UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

DIPARTIMENTO DI FARMACIA

DOTTORATO DI RICERCA IN

"SCIENZA DEL FARMACO"

XXV CICLO 2010/2013

Natural ligands of nuclear receptors. Isolation, design, synthesis, biochemical decodification and potential therapeutic applications.

Dr. Raffaella Ummarino

Tutor

Prof.ssa A. Zampella

Coordinatore

Prof.ssa M.V. D'Auria

"I think I can affirm that in scientific research
neither intelligence grade nor the capacity to do
and to bring to the end the assignment undertaken
are essential factors for success and personal satisfaction.
In both cases total dedication and to close your eyes
in front of difficulties mostly count: in this way we can face
problems that others, more incisive and sharp,
would not face."

"Credo di poter affermare che nella ricerca scientifica
né il grado di intelligenza né la capacità di eseguire
e portare a termine il compito intrapreso siano fattori essenziali
per la riuscita e per la soddisfazione personale.
Nell'uno e nell'altro contano maggiormente la totale dedizione
e il chiudere gli occhi davanti alle difficoltà:
in tal modo possiamo affrontare i problemi che altri,
più critici e più acuti, non affronterebbero."

Rita Levi Montalcini

INDEX

ABSTRACT

Natural products have historically been a rich source of lead compounds in drug discovery. The biochemical investigation of marine organisms, through the deep collaboration between chemists and pharmacologists, focused on searching of new biologically active compounds, is a central issue of this kind of studies.

My research work, described in this PhD thesis, has been developed in this research area and was addressed to the identification of new ligands of nuclear receptors, discovering potent and selective modulators of farnesoid-X-receptor (FXR) and pregnane-X-receptor (PXR), regulators of various processes including reproduction, development, and metabolism of xeno- and endobiotics.

First, analysis of the polar extract of the sponge *Theonella swinhoei* afforded two new sulfated sterols, solomonsterols (SA and SB), the first example of marine PXR agonists. Both have been synthesized and characterized in animal models of inflammation. Administration of synthetic solomonsterol A effectively protects against development of clinical signs and symptoms of colitis; therefore SA holds promise in the treatment of inflammatory bowel deseases (IBDs).

To overcome a limitation of SA in clinical settings, a small library of SA derivatives has been designed and prepared. Indeed, SA could be absorbed from the GIT causing severe systemic side effects resulting from the activation of PXR in the liver. This study disclosed cholestan disulfate (Coldisolf) as a new, simplified agonist of PXR, currently in pharmacological evaluation on animal models of liver fibrosis induced by HIV infection.

Simultaneously, a wide family of 4-methylene steroids were isolated from the apolar extracts of *Theonella swinhoei*. These marine steroids are endowed with a

potent agonistic activity on PXR while antagonize the effects of natural ligands for FXR.

Among this rich family, we have identified theonellasterol as the first example of a sponge derived highly selective FXR antagonist demonstrating its pharmacological potential in the treatment of cholestasis. Using this compound as a novel FXR antagonist hit, we have prepared a series of semi-synthetic derivatives in order to gain insights into the structural requirements for exhibiting antagonistic activity. These molecules could be used for the pharmacological treatment of cholestasis but also in chemotherapy of carcinoma characterized by over-expression of FXR.

In summary, Nature continues to be one of the best sources not only of potential chemotherapeutic agents but also of lead compounds that could represent an inspiration for the discovery of new therapeutic strategies.

ABSTRACT (Italian)

Le sostanze naturali sono da sempre un'ispirazione per la scoperta di nuove strategie terapeutiche. Lo studio chimico di organismi marini in combinazione con la valutazione della loro attività biologica costituisce il fulcro della Chimica delle Sostanze Naturali. In tale ambito, l'attività di ricerca condotta durante il corso di Dottorato, i cui risultati sono riportati nella seguente tesi, è stata focalizzata principalmente sull'identificazione di ligandi di recettori nucleari metabolici, individuando potenti e selettivi modulatori del recettore dei farnesoidi (FXR) e del recettore dei pregnani (PXR), regolatori di processi di detossificazione di metaboliti endogeni (acidi biliari) e/o esogeni.

In particolare, dall'estratto polare della spugna *Theonella swinhoei* sono stati isolati due nuovi steroli solfatati, i solomonsteroli (SA e SB), il primo esempio di agonisti di PXR a struttura steroidica dal mare. Per entrambe le molecole si è proceduto alla sintesi totale in larga scala e al conseguente approfondimento farmacologico in modelli animali di infiammazione. Il SA si è rivelato efficace nel prevenire i sintomi associati alla colite nonché nel migliorare i segni clinici e si propone quindi come nuovo lead per il trattamento delle IBDs (Inflammatory Bowel Diseases).

Dalla scoperta dei solomonsteroli, si è poi passati alla progettazione e sintesi di derivati ad azione colon-specifica cercando di superare i limiti del lead naturale ampiamente assorbito a livello intestinale e quindi potenzialmente tossico per effetto su PXR epatico. Questo lavoro ha portato all'identificazione di una nuova molecola il colestan disolfato (Coldisolf), di facile sintesi e al momento in fase di sperimentazione farmacologica sulla fibrosi epatica indotta da infezione da HIV.

Parallelamente dagli estratti apolari della spugna *Theonella swinhoei* è stata, invece, isolata un'ampia famiglia di 4-metilensteroli con un range di attività che spazia dall'agonismo su PXR all'antagonismo su FXR passando per la modulazione duale. Tra queste molecole, il theonellasterolo, rappresenta il primo esempio di antagonista selettivo di FXR di origine naturale e quindi promettente lead per il trattamento farmacologico della colestasi.

Usando questa molecola come nuovo hit, si è proceduto alla progettazione e sintesi di una nutrita serie di derivati, che sottoposti ad una robusta sperimentazione farmacologica in vitro, hanno contribuito a delineare la prima SAR su questo nuovo chemotipo di antagonista e soprattutto a tracciare le linee guida per l'ottenimento di molecole a potenziale uso per il trattamento della colestasi e la chemioterapia di carcinomi caratterizzati da over-espressione di FXR.

In conclusione, la Natura è, e continua ad essere, la maggiore fonte di ispirazione di nuovi *lead* da utilizzare per la progettazione di nuovi farmaci.

Dunque, la chimica delle sostanze naturali offre ancora entusiasmanti prospettive.

INTRODUCTION

Today about 40% of modern pharmaceuticals are derived from biological sources.[1,2] This simply observation can give an idea of the incredible biomedical potential represented by the chemical analysis of the biodiversity of natural organisms.[3,4] Secondary metabolites contained in these organisms are the result of millions of years of evolution and natural selection: even a single species constitutes a library of metabolites that is validated for the bioactivity. As the results of enzymatic reactions, natural products have an intrinsic capacity to recognize and bind macromolecules, perturb their activity, and modulate biological processes. Besides their potential use as pharmaceutical drugs, natural products have and will continue to play critical roles as biological probes, to wield temporal control over biochemical pathways, and ultimately, to identify novel therapeutic targets.[5] Surely among Nature, plants represent a rich source of novel compounds to be used as lead to design new drugs and, notably, several drugs, from aspirin to morphine, currently in use for human diseases, have this origin. Particularly rich is also the marine environment. Ocean cover seventy percent of the surface of the planet and represents a wealthy source of plants, animals and micro-organisms which, due to their adaptation to this unique habitat, produces a wide variety of secondary metabolites unlike those found in terrestrial species.[6] Today, with the modern tools of molecular biology and advanced technology, the potential of marine environment, with its vast reservoir of original molecules, represents a great promise to provide new drugs. Beside in the past century the high-throughput screening of natural sources has long been recognized as an invaluable source of new lead structures, today targeted oriented discovery, focused on the

identification of natural products as ligands of specific proteins or enzymes, is considered the best rationale approach for the identification of novel therapeutic agents from Nature. Indeed natural products are being biosynthesized by their hosts to interact with proteins, such as enzymes or receptors, and many human protein targets contain structural domains similar to the targets with which small ligands (or natural products) have coevolved.

Nuclear receptors (NRs) represent one of the most important drug targets in terms of potential therapeutic application,[7] playing a role in every aspect of development, physiology and disease in humans. They are ubiquitous in the animal kingdom suggesting that they may have played an important role in their evolution. NRs have a rich and long-standing history in drug discovery for two fundamental reasons. First of all, they have been designed by nature to selectively bind small lipofilic molecules, and then they are able to regulate a diverse set of biologically important functions. NRs share considerable amino acid sequence similarity in two highly conserved domains, the N-terminal DNA-binding domain (DBD) and the C-terminal ligand-binding domain (LBD), responsible for binding specific DNA sequences and small lipophilic ligands, respectively (Figure 1). Upon ligand binding, NR induces conformational changes that lead to the release of the co-repressors and recruitment of a co-activators, thus providing a chromatin remodeling and subsequent activation of transcriptional machinery.

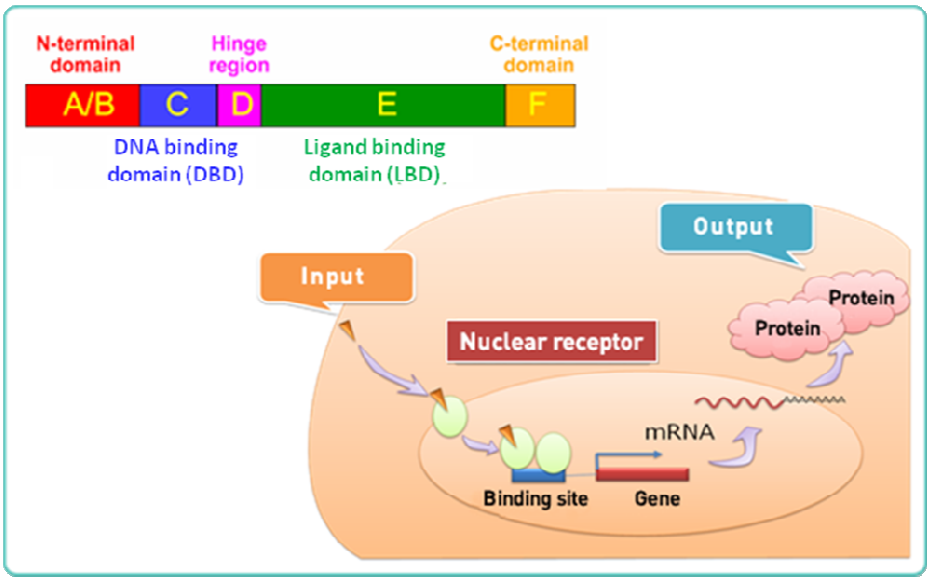


Figure 1. General structure of nuclear receptor

There are 48 genes in the human genome coding for the NRs superfamily. Most of them has been discovered in the last twenty years and several still require a de-orphanization and a complete and detailed clarification of their physiological role. Nevertheless in the last two decades an huge experimentation has been focus on the discovery of selective NRs modulators. There are three subfamilies of nuclear receptors: NR1, NR2, and NR3. NR3 subfamily, also known as classical homodimer steroidal receptors, includes estrogen receptors α and β (ERα and ERβ), glucocorticoid receptor (GR), progesterone receptor (PR), androgen receptor (AR), and mineralocorticoid receptor (MR). Nuclear receptors of class 1 and 2, unlike steroidal receptors, function as heterodimers with the retinoid X receptor (RXR) (Figure 2). Importantly, these receptors, including the peroxisome proliferator-activated receptor (PPAR), liver X receptor (LXR), farnesoid X receptor (FXR), vitamin D3 receptor (VDR), retinoic acid receptor (RAR) and thyroid hormone receptor (TR), serve as endogenous sensors for fatty acids, oxysterols, thyroid hormones and bile acids. These classes also include pregnane X receptor (PXR) and constitutive androstane receptor (CAR) for which no

physiological ligands have been so far identified. PXR and CAR are defined the xenobiotic NRs, master regulators of Phase I and Phase II enzymes and drug transporters.[8]

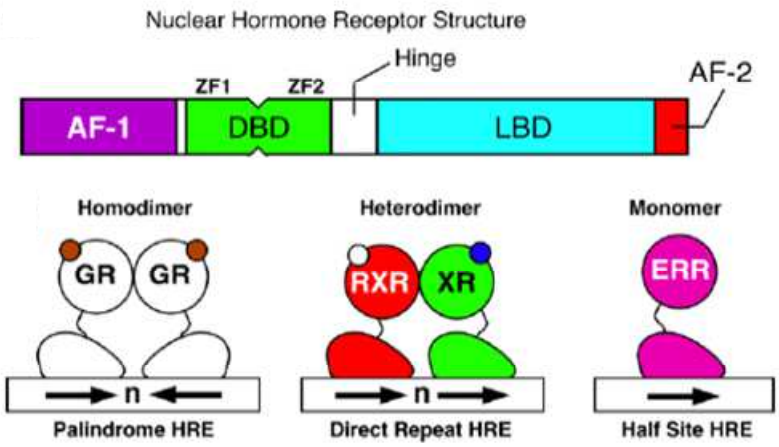


Figure 2. Formation of an heterodimer with the retinoid X receptor (RXR)

PXR is a master gene orchestrating the expression of a wide family of genes involved in uptake, metabolism and disposal of a number of endo- and xeno-biotics, including drugs, bile acids, steroid hormones, environmental toxicants and metabolic intermediates in mammalian cells.[9] It is almost exclusively expressed in the gastrointestinal tract and liver, with lower levels in the kidney and ovary. Following ligand binding, PXR forms an heterodimer with RXR that binds to specific PXR response elements (PXREs), located in the 5′-flanking region of PXR target genes, resulting in their transcriptional activation. Among these, P450 enzymes (CYP3A, CYP2C, and CYP2B) that promote oxidative (phase I) drug metabolism,[10,11] phase II-conjugating enzymes that improve solubility of phase I metabolites (glutathione *S*-transferases, sulfotransferases, and UDP-glucoronosyltransferases)[12,13] and xenobiotic transporters (MDR1, MRP2, MRP3, and OATP2) mediating excretion of the above compounds (Figure 3). In addition

to its involvement in detoxification and metabolism of xenobiotics, recent studies have indicated that this receptor plays a regulatory role in various physiological and pathophysiological processes, such as lipid metabolism,[14] glucose homeostasis, and inflammatory response.[15] To date, several evidences suggest that PXR may be an useful target for pharmacological therapies in various conditions, including liver disease,[16] and inflammatory bowel diseases (IBDs), encompassing Crohn's disease (CD), ulcerative colitis (UC) and liver fibrosis (LF).[17]

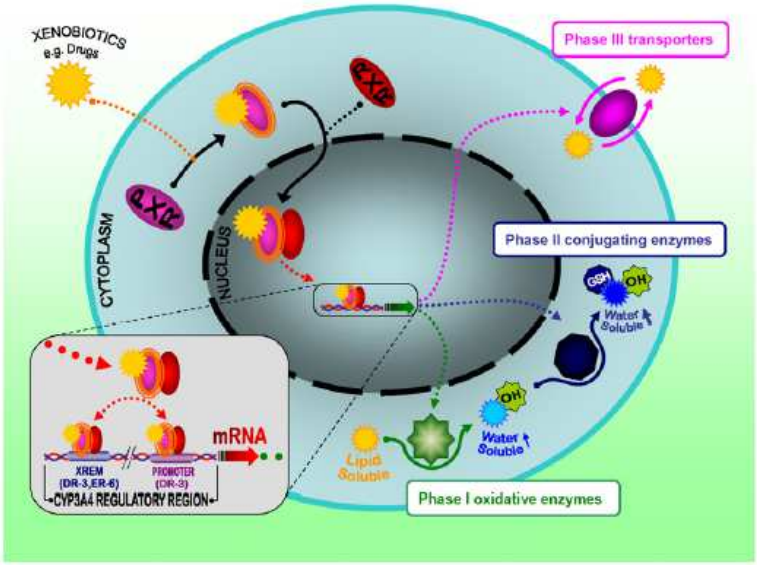


Figure 3. Functions of PXR on hepatic metabolism

Besides PXR shows the typical NRs organization, X-ray crystallography revealed an LBD larger than those of many other nuclear receptors, including the steroidal hormone receptor.[18] As a consequence, hPXR binds both small and large ligands and the number of chemicals that are reported to activate PXR has grown rapidly including many drugs currently in use such as statins, the antibiotic rifampicin and its semisynthetic derivative rifaximin, antihypertensive drugs nifedipine and spironolactone, anticancer compounds, HIV protease inhibitors, calcium channel

modulators as well as diverse environmental toxicant, plasticizers and pesticides, and agonists of additional nuclear receptors.[19]

Rifaximin, (Figure 4) a nonabsorbable structural analog of rifampicin used in the treatment of traveler's diarrhea, IBDs, and hepatic encephalopathy, is a gut-specific hPXR agonist. When fed to transgeneic mice expressing hPXR, rifaximin attenuates inflammation induced by dextran sulfate sodium (DSS) and trinitrobenzene sulfonic acid (TNBS), two classical models of IBDs. For a molecular point of view, amelioration of IBDs symptoms in hPXR mice by rifaximin has been linked to NF-kB and the negative cross-talk PXR-NFkB has recently demonstrated.[20]

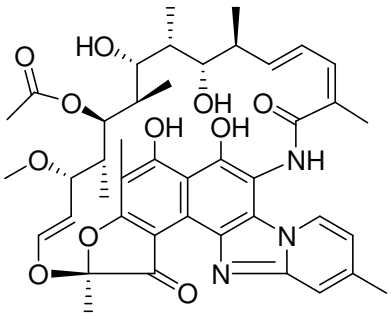


Figure 4. Rifaximin

Pregnenolone-16α-carbonitrile (PCN) (Figure 5) is a potent and specific agonist for murine PXR, with no activity for human PXR. It significantly decreased CYP7A1 expression, with competition between PXR and PGC-1α for binding to HNF4α, thereby blocking PGC-1α-stimulated activation of CYP7A1 by HNF4α.[21]

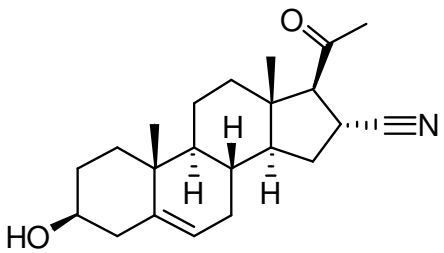


Figure 5. Pregnenolone-16α-carbonitrile (PCN)

As concern natural products, hyperforin (Figure 6), the psychoactive constituent of the widely used antidepressant herbal *H. perforatum*, commonly known as St. John's wort, was the first potent agonist of PXR reported from plants.[22,23] To date hyperforin is one of the most potent activators of human PXR with nanomolar EC_{50} (0.023 μM). Hyperforin competes with 3HSR12813 for binding to human PXR and stimulates the interaction between human PXR and the co-activator SRC-1. After the discovery of hyperforin, herbal medicines (*e.g.*, Ayurvedic medicine and traditional Chinese medicine) have attracted the interest of scientific community in order to identify the chemical constituents responsible for biological effects of their extracts and various chemicals have been characterized as ligands for PXR.[24]

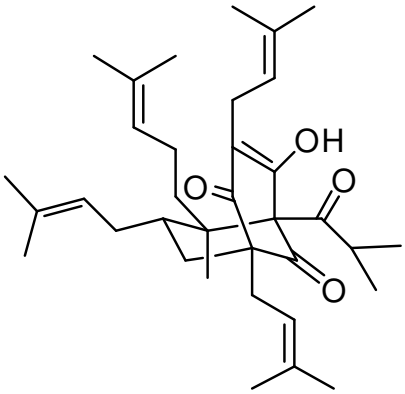


Figure 6. Hyperforin

Ginkgolide A (Figure 7) isolated from *G. biloba* has been identified as a PXR activator, increasing the expression of target genes in LS180 human colon adenocarcinoma cell (*CYP3A4*, *CYP3A5*,and *ABCB1*) and cultured human hepatocytes. Ginkgolide A contributed to the increase in hPXR target genes expression (CYP3A4 mRNA and CYP3A-mediated testosterone 6β-

hydroxylation), moreover in a cell-based reporter gene assay ginkgolide A treatment results in increment of SRC-1 recruitment on PXR.[25]

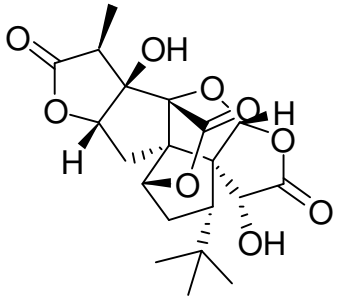


Figure 7. Ginkgolide A

It should be noted that all these compounds are PXR agonists whereas to date only few PXR antagonists have been described from vegetal sources.[26]

Coumestrol (Figure 8), a coumestan phytoestrogen present in soy sprouts and alfalfa endowed with estrogen-like structure and actions, has been reported as an antagonist of the human nuclear receptor PXR without effects on mouse PXR. In primary human hepatocytes, coumestrol suppresses the effects of PXR agonists on the expression of CYP3A4 and CYP2B6 as well as inhibits metabolism of tribromoethanol in humanized PXR mice and antagonizes the recruitment of SRC-1 on PXR.

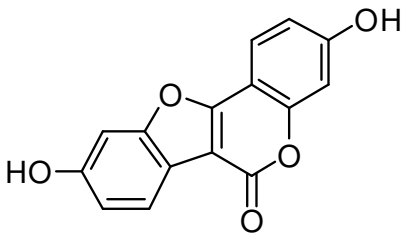


Figure 8. Coumestrol

The first marine ligand to be described was ecteinascidin 743 (ET-743) (Figure 9), isolated from the tunicate *Ecteinascidia turbinata*. Nanomolar concentrations of this potent marine-derived anticancer blocked activation of human PXR by either

SR12813, a synthetic agonist, or paclitaxel in cell-based reporter assays.[27] ET-743 also blocked the induction of the PXR target genes CYP3A4 and MDR1 in a human intestinal cell lines.

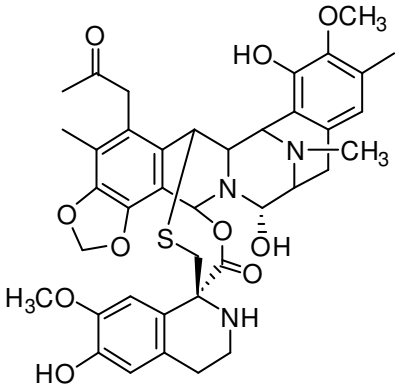


Figure 9. Ecteinascidin 743 (ET-743)

Among metabolic NRs, also FXR (Figure 10) has emerged as a valuable pharmacological target[28,29] in several human deseases for its regulatory function on bile acids (BAs), lipid and glucose homeostasis. Activation of FXR, highly expressed in the liver, intestine, kidney and adrenals, leads to complex responses, the most relevant of which is the inhibition of bile acids synthesis through the indirect repression of the expression of cytochrome 7A1 (CYP7A1), the rate limiting enzyme of this pathway. It forms part of a complex network encompassing PXR and PPARs that regulates the essential steps of bile acid and xenobiotic uptake, metabolism and excretion by hepatocytes, cholangiocytes and kidney cells.[30,31] The FXR gene is conserved from humans to fish[32] and, in humans and primates, encodes four FXRα isoforms (FXRα1, FXRα2, FXRα3 and FXRα4).[33] As for other non-steroid hormone NRs, FXRα binds to specific DNA response elements as an heterodimer with RXR.[34] Upon ligand binding, FXR undergoes conformational changes to release co-repressors such as NCor (Nuclear Co-repressor) and to recruit co-activators, such as SRC-1 (Steroid Receptor Co-

activator-1), PRMT (Protein Arginine(R) Methyl Transferase-1), CARM (Coactivator-Associated Arginine Methyltransferase-1), PGC (PPAR-γ Coactivator-1α) and DRIP (vitamin D Receptor-Interacting Protein-205). The mechanisms that regulate recruitment of these co-activators by FXR ligands and the relevance of these molecules to the regulation of specific genes by FXR ligands is still unknown.

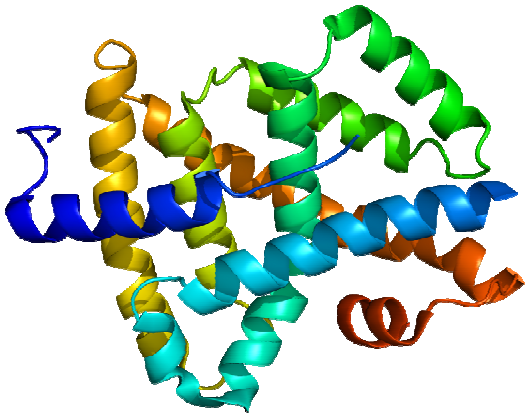


Figure 10. Structure of nuclear receptor FXR

After FXR discovery, specific bile acids (BAs) were identified[35,36,37] as endogenous ligands (Figure 11). The amphipatic properties of the bile acid skeleton displaying a convex hydrophobic face and a concave hydrophilic face are essential for their recognition in the FXR-LBD.[38] In contrast to other endogenous steroids, BAs nucleus adopts a bent shape due to the A/B *cis* ring juncture that forces ring A to lie outside of the plane of the BCD ring system, giving to BAs a profile that allows a close fit with respect to the pocket in FXR. Besides the β hydrophobic face is common in all BAs, the differences between the primary and secondary BAs are in the α face and in their specific pattern of hydroxylation at the 7 and 12 positions. Chenodeoxycholic acid (CDCA), the most effective activator of FXR, with its two hydroxyl groups at C-3 and C-7 oriented in a *cis* relationship transactivates FXR, whereas ursodeoxycholic acid (UDCA), with its

two hydroxyl groups at C-3 and C-7 oriented in a *trans* relationship does not activate this receptor. It creates a more open ligand binding pocket, and this arrangement may force a suboptimal orientation of helix 12 and results in partial inhibition.

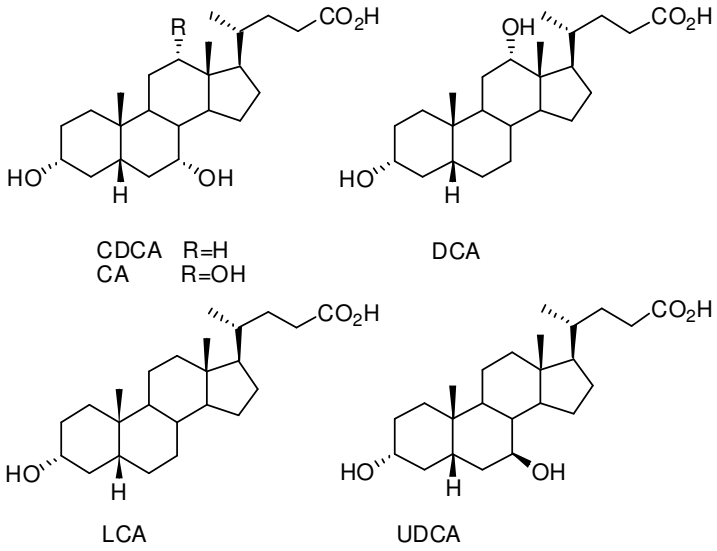


Figure 11. Structures of endogenous BAs as FXR ligands

Shortly, after FXR de-orphanization by BAs, potent FXR agonists have been generated to target liver and metabolic disorders. The most used is the non-steroidal isoxazole analog GW4064,[39] 3-(2,6-Dichlorophenyl)-4-(3'-carboxy-2-chlorostilben-4-yl)oxymethyl-5-isopropylisoxazole (Figure 12), a nanomolar nonsteroidal activator of FXR,[40] reducing the extent of hepatic injury when administred to rats rendered cholestatic by bile duct ligation or chemical intoxication with α-naphthyl-isothiocyanate. Because the clinical utility of GW4064 turned out to be limited because its short terminal half-life and limited oral exposure (< 10%), several derivatives modified in the stilbene functionality, recognized as toxic pharmacophore, have been designed and prepared. So obtained the 6-substituted 1-naphthoic acid is a full agonist essentially equipotent

to GW 4064. In a rodent model of chemically-induced cholestasis, both compounds increased Bsep and SHP and reduced Ntcp, Cyp7A1, alkaline phosphatase, alanine amino-transferase, total bile acids and direct bilirubin levels.

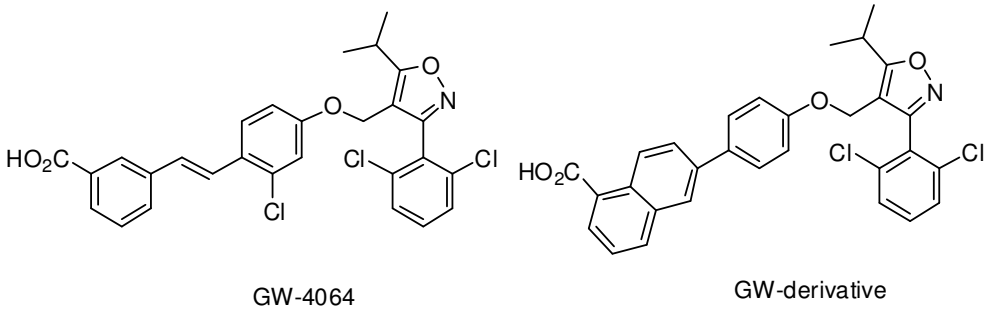
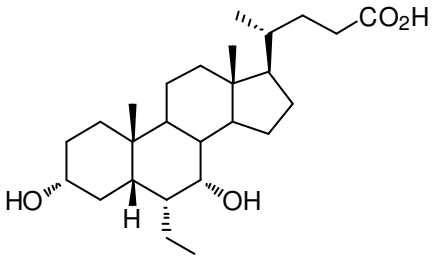


Figure 12. Structures of GW-4064 and 6-substituted 1-naphthoic acid

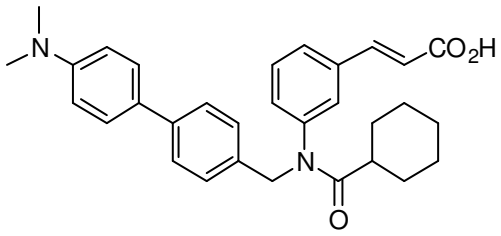
In the FXR-LBD the semisynthetic BA, 6-ethyl-CDCA[41] (Figure 13) places the 6α-ethyl group into one and additional hydrophobic cavity that exists between the side chains of Ile359, Phe363, and Tyr366, accounting for its higher affinity. It is bound to LBD with ring A directed toward Helix 11 and 12 of the LBD, while the carboxylic acid function of the side chain approaches the entry pocket at the back. 6-ECDCA was found effective in protecting against bile flow impairment induced by administration of estrogen $E_2$17α, a model of intrahepatic cholestasis with minimal or absent alteration of liver morphology. Similarly to GW4064, 6-ECDCA increased the liver expression of Bsep and SHP, while reduced Ntcp and Cyp7A1. In aggregate these preclinical observations support the notion that administration of potent FXR ligands in a cholestatic setting would induce a pattern of genes involved in hepatic detoxification and apical secretion of BA as well as inhibition of BAs uptake and BA synthesis. However, with the except of the estrogen model, FXR ligands are only partially effective in reducing cholestasis.



6-ECDCA

Figure 13. Structures of 6-ECDCA

Optimization of a benzopyrane-based combinatorial derived libraries had led to the identification of fexaramine,[42] 3-[3-[(Cyclohexylcarbonyl)-[[4'-(dimethylamino)-[1,1'-biphenyl]-4-yl]methyl]amino]phenyl]-2-propenoic acid methyl ester (Figure 14), as a new chemotype of FXR agonist, also endowed with nanomolar potency. In vitro assays established that fexaramine and related ligands robustly recruit the coactivator SRC-1 to FXR in a manner comparable to that of GW4064.



Fexaramine

Figure 14. Structure of Fexaramine

Despite the good results obtained with FXR agonists, a growing body of evidence is emerging about the negative impact of FXR activation on adaptation to cholestasis. FXR activation downregulates CYP7A1 inhibiting BAs synthesis eventually decreasing BAs pool size, the most important determinant of BAs secretory rate. In addition FXR activation reduces the expression/activity of those basolateral transporters such as MRP4, essential for BAs secretion in the systemic circulation. These observations suggest that FXR activation might impair the BAs

efflux, one of the key adaptative changes observed in cholestasis and therefore FXR antagonists might hold utility in the treatment of this desease. To date only few FXR antagonists are known and the main contribute is derived by natural compounds. Guggulsterone, isolated from the resin extract of the tree *Commiphora mukul*,[43] and Xanthohumol, the principal prenylated chalcone from beer hops *Humulus lupulus L.* [44] (Figure 15) were the first FXR antagonists to be reported from "Nature". However, guggulsterone is a promiscous agent wich binds and actives PXR, the glucocorticoid receptor and the progesterone receptor at concentrations that are approximately 100-fold lower than that required for FXR antagonism. [45,46]

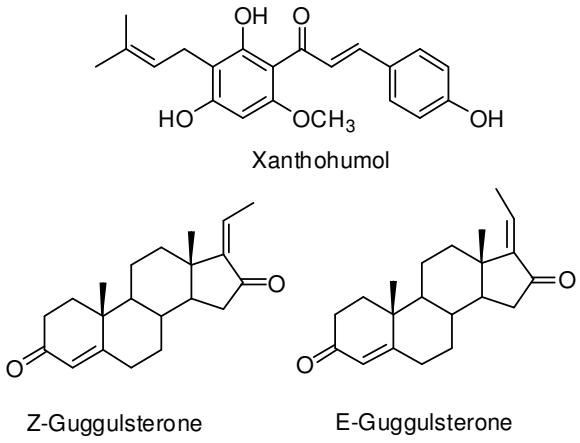


Figure 15. Structures of natural FXR antagonists.

In this context, the sea, with its extraordinary variety of organisms, has recently emerged as an evaluable source of FXR antagonists. As reported in this dissertation, my research work afforded the identification, for the first time, of several compounds endowed with promising activity on human NRs encompassing the first example of FXR antagonists from the "sea".

CHAPTER 1

STEROLS from THEONELLA SWINHOEI

In the last 30 years many sterols with unprecedented structures have been isolated from marine sources. Initially carbon skeleton modifications ranged from C27 to C29, with variation occurring exclusively in the side chain at C24.[47] After the discovery of the C26-sterols, first detected in 1970 from the mollusk *Placopecten magellanicus*[48] and later found widespread in marine invertebrates and also in a marine phytoplankton,[49] a number of "nonconventional" sterols have been reported. Unconventional steroids often co-occur with the conventional ones and are sometimes present in small amounts; however, many exceptions are reported for sponges producing unusual structures as the predominant steroids rather than cholesterol or the conventional 3β-hydroxy sterols.[50,51,52] When a sponge contains unusual steroids in large quantities, probably they play a functional (rather than metabolic) role in maintaining the integrity of membranous structures. It has been hypothesized and, to some extent, documented that the uniqueness of sterols in cell membranes of sponges is related to other components, particularly the phospholipids. These latter are formed by head groups and fatty acids very different from those of higher animals; therefore, the structural modifications exhibited by the sponge sterols may be a sort of structural adjustments for a better fit with other membrane components.[53,54,55] The sterols isolated from sponges are sometimes very complex mixtures of highly functionalized compounds, many of which have no terrestrial counterpart. These include sterols having side chains modified by the apparent loss of carbon atoms or by the addition of extra carbon atoms at biogenetically unprecedented positions of a normal Cα side chain, as

well sterols with unusual nuclei, containing a variety of oxygenated functionalities such as polyhydroxy, epoxide, epidioxy, and mono or polyenone systems. A plethora of inusual functional groups such as quaternary alkyl groups, cyclopropane and cyclopropene rings, allenes, and acetylenes has been found in the side chains of marine sterols and in figure 16 are reported the most representative but it is not an exhaustive list. [56,57]

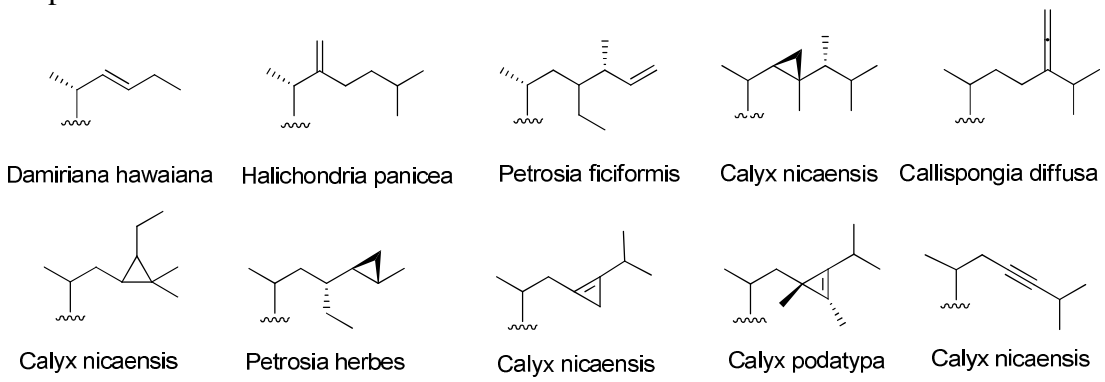
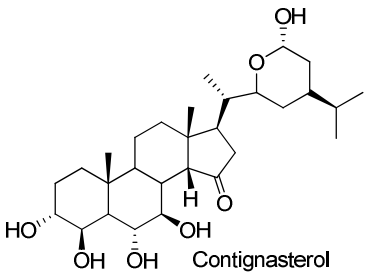
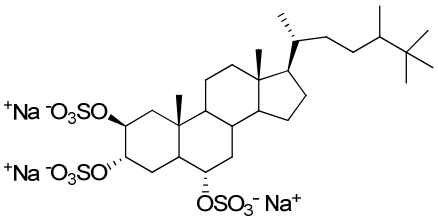


Figure 16. Examples of nonconventional side chains of sponge monohydroxysterols.

Highly functionalized steroids have attracted considerable attention because of their biological and pharmacological activities. A remarkable example is the potent inhibitor of histamine release from rat mast cells induced by anti-IgE. [58,59] Contignasterol that represents the first marine steroid found to have a *cis* C/D ring junction as well as a cyclic hemiacetal functionality at C-29 of the side-chain.



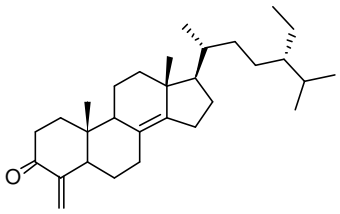
Halistanol sulfate, present in *Halichondriidae* sponges and characterized by the 2β,3α,6α-trisulfoxy functionalities and alkylation on the side chain, is the first



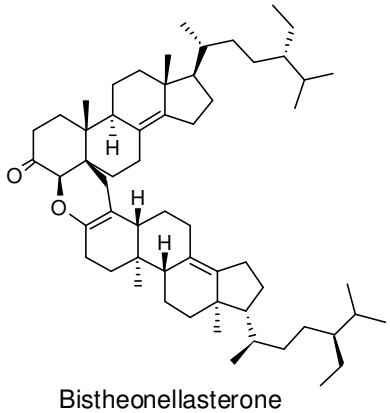
Halistanol Sulfate

example of sulfated sterol isolated from *Porifera*, with a potent anti-HIV activity.[60] Successively, several new sulfate sterols have been reported.

Other examples of sterols with unconventional nuclei are theonellasterone and bistheonellasterone, isolated from an Okinawan collection of *Theonella swinhoei*; bistheonellasterone represents a dimeric steroid biosynthesized from theonellasterone through a Diels-Alder cycloaddition with its Δ^4-isomer.



Theonellasterone



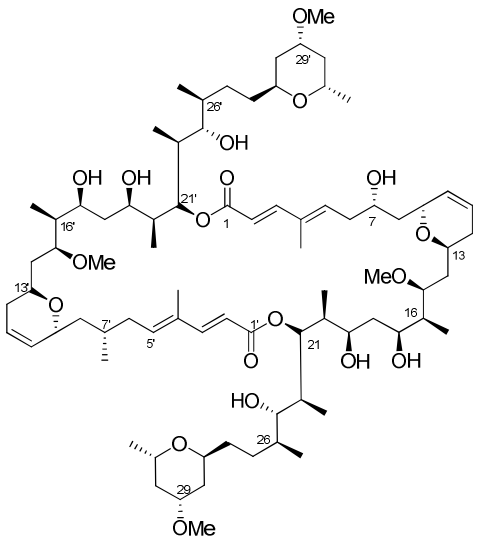
Bistheonellasterone

Indeed theonellasterone is the oxidized derivative of theonellasterol, the ideal biomarker of sponges of *Theonella* genus containing the rare 4-methylene steroids as exclusive components of the steroidal biogenetic class.

Theonella swinhoei

Theonella genus belongs to order Lithistida, an evolutionary ancient lineage that is typically found in deeper waters and caves of tropical oceans. Lithistid sponges have a structurally massive, rigid or "rock-like" morphology and are well known among the scientific community for the extraordinary chemio-diversity so far exhibited. Notably over half of the compounds reported for litisthid sponges were isolated from *Theonella* (family Theonellidae). *Theonella* species have been reported to contain a wide variety of diverse secondary metabolites with intriguing structures and promising biological

activities, which have been calculated to represent more than nine biosynthetic classes.[61] In particular, *Theonella swinhoei* represents one of the most prolific source of innovative and bioactive metabolites, which include complex polyketides as swinholide A and misakinolide A,[62,63] showing potent cytotoxic activity through the distruption of functionality of the actine cytoskeleton; tetramic acid glycosides as the antifungal aurantosides.[64,65,66]

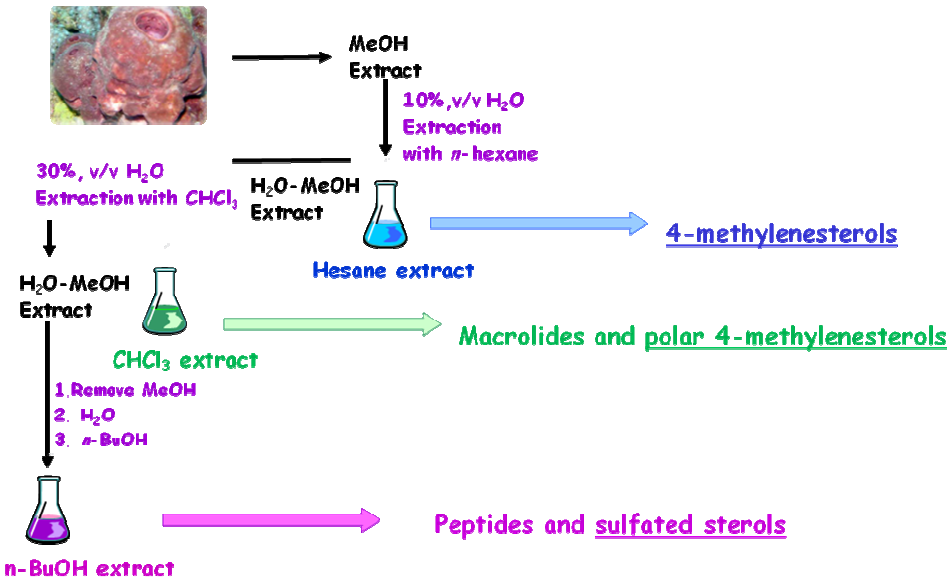


Swinholide A

The exceptional chemical diversity found in the metabolites isolated from *Theonella* sponges may in part be due to the biosynthetic capacity of bacteria that they host.[67] This hypothesis has been convincingly supported in the case of swinholide A, omnamides and theopederins. In 2005, Gerwich[68] reported the direct isolation of swinholide A and related derivatives from two different cyanobacteria, thus unequivocally demonstrating that marine cyanobacteria are the real productors of this class. Moreover, from the highly complex metagenome of *Theonella swinhoei*, the prokaryotic gene cluster,[69] likely responsible for the biosynthesis of omnamides and theopederins has been recently identified.[70,71]

In the course of a search for novel metabolites from marine sponges belonging to Lithistida order, I had the opportunity to study the sponge *Theonella swinhoei*. A specimen of sponge *Theonella swinhoei* was collected on the barrier reef of Vangunu Island, Solomon Islands, in July 2004. The samples were frozen immediately after collection and lyophilized to yield 207 g of dry mass.

Taxonomic identification was performed by Prof. John Hooper of Queensland Museum, Brisbane, Australia, and reference specimens are on file (R3170) at the ORSTOM Centre of Noumea. The lyophilized material was extracted with methanol and the crude methanolic extract was subjected to a modified Kupchan's partitioning procedure (Scheme 1).[72] Purification on the apolar extracts afforded macrolides and many polyhydroxylated sterols which have been demonstrated potent ligands of human nuclear pregnane receptor (PXR) and modulator of farnesoid-X-receptor (FXR). On the other hands, polar extract afforded the isolation of two new sulfated sterols, solomonsterols A and B, the first example of C-24 and C-23 sulfated sterols from a marine source endowed with a PXR agonistic activity;[73] a large family of cyclical peptides. Perthamides B-K, encompassing endowed with a potent anti-inflammatory and immunosuppressive activities[74] and two minor peptides, solomonamides A and B with an interesting anti-inflammatory activity and an unprecedented chemical skeleton.[75]

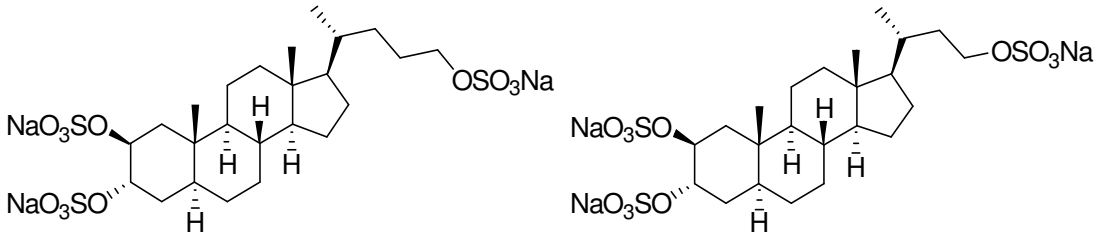


Scheme 1. Modified Kupchan's partitioning methodology applied to the sponge *Theonella swinhoei*.

CHAPTER 2

PXR AGONISTS

Sulfated steroids are a family of secondary metabolites often found in sponges and echinoderms. They are interesting not only from a structural point of view, but also because they often exhibite a variety of biological activities including anti-viral,[76,77] antifungal,[78] antifouling,[79] and action on specific enzymatic targets.[80,81,82,83] In a recent work, my group of research worked on the purification of the most polar fractions of *n*-BuOH extract of the sponge *Theonella swinhoei,* that afforded two new sulfated sterols with a 5-α-cholane and 24-nor-5-α-cholane skeleton**,** named solomonsterols A and B.[73] They possess a truncated side chain at C24 and C23 respectively, and three sulfoxy groups, two secondary sulfoxy groups, positioned on ring A at C2 and C3 of the steroidal nucleus, and one primary sulfoxy group on the side chain at C24 for solomonsterol A and at C23 for solomonsterol B. The A/B *trans* ring juncture represented the main structural difference respect to BAs with A/B *cis* ring juncture. (This A/B *cis* ring juncture is fundamental for activation of FXR.)



Solomonsterol A (**1**) Solomonsterol B (**2**)

Figure 17. Solomonsterols A (1) and B (2) from Theonella swinhoei.

Despite this difference they have been valuated as potential ligands for nuclear receptors. The results of these studies demonstrated that, while solomonsterols A

and B did not activate the farnesoid-X-receptor (FXR, data not shown), both agents were effective ligands for PXR, an evolutionary conserved nuclear receptor. The agonistic behavior of solomosterols toward PXR and PXR regulated genes, therefore was assisted by a transactivation in a cell based luciferase assay using an human hepatocyte cell line (HepG2 cells). Since PXR functions as an heterodimer with the retinoid-X-receptor (RXR), HepG2 cells were transfected with a PXR and RXR expressing vectors (pSG5-PXR and pSG5-RXR), with a reporter vector containing the PXR target gene promoter (CYP3A4 gene promoter) cloned upstream of the luciferase gene (pCYP3A4promoter-TKLuc) and with a β-galactosidase expressing vector as internal control of transfection efficiency (pCMV-β-gal). As illustrated in Figure 18, solomonsterols were potent inducers of PXR transactivation, boosting the receptor activity by 4-5 folds (n=4; $P<0.05$ versus untreated) and were at least as potent as rifaximin, a well characterized ligand for the human PXR.[84] This activity was then confirmed in a RT-PCR assay on PXR target genes; both agents effectively stimulated the expression of CYP3A4 and MDR1 in the same cell line (Figure 18; n=4; $P<0.05$).

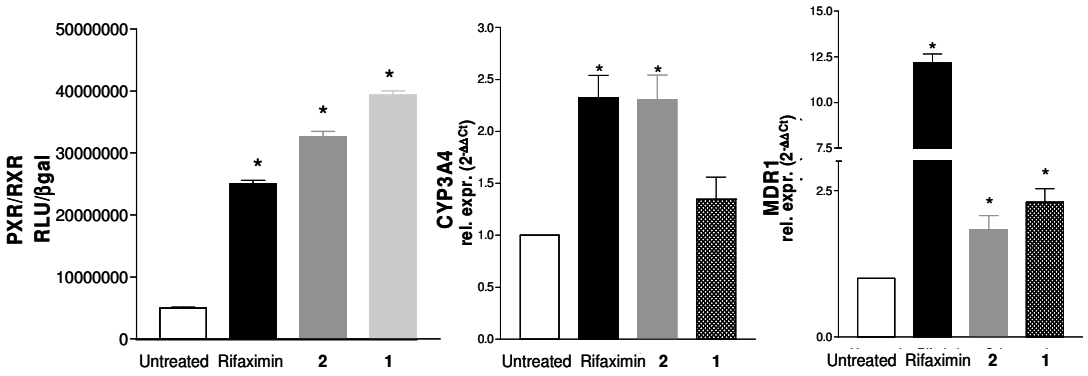


Figure 18. *Panel A*. Transactivation of PXR by compounds 1 and 2 in HepG2 cells. *Panel B and C*. Compounds 1 and 2 effectively increased the expression of CYP3A4 and MDR1, two PXR regulated genes in HepG2 cells. Expression of target genes was measured by RT-PCR.

Considering the well known relationship between PXR and immunity,[85] it was investigated whether Solomonsterols exert any effect on cells of innate immunity, the first line and the most ancient line of defence of mammalians against bacteria and viruses.[86] For this purpose, RAW264.7 cells, a murine macrophage cell line, were incubated with these compounds at the concentration of 10 and 50 μM in the presence of bacterial endotoxin (LPS) and expression of mRNA encoding for pro-inflammator mediators was measured by real-time (RT) polymerase chain reaction (PCR). As illustrated in Figure 19, at the concentration of 50 μM solomonsterols A and B effectively inhibited induction of the expression of interleukin-(IL)-1β mRNA (Figure 19; N=4;P<0.05 versusLPS alone) induced by challenging RAW264.7 cells with lipopolysaccharide (LPS), a tool like receptor (TLR)-4 ligand. By contrast both compounds exerted no effect on expression of tumor necrosis factor (TNFα) mRNA (Figure 19, panel B). At the end of the experiment cell viability was valued using the trypan blue (a diazo dye used to selectively colour dead tissues or cells blue). In fact both compounds had no effect on cell viability measured by the trypan blue excluding test (viability was greater than 95%) excluding that inhibition of IL-1β production was due to a cytotoxic effect.

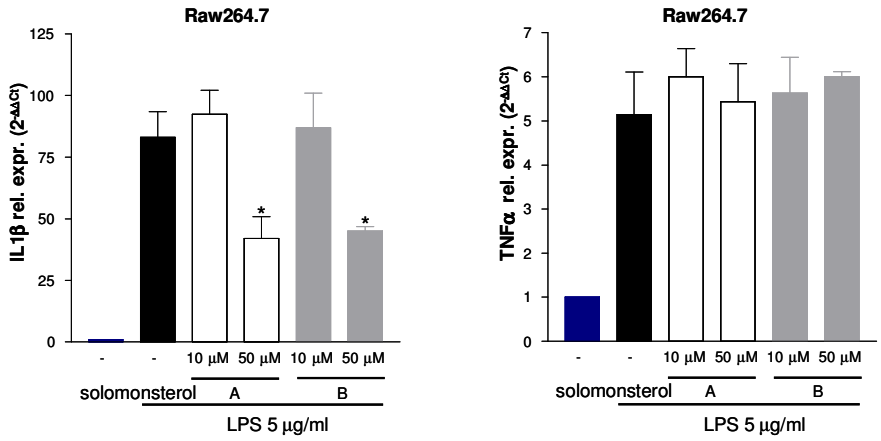


Figure 19. Relative mRNA expression of IL-1β and TNFα in RAW264.7 macrophages treated with LPS alone or with LPS in combination with Solomonsterol A and B.

Because IL-1β is a key cytokine and high in the hierarchy that drives innate immune response, these results highlight the potential for the use of solomonsterols in clinical conditions characterized by a dysregulation of innate immunity. To have details for what concerns the binding mode of solomonsterols A and B to PXR at atomic level, molecular docking studies were performed on solomonsterol A with PXR using Autodock Vina 1.0.3 software.[87] The docking results positioned solomonsterol A within the PXR binding pocket, and among the 9 docked conformations generated, the lowest binding energy displayed an affinity of -10.0 Kcal/mol (Figure 20). In this model, the steroidal nucleus establishes hydrophobic interactions with Leu206, Leu209, Val211, Ile236, Leu239, Leu240, Met243, Met246, confirming the binding mode already reported for a set of analogous compounds.[88] Moreover, the sulfate groups exert hydrogen bonds with Ser247 (3-*O*-sulfate), His407 (2-*O*sulfate), and Lys210 (24-*O*-sulfate, also protruding toward the solvent), providing the complex with an increased predicted stability fully compatible with the experimental biological assays.

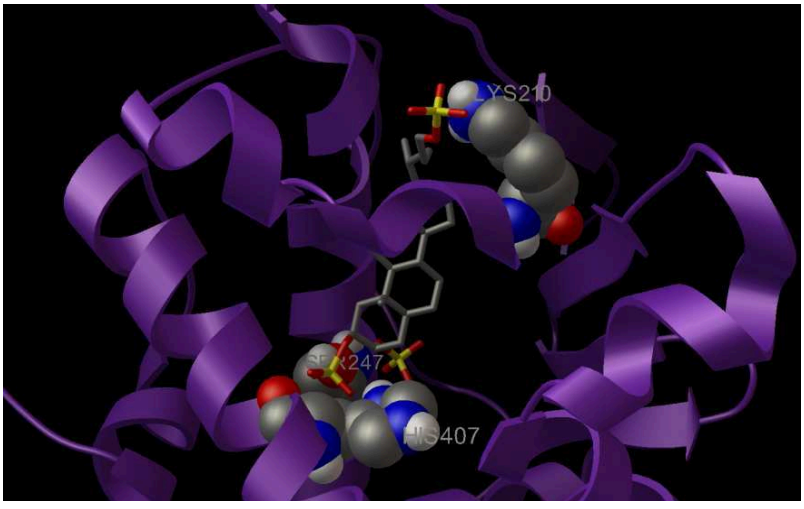


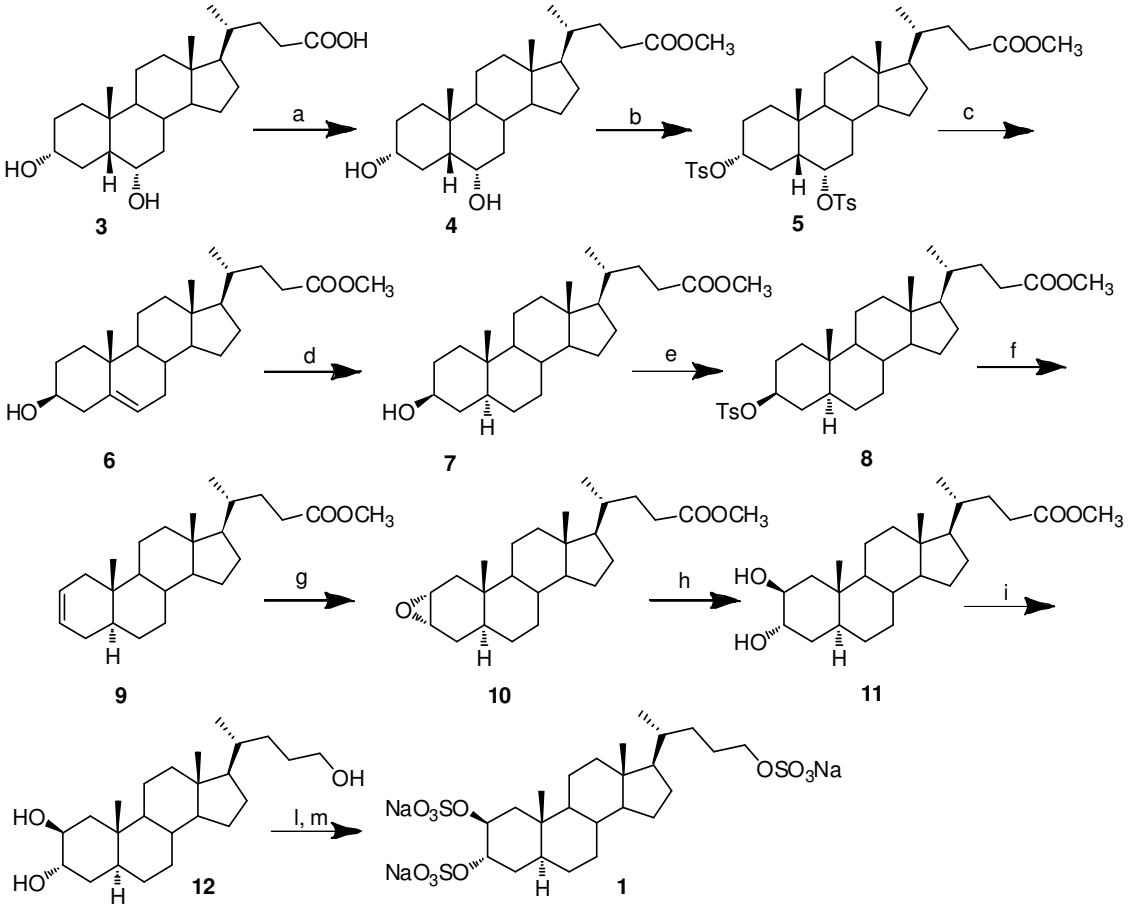
Figure 20. Docked model of solomonterol A bound to PXR model (pdb code: 1M13, displayed as purple ribbon); solomonsterol A is displayed as sticks coloured by atom type, while HIS407, SER247, and LYS210 are depicted as atom type coloured CPK models.

In conclusion, solomonsterols A and B are a novel class of PXR agonsts, isolated from *Theonella swinhoei*; such compounds could have a pharmacological potential for the treatment of human disorders characterized by dysregulation of innate immunity and with inflammation. SA and SB have been isolated in very small amounts from the biological source. To a further and detailed pharmacological evaluation, total synthesis of the two natural leads was accomplished.

2.1 Total synthesis of solomonsterol A

Key structural features of solomonsterol A (**1**) are the presence of a truncated C24 side chain, and three sulfated groups at C2, C3 and C24. We envisaged that the commercially available hyodeoxycholic acid (**3**) could be a suitable starting material to set up a robust route to prepare solomonsterol A in large amount.[89] Thus the total synthesis of solomonsterol A (**1**) started with **3**, which was methylated with diazomethane and treated with tosyl chloride in pyridine to give the corresponding 3,6-ditosylate (**5**) in nearly quantitative yield (Scheme 2). When **5** was treated with boiling DMF in the presence of CH_3COOK for 1 h, simultaneous inversion at the C-3 position and elimination at the C-6 position took place to give methyl 3-hydroxy-5-cholen-24-oate (**6**),[90,91] which in turn was hydrogenated to give the required A/B *trans* ring junction in **7**.[92] The simultaneous introduction of the 2β,3α-dihydroxy functionality was achieved by the following three-step sequence:[93,94] a) elimination at C3-position and consequent introduction of Δ-2 double bond; b) epoxidation with *m*-CPBA; c) acid catalyzed ring opening of the epoxide to afford diol **11**. β-Elimination and epoxydation were found to proceed with excellent regioselectivity and stereoselectivity, respectively, as determined by analysis of NMR spectra and

comparison of the NMR data of **9** and **10** with previously reported compounds. According to the Fürst–Plattner rule,[95] epoxide ring opening with sulfuric acid in THF provided the desired 2β,3α-diol **11** exclusively. The ^{1}H NMR signals of 2-H and 3-H (broad singlet at 3.89 ppm and broad singlet at 3.85 ppm) also confirmed the trans-diaxial disposition of the two hydroxy groups in **11**. Reduction of methyl ester at C24 with $LiBH_4$ afforded triol **12** in 92% yield. Treatment of **12** with 10 equivalents of triethylammonium–sulfur trioxide complex at 95 °C afforded the ammonium sulfate salt of solomosterol A, which was transformed via ion exchange into the desired target trisodium salt **1** (Scheme 2). The complete match of optical rotation, NMR and HRMS data of solomonsterol A with that of the natural product secured the identity of the synthetic derivative. This synthesis was completed in a total of ten steps starting from commercially available hyodeoxycholic acid (**3**) and had an overall yield of 31%. This route enabled us to prepare sufficient quantities of solomonsterol A to be further characterized in pharmacological tests. [89]



Scheme 2. Reagents and conditions: (a) CH_2N_2, quantitative; (b) *p*-TsCl, pyridine, quantitative; (c) CH_3COOK, DMF/H_2O 9:1, reflux, 78%; (d) H_2 (1 atm), Pd/C, THF/MeOH 1:1, 80%; (e) *p*-TsCl, pyridine; (f) LiBr, $LiCO_3$, DMF, reflux, 83% over two steps; (g) mCPBA, Na_2CO_3, CH_2Cl_2/H_2O 1:0.7; (h) H_2SO_4 1N, THF, 73% over two steps; (i) $LiBH_4$, MeOH/THF, 0 °C, 92%; (l) $Et_3N.SO_3$, DMF, 95 °C; (m) Amberlite CG-120, sodium form, MeOH, 90% over two steps.

2.1.1 Pharmacological evaluation *in vivo*

We have first investigated whether the synthetic solomonsterol A (**1**) transactivates hPXR in PXR transactivation assay. As illustrated in Figure 21, solomonsterol A (**1**) was equally effective as rifaximin in transactivating the hPXR in HepG2 cells. The relative EC_{50} was 2.2 ± 0.3 μM for rifaximin and 5.2 ± 0.4 μM for solomonsterol A (n=3).

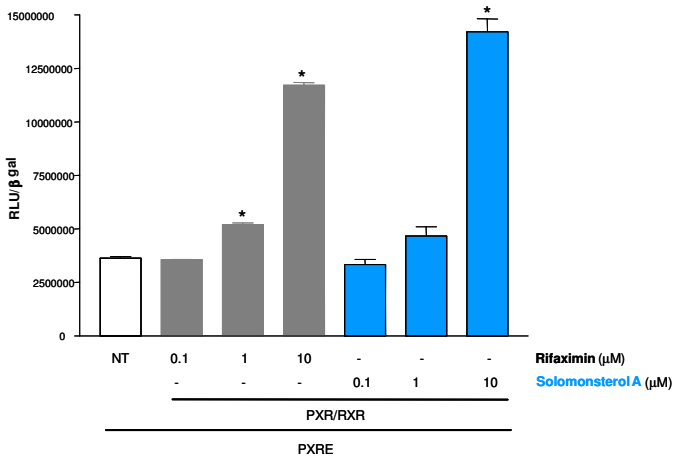


Figure 21. Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-βgal, and p(cyp3a4)TKLUC vectors and stimulated 18 h with (A) rifaximin or solomonsterol A (0.1, 1 and 10 μM). **P* < 0.05 versus not treated (NT)(*n* = 4).

Colon inflammation that develops in mice administered TNBS (trinitrobenzenesulfonic acid) is a model of a Th1-mediated disease with dense infiltrations of lymphocytes/macrophages in the *lamina propria* and thickening of the colon wall.[96,97] In order to assess whether solomonsterol A would exert immune-modulatory activity, TNBS was administered to C57Bl/6 transgenic mice expressing the human PXR. In these experiments, mice were treated with solomonsterol A and rifaximin for 7 days starting 3 days before intrarectal administration of TNBS.

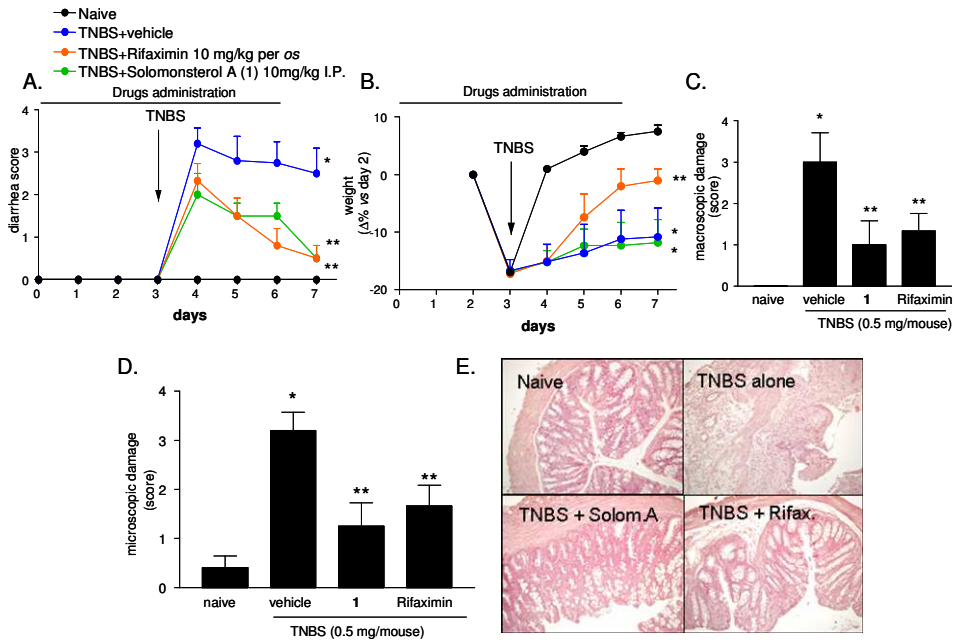


Figure 22. Colitis was induced by intrarectal administration of 0.5 mg of TNBS per hPXR mouse, and animals were sacrificed 4 days after TNBS administration. Solomonsterol A (**1**) and rifaximin were administered intraperitoneally (I.P.) and orally (per *os*), respectively, for 3 days before TNBS. The severity of TNBS-induced inflammation (**A**, diarrhea score, **B**, weight loss, **C** macroscopic colon damage) is modulated by rifaximin and solomonsterol A (**1**) administration. **D** microscopic colon damage, **E** histological analysis of colon samples (original magnification 40×, H&E staining). TNBS administration causes colon wall thickening and massive inflammatory infiltration in the *lamina propria*.

As shown in Figure 22, administering hPXR transgenic mice with solomonsterol A (**1**) effectively attenuated colitis development as measured by assessing local and systemic signs of inflammation. Thus, similarly to rifaximin, treatment with **1** at the dose of 10 mg/kg protected against the development colitis, as measured by diarrhea score and the weight loss (Figure 22A and B, n=6-7; *p<0.05 versus naïve; **p<0.05 versus TNBS group) and reduced the macroscopic score of colitis as well as the microscopic score (Figure 22C-D, n=6-7; *p<0.05 versus naïve; **p<0.05 versus TNBS group). As shown in Figure 22E, histopathological samples of colon removed from hPXR transgenic mice administered TNBS had an extensive cellular infiltrate, submucosal edema, and large areas of epithelial erosions. These changes were robustly attenuated by pre-treatment with solomonsterol A (**1**) or rifaximin.These effects were supported by an attenuation

of signs of inflammation-driven immune dysfunction induced by TNBS administration. Thus, similarly to rifaximin, solomonsterol A (**1**) reduced neutrophils accumulation in the colonic mucosa as assessed by measuring MPO (myeloperoxidase) activity, as well as the expression of a number of signature cytokines and chemokines including TNFα (tumor necrosis factor alfa), IFNγ (interferon gamma), IL-12p70 (interleukin-12 p70 subunit) and MIP-1α (macrophage inflammatory protein-1α) (Figure 23). Of interest, both rifaximin and solomonsterol A (**1**) effectively increased the colon expression of IL-10 (interleukin-10), a key counter-regulatory cytokine. A similar pattern, thought non significant for solomonsterol A (**1**), was observed for TGFβ (transforming growth factory beta) mRNA, a growth factor whose colon expression is linked to generation of a subset of regulatory T cells (Treg)[98] (Figure 23F and G, n=6-7; *p< 0.05 versus naïve; **p<0.05 versus TNBS).

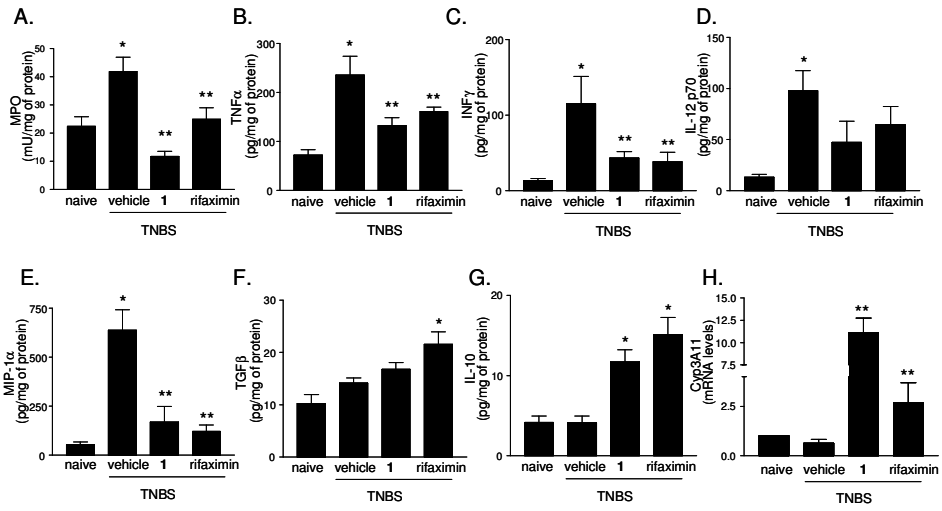


Figure 23. MPO (**A**), TNFα (**B**), INFγ (**C**), IL-12 p70 (**D**), MIP-1α (**E**), TGF-β1 (**F**) and IL-10 (**G**) protein levels in colons obtained 5 days after administration of TNBS alone or in combination with rifaximin and **1**. Treatments inhibit the increase of neutrophil infiltration (MPO levels) and proinflammatory factors induced by intrarectal administration of TNBS. Data represent the mean ± SE of 6-7 mice per group. (*p< 0.05 *vs* naïve; **p<0.05 *vs* TNBS). (**H**) Effects of rifaximin and solomonsterol A in colon mRNA levels of a PXR target gene such as Cyp3A11 (data are mean ± SE of 5 animal per group *p< 0.05 solomonsterol A *vs* rifaximin; **p<0.05 *vs* TNBS).

Finally we found that administering hPXR mice with solomonsterol A (**1**) effectively triggered PXR activation *in vivo*. Indeed, as shown in Figure 23H, both solomonsterol A (**1**) and rifaximin caused a potent induction in the expression of Cyp3A11. In the mice, Cyp3A11 is the orthologue of the human CYP3A4 gene in and it is a PXR regulated gene highly expressed in the intestine. These data strongly indicated that solomonsterol A and rifaximin are PXR agonists *in vivo* (Figure 23H, n=6-7; *p< 0.05 versus naïve; **p<0.05 versus TNBS). In addition we have evaluated the effect of solomonsterol A on NF-κB (nuclear factor-kappaB) DNA binding activity in LPS (lipopolysaccharide)-activated monocyte cell line. As shown in Figure 24, compared to LPS alone, co-treatment with solomonsterol A resulted in robust reduction of LPS induced NF-κB DNA binding activity as measured in EMSA assay. These results suggest that solomonsterol A inhibits the LPS-induced cytokines and chemokines expression by suppressing NF-κB DNA-binding activity.

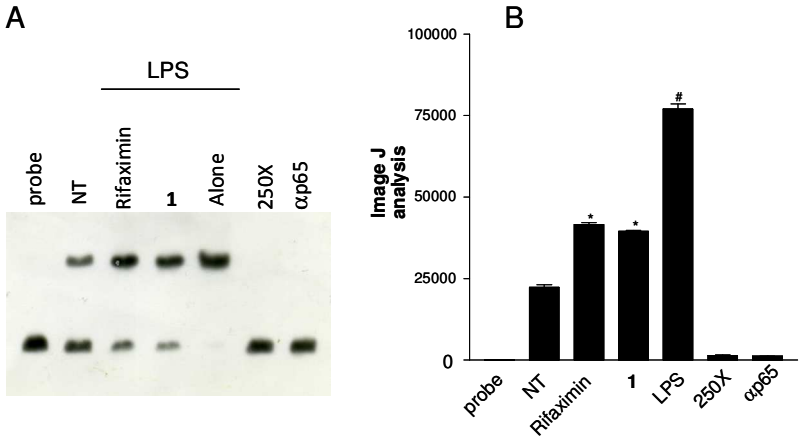


Figure 24: Rifaximin and solomonsterol A (**1**) suppress LPS-induced NF-κB DNA binding activity in THP-1 (human acute monocytic leukemia) cells that are a monocyte-like cell line. The cells were pre-treated with rifaximin (10 μM) or 1 (10 μM) with vehicle for 3 h and then stimulated with LPS (1 μg/mL) for 20 h. (**A**) DNA binding activity of NF-κB p65 subunit was analyzed by EMSA. For the competition assay, a 250-fold excess of unlabeled probe was added together with the labelled probe. For the supershift assay, 1 μg of antibody against NF-κB p65 was added together with the nuclear extract. Similar results were observed in three independent experiments. (**B**) Desintometric analysis of p65 binding to NF-κB responsive elements shown in the left panel.

Because these data demonstrate that prophylactic treatment with solomonsterol A (**1**) effectively protects against colitis development, we have investigated whether this agent is effective in driving the healing of an established active colitis. For this purpose, solomonsterol A was administered in a therapeutic manner in mice rendered colitic by TNBS administration. As illustrated in Figure 25, when administered to mice on day 1 after TNBS administration, solomonsterol A (**1**) effectively attenuated clinical signs of colitis (Figure 25A and B), including the diarrhea score and wasting disease. In addition, solomonsterol A (**1**) attenuated the macroscopic and microscopic scores as well as the MPO activity, a measure of neutrophil infiltration into the colonic mucosa (Figure 25C-E).

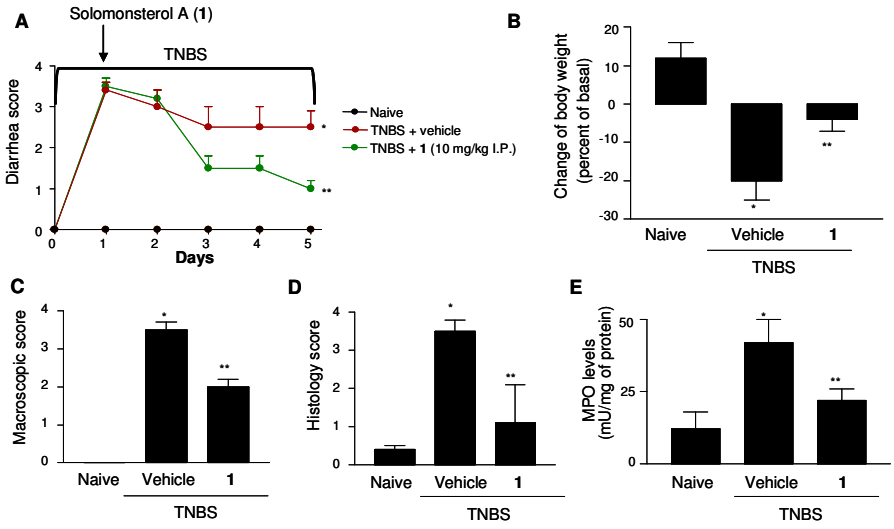


Figure 25. Colitis was induced by intrarectal administration of 0.5 mg of TNBS per mouse, and animals were sacrificed 5 days after TNBS administration. Solomonsterol A (**1**) was administered on day 1 after TNBS administration. The severity of TNBS-induced inflammation (A, diarrhea score, B, weight loss, C, macroscopic colon damage, D, microscopic score damage and E, MPO activity) was reduced by solomonsterol A (1) administration. Body weight is expressed as delta percentage versus the weight of mice on the day before TNBS administration. Data represent the mean ± SE of 4-6 mice per group (*p< 0.05 vs naïve; **p<0.05 vs TNBS).

In conclusion SA acts as a potent hPXR agonist as demonstrated by the induction of a canonical target gene of PXR in colon, such as Cyp3A11 (Figure 23H), effectively it protects against development of clinical signs and symptoms of

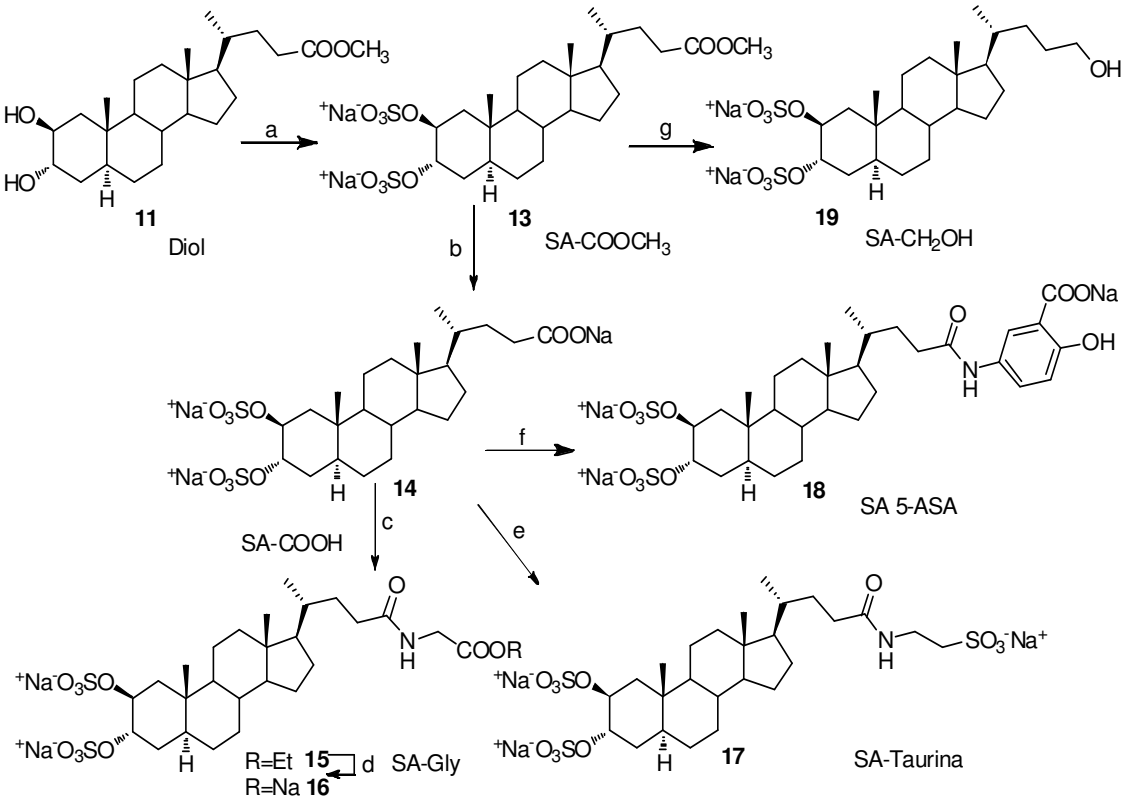
colitis, reduces the generation of TNFα and enhances the expression of TGFβ and IL-10, two potent counter-regulatory cytokines in IBD, via inhibition of NF-κB activation in a PXR dependent mechanism.

2.2 Modifications in the side chain of solomonsterol A

We have identified solomonsterol A (**1**) as as a new lead in the treatment of IBD.[89] However one of the possible limitation to its use in clinical settings is that, when administered *per os*, solomonsterol A could undergo absorption from the GIT before reaching the colon causing severe systemic side effects resulting from the activation of PXR in the liver. One of the best approaches used for colon specific drug delivery is based on the formation of a prodrug through chemical modification of the drug structure, usually by the conjugation with a suitable carrier, such as amino acids, sugars, glucuronic acid, dextrans or polysaccharides. Since the luxuriant microflora presents in the colon, the prodrug undergoes enzymatic biotransformation in the colon thus releasing the active drug molecule. Another challenging task is the design of a dual-drug able to release in the colon two molecules acting in a synergic manner. For example the possible eventual chemical linkage of solomonsterol A (**1**) to 5-ASA (5-aminosalycilic acid), one of the oldest anti-inflammatory agents in use for the treatment of IBD, could produce a dual-drug with enhanced potency. Upon enzymatic hydrolysis in the colon, this kind of molecule could release solomonsterol A and 5-ASA, potent agonists of PXR and PPARγ,[99] respectively, two nuclear receptors playing a key role in colon inflammation diseases. When synthesizing prodrugs, the first step is the introduction of a functional group on the drug molecule suitable of conjugation with a selected carrier (*e.g.*, an hydroxyl group that could enter into a glycoside linkage with various sugars, or alternatively a carboxyl group to form ester e/o

amide conjugates with cyclodextrins, amino acids etc). Inspection of chemical structure of solomonsterol A (**1**) revealed that the presence of three sulfate groups hampered any further derivatization e/o conjugation. In order to introduce a function group suitable for further derivatization, we decided to prepare several solomonsterol A (**1**) derivatives with a modified side chain but preserving the steroidal tetracyclic nucleus.[100] Our synthetic route started from the advanced intermediate **11**[89] that was sulfated with 10 equivalents of triethylammonium-sulfur trioxide complex and transformed in the sodium sulfate salt **13** through Amberlite CG-120 treatment. The crude product was subsequently hydrolyzed with methanolic NaOH (5%) to remove the protecting group at the C-24 methyl ester on the side chain affording the desired carboxylic acid functional group. The reaction mixture was adjusted to pH 5 with HCl 1N, and loaded onto a C18 cartridge for the reversed-phase solid extraction. Elution with 30% aqueous methanol gave the carboxylic acid **14** as a 2,3,24-trisodium salt in satisfactory yield (85% over two steps). Having obtained the carboxyl acid at C24, we decided to carry on with the reaction of amidation with glycine ethyl ester, taurine and 5-ASA. Using the versatile coupling agent, DMT-MM [4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride],[101] the amidation reaction proceeded nearly quantitatively, requiring the activation of the carboxylate sodium salt by DMT-MM and triethylamine in DMF at room temperature and subsequent condensation of the resulting acyloxytriazine with glycine ethyl ester hydrochloride, taurine and 5-ASA affording the amide derivatives **15**, **17** and **18** respectively, as ammonium sulfate salts. Alkaline hydrolisis of ethylester **15** with NaOH 5% in MeOH/H_2O 1:1 afforded the sodium carboxylate **16**. Amide derivatives with taurine and 5-ASA were transformed via ion exchange

(Amberlite CG-120, sodium form, MeOH) into the desired target trisodium salts **17** and **18** in nearly quantitative yields (Scheme 3).



Scheme 3. a) $Et_3N{\cdot}SO_3$, DMF, 95 °C; b) NaOH 5% in MeOH:H_2O 5:1 v/v, 85% over two steps; c) DMT-MM, Et_3N, GlyOEt, DMF dry; d) NaOH 5% in MeOH:H_2O 5:1 v/v, 58% over two steps; e) DMT-MM, Et_3N, taurine, DMF dry. Then Amberlite CG-120, MeOH, 67%. f) DMT-MM, Et_3N, 5-ASA, DMF dry. Then Amberlite CG-120, MeOH, 72%; g) $LiBH_4$, MeOH, THF, 0 °C, 75%. Elution with 30% aqueous methanol gave the carboxylic acid 3 as a 2,3,24-trisodium salt in satisfactory yield (85% over two steps).

To prove the ability of these compounds to activate PXR and eventually PXR regulated genes, a luciferase reporter assay on human hepatocyte cell line (HepG2 cells) transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-βgalactosidase, and p(CYP3A4)-TK-Luc vectors (Figure 26), has been performed. Cells were then stimulated with rifaximin, SA and with compounds **14-18** at the concentration of 10 μM each. As shown in Figure 26A, beside the closely structural resemblance with solomonsterol A (**1**), only carboxylate (**14**) showed a slight activity in transactivating PXR. Besides at first sight this behaviour should

be ascribable to a scarce bioavailability, the scarce activity also for the methyl ester **13** (Figure 26A) and the complete loss of activity for C-24 alcohol **19** (Figure 26A), obtained through $LiBH_4$ reduction of **13** (75% yield), pointed towards unfavourable pharmacodinamic features. Indeed, although compounds **14-18** possess a negative charge on their side chains, most likely they are less able to form polar interactions with Lys210[89] or alternatively with other polar amino acids of PXR LBD.

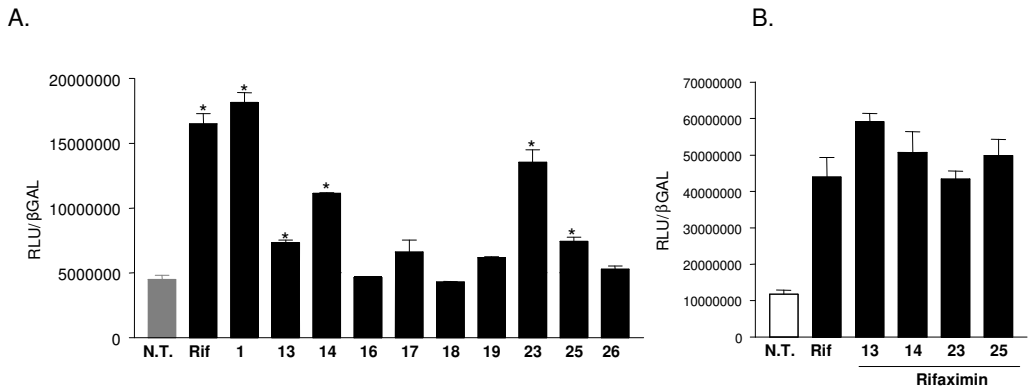
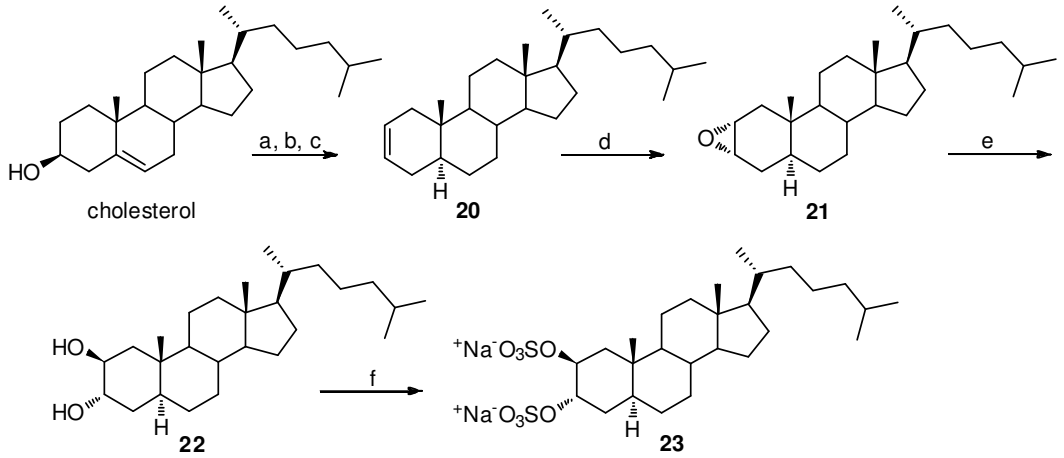


Figure 26. Luciferase reporter assay. HepG2 cells, a hepatocarcinoma cell line, were transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-βgalactosidase and p(CYP3A4)-TK-Luc vectors and then stimulated with (**A**) 10 μM rifaximin or compounds **1, 13, 14, 16, 17, 18, 19, 23, 25 and 26** for 18 h, or (**B**) 10 μM rifaximin alone or in combination with 50 μM of compounds **13, 14, 23 and 25** . N.T., not treated. Rif, rifaximin. *P<0.05 versus cells left untreated. Data are mean± SE of three determinations.

2.2.1 Discovery of cholesterol disulfate and pharmacological investigations

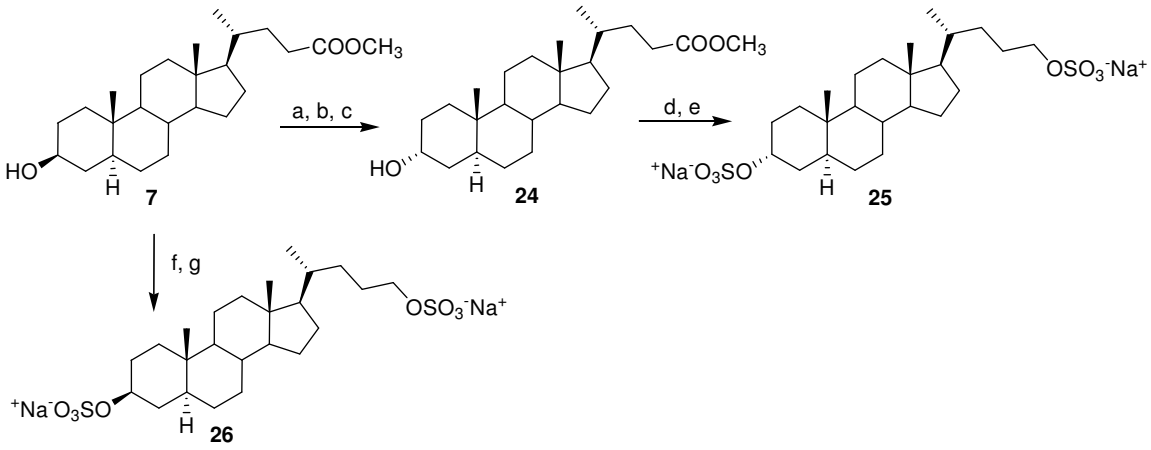
As previously reported, PXR presents a large ligand binding cavity[102] allowing the accommodation of different kind of molecules and the possible binding modes are characterized an adequate balance between hydrogen bond and Van der Waals interactions established between a small molecule and the receptor. Therefore the lack of a polar interaction should be get over by increasing the contribution of the hydrophobic interactions on the side chain, so the derivative **23**, 2β,3α-cholestan disulfate, has been prepared according the synthetic procedure reported in Scheme 4.[101]



Scheme 4. a) H_2, Pd/C, THF:MeOH 1:1, room temperature, 90%; b) *p*-TsCl, pyridine; c) LiBr, Li_2CO_3, DMF, reflux, 87% over two steps; d) *m*-CPBA, $CHCl_3$ room temperature; e) H_2SO_4 1N, THF, room temperature, 78% over two steps; f) $Et_3N{\cdot}SO_3$, DMF, 95 °C. Then Amberlite CG-120, MeOH, 90%.

Δ^5 cholesterol reduction (H_2, Pd/C, THF:MeOH 1:1) followed by tosylation and LiBr elimination afforded Δ^2-cholestane derivative **20** (78% yield in three steps). The introduction of the 2β,3α-dihydroxy functionality was achieved by epoxidation with *m*-CPBA followed by acid catalyzed ring opening on epoxide **21**.[103] β-Elimination and epoxidation proceeded with excellent regioselectivity and stereoselectivity, providing exclusively the desired 2β,3α-diol **22** in excellent yields (78% over two steps). Sulfation of diol **22** followed by Amberlite CG-120 treatment and RP-18 chromatography afforded the disodium salt **23** in good yields. As shown in Figure 26A, compound **23** with its hydrophobic side chain is able to transactivate PXR with a potency comparable with the parent solomonsterol A (**1**). Having set a flexible synthetic strategy, we decided to speculate the pharmacophoric role played by the sulfate groups on ring A in the PXR agonistic activity of solomonsterol A (**1**). Tosylation of methyl 3β-hydroxy-5α-cholan-24-oate (**7**)[89] followed by inversion of configuration at C-3 with potassium acetate in DMF/H_2O and de-acetylation in acidic condition (Scheme 5) afforded the 3α-hydroxy derivative **24** (75% over three steps). Reduction at C-24, sulfatation/Amberlite ion exchange gave **25** as disodium

salt. Methyl 3β-hydroxy-5α-cholan-24-oate (**7**) was also used as starting material for the easy transformation in derivative **26** through $LiBH_4$ reduction of C-24 methyl ester and successive sulfation of the alcoholic functions at C-3 and C-24.



Scheme 5. a) *p*-TsCl, pyridine; b) CH_3COOK, DMF:H_2O 9:1, reflux; c) *p*-TsOH, $CHCl_3$:MeOH 5:3, 75% over three steps; d) $LiBH_4$, MeOH, THF, 0 °C, 85%; e) $Et_3N{\cdot}SO_3$, DMF, 95 °C; then Amberlite CG-120, MeOH, 63%; f) $LiBH_4$, MeOH, THF, 0 °C, 72%; g) $Et_3N{\cdot}SO_3$, DMF, 95 °C; then Amberlite CG-120, MeOH, 78%.

As indicated in Figure 26A, besides compound **25** induces a slight PXR transactivation, the lack of sulfate group at C-2 as well as the inversion of configuration at C-3 are responsible of a general loss in the agonistic activity towards PXR. To investigate whether these compounds could act as potential antagonists of PXR we have carried out a transactivation experiment in HepG2 cells stimulated with rifaximin (10 μM) and compounds **13**, **14**, **23** and **25** at the concentration of 50 μM each. As shown in Figure 26B, all compound failed to reverse the induction of luciferase caused by rifaximin, indicating that none of these solomonsterol A derivatives is a PXR antagonist. To further examine the activity of compound **23** as PXR activator and further clarify the behavior of compounds **13**, **14** and **25**, we have tested the effects of all members of our series on the expression CYP3A4, a canonical PXR target gene (Figure 27). Despite compounds **13**, **14** and **25** caused a slight transactivation of PXR, they failed to modulate the expression of

CYP3A4 at the concentration of 10 μM. In contrast, confirming data shown in Figure 26, compound **23** effectively increased the expression of CYP3A4 (Figure 27) in HepG2 cells, with a magnitude similar to that of rifaximin and solomonsterol A (**1**).

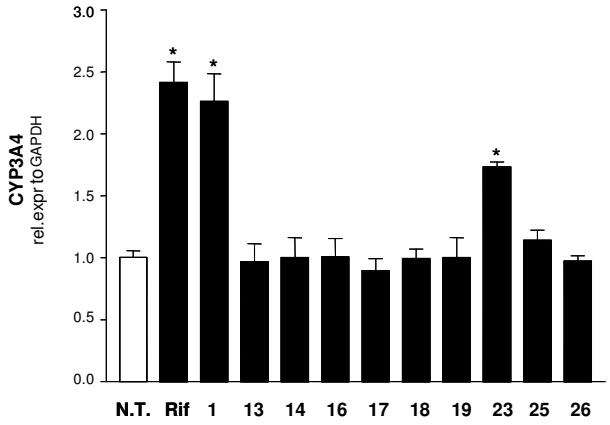


Figure 27. Real-Time PCR of CYP3A4 carried out on cDNA isolated from HepG2 not stimulated or primed with 10 μM rifaximin, and compounds **1,13,14,16,17,18,19,23,25 and 26**. N.T., not treated. Rif, rifaximin. *P<0.05 versus N.T. cells.

To investigate whether compound **23** plays an immunomodulatory activity, the effect of **23** in regulating immune response using THP1 cells, a human macrophage/monocytic cell line, challenged with lipopolysaccharide (LPS), has been evaluated. Previous studies[94] have shown that activation of PXR in this setting attenuates immune response triggered by LPS. Data shown in Figure 28, demonstrate that compound **23** effectively attenuates induction of IL-1β, TNFα and MCP-1 induced by LPS.

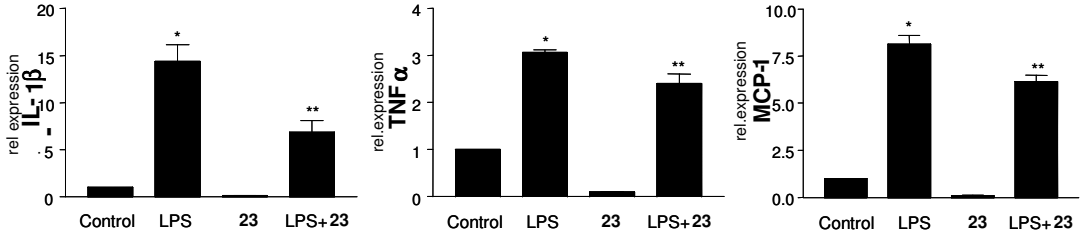


Figure 28. Effect of compound **23** on cytokine release induced by LPS in THP1 cells. $3x10^6$ THP were starved for 24 h and then pre-treated with 10 μM of compound **23** for 3 h and then stimulated for 18 h with LPS 1 μg/mL. Cytokine expression was assessed by RT-PCR. Data shown are mean ± SE of 9 assays from three different sets of experiments. *P<0.05 versus control cells; ** P<0.05 versus LPS alone.

Because the above mentioned data indicate that compound **23** effectively modulates immune response in human monocytes, additional experiments were carried out to investigate the effect of this compound in another model of inflammation-driven activation, using hepatic stellate cells (HSCs). HSCs are a liver-resident cell population that proliferates in response to liver injury. In response to immune activation, HSC undergoes a complex phenotype's rearrangement characterized by resetting expression of nuclear receptors, including PXR, and acquisition of an activated, myofibroblast-like phenotype whose main characteristic is the ability to express α-smooth muscle actin (αSMA). HSCs are recognized as the main source of extracellular matrix production in the fibrotic liver. Previous studies have shown that, along with other nuclear receptors, PXR ligands reverse this phenotype and reduce α-SMA expression.[104,105] For this purpose HSCs were exposed to thrombin, a proteinase activated receptor (PAR)-1 agonist alone or in combination with compound **23**. Previous studies have shown that thrombin drives HSCs trans-differentiation and its inhibition reverses HSCs from an activated to a quiescent phenotype.[106] Results shown in Figure 29, demonstrate that not only, similarly to solomonsterol A (**1**), compound **23** effectively reduces basal expression of αSMA, but it also attenuates HSCs trans-differentiation (i.e. induction of αSMA expression) triggered by thrombin.

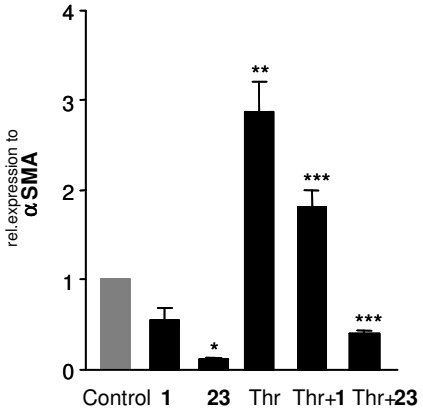


Figure 29. HSC-T6 cells were starved for 72 h and then stimulated with thrombin, 10 U/mL, in the presence of solomonsterol A (**1**) or compound **23**, 10 μM each. αSMA expression was assessed by RT-PCR. Data shown are mean ± of three experiments.* P<0.05 versus control cells; ** P<0.05 thrombin versus control cells; ***P<0.05 versus thrombin alone.

2.2.2 Docking studies

In order to clarify the different activities of described compounds at a molecular level,[107] their positioning in PXR-LBD in comparison with SA was evaluated by docking calculations.[108] As previously described, the three sulfate groups of solomonsterol A (**1**) act as key points of interactions with polar amino acids in LBD (Ser247, His407, Lys210), and contribute to accommodate the steroid nucleus in a mostly hydrophobic part of the binding site of PXR. Solomonsterol A (**1**) establishes hydrogen bonds (Figure 30) with Cys284 (2-*O*-sulfate) and with the Lys210 (24-*O*-sulfate) and electrostatic interactions with the Ser247 (2-*O*-sulfate) and His407(3-*O*-sulfate).[89]

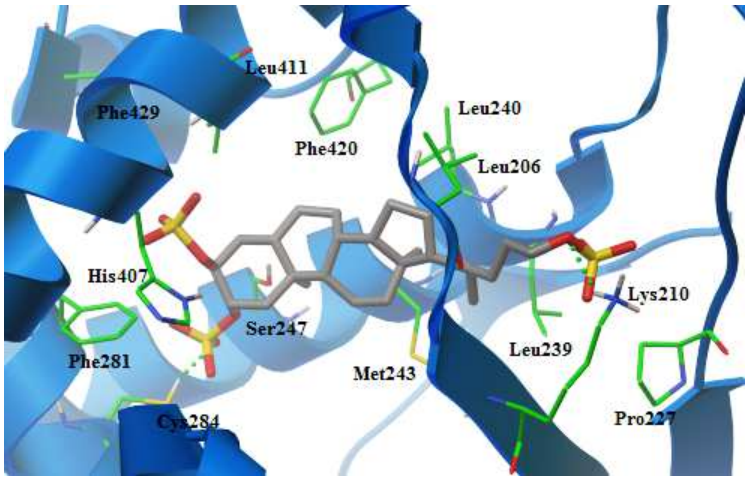


Figure 30. Solomonsterol A (**1**) (coloured by atom types: C grey, O red, S yellow) in docking with PXR-LBD (residues are coloured by atom type: C green, H light grey, O red, N blue). Hydrogen bonds are displayed with green spheres.

Compound **23**, featuring the C8 aliphatic side chain of cholesterol, is well superimposed with the binding pose of **1**, and is able to interact with the Ser247, Cys284 and the His407 through its two sulfate groups in the ring A (Figure 31). Moreover, **23** establishes hydrophobic interactions with almost all the residues observed for solomonsterol A (**1**) (Leu209, Val211, Pro228, Leu239, Met243, Phe281, Phe288, Leu411). The presence of an hydrophobic chain allows to gain two more Van der Waals interactions (with the Leu209 and Val211) that may counter the loss of electrostatic interaction observed for the sulfate group at C24 of parent solomonsterol A (**1**). Nevertheless, the weaker nature of these Van der Waals interactions could explain the decrease of the activity of **23** on PXR (difference of predicted binding energies **1**-**23**=1.05 kcal/mol).

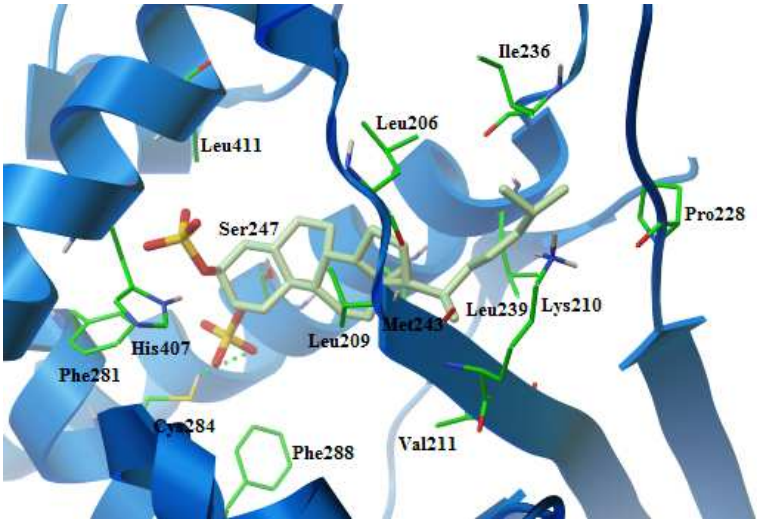


Figure 31. Compound **23** (coloured by atom types: C light green, O red, S yellow) in docking with PXR-LBD (residues are coloured by atom type: C green, H light grey, O red, N blue). Hydrogen bonds are displayed with green spheres.

On the other hand, the absence of the sulfate group at C-2 in the steroid nucleus causes the observed decrease of activity, due to an inability to interact simultaneously with the three key points of contact previously described. For example, compounds **25** and **26** are able to interact with the Lys210 but they fail to respect the key interactions involving the internal part of the binding site (Figure 32). As concern compound **14**, its tetracyclic nucleus is well superimposed with **1**, but its shorter side chain causes a poor interaction with the nitrogen of Lys210. The two oxygens of its terminal carboxylic part are not well overlapped with the oxygens of the 24-*O*-sulfate of the **1**, and the different arrangement of the side chain causes also a loss of two Van der Waals interactions with the Leu239 and Pro227 (Figure 32). The rings A of compounds **13** and **19** are in the place occupied by the ring B of **1** and, as a consequence, the 2-*O*-sulfate and/or 3-*O*-sulfate are in a less deep position (Figure 32). Compounds **16**, **17** and **18** present a longer and more functionalized side chain (Figure 32) compared with the previous derivatives, but also in this case the steroid nucleus are placed toward the external part of the binding site of PXR (**16**, **18**). Moreover, compound **17** is

unable to bind in the above described fashion and accommodates in a reverse orientation (a flipping of ~ 180° along the major axis of the steroid nucleus) of its steroid nucleus (Figure 32). The overall result is an inverted disposition of all the chemical groups (sulfates/methyl groups, and side chain) in the binding pocket of PXR and then a different pattern of interactions.

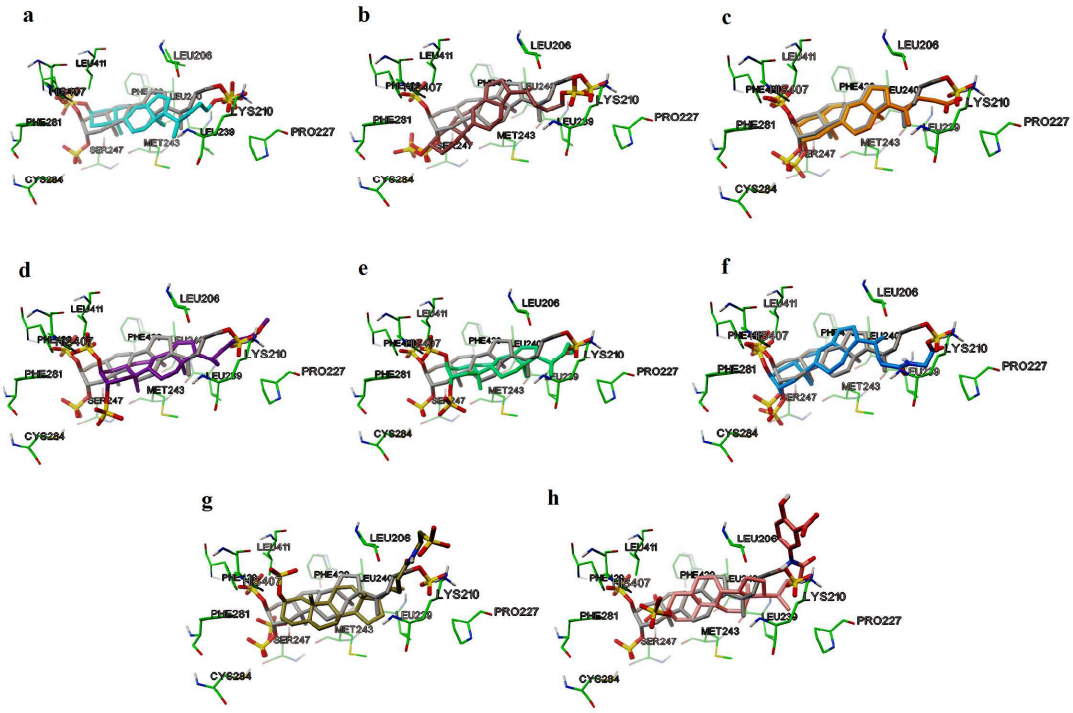


Figure 32. Superimposition between **1** (coloured by atom types: C grey, O red, S yellow) and: a) **25** (coloured by atom types: C sky-blue, O red, S yellow); b) **26** (coloured by atom types: C brown, O red, S yellow); c) **14** (coloured by atom types: C orange, O red, S yellow); d) **13** (coloured by atom types: C purple, O red, S yellow); e) **19** (coloured by atom types: C turquoise green, O red, S yellow); f) **16** (coloured by atom types: C dodger blue, O red, S yellow); g) **17** (coloured by atom types: C dark green, O red, S yellow); h) **18** (coloured by atom types: C pink, O red, S yellow) in PXR-LBD (residues are coloured by atom type: C green, H light grey, O red, N blue).

In summary, compound **23** is a robust PXR agonist that modulates immune response in human macrophages and liver fibrosis in epatocites. Because its simplified structure, compound **23** is a suitable candidate for further development in preclinical models of inflammatory diseases and in liver fibrosis induced by HIV infection. Further studies aimed to the evaluation of efficacy of **23** in animal

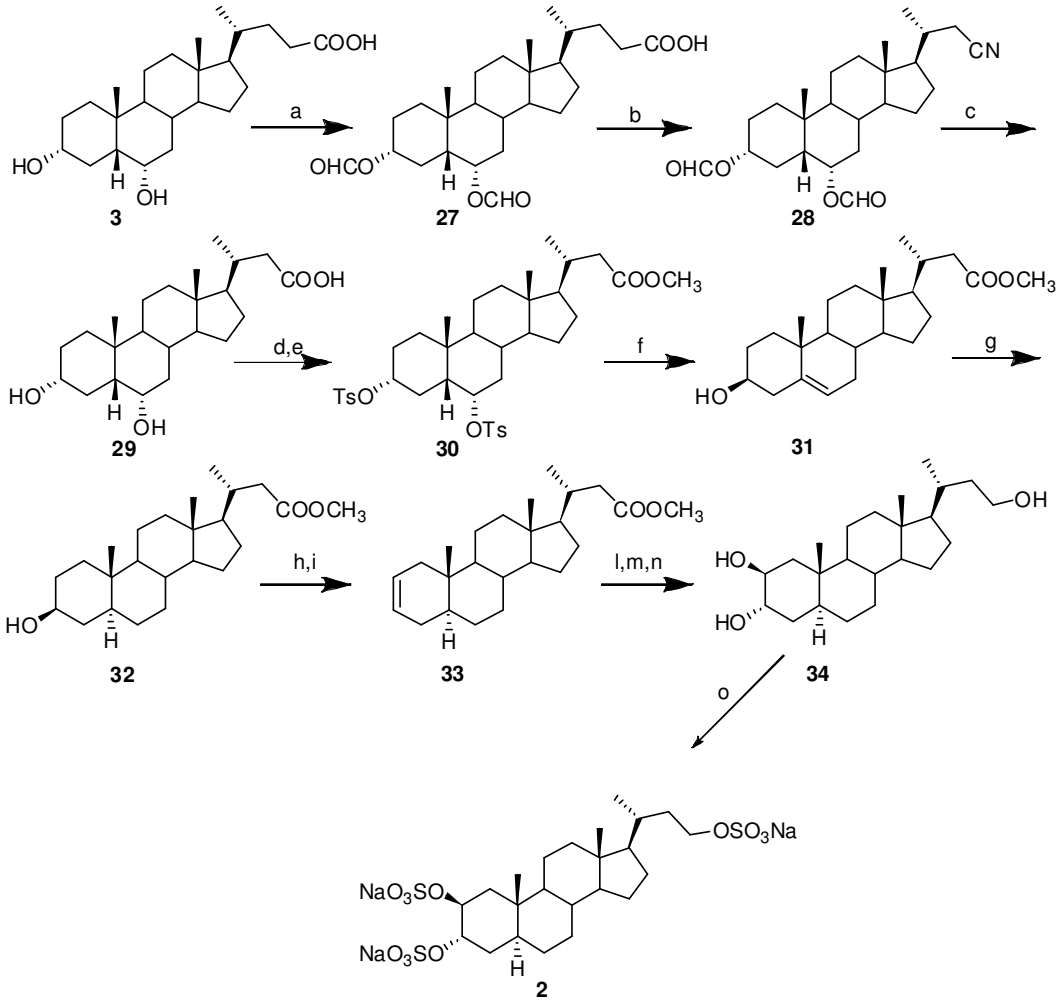
models, together with the determination of its chemical-physical proprieties are currently in progress.

2.3 Total synthesis of Solomonsterol B

Solomonsterol B (**2**) shares the same tetracyclic nucleus with **1**, but differs in the length of the side chain. Besides was proved that this modification has no influence on the binding within the LBD of PXR,[73] and therefore on the ability to transactivate PXR, recent reports have demonstrated that in several potential drugs the length of the side chain could exert dramatic effects. For example, *nor*-ursodeoxycholic acid, the C23 homologue of ursodeoxycholic acid (UDCA), has been shown to be more potent that the parent UDCA in pharmacological treatments for cholangiopathies and cholestatic liver diseases, demonstrating that its therapeutic effects are related to the side chain structure, which strongly influences the metabolism and consequently the pharmacokinetic behavior of this molecule.[109] Unfortunately, any further pharmacological in vivo experimentation or evaluation of the pharmacokinetic properties of solomonsterol B (**2**) was hampered by the scarcity of biological material isolated from the marine sponge. So we designed and realized the first total synthesis of solomonsterol B (**2**)[110] starting from commercially available hyodeoxycholic acid (**3**). The synthetic procedure also allowed the preparation of a derivative modified in the side chain, and thus a preliminary structure–activity relationship on the interaction between solomonsterol B and PXR was established. As depicted in Scheme 6, the key steps of our synthetic protocol are the one-carbon degradation at C24 and the modification of the functionalities of the A and B rings to establish the desired *trans* junction and the two hydroxyl groups at C2 and C3. Hyodeoxycholic acid (HDCA, **3**) was protected as performate derivative **27** by Fischer esterification

with formic acid, followed by acetic anhydride treatment to shift the equilibrium towards the complete formylation of **3**.[111] Intermediate **27** was subjected to the so-called second-order "Beckmann rearrangement" by treatment with sodium nitrite in a mixture of trifluoroacetic anhydride and trifluoroacetic acid.[112] Prolonged alkaline hydrolysis of resulting 23- nitrile intermediate **28** gave 24-*nor*-HDCA (**29**) in an isolated yield of 60% over the three-step sequence. Esterification of the carboxylic acid at C23 with methanol and *p*-toluenesulfonic acid (*p*TsOH), and tosylation of the resulting methyl ester with tosyl chloride in pyridine gave methyl 3α,6α-ditosyloxy-24-*nor*-5α-cholan-23-oate (**30**) in a satisfactory yield (75%, two steps). Heating **30** with CH_3COOK in refluxing DMF for 1 h resulted in simultaneous inversion at C3 and elimination at C6, with the formation of a mixture of **31** and its 3-*O*-acetyl derivative. Hydrolysis with *p*TsOH gave methyl 3-hydroxy-5-cholen-24-oate (**31**),[90,91] which in turn was hydrogenated to give **32**, with the required A/B *trans* ring junction.[92] Tosylation and elimination at C3 yielded the corresponding Δ2 ester **33** in 81% isolated yield after chromatographic purification on silica gel. The introduction of three hydroxyl groups in **34**, two on ring A in a *trans*-diaxial disposition and one in the side chain, was obtained by epoxidation of double bond, subsequent acid-catalyzed epoxide opening[94,95] with sulfuric acid in THF, and finally reduction of the methyl ester at C23 with $LiBH_4$ (56% yield over three steps). Sulfation of triol **34** gave the ammonium sulfate salt of solomonsterol B in 72% isolated yield over two steps, and this compound was transformed by ion exchange into desired target trisodium salt **2** and purified by reversed-phase solid extraction on a C18 cartridge. The complete match of optical rotation, NMR spectroscopic data and HRMS data of synthetic solomonsterol B (**2**) with that of the natural product confirmed the identity of the synthetic

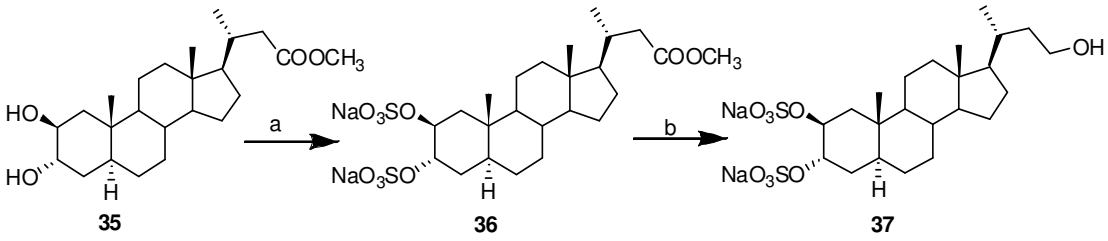
derivative. This synthesis was completed in a total of 13 steps, starting from commercially available hyodeoxycholic acid (**3**), in an overall yield of 10%. This route enabled us to prepare sufficient quantities of solomonsterol B (**2**) to be further characterized in pharmacological tests.



Scheme 6. Reagents and conditions. (a) HCOOH, $HClO_4$ 50 °C, then $(Ac)_2O$, 97%; (b) CF_3COOH, $(CF_3CO)_2O$, $NaNO_2$, 1 h, 0 °C, then 40 °C for 1.5 h, 80%; (c) 30% KOH, EtOH:H_2O 1:1, reflux, 78%; (d) *p*-TsOH, MeOH dry, 96%; (e) *p*-TsCl, pyridine, 78%; (f) CH_3COOK, DMF/H_2O 9:1, reflux; then *p*-TsOH, $CHCl_3$:MeOH 5:3, 80% over two steps; (g) H_2, Pd/C, THF/MeOH 1:1, room temperature, 84%; (h) *p*-TsCl, pyridine; (i) LiBr, Li_2CO_3, DMF, reflux, 81% over two steps; (l) *m*-CPBA, $CHCl_3$. room temperature; (m) H_2SO_4 1N, THF, room temperature, 81% over two steps; (n) $LiBH_4$, MeOH/THF, 0 °C, 69%; (o) $Et_3N{\cdot}SO_3$, DMF, 95 °C. Then Amberlite CG-120, MeOH, 72%.

Advanced intermediate **35** was also used as the starting material to obtain alcohol **37** (Scheme 7), which was judged instrumental to investigate the pharmacophoric role played by the side chain sulfate group in the PXR-agonistic activity of

solomonsterol B (**2**). As already discussed for SA, the replacement of the sulfate group at C23 of SB with a polar group, such as a hydroxy group as in **37**, could preserve the key interaction with Lys210, while at the same time introducing a functional group suitable for conjugation to a carrier for colon specific drug delivery. In fact, a PXR agonist, when administered by mouth in the pharmacological treatment of colon diseases (IBD, UC, CD), could undergo absorption before reaching the colon, and thus cause severe systemic side-effects resulting from the activation of PXR in the liver. Thus, as reported in Scheme 7, diol **35** was sulfated with triethylammonium–sulfur-trioxide complex, and transformed in sodium sulfate salt **36** by Amberlite treatment and purification on a C18 column. LiBH4 reduction gave C23 alcohol **37** in 78% yield from diol **35**.



35 **36** **37**

Scheme 7. Reagents and conditions: (a) Et3N·SO3, DMF, 95 °C; then Amberlite CG-120, MeOH; (b) LiBH4, MeOH/THF, 0 °C, 78% over two steps.

2.3.1 Pharmacological evaluation

Synthetic solomonsterol B (**2**) and alcohol **37** were tested in a luciferase reporter assay on a human hepatocyte cell line (HepG2 cells), transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-βgalactosidase, and p(CYP3A4)-TKLuc vectors (Figure 33).[113,114,115] Cells were then stimulated with rifaximin and with compounds **2** and **37** at a concentration of 10 μm each. As shown in Figure 33, solomonsterol B (**2**) was able to transactivate PXR with a potency comparable to rifaximin, whereas C23 alcohol **37** was inactive, thus demonstrating again the

pharmacophoric role played by the sulfate group in the side chain of solomonsterol B (**2**). Moreover compounds **2** and **37** were tested in a RT-PCR assays on the PXR target genes CYP3A4, CYP3A7, SULT2A1 and MDR1. As illustrated in Figure 34, with the exception of CYP3A7, solomonsterol B (**2**) was able to induce CYP3A4, SULT2A1 and MDR1, while, as expected, alcohol **37** failed to induce these PXR target genes on HepG2 cells.

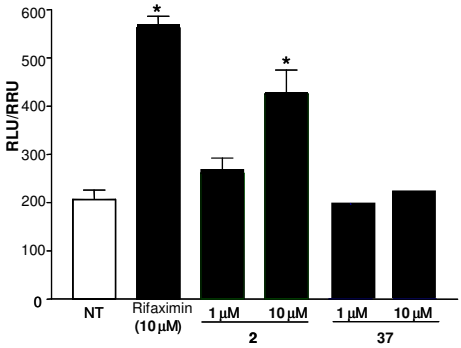


Figure 33. Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-βgalactosidase, and p(CYP3A4)-TK-Luc vectors, and stimulated for 18 h with rifaximin (10 μm), **2** (1 and 10 μm) and **37** (1 and 10 μm).

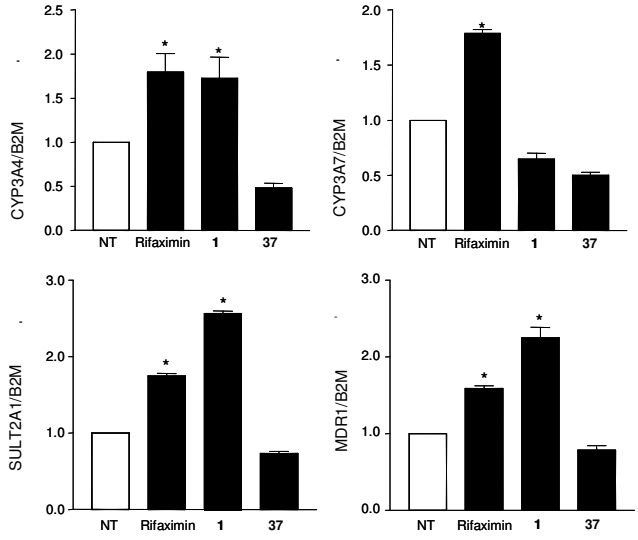


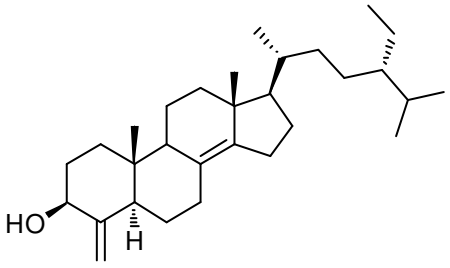
Figure 34. Real-time PCR analysis of PXR target genes CYP3A4, CYP3A7, SULT2A1 and MDR1 carried out on cDNA isolated from HepG2 not stimulated or primed with 10 μm rifaximin, **2** and **37**. Values are normalized relative to B2M mRNA, and are expressed relative to those of untreated cells, which are arbitrarily set to 1.

The results of luciferase experiments as well as PCR analysis of canonical PXR target genes clearly demonstrated that solomonsterol B (**2**) is a PXR agonist providing also information on the pharmacoforic role of C23 sulfate group.

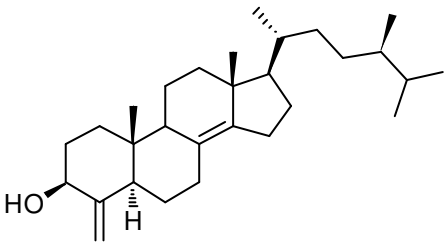
CHAPTER 3

DUAL PXR/FXR LIGANDS

Steroids bearing a 4-methylene group are relatively rare metabolites in nature. They have been exclusively isolated from sponges of the genus *Theonella*, mainly from *T. Swinhoei,* unaccompanied by conventional steroids and therefore proposed as ideal taxonomic markers for sponges of this genus.[116] Since the isolation, by Djerassi et al., of conicasterol and theonellasterol (Figure 35) from *T. conica* and *T. swinhoei*,[65] respectively, about twenty new 4-methylene-steroids were isolated from *Theonella* sponges.[117,118,119,120] Common structural features are a 24-methyl and/or 24-ethyl side chain, the presence of oxygenated functions at C(3), C(7), or C(15), of a $\Delta^{8,14}$ double bond rarely replaced by a 8(14)-seco-skeleton.



Theonellasterol (38)



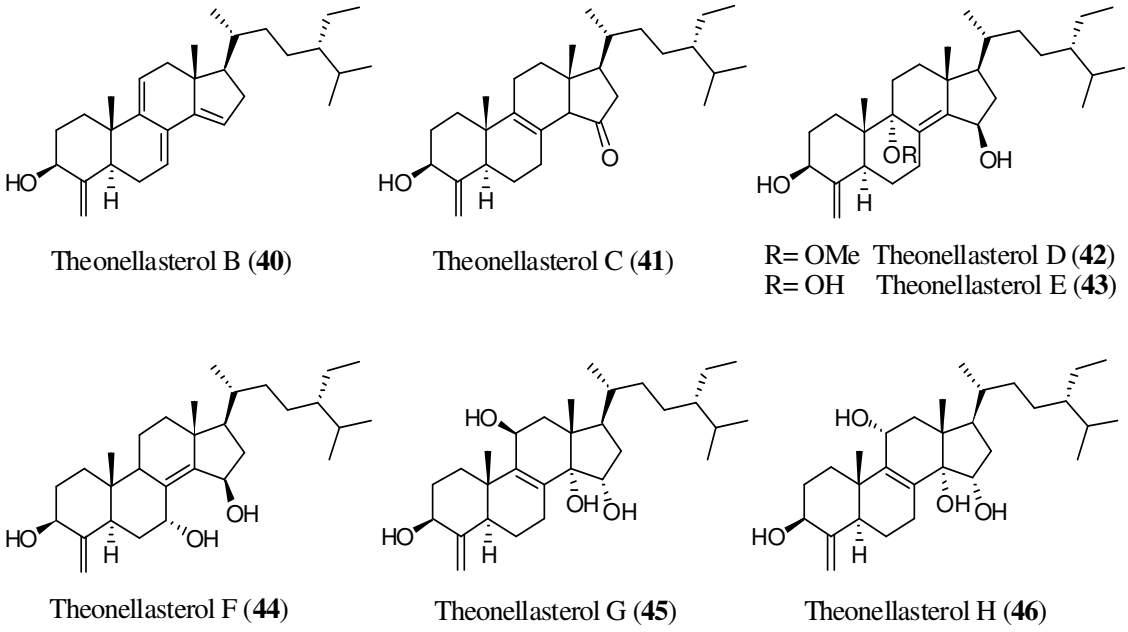
Conicasterol (39)

Figure 35. Theonellasterol and conicasterol previously isolated from *Theonella* species.

Pursuing our systematic study on the chemical diversity and bioactivity of secondary metabolites from marine organisms collected at Solomon Islands, we studied the less polar extracts, which resulted in the isolation and identification of theonellasterol[65] together with ten new polyoxygenated steroids, which we named theonellasterols B-H (**40**-**46**) and conicasterols B-D (**47**-**49**)[121] (Figures 36 and 38). These marine steroids are endowed with potent agonistic activity on the human pregnane-X-receptor (PXR) while antagonize the effect of natural ligands

for the human farnesoid-X-receptor (FXR). Exploiting these properties, we have identified theonellasterol G (**45**) as the first example of PXR agonist and FXR modulator from marine origin, that might have utility in treating liver disorders.

3.1 Structural determination of theonellasterols B-H



Theonellasterol B (**40**) Theonellasterol C (**41**) R= OMe Theonellasterol D (**42**)
R= OH Theonellasterol E (**43**)

Theonellasterol F (**44**) Theonellasterol G (**45**) Theonellasterol H (**46**)

Figure 36. Theonellasterols from *Theonella swinhoei.*

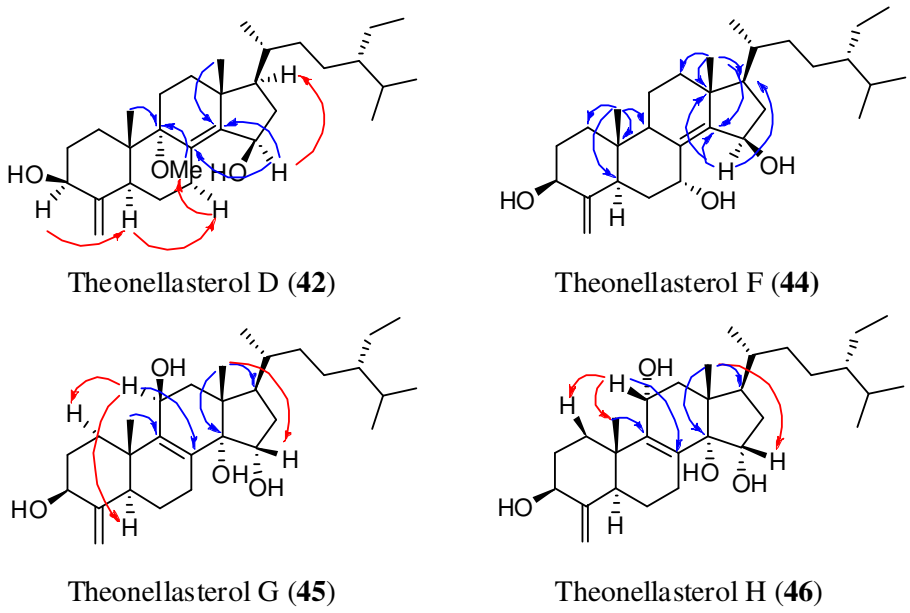
Theonellasterol B (**40**) was isolated as pale yellow oil. The molecular formula of $C_{30}H_{46}O$ was established by HR ESIMS based on the pseudo molecular ion $[M+Li]^+$ at *m/z* 429.3729 (calculated 429.3709), indicating eight degrees of unsaturation. The ^{1}H NMR spectrum of **40** (Table 2) showed signals characteristic of a 4-methylene-24-ethyl steroidal system: two methyl singlets (δ_H 0.89 and 0.98), three methyl doublets (δ_H 0.90, 0.92 and 1.01), one methyl triplet (δ_H 0.95), two broad singlets at δ_H 4.77 and 5.33, and one methine proton on an oxygenated carbon at δ_H 3.67. The low-field portion of the ^{1}H NMR spectrum also contained signals relative to three olefinic protons at δ_H 6.10 (1H, br d, J = 5.6 Hz, H-7), δ_H 5.85 (1H, br s, H-15) and δ_H 5.50 (1H, br d, J = 6.5 Hz, H-11). The ^{13}C NMR (Table 1) interpreted with the help of the HSQC experiment, evidenced a C30

steroidal theonellasterol skeleton with three cross conjugated trisubstituted double bonds [UV (MeOH): λmax (log ε) 275 nm (3.62)]. The localisation of the double bonds follows from the analysis of COSY and HMBC data. In particular, COSY correlations delineated the spin system H-1 through H-7, which included one hydroxyl group at C3, the exocyclic double bond at C4 and a double bond at C7 position. The olefinic proton H-7 at δ_H 6.10 showed a long range coupling with H-11 at δ_H 5.50 that was consistent with a double bond at the C9/C11 position, further supported by HMBC cross-peaks from Me-19 to C9 and from H-11 to C10. Finally, the last trisubstituted double bond was placed at C14/C15 on the basis of diagnostic HMBC cross-peak from Me-18 to C14. The configuration at C24 was determined by comparison of ^{13}C-NMR data (Table 1) with the epimeric steroidal side chain.[122,123] Therefore, theonellasterol B (**40**) was elucidated as 4-methylene-24(*S*)-ethylcholest-7,9(11),14-trien-3β-ol. The molecular formula of theonellasterol C (**41**) was determined to be $C_{30}H_{48}O_2$ by HR ESIMS with [M+Li]+ at *m/z* 447.3834 (calculated 447.3814), possessing seven degrees of unsaturation. Analysis of 2D NMR spectra clearly evidenced the same A ring and C24-ethyl side chain as in **40**.^{13}C NMR data (Table 1) indicated the presence of a tetrasubstituted double bond (δ_C 122.5, s and 138.6, s) and one carbonyl (δ_C 214.8). Proton signals due to H_2-7 (δ_H 2.15 and 2.51), H_2-11 (δ_H 1.82 and 2.04), and H-14 (δ_H 2.30) showed chemical shifts consistent with allylic hydrogens and a double bond at the C8/C9 position that was further supported by an HMBC cross-peak from Me-19 to C-9. The carbonyl group was placed at C15 on the basis of ^{13}C chemical shifts of C14 and C16 and of diagnostic HMBC correlations from H_2-16 and H-14 to C15. Hence, the structure of theonellasterol C (**41**) was determined as 3β-hydroxy-4-methylene-24(*S*)-ethylcholest-8(9)-en-15-one. The

molecular formula of theonellasterol D (**42**), $C_{31}H_{52}O_3$, indicated one carbon more than a conventional theonellasterol-like skeleton. ^{1}H and ^{13}C NMR data (Tables 1 and 2) evidenced a methoxy group (δ_H 3.07, δ_C 51.1), one tetrasubstituted double bond (δ_C 132.2, s and 153.7, s) and one secondary hydroxyl group (δ_H 4.60, δ_C 70.0). The tetrasubstituted double bond was placed at 8(14) position on the basis of diagnostic HMBC correlation (Figure 37) from Me-18 to C14 and confirmed by the long range correlations H-15 and H_2-7 to C8 and C14. Analogously, the HMBC correlations Me-19/C9 and OMe/C9 support the localization of the methoxy group at C9, whereas the 15-hydroxy function follows from $^{1}H,^{1}H$-COSY correlations. The ROESY correlations H-5α/H-7α, H-7α/OMe, implied that the methoxy group at C9 was α-oriented, which was further supported by the syn-axial γ-effects exerted by the methoxy substituent on C1, C5 and C7, and by the downfield shift of H-5. The strong ROE effect between H-15 and H-17α (Figure 37) indicated that the hydroxy group at C15 is in a β-position, as confirmed by the chemical shifts of the nuclei of ring D.[124] On the basis of foregoing analysis the structure of theonellasterol D (**42**) was determined as 4-methylene-9α-methoxy-24(*S*)-ethylcholest-8(14)-en-3β,15β-diol. The molecular formula of theonellasterol E (**43**) was determined as $C_{30}H_{50}O_3$, equivalent to theonellasterol D (**42**) less CH_2. Comparison with spectral data of **42**, chemical shift arguments (Tables 1 and 2), 2D NMR analysis indicated that theonellasterol E (**43**) differs from **40** for a 9α-hydroxy group replacing the 9α-methoxy group, disclosing a 4-methylene-24(*S*)-ethylcholest-8(14)-en-3β,9α,15β-triol structure for theonellasterol E (**43**). As determined by HR ESIMS measurements, theonellasterol F (**44**) is isomeric with **43**. From 1D NMR spectra it was possible to assign a tetrasubstituted double bond and two oxygen bearing methine groups.

The same $\Delta^{8,14}$-15β-hydroxy substructure as in theonellasterols D and E was inferred from 2D NMR analysis. COSY correlations delineated the spin system H-1 through H-7, the ^{13}C chemical shift for C7 (δ_C 66.7) implied an OH substitution. The relative stereochemistry at C7 was deduced from the small vicinal coupling constant of H-7, which is consistent with an equatorial disposition for this proton, thereby placing the hydroxyl group in an axial α-orientation. This was confirmed by the relatively low-field shifts of H-5 and H-9 caused by the 1,3-diaxial relationship of these protons to the hydroxy group at C7 (Table 2). Thus, theonellasterol F (**44**) was defined as 4-methylene-24(*S*)-ethylcholest-8(14)-en-3β,7α,15β-triol. Theonellasterol G (**45**) was isolated as an optically active pale yellow oil with a molecular formula of $C_{30}H_{50}O_4$. In addition to the signals relative to 4-methylene-3β-hydroxy structure, the ^{13}C NMR spectrum of **45** (Table 1) showed the presence of two quaternary sp^2 carbons, (δ_C 133.6 and 142.5), two oxygen-bearing methine carbons (δ_C 71.2 and 71.3) and one oxygen-bearing quaternary carbon (δ_C 89.7). The HMBC correlation between the angular methyl singlet Me-18 and a quaternary oxygenated carbon at δ_C 89.7 allowed to place an hydroxyl group at C14 position, whereas the correlation between the Me-19 and the olefinic carbon at δ_C 142.5 evidenced the presence of a $\Delta^{8,9}$ double bond. Moreover, the localization of the secondary alcoholic function at C11 follows from the HMBC correlations between H-11 at δ_H 4.39 and C8, C9 and C13 (Figure 37) and an hydroxyl group was placed at C-15 by COSY analysis. The relative configuration for the steroidal nucleus could be assigned by ROESY (Figure 37) and analysis of coupling constant data and on the basis of chemical shift considerations. The equatorial disposition of H-11 was evident from the small $^3J_{HH}$ vicinal coupling to H_2-12, and from the ROESY correlations to axial

protons H-1 and H-5. The strong downfield shift exhibited by the H-17 was indicative of the α configuration of the 14 hydroxy group, whereas the α orientation of 15-OH was assigned on the basis of ROESY correlation between H-15 and Me-18 (Figure 37) and confirmed by ^{13}C chemical shift pattern of ring D nuclei.[124] Thus theonellasterol G (**45**) was defined as 4-methylene-24(*S*)-ethylcholest-8(9)-en-3β,11β,14α,15α-tetraol. Theonellasterol H (**46**) showed the same molecular formula $C_{30}H_{50}O_4$ by HR ESIMS as **45**. Comparison of 1D NMR data of both compounds showed that they differed in ring C (Tables 1 and 2). When compared with **45**, the chemical shifts of C9, C11 and C14 in theonellasterol H (**46**) were significatively upfield shifted, whereas those of C10, C12 and C13 were downfield shifted. These data suggested that the pseudo-axial 11-OH group in **45** has been replaced by a pseudo-equatorial OH, as confirmed from ROESY correlations H-11/Me-19 and H-11/H-1β (Figure 37). As for **45**, the α orientation of 15-OH and 14-OH was established by ROESY correlation H-15/Me-18 and on the basis of H-17 resonance, respectively. Therefore theonellasterol H (**46**) was defined as 4-methylene-24(*S*)-ethylcholest-8(9)-en-3β,11α,14α,15α-tetraol.

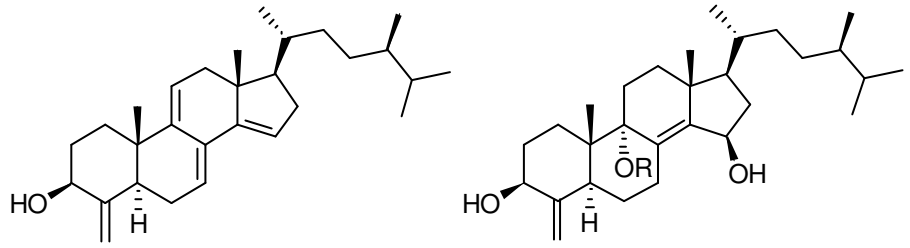


Theonellasterol D (**42**) Theonellasterol F (**44**)

Theonellasterol G (**45**) Theonellasterol H (**46**)

Figure 37. Key HMBC (blue arrows) and ROESY (red arrows) correlations for theonellasterols D (**42**), F (**44**), G (**45**) and H (**46**).

3.2 Structural determination of conicasterols B-D



Conicasterol B (**47**) R= OMe Conicasterol C (**48**)
R= OH Conicasterol D (**49**)

Figure 38. Conicasterols from *Theonella swinhoei*

Conicasterol B (**47**) was isolated as an optically active pale yellow oil. The molecular formula of $C_{29}H_{44}O$, equivalent to theonellasterol B (**40**) less CH_2, was established by high resolution mass spectrometry. The ^{1}H and ^{13}C NMR spectra of **47** were very similar to **40** (Tables 1 and 3), with the steroid nucleus being identical. The only difference between compounds **40** and **47** appears to be in the steroidal side chain, with a methyl group at δ_H 0.85 (3H, d, J = 7.0 Hz) replacing the C24 ethyl group in **40**, thus accounting for the difference in molecular formula between the two molecules. The stereochemistry of the side chain was determined

to be the same of conicasterol by comparison of ^{1}H[125] and ^{13}C chemical shifts.[122] Thus, conicasterol B (**47**) is 4-methylene-24(*R*)-methylcholest-7,9(11),14-trien-3β-ol. Inspection of NMR data of conicasterols C (**48**) and D (**49**) (Tables 1 and 3) clearly evidenced a strong resemblance with theonellasterols D and E, respectively. Once again the differences regarded the side chain with a methyl group (δ_H 0.88, 3H, d, J = 6.7 Hz in **48**; δ_H 0.89, 3H, d, J = 6.8 Hz in **49**) replacing the C24 ethyl group in compounds **42** and **43**. HRESIMS analysis gave definitively confirmation to the structures reported in Figure 38 as a 4-methylene-9α-methoxy-24(*R*)-methylcholest-8(14)-en-3β,15β-diol for conicasterol C (**48**) and a 4-methylene-24(*R*)-methylcholest-8(14)-en-3β,9α,15β-triol for conicasterol D (**49**).

Table 1. ^{13}C NMR Spectroscopic Data (175 and 125 MHz, C_6D_6) for compounds **40**-**49**.[a]

	Compounds									
C	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**
1	35.4	34.9	29.6	30.1	36.9	34.3	34.6	35.4	29.5	30.0
2	33.4	32.9	33.3	33.3	33.5	32.9	33.0	33.3	33.3	33.4
3	73.0	72.9	73.2	73.1	73.2	72.9	72.8	73.1	73.1	73.1
4	154.0	152.9	154.3	154.2	153.4	152.4	152.7	154.2	154.3	154.1
5	43.9	46.9	41.6	41.7	43.2	46.4	46.2	44.0	41.6	41.7
6	26.3	21.3	25.3	24.8	31.6	20.8	20.7	26.2	25.3	24.7
7	121.1	29.9	27.2	27.1	66.7	25.9	26.9	121.1	27.2	27.2
8	128.4	122.5	132.2	135.4	136.7	133.6	138.4	128.3	132.1	135.3
9	142.5	138.6	77.8	74.5	45.6	142.5	141.4	142.7	77.8	74.5
10	38.5	40.1	45.0	43.7	40.0	37.6	38.7	38.3	45.0	43.8
11	118.6	21.9	26.5	29.2	20.3	71.2	64.9	118.6	26.5	29.1
12	42.7	32.3	37.5	34.5	38.1	39.9	45.0	42.7	37.4	34.4
13	45.6	41.0	43.6	44.2	43.8	45.0	49.8	45.5	43.4	44.2
14	148.3	65.7	153.7	150.2	151.5	89.7	83.7	148.3	153.7	150.3
15	118.0	214.8	70.0	69.8	70.2	71.3	71.6	118.1	70.0	69.8
16	36.2	40.7	39.4	39.4	39.5	40.0	39.6	36.2	39.4	39.4
17	58.6	42.6	54.3	53.7	53.5	51.8	52.0	58.4	54.2	53.7
18	18.0	22.6	18.1	18.2	19.9	20.6	19.9	18.0	18.1	18.3
19	19.8	18.3	17.1	16.9	12.9	18.6	18.0	19.8	17.1	17.0
20	35.0	35.4	34.6	34.5	34.6	36.9	36.9	34.9	35.0	34.9
21	19.2	19.3	19.3	19.5	19.4	18.8	18.6	19.3	19.1	19.2
22	34.3	33.7	34.2	34.3	34.2	34.1	34.1	33.8	33.8	33.8
23	26.7	27.5	26.8	26.7	26.8	27.4	27.3	30.6	30.6	30.6
24	46.6	46.3	46.6	46.6	46.7	46.6	46.5	39.2	39.2	39.2
25	29.3	29.3	29.2	29.4	29.5	29.6	29.3	32.7	32.7	32.7
26	19.3	19.3	19.3	19.3	19.3	19.4	19.2	18.3	18.5	18.4
27	19.7	19.7	19.7	19.8	19.6	20.1	19.9	20.3	20.3	20.3
28	23.5	23.4	23.5	23.5	23.5	23.6	23.4	15.5	15.5	15.6
29	12.7	12.6	12.6	12.6	12.6	12.7	12.6	104.1	103.5	103.7
30	104.1	103.3	103.7	103.8	103.3	103.4	103.3			
OMe			51.1						51.0	

[a] ^{13}C assignments aided by COSY, HSQC and HMBC experiments.

Table 2. Selected ^{1}H NMR Spectroscopic Data (700 and 500 MHz, C_6D_6) for compounds **40**-**46.**

H	**40**	**41**	**42**	**43**	**44**	**45**	**46**
3	3.67 m	3.71 dd (5.4, 11.0)	3.89 dd (5.3, 11.2)	3.90 m	3.79 m	3.73 dd (5.7, 11.0)	3.79 m
5	1.95 m	1.74 m	2.89 br d (12.4)	2.78 br d (12.3)	2.42 m	1.85 m	1.78 m
7	6.10 br d (5.6)	2.15 dd (5.6,18.1) 2.51 m	2.16 ddd (5.3,8.2, 13.6) 2.52 m	2.41 m 2.53 ddd (1.8,4.8, 14.5)	4.63 br t (2.9)	2.02 dd (5.7,18.3) 2.61 m	2.10 dd (5.8,18.3 2.47 m
9	-	-	-	-	2.41 m	-	-
11	5.50 br d (6.5)	1.82 m 2.04 m	1.60 m 1.75 m	1.55 m 1.81 m	1.54 m	4.39 br t (2.9)	4.07 d (7.9)
15	5.85 br s	-	4.60 br t (4.4)	4.50 br d (6.0)	4.58 br d (4.8)	4.47 dd (6.0, 9.6)	4.19 dd (4.9,9.6)
17	1.73 m	2.00 m	1.64 m	1.62 m	1.64 m	2.40 m	2.35 m
18	0.98 s	0.81 s	0.73 s	0.73 s	0.79 s	0.43 s	0.49 s
19	0.89 s	0.82 s	0.72 s	0.71 s	0.59 s	0.59 s	0.66 s
21	1.01 d (6.4)	0.94 d (6.6)	1.01 d (6.6)	1.04 d (6.5)	1.02 d (6.2)	0.88 d (6.4)	0.90 d (6.4)
26	0.90 d (7.0)	0.84 d (6.9)	0.91 d (6.8)	0.93 d (6.9)	0.93 d (6.2)	0.91 d (6.8)	0.91 d (6.8)
27	0.92 d (7.0)	0.87 d (6.9)	0.93 d (6.8)	0.94 d (6.9)	0.94 d (6.2)	0.94 d (6.8)	0.93 d (6.8)
28	1.22 m 1.39 m	1.14 m 1.33 m	1.23m 1.41 m	1.24 m 1.42 m	1.23 m 1.41 m	1.22 m 1.42 m	1.22 m 1.42 m
29	0.95 t (7.4)	0.89 t (7.4)	0.96 t (7.4)	0.96 t (7.3)	0.96 t (7.4)	0.96 t (7.3)	0.97 t (7.4)
30	4.77 br s 5.33 br s	4.69 br s 5.24 br s	4.75 br s 5.36 br s	4.76 br s 5.36 br s	4.73 br s 5.33 br s	4.67 br s 5.28 br s	4.71 br s 5.30 br s
OMe			3.07 s				

[a]Coupling costants are in parentheses and given in hertz. ^{1}H assignments aided by COSY experiments.

Table 3. Selected ^{1}H NMR Spectroscopic Data (700 and 500 MHz, C_6D_6) for compounds **47-49.**

H	**47**	**48**	**49**
3	3.65 m	3.89 dd (5.3, 11.3)	3.89 m
5	1.92 m	2.89 br d (12.4)	2.73 br d (12.3)
7	6.08 br d (5.6)	2.15 ddd (5.3, 8.2, 13.6) 2.51 m	2.42 m 2.55 ddd (1.8, 4.7, 14.6)
9	-	-	-
11	5.49 br d (6.5)	1.60 m 1.75 m	1.56 m 1.83 m
15	5.84 br s	4.61 br t (4.3)	4.51 br d (6.0)
17	1.73 m	1.65 m	1.62 m
18	0.97 s	0.73 s	0.73 s
19	0.87 s	0.72 s	0.70 s
21	0.97 d (6.4)	0.98 d (6.3)	1.01 d (6.6)
26	0.85 d (7.0)	0.88 d (6.5)	0.89 d (6.8)
27	0.90 d (7.0)	0.93 d (6.5)	0.93 d (6.8)
28	0.85 d (7.0)	0.88 d (6.7)	0.89 d (6.8)
29	4.77 br s 5.33 br s	4.74 br s 5.35 br s	4.76 br s 5.36 br s
OMe		3.07 s	

[a]Coupling costants are in parentheses and given in hertz. ^{1}H assignments aided by COSY experiments.

3.2.1 Pharmacological evaluation

We have investigated whether this family of hydroxylated sterols might act as modulators of two well characterized nuclear receptors, the farnesoid-X-receptor (FXR) and pregnane-X-receptor (PXR), highly expressed in the mammalian livers. For this purpose compounds **40**-**47** and compound **49** were challenged in a reporter gene assay using HepG2 cells, a human hepatoma cell line. As illustrated in Figure 39A, all compounds, except theonellasterol G (**45**), that partially activated FXR, failed to activate FXR at the concentrations of 10 μM. By contrast, all compounds, with the exception of theonellasterols D (**42**) and H (**46**), effectively antagonized FXR transactivation induced by CDCA. It is noteworthy that theonellasterol G (**45**) behaves as an antagonist in the presence of CDCA but

is able to partially transactivate FXR, indicating that this agent is an FXR modulator.

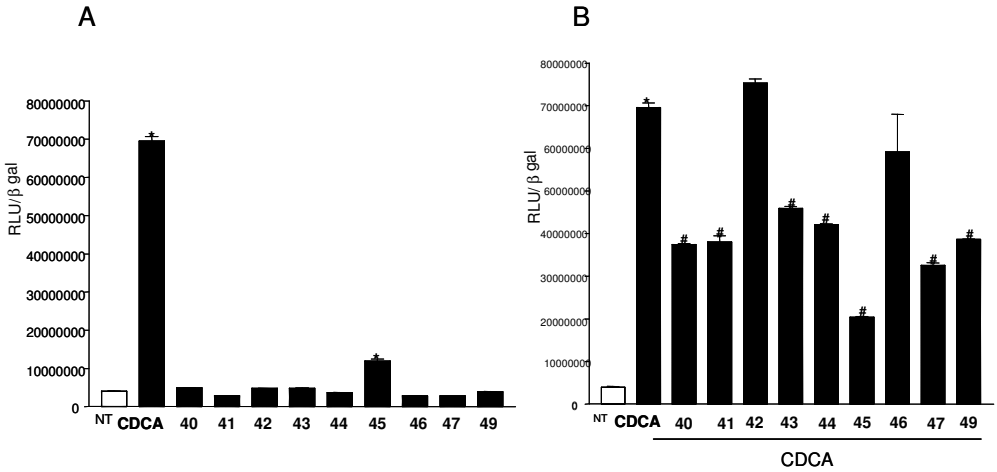


Figure 39. Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-FXR, pSG5-RXR, pCMV-βgalactosidase, and p(HSP27)-TK-Luc vectors and stimulated 18 h with (A) CDCA (10 μM) and compounds **40**-**49** (10 μM). (B) CDCA (10 μM) alone or in combination with compounds **40**-**49** (50 μM). *P < 0.05 versus not treated (NT). **P < 0.05 versus CDCA (n = 4).

As shown in Figure 40A, at the concentration of 10 μM, all these compounds were PXR agonists. Interestingly, also theonellasterol G (**45**) acted as a PXR ligand, suggesting that this compound might be considered the first FXR modulator and PXR ligand so far identified.

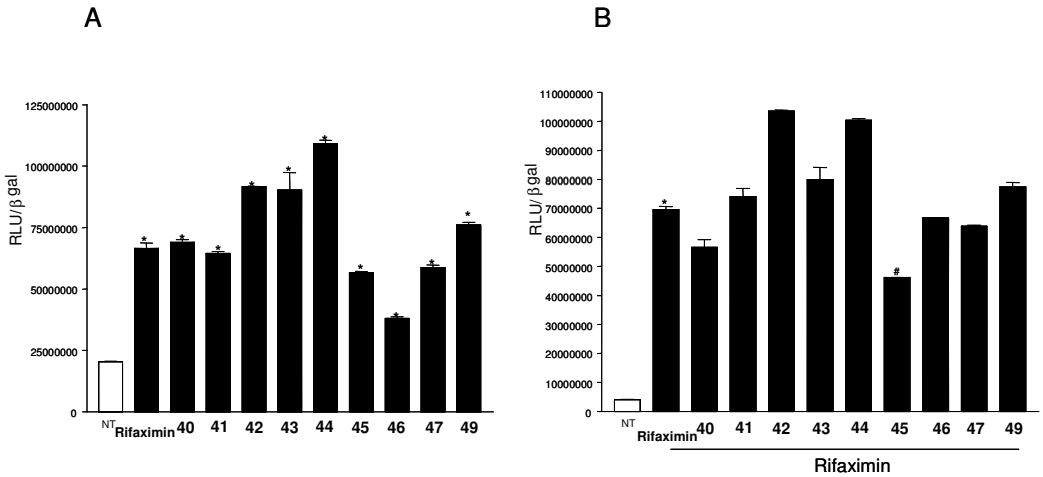


Figure 40. Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-bgal, p(cyp3a4)TKLUC vectors and stimulated 18 hours with (A) rifaximin, 10 μM, and of compounds **40**-**47** and compound **49**, 10 μM; and (B) rifaximin, 10 μM, alone or in combination with compounds **40**-**47** and compound **49**, 50 μM. *P<0.05 versus not treated (NT). #P<0.05 versus rifaximin (n=4).

Indeed, when incubated with human hepatocytes, theonellasterol G (**45**) increased the expression of OSTα, a well characterized FXR target, along with SULT2A1 and MDR1, two well characterized PXR responsive genes (Figure 41). Thus, in agreement with transactivation experiments, the PCR data demonstrated that theonellasterol G (**45**) is an FXR modulator and PXR agonist.

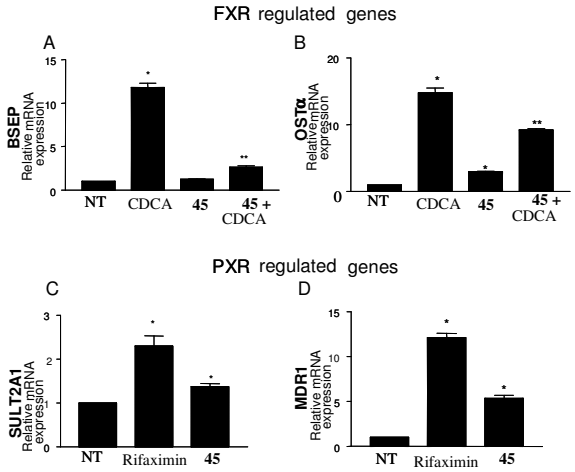


Figure 41. Real-Time PCR analysis of (A) BSEP; (B) OSTα; (C) SULT2A1 and (D) MDR1 mRNAs expression carried out on cDNA isolated from HepG2 cells not treated (NT) or primed with 10 μM CDCA alone or in combination with 50 μM of **45** (A-B) or treated with rifaximin 10 μM (C-D). *P<0.05 versus not treated. **P<0.05 versus cells treated with CDCA.

3.2.2 Docking studies

Based on the above mentioned results, we have then analyzed, by means of molecular docking calculations, the interactions of polyhydroxylated steroids **40-47** and **49** with human FXR and PXR, in order to generate a structure-activity relationship and obtain informations on their binding mode at atomic level. All the calculations were run by Autodock4.2 software.[109] As shown in the Figure 42, compounds **40-47** and **49** adopt the same positioning in the FXR binding site when compared to 6-ECDCA. Even if the junction between A/B rings is *trans* and the OH group at position 3 is in the β position with respect to the co-crystallized molecule, for all the complex models relative to compounds **40-47** and **49**, the fundamental hydrogen bond contacts with the two amino acids of the catalytic

triad[38] (namely Tyr358 in helix 7, His444 in helix 10/11) are maintained. Consequently, it is possible to single out two different docking poses for these molecules: a) the first one regards compounds containing the OH group at position 3 and/or the OH group at position 15 in a *trans* relationship (**40, 41, 45-47**, Figure 42A); b) the second family includes compounds having the 3- and 15-OHs in a relative *cis* arrangement (**42-44**, and **49**, Figure 42B). Nevertheless for all compounds **40-47** and **49**, the steroid nucleus is able to accommodate in the ligand binding site, establishing hydrophobic interactions with the cavity pocket formed between the Helix 2, 3, 5-7, and 10/11 (Figure 42).[38] Concerning the interactions regarding other positions, the α-OH at position 9 of **43** and **49** forms a hydrogen bond with Ser329 (Helix 5) and this interaction is lost when the hydroxyl group is replaced by a methoxy group, as in compound **42**. Moreover, OH groups at 14α and 15α positions (compounds **45** and **46**, respectively) are able to establish hydrogen bonds with Ser329 (Helix 5), whereas OH group at position 15β (compounds **42**-**44**, and **49**) interacts with Leu284 (Helix 3). The carbonyl group at position 15 (compound **41**) and the α-OH group at position 7 (compound **44**) do not exert further polar interactions with the FXR ligand binding site.

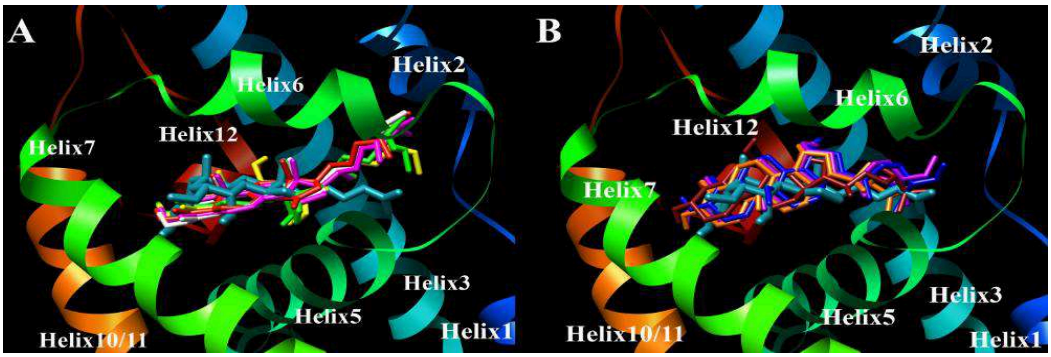


Figure 42. The two different spatial arrangements of polyhydroxylated steroids **40-47** and **49** in the binding site of FXR (chain A of crystal structure 1OSV). A) Superposition of 6-ECDCA (light blue) with compounds **40** (pink), **41** (red), **45** (yellow), **46** (green), and **47** (white). B) Superposition of 6-ECDCA (light blue) with compounds **42** (blue), **43** (purple), **44** (dark red), and **49** (orange).

Regarding the side chain, the presence of a methyl group at position 24 for **45** and **49** allows stronger interactions with amino acids Met287 (Helix 3), Met262 (Coil 2), and His291 (Helix 3), present on the external part of the target molecular surface, with respect to compounds **40** and **43** bearing an ethyl group with different configuration (Figure 43C). As previously reported,[126-128] the activation of FXR receptor by bile acids and bile acid analogs is affected by simultaneous presence of two α-OH at position 3 and 7 and by a good balance between hydrophobic and hydrophilic substituents at the α and β face of the nucleus. This kind of interactions are missing for our compounds, and in fact all compounds, except theonellasterol G (**45**), failed in the activation of FXR at the concentration of 50 μM (see biological section). In particular, as depicted in Figure 43A, the β-OH group at position 11 of compound **45** is involved in an additional hydrogen bond with Leu284 (Helix 3) with respect to its epimer **46** (Figure 43C), where this interaction is lost. Moreover, even if also **43** and **49** are able to interact with Ser329 (Helix 5) and Leu284 (Helix 3), they present a different spatial arrangement (Figures 42 and 43C) with respect to theonellasterol G (**45**), lacking some hydrophobic interactions with the amino acids of FXR ligand binding pocket (e.g. Trp466, Ile270, Thr267, Leu345, Figure 43). So, our docking calculations point out that the simultaneous interactions[129] of **45** with Helix 7 (Tyr358) and Helix 3 (Leu284) and its optimal hydrophobic contacts with LBD compared to compounds **46** (only interacting with Helix 7), and **43** and **49** respectively (Figures 41 and 42), cause a great difference in activity between these sterols, suggesting that a correct orientation of the OH group at position 11 and the hydrophobic contacts with the receptor are critical for the FXR modulator activity.

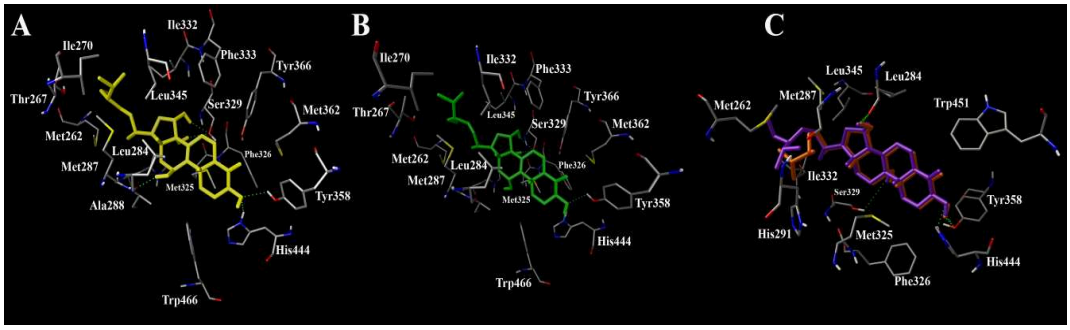


Figure 43. Three dimensional models of docking pose of **45** (**A**, yellow), **46** (**B**, green), **43** (**C,** purple) and **49** (**C**, orange) with FXR (see text for details).

For what concerns the PXR[130] agonist activity, the first observation suggested by molecular docking analysis regards the positioning of compounds **40**-**47** and **49** in the PXR-LBD. This region is formed by hydrophobic (Phe251, Phe288, Phe429, Cys284, Leu206, Leu209, Leu324, Leu411, Met243, Met243, Met425, Trp299), polar and charged (Ser247, Ser208, His407, Asp205, Arg410, Gln285) amino acids; particularly crucial is the interaction of our sterols with the Ser247, previously reported to be involved in key interactions for the PXR agonist activity. In fact, this interaction is associated to the activity modulation observed in our series of compounds that is relative to substitution pattern of steroidal nucleus. It is noteworthy that 15α-OH, 15β-OH and 15-keto substitutions are all possible H-bond acceptor sites interacting with H-bond donor Ser247. The further interactions of OH group at position 7 with Ser247 for compound **44**, of OH group at position 9 with Gln285 for **43**, and of OMe group at position 9 with Cys284 for **42**, may be responsible for the increased activity observed for these derivatives (Figure 44). Finally, compounds **47** and **49**, featuring a methyl group at position 24, are endowed with a weaker agonist activity with respect to **40** and **43**, bearing an ethyl group with different configuration. Even though compounds **47** and **49** are able to interact by their side-chain with Leu209, our models suggest

that their decreased activity is due to lacking interactions with Phe420, Leu411, and Phe429 (Figure 44).

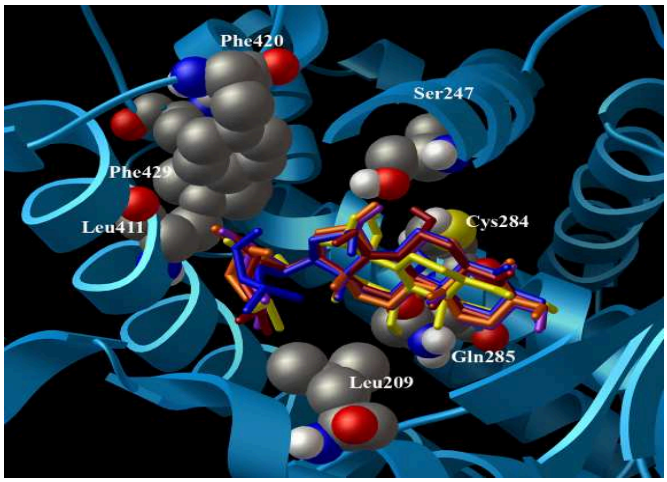


Figure 44. Three dimensional model of the most representative polyhydroxylated steroids **42** (blue), **43** (purple), **44** (dark red), **45** (yellow) and **49** (orange) with the PXR binding pocket.

In summary, a novel class of FXR/PXR modulators of marine origin has been discovered. Because PXR is a gene orchestrating the metabolism of xeno- and endo-biotics by liver and intestinal epithelial cells,[9] and, usually, PXR ligands are not ligand for FXR, the discovery of an ancestral dual activator highlights the potential for the existence of similar ligands in the mammalian body. From the pharmacology point of view, a dual ligand holds potential in the treatment of liver disorders characterized by cholestasis and/or impaired metabolism of xenobiotics. Of relevance, because both FXR and PXR exert anti-inflammatory effects in the intestine, a dual ligand holds potential in the treatment of IBDs.

3.3 Analysis of the third specimen of *Theonella swinhoei*

In a recent report Crews *et al.*[61] pointed on the existence of at least three phenotypes of *Theonella swinhoei* [131], and observed that the morphology of the sponge has some influence on the chemical composition. In the course of my PhD experimental work, I have the opportunity of study three *Theonella swinhoei* specimens. Two of them, whose macroscopic morphology could be ascribed to

phenotype I, have been subjected by us to extensive chemical investigation, previously described, that disclosed those specimens as invaluable sources of new secondary metabolites, peptides, macrolides etc. On the other hand, when we investigated the apolar extracts, we found, in all three specimens, a great variety of polyhydroxysteroids all characterized by a 4-methylene functionality. In particular 4-methylene-steroids[122,132] from phenotype I specimens mainly possessed a 24-ethyl side chain (Figure 36) and have been demonstrated modulators of FXR and PXR (see above). The analysis of the apolar extract of the third specimen evidenced again the presence of a family of polyhydroxy steroids but in this case most of them featuring a 24-methyl side chain (Figure 45, conicasterols G-K, **50**-**54**), including known conicasterol (**39**)[65] with its 7- and 15-hydroxy derivatives (**57** and **58**),[120] dehydroconicasterol (**56**),[117] and the 8-14-seco, swinhosterol B (**59**)[133] (Figure 46). Trace of a new theonellasterol-like derivative, theonellasterol J (**55**), was also isolated.[134]

3.3.1 Structural determination of 4-methylenesterols (phenotype III).

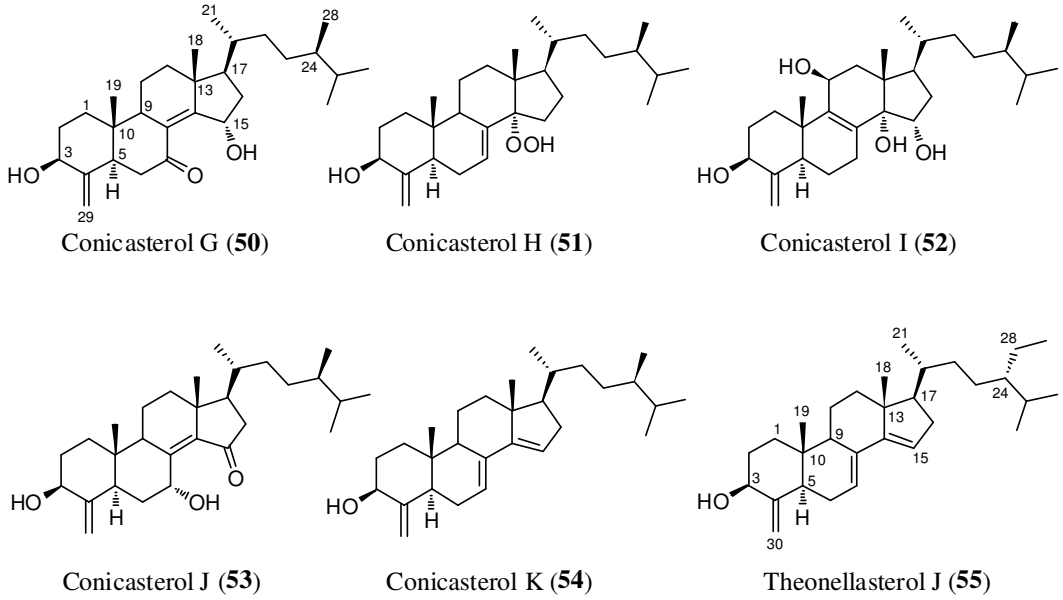


Figure 45. New 4-methylenesterols isolated from *Theonella swinhoei*, phenotype III.

A molecular formula $C_{29}H_{46}O_3$ was deduced from HR ESIMS (*m/z* 449.3611 $[M+Li]^+$) and ^{13}C NMR data for conicasterol G (**50**). Seven degree of unsaturation implied by the molecular formula were ascribed to four rings, one C=O bond (δ_C 203.6), also supported by strong IR adsorption at ν=1710 cm^{-1}, one tetrasubstituted double bond (δ_C 167.5 and 130.9) and the usual methylidene functionality at C-4 (δ_C 150.6 and 104.1). The low field region of the ^1H NMR spectrum contained, in addition to a double doublet at δ_H 3.59 assigned to the 3α-proton, a broad doublet proton signal (δ_H 4.76) ascribable to a secondary carbinol. The carbonyl group was placed at C-7 on the basis of HMBC correlations (Figure 46) from H_2-6 to C-7 (δ_C 203.6), and of diagnostic chemical shift values and multiplicity of resonances of H_2-6 [δ_H 2.16 (t, *J* = 16.1 Hz) and 2.37 (dd, *J* = 16.1, 12.3 Hz). The placement of a conjugated double bond at 8(14) position was inferred by HMBC correlation (Figure 46) from Me-18 to C-14 (δ_C 167.5) and supported by ^{13}C chemical shifts of C-9 and C-14. Finally the presence of a secondary hydroxy group at C-15 follows from $^1H,^1H$-COSY correlations and from the key HMBC correlation H-15 (δ_H 4.76)/C-17 (δ_C 53.8). As previously reported,[122] the configuration of the 15-OH (as depicted in Figure 45) was established by downfield shift exhibited by H-17 (δ_H 2.07). Conicasterol H (**51**) was obtained as a white amorphous powder. The molecular formula of **51** was determined as $C_{29}H_{48}O_3$ by HR ESIMS (*m/z* 451.3770 $[M+Li]^+$), 32 mass higher than parent conicasterol (**39**). Respect to **39**, the ^1H NMR spectrum showed the presence of an additional olefinic proton signal at δ_H 5.52. This proton showed, in the ^1H-^1H COSY spectrum ^3J correlations with H_2-6 at δ_H 1.85 and 1.99 and a long range coupling with H-9 at δ_H 2.06, consistent with a trisubstituted double bond at the C-7/C-8 position. The presence in the ^{13}C NMR spectrum of only one

oxygen-bearing quaternary carbon, its low-field resonance (δ_C 97.8) and mass data indicated the presence of an hydroperoxy group in the molecule. As shown in Figure 46 the placement of this group at C-14 was inferred by diagnostic HMBC correlations and also confirmed by chemical shift values of ring C and D nuclei (Table 4). The downfield shift exhibited by H-17 (δ_H 2.09) was indicative of the α configuration of the 14 hydroperoxy group. The ^{1}H NMR spectrum of conicasterol I (**52**, $C_{29}H_{48}O_4$, as determined by HRESIMS) is almost superimposable to those of theonellasterol G[122] with the exception of the 0.4-1.0 ppm region where a methyl doublet [δ_H 0.88 (d, J= 6.8 Hz)] replaced the methyl triplet of a 24-ethyl theonellasterol-like side chain. Conicasterol J (**53**) was isomeric with conicasterol G (**50**). As for **50**, NMR analysis revealed the presence of a tetrasubstituted double bond (δ_C 149.5 and 142.4), one carbonyl carbon (δ_C 205.7) and one oxygen bearing methine (δ_C 60.9 and δ_H 6.27). COSY correlations delineated the spin system H-1 through H-7, in which the chemical shift of H-7 (δ_H 6.27) implied OH substitution, whereas HMBC correlations H_3-18 to C-14 (δ_C 142.4), H-9 to C-8 (δ_C 149.5) and C-14 (Figure 46) pointed out the location of double bound at 8(14) position. Finally, a conjugate carbonyl was placed at C-15 on the basis of HMBC cross-peaks between H_2-16 (δ_H 1.94 and 2.38) and the carbon resonance at δ_C 205.7 and also confirmed by chemical shift of C-16 (δ_C 42.3). The unusual downfield shift exhibited by the carbinol proton at C-7 position (δ_H 6.27) could be explained involving the anisotropic effect of carbonyl group at C-15 position, as already reported for H-7β in 9α-hydroxy-15-oxoconicasterol, previously isolated from a Chinese collection of *Theonella swinhoei*[116.] Following this consideration, the α configuration of the hydroxyl group at C-7 was established and confirmed by small vicinal coupling constant of

H-7 (br t, J = 2.9 Hz) and the relatively low-field shifts of H-5 (δ_H 2.46) and H-9 (δ_H 2.57). The molecular formula of conicasterol K (**54**) was determined to be $C_{29}H_{46}O$, two hydrogen less respect to conicasterol (**39**). NMR data revealed the presence of two trisubstituted double bonds (Table 4). The placement of a Δ^7 double bond derived from COSY analysis that disclosed the spin system from H-1 to H-7, in which the chemical shift of H-7 proton was indicative of its olefinic nature (δ_H 5.45, δ_C 102.7). The second double bond was placed at C-14 on the basis of HMBC correlations (Figure 46) from H-9, Me-18 and H_2-16 with carbon resonance at δ_C 163.1 (C-14). Definitive confirmation to the structure **54** reported in Figure 45 derived from chemical shift values of the protons at C-16 (δ_H 1.87 and 2.30) and their COSY correlations with the olefinic proton at δ_H 5.27 (H-15). NMR (Tables 4 and 5) and mass data clearly evidenced that theonellasterol J (**55**) shared the same tetracyclic nucleus as conicasterol K (**54**), while featured a 24-ethyl side chain. The 24*R* stereochemistry of the side chains of conicasterols G-K (**50**-**54**) and 24*S* of theonellasterol J (**55**) were determined by comparison of ^1H[126] and ^{13}C chemical shifts[123].

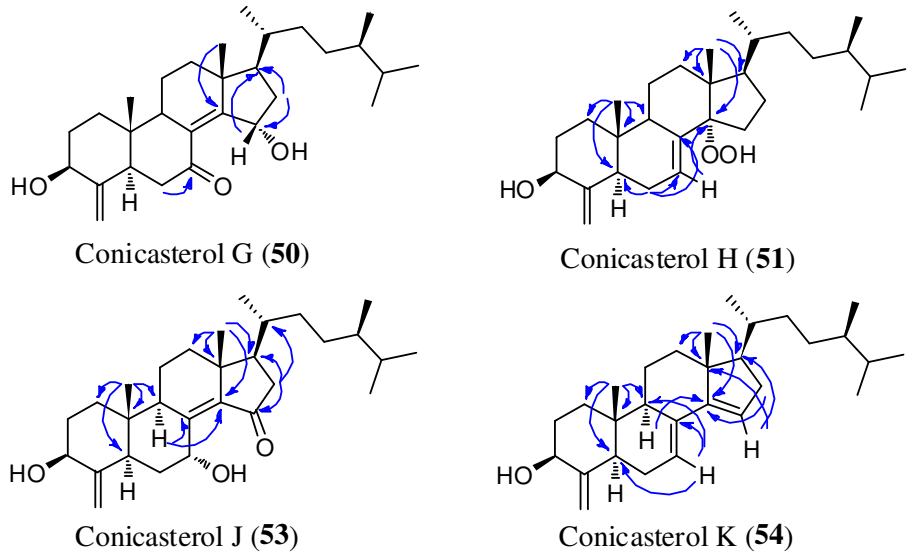
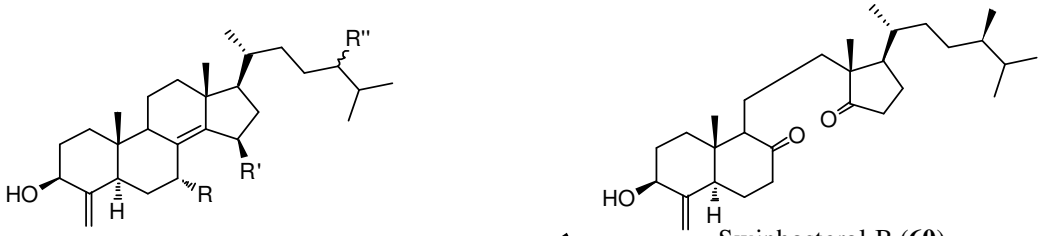


Figure 46. Key HMBC correlations for conicasterols G (**50**), H (**51**), J (**53**) and K (**54**).

The structures of compounds **39**, **56**-**59** as conicasterol, dehydroconicasterol, 7α-hydroxyconicasterol, 15β-hydroxyconicasterol, and swinhosterol B, respectively, were deduced by NMR and mass analysis and by direct comparison of their chemical shift data with existing literature values (Figure 47). [65,117,120,134]



Swinhosterol B (**60**)

Conicasterol (**56**) R=H R'=H R''=
Dehydroconicasterol (**57**) R=H R'=H R''=
7a-hydroxyconicasterol (**58**) R=OH R'=H R''=
15b-hydroxyconicasterol (**59**) R=H R'=OH R''=

Figure 47. Known 4-methylenesterols isolated from *Theonella swinhoei*, phenotype III.

Table 4. ^{13}C NMR data for compounds **50**-**55** (175 and 125 MHz, C_6D_6).[a]

C	**50**	**51**	**52**	**53**	**54**	**55**
1	36.1	37.4	34.4	36.8	36.8	36.9
2	33.1	33.2	32.9	33.1	33.3	33.1
3	72.7	73.2	72.7	73.1	73.0	73.2
4	150.6	152.4	152.4	152.3	152.1	152.0
5	46.4	45.2	46.4	42.0	44.8	44.7
6	41.8	25.4	20.9	32.5	24.2	24.1
7	203.6	124.8	25.9	60.9	102.7	102.8
8	130.9	136.0	133.6	149.5	154.7	154.9
9	48.9	45.7	142.5	44.3	52.1	52.0
10	38.9	37.4	37.7	41.7	39.3	39.3
11	20.3	21.3	71.1	19.7	22.7	22.7
12	36.1	31.2	39.9	37.1	46.3	46.2
13	45.5	47.8	45.0	42.4	47.4	47.3
14	167.5	97.8	89.7	142.4	163.1	163.1
15	70.3	25.2	71.4	205.7	110.8	110.9
16	37.1	27.6	40.0	42.3	31.7	31.6
17	53.8	51.6	51.8	51.0	56.1	56.2
18	20.4	17.1	20.5	18.7	18.5	18.6
19	12.5	13.9	18.6	12.8	12.7	12.6
20	34.3	36.5	35.0	35.0	35.4	35.3
21	19.5	19.2	18.7	19.4	19.6	19.6
22	34.0	34.2	33.7	33.7	33.4	33.1
23	31.2	30.9	30.7	30.5	30.3	26.8
24	39.4	39.3	39.4	39.3	39.3	46.7
25	32.8	32.8	32.7	32.8	32.6	29.5
26	18.5	18.6	18.6	18.5	18.6	19.4
27	20.5	20.5	20.6	20.4	20.3	19.8
28	15.8	15.7	15.8	15.7	15.9	23.6
29	104.1	104.0	103.5	103.8	104.5	12.7
30						104.5

[a] ^{13}C assignments aided by COSY, HSQC and HMBC experiments.

Table 5. ^{1}H NMR data for tetracyclic nucleus of compounds **50**-**55** (700 and 500 MHz, C_6D_6).[a]

H	**50**	**51**	**52**	**53**	**54**	**55**
3	3.59 dd (4.7, 11.2)	3.75 m	3.71 dd (5.8, 11.3)	3.75 dd (5.2, 11.5)	3.68 dd (5.2, 11.1)	3.69 dd (5.3, 11.1)
5	1.63 m	1.69 dd (3.8, 11.4)	1.85 ovl	2.46 br d (13.6)	1.79 dd (4.1, 11.3)	1.78 dd (4.0, 11.4)
6	2.16 t (16.1)	1.85 m	1.44 ovl	1.59 m	1.90 m	1.91 m
	2.37 dd (12.3, 16.1)	1.99 m	1.59 m	1.72 dt (3.1, 13.8)	2.09 br t (13.8)	2.09 m
7	-	5.52 br d (5.6)	2.01 dd (5.7, 18.0)	6.27 br t (2.9)	5.45 m	5.46 m
			2.61 ddd (6.5, 11.3, 18.0)			
9	1.70 m	2.06 ovl	-	2.57 dd (7.6, 10.2)	1.98 ovl	1.97 ovl
11	1.35 ovl	1.37 ovl	4.37 br t (2.6)	1.23 ovl	1.45 m	1.46 m
		1.51 m		1.42 ovl	1.58 ovl	1.58 ovl
12	1.40 ovl	1.65 br d (13.0)	2.16 dd (2.6, 13.4)	1.11 m	1.38 m	1.37 br t (6.7)
	1.92 br d (11.9)	2.07 ovl	0.81 dd (1.5, 13.4)	1.88 dt (3.3, 12.3)	1.99 ovl	1.99 ovl
15	4.76 br d (6.2)	1.56 ovl	4.46 dd (6.2, 9.8)	-	5.27 t (2.3)	5.28 t (2.3)
		2.23 m				
16	1.66 m	1.35 ovl	1.88 m	1.94 dd (12.6, 18.3)	1.87 m	1.89 m
	2.33 m	2.17 m	1.92 m	2.38 dd (7.7, 18.2)	2.30 dd (8.1, 15.3)	2.31 ddd (2.8, 8.1, 15.4)
17	2.07 m	2.09 ovl	2.41 m	1.33 m	1.75 ovl	1.75 ovl
18	0.77 s	0.72 s	0.43 s	0.80 s	1.16 s	1.15 s
19	0.49 s	0.66 s	0.58 s	0.47 s	0.64 s	0.66 s

[a] Coupling costants are in parentheses and given in hertz. ^{1}H assignments aided by COSY experiments.

3.3.2 Pharmacological evaluation of 4-methylenesterols (phenotype III).

As mentioned before, we demonstrated that specimens belonging to *Theonella swinhoei* phenotype I produce almost exclusively 24-ethylsterols (Figure 36), most of them endowed with potent activity towards to FXR and PXR.[122] Within this library of compounds, we have also established that the presence of a methyl group at position 24 (conicasterols B-D, **47-49**) allows stronger interactions in the external part of FXR molecular surface, with respect to compounds bearing an ethyl group (theonellasterols B-H, **40-46**). Therefore, following this acquisition, all compounds isolated from *Theonella swinhoei* phenotype III (Figure 45,47) were tested *in vitro*, in a transactivation assay (Figure 48).

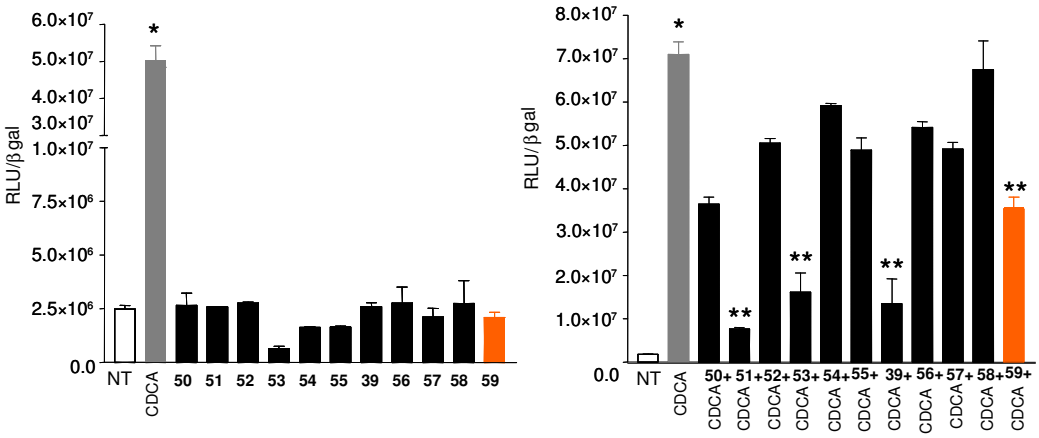


Figure 48. Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-FXR, pSG5-RXR, pCMV-βgalactosidase, and p(HSP27)-TK-Luc vectors and stimulated 18 h with (A) CDCA (10 μM) and compounds **50-55, 39, 56-59** (10 μM). (B) CDCA (10 μM) alone or in combination with compounds **50-55, 39, 56-59** (50 μM). *P < 0.05 versus not treated (NT). **P < 0.05 versus CDCA (n = 4).

HepG2 cells were stimulated with compounds **39, 50-59** alone or in combination with CDCA and, as shown in Figure 48, even if none of these compounds appears to be an FXR agonist in the transactivation assay, several compounds, tested at the concentration of 50 μM, showed a slight inhibitory activity against FXR transactivation induced by 10 μM of CDCA with conicasterol H (**51**), conicasterol J (**53**), swinhosterol B (**59**) and the parent conicasterol (**39**) the most potent of this

series. In addition, many of these steroids effectively induced PXR expression with compounds **54, 55, 39** and compounds **57**-**59** being as effective as rifaximin in inducing PXR transactivation (Figure 49).

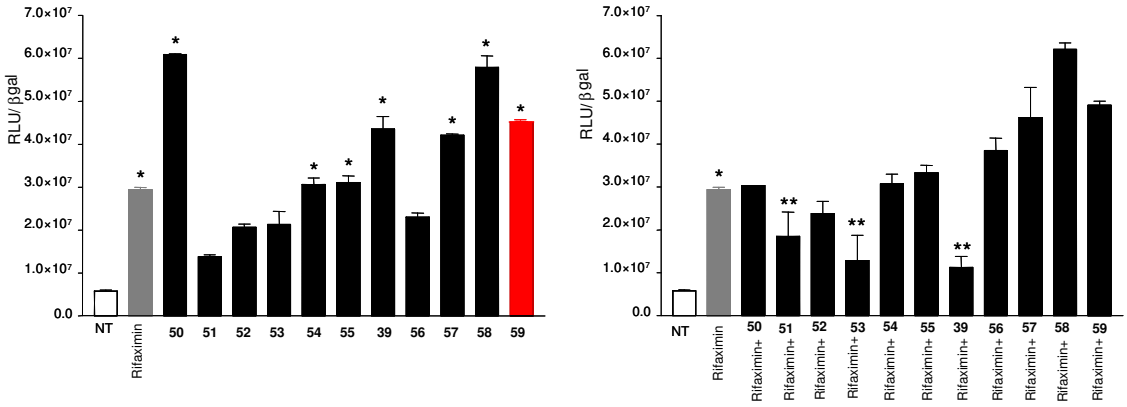


Figure 49. Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-βgalactosidase, and p(CYP3A4)-TK-Luc vectors and stimulated 18 h with (A) rifaximin (10 μM) and with **50**-**55, 39, 56**-**59** (10 μM). (B) Rifaximin (10 μM) alone or in combination with **50**-**55, 39, 56**-**59** (50 μM). *P < 0.05 versus not treated (NT).**P< 0.05 versus rifaximin (n = 4).

Again, these data indicated that 4-methylenesterols from *Theonella* hold the potential to act as PXR agonists and FXR antagonists, and therefore the activity of a select sample against expression of a whole family of nuclear receptors has been evaluated. Among all molecules showing this dual behavior, swinhosterol B (**59**) was selected because it is a potent PXR agonist endowed with a robust FXR antagonism. By profiling the expression of 86 genes using a microarray system (Figure 50) we found that exposure of human HepG2 cells to 10 μM of swinhosterol B (**59**) had no effect on the expression of the waste majority of these genes. However, **59** induced the expression of ESR1, NC0A6, NR1D2, NR2C1, PPARGC1A, PSMC3, PSMC5, RARG, RORA and RXRA. None of these genes were known target for FXR or PXR.

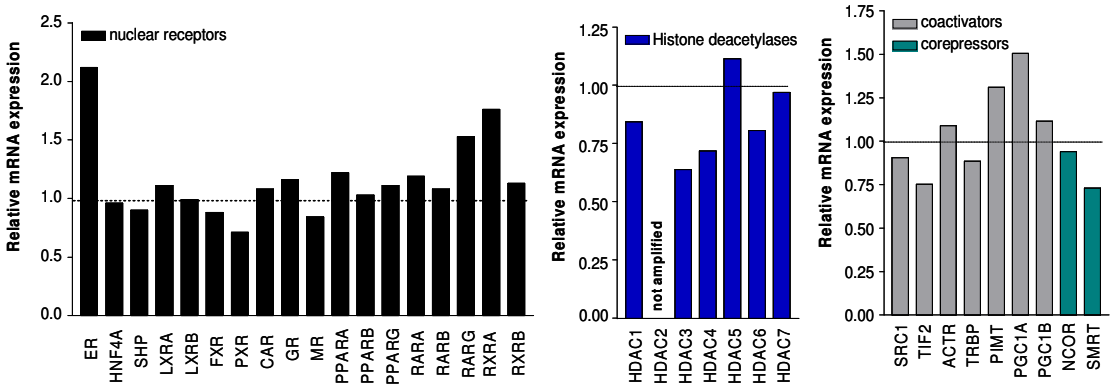


Figure 50. RT^2 profile PCR array analysis showing the relative mRNA expression of (A) various nuclear receptors, (B) histone deacetylases and (C) transcriptional co-regulators (co-activators and co-repressors) following stimulation of HepG2 cells with 10 μM swinhosterol B (**59**). Data are the mean ± S.E. of three experiments.

Because both PXR and FXR have been shown to exert immunomodulatory effects on macrophages we have then investigated whether **59** modulates immune response of macrophages, using cells from the spleen of wild type mice, transgenic mice expressing the human PXR (hPXR) and the spleen of FXR deficient mice ($FXR^{-/-}$). Results of these experiments are shown in Figures 51A-C and demonstrate that swinhosterol B (**59**) effectively counteracts stimulation of hPXR-macrophages caused by LPS. Indeed, at the concentration of 10 μM, swinhosterol B (**59**) causes a robust attenuation of TNFα, IL-1β and IL-6 generation induced by LPS.

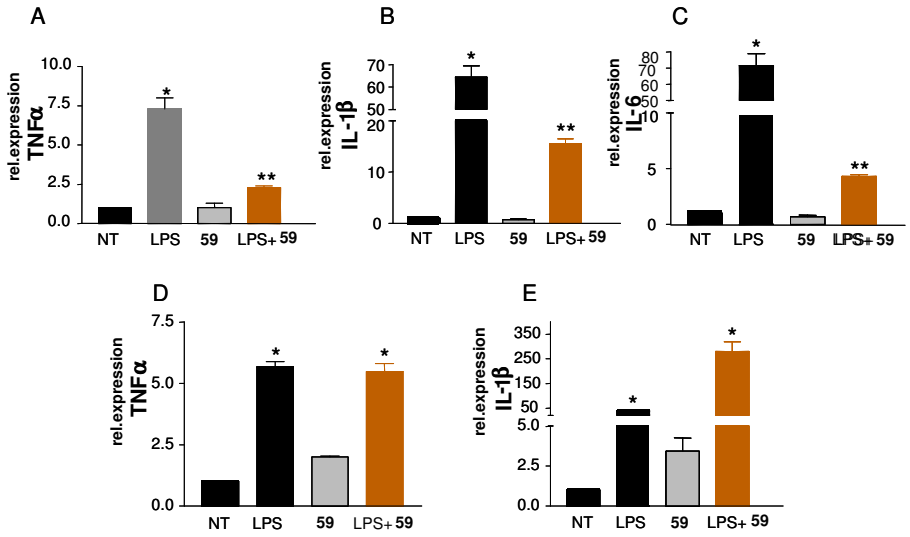


Figure 51. Swinhosterol B (**59**) abrogates stimulation of macrophages caused by LPS. Panel A-C. Transgenic mice expressing the human PXR were used for this experiments. Cells were stimulated with 5 µM LPS alone or in combination with 10 µM of **59**. Data are mean ± SE of 4 spleens. *P<0.01 versus untreated cells (NT). **P<0.01 versus hPXR macrophages exposed to LPS alone. Panel D and E. Swinhosterol B (**59**) fails to recognize the murine PXR in macrophages prepared from wild type mice. Cells were stimulated with 5 µM LPS alone or in combination with 10 µM of **59**. Data are mean ± SE of 4 spleens. *P<0.01 versus untreated cells (NT).

To further investigate if the FXR antagonistic profile was maintained in macrophages expressing the murine PXR, we have challenged spleen-derived monocytes with swinhosterol B (**59**) in the presence of LPS. As shown in Figure 52, in this context swinhosterol B (**59**) fails to inhibit cytokine generation caused by LPS indicating that this compound activates selectively the human PXR and fails to recognize the murine PXR. We have then investigated whether these immunological effects extended to cells of adaptive immunity. For these purposes $CD4^+$ T cells were prepared from the spleen on transgenic mice expressing the hPXR and wild type mice expressing the murine PXR. Data shown on Figure 52 illustrate that, similarly to solomonsterol A (**1**), swinhosterol B (**59**) had no effect on generation of IFNγ (Figure 52A), but potently stimulated the production of IL-10, an anti-inflammatory cytokine from cells isolated from hPXR transgenic mice [Figure 52B; n=4; *P<0.05 versus not treated (NT)]. The relative potency of the effect of swinhosterol B (**59**) at the concentration of 10 µM was comparable to

that of solomonsterol A, 10 µM, and to that of T cell activator and mitogen concanavallin A, 2 µg/ml. Of interest, the ability to induce IL-10 mRNA was lost in $CD4^+$ T cells prepared from wild type mice, *i.e.* mice expressing the murine PXR [Figures 52C and D; n= 4; *P<0.05 versus untreated (NT)]. Thus swinhosterol B (**59**) induces generation of anti-inflammatory IL-10 via induction of h-PXR.

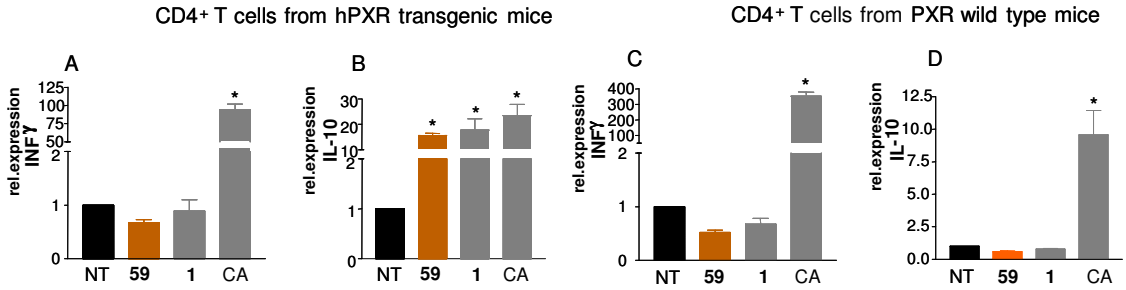


Figure 52. Swinhosterol B (**59**) modulates T cell functions. Panel A and B. Transgenic mice expressing the human PXR were used for these experiments. $CD4^+$ T cells were stimulated with 10 µM of **59** or solomonsterol A (**1**), a potent agonist of hPXR or 2 µg/ml concanavallin A (CA). Data are mean ± SE of 4 spleens. *P<P0.01 versus untreated cells (NT). Panel C and D. Swinhosterol B (**59**) fails to recognize the murine PXR in $CD4^+$ T cells prepared from wild type mice. Data are mean ± SE of 4 spleens. *P<0.01 versus untreated cells (NT).

We finally assayed the FXR antagonistic effect of swinhosterol B (**59**) in HepG2 cells stimulated with the FXR agonist CDCA. As shown in Figure 53, stimulation of HepG2 with **59** was itself sufficient to inhibit FXR target genes such as OSTα, BSEP and SHP [Figures 53A, B and C; n=3; *P<0.05 versus untreated cells (NT)]. In addition, when **59** was combined with CDCA, a robust down-regulation of FXR target genes mediated by CDCA was reported (Figures 53A, B and C; n=3; #P<0.05 versus CDCA stimulated cells). The antagonistic activity of **59** was maintained also for CYP7A1. Indeed, this gene was down-regulated by CDCA and this inhibition was significantly reversed by **59** co-treatment (Figure 53D; n=3; #P<0.05 versus CDCA stimulated cells).

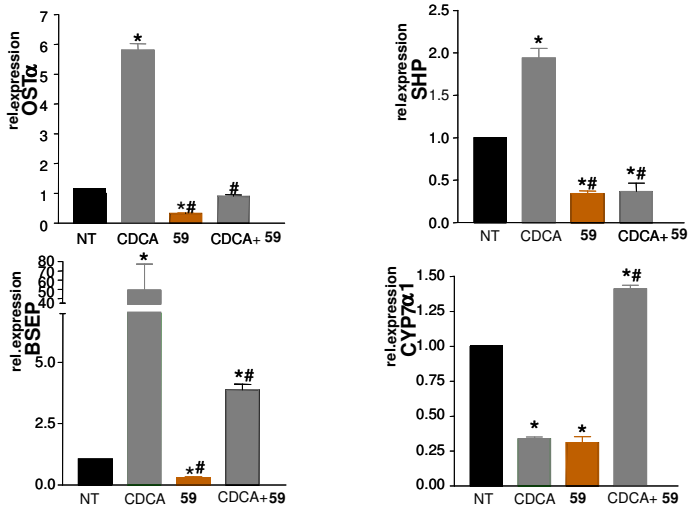


Figure 53. Swinhosterol B (**59**) antagonizes FXR activity in HepG2 cells. HepG2 cells were stimulated with 10 µM of **59**, 10 µM of CDCA or with the combination of the two. Relative mRNA expression of FXR targets (A: OSTα; B: SHP; C: BSEP; D: CYP7α1) was assayed by quantitative Real-Time PCR. Data are mean ± SE of 3 experiments. *P<0.05 versus untreated cells (NT); #P<0.05 versus CDCA stimulated cells.

3.3.3 Docking studies (phenotype III).

Also in this case the binding mode of the 4-methylenesterols isolated from *Theonella swinhoei* sponge (phenotype III) on FXR receptor, was rationalized by molecular docking experiments using Autodock 4.2 software.[108] In particular, we used the detailed binding mode analysis of sterols with new substitution patterns (**50** and **53**) and/or with different or original nuclei (**39**, **54**-**55** and **59**) to obtain useful information for tracing an accurate profile of new potential steroid-based FXR antagonists. Considering the possible hydrophilic interactions with the FXR ligand binding domain, we have focused our attention on **50**, **53** and **59** presenting both carbonyl and hydroxyl groups in different position of tetracyclic nucleus. In this way, comparing docking results, we aimed to rationalize the influence of the different H-bond donor and acceptor pattern on steroids biological activity. For these reasons, we excluded the detailed analysis of compounds with only one OH group at C-7 as in compound **57** or at C-15 as in compound **58** or, as the case of conicasterol I (**52**), presenting the same nucleus of theonellasterol G, already

described above.[121] For what concern the other compounds, comparing the diverse rigidity of the nucleus of conicasterol (**39**) *vs* conicasterol K (**54**), and swinhosterol B (**59**) combined with the different substitution at the C-24 (**39**, **55** and **56**), we have tried to rationalize the influence of the hydrophobic interactions on FXR antagonist behavior of marine steroids. On this basis, compounds **39**, **50**, **53-56,** and **59** are able to interact by the OH at C-3 with FXR binding triad (namely Tyr358 in Helix 7, His444 in Helix 10/11, Trp466 Helix12)[38] (Figure 54A), that, as previously described,[38,121,135] represents one of the principal factor responsible of activity on this nuclear receptor. In particular conicasterol (**39**) and swinhosterol B (**59**), both able to antagonize CDCA (see biological section), form hydrogen bonds with Tyr358 (Helix 7) and His444 (Helix 10/11); on the other hand conicasterol G (**50**), conicasterol J (**53**), theonellasterol J (**55**), and dehydroconicasterol (**56**) interact only with the hydroxyl group of Tyr358 (Helix 7). Comparing the docking pose of our series with respect to the co-crystallized agonist 6-ECDCA and the isomers *E* and *Z* of gugglusterone, it is evident that all the molecules bind FXR-LBD between the Helix 2, 3, 5-7, and 10/11.

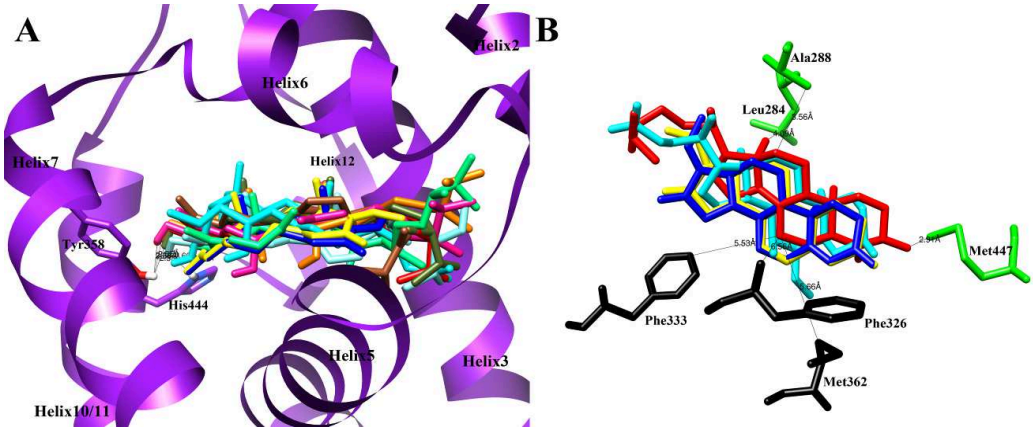


Figure 54. (A) Superimposition of 6-ECDCA (sky blue), and Z (yellow)/E (blue) gugglusterone with **50** (orange), **53** (pink), **54** (dark green), **55** (light blue), **39** (red), **56** (brown), and **59** (emerald) in the binding pocket of FXR (pdb code:1OSV). The FXR molecule is depicted by purple ribbon and the crucial amino acids by sticks (by atom type: C, purple; O, red; N, dark blue, H, white). (B) Amino acids interacting with 6-ECDCA (sky blue) and *Z* (yellow)/*E* (blue) gugglusterone are depicted in black; amino acids interacting with **39** (red) are depicted in green.

Also the molecule side chain, as reported above, influences the activity on FXR; in fact, the methyl group at position 24 of conicasterol (**39**) (red), conicasterol K (**54**) and swinhosterol B (**59**) is in close contact with Ile332 on the surface receptor with respect to theonellasterol J (**55**), bearing an ethyl group with different configuration at the same position (Figure 55A). On the other hand, the major rigidity of the side chain of dehydroconicasterol (**56**) in comparison with the parent conicasterol (**39**) (red) does not allow further hydrophobic interactions with Met262 (Coil 2) and His291 (Helix 3). Moreover, the different positioning of the unsaturations of conicasterol K (**54**) and theonellasterol J (**55**) with respect to conicasterol (**39**) and dehydroconicasterol (**56**) causes a further loss of hydrophobic interactions with the amino acids Leu284, Ile349, and Ile354 (Figure 55A). On the other hand, the unusual open nucleus of swinhosterol B (**59**) maintains the same hydrophobic interactions with respect to **39** with the exception of Leu284 and Met262 (Figure 55A). Considering the other nucleus substitutions, the inverted positions of CO and OH groups at C-7 and C-15 of conicasterols G (**50**) and J (**53**) cause a different pattern of hydrogen bonds; in fact conicasterol G (**50**) is in close contact only with Ser329, while the α-OH at C-7 of conicasterols J (**53**) forms two hydrogen bonds with FXR binding pocket, as H-bond acceptor with OH of Ser329 and as H-bond donor with the OH group of Tyr366. Moreover its CO at C-15, as well as the carbonyl group at C-14 of **59**, establishes a further weak interaction with Ser329 (Figure 55B).

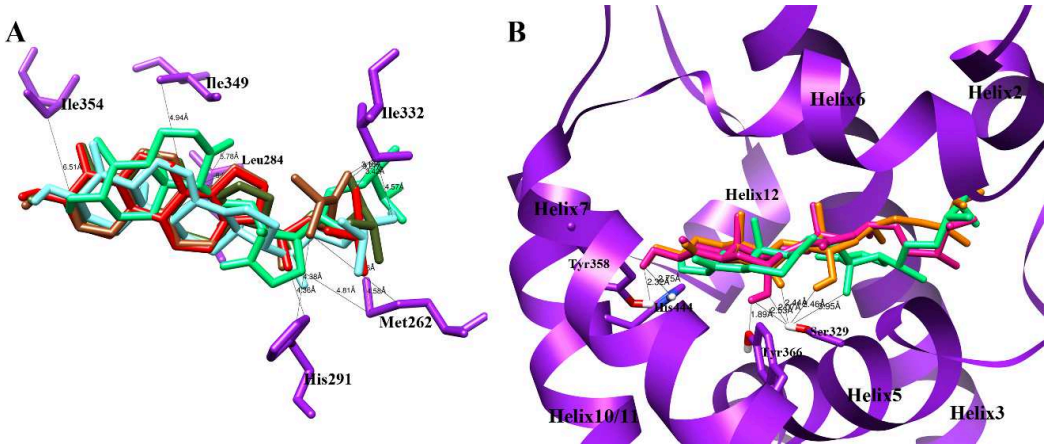


Figure 55. (A) Superimposition between **54** (dark green), **55** (light blue), **39** (red), **56** (brown), and **59** (emerald) in the FXR binding site. (B) Three dimensional model of the different hydrogen bond pattern of **53** (pink), **50** (orange) and **59** (emerald) with FXR. In both figures the crucial amino acids of FXR receptor are depicted by purple sticks.

Swinhosterol B (**59**) was studied in PXR-LBD and three different docking poses are possible. As shown in Figure 56, the OH at C-3 forms hydrogen bonds with CO of Gln285, OH of Ser247, and with CO of His407 for the docking poses A, B, C, respectively. The CO at C-8 establishes further hydrogen bonds with OH of Ser 247 for pose A, while the CO at C-14 interacts with NH of His407 for poses A and B. Besides the different pattern of hydrogen bonds, all the three poses establish different Van der Waal interactions with PXR large ligand binding pocket (Figure 56) formed by hydrophobic (Cys284, Leu206, Leu209, Leu239, Leu240, Leu411, Met243, Met423, Met425, Phe288, Phe420, Trp299, Val211), polar, and charged (Arg410, Tyr306) amino acids. [136,137,138]

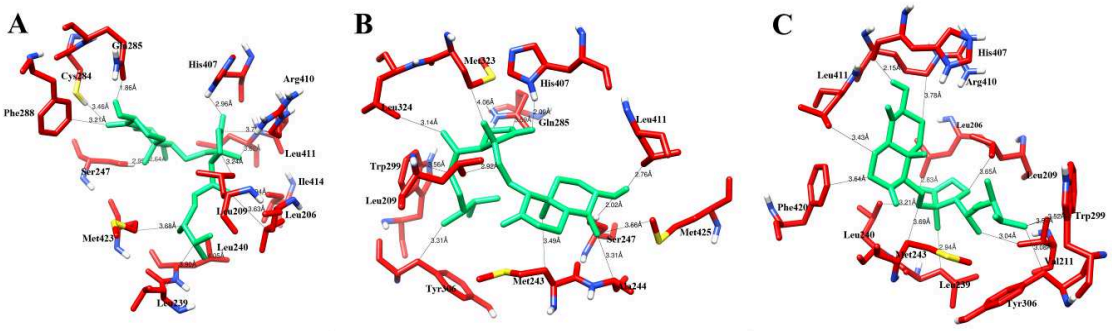


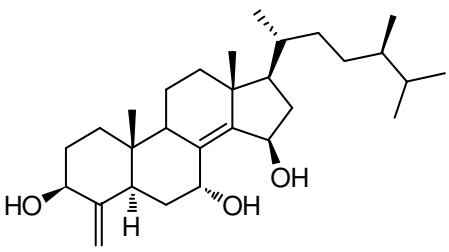
Figure 56. Three dimensional models of the possible docking poses (A-C) of swinhosterol B (**59**) with hPXR.

CHAPTER 4

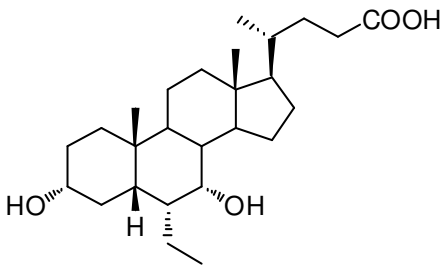
FXR MODULATORS

Among nuclear receptors, farnesoid-X-receptor (FXR) has emerged as a valuable pharmacological target[28,139-142] due to its role in regulating bile acids (BAs), lipid and glucose homeostasis. After its de-orphanization[35,37] a number of nonsteroidal[39] and steroidal compounds,[126] have been shown to interact with the ligand binding domain (LBD) of the receptor and to promote FXR mediated gene transcription. Among these, 6-ECDCA has emerged as a potent, orally bioavailable, FXR agonist[41] and ongoing clinical trials have shown its utility in the treatment of type 2 diabetes.[143] In this scenario the discovery of FXR ligands represents an important answer to the urgent demand of new drugs for the treatment of relevant human diseases including dyslipidemia, cholestasis, non-alcoholic steatohepatitis (NASH) and type-2 diabetes. Nevertheless the use of potent FXR agonists holds some potential risk. Indeed, it has been shown that FXR activation in mammalian cells and tissues inhibits biosynthesis of endogenous bile acids by indirect transrepression of cholesterol-7alpha-hydroxylase (CYP7A1), a gene encoding for the first and rate limiting enzyme involved in their biosynthesis. This effect is indirect and mediated by activation of SHP (small heterodimer partner), an atypical nuclear receptor lacking the DNA binding domain, that binds to liver-X-receptor (LXR) causing its displacement from a positive regulatory element in the CYP7A1 promoter. Despite the effect of SHP has been shown to be dispensable in some settings, it is well recognized that SHP activation amplify the effects of FXR on bile acids uptake and biosynthesis, strongly suggesting that identification of SHP-sparing FXR modulators and FXR

antagonists might have the potential to promote bile acid detoxification without interfering on the biosynthesis. Indeed, FXR antagonists/modulators are rare and to date only few molecules are endowed with this pharmacological profile. The main contribute has been derived from natural compounds. Guggulsterone, the active components of the resin extract of the tree *Commiphora mukul*,[43] and xanthohumol,[144] the principal prenylated chalcone from beer hops, were the first examples of FXR antagonists to be reported from Nature. However Guggulsterone is a promiscuous agent which binds and actives PXR and GR at concentrations that are approximately 100 fold lower than required for FXR antagonism. In this context the sea has recently emerged as an available source of scalarane sesterterpenes,[145] isoprenoids[146] and polyhydroxylated sulfated steroids[135] as FXR antagonists. Once again a considerable contribute derivated from the sponge *Theonella swinhoei*. As mentioned in chapter 3, analysis of the apolar extracts allowed the isolation of a small library of 4-methylenesteroids proved to be potent PXR agonist able to antagonize the effect of CDCA on human FXR. [121,134] Within this series, we found also conicasterol E (**60**), a 7α,15β-dihydroxyconicasterol derivative, as the first example of SHP-sparing FXR modulator endowed with PXR agonistic activity from a marine source. Furthermore, in order to prove its efficiency in transactivation assays in comparison to a well validated FXR agonist, we have generated a novel synthetic strategy to obtain 6-ECDCA, a widely used CDCA derivative.[147]



Conicasterol E (**60**)

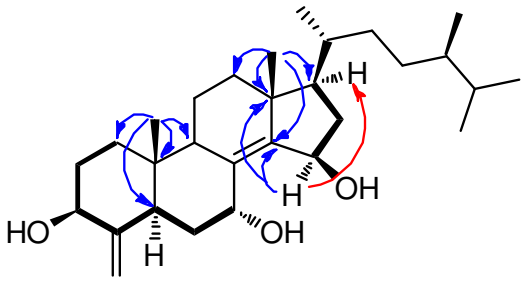


6-ECDCA (**61**)

Figure 57. Conicasterol E (**60**), the first example of SHP-sparing marine FXR modulator, and 6α-ethyl-chenodeoxycholic acid (**61**).

4.1 Isolation and structural characterization of conicasterols E (60)

The initial processing of the *Theonella swinhoei* (coll. No. R3159) was conducted according to procedures previously described.[74] The *n*-hexane extract from a solvent partitioning Kupchan procedure was cromatographed by silica gel and the fraction eluted with CH_2Cl_2:MeOH (96:4) was further purified by reverse phase HPLC to give 2.1 mg of conicasterol E (**60**) as a colourless amorphous solid ($[\alpha]_D^{25}$ +59.6). The molecular formula of $C_{29}H_{48}O_3$, established by HR ESIMS, $[M+Li]^+$ at *m/z* 451.3769 (calculated 451.3763) and NMR data (Table 18) were compatible with a steroidal tetracyclic nucleus, two double bond and three hydroxyl groups in the molecule. COSY correlations delineated the spin system H-1 through H-7 and the spin system H-15/H-17 with OH substitutions at C-7 (δ_H 4.63) and at C-15 (δ_H 4.57) whereas the presence of a Δ8(14) double bound was inferred from careful analysis of HMBC data reported in the Table 18 and in Figure 58. The small vicinal coupling constant of H-7 (br t, *J*=2.9 Hz) allowed us to establish an equatorial disposition for this proton, thereby placing the hydroxyl group in an axial α-orientation, while the ROE effect H-15/H-17α (Figure 58) was indicative of a β-orientation of the hydroxy group at C15.



60

Figure 58. COSY connectivities (bold bonds), HMBC (blue arrows) and ROESY correlations (red arrows) for conicasterol E (**60**).

Comparison with NMR data of theonellasterol F (**44**) evidenced a strong resemblance with conicasterol E. The difference between **60** and theonellasterol F (**44**) lies in the side chain with a methyl group (δ_H 0.88, 3H, d, J = 6.7 Hz) replacing the C24 ethyl group present in all theonellasterol -like compounds.

Table 6. NMR data (500 MHz, C_6D_6) for **60**.

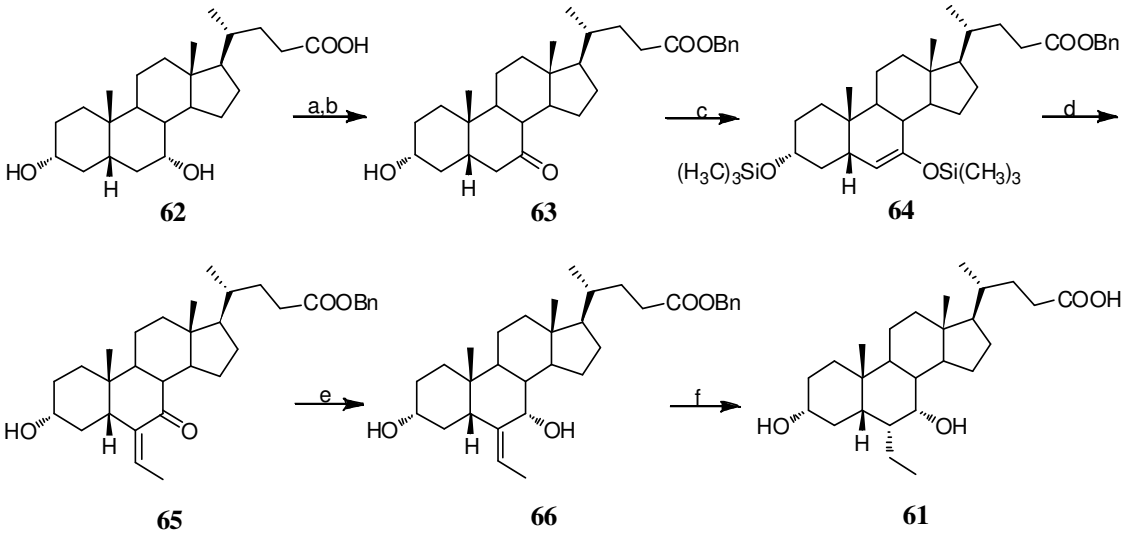
position	δ_H	δ_C	Key HMBC
1	1.11 m, 1.52 ovl	36.7	
2	1.30 ovl, 1.82 m	33.4	
3	3.77 m	73.4	
4	-	153.3	
5	2.41 m	43.2	
6	1.65 ovl, 1.91 ovl	31.7	
7	4.63 br t (2.9)	66.6	
8	-	136.7	
9	2.40 m	45.5	
10	-	40.0	
11	1.54 ovl	20.3	
12	1.31 ovl, 1.92 ovl	38.1	
13	-	43.8	
14	-	151.5	
15	4.57 br d (4.8)	70.1	C13, C14, C17
16	1.50 ovl, 1.66 ovl	39.5	C13, C14, C15, C17
17	1.64 ovl	53.4	C13, C14, C15
18	0.79 s	19.8	C12, C13, C14, C17
19	0.59 s	12.7	C1, C5, C9, C10
20	1.43 m	34.6	
21	1.00 d (6.3)	19.1	
22	1.20 m, 1.48 m	33.7	
23	1.16 m, 1.47 m	30.6	
24	1.33m	39.3	
25	1.59 m	32.8	
26	0.88 d (6.7)	18.6	C24, C25
27	0.94 d (6.7)	20.4	C24, C25
28	0.87 d (6.8)	15.7	C23, C24, C25
29	4.72 br s, 5.32 br s	103.3	C3, C4, C5

Coupling constants are in parentheses and given in hertz. ^{1}H and ^{13}C assignments aided by COSY, TOCSY, ROESY, HSQC and HMBC experiments.

4.2 New synthetic strategy of 6-ECDCA (61)

So far two synthetic procedures of 6-ECDCA have been reported. The first process was based on the alkylation of the 3-tetrahydropyranyloxy derivative of 7-keto-lithocholic acid (LCA) with lithium diisopropylamide and ethyl bromide followed by standard reduction and hydrolysis steps.[41,148] Our preliminary screening of this synthesis revealed low-yielding steps, especially in the alkylation of 7-keto-LCA, and the need for chromatographic purification of each step. Recently an alternative procedure[149] via aldol-type addition of a silyl enolether derivative of 7-keto-LCA methyl ester with acetaldehyde, followed by hydrogenolysis with PtO_2, alkali hydrolysis (10% NaOH in refluxing methanol) and selective reduction of the C7-ketone with sodium borohydride has been reported (58% overall yield from 7-keto-lithocholic acid). Despite our extensive effort, we faced several problems in reproducing the reported high yield in the hydrolysis and reduction steps. In fact the hydrolysis of methyl ester in alkaline condition proceeded with low yields and extensive epimerization at C-6 position due to the presence of the carbonyl group at C-7. Alternatively we tested the possibility to invert the last two steps of the reported protocol and to perform firstly the C-7 reduction followed by methyl ester hydrolysis. Despite the reported regio- and stero-selectivity of $NaBH_4$ reduction, in our hand the reduction of methyl 3α-hydroxy-6α-ethyl-7-keto-5β-cholan-24-oate proceeded with the formation of large amount of the over-reduced product with the primary alcoholic function at C-24 and with scarce stereoselectivity in the introduction of hydroxyl group at C-7 with the required α-configuration. To overcome the inconvenience relaying on the coexistence of C-7-ketone and C-24-methyl ester, we initially attempted to perform the generation of silyl enolether on the 7-keto-LCA without

the protection of the carboxyl function at C-24 but, although our several efforts (DIPA, BuLi, TMSCl, Et_3N, THF, -78 °C one pot procedure; generation of 3α-silyl ether with Et_3N and TMSCl followed by LDA/TMSCl formation of enolether), no transformation occurred. At this point we decided to protect the carboxyl function at C-24 as benzyl ester. The choice of benzyl as protecting group answered to two essential demands: reduction of synthetic steps in order to achieve better chemical yield and, importantly, improvement of regio- and stereoselectivity of the entire process. The synthesis of **61** starting from commercially available chenodeoxycholic acid is outlined in Scheme 8. Oxidation with sodium hypochlorite solution/NaBr and tetrabutylammonium bromide in a mixture of methanol/acetic acid/water/ethyl acetate as solvent afforded 7-keto-lithocholic acid in nearly quantitative yield and without trace of 3-keto regioisomer and di-oxidated product. Benzylation, generation of silyl enolether followed by aldol addition with acetaldehyde in the presence of $BF_3{\cdot}OEt_2$ gave the desired benzyl 3α-hydroxy-6-ethylidene-7-keto-5β-cholan-24-oate (**65**) in 42% yield over four steps. High selective reduction of the C7-ketone with $NaBH_4/CeCl_3$ in a mixture of THF/MeOH at room temperature followed by concomitant removal of the benzyl-protecting group and reduction of exocyclic double bond (H_2 on Pd/C Degussa type) afforded the desidered 6-ECDCA (**61**). In vitro pharmacological screening demonstrated that the so obtained 6-ECDCA (**61**) is a potent FXR agonist (Figure 59).



Scheme 8. Reagents and conditions: (a) $NaOCl/Bu_4N^+Br^-$, NaBr, 0 °C; (b) BnBr, Cs_2CO_3, CH_3CN, reflux; (c) DIPA, *n*-BuLi, TMSCl, Et_3N, THF, -78 °C; (d) MeCHO, CH_2Cl_2, $BF_3{\cdot}OEt_2$, -60 °C; (e) $NaBH_4/CeCl_3$, THF/MeOH 4:1; (f) H_2, Pd/C, THF/MeOH 1:1.

4.2.1 Pharmacological evaluation

Conicasterol E (**60**) was tested in vitro using a luciferase assay on HepG2 cells transfected with FXR. As shown in Figure 59, conicasterol E was almost as potent as the natural FXR ligand CDCA in transactivaring FXR, but less potent than the synthetic FXR ligand 6-ECDCA (**61**) (Figure 59). Conicasterol E (**60**) was devoid of any antagonistic activity when co-administered with CDCA (Figure 59B) and 6-ECDCA (data not shown).

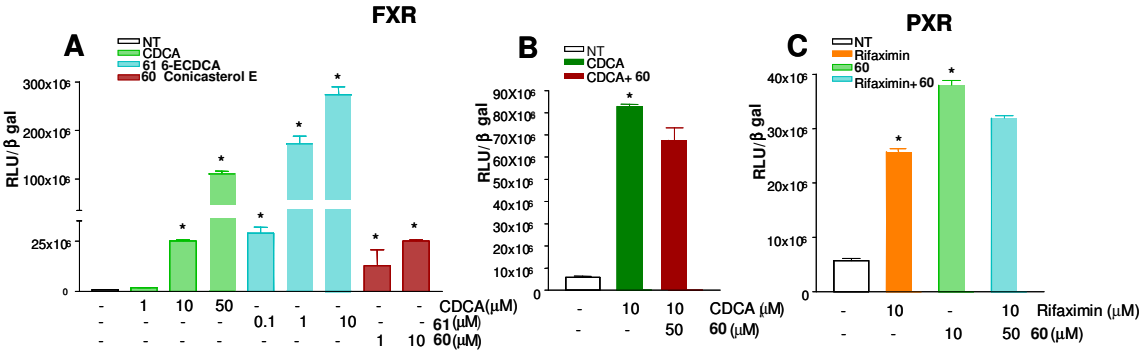


Figure 59. Panel A. Relative potency of FXR activation by CDCA, 10 µM, conicasterol E (**60**), 10 µM, and 6-ECDCA (**61**), 1 µM, as measured by transactivation assay in HepG2 cells. Panel B. Conicasterol E (**60**), 10 µM, does not revert the effect of CDCA, 10 µM, on FXR transactivation in HepG2 cells. Panel C. Relative potency of PXR activation by rifaximin, 10 µM, and conicasterol E (**60**) alone, 10 µM, or in combination, 50 µM. Data are mean ± SE of 4 experiments.

In addition to an FXR agonistic activity, conicasterol E (**60**) effectively induced PXR expression, being as effective as rifaximin in inducing PXR transactivation (Figure 59C). Thus conicatserol E (**60**) is an FXR and PXR agonist. To further characterize the biological activity of the conicasterol E (**60**), we have examined the effect of this agent on the expression of canonical FXR and PXR target genes in hepatocytes, and, as shown in Figure 60, we found that exposure to conicasterol E slightly increased the expression of OSTα and BSEP mRNAs (two FXR regulated genes) and the expression of CYP3A4 mRNA (a PXR-regulated gene), while no effect was observed on SHP mRNA expression. In addition, in contrast to CDCA, **60** failed to repress CYP7A1. Thus, while the expression of this gene was reduced by 30% by CDCA, exposure to conicasterol E (**60**) increased CYP7A1 mRNA by 2-3 folds. These data are a further evidence that in HepG2 cells repression of CYP7A1 by FXR is indirect and requires induction of SHP.

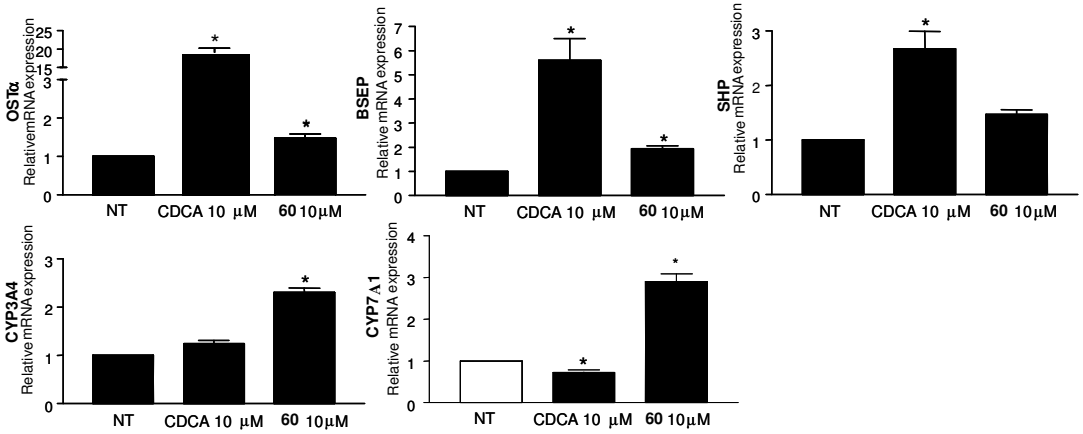


Figure 60. RT-PCR analysis of effects of CDCA, 10 μM, and conicasterol E (**60**), 10 μM, on expression of FXR-regulated genes in HepG2 cells. Conicasterol E (**60**) does not induce SHP and fails to repress the expression of CYP7A1. Data are mean ± SE of 4 experiments.

Further on, when administered in combination with a concentration of CDCA of 10 μM, conicasterol E exerted an additive effect with CDCA on the expression of OSTα and BSEP while no further changes were observed in the expression of

SHP (Figure 61). Taken together, these data highlight that conicasterol E (**60**) is a FXR modulator whose potency on selective target genes is very close to that of the endogenous mammalian ligand CDCA and lower than that of a synthetic agonist 6-ECDCA. Interestingly, conicasterol E (**60**) failed to stimulate SHP even when co-administered in combination with CDCA.

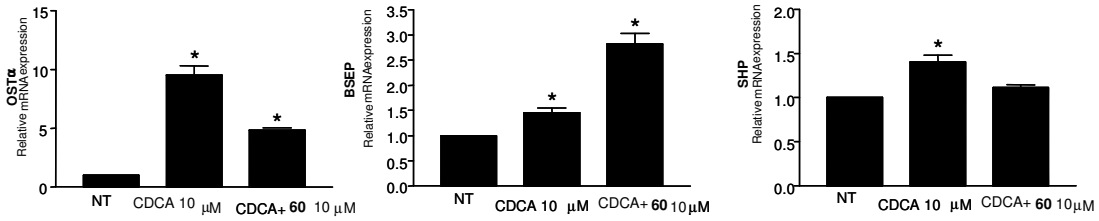


Figure 61. RT-PCR analysis of effects of CDCA, 10 μM, alone or in combination with conicasterol E (**60**), 10 μM, on expression of FXR-regulated genes in HepG2 cells. Conicasterol E (**60**) does not induce SHP even when cells were co-incubated with CDCA while the association of the two agents induce a robust expression of OSTα and BSEP. Data are mean ± SE of 4 experiments.

Finally, analysis of CYP3A4 expression, shown in Figure 62, demonstrated that conicasterol E (**60**) has no antagonistic effects on expression of CYP3A4 mRNA induced by rifaximin, a potent PXR agonist.

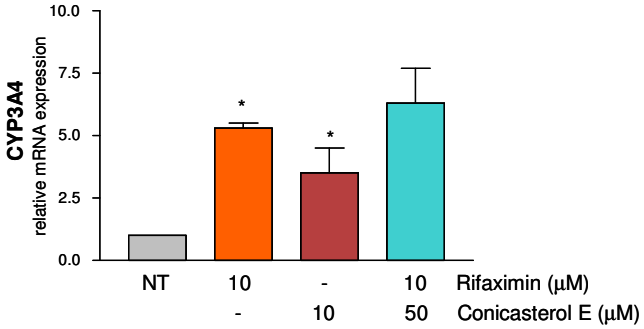


Figure 62. Activation of CYP3A4 by the PXR agonist rifaximin, 10 μM, is not modulated by conicasterol E (**60**) 50 μM . Data are mean ± SE of 4 experiments.

4.2.2 Docking studies

As reported in previous chapter, 4-methylenesterols isolated from *Theonella swinhoei* are able to modulate in different ways the FXR activity depending on the steroid skeleton substitutions. As shown in Figure 63A, the FXR binding site,

located between the Helix 2, 3, 5-7 and 10/11, is occupied by **60** and, as previously reported,[121] the β-OH groups at position 3 and 15, and the *trans* junction between A/B rings cause a different positioning with respect to the co-crystallized molecule 6-ECDCA. In particular (Figure 63B) conicasterol E**,** compared to the synthetic agonist 6-ECDCA, is able to interact with two amino acids of the catalytic triad formed by Tyr358 in Helix 7, His444 in Helix 10/11, Trp466 Helix12, responsible of the activation of FXR.[38] Specifically, the 3-OH group at β position forms a hydrogen bond with Tyr358 in Helix 7, while the *trans* junction between the A/Bring allows the hydrophobic interaction with His444 (Helix 10/11). In previous chapter, we have described the influence of the side chain on the FXR binding; in fact, the methyl at position 24 of conicasterol E (yellow, Figure 63B) relating to 6-ECDCA (red, Figure 63B) is able to simultaneously interact with the Met262 (Coil 2), His291 (Helix 3), and Met287 (Helix 3), present on the external part of the receptor molecule. Moreover, the OH at position 15β in **60** forms an additional hydrogen bond with the CO of Leu284 (Helix 3), and the steroid skeleton is in close contact with Leu345, Ala288, Met447, Phe326 and Trp451 relating to the 6-ECDCA. On the other hand, the OH at 7α position does not seem to exert further polar interactions with the FXR binding site.

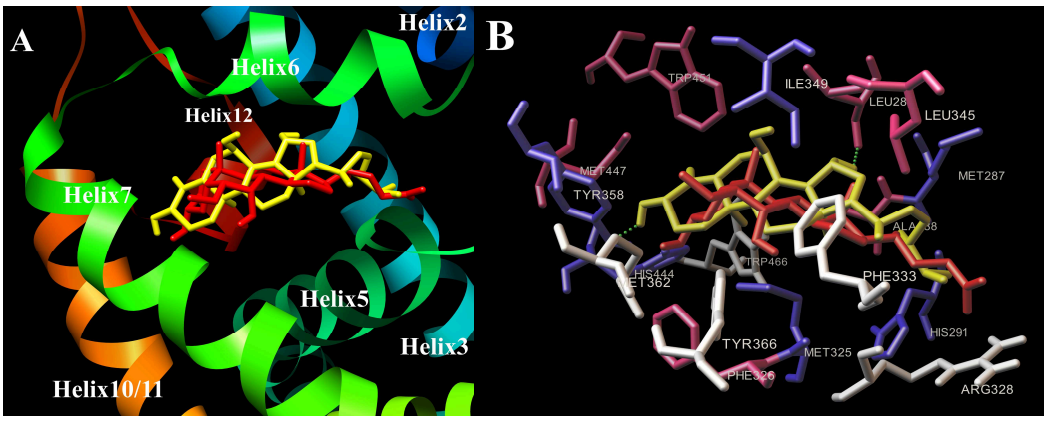


Figure 63. (A) Superimposition of **60** (yellow) with 6-ECDCA (red) in the binding pocket of FXR (pdb code:1OSV).[Errore. Il segnalibro non è definito.] (B) Amino acids interacting with 6-ECDCA (red) are depicted in white, amino acids interacting with **60** (yellow) are depicted in pink, and amino acids interacting with both molecules are depicted in blue.

In summary, respect to the other 4-methylene-steroids endowed with FXR antagonistic profiles, the exclusive amino acid interactions exerted by conicasterol E (**60**) might support the notion that the compound is FXR agonist endowed with the ability to activate OSTα and BSEP without effect on SHP expression. These studies pave the way to further elaborating on the critical interactions on the FXR-LBD aimed to the identification of site-specific ligand that could be used to induce selective genes.

4.3 Theonellasterol, a new lead in cholestasis

Cholestasis is a liver disorder that occurs primarily in the context of genetic mutation of basolateral or apical membrane transporters in hepatocytes. Cholestasis represents the main biochemical feature of primary biliary cirrhosis[141,150] (PBC) and sclerosing cholangitis (PSC), two immune-mediated disorders characterized by progressive bile duct destruction for which medical therapy is still poorly effective and investigations are ongoing to identify novel therapeutic approaches. Theoretically, because PBC and PSC are characterized by bile duct destruction, therapy should be aimed at activating bile acid secretion

from the basolateral membrane of hepatocytes, while stimulation of bile acid secretion from the apical membrane is likely to worsens liver injury due to the obstruction of bile flow.[141,150] While FXR activation favours bile acid detoxification by hepatocytes and FXR agonists have been proposed in the treatment of PBC patients,[141] results from models of obstructive cholestasis in FXR2/2 mice have shown that FXR gene ablation protects against liver injury caused by ligation of common bile duct (BDL).[151] Molecular decoding of the BDL model has lead to the demonstration that FXR functions as a negative regulator of multidrug resistance-associated protein (MRP)-4, a gene mediating basolateral secretion of bile acids. In vitro characterization of interaction of FXR with MRP-4 has lead to the demonstration that FXR functions as a braking signals for MRP-4 induction caused by activation of Constitutive Androstane Receptor (CAR).[152] In aggregate, these data suggest that FXR activation in obstructive cholestasis might worsen liver injury by hijacking a protective mechanism regulated by CAR, i.e induction of MRP-4.[143] While these data strongly advocate the utility of an FXR antagonist in the treatment of obstructive cholestasis, this concept has remained unproved for long time because the lack of a selective FXR antagonist. As detailed before, the major component of the steroidal fraction of *Theonella swinhoei* is theonellasterol (**38**),[153] which is proposed as a taxonomic marker for the Theonella sponge phenotypes[65] and the biogenetic precursor of all theonellasterols like 4-methylene steroids so far isolated. During the extraction procedure of 4-methylene steroids, theonellasterol (**38**) was isolated in high amounts following a very simple procedure. Following our acquisition of the ability of 4-methylene steroids to act as NRs ligands, theonellasterol (**38**) was evaluated on FXR in a luciferase transactivation assay. As shown in Figure 64A

and B, stimulation of HepG2 cells with 10μM of theonellasterol (**38**) failed to activate FXR. By contrast theonellasterol (**38**) effectively antagonized FXR transactivation induced by CDCA.

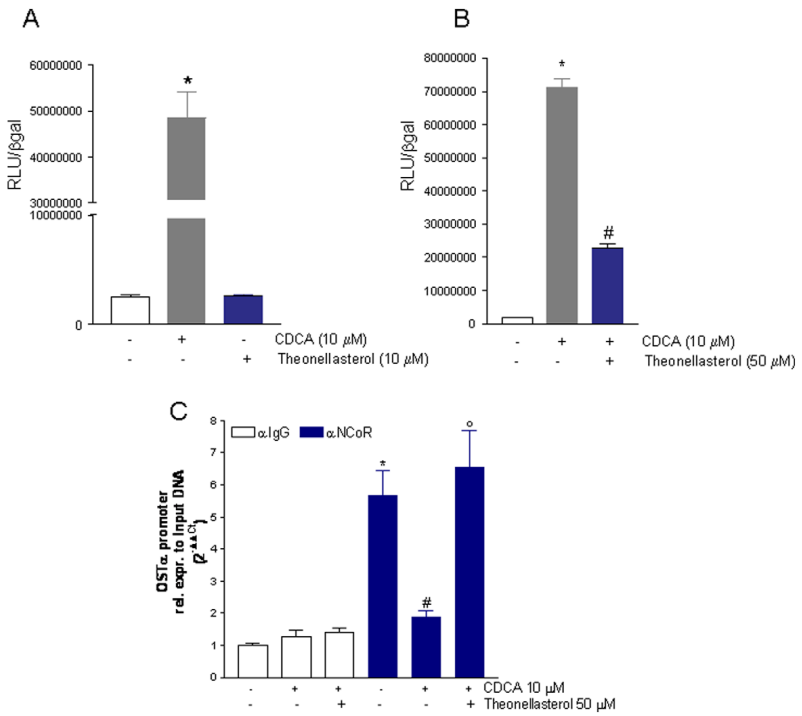


Figure 64. Luciferase reporter assay performed in HepG2 transiently transfected with pSG5-PXR, pSG5-RXR, pCMV-bgal, pCYP3A4promoter-TKLuc vectors and stimulated 18 h with (A) 10 mM of CDCA or theonellasterol (**38**) and (B) 10 mM of CDCA alone or in combination with theonellasterol (**38**) 50 mM. (C) CHiP assay of NCoR binding to the OSTa promoter. RT-PCR analysis of proteins immune-precipitated with a control IgG are shown as control.

The antagonistic activity was further confirmed examining the effect of theonellasterol (**38**) on the expression of canonical FXR target genes[150]. Theonellasterol (**38**) reversed the effect of CDCA on the expression of OSTα, BSEP, SHP and MRP-4 (Figure 65A–C; n =4; P,0.05 versus CDCA alone). Because the regulatory activity of theonellasterol (**38**) on MRP-4 holds promise for its potential therapeutic use in obstructive cholestasis, the antagonistic activity of theonellasterol (**38**) on MRP-4 was examined from a molecular point of view. For this purpose a ChIP assay was carried by immune-precipitating nuclear extracts from HepG2 cells left untreated or primed with CDCA alone or with the

combination of CDCA plus theonellasterol (**38**) with an anti-FXR antibody. As shown in Figure 65E, results of Real-Time PCRs demonstrated that while in basal conditions FXR is not constitutively bound to the MRP-4 promoter, but is recruited on the promoter following activation with CDCA. Recruitment of FXR to the MRP-4 promoter in the presence of CDCA was robustly attenuated by co-incubating the cells with the theonellasterol (**38**) (n= 4; P,0.05 versus CDCA alone). All together these results indicate that theonellasterol (**38**) exerts its antagonistic activity by reducing the binding of FXR on the MRP-4 promoter, thus preventing its down-regulation caused by CDCA (Figures 65D and E). Thus, as illustrated in Figure 64C, the ChIP analysis demonstrates that while exposure to CDCA, 10 μM, released NCoR from the OSTα promoter, co-treating cells challenged with CDCA with the theonellasterol (**38**), 50 μM, abrogated this pattern (n= 3; P,0.05 versed CDCA alone). Finally, exposure of HepG2 cells to 50 μM theonellasterol (**38**) effectively stabilized the nuclear corepressor NCoR at its binding site in the promoter of OSTα, a well characterized FXR-regulated gene.

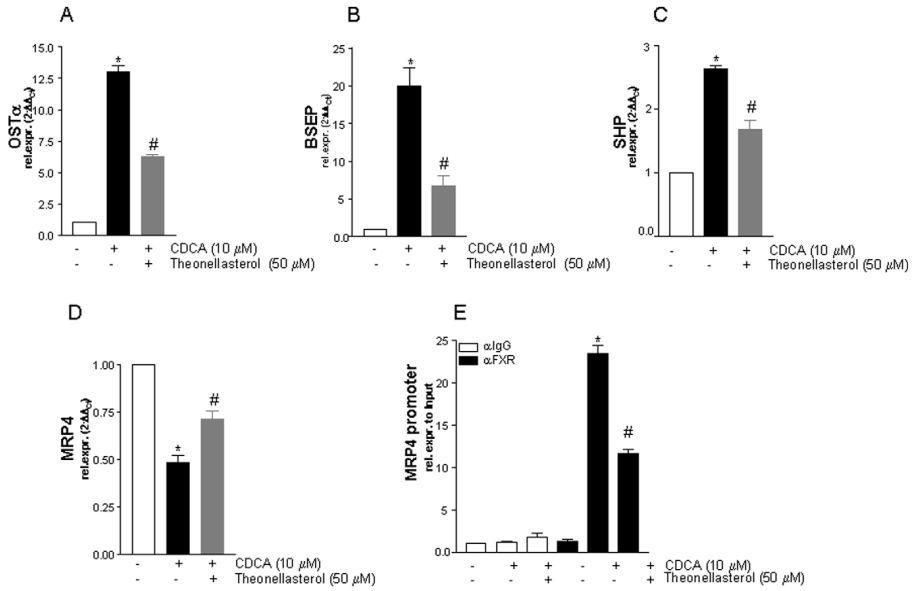


Figure 65. Relative mRNA expression of (A) OSTa, (B) BSEP, (C) SHP and (D) MRP4 in HepG2 cells treated with 10 mM CDCA alone or with the combination of CDCA plus theonellasterol (**38**) 50 mM. (E) CHiP assay performed in HepG2 cells not stimulated or primed with CDCA, 10 mM, alone or in combination with theonellasterol (**38**) , 50 mM.

Moreover theonellasterol (**38**) was demonstrated the first example of selective FXR antagonist. As shown in Figure 66A–D, theonellasterol (**38**) at the concentration of 10 μM failed to transactivate PPARγ, PXR, VDR and GR, nor it inhibited the activation of these receptor promoted by specific ligands, i.e. rosiglitazone, rifaximin, 1,25 dihydroxy colecalciferol and dexamethasone when co-incubated with these selective agonists at the concentration of 50 mmol/L. In addition microarray analysis demonstrated that theonellasterol (**38**) itself had no effect on the expression of any of these regulatory factors (not shown) and the sponge steroid caused no changes in the expression of these nuclear receptors and regulatory factors in cells challenged with CDCA.

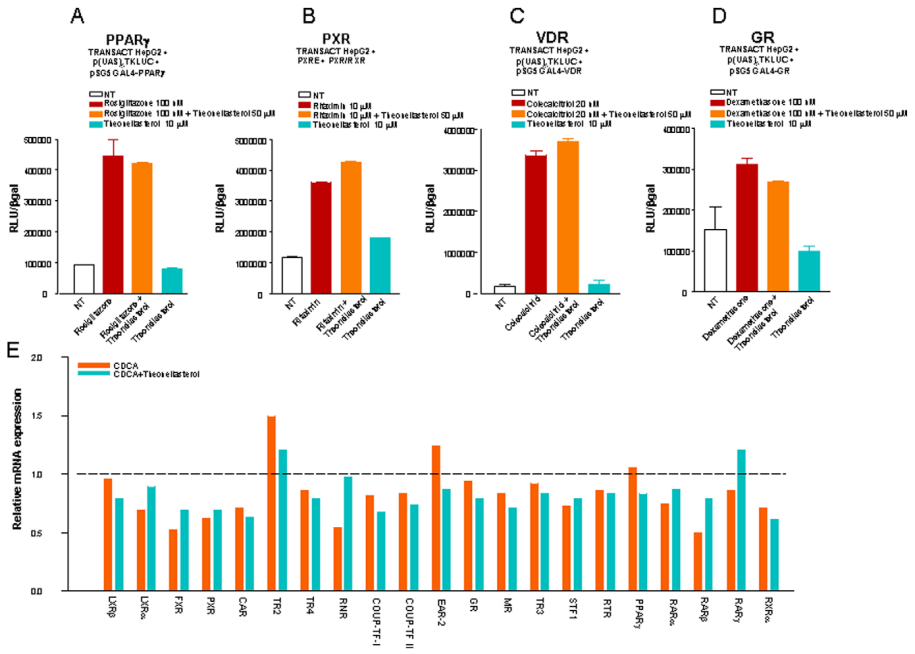


Figure 66. (A) (C) and (D) HepG2 cells were co-transfected with the Gal4 luciferase reporter and a series of chimeras in which the Gal4 DNA binding domain is fused to the LBD of the indicated nuclear receptors. Cells were treated with the appropriate agonists or specific agonists in combination with theonellasterol (**38**). (B) HepG2 cells were co-transfected with pSG5-PXR, pSG5-RXR and with the reporter pCYP3A4promoter-TKLuc and then stimulated with rifaximin, a PXR agonist, alone or in combination with theonellasterol (**38**). E) Microarray analysis showing the relative mRNA expression of various nuclear receptors and nuclear receptors co-activators following stimulation of HepG2 cells with CDCA, 10 mM, alone or in combination with theonellasterol (**38**) 50 mM.

Because these data illustrate that theonellasterol (**38**) is a selective FXR antagonist, it was designed a proof-of-concept study to ascertain whether this compound was effective in attenuating liver injury caused by BDL in mice, a model of obstructive cholestasis that is attenuated by FXR gene ablation.[151] For this purpose BDL mice were administered the theonellasterol (**38**), 10 mg/kg, or an FXR agonist, 6-ECDCA (**61**), 30 mg/kg, for 3 days. In this model theonellasterol (**38**) reduced intrahepatic bile duct pressure and attenuated liver injury caused by bile duct ligation as measured by serum alanine aminotransferase (ALT) levels and extent of liver necrosis at histopatology. On the contrary administering bile duct legated mice with 6-ECDCA (**61**) failed to rescue from liver injury (Figure 67-68).

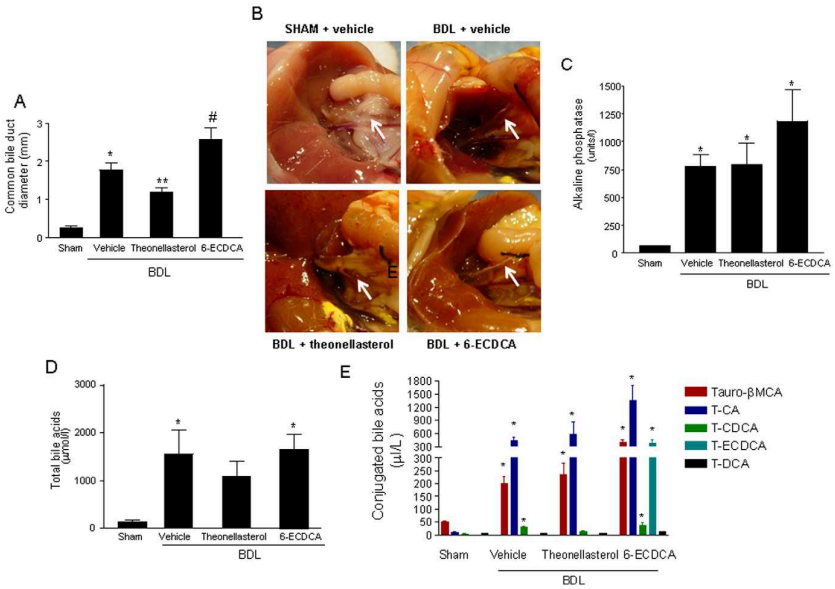


Figure 67. (A) Common bile duct dilation, 3 days after BDL, is worsened by 6-ECDCA (**61**) and attenuated by theonellasterol (**38**). (B) Representative macroscopic features of common bile duct (white arrows) in 3-day BDL mice administered theonellasterol (**38**) and 6-ECDCA (**61**). (C) BDL increases serum levels of alkaline phosphatase, a marker of bile duct obstruction. (D) Serum concentration of total bile acids. (E) Quantitative analysis of tauro-conjugated bile acids in BDL animals.

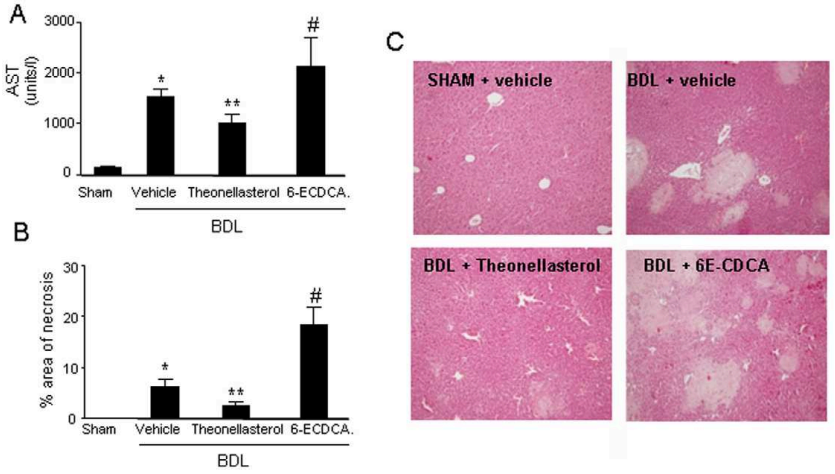
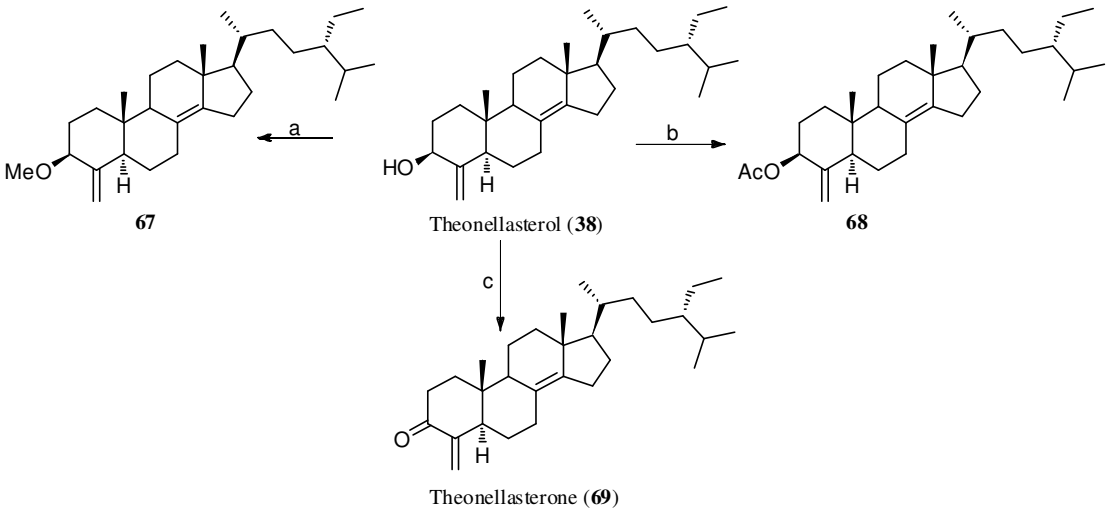


Figure 68. (A and B) Liver necrosis was robustly attenuated by theonellasterol (**38**) as confirmed by assessment of ALT and histopathology analysis. C) Representative liver histology from an individual mice per group. Liver sections were stained with H&E, original magnification 106.

4.4 Preliminary Structure-Activity Relationship on Theonellasterol

Having established the therapeutic potential of theonellasterol (**38**) in cholestasis, the effect of chemical transformations on its biological activity and the first structure-activity relationship (SAR) study have been performed in the last period of my research work.[154] As the highly hindered 8,14 double bond was found to be chemically unreactive toward most chemical reagents, two are the points for

chemical modification in the structure of theonellasterol (**38**): the exocyclic carbon-carbon double bond at C-4 and the hydroxyl group at C-3. These functionalities were subjected to simple chemical reactions and the products obtained (**67**–**77**) were fully characterized by means of MS, and 1D and 2D NMR spectroscopy. The methyl ether derivative (**67**), the 3-*O*-acetyl derivative (**68**) and the α,β-unsaturated ketone (**69**), already known as theonellasterone,[65] were prepared from theonellasterol (**38**) (Scheme 9) to explore the pharmacophoric role of the hydroxyl group at C-3 as hydrogen bond donor in the FXR-LBD (Scheme 9).



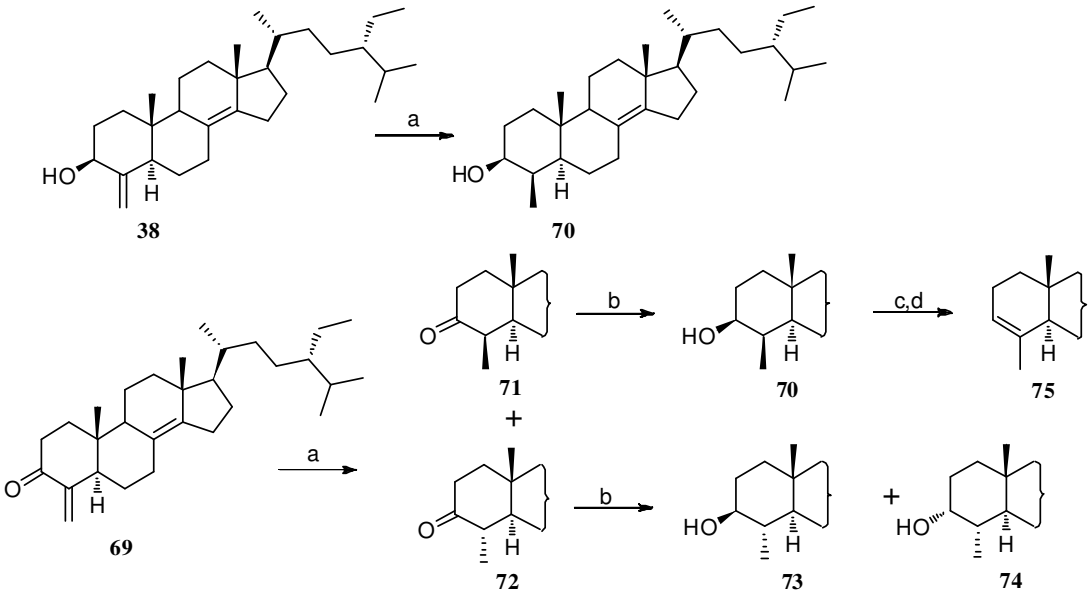
67 Theonellasterol (**38**) **68**

Theonellasterone (**69**)

Scheme 9. Modification at C-3 hydroxyl group. Reagents and conditions: (**a**) NaH, THF, MeI, 0 °C, 79%; (**b**) Ac_2O, pyridine (pyr), room temperature (rt), 96%; (**c**) PCC, CH_2Cl_2, quantitative yield.

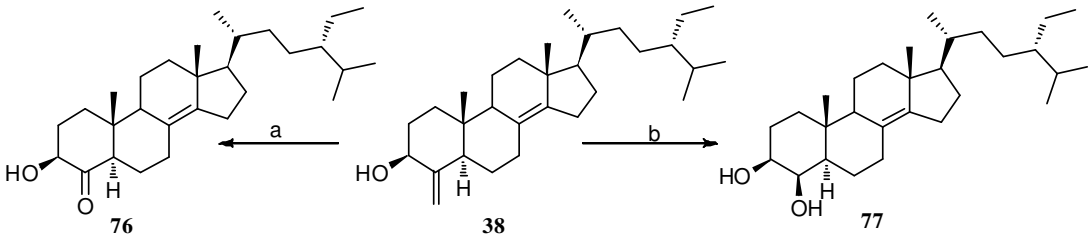
As concern the exomethylene on ring A, hydrogenation of theonellasterol (**38**) on different catalysts (Pt/C, $Pd(OH)_2$ Degussa type, PtO_2) produced exclusively the 4β-methyl derivative (**70**) through the approach of the hydrogen from the α-face of the steroid nucleus (Scheme 10). To obtain the 4α-methyl theonellasterol derivative, a synthetic procedure via theonellasterone (Scheme 10) was set up. Theonellasterone (**69**) was reduced to a mixture of the two 4-methyl

diasteroisomers (H_2, 10% Pt/C, THF/MeOH), **71** and **72**, which were efficiently separated by HPLC. In the ^{1}H NMR spectrum of **72**, H-4 was observed as a double quartet (2.05, dq, J = 14.2, 6.1 Hz) and the large coupling constant with H-5 clearly pointed towards its axial position, thus implying the α-orientation of the methyl group at C-4. Chemical correlation gave definitive confirmation of the above stereochemical assignment. As depicted in Scheme 10, for the concomitant steric effect played by Me-19 and Me-30, both orientated on the β-face of the steroidal nucleus, $NaBH_4$ reduction of **71** afforded exclusively 3β-hydroxy-4β-methyl steroisomer (**70**). On the other hand, reduction of **72** gave a mixture of 3β-hydroxy-4α-methyl theonellasterol derivative (**73**) with its C-3 epimer, 3α-hydroxy-4α-methyl- derivative (**74**). As previously reported for several natural and synthetic 4-methyl cholestane derivatives,[155] in the 4α-methyl-3β-ol derivative (**73**), the 3α-proton resonance is consistently shifted upfield with respect to the corresponding resonances in the 4α-methyl-3α-ol (**74**) and 4β-methyl-3β-ol diasteroisomers (**70**) (δH 2.93 in **73**, δH 3.55 in **74**, δH 3.54 in **70**), thus substantiating the stereochemical assignment reported in Scheme 10. Moreover in the ^{1}H NMR of **74**, H-3 was observed as a broad multiplet, allowing its assignment as equatorial and therefore establishing the orientation of the hydroxyl group at C-3 on the α-face of the molecule. To access the 4β-methyl-3α-ol derivative, the 4β-methyl theonellasterol derivative (**70**) was subjected to a two step sequence involving treatment with tosyl chloride in pyridine followed by potassium acetate in DMF/H_2O (Scheme 10). Unfortunately the basic treatment of the 3-*O*-tosyl intermediate produced β-elimination with the formation of derivative **75** with the 3,4 double bond. Nevertheless, **75** could be instrumental in the evaluation of the pharmacophoric role played by the oxygen atom on ring A.



Scheme 10. Modification at C-4 double-bond. Reagents and conditions: (**a**) H_2, 10% Pt/C, THF:MeOH 1:1; (**b**) $NaBH_4$, absolute MeOH, 0 °C; (**c**) TsCl, pyr, rt; (**d**) CH_3COOK, DMF:H_2O 7:1, reflux, 75%, over two steps.

Finally we decided to investigate the effects of the introduction of a polar group at C-4 in the binding of theonellasterol (**38**) in FXR-LBD. Oxidative cleavage with ozone (O_3, CH_2Cl_2, −78 °C, 5 min) followed by dimethylsulfide or $NaBH_4$ work-up afforded the 4-keto derivative (**76**) and the 4-hydroxy derivative (**77**), respectively (Scheme 11). The presence in the ^{13}C NMR spectrum of a signal at δC 212.1 clearly inferred the presence of a ketone at C-4 in **76** that was also confirmed by the chemical shift value of the H-3 resonance, shifted downfield with respect to **70** (δH 3.80 in **76**, δH 3.54 in **70**). In agreement with the steric influence played by Me-19, ^1H NMR spectrum analysis revealed that the sodium borohydride work-up proceeded in a stereoselective manner affording the exclusive formation of 4β-hydroxy derivative (**77**) as judged by the shape of H-4 as a broad singlet. This is consistent with an equatorial disposition for this proton, and therefore with the axial β-orientation of the hydroxyl group. It was confirmed by the strong downfield shift exhibited by Me-19 (δH 1.12 in **77**, δH 0.63 in **38**) caused by the 1,3-diaxial relationship with the hydroxy group at C-4.

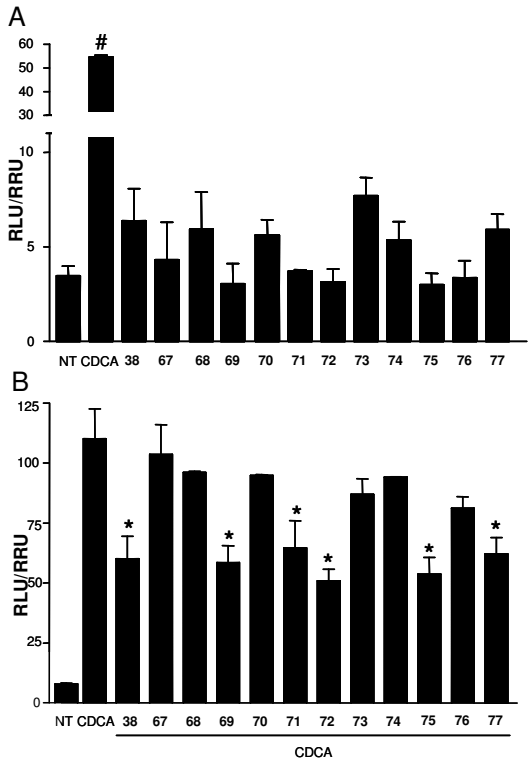


76 **38** **77**

Scheme 11. Modification at C-4 double-bond. Reagents and conditions: (**a**) O_3 solution in CH_2Cl_2, −78 °C, then DMS, 84%; (**b**) O_3 solution in CH_2Cl_2, −78 °C, then $NaBH_4$, 93%.

4.4.1 Pharmacological evaluation *in vitro*

All derivatives of this small library were tested *in vitro* on FXR. HepG2 cells were stimulated with compounds theonellasterol and with its semi-synthetic derivatives (**67–77)** in the presence or in the absence of CDCA (10 μM). As shown in Figure 69A, none of these compounds appears to be an FXR agonist in the transactivation assay. However, when HepG2 cells transfected with FXR vectors were treated with compounds **67–77** (Figure 69B) in the presence of 10 μM CDCA, several derivatives, in accordance with the antagonistic behaviour of theonellasterol (**38**), showed inhibitory activity against FXR transactivation induced by CDCA.



C

Compound	% inhibition	Compound	% inhibition
1	45.5 ± 14.5	7	53.7 ± 7.5
2	5.8 ± 19.4	8	21.0 ± 10.1
3	12.7 ± 0.7	9	14.5 ± 0.5
4	46.9 ± 11.0	10	51.2 ± 10.8
5	13.9 ± 0.6	11	26.35 ± 7.5
6	41.3 ± 17.8	12	43.6 ± 10.8

Figure 69. (**A**) Luciferase reporter assay performed in HepG2 transiently transfected with pCMVSPORT-FXR, pSG5-RXR, pGL4.70-Renilla, and p(hsp27)TKLUC vectors and stimulated 18 h with chenodeoxycholic acid (CDCA), 10 μM, and compounds **38, 67**–**77**, 10 μM. # $P < 0.05$ *vs.* NT ($n = 4$); (**B**) Luciferase reporter assay performed in HepG2 transiently transfected with pCMVSPORT-FXR, pSG5-RXR, pGL4.70-Renilla, and p(hsp27)TKLUC vectors and stimulated 18 h with CDCA, 10 μM, alone or in combination with compounds **38, 67**–**77**, 50 μM. * $P < 0.05$ *vs.* CDCA ($n = 4$); (**C**) Antagonism reported as percent of inhibition normalized to CDCA as 100%.

Among these synthetic derivatives, compounds (**69**, **71**, **72**, **77**) were judged the most active and their relative efficacy in inhibiting FXR transactivation caused by CDCA was measured in a luciferase reporter assay. Data shown in Figure 70 demonstrated that, in comparison to theonellasterol (**38**) (EC_{50} approximately 50 μM), the selected derivatives had an EC_{50} ranging from 35 to 50 μM with a relative potency in inhibiting FXR transactivation similar to that of the parent theonellasterol (**38**) (*i.e.*, 50%–60%).

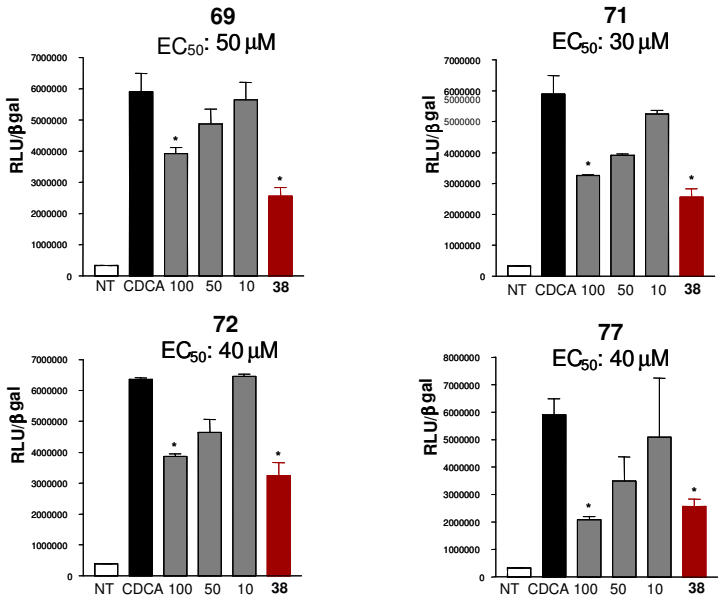


Figure 70. Luciferase reporter assay on HepG2 transiently transfected with pCMVSPORT-FXR, pSG5-RXR, pGL4.70-Renilla, and p(hsp27)TKLUC vectors and stimulated with 10 µM CDCA alone (black bar), or in combination with increasing concentrations (10, 50 and 100 µM, grey bars) of derivatives **69**, **71**, **72**, **77** and with theonellasterol (**38**) 100 µM (red bar). # $P < 0.05$ *vs.* CDCA ($n = 3$).

4.4.2 Docking studies

To clarify the effects of the chemical modifications at C-3 and at C-4 of theonellasterol (**38**) at the atomic level, molecular docking calculations[108] were performed. In our three dimensional models (Figure 71), all synthetic derivatives **67**–**77** adopt the same positioning in the FXR-LBD with respect to the parent compound theonellasterol (**38**) maintaining hydrophobic contacts of their tetracyclic cores with the receptor (Leu345, Met287, Met325, Met262, Ser329, Trp466), and interacting in a different manner with the key aminoacids Tyr358, His444, and Trp466 (Figure 71).[38]

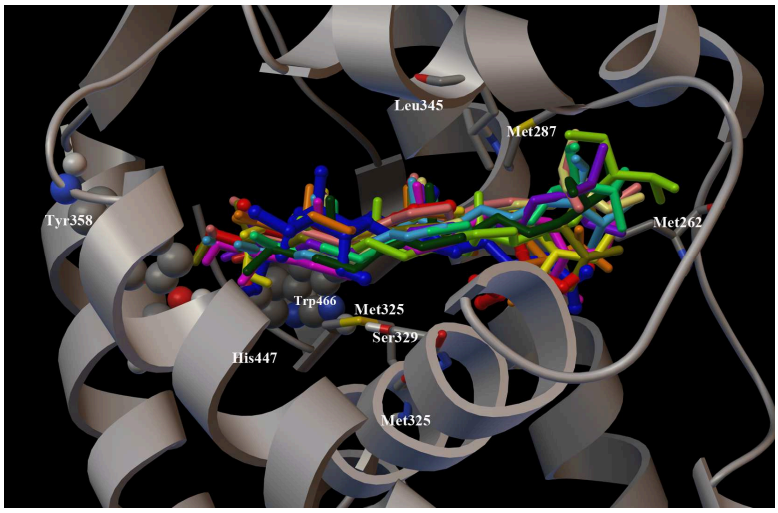


Figure 71. Superimposition of the binding modes of **38, 67**–**77**, and of the co-crystallized agonist 6-ECDCA (**61**) in the FXR-LBD. The chain A of FXR (pdb code: 1OSV) and the key amino acids (see main text for details) are represented by grey ribbon, cpk and stick and balls respectively colored by atom type (C, grey; O, red; N, dark blue; S, yellow; H, light grey). 6-ECDCA (dark blue), **38** (red), **67** (light yellow), **68** (green), **69** (emerald green), **70** (olive green), **71** (purple), **72** (dark green), **73** (pink), **74** (yellow), **75** (orange), **76** (cyan), and **77** (light pink) are represented by stick and balls.

Interestingly, when a hydroxyl group was introduced in place of the exomethylene functionality (Figure 72, panels B and C), the resulting derivative **77** showed an antagonistic activity comparable with **38** thus suggesting a role of hydrogen bond donor for this group in FXR-LBD as confirmed by the inactivity of derivative **76** with a ketone functionality at C-4. In fact, with respect to **38** and **76**, in the derivative **77** the OH group forms an additional hydrogen bond with the side chain of Met447 (Figure 72A). Concerning the C-3 hydroxyl group, the complete inefficacy of derivatives **67** and **68** (Figure 72, panels B and C), confirms its role as hydrogen bond donor as demonstrated for **38** (Figure 72B).[153] On the other hand, theonellasterone (**69**) maintained the antagonistic activity, with the ketone functionality at C-3 acting as H bond acceptor through the interactions with the Tyr358 and the His444 as for theonellasterol. As shown in Figure 72B these contacts are hampered in **67** and **68** for the steric effects played by the methyl or acetyl substitution, respectively. Moreover, the binding of theonellasterone (**69**) in

the FXR-LBD is also stabilized by additional interactions, mainly a carbonyl-π contact with Tyr358 and His444 and an exomethylene-π interaction with Tyr358 of the aromatic pocket formed by His444, Phe326, Phe363, Tyr358, Tyr366, Trp466 and Trp451 (Figure 72B).

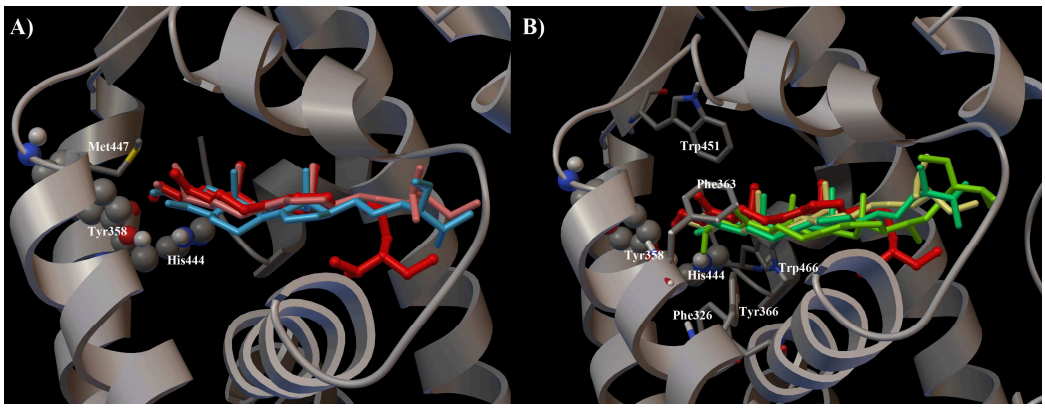


Figure 72. (**A**) Superimposition of the docking poses of **38** (red), **76** (cyan), and **77** (light pink) in the FXR-LBD. (**B**) Three dimensional model of the different interaction pattern of **38** (red), **67** (light yellow), **68** (green), **69** (emerald green), with FXR. In both figures the crucial amino acids (see main text for details) of the receptor are depicted by grey ribbon, cpk, and stick and balls respectively colored by atom type (C, grey; O, red; N, dark blue; S, yellow; H, light grey).

These interactions may be also responsible for the retained FXR antagonist activity of **71** and **72** (Figure 73A) with respect to the 4-methyl theonellasterol derivatives (**70**, **73** and **74**), which were found to be inactive towards FXR (Figure 73B). In other words these data suggest that in the 3-keto derivatives the antagonistic activity is also retained when the exomethylene at C-4 is replaced by a methyl group. Further on, the presence of the above aromatic pocket in the LBD is also responsible for the ability of the derivative **75** to antagonize CDCA in FXR transactivation suggesting that modifications at C-3 and C-4 with concomitant elimination of the hydroxyl and the exomethylene group could produce active compounds if the π interactions with Tyr358, His444, and Trp451 are maintained (Figure 73C). Therefore, the ability to form hydrogen bonds and π interactions

with the FXR-LBD seems to represent the main driving force to obtain stable and efficient sterol-receptor complexes.

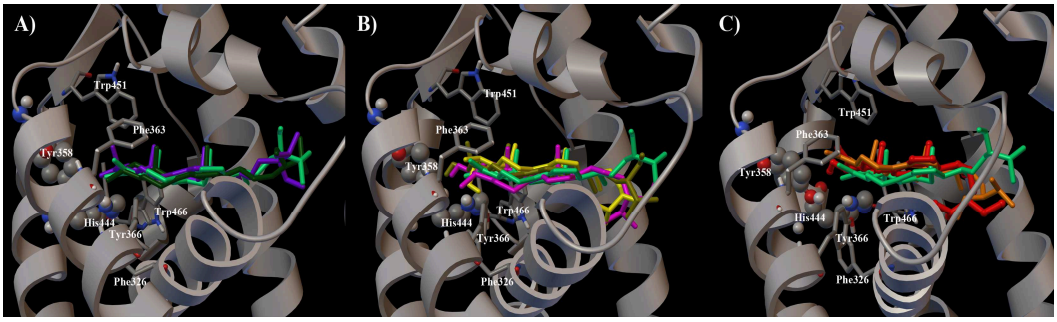


Figure 73. (**A**) Superimposition of the binding modes of **69** (emerald green) with **71** (purple) and **72** (dark green). (**B**) **69** (emerald green) with **70** (olive green), **73** (pink) and **74** (yellow). (**C**) **69** (emerald green) with **38** (red) and **77** (orange) in the FXR-LBD. In all the figures the crucial aminoacids (see main text for details) of the receptor are depicted by grey ribbon, cpk, and stick and balls respectively colored by atom type (C, grey; O, red; N, dark blue; H, light grey).

CHAPTER 5

Stereochemical studies of Perthamide C

The study of the polar n-BuOH extract of *Theonella swinhoei* sponge afforded several new cyclic peptides, perthamides C-K.[74,75,156] Perthamide C was isolated in gram scale in our laboratories and when it was tested in a well characterised model of inflammation in vivo, i.e., mouse paw oedema,[157] significantly reduced carrageenan-induced paw oedema, thus displaying potent dose-dependent anti-inflammatory activity. Structurally, perthamide C comprises a 25-membered macrocycle with several non-proteinogenic amino acid residues, such as *γ*-methylproline, *β*-hydroxyasparagine, *o*-tyrosine and 3-amino-2-hydroxy-6-methylheptanoic acid (AHMHA). Whereas the configuration of non-conventional residues was secured through an integrated approach, which combined NMR analysis, chemical degradation, stereoselective synthesis and LC/MS analysis, the remaining stereochemical uncertainty of perthamide C concerns the configuration of the two stereogenic centres in the AHMHA unit. Previously a *threo* relative disposition of the hydroxyl and the amino group on the α- and β-position, respectively, was proposed on the basis of the *J* coupling analysis, whereas, due to the lack of adequate standards the absolute configuration remained unassigned. Subsequently, the results of a stereochemical investigation on the AHMHA unit in perthamides allowed the revision of the relative configuration originally proposed and the unambiguous definition of the absolute configuration (Figure 74).[158]

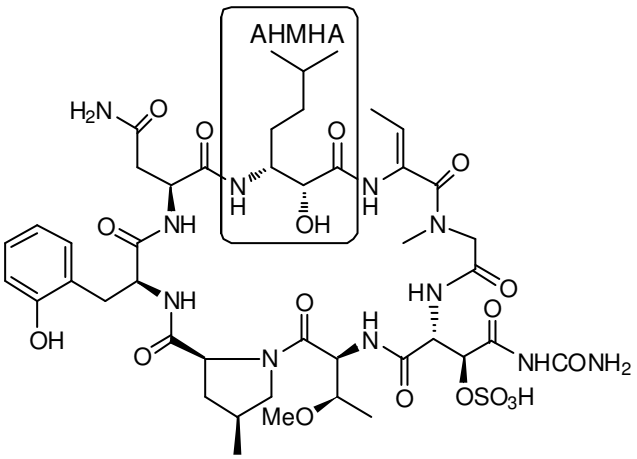


Figure 74. Revised structure for perthamide C

5.1 Application of quantitative QM-*J* method

Previously,[74] the application of the conventional *J*-based NMR method[159] evidenced some intrinsic difficulties, arising from the qualitative comparison of the predicted and some approximation in the measurement of the experimental coupling constants, which had been determined on the basis of the sole analysis of phase sensitive-HMBC spectra. For this reason an accurate determination of the heteronuclear *J* values, was afforded through the interpretation of a 2D-HETLOC[160] experiment. In addition, for the comparison of the experimental couplings, the more reliable from a quantitative point of viewd QM-NMR integrated approach[161,162,163] was applied. In particular, we considered for our calculations the *threo* and the *erythro* arrangements of a simplified fragment representative of the AHMHA unit (Figure 75), containing, together with the β-aminoacid unit, a N-methyl group and an acetyl group in the carboxy and amino terminal positions, respectively.

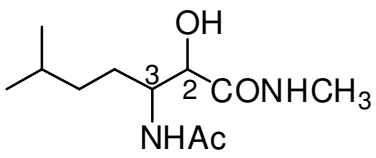


Figure 75. Molecular fragment representing AHMHA; calculations for *erythro* and *threo* arrangements were performed varying C2 and C3 centres.

Following the protocol previously reported in literature[161,162,163] the three staggered arrangements (*gauche* **+**, *gauche*- and *anti*) for both the *threo* and *erythro* configurations was optimized at QM chemical level. On the optimised geometries, *J* values at the MPW1PW91 level, using the 6-31G(d,p) basis set were calculated[164] and the results of the calculation compared to the experimental data. Surprisingly, differently from what was found using the *J*-based method,[74] the quantitative QM-*J* method evidenced a better superimposition between the experimental and calculated data for the *erythro* arrangement. In fact, as displayed in Table 7, the lowest total absolute deviation (TAD) is observed for the *erythro gauche***+** arrangement (4.8 Hz), even though a comparable value of 6.4 Hz is observed for the *gauche*-conformation of the *threo* arrangement.

Table 7. Calculated and experimental *J* values (Hz) for a fragment corresponding to AHMHA in perthamide C.

	Calcd.						**Exp**.
	threo			*erythro*			
	g^+	*anti*	g^-	g^+	*anti*	g^-	
$^3J_{H2\text{-}H3}$	5.3	8.7	1.5	3.0	7.8	3.4	1.8
$^3J_{H2\text{-}C4}$	4.2	1.2	1.7	4.1	4.3	1.0	1.9
$^3J_{H3\text{-}C1}$	6.4	1.9	1.0	0.4	0.6	6.7	0.7
$^2J_{H2\text{-}C3}$	-1.2	-1.5	1.8	-3.1	-3.0	0.5	-2.5
$^2J_{H3\text{-}C2}$	-2.0	-3.0	1.3	-0.6	-3.7	-2.3	0.0
TAD[a]	14.9	12.7	6.4	4.8	12.6	13.7	

[a]Total absolute deviation values (Σ|Jcalcd.-Jexp.|).

5.2 Stereoselective synthesis of AHMHA

Definitive confirmation to above results concerning the relative arrangement and for the complete and unambiguous definition of the absolute configuration of AHMHA residue, the stereoselective synthesis of all diastereomeric possibilities for the AHMHA residue was undertaken.[165] For our purposes, *trans* α-β-epoxy esters, which are easily obtained in optically enriched form (ee >90%) by Sharpless asymmetric epoxidation (AE), were judged valuable intermediates to

access to both *erythro* and *threo* adducts. As shown in our retrosynthetical analysis (Figure 76), the key intermediate epoxy acid **79** would arise from allylic alcohol **80**, which in turn could be prepared through a HWE olefination on aldehyde **81** and subsequent chemoselective reduction.

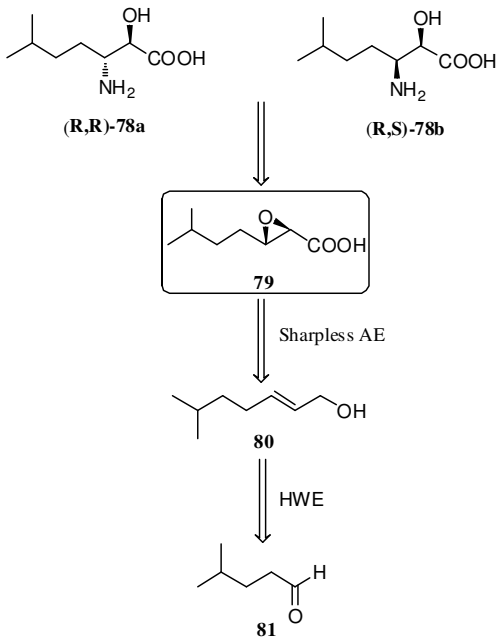
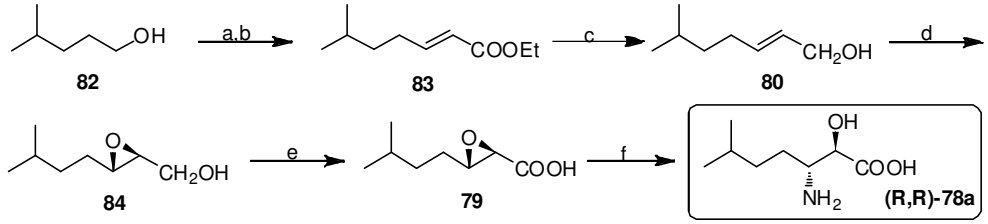


Figure 76. Retrosynthetic analysis of (2*R*,3*R*)- and (2*R*,3*S*)-3-amino-2-hydroxy-6-methylheptanoic acids (**78a** and **b**).

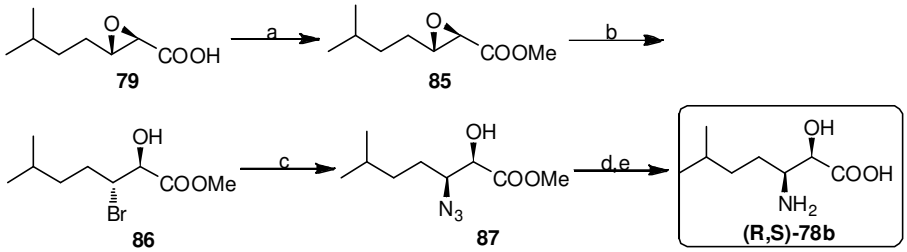
The synthetic sequence is outlined in Scheme 12, starting with commercially available alcohol **82**. Primary alcohol **82** was submitted to oxidation under Swern condition[166] and the unpurified aldehyde was subjected to standard HWE two-carbon homologation giving (E)-unsaturated ester **83** in 72% yield (>98% de, as judged by NMR data) over two steps. Chemoselective reduction with DIBAL-H to allylic alcohol **80**, followed by AE, afforded epoxy alcohol **84** in 76% yield over two steps, whose optical purity was judged to be >98% by the application of the modified Mosher's method. Treatment of (2*S*,3*S*)-**84** with ruthenium chloride and sodium periodate[167] produced carboxylic acid (2*R*,3*S*)-**79** in nearly quantitative yield. With optically active key intermediate **79** on our hand, the

synthesis of *erythro* diastereoisomer was completed through the one-pot metal catalyzed azidolysis/reduction of α-β-epoxycarboxylic acids developed by Fringuelli et al.[168] Treatment of **79** with sodium azide in the presence of 10 mol % $Cu(NO_3)_2$, followed by in situ $NaBH_4$ reduction furnished *erythro* α-hydroxy-β-amino-6-methylheptanoic acid **78a** in a 75% yield (Scheme 12). The reaction proceeded smoothly with excellent regio- and stereoselectivity as judged by NMR data of **78a** when compared with those reported for *allo*-ethylnorstatine.[168]



Scheme 12. Reagents and conditions: (a) $COCl_2$, DMSO, TEA, CH_2Cl_2 dry, -78 °C; (b) TEPA, LiOH, THF dry, 72% two steps; (c) DIBAL-H, toluene dry, -78 °C; (d) $Ti(Oi\text{-}Pr)_4$, (+)-DET, TBHP,4 A MS, CH_2Cl_2 dry, -20 °C, 76% two steps; (e) $NaIO_4$, $RuCl_3$, $CH_3CN/CCl_4/H_2O$ 2:2:3; (f) NaN_3, $Cu(NO_3)_2$, H_2O, 65 °C, then $NaBH_4$, 0 °C, 75% two steps.

Using Bonini's methodology,[169] *threo* diastereomer was synthesized with similar stereoselectivity and chemical yields starting from α-β -epoxycarboxylic acid **79** (Scheme 13). Diazomethane esterification produced *trans* epoxy methyl ester **85** that in turn was subjected to $MgBr_2$-mediated epoxy-opening to give bromohydrin **86**. Azide substitution followed by hydrogenation and acid hydrolysis produced the targeted *threo* α-hydroxy-β-amino acid **78b**. Also in this case the reaction sequence proceeded in good yields and in highly regio- and stereoselective manner in all steps.



Scheme 13. Reagents and conditions: (a) CH_2N_2, Et_2O, 82% two steps from **84**; (b) $MgBr_2{\cdot}OEt_2$, Et_2O, 98%; (c) NaN_3, DMF, 65 °C, 92%; (d) H_2, Pd/C; (e) 6N HCl, 120 °C, 75% two steps.

Comparison of NMR spectra of the AHMHA obtained from the acid hydrolysis of perthamide C with those of the two synthetic analogues unambiguously indicates the configurational assignment of the natural fragment as either *R*,*R* or *S*,*S* (Table 8). Besides slight differences in the chemical shift values of the methyl signals of (*R*,*S*)-**78b** and the natural fragment, the ^{1}H NMR spectrum of (*R*,*S*)-**78b** displays two well distinguished proton signals (1.58 and 1.80 ppm) for the diastereotopic methylene at C-4 where the natural fragment displays a single broad 2H multiplet (1.62e1.69 ppm). Also ^{13}C NMR data confirmed the stereochemical assignment made. Once again, while the spectrum of diastereoisomer **78a** completely matches that of the natural material, ^{13}C NMR data for *threo* stereoisomer differ especially in the chemical shift of carbon at position-4.

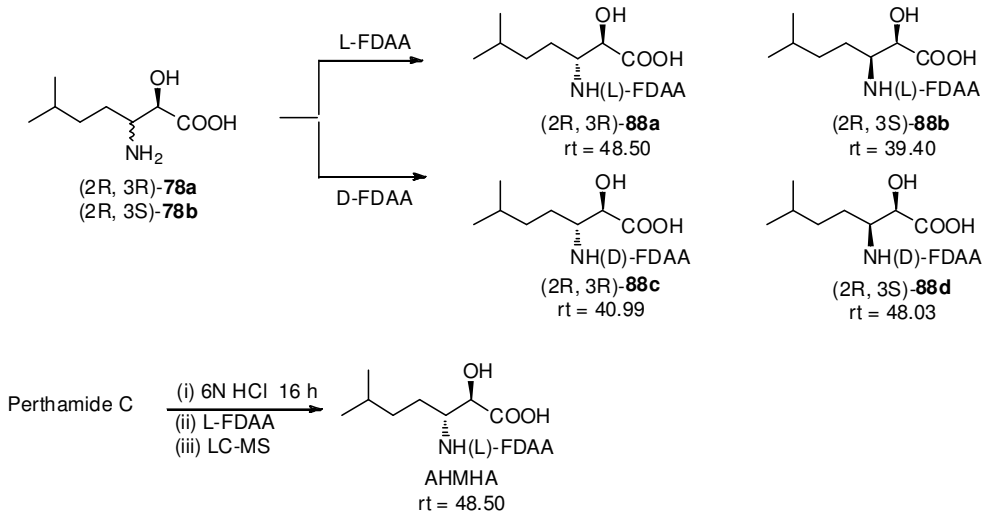
Table 8. Comparison of the NMR data of synthetic 2-hydroxy-3-amino-6-methylheptanoic acid diastereomers **78a- 78b** with the natural fragment.[a]

	2*R*,3*R*-**78a** *δ*, multiplicity (*J*), H -		2*R*,3*S*-**78b** *δ*, multiplicity (*J*), H		natural fragment *δ*, multiplicity (*J*), H	
	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1		172.7		172.1		172.5
2	4.01 m, 1H	72.5	4.00 m, 1H	71.3	4.04 m, 1H	72.5
3	3.35 m, 1H	55.7	3.31 m, 1H	55.4	3.34 m, 1H	55.7
4	1.63, 1.70 m, 2H	26.4	1.58, 1.80 m, 2H	28.4	1.62, 1.69 m, 2H	26.8
5	1.30 m, 1H	35.6	1.34 m, 1H	35.4	1.29 m, 1H	35.6
6	1.57 m, 1H	28.8	1.60 m, 1H	29.1	1.56 m, 1H	28.9
7	0.9 d (6.7), 3H	22.5	0.95 d (6.7), 3H	22.5	0.93 d (6.5), 3H	22.6
Me-6	0.92 d (6.7), 3H	22.2	0.95 d (6.7), 3H	22.5	0.92 d (6.5), 3H	22.1

[a]All chemical shifts are reported in ppm and were measured in CD_3OD. Coupling constants (*J*) are expressed in Hz.

With two diastereoisomers of AHMHA in our hands, it was then possible to proceed to the determination of the absolute configuration of this unit in perthamide C using a pre-column derivatization method. A small sample (1 mg each) of synthetic diastereomers **78a-78b** was derivatized with both enantiomers of Marfey's reagent (N-(3- fluoro-4,6-dinitrophenyl)-alaninamide; L-FDAA and D-FDAA). The L- and D-FDAA derivatives **88a-88d** were analysed using ESI LC/MS in the positive ion mode. By monitoring for FDAA/AHMHA at m/z 428, they were detected as separate peaks at 48.50, 39.40, 40.99, 48.03 min, respectively (Scheme 14). The L-FDAA derivative of AHMHA unit in perthamide C was co-eluted with the L-FDAA derivative of (2*R*,3*R*)-2-hydroxy-3-amino-6-methylheptanoic acid **78a**. Thus, the (2*R*,3*R*) configuration for the AHMHA residue in perthamide C was unambiguously established (Scheme14), as

further confirmed by the optical rotation data [synthetic **78a**: $[\alpha]_D^{20}$ +8.6 (c 0.5, MeOH); natural AHMHA: $[\alpha]_D^{20}$ +8.8 (c 0.55, MeOH)].



Scheme 14. HPLC retention times (rt) of FDAA derivatives of synthetic (2*R*, 3*R*)- and (2*R*, 3*S*)-3-amino-2-hydroxy-5-methylheptanoic acid (**78a** and **78b**) and configurational assignment of AHMHA in perthamide C.

CONCLUSIONS

Living marine organisms, plants, animal and micro-organisms, were appeared in the sea more than 3500 million years ago and, in order to survive in a hostile milieu, they have developed exquisitely complex biological mechanisms able to produce a wide variety of secondary metabolites unlike those found in terrestrial species. Marine natural products cover a far greater area of chemical space than synthetic compounds, and have property distributions that are similar to those of drugs currently in use. When compared to synthetic compounds, marine products, on average, have higher molecular weights, are sterically more complex, with more bridgehead atoms, rings, and chiral centres. This simple observation can give an idea of the incredible potential represented by marine organisms. Secondary metabolites contained in these organisms are the result of millions of years of evolution and natural selection: even a single species constitutes a library of metabolites that is validated for the bioactivity. As the results of enzymatic reactions, natural products have an intrinsic capacity to recognize and bind macromolecules, perturb their activity, and modulate biological processes. Beside in the past century the high-throughput screening of natural sources has long been recognized as an invaluable source of new lead structures, today targeted oriented discovery, focused on the identification of natural products as ligands of specific proteins or enzymes, is considered the best rationale approach for the identification of novel therapeutic agents from Nature. Indeed natural products are being biosynthesized by their hosts to interact with proteins, such as enzymes or receptors, and many human protein targets contain structural domains similar to the targets with which small ligands (or natural products) have coevolved. Surely,

nuclear receptors (NRs) represent one of the most important drug targets in terms of potential therapeutic application, playing a role in every aspect of development, physiology and disease in humans. They are ubiquitous in the animal kingdom, suggesting that they may have played an important role in their evolution, and today are recognized important target for the discovery of new therapeutic opportunity in inflammation, cancer and several metabolic diseases.

In this context my research work, described in this PhD thesis, has been addressed to the identification of new ligands of nuclear receptors, discovering potent and selective modulators of farnesoid-X-receptor (FXR) and pregnane-X-receptor (PXR). Specifically, structural characterization and biochemical decodification on steroidal composition of the marine sponge *Theonella swinhoei* allowed the identification, for the first time, of several molecules with promising activity on human nuclear receptors. Depending on the side chain and on the substitution of the tetracyclic nuclei, these molecules showed a pharmacological activity ranging from pure and potent PXR agonism to FXR antagonism, through dual modulation on these NRs. Solomonsterols are sulfated sterols with agonistic activity on the PXR and judged potential leads in inflammatory bowel diseases (IBDs), such as Crohn's disease (CD) and ulcerative colitis (UC). In animal model, solomonsterol A effectively protects against the development of signs and symptoms of colitis by a mechanism involving inhibition of NF-κB. In addition, solomonsterol A attenuated the release of TNFα in the inflamed tissue and induced the expression of IL-10 and TGFβ, two key counter-regulatory cytokines. Total syntheses of the natural leads, as well as design and synthesis of several side chain modified derivatives, have been realized during my research work providing sufficient amounts for pharmacological characterization and discovery of cholestan disulfate

(Coldisolf) as a new, simplified agonist of PXR, currently in pharmacological evaluation on animal models of liver fibrosis induced by HIV infection.

Theonellasterol was identified as the first example of a sponge derived highly selective FXR antagonist demonstrating its pharmacological potential in the treatment of cholestasis. Using this compound as a novel FXR antagonist hit, a series of semi-synthetic derivatives has been prepared in order to gain insights into the structural requirements for exhibiting antagonistic activity. These molecules could be used for the pharmacological treatment of cholestasis but also in chemotherapy of human carcinoma characterized by over-expression of FXR.

Finally a library of 4-methylene-polyhydroxylated steroids have been isolated and identified as the first example of dual PXR/FXR ligands. Because PXR ligands are not ligand for FXR, the discovery of ancestral dual activators highlights the potential for the existence of similar ligands in the mammalian body. From the pharmacology stand point, a dual ligand holds potential in the treatment of liver disorders characterized by cholestasis and/or impaired metabolism of xenobiotics such as primary biliary cirrhosis (PBC) and sclerosing cholangitis (PSC).

In conclusion, this study reaffirms *Nature* as one of the best sources of new chemical entities representing an essential component of today's research arsenal to shed light on complex biological processes, biochemical pathways and, in prospective, an inspiration for the discovery of new therapeutic strategies.

EXPERIMENTAL SECTION

I. General experimental procedures.

Chemistry.

Specific rotations were measured on a Jasco P-2000 polarimeter. High-resolution ESI-MS spectra were performed with a Micromass Q-TOF mass spectrometer. ESI-MS experiments were performed on an Applied Biosystem API 2000 triple-quadrupole mass spectrometer. NMR spectra were obtained on Varian Inova 400, Varian Inova 500 and Varian Inova 700 NMR spectrometers (^{1}H at 400, 500 and 700 MHz, ^{13}C at 100, 125 and 175 MHz, respectively) equipped with a SUN micro system ultra5 hardware and recorded in $CDCl_3$ (δ_H =7.26 and δ_C =77.0 ppm), in CD_3OD (δ_H =3.30 and δ_C =49.0 ppm) and in C_6D_6 (δ_H =7.16, δ_C =128.4). All of the detected signals were in accordance with the proposed structures. Coupling constant (J values) are given in Hertz (Hz) and chemical shifts (δ) are reported in ppm and referred to $CHCl_3$, CHD_2OD and C_6HD_5 as internal standards. Spin multiplicities are given as: s (singlet), br s (broad singlet), d (doublet), or m (multiplet). HPLC was performed with a Waters Model 510 pump equipped with Waters Rheodine injector and a differential refractometer, model 401. The purity of all of the intermediates, checked by ^{1}H NMR and HPLC, was greater than 95%. Through-space ^{1}H connectivities were evidenced using a ROESY experiment with mixing times of 200 and 500 ms, respectively. Reaction progress was monitored via thin-layer chromatography (TLC) on Alugram silica gel G/UV254 plates. Silica gel MN Kieselgel 60 (70-230 mesh) from Macherey-Nagel Company was used for column chromatography. All chemicals were obtained from Sigma-Aldrich, Inc. Solvents and reagents were used as supplied from commercial sources with the following exceptions. Tetrahydrofuran, toluene, dichloromethane, ether and triethylamine were distilled from calcium hydride immediately prior to use. All reactions were carried out under argon atmosphere using flame-dried glassware.

Biological assays.

Plasmids, Cell Culture, Transfection, and Luciferase Assays. All transfections were made using Fugene HD transfection reagent (Roche). For FXR mediated transactivation, HepG2 cells, plated in a 6-well plate at $5x10^5$ cells/well, were transfected with 100 ng pSG5-FXR, 100 ng pSG5-RXR, 200 ng pCMV-βgalactosidase and with 500 ng of the reporter vector p(hsp27)-TK-LUC containing the FXR response element IR1 cloned from the promoter of heat shock protein 27 (hsp27). At 48 h post-transfection, cells were stimulated 18 h with 10 μM CDCA or with tested compounds (50 μM) alone or in combination with CDCA. For PXR mediated transactivation, HepG2 cells, plated in a 6-well plate at $5x10^5$ cells/well, were transfected with 100 ng pSG5-PXR, 100 ng pSG5-RXR, 200 ng pCMV-βgalactosidase and with 500 ng of the reporter vector containing the PXR target gene promoter (CYP3A4 gene promoter) cloned upstream of the luciferase gene (pCYP3A4promoter-TKLuc). At 48 h post-transfection, cells were stimulated 18 h with 10 μM Rifaximin or with tested compounds (50 μM) alone or in combination with Rifaximin. Cells were lysed in 100 μL diluted reporter lysis buffer (Promega) and 0.2 μL of cellular lysates was assayed

for luciferase activity using the Luciferase Assay System (Promega). Luminescence was measured using an automated luminometer. Luciferase activities were normalized for transfection efficiencies by dividing the relative light units by β-galactosidase activity expressed from cells cotransfected with pCMV-βgal.

Real-Time PCR. Total RNA was extracted using the TRIzol reagent (Invitrogen), purified of the genomic DNA by DNAase I treatment (Invitrogen) and random reverse-transcribed with Superscript II (Invitrogen). Fifty ng template was amplified using the following reagents: 0.2 µM of each primer and 12.5 µL of 2X SYBR Green qPCR master mix (Invitrogen). All reactions were performed in triplicate and the thermal cycling conditions were: 2 min at 95 °C, followed by 40 cycles of 95 °C for 20 s, 55 °C for 20 s and 72 °C for 30 s in iCycler iQ instrument (Biorad). The relative mRNA expression was calculated and expressed as $2^{-(\Delta\Delta Ct)}$. Primers used for qRT-PCR were: mGAPDH: CTGAGTATGTCGTGGAGTCTAC and GTTGGTGGTGCAGGATGCATTG; mCyp3A11: TGAAACCACCAGTAGCACAC and CCATATCCAGGTATTCCATCTCC; mIL1β: TCACAGCAGCACATCAACAA and TGTCCTCATCCTCGAAGGTC; mIL-6: CCGGAGAGGAGACTTCACAG and TCCACGATTTCCCAGAGAAC; mTNFα: ACGGCATGGATCTCAAAGAC and GTGGGTGAGGAGCACGTAGT; rGAPDH: ATGACTCTACCCACGGCAAG and TACTCAGCACCAGCATCACC; rαSMA: GCTCCATCCTGGCTTCTCTA and TAGAAGCATTTGCGGTGGAC. Human primers were as follows: hSHP: GCTGTCTGGAGTCCTTCTGG and CCAATGATAGGGCGAAAGAAGAG; hGAPDH: GAAGGTGAAGGTCGGAGT and CATGGGTGGAATCATATTGGAA; hCYP3A4: CAAGACCCCTTTGTGGAAAA and CGAGGCGACTTTCTTTCATC; hIL1β: GGACAAGCTGAGGAAGATGC and TCGTTATCCCATGTGTCGAA; hTNFα: AACCTCCTCTCTGCCATCAA and GGAAGACCCCTCCCAGATAG; hMCP-1: CCCCAGTCACCTGCTGTTAT and TCCTGAACCCACTTCTGCTT; hB2M: TGCTATGTGTCTGGGTTTCATC and TGACAAAGTCACATGGTTCACA; hCYP3A7: CAAGACCCCTTTGTGGAAAA and TGTCTCTTTGAGGCGACCTT; hSULT2A1: GATCCAATCTGTGCCCATCT and TAAATCACCTTGGCCTTGGA; hMDR1: GTGGGGCAAGTCAGTTCATT and TCTTCACCTCCAGGCTCAGT; hBSEP: GGGCCATTGTACGAGATCCTAA and TGCACCGTCTTTTCACTTTCTG; hCYP7A1: CACCTTGAGGACGGTTCCTA and CGATCCAAAGGGCATGTAGT; hOSTα: TGTTGGGCCCTTTCCAATAC and GGCTCCCATGTTCTGCTCAC.

Computational details.

We performed molecular docking experiments by Autodock 4.2 software on 4 x AMD Opteron SixCore 2.4 GHz. We used a grid box size of 94 x 96 x 68 for chain A of FXR (pdb code:1OSV), and 90 x 106 x 92 for PXR (pdb code: 1M13) with spacing of 0.375 Å between the grid points, and centered at 20.689 (x), 39.478 (y), 10.921 (z) between the SCH_3 of Met262 and the OH group of Thr267 for FXR, and at 14.282 (x), 74.983 (y), 0.974 (z) between the ring of His407 and the side chain of Leu209 for PXR, covering the active site of both the receptors. To achieve a representative conformational space during the docking studies and for taking into account the

variable number of active torsions, 10 calculations consisting of 256 runs were performed, obtaining 2560 structures for each ligand. The Lamarckian genetic algorithm (LGA) was employed for docking experiments, choosing an initial population of 600 randomly placed individuals. The maximum number of energy evaluations and of generations was set up to 5 x 10^6 and to 6 x 10^6 respectively. For all the docked structures, all bonds were treated as active torsional bonds except the bonds in cycles, which are considered fixed together with the receptors. Results differing by less than 3.5 Å in positional root-mean-square deviation (RMSD) were clustered together and represented by the result with the most favorable free energy of binding. Illustrations of the 3D models were generated using the Python[115] and Chimera software.[137]

II. Experimental section of PXR agonists

Synthesis of solomonsterol A (1)

Methyl hyodeoxy-5-cholan-24-oate (**4**). To a solution of hyodeoxycholic acid **3** (2.5 g, 6.4 mmol) in dry diethyl ether (15 mL) a solution of diazomethane in dry diethyl ether was added until yellow persistent coloration. The mixture was stirred at room temperature for four hours. The solution was dried under vacuum without any further purification to give the methyl hyodeoxycholate **4** (2.5 g, quantitative yield). $[\alpha]_{24}^{D}$= +1.68 (*c* 6.86, $CHCl_3$); selected ^{1}H NMR (400 MHz $CDCl_3$): δ 4.06 (1H, dt, J= 4.86, 9.7 Hz), 3.63 (1H, m), 3.66 (3H, s), 2.36 (1H, m), 0.64 (3H, s), 0.90 (3H, s), 0.91(3H, d, J=5.9 Hz). ^{13}C NMR (100 MHz $CDCl_3$): δ 75.2, 71.7, 68.3, 56.4, 56.2, 54.6, 51.7, 48.7, 43.0, 40.2, 40.1, 36.1, 35.8, 35.6, 35.0, 31.8, 31.2, 30.2, 29.4, 28.3, 24.4, 21.0, 18.5, 14.1, 12.2; HRMS-ESI *m/z* 407.3173 ($[M + H]^+$, $C_{25}H_{43}O_4$ requires 407.3161.

Methyl 3,6-ditosyloxy-5-cholan-24-oate (**5**). To a solution of methyl hyodeoxycholate **4** (2.5 g, 6.1 mmol) in dry pyridine (14 ml) a solution of tosyl chloride (3.4 g, 18.3 mmol) in dry pyridine (6 ml) was added, and the mixture was stirred at room temperature for 2 days. It was poured into cold water (20 mL) and extracted with CH_2Cl_2 (3 × 30 mL). The combined organic layer was washed with saturated $NaHCO_3$ solution (30 mL), and water (30 mL), and then dried over anhydrous $MgSO_4$ and evaporated in vacuo to give 4.2 g of **5** (quantitative yield) in the form of colorless needles, that was subjected to next step without any purification. $[\alpha]_{24}^{D}$= +7.2 (*c* 6.84, $CHCl_3$) selected ^{1}H NMR (400 MHz $CDCl_3$): δ 0.56 (3H, s), 0.80 (3H, s), 0.88 (3H, d, J= 5.9 Hz), 2.46 (6H, s), 2.33 (1H, m), 2.22 (1H, m), 3.66 (3H, s), 4.30 (1H, m), 4.78 (1H, m), 7.34, 7.72 e 7.78 (8H, m). ^{13}C NMR (100 MHz $CDCl_3$): δ 174.9, 144.9 (2C), 144.0 (2C), 130.0 (4C), 127.9 (4C), 82.1, 80.0, 56.0, 55.9, 54.6, 51.7, 46.6, 43.0, 39.8, 39.6, 36.4, 35.5, 35.0 (2C), 32.3, 31.2, 31.1, 28.2, 26.7, 24.2, 21.9 (2C), 20.7, 18.4, 14.4, 12.2. HRMS-ESI *m/z* 715.3350 ($[M + H]^+$, $C_{39}H_{55}O_8S_2$ requires 715.3338.

Methyl 3-hydroxy-5-cholen-24-oate (**6**). A solution of the ditosylate **5** (4.2 g, 5.9 mmol) and CH_3COOK (579 mg, 5.9 mmol) dissolved in water (3 ml) and *N*,*N*'- dimethylformamide (DMF 27 ml) was refluxed for 4h. The solution was cooled at room temperature and poured into cold water (20 mL) and extracted with ethyl acetate (3 × 30 mL). Then, it was dried over anhydrous $MgSO_4$ and evaporated in vacuo to give a oleos residue that was purified on a silica gel column by eluting with hexane-EtOAc (9:1, 0.5% TEA) to obtain 2.0 g of **6** (78% yield); $[\alpha]_{24}^{D}$= -9.0 (*c* 0.73,

$CHCl_3$). Selected ^1H NMR (400 MHz $CDCl_3$): δ 5.34 (1H, d, J= 4.8 Hz), 3.65 (3H, s), 3.51 (1H, m), 2.33 (1H, m), 2.23 (1H, m), 1.00 (3H, s), 0.92 (3H, d, J= 6.3 Hz), 0.67 (3H, s). ^{13}C NMR (100 MHz $CDCl_3$): δ 174.8, 140.8, 121.5, 71.6, 56.6, 55.7, 51.4, 50.0, 42.2, 41.1, 39.7, 37.2 (2C), 35.2, 31.7 (2C), 31.4, 30.9, 30.8, 28.0, 24.2, 20.9, 19.3, 18.2, 11.7; HRMS-ESI *m/z* 389.3062 ($[M + H]^+$, $C_{25}H_{41}O_3$ requires 389.3056.

Methyl 3β-hydroxy-5α-cholan-24-oate (**7**). An oven-dried 250 mL flask was charged with 10% palladium on carbon (100 mg) and compound **6** (2.00 g, 5.15 mmol) and the flask was evacuated and flushed with argon. Absolute methanol (50 mL) and dry THF (50 mL) were added, and the flask was flushed with hydrogen. The reaction was stirred at room temperature under H_2 (1 atm) for 4 h. The mixture was filtered through celite, and the recovered filtrate was concentrated to give 1.70 g of crude product. The residue was subjected to column chromatography on silica gel eluting with hexane-EtOAc (9:1, 0.5% TEA) to give 1.60 g of pure **7** (80%). $[\alpha]_{24}^D$= +3.4 (*c* 0.54, $CHCl_3$); selected ^1H NMR (400 MHz, $CDCl_3$): δ 3.64 (3H, s), 3.56 (1H, m), 2.33 (1H, m), 2.19 (1H, m), 0.89 (3H, d, J= 6.3 Hz), 0.78 (3H, s), 0.63 (3H, s). ^{13}C NMR (100 MHz, $CDCl_3$): δ 174.8, 71.3, 56.4, 55.8, 54.3, 51.5, 44.8, 42.6, 41.1, 38.0, 36.9, 35.4, 35.3 (2C), 32.0, 31.3, 31.0, 30.9, 28.6, 28.1, 24.1, 21.2, 18.2, 12.5, 12.3; HRMS-ESI *m/z* 391.3227 ($[M + H]^+$, $C_{25}H_{43}O_3$ requires 391.3212.

Methyl 3β-tosyloxy-5α-cholan-24-oate (**8**). To a solution of **7** (1.50 g, 3.84 mmol) in dry pyridine (30 mL), *p*-toluenesulfonyl chloride (2.20 g, 11.52 mmol) was added. The solution was stirred at room temperature for 24 h and then poured into cold water (20 mL) and extracted with CH_2Cl_2 (3 × 15 mL). The combined organic layer was washed saturated $NaHCO_3$ solution (30 mL), and water (30 mL), and then dried over anhydrous $MgSO_4$ and evaporated in vacuo to give 2.09 g of **8**, that was subjected to next step without any purification. $[\alpha]_{24}^D$= +15.54 (*c* 0.47, $CHCl_3$) Selected ^1H NMR (400 MHz $CDCl_3$): δ 7.79 (2H, d, J= 8.2 Hz), 7.33 (2H, d, J= 8.2 Hz), 4.41 (1H, m), 4.23 (1H, m), 3.66 (3H, s), 2.45 (3H, s), 2.34 (1H, m), 2.22 (1H, m), 0.89 (3H, d, J= 5.9 Hz), 0.77 (3H, s), 0.63 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 174.8, 148.8, 129.8 (2C), 127.7 (2C), 124.3, 82.7, 56.6, 56.1, 54.2, 51.7, 45.0, 41.6, 40.1, 37.0, 35.5 (2C), 35.2, 35.0, 32.1, 31.3, 31.2, 29.9, 28.6, 28.3, 24.4, 21.9, 21.4, 18.5, 14.4, 12.3; HRMS-ESI *m/z* 545.3311 ($[M + H]^+$), $C_{32}H_{49}O_5S$ requires 545.3301.

Methyl 5α-chol-2-en-24-oate (**9**). Lithium bromide (3.30 g, 38.4 mmol) and lithium carbonate (2.8 g, 38.4 mmol) were added to a solution of **8** (2.08 g, 3.84 mmol) in dry DMF (150 mL), and the mixture was refluxed for 1.5 h. After cooling to room temperature, the mixture was slowly poured into 10% HCl solution (150 mL) and extracted with CH_2Cl_2 (3 × 150 mL). The combined organic layer was washed successively with water, saturated $NaHCO_3$ solution and water, and then dried over anhydrous $MgSO_4$ and evaporated to dryness to give a white solid residue that was purified on a silica gel column by eluting with hexane-EtOAc (9:1, 0.5% TEA) to obtain pure **9** (1.19 g, 83% over two steps). $[\alpha]_{24}^D$= +18.6 (*c* 2.07, $CHCl_3$); selected ^1H NMR (400 MHz $CDCl_3$): δ 5.48 (2H, br s), 2.25 (1H, m), 2.11 (1H, m), 0.83 (3H, d, J= 6.2 Hz), 0.65 (3H, s), 0.56 (3H, s). ^{13}C NMR (100 MHz $CDCl_3$): δ 175.2, 126.2 (2C), 56.7, 56.0, 54.2, 51.7, 42.8, 41.6, 40.2, 39.9, 35.8,

35.5, 34.3, 32.0, 31.3, 31.1, 30.5, 28.9, 28.3, 24.4, 21.1, 18.5, 12.2 (2C); HRMS-ESI *m/z* 373.3117 ($[M + H]^+$, $C_{25}H_{41}O_2$ requires 373.3107.

Methyl 2α,3α-epoxy-5α-cholan-24-oate (**10**). To a solution of **9** (1.09 g, 2.93 mmol) in CH_2Cl_2 (66 mL) were added water (46.2 mL) and Na_2CO_3 (1.15 g, 10.8 mmol). The reaction mixture was stirred vigorously and *m*-chloroperbenzoic acid (708 mg, 4.1 mmol) was added slowly. The mixture was stirred for 4 h at room temperature, and then the aqueous layer was extracted with CH_2Cl_2 (3 × 35 mL). The combined CH_2Cl_2 extracts were washed successively with 5% Na_2SO_3 solution (100 mL), saturated $NaHCO_3$ solution (100 mL), and water (100 mL), dried over anhydrous $MgSO_4$ and evaporated to dryness to give 1.13 g of crude epoxide **10**, that was subjected to next step without any purification. $[\alpha]_{24}^D$= +13.7 (*c* 0.42, $CHCl_3$) Selected ^{1}H NMR (400 MHz $CDCl_3$): δ 3.65 (3H, s), 3.14 (1H, br s), 3.09 (1H, br s), 2.34 (1H, m), 2.20 (1H, m), 0.90 (3H, d, J = 6.7 Hz), 0.74 (3H, s), 0.63 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 175.2, 56.5, 56.0, 53.8, 52.6, 51.7, 51.3, 42.6, 40.2, 40.7, 38.5, 36.4, 35.8, 35.6, 31.9, 31.3, 31.2, 30.9, 28.6, 28.3, 24.4, 21.1, 18.5, 13.2, 12.2; HRMS-ESI *m/z* 389.3036 ($[M + H]^+$, $C_{25}H_{41}O_3$ requires 389.3056.

Methyl 2β,3α-dihydroxy-5α-cholan-24-oate (**11**). A solution of epoxide **10** (1.13 g, 2.93 mmol) in THF (70 mL) was treated with 1 N H_2SO_4 (7.32 mL, 7.32 mmol) solution and stirred for 24 h at room temperature. After neutralization with saturated $NaHCO_3$ solution, the mixture was evaporated to a fifth of the initial volume, diluted with water (50 mL), and extracted with ethyl acetate (3 x 50 mL). The combined organic extracts were washed with water, dried over anhydrous $MgSO_4$, filtered and evaporated to dryness. Purification on a silica gel column by eluting with CH_2Cl_2-MeOH (9:1) afforded pure **11** (870 mg, 73% over two steps). $[\alpha]_{24}^D$= +0.67 (*c* 0.46, $CHCl_3$); selected ^{1}H NMR (400 MHz $CDCl_3$): δ 3.89 (1H, br s), 3.85 (1H, br s), 3.64 (3H, s), 2.34 (1H, m), 2.20 (1H, m), 0.97 (3H, s), 0.89 (3H, d, J = 6.7 Hz), 0.63 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 175.4, 72.1, 70.9, 56.6, 56.0, 53.7, 51.7, 42.6, 40.7, 40.2, 39.2, 36.3, 35.6, 35.1, 32.1, 31.9, 31.3, 31.2, 28.4, 28.3, 24.4, 21.1, 18.5, 14.9, 12.4; HRMS-ESI *m/z* 407.3181 ($[M + H]^+$, $C_{25}H_{43}O_4$ requires 407.3161.

5α-Cholan-2β,3α,24-triol (**12**). Dry methanol (320 µL, 7.88 mmol) and $LiBH_4$ (3.94 mL, 2 M in THF, 7.88 mmol) were added to a solution of the methyl ester **11** (800 mg, 1.97 mmol) in dry THF (10 mL) at 0 °C under argon and the resulting mixture was stirred for 4 h at 0 °C. The mixture was quenched by addition of NaOH (1 M, 4 mL) and then allowed to warm to room temperature. Ethyl acetate was added and the separated aqueous phase was extracted with ethyl acetate (3 ×1 5 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and concentrated. Purification by silica gel eluting with CH_2Cl_2-MeOH (9:1) gave the alcohol **12** as a white solid (685 mg, 92%). $[\alpha]_{24}^D$= -1.2 (*c* 0.12, $CHCl_3$); selected ^{1}H NMR (400 MHz CD_3OD): δ 3.81 (1H, br s), 3.77 (1H, m), 3.67 (1H, br s), 3.52 (1H, m), 1.01 (3H, s), 0.96 (3H, d, J = 6.0 Hz), 0.70 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 73.3, 70.5, 63.1, 57.0, 56.6, 55.6, 43.0, 41.6, 40.5, 40.2, 36.0, 35.3, 35.2, 32.3, 31.7, 29.6, 29.5, 28.7, 28.6, 24.6, 21.1, 18.9, 14.2, 12.3; HRMS-ESI *m/z* 379.3154 ($[M + H]^+$, $C_{24}H_{43}O_3$ requires 378.3134.

5α-Cholan-2β,3α,24-tryl-2,3,24-sodium trisulfate (**1**). The triethylamine-sulfur trioxide complex (1.81 g, 10.0 mmol) was added to a solution of triol **12** (600 mg, 1.58 mmol) in DMF dry (2 mL) under an argon atmosphere, and the mixture was stirred at 95 °C over the week-end. Then, the reaction mixture was quenched with water (1.6 mL). The solution was poured over a silica gel column to remove excess $SO_3 \cdot NEt_3$. The product was eluted by MeOH and followed by evaporation of the solvent to yield a yellow solid [tris-(triethylammonium sulphate) salt]. To the solution of the solid in methanol (15 mL) was added Amberlite CG 120 sodium form (20 g). The mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated to obtain compound **1** as a white solid (972 mg, 90%). $[\alpha]_{24}^{D}$= +4.6 (*c* 0.8, CH_3OH); selected ^{1}H NMR (400 MHz CD_3OD): δ 4.73 (1H, br s), 4.70 (1H, br s), 3.69 (2H, m), 1.00 (3H, s), 0.95 (3H, d, J= 6.3 Hz), 0.69 (3H, s) ; ^{13}C NMR (100 MHz CD_3OD): δ 76.2, 75.6, 69.3, 57.5, 57.3, 56.4, 43.2, 41.3, 39.7, 38.5, 36.4, 36.2 (2C), 32.8, 32.6, 30.3, 28.8, 28.7, 26.9, 24.7, 21.5, 18.6, 14.4, 12.3. HRMS-ESI *m/z* 661.1415 $[M-Na]^-$, $C_{24}H_{39}Na_2O_{12}S_3$ requires 661.1399.

Specific biological assays.

Electrophoretic mobility shift assay (EMSA). The NF-κB DNA-binding activity was determined by EMSA. After treatment, nuclear and cytoplasmic extracts were prepared using the NE-PER Nuclear and Cytoplasmic Extraction Reagents (Pierce Biotechnology, Inc.). EMSA probes were created by biotinylating the 3' end of the single stranded oligonucleotides using a biotin 3' end DNA labelling kit (Pierce Biotechnology, Inc.) according to the manufacturer's protocol. The biotinylated oligonucleotides were annealed by boiling for 1 min and then allowing them to slowly cool to room temperature. The consensus nucleotide sequence for NF-κB was 5'-AGA GAT TGC CTG ACG TCA GAC AGC TAG-3'. The EMSA binding reaction was performed by utilizing a LightShift chemiluminescent EMSA kit (Pierce Biotechnology, Inc.). A nuclear extract was incubated in 1× binding reaction buffer containing 50 mM KCl, 10 mM EDTA, 25 ng/mL poly dI-dC, 5 mM $MgCl_2$, and the biotinylated probe. After a 20 min incubation at room temperature, the reaction mixture was electrophoresed on a non-denaturing 6% polyacrylamide gel and then transferred to a nylon membrane. The transferred mixture was UV-cross-linked to the membrane and detected by chemioluminescent reagents (Pierce Biotechnology, Inc.). For the competition assay, a 200-fold excess of unlabeled probe was added together with the labelled probe. For the supershift assay, 1 μg of antibody against NF-κB p65 was added together with the nuclear extract.

Animals. For the TNBS studies humanized (h)PXR mice, 8–10 weeks of age, were provided by Frank J. Gonzalez (Laboratory of Metabolism, Centre for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland). hPXR, male mice were housed in temperature- and light-controlled rooms and were given tap water and pelleted standars mouse chow *ad libitum*. The hPXR mice express the human PXR in the PXR-null background. Mice were housed under controlled temperatures (22 °C) and photoperiods (12:12-h light/dark cycle), allowed unrestricted access to standard mouse chow and tap water. The present protocol was approved by the Italian Minister Health and conforms to national guidelines. (The ID for this project is

#11/2010-B and authorization released to Prof. Stefano Fiorucci, as a principal investigator, lastly on January 25, 2010.)

Induction of colitis. Mice were fasted for 16 h and lightly anesthetized by intraperitoneal injection of 100 μL of ketamine/xylazine solution[96] for 10 g body weight and then administered intrarectally (i.r.) with the haptenating agent TNBS (0.5 mg/mouse) dissolved in ethanol 50%, via a 3.5 French (F) catheter equipped with a 1 mL syringe. The catheter was advanced into the rectum for 4 cm and then the haptenating agent was administered in a total volume of 150 μL. To ensure distribution of the agent within the entire colon and cecum, mice were held in a vertical position for 30 s. Solomonsterol A (10 mg/kg) and rifaximin (10 mg/kg) were dissolved in DMSO (10 mg/100 μL), diluted in methylcellulose 1% and administered intraperitoneally (i.p.) and orally, respectively, at the final volume of 100 μL/mouse, 3 days before the induction of colitis. At this time, the TNBS-alone group received the vehicle alone (methylcellulose 1% in a final volume of 100 μL/mouse), every day. In another experiment, a therapeutic model, solomonsterol A (10 mg/kg) was administrated the day after colitis induction. Mice were analyzed for the presence of diarrhea, body weight, and survival. The body weight was expressed as delta percentage versus the weight of mice on the day before TNBS administration. Mice were analyzed for the presence of diarrhoea, body weight, and survival. Four days after TNBS administration, surviving mice were sacrificed, colons were removed and either immediately snap-frozen in liquid nitrogen and stored at −80 °C until use or formalin fixed. The macroscopic appearance was analyzed considering the presence of indurations, oedema, thickness and evidence of mucosal haemorrhage. Grading was performed in a blinded fashion. For histological examination, tissues were fixed in 10% buffered formalin phosphate, embedded in paraffin, sectioned, and stained with hematoxylin and eosin (H&E). Histology images were captured by a digital camera (SPOT-2; Diagnostic Instruments Inc, Burroughs, MI) and analyzed by specific software (Delta Sistemi, Rome, Italy). The degree of colon inflammation was examined microscopically in transversal sections[97] and graded semi-quantitatively from 0 to 4 (0, no signs of inflammation; 1, very low level; 2, low level of leukocyte infiltration; 3, high level of leukocyte infiltration, high vascular density, and thickening of the colon wall; and 4, transmural infiltration, loss of goblet cells, high vascular density, and thickening of the colon wall). Grading was performed by blinded observers.

Colon myeloperoxidase (MPO) and cytokine levels. Colon samples were lysed in 1 mL of lysis buffer T-PER (Pierce, Rockford, USA) and finely minced. After the tissues were centrifuged at 10,000 g for 15 min at 4 °C. Colon homogenates were used to determine MPO activity, after two freeze/thaw cycles, using a spectrophotometric assay with trimethylbenzidine (TMB) as a substrate and normalized for the protein levels. Colon TNFα and IL-10 levels in tissue homogenates were quantified by ELISA (SABioscences) according to the manufacturer's instructions, and normalized for the protein levels.

Statistical analysis. All values are expressed as the mean ± SE. Number (n) of experiments or mice used in the experiments is shown. Comparisons of more than 2 groups were made with a one-

way analysis of variance with post hoc Tukey tests. Differences were considered statistically significant if p was <0.05.

Synthesis of derivatives of solomonsterol A

Diol **11** was prepared according to the synthetic procedure previously described.[89]

Methyl 2β,3α-disulfate-5α-cholan-24-oate (**13**). Triethylamine-sulfur trioxide complex (1.34 g, 7.4 mmol) was added to diol **11** (300 mg, 0.74 mmol) in DMF dry (10 mL) under an argon atmosphere, and the mixture was stirred at 95 °C for 48 h. The reaction mixture was quenched with water (1.6 mL) and the solution was poured over a C18 silica gel column to remove excess $Et_3N{\cdot}SO_3$. The product was eluted by MeOH and followed by evaporation of the solvent to yield a yellow solid [bis-(triethylammonium sulfate) salt]. To the solid dissolved in methanol (30 mL) was added Amberlite CG 120 sodium form (30 g) and the mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated to obtain compound **13** as a white solid (270 mg, 60%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 34:66, flow rate 1.5 mL/min, t_R=3.6 min). $[\alpha]_{24}^{D}$=+9.6 (*c* 0.08, CH_3OH); selected ^{1}H NMR (400 MHz CD_3OD): δ 4.69 (1H, br s), 4.66 (1H, br s), 3.61 (3H, s), 2.31 (1H, m), 2.17 (1H, m), 0.96 (3H, s), 0.90 (3H, d, J = 6.5 Hz), 0.66 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 176.5, 76.5, 76.1, 57.8, 57.4, 56.3, 51.8, 43.8, 41.4, 40.3, 39.1, 36.8, 36.5, 36.4, 33.2, 32.5, 32.4, 30.5, 29.1 (2C), 25.2, 21.9, 18.8, 14.3, 12.5. HR ESIMS *m/z* 587.1975 [M-Na], $C_{25}H_{40}NaO_{10}S_2$ requires 587.1961.

Disodium 2β,3α-disulfate-5α-cholan-24-oic acid (**14**). Compound **13** (160 mg, 0.26 mmol) was dissolved in 5% methanolic NaOH (10 mL) and water (2 mL) and the solution was refluxed for 5 h. The resulting solution was adjusted to pH 5 and concentrated under reduced pressure. Purification by C18 silica gel column eluting with H_2O:MeOH (99:1) gave the carboxylate **14** as a white solid (140 mg, 90%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 30:70, flow rate 1.5 mL/min, t_R=3.6 min). $[\alpha]_{24}^{D}$= +2.7 (*c* 0.3, CH_3OH); selected ^{1}H NMR (400 MHz CD_3OD): δ 4.73 (1H, br s), 4.69 (1H, br s), 2.31 (1H, m), 2.17 (1H, m), 0.99 (3H, s), 0.94 (3H, d, J = 6.0 Hz), 0.69 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 178.9, 76.4, 76.1, 57.8, 57.4, 56.5, 43.8, 41.4, 40.2, 39.2, 36.8, 36.4, 36.3, 33.2, 32.5 (2C), 30.5, 29.2 (2C), 25.2, 22.0, 18.7, 14.3, 12.6. ESIMS *m/z* 595.1636 $[M-Na]^-$, $C_{24}H_{37}Na_2O_{10}S_2$ requires 595.1624.

2β,3α-disulfate-5α-cholan-24-oyl glycine trisodium salt (**16**). Carboxylate **14** (50 mg, 0.08 mmol) in DMF dry (2 mL) was treated with DMT-MM (66 mg, 0.24 mmol) and triethylamine (278 µL) and the mixture was stirred at room temperature for 10 min. Glycine ethyl ester (49.5 mg, 0.48 mmol) was then added to the mixture and stirring was continued for 24 h. After adding 5% methanolic NaOH (5 mL) and stirring for 5 h at room temperature, the alkaline solution was adjusted to pH 5, diluted with water and loaded onto a C18 silica gel column, which was washed with water (50 mL). Elution with 10% methanol gave the compound **16** (31 mg, 58% over two steps). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 30:70, flow rate 1.5 mL/min, t_R=3.4 min). $[\alpha]_{24}^{D}$= +10.8 (c 0.15, CH_3OH);

selected ^{1}H NMR (400 MHz CD_3OD): δ 4.73 (1H, br s), 4.70 (1H, br s), 3.72 (2H, s), 2.28 (1H, m), 2.13 (1H, m), 1.00 (3H, s), 0.97 (3H, d, J = 6.0 Hz), 0.69 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 176.5, 172.9, 76.5, 76.2, 57.8, 57.4, 56.6, 44.5, 43.8, 41.4, 40.2, 39.1, 36.9, 36.5, 36.4, 36.3, 34.1, 33.2, 33.1, 30.5, 29.2, 25.2, 22.0, 18.8, 14.3, 12.6. HR ESIMS m/z 652.1855 $[M-Na]^-$, $C_{26}H_{40}NNa_2O_{11}S_2$ requires 652.1838.

2β,3α-disulfate-5α-cholan-24-oyl taurine trisodium salt (**17**). Carboxylate **14** (50 mg, 0.08 mmol) in DMF dry (2 mL) was treated with DMT-MM (66 mg, 0.24 mmol) and triethylamine (278 µL) and taurine (60 mg, 0.48 mmol) was then added to the mixture, which was further stirred for 24 h. Then, the reaction mixture was concentrated under vacuo and dissolved in water (5 mL). The solution was poured over a C18 silica gel column. The product was eluted with H_2O:MeOH 99:1. To the solution of the solid in methanol (2 mL) was added Amberlite CG 120 sodium form (1 g)and the mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated to obtain compound **17** as a white solid (39 mg, 67%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 30:70, flow rate 1.5 mL/min, t_R=3.4 min). $[\alpha]_{24}^{D}$= +7.1 (c 0.22, CH_3OH); selected ^{1}H NMR (400 MHz CD_3OD): δ 4.73 (1H, br s), 4.70 (1H, br s), 3.57 (2H, t, J = 7.0 Hz), 2.95 (2H, t, J = 7.0 Hz), 2.23 (1H, m), 2.07 (1H, m), 0.99 (3H, s), 0.95 (3H, d, J = 6.4 Hz), 0.68 (3H, s) ; ^{13}C NMR (100 MHz CD_3OD): δ 176.7, 76.5, 70.5, 57.9, 57.3, 56.5, 51.5, 43.9, 41.4, 40.2 (2C), 39.2, 36.9, 36.6, 36.4, 36.3, 34.2, 33.2, 33.1, 30.5, 29.1, 25.1, 22.0, 18.8, 14.3, 12.6. HR ESIMS *m/z* 702.1685 $[M-Na]^-$, $C_{26}H_{42}NNa_2O_{12}S_3$ requires 702.1665.

Trisodium 2β,3α-disulfate-5α-cholan-24-oyl 5-aminosalycilic acid (**18**). Carboxylate **14** (50 mg, 0.08 mmol) in DMF dry (2 mL) was treated with DMT-MM (66 mg, 0.24 mmol), triethylamine (278 µL) and 5-aminosalicylic acid (5-ASA) (73 mg, 0.48 mmol) as described for compound **17** to obtain **18** as a white solid (43.3 mg, 72%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 30:70, flow rate 1.5 mL/min, t_R=3.4 min). $[\alpha]_{24}^{D}$= +8.2 (*c* 1.8, CH_3OH); selected ^{1}H NMR (400 MHz CD_3OD): δ 7.71 (1H, δ, J =3.1Hz), 7.6 (1H, dd, J = 3.1 and 8.7 Hz), 6.7 (1H, d, J = 8.7 Hz), 4.75 (1H, br s), 4.72 (1H, br s), 2.39 (1H, m), 2.25 (1H, m), 1.02 (3H, d, J = 6.0 Hz), 1.01 (3H, s), 0.72 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 181.2, 171.4, 142.1, 141.7, 131.8, 131.1, 129.8, 127.0, 76.5, 76.0, 57.9, 57.6, 56.7, 43.5, 41.4, 40.2, 39.2, 37.2, 36.5, 36.4, 36.3, 34.0, 33.2, 33.1, 30.5, 29.2, 24.2, 22.0, 18.9, 14.3, 12.6. HR ESIMS *m/z* 708.2135 $[M-Na]^-$, $C_{31}H_{43}NNaO_{12}S_2^-$ requires 708.2124. HR ESIMS *m/z* 730.1927 $[M-Na]^-$, $C_{31}H_{42}NNa_2O_{12}S_2$ requires 730.1944.

Sodium 2β,3α-disulfate-5α-cholan-24-ol (**19**). Dry methanol (26 µL, 0.65 mmol) and $LiBH_4$ (325 µL, 2 M in THF, 0.65 mmol) were added to a solution of the methyl ester **13** (100 mg, 0.16 mmol) in dry THF (5 mL) at 0 °C under argon and the resulting mixture was stirred for 4 h at 0 °C. The mixture was quenched by addition of MeOH (2 mL) and then concentrated under *vacuo*. Purification by C18 silica gel eluting with H_2O:MeOH (7:3) gave the product **19** as a white solid (70 mg, 75%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 32:68, flow rate 1.5 mL/min, t_R=3.2 min). $[\alpha]_{24}^{D}$= +7.8 (*c* 0.24,

CH_3OH); selected ^1H NMR (400 MHz CD_3OD): δ 4.69 (1H, br s), 4.74 (1H, br, s), 4.70 (1H, br s), 3.65 (1H, m), 3.50 (1H, m), 2.10 (1H, m), 2.00 (1H, m), 1.00 (3H, s), 0.95 (3H, d, J = 6.5 Hz), 0.69 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 76.5, 76.2, 63.6, 57.9, 57.6, 56.6, 43.8, 41.5, 40.2, 39.1, 37.0, 36.4, 36.3, 33.2 (2C), 30.4, 30.3, 29.2, 29.1, 25.2, 22.0, 19.2, 14.3, 12.5; HR ESIMS *m/z* 559.2037 $[M-Na]^-$, $C_{24}H_{40}NaO_9S_2^-$ requires 559.2011.

5α-Cholest-2-ene (**20**). An oven-dried 250 mL flask was charged with 10% palladium on carbon (100 mg) and cholesterol (1.00 g, 2.6 mmol) and the flask was evacuated and flushed with argon. Absolute methanol (50 mL) and dry THF (50 mL) were added, and the flask was flushed with hydrogen. The reaction was stirred at room temperature under H_2 (1 atm) for 4 h. The mixture was filtered through celite and the recovered filtrate was concentrated. The residue was subjected to column chromatography on silica gel eluting with hexane:AcOEt (9:1) to give 950 mg of pure 5α-cholestan-3β-ol (90%). To a solution of this latter (950 mg, 2.4 mmol) in dry pyridine (50 mL), *p*-toluenesulfonyl chloride (1.3 g, 7.2 mmol) was added. The solution was stirred at room temperature for 24 h and then poured into cold water (30 mL) and extracted with CH_2Cl_2 (3×50 mL). The combined organic layer was washed with saturated $NaHCO_3$ solution (50 mL), and water (50 mL), and then dried over anhydrous $MgSO_4$ and evaporated in *vacuo* to give 1.2 g of residue, that was subjected to next step without any purification. Lithium bromide (1.9 g, 22.0 mmol) and lithium carbonate (1.6 g, 22.0 mmol) were added to a solution of tosylate (1.2 g, 2.2 mmol) in dry DMF (150 mL), and the mixture was refluxed for 1.5 h. After cooling to room temperature, the mixture was slowly poured into 10% HCl solution (150 mL) and extracted with CH_2Cl_2 (3×150 mL). The combined organic layer was washed successively with water, saturated $NaHCO_3$ solution and water, and then dried over anhydrous $MgSO_4$ and evaporated to dryness to obtain pure **20** (773.3 mg, 87% over two steps). $[\alpha]_{24}^D$= +52.5 (*c* 0.11, $CHCl_3$); selected ^1H NMR (400 MHz $CDCl_3$): δ 5.6 (2H, m), 0.91 (3H, d, J = 6.4 Hz), 0.87 (6H, d, J = 6.4 Hz), 0.75 (3H, s), 0.66 (3H, s). ^{13}C NMR (100 MHz $CDCl_3$): δ 126.0, 125.8, 56.4, 56.2, 54.0, 41.4, 40.0, 39.7, 39.5 (2C), 36.5, 36.1, 35.7, 35.6, 34.6, 31.8, 30.3, 28.7, 28.2, 28.0, 24.2, 22.8, 22.5, 20.9, 18.6, 12.0, 11.6; HR ESIMS *m/z* 393.3487 $[M+Na]^+$, $C_{27}H_{46}Na$ requires 393.3497.

2α,3α-Epoxy-5α-cholestane (**21**). To a solution of **20** (750 mg, 2.02 mmol) in $CHCl_3$ (50 mL) was added slowly *m*-chloroperbenzoic acid (488 mg, 2.8 mmol). The mixture was stirred for 4 h at room temperature, and then CH_2Cl_2 (50 mL) and 5% Na_2SO_3 solution (70 mL) were added. The CH_2Cl_2 phase was washed with water (70 mL), dried over anhydrous $MgSO_4$ and evaporated to dryness to give 560 mg of crude epoxide **21**, that was subjected to next step without any purification. Selected ^1H NMR (400 MHz $CDCl_3$): δ 3.16 (1H, br s), 3.11 (1H, br s), 0.88 (3H, d, J = 6.5 Hz), 0.85 (6H, d, J = 6.5 Hz), 0.73 (3H, s), 0.62 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 56.2, 56.1, 53.6, 52.6, 51.3, 42.4, 40.0, 39.8, 39.5, 38.2, 36.2, 36.1, 35.7, 35.6, 33.6, 31.6, 29.0, 28.4, 28.2, 28.0, 24.2, 22.8, 22.5, 20.8, 18.6, 12.9, 11.9; HR ESIMS *m/z* 409.3458 ($[M+Na]^+$, $C_{27}H_{46}NaO$ requires 409.3446.

5α-Cholestan-2β,3α-diol (**22**). A solution of epoxide **21** (560 mg, 1.45 mmol) in THF (15 mL) was treated with 1N H_2SO_4 (3.6 mL, 3.6 mmol) solution and stirred for 24 h at room temperature. After

neutralization with saturated $NaHCO_3$ solution, the mixture was evaporated to a fifth of the initial volume, diluted with water (50 mL), and extracted with ethyl acetate (3x50 mL). The combined organic extracts were washed with water, dried over anhydrous $MgSO_4$, filtered and evaporated to dryness. Purification on a silica gel column by eluting with hexane:AcOEt 7:3 afforded pure **22** (637 mg, 78% over two steps). $[\alpha]_{24}^{D}$= +13.2 (*c* 0.2, CH_3OH); selected ^1H NMR (400 MHz $CDCl_3$): δ 3.90 (1H, br s), 3.87 (1H, br s), 0.90 (3H, d, J = 6.7 Hz), 0.87 (6H, d, J = 6.7 Hz), 0.80 (3H, s), 0.65 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 72.1, 71.0, 56.7, 56.5, 55.4, 42.7, 40.5, 39.9, 39.5, 38.9, 36.2, 36.1, 35.8, 35.7, 35.3, 31.3, 31.0, 28.5, 28.2, 28.0, 24.1, 22.8, 22.5, 20.8, 18.6, 14.6, 12.0; HR ESIMS *m/z* 427.3572 ($[M + Na]^+$, $C_{27}H_{48}NaO_2$ requires 427.3552.

2β,3α-cholestane disulfate (**23**). Triethylamine-sulfur trioxide complex (4.07 g, 22.5 mmol) was added to a solution of diol **22** (600 mg, 1.5 mmol) in DMF dry (70 mL) under an argon atmosphere, and the mixture was stirred at 95 °C for 48 h. Then, the reaction mixture was quenched with water (30 mL). The solution was poured over a C18 silica gel column to remove excess $Et_3N{\cdot}SO_3$. The product was eluted by MeOH and followed by evaporation of the solvent to yield a yellow solid [tris-(triethylammonium sulfate) salt]. To the solution of the solid in methanol (60 mL) was added Amberlite CG 120 sodium form (60 g). The mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated to obtain compound **23** as a white solid (821 mg, 90%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 87:13, flow rate 1.5 mL/min, t_R=2.6 min). $[\alpha]_{24}^{D}$= +16 (*c* 0.2, CH_3OH); selected ^1H NMR (400 MHz $CDCl_3$): δ 4.75 (1H, m), 4.71 (1H, m), 0.97 (3H, s), 0.93 (3H, d, J = 6.4 Hz), 0.88 (6H, d, J = 6.4 Hz), 0.68 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 76.5, 76.1, 57.8, 57.6, 56.6, 43.8, 41.4, 40.7, 40.2, 39.1, 37.3, 37.1, 36.4, 36.3, 31.8, 30.7, 30.5, 29.3, 29.1 (2C), 25.2, 23.2, 22.9, 22.0, 19.2, 14.3, 12.5; HR ESIMS *m/z* 585.2527 $[M-Na]^-$, $C_{27}H_{49}NaO_8S_2$ requires 585.2532.

Methyl 3α-hydroxy-5α-cholan-24-oate (**24**). To a solution of methyl 3β-hydroxy-5α-cholan-24-oate **7**[89] (173 mg, 0.44 mmol) in dry pyridine (30 mL), *p*-toluenesulfonyl chloride (420 mg, 2.2 mmol) was added. The solution was stirred at room temperature for 24 h and then poured into cold water (30 mL) and extracted with CH_2Cl_2 (3×50 mL). The combined organic layer was washed saturated $NaHCO_3$ solution (50 mL), and water (50 mL), and then dried over anhydrous $MgSO_4$ and evaporated in vacuo to give 200 mg of residue, that was subjected to next step without any purification. A solution of methyl 3β-tosyl-5α-cholan-24-oate (200 mg, 0.36 mmol) and CH_3COOK (36 mg, 0.36 mmol) dissolved in water (1 mL) and DMF (7 mL) was refluxed for 5 h. The solution was cooled at room temperature and then ethyl acetate and water were added. The separated aqueous phase was extracted with ethyl acetate (3×30 mL). The organic phase was washed with water, dried (Na_2SO_4) and evaporated to dryness to give 190 mg of mixture, that was subjected to next step without any purification. This residue was dissolved in 32 mL of mixture $CHCl_3$:MeOH (5:3) and *p*-toluensulfonic acid (*p*-TsOH) (619 mg, 3.6 mmol) was added. The mixture was quenched by addition of $NaHCO_3$ solution (30 mL) and then concentrated under vacuo. Ethyl acetate and water were added and the separated aqueous phase was extracted with

ethyl acetate (3×50 mL). The combined organic phase was washed with water, dried (Na_2SO_4) and concentrated. Purification by silica gel eluting with hexane-AcOEt 99:1 gave the alcohol **24** as a white solid (130 mg, 75% over three steps). $[\alpha]_{24}^{D}$= +7.2 (*c* 0.05, CH_3OH); selected ^{1}H NMR (400 MHz $CDCl_3$): δ 4.03 (1H, br s), 3.66 (3H, s), 2.35 (1H, m), 2.20 (1H, m), 0.90 (3H, d, J = 6.0 Hz), 0.77 (3H, s), 0.64 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 175.3, 66.8, 56.8, 56.1, 54.6, 51.8, 42.8, 40.2, 40.0, 39.3, 36.1, 35.7, 35.6, 32.4, 32.2, 31.3, 31.2, 29.2, 28.8, 28.3, 24.4, 21.0, 18.5, 14.2, 12.3; HR ESIMS *m/z* 413.5867 $[M+Na]^+$, $C_{25}H_{42}NaO_3$ requires 413.5884.

5α-cholan-3α,24-diyl-3,24-sodium disulfate (**25**). Dry methanol (70 µL, 1.8 mmol) and $LiBH_4$ (900 µL, 2 M in THF, 1.8 mmol) were added to a solution of the methyl ester **24** (100 mg, 0.25 mmol) in dry THF (10 mL) at 0 °C under argon and the resulting mixture was stirred for 4 h at 0 °C. The mixture was quenched by addition of MeOH (4 mL) and then concentrated under vacuo. Ethyl acetate and water were added and the separated aqueous phase was extracted with ethyl acetate (3×30 mL). The combined organic phase was washed with water, dried (Na_2SO_4) and concentrated to obtain 5α-cholan-3α,24-diol as a white solid (77 mg, 85%). The triethylamine-sulfur trioxide complex (350 mg, 2 mmol) was added to a solution of diol (70 mg, 0.2 mmol) in DMF dry (2 mL) under an argon atmosphere, and the mixture was stirred at 95 °C for 24 h. Then, the reaction mixture was quenched with water (1.6 mL) and the solution was poured over a C18 silica gel column to remove excess $Et_3N{\cdot}SO_3$. The product was eluted by MeOH and followed by evaporation of the solvent to yield a yellow solid [tris-(triethylammonium sulfate) salt]. To the solution of the solid in methanol (4 mL) was added Amberlite CG 120 sodium form (2 g). The mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated to obtain a residue that by purification on C18 [H_2O:MeOH (7:3)] gave the product **25** as a white solid (71.3 mg, 63%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 25:75, flow rate 1.5 mL/min, t_R=2.4 min). $[\alpha]_{24}^{D}$= + 0.42 (*c* 1.2, CH_3OH); selected ^{1}H NMR (400 MHz CD_3OD): δ 4.55 (1H, br s), 3.96 (2H, m), 0.90 (3H, d, J = 6.5 Hz), 0.80 (3H, s), 0.65 (3H, s) ; ^{13}C NMR (100 MHz CD_3OD): δ 76.5, 69.7, 57.9, 57.6, 55.1, 43.8, 41.6, 41.4, 40.8, 36.9, 36.8, 34.7, 33.8, 33.2, 33.1, 29.6, 29.2, 27.9, 27.2, 25.2, 21.9, 19.0, 14.4, 12.5. HR ESIMS *m/z* 543.2078 $[M-Na]^-$, $C_{24}H_{40}NaO_8S_2$ requires 543.2062.

5α-Cholan-3β,24-diyl-3,24-sodium disulfate (**26**). Dry methanol (10 µL, 0.20 mmol) and $LiBH_4$ (100 µL, 2 M in THF, 0.20 mmol) were added to a solution of the methyl ester **7** (20 mg, 0.05 mmol) in dry THF (2 mL) at 0 °C under argon and the resulting mixture was stirred for 4 h at 0 °C. The mixture was quenched by addition of MeOH (2 mL) and then allowed to warm to room temperature and concentrated under vacuo. Ethyl acetate and water were added and the separated aqueous phase was extracted with ethyl acetate (3×30 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and concentrated to obtain 5α-cholan-3β,24-diol as a white solid (13 mg, 72%). The triethylamine-sulfur trioxide complex (63 mg, 0.35 mmol) was added to a solution of diol (13 mg, 0.035 mmol) in DMF dry (2 mL) under an argon atmosphere, and the mixture was stirred at 95 °C for 3 h. The resulting solution was concentrated under reduced

pressure. To the solution of the solid in methanol (3 mL) was added Amberlite CG 120 sodium form (1 g). The mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated and then diluted with water and loaded onto a C18 silica gel column, which was washed with water (50 mL). Elution with 30% aqueous methanol gave the compound **26** as a white solid (15 mg, 78%). An analytical sample was obtained by HPLC on a Nucleodur 100-5 C18 (5µm; 4.6 mm i.d. x 250 mm, MeOH:H_2O 25:75, flow rate 1.5 mL/min, t_R=2.4 min). $[\alpha]_{24}^{D}$= +6.4 (*c* 1.0, CH_3OH); selected ^{1}H NMR (400 MHz CD_3OD): δ 4.59 (1H, m), 3.95 (1H, br s), 0.95 (3H, d, J = 6.0 Hz), 0.82 (3H, s), 0.78 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 79.9, 64.3, 57.9, 57.6, 55.4, 43.8, 41.4 (2C), 38.2, 36.8, 36.7, 36.3, 33.3, 33.1, 32.2, 30.7, 29.2, 28.7, 27.1, 25.2, 22.3, 19.0, 14.4, 12.5. HR ESIMS *m/z* 543.2075 $[M\text{-}Na]^-$, $C_{24}H_{40}NaO_8S_2$ requires 543.2062.

Synthesis of solomonsterol B

3α,6α-Diformyloxy-5β-cholan-24-oic Acid (**27**): A solution of hyodeoxycholic acid (**3**) (2 g, 5.09 mmol) in 90% formic acid (30 mL) containing 70% perchloric acid (70 µL) was stirred at 50 °C for 1.5 h. The temperature was lowered to 40 °C. Acetic anhydride (24 mL) was added over 10 min, and the mixture was stirred for an additional 10 min. The solution was cooled to room temperature and poured into water (50 mL) with vigorous stirring. The precipitate was filtered, washed with water (20 mL), and dissolved in diethyl ether. The solution was dried with sodium sulfate, and evaporated to give of **27** (1.9 g, 97 %), which was used in the next step without further purification. An analytical sample was obtained by silica gel chromatography, eluting with hexane/EtOAc (6:4). $[\alpha]_{24}^{D}$ = +5.8 (*c* = 1.3, CH3OH). Selected 1H NMR (400 MHz, CDCl3): *δ* 7.90 (s, 1 H), 7.87 (s, 1 H), 5.15 (m, 1 H), 4.67 (m, 1 H), 2.24 (m, 1 H), 2.10 (m, 1 H), 0.85 (s, 3 H), 0.77 (d, *J* = 6.5 Hz, 3 H), 0.5 (s, 3 H) ppm. 13C NMR (100 MHz, CDCl3): *δ* 179.7, 160.7, 160.6, 73.4, 70.8, 55.7, 55.6, 50.0, 45.1, 42.6, 39.6, 35.8 (2 C), 35.0, 34.6, 34.4, 30.8, 30.5, 27.8, 26.2, 25.9, 23.9, 23.1, 20.5, 18.1, 11.8 ppm. HRMS-ESI *m/z* 449.2913 ($[M + H]^+$, $C_{26}H_{41}O_6^+$ requires 449.2903.

3α,6α-Diformyloxy-24-nor-5β-cholane-23-nitrile (**28**). 1.9 g of 3α,6α-diformyloxy-5β-cholan-24-oic acid, **27** (4.6 mmol), in 6.7 mL of trifluoroacetic acid and 1.8 mL (12.7 mmol) of trifluoroacetic anhydride were stirred at 5 °C until dissolution. Sodium nitrite (321 mg, 4.65 mmol) was added in small portions and the reaction mixture was stirred at 5 °C for 1 h and then at 40 °C for 1.5 h. The solution was cooled to room temperature and poured into 30 mL of 2N NaOH. The mixture was extracted with ethyl ether and the organic layer was washed with 1N NaOH and water, dried with sodium sulfate, and evaporated to give 1.66 g of pure **28** (80%). $[\alpha]_{24}^{D}$= +8.8 (*c* 0.22, $CHCl_3$); ^{1}H NMR (400 MHz, CD_3OD): δ 8.08 (1H, s), 8.07 (1H, s), 5.27 (1H, m), 4.73 (1H, m), 2.46 (1H, dd, J= 4.1, 16.9 Hz), 2.34 (1H, dd, J= 6.0, 16.9 Hz), 0.93 (3H, s), 0.72 (3H, d, J= 6.9 Hz), 0.71 (3H, s). ^{13}C NMR (100 MHz, CD_3OD): δ 162.7, 162.6, 120.3, 72.3, 68.6, 57.3, 56.3, 49.6, 46.8, 43.9, 41.1, 40.9, 36.6, 36.4, 36.0, 35.3, 34.5, 29.8, 28.9, 26.3, 25.0, 24.1, 21.8, 19.7, 12.5; HRMS-ESI *m/z* 416.2821 $[M + H]^+$, $C_{25}H_{39}NO_4^+$ requires 416.2801.

24-nor-hyodeoxycholic acid (**29**). 1.66 g of crude **28** was refluxed with 30% KOH in ca. 50 mL of ethanol-water 1:1 for 48 h. The ethanol was evaporated, the solution was saturated with sodium chloride and extracted with ethyl ether. The organic phase was washed with water, dried, and evaporated to give white solid residue **29** (1.18 g, 78%). $[\alpha]_{24}^{D}$= +3.0 (*c* 2.9, CH_3OH); ^{1}H NMR (400 MHz, CD_3OD): δ 4.01 (1H, dt, J= 4.6, 12.1 Hz), 3.52 (1H, m), 2.40 (1H, m), 1.99 (1H, m), 0.98 (3H, d, J= 6.0 Hz), 0.91 (3H, s), 0.70 (3H, s). ^{13}C NMR (100 MHz, CD_3OD): δ 177.8, 72.4, 68.7, 57.4, 57.2, 49.5, 43.9, 42.4, 41.1, 41.0, 36.8, 36.6, 36.0, 35.3, 34.8, 30.9, 29.7, 29.1, 25.1, 23.9, 21.7, 19.9, 12.4; HRMS-ESI *m/z* 379.2858 $[M + H]^+$, $C_{23}H_{39}O_4^+$ requires 379.2848.

Methyl 3α,6α-dihydroxy-24-nor-5β-cholan-23-oate. A mixture of the 24-*nor*hyodeoxycholic acid **29** (1.18 g, 3.11 mmol) and *p*-toluenesulfonic acid (1.2 g, 6.22 mmol) in methanol dry (50 mL) was stirred at room temperature for over night. The mixture was quenched with $NaHCO_3$ solution, most of the solvent was evaporated, and the residue was extracted with EtOAc. The combined extract was washed with brine, dried with Na_2SO_4, and evaporated to give the desired methyl ester as colorless amorphous solids: 1.0 g (96%). $[\alpha]_{24}^{D}$= -1.0 (*c* 0.69, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.98 (1H, m), 3.61 (3H, s), 3.54 (1H, m), 2.38 (1H, m), 1.98 (1H, m), 0.92 (3H, d, J= 6.0 Hz), 0.84 (3H, s), 0.62 (3H, s). ^{13}C NMR (100 MHz, $CDCl_3$): δ 174.1, 71.3, 67.8, 56.1, 56.0, 51.3, 48.2, 42.8, 41.3, 39.7 (2C), 35.8, 35.4, 34.6, 34.4, 33.7, 29.7, 28.9, 28.1, 24.1, 23.3, 20.6, 19.4, 11.9; HRMS-ESI *m/z* 393.3015 $[M + H]^+$, $C_{24}H_{41}O_4^+$ requires 393.3005.

Methyl 3,6-ditosyloxy-24-nor-5α-cholan-23-oate (**30**). To a solution of methyl ester (1.0 g, 2.55 mmol) in dry pyridine (15 mL), a solution of tosyl chloride (2.4 g, 12.7 mmol) in dry pyridine (15 mL) was added, and the mixture was stirred at room temperature for 4 h. The solvent was evaporated and the precipitate was re-dissolved in CH_2Cl_2, washed with $NaHCO_3$ and water, dried with Na_2SO_4 and then evaporated to dryness. The residue was passed through a short column of silica gel (120 g) eluting with EtOAc/hexane (1:1, v/v) to give **30** as a white solid, 1.4 g (78%). $[\alpha]_{24}^{D}$= +3.9 (*c* 1.2, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.78 (2H, d, J= 8.3 Hz), 7.71 (2H, d, J= 8.3 Hz), 7.35 (2H, d, 8.3 Hz), 7.33 (2H, d, J= 8.3 Hz), 4.77 (1H, m), 4.30 (1H, m), 3.65 (3H, s), 2.5 (3H, s), 2.40 (1H, dd, J= 3.4, 15.0 Hz), 2.02 (1H, m), 0.93 (3H, d, J= 6.2 Hz), 0.79 (3H, s), 0.62 (3H, s). ^{13}C NMR (100 MHz, $CDCl_3$): δ 174.2, 145.0 (2C), 134.7 (2C), 130.2 (4C), 127.8 (4C), 82.0, 79.9, 56.1, 56.0, 51.7, 46.6, 43.2, 41.6, 39.7, 39.6, 36.4, 35.8, 35.0 (2C), 33.9, 28.4, 27.6, 26.7, 24.2, 23.2 (2C), 21.9, 20.7, 19.7, 12.3; HRMS-ESI *m/z* 701.3192 $[M + H]^+$, $C_{38}H_{53}O_8S_2^+$ requires 701.3182.

Methyl 3β-hydroxy-24-nor-5-cholen-23-oate (**31**). A solution of methyl 3,6-ditosyloxy-24-nor-5α-cholan-23-oate (1.4 g, 2 mmol), CH_3COOK (196.3 mg, 2 mmol) dissolved in water (2 mL) and N,N'-dimethylformamide (DMF, 14 mL) was refluxed for 4 h. The solution was cooled at room temperature and then ethyl acetate and water were added. The separated aqueous phase was extracted with ethyl acetate (3 ×30 mL) and the combined organic phases were washed with water, dried (Na_2SO_4) and evaporated to dryness to give 1 g of mixture, that was subjected to next step without any purification. The mixture was dissolved in 32 mL of $CHCl_3$:MeOH (5:3) and to the solution was added *p*-toluensulfonic acid (*p*-TsOH) (1.14 g, 6 mmol). The mixture was quenched

by addition of $NaHCO_3$ solution (30 mL) and then concentrated under vacuo. Ethyl acetate and water were added and the separated aqueous phase was extracted with ethyl acetate (3×50 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and concentrated. Purification by silica gel eluting with hexane/AcOEt (97:3) gave the alcohol **31** as a white solid (600 mg, 80% over two steps). $[\alpha]_{24}^{D}$= -2.4 (*c* 1.45, $CHCl_3$); selected ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.30 (1H, d, J= 2.4 Hz), 3.62 (3H, s), 3.48 (1H, m), 2.39 (1H, dd, J= 2.6, 13.7 Hz), 2.00 (1H, m), 0.96 (3H, s), 0.94 (3H, d, J = 6.5 Hz), 0.67 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$): δ 174.4, 141.1, 121.8, 71.9, 56.9, 56.1, 51.7, 42.6, 41.6, 41.1, 39.8, 37.5, 36.7, 34.0, 32.1, 32.0 (2C), 31.6, 28.4, 24.4, 24.1, 21.2, 19.7, 12.3; HRMS-ESI *m/z* 375.2908 $[M+H]^+$, $C_{24}H_{39}O_3^+$ requires 375.2899.

Methyl 3β-hydroxy-24-nor-5α-cholan-23-oate (**32**). An oven-dried 100 mL flask was charged with 10% palladium on carbon (50 mg) and compound **31** (600 mg, 1.6 mmol) and the flask was evacuated and flushed with argon. Absolute methanol (30 mL) and dry THF (30 mL) were added and the reaction was stirred at room temperature under H_2 for 4 h. The mixture was filtered through celite and the recovered filtrate was concentrated to give 506 mg of pure **32** (84%). $[\alpha]_{24}^{D}$= +1.0 (*c* 0.17, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.63 (3H, s), 3.56 (1H, m), 2.40 (1H, dd, J= 2.8, 14.3 Hz), 1.98 (1H, m), 0.94 (3H, d, J= 6.3 Hz), 0.77 (3H, s), 0.66 (3H, s). ^{13}C NMR (100 MHz, $CDCl_3$): δ 174.4, 71.6, 56.7, 56.3, 54.5, 51.6, 42.9, 41.7, 40.1, 38.4, 37.2, 35.7 (2C), 34.1, 32.3 (2C), 31.7, 28.9, 28.5, 24.4, 21.4, 19.7, 12.5, 12.3; HRMS-ESI *m/z* 377.3060 ($[M + H]^+$, $C_{24}H_{41}O_3^+$ requires 377.3050.

Methyl 24-nor-5α-chol-2-en-23-oate (**33**). To a solution of **32** (500 mg, 1.3 mmol) in dry pyridine (25 mL), *p*-toluenesulfonyl chloride (1.2 g, 6.6 mmol) was added. The solution was stirred at room temperature for 24 h and then poured into cold water (20 mL) and extracted with CH_2Cl_2 (3 × 15 ml). The combined organic layer was washed with saturated $NaHCO_3$ solution (30 mL) and water (30 mL), dried over anhydrous $MgSO_4$ and evaporated in vacuo to give 665 mg of 3-tosylate derivative that was subjected to next step without any purification. ^{1}H NMR (400 MHz $CDCl_3$): δ 7.78 (2H, d, J= 8.1 Hz), 7.32 (2H, d, J= 8.1 Hz), 4.41 (1H, m),. 3.65 (3H, s), 2.44 (3H, s), 2.41 (1H, dd, J= 2.7, 12.0 Hz), 1.96 (1H, m), 0.95 (3H, d, J= 6.3 Hz), 0.77 (3H, s), 0.66 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 174.3, 144.6, 135.0, 130.0 (2C), 127.7 (2C), 82.7, 56.5, 56.2, 54.1, 51.6, 42.9, 41.6, 41.2, 39.9, 36.9, 35.5 (2C), 35.3, 35.0, 34.0, 32.0, 28.5, 28.4, 24.3, 24.1, 21.4, 21.3, 19.7, 12.2; HRMS-ESI *m/z* 531.3164 $[M + H]^+$, $C_{31}H_{47}O_5S^+$ requires 531.3144. Lithium bromide (217 mg, 2.5 mmol) and lithium carbonate (184 mg, 2.5 mmol) were added to a solution of 3-tosylate derivative (665 mg, 1.25 mmol) in dry DMF (30 mL), and the mixture was refluxed for 5 h. After cooling to room temperature, the mixture was slowly poured into 10% HCl solution (150 mL) and extracted with CH_2Cl_2 (3 × 150 mL). The combined organic layer was washed successively with water, saturated $NaHCO_3$ solution, and water, and then dried over anhydrous $MgSO_4$ and evaporated to dryness to obtain pure **33** (380 mg, 81% over two steps). $[\alpha]_{24}^{D}$= +3.1 (*c* 0.76, $CHCl_3$); ^{1}H NMR (400 MHz $CDCl_3$): δ 5.58 (2H, m), 3.65 (3H, s), 2.42 (1H, dd, J= 2.7, 14.0 Hz), 1.96 (1H, m), 0.96 (3H, d, J= 6.5 Hz), 0.74 (3H, s), 0.69 (3H, s). ^{13}C NMR (100 MHz $CDCl_3$): δ 174.4, 126.2, 126.1, 56.7, 56.3, 54.2, 51.6, 42.8, 41.7, 41.6, 40.1, 40.0, 39.9, 35.8, 34.7,

34.1, 31.9, 30.5, 28.4, 24.4, 21.1, 20.9, 19.7, 12.2; HRMS-ESI *m/z* 359.2963 ($[M + H]^+$, $C_{24}H_{39}O_2^+$ requires 359.2950.

Methyl 2β,3α-dihydroxy-24-nor-5α-cholan-23-oate (**35**). To a solution of **33** (380 mg, 1.06 mmol) in $CHCl_3$ (20 mL) was added slowly *m*-chloroperbenzoic acid (256 mg, 1.5 mmol). The mixture was stirred for 4 h at room temperature, and then CH_2Cl_2 (3 × 50 mL) and 5% Na_2SO_3 solution (70 mL) were added. The combined CH_2Cl_2 extracts were washed with water (70 mL), dried over anhydrous $MgSO_4$ and evaporated to dryness to give 400 mg of crude 2,3-epoxide, that was subjected to next step without any purification. A solution of epoxide (400 mg, 1.06 mmol) in THF (20 mL) was treated with 1N H_2SO_4 (5.3 mL, 2.65 mmol) solution and stirred for 24 h at room temperature. After neutralization with saturated $NaHCO_3$ solution, the mixture was concentrated, diluted with water (50 mL), and extracted with ethyl acetate (3 x 50 mL). The combined organic extracts were washed with water, dried over anhydrous $MgSO_4$, filtered and evaporated to dryness. Purification on silica gel column eluting with CH_2Cl_2-MeOH (9:1) afforded pure **35** (340 mg, 81% over two steps). $[\alpha]_{24}^D$= -0.6 (*c* 0.5, $CHCl_3$); ^{1}H NMR (400 MHz $CDCl_3$): δ 3.85 (1H, br s), 3.80 (1H, br s), 3.63 (3H, s), 2.38 (1H, dd, J= 2.7, 14.0 Hz), 0.95 (3H, s), 0.90 (3H, d, J = 6.5 Hz), 0.60 (3H, s); ^{13}C NMR (100 MHz $CDCl_3$): δ 171.3, 71.6, 70.5, 56.6, 56.2, 55.2, 51.6, 41.7, 40.4, 40.0, 39.9, 39.1, 35.1 (2C), 34.1, 32.0, 31.8, 31.7, 28.4, 24.3, 21.0, 19.7, 14.5, 12.3; HRMS-ESI *m/z* 393.3015 ($[M + H]^+$, $C_{24}H_{41}O_4^+$ requires 393.3005.

24-Nor-5α-cholan-2β,3α,23-triol (**34**). Dry methanol (122 µL, 3.03 mmol) and $LiBH_4$ (1.5 mL, 2 M in THF, 3.03 mmol) were added to a solution of the methyl ester **35** (170 mg, 0.43 mmol) in dry THF (10 mL) at 0 °C under argon and the resulting mixture was stirred for 8 h at 0 °C. The mixture was quenched by addition of NaOH (1N, 4 mL) and then allowed to warm to room temperature. Ethyl acetate was added and the separated aqueous phase was extracted with ethyl acetate (3×15 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and concentrated. Purification by silica gel eluting with CH_2Cl_2-MeOH (95:5) gave the alcohol **34** as a white solid (110 mg, 69%). $[\alpha]_{24}^D$= -9.8 (*c* 0.05, CH_3OH); ^{1}H NMR (400 MHz CD_3OD): δ 3.80 (1H, br s), 3.75 (1H, br s), 3.46 (2H, m), 0.99 (3H, s), 0.95 (3H, d, J = 6.3 Hz), 0.69 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 72.2, 71.2, 60.8, 59.9, 57.9, 56.7, 41.5, 40.9, 40.1, 39.8, 36.9, 36.3, 34.1, 33.3, 32.4, 30.7, 29.5, 29.3, 25.2, 22.0, 19.3, 14.6, 12.5; HRMS-ESI *m/z* 365.3078 ($[M + H]^+$, $C_{23}H_{41}O_3^+$ requires 365.3056.

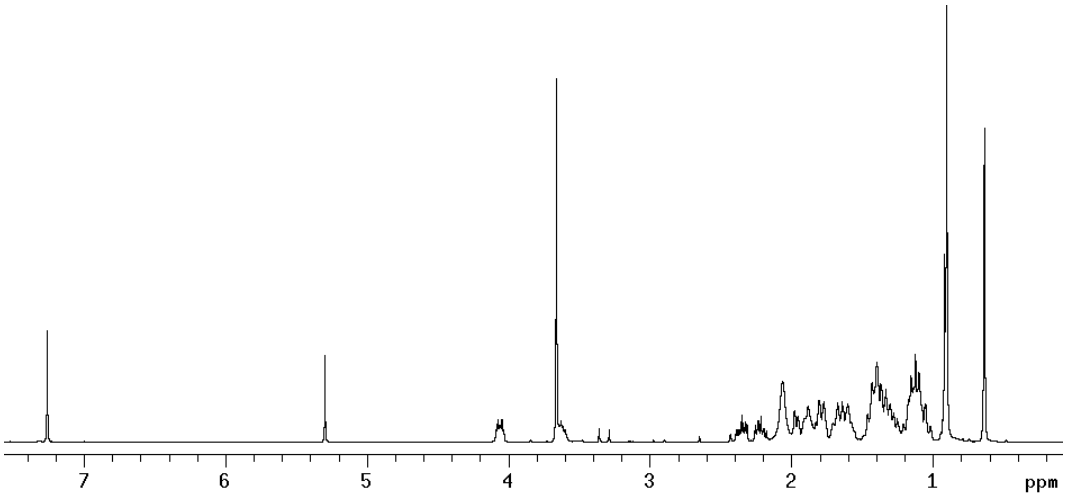
24-Nor-5α-cholan-2β,3α,23-tryl-2,3,23-sodium trisulfate (**2**). The triethylamine-sulfur trioxide complex (272 mg, 1.5 mmol) was added to a solution of triol **34** (110 mg, 0.3 mmol) in DMF dry (5 mL) under argon and the mixture was stirred at 95 °C for 24 h. The reaction mixture was quenched with water (800 µL) and the resulting solution was poured over a silica gel column to remove the excess of $SO_3 \cdot NEt_3$. Elution with MeOH, followed by evaporation of the solvent, yielded a yellow solid (diammonium trisulfate salt). To the solution of the solid in methanol (15 mL) was added Amberlite CG 120 sodium form (10 g), and the resulting mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated to obtain a mixture that was further purified by HPLC on a Nucleodur 100-5 C18 (5 µm, 10 mm i.d.

x 250 mm) with MeOH:H_2O (34:66) as eluent (flow rate 3 mL/min) to give 144 mg (72%) of solomonsterol B (**2**) (t_R=4.4 min), $[\alpha]_{24}^{D}$= +2.1 (*c* 0.08, CH_3OH); ^1H NMR (700 MHz CD_3OD): δ 4.73 (1H, br s), 4.70 (1H, br s), 4.05 (1H, m), 4.01 (1H, m), 1.00 (3H, s), 0.98 (3H, d, J= 6.5 Hz), 0.70 (3H, s); ^{13}C NMR (100 MHz CD_3OD): δ 76.5, 71.1, 67.3, 57.8, 56.6, 56.1, 43.5, 41.4, 39.8, 38.7, 36.6, 36.4, 36.3, 34.2, 33.2, 30.5, 29.2, 29.1, 25.2, 22.0, 19.1, 14.3, 12.5. HRMS-ESI *m/z* 647.1298 ([M - Na]$^-$,$C_{23}H_{37}Na_2O_{12}S_3^-$, requires 647.1243.

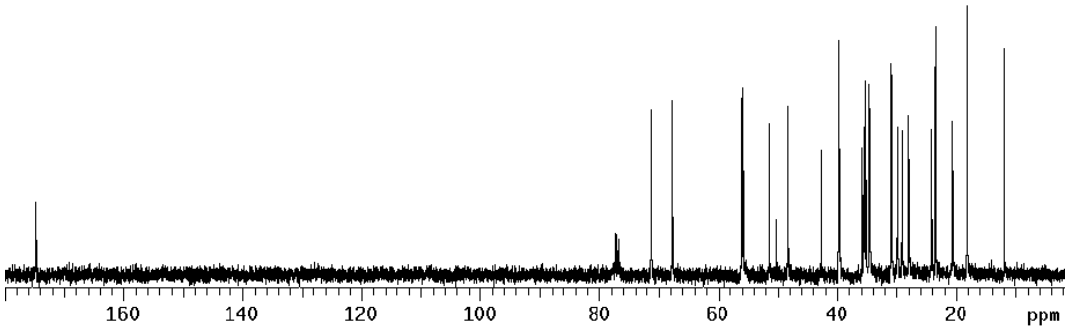
Sodium 2β,3α-disulfate-24-nor-5α-cholan-23-ol (**37**). Triethylamine-sulfur trioxide complex (780 mg, 4.3 mmol) was added to diol **35** (170 mg, 0.43 mmol) in DMF dry (10 mL) under an argon atmosphere, and the mixture was stirred at 95 °C for 48 h. The reaction mixture was quenched with water (1.6 mL) and the solution was poured over a C18 silica gel column to remove excess $SO_3 \cdot NEt_3$. The product was eluted by MeOH and the evaporation of the solvent yielded a yellow solid [bis-(triethylammonium sulfate) salt]. To the solid dissolved in methanol (30 mL) was added Amberlite CG 120 sodium form (30 g) and the mixture was stirred for 5 h at room temperature. The resin was removed by filtration, and the filtrate was concentrated to obtain 250 mg methyl 2β,3α-disulfate-24-*nor*-5α-cholan-23-oate **36**, that was subjected to next step without any purification. Dry methanol (54 μL, 2.15 mmol) and $LiBH_4$ (1.0 mL, 2 M in THF, 2.15 mmol) were added to a solution of **36** (250 mg, 0.43 mmol) in THF dry (5 mL) at 0 °C under argon and the resulting mixture was stirred for 3 h at 0 °C. The mixture was quenched by addition of MeOH (2 mL) and then concentrated under vacuo. Purification by C18 cartige eluting with MeOH:H_2O (1:99) gave the product **37** as a white solid (190 mg, 78% over two steps). $[\alpha]_{24}^{D}$= +19.8 (*c* 0.02, CH_3OH); selected ^1H NMR (400 MHz CD_3OD): δ 4.72 (1H, br s), 4.70 (1H, br, s), 3.64 (1H, m), 3.63 (1H, m), 1.00 (3H, s), 0.95 (3H, d, J = 6.5 Hz), 0.70 (3H, s); HRMS-ESI *m/z* 545.1875 [M-Na]$^-$, $C_{23}H_{38}NaO_9S_2^-$ requires 545.1855.

Spectroscopic Data

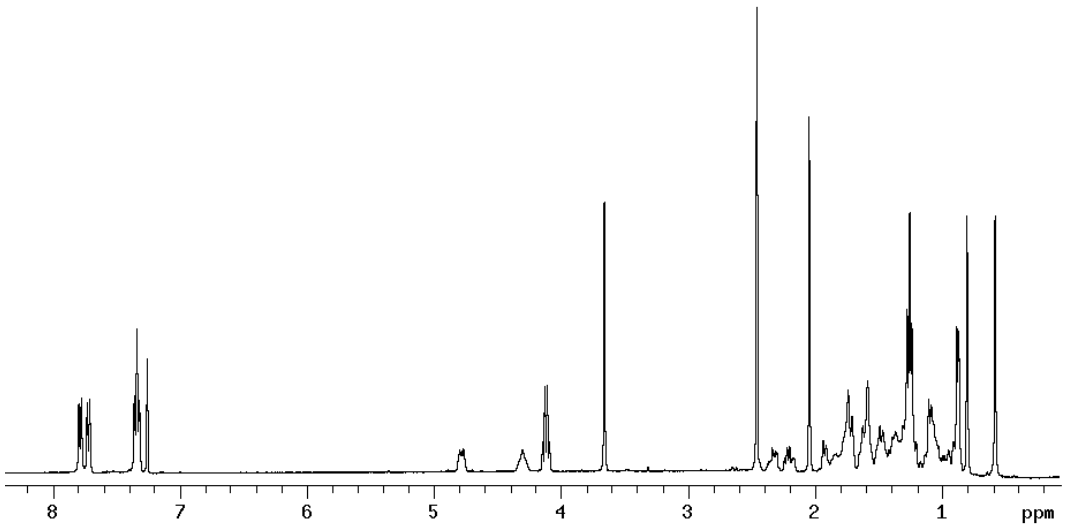
^1H NMR spectrum of Compound **4** in $CDCl_3$ at 400 MHz.



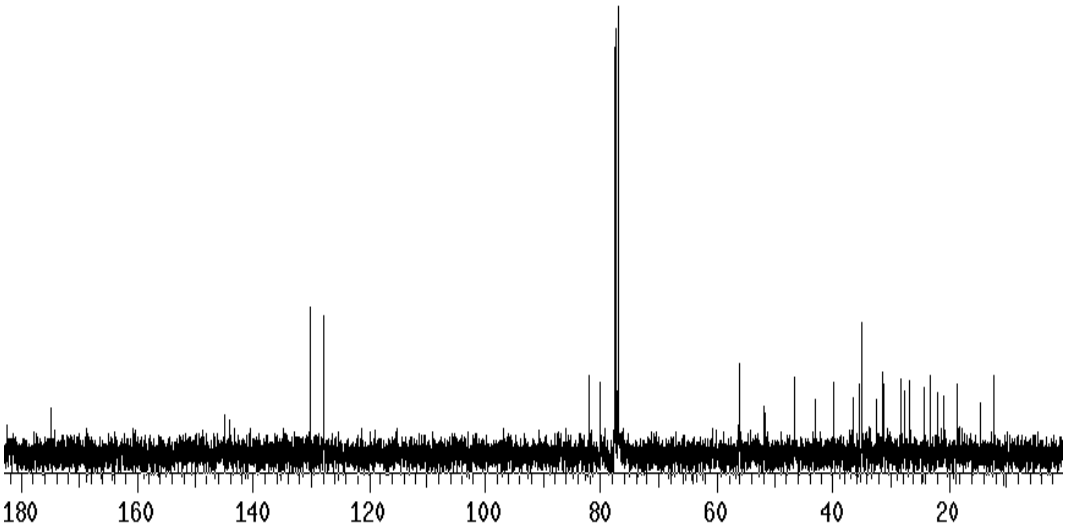
^{13}C NMR spectrum of Compound **4** in $CDCl_3$ at 100 MHz.



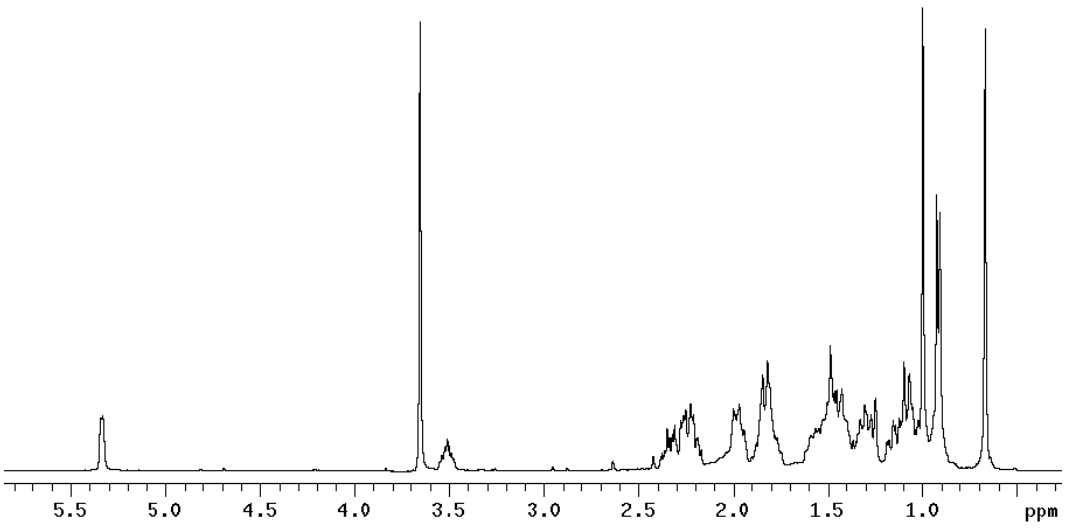
^{1}H NMR spectrum of Compound **5** in $CDCl_3$ at 400 MHz.



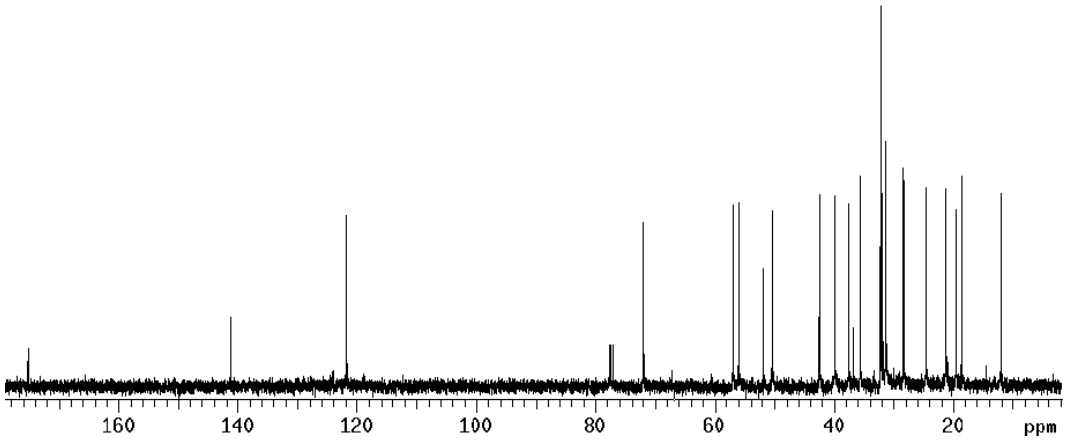
^{13}C NMR spectrum of Compound **5** in $CDCl_3$ at 100 MHz.



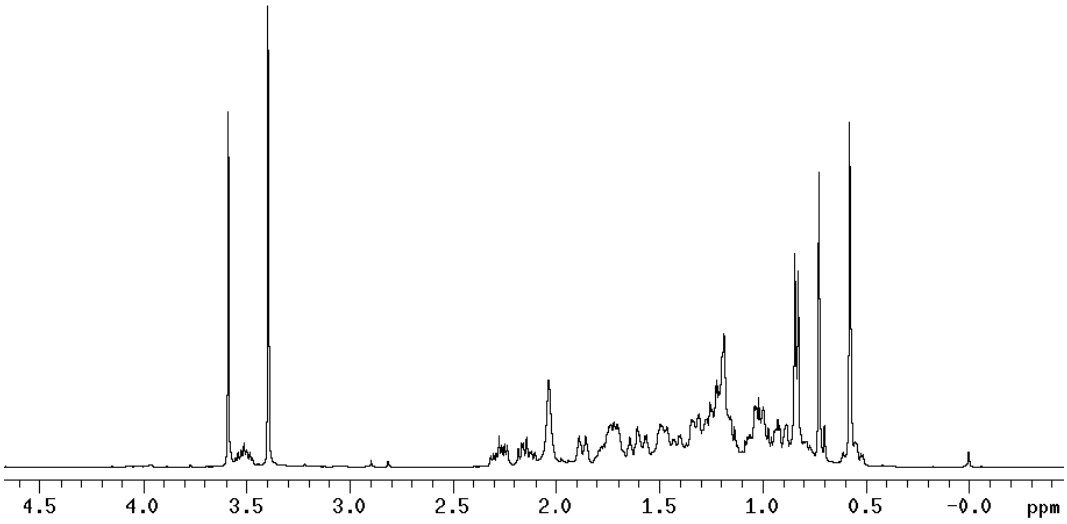
^{1}H NMR spectrum of Compound **6** in $CDCl_3$ at 400 MHz.



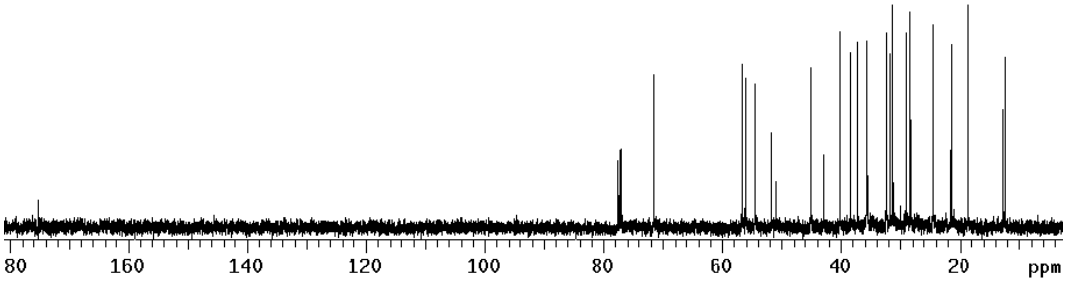
^{13}C NMR spectrum of Compound **6** in $CDCl_3$ at 100 MHz.



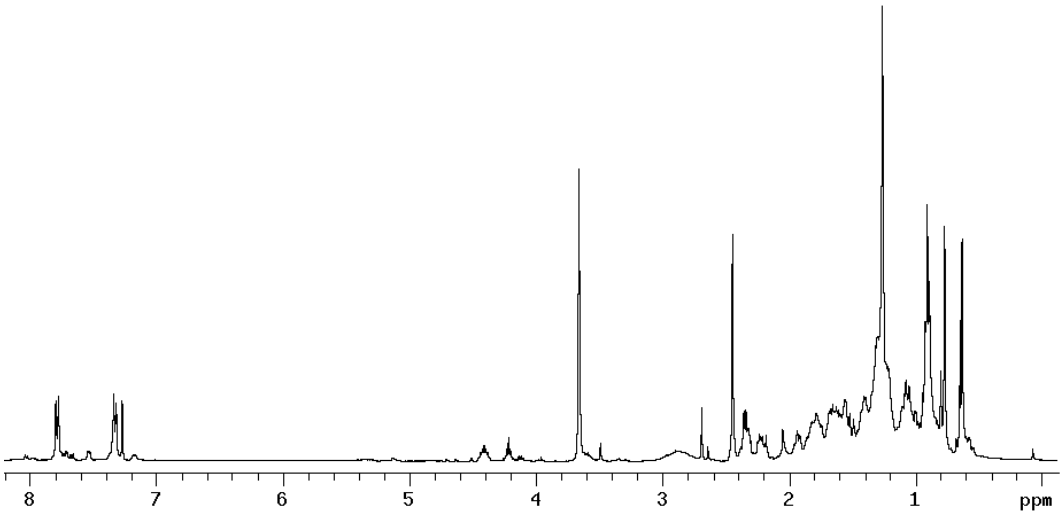
^{1}H NMR spectrum of Compound **7** in $CDCl_3$ at 400 MHz.



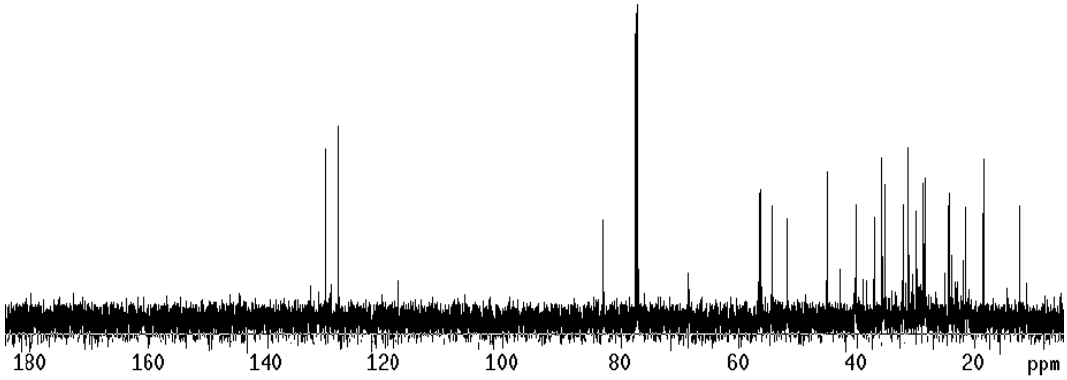
^{13}C NMR spectrum of Compound **7** in $CDCl_3$ at 100 MHz.



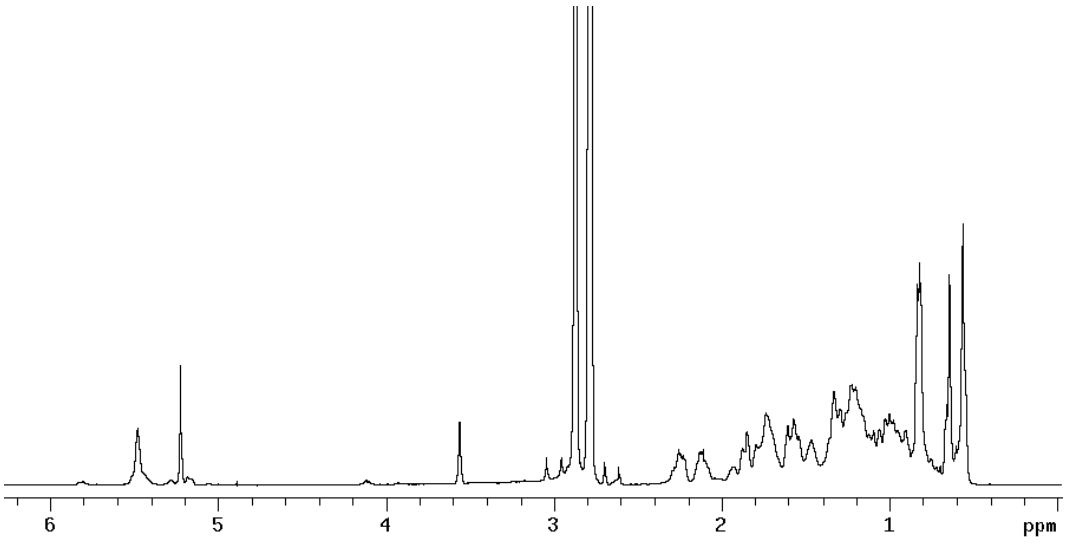
^{1}H NMR spectrum of Compound **8** in $CDCl_3$ at 400 MHz.



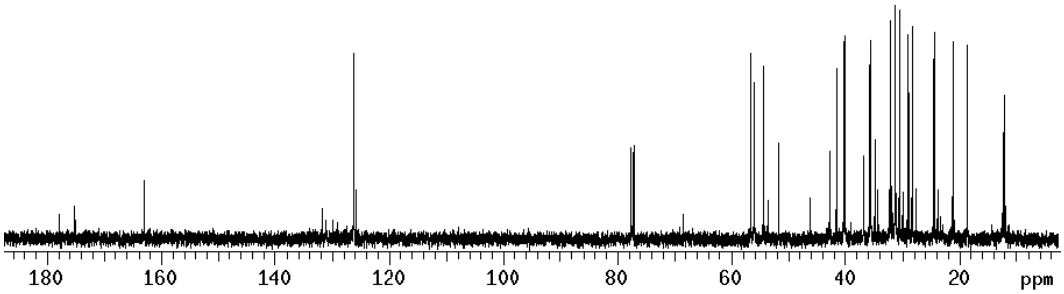
^{13}C NMR spectrum of Compound **8** in $CDCl_3$ at 100 MHz.



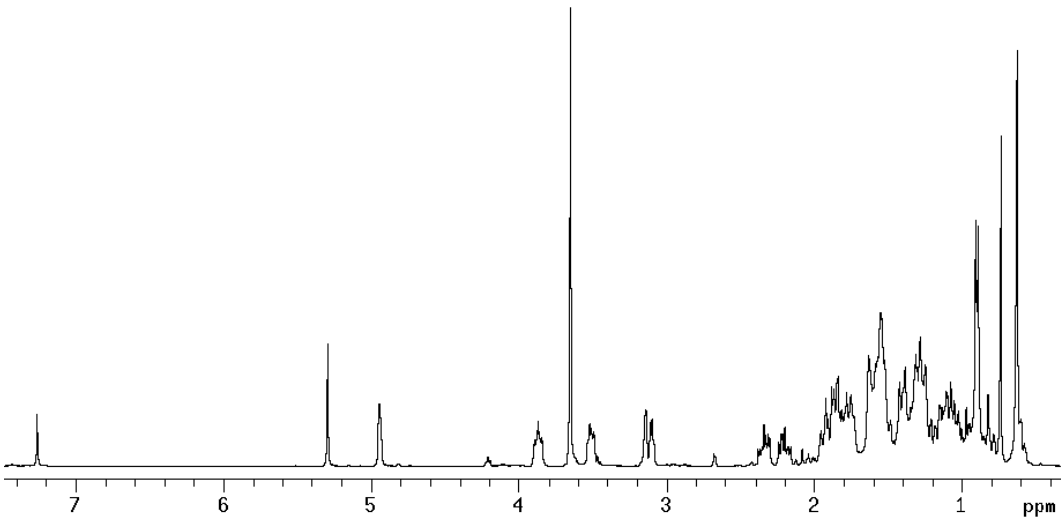
^{1}H NMR spectrum of Compound **9** in $CDCl_3$ at 400 MHz.



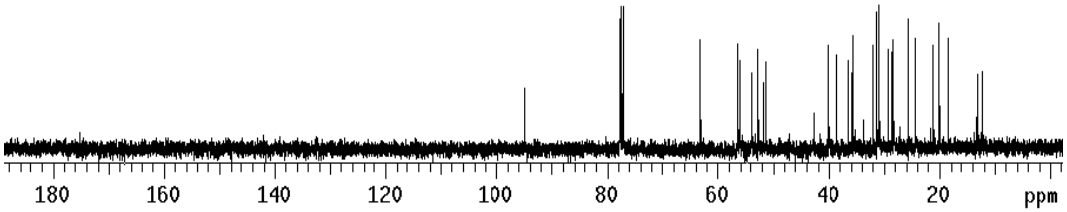
^{13}C NMR spectrum of Compound **9** in $CDCl_3$ at 100 MHz.



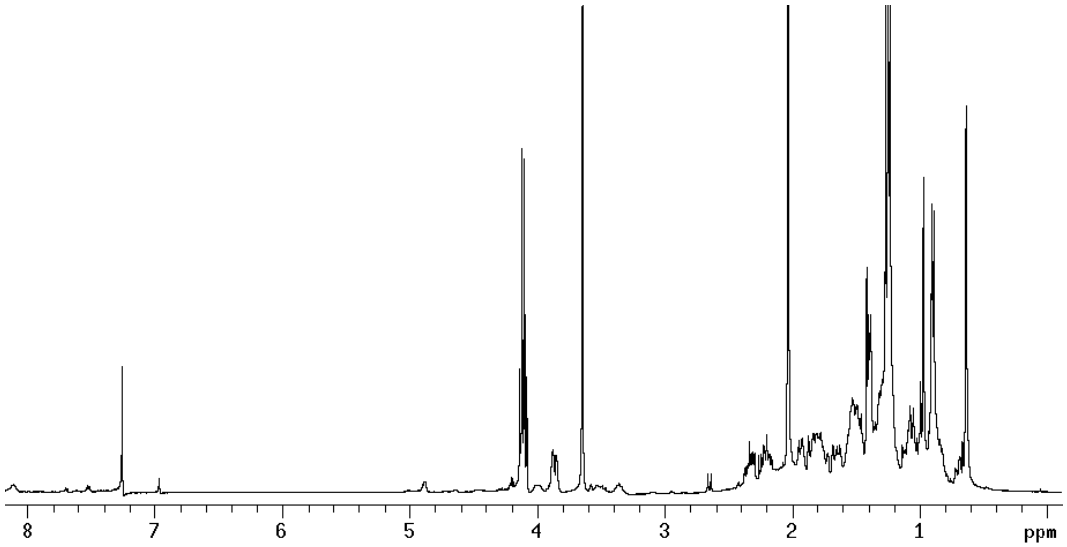
^{1}H NMR spectrum of Compound **10** in $CDCl_3$ at 400 MHz.



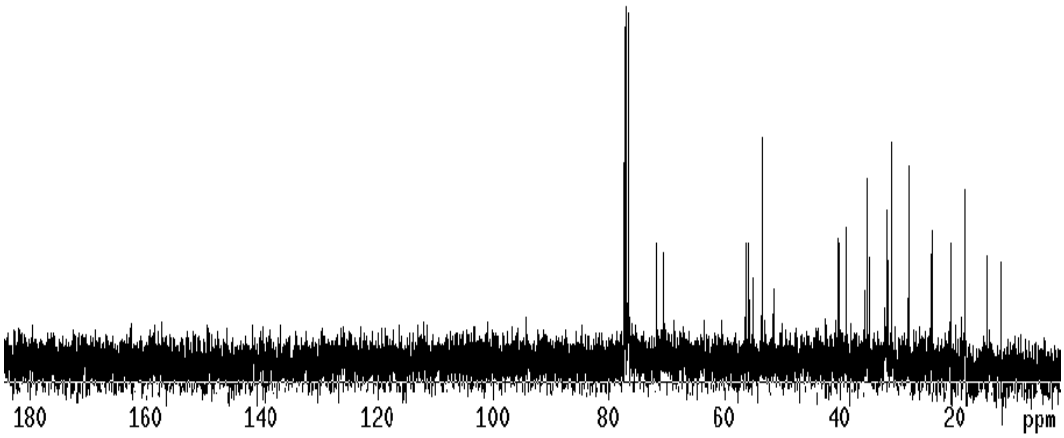
^{13}C NMR spectrum of Compound **10** in $CDCl_3$ at 100 MHz.



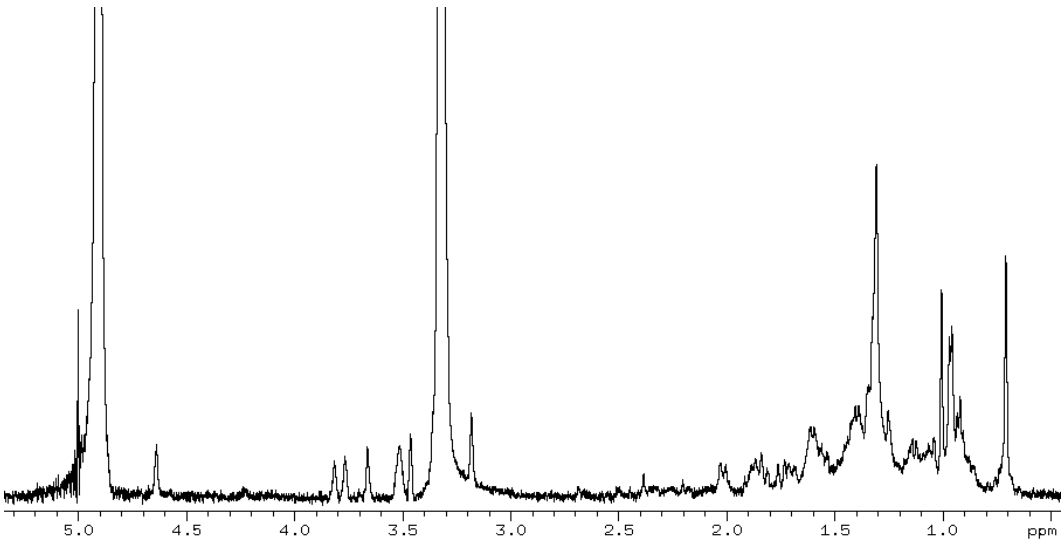
^{1}H NMR spectrum of Compound **11** in $CDCl_3$ at 400 MHz.



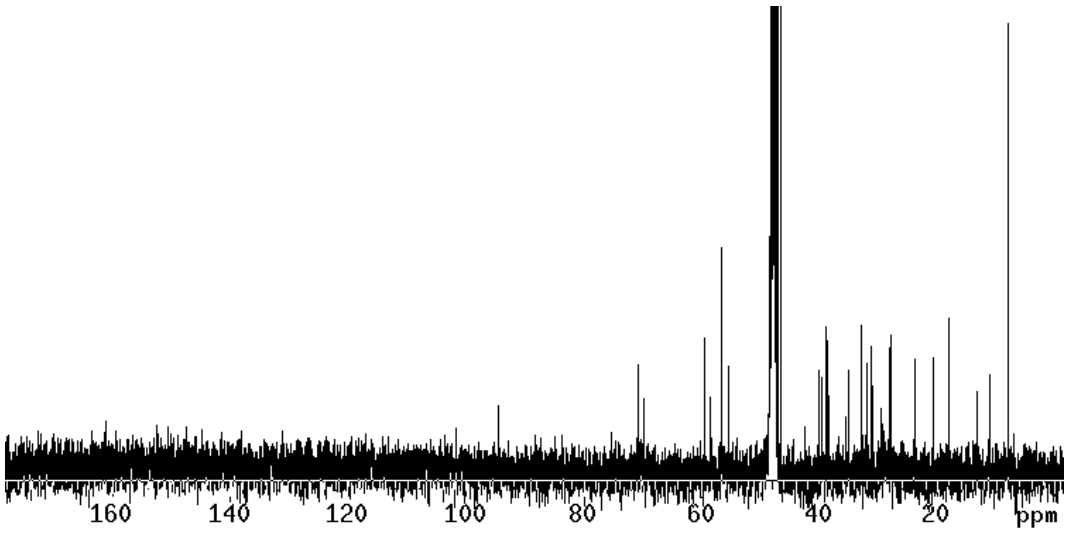
^{13}C NMR spectrum of Compound **11** in $CDCl_3$ at 100 MHz.



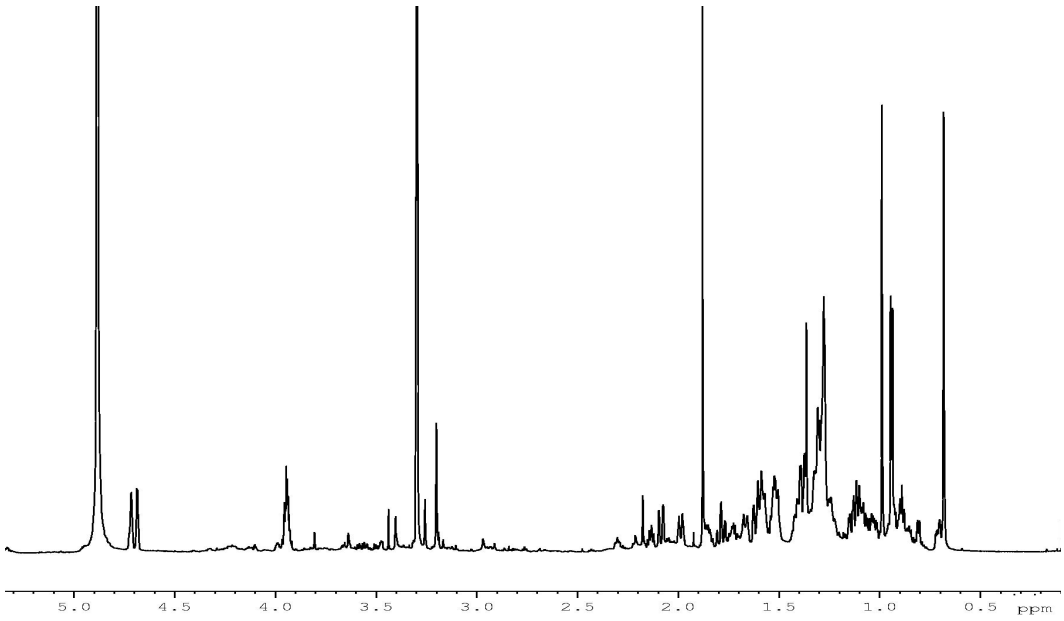
^{1}H NMR spectrum of Compound **12** in CD_3OD at 400 MHz.



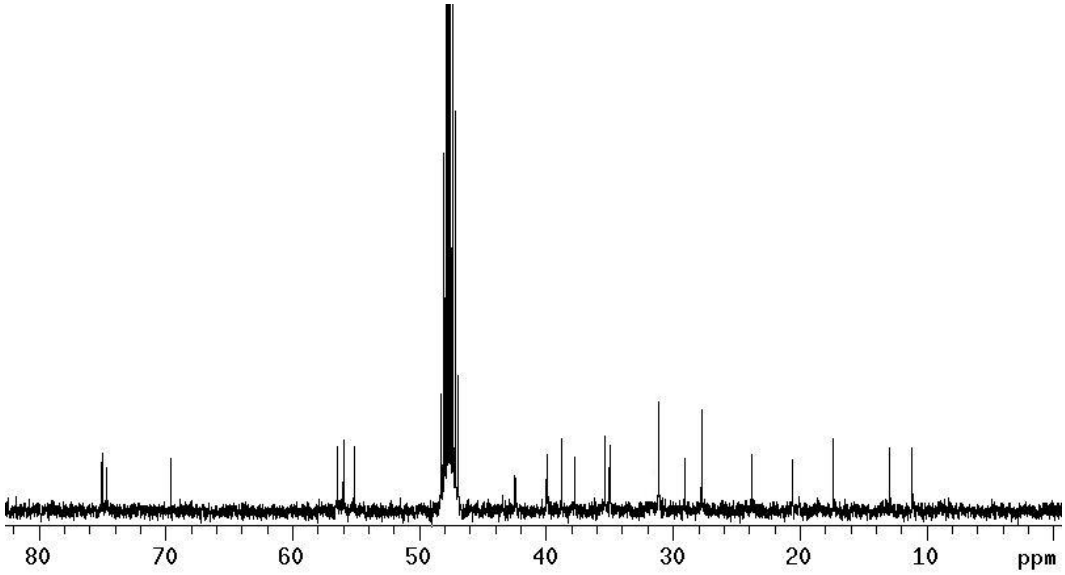
^{13}C NMR spectrum of Compound **12** in CD_3OD at 100 MHz.



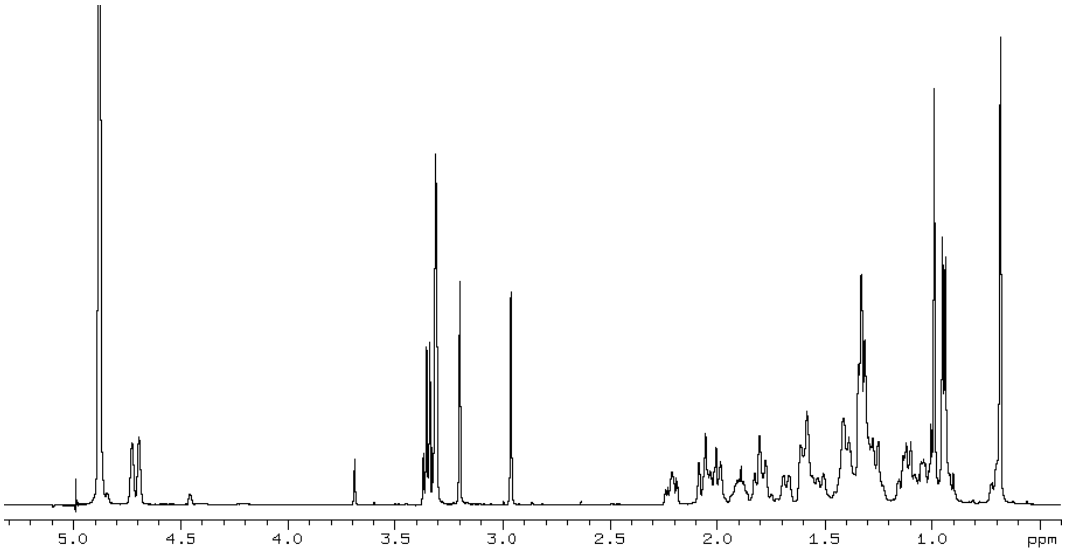
^{1}H NMR spectrum of Compound **1** in CD_3OD at 400 MHz.



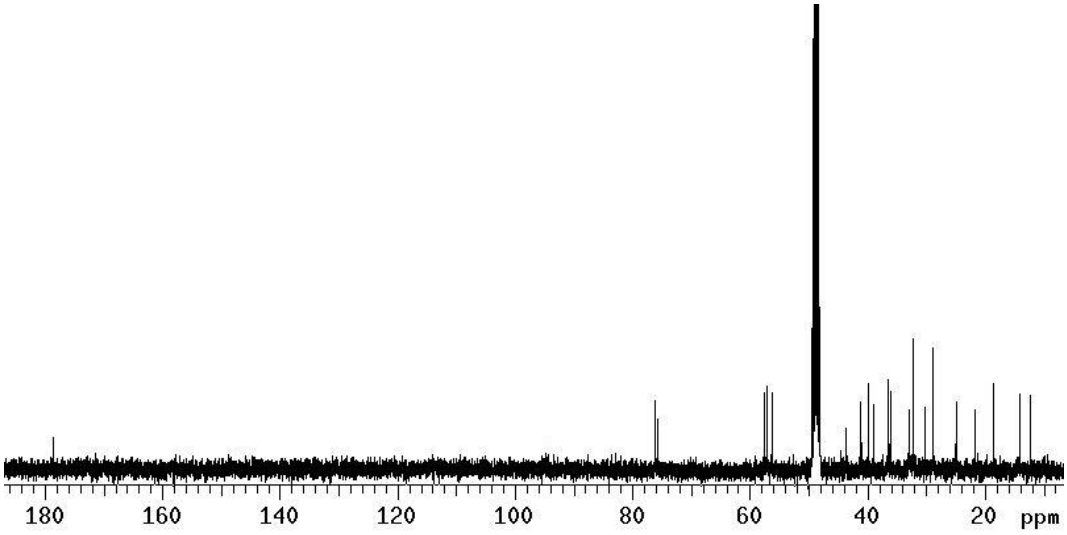
^{13}C NMR spectrum of Compound **1** in CD_3OD at 100 MHz.



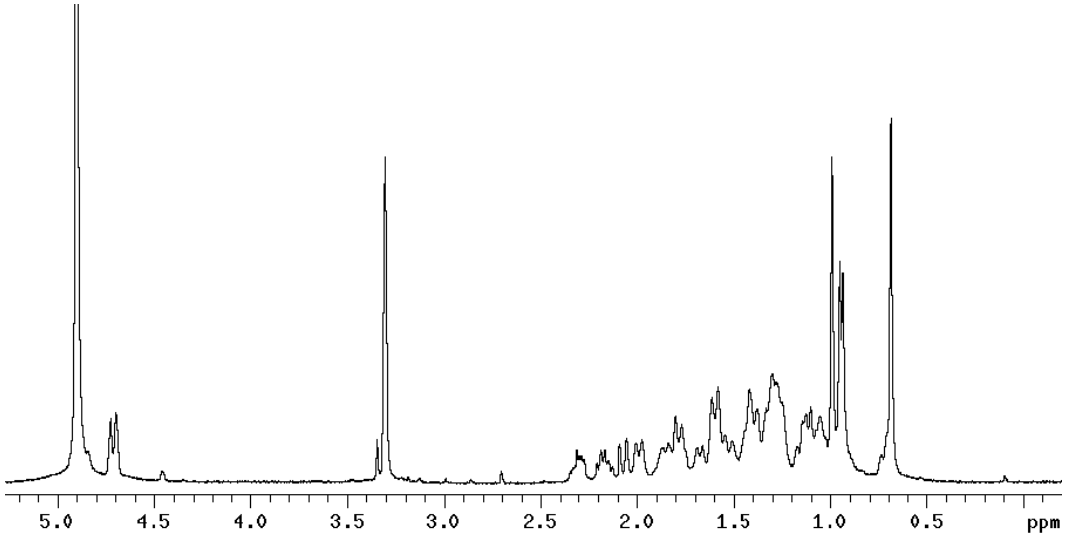
^{1}H NMR spectrum of Compound **13** in CD_3OD at 400 MHz.



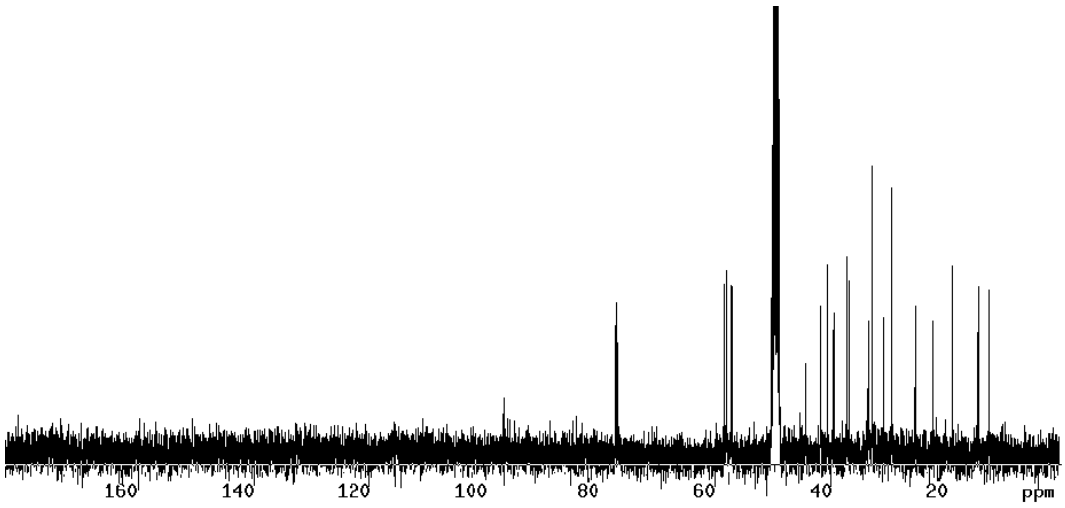
^{13}C NMR spectrum of Compound **13** in CD_3OD at 100 MHz.



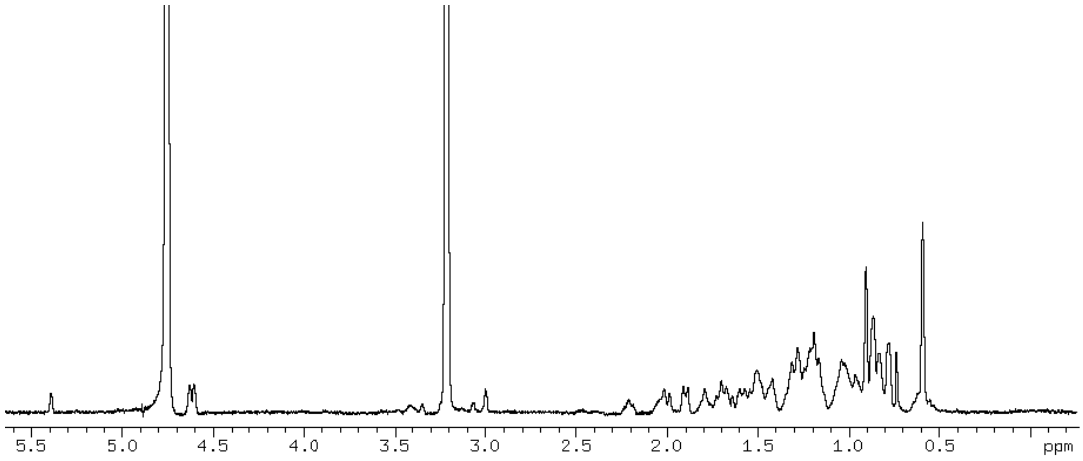
^{1}H NMR spectrum of Compound **15** in CD_3OD at 400 MHz.



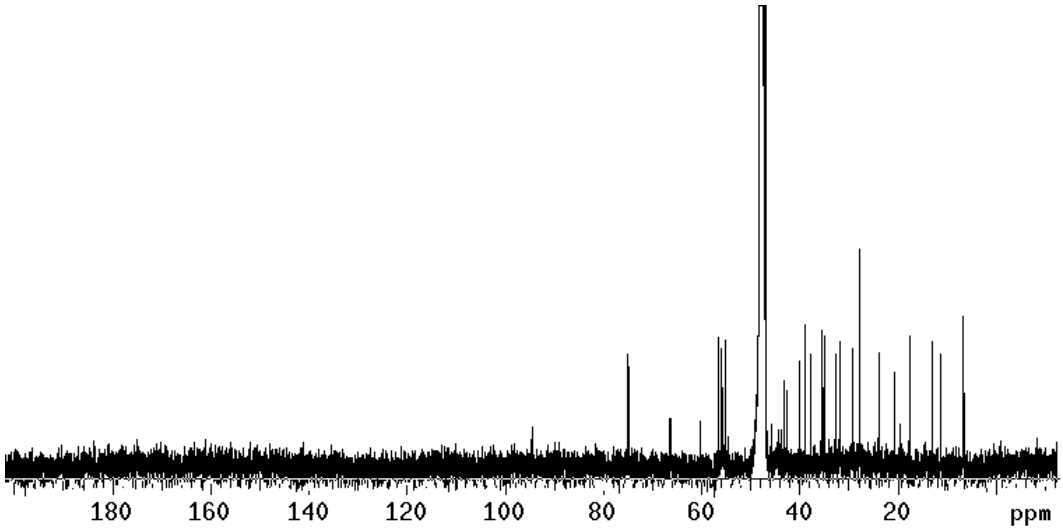
^{13}C NMR spectrum of Compound **15** in CD_3OD at 100 MHz.



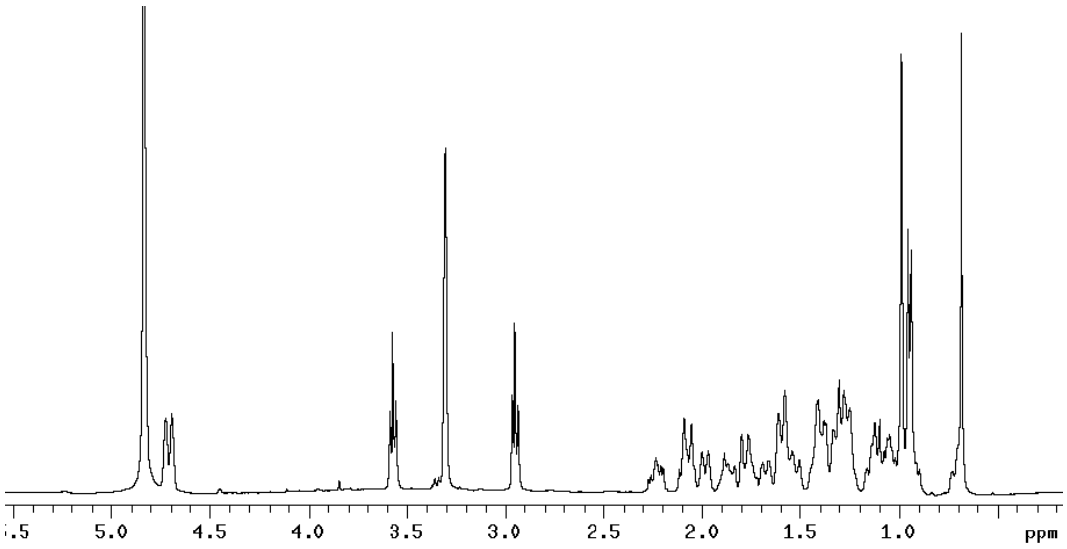
^{1}H NMR spectrum of Compound **16** in CD_3OD at 400 MHz.



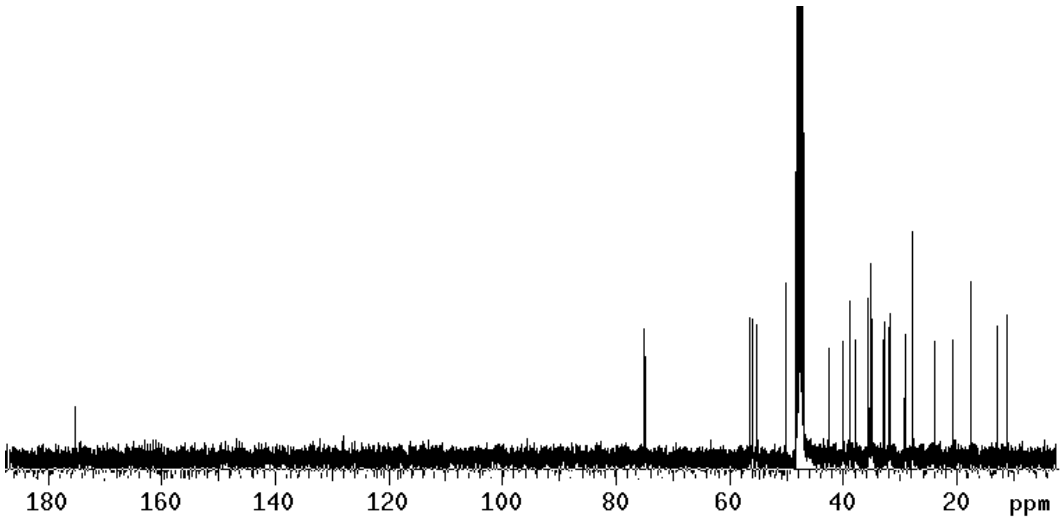
^{13}C NMR spectrum of Compound **16** in CD_3OD at 100 MHz.



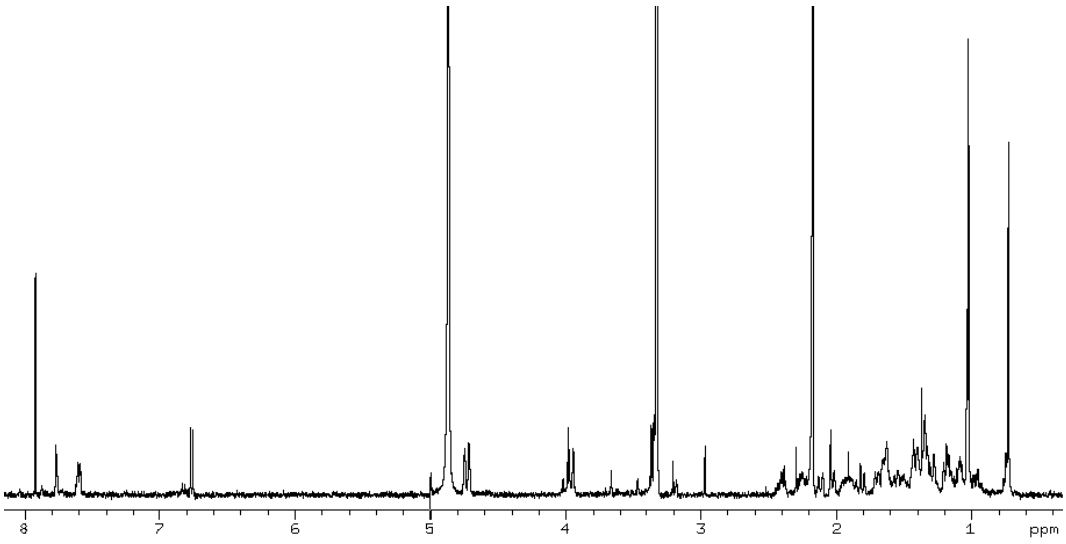
^{1}H NMR spectrum of Compound **17** in CD_3OD at 400 MHz.



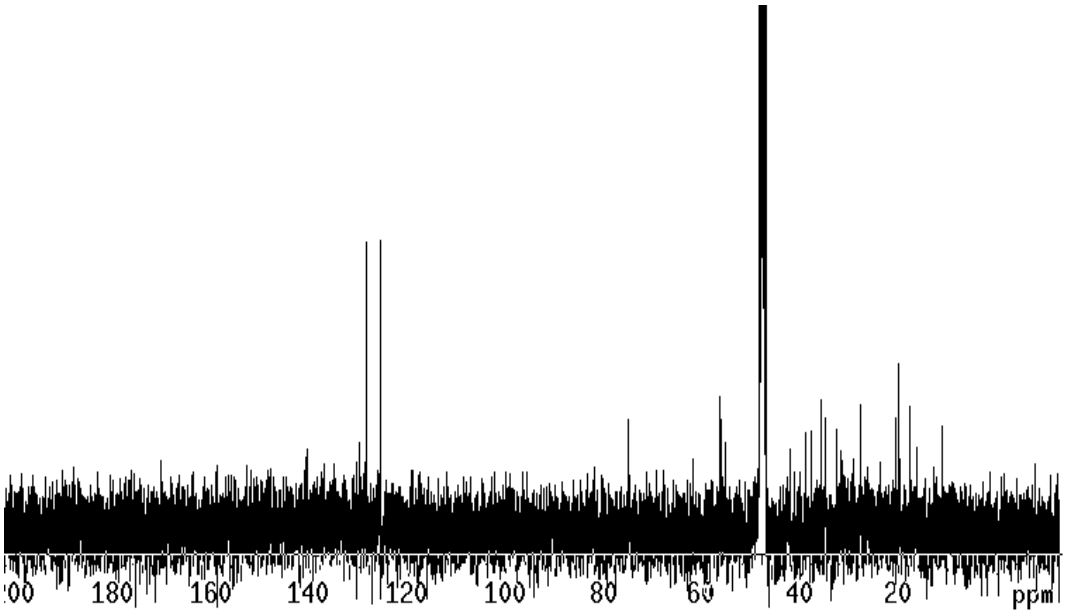
^{13}C NMR spectrum of Compound **17** in CD_3OD at 100 MHz.



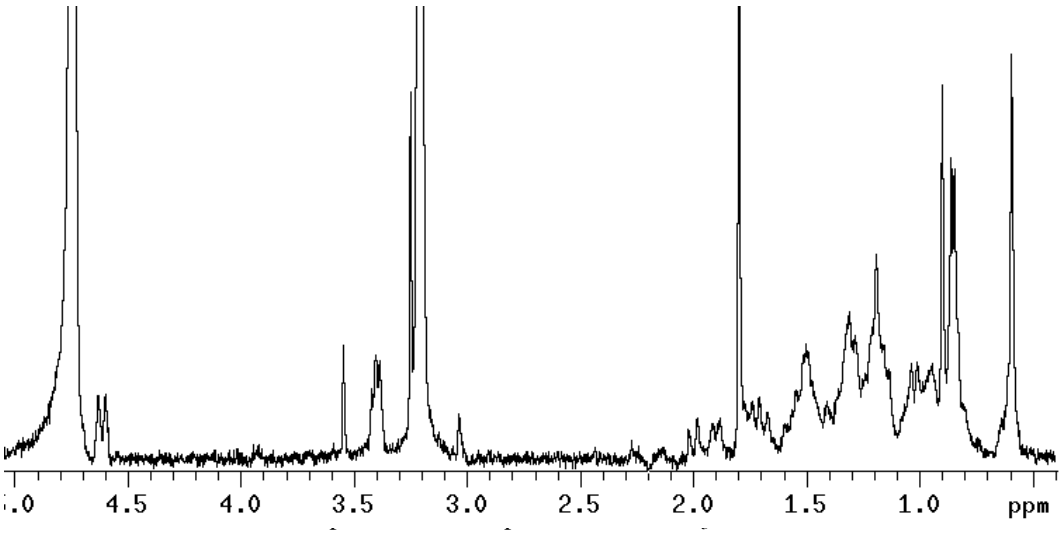
^{1}H NMR spectrum of Compound **18** in CD_3OD at 400 MHz.



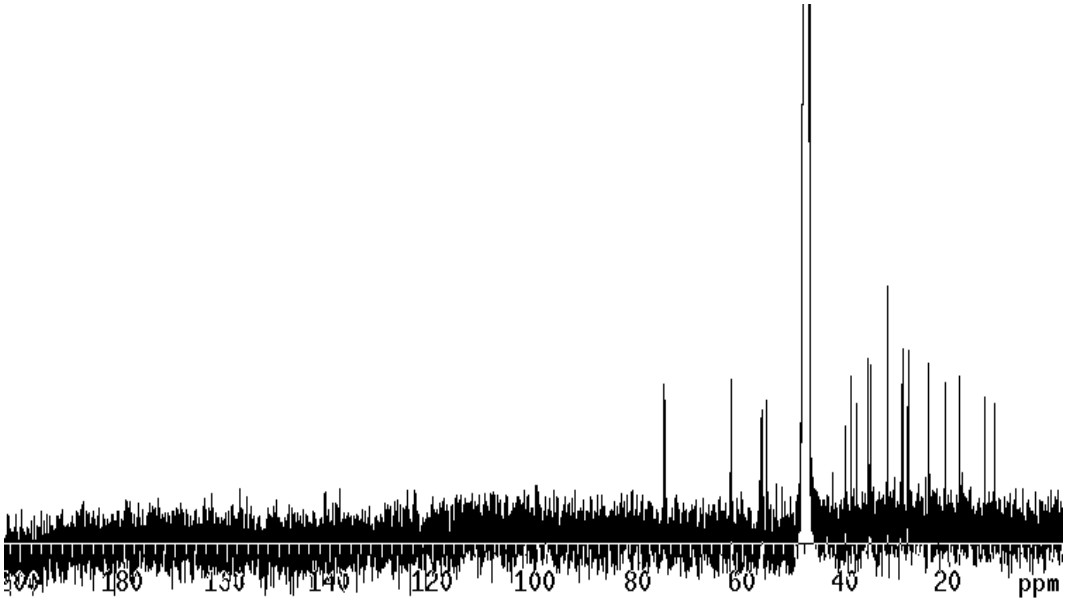
^{13}C NMR spectrum of Compound **18** in CD_3OD at 100 MHz.



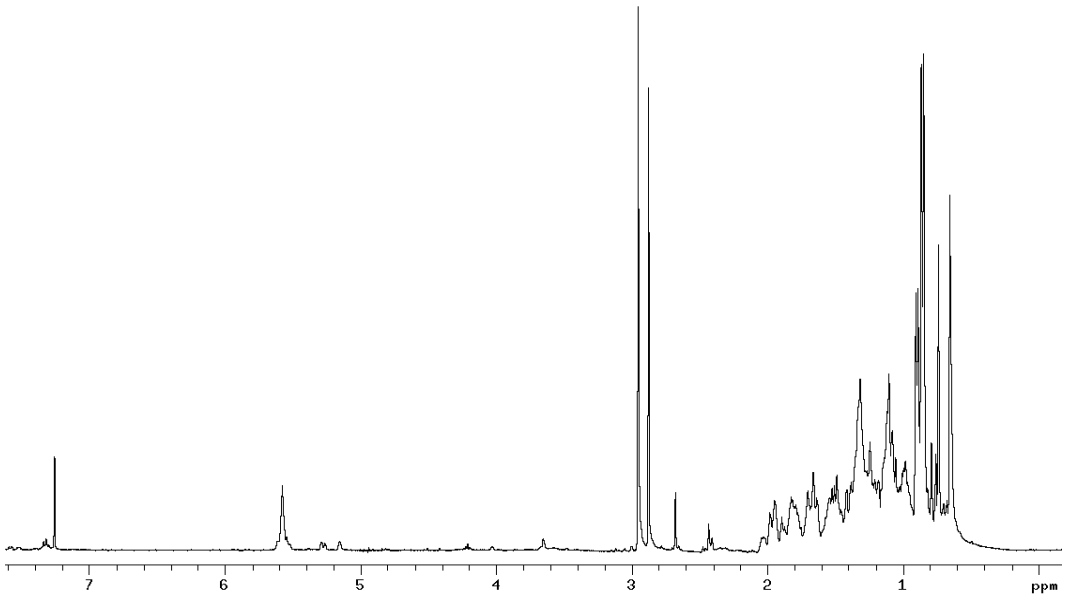
^{1}HNMR spectrum of Compound **19** in CD_3OD at 400 MHz.



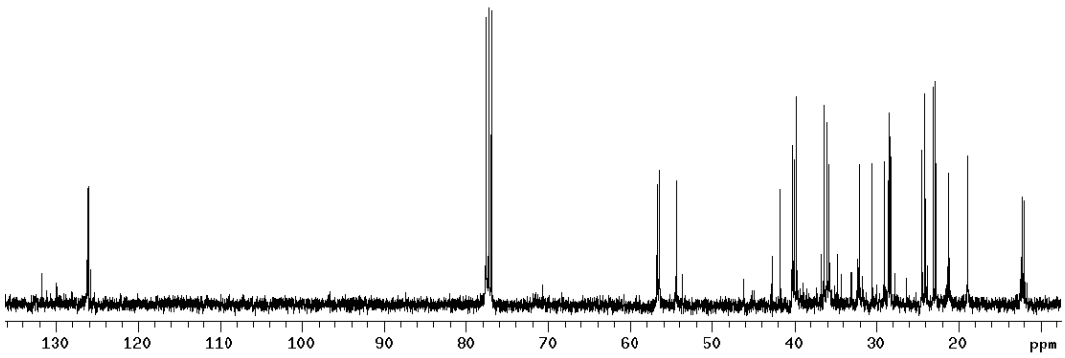
^{13}CNMR spectrum of Compound **19** in CD_3OD at 100 MHz.



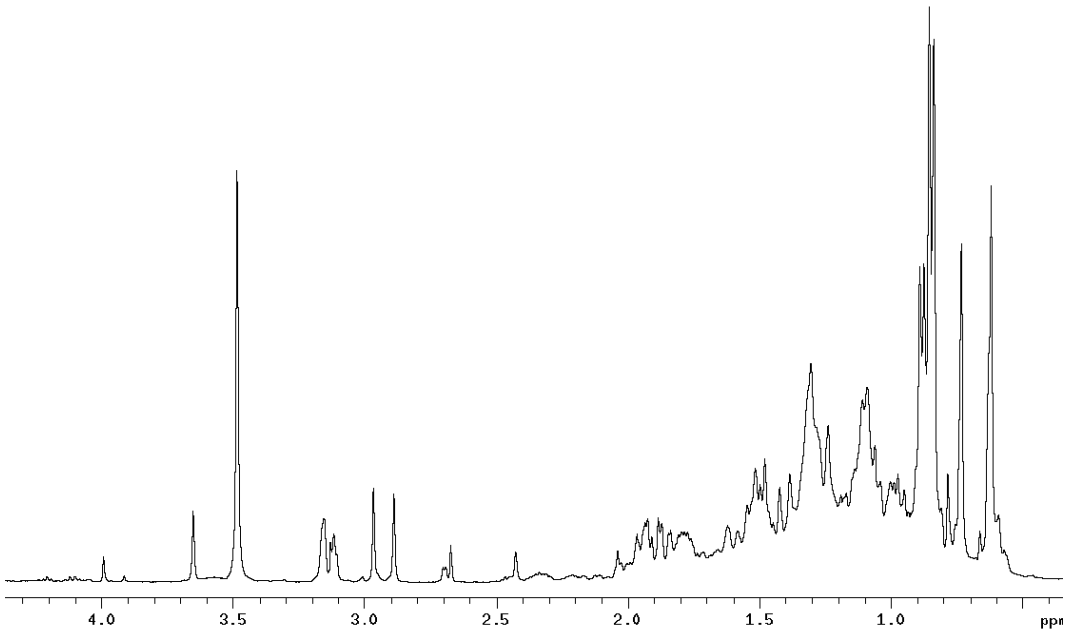
^{1}H NMR spectrum of Compound **20** in $CDCl_3$ at 400 MHz.



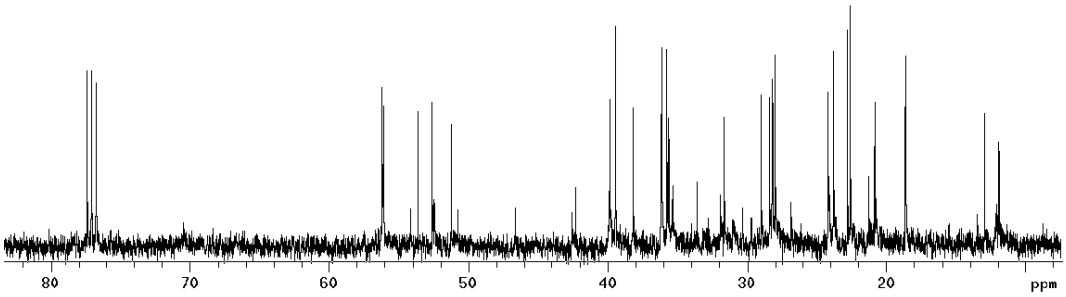
^{13}C NMR spectrum of Compound **20** in $CDCl_3$ at 100 MHz.



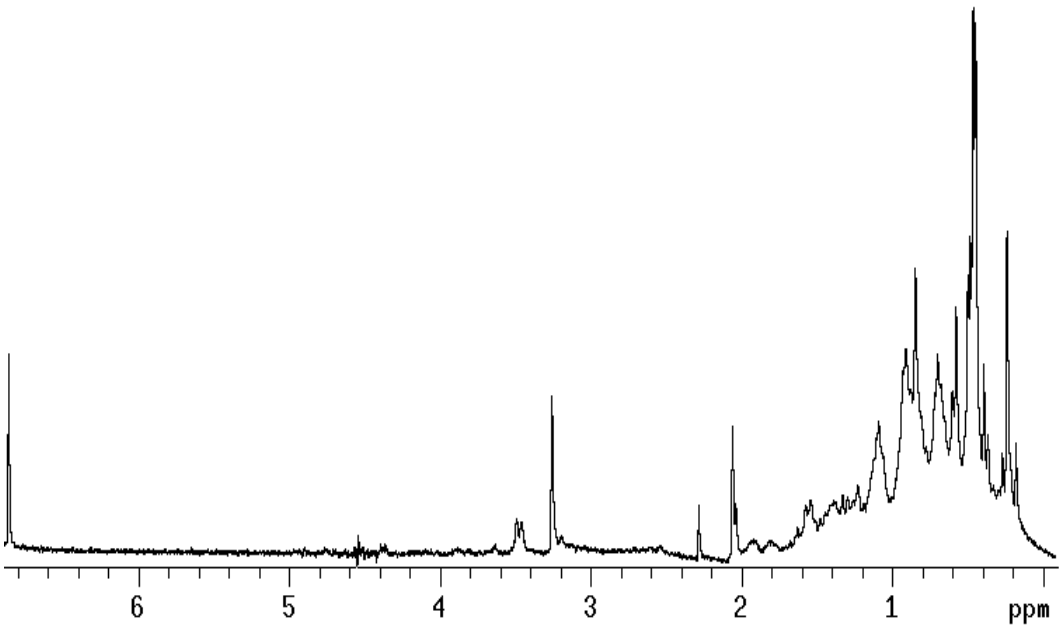
^{1}H NMR spectrum of Compound **21** in $CDCl_3$ at 400 MHz.



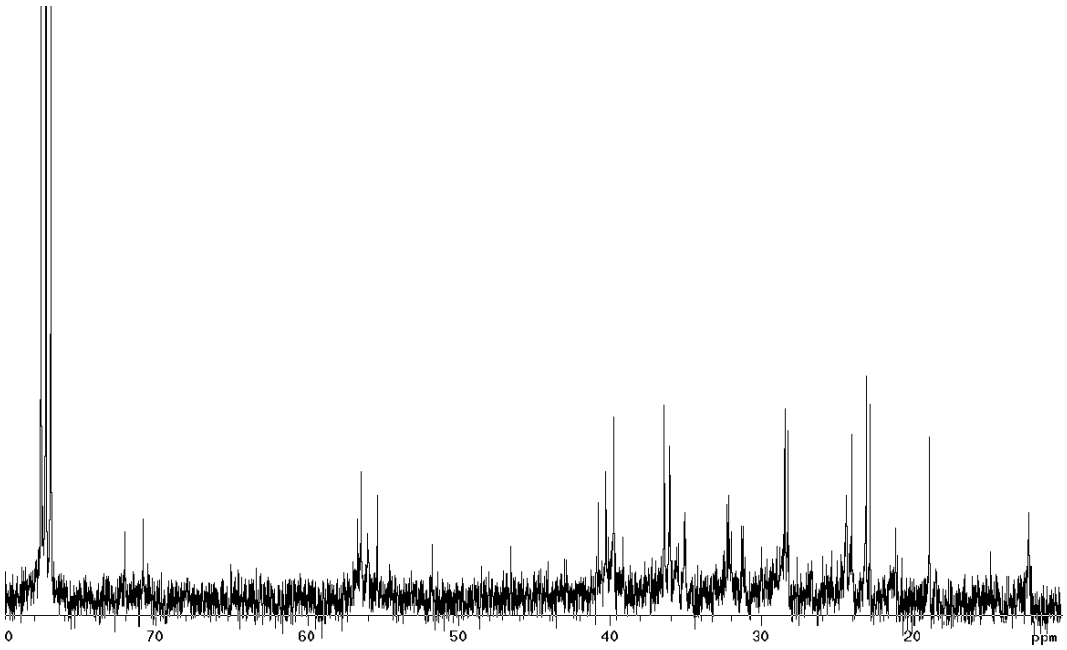
^{13}C NMR spectrum of Compound **21** in $CDCl_3$ at 100 MHz.



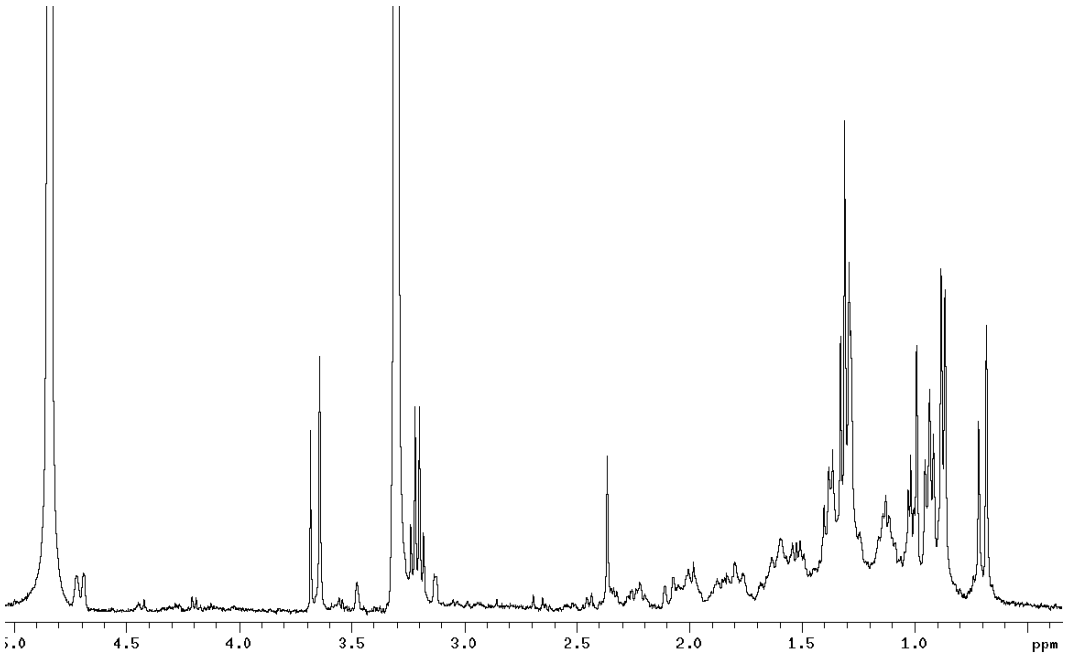
^{1}H NMR spectrum of Compound **22** in $CDCl_3$ at 400 MHz.



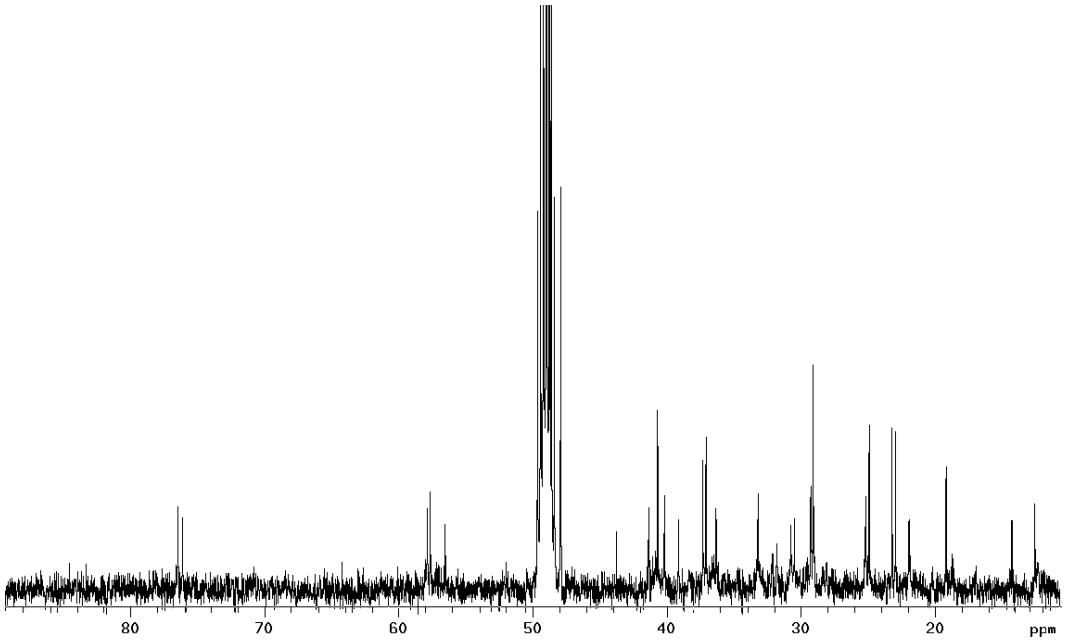
^{13}C NMR spectrum of Compound **22** in $CDCl_3$ at 100 MHz.



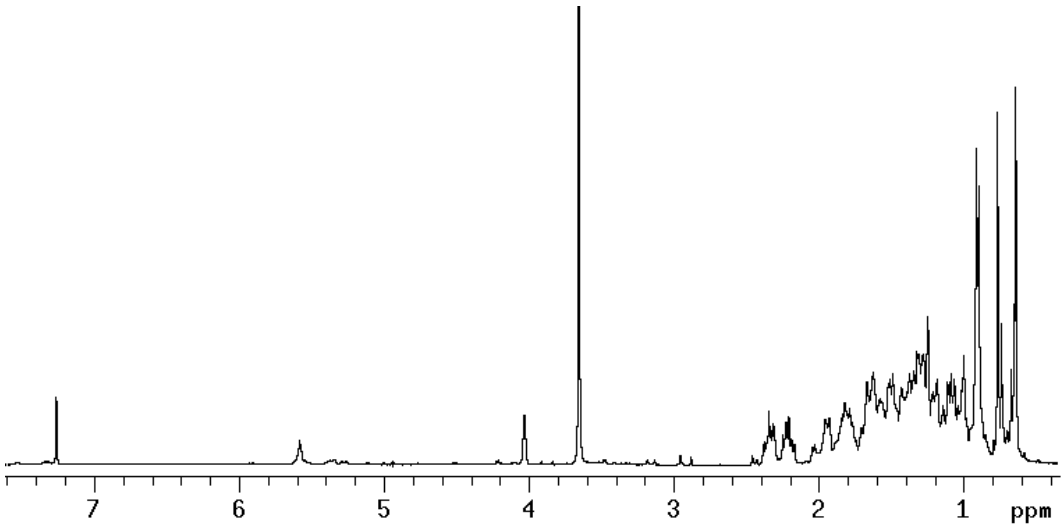
^{1}H NMR spectrum of Compound **23** in CD_3OD at 400 MHz.



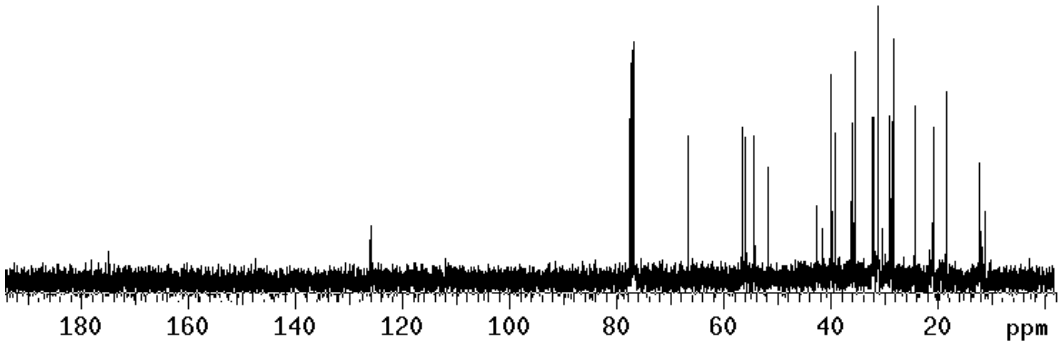
^{13}C NMR spectrum of Compound **23** in CD_3OD at 100 MHz.



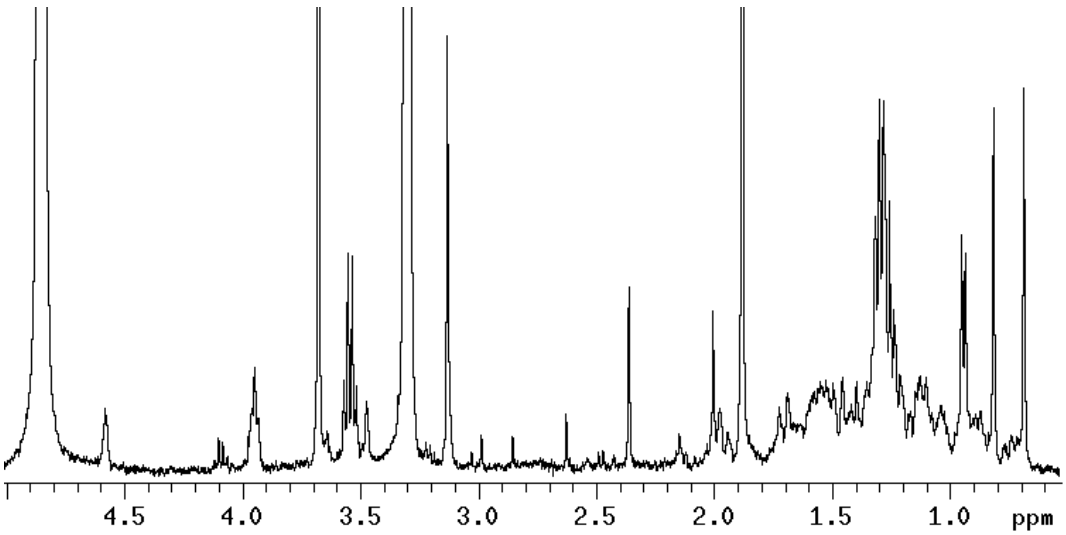
^{1}H NMR spectrum of Compound **24** in $CDCl_3$ at 400 MHz.



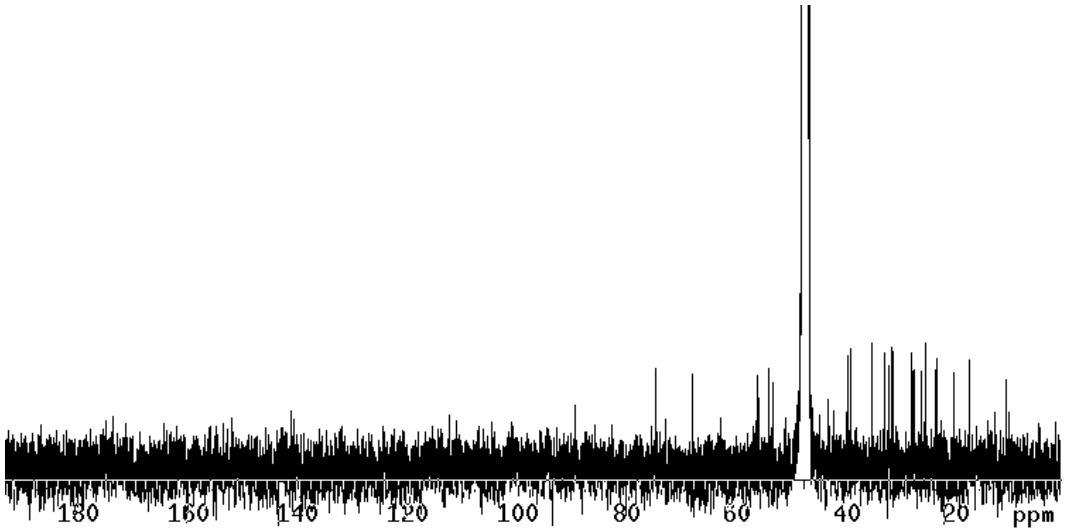
^{13}C NMR spectrum of Compound **24** in $CDCl_3$ at 100 MHz.



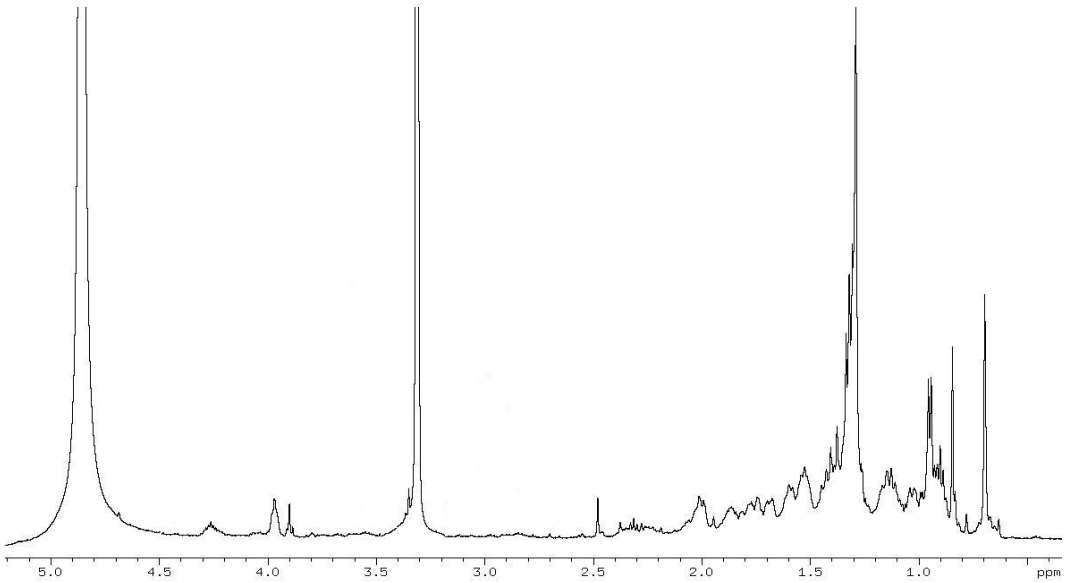
^{1}H NMR spectrum of Compound **25** in CD_3OD at 400 MHz.



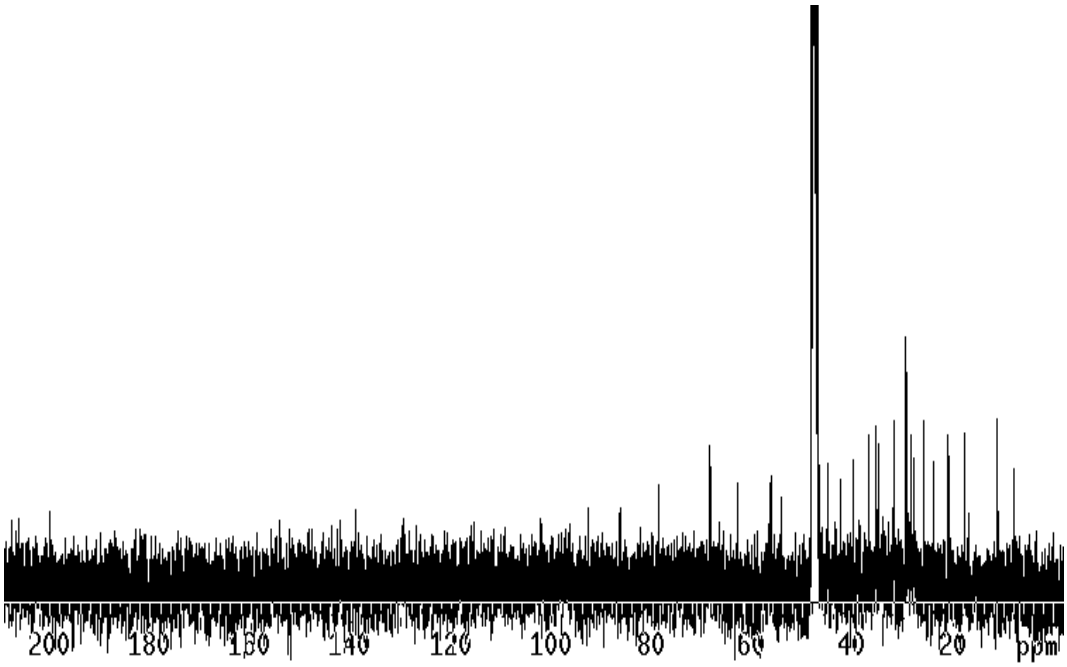
^{13}C NMR spectrum of Compound **25** in CD_3ODat 100 MHz



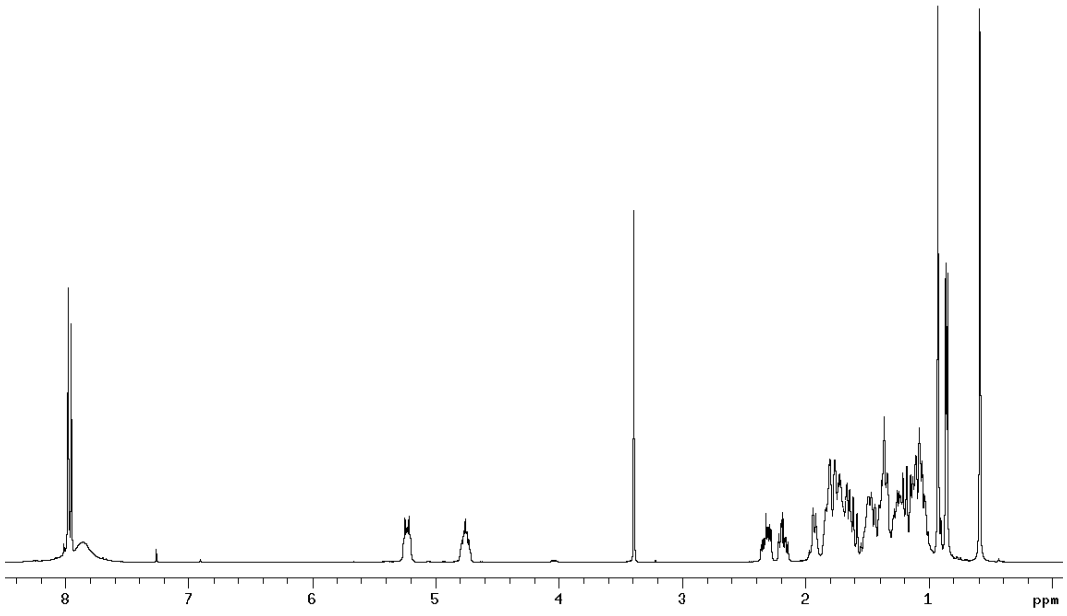
^{1}H NMR spectrum of Compound **26** in CD_3ODat 400 MHz.



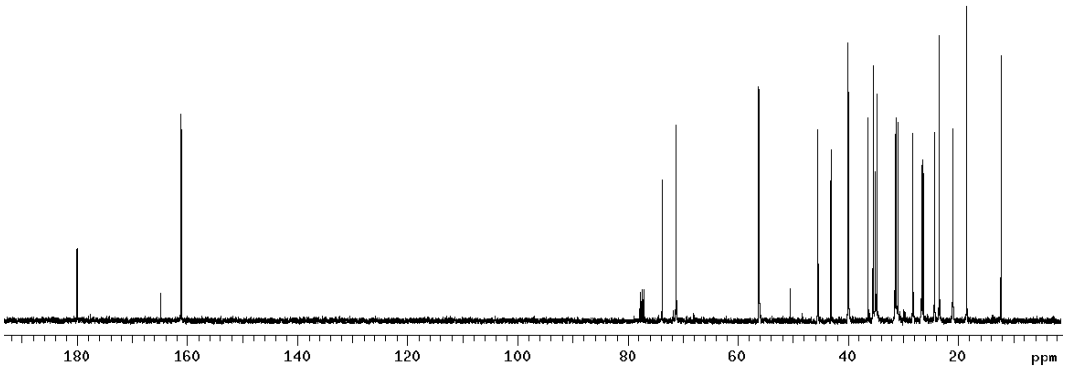
^{13}C NMR spectrum of Compound **26** in CD_3ODat 100 MHz.



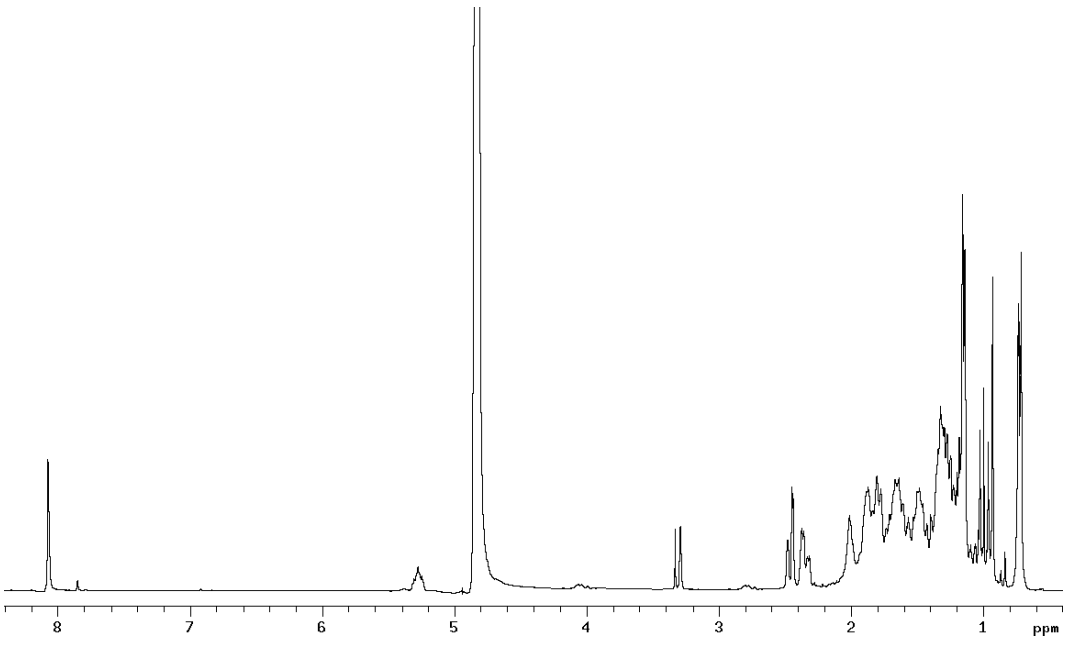
^{1}H NMR spectrum of Compound **27** in $CDCl_3$ at 400 MHz.



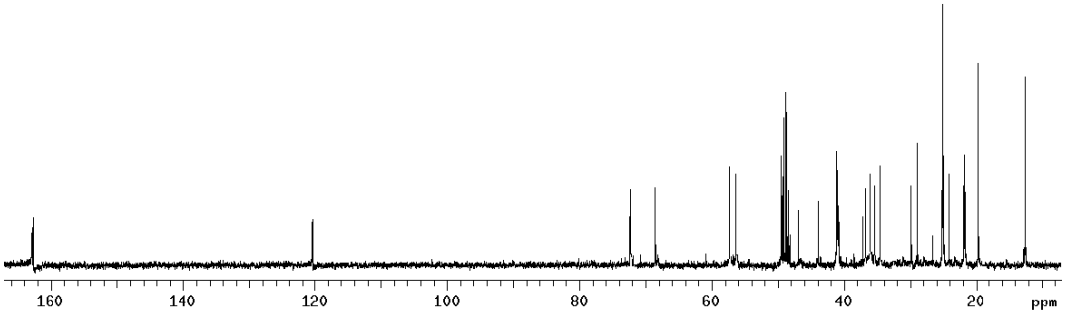
^{13}C NMR spectrum of Compound **27** in $CDCl_3$ at 100 MHz.



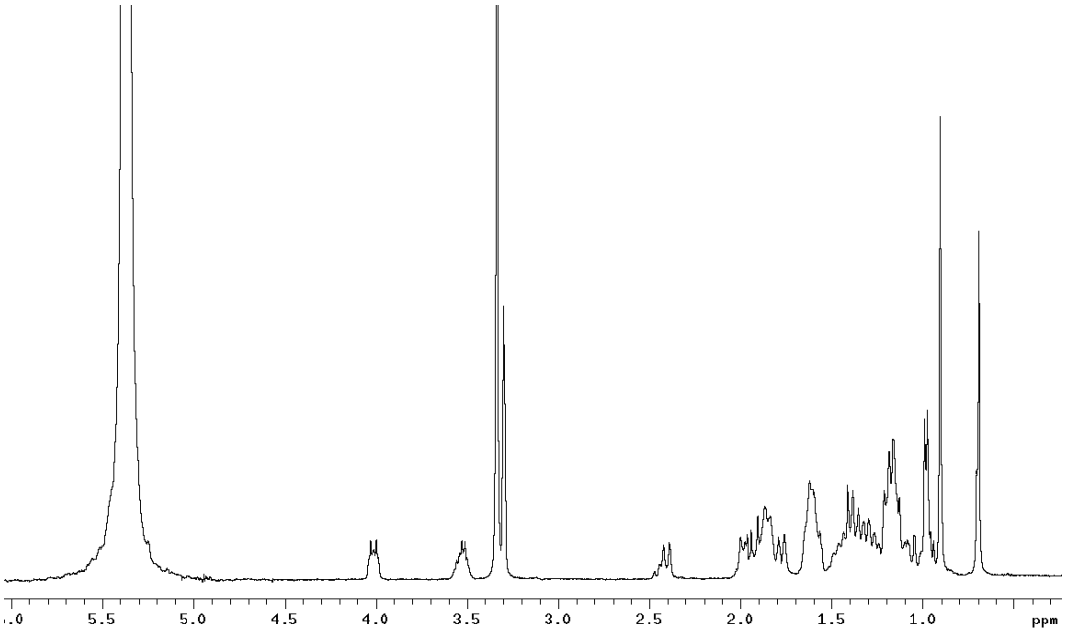
^{1}H NMR spectrum of Compound **28** in CD_3ODat 400 MHz.



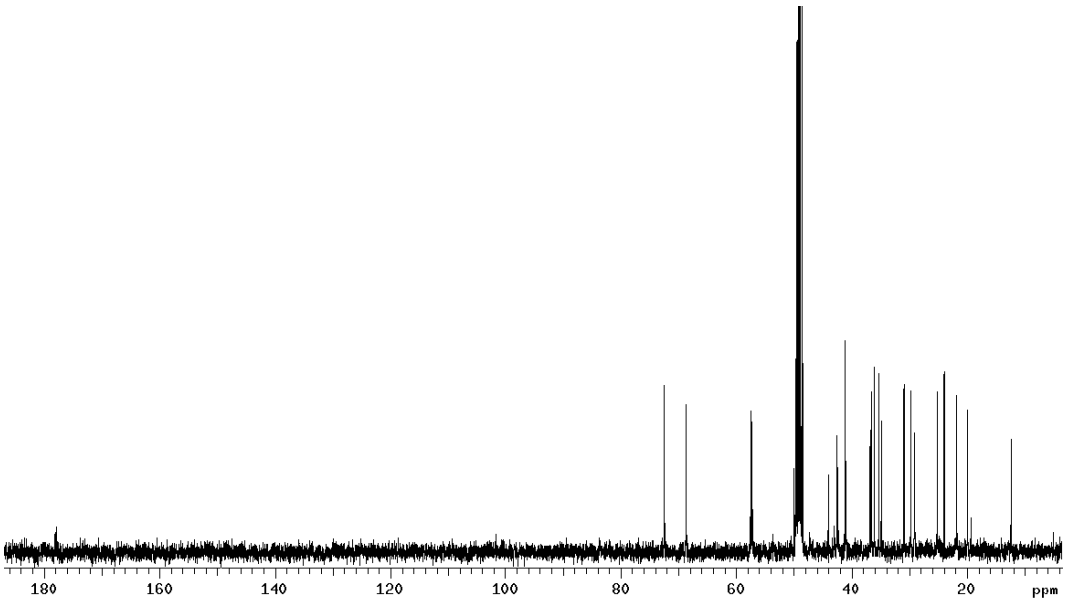
^{13}C NMR spectrum of Compound **28** in CD_3ODat 100 MHz.



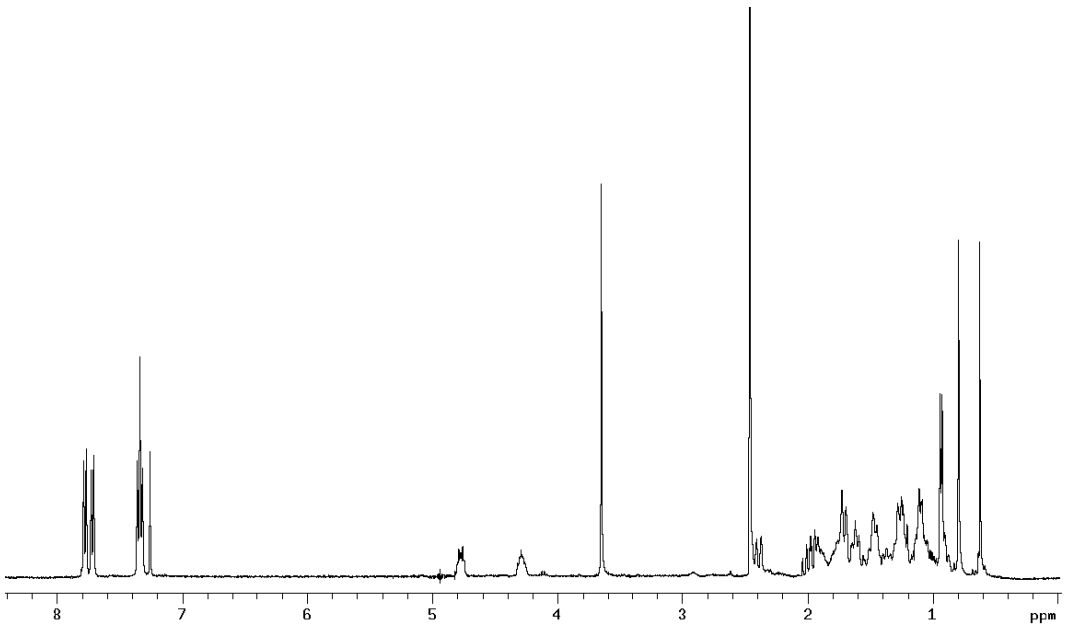
^{1}H NMR spectrum of Compound **29** in CD_3ODat 100 MHz.



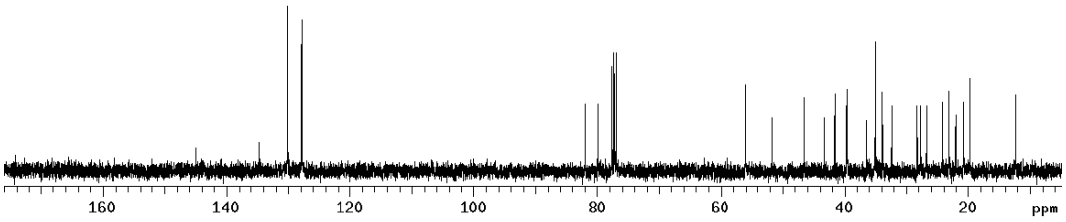
^{13}C NMR spectrum of Compound **29** in CD_3ODat 100 MHz.



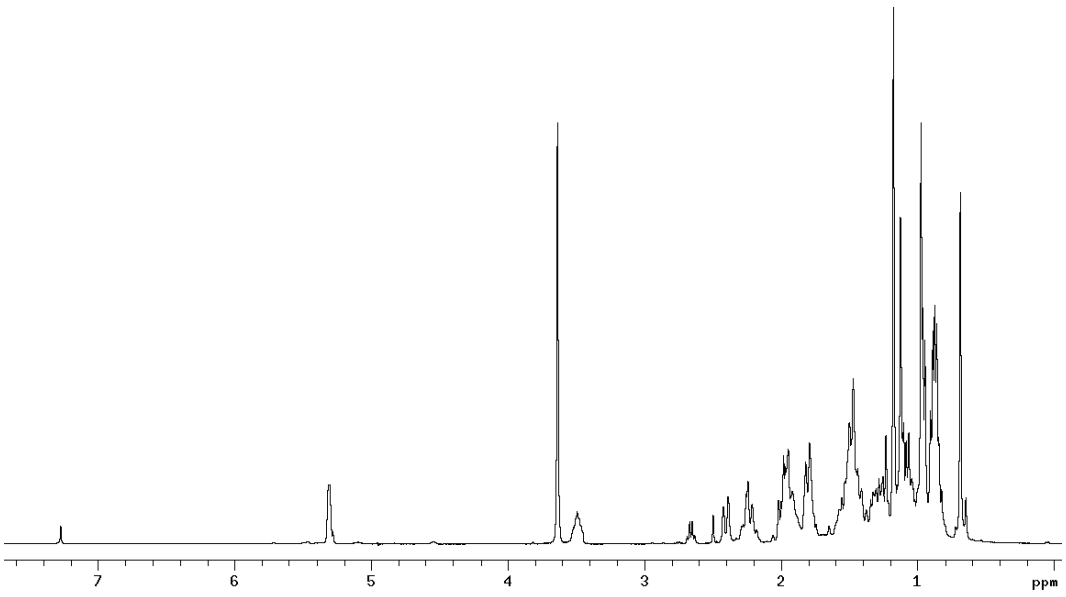
^{1}H NMR spectrum of Compound **30** in $CDCl_3$ at 400 MHz.



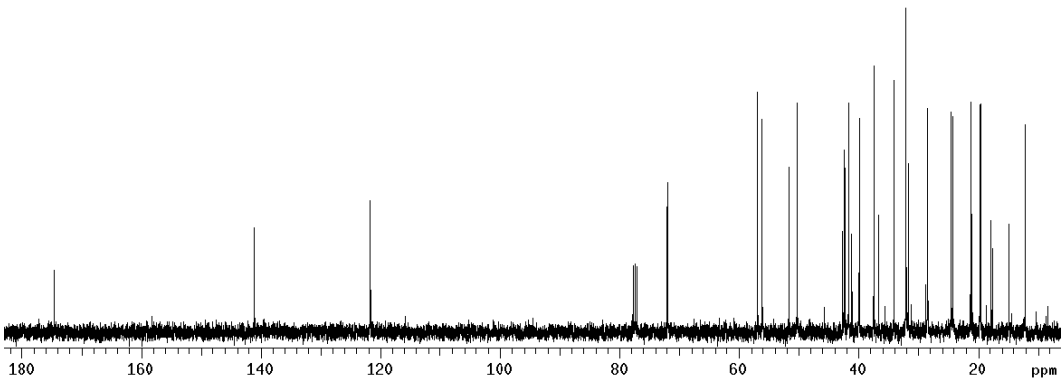
^{13}C NMR spectrum of Compound **30** in $CDCl_3$ at 100 MHz.



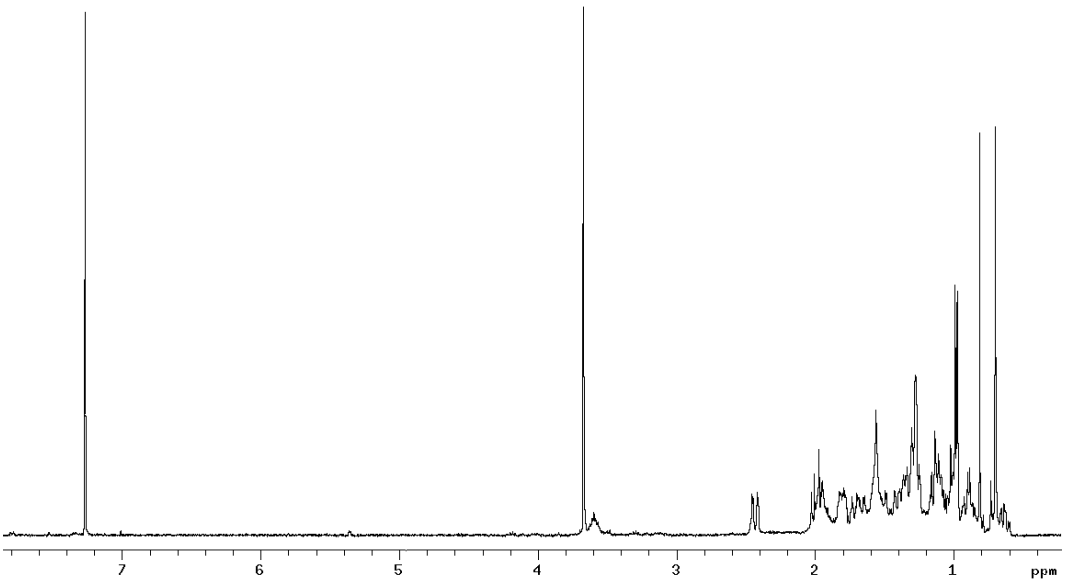
^{1}H NMR spectrum of Compound **31** in $CDCl_3$ at 400 MHz.



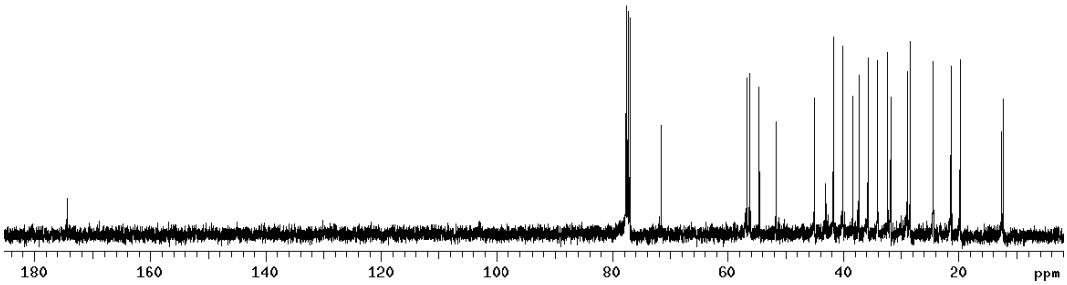
^{13}C NMR spectrum of Compound **31** in $CDCl_3$ at 100 MHz.



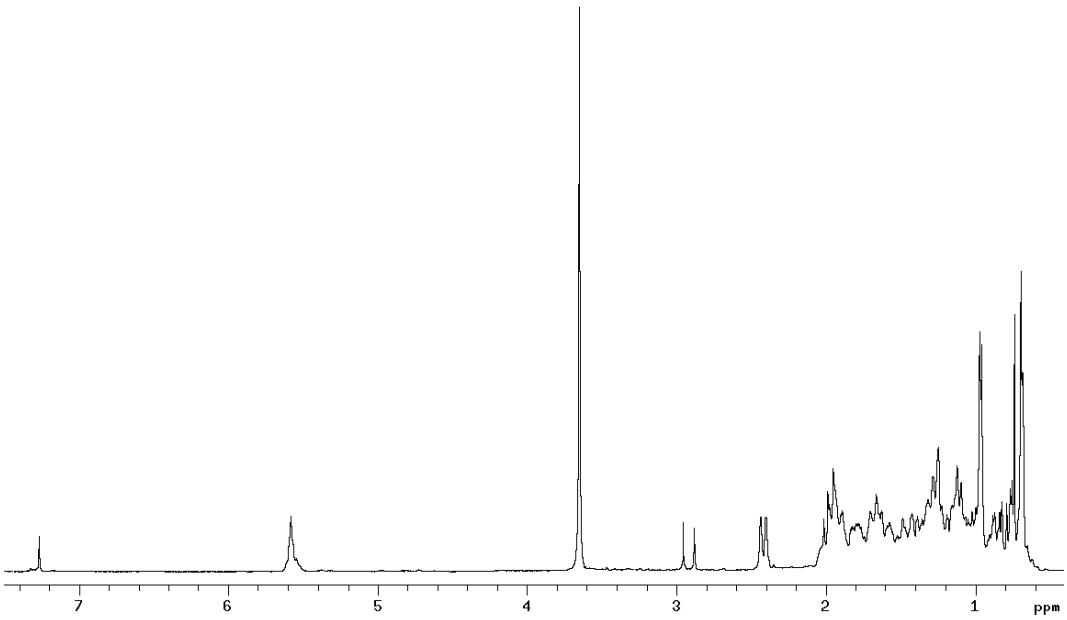
^{1}H NMR spectrum of Compound **32** in $CDCl_3$ at 400 MHz.



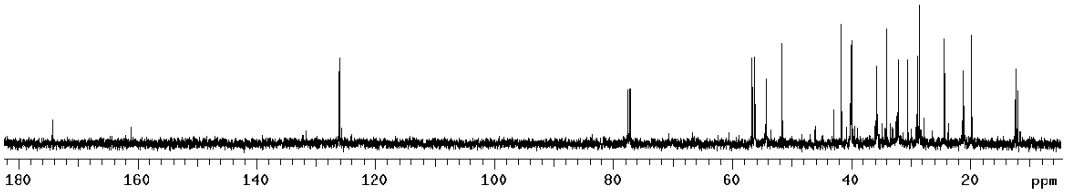
^{13}C NMR spectrum of Compound **32** in $CDCl_3$ at 100 MHz.



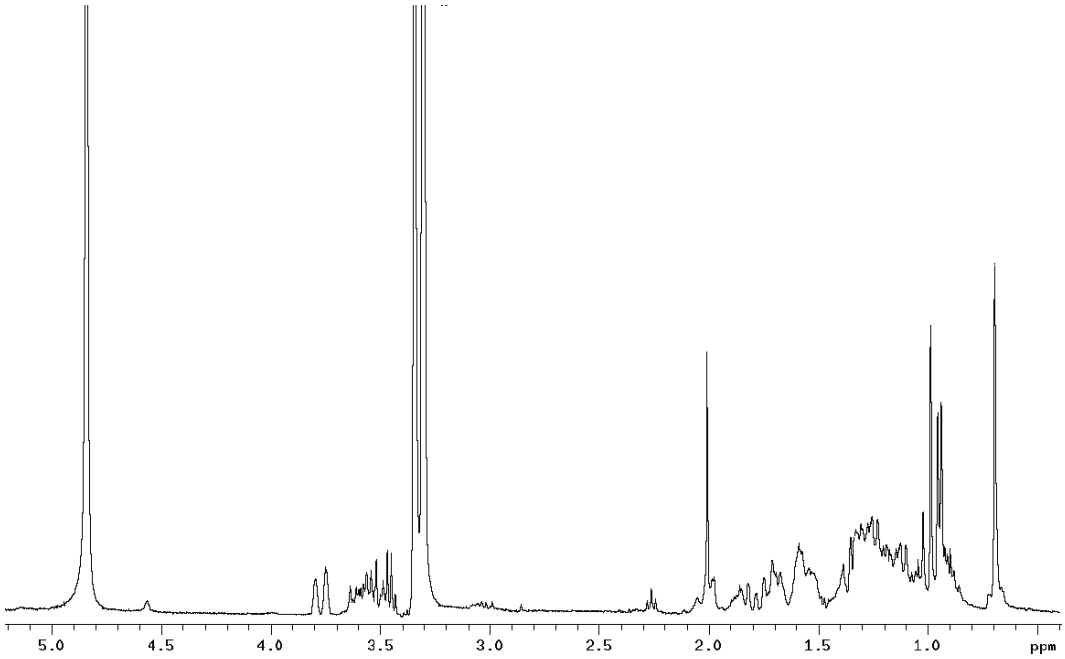
^{1}H NMR spectrum of Compound **33** in $CDCl_3$ at 400 MHz.



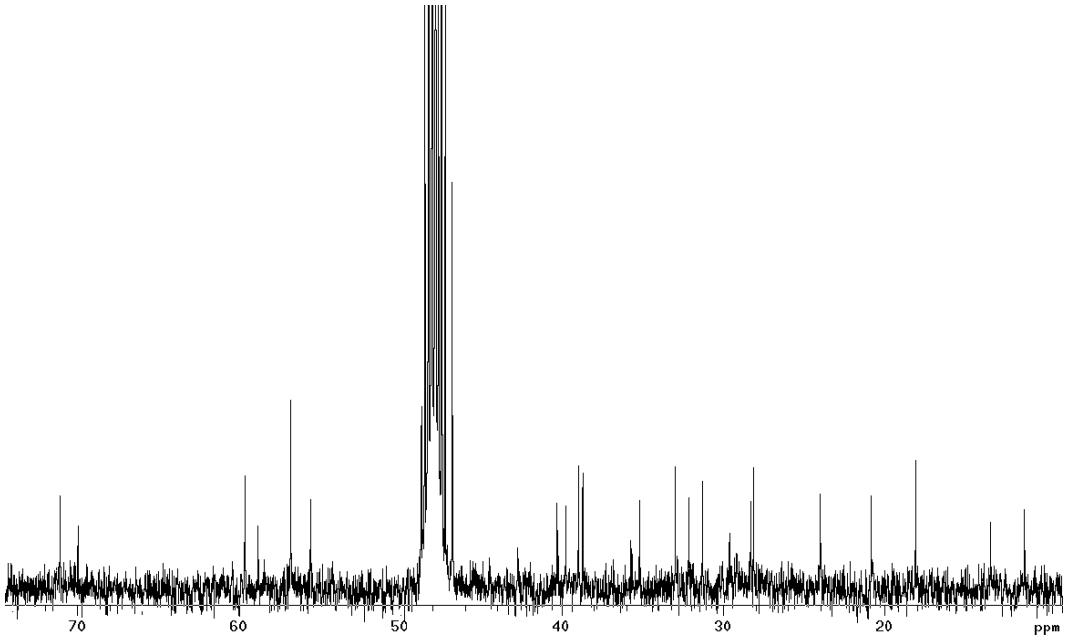
^{13}C NMR spectrum of Compound **33** in $CDCl_3$ at 100 MHz.



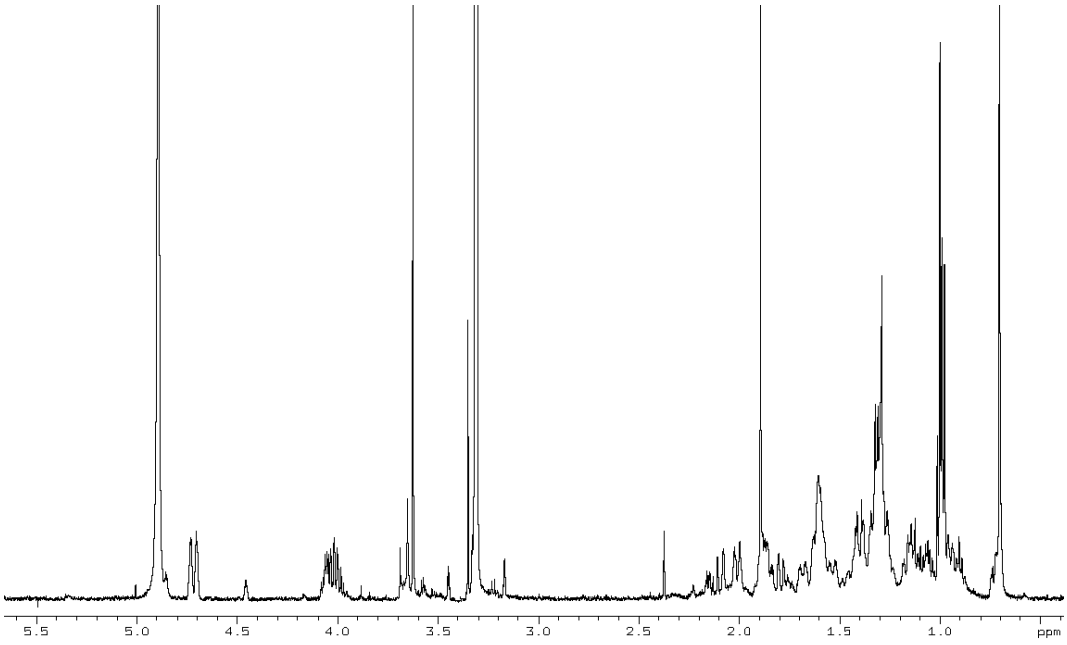
^{1}H NMR spectrum of Compound **34** in CD_3OD at 400 MHz.



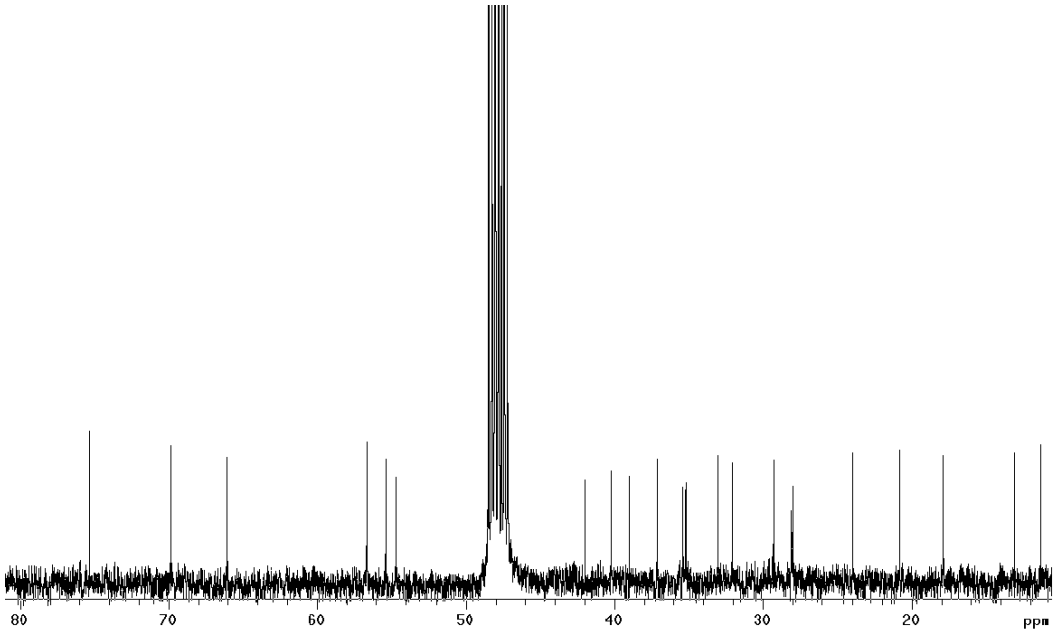
^{13}C NMR spectrum of Compound **34** in CD_3OD at 100 MHz.



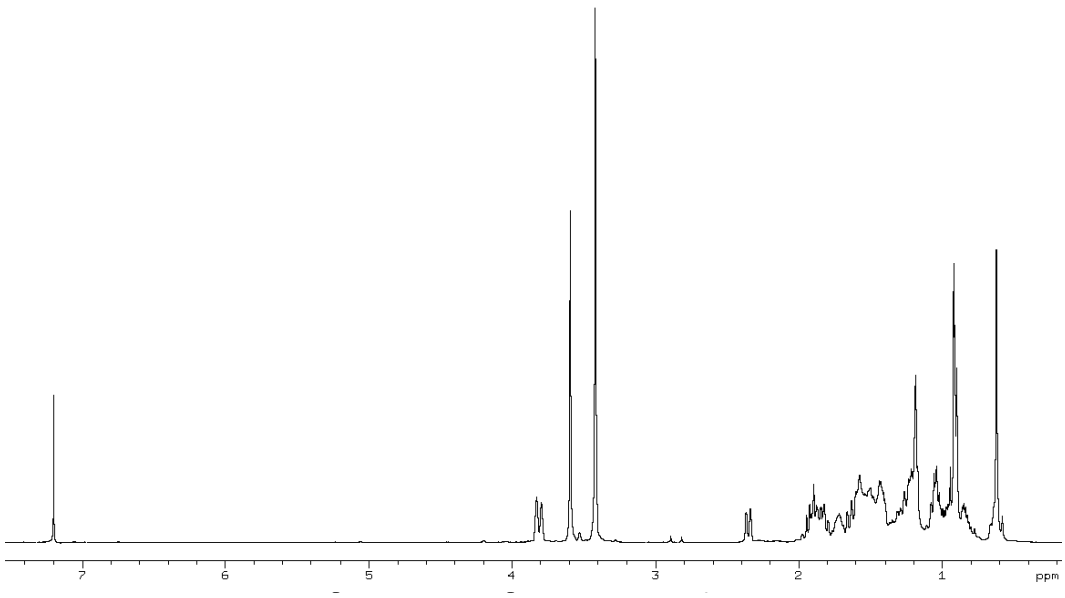
^{1}H NMR spectrum of Compound **2** in CD_3OD at 700 MHz.



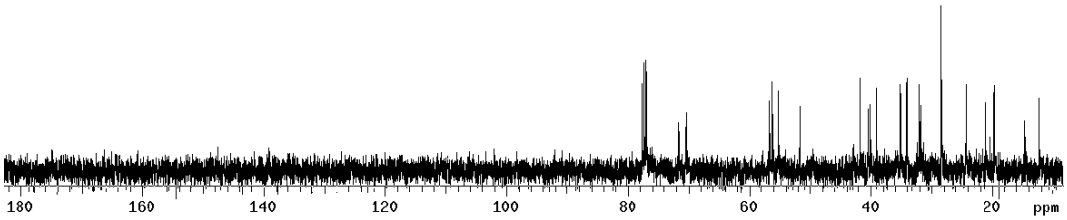
^{13}C NMR spectrum of Compound **2** in CD_3OD at 100 MHz.



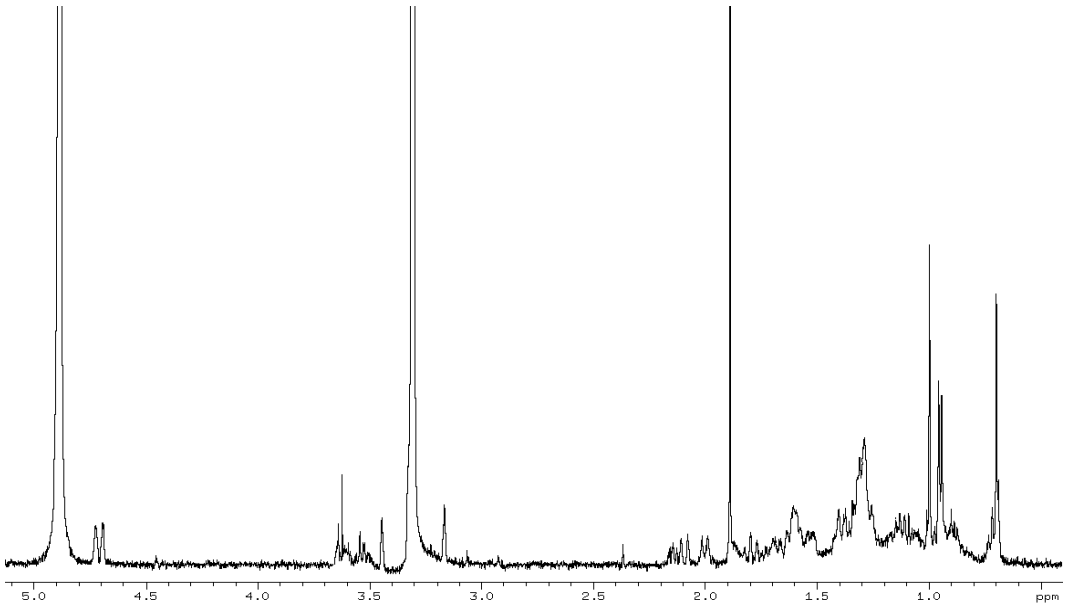
^{1}H NMR spectrum of Compound **35** in $CDCl_3$ at 400 MHz.



^{13}C NMR spectrum of Compound **35** in $CDCl_3$ at 100 MHz.



^{1}H NMR spectrum of Compound **37** in CD_3OD at 400 MHz.



III. Experimental section of dual PXR/FXR ligands

Sponge material and separation of individual sterols (Compounds 40-49). Two taxonomic vouchers (field collection references R3170 and R3144) both assigned to the widespread species *Theonella swinhoei* (order Lithistida, family Theonellidae) were collected on the barrier reef of Vangunu and Malaita Island, Solomon Islands, in July 2004. The samples were frozen immediately after collection and lyophilized to yield 600 g and 355 g (dry mass) of R3170 and R3144, respectively. Taxonomic identification was performed by Dr John Hooper of Queensland Museum, Brisbane, Australia, where specimens are deposited under the accession number G3122662 and G323969, respectively.

Theonella swinhoei (R3170). The lyophilized material (600 g) was extracted with methanol (3×2.7 L) at room temperature and the methanolic extract, taken to dryness, was subjected to a modified Kupchan's partitioning procedure as follows. The methanol extract was dissolved in a mixture of MeOH/H_2O containing 10% H_2O and partitioned against *n*-hexane (19.7 g). The water content (% v/v) of the MeOH extract was adjusted to 30% and partitioned against $CHCl_3$ (17.8 g). The aqueous phase was concentrated to remove MeOH and then extracted with *n*-BuOH (10 g). The hexane extract was chromatographed in two runs by silica gel MPLC using a solvent gradient system from CH_2Cl_2 to $CHCl_2$:MeOH 1:1. Fractions eluted with CH_2Cl_2:MeOH 96:4 (419 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 1.6 mg of theonellasterol H (**46**) (t_R=9.5 min), 12 mg of theonellasterol G (**45**) (t_R=11.6), 3.5 mg of theonellasterol F (**44**) (t_R=12.4 min), 0.8 mg of conicasterol D (**49**) (t_R=12.8 min), 3.3 mg of theonellasterol D (**42**) (t_R=14.2 min) and 3.6 mg of theonellasterol E (**43**) (t_R=16 min). Fractions eluted with CH_2Cl_2:MeOH 95:5 (40 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 0.6 mg of conicasterol B (**47**) (t_R=7.8 min) and 0.8 mg of theonellasterol B (**40**) (t_R= 8.4 min).

Theonella swinhoei (R3144) was extracted with methanol (3×1.7 L) at room temperature and the crude methanolic extract (32 g) was subjected to a modified Kupchan's partitioning procedure as described for *Theonella swinhoei* R3170, to obtain 3.9 g of *n*-hexane extract. The hexane extract was chromatographed in two runs by silica gel MPLC using a solvent gradient system from CH_2Cl_2 to $CHCl_2$:MeOH 1:1. Fractions eluted with CH_2Cl_2:MeOH 97:3 (15.5 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 3.4 mg of theonellasterol D (**42)** and 1.1 mg of theonellasterol C (**41**) (t_R=20.4 min). Fractions eluted with CH_2Cl_2:MeOH 94:6 (12.0 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 1.2 mg of theonellasterol D (**42**), 4.0 mg of theonellasterol E (**43**), 0.5 mg conicasterol C (**48**) (t_R=15 min) and 0.8 mg of conicasterol D (**49**).

Sponge material and separation of individual sterols (Compounds 39, 50-59). *Theonella swinhoei* (order Lithistida, family Theonellidae) was collected at a depth of 22 m, on an isolated

reef off the western coast of Malaita Island, Solomon Islands, in July 2004 and reference specimens are on file (R3159) at the ORSTOM, Centre of Noumea. The samples were frozen immediately after collection and lyophilized to yield 207 g of dry mass. Taxonomic identification was performed by Dr John Hooper at Queensland Museum, Brisbane, Australia. The lyophilized material (207 g) was extracted with methanol (3×1.5 L) at room temperature and the crude methanolic extract was subjected to a modified Kupchan's partitioning procedure as follows. The methanol extract was dissolved in a mixture of MeOH/H_2O containing 10% H_2O and partitioned against *n*-hexane (4.5 g). The water content (% v/v) of the MeOH extract was adjusted to 30% and partitioned against $CHCl_3$ (6.0 g). The aqueous phase was concentrated to remove MeOH and then extracted with *n*-BuOH (10.3 g). The hexane extract (4.5 g) was chromatographed by silica gel MPLC using a solvent gradient system from CH_2Cl_2 to $CHCl_2$:MeOH 1:1. Fractions eluted with CH_2Cl_2:MeOH 99:1 (340.4 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (95:5) as eluent (flow rate 5 mL/min) to give 1.5 mg of conicasterol (**39**) (t_R=4.2 min) and 2.0 mg of conicasterol H (**51**) (t_R=15.8 min). Fractions eluted with CH_2Cl_2:MeOH 97:3 (242.5 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 5.8 mg of swinhosterol B (**59**) (t_R=7.5 min), 1.2 mg of conicasterol I (**52**) (t_R=8.5 min), 2.5 mg of conicasterol G (**50**) (t_R=11.0 min), 1.2 mg of 7α-hydroxyconicasterol (**57**) (t_R=17 min) and 1.0 mg of 15β-hydroxyconicasterol (**58**) (t_R=17.2 min). Fractions eluted with CH_2Cl_2:MeOH 99:1 (351 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (998:2) as eluent (flow rate 3 mL/min) to give 9.3 mg of conicasterol K (**54**) (t_R=7.8 min), 8.9 mg of theonellasterol J (**55**) (t_R=8.2 min), 7.5 mg of dehydroconicasterol (**56**) (t_R=12.0 min). The $CHCl_3$ extract (6.0 g) was chromatographed by silica gel MPLC using a solvent gradient system from CH_2Cl_2 to $CHCl_2$:MeOH 3:7. Fractions eluted with CH_2Cl_2:MeOH 96:4 (80 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 3 mL/min) to give 1.3 mg of conicasterol I (**53**) (t_R=8.0 min).

Characteristic data for each compounds

Theonellasterol B (**40**): pale yellow oil; $[\alpha]_D^{25}$ +1.6 (*c* 0.08, MeOH); UV (MeOH): λ_{max} (log ε) 275 nm (3.62); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 429.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{30}H_{46}LiO$: 429.3709; found 429.3729 $[M+Li]^+$.

Theonellasterol C (**41**): pale yellow oil; $[\alpha]_D^{25}$ +5.7 (*c* 0.16, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 447.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{30}H_{48}LiO_2$: 447.3814; found 447.3834 $[M+Li]^+$.

Theonellasterol D (**42**): white amorphous solid; $[\alpha]_D^{25}$ +3.3 (*c* 0.17, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 479.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{31}H_{52}LiO_3$: 479.4076; found 479.4093 $[M+Li]^+$.

Theonellasterol E (**43**): white amorphous solid; $[\alpha]_D^{25}$ +34 (*c* 0.1, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 465.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{30}H_{50}LiO_3$: 465.3920; found 465.3244 $[M+Li]^+$.

Theonellasterol F (**44**): white amorphous solid; $[\alpha]_D^{25}$ +4.3 (*c* 0.23, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 465.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{30}H_{50}LiO_3$: 465.3920; found 465.3247 $[M+Li]^+$.

Theonellasterol G (**45**): pale yellow oil; $[\alpha]_D^{25}$ +68 (*c* 1.0, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 481.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{30}H_{50}LiO_4$: 481.3869; found 481.3889 $[M+Li]^+$.

Theonellasterol H (**46**): pale yellow oil; $[\alpha]_D^{25}$ +16 (*c* 0.1, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 481.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{30}H_{50}LiO_4$: 481.3869; found 481.3875 $[M+Li]^+$.

Conicasterol B (**47**): pale yellow oil; $[\alpha]_D^{25}$ +29 (*c* 0.06, MeOH); UV (MeOH): λ_{max} (log ε) 275 nm (3.56);^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 415.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{44}LiO$: 415.3552; found 415.3532 $[M+Li]^+$.

Conicasterol C (**48**): white amorphous solid ; $[\alpha]_D^{25}$ +43 (*c* 0.13, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 465.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{30}H_{50}LiO_3$: 465.3920; found 465.3942 $[M+Li]^+$.

Conicasterol D (**49**): white amorphous solid; $[\alpha]_D^{25}$ +33 (*c* 0.08, MeOH); ^1H and ^{13}C NMR data in C_6D_6 given in **Table 1-3**; ESI-MS: *m/z* 451.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{48}LiO_3$: 451.3763; found 451.3739 $[M+Li]^+$.

Conicasterol G (**50**): white amorphous solid; $[\alpha]_D^{25}$ +17.4 (*c* 0.2, MeOH); selected ^1H and ^{13}C NMR data in C_6D_6 given in **Tables 4** and **5**; ESI-MS: *m/z* 449.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{46}LiO_3$: 449.3607; found 449.3611 $[M+Li]^+$.

Conicasterol H (**51**): white amorphous solid; $[\alpha]_D^{25}$ +70.0 (*c* 0.2, MeOH); selected ^1H and ^{13}C NMR data in C_6D_6 given in **Tables 4** and **5**; ESI-MS: *m/z* 451.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{48}LiO_3$: 451.3763; found 451.3770 $[M+Li]^+$.

Conicasterol I (**52**): pale yellow oil; $[\alpha]_D^{25}$ +58.2 (*c* 0.12, MeOH); selected ^1H and ^{13}C NMR data in C_6D_6 given in **Tables 4** and **5**; ESI-MS: *m/z* 467.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{48}LiO_4$: 467.3713; found 467.3724 $[M+Li]^+$.

Conicasterol J (53): pale yellow oil; $[\alpha]_D^{25}$ +37.2 (*c* 0.13, MeOH); selected ^1H and ^{13}C NMR data in C_6D_6 given in **Tables 4** and **5**; ESI-MS: *m/z* 449.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{46}LiO_3$: 449.3607; found 449.3620 $[M+Li]^+$.

Conicasterol K (**54**): white amorphous solid; $[\alpha]_D^{25}$ +5.9 (*c* 0.93, MeOH); selected ^1H and ^{13}C NMR data in C_6D_6 given in **Tables 4** and **5**; ESI-MS: *m/z* 417.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{46}LiO$: 417.3709; found 417.3723 $[M+Li]^+$.

Theonellasterol J (**55**): pale yellow oil; $[\alpha]_D^{25}$ +6.9 (*c* 0.37, MeOH); selected ^1H and ^{13}C NMR data in C_6D_6 given in **Tables 4** and **5**; ESI-MS: *m/z* 431.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{48}LiO$: 431.3865; found 431.3875 $[M+Li]^+$.

Conicasterol (**39**): white amorphous solid; $[\alpha]_D^{25}$ +52.5 (*c* 0.15, MeOH); ESI-MS: *m/z* 419.4 $[M+Li]^+$.

Dehydroconicasterol (**56**)**:** white amorphous solid; $[\alpha]_D^{25}$ +31.7 (*c* 0.31, MeOH); ESI-MS: *m/z* 417.4 $[M+Li]^+$.

7α-Hydroxyconicasterol (**57**): yellow amorphous solid; $[\alpha]_D^{25}$ +3.5 (*c* 0.12, MeOH); ESI-MS: *m/z* 435.4 $[M+Li]^+$.

15β-Hydroxyconicasterol (**58**): white amorphous solid; $[\alpha]_D^{25}$ +5.1 (*c* 0.1, MeOH); ESI-MS: *m/z* 435.4 $[M+Li]^+$.

Swinhosterol B (**59**): white amorphous solid; $[\alpha]_D^{25}$ -14.6 (*c* 0.58, MeOH); ESI-MS: *m/z* 451.4 $[M+Li]^+$.

Specific biological assays.

RT^2 Profiler PCR array. Total RNA from HepG2 cells left untreated or stimulated with swinhosterol B (**59**) was extracted with Trizol reagent (Invitrogen) and reverse transcribed with Superscript-II reverse transcriptase (Invitrogen) following the manual instructions. 100 ng cDNA was pipetted in each well of a 96 well gene array plate (Human Nuclear Receptors and Coregulators RT² Profiler™ PCR Array- http://www.sabiosciences.com/rt_pcr_product/HTML/PAHS-056A.html -Superarray Bioscence, Frederick, MD, USA) and amplified following the manual instructions. Genes selected for this array encode several classes of nuclear receptors and co-regulators of transcription, including co-activators and co-repressors. Array analysis was carried out with the on-line software RT2 Profiler PCR Array Data Analysis (http://pcrdataanalysis.sabiosciences.com/pcr/arrayanalysis.php).

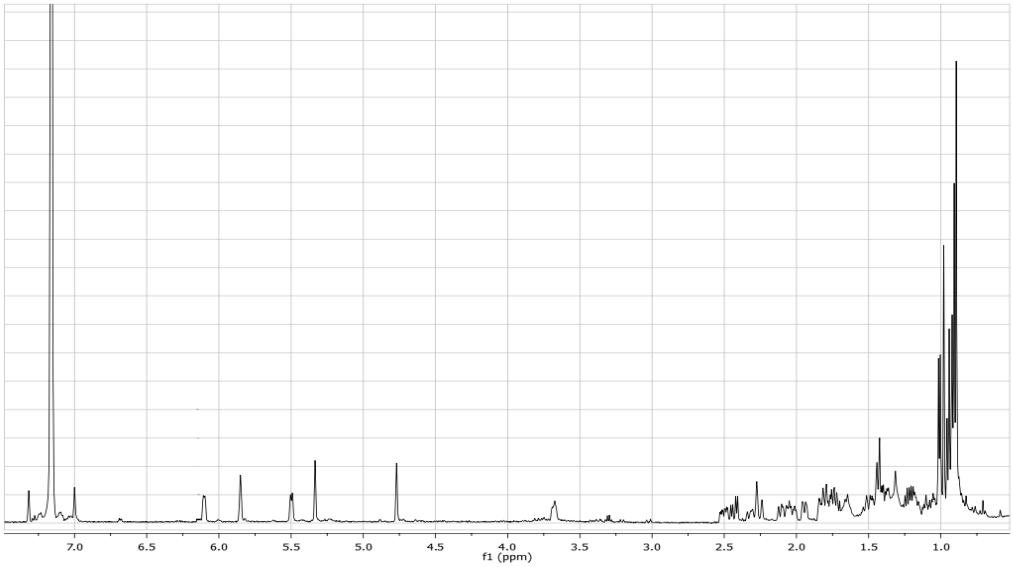
Macrophages preparation and in vitro testing. Animals. Humanized (h)PXR mice, 8-10 weeks of age, were provided by Frank J. Gonzalez (Laboratory of Metabolism, Centre for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland). The hPXR mice express the human PXR in the mice PXR-null background.[17] FXR-/-mice (C57BL/6BJ6 background) were originally provided by F. Gonzales (NIH, Bethesda). All mice were housed in a temperature-controlled room and had free access to food and water.[136] The present protocol was approved by the Italian Minister Health and conforms to national guidelines. The ID for this project is #11/2010-B and authorization released to Prof. Stefano Fiorucci, as a principal investigator, lastly on January 25, 2010.

Isolation of mouse macrophages. Monocytes were isolated from spleen of hPXR and FXR-/-mice (N=4); monocytes were obtained by means of a magnetic cell separation isolation system as previously described using mouse CD11b Microbeads (Milteny Biotech; Germany). CD11b+ cells were then suspended in complete medium (RPMI 1640, 10% heat-inactivated fetal bovine serum, 3 mmol/L L-glutamine, 100 U/mL penicillin, 100 U/mL streptomycin) and cultured at a concentration of 1 × 106 cells/mL. The isolated cells were placed into culture wells (Tissue Culture Plate, 24 wells) with or without LPS (5 μg/mL) alone or in combination with **59** (10 μM; 2 h of pre-treatment) and with **59** (10 μM) alone, and maintained in an atmosphere of 5% CO_2 and 95% air at 37 °C for 18 h.

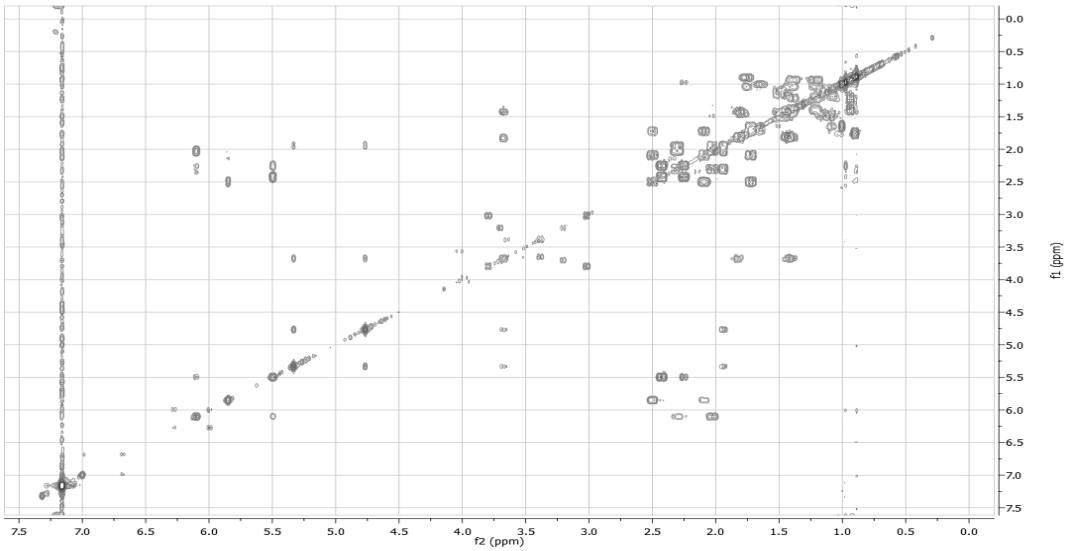
Statistical Analysis. All values were expressed as the mean ± SE. The analysis of variance (ANOVA) with Bonferroni correction for multiple comparisons (Graphpad Software. Inc., San Diego, CA) was used to assess significant statistical difference between groups.

Spectroscopic Data

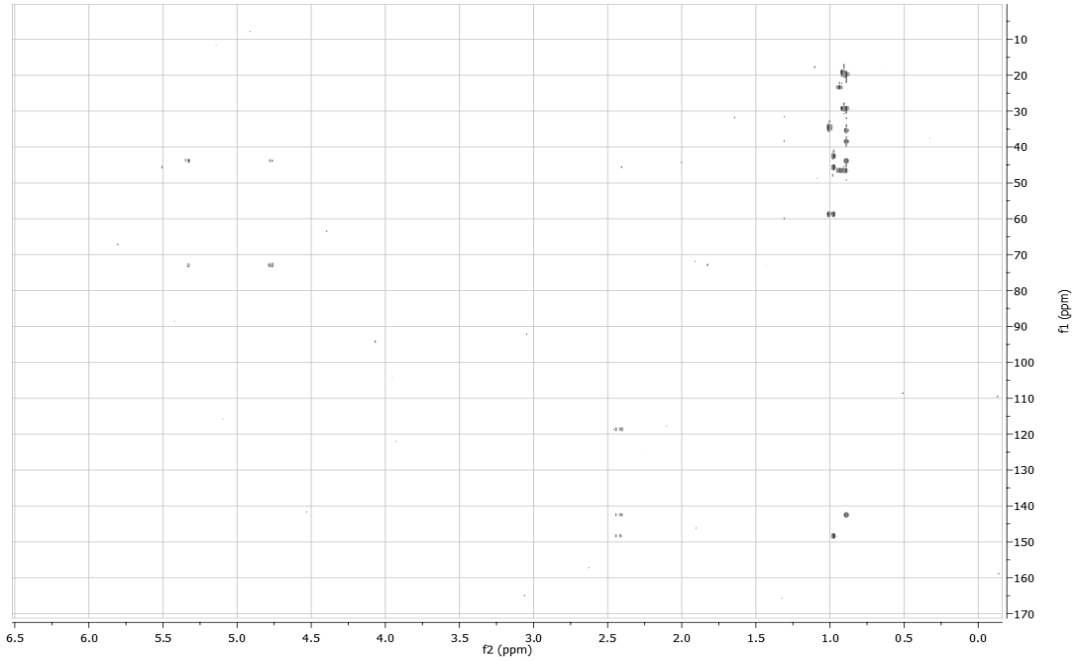
^{1}H NMR (500 MHz, C_6D_6) of Theonellasterol B (**40**)



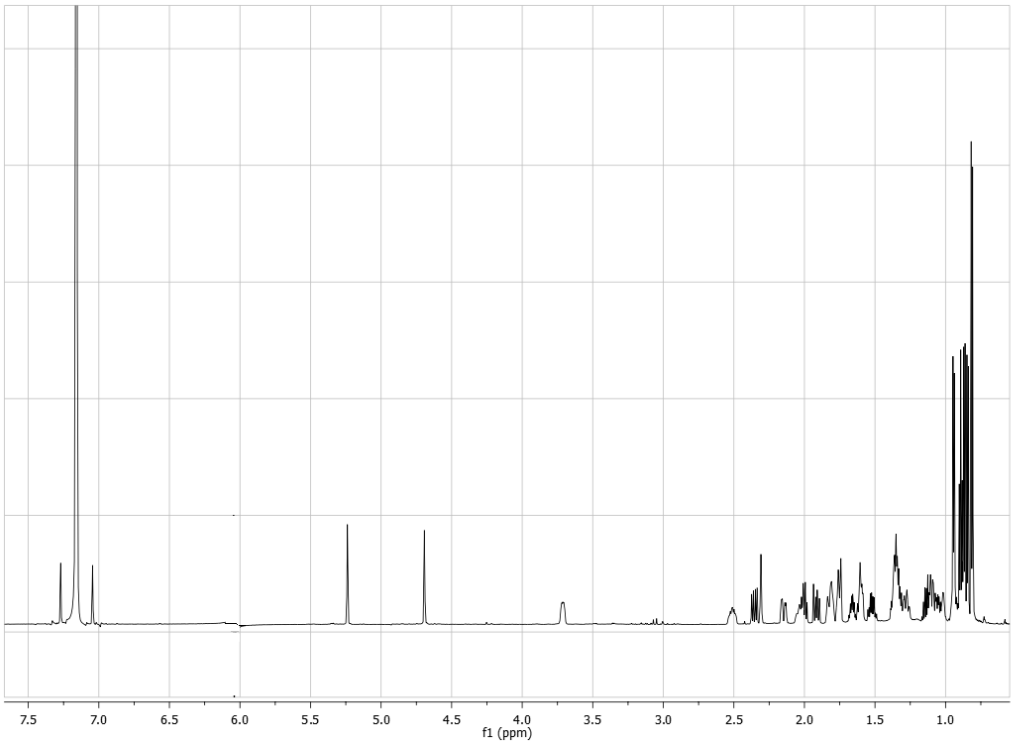
COSY spectrum (500 MHz, C_6D_6) of Theonellasterol B (**40**)



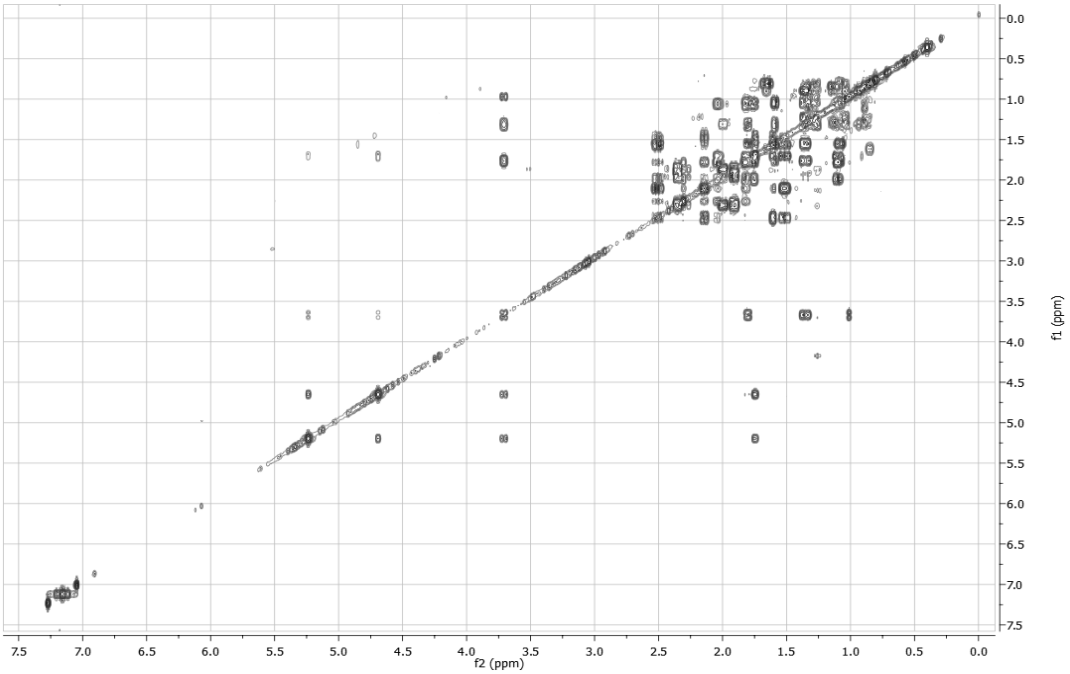
HMBC spectrum (500 MHz, C_6D_6) of Theonellasterol B (**40**)



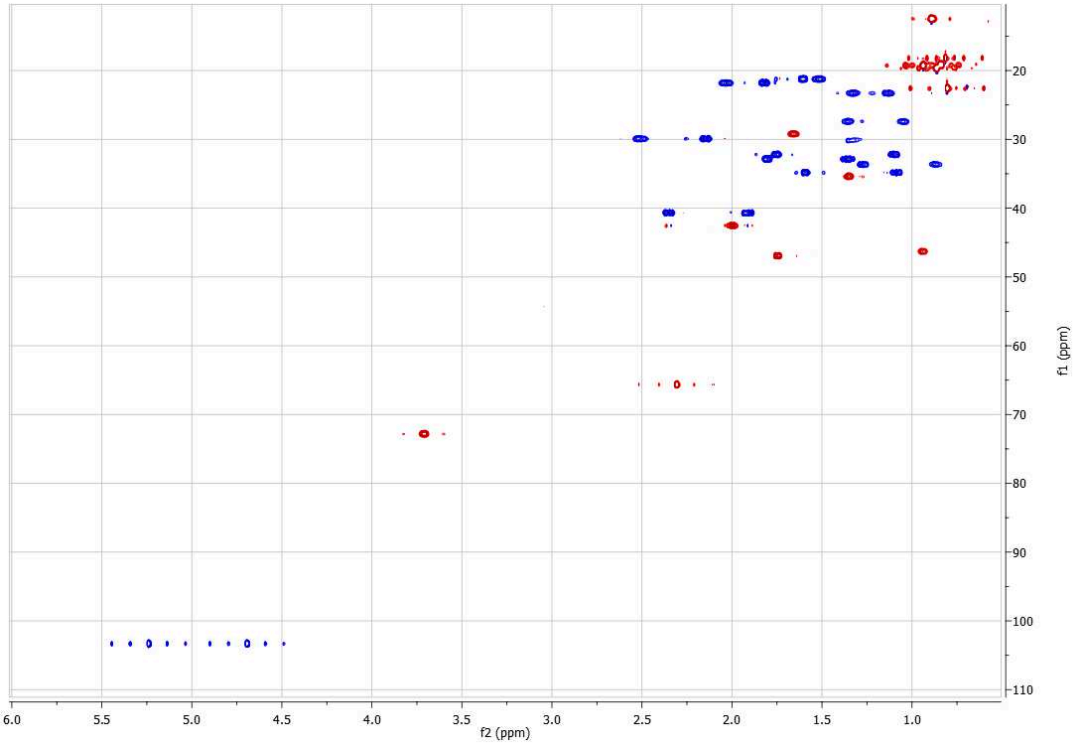
^{1}H NMR (700 MHz, C_6D_6) of Theonellasterol C (**41**)



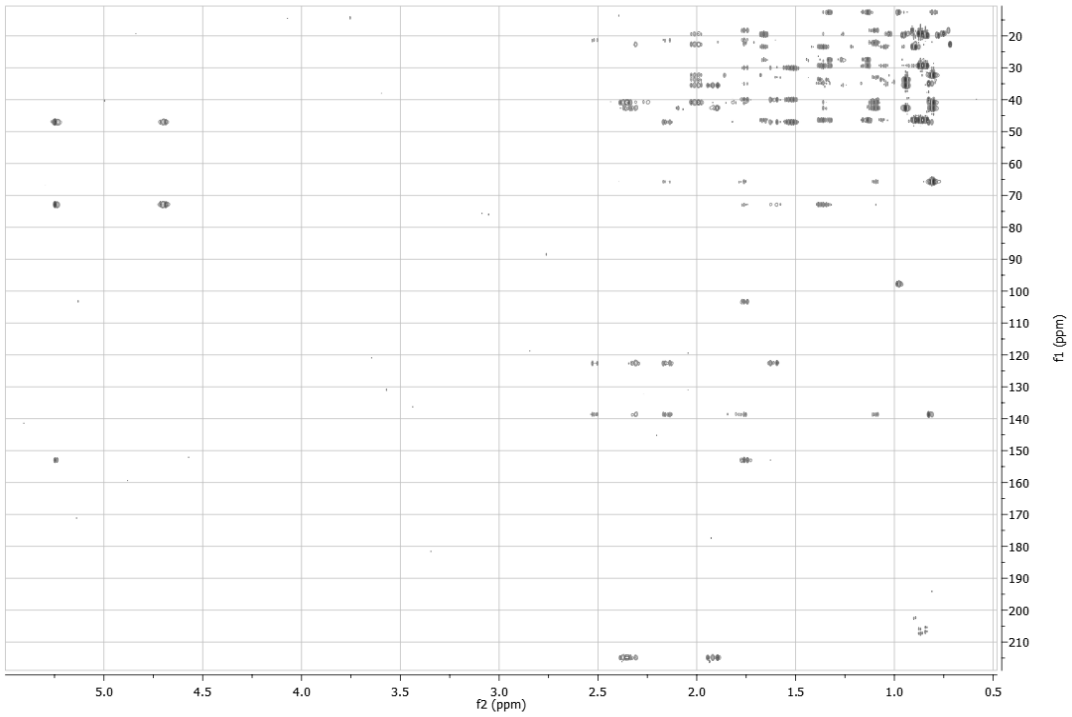
COSY spectrum (700 MHz, C_6D_6) of Theonellasterol C (**41**)



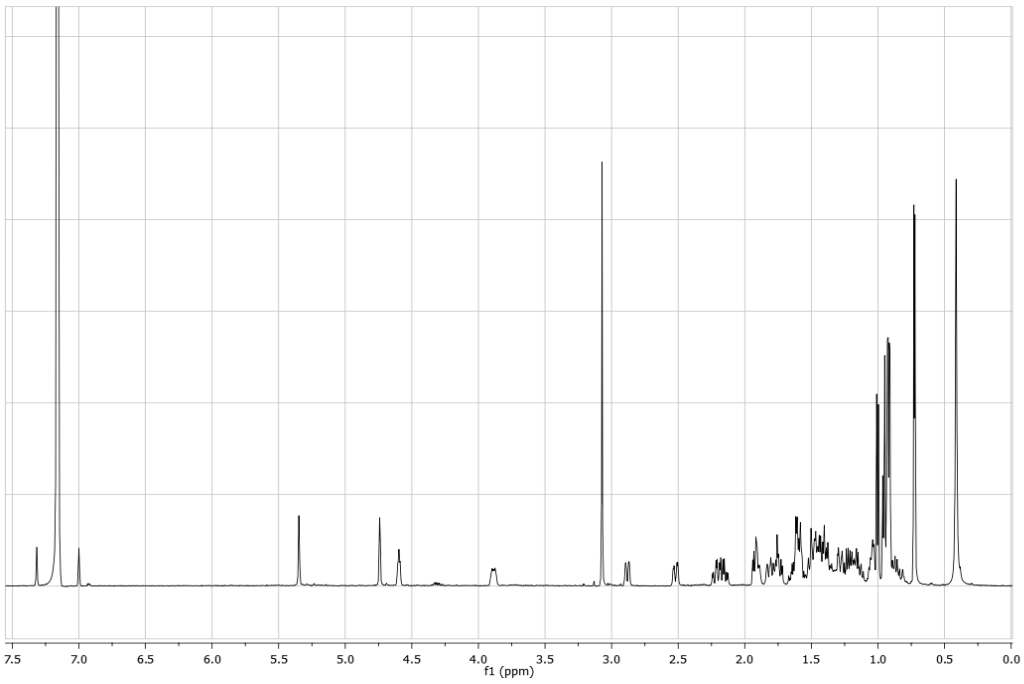
HSQC spectrum (700 MHz, C_6D_6) of Theonellasterol C (**41**)



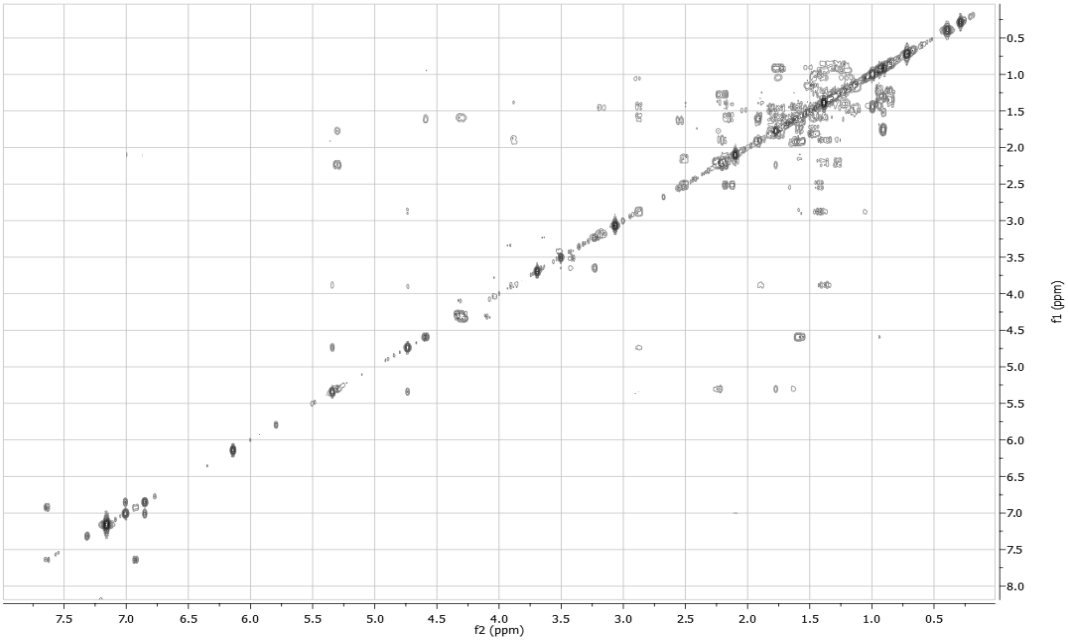
HMBC spectrum (700 MHz, C_6D_6) of Theonellasterol C (**41**)



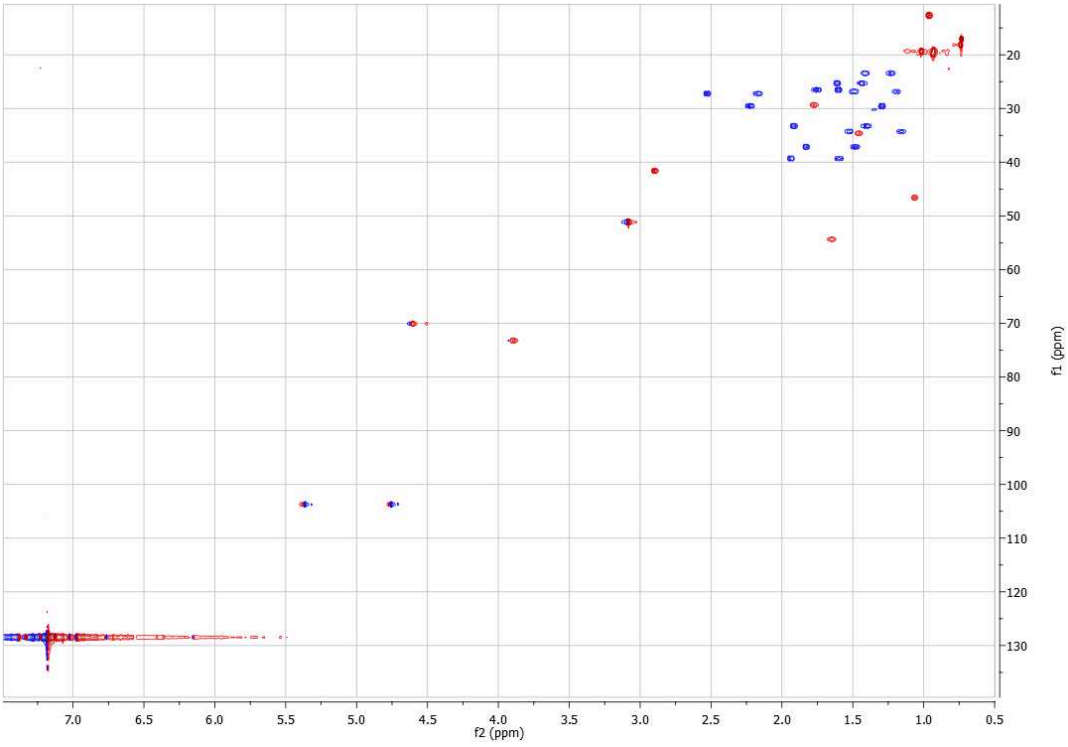
^{1}H NMR (500 MHz, C_6D_6) of Theonellasterol D (**42**)



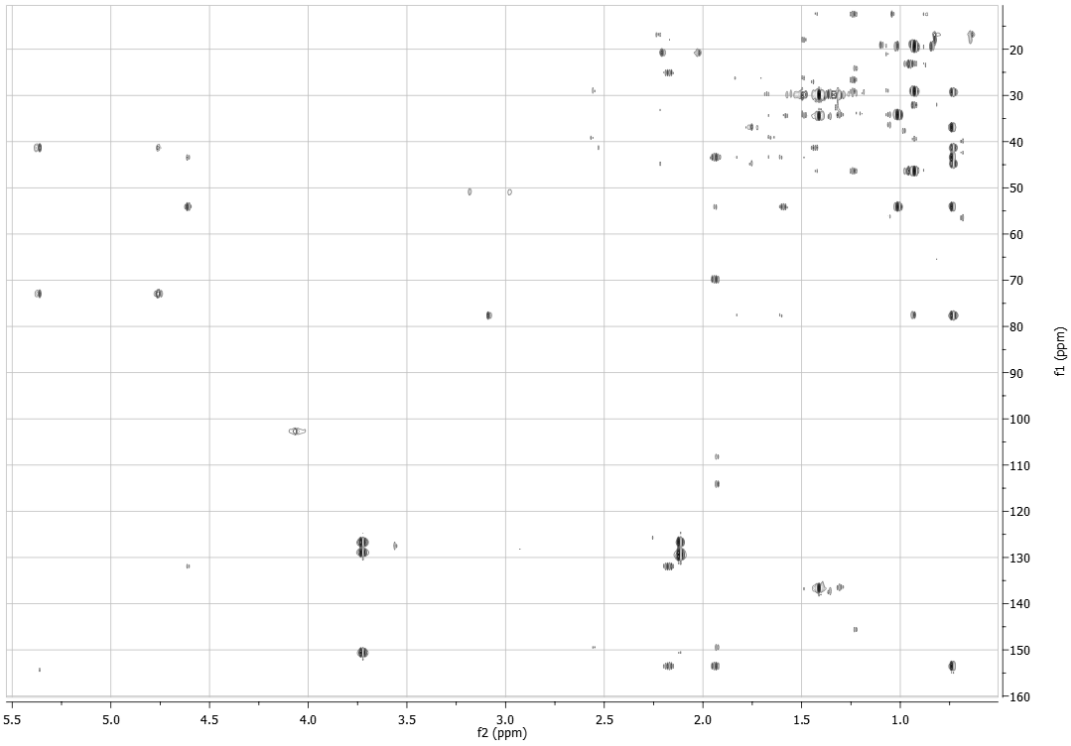
COSY spectrum (500 MHz, C_6D_6) of Theonellasterol D (**42**)



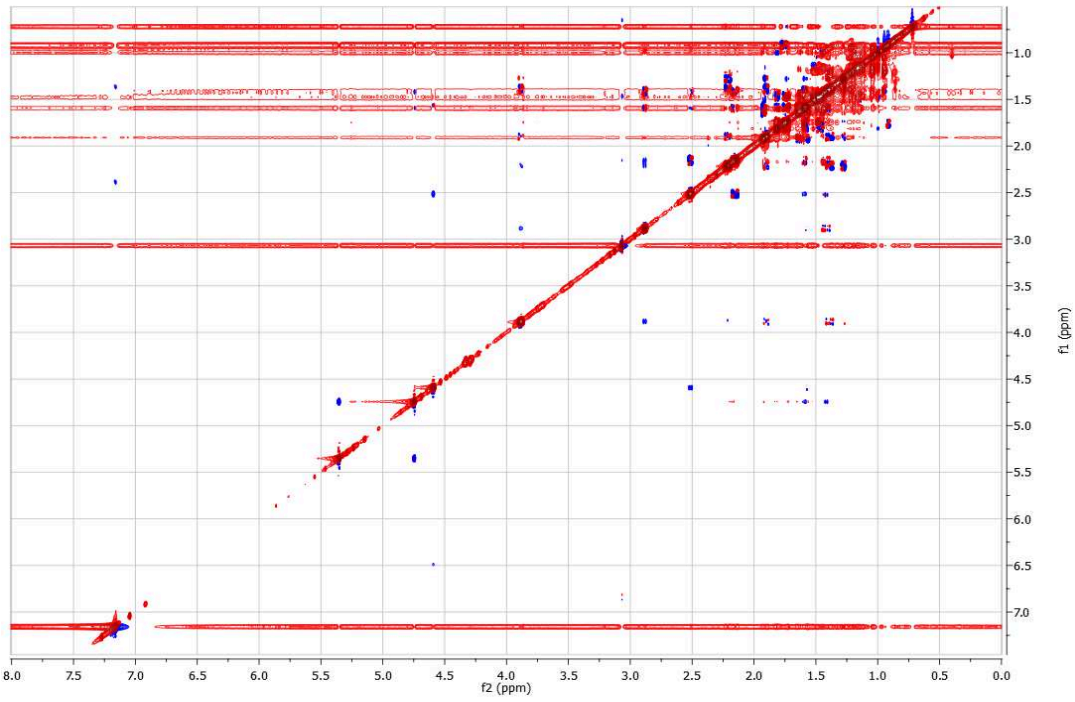
HSQC spectrum (700 MHz, C_6D_6) of Theonellasterol D (**42**)



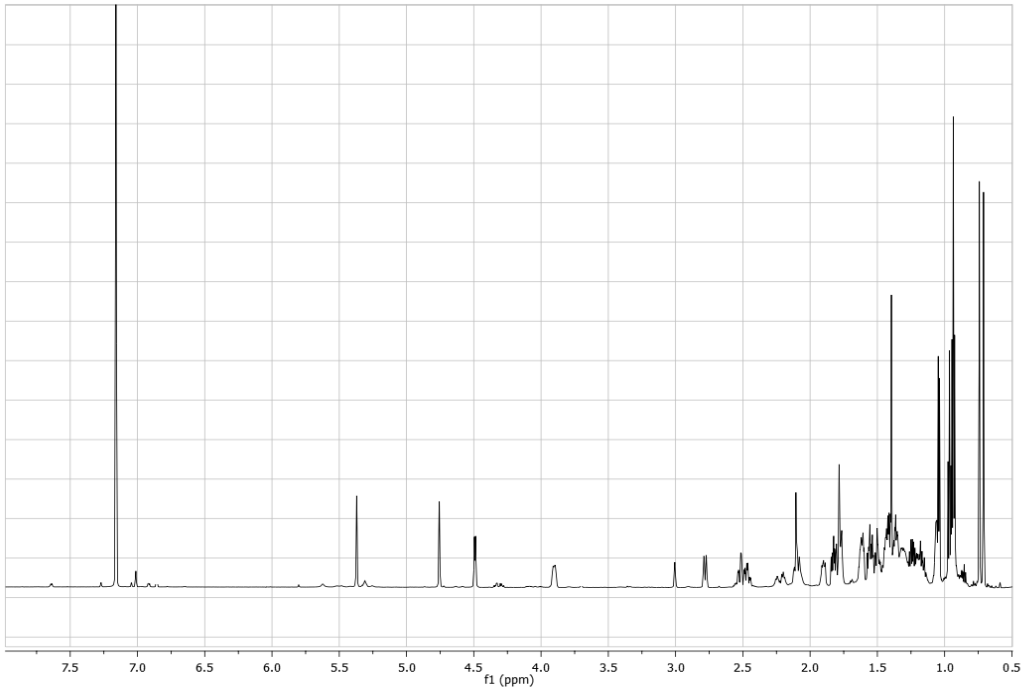
HMBC spectrum (700 MHz, C_6D_6) of Theonellasterol D (**42**)



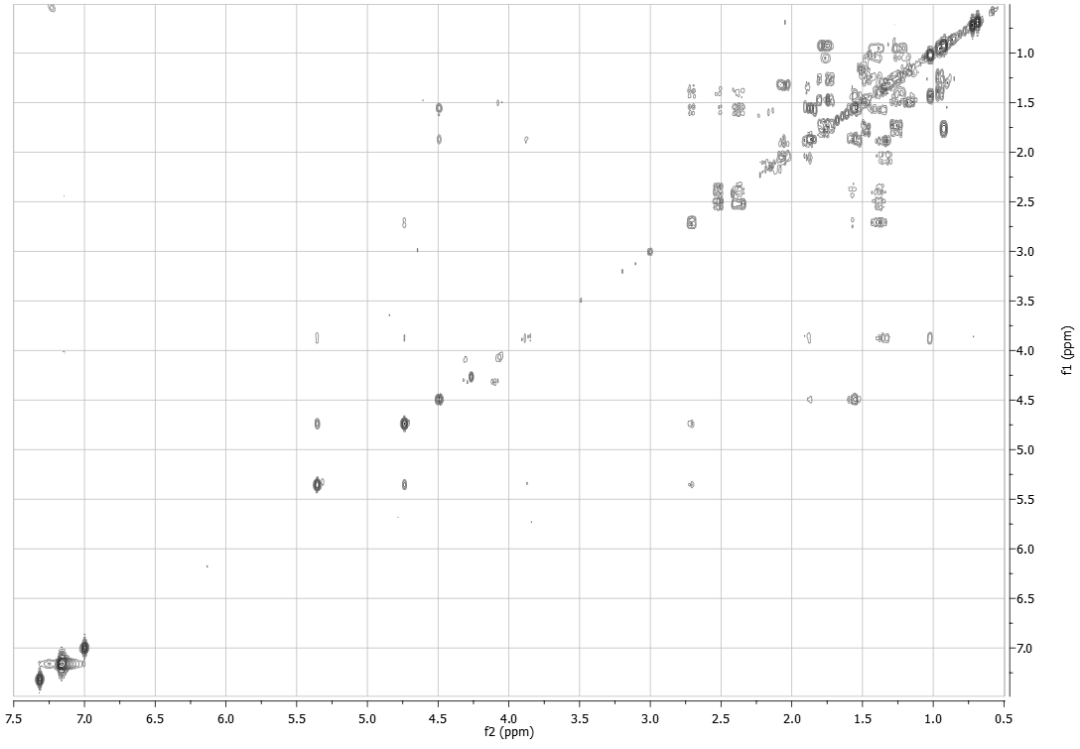
ROESY spectrum (700 MHz, C_6D_6, 200 ms) of Theonellasterol D (**42**)



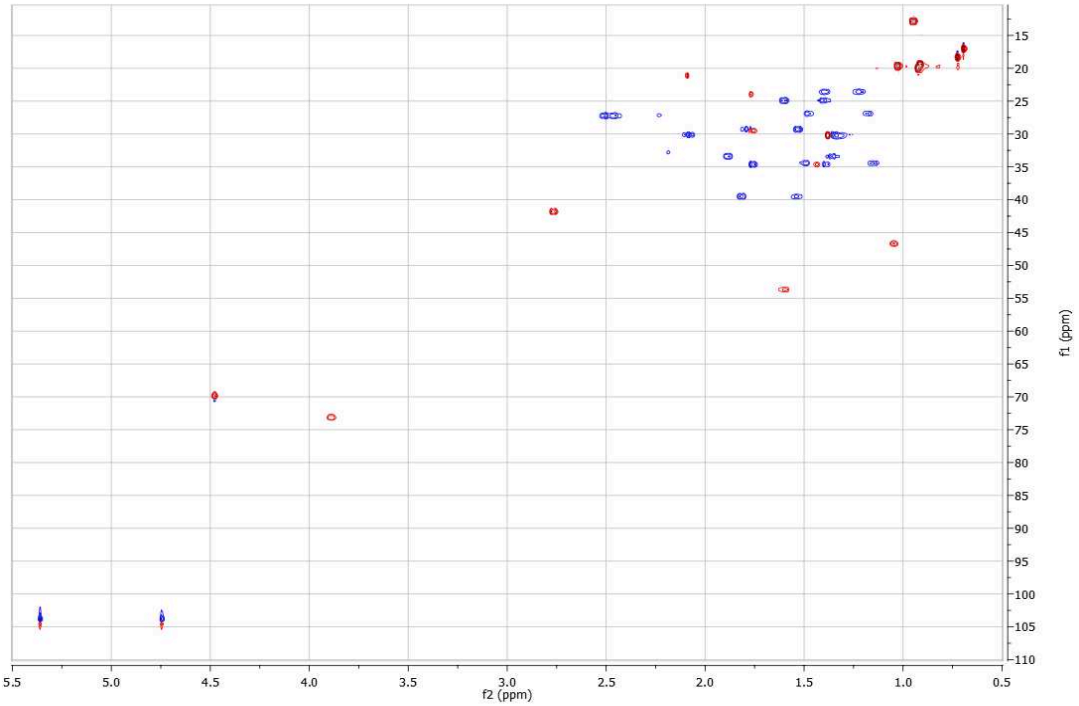
^{1}H NMR (700 MHz, C_6D_6) of Theonellasterol E (**43**)



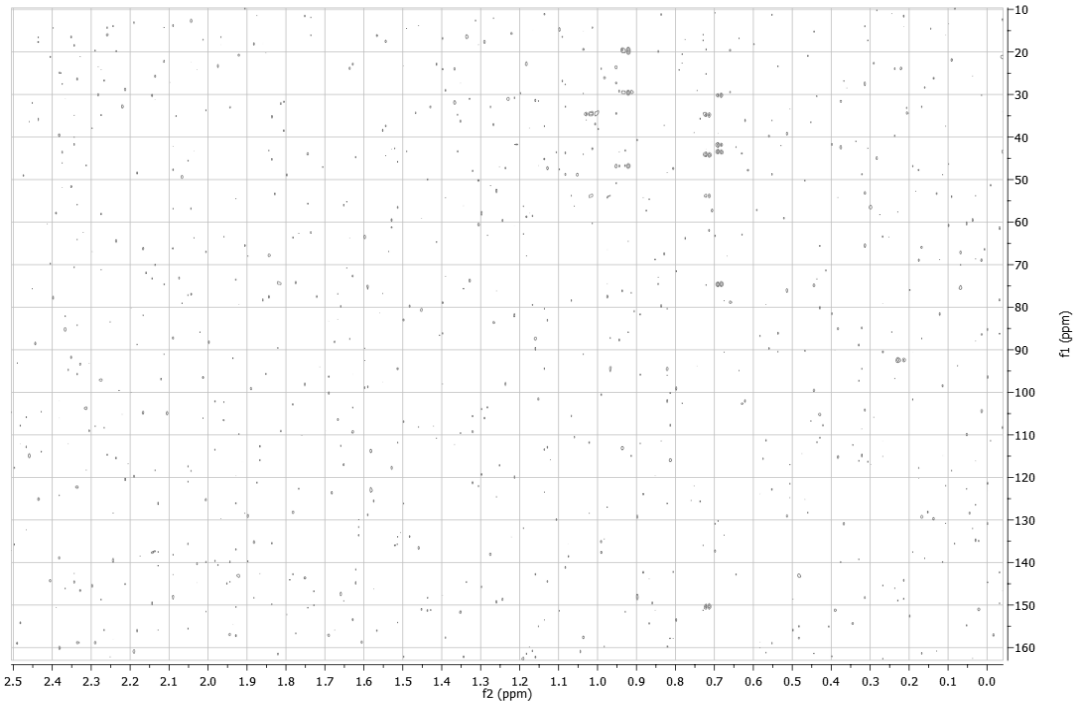
COSY spectrum (500 MHz, C_6D_6) of Theonellasterol E (**43**)



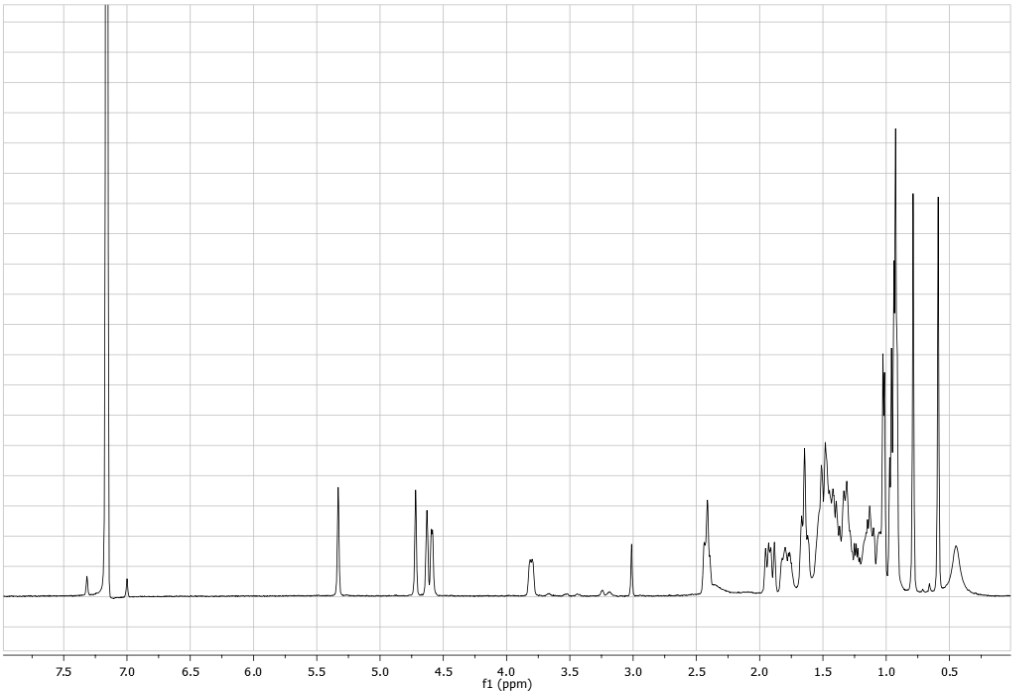
HSQC spectrum (700 MHz, C_6D_6) of Theonellasterol E (**43**)



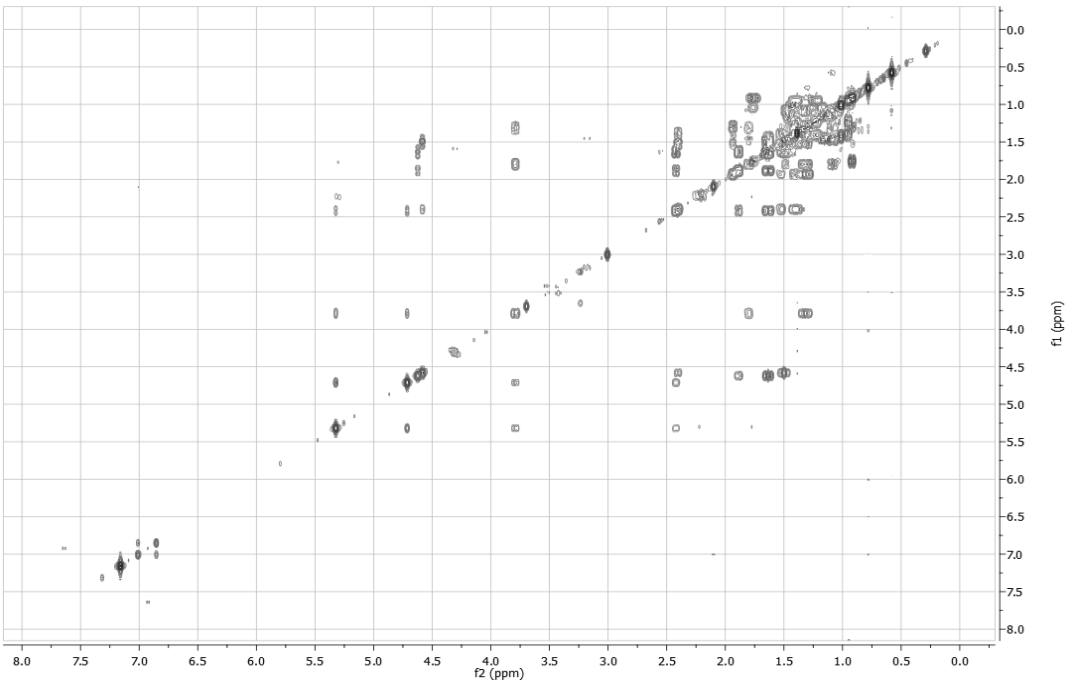
Expanded HMBC spectrum (500 MHz, C_6D_6) of Theonellasterol E (**43**)



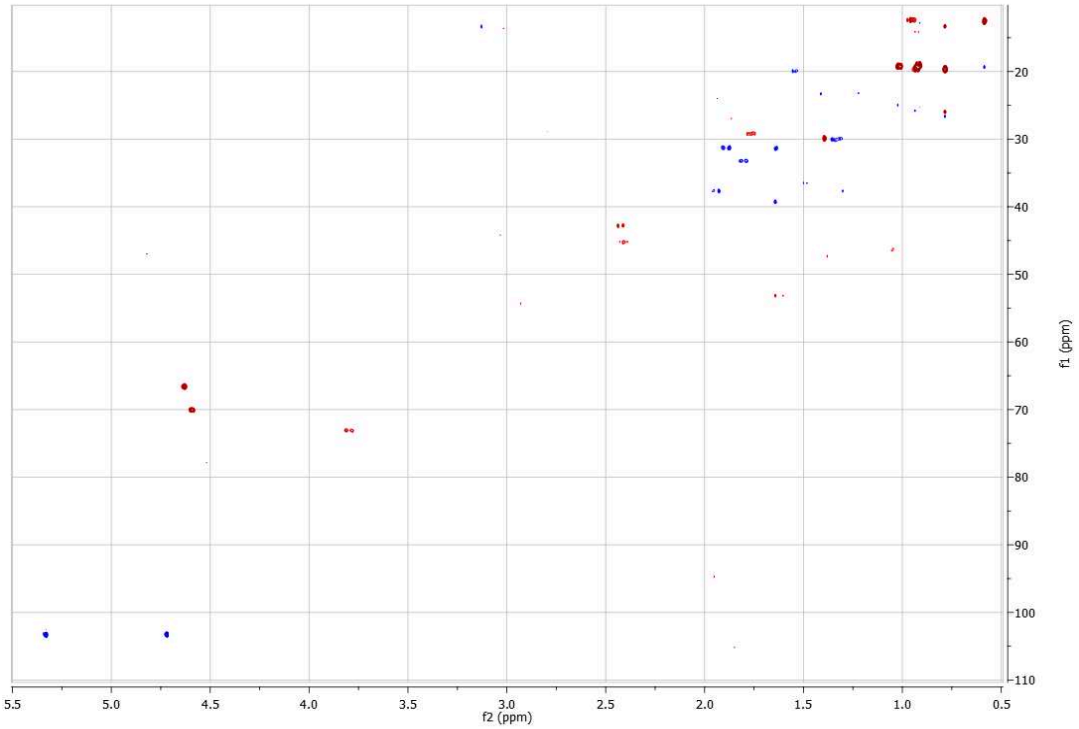
^{1}H NMR (500 MHz, C_6D_6) of Theonellasterol F (**44**)



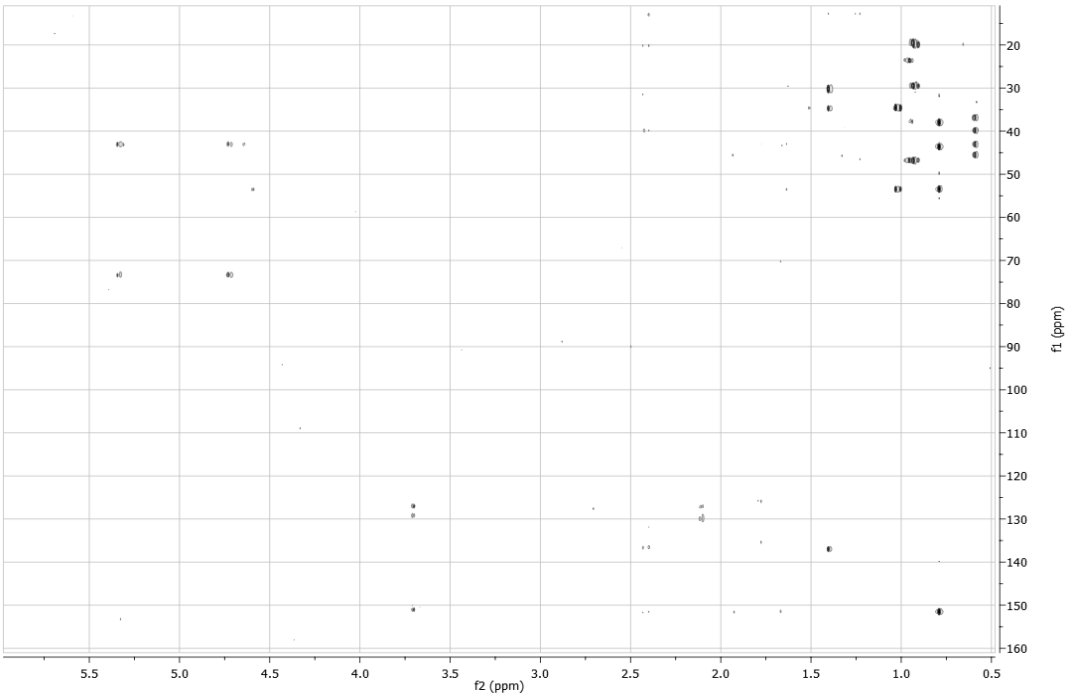
COSY spectrum (500 MHz, C_6D_6) of Theonellasterol F (**44**)



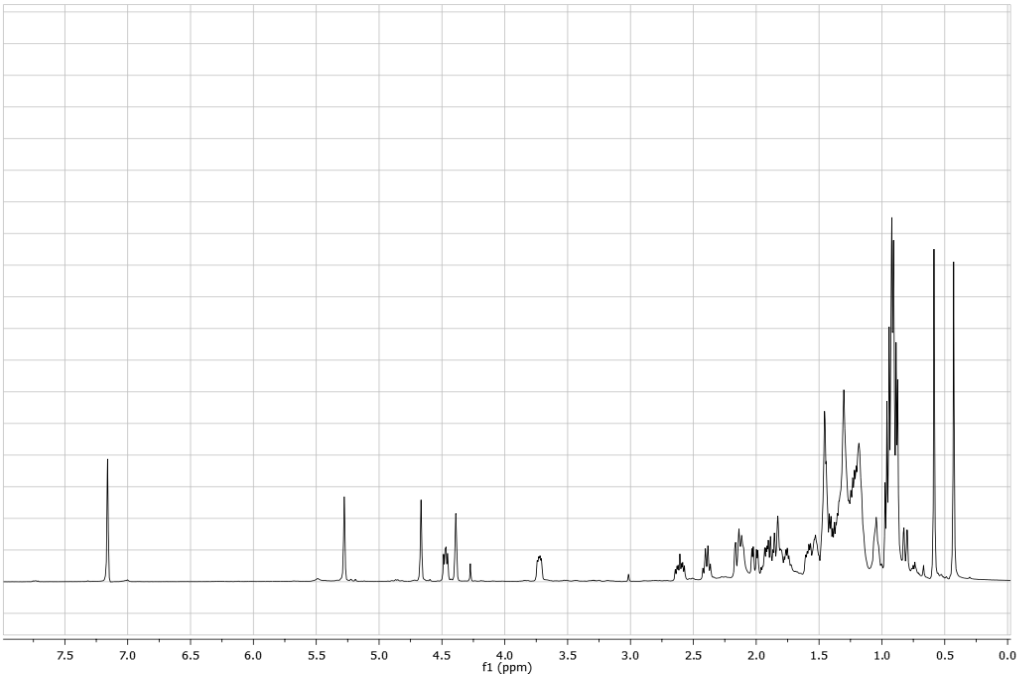
HSQC spectrum (500 MHz, C_6D_6) of Theonellasterol F (**44**)



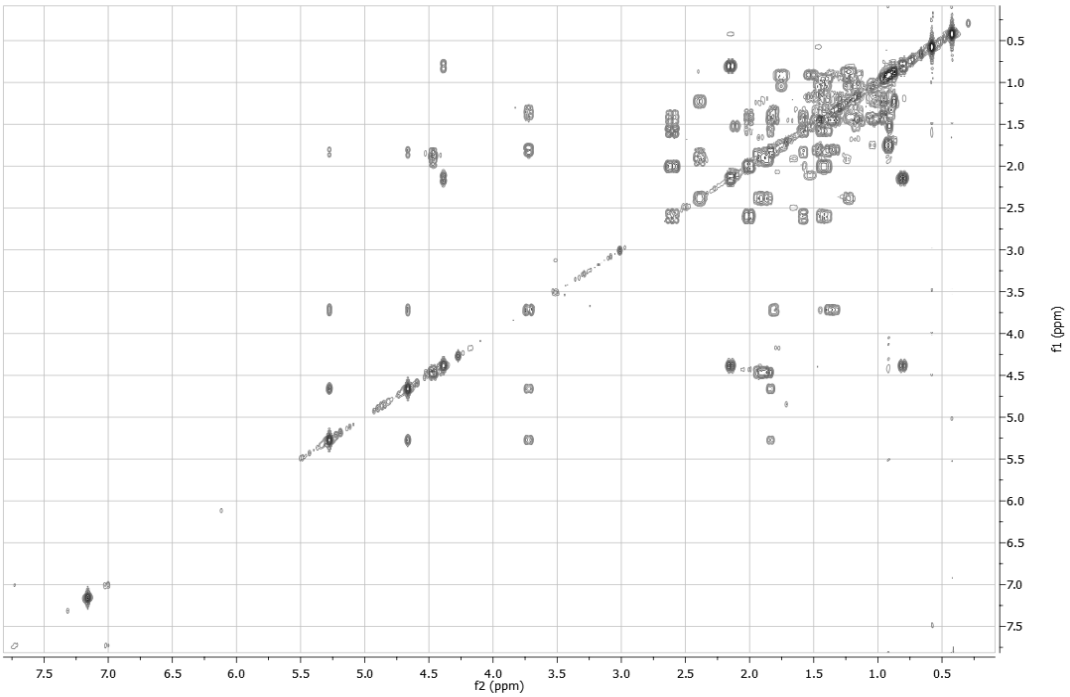
HMBC spectrum (500 MHz, C_6D_6) of Theonellasterol F (**44**)



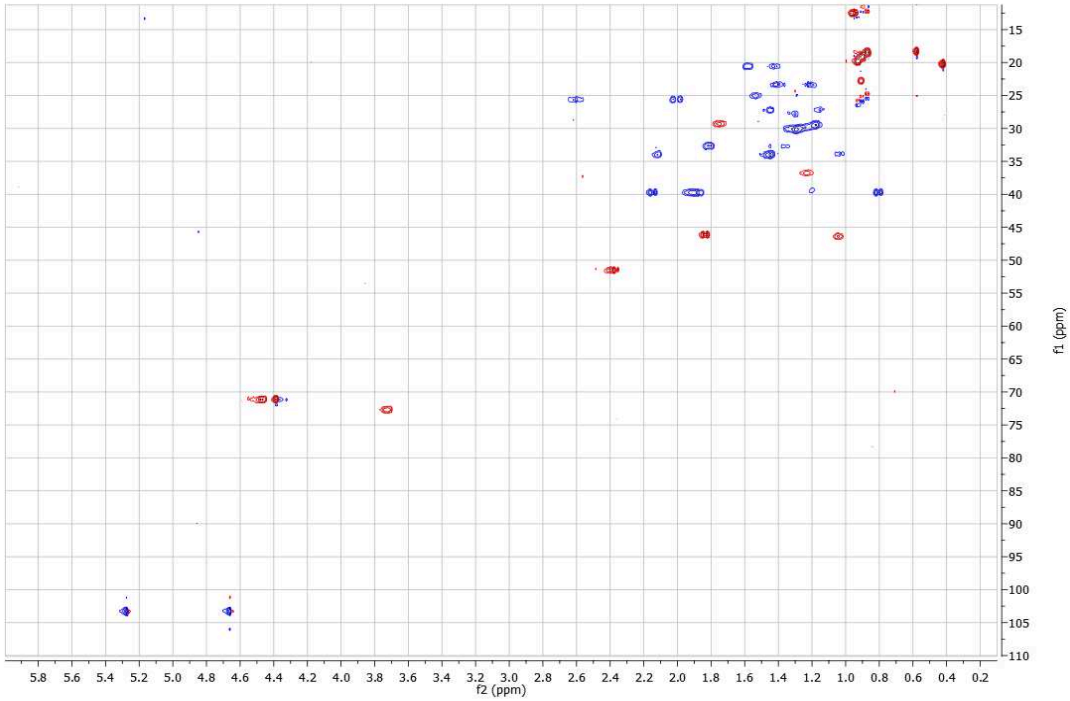
^{1}H NMR (500 MHz, C_6D_6) of Theonellasterol G (**45**)



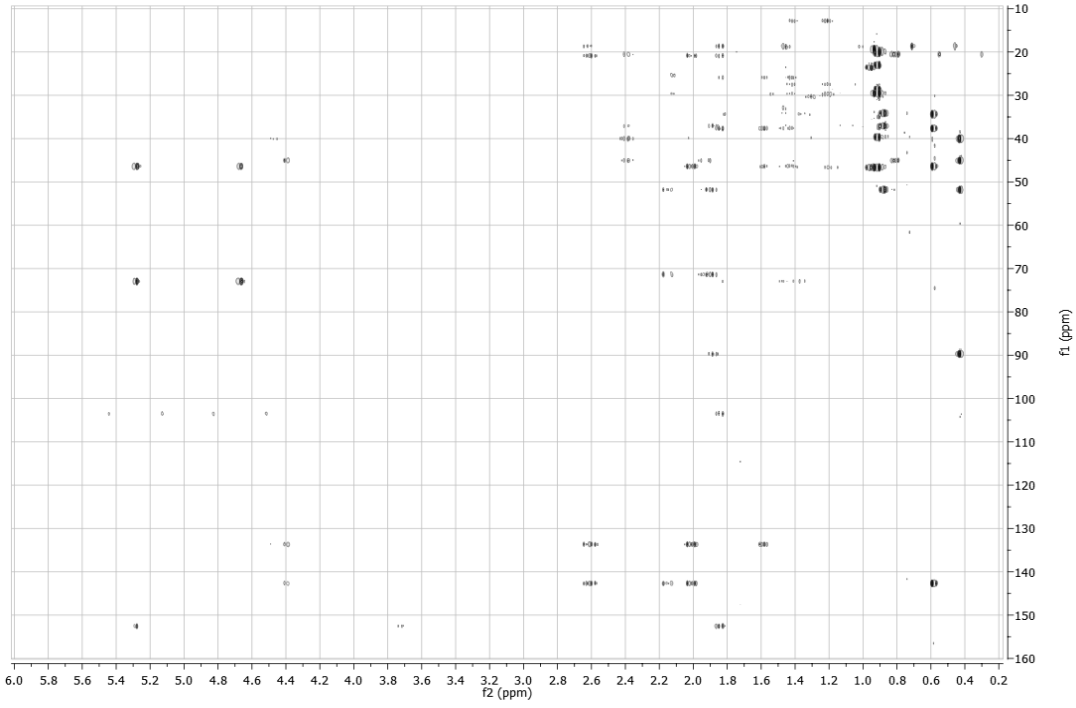
COSY spectrum (500 MHz, C_6D_6) of Theonellasterol G (**45**)



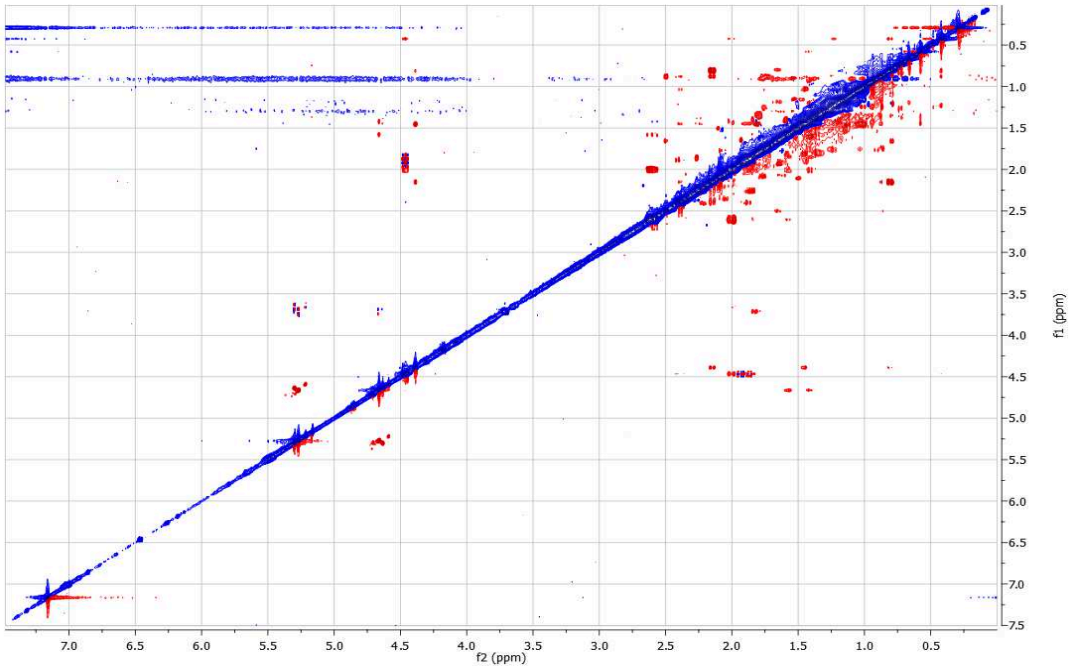
HSQC spectrum (500 MHz, C_6D_6) of Theonellasterol G (**45**)



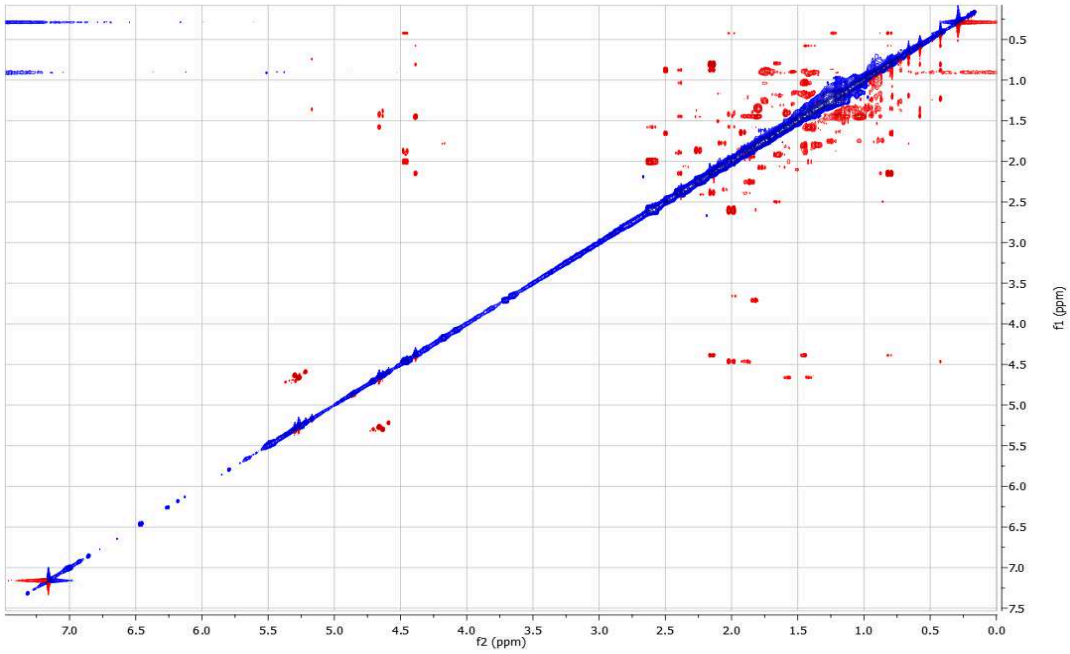
HMBC spectrum (500 MHz, C_6D_6) of Theonellasterol G (**45**)



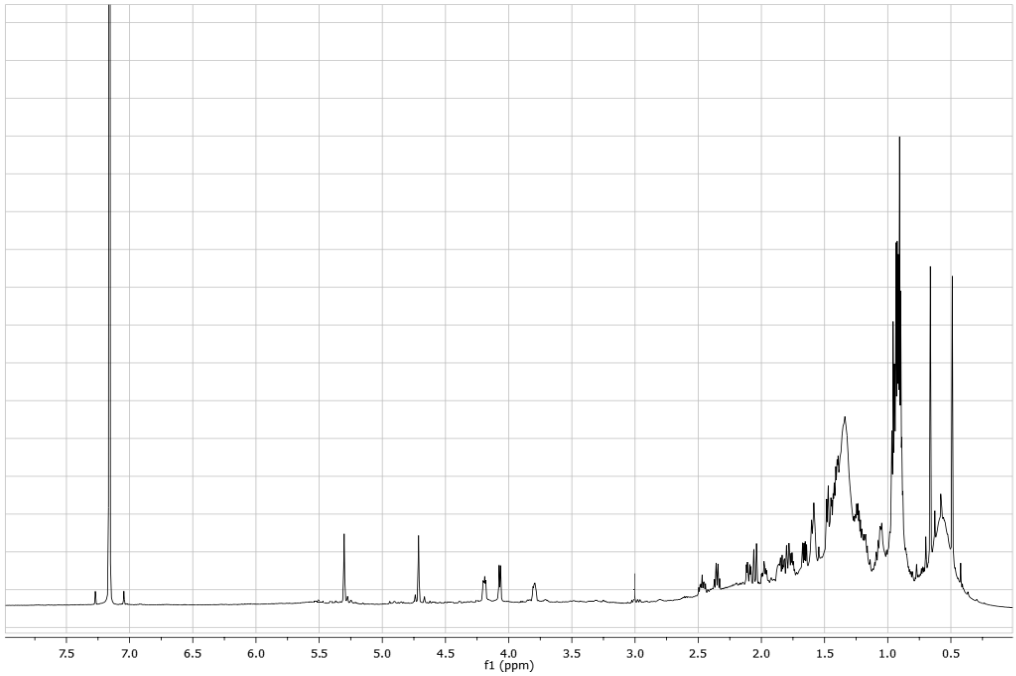
ROESY spectrum (500 MHz, C_6D_6, 200 ms) of Theonellasterol G (**45**)



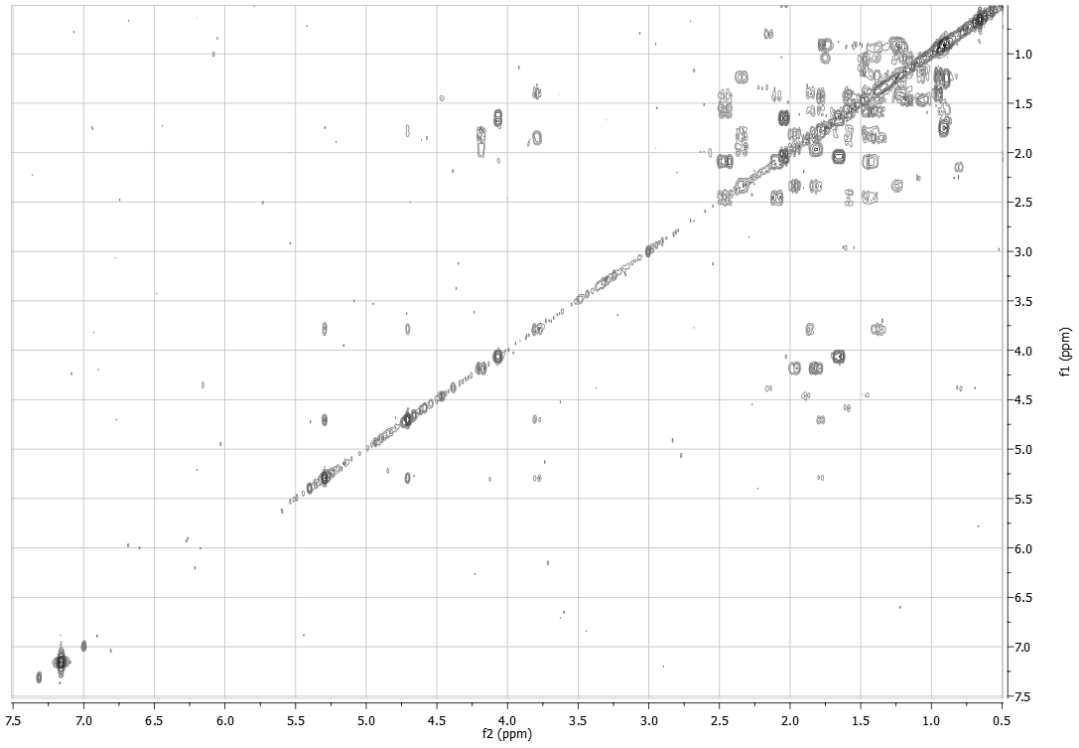
ROESY spectrum (500 MHz, C_6D_6 ,500 ms) of Theonellasterol G (**45**)



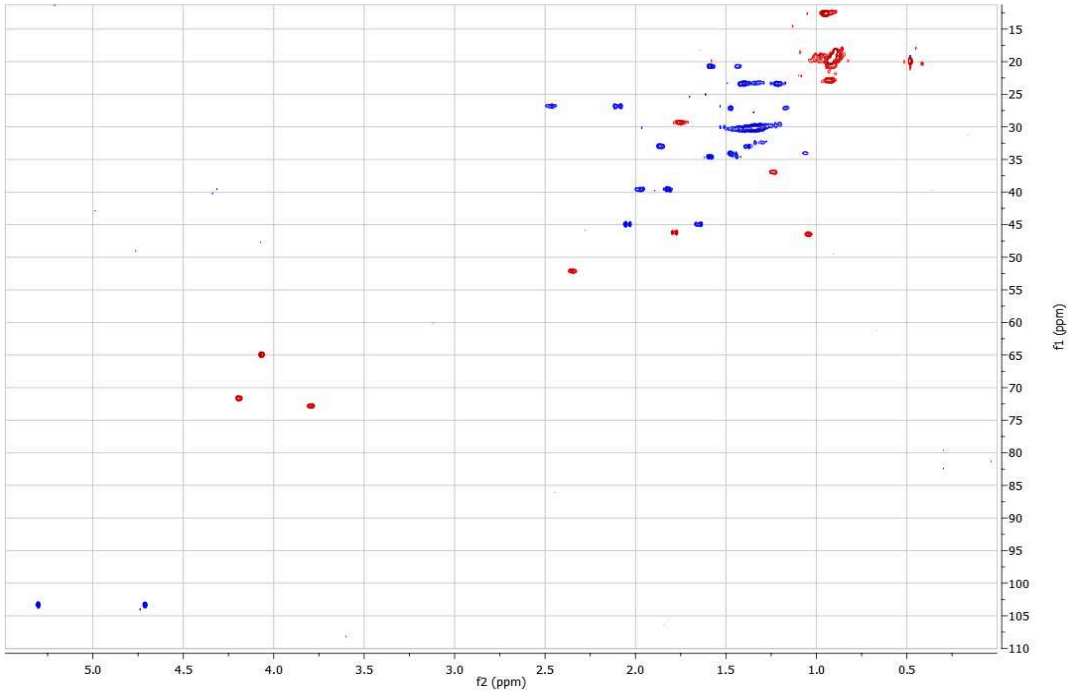
^{1}H NMR (700 MHz, C_6D_6) of Theonellasterol H (**46**)



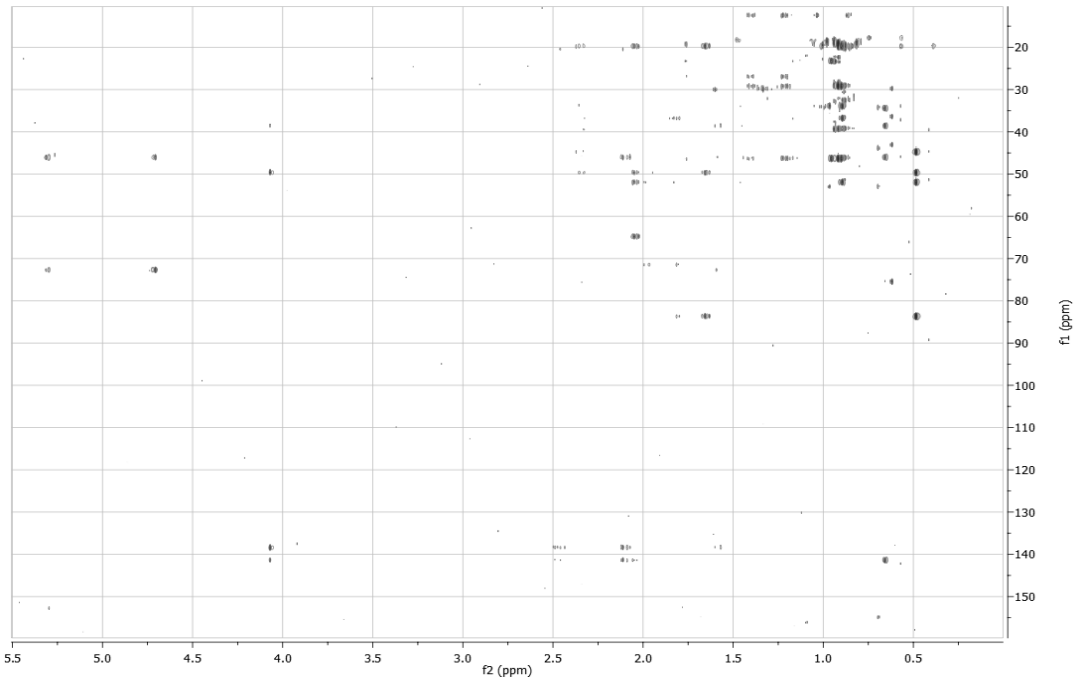
COSY spectrum (700 MHz, C_6D_6) of Theonellasterol H (**46**)



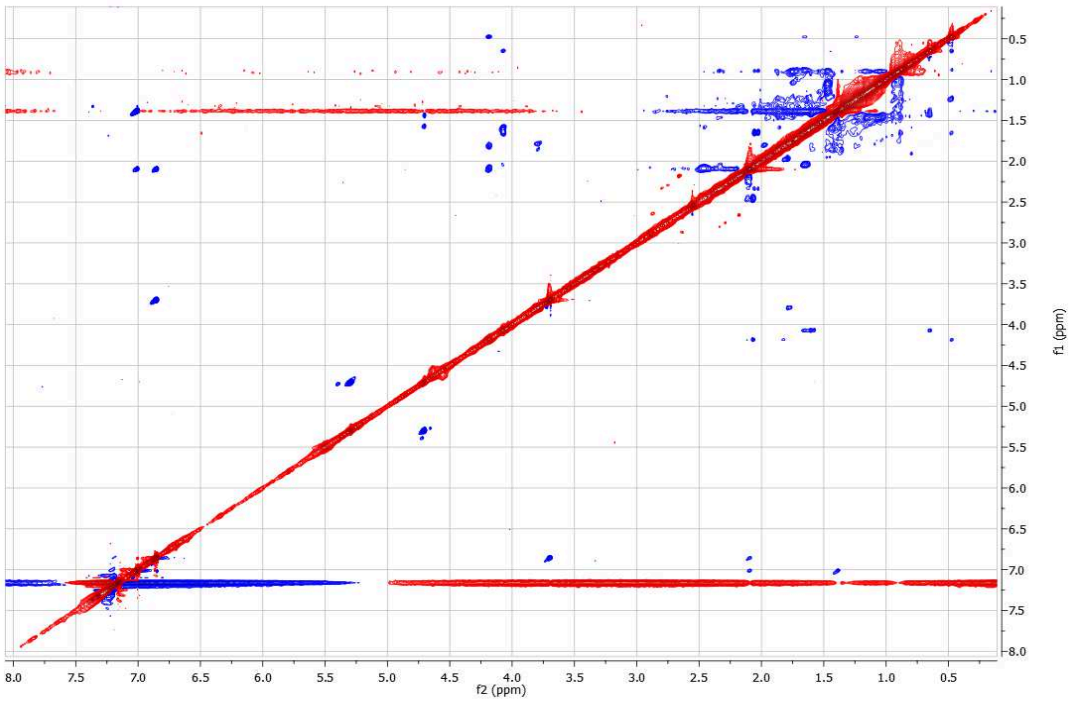
HSQC spectrum (700 MHz, C_6D_6) of Theonellasterol H (**46**)



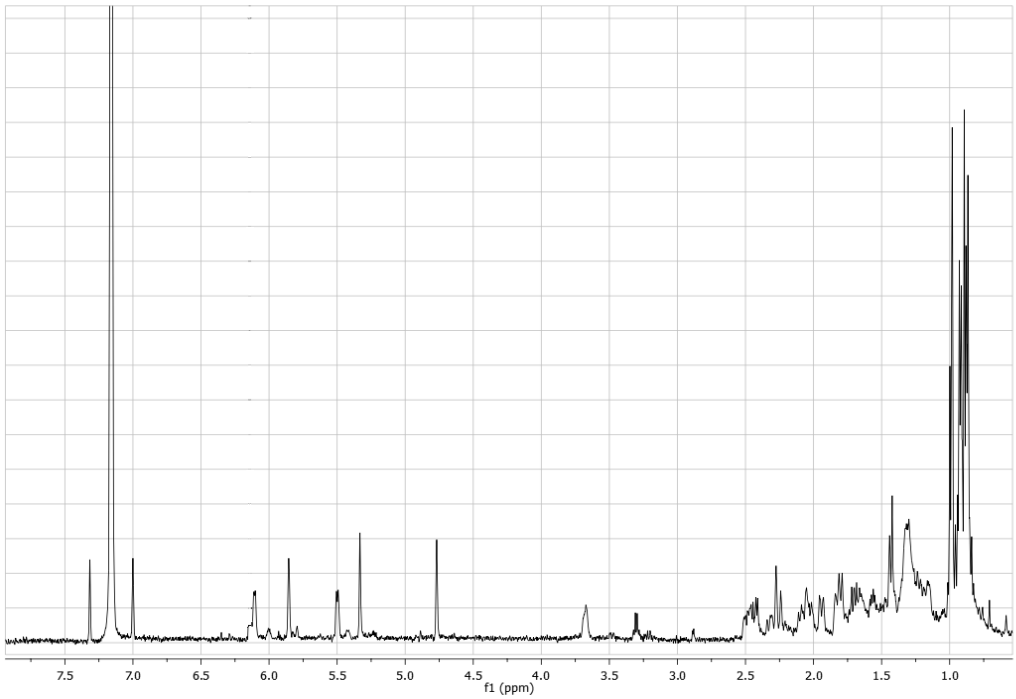
HMBC spectrum (700 MHz, C_6D_6) of Theonellasterol H (**46**)



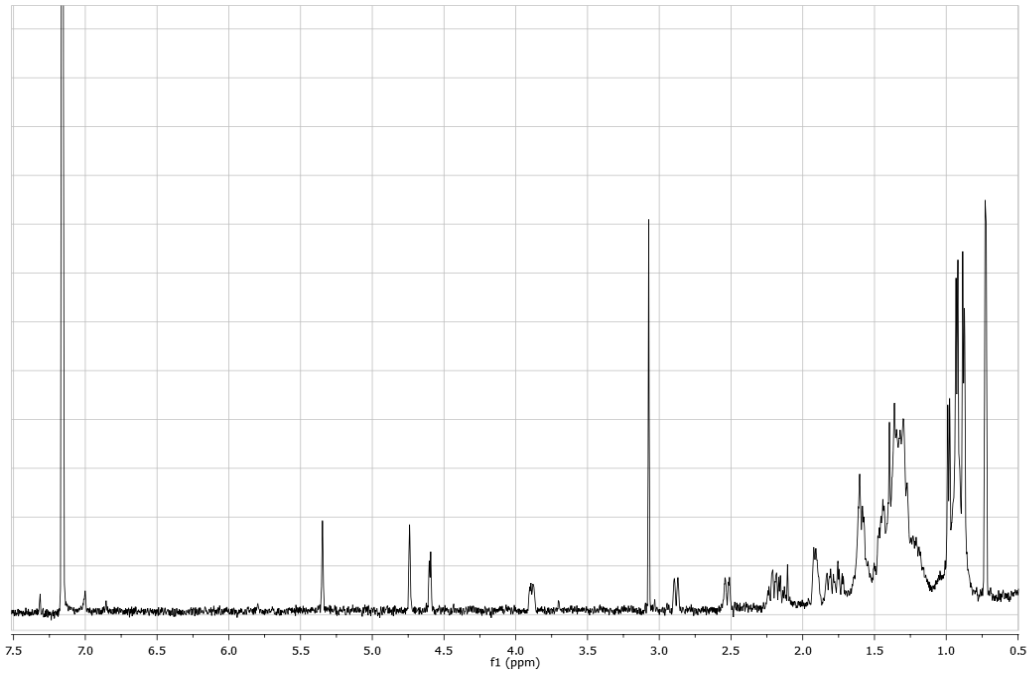
ROESY spectrum (500 MHz, C_6D_6, 500 ms) of Theonellasterol H (**46**)



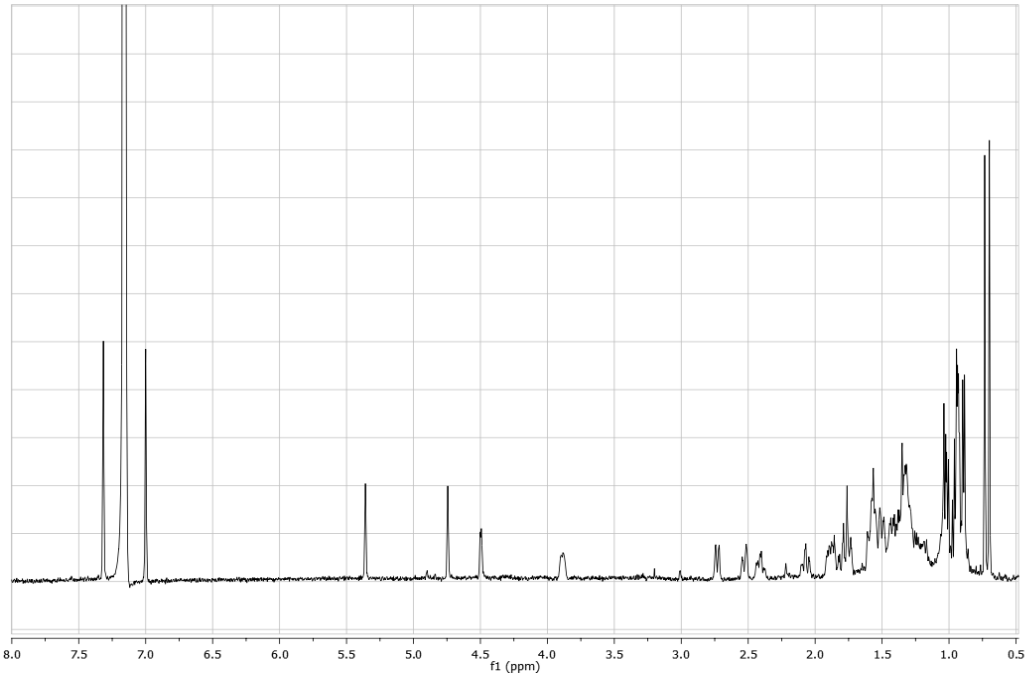
^{1}H NMR (500 MHz, C_6D_6) of Conicasterol B (**47**)



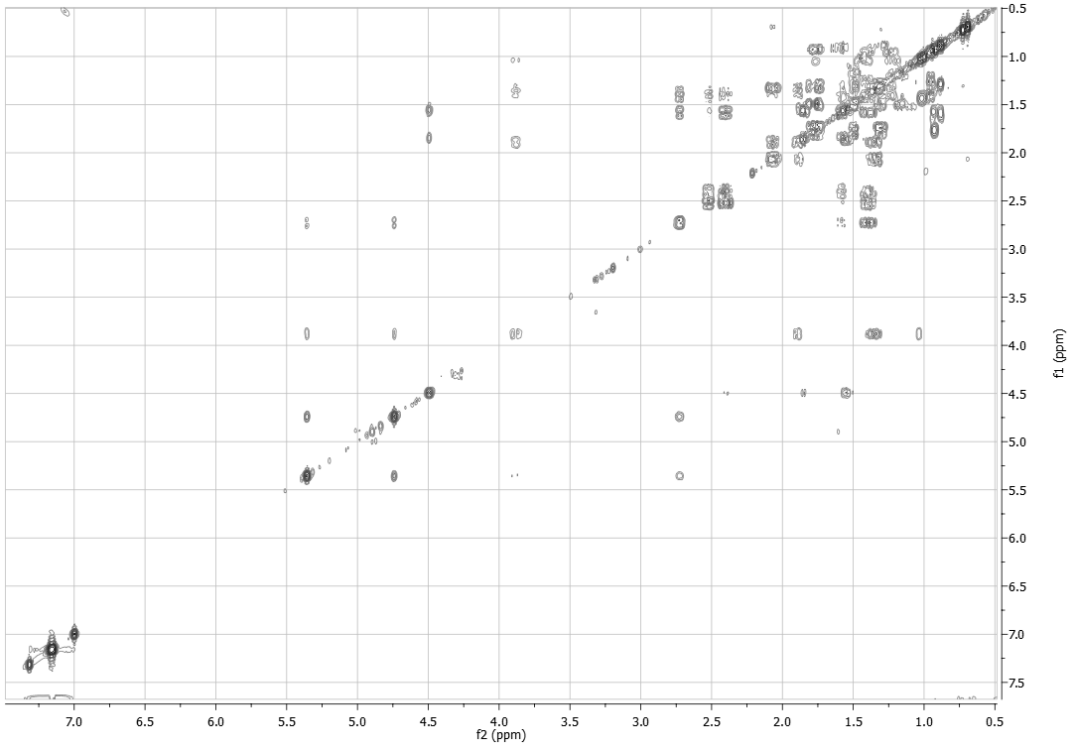
^{1}H NMR (500 MHz, C_6D_6) of Conicasterol C (**48**)



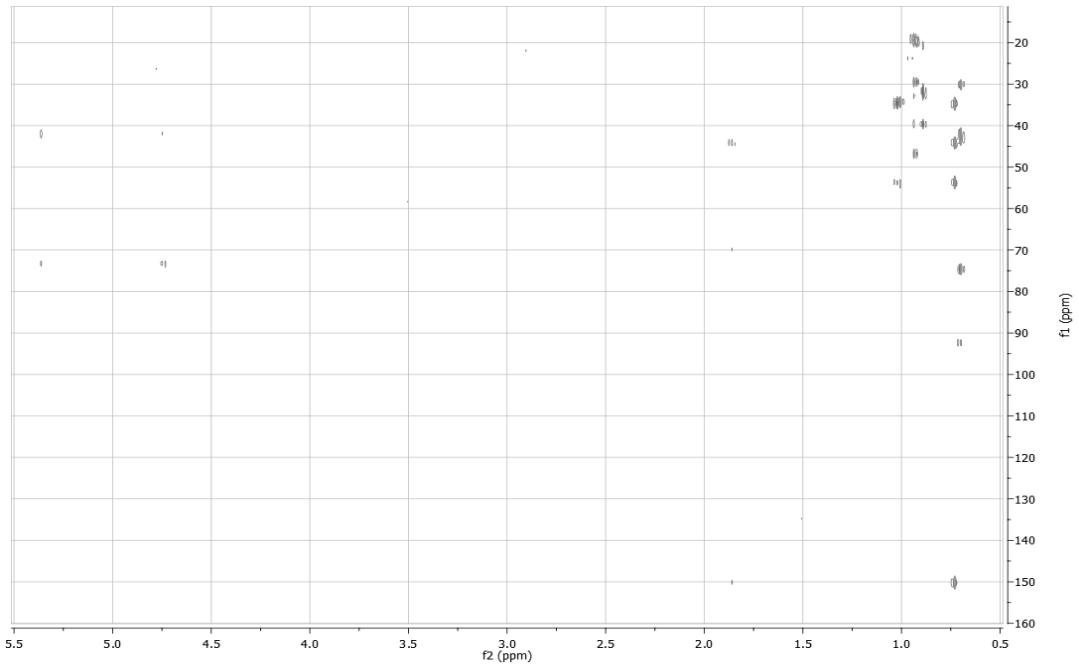
^{1}H NMR (500 MHz, C_6D_6) of Conicasterol D (**49**)



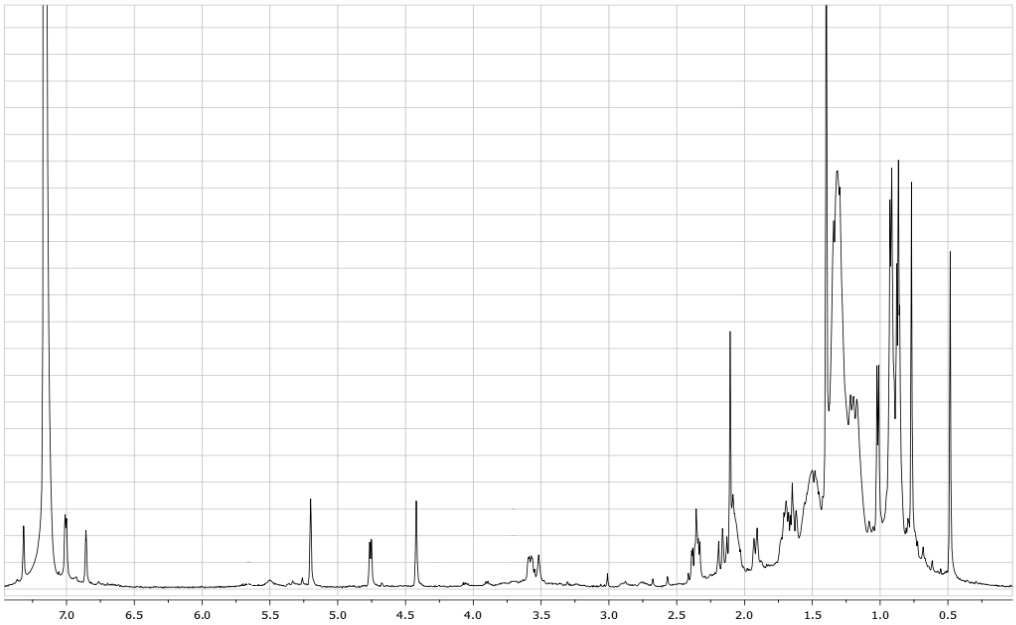
COSY spectrum (500 MHz, C_6D_6) of Conicasterol D (**49**)



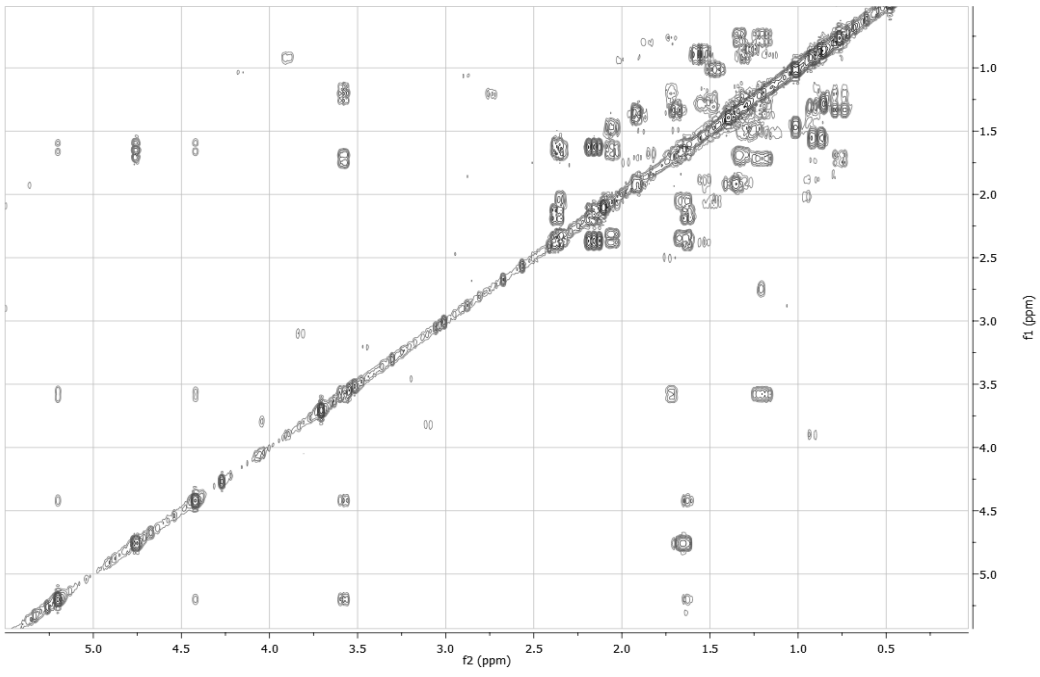
HMBC spectrum (500 MHz, C_6D_6) of Conicasterol D (**49**)



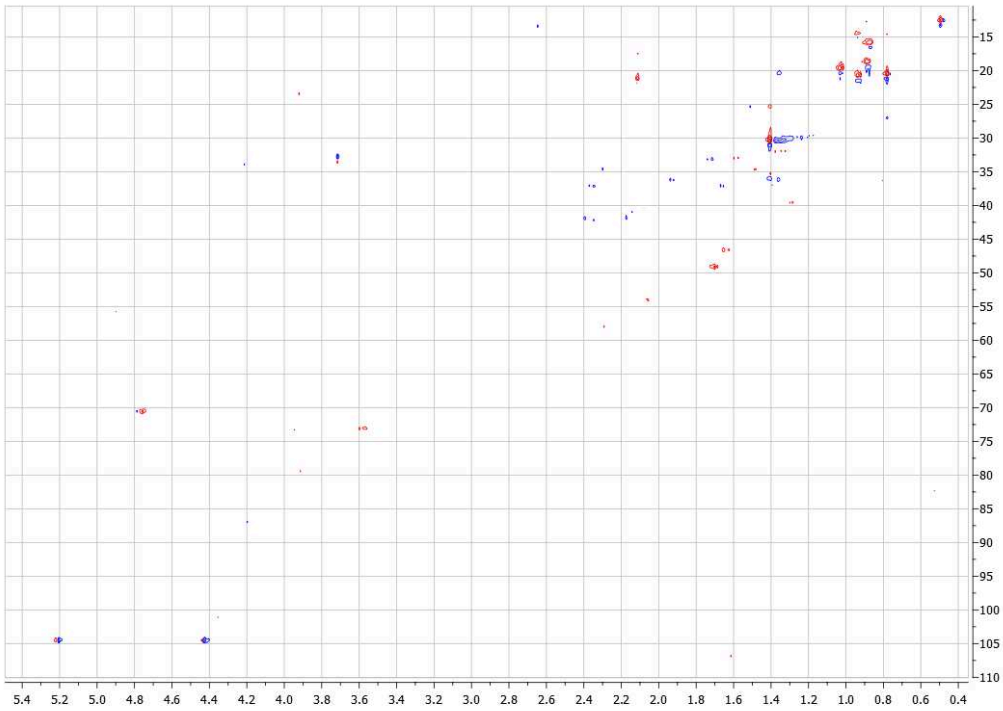
^{1}H NMR (C_6D_6, 500 MHz) of conicasterol G (**50**)



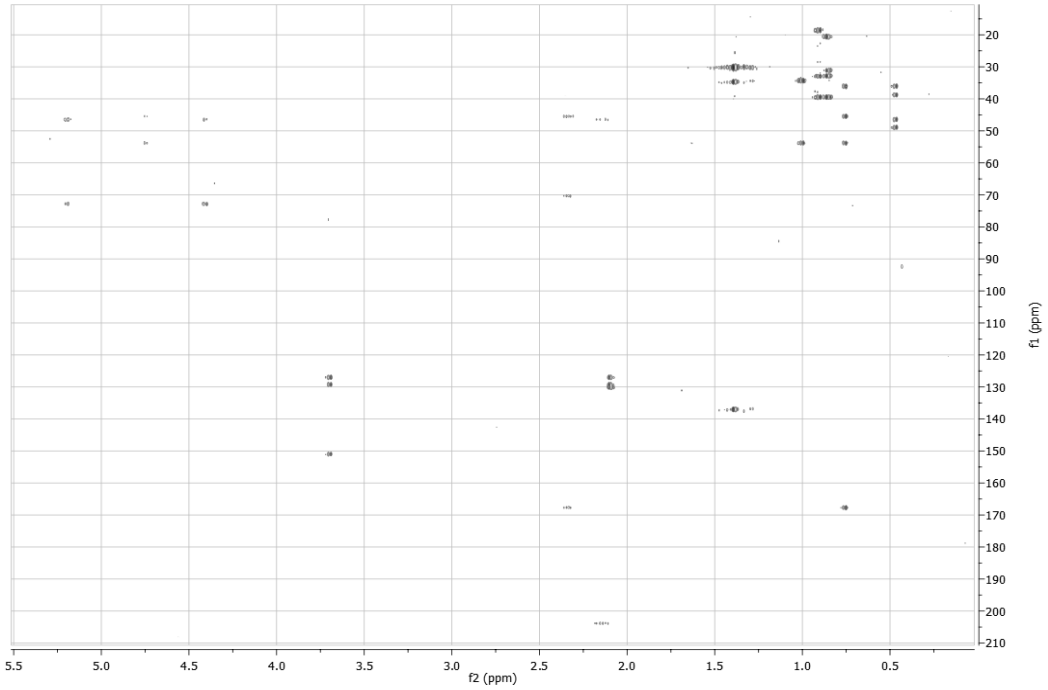
COSY spectrum (C_6D_6, 500 MHz) of conicasterol G (**50**)



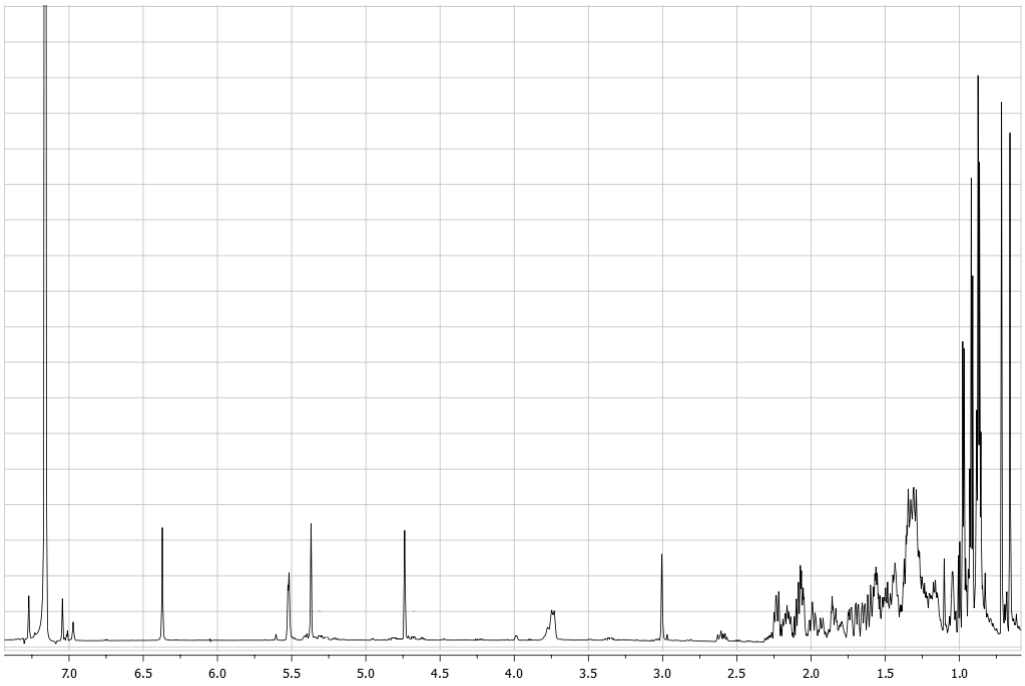
HSQC spectrum (C_6D_6, 500 MHz) of conicasterol G (**50**)



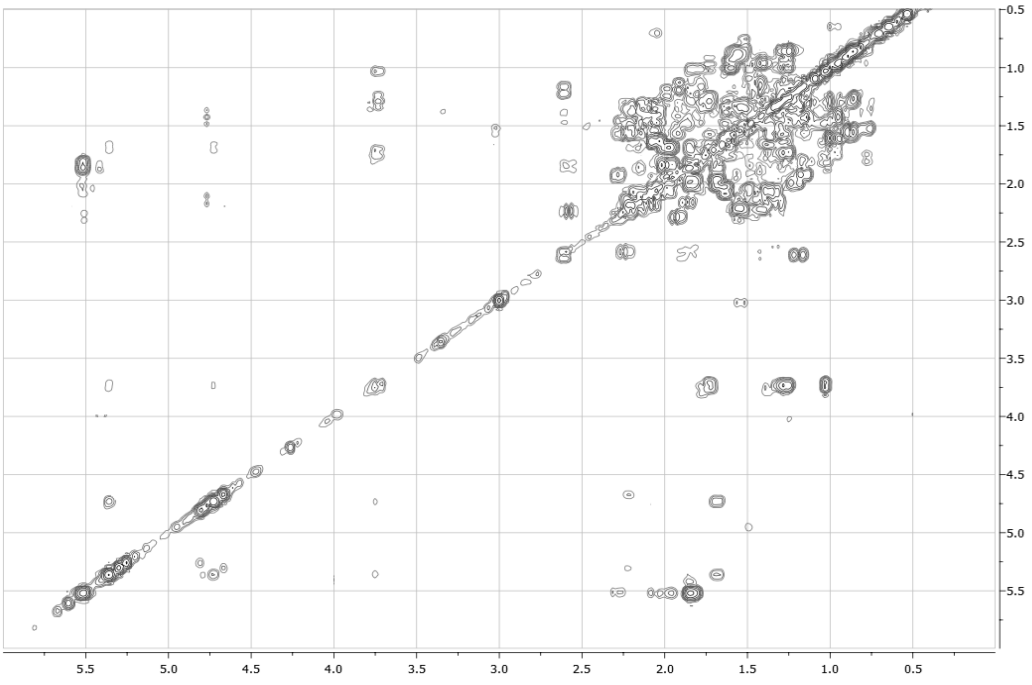
HMBC spectrum (C_6D_6, 500 MHz) of conicasterol G (**50**)



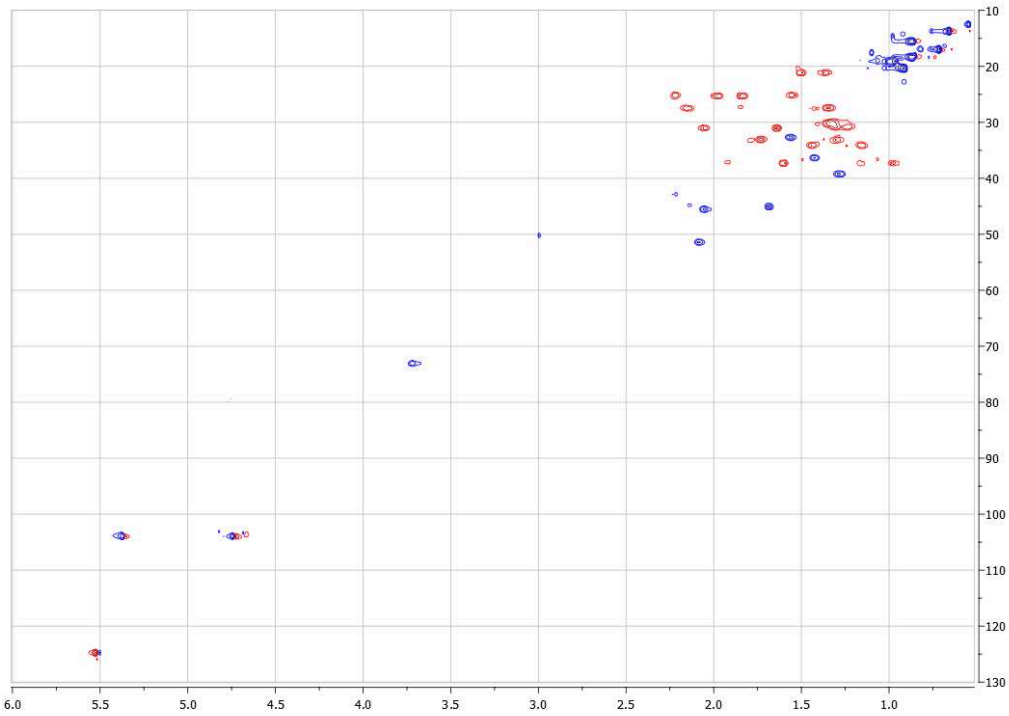
^{1}H NMR (C_6D_6, 700 MHz) of conicasterol H (**51**)



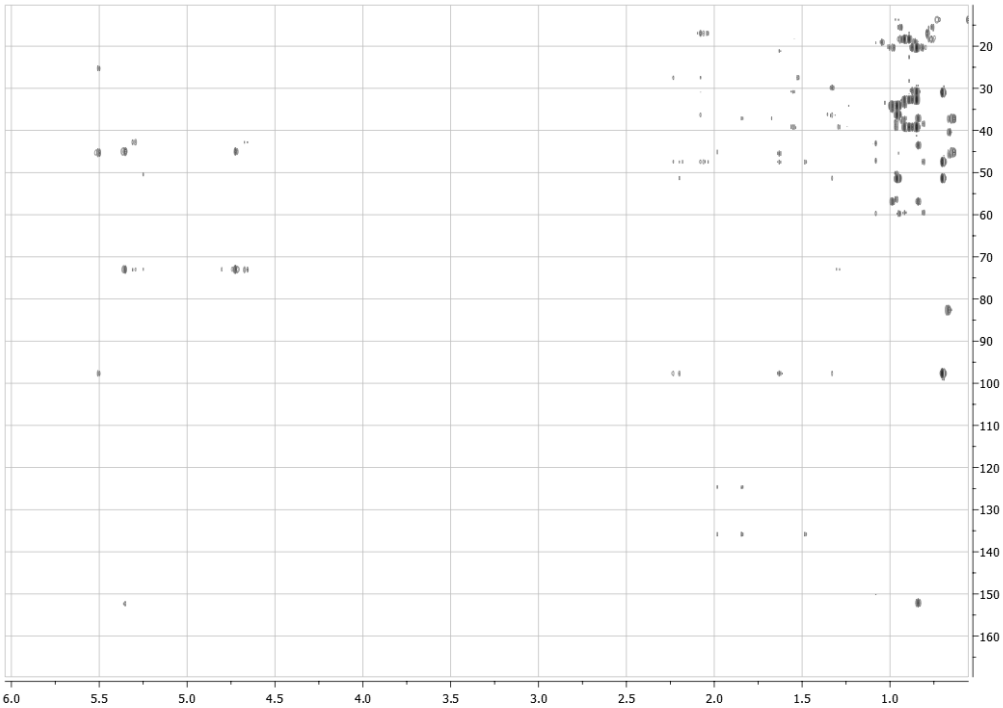
COSY spectrum (C_6D_6, 500 MHz) of conicasterol H (**51**)



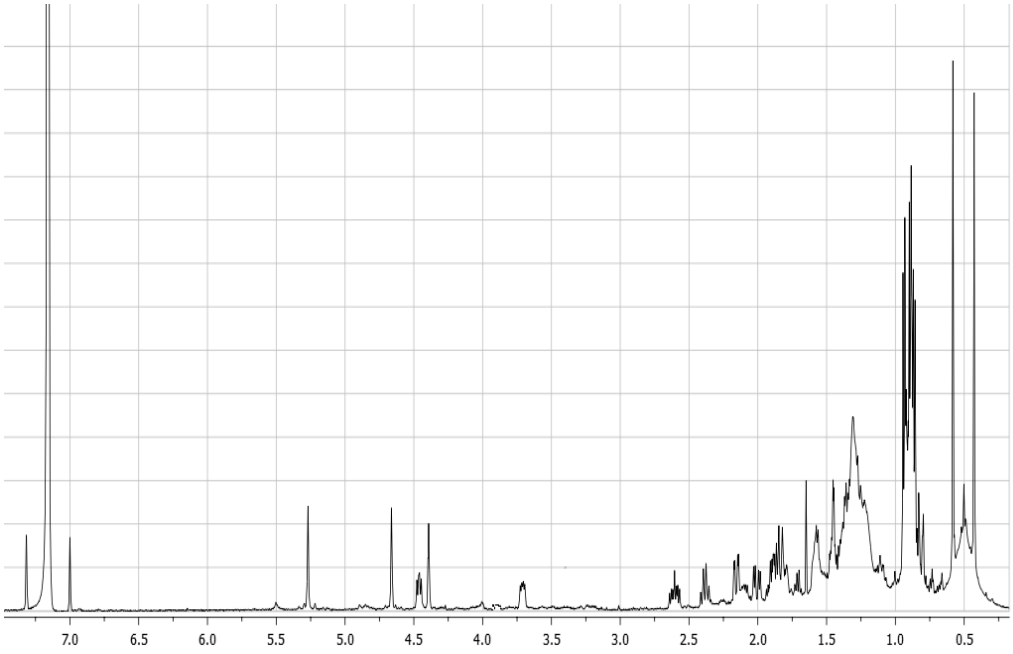
HSQC spectrum (C_6D_6, 700 MHz) of conicasterol H (**51**)



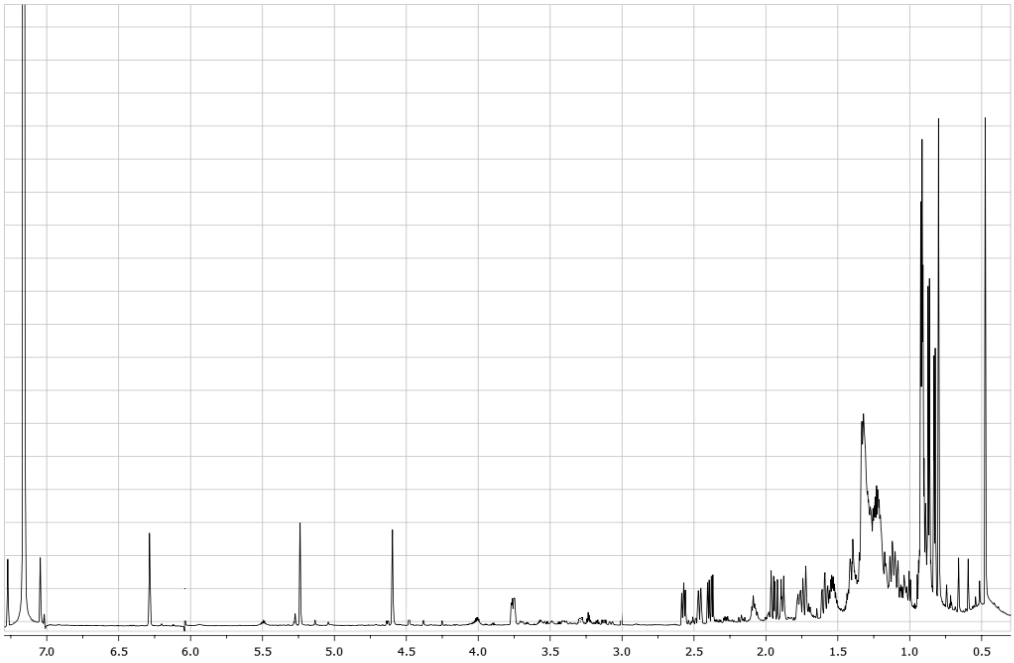
HMBC spectrum (C_6D_6, 700 MHz) of conicasterol H (**51**)



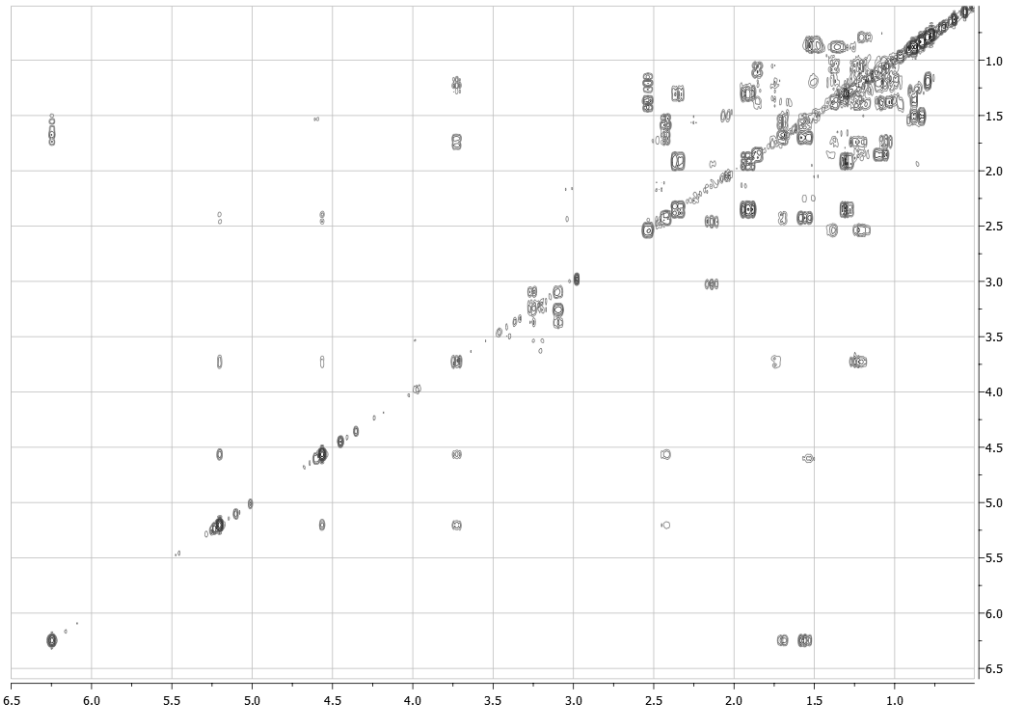
^{1}H NMR (C_6D_6, 500 MHz) of conicasterol I (**52**)



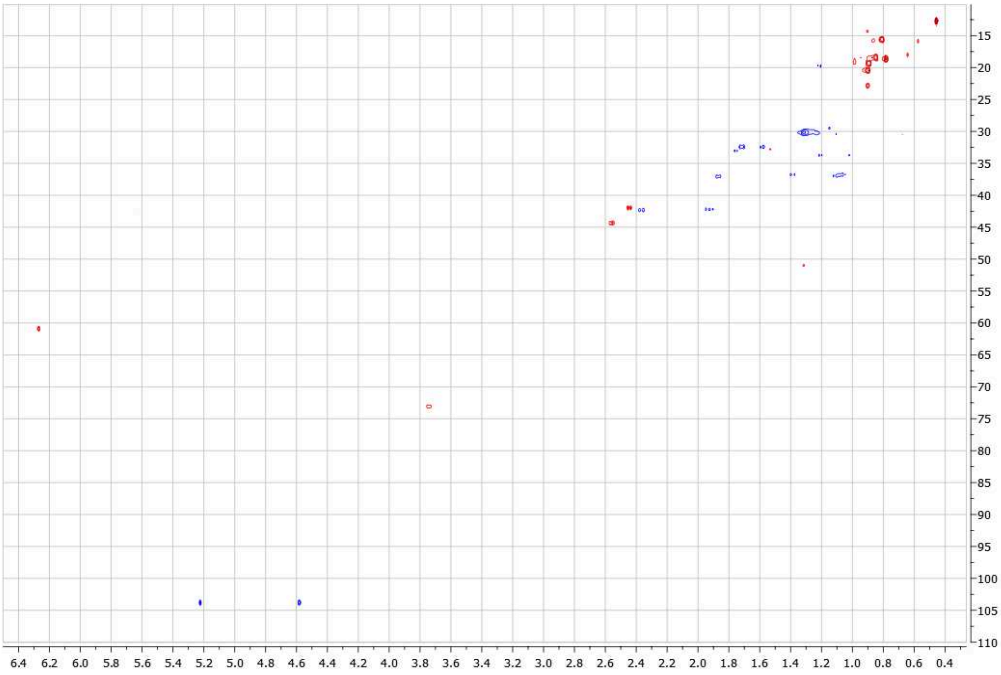
^{1}H NMR (C_6D_6, 500 MHz) of conicasterol J (**53**)



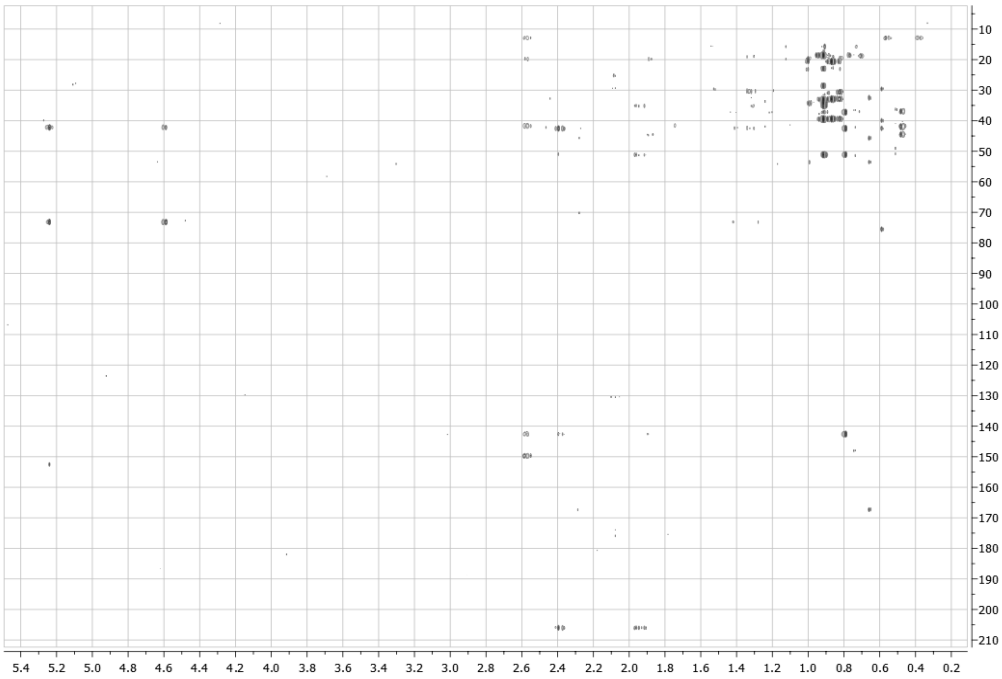
COSY spectrum (C_6D_6, 500 MHz) of conicasterol J (**53**)



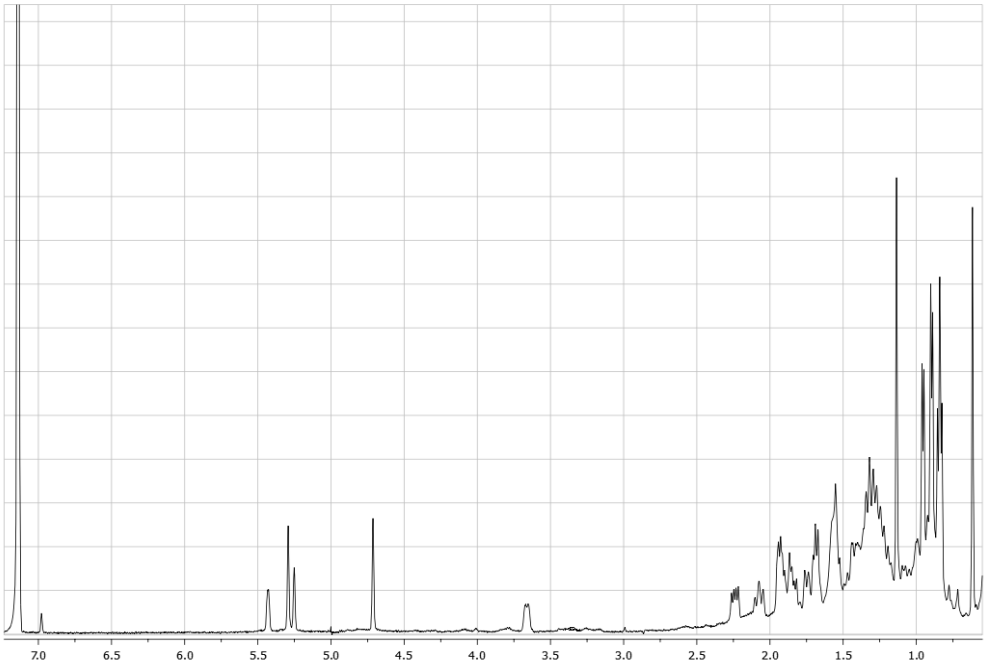
HSQC spectrum (C_6D_6, 500 MHz) of conicasterol J (**53**)



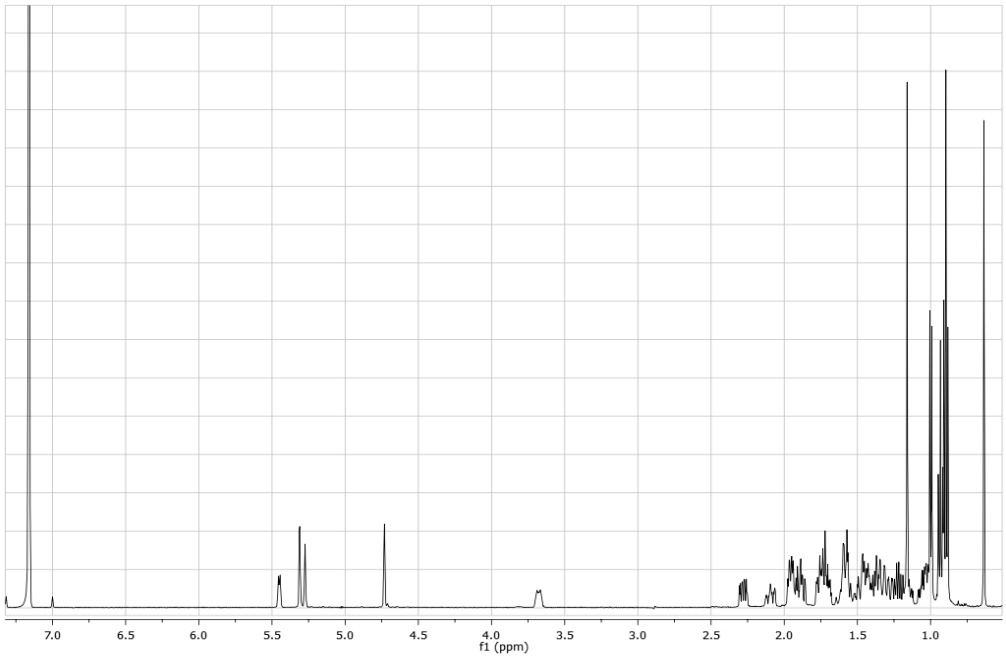
HMBC spectrum (C_6D_6, 500 MHz) of conicasterol J (**53**)



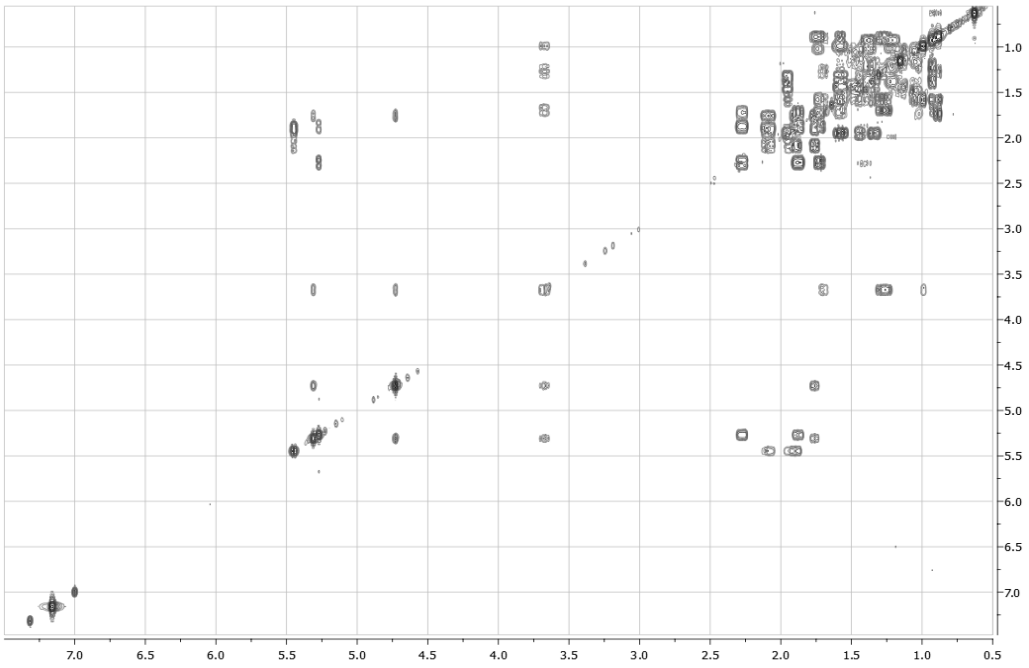
^{1}H NMR (C_6D_6, 500 MHz) of conicasterol K (**54**)



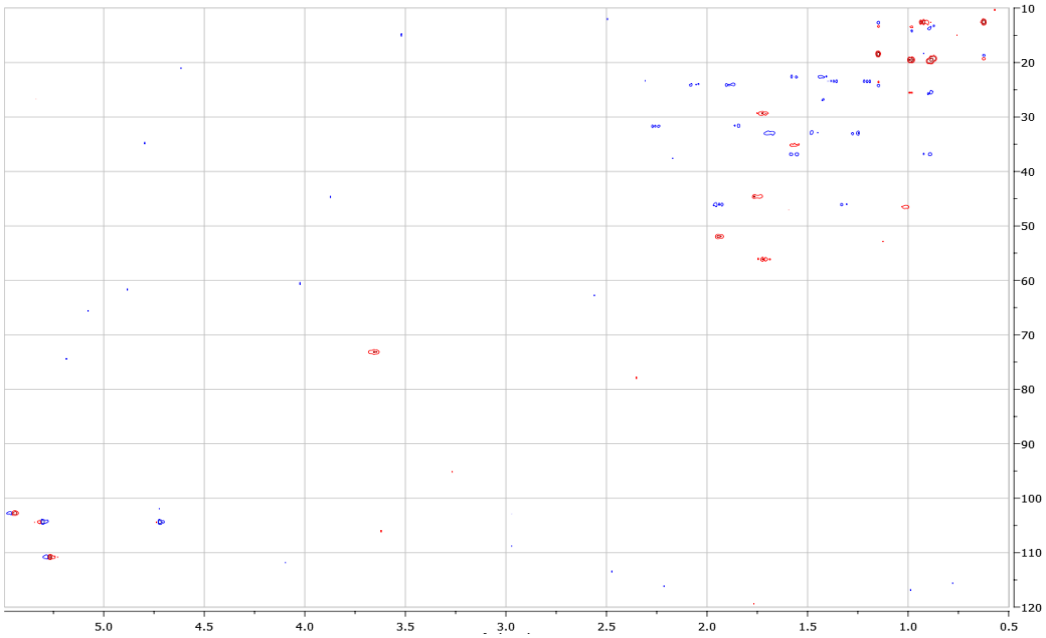
^{1}H NMR (C_6D_6, 500 MHz) of Theonellasterol J (**55**)



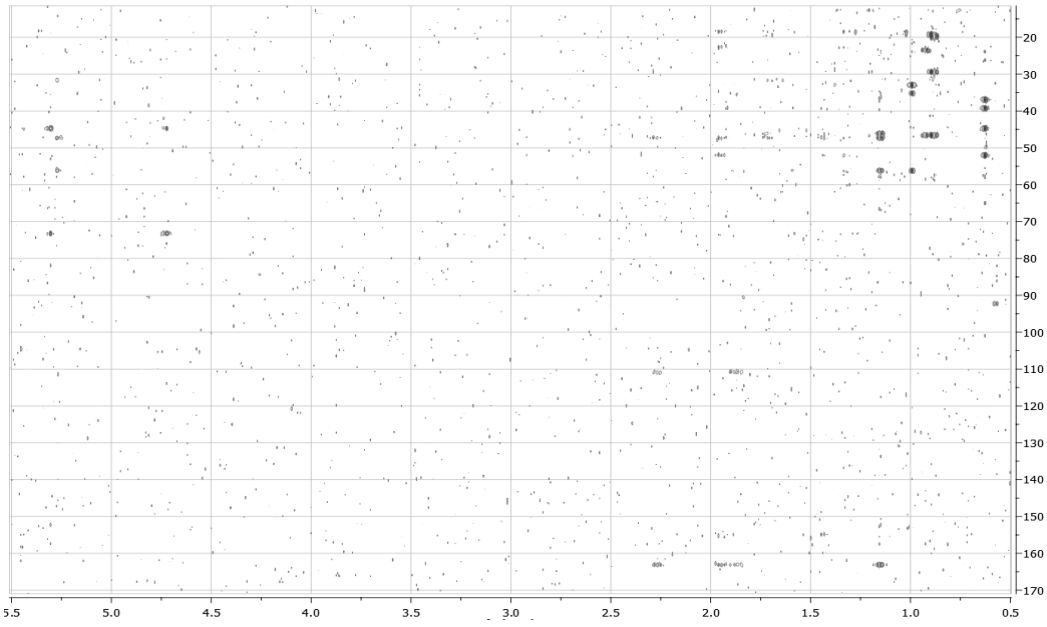
COSY spectrum (C_6D_6, 500 MHz) of theonellasterol J (**55**)



HSQC spectrum (C_6D_6, 500 MHz) of theonellasterol J (**55**)



HMBC spectrum (C_6D_6, 500 MHz) of Theonellasterol J (**55**)



IV. Experimental section of FXR modulators

Sponge material and isolation of conicasterol E (60).

The hexane extract (4.5 g) of sponge *Theonella swinhoei* (R3159) was chromatographed in two runs by silica gel MPLC using a solvent gradient system from CH_2Cl_2 to $CHCl_2$:MeOH 1:1. Fractions eluted with CH_2Cl_2:MeOH 96:4 (18.5 mg) were further purified by HPLC on a Nucleodur 100-5 C18 (5µm; 10 mm i.d. x 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 2.1 mg of conicasterol E (**60**) (t_R=14.5 min).

Conicasterol E (**60**): white amorphous solid; $[\alpha]_D^{25}$ +59.6 (*c* 0.06, MeOH); ^{1}H and ^{13}C NMR data in C_6D_6 given in Table 18; ESI-MS: *m/z* 451.4 $[M+Li]^+$. HRMS (ESI): calcd. for $C_{29}H_{48}LiO_3$: 451.3763; found 451.3769 $[M+Li]^+$.

Total synthesis of 6-ECDCA (61).

Benzyl 3α-hydroxy-7-keto-5β-cholan-24-oate (**63**). An oven-dried 250 mL flask was charged with chenodeoxycholic acid (**62**) (2.00 g, 5.1 mmol), sodium bromide (30 mg, 0.25 mmol), tetrabutylammonium bromide (5.4 g, 16.8 mmol) and 43 mL of a solution of MeOH:CH_3COOH:H_2O:AcOEt 3:1:0.25:6.5 v/v. The mixture was stirred at room temperature until a homogeneous solution formed and then cooled at 0 °C. Sodium hypochlorite solution (10%, 5 mL, 5.6 mmol) was added, until the test for hypochlorite (peroxide test paper) was positive and the yellow suspension was stirred. The reaction was stirred at room temperature for 6 h. Aqueous sodium bisulfite (3.3%) was added to afford a white suspension (negative test for peroxide). Water (50 ml) was added and mixture stirred at 15 °C for 5 min. Aqueous solution was extracted with AcOEt (3 × 50 mL). The combined organic layer was washed with aqueous sodium bisulfite (50 mL) and water (50 mL), and then dried over anhydrous $MgSO_4$ and evaporated in vacuo to give 2.0 g of 3α-hydroxy-7-keto-5β-cholan-24-oic acid, that was subjected to next step without any purification. To a solution of this latter intermediate (2.0 g, 5.1 mmol) in CH_3CN dry (30 mL), Cs_2CO_3 (2.5 g, 7.6 mmol) was added. The solution was heated to 150 °C and BnBr (3.0 mL, 25.5 mmol) was added under reflux. The solution was stirred at this temperature for 24 h and then after cooled to room temperature and after removal CH_3CN in the rotavapor, poured into saturated $NaHCO_3$ solution (50 mL) and extracted with AcOEt (3 × 30 mL). The combined organic layer was washed with water (30 mL), and then dried over anhydrous $MgSO_4$ and evaporated in vacuo. Purification on silica gel column eluting with hexane:AcOEt (7:3) and 0.5% Et_3N, afforded pure **63** (1.5 g, 60% over two steps). $[\alpha]_{24}{}^{D}$= -10.9 (*c* 0.99, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.33 (5H, m), 5.09 (2H, dd, *J* = 12.9, 15.6 Hz), 3.57 (1H, m), 2.83 (1H, m), 2.37 (1H, m, ovl), 2.26 (1H, m), 2.15 (1H, m), 1.37 (3H, s), 0.89 (3H, d, *J*= 6.0 Hz), 0.61 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$): δ 212.0, 174.0, 136.2, 128.5 (2C), 128.2 (2C), 128.1, 70.7, 66.0, 54.7, 49.4, 48.8, 46.0, 45.3, 42.7, 42.5, 38.9, 37.3, 35.1, 34.1, 31.2, 30.9, 29.8 (2C), 28.2, 24.7, 23.0, 21.6, 18.3, 11.9; HRMS-ESI *m/z* 480.3240 ($[M + H]^+$, $C_{31}H_{44}O_4$ requires 480.3235.

Benzyl 3α-trimethylsilyloxy-5β-cholan-6-en-24-oate (**64**). To a solution of diisopropylamine (5.5 mL, 39 mmol) in dry THF (50 mL) was added dropwise a solution of *n*-butyllithium (15 mL, 2.5M in hexane, 37.2 mmol) at -78°C under nitrogen atmosphere. After 30 min, trimethylchlorosilane

(3.9 mL, 31 mmol) was added and the resulting mixture was reacted for additional 20 min. A solution of benzyl 3α-hydroxy-7-keto-5β-cholan-24-oate (1.5 g, 3.1 mmol) in dry THF (20 mL) was added dropwise in 10 min. The reaction was stirred at -78 °C for an additional 45 min and then triethylamine (7.8 mL, 56 mmol) was added. After 1 h, the reaction mixture was allowed to warm to -20 °C, treated with aqueous saturated solution of $NaHCO_3$ (10 mL) and brought up to room temperature in 2 h. The organic phase was separated, and aqueous phase was extracted with ethyl acetate (3x50 mL). The combined organic phases were washed several times with saturated solution of $NaHCO_3$, water and brine. After drying over anhydrous Na_2SO_4, the residue was evaporated under vacuum to give 1.8 g of yellow residue, that was subjected to next step without any purification. $[\alpha]_{24}{}^{D}$= +4.3 (*c* 0.58, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.35 (5H, m), 5.10 (1H, dd, *J*= 15.8, 12.4 Hz), 4.72 (1H, d, *J*= 6.1 Hz), 3.50 (1H, m), 2.40 (1H, m), 2.28 (1H, m), 0.90 (3H, d, *J*= 6.3 Hz), 0.81 (3H, s), 0.64 (3H, s), 0.18-0.001 (18H, ovl). ^{13}C NMR (100 MHz $CDCl_3$): δ 174.0, 151.7, 136.2, 128.5 (2C), 128.2 (2C), 128.1, 108.8, 71.5, 66.1, 54.8, 54.1, 44.3, 42.6, 41.0, 40.9, 40.3, 40.1, 35.2, 34.6, 32.9, 31.4, 31.0, 30.7, 28.6, 27.0, 22.5, 20.9, 18.4, 12.4, 1.4 (3C), 0.4 (2C), 0.2; HRMS-ESI m/z 624.4030 ($[M + H]^+$, $C_{37}H_{60}O_4Si_2$ requires 624.4047.

Benzyl 3α-hydroxy-6-ethyliden-7-keto-5β-cholan-24-oate (**65**). To a cooled (-60 °C) and stirred solution of acetaldehyde (328 μL, 5.86 mmol) and benzyl 3α-trimethylsilyloxy-5β-cholan-6-en-24-oate (**64**) (1.8 g, 2.93 mmol) in dry CH_2Cl_2 (30 mL) was added dropwise $BF_3{\cdot}OEt_2$ (3.7 mL, 29.3 mmol). The reaction mixture was stirred for 2 h at -60 °C and allowed to warm to room temperature. The mixture was quenched with saturated aqueous solution of $NaHCO_3$ and extracted with CH_2Cl_2. The combined organic phases were washed with brine, dried over anhydrous $MgSO_4$ and concentrated under *vacuum*. Purification on silica gel column eluting with hexane:AcOEt (9:1) and 0.5% Et_3N, afforded pure **65** (1.09 g, 70% over two steps). $[\alpha]_{24}{}^{D}$= -42.5 (*c* 0.12, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.34 (5H, m), 6.16 (1H, q, *J* = 6.7 Hz), 5.10 (2H, dd, *J* = 13.0, 17.8 Hz), 3.64 (1H, m), 2.56 (1H, m), 2.38 (1H, m), 2.27 (1H, m), 1.67 (3H, d, *J* = 6.7 Hz), 0.99 (3H, s), 0.91 (3H, d, *J* = 6.0 Hz), 0.60 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$): δ 215.2, 174.0, 143.3, 140.9, 129.8, 128.5 (2C), 128.2 (2C), 128.1, 70.5, 66.1, 54.5, 50.6, 48.6, 45.5, 43.5, 39.1, 38.9, 37.5, 35.1, 34.4, 31.2, 30.9, 29.6 (2C), 28.4, 25.9, 22.8, 21.3, 18.4, 12.6, 12.0; HRMS-ESI *m/z* 506.3396 ($[M + H]^+$, $C_{33}H_{46}O_4$ requires 506.3378.

Benzyl 3α,7α-dihydroxy-6-ethyliden-5β-cholan-24-oate (**66**). The residual (1.00 g, 1.97 mmol) was dissolved in a solution of dry tetrahydrofuran/dry methanol (50 mL, 4/1 v/v) and treated with $CeCl_3$ (1.46 g, 5.93 mmol) and $NaBH_4$ (667 mg, 2.36 mmol). After 3 h, water and MeOH was added. Then after evaporation of the solvents, the residue was diluted with water and extracted with ether (3x50 mL). The combined organic phases were washed with brine, dried over Na_2SO_4 anhydrous and evaporated under reduced pressure. The crude residue was purified by flash chromatography on silica gel using dichloromethane/ methanol (7:3 v/v) and 0.5% Et_3N as eluent, to afford 950 mg of **66** (95 % yield). $[\alpha]_{24}{}^{D}$= + 18 (*c* 0.06, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.34 (5H, m), 5.64 (1H, q, *J* = 6.2 Hz), 5.10 (2H, dd, *J* = 12.5, 17.3 Hz), 3.98 (1H, m), 3.64 (1H, m), 2.47 (1H, m), 2.39 (1H, m), 2.27 (1H, m), 1.60 (3H, d, *J* = 6.6 Hz), 0.91 (3H, d, *J* = 6.2 Hz),

0.77 (3H, s), 0.61 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$): δ 174.0, 141.6, 136.1, 128.5 (2C), 128.3 (2C), 128.2, 114.4, 73.3, 71.2, 66.1, 56.0, 55.0, 45.5, 44.0, 40.0, 39.5, 36.3, 35.4, 35.2, 34.7, 31.3, 31.0, 30.2, 29.6, 28.6, 27.1, 22.8, 21.2, 18.4, 12.4, 12.2; HRMS-ESI *m/z* 508.3553 ($[M + H]^+$, $C_{33}H_{48}O_4$ requires 508.3565.

6-ECDCA (**61**). An oven-dried 50 mL flask was charged with 10% palladium on carbon (50 mg) and compound **66** (950 mg, 1.87 mmol) and the flask was evacuated and flushed with argon. Absolute methanol (10 mL) and dry THF (10 mL) were added, and the flask was flushed with hydrogen. The reaction was stirred at room temperature under H_2 for 4 h. The mixture was filtered through celite, and the recovered filtrate was evaporated under vacuum to give pure 6α-ethyl-chenodeoxycholic acid **61** (630 mg, 80%). $[\alpha]_{24}{}^{D}$= +5.11 (*c* 1.8, CH_3OH); ^{1}H NMR (400 MHz, CD_3OD): δ 3.60 (1H, brs), 3.32 (1H, m), 2.33 (1H, m), 2.20 (1H, m), 0.97 (3H, d, *J*= 6.2 Hz), 0.91 (3H, s), 0.90 (3H, t, *J*= 7.0 Hz), 0.69 (3H, s). ^{13}C NMR (100 MHz, CD_3OD): δ 178.4, 73.3, 71.2, 57.4, 51.7, 46.9, 43.8, 43.2, 41.6, 41.1, 36.8 (2C), 36.6, 34.5, 34.4, 32.4, 32.2, 31.2, 29.3, 24.6, 23.7, 23.5, 21.9, 18.8, 12.2, 12.0; HRMS-ESI m/z 420.3240 ($[M + H]^+$, $C_{26}H_{44}O_4$ requires 420.3237.

Synthesis of derivatives of theonellasterol.

3β-O-Methyl-theonellasterol (**67**). To a solution of theonellasterol (**38**) (10 mg, 0.023 mmol) in dry THF (5 mL) at 0 °C was added NaH (5.5 mg, 0.23 mmol). After 10 min methyl iodide (28.6 μL, 0.46 mmol) was added and the mixture was left to stand at room temperature for 4 h. The mixture was quenched by addition at 0 °C of methanol (2 mL) and then concentrated *in vacuo*. Ethyl acetate and water were added and the separated aqueous phase was extracted with ethyl acetate (3 × 50 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and concentrated. Purification by silica gel eluting with CH_2Cl_2 gave the methyl ether **67** as an amorphous solid (8.0 mg, 79%). $[\alpha]_{25}{}^{D}$ = +4.0 (*c* 0.75, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): *δ* 5.49 (1H, s, H-30a), 4.77 (1H, s, H-30b), 3.39 (1H, dd, *J* = 4.5, 11.7 Hz, H-3), 3.33 (3H, s, OCH_3), 1.06 (3H, d, *J* = 6.9 Hz, H_3-21), 0.94 (3H, t, *J* = 7.0 Hz, H_3-29), 0.93 (3H, s, H_3-18), 0.91 (3H, d, *J* = 7.0 Hz, H_3-26), 0.90 (3H, d, *J* = 7.0 Hz, H_3-27), 0.67 (3H, s, H_3-19). ^{13}C NMR (100 MHz C_6D_6): *δ* 150.7 (C-4), 143.3 (C-14), 126.8 (C-8), 104.5 (C-30), 83.4 (C-3), 57.6 (C-17), 57.4 (-OCH_3), 50.4 (C-9), 50.1 (C-5), 47.0 (C-24), 43.5 (C-13), 40.8 (C-10), 38.2 (C-12), 37.5 (C-1), 35.7 (C-20), 34.6 (C-22), 31.0 (C-2), 30.2 (C-6), 29.7 (C-25), 27.9 (C-7), 27.0 (C-23), 26.5 (C-16), 25.4 (C-15), 23.8 (C-28), 21.2 (C-11), 20.1 (C-27), 19.9 (C-21), 19.6 (C-26), 18.9 (C-18), 13.7 (C-19), 13.0 (C-29); HRESI MS *m/z* 441.4092 (calcd for $C_{31}H_{53}O$ 441.4096).

3β-O-Acetyl-theonellasterol (**68**). A mixture of theonellasterol (**38**) (10 mg, 0.023 mmol) and acetic anhydride (55 μL, 0.575 mmol) in dry pyridine (10 mL) was left to stand at room temperature for 8 h. Then the solvent was evaporated and purification by silica gel eluting with CH_2Cl_2 gave **68** as an amorphous solid (10.3 mg, 96%). $[\alpha]_{25}{}^{D}$ = +0.5 (*c* 2.4, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): *δ* 5.37 (1H, dd, *J* = 4.7, 11.5 Hz, H-3), 5.15 (1H, s, H-30a), 4.69 (1H, s, H-30b), 1.80 (3H, s, CH_3CO), 1.05 (3H, d, *J* = 6.6 Hz, H_3-21), 0.94 (3H, t, *J* = 7.0 Hz, H_3-29), 0.91 (3H, d, *J* = 6.5 Hz, H_3-26), 0.91 (3H, s, H_3-18), 0.90 (3H, d, *J* = 6.5 Hz, H_3-27), 0.63 (3H, s, H_3-

19). ^{13}C NMR (100 MHz C_6D_6): δ 169.7 (CH_3CO), 149.1 (C-4), 143.4 (C-14), 126.6 (C-8), 104.4 (C-30), 75.1 (C-3), 57.6 (C-17), 50.0 (C-9), 49.6 (C-5), 46.9 (C-24), 43.5 (C-13), 40.3 (C-10), 38.1 (C-12), 36.9 (C-1), 35.7 (C-20), 34.6 (C-22), 30.4 (C-6), 29.9 (C-2), 29.7 (C-25), 27.9 (C-7), 27.0 (C-23), 26.6 (C-16), 25.2 (C-15), 23.8 (C-28), 21.1 (2C, C-11 and CH_3CO), 20.1 (C-27), 19.9 (C-21), 19.6 (C-26), 18.8 (C-18), 13.6 (C-19), 13.0 (C-29); HRESI MS *m/z* 469.4098 (calcd for $C_{32}H_{53}O$ 469.4096).

Theonellasterone (**69**). To the solution of theonellasterol (**38**) (100 mg, 0.23 mmol) in dichloromethane (5 mL) was added pyridinium chlorochromate (99 mg, 0.46 mmol). The reaction mixture was stirred at room temperature for 3 h, and then dichloromethane and water were added. The separated aqueous phase was extracted with dichloromethane (3 × 30 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and evaporated to dryness. The brown oily residue was passed through a short column of silica gel (10 g) and eluted with CH_2Cl_2 to give **69** (95 mg, quantitative yield) as an amorphous solid. $[\alpha]_{25}^{D}$ = +2.2 (*c* 0.07, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 6.04 (1H, s, H-30a), 4.86 (1H, s, H-30b), 2.39 (1H, dd, *J* = 2.1, 13.8 Hz, H-2a), 1.06 (3H, d, *J* = 6.5 Hz, H_3-21), 0.95 (3H, t, *J* = 7.3 Hz, H_3-29), 0.92 (3H, d, *J* = 6.6 Hz, H_3-26), 0.90 (3H, s, H3-18), 0.90 (3H, d, *J* = 6.6 Hz, H_3-27), 0.57 (3H, s, H_3-19); ^{13}C NMR (100 MHz C_6D_6): δ 200.9 (C-3), 150.3 (C-4), 143.7 (C-14), 127.4 (C-8), 118.2 (C-30), 57.6 (C-17), 50.0 (C-9), 49.0 (C-5), 46.9 (C-24), 43.4 (C-13), 40.0 (C-10), 38.0 (C-12), 37.5 (C-1), 35.9 (C-2), 35.7 (C-20), 34.6 (C-22), 34.5 (C-6), 29.8 (C-25), 27.9 (C-7), 27.1 (C-3), 26.6 (C-16), 25.3 (C-15), 23.8 (C-28), 20.9 (C-11), 20.1 (C-27), 19.9 (C-21), 19.6 (C-26), 18.8 (C-18), 12.9 (C-29), 12.8 (C-19). HRESI MS *m/z* 425.3795 (calcd for $C_{30}H_{49}O$ 425.3783).

(24S)-24-Ethyl-4β-methyl-5α-cholestan-3β-ol (**70**). An oven-dried 50 mL flask was charged with 10% platinum on carbon (20 mg) and theonellasterol (**38**) (100 mg, 0.23 mmol) and the flask was evacuated and flushed with argon. Absolute methanol (10 mL) and dry THF (10 mL) were added, and the flask was flushed with hydrogen. The reaction was stirred at room temperature under H_2 for 5 h. The mixture was filtered through Celite, and the recovered filtrate was concentrated to give 85 mg of pure **70** as an amorphous solid (86%). $[\alpha]_{25}^{D}$ = +6.7 (*c* 0.25, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 3.54 (1H, m, H-3), 1.05 (3H, d, *J* = 6.2 Hz, H_3-21), 0.95 (3H, d, ovl, H_3-30), 0.94 (3H, s, H_3-18), 0.94 (3H, t, ovl, H_3-29), 0.91 (3H, d, *J* = 7.0 Hz, H_3-26), 0.89 (3H, d, *J* = 7.0 Hz, H_3-27), 0.75 (3H, s, H_3-19). ^{13}C NMR (100 MHz C_6D_6): δ 143.8 (C-14), 127.0 (C-8), 74.1 (C-3), 57.7 (C-17), 51.7 (C-9), 48.6 (C-5), 46.9 (C-24), 43.8 (C-13), 40.9 (C-4), 40.5 (C-10), 38.1 (C-12), 37.9 (C-1), 37.8 (C-2), 35.7 (C-20), 34.6 (C-22), 30.8 (C-6), 29.7 (C-25), 27.9 (C-7), 27.1 (C-23), 26.6 (2C, C-15 and C-16), 23.8 (C-28), 20.1 (2C, C-11 and C-27), 19.9 (C-21), 19.6 (C-26), 19.1 (C-18), 15.5 (C-30), 15.4 (C-19), 12.9 (C-29); HRESI MS *m/z* 429.4076 (calcd for $C_{30}H_{53}O$ 429.4096).

(24S)-24-Ethyl-4β-methyl-5α-cholestan-3-one (**71**) *and (24S)-24-Ethyl-4α-methyl-5α-cholestan-3-one* (**72**). An oven-dried 50 mL flask was charged with 10% platinum on carbon (20 mg) and theonellasterone (100 mg, 0.23 mmol) and the flask was evacuated and flushed with argon. Absolute methanol (10 mL) and dry THF (10 mL) were added, and the flask was flushed with

hydrogen. The reaction was stirred at room temperature under H_2 for 5 h. The mixture was filtered through Celite, and the recovered filtrate was concentrated. The mixture was purified by HPLC on a Nucleodur Isis 100-5 C18 (5 μm; 4.5 mm internal diameter × 250 mm) with MeOH:H_2O (999.5:0.5) as eluent (flow rate 1 mL/min) to give 39 mg (40% from 5) of 71 (*t*R = 55 min) and 30 mg (31% from 5) of 72 (*t*R = 60 min) as amorphous solids. *(24S)-24-Ethyl-4β-methyl-5α-cholestan-3-one* **71**. $[\alpha]_{25}^{D}$ = −0.7 (*c* 0.06, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 2.38 (1H, m, H-4), 1.05 (3H, d, *J* = 6.8 Hz, H_3-21), 0.95 (3H, t, *J* = 7.0 Hz, H_3-29), 0.95 (3H, d, *J* = 7.0 Hz, H_3-30), 0.93 (3H, s, H_3-18), 0.91 (3H, d, *J* = 7.0 Hz, H_3-26), 0.90 (3H, d, *J* = 7.0 Hz, H_3-27), 0.73 (3H, s, H_3-19). ^{13}C NMR (100 MHz C_6D_6): δ 196.0 (C-3), 143.4 (C-14), 126.6 (C-8), 57.6 (C-17), 50.8 (C-9), 49.6 (C-5), 48.7 (C-4), 46.9 (C-24), 43.4 (C-13), 39.0 (C-10), 38.3 (C-12), 37.9 (C-1), 37.7 (C-2), 35.6 (C-20), 34.6 (C-22), 30.4 (C-6), 29.7 (C-25), 27.8 (C-7), 27.0 (C-23), 26.6 (C-16), 26.5 (C-15), 23.8 (C-28), 20.1 (C-27), 19.9 (C-21), 19.7 (C-11), 19.6 (C-26), 19.0 (C-18), 14.6 (C-19), 14.5 (C-30), 12.9 (C-29); HRESI MS *m/z* 427.3936 (calcd for $C_{30}H_{51}O$ 427.3940). *(24S)-24-Ethyl-4α-methyl-5α-cholestan-3-one* **72**. $[\alpha]_{25}^{D}$ = −12.3 (*c* 0.10, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 2.05 (1H, dq, *J* = 6.1, 14.2 Hz, H-4), 1.09 (3H, d, *J* = 6.1 Hz, H_3-30), 1.06 (3H, d, *J* = 6.7 Hz, H_3-21), 0.94 (3H, s, H_3-18), 0.94 (3H, t, *J* = 7.4 Hz, H_3-29), 0.91 (3H, d, *J* = 6.7 Hz, H_3-26), 0.89 (3H, d, *J* = 6.7 Hz, H_3-27), 0.70 (3H, s, H_3-19). ^{13}C NMR (100 MHz C_6D_6): δ 199.1 (C-3), 143.2 (C-14), 126.4 (C-8), 57.5 (C-17), 53.8 (C-9), 49.5 (C-5), 46.9 (C-24), 45.5 (C-4), 43.4 (C-13), 39.0 (C-10), 38.5 (C-12), 38.1 (2C, C-1 and C-2), 35.7 (C-20), 34.6 (C-22), 30.1 (C-6), 29.7 (C-25), 27.9 (C-7), 27.0 (C-23), 26.8 (C-15), 26.6 (C-16), 23.8 (C-28), 20.7 (C-11), 20.1 (C-27), 19.9 (C-21), 19.6 (C-26), 18.9 (C-18), 13.4 (C-30), 12.9 (C-29), 12.3 (C-19); HRESI MS *m/z* 427.3944 (calcd for $C_{30}H_{51}O$ 427.3940).

(24S)-24-Ethyl-4α-methyl-5α-cholestan-3β-ol (**73**) *and (24S)-24-Ethyl-4α-methyl-5α-cholestan-3α-ol* (**74**). To a solution of **72** (30 mg, 0.070 mmol) in dry methanol (5 mL) was added $NaBH_4$ (13 mg, 0.35 mmol) at 0 °C. After 30 min the reaction was quenched by addition of MeOH (3 mL) and then concentrated *in vacuo*. Ethyl acetate and water were added and the separated aqueous phase was extracted with ethyl acetate (3 × 30 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and concentrated. The mixture was purified by HPLC on a Nucleodur Isis 100-5 C18 (5 μm; 4.5 mm internal diameter × 250 mm) with MeOH:H_2O (999.5:0.5) as eluent (flow rate 1 mL/min) to give 15 mg (50% from **72**) of **73** (*t*R = 47.5 min) and 10 mg of **74** (34% from **72**) (*t*R = 50 min) as amorphous solids. *(24S)-24-Ethyl-4α-methyl-5α-cholestan-3β-ol* (**73**). $[\alpha]_{25}^{D}$ = −2.1 (*c* 0.02, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 2.93 (1H, m, H-3), 2.01 (1H, dt, *J* = 6.0, 12.6 Hz, H-4), 1.06 (3H, d, *J* = 6.5 Hz, H-21), 1.02 (3H, d, *J* = 6.0 Hz, H-30), 0.95 (3H, s, H-18), 0.94 (3H, t, *J* = 7.3 Hz, H-29), 0.91 (3H, d, *J* = 7.3 Hz, H-26), 0.90 (3H, d, *J* = 7.3 Hz, H-27), 0.71 (3H, s, H-19). HRESI MS *m/z* 429.4088 (calcd for $C_{30}H_{53}O$ 429.4096). *(24S)-24-Ethyl-4α-methyl-5α-cholestan-3α-ol* (**74**). $[\alpha]_{25}^{D}$= +4.1 (*c* 0.04, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 3.55 (1H, br m, H-3), 1.05 (3H, d, *J* = 6.5 Hz, H_3-21), 0.95 (3H, d, *J* = 6.5 Hz, H_3-30), 0.95 (3H, s, H_3-18), 0.94 (3H, t, *J* = 6.8 Hz, H_3-29), 0.91 (3H, d, *J* = 7.0 Hz, H_3-26), 0.89 (3H, d, *J* = 7.0 Hz, H_3-27), 0.75 (3H, s, H_3-19). HRESI MS *m/z* 429.4084 (calcd for $C_{30}H_{53}O$ 429.4096).

(24S)-24-Ethyl-4-methyl-5α-cholest-3-ene (**75**). To a solution of **74** (30 mg, 0.070 mmol) in dry pyridine (5 mL), a solution of tosyl chloride (66 mg, 0.35 mmol) in dry pyridine (5 mL) was added. The solution was stirred at room temperature for 2 h and then concentrated *in vacuo*. The precipitate was re-dissolved in CH_2Cl_2, washed with $NaHCO_3$ saturated solution and water, dried with Na_2SO_4, and then evaporated to dryness to give the 3β-tosylate, which was subjected to the next step without any purification. A solution of 3β-tosylate and CH_3COOK (7.5 mg, 0.077 mmol) dissolved in DMF (3.5 mL) and water (0.5 mL) was refluxed for 2 h. The solution was cooled at room temperature and then ethyl acetate and water were added. The separated aqueous phase was extracted with ethyl acetate (3 × 30 mL). The combined organic phases were washed with water, dried (Na_2SO_4) and evaporated to dryness. Purification by HPLC on a Nucleodur Isis 100-5 C18 (5 μm; 4.5 mm internal diameter × 250 mm) with MeOH:H_2O (999.5:0.5) as eluent (flow rate 1 mL/min) gave **75** (21 mg, 75% over two steps) as an amorphous solid. $[\alpha]_{25}^{D}$ = +43 (*c* 0.01, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 5.44 (1H, br s, H-3), 1.64 (3H, s, H_3-30), 1.06 (3H, d, *J* = 6.4 Hz, H_3-21), 0.96 (3H, s, H_3-18), 0.94 (3H, t, *J* = 7.5 Hz, H_3-29), 0.91 (3H, d, *J* = 6.8 Hz, H_3-26), 0.90 (3H, d, *J* = 6.8 Hz, H_3-27), 0.81 (3H, s, H_3-19); HRESI MS *m/z* 411.3985 (calcd for $C_{30}H_{51}$ 411.3991).

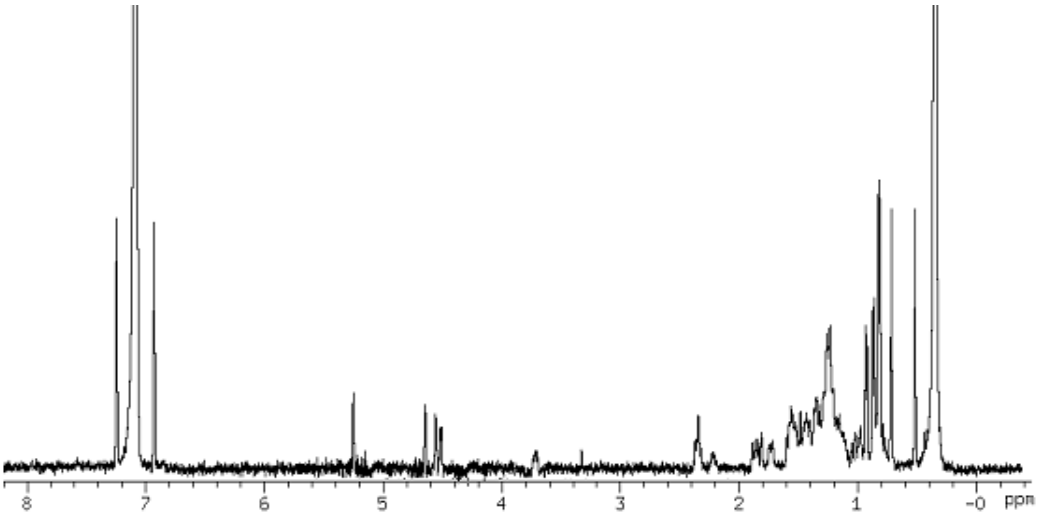
(24S)-24-Ethyl-3β-hydroxyl-5α-cholest-4-one (**76**). A stream of O_3 was bubbled into CH_2Cl_2 (5 mL) at −78 °C until a blue-colored solution resulted. A portion of this solution (4 mL) was added to a solution of **38** (10 mg, 0.023 mmol) in CH_2Cl_2 kept under argon at −78 °C. After stirring for 1 h, excess of ozone was removed upon bubbling with N_2 and the solution was treated with excess dimethylsulfide (2 mL). After 5 h, the solution was concentrated *in vacuo* to remove the solvent and the mixture was purified by HPLC on a Nucleodur Isis 100-5 C18 (5 μm; 4.5 mm internal diameter × 250 mm) with MeOH:H_2O (999.5:0.5) as eluent (flow rate 1 mL/min) to give 8.4 mg (84%) of **76** (*t*R = 27.5 min) as an amorphous solid. $[\alpha]_{25}^{D}$ = +3.2 (*c* 0.37, CH_3OH); selected ^{1}H NMR (400 MHz C_6D_6): δ 3.80 (1H, m, H-3), 1.06 (3H, d, *J* = 6.4 Hz, H_3-21), 0.94 (3H, t, *J* = 7.4 Hz, H_3-29), 0.91 (3H, d, *J* = 7.0 Hz, H_3-26), 0.90 (3H, d, *J* = 7.0 Hz, H_3-27), 0.83 (3H, s, H_3-18), 0.45 (3H, s, H_3-19); ^{13}C NMR (100 MHz C_6D_6): δ 212.1 (C-4), 144.3 (C-14), 125.3 (C-8), 75.2 (C-3), 57.6 (C-17), 56.6 (C-9), 49.6 (C-5), 47.0 (C-24), 44.3 (C-10), 43.4 (C-13), 38.0 (C-1), 35.7 (C-20), 35.1 (C-12), 34.5 (C-22), 33.4 (C-2), 29.8 (C-25), 28.8 (C-7), 27.8 (C-15), 27.1 (C-23), 26.6 (C-16), 23.8 (C-28), 21.5 (C-6), 20.9 (C-11), 20.1 (C-27), 19.9 (C-26), 19.6 (C-21), 18.8 (C-18), 14.4 (C-19), 12.9 (C-29); HRESI MS *m/z* 429.3741 (calcd for $C_{29}H_{49}O_2$ 429.3733).

(24S)-24-Ethyl-5α-cholestan-3β,4β-diol (**77**). A stream of O_3 was bubbled into CH_2Cl_2 (5 mL) at −78°C until a blue-colored solution resulted. A portion of this solution (4 mL) was added to a solution of theonellasterol (10 mg, 0.023 mmol) in CH_2Cl_2 kept under argon at −78 °C. After stirring for 1 h, excess O_3 was removed upon bubbling with N_2. To the solution was added methanol (2 mL) and then treated with an excess of $NaBH_4$. The solution was stirred at room temperature for 3 h and then concentrated *in vacuo*. The precipitate was re-dissolved in ethyl acetate, washed with water, dried with Na_2SO_4 and then evaporated to dryness. The mixture was purified by HPLC on a Nucleodur Isis 100-5 C18 (5 μm; 4.5 mm internal diameter × 250 mm)

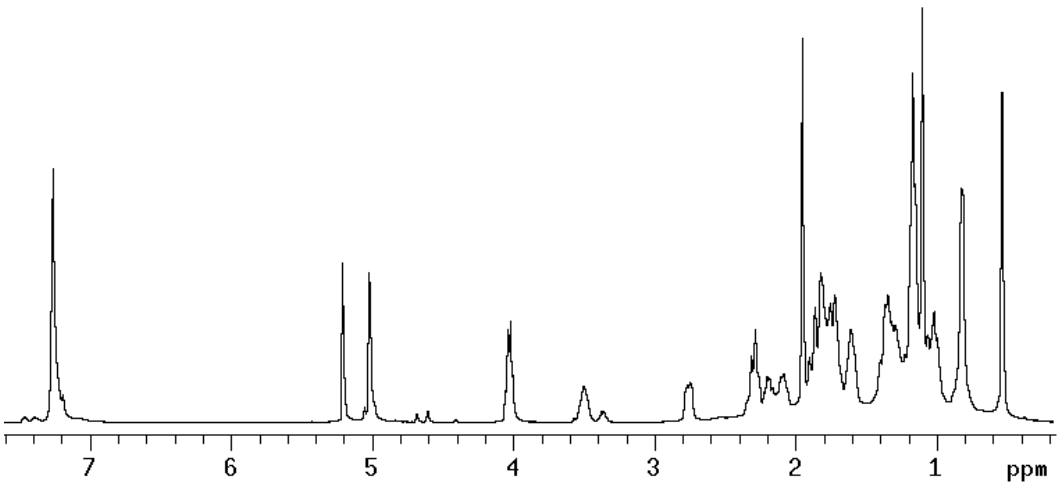
with MeOH:H_2O (999.5:0.5) as eluent (flow rate 1 mL/min) to give 9.2 mg (93%) of **77** (*t*R = 31.5 min) as a amorphous solid. $[\alpha]_{25}^{D}$ = +6.9 (*c* 0.5, CH_3OH); selected ^{1}H NMR (400 MHz, C_6D_6): δ 3.52 (1H, br s, H-4), 3.28 (1H, dd, *J* = 5.2, 10.6 Hz, H-3), 1.12 (3H, s, H-19), 1.05 (3H, d, *J* = 6.8 Hz, H-21), 0.94 (3H, s, H-18), 0.93 (3H, t, *J* = 7.5 Hz, H-29), 0.90 (3H, d, *J* = 7.2 Hz, H-26), 0.89 (3H, d, *J* = 7.2 Hz, H-27); ^{13}C NMR (100 MHz C_6D_6): δ 143.1 (C-14), 127.3 (C-8), 74.8 (C-3), 72.7 (C-4), 57.6 (C-17), 51.0 (C-9), 49.0 (C-5), 46.9 (C-24), 43.6 (C-13), 38.3 (C-10), 38.2 (C-12), 37.2 (C-1), 35.8 (C-20), 34.6 (C-22), 30.6 (C-6), 29.7 (C-25), 27.9 (C-7), 27.0 (C-23), 26.6 (3C, C-2, C-15 and C-16), 23.8 (C-28), 20.1 (C-11), 20.0 (C-27), 19.9 (C-21), 19.6 (C-26), 19.0 (C-18), 15.4 (C-19), 13.0 (C-29); HRESI MS *m/z* 431.3883 (calcd for $C_{29}H_{51}O_2$ 431.3889).

Spectroscopic Data

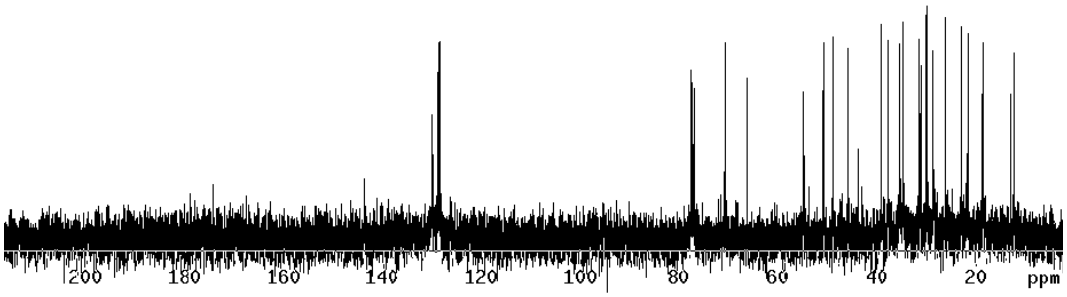
^{1}H NMR spectrum of Conicasterol E (**60**) in C_6D_6 at 500 MHz.



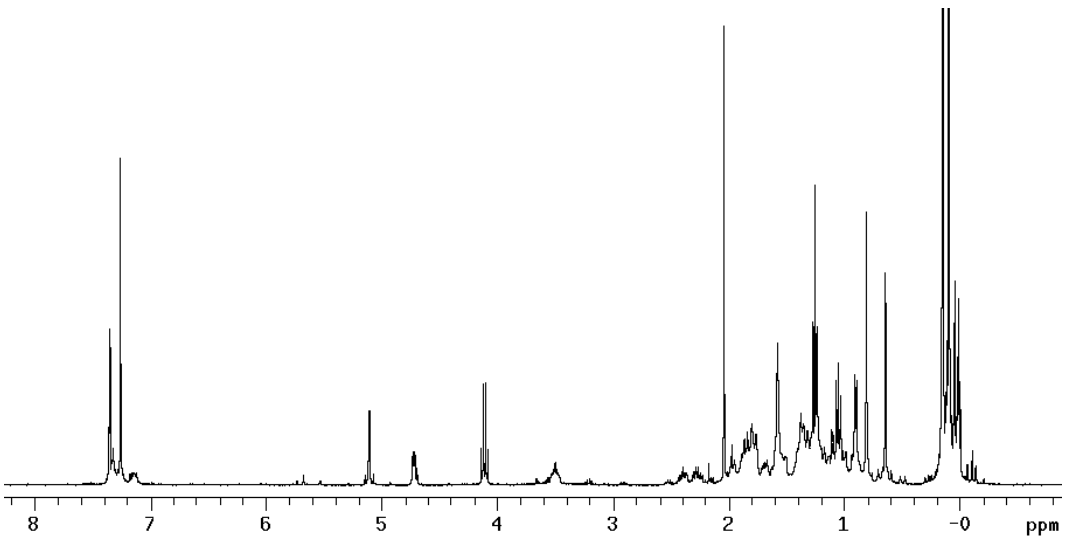
^{1}H NMR spectrum of Compound **63** in $CDCl_3$ at 400 MHz.



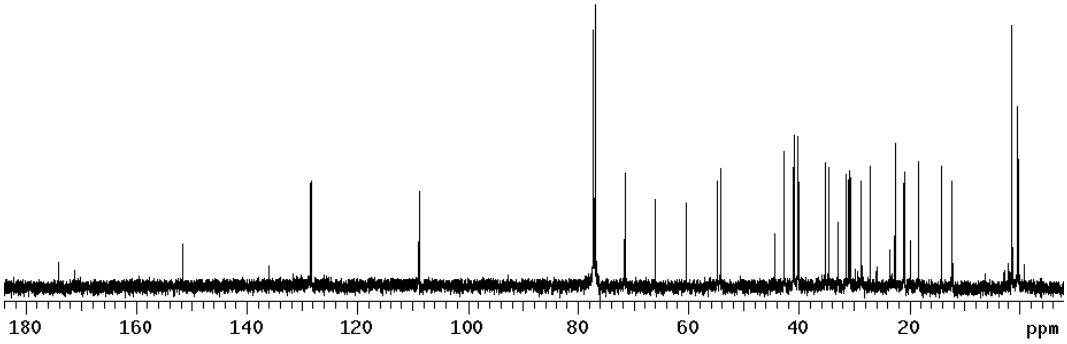
^{13}C NMR spectrum of Compound **63** in $CDCl_3$ at 100 MHz.



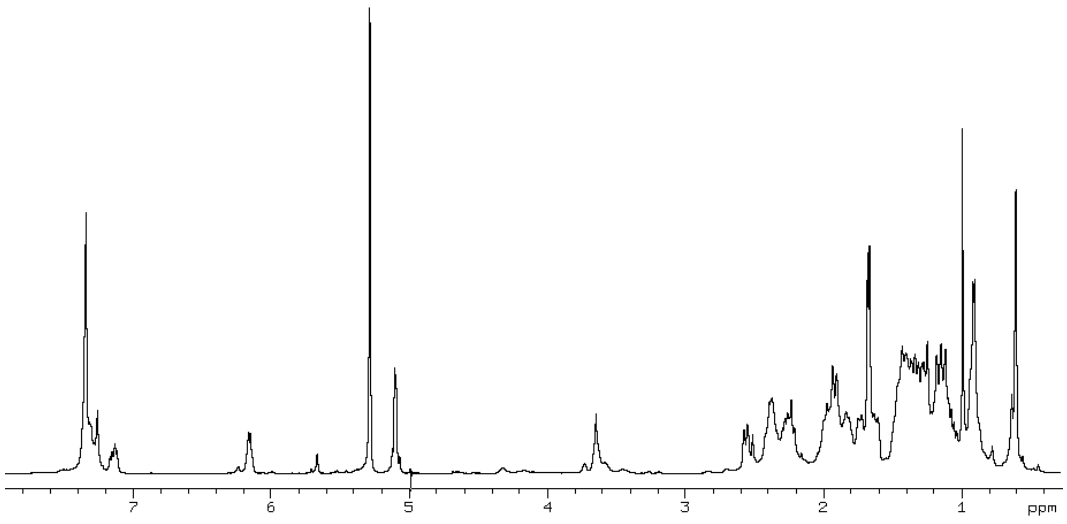
^{1}H NMR spectrum of Compound **64** in $CDCl_3$ at 400 MHz.



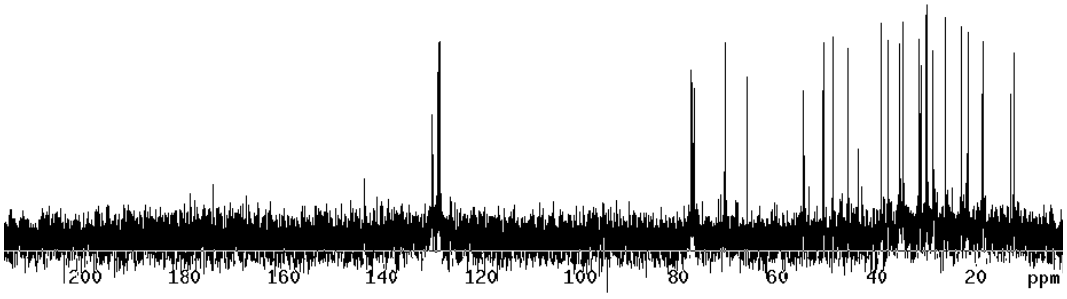
^{13}C NMR spectrum of Compound **64** in $CDCl_3$ at 100 MHz.



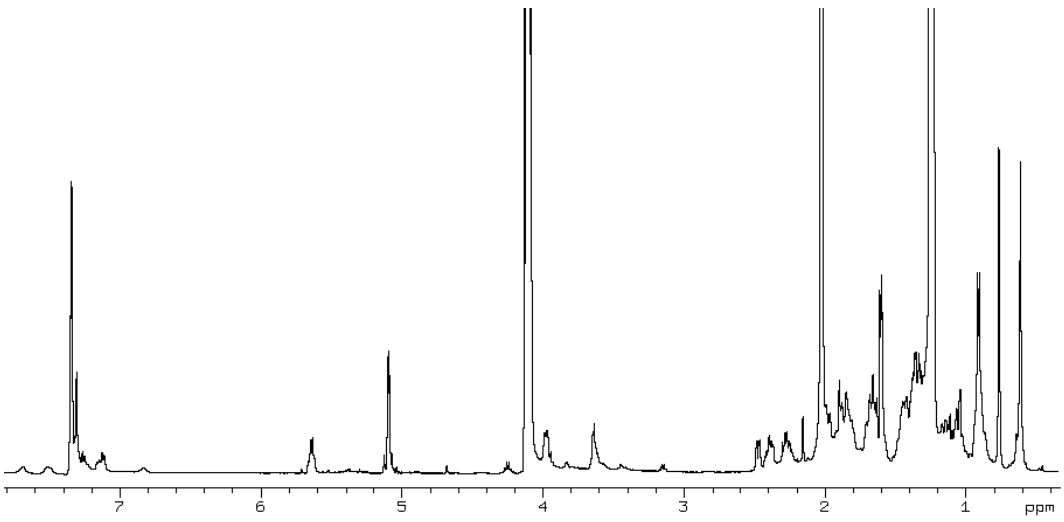
^{1}H NMR spectrum of Compound **65** in $CDCl_3$ at 400 MHz.



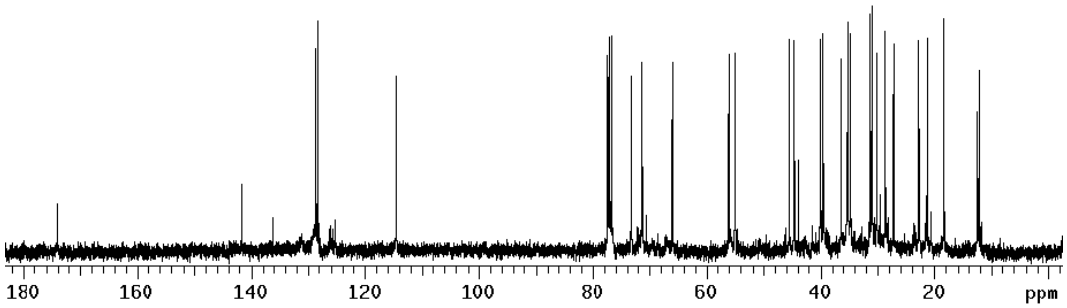
^{13}C NMR spectrum of Compound **65** in $CDCl_3$ at 100 MHz.



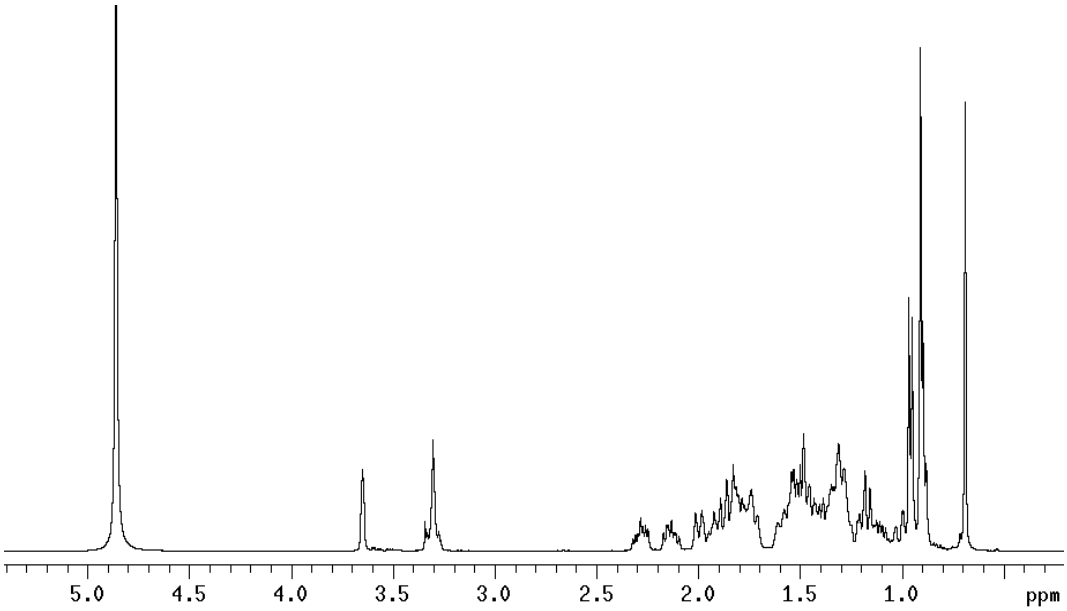
^{1}H NMR spectrum of Compound **66** in $CDCl_3$ at 400 MHz.



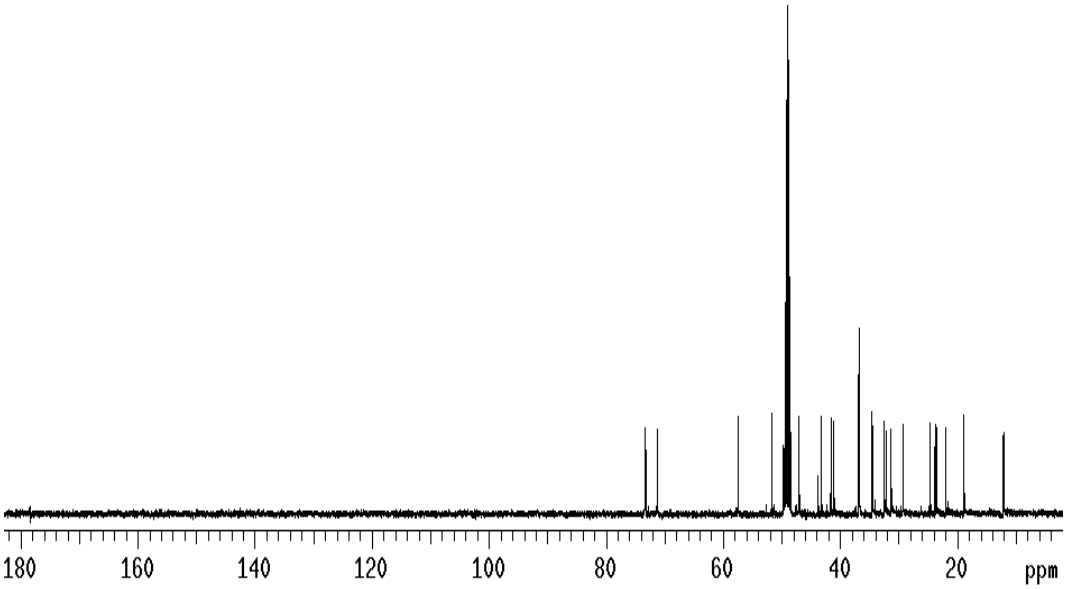
^{13}C NMR spectrum of Compound **66** in $CDCl_3$ at 100 MHz.



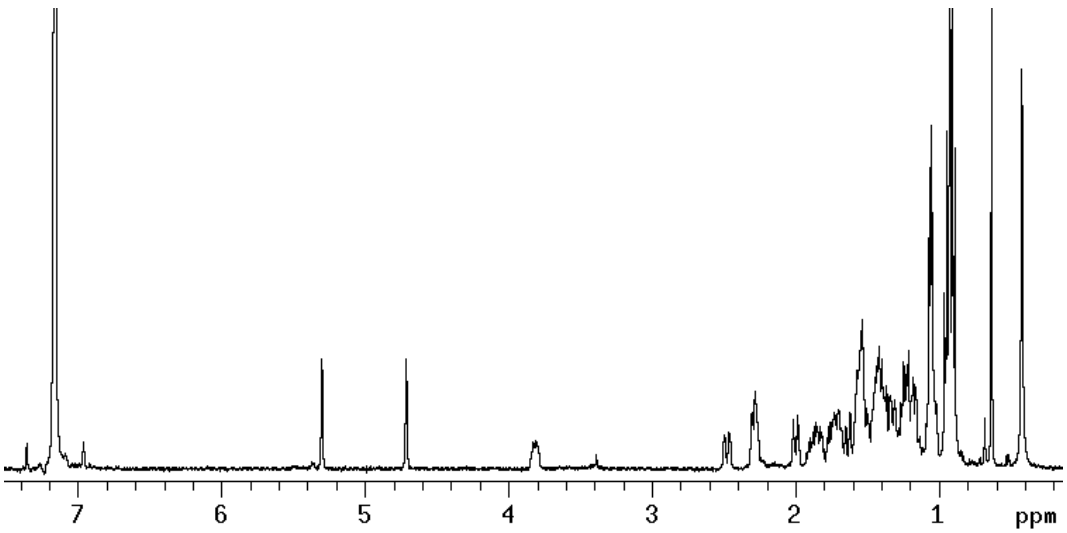
^{1}H NMR spectrum of 6-ECDCA (**61**) in CD_3OD at 400 MHz.



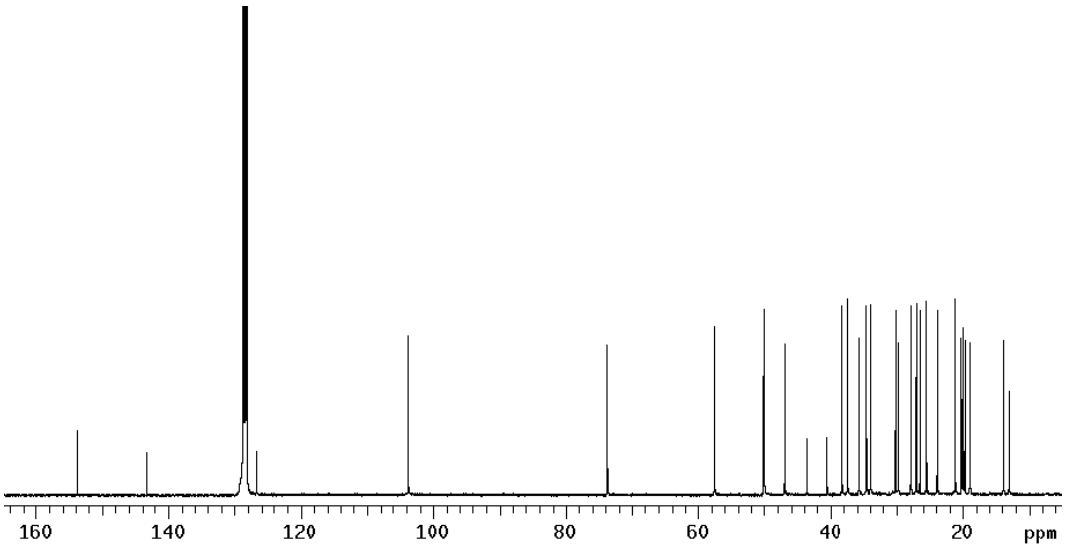
^{13}C NMR spectrum of 6-ECDCA (**61**) in CD_3OD at 100 MHz.



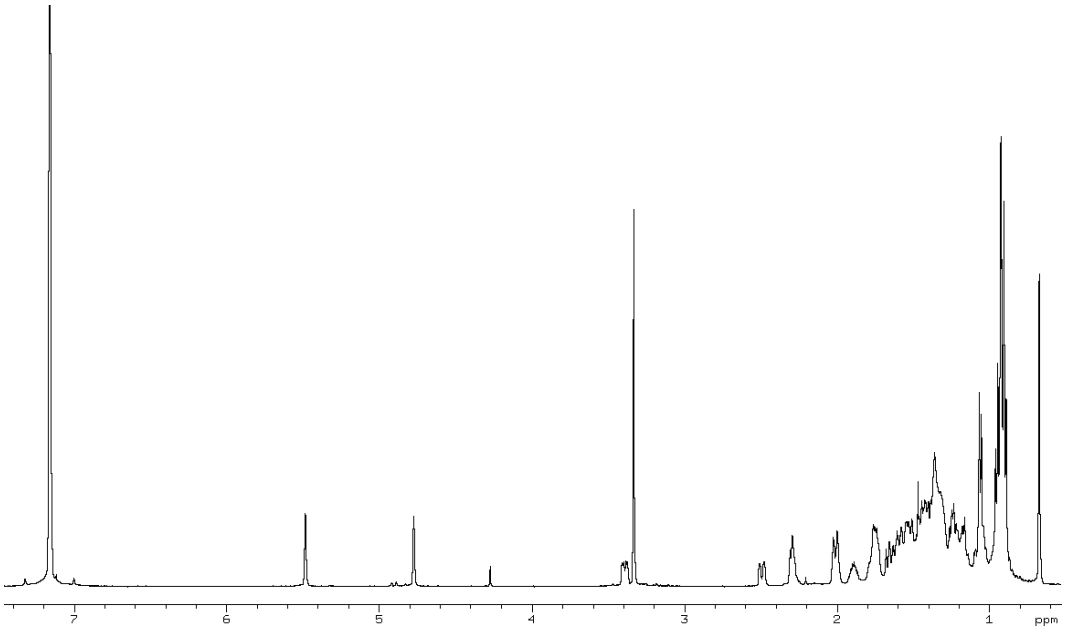
^{1}H NMR of Theonellasterol (**38**) in C_6D_6 at at 500 MHz.



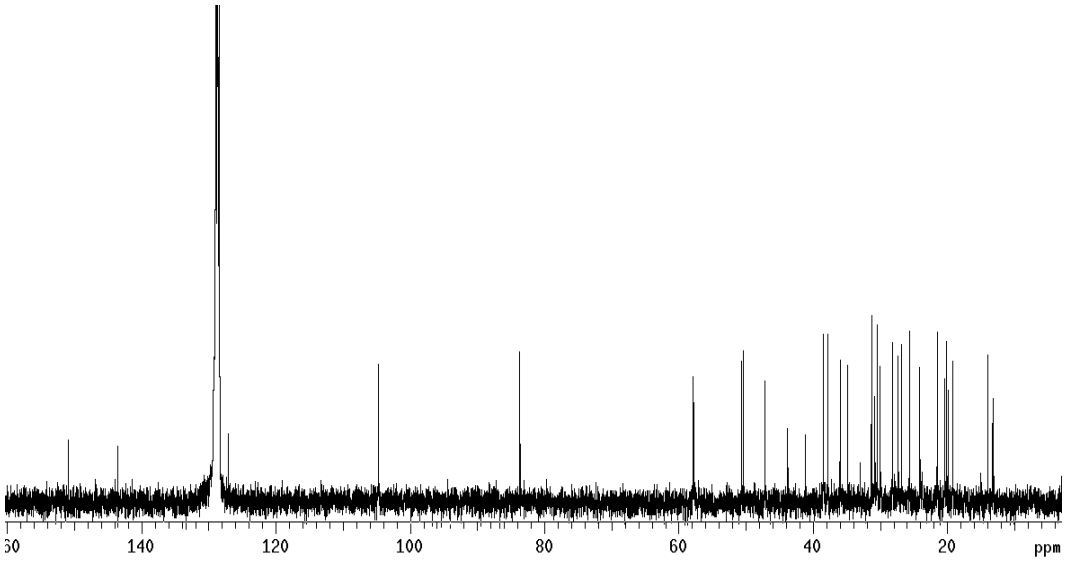
^{13} C NMR of Theonellasterol (**38**) in C_6D_6 at at 100 MHz.



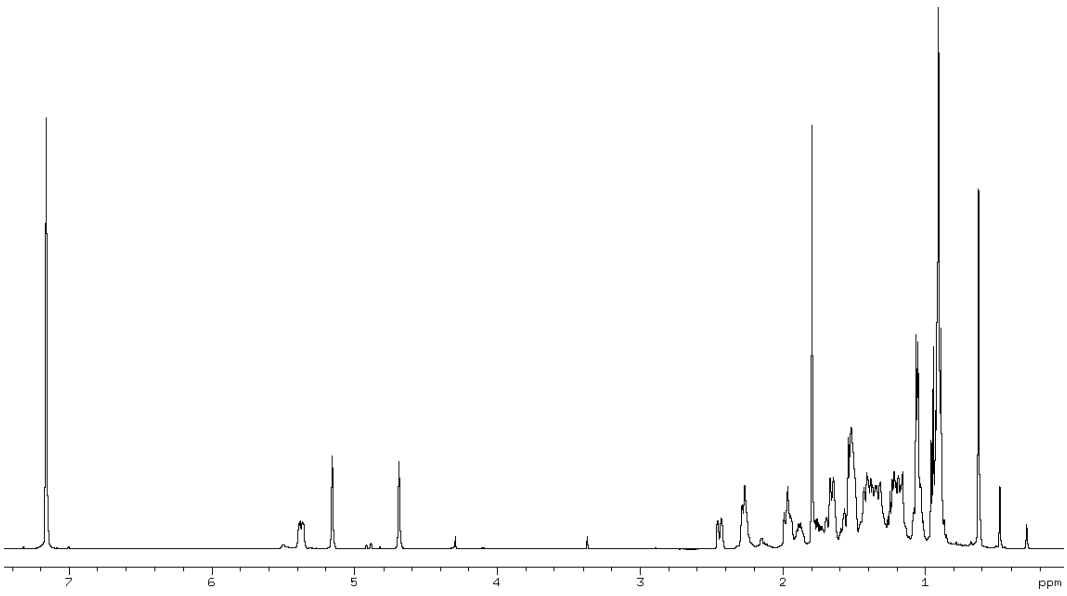
^{1}H NMR of Compound (**67**) in C_6D_6 at at 400 MHz.



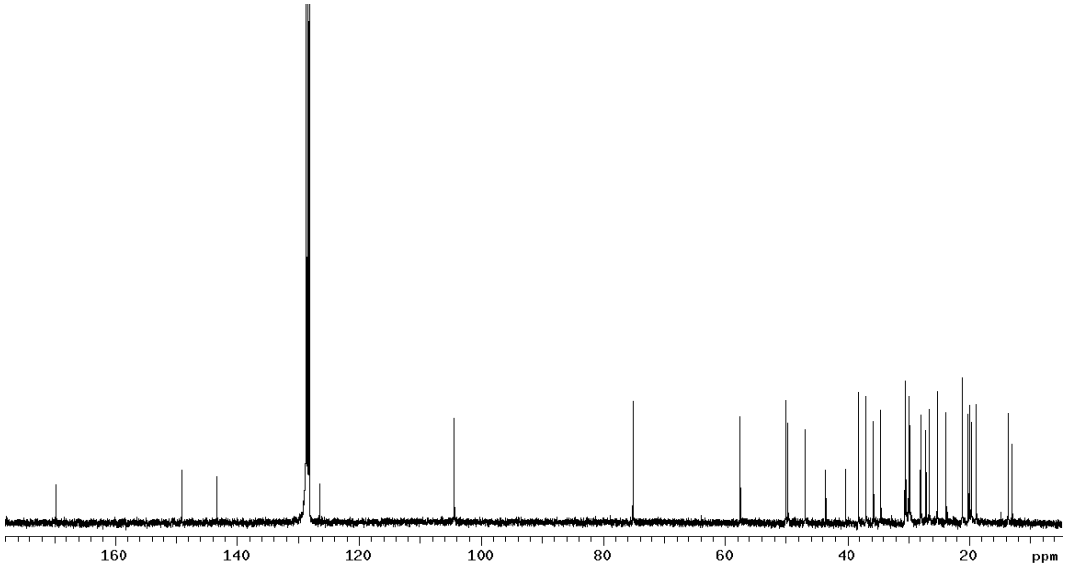
^{13}C NMR of Compound (**67**) in C_6D_6 at at 100 MHz.



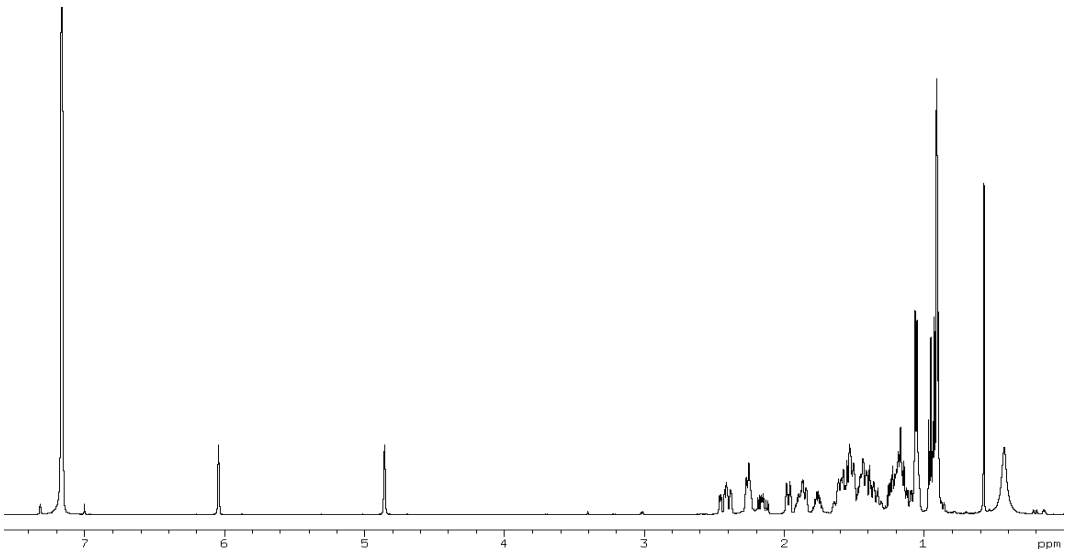
^{1}H NMR of Compound (**68**) in C_6D_6 at at 400 MHz.



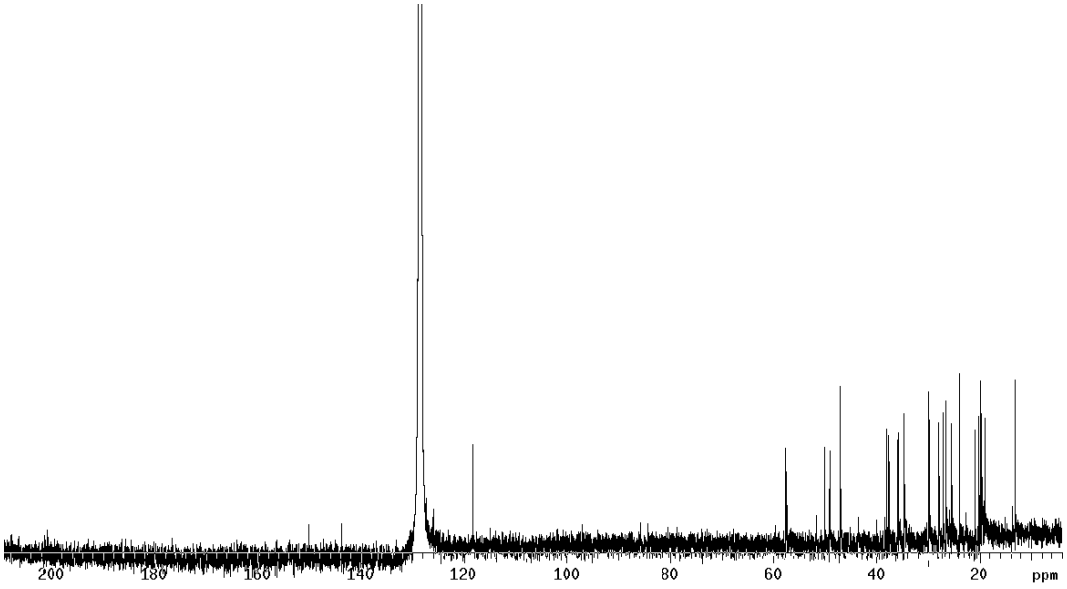
^{13}C NMR of Compound (**68**) in C_6D_6 at at 100 MHz.



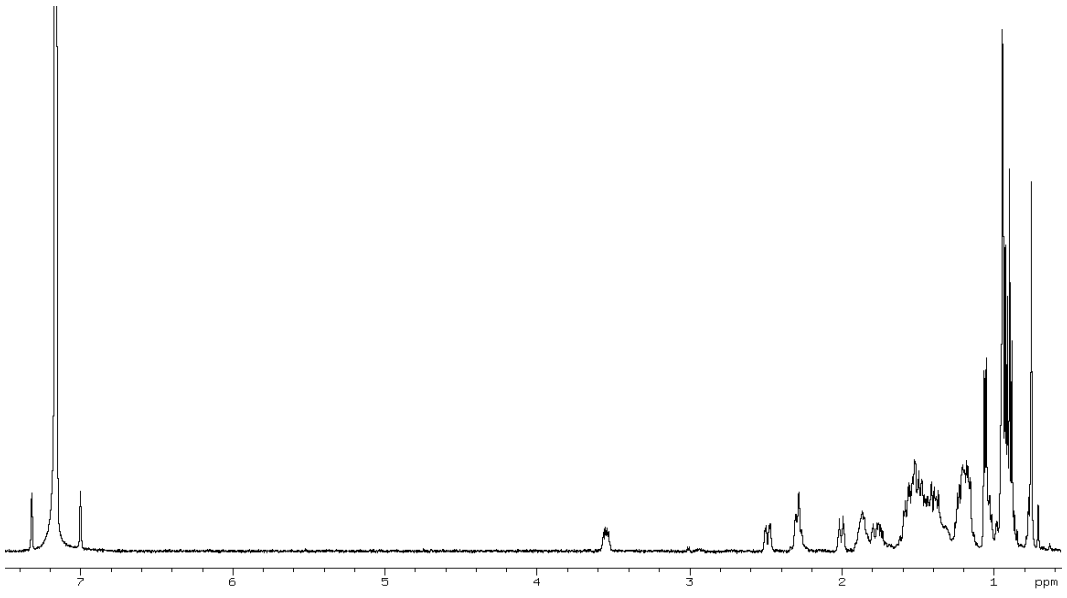
^{1}H NMR of Compound (**69**) in C_6D_6 at at 400 MHz.



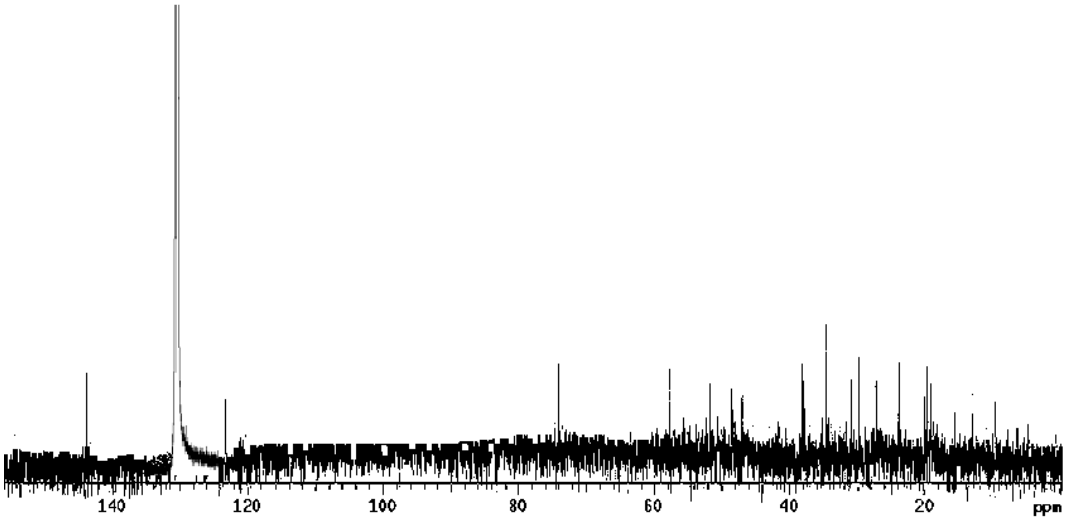
^{13}C NMR of Compound (**69**) in C_6D_6 at at 100 MHz.



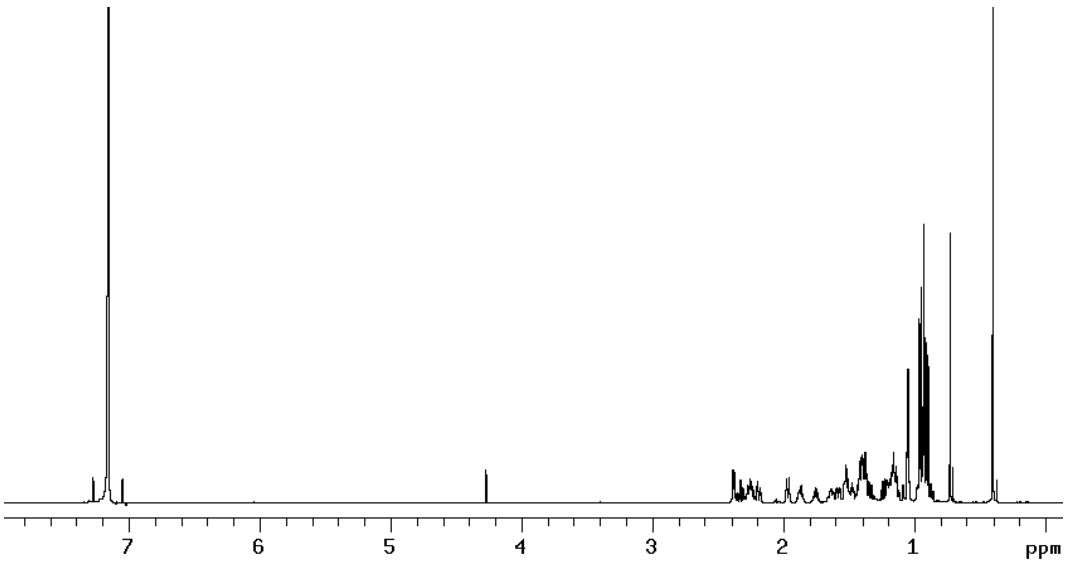
^{1}H NMR of Compound (**70**) in C_6D_6 at at 400 MHz.



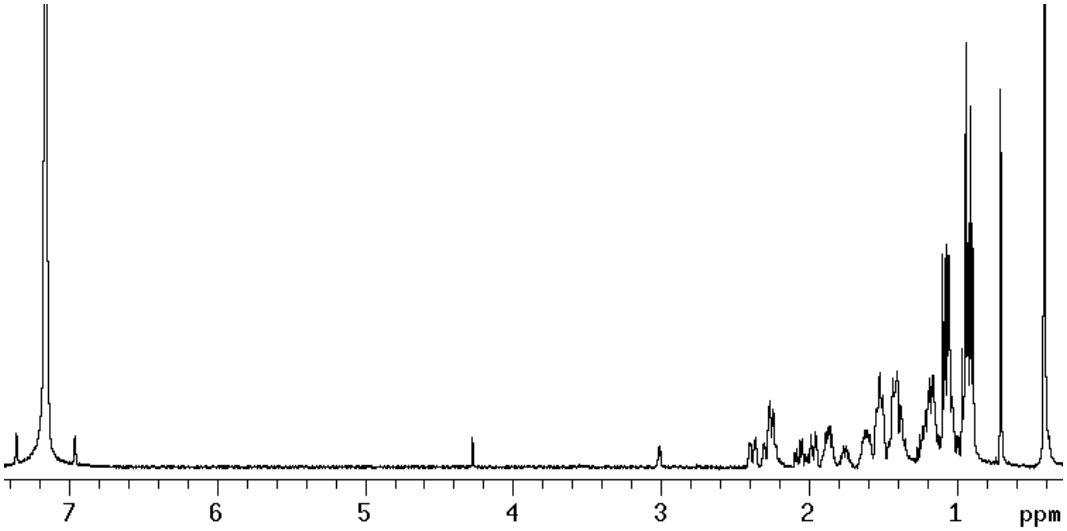
^{13}C NMR of Compound (**70**) in C_6D_6 at at 100 MHz.



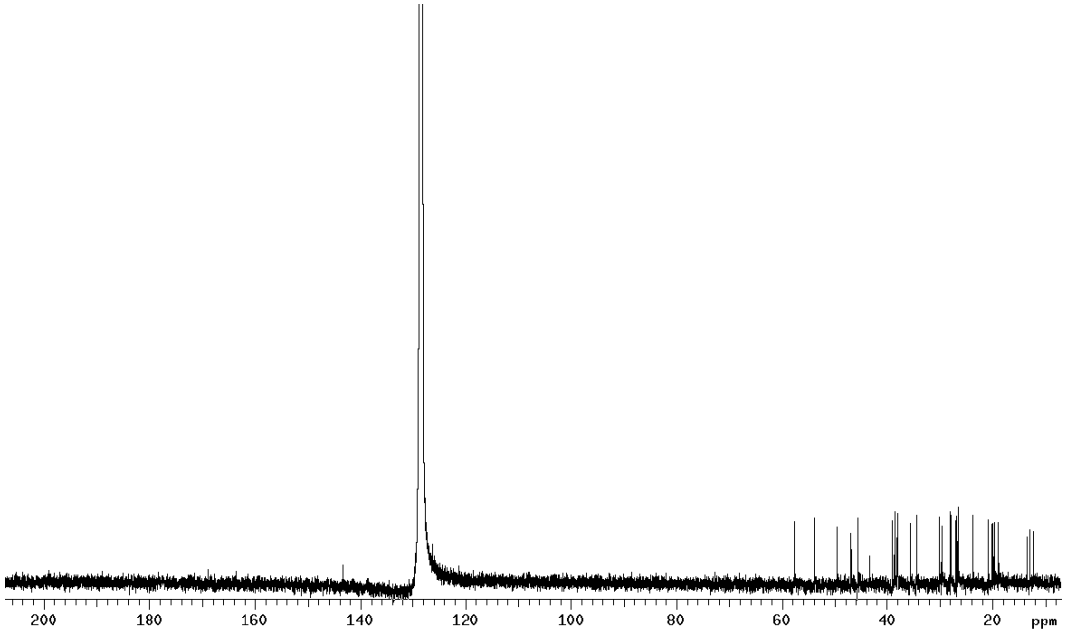
^{1}H NMR of Compound (**71**) in C_6D_6 at at 400 MHz.



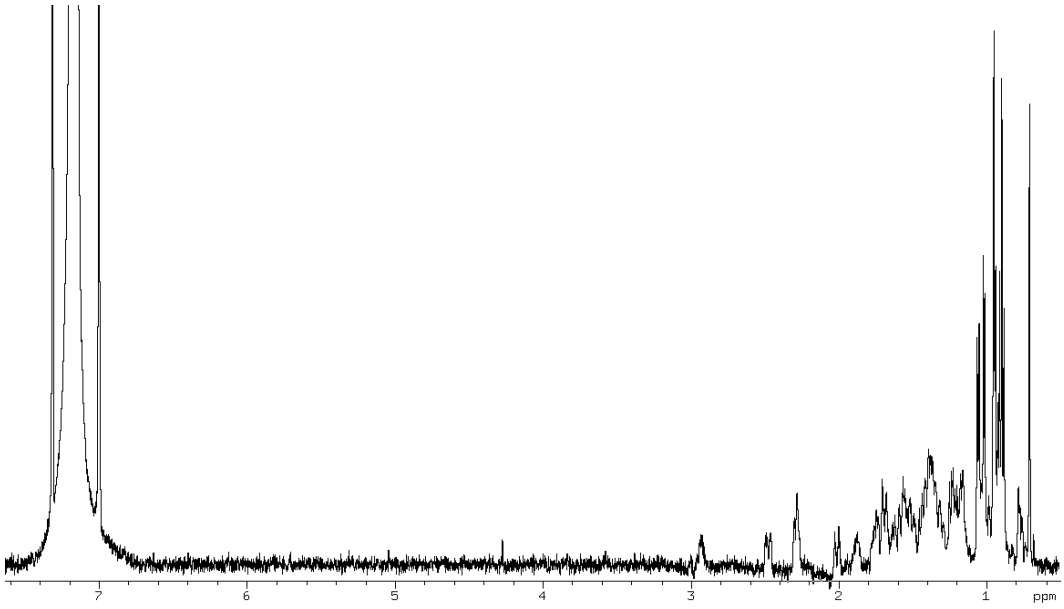
^{1}H NMR of Compound (**72**) in C_6D_6 at at 400 MHz.



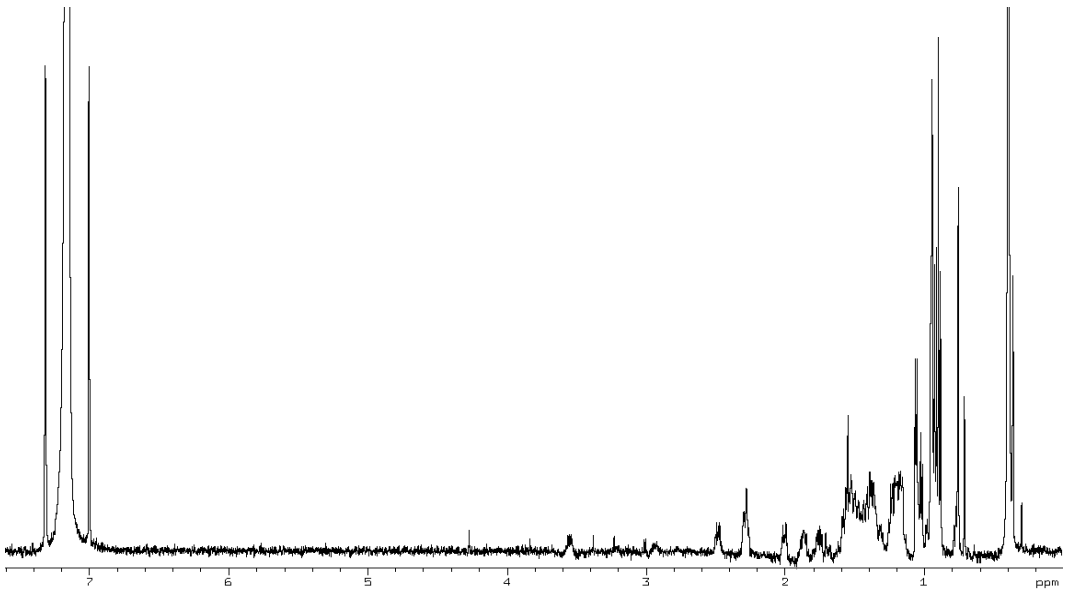
^{13}C NMR of Compound (**72**) in C_6D_6 at at 100 MHz.



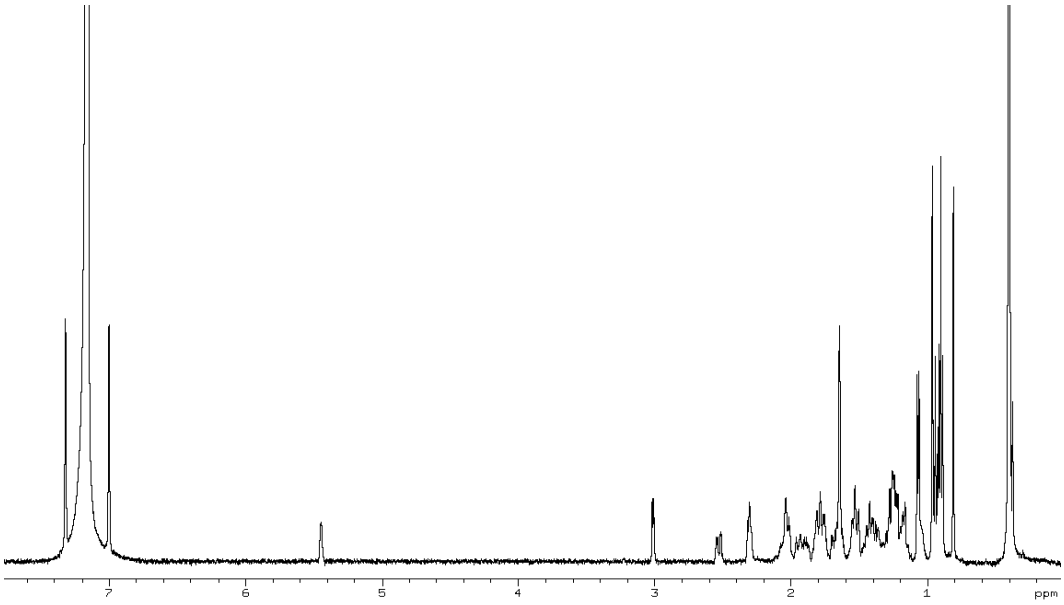
^{1}H NMR of Compound (**73**) in C_6D_6 at at 400 MHz.



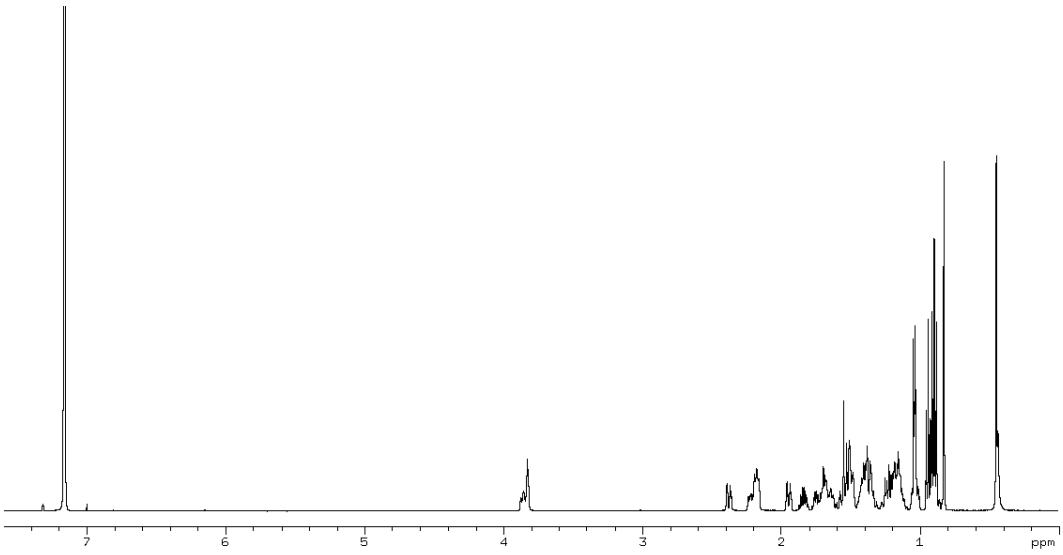
^{1}H NMR of Compound (**74**) in C_6D_6 at at 400 MHz.



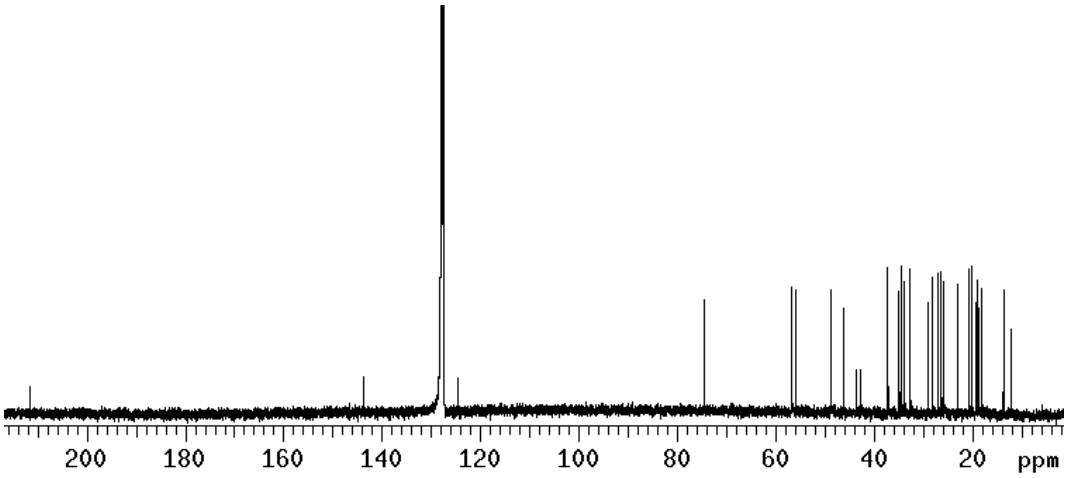
^{1}H NMR of Compound (**75**) in C_6D_6 at at 400 MHz.



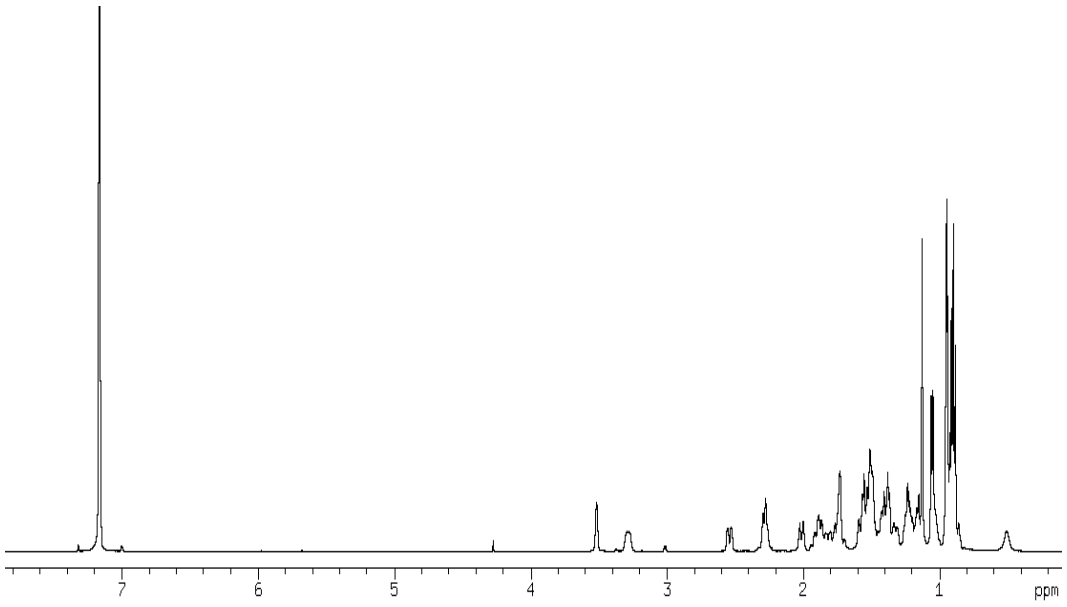
^{1}H NMR of Compound (**76**) in C_6D_6 at at 400 MHz.



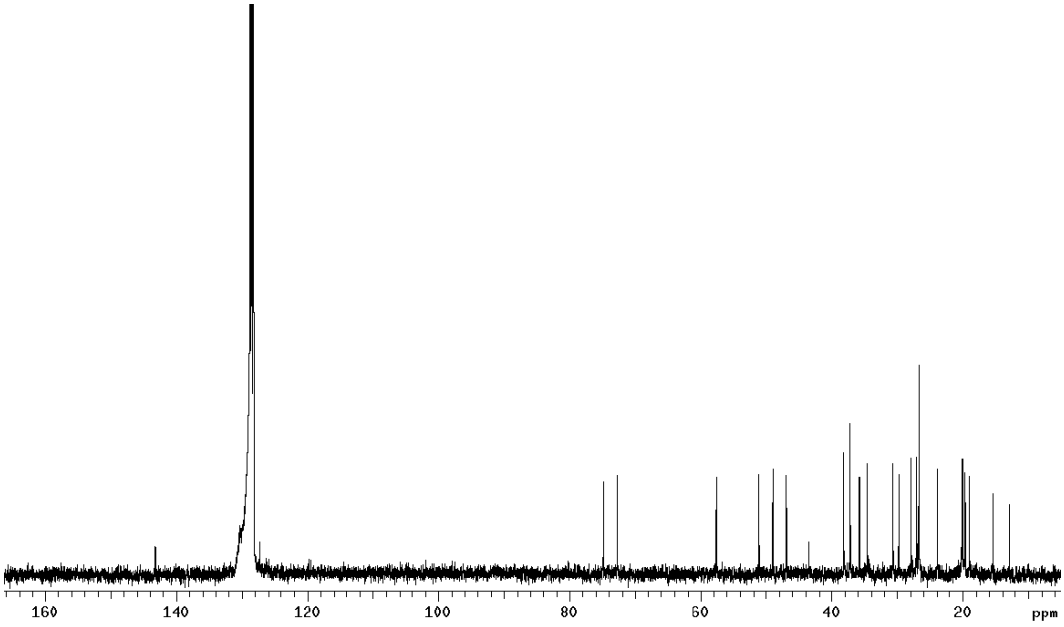
^{13}C NMR of Compound (**76**) in C_6D_6 at at 100 MHz.



^{1}H NMR of Compound (**77**) in C_6D_6 at at 400 MHz.



^{13}C NMR of Compound (**77**) in C_6D_6 at at 100 MHz.



V. Experimental section of stereochemical studies of perthamide C

General experimental procedures.

Specific rotations were measured on a Jasco P-2000 polarimeter. High-resolution ESI-MS spectra were performed with a Micromass Q-TOF mass spectrometer. ESI-MS experiments were performed on a Applied Biosystem API 2000 triple-quadrupole mass spectrometer. HPLC was performed using a Waters Model 6000-A pump equipped with U6 K injector and a differential refractometer, model 401. NMR spectra were obtained on a Varian Mercury-400 and Inova-500 NMR spectrometers, δ in parts per million, *J* in hertz, spectra referred to CD_2HOD or $CHCl_3$ as internal standards (δH=3.31 and 7.26, respectively). For an accurate measurement of the coupling constants, NMR spectra were recorded on Varian 700 MHz and Bruker DRX-600 spectrometers, equipped with cryo-probe. All spectra were acquired in the phase-sensitive mode and the TPPI method was used for quadrature detection in the ω1 dimension. NMR sample was obtained dissolving 20 mg of perthamide C in DMSO-d_6. $^3J_{H\text{-}H}$ values were extracted from 1D ^{1}H NMR and 2D E.COSY spectra. For the E.COSY spectrum 32 scans for t_1 value were acquired with a t_{1max} of 65 ms. $^{2,3}J_{C\text{-}H}$ values were obtained from phasesensitive PFG-PS-HMBC spectra and ^{13}C coupled and decoupled HSQC-TOCSYaccording the following conditions. The PFG-PS-HMBC spectrum was recorded using 2 K points in ω2, setting the delay for long-range coupling evolution (Δ) at 50 ms, with 32 scans/t_1 (t_{1max} 15.2 ms). Zero-filling (8x1 K) was carried out in ω2 and ω1, respectively, to obtain a digital resolution of 0.9 Hz in ω2. The ^{13}C coupled and decoupled HSQC-TOCSY were recorded using 2 K points in ω2, setting the mixing time 80 ms and $^1J_{CH}$ of 140 Hz, with 120 scans/t_1 (t_{1max} 9.7 ms). Zero-filling (8Kx1K) was carried out in ω2 and ω1, respectively, to obtain a digital resolution of 0.5 Hz in ω2; the protone-carbon $^{2,3}J$ coupling were extracted through a computer-aided analysis of the heteronuclear coupled and decoupled multiplets acquired in two separate experiments. The reverse single-quantum heteronuclear correlation (HSQC) spectra were recorded by using a pulse sequence with a Wurst pulse 0.15 s before each scan to suppress the signal originating from protons not directly bound to ^{15}N; the interpulse delays were adjusted for an average $^1J_{NH}$ of 90 Hz. For the phase-sensitive PFG-HETLOC spectrum, a total of 160 scans/t_1 were acquired using 4 K points in ω2, with a spin lock of 50 ms and a t_{1max} of 41.7 ms. The data matrices were zero-filled to 4Kx1k affording a digital resolution of 0.4 Hz in ω2. Through-space ^{1}H connectivities were evidenced using a ROESY experiment with mixing times of 400, 100 and 50 ms, respectively, and NOESYexperiment acquired with 200 and 100 ms mixing times.

Computational details.

Molecular mechanics (MM) calculations were performed using the MacroModel 8.5 software package and the MMFFs force fields. MonteCarlo Multiple Minimum (MCMM) method (10,000 steps) of the MacroModel package was used in order to allow a full exploration of the conformational space. All the structures, so obtained, were optimised using the PolakeRibier Coniugated Gradient algorithm (PRCG, 1000 steps, maximum derivative less than 0.05 kcal/mol). The initial geometries of the minimum energy conformers were optimised at the hybrid DFT

MPW1PW91 level using the 6-31G(d) basis set (Gaussian 03 software package). GIAO *J*-coupling calculations were performed using the MPW1PW91 functional and the 6-31G(d) basis set, using as the input the geometry previously optimised at MPW1PW91/6-31G(d) level.

Synthetic procedures.

(E)-Ethyl 6-methylhept-2-enoate (**83**). DMSO (4.18 mL, 58.8 mmol) was added dropwise for 15 min to a solution of oxalyl chloride (14.7 mL, 29.4 mmol) in dry dichloromethane (50 mL) at -78 °C under argon atmosphere. After 30 min a solution of the alcohol **82** (1.00 g, 9.80 mmol) in dry CH_2Cl_2 was added via cannula and the mixture was stirred at -78 °C for 1 h. Et_3N (6.83 mL, 49.0 mmol) was added dropwise and the mixture was allowed to warm to room temperature. The reaction was quenched by addition of aqueous $NaHSO_4$ (1 M, 50 mL). The layers were separated and the aqueous phase was extracted with CH_2Cl_2 (3x50 mL). The combined organic layers were washed with saturated aqueous $NaHSO_4$, saturated aqueous $NaHCO_3$ and brine. The organic phase was then dried over Na_2SO_4 and concentrated to give the corresponding aldehyde **81** (0.951 g, 97%) as a colourless oil, which was used without any further purification. δH (400 MHz, $CDCl_3$) 9.70 (1H, s, CHO), 2.44 (2H, m, CH_2CHO), 1.59 (1H, m, $CHMe_2$), 1.25 (2H, m, CH_2), 0.92 (3H, d, J 6.6 Hz, Me), 0.91 (3H, d, J 6.6 Hz, Me); δC (100 MHz $CDCl_3$) 203.5, 40.2, 31.1, 27.9, 22.7, 22.5. To a solution of compound **81** (0.950 g, 9.50 mmol) and LiOH (250 mg, 10.5 mmol) in THF (10 mL) was added TEPA (triethyl phosphonoacetate, 2.07 mL, 10.5 mmol). The reaction mixture was stirred for 24 h at room temperature and then quenched withwater (10 mL). The mixture was then extracted with EtOAc (3x30 mL), and the organic phase was concentrated in vacuo. Flash chromatography (hexane/EtOAc, 99:1) afforded pure **83** (1.21 g, 75%). δH (400 MHz, $CDCl_3$) 6.95 (1H, dt, J 15.4, 6.6, CH=CHCO), 5.80 (1H, d, J 15.4 Hz, CH=CHCO), 4.17 (2H, q, J 7.4, OCH_2), 2.20 (2H, m, $CH_2CH_=$), 1.56 (1H, m, $CHMe_2$), 1.32 (2H, m, CH_2), 1.24 (3H, t, J 7.4, OCH_2Me), 0.88 (3H, d, J 6.6 Hz, Me), 0.89 (3H, d, J 6.6 Hz, Me); δC (100 MHz $CDCl_3$) 165.4, 150.8, 121.0, 60.4, 37.2, 30.4, 27.8, 22.5 (2C), 14.5; HRMS (ESI): calcd for $C_{10}H_{19}O_2$: 171.1382; found 171.1389 $[M+H]^+$.

(E)-6-Methylhept-2-en-1-ol (**80**). To a solution of compound **83** (1.10 g, 6.47 mmol) in toluene (35 mL) at -78 °C, under a nitrogen atmosphere and with stirring, was slowly added DIBAL-H (7.59 mL of a 1.7 M solution in toluene, 12.9 mmol). After 1 h (TLC monitoring), the reaction was quenched with saturated NH_4C1 (25 mL) and the mixture was extracted with Et_2O (50 mL). Organic layer was dried over Na_2SO_4, concentrated in vacuo and purified on SiO_2 chromatography (hexane/EtOAc, 90:10) affording pure compound **80** (787 mg, 95%). δH (400 MHz, $CDCl_3$) 5.67 (2H, m, CH=CH), 4.09 (2H, d, J 5.2 Hz, CH_2OH), 2.05 (2H, m, $CH_2CH_=$),1.63 (1H, m, $CHMe_2$), 1.27 (2H, m, CH_2), 0.82 (3H, d, J 6.6 Hz, Me), 0.81 (3H, d, J 6.6 Hz, Me); δC (100 MHz $CDCl_3$) 133.8., 128.6, 63.8, 38.3, 30.1, 27.5, 22.4 (2C); HRMS (ESI): calcd for $C_8H_{17}O$: 129.1279; found 129.1274 $[M+H]^+$.

(2S,3S)-2,3-Epoxy-6-methylheptan-1-ol (**84**). To a 100 mL round bottom flask under argon equipped with a magnetic stirrer were added molecular sieves (4A°, 164 mg) in CH_2Cl_2 (50mL). At -23 °C were then added $Ti(Oi\text{-}Pr)_4$ (77.7 mg, 0.273mmol) and L-(+)-DET (67.7mg,

0.328mmol). The solution was allowed to stir for 5 min, then compound **80** (0.700 g, 5.47mmol) and TBHP (1.98mL of a 5.5 M solution in decane, 10.9mmol) were added successively. After 24 h a solution of tartaric acid (492 mg, 3.28mmol) and $FeSO_4$ (1.8 g, 6.56mmol) in 20mL of water was added and was stirred at -23 °C. After 30min, the cooling bath was removed and stirring was continued at room temperature for 1 h until the aqueous layer become clear. After separation of the aqueous layer, the organic layer was washed once with water, dried over Na_2SO_4 and concentrated. This oil was diluted with Et_2O (50mL) and cooled in a ice bath, and then NaOH (25mL of 1 N solution in brine) was added; the two phase mixture was stirred at 0 °C for 0.5 h, and then the ether phase was washed with brine, dried over Na_2SO_4 and concentrated. SiO_2 chromatography (hexane/EtOAc, 90:10) afforded pure compound **84** (709mg, 90%). $[\alpha]_D^{20}$ -10.7 (c 0.3, $CHCl_3$); δH (400 MHz, $CDCl_3$) 3.89 (1H, br d, J 12.6 Hz, O-CHCH_2OH), 3.60 (1H, br d, J 12.6 Hz, CH-O), 2.92 (2H, m, CH_2OH), 1.58 (1H, m, $CHMe_2$), 1.32 (2H, m, CH_2CH-O), 1.25 (2H, m, CH_2CH_2CH-O), 0.88 (3H, d, J 6.4 Hz, Me), 0.87 (3H, d, J 6.4 Hz, Me); δC (100 MHz $CDCl_3$) 61.9, 58.8, 53.7, 35.1, 29.6, 28.0, 22.7, 22.6; HRMS (ESI): calcd for $C_8H_{17}O_2$: 145.1229; found 145.1224 $[M+H]^+$. Compound **84** (0.5-1.0 mg) was dissolved in freshly distilled CH_2Cl_2 and treated with triethylamine (10 mL) and (+)-α-methoxy-α-(trifluoromethyl)phenylacetyl chloride [(*S*)-MTPA-Cl] and a catalytic amount of 4-(dimethylamino)pyridine. The mixturewas left to stand at room temperature for 2 h. After this period, the mixture was concentrated in vacuo affording pure (*R*)-MTPA ester in quantitative yield. δH (400 MHz, $CDCl_3$) 7.47 (2H, m, ArH), 7.36 (3H, m, ArH), 4.47 (1H, dd, J 12.0, 3.0 Hz CH_2O), 4.16 (1H, dd, J 12.0, 6.0 Hz CH_2O), 3.50 (3H, s, -OMe), 2.94 (1H, m, CHO), 2.76 (1H, m, CHO) 1.47 (4H, m, CH_2CH_2CH-O), 1.22 (1H, m, $CHMe_2$), 0.81 (3H, d, J 6.6 Hz, Me), 0.80 (3H, d, J 6.6 Hz, Me); HRMS (ESI): calcd for $C_{18}H_{24}F_3O_4$: 361.1627; found 361.1632 $[M+H]^+$.

(2R,3S)-2,3-Epoxy-6-methylheptanoic acid (**79**). To a vigorously stirred mixture of compound **84** (0.700 g, 4.86 mmol) were added $NaIO_4$ (4.25 g, 19.9 mmol) in CCl_4 (18.5 mL), CH_3CN (18.5 mL), H_2O (27.8 mL) and $RuCl_3{\cdot}H_2O$ (26.2 mg, 0.0972 mmol). The mixture was stirred at 20 °C for 2 h; then the acidic material was carefully extracted at 0 °C into ether, dried briefly over Na_2SO_4 and evaporated in vacuo to give a crude residue **79** (750 mg), that was subjected to next steps without further purification. δH (400 MHz, $CDCl_3$) 3.50 (1H, m, O-CHCOOH), 3.19 (1H, m, CH-O), 1.62 (1H, m, $CHMe_2$), 1.36 (2H, m, CH_2CH-O), 1.26 (2H, m, CH_2CH_2CH-O), 0.91 (3H, d, J 6.3 Hz, Me), 0.90 (3H, d, J 6.3 Hz, Me); δC (100 MHz $CDCl_3$) 173.0, 59.4, 52.8, 34.7, 29.6, 27.9, 22.6 (2C); HRMS (ESI): calcd for $C_8H_{13}O_3$: 157.0865; found 157.0861 $[M+H]^+$.

(2R,3S)-Methyl 2,3-epoxy-6-methylheptanoate (**85**). To an aliquot of **79** (200 mg) in CH_2Cl_2 was added an ethereal solution of diazomethane until the solution become yellow. Evaporation and silica gel chromatography (hexane/EtOAc, 99:1) afforded pure **85** (196 mg, 82% over two steps from **84**). $[\alpha]_D^{20}$ +27.0 (c 1.1, $CHCl_3$); δH (400 MHz, $CDCl_3$) 3.76 (3H, s,OMe), 3.50 (1H, m, O-CHCOOMe), 3.13 (1H, m, CH-O),1.63 (1H, m, $CHMe_2$),1.36 (2H, m, CH_2CH-O),1.26 (2H, m, CH_2CH_2CH-O), 0.88 (3H, d, J 6.3 Hz, Me), 0.87 (3H, d, J 6.3 Hz, Me); δC (100 MHz $CDCl_3$)

170.8, 58.9, 53.2, 52.6, 34.7, 29.5, 27.9, 22.5, 22.6; HRMS (ESI): calcd for $C_9H_{17}O_3$: 173.1178; found 173.1172 $[M+H]^+$.

(2S,3R)-3-Bromo-2-hydroxy-6-methylheptanoate (**86**). To a solution of compound **85** (160 mg, 0.930 mmol) in Et_2O (6 mL) was added $MgBr_2{\cdot}Et_2O$ (361 mg, 1.40 mmol). The solutionwas stirred at room temperature for 1 h and then washed with brine, and the organic layers were evaporated in vacuo. The crude mixture was purified by SiO_2 chromatography (hexane/EtOAc, 99:1), affording pure **86** (230 mg, 98%). $[\alpha]_D^{20}$ -6.4 (c 0.3, $CHCl_3$); δH (400 MHz, $CDCl_3$) 4.40 (1H, d, J 3.3 Hz, CHOH), 4.16 (1H, m, CH-Br), 3.81 (3H, s, OMe), 1.90 (1H, m, CH_aH_bCHBr), 1.79 (1H, m, CH_aH_bCHBr), 1.55 (1H, m, $CHMe_2$), 1.44 (1H, m, $CH_aH_bCH_2CHBr$),1.29 (1H, m, $CH_aH_bCH_2CHBr$), 0.88 (3H, d, J 6.3 Hz; Me), 0.86 (3H, d, J 6.3 Hz, Me); δC (100 MHz $CDCl_3$) 171.5, 74.7, 53.0, 37.0, 32.0, 27.7, 23.0, 22.4, 21.2; HRMS (ESI): calcd for $C_9H_{18}{}^{79}BrO_3$: 253.0434; found 253.0439 $[M+H]^+$.

(2R,3S)-3-Azido-2-hydroxy-6-methylheptanoate (**87**). A mixture of **86** (189 mg, 0.750 mmol) and NaN_3 (195 mg, 3.00 mmol) in DMF (5.0 mL) was stirred at 65 °C for 24 h. The mixture was then diluted with EtOAc (10.0 mL), washed with water (10.0 mL), and concentrated in vacuo. Silica gel chromatography (hexane/ether, 99:1) afforded pure **87** (148 mg, 92%). $[\alpha]_D^{20}$ +1.5 (c 0.1, $CHCl_3$); δH (400 MHz, $CDCl_3$) 4.25 (1H, d, J 5.6 Hz, CHOH), 3.82 (3H, s, OMe), 3.47 (1H, t, J 7.0, CH-N_3), 3.03 (1H, d, J 5.6 Hz, CHOH), 1.83 (1H, m, $CH_aH_bCHN_3$), 1.81 (1H, m, $CH_aH_bCHN_3$), 1.61 (1H, m, $CHMe_2$), 1.34 (2H, m, $CH_2CH_2CHN_3$), 0.93 (3H, d, J 6.3 Hz, Me), 0.91 (3H, d, J 6.3 Hz, Me); δC (100 MHz $CDCl_3$) 172.6, 68.4, 59.0, 53.7, 30.6, 29.2, 24.0, 23.2 (2C); HRMS (ESI): calcd for $C_9H_{18}N_3O_3$: 216.1348; found 216.1352 $[M+H]^+$.

(2R,3S)-3-Amino-2-hydroxy-6-methylheptanoic acid (**78b**). A solution of **87** (130 mg, 0.604 mmol) in ethyl acetate was hydrogenated in the presence of Pd/C catalyst (2 mg) for 24 h at room temperature. The mixture was filtered through Celite and concentrated in vacuo to give a 125 mg residue that was subjected to under vacuum vapour-phase hydrolysis (HCl 6 N, 110 °C for 18 h). HPLC purification on the reversed-phase Phenomenex Hydro (4 m, 250x4.6 mm) column eluting with 85% H_2O/MeOH (0.1% TFA) furnished compound **78b** (79.3 mg, rt=19.5 min) in 75% yield over two steps. $[\alpha]_D^{20}$ +3.3 (c 0.3, MeOH); δH (400 MHz, CD_3OD) see Table 8 in the text; ^{13}C NMR (100 MHz, CD_3OD) δC: see Table 8 in the text; HRMS (ESI): calcd for $C_8H_{18}NO_3$: 176.1286; found 176.1293 $[M+H]^+$.

(2R,3R)-3-Amino-2-hydroxy-6-methylheptanoic acid (**78a**). Compound **79** (200 mg, 1.27 mmol) was dissolved in water (2 mL). Powdered NaN_3 (124 mg, 1.91 mmol) and 1.3 mL of an aqueous solution 0.1 M of $Cu(NO_3)_2$ were added under stirring (resulting pH 4.3e4.5), and the mixture was warmed to 65 °C. After 1.5 h (ca. pH 5.5) the reaction mixture was cooled to 0 °C, and $NaBH_4$ was added portion-wise (96.1 mg, 2.54 mmol). After 30 min at 0 °C the reaction mixture was filtered and concentrated in vacuo to give a residue that, subjected to HPLC purification on the reversedphase Phenomenex Hydro (4 m, 250x4.6 mm) column eluting with 85% H_2O/MeOH (0.1% TFA), furnished compound **78a** (167 mg, rt=21 min) in 75% yield over two steps. $[\alpha]_D^{20}$ +8.6 (c 0.5, MeOH); δH (400 MHz, CD_3OD) see Table 8 in the text; ^{13}C NMR (100 MHz,

CD_3OD) δ: see Table 8 in the text; HRMS (ESI): calcd for $C_8H_{18}NO_3$: 176.1286; found 176.1291 $[M+H]^+$.

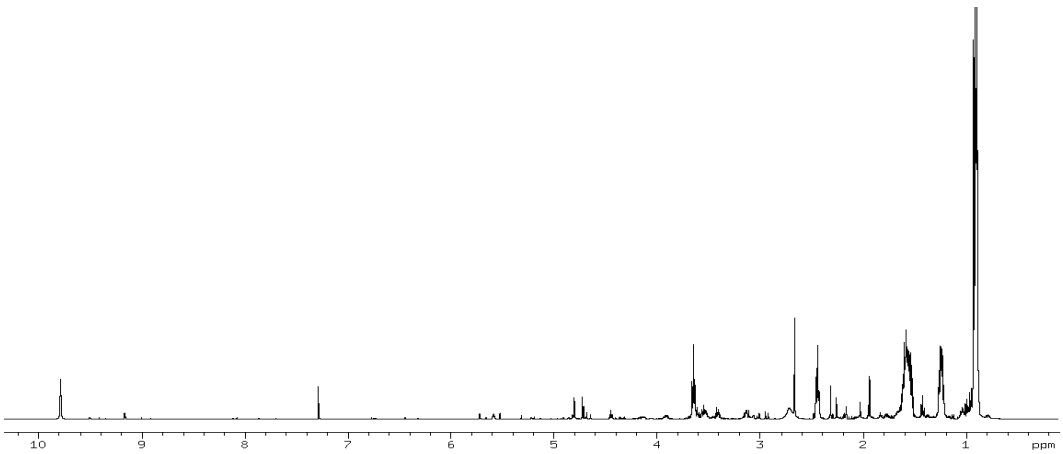
Determination of the absolute configuration

Peptide hydrolysis and AHMHA isolation. A 30 mg sample of perthamide C was dissolved in 6 N HCl (3 mL) and heated in vacuo at 130 °C for 12 h. The crude residue was fractionated by HPLC on the reversed-phase Phenomenex Hydro (4 m, 250x4.6 mm) column eluting with 85% H_2O/MeOH (0.1% TFA) (flow rate 0.5 mL/min) to give 3 mg of pure ADMHA (rt=21 min). $[\alpha]_D^{20}$ +8.8 (c 0.55, MeOH); δ_H (400 MHz, CD_3OD) see Table 8 in the text; ^{13}C NMR (100 MHz, CD_3OD) δ: see Table 8 in the text.

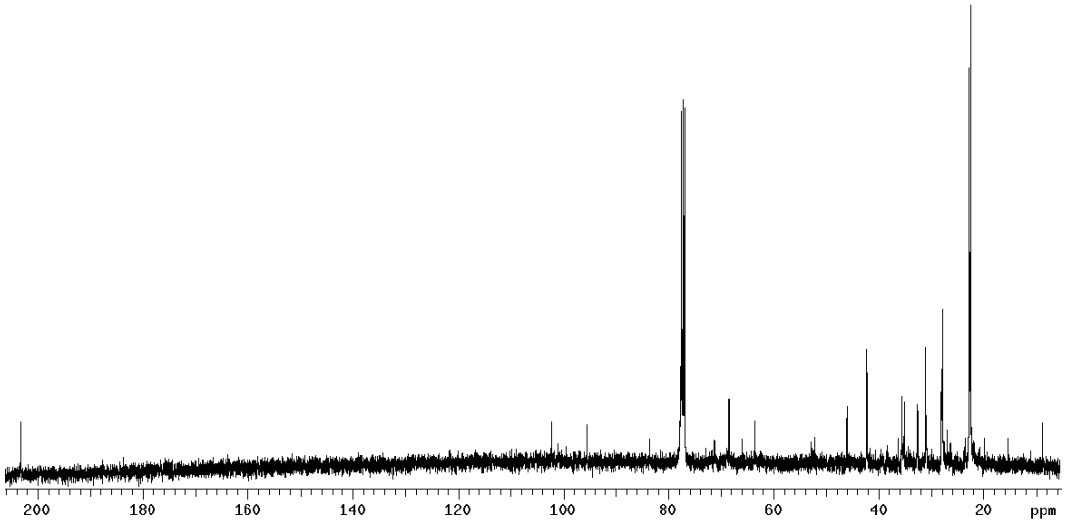
LC-MS analysis of Marfey's (FDAA) derivatives. A portion of pure AHMHA or the amino acid standards **78a** and **78b** (500 mg) was dissolved in 80 mL of a 2:3 solution of TEA/MeCN and this solution was then treated with 75 mL of 1% 1-fluoro-2,4-dinitrophenyl-5- alaninamide (L- or D-FDAA) in 1:2 MeCN/acetone. The vials were heated at 70 °C for 1 h, and the contents were neutralised with 0.2 N HCl (50 mL) after cooling to room temperature. An aliquot of the FDAA derivative was dried under vacuum, diluted with 50% aqueous acetonitrile containing 5% formic acid, and separated on a Proteo C18 (25x1.8 mm i.d.) column by means a linear gradient from 10% to 50% aqueous acetonitrile containing 5% formic acid and 0.05% trifluoroacetic acid, over 45 min at 0.15 mL/min. The RP-HPLC system was connected to the electrospray ion source by inserting a splitter valve and the flow going into the mass spectrometer source was set at a value of 100 mL/min. Mass spectra were acquired in positive ion detection mode (m/z interval of 320e900) and the data were analysed using the suite of programs Xcalibur; all masses were reported as average values. Capillary temperature was set at 280 °C, capillary voltage at 37 V, tube lens offset at 50 V and ion spray voltage at 5 V. Retention times of L-FDAA-3-amino-2-hydroxy-6-methylheptanoic acid (min): (2*R*, 3*R*)-**88a** (48.50), (2*R*, 3*S*)-**88b** (39.40), AHMHA-naturale (48.50). Retention times of D-FDAA-3-amino-2-hydroxy-6-methylheptanoic acid (min): (2*R*, 3*R*)-**88c** (40.99), (2*R*, 3*S*)-**88d** (48.30).

Spectroscopic data

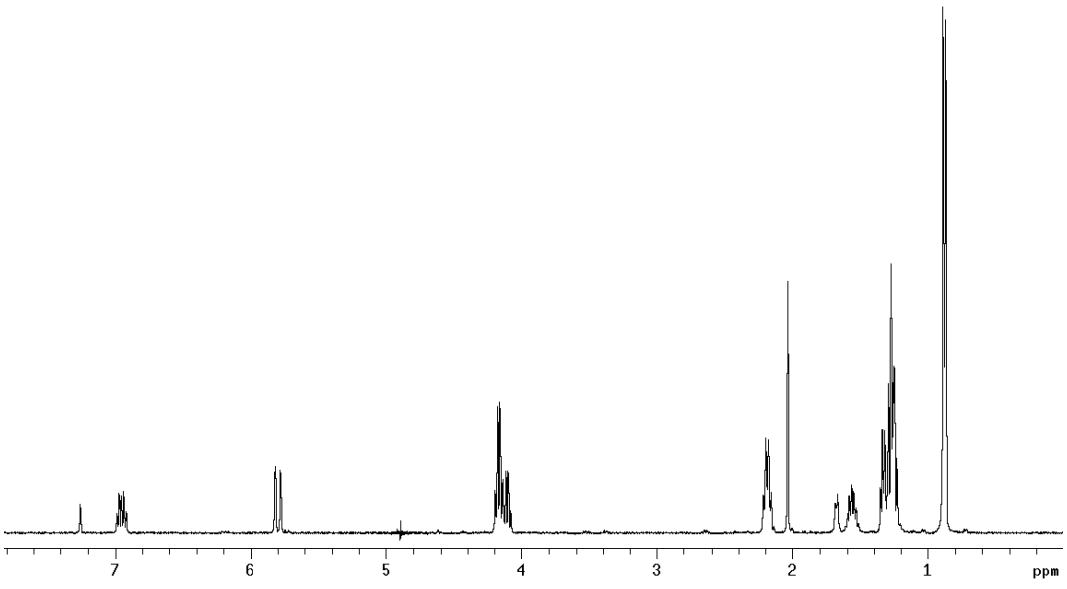
^1H NMR of Compound (**81**) in $CDCl_3$ at at 400 MHz.



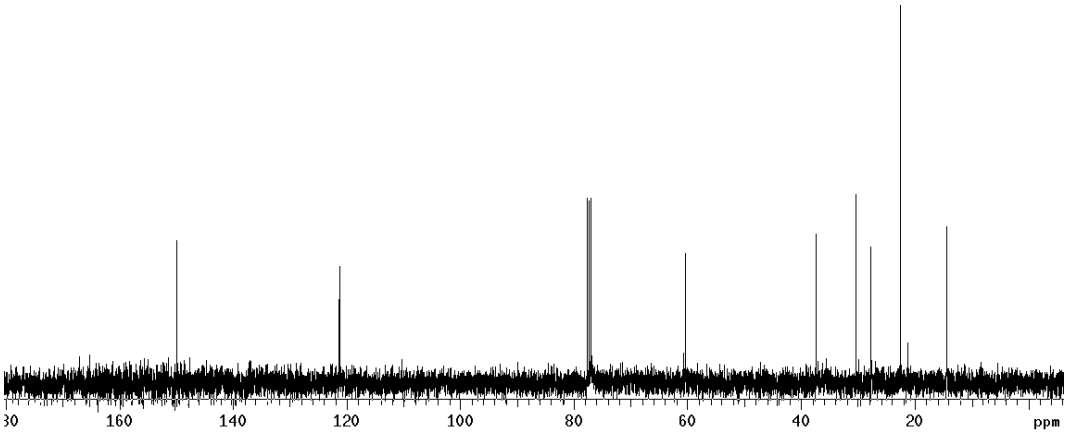
^{13}C NMR of Compound (**81**) in $CDCl_3$ at at 100 MHz.



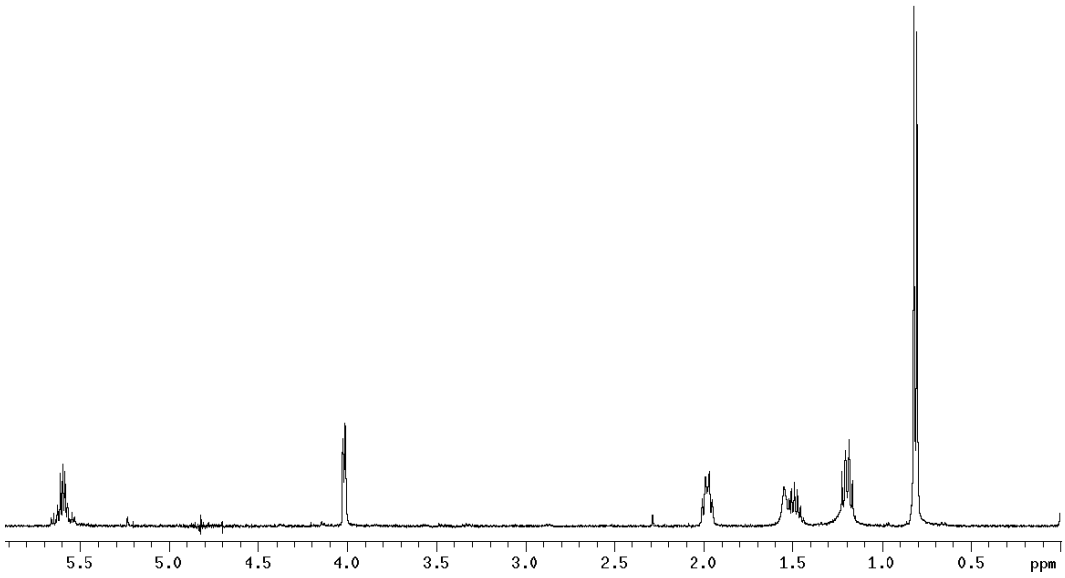
^{1}H NMR of Compound (**83**) in $CDCl_3$ at at 400 MHz.



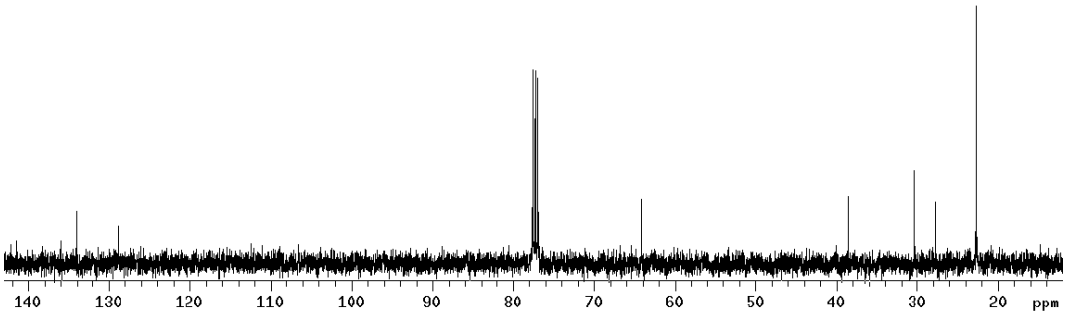
^{13}C NMR of Compound (**83**) in $CDCl_3$ at at 100 MHz.



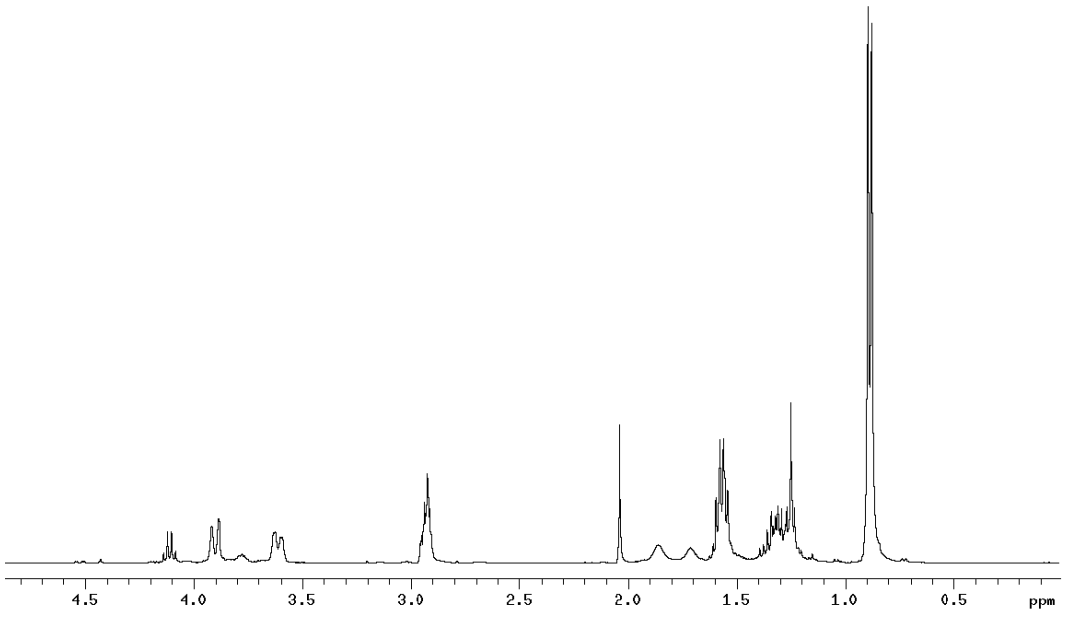
^{1}H NMR of Compound (**80**) in $CDCl_3$ at at 400 MHz.



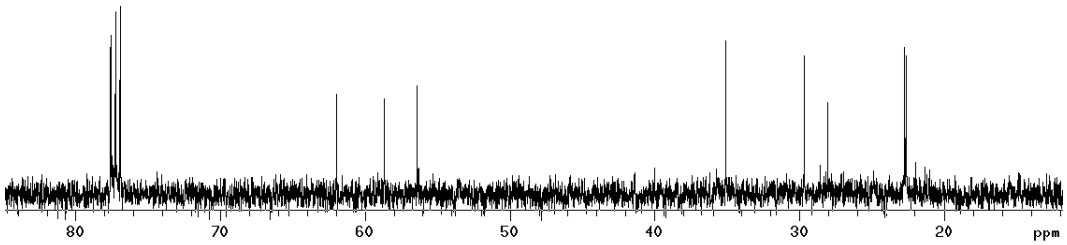
^{13}C NMR of Compound (**80**) in $CDCl_3$ at at 100 MHz.



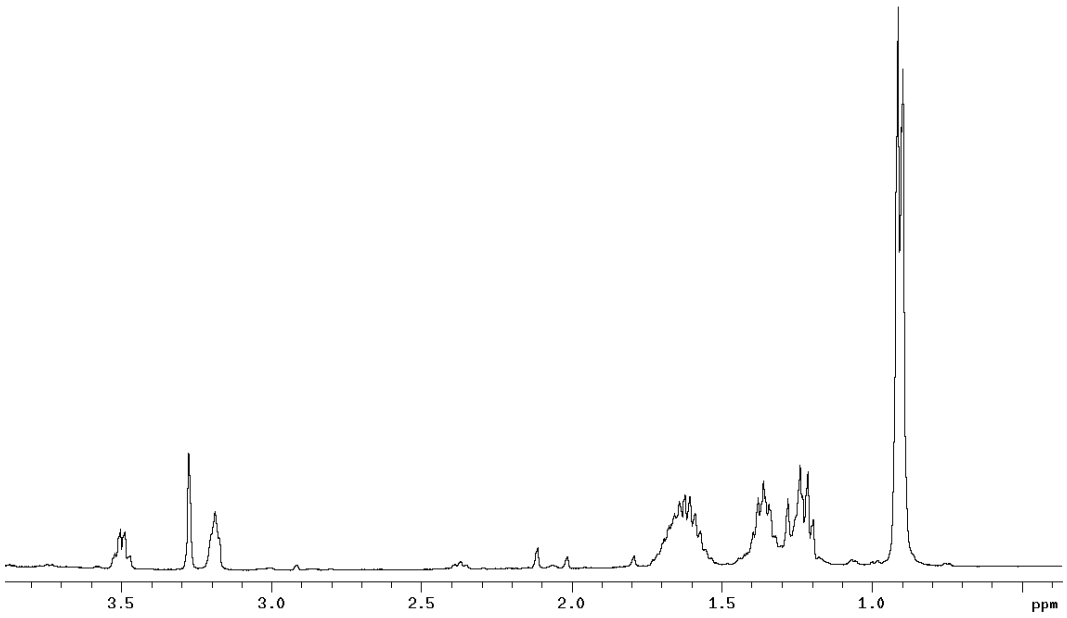
^{1}H NMR of Compound (**84**) in $CDCl_3$ at at 400 MHz.



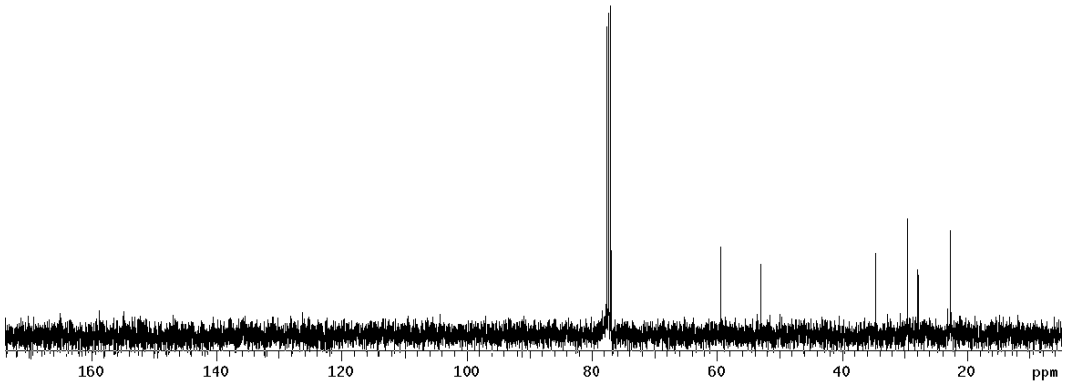
^{13}C NMR of Compound (**84**) in $CDCl_3$ at at 100 MHz.



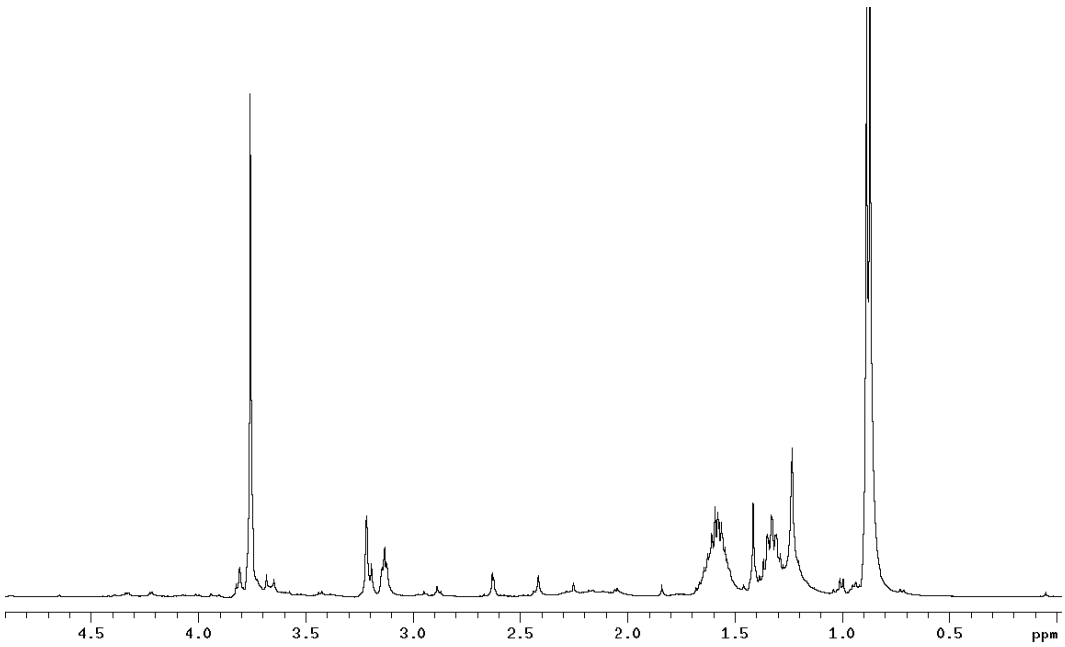
^{1}H NMR of Compound (**79**) in $CDCl_3$ at at 400 MHz.



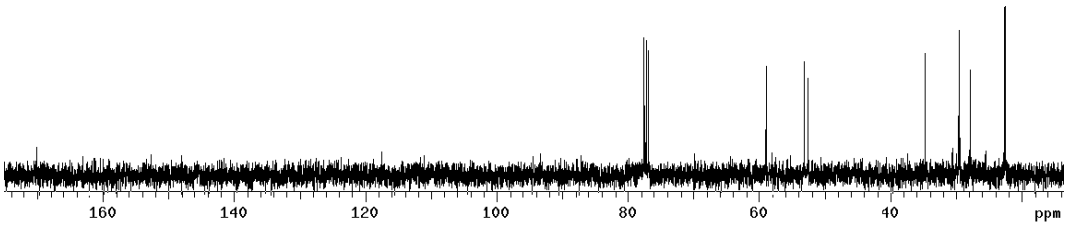
^{13}C NMR of Compound (**79**) in $CDCl_3$ at at 100 MHz.



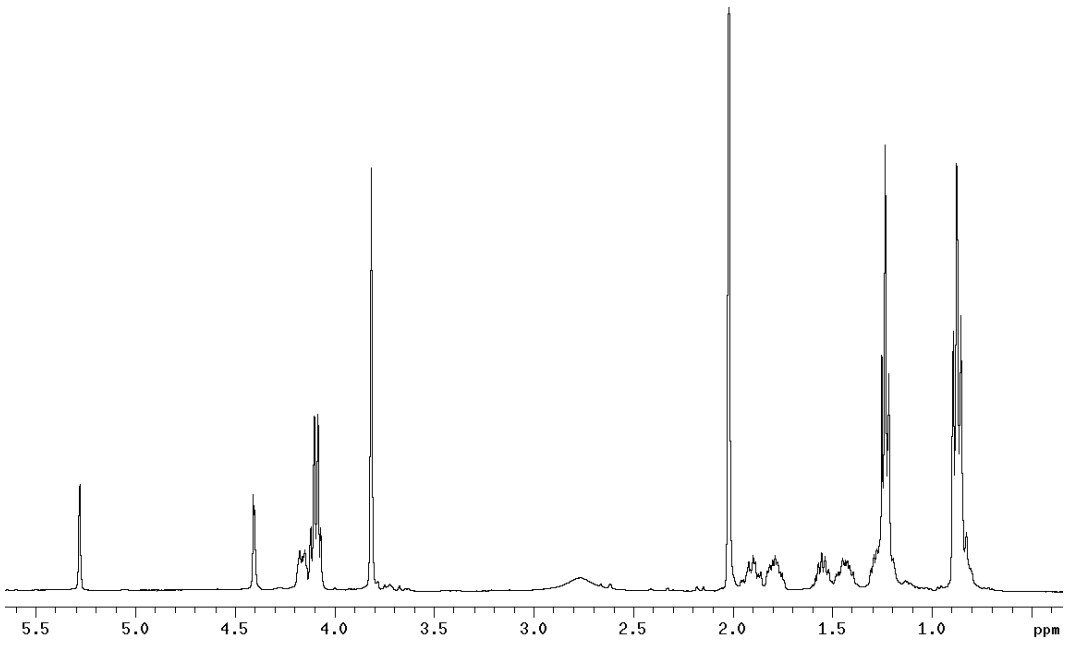
^{1}H NMR of Compound (**85**) in $CDCl_3$ at at 400 MHz.



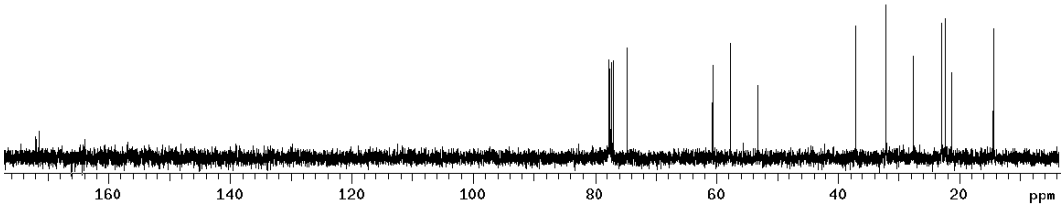
^{13}C NMR of Compound (**85**) in $CDCl_3$ at at 100 MHz.



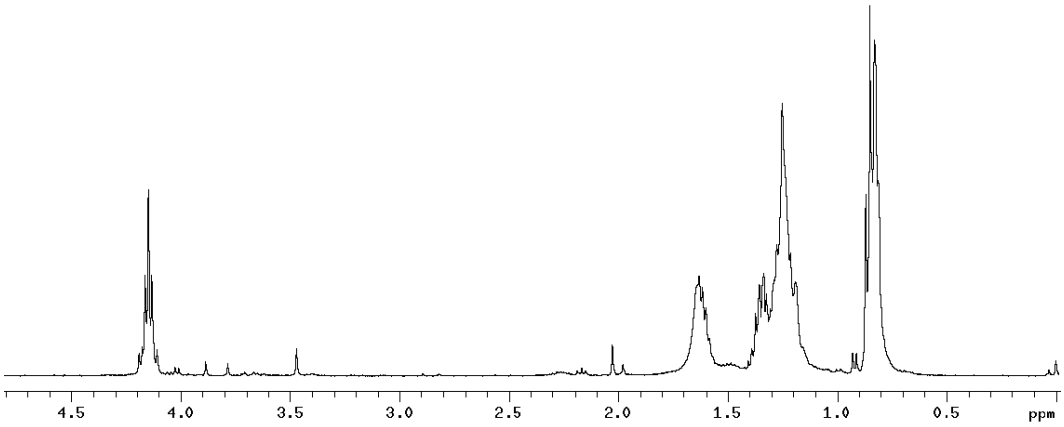
^{1}H NMR of Compound (**86**) in $CDCl_3$ at at 400 MHz.



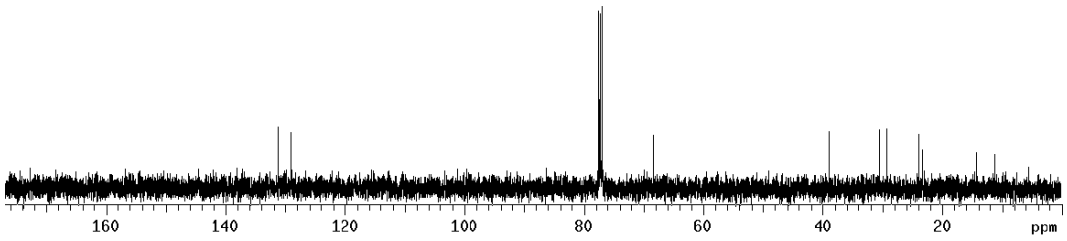
^{13}C NMR of Compound (**86**) in $CDCl_3$ at at 100 MHz.



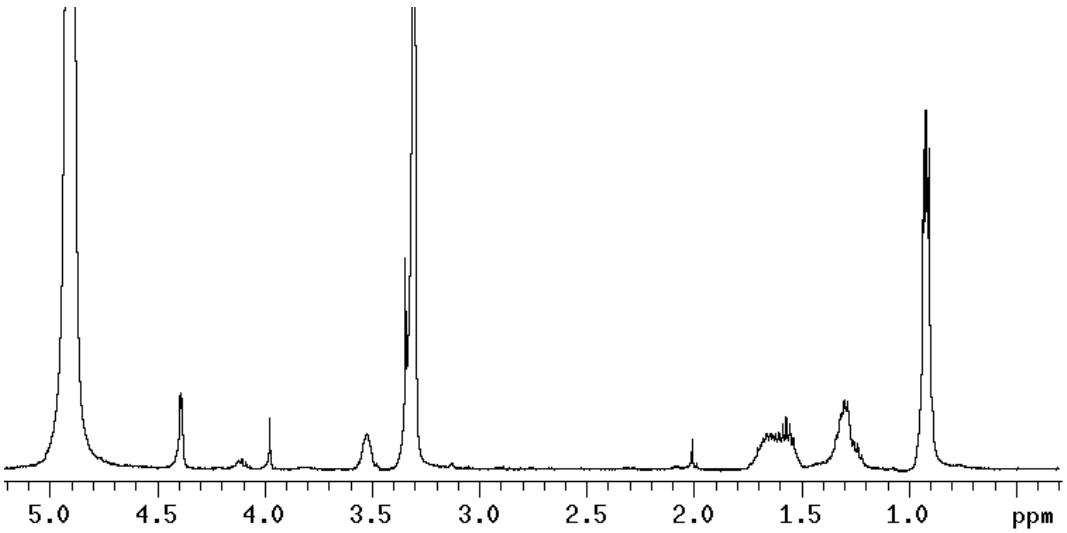
^{1}H NMR of Compound (**87**) in $CDCl_3$ at at 400 MHz.



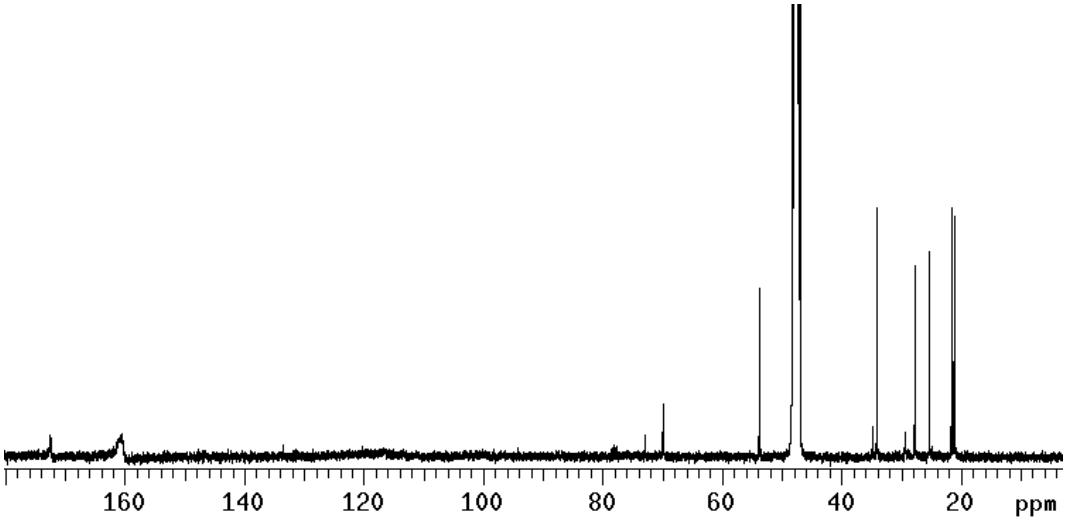
^{13}C NMR of Compound (**87**) in $CDCl_3$ at at 100 MHz.



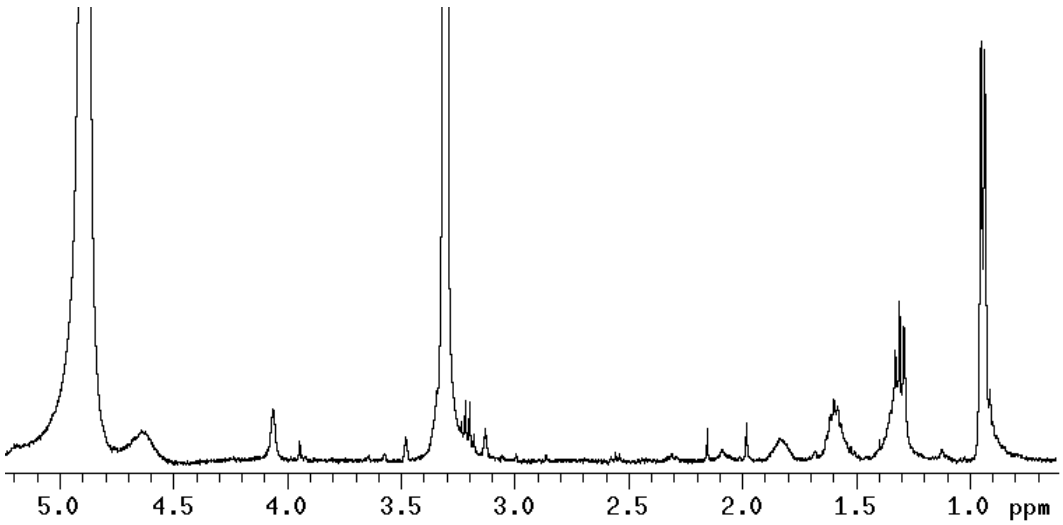
^{1}H NMR of Compound (**78a**) in CD_3OD at at 400 MHz.



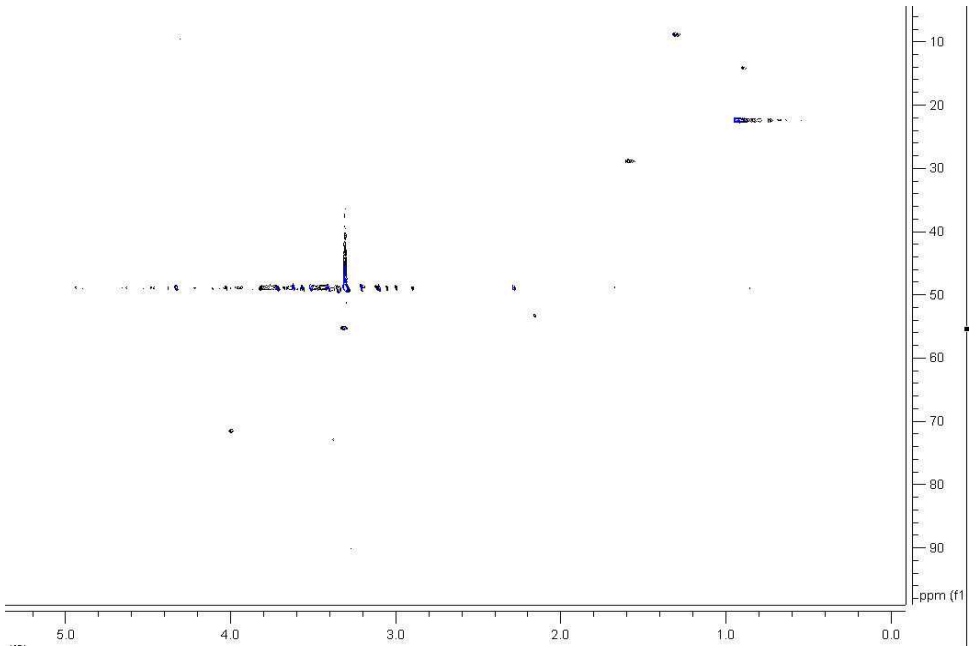
^{13}C spectrum of Compound (**78a**) in CD_3OD at at 100 MHz.



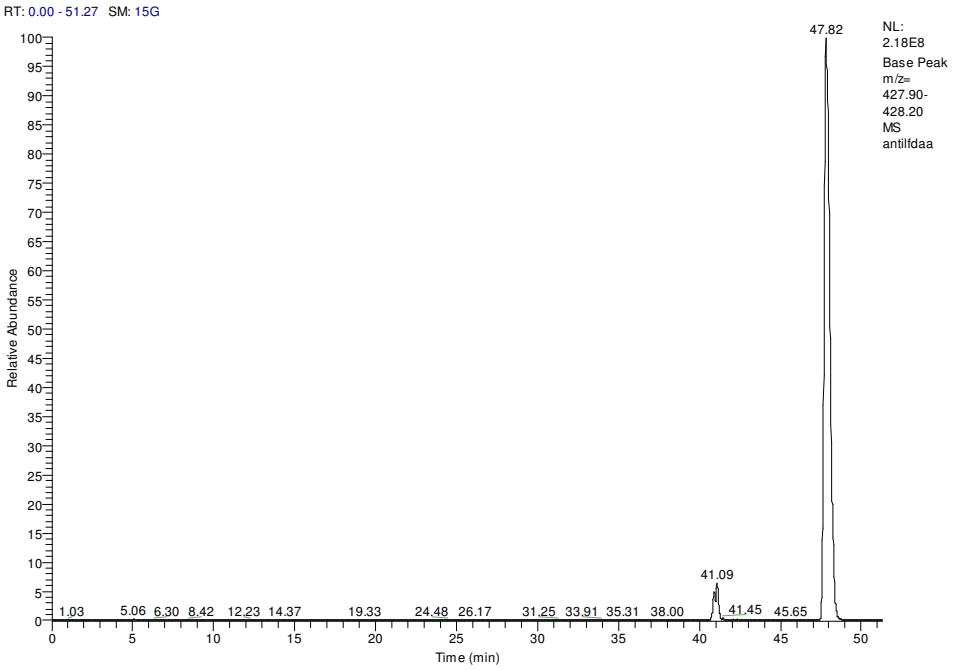
^{1}H NMR of Compound (**78b**) in CD_3OD at at 400 MHz.



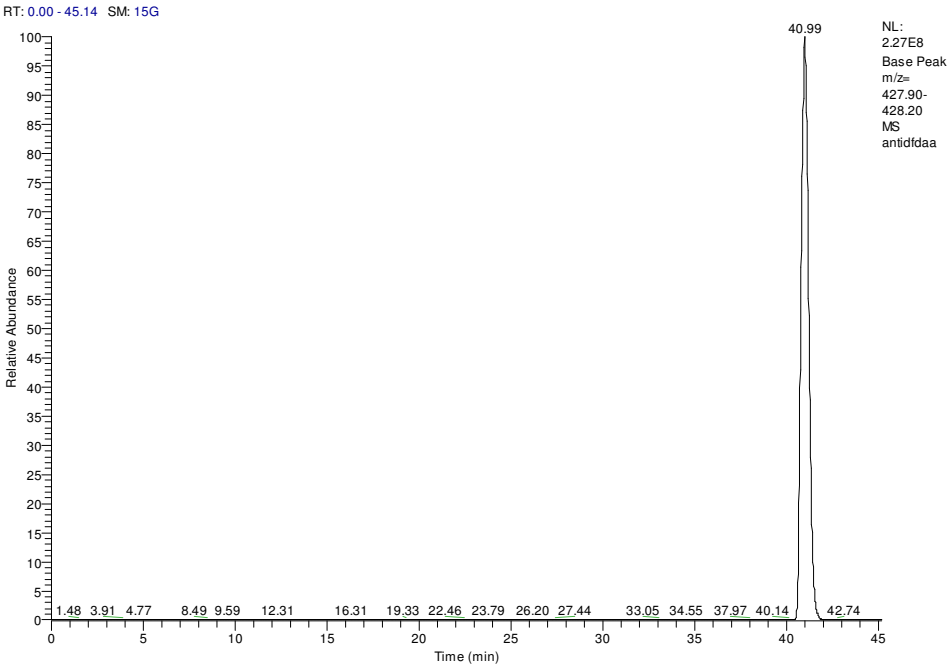
HSQC spectrum of Compound (**78b**) in CD_3OD at at 100 MHz.



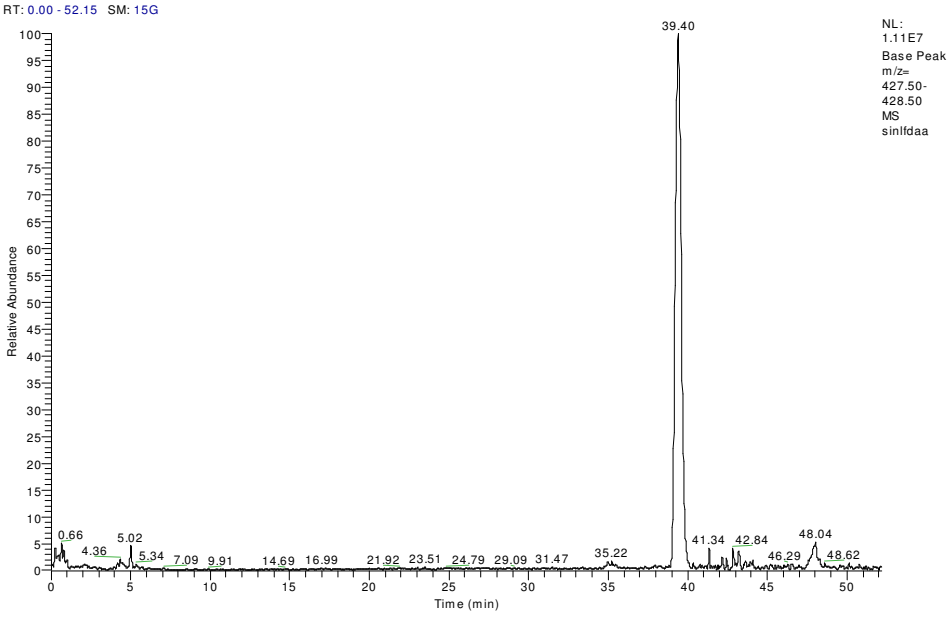
ESI-LC/MS spectrum of (**88a**).



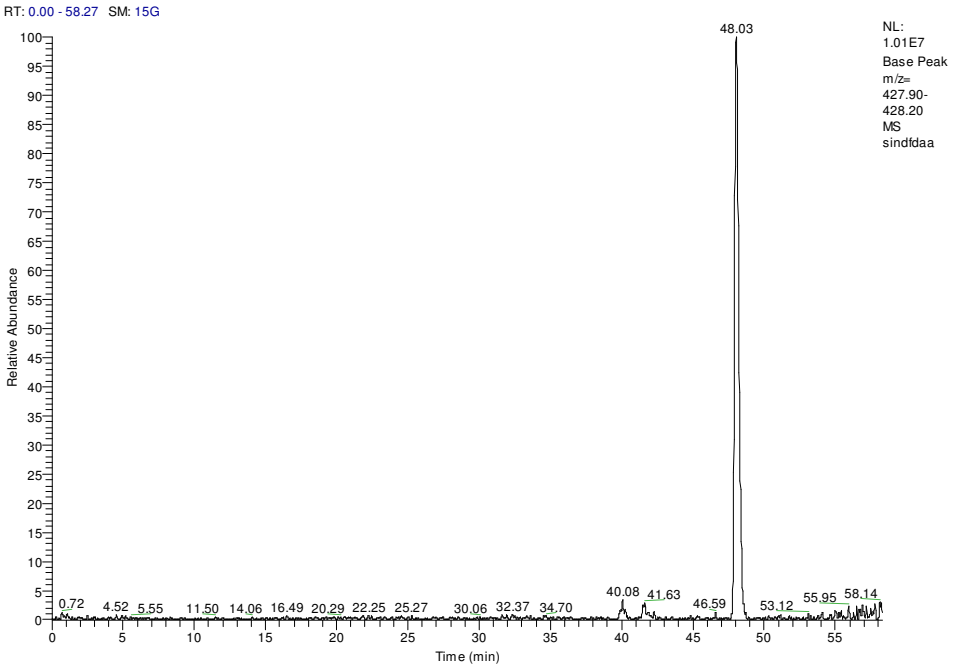
ESI-LC/MS spectrum of (**88c**).



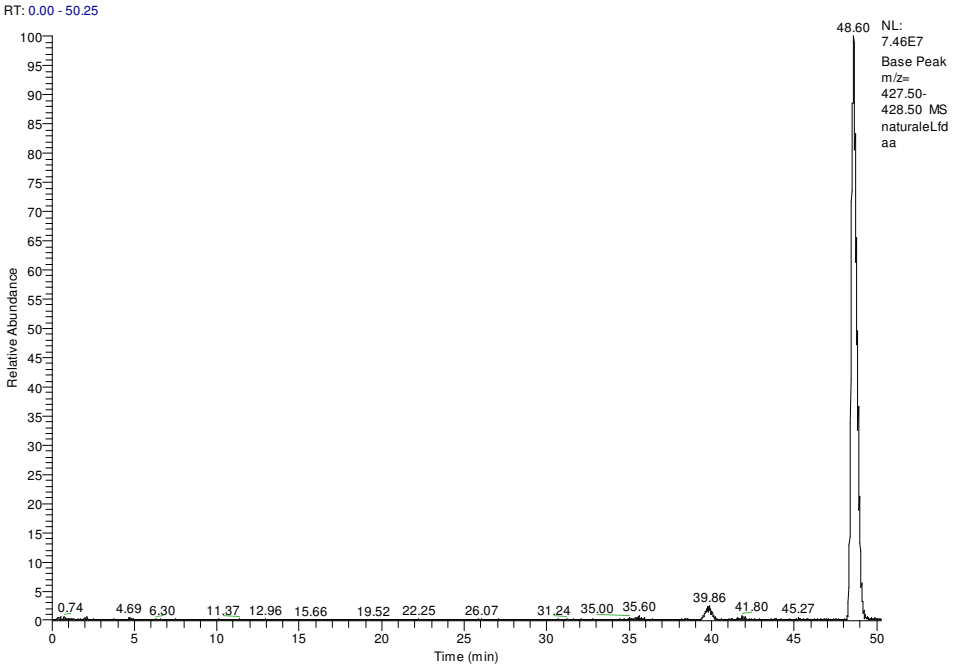
ESI-LC/MS spectrum of (**88b**).



ESI-LC/MS spectrum of (**88d**).



ESI-LC/MS spectrum of (L)-FDAA derivative of AHMHA.



REFERENCES

[1] Li, J. W. H.; Vederas, J. C. *Science,* **2009**, *325*, 161-165.

[2] Newman, D. J.; Cragg, G. M.; Snader, K. M. *J. Nat. Prod.*, **2003**, *66*, 1022-1037.

[3] Clardy, J.; Walsh, C. *Nature*, **2004**, *432*, 829-837.

[4] Koehn, F. E.; Carter, G. T. *Nat. Rev. Drug Discov.*, **2005**, *4*, 206-220.

[5] Carlson, E. E. *ACS Chem. Biol.*, **2010**, *5*, 639-653.

[6] Murti, Y.; Agrawal, T. *Int. J. ChemTech Res.*, **2010**, *2*, 2198-2217.

[7] Landry, Y.; Gies, J. P. *Fundam. Clin. Pharmacol.*, **2008**, *22*, 1-18.

[8] D'Auria, M. V.; Sepe, V.; Zampella, A. *Curr Top Med Chem.,* **2012,** 12 (6): 637-69.

[9] Rock, K. L.; Latz, E.; Ontiveros, F.; Kono, H. *Annu. Rev. Immunol.*, **2010**, *28*, 321-342.

[10] Synold, T. W., Dussault, I.; Forman, B. M. *Nat. Med.,* **2001**, *7***,** 584-590.

[11] Maglich, J. M.; Stoltz, C. M.; Goodwin, B.; Hawkins-Brown, D.; Moore, J. T.; Kliewer, S. A. *Mol. Pharmacol.*, **2002**, *62***,** 638-646.

[12] Rae, J. M.; Johnson, M. D.; Lippman, M. E.; Flockhart, D. A. *J. Pharmacol. Exp. Ther.,* **2001**, *299***,** 849-857.

[13] Sonoda, J.; Xie, W.; Rosenfeld, J. M.; Barwick, J. L.; Guzelian, P. S.; Evans, R. M. *Proc. Natl. Acad. Sci. U.S.A.*, **2002**, *99*, 13801-13806.

[14] Moreau, A.; Vilarem, M. J.; Maurel, P.; Pascussi, J. M. *Mol. Pharmacol.*, **2008**, *5*, 35-41.

[15] Mencarelli, A.; Migliorati, M.; Barbanti, M.; Cipriani, S.; Palladino, G.; Distrutti, E.; Renga, B.; Fiorucci, S. *Biochem. Pharmacol.,* **2010**, *80*, 1700-1707.

[16] Kakizaki, S.; Yamazaki, Y.; Takizawa, D.; Negishi, M. *Curr. Drug Metab.*, **2008**, *9*, 614-621.

[17] Cheng, J.; Shah, Y. M.; Ma, X.; Pang, X.; Tanaka, T.; Kodama, T.; Krausz, K. W.; Gonzalez, F. J. *J. Pharmacol. Exp. Ther.,* **2010**, *335*, 32-41.

[18] Xue, Y.; Moore, L. B.; Orans, J.; Peng, L.; Bencharit, S.; Kliewer, S. A.; Redinbo, M. R. *Mol. Endocrinol.*, **2007**, *21*, 1028-1038.

[19] Ekins, S.; Chang, C.; Mani, S.; Krasowski, M. D.; Reschly, E. J.; Iyer, M.; Kholodovych, V. A. N.; Welsh, W. J.; Sinz, M.; Swaan, P. W.; Patel, R.; Bachmann, K. *Mol. Pharmacol.*, **2007**, *72*, 592-603.

[20] Cheng, J.; Shah, Y. M.; Gonzalez, F. J. *TIPS,* **2012**, 33, 323-330.

[21] Fiorucci, S.; Distrutti, E.; Bifulco, G.; D'Auria, M. V.; Zampella, A. *TIPS*, **2012,** 33, 591-601.

[22] Wentworth, J. M.; Agostini, M.; Love, J.; Schwabe, J. W.; Chatterjee, V. K. K. *J. Endocrinol.,* **2000**, *166*, R11-16.

[23] Moore, L. B.; Goodwin, B.; Jones, S. A.; Wisely, G. B.; Serabjit-Singh, C. J.; Willson, T. M.; Collins, J. L.; Kliewer, S. A. *Proc. Natl. Acad. Sci. U.S.A*, **2000**, *97*, 7500-7502.

[24] Chang, T. K. H. *The AAPS,* **2009**, *11* (3), 590-601.

[25] Lau, A. J.; Yang, G.; Rajaraman, G.; Baucom, C. C.; Chang T. K. H. *JPET,* **2010**, 335, 771-780.

[26] Zhou, C.; Poulton, E. J.; Grun, F.; Bammler, T. K.; Blumberg, B.; Thummel, K. E.; Eaton, D. L. *Mol. Pharmacol*., **2007**, *71*, 220-229.

[27] Synold, T. W., Dussault, I.; Forman, B. M. *Nat. Med.,* **2001**, *7***,** 584-590.

[28] Fiorucci, S.; Mencarelli, A.; Distrutti, E.; Palladino, G.; Cipriani S. *Curr. Med. Chem*., **2010**, *17,* 139-159.

[29] Fiorucci, S.; Rizzo, G.; Donini, A.; Distrutti, E.; Santucci, L. *Trends Mol. Med.,* **2007**, *13*, 298-309.

[30] Modica, S.; Bellafante, E.; Moschetta, A. *Front Biosci.,* **2009**, *14*, 4719-4745.

[31] Trauner, M.; Fickert, P.; Halilbasic, E.; Moustafa, T. *Wien. Med. Wochenschrift,* **2008**, *158*, 542-548.

[32] Huber, R. M.; Murphy, K.; Miao, B.; Link, J. R., Cunningham, M. R.; Rupar, M. J.; Gunyuzlu, P. L.; Haws, T. F. J.; Kassam, A.; Powell, F.; Hollis, G. F.; Young, P. R.; Mukherjee, R.; Burn, T. C. *Gene*, **2002**, *290*, 35-43.

[33] Moore, D. D.; Kato, S.; Xie, W.; Mangelsdorf, D. J.; Schmidt, D. R.; Xiao, R.; Kliewer, S. A. *Pharmacol. Rev.*, **2006**, *58*, 742-759.

[34] Forman, B. M.; Goode, E.; Chen, J.; Oro, A. E.; Bradley, D. J.; Perlmann, T.; Noonan, D. J.; Burka, L. T.; McMorris, T.; Lamph, W. W.; Evans, R. M.; Weinberger, C. *Cell*, **1995**, *81*, 687-693.

[35] Wang, H.; Chen, J.; Hollister, K.; Sowers, L. C.; Forman, B. M. *Mol. Cell.,* **1999**, *3*, 543-553.

[36] Makishima, M.; Okamoto, A. Y.; Repa, J. J.; Tu, H.; Learned, R. M.; Luk, A.; Hull, M. V.; Lustig, K. D.; Mangelsdorf, D. J.; Shan, B. *Science*, **1999**, *284*, 1362-1365.

[37] Parks, D. J.; Blanchard, S. G.; Bledsoe, R. K.; Chandra, G.; Consler, T. G.; Kliewer, S. A.; Stimmel, J. B.; Willson, T. M.; Zavacki, A. M.; Moore, D. D.; Lehmann, J. M. *Science*, **1999**, *284*, 1365-1368.

[38] Mi, L.-Z.; Devarakonda, S.; Harp, J. M.; Han, Q.; Pellicciari, R.; Willson, T. M.; Khorasanizadeh, S.; Rastinejad, F. *Mol. Cell.,* **2003**, *11*, 1093-1100.

[39] Maloney, P. R.; Parks, D. J.; Haffner, C. D.; Fivush, A. M.; Chandra, G.; Plunket, K. L.; Creech, K. D.; Moore, L. B.; Wilson, J. G; Lewis, M. C.; Jones, S. A.; Willson, T. M. *J. Med. Chem.*, **2000**, *43*, 2971-2974.

[40] Fiorucci, S.; Mencarelli, A.; Distrutti, E.; Zampella A. *Future Medicinal Chemistry,* **2012**, 4, 877-891.

[41] Pellicciari, R.; Fiorucci, S.; Camaioni, E.; Clerici, C.; Costantino, G.; Maloney, P. R.; Morelli, A.; Parks, D. J.; Willson, T. M. *J. Med. Chem.*, **2002**, *45*, 3569-3572.

[42] Nicolaou, K. C.; Evans, R. M.; Roecker, A. J.; Hughes, R.; Downes, M.; Pfefferkorn, J. A. *Org. Biol. Chem.,* **2003**, *1*, 908-920.

[43] Singh, R. B.; Niaz, M. A.; Ghosh, S. *Drugs Ther.,* **1994**, *4*, 659-664.

[44] Stevens J. F.; Page J. E. *Phytochemistry*, **2004**, *65*, 1317-1330.

[45] Szapary, P. O.; Wolfe, M. L.; Bloedon, L. T.; Cucchiara, A. J.; Der Marderosian, A. H.; Cirigliano, M. D.; Rader, D. J. *J. Am. Med. Assoc*., **2003**, *290*, 765-772.

[46] Burris, T. P.; Montrose, C.; Houck, K. A.; Osborne, H. E.; Bocchinfuso, W. P.; Yaden, B. C.; Cheng, C. C; Zink, R. W.; Barr, R. J.; Hepler, C. D.; Krishnan, V.; Bullock, H. A.; Burris, L. L.; Galvin, R. J.; Bramlett, K.; Stayrook, K. R. *Mol. Pharmacol*., **2005**, *67*, 948-954.

[47] Lederer, E. *Quart. Rev.* **1969**, *23*, 453.

[48] Idler, D. R.; Wiseman, P. M.; Safe, L. M. *Steroids*, **1970**, *16*, 451.

[49] Ferezon, J. P.; Devys, M.; Allois, J. P.; Barbier, M. *Phytochemistry*, **1974**, *13*, 593.

[50] Fattorusso, E.; Magno, S.; Mayol, L.; Santacroce, C.; Sica, D. *Tetrahedron*, **1975**, 31, 1715–6.

[51] Ker, G.; Kerr, S. L.; Pettit, G. R.; Herald, D. L.; Groy, T. L.; Djerassi C. *J. Org. Chem.*, **1991**, *52*, 58–62.

[52] Aiello, A.; Fattorusso, E.; Menna, M.; Carnuccio, R.; Iuvone, T. *Steroids*, **1995**, *60*, 666–73.

[53] Hahn, S.; Stoilov, I. L.; Raedersdorff, D.; Djerassi, C. *J. Am. Chem. Soc.*, **1988**, *110*, 8117–24.

[54] Lam, W. K.; Hahn, S.; Ayanoglu, E.; Djerassi, C. *J. Org. Chem.*, **1989**, *54*, 3428–32.

[55] Djerassi, C, Lam WK. *Acc. Chem. Res*., **1991**, *24*, 69–75.

[56] Burgoyne, D. L.; Andersen, R. J,; Allen, T. M. *J. Org. Chem.*, **1992**, *57*, 525–8.

[57] Giner, J. L. *Chem Rev,* **1993**, *93*, 1735–52.

[58] Capon, R. J.; Faulkner, D. J. *J. Org. Chem.*, **1985**, *50*, 4771–3.

[59] Kerr, R. G.; Kerr, S. L.; Malik, S.; Djerassi, C. *J. Am. Chem. Soc.,* **1992**, *114*, 299 –303.

[60] Fusetani, N.; Matsunaga, S.; Konosu, S. *Tetrahedron Lett.*, **1981**, *21*, 1985-1988.

[61] Wegerski, C. J.; Hammond, J.; Tenney, K.; Matainaho, C.; Crews, P. *J. Nat. Prod*., **2007**, *70*, 89-94.

[62] Kitagawa, I.; Kobayashi, M.; Katori, T.; Yamashita, M.; Tanaka, J.; Doi, M.; Ishida, T. *J. Am Chem Soc*, **1990**, 112, 3710-3712.

[63] Kobayashi, M.; Tsukamoto, S.; Tanabe, A.; Sakai, T.; Ishibashi, M. J.;, *T. J. Chem. Soc. Perkin trans.*, **1991**, 1, *112*, 2379-2389.

[64] Ratnayake, A. S.; Davis, R. A.; Harper, M. K.; Veltri, C. A.; Andjelic, C. D.; Barrows, L. R.; Ireland, C. M., *J. Nat. Prod*., **2005**, *68*, 104-107.

[65] Kho, E.; Imagawa, D. K.; Rohmer, M.; Kashman, Y.; Djerassi, C., *J. Org. Chem*., **1981**, *46*, 1836-1839.

[66] Hamada, T.; Sugawara, T.; Matsunaga, S.; Fusetani, N., *Tetrahedron Let*., **1994**, *35*, 609-612.

[67] Bewley, C. A.; Faulkner, D. J. *Angew. Chem. Int. Ed*., **1998**, *37*, 2162-2178.

[68] Andrianasolo, E. H.; Gross, H.; Goeger, D.; Musafija-Girt, M.; McPhail, K.; Leal, R. M.; Mooberry, S. L.; Gerwick, W. H., *Org. Lett*., **2005**, *7*, 1375-1378.

[69] Piel, J.; Butzke, D.; Fusetani, N.; Hui, D.; Platzer, M.; Wen, G.; Matsunaga, S., *J. .Nat. Prod*., **2005**, *68*, 472-479.

[70] Fusetani, N.; Sugawara, T.; Matsunaga, S., *J. Am. Chem. Soc*., **1991**, *113*, 7811-7812.

[71] Tsukamato, S.; Matsunaga, S.;Fusetani, N.; Toh-E, A., *Tetrahedron*, **1999**, *55*, 13697-13702.

[72] Kupchan, S. M., Britton, R. W., Ziegler, M. F., Sigel, C. W., *J. Org. Chem*., **1973**, *38*, 178.

[73] Festa, C.; De Marino, S.; D'Auria, M. V.; Bifulco, G.; Renga, B.; Fiorucci, S.; Petek, S.; Zampella, A., *J. Med. Chem*. **2011**, *54*, 401-405.

[74] Festa, C.; De Marino, S.; Sepe, V.; Monti, M. C.; Luciano, P.; D'Auria, M. V.; Debitus, C.; Bucci, M.; Vellecco, V.; Zampella, A., *Tetrahedron,* **2009**, *65*,10424-10429.

[75] Festa, C.; De Marino, S.; Sepe, V.; D'Auria, M. V.; Bifulco, G.; Debitus, C.; Bucci, M.; Vellecco, V.; Zampella, A. *Org. Lett.* **2011**, *13 (6)*, 1532–1535.

[76] Koehn, F. E.; Gunasekera, M.; Cross, S. S. *J. Org. Chem*., **1991**, *56*, 1322-1325.

[77] Sun, H. H.; Cross, S. S.; Gunasekera, M.; Koehn, F. E. *Tetrahedron*, **1991**, *47*, 1185- 1190.

[78] Yang, S. W.; Chan, T. M.; Pomponi, S. A.; Chen, G.; Loebenberg, D.; Wright, A.; Patel, M.; Gullo, V.; Pramanik, B.; Chu, M. *J. Antib*., **2003**, *56*, 186-189.

[79] Nakatsu, T.; Walker, R. P.; Thompson, J. E.; Faulkner, D. J. *Experientia*, **1983**, *39*, 759-761.

[80] Gunasekera, S. P.; Sennett, S. H.; Kelly-Borges, M.; Bryant, R. W. *J. Nat.Prod*., **1994**, *57*, 1751-1754.

[81] Fujita, M.; Nakao, Y.; Matsunaga, S.; Seiki, M.; Itoh, Y.; van Soest, R. W. M.; Heubes, M.; Faulkner, D. J.; Fusetani, N. *Tetrahedron*, **2001**, *57*, 3885-3890.

[82] Aoki, S.; Naka, Y.; Itoh, T.; Furukawa, T.; Rachmat, R.; Akiyama, S.-i.; Kobayashi, M. *Chem. Pharm. Bull*., **2002**, *50*, 827-830.

[83] Yang, S.-W.; Buivich, A.; Chan, T.-M.; Smith, M.; Lachowicz, J.; Pomponi, S. A.; Wright, A. E.; Mierzwa, R.; Patel, M.; Gullo, V.; Chu, M. *Bioorg. Med. Chem. Lett*., **2003**, *13*, 1791-1794.

[84] Ma, X.; Shah, Y. M.; Guo, G. L.; Wang, T.; Krausz, K. W.; Idle, J. R.; Gonzalez, F. J. *J. Pharmacol. Exp. Ther.* **2007**, *322*, 391–398.

[85] Youssef, D. T.; Ibrahim, A. K.; Khalifa, S. I.; Mesbah, M. K.; Mayer, A. M.; Van Soest, R. W. *Nat. Prod. Commun.* **2010**, *5*, 27–31.

[86] Rock, K. L.; Latz, E.; Ontiveros, F.; Kono, H. *Annu. Rev. Immunol.* **2010**, *28*, 321–342.

[87] Trott, O.; Olson, A. J. *J. Comput. Chem*., **2010**, *31*, 455-461.

[88] Ekins, S.; Kortagere, S.; Iyer, M.; Reschly, E. J.; Lill, M. A.; Redinbo, M. R.; Krasowski, M. D. *PLoS Comput. Biol*., **2009**, *5*, e1000594.

[89] Sepe, V.; Ummarino, R.; D'Auria, M. V.; Mencarelli, A.; D'Amore, C.; Renga, B.; Zampella, A.; Fiorucci, S. *J. Med. Chem.*, **2011**, *54*, 4590-4599.

[90] Iida, T.; Kakiyama, G.; Hibiya, Y.; Miyata, S.; Inoue, T.; Ohno, K.; Goto, T.; Mano, N.; Junichi Goto, J.; Nambara, T.; Hofmann, A. F. *Steroids* **2006**, *71*, 18-29.

[91] Iida, T.; Momose, T.; Tamura, T.; Matsumoto, T.; Chang, F. C.; Goto, J.; Nambara, T. *J. Lipid. Res.* **1988**, *29*, 165-171.

[92] MacNevin, C. J.; Atif, F.; Sayeed, I.; Stein, D. G.; Liotta, D. C. *J. Med. Chem.* **2009**, *52*, 6012-6023.

[93] Santos, G. A. G.; Murray, A. P.; Pujol, C. A.; Damonte, E. B.; Maier, M. S. *Steroids* **2003**, *68*, 125-132.

[94] Cong, R.; Zhang, Y.; Tian, W. *Tetrahedron Lett.* **2010**, *51*, 3890-3892.

[95] Voigt, B.; Porzel, A.; Adam, G.; Golsch, D.; Adam, W.; Wagner, C.; Merzweiler, K. *Periplaneta americana Collect. Czech. Chem. Commun*. **2002**, *67*, 91-102.

[96] Fiorucci, S.; Mencarelli, A.; Palazzetti, B.; Sprague, A. G.; Distrutti, E.; Morelli, A.; Novobrantseva, T. I.; Cirino, G.; Koteliansky, V. E.; de Fougerolles, A. R. *Immunity* **2002**, *17*, 769-780.

[97] Mencarelli, A.; Renga, B.; Palladino, G.; Distrutti, E.; Fiorucci, S. *Biochem. Pharmacol.* **2009**, *78*, 1214-1223.

[98] Rimoldi, M.; Chieppa, M.; Salucci, V.; Avogadri, F.; Sonzogni, A.; Sampietro, G. M.; Nespoli, A.; Viale, G.; Allavena, P.; Rescigno, M. *Nat. Immunol.* **2005**, *6*, 507-514.

[99] Rousseaux, C.; Lefebvre, B.; Dubuquoy, L.; Lefebvre, P.; Romano, O.; Auwerx, J.; Metzger, D.; Wahli, W.; Desvergne, B.; Naccari, G. C.; Chavatte, P.; Farce, A.; Bulois, P.; Cortot, A.; Colombel, J.; Desreumaux, F. P. *J. Exp. Med.*, **2005**, *201*,1205-1215.

[100] Sepe, V.; Ummarino, R.; D'Auria, M. V.; Lauro, G.; Bifulco, G.; D'Amore, C.; Renga, B.; Fiorucci, S.; Zampella, A. *Org. Biomol. Chem.*, **2012**, *10*, 6350-6362.

[101] Kunishima, M.; Kawachi, C.; Hioki, K.; Terao, K.; Tani, S. *Tetrahedron*, **2001**, *57*, 1551-1558.

[102] Watkins, R. E.; Wisely, G. B.; Moore, L. B.; Collins, J. L.; Lambert, M. H.; Williams, S. P.; Willson, T. M.; Kliewer, S. A.; Redinbo, M. R. *Science*, **2001**, *292*, 2329-2333.

[103] Santos, G. A. G.; Murray, A. P.; Pujol, C. A.; Damonte, E. B.; Maier, M. S. *Steroids* **2003**, *68*, 125-132.

[104] Wang, K.; Damjanov, I.; Wan, Y. J. *Lab. Invest.*, **2010**, *90*, 257-265.

[105] Renga, B.; Mencarelli, A.; Migliorati, M.; Cipriani, S.; D'Amore, C.; Distrutti, E.; Fiorucci, S. *Inflamm. Res.*, **2011**, *60*, 577-587.

[106] Fiorucci, S.; Antonelli, E.; Rizzo, G.; Renga, B.; Mencarelli, A.; Riccardi, L.; Orlandi, S.; Pellicciari R.; Morelli, A. *Gastroenterology*, **2004**, *127*, 1497-1512.

[107] Fiorucci, S.; Antonelli, E.; Distrutti, E.; Severino, B.; Fiorentina, R.; Baldoni, M.; Caliendo, G.; Santagada, V.; Morelli A.; Cirino, G. *Hepatology*, **2004**, *39*, 365-375.

[108] Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.*, **2009**, *30*, 2785-2791.

[109] Hofmann, A. F.; Zakko, S. F.; Lira, M. ; Clerici, C.; Hagey, L. R.; Lambert, K. K.; Steinbach, J. H.; Schteingart, C. D.; Olinga, P.; Groothuis, G. M. *Hepatology* **2005**, *42*, 1391–1398.

[110] Sepe, V.; Ummarino, R.; D'Auria, M. V.; Renga, B.; Fiorucci, S.; Zampella A. *Eur. J.O.C.,* **2012**, 5187–5194.

[111] Tserng, K. Y.; Klein, P. D. *Steroids* **1977**, *29*, 635–648.

[112] Schteingart, C. D.; Hofmann, A. F. *J. Lipid Res.* **1988**, *29*, 1387–1395.

[113] Ma, X.; Shah, Y. M.; Guo, G. L.; Wang, T.; Krausz, K. W.; Idle, J. R.; Gonzalez, F. J. *J. Pharmacol. Exp. Ther.* **2007**, *322*, 391–398.

[114] Cheng, J.; Shah, Y. M.; Ma, X.; Pang, X.; Tanaka, T.; Kodama, T.; Krausz, K. W.; Gonzalez, F. J. *J. Pharmacol. Exp. Ther.* **2010**, *335*, 32-41.

[115] Sanner, M. F. *J. Mol. Graphics Modell.*, **1999**, *17*, 57-61.

[116] Zhang, H. J.; Yi, Y. H.; Lin, H. W. *Helv. Chim. Acta* **2010**, *93*, 1120-1126.

[117] Angawi, R. F.; Calcinai, B.; Cerrano, C.; Dien, H. A.; Fattorusso, E.; Scala, F.; Taglialatela-Scafati, O. *J. Nat. Prod.* **2009**, *72*, 2195-2198.

[118] Kobayashi, M.; Kawazoe, K.; Katori, T.; Kitagawa, I. *Chem. Pharm. Bull.* **1992**, *40*, 1773-1778.

[119] Qureshi, A.; Faulkner, D. J. *J. Nat. Prod.* **2000**, *63*, 841-842.

[120] Sugo, Y.; Inouye, Y.; Nakayama, N. *Steroids* **1995**, *60*, 738-742.

[121] De Marino, S.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; Renga, B.; D'Amore, C.; Fiorucci, S.; Debitus, C.; Zampella, A. *J. Med. Chem.,* **2011**, *54*, 3065-75.

[122] Wright, J. L. C.; Mclnnes, A. G.; Shimizu, S.; Smith, D. G.; Walter, J. A.; Idler, D.; Khalil, W. *Can. J. Chem.* **1978**, *56*, 1898-1903.

[123] Horibe, I.; Nakai, H.; Sato, T.; Seo, S.; Takeda, K.; Takatsuto, S. *J. Chem. Soc. Perkin 1 Trans. 1*, **1989**, 1957-1967.

[124] Wang, F.; Fang, Y.; Zhang, M.; Lin, A.; Zhu, T.; Gu, Q.; Zhu, W. *Steroids* **2008**, *73*, 19-26.

[125] Rubinstein, I.; Goad, L. J.; Clague, A. D. H.; Mulheirn, L. J. *Phytochemistry* **1976**, *15*, 195-200.

[126] Fujino, T.; Une, M.; Imanaka, T.; Inoue, K.; Nishimaki-Mogami, T. *J. Lipids Res.* **2004**, *45*, 132-138.

[127] Modica, S.;. Gadaleta, R. M; Moschetta, A. *NRS* **2010**, *8*, 1-28.

[128] Downes, M.; Verdecia, M. A.; Roecker, A. J.; Hughes, R.; Hogenesch, J. B.; Kast-Woelbern, H. R.; Bowman, M. E.; Ferrer, J. –L.; Anisfeld, A. M.; Edwards, P. A.; Rosenfeld, J. M.; Alvarez, J. G. A.; Noel, J. P.; Nicolaou, K. C; Evans; R. M. *Mol. Cell.* **2003**, *11*, 1079-1092.

[129] Kuipers, F.; Claudel, T.; Sturm, E.; Staels, B. *Reviews in Endocrine & Metabolic Disorders* **2004**, *5*, 319-326.

[130] Watkins, R. E.; Maglich, J. M.; Moore, L. B.; Wisely, G. B.; Noble, S. M.; Davis-Searles, P. R.; Lambert, M. H.; Kliewer, S. A.; Redinbo. M. R. *Biochemistry* **2003**, *42*, 1430-1438.

[131] (a) Type I (previously identified by R. W. M. van Soest): van Soest R.W. M. *Neth J Sea Res* **1989**, *23*, 223-30; (b) Type II (previously identified by R. W. M. van Soest): Hamada, T.; Matsunaga, S.; Yano, S.; Fusetani, N. *J. Am. Chem. Soc.* **2005**, *127*, 110-8; (c) Type III (previously identified by M. K. Haper): Ratnayake, A. S.; Davis, R. A.; Harper, M. K.; Veltri, C. A.; Andjelic, C. D.; Barrows, L. R.; Ireland, C. M. *J. Nat. Prod.* **2005,** *68*, 104-7.

[132] De Marino, S.; Sepe, V.; D'Auria, M. V.; Bifulco, G.; Renga, B.; Petek, S.; Fiorucci, S.; Zampella, A. *Org Biomol Chem* **2011**, *9*, 4856-62.

[133] Umeyama, A.; Shoji, N.; Enoki, M.; Arihara, S. *J Nat Prod* **1997**, *60*, 296-8.

[134] De Marino, S.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; D'Amore, C.; Renga, B.; Mencarelli, A.; Petek, S.; Fiorucci, S.; Zampella, A. *Steroids,* **2012**, *77*, 484–495.

[135] Sepe, V.; Bifulco, G.; Renga, B.; D'Amore, C.; Fiorucci, S.; Zampella, A. *J Med Chem* **2011**, *54*, 1314-20.

[136] Fiorucci, S.; Cipriani, S.; Mencarelli, A.; Renga, B.; Distrutti, E.; Baldelli, F. *Curr Mol Med* **2010**, *10*, 579–95.

[137] Sinal, C. J.; Tohkin, M.; Miyata, M.; Ward, J. M.; Lambert, G.; Gonzalez, F. J. *Cell **2000***, *102*, 731-44.

[138] Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—*J. Comput. Chem.* **2004**, *25*, 1605–1612.

[139] Fiorucci, S.; Cipriani, S.; Baldelli, F.; Mencarelli, A. *Prog. Lipid Res.* **2010**, *49*, 171-185.

[140] Fiorucci, S.; Mencarelli, A.; Palladino, G.; Cipriani, S. *Trends Pharmacol. Sci*. **2009**, *30*, 570-580.

[141] Fiorucci, S.; Baldelli, F. *Curr. Opin. Gastroenterol.* **2009**, *25*, 252-259.

[142] Fiorucci, S.; Rizzo, G.; Donini, A.; Distrutti, E.; Santucci, L. *Trends Mol. Med.* **2007**, 13, 298-309.

[143] Fiorucci, S; Cipriani, S; Mencarelli, A; Baldelli, F; Bifulco, G; Zampella, A. *Mini Rev. Med. Chem.* **2011**, *11(9)*, 753-762.

[144] Nozawa, H. *Biochem. Biophys. Res. Comm.* **2005**, *336*, 754-761.

[145] Nam, S. J.; Ko, H.; Shin, M.; Ham, J.; Chin, J.; Kim, Y.; Kim, H.; Shin, K.; Choi, H.; Kang, H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5398-5402.

[146] Choi, H.; Hwang, H.; Chin, J.; Kim, E.; Lee, J.; Nam, S. J.; Lee, B. C.; Rho, B. J.; Kang, H. *J. Nat. Prod.* **2011**, *74*, 90-94.

[147] Sepe, V.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; Renga, B.; D'Amore, C.; Debitus, C.; Fiorucci, S.; Zampella, A. *J. Med. Chem.* **2012**, *55*, 84-93.

[148] Pellicciari, R.; Costantino, G.; Camaioni, E.; Sadeghpour, B. M.; Entrena, A.; Willson, T. M.; Fiorucci, S.; Clerici, C.; Gioiello, A. *J. Med. Chem.* **2004**, *47*, 4559-4569.

[149] Gioiello, A.; Macchiarulo, A.; Carotti, A.; Filipponi, P.; Costantino, G.; Rizzo, G.; Adorini, L.; Pellicciari, R. *Bioorg. Med. Chem.* **2011**, *19*, 2650-2658.

[150] Jonker, J. W.; Liddle, C.; Downes, M. *J. Steroid Biochem. Mol. Biol.* **2011**, doi:10.1016/j.jsbmb. 2011.06.012.

[151] Stedman, C.; Liddle, C.; Coulter, S.; Sonoda, J.; Alvarez, J. G. et al. *Proc. Natl. Acad. Sci. USA* **2006,** *103*(*30*), 11323–11328.

[152] Stedman, C. A.; Liddle, C.; Coulter, S. A.; Sonoda, J.; Alvarez J. G. et al. *Proc. Natl. Acad. Sci. USA* **2005,** *102*, 2063–2068.

[153] Renga, B.; Mencarelli, A.; D'Amore, C.; Cipriani, S.; D'Auria, M. V.; Sepe, V.; Chini, M. G.; Monti, M. C.; Bifulco, G.; Zampella, A.; Fiorucci S. *PloSone,* **2012***,* 7, e30443.

[154] Sepe, V.; Ummarino, R.; D'Auria, M. V.; Taglialatela Scafati, O.; D'Amore, C.; Renga, B.; Nakao, Y.; Fusetani, N.; Fiorucci, S.; Zampella, A. *Mar. Drugs* **2012**, *10*, 2448-2466.

[155] Schmidt, A. W.; Doert, T.; Goutal, S.; Gruner, M.; Mende, F.; Kurzchalia, T. V.; Knölker, H. J. *Eur. J. Org. Chem.* **2006**, *16*, 3687–3706.

[156] Gulavita, N. K.; Pomponi, S. A.; Wright, A. E.; Yarwood, D.; Sills, M. A. *Tetrahedron Lett*., **1994**, *35*, 6815-18.

[157] Posadas, I.; Bucci, M.; Roviezzo, F.; Rossi, A.; Parente, L.; Sautebin, L.; Cirino, G. *Br. J. Pharmacol.* **2004**, *142*, 331-338.

[158] Sepe, V.; D'Auria, M. V.; Bifulco, G.; Ummarino, R.; Zampella, A. *Tetrahedron* **2010**, *66*, 7520-7526.

[159] Matsumori, N.; Kaneno, D.; Murata, M.; Nakamura, H.; Tachibana, K. *J. Org. Chem.* **1999**, *64*, 866-876.

[160] Kurz, M.; Schmieder, P.; Kessler, H. *Angew. Chem., Int. Ed. Engl.* **1991**, *10*, 1329-1331.

[161] Bifulco, G.; Bassarello, C.; Riccio, R.; Gomez-Paloma, L. *Org. Lett.* **2004**, *6*, 1025-1028.

[162] Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. *Chem. Rev.* **2007**, *107*, 3744-3779.

[163] Di Micco, S.; Chini, M. G.; Riccio, R.; Bifulco, G. *Eur. J. Org. Chem.* **2010**, 1411-1434.

[164] Tomasi, J.; Mennucci, B.; Cances, E. *THEOCHEM* **1999**, *464*, 211-226.

[165] Bonini, C.; Righi, G. *Tetrahedron* **2002**, *58*, 4981-5021.

[166] Mancuso, A. J.; Swern, D. *Synthesis* **1981**, 165-185.

[167] Zampella, A.; D'Orsi, R.; Sepe, V.; Casapullo, A.; Monti, M. C.; D'Auria, M. V. *Org. Lett.* **2005**, *7*, 3585-3588.

[168] Fringuelli, F.; Pizzo, F.; Rucci, M.; Vaccaro, L. *J. Org. Chem.* **2003**, *68*, 7041-7045.

[169] Righi, G.; Rumboldt, G.; Bonini, C. *J. Org. Chem.* **1996**, *61*, 3557-3560.

Aknowledgements

I would like to express my sincere gratitude to the following people, whose contribution in assorted ways to the research deserves special mention:

Professor Angela Zampella, my supervisor, who has been a great intellectual support with her expertise and research insight. She gave me a constant encouragement and enriched considerably my Ph.D. experience;

Professor Maria Valeria D'Auria for her constant contribution and for involving me with enthusiasm in the projects carried out;

Professor Stefano Fiorucci (Dipartimento di Medicina Clinica e Sperimentale, Università di Perugia, Perugia, Italy) and his research group for pharmacological studies in vitro and in vivo;

Professor Giuseppe Bifulco (Dipartimento di Scienze Farmaceutiche e Biomediche, Università di Salerno, Fisciano (SA), Italy) and his research group for docking studies;

Professor Franco Zollo, for his generous support and academic guidance;

Dr. Simona De Marino, for her collaboration, patience, helpfulness and useful advice;

Dr. Valentina Sepe, valuable guidance and support during my research work, for having answers to all my questions;

Dr. Carmen Festa and Chiara Del Gaudio for helping me during PhD period, for nice time spent together and their sincere friendship;

Dr. Gerardo Della Sala, Sveva Pelliccia and Stefano De Tito with whom I shared this Ph.D. experience;

All the members of Dipartimento di Chimica delle Sostanze Naturali and all students, who spent a period in laboratories N20 and N22, with love I will remember all of them.

I wish to thank also my entire family for providing a loving environment for me. Thank to my mother and my father for their unconditional support and attentive care. My brothers, Gennaro and Angelo, and my cousin Giusy who have always motivated me and shared with me the challenges of the life.

And last but not least I want to thank my boyfriend, Giuseppe, for his love, encouragement and patience during these years.

Above all, a thought and a thanks to you, my strength. To you I dedicate this thesis.

Thanks all

Raffaella Ummarino

Is pleased to recognize

Simona De Marino, Raffaella Ummarino, Maria Valeria D'Auria, Maria Giovanna Chini, Giuseppe Bifulco, Barbara Renga, Claudio D'Amore, Stefano Fiorucci, Cécile Debitus, and Angela Zampella

For the

Highly Cited Article of 2011

Theonellasterols and Conicasterols from Theonella swinhoei. Novel Marine Natural Ligands for Human Nuclear Receptors

Journal of Medicinal Chemistry, Vol 54, *Issue 8*, 2011

Shaomeng Wang
Editor-in-Chief

February 21, 2013

ACS Publications
MOST TRUSTED. MOST CITED. MOST READ.

Gunda Georg
Editor-in-Chief