Dynamic covalent chemistry on self-templating PNA oligomers: formation of a bimolecular PNA quadruplex†

Yamuna Krishnan-Ghosh,[ab] Andrew M. Whitney[a] and Shankar Balasubramanian*[a]

Received (in Cambridge, UK) 10th March 2005, Accepted 21st April 2005
First published as an Advance Article on the web 11th May 2005
DOI: 10.1039/b503578c

The tetramolecular PNA quadruplex motif has been probed using a dynamic covalent chemistry (DCC) approach to create and characterize a bimolecular PNA quadruplex.

Covalent synthesis, facilitated by the supramolecular interaction between reactants and a template, has been used to construct complex assemblies such as catenanes,[1] rotaxanes[2] and molecular capsules.[3] A specialized case of template assisted synthesis, where the template becomes an integral part of the structure, is termed 'self-templating'.[4] Self-templating has been used in conjunction with DCC to probe secondary structure in peptides.[4,5] However, the potential of DCC to analogously report on secondary structure in nucleic acids has been largely unexplored. This paper demonstrates proof of concept using fragments of peptide nucleic acids (PNA), where nucleobase recognition results in the formation of secondary structure, *i.e.*, a PNA quadruplex.[6] Because DCC covalently traps the templating secondary structure, it is possible to use DCC amplification as a tool to address the stability of the templating assembly. We have thus used DCC as a probe for PNA quadruplex formation where covalent trapping of the templating tetramolecular PNA quadruplex results in DCC amplification of a novel bimolecular PNA quadruplex.

Certain guanine-rich DNA sequences can self-assemble *via* Hoogsteen bonding to yield four-stranded structures called G-quadruplexes.[7] Our objectives were to explore whether quadruplex assembly could drive DCC amplification. Quadruplex formation has been demonstrated for iso-G[8] and G repeats on DNA,[9] RNA[10] and PNA[6,11] backbones. We have chosen to study PNA backbones owing to the synthetic ease with which chemical functionality can be incorporated. The design of our system comprises two PNA strands $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$, comprising three tandem guanine residues and four tandem thymine residues respectively (Fig. 1). A C-terminal lysine was incorporated to ensure solubility over a wide range of concentrations. A thymine at the C terminus of $\mathbf{G_{SH}}$ served to prevent higher order structure formation.[12] For the reversible covalent coupling chemistry we chose the thiol–disulfide system because it is water compatible, fast and can be switched off by a pH change.[13] Thiols were incorporated as an N-terminal Cys residue *via* a Gly–Gly spacer to allow some conformational flexibility for disulfide formation.[5]

A mixture of $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ can covalently combine to form two possible homodimers ($\mathbf{G_{SS}G}$ or $\mathbf{T_{SS}T}$) and a heterodimer ($\mathbf{G_{SS}T}$). Product distributions from all experiments were analysed from UV-HPLC traces of aliquots of the reaction mixtures and quantified from a comparison of peak areas with those of purified standards. We first allowed $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ to combine under irreversible conditions using chemical oxidation, in order to explore whether non-covalent interactions could influence product formation. Sodium perborate has been shown to rapidly oxidize thiols, promoting conditions of kinetic control.[14] When a solution of equimolar $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ (each 500 μM) in buffer,[15] was chemically oxidized using 20 mM sodium perborate, dimers $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$ were obtained in a 1.1:1.9:1.0 molar ratio (Fig. 2A). This near-statistical distribution reflects that, under these conditions, disulfide formation is uninfluenced by non-covalent interactions.

Air oxidation is slower than perborate oxidation and so, at high (mM) thiol concentrations thiol–disulfide exchange occurs at a much faster rate than the air oxidation process. Thus, air oxidation of thiols at millimolar concentration would be expected to yield a product distribution reflective of a thermodynamic equilibrium.[16] Indeed, air oxidation of equimolar solutions of 500 μM $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ conducted in buffer showed that disulfide exchange did influence the product distribution. Product distributions were analyzed at various time points. At t = 14 h 50% of $\mathbf{G_{SH}}$ has

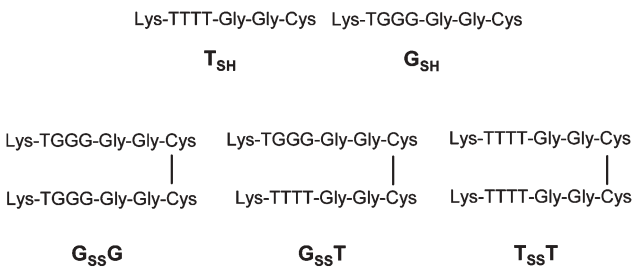

Fig. 1 Sequences of peptide nucleic acids used in this study, written from their C to N termini, and their corresponding disulfides.

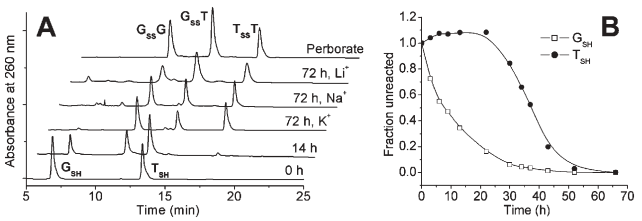

Fig. 2 **A**) HPLC profiles at various time points of equimolar 500 μM $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ in buffer, undergoing air oxidation and also chemical oxidation with $NaBO_3$ after 5 min. **B**) Rates of disappearance of $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ during the air oxidation of 500 μM $\mathbf{G_{SH}}$ and 500 μM $\mathbf{T_{SH}}$ in buffer, pH 7.4.

† Electronic supplementary information (ESI) available: synthesis, control DCC experiments, rates of oxidation, mass spectra and CD spectra. See http://www.rsc.org/suppdata/cc/b5/b503578c/

*sb10031@cam.ac.uk

reacted to form $\mathbf{G_{SS}G}$ while $\mathbf{T_{SH}}$ is still to be oxidized. At $t = 72$ h all of $\mathbf{T_{SH}}$ has either oxidized or exchanged with $\mathbf{G_{SS}G}$ to give a final product distribution of 2:1:2 $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$ (Fig. 2A). A plot of the rates of disappearance of $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ during air oxidation (Fig. 2B) reveals that $\mathbf{G_{SH}}$ is almost entirely reacted to form $\mathbf{G_{SS}G}$ before $\mathbf{T_{SH}}$ starts to be consumed. This indicates that $\mathbf{G_{SS}G}$ is either stable by itself or is part of a particularly stabilized structure.

In order to test the hypothesis that $\mathbf{G_{SS}G}$ is stabilized, its structure was studied by nanoelectrospray ionisation mass spectrometry, (Q-TOF-1 mass spectrometer Micromass, Manchester, UK) (Table 1).[6,17] An aqueous solution of 250 μM $\mathbf{G_{SS}G}$ was heated to 90 °C, annealed and incubated at 5 °C for 8 h. Nano-ESI-MS showed a peak centered at *m*/*z* 1202.8 corresponding to $[M_2 + 5H]^{5+}$. The separation between sodiated and potassiated species at *m*/*z* 1207.9, 1215.2, 1219.7 (see ESI†) confirmed that these peaks corresponded to a quintuply charged species of MW 6008.9 ± 2 Da, indicating dimer formation by $\mathbf{G_{SS}G}$. Solution phase H/D exchange experiments combined with ESI-MS were used to ascertain whether dimer formation occurred in solution *via* a specific interaction as opposed to being non-specifically formed *in vacuo*. If complexation occurs *via* specific H-bonding interactions, a comparison of the number of exchangeable protons in uncomplexed $\mathbf{G_{SS}G}$ with that of complexed $\mathbf{G_{SS}G}$ would not only indicate whether the complex was formed in solution but also provide insight into its structure.

An aliquot of 250 μM $\mathbf{G_{SS}G}$ in water was lyophilized and resuspended in an equal volume of D_2O to deuterate exchangeable protons. Nano-ESI-MS showed a major peak centered at *m*/*z* 1215.9 corresponding to $[M_2 + 5H]^{5+}$ indicating a complex MW of 6074.5 ± 2.3 Da (see ESI†). The MW per $\mathbf{G_{SS}G}$ monomer in the complex increased to 3037 ± 1.2 Da after H/D exchange indicating only partial deuteration (32 out of 44 exchangeable protons) of the complex (Table 1). A complementary experiment was conducted where 250 μM $\mathbf{G_{SS}G}$ was complexed in D_2O instead of H_2O. The MW of this complex (Table 1) confirmed that all the 44 exchangeable protons in $\mathbf{G_{SS}G}$ had undergone H/D exchange, yielding the fully deuterated complex. An aliquot of fully deuterated complex was lyophilized and resuspended in an equal volume of H_2O. Nano-ESI-MS gave a MW: 6035.5 ± 2.3 Da for the complex, indicating partial D/H exchange corresponding to 32 out of 44 sites per $\mathbf{G_{SS}G}$ [Table 1, ESI Figure B(iv)†]. This consistent pattern in the forward and reverse H/D exchange confirmed that dimer formation did indeed occur in solution *via* a specific type of interaction. Furthermore, this dimer had 12 protons per $\mathbf{G_{SS}G}$ that were exchange inert, possibly due to their involvement in hydrogen bonds. A likely model to explain these observations was Hoogsteen H-bonding by the guanine nucleobases in a cyclic fashion in the dimer $(\mathbf{G_{SS}G})_2$. This would protect 2 protons per guanine or 12 protons per $\mathbf{G_{SS}G}$. This implies tetrad formation by the guanines of $(\mathbf{G_{SS}G})_2$, leading to a PNA analog of a hairpin-type DNA quadruplex[18] with the Gly–Cys–Cys–Gly motif acting as a loop.

In order to verify this model, a solution of $(\mathbf{G_{SS}G})_2$ in buffer, was subjected to a temperature dependent UV absorption study (Varian Cary 1E UV/Vis spectrophotometer) at 305 nm. Quadruplexes resulting from tetrad formation are characterized by an inverse sigmoidal UV-melting profile at 305 nm.[19] The UV trace of 100 μM $(\mathbf{G_{SS}G})_2$ in buffer, showed a characteristic inverse sigmoidal shape at 305 nm, confirming that it indeed forms a quadruplex (see ESI†). A concentration dependent study of the melting temperature, $T_{1/2}$, of $(\mathbf{G_{SS}G})_2$ revealed an increase in $T_{1/2}$ with an increase in strand concentration (Table 2), reaffirming bimolecular quadruplex formation by $\mathbf{G_{SS}G}$.

Disulfide bond formation could precede nucleobase recognition (Pathway **A**), or the disulfide bond could be formed after nucleobase recognition (Pathway **B**). In order to resolve which pathway predominated, equimolar $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ (500 μM each) in buffer were equilibrated for 10 h to allow quadruplex formation, if any, by $\mathbf{G_{SH}}$. During this equilibration, ~65% $\mathbf{G_{SH}}$ was oxidized to $\mathbf{G_{SS}G}$. Fig. 3A(i) shows the distribution of only the *remaining* thiol reactants $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ in the mixture at this time point, which is 1.0:3.2 $\mathbf{G_{SH}}$:$\mathbf{T_{SH}}$. Sodium perborate was added to this partially equilibrated mixture, to oxidize the *remaining* $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$. In the case of pathway **A**, $\mathbf{G_{SH}}$ is not pre-organized and hence perborate oxidation should yield a disulfide distribution obtained by the statistical combination of 1.0:3.2 $\mathbf{G_{SH}}$:$\mathbf{T_{SH}}$ (*i.e.*, a 0.16:1.0:1.6 distribution of $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$). Instead, the observed product distribution upon perborate oxidation, Fig. 3A(ii), was 1.32:1.0:4.55 $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$ (see ESI for

Table 1 Molecular weight of $\mathbf{G_{SS}G}$ as determined by nanoelectrospray ionisation mass spectrometry (Nano-ESI-MS) before and after H/D exchange[a]

Trial	Observed *m*/*z* (ion)	MW of M_2 (Da) observed [calcd][b]	Av. MW of M_2 (Da) observed [calcd][b]	No. of H/D exchanges per M
1[c]	1202.8 ($[M_2 + 5H]^{5+}$); 1207.9 ($[M_2 + Na + 4H]^{5+}$); 1215.2 ($[M_2 + Na + K + 2H]^{5+}$)	6008.9 ± 2 [6009.9]	3004.5 ± 1 [3004.9]	0
2[d]	1215.9 ($[M_2 + 5H]^{5+}$); 1235.5 ($[M_2 + 2K + Na + 2H]^{5+}$); 1239.9 ($[M_2 + 2Na + 2K + H]^{5+}$)	6074.5 ± 2.3 [6074.4]	3037.3 ± 1.2 [3037.2]	32 ± 1
3[e]	1220.4 ($[M_2 + 5H]^{5+}$); 1224.8 ($[M_2 + 2K + Na + 2H]^{5+}$); 1227.8 ($[M_2 + 2Na + 2K + H]^{5+}$)	6097.0 ± 2.1 [6098.4]	3048.5 ± 1.1 [3049.2]	44 ± 1
4[f]	1208.1 ($[M_2 + 5H]^{5+}$); 1212.9 ($[M_2 + 4H + Na]^{5+}$); 1228.1 ($[M_2 + 2H + 2K + Na]^{5+}$)	6035.5 ± 2.3 [6034.1]	3017.8 ± 1.2 [3017.1]	32 ± 1[g]

[a] Nano-ESI-MS carried out with a cone voltage of 60 eV, source temperature of 30 °C and analyzer pressure of 10^{-5} bar. [b] Molecular weight calculated assuming the number of exchanges listed in column. [c] 250 μM $\mathbf{G_{SS}G}$ in H_2O equilibrated for 8 h (Sample 1). [d] Sample 1 lyophilized and resuspended in D_2O. [e] 250 μM $\mathbf{G_{SS}G}$ in D_2O equilibrated for 8 h (Sample 2). [f] Sample 2 resuspended in an equal volume of H_2O. [g] D/H exchange.

Table 2 UV-melting temperature[a], $T_{1/2}$, of tetramolecular and bimolecular PNA quadruplexes from $\mathbf{G_{SH}}$ and $\mathbf{G_{SS}G}$ respectively at various PNA strand concentrations and ions[b]

$\mathbf{G_{SH}}$ (200 μM)	$T_{1/2}$ (°C)	$\mathbf{G_{SH}}$ (100 mM KCl)	$T_{1/2}$ (°C)	$\mathbf{G_{SS}G}$ (250 μM)	$T_{1/2}$ (°C)	$\mathbf{G_{SS}G}$ (100 mM KCl)	$T_{1/2}$ (°C)
100 mM K^+	28	50 μM	23	100 mM K^+	38	25 μM	24
100 mM Na^+	25	100 μM	26	100 mM Na^+	32	100 μM	31
100 mM Li^+	—[c]	400 μM	33	100 mM Li^+	—[c]	400 μM	44

[a] Absorbance followed at 305 nm; heating rate 1 °C min^{-1}, errors ±1 °C. [b] Samples prepared in 50 mM Tris, pH 7.4. [c] Quadruplex formation was not detectable.

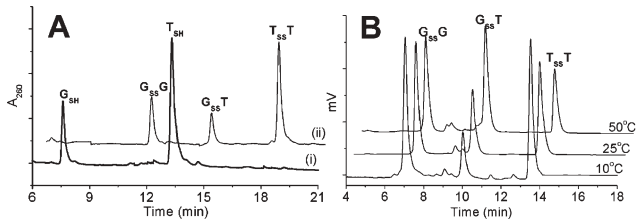

Fig. 3 **A)** (i) Distribution of only the *remaining* thiol reactants $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ in a partially, equilibrated reaction mixture (t = 10 h) starting with 500 μM $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ in buffer. (ii) Disulfide distribution obtained following perborate oxidation of remaining thiol reactants $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ [Trace (i)]. **B)** HPLC profiles of equimolar 500 μM $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ in buffer that has undergone air oxidation at various temperatures.

details†). This corresponds to a >8-fold amplification of $\mathbf{G_{SS}G}$ relative to $\mathbf{G_{SS}T}$ as compared to the statistical ratio. Such a pronounced deviation from the statistical ratio in favor of $\mathbf{G_{SS}G}$ indicates significant pre-organization of $\mathbf{G_{SH}}$ prior to oxidation, revealing that it is pathway **B** that predominates.

Air oxidation of 500 μM $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ in buffer was carried out at 10, 25 and 50 °C (Fig. 3B). At 25 °C, a 2:1:2 $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$ ratio was obtained. Lowering the temperature to 10 °C resulted in a 3.3:1.0:3.4 $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$ product distribution, showing a further enhancement in DCC amplification of $\mathbf{G_{SS}G}$. This is consistent with the UV-melting experiments (see ESI†) which show that at 10 °C, nearly all $\mathbf{G_{SS}G}$ formed exists as a quadruplex. Using the same argument, at 50 °C, neither $\mathbf{G_{SH}}$ nor $\mathbf{G_{SS}G}$ should be able to form a quadruplex (Table 2). An HPLC profile of the end point of air oxidation conducted at 50 °C evidenced a near-statistical product distribution (1.0:1.7:1.0 $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$). Increasing temperature progressively destabilizes both the templating tetramolecular PNA quadruplex as well as the covalently trapped bimolecular PNA quadruplex and this is reflected by the concomitant decrease in DCC amplification of $\mathbf{G_{SS}G}$.

Furthermore, quadruplex stability shows a characteristic ion dependence.[20] Therefore, the effect of cations on the air oxidation of 500 μM $\mathbf{G_{SH}}$ and $\mathbf{T_{SH}}$ was also investigated. In the presence of 100 mM NaCl, rather than K^+, a 1:1:1 product distribution was observed (Fig. 2A), suggestive of a reduced self-templating effect. The presence of 100 mM LiCl completely destroyed templating, giving a statistical product distribution 1.0:2.01:1.04 $\mathbf{G_{SS}G}$:$\mathbf{G_{SS}T}$:$\mathbf{T_{SS}T}$. The thermal stability of the tetramolecular PNA quadruplex decreases as $K^+ > Na^+ > Li^+$, analogous to DNA quadruplexes (Table 2).[20] PNA quadruplex formation was not detectable in the presence of Li^+. Templating also showed a pronounced cation dependence ($K^+ > Na^+ > Li^+$) consistent with quadruplex stability. This reaffirms that, for this system, greater stabilization of the templating assembly translates into a higher product amplification using DCC.

This shows that PNA quadruplex formation can act as a thermodynamic sink and drive product amplification by dynamic covalent chemistry (DCC). The mechanism of product amplification here involves the formation of the templating quadruplex assembly prior to covalent bond formation.

We thank the BBSRC for funding and Dr S. Otto for helpful discussions. Y. K. G. is an 1851 Research Fellow.

Yamuna Krishnan-Ghosh,[ab] Andrew M. Whitney[a] and Shankar Balasubramanian*[a]
[a]*University Chemical Laboratories, Department of Chemistry, Lensfield Road, University of Cambridge, Cambridge, UK CB2 1EW. E-mail: sb10031@cam.ac.uk; Fax: +44 1223 336913; Tel: +44 1223 336347*
[b]*National Centre for Biological Sciences, TIFR, UAS-GKVK Campus, Bellary Road, Bangalore 560 065, India. E-mail: yamuna@ncbs.res.in; Fax: +91 80 23636675; Tel: +91 80 23636421*

Notes and references

1 (*a*) M. Fujita, F. Ibukoro, H. Hagihara and K. Ogura, *Nature*, 1994, **367**, 720–723; (*b*) T. J. Kidd, D. A. Leigh and A. J. Wilson, *J. Am. Chem. Soc.*, 1999, **121**, 1599–1600.
2 (*a*) S. J. Cantrill, S. J. Rowan and J. F. Stoddart, *Org. Lett.*, 1999, **1**, 1363–1366; (*b*) Y. Furusho, T. Hasegawa, A. Tsuboi, N. Kihara and T. Takata, *Chem. Lett.*, 2000, 18–19.
3 (*a*) S.-Y. Tam-Chang, J. S. Stehouwer and J. Hao, *J. Org. Chem.*, 1999, **64**, 334; (*b*) S. Ro, S. J. Rowan, A. R. Pease, D. J. Cram and J. F. Stoddart, *Org. Lett.*, 2000, **2**, 2411.
4 Y. Krishnan-Ghosh and S. Balasubramanian, *Angew. Chem. Int. Ed.*, 2003, **42**, 2173.
5 A. Saghatelian, Y. Yokobayashi, K. Soltani and M. R. Ghadiri, *Nature*, 2001, **409**, 797; B. Bilgiçer, X. Xing and K. Kumar, *J. Am. Chem. Soc.*, 2001, **123**, 11815; M. A. Case and G. L. McLendon, *J. Am. Chem. Soc.*, 2000, **122**, 8089.
6 Y. Krishnan-Ghosh, E. Stephens and S. Balasubramanian, *J. Am. Chem. Soc.*, 2004, **126**, 5944.
7 R. Williamson, M. K. Raghuraman and T. Cech, *Cell*, 1989, **59**, 871; W. Guschlbauer, J. F. Chantot and D. Thiele, *J. Biomol. Struct. Dyn.*, 1990, **8**, 491–511; J. T. Davis, *Angew. Chem. Int. Ed.*, 2004, **43**, 668.
8 F. Seela, C. Wei and A. Melenewski, *Nucleic Acids Res.*, 1996, **24**, 4940.
9 Y. Wang and D. J. Patel, *Biochemistry*, 1992, **31**, 8112.
10 C. J. Cheong and P. B. Moore, *Biochemistry*, 1992, **31**, 8406.
11 B. Datta, C. Schmitt and B. A. Armitage, *J. Am. Chem. Soc.*, 2003, **125**, 4111.
12 D. Sen and W. Gilbert, *Biochemistry*, 1992, **31**, 65; Y. Krishnan-Ghosh, D. Liu and S. Balasubramanian, *J. Am. Chem. Soc.*, 2004, **126**, 11009.
13 R. P. Szajewski and G. M. Whitesides, *J. Am. Chem. Soc.*, 1980, **102**, 2011.
14 A. McKillop and J. A. Tarbin, *Tetrahedron Lett.*, 1983, **24**, 1505.
15 Unless otherwise mentioned, buffer used was 50 mM Tris, 100 mM KCl, pH 7.4.
16 S. Otto, R. L. E. Furlan and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2000, **122**, 12063.
17 D. R. Goodlett, D. G. Camp, C. C. Hardin, M. J. Corregan and R. D. Smith, *Biol. Mass Spectrom.*, 1993, **22**, 181; S. A. Hofstadler and R. H. Griffey, *Chem. Rev.*, 2001, **101**, 377.
18 C. Kang, X. Zhang, R. Ratliff, R. Moyzis and A. Rich, *Nature*, 1992, **356**, 126–131; F. W. Smith and J. Feigon, *Nature*, 1992, **356**, 164–168.
19 J.-L. Mergny, A.-T. Phan and L. Lacroix, *FEBS Lett.*, 1998, **435**, 74.
20 H. Deng and W. D. Braunlin, *J. Mol. Biol.*, 1996, **255**, 476.

