Birkbeck
UNIVERSITY OF LONDON

BIROn - Birkbeck Institutional Research Online

Journal Name

COMMUNICATION

XNA ligation using T4 DNA ligase in crowding conditions

Donaat Kestemont[a][†][*], Marleen Renders[a][†], Piotr Leonczak[a], Mikhail Abramov[a], Guy Schepers[a], Vitor Pinheiro[b,c], Jef Rozenski[a], Piet Herdewijn[a,d]

T4 DNA ligase is capable of ligating 2'OMe-RNA duplexes, HNA, LNA and FANA mixed sequences in the presence of 10% w/v PEG8000 and 3 M betaine. The enzymatic joining of oligonucleotides containing multiple consecutive XNA nucleotides at the ligation site has not been reported before.

Backbone-modified nucleic acids are often enzymatically and chemically more stable than DNA and RNA, making them relevant in diagnostics or therapeutics as probes, aptamers, aptazymes, siRNAs or antisense RNAs,[1-5] as orthogonal genetic materials for the development of safe genetically contained organisms (or 'GCOs'),[6,7] or in nanotechnology.[8,9]

It is a significant technical challenge to synthesise a backbone-modified nucleic acid or an oligonucleotide with a diversity of backbone chemistries. DNA-dependent xeno nucleic acid (XNA) polymerases have been engineered and XNA synthesis is possible.[1] Nonetheless, the enzymes are not processive and can introduce errors during synthesis.[10] Some backbone-modified nucleic acids can be chemically synthesised, but the length of the modified fragments is often limited due to the lower coupling efficiency of the monomer phosphoramidites and between different chemistries on solid-phase, and for some chemistries, no solid-phase synthesis method is available.

The problems associated with the synthesis of longer, mixed-chemistry nucleic acid fragments, could be circumvented by using ligases to join shorter pieces of modified nucleic acids. This strategy has been used for the generation of highly functionalized base-modified DNA aptamers.[11,12] T4 DNA ligase catalyses the phosphodiester bond formation between the 3'-hydroxyl group ('acceptor') and the 5'-phosphate terminus ('donor') of juxtaposed oligonucleotides in nicked DNA or, in some cases, in a hybrid DNA/RNA or RNA duplex.[13-17] Additionally, it can join blunt and cohesive ends.[18] T4 DNA ligase has shown to tolerate some mismatches in certain conditions,[14,19-26] some modified bases[27] and has recently been reported to accept oligonucleotides with single backbone-modified nucleotide substitutions as a substrate, albeit with reduced efficiency.[28-30] The enzyme has shown to ligate DNA fragments unmodified at the ligation site but containing several PNA nucleotides.[31]

Crowding agents such as high molecular weight PEG, BSA, glycogen, Ficoll PM 70 and hexaminecobalt(III) chloride and organic solvents such as DMSO and formamide can be used to stimulate ligations using T4 DNA ligase, whether enabling ligation across mismatches or fine tuning the specificity of ligase detection reaction.[19,32-40] Additionally, a number of small molecules (MW < 1000 Da) such as 1,2-propanediol, ethylene glycol and straight or branched chain alcohols (e.g. ethanol and isopropanol) have been patented as potential enhancers of T4 DNA ligation activity.[41]

Despite the effect that additives have shown on the activity of T4 DNA ligase, the influence of these factors on the substrate spectrum of nucleic acid ligases has not been previously explored. We set out to screen a panel of commercially available DNA and RNA ligases using crowding conditions and cosolutes to test the potential impact on their substrate specificity. We focused our screen on the ligation of 2'OMe RNA, LNA ('locked' nucleic acids), FANA (2'-fluoroarabino nucleic acids) and HNA (1,5-dianhydrohexitol nucleic acids) as model XNAs as they cover a range of structures in the vicinity of both DNA and RNA. An overview of the oligonucleotides is given in Table S1.

In initial experiments used to measure the basal ligase activity to join DNA-2'OMe RNA chimeric duplexes, only T3 DNA ligase, T4 DNA ligase, T7 DNA ligase and SplintR DNA ligase showed ligation activity (Figure 1 left). T4 DNA consistently

a. *Laboratory for Medicinal Chemistry, Rega Institute for Medical Research, Herestraat 49 box 1041, 3000 Leuven, Belgium.*
b. *Structural and Molecular Biology Department, University College London, Gower Street, London, WC1E B6T, United Kingdom.*
c. *Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck College, University of London, Malet Street, London, WC1E 7HX, United Kingdom.*
d. *Université d'Evry, CNRS-UMR8030/Laboratoire iSSB, CEA, DRF, IG, Genoscope, Université Paris-Saclay, Evry 91000, France.*
*Tel: +3216322666; Fax: +3216330026; Email: marleen.renders@kuleuven.be
†These authors contributed equally. Electronic Supplementary Information (ESI) available: [xxx]. See DOI: 10.1039/x0xx00000x*

outperformed other ligases (Figure 1 right) and was selected for further optimization.
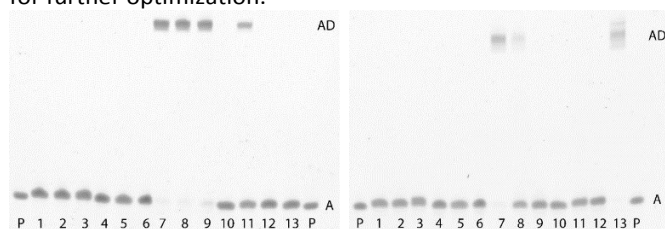

Figure 1. The ligation of A1 to D2 opposite T1 (left) or A2 to D2 opposite T1 (right) overnight with different ligases at the reported optimal temperature for each ligase, using 0.1 µM for the acceptor, 0.2 µM for the donor and the template oligonucleotides, and 1 µl of commercial enzyme per 10 µl reaction volume. P indicates the primer control, A indicates the position of the FAM-labelled acceptor oligonucleotide on the gel, whereas AD shows the position of the reaction product. The lanes in the gel contain the reaction catalysed by the following ligases: 1. *Rhodothermus marinus* DNA ligase, 2. *Thermus scotoductus* DNA ligase, 3. Taq DNA ligase, 4. Ampliligase, 5. 9°N DNA ligase, 6. Thermostable RNA ligase, 7. T4 DNA ligase, 8. T3 DNA ligase, 9. T7 DNA ligase, 10. *E. coli* DNA ligase, 11. SplintR DNA ligase, 12. T4 RNA ligase 1, 13. T4 RNA ligase 2

We screened the impact of molecular crowding agents and co-solutes on the ligation of 2'OMe RNA molecules (both donor and acceptor) against a DNA template, which are poor ligation substrates for the enzyme in standard conditions (Figure S1). Additionally, 5' deadenylase was tested because it has been reported to increase the ligation yield in reactions where the yield is limited due to the abortive formation of 5'-adenylated DNA end product. Betaine, PEG8000 and DMSO (Figure 2) had the biggest impact in enhancing the ligation, with more than 50% of the products ligated in the presence of betaine or PEG8000 in stark contrast to the non-detectable ligation levels in the absence of these enhancers, and were chosen for further optimization. DMSO and betaine enhanced the reaction even at the highest concentrations tested (40% v/v and 3 M respectively), while PEG8000 enhancement was optimum at around 10% w/v (See Table S2 for results). This is similar to the effect observed with PEG8000 and T7 DNA polymerase[42] and could, likewise, be explained by the increased binding of the enzyme to the nicked nucleic acid duplex as the PEG8000 concentration increases, together with a diminishing catalytic activity, hampering the ligation event at higher PEG8000 concentrations. We further optimised the reaction to identify the best ratio of the three components. Near-complete (>90%) ligation of two 2'OMe-RNA oligos against a 2'OMe-RNA template were achieved in 10% PEG8000, 3 M betaine, 10 mM Mg^{2+} and 10 µM ATP, in the presence of 4 U/µl enzyme and 0.1 µM of acceptor (A2) and 0.2 µM of the donor (D3) and template (T2) oligonucleotides in reactions carried out at 37°C for 16 hours (Figure 3). No ligation was observed in standard ligase reaction buffer in those conditions.

Having demonstrated that the PEG8000/betaine supplemented buffer enhanced 2'OMe RNA ligation, we set out to investigate whether it could be used to enhance ligation of a range of XNAs. As shown in Figure 3 and Figure S2, the optimized buffer enables ligation of a variety of XNAs in the close structural vicinity of DNA and RNA, generating molecules whose chemical synthesis (including multiple different

chemistries) would be challenging. The identity of the ligation products was confirmed by mass spectrometry (MS spectra in Figure 3) using biotinylated template strands for capture on streptavidin magnetic beads and elution of the ligated sequences using NaOH prior to the analysis. An alternative series of 2'OMe RNA and HNA oligonucleotides was ligated to demonstrate that the T4 DNA ligase-catalyzed ligation in these conditions is possible in multiple sequence contexts.
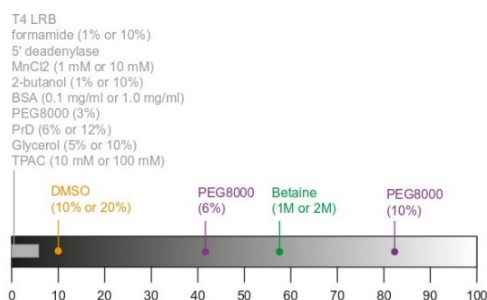

Figure 2. The bar represents the percentage of ligation of A2 to D3 using T1 as a template at a 0.1 µM (A2) or 0.2 µM (D3 and T1) concentration using T4 DNA ligase (4 U/µl concentration final) in the presence of (different concentrations of) the additives. The reaction in standard T4 DNA ligase reaction buffer is indicated by 'T4 LRB'. PrD represents 1,2-propanediol, TPAC signifies tetrapropyl ammonium chloride. The reactions were incubated at 25°C for 16 hours. All reactions were carried out in triplicate.

In a parallel strategy, we investigated whether the addition of DNA binding domains to T4 DNA ligase could also contribute to extend its substrate range.[43,44] Chimeric proteins were screened but no significant further reaction enhancement was observed (data not shown).

The activity of the enzyme was measured in a time course assay in the presence and absence of 10% PEG8000 and 3 M betaine. No increase in enzymatic activity could be observed in the presence of the crowding agents. On the contrary, a slight decrease in the ligation activity could be observed (Figure S3). An increased (and concentration dependent) binding of the enzyme to its substrate is observed in the presence of 10 % PEG8000, likely enabling the ligation of the unnatural substrates (data not shown).

Conclusions

Molecular crowding has been previously exploited to increase the catalytic activity of nucleic acid processing enzymes. In particular, polyethylene glycol (PEG) has been extensively used to enhance molecular interactions, including DNA binding proteins,[45] nucleic acid processing enzymes[32-36, 46-49] and DNA polymerases.[42, 50-52] Similarly, DMSO and betaine have shown a positive effect on DNA polymerase reactions, potentially by destabilizing unwanted secondary structures in double stranded DNA and reducing the base pair composition dependency of DNA melting in PCR.[53-55] Betaine could function as a protein stabiliser.[56-59] It has shown to improve Ligase Cycling Reactions (LCR), using a thermostable ligase, in the past.[60] However, it has also proven destabilizing to some proteins in certain conditions.[61-64]
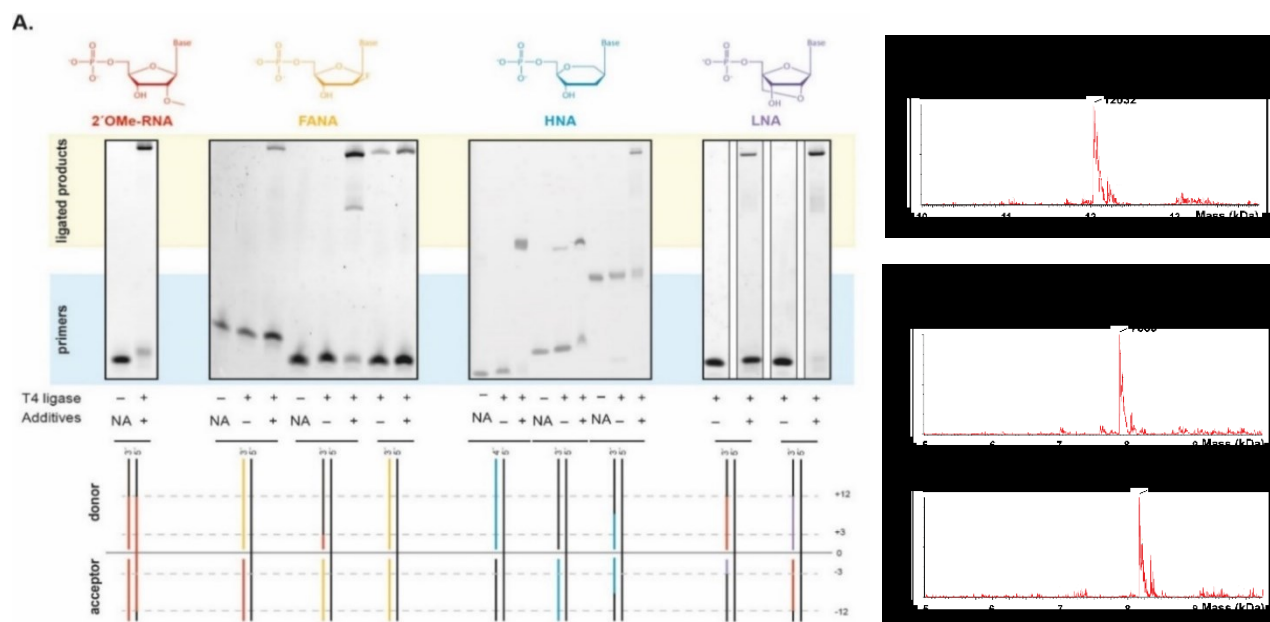
COMMUNICATION

Figure 3 The ligation of a 2'OMe RNA duplex and FANA, LNA and HNA chimeras in T4 DNA ligase reaction buffer with or without PEG8000 10% and 3 M betaine, in the optimized conditions. The ligation substrates are shown schematically under the figure, where red indicates 2'OMe RNA, yellow FANA, blue HNA and purple LNA. The oligos used were A2-D3/T2 (2'OMe RNA); A2-D4, A3-D2 and A3-D4/T1 for FANA; A7-D7/T4, A6-D9/T4 and A7-D8/T5, respectively for HNA; and A4-D3, and A2-D6/T1 for LNA. The deconvoluted mass spectra of the ligation of a 2'OMe RNA duplex (B, theoretical mass 12041) HNA as either the acceptor (C top, theoretical mass 7894) or as the donor (C bottom, theoretical mass 8180) in the reaction are shown in the right.

Here, we show that in crowding conditions T4 DNA ligase can function as an XNA ligase for a range of chemistries that have already been validated as synthetic genetic materials: 2'OMe RNA,[65] HNA, FANA and LNA.[1] Mixed chemistry polymers cannot be delivered by polymerases, but can be read out by a single XNA reverse transcriptase, enabling aptamer/aptazyme selections using mixed-chemistry oligomers.[1]

We observed that T4 DNA ligase accepts 2'OMe RNA and FANA much more as a substrate in the acceptor than in the donor position – in agreement with findings that T4 DNA ligase reaction rate decreases 32,000-fold from a RNA-DNA/DNA duplex to a DNA-RNA/DNA hybrid ligation at 37°C[66] and the differential effect of the substitution of a single DNA nucleotide by a 2'OMe RNA building block on either side of the nick[28]. This effect can be explained by the larger area of interaction between the ligase and its substrate on the 5'-end of the nick. The contact between T4 DNA ligase and its substrate has been determined to be seven to nine bases on the donor side of the nick and three to five bases on the acceptor side.[67]

The above findings allow the construction of long XNA or mixed chemistry oligonucleotides, which could be useful for producing non-toxic origami structures with an enhanced stability for in-cell delivery,[8] circular mRNA, aptamers or aptazymes with increased stability and enhanced reproducibility of binding due to a decreased tendency to fold into alternative secondary structures[68-70] or for siRNA applications. Additionally, in the field of xenobiology, the initial ligation of XNA-containing oligonucleotides is an important stepping stone towards the in vitro evolution of an XNA ligase and paves the way towards XNA applications in vivo (an XNA episome for the construction of safe GCOs). Finally, the above results show that it may be possible to use the strategy presented here to extend the substrate and reaction range of other nucleic acid processing enzymes greatly decreasing the engineering challenge for the development of XNA molecular biology.

Notes and references

1. V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S. Y. Peak-Chew, S. H. McLaughlin, P. Herdewijn and P. Holliger, *Science*, 2012, **336**, 341-344.
2. H. Yu, S. Zhang and J. C. Chaput, *Nat. Chem.*, 2012, **4**, 183-187.
3. P. E. Burmeister, S. D. Lewis, R. F. Silva, J. R. Preiss, L. R. Horwitz, P. S. Pendergrast, T. G. McCauley, J. C. Kurz, D. M. Epstein, C. Wilson and A. D. Keefe, *J. Chem. Biol.*, 2005, **12**, 25-33.
4. A. I. Taylor, V. B. Pinheiro, M. J. Smola, A. S. Morgunov, S. Peak-Chew, C. Cozens, K. M. Weeks, P. Herdewijn and P. Holliger, *Nature*, 2015, **518**, 427-430.
5. I. Alves Ferreira-Bravo, C. Cozens, P. Holliger and J. J. DeStefano, *Nucleic Acids Res.*, 2015, **43**, 9587-9599.
6. L. Torres, A. Kruger, E. Csibra, E. Gianni and V. B. Pinheiro, *Essays Biochem.*, 2016, **60**, 393-410.
7. Y. Zhang, J. L. Ptacin, E. C. Fischer, H. R. Aerni, C. E. Caffaro, K. San Jose, A. W. Feldman, C. R. Turner and F. E. Romesberg, *Nature*, 2017, **551**, 644-647.

8. A. I. Taylor, F. Beuron, S. Y. Peak-Chew, E. P. Morris, P. Herdewijn and P. Holliger, *ChemBioChem*, 2016, **17**, 1107-1110.

9. V. B. Pinheiro and P. Holliger, *Trends Biotechnol.*, 2014, **32**, 321-328.

10. V. Pezo, F. W. Liu, M. Abramov, M. Froeyen, P. Herdewijn and P. Marliere, *Angew. Chem. Int. Ed. Engl.*, 2013, **52**, 8139-8143.

11. Z. Chen, P. A. Lichtor, A. P. Berliner, J. C. Chen and D. R. Liu, *Nat. Chem.*, 2018, **10**, 420-427.

12. D. Kong, W. Yeung and R. Hili, *J. Am. Chem. Soc.*, 2017, **139**, 13977-13980.

13. R. Rossi, *Nucleic Acids Res.*, 1997, **25**, 2106-2113.

14. D. Y. Wu and R. B. Wallace, *Gene*, 1989, **76**, 245-254.

15. O. Westergaard, D. Brutlag and A. Kornberg, *J. Biol. Chem.*, 1973, **248**, 1361-1364.

16. K. Nath and J. Hurwitz, *J. Biol. Chem.*, 1974, **249**, 3680-3688.

17. K. Kleppe, J. H. van de Sande and H. G. Khorana, *Proc. Natl. Acad. Sci. U. S. A.*, 1970, **67**, 68-73.

18. M. J. Engler and C. C. Richardson, in *The Enzymes*, ed. P. D. Boyer, Academic Press, New York, 1982, vol. XV, pp. 3-29.

19. R. C. Alexander, *Nucleic Acids Res.*, 2003, **31**, 3208-3216.

20. C. M. Tsiapalis and S. A. Narang, *Biochem. Biophys. Res. Commun.*, 1970, **39**, 631-636.

21. V. Sgaramella and H. G. Khorana, *J. Mol. Biol.*, 1972, **72**, 427-444.

22. C. Goffin, V. Bailly and W. G. Verly, *Nucleic Acids Res.*, 1987, **15**, 8755-8771.

23. K. Harada and L. E. Orgel, *Nucleic Acids Res.*, 1993, **21**, 2287-2291.

24. U. Landegren, R. Kaiser, J. Sanders and L. Hood, *Science*, 1988, **241**, 1077-1080.

25. Z. Tang, *Nucleic Acids Res.*, 2003, **31**, 148e.

26. A. Cherepanove, E. Yildirim and S. de Vries, *J. Biochem.*, 2001, **129**, 61-68.

27. D. Loakes, A. V. Aerschot, D. M. Brown and D. Hill, *Nucleosides Nucleotides*, 1996, **15**, 1891-1904.

28. B. Zhao, Z. Tong, G. Zhao, R. Mu, H. Shang and Y. Guan, *Acta Biochim. Biophys. Sin*, 2014, **46**, 727-737.

29. N. Paul, G. Zon, R. Hogrefe and A. Lebedev, *J. Biomol. Tech.*, 2013, **24**, S58.

30. T. Mikita and G. P. Beardsley, *Biochemistry*, 1988, **27**, 4698-4705.

31. M. Egholm and C. Chen, WO2001027326 A3, 2001.

32. K. I. Hayashi, M. Nakazawa, Y. Ishizaki, N. Hiraoka and A. Obayashi, *Nucleic Acids Res.*, 1986, **14**, 7617-7631.

33. B. H. Pheiffer and S. B. Zimmerman, *Nucleic Acids Res.*, 1983, **11**, 7853-7871.

34. Z. X. Xiao, H. M. Cao, X. H. Luan, J. L. Zhao, D. Z. Wei and J. H. Xiao, *Mol. Biotechnol.*, 2007, **35**, 129-133.

35. K. I. Hayashi, M. Nakazawa, Y. Ishizaki and A. Obayashi, *Nucleic Acids Res.*, 1985, **13**, 3261-3271.

36. D. Louie and P. Serwer, *Nucleic Acids Res.*, 1991, **19**, 3047-3054.

37. J. R. Rusche and P. Howard-Flanders, *Nucleic Acids Res.*, 1985, **13**, 1997-2008.

38. S. B. Zimmerman and B. H. Pheiffer, *Proc. Natl. Acad. Sci. U. S. A.*, 1983, **80**, 5852-5856.

39. B. Harrison and S. B. Zimmerman, *Nucleic Acids Res.*, 1984, **12**, 8235-8251.

40. K. I. Hayashi, M. Nakazawa, Y. Ishizaki, N. Hiraoka and A. Obayashi, *Nucleic Acids Res.*, 1985, **13**, 7979-7992.

41. R. Kucera and T. C. Evans, US8697408 B2, 2014.

42. Y. Sasaki, D. Miyoshi and N. Sugimoto, *Biotechnol. J.*, 2006, **1**, 440-446.

43. R. H. Wilson, S. K. Morton, H. Deiderick, M. L. Gerth, H. A. Paul, I. Gerber, A. Patel, A. D. Ellington, S. P. Hunicke-Smith and W. M. Patrick, *Protein Eng. Des. Sel.*, 2013, **26**, 471-478.

44. C. Cozens, V. B. Pinheiro, A. Vaisman, R. Woodgate and P. Holliger, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 8067-8072.

45. L. D. Murphy and S. B. Zimmerman, *Biochim. Biophys. Acta*, 1994, **1219**, 277-284.

46. F. Leng and R. McMacken, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9139-9144.

47. P. Forterre, G. Mirambeau, C. Jaxel, M. Nadal and M. Duguet, *EMBO J.*, 1985, **4**, 2123-2128.

48. B. Harrison and S. B. Zimmerman, *Nucleic Acids Res.*, 1986, **14**, 1863-1870.

49. S. B. Zimmerman and S. O. Trach, *Nucleic Acids Res.*, 1988, **16**, 6309-6326.

50. Y. Sasaki, D. Miyoshi and N. Sugimoto, *Nucleic Acids Res.*, 2007, **35**, 4086-4093.

51. R. R. Lareu, K. S. Harve and M. Raghunath, *Biochem. Biophys. Res. Commun.*, 2007, **363**, 171-177.

52. K. Koumoto, H. Ochiai and N. Sugimoto, *Nucleic Acids Symp. Ser.*, 2008, **52**, 257-258.

53. C. S. V. Rajendrakumar, T. Suryanarayana and A. R. Reddy, *FEBS Lett.*, 1997, **410**, 201-205.

54. W. A. Rees, T. D. Yager, J. Korte and P. H. Von Hippel, *Biochemistry*, 2002, **32**, 137-144.

55. C. H. Spink, N. Garbett and J. B. Chaires, *Biophys. Chem.*, 2007, **126**, 176-185.

56. T. Caldas, N. Demont-Caulet, A. Ghazi and G. Richarme, *Microbiology*, 1999, **145**, 2543-2548.

57. F. Anjum, V. Rishi and F. Ahmad, *Biochim. Biophys. Acta*, 2000, **1476**, 75-84.

58. L. R. Singh, T. A. Dar, S. Rahman, S. Jamal and F. Ahmad, *Biochim. Biophys. Acta*, 2009, **1794**, 929-935.

59. M. M. Santoro, Y. Liu, S. M. A. Khan, L. X. Hou and D. W. Bolen, *Biochemistry*, 2002, **31**, 5278-5283.

60. S. de Kok, L. H. Stanton, T. Slaby, M. Durot, V. F. Holmes, K. G. Patel, D. Platt, E. B. Shapland, Z. Serber, J. Dean, J. D. Newman and S. S. Chandran, *ACS synthetic biology*, 2014, **3**, 97-106.

61. L. R. Singh, N. K. Poddar, T. A. Dar, R. Kumar and F. Ahmad, *Life Sci.*, 2011, **88**, 117-125.

62. A. Natalello, J. Liu, D. Ami, S. M. Doglia and A. de Marco, *Proteins*, 2009, **75**, 509-517.

63. M. K. Chow, G. L. Devlin and S. P. Bottomley, *Biol. Chem.*, 2001, **382**, 1593-1599.

64. S. Knapp, R. Ladenstein and E. A. Galinski, *Extremophiles*, 1999, **3**, 191-198.

65. Z. Liu, T. Chen and F. E. Romesberg, *Chem. Sci*, 2017, **8**, 8179-8182.

66. D. R. Bullard and R. P. Bowater, *Biochem. J.*, 2006, **398**, 135-144.

67. A. J. Doherty and T. R. Dafforn, *J. Mol. Biol.*, 2000, **269**, 43-56.

68. L. Wang and D. E. Ruffner, *Nucleic Acids Res.*, 1998, **26**, 2502-2504.

69. H. Kuai, Z. Zhao, L. Mo, H. Liu, X. Hu, T. Fu, X. Zhang and W. Tan, *J. Am. Chem. Soc.*, 2017, **139**, 9128-9131.

70. J. Li, J. Zhou, T. Liu, S. Chen, J. Li and H. Yang, *Chem. Commun.*, 2018, **54**, 896-899.