ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2017, 53, 12704

Received 13th September 2017,
Accepted 2nd November 2017

DOI: 10.1039/c7cc07173f

rsc.li/chemcomm

Structural basis for the selective incorporation of an artificial nucleotide opposite a DNA adduct by a DNA polymerase†

K. Betz,[a] A. Nilforoushan,[b] L. A. Wyss,[b] K. Diederichs,[c] S. J. Sturla *[b] and A. Marx *[a]

The possibility to sequence cytotoxic O^6-alkylG DNA adducts would greatly benefit research. Recently we reported a benzimidazole-derived nucleotide that is selectively incorporated opposite the damaged site by a mutated DNA polymerase. Here we provide the structural basis for this reaction which may spur future developments in DNA damage sequencing.

There have been major recent advances addressing the distribution of epigenetically encoded DNA methylation in the genome due to the application of bisulfite sequencing,[1] however, knowledge concerning the locations of chemically induced base alkylation lags far behind due to a lack of strategies for mapping it. An envisioned basis for DNA alkylation sequencing involves the use of artificial nucleotides incorporated opposite alkylated guanine by a DNA polymerase. Chemical alkylation of DNA by carcinogens or chemotherapeutics may occur, for example, at the O^6 position of guanine resulting in O^6-alkylG DNA adducts that induce toxicity and promote the misincorporation of T, leading to G-to-A mutations prevalent in human cancers.[2–5]

Recent DNA damage sequencing advances include versions of single molecule real time sequencing,[6] and the use of DNA repair enzymes for excising oxidized or cross-linked bases[7–10] and mapping these locations in the genome.[11,12] For DNA polymerase-mediated selective amplification from O^6-alkylG adducts,[13] KlenTaq DNA polymerase (large fragment of *Thermus aquaticus* DNA polymerase I) has been shown to process an artificial benzimidazole-derived nucleoside triphosphate (Fig. 1, BenziTP) opposite O^6-methylguanine (O^6-MeG) more efficiently than opposite guanine.[13,14] A mutated KlenTaq DNA polymerase

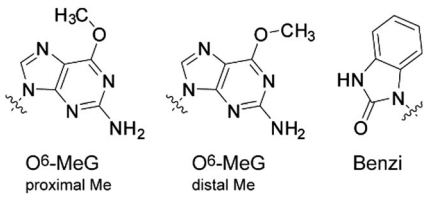

Fig. 1 Structure of O^6-methylguanosine with the methyl group in the proximal *vs.* distal orientation, and the artificial nucleotide Benzi. The wavy line indicates attachment in DNA or on 2′-deoxynucleoside-5′-*O*-triphosphate.

was even more proficient and incorporated BenziMP up to 150-fold more efficiently opposite O^6-MeG than opposite guanine.[13,14] The basis for the proficiency of the wildtype enzyme and the increased efficiency by the mutant DNA polymerase are not known, however, therefore limiting the knowledge needed to further engineer polymerase-based strategies for sequencing adducts.

Engineering DNA polymerases has led to the creation of enzymes with various new functions and applicability.[15] The engineered KlenTaq DNA polymerase (KlenTaq M747K) found to replicate O^6-alkylG adducts carries a methionine to lysine mutation at amino acid position 747 that makes it more efficient in bypassing various DNA lesions.[13,16,17] The mutation site is located next to the negatively charged DNA template backbone near the active site of the enzyme and substitution by the lysine residue generated an increased positively charged surface potential that is assumed to promote translesion synthesis.[16,17]

KlenTaq M747K discriminated between O^6-benzylG *vs.* G, in the insertion of BenziMP (Fig. 1) with significantly higher efficiency than the wild type enzyme, and also exhibited a low rate of misincorporation of natural nucleotides opposite O^6-MeG, a property that is required for adduct sequencing. Compared to BenziMP the most efficiently incorporated natural nucleotide opposite O^6-MeG is dTMP (19-fold less efficiently incorporated by M747K than BenziMP).[13] The effect mainly relies on the lower K_M for BenziTP (18 μM *vs.* 410 μM) as the k_{cat} is almost the same for BenziTP and dTTP (15.6 min^{-1} *vs.* 13.2 min^{-1}). The efficiency of BenziMP incorporation opposite O^6-MeG is

[a] *Department of Chemistry & Konstanz Research School Chemical Biology, University of Konstanz, Universitätsstr. 10, D-78457 Konstanz, Germany. E-mail: andreas.marx@uni-konstanz.de; Tel: +49 7531 885139*
[b] *Department of Health Sciences and Technology, ETH Zurich, Schmelzbergstrasse 9, CH-8092 Zurich, Switzerland. E-mail: sturlas@ethz.ch; Tel: +41 44 6329175*
[c] *Department of Biology & Konstanz Research School Chemical Biology, University of Konstanz, D-78457 Konstanz, Germany*

† Electronic supplementary information (ESI) available: Experimental details, data collection statistics. See DOI: 10.1039/c7cc07173f

still lower compared to the natural incorporation of dCMP opposite dG (260 fold less efficient for M747K[14]) The difference relies mainly on the higher K_M of BenziTP while k_{cat} is almost the same. Structural data would be highly beneficial for optimizing nucleotide design towards higher incorporation efficiency which is crucial for furthering the sequencing approach. Here, we provide structural insights into how the O^6-MeG:Benzi pair is processed by KlenTaq M747K.

To investigate the preferential incorporation of BenziTP opposite O^6-MeG on a structural level we crystallized KlenTaq M747K in a ternary complex with the substrate BenziTP paired opposite a templating O^6-MeG in the active site (structure is termed $M747K_{O^6\text{-MeG:Benzi}}$).‡ For comparison we also solved the structure of KlenTaq M747K complexed with a natural dG:dCTP pair which is termed $M747K_{G:C}$. Analysis of the structures revealed that BenziTP forms a Watson–Crick-like base pair together with O^6-MeG with some variations in positioning of the adduct methyl group and the substrate sugar pucker.

The structure $M747K_{O^6\text{-MeG:Benzi}}$ was obtained by crystallizing KlenTaq M747K in a binary complex with a primer/template duplex carrying O^6-MeG at the templating position. The crystals were soaked with BenziTP together with up to 100 mM $MnCl_2$. Soaking without manganese ions did not result in a structure of a closed ternary complex. It was described that presence of Mn^{2+} can lower the K_M of nucleotides binding to a DNA polymerase[18] which might aid to form a stable complex since the K_M of BenziTP opposite O^6-MeG is two orders of magnitude higher than the K_M of the unmodified nucleotide in the same context (*e.g.* 18 μM for BenziTP opposite O^6-MeG *versus* 0.7 μM for dCTP opposite dG as determined in[14]). Ternary DNA polymerase complexes with manganese ions have already been obtained before and the position as well as the coordination patterns of the ions in the active sites were highly similar compared to the smaller magnesium ions.[19,20]

The $M747K_{G:C}$ structure was solved by cocrystallization of KlenTaq M747K with a natural primer/template complex and dCTP in the presence of magnesium ions. It is very similar to the KlenTaq WT structure with the same primer/template and substrate complexed (PDB ID. 3RTV;[21] rmsd for Cα atoms: 0.224, Fig. S1A and B, ESI†). Just as 3RTV, the structure $M747K_{G:C}$ is in a closed conformation and coordination of the two Mg^{2+} ions and stabilization of the triphosphate in the active site is the same. (Fig. S1B, ESI†) The only significant difference is the Met747Lys mutation, which points towards the negatively charged template strand (Fig. S1C and D, ESI†). The electrostatic map in Fig. S1C and D (ESI†) visualizes the gained positive charge that may help to stabilize the template strand and promote selective lesion-bypass synthesis as was described before.[16]

The structure $M747K_{O^6\text{-MeG:Benzi}}$ also shows a closed enzyme conformation and superposes well with $M747K_{G:C}$ (rmsd for Cα atoms: 0.182). The substrate BenziTP is bound in the active site and adopts an *anti* conformation.[14,22] It pairs with the O^6-MeG templating nucleotide *via* two hydrogen bonds: one between the N1 of O^6-MeG and the –NH donor on BenziTP (3.2 Å) and the second between the $-NH_2$ donor on O^6-MeG and the carbonyl group of BenziTP (2.9 Å, Fig. 2). After superimposition of the nascent base-pairs in $M747K_{O^6\text{-MeG:Benzi}}$ and $M747K_{G:C}$ based on the templating nucleotides the slightly larger propeller twist of the O^6-MeG–Benzi pair becomes apparent (Fig. 2C). The simulated annealing omit map of both modified nucleotides is shown in Fig. 2A. For O^6-MeG the density did not unambiguously define if the methyl group points towards or away from the BenziTP. We modelled both possibilities at the same time and refined their occupancies. Thereby the proximal O^6-MeG refined to an occupancy of only 0.3, whereby the distal O^6-MeG had an occupancy of 0.7. Therefore we decided to model the distal conformation keeping in mind that the alternative proximal conformation may also be populated in the crystal.

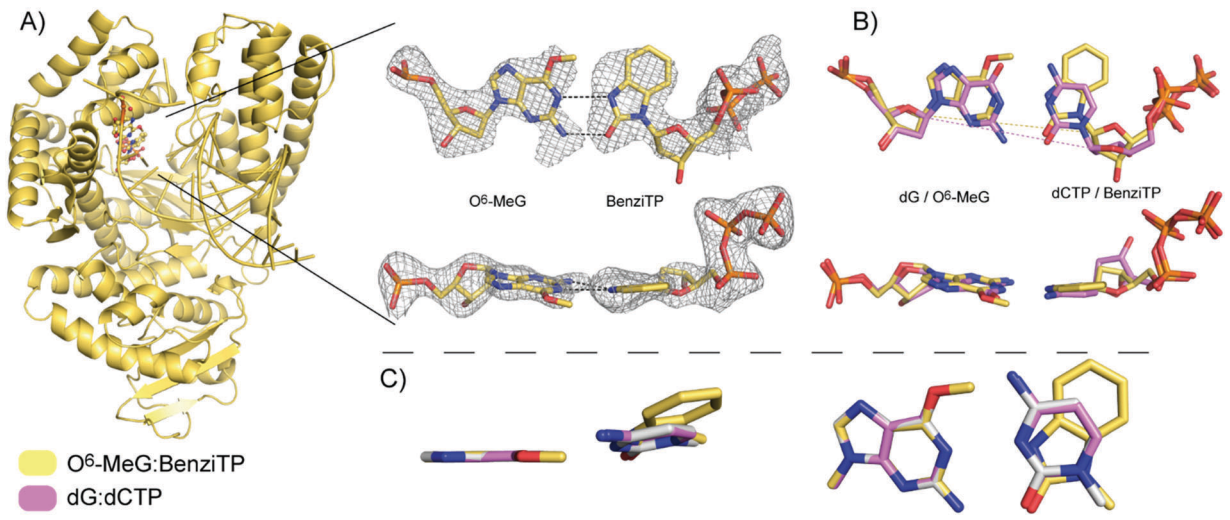

Fig. 2 (A) Structure of $M747K_{O^6\text{-MeG:Benzi}}$. The zoomed in portion shows the simulated annealing omit map for O^6-MeG and BenziTP contoured at 3σ. Hydrogen bonds are shown as dashed lines and distances are given in the text. (B) Overlay with the natural dG:dCTP base pair of $M747K_{G:C}$. The C1′–C1′ distances are indicated by dashed lines. (C) Detailed view of two different orientations of the base pairs superimposed on the templating nucleobases with larger propeller twist of O^6-MeG:Benzi is visible.

Depending on the position of the methyl group, the BenziTP is slightly shifted. As the distal methyl group comes closer to the phenyl ring of BenziTP (closest distance 3.4 Å) BenziTP is shifted further away from the pairing partner resulting in a little enlarged base-pair width compared to the natural case (C1′–C1′-distance 11.0 Å *vs.* 10.6 Å, Fig. 2B).

The electron density profiles for the ribose moiety suggest that several sugar conformations may be present at the same time in the crystal. We placed two different conformations (C2′-*endo* and C3′-*endo*), performed one round of occupancy refinement for the two BenziTP ligands and found a 1/3 to 2/3 occupancy ratio for C3′-*endo versus* C2′-*endo*. For natural substrates the common sugar mostly observed is C3′-*endo* in the KlenTaq DNA polymerase active site.[23] The triphosphate moiety is very well defined in the electron density and interactions with polar residues of the O-helix allow the same stabilization as in the natural case (Fig. 3A and B). Additional stabilization may be possible with Asn750 and Gln754 and the carbonyl group of Benzi *via* a water molecule. The distance to the water molecule, however, is larger than in the natural case (3.4 Å *versus* 2.8 Å).

Characteristic for an active complex prior to catalysis is that the finger domain is in a closed conformation and the triphosphate and amino acid side chains within the palm domain coordinate two divalent metal ions. The residues involved in coordination in M747K$_{O^6MeG:Benzi}$ are the same as with a natural base-pair but the coordination shows a slightly different geometry (Fig. 3). The overlay in Fig. 3C shows the metal coordination in M747K$_{O^6MeG:Benzi}$ and M747K$_{G:C}$. Metal ion B which is coordinated in the typical octahedral geometry by the triphosphate and Asp610, Asp785 and the backbone carbonyl of Tyr611 is at the same position. This ion is postulated to stabilize the pentacovalent intermediate during the insertion reaction and neutralize the negative charge of the leaving pyrophosphate.[24] The second ion at position A is 3.7 Å away from metal ion B (distance between Mg^{2+}-ions in natural case: 3.6 Å) and displaced with respect to the usual metal position A by 1.3 Å (Fig. 3C). It is still coordinated by the alpha phosphate, Asp610 and two water molecules but the two water molecules are at different positions resulting in a distorted coordination geometry. As Asp785 is too far away for the displaced ion it shows only four coordination partners. In both complexes an additional coordination of metal ion A would be mediated by the 3′-OH group of the primer end which is not present in our structures since we are using a dideoxy-terminated primer as a common strategy to capture the ternary complex in crystallization. The assigned role of metal ion A is to lower the pK_a of that terminal OH group for the nucleophilic attack on the alpha phosphate.[25,26] Why this different ion binding (along with the different sugar conformation) occurs and if it still displays an active complex from which the catalytic reaction would occur or if we trapped the structure in a pre-catalytic state from which catalysis can take place only after an additional conformational change is not clear.

An additional difference to the natural complex is the fact that the enzyme is not as tightly closed, which is indicated by higher flexibility of the whole finger domain (high B-factors and weak electron density in that region, Fig. S2A, ESI†). This might be due to the larger size of the nucleobase-pair where the methyl group attached to dG and the phenyl ring of BenziTP both extend beyond the consensus pocket of natural base-pairs (Fig. S2B, ESI†).[27] Similar flexibility has already been observed in KlenTaq DNA polymerase complexes with artificial base-pairs and non cognate substrates.[21,28,29] An anticipated outcome may be reduced capacity to stabilize the substrate/enzyme complex, thus allowing some deviation from the optimal geometry for catalysis. Finally this conformational heterogeneity might account for the observed impaired enzymatic efficiency of the DNA polymerase when processing modified substrates in comparison to the cognate substrates.

Another difference in M747K$_{O^6\text{-}MeG:Benzi}$ compared to M747K$_{G:C}$ is the interaction of Arg600 with the 3′-primer terminus. The Arg660 side chain is rotated compared to its position in the natural dG:dCTP case and loses its interaction with the primer strand (Fig. S3, ESI†). Instead, Arg587 is rotated towards the primer 3′-end, taking over the lost interaction. The size of the BenziTP nucleobase however does not explain, why Arg660 is shifted as it would not clash with the side chain in a position observed in M747K$_{G:C}$ (Fig. S3C, black arrow, ESI†). Concerning the position of Arg660 and Arg587 M747K$_{O^6MeG:Benzi}$ seems to be similar to a KlenTaq complex with a bound dGTP (PDB ID:1QSS,[30] Fig. S3D, ESI†)

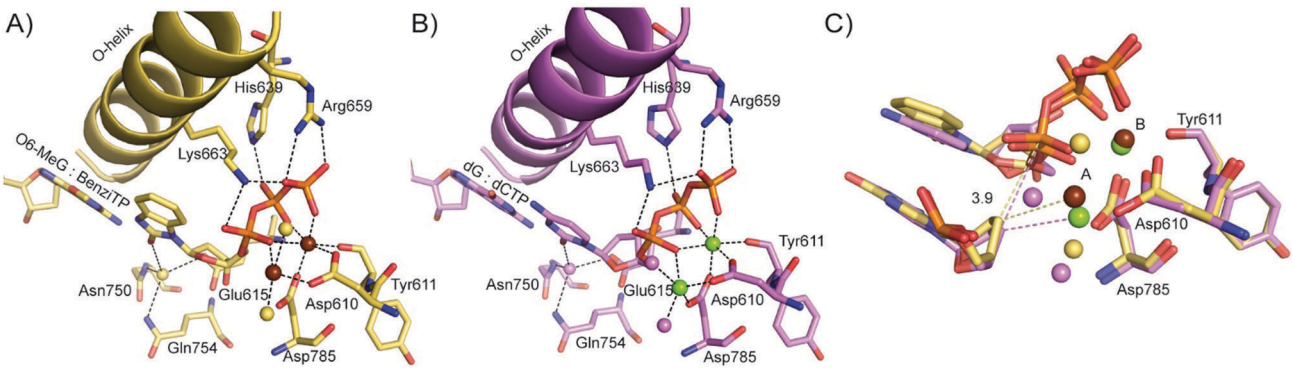

Fig. 3 (A and B) Interactions with the triphosphate and coordination of active site metal ions for M747K$_{O^6\text{-}MeG:Benzi}$ (yellow) and M747K$_{G:C}$ (violet). Interactions are shown as black dashed lines. Manganese ions are shown in brown and magnesium ions are shown in green. Water molecules are shown as yellow or violet spheres. (C) Overlay of the metal ions and interacting residues visualizes the shifts described in the text. The distance of the primer 3′-end and the alpha phosphate is 3.9 Å in both complexes.

as this is the only nucleobase that was found to cause a different positioning of these residues in the KlenTaq DNA polymerase structures elucidated by Waksman *et. al.*[30,31] Based on the position of Arg660 in 1QSS the different shape of BenziTP would indeed cause a clash with Arg660 (Fig. S3D, ESI†) and therefore explain why it is rotated away.

In summary we could show why the synthetic adduct-directed nucleotide BenziTP is specifically incorporated opposite the alkylation-induced DNA adduct O^6-MeG. The structural composition of both components allow formation of a Watson–Crick-like base-pair mediated by two hydrogen bonds with a size and geometry similar to the cognate base-pairs to be readily accepted by KlenTaq DNA polymerase and its M747K mutant. Neither should the other four natural nucleotides be able to form an equally similar base-pair with O^6-MeG nor should Benzi pair in the same way with natural G. The study shows that KlenTaq DNA polymerase forms an active site that is able to accommodate the unnatural base-pair. While the mutated amino acid seems to be not directly involved in selecting the nucleotide opposite O^6-MeG *e.g.*, by forming hydrogen bonds to BenziTP, it might contribute nonetheless to stabilizing the conformations at the active site that favour selective incorporation by increasing the positive electrostatic potential at the negatively charged template phosphate backbone site (Fig. S1C and D, ESI†). The fact that well-diffracting crystals could only be obtained in the presence of manganese ions, and the higher flexibility of the substrates and the finger domain in the complex might be indicative for the lower incorporation efficiency of BenziMP opposite O^6-MeG compared to a canonical base pair. The structural data provided here might guide optimization of the processes *e.g.*, by altering the chemical structure of the triphosphate or further mutating the enzyme.

We thank the beamline staff of the Swiss Light Source at the Paul Scherrer Institute for access and help at the beamline. We also acknowledge the European Research Council (Grant 260341) and the Swiss National Science Foundation (Grant 156280) (S. J. S.) for funding.

Conflicts of interest

There are no conflicts to declare.

Notes and references

‡ Atomic coordinates and structure factors have been deposited in the Protein Data Bank with PDB codes 5O7T for $M747K_{G:C}$ and 5OXJ for $M747K_{O^6\text{-MeG:Benzi}}$.

1 H. Heyn and M. Esteller, *Nat. Rev. Genet.*, 2012, **10**, 679–692.
2 G. T. Pauly and R. C. Moschel, *Chem. Res. Toxicol.*, 2001, **14**, 894–900.
3 M. G. Pence, J.-Y. Choi, M. Egli and F. P. Guengerich, *J. Biol. Chem.*, 2010, **285**, 40666–40672.
4 G. P. Pfeifer, M. F. Denissenko, M. Oliver, N. Tretyakova, S. S. Hecht and P. Hainaut, *Oncogene*, 2002, **21**, 7435–7451.
5 The Cancer Genome Atlas Network, *Nature*, 2012, **487**, 330–337.
6 T. A. Clark, K. E. Spittle, S. W. Turner and J. Korlach, *Genome Integr.*, 2011, **2**, 10.
7 A. G. Zavala, R. T. Morris, J. J. Wyrick and M. J. Smerdon, *Nucleic Acids Res.*, 2014, **42**, 893–905.
8 D. S. Bryan, M. Ransom, B. Adane, K. York and J. R. Hesselberth, *Genome Res.*, 2014, **24**, 1534–1542.
9 J. Hu, S. Adar, C. P. Selby, J. D. Lieb and A. Sancar, *Genes Dev.*, 2015, **29**, 948–960.
10 J. Hu, J. D. Lieb, A. Sancar and S. Adar, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 11507–11512.
11 J. Riedl, Y. Ding, A. M. Fleming and C. J. Burrows, *Nat. Commun.*, 2015, **6**, 8807.
12 J. Riedl, A. M. Fleming and C. J. Burrows, *J. Am. Chem. Soc.*, 2016, **138**, 491–494.
13 L. A. Wyss, A. Nilforoushan, F. Eichenseher, U. Suter, N. Blatter, A. Marx and S. J. Sturla, *J. Am. Chem. Soc.*, 2015, **137**, 30–33.
14 L. A. Wyss, A. Nilforoushan, D. M. Williams, A. Marx and S. J. Sturla, *Nucleic Acids Res.*, 2016, **44**, 6564–6573.
15 G. Houlihan, S. Arangundy-Franklin and P. Holliger, *Acc. Chem. Res.*, 2017, **50**, 1079–1087.
16 C. Gloeckner, K. Sauter and A. Marx, *Angew. Chem., Int. Ed.*, 2007, **46**, 3115–3117.
17 S. Obeid, A. Schnur, C. Gloeckner, N. Blatter, W. Welte, K. Diederichs and A. Marx, *ChemBioChem*, 2011, **12**, 1574–1580.
18 M. F. Goodman, S. Keener, S. Guidotti and E. W. Branscomb, *J. Biol. Chem.*, 1983, **258**, 3469–3475.
19 V. K. Batra, W. A. Beard, D. D. Shock, L. C. Pedersen and S. H. Wilson, *Mol. Cell*, 2008, **30**, 315–324.
20 E. Y. Wu, A. R. Walsh, E. C. Materne, E. P. Hiltner, B. Zielinski, B. R. Miller, L. Mawby, E. Modeste, C. A. Parish, W. M. Barnes and M. B. Kermekchiev, *Biochemistry*, 2015, **54**, 881–889.
21 K. Betz, D. A. Malyshev, T. Lavergne, W. Welte, K. Diederichs, T. J. Dwyer, P. Ordoukhanian, F. E. Romesberg and A. Marx, *Nat. Chem. Biol.*, 2012, **8**, 612–614.
22 A. Nilforoushan, A. Furrer, L. A. Wyss, B. van Loon and S. J. Sturla, *J. Am. Chem. Soc.*, 2015, **137**, 4728–4734.
23 K. Betz, F. Streckenbach, A. Schnur, T. Exner, W. Welte, K. Diederichs and A. Marx, *Angew. Chem., Int. Ed.*, 2010, **49**, 5181–5184.
24 V. Sosunov, E. Sosunova, A. Mustaev, I. Bass, V. Nikiforov and A. Goldfarb, *EMBO J.*, 2003, **22**, 2234–2244.
25 L. S. Beese and T. A. Steitz, *EMBO J.*, 1991, **10**, 25.
26 S. Xia, M. Wang, G. Blaha, W. H. Konigsberg and J. Wang, *Biochemistry*, 2011, **50**, 9114–9124.
27 E. T. Kool, *Annu. Rev. Biochem.*, 2002, **71**, 191–219.
28 K. Betz, M. Kimoto, K. Diederichs, I. Hirao and A. Marx, *Angew. Chem., Int. Ed.*, 2017, **56**, 12000–12003.
29 S. Obeid, W. Welte, K. Diederichs and A. Marx, *J. Biol. Chem.*, 2012, **287**, 14099–14108.
30 Y. Li, V. Mitaxov and G. Waksman, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 9491–9496.
31 Y. Li, S. Korolev and G. Waksman, *EMBO J.*, 1998, **17**, 7514–7525.