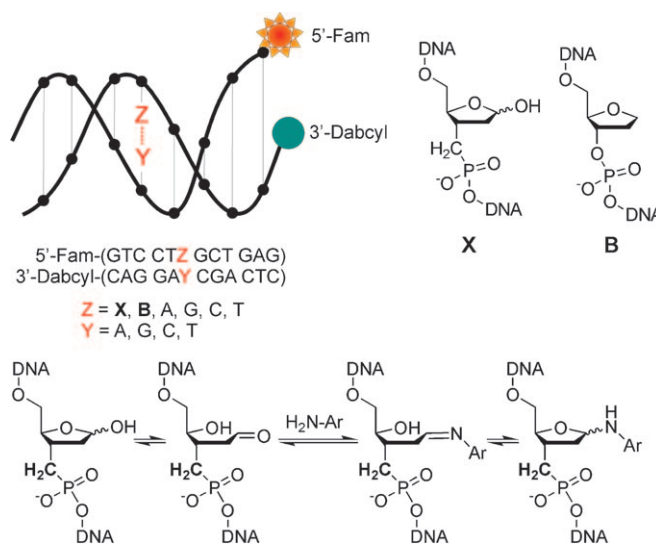


# Ein paralleles Testverfahren zur Entdeckung neuer DNA-Basenpaare\*\*

Oezlem Yaren, Markus Mosimann und Christian J. Leumann\*

In der DNA ist der genetische Code durch die Abfolge der vier Nucleobasen festgelegt, und die Weitergabe der genetischen Information erfolgt über die hochspezifischen Watson-Crick-Paarungen A-T und G-C. In den letzten Jahren entstand ein reges Interesse, dem bereits bestehenden genetischen Alphabet zusätzliche, orthogonale Basenpaare hinzuzufügen. Solche zusätzlichen Basenpaare könnten in der Biotechnologie zur Einführung nichtnatürlicher Aminosäuren in Proteine herangezogen<sup>[1–3]</sup> oder zur Entwicklung neuer Werkzeuge zur zuverlässigen Erkennung fehlerhafter oder fremder Nucleinsäuren verwendet werden.<sup>[4–8]</sup> Allerdings sind die Freiheitsgrade beim Entwurf neuer Basenpaare durch das gegebene Strukturgerüst der Doppelhelix beschränkt. Ein naheliegender Ansatz basiert deshalb auf strukturell isomorphen Basen mit alternativem Watson-Crick-ähnlichem Wasserstoffbrückenmuster.<sup>[6,9,10]</sup> Doch auch nichtisomorphe aromatische Heterocyklen, die sich gegenseitig nicht durch Wasserstoffbrücken, sondern durch andere intermolekulare Phänomene erkennen, sind mit hoher Präzision und mit kinetischen Eigenschaften ähnlich denen der natürlichen Basen replizier- und transkribierbar.<sup>[11–16]</sup> Durch diesen Ansatz ist die Bandbreite an möglichen Strukturen erheblich größer geworden, doch es fehlen ihm leider Regeln für den Basenentwurf, was das Durchprüfen einer großen Zahl potenzieller Kandidaten nötig macht. Von jeder potenziellen Base muss dazu zuerst das entsprechende Nucleosid synthetisiert und in DNA eingebaut werden, und erst danach können dessen Eigenschaften getestet werden – ein zeitraubendes Unterfangen.

Um hier Abhilfe zu schaffen, haben wir ein Verfahren entwickelt, mit dem eine Bibliothek von heterocyclischen Aminen im Parallelverfahren auf potenzielle Basenkandidaten getestet werden kann. In einem ersten Schritt suchten wir nach Molekülen, die selektiv und hoch effizient natürliche Nucleobasen erkennen. Der Test beruht auf DNA-Dodecamer-Duplexen mit einer abasischen Stelle **X** im Zentrum und einem Fluoreszenzlöscherpaaar an einem Ende (Schema 1). **X** unterscheidet sich strukturell von einer natürlichen abasi-



**Schema 1.** Funktionsweise des Testverfahrens; für Einzelheiten siehe Text. Dabcyl: 4-[4-(Dimethylamino)phenylazo]benzoesäure, Fam: 6-Carboxyfluorescein.

schen DNA-Stelle durch Ersatz des O(3') durch eine CH<sub>2</sub>-Gruppe. Dies verhindert einen basen- oder wärmeinduzierten Strangbruch am 3'-Ende.<sup>[17]</sup> Diese Duplexe wurden parallel mit einer Vielfalt an heterocyclischen Aminen inkubiert, wobei die Amine unter Bildung einer Halbaminallfunktion kovalent an **X** banden. Die daraus resultierenden *exo*-Aminonucleoside sind den natürlichen Nucleosiden strukturell ähnlich. Durch Messung der thermischen Stabilität der Duplexe mittels Fluoreszenzschmelzkurven kann die relative Affinität jedes Amins zu jeder der vier natürlichen Nucleobasen bestimmt werden (siehe die Hintergrundinformationen).

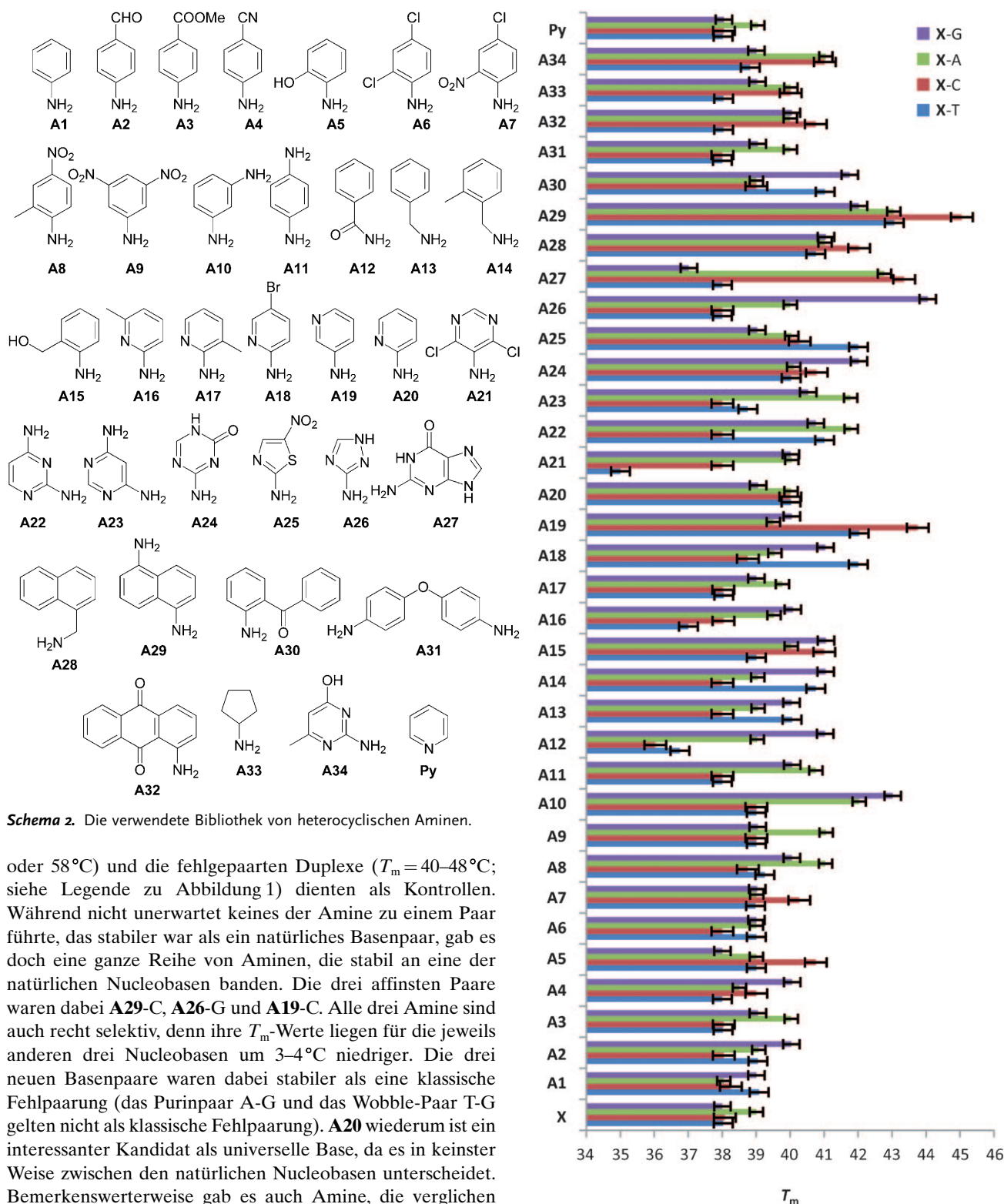
Zur Validierung des Testverfahrens suchten wir in einer Bibliothek von 34 willkürlich gewählten, kommerziell erhältlichen heterocyclischen Aminen (Schema 2) nach Kandidaten, die selektiv und mit hoher Affinität an alle vier natürlichen Nucleobasen binden. Um den Gleichgewichtszustand in Richtung der Halbaminale zu verschieben, wurden die Duplexe (0.6  $\mu$ M) in einem pH-8-Puffer zwei Tage lang bei 55 °C mit einem 1000-fachen Überschuss an Amin inkubiert. Von jedem Gemisch wurde anschließend eine Fluoreszenz-*T<sub>m</sub>*-Analyse durchgeführt. Die erhaltenen *T<sub>m</sub>*-Werte sind in Abbildung 1 wiedergegeben; die numerischen Daten befinden sich in den Hintergrundinformationen.

Aus Abbildung 1 wird ersichtlich, dass alle *T<sub>m</sub>*-Werte zwischen 35 und 45 °C liegen. Die abasische Einheit **X** allein (*T<sub>m</sub>* = 38 oder 39 °C) sowie die korrekt gepaarten (*T<sub>m</sub>* = 54

[\*] O. Yaren, Dr. M. Mosimann, Prof. C. J. Leumann  
Department für Chemie und Biochemie, Universität Bern  
Freiestrasse 3, 3012 Bern (Schweiz)  
Fax: (+41) 31-631-3422  
E-Mail: leumann@ioc.unibe.ch

[\*\*] Wir danken für großzügige finanzielle Unterstützung dieses Projektes durch den Schweizerischen Nationalfonds (Gesuchs-Nr.: 200020-115913).

Hintergrundinformationen zu diesem Beitrag (z. B. experimentelle Details) sind im WWW unter <http://dx.doi.org/10.1002/ange.201005300> zu finden.



**Schema 2.** Die verwendete Bibliothek von heterocyclischen Aminen.

oder 58°C) und die fehlgepaarten Duplexe ( $T_m = 40\text{--}48^\circ\text{C}$ ; siehe Legende zu Abbildung 1) dienen als Kontrollen. Während nicht unerwartet keines der Aminen zu einem Paar führte, das stabiler war als ein natürliches Basenpaar, gab es doch eine ganze Reihe von Aminen, die stabil an eine der natürlichen Nucleobasen banden. Die drei affinsten Paare waren dabei **A29-C**, **A26-G** und **A19-C**. Alle drei Aminen sind auch recht selektiv, denn ihre  $T_m$ -Werte liegen für die jeweils anderen drei Nucleobasen um 3–4°C niedriger. Die drei neuen Basenpaare waren dabei stabiler als eine klassische Fehlpaarung (das Purinpaar A-G und das Wobble-Paar T-G gelten nicht als klassische Fehlpaarung). **A20** wiederum ist ein interessanter Kandidat als universelle Base, da es in keinsten Weise zwischen den natürlichen Nucleobasen unterscheidet. Bemerkenswerterweise gab es auch Aminen, die verglichen mit der abasischen Stelle **X** allein die Doppelhelix destabilisieren. Ein Beispiel dafür ist das **A21-T**-Paar. Wie erwartet beeinflusst Pyridin, das keine kovalente Bindung mit **X** eingehen kann, den Schmelzpunkt nicht. Dies zeigt, dass eine kovalente Verknüpfung der heterocyclischen Aminen mit der abasischen Stelle wichtig ist.

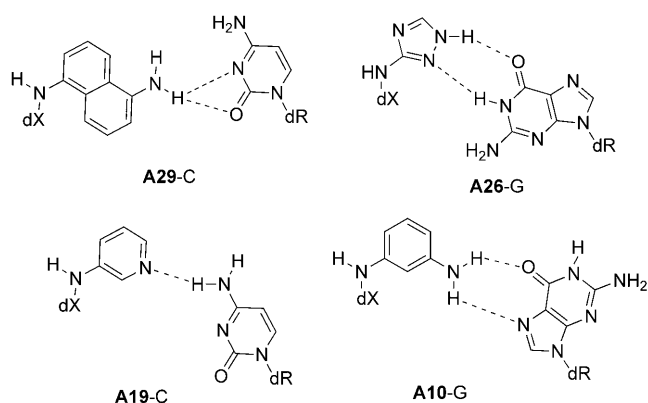
Mithilfe einer Reihe von Kontrollexperimenten wurde die korrekte Funktionsweise des Testverfahrens belegt. Um

**Abbildung 1.**  $T_m$ -Werte der Duplexe mit den heterocyclischen Aminen **A1–A34**. Es wurden Doppellexperimente durchgeführt. Die  $T_m$ -Werte der korrekt gepaarten Duplexe sind: T-A 54, C-G 58, A-T 54, G-C 58°C; die  $T_m$ -Werte der fehlgepaarten Duplexe sind: T-C und C-T 41, A-C 40, A-G 47, G-T 48°C.

auszuschließen, dass die unterschiedlichen Duplexstabilitäten die Folge von zufälligem Binden der in großem Überschuss vorliegenden Amine an die Doppelhelix waren, wurde der native Duplex mit einem natürlichen Basenpaar (Z-Y = A-T) mit einem Teil der Aminbibliothek inkubiert. Keines dieser Amine veränderte jedoch die  $T_m$ -Werte wesentlich (siehe die Hintergrundinformationen), womit Interkalation und unspezifisches Binden in den beiden Furchen der DNA als Ursache der Stabilitätsunterschiede ausgeschlossen werden konnte. In einem nächsten Schritt untersuchten wir, ob die Halbacetal-einheit von **X** an der kovalenten Bindung der Amine tatsächlich beteiligt war. Dazu wurde das Testoligosoxynucleotid d(TTTXTTT) mit Anilin (**A1**, Schema 2) unter den Bedingungen des Testverfahrens behandelt. Durch ESI-Massenspektrometrie konnte die Bildung des entsprechenden Halbaminals nachgewiesen werden (siehe die Hintergrundinformationen). In einem weiteren Experiment wollten wir die Notwendigkeit einer kovalenten Bindung zwischen den Aminen und **X** beweisen. Wir wiederholten die Experimente mit der ganzen Bibliothek von Aminen mit einem Duplex, der statt der abasischen Einheit **X** die abasische Einheit **B** gegenüber der Nucleobase T aufwies (Z-Y = **B**-T, Schema 1). Auch in diesem Fall bewegten sich alle gemessenen  $T_m$ -Werte in der engen Bandbreite von  $(38 \pm 0.5)^\circ\text{C}$  und wurden somit durch die Anwesenheit der Amine nicht beeinflusst (siehe die Hintergrundinformationen). Die beiden letzten Experimente zeigen klar die Wichtigkeit der chemischen Reaktivität am Anomerzentrum von **X** und stützen die Hypothese der Halbaminallbildung.

Es ist nicht einfach, die drei aufgefundenen hochaffinen Basenpaare mithilfe von Strukturargumenten zu erklären, da die Amine sowohl die  $\alpha$ - als auch die  $\beta$ -Konfiguration am anomeren Zentrum von **X** annehmen können. Zusätzlich weisen einige der Amine mehrere nichtsymmetrische Aminogruppen auf, was zu zusätzlicher struktureller Vielfalt führen kann. In Schema 3 haben wir mögliche Basenpaare für die vier stabilsten Anordnungen aufgezeigt. Diese sind jedoch rein hypothetisch.

Bei der Interpretation der Resultate sind mehrere Punkte zu berücksichtigen. So zeigen die Amine der verwendeten Bibliothek eine große Variabilität in der Nucleophilie der



**Schema 3.** Mögliche Strukturen der vier stabilsten im Testverfahren identifizierten Basenpaare (dR = 2'-Desoxyribose, dX = C(3')-Methylen-2'-desoxyribose).

Aminofunktion, die die Lage des Halbacetal-Halbaminall-Gleichgewichts und die Stabilität der glycosidischen Bindung stark beeinflussen dürfte. Es ist daher wahrscheinlich, dass trotz des großen Überschusses an Amin nicht alle abasischen Stellen zu 100% belegt sind. Die gemessenen  $T_m$ -Werte reflektieren deshalb nicht notwendigerweise die größtmögliche Stabilisierung, was glücklicherweise nur zu falsch negativen, nicht aber zu falsch positiven Resultaten führen kann.

In einem Versuch, das Ausmaß der Halbaminallbildung abzuschätzen, erhoben wir  $T_m$ -Werte bei verschiedenen pH-Werten (5.5–8.0) am Beispiel der Amine **A10**, **A19**, **A26** und **A29**, da wir eine einfachere Reaktion bei niedrigerem pH-Wert erwarteten. Innerhalb der durch Fluorescein gegebenen Grenzen (schwache Fluoreszenz unterhalb pH 6) konnten wir jedoch keine pH-Abhängigkeit der  $T_m$ -Werte finden (siehe die Hintergrundinformationen), was entweder auf eine fehlende pH-Empfindlichkeit der Reaktion in diesem pH-Bereich oder auf vollständige Halbaminallbildung für diese Basen schließen lässt.

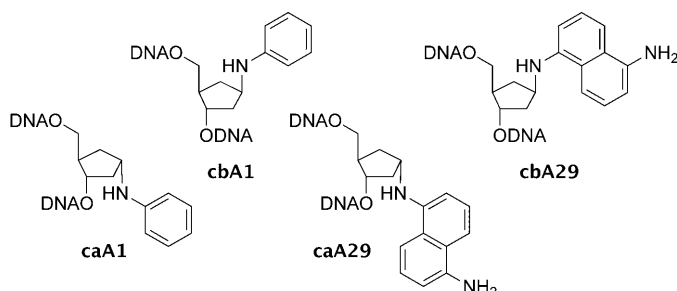
Das hier beschriebene Testverfahren fußt auf der abasischen Einheit **X**, die von ihrem natürlichen Pendant durch Ersatz von O an C(3') durch  $\text{CH}_2$  abweicht. Es ergibt sich deshalb die Frage, inwiefern diese chemische Mutation den  $T_m$ -Wert beeinflusst. In Arbeiten der Firma ISIS war gezeigt worden, dass C(3')-Methylen nucleoside DNA-Duplexe leicht stabilisieren.<sup>[18]</sup> Deshalb dürfte diese chemische Änderung höchstens einen kleinen Einfluss auf das Testverfahren haben.

Aus früheren Arbeiten über Ribozyme weiß man, dass für den katalytischen Schritt kritische Nucleoside in einen abasischen Rest **B** und die freie Base getrennt werden können. Während die abasische Stelle allein zu einem inaktiven Ribozym führt, kann die Aktivität zumindest teilweise durch Zusatz der Base oder einer anderen geeigneten Einheit wiederhergestellt werden.<sup>[19–21]</sup> Es wird angenommen, dass die Base die Strukturlücke im Ribozym füllt und damit die Aktivität wiederherstellt. Wie bereits erwähnt spielt diese Wiederherstellung der Aktivität in unserem Testverfahren keine Rolle, da keine unterschiedlichen  $T_m$ -Werte im **B** enthaltenen Duplex in Gegenwart der Amine zu beobachten waren (siehe die Hintergrundinformationen). Dies ist nicht überraschend, da die Katalyse (Kinetik) nicht unbedingt eine hohe Affinität der Base zur Lücke verlangt, während für die Stabilität (Thermodynamik) genau dies (bestenfalls als kovalente Verknüpfung) erwartet werden muss.

Obwohl die glycosidische Bindung zwischen einem Zucker und einem primären heterocyclischen Amin im Allgemeinen weniger stabil ist als diejenige eines natürlichen Nucleosids in DNA und RNA, ist eine ganze Reihe solcher *exo*-Aminonucleoside aus natürlichen oder synthetischen Quellen bekannt.<sup>[22,23]</sup> Sollte jedoch ein im Testverfahren identifiziertes potenzielles *exo*-Aminonucleosid chemisch zu wenig stabil sein, bietet sich die Alternative, die 2'-Desoxyribofuranoseeinheit durch eine carbocyclische Zuckereinheit zu ersetzen. Solche Nucleoside sind präparativ leicht zugänglich,<sup>[24]</sup> strukturell zu natürlichen Nucleosiden isomorph und dafür bekannt, dass sie mit komplementären Nucleosiden in gleicher Weise wechselwirken wie ihre natürlichen Pendant.<sup>[25,26]</sup> Zusätzlich bieten solche carbocyclischen Nucleoside eine einfache und sichere analytische Möglichkeit zu

testen, welche anomere Form für die beobachteten Affinitäten verantwortlich sind.

Um diese letzte Hypothese zu prüfen, synthetisierten wir die Fam enthaltenden Oligodesoxynucleotide mit den carbocyclischen Nucleosiden der Amine **A1** und **A29** sowohl in der  $\alpha$ - als auch in der  $\beta$ -Konfiguration in Position Z (Schema 4, Tabelle S2 in den Hintergrundinformationen) und ermittelten die  $T_m$ -Werte mit den komplementären Oligodesoxynucleotiden (Y = A, T, C, G) an der gegenüberliegenden Stelle (Tabelle 1).



**Schema 4.** Oligonucleotide (5'-Fam-(GTCCTZGCTGAG), Schema 1) mit carbocyclischen Nucleosiden der Amine **A1** und **A29** in  $\alpha$ - oder  $\beta$ -Konfiguration.

**Tabelle 1:** Fluoreszenz- $T_m$ -Werte [°C] von Duplexen (Schema 1) mit den carbocyclischen Nucleosiden der Amine **A1** und **A29** (Schema 4).

Y	Z <sup>[a]</sup>			
	caA1 <sup>[b]</sup>	cbA1	caA29	cbA29
A	–	40	44	39 <sup>[c]</sup>
C	–	39	45	39
G	–	41	43	40
T	–	40	44	39

[a] 0.6  $\mu$ M Duplex in 3.5 mM MgCl<sub>2</sub>, 50 mM KCl, 10 mM Tris, pH 8.0.

[b] Die Kombinationen mit **caA1** wurden nicht detektiert. [c]  $T_m$ -Wert aus der UV-Schmelzkurve (260 nm).

Mit **cbA1** waren die  $T_m$ -Werte leicht höher (ca +1 °C) als beim Testverfahren, und wie im Testverfahren war auch hier fast keine Basendiskriminierung zu erkennen. Mit **caA1** konnten keine kooperativen Übergänge in den Schmelzkurven gefunden werden. Im Falle des Amins **A29** ergab sich eine andere Situation. Beim  $\alpha$ -Nucleosid (**caA29**) lagen die  $T_m$ -Werte wiederum im selben Bereich wie im Testverfahren, die Paarungselektivitäten allerdings waren etwas weniger ausgeprägt, folgten aber derselben Stabilitätsreihenfolge. Dagegen führte das entsprechende  $\beta$ -Nucleosid (**cbA29**) zu  $T_m$ -Werten, die um 5–6 °C niedriger waren. Dies zeigt, dass im Fall von **A29** die  $\alpha$ -Form für die im Testverfahren gefundenen  $T_m$ -Werte verantwortlich ist. In allen Fällen konnten im Übrigen die  $T_m$ -Werte mittels UV-Schmelzkurven bestätigt werden.

Diese Reproduktion der  $T_m$ -Werte und im Großen und Ganzen auch der Paarungselektivitäten belegt zweifelsfrei die korrekte Funktionsweise unseres Testverfahrens. Zusätzlich zeichnet sich ab, dass erfolgreiche Kandidaten am basenträgenden Zentrum sowohl  $\alpha$ - als auch  $\beta$ -konfiguriert sein können.

Wir haben hier ein funktionierendes Testverfahren zur Erkennung von neuen DNA-Basenpaaren vorgestellt. Zugegebenermaßen wurde bisher nur eine sehr beschränkte Zahl an heterocyclischen Aminen getestet, die nur einen kleinen Teil der möglichen Strukturen abdecken. Das Verfahren kann jedoch weiter optimiert und seine Möglichkeiten auf mehrere Arten erweitert werden. Beispielsweise könnten die Fluoreszenzschmelzkurven in einem multiparallelen Format aufgezeichnet werden, was den Durchsatz deutlich erhöhen und die Verwendung größerer Amin-Bibliotheken zulassen würde. Des Weiteren bestünde die Möglichkeit, unter Verwendung eines Duplex, in dem sich zwei abasische Einheiten gegenüber stehen (Z = Y = X), nach ganz neuen Homobasenpaaren zu suchen.

Eingegangen am 25. August 2010,  
veränderte Fassung am 15. November 2010  
Online veröffentlicht am 21. Januar 2011

**Stichwörter:** DNA-Erkennung · Fluoreszenzsonden · Kombinatorische Chemie · Nucleoside · Parallelsynthesen

- [1] J. D. Bain, C. Switzer, A. R. Chamberlin, S. A. Benner, *Nature* **1992**, 356, 537–539.
- [2] I. Hirao, M. Kimoto, T. Mitsui, T. Fujiwara, R. Kawai, A. Sato, Y. Harada, S. Yokoyama, *Nat. Methods* **2006**, 3, 729–735.
- [3] A. A. Henry, F. E. Romesberg, *Curr. Opin. Chem. Biol.* **2003**, 7, 727–733.
- [4] M. Kimoto, R. Kawai, T. Mitsui, S. Yokoyama, I. Hirao, *Nucleic Acids Res.* **2009**, 37, e14.
- [5] D. A. Malyshev, Y. J. Seo, P. Ordoukhanian, F. E. Romesberg, *J. Am. Chem. Soc.* **2009**, 131, 14620–14621.
- [6] Z. Yang, F. Chen, S. G. Chamberlin, S. A. Benner, *Angew. Chem.* **2010**, 122, 181–184; *Angew. Chem. Int. Ed.* **2010**, 49, 177–180.
- [7] S. Hoshika, F. Chen, N. A. Leal, S. A. Benner, Z. Yang, S. G. Chamberlin, *Angew. Chem.* **2010**, 122, 5686–5689; *Angew. Chem. Int. Ed.* **2010**, 49, 5554–5557.
- [8] A. T. Krueger, H. Lu, A. H. F. Lee, E. T. Kool, *Acc. Chem. Res.* **2007**, 40, 141–150.
- [9] J. A. Piccirilli, T. Krauch, S. E. Moroney, S. A. Benner, *Nature* **1990**, 343, 33–37.
- [10] C. Switzer, S. E. Moroney, S. A. Benner, *J. Am. Chem. Soc.* **1989**, 111, 8322–8323.
- [11] M. Berger, A. K. Ogawa, D. L. McMinn, Y. Q. Wu, P. G. Schultz, F. E. Romesberg, *Angew. Chem.* **2000**, 112, 3069–3071; *Angew. Chem. Int. Ed.* **2000**, 39, 2940–2942.
- [12] S. Matsuda, J. D. Fillo, A. A. Henry, P. Rai, S. J. Wilkens, T. J. Dwyer, B. H. Geierstanger, D. E. Wemmer, P. G. Schultz, G. Spraggon, F. E. Romesberg, *J. Am. Chem. Soc.* **2007**, 129, 10466–10473.
- [13] M. Kimoto, I. Hirao, *Methods Mol. Biol.* **2010**, 634, 355–369.
- [14] A. T. Krueger, E. T. Kool, *Curr. Opin. Chem. Biol.* **2007**, 11, 588–594.
- [15] Y. J. Seo, S. Matsuda, F. E. Romesberg, G. T. Hwang, P. Ordoukhanian, *J. Am. Chem. Soc.* **2009**, 131, 5046–5047.
- [16] Y. J. Seo, G. T. Hwang, P. Ordoukhanian, F. E. Romesberg, *J. Am. Chem. Soc.* **2009**, 131, 3246–3252.
- [17] M. Mosimann, P. A. Küpfer, C. J. Leumann, *Org. Lett.* **2005**, 7, 5211–5214.
- [18] H. An, T. Wang, M. A. Maier, M. Manoharan, B. S. Ross, P. D. Cook, *J. Org. Chem.* **2001**, 66, 2789–2801.
- [19] Y. I. Kuzmin, C. P. Da Costa, M. J. Fedor, *J. Mol. Biol.* **2004**, 340, 233–251.

- [20] A. Peracchi, L. Beigelman, N. Usman, D. Herschlag, *Proc. Natl. Acad. Sci. USA* **1996**, 93, 11522–11527.
  - [21] A. T. Perrotta, T. S. Wadkins, M. D. Been, *RNA* **2006**, 12, 1282–1291.
  - [22] R. J. Moss, C. R. Petrie, R. B. Meyer, L. D. Nord, R. C. Willis, R. A. Smith, S. B. Larson, G. D. Kini, R. K. Robins, *J. Med. Chem.* **1988**, 31, 786–790.
  - [23] C. V. Varaprasad, Q. Habib, D. Y. Li, J. F. Huang, J. W. Abt, F. Rong, Z. Hong, H. Y. An, *Tetrahedron* **2003**, 59, 2297–2307.
  - [24] O. R. Ludek, C. Meier, *Synthesis* **2003**, 2101–2109.
  - [25] J. Sagi, A. Szemző, J. Szécsi, L. Ötvös, *Nucleic Acids Res.* **1990**, 18, 2133–2140.
  - [26] a) K. H. Altmann, M. O. B'Vierre, A. De Mesmaeker, H. E. Moser, *Bioorg. Med. Chem.* **1995**, 5, 431–436; b) Y. Xu, K. Kino, H. Sugiyama, *J. Biomol. Struct. Dyn.* **2002**, 20, 437–446.
-