Supporting Information

Affinity Crystallography: A New Approach To Extracting High Affinity Enzyme Inhibitors From Natural Extracts

Adeleke H. Aguda [†,‡], Vincent Lavallee [‡], Ping Cheng [§], Tina M. Bott [§], Labros G. Meimetis [§], Simon Law [‡], Nham T. Nguyen [‡], David E. Williams [§], Jadwiga Kaleta [†], Ivan Villanueva [⊥], Julian Davies [⊥], Raymond J. Andersen [§], Gary D. Brayer [‡] & Dieter Brömme [†,‡,▽*]

[†] *Department of Oral Biological and Medical Sciences, Faculty of Dentistry, University of British Columbia, Vancouver, B. C., CANADA, V6T 1Z3*

[‡] *Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of British Columbia, Vancouver, B. C., CANADA, V6T 1Z3*

[§] *Department of Chemistry and Earth, Ocean & Atmospheric Sciences, Faculty of Science, University of British Columbia, Vancouver, B. C., CANADA, V6T 1Z1*

[⊥] *Department of Microbiology, Faculty of Science, University of British Columbia, Vancouver,* B. C., CANADA, V6T 1Z3

[▽] *Center for Blood Research, University of British Columbia, Vancouver, B. C., CANADA, V6T 1Z3*

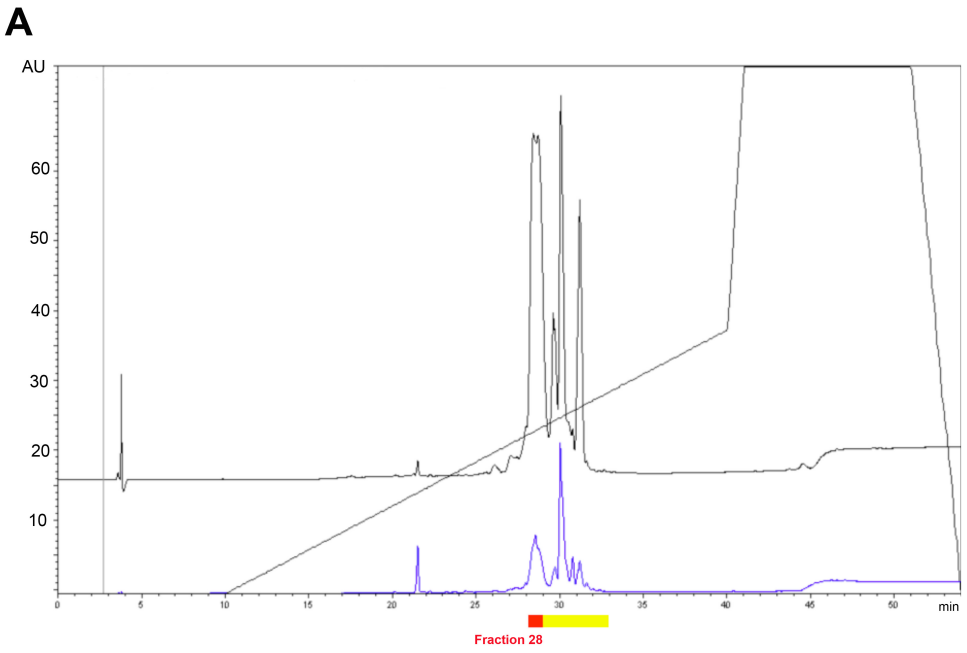
Table of contents

Page SI-3: Figure S1. Purification of CatK inhibitors from Actinomycetes strain.
Page SI-4: Structural Elucidation of Purified CatK Inhibitors V4 (**1**) and V2 (**2**).
Pages SI-5 to SI-6: Table S1. NMR Data for V2 (**2**) and ^{13}C NMR Assignment Comparison with V4 (**1**) Recorded in DMSO d_6.
Page SI-7: ^{1}H NMR Spectrum of V2 (**2**) Recorded at 600 MHz in DMSO-d_6.
Page SI-8: Expanded ^{1}H NMR Spectrum of V2 (**2**) with Peak Picking Recorded at 600 MHz in DMSO-d_6.
Page SI-9: Expanded ^{1}H NMR Spectrum of V2 (**2**) with Peak Picking Recorded at 600 MHz in DMSO-d_6.
Page SI-10: Expanded ^{1}H NMR Spectrum of V2 (**2**) with Peak Picking Recorded at 600 MHz in DMSO-d_6.
Page SI-11: ^{13}C NMR Spectrum of V2 (**2**) Recorded at 150 MHz in DMSO-d_6.
Page SI-12: Figure S2: K_i Value Determination of V2 (**1**), V4 (**2**), Antipain and Lichostatinal (**3**).
Page SI-13: Molecular Docking Methodology for Figure S3.
Page SI-13: Figure S3. Molecular Docking of V2 (**2**) and V4 (**1**) into the Active Site of CatK.
Page SI-14: Table S3. Crystallographic Data Statistics for CatK-Inhibitor Complexes.
Page SI-15: Table S3. Amino Acid Residues within 3.6 Å of Lichostatinal in Sample 2 Crystal.
Page SI-16: Table S4. NMR Data for Lichostatinal (**3**) Recorded in DMSO d_6.
Page SI-17: ^{1}H NMR NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6.

Page SI-18: Expanded ^{1}H NMR Spectrum of Lichostatinal (**3**) with Peak Picking Recorded at 600 MHz in DMSO-d_6.
Page SI-19: Expanded ^{1}H NMR Spectrum of Lichostatinal (**3**) with Peak Picking Recorded at 600 MHz in DMSO-d_6.
Page SI-20: Expanded ^{1}H NMR Spectrum of Lichostatinal (**3**) with Peak Picking Recorded at 600 MHz in DMSO-d_6.
Page SI-21: ^{13}C NMR NMR Spectrum of Lichostatinal (**3**) Recorded at 150 MHz in DMSO-d_6.
Page SI-22: gradCOSY60 NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6.
Page SI-23: gradHSQC NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6.
Page SI-24: gradHMBC NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6.
Page SI-25: References.

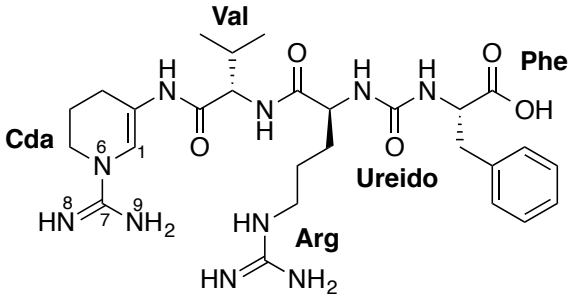
Figure S1. Purification of CatK Inhibitors from Actinomycetes Strain.

(A) Following the LH-20 size exclusion, the L-91-3 sample was run on semi-preparative HPLC at an isocratic elution of H_2O with 0.1% TFA for 10 min followed by a 30 min gradient to 50% MeCN with 0.1% TFA. Absorbance was measured at two wavelengths, 220 nm (blue) and 260 nm (black). Active fractions against CatK were found between 28-33 minutes. Fraction #28 (red) was isolated and further purified using RP C18 preparative HPLC. **(B)** Separation of 4 active peaks (V1-V4) after preparative HPLC of fraction 28.

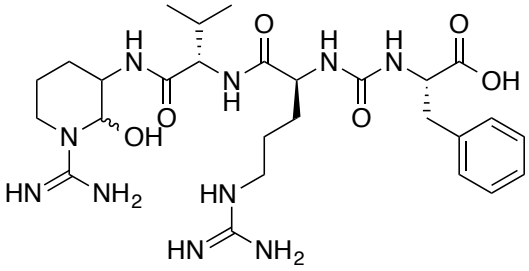


Structural Elucidation of Purified CatK Inhibitors V4 (1) and V2 (2). V4 (**1**) was readily identified as the cycloarginal (the hemiaminal) tautomer of antipain with an $[M+H]^+$ ion in the HRESIMS at *m/z* 605.3511 appropriate for a molecular formula of $C_{27}H_{44}N_{10}O_6$ and NMR data that allowed for an unambiguous assignment of the structure upon comparison with the literature[1-4]. V2 (**2**) gave an $[M+Na]^+$ ion in the HRESIMS at *m/z* 587.3403 appropriate for a molecular formula of $C_{27}H_{42}N_{10}O_5$ that differs from that of V4 (**1**) by the loss of H_2O, and requiring one additional site of unsaturation. The ^1H and ^{13}C NMR spectra of V2 (**2**) (Table S1) were markedly similar to those of V4 (**1**) but with one significant structural and functional difference. The difference was to be found in the terminal cycloarginal residue of antipain (**1**) because resonances appropriate for a trisubstituted double bond (δ 122.7 (C-2), 113.2 (C-1), 7.34 (H-1)) had to be incorporated. Dehydration across the C-1/C-2 bond would account for the loss of H_2O and the additional site of unsaturation. Examination of the 1D and 2D NMR data obtained for V2 (**2**) revealed gHMBC and g^{15}NlrHMQC correlations between H-1 (δ 7.34) and C-2 (δ 122.7), C-3 (δ 23.9), C-5 (δ 43.3), N-6 (δ -282.5) and C-7 (δ 154.1). In addition, the downfield amide proton singlet at δ 9.24 of the cyclo-dehydrated arginine residue (Cda) correlated to C-1 (δ 113.2), C-2 (δ 122.7) and C-3 (δ 23.9) and to the amide carbonyl of the valine residue at δ 169.7 that established the complete sequence and planar structure of V2 (**2**).

Table S1. NMR Data for V2 (**2**) and ^{13}C NMR Assignment Comparison with V4 (**1**) Recorded in DMSO d_6.



2



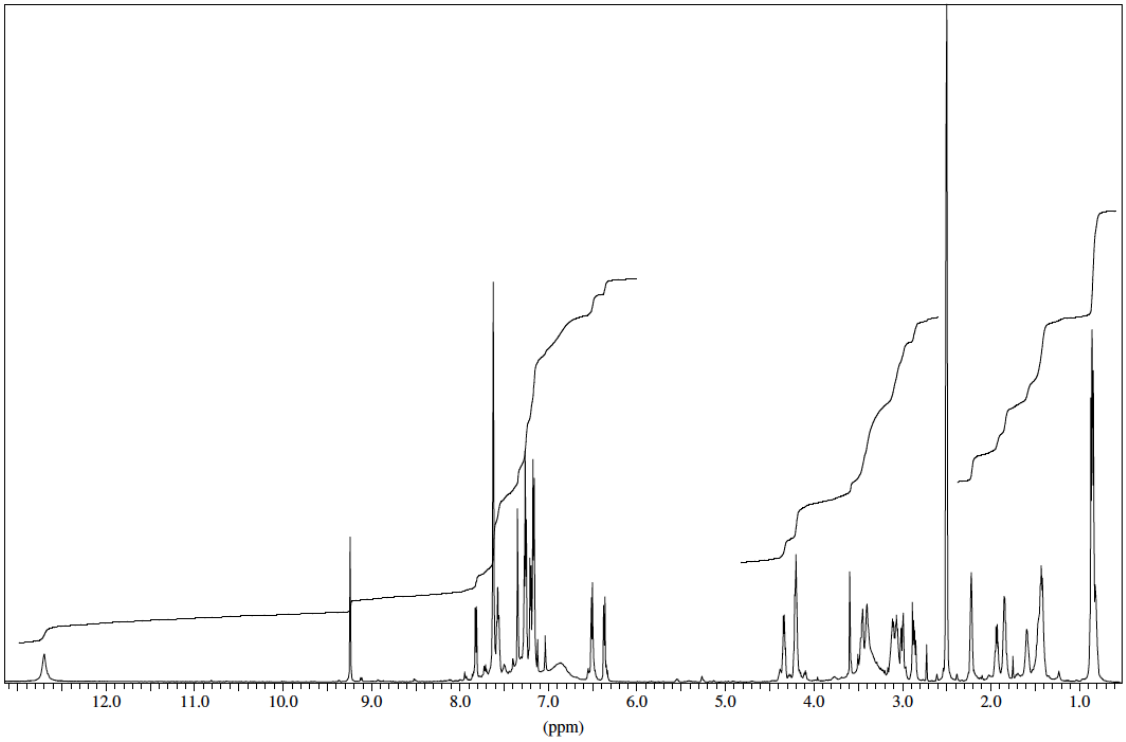
1

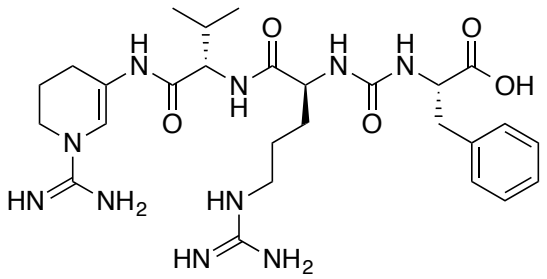
		2		**1**
residue	position	^{13}C/^{15}N[a] (δ)	^{1}H (δ)	^{13}C/^{15}N[a] (δ)
Phe	CO (1)	173.6		173.6
	α (2)	53.8	4.33 m	53.9
	β (3)	37.4	2.86 dd *J*=13.7, 7.7 Hz 3.00 dd *J*=13.7, 4.7 Hz	37.4
	i (4)	137.2		137.2
	o (5/9)	129.2	7.16 d *J*=7.2 Hz	129.3
	m (6/8)	128.1	7.26 t *J*=7.2 Hz	128.2
	p (7)	126.4	7.19 t *J*=7.2 Hz	126.4
	NH	-293.9	6.36 d *J*=7.9 Hz	-293.7
	COOH		12.70 bs	
Ureido	CO (1)	157.1		157.2
Arg	CO (1)	172.1		171.9
	α (2)	52.2	4.20[b]	52.4
	β (3)	30.1	1.41[b] 1.58 m	30.0
	γ (4)	25.0	1.42-1.47	25.0
	δ (5)	40.4	3.07 m 3.10 m	40.4
	δ–NH (6)	no	7.57 bt *J*=7.6 Hz	-265.2
	7	156.6		156.7
	NH-8	no	no	no
	NH_2-9	no	no	no
	NH	-292.5	6.50 d *J*=8.0 Hz	-291.8
Val	CO (1)	169.7		170.2
	α (2)	57.8	4.20[b]	56.9

	β (3)	30.8	1.93 m	31.1
	β Me (4)	19.1	0.85 d *J*=6.8 Hz	19.1
	β Me (5)	18.0	0.84 d *J*=7.1 Hz	17.7
	NH	-265.0	7.82 d *J*=8.5 Hz	-266.2
Cda	1	113.4	7.34 s	76.3
	2	122.7		48.7
	3	23.9	2.21 m 2.21 m	23.8
	4	20.6	1.83 m 1.83 m	23.3
	5	43.3	3.34 m 3.39 m	39.3
	N-6	-282.5		no
	7	154.1		156.6
	NH-8	no	no	no
	NH_2-9	-299.9	7.62 bs 7.62 bs	-299.7
	2-NH	-249.5	9.24 s	-257.4

[a]The ^{15}N assignments were not calibrated with an external standard. The δ value has an accuracy of about 1 ppm in reference to CH_3NO_2 (0 ppm) and are assigned on the basis of ^{15}NHSQC and ^{15}NlrHMQC correlations.
[b]Multiplicity not determined due to overlapping signals/chemical shifts determined from 2D data. no – not observed.

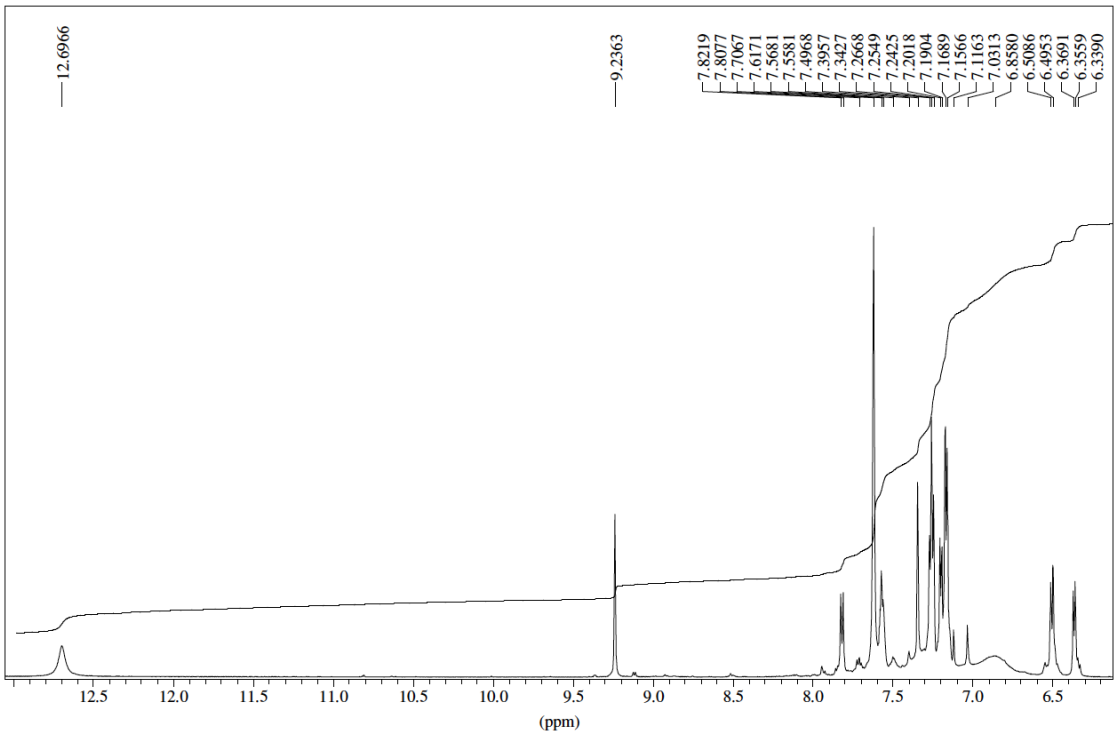
^{1}H NMR Spectrum of V2 (**2**) Recorded at 600 MHz in DMSO-d_6

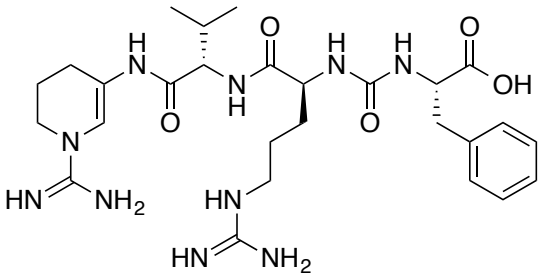




2

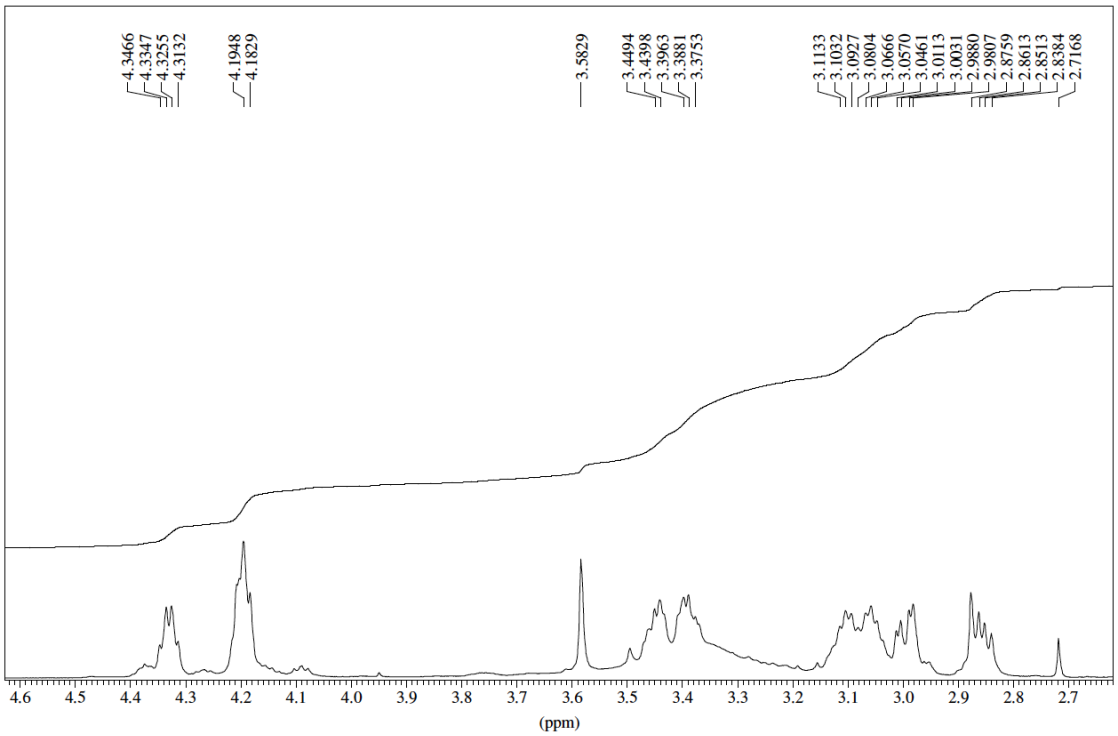
Expanded ^{1}H NMR Spectrum of V2 (**2**) with Peak Picking Recorded at 600 MHz in DMSO-d_6

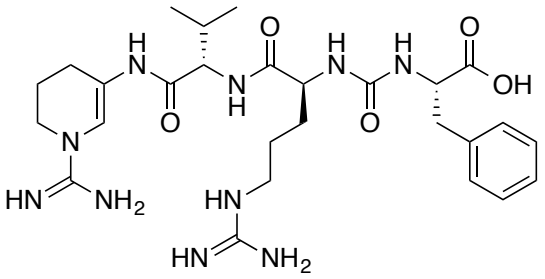




2

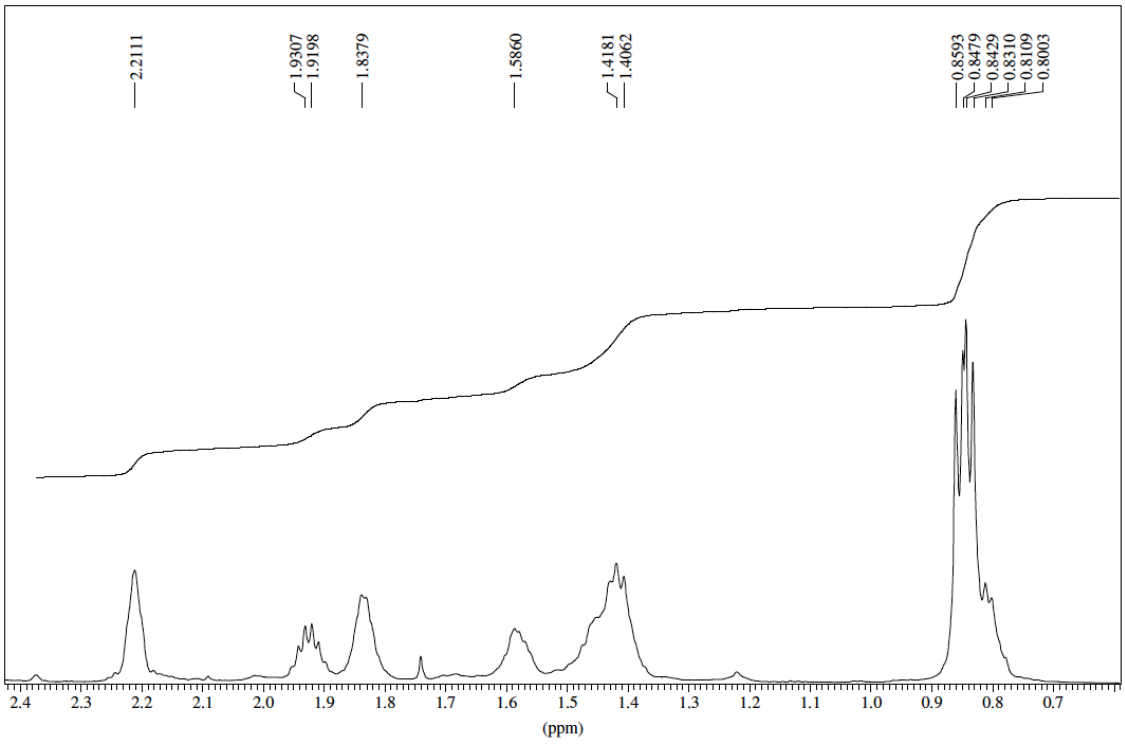
Expanded ^{1}H NMR Spectrum of V2 (**2**) with Peak Picking Recorded at 600 MHz in DMSO-d_6

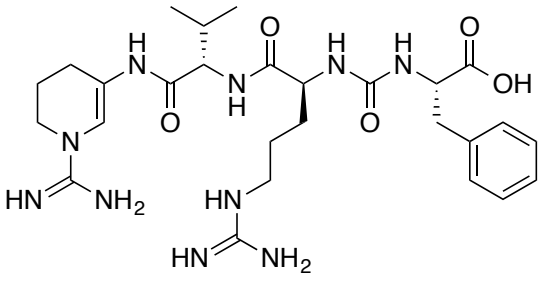




2

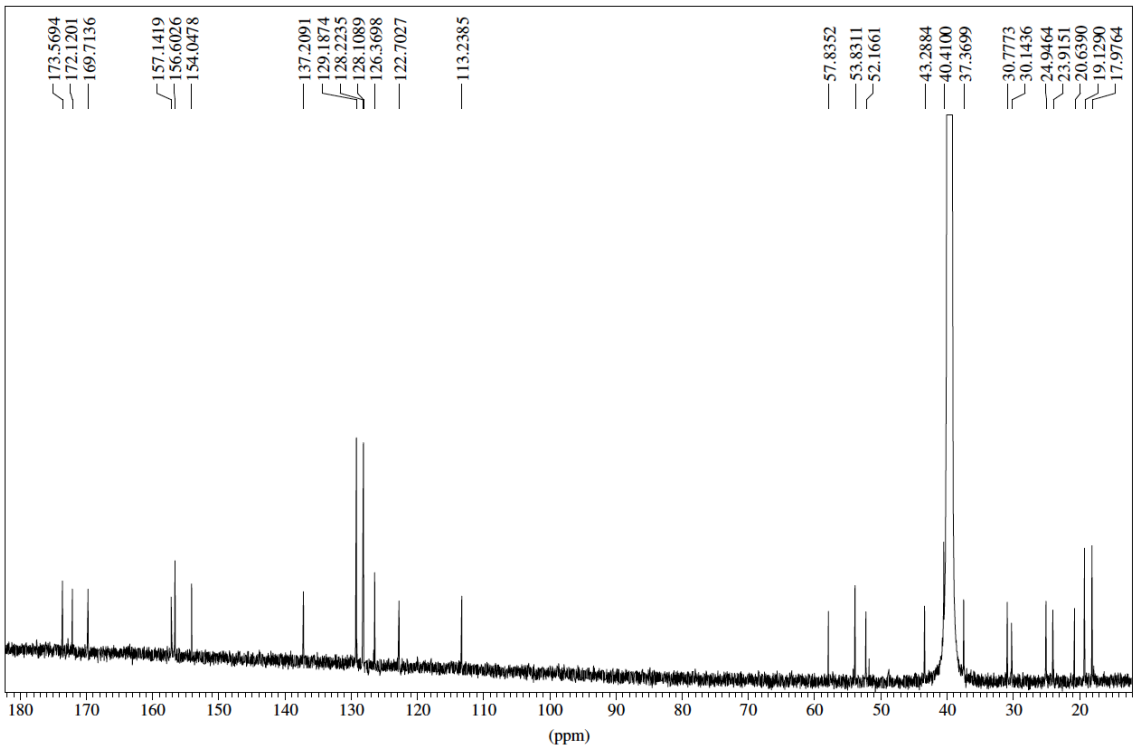
Expanded ^{1}H NMR Spectrum of V2 (**2**) with Peak Picking Recorded at 600 MHz in DMSO-d_6

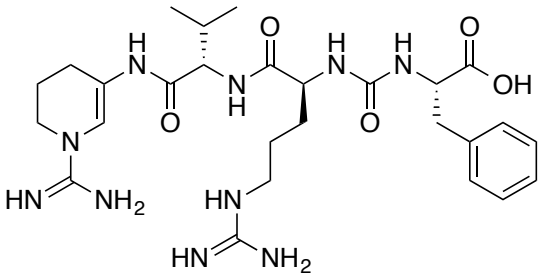




2

^{13}C NMR Spectrum of V2 (**2**) Recorded at 150 MHz in DMSO-d_6

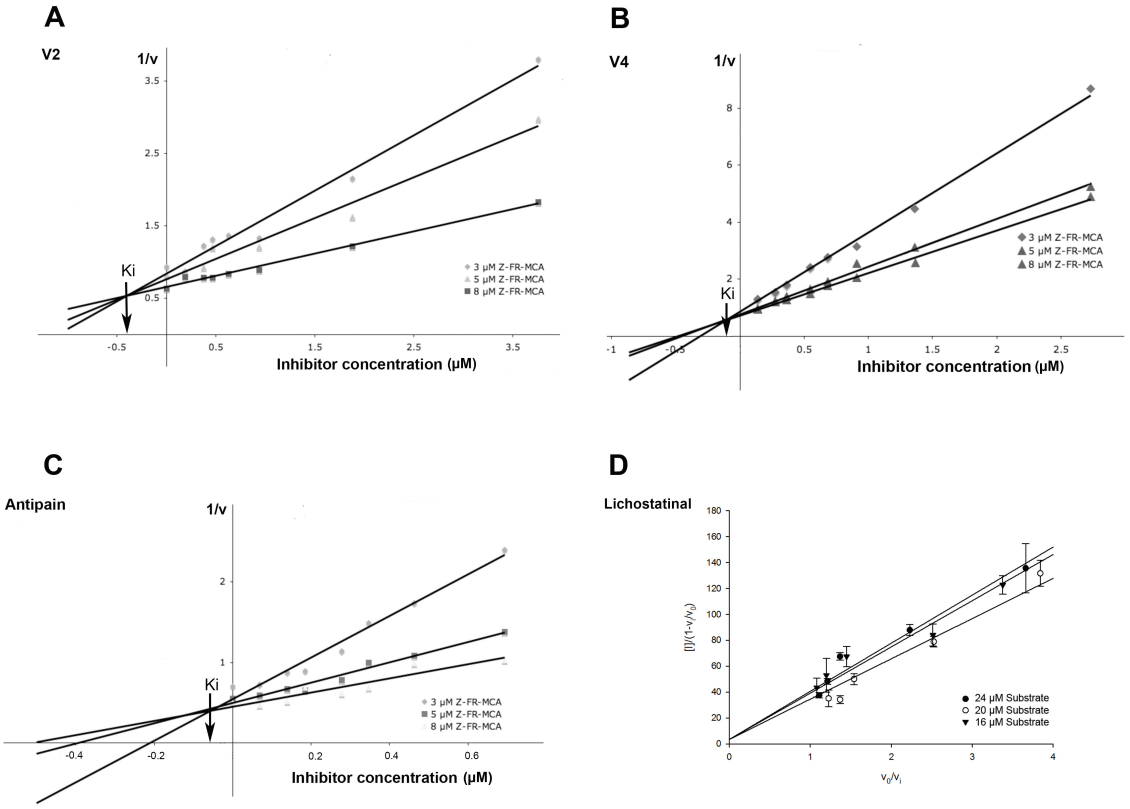




2

Figure S2. K_i Value Determination of V2 (1), V4 (2), Antipain and Lichostatinal (3)

K_i values for the inhibition of the hydrolysis of Z-FR-MCA by CatK by the inhibitors V2 (**2**) (A), V4 (**1**) (B), and antipain (C) were determined using Dixon plots. The Ki value for lichostatinal (**3**) was determined using the Henderson plot (D). The concentration of Z-FR-MCA ranged from 3 μM to 8 μM per reaction well while the concentration of the inhibitors ranged from 200 nM to 8 μM. All inhibitors were found to inhibit CatK in a competitive manner with K_i values of 393 nM for V2 (**2**) (A), 105 nM for V4 (**1**) (B), 69 nM for antipain (C), and 11 nM for lichostatinal (**3**) (D). Compound V1, an isoform of antipain, had a K_i value of 163 nM (data not shown).



Molecular Docking Methodology for Figure S3. Probable binding conformations of the V2 (**2**) and V4 (**1**) compounds were generated using Surflex-Dock. The 3D structures for V2 (**2**) and V4 (**1**) were initially calculated using SYBYL-X (version 2.1, Tripos, 2014). Charges and ligand force fields were assigned to the molecules using Tripos and Gasteiger-Hückel methods, respectively. The resulting output structures were used to perform the docking. The CatK molecule (PDB ID: 1ATK) was also prepared using SYBYL-X (version 2.1, Tripos, 2014). Ligands and water molecules were removed from the binding pocket and hydrogen atoms were added using standard geometry. Side chain amides were also checked for orientation and protonation. The protomol was generated via residue mode using Cys25 in the active site with a threshold of 0.5 and a bloat value of 3. Docking was performed with both protein and ligand flexibility enabled at the highest precision. Results were then analyzed using PyMol 1.7. **Figure S3** shows the putative binding of V2 (**2**) and V4 (**1**) to the active site of CatK.

Figure S3. Molecular Docking of V2 (2) and V4 (1) into the Active Site of CatK

Compounds V2 (**2**) and V4 (**1**) were docked to CatK using Surflex-dock. The active site cysteine residue Cys25 is depicted in yellow. 1ATK [5] was used for the CatK structure.

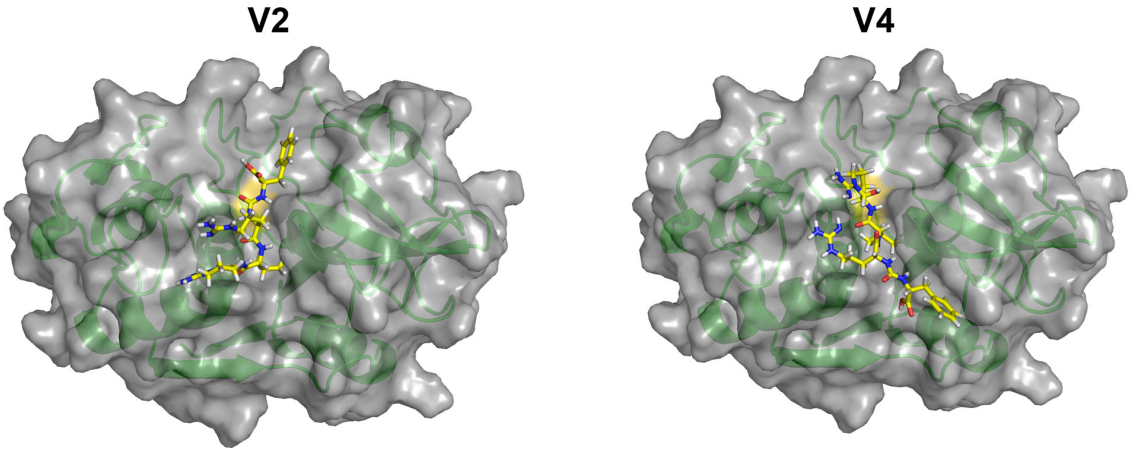


Table S2. Crystallographic Data Statistics for CatK-Inhibitor Complexes

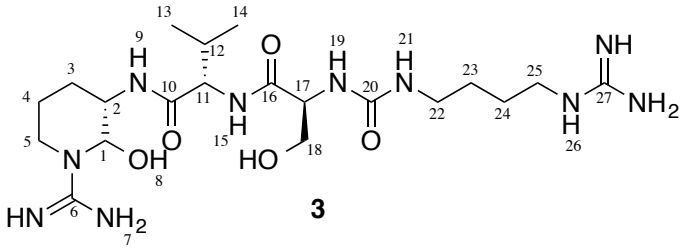
Data Collection Parameters	**Crude (sample 1)**	**Semi-purified (sample 2)**
Space group	$P4_32_12$	$P4_32_12$
Unit cell dimension (Å)	a = b = 56.72, c = 129.91, α = β = γ = 90.00	a = b = 56.72, c = 130.36, α = β = γ = 90.00
No. of total reflections[a]	199485 (25681)	134470 (20085)
No. of unique reflections[a]	20402 (2913)	15131 (2173)
Mean I/σI[a]	18.2 (4.2)	18.9 (7.7)
Multiplicity[a]	9.8 (8.8)	8.9 (9.2)
Merging R-factor (%)[a]	7.1 (49.5)	7.4 (28.0)
Maximum resolution (Å)	1.80	2.0
Structure Refinement Values		
Resolution range (Å)	51.98 - 1.80	38.3 – 2.0
Completeness (%)[a]	99.7 (99.5)	99.7 (100.0)
No. of protein atoms	1636	1636
No. of inhibitor atoms	0	35
No. of solvent atoms	133	173
Average thermal factors (all/protein/inhibitor/solvent) ($Å^2$)	31.6/30.1/-/38.5	22.8/21.1/42.7/36.6
R-factor (%)	18.2	17.5
R-free (%)	20.5	20.4
Structure Stereochemistry		
Rmsd bonds (Å)	0.004	0.006
Rmsd angles (deg)	0.801	0.845

[a]Values in parentheses are for the highest resolution shell of each data set.

Table S3. Amino Acid Residues within 3.6 Å of Lichostatinal in Sample 2 Crystal

Substrate Residue	Ligand Atom	Interacting Atom on CatK	Interaction Mediated Through	Distance in Å
P1	Arg NH1	Gly64 O	H_2O	3.57
P1	Arg NH2	Cys63 O	H_2O	3.20
P1	Arg NH2	Cys23 N	H_2O	3.18
P1	Arg O	SO_4 O1		2.74
P1	Arg O	Arg127 NH1 (sym)	SO_4 O4	2.96
P1	Arg O	Trp184 NE1	SO_4 O3	2.74
P1	Arg O	His162 ND1	SO_4 O3	3.12
P1	Arg O	Gln19 NE2		3.04
P1	Arg O	Cys25 N		3.00
P1	Arg N	Asn161 O		3.16
P2	Val O	Gly66 N		3.00
P2	Val O	Gly66 O		3.46
P2	Val O	Trp26 CD1		3.52
P2	Val CB	Gly66 O		3.43
P2	Val CG1	Ala163 CB		3.49
P2	Val CG2	Leu160 O		3.35
P3	Ser O	Leu160 O		3.17
P3	Ser O	ASN161 O		3.18
P3	URE O	Lys44 NZ (sym)		2.99
P3	URE N1	Leu160 O		3.54
P3	URE N1	Leu160 N	H_2O	2.82
P3	URE N1	Asp158 O	H_2O	3.13
P4	Agm NH1	Asp85 OD2 (sym)	H_2O	2.68
P4	Agm NH1	Ser83 OG (sym)	H_2O	2.62
P4	Agm NH2	Glu115 OE2		3.09

Table S4. NMR Data for Lichostatinal (**3**) Recorded at 600 MHz in DMSO d_6.

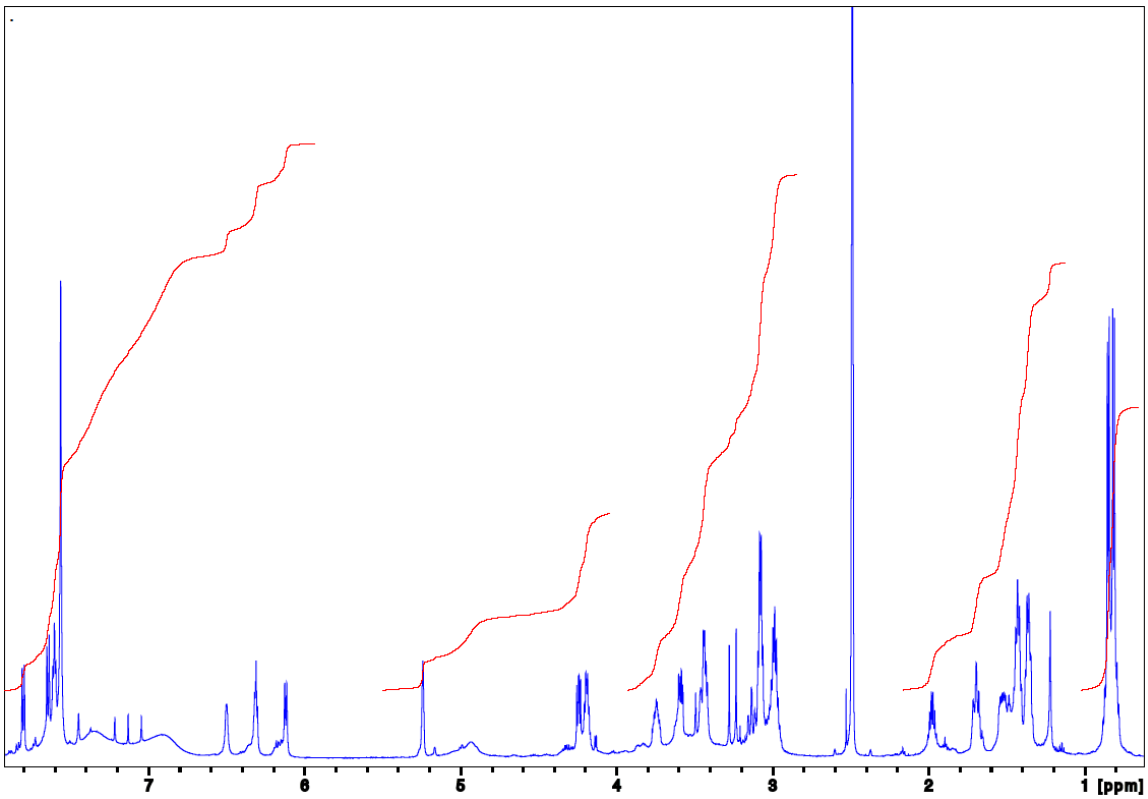


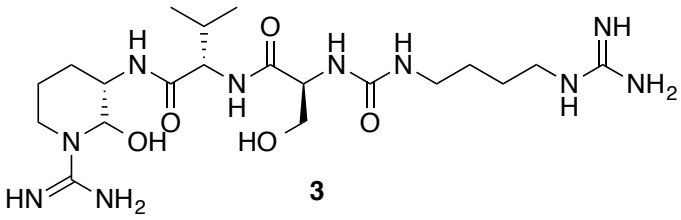
3

Atom number	^{13}C (δ)	^{1}H (δ)
1	76.2	5.24 bs
2	48.9	3.75 m
3	23.5	1.53 m 1.70[a]
4	23.3	1.53 m 1.70[a]
5	39.3	3.14 t *J*=11.2 Hz 3.45 [a]
6	156.7	
7		7.56 s
8		6.50 bs
9		7.81 d *J*=8.2 Hz
10	170.2	
11	57.1	4.25 dd *J*=6.0, 8.8 Hz
12	30.7	1.98 m
13	17.7	0.82 d *J*=6.8 Hz
14	19.2	0.85 d *J*=6.8 Hz
15		7.65 d *J*=8.8 Hz
16	171.0	
17	55.0	4.19 m
18	62.5	3.43 dd *J*=5.2, 10.3 Hz 3.59 dd *J*=5.2, 10.3 Hz
19		6.12 d *J*=7.7 Hz
20	157.7	
21		6.31 t *J*=5.5 Hz
22	38.6	2.99 m
23	27.2	1.37 m
24	25.9	1.43 m
25	40.4	3.08 m
26		7.61 t *J*=5.7 Hz
27	156.7	

[a]Multiplicity not determined due to overlapping signals/chemical shifts determined from 2D data.

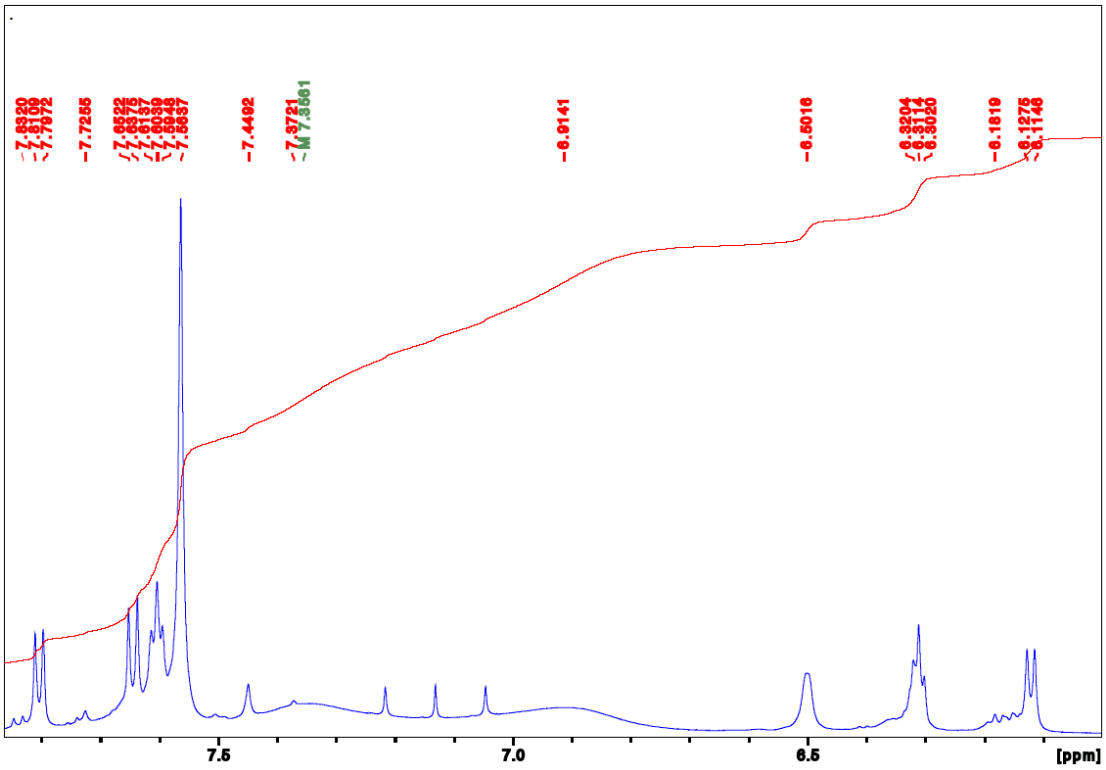
^{1}H NMR NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6

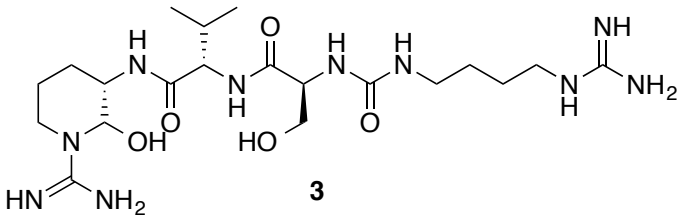




3

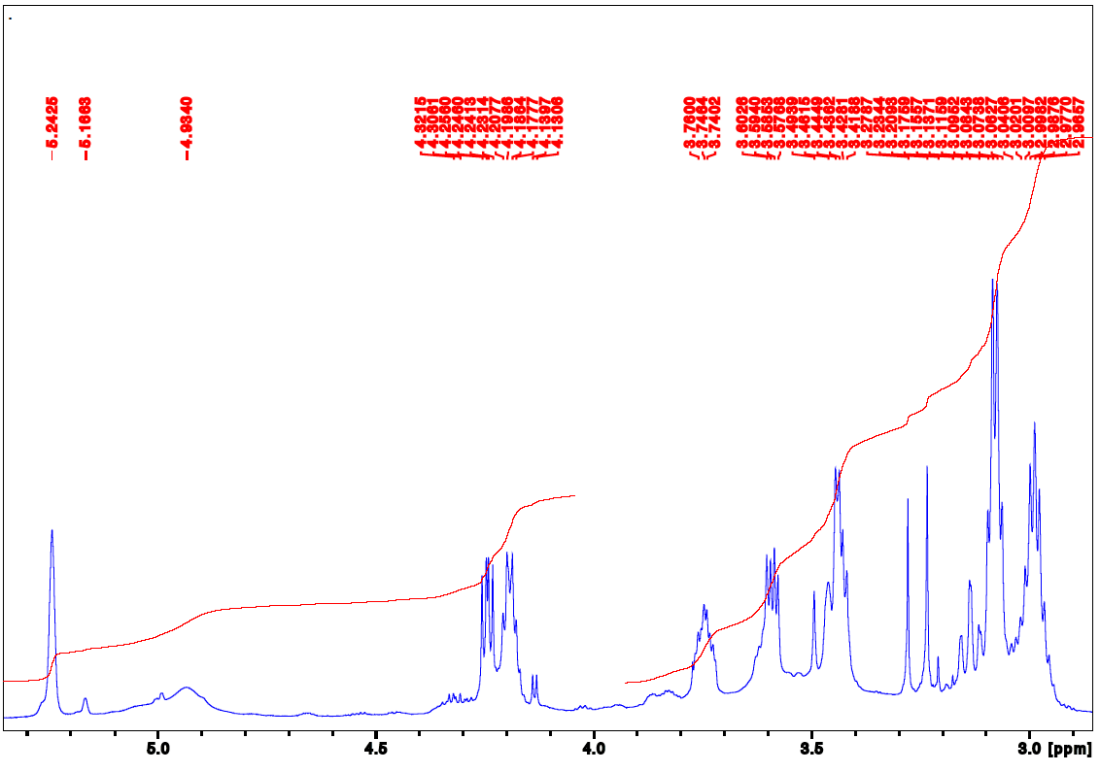
Expanded ^{1}H NMR Spectrum of Lichostatinal (**3**) with Peak Picking Recorded at 600 MHz in DMSO-d_6

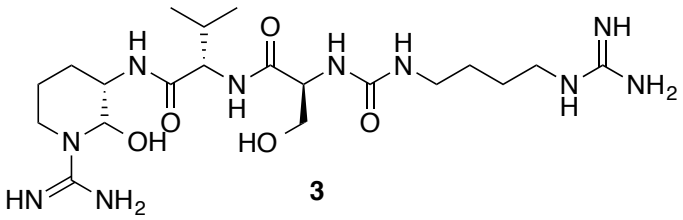




3

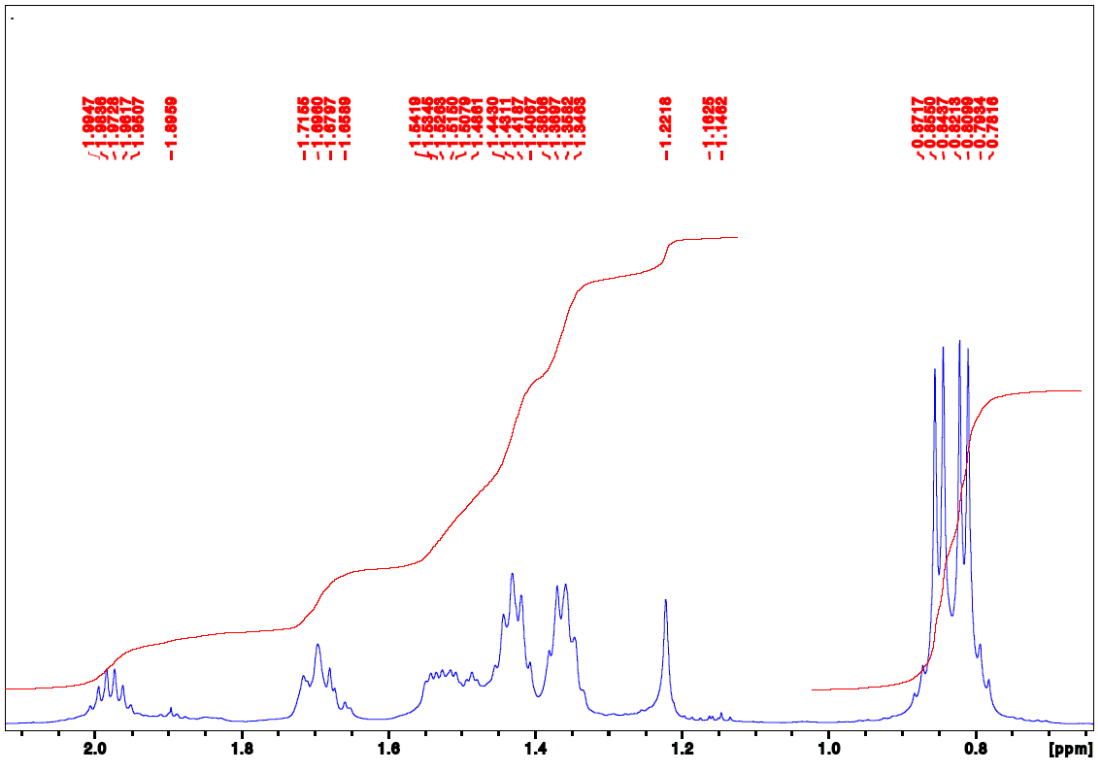
Expanded ^{1}H NMR Spectrum of Lichostatinal (**3**) with Peak Picking Recorded at 600 MHz in DMSO-d_6

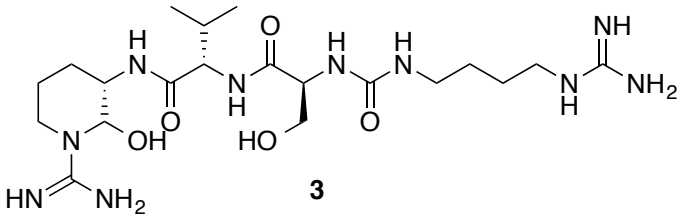




3

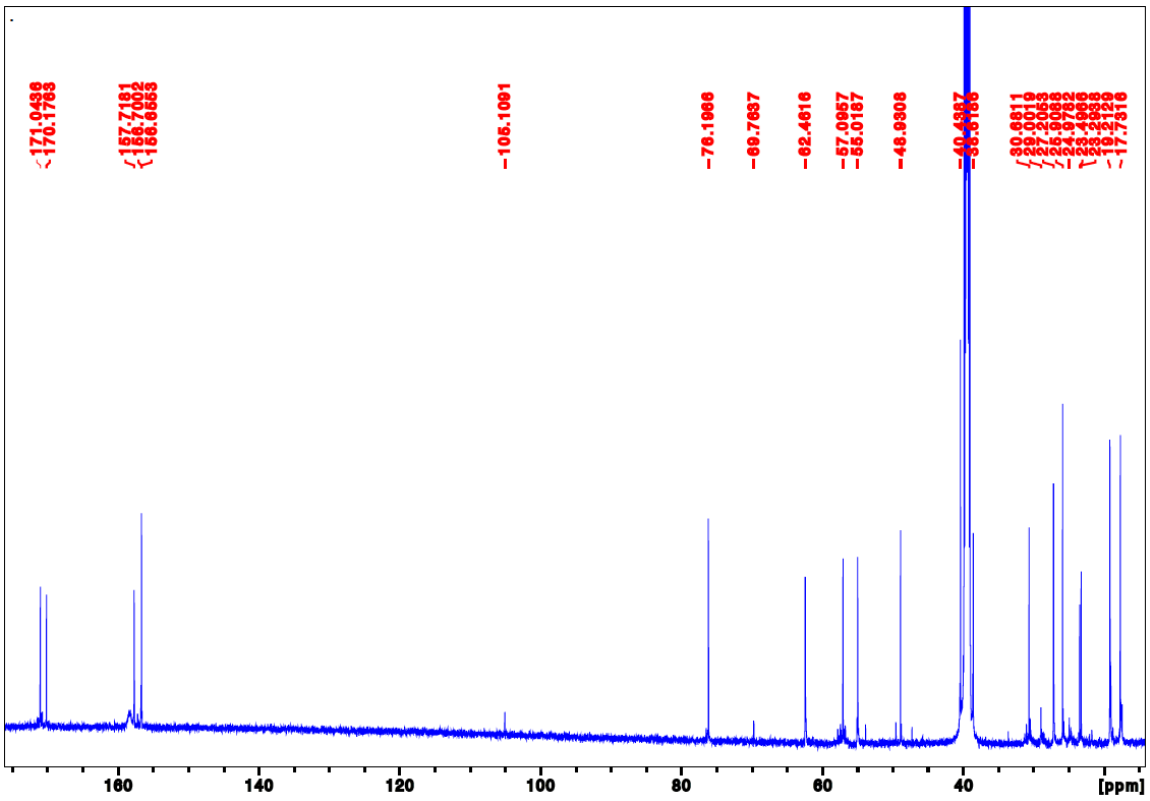
Expanded ^{1}H NMR Spectrum of Lichostatinal (**3**) with Peak Picking Recorded at 600 MHz in DMSO-d_6

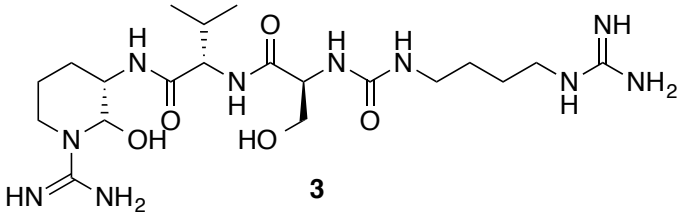




3

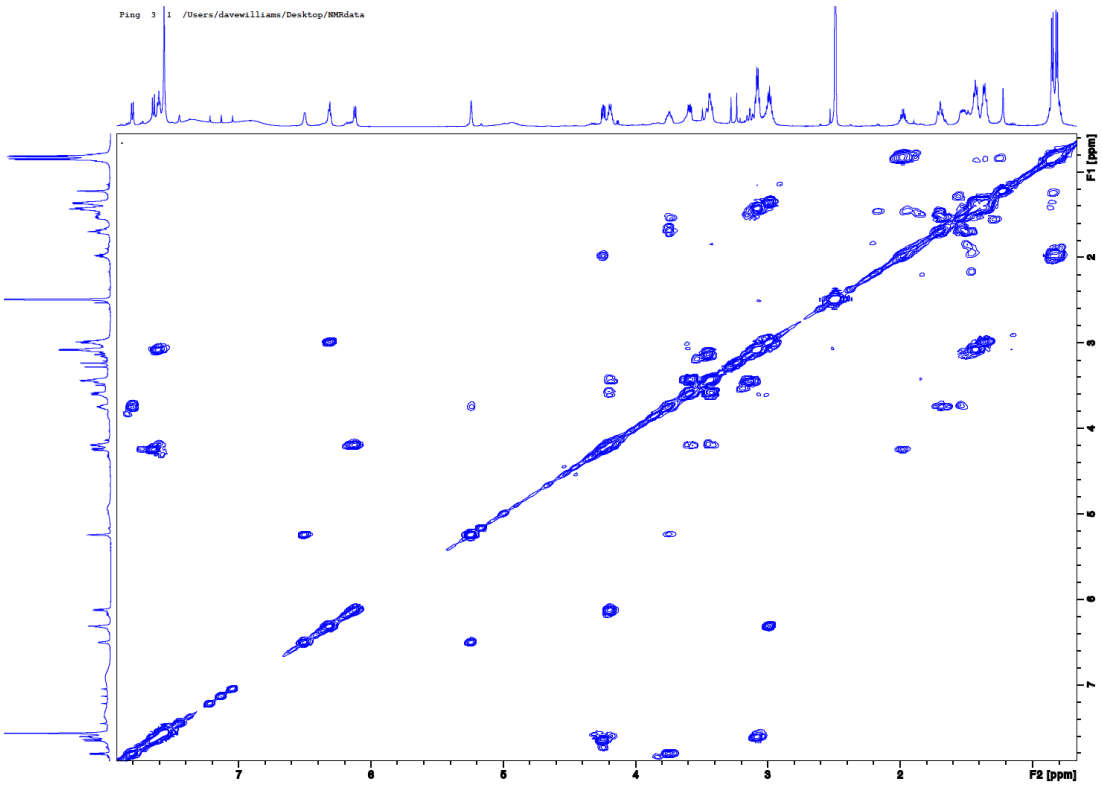
^{13}C NMR NMR Spectrum of Lichostatinal (**3**) Recorded at 150 MHz in DMSO-d_6

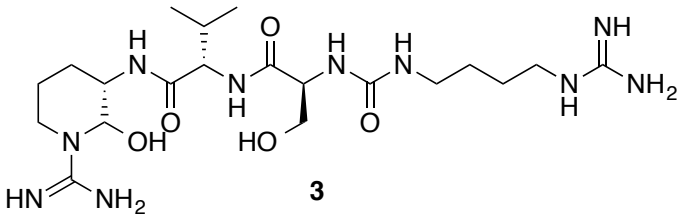




3

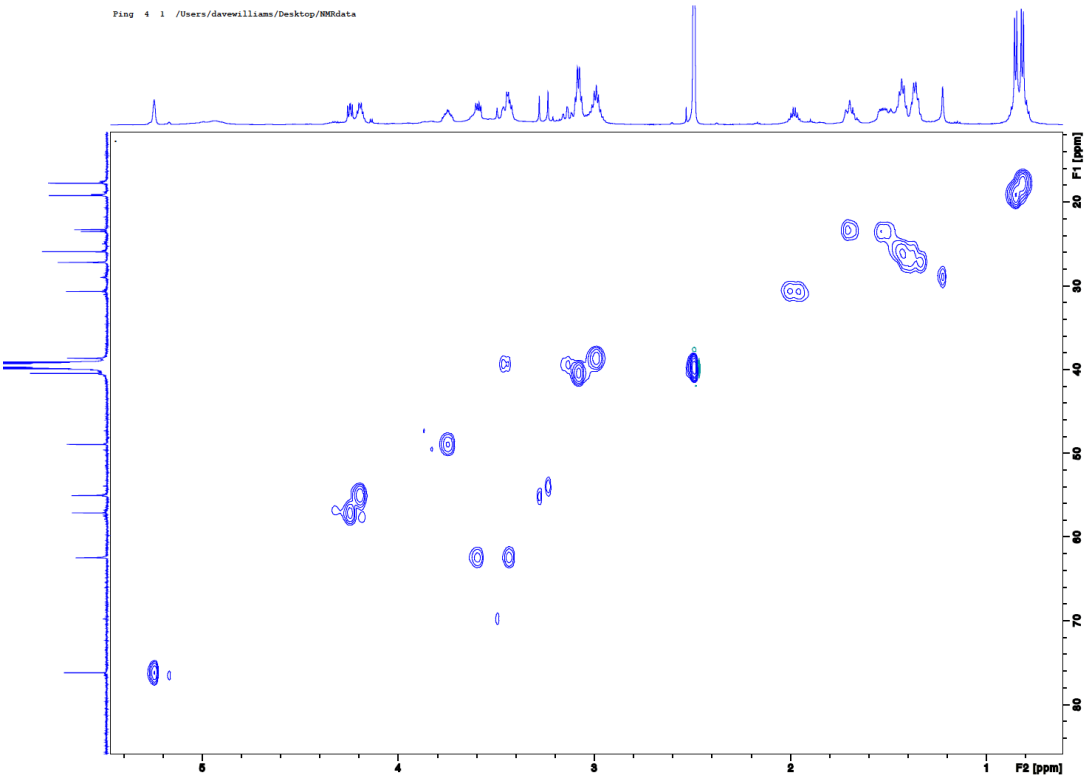
gradCOSY60 NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6

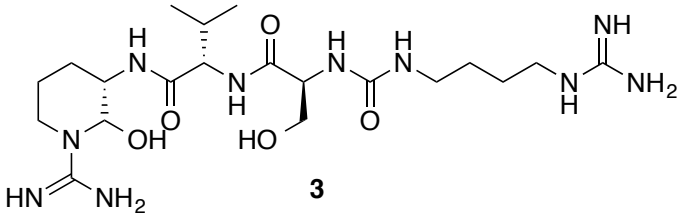




3

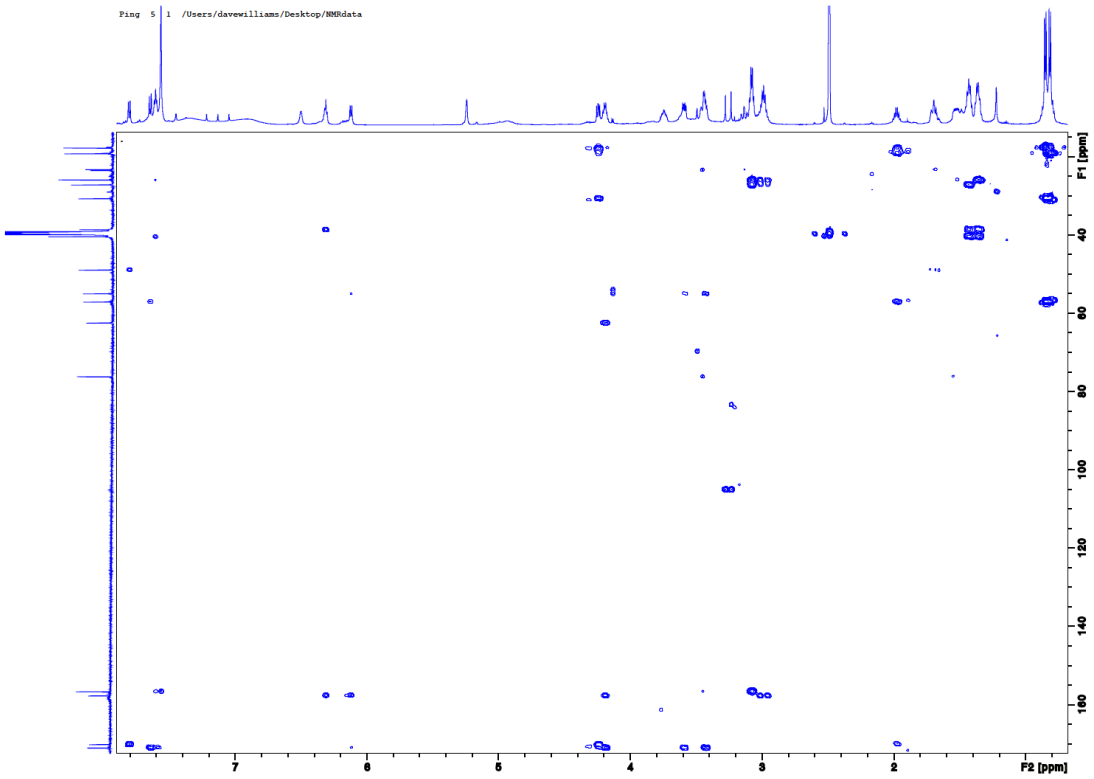
gradHSQC NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6

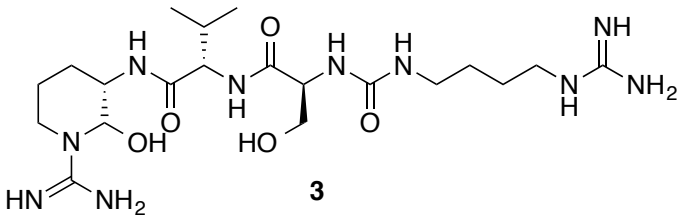




3

gradHMBC NMR Spectrum of Lichostatinal (**3**) Recorded at 600 MHz in DMSO-d_6





3

References:

(1) Suda, H.; Aoyagi, T.; Hamada, M.; Takeuchi, T.; Umezawa, H. *J Antibiot (Tokyo)* **1972**, *25*, 263.

(2) Umezawa, S.; Tatsuta, K.; Fujimoto, K.; Tsuchiya, T.; Umezawa, H. *J Antibiot (Tokyo)* **1972**, *25*, 267.

(3) Maeda, K.; Kawamura, K.; Kondo, S.; Aoyagi, T.; Takeuchi, T. *J Antibiot (Tokyo)* **1971**, *24*, 402.

(4) Hoebeke, L.; Bussatto-Samsoen, C.; Davoust, D.; Lebrun, E. *Magn. Reson. Chem.* **1994**, *32*, 220.

(5) Zhao, B.; Janson, C. A.; Amegadzie, B. Y.; D'Alessio, K.; Griffin, C.; Kurdyla, J.; McQueney, M.; Qiu, X.; Smith, W. W.; Abdel-Meguid, S. S. *Nature Struct Biol* **1997**, *4*, 109.