Drug Affinity Responsive Target Stability (DARTS) Identifies Laurifolioside as a New Clathrin Heavy Chain Modulator

Fabrizio Dal Piaz[†], Mariela Beatriz Vera Saltos[‡], Silvia Franceschelli[†], Giovanni Forte[†], Stefania Marzocco[†], Tiziano Tuccinardi[§], Giulio Poli[§], Samad Nejad Ebrahimi[⊥], Matthias Hamburger[‖], Nunziatina De Tommasi[*,†], and Alessandra Braca[§]

[†]*Dipartimento di Farmacia, Università di Salerno, Via Giovanni Paolo II 132, 84084 Fisciano (SA), Italy*

[‡]*Departamento de Ciencias de la Vida, Universidad de las Fuerzas Armadas, ESPE, Av. General Rumiñahui s/n, Sangolqui, Ecuador*

[§]*Dipartimento di Farmacia, Università di Pisa, Via Bonanno 33, 56126 Pisa, Italy*

[⊥]*Department of Phytochemistry, Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, G. C., Evin, Tehran, Iran*

[‖] *Division of Pharmaceutical Biology, Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland*

ABSTRACT: Five new diterpenes (**1**-**5**) and a megastigmane derivative (**6**) were isolated from the aerial parts of *Euphorbia laurifolia*, along with several known compounds. Their structures were elucidated by NMR, MS and ECD, and by chemical methods. A chemical proteomics Drug Affinity Responsive Target Stability (DARTS) approach to investigate the lathyrane diterpene **1**, laurifolioside, on its putative cellular target(s) was performed. Clathrin heavy chain 1, a protein mainly involved in selective uptake of proteins, viruses, and other macromolecules at the plasma membrane of cells, was identified as the major interaction partner of compound **1**. The modulation of clathrin activity by **1** was studied through microscopy, molecular docking, and molecular dynamics studies, suggesting a new activity of lathyrane diterpenes in the modulation of trafficking pathways.

Over the last two decades, the discovery of macrocyclic diterpenes in *Euphorbia* species (Euphorbiaceae) in acting as modulators of multidrug resistance (MDR) has stimulated research on the phytochemical study of plants belonging to this genus. *Euphorbia* species biosynthesize skin-irritant diterpenes, such as ingenane, tigliane, and daphnane derivatives, together with non-irritant polyoxygenated macrocyclic and polycyclic diterpenoids.[1,2] Besides MDR inhibitory properties, macrocyclic diterpenes from *Euphorbia* species have been shown to possess antitumor, anti-inflammatory, vasorelaxant, neuroprotective, molluscicidal, antiviral, and antimicrobial activities.[3,4] MDR occurs in blood cancers and many solid tumors allowing drug-resistant cells to survive and drive tumor growth. This phenomenon is conveyed via several mechanisms, including the overexpression of P-glycoprotein, which serves as an efflux pump for chemotherapeutics and other drugs. In principle, modulators of P-glycoprotein activity could overcome MDR, but the search for therapeutically useful inhibitors has met with serious difficulties. Recently, macrocyclic diterpenes and some polycyclic derivatives from *Euphorbia* species were found to be strong inhibitors of P-glycoprotein and related enzymes.[2,4]

The biological properties of plant secondary metabolites very frequently are not due to the modulation of a single protein or pathway, but some are known to interact with a plethora of cellular components.[5] Thus, a full investigation of the cellular interactome of such compounds is important to better understand their modes of action, and to discover new hit or lead compounds for medicinal chemistry. The way to study the polypharmacological activity of plant molecules is the characterization of their interaction with cellular biomolecules. The discovery of bioactive molecules endowed with interesting pharmacological profiles is of major importance in the medicinal chemistry, molecular medicine, and pharmacology fields. In recent years, a number of approaches have been developed for the target identification of small molecules, of which the most advanced and widely used are direct and indirect chemical proteomic methods. Direct chemical proteomic analysis comprises all procedures that directly identify proteins, by using their binding to the small molecule ligand. Indirect methods do not characterize bound proteins, but, based on indirect cellular or biochemical perturbation produced by small molecules, identify the protein

target. Drug Affinity Responsive Target Stability (DARTS) is an emerging general methodology for identifying and studying protein-ligand interactions.[6,7] DARTS uses unmodified small molecules, and this is an advantage with respect to affinity chromatography-based proteomic approaches. The technique is based on the principle that binding of a small molecule to a target protein stabilizes the target protein by increasing its resistance to proteases. DARTS is particularly useful for the initial identification of protein targets for small molecules, but can also be used to validate potential protein-ligand interactions predicted or identified by other means, and also to estimate the strength of interactions.

As a part of an investigation on the construction of a diterpene and diterpene-like library to screen against biological targets, *Euphorbia* species were selected, which are well known for their content of this compound class. The phytochemical studies on *Euphorbia laurifolia* Lam. (syn. *Euphorbia latazi* Kunth., *Euphorbiua lehmanniana* Pax) aerial parts led to the isolation of five new diterpenes (**1**-**5**) and one megastigmane derivative (**6**), along with several known compounds. Structural elucidation was conducted by analysis of 1D, 2D NMR and ECD spectroscopic and MS spectrometry data. Previous studies on *E. laurifolia* latex reported the presence of lathyrane diterpenes having acetyl, isobutyryl, and benzoyl moieties.[8,9] Aiming to characterize the lathyrane diterpene target in a cellular system, a DARTS strategy was employed for compound **1**, named laurifolioside, and this showed that clathrin heavy chain 1 is the main partner. The ability of compound **1** to modulate clathrin heavy chain activity was assessed by confocal microscopy, molecular docking, and molecular dynamics analyses, suggesting a new activity of lathyrane diterpenes in the modulation of cellular trafficking pathways.

RESULTS AND DISCUSSION

The aerial parts of *E. laurifolia* were extracted with solvents of increasing polarity. Compounds **1**-**6** were isolated by chromatography on silica gel, Sephadex LH-20, followed by reversed-phase HPLC.

Compounds **1** and **2** were obtained as yellow oils. Their HRESIMS data exhibited molecular ions at *m/z* 501.2439 $[M+Na]^+$ for **1**, and 501.2437 $[M+Na]^+$ for **2**, indicative of a common molecular formula of $C_{26}H_{38}O_8$. The two compounds were thus found to be isomers. A product ion detected in their HRESIMS at *m/z* 317.2104 $[M+H-162]^+$ and 317.2101 $[M+H-162]^+$, respectively, suggested the presence of a hexose unit. The ^1H NMR, ^{13}C NMR (Table 1), and ^{13}C DEPT spectra of **1** indicated, besides signals attributable to a sugar moiety, the presence of two ketones, two double bonds (one trisubstituted and one tetrasubstituted), and 14 sp^3 carbons, including four methyls, five methylenes, four methines, and one quaternary carbon. DQF-COSY, 1D-TOCSY and HSQC experiments of **1** established the connectivities of H-2—H-3 in ring A, and H-7—H-20 in ring B, which were indicative of a lathyrane diterpene.[8] The lathyrane skeleton was confirmed by the presence in the ^1H NMR spectrum of signals at δ 0.56 (1H, br t, J = 9.0 Hz) and 0.77 (1H, ddd, J = 13.0, 8.0, 3.0 Hz), characteristic of a cyclopropane ring, as present in many types of diterpenes in the genus *Euphorbia*.[10] Chemical shifts, signal multiplicities, and *J*-values in the ^1H NMR spectrum, and ^{13}C NMR chemical shifts indicated the presence of a β-glucopyranosyl moiety. The absolute configuration of the sugar moiety was determined to be D by hydrolysis, trimethylsilylation, and GC analysis. The chemical shift assignments of the carbon atoms were established from the HSQC and HMBC spectra. Key HMBC correlations between H_2-3—C-1, H_2-3—C-4, H_2-3—C-15; H_2-7—C-5, H_2-7—C-9, H_2-7—C-17; H-9—C-7, H-9—C-18; H-11—C-9, H-11—C-13; Me-16—C-1, Me-16—C-3; Me-17—C-4, Me-17—C-5; H_2-18—C-1_{glc}; Me-20—C-11, Me-20—C-14, were observed and allowed the locations to be made of the α,β-unsaturated carbonyl groups at C-1 and C-14, the conjugated double bonds at C-4/C-15 and C-5/C-6, and the glucopyranosyl moiety at C-18. The NMR spectra of **2** (Table 1) were almost superimposable with those of **1**. Differences were observed for the chemical shifts of H-2/C-2 (δ 2.62 and 40.9 in **1**, and δ 2.58 and 41.0 in **2**), H-3/C-3 (δ 2.97 and 2.53 and 37.2 in **1**, and δ 2.30 and 3.19 and 37.2 in **2**), and Me-16 (δ 1.22 and 16.6 in **1**, and δ 1.29 and 16.5 in **2**). Therefore, it was possible to hypothesize that the configuration at C-2 of ring A differed between these compounds. The

presence of chromophoric systems (α, β, γ, δ-diene carbonyl and ketone groups) next to stereogenic centers suggested that the configuration possibly could be solved by ECD spectroscopy. The ECD spectra of compounds **1** and **2** showed two negative Cotton effects (CE) at 350 and 202 nm, along with positive CEs at 254 and 227 nm. These CEs were due to π→π* transitions of the α,β-γ,δ-diene carbonyl group. The absolute configurations of **1** and **2** were established by comparison with calculated ECD data. For quantum chemical calculation, the sugar moiety was omitted since it has no influence on the ECD spectra. A conformational search followed by geometrical optimization using density function theory (DFT) with the B3LYP function and 6-31G** indicated the presence of a predominant conformer for the selected stereoisomers as (2*R*,9*S*,10*R*,11*R*,13*S* and 2*S*,9*S*,10*R*,11*R*,13*S*) (Figure 1). Comparison of calculated ECD spectra of these two possible stereoisomers with the experimental spectra of **1** and **2** showed a close match (Figure 1), in particular two positive CEs around 225 and 260 nm, along with a negative CE at 350 nm. The results revealed that the calculated and experimental data were identical for these stereoisomers and changing of the stereochemistry at position C-2 had no influence on the ECD spectra. This is due to the presence of a weak chromophore (ketone group) next to C-2 and the weak n→π* transitions of the ketone group were buried by strong π→π* transition of extended α,β,γ,δ-diene carbonyl group. Also, the calculation of ECD spectra for other stereoisomers resulted in completely different outcomes (Figure S1, Supporting Information). Therefore, a final conclusion was achieved on the basis of NMR and ECD data. The structure of **1** was thus elucidated as (2*R*,9*S*,10*R*,11*R*,13*S*)-18-hydroxy-9α*H*,11α*H*-lathyra-4(15),5(6)-dien-1,14-dione-18-β-D-glucopyranoside, a new natural product, and named laurifolioside. In turn, compound **2** was characterized as (2*S*,9*S*,10*R*,11*R*,13*S*)-18-hydroxy-9α*H*,11α*H*-lathyra-4(15),5(6)-dien-1,14-dione-18-β-D-glucopyranoside and named 2-*epi*-laurifolioside.

Compound **3** gave a molecular formula of $C_{30}H_{42}O_9$, as established from a molecular ion peak at *m/z* 569.2344 $[M+Na]^+$ in the HRESIMS. HRESIMS/MS analysis showed fragment ions at *m/z* 509.2666 $[M+Na-60]^+$, 449.3008 $[M+Na-60-60]^+$, and 389.3351 $[M+Na-60-60-60]^+$ due to consecutive losses of three acetyl groups. The molecular formula and ^{13}C NMR data accounted for

ten degrees of unsaturation, of which seven were attributable to four ester functionalities, two double bonds, and one carbonyl group. Analysis of the NMR spectroscopic data (Table 1) showed that **3** is a tricyclic lathyrane diterpene, and the spectroscopic data of the diterpenoid core matched with those reported for latazienone.[8] The only difference from latazienone was the replacement of the benzoyl group at C-7 by an acetyl group in **3**. Therefore, compound **3** was assigned as 7β,8α,15β-triacetoxy-3β-(2-methylpropanoyloxy)-4α,9α*H*,11α*H*-lathyra-5*E*,12*E*-dien-14-one.

The ^{13}C NMR spectrum of **4** (Table 2) showed 26 carbon resonances, which were identified as four methyls, three methylenes, thirteen methines, and six quaternary carbons. Moreover, the ^{13}C NMR chemical shifts indicated the presence of eleven oxygen-bearing carbon atoms, including one carbonyl, and four sp^2 carbons. The molecular formula of $C_{26}H_{38}O_{10}$ was established from ^{13}C NMR data and the HRESIMS (*m/z* 509.2432 $[M-H]^-$). The HRESIMS/MS fragmentation pattern suggested the presence of a hexose unit (*m/z* 347.3342 $[M-H-162]^-$ and 329.3456 $[M-H-162-18]^-$). The ^{1}H NMR resonances for the aglycone portion revealed the presence of two olefinic protons at δ 5.70 (1H, br d, *J* = 4.0 Hz) and 5.88 (1H, s), two methyl groups bearing a double bond at δ 1.77 (3H, s) and 1.87 (3H, s), two hydroxymethines at δ 3.34, as overlapped signals, and 4.50 (1H, s), and one hydroxymethylene at δ 3.70 and 3.71 (each 1H, d, *J* = 11.0 Hz). Additionally, two methines at δ 0.84 (1H, br t, *J* = 8.5, 4.2 Hz) and 0.95 (overlapped signal) attributable to a cyclopropane ring were observed. DQF-COSY, 1D-TOCSY, and HSQC spectroscopic data indicated that **4** is an ingenol derivative.[11] In the HMBC spectrum, H-3 showed correlations with C-1, C-5, and C-10, H-7 with C-8, C-9 and Me-20, H-11 with C-9, C-13, and Me-18. Correlations of Me-20 with C-5, C-6, and C-9, and of Me-16 with C-13, C-15, and C-17 established the remaining connectivities. The position of the sugar moiety was assigned at C-3, as deduced from a HMBC cross peak between H-1_{glc} (δ 4.52) and C-3 (90.0 ppm). The relative configuration of the stereogenic centers was established by a ROESY spectrum. Cross peaks between H-8 and H-11 and H_2-17, between H-3 and H-5, and between H-11 and H_2-17 were in agreement with the relative stereochemistry reported for ingenol derivatives.[11,12] The configuration of the glucopyranosyl moiety was determined as in **1**. Consequently, compound **4** was identified as 20-deoxy-16-hydroxyingenol-3-β-D-glucopyranoside.

Compound **5** showed a molecular formula of $C_{20}H_{28}O_3$ (HRESIMS at *m/z* 317.2920 $[M+H]^+$). The ^1H NMR spectrum (Table 2) revealed the presence of four methyls, of which one was linked to a double bond, one olefinic proton, and two hydroxymethines. The ^{13}C NMR spectrum (Table 2) displayed twenty resonances, including four sp^2 carbons and one lactone functionality. With seven degrees of unsaturation, compound **5** was deduced as being tetracyclic. Careful inspection of NMR data revealed signals attributable to a helioscopinolide diterpene.[13,14] The position of a hydroxy group at C-2 was derived from the chemical shifts of C-1, C-2, and C-3, and from HMBC correlations between H-1 and C-2, H-1 and C-3, H-3 and C-1, H-3 and C-2. The relative configuration of the hydroxy group at C-2 was established as 2α, based on the chemical shift and coupling constants of H-2 (δ 3.84, br dd, J = 4.5, 3.5 Hz). Thus, **5** was characterized as 2α-hydroxy-*ent*-abieta-8,13-dien-12,16-olide.

Compound **6** exhibited a molecular formula of $C_{13}H_{22}O_4$, as deduced from the HRESIMS (*m/z* 507.3278) and NMR data. The ^1H and ^{13}C NMR spectra (Table 2) indicated that compound **6** is a megastigmane derivative with a carbonyl group at C-3, a hydroxy group at C-6, and a disubstituted double bond. Elucidation of the skeleton was achieved by analysis of the 2D NMR spectroscopic data. The HMBC spectrum indicated a long-range correlation of H_2-13 with C-4, C-5, and C-6, while H-7 exhibited a long-range correlation with C-6 and C-9. Hence, a hydroxymethylene function could be located at C-5, and a double bond between carbons C-7 and C-8. Thus, compound **6** was identified as deglucosyl lauroside B. The glucoside has been previously isolated from *Laurus nobilis* leaves.[15]

In addition, latazienone,[8] *ent*-16α,17-dihydroxykauran-3-one,[16] vomifoliol,[17] and *ent*-16α,17-dihydroxyatisan-3-one[18] were isolated and characterized as known compounds.

Many studies have been carried out to characterize plant-derived small molecules as MDR modulators. Among these, lathyrane diterpene derivatives were found to be toxic to drug-resistant phenotypes, mainly via modulation of P-glycoprotein in resistant cancer cells. However, the intracellular partner(s) of these natural products was not fully clarified.[2,19,20] The most abundant compound **1** was selected as a candidate molecule for target identification through DARTS

experiments. DARTS involves the incubation of small molecules with cells and/or cell lysates or other complex protein mixtures (without requiring purified proteins), followed by proteolysis, SDS-PAGE, and LC-MS analysis. First, the antiproliferative activity of **1** was assayed in human prostate cancer (PC-3) and human breast adenocarcinoma (MCF-7) cells. The cells were incubated for 24 h with different concentrations of **1** (5-50 μM), and cell viability was determined by the MTT proliferation assay. IC_{50} values of 25.0 ± 0.8 μM for PC-3, and 32.0 ± 0.9 μM for MCF-7 cells were obtained. DARTS experiments were firstly performed incubating PC-3 cell lysates with 20 and 40 μM of compound **1** for 1 h. Following the treatment, subtilisin was added, and, after 30 min of digestion, the reaction mixture was analyzed by SDS-PAGE. A comparison between the gel lanes of treated and control lysates allowed to detect some bands showing different intensities (Figure 2a). MS spectrometry-based examination of these bands permitted to define compound **1** interacting proteins (Table 3A). To study the effect of **1** inside the cells, this procedure was performed again incubating the compound (20 and 40 μM) for 2 h with intact PC-3 cells; following this treatment, cells were lysate under conditions that were not denaturing. The protein extract was subjected to subtilisin digestion and high-resolution LC-MS/MS analysis of the bands that differed in abundance between treated and control cells (Table 3B). Comparing the proteins identified in these two experiments, clathrin heavy chain emerged as a putative target of **1**. The same experiments were also performed on MCF7 cell lines (Figure 2b and Tables 3C and 3D), confirming the interaction between **1** and clathrin heavy chain. This protein has critical roles in intracellular membrane trafficking including endocytosis. It regulates cell surface levels and uptake of plasma membrane proteins such as growth factor receptors, transporters, ion channels, and adhesion proteins. Thus, compound **1** was studied for clathrin heavy chain 1 interaction.

To evaluate the effect of **1** binding to clathrin, Clathrin Mediated Endocytosis (CME) was monitored in PC-3 cells. Towards this aim, fluorescence microscopy was used to observe endocytosis of transferrin conjugates under different experimental conditions. According to the literature,[21] after 15 min of endocytosis, internalized transferrin is localized in the perinuclear region (Figure 3a), while treatment of PC-3 cells with 30 μM of the clathrin inhibitor Pitstop2 led to

an almost complete block in internalization of the transferrin receptor (Figure 3b). Treatment of PC-3 cells with sub-toxic concentrations of compound **1** modified significantly the endocytosis of transferrin (Figures 3d-f), changing both the amount of internalized protein and its cellular localization. Thus, by interacting with clathrin heavy chain 1, the compound affected its activity, confirming that this protein represents a molecular target of **1**.

To rationalize how compound **1** can interact with clathrin heavy chain 1, molecular docking studies followed by molecular dynamic (MD) simulations and binding energy evaluations were conducted. The β-propeller terminal domain (TD) of human clathrin, corresponding to the amino terminus of its heavy chain, was used for these studies, since it is a well-known binding domain for several proteins that interact with clathrin and are necessary for CME. The groove between the first and the second of the seven β-stranded "blades" forming the clathrin β-propeller domain appears to be the principal binding site for such proteins, like clathrin adaptor proteins (APs), and endocytic accessory proteins (EAP) such as amphiphysin, β-arrestines, epsins, and others. A considerable number of proteins involved in the CME process present a well-characterized peptide motif, usually referred to as the "clathrin box", which can bind to this specific site, as was shown by the crystal structures of clathrin TD in complex with the clathrin box-containing peptides of β-arrestin 2 and AP-3 proteins.[22,23] Moreover, X-ray crystallography revealed that small-molecule inhibitors of clathrin interactions with EAPs bind to the same groove of clathrin TD, and thereby compete with clathrin box peptides.[24] These ligands inhibited clathrin association with amphiphysin, synaptojanin 1, the *C*-terminal domain of AP180, and the PH domain of OCRL, with IC_{50} values ranging from 12 to 40 μM, as well as the CME of transferrin and epidermal growth factor. Notably, it was shown that these inhibitors were able of blocking cancer cell proliferation and causing cell death, especially in dividing cells.[25] For these reasons, docking studies were focused on the clathrin box site of the β-propeller TD, and compound **1** was docked into the crystal structure of the human clathrin TD complexed with one of these inhibitors (PDB code 2XZG).[24] For this analysis a robust AUTODOCK[26] procedure was employed that had shown previously good results in a virtual screening study on protein-protein interaction inhibitors.[27] The 200 different docking results

generated were clustered using a root-mean square deviation (RMSD) threshold of 2.0 Å, and the thus obtained four clusters of solutions were considered for further studies (see "Experimental Section" for details). For each cluster, the docking pose associated with the best estimated binding energy was selected as a representative binding mode. The stability of the four different binding modes was then assessed through MD simulation. The MD protocol was set up using the reference X-ray complex of clathrin inhibitor bound to the β-propeller TD, which was then subjected to a total of 30 ns MD simulation. Analysis of the total energy of the system during the simulation showed that after about 0.8 ns the system reached an equilibrium, since it maintained a steady energy value for the remaining 29.2 ns (Figure S2, Supporting Information). The complex remained stable throughout the whole MD simulation, since the ligand maintained its binding mode, and the protein α carbons showed an average RMSD of their position with respect to the crystallographic coordinates of approx. 1.4 Å (Figure S2, Supporting Information). The same MD protocol was then applied to the four clathrin TD-compound **1** complexes predicted by docking. As shown in Figure 4, the protein showed good stability, maintaining its conformation in all MD simulations. After about 10 ns the average RMSD of the protein α carbons was found to be rather constant in all complexes, oscillating around a value between 1.2 and 1.4 Å.

Analysis of the RMSD of the position of the ligand with respect to the initial docking pose highlighted that, although in each complex the ligand showed at least some adjustment of its binding pose due to the protein flexibility, in the case of pose 4, the compound completely lost most of its interactions with the protein after major changes in its binding conformation and orientation. For this reason, pose 4 was considered as not sufficiently reliable as compared to the other binding modes for compound **1**, and was thus discarded. The three remaining binding poses were subjected to binding energy evaluations using the MD trajectories relative to the last 15 ns of simulation. Calculations were carried out using the Molecular Mechanic-Generalized Born surface area (MM-GBSA) and the Molecular Mechanic-Poisson Boltzmann surface area (MM-PBSA) methods.[28] These approaches analyze MD simulation snapshots calculating the contributions of both gas phase and solvation free energies for unbound ligand, unbound protein, and bound complex. The average

contribution of each component was then used to calculate the ligand-protein interaction energy. The analysis identified pose 3 as the most reliable binding mode, since it showed the best binding energy according to both evaluation methods (ΔGBSA = -46.0 kcal/mol; ΔPBSA = -39.5 kcal/mol), and exceeded by about 13 kcal/mol the interaction energies associated with the other poses (Table S1, Supporting Information).

Figure 5 shows the energy minimized average structure of clathrin TD complexed with compound **1** in the predicted binding mode obtained from the last 15 ns of MD simulation. The ligand was sandwiched between the first and the second blade of clathrin β-propeller. Its tricyclic lipophilic portion was placed in the large hydrophobic pocket formed by Val50, Val51, Ile52, Ile62, Ile80, Phe91 and Ile93, and it interacted favorably with the hydrophobic portion of the Arg64 and Lys96 side chains. Interestingly, the carbonyl oxygen at C-14 of the ligand formed a hydrogen bond with Arg64. The glycosidic linker was placed among Ile66, Leu82 and Phe91, where the lipophilic pocket opens to the solvent, whereas the sugar moiety that was partially solvent-exposed took contact with these residues and interacted also with Gln89 and Lys98, forming a hydrogen bond with the latter residue. Although the compound mainly established hydrophobic interactions with the protein, in agreement with the high VDW contribution to its estimated binding energy (Table S1), it also showed two hydrogen bonds with the protein. The sugar portion of the molecule formed several additional hydrogen bonds with the solvent, suggesting the possibility of additional water-bridged interactions with the surrounding residues.

Therefore, five new diterpene derivatives from the aerial parts of *Euphorbia laurifolia* were identified and a DARTS chemical proteomics approach was applied to identify target proteins of laurifolioside (**1**), a new lathyrane derivative. Lathyrane diterpenes have shown previously a wide range of relevant bioactivities, such as antitumor, anti-inflammatory, MDR-reversing, and antiviral properties. In the present experiments, clathrin heavy chain 1 was found to be the binding partner of **1**, confocal microscopy was used to validate the ability of **1** to bind clathrin, and to modulate the clathrin dependent internalization process. Docking and molecular dynamic simulation studies were then performed to rationalize how **1** could interact with clathrin heavy chain. The present findings

suggest that lathyrane diterpenes such as **1** may modulate cell trafficking pathways. On the other hand, the CME machinery is also used by bacteria and viruses for cell entry. The reported antiviral activity for this diterpene class could also be possibly explained by the ability of modulating clathrin activity. Small molecules able to inhibit CME have attracted significant attention recently as a means for disrupting protein complexes and functions. The development of leads able to block this process is still at the beginning, due to an incomplete understanding of CME mechanisms.[29,30] Since few organic compounds are known to modulate CME, laurifolioside (**1**) represents a new scaffold for the modulation and investigation of clathrin function and CME.

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotations were measured on a Perkin-Elmer 241 polarimeter equipped with a sodium lamp (589 nm) and a 1 dm microcell. UV spectra were recorded on a Perkin-Elmer-Lambda spectrophotometer. ECD spectra were measured at room temperature in MeOH on a Chirascan (Applied Photophysics, Surrey, UK) spectrometer in a 0.1 cm cell using the following conditions: speed 50 nm/min, time constant 1 s, bandwidth 1.0 nm. NMR experiments were recorded at 300 K in CD_3OD on a Bruker DRX-600 spectrometer (Bruker BioSpin GmBH) equipped with a Bruker 5 mm TCI CryoProbe. Bruker standard pulse sequences and phase cycling were used for DQF-COSY, TOCSY, HSQC, HMBC, and NOESY experiments. HRESIMS data were acquired in the positive-ion mode on a Q-TOF premier spectrometer equipped with a nano-electrospray ion source (Waters, Milford, MA, USA). A ≥95% purity of all compounds was inferred from HPLC analysis on an Agilent 1200 series system with UV detection at 220 nm (Phenomenex Kinetex C_{18} 2.1 × 150 mm, 5 μm, 10−90% CH_3CN in H_2O with 0.1% TFA for 20 min; flow rate 0.5 mL/min). Preparative separation was performed on silica gel, Sephadex LH-20, and by HPCPC (high-performance centrifugal partition chromatography). Semi-preparative HPLC was performed with a Shimadzu LC-8A series pumping system equipped with a Shimadzu RID-

10A refractive index detector and Shimadzu injector, using a C_{18} μ-Bondapak column (30 cm x 7.8 mm, 10 μm; Waters) at a flow rate of 2.0 mL/min.

Plant Material. The aerial parts of *Euphorbia laurifolia* were collected in Tumbaco, Ecuador, in September 2011. The plant was identified at the Herbarium of Jardin Botanico de Quito, Quito, Ecuador. A voucher specimen (no. 4498 *Euphorbia laurifolia*/1) was deposited at Herbarium Horti Botanici Pisani, Pisa, Italy.

Extraction and Isolation. The dried powdered leaves of *E. laurifolia* (527 g) were extracted successively for 48 h with *n*-hexane, $CHCl_3$, $CHCl_3$-MeOH (9:1) and MeOH, by exhaustive maceration (3 x 2 L), to give 15.2, 14.6, 2.7 and 20.2 g of dry residues. The *n*-hexane extract (15.2 g) was partitioned between *n*-hexane and MeOH-H_2O (3:2), to afford a polar fraction (0.473 g) that was chromatographed over silica gel column (25 g silica SNAP cartridges, flow rate 25 mL/min) using a Biotage® Isolera™ Spektra flash purification system, eluting with *n*-hexane-$CHCl_3$ (1:1) followed by increasing concentrations of $CHCl_3$ in *n*-hexane (between 50% and 100%), and increasing concentrations of MeOH in $CHCl_3$ (between 3% and 10%). Fractions of 12 mL were collected, analyzed by TLC, and grouped into three major fractions (A-C). Fraction B (295.6 mg) was purified by HPCPC with *n*-hexane-EtOAc-MeOH-H_2O (8:5:5:2). The stationary phase consisted of the upper phase (descending mode, flow rate 3 mL/min). Three fractions (A-C) were grouped by TLC. Fraction B (20.2 mg) was purified by RP-HPLC with MeOH-H_2O (7.5:2.5) as eluent to afford **3** (0.8 mg, t_R 10 min) and latazienone (0.7 mg, t_R 8 min). Part of the $CHCl_3$ extract (5 g) was chromatographed over silica gel column (340 g silica SNAP cartridges, flow rate 100 mL/min) using a Biotage® Isolera™ Spektra flash purification system, eluting with *n*-hexane-$CHCl_3$ (1:1), followed by increasing concentrations of $CHCl_3$ in *n*-hexane (between 50 % and 100%), and MeOH in $CHCl_3$ (between 5% and 20%). Fractions of 27 mL were combined into ten major fractions (A-J). Fraction E (486 mg) was purified by RP-HPLC with MeOH-H_2O (8.5:1.5) as eluent to give **5** (1.5 mg, t_R 7 min). Fraction F (180.5 mg) was separated by RP-HPLC with MeOH-H_2O (7:3) as eluent to afford *ent*-16α,17-dihydroxykauran-3-one (2.2 mg, t_R 12 min). Fraction G (485.6 mg) was subjected to RP-HPLC with MeOH-H_2O (6.5:3.5) as eluent to give vomifoliol (3.1 mg, t_R

6 min) and *ent*-16α,17-dihydroxyatisan-3-one (1.3 mg, t_R 17 min). Fraction H (220.9 mg) was purified over RP-HPLC with MeOH-H_2O (2:3) as eluent to give compound **6** (2.5 mg, t_R 24 min). Fraction I (190.9 mg) was subjected to RP-HPLC with MeOH-H_2O (5.5:4.5) as eluent to yield compound **1** (7.8 mg, t_R 13 min). The MeOH extract (20.1 g) of *E. laurifolia* was partitioned between *n*-BuOH and H_2O, to afford a *n*-BuOH residue (2.9 g). The *n*-BuOH fraction was submitted to passage over Sephadex LH-20 using MeOH as eluent to obtain eight major fractions (A-H). Fraction B (407.4 mg) was purified over RP-HPLC with MeOH-H_2O (4.5:5.5) as eluent to give compounds **1** (1.6 mg, t_R 13 min), **4** (1.1 mg, t_R 23 min), and **2** (1.2 mg, t_R 60 min).

*Compound **1***: pale yellow oil; $[\alpha]_D^{25}$ +19 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 245 (3.93), 230 sh (3.07) nm; ECD (c = 7.30 x 10^{-4} M, 0.1 cm, MeOH) $[\theta]_{202}$ –3206, $[\theta]_{227}$ +12189, $[\theta]_{254}$ +11382, $[\theta]_{350}$ –2674; ^{1}H and ^{13}C NMR data, see Table 1; ESIMS *m/z* 477 $[M - H]^-$, 315 $[M - H - 162]^-$; HRESIMS *m/z* 501.2439 $[M + Na]^+$ (calcd for $C_{26}H_{38}O_8Na$ 501.2463), 317.2104 $[M + H - 162]^+$.

*Compound **2***: pale yellow oil; (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 245 (3.92), 230 (3.06) sh nm; ECD (c = 8.36 x 10^{-4} M, 0.1 cm, MeOH) $[\theta]_{202}$ –4525, $[\theta]_{227}$ +9885, $[\theta]_{254}$ +9590, $[\theta]_{350}$ –2569; ^{1}H and ^{13}C NMR data, see Table 1; HRESIMS *m/z* 501.2437 $[M + Na]^+$ (calcd for $C_{26}H_{38}O_8Na$ 501.2463), 317.2101 $[M + H - 162]^+$.

*Compound **3**:* pale yellow oil; ^{1}H and ^{13}C NMR data, see Table 1; HRESIMS *m/z* 569.2744 $[M + Na]^+$ (calcd for $C_{30}H_{42}O_9Na$ 569.2727), 509.2666 $[M + Na - 60]^+$, 449.3008 $[M + Na - 60 - 60]^+$, 389.3351 $[M + Na - 60 - 60 - 60]^+$.

*Compound **4**:* pale yellow oil; ^{1}H and ^{13}C NMR data, see Table 2; HRESIMS *m/z* 509.2432 $[M - H]^-$ (calcd for $C_{26}H_{37}O_{10}$ 509.2387), 347.3342 $[M - H - 162]^-$, 329.3456 $[M - H - 162 - 18]^-$, 533.2091 $[M + Na]^+$, 353.3166 $[M + Na - 180]^+$.

*Compound **5***: amorphous powder; ^{1}H and ^{13}C NMR data, see Table 2; HRESIMS *m/z* 317.2920 $[M + H]^+$ (calcd for $C_{20}H_{29}O_3$ 317.2908).

*Compound **6***: amorphous powder; $[\alpha]_D^{25}$ +8 (*c* 0.13, MeOH); ^{1}H and ^{13}C NMR data, see Table 2; HRESIMS *m/z* 507.3278 $[2M + Na]^+$ (calcd for $C_{13}H_{22}O_4$ 242.1518).

Acid Hydrolysis of Compounds 1 and 4. Acid hydrolysis of compounds **1** and **4** was carried out as reported in a previous paper.[31] D-Glucose was identified as sugar moiety by comparison with the retention time of an authentic sample (Sigma Aldrich).

Computational Methods. Conformational analysis of compound **1** and **2** were performed with MacroModel 9.1 (Schrödinger, LLC, New York) employing the OPLS-2005 (*optimized potential for liquid simulations*) force field in H_2O. Conformers within a 2 kcal/mol energy window from the global minimum were selected for geometrical optimization and energy calculation applying DFT with the B3LYP/6-31G** level of theory with the Gaussian 09 program package.[32] Vibrational evaluation was done at the same level to confirm minima. Excitation energy (denoted by wavelength in nm), rotator strength dipole velocity (R_{vel}), and dipole length (R_{len}) were calculated in MeOH by TDDFT/B3LYP/6-31G**) using the SCRF (*self-consistent reaction field*) method with the CPCM (c*onductor-like polarizable continuum*) model. ECD curves were constructed on the basis of rotatory strengths with a half-band of 0.35 eV using SpecDis v1.61.[33]

Cell Culture, Proliferation, and Viability. PC-3 (human prostate cancer cell line) and MCF-7 cells (human breast adenocarcinoma cell line) were cultured and treated as previously reported.[34] In brief, cells were seeded in 96 well-plates at a cell density of 1 x 10^4/ well (100 µL of a 1 x 10^5 cells/mL) and allowed to growth in the absence and in the presence of different concentrations of compound **1**. At 24, 48 h and 72 h, the number of cells was quantified by using an MTT conversion assay.

DARTS Assay. Human prostate cancer (PC-3) and human breast adenocarcinoma (MCF-7) cells were treated with 20 µM or 40 µM compound **1** or DMSO control for 1 h. Cells were lysed in PBS containing 0.1% Igepal (lysis buffer) and phosphatase inhibitor cocktail (P8340, Sigma-Aldrich, St. Louis, MO, USA). Protein concentrations were determined by a Bio-Rad DC Protein Assay (Bio-Rad, Hercules, CA, USA) using bovine albumin as standard. All steps were performed on ice or at 4 °C. Samples were warmed to room temperature and digested enzymatically with subtilisin (enzyme: lysate 1:750 w/w for 30 min at 30 °C). The resulting mixtures were separated by SDS-PAGE, and stained with Coomassie blue. Gel lanes showing significant differences in intensity in the different

samples were excised manually and subjected to an in-gel digestion procedure.[35] Peptides were analyzed by high-resolution LC-MS/MS, using a Orbitrap XL mass spectrometer (Thermo Fisher Scientific Inc., Rockford, IL USA) equipped with a nanospray ion source and coupled to a nano-Acquity capillary UPLC system (Waters, Milford, MA, USA). Mass spectra were acquired in a *m*/*z* range from 400 to 1800, and MS/MS spectra in a *m/z* range from 25-2000. Mass and MS/MS spectra calibrations were performed using a mixture of angiotensin and insulin as external standard, and human [Glu]-fibrinopeptide B as lock mass standard. MS and MS/MS data were used by Mascot (Matrix Science) to interrogate the Swiss Prot non-redundant protein database. Settings were as follows: mass accuracy window for parent ion, 10 ppm; mass accuracy window for fragment ions, 200 millimass units; fixed modification, carbamidomethylation of cysteines; variable modifications, oxidation of methionine. Proteins with more than two peptides and program scores > 100 were considered as reliable proteins. The same experiments were also performed with PC-3 and MCF-7 protein extracts after incubations with 20 μM or 40 μM compound **1**, or DMSO as control.

Transferrin Internalization and Fluorescence Microscopy. PC-3 human pancreatic adenocarcinoma cells were cultured at 37 °C in a 5% CO_2 humidified atmosphere in RPMI 1640 containing 10% heat-inactivated fetal bovine serum, and supplemented with antibiotics (10000 U/mL penicillin and 10 mg/mL streptomycin). Cells were seeded in a six-well plastic plate at 3 x 10^4 cells per well, on a 12-mm glass coverslip. After incubation with starvation medium (RPMI 1640/0.1% FBS), cells were pretreated with 30 μM pitstop2 (Abcam Biochemicals), DMSO 0.1%, or different concentrations of compound **1** (25, 10, 5 μM) for 15 min. Cells were then washed with PBS/HEPES/20 mM glucose/1% BSA, 20 μg/mL transferrin conjugate (Texas Red® conjugate, Molecular Probes) for 15 min at 37 °C, in the presence or absence of inhibitors. Then coverslips were fixed in *p*-formaldehyde (4% v/v in PBS) for 30 min, washed with 1X PBS, and mounted on microscope slides. Single planes were performed with a Zeiss LSM510 laser scanning confocal microscope (Carl Zeiss Microimaging GmbH, Germany). Images were acquired in sequential scan mode by using the same acquisition parameters when comparing experimental and control material.

Docking Studies. The ligand was built using Maestro,[36] and minimized into an aqueous environment with Macromodel[37] (employing the generalized Born/surface area model). Minimization was carried out by means of the conjugate gradient, the Merck molecular force fields, and a distance-dependent dielectric constant of 1.0, until a convergence value of 0.05 kcal/(Å•mol) was reached. The ligand was docked into the crystal structure of the terminal domain (TD) of human Clathrin complexed with an inhibitor (PDB code 2XZG)[24] using AUTODOCK4.2.[26] AUTODOCK TOOLS[38] were employed to define the torsion angles in the ligands, to add the solvent model, and to assign partial atomic charges (Gasteiger for the ligands, and Kollman for the receptors). The docking site used for calculations was defined considering the bound ligand as the central group of a grid of 60, 70, and 60 points in the x, y, and z directions, respectively. The energetic map calculations were carried out by using a grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant. The ligand was subjected to 200 runs of the AUTODOCK search using the Lamarckian genetic algorithm with 10,000,000 steps of energy evaluation. The number of individuals in the initial population was set to 500, and a maximum of 10,000,000 generations were simulated during each docking run. An rms tolerance of 2.0 Å was used to carry out the cluster analysis of the docking solutions, and all the other settings were left as their defaults. The four clusters of solutions with a population higher than 5%, i.e., including more than 5% of all the generated docking poses, were taken into account.

Molecular Dynamic Simulations. All simulations were carried out using AMBER 14.[39] The simulation protocol was set up using the X-ray complex Clathrin TD-inhibitor employed for docking studies (PDB code 2XZG), which was used as a reference. The different ligand-protein complexes were placed in a cubic water-box and solvated with a 15 Å water cap; sodium ions were then added as counter ions to neutralize the system. General amber force field (GAFF) parameters were assigned to the ligands, while partial charges were calculated using the AM1-BCC method. Prior to the molecular dynamic (MD) simulations, the complexes were energy minimized through 5,000 steps of steepest descent, followed by conjugate gradient until a convergence of 0.05 kcal/(mol·Å^2) was reached. At this stage, a position restraint of 10 kcal/(mol·Å^2) was applied to the

protein α carbons. The minimized systems were used as starting point for the MD simulations which were performed in three different steps. Each step was run using particle mesh Ewald electrostatics and periodic boundary conditions,[40] a cutoff of 10 Å for the nonbonded interactions, and employing SHAKE algorithm to keep rigid every bond involving hydrogen. The first MD step consisted in 0.5 ns of constant-volume simulation in which the temperature of the system was raised from 0 to 300 K. In the second step a 3 ns constant-pressure simulation was carried out to equilibrate the system, and the temperature of the system was kept constant at 300 K by using the Langevin thermostat. In both the first and second step, a harmonic potential of 10 kcal/(mol·Å^2) was applied to the protein α carbons, as in the minimization stage. Step 3 consisted of a 26.5 ns simulation that was performed using the same conditions as used in step 2, but without applying any position restraint in order to leave the system completely free. A total of 30 ns MD simulation was thus performed for each analyzed ligand-protein complex.

Evaluation of Binding Energy. Evaluation of the binding energy of the different ligand-protein complexes analyzed through MD simulations was carried out using AMBER 14. The trajectories relative to the last 15 ns of each simulation were extracted, and used for the calculation for a total of 150 snapshots (at time intervals of 100 ps). Van der Waals, electrostatic and internal interactions were calculated with the SANDER module of AMBER 14, whereas polar energies were calculated using both the Generalized Born and the Poisson−Boltzman methods with the MM-PBSA module of AMBER 14. Dielectric constants of 1 and 80 were used to represent the gas and water phases, respectively, while the MOLSURF program was employed to estimate the nonpolar energies. The entropic term was considered as approximately constant in the comparison of the ligand−protein energetic interactions.

ASSOCIATED CONTENT

Supporting Information. NMR spectra of compounds **1**-**6**; calculated ECD spectra for compounds **1** and **2**; MD simulations analysis; MM-GBSA and MM-PBSA method results. This material is available via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Tel: +39-089-969754. Fax: +39-089-969602. E-mail: detommasi@unisa.it.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Engi, H.; Vasas, A.; Redei, D.; Molnar, J.; Hohmann, J. *Anticancer Res*. **2007**, *27*, 3451-3458.

(2) Ferreira, M-J. U.; Duarte, N.; Reis, M.; Madureira, A. M.; Molnar, J. *Phytochemistry Rev.* **2014**, *13*, 915-935.

(3) Vasas, A.; Hohmann, J. *Chem Rev*. **2014**, *114*, 8579-8612.

(4) Reis, M.; Ferreira, R. J.; Santos, M. M. M.; dos Santos, D. J. V. A.; Molnar, J.; Ferreira, M-J. U. *J. Med. Chem*. **2013**, *56*, 748-760.

(5) Baell, J. *J. Nat. Prod*. **2016**, *79*, 616-628.

(6) Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A.; Wohlschlegel, J. A.; Vondriska, T. M.; Pellletier, J.; Herschmann, H. R.; Clardy, J.; Clarke, C. F.; Huang, J. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 1894-1989.

(7) Lomenick, B.; Olsen, R. W.; Huang, J. *ACS Chem. Biol*. **2011**, *6*, 34-46.

(8) Rondon, M.; Morales, A.; Amaro-Luis, J. M.; Bahsas, A.; Rojas, J.; Buitrago, D. *Nat. Prod. Res*. **2005**, *19*, 597-602.

(9) Avila, L.; Perez, M.; Sanchez-Duffhues, G.; Hernandez-Galan, R.; Muñoz, E.; Cabezas, F.; Quiñones, W.; Torres, F.; Echeverri, F. *Phytochemistry* **2010**, *71*, 243-248.

(10) Liao, S.-G.; Zhan, Z.-J.; Yang, S.-P.; Yue, J.-M. *Org. Lett.* **2005**, *7*, 1379-1382.

(11) Marco, J. A.; Sanz-Cervera, J. F.; Yuste, A. *Phytochemistry* **1997**, *45*, 563-570.

(12) Bakhsh Baloc, I.; Kaleem Baloch, M.; Khan Baloch, A. *Eur. J. Med. Chem*. **2009**, *44*, 3188-3194.

(13) Yu, C.-H.; Hsieh, C.-R.; Hsiao, G.; Chen, P.-Y.; Chang, M.-L.; Yin, H.-W.; Lee, T.-H.; Lee, C.-K. *Molecules* **2012**, *17*, 2082-2090.

(14) Lu, Z.-Q.; Guan, S.-H.; Li, X.-N.; Chen, G.-T.; Zhang, J.-Q.; Huang, H.-L.; Liu, X.; Guo, D.-A. *J. Nat. Prod.* **2008**, *71*, 873-876.

(15) De Marino, S.; Borbone, N.; Zollo, F.; Ianaro, A.; Di Meglio, P.; Iorizzi, M. *J. Agric. Food Chem*. **2004**, *52*, 7525-7531.

(16) Ding, Y.-L.; Jia, Z.-J. *Phytochemistry* **1991**, *30*, 2413-2415.

(17) Calis, I.; Kuruuzum-Uz, A.; Lorenzetto, P. A.; Ruedi, P. *Phytochemistry* **2002**, *59*, 451-457.

(18) Lal, A. L.; Cambie, R. C.; Rutledge, P. S.; Woodgate, P. D. *Phytochemistry* **1990**, *29*, 1925-1935.

(19) Sousa, I. J.; Ferreira, M-J. U.; Molnar, J.; Fernandes, M. X. *Eur. J. Pharm. Sci*. **2013**, *48*, 542-553.

(20) Barile, E.; Borriello, M.; Di Pietro, A.; Doreau, A.; Fattorusso, C.; Fattorusso E.; Lanzotti, V. *Org. Biomol. Chem*. **2008**, *6*, 1756-1762.

(21) Banerjee, D.; Liu, A. P.; Voss, N. R.; Schmid, S. L.; Finn, M. G. *Chembiochem* **2010**, *11*, 1273-1279.

(22) Haar, E.; Harrison, S. C.; Kirchhausen, T. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1096-1100.

(23) Lafer, M. *Traffic* **2002**, *3*, 513-520.

(24) Kleist, L.; Stahlschmidt, W.; Bulut, H.; Gromova, K.; Puchov, D.; Robertson, M. J.; MacGregor, K. A.; Tomilin, N.; Pechtstein, A.; Chau, N.; Chircop, M.; Sakoff, J.; Kries, J. P.; Saenger, W.; Kräusslich, H.; Shupliakov, O.; Robinson, P. J.; McCluskey, A.; Haucke, V. *Cell* **2011**, *146*, 471-484.

(25) Smith, C. M.; Haucke, V.; McCluskey, A.; Robinson, P. J.; Chircop, M. *Mol. Cancer* **2013**, *12,* 4.

(26) 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.

(27) Poli, G.; Gelain, A.; Porta, F.; Asai, A.; Martinelli, A.; Tuccinardi, T. *J Enzyme Inhib. Med. Chem*. **2015**, 1-7.

(28) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889-897.

(29) Ybe, J. A. *BioMol. Concepts* **2014**, *5*, 175-182.

(30) Kirchhausen T.; Owen, D.; Harrison, S. C. *Cold Spring Harb. Perspect. Biol*. **2014**, *6*, a016725.

(31) Dal Piaz, F.; Vassallo, A.; Temraz, A.; Cotugno, R.; Belisario, M. A.; Bifulco, G.; Chini, M. G.; Pisano, C.; De Tommasi, N.; Braca, A. *J. Med. Chem.* **2013**, *56*, 1583-1595.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision A02; Gaussian, Inc: Wallingford, CT, 2009.

(33) Bruhn, T.; Schaumlöffel, A.; Hemberger, Y.; Bringmann, G. SpecDis version 1.61; University of Wuerzburg: Germany, 2013.

(34) Dal Piaz, F.; Malafronte, N.; Romano, A.; Gallotta, D.; Belisario, M. A.; Bifulco, G.; Gualtieri, M. J.; Sanogo, R.; De Tommasi, N.; Pisano, C. *Phytochemistry* **2012**, *75*, 78-89.

(35) Dal Piaz, F.; Cotugno, R.; Lepore, L.; Vassallo, A.; Malafronte, N.; Lauro, G.; Bifulco, G.; Belisario, M. A.; De Tommasi, N. *J. Proteom*. **2013**, *82*, 14-26.

(36) *Maestro*, version 9.0. Schrödinger, Inc: Portland, OR, 2009.

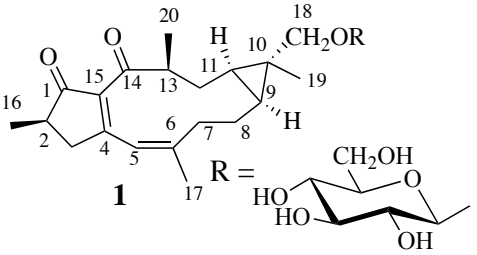
(37) *Macromodel*, version 9.7. Schrödinger, Inc: Portland, OR, 2009.

(38) Sanner, M. F. *J. Mol. Graphics Modell*. **1999**, *17*, 57-61.

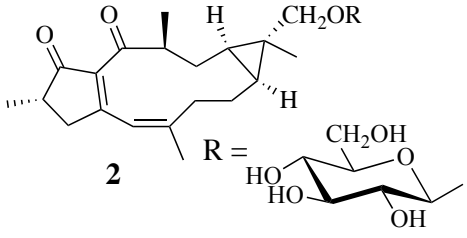
(39) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E. III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus,

J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Salomon-Ferrer, R.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D.M.; Kollman, P. A. *AMBER*, version 14. In University of California: San Francisco, CA, 2015.

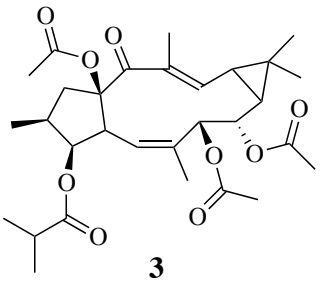
(40) York, D. M.; Darden, T. A.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345-8348.


20
18
O
O
H
10
CH2OR
1
15
14
13
11
19
16
9
H
6
7
8
2
4
5
17
R =
CH2OH
O
HO
HO
OH

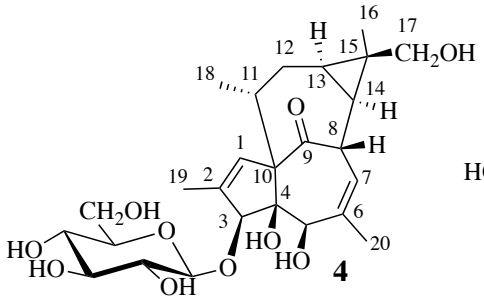
1


O
H
CH2OR
H
R =
CH2OH
O
HO
HO
OH

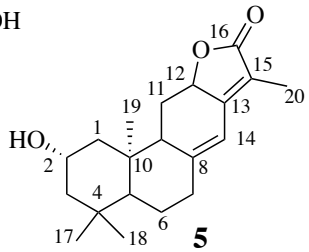
2


O
O
O
O
O
O
O
O

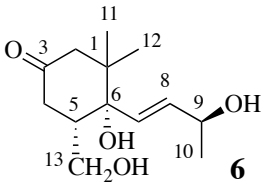
3


16
H
17
12
15
CH2OH
18
11
13
14
H
O
8
H
1
9
19
2
10
7
4
6
CH2OH
O
3
HO
HO
HO
HO
O
OH
20

4


16
O
O
12
15
11
13
20
19
1
14
HO
2
10
8
4
6
17
18

5


11
O
3
1
12
8
6
9
OH
5
OH
13
CH2OH
10

6

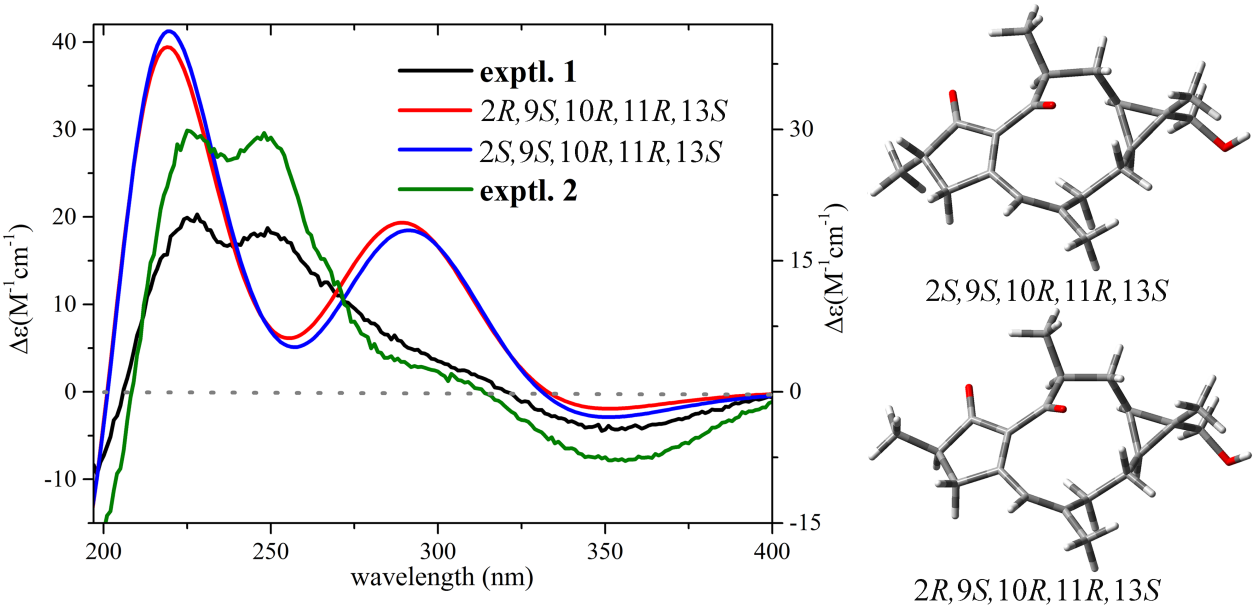


Figure 1. Comparison of calculated and experimental ECD spectra (left), and DFT optimized structure of two possible stereoisomers (right). Calculated spectra were obtained by using TDDFT at the B3LYP/6-31G** level of theory in MeOH.

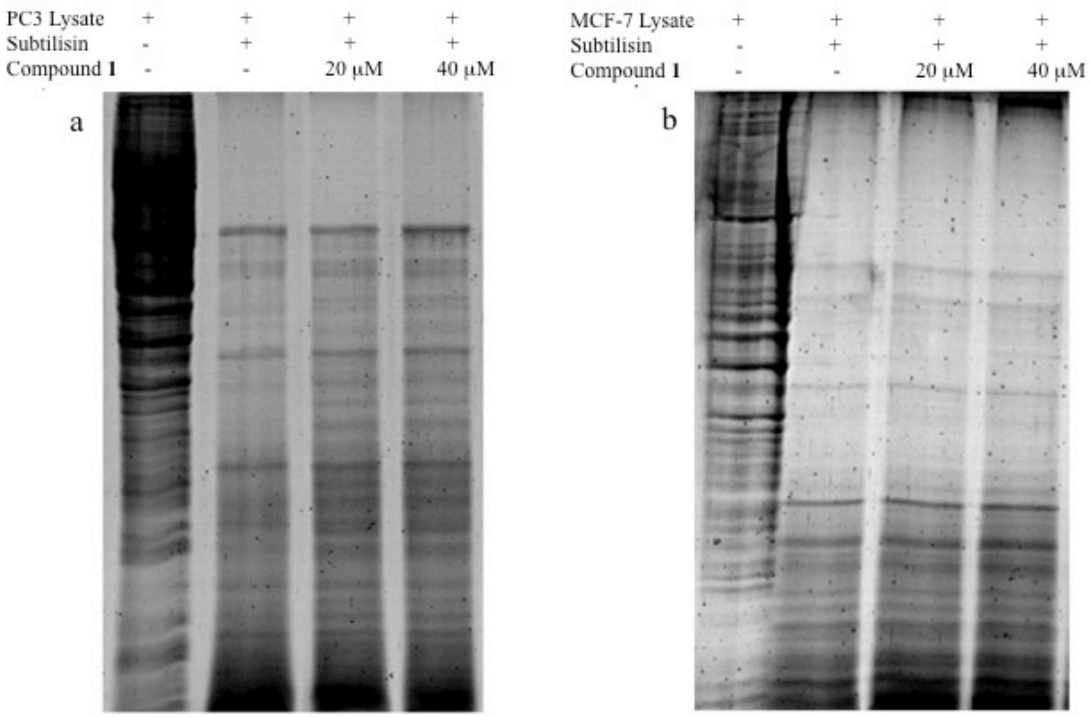


Figure 2. SDS-PAGE analysis of the cellular lysates subjected to subtilisin digestion and Coomassie (SimplyBlue) staining. (a) PC-3 cell lysates with subtilisin with or without 20/40 μM of compound **1**. (b) MCF-7 cell lysates with subtilisin with or without 20/40 μM of **1**.

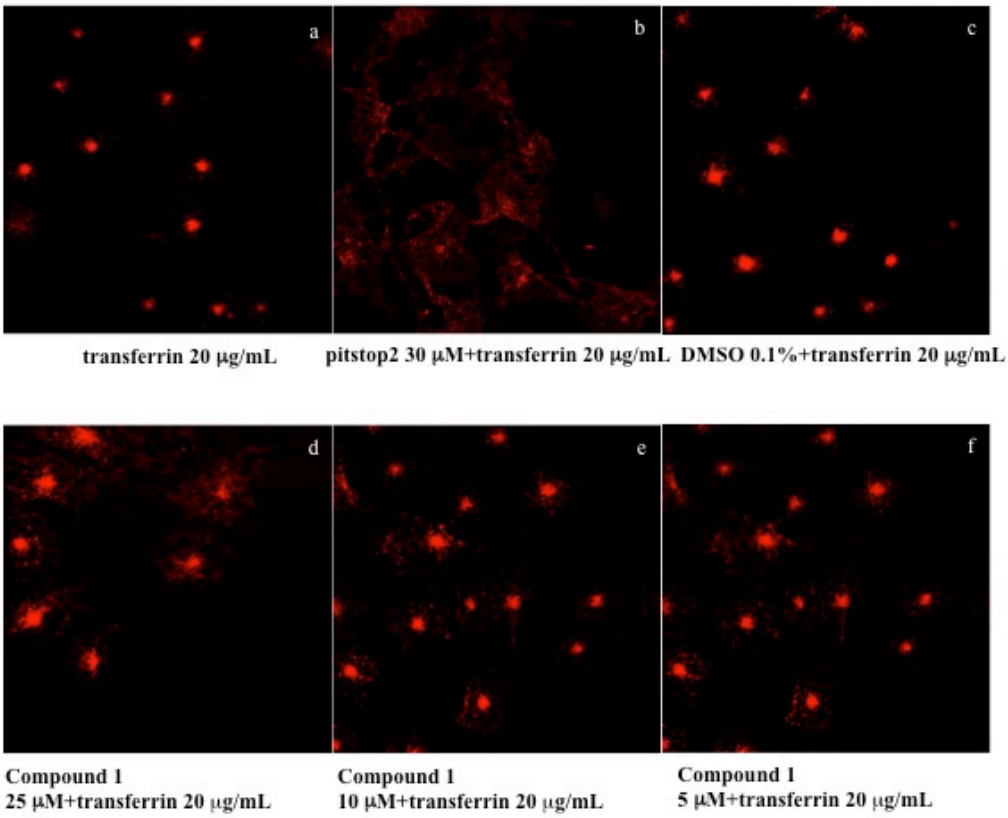


Figure 3. Compound **1** interferes with transferrin internalization. (a) 15 min incubation with transferrin; (b) 15 min preincubation with pitstop2 and 15 min incubation with transferrin; (c) 15 min preincubation with DMSO 0.1% and 15 min incubation with transferrin; (d)-(f) 15 min preincubation with compound **1** (25, 10, 5 μM) and 15 min incubation with transferrin.

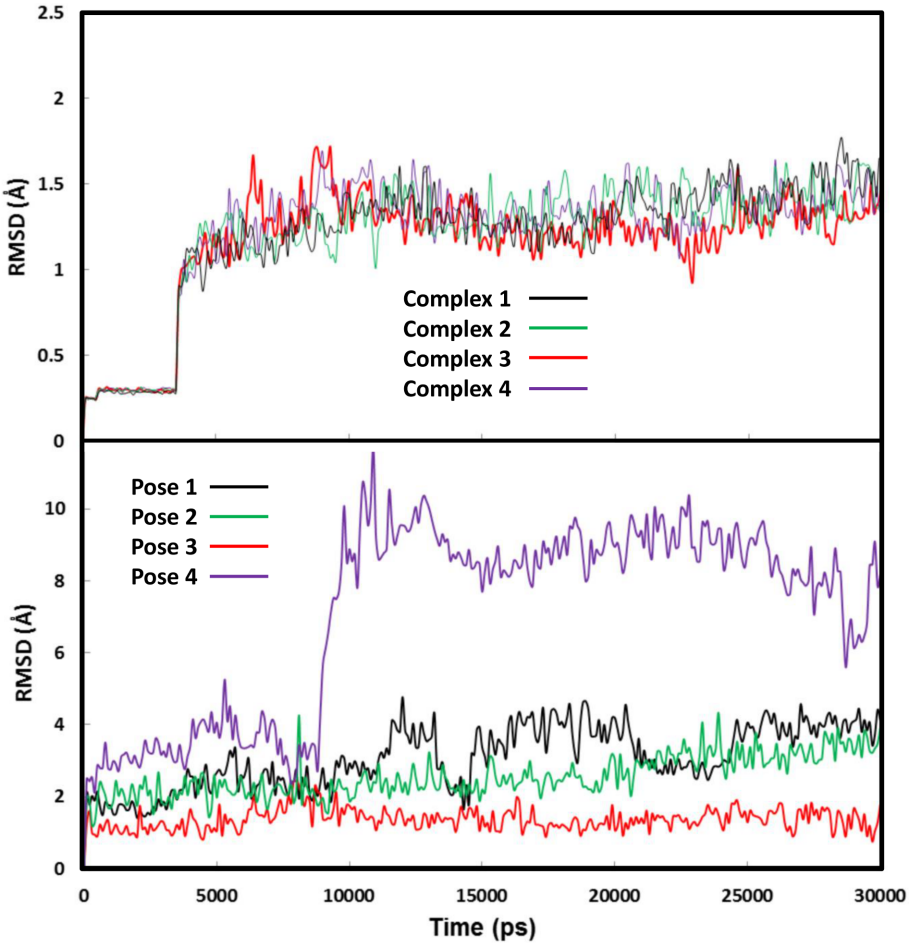


Figure 4. Analysis of the MD simulations of the four different clathrin TD-compound **1** complexes. The first plot shows the RMSD of the protein α carbons from their crystallographic coordinates during the simulation; in the second plot the RMSD of the position of the ligand with respect to its initial docking pose is shown.

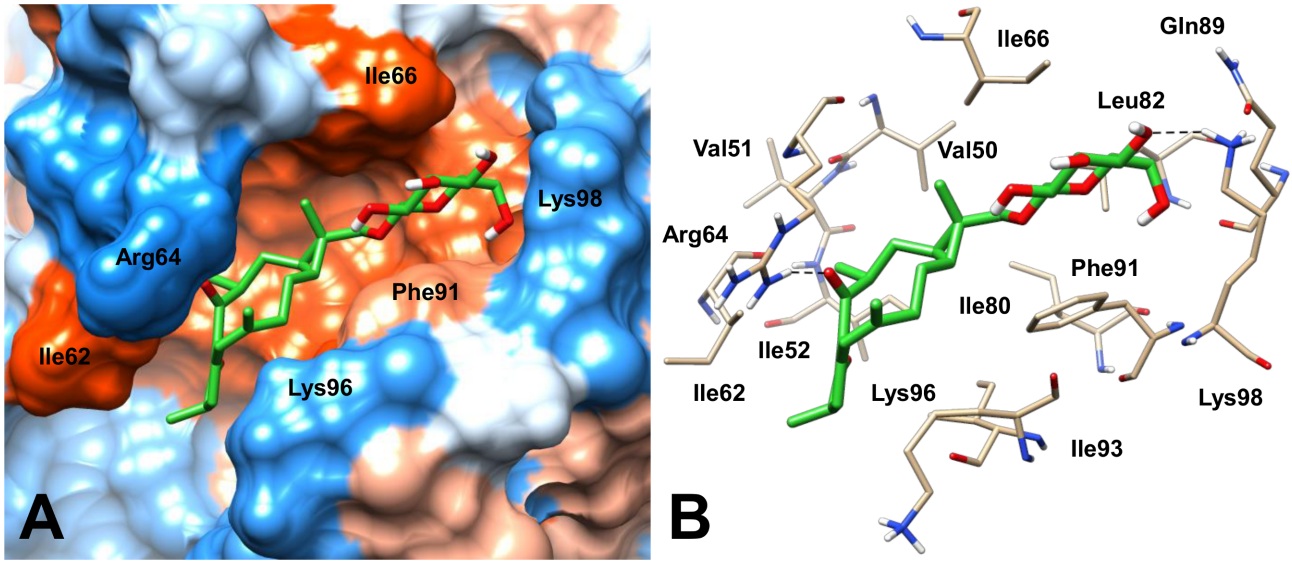


Figure 5. Minimized average structure of compound **1** bound to clathrin TD in pose 3, derived from the last 15 ns of MD simulation. (A) The protein hydrophobicity surface is shown; the most lipophilic regions are colored in orange while the most polar ones are colored in blue. (B) The protein residues directly interacting with compound **1** are shown. Hydrogen bonds are represented as black dashed lines.

Table 1. ^{1}H and ^{13}C NMR Data of Compounds 1 and 2 (Methanol-d_4, 600 MHz)[a]

	1		2		3	
position	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1a		213.4		213.2	2.81 dd (15.0, 6.2)	42.2
1b					2.19 dd (15.0, 7.8)	
2	2.62 m	40.9	2.58 m	41.0	2.32 m	39.0
3a	2.97 dd (14.0, 3.5)	37.2	3.19[b]	37.2	4.85[b]	83.0
3b	2.53 dd (14.0, 2.0)		2.30 dd (18.0, 16.0)			
4		169.3		169.4	2.84 dd (6.2, 3.4)	47.8
5	5.71 s	116.6	5.71 s	116.5	5.85 d (10.7)	124.6
6		145.9		145.4		140.8
7a	2.37 m	39.4	2.36 m	39.4	4.88[b]	77.8
7b	2.25 br t (12.0)		2.22 br t (12.0)			
8a	1.89[b]	23.6	1.88[b]	23.5	5.29 br t (10.5)	73.8
8b	1.18 m		1.18 m		1.18 m	
9	0.56 br t (9.0)	26.2	0.55 br t (9.2)	26.1	1.43 br t (8.5)	34.6
10		21.7		21.7		26.4
11	0.77 ddd (13.0, 8.0, 3.0)	22.7	0.76 ddd (13.5, 9.0, 5.0)	21.4	1.80 dd (12.0, 8.5)	30.7
12a	1.89[b]	27.5	1.88[b]	27.4	6.66 d (12.0)	145.6
12b	1.40 ddd (15.0, 8.0, 3.0)		1.40 ddd (15.0, 7.5, 3.0)			
13	3.32[b]	44.6	3.33[b]	44.2		134.3
14		211.9		212.0		195.8
15		142.0		142.0		96.0
16	1.22 d (6.5)	16.6	1.29 d (6.5)	16.5	1.05 d (6.5)	18.4
17	1.68 s	20.7	1.66 s	20.6	1.61 s	18.2
18a	3.59 d (11.0)	80.4	3.58 d (11.0)	80.4	1.24 s	28.5
18b	3.40 d (11.0)		3.39 d (11.0)			
19	1.02 s	11.5	1.02 s	11.0	1.11 s	17.0
20	1.07 d (6.5)	17.3	1.08 d (6.5)	17.9	1.83 s	12.3
glc-1'	4.31 d (7.5)	103.6	4.31 d (7.5)	103.2		
2'	3.22 dd (9.0, 7.5)	75.0	3.21 dd (9.0, 7.5)	74.7		
3'	3.29 t (9.0)	77.6	3.27 t (9.0)	77.7		
4'	3.34 t (9.0)	71.3	3.34 t (9.0)	71.4		
5'	3.37 m	78.0	3.34 m	78.0		
6'a	3.88 dd (12.0, 3.5)	62.5	3.88 dd (12.0, 3.0)	62.3		
6'b	3.71 dd (12.0, 5.0)		3.72 dd (12.0, 4.5)			
$\underline{CH_3}CO$-7					2.00 s	20.7
$CH_3\underline{CO}$						171.3
$\underline{CH_3}CO$-8					2.12 s	21.7
$CH_3\underline{CO}$						171.6
$\underline{CH_3}CO$-15					2.12 s	21.7
$CH_3\underline{CO}$						171.6
isobut-1						177.3
2					2.35 sept (7.0)	35.6
3/4					1.26 d (6.5)	19.5

[a]*J* values are in parentheses and reported in Hz; chemical shifts are given in ppm; assignments were confirmed by DQF-COSY, 1D-TOCSY, HSQC, and HMBC experiments. [b]Overlapped signal.

Table 2. ^{1}H and ^{13}C NMR Data of Compounds 4-6 (Methanol-d_4, 600 MHz)[a]

	4		5		6	
position	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
1a	5.88 s	131.0	2.23 br d (14.0)	49.6		43.6
1b			1.08 dd (14.0, 4.5)			
2a		140.0	3.84 br dd (4.5, 3.5)	66.8	2.96 d (13.5)	53.0
2b					1.86 d (13.5)	
3a	4.50 s	90.0	1.79 dd (12.5, 5.0)	51.9		215.2
3b			1.21 dd (12.5, 3.5)			
4a		87.6		35.0	2.80 t (13.5)	41.5
4b					2.29 dq (13.5, 6.7, 4.0, 2.0)	
5	3.34[b]	76.4	1.26 dd (12.0, 3.0)	56.0	2.21 m	43.6
6a		140.6	1.86 m	27.0		78.1
6b			1.57 m			
7a	5.70 br d (4.0)	123.6	2.58 br d (13.0)	37.0	5.80 d (15.5)	132.0
7b			2.29 m			
8a	4.38 br d (11.0)	44.4		152.2	5.98 dd (15.5, 6.2)	137.6
8b						
9		211.0	2.39 d (8.5)	53.0	4.37 q (12.6, 11.5, 6.0, 4.7)	69.0
10		73.6		42.0	1.29 d (6.0)	24.4
11a	2.44 m	40.6	2.63 dd (13.0, 7.0)	28.8	0.92 s	24.0
11b			1.52 m			
12a	2.40 m	31.3	4.99[b]	77.2	0.92 s	24.0
12b	1.86[b]					
13a	0.84 br t (8.5, 4.2)	24.6		157.7	3.85 dd (11.0, 5.0)	64.0
13b					3.61 dd (11.0, 2.5)	
14	0.95[b]	24.0	6.49 s	113.4		
15		31.0		116.8		
16	1.14 s	24.6		176.0		
17a	3.71 d (11.0)	62.8	0.94 s	23.3		
17b	3.70 d (11.0)					
18	0.95 d (6.5)	17.0	1.00 s	34.0		
19	1.87 s	15.6	1.03 s	34.7		
20	1.77 s	22.3	1.70 s	8.9		
glc-1'	4.52 d (7.5)	104.8				
2'	3.20 dd (9.5, 7.5)	75.0				
3'	3.38 t (9.5)	78.0				
4'	3.36 t (9.5)	71.0				
5'	3.40 m	78.0				
6'a	3.88 dd (12.0, 3.0)	62.8				

[a]*J* values are in parentheses and reported in Hz; chemical shifts are given in ppm; assignments were confirmed by DQF-COSY, 1D-TOCSY, HSQC, and HMBC experiments. [b]Overlapped signal.

Table 3A. Proteins Identified by DARTS on PC-3 Protein Extracts as Compound 1 Molecular Targets

swiss-prot code	protein	mr	sequence coverage (%)	matches	sequences	mascot score
ADT2_HUMAN	ADP/ATP translocase 2	32852	26	8	7	562
CALR_HUMAN	calreticulin	48142	18	8	8	531
CLH1_HUMAN	clathrin heavy chain 1	187030	8	9	8	483
DSC1_HUMAN	desmocollin-1	99987	5	5	4	316
FLNA_HUMAN	filamin-A	280739	3	4	4	148
TRXR1_HUMAN	thioredoxin reductase 1	70906	7	4	4	146

Table 3B. Proteins Identified by DARTS on Intact PC-3 Cells as Compound 1 Molecular Targets

swiss-prot code	protein	mr	sequence coverage (%)	matches	sequences	mascot score
ANXA1_HUMAN	annexin 1	38714	25	7	7	605
CLH1_HUMAN	clathrin heavy chain 1	187030	10	10	9	594
HSP7C_HUMAN	heat shock cognate 71KDa	70898	12	6	4	419
MOES_HUMAN	moesin	67820	15	4	4	320
PSA2_HUMAN	proteasome subunit α2	25899	22	5	3	288

Table 3C. Proteins Identified by DARTS on MCF7 Protein Extracts as Compound 1 Molecular Targets

swiss-Prot code	protein	mr	sequence coverage (%)	matches	sequences	mascot score
ANXA3_HUMAN	annexin A3	36375	11	4	4	259
CLH2_HUMAN	clathrin heavy chain 1	187030	8	7	7	430
CYFP1_HUMAN	cytoplasmic FMR1-interacting protein 1	145182	7	5	5	296
HSP90A_HUMAN	heat shock protein HSP 90α	84660	14	8	7	591
MBB1A_HUMAN	myb-binding protein 1A	148855	5	3	2	96
PSME2_HUMAN	proteasome activator complex subunit 2	27402	19	5	5	384
SYEP_HUMAN	bifunctional glutamate/proline-tRNA ligase	170591	9	6	4	157

Table 3D. Proteins Identified by DARTS on Intact MCF-7 Cells as Compound 1 Molecular Targets

swiss-prot code	protein	mr	sequence coverage (%)	matches	sequences	mascot score
CLH1_HUMAN	clathrin heavy chain 1	187030	12	10	10	616
1433S_HUMAN	14-3-3 protein sigma	27774	28	5	5	448
CYFP1_HUMAN	cytoplasmic FMR1-interacting protein 1	145182	7	5	4	215
XPOT_HUMAN	exportin-T	109964	11	6	4	197
MBB1A_HUMAN	myb-binding protein 1A	148855	6	4	4	109