Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tetrahedron Letters 45 (2004) 6783–6786
Tetrahedron Letters

Synthesis and properties of a pyrrole–imidazole polyamide having a ferrocene dicarboxylic amide linker

Kohji Seio,[a,c] Masahiro Mizuta,[b,c] Takeshi Terada[b] and Mitsuo Sekine[b,c,*]

[a]*Frontier Collaborative Research Center, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan*
[b]*Department of Life Science, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501, Japan*
[c]*CREST, JST (Japan Science and Technology Agency), 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501, Japan*

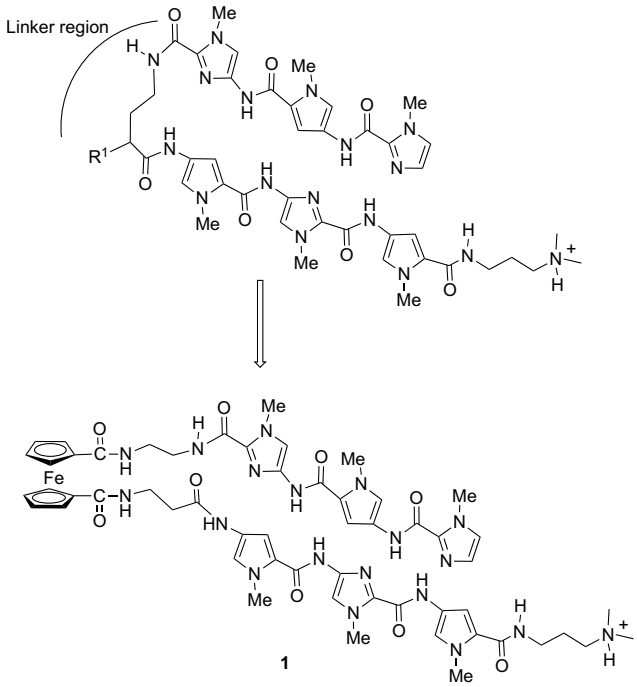
Received 22 April 2004; revised 30 June 2004; accepted 8 July 2004
Available online 3 August 2004

Abstract—A new minor groove binder **1** having a redox active ferrocene dicarboxamide linker was synthesized. The affinity of the newly synthesized compound to the target DNA was confirmed by CD titration experiments. Moreover, the redox property of **1** was also demonstrated by cyclic voltammetry. These results suggest the potential of **1** as a new tool for electrochemical gene detection technology.

The hairpin pyrrole–imidazole polyamide (PIPA)[1] is the prototype of an artificial organic molecule that can bind to DNA with a highly sequence specific manner. The potential of such PIPA compounds as the gene regulator was proposed earlier.[1,2] Afterward, a lot of structural modifications on the original PIPA have been reported in order to modify the functions and properties. For example, several research groups reported the structural modification of the original PIPAs by fluorescent dyes[3] or alkylating agents,[4] and the application of such molecules as the gene detecting tools and the potential anticancer agents.

Apart from such functional modification, several research groups reported the structural modifications of the pyrrole and imidazole structures by use of other heteroaromatic compounds in order to improve the affinity and sequence specificity of PIPAs to DNA duplexes, and several of these modifications succeeded in the enhancement of the sequence selectivity.[5]

Schematically, the structure of the original PIPAs can be described as a 6- or 8-mer of pyrrole–imidazole polyamide with a linker region at the center of the molecule (Scheme 1). Therefore, besides addition of the pendant groups and the heteroaromatic modifications described above, the linker region can be another site for modification. As usual, 4-aminobutylic acid has been used as a linker.



Scheme 1.

* Corresponding author. Tel.: +81-45-924-5706; fax: +81-45-924-5144; e-mail: msekine@bio.titech.ac.jp

0040-4039/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2004.07.060

We are interested in the PIPA derivative **1** having a ferrocene moiety as a linker (Scheme 1). Ferrocene derivatives have been widely used as the building blocks in the field of supramolecular chemistry because of their unique 3D structure and the conformational properties.[6] In addition, the typical distance between the two Cp rings of ferrocene derivatives is approximately 3.3 Å,[7] which is close to the distance between the two PIPA moieties in hairpin PIPAs that bind to DNA minor grooves in the U shape structure.[8] Therefore, it is expected that the ferrocene moiety of **1** fits to the minor groove when it binds to DNA duplex in a similar conformation to that of conventional hairpin PIPA.

Moreover, because the electrochemical properties of ferrocene derivatives has been attractive from the view point of chemosensor[9] and biosensor especially in the gene detection,[10] the ferrocene modified PIPA molecules can be a new sequence specific ligand for the electrochemical detection of genes by use of minor groove binder.[11]

Synthesis of the target molecule **1** is shown in Scheme 2. The structure of **1** was designed considering both the availability of the starting materials, and the chemical stability of the products and the intermediates. For example, our preliminary attempt to use 1′-aminoferrocene 1-carboxylic acid[12] in place of the ferrocene dicarboxylic acid failed because of the low stability of the compounds containing this moiety.

The sequence of the PIPA region of **1** was designed to bind CGC/GCG sequence targeting a single nucleoside mutation in human genome, which could be a factor that determined the interferon resistance of the patients infected by hepatitis C virus.[13]

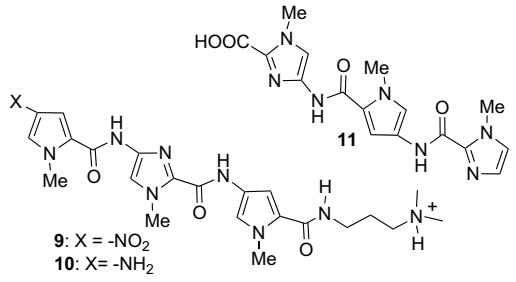
Methyl ferrocene dicarboxylic acid **2**[14] was condensed with *N*-*t*-butoxycarbonyl-1,3-propanediamine by the use of 2-chloro-1-methylpyridinium iodide (CMPI)[15] in 87% yield. Compound **3**, thus obtained was converted to the carboxylic acid **4** by treatment with NaOH, and **4** was condensed with ethyl 3-aminopropionic acid to give the ferrocene building unit **5** in 62% yield.

In order to introduce the pyrrole–imidazole moiety, the ethyl ester of **5** was removed by treatment with NaOH to give **6** in 93%. The liberated carboxylic group of **6** was coupled with 3-aminopyrrole derivative **10** prepared in situ by the hydrogenation of 3-nitropyrrole derivative **9**[16] by use of chloro-1-methylpyridinium iodide to give **7** in 62%. Compound **7** thus obtained was further hydrolyzed by treatment with 10% TFA in CH_2Cl_2 to give the primary amine **8** in 71% yield (Scheme 3).

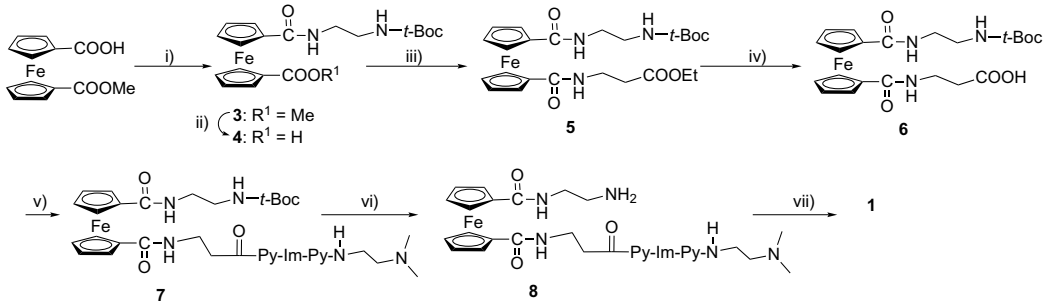
The condensation of **8** with the separately synthesized carboxylic acid **11**[17] was examined to obtain **1** by use of various coupling reagents and catalysts. All attempts to use the reagents such as EDC/DMAP, di(2-pyridyl)carbonate/DMAP, diphenylphosphoryl chloride, and CMPI failed probably because of the steric hindrance of **8**. An attempt to obtain **1** via the corresponding acid chloride of **11** was also unsuccessful.

Finally, we found that *O*-(benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HBTU)[18] was the best condensing agent to activate **11** and the desired target **1** could be obtained in 61% yield. The structure of **1** was confirmed by ^{1}H NMR, ^{13}C NMR, and ESI-MS spectroscopy.[19]

In order to clarify the usefulness of **1** as a ligand for double helical DNA, the binding of **1** to the DNAs with its target sequence, **DNA1**: 5′-GACTGCGTAGG-3′/3′-CTGACGCATCC-5′ and non-target sequence, **DNA2**: 5′-GACTGAGTAGG-3′/3′-CTGACTCATCC-5′ were examined by use of CD spectra. Shown in Figure 1 are the CD profiles of the titration of **DNA1** and **DNA2** with various amounts of **1**. It is well known that the interaction of minor groove binder induces positive elip-



Scheme 3.



Scheme 2. Reagents and conditions: (i) *N*-*t*-Butoxycarbonyl-1,3-propanediamine (1.0 equiv), CMPI (1.2 equiv), N(*n*-Bu_3) (2.4 equiv), CH_2Cl_2, 0.5 h, 87%. (ii) 0.5 M NaOH/MeOH (1:1, v/v), 80 °C, 12 h, 91%; (iii) β-alanine ethyl ester hydrochloride (1.0 equiv), CMPI (1.2 equiv), N(*n*-$Bu)_3$ (3.4 equiv), CH_2Cl_2, 2.5 h, 62%; (iv) 0.1 M NaOH aq/MeOH (1:1, v/v), 1 h, then H^+, 93%; (v) **10** (1.0 equiv), DCC (1.0 equiv), HOBt (1.0 equiv), CH_2Cl_2, 4 h, 62%. (vi) 10% TFA/CH_2Cl_2, 6 h, 88%; (vii) HBTU (2.0 equiv), *N*,*N*-diisopropylethylamine (2.0 equiv), DMF, 1 h, 61%.

ticity at 300–360 nm.[20] In the combination of **1** and **DNA1**, the positive cotton effect increases until 1:1.8 **DNA1**:**1** stoichiometry is reached. The maximum magnitude of the induced signal was 5.3×10^5. Addition of **1** beyond 1:1.8 stoichiometry did not lead to significant increase of the peak intensity (Fig. 1a).

On the other hand, in the case of the titration of **DNA2** by **1**, the saturation of the positive cotton effects required 1:6.0 stoichiometry with the maximum induced signals of 2.4×10^5, which was much smaller than that of the combination of **DNA1** and **1** (Fig. 1b). Previously, Breslauer et al. studied the sequence specific binding of a hairpin polyamide to DNA duplexes by CD titration.[20] In their analysis, a positive correlation between the magnitudes of the induced CD signals and binding constants were observed. The empirical correlations together with the CD titration profiles in Figure 1 indicated that **1** bound to **DNA1** more tightly than to **DNA2**. Accurate binding constants could not be determined for **1** because of the presence of multiple binding modes at high **1** to **DNA1** or **DNA2** ratios indicated by the disappearance of the isosbestic points. The complexity of the binding mode might be partially attributable to the palindromic GCG target sequence. Although the details of the binding modes and the quantitative binding constants are yet to be clarified, it was revealed that the pyrrole–imidazole polyamide compound **1** could bind to the double strand DNA in a sequence specific manner.

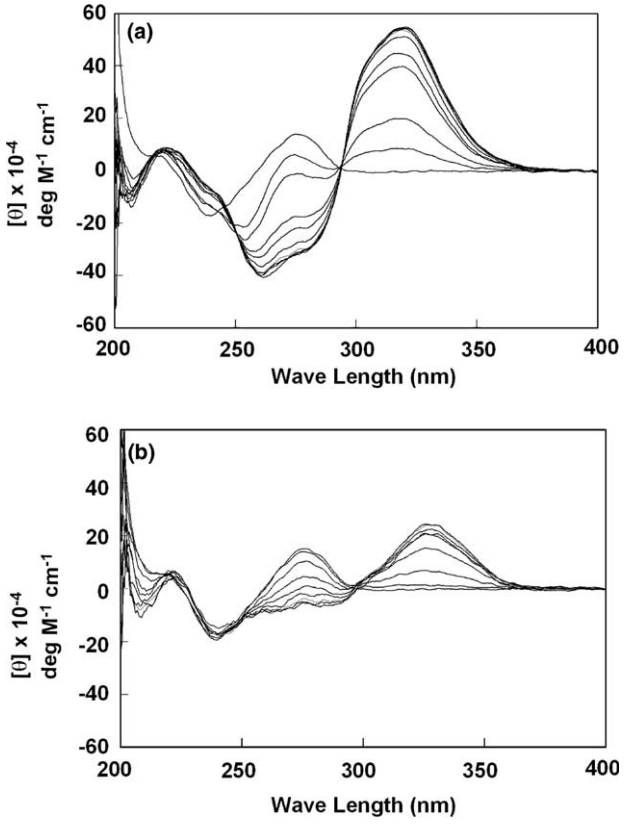


Figure 1. CD profiles of the titration of the 5 μM of (a) **DNA1** and (b) **DNA2** by **1** in 10 mM sodium cacodylate (pH 6.9) containing 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

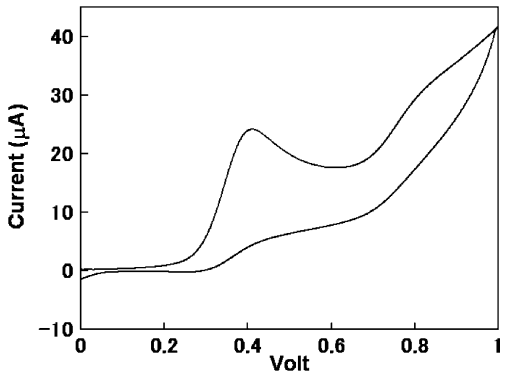


Figure 2. Cyclic voltammogram of **1** (5 mM) in *N*,*N*-dimethylformamide. Supporting electrolyte: 0.2 M $NaClO_4$; reference electrode: Ag/$AgCl_2$; counter electrode: Pt wire; working electrode: Pt plate.

Finally, we examined the electrochemical properties of **1** by use of cyclic voltammetry (CV). The result is shown in Figure 2. Apparently, compound **1** showed clear oxidation wave which was attributable to the oxidation of the ferrocene moiety and did not show any clear reduction wave. Moreover, when the voltammograms were recorded repeatedly, the oxidation peak disappeared gradually. These behaviors indicated the irreversible properties of **1** because of the conversion of the electrochemically oxidized products to redox-inactive species before being reduced to **1**.

The electrochemical activity of **1** together with the affinity toward the target DNA described above indicates the potential usefulness of **1** as a duplex specific redox-active ligand[10c,11] for the electrochemical detection of genes. There have been reported several systems to detect DNA–DNA hybridization electrochemically.[21] However, the sequence selectivity of these systems are completely dependent on the ability of the probe DNAs to discriminate the non-complementary sequences. It is well known that canonical nucleobases can form considerably stable mismatch base pairs[22] such as G–T, A–G, and G–G, which can increase the false-positive signals in the gene detections using oligonucleotide probe. In contrast, if compound **1**, which has both redox activity and ability to recognize Watson–Crick base pair, were used in electrochemical gene-detection systems in the combination with appropriate probe DNAs, the sequence selectivity of the system must be very high because the sequence selectivity of the probe DNA and **1** must be multiplied. Although the CD titration experiments shown in Figure 1 revealed the potential sequence selectivity of **1**, the binding modes, and binding constants to high and low affinity sites should be clarified in more detail to achieve this purpose. The application of **1** in this direction is in progress and will be reported elsewhere.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research on Priority Areas (C) 'Medical Genome Science' from the Ministry of Education, Culture, Sports,

Science and Technology of Japan and by CREST of JST (Japan Science and Technology Agency). This work was also supported by The Sumitomo Foundation. We thank Dr. Yoshio Ishimori (Toshiba Co. Ltd) for fruitful discussion throughout this study.

References and notes

1. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983–7988.
2. Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202–205.
3. Rucker, V. C.; Foister, S.; Melander, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 1195–1202.
4. (a) Wang, Y. D.; Dziegielewski, J.; Wurtz, N. R.; Dzielewska, B.; Dervan, P. B.; Beerman, T. A. *Nucleic Acids Res.* **2003**, *31*, 1208–1215; (b) Bando, T.; Narita, A.; Saito, I.; Sugiyama, H. *Chemistry* **2002**, *18*, 4781–4790.
5. (a) Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B. *Bioorg. Med. Chem. Lett.* **2003**, *11*, 4333–4340; (b) Urbach, A. R.; Szewczyk, J. W.; White, S.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 11621–11629.
6. (a) Tomapatanaget, B.; Tuntulani, T.; Chailapakul, O. *Org. Lett.* **2003**, *9*, 1539–1542; (b) Ion, A.; Buda, M.; Moutet, J.-C.; Saint-Aman, E.; Royal, G.; Gautier-Luneau, I.; Bonin, M.; Ziessel, R. *Eur. J. Inorg. Chem.* **2002**, *22*, 1357–1366; (c) Plumb, K.; Kraatz, H.-B. *Bioconjugate Chem.* **2003**, *14*, 601–606.
7. Fey, N. J. *Chem. Technol. Biotechnol.* **1999**, *74*, 852–862.
8. (a) Kielkopf, C. L.; Bremer, R. E.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *J. Mol. Biol.* **2000**, *295*, 557–567; (b) Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104–109.
9. (a) Linjing, H. H.; Mu, L.; He, J.; Cheng, J. P. *J. Org. Chem.* **2003**, *68*, 7605–7611; (b) Saji, T. *Chem. Lett.* **1986**, *22*, 275–276.
10. (a) Fan, C.; Plaxco, K. W.; Heeger, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9134–9137; (b) Ihara, T.; Maruo, Y.; Takenaka, S.; Takagi, M. *Nucleic Acids Res.* **1996**, *24*, 4273–4280; (c) Sato, S.; Maeda, Y.; Nojima, T.; Kondo, H.; Takenaka, S. *Nucleic Acids Res. Suppl.* **2003**, *3*, 169–170.
11. Hashimoto, K.; Ito, K.; Ishimori, Y. *Anal. Chem.* **1994**, *66*, 3830–3833.
12. Barisic, L.; Rapic, V.; Kovac, V. *Croat. Chem. Acta* **2002**, *75*, 199–210.
13. Hijikata, M.; Ohta, Y.; Mishiro, S. *Intervirology* **2000**, *43*, 124–127.
14. Barisic, L. R.; Vladimir, K. V. *Croat. Chem. Acta* **2002**, *75*, 199–210.
15. Bald, E.; Saigo, K.; Mukaiyama, T. *Chem. Lett.* **1975**, *22*, 1163–1166.
16. Coumpound **9** was prepared by use of the monomer units, 1-methyl-3-nitro-5-trichloroacetylpyrrole and 1-methyl-4-nitro-2-trichloroacetylimidazole reported in Ref. 17. Compound **11** were prepared by use of 2-ethoxycarbonyl-1-methyl-4-nitropyrrole, 1-methyl-2-trichloroacetylimidazole and 2-ethoxycarbonyl-1-methyl-4-nitroimidazaole.
17. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6146–6414.
18. (a) Dourtoglou, V.; Ziegler, J.-C.; Gross, B. *Tetrahedron Lett.* **1978**, *15*, 1269–1272; (b) Dourtoglou, V.; Gross, B. *Synthesis* **1969**, 572–574.
19. H NMR (DMSO-d_6): δ 1.59–1.65 (2H, m), 2.15 (6H, s), 2.26–2.28 (2H, t, J = 7.1 Hz), 2.56–2.59 (2H, t, J = 7.0 Hz), 3.18–3.22 (2H, m), 3.37–3.50 (6H, m), 3.81 (3H, s), 3.83 (3H, s), 3.85 (3H, s), 3.94 (3H, s), 3.97 (3H, s), 3.99 (3H, s), 4.27–4.30 (4H, m), 4.73 (4H, s), 6.88 (1H, d, J = 1.7 Hz), 6.94 (1H, d, J = 1.5 Hz), 7.02 (1H, s), 7.16 (1H, d, J = 1.7 Hz), 7.19 (1H, d, J = 1.5 Hz), 7.26 (1H, d, J = 1.7 Hz), 7.35–7.36 (2H, m), 7.46 (1H, s), 7.49 (1H, s), 7.86–7.87 (1H, t, J = 5.6 Hz), 7.91–7.93 (1H, t, J = 5.5 Hz), 7.98–7.80 (1H, t, J = 5.5 Hz), 8.05–8.08 (1H, t, J = 5.7 Hz), 9.77 (1H, s), 9.88 (1H, s), 10.03 (1H, s), 10.06 (1H, s), 10.19 (1H, s); ^{13}C NMR (DMSO-d_6): δ 27.2, 35.1, 35.1, 35.3, 36.0, 36.1, 36.2, 36.4, 36.5, 37.4, 38.8, 39.0, 45.3, 57.3, 69.8, 71.7, 71.8, 77.9, 78.0, 79.4, 104.1, 104.8, 106.1, 114.6, 114.9, 118.0, 119.4, 119.6, 121.4, 121.7, 122.0, 122.3, 122.4, 123.5, 126.6, 127.2, 134.1, 134.4, 136.2, 138.9, 155.9, 156.3, 158.9, 158.9, 159.2, 161.2, 168.2, 168.6, 169.0; MS *m/z* calcd for $C_{55}H_{66}FeN_{19}O_9$ $[M + H]^+$: 1191.45, found 1191.99.
20. Pilch, D. S.; Poklar, N.; Baird, E. E.; Dervan, P. B.; Breslauer, K. J. *Biochemistry* **2002**, *38*, 2143–2151.
21. Drummond, T. G.; Hill, M. G.; Barton, J. K. *Nat. Biotechnol.* **2003**, *21*, 1192–1199.
22. Peyret, N.; Seneviratne, P. A.; Allawi, H. T.; SantaLucia, J., Jr. *Biochemistry* **1999**, *38*, 3468–3477.