

Supplementary Information

Identification of potential novel insect TRPV channel modulators by homology modeling, binding mode analysis, virtual screening studies and chemical optimization

Xiaoyang Li^{a#}, Cong Zhou^{a#}, Lujue He^a, Zhiping Xu^a, Zhong Li^a and Jiagao Cheng^{a*}

^aShanghai Key Laboratory of Chemical Biology, School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China

***Corresponding author:**

Jiagao Cheng

Address: Shanghai Key Laboratory of Chemical Biology, School of Pharmacy, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237,

PO Box 544, P.R. China

E-mail: jgcheng@ecust.edu.cn

Tel: +86-21-64251348

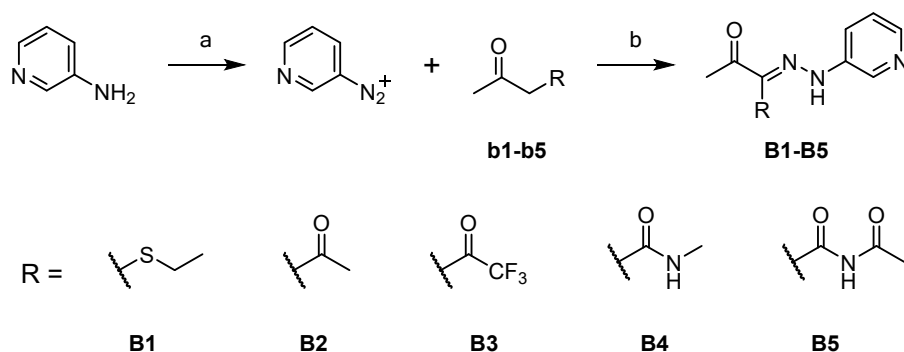
Fax: +86-21-64252603

[#]These authors contributed equally to this paper

Contents

1. Synthetic procedures and characterization of compounds
2. Sequence alignment of *N. lugens* Nanchung with rabbit TRPV5 (PDB code: 6B5V)
3. Sequence alignment of *N. lugens* Inactive with human TRPV6 (PDB code: 7S8C).
4. PROCHECK Ramachandran plot and Prosa-web z-score for *N. lugens* Nanchung and Inactive
5. ¹H NMR and ¹³C NMR spectra of the compounds **B1-B5**

1. Synthetic procedures and characterization of compounds



Scheme 1 Synthetic route for the preparation of compounds **B1-B5**. (a) HCl, NaNO₂, 0 °C, (b) NaOAc, EtOH, 0-5 °C.

3-aminopyridine (3 mmol, 1.0 eq) was dissolved in 6 M HCl (2 mL), the solution was cooled to 0°C and sodium nitrite (270 mg, 1.3 eq) was added. The reaction mixture was stirred for 30 min at 0°C to obtain 3-pyridyl diazonium salt, which was used without further purification.

The solution of diazonium salt was added dropwise to a solution of compound **b1-b5** (3 mmol, 1 eq) in NaOAc aqueous (10 %, 25 mL) and EtOH (3 mL) at 0-5 °C. The mixture was stirred for 5 h at 0-5 °C, and the completion of the reaction was confirmed by TLC analysis. The aqueous was then extracted with CH₂Cl₂ (3 × 30 mL), dried over Na₂SO₄, filtered, concentrated in vacuo and purified by flash column chromatography to yield compound **B1-B5**.

Data for Ethyl-2-oxo-*N*-(pyridin-3-yl)propanehydrazonothioate (B1**):** yellow solid; yield 57.0%; mp 59.4 - 60.2°C; ¹H NMR (400 MHz, CDCl₃) δ 9.23 (s, 1H), 8.57 (d, *J* = 2.6 Hz, 1H), 8.31 (dd, *J* = 4.8, 1.4 Hz, 1H), 7.60 (ddd, *J* = 8.4, 2.6, 1.4 Hz, 1H), 7.29 (dd, *J* = 8.4, 4.8 Hz, 1H), 2.95 (q, *J* = 7.4 Hz, 2H), 2.54 (s, 3H), 1.24 (t, *J* = 7.4 Hz, 3H) ppm. ¹³C NMR (100 MHz, CDCl₃) δ 193.67, 144.22, 138.56, 137.56, 136.90, 124.13, 121.31, 27.70, 25.57, 15.97 ppm. HRMS (ESI) calc. for C₁₀H₁₃N₃OS [M+H]⁺: 224.0852, found: 224.0855.

Data for 3-(2-(pyridin-3-yl)hydrazono)pentane-2,4-dione (B2**):** yellow solid; yield 88.9%; mp

84.6 – 85.7 °C; ^1H NMR (400 MHz, CDCl_3) δ 14.62 (s, 1H), 8.72 (d, $J = 2.8$, Hz, 1H), 8.47 (dd, $J = 4.8$, 1.6 Hz, 1H), 7.78 (ddd, $J = 8.4$, 2.8, 1.6 Hz, 1H), 7.38 (dd, $J = 8.4$, 4.8 Hz, 1H), 2.64 (s, 3H), 2.52 (s, 3H) ppm. ^{13}C NMR (100 MHz, CDCl_3) δ 198.43, 196.85, 146.60, 138.59, 138.13, 134.41, 124.15, 122.90, 31.69, 26.58 ppm. HRMS (ESI) calc for $\text{C}_{10}\text{H}_{11}\text{N}_3\text{O}_2$ ($\text{M}+\text{H}$) $^+$: 206.0924, found: 206.0923.

Data for 1,1,1-trifluoro-3-(2-(pyridin-3-yl)hydrazono)pentane-2,4-dione (B3): yellow solid; yield 76.4%; mp 76.6 – 77.6°C; ^1H NMR (400 MHz, CDCl_3) δ 15.10 (s, 1H), 8.72 (d, $J = 2.6$ Hz, 1H), 8.55 (dd, $J = 4.8$, 1.4 Hz, 1H), 7.88 (ddd, $J = 8.4$, 2.6, 1.4 Hz, 1H), 7.42 (dd, $J = 8.4$, 4.8 Hz, 1H), 2.66 (s, 3H) ppm. ^{13}C NMR (100 MHz, CDCl_3) δ 198.00, 177.22 (q, $^2J_{\text{CF}} = 32.6$ Hz), 148.28, 139.52, 137.46, 129.12, 124.56, 123.89, 117.23 (q, $^1J_{\text{CF}} = 290.5$ Hz), 31.06 ppm. HRMS (ESI) calc for $\text{C}_{10}\text{H}_8\text{F}_3\text{N}_3\text{O}_2$ ($\text{M}+\text{H}$) $^+$: 260.0641, found: 260.0639.

Data for *N*-methyl-3-oxo-2-(2-(pyridin-3-yl)hydrazono)butanamide (B4): white solid; yield 90.7%; mp 178.8 – 179.6°C; ^1H NMR (400 MHz, CDCl_3) δ 14.80 (s, 1H), 9.24 (s, 1H), 8.64 (d, $J = 2.6$ Hz, 1H), 8.37 (dd, $J = 4.8$, 1.5 Hz, 1H), 7.66 (ddd, $J = 8.4$, 2.6, 1.5 Hz, 1H), 7.30 (dd, $J = 8.4$, 4.8 Hz, 1H), 2.89 (d, $J = 4.9$ Hz, 3H), 2.49 (s, 3H) ppm. ^{13}C NMR (100 MHz, CDCl_3) δ 199.04, 165.41, 145.75, 138.32, 138.18, 127.76, 124.01, 122.18, 25.98, 25.17 ppm. HRMS (ESI) calc for $\text{C}_{10}\text{H}_{12}\text{N}_4\text{O}_2$ ($\text{M}+\text{H}$) $^+$: 221.1033, found: 221.1031.

Data for *N*-acetyl-3-oxo-2-(2-(pyridin-3-yl)hydrazono)butanamide (B5): yellow solid; yield 76.4%; mp 157.3– 158.4°C; ^1H NMR (400 MHz, CDCl_3) δ 14.43 (s, 1H), 11.72 (s, 1H), 8.72 (dd, $J = 2.6$, 0.8 Hz, 1H), 8.49 (dd, $J = 4.8$, 1.4 Hz, 1H), 7.75 (ddd, $J = 8.4$, 2.6, 1.4 Hz, 1H), 7.39 (ddd, $J = 8.4$, 4.8, 0.8 Hz, 1H), 2.57 (s, 3H), 2.43 (s, 3H) ppm. ^{13}C NMR (100 MHz, CDCl_3) δ 198.90, 171.03, 163.01, 147.13, 138.68, 137.71, 126.28, 124.23, 122.96, 26.28, 26.16 ppm. HRMS (ESI) calc for

C₁₁H₁₂N₄O₃ (M+H)⁺: 249.0982, found: 249.0980.

2. Sequence alignment of *N. lugens* Nanchung with rabbit TRPV5 (PDB code: 6B5V)

```
Nan : MGNTESNVTSGVKKQTDTSSTIMYKLVLDLKGGLLVEELMKRATQNKQYAEILDHAIKTKVEPFYLNKG : 67
TRPV5 : -----

Nan : NGRI.TSTAKI.VI.I.RNKDRPPTKMTIPPLKMNENPDDFDVIGSNIPGIPEPK--ETDDPSTYRDVVCWDL : 132
TRPV5 : -----WEQYRDRVNMIL-----QQER---TRDSPLLQAAKENDLRLLLKILLNQSCLF : 44

Nan : NERGAVGETI.LHTCLINATSVHADLAKRI.LKIFYPKIINDIYMCDEYYGESV.LHTATVNE.DPGMVKFTL : 199
TRPV5 : QORGAVGETALHVAALYDN---LEAATLLMEAAPEI.LAKEPALCEPEVGGTALHIAVMNQNLNLRAL : 108

Nan : I.NSGAD.FHER.CFCGNFMCPEDQKASRSDSFDHEWVNLWPVTNYEGYVYWGEYPLSFAACLGOEFCYRL : 266
TRPV5 : LARGASVSARATG-----AAFRRSP-----HNL.IYYGEHPLSFAACVGSSEEIVRL : 153

Nan : MIARGANPDNODTNGNTVI.HMI.VIY.EKMGTFDTAYE.I.GANVSVR-----N.IONT.TPLT.TAA : 322
TRPV5 : LIEHGADIRAQDSLGNLTVLHILILIQPNKTFACQMYNLLLSYDEHSDHLQSLELVPNHQGLTPFKLAG : 220

Nan : KLARIEMFFHIMNIERETIYQIIGSITCAAYPTMLIDTIDIITGNISKDSAINLVVFGDKDEHLDLME : 389
TRPV5 : VEGNTVMFQHLMQKRKHVQWTCGPIITSTLYDITELDSWGEEL-----SFEELVVSSKREARQILE : 281

Nan : G-VI.VD.I.HAKWNAFVVKFRFYRQFIMFQYETLISCVCETLRPGPPEAHDGMNITANATLNLISFGVLE : 455
TRPV5 : QTPVKELV.SFKWKKYGKPYFCVLASLYILYMLCFTTCCITRPLKLE----- : 326

Nan : VRPEAL.OALLYRIEKL.SL.POSGDLERFILE.SLADEKSI.SEAAL.ENDENREPGWNNNSVFSTYETLARD : 522
TRPV5 : -----

Nan : TGYOTDESPHDGLSSSMDYAEVAERLTKLMIDVNLTERSGAVNEIILKALPSLEEKEDGLEEWWSGI : 589
TRPV5 : -----

Nan : AGDCTT.LQVNDLYL.LART.CAFVALEI.GAETIYITAAI.REARETGLKMMIEN-IMTAPSRVMT.FSCCT : 655
TRPV5 : -----RAYVTHQDNIRLVGELVTVTGAVIILILELIPDIFRVGASRYFGQTI.LGGPFHVIIITTYASL : 387

Nan : MMTMPFFRFITCODETE.DI.LAVVIMLITGPFYI.FFCRGEKTVGPFVMTYRMVMGDLIRFASTYLVFV : 722
TRPV5 : VLLTMVMRLITNM-NGEVVPLSFALV.LGWCSVMYFARGEQMLGPFTIMIQKMI.FGDLMREFCWLMAVVI : 453

Nan : MGF.SQAYYIIFLTYDNPTTPEGTDDTTSNPLPSPTFEAIMAMFMSITNFGDYYSAFENTEHEIEAKA : 789
TRPV5 : LGFASAFHITFQTED----PNNLGEFS--YPTALFSTFELFELTIIDGPANYSVLDLP-----FMYCI : 509

Nan : FFVIY.MVTVAT.I.T.NM.LTAMMGNTYOKTAFKKNRWRQWARIVLVVERGVSPKERLKLMSYSQPMS : 856
TRPV5 : TYAAFAIATLLMLNLFIAMMGDTHWRVAQERDELWRQAQVVATTVMLEKMPRFLWPRSGICGYEYG : 576

Nan : DGRRALVLRLLNMTFEEDKEEMKEILEMKRVHDKIVMKRKKGLKLLGTTTSPPKRMKGRETPI : 917
TRPV5 : LGDRWFLRVENHH----- : 589
```

Figure S1 Sequence alignment of *N. lugens* Nanchung with rabbit TRPV5 (PDB ID 6B5V).

Conserved residues are marked with purple

3. Sequence alignment of *N. lugens* Inactive with human TRPV6 (PDB code: 7S8C).

```

Iav : MGNLLSGGANLNAGSVLDRVISQGSTEDCELLYRLANYRNGGELIDAYNVGGOTEVEKLIKQFGVLI : 67
TRPV6 : -----SWAQSREQNLLQO-----KRIWESPLLI : 24

Iav : MYKDGKGOVINRSEYLRWKFRDTRKQVRLPIEASLSPHDPAAKWEDHDAWQMOYRGSLETLLHVLI : 134
TRPV6 : AAKDNDVQALNK-----LLKYED---CKVHQ-RGAMGETALHIAA : 60

Iav : TCDTKLHTRLARTITKCFPNIAMDVVEGFYLCASATHIATAYFNNETVQETVEAGANVMORAIGSF : 201
TRPV6 : LYDN---LEAAMVLMEEAPELVFEPMTSELYEGQTALHIAVNVQNMNLVRAALLARRASVSARATGTA : 124

Iav : FLPRDQQRARPSRHTDYEGTAYLGFEPILAWAACCANFQSVYNLLIDSGAHPDHODSEFGNMTLHMVVV- : 267
TRPV6 : FR-----RS-----PRNLIYFGEHPLSFAACVNSEETVRLLIETHGADIRAQDSLGNITVLHILILQ : 179

Iav : -----CDKLDMEFGYALRHP-KIPASNGITNNEGITPLTLACKLGRAEVRFMLELSAREFWRYSNIT : 327
TRPV6 : PNKTFAQOMYNLLLSYDRHGDHLQPLDLVFNHOGLETPEFKLAGVECNTVMFOHLMQKRKHTQWTYGPL : 246

Iav : TCSAYPIINALDTLLPDGRTNWNSAIFITTINGTKEFHLDMLDGGITQRLIEEKWKTFARLQFLKRLCI : 394
TRPV6 : TSTLYDLTEIDSSG-D----EQSLLELLITTKKREARQLLDQTPVKELVSLKWKRYGRPYECMLGAI : 308

Iav : LAIHLIIMLSITAVYLRPIID-RSKSLLEGDD-----WADYARYGFETGTVIGVLSYLVV : 445
TRPV6 : YLLYIICFTMCCIYRPIKPRITNNRTSPRDNTLLQKLLQEAYVTPKDDIRLVGELVTVIGATIIILLV : 375

Iav : QQGGEIKNQGFSSFIKQLDPAKATETLISNITLACTPEFRIAGDILMFEATVVAIPGSWFFTFMFFAG : 512
TRPV6 : EVP-----EFHVLIITTYAFMVLVTVMMRLIS-ASGEVVPMSFALVLGWCNVMYFAR : 425

Iav : AVRIITGPFVTMVYSMITGDMITFGIITYTVLFGFSQTFEFVYKGFPGVKNTLYGTYPITWMALEFOIT : 579
TRPV6 : GFQMLGPFTIMIQKMIFFGDLMRFCWLMMAVVILGFASAFYITFQTEDPEELGHEFYDYPMALESTFELF : 492

Iav : TGDYNYAELANTTYPSVAKTVFAIFMVFVPTLLLNMLTAMMGNTYAHVIEOSEKEWMKQWAKIVVSTI : 646
TRPV6 : LTIIDGPANYNVDLPEMYSITYAAEATLTLMLNLLIAMMGDTHWRVAHERDELWRAQVVATTVML : 559

Iav : FRAVPHEDCHNYLQEYSIKLGPGETPGTEQRAVMVIKSKSKTRAKQRKGALSNNWKVKRKFVLLLVGG : 713
TRPV6 : ERKLPRLCWPRSGICGREYGLGDRWFLRVEDRQD----- : 593

Iav : ECRSLOEGEPOLKIIIIOMGDIHOVSTIFTI : 743
TRPV6 : ----- : -

```

Figure S2 Sequence alignment of *N. lugens* Inactive with human TRPV6 (PDB ID 7S8C).

Conserved residues are marked with purple

4. PROCHECK Ramachandran plot and Prosa-web z-score for *N. lugens* Nanchung and

Inactive

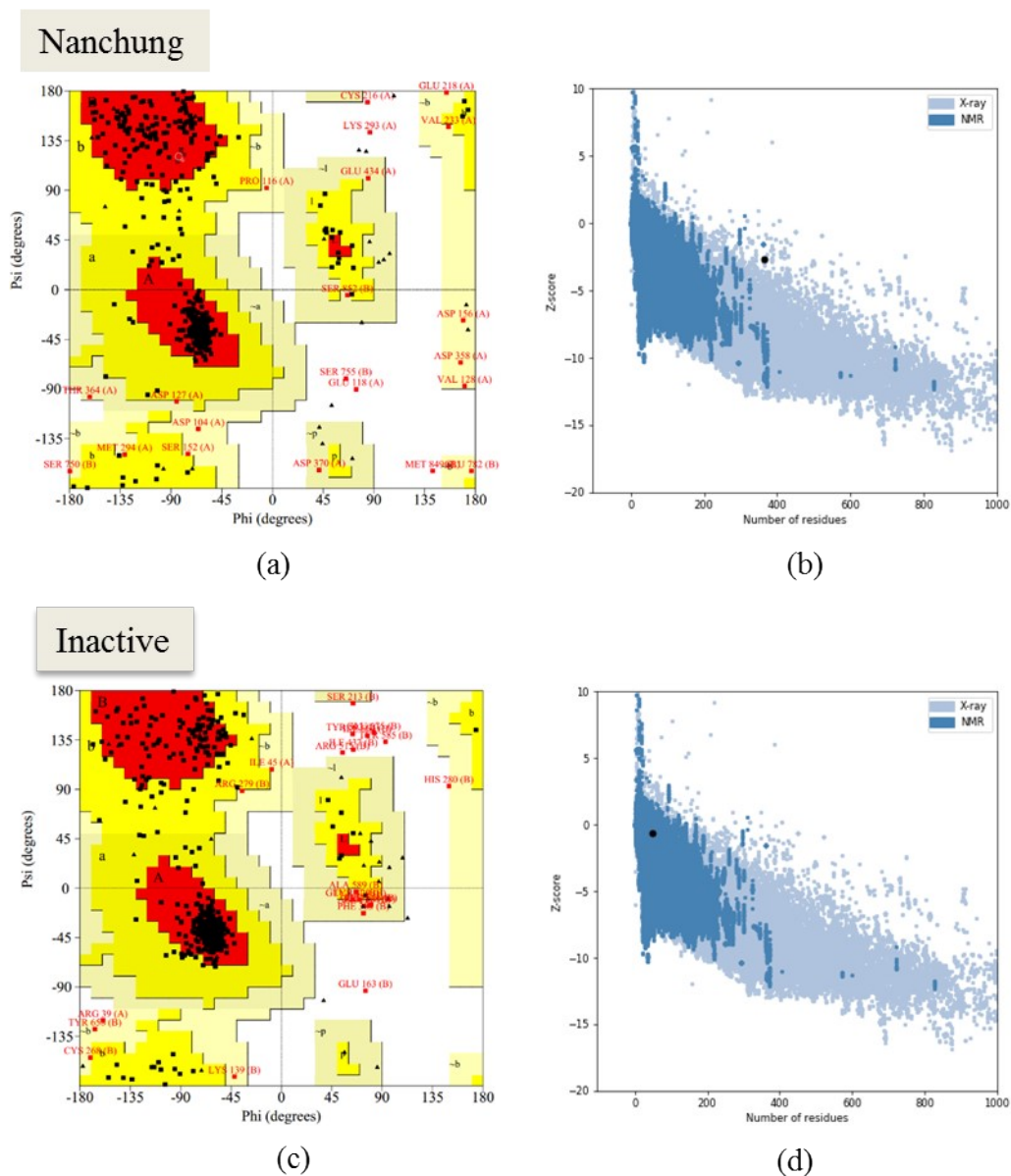


Figure S3 The Ramachandran plots (a), the Z-Score plots (b) results of *N. lugens* Nanchung model and the Ramachandran plots (c), the Z-Score plots (d) results of *N. lugens* Inactive model. In Ramachandran plots, residues are color-coded as red (favorable), yellow (allowed), light yellow (generously allowed), and white (disallowed) based on their ψ and ϕ angles.

5. ^1H NMR and ^{13}C NMR spectra of the compounds B1-B5

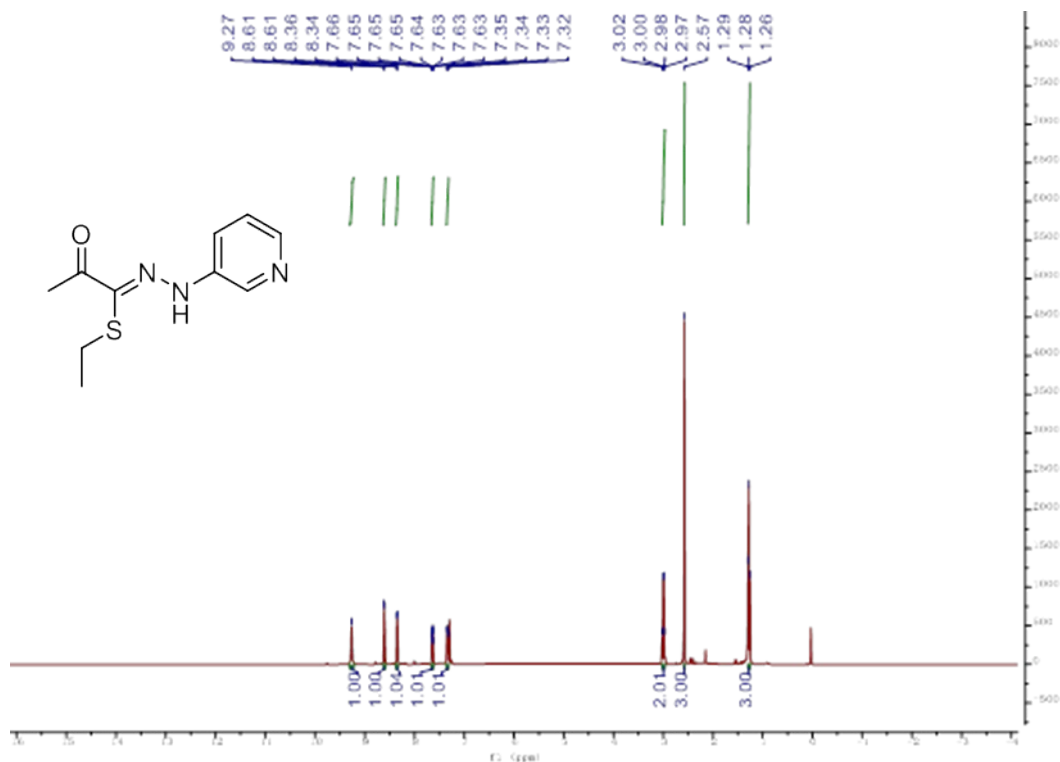


Figure S4 ^1H NMR spectrum of B1 (400 MHz, CDCl_3)

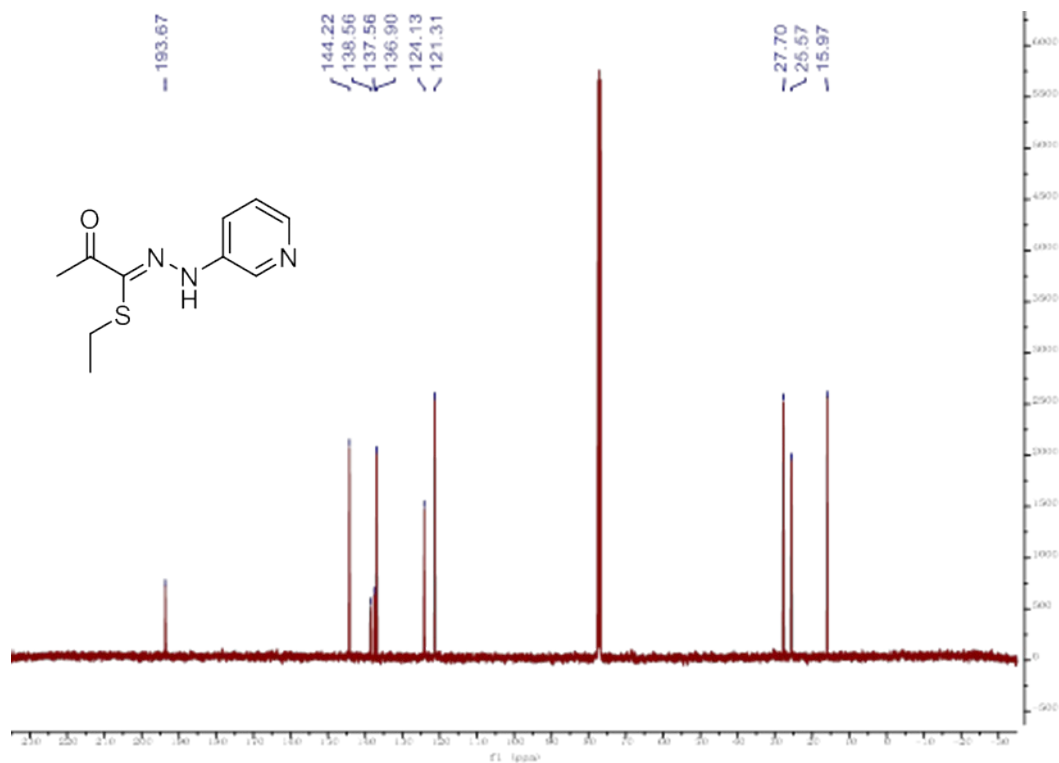


Figure S5 ^{13}C NMR spectrum of B1 (100 MHz, CDCl_3)

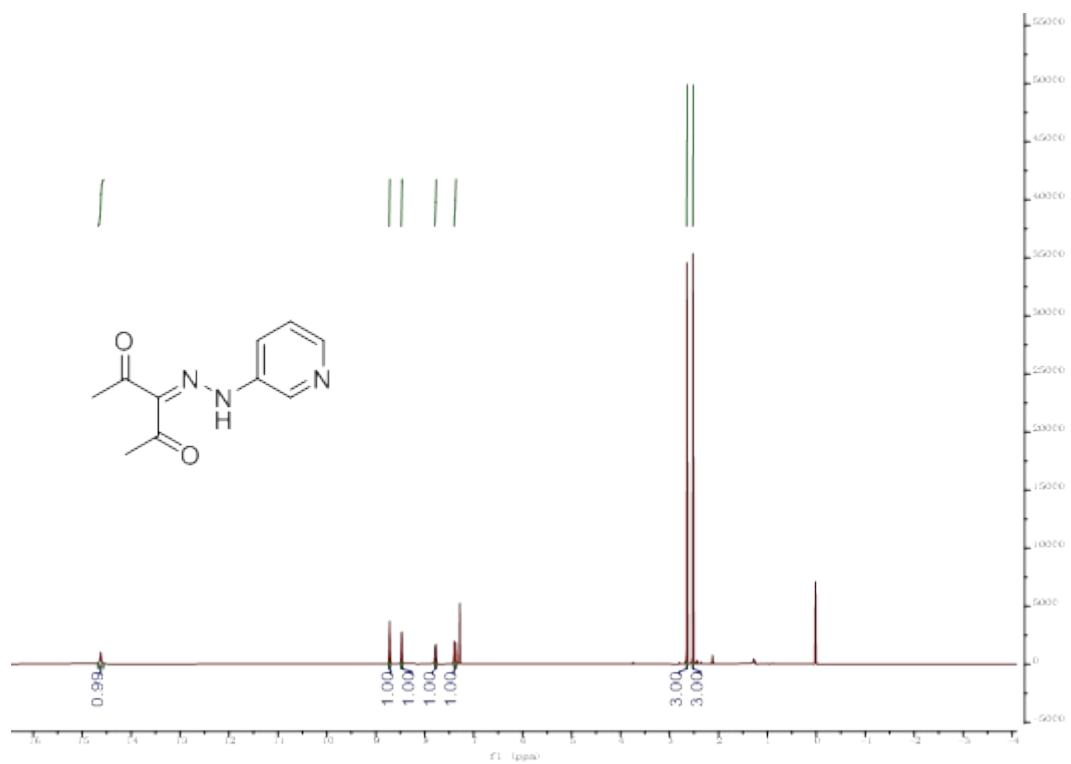


Figure S6 ¹H NMR spectrum of B2 (400 MHz, CDCl₃)

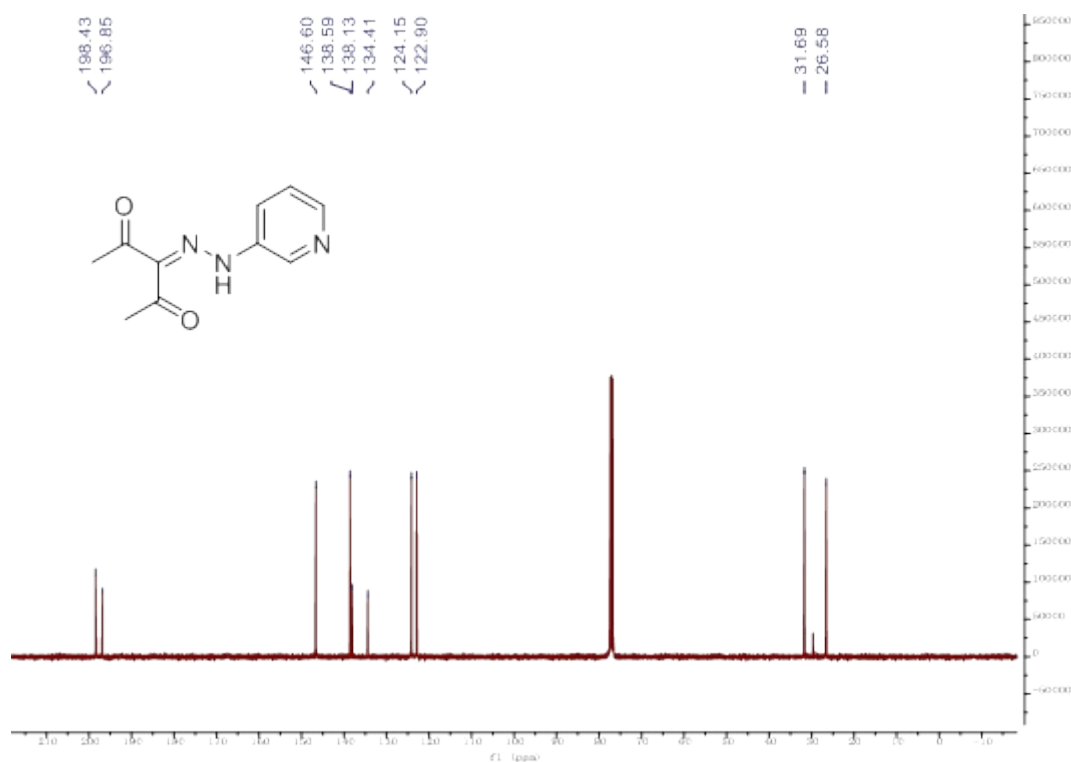


Figure S7 ¹³C NMR spectrum of B2 (100 MHz, CDCl₃)

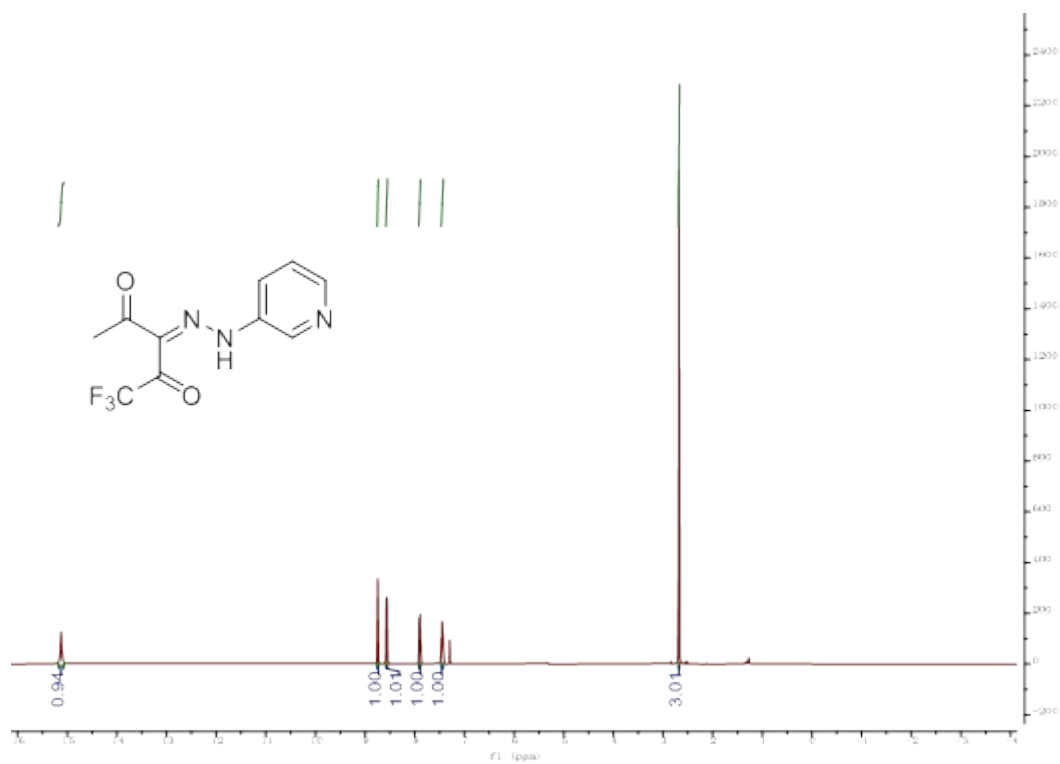


Figure S8 ¹H NMR spectrum of **B3** (400 MHz, CDCl₃)

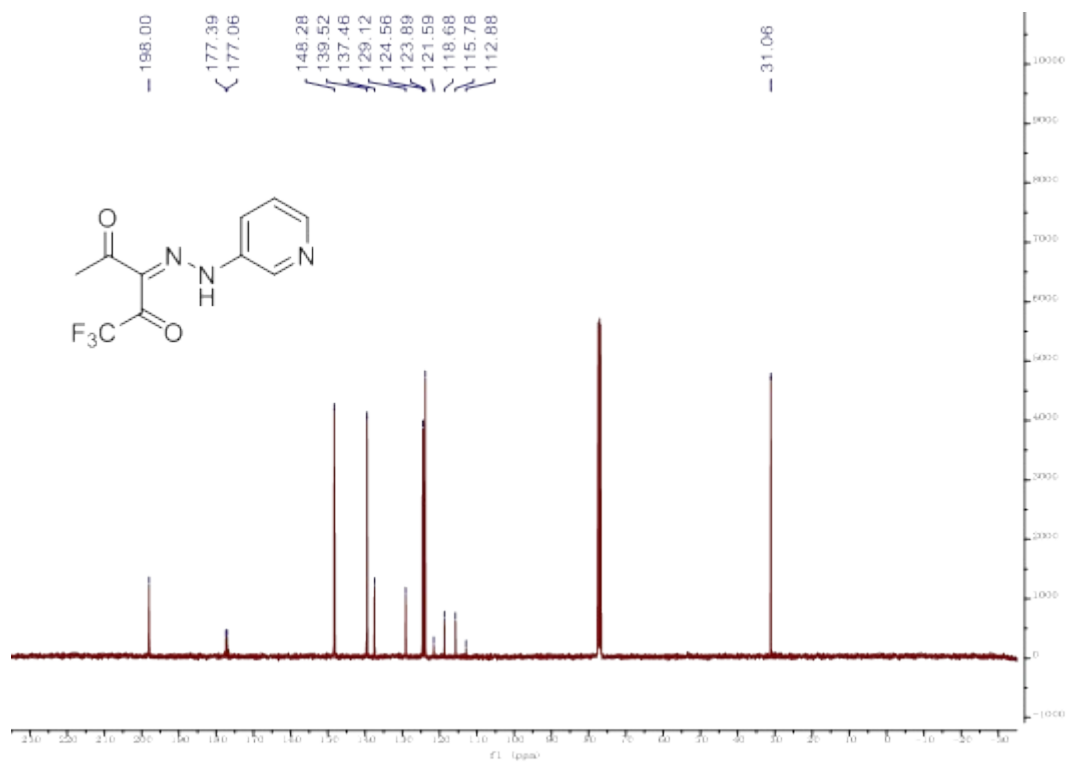


Figure S9 ¹³C NMR spectrum of **B3** (100 MHz, CDCl₃)

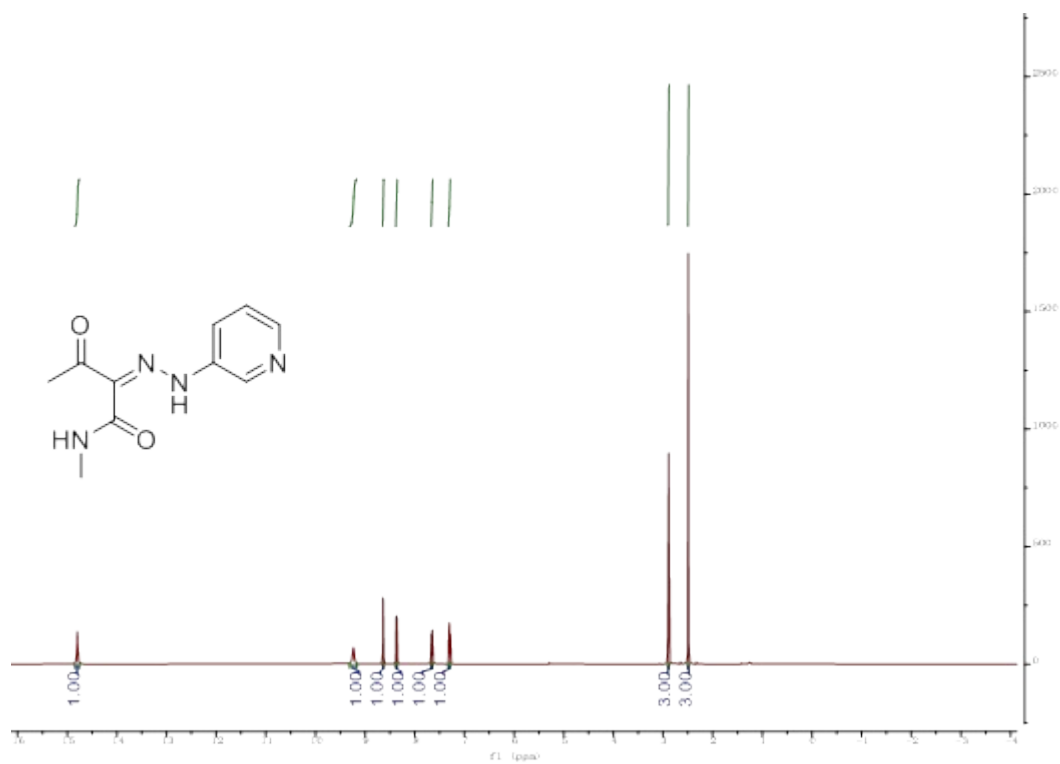


Figure S10 ¹H NMR spectrum of B4 (400 MHz, CDCl₃)

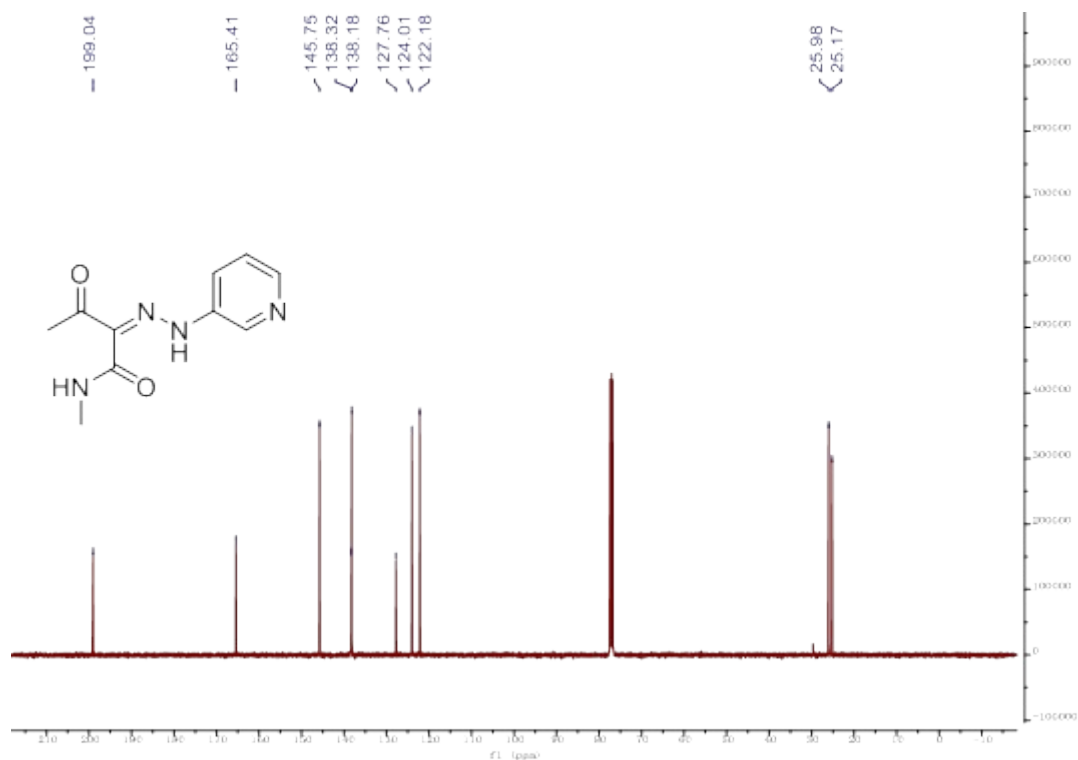


Figure S11 ¹³C NMR spectrum of B4 (100 MHz, CDCl₃)

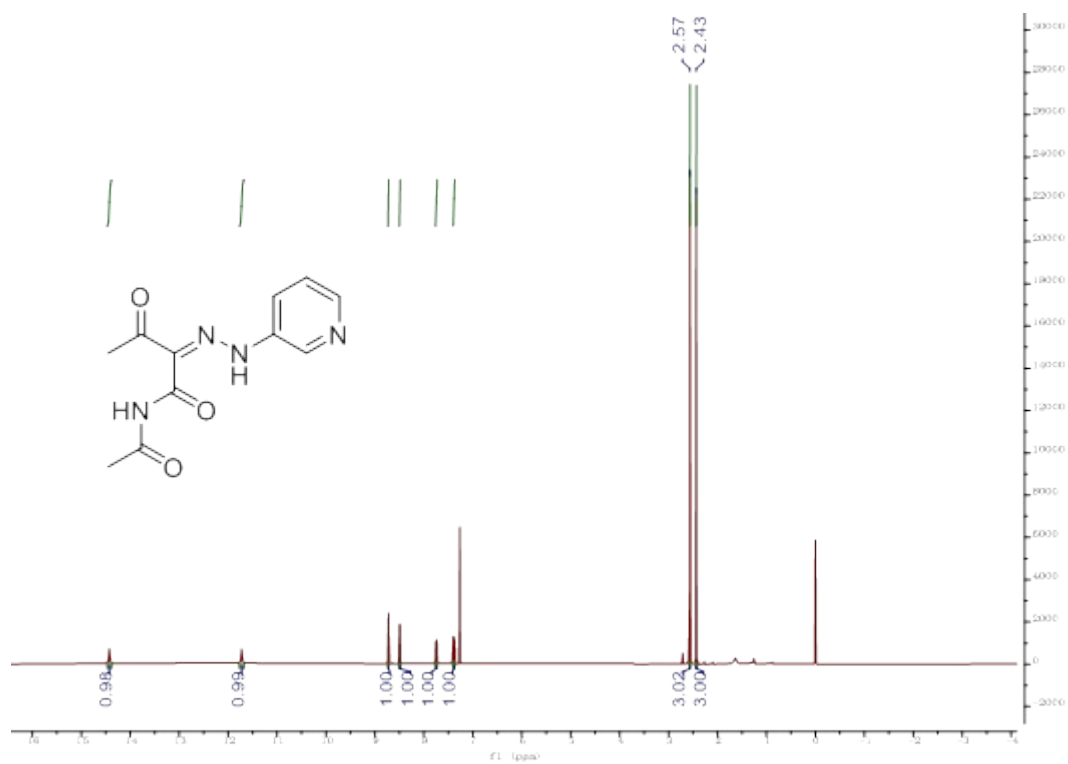


Figure S12 ¹H NMR spectrum of **B5** (400 MHz, CDCl₃)

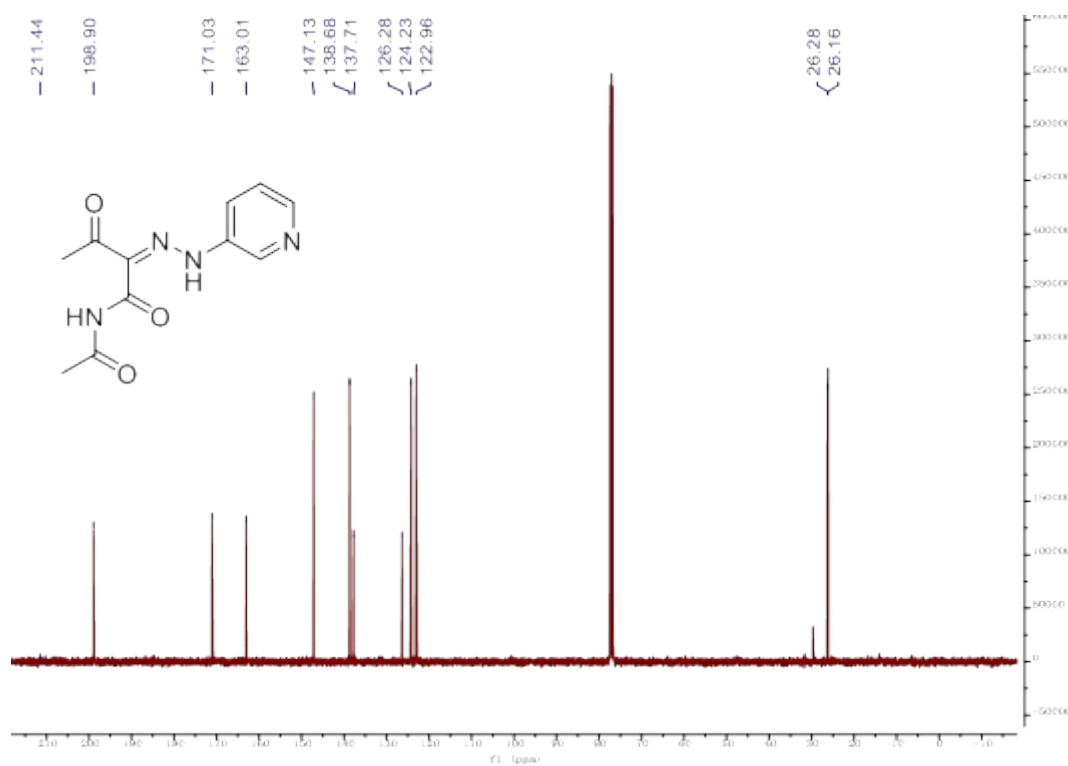


Figure S13 ¹³C NMR spectrum of **B5** (100 MHz, CDCl₃)