Pergamon

Tetrahedron, Vol. 52, No. 6, pp. 2229-2234, 1996

0040-4020/96 $15.00 + 0.00

0040-4020(95)01053-X

Synthesis of a Peptide-Intercalator Hybrid Based on the bZIP Motif From GCN4

Alethea B. Tabor

University College London, Chemistry Department, Christopher Ingold Laboratories, 20, Gordon Street, London WC1H 0AJ. UK.

Abstract: An artificial peptide, designed to combine the DNA-recognition portion of the bZIP motif from the yeast transcription factor GCN4 with an intercalating portion, 9-aminoacridine, has been synthesised.

INTRODUCTION

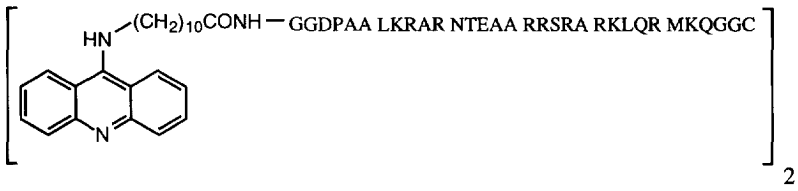
The sequence-specific recognition of double-stranded DNA has been the subject of extensive studies. The design of ligands specific to a given DNA sequence is an important area of research, as such ligands have the potential to control transcription[1] and replication.[2] Research to date has mainly focused on the recognition and modification of double-stranded DNA by oligonucleotides containing both natural and synthetic base pairs; these bind into the major groove of DNA by forming Hoogsteen hydrogen bonds, resulting in a triple helix structure.[3] The binding of these oligonucleotides has been further enhanced by the attachment of intercalating agents.[4,5] A limitation of this approach is that such oligonucleotides are susceptible to attack by nucleases.

A complementary approach to this problem would be to use peptide or peptide mimetics as sequence-specific DNA ligands. Until recently, it was impossible to design peptides which recognise double-stranded DNA sequences with the degree of specificity obtainable with oligonucleotides. However, extensive work on the DNA-binding motif of the yeast transcription factor GCN4[6] has begun to yield smaller peptides which can be tailored to recognise a range of DNA sequences.

The DNA-binding domain of GCN4 is a member of the bZIP family of DNA-recognition motifs. It is about 60 amino acids long, divided up into two regions of equal size, and binds to DNA as a dimer. The C-terminal region enables the bZIP motif to dimerise, via formation of a parallel coiled-coil structure known as the leucine zipper, and the N-terminal region, which contains a large number of basic residues, contacts the DNA directly.[7] Random mutagenesis experiments have been used to alter the basic region of the bZIP motif from GCN4 to recognise different sequences of double-stranded DNA.[8,9] Synthetic studies have also been carried out on this peptide motif; these established that the leucine zipper is not necessary for dimerisation, and can be replaced with a cysteine disulfide linker[10] or a transition metal complex.[11] Although the binding of these truncated peptides to the consensus DNA sequence recognised by GCN4 was highly selective, it was also weaker and binding was not detected above 0 °C.

In this paper, the synthesis of a GCN4 basic region-intercalator hybrid peptide (**1**) is described. This molecule has been designed to combine the minimum features of the peptide necessary for recognition of the consensus site[10] with the known binding enhancement of an intercalating acridine moiety.[4,5] In this way, it is

hoped to achieve both tight binding and a high degree of molecular recognition with a comparatively small peptide.

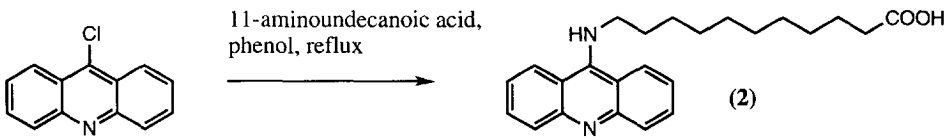


(1)

RESULTS AND DISCUSSION

The hybrid peptide (**1**) consists of the N-terminal basic region of GCN4, with a Gly-Gly-Cys motif added at the C-terminus to enable the motif to be dimerised.[10] At the N-terminus the intercalator, 9-aminoacridine, is attached via an 11-aminoundecanoic acid linker and two further glycine residues. Molecular modelling studies indicated that this would make the linker flexible and long enough to allow intercalation to occur between the second and third base pairs after the binding site; closer proximity of the intercalator to the binding site might lead to reduced sequence specificity, as intercalating groups are known to introduce local distortions into the structure of double-stranded DNA.[12]

The linker-acridine unit (**2**) was readily synthesised from 9-chloroacridine[13] and 11-aminoundecanoic acid via a modification of Hélène's method.[14]

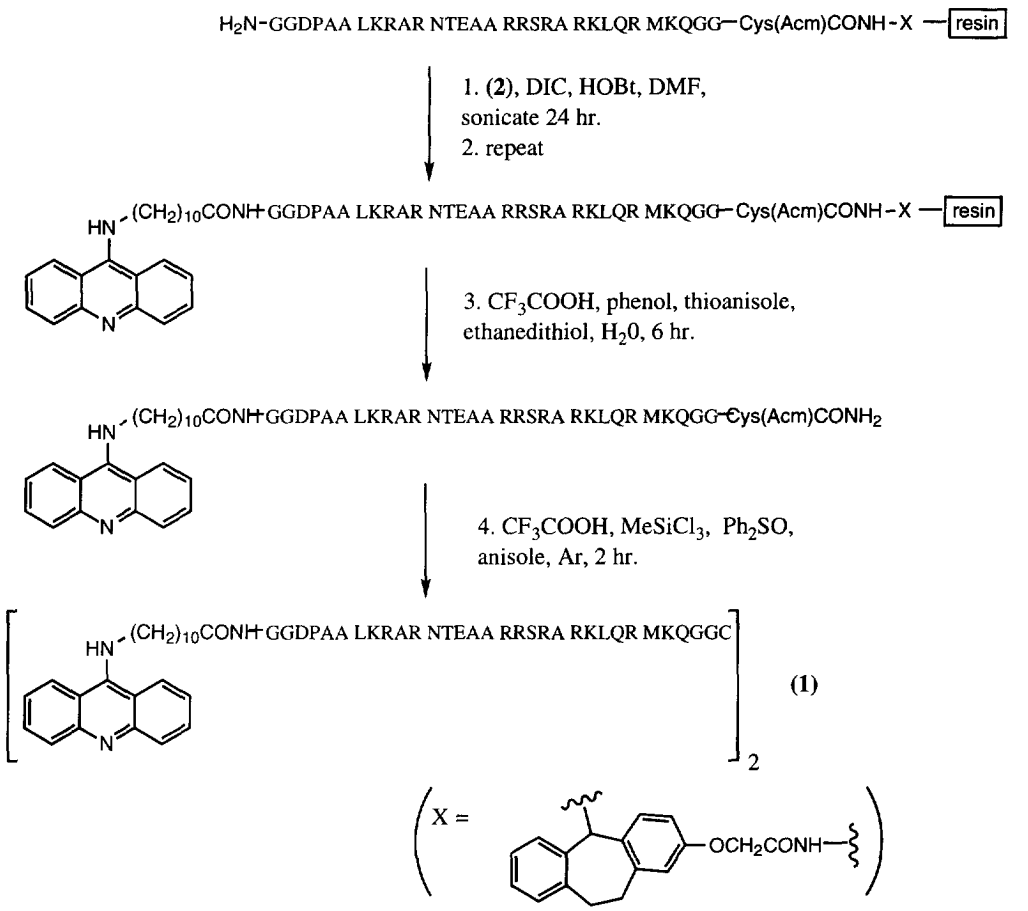


(2)

Preliminary investigations with model peptides suggested that it would be possible to couple (**2**) directly to the end of the peptide chain using standard coupling conditions (DIC, HOBt, DMF)[15] although sonication and prolonged reaction times were necessary, as the solubility of (**2**) in DMF was poor. The peptide component of (**1**) was therefore synthesised using FMOC methodology on an Applied Biosystems 430A peptide synthesiser, using a modification of the Rink resin recently reported[16] (Scheme 1). The side-chains were protected as the t-butyl ester or t-butyloxycarbonyl derivatives, where necessary, except for the Cys sulfhydryl which was protected as the Acm derivative. After the deprotection of the final FMOC group, the resin was removed from the synthesiser and double-coupled to (**2**). The peptide was then cleaved from the resin using standard conditions and purified.

Deprotection of the Acm group using silver triflate[17] proceeded smoothly, but dimerisation of the peptide to give (**1**) proved to be problematic. Aerial oxidation at pH 8, in a variety of buffers, was unsuccessful, as the peptide is insoluble in aqueous solvents at this pH. Oxidation in neat DMSO[18] proceeded very slowly. Removal of the Acm group using I_2, accompanied by *in situ* oxidation, was then attempted;[19] short reaction times resulted in the recovery of the protected starting material, and longer reaction times resulted in degradation of the sample. Finally, simultaneous deprotection and oxidation was achieved using diphenyl sulfoxide-methyl trichlorosilane.[20]

In summary, the synthesis of a novel peptide-intercalator hybrid has been described. The route is short and high-yielding, and can be easily adapted to incorporate other linker/intercalator combinations.[21] Biological tests are in progress, and the results will be reported elsewhere.



Scheme 1

EXPERIMENTAL

General Procedures.

IR spectra were determined with a Perkin-Elmer 1600 FTIR spectrometer. UV spectra were obtained with a Shimadzu UV-160A spectrometer. ^{1}H NMR (400 MHz) spectra were recorded using a Varian VXR400, and ^{13}C NMR (90.5 MHz) spectra using a Brucker WH360 spectrometer. Chemical shifts (δ) are given in ppm. Column chromatography was performed using silica gel (40-63 μm, BDH) and preparative thin layer chromatography was performed using silica gel 60 F_{254} coated on aluminium plates (Merck). FAB mass spectra were determined using a Kratos MS50.

The modified Rink resin (0.24 mmol/g loading) was synthesised as previously described by Ramage et al.[16] Fmoc amino acids were purchased from Bachem. The following amino acid side-chain protecting groups were used: Arg(Pmc); Asp(OBu^t); Cys(Acm); Lys (Boc); Ser(Bu^t); Thr(Bu^t). Peptide synthesis solvents DMF, dioxane and piperidine were purchased from Rathburn, and were of peptide synthesis grade. Diisopropylcarbodiimide (DIC), acetic anhydride and pyridine were purchased from Aldrich and used as supplied. *N,N*-Dimethylaminopyridine (DMAP) and *N*-hydroxybenzotriazole (HOBt) were purchased from Fluka. Analytical HPLC was carried out using ABI equipment on a Vydac C_{18} (5 μm, 218TP54) 4.6 x 250 mm column at 1 ml/min, monitoring at 254 nm. Preparative HPLC was carried out using ABI or Gilson equipment, on a Vydac C_{18} (10 μm, 218TP1022) 22 x 250 mm column at 18 ml/min, monitoring at 254 nm. Solvents used were of HPLC grade; A: H_2O (0.1% TFA); B: 95% MeCN, 5% H_2O (0.1% TFA). Amino acid analysis was carried out using a Pharmacia Alpha Plus system. Linear MALDI TOF mass spectra were determined using a Voyager Biospectrometry instrument (PerSeptive Biosystems) (Department of Chemistry, University of Edinburgh) and a VG TofSpec instrument (London School of Pharmacy). Electrospray spectra were acquired using a VG Platform (NIMR, Mill Hill); 10 μl of the peptide solution recovered from the HPLC experiment was injected via an Applied Biosystems syringe pump running 50% MeCN at 15 μl/min into the mass spectrometer. The spectra were acquired over a mass range of 450 - 800 m/z at 5 s/scan in continuum mode, with 10 scans averaged. The data were calibrated using polyethylene glycol.

Synthesis of N-(9-acridinyl)-11-aminoundecanoic acid (2)

9-Chloroacridine[13] (206 mg, 0.96 mmol), 11-aminoundecanoic acid (391 mg, 1.9 mmol, 2 eq.) and phenol (500 mg) were heated at 100 °C for 4 h. The bulk of the phenol was then removed *in vacuo* to give a black tar. This was purified by flash chromatography (20% MeOH in $CHCl_3$) and then recrystallised from isopropanol/water to give (**2**) as a yellow powder (297 mg, 81%), m. pt. 148 °C (dec.). R_f 0.14 (20% MeOH in $CHCl_3$); ^{1}H NMR (400 MHz, CD_3OD) 1.30-1.41 (10H, m), 1.47-1.50 (2H, m), 1.59-1.61 (2H, m), 1.99-2.02 (2H, m), 2.24 (2H, t, J = 7.4 Hz), 4.16 (2H, t, J = 7.3 Hz), 7.57 (2H, m), 7.86 (2H, d, J = 8.6 Hz), 7.93 (2H, m), 8.51 (2H, d, J = 8.7 Hz); ^{13}C NMR (90.5 MHz, $(CD_3)_2SO$) 22.6, 25.1, 26.5, 28.8, 28.9, 29.0, 30.9, 35.2, 50.3, 116.6, 121.7, 125.0, 126.0, 130.2, 146.9, 152.1, 174.7; IR ν_{max} (Nujol) 3385, 2920, 2850, 1635, 1560, 1455, 1375 cm^{-1}; UV λ_{max} nm (MeOH) (ε) 220 (11,892), 267 (28,793), 410 (5550), 434 (4865); FAB-MS m/z 391 ((M + Na)$^+$, 90%) 379 (M$^+$, 100%), 279 (100%),. 205 (85%); high resolution FAB-MS calcd. for $C_{24}H_{31}N_2O_2$ 379.23854, found 379.23855.

Synthesis of peptide (1)

The synthesis of peptide (1) was carried out on a 0.25 mmol scale on an Applied Biosystems 430A peptide synthesiser, equipped with an Applied Biosystems 757 absorbance detector linked to a Hewlett-Packard HP3396A integrator for on-line monitoring of the deprotection solution containing the piperidine-Fmoc adduct. Glycine was single-coupled using 1 mmol of symmetrical anhydride. Asparagine and glutamine were both double-coupled as their HOBt-active esters (2 x 0.5 mmol amino acid). All other amino acids were double-coupled, first as the symmetric anhydride (0.5 mmol anhydride) followed by the HOBt active ester (0.5 mmol amino acid). All couplings were carried out using DIC in dioxan as the activating agent. Each coupling step lasted 30 min, and was followed by a capping step using acetic anhydride in DMF for 6 min.

The Fmoc group was then removed with successive 5, 3, 3 and 1 min treatments with 20% piperidine in DMF, with extensive washing with DMF between each deprotection.

After the final glycine residue had been installed, and the Fmoc group removed, the resin was removed from the machine. The N-(9-acridinyl)-11-aminoundecanoic acid (**2**) (189 mg, 0.5 mmol) was suspended in DMF (4 ml) and a solution of HOBt (34 mg, 0.25 mmol) in DMF (2 ml) added, followed by a solution of DIC (32 mg, 0.25 mmol) in dioxane (2 ml). The mixture was sonicated for 15 min., then the resin added and the sonication continued for 24 h in the dark. The mixture was then filtered, the resin washed with DMF, and then subjected to a second coupling with N-(9-acridinyl)-11-aminoundecanoic acid (**2**) under the same conditions. The resin was then washed again with DMF, partially dried and stored under nitrogen at 4 °C in the dark until required.

A sample of the resin (368 mg) was placed in a sinter funnel attached to a round-bottomed flask and a 95:5 mixture of trifluoroacetic acid and water (5 ml) added. The mixture was allowed to stand for 15 min, then drawn through the sinter under suction onto a mixture of crystalline phenol (300 mg) ethanedithiol (800 µl) and thioanisole (200 µl) in the flask. The washing was repeated twice with the 95% TFA/water mixture (5 ml). The flask was then purged with N_2 and allowed to stand at room temperature for 7 hr. The resulting yellow solution was then concentrated in vacuo to about 1 ml, and Et_2O (20 ml) added. The peptide was allowed to precipitate at 4 °C for 24 h, then the precipitate was collected by centrifugation, dissolved in acetic acid and lyophilised. The peptide was then purified by preparative HPLC (20-45% B in 25 min, with the peptide eluting at 15 min), giving 4 mg product. m/z (MALDI TOF) 3942 (M+H). Amino acid analysis: found (required) Asp + Asn 1.935 (2) Thr 0.994 (1) Ser 1.568 (1) Glu + Gln 3.735 (3) Pro 0.86 (1) Gly 4.223 (4) Ala 5.778 (6) Cys 0.725 (1) Met 0.793 (1) Leu 1.83 (2) Lys 3.378 (3) Arg 6.780 (7).

Dimerisation of the peptide

To a solution of peptide (1) (4.8 mg, 1.2 µmol) and diphenyl sulfoxide (2.5 mg, 12 µmol, 10 eq) in trifluoroacetic acid (20 µl) and anisole (13.2 µl, 120 µmol, 100 eq) under Ar was added methyltrichlorosilane (13 µl, 120 µmol, 100 eq). The reaction was stirred at room temperature under Ar for 2 h, then cooled to 0 °C and ether (500 µl) and 10% acetic acid/90% water (500 µl) added. The mixture was vortexed, centrifuged and the layers separated. The aqueous layer was lyophilised to give a yellow powder (6.1 mg). This was purified by reverse phase HPLC (20-45% B in 25 min, with the dimerised peptide eluting at 20 min) and lyophilised to give 1.5 mg dimerised peptide. Electrospray gave the mass of this peptide as 7738.45 ± 0.52.

ACKNOWLEDGEMENTS

I would like to thank Mr. Martin Andrews and Ms Sarah Devine (University of Edinburgh) for preliminary work; Professor Robert Ramage, Mr. Kevin Shaw and Mr. Brian Whigham (University of Edinburgh) for assistance with the peptide synthesis, MALDI TOF mass spectrometry and amino acid analysis; Mr. Mark Domin (London School of Pharmacy) for MALDI TOF mass spectrometry; and Dr. Steven Howell (NIMR, Mill Hill) for electrospray mass spectrometry. Financial support from the Royal Society and the Nuffield Foundation (Undergraduate Research Bursary to M. A.) is also gratefully acknowledged.

REFERENCES

1. Thompson, C. C., McKnight, S. L., *Trend Genet.*, **1992**, *8*, 232-236.
2. Landers, J. P., Spelsberg, T. C., *CRC Crit. Rev. Eukaryotic Gene Expression*, **1992**, *2*, 19-63.
3. For a review, see: Thuong, N. T., Hélène, C., *Angew. Chem. Int. Ed. Engl.*, **1993**, *32*, 666-690.
4. Collier, D. A., Thuong, N. T., Hélène, C., *J. Am. Chem. Soc.*, **1991**, *113*, 1457-1458; Perroualt, L., Asseline, U., Rivalle, C., Thuong, T., Bisagni, E., Giovannangeli, C., Le Doan, T., Hélène, C., *Nature*, **1990**, *344*, 358-360.
5. Orson, F. M., Kinsey, B. M., McShan, W. M., *Nucleic Acids Res.*, **1994**, *22*, 479-484.
6. Hinnebusch, A. G., *Proc. Natl. Acad. Sci. U. S. A.*, **1984**, *81*, 6442-6446.
7. Ellenberger, T. E., Brandl, C. J., Struhl, K., *Cell*, **1992**, *71*, 1223-1237.
8. Kim, J., Tzamerias, D., Ellenberger, T., Harrison, S.C., Struhl, K., *Proc. Natl. Acad. Sci. U. S. A.*, **1993**, *90*, 4513-4517.
9. Suckow, M., von Wilken-Bergmann, B., Müller-Hill, B., *EMBO J.*, **1993**, *12*, 1193-1200.
10. Talanian, R. V., McKnight, C. J., Rutkowski, R., Kim, P. S., *Biochemistry*, **1992**, *31*, 6871-6875.
11. Cuenoud, B., Schepartz, A., *Proc. Natl. Acad. Sci. U. S. A.*, **1993**, *90*, 1154-1159.
12. Neidle, S., Jenkins, T. C., *Meth. Enzymol.*, **1991**, *203*, 433.
13. Albert, A., Ritchie, B., *Org. Syn. Coll. III*, 53-56.
14. Thuong, N. T., Asseline, U. *Oligonucleotides and Analogues: A Practical Approach;* Eckstein, F. Ed.; IRL Press: Oxford: 1989.
15. Atherton, E., Sheppard, R. C., *Solid Phase Peptide Synthesis; A Practical Approach*; IRL Press: Oxford: 1989, pp. 283-287.
16. Ramage, R., Irving, S. L., McInnes, C., *Tetrahedron Lett.*, **1993**, *34*, 6599.
17. N. Fujii, A. Otaka, T. Watanabe, A. Okamachi, H. Tamamura, H. Yajima, Y. Inagaki, M. Nomizu, K. Asano, *J. Chem. Soc., Chem. Commun.*, **1989**, 283.
18. J. P. Tam, C.-R. Wu, W. Liu, J.-W. Zhang, *J. Am. Chem. Soc.*, **1991**, *113*, 6657-6662.
19. Kamber, B., *Helv. Chem. Acta.*, **1971**, *54*, 927-930.
20. K. Akaji, T. Tatsumi, M. Yoshida, T. Kimura, Y. Fujiwara, Y. Kiso, *J. Am. Chem. Soc.*, **1992**, *114*, 4137-4143.
21. See, for example; Fkyerat, A., Demeunyck, M., Constant, J-F., Lhomme, J., *Tetrahedron* , **1993**, *49*, 11237-11252, and ref. 5.

(*Received in UK* 1 *November* 1995; *revised* 27 *November* 1995; *accepted* 30 *November* 1995)