# Design and Synthesis of a Cell Permeable Synthetic Transcription Factor Mimic

**Xiangshu Xiao**, **Peng Yu**, **Hyun-Suk Lim**, **Devanjan Sikder**, and **Thomas Kodadek**[*]
*Division of Translational Research, Departments of Internal Medicine and Molecular Biology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-9185*

## Abstract

Synthetic molecules capable of activating the expression of specific genes are of great interest as tools for biological research and, potentially, as a novel class of pharmaceutical agents. It has been demonstrated previously that such synthetic transcription factor mimics (STFMs) can be constructed by connecting a sequence-specific DNA-binding module to a molecule capable of binding to the transcriptional machinery via a suitable linker. These chimeras mimic the two basic properties of native transcription factors, which are able to recognize a promoter sequence specifically and to recruit the transcriptional machinery to that promoter. However, none of the compounds of this type reported to date have been shown to function in living cells. We report here the first example of a cell permeable STFM that activates the transcription of a reporter gene in mammalian cells. The compound is comprised of a cell permeable coactivator-binding peptoid fused to a DNA-binding hairpin polyamide. The peptoid was identified by screening a combinatorial library of µ 50,000 compounds for binding to the KIX domain of the CREB-binding protein (CBP), a mammalian transcription coactivator. When incubated with cultured HeLa cells carrying a luciferase reporter plasmid bearing several hairpin polyamide-binding sites, a five-fold increase in luciferase expression was observed. These experiments set the stage for the identification of hairpin polyamide-peptoid conjugates that are targeted to native genes.

## Introduction

Tightly regulated gene transcription and expression are important for the physiological development of organisms and deregulation of these processes often contributes to the pathogenesis of a variety of human diseases including cancer.[1] Therefore, much effort has been devoted to the development of technologies to control the expression of specific genes in living cells and animals to either study the function of the gene of interest or provide novel therapeutic agents to those diseases caused by aberrant gene expression. There are two ways to control the specific gene expression by up-regulation or down-regulation. For down-regulation of gene expression, the development of RNA interference (RNAi) technology has provided a relatively general way to knock down the expression of any gene.[2] On the other hand, it is much more difficult to specifically up-regulate gene expression and general methods for this purpose are still yet to be developed.[3–5]

Most efforts to develop reagents for gene-specific activation have focused on constructing compounds that reproduce the basic properties of native transcriptional activators. These proteins are minimally comprised of two domains, a sequence-specific DNA binding domain and a transcriptional activation domain (AD) capable of binding various proteins in the

---

*To whom correspondence should be addressed. Phone: 214-648-1239. FAX: 214-648-4156. E-mail: Thomas.Kodadek@utsouthwestern.edu.

transcription complex and chromatin remodeling and modification complexes.[6] These two activities allow transcriptional activators to recruit these complexes to specific promoters, thus triggering activation of gene expression.

Consistent with this view, stitching together heterologous DNA-binding and activation domains has been shown to produce functional transcription activators.[7, 8] The most recent elaboration of this biologically-inspired approach to the creation of designer transcription factors has been to identify zinc finger-type DNA-binding domains that target a promoter of interest and express these as fusion proteins with a potent AD such as VP16.[9–11] These chimeric proteins have been shown to act as potent and selective activators of endogenous genes in cultured cells. However, protein-based artificial transcription activators have the inherent limitation that the DNA that encodes them must be introduced into cells or animals, requiring transfection or related procedures for use in cultured cells and gene therapy in the case of eventual therapeutic applications. Therefore, the challenge of creating synthetic molecules capable of gene-specific activation has been taken up by several laboratories.

The first generation of synthetic transcription factor mimics (STFMs) were comprised of a sequence specific DNA-binding hairpin polyamide, peptide nucleic acid (PNA) or triplex-forming oligonucleotide (TFO) molecule conjugated to a peptide that was derived either from native AD sequences or identified via *in vitro* library screening procedures.[12–17] These molecules were important in demonstrating the proof of principle that synthetic molecules could activate gene expression *in vitro*, but the low cell permeability and stability of peptides rendered these chimeras poor candidates for *in vivo* applications. More recently, non-peptidic AD mimics have been described.[18–20] However, the only STFM reported to date that contained one of these molecules and a DNA-binding hairpin polyamide failed to activate transcription in cultured cells.[21] This was presumably due to the poor cell permeability of this hairpin polyamide-"wrenchnolol" conjugate since it did show activity in a cell-free system.[21]

In this article, we describe the isolation of six peptoids (oligo-N-substituted glycines[22]) that bind the KIX domain of the mammalian transcription coactivator CREB-Binding Protein (CBP).[23] These peptoids are shown to be cell permeable and to have activity as activation domain surrogates. In particular, we show that when a derivative of one of these peptoids is fused to a hairpin polyamide, the resultant chimera is capable of activating transcription in living cells.[24]

## Results

### Isolation of Core KIX Domain-binding Peptoids from a Combinatorial Library

The KIX domain of the mammalian transcription coactivator CBP has been validated as a target for the isolation of peptides that have activation domain-like activity.[25, 26] Therefore, following protocols published previously,[27, 28] a one-bead one-compound library of ~50,000 hexameric peptoids (Figure 1A, B; see Experimental Section) was screened for molecules that bind to a His$_6$-tagged derivative of murine CBP(586-672) (the core KIX domain).[29] After incubation of the beads with the biotinylated protein and thorough washing, Streptavidin-modified, red-emitting quantum dots were employed to identify beads that had retained significant amounts of protein (Figure 1C). Six putative hits, called BHK1-6 (BHK = Binders of His$_6$KIX), were identified and their sequences were determined by on-bead Edman degradation (Figure 1D and Figure S1).[27]

Examination of the sequences of these peptoids revealed significant sequence similarity (Figure 1D). Specifically, all six peptoids contain an Npip residue (see Figure 1B for nomenclature convention) at the C-terminus and an aromatic residue at the N-terminus. Furthermore, BHK2 and BHK5 are almost identical, differing only in the first position. BHK1, BHK3, BHK4, and

BHK6 all contain the sequence HYD-Nlys-Nlys-Nmba/Nall-HYD (where HYD represents a hydrophobic residue) (Figure 1D).

To determine if these compounds indeed bind KIX *in vitro*, a fluorescence polarization experiment was performed to monitor the binding of fluorescein-labeled BHKs to a GST-KIX fusion protein (Figure 1E). Five out of the six peptoids isolated show saturation binding to GST-KIX with apparent equilibrium dissociation constants in the low μM range. The binding of fluoresceinated BHK4 to GST-KIX could not be fit to a saturation binding curve (data not shown) in spite of its high sequence similarity with BHK3 and BHK6. None of the peptoids exhibited detectable binding to GST, showing that they recognize the KIX domain of CBP (Figure S2).

## Cellular Activities of the KIX-binding Peptoids

We then focused on whether these peptoids were capable of acting as AD surrogates in mammalian cells. To do this, we employed an assay reported previously (Figure 2A).[19] This involves incubating a peptoid-steroid conjugate with cells that contain a Gal4-responsive luciferase reporter gene and express a Gal4 DNA-Binding Domain-Glucocorticoid Receptor Ligand-Binding Domain (Gal4 DBD-GR LBD) fusion protein, which is trapped in the cytoplasm in the apo form.[30] The steroid employed in these conjugates is a dexamethasone derivative called OxDex,[31] which is a strict antagonist of the GR.[32, 33] Thus, if the OxDex-peptoid conjugate is able to diffuse through the cell membrane, binding of the steroid to the GR LBD will result in translocation of the Gal4 DBD-GR LBD complex into the nucleus, but expression of the luciferase reporter will not be stimulated unless the linked peptoid is capable of acting as an AD, since the fusion protein lacks this activity (see reference[34] for the first report of an assay of this type). In order to realize this scenario, it is critical that linkage of the peptoid to OxDex does not compromise binding of the steroid to the GR LBD and that the OxDex-peptoid conjugate is cell permeable. The binding affinities of the OxDex-BHK molecules (see Figure S3 for structures) to GR were assessed using a competition fluorescence polarization assay where increasing amounts of the OxDex-peptoid conjugates were added to a GR-fluormone™ complex (Invitrogen) and the subsequent decrease in polarization of the displaced fluormone ligand was monitored.[31] As presented in Table 1, all six OxDex-peptoid conjugates exhibit similar binding affinities to GR with $IC_{50}$ values ranging from 0.19 to 0.79 μM.

To test the relative cell permeability of all of the OxDex-peptoid conjugates, we employed the cellular assay described previously.[31] This is quite similar to that depicted in Figure 2A, except that the Gal4 DBD-GR LBD fusion protein also contains a potent VP16 AD.[35] Thus, entry of the OxDex-peptoid conjugate into cells results in concentration-dependent activation of the Gal4-responsive luciferase gene irrespective of the AD-like activities of the peptoid. As shown in Figure 2B, all of the OxDex-peptoid conjugates except OxDex-BHK3 entered cells to some extent, resulting in induction of luciferase expression. However, the results show clear differences in relative cell permeability. For example, the level of luciferase induction by OxDex-BHK1 is comparable to that of the simple amide OxDex-AEEA-CONH$_2$,[31] indicating that the peptoid does not hinder uptake of the steroid, but the other four conjugates show reduced induction level, indicating they are only modestly cell permeable.

Having established that five of the six BHKs bind KIX *in vitro* and that five of the six are cell permeable to some extent, we then asked whether these peptoids can function in cells as AD surrogates using the Figure 2A assay. Consistent with the permeability assay, OxDex-BHK3 was not able to activate the transcription of luciferase. The other five cell permeable OxDex-BHKs displayed measurable AD activity (Figure 2C; also see Table 1). There was no obvious correlation between KIX binding affinity and transactivation potency, which was also the case

for KIX-binding peptides.[26] It is also noteworthy that BHK4, which displayed non-ideal binding to the KIX domain, nonetheless exhibits significant AD-like activity in cells.

## Cellular and in vitro Binding Activities of Truncated BHKs

Only OxDex-BHK1 showed moderate AD-like activity (25-fold activation of luciferase) at 1.0 μM. None of the peptoid conjugates displayed measurable activity at 0.1 μM. Based on the data shown in Figure 2B, it seems likely that this is due to the limited cell permeability of the peptoids. The issue of cell permeability is likely to become even more serious when a peptoid AD surrogate is conjugated to a DNA-binding molecule to create an STFM. Presumably, the simplest way to increase the cell permeability of the peptoids is to make them smaller. Therefore, we have begun to examine the activity of truncated derivatives of the BHKs.

The C-terminal Npip residue was deleted in the most potent compounds BHK1, BHK5 and BHK6 resulting in TBHK1, TBHK5 and TBHK6 (**t**runcated **BHK**). As expected, the permeability of OxDex-TBHK6 was significantly increased compared to OxDex-BHK6 (Figure 3A). Surprisingly, however, the permeability of OxDex-TBHK1 was reduced drastically relative to the parent compound, whereas OxDex-TBHK5 was about as permeable as the parent compound (Figure 3A). These data indicate that predicting the cell permeability of compounds based on simple concepts such as molecular mass may not be straightforward.

The binding of fluorescein conjugates of the TBHKs to GST-KIX was assessed by fluorescence polarization spectroscopy (Figure 3B). TBHK1 and TBHK6 bind GST-KIX ($K_D$s of 2.5 and 11.6 μM, respectively) *in vitro*. On the other hand, TBHK5 did not show saturation binding to GST-KIX (data not shown). None of these three TBHK conjugates bound detectably to GST (Figure S2). The transcriptional activation potency of these three truncated OxDex-peptoid conjugates was also measured using the Figure 2A assay. Consistent with the permeability data, OxDex-TBHK1 was significantly less potent than the parent OxDex-BHK1, whereas both OxDex-TBHK5 and OxDex-TBHK6 were more potent than their parent compounds (Figure 3C). Of particular interest was the fact that OxDex-TBHK6 displayed detectable activity (9-fold activation) at a concentration as low as 0.1 μM. Based on this finding, we chose to evaluate TBHK6 as an AD surrogate in the context of a STFM.

## A Hairpin Polyamide-TBHK6 Conjugate Activates Gene Transcription in Living Cells

TBHK6 was conjugated via a flexible linker to a hairpin polyamide ImPy7,[12, 36] which was designed to bind the DNA sequence 5′-WGWWWW-3′ (W = A or T[37]), to give ImPy7-TBHK6 (Figure 4A; see Scheme S1 for the synthesis). This particular polyamide was employed because it has been shown previously that it is capable of recognizing its target sequence in living cells.[38]

This full conjugate was shown to be able to bind both GST-KIX and DNA with ImPy7-binding sites in vitro.[24] To test whether this ImPy7-TBHK6 conjugate can activate gene expression in living cells, it was added to HeLa cells transiently transfected with a plasmid-borne luciferase reporter gene containing six ImPy7 binding sites in its promoter (pGL3-6 × HPB).[38] A dose-dependent increase in luciferase expression was observed (Figure 4B), with the maximum effect being a 5-fold increase at an extracellular concentration of 3.0 μM. When acetylated ImPy7 (ImPy7Ac, Figure S4), lacking a conjugated peptoid, was employed, no activation was observed. Similar results were obtained with a conjugate containing ImPy7 and a non-KIX-binding peptoid.[24] Likewise, little effect was observed when these cells were incubated with OxDex-TBHK6 lacking a hairpin polyamide moiety. Finally, activation required the presence of multiple tandem ImPy7 binding sites in the promoter of the reporter gene, as evidenced by the fact that incubation of ImPy7-TBHK6 with cells transfected with a luciferase reporter plasmid containing only two ImPy7 binding sites (pGL3-2 × HPB)[38] failed to induce reporter

gene expression (Figure 4B). This result was essentially identical to that obtained using a reporter plasmid lacking ImPy7 binding sites (pGL3-E1b) in the promoter. These data demonstrate that activation of the reporter gene requires both the presence of multiple copies of the TBHK6 AD surrogate and its delivery to the promoter of the reporter gene via sequence-specific ImPy7-DNA interactions.

## The Action of ImPy7-TBHK6 on Endogenous Genes

In order to better understand the activities of the polyamide-peptoid conjugate in HeLa cells, an expression profiling experiment was conducted. HeLa cells were treated with ImPy7Ac (3.0 μM) or ImPy7-TBHK6 (3.0 μM) for 40 h. Total mRNA was isolated from the treated cells and the levels of ~46,000 transcripts were compared to identify genes that were up-regulated at least three fold in the cells treated with ImPy7-TBHK6 relative to those treated with the polyamide alone.[39–41] This analysis identified 45 genes that were up-regulated specifically by ImPy7-TBHK6 (Table S2). To determine if this activation likely reflected the binding of multiple polyamide-peptoid molecules to the promoters of these genes, the DNA up to 1000 base pairs 5′ of the transcriptional start site was searched for sequences fitting the polyamide target site (5′-WGWWWW-3′). As shown in Table S2, all these up-regulated genes contain multiple copies of ImPy7 binding sites in their putative promoter regions. This is consistent with the idea that the STFM is capable of activating endogenous genes by binding to promoters with multiple target sites.

## Cell Permeability of the Hairpin Polyamide-peptoid Conjugate

To the best of our knowledge, the above results represent the first report of a chimeric molecule comprised of a DNA-binding and an activation domain-like module with activity in living cells and, as such, represents a significant milestone in the field. However, the five-fold level of induction seems modest relative to that achieved in Figure 3C where the peptoid was delivered to the Gal4 DNA-binding domain via the steroid-GR LBD interaction. There are several possible explanations for this. One is that ImPy7-TBHK6, being a relatively large molecule, is only slightly cell permeable. To test this hypothesis, OxDex-ImPy7-TBHK6 (see Scheme S2 for structure and synthesis) was prepared and its cell permeability was compared with that of OxDex-TBHK6 employing the assay shown in Figure 2B. As shown in Figure 5, OxDex-ImPy7-TBHK6 was apparently less permeable than OxDex-TBHK6 by approximately 20-fold (Figure 5). However, as explained above, it is important to normalize these data to the relative affinity of the two OxDex conjugates for the GR LBD. OxDex-ImPy7-TBHK6 binds GR *in vitro* with an $IC_{50}$ of $0.71 \pm 0.4$ μM and OxDex-TBHK6 with an $IC_{50}$ of 0.09 μM (Table 1), an approximately 8-fold difference. Assuming that these values represent the dissociation constants for these steroid conjugates with the GR LBD in cells, the data suggest that the polyamide-peptoid conjugate is indeed less permeable than the peptoid alone, but only modestly so ($\approx$ 2.5-fold). However, this estimate should be taken with a grain of salt since it is based on the assumption that the apparent affinities of the steroid conjugates for GR observed in vitro correspond to the affinity of these molecule for Gal4 DBD-GR-LBD-VP16 inside the cell, which seems reasonable but is difficult to prove. Nonetheless, it is likely that other factors, perhaps including the efficiency with which the polyamide binds in the promoter in the cell, also contribute to the limited activity observed with ImPy7-TBHK6.

## Discussion

As discussed in the introduction, there is considerable interest in the development of pharmacological methods for the activation of specific genes. In this report, we describe the synthesis and characterization of a hairpin polyamide-peptoid conjugate that is capable of activating genes in living cells. Six hexameric peptoids were isolated by screening a library of ~50,000 compounds for binding to the core KIX domain of the murine coactivator CBP (Figure

1). A series of experiments to characterize their properties (Figures 2 and 3) revealed that while most of the peptoids indeed exhibited the desired AD-like properties in cells, their permeability was likely to be a limiting factor in their use as AD surrogates in the context of STFMs. Therefore, some preliminary work was done on the development of derivatives with improved permeability and a truncated pentameric derivative of one of the initial hits was chosen for evaluation in the context of a hairpin polyamide conjugate. As shown in Figure 4A, such a conjugate (which we call ImPy7-TBHK6) was synthesized. When incubated with HeLa cells, this compound supported a 5-fold activation of the transcription of a reporter gene carrying six ImPy7 binding sites in its promoter (Figure 4B).

Of even greater interest was the observation that the expression of 45 endogenous genes was activated at least three-fold by the ImPy7-TBHK6 chimera relative to the results observed when the cells were treated with the polyamide alone. This modest number indicates that the compound does not have widespread, non-specific effects on the transcriptome. Indeed, the finding that all these up-regulated genes bear multiple number of expected ImPy7 binding sites within 1000 base pairs of the transcriptional start site suggests that the polyamide-peptoid conjugate is acting as a promoter-specific activator. This result also suggests that significant activation requires the delivery of several copies of STFMs to a target promoter. This hypothesis is also consistent with the observation that only two polyamide binding sites in the promoter of the reporter gene were insufficient to support detectable activation by ImPy7-TBHK6 (Figure 4B).

Of course, this demonstration that a polyamide-peptoid can activate endogenous genes with appropriate binding sites in their promoters is merely the first step in a longer process to develop molecules with real utility as tools for biological studies or possible therapeutics. The activity of ImPy7-TBHK6 is modest compared to that of a protein-based transcription factor. For example, the Gal4 DBD-GR LBD-VP16 fusion protein is capable of stimulating the expression of a similar reporter gene with five Gal4 binding sites in the promoter by several hundred-fold in the presence of saturating steroid.[31] The relatively low activity observed with ImPy7-TBHK6 may be at least partially due to its limited cell permeability (Figure 5) and also may be limited by the efficiency with which ImPy7 recognizes the promoter.[38] Thus, the focus of future work will be to develop molecules with enhanced permeability and improved activity.

One route to improving permeability will be to minimize the size of these molecules. It may also be worthwhile to identify peptoids with higher affinity for coactivators. Since Wright and co-workers have published an NMR structure of the KIX domain,[29] it may be possible to characterize the peptoid-KIX complex structurally and use this information to design smaller and more effective derivatives. Efforts directed towards this goal are in progress. In this vein, it is worthwhile mentioning that we had previously isolated a different peptoid that binds CBP and which also exhibits AD-like activity in the context of a steroid fusion.[19] However, we have shown recently that this peptoid binds a region outside the KIX core domain[42] and thus it may be more difficult to optimize this peptoid-coactivator interaction in a structure-based fashion. This was one of the reasons for conducting the new screen focused on the core KIX domain reported here. However, this earlier CBP-binding peptoid might be useful in a synergy-based strategy to improve activator efficacy. Most genes are regulated by more than one native activator and synergistic interactions between these species are often observed. In some cases, this is due to cooperative binding to the promoter but in others, the basis of synergy is thought to be due to different activators recruiting different components of the transcriptional apparatus.[43, 44] Alternatively, two activators that bind the same coactivator, each with modest affinity, might recruit that protein more efficiently through cooperative interactions.[45, 46] Thus, a promising strategy to stimulate the level of activation and perhaps achieve greater specificity would be to deliver several different polyamide-peptoid conjugates to a target promoter that bind different sequences near the start site and which carry peptoids targeted to different

transcription proteins or to different surfaces of the same coactivator. We are currently testing whether the peptoids described here might function synergistically with our original CBP-binding molecule. We are also initiating efforts to identify peptoids targeted to other likely targets of native activators, such as chromatin remodeling and modification complexes.

# Experimental Section

## Peptoid Library Synthesis

The hexameric peptoid library was synthesized by split-and-pool method on TentaGel macrobeads (Rapp Polymere) using sub-monomer approach[47] by a microwave assisted protocol.[48] Briefly, TentaGel macrobeads (1.0 g, 0.51 mmol) were swollen in DMF for 60 minutes and washed with DMF. The bromoacetylation step was carried out using 2.4 mL each of 2.0 M bromoacetic acid in DMF and 3.2 M diisopropylcarbodiimide (DIC) in DMF. After washing the beads thoroughly with DMF, the beads were partitioned equally into seven portions. Each portion of the beads was treated with one of the seven amines (Figure 1B, 0.4 mL, 2 M in DMF), where the additional amino group in 1,4-diaminobutane and hydroxyl group in ethanolamine was protected as Boc and *t*-butyl group respectively. The procedures were repeated until the hexameric peptoid library with two constant C-terminal Npip was created. At the end of the synthesis, the protecting groups were removed by treating the beads with a 5 mL solution of 95% trifluoroacetic acid (TFA), 2.5% water and 2.5% triisopropylsilane at room temperature for 1.5 h. The beads were then washed thoroughly with dichloromethane, followed by DMF, neutralized with 20% diisopropylethylamine (DIPEA) in DMF, rinsed with dichloromethane again.

## Plasmids and Protein Expression

pEGal4 DBD-GR LBD, pEGal4 DBD-GR LBD-VP16, and pEG5B were described previously. [31] $His_6$KIX expression plasmid pHis$_6$KIX was constructed by ligating KIX coding region (586-672), digested from pET-21a containing KIX insert[29] (a gift from Peter Wright, Scripps Research Institute) with NdeI and BamHI, with NdeI- and BamHI-digested pET-15b (Novagen) using T4 DNA ligase (New England Biolabs). The GST-KIX expression plasmid pGST-KIX was constructed by amplifying the KIX coding region (586-672) from pET-21a containing KIX insert using forward primer 5′-GCTATCGGATCCGGTGTTCGAAAAGGCTGG-3′ and reverse primer 5′-GTCATCGAATTCTCATAAACGTGACCTCCG-3′. After digestion, the amplified DNA was ligated into EcoRI and BamHI-digested pGEX-2T (Amersham Biosciences) using T4 DNA ligase. The expression and purification of $His_6$KIX and GST-KIX fusions were carried out according to standard methods using affinity fast protein liquid chromatography (FPLC).

## Cell Culture, Transient Transfection, and Luciferase Assay

HeLa cells ($2 \times 10^4$) were plated in 96-well plate 24 h before transfection at 37 °C under 5% $CO_2$ in Dulbecco's modified Eagle medium (Invitrogen) supplemented with 10% (v/v) fetal calf serum (Invitrogen), 10 μg/mL penicillin and 10 μg/mL streptomycin (Invitrogen). Cells were then transfected by Lipofectamine™ 2000 (Invitrogen) according to manufacturer's protocol. After 3 h, the transfection mixture was removed and replaced with fresh normal medium and then peptoid derivatives were added to the tissue culture medium and the cells were then further incubated at 37 °C under 5% $CO_2$ for 40 h. Luciferase assay was done with the cell lysates using dual-luciferase assay systems (Promega).

## Fluorescence Polarization Assay

The fluoresceinated BHKs or TBHKs (5–10 nM) was incubated with indicated concentrations of GST-KIX or GST in the presence of 1.0 μM BSA in PBS (1×, pH 7.4) in a final volume of

100 μL at room temperature for 2 h at dark. The fluorescence polarization values were then measured on a Panvera Beacon 2000 instrument (Invitrogen). The competition fluorescence polarization to monitor the binding of OxDex-peptoids to Glucocorticoid Receptor (GR) was done as described previously.[31]

### Biotinylation of His$_6$KIX

The His$_6$KIX fusion in 0.1 N NaHCO$_3$ buffer (pH 8.3) and water soluble biotin-NHS ester sulfo-NHS-LC-Biotin (8 eq., Invitrogen) were mixed together at room temperature and the mixture was incubated with tumbling at room temperature for 1 h. The reaction was quenched by adding 1.5 N hydroxylamine and incubated at room temperature for another hour. The unreacted biotin derivative was removed by dialysis against a buffer containing 50 mM Tris (pH 8.0), 100 mM KCl, 12.5 mM MgCl$_2$, 1 mM EDTA and 0.05% Tween-20. The degree of labeling was determined by FluoReporter® Biotin Quantitation Assay Kit (Invitrogen) and was found to be 3.

### Global Gene Expression Analysis

HeLa cells were either treated with 3.0 μM ImPy7-TBHK6 or ImPy7Ac for 40 h at 37 °C under 5% CO$_2$ in Dulbecco's modified Eagle medium (Invitrogen) supplemented with 10% (v/v) fetal calf serum (Invitrogen), 10 μg/mL penicillin and 10 μg/mL streptomycin (Invitrogen). Total RNA was isolated and processed using TotalPrep RNA amplification and labeling kit as suggested (Illumina, catalog # IL1791). Each sample was hybridized in duplicate to Sentrix® Human-6 expression beadchip for profiling over 46,000 RNA transcripts and the data were analyzed by Beadstudio (Illumina) in UT Southwestern Microarray Core Facility.

### Screening of the Peptoid Library for Ligands to His$_6$KIX

The hexameric peptoid library displayed on TentaGel macrobeads (~50,000 beads) were swollen in TBST (50 mM Tris, pH 8.0, 150 mM NaCl, 0.1% Tween-20) overnight. TBST was decanted and the resulting beads were incubated with Qdot® 655 Streptavidin (Invitrogen, 5 nM) in the presence of *E. coli.* lysate (6.5 mg/mL) and BSA (10 mg/mL) in TBST at room temperature for 2 h. The buffer was removed and washed with TBST (5 × 1 mL). The peptoids that bind the Streptavidin-coated quantum dots directly were visualized by irradiation of the beads through a DAPI filter (360 nm), and the red-emitting beads were removed under a fluorescence microscope using a micropipette. The remaining beads were then incubated with biotin-labeled His$_6$KIX (500 nM) in the presence of *E. coli.* lysates (6.5 mg/mL) and BSA (10 mg/mL) in TBST at room temperature for 1 h. The buffer was removed and the beads were washed with TBST (5 × 1 mL). Then the beads were incubated with Qdot® 655 Streptavidin (5 nM) in the presence of *E. coli.* lysate (6.5 mg/mL) and BSA (10 mg/mL) in TBST at room temperature for 2 h. After removal of buffer, the beads were washed with TBST (5 × 1 mL). The red-shining beads were again visualized and picked as described above. The six beads scored as hits were heated at 95 °C in 1% SDS for 5 min. The beads were then washed successively with PBS (pH 7.4, twice) and H$_2$O (three times). The washed beads were then subjected to automated single-bead Edman degradation analysis to derive the peptoid sequence on each individual bead.

### Chemical Synthesis of Individual Compounds

The detailed structures and their syntheses are presented in Supporting Information.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

# References

1. Hanahan D, Weinberg RA. Cell 2000;100:57–70. [PubMed: 10647931]

2. Novina CD, Sharp PA. Nature 2004;430:161–164. [PubMed: 15241403]

3. Denison C, Kodadek T. Chem Biol 1998;5:R129–R145. [PubMed: 9653545]

4. Ansari AZ, Mapp AK. Curr Opin Chem Biol 2002;6:765–772. [PubMed: 12470729]

5. Arndt HD. Angew Chem-Int Ed 2006;45:4552–4560.

6. Ptashne M. Nature 1988;335:683–689. [PubMed: 3050531]

7. Brent R, Ptashne M. Cell 1985;43:729–736. [PubMed: 3907859]

8. Sadowski I, Ma J, Triezenberg S, Ptashne M. Nature 1988;335:563–564. [PubMed: 3047590]

9. Tan SY, Guschin D, Davalos A, Lee YL, Snowden AW, Jouvenot Y, Zhang HS, Howes K, McNamara AR, Lai A, Ullman C, Reynolds L, Moore M, Isalan M, Berg LP, Campos B, Qi H, Spratt SK, Case CC, Pabo CO, Campisi J, Gregory PD. Proc Natl Acad Sci U S A 2003;100:11997–12002. [PubMed: 14514889]

10. Lund CV, Blancafort P, Popkov M, Barbas CF. J Mol Biol 2004;340:599–613. [PubMed: 15210357]

11. Beerli RR, Barbas CF. Nat Biotechnol 2002;20:135–141. [PubMed: 11821858]

12. Mapp AK, Ansari AZ, Ptashne M, Dervan PB. Proc Natl Acad Sci U S A 2000;97:3930–3935. [PubMed: 10760265]

13. Ansari AZ, Mapp AK, Nguyen DH, Dervan PB, Ptashne M. Chem Biol 2001;8:583–592. [PubMed: 11410377]

14. Arora PS, Ansari AZ, Best TP, Ptashne M, Dervan PB. J Am Chem Soc 2002;124:13067–13071. [PubMed: 12405833]

15. Liu B, Han Y, Corey DR, Kodadek T. J Am Chem Soc 2002;124:1838–1839. [PubMed: 11866581]

16. Liu B, Han Y, Ferdous A, Corey DR, Kodadek T. Chem Biol 2003;10:909–916. [PubMed: 14583257]

17. Stanojevic D, Young RA. Biochemistry 2002;41:7209–7216. [PubMed: 12044151]

18. Minter AR, Brennan BB, Mapp AK. J Am Chem Soc 2004;126:10504–10505. [PubMed: 15327284]

19. Liu B, Alluri PG, Yu P, Kodadek T. J Am Chem Soc 2005;127:8254–8255. [PubMed: 15941237]

20. Shimogawa H, Kwon Y, Mao Q, Kawazoe Y, Choi Y, Asada S, Kigoshi H, Uesugi M. J Am Chem Soc 2004;126:3461–3471. [PubMed: 15025473]

21. Kwonj Y, Arndt HD, Qian M, Choi Y, Kawazoe Y, Dervan PB, Uesugi M. J Am Chem Soc 2004;126:15940–15941. [PubMed: 15584709]

22. Simon RJ, Kania RS, Zuckermann RN, Huebner VD, Jewell DA, Banville S, Ng S, Wang L, Rosenberg S, Marlowe CK, Spellmeyer DC, Tan RY, Frankel AD, Santi DV, Cohen FE, Bartlett PA. Proc Natl Acad Sci U S A 1992;89:9367–9371. [PubMed: 1409642]

23. Chrivia JC, Kwok RPS, Lamb N, Hagiwara M, Montminy MR, Goodman RH. Nature 1993;365:855–859. [PubMed: 8413673]

24. Xiao X, Yu P, Lim H, Sikder D, Kodadek T. Angew Chem-Int Ed 2007;46:2865–2868.

25. Frangioni JV, LaRiccia LM, Cantley LC, Montminy MR. Nat Biotechnol 2000;18:1080–1085. [PubMed: 11017047]

26. Volkman HM, Rutledge SE, Schepartz A. J Am Chem Soc 2005;127:4649–4658. [PubMed: 15796530]

27. Alluri PG, Reddy MM, Bachhawat-Sikder K, Olivos HJ, Kodadek T. J Am Chem Soc 2003;125:13995–14004. [PubMed: 14611236]

28. Olivos HI, Bachhawat-Sikder K, Kodadek T. ChemBioChem 2003;4:1242–1245. [PubMed: 14613120]

29. Radhakrishnan I, Perez-Alvarado GC, Parker D, Dyson HJ, Montminy MR, Wright PE. Cell 1997;91:741–752. [PubMed: 9413984]

30. Picard D. Methods Enzymol 2000;327:385–401. [PubMed: 11044998]

31. Yu P, Liu B, Kodadek T. Nat Biotechnol 2005;23:746–751. [PubMed: 15908941]

32. Rousseau GG, Kirchhoff J, Formstecher P, Lustenberger P. Nature 1979;279:158–160. [PubMed: 35753]

33. Licitra EJ, Liu JO. Proc Natl Acad Sci U S A 1996;93:12817–12821. [PubMed: 8917502]

34. Nyanguile O, Uesugi M, Austin DJ, Verdine GL. Proc Natl Acad Sci U S A 1997;94:13402–13406. [PubMed: 9391036]

35. Triezenberg SJ, Kingsbury RC, McKnight SL. Genes Dev 1988;2:718–729. [PubMed: 2843425]

36. Dickinson LA, Burnett R, Melander C, Edelson BS, Arora PS, Dervan PB, Gottesfeld JM. Chem Biol 2004;11:1583–1594. [PubMed: 15556009]

37. Dervan PB, Edelson BS. Curr Opin Struct Biol 2003;13:284–299. [PubMed: 12831879]

38. Liu B, Yu P, Alluri PG, Kodadek T. Mol BioSyst 2005;1:307–317. [PubMed: 16880996]

39. Olenyuk BZ, Zhang GJ, Klco JM, Nickols NG, Kaelin WG, Dervan PB. Proc Natl Acad Sci U S A 2004;101:16768–16773. [PubMed: 15556999]

40. Dudouet B, Burnett R, Dickinson LA, Wood MR, Melander C, Belitsky JM, Edelson B, Wurtz N, Briehn C, Dervan PB, Gottesfeld JM. Chem Biol 2003;10:859–867. [PubMed: 14522056]

41. Supekova L, Pezacki JP, Su AI, Loweth CJ, Riedl R, Geierstanger B, Schultz PG, Wemmer DE. Chem Biol 2002;9:821–827. [PubMed: 12144926]

42. Alluri PG, Liu B, Yu P, Xiao X, Kodadek T. Mol BioSyst 2006;2:568–579. [PubMed: 17216038]

43. Chi TH, Lieberman P, Ellwood K, Carey M. Nature 1995;377:254–257. [PubMed: 7675113]

44. Herschlag D, Johnson FB. Genes Dev 1993;7:173–179. [PubMed: 8436289]

45. Hiroi M, Ohmori Y. J Biol Chem 2003;278:651–660. [PubMed: 12403783]

46. Mink S, Haenig B, Klempnauer KH. Mol Cell Biol 1997;17:6609–6617. [PubMed: 9343424]

47. Zuckermann RN, Kerr JM, Kent SBH, Moos WH. J Am Chem Soc 1992;114:10646–10647.

48. Olivos HJ, Alluri PG, Reddy MM, Salony D, Kodadek T. Org Lett 2002;4:4057–4059. [PubMed: 12423085]
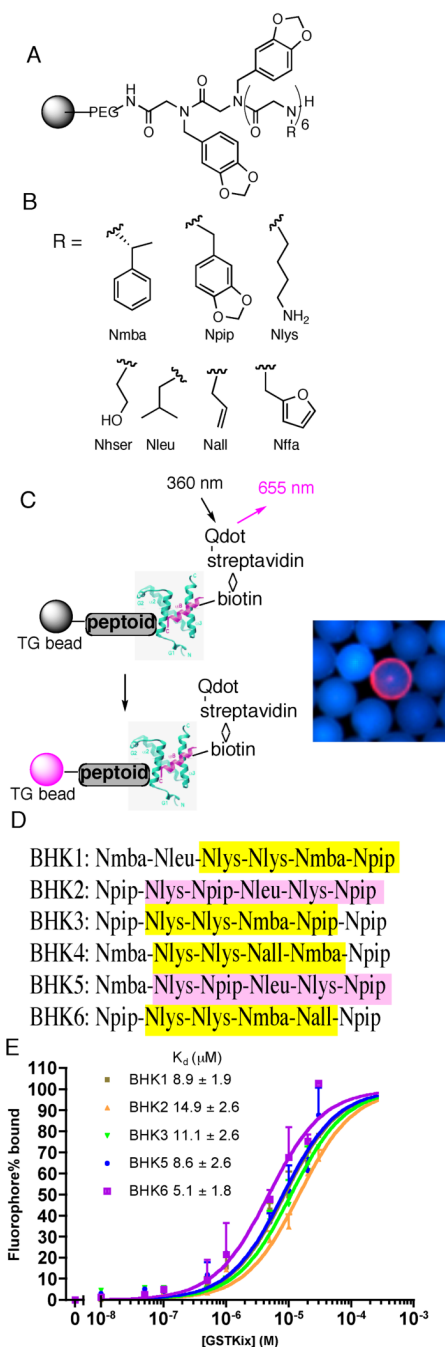
**Figure 1.**
Isolation and characterization of KIX-binding peptoids BHK1-6. (A) General structure of the hexameric peptoid library. (B) The structure and code of the seven residues in the library. The codes are either the short abbreviations of the amines used in the synthesis (<u>pip</u>eronylamine, (*R*)-(+)-α-<u>m</u>ethyl<u>b</u>enzyl<u>a</u>mine, <u>all</u>ylamine, <u>f</u>ur<u>f</u>uryl<u>a</u>mine) or derived from the side chain similarity with natural amino acids (<u>lys</u>ine, <u>leu</u>cine, <u>ho</u>mo<u>ser</u>ine). (C) Left: A schematic presentation of screening protocol utilizing red-emitting streptavidin-modified quantum dots. Right: A typical fluorescence micrograph of a bead scored as a hit (red halo) in the midst of large number of beads scored as negatives (blue). (D) The sequences of the six hits as determined by on-bead Edman degradation analysis (see Figure S1 for chemical structures).

Similarities in the sequences are highlighted (see text for a discussion). (E) *In vitro* binding of carboxyfluorescein-substituted BHKs (see Figure S3 for structures) to GST-KIX monitored by fluorescence polarization. The observed polarization values ($P$) were converted to percentage of fluorophore binding [$(P-P_{min})/(P_{max}-P_{min}) \times 100\%$] using fluorophore only as $P_{min}$, whereas $P_{max}$ was derived by fitting the raw $P$ data to a one-site binding model in Prism4.0 (GraphPad Software, San Diego, CA). The apparent $K_D$s, presented as mean $\pm$ SEM, were derived from the best fit of the data to the sigmoidal dose-response model in Prism4.0. Each data point, shown as mean + SD, represents an average of three independent measurements.
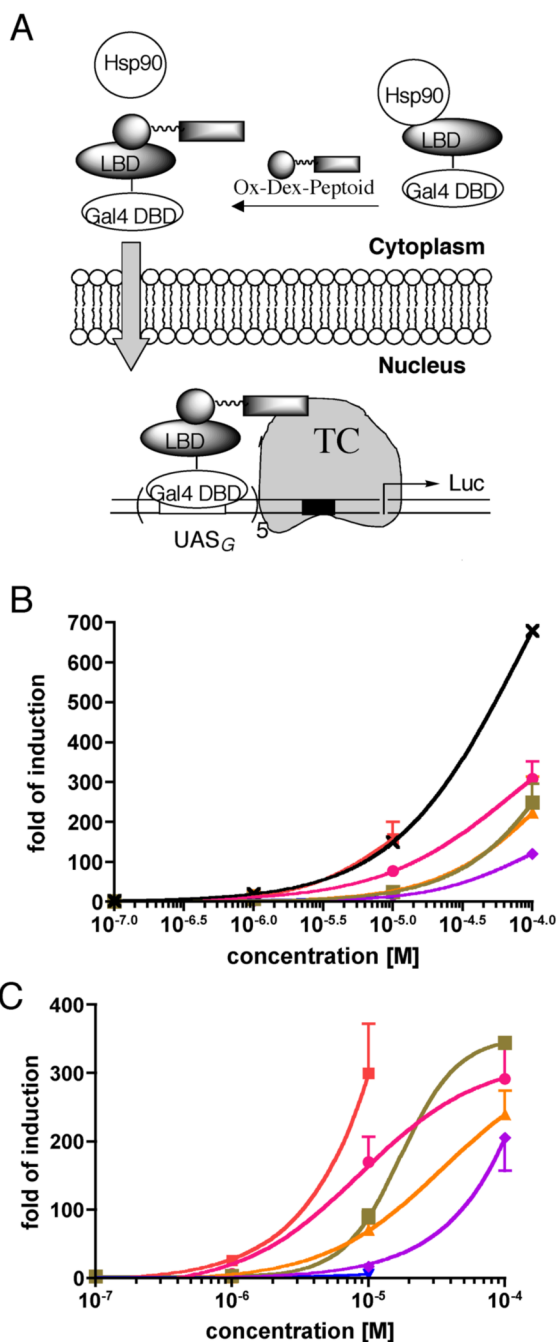
**Figure 2.**
Cellular activities of OxDex-BHK1-6 (see Figure S3 for structures). (A) Cartoon of the assay employed to measure the AD-like activity of the peptoid (see text for details). TC = transcription complex. (B) HeLa cell permeability of the OxDex-peptoid conjugates. OxDex-AEEA-CONH$_2$ (×), OxDex-BHK1 ( ■ ), OxDex-BHK2( ▲ ), OxDex-BHK3( ▼ ), OxDex-BHK4( ◆ ), OxDex-BHK5( ● ) or OxDex-BHK6( ■ ). This assay[31] is almost identical to that shown in A, except that the Gal4 fusion protein also contains a VP16 AD. The values shown are ratios of the level of induction measured using dual-luciferase assay system (Promega; Gal4-responsive/constitutive). Data points for OxDex-BHK1 and OxDex-BHK3 at 100 μM could not be obtained due to significant cytotoxicity at this concentration. (C) Activation

domain-like activity of OxDex-BHK1-6 in the assay shown schematically in (A). The symbol designations are the same as in (B). The data, shown as the mean + SD or mean - SD, are representative of at least three independent experiments with each point measured in duplicate. The curves are the best fit of the data to the sigmoidal dose-response model in Prism4.0.
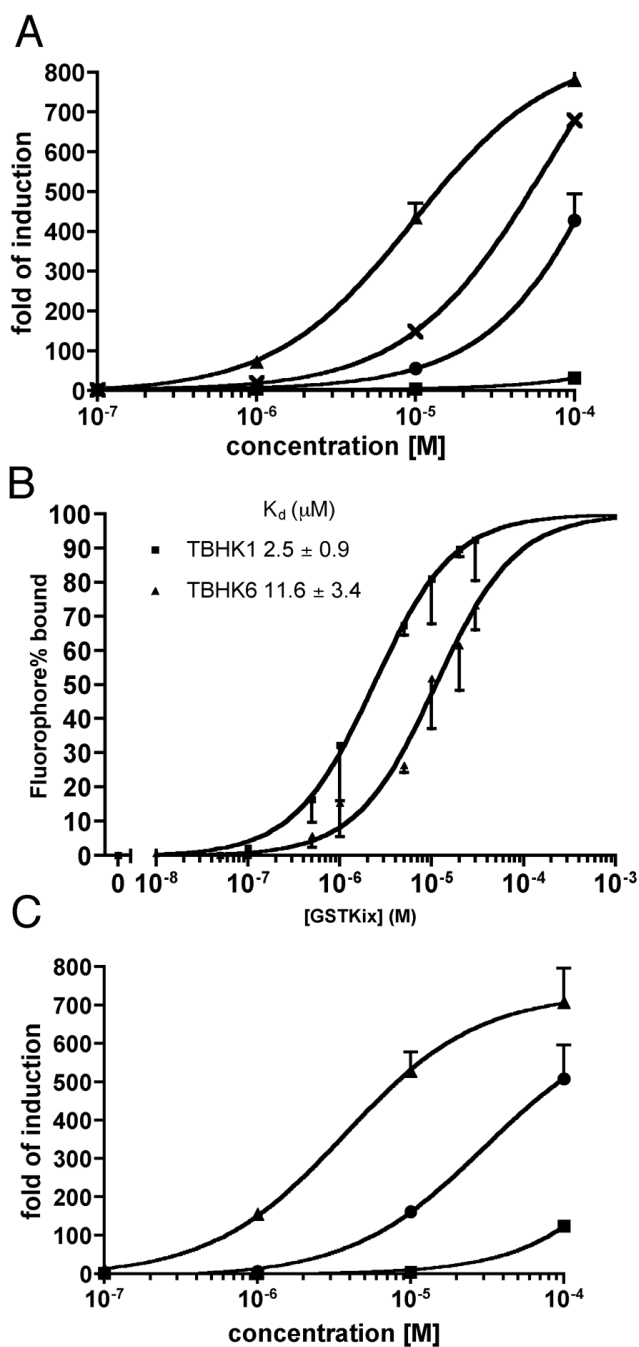
**Figure 3.**
HeLa cell permeability, transcriptional activities of OxDex-TBHKs and binding affinity of carboxyfluorescein substituted TBHKs to GST-KIX (see Figure S3 for structures). (A) HeLa cell permeability of OxDex-TBHK1,5,6. OxDex-AEEA-$CONH_2$ (×), OxDex-TBHK1(■), OxDex-TBHK5(●) or OxDex-TBHK6 (▲). (B) *In vitro* binding of carboxyfluorescein-substituted TBHKs to GST-KIX monitored by fluorescence polarization. Data processing is the same as that in Figure 1E. Each data point, shown as mean - SD, represents an average of three independent measurements. (C) Activation domain-like activity of OxDex-TBHK1,5,6. The symbol designations are the same as in (A). The data, shown as mean + SD, are

representative of at least three independent experiments performed in duplicate. The curves are the best fit of the data to the sigmoidal dose-response model in Prism4.0.
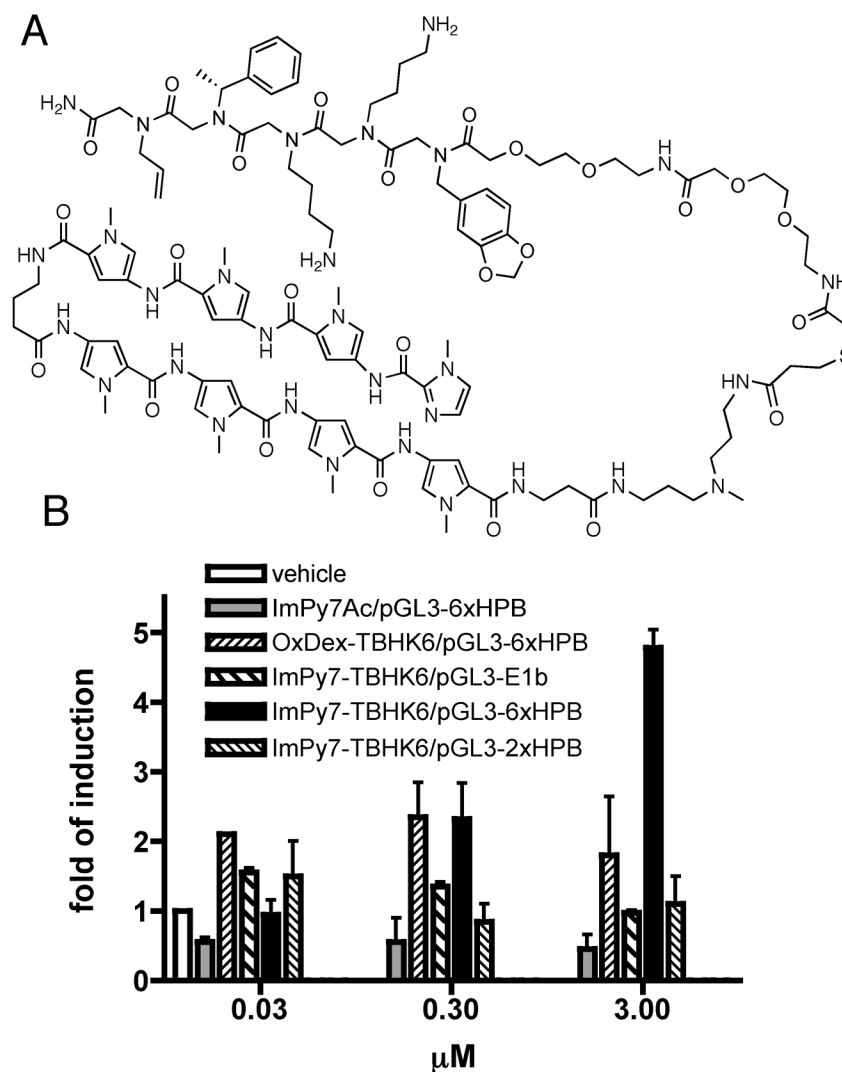
**Figure 4.**
Structure and activity of ImPy7-TBHK6. (A) Structure of ImPy7-TBHK6 (see Scheme S1 for its synthesis). (B) Transcriptional activity of ImPy7-TBHK6 in HeLa cells. Reporter plasmids pGL3-6×HPB, pGL3-2×HPB, pGL3-E1b have six, two, and zero ImPy7 binding sites upstream of luciferase gene, respectively. The fold of induction represents the ratio of the inducible luciferase activity normalized to the constitutively expressed *Renilla* luciferase activity. The activities were measured after incubating the cells with the compound indicated for 40 hours.
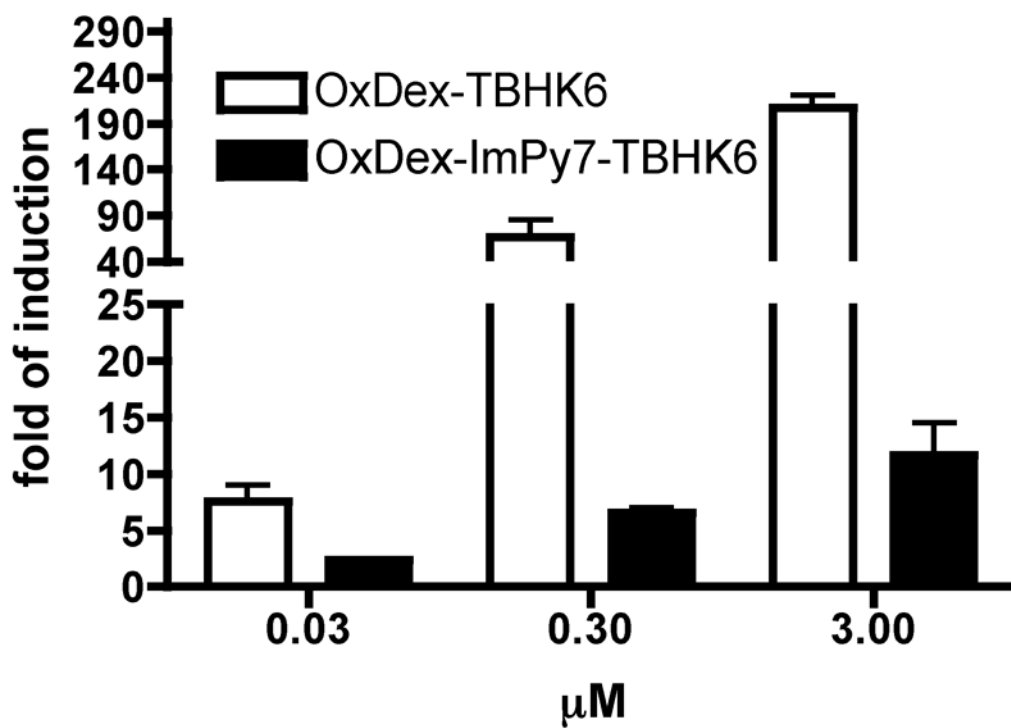
**Figure 5.**
HeLa cell permeability of OxDex-TBHK6, OxDex-ImPy7-TBHK6 (see Scheme S2 for its synthesis). The data, shown as mean + SD, are representative of three independent experiments performed in duplicate.

**Table 1**

Summary of cell permeability, transcriptional activity and *in vitro* binding to GR of various OxDex conjugates. Data are presented as mean ± SEM.

| compound | Luc induction in cells with Gal4DBD-GRLBD-VP16 | | Luc induction in cells with Gal4DBD-GRLBD | | *in vitro* binding affinity[a] |
|---|---|---|---|---|---|
| | $EC_{50}$ (μM) | $IND_{max}$ (fold) | $EC_{50}$ (μM) | $IND_{max}$ (fold) | $IC_{50}$ (μM) |
| OxDex-BHK1 | Nd[b] | >158 | Nd[b] | >300 | 0.42 ± 0.07 |
| OxDex-BHK2 | Nd[b] | >223 | Nd[b] | >230 | 0.23 ± 0.07 |
| OxDex-BHK3 | Nd[c] | <20 | Nd[c] | <6 | 0.23 ± 0.08 |
| OxDex-BHK4 | Nd[b] | >120 | Nd[b] | >206 | 0.79 ± 0.20 |
| OxDex-BHK5 | Nd[b] | >310 | 11.4 ± 7.6[d] | 295 ± 83[d] | 0.21 ± 0.02 |
| OxDex-BHK6 | Nd[b] | >250 | 78 ± 56[d] | 354 ± 150[d] | 0.19 ± 0.06 |
| OxDex-TBHK1 | Nd[b] | >30 | Nd[b] | >125 | 0.20 ± 0.09 |
| OxDex-TBHK5 | Nd[b] | >428 | Nd[b] | >508 | 0.23 ± 0.10 |
| OxDex-TBHK6 | 10 ± 3[d] | 846 ± 85 | 4.7 ± 2.7[d] | 734 ± 119[d] | 0.09 ± 0.03 |
| OxDex-ImPy7-TBHK6 | Nd[b] | >12 | NT[e] | NT[e] | 0.71 ± 0.4 |

[a] *in vitro* binding affinity of OxDex-peptoids to GR as determined by competition fluorescence polarization employing GR competitor assay kit (Invitrogen).

[b] not determined due to the fact that the dose-response curve is not saturated at the highest test concentration.

[c] not determined due to the limited activity seen with this compound.

[d] obtained from the best fit of the data to the sigmoidal dose-response model in Prism4.0.

[e] not tested in this system.