

# Spezifische Proteinerkennung durch Array-gebundene Hexanukleotide\*\*

Alessandra Mescalchin, Winfried Wünsche und Georg Sczakiel\*

Hexanukleotide können typische Eigenschaften von Aptameren einnehmen. Ihre spezifischen Proteinbindungseigenschaften legen nahe, sie in biologischen Studien und als pharmazeutische Leitsubstanzen zu erschließen. Die systematische Identifizierung von Hexanukleotid-Protein-Wechselwirkungen ist durch ein neues Array-Hochdurchsatzverfahren möglich. Diese Technik schließt den gesamten Hexanukleotid-Sequenzraum ein, und sie zeigt hohe Spezifität, was durch das Auffinden von Spezies unterlegt wird, die selektive Bindung an die hoch konservierten Polymerasen von HIV-1 und HIV-2 aufweisen.

Klassischerweise kann die spezifische und hochaffine Erkennung von Zielmolekülen durch gefaltete Oligonukleotide (Aptamere) von 30 bis 80 Nukleotiden Länge in SELEX-Protokollen erfolgen.<sup>[1,2]</sup> Auch ungewöhnlich kurzkettige Oligonukleotide von weniger als 10 Nukleotiden Länge können eine feste, spezifische und biologisch funktionelle Proteinerkennung eingehen.<sup>[3–5]</sup> Dies schließt die Hexanukleotide ein, für die Aptamer-vergleichbare Eigenschaften gezeigt wurden.<sup>[3]</sup> Mit Blick auf die pharmazeutische Wirkstoffentwicklung können daher Hexanukleotide als neue Klasse von Leitsubstanzen betrachtet werden. Dies wird durch Spezies unterstrichen, die virale Funktionen inhibieren können, wie am Beispiel von HIV-1<sup>[3]</sup> und HAV<sup>[10]</sup> gezeigt wurde. Zudem sind Hexanukleotide als Wirkstoffe interessant, weil ihre chemische Synthese, die Einführung chemischer Modifikationen und ihre Modellierung weit fortgeschritten und relativ einfach sind.

Sind Hexanukleotid-basierte Leitsubstanzen identifiziert, so können diese durch rationale Ansätze wie Docking-Studien<sup>[6]</sup> in Kombination mit ortsspezifischen chemischen Modifikationen oder strukturbasierten Ansätzen wie etwa der

Sättigungstransfer-Differenz(STD)-NMR-Spektroskopie verbessert werden.<sup>[7]</sup>

Hexanukleotide nehmen keine stabile Sekundärstruktur ein. Im Fall der reversen Transkriptase von HIV-1 (HIV-1-RT)<sup>[3]</sup> scheint die Erkennung von quasi eindimensionalen Oligonukleotidketten zu erfolgen, die erst nach der Bindung an ihr Ziel eine definierte Struktur einnehmen. Alternativ hierzu können Hexanukleotide auch höhere geordnete und biologisch aktive Strukturen aufbauen, wie das G-Quadruplex-bildende G<sub>5</sub>T im Fall der 3C-Proteinase von HAV.<sup>[10]</sup> Die Selektion spezifischer Hexanukleotide aus dem vollständigen Sequenzraum von 4096 (= 4<sup>6</sup>) linearen Spezies ist durch eine einzige Selektionsrunde denkbar, die erheblich weniger arbeitsintensiv ist als Methoden, die Vervielfältigungsrunden einschließen. Kombinatorische Ansätze in Lösung sind grundsätzlich von kaum kontrollierbaren Unwägbarkeiten betroffen wie die Bildung von aktiven und spezifischen höheren Komplexen aus gleichen oder unterschiedlichen Spezies, die Ungleichverteilung von Spezies, oder technische Hürden wie Klonierungsschritte, die für kurze Oligonukleotide extrem aufwendig sein können.

Da Hexanukleotide spezifisch und mit  $K_d$ -Werten im mikromolaren Bereich mit Proteinen wechselwirken können,<sup>[3]</sup> sind kombinatorische Ansätze mit dem vollständigen Hexanukleotid-Sequenzraum in Array-Form attraktiv, weil Nachteile von Protokollen in Lösung umgangen werden. Weitere Vorteile von Array-Ansätzen sind eine einfache Methodik, niedrige Kosten und computerbasierte und quantitative Analysen der Ergebnisse.

Im Folgenden werden der Entwurf und die Eigenschaften eines Hexanukleotid-Arrays beschrieben. Zunächst wurde ein Versuchsarray analysiert, das Oligonukleotide mit bekannten Bindungseigenschaften in Lösung zu HIV-1-RT enthielt<sup>[3]</sup> (siehe die Hintergrundinformationen, Tabelle S1). Hierzu zählten RT-bindende und nichtbindende Hexanukleotide, langkettige Oligonukleotide, die als Templatstrang an Polymerasen binden, sowie ein RT-gerichtetes DNA-Aptamer.<sup>[8]</sup> Alle Oligonukleotide wurden über nichtnukleosidische Linker mit variierender Länge an die feste Array-Oberfläche gebunden (Abbildung S1a), um unerwünschte zusätzliche und unspezifische Wechselwirkungen mit der Zielstruktur auszuschließen. Zunächst wurden Desoxyribonukleotide gewählt, die bisher vergleichbare Bindungseigenschaften gezeigt hatten wie Ribonukleotide und nur geringfügig von anderen 2'-OH-modifizierten Strängen abweichen. Dagegen zeigten Phosphorthioat-modifizierte Oligonukleotide eine höhere Affinität, aber auch einen Verlust an Spezifität.<sup>[3]</sup> Hierbei muss beachtet werden, dass zunächst Leitmoleküle gesucht werden, die anschließend mit gängigen

[\*] Dr. A. Mescalchin, W. Wünsche, Prof. Dr. G. Sczakiel  
Institut für Molekulare Medizin  
Zentrum für medizinische Struktur- und Zellbiologie (ZMSZ)  
Universität zu Lübeck und Universitätsklinikum Schleswig-Holstein  
Ratzeburger Allee 160, 23538 Lübeck (Deutschland)  
Fax: (+49) 451-500-2729  
E-Mail: sczakiel@imm.uni-luebeck.de  
Homepage: <http://www.molmed.uni-luebeck.de>  
Dr. A. Mescalchin, Prof. Dr. G. Sczakiel  
Kompetenzzentrum für Drug Design/Target Monitoring (KDDTM)  
Maria-Goeppert-Straße 1, 23562 Lübeck (Deutschland)

[\*\*] Wir danken T. Restle für gereinigte Polymerasen von HIV-1 und HIV-2 sowie P. Zeltz (BadenBioTec GmbH) für Tipps bei der Anwendung der Arrays. Diese Arbeit wurde durch die Medizinische Fakultät der Universität zu Lübeck (A.M.; FKZ: A26-2007), die Possehl-Stiftung (Lübeck) und das EC-Projekt LSHG-CT-2003-503480 unterstützt.

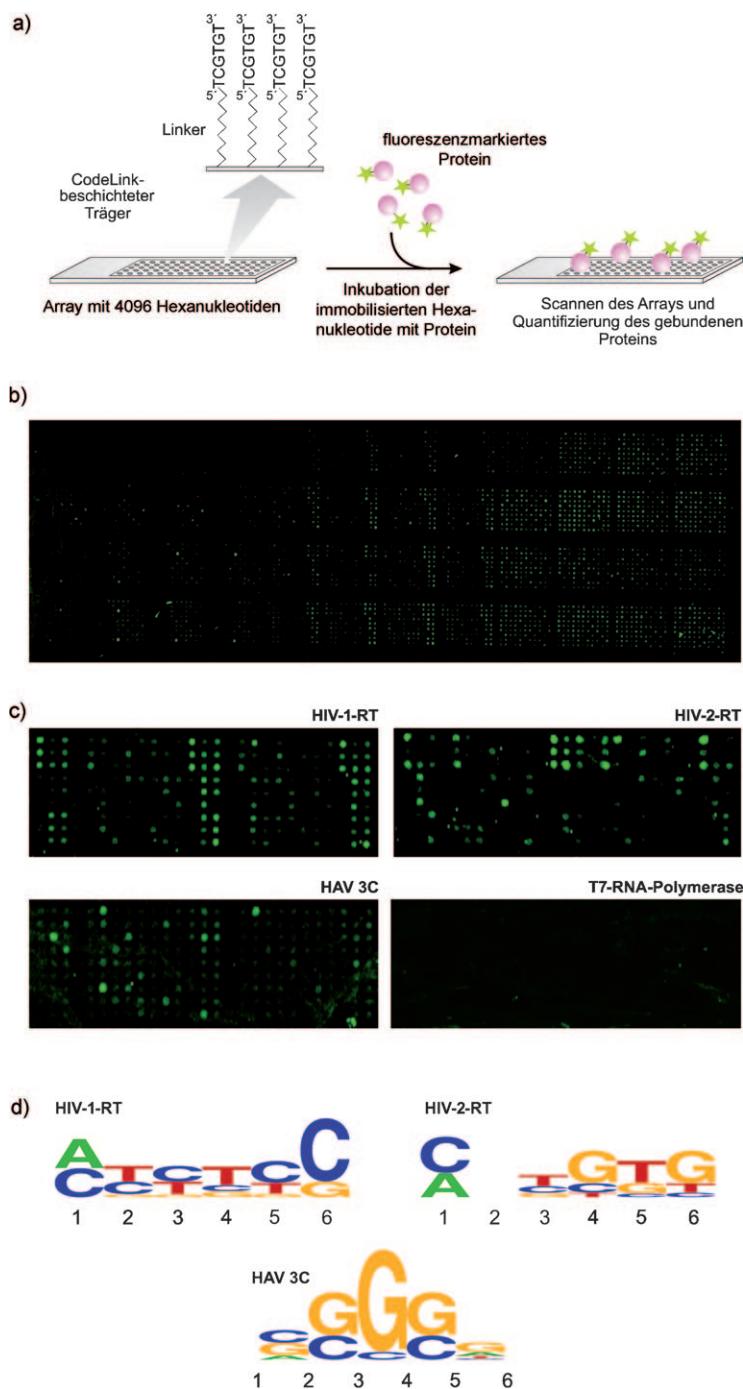
 Hintergrundinformationen zu diesem Beitrag sind im WWW unter <http://dx.doi.org/10.1002/ange.201003462> zu finden.

Verfahren der Nukleosid- und Nukleotidchemie variiert werden können.

Die Exposition von Versuchsarrays gegen HIV-1-RT in Konzentrationen von 5 nm bis 5  $\mu$ m zeigte die höchste Signalstärke und das beste Signal/Rausch-Verhältnis oberhalb von 5 nm (Abbildung S1b), was gemessen an Bindungsstudien in Lösung erstaunlich niedrig ist und einen technischen Vorteil bedeutet, wenn man den  $K_d$ -Wert von 5.7  $\mu$ m für die Hex-S3-Bindung an HIV-1-RT in Lösung bedenkt.<sup>[3]</sup> Weiterhin reflektiert das Hexanukleotid-Bindungsmuster des Arrays das Bindungsverhalten von Hexanukleotiden zu RT in Lösung,<sup>[3]</sup> wobei Linker von 21 und 38 Bindungen Länge die besten Ergebnisse lieferten, während längere Linker mit einem Verlust an Signalstärke einhergingen (Abbildung S1c). Dies wurde auch beobachtet, wenn gebundene Proteine mit Antikörpern nachgewiesen wurden, sodass insgesamt die Durchführbarkeit von Proteinbindungsstudien für HIV-1-RT mit dem vollständigen Hexanukleotid-Sequenzraum in Form eines Arrays naheliegt.

Ein Array mit dem vollständigen Hexanukleotid-Sequenzraum (4096 Spezies) wurde prozessiert wie in Abbildung 1a gezeigt. Dabei wurden zunächst die auf jedem Punkt gebundenen Spezies quantitativ durch 3'-Endmarkierung mit terminaler Desoxy-nucleotidyl-Transferase oder durch Anfärben mit SYBR-Gold quantifiziert. Außerdem wurden bezüglich HIV-1-RT-Bindung gut charakterisierte Hexanukleotide, langkettige bindende Oligonukleotide sowie Nichtbinden<sup>[3]</sup> in das Array aufgenommen (Tabelle S1). Ein typischer Bindungsversuch mit fluoreszenzmarkierter HIV-1-RT (Abbildung 1b) zeigt dynamische Signale, was auf eine ausgeprägte Diskriminierung der RT für Array-gebundene Hexanukleotide hinweist. Dieses selektive Bindungsverhalten ist weiterhin konsistent mit dem Bindungsverhalten von Positiv- und Negativkontrollen im Versuchsarray sowie in Lösung. Bemerkenswert ist, dass für einige Proteine keine bindenden Hexanukleotide erkannt wurden, zu denen die Leuchtkäfer-Luciferase zählt. Dies ist kompatibel mit einem SELEX-Versuch in Lösung, der keine Aptamere gegen diese Luciferase erkennen ließ.

Eine Überprüfung der Bindungsspezifität von Hexanukleotiden ist weiterhin mithilfe der RT von HIV-2 möglich. Die reversen Transkriptasen von HIV-1 und HIV-2 sind bezüglich ihrer Primär- und Tertiärstruktur stark homolog,<sup>[9]</sup> weshalb HIV-2-RT als eine besonders stringente Spezifitätskontrolle gilt. Das gefundene Bindungsmuster für beide Polymerasen ist ausgeprägt unterschiedlich (Abbildung 1b). Eine Analyse der am stärksten bindenden Hexanukleotide (Tabelle S2) zeigt unterschiedliche Konsensussequenzen, was vor allem für die Positionen 4 und 6 der Hexanukleotide auffällt (Abbildung 1b).



**Abbildung 1.** a) Prinzip des Screenings auf Protein-Hexanukleotid-Wechselwirkungen: Die Hexanukleotide sind über einen nichtnukleosidischen Linker kovalent an die Array-Oberfläche gebunden, wodurch nicht sequenzspezifische Wechselwirkungen des Linkers mit dem Zielprotein vermieden werden (Abbildung S1a). Zur Visualisierung von Bindungereignissen können die Proteine radioaktiv markiert oder fluoreszenzmarkiert werden. Sie können auch über Biotin oder ihre antigenen Domänen durch Immundetektion nachgewiesen werden. Hier wurde die Proteinbindung nach Blockierung des Arrays, dem Bindungsschritt, dem Waschen und Trocknen des Arrays durch Fluoreszenzmessung mit einem Typhoon-8600-Scanner gemessen. b) Beispielhaftes Bindungsmuster des Arrays bei 100 nm p66/p51<sup>K281C</sup>Alexa<sup>488</sup> (oben). c) Illustration der spezifischen Signaturen für HIV-1-RT, HIV-2-RT, HAV 3C und T7-RNA-Polymerase durch vergrößerte Teilflächen des Arrays. d) Konsensus-Bindungsmotive (5' nach 3') wurden auf der Grundlage aller 4096 Signale des Arrays für HIV-1-RT, HIV-2-RT und HAV 3C abgeleitet. Für T7-RNA-Polymerase konnte kein signifikantes Bindungsmotiv erkannt werden.

**Tabelle 1:** Wechselwirkung von Zielproteinen mit Hexanukleotid-Arrays.

Protein <sup>[a]</sup>	Array <sup>[b]</sup> (4096)	Array <sup>[b]</sup> (Versuch)	Zahl an Bindern	Konsensus- Sequenz <sup>[c]</sup>
HIV-1-RT	getestet	getestet	>15	YYYYYC
HIV-2-RT	getestet	getestet	>15	MNYSKK
HAV 3C	getestet	n.t.	14	SGGGNN
T7-RNA- Polymerase	getestet	getestet	0	–
3CL <sup>pro</sup>	getestet	n.t.	4	CTCTYN
ICAM	getestet	n.t.	0	–
Leuchtkäfer- Luciferase	getestet	getestet	0	–
BSA	n.t.	getestet	0	–

[a] HAV 3C = Hepatitis-A-Virus-Protease 3C; 3CL<sup>pro</sup> = Coronavirus-Hauptproteinase. [b] n.t. = nicht getestet. [c] Nukleotide: M (A, C), Y (C, T), N (G,C,A,T), S (G, C), K (G, T).

Weitere Zielproteine und ihre bindenden Hexanukleotide sind in Tabelle 1 zusammengefasst. Auffällig sind dabei die zum Teil unerwartet kleine Zahl bindender Hexamere im 4096er-Array sowie das völlige Ausbleiben von messbarer Bindung wie bei der T7-RNA-Polymerase, was insgesamt eine hohe Spezifität von Hexanukleotid-Protein-Wechselwirkungen anzeigt.

Das hier beschriebene, extrem zeitsparende und robuste Array-Verfahren ermöglicht die Identifizierung spezifischer Hexanukleotid-Protein-Wechselwirkungen in einem Versuch. Diese Technik kann sowohl auf Oligonukleotide mit anderen Kettenlängen als auch auf andere Zielstrukturen ausgeweitet werden. Dabei verhindert die kovalente Fixierung der Hexanukleotide mögliche Interferenzen zwischen unterschiedlichen Hexanukleotiden bei der Bindung an die Zielstruktur. Weiterhin ist dieser Screening-Ansatz empfindlich (Abbildung S1b) und schnell, und er vermeidet Klonierungs- und Vervielfältigungsschritte, die das Ergebnis verfälschen können.

Berücksichtigt man die natürliche Prädisposition von RT für die Wechselwirkung mit Polynukleotiden, so könnte man eine größere Zahl bindender Hexanukleotide erwarten. Umso auffallender ist die geringe Zahl an beobachteten Hexanukleotid-Bindern, die sich zudem stark zwischen HIV-1- und HIV-2-RT unterscheiden. Dies lässt darauf schließen, dass an den beobachteten Wechselwirkungen vor allem funktionelle Gruppen der Nukleobasen beteiligt sind. Nach einem ersten Array-Screening kann eine Optimierung des Bindungsverhaltens von Hexanukleotiden mithilfe etablierter chemischer Modifikationen der Ribose, der Nukleobasen

oder der Phosphatbrücken sowie durch Einführung zusätzlicher Substituenten erfolgen. Damit lässt sich der ursprüngliche Hexanukleotid-Sequenzraum von 4096 Spezies extrem erweitern, sodass er in puncto Komplexität sukzessive an etablierte systematische Methoden heranreichen kann, die von deutlich größeren Sequenzräumen ausgehen. Auch für ähnlich kurzkettige Oligonukleotide wie Pentanukleotide oder Heptanukleotide gelten die Vorteile einer kontrollierbaren und einfachen chemischen Synthese gegenüber langkettigen Oligonukleotiden wie klassischen Aptameren.

Die hier beschriebene Technik unterstützt die pharmazeutische Wirkstoffsuche auf der Ebene der Identifizierung von Leitsubstanzen und kann für diagnostische Zwecke sowie zu einem Aptamer-ähnlichen Hilfsmittel weiterentwickelt werden. Dieser Ansatz könnte auch in neuen Anwendungen genutzt werden, etwa zur Identifizierung regulatorischer Sequenzen, die von spezifischen Proteinen erkannt werden, zur Analyse von Änderungen der Bindungsfähigkeit von Proteinen in wechselnder Umgebung oder zur Untersuchung der Wechselwirkungen zwischen Hexanukleotiden und anderen, komplizierter aufgebauten Nukleinsäuren.

Eingegangen am 7. Juni 2010,  
veränderte Fassung am 4. August 2010  
Online veröffentlicht am 22. Dezember 2010

**Stichwörter:** HIV-1 · Mikroarrays · Oligonukleotide · Proteine · Reverse Transkriptase

- [1] A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818–822.
- [2] C. Tuerk, L. Gold, *Science* **1990**, *249*, 505–510.
- [3] A. Mescalchin, W. Wünsche, S. D. Laufer, D. Grohmann, T. Restle, G. Sczakiel, *Nucleic Acids Res.* **2006**, *34*, 5631–5637.
- [4] M. Pinskaya, E. Romanova, E. Volkov, E. Deprez, H. Leh, J. C. Brochon, J. F. Mouscadet, M. Gottikh, *Biochemistry* **2004**, *43*, 8735–8743.
- [5] J. R. Wyatt, T. A. Vickers, J. L. Roberson, R. W. Buckheit, T. Klimkait, E. DeBaets, P. W. Davis, B. Rayner, J. L. Imbach, D. J. Ecker, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 1356–1360.
- [6] A. D. van Dijk, R. Boelens, A. M. Bonvin, *FEBS J.* **2005**, *272*, 293–312.
- [7] B. Meyer, T. Peters, *Angew. Chem.* **2003**, *115*, 890–918; *Angew. Chem. Int. Ed.* **2003**, *42*, 864–890.
- [8] D. J. Schneider, J. Feigon, Z. Hostomsky, L. Gold, *Biochemistry* **1995**, *34*, 9599–9610.
- [9] J. Ren, L. E. Bird, P. P. Chamberlain, G. B. Stewart-Jones, D. I. Stuart, D. K. Stammers, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14410–14415.
- [10] Blaum et al., unveröffentlichte Ergebnisse.