Anti-Colon Cancer Activity of Largazole, a Marine-Derived Tunable Histone Deacetylase Inhibitor

Yanxia Liu, Lilibeth A. Salvador, Seongrim Byeon, Yongcheng Ying, Jason C. Kwan, Brian K. Law, Jiyong Hong, and Hendrik Luesch

Department of Medicinal Chemistry, University of Florida, 1600 SW Archer Road, Gainesville, Florida 32610 (Y.L., L.A.S., J.C.K., H.L.), Department of Chemistry, Duke University, 124 Science Drive, Durham, North Carolina 27708 (S.B., Y.Y., J.H.), Department of Pharmacology & Therapeutics, University of Florida, 1600 SW Archer Road, Gainesville, Florida 32610 (B.K.L.)
Running title: Anti-Colon Cancer Activity of the HDAC Inhibitor Largazole

Corresponding author:

Dr. Hendrik Luesch,

Department of Medicinal Chemistry, University of Florida, 1600 SW Archer Road, P.O. Box 100485, Gainesville, Florida 32610

Tel: (352) 273-7738; Fax: (352) 273-7741; E-mail: luesch@cop.ufl.edu

Number of pages: 37

Number of figures: 7

Number of tables: 0

Number of words in the Abstract: 247

Number of words in the Introduction: 531

Number of words in the Discussion: 1,113

Abbreviations: APC, adenomatous polyposis coli; CDK6, cyclin-dependent kinase 6; CCND1, cyclin D1; EGFR, epidermal growth factor receptor; HDAC, histone deacetylase; HER-2, ERBB2 (v-erb-b2 erythroblastic leukemia viral oncogene homolog 2); IRS-1, insulin receptor substrate 1; MET, met proto-oncogene (hepatocyte growth factor receptor); MLH1, mutL homolog 1; MRM, multiple reaction monitoring; MSH6, mutS homolog 6; NCI, National Cancer Institute; PI3K, phosphoinositide 3-kinase; PP1, protein phosphatase 1; PSF, antibiotic-antimycotic formulation (penicillin, streptomycin, and fungizone = amphotericin B); RTK, receptor tyrosine kinase; SAHA, suberoylanilide hydroxamic acid; SAR, structure–activity relationship

Section: Chemotherapy, Antibiotics, and Gene Therapy
Abstract

Histone deacetylases (HDACs) are validated targets for anticancer therapy as attested by the approval of SAHA and FK228 to treat cutaneous T-cell lymphoma. We recently described the bioassay-guided isolation, structure determination, synthesis, and target identification of largazole, a marine-derived antiproliferative natural product which is a pro-drug that releases a potent HDAC inhibitor, largazole thiol. Here we characterize the anticancer activity of largazole using *in vitro* and *in vivo* cancer models. Screening against the NCI’s 60 cell lines revealed that largazole is particularly active against several colon cancer cell types. Consequently we tested largazole, along with several synthetic analogues, for HDAC inhibition in human HCT116 colon cancer cells. Enzyme inhibition strongly correlated with the growth inhibitory effects, and differential activity of largazole analogues was rationalized by molecular docking to an HDAC1 homology model. Comparative genome-wide transcript profiling revealed a close overlap of genes that are regulated by largazole, FK228 and SAHA. Several of these genes can be related to largazole’s ability to induce cell cycle arrest and apoptosis. Stability studies suggested reasonable bioavailability of the active species, largazole thiol. We established that largazole inhibits HDACs in tumor tissue *in vivo* using a human HCT116 xenograft mouse model. Largazole strongly stimulated histone hyperacetylation in the tumor, showed efficacy in inhibiting tumor growth and induced apoptosis in the tumor. This effect is likely mediated by modulation of levels of cell cycle regulators, by antagonism of the AKT pathway through IRS-1 downregulation, and by reduction of epidermal growth factor receptor levels.
Introduction

Oceans cover most of the Earth’s surface and contain the greatest biodiversity in the world, unmatched by their terrestrial counterpart. Terrestrial organisms such as plants, fungi and soil bacteria have provided mankind at least with the inspiration for approximately half of the currently marketed drugs (Newman and Cragg, 2007). The marine environment, however, is still a largely untapped resource for biomedicine. The unique marine biodiversity surely harbors unprecedented chemical structures awaiting discovery. All these natural products have evolved to effectively modulate their respective targets and thus are biologically validated in a complex bioassay: life and survival in an intricate ecosystem. Humans are only beginning to clinically exploit some chemical secrets found in marine organisms; the first marine anticancer drug (ET-743, trabectedin/Yondelis®) was recently approved in Europe (Casali et al, 2010). Salinosporamide A is a recent example of a promising marine natural product in cancer clinical trials (Fenical et al., 2009; Mayer et al., 2010).

We recently described the isolation and structure determination of largazole from a marine cyanobacterium, an antiproliferative natural product with selectivity for cancer cells over nontransformed cells (Taori et al., 2008). We devised an efficient total synthesis that can provide gram-scale quantities of largazole for rigorous biological evaluation (Ying et al., 2008a). We already established the mode of action of largazole. Largazole is a pro-drug and liberates largazole thiol that can chelate Zn$^{2+}$ in the active site of class I HDACs (Fig. 1) (Ying et al., 2008a). Subsequently we have initiated structure–activity relationship (SAR) studies (Ying et al., 2008b).

HDACs regulate gene transcription by increasing the charge status of histone lysine residues through deacetylation of ε-amino groups, increasing the compactness of the chromatin complex.
Conversely, inhibiting HDACs increases the accessibility of DNA to the transcriptional machinery, thereby modulating gene expression. HDACs can be grouped into Zn$^{2+}$- (class I, II and IV) and NAD$^{+}$- (class III) dependent isoforms (Dokmanovic et al., 2007). Class I HDAC isoforms such as HDAC1 and HDAC3 are overexpressed in various cancers, including colon, and linked to cellular proliferation (Ishihama et al., 2005; Wilson et al., 2006; Senese et al., 2007; Spurling et al., 2008). Despite seemingly nonspecific global transcriptional effects, inhibition of Zn$^{2+}$-dependent HDACs has been shown to produce significant anticancer effects \textit{in vitro} and \textit{in vivo}. SAHA, a broad-spectrum HDAC inhibitor, and more recently FK228 have already been approved for treating cutaneous T-cell lymphoma (Fig. 1) (Marks and Breslow, 2007; Lansigan and Foss, 2010). FK228 has superior potency and selectivity for class I isoforms and requires metabolic activation by disulfide reduction (Fig. 1). Various other HDAC inhibitors are undergoing clinical trials (Senese et al., 2007, Paris et al., 2008). Thus, we have already addressed challenges in drug development associated with marine natural products: supply problem, target identification and target validation. Largazole has attracted tremendous interest from the synthetic and medicinal chemistry communities that rely on natural products groups to discover new promising leads, confirming our findings but also attesting to the potential of largazole and marine natural products in general (Newkirk et al., 2009; Zeng et al., 2010, and references cited therein). Here we characterize the anticancer and HDAC inhibitory activity of largazole in cellular and \textit{in vivo} cancer assay model systems.
Material and Methods

Chemicals and Chemical Synthesis

Chemical reagents. Largazole was synthesized as described (Ying et al., 2008a). Largazole analogues were made by adaptation of this method, which we previously used to synthesize analogues (Ying et al., 2008b). FK228 was synthesized as described (Wen et al., 2008). SAHA was purchased from Selleck (Shanghai, China).

Synthesis of N-methylated analogues of largazole and Synthesis of Phe, His, Asp, and Tyr analogues of largazole. The synthesis of the N-methylated analogues of largazole and synthesis of Phe, His, Asp, and Tyr analogues of largazole can be found in the Supplemental data (Supporting Schemes S1 and S2).

In vitro Assays

Recombinant HDAC1 enzymatic assays. The assays were carried out by Reaction Biology Corp. (Malvern, PA) as previously described (Ying et al., 2008b). Peptide substrate, p53 residues 379–382 (RHKKAc), conjugated with AMC (7-acetoxy-4-methylcoumarin) was used as the fluorogenic substrate at 50 μM assay concentration. Briefly, compounds dissolved in DMSO together with HDAC1 were incubated at 30 °C for 2 h in the reaction buffer which contained 25 mM Tris-HCl, pH 8.0, 137 mM NaCl, 2.7 mM KCl, 1 mM MgCl₂, and 0.25 mg/mL BSA, before adding developer reagent. The free AMC is detected with excitation of 360 nm and emission 460 nm in kinetic mode for 90 min. The initial velocity of an enzyme reaction was normalized and plotted with GraphPad Prism (GraphPad Software, Inc. La Jolla, CA) to derive the IC₅₀ values.

Cell culture. Human colorectal carcinoma cells (HCT116, HT29 and HCT15) were purchased from ATCC (Manassas, VA) and cultured in Dulbecco’s Modified Eagle Medium
(Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (HyClone, Logan, UT) and 1% PSF at 37 °C humidified air and 5% CO₂.

**Cell cycle analysis.** HCT116, HT29 and HCT15 cells were incubated with largazole at various concentrations for 24 h. Cells were pelleted by centrifugation and fixed in ice-cold 70% ethanol. DNA was stained with 10 μg/mL propidium iodide (Invitrogen) in a reaction solution containing 100 μg/mL RNase A (Sigma). Fluorescence emitted from the propidium iodide–DNA complex was quantified using FACScan (Becton Dickinson Medical Systems, Sharon, MA).

**Cell viability assay (MTT).** HCT116, HT29 and HCT15 cells (1 × 10⁴ cells per well) were seeded in 96-well clear bottom plates, and 24 h later the cells were treated with various concentrations of largazole, largazole analogues (100 pM–10 μM) or solvent control. After 48 h of incubation, cell viability was detected using MTT according to the manufacturer’s instructions (Promega, Madison, WI).

**Caspase 3/7 assays.** HCT116 cells were seeded in 96-well white assay plates (1 × 10⁴/well) and 24 h later treated with largazole at various concentrations. After 48 h of treatment, the Caspase-Glo 3/7 assay (Promega) was used to measure caspase 3/7 activity. The reagent was prepared immediately before use. The lysis buffer and luciferase substrate were equilibrated to room temperature and mixed together. The assay plate was equilibrated to room temperature (~10 min). The same volume of Caspase-Glo 3/7 reagent as culture medium was added to each well, the plate was mixed on a plate shaker for ~1 min and incubated at room temperature for 30 min. The luminescence was read using luminescence plate reader (SpectraMax M5, Molecular Devices, Sunnyvale, CA).

**RT-qPCR.** 2 × 10⁵ HCT116 cells (per well) were seeded in 6-well dishes. 24 h later, the cells were treated with largazole (20 nM), SAHA (2 μM), FK228 (20 nM) or solvent control and
incubated for various time points (1–20 h). RNA was isolated using the RNeasy mini kit (Qiagen, Valencia, CA) and cDNA was synthesized from 2 µg of total RNA by using SuperScript II Reverse Transcriptase (Invitrogen) and Oligo(dT)\textsubscript{12–18} Primer (Invitrogen). Real-time PCR was performed in a total reaction volume of 25 µL which included 12.5 µL of TaqMan 2x universal master mix (Applied Biosystems, Foster City, CA), 1.25 µL of 20x TaqMan gene expression assay mix, 2 µl of cDNA and 9.25 µL of sterile water in a 96-well plate (Applied Biosystems) by using the ABI 7300 sequence detection system (Applied Biosystems). The thermocycler program consisted of a 2 min at 50 °C, 10 min at 95 °C, and 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Each assay was carried out in triplicate. For normalization, \textit{GAPDH} expression was used as internal control.

**GeneChip analysis.** RNA samples were subsequently used for GeneChip analysis. RNA concentration was further determined on a NanoDrop Spectrophotometer (NanoDrop Technologies, Inc.) and sample quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.). GeneChip® 3’ IVT Express kit (Affymetrix, Inc.) was used for duplicate sample preparation and the reactions were carried out according to the protocol detailed by Affymetrix (Santa Clara, CA, USA). Briefly, cDNA was synthesized from 250 ng of total RNA and template for \textit{in vitro} transcription (IVT) during which a biotin-modified nucleotide was incorporated. The biotin-labeled aRNA was then purified and fragmented. Samples were hybridized with rotation at 45 °C for 16 h to the Affymetrix GeneChip® Human Genome U133 plus 2.0 arrays. The arrays were washed and stained with the reagents supplied in GeneChip® Hybridization Wash and Stain kit (Affymetrix, Inc.) on an Affymetrix Fluidics Station 450, and scanned with a GeneChip® 7G Scanner (Affymetrix, Inc). Statistical tests were performed using Bioconductor statistical software (http://www.bioconductor.org/) and R
program (R: A language and environment for statistical computing) (Gentleman et al., 2004). Raw data were normalized by Robust Multichip Analysis (RMA) approach (Bolstad et al., 2003). Then the probe set’s detection call was estimated using the Wilcoxon signed rank-based algorithm. Probe sets that are absent in all the study samples were removed from further analyses. A linear modeling approach and the empirical Bayes statistics as implemented in the limma package (Smyth, 2004) in the R software were employed for differential expression analysis. The \( p \)-values obtained were controlled for multiple testing (FDR, false discovery rate) using Benjamini-Hochberg method (Benjamini and Hochberg, 1995). Differentially expressed genes were then ranked by the \( p \)-values, and genes with \( p \)-value less than 0.01 and 2-fold change were considered as differentially expressed genes at a statistically significant level. Hierarchical clustering of the data was computed on log-transformed and normalized data using complete linkage and Pearson correlation distances. Computations and visualization were done using R packages.

**Immunoblot analysis.** HCT116 cells (2 \( \times \) 10\(^5\) per well), HT29 cells (4 \( \times \) 10\(^5\) per well) and HCT15 cells (3 \( \times \) 10\(^5\) per well) were seeded in 60-mm dishes. The next day, largazole, SAHA or FK228 were added to final concentrations from 0 to 10 µM and incubated for 8 h or 24 h. Whole cell lysates were collected using PhosphoSafe buffer (Novagen, Madison, WI). The protein concentration was measured using the BCA Protein Assay kit (Pierce). Lysates containing equal amounts of protein were separated by SDS-PAGE gel (4–12%), transferred to PVDF membranes, probed with antibodies and detected with the SuperSignal West Femto Maximum Sensitivity Substrate (Pierce). Anti-acetyl histone H3 antibody was obtained from Millipore (Billerica, MA). Anti-acetyl-\( \alpha \)-tubulin, \( \alpha \)-tubulin, \( \beta \)-actin, cyclin D1, EGFR, MET, HER-2, p21, as well as secondary (anti-rabbit, anti-mouse) antibodies were from Cell Signaling (Beverly,
MA). Anti-BCL2L11 was from Abcam (Cambridge, MA). Anti-CDK6 and IRS-1 were from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA).

Stability Studies

Materials and general procedures. HPLC-MS was done on a 3200 QTRAP (Applied Biosystems) equipped with a Shimadzu UFLC System. Mouse serum (S7273) and harmine (286044) were purchased from Sigma Aldrich. Pooled CD1 mouse liver (female) microsomes (M1500) were purchased from XenoTech (Lenexa, KS) with protein concentrations of 0.5 mg/mL. HCT116 cell lysates were prepared using PhosphoSafe lysis buffer (Novagen). Protein concentration was determined using the BCA Assay.

Sample preparation for stability studies. Stock solutions of largazole and thiol were prepared by dissolving the compounds in ethanol to give a 1 mg/mL solution. Aliquots of this stock solution were then obtained to afford a 40 μg/mL solution in methanol. Serial dilution of the 40 μg/mL solution in methanol gave standard solutions with concentration of 25, 12.5, 2.5, 1.25, 0.25, 0.125, 0.025, 0.0125 μg/mL. A 1 mg/mL stock solution of the internal standard harmine was prepared in ethanol, which was subsequently used to prepare a 100 μg/mL solution with ethanol. An aliquot of the 100 μg/mL harmine solution was diluted to 100 ng/mL with ethyl acetate to serve as the working internal standard solution.

Plasma stability. In vitro plasma stability of largazole was done using a modification of a published method (Chen et al., 2008). Ten microliters of largazole (25 μg/mL) were added to 100 μL of mouse serum, the solution was vortex-mixed for 15 s and incubated for 0.25 min to 24 h (11 time points). At the end of each incubation period, 400 μL of ethyl acetate were added to each tube, followed by 200 μL harmine to quench the reaction and extract largazole and its
metabolites. Samples were further incubated in a thermomixer at 27 °C, (750 rpm, 5 min) and later centrifuged for 5 min at 1,643 g. The ethyl acetate layer was collected and evaporated to dryness under nitrogen. Samples were reconstituted in 50 μL methanol. A volume of 10 μL of the reconstituted solution was injected into the HPLC-MS system.

**Microsomal stability.** The stability of largazole in the presence of mouse microsomes was determined using an adaptation of a published procedure (Ackley et al., 2004). In brief, microsomes were added to pre-warmed phosphate buffer (100 mM, pH 7.4) at 37 °C. Largazole (3 μL) was added to the microsomal preparation followed by NADPH cofactor solution (1.3 mM NADP, 3.3 mM glucose-6-phosphate, 0.4 U/mL glucose-6-phosphate dehydrogenase, 3.3 mM MgCl₂). The reaction was allowed to proceed for 5, 15, and 30 min at 37 °C (thermomixer, 1050 rpm). The reaction was quenched by addition of ethyl acetate and subsequently spiked with harmine. The zero time point was defined by denaturing the microsomes with ethyl acetate prior to addition of largazole. Incubation of largazole with microsomes alone and denatured microsomes (heated at 95 °C, 5 min) with NADPH cofactor were also performed following the same procedure to determine NADPH dependent metabolism. The maximum amount of the thiol adduct formed from largazole was defined as the amount of the metabolite formed after 5 min in the presence of microsomes alone. The final concentration of the incubation mixture contained 0.5 mg/mL protein concentration and 1 μM largazole.

**Cellular stability.** Aliquots of HCT116 cell lysates were diluted with 25 mM Tris-HCl buffer (pH 8.0) to give a final reaction volume of 100 μL and protein concentration of 0.7 mg/mL. Cell lysate solutions were incubated with 10 μL of largazole (25 μg/mL) for 0.25 min to 24 h (9 time points). Remaining largazole and its metabolites were extracted from the reaction
solution at the end of the incubation periods with ethyl acetate using the same procedure as described for the plasma stability assay.

**Aqueous stability.** The stability of largazole in aqueous solution was determined in 100 mM phosphate buffer pH 4.0, 100 mM phosphate buffer pH 7.0, 100 mM phosphate buffer pH 8.0, Dulbecco’s Modified Eagle Medium containing 10% fetal bovine serum and 1% PSF. One hundred microliter portions of each solution were spiked with 10 μL of largazole solution (25 μg/mL) and allowed to incubate for 0.25 min to 24 h (8 time points). The reaction was quenched at the end of each time point and the remaining largazole was extracted following the ethyl acetate extraction procedure, as in the plasma stability study.

**HPLC-MS parameters.** Analysis of largazole and its metabolites was done using HPLC-MS [column, Onyx Monolithic C18 (3.0 × 100 mm), Phenomenex; solvent, 0.1% aqueous formic acid (solvent A) 0.1% formic acid in MeOH (solvent B); flow rate, 0.5 mL/min; detection by ESIMS in positive ion mode (MRM scan)]. A step-wise gradient elution was employed starting at 60% B and 40% A, then increasing to 100% B at 7 min and maintained at this condition for 8 min. Parameters were optimized prior to analysis using direct syringe infusion. The retention times ($t_R$, min; MRM ion pair) of the analytes and internal standard are as follows: harmine (1.5; 213→171), thiol adduct (2.2; 497→212), free thiol (4.7; 497→212), largazole (6.7; 623→497). Compound dependent parameters used were as follows. Largazole: DP 30.95, EP 9.16, CE 30.95, CXP 9.41, CEP 58.15; thiol adduct and free thiol: DP 49.89, EP 5.32, CE 45.39, CXP 4.13, CEP 17.1; harmine: DP 45.0, EP 9.8, CE 45.0, CXP 3.9, CEP 17.1. Source gas parameters used were as follows: CUR 30.0, CAD Low, IS 4500, TEM 450.0, GS1 60.0, GS2 60.0.
Data analysis. Calibration curves for largazole and its metabolites in the presence of mouse serum, HCT116 cell lysates, and aqueous solutions were generated by least-square linear regression analysis of the analyte peak area and internal standard peak area ratio against the nominal concentration of the standard solutions. A linear regression analysis was performed and the concentration of largazole, thiol, and thiol adduct at each time point was determined through interpolation for plasma, cellular and aqueous stability experiments. The amount of remaining largazole with microsome incubation was determined from the peak area ratio of largazole at \( t_x \) (5, 15, 30 min) and \( t_0 \). The amount of thiol adduct remaining during the microsomal stability assay was determined by calculating the peak area ratio of the thiol adduct at \( t_x \) (5, 15, 30 min) and \( t_0 \). All calculations were done using Analyst 1.4.2 Quantitate Mode.

Tumor Studies

In vivo activity. Female nude mice (nu/nu) of 3–5 weeks old were obtained from Charles River Laboratory (Wilmington, MA) and used for human tumor xenografts. Tumors were established by subcutaneous injection of \( 1 \times 10^6 \) HCT116 cells on the left rear flank of a nude mouse in a volume of 100 µl of sterile saline. Tumor dimensions were measured daily using calipers and tumor volumes were calculated using the formula \( W^2 \times L \times 0.5 \), where width (\( W \)) ≤ length (\( L \)). Tumor growth rate was analyzed based on slopes. The slopes were calculated by the difference between the final and initial tumor volumes and then divided by the number of treatment days: \( (V_n - V_0) / n \). \( V_0 \) is the tumor volume (in mm\(^3\)) when it first becomes apparent (20–100 mm\(^3\)), \( V_n \) is the tumor volume \( n \) days later and \( n \) is the number of treatment days. To determine the largazole dose for efficacy studies, mice were treated with various doses of largazole (0, 1.5, 5, 15, 50 mg/kg) by intraperitoneal injection, 4 h later tumors were harvested, 50 mg of tumor tissue was sonicated in PhosphoSafe lysis buffer and used for immunoblot
analysis probing with anti-acetyl histone H3 antibody as described above. Mice were treated with the optimized dose of 5 mg/kg of largazole or solvent control every day until the tumor size in one dimension reached 15 mm (or the tumor volume exceeded 1,000 mm$^3$) and tumor tissue was harvested the day after final injection. Immunoblot analysis was performed as described above. Tumor tissue for immunohistochemistry was immediately kept in ice-cold 4% paraformaldehyde.

All studies were carried out under the protocol approved by Institutional Animal Care and Use Committee at the University of Florida.

**Immunohistochemistry.** Xenograft tumors were fixed in 4% paraformaldehyde, embedded in paraffin and sectioned. Sections were incubated with a 1:100 diluted caspase 3 primary antibody (Cell Signaling) and a biotinylated 1:200 diluted secondary antibody (Cell Signaling). After incubation with 100–400 μL ABC reagent, the slides were developed in 100–400 μL DAB. The slides were further counterstained with hematoxylin, dehydrated in ethanol and xylene. After the coverslips were mounted, the slides were scanned using ScanScope CS (Aperio Technologies, Vista, CA) and the images were captured using ImageScope (Aperio Technologies). The images were analyzed using MetaMorph (Olympus, Tokyo, Japan). Caspase 3 positive stained cells were quantified on 12 randomly selected fields from two independent tumors.

**Molecular Docking**

Docking studies utilized a previously validated homology model of human HDAC1 (Wang et al., 2005). Structures for the largazole thiols were constructed based on a reported crystal structure of a synthetic intermediate (Seiser et al., 2008). The appropriate side chain atoms were added in
Sybyl (Tripos); then the structure was minimized. This process did not appreciably alter the conformation of the macrocyclic ring. In the case of the Tyr and His analogues, the isopropyl group of the Val side chain was replaced with the appropriate side chain in Sybyl before further minimization. Docking studies were carried out using AutoDock Vina (Trott and Olson, 2010), allowing rotation of the side chains but not within the macrocycle, and using an exhaustiveness value of 25. Figures of docked poses were prepared in PyMol.
Results

Largazole exhibits broad-spectrum yet differential activity in the NCI’s 60 cell line screen. To determine an appropriate cancer model system for initial biological studies, largazole was tested for antiproliferative activity in the National Cancer Institute’s cell viability screen using approximately 60 cancer cell lines derived from various solid and liquid tumors (Paull et al., 1995). While we observed growth inhibition across all cancer cell lines, there was increased cytotoxicity in several cell lines that originated from melanoma, renal and particularly colon cancer (Fig. 2A). We had previously observed potent activity of largazole during our early SAR studies using HCT116 colorectal carcinoma cells (Ying et al., 2008b). Consequently, we used this cell line for further biological studies.

Largazole induces dose-dependent cell cycle arrest and apoptosis in HCT116 colon cancer cells. To rationalize the antiproliferative activity we determined largazole’s effect on the cell cycle using FACS-based DNA content analysis after 24 h treatment. While low concentrations of largazole induced G1 arrest (1–3.2 nM), higher concentrations (≥10 nM) shifted the population distribution significantly to an apparent G2/M arrest (Fig. 2B). The higher concentrations also caused apoptosis as evidenced by sub-G1 and ‘sub-G2’ populations. This result was consistent with the induction of caspase 3/7 activity (Fig. 2C), downstream mediators of programmed cell death.

SAR study identifies critical and tunable elements in the largazole molecule. When we replaced Val with Ala in the largazole structure we previously observed that largazole’s activity is tunable at that position upon appropriate chemical modification (Ying et al., 2008b).
extended these studies by introducing aromatic (Phe, Tyr), acidic (Asp) or basic (His) amino acid residues at that position (Fig. 3A). All of the analogues exhibited reduced activity in HDAC1 and cell viability assays, but to different extents. The antiproliferative activity correlated with hyperacetylation of histone H3 (Lys9/14) and thus cellular HDAC inhibition in HCT116 cells (Fig. 3B). According to our docking studies to an HDAC1 homology model (Wang et al., 2005) with the corresponding thiols, the aromatic moieties (Phe, His, Tyr) were able to fill an adjacent pocket next to Glu-98 of HDAC1 (Fig. 3C–H). While the Phe analogue was the least active largazole analogue of these three compounds, the His analogue was slightly more active likely due to hydrogen bonding interactions with the ω-carboxylic acid unit of Glu-98. The Tyr analogue showed even better activity, presumably due to hydrogen bonding with the amide carbonyl of Glu-98, which is not possible for the Phe analogue (Supplementary Fig. S1). The Tyr analogue had been prepared previously by others (Zeng et al., 2010); however, those studies used a different methodology rather than a validated HDAC1 homology model to rationalize activity, consequently leading to different docking results. In all of our structures, the mercapto group of each largazole thiol analogue chelates the Zn^{2+} located in a tunnel in the active site of HDACs.

To probe the contribution of the two secondary amide hydrogens to largazole’s HDAC inhibitory activity, we synthesized the two corresponding \( N \)-methylated largazoles and subjected them to HCT116 cells to test their effect on cell viability and cellular HDAC activity. Both compounds were 100- to 1,000-fold less active (Fig. 3B). The crystal structure of the largazole macrocycle suggests hydrogen bonding of the Val-NH with the thiazoline substructure (Seiser et al., 2008), which would be lost upon methylation, potentially leading to a conformational change and thereby reduced activity. Furthermore, in all structures obtained with secondary amides, the NH of the Gly-derived unit can form a hydrogen bond to Leu-271 in the enzyme’s active site.
(Fig. 3C–H). The drastically reduced activity of the \(N\)-methylated analogue at that position suggests that this interaction is crucial for HDAC inhibitory activity of largazole.

Expansion of the Phe-containing macrocycle resulted in loss of activity; we tested a dimeric macrocyclic product (Fig. 3A) that was obtained as a minor product during the macrocyclization reaction. The 400-fold drop in activity (Fig. 3B) indicates the importance of the macrocycle size of our HDAC inhibitor. Residual activities for the \(N\)-methylated analogues and the expanded macrocycle are likely due to the presence of war head (upon thioester hydrolysis) and the proper length of the spacer, retaining some (albeit not strong) ability to chelate the zinc ion in the active site of HDACs.

**Transcriptional changes induced by largazole.** Our next goal was to identify key transcriptional targets related to largazole’s antiproliferative activity. To find optimal conditions for a genomic transcript analysis, we first established the optimal dose that can prevent deacetylation of histone H3 after 8 h. We verified by immunoblot that largazole (like FK228 and in contrast to pan-inhibitor SAHA) does not inhibit cellular HDAC6 (class IIb) responsible for tubulin deacetylation (Fig. 4A). The endogenous cell cycle inhibitor p21 is one of the putative markers whose increased expression upon HDAC inhibitor treatment has been linked to the growth-arrest inducing ability of this class of agents (Wilson et al., 2006). We then used this marker to establish the optimal time point for genome-wide transcriptional profiling, monitoring \(p21\) expression in HCT116 cells upon largazole treatment (Supplementary Fig. S2A). For comparison, we also tested side-by-side SAHA and FK228 which we synthesized via a published route (Wen et al., 2008). Largazole and FK228 were about 100-fold more potent in these assays and \(p21\) transcript levels reached a maximum after 10 h (2.55- and 2.13-fold increases,
respectively; Supplementary Fig. S2A). Therefore, we used RNA derived from 10-h treated cells with all three drugs for a comparative global gene expression analysis. At the two-fold threshold, largazole, SAHA and FK228 (directly or indirectly) modulated transcript levels for hundreds of genes in either direction (Fig. 4B). Several upregulated genes are cell cycle inhibitors other than p21 that can be linked to the drug-induced G1 growth arrest, such as p19, p15 and p57 (Supplementary Table S1). Cyclin-dependent kinase 6 (CDK6) and cyclin D1 (CCND1) – encoding its regulatory subunit – were downregulated, which may also contribute to the antiproliferative effects. Elevation of transcript levels of BCL2L11, which codes for a proapoptotic BCL2-protein family member, may in part explain the induction of apoptosis (Supplementary Table S1). Other downregulated genes (Supplementary Table S1) include growth factor receptors (EGFR, HER-2 and MET) and insulin-receptor substrate-1 (IRS-1), which is downstream of insulin growth factor signaling and upstream of PI3K/AKT. The comparative transcriptional profiling revealed a close overlap of genes regulated by these three anticancer agents in HCT116 cancer cells (Supplementary Fig. S2B). Closer inspection of the relatively few genes tentatively categorized as uniquely regulated (Supplementary Fig. S2B) suggested that differences arise mainly from differences in the magnitude of fold-change and may not reflect biological significance (Supplementary Table S2). Gene ontology analysis revealed that overrepresented functional classes of genes induced by largazole include those involved in chromatin assembly, negative regulation of cell cycle and transcription. Using the same analysis, overrepresented classes of genes downregulated by largazole include genes with function in positive regulation of transcription, intracellular protein cascade, positive regulation of nucleic acid metabolism, RNA biosynthesis and metabolism, and epidermal growth factor
receptor activity. We validated several regulated genes by Western blot analysis in a dose-response study after 24 h of treatment (Fig. 4C).

The antiproliferative effects of largazole on other colon cancer cell types correlate with the induction of histone H3 hyperacetylation through HDAC inhibition. To exclude that the inhibition of HDACs in colon cancer cells by largazole is context dependent, we assessed histone H3 acetylation in another susceptible colon cancer cell line (HT29) and another less susceptible cell line (HCT15) based on the NCI-60 data (Fig. 2A). While the acquisition of a malignant phenotype in HCT116 cells is associated with microsatellite instability due to a defect in DNA mismatch repair because of \textit{MLH1} mutations, HT29 cells are characterized by \textit{APC} mutations (among others), leading to chromosomal instability. HCT15 cells exert both microsatellite and chromosomal instability due to mutations in another DNA mismatch repair gene (\textit{MSH6}) but also in the \textit{APC} gene (among others). As seen for HCT116 cells (Fig. 3B), in both HT29 and HCT15 cells the inhibition of histone H3 deacetylation correlates with the antiproliferative effects of largazole (Fig. 5A,B). Additionally, we observed a dose-dependent cell cycle arrest paralleling the effects of largazole on HCT116 cells (Fig. 2B). Concentrations near but slightly below the GI$_{50}$ caused G1 arrest, while higher concentrations induced G2/M arrest (Fig. 5C,D).

Aqueous, plasma, microsomal, and cellular stability of largazole. Stability studies are a prerequisite to predict bioavailability \textit{in vivo}. For largazole, those studies must also take into account the reactive species, largazole thiol, as this is the primary metabolite interacting with the HDACs. We monitored levels of largazole and largazole thiol after extraction from aqueous reaction mixtures with ethyl acetate followed by LC-MS analysis in the MRM mode. Largazole
was remarkably stable in aqueous solution (Supplementary Fig. S3). However, in mouse serum largazole rapidly converted to the largazole thiol (Fig. 6A), suggesting that the thioester hydrolysis is mediated by plasma proteins. We also noticed the formation of a largazole thiol adduct by observing a peak at a retention time different from that of largazole thiol but with the same MRM transition as “free” largazole thiol (Fig. 6B), which initially signaled to us that the adduct formation was reversible. When largazole was incubated with mouse microsomes, we observed a similar trend in that >99% hydrolyzed within 5 minutes. However, it did not hydrolyze when the microsomes were inactivated (95 °C, 5 min) prior to largazole exposure. This suggests that denatured proteins are unable to assist in the hydrolysis reaction. Active microsomal proteins also formed adducts with largazole thiol as observed for other proteins (Fig. 6), and the adduct was stable in the absence of NADPH. In the presence of NADPH, largazole thiol was metabolized by microsomes (approximately 19%, 3% and 1% remaining after 5, 15 and 30 min, respectively; Supplementary Table S3). To establish the stability of largazole in a cellular context, we incubated largazole with whole-cell protein lysate derived from HCT116 cells and measured metabolite levels over time as described above. Largazole levels dropped concomitant with a rise in concentrations of a largazole thiol adduct, suggesting that protein-assisted hydrolysis and subsequent reversible protein binding had occurred (Fig. 6C). These experiments confirmed that – when exposed to cellular proteins – largazole hydrolyzes and generates the reactive largazole thiol.

**In vivo activity and efficacy of largazole.** Acute toxicity studies in *nu/nu* mice indicated that largazole was well tolerated up to the highest concentration tested (50 mg/kg i.p.). The lack of toxicity in these studies prompted us to establish the appropriate largazole dose by examining the
Histone acetylation status in the HCT116 tumor xenograft as a biochemical marker of largazole activity rather than determining the maximum tolerated dose (MTD). Largazole induced hyperacetylation of histone H3 in the tumor after 4 h of treatment in a dose-response manner (Fig. 7A). These studies indicated that largazole at 5 mg/kg inhibited HDAC in vivo in the tumor and that the inhibitor is sufficiently bioavailable. Having established the largazole dose required to cause a biochemical response, we subsequently carried out efficacy studies (daily i.p.). We observed retarded HCT116 tumor growth in vivo (Fig. 7B,C). Immunohistochemistry of control tumors and residual tumors from largazole-treated mice, staining for caspase 3, indicated that largazole induced apoptosis by ca. 10-fold in vivo (Fig. 7D,E). To further compare our results with the mode of action of largazole in vitro, we analyzed protein extracts derived from control and largazole tumors. Consistent with the transcriptional effects observed in cell culture, largazole also induced levels of p15 and decreased levels of cyclin D1 and growth factor receptors such as HER-2 in vivo (Fig. 7F). Furthermore, IRS-1 was strongly downregulated concomitant with inhibition of AKT (T308 and S473) phosphorylation, suggesting that the pro-survival PI3K/AKT pathway is inhibited in vivo (Fig. 7F). This is one potential explanation for the pro-apoptotic effect of largazole leading to caspase 3 activation. Thus, we confirmed that largazole’s mechanisms of action in vitro and in vivo are at least partially overlapping.
Discussion

The marine cyanobacterial metabolite largazole is among the most potent HDAC inhibitors found to date, displaying similar potency as FK228. Both compounds require metabolic activation through thioester hydrolysis or disulfide reduction, respectively. Largazole thiol possesses sub-nanomolar activity (Ying et al., 2008b) against recombinant HDAC1 and is more active than any other marine-derived HDAC inhibitors such as psammaplin (Piña et al., 2003) and azumamides (Nakao et al., 2006; Nakao et al., 2008). In a systematic approach we determined that certain colon cancer cell lines are particularly susceptible to the cytotoxic effects of largazole and presented mechanistic, cellular SAR and in vivo efficacy results in HCT116 colon cancer. We showed that we can tune HDAC1 inhibitory and antiproliferative activities through systematic chemical mutagenesis of the Val residue. We have been able to decipher interactions between the molecule and HDAC1 enzyme by molecular docking. As we had originally suggested (Ying et al., 2008b), the HDAC1 pocket can tolerate various amino acid residues in this position without sudden loss of activity. Such a continuous spectrum of less potent compounds with differential activities may be exploited like SAHA, where more potent analogues proved to be too toxic (Marks and Breslow, 2007). Furthermore, we are now positioned to explore effects on isoform specificity. Our previous and present studies indicated that altering the macrocycle conformation, e.g., by inversion of C17 stereocenter (Ying et al., 2008b) or N-methylation, is detrimental to largazole’s activity. In contrast, there is only limited SAR available for FK228 (Yurek-George et al., 2007).

Investigation of largazole’s effect on the HCT116 transcriptome revealed close similarity to the effects of SAHA and FK228 tested in parallel. It appears likely that any major differences between these HDAC inhibitors would rather manifest themselves in other cell types, in their
effects on non-histone targets and in their metabolic stability and bioavailability. While the cell cycle inhibitor p21 is a standard biochemical marker for HDAC inhibition, we have shown that largazole also induces other cell cycle inhibitors and downregulates CDK6 and its regulatory subunit, cyclin D1, which is overexpressed in many cancers and plays a critical role in tumor development (Tashiro et al., 2007). Thus largazole affects the cell cycle through various mechanisms, which is reflected in the dose-dependent effects on the cell cycle profile. At lower concentrations, modulation of genes involved in G1-S transition appears to dominate the cell cycle behavior while higher concentrations of largazole caused a pronounced G2/M arrest. These differences may be indicative of target-dependent dynamics or secondary effects and may be cell-type dependent. Largazole also downregulates several cancer-associated receptor tyrosine kinases. Specifically, levels of EGFR, HER-2 and MET, all of which are master regulators of many aberrantly activated signaling pathways in cancers, are substantially reduced in a dose-dependent manner. This effect on growth factor receptors is likely one major mechanism by which largazole exerts its potent antiproliferative activity. Numerous specific inhibitors of RTKs, including EGFR (e.g., erlotinib), have been developed to treat different types of cancers. Resistance due to mutations and cross-talk between signaling pathways, however, has caused a problem for the efficacy of these inhibitors and, consequently, combination therapy with several RTK inhibitors is one promising clinical strategy (Baselga, 2006; Kling, 2006). Largazole regulates a subset of these receptors, which may be an advantage over single agent therapy with specific growth factor receptor inhibitors. Another antiproliferative mechanism particularly related to the pro-apoptotic activity of largazole is the inhibition of AKT phosphorylation and thereby inhibition of a prosurvival pathway important to cancer cells. Other studies have shown that a subset of known HDAC inhibitors affects AKT independent of histone acetylation by
disrupting HDAC–protein phosphatase 1 (PP1) complexes (Chen et al., 2005). We have shown here that largazole strongly downregulates IRS-1 in vitro and in vivo, which may be at least partly responsible for AKT inhibition. IRS-1 can activate PI3K and therefore AKT, but also acts as a transcriptional co-factor in promoters of cell cycle genes (Baserga, 2009). IRS-1 was found by ChIPs to associate with the cyclin D1 promoter (Sun and Baserga, 2008), providing a link to why largazole may downregulate cyclin D1. Furthermore, IRS-1 expression is often increased in tumors and IRS-1 plays a crucial role in cancer and transformation. The most effective miRNA (miR145) that targets the 3’ UTR of IRS-1 (among others) dramatically inhibits colon cancer cell growth (Shi et al., 2007).

Largazole’s activity and therefore its therapeutic applicability are not limited to colon cancer cells with mutations in DNA mismatch repair genes (e.g., HCT116). HT29 cells which display chromosomal instability are similarly susceptible as HCT116 cells, and in all cases HDAC inhibition also correlated with largazole’s antiproliferative activity even in less susceptible colon cells such as HCT15.

We have shown that largazole is relatively stable in aqueous solution; however, its hydrolysis appears to be protein-assisted; it occurs in plasma and in the presence of active microsomal and cellular proteins. In comparison with FK228, which is reduced to redFK228 intracellularly, largazole thioester hydrolysis expectedly occurs faster to generate largazole thiol. However, in contrast to redFK228 (Furumai et al., 2002), largazole thiol is fairly stable, which is consistent with our observation that direct treatment of cancer cells with largazole thiol has strong antiproliferative activity. It has been reported that redFK228 and a related thiol analogue exhibit much weaker activity due to limited bioavailability and require delivery as a prodrug (Yurek-George et al., 2007; Furumai et al., 2002). Interestingly, largazole thiol is largely bound to
proteins which may serve dual function: adduct formation (via thioester or disulfide formation, or non-covalently) could “re-protect” largazole thiol (deprotection – reprotection or “trans-protection” strategy), but protein binding partners can also act as carrier proteins that release the active largazole thiol at the site of action, which (depending on the nature of the adduct) may occur hydrolytically, reductively, or by shifting the equilibrium towards formation of an HDAC–thiol complex because of the high enzyme affinity of the thiol. While reversibility is only suggested indirectly since HDAC inhibition was observed in the tumor, reversible formation of a thiol adduct is an interesting strategy which may be explored for the development of other drugs as well. We have demonstrated that largazole thiol reaches the tumor and strongly inhibits HDACs responsible for histone H3 acetylation, validated the regulation of several proliferation and apoptosis related target genes in vivo and demonstrated in vivo efficacy.

In ongoing interactive chemical and biological studies we are systematically exploring the application of largazole and its derivatives to treat other types of cancer and study the detailed mechanisms of action downstream of HDAC inhibition. We are also extending our discovery beyond cancer and comprehensively apply largazole-based HDAC inhibitors to other disease indications where cellular reprogramming of transcription may be advantageous.
Acknowledgments

We thank Jason Li for assistance with bioinformatics.
References


Footnotes:

This work was supported by the National Institutes of Health National Cancer Institute [Grant R01CA138544].

Data Deposition: The microarray data has been deposited in the Gene Expression Omnibus (GEO) database (accession no. GSE22061).
**Legends for Figures**

Fig. 1. Structures of largazole, FK228 and SAHA and modes of activation of largazole and FK228 to generate largazole thiol and redFK228, respectively.

Fig. 2. Antiproliferative activity of largazole. (A) Performance of largazole across cell lines in the NCI-60 screen using three different measures of antiproliferative activity (growth-inhibitory effect, GI$_{50}$; cytostatic effect, TGI; cytotoxic effect, LC$_{50}$; concentration in M) (Paull et al., 1995). Color represents activity on the indicated continuous color scale based on log$_{10}$ transformed values. (B) Largazole induces G1 arrest at lower concentration and G2/M arrest at higher concentration in HCT116 cells (24 h treatment). (C) Largazole induces apoptosis in HCT116 cells. HCT116 cells were treated for 48 h with largazole and apoptosis measured as function of caspase 3/7 activity using Caspase-Glo 3/7. Error bars indicate S.D. from triplicate experiments.

Fig. 3. SAR studies. (A) Structures of largazole and its analogues. (B) Antiproliferative activity in HCT116 cells, recombinant HDAC1 and cellular HDAC inhibition (8 h treatment) by largazole and its analogues. [Note: Thioesters were tested in the HDAC1 assay. The corresponding thiols are approximately 10-times more potent since the activity arises from the fraction that is hydrolyzed under assay conditions (equal hydrolysis rate assumed) (Ying et al., 2008b)] (C–H) Molecular docking showing binding modes from different angles (C–E vs. F–H). (C,F) Largazole thiol, (D,G) His analogue of largazole thiol, and (E,H) Tyr analogue of largazole thiol docked into a homology model of HDAC1 (Wang et al., 2005) using AutoDock Vina (Trott
and Olson, 2010). In all cases the side chain bearing the thiol group interacts with the zinc ion (purple) resident at the active site of the enzyme. In all structures, the NH of the Gly derived unit is able to form a hydrogen bond with Leu-271 (solid red line). In addition, the His and Tyr residues of the analogues are able to form hydrogen bonds to the ω-carboxylic acid functionality (side chain) and amide carbonyl oxygen of Glu 98, respectively (D,G vs. E,H; solid red line).

**Fig. 4.** Transcriptional analysis of HDAC inhibition in HCT116 cells and validation. (A) Dose optimization. Cells were treated with various concentrations of largazole, SAHA or FK228 for 8 h, protein lysates collected and analyzed by immunoblot analysis for histone H3 acetylation as a measure for class I HDAC inhibition, and for acetyl-α-tubulin as a measure of HDAC6 (class IIb) inhibition. Based on this analysis, it was estimated that largazole and FK228 are effective at 20 nM and SAHA at 2 μM. (B) Hierarchical cluster analysis of probe sets with ≥2-fold mRNA changes for any of the three HDAC inhibitors. Transcriptome analysis (Affymetrix GeneChip® Human Genome U133 plus 2.0 arrays) was carried out with duplicate biological samples. (C) Validation of transcriptional changes induced by largazole on the protein level (24 h treatment).

**Fig. 5.** Effects of largazole on other colon cancer cell types with different susceptibility. (A,B) Effect of largazole on histone H3 acetylation in (A) HT29 cells and (B) HCT15 cells upon 8 h treatment, assessed by immunoblot analysis. (C,D) DNA content analysis. As observed for HCT116 cells, largazole induces G1 arrest at lower concentration and G2/M arrest at higher concentration in (C) HT29 cells and in (D) HCT15 cells (24 h treatment). Required largazole concentrations differ but correspond to the concentration range that induces histone H3 hyperacetylation (A,B).
Fig. 6. Largazole stability and activation studies. (A–C) Largazole and its metabolites were extracted with ethyl acetate, subjected to LC-MS and monitored using compound-specific MRM mode with harmine as internal standard. (A) Stability of largazole, liberated thiol, and thiol adduct in mouse serum. Largazole hydrolyzed in the presence of mouse serum to yield the largazole thiol adduct at earlier time points. Largazole thiol was generated from the largazole thiol adduct and peaked with 1 h incubation. (B) LC-MS profile of largazole and metabolites. On the basis of specific MRM transitions and retention times, largazole and metabolites were identified. Largazole thiol and largazole thiol adduct showed the same MRM transition with different retention times. An authentic standard of largazole thiol eluted at $t_R$ 4.7 min and also formed the adduct ($t_R$ 2.2 min) when added directly to serum. (C) Cellular stability of largazole measured upon exposure to HCT116 protein lysate (0.7 mg/mL).

Fig. 7. *In vivo* activity using a HCT116 xenograft mouse model. (A) *In vivo* dose optimization. Largazole was dissolved in DMSO and injected i.p. into subcutaneous HCT116 tumor-bearing *nu/nu* mice. After 4 h, tumors were removed and analyzed for histone H3 acetylation by immunoblot analysis. (B,C) Efficacy studies. Tumor-bearing mice ($n = 9$) were injected with largazole (5 mg/kg) by daily i.p. till the end point was reached in the control group (tumor length $\geq$ 15 mm or tumor volume $\geq$ 1,000 mm$^3$). Graphs depict (B) time-dependent tumor volume and (C) overall growth rate after 2 weeks of treatment. Error bars indicate S.E.M. (D,E) Immunohistochemistry. (D) Xenograft tumor sections were mounted on glass slides, stained for caspase-3 and counterstained with hematoxylin. Representative images are shown. (E) Caspase-3 positive stained cells were quantified on 12 randomly selected fields from two independent
tumors from each group. Error bars indicate S.E.M. (F) Biochemical analysis of largazole-induced protein changes in the tumor. Tumor tissue was harvested after 2 weeks of largazole or vehicle treatment and protein extracts analyzed by immunoblot analysis.
Figure 3
Figure 4
Figure 5
Figure 6
Figure 7

A. *In vivo* dose optimization

Largazole-treated HCT116 tumor

B. Efficacy (HCT116 tumor)

![Graph showing tumor volume (mm^3) over days for Vehicle and Largazole treatments.](image)

C. Tumor growth rate (mm^3/day)

![Bar graph comparing Vehicle and Largazole treatments.](image)

D. **Vehicle**

![Image showing caspase 3 staining in Vehicle treatment.](image)

**Largazole**

![Image showing caspase 3 staining in Largazole treatment.](image)

E. Relative caspase 3 staining

![Bar graph comparing Vehicle and Largazole treatments.](image)

F. Immunoblot analysis

![Immunoblot images for HER-2, p15, CCND1, IRS-1, pAKT (T308), pAKT (S473), AKT, β-Actin.](image)
SUPPLEMENTAL DATA

Anti-Colon Cancer Activity of Largazole, a Marine-Derived Tunable Histone Deacetylase Inhibitor

Yanxia Liu, Lilibeth A. Salvador, Seongrim Byeon, Yongcheng Ying, Jason C. Kwan,
Brian K. Law, Jiyong Hong, and Hendrik Luesch

Department of Medicinal Chemistry, University of Florida, 1600 SW Archer Road, Gainesville, Florida 32610 (Y.L., L.A.S., J.C.K., H.L.), Department of Chemistry, Duke University, 124 Science Drive, Durham, North Carolina 27708 (S.B., Y.Y., J.H.), Department of Pharmacology & Therapeutics, University of Florida, 1600 SW Archer Road, Gainesville, Florida 32610 (B.K.L.)
Synthesis of N-methylated analogues of largazole. The N-methylated analogues of largazole were synthesized as shown in Scheme S1.

Scheme S1. Synthesis of the N-methylated analogues of largazole. *Reagents and conditions:* (a) NaH, Mel, DMF, 25 °C, 48 h; (b) TFA, CH₂Cl₂, 25 °C, 1 h; (c) 3, DMAP, CH₂Cl₂, 25 °C, 1 h, 93% for three steps (4b); (d) 2,4,6-trichlorobenzoyl chloride, Et₃N, THF, 0 °C, 1 h; then N-Boc-N-methyl-L-valine (for 5a), N-Boc-L-valine (for 5b), DMAP, 25 °C, 10 h, 95% (5a), 99% (5b); (e) 0.5 N LiOH, THF, H₂O, 0 °C, 3 h; (f) TFA, CH₂Cl₂, 25 °C, 2 h; (g) HATU, HOAt, i-Pr₂NEt, CH₂Cl₂, 25 °C, 32 h, 35% for three steps (8a), 53% for three steps (8b); (h) 9, Grubbs' second-generation catalyst (50 mol%), toluene, reflux, 4 h, 35% (10a), 49% (10b).

**Compound 10a.** $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 7.63 (s, 1H), 6.70 (brs, 1H), 6.29 (d, $J$ = 6.8 Hz, 1H), 5.82 (dt, $J$ = 15.2, 6.8 Hz, 1H), 5.69 (t, $J$ = 9.4 Hz, 1H), 5.47 (dd, $J$ = 15.2, 8.4 Hz, 1H), 4.97 (dd, $J$ = 16.8, 7.2 Hz, 1H), 4.37 (dd, $J$ = 18.8, 4.8 Hz, 1H), 4.34 (d, $J$ = 11.2 Hz, 1H), 3.38 (d, $J$ = 11.6 Hz, 1H), 2.95 (s, 3H), 2.88 (t, $J$ = 6.8 Hz, 2H), 2.71 (dd, $J$ = 14.8, 10.4 Hz, 1H), 2.55-2.50 (m, 3H), 2.42-2.34 (m, 1H), 2.29 (q, $J$ = 7.2 Hz, 2H), 1.70 (s, 3H), 1.66-1.61 (m, 2H), 1.29-1.25 (m, 8H), 1.25 (d, $J$ = 6.8 Hz, 3H), 1.10 (d, $J$ = 7.2 Hz, 3H), 0.88 (t, $J$ = 7.2 Hz, 2H); HRMS (FAB) found 636.2465 [calcd for C$_{30}$H$_{44}$N$_4$O$_5$S$_3$ (M)$^+$ 636.2474].

**Compound 10b.** $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 7.74 (s, 1H), 7.51 (d, $J$ = 8.4 Hz, 1H), 6.01 (dd, $J$ = 10.8, 6.8 Hz, 1H), 5.85 (dt, $J$ = 15.6, 6.8 Hz, 1H), 5.60 (dd, $J$ = 15.6, 7.2 Hz, 1H), 5.23 (d, $J$ = 16.0 Hz, 1H), 4.46 (dd, $J$ = 8.0, 3.6 Hz, 1H), 4.20 (d, $J$ = 16.0 Hz, 1H), 4.05 (d, $J$ = 11.2 Hz, 1H), 3.32 (d, $J$ = 11.6 Hz, 1H), 2.94-2.88 (m, 2H), 2.86 (s, 3H), 2.56-2.45 (m, 2H), 2.52 (t, $J$ = 7.6 Hz, 2H), 2.31 (dt, $J$ = 14.0, 7.2 Hz, 2H), 2.11-2.07 (m, 1H), 1.85 (s, 3H), 1.68-1.63 (m, 2H), 1.32-1.26 (m, 8H), 0.88 (t, $J$ = 7.2 Hz, 3H), 0.70 (d, $J$ = 6.8 Hz, 3H), 0.66 (d, $J$ = 6.8 Hz, 3H); HRMS (FAB) found 636.2469 [calcd for C$_{30}$H$_{44}$N$_4$O$_5$S$_3$ (M)$^+$ 636.2474].
Synthesis of Phe, His, Asp, and Tyr analogues of largazole. The Phe, His, Asp, and Tyr analogues of largazole were synthesized as shown in Scheme S2.

**Scheme S2.** Synthesis of the Phe, His, Asp, and Tyr analogues of largazole. *Reagents and conditions:* (a) 2,4,6-trichlorobenzoyl chloride, Et$_3$N, THF, 0 °C, 1 h; then Boc-L-Phe-OH (for 11a), N-Boc-L-His-OH (for 11b), Fmoc-Asp(OrBu)-OH (for 11c), Fmoc-Try(tBu)-OH (for 11d), DMAP, 25 °C, 10 h, 100% (11a), 62% (11b), 100% (11c), 96% (11d); (b) 0.25 N LiOH, THF, H$_2$O, 0 °C, 3 h; (c) For 13a and 13b: TFA, CH$_2$Cl$_2$, 25 °C, 2 h; For 13c and 13d: Et$_2$NH, CH$_3$CN, 25 °C, 2 h; (d) HATU, HOAt, i-Pr$_2$NEt, CH$_2$Cl$_2$, 25 °C, 32 h, 30% for three steps (14a), 12% for three steps (14b), 18% for three steps (14c), 18% for three steps (14d); (e) (Boc)$_2$O, Et$_3$N, DMAP, 50 °C, 1 h, 41%; (f) 9, Grubbs’ second-generation catalyst (50 mol%), toluene, reflux, 4 h, 59% (15a), 27% (16b), 34% (15c), 47% (15d); (g) TFA, CH$_2$Cl$_2$, 25 °C, 1–4 h, 39% (17b), 41% (16c), 100% (16d).
Compound 15a. $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 7.65 (s, 1H), 6.88 (d, $J$ = 6.8 Hz, 2H), 6.78-6.62 (m, 3H), 6.01 (brs, 1H), 5.81 (dt, $J$ = 14.4, 6.8 Hz, 1H), 5.72 (m, 1H), 5.53 (dd, $J$ = 15.6, 6.8 Hz, 1H), 4.90 (m, 1H), 4.80 (dd, $J$ = 17.6, 6.4 Hz, 1H), 4.37 (d, $J$ = 16.8 Hz, 1H), 4.13 (d, $J$ = 11.6 Hz, 1H), 3.26 (d, $J$ = 11.6 Hz, 1H), 3.20 (dd, $J$ = 13.6, 2.8 Hz, 1H), 3.06 (dd, $J$ = 13.6, 5.2 Hz, 1H), 2.91-2.81 (m, 2H), 2.66-2.52 (m, 2H), 2.49 (t, $J$ = 7.2 Hz, 2H), 2.27 (q, $J$ = 6.8 Hz, 2H), 1.83 (s, 3H), 1.63 (m, 2H), 1.28-1.23 (m, 8H), 0.87 (t, $J$ = 6.8 Hz, 3H); HRMS (FAB) found 670.2316 [calcd for C$_{33}$H$_{42}$N$_4$O$_5$S$_3$ (M)$^+$ 670.2317].

Phe-dimer. $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 7.81 (s, 2H), 7.11-6.89 (m, 10H), 5.88-5.68 (m, 1H), 5.59-5.55 (m, 3H), 5.19 (d, $J$ = 10.4 Hz, 1H), 5.16 (d, $J$ = 17.2 Hz, 2H), 4.93 (dd, $J$ = 15.2, 6.4 Hz, 2H), 4.59 (dd, $J$ = 15.6, 4.4 Hz, 2H), 4.19 (q, $J$ = 5.6 Hz, 2H), 3.62 (d, $J$ = 11.6 Hz, 2H), 3.14 (d, $J$ = 11.6 Hz, 2H), 2.87 (t, $J$ = 7.6 Hz, 2H), 2.87-2.82 (m, 4H), 2.69-2.56 (m, 4H), 2.52 (t, $J$ = 7.2 Hz, 2H), 2.32-2.24 (m, 2H), 1.62-1.56 (m, 8H), 1.32-1.22 (m, 8H), 0.88 (t, $J$ = 6.8 Hz, 3H); MS (ESI) found 1177.4 [calcd for C$_{56}$H$_{66}$N$_8$O$_9$S$_5$Na (M+Na)$^+$ 1177.36].

Compound 17b. $^1$H NMR (400 MHz, CD$_3$OD) $\delta$ 8.01 (s, 1H), 7.72 (s, 2H), 7.54 (brs, 1H), 7.17 (brs, 1H), 5.81 (dt, $J$ = 15.2, 7.6 Hz, 1H), 5.76-5.71 (m, 1H), 5.62 (dd, $J$ = 15.2, 6.8 Hz, 1H), 5.16 (d, $J$ = 17.6 Hz, 1H), 4.96 (m, 1H), 4.39 (d, $J$ = 17.6 Hz, 1H), 3.75 (d, $J$ = 11.6 Hz, 1H), 3.48 (m, 1H), 3.14 (m, 1H), 3.04 (dd, $J$ = 16.8, 10.8 Hz, 1H), 2.93 (m, 2H), 2.72 (dd, $J$ = 16.8, 2.0 Hz, 1H), 2.55 (t, $J$ = 7.6 Hz, 2H), 2.31 (dt, $J$ = 12.0, 6.8, 2H), 1.76 (s, 3H), 1.68-1.59 (m, 2H), 1.32-1.29 (m, 8H), 0.90 (t, $J$ = 7.2 Hz, 3H); HRMS (FAB) found 661.2305 [calcd for C$_{30}$H$_{41}$N$_6$O$_5$S$_3$ (M+H)$^+$ 661.2300].
Compound 16c. $^1$H NMR (500 MHz, CD$_3$OD) δ 8.03 (s, 1H), 7.69 (d, $J$ = 6.0 Hz, 1H), 5.80 (dt, $J$ = 15.0, 6.5 Hz, 1H), 5.63-5.59 (m, 2H), 4.98 (dd, $J$ = 17.5, 7.0 Hz, 1H), 4.59 (m, 1H), 4.48 (d, $J$ = 17.5 Hz, 1H), 3.85 (d, $J$ = 11.5 Hz, 1H), 3.41 (d, $J$ = 11.0 Hz, 1H), 2.92 (ddd, $J$ = 7.5, 7.5, 4.0 Hz, 2H), 2.85-2.67 (m, 2H), 2.55 (t, $J$ = 7.5 Hz, 2H), 2.31 (q, $J$ = 6.5 Hz, 2H), 1.77 (s, 3H), 1.64 (t, $J$ = 7.0 Hz, 2H), 1.32-1.29 (m, 8H), 0.91 (t, $J$ = 7.0 Hz, 3H); HRMS (ESI) found 639.2105 [calcd for C$_{28}$H$_{39}$N$_4$O$_7$S$_3$ (M+H)$^+$ 639.1981].

Compound 16d. $^1$H NMR (500 MHz, CDCl$_3$) δ 7.71 (s, 1H), 7.30 (d, $J$ = 6.0 Hz, 1H), 6.71 (d, $J$ = 8.0 Hz, 2H), 6.26 (brs, 1H), 6.23 (d, $J$ = 8.5 Hz, 2H), 5.84 (dt, $J$ = 14.5, 7.0 Hz, 1H), 5.70 (m, 1H), 5.55 (dd, $J$ = 15.5, 6.5 Hz, 1H), 4.94-4.89 (m, 2H), 4.37 (dd, $J$ = 17.5, 3.5 Hz, 1H), 4.15 (d, $J$ = 11.5 Hz, 1H), 3.28 (d, $J$ = 11.5 Hz, 1H), 3.14 (dd, $J$ = 13.5, 2.0 Hz, 1H), 2.98 (ddd, $J$ = 14.5, 5.5 Hz, 1H), 2.92-2.81 (m, 1H), 2.67 (m, 2H), 2.50 (t, $J$ = 7.0 Hz, 1H), 2.29 (q, $J$ = 7.0 Hz, 1H), 1.84 (s, 3H), 1.61 (m, 2H), 1.30-1.26 (m, 8H), 0.88 (t, $J$ = 7.5 Hz, 3H); HRMS (FAB) found 687.2496 [calcd for C$_{33}$H$_{43}$N$_4$O$_6$S$_3$ (M+H)$^+$ 687.2344].
Supplementary Tables

**Supplementary Table S1.** Relative mRNA expression of selected genes in HCT116 cells upon largazole (20 nM) treatment (10 h)

<table>
<thead>
<tr>
<th>Category</th>
<th>Symbol</th>
<th>Fold change$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell cycle</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>CDKN1A (p21)</em></td>
<td>+1.82 (+2.45$^b$)</td>
</tr>
<tr>
<td></td>
<td><em>CDKN1C (p57)</em></td>
<td>+3.19, +2.99, +2.78, +2.73, +2.54</td>
</tr>
<tr>
<td></td>
<td><em>CDKN2B (p19)</em></td>
<td>+3.71</td>
</tr>
<tr>
<td></td>
<td><em>CDKN2D (p15)</em></td>
<td>+3.26</td>
</tr>
<tr>
<td></td>
<td><em>CDK6</em></td>
<td>–2.05</td>
</tr>
<tr>
<td></td>
<td><em>CCND1</em></td>
<td>–2.01</td>
</tr>
<tr>
<td>Growth factor receptor signaling</td>
<td><em>EGFR</em></td>
<td>–2.54</td>
</tr>
<tr>
<td></td>
<td><em>ERBB2 (HER-2)</em></td>
<td>–2.28</td>
</tr>
<tr>
<td></td>
<td><em>ERBB3</em></td>
<td>–2.02, –2.00</td>
</tr>
<tr>
<td></td>
<td><em>HGFR (MET)</em></td>
<td>–2.66, –2.53</td>
</tr>
<tr>
<td></td>
<td><em>IRS-1</em></td>
<td>–7.90, –5.41, –4.36</td>
</tr>
<tr>
<td>Apoptosis</td>
<td><em>BCL2L11</em></td>
<td>+3.60, +3.10, +3.07, +2.38</td>
</tr>
</tbody>
</table>

$^a$ For some transcripts there were several probes. $^b$ Value from RT-qPCR.
Supplementary Table S2. Genes regulated ≥2-fold by largazole but to a lesser extent by SAHA and FK228 in HCT116 cells (see Supplementary Fig. S2B)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene ID</th>
<th>Relative mRNA expression (fold change)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Largazole</td>
</tr>
<tr>
<td>CD24</td>
<td>CD24 molecule</td>
<td>2.54</td>
</tr>
<tr>
<td>COL1A1</td>
<td>collagen, type I, alpha 1</td>
<td>2.42</td>
</tr>
<tr>
<td>VAMP1</td>
<td>vesicle-associated membrane protein 1 (synaptobrevin 1)</td>
<td>2.26</td>
</tr>
<tr>
<td>HIST2H2BE</td>
<td>histone cluster 2, H2be</td>
<td>2.15</td>
</tr>
<tr>
<td>PRAGMIN</td>
<td>homolog of rat pragma of Rnd2</td>
<td>2.14</td>
</tr>
<tr>
<td>DDX26B</td>
<td>DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26B</td>
<td>2.12</td>
</tr>
<tr>
<td>MSRB3</td>
<td>methionine sulfoxide reductase B3</td>
<td>2.06</td>
</tr>
<tr>
<td>PPP4R4</td>
<td>protein phosphatase 4, regulatory subunit 4</td>
<td>2.05</td>
</tr>
<tr>
<td>HSPA1A</td>
<td>heat shock 70kDa protein 1A</td>
<td>2.03</td>
</tr>
<tr>
<td>LOC100288387</td>
<td>similar to e-jun</td>
<td>2.03</td>
</tr>
<tr>
<td>TLE3</td>
<td>transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)</td>
<td>2.03</td>
</tr>
<tr>
<td>ZNF18</td>
<td>zinc finger protein 18</td>
<td>2.03</td>
</tr>
<tr>
<td>EGLN3</td>
<td>egl nine homolog 3 (C. elegans)</td>
<td>2.01</td>
</tr>
<tr>
<td>SLC46A3</td>
<td>solute carrier family 46, member 3</td>
<td>2.01</td>
</tr>
<tr>
<td>MPP7</td>
<td>membrane protein, palmitoylated 7 (MAGUK p55 subfamily member 7)</td>
<td>−2.02</td>
</tr>
<tr>
<td>IPP</td>
<td>intracisternal A particle-promoted polypeptide</td>
<td>−2.03</td>
</tr>
<tr>
<td>C8orf58 ///</td>
<td>chromosome 8 open reading frame 58 ///</td>
<td>−2.07</td>
</tr>
<tr>
<td>PDLIM2</td>
<td>PDZ and LIM domain 2 (mystique)</td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Table S3.** Microsomal stability of largazole and largazole thiol adduct

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Largazole</th>
<th>Largazole thiol adduct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Microsomes only</td>
<td>Denatured microsomes + NADPH</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>0.13 ± 0.05</td>
<td>90 ± 17</td>
</tr>
<tr>
<td>15</td>
<td>0.07 ± 0.02</td>
<td>76 ± 21</td>
</tr>
<tr>
<td>30</td>
<td>0.03 ± 0.00</td>
<td>69 ± 14</td>
</tr>
</tbody>
</table>

*a* Assays were done in triplicate. Values are expressed as % remaining. Mean values are shown ± S.D. *b* Defined as maximum amount of thiol adduct formed from largazole.
Supplementary Figures

Supplementary Fig. S1. Molecular Docking with selected largazole thiol analogues. The Phe and Tyr analogue of largazole thiol are docked into a homology model of HDAC1 and an overlay of both structures is shown. The Phe analogue of largazole thiol is able to bind in a similar conformation as the Tyr analogue (see Fig. 3H), but is unable to form a hydrogen bond with Glu 98.
Supplementary Fig. S2. Transcriptional analysis of HCT116 cells treated with largazole, SAHA or FK228. (A) Time point optimization for transcriptomic analysis based on p21 marker expression. Cells were treated for various time periods with drug concentrations that effectively inhibit cellular HDACs responsible for histone H3 hyperacetylation (see Fig. 4A). RNA was isolated and subjected to TaqMan-based RT-qPCR analysis for p21. Highest levels of p21 transcript were induced after 10 h of treatment by all three HDAC inhibitors. (B) The Venn diagram depicts overlapping and non-overlapping genes that are modulated (directly or indirectly) by these HDAC inhibitors (20 nM largazole, 20 nM FK228, 2 μM SAHA) upon 10 h of treatment (≥ 2-fold change, p < 0.01).
Supplementary Fig. S3. Stability of largazole in aqueous solution. Largazole was added to cell growth medium or phosphate buffer at pH 4.0, 7.0, 8.0. Largazole was stable over a period of 24 h and no significant hydrolysis was observed. Largazole thiol was not detected by LC-MS. Error bars indicate S.D. from duplicate experiments.