Article

Cell-based selection expands the utility of DNA-encoded small-molecule library technology to cell surface drug targets: Identification of novel antagonists of the NK3 tachykinin receptor

Zining Wu, Todd L. Graybill, Xin X Zeng, Michael Platchek, Jean Zhang, Vera Q Bodmer, David D Wisnoski, Jianghe Deng, Frank T. Coppo, Gang X Yao, Alex Tamburino, Genaro S Scavello, G Joseph Franklin, Sibongile Mataruse, Katie L Bedard, Jeffrey A Messer, Andrew J Pope, and David I Israel

ACS Comb. Sci., **Just Accepted Manuscript** • DOI: 10.1021/acscombsci.5b00124 • Publication Date (Web): 12 Nov 2015

Cell-based selection expands the utility of DNA-encoded small-molecule library technology to cell surface drug targets: Identification of novel antagonists of the NK3 tachykinin receptor

Zining Wu[1], Todd L Graybill[1], Xin Zeng[1], Michael Platchek[1], Jean Zhang[2], Vera Q. Bodmer[1], David D. Wisnoski[1], Jianghe Deng[1], Frank T. Coppo[1], Gang Yao[2], Alex Tamburino[2], Genaro Scavello[1], G. Joseph Franklin[2], Sibongile Mataruse[2], Katie L. Bedard[2], Jeffrey A Messer[2], Andrew J Pope[1] and David I Israel[2]

[1]Molecular Discovery Research, GlaxoSmithKline, Collegeville, Pennsylvania;

[2]Molecular Discovery Research, GlaxoSmithKline, Waltham, Massachusetts;

Keywords

DNA encoded chemical library

Affinity selection

Drug discovery

Combinatorial chemistry

G-protein coupled receptor

Cell-associated target protein

Abstract

DNA-encoded small-molecule library technology has recently emerged as a new paradigm for identifying ligands against drug targets. To date, this technology has been used with soluble protein targets that are produced and used in a purified state. Here, we describe a cell-based method for identifying small-molecule ligands from DNA-encoded libraries against integral membrane protein targets. We use this method to identify novel, potent, and specific inhibitors of NK3, a member of the tachykinin family of G-protein coupled receptors (GPCRs). The method is simple and broadly applicable to other GPCRs and integral membrane proteins. We have extended the application of DNA-encoded library technology to membrane-associated targets and demonstrate the feasibility of selecting DNA-tagged, small-molecule ligands from complex combinatorial libraries against targets in a heterogeneous milieu such as the surface of a cell.

Introduction

With the recent advent and description of DNA-encoded small-molecule library technology (ELT) (1-5), library collections containing billions of unique compounds have been used to identify potent and selective ligands against biological targets. These libraries are synthesized using combinatorial methods and exist as mixtures containing millions to billions of different compounds per library. Ligands that bind to a target protein are identified using affinity based selection methods from the mixture of compounds that comprise the library, and the chemical structures of selected compounds are deduced from the sequence of the attached DNA (1). While massively parallel DNA sequence analysis generates a large data set of selected molecules, only a small number of the more enriched and chemically interesting small-molecule structures are synthesized discretely and tested for activity against the target. Rapid screening of mixtures (selection) coupled with focused synthesis greatly reduces the time and cost of interrogating the DNA-encoded library collection compared to many conventional screening methods.

Previously, encoded small-molecule libraries have been used with soluble, tagged forms of target proteins used at high concentrations and in a purified state (1-8). Many drug targets, however, are integral membrane proteins that reside on the cell surface, such as G-protein coupled receptors (GPCR) and ion channels. The conventional means to discover ligands against these types of drug targets is high throughput screening (HTS), where at most several million compounds can be interrogated for activity against a target. Adapting ELT methods developed for use with

soluble protein targets to integral membrane protein targets has been extremely challenging. Detergent solubilization of integral membrane proteins often results in denaturation and/or aggregation of the protein, making their purification in a native and functional conformation nearly impossible. Our attempts using detergent solubilized GPCRs as ELT targets have all yielded negative results. Furthermore, it is difficult to express integral membrane proteins at a sufficiently high level and in a native state that they comprise the majority of protein in a cell membrane. Svensen et al. have described selections of a PNA-encoded peptide library against cells overexpressing $\alpha_v\beta_5$ and $\alpha_v\beta_3$ integrins (9) and the chemokine receptor CCR6 (10). This library, however, contained only 10,000 different peptidic compounds and is unlikely to generate starting points for small molecule drug discovery. Several examples of ligand discovery from small molecule DEL libraries against soluble domains of cell surface proteins have also been reported (11, 12). This strategy might be successful for a subset of membrane proteins, particularly those with large extracellular domains and a single transmembrane domain. However, it is unlikely that integral membrane proteins containing multiple transmembrane domains can be successfully interrogated using this strategy.

Here, we describe a cell-based selection method for identifying high affinity and selective ligands against cell surface targets from small-molecule DNA-Encoded Libraries (DELs). We used this method to interrogate a library collection containing approximately 15 billion compounds and identified multiple chemotypes of antagonists with high affinities and specificity to the NK3 receptor, a member of the tachykinin family of GPCR. We provide detailed pharmacological characterization of several novel chemotypes that are potent and specific antagonists of the NK3 receptor.

Results

Summary of the method. ELT is a robust methodology for identification of biologically active low molecular weight compounds that bind to purified biochemical targets (1-8). One objective of our continuing development of the technology has been to apply it to membrane-associated targets. We modified the previously described methods for soluble protein targets to integral membrane targets expressed on the cell surface. Briefly, after evaluation of various host cell/vector expression systems and selection protocols, we developed an optimized protocol that uses HEK293-derived cells transduced with recombinant BacMam viruses encoding the targets of interest. Intact, viable cells were used in several rounds of selections to purify molecules in the DELs that bind to the cell surface. Cells transduced with target-negative control viruses were used as non-target controls to identify DEL molecules that bind to endogenous cell surface molecules. Informatics tools were used to subtract cell-binding molecules present in the target-negative sample from the molecules obtained against cells expressing the recombinant GPCR target, revealing a population of target-specific ligands. Compounds lacking the encoding DNA tags were then synthesized and tested for biological activities in biochemical or cell-based functional assays. This method is rapid, simple and broadly applicable to various cell surface proteins.

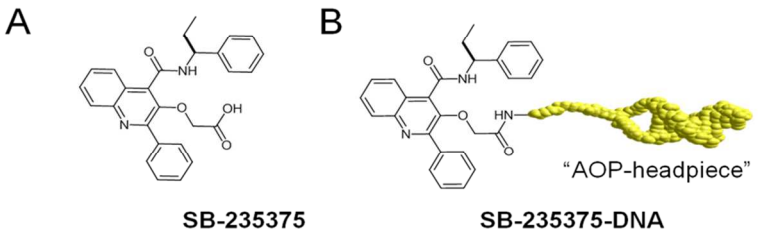
Methods development using a DNA-tagged tool compound. We chose the tachykinin receptor neurokinin-3 (NK3) to develop and optimize a cell-based ELT selection method for integral membrane proteins because of its relevance as a therapeutic target, and the availability of small-molecule tool compounds, expression

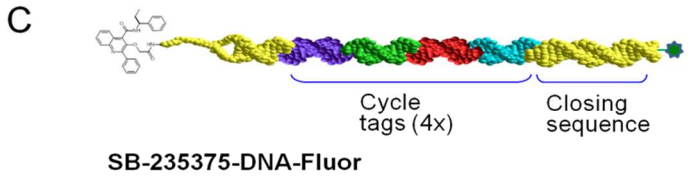
vectors, and quantitative binding assays. The tachykinin receptor family and their natural neuropeptide ligands substance P (SP), neurokinin A (NKA) and neurokinin B (NKB), are described in detail elsewhere (13-16). A number of non-peptide tachykinin receptor antagonists have been identified as research tools and clinical candidates (17-20).

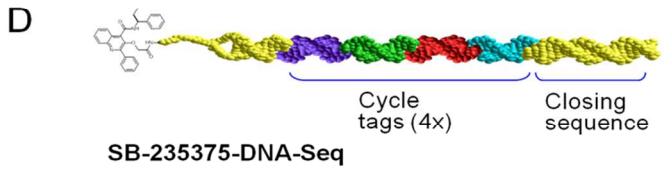
We first transduced HEK293 cells with a recombinant BacMam virus encoding NK3 at various virus to cells ratios. Receptor expression levels on the cell surface were measured with a filter binding assay using ^{125}I-NKB. Cells expressing approximately 500,000 receptors per cell, the highest level achieved with this expression system for NK3, were used in protocol optimization and selection experiments. HEK293 cells transduced with a NK3-negative BacMan virus were used as non-target control (NTC) in parallel selections.

Using HEK293 cells expressing the maximum obtainable levels of NK3, we optimized and validated a general cell-based ELT selection method with a tool compound synthesized by conjugating a known NK3 antagonist to DNA. **SB-235375-DNA** (Fig. 1*B*) is a modified, DNA-tagged derivative of a quinoline NK3 antagonist previously used in pre-clinical studies of disease modulation by NK3 antagonists (21). SAR for this antagonist series (22) suggested the carboxylic acid group is a rational point of attachment to the starting DNA molecule of ELT libraries (1). While **SB-235375-DNA** had slightly lower affinity to NK3 than the starting compound (K_i of 20 nM in competition binding assay and compared to 3 nM of unmodified **SB-235375**), it nevertheless remained a high affinity NK3 ligand. The DNA portion of **SB-235375-DNA** was further modified by ligation with a fluorescence-labeled oligonucleotide or

alternatively, with a double-strand DNA (dsDNA) fragment similar in size to the DNA tags used in DEL library molecules. These DNA-tagged tool compounds provided sensitive reagents for use in development and optimization of cell-based ELT selection methods. The resulting **SB-235375-DNA-Fluor** (Fig. 1*C*) and **SB-235375-DNA-Seq** (Fig. 1*D*) were used in fluorescence binding assays or as a sequencing probe in selection experiments, respectively**. SB-235375-DNA-Fluor** was added to $1x10^7$ cells expressing NK3 or NTC, and cell-based selections were performed as described in the Methods section. Fluorescence intensities were measured on samples after exhaustive washing. Typically, about 10-15 fold higher fluorescence intensities were obtained from cells expressing NK3 than that with target-negative control cells, suggesting an enrichment of compound **SB-235375-DNA-Fluor** that is NK3 specific.







E

fold dilution into DEL35 (1 in x)	1.0×10^6	5.0×10^6	2.5×10^7
enrichment factor	2500 ± 800	3400 ± 1900	3300 ± 2000

Figure 1. Positive target specific enrichment of DNA-tagged tool compounds with a cell-based ELT selection protocol. (A) a quinoline NK3 antagonist. (B) a modified, DNA-tagged derivative of SB-235375. The DNA portion of **SB-235375-DNA** was further modified by ligation with a fluorescence-labeled oligonucleotide (C) or with a double-strand DNA (dsDNA) fragment similar in size to the DNA tags used in DEL library molecules (D). **SB-235375-DNA-Seq** was spiked into a DEL library at different ratios and subjected to three rounds of selection. The recovered library DNA molecules were PCR amplified and sequenced to decode the dsDNA tags. The enrichment factors for

SB-235375-DNA-Seq, i.e. fold increase of copy counts relative to all DEL molecules in the sample, was calculated and summarized in the table (E).

SB-235375-DNA-Seq was next spiked into a DEL library at different ratios (chosen to mimic the concentration of an individual member in DELs), and subjected to three rounds of selection. Typically, a total of $1.5x10^{15}$ DEL molecules (2 nMoles of library) were used as input, and approximately 0.1-1% of library molecules were recovered after each round of selection. After three rounds of selection, between $1x10^{7}$ and $1x10^{9}$ DEL molecules were typically recovered. The frequencies of DEL library molecules and **SB-235375-DNA-Seq** were measured after PCR amplification and next generation DNA sequence analysis. The ratios of **SB-235375-DNA-Seq** to DEL molecules allowed us to calculate enrichment of the DNA-tagged tool compound relative to naïve library. The enrichment factors for **SB-235375-DNA-Seq**, i.e. fold increase in copy counts relative to all DEL molecules in the sample, were calculated. Fig. 1*E* summarizes the outcome of 5 independent selection experiments with **SB-235375-DNA-Seq** spiked into DELs, and demonstrates reproducible enrichment of the tool compound that is independent of the concentration of the tool in the starting DELs. The average enrichment factor was approximately 3000-fold after three rounds of selections. This is equivalent to an enrichment factor of approximately 15 per round for each of the three rounds of selection, a value similar to the 10-15 fold signal observed with **SB-235375-DNA-Fluor** on NK3-positive versus NK3-negative cells using a fluorescence binding assay. The finding is also consistent with the principle that the ELT selection was driven by the concentrations of target, i.e. NK3 receptor in this case, but not the

concentration of ligand, as the enrichment factor is independent of the amount of **SB-235375-DNA-Seq** spiked into the sample.

Selection of novel NK3 ligands. We next conducted cell-based ELT selections with a number of DELs, ranging in diversity from $5x10^5$ to $4x10^9$ unique small-molecule compounds per library. After three rounds of selection, DNA was amplified by PCR, sequenced, and decoded as described previously (1). DEL35, a three-cycle library based on a triazine scaffold (Fig 2*A*), was among the set of DELs employed. We visualize a selection outcome as a cubic scatter plot (Fig 2B). Each axis of the plot represents one cycle of synthesis, and contains a number of points equal to the number of building blocks used in that cycle of library synthesis. Thus, each unique molecule of the three-cycle DEL35 library is represented at a unique location in the cubic space. This visualization allows families of structurally-related molecules to appear as lines (two building blocks in common) or planes (one building block in common). NK3-specific patterns (called “features” and highlighted in RED color) became visible in the DEL35 selection cube after single copy molecules are removed from the visualization. The cubic scatter plots from affinity selections against NK3 or NTC are shown in Fig. 3*A*

and 3*B*, respectively.

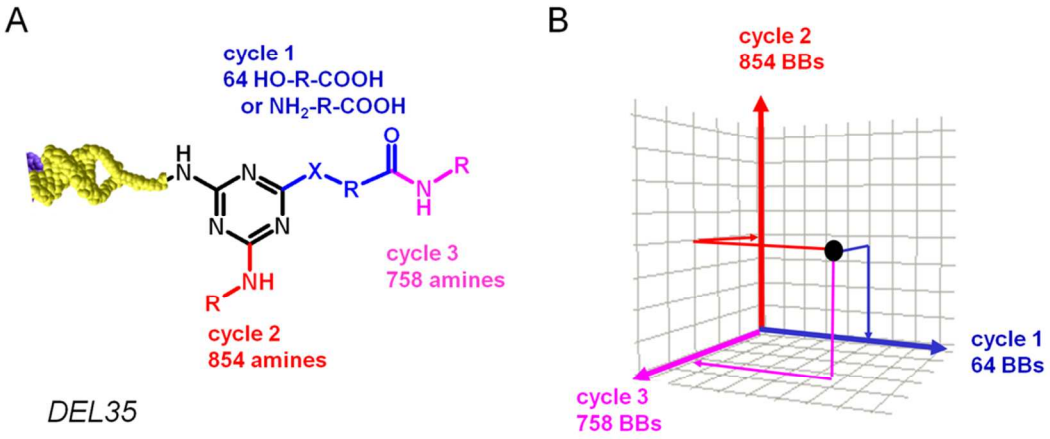


Figure 2. (A) DEL35 is a library consisting of 64, 854, and 758 building blocks (BBs) at cycle 1, 2 and 3, respectively, yielding a total of 41 million unique compounds on DNA. Methods for library synthesis are described in detail in **Supplemental Data**. (B) Interpretation of ELT selection outcomes for DEL35 as selection cube where a dot represents a unique compound, with the building block for individual cycle mapped to the corresponding axis. Dots are scaled by the frequency the specific encoding tag sequence is found in the DNA sequencing output.

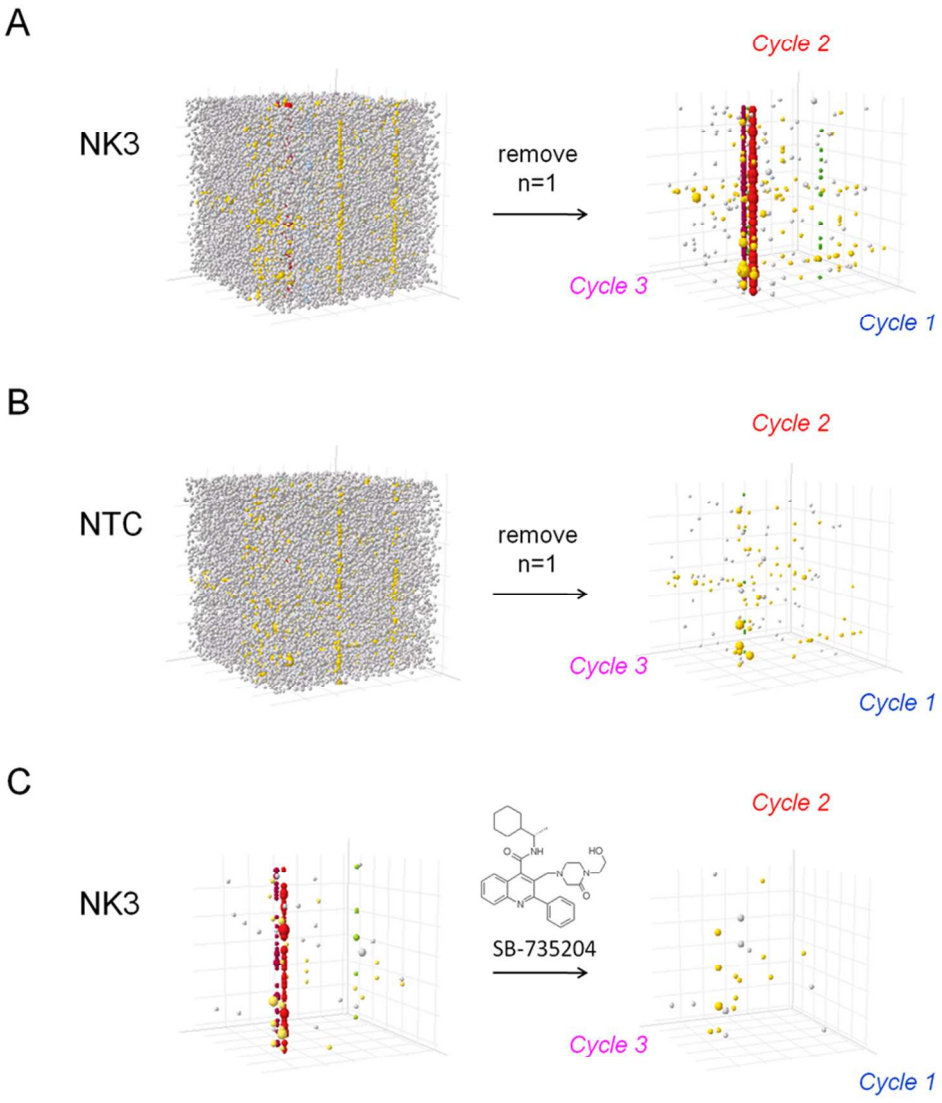


Figure 3. Results of DEL35 whole cell selection experiments with cell expressing NK3 or NTC. A total of 34,746 and 28,637 unique DEL35 molecules were identified and visualized in the cubes for selection against NK3 (A, left) or NTC (B, left), respectively. Some library molecules are more abundant than others in the selected population, as reflected by their corresponding DNA sequence appearing more often in the selection outcome. In this dataset the probability due to random chance of having 2 or more copies of the same warhead is less than 1 in one million (23). In order to more easily

visualize selected molecules with a copy count greater than one, the single copy molecules were removed from visualization, revealing several prominent lines ("features") in the NK3 selection (A, right), but absent in NTC selection (B, right). DEL35 features reproduced and competed away by a known NK3 antagonist (C).

The features consisting of a set of parallel lines (Fig. 3*A*, right) demonstrate that related molecules are more abundant in the population of molecules selected against NK3-expressing cells. These putative NK3-binding pharmacophores are composed of several specific combinations of library building blocks at synthesis cycle 1 and 3 with many building blocks tolerated at cycle 2.

To elucidate the binding site of the selected DEL molecules, the experiment was repeated in the presence or absence of a known NK3 antagonist from the same structural series (SB-735204, Ki of 10 nM). The selection results against NK3 were highly reproducible, with the same features observed again when no inhibitor was added (compare Fig. 3*A* right and 3*C* left). In contrast, SB-735204 completely eliminated the NK3-specific features in the selected DEL35 population (Fig. 3*C*, cube on right). The competition of these lines by SB-735204 suggested that the library molecules that comprise these selected features may directly interact with NK3 at the same site as the known inhibitor.

The DEL35 NK3 selection was repeated with more in-depth sequencing (to a total of about 114,000 unique compounds identified). The resulting DEL35 selection cube (Fig. 4) shows additional clearly-defined features and allowed us to formulate

selection SAR hypotheses. The most prominent features in each cubic scatter-plot are highlighted in shades of RED with GREEN, YELLOW, and BLUE highlight indicating decreasing levels of copy count of each feature.

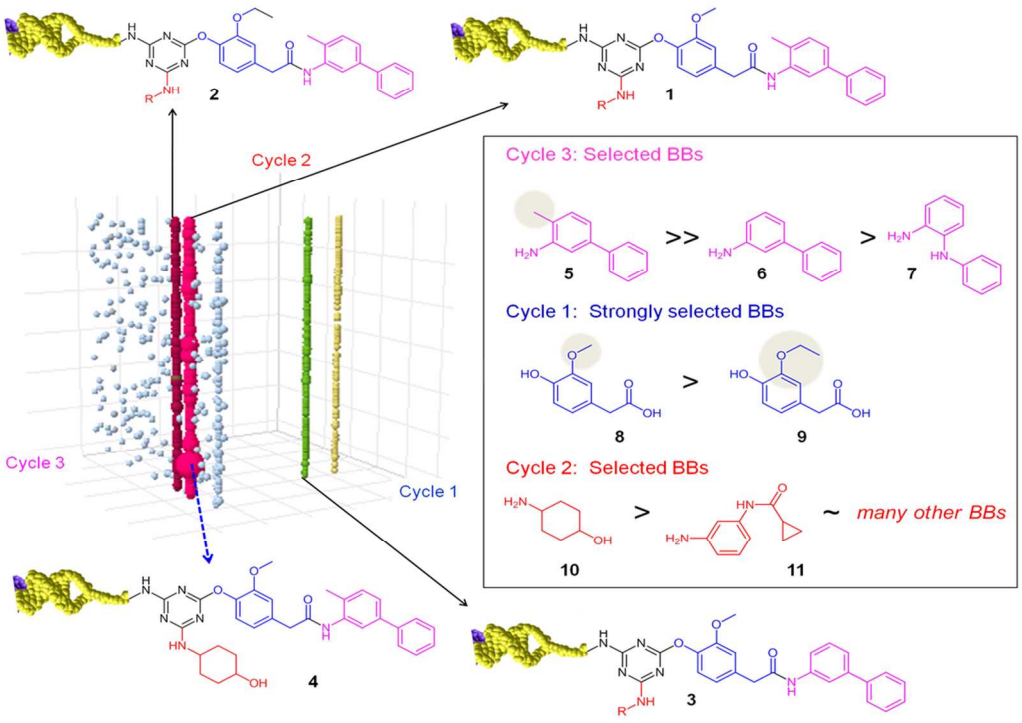


Figure 4. Features from cell-based ELT selection provide testable SAR hypotheses for NK3-binding pharmacophores. Each individual dot in the cube represents a DEL35 molecule, while the size of the dot corresponds to the total copy number identified for that library molecule. Three types of features are seen in the cube: a plane, prominent lines (exemplified as **1**, **2** & **3**), and high-copy-number dots (exemplified by **4**). The plane and lines highlight shared building blocks. For example, dots in the light BLUE plane share the same cycle 3 building block **5**. In a colored line, all dots share the same cycle 1 and cycle 3 building blocks.

The line features (RED, GREEN, YELLOW) suggested only a small number of cycle-1 and cycle-3 building blocks were selected with NK3-positive cells, while variability in cycle-2 was more widely tolerated. The plane feature (BLUE highlight) defined by cycle-3 building block **5** suggests this building block is important for NK3 affinity, and two of the most prominent feature lines (RED highlight) lie on this plane. The encoded structure **1** of the most prominent line predicts the combination of the highly preferred cycle-3 building block **5** and the cycle-1 building block **8** may be especially important for binding. Structure **2**, which corresponds to the other feature line on the same plane, differed from structure **1** only by a single methyl group in cycle-1 building blocks (compare **8** with **9**). For cycle-3, in addition to building block **5**, building blocks **6** (GREEN line) and **7** (YELLOW line) were the only other prominent building blocks in the NK3 selection among a total of 758 cycle-3 amines. Interestingly, these cycle-3 building blocks are also structurally related. Several cycle-2 building blocks, such as building blocks **10** and **11**, occurred more often than the rest of the cycle-2 building blocks found in the line encoding structure **1**, suggesting certain cycle-2 building blocks could further enhance the affinity to NK3. Structure **4** is a compound representing a combination of highest copy count building blocks from all three cycles, building block **8**, **10** & **5** from cycle-1, -2 and -3, respectively. In contrast to cycles 1 & 3, almost all 854 cycle-2 amines were found in the line with encoded structure **1,** allowing us to predict that cycle-2 building blocks are not required for binding to NK3.

Synthesis and characterization of selected small-molecule compounds. A small set of compounds were synthesized, without the DNA tags and subjected to

biochemical and functional assays. These compounds have biological activity and the assay results support the SAR hypotheses. Figure 5 summarizes the potency data for a representative set of compounds tested in a ^{125}I-NKB competition binding assay and cell-based NK3, NK2, and NK1 functional FLIPR assays. There were consistent potency trends amongst the compounds in both binding and functional assays. All of the compounds listed in Figure 5 were characterized as functional antagonists. No agonist or positive modulator activities were observed with any of these compounds. Compound **12**, which corresponded to the on-DNA compound **4** and had the highest copy count in the NK3 selection, demonstrated low nano-molar potency in both the NK3 binding and functional cell-based assays. In addition, greater than 1000-fold selectivity for NK3 was seen versus the related NK1 and NK2 receptors. Furthermore, the compound that corresponds to the library molecule with the second highest copy count, compound **13**, was nearly as potent and selective. Consistent with the hypothesis based on selection SAR, a cycle-2 building block was not required for NK3 activity in a functional assay (compound **14**; fKi 25 nM). Furthermore, the group that mimics the attachment point to the DNA construct (compound **15**; fKi 160 nM) was dispensable for activity. These truncations provided lower molecular weight compounds that remained potent NK3 antagonists and retained selectivity versus NK1 and NK2. An additional loss of 10-fold potency was seen with compound 16 (a simple combination of strongly selected building blocks 8 and 5 and lacking the library scaffold). The SAR trends predicted by the selection cube are largely consistent with rank order potency of the “off-DNA” compounds, i.e. compound **14** vs **17** corresponding to cycle-3 building block **5** vs **6** and compound **14** vs **18** corresponding to cycle-3 building block **8** vs **9**, respectively.

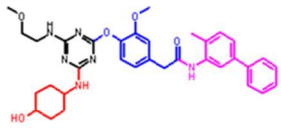
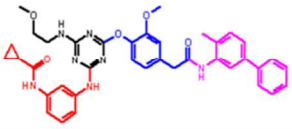
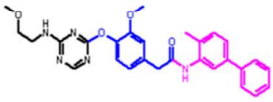
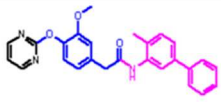
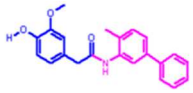
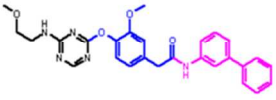
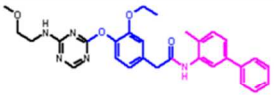
Structure	Compound	Activities on NK3 Receptor			Selectivity	
		NK3 Ki (nM)	NK3 fKi (nM)	LE	NK1 fKi (nM)	NK2 fKi (nM)
	12	1.8	0.6	0.28	>5000	625
	13	3.8	4.5	0.23	>5000	>5000
	14	190	25	0.28	3,700	1,300
	15	900	160	0.29	2,800	1,700
	16	n/a	1,700	0.30	3,200	1,700
	17	890	220	0.25	980	1500
	18	n/a	42	0.27	730	270

Figure 5. Activity of off-DNA exemplars support SAR hypotheses derived from selection cube. Ki values for each compound were determined in competition binding assays, while fKi values were calculated from IC50 values determined in FLIPR functional assays as described in Experimental Procedures.

Finally, a FACS assay was used to demonstrate direct binding of a fluorescently-labeled DNA conjugate of compound 12 to NK3-positive cells (Figure 6). The DNA conjugate of compound **12-DNA** was made fluorescent by ligation of a 5'-FAM-labeled

DNA oligonucleotide, providing **12-DNA-Fluor** (Figure 6*A*). NK3-expressing cells probed with **12-DNA-Fluor** showed greater than 10-times the fluorescent signal than the parental NK3-negative cell line probed with the same molecule (Figure 6*B*), similar to that observed with the tool compound **SB-235375-DNA-Fluor**. This result demonstrates that binding of **12-DNA-Fluor** requires the presence of NK3 on the cell surface, and recapitulates enrichment of this molecule from the library with a pure molecule using a whole cell fluorescence binding assay.

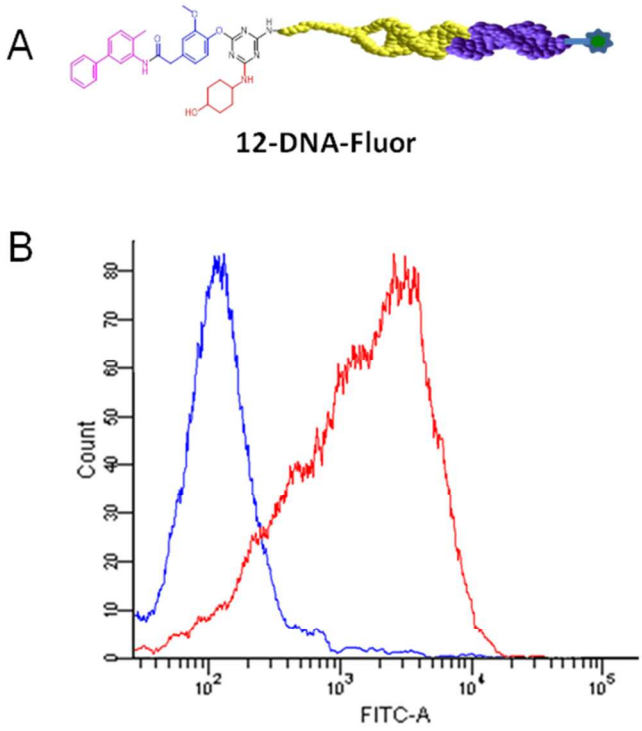


Figure 6. FACS assay demonstrates direct binding of a DNA-tagged DEL35 molecule to NK3 expressing cells. A fluorescent probe **12-DNA-Fluor** (A) was used in the binding experiment on NK3 positive (red histogram) or NK3 negative (blue histogram) cells (B).

Cells were analyzed by FACS. The mean channel fluorescence of NK3 positive cells was 12-fold greater than control NK3 negative cells (2921 versus 244, respectively).

Identification of additional NK3 antagonist chemotypes. To determine if ligands derived from other DNA-encoded libraries could be identified against NK3, additional selection experiments were performed. Approximately 30 libraries containing greater than 15 billion compounds in total were selected against NK3. NK3-specific features were observed from multiple libraries with various scaffolds. Representative small-molecules were synthesized off-DNA and tested for NK3 receptor activities. Fig. 7 highlights representative features and active molecules found from three additional libraries. These chemotypes are structurally-distinctive from each other and from the original DEL35 feature. Full descriptions of these additional DNA-encoded libraries will be communicated at a later time.

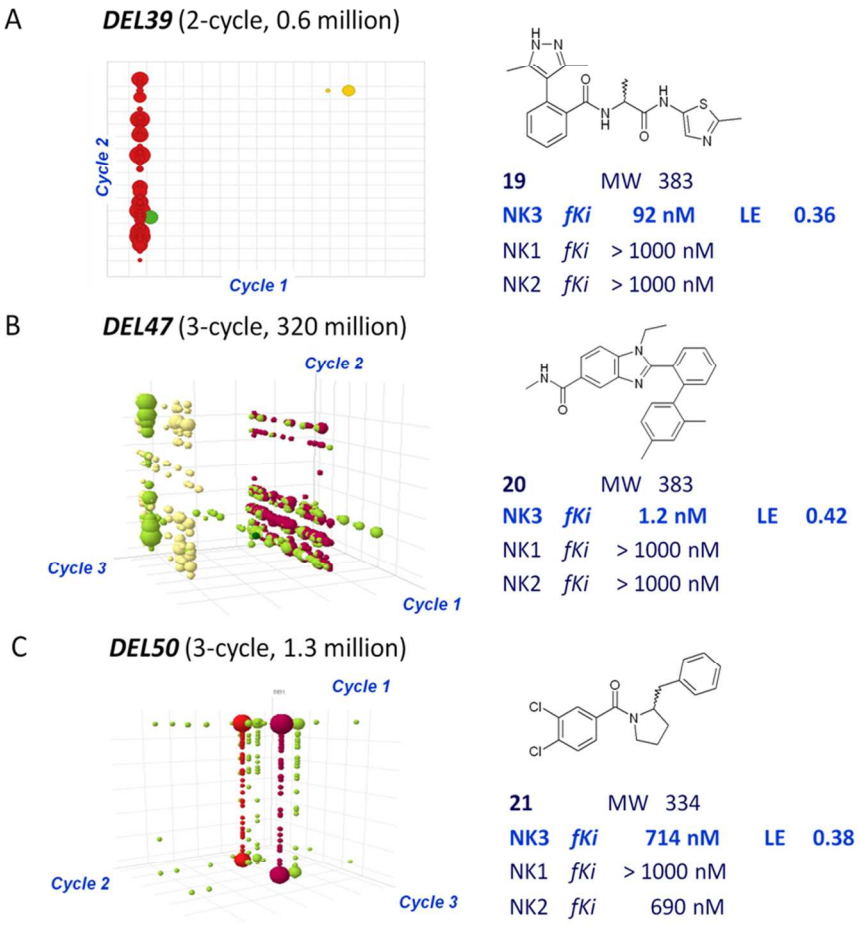


Figure 7. Additional NK3-specific features from cell-based ELT selection provide novel NK3 antagonist pharmacophores. From libraries DEL39 (A), DEL47 (B) and DEL50 (C), representative selection features and small-molecule feature exemplars are shown. NK3 antagonist and tachykinin receptor selectivity data are also summarized.

Selection with the 2-cycle library DEL39 (600,000 unique compounds) provided the 2-D plot shown in Fig. 7*A*, indicating a single cycle-1 building block was strongly preferred among approximately 2,000 building blocks used in the library synthesis. Compounds containing this cycle-1 building block and a few selected cycle-2 building

blocks were prepared to confirm that the 3,4,5-substituted pyrazole (cycle-1 building block) was the critical pharmacophore driving NK3 receptor binding. Potency and selectivity data for compound **19**, is also shown in Fig. 7*A*. Similarly, features were identified from DEL47, a 3-cycle, 320 million member library based on the benzimidazole core. Guided by selection SAR (Fig. 7*B*), compound **20** was prepared. Low nanomolar potency, drug-like ligand efficiency (24), and excellent selectivity against NK1 and NK2 were found for this compound.

Additionally, a set of features from DEL50, a 3-cycle library with 1.3 million unique compounds, yielded compound 21 (Fig. 7*C*). This compound represents a distinct chemotype and is also a ligand efficient NK3 antagonist.

Summary

The DNA-tagged tool compounds and proof of concept experiments described herein enabled us to establish a simple protocol for cell-based affinity selection using DNA-encoded combinatorial libraries. Selection on cells expressing NK3 yielded enriched features encoding multiple series of novel chemotypes. Structure activity relationships, deduced from the selection outputs, effectively guided decision-making regarding which compounds to synthesize and test for activity. Multiple series of potent, ligand-efficient, and selective NK3 antagonists were identified that are structurally-distinct from known NK3 antagonists (25, 26), and to our knowledge have not been reported in the literature.

Initial selection experiments with DEL35 enabled proof of concept and provided potent and NK3-selective triazine-based antagonists **12**-**18**. Thereafter, hit confirmation efforts focused on features found from structurally-diverse, non-triazine 2- and 3-cycle libraries. The lower molecular weight and higher ligand efficiency (24) make these chemotypes more attractive starting points for medicinal chemistry. The majority of these new chemotype exemplars (i.e. **19**, **20**, **21**) have ligand efficiency values that are in excess of 0.3 and thus consistent with quality leads and drugs (see Fig. *S*8 in Supplemental Data). The potency/molecular weight profile of the DEL47 exemplar, compound **20,** compares well to that of Talnetant **22** and Osanetant **23**, two NK3 antagonists with the most clinical studies for disease interventions (18).

The described method is based on affinity enrichment of DNA-tagged compounds on a target of interest. Since the concentration of any individual compound in the DEL library is very low, often in the pM or fM range, binding of a DNA-linked compound to a target protein is driven by target concentration, cell surface receptor density in this case. As such, a critical factor for successful enrichment is the receptor expression level. Using a DNA-tagged NK3 antagonist as a tool compound, we determined the minimum level of NK3 expression for successful enrichment was around 25,000 receptors per cell, an expression level that is achievable for most GPCRs. The enrichment of target-specific DNA-tagged compounds was directly correlated to the surface receptor density (see Fig.*S*9 in ***Supplemental Data***).

Conclusions and Implications

Identification of high quality and novel chemical templates as lead compounds is essential for the success of small molecule drug discovery. Most lead discovery efforts are carried out with traditional HTS campaigns that utilize biochemical or cell-based functional assays, in micro-titer plates with one compound per well. In this paper, we describe a cell-based selection methodology and representative screening results of DNA-encoded small-molecule library technology against a G-protein coupled receptor, NK3. To our knowledge, this is the first reported application of small molecule DNA encoded library technology to integral membrane protein targets, such as GPCRs, and represents a new paradigm for discovery of low molecular weight ligands that bind to integral membrane proteins.

Our results illustrate the exciting potential of ELT for membrane based targets, such as GPCRs (27). Combined with the very simple BacMam expression technology (28, 29), the selection with multiple DEL libraries can be done quickly with target expression on intact cells and without the requirement of protein purification. The selection method is simple and easily amendable for modification to include additional selection conditions, such as co-factors or competitors. Libraries with combined diversity of billions or more can be selected simultaneously.

Since the original proof-of-concept experiment with NK3, a number of other GPCRs have gone through the selection process. Novel chemotypes were successfully identified for some of these receptors, ranging from chemokine receptors, peptides/protein hormone receptors to lipid sensing receptors. Each receptor yielded a unique set of enriched features that were specific to that receptor. Off DNA resynthesis and testing demonstrated that all compounds were antagonists. Interestingly, to date we

have had little success with receptors whose natural ligands are low molecular weight compounds, such as the aminergic receptor sub-families in the family A of GPCR. This is in contrast with our experience with traditional plate-based HTS, in which high hit rates were often found with aminergic receptors. The cause of this discrepancy is not fully understood, although compounds with structures similar to antagonists of amine receptors are expected to be widely present in the DEL libraries.

Not surprisingly, all compounds we identified here were antagonists and none were active as agonists or positive modulators. This is analogous to our experience with other lead identification strategies and is probably due to the difficulty of mimicking biological activity of a peptide ligand with low molecular weight synthetic compounds. However, encoded library technology should be able to identify agonist ligands if contained in the library and suitably presented to cells. For example, we were able to significantly enrich a DNA-tagged tool compound prepared from a known agonist peptide ligand of NK3 using the same cell-based selection methodology described here (unpublished data), demonstrating that an agonist-DNA conjugate can bind to the receptor.

The NK3 antagonists discovered from DEL35, DEL39, DEL47 and DEL50 (compounds **12**, **19**, **20**, **21**, respectively) are structurally distinct from one another. A retrospective analysis revealed that no closely related structural analogs of these compounds were present in the approximately 2 million member screening collection in use at that time (entire HTS set < 0.85 Tanimoto similarity using Chemaxon chemical fingerprints) (30). Our results with the on-DNA tool compound SB-235375-DNA-Seq (Figure 1) demonstrate that a DNA conjugate derived from Talnetant is capable of

binding to NK3. However, we did not discover analogues of Talnetant (or Osanetant) in our ELT screen. In this regard, ELT and traditional plate-based HTS methods are complementary hit identification approaches because different chemical templates were found with each platform.

All of the chemotypes described here represent primary hits directly selected from the DEL library collection (or are lower molecular weight, truncated derivatives) and are un-optimized by medicinal chemistry. Several of these compounds have potency, specificity and ligand efficiency properties comparable to Talnetant and Osanetant (i.e. DEL35-derived **12** and DEL47-derived **20**). These results demonstrate the capability of ELT to efficiently provide high quality, lead- and drug-like molecules against integral membrane protein targets.

Experimental Procedures

Preparation of DNA-encoded tool and library molecules. For the synthesis of all DNA-encoded tool compounds and library molecules, a common AOP-headpiece construct was employed as starting material. The structure, synthesis, characterization, and several common synthetic protocols used to modify this construct were reported previously (1). Briefly, DNA-tagged NK3 antagonist tool compound **SB-235375-DNA** was obtained by DIC/HOAt mediated coupling in water of the known NK3 antagonist SB-235375 to the AOP-headpiece. Synthesis of a 4-cycle triazine-based library has already been described in our earlier work (see DEL-B in reference 1). The simpler 3-cycle DEL35 library (Fig. 2*A*) used here was prepared using only the last three of the

four encoded diversity cycles of DEL-B. Triazine was linked to the DNA headpiece through a short PEG-based spacer. Cycle-1 and cycle-2 building blocks were directly linked to the other two coupling sites on triazine. Cycle 1 contains 64 amino acid or phenolic acid building blocks, while cycle 2 contains 854 amines. Cycle 3, which was linked to cycle 1, contains 758 amines, giving a library with a theoretical total of 41,429,248 components. Validation of library building blocks and installation of the DNA-encoding of DEL35 were done with only minor modification to the published methods and protocols (for DEL35 reaction scheme and more details, see **Supplementary Methods**).

Preparation of cells. HEK293 cells were grown as a monolayer culture and transduced with recombinant BacMam virus (31) encoding NK3 receptors at various virus to cell ratios. After 24 hours, cells were harvested with Versene® treatment and receptor expression levels on the cell surface were determined with a filter binding assay using radioactive-labeled NKB as described below. For cell-based DNA-encoded small-molecule library selection and competition binding assays, HEK293 cells transduced with the appropriate virus to cell ratios were prepared as described above.

Cell-based DNA-encoded small-molecule library selection. Selection experiments were started with 1.5×10^{15} of input DEL molecules with or without the on-DNA tool compound spike-in. DEL library (either a single library or a pool of libraries) was incubated with 10^7 cells at 37°C for 30 minutes in 1 mL PBS containing 0.1% (g/100 mL) sodium azide to inhibit receptor internalization. A blocking agent, such as sheared salmon sperm DNA, was also included at a final concentration 1 mg/mL to reduce nonspecific binding. Cells were then washed with ice cold PBS (1 mL) ten times by

gentle centrifugation and re-suspension to remove unbound library molecules. Bound molecules were eluted by heating cells in PBS to 95^{o}C for 10 minutes, followed by centrifugation at 14,000 rpm to remove cells. The recovered DEL molecules in the supernatant were used for the next round of selection with freshly prepared cells using the same protocol. Typically, this process was repeated one more time for a total of three rounds of selection. For competitive selections, known NK3 antagonist SB-735204 was added to the selection buffer to a final concentration of 10 μM at each selection cycle, and the selection was performed using the same method.

DNA sequencing, tag decoding, and data visualization. The amount of total DEL library recovered after each round of selection was measured by quantitative polymerase-chain-reaction (qPCR) using a DEL specific primer pair. Selection outputs after the final round of selection were also amplified using preparative PCR, which was then used for high-throughput sequencing using either the Roche 454 Genome Sequencer or Illumina Genome Analyzer. Sequences were parsed into individual libraries using a DNA library tag that was unique to each library. Each selected molecule was “translated” into the chemical building blocks used during library synthesis, encoded binding motifs structures virtually enumerated, and analyzed as described previously (1). All PCR and sequencing reactions were conducted using standard molecular biology methods as recommended by the reagent suppliers. The resulting data were visualized as a cubic scatter plot using Spotfire DecisionSite (Spotfire, Inc., Somerville, Massachusetts).

Calculation of Enrichment Factor, i.e. fold increase of copy-number relative to all DEL molecules in the sample, is calculated as: **Enrichment Factor = (*Copy-number A***

/ Total*)** $_{\textbf{after selection}}$ **/ (*Copy-number A / Total*)** $_{\textbf{before selection}}$, where ***Copy-number A is the copy-number of sequence that represent DEL molecule of interest, such as **SB-235375-DNA**, and ***Total*** is the sum of copy numbers for all variable sequences present in the samples.

Filter binding assays with radioactive-labeled NK receptor ligands. Radioactive-labeled tachykinin receptor ligands, ^{125}I-Substance P, ^{125}I-NKA and ^{125}I-[MePhe7]-NKB for NK1, NK2 and NK3 receptor, respectively, were used in the filter binding assays as described previously (32-34).

Measurement of NK functional activity. Compounds were characterized in a functional assay for calcium release in response to NK receptor ligands. Changes in intracellular Ca^{2+} concentrations were monitored using a fluorescence imaging plate reader (FLIPR, Molecular Devices Corp., Sunnyvale, California) by a modified method described previously (32, 33). The functional affinities of antagonist compounds fK_i were calculated using the following equation, where the concentration of agonist used is ***[A]***, the concentration of agonist producing 50% maximal response is EC_{50} and the ***n*** is the Hill coefficient of the agonist dose-response curve (35).

$$fK_i = IC_{50} / (2 + ([A]/EC_{50})^n)^{1/n} - 1$$

Acknowledgements

The authors wish to acknowledge Matthew Clark and Steven Skinner for guidance and help during synthesis of the DEL libraries and DNA-tagged tool

compounds and Ming Jiang for testing the compounds in competitive binding and FLIPR functional assays. The authors want to also acknowledge Barry A Morgan and Robert P Hertzberg for their sponsorship of this project.

Supporting Information (summary)

The Supporting Information section describes in detail syntheses of the DEL35 library, on-DNA tool compounds, and molecules derived from the libraries that were re-synthesized (off DNA) and tested for biological activity. There is also more detail on ligand efficiency and selection enrichment calculations as it relates to receptor expression levels, topics that are only briefly alluded to in the manuscript.

Abbreviations

NK – neurokinin

GPCR - G-protein coupled receptor

ELT – encoded library technology

HTS – high throughput screening

DEL – DNA encoded library

SAR – structure activity relationship

NTC – no target control

FACS – fluorescence activated cell sorter

FAM – fluorescein

PBS – phosphate buffered saline

References

1. Clark, M.A.; Acharya, R.A.; Arico-Muendel, C.C.; Belyanskaya, S.L.; Benjamin, D.R.; Carlson, N.R.; Centrella, P.A.; Chiu, C.H.; Creaser, S.P.; Cuozzo, J.W.; Davie, C.P.; Ding, Y.; Franklin, G.J.; Franzen, K.D.; Gefter, M.L.; Hale, S.P.; Hansen, N.J.; Israel, D.I.; Jiang, J.; Kavarana, M.J.; Kelley, M.S.; Kollmann, C.S.; Li, F.; Lind, K.; Mataruse, S.; Medeiros, P.F.; Messer, J.A.; Myers, P.; O'Keefe, H.; Oliff, M.C.; Rise, C.E.; Satz, A.L.; Skinner, S.R.; Svendsen, J.L.; Tang, L.; van Vloten, K.; Wagner, R.W.; Yao, G.; Zhao, B.; Morgan, B.A. Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat. Chem. Biol.* **2009**, *5*, 647-654.

2. Franzini, R.M.; Neri, D, Scheuermann, J. DNA-encoded chemical libraries: advancing beyond conventional small-molecule libraries. *Acc. Chem. Res.* **2014**, *47*, 1247-1255.

3. Kleiner, R.E.; Dumelin, C.E.; Liu, D.R. Small-molecule discovery from DNA-encoded chemical libraries. *Chem. Soc. Rev*. **2011**, *40*, 5707-5017.

4. Halpin, D. R.; Harbury, P. B. DNA display II. Genetic manipulation of combinatorial chemistry libraries for small molecule evolution *PLoS Biol*. **2004**, *2*, 1022– 2030

5. Hansen, M.H.; Blakskjaer, P; Petersen, L.K.; Hansen, T.H.; Højfeldt, J.W.; Gothelf, K.V.; Hansen, N.J. A yoctoliter-scale DNA reactor for small-molecule evolution. *J. Am. Chem. Soc.* **2009**,*131*, 1322-1327.

6. Yang, H.; Medeiros, P.F.; Raha, K.; Elkins, P.; Lind, K.E.; Lehr, R.; Adams, N.D.; Burgess, J.L.; Schmidt, S.J.; Knight, S.D.; Auger, K.R.; Schaber, M.D.; Franklin,

G.J.; Ding, Y.; DeLorey, J.L.; Centrella, P.A.; Mataruse, S.; Skinner, S.R.; Clark, M.A.; Cuozzo, J.W.; Evindar, G. Discovery of a Potent Class of PI3Kα Inhibitors with Unique Binding Mode via Encoded Library Technology (ELT). *ACS Med. Chem. Lett.* **2015**, *6*, 531-536.

7. Deng, H.; Zhou, J.; Sundersingh, F.S.; Summerfield, J.; Somers, D.; Messer, J.A.; Satz, A.L.; Ancellin, N.; Arico-Muendel, C.C.; Sargent Bedard K.L.; Beljean, A.; Belyanskaya, S.L.; Bingham, R.; Smith, S.E.; Boursier, E.; Carter, P.; Centrella, P.A.; Clark, M.A.; Chung, C.W.; Davie, C.P.; Delorey, J.L.; Ding, Y.; Franklin, G.J.; Grady, L.C.; Herry, K.; Hobbs, C.; Kollmann, C.S.; Morgan, B.A.; Pothier Kaushansky L.J.; Zhou, Q. Discovery, SAR, and X-ray Binding Mode Study of BCATm Inhibitors from a Novel DNA-Encoded Library. *ACS Med. Chem. Lett.* **2015**, *6*, 919-924.

8. Litovchick, A.; Dumelin, C.E.; Habeshian, S.; Gikunju, D.; Guié, M.A.; Centrella, P.; Zhang, Y.; Sigel, E.A.; Cuozzo, J.W.; Keefe, A.D.; Clark, M.A. Encoded Library Synthesis Using Chemical Ligation and the Discovery of sEH Inhibitors from a 334-Million Member Library. *Sci. Rep.*, **2015**, *5*, 10916.

9. Svensen, N.; Díaz-Mochón, J. J.; Bradley, M. Decoding a PNA encoded peptide library by PCR: the discovery of new cell surface receptor ligands. *Chem. Biol.* **2011**, *18*, 1284-1289.

10. Svensen, N.; Díaz-Mochón, J. J.; Bradley, M. Encoded peptide libraries and the discovery of new cell binding ligands. *Chem. Commun.***2011**, *47*, 7638-7640

11. Kollmann, C.S.; Bai, X,; Tsai, C.H.; Yang, H,; Lind, K.E.; Skinner, S.R.; Zhu, Z.; Israel, D.I.; Cuozzo, J.W.; Morgan, B.A.; Yuki, K.; Xie, C.; Springer, T.A.; Shimaoka, M.; Evindar, G. Application of encoded library technology (ELT) to a protein-protein interaction target: discovery of a potent class of integrin lymphocyte function-associated antigen 1 (LFA-1) antagonists. *Bioorg. Med. Chem*. **2014**, *22*, 2353-2365.

12. Buller, F.; Steiner, M,; Frey, K., Mircsof, D,; Scheuermann, J.; Kalisch, M,; Buhlmann, P.; Supuran, C.T,; Neri, D. Selection of Carbonic Anhydrase IX Inhibitors from One Million DNA-Encoded Compounds. *ACS Chem. Biol.* **2011**, *15*, 336-344.

13. Maggi, C.A. The mammalian tachykinin receptors. *Gen. Pharmacol.* 1995, *26*, 911-944.

14. Longmore, J.; Hill, R. G.; Hargreaves, R. J. Neurokinin-receptor antagonists: pharmacological tools and therapeutic drugs. *Can. J. Physiol Pharmacol.* **1997**, *75*, 612-621.

15. Sanger, G. J. Neurokinin NK1 and NK3 receptors as targets for drugs to treat gastrointestinal motility disorders and pain. *Br. J. Pharmacol.* **2004**, *141*, 1303-1312.

16. Steinhoff, M.S.; von Mentzer, B.; Geppetti, P.; Pothoulakis, C.; Bunnett, N.W. Tachykinins and their receptors: contributions to physiological control and the mechanisms of disease. *Physiol. Rev.* 2014, *94*, 265–301.

17. Joos, G. F.; Pauwels, R. A. Tachykinin receptor antagonists: potential in airways diseases. *Curr. Opin. Pharmacol.* **2001**, *1*, 235-241.

18. Dawson, L. A.; Smith, P. W. Therapeutic utility of NK3 receptor antagonists for the treatment of schizophrenia. *Curr. Pharm. Des.* **2010**, *16*, 344-357.

19. Smith, P. W.; Dawson, L. A. Neurokinin 3 (NK3) receptor modulators for the treatment of psychiatric disorders. *Recent Pat CNS. Drug Discov.* **2008**, *3*, 1-15.

20. Dawson, L.A.; Langmead, C.J.; Dada, A.; Watson, J.M.; Wu, Z.; de la Flor, R.; Jones, G.A.; Cluderay, J.E.; Southam, E.; Murkitt, G.S.; Hill, M.D.; Jones, D.N.; Davies, C.H.; Hagan, J.J.; Smith, P.W. *In vitro* and *in vivo* comparison of two non-peptide tachykinin NK3 receptor antagonists: Improvements in efficacy achieved through enhanced brain penetration or altered pharmacological characteristics *Eur. J. Pharmacol.* **2010**, *627*, 106–114.

21. Hay, D.W.; Giardina, G.A.; Griswold, D.E.; Underwood, D.C.; Kotzer, C.J.; Bush, B.; Potts, W.; Sandhu, P,; Lundberg, D.; Foley, J.J.; Schmidt, D.B.; Martin, L.D.; Kilian, D,; Legos, J.J.; Barone, F.C.; Luttmann, M.A.; Grugni, M.; Raveglia, L.F.; Sarau, H.M. Nonpeptide tachykinin receptor antagonists. III. SB 235375, a low central nervous system-penetrant, potent and selective neurokinin-3 receptor antagonist, inhibits citric acid-induced cough and airways hyper-reactivity in guinea pigs. *J. Pharmacol. Exp. Ther.* **2000**, *300*, 314-323.

22. Giardina, G.A.; Raveglia, L.F.; Grugni, M.; Sarau, H.M.; Farina, C.; Medhurst, A.D.; Graziani, D.; Schmidt, D.B.; Rigolio, R.; Luttmann, M.; Cavagnera, S.; Foley, J.J.; Vecchietti, V.; Hay, D.W. Discovery of a novel class of selective non-peptide antagonists for the human neurokinin-3 receptor. 2. Identification of (S)-N-(1-

phenylpropyl)-3-hydroxy-2-phenylquinoline-4-carboxamide (SB 223412). *J. Med. Chem.* **1999**, *42*, 1053-1065.

23. Wadsworth, G. P.; Bryan, J. G. Discrete Random Variables. In *Introduction to probability and random variables,* 1st ed.; Blackwell, D. Ed.; McGraw-Hill; New York, 1960; p. 52

24. Hopkins, A. L.; Groom, C. R.; Alex, A. (2004) Ligand efficiency: a useful metric for lead selection. Drug Discov. Today **2004**, *9*, 430-431.

25. Simonsen, K. B.; Juhl, K.; Steiniger-Brach, B.; Nielsen, S. M. Novel NK(3) receptor antagonists for the treatment of schizophrenia and other CNS indications. *Curr. Opin. Drug Discov. Devel.* **2010**, *13*, 379-388.

26. Dawson, L.A.; Porter, R.A. Progress in the development of neurokinin 3 modulators for the treatment of schizophrenia: molecule development and clinical progress. *Future Med. Chem.* **2013**, *5*, 1525-1546

27. Overington, J.P.; Al-Lazikani, B.; Hopkins, A.L. How many drug targets are there? *Nat. Rev. Drug Discov.* **2006**, *5*, 993-996.

28. Ames, R.; Fornwald, J.; Nuthulaganti, P.; Trill, J.; Foley, J.; Buckley, P.; Kost, T.; Wu, Z.; Romanos, M. BacMam recombinant baculoviruses in G protein-coupled receptor drug discovery. *Receptors. Channels* **2004**, *10*, 99-107.

29. Kost, T. A.; Condreay, J. P.; Ames, R. S. (2010) Baculovirus gene delivery: a flexible assay development tool. *Curr. Gene Ther.* **2010**, *10*, 168-173.

30. JChem Base Similarity search https://docs.chemaxon.com/display/jchembase/Similarity+search (accessed October 9, 2015).

31. Condreay, J. P.; Kost, T. A. Baculovirus expression vectors for insect and mammalian cells. *Curr. Drug Targets* **2007**, *8*, 1126-1131.

32. Sarau, H.M.; Griswold, D.E.; Potts, W.; Foley, J.J.; Schmidt, D.B.; Webb, E.F.; Martin, L.D.; Brawner, M.E.; Elshourbagy, N.A.; Medhurst, A.D.; Giardina, G.A.; Hay, D.W. Nonpeptide tachykinin receptor antagonists: I. Pharmacological and pharmacokinetic characterization of SB 223412, a novel, potent and selective neurokinin-3 receptor antagonist. *J. Pharmacol. Exp. Ther.* **1997**, *281*, 1303-1311.

33. Sarau, H.M.; Griswold, D.E.; Bush, B,; Potts, W.; Sandhu, P,; Lundberg, D.; Foley, J.J.; Schmidt, D.B.; Webb, E.F.; Martin, L.D.; Legos, J.J.; Whitmore, R.G.; Barone, F.C.; Medhurst, A.D.; Luttmann, M.A.; Giardina, G.A.; Hay, D.W. Nonpeptide tachykinin receptor antagonists. II. Pharmacological and pharmacokinetic profile of SB-222200, a central nervous system penetrant, potent and selective NK-3 receptor antagonist. *J. Pharmacol. Exp. Ther.* **2000**, *295*, 373-381.

34. Blaney, F.E.; Raveglia, L.F.; Artico, M.; Cavagnera, S.; Dartois, C.; Farina, C.; Grugni, M.; Gagliardi, S.; Luttmann, M.A.; Martinelli, M.; Nadler, G.M.; Parini, C.; Petrillo, P.; Sarau, H.M.; Scheideler, M.A.; Hay, D.W.; Giardina, G.A. Stepwise modulation of neurokinin-3 and neurokinin-2 receptor affinity and selectivity in quinoline tachykinin receptor antagonists. *J. Med. Chem.* **2001**, *44*, 1675-1689.

35. Kenakin, T .P. The Schild regression in the process of receptor classification. *Can. J. Physiol Pharmacol.* **1982**, *60*, 249-265.

Table of Contents graphical image