Journal of Medicinal Chemistry Article

J. Med. Chem. **2010,** *53,* 5121–5128 **5121**
DOI: 10.1021/jm9014179

Isolation, Structure, and Functional Elucidation of a Modified Pentapeptide, Cysteine Protease Inhibitor (CPI-2081) from *Streptomyces Species* 2081 that Exhibit Inhibitory Effect on Cancer Cell Migration

Jay Prakash Singh,[†,⊥] Sudarsan Tamang,[‡,#] P. R. Rajamohanan,*[,§] N. C. Jima,[§] Goutam Chakraborty,[∥,▽] Gopal C. Kundu,[∥] Sushma M. Gaikwad,[†] and Mohamad I. Khan*[,†]

[†]*Division of Biochemical Sciences, National Chemical Laboratory, Dr. Homi Bhabha Road, Maharastra, Pune 411008, India,* [‡]*Physical Chemistry Division, National Chemical Laboratory, Pune 411008, India,* [§]*Central NMR Facility, National Chemical Laboratory, Pune 411008, India, and* [∥]*National Centre for Cell Science, NCCS Complex, Pune University Campus, Pune 411007, India.* [⊥]*Present address: Section of Comparative Medicine, Yale University School of Medicine, New Haven, Connecticut 06511.* [#] *Present address: CEA Grenoble, LEMOH/INAC/SPrAM, 17 rue des Martyrs, 38054 Grenoble, France.* [▽]*Present address: Cell Biology Division, Memorial Sloan Kettering Cancer Center (MSKCC), New York, New York*

Received September 23, 2009

Cysteine proteases play an important role in cell migration and tumor metastasis. Therefore, their inhibitors are of colossal interest, having potential to be developed as effective antimetastatic drugs for tumor chemotherapy. Traditionally, secondary metabolites from streptomyces show a wide range of diversity with respect to their biological activity and chemical nature. In this article, we have described the characterization of small molecule cysteine protease inhibitor, CPI-2081 (compound **1**), a mixture of two novel pentapeptides, compound **1a** (736.71 Da), and compound **1b** (842.78 Da), isolated from *Streptomyces species* NCIM2081, following solvent extraction and repeated HPLC based on C18 chemistry, and completely characterized using a variety of both 1D and 2D NMR spectroscopy. Further, it was found that nanomolar concentration of compound **1** is able to inhibit papain hydrolytic activity. Also, compound **1** significantly inhibits tumor cell migration at sub cytotoxic concentration, indicating its remarkable potential to be developed as antimetastatic drug, which will make chemotherapy more localized and specific, thereby minimizing the hazardous side effects on normal tissues.

Introduction

Deregulated expression of proteases is convincingly reported to be involved in tumor progression and metastasis.[1−4] Cysteine proteases such as cathepsin B expression is increased in many human cancers.[5,6] The predominant expression of cathepsin K in osteoclasts has rendered the enzyme a major target for the development of novel antiresorptive drugs, and cathepsin S appears to be a considerable drug target for various inflammatory diseases including rheumatoid arthritis.[7−9] Several studies have shown that parasites such as *Trypnosoma*, *Leishmania*, and *Plasmodium* also take advantage of cysteine protease activity to survive the hostile microenvironment by immune-evasion, hydrolysis of host proteins, enzyme activation, and cellular invasion.[10−15] Therefore, it is reasonable to assume that a good cysteine protease inhibitor can be a useful compound to target the aforementioned disorders. Among all kind of sources for natural bioactive compounds, actimomycetes always had a competitive edge over others with respect to their ability to produce bioactive small molecules for drug development. There are many novel bioactive compounds isolated from actinomycetes in recent past. Lanophylins, A1, A2, B1, and B2 were isolated from *Streptomyces* sp.[16] Lanophylins were the first naturally occurring (3*E*)-hexadecylmethyllidine-2-methyl-1-pyrroline backbone that inhibits lanosterol synthease activity important for cholesterol maintenance. Some alkaloids isolated from *Dactylosporangium* sp. inhibited recombinant human kinase and serine proteases, cathepsin G, and chymotrypsin.[17,18] Mannopeptimycins glycopeptides exhibiting antibacterial activities were isolated from *Streptomyces hygroscopicus*.[19] A few small molecule nonproteinacious cysteine protease inhibitors, which are used in biomedical research as protease inhibitor, have also been isolated from *Streptomyces* sp.[20,21] FA-70C, an antipain (cysteine protease inhibitor) analogue, was isolated from the culture supernatant of *Streptomyces* species strain FA-70. It inhibits Arg-gingipain (Rgp), a key cysteine protease produced by *Porphyromonas gingivalis*, a major pathogen of advanced periodontal diseases.[22] A new compound migrastatin, an inhibitor of cell migration, was isolated from *Streptomyces* species Mk929-43F1.[23] A series of successful attempts to isolate protease inhibitors from *Streptomyces* species led to discovery of few small molecule cysteine and serine protease inhibitors such as leupeptin, E64, and antipain.[20,21,24] Because of the requirement of high molar concentration for inhibitory activity, these compounds have not found their applicability in therapeutic formulations to target proteases in human diseases and are maximally restricted to use in research laboratories as protease inhibitor cocktails. We report identification, isolation, and structure−functional characterization of a novel small molecule cysteine protease inhibitor

*To whom correspondence should be addressed. For M.I.K.: phone, +91-020-2590-2241; fax, +91-020-2590-2648; E-mail, mi.khan@ncl.res.in For P.R.R.: phone, +91-020-25890-2078; fax, +91-020-2590-261; E-mail, pr.rajamohanan@ncl.res.in.

Published on Web 06/16/2010

that is effective in nanomolar concentrations. We have also demonstrated its inhibitory effect on cancer cell migration.

Results

Identification and Isolation of Compound 1. The design of the bioassay consists of screening various *Streptomyces* species fermentation broth on the basis of its ability to inhibit cysteine protease activity in a cell free system (Supporting Information (SI) Figure S1). The fermentation broth of strains that are active and exert their cysteine protease inhibitory activity were selected for identification, extraction, and isolation of specific compound. Aliquots of the dichloromethane extracts of *Streptomyces* species NCIM 2081 fermentation broths in this assay were able to inhibit the substrate hydrolysis by papain (EC 3.4.22.2). The strain, *Streptomyces* species NCIM 2081, was cultured in shake flasks in submerged conditions for 72 h, and culture supernatant was harvested, followed with lyophilization to dehydrate. The anhydrous powder thus obtained was dissolved in aqueous methanol to precipitate all proteins and extracted with *n*-hexane followed by dichloromethane (SI Scheme 1). The dichloromethane was removed under vacuum after filtration and chromatographed on semipreparative Delta pack C18 column equipped with reverse-phase HPLC resin and eluted with 40 min aqueous acetonitrile gradient (5% to 95%) containing 0.1% TFA (SI Figure S2). The fractions were analyzed in the cell free protease inhibition assay as described in the Experimental Section. The active fractions were pooled and rechromatographed by reversed-phase HPLC that showed many closely eluted compounds (SI Figure S3). Chromatography of this fraction on analytical, symmetry-C18 reversed-phase HPLC using gradient of aqueous acetonitrile containing 0.1% TFA led to separation of compounds, which were further purified by a second round of reversed-phase HPLC to yield 9 mg of compound **1** from 35 L of the original fermentation broth (SI Figure S4). ESI-MS analysis of purified compound showed a set of fragmented masses roughly in the *m*/*z* range of 737−890. The interpretations of structural information for the compound were annotated mainly by NMR spectroscopy. The two peptides with mass *m*/*z* 736.40 (compound **1a**) and *m*/*z* 842.22 (compound **1b**) have been (SI Figure S5) confirmed by complete assignments of NMR spectra such as ^{1}H, ^{13}C COSY, TOCSY, ROSEY, ^{1}H−^{13}C−HSQC, HMBC, ^{1}H−^{15}N−HSQC performed at 500/400 MHz NMR spectrometers in DMSO-d_6.

The UV spectrum of compound **1** displayed absorption maxima at λ_{max} 280 and 220 nm, which is analogous to the UV spectrum of peptides containing aromatic amino acids. The FT-IR spectra in solid KBr revealed that it contains peptide amide core (SI Figure S6) and, the biochemical analysis of amino acid composition indicated the presence of five amino acids in nearly equal molar concentrations (Table 1). The NMR spectra indicated the presence of small amount of minor component along with the major peptide. ^{1}H NMR showed that they are present in the ratio of ∼65:35, and complete structure of both the components were assigned using various 2-D NMR.

NMR Analysis and Structure Elucidation of Major Component (Compound 1a). Analysis of ^{1}H and ^{13}C NMR spectra clearly indicated the presence of two peptides with more or less identical amino acid residues (SI Figures S7A−C and S8). In addition, the ^{1}H NMR showed that they are present in the ratio of ∼65:35. The presence of six major carbonyl signals in the ^{13}C spectrum (SI Figure S8), five protons in the NH (8.3−7.8δ) and Cα (4.5−4.0δ) regions, suggested a pentapeptide structure for the major compound present. The presence of an extra carbonyl group (169.88δ) with HMBC correlation to a methyl group at 1.84δ indicated the presence of an acetyl group. A singlet at 1.23δ corresponding to nine protons having HMBC connectivity to a quaternary carbon at 42.52δ and to its own carbon at 31.23δ suggested the presence of a *t*-butyl group (SI Figure S12C). Thus modification of at least two of the amino acids present in the major compound by an acetyl and *t*-butyl group is very evident. In addition, the presence of more than one aromatic amino acid is also borne out from the ^{1}H and ^{13}C spectra. The ^{13}C DEPT spectrum had only four major CH_2 signals in the region expected for β carbons, indicating that only four out of five amino acids have β CH_2 groups. The existence of a pentapeptide is also supported by the ^{15}N HSQC data, which showed six distinct NH proton signals, five of which fall in the peptide region and the remaining most deshielded signal (−248.74 δ) corresponds to an indole NH of a tryptophan residue and the other five of them originating from the five peptide NH of the pentapeptide (SI Figure S9). The chemical shift extracted from the ^{1}H−^{15}N HSQC spectrum is given in Table 2.

Table 1. Amino Acid Composition of Compound **1**

amino acid	mol (%) ± SEM
alanine	18.07 ± 3.2
cysteine	22.1 ± 2.4
leucine	19.79 ± 1.9
phenylalanine	21.04 ± 1.5
tryptophan	20.09 ± 1.7

From the COSY (SI Figure S10A−C), TOCSY (SI Figure S10D,E), ^{1}H−^{13}C HSQC (SI Figure S11, Table S1), and ^{1}H−^{13}C HMBC (SI Figure S12A−F and Tables S2−S8), the amino acid residues of the major peptide has been unambiguously identified as alanine, cysteine, leucine, phenylalanine, and tryptophan, which correlates well with the hydrophobic nature of the peptide. The details of chemical shift and coupling constants for the major peptide are given in Table 2. HMBC correlations of the quaternary carbon of the *t*-butyl group (42.52δ) with the β CH_2 protons of cysteine (2.78 δ, 2.66δ) provided unequivocal evidence for modification of the thiol proton of cysteine by a *t*-butyl group, in other words, the presence of a cysteine-*S*-*t*-butyl moiety (SI Figure S12C). Similarly, the HMBC connectivity of acetyl methyl protons (1.84δ) to leucine α carbon at 51.68δ and leucine-NH (8.03δ) to acetyl carbonyl (169.88δ) confirmed the existence of leucine-NH-$COCH_3$ as the *N* terminal amino acid (SI Figure S12C). The sequence of the major pentapeptide was identified with the help of ^{13}C HMBC connectivity and sequential NOEs of the type H(i)-NH(i + 1), and NH(i)-NH(i + 1) (SI Figure S13A, B) as AcNH-leucine-cysteine (*S*-*t*Bu)-tryptophan-alanine-phenylalanine, which is also supported by mass spectral data (*m*/*z* = 737.71 Da; SI Figure S5). The HMBC, COSY/TOCSY, and NOESY correlations used for the structural elucidation are shown schematically in Figure 1A.

NMR Analysis and Structure Elucidation of Minor Component (Compound 1b). It is clear from the ^{1}H and ^{13}C spectra that has weaker signal (∼35 mole percent) along with some of the signal of the major peptide (SI Figure S7B). The detailed NMR investigation carried out showed that the minor peptide is also likely to be a pentapeptide containing more or less the same amino acid residues. The intensities observed for the α, β, and aromatic protons compared to the peptide NH in the ^{1}H spectrum and overall patterns obtained for the

Table 2. NMR Data for Compound **1** at 500 MHz (^{1}H), 125.75 MHz (^{13}C), 40.56 MHz (^{15}N) in DMSO-d_6[a,b]

	compound **1a** *				compound **1b** #		
Amino acid	Position	δH(int; mult; J(Hz)	δ^{13}C	δ^{15}N	δH(int; mult; J(Hz)	δ^{13}C	δ^{15}N
Leu	α	4.28 (1H, m)	51.65		4.28	51.65	
	β	1.41(2H, m)	41.27		1.41	41.27	
	γ	1.58(1H, sep, 6.72)	24.61		1.58	24.61	
	δ	0.86(d), (3H, d, 6.72)	23.48		0.86	23.48	
		0.83(d) (3H, d, 6.72)	22.08		0.83	22.08	
	C=O	-	172.74		-	172.77	
	NH	8.03(1H, d, 8.05)	-	-256.25	8.03		-256.25
Acetyl	Ac Me)	1.84 (3H, s)	22.96		1.84	22.96	
	Ac (C=O)	-	169.88		-	169.91	
Cys	α	4.28 (1H, m)	53.87		4.28	53.87	
	β	2.75(1H, dd,5.08, 12.72)	30.30		2.70 (dd,5.08,12.71)	30.24	
		2.64(1H, dd 8.64, 12.72)			2.63(dd,8.34,12.72)		
	C=O	-	170.43		-	170.21	
	NH	8.11 (1H, d, 7.79)	-	-263.68	8.16 (1H,d.7.79)	-	-263.29
t-Bu	Me	1.22 (9H, s)	31.10		1.23	31.09	
	C	-	42.42		-	42.44	
Tryptophan	α	4.51 (1H, m)	53.78		4.51	54.44	
	β	3.12(1H, dd, 4.92,14.86)	27.86		3.15	27.36	
		2.96 (1H, dd,8.10,14.86)			2.94		
	NH(Indole)	10.76 (1H,d,1.8)	-	-248.91	10.61 (1H,s)	-	-245.60
	2	7.11 (1H, d, 2.12)	124.03		--	136.99	
	3	-	110.31		-	106.55	
	3a	-	127.91		-	128.33	
	4	7.53 (1H, d, 7.88)	118.80		7.50 (1H,d,7.88)	118.80	
	5	6.93(1H, t, 7.44)	118.62		6.87 (1H,t, 7.6)	118.52	
	6	7.03 (1H, t,7.45)	121.23		6.95(1H,t, 7.5)	120.50	
	7	7.30 (1H, d, 8.1)	111.67		7.20	111.01	
	7a	-	136.53		-	135.99	
	8	-	-		3.88,3.93(2H,ABq,15.7)	31.1	
	9	-	-		-	130.38	
	10	-	-		6.98 (1H,d,8.30)	129.69	
	11	-	-		6.66 (1H,d,8.16)	115.60	
	12	-	-		-	156.06	
	13	-	-		6.66	115.60	
	14	-	-		6.98	129.69	
	C=O	-	170.89		-	170.78	
	NH	7.97 (1H, d, 8.00)		-261.38	7.97		-261.38
	OH	-			9.18 (1H,br)		
Ala	α	4.28 (1H, m)	48.63		4.30	48.61	
	β	1.15 (3H, d, 7.26)	18.69		1.15	18.83	
	C=O	-	172.20		-	172.07	
	NH	7.88 (1H, d, 7.44)		-260.05	7.80 (1H, d,7.44)		-259.52
Phe	α	4.38 (1H, m)	54.13		4.36	54.13	
	β	3.05(1H, dd, 5.38,3.82)	37.17		3.03	37.17	
		2.93(1H,dd,8.28,13.82)			2.90		
	1	-	138.05		-	138.05	
	2	7.23	129.63		7.23	129.63	
	3	7.25	128.57		7.25	128.57	
	4	7.18 } 5H	126.80		7.23 } 5H	126.80	
	5	7.25	128.57		7.25	128.57	
	6	7.22	129.63		7.22	129.63	
	NH	8.00 (1H, d, 8.17)		-264.38	8.00	-	-264.38
	COOH	12.72(1H, br)	173.14		12.72	173.14	

[a] *The intensities are calculated with respect to indole NH at 10.75δ. The notations s, d, t, q, m, dd, br in the table stands for singlet, doublet, triplet, multiplet, doublet of doublet, and broad signal, respectively. [b] #The intensities are calculated with respect to indole NH at 10.60δ. Coupling constants are shown only for signals with minimum overlap with the major signals. For some of the protons, the chemical shifts are extracted from the ^{1}H−^{13}C HSQC/HMBC spectrum.

^{1}H and ^{13}C spectra can be rationalized only by considering a pentapeptide having identical residues with modification in only one amino acid residues of the major pentapeptide (compound **1a**).

The multiplicities of the tryptophan residue signals observed in the ^{1}H spectrum suggest the presence of a modified tryptophan residue in the minor peptide (~35 mol %). A splitting of ~2 Hz for the major indole NH at 10.76δ, due to weak scalar coupling with C2 proton of the tryptophan ring at 7.15δ, is clearly visible in the expanded ^{1}H spectrum (SI Figure S7B) while the weak NH signal at 10.6δ appear as a singlet. This clearly suggests the absence of proton on tryptophan C2 carbon in the minor pentapeptide (compound **1b**). The absence of *J* coupling of the minor NH proton (10.6δ) of the indole ring in the COSY and TOCSY spectra was also noticeable (SI Figure S10B and S10D). The most interesting observation is the presence of a tyrosine type of moiety (doublets at 6.98δ, 6.65δ, broad peak at 9.2δ), which also corresponds to the minor signals in the ^{1}H spectrum (SI Figure S7A and B) ^{13}C NMR also confirms the presence of a tyrosine

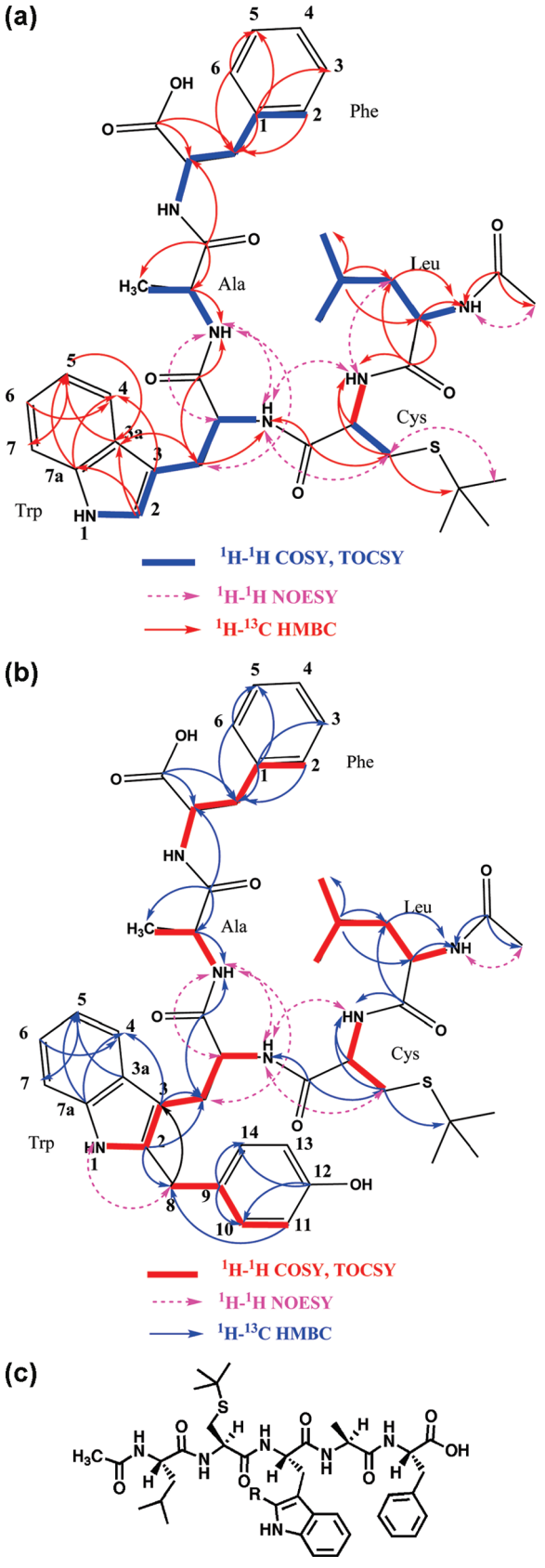

Figure 1. (A) Selected HMBC, COSY, TOCSY, and NOESY correlations of compound-**1a** (Major). (B) Selected HMBC, TOCSY, COSY, and NOESY correlations of compound-**1b** (Minor). (C) Final chemical structure, elucidated with the help of UV/Vis, FT-IR and NMR spectroscopy of cysteine protease inhibitor (CPI-2081), compound **1** (R = H or $CH_2-C_6H_4-OH$).

type of moiety in the compound **1b** (quaternary carbons at 156.6δ and 130.6δ and aromatic CH carbons at 129.7δ and 115.6δ). Apart from this, an additional quartet (J = 15.7 Hz), corresponding to the minor set of signals, is seen at 3.90δ. The $^{13}C-^1H$ HSQC spectrum correlates this quartet to a carbon at ~31.1 ppm, where the ^{13}C signal from the **t**-butyl group also incidentally appears (SI Figure S11C). These protons were found to show weak J couplings in the COSY and TOCSY spectra to the protons on $-C_6H_4$-OH moiety (6.98δ, 6.65δ), COSY and TOCSY expanded (SI Figure S14A). The AB quartet pattern observed with a large coupling constant (15.7 Hz) is indicative of a methylene group. The absence of any $^1H-^1H$ J couplings of it to other protons rules out the possibility of a tyrosine residue, as it should have shown further J couplings to its α proton.

All these observations can be rationalized by considering the presence of only a tyrosine type side chain ($-CH_2-C_6H_4-$OH) group (Figure 1B). The presence of such a group has been confirmed from the $^{13}C-^1H$ HMBC spectrum where connectivities of this CH_2 group to the carbons of the phenolic residue (CH carbon at 129.60δ and quaternary carbon at 130.36δ) are seen (SI Figure S14B). More information regarding the linking of the other side of this CH_2 group is obtained from the $^{13}C-^1H$ HMBC spectrum. This quartet (centered at 3.98δ) shows correlation to a weak signal at 106.6 ppm, a quaternary carbon that in turn correlates to the α-proton of tryptophan residue. (SI Figure S14B). This also suggests that the α-proton of both major and minor signal has nearly same chemical shift. The major α proton correlates to the tryptophan C3 carbon at 110.4δ. The correlation of the weak signal at 106.6 δ to the weaker CH_2 protons (3.14, 2.95δ) appears very close to the β protons of tryptophan residue. This in turn confirms that the tryptophan moiety is intact and the modification has taken place at C2 position by substitution of the indole ring proton by a $CH_2-C_6H_4-OH$ group (Figure 1B). The C2 carbon of the modified tryptophan residue appears at 136.93δ as a quaternary carbon, which not only shows connectivity to the CH_2 group at 3.98δ but also to the weak β protons (3.14, 2.95δ) appearing close to the β protons of the unmodified tryptophan residue (SI Figure S14D).

The HMBC connectivity of the NH group of indole ring of tryptophan also shows unambiguous evidence for this modification (SI Figure S14C). Both the major and minor NH protons show correlation to four different carbons viz., C2, C3, C3a, C7, and C7a. The changes in ^{13}C chemical shifts for these carbons are very much evident from this. Because the modification is only on one of the amino acids of the pentapeptide, the ^1H and ^{13}C chemical shifts observed for the minor component (compound **1b**) is very close to that of the major signals (compound **1a**) except for the signals arising form the tryptophan residue. The ^1H and ^{13}C chemical shifts of the minor pentapeptide (compound **1b**) are given in Table 2, and the connectivities observed are shown in Figure1B. The presence of two very similar pentapeptides have also been confirmed by 2D DOSY experiment performed on a sample dissolved in methanol-d_4, which clearly indicated the presence of two compounds having very close but distinct diffusion coefficients (SI Figure S15).

Thermodynamic Study of Proteolytic Inhibition Activity of Compound 1. Because the compound **1** was isolated from culture supernatant of *Streptomyces* sp. NCIM2081 following activity guided fractionation, and the enzyme used for the purpose was papain, a comprehensively studied cysteine

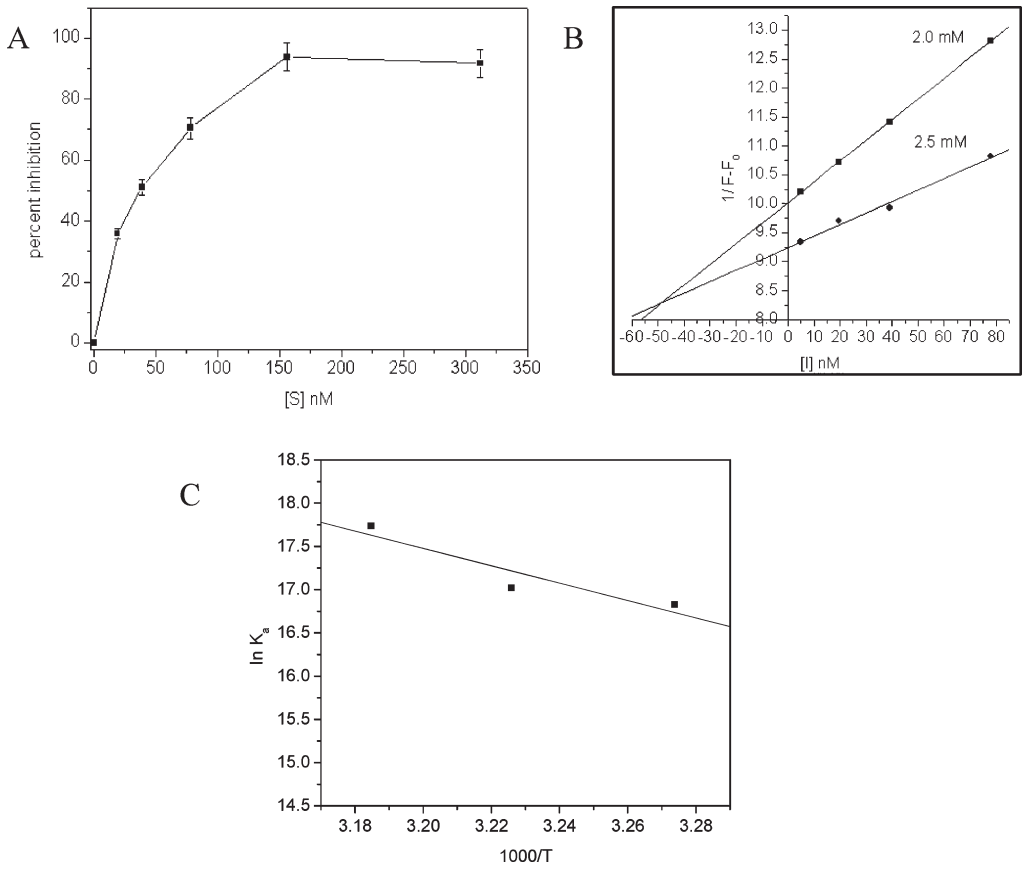

Figure 2. Determination of IC_{50}, K_i for inhibition of hydrolytic activity of papain and thermodynamic parameters of compound **1** and papain binding. (A) Determination of IC_{50} (36.9 ± 1.8 nM) for inhibition of hydrolysis of substrate by papain. Percent inhibition ± SEM is plotted against concentration of compound **1** in nanomolar (nM). (B) Dixon plot for determination of inhibitory constant (K_i = 49.14 ± 2.45 nM) for inhibition of proteolysis by compound **1**. Reciprocal rate of reaction is plotted against concentration of compound **1** in nanomolar (nM). (C) Vant Hoff plot for association of compound **1** with papain (r = 0.9336 and N = 3).

protease enzyme. Therefore, we decided to examine the thermodynamic feasibility for the inhibition of papain's hydrolytic activity. Also, we determined the inhibitory constant (K_i), which is an important parameter that shows the potency of inhibitory compounds. We analyzed the initial kinetics of substrate hydrolysis inhibition by compound **1**. Substrate hydrolysis decreased as a function of concentration of compound **1** in a dose dependent manner displaying the IC_{50} value of 36.9 ± 1.8 nM (Figure 2A). The equilibrium constants for association (K_i) was calculated with the help of a Dixon plot using ORIGIN 6.1 software, where inverse rate of synthetic substrate, BAPNA (*N*-benzoyl-DL-arginine-*p*-nitroanilide hydrochloride) hydrolysis was plotted against a range of concentration of compound **1** (from 4.87 to 78 nM) at 2 and 2.5 mM BAPNA. Dixon plot demonstrated the competitive mode of association of compound **1** with papain, showing K_i value of 49.14 ± 2.45 nM (Figure 2B), which is close to IC_{50} value. Thermodynamic analysis of interaction between compound **1**, and papain were studied as described in literature.[25] In brief, the effect of temperature on K_i was investigated and thermodynamic parameters for the association of compound **1** with papain was calculated using the Vant Hoff plot (Figure 2C and Table 3).

Compound 1 Inhibits Cancer Cell Migration. Cancer cells spread from the primary tumor either as individual cells, using amoeboid, or as cell sheets, strands, and clusters by means of collective migration. Cysteine protease antagonists such as leupeptin and antinpain have been demonstrated to retard the cell translocation during wound repair process.[26] We investigated the ability of compound **1** to inhibit cancer cell migration. To examine this, scratch wound healing assays were performed in various cancer cell lines as described by Liang et al.[27] Briefly, the motility inhibition of tumor cells was determined by capability of compound **1** to inhibit the wound closure over the time in monolayer confluent cells.[28] Postconfluent MDA-MB-231 (human breast carcinoma), B16F10 (murine melanoma), or A-375 (human melanoma) cells with the typical cobblestone morphology were used for this experiment. Compound **1** (25 μM) significantly inhibited the wound closure in A-375 and MDA-MB-231 cells as compared to vehicle control (Figure 3A−C). Similarly, 25 μM compound **1** was enough to inhibit the wound closure in B16F10 by 20% (Figure 3A,D). MTT cell viability assay revealed that the wound closure inhibition was not due to cytotoxic effect of compound **1** (Figure 4). A-375 cells were approximately 100% viable as compared to control even in 50 μM compound **1** in culture conditions, which is much lower than the compound **1** concentration needed for 30% inhibition

Table 3. Thermodynamic Parameters for Binding of Compound **1** with Papain

temp (°C)	k_i (nM)	ΔG (kJ/mol)	ΔH^a (kJ/mol)	ΔS J/mol/K
32	−49.15	−42.67	83.84	419.36
37	−40.58	−43.86	83.84	411.61
41	−19.8	−46.3	83.84	411.43

[a] All values are calculated at 37 °C.

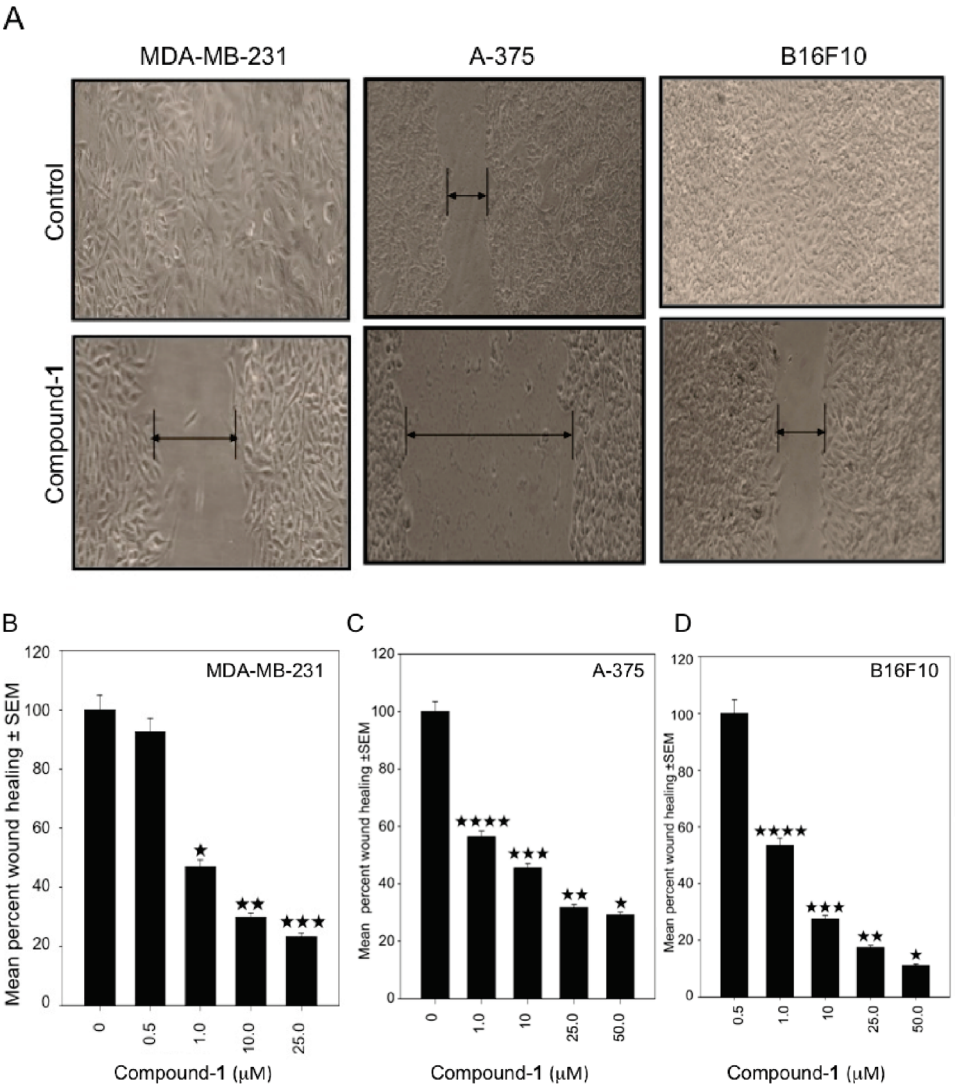

Figure 3. Compound **1** inhibits cancer cell migration. (A) Representative bright phase images showing the wound closure after 24 h upon compound **1** (50 μg/mL) treatment in MDA-MB-231, A-375, and B16F10 cells as compared to control. Mean percent wound migration (standard error of mean) is plotted as a function of concentration of compound **1** as indicated. (B) MDA-MB-231 human breast carcinoma cells, statistical analysis, *P* values for significantly different means, $P^* = 0.010$, $P^{**} = 0.005$, $P^{***} = 0.004$ vs control (0 μg/mL compound **1**). (C) A375 human melanoma cells, statistical analysis, *P* values for significantly different means, $P^{****} = 0.008$, $P^{***} = 0.004$, $P^{**} = 0.002$, $p^* = 0.002$ vs control (0.5 μg/mL compound **1**). (D) B16F10 murine melanoma cells, statistical analysis, *P* values for significantly different means, $P^{****} = 0.013$, $P^{***} = 0.004$, $P^{**} = 0.003$, $p^* = 0.002$ vs 0.5 μg/mL.

of cell migration (Figure 4). Similarly, compound **1** was able to inhibit the cell migration of MDA-MB-231 and B16F10 cells at the concentrations lower than the in vitro cytotoxic dose.

Discussion

Compounds bearing a *t*-butyl group are very rare in nature and are mostly isolated from marine sponges and include peptides, terpenes, carbinols, esters, and ketone.[29−32] A very unusual compound containing tertiary butyl group, *t*-bu ketone coumarin swietenone, was isolated from *Chloroxylon swietenia*.[33] Here, we describe the isolation and structure elucidation of novel tertiary butyl containing cysteine protease inhibitor (compound **1**), a mixture of two very similar but distinctly modified peptides in approximate molar proportion of 35:65. As evidenced both by biochemical analysis (Table 1) and NMR analysis (Table 2, Figure 1A, Figure 1B). Four out of five amino acids in compound **1** (compound **1a** and compound **1b**) are common, therefore, exhibiting very similar physiochemical properties, making it nearly impossible to isolate individual components of compound **1** with the help of reverse phase HPLC equipped with C18 chemistry using various solvent conditions. The complete structure elucidated with the help of UV−vis, FT-IR, MS-MS, and extensive NMR spectroscopic techniques has shown that compound **1** contains unique groups and modifications such as acetylation of N-terminal leucine and tertiary butylation of cysteine residue. N-Terminal acetylation and *t*-butylation of bioactive peptide confers the extra protection from endogenous peptidases. The thermodynamic analysis shows that the K_i kept on increasing with respect to the increase in temperature between 32 and 41 °C (Table 3), and the Vant Hoff plot was a straight line having a negative slope (Figure 2C). As shown in Table 3, the Gibbs free energy (ΔG) is negative, indicating that the interaction between compound **1** and papain is spontaneous while positive ΔH suggests that the overall reaction is endothermic. Also, significantly high positive entropy (ΔS) shows that the

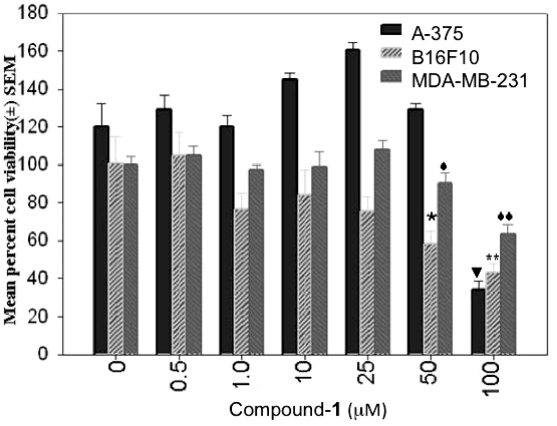

Figure 4. MTT cell viability assay upon compound **1** treatment: MDA-MB-231 and B16F10 cells were treated with increasing concentration of compound **1** for 72 h and A-375 cells were treated for 48 h. Cells were incubated with MTT for 4 h and an ELISA reader measured absorbance at 570 nm. Mean percent cell viability (±standard error of mean) is plotted as a function of compound **1** concentration in μM. Statistical analysis, P values for significantly different means, $P^{\blacktriangledown} = 0.003$, $P^{*} = 0.012$, $P^{**} = 0.017$, $P^{\blacklozenge\blacklozenge} = 0.0009$ vs control.

inhibition of papain by compound **1** is due to hydrophobic interaction and association is driven by increase in randomness. All these kinetic and thermodynamic data suggest that the newly isolated compound **1** could be a potential scaffold to study and can be used to develop new potent drug candidates for the diseases associated with overactive and/or deregulated expression of cysteine proteases.

Because tumor metastasis is observed in almost all kind of cancer, it is a leading problem in cancer prevention strategies. Because of the diverse nature of tumor cells with respect to type, grade, localization, and close similarity to cellular composition of normal tissue, it is very difficult to have anticancer agent without having any side affect on normal tissues. Therefore, to achieve optimal chemotherapy with anticancer drugs, it very important to develop antimetastatic drugs which could arrest the tumor growth and dispersal. We are reporting isolation and characterization of a small molecule cysteine protease inhibitor, compound **1** that has a potential to be developed as an antimetastatic drug.

Experimental Section

Bacterial Strain Culture and Fermentation. *Streptomyces* sp. NCIM2081 cultures were seeded in 1000 mL flask containing MGYP culture media (malt extract, glucose, yeast extract, and peptone) containing 2% soybean meal and incubated at 28 °C with shaking at 200 rpm. After 72 h, three flasks were removed and kept at 4 °C for harvesting culture supernatant, followed by protease inhibition activity assay using papain as the model enzyme.

Inhibitory Activity for Compound 1 against Papain. The inhibitory activity of compound **1** against papain was determined by assaying the proteolytic activity of 1.28 μM of papain in Tris-HCL buffer, pH 6.5, in the presence of 10 mM DDT (dithiothreitol) and 2 mM EDTA (ethelene diamine tetra acetic acid) using 1.5 mM BAPNA (*N*-benzoyl-DL-arginine-*p*-nitroanilide hydrochloride) as the substrate in the presence and absence of compound **1** at 37 °C. To determine the IC_{50} value and K_i, the increasing concentration of compound **1** (ranging from 19.4 to 312 nM) was preincubated with papain and 1.5 mM BAPNA was added to initiate the reaction. The percent substrate hydrolysis was calculated using an end point assay in which the release of *p*-nitroanilide from BAPNA on account of substrate hydrolysis was monitored by recording the absorption maxima at 405 nm.

Amino Acid Analysis Using Fluorescent HPLC. A total of 2 mg of compound **1** was digested in 6N HCl at 110 °C for 24 h in vacuum-sealed tubes. The hydrolysate was then used for amino acid analysis with an AccQ-Fluor kit (Waters Corporation, USA) following the manufacturer's instruction. The digested sample was derivatized with 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (AQCC) following manufacturer's instruction. In brief, five picomole of the sample was loaded onto AccQ-tag HPLC column and eluted with acetonitrile gradient (5−95%). Elute was monitored with a florescence detector (Waters Corporation, USA). To calculate the molar proportion of constituents, the peak areas of individual amino acids were compared with standards running under identical conditions. Total cysteine and tryptophan were estimated with intact peptide according to the method of Cavallini et al.[34] and Spande and Witkop,[35] respectively.

NMR Measurements. All the ^1H and ^{13}C NMR measurements were carried out on Bruker AV 500 spectrometer operating at 500.13 and 125.75 MHz, respectively. Ten mg of the sample isolated was dissolved in DMSO-d_6 in a standard 5 mm NMR tube and the ^1H, COSY, NOESY, ROESY, ^{13}C CPD, ^{13}C DEPT, $^{13}C-^1H$ HSQC, $^{13}C-^1H$ HMBC, and $^{15}N-^1H$ HSQC experiments were performed. The one-dimensional ^{13}C experiments were performed on a 5 mm quadra nuclei probe at ambient temperature (~28 °C). Twenty thousand and 11000 transients were collected for ^{13}C CPD and DEPT135 respectively. All the 2D experiments except the $^{15}N-^1H$ HSQC were conducted on a 5 mm broad band inverse gradient probe. The $^{15}N-^1H$ HSQC experiment was carried out on Bruker AV 400 NMR spectrometer operating at 40. 56 MHz for ^{15}N using a broad band observe (BBO) gradient probe. Gradient spectroscopic techniques were employed for all the 2D experiments. Four hundred experiments (t_1 increments) of 24 scans were used for COSY, NOESY, and ROESY measurements. The COSY and the HMBC spectra were collected in a magnitude mode, while a phase sensitive (States-TPPI) mode was used for HSQC, NOSEY, TOCSY, and ROESY measurements. A mixing time of 1 s and 300 m s was employed for NOESY and ROESY experiments, respectively. The numbers of scans used for each t_1 increment for other 2D experiment were as follows: 24 (^{13}C HSQC,), 80 (^{13}C HMBC), 64 (^{15}N HSQC). The ^{13}C HMBC data was optimized for a long-range coupling constant of 6 Hz. A pulse sequence employing a double low pass filter gave better results for ^{13}C HMBC due to spread in $^1J_{C-H}$ values (160−135 Hz). The HMBC spectra were acquired without proton decoupling during detection. The 90° pulse lengths for ^1H, ^{13}C, and ^{15}N were 13.5, 10, and 14 per s, respectively. Appropriate window functions viz. sine squared bell with no phase shift for all magnitude mode and phase shifted (ssb = 2) sine squared bell for phase sensitive mode were used for data processing. In general a 2K × 2K data matrix size was used for the 2D experiments. The ^1H and 13C chemical shifts were referred to the residual solvent peak (2.50 and 39.95 the central signal of the solvent, respectively for ^1H and ^{13}C). The ^{15}N chemical shifts were referred to an external sample of nitro methane (0 ppm). DOSY experiments were carried out on Burker AV 500 NMR spectrometer equipped with a 5 mm broad band inverse gradient probe by systematic variation of the strength of the gradient amplitude and data was processed using standard Bruker software.

Mammalian Cell Culture. Human breast adenocarcinoma MDA-MB-231 was cultured in L-15 media supplemented with 10% fetal calf serum, 100 units/mL penicillin, 100 μg/mL streptomycin, and 2 mM glutamine in a humidified atmosphere of air containing 5% CO_2 at 37 °C. Human melanoma A-375 and murine melanoma B16F10 was cultured in DMEM supplemented with fetal calf serum 10%, penicilin100 units/mL, streptomycin 100 μg/mL, and 2 mM glutamine in a humidified atmosphere containing 5% CO_2 and 95% air at 37 °C.

Cell Survival Assay. MDA-MB-231, B16F10, or A-375 cells were plated in 96-well plates. After 24 h, cells were treated with compound **1** in indicated concentrations and time. After treatment, media was removed and 0.5 mg/mL MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide] solution was added to the cells and A_{570} was recorded after 4 h of incubation.

Wound Healing Assay. The migration of tumor cells was determined by scratch wound healing assay. Postconfluent cells showing the typical cobblestone morphology were used for this experiment. Scratch wounds with a constant diameter were made with a 10 μL micropipet tip on monolayer and cells were maintained in media supplemented with 5% FBS. Then cells were treated with compound **1** in indicated concentration. Cells were incubated at 37 °C for 24 h. After termination of experiments, wound photographs were taken under phase contrast microscope (Nikon). The photographs were analyzed for cell wound healing over the time by measuring the horizontal grid length of wound (scratch) in Microsoft PowerPoint.

Acknowledgment. We thank Dr. G. C. Mishra for encouragement, Dr. S. S. Deshmukh for technical assistance in microbial work, and Drs. Mahesh Kulakarni and K. V. S. Rao for mass spectroscopic analysis. J. P. S. is a senior research fellow of CSIR, Government of India. The financial assistance was provided by the Department of Biotechnology, Government of India, to M. I. K.

Supporting Information Available: Experimental procedure, production and purification of compound, mass spectroscopy, detailed 1D-NMR ^{15}N NMR, 2D-NMR, DOSY are provided as supplementary information. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Lah, T. T.; Kos, J. Cysteine proteinases in cancer progression and their clinical relevance for prognosis. *Biol. Chem.* **1998**, *379*, 125–130.
(2) Premzl, A.; Puizdar, V.; Zavasnik-Bergant, V.; Kopitar-Jerala, N.; Lah, T. T.; Katunuma, N.; Sloane, B. F.; Turk, V.; Kos, J. Invasion of ras-transformed breast epithelial cells depends on the proteolytic activity of cysteine and aspartic proteinases. *Biol. Chem.* **2001**, *382*, 853–857.
(3) Sameni, M.; Moin, K.; Sloane, B. F. Imaging proteolysis by living human breast cancer cells. *Neoplasia* **2000**, *2*, 496–504.
(4) Wang, B.; Sun, J.; Kitamoto, S.; Yang, M.; Grubb, A.; Chapman, H. A.; Kalluri, R.; Shi, G. P. Cathepsin S controls angiogenesis and tumor growth via matrix-derived angiogenic factors. *J. Biol. Chem.* **2006**, *281*, 6020–6929.
(5) Heidtmann, H. H.; Salge, U.; Abrahamson, M.; Bencina, M.; Kastelic, L.; Kopitar-Jerala, N.; Turk, V.; Lah, T. T. Cathepsin B and cysteine proteinase inhibitors in human lung cancer cell lines. *Clin. Exp. Metastasis* **1997**, *15*, 368–381.
(6) Sloane, B. F.; Moin, K.; Krepela, E.; Rozhin, J. Cathepsin B and its endogenous inhibitors: the role in tumor malignancy. *Cancer Metastasis Rev.* **1990**, *9*, 333–352.
(7) Desai, S. N.; White, D. M.; O'Shea, K, M.; Brown, M. L.; Cywin, C. L.; Spero, D. M.; Panzenbeck, M. J. An orally active reversible inhibitor of cathepsin S inhibits human trans vivo delayed-type hypersensitivity. *Eur. J. Pharmacol.* **2006**, *538*, 168–174.
(8) Wang, D.; Li, W.; Pechar, M.; Kopeckova, P.; Bromme, D.; Kopecek, J. Cathepsin K inhibitor−polymer conjugates: potential drugs for the treatment of osteoporosis and rheumatoid arthritis. *Int. J. Pharm.* **2004**, *277*, 73–79.
(9) Weidauer, E.; Yasuda, Y.; Biswal, B. K.; Cherny, M.; James, M. N.; Bromme, D. Effects of disease-modifying anti-rheumatic drugs (DMARDs) on the activities of rheumatoid arthritis-associated cathepsins K and S. *Biol. Chem.* **2007**, *388*, 331–336.
(10) Costales, J.; Rowland, E. C. A role for protease activity and host-cell permeability during the process of *Trypanosoma cruzi* egress from infected cells. *J. Parasitol.* **2007**, *93*, 1350–1359.
(11) Hernandez, A. A.; Roush, W. R. Recent advances in the synthesis, design and selection of cysteine protease inhibitors. *Curr. Opin. Chem. Biol.* **2002**, *6*, 459–465.
(12) Pandey, K. C.; Singh, N.; Arastu-Kapur, S.; Bogyo, M.; Rosenthal, P. J. Falstatin, a cysteine protease inhibitor of Plasmodium falciparum, facilitates erythrocyte invasion. *PLoS Pathog.* **2006**, *2*, 1031–1041.
(13) Sajid, M.; McKerrow, J. H. Cysteine proteases of parasitic organisms. *Mol. Biochem. Parasitol.* **2002**, *120*, 1–21.
(14) Shenai, B. R.; Sijwali, P. S.; Singh, A.; Rosenthal, P. J. Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. *J. Biol. Chem.* **2000**, *275*, 29000–29010.
(15) Sijwali, P. S.; Shenai, B. R.; Gut, J.; Singh, A.; Rosenthal, P. J. Expression and characterization of the *Plasmodium falciparum* haemoglobinase falcipain-3. *Biochem. J.* **2001**, *360*, 481–489.
(16) Sakano, Y.; Shibuya, M.; Matsumoto, A.; Takahashi, Y.; Tomoda, H.; Omura, S.; Ebizuka, Y. Lanopylins A1, B1, A2, and B2, novel lanosterol synthase inhibitors from *Streptomyces* sp. K99−5041. *J. Antibiot.* (*Tokyo*) **2003**, *56*, 817–826.
(17) Tani, M.; Gyobu, Y.; Sasaki, T.; Takenouchi, O.; Kawamura, T.; Kamimura, T.; Harada, T. SF2809 compounds, novel chymase inhibitors from *Dactylosporangium* sp. 1. Taxonomy, fermentation, isolation and biological properties. *J. Antibiot.* (*Tokyo*) **2004**, *57*, 83–88.
(18) Tani, M.; Harimaya, K.; Gyobu, Y.; Sasaki, T.; Takenouchi, O.; Kawamura, T.; Kamimura, T.; Harada, T. SF2809 compounds, novel chymase inhibitors from *Dactylosporangium* sp. 2. Structural elucidation. *J. Antibiot.* (*Tokyo*) **2004**, *57*, 89–96.
(19) Singh, M. P.; Petersen, P. J.; Weiss, W. J.; Janso, J. E.; Luckman, S. W.; Lenoy, E. B.; Bradford, P. A.; Testa, R. T.; Greenstein, M. Mannopeptimycins, new cyclic glycopeptide antibiotics produced by *Streptomyces hygroscopicus* LL-AC98: antibacterial and mechanistic activities. *Antimicrob. Agents Chemother.* **2003**, *47*, 62–69.
(20) Aoyagi, T.; Takeuchi, T.; Matsuzaki, A.; Kawamura, K.; Kondo, S. Leupeptins, new protease inhibitors from Actinomycetes. *J. Antibiot.* (*Tokyo*) **1969**, *22*, 283–286.
(21) Suda, H.; Aoyagi, T.; Hamada, M.; Takeuchi, T.; Umezawa, H. Antipain, a new protease inhibitor isolated from actinomycetes. *J. Antibiot.* (*Tokyo*) **1972**, *25*, 263–266.
(22) Kadowaki, T.; Kitano, S.; Baba, A.; Takii, R.; Hashimoto, M.; Katunuma, N.; Yamamoto, K. Isolation and characterization of a novel and potent inhibitor of Arg-gingipain from *Streptomyces* sp. strain FA-70. *Biol. Chem.* **2003**, *384*, 911–920.
(23) Nakamura, H.; Takahashi, Y.; Naganawa, H.; Nakae, K.; Imoto, M.; Shiro, M.; Matsumura, K.; Watanabe, H.; Kitahara, T. Absolute configuration of migrastatin, a novel 14-membered ring macrolide. *J. Antibiot.* (*Tokyo*) **2002**, *55*, 442–444.
(24) Hanada, K.; T., M.; Yamagishi, M. Isolation and charrecterization of E64, a new cysteine protease inhibitor. *Agric. Biol. Chem.* **1978**, *42*, 523–528.
(25) Hymes, A. J.; Robinson, D. A.; Canady, W. J. Thermodynamics Of The Solution Process. II. The Use Of An Extraction Model For Enzyme−Inhibitor Complex Formation. *J. Biol. Chem.* **1965**, *240*, 134–138.
(26) Gordon, S. R.; DeMoss, J. Exposure to lysosomotropic amines and protease inhibitors retard corneal endothelial cell migration along the natural basement membrane during wound repair. *Exp. Cell Res.* **1999**, *246*, 233–242.
(27) Liang, C.-C.; P., A. Y.; Guan., J.-L. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. *Nature Protoc.* **2007**, *2*, 329–333.
(28) Li, W.; Henry, G.; Fan, J.; Bandyopadhyay, B.; Pang, K.; Garner, W.; Chen, M.; Woodley, D. T. Signals that initiate, augment, and provide directionality for human keratinocyte motility. *J. Invest. Dermatol.* **2004**, *123*, 622–633.
(29) Chevallier, C.; Richardson, A. D.; Edler, M. C.; Hamel, E.; Harper, M. K.; Ireland, C. M. A new cytotoxic and tubulin-interactive milnamide derivative from a marine sponge *Cymbastela* sp. *Org. Lett.* **2003**, *5*, 3737–3739.
(30) Geonseek Ryu, S. M.; Nobuhiro, F.. Discodermin E, a cytotoxic and antimicrobial tetradecapeptide, from the marine sponge *Discodermia kiiensis*. *Tetrahedron Lett.* **1994**, *35*, 8251–8254.
(31) Hamel, E. Interactions of antimitotic peptides and depsipeptides with tubulin. *Biopolymers* **2002**, *66*, 142–160.
(32) Matsunaga, S.; F., N.; Konosu, S. Bioactive marine metabolites VII. Structures of discodermins B, C, and D, antimicrobial peptides from the marine sponge *Discodermia kiiensis*. *Tetrahedron Lett.* **1985**, *26*, 855–856.
(33) Mujumdar, R. B.; R., A. V. R.; Rathi, S. S.; Venkataraman, K. Swietenone, the first natural-butyl ketone, from *Chloroxylon swietenia*. *Tetrahedron Lett.* **1975**, *16*, 867–868.
(34) Cavallini, D.; Graziani, M. T.; Dupre, S. Determination of disulphide groups in proteins. *Nature* **1966**, *212*, 294–298.
(35) Spande, T. F.; Witkop, B. Determination of the tryptophan content of proteins with *N*-bromosuccinimide. *Methods Enzymol.* **1967**, *11*, 498–505.