

DISS. ETH Nr. 22697

**DNA-ENCODED SELF-ASSEMBLING CHEMICAL LIBRARY  
TECHNOLOGY FOR *DE NOVO* LIGAND DISCOVERY AND  
TUMOR TARGETING APPLICATIONS**

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

MORENO ATTILIO WICHERT

MSc, ETH Zurich

born on 10.05.1984

citizen of Altendorf SZ

accepted on the recommendation of

Prof. Dr. Dario Neri (examiner)  
Prof. Dr. Gisbert Schneider (co-examiner)  
Dr. Jörg Scheuermann (co-examiner)

2015

## I. SUMMARY

Modern drug discovery research projects start with the identification of potent and selective binding molecules to a validated biological target of pharmaceutical interest. While traditionally such molecules have been identified by large high-throughput screening campaigns, new screening methods and innovative library technologies are increasingly being implemented in the drug discovery process. In this context DNA-encoded chemical libraries (DECLs) have emerged as a promising tool to facilitate large high-throughput compound screening campaigns and lead identification.

DECLs are composed of small organic molecules which are individually coupled to oligonucleotides serving as amplifiable identification barcodes. The DNA tagging procedure facilitates the construction and screening of compound collections of unprecedented size.

DNA-encoded self-assembling chemical (ESAC) libraries are a specific implementation of DECLs featuring the combinatorial formation of DNA heteroduplexes from two sub-libraries consisting of oligonucleotide–organic compound conjugates. Owing to the combinatorial self-assembly process, a very broad spectrum of structural diversity within the chemical space can be rapidly generated and used in a selection process to identify ligands for target proteins of choice. After an affinity capture step on an immobilized target protein, the DNA barcodes of the binding molecules are amplified by polymerase chain reaction (PCR) and subsequently identified by high-throughput DNA sequencing.

In contrast to standard fragment-based drug discovery approaches, dual-pharmacophore ESAC libraries have the potential to identify fragment pairs that bind simultaneously to the target protein of interest, benefiting from the chelate effect. However, the technology has so far been limited by the difficulty in unambiguously decoding the ligand pairs from large combinatorial libraries. This thesis presents a strategy that overcomes this limitation and enables the efficient identification of ligand pairs that bind to a target protein.

Small organic molecules were conjugated to the 5' and 3' ends of complementary DNA strands that contain a unique identifying code. DNA hybridization followed by an inter-strand code-transfer created a stable dual-display DNA-encoded chemical library of 111,100 encoded members. Using this approach we have discovered a low micromolar binder to alpha-1-acid glycoprotein (AGP) and have generated a high-affinity ligand to carbonic anhydrase IX (CAIX), a validated marker of renal cell carcinoma. The newly discovered subnanomolar CAIX binder dramatically improved tumor targeting performance *in vivo* compared to previously described CAIX ligands.

In addition to CAIX, we focused on a second promising tumor marker: prostate-specific membrane antigen (PSMA). This protein was recombinantly expressed in insect cells and isolated with high purity. Affinity maturation selection experiments with the ESAC library, containing a high-affinity lead against PSMA on one DNA strand, revealed the preferential enrichment of complementary pharmacophores. Selections with a structurally compact single-pharmacophore DECL against PSMA yielded a novel nanomolar ligand for this protein. Moreover, antibody phage-display selections led to the identification of high-affinity anti-PSMA antibodies.

The results of this thesis indicate that ESAC libraries may represent a useful source of binding specificities for pharmaceutical applications. In particular, small organic molecules capable of high-affinity binding and selective recognition of accessible target proteins as presented in this thesis may have great potential for the directed delivery of cytotoxic drugs and other payloads in cancer therapy.

## II. ZUSAMMENFASSUNG

Die Entwicklung neuer und moderner Wirkstoffe beginnt mit der Identifikation von potenten, selektiven Liganden für ein validiertes biologisches Zielmolekül von pharmazeutischem Nutzen. Während solche Wirkstoffe traditionell in Screening-Verfahren mit hohem Durchsatz identifiziert wurden, werden zunehmend neue Screening-Methoden und innovative Substanzbibliotheken in den Prozess der Wirkstoffsuche integriert. Besonders DNA-codierte Substanzbibliotheken erweisen sich als vielversprechendes Instrument, um gross angelegte Hochdurchsatz-Screening-Kampagnen und die darauf folgende Leitstrukturentwicklung zu vereinfachen.

DNA-codierte Substanzbibliotheken bestehen aus niedermolekularen Verbindungen, welche einzeln mit amplifizierbaren Oligonucleotiden verknüpft und durch diese eindeutig identifizierbar sind. Diese Methode der DNA-Etikettierung vereinfacht den Aufbau und das Screening von Kollektionen chemischer Verbindungen ungeahnter Grösse.

Sich selbst formierende DNA-codierte chemische (*engl. encoded self-assembling chemical, ESAC*) Bibliotheken sind ein Spezialfall von DNA-codierten Substanzbibliotheken. Sie entstehen aus der kombinatorischen Bildung von DNA-Heteroduplexen zweier komplementärer Teilbibliotheken. Durch die spontane Selbstassoziation kann sehr schnell eine Bibliothek von Molekülen mit grosser chemischer Diversität erzeugt werden, welche es erlaubt, in Selektionsexperimenten Liganden gegen gewünschte Zielproteine zu identifizieren. Im Anschluss an ein Affinitätsselektionsexperiment mit einem immobilisierten Zielprotein werden die DNA-Etiketten der Bindungspartner durch die Polymerase-Kettenreaktion amplifiziert und schliesslich mittels Hochdurchsatz-DNA-Sequenzierungsmethoden decodiert.

Im Gegensatz zur klassischen fragmentbasierten Wirkstoffsuche erlauben DNA-codierte Dual-Pharmakophorbibliotheken, chemische Fragmentpaare zu identifizieren, welche gleichzeitig, unter Ausnützung des Chelat-Effekts, an ein Zielprotein binden. Die Schwierigkeit dieser Technologie bestand bisher jedoch darin, die Ligandenpaare aus grossen kombinatorischen Bibliotheken eindeutig zu decodieren. Die vorliegende Dissertation beschreibt eine Strategie, diese Limitierung aufzuheben und die effiziente Identifikation von Ligandenpaaren zu ermöglichen, die an ein Zielprotein binden.

Niedermolekulare Verbindungen wurden an das 5'- und 3'-Ende von komplementären DNA-Strängen geknüpft, die einen eindeutigen Identifizierungscode enthalten. DNA-Hybridisierung gefolgt von einem Code-Transfer zwischen den beiden DNA-Strängen resultierte in einer DNA-codierten chemischen Substanzbibliothek von 111'100 codierten Verbindungen. Unter Benutzung dieses Ansatzes fanden wir einen tief-mikromolaren Bindungspartner für das Protein Orosomucoid und entwickelten einen hochaffinen Liganden gegen den validierten Nierenzellkrebsmarker Carboanhydrase IX (CAIX). Der neuentdeckte subnanomolare CAIX-Binder verbesserte die *in vivo* Tumoranreicherung im Vergleich zu bereits beschriebenen CAIX-Liganden deutlich.

Zusätzlich fokussierten wir uns auf einen zweiten vielversprechenden Tumormarker, das Prostata-spezifische Membranantigen (PSMA). Dieses Protein wurde rekombinant in Insektenzellen exprimiert und in grosser Reinheit isoliert. Affinitätssteigerungsselektionen mit der ESAC-Bibliothek und einer hochaffinen Leitstruktur gegenüber PSMA führten zu bevorzugt angereicherten komplementären Pharmakophoren. Selektionen gegen PSMA mit einer strukturell kompakten DNA-codierten Single-Pharmakophorbibliothek ermöglichte die Identifikation eines neuen nanomolaren Liganden für dieses Protein. Ausserdem wurden in Phagen-Display Selektionen hochaffine anti-PSMA Antikörper gefunden.

Die Resultate der vorliegenden Dissertation zeigen, dass ESAC-Bibliotheken eine nützliche Quelle von Bindungsspezifitäten für pharmazeutische Anwendungen darstellen. Besonders niedermolekulare Verbindungen, welche im Körper zugängliche Zielproteine selektiv erkennen und hochaffin binden, haben ein grosses Potential für die zielgerichtete Abgabe von Cytostatika und anderen Wirkstoffen in der Krebstherapie.