Electronic Supplementary Material (ESI) for ChemComm.

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]

[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]University 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

TABLE OF CONTENTS

MATERIALS AND METHODS

Materials

The following chemicals and biologicals were used: bidistilled glycerol 99.5 % AnalaR NORMAPUR® (VWR), deionized molecular biology grade formamide (Applichem A2156, 0100), Acrylamide-bisacrylamide 40% (19:1, Merck), Illustra NAP-25 columns (GE Healthcare), 5′ deadenylase (NEB M0331), T4 DNA ligase (NEB, M0202), SplintR ligase (NEB, M0375), T3 DNA ligase (NEB, M0317), T7 DNA ligase (NEB, M0318), *E. coli* DNA ligase (NEB, M0205), Taq DNA ligase (NEB, M0208), 9°N DNA ligase (NEB, M0238), Pfu DNA ligase (Agilent, 600191), Ampligase® DNA ligase (Epicentre, A3210K), Tsc DNA ligase (Prokazyme, Dlig119S), Rma DNA ligase (Prokazyme, Dlig120S), T4 RNA ligase 1 (NEB, M0204) and T4 RNA ligase 2 (NEB, M0239). DNA and 2′OMe RNA oligonucleotides were purchased from IDT DNA technologies. All other chemicals were purchased from Sigma Aldrich.

$MnCl_2$ was freshly prepared as a 10 mM solution. The ligation buffer ('quick ligation reaction buffer') was adapted from the Quick Ligation™ Kit from NEB and contained 66 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT and 1 mM ATP, pH 7.6 at 25°C. HNA and FANA oligonucleotides were prepared chemically from the HNA nucleoside phosphoramidites (43) and FANA nucleoside building blocks (Metkinen Chemistry) using a DNA synthesizer following the method described by Hendrix *et al.* (44), HPLC purified and verified by mass spectrometry (the MS spectra of the oligonucleotides synthesized in-house are provided in the supporting information; Figures S6-S15). Constructs for T4 DNA ligase fused to the DNA binding domains sso7d and NF-κB p50 (N-terminally) and cTF and PprA (C-terminally), were obtained as a kind gift from Dr. Wayne Patrick (Laboratory for Enzyme Engineering and Evolution, School of Biological Sciences, Victoria University, Wellington, New Zealand).

Methods

Purification of the oligonucleotides. All the oligonucleotides were purified by 15% denaturing PAGE. Subsequently, the oligonucleotides were eluted from the gel using 0.3 M sodium acetate (pH 5.4) overnight at 37°C. The sodium acetate eluates were desalted on an Illustra NAP-25 column and ethanol precipitated.

Ligation reactions. The donor, acceptor and template oligonucleotides (Table S1) were mixed to obtain a final concentration of 0.1 μM for the acceptor and 0.2 μM for the donor and template oligonucleotides (unless indicated otherwise) in DNA ligase reaction buffer 1x or quick ligation reaction buffer 1x together with the additives and the ligase, and incubated between 4 and 60°C for either 1 or 16 hours.

Analysis of the ligation reactions. The ligation reactions were mixed with an equal volume of gel loading dye (90% v/v formamide, 50 mM EDTA, 0.05% w/v orange G and bromophenol blue) and loaded onto a 15% denaturing PAGE (20x20 cm). Gels were run at a constant voltage of 400 V and analysed using a Typhoon FLA9500 imager (GE Healthcare). Gel band intensities were analysed using the ImageQuant™ TL 1D v8.1 software. For mass spectrometry analysis, the ligation mixture was heated to 95°C for 10 minutes to denature the ligase. Afterwards the ligated oligonucleotides were captured on Dynabeads™ M-280 Streptavidin (Invitrogen) following the recommended nucleic acid capturing protocol. The biotinylated template and the ligated oligonucleotide were separated through suspending the Dynabeads™ in 0.15 M NaOH. The mixture was incubated for 10 minutes at room temperature, after which the supernatant containing the non-biotinylated ligated oligonucleotide was collected and neutralized using acetic acid. This mixture was subsequently purified by phenol/chloroform extraction and ethanol precipitation. The oligonucleotide was lyophilized and desalted in an additional purification step using a Ziptip (C18, Millipore) before mass spectrometry.

Mass spectrometry analysis. Mass spectra were obtained in negative ion mode on an electrospray ionization quadrupole/time-of-flight mass spectrometer (Q-TOF 2, Micromass). Molecular weights for the oligonucleotides were obtained by deconvolution of the spectra using the MaxEnt1 algorithm

(MassLynx 3.4, Micromass). HPLC was performed on a C18 reversed-phase column (PepMap 0.5 × 15 mm, particle size 3 µm LC Packings) using a buffer containing N,N-dimethylaminobutane (DMAB, Acros) as ion pairing reagent and 1,1,1,3,3,3-hexafluoro-2-propanol (hexafluoroisopropanol, HFiP, Acros). In brief, the solvent system consisted of acetonitrile 84% in water (v/v) as organic phase and DMAB 0.05% (v/v) with HFiP 1% (v/v) in water as the aqueous phase (pH 8.0). Oligonucleotides were eluted with a flow rate of 12 µl/min applying a gradient starting at 2% organic phase and increasing by 2% per minute for 15 minutes. Alternatively, UPLC was performed on an ACQUITY UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 x 50 mm) from Waters, installed on an ACQUITY UPLC H-Class System (SQ Detector 2). A buffer containing 20 mM TEA and 400 mM HFiP in H_2O was used with a linear gradient from 18 to 31% methanol within 5 minutes and a flow rate of 0.3 ml/min for elution of the oligonucleotides. The concentration of the samples was 100 µM and 0.5 µL (50 pmol) product was injected per run.

Table S1. Overview of the oligonucleotide sequences. Acceptor oligonucleotides were FAM-labelled at the 5'-end to allow for in-gel visualisation of the ligation reactions. The donor and template oligonucleotides were modified with a C3-spacer at the 3'-end to avoid ligation at this end of the oligo. 2'OMe RNA stretches are indicated by 'm', HNA sequences are preceded by 'h', FANA is preceded by 'f' and LNA is preceded by 'l' and shown in between brackets. Nts signifies 'nucleotides'.

Name	Oligonucleotide sequence	N° nts
A1	5'-FAM-AACAGGATTAGCAG-3'	14
A2	5'-FAM-m(AACAGGAUUAGCAG)-3'	14
A3	5'-FAM-f(AACAGGATTAGCAG)-3'	14
A4	5'-FAM-AACAGGATTAGl(CAG)-3'	14
A5	5'-FAM-AACAGl(GATTAGCAG)-3'	14
A6	5'-FAM-ACTGGCTTCAGC-3'	12
A7	5'-FAM-h(ACTGGCTTCAGC) -3'	12
A8	5'-ACTGGC h(TTCAGC)-3'	12
A9	5'-FAM-ATTGCGACCATT h(GCGACC)-3'	18
A10	5'-FAM-GATTCCGATTCCGA h(TAGTGG)-3'	20
D1	5'-P-AGCGAGGTATGTAGGCGGTG-C3spacer-3'	20
D2	5'-P-m(AGC) GAGGTATGTAGGCGGTG-C3spacer-3'	20
D3	5'-P-m(AGCGAGGUAUGU) AGGCGGTG-C3spacer-3'	20
D4	5'-P-f(AGCGAGGTATGTAGGCGGTG)-C3spacer-3'	20
D5	5'-P-l(AGC)GAGGTATGTAGGCGGTG-C3spacer-3'	20
D6	5'-P-l(AGCGAGGTA)TGTAGGCGGTG-C3spacer-3'	20
D7	5'-P-AGAGCGCAGATAC-C3spacer-3'	13
D8	5'-P-h(GT)GTGGTGCA-3'	10
D9	5'-P-h(AGAGCGCAGATAC)-3'	13
D10	5'-P-h(GGTCGC)AATGGTCGCAAT-C3spacer-3'	18
D11	5'-P-h(TT)GTGGTGCA-3'	10
D12	5'-P-CTAGCGCCG h(TGC)CATGCA-3'	18
D13	5'-P-CTAGCG h(CCGTGC)CATGCA-3'	18
D14	5'-P-CTAG h(CGCCGTGC)CATGCA-3'	18
D15	5'-P-m(UGCGAGGUAUGU) AGGCGGTG-C3spacer-3'	20
D16	5'-P-m(GGCGAGGUAUGU) AGGCGGTG-C3spacer-3'	20
D17	5'-P-m(CGCGAGGUAUGU) AGGCGGTG-C3spacer-3'	20
T1	5'-CACCGCCTACATACCTCGCTCTGCTAATCCTGTT-C3spacer-3'	34
T2	5'-CACCGCCT m(ACAUACCUCGCUCUGCUAAUCCUG)TT-C3spacer-3'	34
T2-B	5'-Biotin-CACCGCCT m(ACAUACCUCGCUCUGCUAAUCCUG)TT-C3spacer-3'	34
T3	5'-f(CACCGCCTACATACCTCGCTCTGCTAATCCTGTT)-C3spacer-3'	34
T4	5'-GTATCTGCGCTCTGCTGAAGCCAGT-C3spacer-3'	25
T4-B	5'-Biotin-GTATCTGCGCTCTGCTGAAGCCAGT-C3spacer-3'	25
T5	5'-TGCACCACACGCTGAAGCCAGT-C3spacer-3'	22
T5-B	5'-Biotin-TGCACCACACGCTGAAGCCAGT-C3spacer-3'	22
T6	5'-ATTGCGACCATTGCGACCGGTCGCAATGGTCGCAAT-C3spacer-3'	36
T7	5'-ATTGCGACCATTGCGACCCCACTATCGGAATCGGAATC-C3spacer-3'	38
T8	5'-CACCGCCTACATACCTCGCGCTGCTAATCCTGTT-C3spacer-3'	34
T9	5'-CACCGCCTACATACCTCGCCCTGCTAATCCTGTT-C3spacer-3'	34
T10	5'-CACCGCCTACATACCTCGCACTGCTAATCCTGTT-C3spacer-3'	34
T11	5'-TGCACCACACGGTCGCAATGGTCGCAAT-C3spacer-3'	28
T12	5'-TGCACCACACCCACTATCGGAATCGGAATC-C3spacer-3'	30
T13	5'-GTATCTGCGCTCTGGTCGCAATGGTCGCAAT-C3spacer-3'	31
T14	5'-GTATCTGCGCTCTCCACTATCGGAATCGGAATC-C3spacer-3'	31
T15	5'-ATTGCGACCATTGCGACCCTGCTAATCCTGTT-C3spacer-3'	33
T16	5'-CACCGCCTACATACCTCGCTCCACTATCGGAATCGGAATC-C3spacer-3'	40
T17	5'-TGCATGGCACGGCGCTAGCCACTATCGGAATCGGAATC-C3spacer-3'	38
T18	5'-TGCACCACAAGCTGAAGCCAGT-C3spacer-3'	22

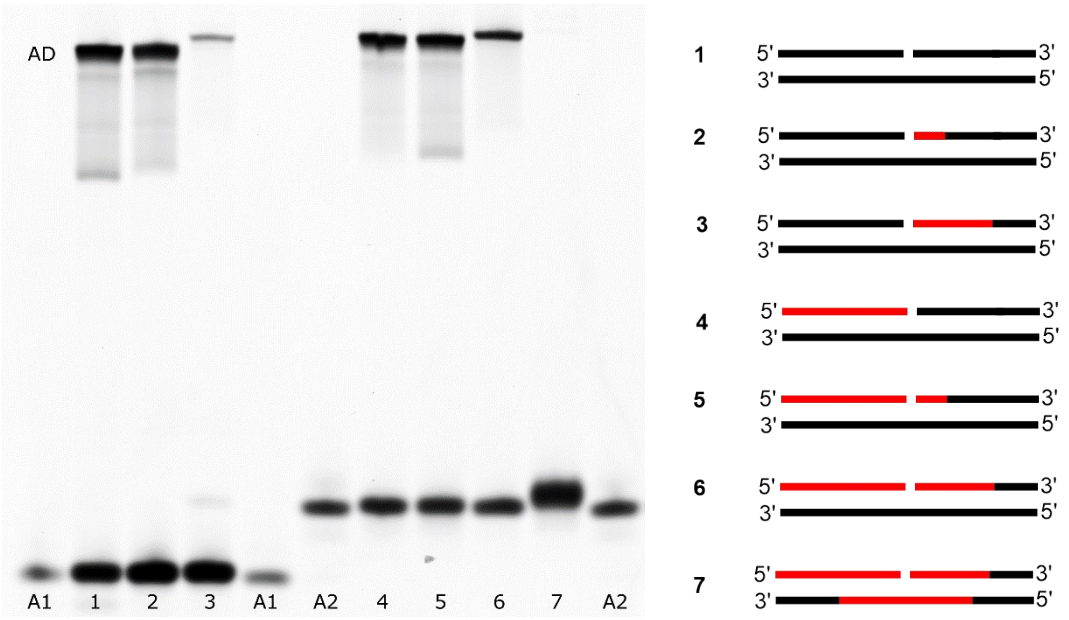
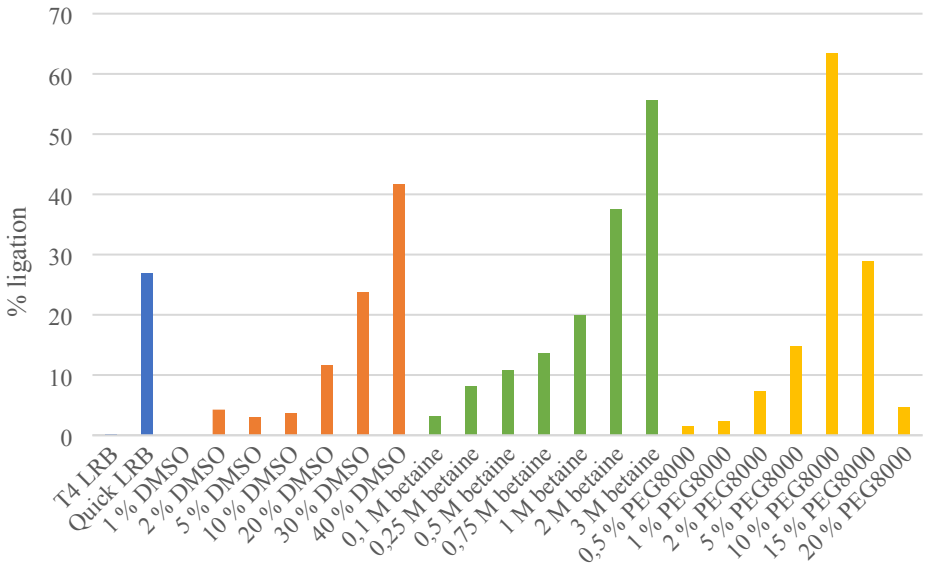


Figure S1. The ligation of 2'OMe RNA-DNA chimers using high concentrations of oligonucleotides (1 µM for the acceptor and 2 µM for the donor and the template) and enzyme (200 U/µl T4 DNA ligase), and long reaction times (16 hours) at 25°C. The primer controls are indicated by 'A1' and 'A2', respectively. On the right-hand side of the figure an illustration is given of the ligation setup. Black indicates DNA, whereas 2'OMe RNA is represented in red. The oligonucleotides used were A1 or A2 as the acceptor, and D1, D2 or D3 as the donor, opposite template T1 or T2. AD indicates the ligated product on the gel.

Table S2. The effect of different concentrations of additives on the 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). The reactions were incubated at 25 °C overnight.



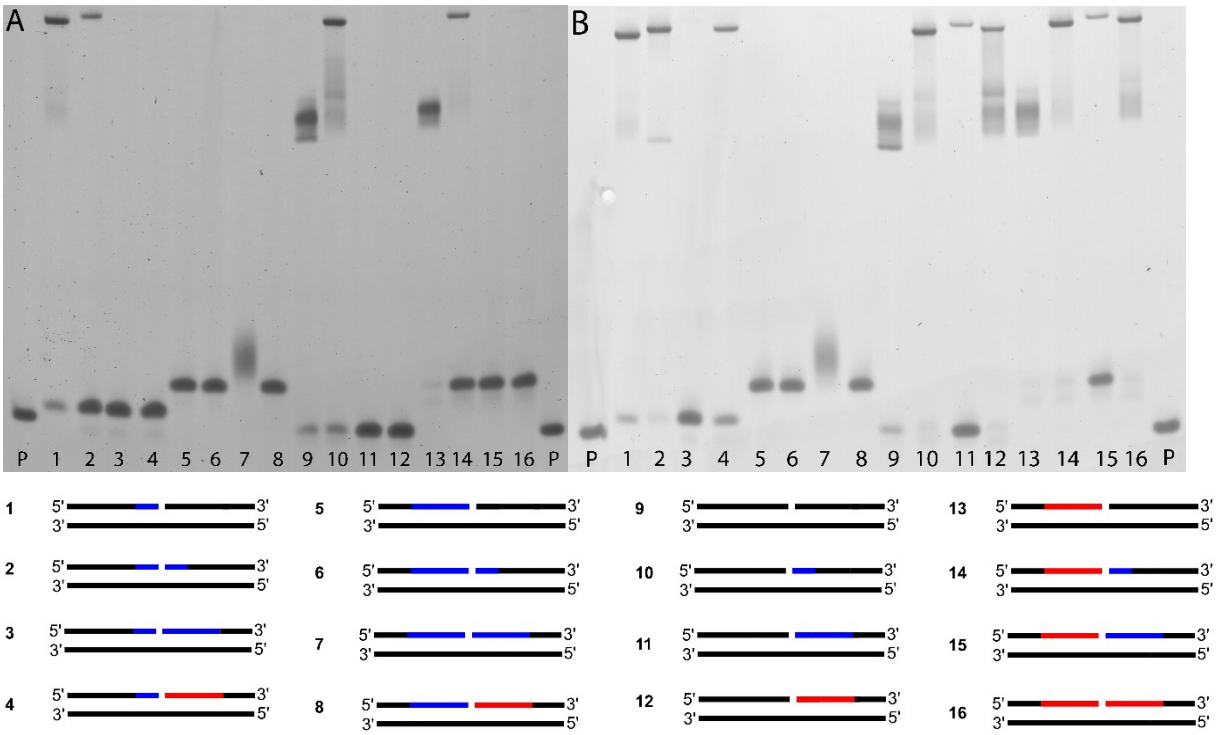


Figure S2. The ligation of LNA, 2'OMe RNA and DNA fragments in T4 LRB (A) and in quick LRB containing 10 % PEG8000 and 3 M betaine (B) in the optimized reaction conditions. The ligation reactions are shown schematically below the figure, where blue indicates the LNA fragments and red signifies the 2'OMe RNA segments. The acceptors used in the reaction were A1, A2, A4 and A5. The donors were D1, D3, D5 and D6. The template oligo that was used throughout the experiment was T1.

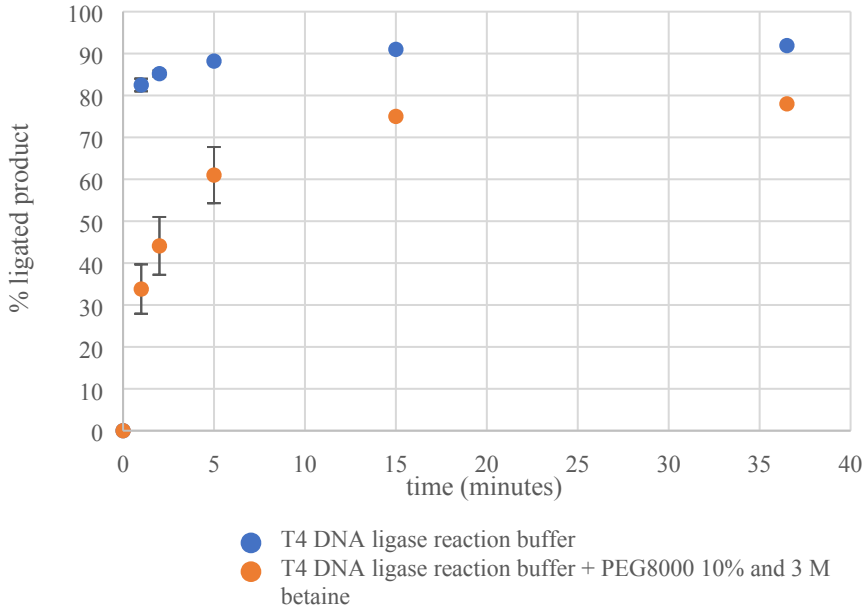


Figure S3. The time course of the ligation of a nicked DNA duplex (A1 to D1 opposite T1) at 25 °C using T4 DNA ligase (0.4 U/µl final), 0.1 µM of the acceptor oligonucleotide and 0.2 µM of the donor and template oligonucleotides in the absence or presence of 10 % PEG8000 and 3 M betaine. Reactions were carried out in triplicate. The standard deviations on the averages are indicated on the graph.

Table S3. The ligation efficiencies of different sequence context pairs in the optimized reaction conditions. Only the acceptor and the donor oligonucleotide sequences are displayed. The corresponding DNA templates from Table S1 were used in the reactions. The oligonucleotides used in the mix are given in the table. The experiments were carried out in triplicate. The standard deviations are given in the table, next to the average yield obtained for each ligation.

combi	mix	Acceptor	Donor	% yield	SD
1	A9-D8/T11	**ATTGCGACCATT h(GCGACC)**	h(GT)GTGGTGCA	90	2
	A10-D8/T12	**GATTCCGATTCCGA h(TAGTGG)**	h(GT)GTGGTGCA	89	4
2	A7-D8/T5	h(ACTGGCTTCAGC)	h(**G**T)GTGGTGCA	94	2
	A7-D11/T18	h(ACTGGCTTCAGC)	h(**T**T)GTGGTGCA	98	1
3	A9-D10/T6	**ATTGCGACCATT h(GCGACC)**	h(GGTCGC)AATGGTCGCAAT	62	10
	A10-D10/T7	**GATTCCGATTCCGA h(TAGTGG)**	h(GGTCGC)AATGGTCGCAAT	64	2
4	A9-D9/T13	**ATTGCGACCATT h(GCGACC)**	h(AGAGCGCAGATAC)	7	3
	A10-D9/T14	**GATTCCGATTCCGA h(TAGTGG)**	h(AGAGCGCAGATAC)	22	10
5	A2-D3/T1	m(AACAGGAUUAGCAG)	m(**A**GCGAGGUAUGU) AGGCGGTG	91	0.2
	A2-D15/T1	m(AACAGGAUUAGCAG)	m(**U**GCGAGGUAUGU) AGGCGGTG	88	2
	A2-D16/T1	m(AACAGGAUUAGCAG)	m(**G**GCGAGGUAUGU) AGGCGGTG	89	3
	A2-D17/T1	m(AACAGGAUUAGCAG)	m(**C**GCGAGGUAUGU) AGGCGGTG	90	0.1

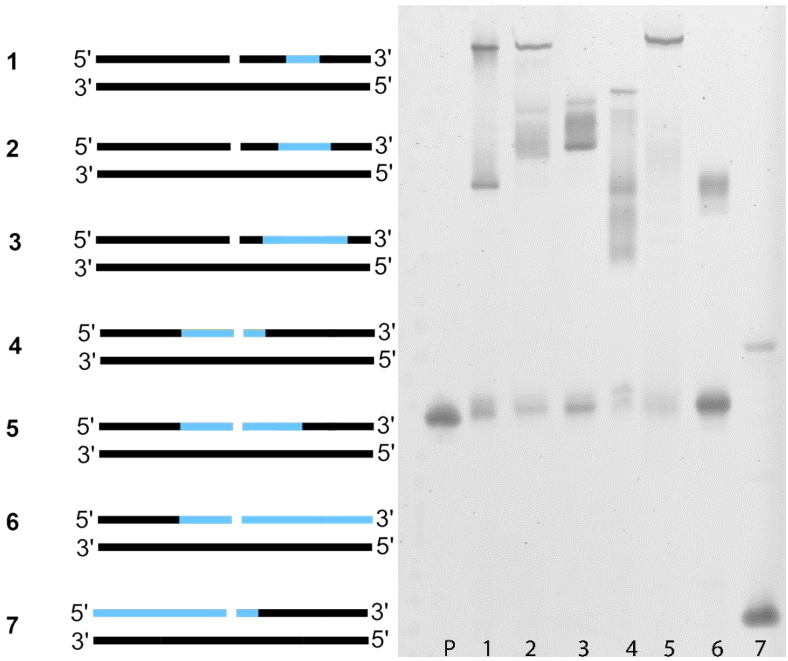


Figure S4. The ligation of a set of oligonucleotides containing HNA at different positions within the duplex in the optimized conditions. The primers control is indicated by 'P'. The oligonucleotides used in each reaction are as follows: 1) A10-D12/T17; 2) A10-D13/T17; 3) A10-D14/T17; 4) A10-D8/T12; 5) A10-D10/T7; 6) A10-D9/T14; 7) A7-D8/T5.

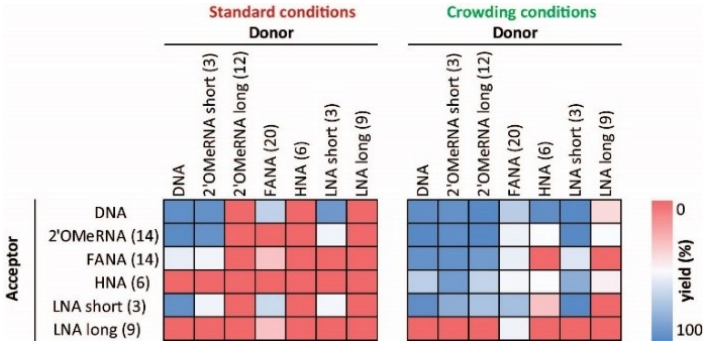


Figure S5. Overview of the XNA ligation yields in T4 DNA ligase buffer ('standard conditions') or in crowding conditions after reaction at 37°C overnight at a concentration of 0.1 µM for the acceptor and 0.2 µM for the donor and the template oligonucleotides. The acceptor oligonucleotides used were A1 (DNA) A2 (2'OMe RNA), A3 (FANA), A9 (HNA), A4 (LNA short) and A5 (LNA long). The donors were D1 (DNA), D2 (2'OMe RNA short), D3 (2'OMe RNA long), D4 (FANA), D10 (HNA), D5 (LNA short) and D6 (LNA long). The corresponding DNA sequences from Table S1 were used as the templates. The number of XNA nucleotides in each oligo is given in between brackets.

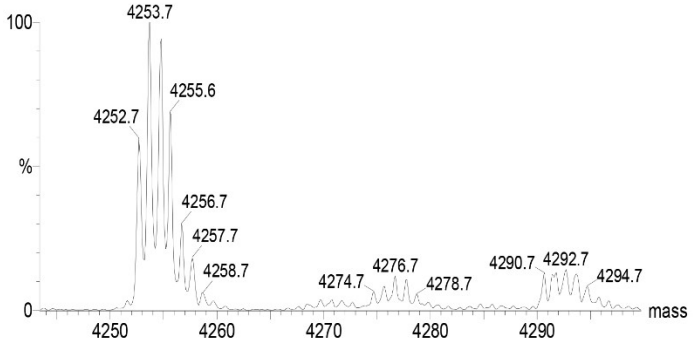


Figure S6. The deconvoluted mass spectrum of D9. Theoretical mass 4252.7

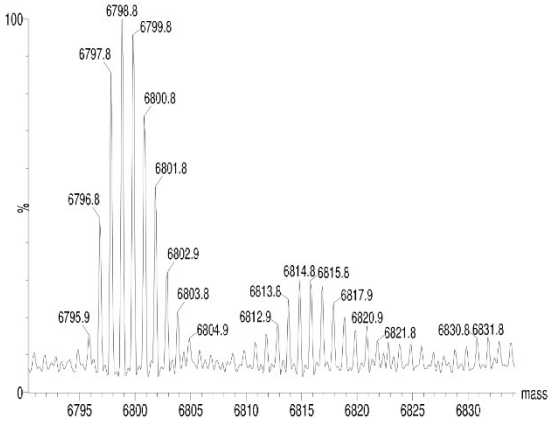


Figure S7. The deconvoluted mass spectrum of D4. Theoretical mass 6796.8

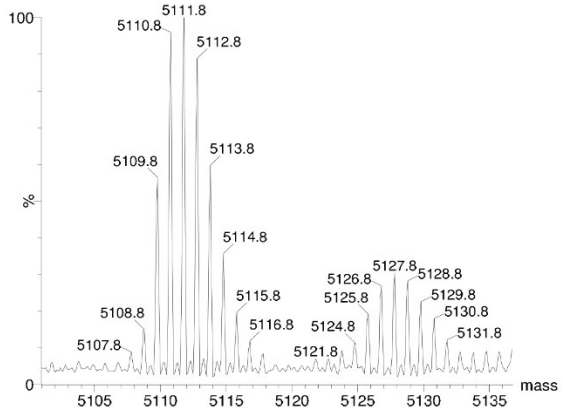


Figure S8. The deconvoluted mass spectrum of A3. Theoretical mass 5109.8

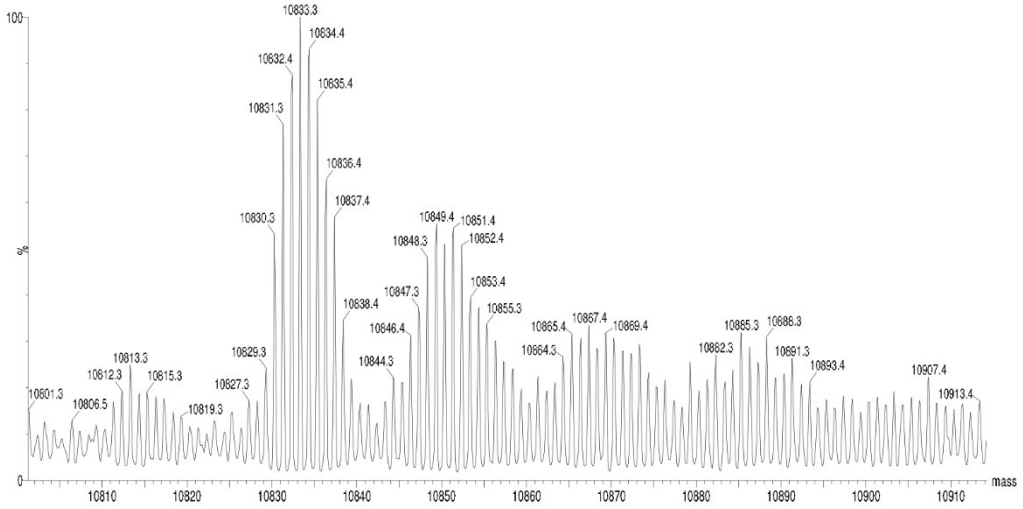


Figure S9. The deconvoluted mass spectrum of T3. Theoretical mass 10829.3

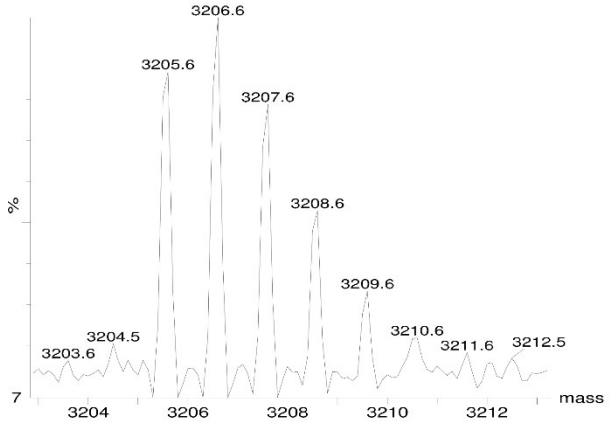


Figure S10. The deconvoluted mass spectrum of D8. Theoretical mass 3206.0

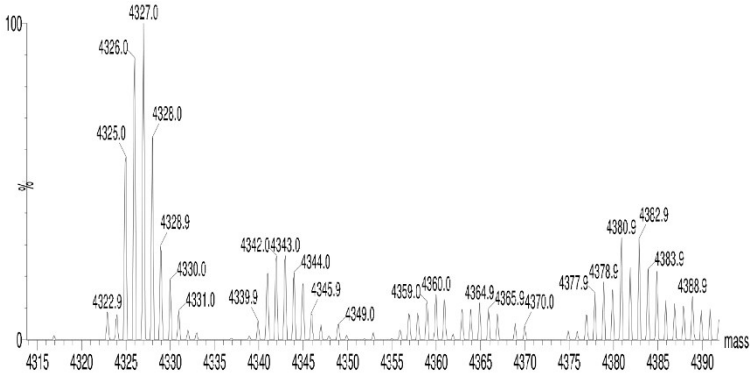


Figure S11. The deconvoluted mass spectrum of A7. Theoretical mass 4325.0

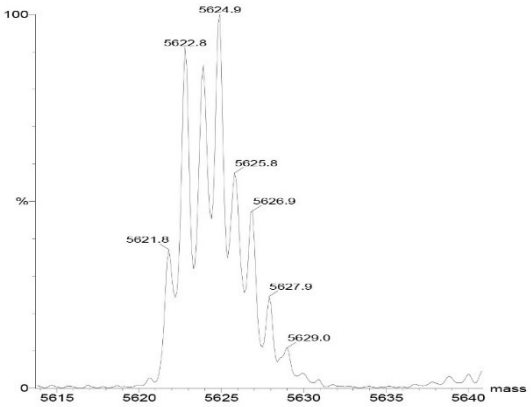


Figure S12. The deconvoluted mass spectrum of D13. Theoretical mass 5622,9

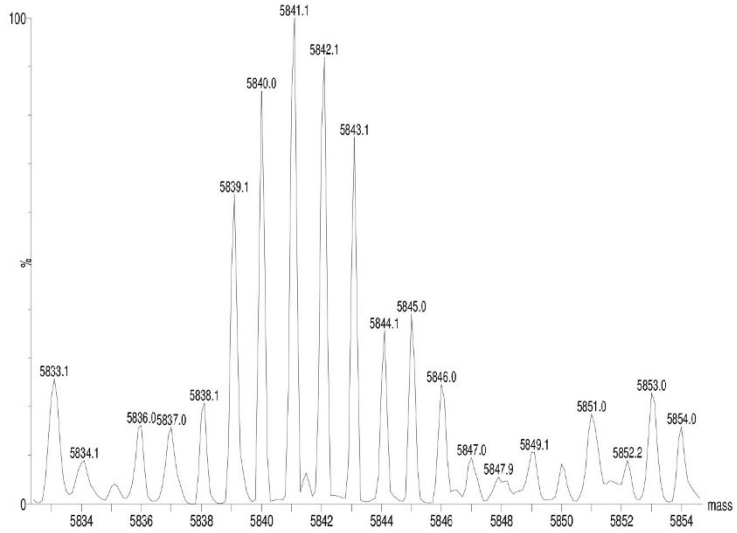


Figure S13. The deconvoluted mass spectrum of D10. Theoretical mass 5839,0

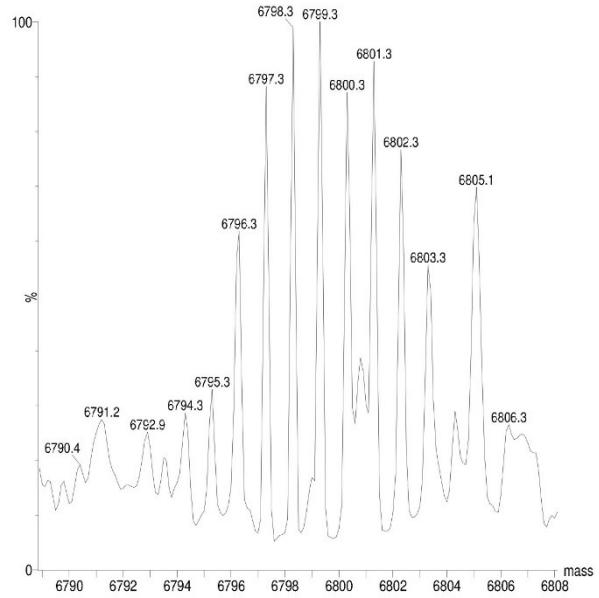


Figure S14. The deconvoluted mass spectrum of A10. Theoretical mass 6797,2

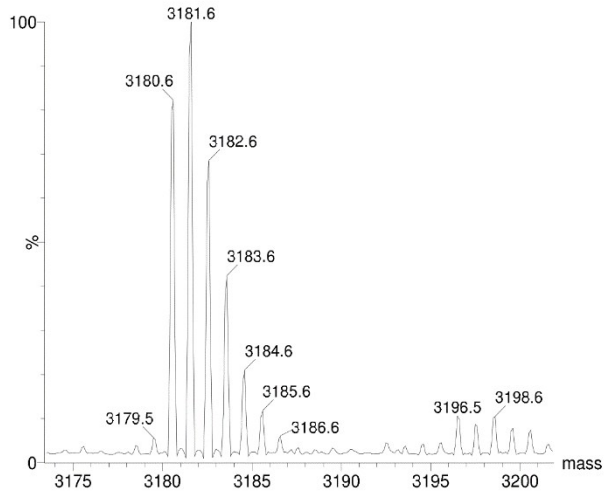


Figure S15. The deconvoluted mass spectrum of D11. Theoretical mass 3180.5