^1$H-NMR and 2D nOe data for compound 12 (700MHz, $D_2O$).**

The nOe data refers to the observed nOe between the protons in column 1 and those indicated in column 3.

For compound **12**: m/z (negative ion electrospray ionisation) 214 [M-H$^-$], 170 [M-COOH$^-$]. The stereochemistry at C-4 of compound **12** was assigned as (*S*)- on the basis of the observation of a strong nOe between H-5 and the methyl group on C-4, together with the observation of a weak nOe between H-5 and H-4.

**Example 13**

**(2*S*,5*S*)-5-((*R*)-1-carboxypropyl)pyrrolidine-2-carboxylic acid, 13**




(2*S*, 5*S*, 6*R*)

(2S,5S)-5-((R)-1-carboxypropyl)pyrrolidine-2-carboxylic acid

| Proton no. | δH compound **13** | δH compound **14** |
|---|---|---|
| H-2 | 4.11 (br. t, J = 8.2 Hz) | 4.01 (br. t, J = 8.3 Hz) |
| H-5 | 3.75 (m) | 3.69 (m) |
| H-6 | 2.52 (m) | 2.35 (m) |
| H-3 | 2.37 (m) | 2.30 (m) |
| H-4 | 2.15 (m) | 2.12 (m) |
| H-3' | 1.97 (m) | 1.88 (m) |
| H-4' | 1.73 (m) | 1.60 (m) |
| H-7 | 1.63 (m) | 1.58 (m) |
| H-7' | | 1.50 (m) |
| 8-Me | 0.89 (t, J = 7.30 Hz) | 0.80 (t, J = 7.52 Hz) |

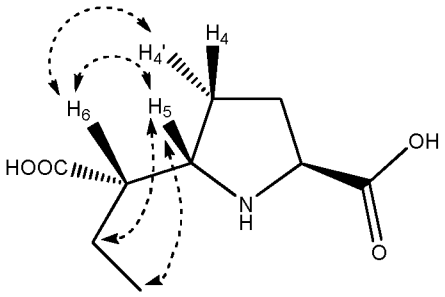
**Table 9: $^1$H-NMR and 2D nOe data for compounds 13 and 14 (500 and 700MHz, $D_2O$).**

For compound **13** and **14**: m/z (negative ion electrospray ionisation) 200 [M-H$^-$], 156 [M-COOH$^-$]. The stereochemistry of C-6 of compound **13** was assigned as (*R*)- based on the following observations:

- The $J_{5,6}$ value of 9.2 Hz together with the weak nOe between H-5 and H-6 indicating an *anti* arrangement of these two protons.
- The observed nOe between H-5 and H-7 and H-7', together with absence of any observed nOe between C-7 protons to any of C-4 protons.

**Example 14**

**(2*S*,5*S*)-5-((*S*)-1-carboxypropyl)pyrrolidine-2-carboxylic acid, 14**




(2*S*, 5*S*, 6*S*)

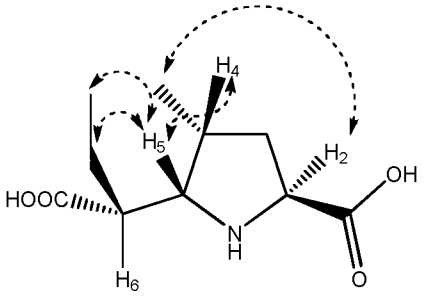
(2S,5S)-5-((S)-1-carboxypropyl)pyrrolidine-2-carboxylic acid

The stereochemistry of C-6 of compound **14** was assigned as (*S*)- based on the following observations:

- The $J_{5,6}$ value of 6.3 Hz together with the strong nOe between H-5 and H-6 indicating a *syn* arrangement of these two protons.
- The strong nOe observed between H-5 and H-7, H-7', coupled to a strong nOe observed between H-5 and C-8 methyl group together with weak nOe between H-6 and H-4'.

**Example 15**

**(2*S*,4*R*,5*S*)-5-((*R*)-1-carboxypropyl)-4,5-dimethylpyrrolidine-2-carboxylic acid, 15**




$J_{5,6}$= 11 Hz,
$J_{5,4}$= 5.3 Hz
(2*S*, 4*R*,5*S*, 6*R*)

(2*S*,4*R*,5*S*)-5-((*R*)-1-carboxypropyl)-4,5-dimethylpyrrolidine-2-carboxylic acid

| Proton no. | δH of compound **15** | δH of compound **16** | δH of compound **17** |
|---|---|---|---|
| H-2 | 4.14 (br. t) | 4.04 (br. t) | 3.98 (br. t) |
| H-5 | 3.71 (m) | 3.37 (dd, 7.3, 9.6 Hz) | 3.23 (dd, 10.7, 2.5 Hz) |
| H-6 | 2.55 (m) | 2.54 (m) | 2.43 (m) |
| H-4 | 2.50 (m) | 2.17 (m) | 1.95 (m) |
| H-3 | 2.07 (m) | 2.46 (m) | 2.43 (m) |
| H-3' | | 1.63 (m) | 1.53 (m) |
| H-7 | 1.68 (m) | 1.58 (m) | 1.59 (m) |
| H-7' | 1.57 (m) | | |
| 4-Me | 0.87 (d, 7.30 Hz) | 0.94 (d, 6.57 Hz) | 0.96 (d, 6.51 Hz) |
| 8-Me | 0.83 (t, 7.56 Hz) | 0.84 (t, 7.42 Hz) | 0.84 (t, 7.51 Hz) |

**Table 10: $^1$H-NMR data of compounds 15, 16 and 17 (700MHz, $D_2O$).**

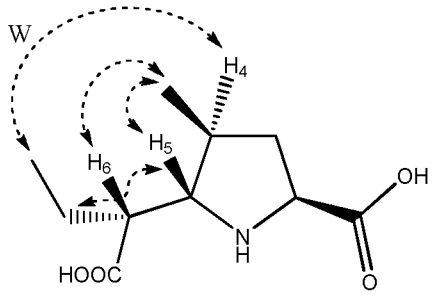
For compound **15**, **16** and **17**: m/z (negative ion electrospray ionisation) 214 [M-H$^-$], 170 [M-COOH$^-$]. The strong nOe between the 4-Me and H-2 coupled to that between H-4 and H-5 confirm the stereochemistry of C-4 as (*R*)-. The stereochemistry of C-6 was assigned as *R* on the basis of the following observations:

1- No nOe between H-5 and H-6 together with the value of $J_{5,6}$ ~ 11 Hz showing *anti* relationship between the two protons.

2- The strong nOe between H-5 and H-7 as well as 8-Me together with the absence of any detectable nOe between H-4 and H-7 nor 4Me to H-7.

**Example 16**

**(2*S*,4*S*,5*S*)-5-((*R*)-1-carboxypropyl)-4,5-dimethylpyrrolidine-2-carboxylic acid, 16**




$J_{56}$= 7.3 Hz,
$J_{54}$= 9.6 Hz
(2*S*, 4*S*,5*S*, 6*R*)

(2*S*,4*S*,5*S*)-5-((*R*)-1-carboxypropyl)-4,5-dimethylpyrrolidine-2-carboxylic acid

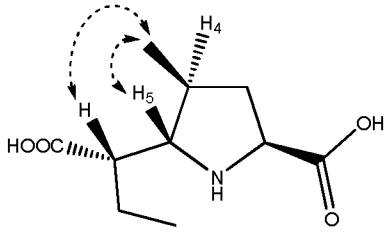
The strong nOe between the 4-Me and H-5 coupled to that between H-4 and H-2 confirm the stereochemistry of C-4 as (S)-. The stereochemistry of C-6 was assigned as (*R*)- on the basis of the following observations:

1- Strong nOe between H-5 and H-6 together with the value of $J_{5,6}$ ~ 7.3 Hz showing *syn* relationship between the two protons.

2- The strong nOe between H-5 and H-7 as well as the weak one between H-5 and 8-Me together with weak nOe between H-4, 4Me and H-7.

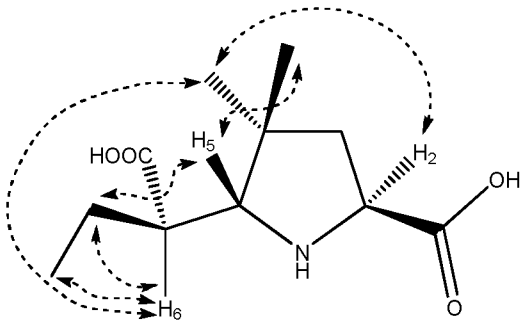
**Example 17**

**(2*S*,4*S*,5*S*)-5-((*S*)-1-carboxypropyl)-4,5-dimethylpyrrolidine-2-carboxylic acid, 17**




$J_{56}$= 2.5 Hz,
$J_{54}$= 10.7 Hz
(2*S*, 4*S*,5*S*, 6*S*)

(2*S*,4*S*,5*S*)-5-((*S*)-1-carboxypropyl)-4,5-dimethylpyrrolidine-2-carboxylic acid



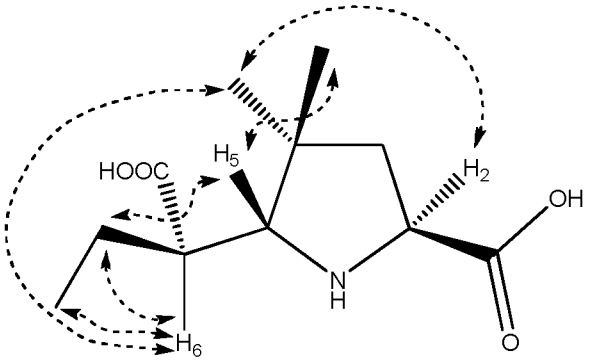

$J_{56}$= 11 Hz,
$J_{54}$= 5.3 Hz
(2*S*,5*S*, 6*R*)

(2*S*,5*S*)-5-((*R*)-1-carboxypropyl)-4,4,5-trimethylpyrrolidine-2-carboxylic acid

The strong nOe between the 4-Me and H-5 reveals the stereochemistry of C-4 as (*S*)-. Assigning the stereochemistry of C-6 was hindered by the fact that H-6 and H-s had the same chemical shift value. However, taking into consideration the unambiguous assignment of the stereochemistry of C-6 of the diasteromers **1** and **2**, one can conclude that of diasteromer **3** as (*S*)-.

**Example 18**

**(2*S*,5*S*)-5-((*R*)-1-carboxypropyl)-4,4,5-trimethylpyrrolidine-2-carboxylic acid, 18**




$J_{56}$= 11 Hz,
$J_{54}$= 5.3 Hz
(2*S*,5*S*, 6*R*)

(2*S*,5*S*)-5-((*R*)-1-carboxypropyl)-4,4,5-trimethylpyrrolidine-2-carboxylic acid

| Proton no. | δH |
|---|---|
| H-2 | 4.06 (br. t, J = Hz) |
| H-5 | 3.44 (d, J = 11.1 HZ) |
| H-6 | 2.48 (m) |
| H-3 | 2.16 (m) |
| H-3' | 1.83 (m) |
| H-7 | 1.58 (m) |
| H-7' | |
| 4,4'-Me | 0.96 (d, J = 4.66 Hz) |
| 8-Me | 0.83 (t, J = 7.32Hz) |

**Table 11: $^1$H-NMR data of compound 18 (700MHz, $D_2O$).**

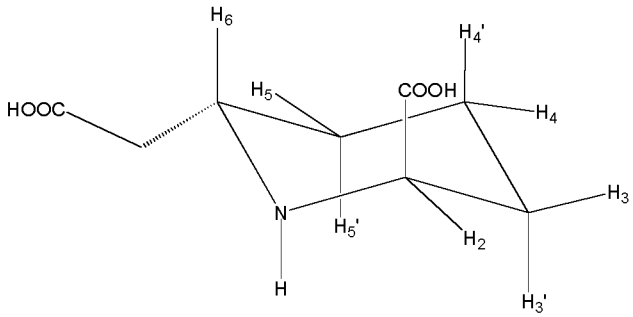
For compound **18**: m/z (negative ion electrospray ionisation) 228 [M-H$^-$], 184 [M-COOH$^-$]. The stereochemistry of C-6 was assigned as (*R*)- on the basis of the following observations:

1- Weak nOe between H-5 and H-6 together with the value of $J_{5,6}$ ~ 11.1 Hz showing *anti* relationship between the two protons.

2- The strong nOe between H-5 and H-7 while it shows nothing to 8-Me together with the absence of any detectable nOe between 4Me and H-7 nor H-7'.

**Example 19**

**(2*S*,6*S*)-6-(carboxymethyl)piperidine-2-carboxylic acid, 19**




(2S,6S)-6-(carboxymethyl)piperidine-2-carboxylic acid

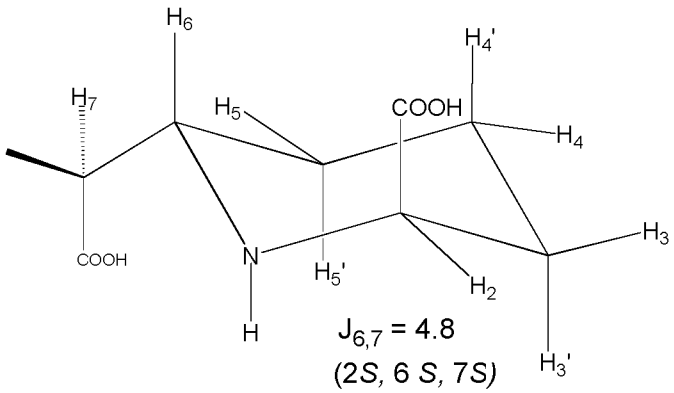
| Proton no. | δH compound **19** | Conformation |
|---|---|---|
| H-2 | 4.01 (br.t, J = 4.8) | eq. |
| H-6 | 3.80 (m) | ax. |
| H-7 | 2.72 (2H, m) | |
| H-3 | 2.10 (m) | eq |
| H-3'<br>H-5 | 1.85 (m) | ax.<br>eq. |
| H-4 | 1.71 (m) | eq. |
| H-5'<br>H-4' | 1.74 (m) | ax.<br>ax. |

**Table 12: $^1$H-NMR data of compound 19 (500MHz, $D_2O$).**

For compound **19**: m/z (negative ion electrospray ionisation) 200 [M-H$^-$], 156 [M-COOH$^-$]. The conformation adopted by compound **19** was confirmed by coupling constants and nOes between different protons around the ring.

**Example 20**

**(2*S*,6*S*)-6-((*S*)-1-carboxyethyl)piperidine-2-carboxylic acid, 20**




(2*S*,6*S*)-6-((*S*)-1-carboxyethyl)piperidine-2-carboxylic acid

| Proton no. | δH compound **20** | δH compound **21** |
|---|---|---|

| | | |
|---|---|---|
| H-2 | 3.90 (br.t, J = Hz) | 3.91 (br.t, J = Hz) |
| H-6 | 3.14 (m) | 3.17 (m) |
| H-7 | 2.47 (dq, J = 4.8,7.4 Hz) | 2.34 (dq, J = 8.1,7.4 Hz) |
| H-3 | 1.77 (bd) | 1.77 (bd) |
| H-3' | 1.34 (m) | 1.45 (m) |
| H-5 | | |
| H-4 | | 1.31(m) |
| H-5' H-4' | 0.96 (m) | 0.99m |
| 7Me | 0.71 (d, J = 7.4 Hz) | 0.75 (d, J = 7.4 Hz) |

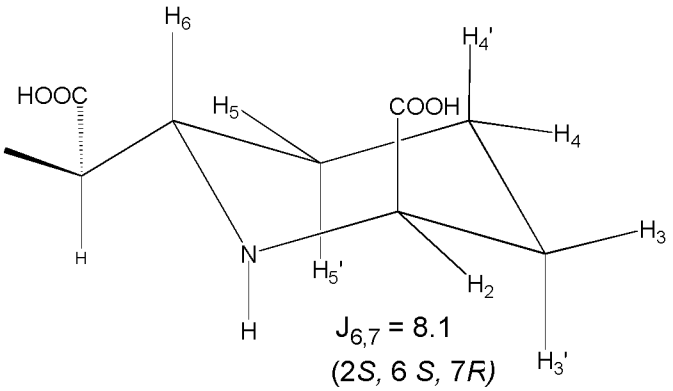
**Table 13: $^1$H-NMR data of compounds 20 and 21 (700MHz, $D_2O$).**

For compounds **20** and **21**: m/z (negative ion electrospray ionisation) 200 [M-H$^-$], 156 [M-COOH$^-$]. The stereochemistry of the side chain (C-7) was assigned as (*S*)- based on:

- The $J_{6,7}$ value of 4.8 Hz as well as the strong nOe between H-6 and H-7 indicating the *gauche* relationship of the two protons.
- The nOe between H-7 and H-5 > H-5' as well as that between the methyl group on C-7 and H-5' > H-5.

**Example 21**

**(2*S*,6*S*)-6-((*R*)-1-carboxyethyl)piperidine-2-carboxylic acid, 21**



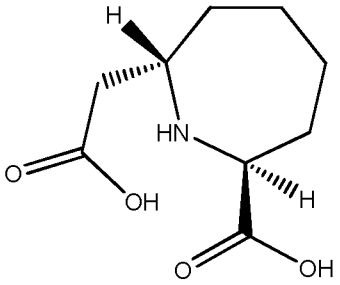

(2*S*,6*S*)-6-((*R*)-1-carboxyethyl)piperidine-2-carboxylic acid

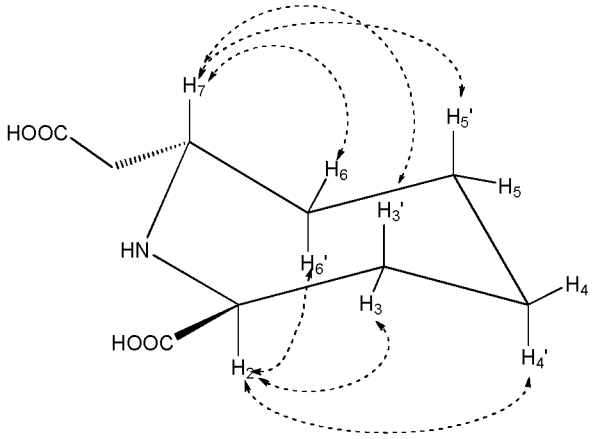
The stereochemistry of C-7 was assigned as (R)- based on:

- The $J_{6,7}$ value of 8.1 Hz as well as the weak nOe between H-6 and H-7 indicating the *anti* relationship of the two protons.
- The nOe between H-7 and H-5' > H-5 as well as that between the methyl group on C-7 and H-5 > H-5'.

**Example 22**

**(2*S*,7*S*)-7-(carboxymethyl)-7-methylazepane-2-carboxylic acid, 22**




(2*S*,7*S*)-7-(carboxymethyl)azepane-2-carboxylic acid

| Proton no. | δH compound 22 | Conformation |
|---|---|---|
| H-2 | 3.76 (dd, J = 11.3, 2.75 Hz) | ax. |
| H-7 | 3.67 (m) | ax. |
| H-8 | 2.67 (2H, m) | |
| H-3 | 2.24 (m) | eq. |
| H-6<br>H-4 | 1.85 (m) | eq.<br>eq. |
| H-5 | 1.77 (m) | eq. |
| H-3' | 1.70 (m) | ax. |
| H-6' | 1.56 (m) | ax. |
| H-4'<br>H-5' | 1.36 (m) | ax.<br>ax. |

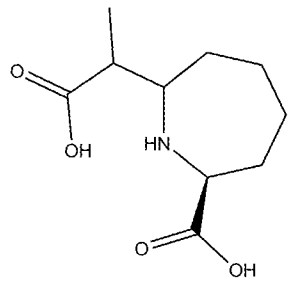
**Table 15: $^1$H-NMR data of compound 22 (700MHz, $D_2O$).**

For compound 22: m/z (negative ion electrospray ionisation) 200 [M-H$^-$], 156 [M-COOH$^-$]. The arrangement and correlation between different protons around the ring was assigned from the $^1$H-NMR, 2D COSY and 2D HSQC. The preferred conformation adopted by

compound **22** was predicted to be chair like on the basis of coupling constants and nOes between different protons around the ring.

**Example 23**

**(2*S*)-7-(1-carboxyethyl)azepane-2-carboxylic acid, 23**



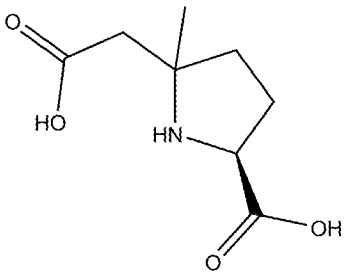

(2*S*)-7-(1-carboxyethyl)azepane-2-carboxylic acid

For compound **23**: m/z (positive ion electrospray ionisation) 216 [$M^+ + 1$], 170 [$M^+$-COOH].

Spectra are shown in Figures 40 and 42.

**Example 24**

**(2*S*)-5-(carboxymethyl)-5-methylpyrrolidine-2-carboxylic acid, 24**




(2*S*)-5-(carboxymethyl)-5-methylpyrrolidine-2-carboxylic acid

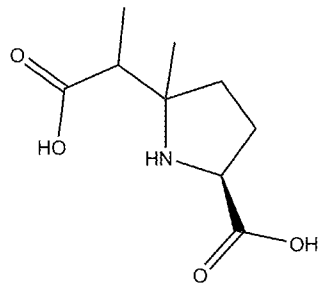
For compound **24**: m/z (positive ion electrospray ionisation) 188 [$M^+ + 1$], 142 [$M^+$-COOH].

Compound **24**, unlike most of other catalytic products in this study, displayed limited solubility in $D_2O$ and other NMR solvents. The $^1$H-NMR spectrum of compound **24** in $D_2O$ revealed characteristic protons e.g. the methylene protons of the side chain which appear as an AB quartet at 2.43 ppm similar in pattern to that observed for the *t*-CMP derivative when all the ring protons were replaced with deuterium. Other $^1$H-NMR signals for **24** occur at: 1.3 (3H, Me), 1.35-1.5 (4H, m. 4 x 3-H and 4-H) and 4.11 (1H, dd, 2-H).

Spectra are shown in Figures 41 and 43.

**Example 25**

**(2*S*)-5-(1-carboxyethyl)-5-methylpyrrolidine-2-carboxylic acid, 25**




(2*S*)-5-(1-carboxyethyl)-5-methylpyrrolidine-2-carboxylic acid

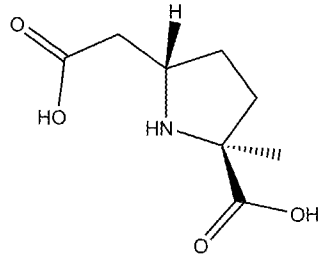
For compound **25**: m/z (positive ion electrospray ionisation) 202 [$M^+ + 1$], 156 [$M^+$-COOH]. Compound **25** displayed very limited solubility in all NMR solvents tried.

This example shows that a quaternary centre has been formed enzymatically.

Spectra are shown in Figure 43.

**Example 26**

**(2*S*,5*S*)-5-(carboxymethyl)-2-methylpyrrolidine-2-carboxylic acid, 26**



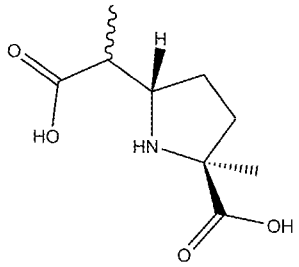

(2*S*,5*S*)-5-(carboxymethyl)-2-methylpyrrolidine-2-carboxylic acid

For compound **26**: m/z (positive ion electrospray ionisation) 188 [$M^+ + 1$], 142 [$M^+$-COOH]. The starting material for compound **26** is a racemic mixture of (2*S*/2*R*)-2-methylglutamate semialdehyde. Assuming the stereochemistry at C-2 is *S* (based on the previous reports on CarB and ThnE which revealed that only L-GSA is a substrate for the enzyme [ Hamed et al ChemBioChem, 2009, 10, 246-250 and Sorensen et al, Chem. Commun., 2005,1155-1157]), the stereochemistry at C-5 was assigned as *S* based on the 2D NOESY correlations between ring proton.

Spectra are shown in Figures 44, 45 and 46.

**Example 27 and 28**

**(2*S*,5*S*, 6*R*)-5-(1-carboxyethyl)-2-methylpyrrolidine-2-carboxylic acid, 27** and **(2*S*,5*S*, 6S)-5-(1-carboxyethyl)-2-methylpyrrolidine-2-carboxylic acid, 28**




(2*S*,5*S*)-5-(1-carboxyethyl)-2-methylpyrrolidine-2-carboxylic acid

For compound **27** and **28**: m/z (positive ion electrospray ionisation) 202 [$M^+ + 1$], 156 [$M^+$-COOH].

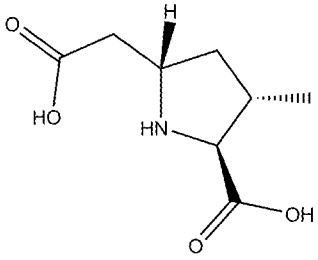
Assuming the stereochemistry at C-2 is *S* (based on the previous reports on CarB and ThnE which revealed that only L-GSA is a substrate for the enzyme ( Hamed et al ChemBioChem, 2009, 10, 246-250 and Sorensen et al, Chem. Commun., 2005,1155-1157), the stereochemistry at C-5 was assigned as *S* based on the 2D NOESY correlations between ring protons.

The NMR analyses of compounds 27 and 28 revealed that, as anticipated, the two compounds are different only at the stereochemistry of C-6. However, due to the free rotation about C5-C6 bond, it was not possible to securely assign the stereochemistry at C-6 of the two diasteromers.

Spectra are shown in Figures 45, 47, 48, 49 and 50.

**Example 29**

**(2*S*,3*S*,5*S*)-5-(carboxymethyl)-3-methylpyrrolidine-2-carboxylic acid, 29**



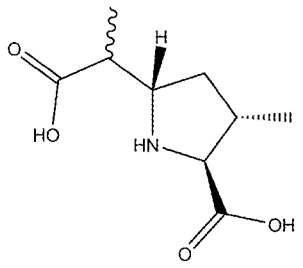

(2*S*,3*S*,5*S*)-5-(carboxymethyl)-3-methylpyrrolidine-2-carboxylic acid

For compound **29**: m/z (positive ion electrospray ionisation) 188 [$M^+ + 1$], 142 [$M^+$-COOH]. The NMR analyses of the compound supported the structural assignment. Based on the known stereochemistry at C-2 and/or C-3, the stereochemistry at C-5 was assigned as *S* based on the 2D NOESY correlations between ring protons.

Spectra are shown in Figures 51, 52, 53 and 54.

**Example 30**

**(2*S*,3*S*,5*S*)-5-(1-carboxyethyl)-3-methylpyrrolidine-2-carboxylic acid, 30**




(2*S*,3*S*,5*S*)-5-(1-carboxyethyl)-3-methylpyrrolidine-2-carboxylic acid

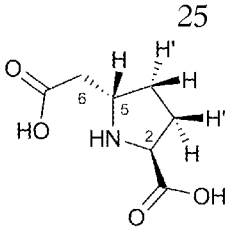
For compound **30**: m/z (positive ion electrospray ionisation) 188 [$M^+ + 1$], 142 [$M^+$-COOH]. The NMR analyses of the compound supported the structural assignment. Due to signals overlap (e.g. H3 and H4, 3Me and 6Me groups), it was not possible to securely assign the stereochemistry at C-6.

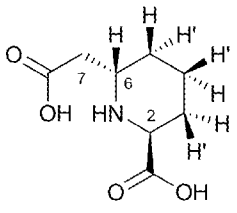
Spectra are shown in Figures 54, 55 and 56.

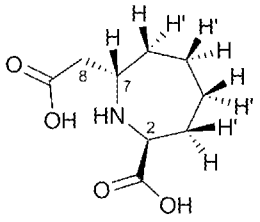
**Example 31**

Carboxymethyl-substituted *N*-heterocycles (5, 6 and 7 membered rings) were produced by the use of CarB (5 and 6 membered rings) and CarBH229A (7 membered ring)

25



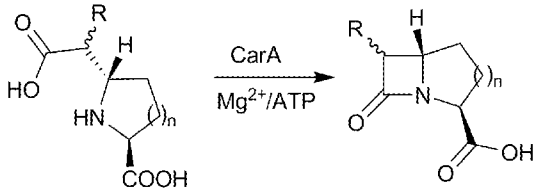




$^1$H-NMR spectra are shown in Figure 57.

**Example 32**

The products from Example 31 were successfully converted to the corresponding bicyclic β-lactams by carbapenam synthetase (CarA) in accordance with the following scheme.



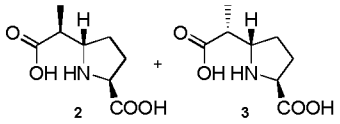
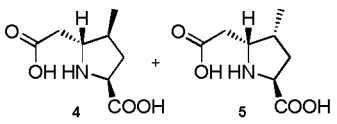
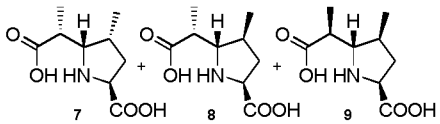
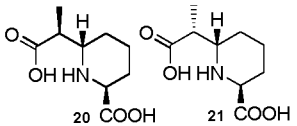
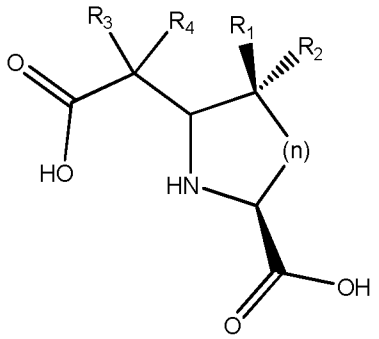
| |  |  |  |  |
|---|---|---|---|---|
| CarB | 45:55 | 50:50 | 50:25:25 | 80:20 |
| CarB M108A | 55:45 | 55:45 | 1:1:1 | 87.5:12.5 |
| CarB M108V | 5:95 | 95:5 | 17:75:8 | little product |
| CarB M108Ile | 5:95 | 95:5 | 19:69:12 | little product |
| CarB M108L | 55:45 | 60:40 | 43:31:26 | 75:25 |
| CarB Q111N | 25:75 | 70:30 | 20:60:20 | little product |
| CarB W79F/W79FM108A | 85:15 | 62:38 | 41:14:45 | 50:50 |
| CarB H229A | 30:70 | 70:30 | 1:1:1 | 50:50 |
| ThnE | 20:80 | 90:10 | No product | No product |
| ThnE V153M | 87:13 | | | |
| ThnE V153A | 91:9 | | | |

**Table 16: The diasteromeric ratios of different carboxymethylproline/piperidine derivatives produced by CarB/CarB variants/ThnE (under standard assay conditions). All ratios were determined by $^1$H NMR analysis of total catalytic product after purification using LC/MS.**




(I)[a]

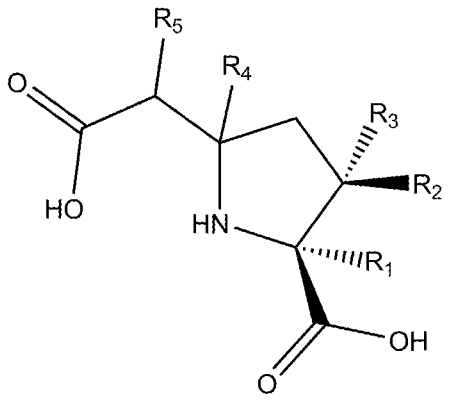
| Substrate | | Compound | $R_1$ | $R_2$ | $R_3$ | $R_4$ | n | Stereochemistry | Biocatalyst (Of highest specificity) |
|---|---|---|---|---|---|---|---|---|---|
| **CoA derivative** | **GSA/ analogue** | | | | | | | | |
| Mal-CoA | GSA | **1** | H | H | H | H | 1 | 2*S*, 5*S* | All |
| Methylmal-CoA | GSA | **2** | H | H | Me | H | 1 | 2*S*, 5*S*, 6*S* | CarBW79F/CarBW79FM108A/ThnEV153A |
| Methylmal-CoA | GSA | **3** | H | H | H | Me | 1 | 2*S*, 5*S*, 6*R* | CarBM108V/I/ThnE |
| Mal-CoA | 4-methylGSA | **4** | Me | H | H | H | 1 | 2*S*, 4*S*, 5*S* | CarBM108V/I/ThnE |
| Mal-CoA | 4-methylGSA | **5** | H | Me | H | H | 1 | 2*S*, 4*R*, 5*S* | CarB |
| Mal-CoA | 4,4-dimethylGSA | **6** | Me | Me | H | H | 1 | 2*S*, 5*S* | All |
| Methylmal-CoA | 4-methylGSA | **7** | H | Me | H | Me | 1 | 2*S*, 4*R*, 5*S*, 6*R* | See Table 16 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Methylmal-CoA | 4-methylGSA | **8** | Me | H | H | Me | 1 | 2*S*, 4*S*, 5*S*, 6*R* | |
| Methylmal-CoA | 4-methylGSA | **9** | Me | H | Me | H | 1 | 2*S*, 4*S*, 5*S*, 6*S* | |
| Methylmal-CoA | 4,4-dimethylGSA | **10** | Me | Me | H | Me | 1 | 2*S*, 5*S*, 6*R* | All except ThnE |
| Dimethylmal-CoA | GSA | **11** | H | H | Me | Me | 1 | 2*S*, 5*S* | All except ThnE |
| Dimethylmal-CoA | 4-methylGSA | **12** | Me | H | Me | Me | 1 | 2*S*, 4*S*, 5*S* | CarB/CarBM108L |
| Ethylmal-CoA | GSA | **13** | H | H | Et | H | 1 | 2*S*, 5*S*, 6*R* | CarBW79F |
| Ethylmal-CoA | GSA | **14** | H | H | H | Et | 1 | 2*S*, 5*S*, 6*S* | CarBW79F |
| Ethylmal-CoA | 4-methylGSA | **15** | H | Me | Et | H | 1 | 2*S*, 4*R*, 5*S*, 6*R* | CarBW79F |
| Ethylmal-CoA | 4-methylGSA | **16** | Me | H | Et | H | 1 | 2*S*, 4*S*, 5*S*, 6*R* | CarBW79F |
| Ethylmal-CoA | 4-methylGSA | **17** | Me | H | H | Et | 1 | 2*S*, 4*S*, 5*S*, 6*S* | CarBW79F |
| Ethylmal-CoA | 4,4-dimethylGSA | **18** | Me | Me | H | Et | 1 | 2*S*, 5*S*, 6*R* | CarBW79F |
| Mal-CoA | AASA | **19** | H | H | H | H | 2 | 2*S*, 6*S* | All |
| Methylmal-CoA | AASA | **20** | H | H | Me | H | 2 | 2*S*, 6*S*, 7*S* | See Table 16 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Methylmal-CoA | AASA | **21** | H | H | H | Me | 2 | 2*S*, 6*S*, 7*R* | |
| Mal-CoA | APSA | **22** | H | H | H | H | 3 | 2*S*, 7*S* | CarB/CarBH229A |
| Methylmal-CoA | APSA | **23** | H | H | H | Me | 3 | 2*S* | CarBH229A |

**Table 17-A: Compound summary demonstrating the substrates used during synthesis and the optimum variant used during synthesis**

5




| Substrate | | Compound | R1 | R2 | R3 | R4 | R5 | Stereochemistry | Biocatalyst (Of highest specificity) |
|---|---|---|---|---|---|---|---|---|---|
| CoA derivative | GSA/ analogue | | | | | | | | |
| Mal-CoA | 5-methylGSA | **24** | H | H | H | Me | H | 2*S* | ThnE/ThnE variants/CarBW79F |
| Methylmal-CoA | 5-methylGSA | **25** | H | H | H | Me | Me | 2*S* | ThnE/ThnE variants |
| Mal-CoA | 2-methylGSA | **26** | Me | H | H | H | H | 2*S*, 5*S* | CarB/CarBH229A |
| Methylmal-CoA | 2-methylGSA | **27** | Me | H | H | H | Me | 2*S*, 5*S* | CarB/CarBH229A |
| Methylmal-CoA | 2-methylGSA | **28** | Me | H | H | H | Me | 2*S*, 5*S* | CarB/CarBH229A |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mal-CoA | 3-methylGSA | **29** | H | H | Me | H | H | 2*S*, 3*S*, 5*S* | ThnE/ThnE variants |
| Methylmal-CoA | 3-methylGSA | **30** | H | H | Me | H | Me | 2*S*, 3*S*, 5*S* | ThnE/ThnE variants |

Table 17-B **Compound summary demonstrating the substrates used during synthesis and the optimum variant used during synthesis**

**Abbreviations:**

**GSA:** glutamate semialdehyde **AASA:** amino adipate semialdehyde **APSA:** amino pimelate semialdehyde **Mal-CoA:** malonyl-coenzyme A

## CLAIMS

1. A process for preparing an enantiomerically enriched compound containing a substituted heterocyclic ring, said process comprising a carbon-carbon bond formation reaction in the presence of a crotonase superfamily protein or a homolog or variant thereof.

2. A process for preparing a substituted heterocyclic ring, said process comprising a carbon-carbon bond formation reaction in the presence of a crotonase superfamily protein or a homolog or variant thereof.

3. The process of claim 1 or claim 2, wherein the prepared compound is a *trans*-carboxymethylproline moiety.

4. The process of claim 3, wherein the stereocentre created by C-C bond formation is *trans* with respect to the C-2 carboxylate group.

5. The process of any one of claims 1 to 4, wherein the crotonase superfamily protein comprises one of the following: 3,5-Dihydroxyphenylglyoxylate synthase (DpgC), Transcarboxylase 12S (TC 12S), *Anabaena* β-Diketone hydrolase (ABDH), 6-Oxocamphor hydrolase (6-OCH), 4-Chlorobenzoyl-CoA dehalogenase (4-CBD), Methylmalonyl-CoA decarboxylase (MMCD), Glutaconyl-CoA de-carboxylase-α subunit (Gcdα), $ECH_2$ Decarboxylase domain of CurF (CurF), Naphthoate synthase (MenB), Adenine, Uracil binding ECH homologue (AUH), Enoyl-CoA hydratase (ECH), Dienyl-DoC isomerase (DCI), Hydroxylcinnamoyl-CoA hydratase-ligase (HCHL), $\Delta^3,\Delta^2$-Enoyl-CoA isomerase (ECI), Acetyl-CoA carboxylase carboxyltransferase subunit from yeast (ACC CT), Carboxymethylproline synthase, the proteolytic subunit of caseinolytic protease (ClpP), Photosystem II D1 CTPase (D1-CTPase), Interphotoreceptor retinoid-binding protein (IRBP) or Tricorn protease (Tricorn).

6. The process of claim 5 wherein the crotonase superfamily protein comprises a carboxymethylproline synthase enzyme.

7. The process of claim 6 wherein the carboxymethylproline synthase enzyme comprises CarB or ThnE or a homolog or variant thereof.

8. The process of claim 7, wherein the crotonase superfamily protein or homolog or variant thereof is wild-type CarB.

9. The process of claim 7, wherein the crotonase superfamily protein variant is a variant of CarB having one or more amino acid substitution between residues 60 and 235.

10. The process of claim 9 wherein the crotonase superfamily protein variant is a variant of CarB having one or more amino acid substitution between residues 75 and 85, and/or between 105 and 115, and/or between 225 and 230.

11. The process of claim 10, wherein the amino acid substitution is at position 108, and/or at position 111, and/or at position 79, and/or at position 229.

12. The process of claim 7, wherein the crotonase superfamily protein or homolog or variant thereof is wild-type ThnE.

13. The process of claim 12, wherein the crotonase superfamily protein variant is a variant of ThnE having one or more amino acid substitution between residues 100 and 290.

14. The process of claim 13, wherein the crotonase superfamily protein variant is a variant of ThnE having one or more amino acid substitution between residues 120 and 130 and/or between 150 and 160 and/or between 270 and 280.

15. The process of claim 14, wherein the amino acid substitution is at position 124 and/or position 153 and/or position 274.

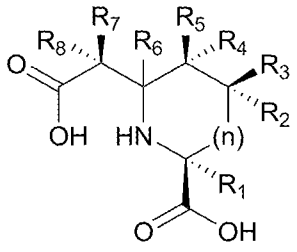
16. The process of any one of the preceding claims, wherein the substituted heterocyclic ring is a monocyclic or bicyclic ring that contains from 3 to 12 ring members.

17. The process of claim 16, wherein the ring is a monocyclic group that contains from 3 to 7 ring members and that contains at least one nitrogen atom.

18. The process of any one of the preceding claims wherein the heterocyclic ring contains a carbon-containing substituent group that is linked to a carbon atom on the heterocyclic ring via a carbon atom in the substituent group.

19. The process of claim 18 wherein the carbon-containing substituent group is present at the C-4 position.

20. The process of any one of the preceding claims, wherein the heterocyclic ring comprises a compound of formula (I):




(I)

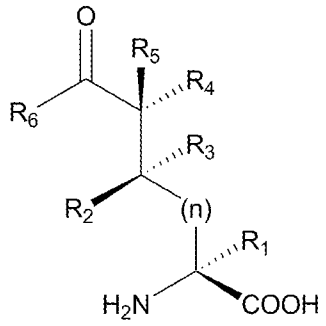
wherein n represents an integer from 0 to 3; and

$R_1$, $R_2$, $R_3$, $R_4$, $R_5$, $R_6$, $R_7$ and $R_8$ independently represent hydrogen or optionally substituted $C_{1-6}$ alkyl.

21. The process of claim 20, wherein $R_1$, $R_2$, $R_3$, $R_4$, $R_5$, $R_6$, $R_7$ and $R_8$ independently represent hydrogen or optionally substituted $C_{1-3}$ alkyl.

22. The process of any one of the preceding claims wherein the process comprises reaction of an amino acid aldehyde, for example a semialdehyde, compound in the presence of a malonyl-CoA compound or derivatives thereof.

23. The process of claim 22, wherein the amino acid aldehyde is an aminoacid semialdehyde compound is a compound of formula (II):

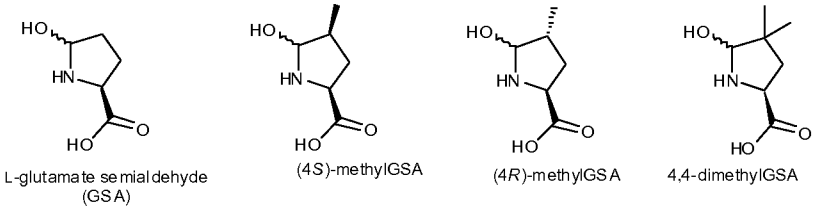


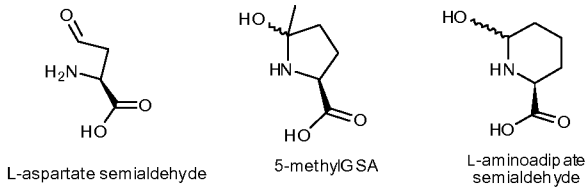

(II)

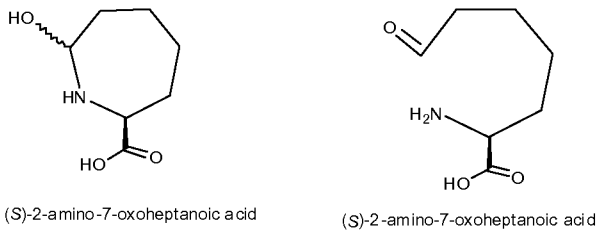
wherein n represents an integer from 0 to 3; and

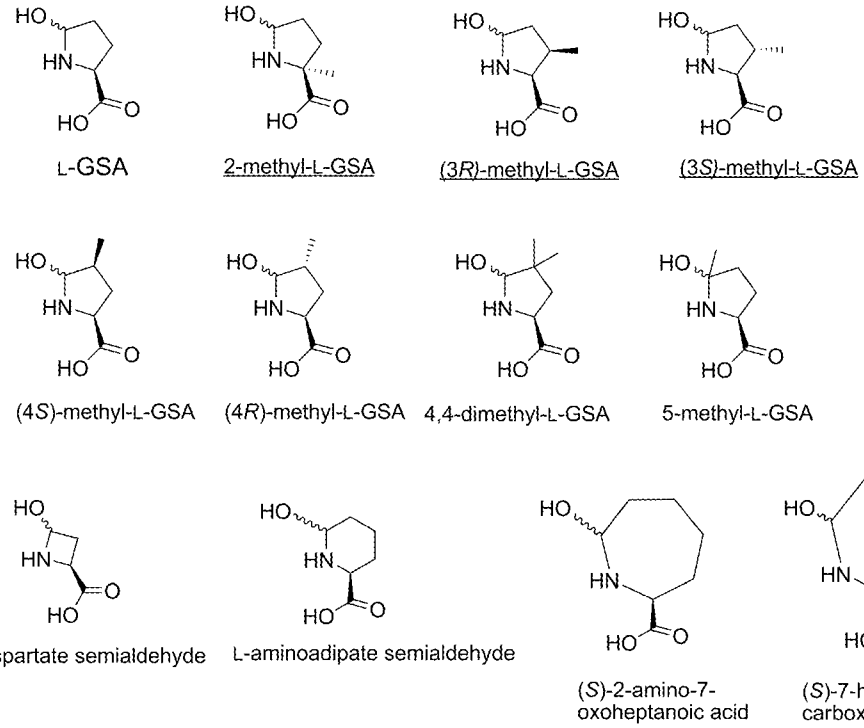
$R_1$, $R_2$, $R_3$, $R_4$, , $R_5$ and $R_6$ independently represent hydrogen or optionally substituted $C_{1-6}$ alkyl.

24. The process of claim 21, wherein the amino acid semialdehyde compound or derivative thereof is selected from any one of:






25. The process of any one of claims 22 to 24, wherein the malonyl-CoA compound or derivative thereof is selected from any one of: malonyl coenzyme A, methylmalonyl coenzyme A, ethylmalonyl coenzyme A, isopropylmalonyl coenzyme A and dimethylmalonyl coenzyme A.

26. A compound containing a substituted heterocyclic ring obtainable by a process of any one of the preceding claims.

27. A process for enhancing the substrate specificity and/or substrate acceptance of a crotonase superfamily protein which comprises the step of preparing a variant of said protein by substituting one or more amino acid residues of the peptide by other amino acid residues and/or deleting one or more amino acid residues from the peptide and/or adding one or more amino acid residues to the peptide.

28. The process of claim 27, wherein the crotonase superfamily protein variant is a variant as defined in any one of claims 9 to 11 or 13 to 15.

29. A crotonase superfamily protein variant which is a variant of CarB as defined in any one of claims 9 to 11 or a variant of ThnE as defined in any one of claims 13 to 15.

30. The variant of claim 29, which is a variant of CarB selected from any one or more of: W79F, W79A, M108A, M108V, M108L, M108I, Q111N and/or H229A or a variant of ThnE selected from any one or more of: V153A, V153M, V153I, V153L, W124F and/or H274A.

31. The variant of claim 30, which is a W79F CarB variant or a W79F/M108A CarB double variant.

32. A heterocyclic ring compound of formula (I) as defined in any one of claims 20 or 21.

33. The use of a wild-type crotonase superfamily protein in the synthesis of compounds containing heterocyclic rings with 6,6'-dialkyl substituents.

34. The use of claim 33 where the wild-type crotonase superfamily protein is selected from those set out in claim 5.

35. The use of claim 33 wherein the wild-type crotonase superfamily protein is wild-type CarB or ThnE.

36. A compound containing a substituted heterocyclic ring as defined in claim 26 or 32 for use in the synthesis of a medicament.

37. The use of claim 36 wherein said medicament is an antibiotic.

38. The use of claim 37, wherein said antibiotic is a substituted proline, a cephem, a carbapenem, or thienamycin.

39. The use of a crotonase superfamily protein variant in the synthesis of thienamycin.

40. The use of claim 39 wherein said variant is a CarB or a ThnE variant.

41. The use of claim 40 wherein said variant is a CarB variant or a ThnE variant as defined in any one of claims 29 or 30.

42. The use of claim 41 wherein said variant is the W79F variant of CarB.

43. A method for preparing a bicyclic beta-lactams by reacting a compound containing a substituted heterocyclic ring as defined in any one of claims 26 or 32 with carbapenam synthetase (CarA).

44. The method of claim 43 wherein the method comprises (i) preparing a compound containing a substituted heterocyclic ring according to a process as defined in any one of claims 1 to 25 and (ii) reacting this compound with carbapenam synthetase (CarA).

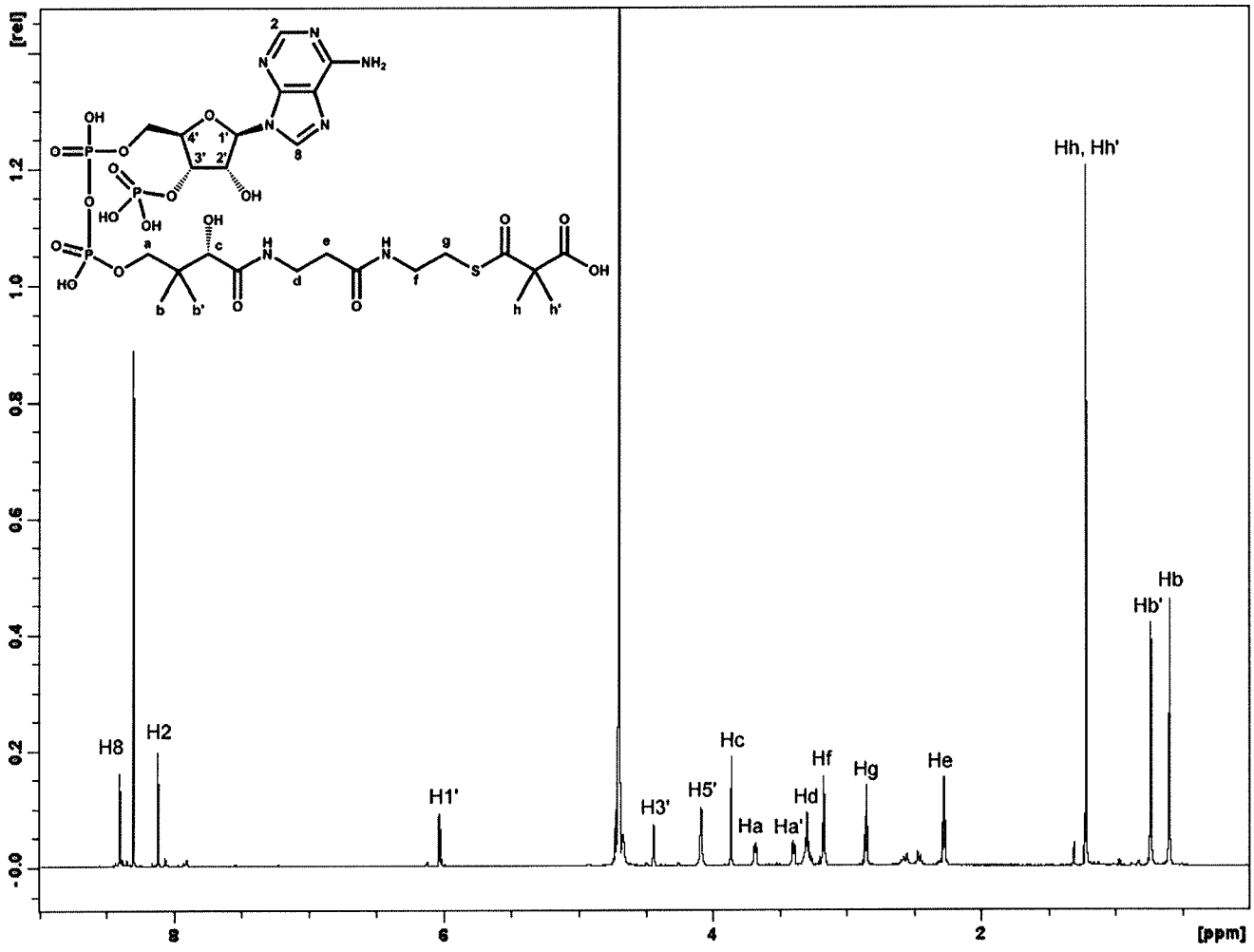



Figure 1: $^{1}$H-NMR spectrum of purified dimethylmalonyl-CoA.

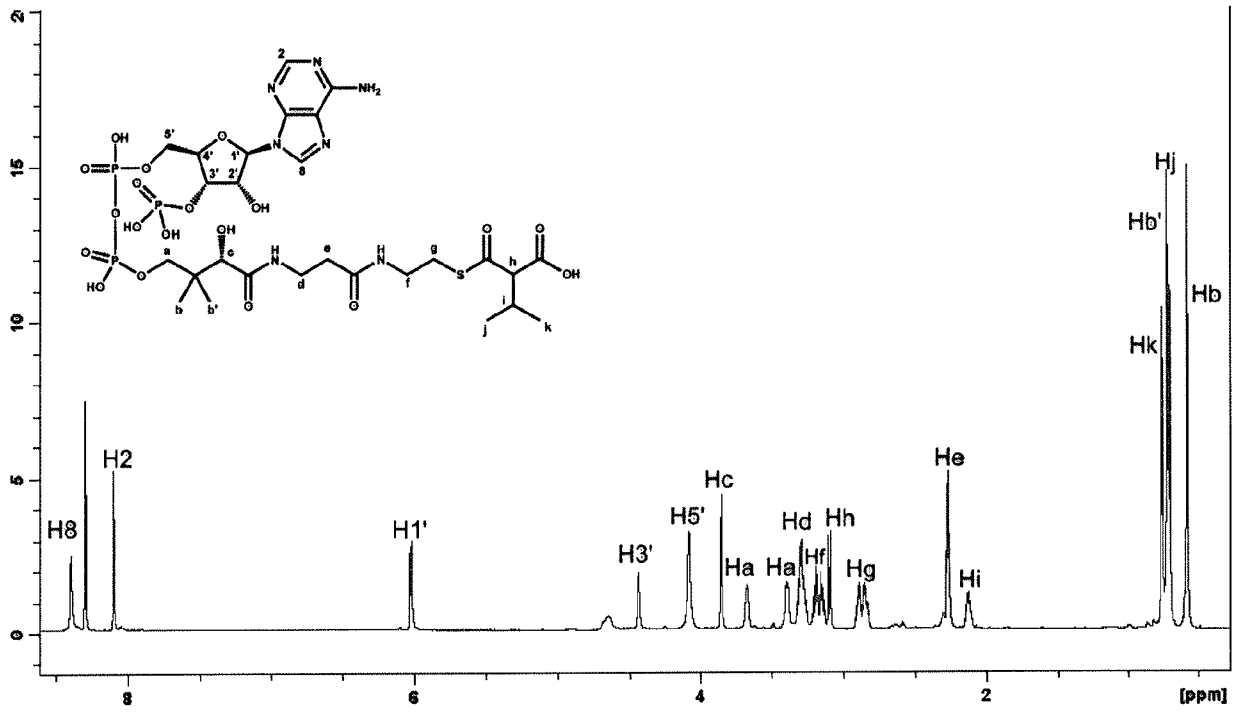



Figure 2: $^{1}$H-NMR spectrum of purified isopropylmalonyl-CoA.

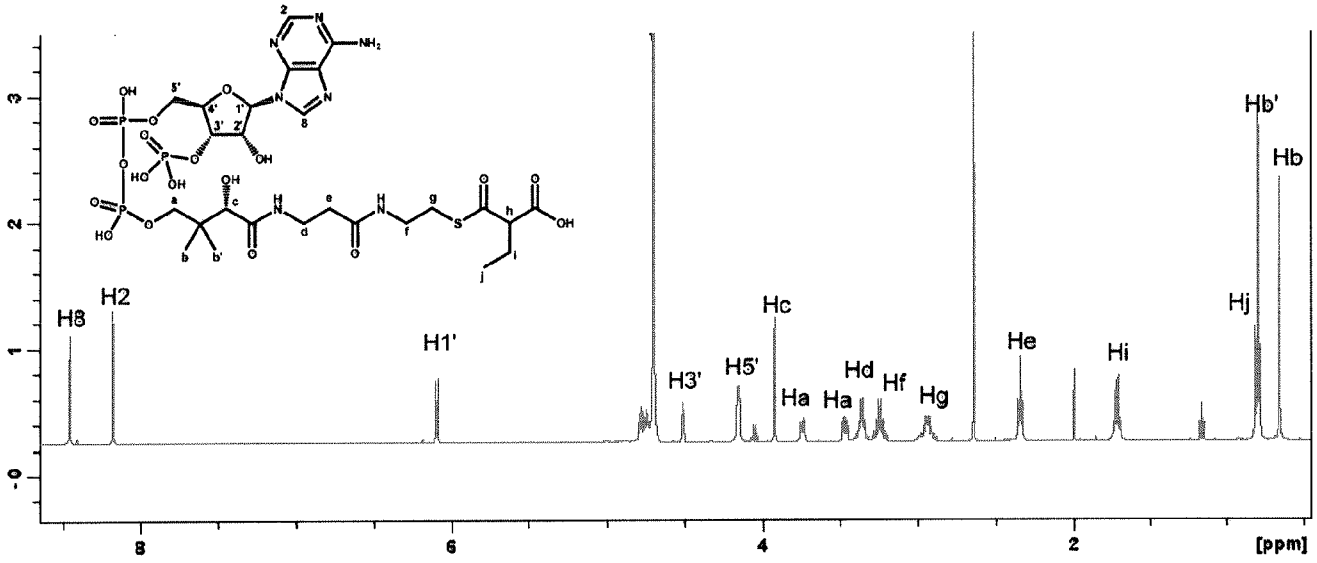



Figure 3: $^{1}$H-NMR spectrum for ethylmalonyl-CoA.

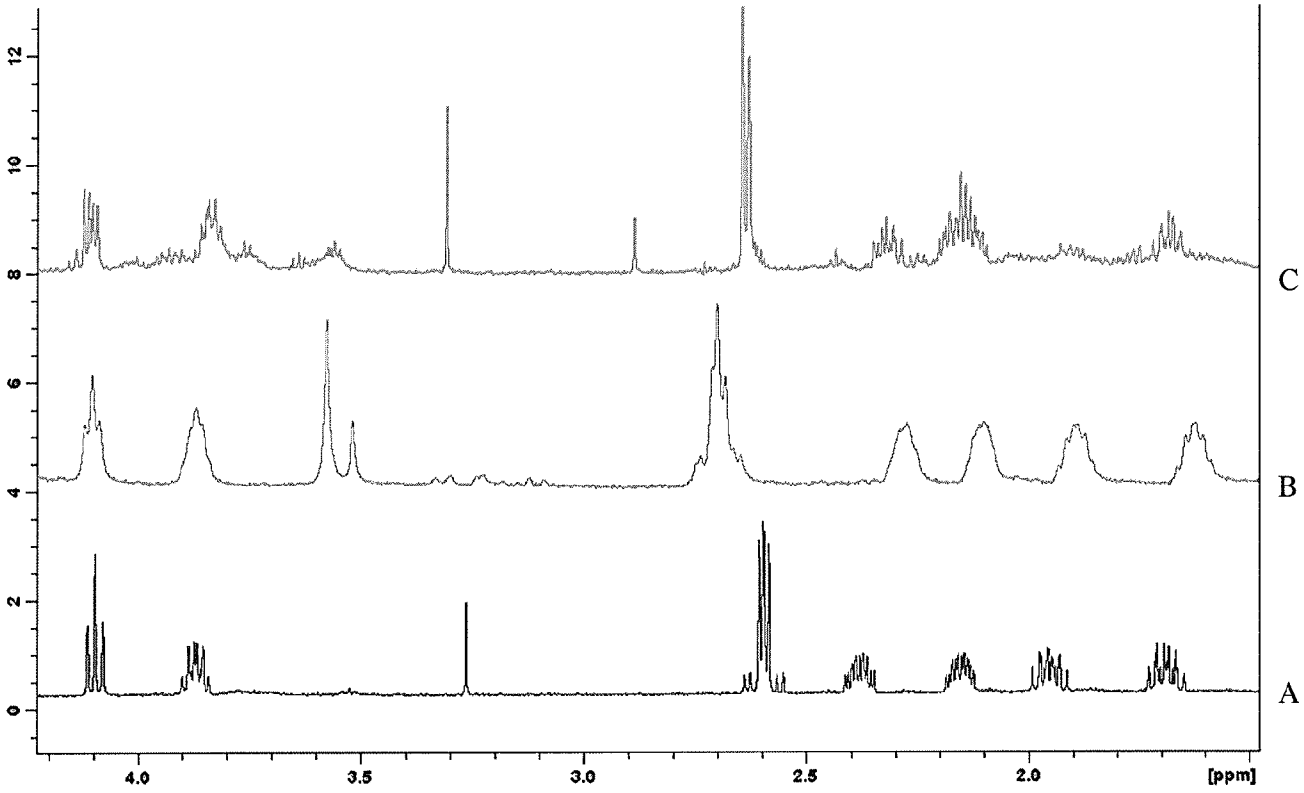



Figure 4: $^{1}$H-NMR spectra of synthetic (2*S*,5*S*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid (A), (2*S*,5*S*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid isolated from a CarB catalysed reaction (B), and synthetic (2*S*,5*R*)-5-(carboxymethyl)pyrrolidine-2-carboxylic acid (C). Relative shifts in positions of certain peaks in A and B are likely due to differences in pH.

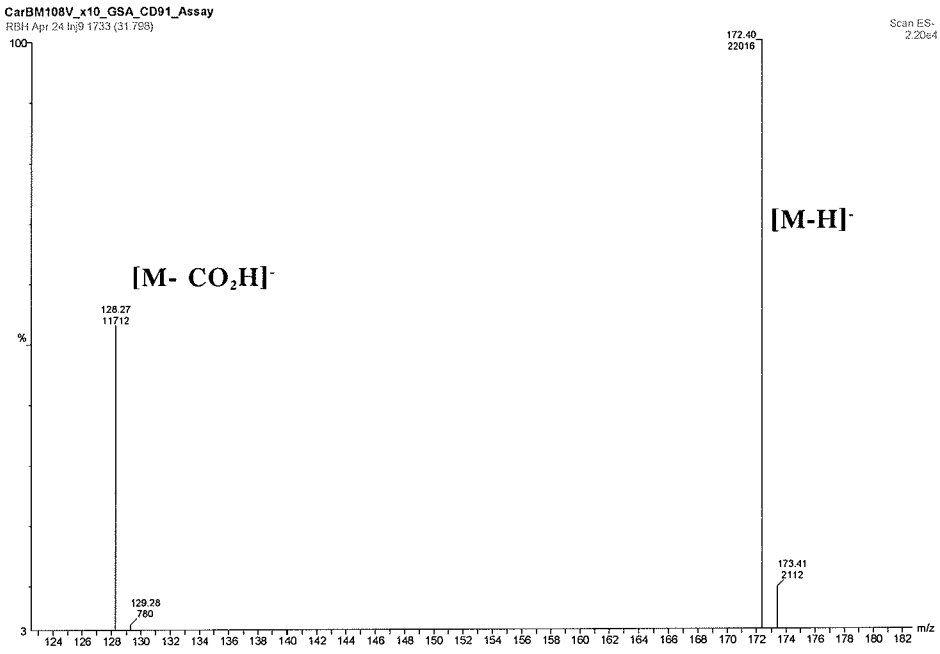



Figure 5: MS spectrum of compound **1** produced in a CarB catalysed reaction.

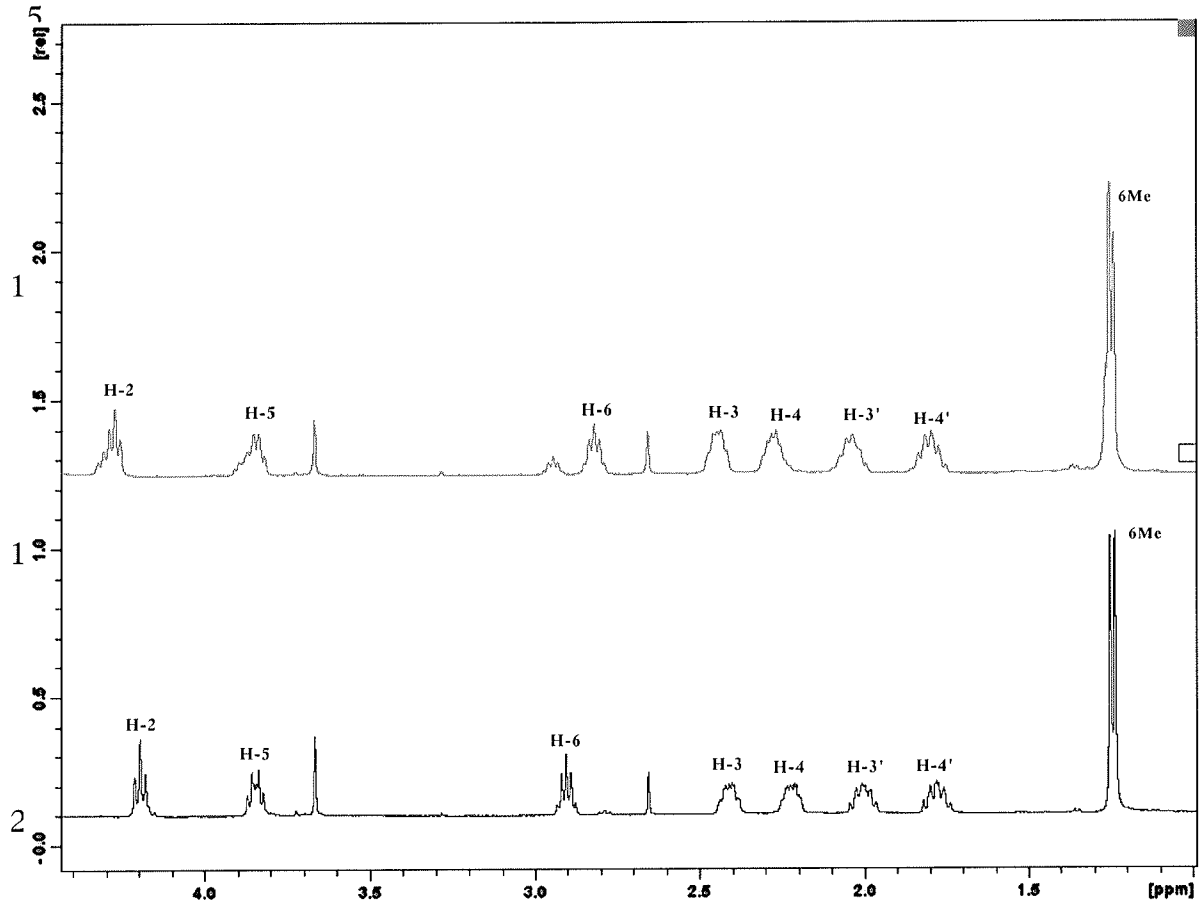



Figure 6: $^{1}$H-NMR spectra for compounds **2** (top) and **3** (bottom) (500MHz, $D_2O$).

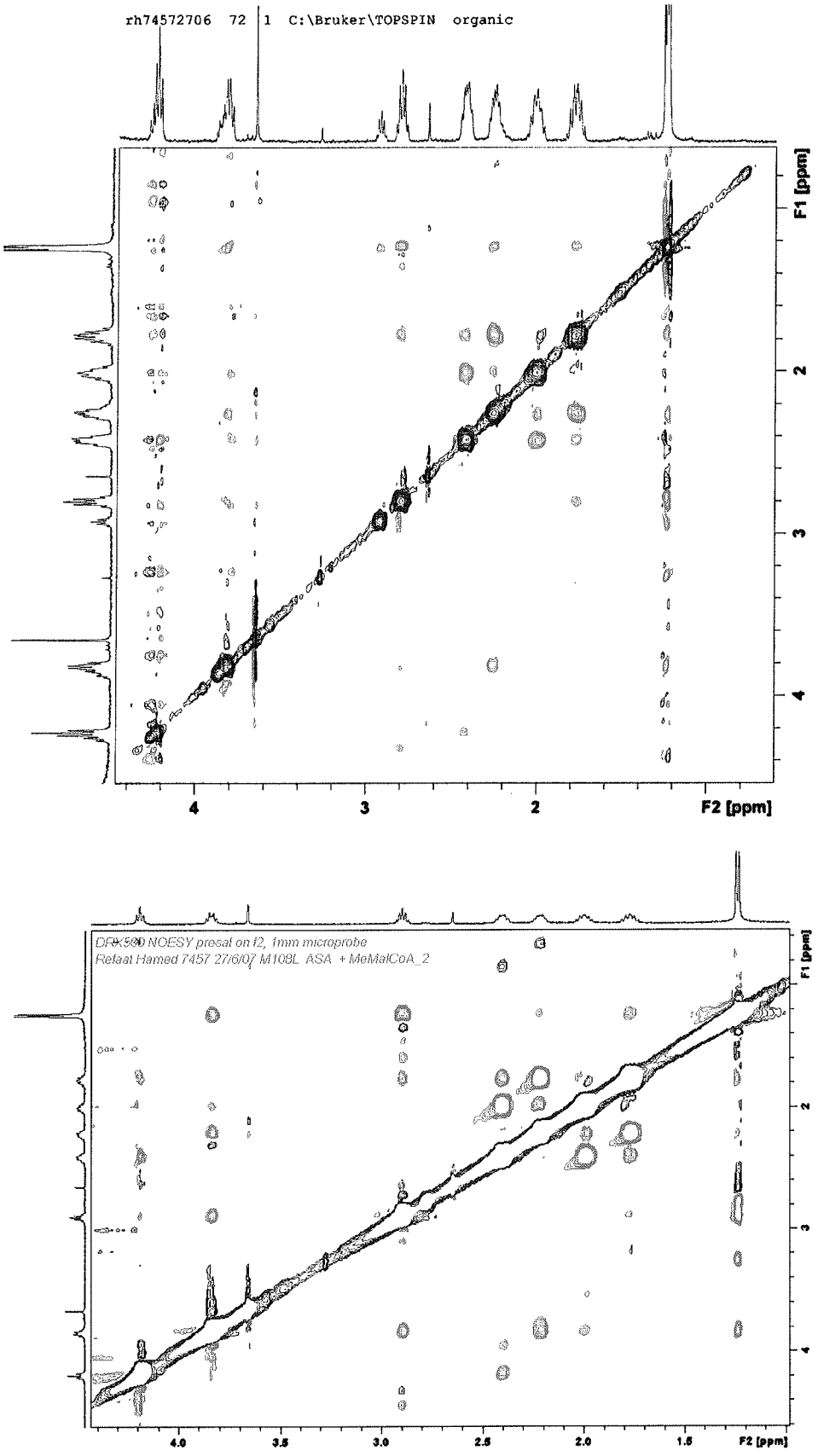



Figure 7: 2D NOESY spectra for compounds **2** (top) and **3** (bottom) (500MHz, $D_2O$).

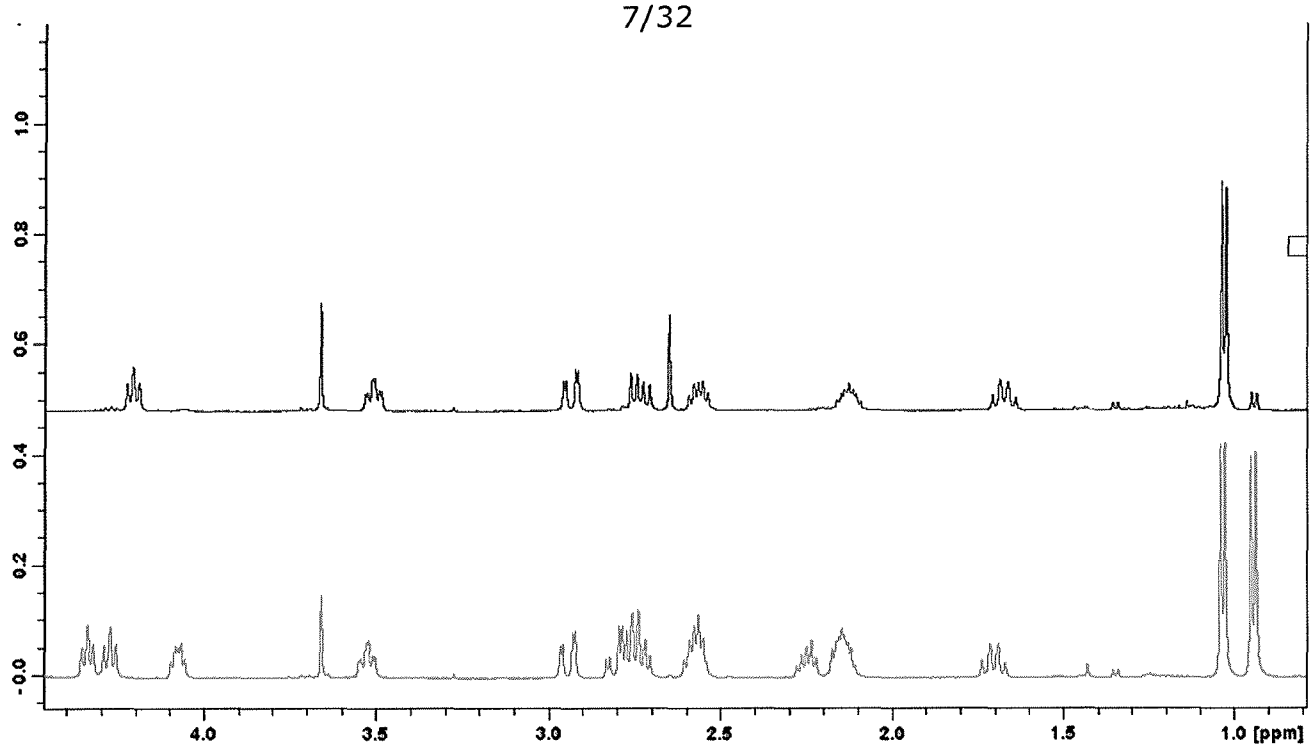



Figure 8: $^{1}$H-NMR spectra for compounds **4** (top) and **5** (bottom, as a mixture with **4**), (500MHz, $D_2O$)

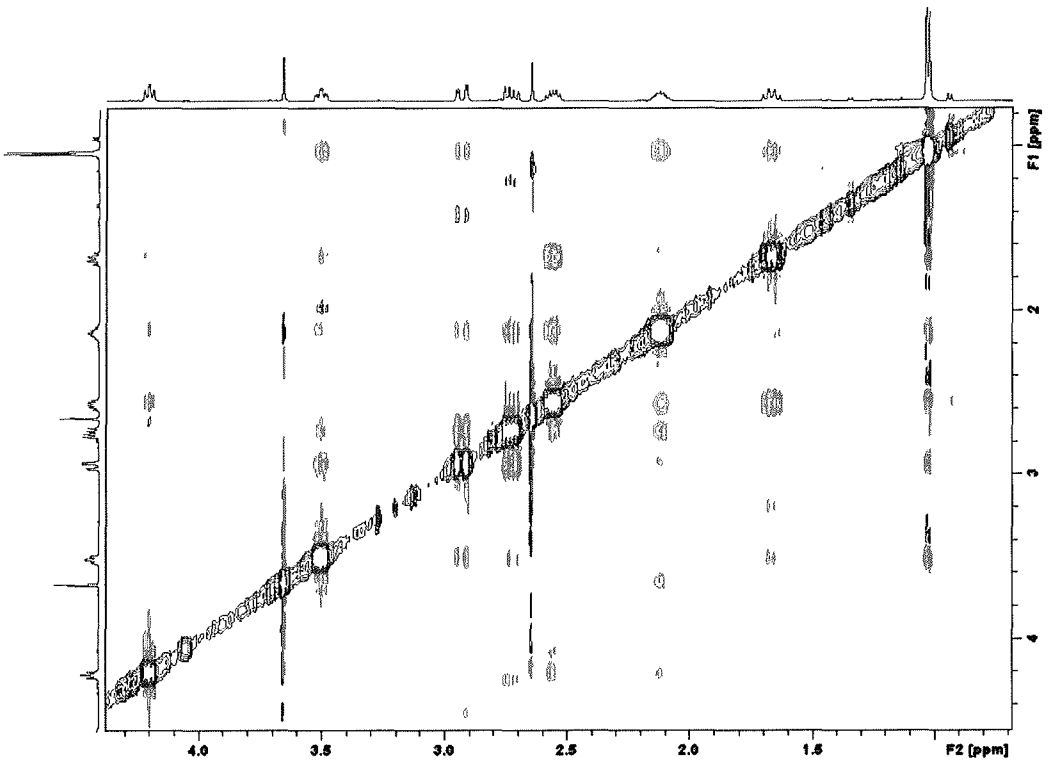



Figure 9: 2D NOESY spectrum for compound **4**, (500MHz, $D_2O$)

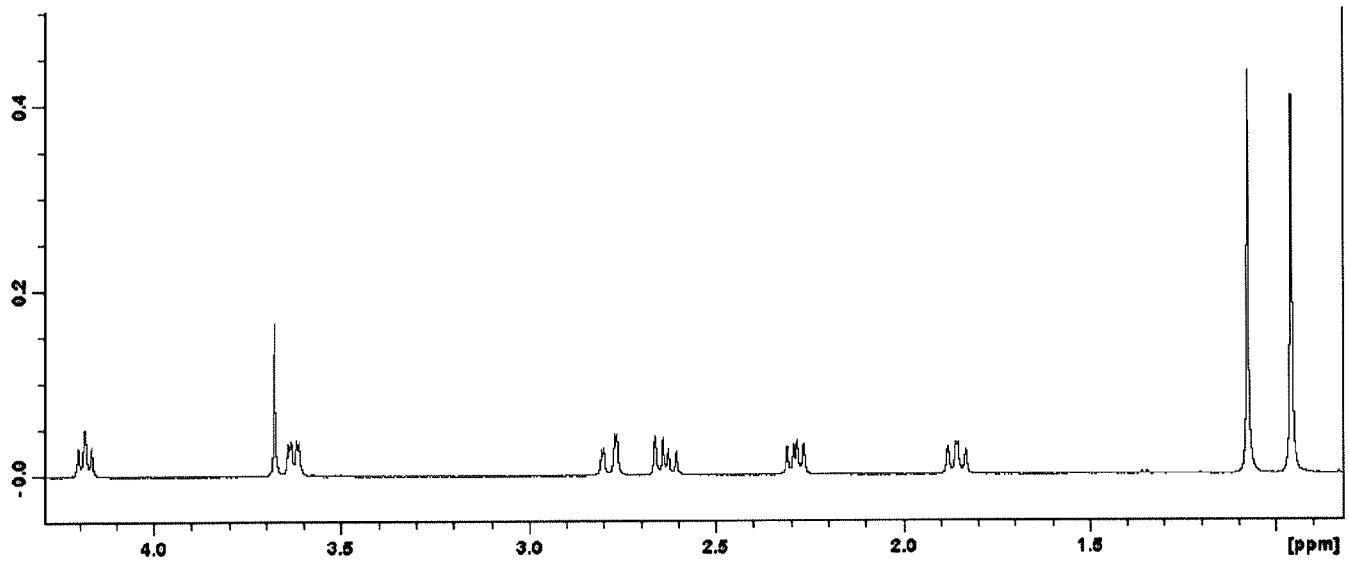



Figure 10: $^{1}$H-NMR spectrum for compound **6**, (500MHz, $D_2O$)

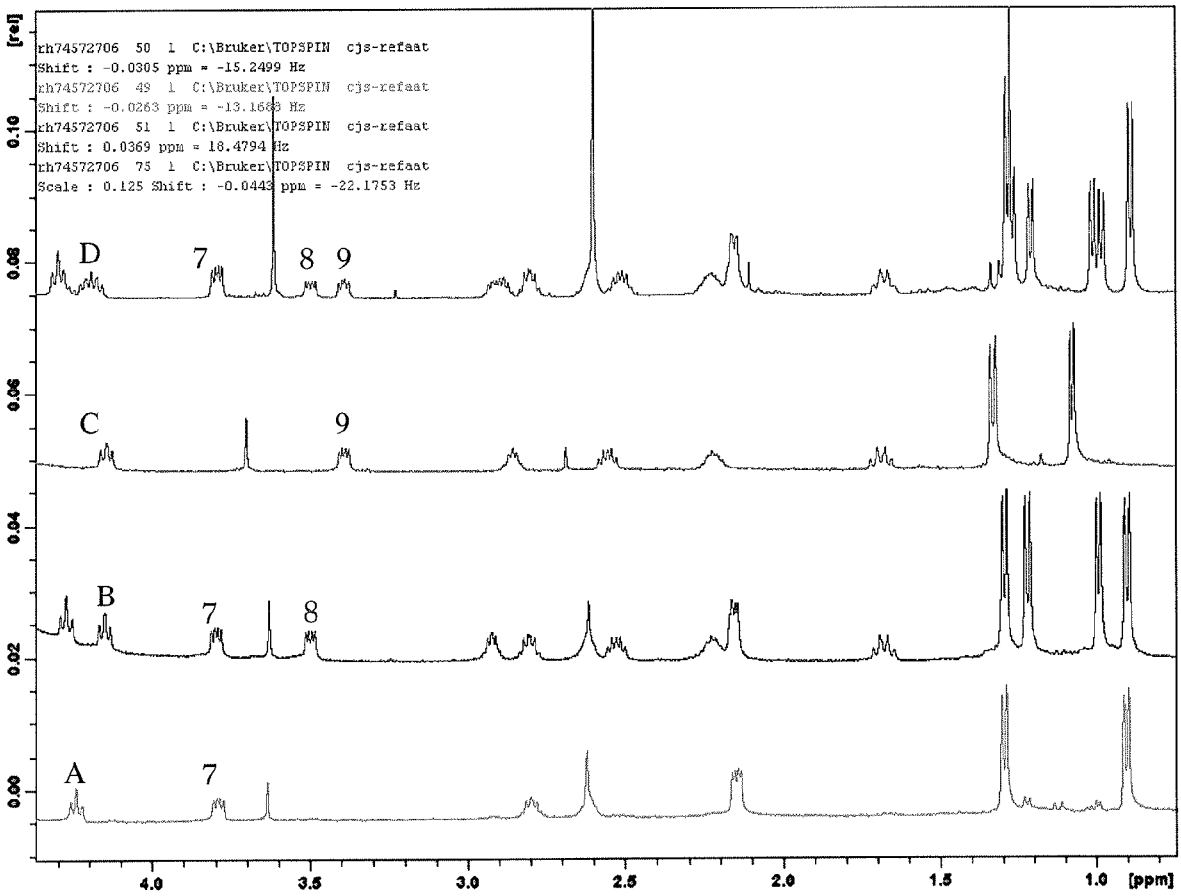



Figure 11: $^{1}$H-NMR spectra for compounds **7**, **8** and **9**. A: purified 7, B: a mixture of 7 and 8, C: purified 9 and D: a mixture of 7, 8 and 9.

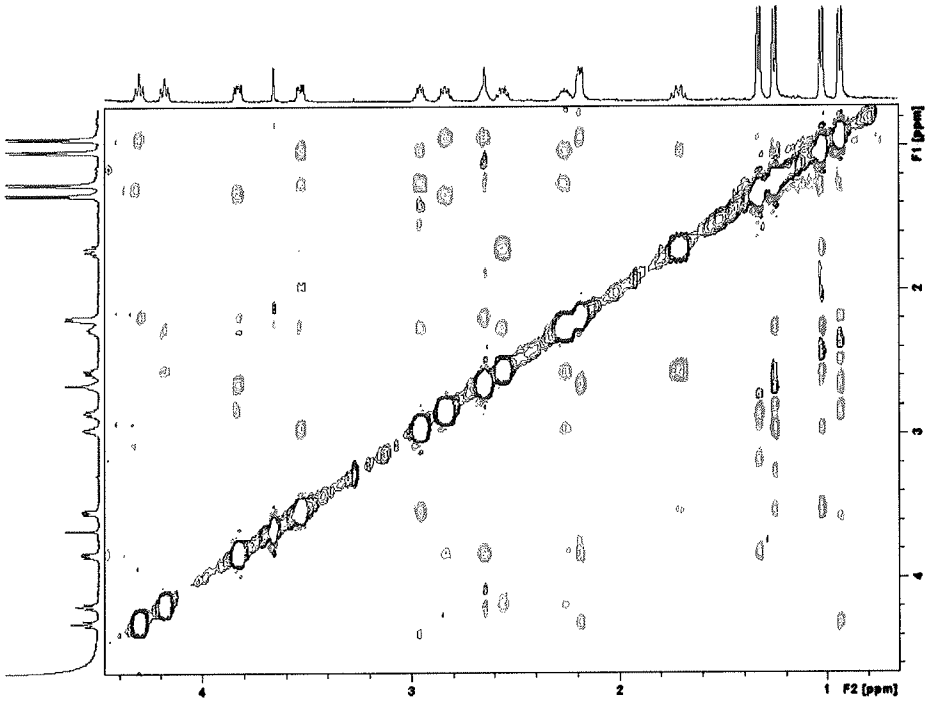



Figure 12: 2D NOESY spectrum for a mixture of compounds **7** and **8**.

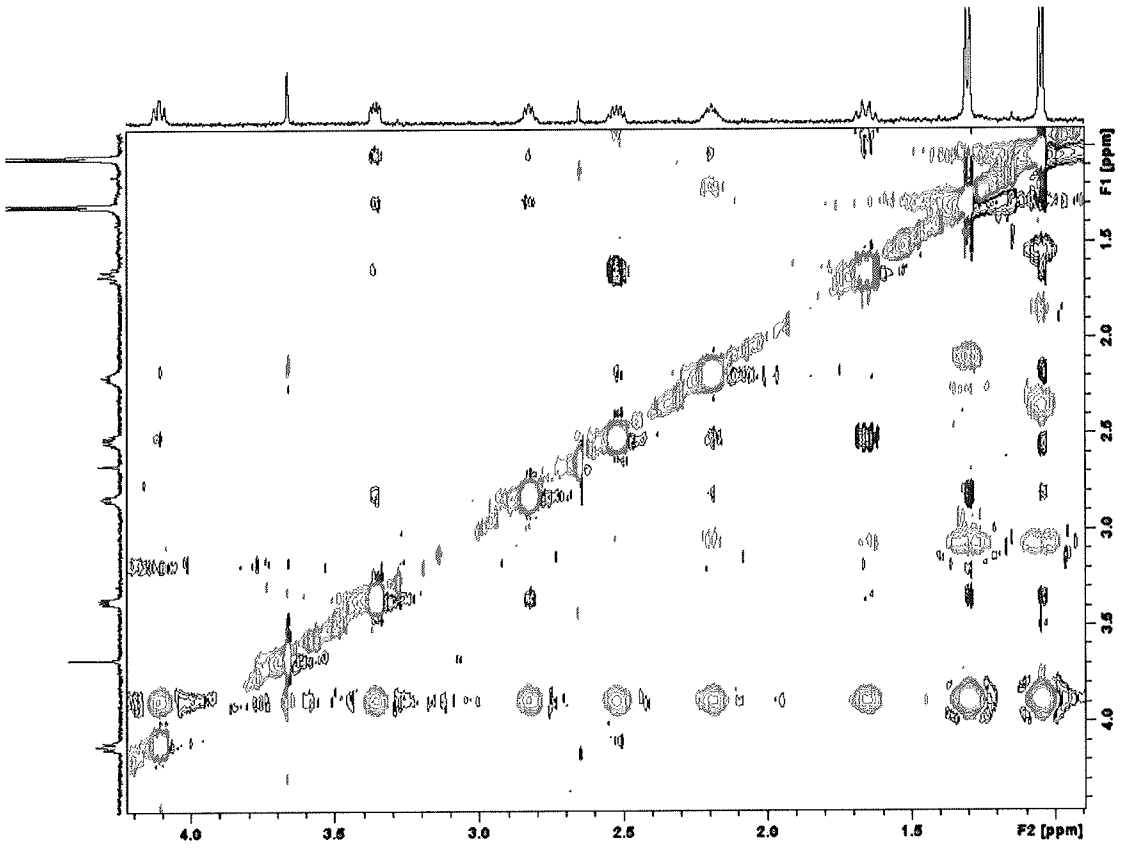



Figure 13: 2D NOESY spectrum for compound **9**.

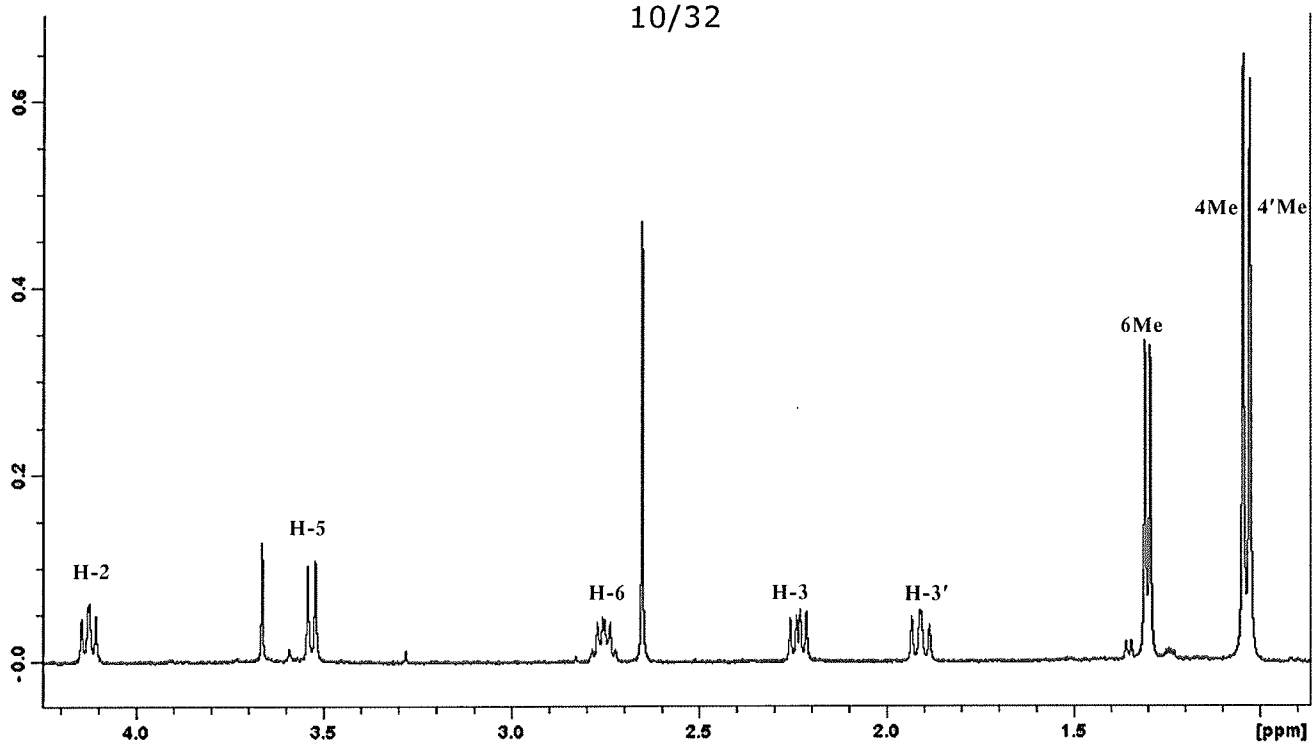



Figure 14: $^{1}$H-NMR spectrum for compound **10**.

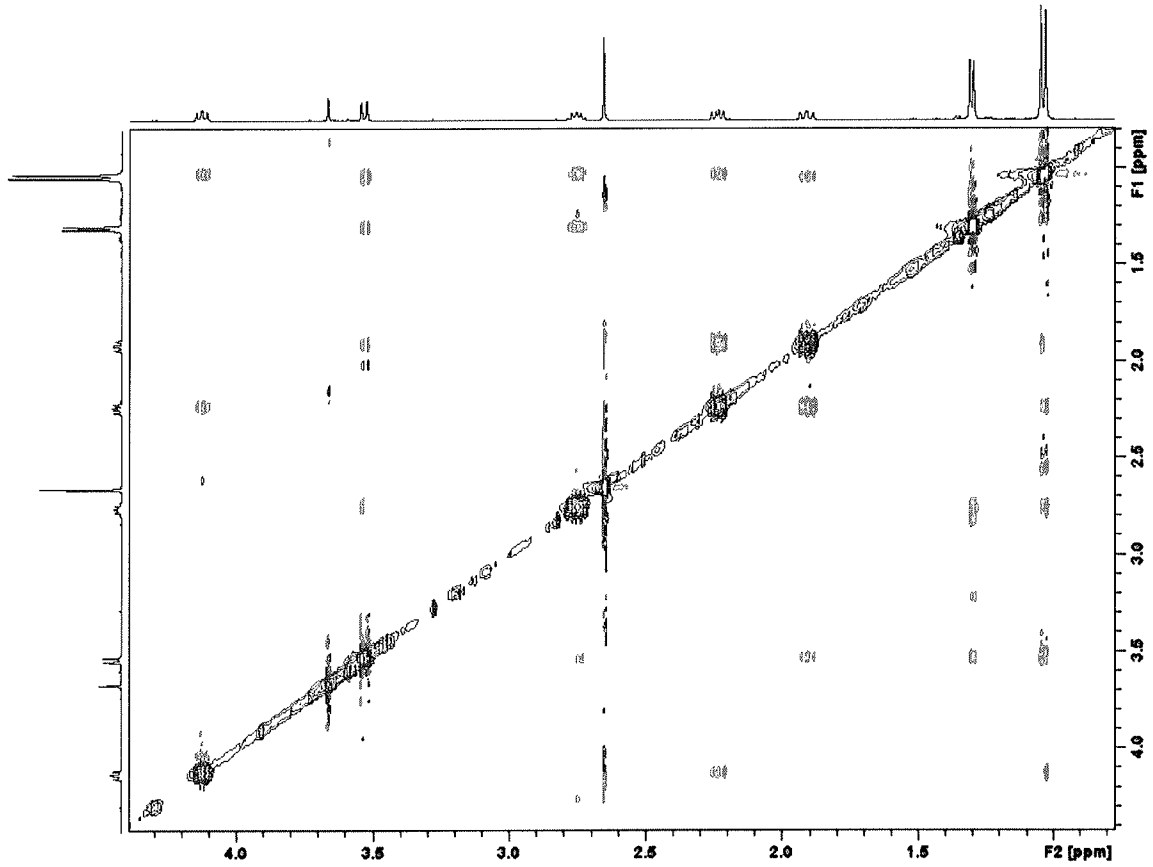



Figure 15: 2D NOESY spectrum for compound **10**.

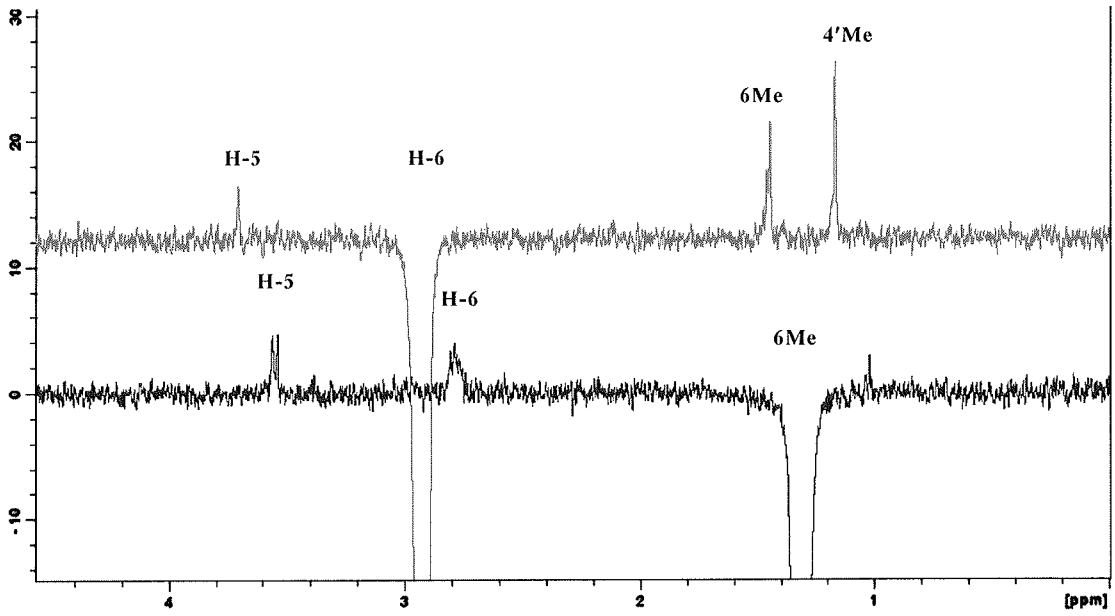



Figure 16: 1D nOe spectra for compound **10** after irradiation of C-6 methyl group (bottom) and H-6 (top), respectively.

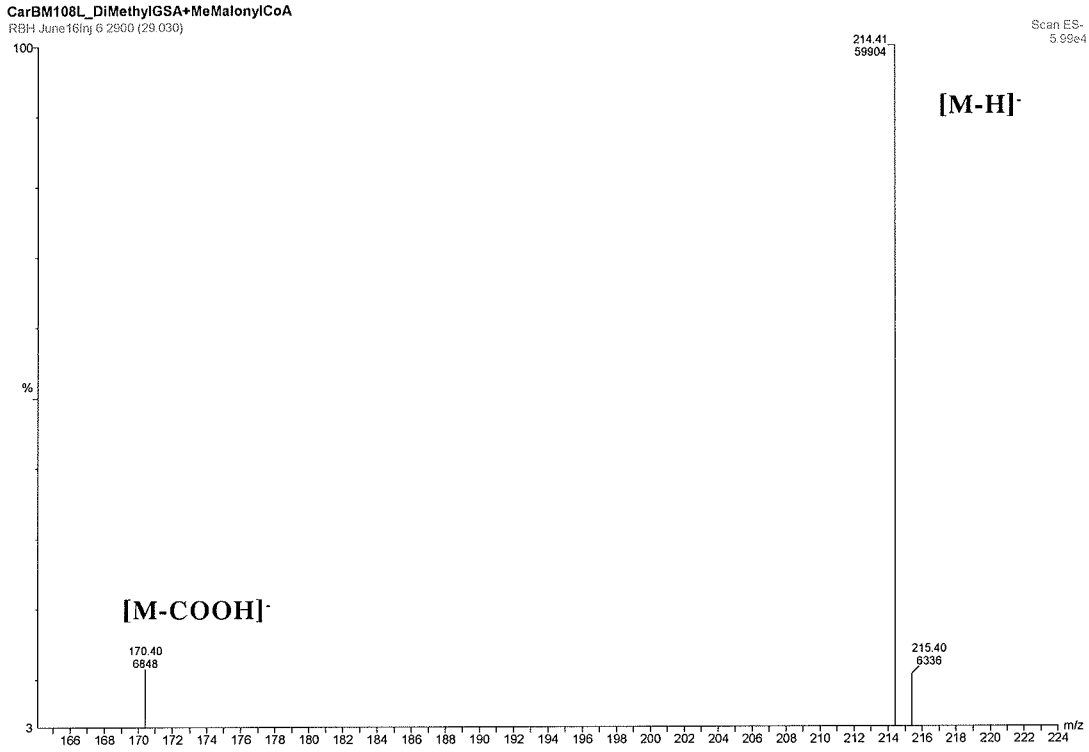



Figure 17: MS spectrum for compound **10**.

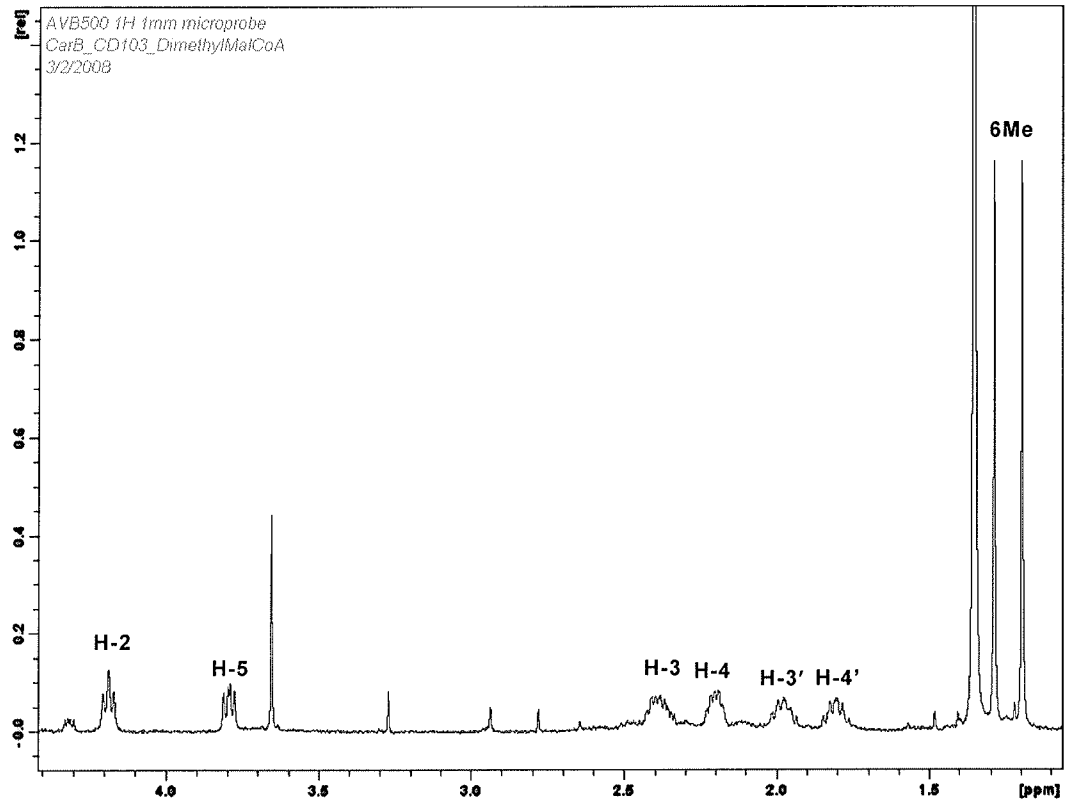



Figure 18: $^{1}$H-NMR spectrum for compound **11**.

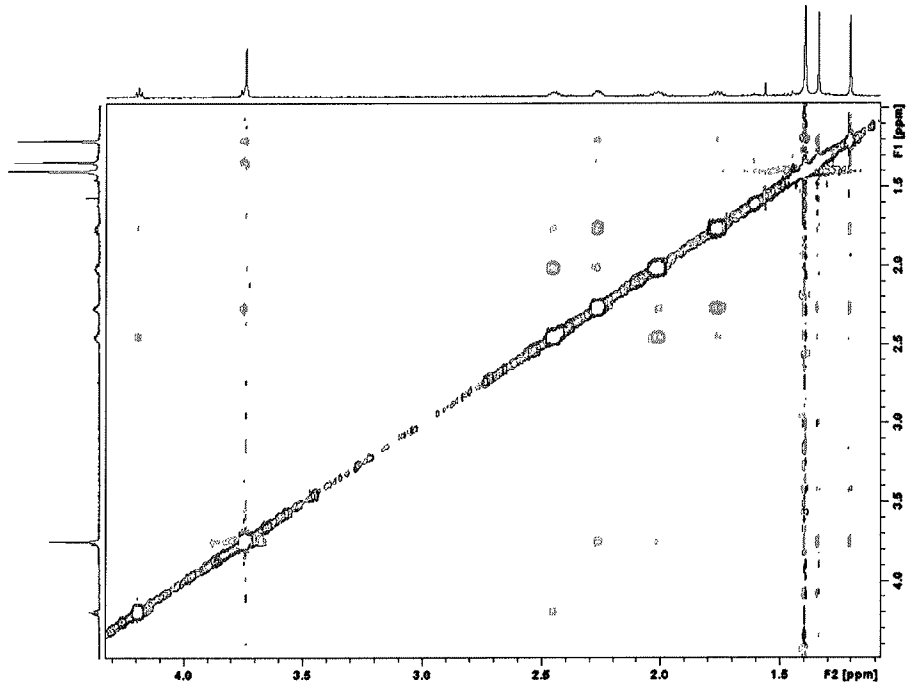



Figure 19: 2D NOESY spectrum for compound **11**.

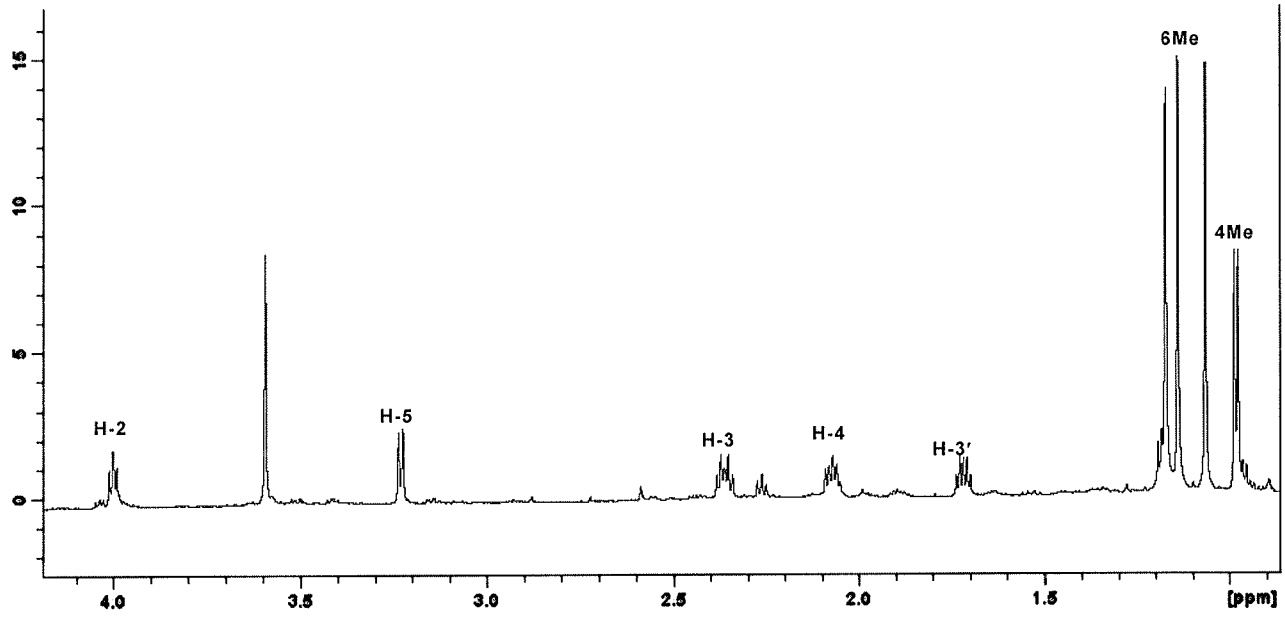



Figure 20: $^{1}$H-NMR spectrum of compound **12**.

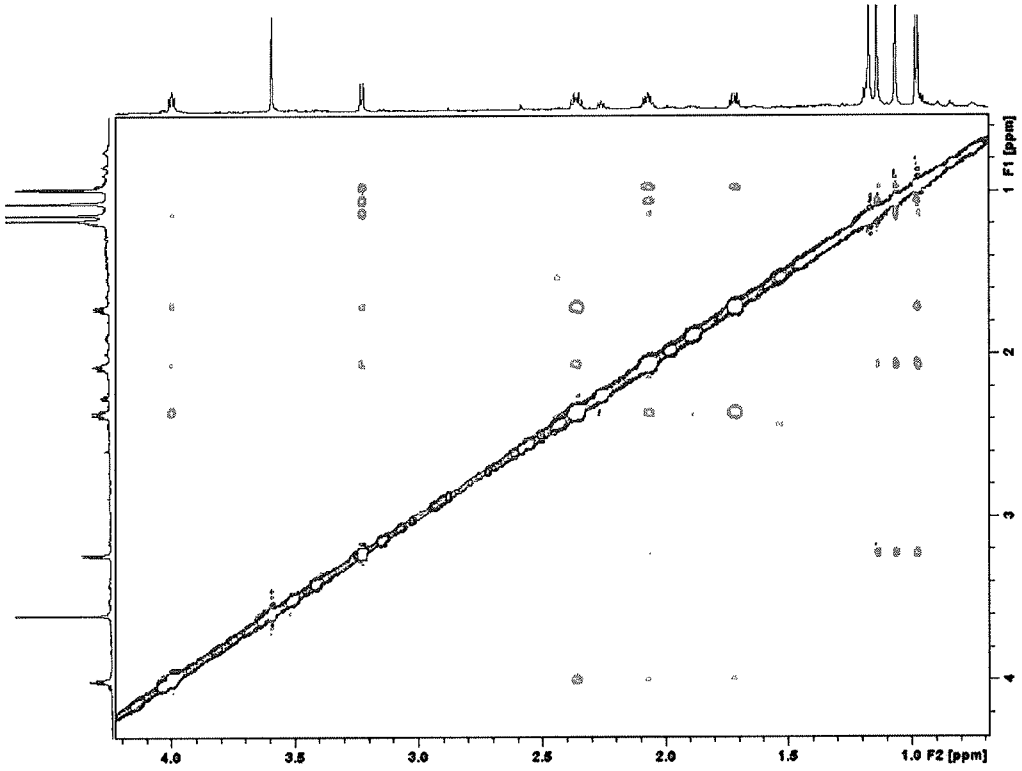



Figure 21: 2D NOESY of compound **12**.

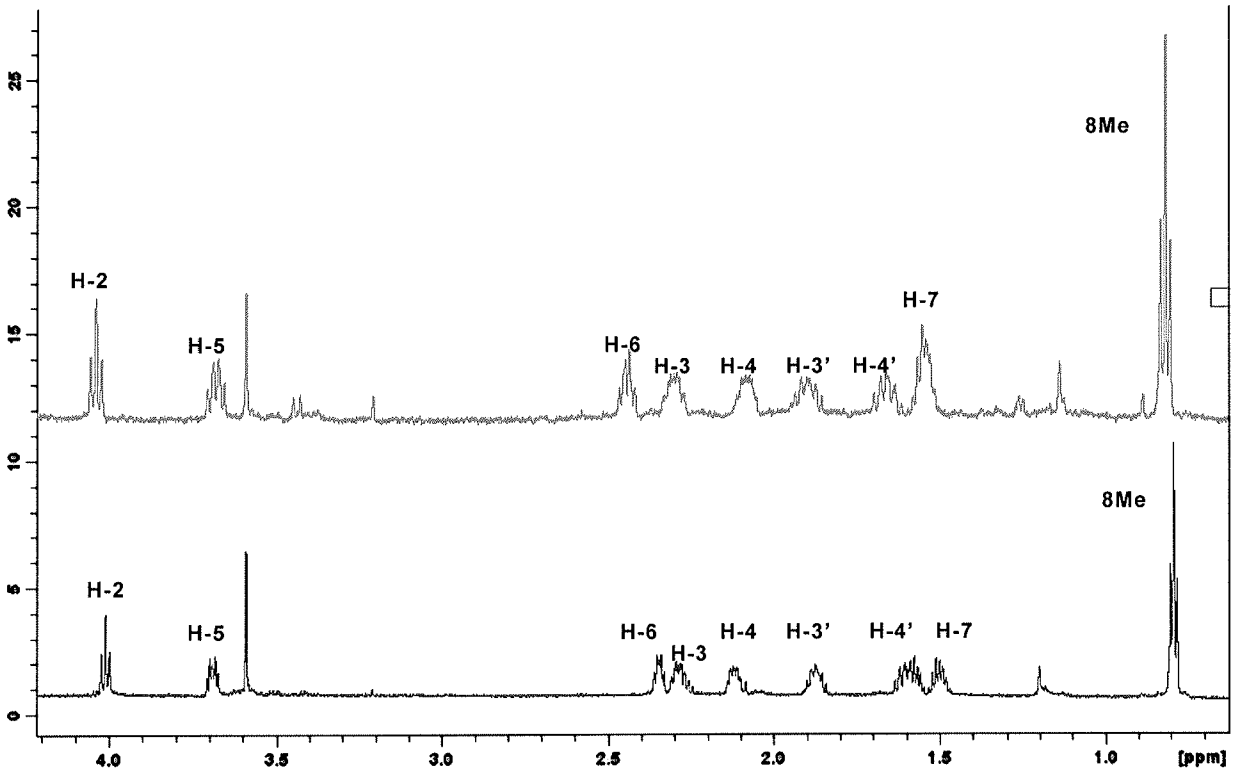



Figure 22: $^{1}$H-NMR spectra for compounds **13** ( top) and **14** (bottom).

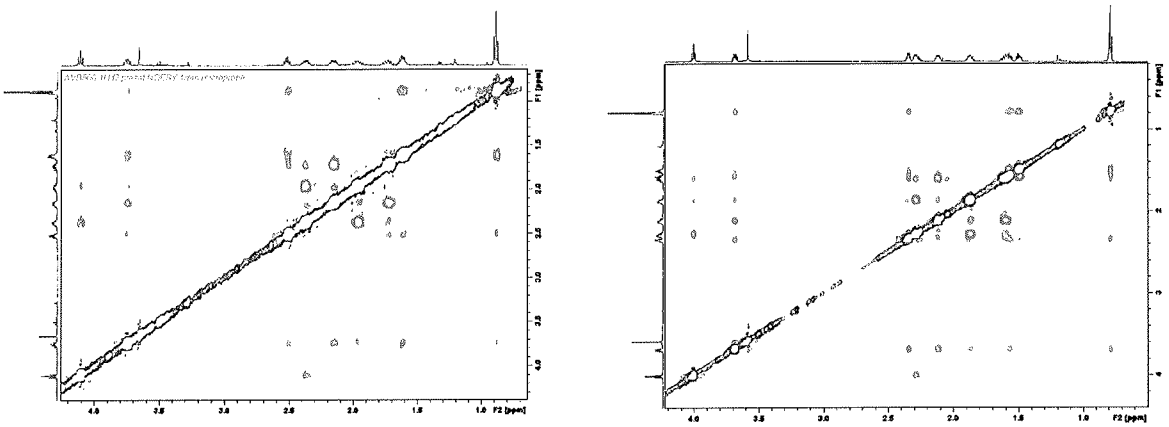


Figure 23: 2 D NOESY spectra for compounds **13** (left) and **14** (right).

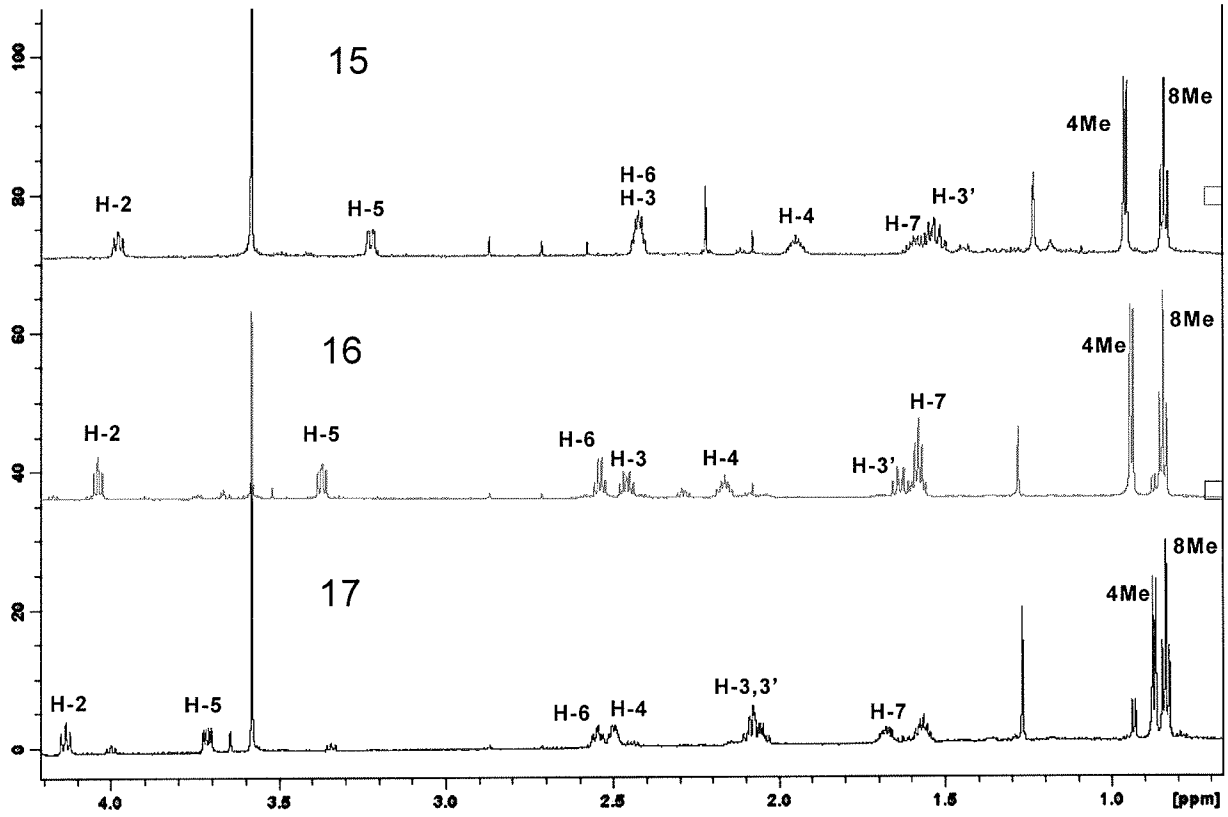



Figure 24: $^{1}$H-NMR spectra for compounds **15** (top), **16** (middle) and **17** (bottom).

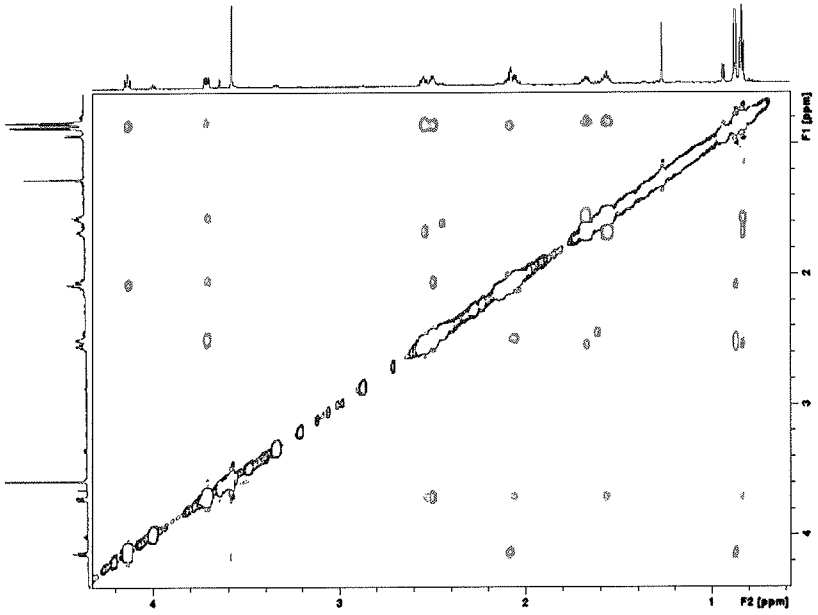



Figure 25: 2D NOESY spectrum for compound **15**.

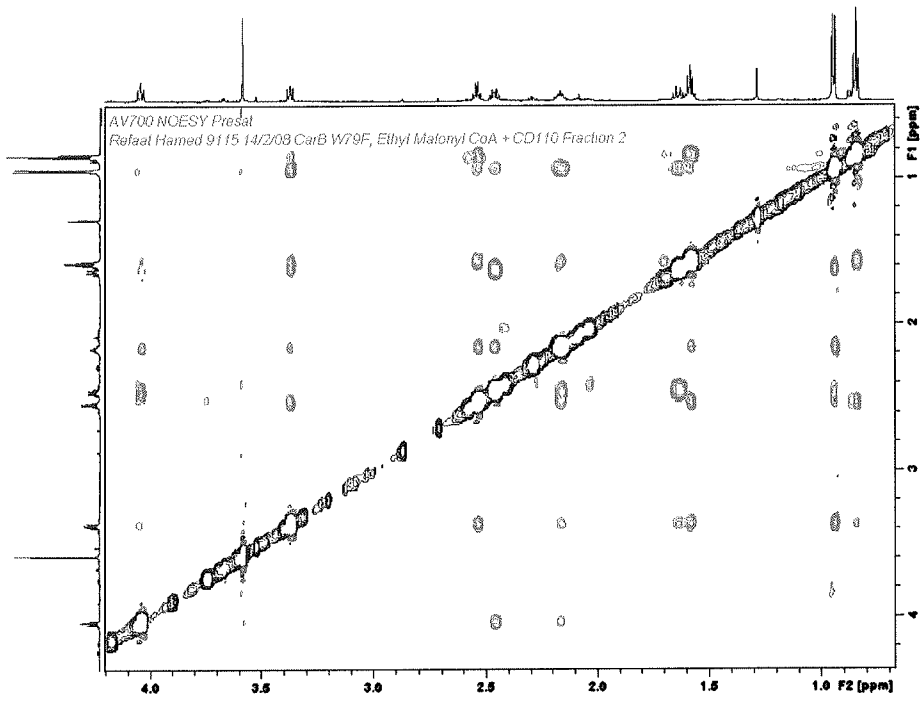



Figure 26: 2D NOESY spectrum for compounds **16**.

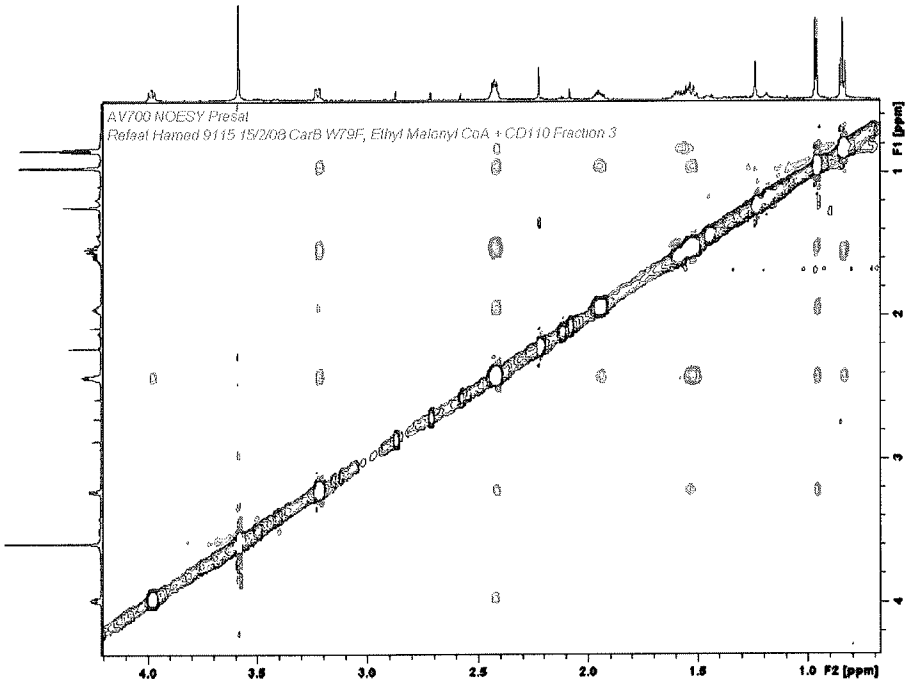



Figure 27: 2D NOESY spectrum for compounds **17**.

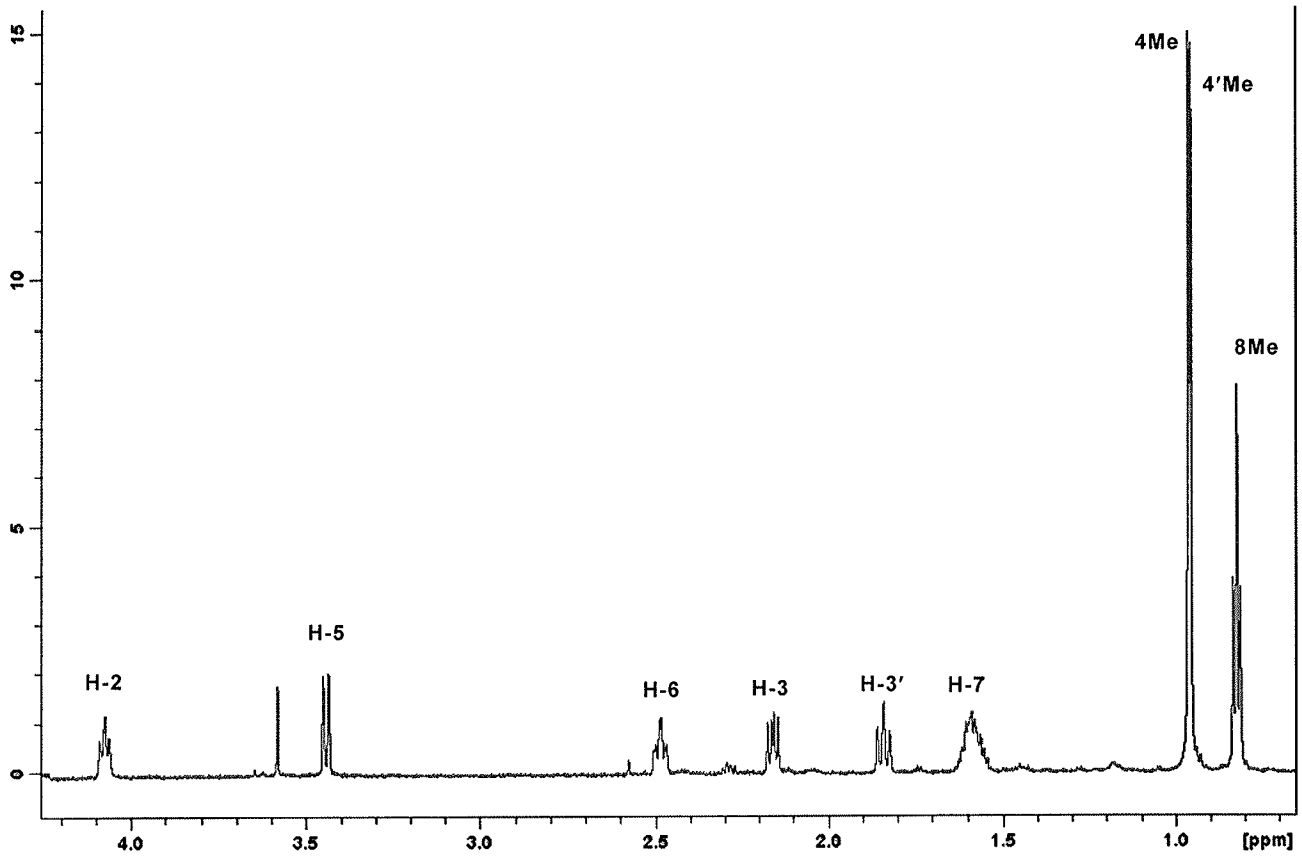



Figure 28: $^{1}$H-NMR spectrum for compound **18**.

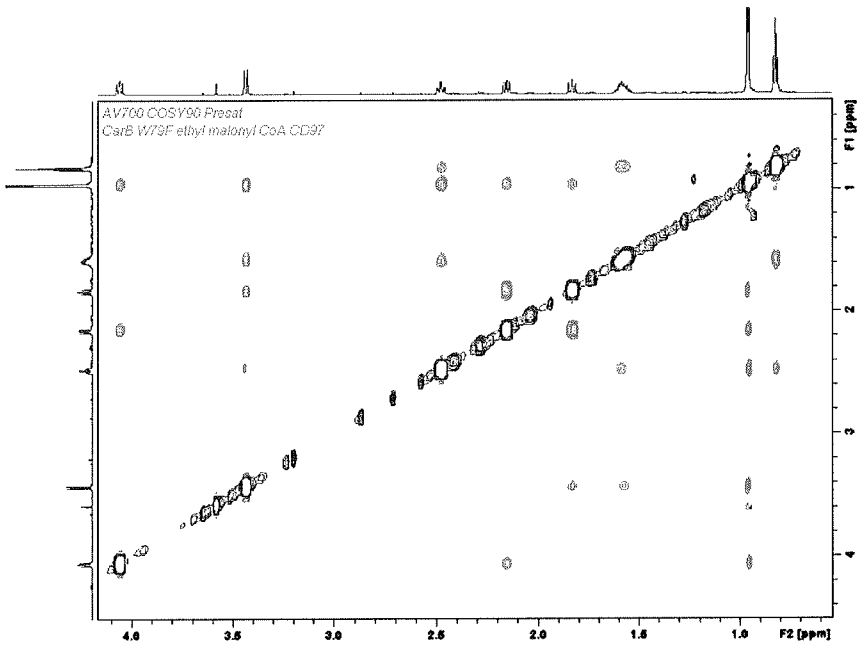



Figure 29: 2D NOESY spectrum for compound **18.**

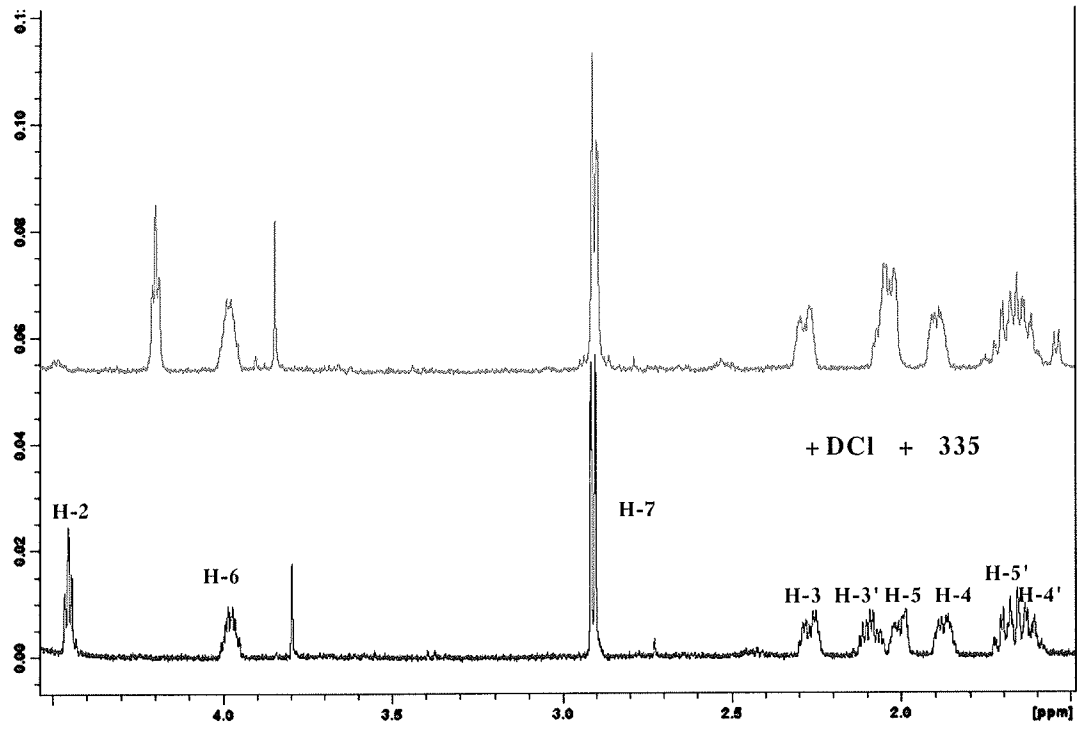



Figure 30: $^{1}$H-NMR spectra of compound **19** (top) and after addition of deuterium chloride and increasing the temperature to 335 K (bottom).

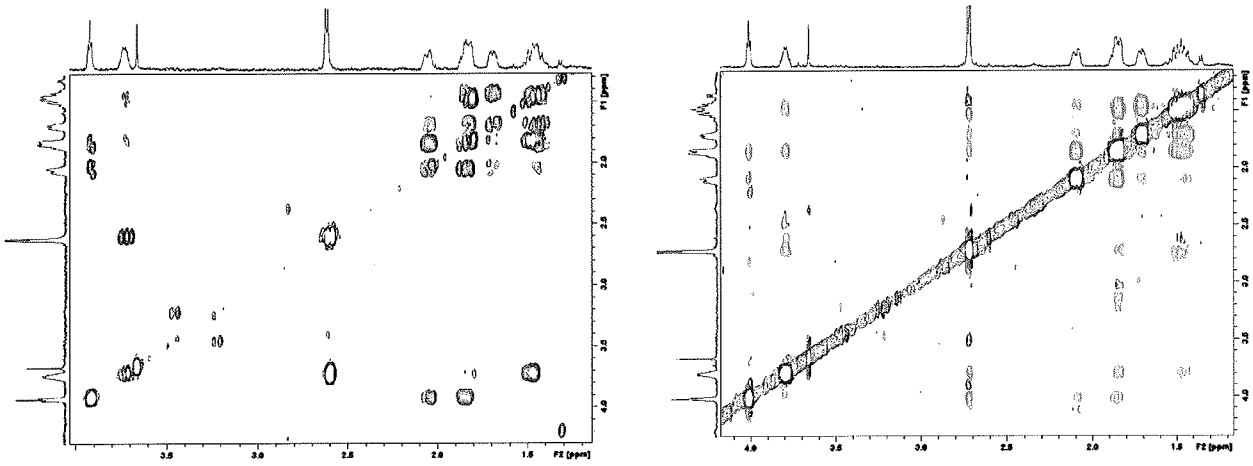


Figure 31: COSY (left) and 2D NOESY (right) spectra for compound **19**.

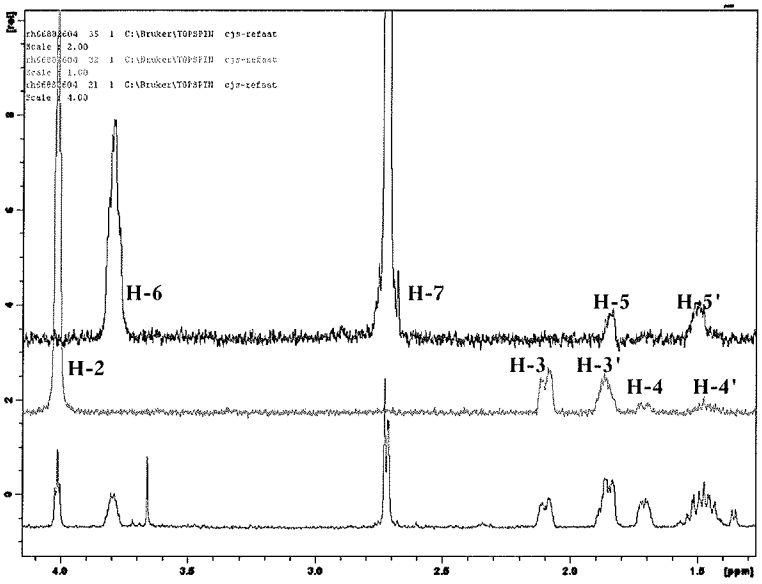



Figure 32: 1D TOCSY spectra for compound **19**. Note the Magnetization transfer between protons after excitation of H-2 (middle) and H-7 (top).

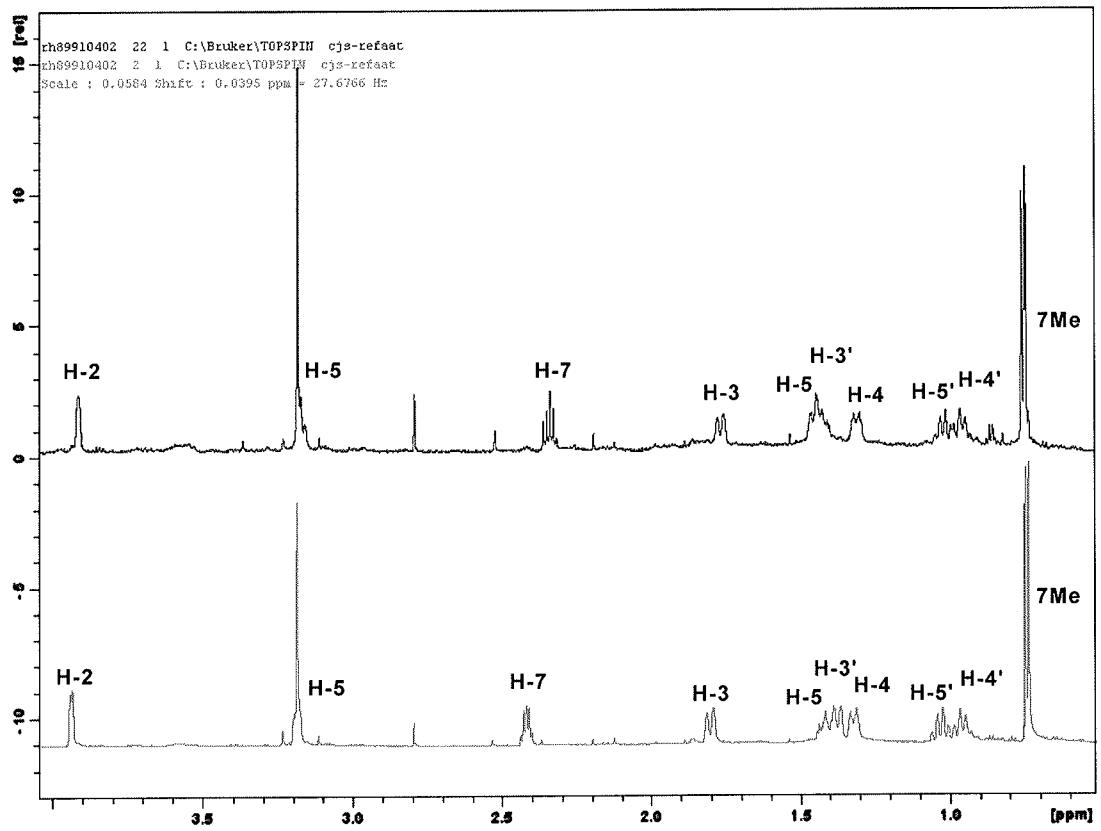



Figure 33: $^{1}$H-NMR spectra for compound **20** (bottom) and **21** (top).

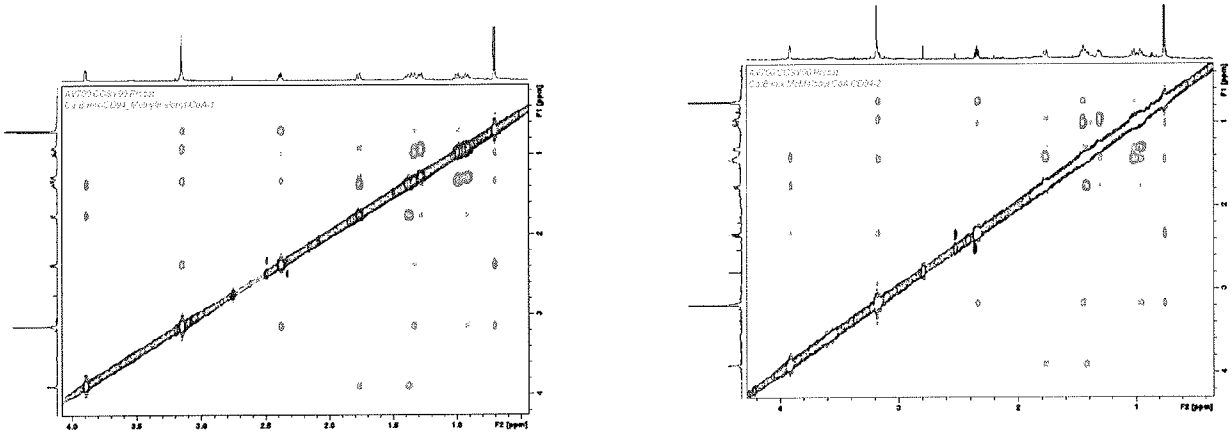


Figure 34: 2D NOESY spectra for compound **20** (left) and **21** (right).

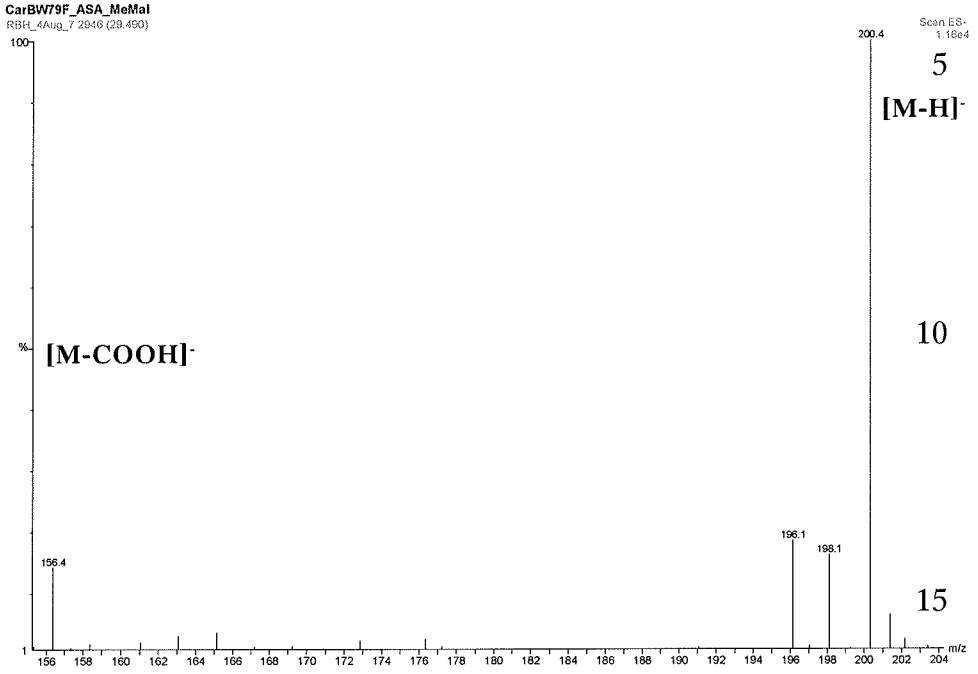



Figure 35: MS spectrum of compound **20**.

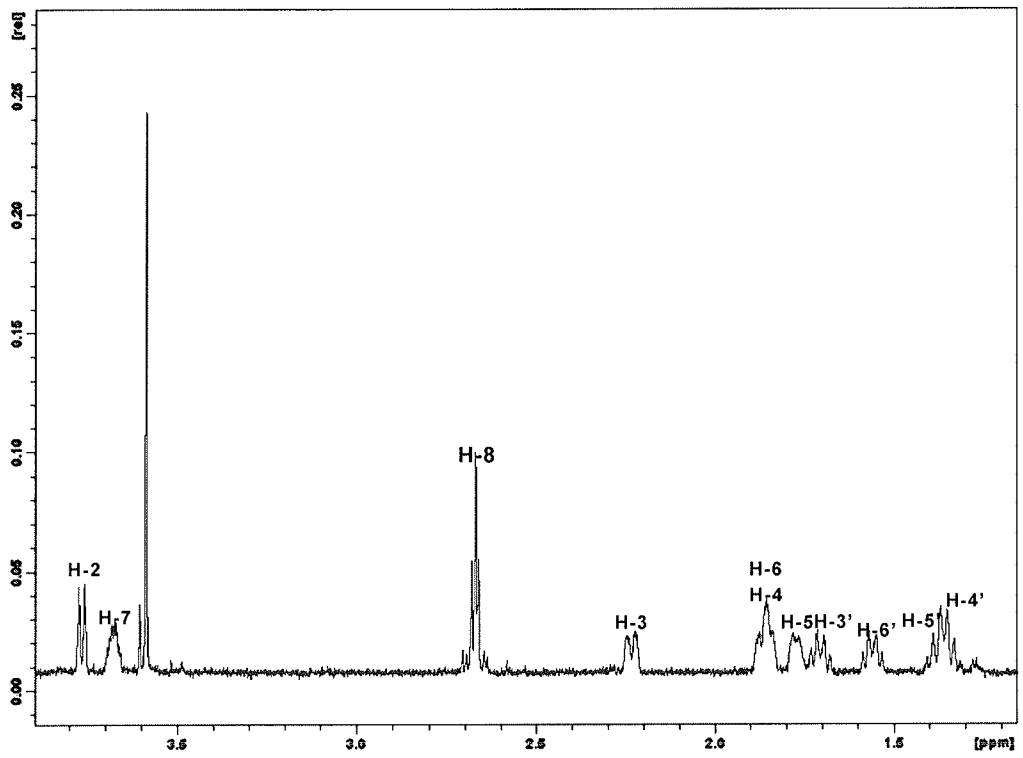



Figure 36: $^{1}$H-NMR spectrum for compound **22.**

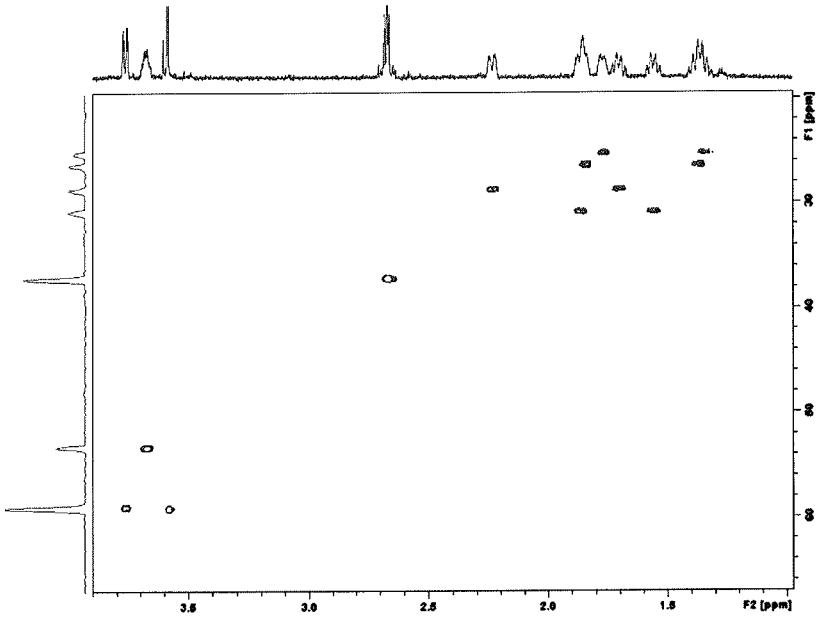



Figure 37: HSQC spectrum for compound **22.**

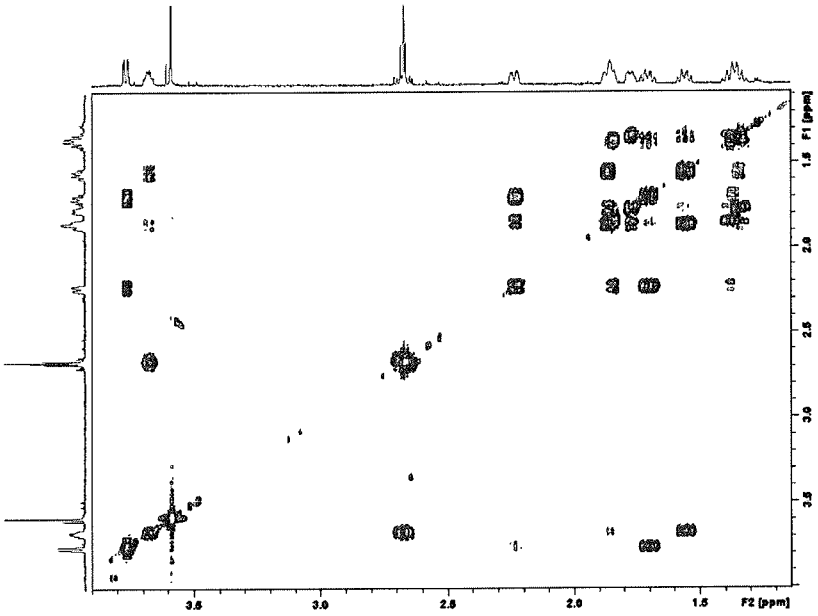



Figure 38: 2D COSY spectrum for compound **22.**

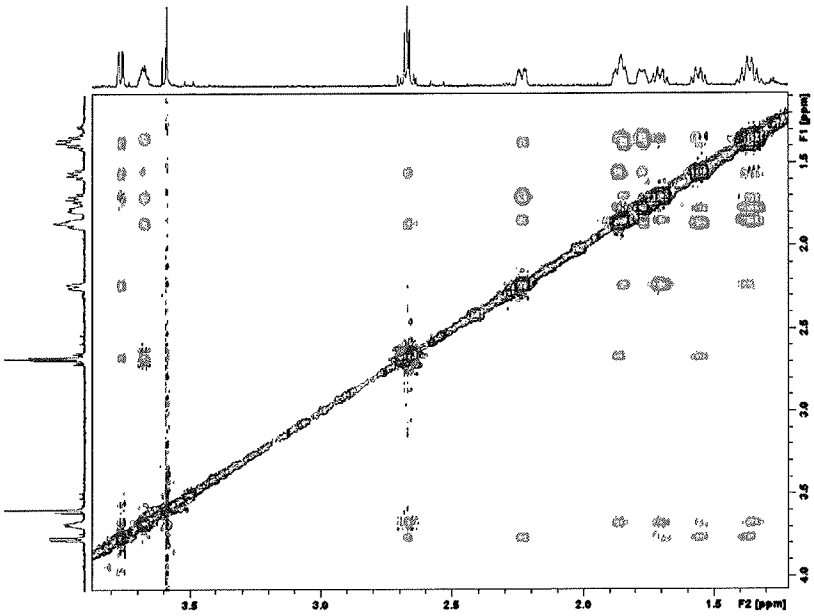



Figure 39: 2D NOESY spectrum for compound **22.**

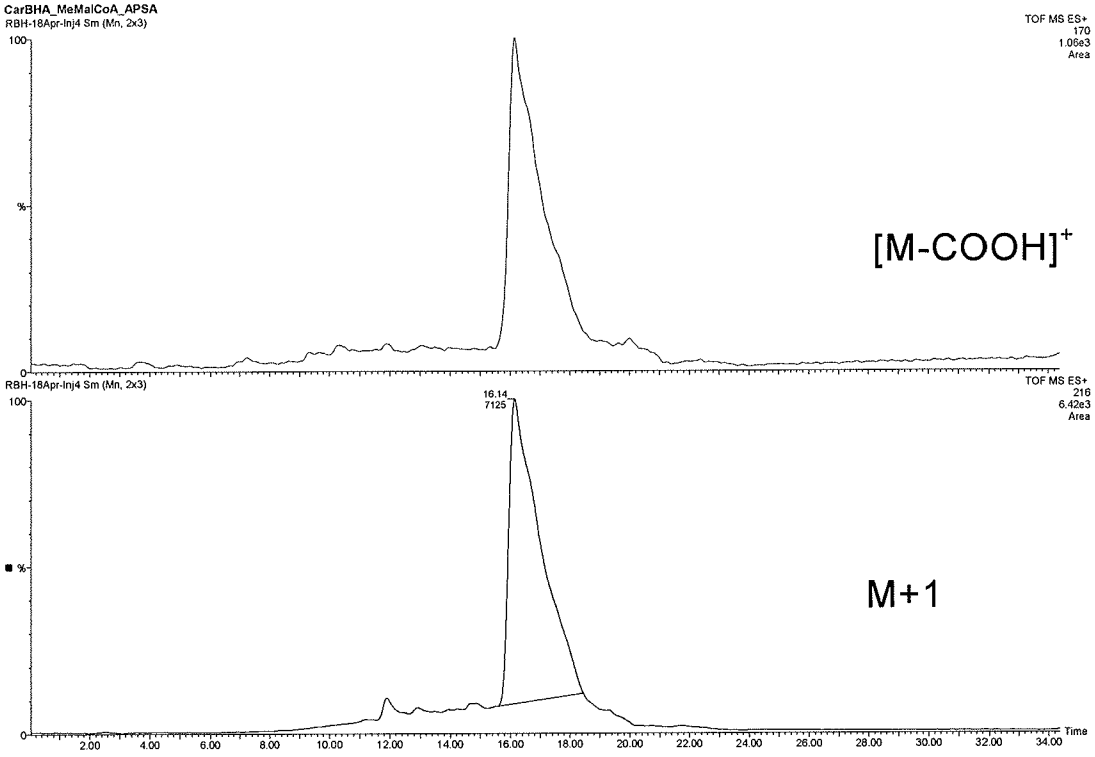



Figure 40: LC/MS spectrum of compound **23**.

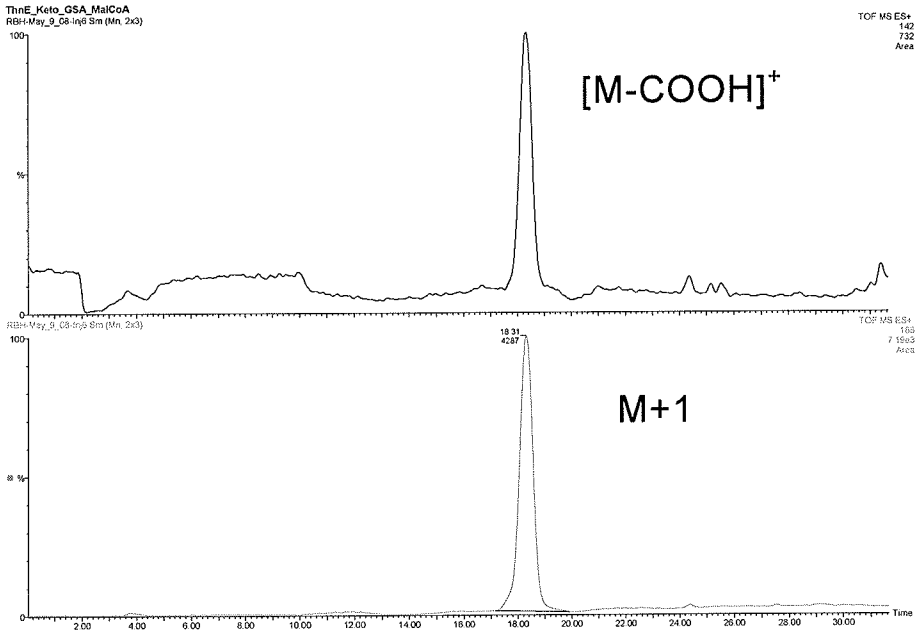



Figure 41: LC/MS spectrum of compound **24**.

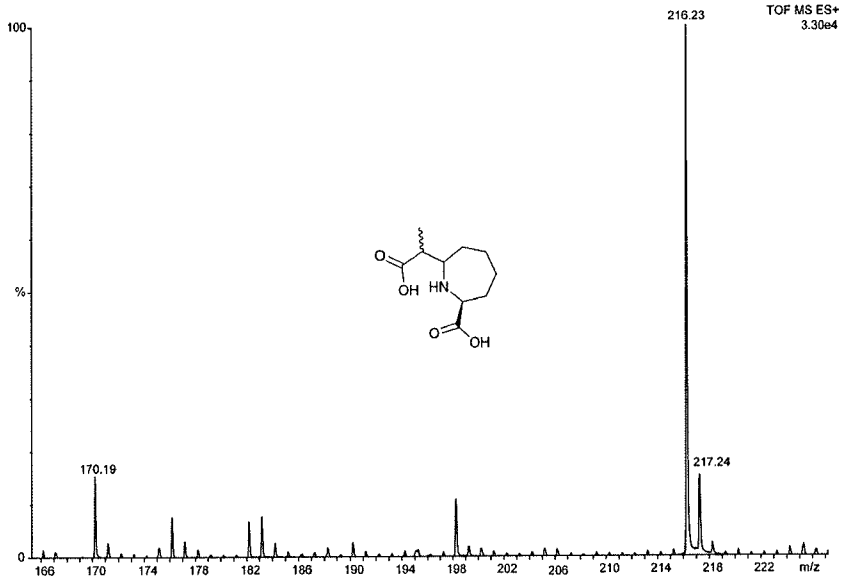



Figure 42: MS spectrum (positive ion electrospray ionization) for compound **23**.

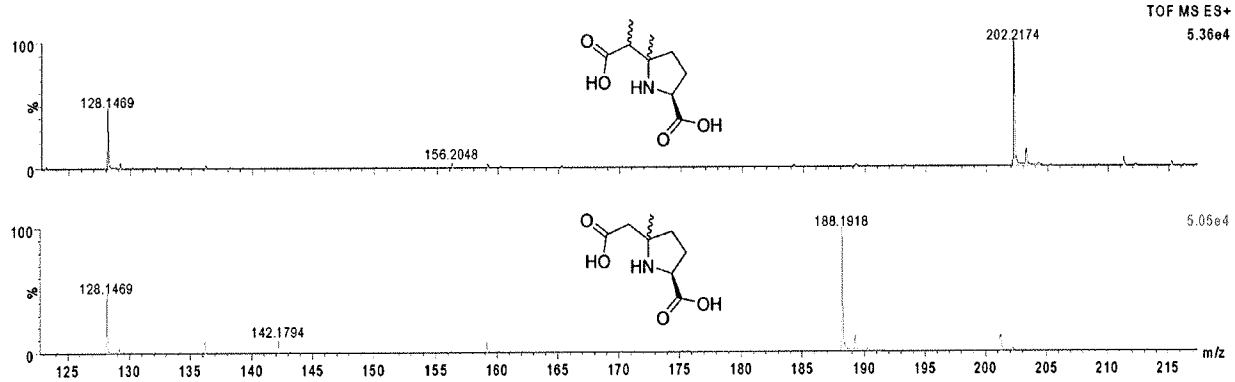



**Figure 43**: MS spectrum (positive ion electrospray ionization) for compound **24** (bottom) and compound **25** (top).

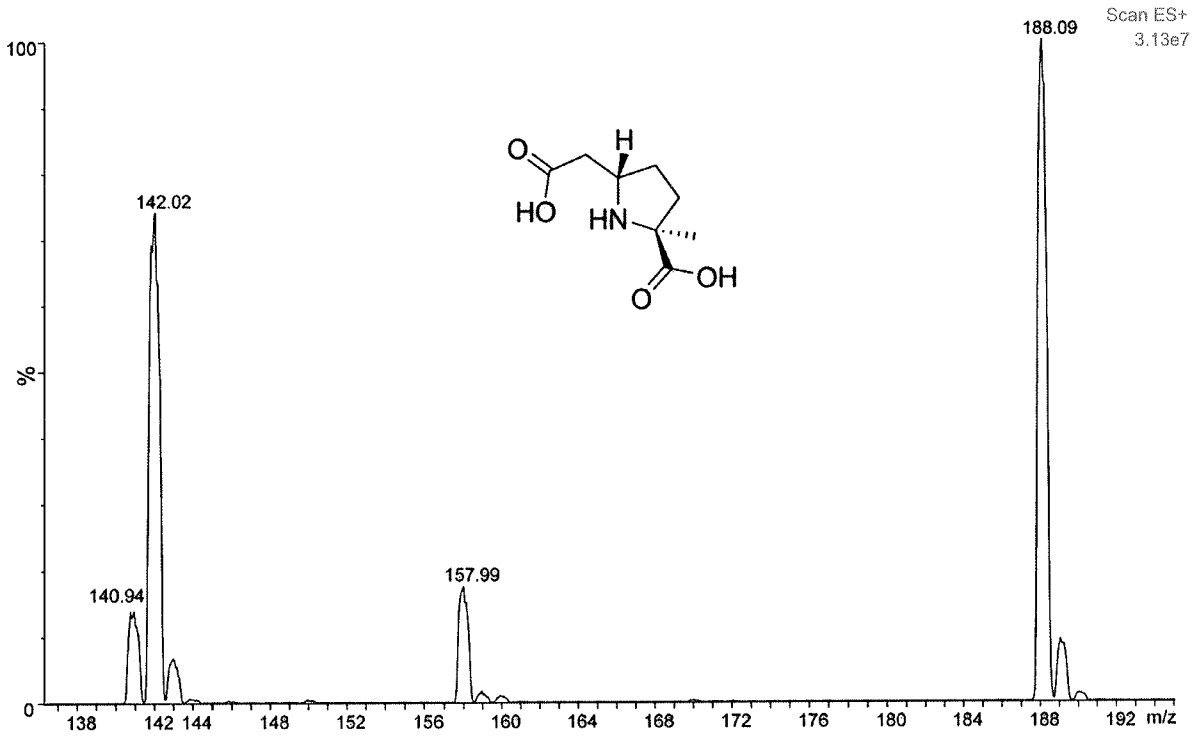



**Figure 44**: MS spectrum for compound **26**.

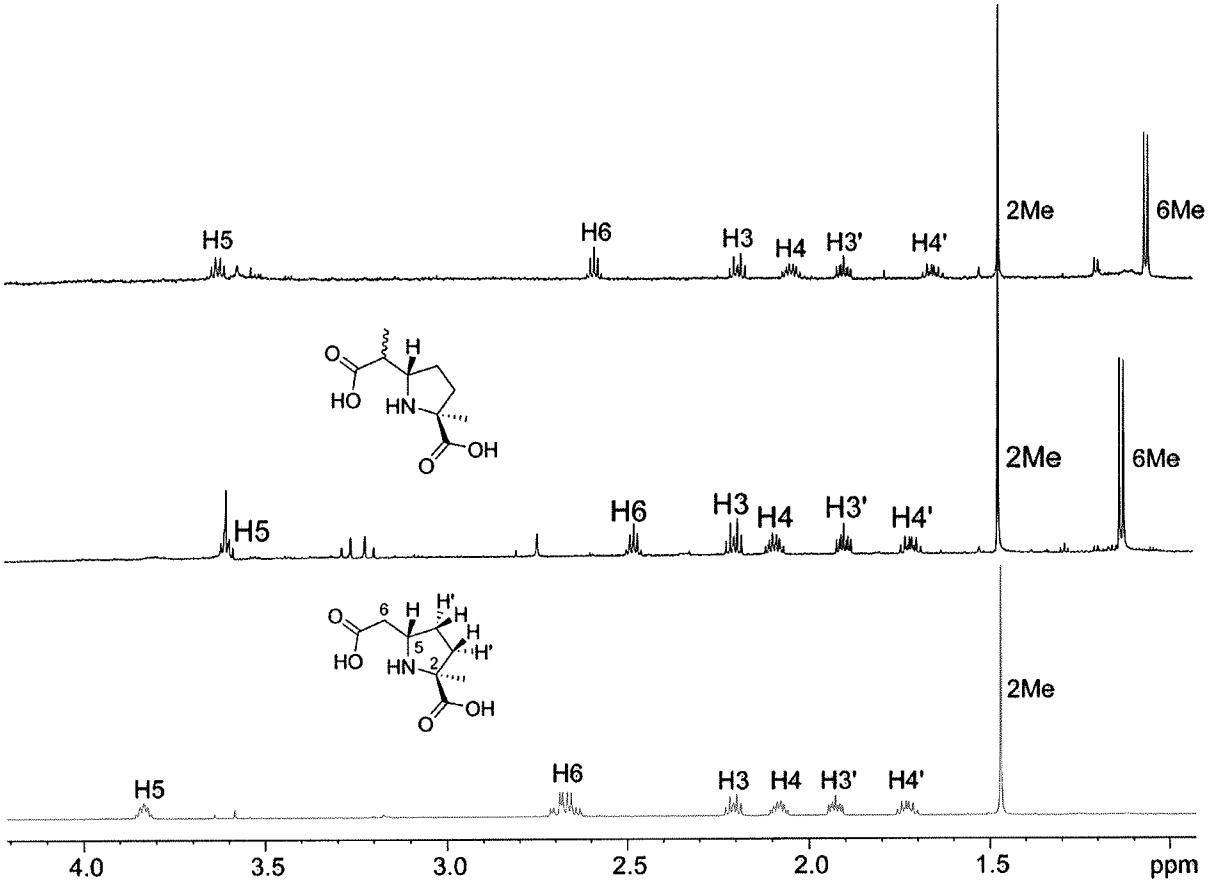



**Figure 45**: $^{1}$H-NMR spectra for compounds **26** (bottom) **27** (middle) and **28** (top).

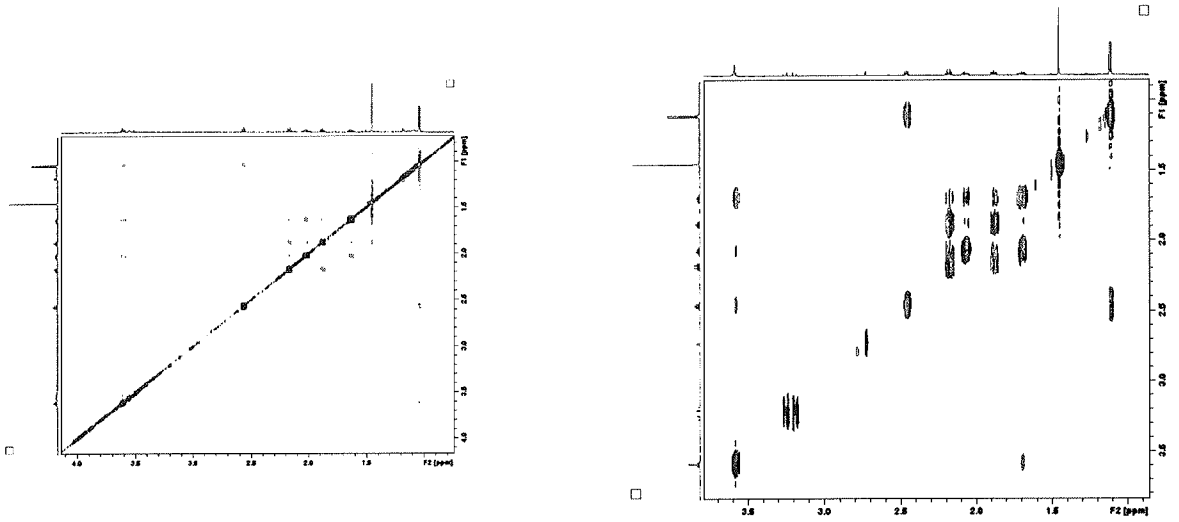


**Figure 46:** COSY and NOESY spectrum for compound **26**

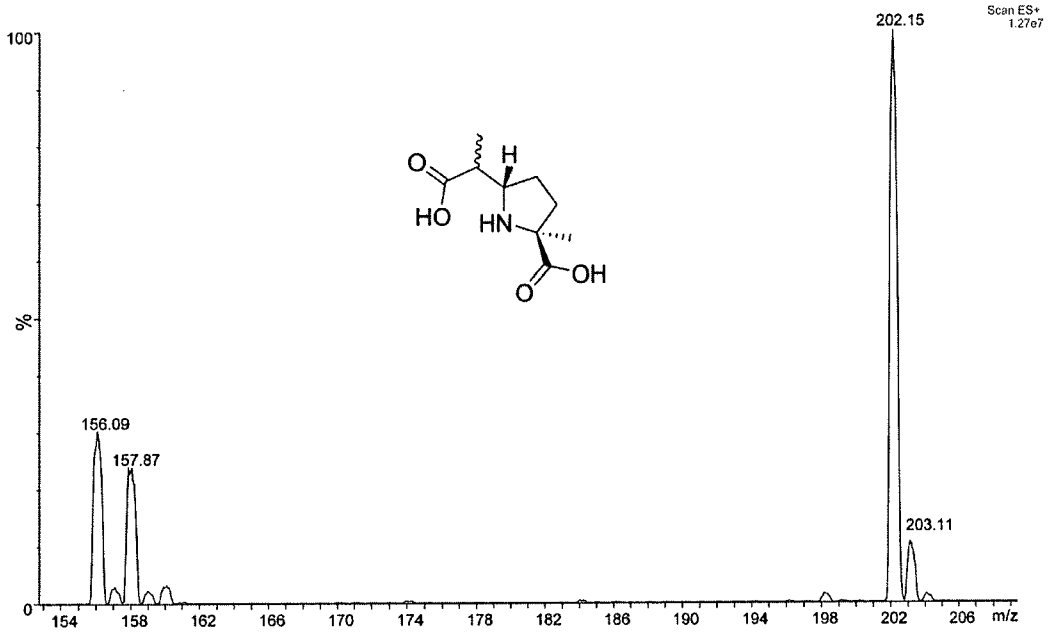



**Figure 47:** MS spectrum for compound **27**

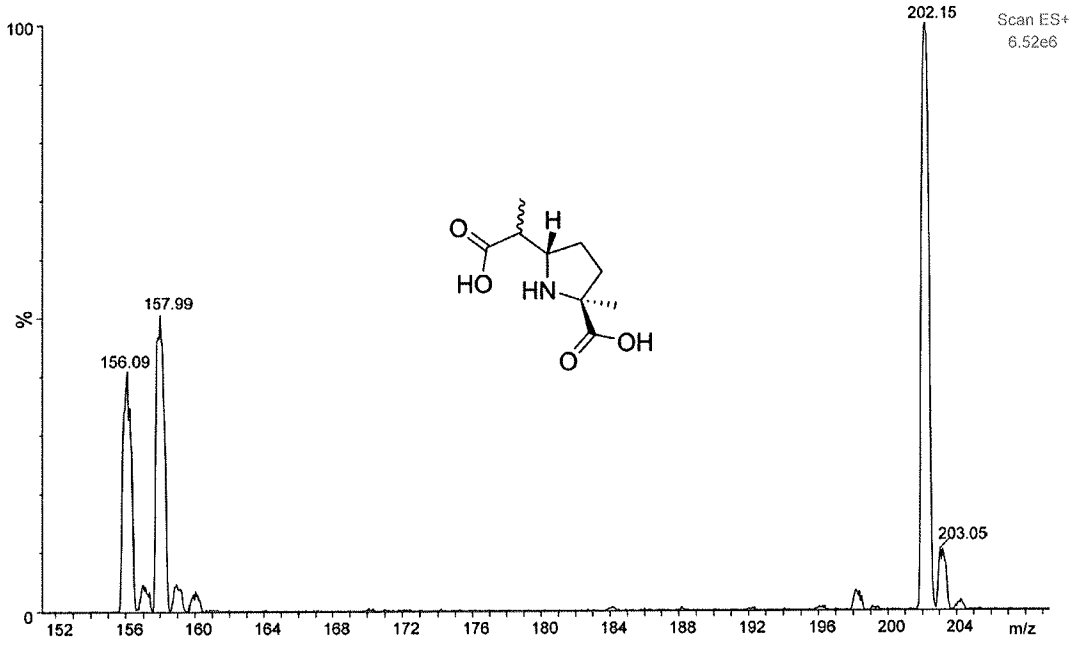



**Figure 48:** MS spectrum for compound **28**.

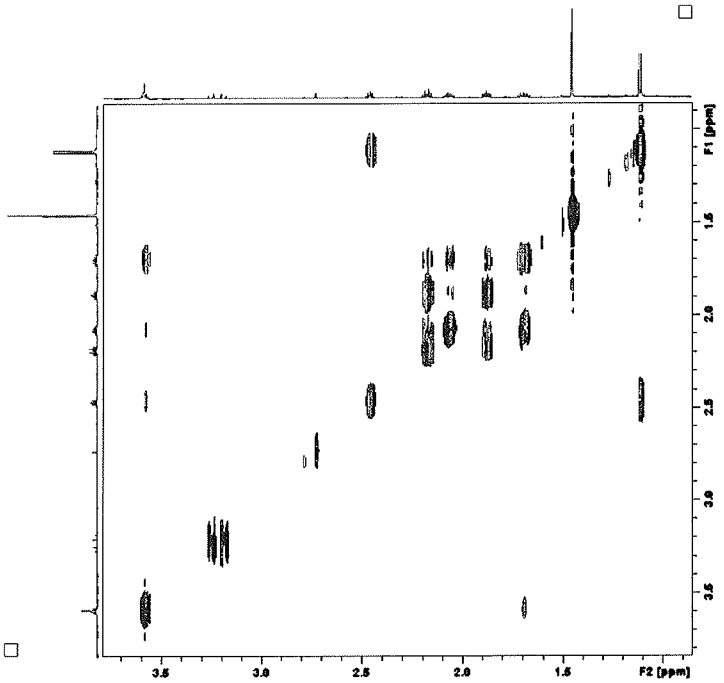



**Figure 49:** COSY spectrum for compound **27**.

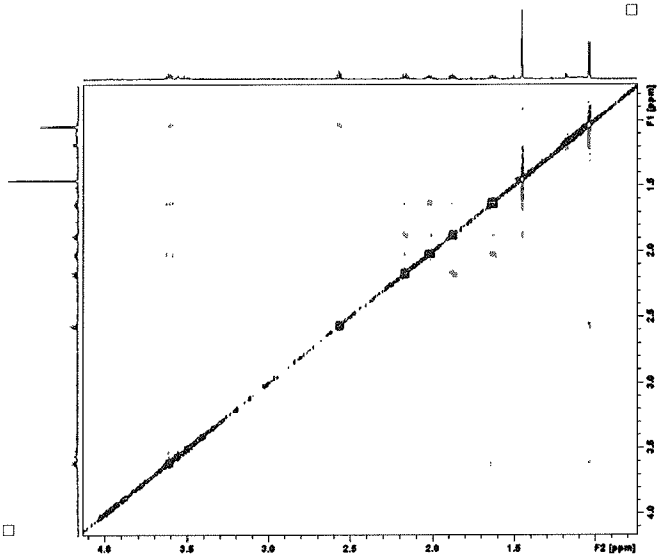



**Figure 50:** NOESY spectrum for compound **28.**

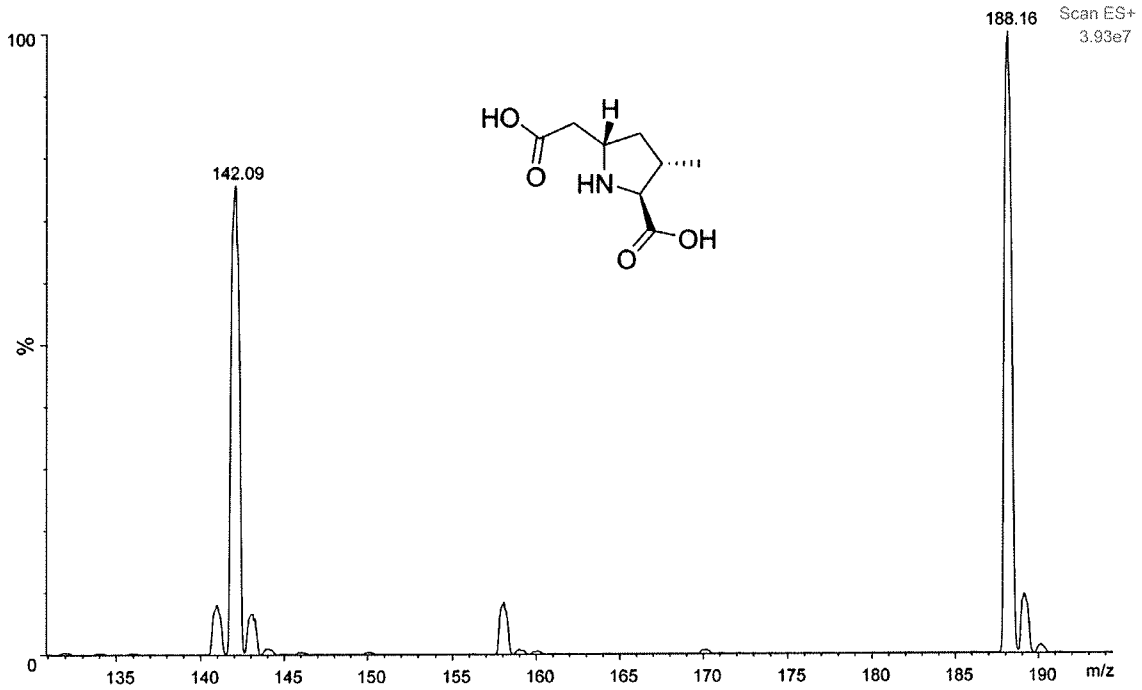



**Figure 51:** MS of compound **29**.

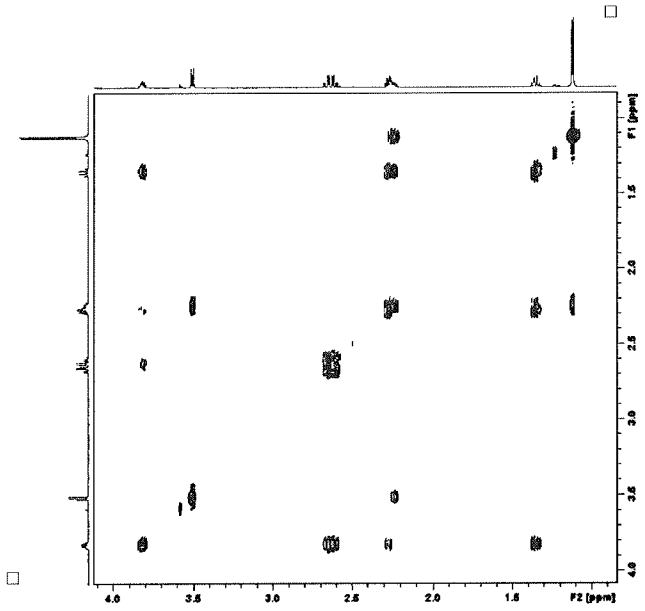



**Figure 52:** COSY spectrum of Compound **29**.

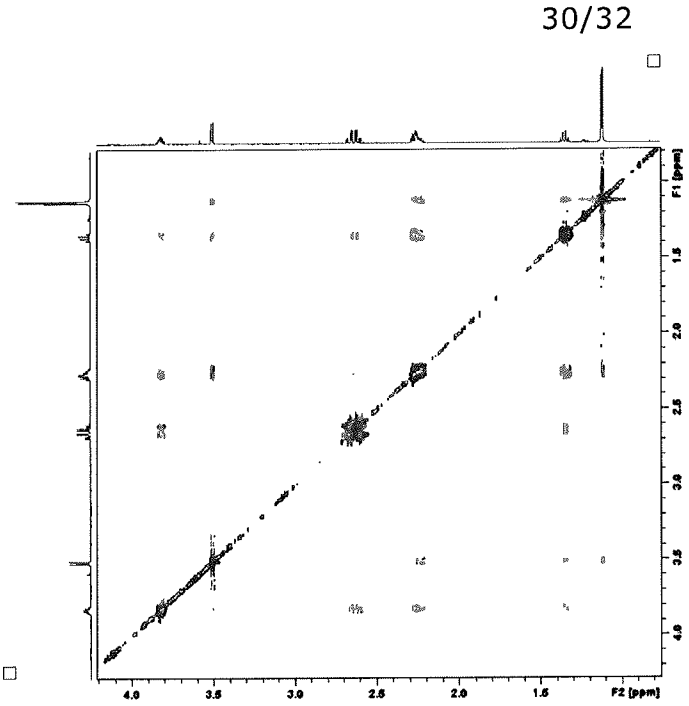



**Figure 53**: NOESY spectrum of Compound **29.**

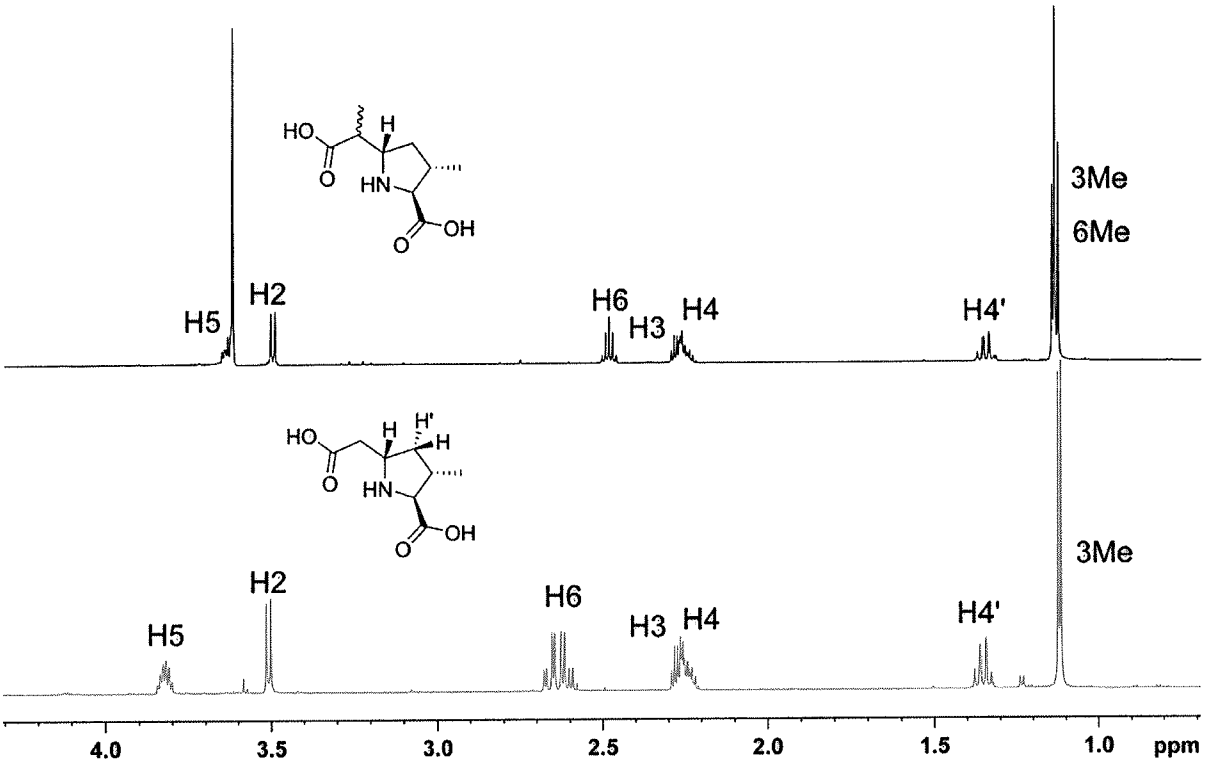



**Figure 54:** $^{1}$H-NMR spectra for compound **29** (bottom) and **30** (top).

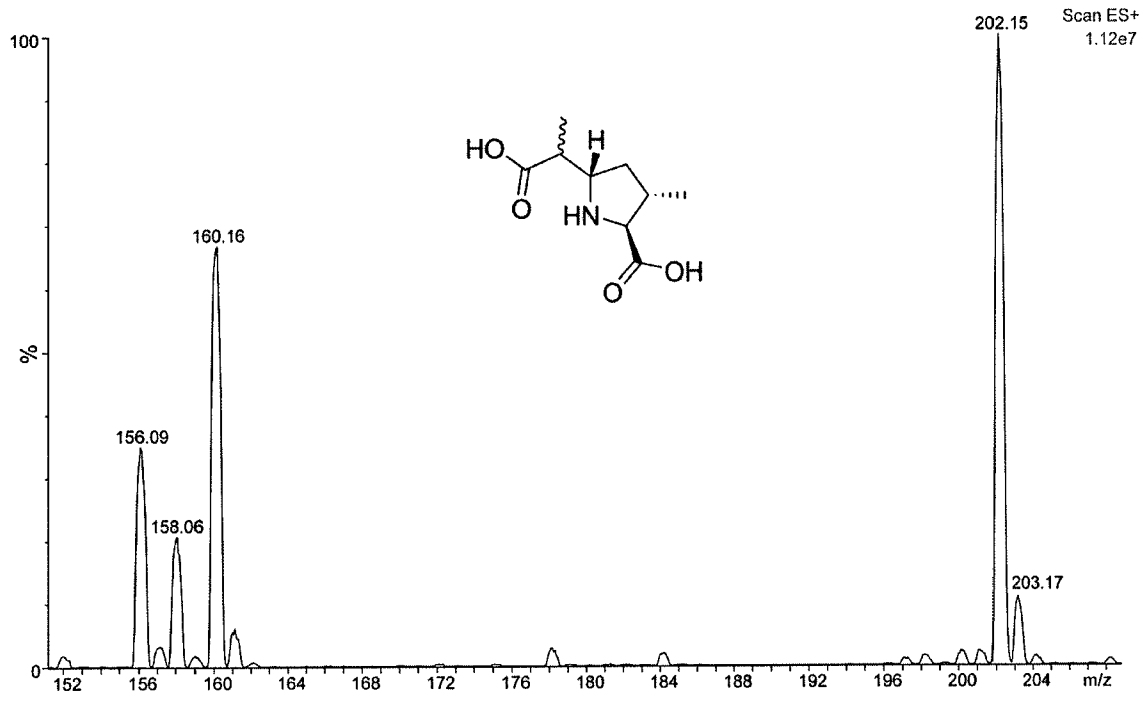



**Figure 55**: MS of compound **30**.

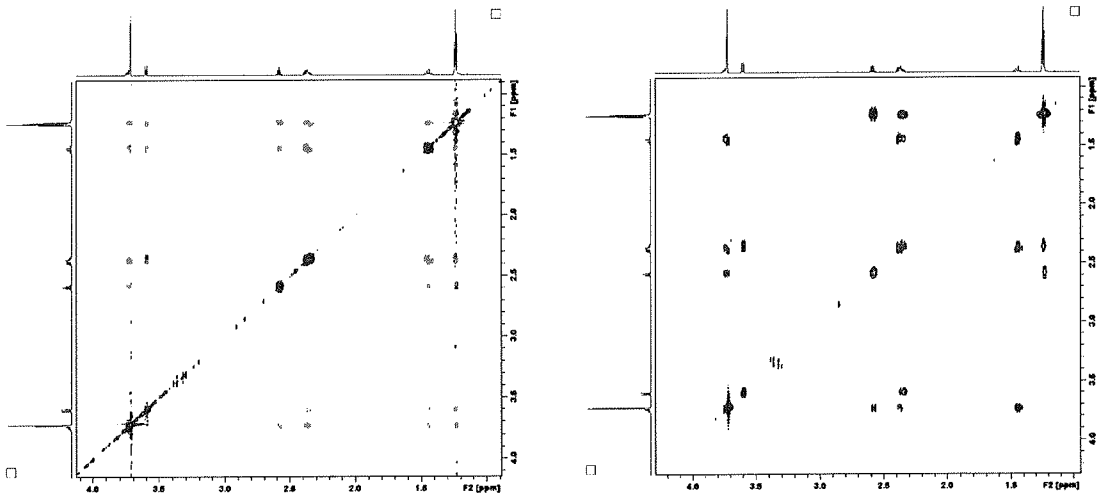


**Figure 56:** COSY and NOESY spectrum of Compound **30.**

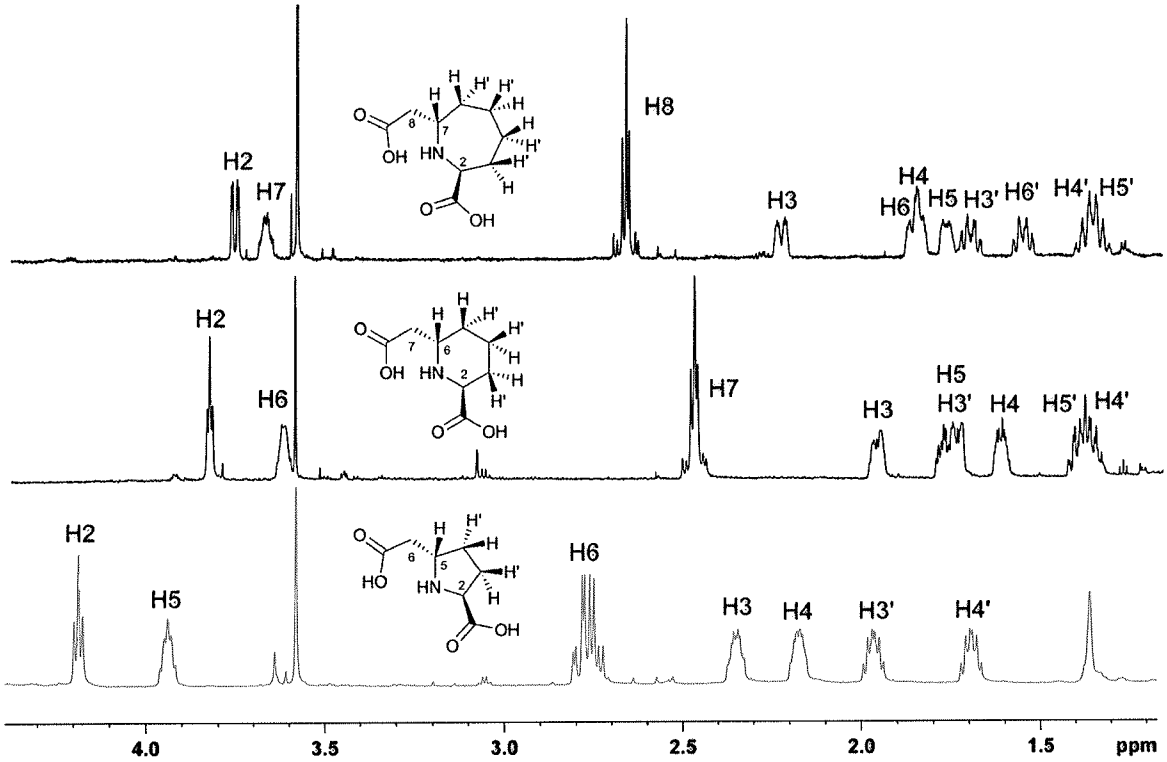



**Figure 57:** $^{1}$H-NMR spectra of purified carboxymethyl-substituted *N*-heterocycles (5, 6 and 7 membered rings) produced by CarB (bottom and middle) and CarBH229A (top) biocatalysis.