ARTICLE

OBC
www.rsc.org/obc

CG base pair recognition by substituted phenylimidazole nucleosides

Wei Wang,[a] Maria G. M. Purwanto[b] and Klaus Weisz*[b]

[a] *Institut für Chemie, Freie Universität Berlin, Takustrasse 3, D-14195 Berlin, Germany*

[b] *Institut für Chemie und Biochemie, Ernst-Moritz-Arndt-Universität Greifswald, Soldmannstrasse 16, D-17487 Greifswald, Germany. E-mail: weisz@uni-greifswald.de*

Received 11th December 2003, Accepted 23rd February 2004
First published as an Advance Article on the web 18th March 2004

Four nonnatural imidazole nucleosides with different substituents were synthesized and studied for their binding to a CG Watson–Crick base pair by NMR spectroscopic techniques in an aprotic solvent. Concentration and temperature dependent measurements allowed the determination of association constants, association enthalpies and entropies. Strong binding was observed with analogues carrying an ureidophenyl substituent and corresponding enthalpies of association are compatible with the anticipated formation of three hydrogen bonds to the CG base pair. In contrast, only weak binding was observed for analogues with an aminophenyl or benzamidophenyl substituent. 2D NOE measurements at low temperatures confirm the proposed binding mode for the high-affinity ligands but indicate binding interactions for the weakly bound analogues different from the expected geometry.

Introduction

Intermolecular triplex formation between single-stranded and double-stranded nucleic acids has been shown to offer a powerful chemical approach for targeting specific sites within the genome.[1] Possible applications of such a methodology include the modulation of gene expression by blocking transcription and thus repress protein production at the DNA level (antigene strategy).[2] In addition, by covalently attaching specific chemical groups to the triplex-forming single strand, manipulations at any given site of the DNA including strand cleavage seems to be conceivable.[3] Up to now, a major drawback of triplex-based methodologies has arisen from the restricted recognition code. Thus, only the purine bases adenine (A) and guanine (G), positioned on the same strand of the duplex, are effectively recognized by thymine (T) and cytosine (C) third strand bases to form T·AT and C^+·GC base triplets in the case of the pyrimidine or parallel triplex motif.[4] Pyrimidine interruptions within an all-purine tract of the double helix generally result in triple helices of low stability due to the lack of their effective recognition. Therefore, the design and synthesis of nonnatural nucleoside analogues to extend the recognition to all four possible base pairs is of major importance for any strategy that is based on triplex formation.

Recently, much attention has been paid to the development of artificial nucleobases as ligands for any given base pair.[5] However, due to the complex interplay of the various interactions within a triple helical complex, predictions as for their affinity and specificity were hardly successful. Here we report on differently substituted phenylimidazole nucleosides as receptors for a CG base pair. A benzamido-substituted phenylimidazole has recently been incorporated into triplex-forming oligonucleotides by Dervan *et al.*,[6] but has been shown to bind *via* intercalation, contrary to what was expected.[7] In this paper we report on a detailed thermodynamic and structural characterization of CG binding by various novel phenylimidazole nucleosides in apolar solvents in order to shed more light on the structure–affinity relationship that is based on the formation of specific hydrogen bonds.

Results and discussion

According to molecular modeling studies, imidazole nucleosides may serve as potential CG recognizing ligands. As shown in Fig. 1, such analogues are able to form a specific hydrogen bond between their endocyclic N3 nitrogen and the Watson–Crick cytosine amino group within a complex closely isomorphous to the canonical T·AT and C^+·GC base triads. Correspondingly, the aminophenyl-substituted imidazole nucleoside **4a** should facilitate the formation of two hydrogen bonds when binding a CG base pair, *i.e.*, to $H4_b$ of cytosine and the O6 carbonyl oxygen of guanosine. In addition to **4a**, the modified analogues **4b**, **4c** and **4d** were synthesized to study the effect of the benzoyl-, ureido- and n-butylureido-substituents on the hydrogen bond mediated CG complex formation. Binding of **4b** should benefit from a more acidic amide proton compared to the amino proton in **4a**, thus constituting a more effective hydrogen bond donor. Nucleosides **4c** and **4d** are expected to form a third hydrogen bond to N7 of the guanine base, thus increasing CG binding affinity.

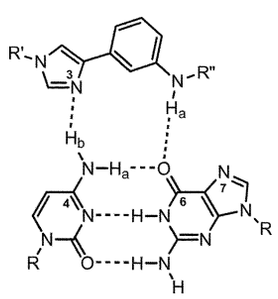

Fig. 1 Proposed binding mode of substituted 4-phenylimidazole nucleosides to a CG Watson–Crick base pair.

The synthesis of 2-unsubstituted imidazoles starting from an α-hydroxy or α-halogenated ketone **1** and formamide is based on the Bredereck imidazole synthesis[8] recently also employed by Griffin *et al.*[6a] A Vorbrüggen-type reaction of the resulting 4-(3-nitrophenyl)-imidazole **2** with 1-*O*-acetyl 2,3,5-tri-O-benzoyl-β-D-ribofuranose afforded the corresponding β-nucleoside **3**. The subsequent reduction of the nitro group with Raney nickel and hydrazine as hydrogen source gave the amino-substituted **4a** (Fig. 2). Benzoylation of **4a** under standard conditions at room temperature afforded **4b**. Analogues **4c** and **4d** were prepared from **4a** using phenyl carbamate and n-butylisocyanate as reagents, respectively (Fig. 3).

DOI: 10.1039/b316077g

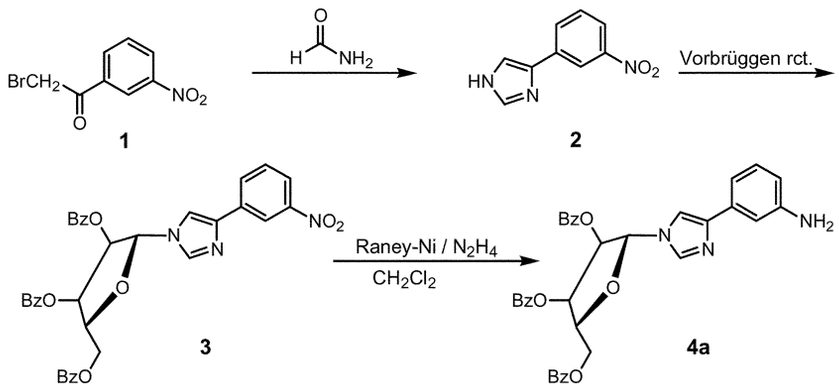

Fig. 2 Scheme for the synthesis of **4a**.

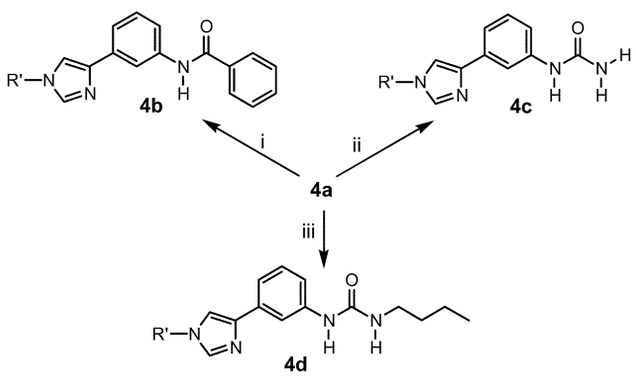

Fig. 3 Synthesis of **4b**, **4c** and **4d** starting from **4a**: (i) BzCl, pyridine, THF, room temperature, 80%; (ii) phenyl carbamate, DMF, 60 °C, 62%; (iii) C_4H_9NCO, THF, 77%; R′ = 2′,3′,5′-tri-*O*-benzoyl-β-D-ribofuranosyl.

NMR titration experiments

The use of *O*-silylated 2′-deoxycytidine and 2′-deoxyguanosine as well as *O*-benzoylated nucleoside analogues **4a**–**d** not only blocks the sugar hydroxyl groups but also greatly enhances solubility in the methylene chloride solvent. Initial experiments on the homoassociation of the four analogues were performed by following the NH_a proton chemical shift of nucleosides **4a**–**d** as a function of nucleoside concentration ranging from 5 to 100 mM. Homoassociation constants were determined by a non-linear least-squares fit of the experimental data and amount to 0.8 M^{-1} for **4a** (290 K), 0.3 M^{-1} for **4b** (290 K), 18 M^{-1} for **4c** (292 K) and 6 M^{-1} for **4d** (300 K). Disregarding some uncertainties in the determination of small association constants within the concentration range employed, homoassociation only occurs to a small extent for all nucleosides and thus justifies its initial neglect in the following analysis of CG complexation. Also, as has been shown before, the strong binding between C and G in apolar solvents ($K_{ass} \sim 10^5$ M^{-1}) allows a 1 : 1 mixture of the two nucleosides to be treated as a single CG species.[9]

Binding of the nucleoside analogues to the CG base pair was followed by changes in the ^{1}H NMR chemical shift of the NH_a proton of **4a**–**d**. Thus, titration of the analogue (2 mM) with a 1 : 1 mixture of C and G (0 to 20 mM) resulted in a downfield shift of the H_a signal with increasing CG concentration as shown in Fig. 4 for the analogue **4d** at four different temperatures. The equilibrium constant for complex formation K_{ass} between each of the nucleoside analogues and the CG base pair was determined by fitting the chemical shift data as a function of concentration to a nonlinear 1 : 1 binding isotherm. The results of measurements at different temperatures are summarized in Table 1.

Clearly, with association constants of 15 M^{-1} (242 K) and 7 M^{-1} (268 K) CG binding of **4a** and **4b** is only weak and does not seem to allow for a specific high-affinity base pair recognition through hydrogen bonding at ambient temperatures. This may be attributed to the weak amino proton donor in **4a** and steric hindrance by the bulky benzamido group in **4b** (*vide infra*).

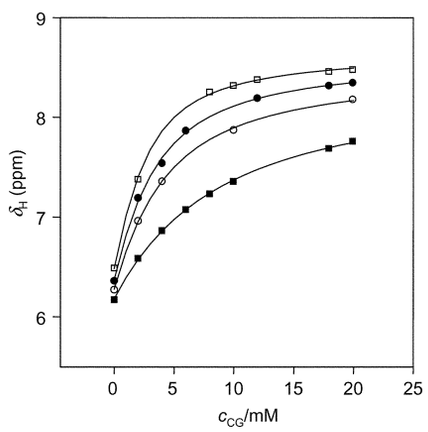

Fig. 4 NH_a proton chemical shift of **4d** as a function of CG concentration at 292 K (solid squares), 268 K (open circles), 255 K (solid circles) and 245 K (open squares). Lines represent the least-squares fit.

Table 1 Summary of K_{ass}, ΔH and ΔS for the association of nucleoside analogues **4a**–**d** with a CG base pair

	T/K	K_{ass}/M^{-1}	ΔH/kJ mol^{-1}	ΔS/J K^{-1} mol^{-1}
4a	242	15 ± 2	−8.7 ± 2	−13.5 ± 4
	219	23 ± 3		
	212	28 ± 3		
4b	268	7 ± 2	−11.5 ± 2	−26.5 ± 6
	245	12 ± 2		
	222	24 ± 2		
	210	28 ± 2		
4c	291	173 ± 25	−24.5 ± 2	−40.8 ± 6
	279	289 ± 18		
	268	459 ± 30		
	256	688 ± 55		
4d	292	131 ± 6	−19.4 ± 2	−25.4 ± 6
	268	316 ± 20		
	255	418 ± 40		
	245	627 ± 41		

In contrast, association constants of 173 M^{-1} (291 K) for **4c** and 131 M^{-1} (292 K) for **4d** compare favorably with an association constant of 70 M^{-1} reported for the canonical AU base pair at ambient temperatures and indicate the formation of three hydrogen bonds to the CG base pair. In line with previous binding studies using other analogues,[9] the ureido group seems to be a major contributor to the strong binding of the free nucleosides in organic solvents, however, *N*-alkylation as in **4d** slightly reduces the binding constant over the temperature range studied.

We have also performed reverse titrations by adding the analogue (0–20 mM) to a 2 mM solution of the CG base pair at 286 K. Based on the proposed binding mode, the non-Watson–Crick bound cytosine amino proton H_b is in all cases expected to undergo a downfield shift upon titration due to the formation of a hydrogen bond to the imidazole N-3. However, the observation of cytosine amino signals is severely hampered by the overlap of resonances, especially with the nucleoside analogue in excess. We have therefore employed specifically 4-^{15}N labeled cytidine, enabling the use of a one-dimensional ^{1}H–^{15}N HMQC experiment as a filter that effectively eliminates all other proton resonances. The downfield shifts observed in such experiments for the C H_b amino resonance when adding **4a**–**d** to the CG mixture were again fitted to a 1 : 1 binding isotherm (Fig. 5). Association constants K_{ass} for the complexation of nucleoside analogues with the CG base pair were found to agree within the estimated margin of error with the association constants obtained from the former titrations (data not shown). Only for **4c**, having the largest self-association constant of all analogues, a more noticeable deviation of about 40% in the heteroassociation constant at 286 K is observed for the two independent titrations and may be attributed to a non-negligible degree of self-association.

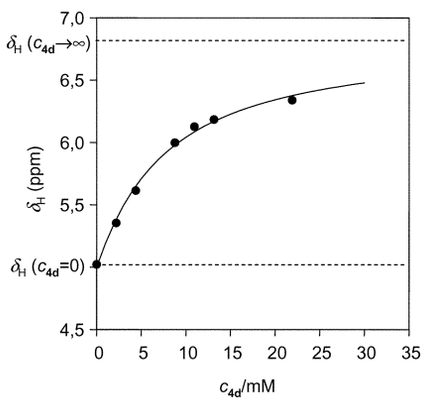

Fig. 5 Representative plot of the cytosine H_b amino proton chemical shift in a CG base pair as a function of the **4d** nucleoside concentration at 286 K. The solid line represents the least-squares fit with limiting chemical shifts indicated by the dotted lines.

The limiting chemical shift at $c_{4a-d} \longrightarrow \infty$ for the cytosine amino proton H_b as also determined from the curve fitting in these reverse titrations (see Fig. 5) corresponds to its shift in a trimolecular complex. Because the chemical environment of this proton is expected to be similar in all complexes, it can be used as a direct measure of the local strength and extent of hydrogen bonding in the various associates. Upon titrating **4d** and **4c** to the CG base pair, limiting chemical shifts of 6.8 ppm and 6.5 ppm for C H_b in **4d**·CG and **4c**·CG were determined. In contrast, smaller downfield shifts upon complexation with limiting chemical shifts of 6.1 ppm and 5.8 ppm for the amino proton signal in **4b**·CG and **4a**·CG may point to the coexistence of additional binding modes with altered interaction sites for the latter two complexes. Unlike the non-Watson–Crick bound C H_b amino proton, the Watson–Crick bound H_a proton signal experienced only a minor shift upon adding **4a**–**d**, indicating that the CG base pair is not disrupted upon binding of the analogue.

Enthalpy and entropy

In order to analyze the complexation data in terms of enthalpic and entropic contributions, the temperature dependent association constants were subjected to a van't Hoff analysis. A plot of $\ln K_{ass}$ as a function of T^{-1} is shown in Fig. 6 for the various complexes and corresponding enthalpies and entropies of association that have been determined by fitting the experimental data are summarized in Table 1. It has to be noted that

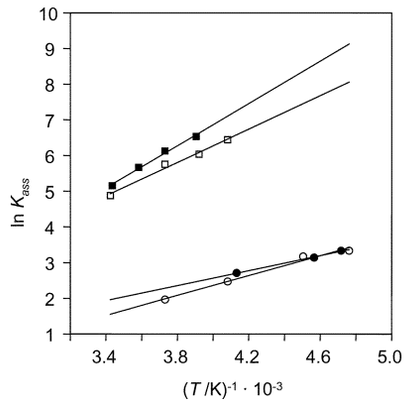

Fig. 6 Van't Hoff plot of the association constants K_{ass} for **4a**·CG (solid circles), **4b**·CG (open circles), **4c**·CG (solid squares) and **4d**·CG (open squares) base triple formation in CD_2Cl_2. Lines represent the least-squares fit.

the enthalpy values not only reflect the formation of hydrogen bonds but also include the enthalpy of solvation, *i.e.*, the entire enthalpy difference between solvated complex and solvated monomers. However, assuming only small differences between the enthalpy of solvation for the dimer and monomers, $-\Delta H$ should reasonably well reflect the hydrogen bond energy. Consequently, values found between -8 and -12 kJ mol^{-1} for **4a**·CG and **4b**·CG base triple formation point to two very weak hydrogen bond contacts or, alternatively, to only one hydrogen bond interaction between the analogue and the base pair. Much larger enthalpic contributions of -24 kJ mol^{-1} and -19 kJ mol^{-1} are determined for **4c**·CG and **4d**·CG, respectively. Assuming the formation of three hydrogen bonds in these cases, the heat of formation averages to about 7 kJ mol^{-1} per hydrogen bond. Although smaller than expected for a single hydrogen bond of medium strength, these values seem reasonable given the non-cooperative nature of the contacts and a weakening of individual interactions upon introducing additional interaction sites without perfect complementarity.

As expected, entropies decrease on complex formation, ranging from -13 J K^{-1} mol^{-1} for **4a**·CG to -41 J K^{-1} mol^{-1} for **4c**·CG. Because a loss in entropy can be partly associated with a restricted conformational flexibility upon CG complexation, the rigid analogue **4a** with a non-acylated amino substituent is expected to lose less degrees of freedom on binding compared to the other analogues. Also, **4d** with its n-butyl substituent might strongly favor a *Z*-conformation thus restricting amide C–N bond rotations even in the uncomplexed monomer and reducing the entropic penalty for the association when compared to **4c**. It must be cautioned, however, that the entropies determined from the temperature dependence of association constants are subject to a relatively large degree of uncertainty. Also, the interpretation of experimental entropies of association at a molecular level is not straightforward and should necessarily also include effects from solvation and desolvation which may be significant.

2D NOE measurements

For a more detailed structural characterization of the complexes, 2D NOE experiments on 1 : 1 mixtures of the non-natural nucleosides and the CG base pair were performed at 213 K in methylene chloride. A portion of the spectrum for the **4d**·CG base triple is shown in Fig. 7. Due to the low temperatures employed for the measurements, equilibria are largely shifted toward the intact base triad. Again, typical contacts observed between guanine and cytosine imino and amino protons for Watson–Crick base pairing indicate that the base pair is not disrupted in the presence of the analogue. Most notably, however, the spectrum displays several intermolecular NOE connectivities between base protons of **4d** and the CG

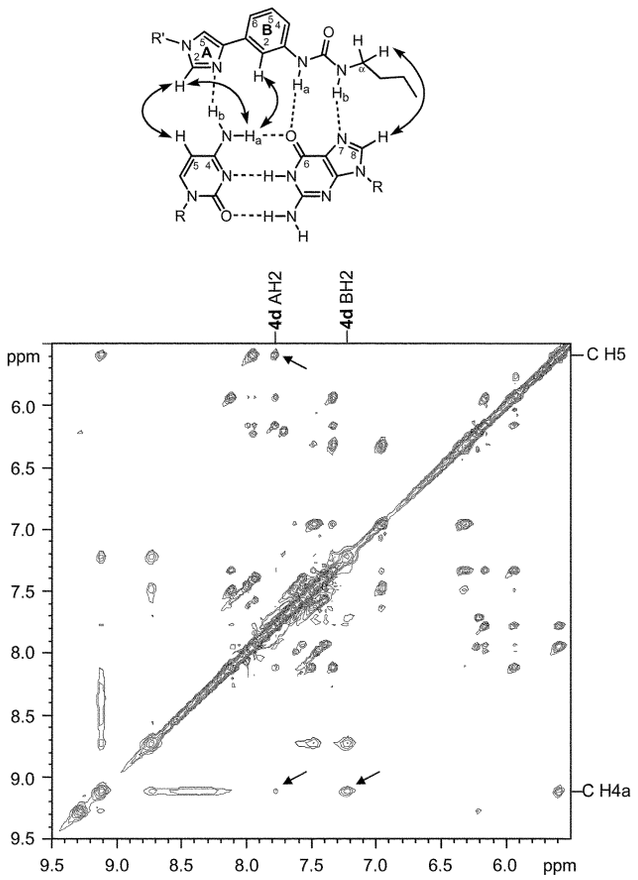

Fig. 7 Portion of a 2D NOE spectrum of a 1 : 1 mixture of **4d** and a CG base pair in CD_2Cl_2 at 213 K (mixing time = 200 ms). Intermolecular crosspeaks between **4d** and the base pair are marked by arrows and indicated in the structure at the top; R = 3′,5′-di-*O*-(triisopropylsilyl)-β-D-2′-deoxyribofuranosyl, R′ = 2′,3′,5′-tri-*O*-benzoyl-β-D-ribofuranosyl.

base pair. Thus, as seen in Fig. 7, intermolecular NOE crosspeaks between cytosine H5 and the **4d** AH2 proton as well as between the Watson–Crick bound cytosine amino proton $H4_a$ and **4d** AH2 and BH2 protons are compatible with the proposed binding mode of the nucleoside analogue to the CG base pair. In addition, a crosspeak between H8 of guanine and the α-methylene protons of the n-butyl residue of **4d** corroborates the formation of a third hydrogen bond to N7 of guanosine. Likewise, corresponding NOE contacts C NH_a–**4c** AH2, C H5–**4c** AH2 and C NH_a–**4c** BH2 are also in line with the corresponding binding mode of **4c**. Expected intermolecular crosspeaks for the non-Watson–Crick bound C $H4_b$ amino proton were not observed possibly due to significant broadening of the signal at this temperature.

In contrast to **4c** and **4d**, we only found a single weak intermolecular NOE contact to the CG base pair for both **4a** and **4b**, namely, a contact between BH2 of **4a** or **4b** and H5 of cytosine (not shown). Without any additional intermolecular crosspeak, the NOE data do not allow an unambiguous characterization of the **4a**·CG and **4b**·CG geometries again reflecting the weak binding of the two analogues to the base pair. However, the presence of a C H5-**4a/b** BH2 contact suggests the formation of complexes with binding modes different from what is expected. In fact, molecular models of a tentative **4b**·CG base triplet indicate that it does not allow coplanarity of the bases due to a steric clash between the benzamido group of **4b** and the guanine purine ring.

Conclusion

The detailed NMR spectroscopic analysis of CG binding with differently substituted imidazole nucleosides not only gives information on the thermodynamics of hydrogen bond mediated association, but also on the binding mode, prerequisite for establishing useful structure–affinity relationships. Our data indicate that, in contrast to the high-affinity binding of analogues carrying an ureidophenyl group, analogues with an unsubstituted or benzoylated aminophenyl substituent only exhibit poor affinity and specificity toward a CG base pair. This can be attributed to the amino group being a weak hydrogen bond donor and the benzamido group giving rise to steric clashes with the guanine purine ring. Within a triple helical complex the benzamido-modified analogue was previously shown to bind through preferential intercalation rather than through the formation of specific hydrogen bonds. Clearly, more studies are needed to gain a better understanding of the contribution of additional non-hydrogen bond interactions in the base pair recognition of nucleic acids. However, although eliminating additional interactions in macromolecular systems, the association of free nucleosides may provide useful information on the binding of ligands within a biological context.

Experimental

General methods

NMR experiments were performed on a Bruker AMX500 spectrometer. ^{1}H chemical shifts in methylene chloride were referenced relative to $CHDCl_2$ (δ_H = 5.32 ppm). Temperatures were calibrated with a standard solution of 4% CH_3OH in CD_3OD and adjusted by a Eurotherm Variable Temperature Unit to an accuracy of ±1.0 °C. Concentration dependent chemical shifts were fitted with an appropriate equation by employing the Marquardt–Levenberg algorithm.

Materials

Reagents of the highest quality available were purchased from Sigma-Aldrich, Deisenhofen, Germany. [4-^{15}N]-2′-deoxycytidine was prepared from unlabeled 2′-deoxyuridine as described previously.[10] Free cytidine and guanosine were *O*-silylated using triisopropylsilyl chloride and subsequently purified by HPLC. All reactions were controlled by TLC on silica gel plates (Merck silica gel 60 F_{254}). If necessary, solvents were dried by standard procedures prior to use.

1-(2′,3′,5′-Tri-*O*-benzoyl-β-D-ribofuranosyl)-4-(3-nitrophenyl)imidazole 3. A mixture of 4-(3-nitrophenyl)imidazole **2**[6a] (4.7 g, 24.84 mmol), 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-β-D-ribofuranose (12.75 g, 24.78 mmol) and potassium nonaflate (21.13 g, 62.38 mmol) were co-distilled to dryness three times each with 30 mL absolute acetonitrile (freshly distilled) and finally dissolved in 250 mL absolute acetonitrile. After addition of hexamethyldisilazane (HMDS, 1.71 mL, 8.10 mmol) and trimethylchlorosilane (TCS, 6.4 mL, 49.56 mmol) under an inert atmosphere the mixture was stirred for 24 hours under reflux. Within the first 4 hours TCS (4 mL total) and HMDS (1 mL total) were added in regular time intervals. After addition of 150 mL acetonitrile and 200 mL methylene chloride the mixture was extracted with a saturated $NaHCO_3$ solution. The organic phase was washed with a saturated NaCl solution, dried over Na_2SO_4 and the solvent was removed under vacuum. Extraction of the solid in hexane/AcOEt 1 : 1 and removal of the solvent afforded a crude product, which was purified by column chromatography (SiO_2, CH_2Cl_2/MeOH 100 : 0 to 95 : 5); yield: 1.6 g (10%); ^{1}H NMR (250 MHz, 293 K, DMSO): δ (ppm) = 4.71–4.90 (3H, m, H4′, H5′, H5″), 5.94 (2H, m, H2′, H3′), 6.45 (1H, d, H1′), 7.41–8.17 (19H, m, ArH), 8.26 (1H, m, H4 Phenyl), 8.52 (1H, s, H2 Phenyl).

1-(2′,3′,5′-Tri-*O*-benzoyl-β-D-ribofuranosyl)-4-(3-aminophenyl)imidazole 4a. To a solution of 1-(2′,3′,5′-tri-*O*-benzoyl-β-D-ribofuranosyl)-4-(3-nitrophenyl)imidazole **3** (1.60 g,

2.5 mmol) in anhydrous methylene chloride, Raney nickel (3.6 g) and hydrazine (11 mL, 0.35 mmol) were added. After the suspension was stirred for 24 hours, the catalyst was filtered off and the solvent was removed. The product was purified by column chromatography (SiO_2, CH_2Cl_2/MeOH 97.5 : 2.5); yield: 0.56 g (37%); ^{1}H NMR (250 MHz, 293 K, $CDCl_3$): δ (ppm) = 3.61 (2H, s, NH_2), 4.55–4.96 (3H, m, H4′, H5′, H5″), 5.92 (2H, m, H2′, H3′), 6.14 (1H, d, H1′), 6.51–8.21 (21H, m, ArH). (Found: C, 69.4; H, 5.1; N, 6.6. Calc. for $C_{35}H_{29}N_3O_7$: C, 69.65; H, 4.8; N, 7.0%).

Synthesis of 1-(2′,3′,5′-tri-*O*-benzoyl-β-D-ribofuranosyl)-4-(3-benzamidophenyl)imidazole 4b. 320 mg **4a** was co-distilled to dryness three times each with 15 mL absolute THF (freshly distilled) and then dissolved in 60 mL absolute THF. 0.24 mL pyridine and 0.2 mL (0.172 mmol) freshly distilled benzoyl chloride were added under argon and the reaction mixture was stirred at room temperature for 3 days. After completion of the reaction, the solvent was removed under vacuum and the product was purified by column chromatography (SiO_2; CH_2Cl_2/MeOH 98 : 2) and HPLC (SiO_2; CH_2Cl_2/MeOH 98.5 : 1.5); yield: 300 mg (80%); ^{1}H NMR (250 MHz, 293 K, $CDCl_3$): δ (ppm) = 4.56–4.81 (3 H, m, H4′, H5′, H5″), 5.89 (2 H, m, H2′, H3′), 6.1 (1 H, d, H1′), 7.15–8.35 (26 H, m, ArH); HRMS (EI, 80 eV 150 °C): *m*/*z* = 707.22583; calculated: 707.226765 $[C_{42}H_{33}N_3O_8]^+$.

Synthesis of 1-(2′,3′,5′-tri-*O*-benzoyl-β-D-ribofuranosyl)-4-(3-ureidophenyl)imidazole 4c. 0.1 g **4a** was dissolved in anhydrous DMF (0.5 mL) at 60 °C. To this solution was added 0.24 g (1.76 mmol) phenyl carbamate and the reaction mixture was stirred at 60 °C until TLC indicated that the educt had been consumed (72 h). The mixture was evaporated to dryness and then purified by column chromatography (SiO_2; CH_2Cl_2/MeOH 95 : 5); yield: 150 mg (62%); ^{1}H NMR (250 MHz, 293 K, $CDCl_3$): δ (ppm) = 4.56–4.82 (3 H, m, H4′, H5′, H5″), 5.65 (2 H, s, NH_2), 5.90 (2 H, m, H2′, H3′), 6.05 (1 H, d, H1′), 7.02–8.08 (22 H, m, ArH and NH); HRMS (EI, 80 eV 150 °C): *m*/*z* = 629.17657; calculated: 629.179815 $[C_{36}H_{30}N_4O_8–NH_3]^+$.

Synthesis of 1-(2′,3′,5′-tri-*O*-benzoyl-β-D-ribofuranosyl)-4-(3-n-butylureidophenyl)imidazole 4d. 300 mg **4a** was co-distilled to dryness three times each with 15 mL absolute THF and then dissolved in 40 mL absolute THF. After addition of 0.15 mL (1.38 mmol) n-butylisocyanate under argon, the mixture was refluxed for 1.5 h and stirred for an additional 24 h at room temperature. After completion of the reaction, the solvent was removed under vacuum and the crude product was purified using column chromatography and HPLC (SiO_2; CH_2Cl_2/MeOH 97 : 3); yield: 270 mg, (77%); ^{1}H NMR (250 MHz, 293 K, $CDCl_3$): δ (ppm) = 0.87 (3 H, t, CH_3), 1.22–1.50 (4 H, m, CH_2CH_2), 3.22 (2 H, q, N–CH_2), 4.54–4.79 (3 H, m, H4′, H5′, H5″), 5.58 (1 H, t, NH), 5.84 (2 H, m, H2′, H3′), 6.1 (1 H, d, H1′), 7.08–8.10 (22 H, m, ArH and NH); HRMS (EI, 80 eV 150 °C): *m*/*z* = 702.26745; calculated: 702.268965 $[C_{40}H_{38}N_4O_8]^+$.

Acknowledgements

We thank the Fonds der Chemischen Industrie for financial support.

References

1 D. Praseuth, A. L. Guieysse and C. Hélène, *Biochim. Biophys. Acta*, 1999, **1489**, 181–206.

2 (*a*) L. J. Maher III, B. Wold and P. B. Dervan, *Science*, 1989, **245**, 725–730; (*b*) J. C. Hanvey, M. Shimizu and R. D. Wells, *Nucleic Acids Res.*, 1990, **18**, 157–161.

3 S. A. Strobel, L. A. Doucette-Stamm, L. Riba, D. E. Housman and P. B. Dervan, *Science*, 1991, **254**, 1639–1642.

4 N. T. Thuong and C. Hélène, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 666–690.

5 (*a*) C.-Y. Huang and P. S. Miller, *J. Am. Chem. Soc.*, 1993, **115**, 10456–10457; (*b*) S. Sasaki, S. Nakashima, F. Nagatsugi, Y. Tanaka, M. Hisatome and M. Maeda, *Tetrahedron Lett.*, 1995, **36**, 9521–9524; (*c*) T. E. Lehmann, W. A. Greenberg, D. A. Liberles, C. K. Wada and P. B. Dervan, *Helv. Chim. Acta*, 1997, **80**, 2002–2022; (*d*) F. Lecubin, R. Benhida, J.-L. Fourrey and J.-S. Sun, *Tetrahedron Lett.*, 1999, **40**, 8085–8088; (*e*) D. Guianvarc'h, R. Benhida, J.-L. Fourrey, R. Maurisse and J.-S. Sun, *Chem. Commun.*, 2001, 1814–1815; (*f*) for reviews see S. O. Doronina and J.-P. Behr, *Chem. Soc. Rev.*, 1997, 63–71; (*g*) D. M. Gowers and K. R. Fox, *Nucleic Acids Res.*, 1999, **27**, 1569–1577; (*h*) M. G. M. Purwanto and K. Weisz, *Curr. Org. Chem.*, 2003, **7**, 427–446.

6 (*a*) L. C. Griffin, L. L. Kiessling, P. A. Beal, P. Gillespie and P. B. Dervan, *J. Am. Chem. Soc.*, 1992, **114**, 7976–7982; (*b*) L. L. Kiessling, L. C. Griffin and P. B. Dervan, *Biochemistry*, 1992, **31**, 2829–2834.

7 K. M. Koshlap, P. Gillespie, P. B. Dervan and J. Feigon, *J. Am. Chem. Soc.*, 1993, **115**, 7908–7909.

8 (*a*) H. Bredereck and G. Theilig, *Chem. Ber.*, 1953, **86**, 88–96; (*b*) G. Theilig, *Chem. Ber.*, 1953, **86**, 96–109; (*c*) M. R. Grimmet, *Adv. Heterocycl. Chem.*, 1970, **12**, 103–183.

9 (*a*) S. C. Zimmerman and P. Schmitt, *J. Am. Chem. Soc.*, 1995, **117**, 10769–10770; (*b*) E. Mertz, S. Mattei and S. C. Zimmerman, *Org. Lett.*, 2000, **2**, 2931–2934; (*c*) D. Lengeler and K. Weisz, *Tetrahedron Lett.*, 2001, **42**, 1479–1481.

10 G. Kupferschmitt, J. Schmidt, T. Schmidt, B. Fera, F. Buck and H. Rüterjans, *Nucleic Acids Res.*, 1987, **15**, 6225–6241.