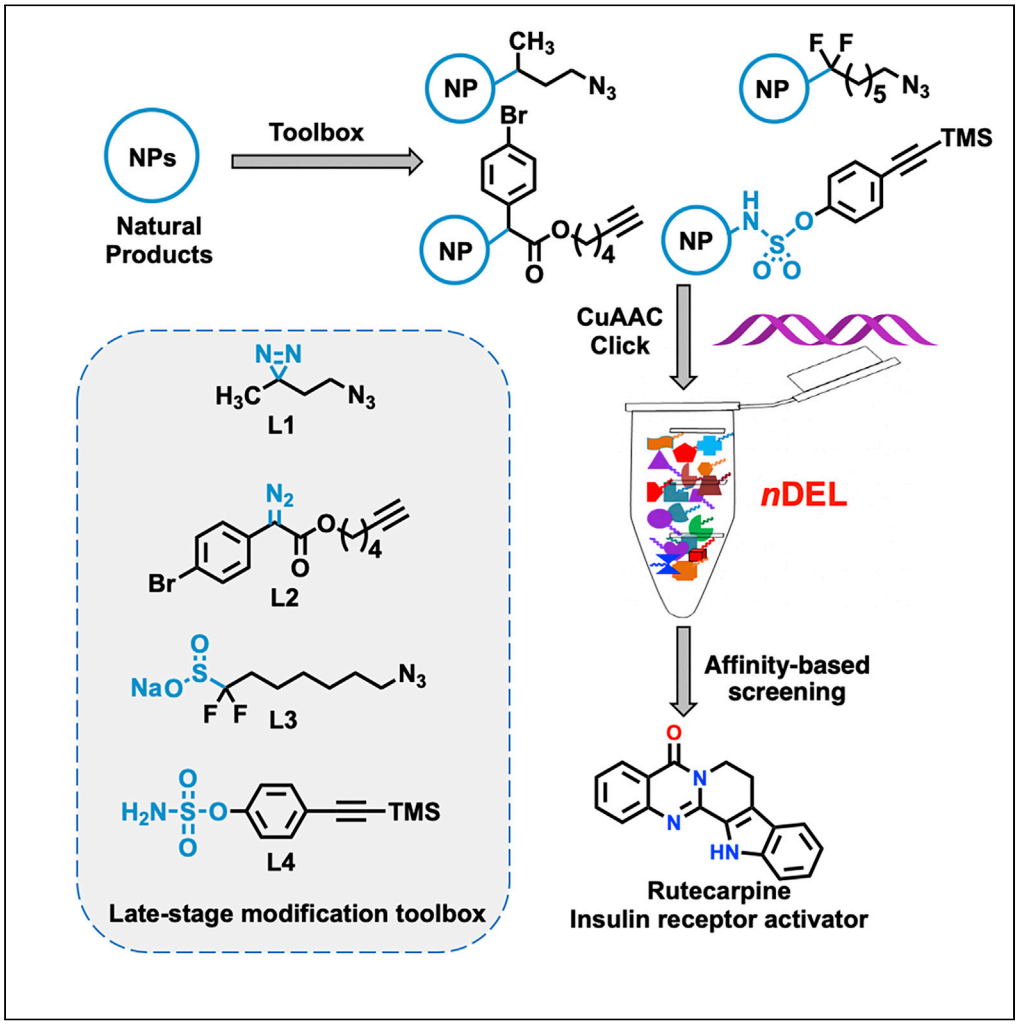
iScience

CellPress
OPEN ACCESS

Article

Selection of Small Molecules that Bind to and Activate the Insulin Receptor from a DNA-Encoded Library of Natural Products



Jia Xie, Shuyue Wang, Peixiang Ma, ..., Hongtao Xu, Guang Yang, Richard A. Lerner

xuht@shanghaitech.edu.cn (H.X.)
yangguang@shanghaitech.edu.cn (G.Y.)
rlerner@scripps.edu (R.A.L.)

HIGHLIGHTS

Annotation of natural products via complementary bifunctional linkers

Function-guided DEL selection using the natural ligand for competitive elution

Identification of Rutaecarpine as a binder and activator of insulin receptor

Xie et al., iScience *23*, 101197
June 26, 2020
https://doi.org/10.1016/j.isci.2020.101197

iScience

CellPress
OPEN ACCESS

Article

Selection of Small Molecules that Bind to and Activate the Insulin Receptor from a DNA-Encoded Library of Natural Products

Jia Xie,[1,7] Shuyue Wang,[2,3,4,5,7] Peixiang Ma,[2,7] Fei Ma,[1] Jie Li,[2,3,4,5] Wei Wang,[2] Fengping Lu,[2] Huan Xiong,[2] Yuang Gu,[2,3,4,5] Shuning Zhang,[2,3,4,5] Hongtao Xu,[2,*] Guang Yang,[2,*] and Richard A. Lerner[1,6,*]

SUMMARY

Although insulin is a life-saving medicine, administration by daily injection remains problematic. Our goal was to exploit the power of DNA-encoded libraries to identify molecules with insulin-like activity but with the potential to be developed as oral drugs. Our strategy involved using a 10^4-member DNA-encoded library containing 160 Traditional Chinese Medicines (nDEL) to identify molecules that bind to and activate the insulin receptor. Importantly, we used the natural ligand, insulin, to liberate bound molecules. Using this selection method on our relatively small, but highly diverse, nDEL yielded a molecule capable of both binding to and activating the insulin receptor. Chemical analysis showed this molecule to be a polycyclic analog of the guanidine metformin, a known drug used to treat diabetes. By using our protocol with other, even larger, DELs we can expect to identify additional organic molecules capable of binding to and activating the insulin receptor.

INTRODUCTION

Since the seminal report of DNA-encoded libraries (DELs) in 1992 (Brenner and Lerner, 1992), academic laboratories, large pharmaceutical corporations, and biotechnology companies have all contributed to great advances in the technology, both in terms of library construction and selection methodology (Chan et al., 2015; Goodnow, 2018; Goodnow et al., 2017; Kleiner et al., 2011; Neri and Lerner, 2018; Scheuermann and Neri, 2015; Shi et al., 2017; Zhao et al., 2019; Zimmermann and Neri, 2016; Zambaldo et al., 2015). This progress has had a profound influence in the pharmaceutical industry by accelerating hit identification in drug discovery, sometimes for heretofore "un-druggable" targets. The popularity of DELs has also led to the development of new chemistries for DNA-compatible reactions, such as diversity-orientated synthesis (DOS) (Christopher et al., 2019), metal- or non-metal-mediated diverse synthesis (Wang et al., 2019; Xiong et al., 2020; Xu et al., 2019, 2020) and late-stage DNA annotation (Ma et al., 2019), all in an effort to expand the chemical space coverage of DELs into more complex molecular structures. Currently, the diversity of DELs contains not only an unprecedented number of simple chemicals but also a collection of highly sophisticated stereo and spatial structures.

The ability to obtain functional molecules is arguably dependent upon the number and diversity of binding molecules in the initial DEL. DEL may be the only method that allows for simultaneous selection of many members of a library on the basis of affinity alone. Late-stage DNA encoding of natural products has been shown to yield selectable libraries with small numbers that are nevertheless rich in structural diversity (Ma et al., 2019). Such encoding technology could be applied to other encoding molecules, such as peptide nucleic acids (PNAs) (Daguer et al., 2011). Traditional Chinese medicine (TCM) molecules are a family of natural products that are usually obtained from plant sources and have been used in medical therapies for more than 3,000 years. Although TCMs have been shown to be effective, the fact that their targets are generally unknown poses a significant challenge for pharmaceutical companies wishing to further improve their efficacy. Nevertheless, over the course of natural evolution these molecules have evolved highly diverse and complicated chemical scaffolds. In an effort to generate a small molecule with insulin-like activity we selected for study 160 TCMs, solely on their availability in pure form. These were encoded with DNA and then incorporated into a relatively small DEL.

Owing to its essential role in glucose homeostasis, insulin signaling has been extensively studied from the viewpoints of both the nature of receptor/ligand structures (Ebina et al., 1985; Gutmann et al., 2018;

[1]Department of Chemistry, Scripps Research Institute, La Jolla, CA 92037, USA

[2]Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China

[3]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

[4]Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

[5]University of Chinese Academy of Sciences, Beijing 100049, China

[6]Lead Contact

[7]These authors contributed equally

*Correspondence: xuht@shanghaitech.edu.cn (H.X.), yangguang@shanghaitech.edu.cn (G.Y.), rlerner@scripps.edu (R.A.L.)

https://doi.org/10.1016/j.isci.2020.101197

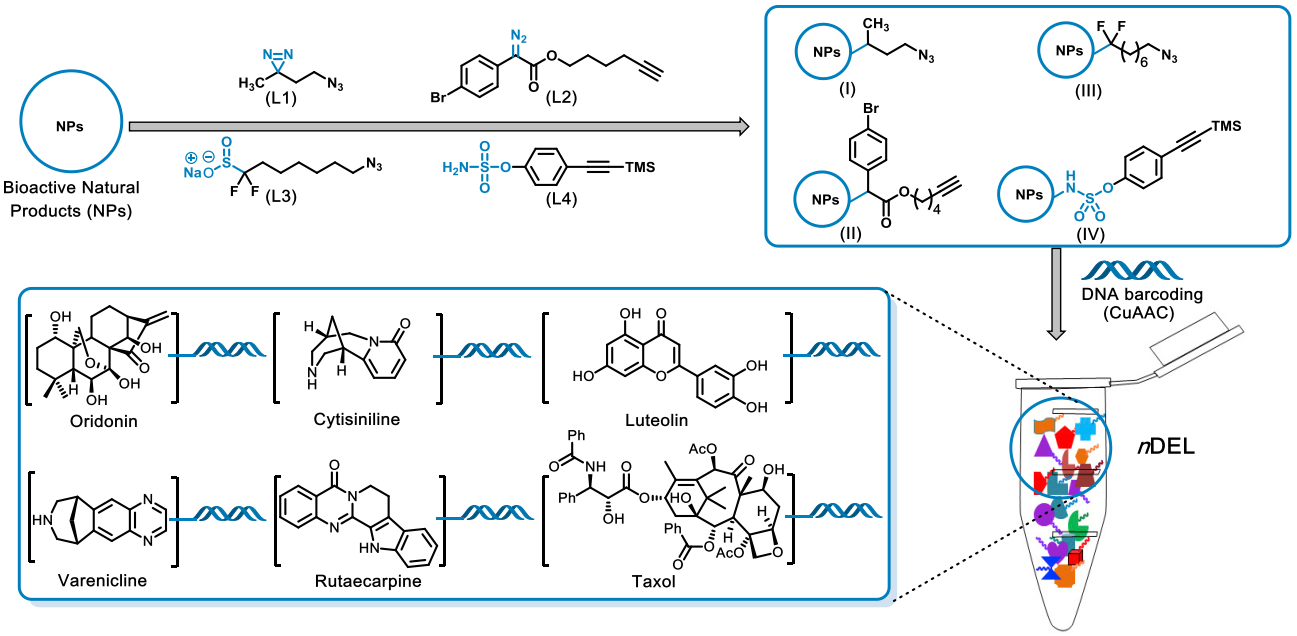


Figure 1. Late-Stage Modification Toolbox to Annotate Natural Products with DNA

Chemical structures highlighted in the frame are the representative DNA-encoded natural products in *n*DEL.

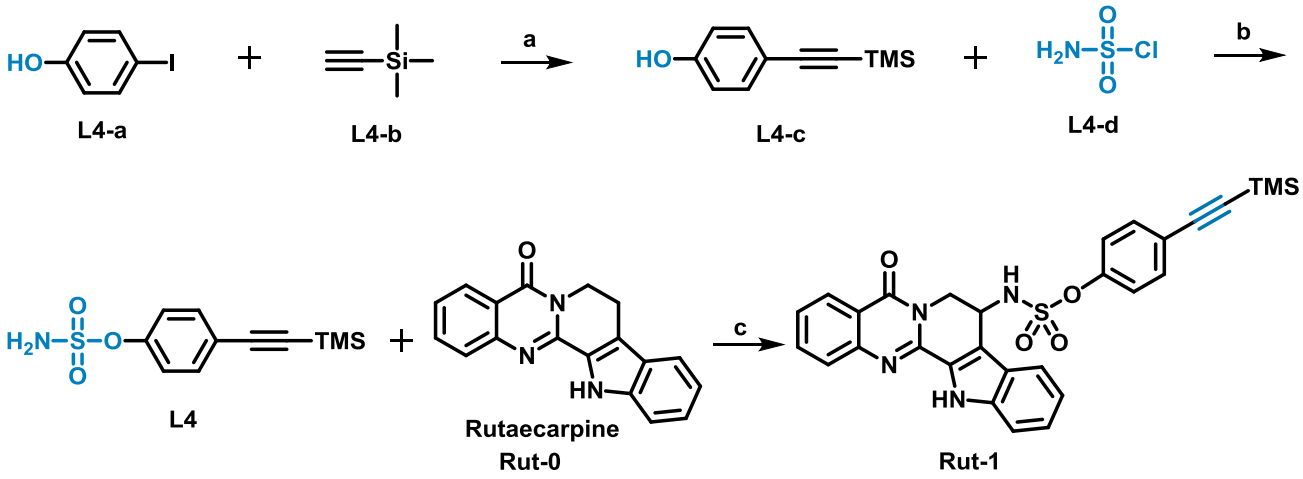
McKern et al., 2006; Smith et al., 2010; Ward and Lawrence, 2009) and the activation mechanisms (De Meyts et al., 1973; Kiselyov et al., 2009; Malaguarnera et al., 2012). The insulin receptor (IR) belongs to a large family of transmembrane tyrosine kinase receptors that is activated by insulin, insulin-like growth factor 1 (IGF-I), and insulin-like growth factor 2 (IGF-II) (Belfiore et al., 2009). These endogenous ligands display a negative cooperative kinetics when binding to and activating the IR (De Meyts et al., 1973). Insulin has been shown to activate the IR through stabilizing the active conformation of the covalent IR dimer, which is pre-formed in a symmetrical and anti-parallel arrangement in the cell membrane, by cross-linking the two binding sites (site 1 and site 2) on the ectodomain of each subunit (Benyoucef et al., 2007; De Meyts and Whittaker, 2002; Fabry et al., 1992; Hao et al., 2006; Huang et al., 2004; Kristensen et al., 2002; Kurose et al., 1994; Mynarcik et al., 1996; Wedekind et al., 1989; Whittaker et al., 2008; Williams et al., 1995). In spite of our detailed understanding of the mechanism by which the IR is activated, attempts to identify orally available small molecule insulin mimetic have met with little success (Garcia-Vicente et al., 2007; Qiang et al., 2014; Tsai and Chou, 2009; Wilkie et al., 2001). A key limiting factor, in addition to issues of toxic side effects and solubility, has been the lack of structural diversity among candidate molecules.

Herein we used a structurally diverse DEL containing 160 TCMs to select a small molecule that binds to and activates the IR. More importantly, we showed that this activation results in insulin-like activity in a cell-based system. In a remarkable coincidence, the selected molecule is a TCM that has been shown to lower blood glucose in murine diabetes models (Chen et al., 2013; Nie et al., 2010, 2016). These findings should encourage attempts to develop a small molecule replacement for insulin. Such a molecule would offer significant advantages in diabetes therapy, as it would be orally available and have an easily standardized dosage.

RESULTS AND DISCUSSION

Late-Stage Modification Toolbox for DNA Encoding of Natural Products

In our previous work, we described a volatile bifunctional linker that could quantitatively annotate a complex organic molecule at any particular site with a DNA barcode sequence (Ma et al., 2019). However, for chemicals lacking functional groups such as amines, hydroxyl groups, and carboxylic acids, this volatile linker alone was not sufficient owing to poor labeling efficiency. To increase the site diversity for the labeling of natural products with DNA barcodes, we expanded the toolbox by developing several new bifunctional linkers with complementary reactivity. As shown in Figure 1, these included carbine precursors (**L1** and **L2**), a radical precursor (**L3**), and a nitrene precursor (**L4**). These linkers can be used to add azide or



Scheme 1. Late-Stage Modification of Rutaecarpine

Reagents and conditions: (a) $Pd(Ph_3P)_2Cl_2$, CuI, Et_3N, DMF, 80°C, 72%; (b) DIPEA, DCM, 63%; (c) $Rh_2(esp)_2$, $PhI(OAc)_2$, CH_3CN, 23%.

alkyne functional groups on to natural products, so that the copper-catalyzed azide-alkyne cycloadditions (CuAAC) (Kolb et al., 2001) can proceed under DNA-compatible reaction conditions. Thus, these chemistries are suitable for efficient late-stage DNA annotation of diverse natural products. The diazoacetate containing linker **L2** introduces alkynes through C-H or C-X (X = OH, or NH_2) insertion (He et al., 2015; Peddibhotla et al., 2007), whereas the difluoroalkyl-sulfinate bifunctional linker **L3** has been shown to directly functionalize the C-H of (hetero)arenes (Zhou et al., 2013). Using sulfonamide as a nitrene precursor (Lu et al., 2018), **L4** was expected to undergo intermolecular C(sp^3)-H amination reactions to add an alkyne group on to natural products. As shown in Scheme 1, the synthesis of **L4** was commenced from Sonogashira cross-coupling of 4-iodophenol **(L4-a)** and ethynyltrimethylsilane (**L4-b)** to afford **L4-c**, which was subsequently reacted with sulfamoyl chloride in the presence of *N,N*-diisopropylethylamine to give the desired **L4** linker. Next, Rutaecarpine (**Rut-0**) was functionalized by **L4** using bis[rhodium (a,a,a′,a′-tetramethyl-1,3-benzene dipropionate)] [$Rh_2(esp)_2$] as the catalyst (Espino et al., 2004) and $PhI(OAc)_2$ as the oxidant to afford the desired amination product **Rut-1** in 23% yield. Using this expanded toolbox, a total of 160 TCMs were annotated and included in the *n*DEL library (Table S1).

Affinity-Based Selection of *n*DEL Members that Bind to the Insulin Receptor

The *n*DEL screening was carried out using the purified recombinant extracellular domain of the human insulin receptor (ECD-*h*IR), which was immobilized on either cobalt beads by the C-terminal polyhistidine tag or streptavidin beads through biotin modification. The bound *n*DEL molecules were collected using conventional heat denaturation. The screening fingerprint of the *n*DEL was plotted as enrichment fold versus normalized sequencing counts as shown in Figure 2. Compared with the negative control, the enrichment pattern was similar in both cobalt bead-based screening and streptavidin bead-based screening, indicative of strong non-specific interactions. The maximum enrichment was also similar, with a 54-fold enrichment using empty cobalt beads versus a 66-fold enrichment with his-tag-insulin receptor-attached cobalt beads, and a 185-fold enrichment using empty streptavidin beads versus a 121-fold enrichment with biotinylated insulin receptor on streptavidin beads. Thus, a more stringent screening strategy was necessary.

Insulin Elution of *n*DEL Members Bound to the Insulin Receptor

As demonstrated above, it is likely that heat denaturation will liberate any compound whose binding depends on a particular secondary or tertiary structure. Since insulin binds to and activates the IR, we reasoned that elution with insulin should liberate compounds with more functional binding to the IR. The IR is a complex dimeric protein consisting of two identical extracellular α subunits each containing two sites that bind to insulin with different affinity (Goldfine, 1987; Moller and Flier, 1991). The binding of the intrinsic ligand, insulin, follows negative cooperative kinetics (De Meyts et al., 1973) suggesting an optimal concentration of insulin that preferentially binds to the active form of the IR. Using the streptavidin beads, the optimal activation concentration of insulin (100 nM) was used to competitively elute bound *n*DEL members. As expected, insulin elution generated a significantly enriched fingerprint pattern relative to the

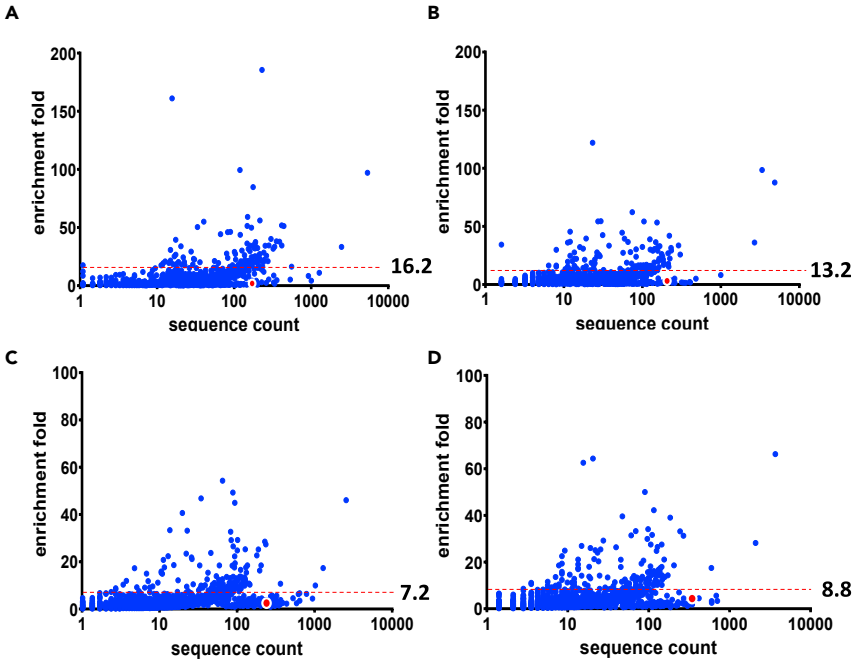


Figure 2. Fingerprint Patterns for *n*DEL Screening of the Insulin Receptor

Fingerprints for *n*DEL selection against streptavidin-beads (A), biotinylated insulin receptor attached to streptavidin-beads (B), cobalt-based beads (C), and his-tag insulin receptor attached to cobalt-based beads (D). Red dashed lines are the cut-offs for hits selection.

negative control (Figure 3). More interestingly, a TCM, Rutaecarpine, was shown to be enriched over 50-fold when eluted with insulin, as compared with less than 5-fold enrichment when eluted with heat denaturation (Table 1), indicating possible specific and functional binding to the IR.

Rutaecarpine Is an Analog of Metformin

Metformin (CSD-JAMRIY01) (Childs et al., 2004) , an oral diabetes medicine, showed striking structural similarity to Rutaecarpine (CSD-OGAXEC)Qin et al., 2018 , in that the arrangement of the five nitrogen hetero atoms of the bis-guanidine appears to be the same in both molecules (Figure 4A).

Both Rutaecarpine and Metformin Bind to the Insulin Receptor

In order to understand the mode of action of Rutaecarpine, we tested if Rutaecarpine directly targets insulin receptor *in vitro* using surface plasmon resonance technology. Rutaecarpine exhibited moderate binding affinity to the IR extracellular domain (amino acids 1–956) with an estimated dissociation constant (K_D) of 14 μM, and indeed, Metformin also bound to the insulin receptor but with a weaker K_D value of 84 μM as shown in Figures 4B–4E.

To determine if the binding of Rutaecarpine triggers conformation change in the IR, we carried out a partial proteolytic digestion of the IR. The human recombinant insulin receptor extracellular domain (ECD-*h*IR) was incubated with 5% DMSO, or 50 μM Rutaecarpine in 5% DMSO, followed by limited trypsin digestion. As shown in Figure S1, SDS-PAGE analysis showed that the pattern of trypsin digestion was altered in the presence of Rutaecarpine. A digestion ladder appeared below the main ECD-*h*IR band, indicating that Rutaecarpine bound to the ECD domain of the IR, changed its confirmation, and facilitated the enzyme digestion. In contrast, Rutaecarpine did not affect the trypsin cleavage pattern of an irrelevant protein, BSA, thus ruling out the possibility that Rutaecarpine is a protease enhancer.

Rutaecarpine and Metformin Activated Autophosphorylation of the Insulin Receptor on Cells

The IR exists on the cell membrane as a homodimer consisting of two extracellular α subunits that bind insulin, as well as two transmembrane β subunits that have intracellular tyrosine kinase activity (Goldfine,

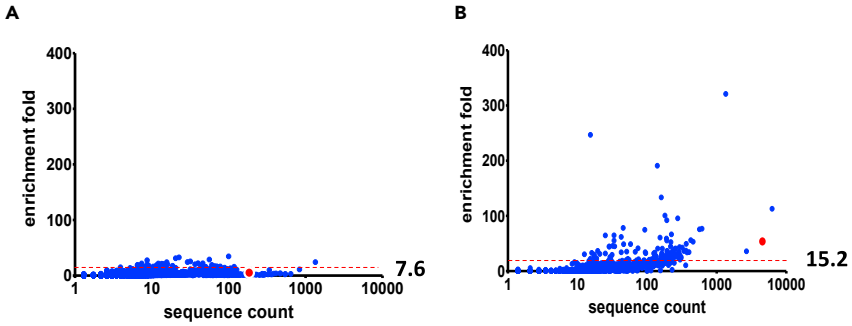


Figure 3. Fingerprint Patterns of *n*DEL Screening of the Insulin Receptor Using Insulin as the Competitive Eluent

Fingerprints for *n*DEL selection using 100 nM insulin elution against streptavidin-beads (A), biotinylated insulin receptor attached to streptavidin-beads (B). Red dashed lines are the cut-offs for hits selection. Red dots represent the DNA-conjugated TCM hit, Rutaecarpine.

1987; Moller and Flier, 1991). When insulin binds to the α subunit of the receptor, the β subunit tyrosine kinase is activated, resulting in autophosphorylation of β subunit tyrosine residues. This autophosphorylation is considered a hallmark of IR activation and, in turn, activates the downstream signaling. CHO cells expressing the human IR were treated with 10 μM Rutaecarpine and Metformin. Five percent DMSO and 100 nM insulin were included as negative and positive controls, respectively. Rutaecarpine and Metformin stimulated significant autophosphorylation of the human insulin receptor as expected (Figure 5). It was noted that Metformin also showed sensitization effect of insulin on the activation of IR, consistent with the literature reports (Meuillet et al., 1999; Tadayyon and Smith, 2003; Kumar and Dey, 2002).

Analysis revealed this compound to be Rutaecarpine, a TCM obtained from the plant *Evodia rutaecarpa*, also called "吴茱萸 (Wu Zhu Yu)" (Moon et al., 1999). "吴茱萸" is a medical herb described in "Divine Farmer's Classic of Materia Medica," the oldest medical book in China during the period of Qin and Han dynasties more than 2,000 years ago. The herb is slightly toxic and has been used to treat respiratory infections, inflammation, pain, hypertension, and diarrhea. It is one of the two active ingredients in the famous medical formulation "Zuojin Pills" developed by Dr. Zhu, Danxi (朱丹溪) during the Yuan dynasty, more than 800 years ago. To this day, in China, this formulation is still being used to treat gastritis and hypertension. Modern research on Rutaecarpine of "Wu Zhu Yu" in animals showed strong glucose regulation and insulin sensitization effects (Nie et al., 2010, 2016; Yeo et al., 2011; Wei, 2008). In spite of its long history, Rutaecarpine's mechanism of action remains unknown. Our structural studies have shown Rutaecarpine to be an analog of the important diabetes drug Metformin, in that all the N hetero atoms are in a similar position relative to each other and "locked down" by rings. Until now the mechanism by which Metformin functions was unknown. Most mechanistic concepts center around the role of Metformin in the activation of AMP-activated protein kinase (AMPK) (Zhou et al., 2001). Our findings indicate that Metformin may bind to and activate the IR, which should allow a better understanding of its mechanism of action and, in turn, could lead to synthesis of better analogs of this important drug. For instance, the binding of Metformin to the IR could be responsible for the known sensitization effect of Metformin on the insulin signaling both in cells and in human (Meuillet et al., 1999; Tadayyon and Smith, 2003; Kumar and Dey, 2002).

One may ask why selection from a small, albeit diverse, DEL should have yielded a molecule with such profound metabolic effects. Although it was not known that any of the DEL components, including Rutaecarpine, would bind the IR, using the IR as the target for binding could be expected to bias the selection process toward molecules having metabolic effects. In this case binding of Rutaecarpine to the IR was

	Heat Denaturation Elution		Insulin Elution
	Cobalt-Based Beads	Streptavidin Beads	Streptavidin Beads
Control	2.95	1.42	2.14
Sample	4.47	2.33	51.5

Table 1. Enrichment of DNA-encoded Rutaecarpine under Different Selection Conditions.

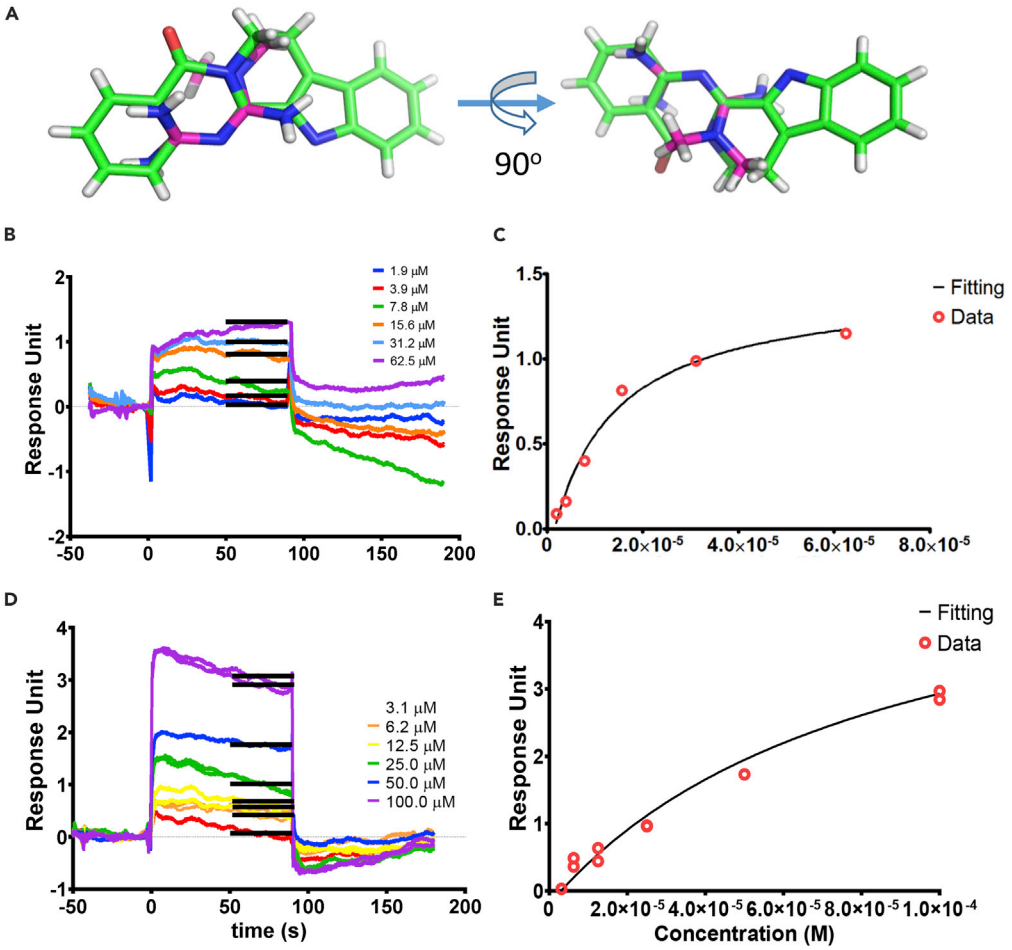


Figure 4. Rutaecarpine and Metformin Directly Bind to the Insulin Receptor

(A) Overlay of crystal structure of Rutaecarpine (green, CSD-OGAXEC) and Metformin (pink, CSD-JAMRIY01), in which oxygen (red), nitrogen (blue), and hydrogen (white) atoms are labeled with different colors. X-ray structures are obtained from the Cambridge Crystallographic Data Center (CCDC) (Groom et al., 2016); (B–E) Rutaecarpine and Metformin bind to the insulin receptor extracellular domain as quantified by surface plasmon resonance measurements (B and D). Graphs of equilibrium response unit versus compound concentrations are plotted (C and E). The estimated K_D is 14 μM for Rutaecarpine and 84 μM for Metformin.

highly specific, as Rutaecarpine was not selected when other proteins were used as the target. The IR itself is a highly evolved protein that is carefully integrated into the plasma membrane and as such may not tolerate perturbation without becoming activated. The binding of hydrophobic organic compounds could easily be a major perturbation. Perhaps membrane receptors that are poised to signal by changing their conformation may be especially susceptible to activation by binding hydrophobic organic molecules. Thus, it is possible that the only requirement to achieve results similar to those reported here is to choose the right receptor target for study and to pair it with its natural ligand for elution (i.e., in this paper *n*DEL members binding to the IR were eluted with insulin). This insight, together with the concept that one needs initially select only for binding, should guide further investigations.

Conclusion

In summary, we developed a tool box of complementary bifunctional linkers that contain carbine precursors, radical precursor, nitrene precursor, and azide or alkyne functional groups to expand the chemical space of DNA-encoded library. By using the natural ligand, insulin, for competitive elution, we discovered a polycyclic nature product named Rutaecarpine, which is capable of both binding to and activating the IR. The experiments reported here rely on the concept that one of the main strengths of DELs is that they allow

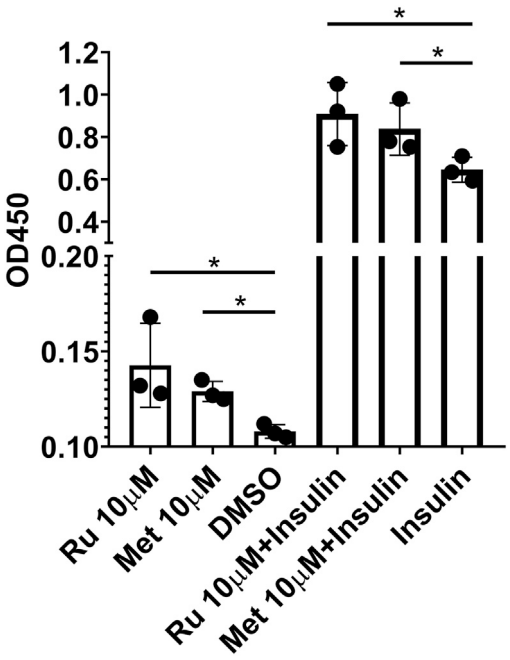


Figure 5. Effects of Rutaecarpine on the Autophosphorylation of the Insulin Receptor.

Comparison of autophosphorylation of the insulin receptor in CHO-*h*IR cells, in which10 μM Rutaecarpine (Ru), 10 μM Metformin (Met), 100 nM Insulin, and 5% DMSO dissolved in media (DMSO) were incubated with the CHO-*h*IR cells. The y axis represents the absorbance at 450 nm developed by the specific antibody targeting phosphorylated insulin receptor. The error bar was determined from the data of three independent experiments. Statistical analysis between two groups was performed using unpaired Student's t test. (*$p < 0.05$, $n = 3$)

for selection based only on binding, with the assumption that once such molecules are found, some of them will be functional. In this case, selection using a relatively small but highly diverse DEL will yield a compound capable of both binding to and activating the target.

Limitations of the Study

This is a highly diverse and small natural product-enriched DNA-encode library (nDEL) that designed for the concept proof; the number of natural products in this nDEL should keep increasing to cover larger chemical space.

Resource Availability

Lead Contact

rlerner@scripps.edu.

Materials Availability

All the materials necessary to reproduce this study are included in the manuscript and Supplemental Information.

Data and Code Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101197.

ACKNOWLEDGMENTS

This work is supported by a grant from JPB Foundation; National Natural Science Foundation of China Grants 21502114, 21977070, U19A2011, and 31500632; and Science and Technology Commission of Shanghai Municipality Grant 16DZ1910200. We thank the biomedical big data platform, analytical

platform, and high-throughput screening platform at Shanghai Institute for Advanced Immunochemical Studies (SIAIS) at ShanghaiTech University for the support of deep-sequencing analyses, mass spectrometry, and flow cytometry experiments.

AUTHOR CONTRIBUTIONS

J.X., S.W., and P.M. planned and carried out most of the experiments and analyzed and summarized the experiments. F.M., J.L., F.L., H.X., Y.G., and S.Z. synthesized the DNA-encoded natural product library (*n*DEL). W.W. performed informatics analyses. R.A.L., G.Y., H.X., P.M., and J.X. supervised the whole research and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 22, 2019
Revised: February 21, 2020
Accepted: May 21, 2020
Published: June 26, 2020

REFERENCES

Belfiore, A., Frasca, F., Pandini, G., Sciacca, L., and Vigneri, R. (2009). Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. Endocr. Rev. *30*, 586–623.

Benyoucef, S., Surinya, K.H., Hadaschik, D., and Siddle, K. (2007). Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation. Biochem. J. *403*, 603–613.

Brenner, S., and Lerner, R.A. (1992). Encoded combinatorial chemistry. Proc. Natl. Acad. Sci. U S A *89*, 5381–5383.

Chan, A.I., McGregor, L.M., and Liu, D.R. (2015). Novel selection methods for DNA-encoded chemical libraries. Curr. Opin. Chem. Biol. *26*, 55–61.

Chen, Y.C., Zeng, X.Y., He, Y., Liu, H., Wang, B., Zhou, H., Chen, J.W., Liu, P.Q., Gu, L.Q., Ye, J.M., and Huang, Z.S. (2013). Rutaecarpine analogues reduce lipid accumulation in adipocytes via inhibiting adipogenesis/lipogenesis with AMPK activation and UPR suppression. ACS Chem. Biol. *8*, 2301–2311.

Childs, S.L., Chyall, L.J., Dunlap, J.T., Coates, D.A., Stahly, B.C., and Stahly, G.P. (2004). A metastable polymorph of metformin hydrochloride: Isolation and characterization using capillary crystallization and thermal microscopy techniques. Crystal Growth & Design *4*, 441–449.

Christopher, G., Mathias, W., Paul, C., and Stuart, S. (2019). DNA barcoding a complete matrix of stereoisomeric small molecules. J. Am. Chem. Soc. *141*, 10225–10235.

Daguer, J.P., Ciobanu, M., Alvarez, S., Barluenga, S., and Winssinger, N. (2011). DNA-templated combinatorial assembly of small molecule fragments amenable to selection/amplification cycles. Chem. Sci. *2*, 625–632.

De Meyts, P., Roth, J., Neville, D.M., Jr., Gavin, J.R., 3rd, and Lesniak, M.A. (1973). Insulin interactions with its receptors: experimental evidence for negative cooperativity. Biochem. Biophys. Res. Commun. *55*, 154–161.

De Meyts, P., and Whittaker, J. (2002). Structural biology of insulin and IGF1 receptors: implications for drug design. Nat. Rev. Drug Discov. *1*, 769–783.

Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J.H., Masiarz, F., Kan, Y.W., Goldfine, I.D., et al. (1985). The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling. Cell *40*, 747–758.

Espino, C.G., Fiori, K.W., Kim, M., and Du Bois, J. (2004). Expanding the scope of C-H amination through catalyst design. J. Am. Chem. Soc. *126*, 15378–15379.

Fabry, M., Schaefer, E., Ellis, L., Kojro, E., Fahrenholz, F., and Brandenburg, D. (1992). Detection of a new hormone contact site within the insulin receptor ectodomain by the use of a novel photoreactive insulin. J. Biol. Chem. *267*, 8950–8956.

Garcia-Vicente, S., Yraola, F., Marti, L., Gonzalez-Munoz, E., Garcia-Barrado, M.J., Canto, C., Abella, A., Bour, S., Artuch, R., Sierra, C., et al. (2007). Oral insulin-mimetic compounds that act independently of insulin. Diabetes *56*, 486–493.

Goldfine, I.D. (1987). The insulin receptor: molecular biology and transmembrane signaling. Endocr. Rev. *8*, 235–255.

Goodnow, R., Jr. (2018). DNA-encoded library technology (DELT) after a quarter century. SLAS Discov. *23*, 385–386.

Goodnow, R.A., Jr., Dumelin, C.E., and Keefe, A.D. (2017). DNA-encoded chemistry: enabling the deeper sampling of chemical space. Nat. Rev. Drug Discov. *16*, 131–147.

Groom, Colin, R, Bruno, Ian, J, Lightfoot, Matthew, P, and Ward, Suzanna, C (2016). The Cambridge Structural Database. Acta Crystallogr B Struct Sci Cryst Eng Mater. *72*, 171–179, https://doi.org/10.1107/S2052520616003954.

Gutmann, T., Kim, K.H., Grzybek, M., Walz, T., and Coskun, U. (2018). Visualization of ligand-induced transmembrane signaling in the full-length human insulin receptor. J. Cell Biol. *217*, 1643–1649.

Hao, C., Whittaker, L., and Whittaker, J. (2006). Characterization of a second ligand binding site of the insulin receptor. Biochem. Biophys. Res. Commun. *347*, 334–339.

He, J., Hamann, L.G., Davies, H.M., and Beckwith, R.E. (2015). Late-stage C-H functionalization of complex alkaloids and drug molecules via intermolecular rhodium-carbenoid insertion. Nat. Commun. *6*, 5943.

Huang, K., Xu, B., Hu, S.Q., Chu, Y.C., Hua, Q.X., Qu, Y., Li, B., Wang, S., Wang, R.Y., Nakagawa, S.H., et al. (2004). How insulin binds: the B-chain alpha-helix contacts the L1 beta-helix of the insulin receptor. J. Mol. Biol. *341*, 529–550.

Kiselyov, V.V., Versteyhe, S., Gauguin, L., and De Meyts, P. (2009). Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation. Mol. Syst. Biol. *5*, 243.

Kleiner, R.E., Dumelin, C.E., and Liu, D.R. (2011). Small-molecule discovery from DNA-encoded chemical libraries. Chem. Soc. Rev. *40*, 5707–5717.

Kolb, H.C., Finn, M.G., and Sharpless, K.B. (2001). Click chemistry: diverse chemical function from a few good reactions. Angew. Chem. Int. Ed. *40*, 2004–2021.

Kristensen, C., Andersen, A.S., Ostergaard, S., Hansen, P.H., and Brandt, J. (2002). Functional reconstitution of insulin receptor binding site from non-binding receptor fragments. J. Biol. Chem. *277*, 18340–18345.

Kumar, N., and Dey, C.S. (2002). Metformin enhances insulin signalling in insulin-dependent and-independent pathways in insulin resistant muscle cells. Br. J. Pharmacol. *137*, 329–336.

Kurose, T., Pashmforoush, M., Yoshimasa, Y., Carroll, R., Schwartz, G.P., Burke, G.T., Katsoyannis, P.G., and Steiner, D.F. (1994). Cross-linking of a B25 azidophenylalanine insulin derivative to the carboxyl-terminal region of the alpha-subunit of the insulin receptor. Identification of a new insulin-binding domain in the insulin receptor. J. Biol. Chem. *269*, 29190–29197.

Lu, X.B., Shi, Y.F., and Zhong, F.R. (2018). Rhodium-catalyzed intermolecular C(sp(3))-H amination in a purely aqueous system. Green. Chem. *20*, 113–117.

Ma, P., Xu, H., Li, J., Lu, F., Ma, F., Wang, S., Xiong, H., Wang, W., Buratto, D., Zonta, F., et al. (2019). Functionality-independent DNA encoding of complex natural products. Angew. Chem. Int. Ed. *58*, 9254–9261.

Malaguarnera, R., Sacco, A., Voci, C., Pandini, G., Vigneri, R., and Belfiore, A. (2012). Proinsulin binds with high affinity the insulin receptor isoform A and predominantly activates the mitogenic pathway. Endocrinology *153*, 2152–2163.

McKern, N.M., Lawrence, M.C., Streltsov, V.A., Lou, M.Z., Adams, T.E., Lovrecz, G.O., Elleman, T.C., Richards, K.M., Bentley, J.D., Pilling, P.A., and Hoyne, P.A. (2006). Structure of the insulin receptor ectodomain reveals a folded-over conformation. Nature *443*, 218–221.

Meuillet, E.J., Wiernsperger, N., Mania-Farnell, B., Hubert, P., and Cremel, G. (1999). Metformin modulates insulin receptor signaling in normal and cholesterol-treated human hepatoma cells (HepG2). Eur. J. Pharmacol. *377*, 241–252.

Moller, D.E., and Flier, J.S. (1991). Insulin resistance mechanisms, syndromes, and implications. N. Engl. J. Med. *325*, 938–948.

Moon, T.C., Murakami, m., Kudo, I., Son, K.H., Kim, H.P., Kang, S.S., and Chang, H.W. (1999). A new class of COX-2 inhibitor, rutaecarpine from *Evodia rutaecarpa*. Inflamm. Res. *48*, 621–625.

Mynarcik, D.C., Yu, G.Q., and Whittaker, J. (1996). Alanine-scanning mutagenesis of a C-terminal ligand binding domain of the insulin receptor alpha subunit. J. Biol. Chem. *271*, 2439–2442.

Neri, D., and Lerner, R.A. (2018). DNA-encoded chemical libraries: a selection system based on endowing organic compounds with amplifiable information. Annu. Rev. Biochem. *87*, 479–502.

Nie, X., Yu, L., Chen, H., Zhao, T., and Bian, K. (2010). Intervention effects of rutaecarpine in type 2 diabetic obese rats. Chin. Pharmacol. Bull. *26*, 872–876.

Nie, X.Q., Chen, H.H., Zhang, J.Y., Zhang, Y.J., Yang, J.W., Pan, H.J., Song, W.X., Murad, F., He, Y.Q., and Bian, K. (2016). Rutaecarpine ameliorates hyperlipidemia and hyperglycemia in fat-fed, streptozotocin-treated rats via regulating the IRS-1/PI3K/Akt and AMPK/ACC2 signaling pathways. Acta Pharmacol. Sin *37*, 483–496.

Peddibhotla, S., Dang, Y., Liu, J.O., and Romo, D. (2007). Simultaneous arming and structure/activity studies of natural products employing O-H insertions: an expedient and versatile strategy for natural products-based chemical genetics. J. Am. Chem. Soc. *129*, 12222–12231.

Qiang, G., Xue, S., Yang, J.J., Du, G., Pang, X., Li, X., Goswami, D., Griffin, P.R., Ortlund, E.A., Chan, C.B., and Ye, K. (2014). Identification of a small molecular insulin receptor agonist with potent antidiabetes activity. Diabetes *63*, 1394–1409.

Qin, Qi, Pin, Zou, Bi, Qun, Hu, Fei, Hong, Huang, Guo, Bao, Wang, Shu, Long, Gu, Yun, Qing, and Tan, Ming, Xiong (2018). Platinum(ii) complexes with rutaecarpine and tryptanthrin derivatives induce apoptosis by inhibiting telomerase activity and disrupting mitochondrial function. Medchemcomm *9*, 1639–1648, https://doi.org/10.1039/c8md00247a.

Scheuermann, J., and Neri, D. (2015). Dual-pharmacophore DNA-encoded chemical libraries. Curr. Opin. Chem. Biol. *26*, 99–103.

Shi, B., Zhou, Y., Huang, Y., Zhang, J., and Li, X. (2017). Recent advances on the encoding and selection methods of DNA-encoded chemical library. Bioorg. Med. Chem. Lett. *27*, 361–369.

Smith, B.J., Huang, K., Kong, G., Chan, S.J., Nakagawa, S., Menting, J.G., Hu, S.Q., Whittaker, J., Steiner, D.F., Katsoyannis, P.G., et al. (2010). Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists. Proc. Natl. Acad. Sci. U S A *107*, 6771–6776.

Tadayyon, M., and Smith, S.A. (2003). Insulin sensitisation in the treatment of Type 2 diabetes. Expert Opin. Investig. Drugs *12*, 307–324.

Tsai, H.J., and Chou, S.Y. (2009). A novel hydroxyfuroic acid compound as an insulin receptor activator. Structure and activity relationship of a prenylindole moiety to insulin receptor activation. J. Biomed. Sci. *16*, 68.

Wang, D.-Y., Wen, X., Xiong, C.-D., Zhao, J.-N., Ding, C.-Y., Meng, Q., Zhou, H., Wang, C., Uchiyama, M., Lu, X.-J., and Zhang, A. (2019). Non-transition metal-mediated diverse aryl–heteroatom bond formation of Arylammonium salts. iScience *15*, 307–315.

Ward, C.W., and Lawrence, M.C. (2009). Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor. Bioessays *31*, 422–434.

Wedekind, F., Baer-Pontzen, K., Bala-Mohan, S., Choli, D., Zahn, H., and Brandenburg, D. (1989). Hormone binding site of the insulin receptor: analysis using photoaffinity-mediated avidin complexing. Biol. Chem. Hoppe Seyler *370*, 251–258.

Wei, Z.X. (2008). Experiences in treating diabetic peripheral neuropathy with traditional Chinese medicine. Chin. J. Integr. Med. *14*, 248–250.

Whittaker, L., Hao, C., Fu, W., and Whittaker, J. (2008). High-affinity insulin binding: insulin interacts with two receptor ligand binding sites. Biochemistry *47*, 12900–12909.

Wilkie, N., Wingrove, P.B., Bilsland, J.G., Young, L., Harper, S.J., Hefti, F., Ellis, S., and Pollack, S.J. (2001). The non-peptidyl fungal metabolite L-783,281 activates TRK neurotrophin receptors. J. Neurochem. *78*, 1135–1145.

Williams, P.F., Mynarcik, D.C., Yu, G.Q., and Whittaker, J. (1995). Mapping of an NH2-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis. J. Biol. Chem. *270*, 3012–3016.

Xiong, H., Gu, Y., Zhang, S., Lu, F., Ji, Q., Liu, L., Ma, P., Yang, G., Hou, W., and Xu, H. (2020). Iridium-catalyzed C-H amidation of s-tetrazines. Chem. Commun. (Camb.) *56*, 4692–4695.

Xu, H., Gu, Y., Zhang, S., Xiong, H., Ma, F., Lu, F., Ji, Q., Liu, L., Ma, P., Hou, W., et al. (2020). A chemistry for incorporation of selenium into DNA-encoded libraries. Angew. Chem. Int. Ed. https://doi.org/10.1002/anie.202003595.

Xu, H., Ma, F., Wang, N., Hou, W., Xiong, H., Lu, F., Li, J., Wang, S., Ma, P., Yang, G., and Lerner, R.A. (2019). DNA-encoded libraries: Aryl fluorosulfonates as versatile electrophiles enabling facile on-DNA suzuki, Sonogashira, and Buchwald reactions. Adv. Sci. *6*, 1901551.

Yeo, J., Kang, Y.M., Cho, S.I., and Jung, M.H. (2011). Effects of a multi-herbal extract on type 2 diabetes. Chin. Med. *6*, 10.

Zambaldo, C., Barluenga, S., and Winssinger, N. (2015). PNA-encoded chemical libraries. Curr. Opin. Chem. Biol. *26*, 8–15.

Zhao, G., Huang, Y., Zhou, Y., Li, Y., and Li, X. (2019). Future challenges with DNA-encoded chemical libraries in the drug discovery domain. Expert Opin. Drug Discov. *14*, 735–753.

Zhou, G., Myers, R., Li, Y., Chen, Y., Shen, X., Fenyk-Melody, J., Wu, M., Ventre, J., Doebber, T., Fujii, N., et al. (2001). Role of AMP-activated protein kinase in mechanism of metformin action. J. Clin. Invest. *108*, 1167–1174.

Zhou, Q., Gui, J., Pan, C.M., Albone, E., Cheng, X., Suh, E.M., Grasso, L., Ishihara, Y., and Baran, P.S. (2013). Bioconjugation by native chemical tagging of C-H bonds. J. Am. Chem. Soc. *135*, 12994–12997.

Zimmermann, G., and Neri, D. (2016). DNA-encoded chemical libraries: foundations and applications in lead discovery. Drug Discov. Today *21*, 1828–1834.

iScience, Volume *23*

Supplemental Information

Selection of Small Molecules that Bind to and Activate the Insulin Receptor from a DNA-Encoded Library of Natural Products

Jia Xie, Shuyue Wang, Peixiang Ma, Fei Ma, Jie Li, Wei Wang, Fengping Lu, Huan Xiong, Yuang Gu, Shuning Zhang, Hongtao Xu, Guang Yang, and Richard A. Lerner

Table of Contents

1. Transparent Methods

1.1 General Methods for Experiments

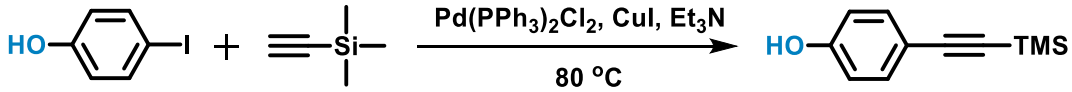
All commercially available organic compounds and DNA headpiece (HP-NH_2, 5'-/5phos/GAGTCA/iSp9/iUniAmM/iSp9/TGACTCCC-3') were obtained from Meilun Biotechnology and BioBioPha with the highest manufacturer grades. Unless otherwise noted, all commercial reagents and solvents were used without additional purification. NMR spectra were recorded on a Bruker AM-500 NMR spectrometer. Chemical shifts were reported as δ (ppm) and coupling constants were reported as J (hertz). Tetramethylsilane (TMS) was used as an internal reference for ^1H NMR and $CDCl_3$ was used as an internal reference for ^{13}C NMR (δ 77.0 ppm). The following abbreviations are used to designate multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, quint = quintet, br = broad. Mass spectra were recorded on an AB SCIEX 4600 mass spectrometer or on a Waters SQD 2 mass spectrometer. The complete DNA encoded chemical diagram followed the previous published scheme (Ma et al.,2019). Primers for PCR and deep-sequencing are showing below:

Forward:5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACG

Reverse:5'-CAAGCAGAAGACGGCATACGAGATGTCGTGATGTGACTGGAGTTC

1.2 Chemical synthesis

The syntheis of 4-((trimethylsilyl)ethynyl)phenol (**L4-c**)



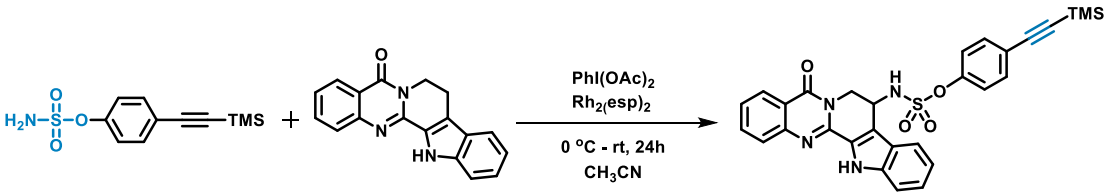
To a dry 25 mL flask, added 4-iodophenol **L4-a** (440.02 mg, 2 mmol), $Pd(PPh_3)_2Cl_2$ (42 mg, 0.06 mmol) and CuI (11.43 mg, 0.06 mmol) in Et_3N (6.5 mL) under nitorgen. Then ethynyltrimethylsilane **L4-b** (0.42 mL, 3 mmol) was added to the mixture and heated to 80 °C overnight. Upon the completion of the reaction, the mixture was filtered with celite and concentrated under vaccum. The filtrate was extracted with 50 mL water and 100 mL ethyl acetate. The organic layer was collected and washed with saturated NaCl aqueous and dried with anhydrous Na_2SO_4. Then the organic layer was concentrated under vacuum and purified with silica gel which gave a desired compound

L4-c as white oil (yield: 72%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.36 (d, *J* = 8.8 Hz, 2H), 6.75 (d, *J* = 8.8 Hz, 2H), 0.23 (s, 9H).

*The syntheis of 4-((trimethylsilyl)ethynyl)phenyl sulfamate (**L4**)*

To a dry 25 mL flask, added 4-((trimethylsilyl)ethynyl)phenol **L4-c** (110 mg, 0.5 mmol) and DIPEA (0.16 mL, 1 mmol) in DCM (2 mL). Then sulfamoyl chloride **L4-d** (69.3 mg, 0.6 mmol) was added to the mixture and stirred at room temperature overnight. Upon the completion of the reaction, the mixture was extracted with 50 mL water and 100 mL DCM. The organic layer was collected and washed with saturated NaCl aqueous and dried with anhydrous Na_2SO_4. Then the organic layer was concentrated under vacuum and purified with silica gel which gave a desired compound **L4** as white solid (yield: 63%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.50 (d, *J* = 8.8 Hz, 2H), 7.26 (d, *J* = 8.8 Hz, 2H),, 5.01 (brs, 2H), 0.25 (s, 9H); ^{13}C NMR (126 MHz, Chloroform-*d*) δ 149.82, 133.67, 122.69, 122.14, 103.59, 95.83, 0.01. HRMS (ESI) calculated for $[M+H]^+$ $[C_{11}H_{16}NO_3SSi]^+$ 270.0620, was 270.0625.

*The synthesis of compound **Rut-1***



A 5 mL sample vial was charged with $Rh_2(esp)_2$ (3 mg, 0.004 mmol) and **L4** (107.6 mg, 0.4 mmol) in 4.0 mL CH_3CN, Rutaecarpine (71.8 mg, 0.2 mmol) was then added. The reaction mixture was cooled to 4 °C, and $PhI(OAc)_2$ (128.8 mg, 0.4 mmol) was added in three portions over 3 hours and the reaction was stirred at 4 °C for 24 h. Water (5 ml) was added and the mixture was extracted with CHCl3 (3 x 10 mL).The organic layers

were dried over Na_2SO_4, filtered, concentrated in vacuum and the residue was purified by chromatography on silica gel which gave the desired compound **Rut-1** as a white solid (23%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 9.31 (s, 1H), 8.59 (s, 1H), 7.85 (d, *J* = 7.7 Hz, 1H), 7.68 (dd, *J* = 8.0, 1.5 Hz, 1H), 7.62 – 7.53 (m, 3H), 7.41 – 7.32 (m, 2H), 7.25–7.21 (m, 1H), 7.17–7.02 (m, 3H), 6.94 (d, *J* = 8.0 Hz, 1H), 5.42 (ddd, *J* = 9.0, 4.4, 1.7 Hz, 1H), 5.35 (dd, *J* = 14.6, 1.7 Hz, 1H), 3.73 (dd, *J* = 14.4, 4.3 Hz, 1H), 0.26 (s, 9H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 160.56, 149.92, 145.78, 144.16, 138.32, 135.17, 133.79, 127.30, 126.84, 126.80, 125.41, 123.49, 122.60, 122.08, 121.69, 120.50, 119.95, 116.86, 112.20, 103.66, 95.87, 47.41, 45.35, 0.00; HRMS (ESI) calculated for $[M+H]^+$ $[C_{29}H_{27}N_4O_4SSi]^+$ 555.1522, was 555.1521.

1.3 General procedure for the labelling of natural products (NPs)

The compounds (NP-Alkyne) were dissolved in DMSO (30 μL, 10 mM in DMSO), and mixed with N_3-HP-DNA (10 μL, 1 mM in water), THPTA (10 μL, 80 mM in DMSO), $CuSO_4 \cdot 5H_2O$ (10 μL, 80 mM in water) and sodium ascorbate (20 μL, 80 mM in water). The resulting mixture was shaked at room temperature overnight, and the products and yields were evaluated by LC-MS upon the reaction finished. After that, the scavenger sodium diethyldithiocarbamic acid (12 μL, 160 mM in water) was added. Then all the HP-DNA conjugated compounds (NP-HP-DNA) were collected and added 5 M NaCl solution (10% by volume) and cold ethanol (2.5 times by volume, ethanol stored at -20 °C). The mixture was stored in a -80 °C freezer for more than 30 minutes. The mixture was centrifuged for 15 minutes at 4°C in a micro-centrifuge at 12000 rpm. The supernatant was removed and the pellet was dissolved in water.

1.4 Combinatorial DEL synthesis

The small 10^4 combinatorial DEL library was synthesized according to previous protocol (Ma et al.,2019), in which combinatorial split-and-pool synthesis was carried out *via* coupling of 6 amine-(PEG)n-acids (building block 1), 46 amino acids (building block 2), and 46 carboxylic acids (building block 3) to afford a total of 12,696

combinatorial compounds.

To the above DEL library, 160 TCMs, FDA approved drugs, and compounds in clinical trials annotated by the late-stage toolbox were spiked based on each compound's cycle threshold (CT) number by quantitative polymerase-chain-reaction (qPCR) to afford the final *n*DEL library.

1.5 nDEL panning for insulin receptor

The biotinylated human insulin receptor extracellular domain (amino acid 1-956) (ECD-*h*IR) (Sino Biological, Cat. # 11081-H08H-B) and his-tagged ECD-*h*IR (Sino Biological, Cat. # 11081-H08H) were used to bind to streptavidin-coated Dynabeads M280 (Thermo Fisher Scientific, Cat. # 11205D) and cobalt based beads (Thermo Fisher Scientific, Cat. # 10103D), respectively. The beads were washed twice with PBST in 5 minute intervals. A magnetic rack was used to separate the beads from the supernatant, then 5ug of protein was added to PBS, at a final volume of 100ul, and incubated with the beads for 30 minutes at room temperature, rotating frequently. Afterward, the beads were washed twice with PBST and the supernatant was removed. 10ul of the DEL pool was added to 90ul of PBS and then used to resuspend the beads. Incubation occurred for 1 hour at room temperature, with rotation. Following incubation, the beads were washed twice with PBST, removing the supernatant each time. Elution occurred two ways. The first way by adding 100ul of PBS to the beads and heating at 95C for 10 minutes, and the second by eluting with 100ul of 50ug/mL insulin for 10 minutes. The supernatant and beads were separated using a magnet, and the supernatant was collected and sent for sequence analysis. The 30 μL final eluent was subject to deep-sequencing analysis.

1.6 Deep-sequencing and data analysis

The *n*DEL library contains a total of 12,856 chemical structures, each of which was encoded with a unique DNA sequence. Deep sequencing of *n*DEL was carried out

using Illumina method. The Illumina adaptor sequences around the DNA coding sequences were trimmed by CLC genomics workbench version 12 (Qiagen). The resulting DNA sequences were 30 base pairs in length corresponding to the DNA sequences of building blocks in 3 rounds of "split-pool" iterations. For each testing sample, the DNA coding sequences were mapped to the reference DEL library. No mismatch was allowed in the mapping. The mapped coding sequences were counted for all compounds across different samples. The total sequencing counts (S_{total}) represents the coding sequences counted for all compounds in a given sample. The sequencing counts (S) for each individual compound were normalized using the following equation (eq. 1), in which S_0 represents the normalized S.

$$S_0 = 100{,}000 \times S / S_{total} \qquad \text{(eq. 1)}$$

An in-house java program was developed to analyze enrichment of *n*DELs during the screening. The fold changes of normalized sequencing counts (i.e. enrichment fold) for each compound in *n*DEL after incubation with target protein were calculated in comparison with that in the reference library as shown in the equation below:

$$\text{enrichment fold} = S_{0,\,sample} / S_{0,\,reference} \qquad \text{(eq. 2)}$$

The hit criteria of *n*DEL screening take into account of both the normalized enrichment fold values (y-axis) and deep sequencing read counts (x-axis). Compounds with read counts less than 10 are considered highly unreliable, thus they are eliminated immediately from DEL before on-target screening. For each DEL, a baseline enrichment fold is recorded in the absence of target protein, and a normalized enrichment fold value can be calculated for each DEL compound in the library. The cutoff for hits identification is based on a simplified statistical analysis of a highly diverse population of data, which is the sum of average value of enrichment-folds of the whole library (μ) plus 3 times of the standard deviation (σ). Any DEL compounds showing enrichment-fold greater than μ+3σ are considered as hits.

1.7 Surface Plasmon Resonance (SPR)

The SPR binding assays were performed on a Biacore 8K instrument (GE Healthcare) The running buffer (PBS-P+ DMSO) contained 20 mM phosphate buffer, 2.7 mM KCl, 137 mM NaCl, 0.05% (v/v) surfactant P20, and 5% DMSO. The ECD-*h*IR (Sinobiological, Cat. # 11081-H08H) was covalently immobilized onto a CM5 chip by a standard LWM Immobilization method supplied by the Biacore 8K Control Software in a 10 mM sodium acetate buffer (pH 4.0). Rutaecarpine and Metformin were serially diluted as indicated using the running buffer and injected at a flow rate of 30 μL/min for 90 seconds for the association step followed by disassociation for an additional 90 seconds using the LWM multi-cycle kinetics / affinity method provided by GE Healthcare. Solvent correction was carried out before and after each analysis with 8 different concentrations of DMSO solution per cycle. The K_D value was derived using Biacore 8K Insight Evaluation Software (GE Healthcare) with predefined LMW multi-cycle kinetics evaluation method. All protein samples and compound samples were centrifuged at 20, 000 g for 10 min.

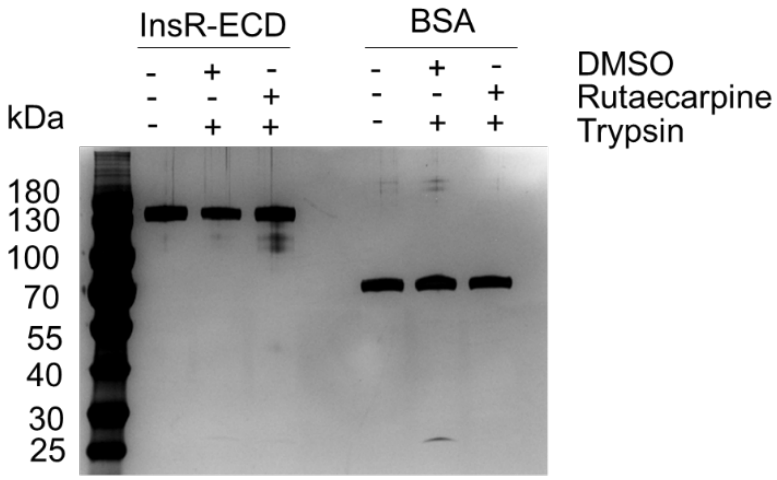
1.8 Partial Proteolysis Assay

200 ng recombinant insulin receptor extracellular domain (Sinobiological, Cat. # 11081-H08H) was mixed with 50 ng trypsin in the presence of 5% DMSO or 50 μM Rutaecarpine with 50 mM Tris/HCl (pH 8.0) and 20 mM calcium chloride ($CaCl_2$) at 37 °C for 5 min. The reaction mixtures were resolved using SDS-PAGE and visualized using silver staining kit (Thermo, Cat. # 24600).

1.9 Enzyme-linked immunosorbent assay for insulin receptor autophosphorylation.

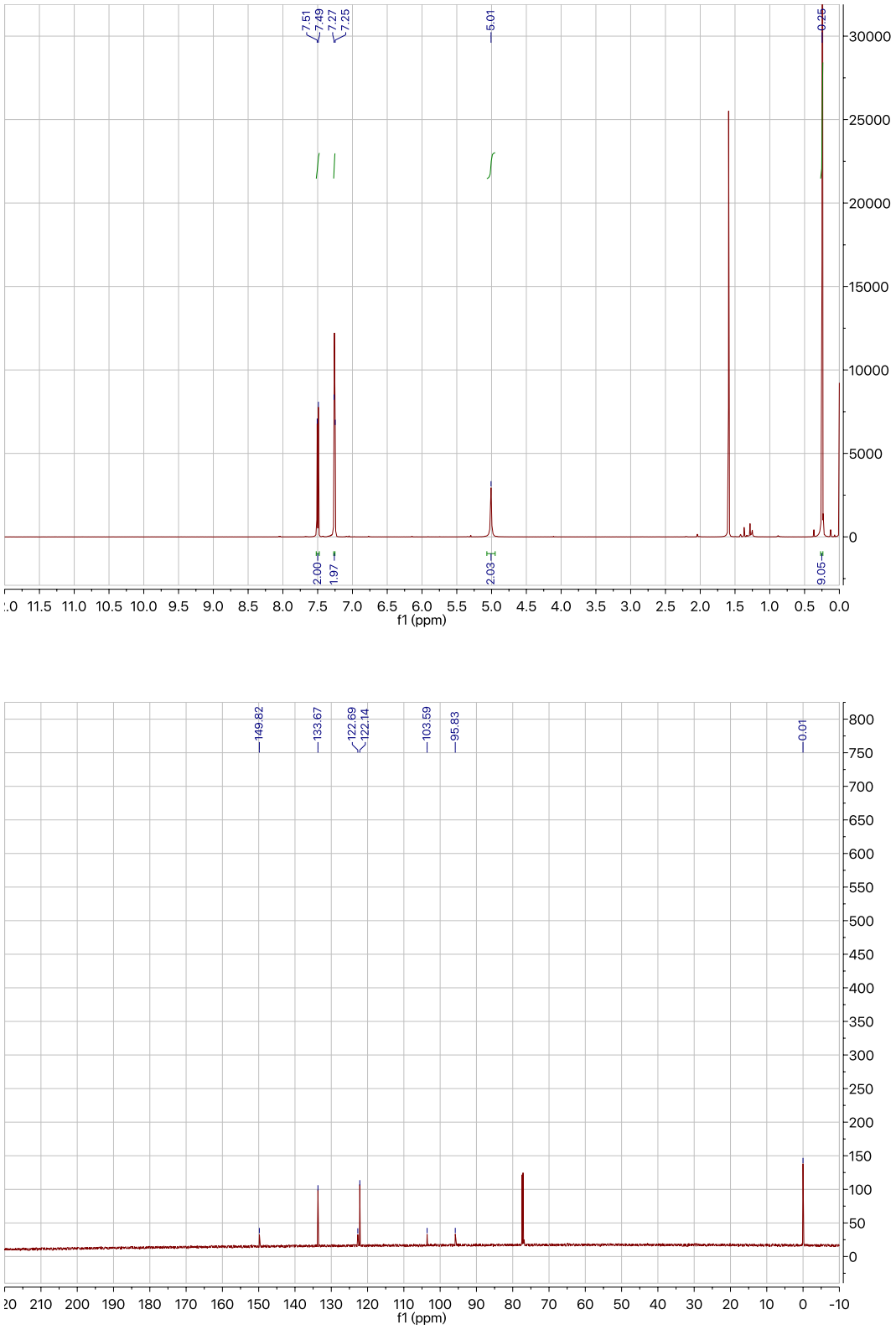
To study the activation of IR and insulin sensitization by selected natural products, a cellular assay to quantify the autophosphorylation was developed. Chinese hamster ovary (CHO) cells were maintained in Ham's F-12K medium with 10% FBS. CHO-*h*IR cells were transfected with plasmids encoding the full length human insulin receptor with C terminal GFP tag (Sinobiological, Cat. # HG11081-ACG) and sorted by flow cytometry. The autophosphorylation of IR was detected using the Phospho-

Insulin Receptor β (Tyr1150/1151) Sandwich ELISA Kit (Cell Signaling, Cat. # 7258C). CHO-*h*IR cells were seeded in a 48-well plate with 120,000 cells in 300 µL Ham's F-12K medium per well. For the activation experiment, cells were treated with 10 µM Rutaecarpine in 5% DMSO or 10 µM Metformin in 5% DMSO for 90 min at 37 °C. For insulin sensitization experiment, cells were first treated with 10 µM Rutaecarpine in 5% DMSO or 10 µM Metformin in 5% DMSO for 90 min, followed by the treatment with 100 nM insulin for 15 min at 37 °C. Cell media (NC) and 5% DMSO (DMSO) were used as negative controls, and 100 nM insulin as the positive control. Mouse monoclonal antibody against the insulin receptor β was coated onto the microwells. After incubation with cell lysates, both phosphorylated and non-phosphorylated insulin receptor proteins were captured by the coated antibody. Following extensive washing, the rabbit monoclonal antibody of phospho-insulin receptor β (Tyr1150/1151) was added to detect the captured phosphorylated insulin receptor (Tyr1150/1151) protein. HRP-linked anti-rabbit IgG was used as the secondary antibody. The HRP substrate, TMB, was added for visualization. The tyrosine phosphorylation was quantified at 450 nm by a plate reader.

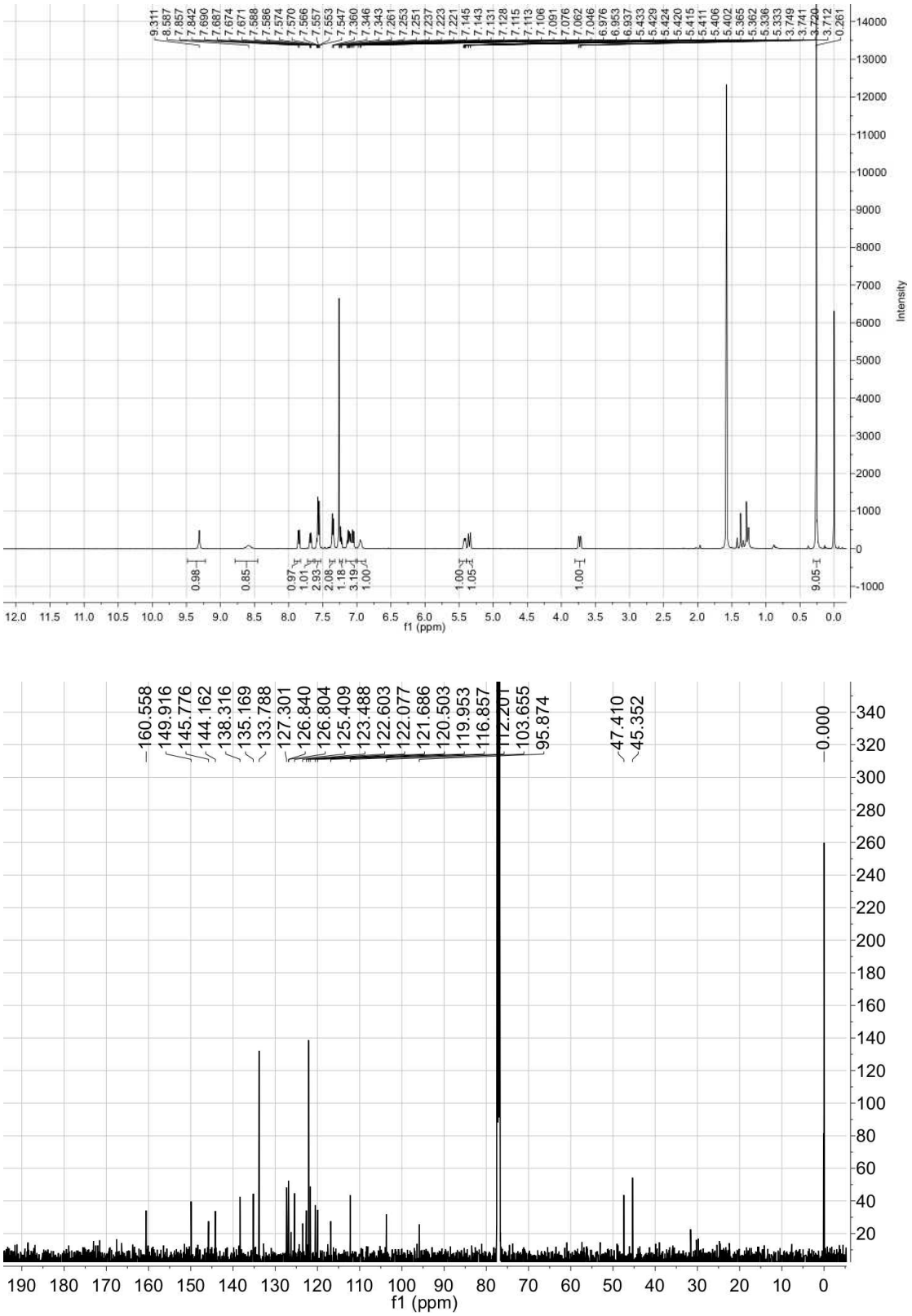


Supplementary Figure 1. Related to Figure 4. Comparison of proteolytic cleavage patterns of ECD-*h*IR and BSA in the presence of Rutaecarpine.

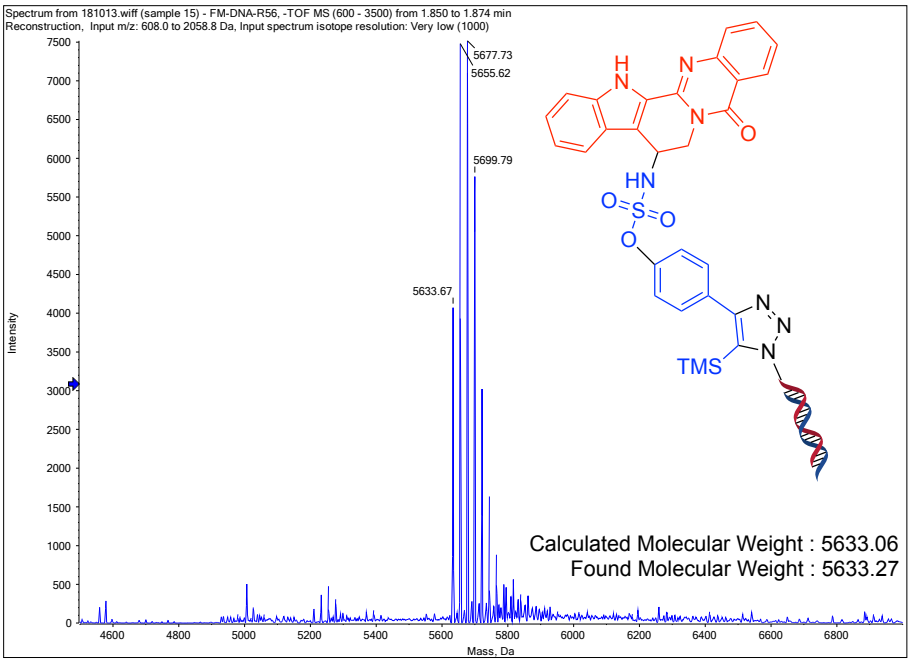
3. Copies of NMR and MS Spectrums for New Compounds.



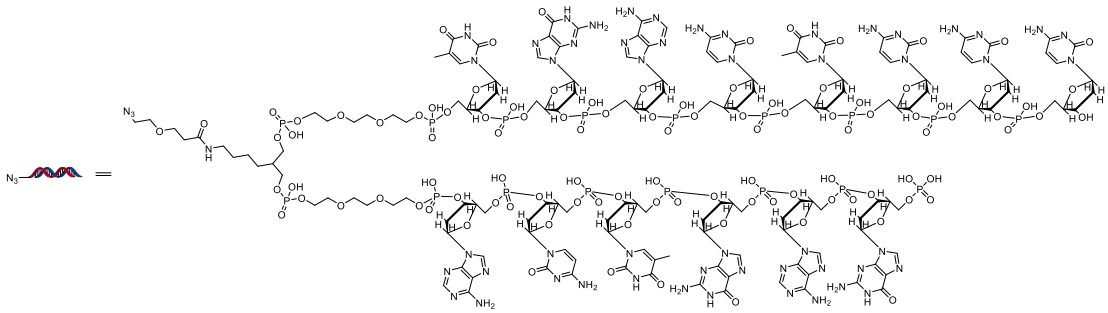
Supplementary Figure 2. Related to Scheme 1. The ^{1}H and ^{13}C NMR spectrum of **L4**.



Supplementary Figure 3. Related to Scheme 1. The ^{1}H and ^{13}C NMR spectrum of **Rut-1.**



Supplementary Figure 4. Related to Scheme 1. The MS spectral of DNA-Conjugated Rutaecarpine.



Supplementary Figure 5. Related to Scheme 1. The structure of N_3-HP-DNA.

3. Reference

Ma, P., Xu, H., Li, J., Lu, F., Ma, F., Wang, S., Xiong, H., Wang, W., Buratto, D., Zonta, F., Wang, N., Liu, K., Hua, T., Liu, Z., Yang, G., Lerner, R.A. (2019). Functionality-Independent DNA Encoding of Complex Natural Products. Angew. Chemie Int. Ed. 58, 9254–9261.

Peddibhotla, S., Dang, Y., Liu, J.O., Romo, D. (2007). Simultaneous Arming and Structure/Activity Studies of Natural Products Employing O−H Insertions: An Expedient and Versatile Strategy for Natural Products-Based Chemical Genetics. J. Am. Chem. Soc. 129, 12222–12231.

Zhou, Q., Gui, J., Pan, C.-M., Albone, E., Cheng, X., Suh, E.M., Grasso, L., Ishihara, Y., Baran, P.S. (2013). Bioconjugation by Native Chemical Tagging of C–H Bonds. J. Am. Chem. Soc. 135, 12994–12997.