Characteristics of triplex-directed photoadduct formation by psoralen-linked oligodeoxynucleotides

Paula J. Bates, Valentine M. Macaulay1, Michael J. McLean2, Terence C. Jenkins, Anthony P. Reszka, Charles A. Laughton and Stephen Neidle*

CRC Biomolecular Structure Unit at The Institute of Cancer Research, Cotswold Road, Sutton, Surrey SM2 5NG, UK, 1CRC Centre for Cell and Molecular Biology, Chester Beatty Laboratories, The Institute of Cancer Research, Fulham Road, London SW3 6JB, UK and 2Cambridge Research Biochemicals, Gadbrook Park, Northwich, Cheshire CW9 7RA, UK

Received August 21, 1995; Revised and Accepted October 4, 1995

ABSTRACT

A triplex-forming oligopyrimidine has been attached at its 5'-end to a photoreactive psoralen derivative and used to target a sequence which forms part of the coding region of the human aromatase gene. The 20 base pair sequence is not a perfect triplex target since it contains three pyrimidine interruptions within the purine-rich strand. Despite this, we have detected triplex-directed photoadduct formation at pH 7.0 between the psoralen-linked oligonucleotide and a 30mer duplex representing the aromatase target. Photoadduct formation was found to be sensitive to pH, temperature, cation concentration and the base composition of the third strand. By varying the base sequence of the target duplex around the psoralen intercalation site, we have characterised the site and mode of psoralen intercalation. The attached psoralen has been found to intercalate at the triplex-duplex junction with a strong preference for one orientation. We have shown that the psoralen will bind at the junction even when there is a preferred TpA step at an adjacent site. We have also compared the binding affinity and photoreactivity of oligodeoxyribonucleotides linked to two different psoralen derivatives and found differences in the rate of crosslinking and the extent of crosslink formation. Finally, we have examined oligodeoxyribonucleotides which are attached to psoralen by polymethylene linkers of different lengths.

INTRODUCTION

Synthetic oligodeoxyribonucleotides (ODNs) can recognise homopurine-homopyrimidine tracts in double-stranded DNA and can bind in the major groove at these sequences to form a local triple helix (triplex). This interaction is highly sequence-specific due to the formation of hydrogen bonds between purine bases (pu) in the duplex and bases in the third strand. Several triplex motifs have been described. In one, an oligopyrimidine strand binds in a parallel orientation with respect to the purine strand of the target duplex sequence, with thymines and protonated cytosines in the third strand recognising Watson–Crick A-T and G-C base pairs respectively (1,2). In a second category, an oligopurine strand is aligned antiparallel to the purine strand of the target duplex, to form a triple helix with A-A-T and G-G-C triplets (3-5). (Throughout this paper, the third strand base will be shown in bold type). In a third motif, ODNs containing thymine and guanine bind to double-stranded DNA with the formation of T-A-T and G-G-C base triplets (6). In this case, the orientation of the third strand appears to be dependent on the number of ApG and GpA steps in the target sequence (7).

Triplex-forming ODNs have been the focus of considerable attention recently, in large part due to their potential as agents for the selective regulation of gene expression (for a review, see 8). In this, the so-called ‘antigene’ approach, ODNs are targeted to specific sequences in double-stranded DNA and have been shown to inhibit gene expression at the transcriptional level (9-12). Antigene ODNs may be targeted to regulatory regions of genes, in order to inhibit the binding of transcription factors or block transcription initiation (for example, 13-16), or to coding sequences where they might inhibit translational elongation (17-19).

One of the limitations of the triplex approach is the weak association of the third strand to the duplex target. In particular, because of the requirement for protonation of third strand cytosines, triplexes containing C+G-C triplets are generally unstable above pH 5, i.e. at physiological pH. A consequence of the relative weakness of triplex interactions is that any biological effects achieved using unmodified ODNs are likely to be transient. The attachment of a reactive intercalating agent to the end of the ODN addresses these problems by markedly increasing the stability of the triplex (20) and allowing covalent modification of the target DNA.

Psoralen-linked ODNs have been widely used as antigene agents (11,17,18,21-23) and to introduce site-specific base mutations (24-27). Psoralens are photoreactive molecules which intercalate into double-stranded DNA where they can form covalent bonds to pyrimidine bases (py), primarily thymines (for a review, see 28). They are bifunctional and have reactive double bonds at the 3,4 (pyrone-side) and 4',5' (furan-side) positions (Fig. 1). On absorption of one UV photon of wavelength >310 nm, a monoadduct (MA) is formed at one of these bonds. If the

* To whom correspondence should be addressed
psoralen is located at an appropriate site, e.g. TpA or ApT, the furan-side MA can, upon absorption of another photon, react at the pyrone-side to produce a bis adduct (XL), thereby crosslinking the target DNA. The pyrone-side MA lacks strong absorption bands at $\lambda > 300$ nm and there is no evidence that it can be converted to XL by further irradiation. Psoralen derivatives which have a methyl group at the 4 position are reported to form <2% pyrone-side MA (29). This is attributed to steric interference between the 4-methyl group of the psoralen and the 5-methyl group of thymine (thymine reacts at the 5,6 double bond). Two possible orientations for a ODN-linked psoralen intercalated in a double helix at a TpA step, which we have called mode I and mode II, are shown in Figure 1.

Aromatase is a cytochrome P450 enzyme which converts androgens to oestrogens and is an important target in the treatment of hormone-dependent breast cancer. We have been examining psoralen-oligonucleotide conjugates in order to target a 20 base pair (bp) sequence (Fig. 2) within the coding region of the human aromatase gene. The third-strand ODN has been attached at the 4' position of the psoralen derivative rather than the more frequently studied 5 position and we have used 4'-hydroxypsoralen (APG-20) which is more reactive and gives a greater yield of crosslinked adducts than psoralen itself (28). The chosen DNA sequence is not an ideal triplex target since it contains three cytosine bases within the purine-rich region. Despite this, we have shown (17) that a psoralen-ODN is able to form a triplex-directed crosslinked adduct in vitro at pH 7 with a 30mer unmodified duplex containing the 20 bp target. We have demonstrated previously (17) that in vivo crosslinking of this ODN to its target within an aromatase expression construct, followed by transfection into COS or MCF-7 human breast cancer cells, leads to a reduction in the level of full-length aromatase mRNA transcripts and reduced aromatase enzyme activity. In this report, we present in vitro studies to complement these biological results. We report here on the effects of pH, temperature, base mismatches and added cations on triplex-directed photoadduct formation in this system. In addition, the site and mode of the triplex-directed intercalation of the psoralen has been characterised. The influence of the length of the hydrocarbon linker joining the psoralen and ODN, and the nature of the psoralen derivative used, have also been investigated.

**MATERIALS AND METHODS**

**Oligonucleotide synthesis**

All oligonucleotides were prepared on an Applied Biosystems 391-EP automated DNA synthesiser using standard phosphoramidite chemistry. Oligonucleotides were purified by ion-exchange HPLC and desalted by passage through a Sephadex column. ODNs linked to 4'-(hydroxymethyl)-4,5',8-trimethylpsoralen via a pentamethylene linker were prepared using an appropriately synthesised phosphoramidite and were purified as above. The presence of psoralen was confirmed subsequent to purification by fluorescence spectroscopy. The 20mer linked to a 5-hydroxypsoralen derivative (APG-20) was obtained from Appligene (Chester-le-Street, Co. Durham, UK). Pso-C2-20 and Pso-C6-20 were prepared using psoralen phosphoramidites from Glen Research (Sterling, VA). The sequences of the ODNs and psoralen-ODN conjugates used are shown in Figure 2.

**Detection of photoadducts and effect of pH, temperature and base composition**

The 30mer purine-rich strand, Arom-30U, was 5' end-labelled using [$\gamma^{32}$P]ATP and T4 polynucleotide kinase (New England Biolabs, Inc.) and separated from unincorporated ATP by
Figure 3. (A) Denaturing gel electrophoretic mobility shift assay for photoadduct formation at pH 5, 6 and 7, at 22°C in the presence of 1 mM spermine. XL marks the position of the slower moving crosslinked adduct; SS represents the faster moving single strand (denatured duplex). The concentrations of Pso-C5-20 are 0.4 μM (lanes 1, 5 and 9), 4 μM (lanes 2, 6 and 10), 40 μM (lanes 3, 7 and 11) and 400 μM (lanes 4, 8 and 12). SS and D are, respectively, single-stranded and duplex markers. (B) Densitometric analysis of the same gel. Photoadduct formation is measured by the integrated intensity of the band corresponding to crosslinked adduct divided by the total integrated intensity (adduct + denatured duplex).

Figure 4. (A) Electrophoretic mobility shift assay for photoadduct formation at various temperatures. Incubation was carried out in a TAM buffer (pH 7.0) in the presence of 1 mM spermine. Concentrations of Pso-C5-20 are 0.1 μM (lanes 1, 4, 7 and 10), 1 μM (lanes 2, 5, 8 and 11) and 10 μM (lanes 3, 6, 9 and 12). (B) Densitometric analysis.

electrophoresis through a 20% acrylamide gel. Formation of the duplex (Arom-30) was achieved by annealing this with a 5-fold excess of the unlabelled Arom-30Y in a TAM buffer (40 mM Tris-acetate, 5 mM MgCl₂ and 10% w/v sucrose, adjusted to the appropriate pH by addition of acetic acid). Duplex formation was confirmed by non-denaturing polyacrylamide gel electrophoresis (PAGE). The psoralen-linked oligonucleotide (Pso-C5-20, Pso-Con or Pso-C5-20B for the base composition experiments) was added to the radiolabelled duplex (10 000 c.p.m., final concentration ~0.5 nM, per reaction) and the reaction mixture was incubated in the presence of 1 mM spermine tetrachloride at 0–37°C in the same TAM buffer.

Following incubation for 60 min, the reaction mixture was irradiated with 366 nm light for 10 min in an open-topped Eppendorf tube at 0°C, at a distance of ~5 cm from the light source. For convenience, the light source used in these experiments was a standard dual-wavelength laboratory UV lamp (Camag) set at 366 nm, fitted with a single 8 W mercury lamp, and giving an intensity of ~3.8 mW/cm² at a distance of 5 cm. Previous experiments (17) used a more intense light source, an inverted long wavelength transilluminator (UVP Ltd) fitted with 6 × 15 W lamps. The lower intensity source was used for all experiments described in this paper, except the base composition and spermine dependence assays. Following irradiation, samples were denatured by heating at 95°C for 5 min and separated by electrophoresis through denaturing 16% polyacrylamide gels containing 7 M urea and 89 mM Tris borate/2 mM EDTA buffer, pH 8.3. Electrophoresis was carried out at 400–450 V for 2–3 h at room temperature.

Effect of added cations

Radiolabelled duplex (10 000 c.p.m. per reaction) was mixed with Pso-C5-20 (final concentration 200 mM or 10 mM) in a TAM pH 7.0 buffer, and spermine tetrachloride, MgCl₂ or NaCl were added to give final concentrations in the range of 0.01–10 mM (spermine³⁺) or 10 μM–1.0 M (Mg²⁺ or Na⁺), in the presence or absence of 1 mM spermine). Samples were then incubated, irradiated and electrophoresed as above.

Site and mode of psoralen intercalation

Six oligonucleotides were synthesised in order to produce 30mer duplexes (named D2, D3, D4; see Fig. 7B) which contained the
RESULTS

Preliminary assays carried out under non-denaturing conditions showed that an unmodified ODN, 20T, could retard the mobility of radiolabelled Arom-30 duplex at pH 5 and 15°C, but not at higher temperature or pH (not shown). With the aim of increasing triplex stability, the 20T oligonucleotide was linked to 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen, to give Pso-C5-20. We have shown previously (17) that Pso-C5-20, forms photoadducts when incubated with radiolabelled Arom-30 duplex at pH 7 and irradiated with near-UV light. To exclude non-specific crosslinking, a control ODN, Pso-Con, was synthesised with a 5′-psoralen modification and the same number of cytosines and thymines as Pso-C5-20, but with an unrelated sequence. No photoadducts were detected when Arom-30 was incubated with ≤400 μM Pso-Con and irradiated as described (not shown).

Effect of pH, temperature and base composition

Triplex-directed photoadduct formation was sensitive to the pH and temperature of incubation. Figures 3 and 4 show the variation in photoadduct formation with the pH and temperature of incubation. The extent of adduct formation decreases as the pH and temperature increase, but significant crosslink formation is detectable at pH 7.0/37°C.

The triplex-forming ODN, Pso-C5-20, was designed so that T-CG triplets would be formed at the three CG mismatch sites in the duplex. We have compared the binding of third-strand psoralen-linked ODNs which could form triplexes containing a single B-C-G (B = T, C, A or G) in order to investigate the effects of different non-canonical triplets. The stability of the triplets, as assessed by the extent of photoadduct formation was found to be T-C-G ~ C-C-G > G-C-G > A-C-G (Fig. 5). The presence of a pu-C-G triplet was highly destabilising and photoadduct formation was almost abolished when such a triplet was introduced.

Effect of added cations

Triplex-directed photoadduct formation was measured as a function of spermine4+, Mg2+ and Na+ concentration and was found to be strongly dependent on the presence of spermine. The results indicated that increasing spermine concentration >1 mM led to no further enhancement of triplex stability in this system.
Figure 7. (A) Denaturing gel showing photoadduct formation after 1 (lanes 1–8) or 15 min (lanes 11–18) irradiation with duplexes Arom-30, D2, D3 and D4, which were radiolabelled on either the purine-rich (odd numbered lanes) or pyrimidine-rich (even numbered lanes) strand. (B) Sequences of Arom-30, D2, D3 and D4 showing possible psoralen intercalation sites. *T indicates the psoralen-linked 5' base of the third strand ODN. (C) Summary of results.

Site and mode of psoralen intercalation

Samples were irradiated under conditions which gave predominantly monoadduct formation (1 min irradiation) or which could convert furan-side monoadduct to crosslinked bis-adduct if at an appropriate site (15 min irradiation). The results using Pso-C5-20 as third strand are shown in Figure 7.

Molecular modelling (Laughton, C.A., unpublished) suggested that when Pso-C5-20 binds to the Arom-30 duplex the psoralen could intercalate either at the triplex-duplex junction (a 5'-TpA step) or at the adjacent 5'-ApT step of the duplex. The most likely conclusion from the experimental data is that the psoralen is intercalated at the triplex-duplex junction. When this consisted of a T-A-T triplet adjacent to a A-T base pair as in duplex D2, the psoralen formed a monoadduct at the TpT step (a non-crosslinkable site) even when there was a preferred TpA crosslinkable site adjacent. Since a photoadduct was observed when Pso-C5-20T was incubated with duplex D3, but not with D4, it is likely that the monoadduct was formed with the T involved in the duplex side of the junction.

After 1 min irradiation, monoadduct formation was observed with Arom-30 only when the purine-rich strand of the duplex (Arom-30U) was labelled. Since this is likely to be a furan-side monoadduct, it would appear that mode I intercalation was favoured. When the psoralen was intercalated at a TpT step (duplexes D2 and D3) it was unclear whether the monoadducts formed were furan-side adducts (requiring mode II intercalation) or pyrone-side adducts (mode I intercalation). This question could be addressed by treating the monoadducts with hot piperidine (21,30) which is reported to lead to cleavage of furan-side but not pyrone-side monoadducts.

Comparison of psoralen derivatives

APG-20 and Pso-C5-20 were found to have very similar binding affinities, as measured by the extent of photoadduct (mono + bis) formation, with the target (not shown). However the photoreaction of Pso-C5-20 was found to be significantly faster than that of the APG-20 as shown in Figure 8. After 10 min irradiation, the Pso-20 photoadduct was almost 100% in the bis form, whereas the APG20 adduct was ~50% mono.

Comparison of linker lengths

Pso-C2-20 and Pso-C6-20 showed the same pattern of adduct formation with duplexes D1-4 as did Pso-C5-20, suggesting a similar site and orientation of the psoralen (not shown). The comparison of photoadduct formation as a function of irradiation time (Fig. 9), indicates that Pso-C2-20 formed crosslinks more slowly than Pso-C5-20.
DISCUSSION

The human aromatase cDNA sequence lacks long stretches of perfectly contiguous purines. The 20 base sequence chosen, which contains eleven adenines, six guanines and three cytosines, represents the closest approach to a perfect triplex target having sufficient length to be potentially unique, at least on a statistically random basis. A search for its occurrence in the Human Genome Databank has not revealed any correspondences, even taking into account all possible bases at the three mismatch positions.

This study has characterised a number of features of in vitro triplex formation between psoralen-linked pyrimidine-rich oligonucleotides and a double-stranded target. The 30mer duplex used here represents bases 821-850 of the aromatase cDNA (31) and incorporates the 20mer target flanked by five bases at each end. Triplex-forming third-strand ODNs were designed by placing thymines to hydrogen bond to AT pairs and cytosines to G-C pairs in the target. In an attempt to minimise the potential destabilising effects of 'mismatched' triplets due to the three pyrimidine interruptions on the purine target strand of the duplex, we placed thymines opposite the C-G pairs, since isolated T-C-G triplets have been reported to be relatively stable in triplexes with a pyrimidine third strand (32). The data presented here are in accord with these results; we also show that more than one such mismatch can be tolerated within a py-pu-py triplex without undue destabilisation, whereas pu-C-G triplets strongly impair triplex formation. The observed order for stability of all four B-C-G triplets is broadly in agreement with previous reports (32,33). The present results demonstrate that conjugation of a triplex-forming ODN to psoralen does not appear to compromise the specificity of triplex recognition.

Attaching psoralen to the triplex-forming ODN permits cross-linking to the target and thereby 'fixes' the triplex, allowing easy detection by a denaturing electrophoretic mobility shift assay. In contrast to unmodified or acridine-linked third strands (unpublished observations), triplex formation at the Arom-30 target can be detected at pH 7 and 37°C by crosslinking with a psoralen-linked ODN, Pso-C5-20. It is difficult to quantitatively compare data between different experiments reported here, as there may be variations in factors such as distance of the sample from the light source during photoactivation, and differences in the densitometric analysis which depend on the exposure of the autoradiogram. However, within a single experiment the data are internally consistent and are thus comparable, since the samples have been prepared and irradiated in parallel. Further, the trends seen in the experiments concerning the dependence of adduct formation on pH, cation concentration, etc. are reproducible. Under the conditions used in most of these experiments, 60–70% of molecules in the triplex form are converted to crosslinked adducts. Using a more intense light source (UVP transilluminator), we were able to achieve >90% photodadduct formation (17 and unpublished observations).
Several aspects of triplet formation studied here have previously been investigated using unmodified triplet-forming ODNs (for examples, see 34–36). Like py-py-py triplet formation between unmodified ODNs, photoadduct formation by psoralen-linked oligopyrimidines is dependent on pH, temperature and the presence of spermine. These results confirm that adduct formation is directed by triplet recognition of the target. The findings concerning the identity and concentrations of added cations are qualitatively similar to other reports (37). The apparent destabilisation of triplet by high concentrations of Na⁺ or Mg²⁺ in the presence of spermine is presumably due to their successful competition with the spermine, which is a stronger stabiliser of triplet helix. In addition, sodium and magnesium ions are also reported to have an inhibitory effect on the dark binding and photoreaction of free psoralens (38). The effect of the various cations on photoadduct formation should therefore be taken into account when the psoralen-oligonucleotides are to be used in cells.

Two previous studies have investigated triplet-directed photoadduct formation by a homopyrimidine ODN linked to 5-hydroxy-psoralen (21) or a homopurine ODN linked to 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (24). Both groups find that the tethered psoralen intercalates at the triplet-duplex junction and that there is a strong preference for the orientation which places the furan-side of the psoralen close to the purine-rich strand of the target duplex. Our results, using an oligopyrimidine linked to 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen, are in agreement with these conclusions. Hélène and co-workers reported that photoaddition of a 5-hydroxypsoralen-ODN (21) produced 85% furan-side monoadduct which could be converted to bis-adduct and 15% of non-crosslinkable pyrone-side monoadduct, and this was confirmed by our own observations of APG-20. In contrast, we found that initial photoaddition by our 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen-linked ODN occurred almost exclusively at the furan-side, leading to a higher yield of bis compared to mono addition.

Comparison of adduct formation by triplet-forming ODNs linked to either 5-hydroxypsoralen (APG-20) or 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (Psy-C5-20), and irradiated at 366 nm, showed that the photoreaction of the former was relatively slow. These results are unsurprising in the light of what is known about the reactions of free psoralen derivatives (28), but since DNA crosslinks are expected to be a more efficient block to transcription or replication than monoadduct formation, these may be important considerations when designing antigenic psoralen-oligonucleotide conjugates.

The therapeutic use of psoralen-linked oligonucleotides may ultimately be limited due to the difficulty of UV activation of the psoralen moiety in an in vivo situation, as well as the possibility of efficient repair of triplet-directed psoralen adducts (18,25). However, they have already proved to be extremely useful tools in molecular biology, and may have potential clinical use in the treatment of superficial skin malignancies or ex vivo marrow purging.

ACKNOWLEDGEMENTS

This work was supported by the Cancer Research Campaign (grant SP1384 to S.N.) and by an Institute of Cancer Research Studentship (to P.J.B.).

REFERENCES