



Synthese eines DNA-Promotorsegments mit allen vier epigenetischen Nukleosiden: 5-Methyl-, 5-Hydroxymethyl-, 5-Formyl- und 5-Carboxy-2'-Desoxycytidin**

Arne S. Schröder, Jessica Steinbacher, Barbara Steigenberger, Felix A. Gnerlich, Stefan Schiesser, Toni Pfaffeneder und Thomas Carell*

Abstract: Ein 5-Formyl-2'-desoxycytidin(fdC)-Phosphoramidit-Baustein wurde entwickelt, der die Synthese fdC-haltiger DNA in ausgezeichneter Reinheit und Ausbeute ermöglicht. In Kombination mit Phosphoramiditen für 5-Methyl-dC, 5-Hydroxymethyl-dC und 5-Carboxy-dC gelang die Herstellung eines Segments des OCT4-Promotors, das alle vier epigenetischen Basen enthält. Aufgrund des großen Interesses an diesen neuen epigenetischen Basen sollte die Fähigkeit, alle vier von ihnen in DNA einzubringen, von besonderer Bedeutung sein.

Neben den vier kanonischen Nukleobasen enthält das Genom einige DNA-Modifikationen, die von epigenetischer Bedeutung sind. Das Nukleosid 2'-Desoxycytidin wird häufig an Position C5 zu 5-Methyl-2'-desoxycytidin (**1**, mdC) methyliert, das an der epigenetischen Gen-Stillelegung beteiligt ist.^[1] In den letzten Jahren wurden die Oxidationsprodukte von mdC entdeckt.^[2–5] Mittlerweile ist bekannt, dass „Teneleven-translocation“-Proteine (TET1–3) mdC zu 5-Hydroxymethyl- (**2**, hmdC), 5-Formyl- (**3**, fdC) und 5-Carboxy-2'-desoxycytidin (**4**, cadC) oxidieren.^[6,7] Diese „neuen“ Modifikationen (Abbildung 1A) treten insbesondere im CpG-Kontext in DNA-Promotorsequenzen auf, wo sie wahrscheinlich an der Regulation der Genexpression involviert sind.^[8] Vor kurzem wurde über die Synthese und Analyse eines tRNA-Segments der menschlichen mitochondrialen tRNA^{Met} berichtet, das 5-Formylcytidin in der Anticodon-Schleife trägt.^[9]

Zur Untersuchung der Funktion dieser neuen Nukleoside ist es besonders wichtig, dass Methoden zur Synthese von hmdC-, fdC- und cadC-haltiger DNA entwickelt werden. Kürzlich berichteten wir über den enzymatischen Einbau der Nukleoside in DNA über deren Triphosphate.^[10] Ein se-

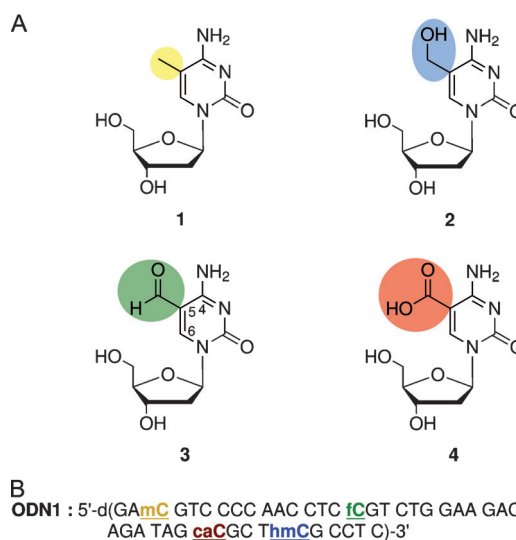


Abbildung 1. A) Die in letzter Zeit gefundenen epigenetisch relevanten Nukleoside 5-Methyl- (**1**, mdC), 5-Hydroxymethyl- (**2**, hmdC), 5-Formyl- (**3**, fdC) und 5-Carboxy-2'-desoxycytidin (**4**, cadC). B) Sequenz eines OCT-4-Promotorsegments, in dem Cytidine im CpG-Kontext mit epigenetischen Nukleosiden **1–4** substituiert wurden.

quenzspezifischer Einbau der Nukleoside, speziell in CpG-Inseln, wie es beispielsweise für proteomische Untersuchungen von Nöten wäre, ist aber mit diesen Methoden nicht möglich.

Die Achillesferse der Synthese von DNA-Strängen mit den vier neuen epigenetisch wichtigen Nukleosiden ist der Einbau von fdC. Probleme bereitet hier die große Reaktivität der Formylgruppe.^[11–13] Der bislang erhältliche fdC-Baustein hat daher eine 1,2-Diol-Einheit quasi als Vorstufe zur Formylgruppe, aus welcher durch Behandlung mit Periodat die Formylgruppe hergestellt wird;^[14] dies dürfte jedoch nicht kompatibel mit z.B. hmdC sein.^[15] Darüber hinaus werden durch die Periodatbehandlung Nebenreaktionen eingeleitet, die es bisher verhindern, DNA-Stränge herzustellen, die mehrere fdCs enthalten. Folglich sind DNA-Stränge, die alle epigenetisch wichtigen Nukleoside mdC, hmdC, fdC und cadC an definierten Positionen enthalten, derzeit nicht zugänglich.^[16]

Hier beschreiben wir die Entwicklung eines neuen fdC-Phosphoramidit-Bausteins, der mehrere Male – auch direkt hintereinander – in DNA-Stränge eingebaut werden kann, wobei Standard-Phosphoramiditbedingungen zum Einsatz

[*] A. S. Schröder, J. Steinbacher, B. Steigenberger, F. A. Gnerlich, S. Schiesser, T. Pfaffeneder, Prof. Dr. T. Carell
Center for Integrated Protein Science
Department für Chemie, Ludwig-Maximilians-Universität
Butenandtstraße 5–13, 81377 München (Deutschland)
E-Mail: Thomas.Carell@lmu.de
Homepage: <http://www.carellgroup.de>

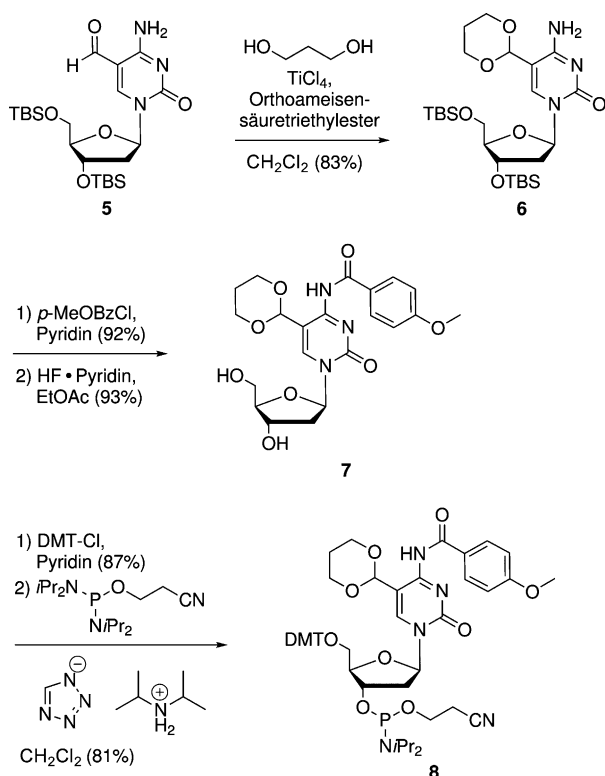
[**] Wir danken der Deutschen Forschungsgemeinschaft (SFB749 und SFB1032) für finanzielle Unterstützung. Zusätzliche Unterstützung erhielten wir durch den Fonds der Chemischen Industrie (Promotionsstipendien für A.S.S., S.S. und T.P.). Wir danken auch den Exzellenzclustern (CiPS^M und NIM).

Hintergrundinformationen zu diesem Beitrag sind im WWW unter <http://dx.doi.org/10.1002/ange.201308469> zu finden.

kommen. Hohe Kupplungsausbeuten und milde Entschützungsbedingungen ermöglichten die Herstellung einer OCT-4-Promotorsequenz mit allen vier epigenetisch relevanten Nukleosiden (Abbildung 1 B, **ODN1**).

Die Haupthürde beim Einbau von fdC ist der elektronenarme Heterocyclus von **3**, der die Spaltung der glykosidischen Bindung vor allem unter leicht sauren Bedingungen begünstigt, wie sie im Rahmen der Oligonukleotidsynthese benötigt werden. Darüber hinaus ist die Formylgruppe oxidationsempfindlich, und natürlich wird die Formylverbindung leicht durch Nukleophile vor allem an Position C6 angegriffen (Abbildung 1 A).^[16] Um die Reaktivität der Formylgruppe zu maskieren, untersuchten wir unterschiedliche Schutzgruppen (siehe die Hintergrundinformationen). Schlussendlich zeigte sich, dass die besten Ergebnisse erzielt wurden, wenn der Aldehyd mit Propan-1,3-diol als 1,3-Dioxan geschützt wird. Dieses Acetal ist spaltbar unter milden Bedingungen (80 % wässr. Essigsäure, 20 °C), es ist aber so stabil, dass es während der Festphasensynthese nicht gespalten wird. Entscheidend ist aber die Entschützungs-temperatur am Ende der Synthese. Bei 15 °C ist die Acetal-spaltung zu langsam, bei ≥ 25 °C hingegen wurde bereits die Spaltung der fdC-glykosidischen Bindung beobachtet. Die Entschützung muss daher strikt bei etwa 20 °C erfolgen.

Die C4-Aminogruppe wurde mit 4-Methoxybenzoylchlorid geschützt, was mit der Acetalgruppe kompatibel ist. Eine herkömmliche Benzoylschutzgruppe stellte sich als labil heraus. Diese fällt während der Festphasensynthese ab, was zur Synthese verzweigter Oligonukleotidnebenprodukte führt.



Schema 1. Synthese des acetalgeschützten fdC-Phosphoramidits **8**.

Die Synthese des fdC-Bausteins gelang wie in Schema 1 dargestellt ausgehend von dem TBS-geschützten fdC **5**, das aus 5-Iod-dC synthetisiert wurde.^[17] Überraschenderweise konnten die üblichen Bedingungen zur Einführung der Acetalschutzgruppe anschließend nicht angewendet werden, da sie nicht das Produkt **6** lieferten. Stattdessen musste die Formylgruppe mit TiCl_4 aktiviert werden, wobei in unseren Händen nur diese Lewis-Säure akzeptable Ausbeuten lieferte. Nachfolgende Schützung der C4-Aminogruppe mit 4-Methoxybenzoylchlorid und Abspaltung der Silylschutzgruppen lieferte das fdC-Derivat **7**, das mithilfe von Standardmethoden in das fdC-Phosphoramidit **8** umgewandelt wurde.^[18]

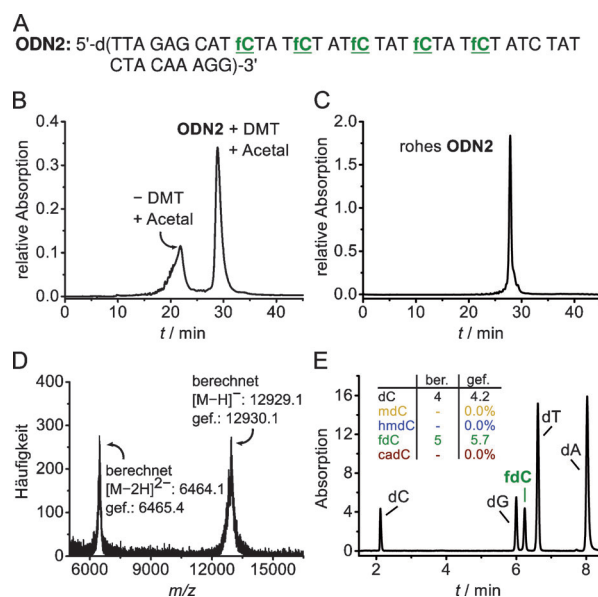


Abbildung 2. A) Sequenz des synthetisierten **ODN2** mit fünffachem Einbau von fdC mithilfe von **8**. B) Umkehrphasen-HPL-Chromatogramm des rohen **ODN2** direkt nach basischer Entschützung (NH_4OH , 23 h, 25 °C, 0–80 % Puffer B in 45 min). C) Umkehrphasen-HPL-Chromatogramm des rohen **ODN2** direkt nach saurer Entschützung der DMT- und Acetalgruppe (80 % wässr. Essigsäure, 6 h, 20 °C, 0–40 % Puffer B in 45 min). D) MALDI-TOF-Spektrum von rohem **ODN2** nach vollständiger Entschützung. E) UHPL-Chromatogramm und Quantifizierungsdaten von verdautem **ODN2**.

Um die Qualität des neuen fdC-Bausteines **8** zu testen, wurde ein Oligonukleotid **ODN2** mit fünf fdCs (Abbildung 2 B) hergestellt. Für die Festphasensynthese wurden Standard-Phosphoramiditbedingungen eingesetzt.^[19,20] Lediglich die Kupplungszeiten wurden für das fdC von 30 auf 60 Sekunden erhöht. Zur Entschützung der Nukleotide und zur Abspaltung vom Harz wurde letzteres mit gesättigter wässriger NH_4OH -Lösung (24 h, 25 °C) behandelt. Das HPL-Chromatogramm in Abbildung 2 B zeigt das Resultat für das Rohprodukt der Synthese des Dimethoxytrityl(DMT)-geschützten Oligonukleotids. In einem zweiten Entschützungsschritt wurde das Oligonukleotid anschließend mit Essigsäure für 6 h bei 20 °C behandelt, was zur Spaltung der Acetal- und DMT-Schutzgruppe führte. Das HPL-Chromatogramm für

das rohe Oligonukleotid **ODN2** mit fünf fdCs im DNA-Strang ist in Abbildung 2C gezeigt. Der einzelne Peak zeigt die Güte der Synthese.

Um zu beweisen, dass sich tatsächlich fünf fdC-Nukleoside im DNA-Strang befinden, wurde ein MALDI-TOF-Spektrum (Abbildung 2D) gemessen, welches die korrekte Masse lieferte. Ein weiterer Beweis wurde durch einen Totalverdau der Roh-DNA erhalten. Zunächst wurde die DNA mit der Nuklease S1 und Schlangengift-Phosphodiesterase I in die 5'-Monophosphate zerlegt. Diese wurden zu den Nukleosiden mithilfe der Antarktischen Phosphatase weiterverdaut. Die erhaltene Nukleosidmischung wurde mit UHPLC-MS/MS (QQQ) analysiert (Abbildung 2E). Das erhaltene UV-Chromatogramm ist sauber und zeigt nur die Signale für die erwarteten Nukleoside. Die fdC-Nukleoside wurden anhand der Retentionszeit und der charakteristischen MS/MS-Fragmentierung (m/z 256.1 \rightarrow 140.1) identifiziert. Wir haben anschließend die Mengen der Basen dC, fdC und dT (Einschub in Abbildung 2E) unter Verwendung synthetischer isotopologer Nukleoside als interne Standards bestimmt.^[16] Wir erhielten ein dC/fdC-Verhältnis von 4:5 in Übereinstimmung mit der Sequenz. Die Tatsache, dass dieser Wert mit dem Roh-DNA-Strang erhalten wurde, zeigt die Qualität der Synthese mit dem neuen fdC-Phosphoramidit. Besonders wichtig ist, dass wir kein α -fdC-Nukleosid detektieren konnten, was zeigt, dass wir die β - zu α -Anomerisierung mit dem neuen Baustein erfolgreich unterdrücken konnten. Wir untersuchten auch die Anwesenheit von hmdC und cadC als Reduktions- und Oxidationsprodukte von fdC ausgelöst durch die obligatorische Iodoxidation, während der DNA-Synthese in jedem Schritt. Wie erwartet wurden diese Verbindungen nicht entdeckt (Einschub in Abbildung 2E). Zusammenfassend kann damit festgehalten werden, dass das fdC-Phosphoramidit **8** die Synthese fdC-enthaltender Oligonukleotide ermöglicht. Ausgehend von diesem Baustein begannen wir anschließend mit der Synthese eines Oligonukleotids, das alle vier epigenetischen Basen enthalten sollte.

Als Zielsequenz wählten wir das Oligonukleotid **ODN1** mit der Sequenz des OCT-4-Promoters (*mus musculus*, Chromosom 17, 35505895-35505943, Abbildung 3B). Alle Cytidine im CpG-Kontext wurden durch die epigenetisch relevanten Nukleoside **1–4** ersetzt. Die besten Ergebnisse wurden erzielt, wenn wir den hmC-Phosphoramiditbaustein einsetzten, der mit einer cyclischen Carbamat-Schutzgruppe ausgestattet ist (für Details zur Oligonukleotidsynthese siehe die Hintergrundinformationen).^[17] Der Carbamat-Baustein und der cadC-Baustein mit der Estergruppe zwangen uns dazu, für die Entschützung NaOH (0.4 M in Methanol/Wasser 4:1) statt NH_4OH im ersten Entschützungsschritt einzusetzen, um die Bildung von Aminomethylgruppen und Amidgruppen zu vermeiden (Abbildung 3A).^[12,17] Nach der Behandlung von synthetisiertem **ODN1** mit NaOH wurde das Oligonukleotid durch Zugabe von Ethanol gefällt. Anschließend wurde der zweite Entschützungsschritt mit Essigsäure durchgeführt. Abbildung 3C zeigt erneut das HPL-Chromatogramm des Rohprodukts nach dem zweiten Entschützungsschritt. Deutlich sichtbar ist, dass das Verfahren ein Hauptprodukt liefert obwohl es sich bei **ODN1** um ein recht langes 49mer mit vier modifizierten Nukleotiden handelt (für

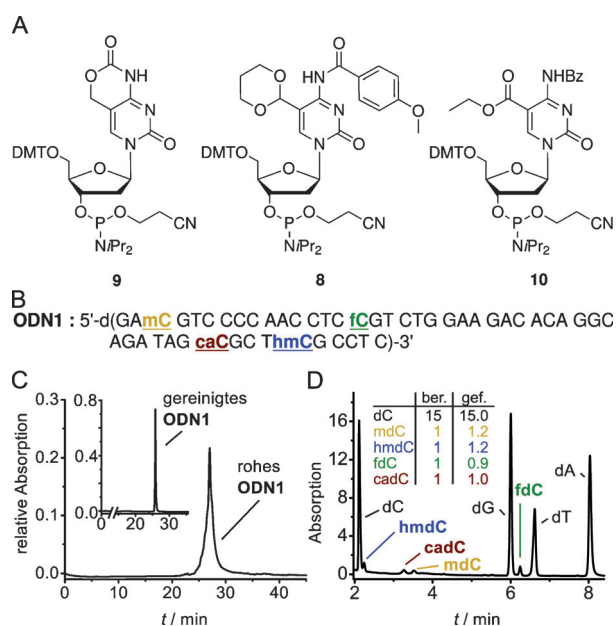


Abbildung 3. A) Die verwendeten Phosphoramidite. B) Sequenz eines OCT-4-Promotorsegments, in dem Cytidine im CpG-Kontext gegen die epigenetischen Nukleoside **1–4** ausgetauscht wurden. C) Umkehrphasen-HPL-Chromatogramm von rohem und gereinigtem (Einschub) **ODN1** nach vollständiger Entschützung (0–40% Puffer B in 45 min). D) UHPL-Chromatogramm und Quantifizierungsdaten von verdautem **ODN1**. Bz = Benzoyl.

detaillierte Daten zu **ODN1** siehe die Hintergrundinformationen). HPLC-Aufreinigung des Oligonukleotids nach dem ersten Entschützungsschritt gefolgt von der zweiten Entschützung lieferte reines **ODN1** in 31% Ausbeute (siehe den Einschub in Abbildung 3C und die Hintergrundinformationen). Nachfolgend verdauten wir das Oligonukleotid und analysierten das Hydrolysat mit UHPLC-MS/MS. Hierbei zeigte sich die Anwesenheit aller vier nichtkanonischen dC-Nukleoside. Isotopenbasierte Quantifizierung bestätigte die richtige Zusammensetzung. Wir erhielten relativ zu dC die Werte 1.2:1.2:0.9:1.0:15.0 in völliger Übereinstimmung mit der Sequenz (Abbildung 3B/D).

Zusammenfassend halten wir fest, dass der neue fdC-Baustein **8** die Synthese von fdC-enthaltender DNA in exzellenten Ausbeuten und Reinheit ermöglicht, was Studien zur biologischen Bedeutung von fdC beschleunigen dürfte. In Kombination mit den Phosphoramiditbausteinen für mdC, hmdC und cadC war es nun auch möglich, Oligonukleotide herzustellen, die alle vier epigenetisch relevanten Cytidinnukleoside **1–4** enthalten. Das wird den Weg für eine detailliertere Analyse ebnen, wie epigenetische Basen biologische Prozesse und die Stammzellenentwicklung beeinflussen.^[8]

Eingegangen am 28. September 2013

Online veröffentlicht am 26. November 2013

Stichwörter: 5-Formylcytosin · DNA-Methylierung · Epigenetik · Festphasensynthesen · Oligonukleotide

- [1] J. A. Law, S. E. Jacobsen, *Nat. Rev. Genet.* **2010**, *11*, 204–220.
- [2] S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929–930.
- [3] M. Tahilian, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, *Science* **2009**, *324*, 930–935.
- [4] T. Pfaffeneder, B. Hackner, M. Truss, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem.* **2011**, *123*, 7146–7150; *Angew. Chem. Int. Ed.* **2011**, *50*, 7008–7012.
- [5] Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, G. L. Xu, *Science* **2011**, *333*, 1303–1307.
- [6] C. Loenarz, C. J. Schofield, *Chem. Biol.* **2009**, *16*, 580–583.
- [7] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, *Science* **2011**, *333*, 1300–1303.
- [8] C. X. Song, C. He, *Trends Biochem. Sci.* **2013**, *38*, 480–484.
- [9] E. M. Gustilo, F. A. P. Vendeix, R. Kaiser, M. O. Delaney, W. D. Graham, V. A. Moye, W. A. Cantara, P. F. Agris, A. Deiters, *Nucleic Acids Res.* **2008**, *36*, 6548–6557.
- [10] B. Steigenberger, S. Schiesser, B. Hackner, C. Brandmayr, S. K. Laube, J. Steinbacher, T. Pfaffeneder, T. Carell, *Org. Lett.* **2013**, *15*, 366–369.
- [11] S. Tardy-Planechaud, J. Fujimoto, S. S. Lin, L. C. Sowers, *Nucleic Acids Res.* **1997**, *25*, 553–559.
- [12] M. Münzel, U. Lischke, D. Stathis, T. Pfaffeneder, F. A. Gnerlich, C. A. Deiml, S. C. Koch, K. Karaghiosoff, T. Carell, *Chem. Eur. J.* **2011**, *17*, 13782–13788.
- [13] Q. Dai, C. He, *Org. Lett.* **2011**, *13*, 3446–3449.
- [14] N. Karino, Y. Ueno, A. Matsuda, *Nucleic Acids Res.* **2001**, *29*, 2456–2463.
- [15] T. M. A. Shaikh, L. Emmanuvel, A. Sudalai, *J. Org. Chem.* **2006**, *71*, 5043–5046.
- [16] S. Schiesser, T. Pfaffeneder, K. Sadeghian, B. Hackner, B. Steigenberger, A. S. Schröder, J. Steinbacher, G. Kashiwazaki, G. Höfner, K. T. Wanner, C. Ochsenfeld, T. Carell, *J. Am. Chem. Soc.* **2013**, *135*, 14593–14599.
- [17] M. Münzel, D. Globisch, C. Trindler, T. Carell, *Org. Lett.* **2010**, *12*, 5671–5673.
- [18] W. Bannwarth, A. Trzeciak, *Helv. Chim. Acta* **1987**, *70*, 175–186.
- [19] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859–1862.
- [20] N. D. Sinha, J. Biernat, J. McManus, H. Köster, *Nucleic Acids Res.* **1984**, *12*, 4539–4557.