私立東海大學應用化學研究所博士論文

指 導 教 授：佘 亮 博士

圓二色光譜及 MALDI-TOF MS 於 DNA-胜肽交互作用之應用

Applications of Circular Dichroism and MALDI-TOF MS Spectroscopy in DNA-Peptide Interactions

研究生：黃則棓

中華民國 一百年七月

摘 要

許多疾病的產生都是起因於基因異常的表現，特別是癌症的成因，幾乎都是源自基因的改變，而具特定序列的胜肽能與 DNA 核酸產生作用力，因此，對胜肽分子與核酸的作用機制及生物功能的研究，在生命科學研究領域中佔有極重要的內容。

本文分為兩大部分：

1. 利用圓二色光譜測定胜肽與 DNA 作用之研究，其中胜肽為序列中含 XP(Hyp)RK 及 4-amino-1-methylpyrrole-2-carboxylic acid (Py) 之結合型胜肽及含 CLB (Chlorambucil) 之對 DNA 有切割作用之胜肽。
2. 利用 MALDI-TOF MS 研究含 CLB (Chlorambucil) 之胜肽對 DNA 切割序列的選擇性及其機轉。

本文中所使用的 DNA 是含有胜肽對 pBR322 DNA 選擇性結合位置 5'-d（AAAA）-d（TTTT）-3'之序列的 U4A-L4T DNA，U4A-L4T DNA 是一有 13 個鹼基對之雙股螺旋 DNA，其序列為 5'-d（TAGGAGAAAATAC）-d（GTATTTTCTCCTA）-3'，而 CD 圖譜顯示 U4A-L4T DNA 為 B-form DNA，而實驗之胜肽在波長~300 nm 均有正吸收峰，顯示均為在 minor groove 位置結合。研究時在[peptide] / [DNA] 對 $\Delta\theta$ 的關係中發現，當[peptide] / [DNA] 比在 0.5 - 2.0 之間時只有 1 分子的胜肽結合在 5'-d（AAAA）-d（TTTT）-3' 之位置上，而當[peptide] / [DNA] 比在 2.0 - 3.0 之間時，則有 2 分子的胜肽結合在選擇性之序列上。含 CLB 之切割型胜肽在橢圓率的變化上都比與之相似序列之結合型胜肽來的小，而部分的切割型胜肽有清楚的飽和現象，是因為被切開的 DNA 其小凹槽構象改變，使得胜肽無法繼續在小凹槽結合。在結合型胜肽的效果上 PyMK-10 最好，而 PySK-10 最差，切割型胜肽則是 CLB-HyQ-10 最好，而 CLB-PyMK-10 最差。因此，圓二色光譜在胜肽與 DNA 交互作用之研究提供以下重要數據: DNA 構象變化依靠胜肽濃度變化，含 XP(Hyp)RK 及 Py 胜肽選擇結合在 DNA 小凹槽，及 DNA-胜肽結合比例。

在含 CLB 之切割型胜肽對 U4A-L4T DNA 切割片段的比對上，利用 MALDI-TOF MS 尋找出精確的 DNA 片段分子量，並找出不同胜肽對 DNA 選擇性切割的位置。CLB-HyM-10 對 U4A-L4T DNA 有很高的專一性切割，位置在 U4A 的 A_5 和 G_6 之間及 L4T 的 T_{20} 和 C_{21} 之間，其中 A_5 和 C_{21} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份，比對時也有發現 L4T 的 T_{15} 和 A_{16} 及 A_{16} 和 T_{17} 之間的切割片段分子量的荷質比訊號，而 A_{16} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份。CLB-HyQ-10 則呈現和 CLB-HyM-10 相似的結果，但多了在 U4A 的

G_3 和 G_4 之間及 L4T 的 C_{23} 和 C_{24} 之間的切割位置。CLB-HyH-10 則切割在 U4A 的 G_4 和 A_5 之間及 L4T 的 C_{23} 和 T_{22} 之間，A_5 的鹼基部位會被胜肽切掉只留下其五碳醣的部份。在 CLB-PyMK-10、CLB-PyHK-10、CLB-PyHK-10、CLB-PyHK-10 對 U4A-L4T DNA 切割片段的比對上有著很高的相似性，他們都只對 U4A 的部份切割，切割位置在 G_4 和 A_5、A_5 和 G_6 之間，而且均不會只去除鹼基而留下五碳醣。本研究顯示不 同序列的胜肽呈現不同的 DNA 序列切割選擇性。

關鍵詞: 胜肽與 DNA 序列選擇性之交互作用研究，
含 CLB (Chlorambucil) 之胜肽對 DNA 切割序列的選擇性 ，
圓二色光譜研究， MALDI-TOF MS 研究。

Abstract

Many diseases are due to the generation of abnormal gene expression, especially cancer causes changes in almost all derived from the gene, and a sequence-specific nucleic acid peptide with DNA force generated, therefore, the peptide molecule and nucleic acids and biological function of the mechanism of action research in the field of life science research plays a very important content.

The contents of this thesis are divided into two parts:

1. Circular dichroism studies of DNA-peptide interactions employing peptides incorporating the XP(Hyp)RK motif and 4-amino-1-methylpyrrole-2-carboxylic acid (Py) residues.

2. Sequence-selective cleavage studies DNA by chlorambucil-peptide conjugates.

The duplex DNA used in these studies is a 13-mer oligomer: 5'-d(TAGGAGAAAATAC)-d(GTATTTTCTCCTA)-3'. CD spectra reveal that all of these peptides induce a strong positive peak around 300 nm, suggesting that the peptides bind strongly to the DNA minor groove. Binding isotherms of $\Delta\theta$ versus [peptide] / [DNA] show that at the [peptide] / [DNA] ratio of 0.5-2.0, only one molecule of peptide binds to the 5'-d(AAAA)-d(TTTT) -3' position, whereas at a ratio of 2.0-3.0, two molecules of peptide bind to this locus. Thus, CD studies of DNA-peptide sequence-selective interactions unveil peptide concentration-dependent DNA conformational changes, minor groove binding preference, and binding stoichiometry.

MALDI-TOF MS spectroscopy was employed to study the DNA cleavage fragments that locate cleavage positions at phosphoester bonds. Strong DNA sequence-selective-cleavage were induced by conjugate CLB-HyM-10 at phosphoester bonds between A_5 and G_6 as well as between T_{20} and C_{21}, of which depurination and depyrimidination occurred on A_5 and C_{21}, respectively. Cleavages also occurred at phosphoester bonds between T_{15} and A_{16}, and between A_{16} and T_{17}, and depurination occurred on A_{16}.

The DNA sequence-selective cleavage pattern of conjugate CLB-HyQ-10 is similar to that of conjugate CLB-HyM-10, with additional cleavage points between G_3 and G_4 of strand U4A, and between C_{23} and C_{24} of strand L4T. Conjugate CLB-HyH-10 prefers to

cleave phosphoester bonds between G_4 and A_5, and between C_{23} and T_{22}, with depurination on A_5.

The DNA sequence-selective cleavages induced by conjugate CLB-PyMK-10, CLB-PyHK-10, CLB-PyQK-10, and CLB-PyWK-10 are very similar, with cleavage preference mainly on the upper U4A strand. Cleavages were assigned between phosphoester bonds between G_4 and A_5, A_5 and G_6, G_6 and A_7. Thus, this study clearly indicates that peptides with different sequence and substitution of amino acid residues afford different DNA sequence-selective cleavage patterns.

Keywords: DNA-peptide sequence-selective interactions, CLB-peptide conjugate induced DNA sequence-selective cleavages, circular dichroism, MALDI-TOF MS spectroscopy

目　錄

第一章　緒論

許多疾病的產生都是起因於基因異常的表現，特別是癌症的成因，幾乎都是源自基因的改變。因此研究合成 DNA 結合小分子或切割 DNA 的專一性載體在生物化學及醫學上是非常重要的領域。研究報導指出，具特定序列的胜肽能與 DNA 核酸產生作用力；因此，對胜肽分子與核酸的作用機制及生物功能的研究，在生命科學研究領域中佔有極重要的內容。測定胜肽與 DNA 互相作用的親和性及結合選擇性，有利於闡明胜肽分子與 DNA 選擇性和作用以及了解其機制，進一步探討胜肽分子結構與 DNA 作用模式及其生物活性的關係。

一般基因表現的機制是在細胞核內有數種蛋白質因子（protein factor）結合到核酸序列中的促進子序列（promoter）上。蛋白質因子會幫助第二型核醣核酸聚合酶（RNA polymerase II）結合至促進子序列上，使得基因的核酸序列被轉錄（transcribe）成訊息核醣核酸（mRNA）。訊息核醣核酸會被送到細胞質（cytoplasma）中並轉譯（translate）成蛋白質。如果可以使得核酸結合分子對去氧核醣核酸的親合力（affinity）大於或接近蛋白質因子並且專一性的結合在促進子序列上，就可以使得蛋白質因子無法結合到促進子序列上進而使得轉錄作用（transcription）被抑制而間接抑制了下游的蛋白質表現[1-4]（圖 1-1A）。另外，核酸結合分子也可以當成將藥物分子或其他化合物運送到去氧核醣核酸的載體（carrier），在核酸結合分子上接上一段活化子（activator）的活性片段並使它專一性的結合到促進子序列上游的活化子結合位（activator binding site）。

活化子的活性片段會幫助蛋白質因子結合到促進子序列上使得轉錄作用因而被活化（圖 1-1B），造成下游的蛋白質表現量增加[5-7]，或是將核酸結合分子和已知會和去氧核醣核酸作用的藥物分子偶合（coupling）。藥物分子可以藉由核酸結合分子的幫助下直接作用在去氧核醣核酸分子上，這樣便可以增加藥物對去氧核醣核酸的專一性，並可以因此而增加藥物的活性並降低藥物的副作用（side effect）[8-10]（圖 1-1C）。

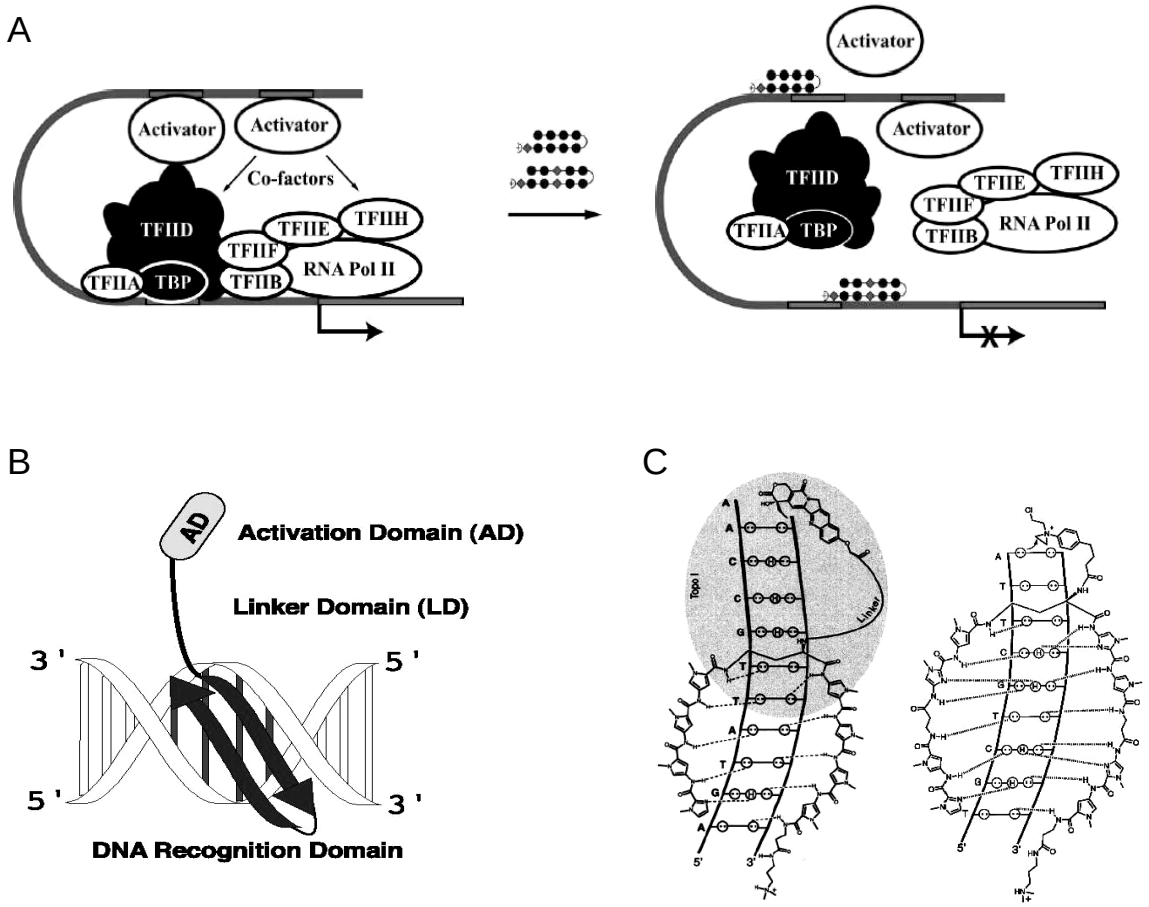


圖 1-1. 核酸結合分子的運用性（A）核酸結合分子如何抑制基因表現[10]（B）核酸結合分子促進基因表現[7]（C）核酸結合分子偶合已知的核酸類藥物[9]

1-1 序列設計來源之 XPRK 模組

Suzuki 教授於 1988 年發表由海膽組織蛋白（Sea Urchin Spermatogenous Histones） H1 及 H2B 中分離出一段含重複六次 SPXX（X 代表鹼基胺基酸殘基）單元（S6 peptide）的蛋白質序列，如圖中 1-2A 框線部分為 SPKK、SPRR、SPRK 等的 SPXX 模組[1-2]。且 H1 的一級結構已經由 Von Holt 等人在 1984 年提出[3]。作者利用圓二色光譜(Circular Dichroism spectrum)發現 Ser-Pro-Arg-Lys(SPRK)所呈現之 DNA 結合結構是的 β 轉折（β-turn）的結構如圖中 1-2B，不同於以往的 α 螺旋(α-helix)、β 摺板(β-sheet)以及鋅手指(Zn-finger)。SPRK 單元結構內含有兩個分子內氫鍵，一個氫鍵是由絲胺酸(Serine) i-th 側鏈烴基上的氧與 i+2-th 胺基上的形成；另一氫鍵是由 i-th 羰基上的氧與 i+3-th 胺基上的氫所形成；因此能形成穩定的 β 轉折(β-turn) 結構有利於與 DNA 結合[4-5]。之後，Churchill 和 Suzuki 也進一步證實 SPKK 單元對於 DNA 小凹槽（minor groove）中富含 A/T （Adenine/ Thymine）序列有的結合性[6]。

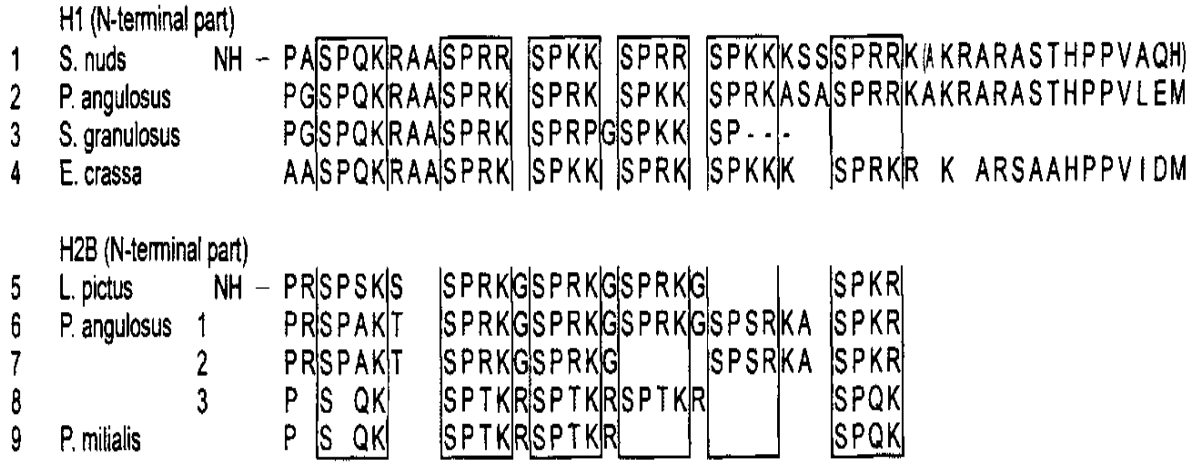


圖 1-2A. 海膽組織蛋白 H1 和 H2B

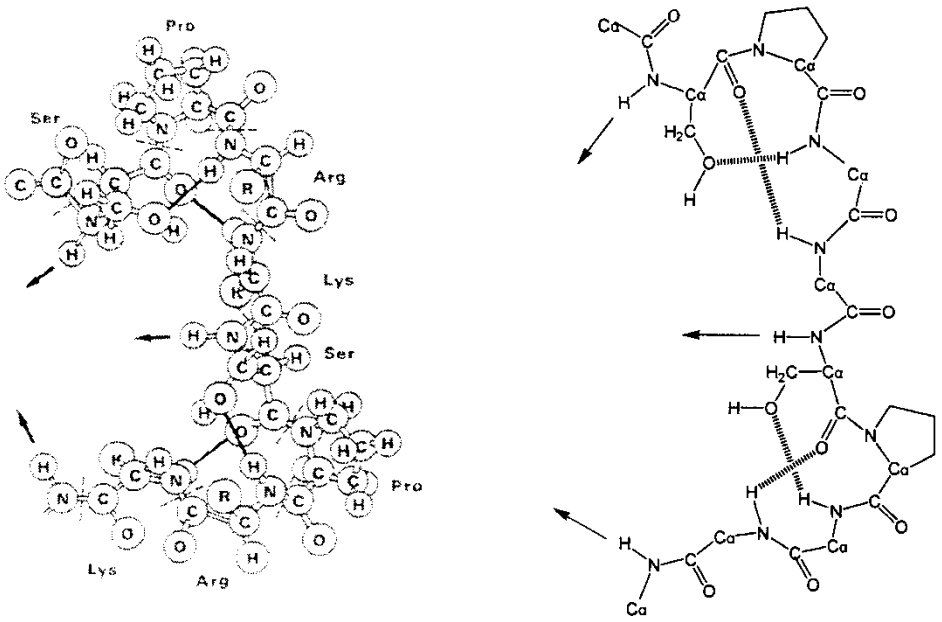


圖 1- 1B.　SPRK 序列 β-turn 結構圖

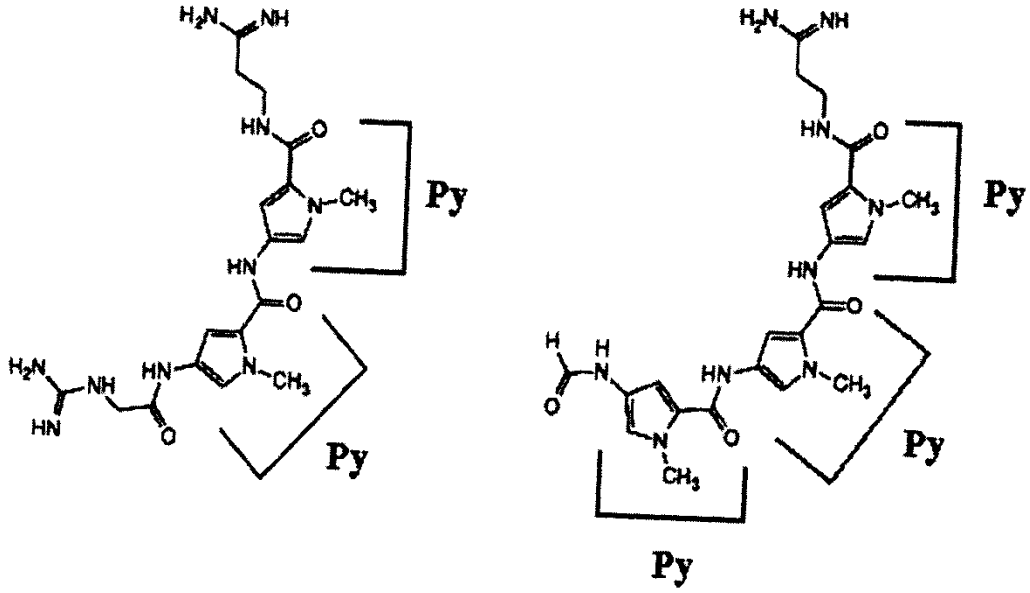
近年來，本實驗小組針對 SPXX 之模組，進行內部胺基酸修飾，於 2003 年提出一種新的結合單元 XPRK[7]；文獻中提出三種含 XPRK 單元胜肽，分別為：

1. YR-12（YPRKYPRKYPRK-NH_2），其中 X 以 Tyrosine 取代
2. WR-12（WPRKWPRKWPRK-NH_2），其中 X 以 Tryptophan 取代
3. SP-30（SPQKRAASPRKSPRKSPKKSPRKASASPRR-NH_2）

其中 SP-30 胜肽選擇性結合富含 A/T 鹼基對之序列，而 YR-12 以及 WR-12 對於 A/T 序列中包含 GC 檢基（Guanine、Cytosine）的 DNA 序列具選擇性結合。因此，本實驗小組近年來依據此構思，更進一步研究修飾或變更 XPRK 模組對於與 DNA 結合性的影響，希望藉此能找出對於 DNA 結合有更高能力之序列。

1-2 含 N-methylpyrrole amino acid（Py）之胜肽

天然抗生素 distamycin（Dst）和 netropsin（Net）由 Streptomyce 提煉出，是第一個被發現對 DNA 小凹槽（minor groove）中富含 AT 序列有專一性鍵結的藥物 [8-11]。從結構中發現 distamycin 有三個 pyrrole（py）結構而 netropsin 有兩個 pyrrole（py）結構。而 Pyrrole（py）結構會與小凹槽中富含 AT 序列位置結合，是因為醯胺中 N 上的 H 質子會與 A （adenine）上的 N（3）及 T （thymine）上的 O（2）產生氫鍵且 Pyrrole ring（py）上的 C-H 會以凡得瓦爾力與 adenine 的 C（2）作用 [12-13]。



Netropsin　　Distamycin

圖 1- 2　Netropsin 與 Distamycin 結構圖

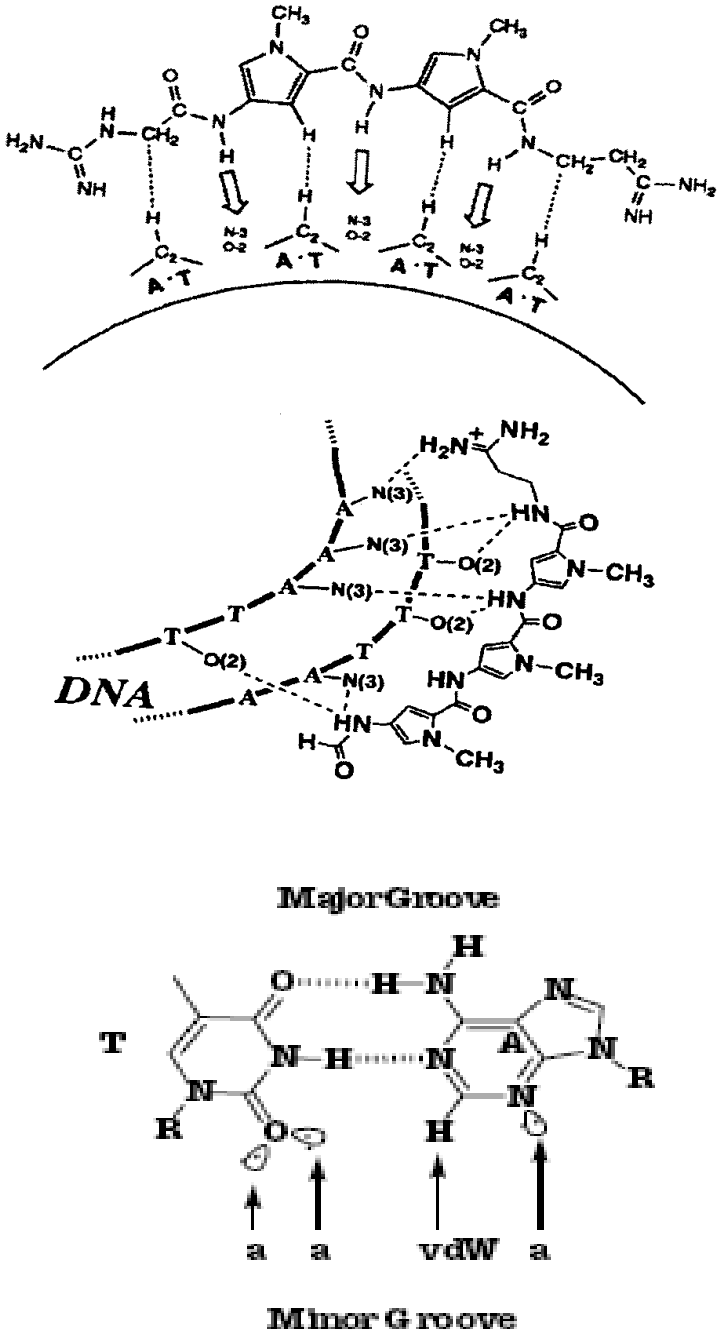


圖 1- 3　Netropsin、Distamycin 與 DNA 結合示意圖

1985 年 Dickerson 分析 netropsin 的晶體結構提出了以一個 imidazole （Im）環狀結構取代一個 pyrrole 用來辨識 GC 鹼基對[14]；imidazole （Im）結構是將 pyrrole 結構中的一個 C-H 以 N 取代。實驗證實 imidazole （Im）確實能對 GC 序列產生辨識專一性結合。Peter Dervan 教授對於 Py、Im 做了許多相關性的研究，Dervan 教授與其他研究人員經由一連串測試及 NMR 圖譜的解釋，證實了 polyamide 與 DNA 的 2：1 反平行的結合模式，發現將此類小分子配對連接成為一 Polyamide 時，能夠高度辨識 DNA 序列與 DNA 小凹槽產生專一性鍵

結，因此發展出 polyamide 的髮夾型（hairpin）結構；利用髮夾型結構使 Py 等小分子對序列產生成對的配對，可得到特定的專一性結合。如 Py 與 Im 相對應時，會有效辨識出 C/G（Cytosine/Guanine）鹼基對；若是 Im 與 Py 相對應時，會辨識出 C/G 鹼基對；而當 Py 與 Py 相對應時，會對 DNA 序列 A/T 或 T/A 進行辨識[15-20]。

在 2004 年，L.Sheh 提出 HPRK 模組與 polyamide 耦合之胜肽序列對 DNA 特定序列具有辨識性，文獻中證實

1. PyH-12 （HPRK（Py）$_4$HPRK-NH_2）胜肽
2. PyH-11（HPRK（Py）$_3$HPRK-NH_2）胜肽
3. PyH-10（HPRK（Py）$_2$HPRK-NH_2）胜肽

大多專一性結合在 158mer Watson 片段連續四個 A 鹼基及 135mer Crick 片段連續四個 T 鹼基的位置，且隨著 Py 的耦合數目不同會對 DNA 的結合能力也有所影響[21]。

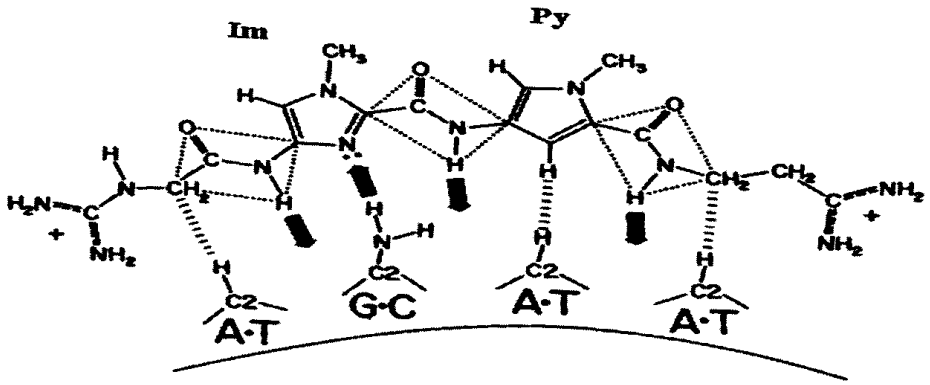


圖 1- 4　Py 用 Im 取代之結構圖

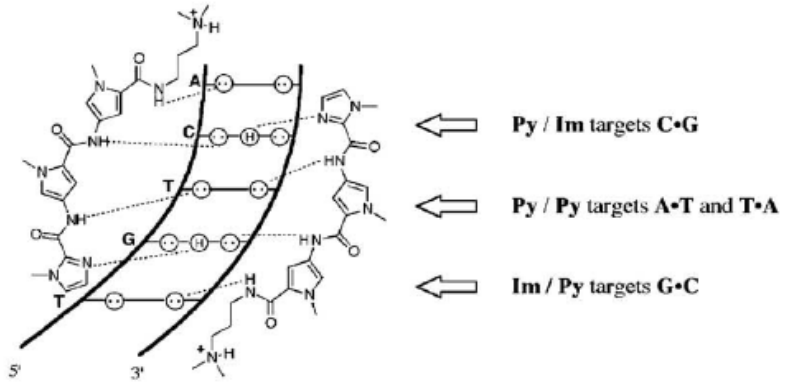


圖 1- 5　2:1 反平行結合模式示意圖

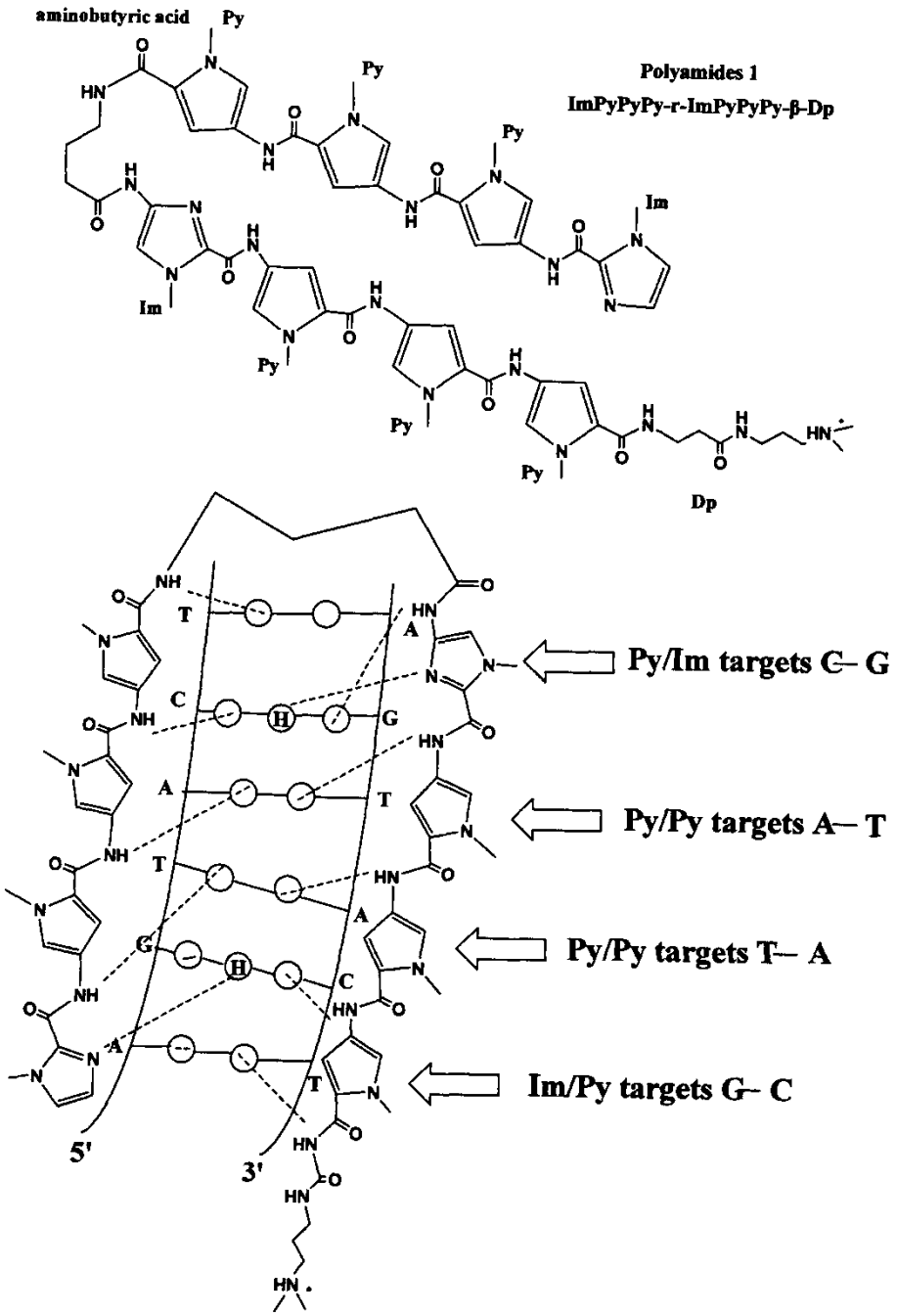


圖 1- 6　Py/Im 之專一性結合圖與髮夾型結構示意圖

1-3 Chlorambucil (CLB) 之胜肽衍生物

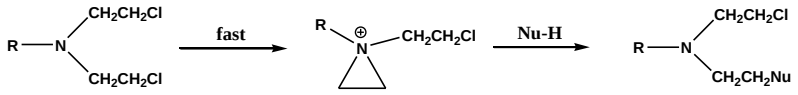
抗腫瘤藥物大致可分為烷化劑(alkylating agent)、抗代謝物(antimetabolites)、抗生素(antibiotics)、有絲分裂抑制劑(mitosisinhibitor)、激素類(hormones)。然而抗腫瘤藥物雖然能夠壞腫瘤細胞並中止細胞的生長、成熟和分裂，但由於這些抗腫瘤藥物並無專一性之辨識能力，所以設計出多肽序列，將之做為一種載體成為一分子目標藥物，特別針對病變部位作治療，如此即可使其他正常細胞不受影響，副作用也相對降低[22-26]。

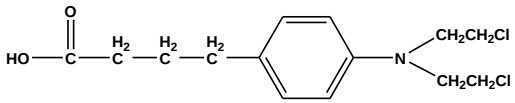
本文所利用的抗腫瘤藥物為氮苯芥丁酸(Chlorambucil ,CLB) 即為烷化劑類藥物。氮苯芥丁酸於 1953 年由 Everett 等人合成[27]，主要用於治療慢性淋巴性血癌(chronic lymphatic leukemia)、淋巴瘤(lymphoma) 及霍奇金氏症(Hodgkin's disease)。此類藥物在化學反應中很容易與親核中心(Nucleophilic center) 作用[28-29]，在人體內與烷基化藥劑結合的核心主要為蛋白質及核酸，而核酸中最容易發生作用以證實在 DNA 之鳥嘌呤(Guanine,G) 上第七位之 Nitrogen(N-7)，此外腺嘌呤(Adenine,A)中的 N-3 及 N-1 位置及胞嘧啶(Cytosine,C) 的 N-1 位置亦可被烷化。

烷化劑與 DNA 作用的方式分為三種:

1. 鳥嘌呤與烷化劑結合後，易發生去嘌呤作用(depurination)，鳥嘌呤部分由核酸鍵中離去，留下非嘌呤酸(apurinic acid)，同時脒唑環(imidazol) 也很容易由位置 8、9 斷裂。
2. 若烷化劑帶兩個官能基，則可能在兩個鳥嘌呤之間發生交叉連結(Cross-linkage)，如此則在 DNA 之雙股螺旋間發生交叉連結，使雙股螺旋(Double strand) 距離拉長，構型(conformation) 改變，而無法成為 RNA 之引子(primer)。
3. DNA 在鳥嘌呤 N-7 位置烷化後，則不在與胞嘧啶成對，而與胸腺嘧啶(Thymine) 成對[30]。

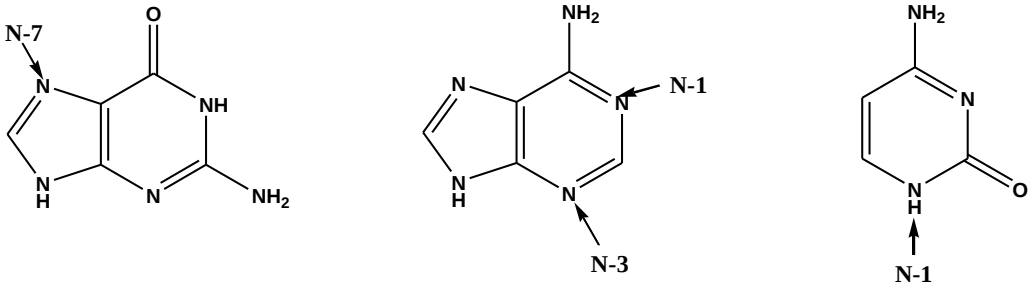
由於本實驗小組研究，純多肽不一定擁有切割能力，但接上 CLB 後即會具有切割效果，其原因是因為 CLB 是一種烷化劑藥物易與 DNA 中的部分鹼基對產生烷化作用，導致 DNA 結構被破壞及斷裂。因此，本實驗小組在 1995 年，經由實驗結果證實與 CLB 耦合之胜肽對 DNA 有選擇性結合及切割作用[31]，並於 2001 年提出含 CLB 與 2,6-Dimethoxyhydroquinone-3-mercaptoacetic acid 之胜肽會對 DNA 進行專一性的切割[32]。





(Chlorambucil,CLB)

圖 1- 7　陽離子形成圖，烷化劑之結構



Guanine　　**Adenine**　　**Cytosine**

圖 1- 8　DNA 鹼基被烷化的位置圖

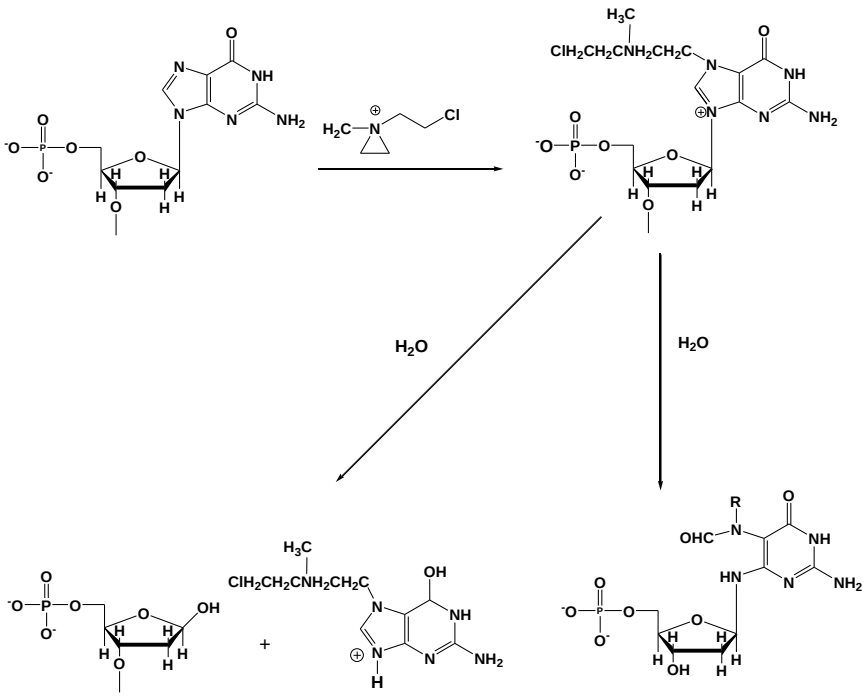
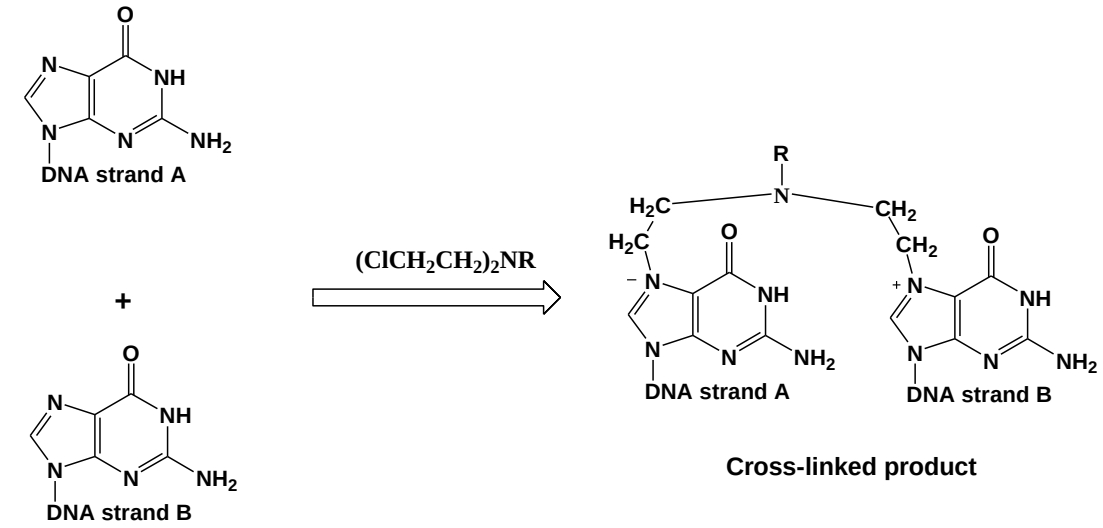


圖 1- 9　Guanine 鹼基被烷化後，水解之產物結構圖



Guanine resudues in DNA

圖 1-11.　Guanine 鹼基被烷化後，cross-linking 結構圖

1-4 本文實驗胜肽之序列

結合型:

RHyp-12 : His-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-Hyp-His

HyM-10 : Met-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

HyQ-10 : Gln-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

HyS-10 : Ser-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

HyH-10 : His-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

HyE-10 : Glu-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

PyHK-10 : Lys-Pro-Arg-Lys-$(Py)_4$-Lys-Arg

PyMK-10 : Lys-Pro-Met-Lys-$(Py)_4$-Lys-Arg

PyQK-10 : Lys-Pro-Gln-Lys-$(Py)_4$-Lys-Arg

PyWK-10 : Lys-Pro-Trp-Lys-$(Py)_4$-Lys-Arg

PySK-10 : Lys-Pro-Ser-Lys-$(Py)_4$-Lys-Arg

PyHR-9 : Lys-Pro-Arg-Lys-$(Py)_4$ –Arg

切割型:

CLB-HyM-10 : CLB-Met-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

CLB-HyQ-10 : CLB-Gln-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

CLB-HyS-10 : CLB-Ser-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

CLB-HyH-10 : CLB-His-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

CLB-HyE-10 : CLB-Glu-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg

CLB-PyHK-10 : CLB-Lys-Pro-Arg-Lys-$(Py)_4$-Lys-Arg

CLB-PyMK-10 : CLB-Lys-Pro-Met-Lys-$(Py)_4$-Lys-Arg

CLB-PyQK-10 : CLB-Lys-Pro-Gln-Lys-$(Py)_4$-Lys-Arg

CLB-PyWK-10 : CLB-Lys-Pro-Trp-Lys-$(Py)_4$-Lys-Arg

本文所使用之胜肽由研究生吳佳容、林筱筑、高君翰、王正一和范淑婷合成及純化。

第二章　材料與方法

2-1 材料

2-1-1 CD 使用之藥品

藥品名稱	藥品來源
1. Sodium Cacodylate	FLUKA
2. DNA	生工合成

2-1-2 MALDI-TOF MS 使用之藥品

藥品名稱	藥品來源
1. Sodium Cacodylate	FLUKA
2. DNA	生工合成
3. Ethanol (99%)	景明
4. NaOAc	景明
5. Methanol	景明
6. trihydroxyacetophenone (THAP)	Acros
7. diammonium hydrogen citrate (AC)	Showa

2-2 實驗原理

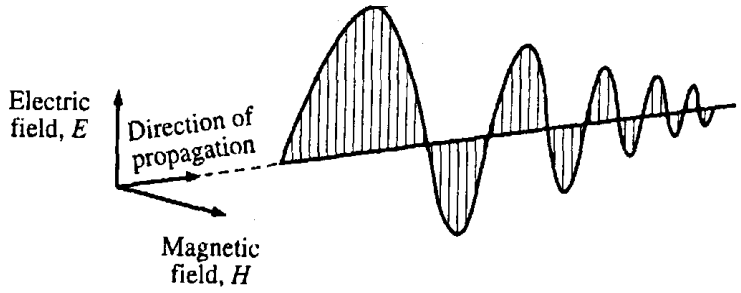
2-2-1 圓二色光譜（Circular Dichroism，CD）

圓二色圖譜是對具有左旋光(levorotatory)和右旋光(dextrorotatory)二種偏轉特性的偏極光(polarization)在通過樣品時有不同的吸收而產生的一種光譜。因此樣品必需有光學活性(optical activity)或含有掌性(chirality)。

一般平面偏極光又稱線偏極光，對著光前進的路線觀察，光的電場或磁場向量都在同一平面上振動。圓偏極光則是對著光前進的路線觀察，光的電場或磁場向量是以光傳播方向為軸的圓形螺旋前進。順時針轉稱右圓偏極光，逆時針轉稱左圓偏極光。橢圓偏極光是對著光前進的路線觀察，光的電場或磁場向量是以光傳播方向為軸的橢圓形螺旋前進。

為了揭示物質的旋光性，科學家菲涅爾提出了以下假設，線偏極光在旋光物質中沿光軸傳播時，分解成為左旋和右旋圓偏極光，它們的傳播速度略有不同，它們相對應的折射率不同，經過旋光物質後產生相位差，而使射出之線偏極光的振動面產生一定偏轉。

如果旋光物質對特定波長的光有吸收，而對左旋和右旋的圓偏極光吸收不同，這種情況下不僅傳播速度不同，振幅也不同。隨著時間的推移，左右圓偏極光合成的光振動向量的末端會呈橢圓軌跡移動。也就是說，速度振幅不同的左右圓偏極光疊加後不在是線偏極光，而是橢圓偏極光，此現象稱為圓二色性。



Linearly (or plane-) polarized light

圖 2-1. 線偏極光示意圖

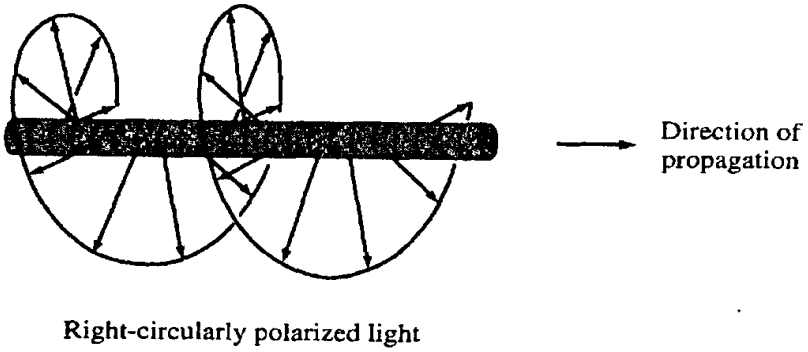


圖 2-2. 右圓偏極光示意圖

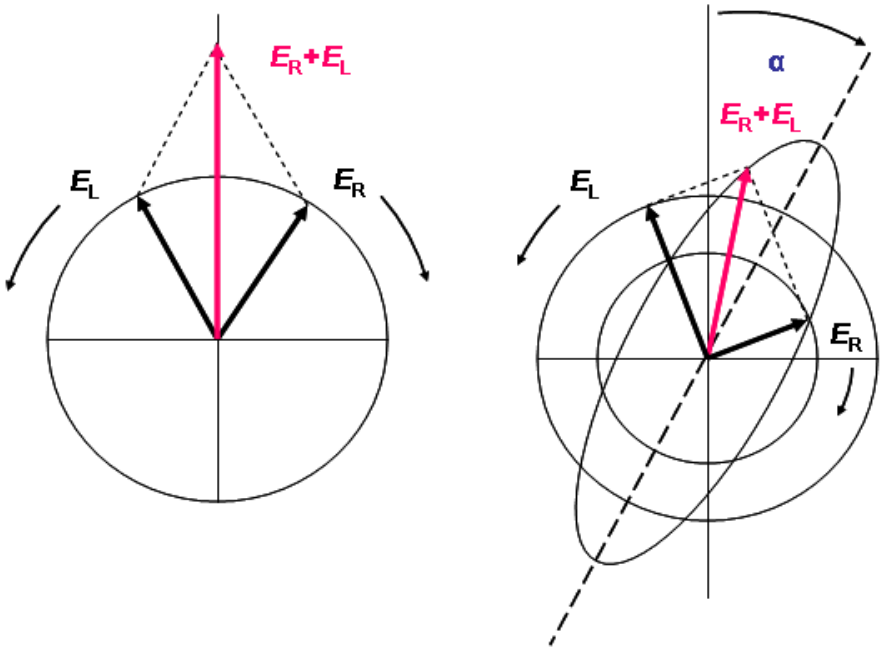


圖 2-3. 橢圓偏極光示意圖

在已知的波長下，光學活性物質對於左右圓偏極光的吸收不同，即 $\Delta A = A_L - A_R$，其中 ΔA 為兩圓偏極光吸收間的差。根據比爾定律 (Beer's Law)，$\Delta A = (\varepsilon_L - \varepsilon_R)*C*l$，其中 ε_L 、ε_R 為莫耳消光係數，C 是莫耳濃度，*l* 是光通過路徑長(cm)。因此，莫耳圓二色性 $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$。莫耳圓二色性與莫耳橢圓率可經由公式推導有 $[\Theta] = 3298\Delta\varepsilon$ 之關係。Θ 為橢圓率，tanΘ = (橢圓長軸-橢圓短軸)。測量不同波長下的 Θ 值或 $\Delta\varepsilon$ 與波長 λ 之間的關係曲線，即為圓二色光譜曲線。

2-2-2 基質輔助電射脫附游離飛行時間質譜法(Matrix Assisted Laser Desoprtion/Ionization Time-of-Flight Mass Spectrometry，MALDI-TOF-MS)

基質輔助雷射脫附游離法是基質將吸收的雷射能量傳遞給分析物，使得分析物進行脫附游離而被偵測到的一種游離方式。由於雷射光的能量是施加在基質上而非分析物，所以可得到完整的分子離子訊號，屬於軟性的游離法。

圖 2-4 所示為 MALDI / TOF 儀器構造示意圖。在進行樣品製備時，大致上的流程為先確認分析物為液體狀，再準備具有吸收雷射能量特性的基質溶液，將兩者以適當比例混合後，取適量點在探針（target）上，經由真空或是大氣的情況下使溶劑揮發，也就是讓兩者形成共結晶，接下來送進 MALDI / TOF 游離源處，在真空下利用氮氣雷射使其進行脫附游離後形成分子離子。接著施加電場在探針上，以給予這些分子離子朝 TOF 管飛行的力量，在經由 TOF 管分離質量差異後，到達偵測器被偵測到。此外，分析物分子若被雷射照射過久會產生熱裂解的現象，所以一般使用具 nanosecond 範圍脈衝頻寬（pulse width）的雷射，且雷射光的聚焦直徑在 10 ~ 300 μm 間，使得 MALDI 分析物分子所感受到的能量約在 10^6 ~ 10^7 W/cm^2 之間。目前較常被使用的 MALDI UV 雷射中，由於 N_2 雷射（波長為 337nm）的價格較低，所以被廣泛的使用。

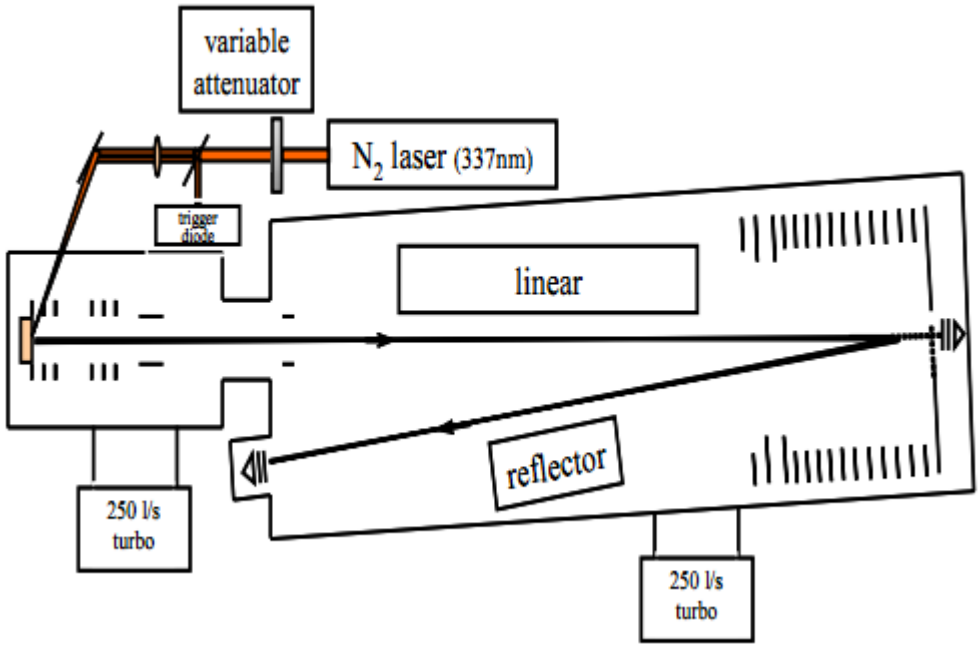


圖 2-4、基質輔助雷射脫附游離飛行時間質譜儀（MALDI - TOF - MS）儀器構造圖

2-3 實驗步驟及儀器操作設定

2-3-1 圓二色光譜

用 5 mM sodium cacodylate buffe r(pH 6.5)分別將 U4A 和 L4T 單股 DNA 溶解，7000 rpm 離心 30 sec，靜置 5 min，待 DNA 完全溶解後，將兩種 DNA 溶液混合均勻，在 94˚C 下加熱 5 min，接著靜置回到室溫，成為 U4A-L4T 雙股 DNA 溶液。將冷卻後的 U4A-L4T DNA 溶液離心（7000 rpm 離心 30 sec），稀釋成 2 μM U4A-L4T DNA 溶液，之後與實驗室合成的胜肽進行反應。將實驗室所合成之胜肽依實驗需要稀釋成不同濃度，取 1 mL U4A-L4T DNA 溶液與 1 mL 胜肽溶液等體積混合，在 37˚C 下反應，不含 CLB 之胜肽反應 1 hr，含 CLB 之胜肽反應 30 min，最後測 CD。

將 CD J-815（Jasco）設定於 High Sensitivity，掃描範圍設在 220 nm ~ 380 nm，Data Pitch 1nm，Continuous Scanning Mode，Scanning Speed 100 nm / min，Response 4 sec，Accumulation 3 times。將 2 mL 樣品加入 3 mL 石英 cell（1 cm × 1 cm）中，在 37˚C 下進行偵測，先分別偵測 DNA only 和各濃度的 peptides only，之後測試 DNA-pepdite 混合溶液之圖譜，最後將所得之數據扣去 DNA only 和 peptides only 的數值，並用 SigmaPlot 10.0 軟體處理之。

2-3-2 Maldi-tof Mass Spectroscopy

用 5 mM sodium cacodylate buffer （pH 6.5）分別將 U4A 和 L4T 單股 DNA 溶解，7000 rpm 離心 30 sec，靜置 5 min，待 DNA 完全溶解後，將兩種 DNA 溶液混合均勻，在 94˚C 下加熱 5 min，接著靜置回到室溫，成為 U4A-L4T 雙股 DNA 溶液。將冷卻後的 U4A-L4T DNA 溶液離心（7000 rpm 離心 30 sec），稀釋成 40 μM U4A-L4T DNA 溶液。取 20 μL 40 μM U4A-L4T duplex DNA 溶液和 20 μL 4μM peptide 溶液混合，10,000 rpm 離心 10 sec，37˚C 反應 30 min。之後再加入 1 mL 0˚C 99% 乙醇和 1.3 μL 5 M NaOAc，10,000 rpm 離心 10 sec，制於-20˚C 冰箱 overnight。4˚C 14,000 rpm 離心 30 min，移除乙醇並且再加入 1 mL 新的 0˚C 99% 乙醇，4˚C 14,000 rpm 離心 30 min，重複上述步驟兩次，最候用 speedvac 800 rpm 25˚C 5 min 將樣品抽乾。

用甲醇配製 0.5 M trihydroxyacetophenone（THAP，Acros）1 mL 和 400 μL 0.5 M diammonium hydrogen citrate（AC，Showa）水溶液，將 THAP：AC ＝ 5：2 的比例混合成為實驗用之基質。將基質與樣品 1:1 混合後點在樣品盤上，待其乾燥後即可使用 MALDI-TOF MS 檢測。

本文使用之基質輔助雷射脫附游離飛行時間質譜儀裝置型號是 microflex（Bruker），具備一波長為 337 nm 的氮氣脈衝式雷射，飛行模式分為直線型與反射型，飛行管長分別為 1.05 m 和 1.96 m，各自有一個偵測器進行測量，本研究皆使用直線型模式操作，操作參數如下：Laser Frequency 20Hz，Ion Source 1 19 kv，Ion Source 2 17.85 kv，Len 8.5 kv，Reflector 20 kv，Pulsed Ion Extraction 400 ns，Number of shots 150，Laser Energy(μJ/pulse) 29~32。

第三章　結果與討論

3-1 圓二色光譜(CD)之研究

本文利用 CD 進行實驗，圓二色光譜對於判定 DNA 的 2 級結構變化是方便且簡易操作的，因此本實驗小組選定數種胜肽，並改變胜肽濃度觀測 DNA 結構上的變化。

3-1-1 結合型胜肽之 CD 結果

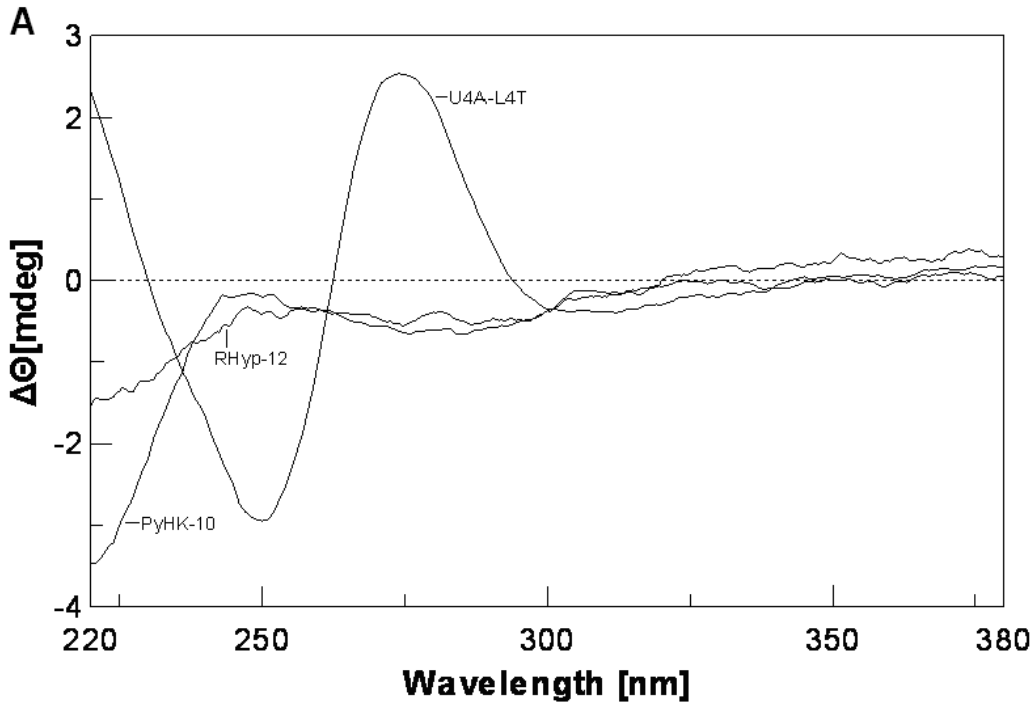


圖 3-1-1A. 1 μM U4A-L4T only、5 μM RHyp-12 only 和 5 μM PyHK-10 only 溶於 5 mM sodium cacodylate buffe r（pH 6.5），37˚C 下之 CD 圖譜。U4A-L4T 在波長 274 nm 有一個向上吸收峰，在波長 251 nm 有一個向下吸收峰，胜肽 RHyp-12 在波長 283 nm 有一個向下吸收峰，PyHK-10 分別在波長 275 nm 和 286 nm 各有一個向下吸收峰。

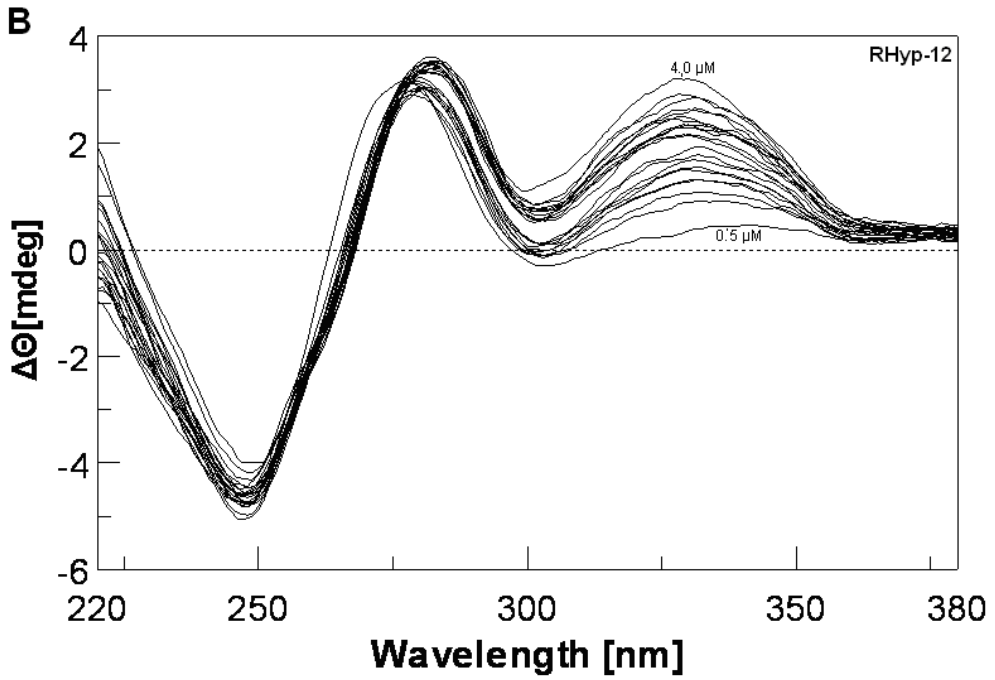


圖. 3-1-1B RHyp-12 與 U4A-L4T DNA 結合之 CD 圖譜。

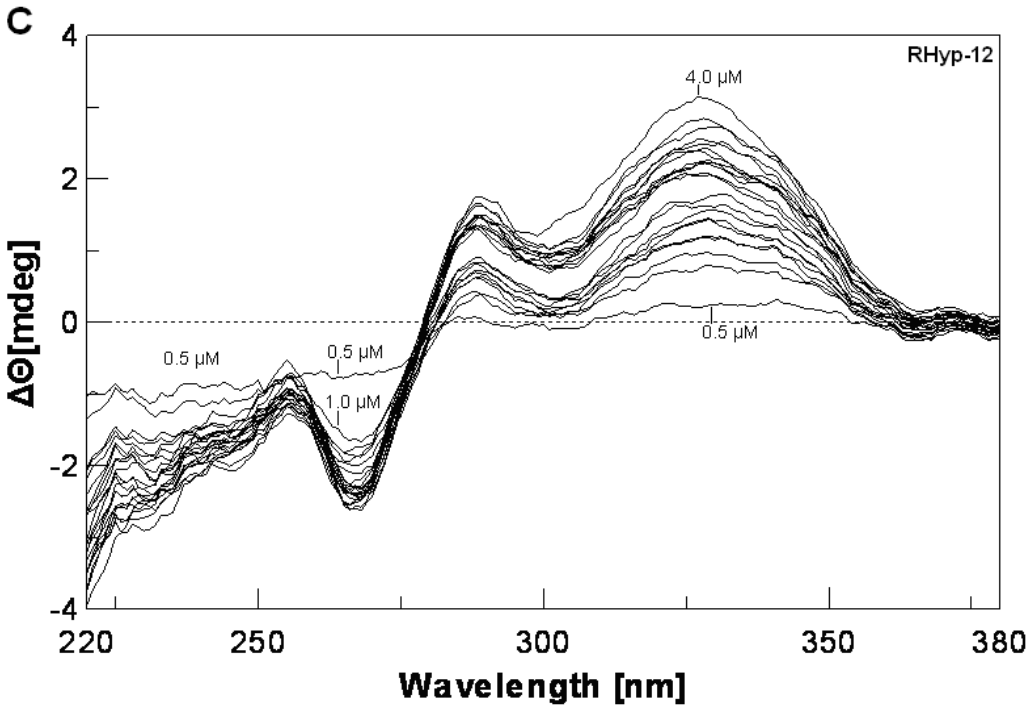


圖 3-1-1C. RHyp-12 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

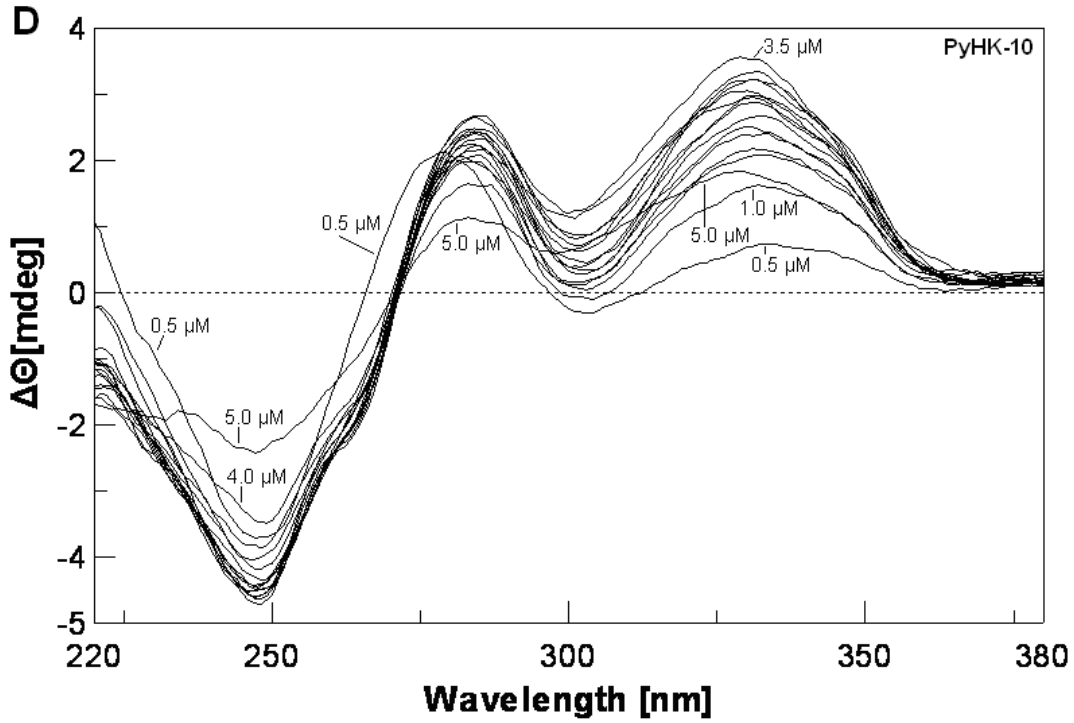


圖 3-1-1D. PyHK-10 與 U4A-L4T DNA 結合之 CD 圖譜。

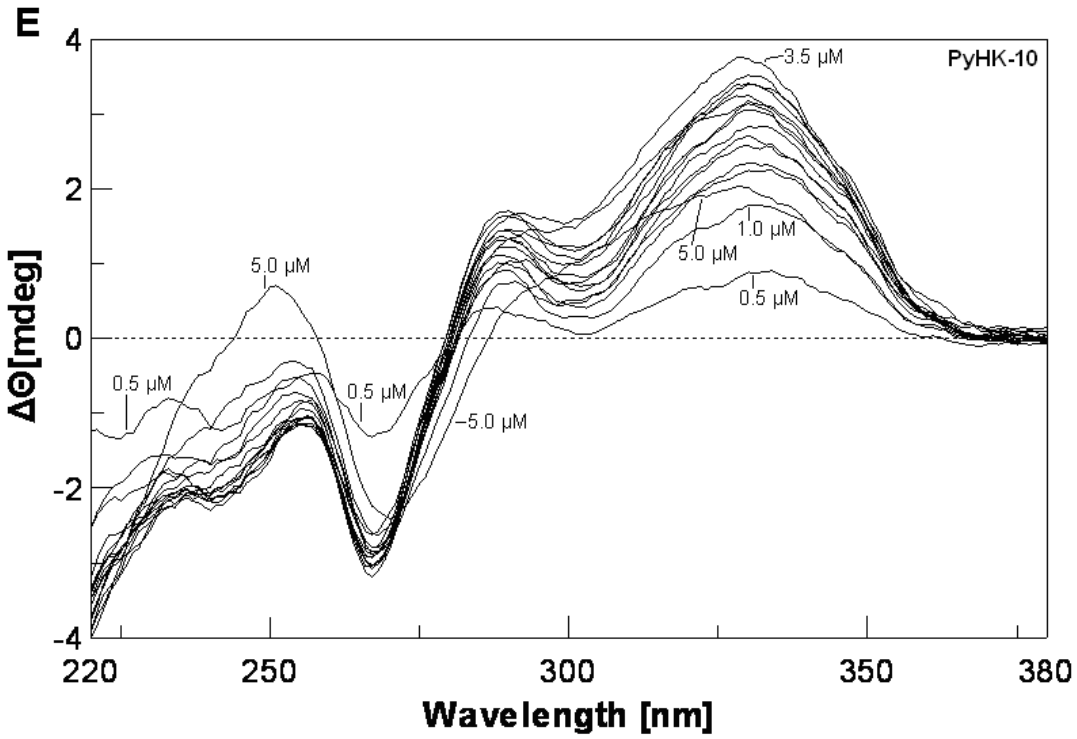


圖 3-1-1E. PyHK-10 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

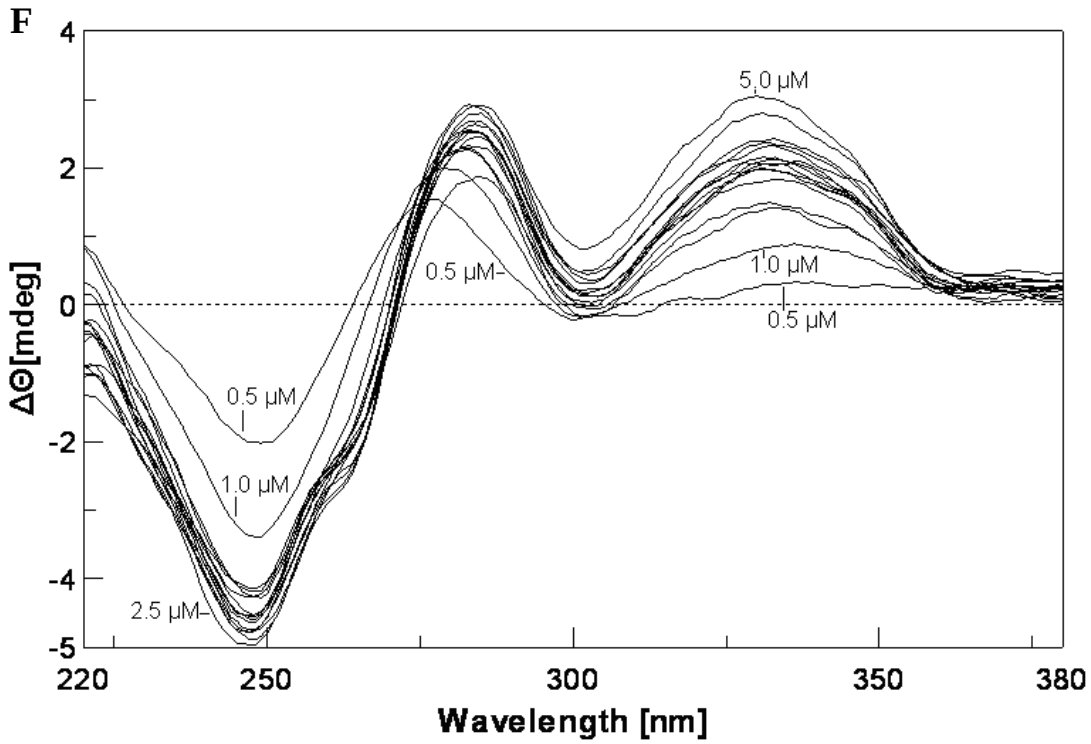


圖 3-1-1F HyH-10 與 U4A-L4T DNA 結合之 CD 圖譜。

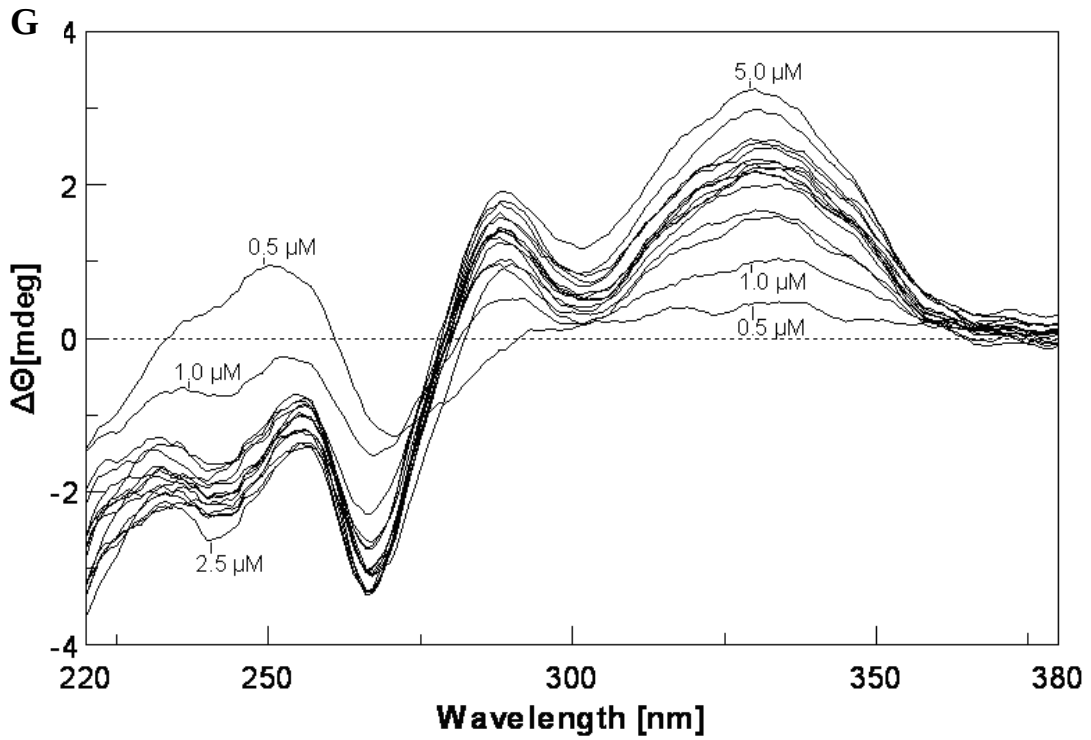


圖 3-1-1G HyH-10 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

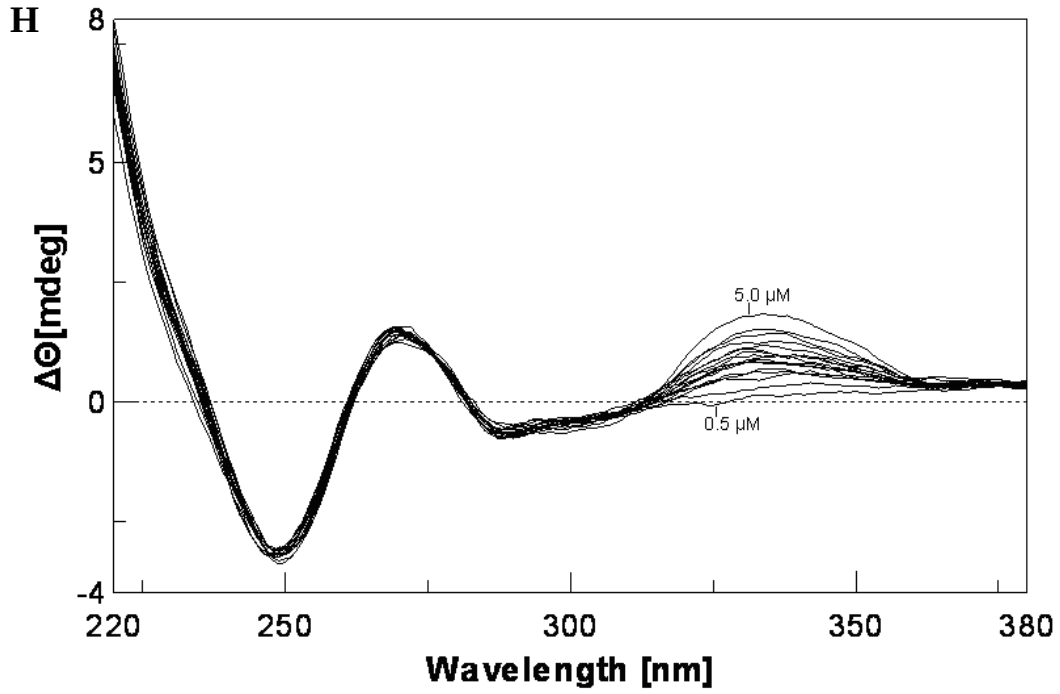


圖 3-1-1H PyHR-9 與 U4A-L4T DNA 結合之 CD 圖譜。

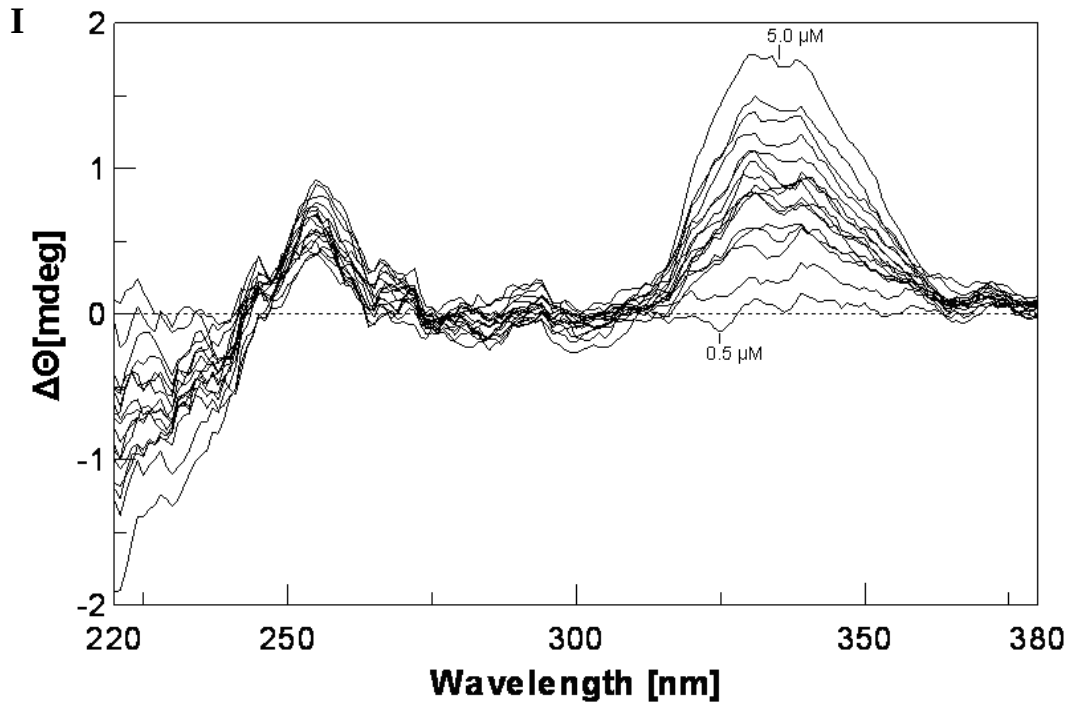


圖 3-1-1I PyHR-9 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

本文從歷年來的足跡法實驗中[31,32]胜肽對 pBR322 DNA 選擇性結合位置 103-114 的序列設計了 U4A-L4T DNA，U4A-L4T DNA 是一有 13 個鹼基對之雙股螺旋 DNA，其序列為 5’-d（TAGGAGAAAATAC）-d（GTATTTTCTCCTA）-3，在序列中包含了 5’-d（AAAA）-d（TTTT）-3’ 的 peptide-DNA binding site。

在 RHyp-12 對 U4A-L4T 的 CD 圖譜中，在波長 247 nm 有一負吸收峰，而在波長 282 nm 和波長 328 nm 則有兩個正吸收峰（圖. 3-1-1B），在其 difference spectra 發現負吸收峰發生了紅位移（red-shifted），吸收波長從 247 nm 變成 269 nm，而兩個正吸收峰則大約在波長 289 nm 和波長 327 nm（圖. 3-1-1C），有一 near-isoelliptic point 發生在波長 280 nm 時，這證明了 2 分子的結合到 DNA 的過程，而在波長 320 nm 附近有正吸收峰產生時，就表示有分子結合在 DNA 的小凹槽（minor groove），顯示 RHyp-12 和實驗室其他胜肽一樣結合在小凹槽。

在 PyHK-10 對 U4A-L4T 的 CD 圖譜中，在波長 248 nm 有一負吸收峰，而在波長 284 nm 一個正吸收峰（圖. 3-1-1D），而在波長 322 nm 則有一個和濃度變化而改變的正吸收峰，在其 difference spectra（圖. 3-1-1E）發現，波長 322 nm 的吸收峰移動到了 329 nm，導致波長 284 nm 的吸收峰位移到 289 nm，波長 322 nm 吸收峰的波長在隨著 PyHK-10 的濃度變化，從濃度 0.5 μM 到 5.0 μM 此吸收峰產生了藍位移（blue-shifted）波長變成 250 nm，這顯示了此吸收峰是 mostly peptide concentration-dependent，使的在 5.0 μM 時的吸收強度反而比較低，這不尋常的橢圓率偏差可能是高濃度的 PyHK-10 突然改變 U4A-L4T 的構像所產生的。

在 HyH-10 對 U4A-L4T 的 CD 圖譜中，在波長 247 nm 有一負吸收峰，而在波長 284 nm 一個正吸收峰（圖. 3-1-1F），在其 difference spectra （圖. 3-1-1G）發現， 330 nm 位置的吸收峰波長並沒有改變，但波長 284 nm 的吸收峰卻產生了紅位移到 288 nm。

在 PyHR-9 對 U4A-L4T 的 CD 圖譜中，在波長 249 nm 有一負吸收峰，而在波長 270 nm 一個正吸收峰（圖. 3-1-1H），還有一個又寬又淺的吸收峰被誘發在 334 nm 附近，在其 difference spectra 中發現一較弱的 255 nm 正吸收峰和原來就有的 334 nm 吸收峰（圖. 3-1-1I）。

在[peptide] / [DNA] 對 $\Delta\theta$ 的關係中發現（圖. 3-1-2），PyHK-10 橢圓率的變化最劇烈，這顯示 PyHK-10 結合在 U4A-L4T DNA 小凹槽時所造成的所造成的構像變化最強烈；換句話說，PyHR-9 在結合 DNA 後所造成的橢圓率變化比 PyHK-10、RHyp-10 和 HyH-10 弱，就表示當 PyHR-9 結合在 U4A-L4T DNA 小凹槽時所造成的構像變化不大。當[peptide] / [DNA] 比在 0.5 - 2.0 之間時只有 1 分子的胜肽結合在 5’-d（AAAA）-d（TTTT）-3’ 之位置上，而當[peptide] / [DNA] 比在 2.0 - 3.0 之間時，則有 2 分子的胜肽結合在選擇性之序列上，之後

[peptide] / [DNA] 比在 4.0 以上時，$\Delta\theta$ 又再進一步的增加，顯示有更多的分子參與對 DNA 的非序列選擇性結合（non-sequence-selective binding）。

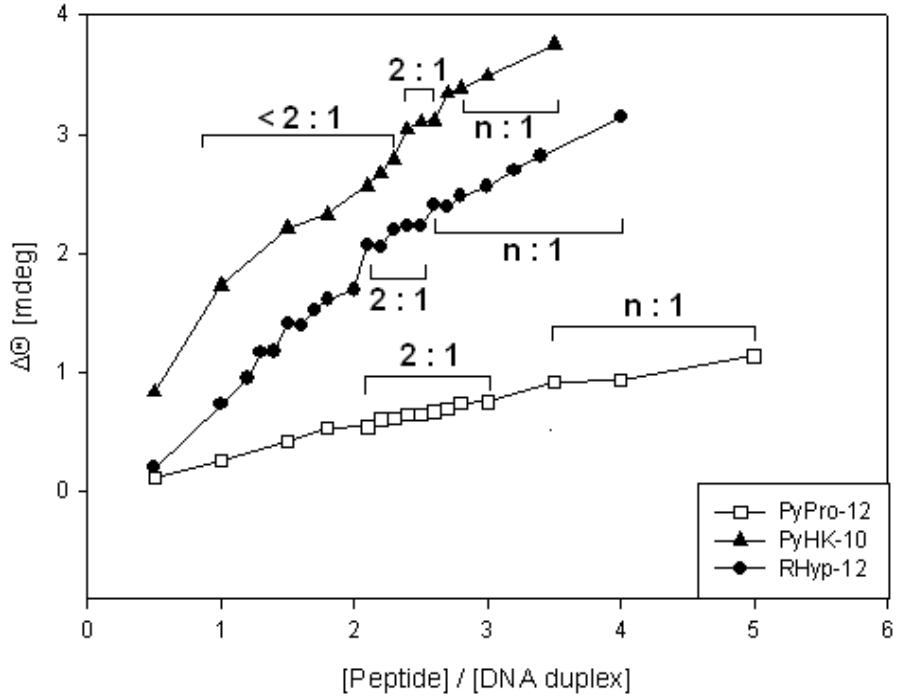


圖 3-1-2. [peptide] / [DNA] 對$\Delta\theta$ 之關係圖。

3-1-2 切割型和結合型胜肽 CD 結果之比較

在序列中含 CLB 之切割型胜肽的 CD 圖譜中，依然可以發現在波長~ 330 nm 代表結合在小凹槽的訊號，因為雖然胜肽切斷了 DNA， 但由於雙股 DNA 之間的氫鍵並未被破壞，所以依舊可以維持一定程度的雙股螺旋 DNA 結構，因此依舊可偵測到橢圓率的變化。

A

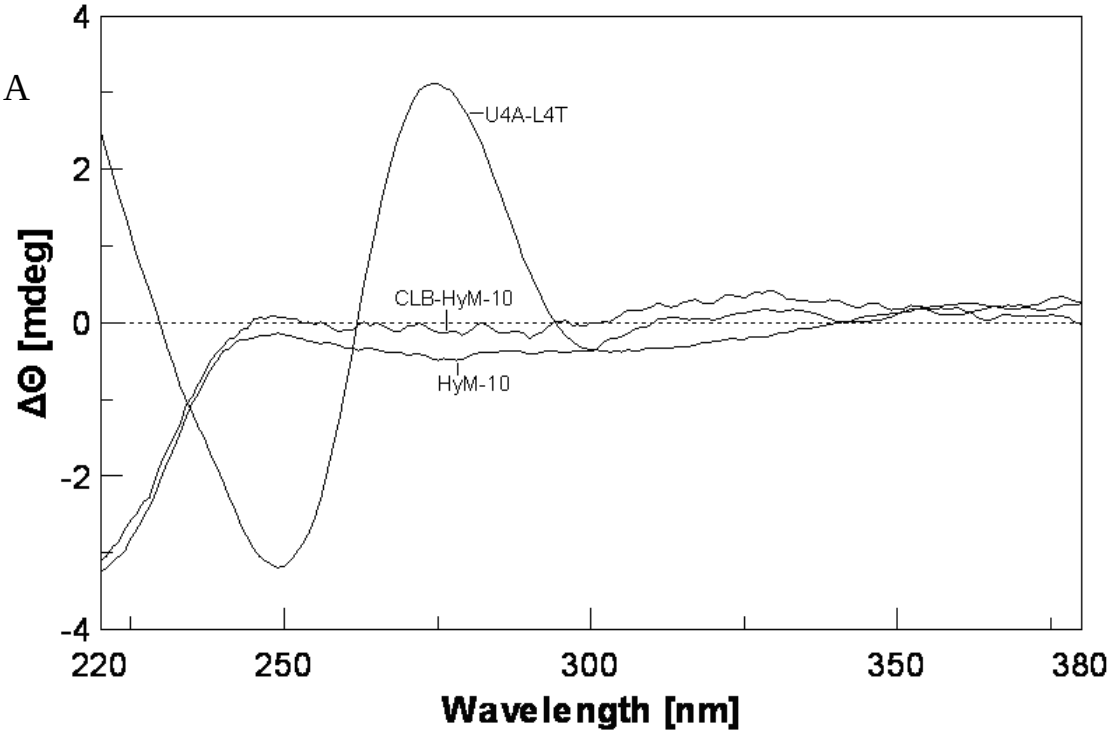


圖 3-1-3A. 1 μM U4A-L4T only、5 μM CLB-HyM-10 only 和 5 μM HyM-10 only 溶於 5 mM sodium cacodylate buffe r(pH 6.5)，37 °C 下之 CD 圖譜。U4A-L4T 在波長 274 nm 有一個向上吸收峰，在波長 251 nm 有一個向下吸收峰，胜肽 CLB-HyM-10 在無明顯吸收峰，HyM-10 分別在波長 275 nm 有一個向下吸收峰。

B

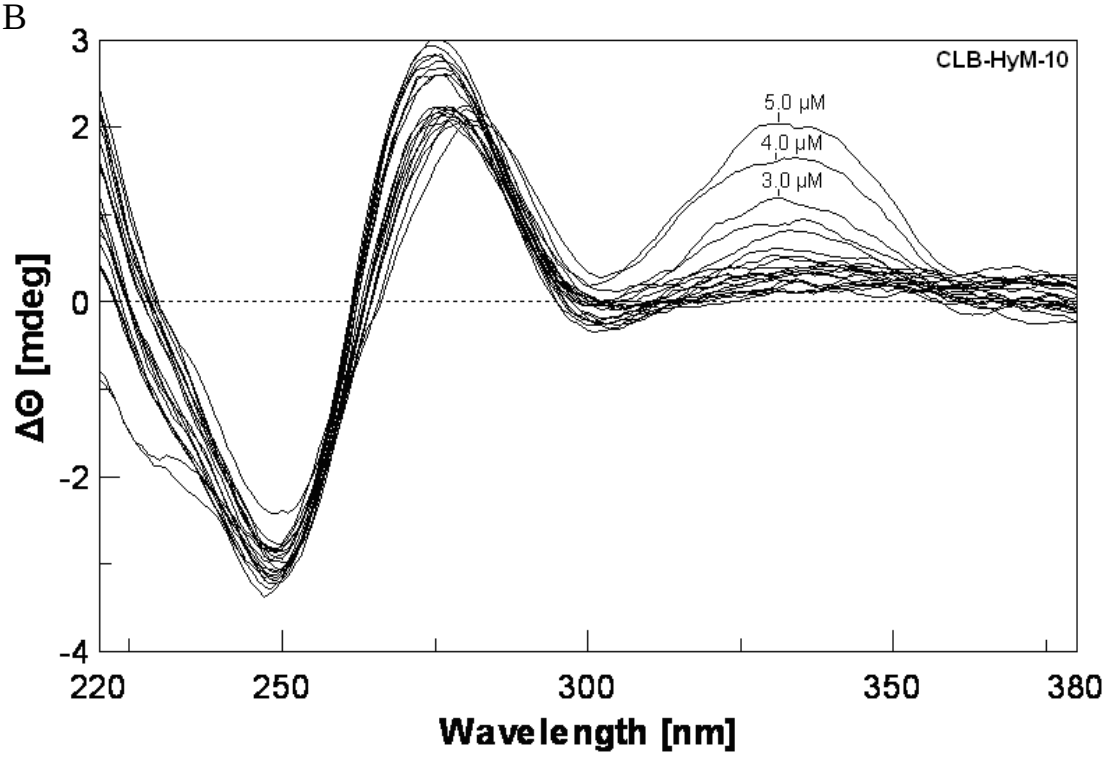


圖 3-1-3B CLB-HyM-10 與 U4A-L4T DNA 結合之 CD 圖譜。

C

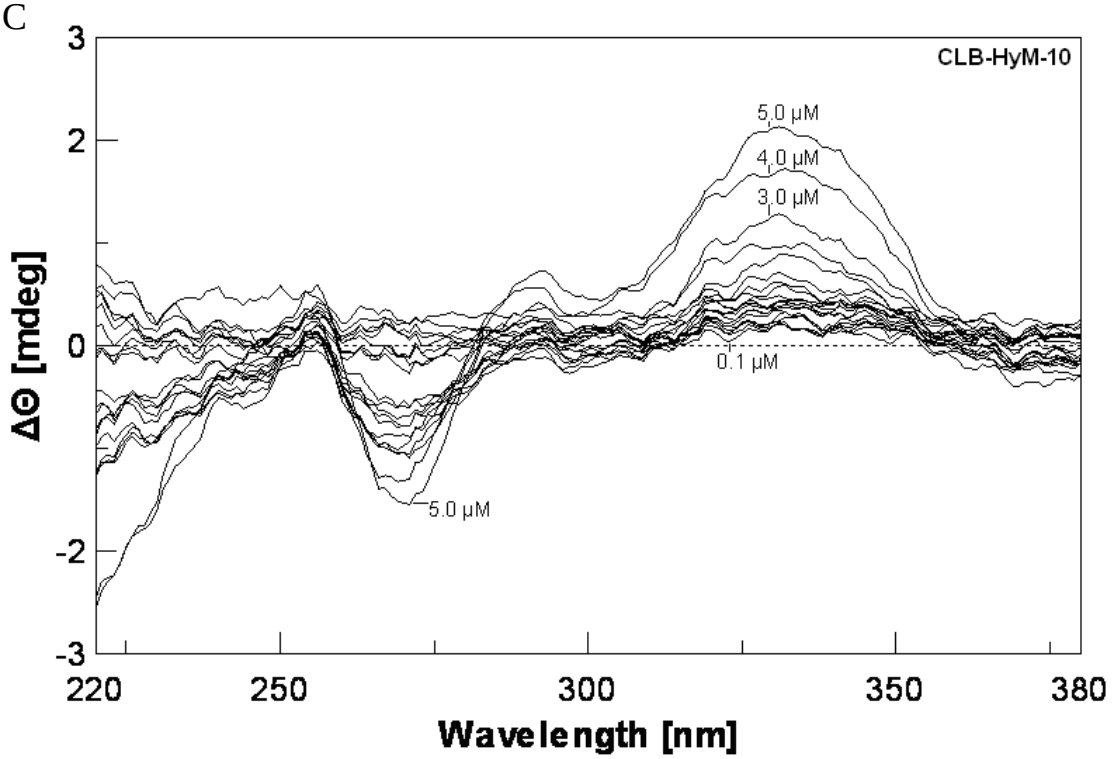


圖 3-1-3C CLB-HyM-10 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

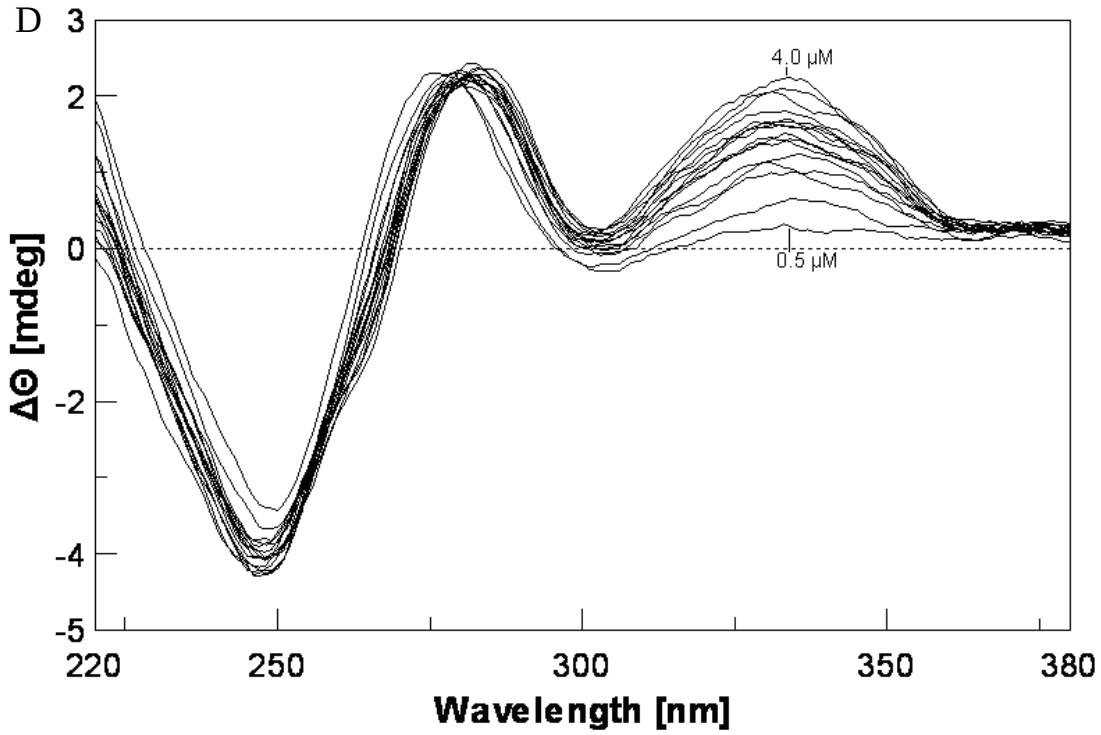


圖 3-1-3D CLB-HyQ-10 與 U4A-L4T DNA 結合之 CD 圖譜。

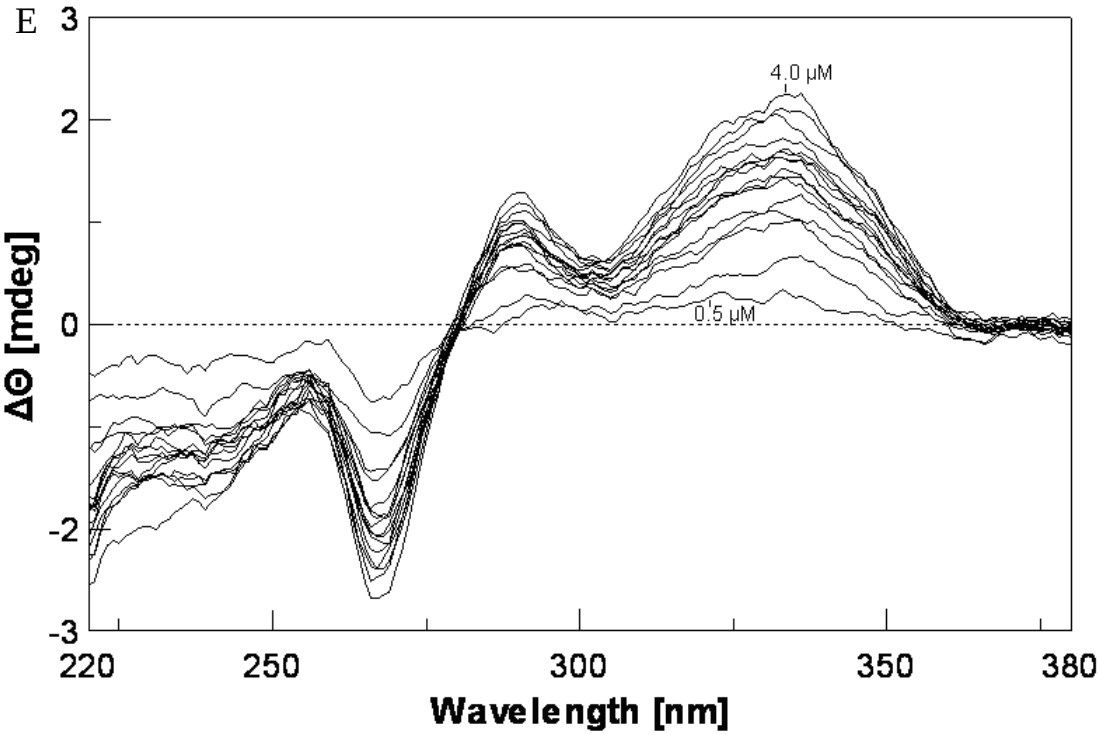


圖 3-1-3E CLB-HyQ-10 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

F

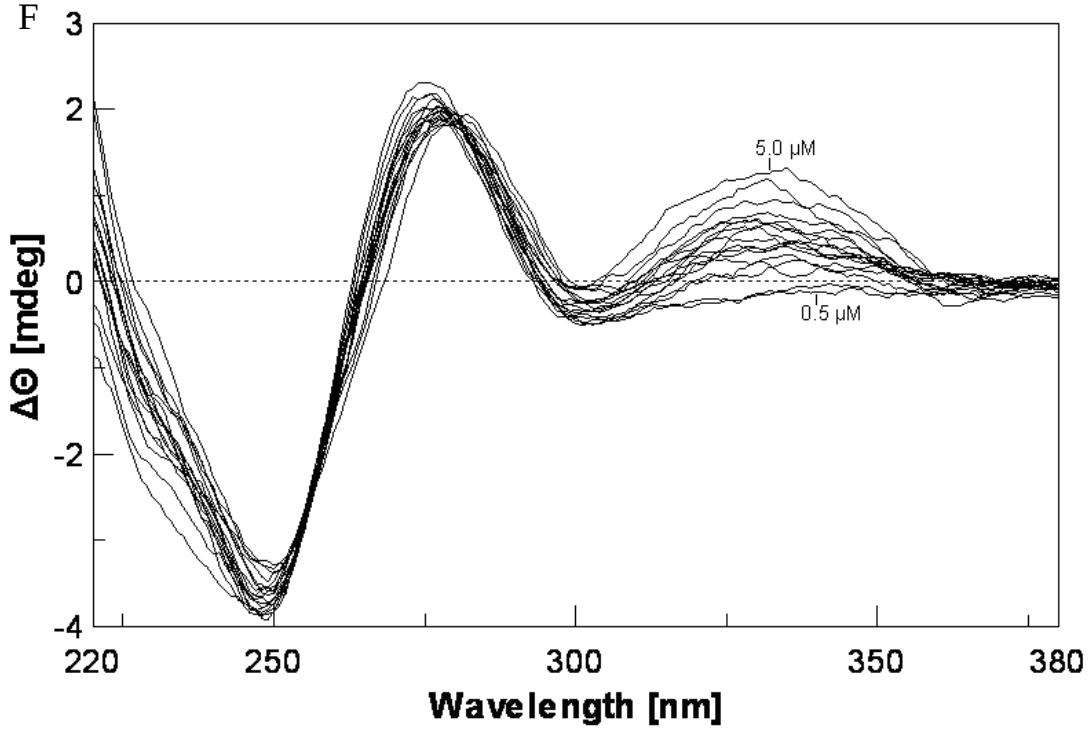


圖 3-1-3F CLB-PyHK-10 與 U4A-L4T DNA 結合之 CD 圖譜。

G

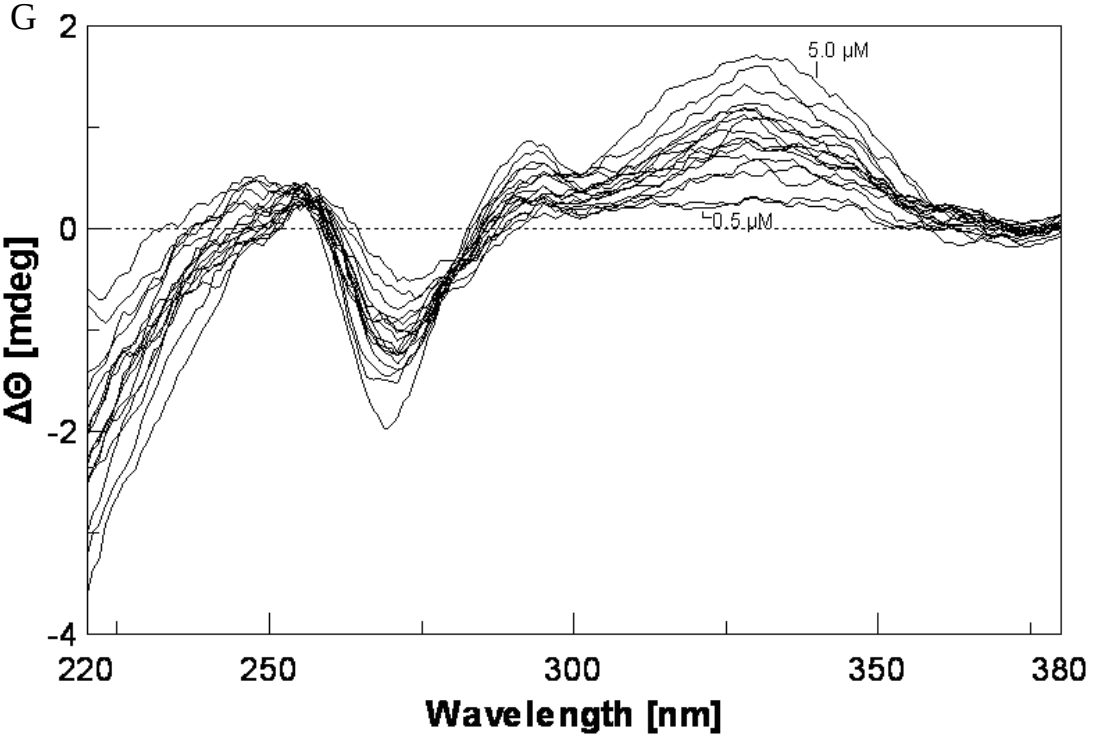


圖 3-1-3G CLB-PyHK -10 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

H

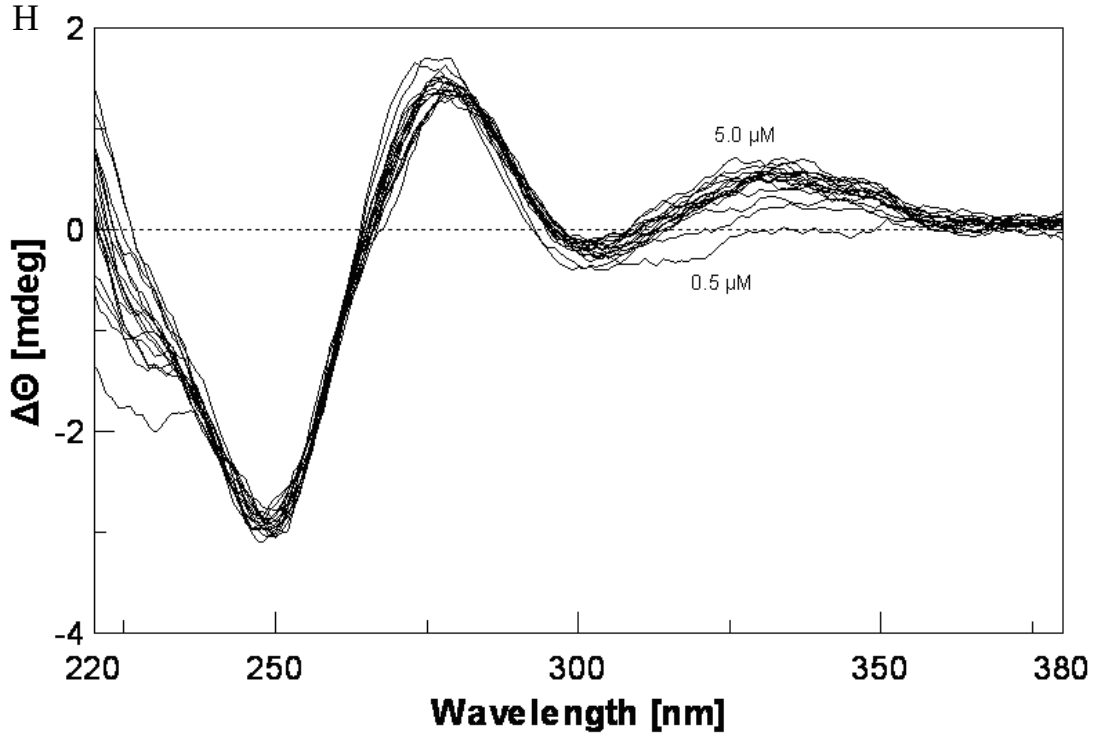


圖 3-1-3H CLB-PyWK -10 與 U4A-L4T DNA 結合之 CD 圖譜。

I

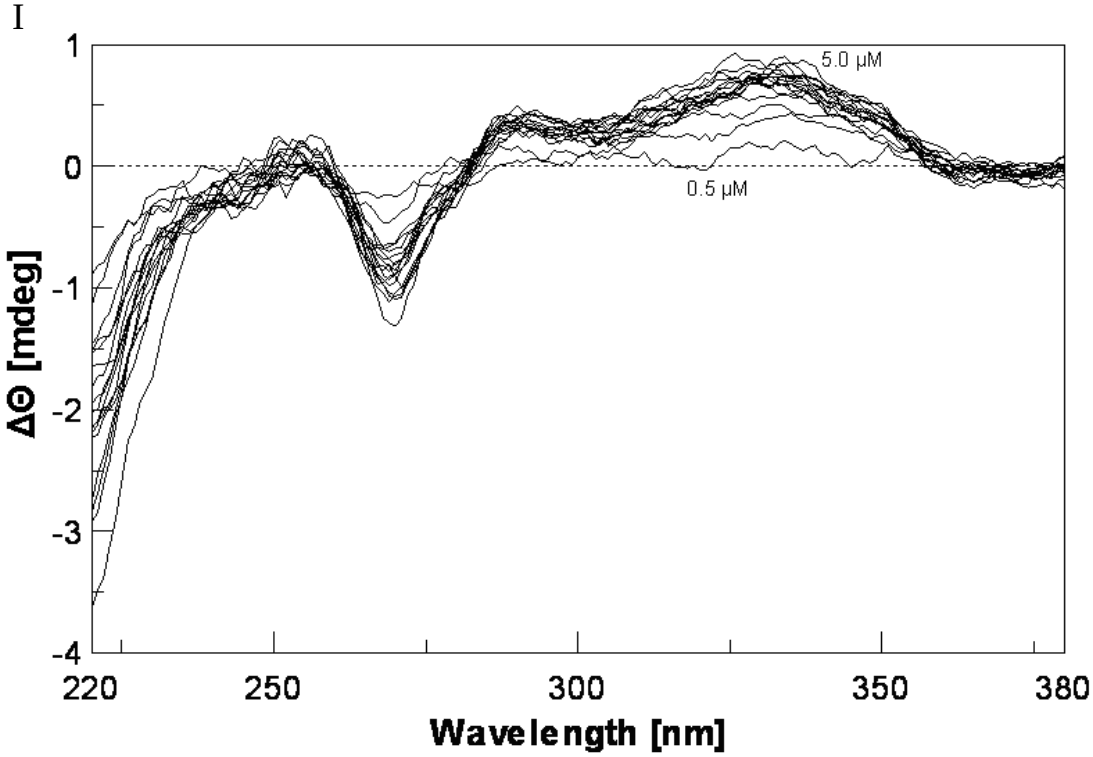


圖 3-1-3I CLB-PyWK -10 與 U4A-L4T DNA 結合之示差光譜 (difference spectra)。

在 CLB-HyM-10 對 U4A-L4T 的 CD 圖譜中，在波長 247 nm 有一負吸收峰，而在波長 278 nm 和波長 337 nm 則有兩個正吸收峰（圖. 3-1-3B），在其 difference spectra 發現負吸收峰消失了，有一吸收波則從正吸收變為負吸收，波長位置不變（圖. 3-1-3C），有一 near-isoelliptic point 發生在波長 287 nm 時，這證明了 2 分子的結合到 DNA 的過程，而在波長 330 nm 附近有正吸收峰產生時，就表示有分子結合在 DNA 的小凹槽（minor groove），顯示切割型和結合型胜肽一樣會結合在小凹槽。

在 CLB-HQ-10 對 U4A-L4T 的 CD 圖譜中，在波長 248 nm 有一負吸收峰，而在波長 280 nm 一個正吸收峰（圖 3-1-3D），而在波長 334 nm 則有一個與濃度變化成正比的正吸收峰，在其 difference spectra（圖. 3-1-3E）發現，波長 280 nm 的正吸收峰變成了 267 nm 的負吸收峰，334 nm 的正吸收峰位置不變。

在 CLB-PyHK-10 對 U4A-L4T 的 CD 圖譜中，在波長 248 nm 有一負吸收峰，而在波長 276 nm 一個正吸收峰（圖 3-1-3F），在其 difference spectra （圖 3-1-3G）發現， 333 nm 位置的吸收峰波長並沒有改變，但波長 276 nm 的吸收峰卻產生了藍位移到 269 nm，而且由正吸收峰變負吸收峰。

在 CLB-PyWK-10 對 U4A-L4T 的 CD 圖譜中，在波長 249 nm 有一負吸收峰，而在波長 278 nm 一個正吸收峰（圖. 3-1-3H），還有一個又寬又淺的吸收峰被誘發在 335 nm 附近，在其 difference spectra 中發現一較弱的 269 nm 負吸收峰和原來就有的 335 nm 吸收峰（圖. 3-1-3I）。

在[peptide] / [DNA] 對$\Delta\theta$ 的關係中發現（圖. 3-1-4ABCD），結合型胜肽橢圓率的變化比切割型劇烈，這顯示結合型胜肽結合在 U4A-L4T DNA 小凹槽時所造成的所造成的構像變化最強烈，而 CLB 結構會對胜肽結合在 U4A-L4T DNA 小凹槽造成負面影響，而被切開的 DNA 亦會使切割型胜肽無法繼續與之結合。橢圓率的變化最劇烈的 HyE-10（圖. 3-1-4A），在接上 CLB 後，由於結合能力強，造成切割的效應大，使 DNA 碎片增加，在 CLB-HyE-10 橢圓率的變化最反而是最小的（圖. 3-1-4B）；橢圓率的變化最小的 HyS-10、HyQ-10（圖. 3-1-4A）也發現相似的桔果，CLB-HyS-10、CLB-HyQ-10（圖. 3-1-4B）橢圓率的變化最反而是最大的。PyHK-10、PyMK-10 結合能力過強（圖. 3-1-4C），至使 PyHK-10 在濃度 3.5 nM，PyMK-10 在濃度 2.5 nM 之後產生非選擇性結合(non-selectivity binding) 扭曲 DNA 構象，遮蔽了 CD 的訊號，而在 CLB-PyHK-10、CLB-PyMK-10 因著切割效應，使得不易產生非選擇性結合，令我們得以觀察到正常的 CD 訊號（圖. 3-1-4D），發現穩定的結合訊號。

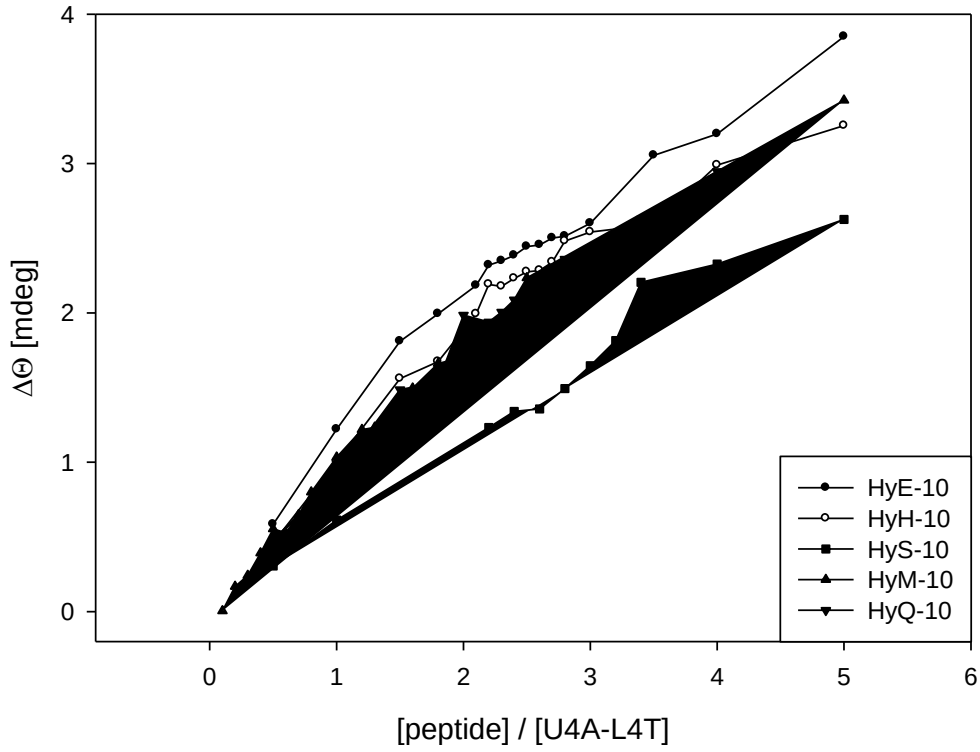


圖 3-1-4A. [peptide] / [DNA] 對Δθ 之關係圖。

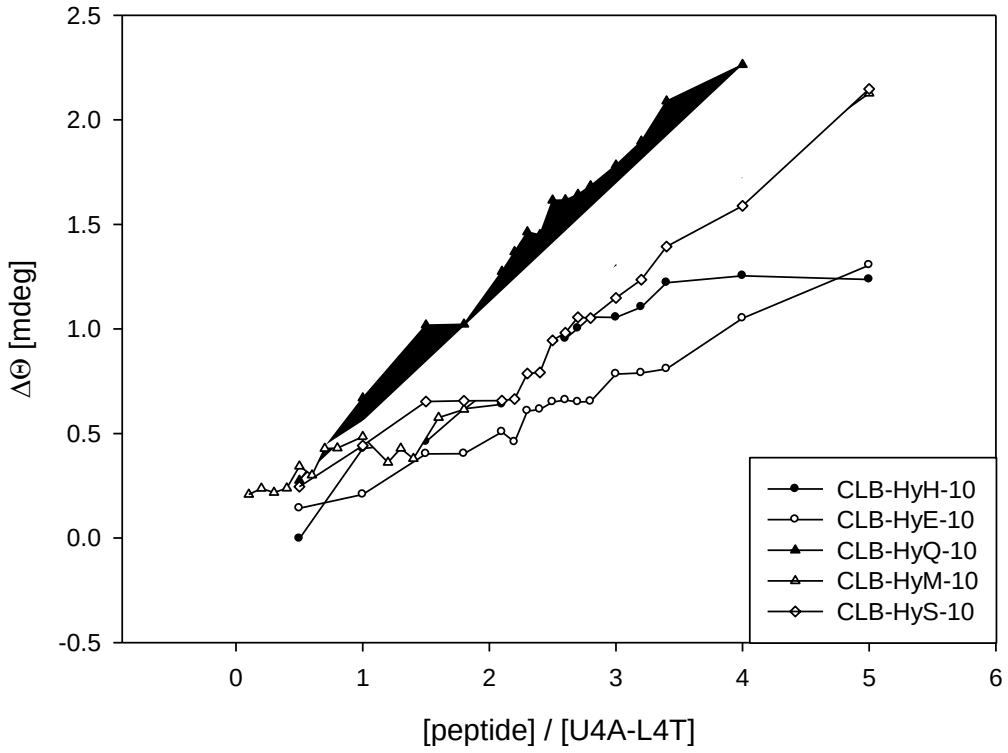


圖 3-1-4B. [peptide] / [DNA] 對Δθ 之關係圖。

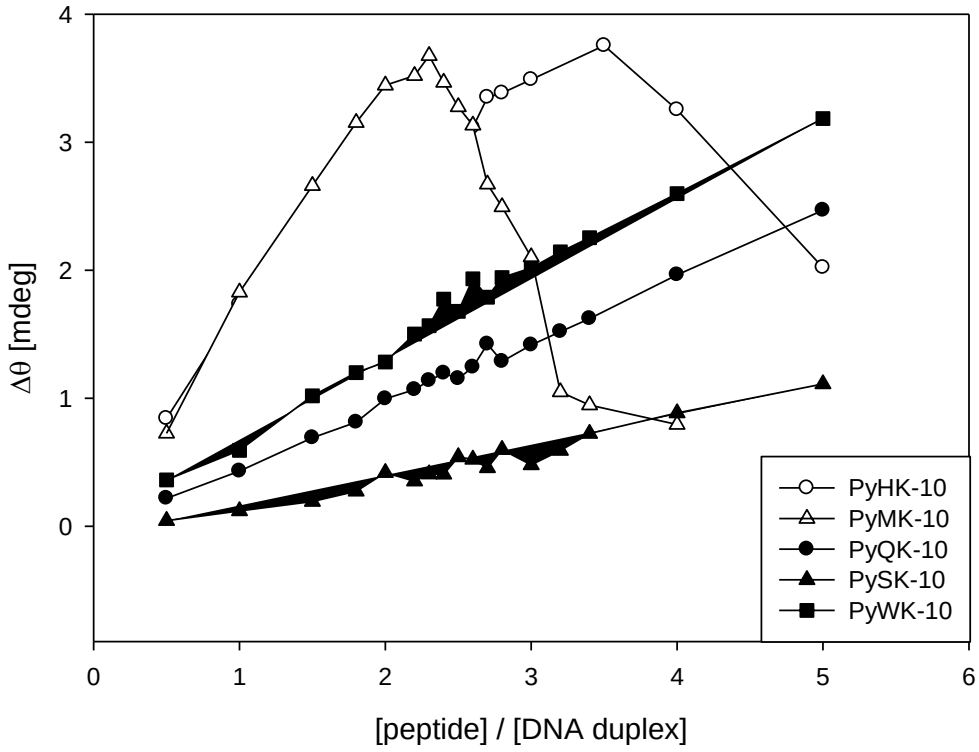


圖 3-1-4C. [peptide] / [DNA] 對$\Delta\theta$ 之關係圖。

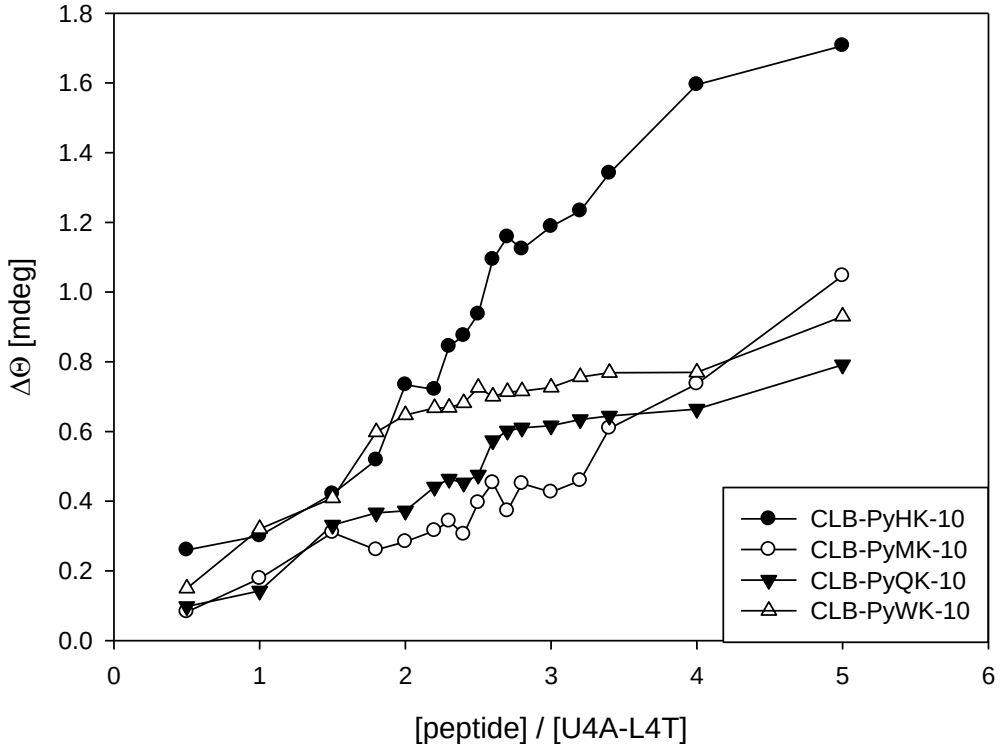


圖 3-1-4D. [peptide] / [DNA] 對$\Delta\theta$ 之關係圖。

3-1-3 胜肽 CD 結果之比較

在結合型胜肽的效果上 PyMK-10 最好，而 PySK-10 最差（圖. 3-1-5A），切割型胜肽則是 CLB-HyQ-10 最好，而 CLB-PyMK-10 最差。

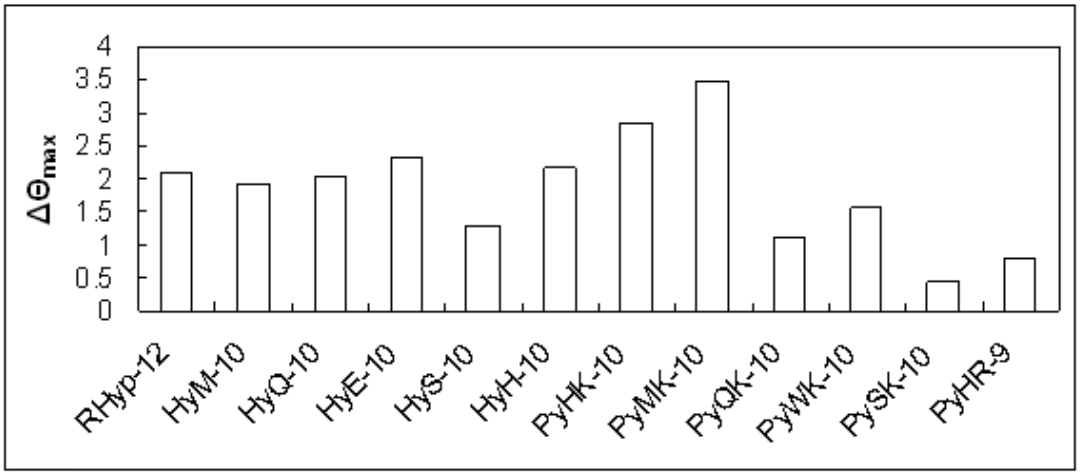


圖 3-1-5A. 結合型胜肽Δθ 之比較圖。

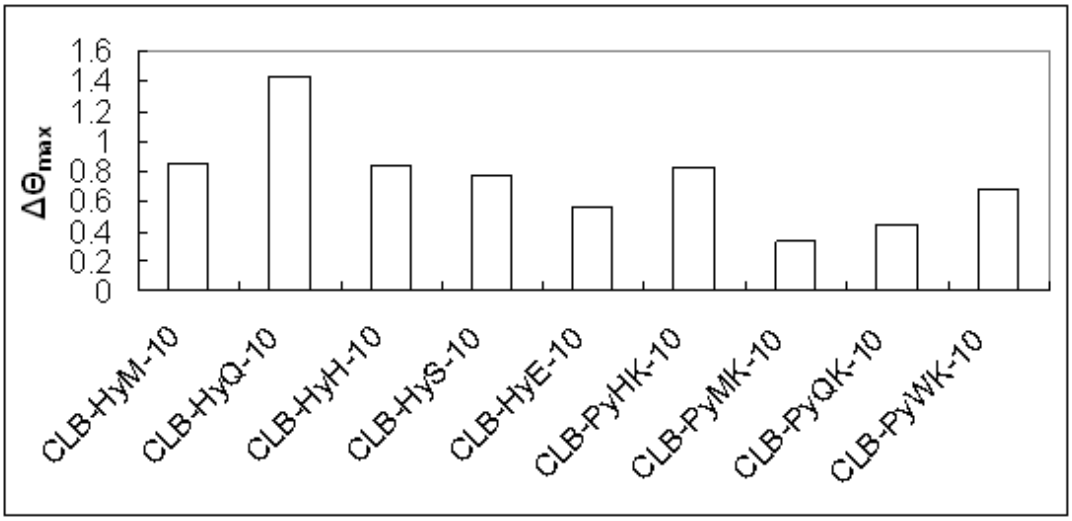


圖 3-1-5B. 切割型胜肽Δθ 之比較圖。

表 1. 各胜肽 $\Delta\Theta_{max}$ 之波長。

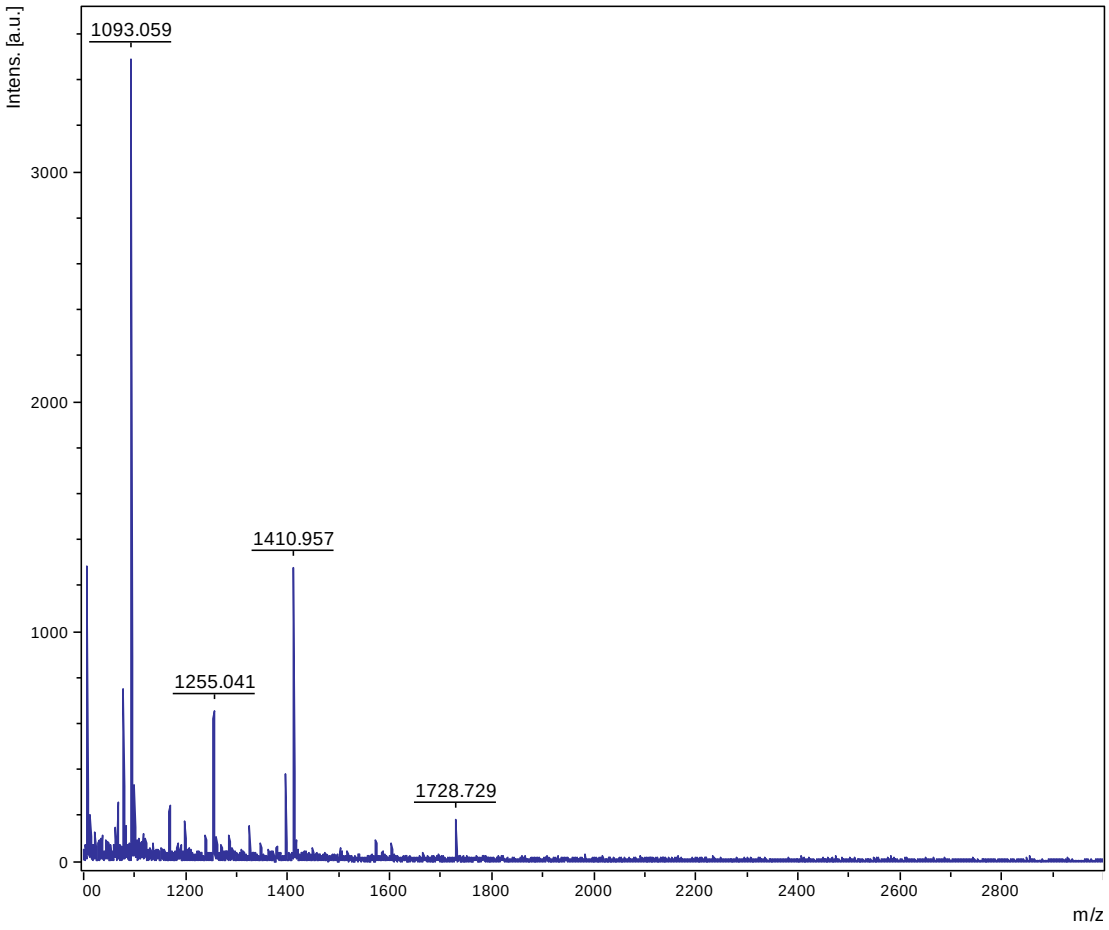
胜肽	$\Delta\Theta_{max}$ (nm)
RHyp-12	320
HyM-10	329
HyQ-10	327
HyE-10	330
HyS-10	334
HyH-10	332
PyHK-10	322
PyMK-10	333
PyQK-10	325
PyWK-10	326
PyHR-9	334
CLB-HyM-10	337
CLB-HyQ-10	334
CLB-HyE-10	333
CLB-HyS-10	330
CLB-HyH-10	321
CLB-PyHK-10	333
CLB-PyMK-10	335
CLB-PyQK-10	325

各胜肽在~330 nm 吸收峰是本文主要探討之光譜位置，以此表列出各胜肽比較十的正確波長。

3-2 MALDI-TOF MS Spectroscopy

含 CLB 之切割型胜肽對 U4A-L4T DNA 切割片段的比對上，本文利用 MALDI-TOF MS 精確得尋找出被胜肽切割後之 DNA 片段分子量，並得知不同胜肽對 DNA 有不同的序列選擇性切割位置。

3-2-1 CLB-HyM-10 MALDI-TOF MS



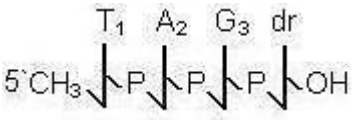
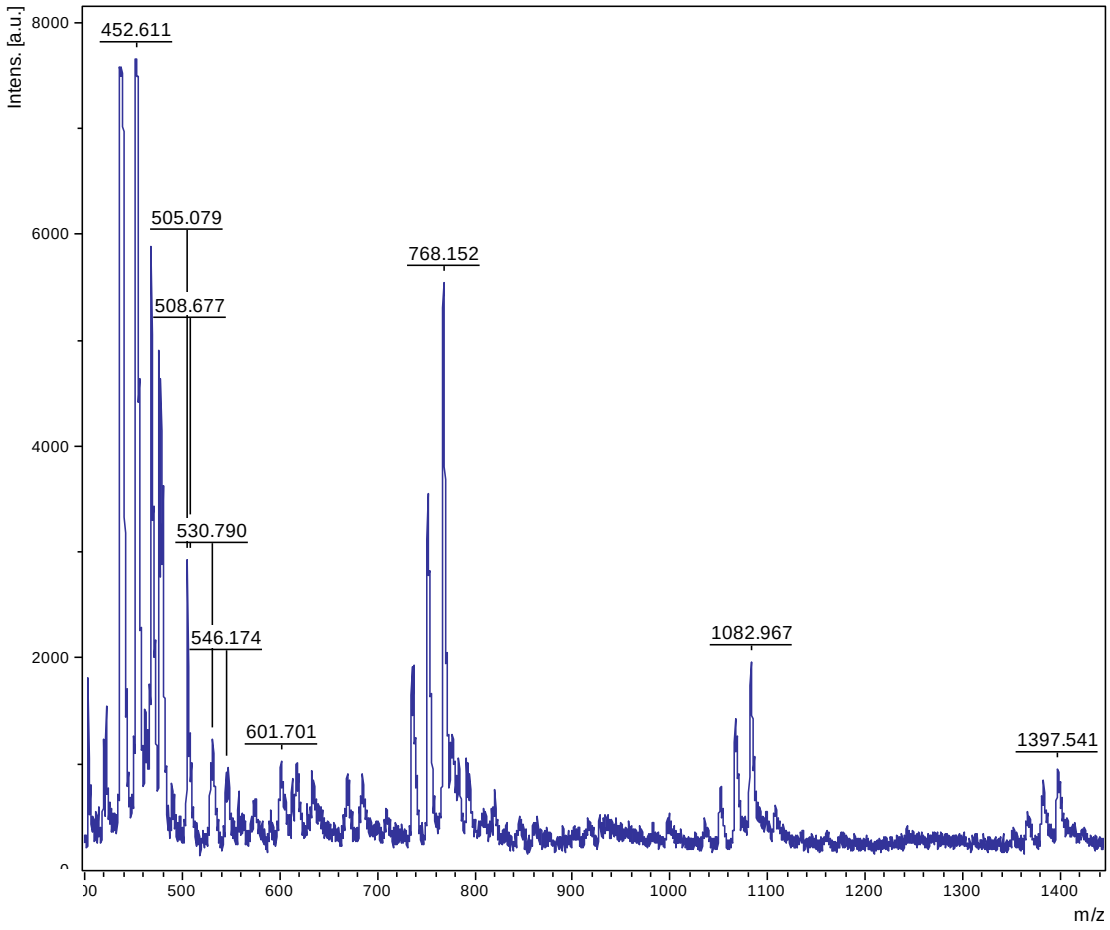
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	1093.72	1093.059

圖 3-2-1A. 胜肽 CLB-HyM-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。



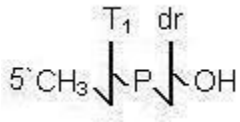
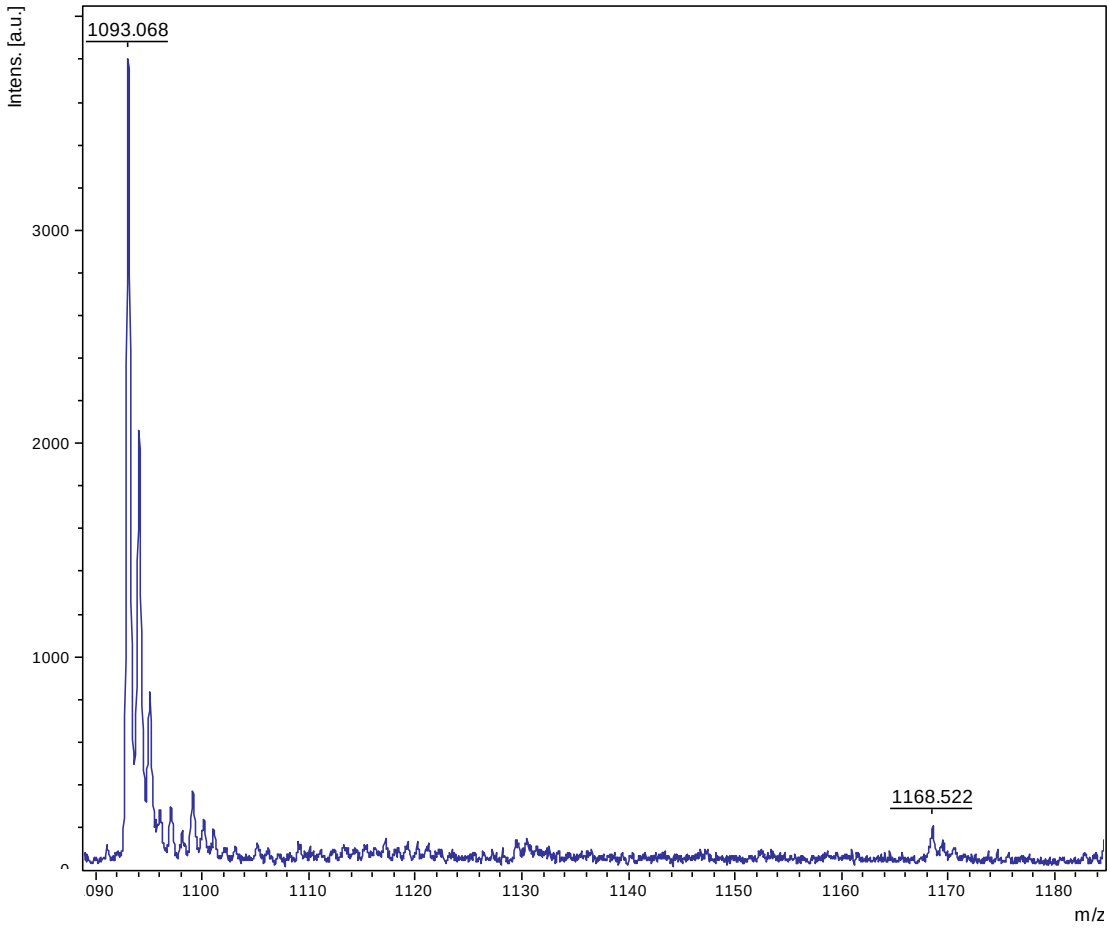
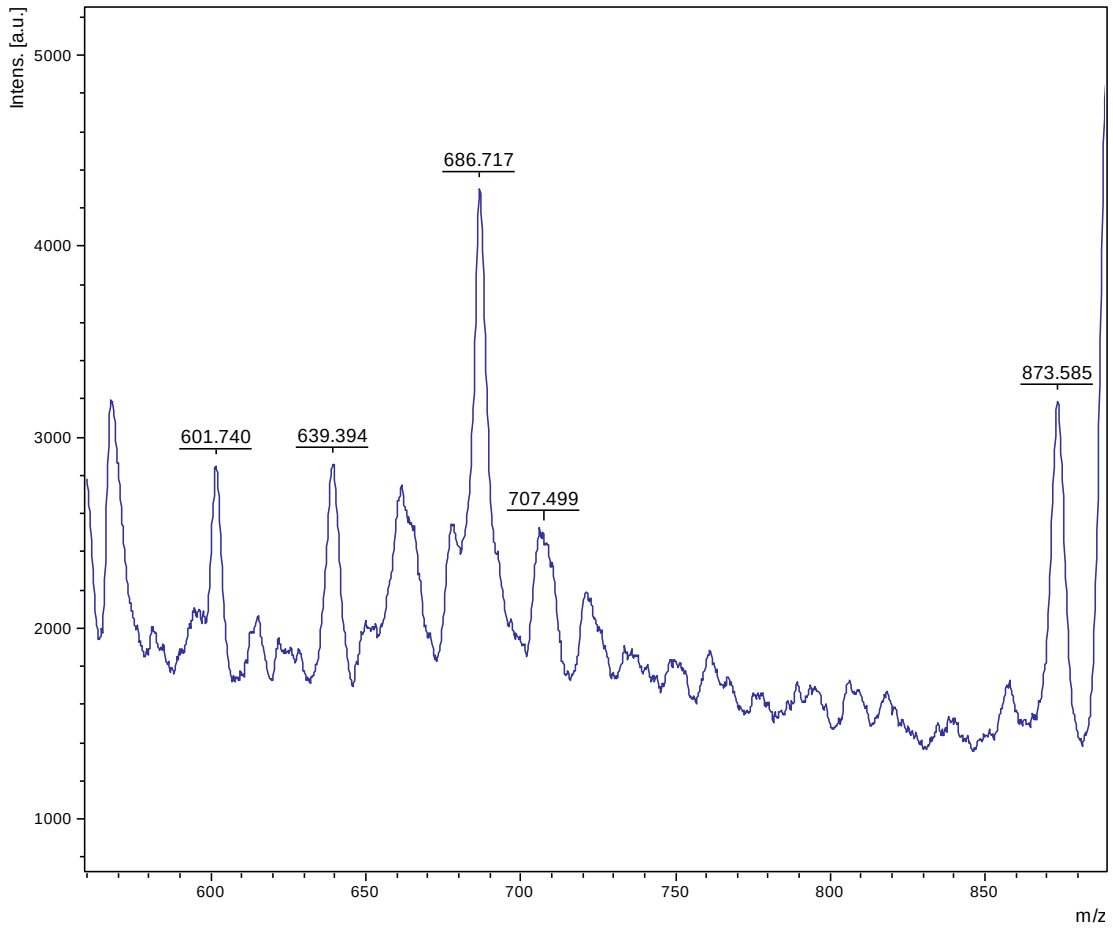
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	451.33	452.611

圖 3-2-B. 胜肽 CLB-HyM-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。



Assigned oligonucleotide fragment	Calcd. $[M-H]^-$	Observed $[M-H]^-$
T_1 A_2 G_3 dr 5`CH_3-P-P-P-OH	1093.72	1093.059

圖 3-2-1C. 胜肽 CLB-HyM-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。



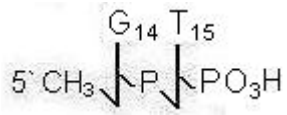
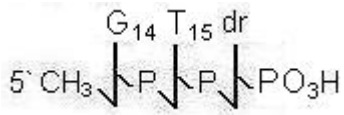
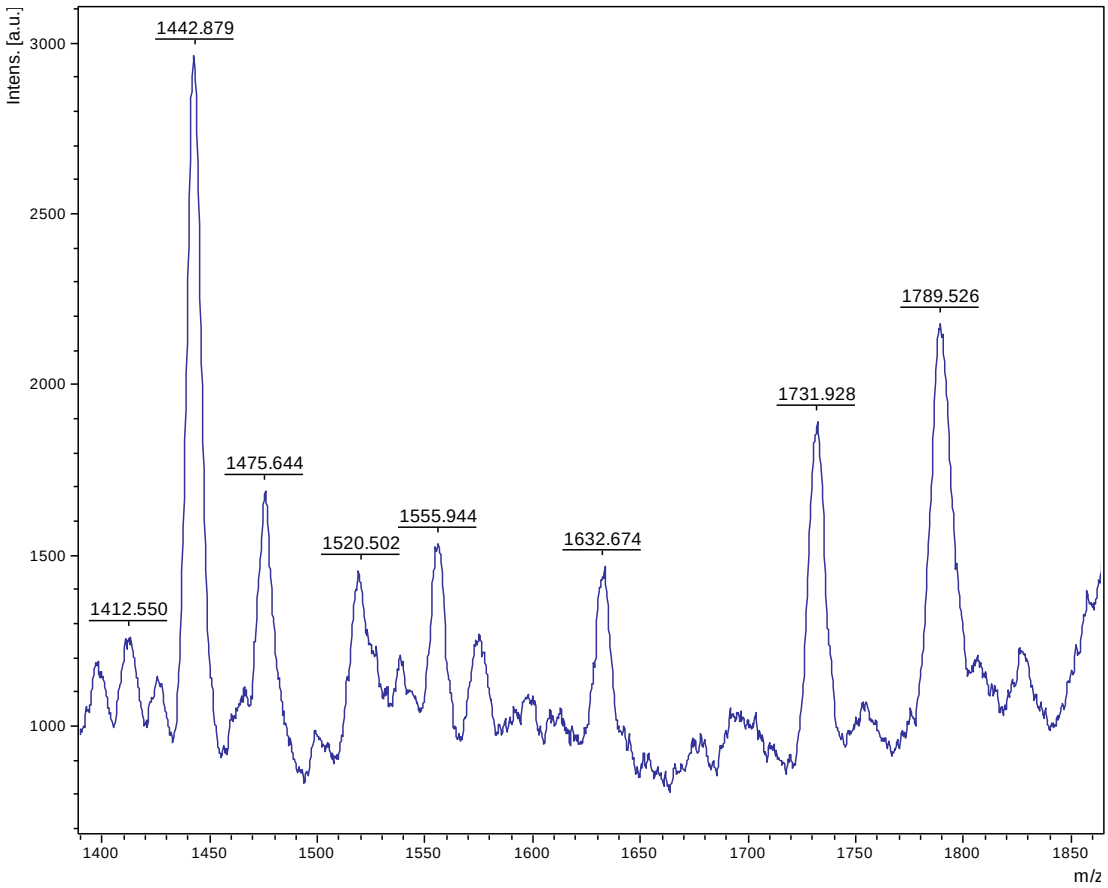
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	686.42 (Na^+)	686.717 (Na^+)
	873.13	873.585

圖 3-2-1D. 胜肽 CLB-HyM-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

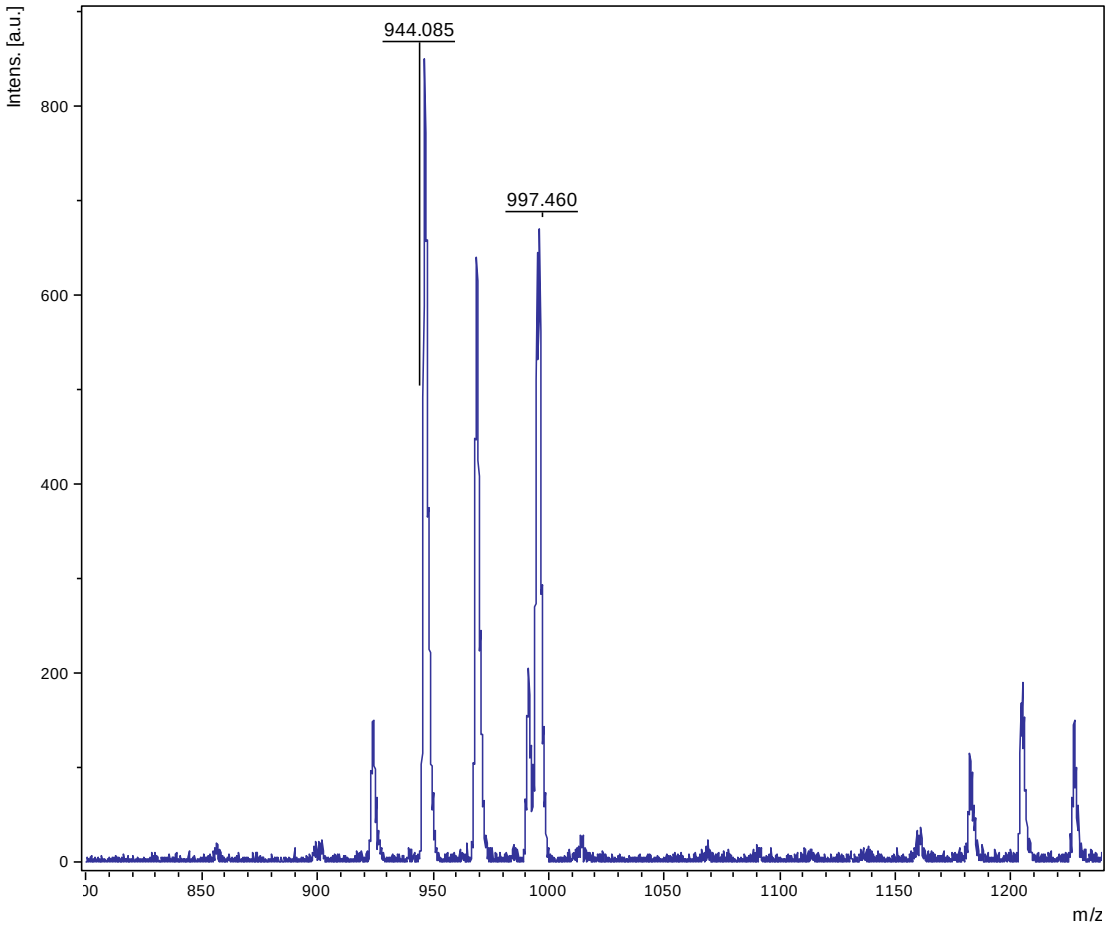


Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
5` CH_3 T_1 A_2 G_3 G_4 dr P P P P OH	1442.12 (Na^+)	1442.879 (Na^+)
HO_3P dr T_{22} C_{23} C_{24} T_{25} A_{26} P P P P P P OH 3`	1731.21 (Na^+)	1731.928 (Na^+)

圖 3-2-1E. 胜肽 CLB-HyM-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

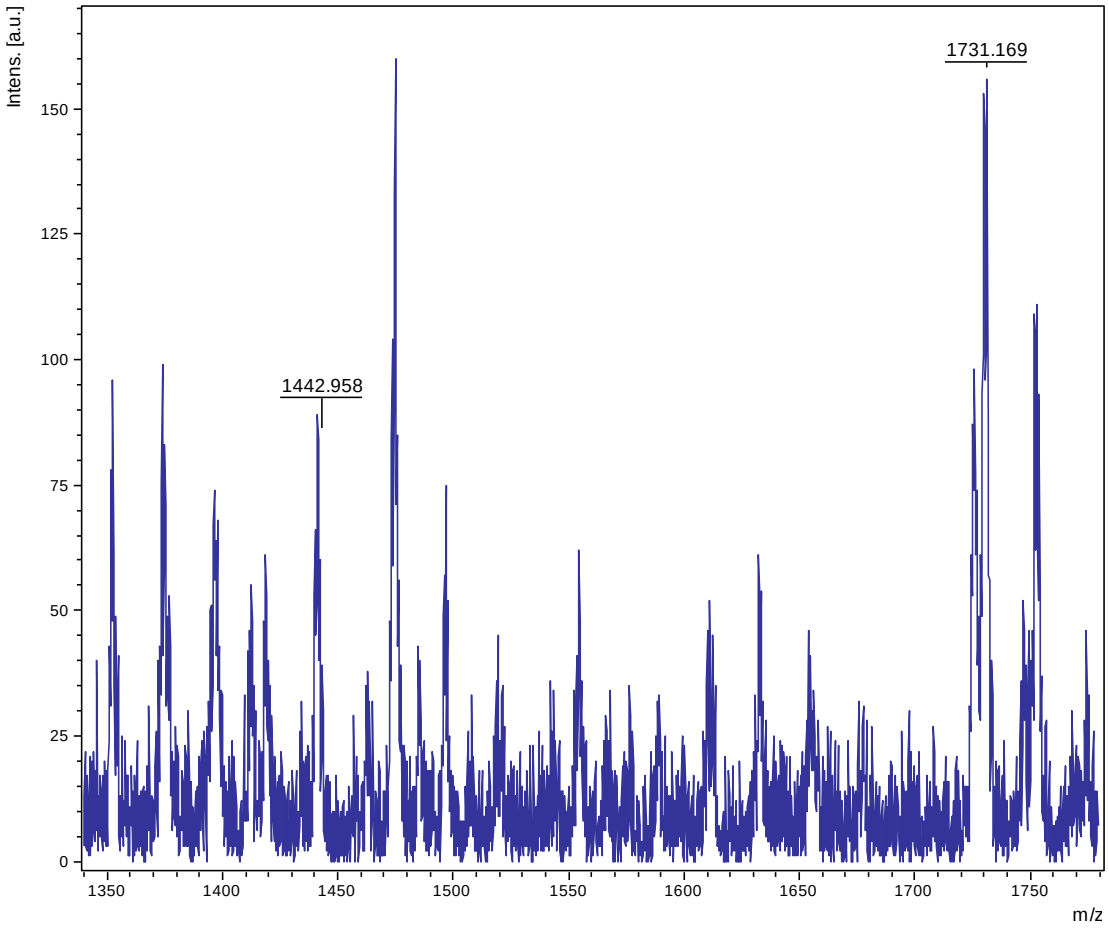
CLB-HyM-10 對 U4A-L4T DNA 有很高的專一性切割，位置在 U4A 的 A_5 和 G_6 之間及 L4T 的 T_{20} 和 C_{21} 之間(圖 3-2-1ABCDE)，其中 A_5 和 C_{21} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份，比對時也有發現 L4T 的 T_{15} 和 A_{16} 及 A_{16} 和 T_{17} 之間的切割片段分子量的荷質比訊號，而 A_{16} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份。

3-2-2 CLB-HyQ-10 MALDI-TOF MS



Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
T_1 A_2 G_3 $5`CH_3$–P–P–PO_3H	997.62 (Na^+)	997.460 (Na^+)
C_{24} T_{25} A_{26} HO_3P–P–P–OH 3`	944.13 (Na^+)	944.085 (Na^+)

圖 3-2-2A. 胜肽 CLB-HyQ-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。



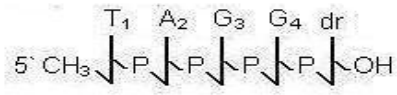
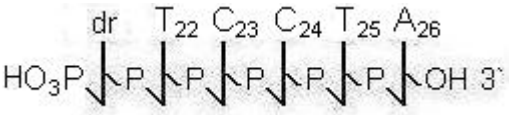
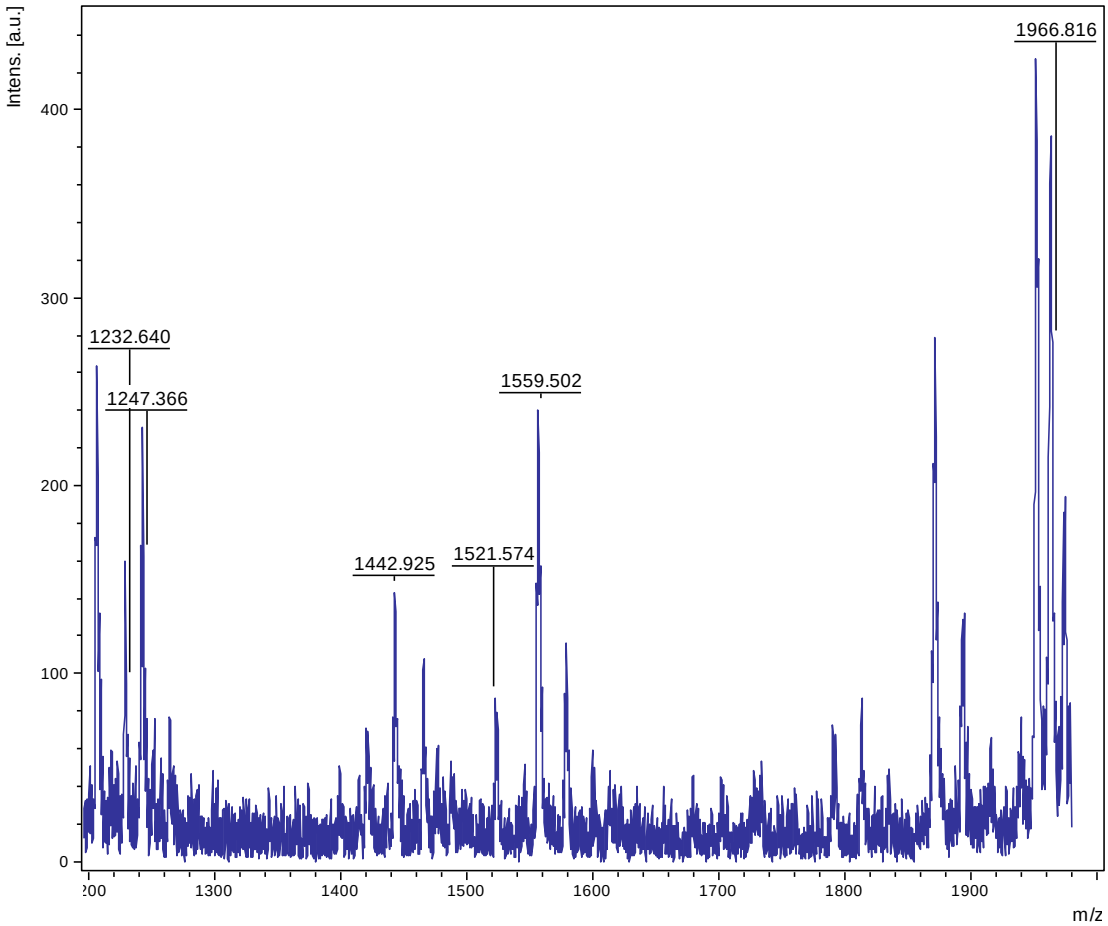
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	1442.12 (Na^+)	1442.958 (Na^+)
	1731.21 (Na^+)	1731.169 (Na^+)

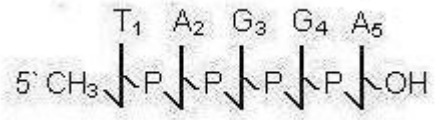
圖 3-2-2B. 胜肽 CLB-HyQ-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

CLB-HyQ-10 則呈現和 CLB-HyM-10 相似的結果，同樣也有很高的序列選擇性切割，切割位置在U4A 的A_5 和G_6 之間及L4T 的T_{20} 和C_{21} 之間 (圖 3-2-2AB)，而且其中 A_5 和 C_{21} 的鹼基部位也被胜肽切掉只留下其五碳醣的部份；CLB-HyQ-10 還有其他切割位置分別在U4A 的G_3 和G_4 之間及 L4T 的C_{23} 和C_{24} 之間。

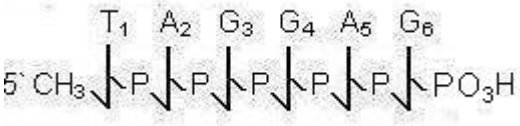
3-2-3 CLB-HyH-10 MALDI-TOF MS



Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
HO_3P-C_{23}-P-C_{24}-P-T_{25}-P-A_{26}-P-OH 3`	1232.17 (Na^+)	1232.640 (Na^+)
5` CH_3-T_1-P-A_2-P-G_3-P-G_4-P-OH	1247.84 (Na^+)	1247.366 (Na^+)
5` CH_3-T_1-P-A_2-P-G_3-P-G_4-P-dr-OH	1442.12 (Na^+)	1442.925 (Na^+)
5` CH_3-T_1-P-A_2-P-G_3-P-G_4-P-dr-PO_4	1521.89 (Na^+)	1521.574 (Na^+)



1559.29 (Na^+)　　1559.502 (Na^+)

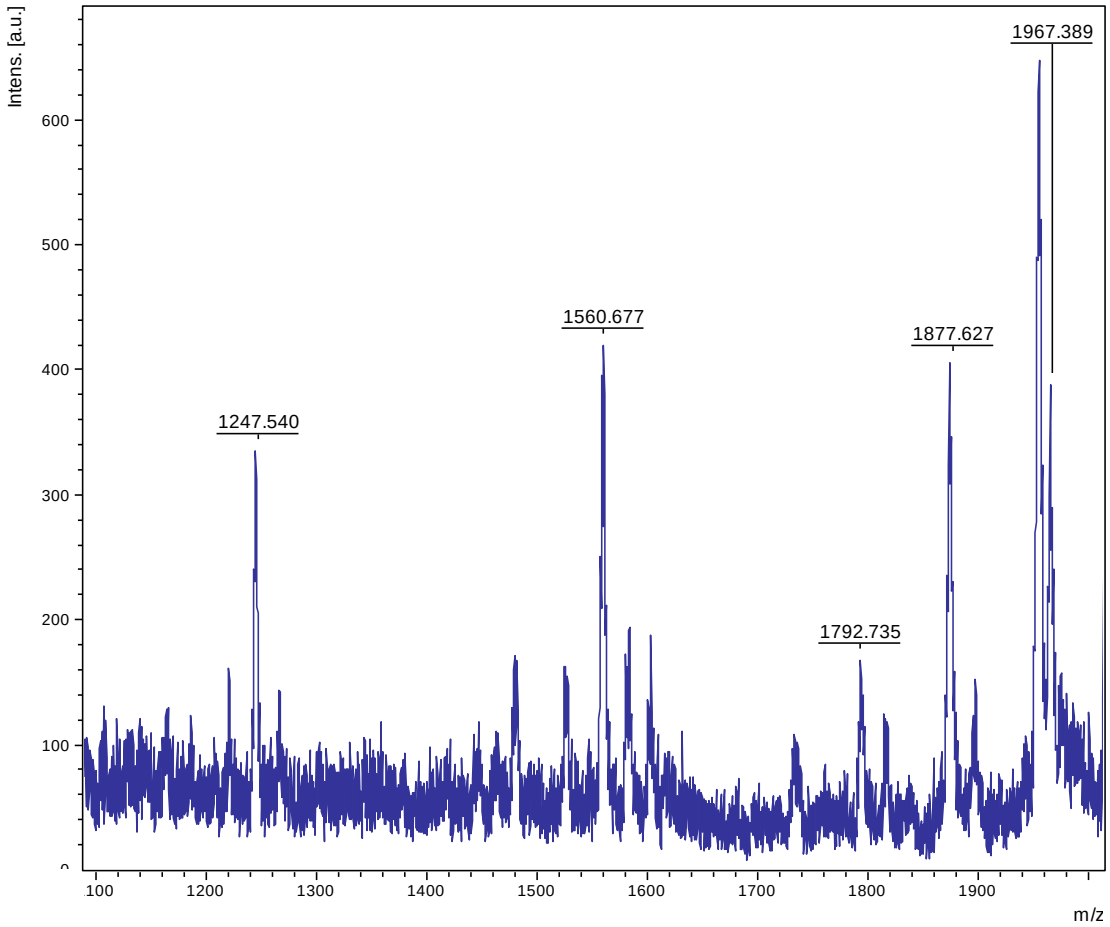


1966.21(Na^+)　　1966.816 (Na^+)

圖 3-2-3. 胜肽 CLB-HyH-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

CLB-HyH-10 對 U4A-L4T DNA 的切割，位置在 U4A 的 G_4 和 A_5、A_5 和 G_6、G_6 和 A_7 之間及 L4T 的 T_{20} 和 C_{21} 之間 (圖 3-2-3)，其中 A_5 鹼基部位會被胜肽切掉只留下其五碳醣的部份。

3-2-4 CLB-PyHK-10 MALDI-TOF MS



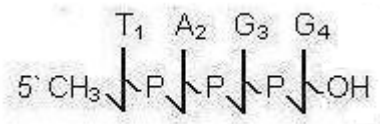
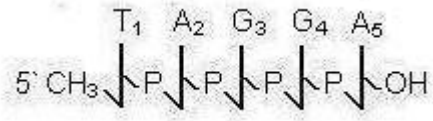
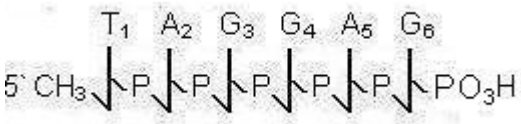
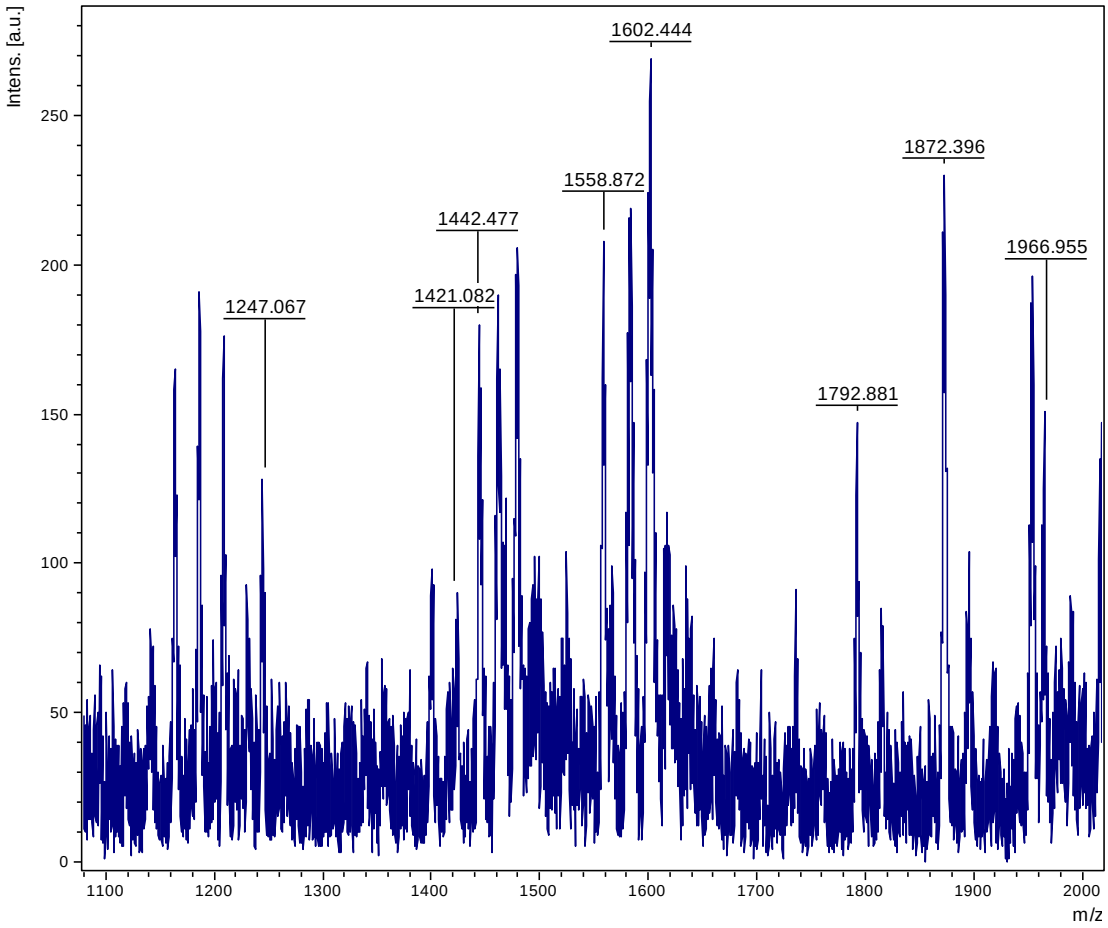
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	1247.12 (Na^+)	1247.540 (Na^+)
	1559.29 (Na^+)	1560.677 (Na^+)
	1966.21(Na^+)	1966.389 (Na^+)

圖 3-2-4A. 胜肽 CLB-PyHK-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。



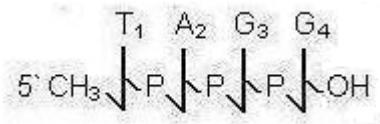
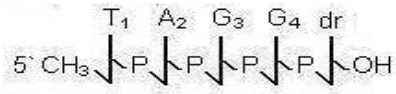
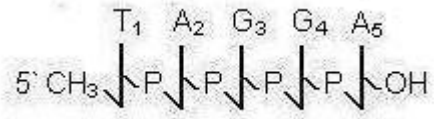
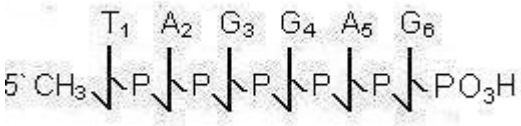
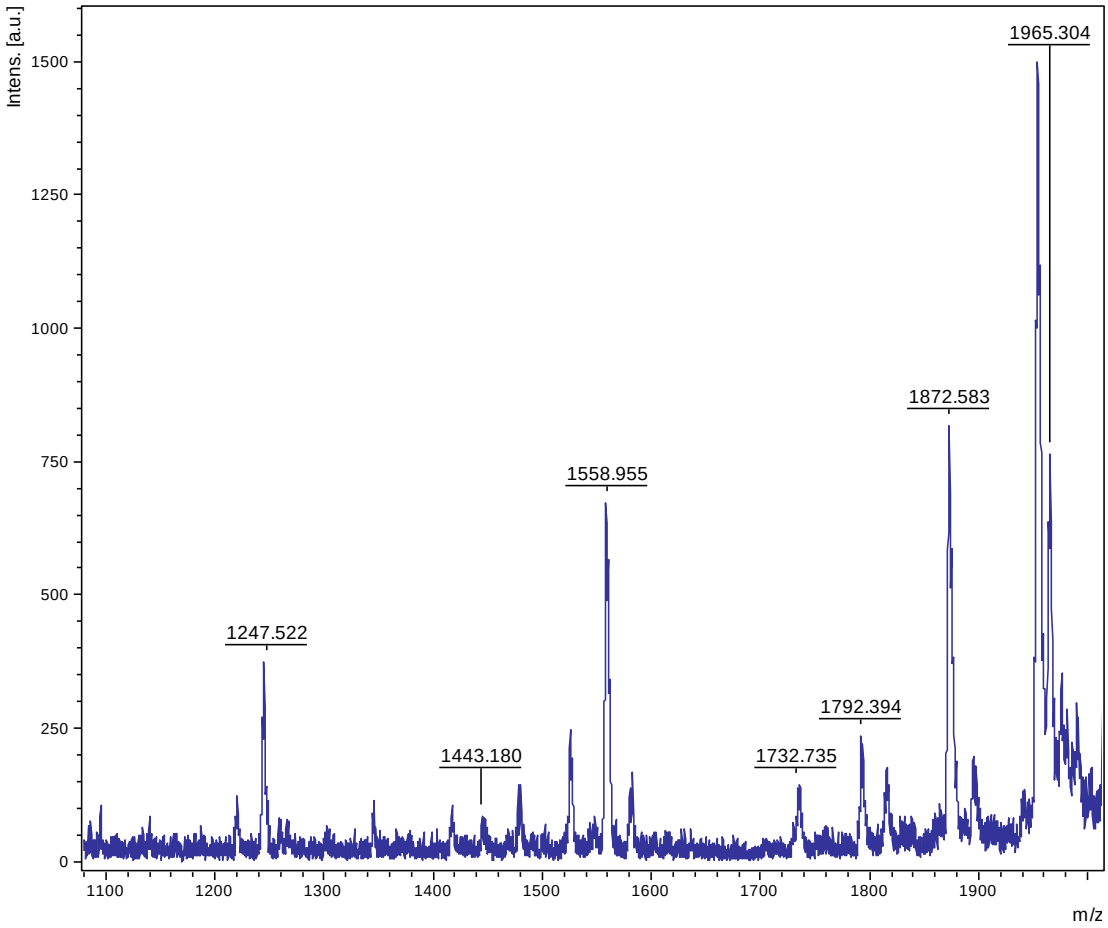
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	1247.12 (Na^+)	1247.067 (Na^+)
	1442.12 (Na^+)	1442.477 (Na^+)
	1559.29 (Na^+)	1558.872 (Na^+)
	1966.21(Na^+)	1966.955 (Na^+)

圖 3-2-4B. 胜肽 CLB-PyHK-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

CLB-PyHK-10 只對 U4A 的部份切割，切割位置在 G_4 和 A_5、A_5 和 G_6、G_6 和 A_7 之間 (圖 3-2-4AB)，而且只有 A_5 的鹼基部位被切除。

3-2-5 CLB-PyMK-10MALDI-TOF MS



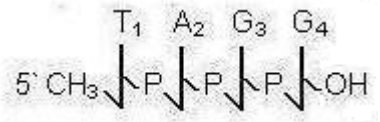
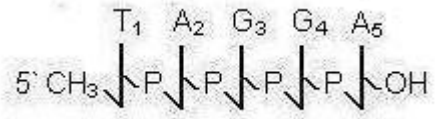
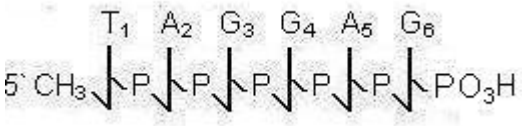
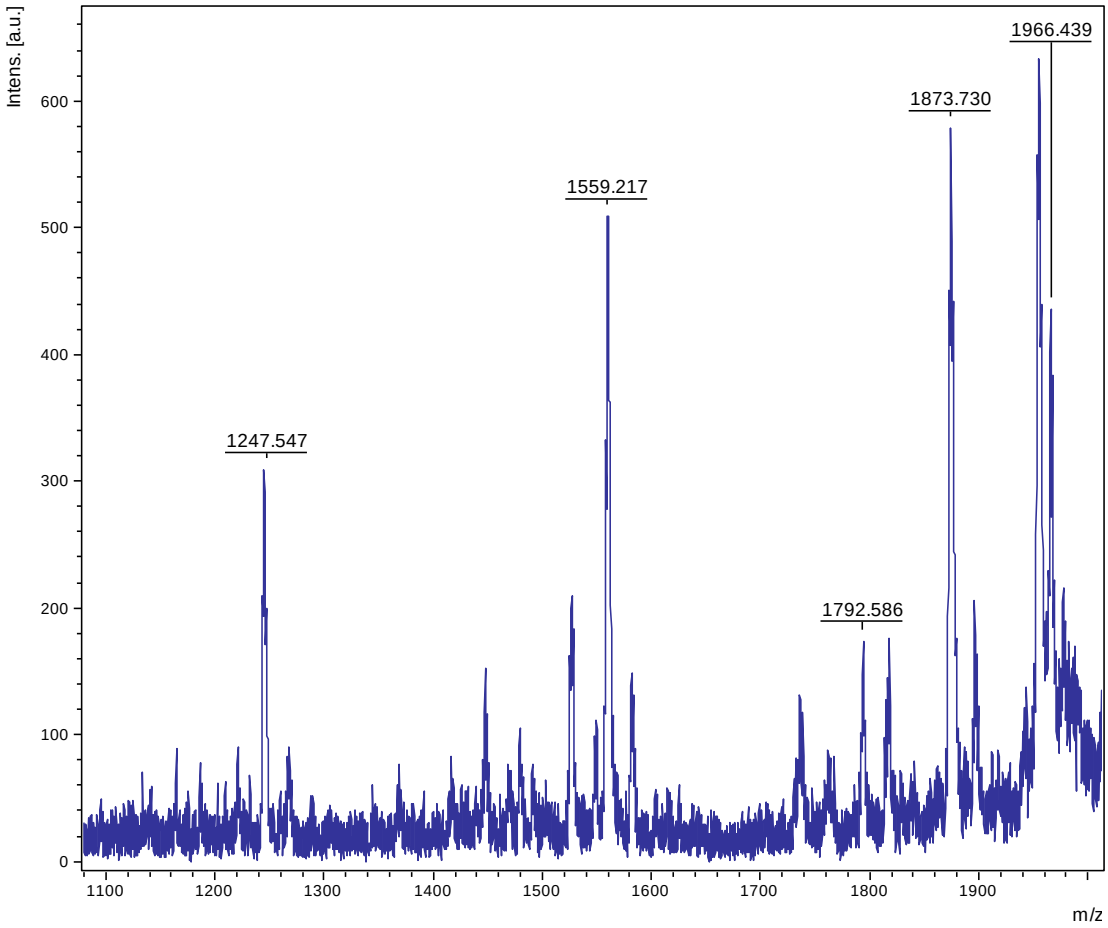
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	1247.12 (Na^+)	1247.522 (Na^+)
	1559.29 (Na^+)	1558.955 (Na^+)
	1966.21(Na^+)	1965.304 (Na^+)

圖 3-2-5A. 胜肽 CLB-PyMK-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。



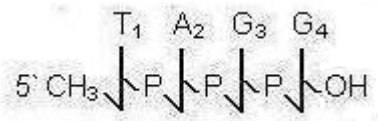
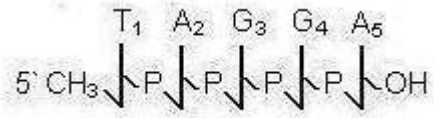
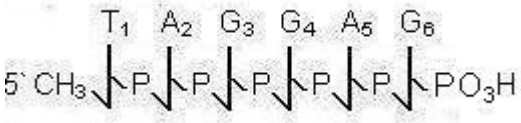
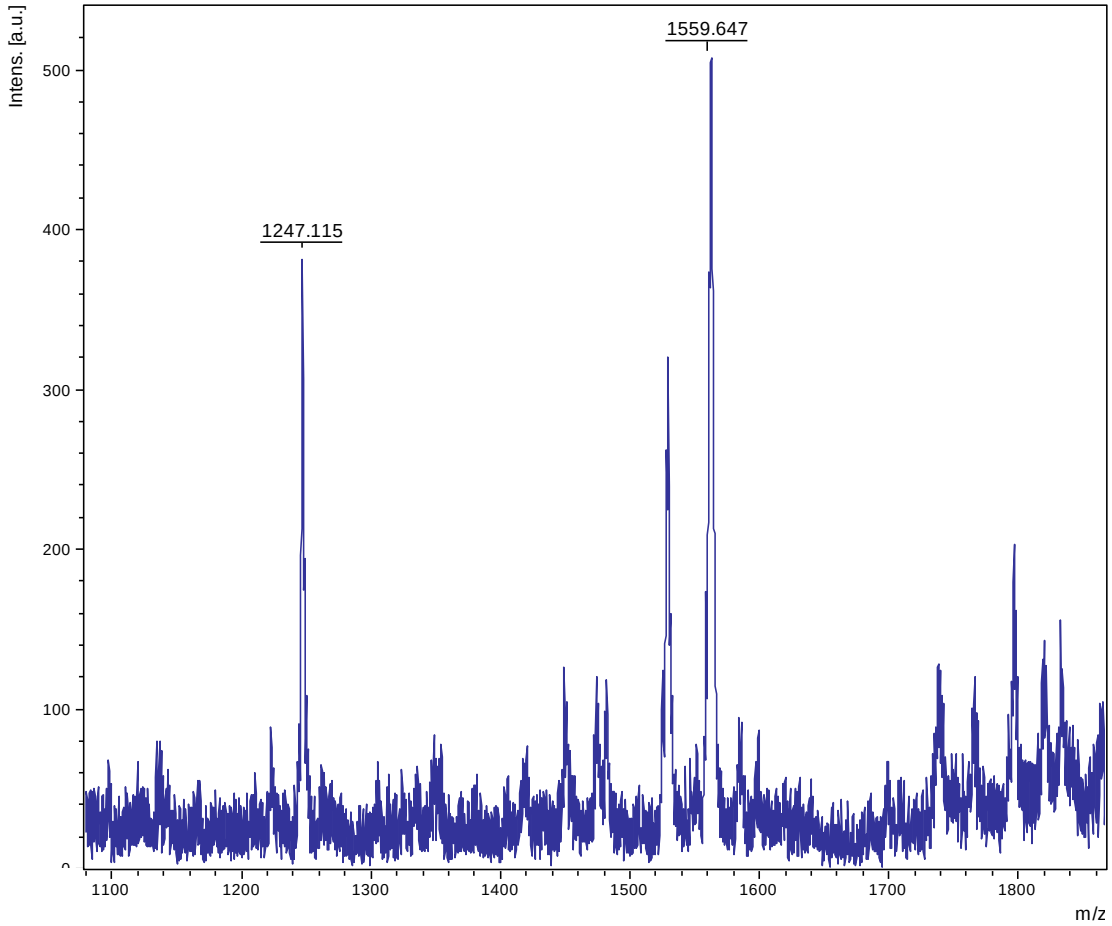
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	1247.12 (Na^+)	1247.547 (Na^+)
	1559.29 (Na^+)	1559.217 (Na^+)
	1966.21(Na^+)	1966.439 (Na^+)

圖 3-2-5B. 胜肽 CLB-PyMK-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

CLB-PyMK-10 對 U4A-L4T DNA 切割片段的比對上與 CLB-PyHK-10 有著很高的相似性，都只對 U4A 的部份切割，切割位置在 G_4 和 A_5、A_5 和 G_6、G_6 和 A_7 之間 (圖 3-2-5AB)，而只有 A_5 的鹼基部位被切除。

3-2-6 CLB-PyQK-10 MALDI-TOF MS



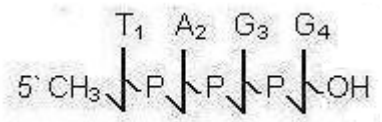
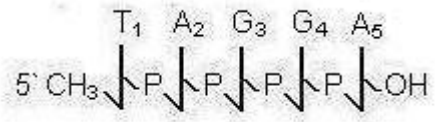
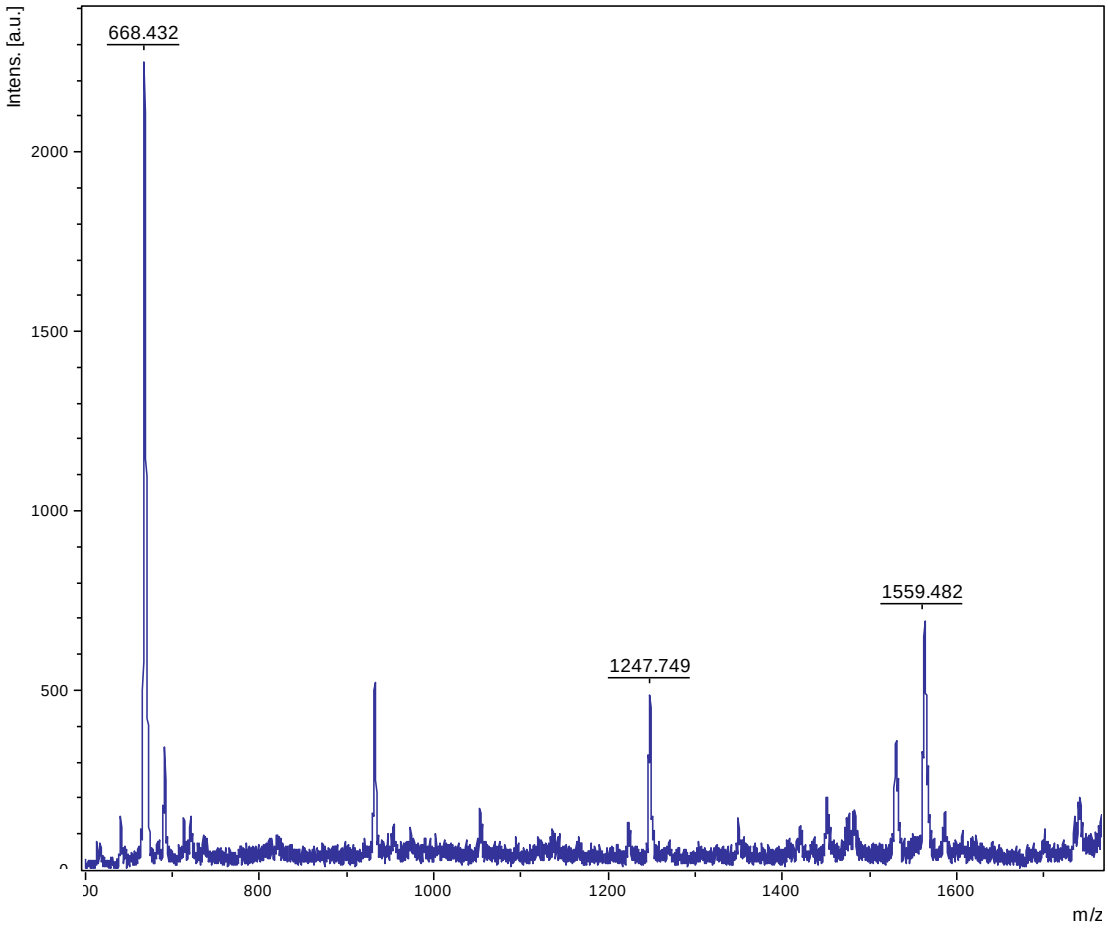
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	1247.12 (Na^+)	1247.115 (Na^+)
	1559.29 (Na^+)	1559.647 (Na^+)

圖 3-2-6. 胜肽 CLB-PyQK-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

CLB-PyQK-10 對 U4A-L4T DNA 的切割，位置在 U4A 的 G_4 和 A_5、A_5 和 G_6 之間 (圖 3-2-6)，而且也沒有找到 L4T 的切割片段。

3-2-7 CLB-PyWK-10 MALDI-TOF MS



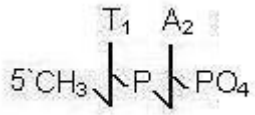
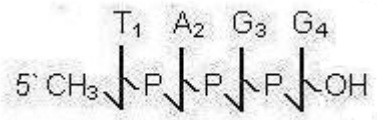
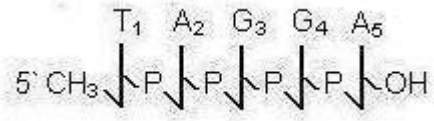
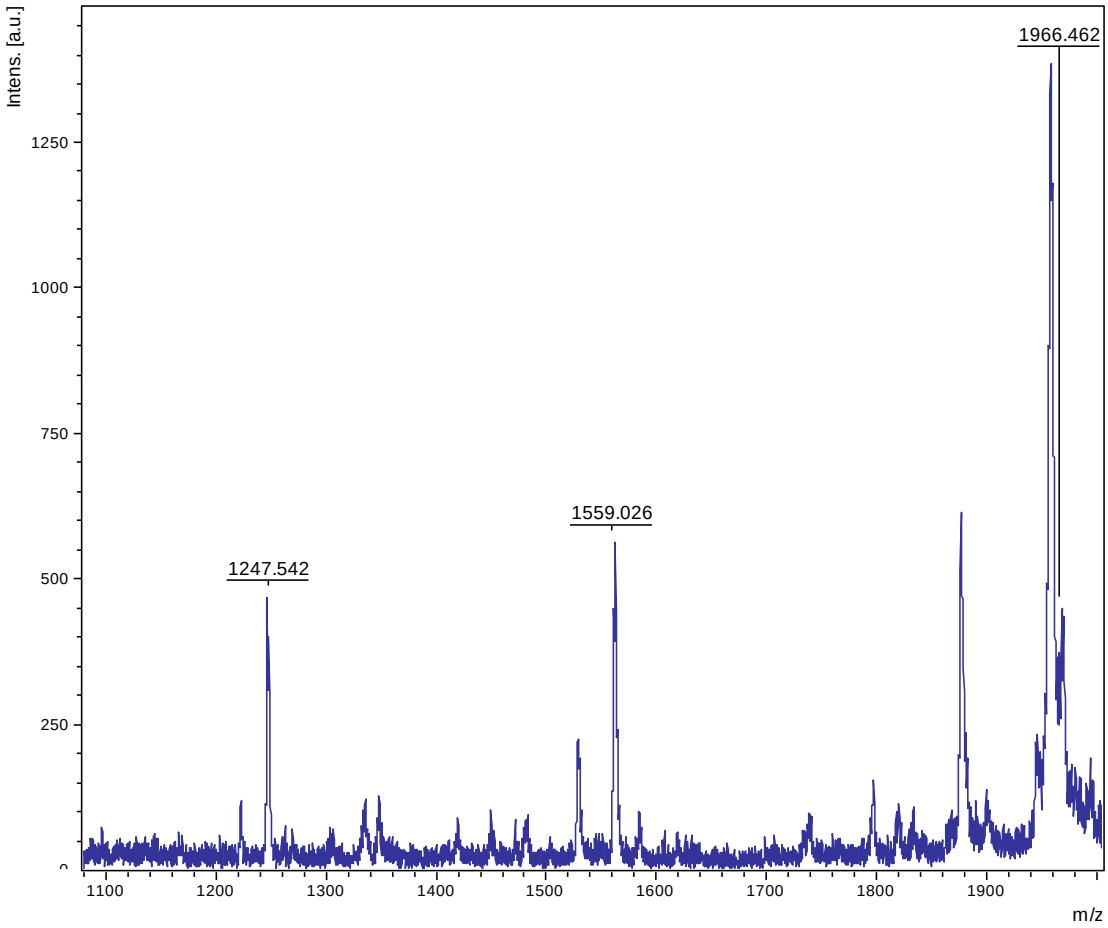
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
	669.42 (Na^+)	668.432 (Na^+)
	1247.12 (Na^+)	1247.749 (Na^+)
	1559.29 (Na^+)	1559.482 (Na^+)

圖 3-2-7A. 胜肽 CLB-PyWK-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。



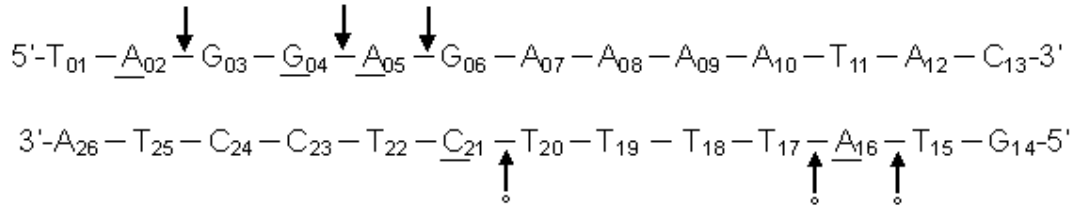
Assigned oligonucleotide fragment	Calcd. $[M-H]^+$	Observed $[M-H]^+$
$T_1\ A_2\ G_3\ G_4$ 5` CH_3-P-P-P-P-OH	1247.12 (Na^+)	1247.542 (Na^+)
$T_1\ A_2\ G_3\ G_4\ A_5$ 5` CH_3-P-P-P-P-P-OH	1559.29 (Na^+)	1559.026 (Na^+)
$T_1\ A_2\ G_3\ G_4\ A_5\ G_6$ 5` CH_3-P-P-P-P-P-P-PO_3H	1966.21(Na^+)	1966.462 (Na^+)

圖 3-2-7B. 胜肽 CLB-PyWK-10 對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

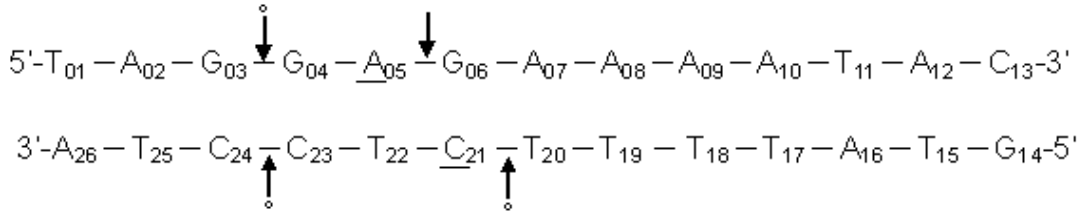
CLB-PyWK-10 對 U4A-L4T DNA 的切割，位置在 U4A 的 A_2 和 G_3、G_4 和 A_5、A_5 和 G_6、G_6 和 A_7 之間 (圖 3-2-7AB)，沒有找到 L4T 的切割片段，其中也沒有鹼基部位被胜肽切掉而只留下其五碳醣部份的鹼基。

3-2-8 CLB 切割型胜肽對 U4A-L4T DNA 位置總整理

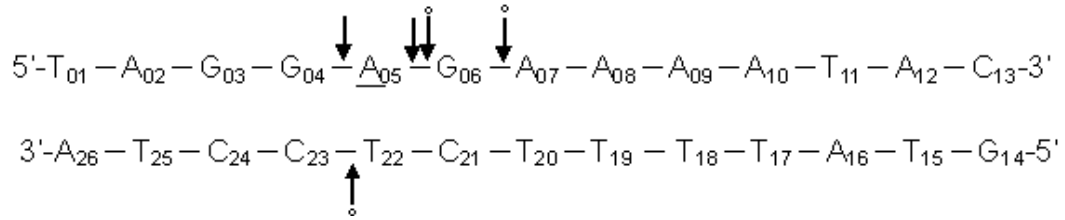
CLB-HyM-10



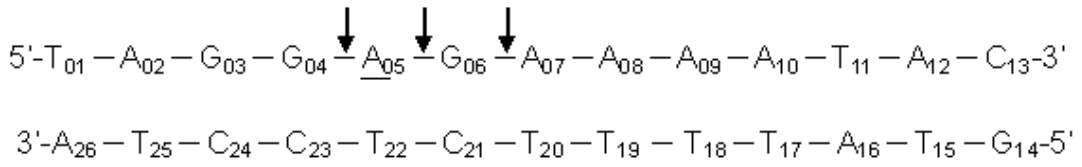
CLB-HyQ-10



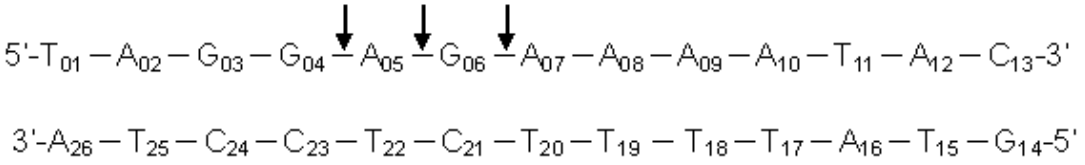
CLB-HyH-10



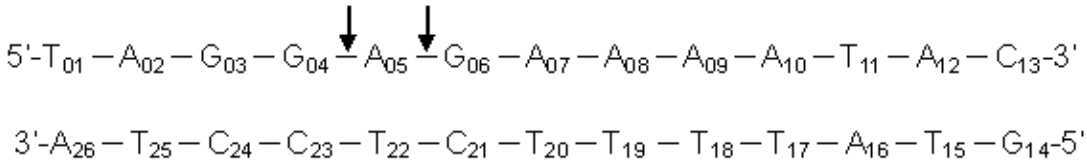
CLB-PyHK-10



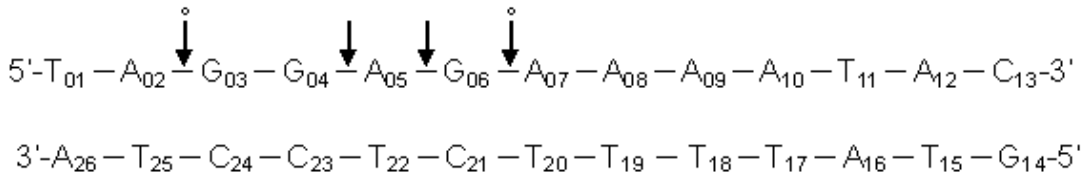
CLB-PyMK-10



CLB-PyQK-10



CLB-PyWK-10



 表示 AA-OH。

 表示 AA-PO_3H。

AA 表示該胺基酸鹼基布費被 CLB 切除，只留下五碳醣。

圖 3-2-8. 胜肽對 U4A-L4T DNA 切割片段之 MALDI-TOF MS 圖譜。

CLB-HyM-10 對 U4A-L4T DNA 有很高的專一性切割，位置在 U4A 的 A_5 和 G_6 之間及 L4T 的 T_{20} 和 C_{21} 之間，其中 A_5 和 C_{21} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份，比對時也有發現 L4T 的 T_{15} 和 A_{16} 及 A_{16} 和 T_{17} 之間的切割片段分子量的荷質比訊號，而 A_{16} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份。CLB-HyQ-10 則呈現和 CLB-HyM-10 相似的結果，但多了在 U4A 的 G_3 和 G_4 之間及 L4T 的 C_{23} 和 C_{24} 之間的切割位置。CLB-HyH-10 則切割在 U4A 的 G_4 和 A_5

之間及 L4T 的 C_{23} 和 T_{22} 之間，A_5 的鹼基部位會被胜肽切掉只留下其五碳醣的部份，A_5 和 G_6、G_6 和 A_7 之間也有切割片段。

在 CLB-PyMK-10、CLB-PyHK-10 對 U4A-L4T DNA 切割片段的比對上有著很高的相似性，他們都只對 U4A 的部份切割，切割位置在 G_4 和 A_5、A_5 和 G_6、G_6 和 A_7 之間。CLB-PyQK-10 則無 G_6 和 A_7 之間的切割，CLB-PyWK-10 則多了 A_2 和 G_3 之間的切割，而這系列的切割型胜肽對沒有 L4T DNA 的切割訊號，也沒有只留下五碳醣而切去鹼基的訊號。

第四章　結論

本文利用圓二色光譜(CD) 探討胜肽與 DNA 作用。由光譜中發現，當 DNA 與胜肽作用後，在波長 300 nm 到 350 nm 產生一個新的吸收峰；因為只有 DNA 存在或只有胜肽存在的光譜中，這段波長區域均未有明顯之吸收峰。因此，本文推測胜肽與 DNA 確實產生結合作用，且造成 DNA 結構上的變化。利用光譜圖可以知道 DNA 與胜肽作用位置是在小凹槽，而且橢圓率的變化會隨著濃度的增加而增強，而 near-isoelliptic point 的發生，亦證明了胜肽與 DNA 的結合作用是 2 個分子的胜肽對 1 個分子的 DNA，也就是二比一的結合。在[peptide] / [DNA] 對 $\Delta\theta$ 的關係中發現，當[peptide] / [DNA] 比在 0.5 - 2.0 之間時只有 1 分子的胜肽結合在 5’-d（AAAA）-d（TTTT）-3’ 之位置上，而當[peptide] / [DNA] 比在 2.0 - 3.0 之間時，則有 2 分子的胜肽結合在選擇性之序列上，這也再次證明了之前的推論。

在切割型胜肽 CD 圖譜的表現上，含 CLB 之胜肽在橢圓率的變化上都比與之相似序列之結合型胜肽來的小，而部分的切割型胜肽有清楚的飽和現象，是因為被切開的 DNA 其小凹槽部位遭到破壞，使得胜肽無法繼續在小凹槽結合。由實驗結果，對於發展找到對 DNA 作用更具選擇性及專一性的胜肽是有很大的幫助。

在含 CLB 之切割型胜肽對 U4A-L4T DNA 切割片段的比對上，本文利用 MALDI-TOF MS 尋找出精確的 DNA 片段分子量，並找到不同胜肽對 DNA 選擇性切割的位置。CLB-HyM-10 對 U4A-L4T DNA 有很高的專一性切割，位置在 U4A 的 A_5 和 G_6 之間及 L4T 的 T_{20} 和 C_{21} 之間，其中 A_5 和 C_{21} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份，比對時也有發現 L4T 的 T_{15} 和 A_{16} 及 A_{16} 和 T_{17} 之間的切割片段分子量的荷質比訊號，而 A_{16} 的鹼基部位會被胜肽切掉只留下其五碳醣的部份。CLB-HyQ-10 則呈現和 CLB-HyM-10 相似的結果，但多了在 U4A 的 G_3 和 G_4 之間及 L4T 的 C_{23} 和 C_{24} 之間的切割位置。CLB-HyH-10 則切割在 U4A 的 G_4 和 A_5 之間及 L4T 的 C_{23} 和 T_{22} 之間，A_5 的鹼基部位會被胜肽切掉只留下其五碳醣的部份，A_5 和 G_6、G_6 和 A_7 之間也有切割片段。

在 CLB-PyMK-10、CLB-PyHK-10 對 U4A-L4T DNA 切割片段的比對上有著很高的相似性，他們都只對 U4A 的部份切割，切割位置在 G_4 和 A_5、A_5 和 G_6、G_6 和 A_7 之間。CLB-PyQK-10 則無 G_6 和 A_7 之間的切割，CLB-PyWK-10 則多了 A_2 和 G_3 之間的切割，而這系列的切割型胜肽對沒有 L4T DNA 的切割訊號，也沒有只留下五碳醣而切去鹼基的訊號。本研究顯示不 同序列的胜肽呈現不同的 DNA 序列切割選擇性。

在本文的研究中發現不論是結合型的胜肽或是切割形的胜肽都有發展成抗癌藥物的潛力，他們對 DNA 都有很高的序列選擇性，而且可

以透過結合或切割的作用使得致病的 DNA 序列失去功能，也可以干擾要與該序列結合的受體，使其無法與 DNA 結合，希望這些研究成果能幫助實驗室往後的研究工作，開發出更有效之胜肽。

第五章 參考文獻

1. Suzuki, M. *Proc. Japan. Acad*. **1988a,** *64B*, 29.
2. Suzuki, M. *J. Mol. Biol.* **1990**, *207*, 61.
3. Von Holt, C., de Groot, P., Schwager, S. and Brandt, W. F. In Stein, G. S., Stein, J. L. and Marzluff, W. F. (eds), *Histone Genes.* Wiley-Interscience, New York, PP. **1984**, *65*
4. Suzuki, M. *The EMBO Journal.* **1989**, *8*, 797
5. Churchill, E. A.; Suzuki, M. *The EMBO Journal.* **1989**, *8*, 4189
6. Suzuki, M. *Proc. Japan. Acad*. **1988b,** *64B*, 33.
7. Yang, C. H.; Chou, P. J.; Luo, Z. L.; Chou, I. C.; Chang, J. C.; Cheng, C. C.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *Bioorg. Med. Chem.* **2003**, *11*, 3279.
8. Ragnarsson, U.; Grehn, L. *J. Org. Chem.* **1981**, *46*, 3492.
9. Patel, D. J. *Proc. Natl. Acad. Sic. U.S.A.* **1982**, *79*, 6424.
10. Pelton, J.G.; Wemmer, D. E. *Proc. Natl. Acad. Sic. U.S.A.* **1982**, *79*, 6424.
11. Pelton, J.G.; Wemmer, D. E. *Proc. Natl. Acad. Sic. U.S.A.* **1989**, *86*, 5723.
12. Bailly, C.; Chaires, B. C. *Bioconjugate Chem*. **1998**, *9*, 513.
13. Kopka, M. L.;Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E.

Proc. Natl. Acad. Sic. U.S.A. **1985**, *82*, 1376.

14. Wurtz, N. R.; Dervan, P. B. *Chemistry & Biology*. **2000**, *7*, 153.
15. Warren, S. W.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *144*, 8783.
16. White, S.; Baird, E. E.; Dervan, P. B. *Biochemistry*. **1996**, *35*, 12532.
17. Trauger, J. W,; Baird, E. E.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6160.
18. Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147.
19. Michelle, E. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.
20. White, S.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 8756.
21. Chang, J. C.; Yang, C. H.; Chau, P. J.; Yang, W. H.; Chau, I. C.; Lu, C. T.; Sheh, L. *Bioorg. Med. Chem*. **2004**, *12*, 53.
22. Bruce B. B.; kanganis, D. L.;Liebes, F.;Robert, S. *Cancer Research*.**1989,** 49, 554-559
23. Hall, A.;Robson, C. N.; Hickson, I. D.;Harris, A.L.; Proctor, S. J.; Cattan. A. R. *Cancer Research*. **1989,** 49, 6265-6268
24. Cullis, P. M.;Davies L. M.;Weaver, R. *J. Am.Chem.* SOC. **1995,**

117, 8033-8034

25. Frahm, A. W.; Heinz H.; Kratz,F. *J.Med. Chem.* **1998,** 41, 2701-2708.

26. Hovinen, J. L.; Silvennoinen, R.; Vilpo. *J. Chem. Res.* Toxicol. **1998,** 11, 91-93.

27. Everett, J. L.;Roberts, J. J.;Ross, W. C. J. *J. Chem. Soc.***1953,** 2386.

28.Hall, A.; Robson, I. D.; Harris, A. L.;Proctor, S. J.; Cattan, A. R.*Cancer Res.* **1989.** 49, 6265

29. Bruce, B. B.; Denise, K.; Leonard, L. F.; Silber, R. *Cancer Res.* **1989.** 49, 554.

30. Hynes, R. O. *Cell.* **1992.** 69, 11-25.

31. Sheh, L. et al.. *Anti-Canceer Drug Design,* **1995,**10, 373

32. Sheh, L. et al.. *Bioconjugate Chem.* **2001,**12, 870

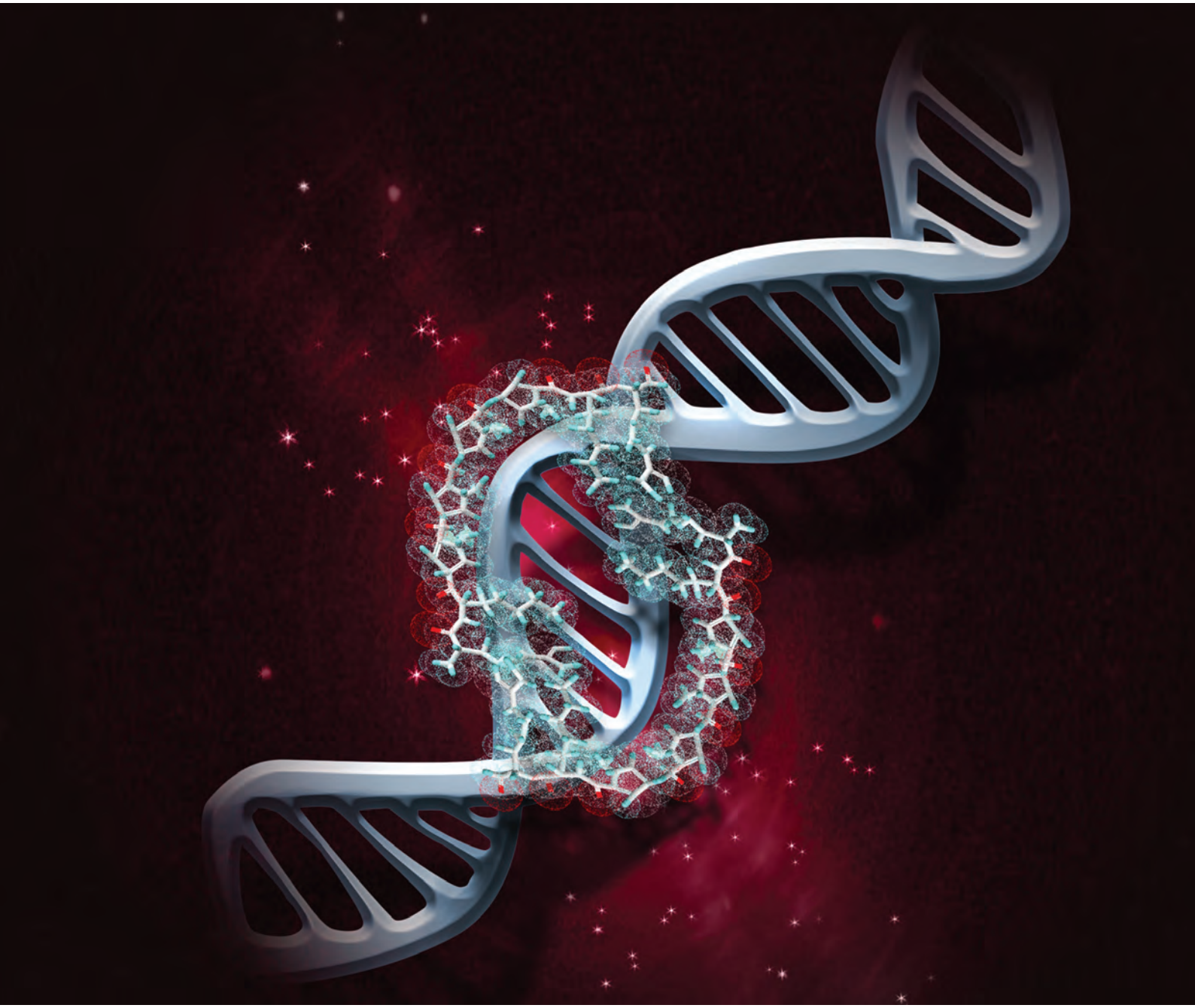
附錄

著　　作

Organic & Biomolecular Chemistry

www.rsc.org/obc

Volume 11 | Number 1 | 7 January 2013 | Pages 1–180



ISSN 1477-0520

RSC Publishing

PAPER

Leung Sheh *et al.*

Energetic studies on DNA–peptide interaction in relation to the enthalpy–entropy compensation paradox

1477-0520(2013)11:1;1-I

Organic & Biomolecular Chemistry

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 48

Energetic studies on DNA–peptide interaction in relation to the enthalpy–entropy compensation paradox†

Robin C. K. Yang,‡[a] Jonathan T. B. Huang,‡[a] Shih-Chuan Chien,[a] Roy Huang,[a] Kee-Ching G. Jeng,[b] Yen-Chung Chen,[a] Mokai Liao,[a] Jia-Rong Wu,[a] Wei-Kang Hung,[a] Chia-Chun Hung,[a] Yu-Ling Chen,[a] Michael J. Waring[c] and Leung Sheh*[a]

This study aims to interpret the energetic basis of complex DNA–peptide interactions according to a novel allosteric interaction network approach. In common with other designed peptides, five new conjugates incorporating the XPRK or XHypRK motif (Hyp = hydroxyproline) attached to a *N*-methylpyrrole (Py) tract with a basic tail have been found to display cooperative binding to DNA involving multiple monodentate as well as interstrand bidentate interactions. Using quantitative DNase I footprinting it appears that allosteric communication *via* cooperative binding to multiple sites on complementary DNA strands corresponds to two different types of DNA–peptide interaction network. Temperature variation experiments using a dodecapeptide RY-12 show that lower temperature (25 °C) favor a circuit type of allosteric interaction network, whereas higher temperatures (31 and 37 °C) afford only a partial-circuit type of network. Circular dichroism studies show that our five peptides induce significant local conformational changes in DNA *via* the minor groove, with apparently dimeric binding stoichiometry. Isothermal titration calorimetry reveals that these peptides, together with another seven for comparison, are strongly exothermic upon binding to a model 13-mer DNA duplex, characterized by ΔH ranging from −14.7 to −74.4 kcal mol^{-1}, and also high $T\Delta S$ ranging from −6.5 to −65.9 kcal mol^{-1}. Multiple monodentate and bidentate interactions, as well as ionic forces that mediate positive cooperativity in sequence recognition, are consistent with a dramatic decrease in entropy and a 'tightening' effect of DNA conformation. Distinctive enthalpy–entropy compensation (EEC) relationships are demonstrated for the interaction of all twelve designed peptides with DNA, affording a straight line of slope close to unity when ΔH is plotted *versus* $T\Delta S$, with a *y*-axis intercept (average ΔG) corresponding to −8.5 kcal mol^{-1}, while the observed ΔG ranges from −8.2 to −9.1 kcal mol^{-1} for the peptides. The EEC seen with peptide RY-12 binding to the model duplex persists throughout various incubation temperatures. The net compensation of energy between the favorable negative ΔH and unfavorable negative ΔS components thus constrains the value of net binding free energy ΔG within a remarkably constant range, as is clearly visible in a 3-dimensional energetic plot. We conclude that the preservation of a rather narrowly-defined ΔG value is central to the EEC in DNA–peptide interactions, illuminating the universal EEC paradox commonly found in diverse biochemical reactions.

Received 11th July 2012,
Accepted 12th September 2012

DOI: 10.1039/c2ob26320c

www.rsc.org/obc

1 Introduction

As research on the sequence-selective binding of drugs to DNA gathers pace, it has raised increasingly sophisticated questions. After more than three decades, during which these studies have laid the foundations of various areas of structural biochemistry and have contributed much to understanding the bioorganic chemistry of nucleic acids, attention is now being turned to applications, particularly the synthesis of potentially useful drugs. Classical DNA–small molecule interaction studies include work on a large variety of small

[a]*Department of Chemistry and Life Science Center, Tunghai Christian University, Taichung, Taiwan 407, R.O.C. E-mail: Lsheh@thu.edu.tw; Fax: +886-4-23590426; Tel: +886-4-23592674*
[b]*Department of Medical Research, Taichung Veterans General Hospital, R.O.C Taiwan 405*
[c]*Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, England*
†Electronic supplementary information (ESI) available: See DOI: 10.1039/c2ob26320c
‡R. C. K. Y. and J. T. B. H. contributed equally to this work.

molecules that serve as effective ligands, namely antitumor antibiotics,[1–4] lexitropsins,[4–6] synthetic polyamides,[4,7–11] heterocycles[12,13] and synthetic peptides.[14–18] One notable outcome is the relatively recent acknowledgment that understanding the molecular basis for the recognition of specific sequences in target DNA by small synthetic molecules will prove useful in future gene therapy.[4] At the same time, efforts to investigate the molecular events underlying cooperativity between DNA–ligand binding sites are expected to provide new insight into chemical rules governing structural biochemistry.

For over ten years we have been studying the sequence-selective recognition of DNA by synthetic peptides incorporating the XP(Hyp)RK motif.[14–18] Peptides incorporating such motifs possess excellent DNA binding capability in the submicromolar concentration range. Further attachment of one or more 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) to XP(Hyp)RK peptides augments their sequence-specificity toward sequences containing consecutive arrays of A or T nucleotides. Our design of XPRK-containing peptides stems from the finding that a SPXX motif[19,20] is often found in repeating sequences in histones, steroid hormone receptors, various segmentation gene products and some oncogene products. The SPXX motif assumes a β-turn stabilized by two hydrogen bonds, and the side chains of the two basic residues engage in salt bridges with the DNA phosphate groups.

It is now generally accepted that allosteric behavior of macromolecules[21] such as proteins is a vital biological control process governing biochemical efficiency and energy expenditure. However, much less is known about the possible allosteric aspects of DNA–ligand binding than about the extensively studied protein–ligand interactions. Notably, over the first decade of the twenty-first century, it has been recognized that interaction networks between bio-molecules occur in many biochemical, biophysical and sub-cellular processes in nature.[22,23] High-throughput techniques have helped to establish interaction maps that expose protein–protein interaction networks, metabolic networks, and transcriptional regulatory networks.[22] Six years ago we proposed a network-based DNA–peptide allosteric interaction model interconnecting multiple sites in fragments of the latent membrane protein (LMP-1) gene from a pathogenic Epstein–Barr virus variant derived from nasopharyngeal carcinoma.[16] Subsequently we refined the model in terms of three different types of network-based allosteric communication between synthetic peptides binding to DNA: circuit type, incomplete circuit type, and non-circuit type characterized by interstrand bidentate interactions.[17,18]

In this study we further explore the allosteric features of DNA–peptide molecular recognition with regard to interaction networks, choosing designed peptides that exhibit significant cooperativity in binding to DNA. Quantitative DNase I footprinting furnishes the prime methodology for studying the allosteric features of DNA interaction networks associated with the binding of two designed decapeptides incorporating the XHypRK motif: Ser-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyS-10), Glu-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyE-10), and two decapeptides incorporating a novel KPXR motif: Lys-Pro-Met-Arg-$(Py)_4$-Lys-Arg-NH_2 (PyMK-10), and Lys-Pro-Gln-Arg-$(Py)_4$-Lys-Arg-NH_2 (PyQK-10). In addition, a closely related dodecapeptide Tyr-Pro-Arg-Lys-$(Py)_4$-Lys-Arg-Pro-Tyr-NH_2 (RY-12) was also investigated. DNase footprinting experiments were carried out using a 5′-[^{32}P]-labeled 158-mer DNA duplex and an essentially complementary 5′-[^{32}P]-labeled 135-mer DNA duplex employed in previous work.[14] Circular dichroism was also brought to bear upon the problem as a means of directly studying conformational changes induced by peptide binding as well as salt effects upon interaction between peptide HyM-10 and a 13-mer deoxyribonucleotide duplex.

We prefer quantitative footprinting as the main method of analysis, because it is difficult for most other techniques to discriminate multiple DNA binding sites engaged in allosteric interactions. In assessing the results, the coefficient determined by the Hill equation has been used as a comparative parameter for expressing cooperativity because of its wide applicability in research.[24–26] Circular dichroism (CD) complements these techniques and is particularly useful for investigating the conformational and stoichiometric aspects of DNA–peptide binding. The influence of temperature upon the allosteric behavior of peptide RY-12 was studied by quantitative DNase I footprinting as well as CD experiments. Further valuable insight into the thermodynamic basis of the allosteric interactions has been gained using isothermal titration calorimetry (ITC).

2 Results and discussion

The major objective of the present study was two-fold: firstly, to examine the footprinting behavior of five designed peptides for the purpose of verifying our recent hypothesis that three different types of network-based allosteric communication occur when synthetic peptides bind to multiple DNA recognition sites.[17,18] Secondly, to gain insight into the thermodynamic basis for the allosteric interaction of peptides incorporating the XP(Hyp)RK motifs with DNA. Hydrogen bonding, broadly categorized as monodentate or interstrand bidentate interaction,[27] between the peptide moieties and DNA bases, is central to sequence-selective DNA–protein[27] and DNA–small ligand[14–18] recognition processes. The interactions of the Arg and Lys residues of the designed peptides with DNA bases are crucial in generating interstrand bidentate interactions as well as monodentate interactions.[27] The "X" amino acid residues of these peptides, such as Ser, Glu, Gln, Tyr and Met, were chosen because their side chains could also engage in monodentate interactions with the DNA bases.[27]

The position of interstrand bidentate interactions can be assigned by connecting bases (with pecked lines) wherever significant simultaneous DNase I blockage can be discerned on complementary strands in the differential cleavage plots (Fig. 1). The peaks of the inhibition profiles are virtually always shifted 2–3 nucleotide pairs in the 3′-direction,

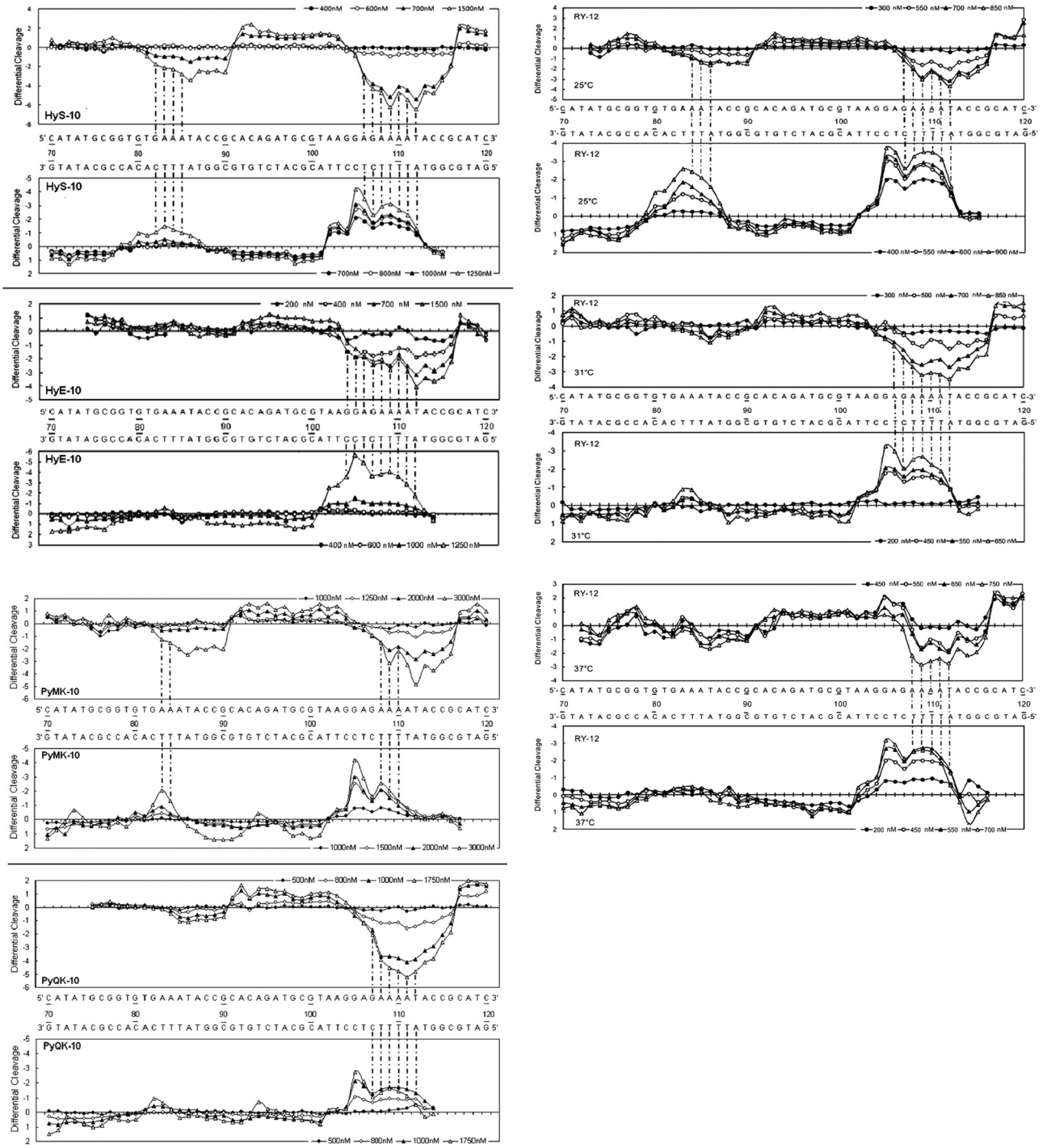


Fig. 1 Differential cleavage plots comparing the susceptibility of DNA fragments to DNase I cleavage after incubation with each peptide in cacodylate buffer at room temperature for 60 min. The upper traces represent the differential cleavage plot for a given peptide bound to the 5′-[^{32}P]-labeled upper strand (158-mer) DNA fragment; the lower traces represent the corresponding plots for the 5′-[^{32}P]-labeled lower strand (135-mer) DNA fragment. Vertical dotted lines between DNA bases represent the assignment of interstrand bidentate interactions where significant coincident H-bonding interactions occur between complementary bases in both strands. The incubation temperature for peptides HyS-10, HyE-10, PyMK-10 and PyQK-10 was 37 °C, whereas the incubation temperatures for peptide YR-12 were as stated.

signifying that the ligand lies in the minor groove of the helix where DNase I cuts it.[17,18]

The two decapeptides Ser-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyS-10) and Glu-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyE-10) were designed such that the number of amino acid residues is minimized, with four Py residues sandwiched between a XHypRK motif and a terminal dipeptide (KR) motif. A single amino acid residue variation (Ser or Glu) at the N-terminal was

introduced to investigate the differences in sequence and network selectivity. Recently, we found that incorporating a new KPXR motif on the N-terminal side also affords designed peptides that bind well to DNA. Two peptide amides of this new series were synthesized, each differing only in the third amino acid residue: Lys-Pro-Met-Arg-$(Py)_4$-Lys-Arg-NH_2 (PyMK-10), and Lys-Pro-Gln-Arg-$(Py)_4$- Lys-Arg-NH_2 (PyQK-10). Footprinting experiments show that these decapeptides bind sequence-selectively with significant positive cooperativity, as expected. The autoradiographs for the footprinting of these peptides are shown in Fig. 2 and Fig. i(a) (ESI†), and the binding parameters are shown in Table 1 (incubation temperature 37 °C).

To explore the temperature-dependence of sequence preference in the XHypRK series of peptides, we went on to examine DNase I footprinting with a dodecapeptide RY-12 using the same DNA fragments at three incubation temperatures (Fig. 1 and Fig. i(b)–i(d), ESI†). RY-12 was designed with its C-terminal tetrapeptide sequence reversed compared to the parent peptide PyPro-12[18] and has two tyrosine residues substituted for histidine. At the usual incubation temperature of 37 °C, like HyE-10, dodecapeptide RY-12 seems to bind at only two strong sites: on the upper strand around position U108-115, corresponding to the sequence 5′-AAAATACC-3′ and showing weak positive cooperativity (n_H = 1.4), and on the lower strand around position L112-105 comprising the sequence 5′-ATTTTCTC-3′ with more significant positive cooperativity (n_H = 2.4). Weak DNase 1 blockages are also observed in position U84-90. Interstrand bidentate interactions are therefore assigned around position 108–112. At 31 °C, the sequence preference of RY-12 remains much the same (Fig. 1 and Fig. i(c), ESI†), with significant positively cooperative binding at position U106-116 (n_H = 2.8) and position L112-105 (n_H = 2.2) indicating interstrand bidentate interactions around position 106–112.

When the incubation temperature is lowered to 25 °C, significant changes occur (Fig. 1 and Fig. i(b), ESI†). Peptide RY-12 now picks up the same four sites as peptides HyM-10, HyQ-10 and HyS-10, namely two on the upper strand, both displaying strong cooperativity: U84-90 (n_H = 3.4) and U107-116 (n_H = 2.9) together with two on the lower strand, L86-80 (n_H = 3.4) and L112-105 (n_H = 2.7). Interstrand bidentate interactions are assigned around positions 84–86 and 107–112.

Based on the footprinting results, network models[17,18] were constructed for each peptide in an effort to interpret the complex communication between binding sites (Fig. 3A and B). It is clear that a broadly similar model applies to all of them, suggesting network-based allosteric communication between practically all recognition sites. Conformational changes at particular binding loci must be readily transmitted to adjacent binding sites and also to sites on the complementary strand, facilitating (or possibly impeding) the binding of further peptide molecules. In this way, occupation of almost any binding locus can be signaled to adjacent subbinding sites and also to sites on the complementary strand, doubtless *via* interstrand bidentate interactions. Allosteric

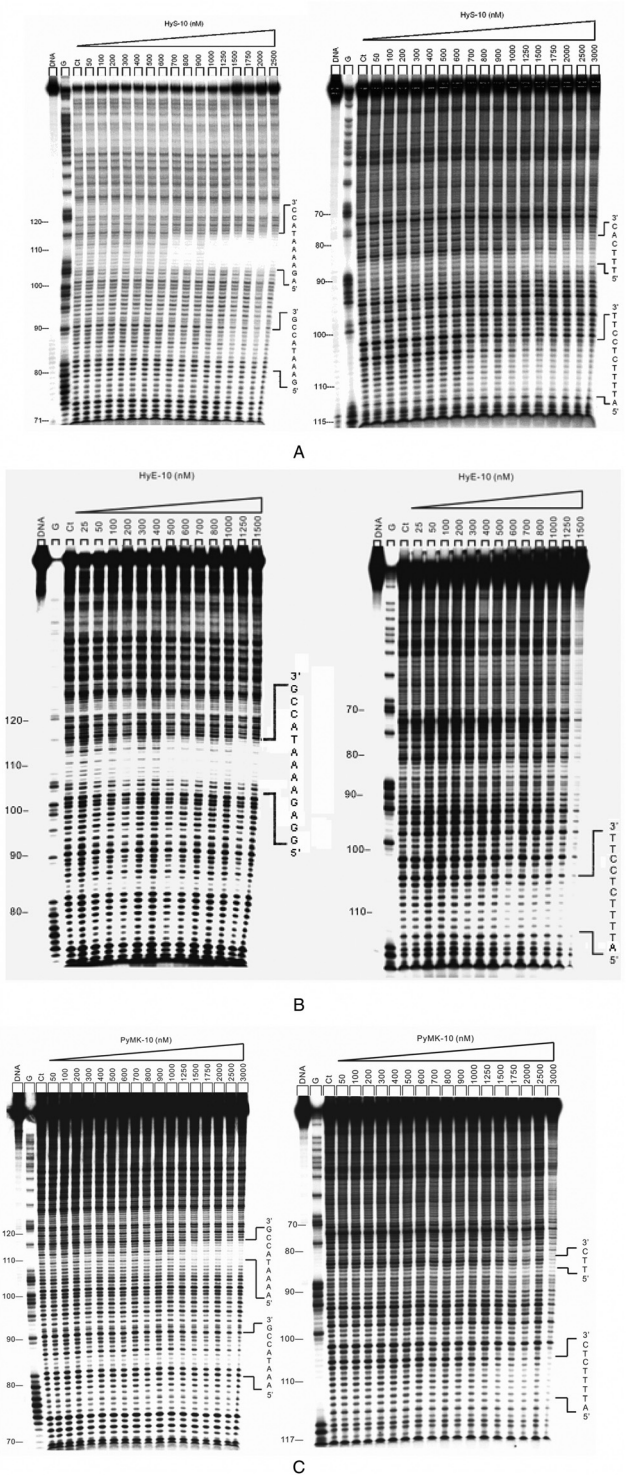


Fig. 2 (A) Autoradiographs showing DNase I footprinting of peptide HyS-10 on DNA duplexes labeled at the 5′ end: 158-mer upper strand, left panel; 135-mer lower strand, right panel. Peptide HyS-10 was equilibrated with the DNA in 5 mM sodium cacodylate buffer, pH 6.5 at 37 °C for 60 min before DNase 1 cleavage. G represents a Maxam–Gilbert guanine sequencing track and Ct shows a DNase I digestion control lane. (B) DNase I footprinting of peptide HyE-10 on DNA duplexes labeled at the 5′ end: upper strand, left panel; lower strand, right panel. (C) DNase I footprinting of peptide PyMK-10 on DNA duplexes labeled at the 5′ end: upper strand, left panel; lower strand, right panel.

Table 1 Sequence-specificity and physicochemical parameters for the binding of peptides to recognition sites in complementary upper (158-mer) and lower (135-mer) DNA strands at 37 °C, determined by quantitative DNase I footprinting.

Ligand	Binding site position	Recognition sequence	K_a	n_H	Position of interstrand bidentate interactions	Type of interaction network
HyS-10	U82-90	5′-GAAATACCG-3′	1.2×10^6	4.0	82–85	Circuit
	U106-115	5′-AGAAAATACC-3′	1.7×10^6	4.5	106–112	
	L85-80	5′-TTTCAC-3′	8.7×10^5	1.5		
	L112-102	5′-ATTTTCTCCTT-3′	1.5×10^6	4.0		
HyE-10	U104-116	5′-GGAGAAAATACCG-3′	1.1×10^7	1.8	104–112	Partial-circuit
	L112-102	5′-ATTTTCTCCTT-3′	3.4×10^6	1.7		
PyMK-10	U83-90	5′-AAATACCG-3′	8.4×10^5	2.4	83–84	Circuit
	U108-116	5′-AAAATACCG-3′	6.4×10^5	4.3	108–110	
	L84-82	5′-TTC-3′	6.5×10^5	3.2		
	L110-105	5′-ATTTTCTC-3′	9.6×10^5	5.4		
PyQK-10	U85-90	5′-ATACCG-3′	9.1×10^5	2.5	107–112	Partial-circuit
	U107-117	5′-GAAAATACCG-3′	9.9×10^5	3.1		
	L112-107	5′-ATTTTC-3′	1.0×10^6	3.1		

K_a and n_H are the apparent association constant and Hill coefficient determined from concentration-dependent DNase I footprinting studies, respectively. Binding sites on the upper and lower strands are identified by U and L, respectively.

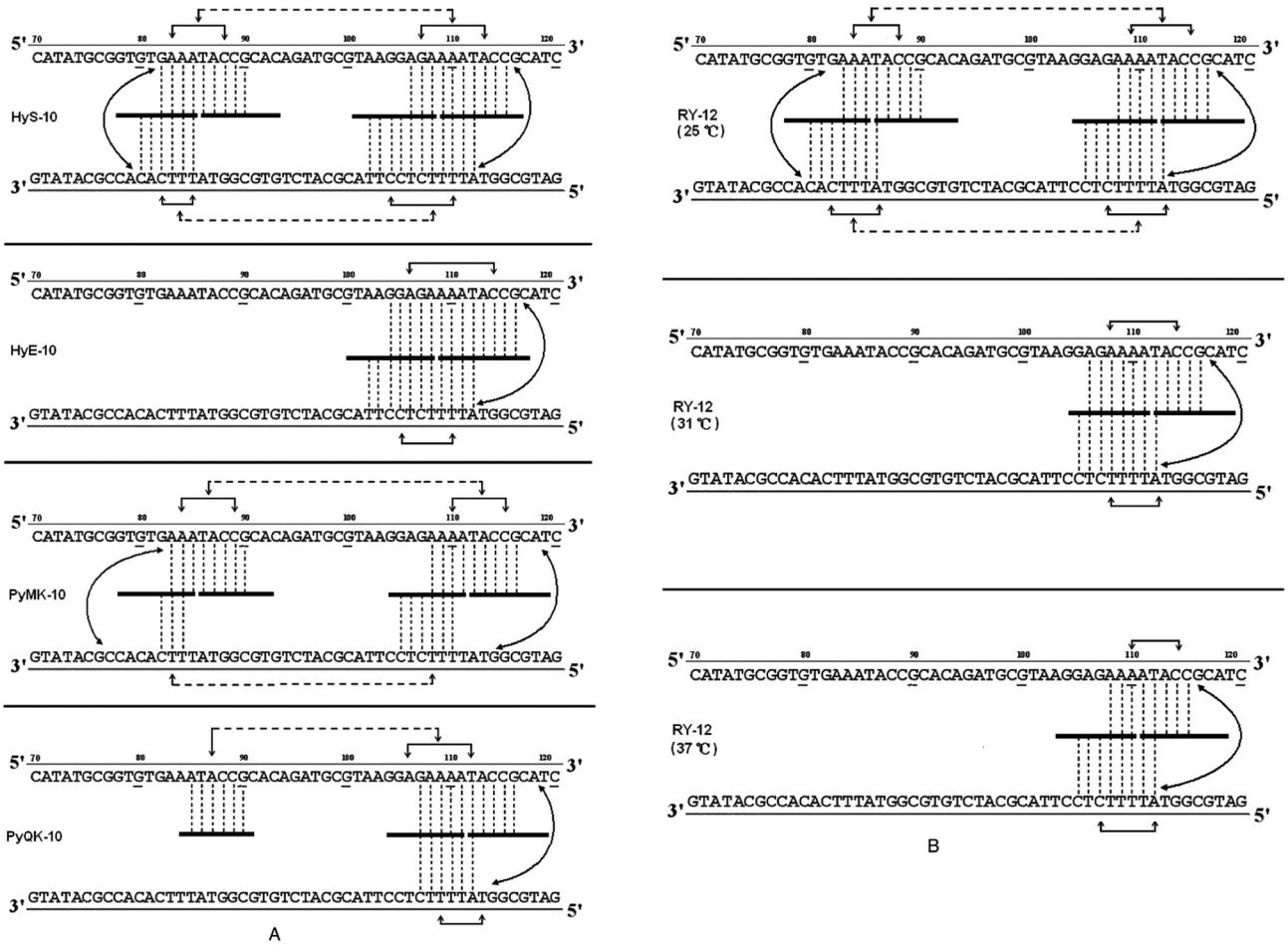


Fig. 3 (A) Proposed allosteric interaction network models for cooperative binding of peptides HyS-10, HyE-10, PyMK-10, and PyQK-10 to pBR322 fragments based on quantitative footprinting experiments. The portion of the ligand binding to each DNA site/sub-site is represented by a thick horizontal line. Monodentate interactions and interstrand bidentate interactions are represented by vertical broken lines. The solid horizontal arrow lines represent communication of the allosteric interaction between DNA sub-binding sites. Broken horizontal arrow lines between neighboring binding loci some 12–16 nucleotides apart are intended to represent moderate or weak cooperative communication. (B) Proposed allosteric interaction network models for the cooperative binding of peptide RY-12 to pBR322 fragments at different incubation temperatures based on quantitative footprinting results. Representations as in the legend to (A).

communications may occasionally be transmitted to a more remote binding site on the same DNA strand, spanning as many as 12–16 base pairs. Recent footprinting experiments[18] show that the n_H values for peptides PyHyp-12 and PyHyp-9 binding to 81-mer duplexes (S-81) containing a single d(AAAA)-d(TTTT) binding locus are significantly higher than those of the 158-mer and 135-mer duplexes, which contain two binding loci d(AAAA)-d(TTTT) and d(AAA)-d(TTT). This finding suggests that the allosteric relay of peptide binding to neighboring d(AAA)-d(TTT) and d(AAAA)-d(TTTT) positions spanning an intervening sequence of 12–14 base pairs on the same DNA strand may be mediated *via* a negative cooperative effect.

The allosteric interaction network models support the idea that binding of decapeptides HyS-10 and PyMK-10 to pBR322 fragments possessing multiple sites proceeds *via* a circuit type[17,18] of process within a network system: the communications form a closed circuit (Fig. 3A). In the circuit type of interaction network, strong intrastrand allosteric communications are conveyed by conformational changes between two peptide molecules bound to two sub-sites of a wide binding site (Fig. 3A and B). Allosteric communications can also be relayed between complementary sites from one strand to another through interstrand bidentate interactions. Weak intrastrand allosteric communications may persist between two widely separated binding sites spanning 12–15 base pairs as, for example, the binding of peptide HyS-10 to two neighboring DNA loci: 5′-AGAAAATACC-3″ and 5′-GAAATACCG-3′ at positions U106-115 and U82-90, respectively (Fig. 3A). On the other hand, peptide HyE-10 fails to bind to the d(AAA)-d(TTT) site, and peptide PyQK-10 refuses to bind the d(TTT) site in the lower strand, both resulting in a partial or incomplete circuit network.[17,18]

It is interesting that temperature has a notable effect on the allosteric interaction behavior of peptide RY-12 (Fig. 3B). At 25 °C, the peptide engages in a circuit type of network interaction. At 31 °C and 37 °C, it switches to a partial-circuit type of network in which the allosteric communications form a partial or incomplete circuit. By contrast, previous studies revealed that the interaction network of peptides $(HPRK)_3NH_2$ [HR-12][16] and $(SPRK)_3NH_2$ [SP-12][16] binding to DNA fragments is devoid of interstrand bidentate interactions and consequently is referred to as a non-circuit type of allosteric communication. Thus, quantitative footprinting results in this work as well as other studies further support our recent hypothesis[17,18] of three different types of allosteric interaction networks in peptide–DNA molecular recognition.

To gain further evidence that conformational changes are involved in the molecular recognition process, we employed circular dichroism spectroscopy (CD) to investigate changes in the helical structure associated with peptide–DNA binding. A 13-mer deoxyribonucleotide duplex was prepared, d(TAGGAGAAAATAC)-d(GTATTTTCTCCTA) (U4A-L4T), in which 12 base pairs correspond to the complementary recognition site of the 158-mer fragment at positions 104–114. With this duplex a negative CD band occurs around 248 nm and a positive band around 276 nm (Fig. 4A). Peptide HyS-10 produces a dose-dependent negative CD enhancement around 248 nm as well as two strong positive CD enhancement bands around 276 nm and 328 nm (Fig. 4B). In the difference spectra the negative band around 248 nm can be seen to be red-shifted to 267 nm and the two strong positive bands appear around 288 nm and 332 nm (Fig. 4C). A near-isoelliptic point occurs at 278 nm, consistent with a two-component binding process. The induced positive band around 332 nm is peptide concentration-dependent and free from interference by native DNA and peptide, so it can be used for plotting binding isotherms. The significant enhancement of CD around 332 nm by HyS-10 and other peptides[17,18] indicates that the principle interaction lies in the minor groove of DNA.[11–13]

The CD spectra seen with decapeptides HyE-10 (Fig. 4D and E), PyMK-10 and PyQK-10 (Fig. ii(a–d), ESI†) are similar to that of peptide HyS-10. All four peptides induce a strong positive CD band around 330 nm, which is correlated with peptide–DNA binding stoichiometry. Fig. 5A and B show the variation of CD with [peptide]/[duplex 4A-4T] ratio. Peptide HyE-10 induces the greatest ellipticity change among the three peptides, suggesting that more drastic conformational changes in the DNA minor groove accompany the binding of this peptide. At peptide/DNA ratios between 0.5 and 2.0 it is likely that one molecule of each peptide binds to the d(AAAA)-d(TTTT) locus of the minor groove, as suggested in some plots by a hint of a plateau. At [peptide]/[duplex] ratios between 2 and 3 small plateau regions do appear in some of the titration curves, indicating that two peptide molecules bind in a dimeric fashion to the binding site in the minor groove. As the [peptide]/[duplex] ratio rises above 3, progressive, usually slower, increases occur in $\Delta\theta$, suggesting that further peptide molecules start to bind non-specifically—probably to the wide major groove, which can loosely accommodate more peptide molecules than the minor groove (Fig. 5A and B). The CD observations concur with the DNase I footprinting results in suggesting that dimeric peptide binding to the d(AAAA)–d(TTTT) locus is often favored, consistent with the very long binding sites spanning over 9–14 base pairs. Subsequent isothermal titration calorimetric experiments also agree that dimeric peptide binding to the d(AAAA)–d(TTTT) locus may be favored, as described below. The dramatic drop in $\Delta\theta$ after the 2 : 1 peptide/DNA molar ratio induced by peptide PyMK-10 is unusual, and might indicate dissociation of this peptide from the minor groove at higher concentrations (Fig. 5B).

Three DNA–peptide incubation temperatures were investigated in CD studies with the dodecapeptide RY-12 (Fig. iii(a)–(f), ESI†). In the plot of $\Delta\theta$ *versus* [peptide]/[DNA] (Fig. 5C), it is evident that the change in ellipticity induced by peptide RY-12 measured at 327 nm increases steadily as the incubation temperature is lowered from 37 °C to 25 °C, indicating that the induced conformational change in the DNA minor groove is more pronounced at lower temperatures. These CD results are again in accord with the quantitative footprinting data. Recall that at 25 °C, peptide RY-12 engages in a circuit type of allosteric interaction network (marked DNA conformational changes for conveying allosteric communications), whereas at

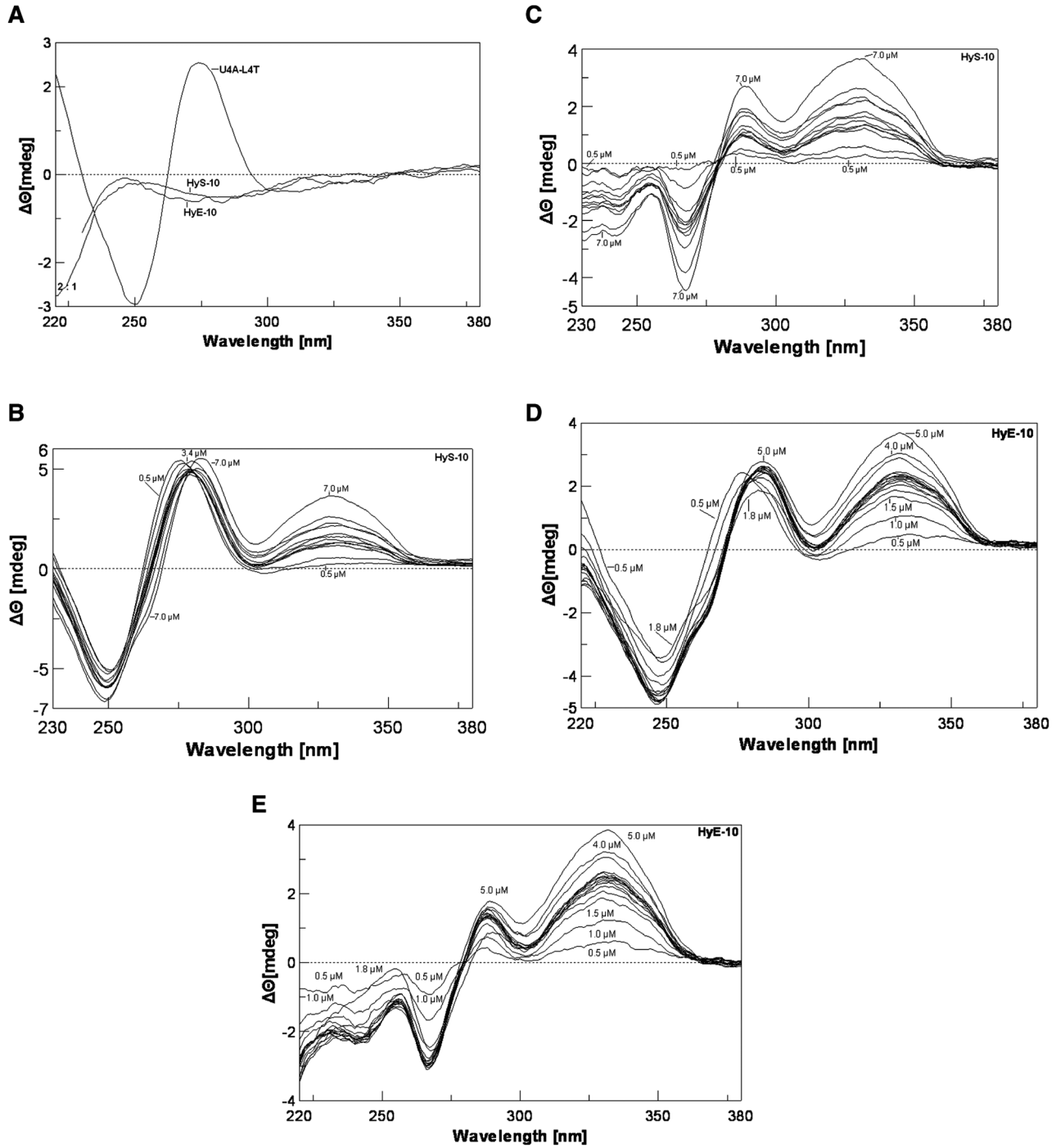


Fig. 4 Panel A: CD spectra of DNA duplex U4A-L4T alone and peptide alone at 37 °C. Panel B: titration of duplex U4A-L4T *versus* peptide HyS-10 at peptide concentrations of 0.5, 1.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, 5.0, 7.0 μM at 37 °C. Panel C: corresponding CD difference spectra with the contribution of free duplex and peptide HyS-10 subtracted. Panel D: titration of U4A-L4T *versus* peptide HyE-10 at peptide concentrations of 0.5, 1.0, 1.5, 1.8, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 3.0, 3.5, 4.0, 5.0 μM at 37 °C. Panel E: corresponding difference spectra with the contribution of free duplex and peptide HyE-10 subtracted.

31 °C and 37 °C the allosteric networks for this peptide switch to partial-circuit type (less DNA conformational change to power allosteric interactions).

To assess the effect of ionic strength on peptide–DNA interactions, we monitored the effect of sodium chloride on the binding of a decapeptide incorporating both XHypRK and Py motifs, Met-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyM-10), to the U4A-L4T duplex using circular dichroism. A significant, monotonic decrease in ellipticity at 320 nm was observed as the concentration of NaCl was increased (Fig. 5D), indicating a near-linear relationship (Fig. 5E). (Note again the hints of small plateaux round about the molar ratios of 1.0 and 2.0.) The

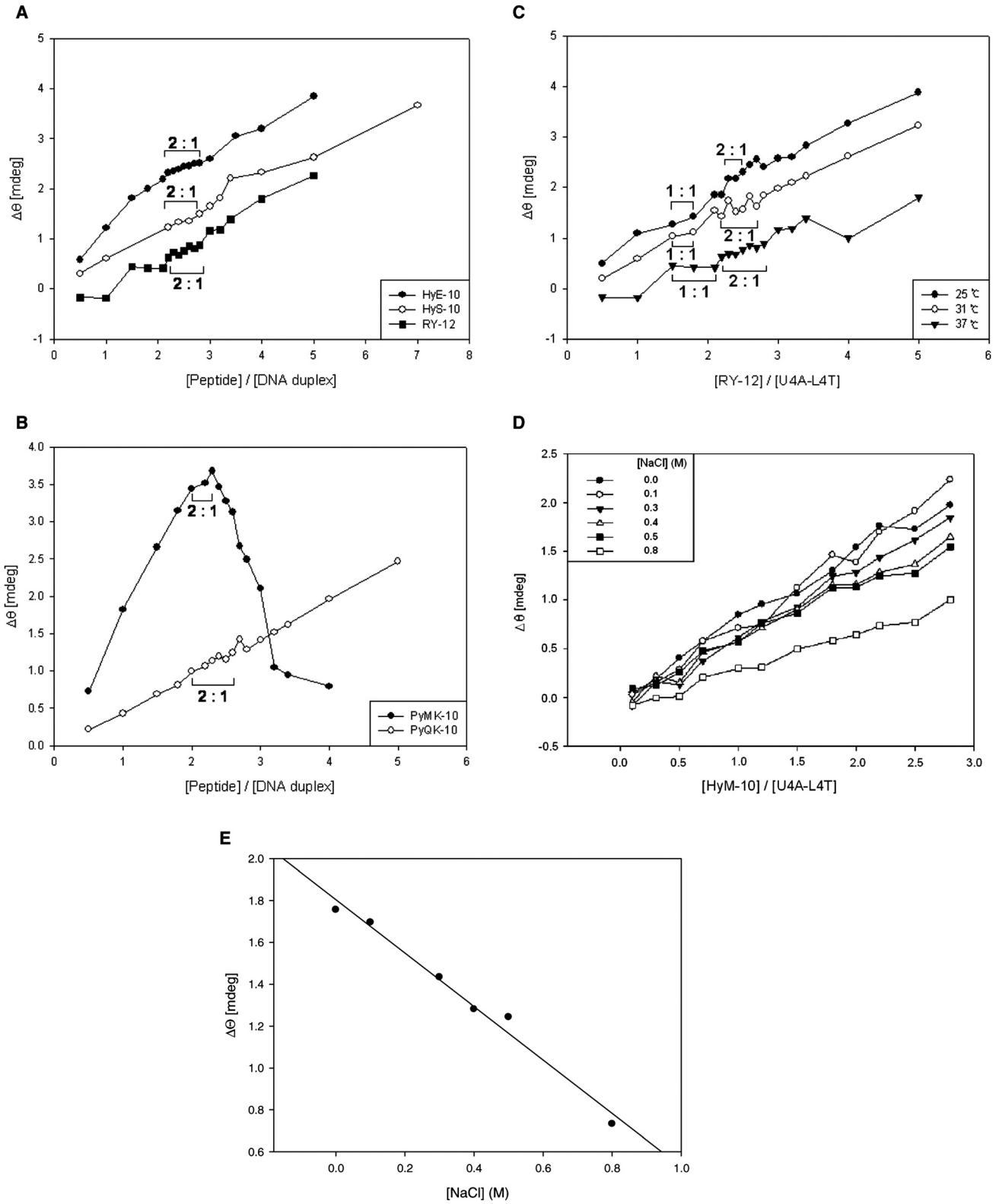


Fig. 5 Panel A: plot of CD intensity at 322 nm *versus* [peptide]/[duplex] for the titration of duplex U4A-L4T with peptides HyS-10, HyE-10, and RY-12 at 37 °C. The proposed stoichiometric binding ratios are as indicated, with binding <2 : 1 considered to be predominantly 1 : 1. Panel B: plot of CD at 322 nm *versus* [peptide]/[duplex] for PyMK-10 and PyQK-10. Panel C: plot of CD at 322 nm *versus* [RY-12]/[duplex] at incubation temperatures of 25 °C, 31 °C and 37 °C. Panel D: plot of CD intensity at 322 nm *versus* [peptide]/[duplex] at various sodium chloride concentrations at 37 °C. Panel E: plot of CD intensity at 322 nm *versus* sodium chloride concentration.

obvious interpretation is that the conformation of the DNA minor groove in the peptide complex is significantly changed upon increasing the salt concentration, supporting the view that ionic interactions play an important role in the sequence-selective binding of peptides incorporating the XHypRK and Py motifs.

With some understanding of the interaction network(s) underlying cooperative sequence-selective binding of XP(Hyp)-RK peptides to DNA, we went on to investigate their energetic basis. Isothermal titration calorimetry (ITC) furnishes a direct means to study the thermodynamics of binding of peptides like the five described here. For purposes of comparison, another seven peptides previously reported were included, and the chosen DNA substrate was the oligonucleotide duplex U4A-L4T (Table 3). The extra peptides included four more decapeptides, HyH-10, PyHK-10, HyM-10 and HyQ-10, and three more dodecapeptides: RPyH-12, PyHyp-12, and RHyp-12 (structures in Table 3). Representative ITC results for peptides HyS-10, HyE-10, PyMK-10 and PyQK-10 are shown in Fig. 6A; other ITC data are presented in Table 3 and in Fig. iv (ESI†). The relative concentrations of peptide and DNA duplex in each titration had to be adjusted carefully in order to produce satisfactory plots of integrated heat *versus* peptide/DNA molar ratio.

The ITC plots for all twelve peptides *versus* U4A-L4T reveal a peptide/DNA binding stoichiometry near 2 : 1, suggesting that the peptides become bound to the model duplex in dimeric fashion, neatly in agreement with the footprinting and circular dichroism results. Moreover, all twelve peptides showed predominantly exothermic behavior in the binding reaction: ΔH was found to vary between −14.7 kcal mol^{-1} and −74.4 kcal mol^{-1}, with most peptides yielding ΔH values greater than −26 kcal mol^{-1} measured at 25 °C (Table 3). Compared to the moderate magnitude of ΔH reported in the pioneering studies of Breslauer[28] and his colleagues on the binding of netropsin, distamycin, ethidium and daunorubicin to duplex poly-[d(A-T)]·[d(A-T)], where ΔH lies within the range −8.9 to −18.5 kcal mol^{-1}, the ITC results in this study suggest that binding of the XP(Hyp)RK and KPXR peptides to DNA is overwhelmingly enthalpy-driven. However, the opposing entropy effects for all of the XP(Hyp)RK and KPXR peptides are strikingly unfavorable: $T\Delta S$ falls within the range −6.5 to −65.9 kcal mol^{-1}. In consequence, apparent net enthalpy–entropy compensation (EEC) effects (Fig. 6B) sustain a binding free energy change ΔG between −8.2 and −9.1 kcal mol^{-1} for all twelve peptide–DNA interactions. The linearity of ΔH *versus* $T\Delta S$ is impressive (Fig. 6B): slope = 1.0061; *y*-intercept = −8.5 kcal mol^{-1}. It originates from unusually high values of ΔH and ΔS that compensate each other. But by contrast, the individual relationship between ΔH and ΔG, and that between ΔG and $T\Delta S$, is non-linear and unpredictable (Fig. v, ESI†).

Much controversy and interest has surrounded the EEC phenomenon,[29] with various theoretical explanations adumbrated, yet the physical basis of EEC has not been convincingly elucidated. The overwhelmingly large magnitudes of ΔH and ΔS measured by ITC for DNA–peptide interaction in this study seem to offer a unique advantage for investigating the energetic basis of the compensation.

According to Williams and co-workers,[30] the binding free energy ΔG can be parsed according to the following equation:

$$\Delta G = \Delta G_{t+r} + \Delta G_r + \Delta G_h + \Sigma\Delta G_p + \Delta G_{conf.} + \Delta G_{vdw} \quad (1)$$

where ΔG_{t+r} refers to the unfavorable cost of association when a ligand binds to a receptor, losing its translational (t) and rotational (r) entropy considerably, relative to the receptor. ΔG_r refers to the energy cost of *n* internal rotors. The term ΔG_h represents the benefit of the net hydrophobic effect and $\Sigma\Delta G_p$ expresses the benefit from the formation of hydrogen bonding and ionic interactions. $\Delta G_{conf.}$ is the total conformational

Table 3 Thermodynamic parameters from titration of peptides *versus* oligonucleotide duplex (U4A-L-4T) by isothermal titration calorimetry

Entry	Peptide (mM)	DNA (mM)	ΔH (kcal mol^{-1})	ΔS (cal mol^{-1})	$T\Delta S$ (kcal mol^{-1})	ΔG (kcal mol^{-1})	K_a (M^{-1})
1	HyM-10 (0.15)	0.0033	−50.5	−140	−41.7	−8.8	2.9×10^6
2	HyQ-10 (0.15)	0.01	−74.4	−221	−65.9	−8.6	1.9×10^6
3	HyS-10 (0.15)	0.006	−31.2	−74.5	−22.2	−9.0	4.2×10^6
4	HyE-10 (0.15)	0.007	−27.8	−64.6	−19.3	−8.5	1.7×10^6
5	HyH-10 (0.05)	0.0033	−51.2	−142	−42.3	−8.8	2.9×10^6
6	PyHK-10 (0.2)	0.0066	−14.7	−21.8	−6.5	−8.2	9.8×10^5
7a (20 °C)	RY-12 (0.05)	0.0033	−17.6	−30.7	−9.0	−8.6	2.7×10^6
7b (25 °C)	RY-12 (0.15)	0.01	−35.0	−89.7	−26.7	−8.2	1.0×10^6
7c (28 °C)	RY-12 (0.15)	0.008	−28.6	−66.4	−20.0	−8.6	1.6×10^6
7d (31 °C)	RY-12 (0.05)	0.0033	−35.0	−86.7	−26.4	−8.6	1.6×10^6
8	RPyY-12 (0.05)	0.0045	−60.5	−173	−51.6	−8.9	3.3×10^6
9	PyHyp-12 (0.05)	0.004	−33.2	−80.9	−24.1	−9.1	4.4×10^6
10	RHyp-12 (0.15)	0.005	−26.0	−59.1	−17.6	−8.4	1.4×10^6
11	PyMK-10 (0.15)	0.005	−19.4	−36.5	−10.9	−8.5	1.8×10^6
12	PyQK-10 (0.15)	0.006	−31.0	−74.9	−22.3	−8.7	2.2×10^6

All ITC experiments were performed at 25 °C unless otherwise stated. The relative concentrations of peptide and DNA in each titration were adjusted in order to produce satisfactory plots of integrated heat *versus* peptide/DNA molar ratio. The amino acid sequences of the peptides are: Met-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (HyM-10), Gln-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (HyQ-10), Ser-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (HyS-10), Glu-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (HyE-10), His-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (HyH-10), His-Pro-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (PyHK-10), Tyr-Pro-Arg-Lys-(Py)$_4$-Lys-Arg-Pro-Tyr-NH$_2$ (RY-12), His-Hyp-Arg-Lys-(Py)$_4$-His-Hyp-Arg-Lys-NH$_2$ (PyHyp-12), His-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-Hyp-His-NH$_2$ (RHyp-12), His-Pro-Arg-Lys-(Py)$_4$-Lys-Arg-Pro-Tyr-NH$_2$ (RPyY-12), Lys-Pro-Met-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (PyMK-10), Lys-Pro-Gln-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (PyQK-10).

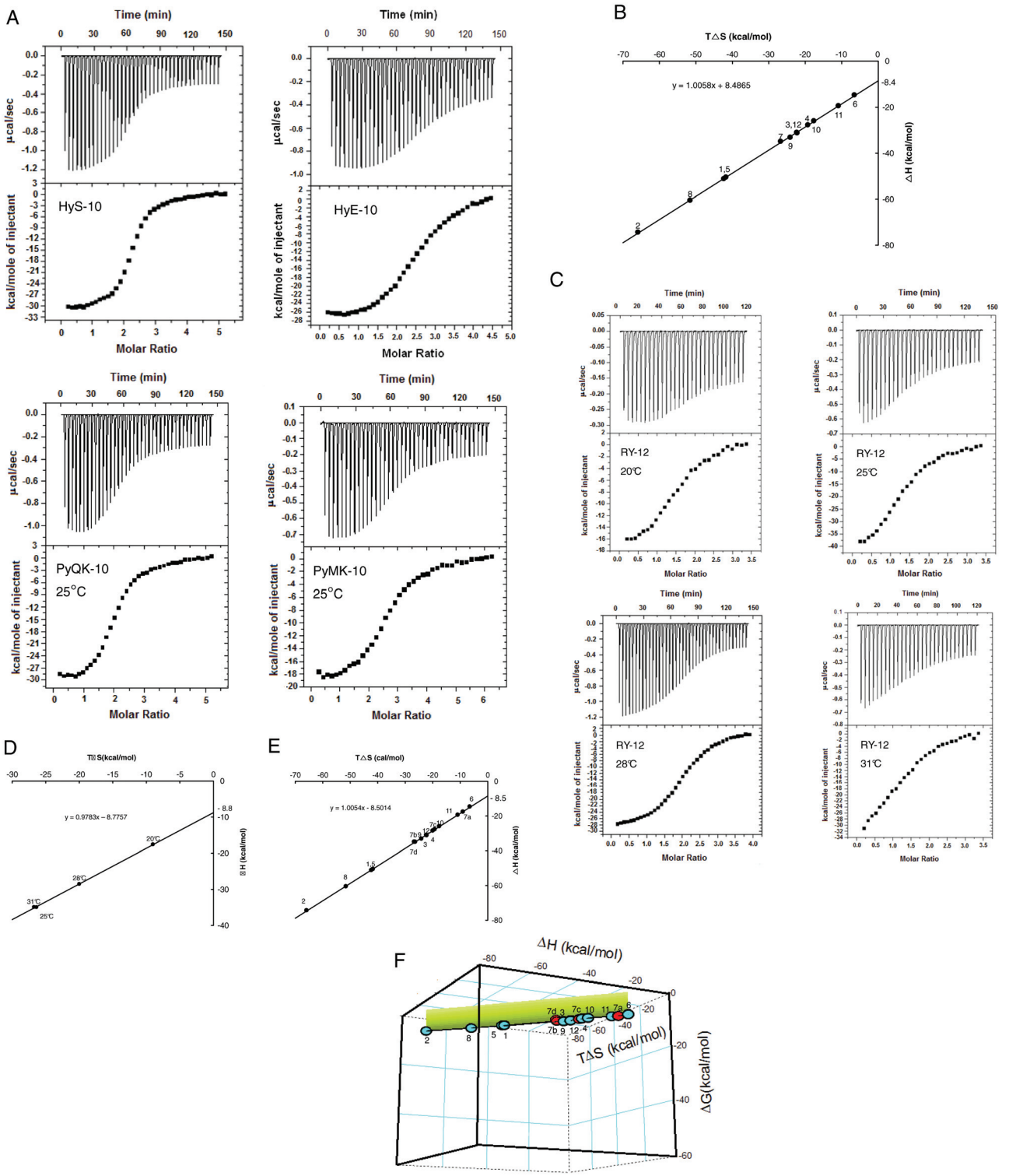


Fig. 6 (A) ITC curves for titration of decapeptides HyS-10, HyE-10, PyMK-10 and PyQK-10 into the U4A-L4T duplex at 25 °C. For each experiment the top panel represents the raw heat of binding generated by successive additions of peptide, and in the bottom panel the integrated heat is plotted *versus* peptide/DNA molar ratio. Data acquisition and analysis were performed using nonlinear least-squares fitting algorithm software (Microcal Origin 7.1). (B) Plot of enthalpy *versus* entropy ($T\Delta S$) from ITC of twelve peptides added to the U4A-L4T duplex at 25 °C (298 K). (C) ITC curves for the titration of peptide RY-12 into the U4A-L4T duplex at various incubation temperatures (legend as in (A)). (D) Plot of enthalpy *versus* entropy ($T\Delta S$) from ITC of peptide RY-12 added to the U4A-L4T duplex at different temperatures. (E) Combined enthalpy–entropy plot from titrating all twelve peptides into the U4A-L4T duplex at 25 °C (with peptide RY-12 at various temperatures: 7a, 20 °C; 7b, 25 °C; 7c, 28 °C; 7d, 31 °C). (F) Three-dimensional plot of ΔH, $T\Delta S$ and ΔG from titration of all twelve peptides into the U4A-L4T duplex. The short vertical projection plane (in green) corresponds to ΔG magnitudes sustained around −8.5 kcal mol^{-1}. The figure represents a view of the 'cube' from below.

strain energy generated upon binding, and ΔG_{vdw} is the van der Waals energy difference between the free and bound states.

Eqn (1) was applied to the present peptide–receptor interaction data. Circular dichroism studies of the DNA–small molecule interaction, in this work as well as others, show that marked conformational changes in the DNA minor groove occur upon peptide binding. Thus, the free energy components ΔG_{t+r} and ΔG_r must include the total translational and rotational entropy of both the DNA duplex and the ligand. The strong negative entropy changes seen with all ten peptides upon binding to DNA suggest that there are significant losses in the rotational and translational motion of the DNA duplex as well as the peptide. Multiple intrastrand and interstrand bidentate interactions between the peptide moieties and DNA bases revealed by our footprinting experiments (Table 1, Fig. 1 and 3), together with ionic interactions between the positively charged Arg and Lys side chains and DNA phosphates as evidenced by CD studies (Fig. 5E), are consistent with considerable loss in translational and rotational entropy of the peptide and DNA duplex upon binding, as will be seen.

Since the net binding free energy ΔG of the twelve peptides titrated against the U4A-L4T duplex lies rather constant within the range −8.2 to −9.1 kcal mol^{-1}, the average value of ΔG for the peptides can be taken as being around −8.5 kcal mol^{-1}. Indeed, in a plot of ΔH *vs.* $T\Delta S$, the enthalpy–entropy compensation ratio k_e (slope of the straight line) for the twelve peptides containing the XP(Hyp)RK and Py motifs approaches unity (slope = 1.006), with the *y*-axis intercept equal to the average ΔG value: −8.53 kcal mol^{-1} (Fig. 6B). Put mathematically,

$$\Delta H = k_e T\Delta S - 8.5$$

When k_e approaches unity,

$$\Delta H = -8.5 + T\Delta S \quad (2)$$

Eqn (2) well describes the strict linear dependence of ΔH on $T\Delta S$ for the twelve peptide–DNA interactions studied and hence underlines the enthalpy–entropy compensation phenomenon. The almost exact compensation of energy between the favorable negative ΔH and unfavorable negative ΔS energy components is what sustains the almost constant value of net binding free energy ΔG. According to eqn (2), maintenance of a negative ΔG of acceptable magnitude seems to be crucial for DNA–peptide binding; any decrease of ΔG (less negative) would require a lowering of ΔH (less negative) and a concomitant change in ΔS (more negative)—both unfavorable for the binding process.

Since the dodecapeptide RY-12 displayed notable differences in binding preference at lower temperatures, an attempt was made to perturb the enthalpy–entropy compensation by varying the incubation temperature (Fig. 6C). Accordingly, peptide RY-12 was titrated against duplex U4A-L4T at 20, 25, 28, and 31 °C (unstable calorimetric readings were recorded at 37 °C). Strikingly, the ΔH and $T\Delta S$ values measured for peptide RY-12 over this temperature range were still found to be linearly related, with a slope of 0.98. Although the temperature variation undoubtedly produces significant deviations in ΔH and $T\Delta S$ individually, the ΔG values still lie within the range −8.2 to −8.6 kcal mol^{-1} (Table 3). Plotting enthalpy *versus* entropy for all twelve peptides binding to the model duplex, including the data for peptide RY-12 at different temperatures (Fig. 6E), still gives a straight line of slope 1.005, so evidently the global value of ΔG does remain remarkably constant. To underline the EEC phenomenon in the DNA–peptide interaction, it is instructive to plot a 3-dimensional figure of ΔH, $T\Delta S$ and ΔG. The result is a near-straight line suspended in thermodynamic space, nicely illustrating the linear enthalpy–entropy-free energy relationship (Fig. 6F). The conclusion is that stringent constraints act in nature to preserve EEC so as to maintain favorable ΔG within a constant or narrow range for successful binding of ligands to DNA.

On the other hand, we know that temperature affects the nature of allosteric interaction networks on the pBR322 fragments because of the quantitative footprinting experiments (Fig. 3), which also establish that binding of peptides to the d(AAAA)-d(TTTT) and d(AAA)-d(TTT) sites on the pBR322 fragments occurs with significant positive cooperativity (Table 1). These DNA–peptide interaction results are consistent with Cooper and Dryden's suggestion[31] that positive cooperativity in ligand binding will induce a 'stiffening' or tightening effect in the protein/receptor, initiated by the loss of many internal vibrational degrees of freedom. Williams and co-workers[30,32] concluded that highly exothermic binding is associated with positive cooperativity, allied to improved bonding within the receptor in its drug-bound state, which would account for the very unfavorable negative entropy change. As predicted by Williams[30] positive cooperativity and negative cooperativity can coexist in the same system. Our previous studies show that DNA binding of the peptide PyPro-12[18] does involve concomitant positively and negatively cooperative site binding, although positive cooperativity is always predominant for peptides incorporating the XPRK motifs. If positive cooperativity is associated with improved bonding within the DNA duplex upon peptide binding, it is most likely related to the existence of multiple intrastrand monodentate interactions and interstrand bidentate interactions between the peptide moieties and the DNA bases. Doubtless ionic forces between the positively charged Arg and Lys side chains of the peptides and DNA phosphates also restrict the conformational mobility of the DNA duplex and thereby help to account for the dramatic decrease in total entropy. According to the law of conservation of energy, any loss in entropic energy must transform to exothermic energy, which in turn causes an increase in enthalpy. On the other hand, high exothermic energy promotes allosteric adjustment of DNA conformation within the binding loci by means of hydrogen bonding and ionic interactions between the DNA duplex and peptide, augmenting the favorable energy components ΔG_h, $\Sigma\Delta G_p$, $\Delta G_{conf.}$, and ΔG_{vdw}, and facilitating positively cooperative binding of approaching

peptide molecules. This is fully consistent with the proposal of Williams[30] that highly exothermic binding is associated with significant positive cooperativity, and is verified by our discovery of positively cooperative peptide binding to multiple DNA recognition sites by DNase I footprinting experiments in this study (Tables 1 and 2) as well as previously.[17,18] This communication of cooperativity between multiple binding loci forms the basis of the DNA–peptide interaction networks proposed here.

3. Conclusion

An important outcome of the present study stems from the ITC experiments on twelve peptides, showing that a dramatic decrease in entropy occurs upon binding to DNA, consistent with the presence of multiple monodentate and bidentate interactions, as well as ionic forces that mediate positively cooperative binding, resulting in a 'tightening' effect of DNA conformation. Allosteric communication between multiple DNA loci *via* cooperative peptide binding establishes the basis of the DNA–peptide interaction networks described here. Temperature-dependence studies with peptide RY-12 using quantitative DNase I footprinting are in accord with the results of CD experiments, namely that a lower temperature (25 °C) favors a circuit type of allosteric interaction network demanding more drastic DNA conformational changes, whereas higher temperatures (31 and 37 °C) afford only a partial-circuit type of interaction network that requires fewer DNA conformational changes. The enthalpy–entropy compensation effects associated with peptide RY-12 binding to the 13-mer DNA duplex persist through various temperatures, while the magnitude of ΔH and $T\Delta S$, as well as the type of allosteric interaction networks seen with RY-12 on the 135-mer and 158-mer DNA duplexes, are changed. The unusually high values of negative ΔH recorded for DNA–peptide binding furnish exothermic energy that facilitates conformational adjustment of DNA binding loci to promote positively cooperative binding of approaching peptide molecules. The net energy compensation between the favorable negative ΔH and unfavorable negative ΔS components sustains the value of net binding free energy ΔG within a remarkably constant range. We derive a linear equation which effectively describes the strict linear dependence of ΔH on $T\Delta S$ for the twelve peptide–DNA interactions studied, and hence underlines the enthalpy–entropy compensation phenomenon: $\Delta H = -8.5 + T\Delta S$.

Thus it appears that the preservation of a negative ΔG value of acceptable magnitude is central to the EEC of many DNA–peptide interactions, and hence the findings illuminate the universal EEC paradox commonly manifested in biochemical reactions. In other words, stringent restraints affecting enthalpy–entropy compensation are conserved in nature to sustain a favorable ΔG value within a narrow range for the successful binding of ligands to DNA. We propose that the emergence of allosteric interaction networks featuring significant positive cooperativity in many DNA–small ligand interactions is also aimed at assuring an optimum binding free energy. These insights will foster the design of new DNA–binding small molecules for further study of DNA–peptide allosteric interactions as well as interaction networks.

4 Experimental

4.1 Chemicals and biochemicals

All the protected amino acid derivatives were purchased from Bachem California (Torrance, CA) and AnaSpec, Inc. (San Jose, CA) or synthesized in our own laboratory according to published procedures. All other analytical reagents were purchased from Acros, Tedia or Sigma. The radiolabelled nucleoside triphosphate [γ-^{32}P]ATP was obtained from NEN Life Science Products at a specific activity of 6000 Ci mmol^{-1}. *Taq* polymerase, T4 polynucleotide kinase, and DNase I were purchased from Promega. All other chemicals were analytical grade reagents, and all solutions were prepared using deionized, Millipore-filtered water.

4.2 Chemical methods

Melting points were determined on a Mel-Temp apparatus (Cambridge, Mass) and are uncorrected. Optical rotations were determined on a Rudolph Autopol II instrument. Semi-preparative and analytical HPLC (Vydac reversed-phase columns,

Table 2 Sequence-selectivity and physicochemical parameters for the binding of peptide RY-12 to recognition sites in complementary upper and lower DNA strands at various incubation temperatures determined by quantitative DNase I footprinting

Incubation temperature	Binding site position	Recognition sequence	K_a	n_H	Position of interstrand bidentate interactions	Type of interaction network
25 °C	U84-90	5′-AATACCG-3′	1.6×10^6	3.4	84–86	Circuit
	U107-116	5′-GAAAATACCG-3′	1.9×10^6	2.9	107–112	
	L86-80	5′-ATTTCAC-3′	2.1×10^6	3.4		
	L112-105	5′-ATTTTCTC-3′	3.3×10^6	2.7		
31 °C	U106-116	5′-AGAAAATACCG-3′	2.1×10^6	2.8	106–112	Partial-circuit
	L112-105	5′-ATTTTCTC-3′	3.5×10^6	2.2		
37 °C	U108-115	5′-AAAATACC-3′	1.9×10^6	1.4	108–112	Partial-circuit
	U84-90	5′-AATACCG-3′	—	—		
	L112-105	5′-ATTTTCTC-3′	1.6×10^6	2.4		

Details as in the footnote to Table 1.

TP201; column 1, 1 × 25 cm; column 2, 0.4 × 25 cm) were performed using a Hitachi L-7100 pump equipped with a gradient elution device and a Soma S-3702 variable wavelength UV detector connected to a PC computer installed with Hitachi HPLC analytical software. Mass spectra were determined with a Finnigan/Thermo Quest MAT 95XL instrument operating in the electrospray ionization (ESI) mode in Chung-Hsing University.

Ser-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyS-10). This peptide was synthesized using solid phase methodology by manual operation of a Protein Technology PS3 peptide synthesizer. The first Fmoc-protected amino acid was coupled to the Nova Rink amide AM resin using PyBOP/NMM in DMF. All of the N_α-Fmoc-protected amino acids (in 4 equivalent ratio excess to the resin) were coupled in a stepwise fashion using PyBOP/NMM in DMF after deprotection of the N_α-Fmoc group by piperidine. The side chains of Arg, Lys and Glu were protected by the Pmc, Boc, and *t*-Bu groups, respectively, whereas the side chains of Ser and Lys were protected by the Trt group. After coupling the last N-terminal Fmoc-amino acid residue, the resin was treated with the cleavage reagent (0.75 g phenol, 10 mL TFA, 0.5 mL thioanisole, 0.25 mL EDT) for 1.5 h, and then lyophilized. The resin was washed with dry ether (2 × 30 mL), filtered, and then washed with 5% acetic acid (200 mL). The combined filtrate was lyophilized and the product was purified by semi-preparative reversed-phase HPLC (column 1) using gradient elution. Eluent A: 5% MeCN, 95% H_2O, 0.1% TFA; eluent B: 95% MeCN, 5% H_2O, 0.1% TFA. A linear gradient was achieved by increasing the MeCN content from eluent A to eluent B in 30 min. t_R (column 2), 14.05 min. m.p. 153–156 °C, $[\alpha]_D^{27}$ −9.25 (*c* 0.072, H_2O); ESIMS requires: 1274.44, found: 1274.0. All other peptides were synthesized and purified to homogeneity using a similar procedure as for peptide HyS-10.

Glu-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (**HyE-10**). This peptide was synthesized and purified by semi-preparative reversed-phase HPLC (column 1), using a similar procedure as for peptide HyS-10. t_R (column 2), 14.89 min, m.p. 150–152 °C, $[\alpha]_D^{27}$ −34.43 (*c* 0.203, MeOH–H_2O, 1 : 1); ESIMS requires: 1316.47, found: 1316.0.

Lys-Pro-Met-Arg-$(Py)_4$-Lys-Arg-NH_2 (**PyMK-10**). This peptide was synthesized and purified by semi-preparative reversed-phase HPLC (column 1), using a similar procedure as for peptide HyS-10. t_R (column 2), 12.03 min, m.p. 155–159 °C, $[\alpha]_D^{31}$ −39.21 (*c* 0.102, H_2O); ESIMS calcd 1301.7, found: 1302.7.

Gln-Pro-Gln-Arg-$(Py)_4$-Lys-Arg-NH_2 (**PyQK-10**). This peptide was synthesized and purified by semi-preparative reversed-phase HPLC (column 1), using a similar procedure as for peptide HyS-10. t_R (column 2), 11.29 min, m.p. 157–159 °C, $[\alpha]_D^{25}$ −41.67 (*c* 0.048, H_2O); ESIMS calcd 1298.72, found: 1299.49.

Tyr-Pro-Arg-Lys-$(Py)_4$-Lys-Arg-Pro-Tyr-NH_2. This peptide was synthesized using a similar procedure as for HyS-10. The crude product was purified by semi-preparative reversed-phase HPLC (column 1), using gradient elution as described for the purification of peptide HyM-10. t_R (column 2), 12.4 min, m.p. 157–161 °C, $[\alpha]_D^{31}$ −51.95 (*c* 0.077, H_2O); ESIMS requires: 1593.84, found: 1594.72.

4.3 Polymerase chain reaction (PCR) and end-labeling of PCR products

The 158-mer DNA duplex (Watson strand 5′-^{32}P-labeled) and 135-mer DNA duplex (Crick strand 5′-^{32}P-labeled) were prepared by PCR amplification in a thermal cycler (ABI model 9700), as reported previously.[14] The DNA concentration was determined by UV spectroscopy to lie in the range 600–800 nM.

4.4 DNase I footprinting

Reactions were carried out following procedures as described in detail previously.[17,18]

4.5 Circular dichroism (CD) studies

CD spectra were measured at 37 °C with a Jasco J-815 instrument in the Institute of Chemistry, Academia Sinica. The duplex DNA was adjusted to 1.0 μM in 5 mM sodium cacodylate buffer (pH 6.5) and peptides, dissolved in the same buffer, were added to maintain final concentrations of 0.2, 1.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0 and 5.0 μM. Spectra were recorded after 60 min incubation at 37 °C.

4.6 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed in the Institute of Chemistry, Academia Sinica using a Microcal VP-ITC (MicroCal, Northampton, MA) with a reaction cell volume of 2 mL for the oligonucleotide duplex (U4A-L4T) at 25 °C or other temperatures, as indicated in Table 3, and all samples were degassed under vacuum for 5 min. The DNA solutions of appropriate concentration in 5 mM sodium cacodylate, pH 6.5 (Table 3) were placed in the calorimeter cell. The peptide, at an appropriate concentration in 5 mM sodium cacodylate, pH 6.5 (Table 3, 273 μL), was placed in the titration syringe and injected in aliquots of 7 μL with 220 s intervals between the individual injections whilst stirring at 304 rpm, for a total of 39 injections. Data acquisition and analysis were performed using nonlinear least-squares fitting algorithm software (single site binding model, Microcal Origin 7.1 software). The K_a values were obtained from the ITC instrument by computer fit of the ITC isotherms and the corresponding ΔG values were calculated from the equation: $\Delta G = -RT\ln K_a$

Acknowledgements

This work is dedicated to the memory of Professor Dudley H. Williams. We thank professors L. S. Kan, S. T. Chen, S. H. Wu, D. K. Chang and Y. T. Tau, Academia Sinica, for generous access to ITC and CD facilities as well as helpful advice, and professors Lin Ma, A. N. Ko, A. Yeh, C. W. Ong, and F. B. Yeh for helpful discussions and encouragement. The work was supported by grant NSC97-2113-M029-005 and funds

from Tunghai Christian University, the Jeng Ching Ho Foundation, and the Inpac International Corporation.

References

1 K. R. Fox and M. J. Waring, *Methods Enzymol.*, 2001, **340**, 412–430.
2 J. B. Chaires, K. R. Fox, J. E. Herrera, M. Britt and M. J. Waring, *Biochemistry*, 1987, **26**, 8227–8236.
3 D. L. Boger, J.-H. Chen and K. W. Saionz, *J. Am. Chem. Soc.*, 1996, **118**, 1629–1644.
4 C. Bailly and J. B. Chaires, *Bioconjugate Chem.*, 1998, **9**, 513–538 and references cited therein.
5 Y. H. Chen and J. W. Lown, *J. Am. Chem. Soc.*, 1994, **116**, 6995–7005.
6 W. L. Walker, E. M. Landaw, R. E. Dickerson and D. S. Goodsell, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 5634–5639.
7 A. Blasko and T. C. Bruice, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 10018–10022.
8 R. P. L. DeClairac, B. N. Geierstanger, M. Mrksich, P. B. Dervan and D. E. Wemmer, *J. Am. Chem. Soc.*, 1997, **119**, 7909–7916.
9 E. J. Fechter and P. B. Dervan, *J. Am. Chem. Soc.*, 2003, **125**, 8476–8485 and references cited therein.
10 M. Mrksich, M. E. Parks and P. B. Dervan, *J. Am. Chem. Soc.*, 1994, **116**, 7983–7988.
11 K. L. Buchmueller, A. M. Staples, C. M. Howard, S. M. Horick, P. B. Uthe, N. Le Minh, K. K. Cox, B. Nguyen, K. A. O. Pacheco, W. D. Wilson and M. Lee, *J. Am. Chem. Soc.*, 2005, **127**, 742–750.
12 T. Brown, H. Mackay, M. Turlinton, A. Sutterfield, T. Smith, A. Sielaff, L. Westrate, C. Bruce, J. Kluza, C. O'Hare, B. Nguyen, W. D. Wilson, J. A. Hartley and M. Lee, *Bioorg. Med. Chem.*, 2008, **16**, 5266–5276.
13 M. Munde, M. A. Ismail, R. Arafa, P. Peixoto, C. J. Collar, l. Yang, L. Hu, M. H. David-Cordonnier, A. Lansiaux, C. Bailly, D. W. Boykin and W. D. Wilson, *J. Am. Chem. Soc.*, 2007, **129**, 13732–13743.
14 J.-C. Chang, C.-H. Yang, P.-J. Chou, W.-H. Yang, I.-C. Chou, C.-T. Lu, P.-H. Lin, R. C.-W. Hou, K.-C. G. Jeng, C.-C. Cheng and L. Sheh, *Bioorg. Med. Chem.*, 2004, **12**, 53–61.
15 C.-H. Yang, W. F. Chen, M.-C. Jong, B.-J. Jong, J.-C. Chang, M. J. Waring, L. Ma and L. Sheh, *J. Am. Chem. Soc.*, 2004, **126**, 8104–8105.
16 C. H. Yang, K. C. G. Jeng, W. H. Yang, Y. L. Chen, C. C. Hung, J. W. Lin, S. T. Chen, S. Richardson, C. R. H. Martin, M. J. Waring and L. Sheh, *ChemBioChem*, 2006, **7**, 1187–1196.
17 K. L. Kao, J. C. T. Huang, C. K. Yang, K. C. G. Jeng, J. C. Chang, W. C. Yao, S. C. Hsien, M. J. Waring, M. H. Chen, L. Ma and L. Sheh, *Bioorg. Med. Chem.*, 2010, **18**, 366–376; J. T. B. Huang, R. C. K. Yang, W. K. Hung, M. J. Waring and L. Sheh, in *Molecular Recognition*, ed. J. A. McEvoy, Nova Science Publishers Inc., New York, 2011. ISBN 978-1-61122-734-5, internet open access available.
18 J. T. B. Huang, Y. C. Chen, J. C. Chang, K. C. G. Jeng, K. L. Kao, C. K. Yang, L. S. Kan, M. T. Wey, M. J. Waring, C. S. Chen, W. J. Chien and L. Sheh, *Bioorg. Med. Chem.*, 2010, **18**, 2575–2585.
19 M. E. A. Churchill and M. Suzuki, *EMBO J.*, 1989, **8**, 4189–4195.
20 M. Suzuki, *Nature*, 1990, **344**, 562–565.
21 D. E. Koshland Jr. and K. Hamadani, *J. Biol. Chem.*, 2002, **277**, 46481–46844 and references cited therein.
22 W. E. Calvano, *et al.*, *Nature*, 2005, **437**, 1032–1037.
23 R. B. Russell and P. Aloy, *Nat. Chem. Biol.*, 2008, **4**, 666–673.
24 M. Ciubotaru, L. M. Ptaszek, G. A. Baker, S. N. Baker, F. V. Bright and D. G. Schatz, *J. Biol. Chem.*, 2003, **278**, 5584–5596.
25 A. A. Ouameur and H.-A. Tajmir-Riahi, *J. Biol. Chem.*, 2004, **279**, 42041–42054.
26 I. Tato, S. Zunzunegui, F. de la Cruz and E. Cabezon, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 8156–8161.
27 N. M. Luscombe, R. A. Laskowski and J. M. Thornton, *Nucleic Acids Res.*, 2001, **29**, 2860–2874.
28 K. J. Breslauer, D. P. Rometa, W. Y. Chou, R. Ferrante, J. Curry, D. Zaunczkowski, J. G. Snyder and L. A. Marky, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 8922–8926.
29 E. B. Starikov and B. Norden, *J. Phys. Chem. B*, 2007, **111**, 14431–14435 and references quoted therein.
30 D. H. Williams, E. Stephens, D. P. O'Brien and M. Zhou, *Angew. Chem., Int. Ed.*, 2004, **43**, 6596–6616.
31 A. Cooper and D. T. F. Dryden, *Eur. Biophys. J.*, 1984, **11**, 103–109.
32 H. Shiozawa, B. C. S. Chia, N. L. Davies, R. Zerella and D. H. Williams, *J. Am. Chem. Soc.*, 2002, **124**, 3914–3919.

Bioorganic & Medicinal Chemistry 18 (2010) 2575–2585

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Novel DNA–peptide interaction networks

Jonathan T. B. Huang [a], Yen-Chung Chen [a], Jung-Cheng Chang [a], Kee-Ching G. Jeng [b], Karen K. L. Kao [a], Robin C. K. Yang [a], Lou-Sing Kan [c,d], Ming-Tsair Wey [c], Michael J. Waring [e], Chee-Shan Chen [f], Wei-Jyun Chien [f], Leung Sheh [a,*]

[a] Department of Chemistry and Life Science Research Center, Tunghai Christian University, Taichung 407, Taiwan, ROC
[b] Department of Medical Research, Taichung Veterans General Hospital, Taichung 405, Taiwan, ROC
[c] Institute of Chemistry, Academia Sinica, Taipei 115, Taiwan, ROC
[d] Department of Bioengineering, Tatung University, Taipei 104, Taiwan, ROC
[e] Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, United Kingdom
[f] Department of Applied Chemistry, Chaoyang University of Technology, Taichung 413, Taiwan, ROC

ARTICLE INFO

Article history:
Received 6 January 2010
Revised 19 February 2010
Accepted 19 February 2010
Available online 2 March 2010

Keywords:
DNA–peptide molecular recognition
Positive cooperative binding
Footprinting
Circular dichroism
Interaction networks

ABSTRACT

Allostery in the binding of peptides to DNA has been studied by quantitative DNase I footprinting using four newly designed peptides containing the XP(Hyp)RK motif and *N*-methylpyrrole (Py) moieties. Apparent binding constants in the micromolar range as well as Hill coefficients were determined for each peptide. The results, together with previous studies on five other peptides support the proposal that interaction network cooperativity is highly preferred in DNA–peptide interactions that involve multiple recognition sites. It is envisaged that interstrand bidentate interactions participate in the relay of conformational changes between recognition sites on the complementary strands. Models for interpreting DNA allostery based upon interaction networks are outlined. Circular dichroism experiments involving the titration of peptides against a short oligonucleotide duplex indicate that some of these peptides bind in a dimeric manner to DNA via the minor groove, inducing characteristic conformational changes. These insights should prompt the design of new DNA-binding peptides for investigating allosteric interactions between peptides and DNA, as well as novel interaction networks, and ultimately may shed light upon the fundamental chemical rules that govern allostery in more complex biological process such as DNA–protein interaction networks.

1. Introduction

An understanding of sequence-specific or sequence-selective binding of small molecules to DNA is important for unraveling molecular aspects of the modulation of gene expression. Its fruits lie in the progress of much structural biochemistry research as well as drug design.[1–3] To this end extensive studies have been carried out on synthetic conjugates containing 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) which are related to the antiviral antibiotics netropsin and distamycin.[4–10]

Early in the study of DNA sequence-specific agents we introduced a novel XPRK[11] tetrapeptide motif which is a modification of the naturally-occurring SPXX motif[12,13] found in repeating sequences in histones, steroid hormone receptors, various segmentation gene products and some oncogene products. It was suggested that the SPXX motif assumes a β-turn stabilized by two hydrogen bonds, and that the side chains of the two basic residues engage in salt bridges with the DNA phosphate groups. Further conjugation of two XPRK motifs with Py residues yielded peptides having high affinity for DNA sequences.[14] We showed that peptides designed in this manner bind preferentially to DNA at sites containing consecutive A's or T's and often engage in interstrand bidentate interactions[14,15] with the bases of DNA (two or more hydrogen bonds formed between a hydrogen donor of an amino acid residue and a base pair in the minor groove). Because peptides incorporating XPRK motifs have binding constants in the micromolar or submicromolar range they fulfill an essential requirement for affording satisfactory DNA footprints.[11,14,16,17,20]

Allostery is recognized as a central process in biological control governing biochemical efficiency and energy expenditure, having been investigated for more than half a century. However, much less is known about allosteric features of DNA–ligand interactions than protein–ligand interactions.[21] In studies on the sequence-selective behavior of many synthetic peptides containing both the XP(Hyp)RK motif and the polyamide motif, we identified a number that show significant positive cooperativity in binding to adjacent recognition sites on the DNA duplex. Among these ligands a newly-synthesized dodecapeptide His-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-Hyp-His-NH$_2$ (**RHyp-12**) has been investigated in detail.

* Corresponding author. Tel.: +886 4 23590248; fax: +886 4 23590426.
E-mail address: Lsheh@thu.edu.tw (L. Sheh).

0968-0896/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2010.02.047

Peptide RHyp-12 is a derivative of a parent compound His-Pro-Arg-Lys-$(Py)_4$-His-Pro-Arg-Lys-NH_2 (**PyPro-12**; formerly named **PyH-12**) in which the proline of PyPro-12 is replaced by a hydroxyproline residue, furnishing an extra hydroxyl group for additional hydrogen bonding. Another peptide His-Pro-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (**PyHK-10**) contains only one X-Pro-Arg-Lys motif, and was designed as a C-terminal truncated derivative of the parent peptide PyPro-12. Similarly, truncation of the C-terminal dipeptide fragment Hyp-His affords the decapeptide His-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (**HyH-10**). Further truncating the penultimate lysyl residue of peptide PyHK-10 affords the nonapeptide His-Pro-Arg-Lys-$(Py)_4$-Arg-NH_2 (**PyHR-9**).

For the past decade, it has been recognized that network-based interactions between bio-molecules exist in many biochemical and physiological processes in nature.[18,19] High-throughput interaction studies have also established protein–protein interaction networks, metabolic networks, and transcription regulatory networks.[18] In 2006 we proposed a network-based DNA–peptide allosteric interaction model interconnecting multiple DNA sites in fragments of the latent membrane protein (LMP-1) gene from a pathogenic Epstein-Barr virus variant derived from nasopharyngeal carcinoma.[17] Recently we proposed three different types of network-based allosteric communication between synthetic peptides binding to DNA: circuit type, incomplete-circuit type, and non-circuit type characterized by interstrand bidentate interactions.[20] In the present study we set out to explore the structural basis for molecular recognition and cooperativity in allosteric DNA interaction networks using the four designed peptides described above, employing quantitative DNase I footprinting as the major methodology. In addition, to investigate molecular aspects of the processes we resorted to circular dichroism experiments which provide an invaluable spectral probe of DNA–peptide interactions.

2. Results and discussion

Quantitative footprinting, using a 5′-^{32}P-labeled 158-mer DNA duplex and a complementary 5′-^{32}P-labeled 135-mer DNA duplex,[14] was employed as the major methodology for this study since most other methods fail to discriminate multiple binding sites on the DNA double helix (binding sites are designated U and L referring to the upper and lower strands, respectively). Monodentate interactions and interstrand bidentate interactions between peptide moieties and the DNA bases have been recognized as an essential basis of DNA–protein[15] and DNA-small ligand[14,20] sequence-specific or sequence-selective recognition. The positions of interstrand bidentate interactions can be assigned by connecting bases (with pecked lines)

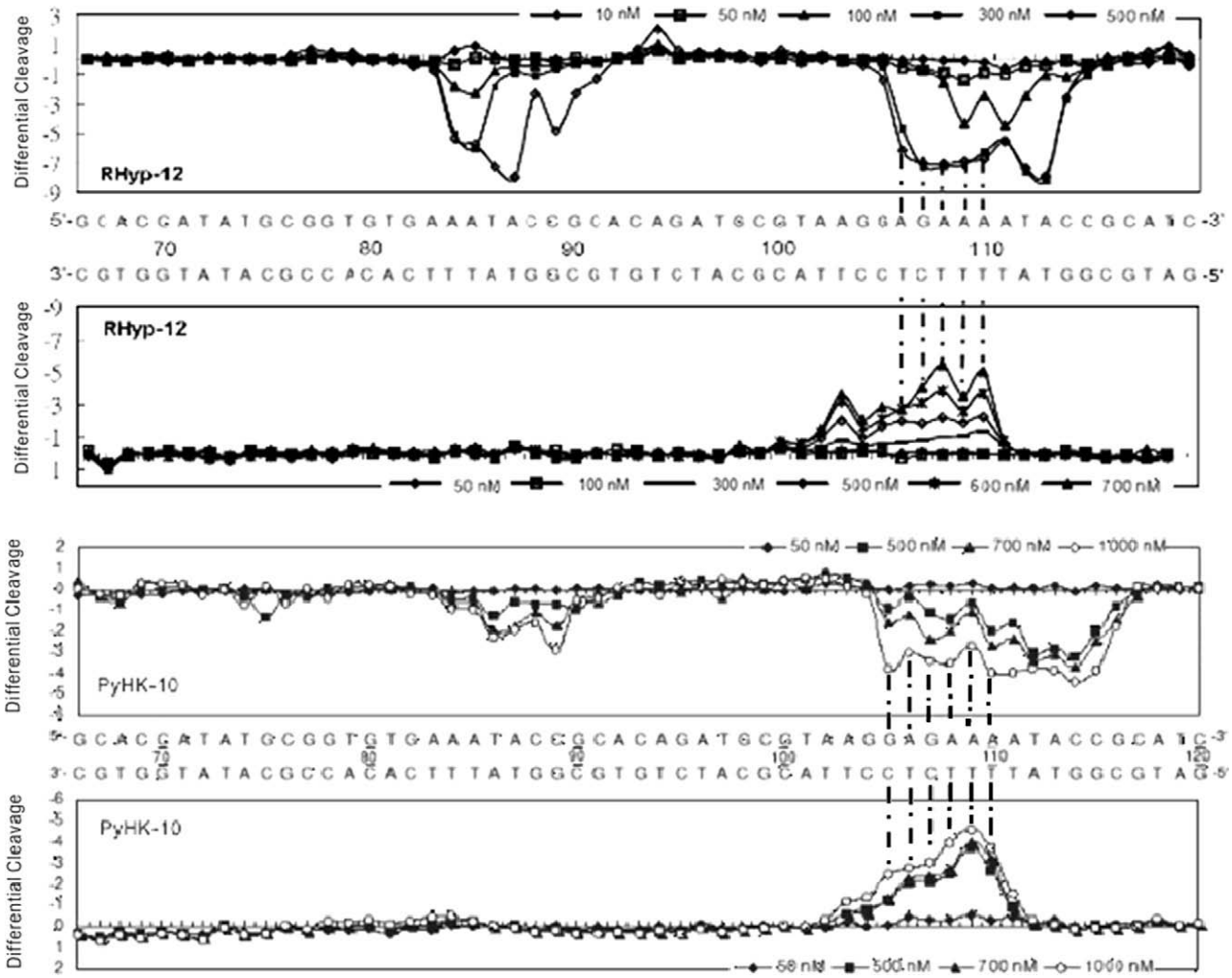


Figure 1. Differential cleavage plots comparing the susceptibility of DNA fragments to DNase I cleavage after incubation with each peptide in cacodylate buffer at room temperature for 60 min. The upper traces represent the differential cleavage plot for a given peptide bound to the 5′-[^{32}P]-labeled upper strand (158-mer) DNA fragment; the lower traces represent the corresponding plots for each peptide bound to the 5′-[^{32}P]-labeled lower strand (135-mer) DNA fragment. The vertical dotted lines between DNA bases represent assignment of interstrand bidentate interactions where significant coincident H-bonding interactions occur between complementary bases in both strands.

where significant simultaneous DNase I blockages on the complementary strands appear in differential cleavage plots (Fig. 1).

On the upper strand of the 158-mer fragment, binding of peptide RHyp-12 produces two major DNase I blockage sites: position U106–113, corresponding to the sequence 5′-AGAAAATA-3′ (Figs. 1 and 2A) and displaying significant positive cooperativity (n_H = 2.1); and position U84–90, comprising the sequence 5′-AATACCG-3′ also showing positive cooperativity (n_H = 1.5). On the lower strand, peptide RHyp-12 elicits a wide DNase I blockage site at position L110–103 corresponding to the sequence 5′-TTTCTCCT-3′, again showing positive cooperativity (n_H = 2.0). From the differential cleavage plots, five interstrand bidentate interactions can be assigned at position 106–110, comprising the self-complementary sequence 5′-AGAAA-TTTCT-3′ (Fig. 1, Table 1).

On the 158-mer upper strand, the decapeptide PyHK-10 produces strong and broad DNase I blockage extending from position U105 to 115, corresponding to the sequence 5′-GAGAAAATACC-3′ and displaying significant positive cooperativity (n_H = 2.4) (Figs. 1 and 2B). At a peptide concentration of 1000 nM, a region of weak and broad DNase I blockage is also observed around position U86–89, corresponding to the sequence 5′-TACC-3′ and displaying weak negative cooperativity. On the 135-mer lower strand, peptide PyHK-10 shows a single DNase I blockage site around position L111–105, corresponding to the sequence 5′-TTTTCTC-3′ and displaying only weak positive cooperativity (n_H = 1.4). Three interstrand bidentate interactions are assigned around position 105–111. Unlike peptide RHyp-12 and the parent peptide PyPro-12,[20] peptide PyHK-10 does not seem to exhibit any preference for

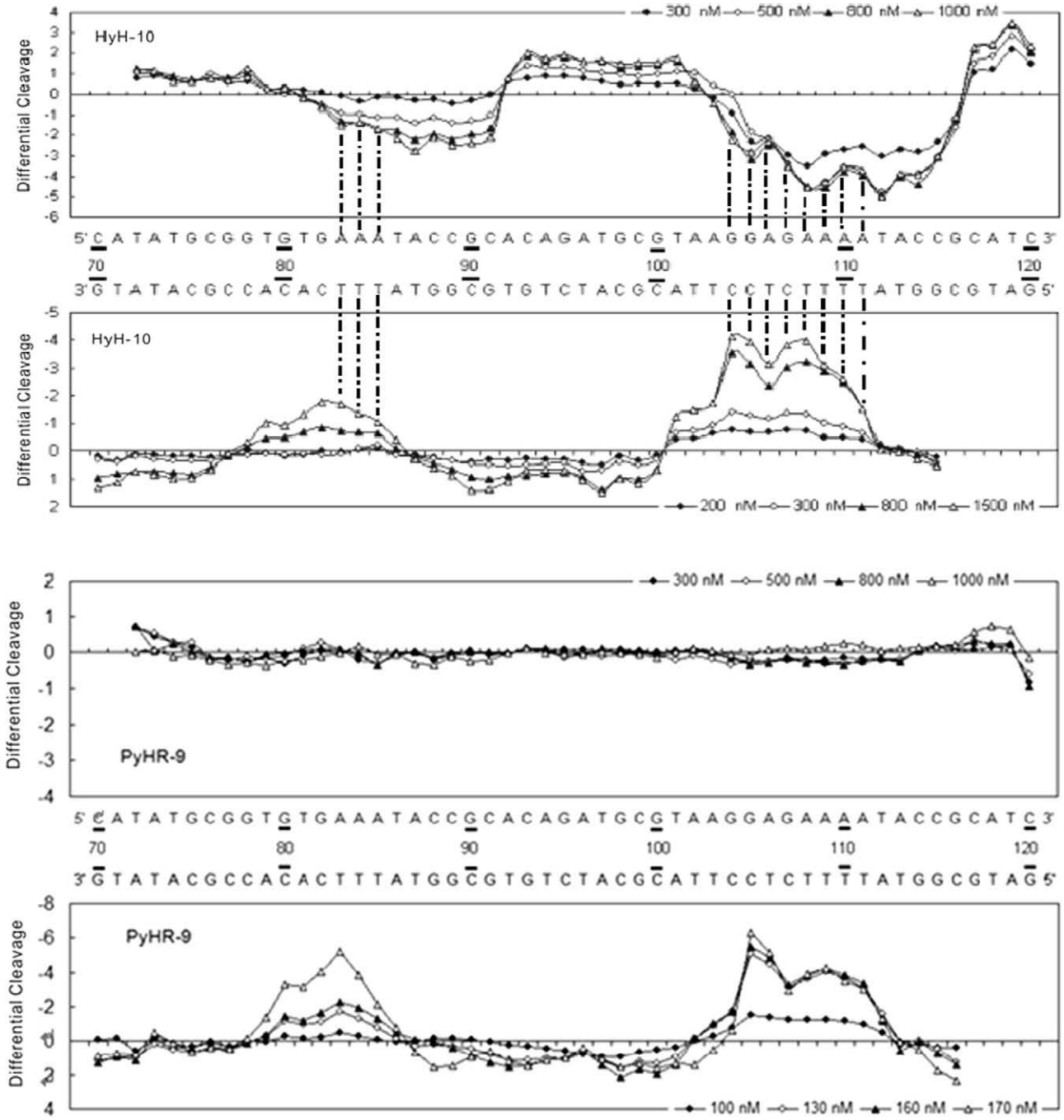


Fig. 1 (*continued*)

binding to two of the three consecutive A residues on the upper strand around position U83–85 but it does in binding to position U86–89 comprising the sequence 5′-TACC-3′ (Figs. 1 and 2B). This discrepancy in DNA sequence preference at the same binding site manifested by PyHK-10 compared to peptides RHyp-12 and Py-Pro-12 can be attributed to the truncated structure of PyHK-10.

The sequence preference of decapeptide HyH-10 (Figs. 1 and 2C) is quite different from that of peptides RHyp-12 and PyHK-10. On the upper strand peptide HyH-10 induces strong and very broad DNase I blockage around position U104–115, corresponding to the sequence 5′-GGAGAAAATACC-3′, displaying positive cooperativity (n_H = 1.7), and there is also a broad DNase I blockage site around position U83–91, comprising the sequence 5′-AAA-TACCGC-3′, where notable positive cooperativity (n_H = 2.6) is evident. On the lower strand, two regions of DNase I blockage appear around positions L85–82, and L111–104. Interstrand bidentate interactions are assigned around positions 83–85 and 104–111 (Fig. 1, Table 1).

Furthermore, the nonapeptide PyHR-9 displays very different DNA sequence preferences from all other peptides studied to date (Fig. i, Supplementary data). At concentrations up to 3000 nM (above which value non-specific peptide binding is evident), no DNase I blockage can be observed on the 158-mer upper strand. Peptide PyHR-9 appears to have a rather strong binding site on the lower strand, where saturation is observed around a peptide concentration of 190 nM (Fig. i, Supplementary data). Two DNase I blockages can be seen around position L85–79, corresponding to the sequence 5′-TTTCACA-3′, displaying high positive cooperativity (n_H = 5.0), and around position L112–104, comprising the sequence 5′-ATTTTCTCC-3′, where unusually high positive cooperativity (n_H = 10.6) is measured. Comparing peptide PyHR-9 with peptide PyHK-10, the only primary structural difference is that the latter peptide bears an additional lysyl residue at the penultimate C-terminal position. As mentioned above, peptide PyHR-9 refuses to bind to the upper strand, and does not engage in interstrand bidentate interactions (Fig. 1). By contrast, peptide

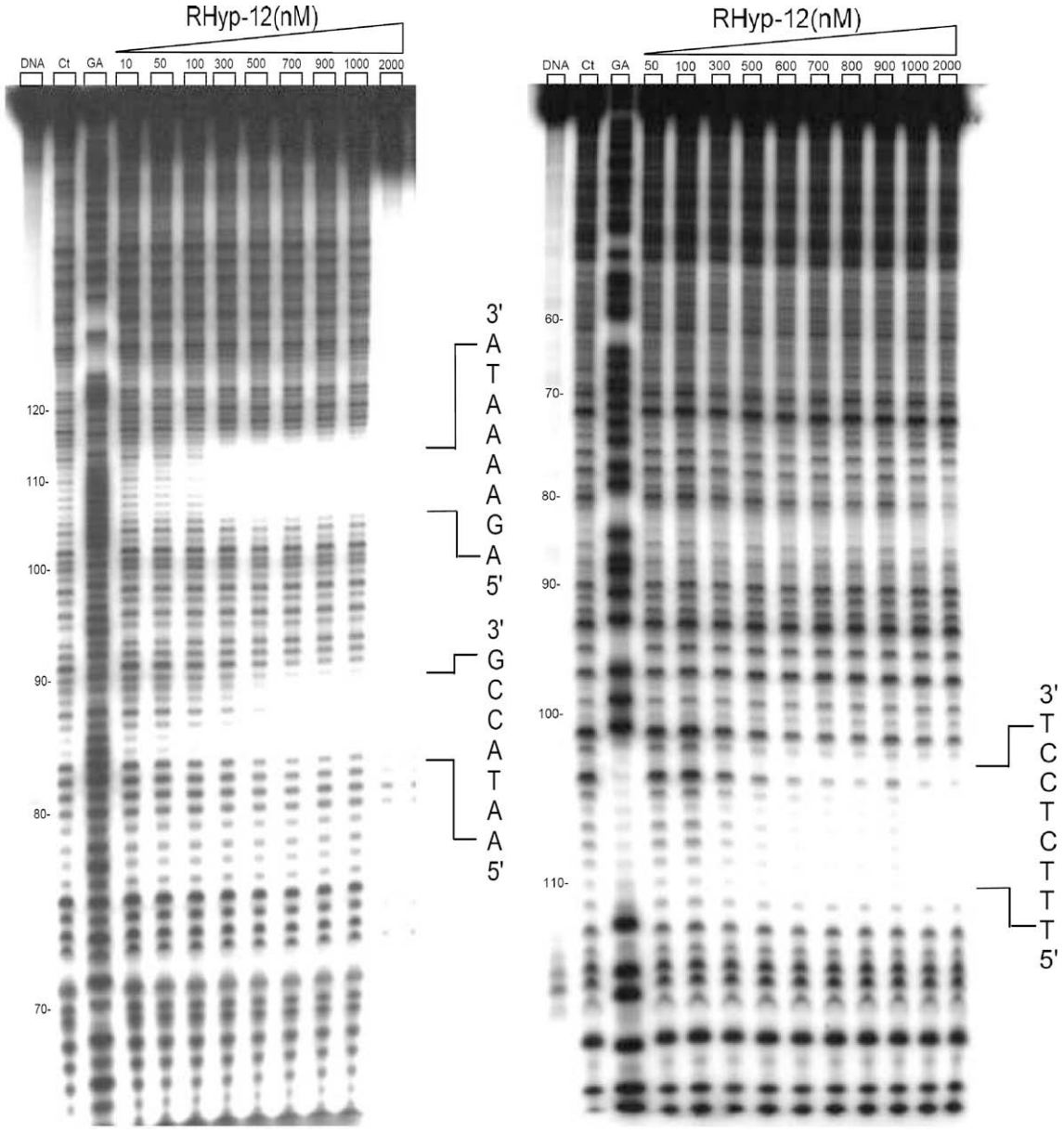


Figure 2A. Autoradiograph showing DNase I footprinting of peptide RHyp-12 on DNA duplexes labeled at the 5′ end: 5′-[^{32}P]-labeled 158-mer upper strand, left panel; and 5′-[^{32}P]-labeled 135-mer lower strand, right panel. Peptide RHyp-12 was equilibrated with the DNA in 5 mM sodium cacodylate buffer, pH 6.5 at 37 °C for 60 min before DNase 1 cleavage. G represents a Maxam–Gilbert guanine sequencing track and Ct shows a DNase I digestion control lane.

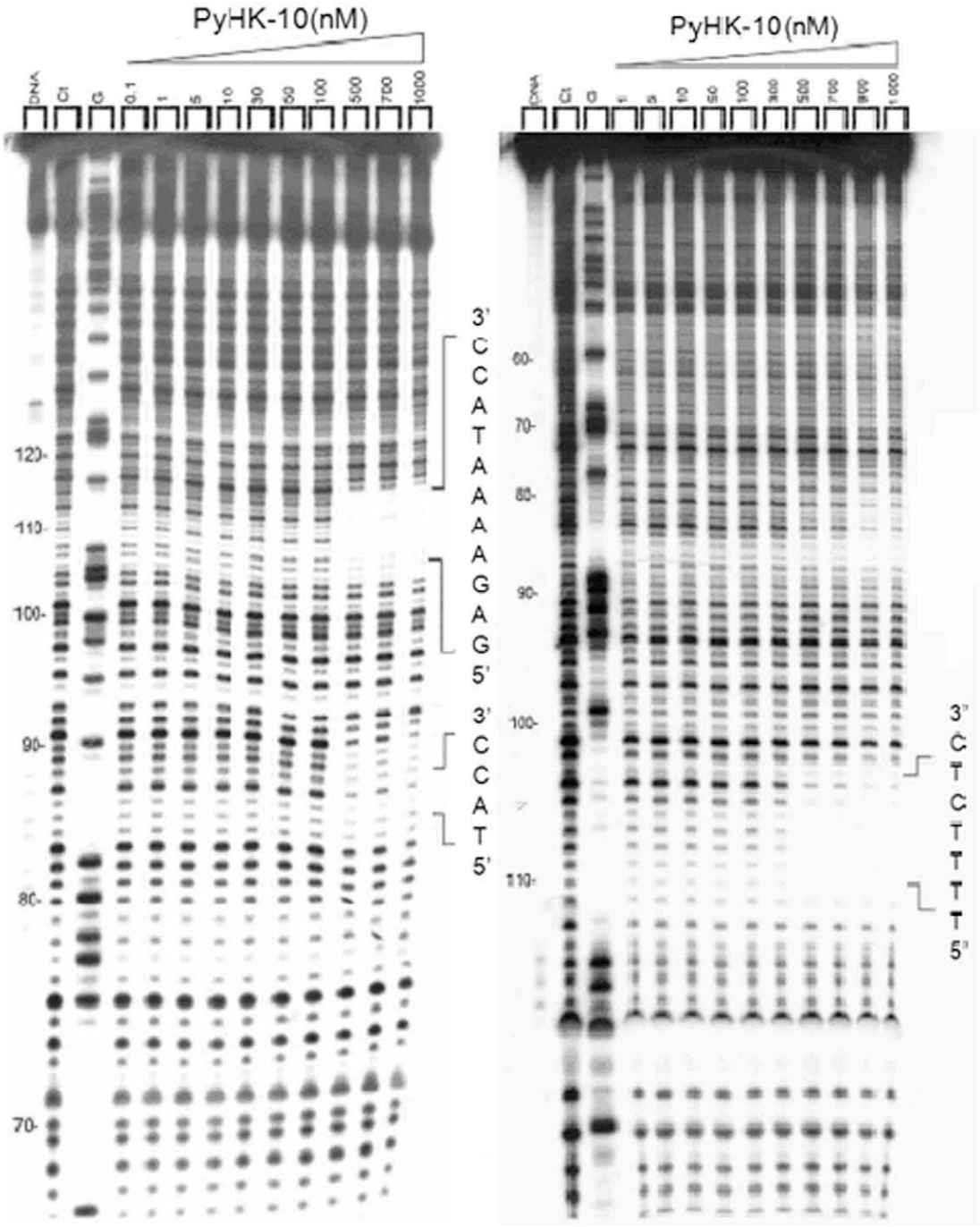


Figure 2B. Autoradiograph showing DNase I footprinting of peptide PyHK-10 on DNA duplexes labeled with [^{32}P] on the 5′ end: 5′-[^{32}P]-labeled 158-mer upper strand, left panel; and 5′-[^{32}P]-labeled 135-mer lower strand, right panel.

PyHK-10 binds strongly to its two major sites on the upper strand and shows interstrand bidentate interactions (Fig. 1). This suggests that at least two Lys or Arg residues near the C-terminal position of the peptide are required for binding to the upper strand binding sites: d(AAAA)- and d(AAA)-. The reason why peptide PyHR-9 appears to engage in unusually strong positive cooperative binding to two sites on the lower DNA strand is unclear and awaits further investigation. However, in a previous investigation,[17] peptides HR-12 and SP-12 that do not manifest interstrand bidentate interactions with LMP-1 gene fragments were also found to show notably strong positive cooperative binding. Thus, we speculate that unusually high positive cooperativity in binding to DNA by small molecules might be related to the absence of interstrand bidentate interactions that probably stabilize local DNA conformation and thus are likely to inhibit strong positive cooperative ligand binding between adjacent sites. The apparent binding constant K_a of the four peptides for the four major binding sites on the pBR322 DNA fragments lie within the range 1.4×10^6–$9.7 \times 10^6\ M^{-1}$. Peptide PyHR-9 gives rise to the two highest apparent binding constants on the 5′-TTTCACA-3′ and 5′-ATTTTCTCC-3′ sites of the lower strand (Table 1).

Based on the footprinting results, we now propose models to interpret the complex communication between DNA–peptide allosteric binding sites for the four designed peptides (Fig. 3). A notable

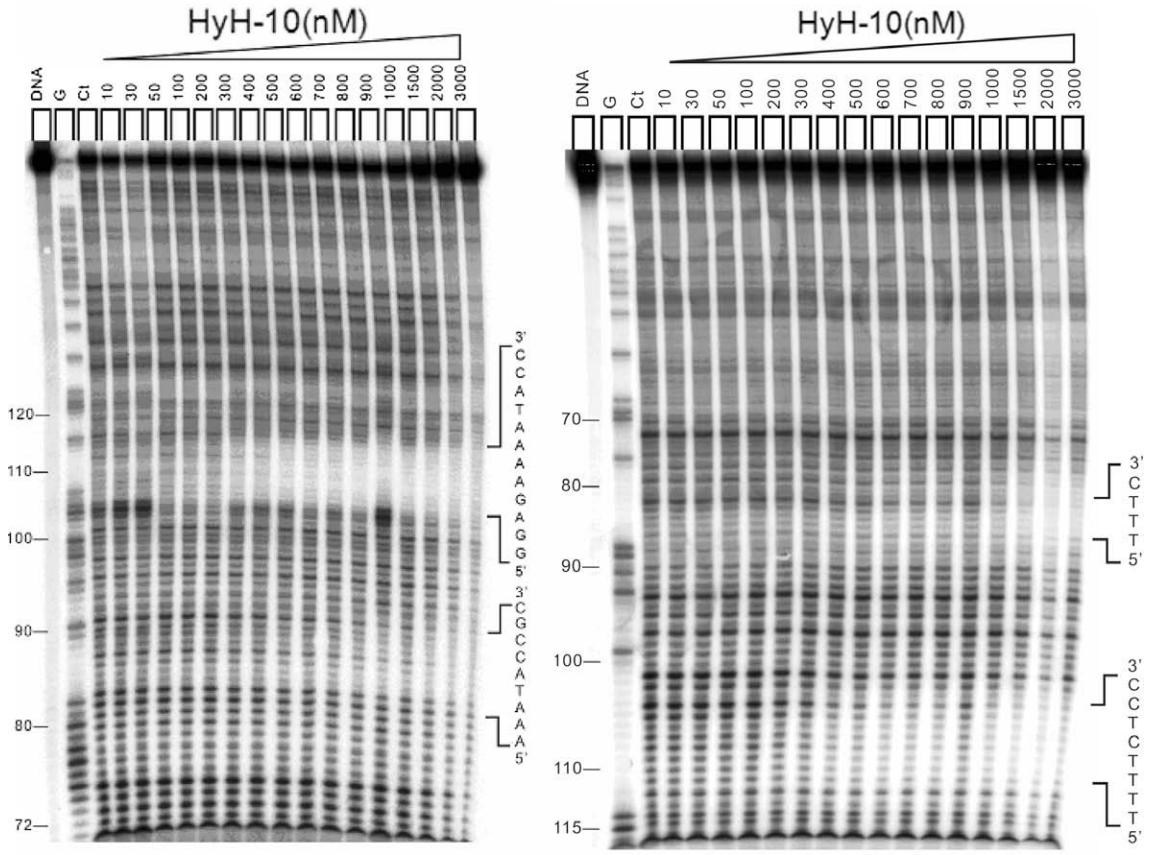


Figure 2C. Autoradiograph showing DNase I footprinting of peptide HyH-10 on DNA duplexes labeled at the 5′ end: 5′-[^{32}P]-labeled 158-mer upper strand, left panel; and 5′-[^{32}P]-labeled 135-mer lower strand, right panel.

Table 1
Binding specificity and physicochemical parameters of sequence-specific binding of peptides to various recognition sites on complementary 5′-[^{32}P]-labeled upper (158-mer) and 5′-[^{32}P]-labeled lower (135-mer) DNA strands at 37 °C determined by quantitative DNase I footprinting

Ligand	Binding site position	Recognition sequence	K_a	n_H	Position of interstrand bidentate interactions
RHyp-12	U84–90	5′-AATACCG-3′	3.0×10^6	1.5	106–110
	U106–113	5′-AGAAAATA-3′	7.6×10^6	2.1	
	L110–103	5′-TTTCTCCT-3′	7.4×10^6	2.0	
PyHK-10	U86–89	5′-TACC-3′	4.9×10^6	0.8	105–110
	U105–115	5′-GAGAAAATACC-3′	2.7×10^6	2.4	
	L111–105	5′-TTTTCTC-3′	3.4×10^6	1.4	
HyH-10	U83–91	5′-AAATACCGC-3′	1.8×10^6	2.6	83–85
	U104–115	5′-GGAGAAAATACC-3′	4.7×10^6	1.7	104–111
	L85–82	5′-TTTC-3′	1.4×10^6	1.3	
	L111–104	5′-TTTTCTCC-3′	3.6×10^6	2.0	
PyHR-9	L85–79	5′-TTTCACA-3′	8.4×10^6	5.0	NIL
	L112–104	5′-ATTTTCTCC-3′	9.7×10^6	10.6	

K_a and n_H are the apparent association constant and Hill coefficient determined from concentration-dependent DNase I footprinting studies, respectively. The binding site positions on the upper and lower strands are abbreviated as U and L, respectively. Interstrand bidentate interactions are assigned where there are coincident effects on

feature is that these models reveal networks of interaction. Put another way, our models point to network-based allosteric connections between recognition sites. Conformational changes occurring at particular binding loci are transmitted to adjacent sub-binding sites and also to sites on the complementary strand, most likely via interstrand bidentate interactions. Allosteric communication may also occur between more remote binding sites, spanning distances up to 12–16 base pairs. Recent footprinting experiments[20] on the 81-mer duplex (S-81) containing a single d(AAAA)–d(TTTT) sequence have shown that the n_H values for peptides PyHyp-12 and PyHyp-9 binding to 81-mer duplexes containing a single binding locus are significantly higher than those for the 158-mer and 135-mer duplexes containing both binding loci. These results suggest that the allosteric relay of peptide binding information to neighboring sites between the d(AAA)–d(TTT) and d(AAAA)–d(TTTT) positions spanning an intervening sequence of 12–14 base pairs may be affected by some negative cooperative effect(s).

It is also notable that the binding of peptides HyH-10 and Py-Pro-12[20] to pBR322 fragments possessing multiple sites reveals what we term a circuit type of allosteric communication within a

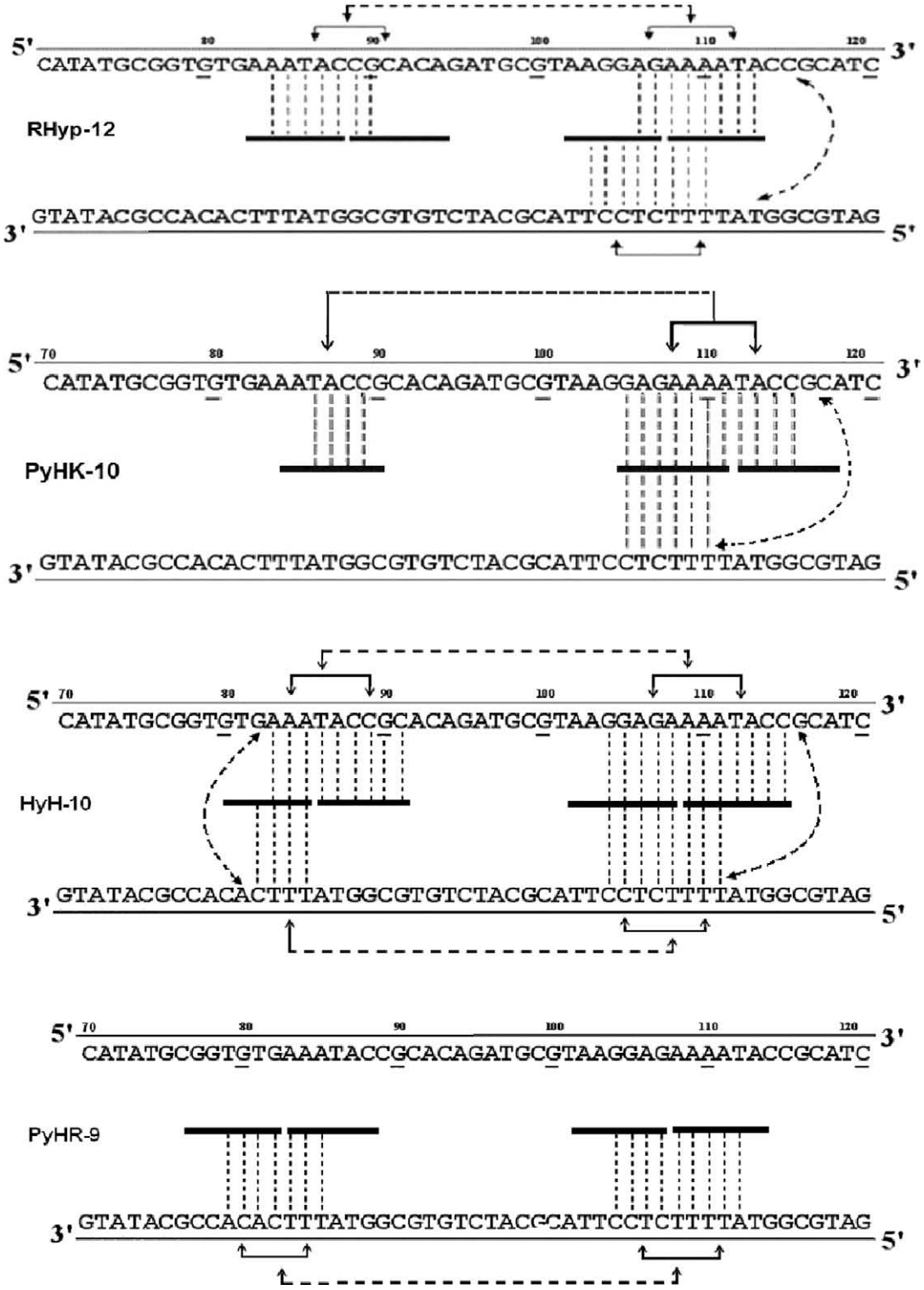


Figure 3. Proposed allosteric models for cooperative binding of peptides RHyp-12, PyHK-10, HyH-10, and PyHR-9 to 158-mer and 135-mer pBR322 fragments based on quantitative footprinting studies. The portion of the ligand binding to each DNA site/sub-site is represented by a thick horizontal line. Monodentate interactions and interstrand bidentate interactions are represented by vertical broken lines. The solid horizontal arrow lines represent communication of allosteric interaction between DNA sub-binding sites. Broken horizontal arrow lines between neighboring binding loci some 12–16 nucleotides apart are intended to represent moderate or weak cooperative communication.

network-based system; that is, the allosteric interactions between binding sites form a closed circuit. For peptides RHyp-12 and PyHK-10, the networks are referred to as incomplete-circuit type; that is, the allosteric communications form a partial or incomplete circuit. On the other hand, peptides PyHR-9, $(HPRK)_3NH_2$ [HR-12],[17] $(SPRK)_3NH_2$ [SP-12][17] whose binding to DNA fragments lacks interstrand bidentate interactions appear to engage in a non-circuit type of allosteric communication.[20]

Thus, the quantitative footprinting results in this work and those reported previously[17,20] support our hypothesis[20] that three

different types of network-based allosteric communication in peptide–DNA molecular recognition can be distinguished: circuit type, incomplete-circuit type and non-circuit type.

To gain insight into conformational changes of DNA associated with the sequence-selective binding of peptides, we carried out circular dichroism studies using the four peptides RHyp-12, PyHK-10, HyH-10 and PyHR-9. A 13-mer deoxyribonucleotide duplex d(TAGGAGAAAATAC)–d(GTATTTTCTCCTA) (U4A-L4T), containing a sequence of 12 base pairs that correspond to the binding site at position 103–114 of the pBR322 fragments was used as substrate. The CD spectra for the DNA and peptides only are shown in Figure 4A.

Titration of peptide RHyp-12 versus the U4A-L4T duplex produces a dose-dependent negative CD enhancement band around 247 nm together with two strong positive CD enhancement bands around 282 nm and 328 nm (Fig. 4B). In the difference spectra the negative band around 247 nm can be seen to be red-shifted to 269 nm and two strong positive bands are seen around 289 nm and 327 nm (Fig. 4C). A near-isoelliptic point occurs at 280 nm, suggesting a two-component binding process. It has been reported in a number of articles[22,23] that molecules which bind to the minor groove typically exhibit an induced positive CD band around 320 nm. Thus, it is safe to conclude that the strong positive CD enhancement bands around 327 nm seen with peptide RHyp-12 and other peptides simply represent binding to the minor groove.

On the other hand, titration of peptide PyHK-10 against the duplex U4A-L4T produces a negative CD band around 248 nm and a positive band around 284 nm (Fig. 4D). The positive CD band extends as a strong and broad dose-dependent positive CD enhancement band up to 332 nm. In the difference spectrum (Fig. 4E) this strong, broad, dose-dependent positive band is shifted to 329 nm whereas the induced positive CD band around 284 nm is shifted to 289 nm. These bands are mostly peptide concentration-dependent, except that the CD band induced by 0.5 μM peptide concentration is blue-shifted while the CD band induced by the highest concentration (5.0 μM) has lowered CD intensity centering around 325 nm. We repeated the CD experiments using 0.5 μM and 5.0 μM peptide concentrations and identical results were obtained. The unusual deviation in ellipticity change of peptide PyHK-10 could be due to an abrupt change in DNA conformation occurring at high peptide concentration (5.0 μM).

The CD spectra seen with peptide HyH-10 are quite similar to those of peptide PyHK-10. Titration of peptide HyH-10 versus the duplex U4A-L4T induces a negative CD band around 247 nm and a positive band around 284 nm (Fig. 4F). A strong, broad, dose-dependent positive CD enhancement appears around 330 nm. In the difference spectrum (Fig. 4G), this CD band around 330 nm is unchanged. Another induced positive CD band is red-shifted from 284 nm to 288 nm.

The CD spectra measured with the nonapeptide PyHR-9 are rather different from those of the decapeptides. Titration of peptide PyHR-9 against the duplex U4A-L4T induces a negative CD band around 249 nm and a positive band around 270 nm (Fig. 4H). A

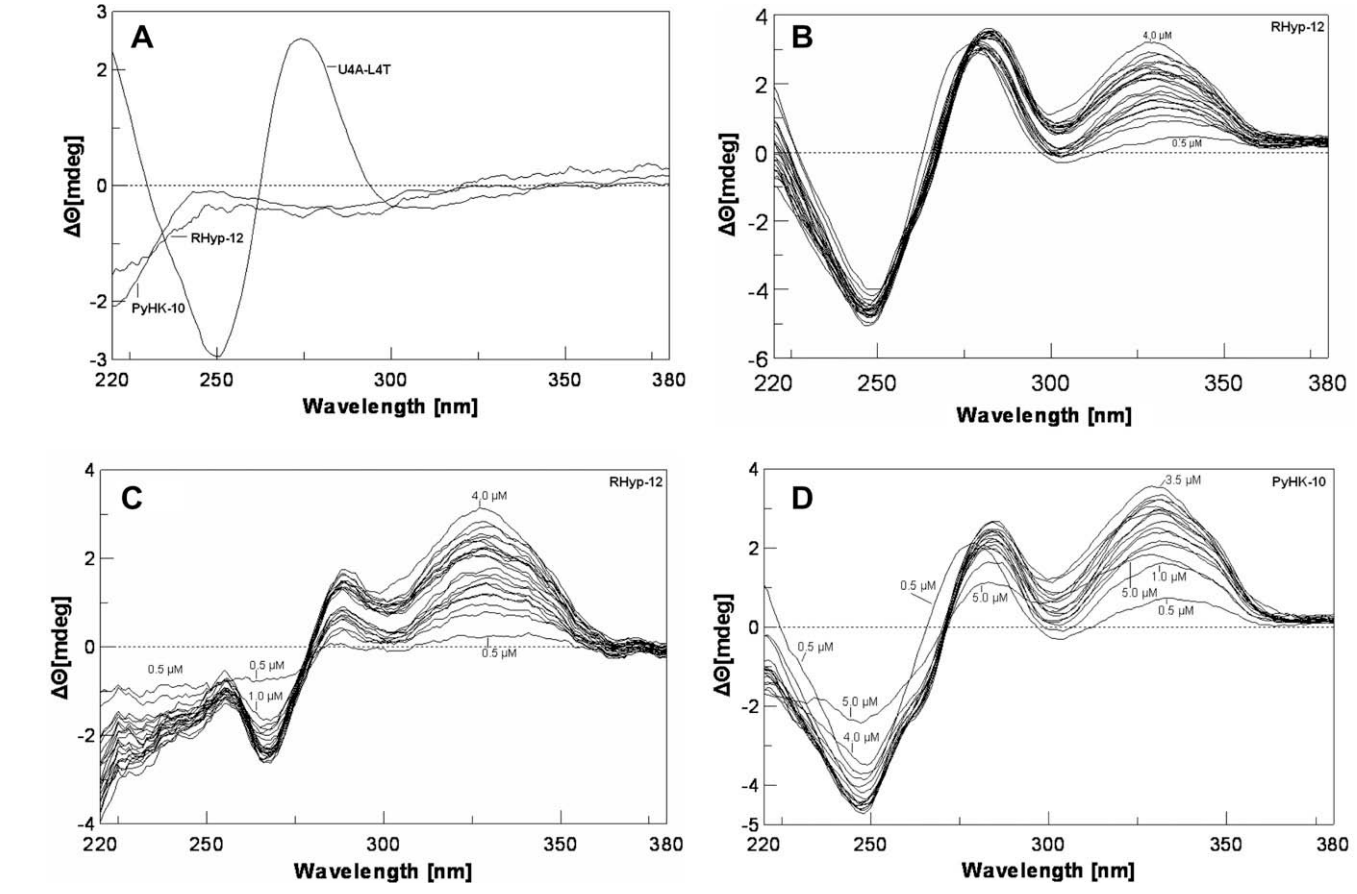


Figure 4. Panel A: CD spectra of DNA duplex U4A-L4T alone and peptide alone. Panel B: Titration of duplex U4A-L4T versus peptide RHyp-12 at peptide concentrations of 0.2, 1.0. 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, 5.0 μM. Panel C: Corresponding CD difference spectra with the contribution of duplex and peptide RHyp-12 subtracted. Panel D: Titration of duplex U4A-L4T versus peptide PyHK-10 at the same peptide concentrations. Panel E: Corresponding CD difference spectra with the contribution of duplex and peptide PyHK-10 subtracted. Panel F: Titration of duplex U4A-L4T versus peptide HyH-10 at the same peptide concentrations. Panel G: Corresponding CD difference spectra with the contribution of duplex and peptide HyH-10 subtracted. Panel H: Titration of duplex U4A-L4T versus peptide PyHR-9 at the same peptide concentrations. Panel K: Corresponding CD difference spectra with the contribution of duplex and peptide PyHR-9 subtracted.

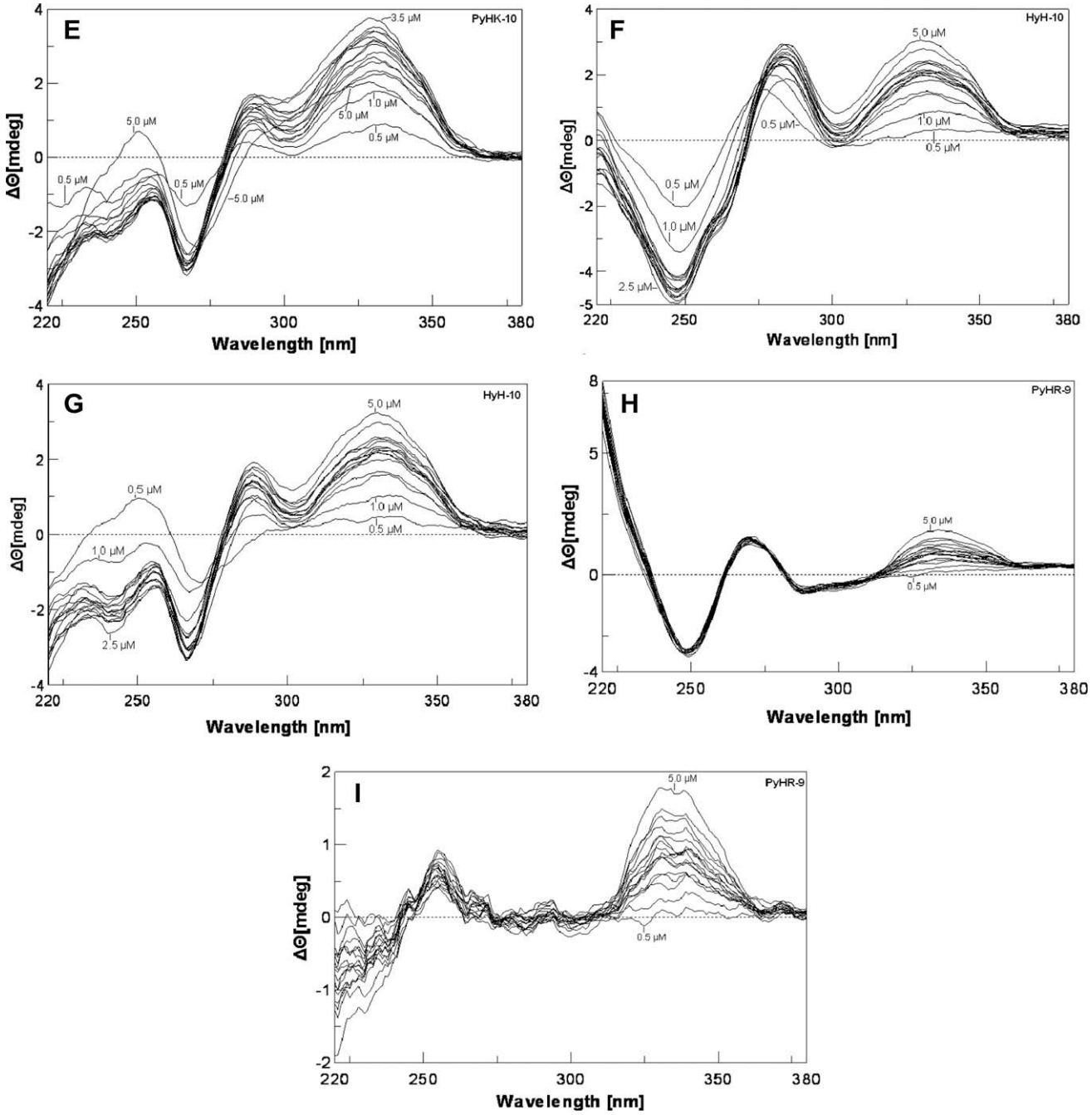


Fig. 4 (*continued*)

wide and shallow positive band is induced around 334 nm. In the difference spectra a weak band occurs around 255 nm and the positive band around 334 nm is unchanged (Fig. 4I).

It is apparent that the peptide-induced enhancement of CD bands is correlated with peptide–DNA binding stoichiometry. Figure 5 shows a plot of [peptide]/[DNA duplex] versus $\Delta\theta$ at 322 nm for peptides RHyp-12, PyHK-10, HyH-10 and PyHR-9. It is evident that the decapeptide PyHK-10 induces the greatest ellipticity change among the four peptides, suggesting that more drastic DNA conformational changes are associated with the binding of this peptide in the minor groove. On the other hand, significantly greater ellipticity changes are induced by peptides PyHK-10, RHyp-12 and HyH-10 compared to nonapeptide PyHR-9, suggesting that much stronger conformational changes are induced in the minor groove by all three former peptides. The ellipticity changes induced by peptides RHyp-12 and HyH-10 are of comparable magnitude.

From the [peptide]/[DNA duplex] plot (Fig. 5), it is apparent that at peptide/DNA ratios of 0.5–2.0, one molecule of peptide binds to the d(AAAA)–d(TTTT) locus. At [peptide]/[DNA duplex] ratios of 2.0–3.0, as indicated by the small plateau region of the titration curve, two peptide molecules seem to bind in a dimeric pattern to the binding site in the minor groove. At [peptide]/[duplex] ratios above 4.0, a progressive increase in $\Delta\theta$ suggests that more peptide molecules begin to engage in some sort of non-sequence-selective binding to DNA. The CD results agree well with other DNase I footprinting results that dimeric peptide binding to the d(AAAA)–d(TTTT) locus is favored, as shown by the very wide binding locus spanning a distance of 9–12 base pairs (Fig. 1, Table 1).

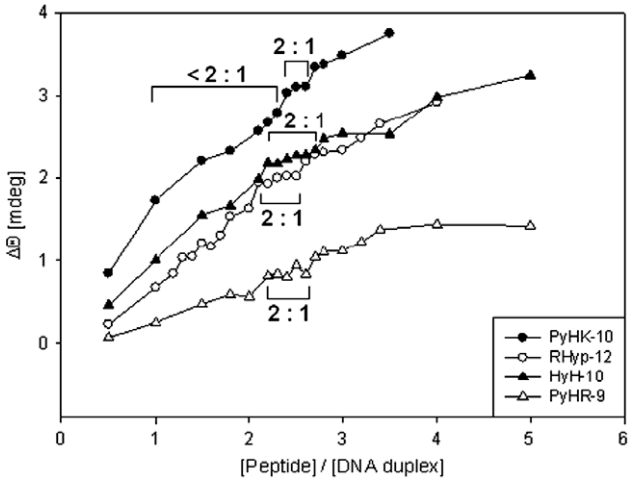


Figure 5. CD intensity at 322 nm as a function of [peptide]/[DNA duplex] The proposed stoichiometric binding ratios are as indicated, with binding below 2:1 (<2:1) considered to be predominantly 1:1.

3. Conclusion

The present study of four peptides, combined with previous studies of five other peptides[17,20] composed of XP(Hyp)RK motifs, supports the proposal that network interaction cooperativity is highly preferred in DNA–peptide interactions that involve multiple recognition sites. It is envisaged that interstrand bidentate interactions participate importantly in the relay of conformational changes between recognition sites on the complementary strands. The present and previous studies[17,20] also lend support to our recent proposal[20] that three different types of allosteric interaction networks can be distinguished in peptide–DNA molecular recognition. Thus, the present work should prompt the design of new DNA-binding peptides for further investigation of DNA–peptide allosteric interactions and DNA interaction networks. In addition, by extending the footprinting approach for studying DNA allosteric interaction networks to multiple DNA-small molecule interactions we may gain useful insights into the fundamental chemical rules that govern allostery in more complex biological process such as DNA–protein interaction networks.

4. Experimental

4.1. Chemicals and biochemicals

All the protected amino acid derivatives were purchased from Bachem California (Torrance, CA) and AnaSpec, Inc. (San Jose, CA). All peptides are synthesized in our laboratory. All other analytical reagents were purchased from Acros, Tedia or Sigma. The radiolabeled nucleoside triphosphate [γ-^{32}P]ATP was obtained from NEN Life Science Products at a specific activity of 6000 Ci/mmol. *Taq* polymerase, T4 polynucleotide kinase, and DNase I were purchased from Promega. All other chemicals were analytical grade reagents, and all solutions were prepared using deionized, Millipore-filtered water.

4.2. Solid-phase peptide synthesis

4.2.1. His-Pro-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (PyHK-10)

This peptide was synthesized using solid phase methodology by manual operation of a Protein Technology PS3 peptide synthesizer. The first Fmoc-protected amino acid was coupled to the Nova Rink amide AM resin using PyBOP/NMM in DMF. All of the N$_\alpha$-Fmoc-protected amino acids (in 4 equiv ratio excess to the resin) were coupled in a stepwise fashion using PyBOP/NMM in DMF after deprotection of the N$_\alpha$-Fmoc group by piperidine. The side chains of Arg, Lys, Tyr and His are protected by the Pmc, Boc and Trt groups, respectively. After coupling the last N-terminal Fmoc-amino acid residue, the resin was treated with the cleavage reagent (0.75 g phenol, 10 mL TFA, 0.5 mL thioanisole, 0.25 mL EDT) for 1.5 h, and then lyophilized. The resin was washed with dry ether (2 × 30 mL), filtered, and then washed with 5% acetic acid (200 mL). The combined filtrate was lyophilized and the product was purified by semi-preparative reversed-phase HPLC (column 1) using gradient elution. Eluent A: 5% MeCN, 95% H_2O, 0.1% TFA; Eluent B, 95% MeCN, 5% H_2O, 0.1% TFA. A linear gradient was achieved by increasing the MeCN content from eluent A to eluent B in 30 min t_R (column 2), 11.93 min, mp 144–146 °C, $[\alpha]_D^{24}$ −46.78 (*c* 0.043, MeOH/H_2O, 1:1); ESIMS requires: 1308.5, found: 1309.0.

4.2.2. His-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-Hyp-His-NH$_2$ (RHyp-12)

This peptide was synthesized and purified to homogeneity using a similar procedure as for peptide PyHK-10. t_R (column 2), 11.65 min mp 159–162 °C, $[\alpha]_D^{24}$ −14.28 (*c* 0.035, MeOH/H_2O, 1:1); ESIMS requires: 1574.75, found: 1574.93.

4.2.3. His-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (HyH-10)

This peptide was synthesized and purified to homogeneity using a similar procedure as for peptide PyHK-10. t_R (column 2), 14.63 min, mp 153–157 °C, $[\alpha]_D^{27}$ −14.0 (*c*, 0.262 MeOH/H_2O, 1:1); ESIMS requires: 1324.50, found: 1325.18.

4.2.4. His-Pro-Arg-Lys-(Py)$_4$-Arg-NH$_2$ (PyHR-9)

This peptide was synthesized and purified to homogeneity using a similar procedure as for peptide PyHK-10. t_R (column 2), 15.29 min, mp 150–153 °C, $[\alpha]_D^{27}$ −28.33 (*c*, 0.20 MeOH/H_2O, 1:1); ESIMS requires: 1180.53, found: 1180.38.

4.3. Polymerase chain reaction (PCR) and end-labeling of PCR products

The 5′-^{32}P-labeled 158-mer DNA duplex and 5′-^{32}P-labeled 135-mer DNA duplex were prepared by PCR amplification in a thermal cycler (ABI model 9700) as reported previously.[14] The DNA concentration was determined by UV spectroscopy to lie in the range 600–800 nM.

4.4. DNase I footprinting

Reactions were conducted in a total volume of 10 μL. Radiolabeled DNA (2 μL) was mixed with varying concentrations of peptide (2 μL) dissolved in 5 mM sodium cacodylate buffer, pH 6.5 and equilibrated at room temperature for 120 min. DNase I (2 μL) was added and the reaction allowed to proceed at 37 °C for 10 min. The DNase I solution (in 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$) was adjusted to yield a final concentration of 0.009 unit/mL so as to limit the digestion to less than 30% of the starting material in order to minimize the incidence of multiple cleavages in any one strand. The digestion was stopped by adding stop solution (4 μL) containing 80% formamide, 10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol. Samples were heated at 90 °C for 4 min and chilled on ice for 4 min prior to electrophoresis. The products of DNase I cleavage were resolved by polyacrylamide gel electrophoresis under denaturing conditions (0.3 mm thick, 8% acrylamide containing 8 M urea). After electrophoresis (about 1.45 h at 70 Watts, 1800 V in TBE buffer, BRL sequencer model S2), gels were soaked in 10% acetic acid/10% methanol for 15 min, transferred to Whatman 3MM paper, dried under vacuum at 80 °C for 45 min. The autoradiographs (footprinting of peptides

RHyp-12 and PyHK-10) were obtained by using Kodak BioMax MR scientific imaging films with intensifying screens for about 5 days. The autoradiographs (footprinting of peptides HyH-10 and PyHR-9) were obtained by placing the dried gels against an imaging plate (BAS-MS2340) of a Fuji phosphoImager FLA-5100 overnight. The electrophoretic band areas were analyzed by a PC computer installed with Viber Lourmat BIO-ID software (Marne La Valle, Cedex 1, France).

The apparent DNA binding site saturation is determined by the equation:[9]

$$Y' = 1 - (I_{tot}/I_{ref})/(I^{o}_{tot}/I^{o}_{ref})$$

where Y' is the fractional saturation, I_{ref} and I^{o}_{ref} are integrated band volumes of 5 bases of a running lane (non-binding site, with peptide) and control lane (without peptide), respectively. I_{tot} and I^{o}_{tot} are integrated band volumes of the binding site locus and corresponding control lane locus, respectively.

The value of Y is optimized by the following equation:

$$Y = (Y' - Y'_{min})/(Y'_{max} - Y'_{min})$$

where Y'_{max} and Y'_{min} are the maximum and minimum site saturation Y' values, respectively.

The Hill coefficients were determined by the Hill equation: $\mathrm{Log}[Y/1-Y] = \log K_a + n_H \log[L]$ where Y is the fractional saturation, [L] the peptide concentration, and n_H the Hill coefficient. A minimum of six data points within or near the linear portion of binding isotherm (Y vs Log [L]) were carefully chosen for the Hill plot ($\mathrm{Log}\, Y/(1-Y)$ vs Log [L]) using a linear least-squares fitting procedure by Sigma Plot software (version 8.0). The Hill coefficient was determined from the slope of the corresponding Hill plot. The apparent binding constant K_a is determined empirically as the peptide concentration at 50% fractional saturation from the binding isotherm.

4.5. Circular dichroism (CD) studies

CD spectra were measured at 37 °C with a Jasco J-815 instrument in the Institute of Chemistry, Academia Sinica. The duplex DNA was adjusted to 1.0 μM in 5 mM sodium cacodylate buffer (pH 6.5) and peptides, dissolved in the same buffer, were added to maintain final concentrations of 0.2, 1.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, 5.0 μM. CD spectra were recorded after 60 min incubation at 37 °C.

Acknowledgments

We thank Miss L.M. Hsu, National Chung-Hsing University for ESI mass analyses. This work was supported by Grant NSC97-2113-M029-005 from the National Science Council, ROC.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2010.02.047.

References and notes

1. Waring, M. J.. In *Antibiotics*; Hahn, F. E., Ed.; Springer: Berlin, 1979; Vol. 5, pp 173–194.
2. Moravek, Z.; Neidle, S.; Schneider, B. *Nucleic Acids Res.* **2002**, *30*, 1182–1191.
3. Fox, K. R.; Waring, M. J. *Methods Enzymol.* **2001**, *340*, 412–430.
4. Hudson, J. S.; Brooks, S. C.; Graves, D. E. *Biochemistry* **2009**, *48*, 4440–4447.
5. Chen, Y. H.; Lown, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 6995–7005.
6. Bailly, C.; Suh, D.; Waring, M. J.; Chaires, J. B. *Biochemistry* **1998**, *37*, 1033–1045.
7. Walker, W. L.; Landaw, E. M.; Dickerson, R. E.; Goodsell, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5634–5639.
8. Chen, F. M.; Sha, F.; Chin, K. H.; Chou, S. H. *Nucleic Acids Res.* **2004**, *32*, 271–277.
9. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983–7988.
10. Fechter, E. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8476–8485. and references cited therein.
11. Yang, C. H.; Chou, P. J.; Luo, Z. L.; Chang, J. C.; Cheng, C. C.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *Bioorg. Med. Chem.* **2003**, *11*, 3279–3288.
12. Churchill, M. E. A.; Suzuki, M. *EMBO J.* **1989**, *8*, 4189–4195.
13. Suzuki, M. *Nature* **1990**, *344*, 562–565.
14. Chang, J. C.; Yang, C. H.; Chou, P. J.; Yang, W. H.; Chou, I. C.; Lu, C. T.; Lin, P. H.; Hou, R. C. W.; Jeng, K. C. G.; Cheng, C. C.; Sheh, L. *Bioorg. Med. Chem.* **2004**, *12*, 53–61.
15. Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860–2869.
16. Yang, C. H.; Chen, W. F.; Jong, M. C.; Jong, B. J.; Chang, J. C.; Waring, M. J.; Ma, L.; Sheh, L. *J. Am. Chem. Soc.* **2004**, *126*, 8104–8105.
17. Yang, C. H.; Jeng, K. C. G.; Yang, W. H.; Chen, Y. L.; Hung, C. C.; Lin, J. W.; Chen, S. T.; Richardson, S.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *ChemBioChem* **2006**, *7*, 1187–1196.
18. Russell, R. B.; Aloy, P. *Nat. Chem. Biol.* **2008**, *4*, 666–673.
19. Palla, G.; Derenyi, E.; Farkas, T.; Vicsek, T. *Nature* **2005**, *435*, 814–818.
20. Kao, K. L.; Jonathan, C. T.; Huang, J. C. T.; Yang, C. K.; Jeng, K. C. G.; Chang, J. C. C.; Yao, W. C.; Hsien, S. C.; Waring, M. J.; Chen, M. H.; Ma, L.; Sheh, L. *Bioorg. Med. Chem.* **2010**, *18*, 366–376.
21. Koshland, D. E., Jr.; Hamadani, K. *J. Biol. Chem.* **2002**, *277*, 46841–46844. and references cited therein.
22. Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Le Minh, N.; Cox, K. K.; Nguyen, B.; Pacheco, K. A. O.; Wilson, W. D.; Lee, M. *J. Am. Chem. Soc.* **2005**, *127*, 742–750.
23. Munde, M.; Lee, M.; Neidle, S.; Arafa, R.; Boykin, D. W.; Liu, Y.; Bailly, C.; Wilson, W. D. *J. Am. Chem. Soc.* **2007**, *129*, 5688–5698.

In: Molecular Recognition
Editor: Jason A. McEvoy
ISBN 978-1-61122-734-5

Chapter 6

RECENT ADVANCES IN DNA-LIGAND MOLECULAR RECOGNITION AND ALLOSTERIC INTERACTIONS

Jonathan T. B. Huang[1], Robin C. K. Yang[1], Wei-Kang Hung[1], Michael J. Waring[2] and Leung Sheh[1] *

[1] Department of Chemistry and Life Science Research Center,
Tunghai Christian University, Taichung 407, Taiwan, R.O.C.
[2] Department of Pharmacology, University of Cambridge,
Tennis Court Road, Cambridge CB2 1PD, England

ABSTRACT

It is generally recognized that elucidating the molecular basis for recognition of specific sequences in target DNA by proteins and small synthetic molecules vitally underpins research on the modulation of gene expression. In this review we discuss the fundamental basis of DNA sequence recognition by small molecules at the atomic bonding level based on recent X-ray diffraction results together with circular dichroism spectra and footprinting experiments on DNA-small ligand binding. Monodentate (single) interactions, intrastrand bidentate interactions and interstrand bidentate interactions are considered not only central to the capabilities of small ligands and proteins to recognize DNA sequences, but also provide the means for allosteric communication between multiple DNA binding loci. Thermodynamic, kinetic and allosteric features of molecular recognition by drugs and small ligands within the minor groove of DNA are reviewed. Allosteric interactions between small synthetic peptides and multiple DNA binding sites are discussed, and hypothetical models are proposed to interpret the complex allosteric communication process. In contrast to protein-protein interaction networks which have been extensively investigated, studies on small ligand-DNA interaction networks have only recently been commenced. In this review, three different types of novel allosteric interaction networks between peptides and DNA are considered, together with hypothetical models featuring monodentate interactions and interstrand bidentate interactions. The new concept of DNA-small ligand interaction networks illuminates

* Correspondence Tel: +886-4-23590248; fax: +886-4-23590426; E-mail: Lsheh@thu.edu.tw

some basic chemical rules of DNA-small ligand allostery and may find applications in future drug design as well as structural biology research.

1. INTRODUCTION

It has long been known that the heredity of living organisms is mediated by genes contained in the nuclei of cells in the form of chromosomes. Genes are long polymers of 2'-deoxyribonucleotides in double helical array, generally stabilized by nuclear proteins. The base pairing of DNA is size-complementary, that is, the large purine bases always hydrogen-bond to the small pyrimidine bases, maintaining the AT and GC pairs having almost identical dimensions. In aqueous solution, the size-complementarity of the bases and the double helical conformation of DNA in the B-form provide a wide (major) groove about 12 Å wide running in parallel with a narrow (minor) groove of about half the width.

Sequence-specific interactions between DNA and transcription factors are central to the implementation and maintenance of genomic expression. Proteins, because of their relatively large size, bind predominantly to the major groove. Well-known DNA binding motifs include the helix-turn-helix motif, zinc finger motif, homeobox domain, and bZip motif. Early studies of interactions between DNA and the amino acid side chains of proteins were pioneered by several research groups. Seeman *et al.* identified hydrogen-bonding atoms on the edges of DNA base pairs and suggested that greater specificity was more likely to arise through amino acid side chain interactions in the major groove rather than the minor groove [1]. Suzuki established early chemical rules for the recognition of DNA bases by amino acid side chains [2]. Mandel-Gutfreund *et al*. presented a systematic analysis of hydrogen bonding between regulatory proteins and DNA [3]. Later, Thornton *et al*. presented a comprehensive study of DNA-protein interactions at the atomic level, investigating hydrogen bonds, van der Waals contacts and water-mediated bonds in 129 DNA-protein complexes [4]. Cheng and Frankel computed *ab initio* interaction energies for 21 hydrogen-bonded amino acid side chain-nucleic acid interactions [5].

Interactions of DNA bases with proteins via the amino acid side chains can be divided into several main categories: monodentate interactions (single interactions), intrastrand bidentate interactions, interstrand bidentate interactions, and complex interactions [4]. In this review we shall concentrate on the role of intra- and inter-strand bidentate interactions in DNA sequence recognition by peptides and in DNA-peptide allosteric interactions.

2. MONODENTATE INTERACTIONS AND BIDENTATE INTERACTIONS

Hydrogen bonds between DNA bases and amino acid residues of proteins or small ligands are paramount for binding and molecular recognition. They are weak non-covalent interactions with bond energy around 4.5 kcal/mol. Their weak bonding character is vital for rapid bond formation and dissociation in molecular recognition processes in biological systems. Monodentate interactions arise when a single hydrogen bond forms between one hydrogen acceptor/donor atom of the DNA base and a corresponding hydrogen donor/acceptor atom of the amino acid side chain [4]. Most proteins interact with DNA in the major groove whereas many conjugates/peptides incorporating 4-amino-1-methylpyrrole-2-

carboxylic acid residues (Py) bind preferentially in the minor groove. In the major groove, the N7 of A and G, O6 of G, and O4 of T frequently act as acceptors for hydrogen bonds, whereas the C6 amino group of A and the C4 amino group of C typically operate as hydrogen bond donors (Figure 1). In the minor groove, the N3 of A and the O2 of T and C can act as acceptors for hydrogen bonding whereas the C2 amino group of G is the only available hydrogen bond donor (Figure 1).

Bidentate interactions are further sub-divided into intrastrand and interstrand modes [4,6,7]. Intrastrand bidentate interactions refer to hydrogen bonds formed between two atoms of a base and one atom of an amino acid side chain, or between one atom of a base and two atoms of an amino acid. Intrastrand bidentate interactions were identified by Cheng and Frankel from *ab initio* interaction energies of 21 hydrogen-bonded amino acid side chains with nucleic acids [5]. The most favorable interactions are Lys-G, Arg-G, Asn-A, Ser-A and Gln-A pairings in the DNA major groove.

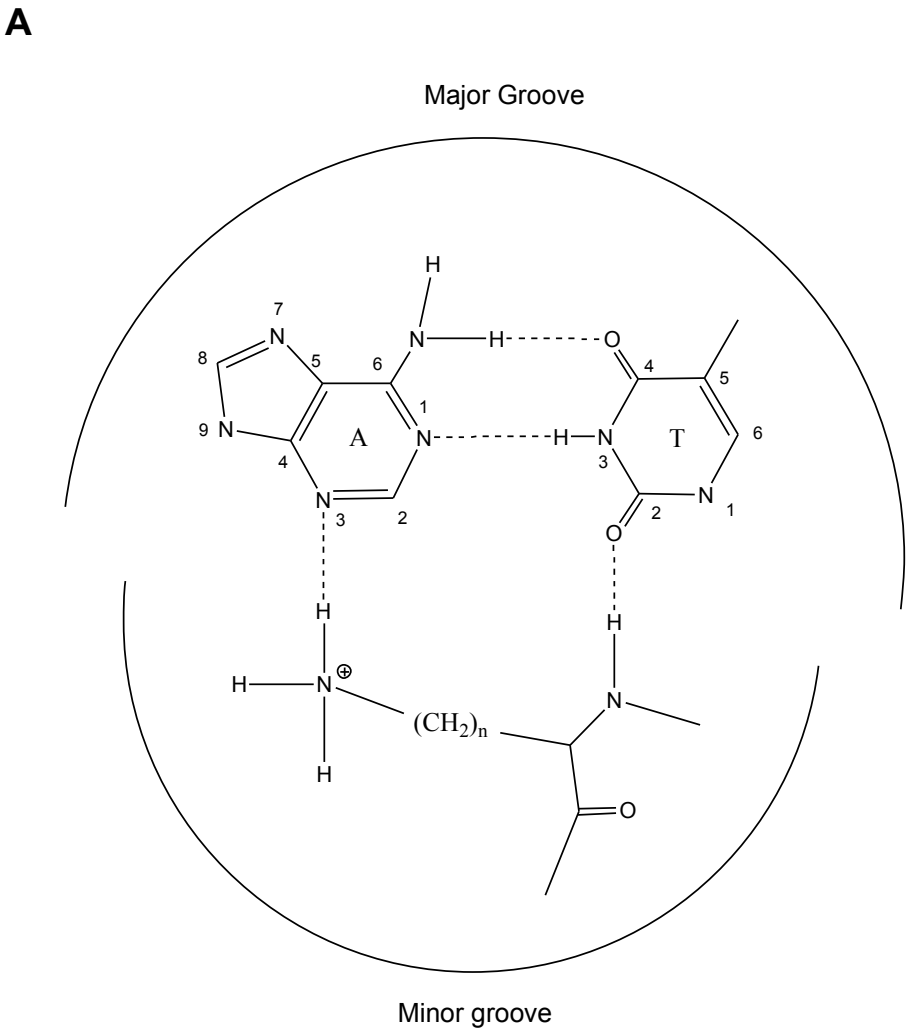


Figure 1. (Continued).

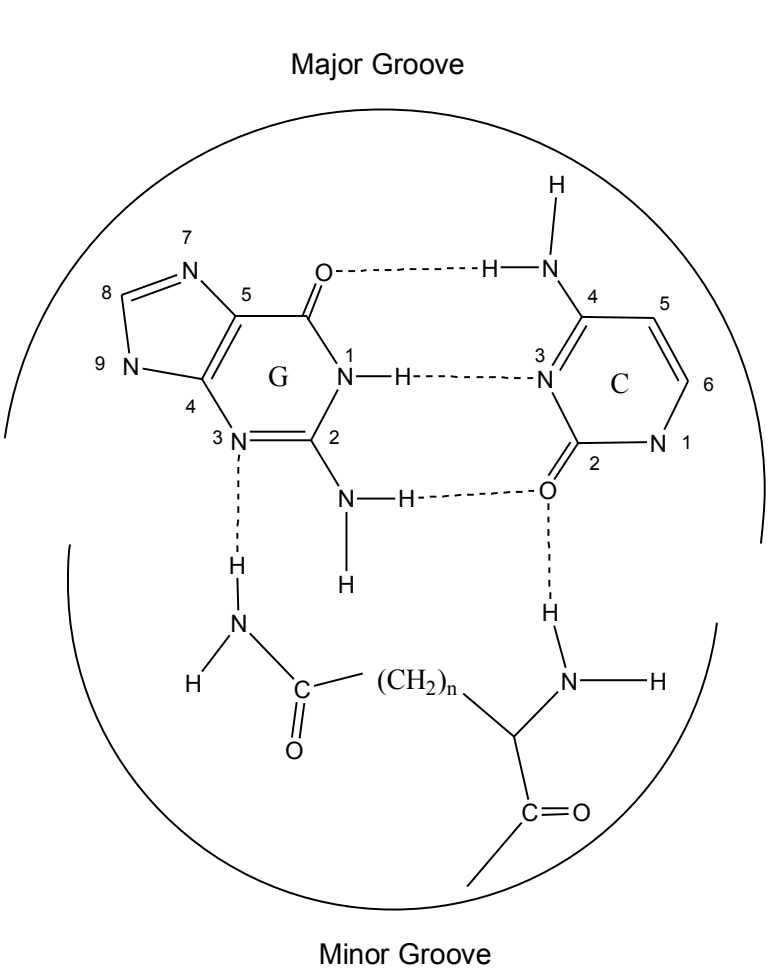


Figure 1. Interstrand bidentate interactions between peptide moieties and bases in the DNA minor groove: proposed interstrand bidentate interaction between an Orn (n = 3) or a Lys residue (n = 4) with a AT bp (panel A); proposed interstrand bidentate interactions between an Asn (n = 1) or a Gln residue (n = 2) with a GC bp (panel B) [proposed figures drawn according to X-ray data of reference [4] (Table 2)].

Interstrand bidentate interactions, on the other hand, are hydrogen bonds that form between two or three donor/acceptor atoms of an amino acid residue (including the α–amino function and side chain functional groups) with two atoms belonging to complementary bases on the two DNA strands [4] (Figure 1). Interestingly, protein residues Asn, Gln, Ser and Tyr can interact with both the major and minor grooves of DNA whereas there are several instances where Arg interacts preferentially with hydrogen bond acceptor atoms belonging to A-T pairs in the minor groove. Some years ago we reasoned that structural information deriving from DNA-protein interactions could be applied to the study of DNA-peptide interactions so far as hydrogen-bonded L-amino acid side chain-DNA interactions in the minor groove are concerned.

3. DNA-SMALL LIGAND MOLECULAR RECOGNITION IN THE MINOR GROOVE

Extensive studies employing diverse methodologies have been carried out with small ligands that are capable of sequence-specific or sequence-selective recognition of DNA. These small ligands include drugs [8-10], antitumor antibiotics [11-14], netropsin/distamycin and its analogues [15-19], and synthetic peptides [6,7, 20-24]. A deep understanding of molecular recognition of DNA by small molecules is important for disclosing molecular aspects of the modulation of gene expression *in vivo* as well as for augmenting the progress of structural biochemistry research and drug design.

An early breakthrough in molecular recognition research affecting the DNA minor groove came with the finding of the bis-intercalation of the two quinoxaline chromophores of the antitumor antibiotic echinomycin into DNA [25, 26]. DNase I footprinting and X-ray diffraction studies showed that echinomycin and its biosynthetic precursor triostin bind preferentially to CpG sequences, with hydrogen bonds forming between the 2-amino groups of guanines and the carbonyl oxygen atoms of the Ala residues, and between the N3 of guanines and the amide protons of Ala [25-27]. Recent studies have gone on to show that echinomycin specifically inhibits the DNA binding activity of hypoxia-inducible factor-1 (a transcription factor that controls genes involved in glycolysis, angiogenesis, migration, and invasion of tumors), suggesting that this antibiotic could have an important prospective in the regulation of gene expression in tumor progression and metastasis [28].

As mentioned before, the molecular basis of sequence-selective or sequence-specific interactions with DNA by drugs or other small ligands is furnished mainly by formation of adequate hydrogen bonds between functional groups of peptides or drugs and DNA bases (monodentate interactions and intrastrand/interstrand bidentate interactions). Other forms of non-covalent interaction may also occur, comprising ionic interactions between positively-charged groups of a ligand and the DNA phosphates, hydrophobic interactions, and π– interactions between intercalating aromatic groups of ligands and the DNA bases. Many synthetic minor groove-binding agents are derivatives or congeners of the antiviral antibiotics netropsin and distamycin that contain 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py). X-ray diffraction studies have shown that the 4-amino group of Py residues in these compounds acts as a hydrogen bond donor to the N3 or O2 of T, while the 3-methine proton of Py is in van der Waals contact with the C2 proton of A. The netropsin molecule fits into the minor groove with guanine stacked on a pyrrole ring of the antibiotic [29]. Lexitropsins containing two thiazole rings with the sulfur atoms directed away from the minor groove bind to alternate purine-pyrimidine sequences such as 5'-TATGAC-3' and 5'-TGCATGC-3' [30]. Similar DNA binding results were obtained with the furan-containing lexitropsins [30]. It has been shown that two distamycin molecules can associate in a side-by-side and antiparallel head-to-tail orientation within the minor groove, and this finding prompted many subsequent designs for minor groove binders such as homo- and hetero-dimeric compounds [31]. Polyamides containing the imidazole residue were found to bind preferentially to GC pairs via hydrogen bonding with the 2-amino group of G; moreover, judicious juxtaposition of Py and imidazole residues allowed G • C pairs to be differentiated from C • G pairs [32].

The thermodynamic basis of molecular recognition between small ligands and DNA has been investigated by several research teams. For many studies isothermal titration calorimetry

(ITC) was the method of choice to provide the thermodynamic parameters. In 1987, Breslauer et al. reported enthalpy-entropy compensation studies based on netropsin, distamycin, ethidium bromide, and daunomycin binding to poly[d(A-T)] • poly[d(A-T)] and poly[d(A)] • poly[d(T)] using calorimetry [33]. Binding energetics of acridine-based antitumor agents were investigated by Graves et al., showing that these drugs exhibit a minor groove binding preference [34]. The enhanced binding enthalpies of C5-substituted analogues were found to be correlated with anticancer activity. Replacing various functional groups on anthracycline antibiotics resulted in a binding free energy penalty, revealing that total ligand binding energies are partitioned among various substituents in these antibiotics [35]. Enthalpy-entropy compensation has been reported in the binding of 7-amino actinomycin D to several short DNA duplexes [36]. Wilson et al. showed that a number of reversed amidine heterocyclic compounds bound to the minor groove of DNA in a mostly entropy-driven manner, with positive $T\Delta S$ values around 4.3-6.0 kcal/mol [37].

Kinetic aspects of the interaction of DNA with drugs or other small molecules were early explored by several research teams employing complementary approaches. As long ago as 1976, Davanloo and Crothers reported the kinetics of formation and dissociation of drug-dinucleotide complexes and actinomycin-nucleotide complexes [38]. In 1981, Fox, Wakelin and Waring reported that the speeding up of rate constants with an increasing level of echinomycin binding to DNA could indicate the occurrence of cooperativity associated with progressive changes in the DNA structure [39]. Stopped-flow kinetic studies on echinomycin-DNA binding showed a single exponential and suggested a molecular mechanism of binding in which both chromophores of the antibiotic become intercalated simultaneously rather than sequentially, and no transition from a mono-intercalated state to a bis-intercalated state could be detected [39]. Using stopped-flow and temperature-jump approaches, Chaires et al. proposed the mechanism of daunomycin-DNA interaction to be two-stepped: a rapid "outside" binding of daunomycin to DNA, followed by intercalation of the drug [40]. Employing a fluorescence-detected stopped-flow assay, Crothers and coworkers studied the kinetics of a series of polyamides binding to DNA and pointed out that covalent linkage of the subunits results in polyamides with dramatically enhanced affinity due to faster association rates [41]. Moreover, the basis for discrimination between matched and mismatched sites for each polyamide was found to arise mainly from differences in the rates of dissociation from these sites. Fletcher and Fox developed a convenient method for studying the dissociation kinetics of echinomycin from CpG binding sites in different sequence environments monitored by the rate of disappearance of DNase I footprints [42]. Surface plasmon resonance (SPR) appears to be another appealing approach for studying DNA-ligand interactions since equilibrium constants, association and dissociation rate constants for monomer and dimer complex formation can be determined [37]. SPR also allows binding rates to be observed in real time for complex formation and dissociation. Wilson, and Lee et al. have been using SPR successfully to analyze the kinetics of reversed amidine heterocycles and polyamides binding to various oligonucleotide duplexes [37].

Against this background we sought to investigate the molecular basis of sequence recognition as well as complex allosteric aspects of DNA-peptide interactions by designing and synthesizing a number of peptides incorporating the XPRK or XHypRK motifs together with a tract of 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) [6,7,20-24]. Synthetic peptides based upon the XPRK or XHypRK motifs possess good DNA sequence-selective binding capability toward DNA sequences incorporating three or four consecutive

W (A or T) base pairs and afford satisfactory DNA footprinting results at sub-micromolar concentrations. These peptides also possess the advantage of good aqueous solubility and obviate the frequent use of organic solvents in DNA-small molecule binding studies.

Our logic for the design of the XPRK motif stems from the fact that a naturally occurring SPXX motif was found in repeating sequences in histones, steroid hormone receptors, various segmentation gene products and some oncogene products [43,44]. It was suggested that the SPXX motif assumes a β-turn stabilized by two hydrogen bonds, and that the side chains of the two basic residues engage in salt bridges with the DNA phosphate groups [43,44].

Quantitative DNase I footprinting originally turned out to be the best methodology for DNA-peptide sequence recognition experiments since most other methods were unable to discriminate peptide binding at multiple recognition sites. In addition, to investigate conformational changes involved in the molecular recognition of DNA we resorted to circular dichroism measurements which provide spectral characterization of DNA-peptide interactions [6,7].

4. DNA-LIGAND ALLOSTERIC INTERACTIONS

Allostery is generally recognized as an indispensable process in biological control regulating biochemical efficiency and energy metabolism in nature. For almost sixty years extensive research has been devoted to understanding the allosteric behavior of functional proteins and regulatory enzymes. The complex network-based allosteric regulation of ATP/GTP and NADH/$FADH_2$ generation in glycolysis and the tricarboxylic acid cycle are representative of 'biological energy economy' that is vital for the maintenance of life in higher animals including man. The term 'allosteric' or 'allostery' was originally used to describe a macromolecule-ligand interaction that results in local conformational change of the macromolecule, thereby facilitating the binding of further ligand molecules to the binding subunits/sub-sites (positive cooperativity), or making it more difficult for subsequent ligand molecules to bind to those subunits/sub-sites (negative cooperativity). The concerted-symmetry allosteric model of Monod, Wyman, and Changeux (MWC model) proposed that all of the protein subunits change shape in a concerted manner to preserve the symmetry of the entire molecule as it is transformed from the low affinity conformation (T state) to the high affinity conformation (R state) [45]. The sequential/induced fit allosteric model of Koshland, Nemethy, and Filmer (KNF model) proposed that each subunit changes shape as ligand binds, so that changes in one subunit lead to distortions in the shape and/or interactions of other subunits within the protein [46]. The KNF model appears to provide a better explanation of the phenomenon of 'negative cooperativity' than the MWC model. Williams et al. used a thermodynamic approach to define positive cooperativity as the decreased dynamic behavior of a receptor system with a benefit in enthalpy and a cost in entropy, and negative cooperativity as the increased dynamic behavior of a receptor system with a cost in enthalpy and a benefit in entropy [47,48]. Ever since 1960, while numerous papers on protein-ligand allosteric interactions have been published, both the MWC and KNF models have been widely used to interpret complex allosteric behavior of proteins.

On the other hand, studies on possible allosteric aspects of DNA-ligand interactions are far fewer, and much less is known about their results than protein-ligand interactions.

However, some early pioneering studies did establish a significant degree of progress towards understanding how allostery might influence DNA-ligand interactions. In 1980 Crothers et al. reported that the binding of distamycin to calf thymus DNA at low levels of saturation is positively cooperative [49]. Three years later Graves and Krugh reported that adriamycin and daunomycin also exhibit positive cooperative binding to various DNAs [50]. The appearance of initial positive curvature in the binding isotherm was found to be dependent on the ionic strength, suggesting a role for DNA flexibility in positive cooperativity. In 1992 Fagan and Wemmer reported the use of ^{1}H NMR spectroscopy to study the cooperative binding of distamycin to d(CGCIIICCGGC) + d(GCCIICCCGCG), where I represents inosine [51]. In 1996 Bailly, Hamy and Waring reported cooperativity in the binding of the antitumor antibiotic echinomycin to DNA using quantitative DNase I footprinting [52]. This study demonstrated that the binding of echinomycin to the sequences ACGTACGT and TCGAACGT is highly cooperative. Also in 1996, Morii et al. reported cooperative binding to DNA of peptides conjugated to adamantyl and β-cyclodextrin groups using gel mobility shift and circular dichroism assays [53]. In 2000 Chaires et al. reported that (+)daunorubicin binds selectively to right-handed DNA and can provoke apparently allosteric conversion of left-handed polynucleotide to a right-handed intercalated form, as well as *vice versa* for a synthetic (-)analogue [54]. Shortly afterwards Laughton et al. used a molecular dynamics approach to investigate cooperativity in drug-DNA recognition [55]. Their results led them to support Cooper and Dryden's hypothesis of allosteric communication without changes [56] in the time-average structure of the macromolecule, though binding free energy can be obtained from changes in conformational flexibility alone. In 2003 Fechter and Dervan used quantitative DNase I footprinting to investigate allosteric inhibition of GCN4 bZip protein binding to the DNA major groove by polyamide-acridine conjugates that bind to the minor groove [57].

In recent years we have been studying possible allosteric features of peptide-DNA interaction using designed peptides targeted to specific recognition sites in DNA [6,7,23,24]. Our experimental protocol is based upon quantitative DNase I footprinting of designed peptides binding to a 5'-^{32}P-labeled 158-mer DNA duplex and a complementary 5'-^{32}P-labeled 135-mer DNA duplex. Binding site positions on the upper and lower strands are designated as U and L, respectively. Like proteins, monodentate interactions and interstrand bidentate interactions between the ligands and the DNA bases have been recognized as essential contributors to molecular recognition of DNA by our synthetic peptides.

To illustrate the power of quantitative DNase I footprinting as one of the best analytical techniques for studying DNA recognition and allostery, we offer the following example. A designed decapeptide His-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyH-10) was footprinted using DNase I and serial peptide dilutions between 10 and 3000 nM [7]. The peptide was equilibrated with the DNA for an hour before enzyme cleavage, then the reaction was stopped, the gel was electrophoresed, and subjected to autoradiography (Figure 2). In previous reports [6,7,24], we describe how the position of interstrand bidentate interactions can be assigned by inspecting differential cleavage plots (experiment versus control) so that regions of significant simultaneous blockage of DNase I cutting on the complementary strands can be cross-correlated; these we indicate by pecked lines connecting base sequences on the two strands (Figure 3).

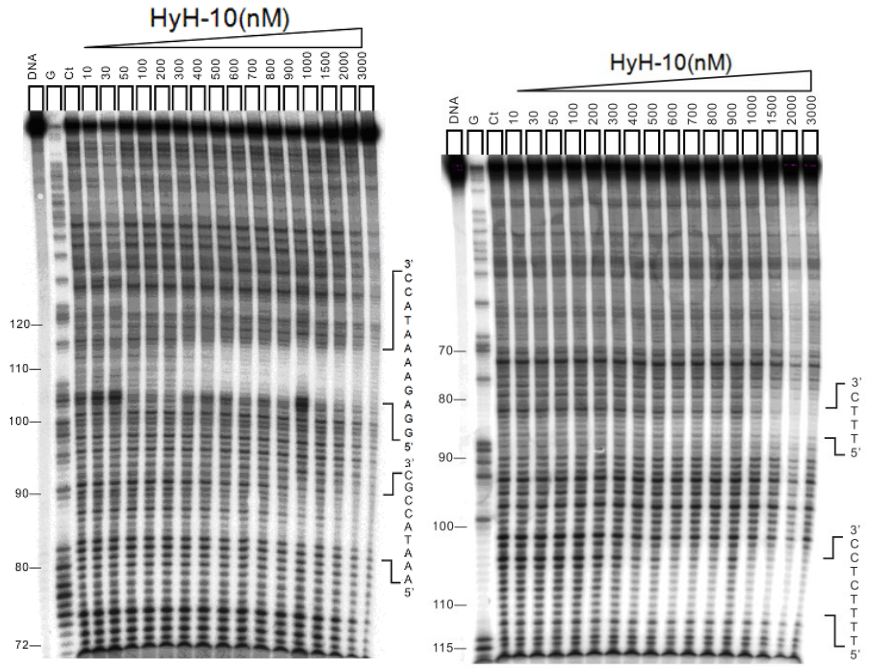


Figure 2. Example of a quantitative DNase footprinting autoradiograph showing the sequence selective binding of synthetic peptide HyH-10 on DNA duplexes labeled at either end: 5'-[^{32}P]-labeled 158-mer upper strand, left panel; and 5'-[^{32}P]-labeled 135-mer lower strand, right panel. Peptide HyH-10 was equilibrated with the DNA in 5 mM sodium cacodylate buffer, pH 6.5 at 37°C for 60 min before DNase 1 cleavage. G represents a Maxam-Gilbert guanine sequencing track and Ct shows a DNase I digestion control lane.

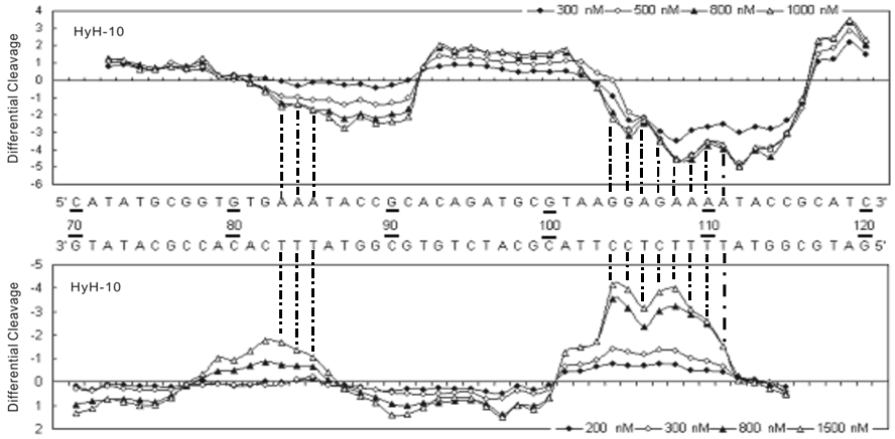
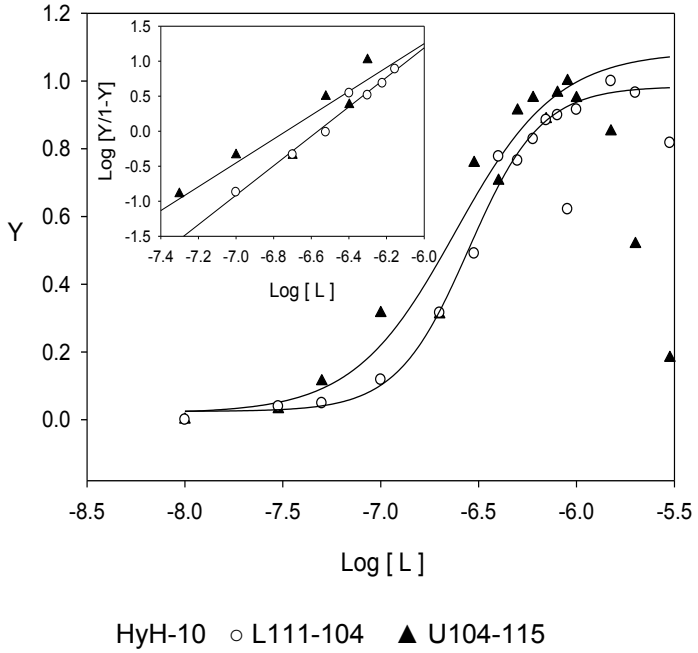


Figure 3. Differential cleavage plots comparing the susceptibility of DNA fragments to DNase I cleavage after incubation with peptide HyH-10 in cacodylate buffer at 37°C for 60 min. The upper traces represent the differential cleavage plot for a given peptide bound to the 5'-[^{32}P]-labeled upper strand (158-mer) DNA fragment; the lower traces represent the corresponding plots for each peptide bound to the 5'-[^{32}P]-labeled lower strand (135-mer) DNA fragment. The vertical dotted lines between DNA bases represent assignment of interstrand bidentate interactions where significant coincident H-bonding interactions occur involving complementary bases in both strands.



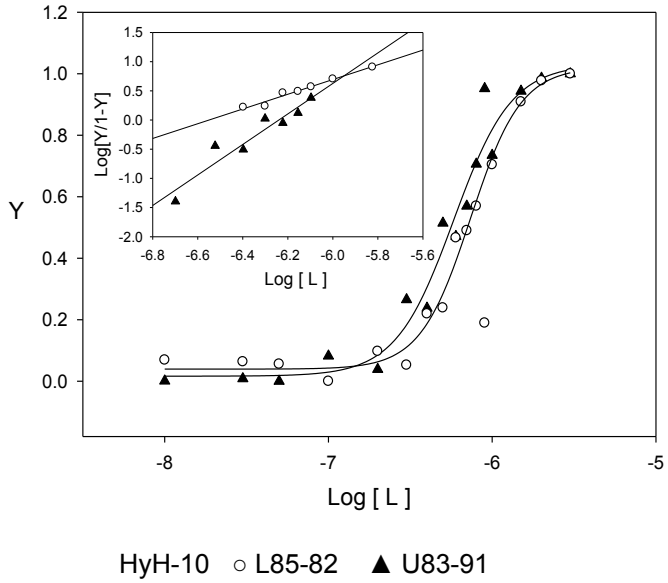


Figure 4. Binding isotherms from DNase I footprinting titrations of peptide HyH-10 (Y, the fractional saturation *versus* log molar concentration) for binding sites L111-104 and U104-115 (upper panel), and sites L85-82 and U83-91 (lower panel). Inset are the respective Hill plots of log[Y/1-Y] *versus* log[L]. The concentration of ligand (peptide) is in nM. The data points of Y at high peptide concentrations are scattered due to over-saturation.

Monodentate interactions are simply assigned by connecting regions of inhibition in the differential cleavage plot profile to bases where there is no simultaneous blockage on the complementary strand.

A Hill plot of log Y/(1-Y) *versus* log [L] yields the Hill coefficient (n_H) which records the degree of cooperativity of the binding process and allows us to compare cooperativity with respect to different binding sites as well as the relative cooperativity of different designed peptides binding to various discrete sites (see section 5, Figure 4).

The autoradiograph (Figure 2) and the differential cleavage plot (Figure 3) reveal that, on the upper strand of the 158-mer pBR322 fragment peptide HyH-10 produces a strong and very broad DNase I blockage around position U104-115, corresponding to the sequence 5'-GGAGAAAATACC-3', displaying positive cooperativity (n_H = 1.7), and there is also a broad DNase I blockage site around position U83-91, comprising the sequence 5'-AAATACCGC-3', where notable positive cooperativity (n_H = 2.6) is evident (The binding site positions on the upper and lower strands are abbreviated as U and L, respectively). On the lower strand, two regions of DNase I blockage appear around positions L85-82 and L111-104. Interstrand bidentate interactions are accordingly assigned around positions 83-85 and 104-111 (Table 1).

The wide areas of DNase cleavage inhibition around positions U104-115 and L111-104 that encompass 12 base pairs suggest that two peptide molecules bind in dimeric fashion here. Consistent with this, the binding isotherms at individual binding sites exhibit sigmoidal curvature, typical of positive cooperativity (Figure 4).

Reviewing hundreds of footprinting experiments, we have frequently observed an abrupt increase in the intensity of footprints as the peptide concentration is raised slightly. This is a clear indication of significant positive cooperativity between peptide molecules binding to adjacent sub-binding sites. Peptides incorporating the XPRK motif interact preferentially with d(AAAA)-d(TTTT) or d(AAA)-d(TTT) sites.

Footprinting studies that indicate a wide binding locus spanning over 12 base-pairs can really only be explained on the basis of closely adjacent sites that might act independently, but are most likely to accommodate two peptide molecules in a dimeric binding mode. A dimeric binding pattern is consistent with independent circular dichroism studies. Thus, in this case the apparent large single binding site appears to be composed of two adjacent sub-binding sites on each complementary strand, representing a total of four sub-sites at each locus (Figure 5).

The conformations of two adjacent sub-binding sites affect each other the most since they are linked covalently. Bonds linking opposite sub-sites are interstrand bidentate hydrogen bonding and exert less conformational influence upon one another than do those linking adjacent sub-sites. Conformational influences between neighboring binding sites also depend on the intervening distance between them. Recent footprinting experiments show that the n_H values for peptides PyHyp-12 and PyHyp-9 binding to 81-mer duplexes (S-81) containing a single d(AAAA)-d(TTTT) binding locus are significantly higher than those of 158-mer and 135-mer duplexes containing two binding loci d(AAAA)-d(TTTT) and d(AAA)-d(TTT) [6]. This was a first indication that the allosteric relay of peptide binding information between neighboring d(AAA)-d(TTT) and d(AAAA)-d(TTTT) positions spanning an intervening sequence of 12-14 base pairs on the same DNA strand might be affected by some negative cooperative effect.

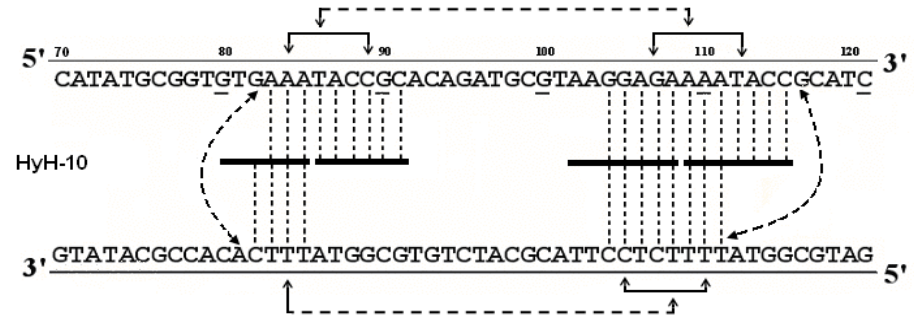
Indeed, it appears that allosteric influences affecting peptide/DNA binding sites may produce either positively or negatively cooperative binding. Most of the allosteric effects found in our studies lead to positive cooperativity.

Table 1. Binding specificity, physicochemical parameters and type of allosteric interaction network for recognition of designed peptides on complementary 5'-[^{32}P]-labeled upper (158-mer) and 5'-[^{32}P]-labeled lower (135-mer) DNA strands at 37°C determined by quantitative DNase I footprinting

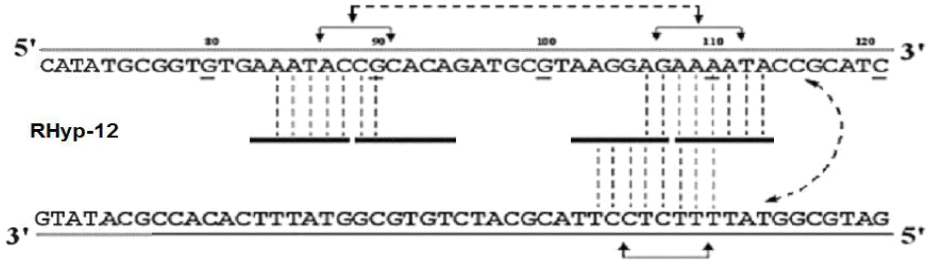
Ligand	Binding Position	Recognition Sequence	K_a	n_H	Position of Interstrand Bidentate interactions	Type of Interaction Network
RHyp-12	U84-90 U106-113 L110-103	5'-AATACCG-3' 5'-AGAAAATA-3' 5'-TTTCTCCT-3'	3.0×10^6 7.6×10^6 7.4×10^6	1.5 2.1 2.0	106-110	Partial Circuit
PyHK-10	U86-89 U105-115 L111-105	5'-TACC-3' 5'-GAGAAAATACC-3' 5'-TTTTCTC-3'	4.9×10^6 2.7×10^6 3.4×10^6	0.8 2.4 1.4	105-110	Partial Circuit
HyH-10	U83-91 U104-115 L85-82 L111-104	5'-AAATACCGC-3' 5'-GGAGAAAATACC-3' 5'-TTTC-3' 5'-TTTTCTCC-3'	1.8×10^6 4.7×10^6 1.4×10^6 3.6×10^6	2.6 1.7 1.3 2.0	83-85 104-111	Circuit
PyHR-9	L85-79 L112-104	5'-TTTCACA-3' 5'-ATTTTCTCC-3'	8.4×10^6 9.7×10^6	5.0 10.6	NIL	Non-circuit
PyHyp-9	U84-89 U107-113 L84-83 L111-103	5'-AATACC-3' 5'-GAAAATA-3' 5'-TT-3' 5'-TTTTCTCCT-3'	2.5×10^6 3.8×10^6 2.7×10^6 2.2×10^6	2.2 2.2 2.0 2.0	84 107-111	Circuit
PyHyp-12	U84-88 U106-114 L110-103	5'-AATAC-3' 5'-AGAAAATAC-3' 5'-TTTCTCCT-3'	2.5×10^6 3.4×10^6 3.4×10^6	1.5 1.4 1.4	106-110	Partial Circuit
PyPro-12	U84-90 U106-116 L85-83 L110-105	5'-AATACCG-3' 5'-AGAAAATACCG-3' 5'-TTT-3' 5'-TTTCTCC-3'	6.7×10^7 3.2×10^6 3.7×10^6 3.8×10^6	0.7 1.2 1.1 1.5	84-85 106-110	Circuit
HypKK-10	U83-90 U107-116 L110-104 L85-81	5'-AAATACCG-3' 5'-AGAAAATACC-3' 5'-TTTCTCC-3' 5'-TTTCA-3'	2.0×10^6 2.4×10^6 3.3×10^6 1.6×10^6	2.7 2.0 2.3 -	107-110	Partial Circuit

K_a and n_H are the apparent association constant and Hill coefficient determined from concentration-dependent DNase I footprinting studies, respectively. The binding site positions on the upper and lower strands are abbreviated as U and L, respectively. Interstrand bidentate interactions are assigned where there are coincident effects on complementary nucleotides in both strands.

Circuit type interaction network



Partial circuit type interaction network



Non-circuit type interaction network

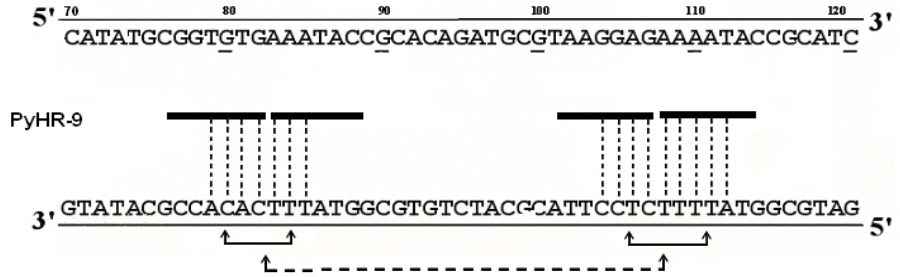


Figure 5. Three different types of allosteric interaction network models for cooperative binding of peptides to DNA fragments based on quantitative footprinting studies. The portion of the ligand binding to each DNA site/sub-site is represented by a thick horizontal line. Monodentate interactions and interstrand bidentate interactions are represented by vertical broken lines. The solid horizontal arrow lines represent communication of allosteric interaction between DNA sub-binding sites. Broken horizontal arrow lines between neighboring binding loci some 12-16 nucleotides apart are intended to represent weak negative cooperative communication.

In these cases, binding of a peptide molecule to a given sub-site generates local DNA conformational changes that enhance the binding affinity of another peptide molecule at an adjacent sub-site. In the negatively cooperative case outlined above, binding of two peptide molecules to a first site produces local DNA conformation changes that decrease the binding affinity of one or two peptide molecules at a different site about 12-18 base-pairs away.

We have also found that interstrand bidentate hydrogen bond interactions may relay conformational changes from one site or sub-site to the opposite site in the complementary strand, and may bestow negative cooperativity in this manner.

Two peptide molecules could be capable of binding in a dimeric head-to-head or head-to-tail fashion to some or all of these sites (Figure 5). Several modes of interaction with complementary or adjacent binding sites can be envisaged that might lead to cooperative binding. Interaction of one peptide molecule in a sequential manner with one sub-site (inside a binding site) on one DNA strand could facilitate the binding of a second molecule to the other sub-site on the same strand. On the other hand, binding of a peptide molecule to a first site might induce local DNA conformational change(s) mediated by interstrand bidentate interactions and facilitate sequential or concerted binding of further peptide molecules to an opposite site on the complementary strand. Of course, binding of peptide molecules to one site could invoke conformational change(s) and facilitate or hinder peptide binding to a neighboring site on the same strand. Thus, cooperative (or anti-cooperative) binding distributed over an array of binding sites may establish a network of cooperativity interconnecting several allosteric sites (Figure 5).

To gain direct insight into possible conformational changes of DNA associated with the sequence-selective binding of peptides we carried out circular dichroism measurements (CD) using a number of synthetic peptides (the amino acid sequences of some of these peptides are shown in Figure 8). As an example we present CD results for a decapeptide, HyH-10: His-Hyp-Arg-Lys-$(Py)_4$-Lys-Arg-NH_2 (HyH-10) [7]. A 13-mer deoxyribonucleotide duplex d(TAGGAGAAAATAC)- d(GTATTTTCTCCTA) (U4A-L4T) was synthesized to contain a sequence of 12 base-pairs that correspond to the binding site at position 103-114 of the pBR322 fragment. The CD spectra for the DNA and peptides alone are shown in Figure 6A. Titration of peptide HyH-10 versus the duplex U4A-L4T induces a negative CD band around 247 nm and a positive band around 284 nm. A strong, broad, dose-dependent positive CD enhancement appears around 330 nm (Figure 6B, C).

In the difference spectrum (Figure 6C), this CD band around 330 nm is unchanged. Another induced positive CD band is red-shifted from 284 nm to 288 nm. It has been reported that molecules which bind to the minor groove typically exhibit an induced positive CD band around 320 nm [37]. Accordingly the spectra indicate strongly that our designed peptides incorporating the XPRK/XHyPRK and polyamide motifs bind preferentially to the minor groove of the double helix.

It is evident that the peptide-induced enhancement of CD bands is correlated with peptide-DNA binding stoichiometry. Figure 7 shows a plot of [peptide]/[DNA duplex] versus $\Delta\theta$ at 322 nm for four synthetic peptides: RHyp-12, PyHK-10, HyH-10 and PyHR-9 [7]. The decapeptide PyHK-10 induces the greatest ellipticity change among the four peptides, suggesting that most far-reaching DNA conformational changes are associated with the binding of this peptide in the minor groove. On the other hand, significantly greater ellipticity changes are induced by peptides PyHK-10, RHyp-12 and HyH-10 compared to PyHR-9,

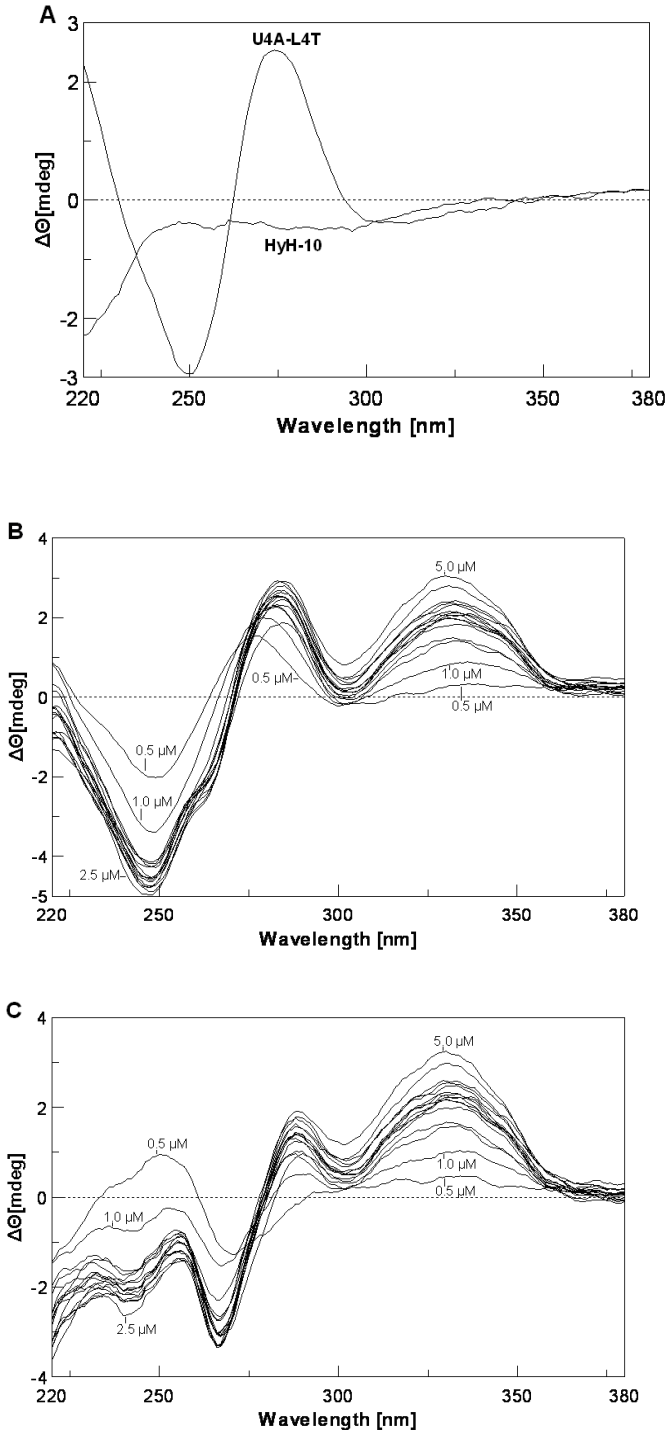


Figure 6. Panel A: CD spectra of DNA duplex d(TAGGAGAAAATAC)- d(GTATTTTCTCCTA) (U4A-L4T) alone and peptide alone. Panel B: Titration of duplex U4A-L4T versus peptide HyH-10 at peptide concentrations of 0.2, 1.0. 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, 5.0 μM. Panel C: Corresponding CD difference spectra with the contribution of duplex and peptide HyH-10 subtracted.

indicating that the nonapeptide is much less able to perturb the conformation of the helical minor groove. The ellipticity changes induced by peptides RHyp-12 and HyH-10 are of comparable magnitude.

From the Δθ *versus* [peptide]/[DNA duplex] plot (Figure 7), it appears that at peptide/DNA ratios of about 0.5 to 2.0, one molecule of peptide binds to the d(AAAA)-d(TTTT) locus. At peptide/DNA ratios of 2 - 3, as indicated by the small plateau region in the titration curve, two peptide molecules seem to bind ostensibly as a dimer in the minor groove. At ratios above 4, a progressive increase in Δθ suggests that more peptide molecules begin to engage in some sort of non-sequence-selective binding to DNA. The CD results agree well with extensive DNase I footprinting results that dimeric peptide binding to the d(AAAA)-d(TTTT) locus is favored, as shown by the very wide binding locus spanning a distance of 9-12 base pairs (Figures 3,5, Table 1).

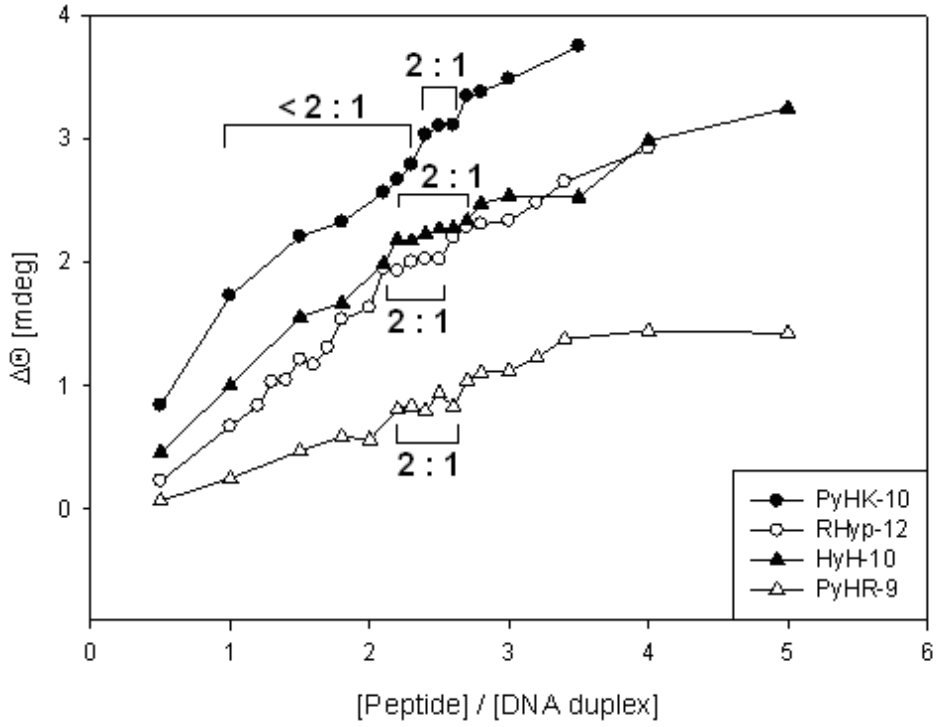


Figure 7. CD intensity at 322 nm as a function of [peptide]/[DNA duplex] The proposed stoichiometric binding ratios are as indicated, with binding below 2:1 (<2:1) considered to be predominantly 1:1.

Isothermal titration calorimetry (ITC) studies of several designed peptides incorporating the XPRK motif revealed that the binding reaction is strongly exothermic [58]. Preliminary surface plasmon resonance (SPR) studies with an oligonucleotide hairpin have shown that two peptides bind to the d(AAAA)-d(TTTT) sequence with $t_{1/2}$ values around 2-2.5 min (depending on the initial concentration of the peptide) and with k_{a1} and k_{a2} around 180-300 $M^{-1}s^{-1}$, suggesting that the binding of these designed peptides to their preferred DNA binding sites is a very slow process compared to other small molecules [58].

5. DETERMINATION OF DNA-PEPTIDE BINDING COOPERATIVITY BY THE HILL PLOT

The apparent DNA binding site saturation is determined by the equation [19].

$$Y' = 1 - (I_{tot}/I_{ref})/(I_{tot}^{o}/I_{ref}^{o})$$

where Y' is the fractional saturation, I_{ref} and I_{ref}^{o} are integrated band volumes over 5 bases of a running lane (non-binding site, with peptide) and control lane (without peptide), respectively. I_{tot} and I_{tot}^{o} are integrated band volumes of the binding site locus and corresponding control lane locus, respectively.

The value of Y is optimized by the following equation:

$$Y = (Y' - Y'_{min})/(Y'_{max} - Y'_{min})$$

where Y'_{max} and Y'_{min} are the maximum and minimum site saturation Y' values, respectively.

The Hill coefficients which indicate the magnitude of cooperativity between binding sub-sites are determined by the Hill equation:

$\log [Y/1-Y] = \log K_a + n_H \log[L]$ where Y is the fractional saturation, [L] is the peptide concentration, and n_H the Hill coefficient. A minimum of six data points within or near the linear portion of the binding isotherm (Y versus log [L]) were chosen for the Hill plot (log Y/(1-Y) versus log [L]) using a linear least-squares fitting procedure. The Hill coefficient was determined from the slope of the corresponding Hill plot (see Figure 4). The apparent binding constant K_a is determined empirically from the binding isotherm as the peptide concentration at 50% fractional saturation.

6. DNA-PEPTIDE ALLOSTERIC INTERACTION NETWORKS

Current studies of biomacromolecule-ligand allostery have progressed to the so-called "interaction network" stage. Allostery appears to interconnect different metabolic and biosynthetic pathways through interaction of various bio-molecules: namely, protein-protein interactions [59,60], protein-small ligand interactions [61], protein-nucleic acid interactions [62], and nucleic acid-small ligand interactions. In addition, many of these complex network interactions may communicate with one another, resulting in significant net biochemical and physiological phenomena. In other words, there are few biochemical processes within the cell that are independent of allosteric or regulatory influences from other metabolic pathways or from other bio-molecules. The past decade has been remarkable for an explosion of interest in network-based interactions between bio-molecules that effectively regulate biochemical and physiological processes in nature. High-throughput studies have contributed much to establishing protein-protein interaction networks, metabolic networks, and transcription regulatory networks [59].

To learn more about the chemical rules that govern DNA-ligand allostery we employed quantitative footprinting to embark upon a study of molecular aspects of cooperative interactions between peptides and multiple DNA binding sites. Thus, in 2006 we proposed an original network-based DNA-peptide allosteric model featuring intercommunicating allosteric binding sites in fragments of the latent membrane protein-1 gene from a pathogenic Epstein-Barr virus variant derived from nasopharyngeal carcinoma [23]. In recent studies we have further explored whether such network-based allosteric behavior can be detected in common DNA motifs containing consecutive A's and T's using our 158-mer and 135-mer pBR322 fragments [6,7].

His-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-Hyp-His-NH$_2$ (RHyp-12)

His-Pro-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (PyHK-10)

His-Hyp-Arg-Lys-(Py)$_4$-Lys-Arg-NH$_2$ (HyH-10)

His-Pro-Arg-Lys-(Py)$_4$-Arg-NH$_2$ (PyHR-9)

His-Hyp-Arg-Lys-(Py)$_3$-His-Hyp-Lys-Arg-NH$_2$ (PyHyp-9)

His-Hyp-Arg-Lys-(Py)$_4$-His-Hyp-Arg-Lys-NH$_2$ (PyHyp-12)

His-Pro-Arg-Lys-(Py)$_4$-His-Pro-Arg-Lys-NH$_2$ (PyPro-12)

His-Hyp-Arg-Lys-(Py)$_4$-Lys-Lys-NH$_2$ (HypKK-10)

Figure 8. Sequences of eight designed peptides used in quantitative DNase 1 footprinting studies (Py = 4-amino-1-methylpyrrole-2-carbonyl-).

Based on quantitative footprinting results with eight designed peptides (Figure 8) on the complementary upper and lower DNA strands summarized in Table 1, we constructed network-based models in an effort to interpret the complex communication between DNA-peptide allosteric binding sites (Figure 5). The duplex fragments used here contain two conspicuous peptide binding loci: around position 83-89, comprising the d(AAA)-d(TTT) sequence, and position 103-116 comprising the d(AAAA)-d(TTTT) sequence. Each locus seems to afford two sites on complementary strands that consist of two to four sub-sites, depending on individual peptides. In our model, conformational changes induced at one binding site affect subsequent peptide binding to a neighboring site, and interstrand bidentate interactions are envisaged as playing an important role in relaying conformational changes between sites on the complementary strands, facilitating or inhibiting the binding of further peptide molecules to sub-sites. The different models for these eight peptides binding to DNA primarily reflect differences in the relay of conformational information between opposite sites on complementary strands, mediated via interstrand bidentate interactions. It can be concluded that peptides HyH-10, PyHyp-9 and PyPro-12 form a circuit type of a communication: that is to say, the allosteric communication between binding sites is complete. For peptides RHyp-12 , PyHK-10, PyHyp-12 and HypKK-10, the networks are referred to as incomplete-circuit type, because the allosteric communications form a partial or incomplete circuit. At the other extreme, synthetic peptides PyHR-9, $(HPRK)_3NH_2$ [HR-12] [23], $(SPRK)_3NH_2$ [SP-12] [23] whose binding to DNA fragments lacks interstrand bidentate interactions altogether, appear to engage in a non-circuit type of allosteric communication.

To summarize, the quantitative footprinting results in these studies support our hypothesis that three different types of network-based allosteric communication can be distinguished in peptide-DNA recognition: circuit type, incomplete-circuit type and non-circuit type [6,7]. The finding that interaction networks of allosteric effects exist in DNA-peptide molecular recognition serves to widen and enlarge our understanding of the fundamental nature of macromolecule-small ligand interactions. Provided that the intervening distance between binding sites lies within about 15 bp, conformational changes can readily be transmitted from one site to another. If more than two sites are sufficiently close, one can envisage that allosteric effects could be propagated over a substantial distance along the DNA duplex. This network concept provides some basic chemical rules of DNA-small ligand allosteric interaction that should find applications in future drug design and structural biology research.

ACKNOWLEDGMENTS

We thank professors L.S. Kan, Y.T. Tau, S.T. Chen, S.H. Wu and D.K. Chang of Academia Sinica for helpful advice and access to ITC, CD and SPR facilities. This work was supported by grants NSC97-2113-M029-005 and TCVGH-T967810.

REFERENCES

[1] Seeman, N.C. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA* 1976, *73,* 804-808.

[2] Suzuki, M. A framework for the DNA-protein recognition code of the probe helix in transcription factors: the chemical and stereochemical rules. *Structure* 1994, *2,* 317-326.

[3] Mandel-Gutfreund, Y.; Schueler, O.; Margalit, H. Comprehensive analysis of hydrogen bonds in regulatory protein–DNA complexes: in search of common principles. *J. Mol. Biol.* 1995, *253,* 370-382 and references cited therein.

[4] Luscombe, N.M.; Laskowski, R.A.; Thornton, J.M. Amino cid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* 2001, *29,* 2860-2869 and references cited therein. .

[5] Cheng, A.C.; Frankel, A.D. Ab initio interaction energies of hydrogen-bonded amino acid side chain-nucleic acid base interactions. *J. Am. Chem. Soc.* 2004, *126,* 434-435 and references cited therein.

[6] Kao, K.L.; Jonathan C.T. Huang; J.C.T.; Yang, C.K.; Jeng, K.C.G.; Chang, J.C.C.; Yao, W.C.; Hsien, S.C.; Waring, M.J.; Chen, M.H.; Ma, L.; Sheh, L. Detection of Multiple Network-based Allosteric Interactions between Peptides and Arrays of DNA Binding Sites. *Bioorg. Med. Chem*. 2010, 18, 366-376.

[7] Huang, T.B.; Chen, Y.C.; Chang, J.C.; Jeng, K.C.; Kao, K.K.L.; Yang, R.C.K., Kan, L.S.; Wey, M.T.; Waring, M.J.; Chen, C.S.; Chien, W.J.; Sheh, L. Novel DNA-peptide interaction networks. *Bioorg. Med. Chem*. 2010, 18, 2575-2585.

[8] Fox, K.R.; Waring, M.J. Meth. Enzymol. Use of DNA molecules substituted with unnatural nucleotides to probe specific drug-DNA interactions. 2001, *340,* 412-430.

[9] Moravek, Z.; Neidle, S.; Schneider, B. Protein and drug interactions in the minor groove of DNA. *Nucleic Acids Res.,* 2002, *30,* 1182-1191.

[10] Bailly, C.; Chaires, J.B. Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues. *Bioconjugate Chem.,* 1998, *9,* 513-538 and references cited therein.

[11] Waring, M.J. (1979) Echinomycin, triostin and related antibiotics. In *Antibiotics,* Hahn, F.E. Eds. Springer-Verlag. Berlin, vol. 5, pp. 173-194.

[12] Chaires, J.B.; Fox, K.R.; Herrera, J.E.; Britt, M.; Waring, M.J. Site and sequence specificity of the daunomycin-DNA interaction. *Biochemistry*, 1987, *26,* 8227-8236.

[13] Boger, D.L.; Chen, J.H.; Saionz, K.W. (-)-Sandramycin: Total synthesis and characterization of DNA binding properties. *J. Am. Chem. Soc.* 1996, *118,* 1629-1644.

[14] Bailly, C.; Suh, D.; Waring, M.J.; Chaires, J.B. Binding of daunomycin to diaminopurine and/or inosine-substituted DNA. *Biochemistry* 1998, *37,* 1033-1045.

[15] Chen, Y.H.; Lown, J.W. A new DNA minor groove binding motif: cross-linked lexitropsins. *J. Am. Chem. Soc.* 1994, *116,* 6995-7005.

[16] Walker, W.L.; Landaw, E.M.; Dickerson, R.E.; Goodsell, D.S. *Proc. Natl. Acad. Sci. U.S.A.* 1997, *94,* 5634-5639.

[17] Satz, A.L.; Bruice, T.C. Recognition in the minor groove of double-stranded DNA by microgonotropens. *Acc. Chem. Res.* 2002, *35,* 86-95.

[18] Bailly, C.; O'hUigin, C.; Houssin, R.; Colson, P.; Houssier, C.; Rivalle, C.; Bisagni, E.; Hénichart, J.P.; Waring, M.J. DNA binding properties of a distamycin-ellipticine hybrid molecule. *Mol. Pharmacol.*. 1992, *41,* 845-855.

[19] Mrksich, M.; Parks, M.E.; Dervan, P.B. Hairpin peptide motif. A new class of oligopeptides for sequence-specific recognition in the minor groove of double-helical DNA. *J. Am. Chem. Soc.* 1994, *116,* 7983-7988; Weyermann, P.; Dervan, P.B. Recognition of ten base pairs of DNA by head-to-head hairpin dimers. *J. Am. Chem. Soc.* 2002, *124,* 6872-6878 and references cited therein.

[20] Yang, C.-H.; Chou, P.-J.; Luo, Z.-L.; Chou, I.-C.; Chang, J.-C.; Cheng, C.-C.; Martin, C.R.H.; Waring, M.J.; Sheh, L. Preferential binding to DNA sequences of peptides related to a novel XPRK motif. *Bioorg. Med. Chem*. 2003, *11,* 3279-3288.

[21] Chang, J.-C.; Yang, C.-H.; Chou, P.-J.; Yang, W.-H.; Chou, I.-C.; Lu, C.-T.; Lin, P.-H.; Hou, R.C.-W.; Jeng, K.-C.G.; Cheng, C.-C.; Sheh, L. DNA sequence-specific recognition of peptides incorporating the HPRK and ployamide motifs. *Bioorg. Med. Chem*. 2004, *12,* 53-61.

[22] Yang, C.-H.; Chen, W.F.; Jong, M.-C.; Jong, B.-J.; Chang, J.-C.; Waring, M.J.; Ma, L.; Sheh, L. Semiquinone footprinting. *J. Am. Chem. Soc.* 2004, *126,* 8104-8105.

[23] Yang, C.-H.; Jeng, K.C.G.; Yang, W.H.; Chen, Y.L.; Hung, C.C.; Lin, J.W.; Chen, S.T.; Richardson, S.; Martin, C.R.H.; Waring, M.J.; Sheh, L. Unusually strong positive cooperativity in binding of peptides to latent membrane protein-1 DNA fragments of the Epstein-Barr viral gene. *ChemBioChem* 2006, *7,* 1187-1196.

[24] Chen, Y.C.; Huang, J.T.B.; Jeng, K.C.G.; Yang, R.C.K.; Liao, M.K.; Chen, C.S.; Chien, W.J.; Wey, M.T.; Kan, L.S.; Sheh, L. Determination of Allosteric Effects and Interstrand Bidentate Interactions in DNA-peptide Molecular Recognition *J. Chin. Chem. Soc.* 2010, *57*, 266-274.

[25] Waring, M.J.; Wakelin, L.P.G. Echinomycin: a bifunctional intercalating antibiotic. *Nature* 1974, *252,* 653-657.

[26] Wakelin, L.P.G.; Waring, M.J. The binding of echinomycin to DNA. *Biochem. J.* 1976, *157,* 721-740.

[27] Waring, M.J. in molecular aspects of anticancer drug-DNA interactions (Neidle, S. and Waring, M.J. eds) 1993, Vol. 1, 213-242, MacMillan, London and references quoted therein.

[28] Kong, D.; Park, E.J.; Stephen, A.G.; Calvani, M.; Cardellina, J.H.; Monks, A.; Fisher, R.J.; Shoemaker, R.H.; Melillo, G. Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity. *Cancer Res.* 2005, *65,* 9047-9055.

[29] Abrescia, N.G.A.; Malinina, L.; Subirana, J.A. Stacking interaction of guanine with netropsin in the minor groove of $d(CGTATATACG)_2$. *J. Mol. Biol.* 1999, *294,* 657-666.

[30] Bailly, C.; Colson, P.; Houssier, C., Houssin, R., Mrani, D.; Gosselin, G.; Imbach, J.L.; Waring, M.J.; Lown, J.W.; Hénichart, J.P. Binding properties and DNA sequence-specific recognition of two bithiazole-linked netropsin hybrid molecules. *Biochemistry* 1992, *31,* 8349-8362.

[31] Chen, X.; Ramakrishnan, T.R.; RAo, S.T.; Sundaralingam, M. Binding of two distamycin A molecules in the minor groove of an alternating B-DNA duplex. *Nature Struct. Biol.* 1994, *1,* 169-175.

[32] Dervan, P.B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem*. 2001, *9,* 2215-2235 and references quoted therein

[33] Breslauer, K.J.; Remeta, D.P.; Chou, W.Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J.G.; Marky, L.A. Enthalpy-entropy compensations in drug-DNA binding studies. *Proc. Natl. Acad. Sci. U.S.A.* 1987, *84,* 8922-8926.

[34] Hutchins, R.A.; Crenshaw, J.M.; Graves, D.E.; Denny, W.A. Influence of substituent modifications on DNA binding energetics of acridine-based anticancer agents. *Biochemistry* 2003, *42,* 13754-13761.

[35] Chaires, J.B.; Satyanarayana, S.; Suh, D.; Fokt, I.; Przewloka, T.; Priebe, W. Parsing the free energy of anthracycline antibiotic binding to DNA. *Biochemistry* 1996, *35,* 2047-2053.

[36] Qu, X.; Ren, J.; Riccelli, P.V.; Benight, A.S.; Chaires, J.B. Enthalpy/entropy compensation; influence of DNA flanking sequence on the binding of 7-amino actinomycin D to its primary binding site in short DNA duplexes. *Biochemistry* 2003, *42,* 11960-11967.

[37] Munde, J.; Lee. M.; Neidle, S.; Arafa, R.; Boykin, D.W.; Liu, Y.; Bailly, C.; Wilson, W.D. Induced fit conformational changes of a reversed amidine heterocycle: optimized interactions In a DNA minor groove complex. *J. Am. Chem. Soc.* 2007, *129,* 5688-5698 references cited therein.

[38] Davanloo, P.; Crothers, D.M. Kinetic studies of drug-dinucleotide complexes. *Biochemistry* 1976, *15,* 5299-5305; Davanloo, P.; Crothers, D.M. Phase partition studies of actinomycin-nucleotide complexes. *Biochemistry* 1976, *15,* 4433-4438.

[39] Fox, K.R.; Waring, M.J. Stopped-flow kinetic studies on the interaction between echinomycin and DNA. *Biochemistry* 1994, *23,* 2627-2633.

[40] Chaires, J.B.; Dattagupta, N.; Crothers, D.M. Kinetics of daunomycin-DNA interactions. *Biochemistry* 1985, *24,* 260-267.

[41] Baliga, R.; Baird, E.E.; Herman, D.M.; Melander, C.; Dervan. P.B.; Crothers, D.M. Kinetic consequences of covalent linkage of DNA binding polyamides. *Biochemistry* 2001, *40,* 3-8.

[42] Fletcher, M.C.; Fox, K.R. Dissociation kinetics of echinomycin from CpG binding sites in different sequence environments. *Biochemistry* 1996, *35,* 1064-1075.

[43] Churchill, M.E.A.; Suzuki, M. SPKK motifs prefer to bind to DNA at A/T –rich sites. *EMBO J.* 1989, *8,* 4189-4195.

[44] Suzuki, M. The heptad repeat in the largest subunit of RNA polymerase II binds by intercalating into DNA. *Nature* 1990, *344,* 562-565.

[45] Monod, J.; Wyman, J.; Changeux, J.P. On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* 1965, *12,* 88-118.

[46] Koshland, D.E. Jr. and Hamadani, K. Proteomics and models for enzyme cooperativity. *J. Biol. Chem.* 2002, *277,* 46481-46844 and references cited therein.

[47] Shiozawa, H.; Chia, B.C.S.; Davies, N.L.; Zerella, R.; Williams, D.H. Cooperative binding interactions of glycopeptide antibiotics. *J. Am. Chem. Soc.* 2002, *124,* 3914-3919.

[48] Williams, D.H.; Davies, N.L.; Zerella, R.; Bardsley, B. Noncovalent interactions: defining cooperativity. Ligand binding aided by reduced dynamic behavior of receptors. Binding of bacterial cell wall analogues to Ristocentin A. *J. Am. Chem. Soc.* 2004, *126,* 2042-2049.

[49] Dattagupta, N.; Hogan, M.; Crothers, D.M. Interaction of netropsin and distamycin with deoxyribonucleic acid: electric dichroism study. *Biochemistry* 1980, *19,* 5998-6005.

[50] Graves, D.E.; Krugh, T.R. Adriamycin and daunorubicin bind in a cooperative manner to deoxyribonucleic acid. *Biochemistry* 1983, *22,* 3941-3947.

[51] Fagan, P.; Wemmer, D.E. Cooperative binding of distamycin-A to DNA in the 2:1 mode. *J. Am. Chem. Soc.* 1992, *114,* 1080-1081.

[52] Bailly, C.; Hamy, F.; Waring, M.J. Cooperativity in the binding of echinomycin to DNA fragments containing closely spaced CpG sites. *Biochemistry* 1996, *35,* 1150-1161.

[53] Morii, T.; Yamane, J.; Aizawa, Y.; Makino, K.; Sugiura, Y. Cooperative oligomerization enhances sequence-selective DNA binding by a short peptide. *J. Am. Chem. Soc.* 1996, *118,* 10011-10017.

[54] Qu, S.; Trent, J.O.; Fokt, I.; Priebe, W.; Chaires, J.B. Allosteric, chiral-selective drug binding to DNA. *Proc. Natl. Acad. Sci. U.S.A.* 2000, *97,* 12032-12037.

[55] Harris, S.A.; Gavathiotis, E.; Searle, M.S.; Orozco, M.; Laughton, C.A. Cooperativity in drug-DNA recognition: a molecular dynamics study. *J. Am. Chem. Soc.* 2001, *123,* 12658-12663.

[56] Cooper. A.; Dryden, D.T.F. Allostery without conformational change. A plausible model. *Eur. Biophys. J.*, 1984, *11,* 103-109.

[57] Fechter, E.J.; Dervan, P.B. Allosteric inhibition of protein-DNA complexes by polyamide-intercalator conjugates. *J. Am. Chem. Soc.* 2003, *125,* 8476-8485.

[58] Yang, R.C.K. et al. paper to be submitted.

[59] Russell, R.B. and Aloy, P. Targeting and tinkering with interaction networks. *Nat. Chem. Biol.* 2008, *4,* 666-673 and references cited therein.

[60] Yu, H. et al. High-quality binary protein interaction map of the yeast interactome network. *Science,* 2008, *322,* 104-110.

[61] Vetter, D. Chemical microarrays, fragment diversity, label-free imaging by plasmon resonance-a chemical genomics approach. *J. Cell. Biochem. Suppl.* 2002, *39,* 79-84.

[62] Berger, M.F.; Bulyk, M.L. protein binding microarrays for rapid, high throughput characterization of the sequence-specificities of DNA binding proteins. *Methods Mol. Biol.* 2006, *338,* 245-260.

Bioorganic & Medicinal Chemistry 18 (2010) 366–376

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Detection of multiple network-based allosteric interactions between peptides and arrays of DNA binding sites

Karen K. L. Kao [a], Jonathan C. T. Huang [a], Chi-Kai Yang [a], Kee-Ching G. Jeng [b], Jung-Cheng Chang [a], Wen-Chen Yao [a], S. C. Hsien [a], Michael J. Waring [c], Ming-Hua Chen [d], Lin Ma [d], Leung Sheh [a,*]

[a] Department of Chemistry and Life Science Center, Tunghai Christian University, Taichung 407, Taiwan, ROC
[b] Department of Medical Research, Taichung Veterans General Hospital, Taichung 405, Taiwan, ROC
[c] Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, England, United Kingdom
[d] Shaw College, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

ARTICLE INFO

Article history:
Received 17 September 2009
Revised 22 October 2009
Accepted 24 October 2009
Available online 31 October 2009

Keywords:
DNA cooperative networks
Footprinting
Circular dichroism
Molecular recognition
Peptides

ABSTRACT

Quantitative DNase I footprinting shows that three designed peptides containing *N*-methylpyrrole (Py) moieties display different types of network-based allosteric communication in binding to DNA: circuit type, incomplete-circuit type, and non-circuit type characterized by interstrand bidentate interactions. Positive cooperative binding of all three peptides to individual DNA binding sites is commonly observed. CD spectral characterization of the interaction between peptides and model undecanucleotide duplexes is consistent with the footprinting results and supports the allosteric model. This study provides insights relating to the interaction network nature of allostery in complex DNA–small molecule interactions.

1. Introduction

Current studies on sequence-selective binding of small ligands to DNA play an essential part in the advance towards understanding molecular aspects of the modulation of gene expression. This work is also important for the advancement of structural biochemistry research and drug design.[1–4] Extensive studies have been carried out with small ligands that are capable of sequence-specific recognition of DNA including drugs,[1–3] antitumor antibiotics,[4–7] netropsin/distamycin and its analogues,[8–13] and synthetic peptides.[14–17]

Previously we introduced a novel XPRK peptide motif as a basis for sequence-selective interaction studies.[14–17] The design of this motif stems from a naturally occurring SPXX motif[18,19] found in repeating sequences in histones, steroid hormone receptors, various segmentation gene products and some oncogene products. It was suggested that the SPXX motif assumes a β-turn stabilized by two hydrogen bonds, and that the side chains of the two basic residues engage in salt bridges with the DNA phosphate groups.[18,19] Further conjugation of two XPRK motifs with a tract of 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) afforded peptides with higher affinity toward DNA sequences incorporating three or four consecutive W (A or T) base pairs.

Extensive research spanning nearly fifty years has been devoted to the allosteric behavior of macromolecules such as proteins[21] because allostery is recognized as a central process in biological control governing biochemical efficiency and energy metabolism. By contrast, studies on allosteric aspects of DNA–ligand interactions have been much less popular than protein–ligand interactions. This prompted us to focus on molecular aspects of cooperative interactions between peptides and DNA. In our early sequence-specificity studies with synthetic molecules containing both the XPRK motif[14–17] as well as the polyamide motif we found a number of peptides that exhibit notable sequence-selectivity in binding to particular sites on the DNA duplex. Among these peptides we identified a nonapeptide His-Hyp-Arg-Lys-$(Py)_3$-Lys-Arg-NH_2 (PyHyp-9) and a dodecapeptide His-Hyp-Arg-Lys-$(Py)_4$-His-Hyp-Arg-Lys-NH_2 (PyHyp-12) that displayed significant cooperativity in binding to DNA. We also established that quantitative footprinting is the method of choice for investigating allosteric interactions between multiple binding sites on complementary DNA strands since it is difficult for most spectroscopic techniques to characterize multiple binding sites.

It is generally recognized that network-based interactions between multiple elements play a part in many biochemical and physiological processes in nature. Recently we proposed a net-

* Corresponding author. Tel.: +886 4 23590248; fax: +886 4 23590426.
E-mail address: Lsheh@thu.edu.tw (L. Sheh).

0968-0896/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2009.10.048

work-based DNA–peptide allosteric model[17,20] that features intercommunicating binding sites in fragments of the latent membrane protein-1 gene from a pathogenic Epstein–Barr virus variant derived from nasopharyngeal carcinoma. In the present study we further explore whether such network-based allosteric behavior can be detected in common DNA motifs containing consecutive A's and T's. To this end we compare the physicochemical binding parameters of the newly designed peptides PyHyp-9 and PyHyp-12 with that of the parent peptide His-Pro-Arg-Lys-$(Py)_4$-His-Pro-Arg-Lys-NH_2 (PyPro-12) using quantitative DNase I footprinting. Cooperativity is assessed using the Hill equation which is well established for that purpose.[17,23–25] To furnish insights into possible mechanisms based on DNA conformational changes circular dichroism spectroscopic results are also reported.

2. Results and discussion

Since the initial objective of this study was to investigate the molecular basis underlying cooperativity between multiple peptide recognition sites, a complementary pair of 5′-^{32}P-labeled pBR322 fragments[15] comprising a 158-mer duplex (upper strand 5′-^{32}P-labeled) and a 135-mer duplex (lower strand 5′-^{32}P-labeled) were employed as substrates. Previous footprinting studies showed that the XPRK series of peptides bind preferentially to four sites on the upper and lower strands where the complementary sequences d(AAAA)–d(TTTT) and d(AAA)–d(TTT) occur around positions 108–111 and 83–85, respectively. The design of the new peptides PyHyp-9 and PyHyp-12 stems from the parent peptide PyPro-12 (previously named PyH-12).[15] In peptides PyHyp-9 and PyHyp-12 proline is replaced by hydroxyproline, furnishing extra hydroxyl groups for additional hydrogen bonding. The nonapeptide PyHyp-9 represents a new design in which the number of Py residues is reduced to three and the *C*-terminal tetrapeptide portion (-His-Hyp-Lys-Arg-NH_2 in peptide PyHyp-12) is truncated to a dipeptide (-Lys-Arg-NH_2).

Peptide PyHyp-9 displays a strong sequence preference for consecutive arrays of A's and T's in DNA (Fig. 1). At present the nonapeptide PyHyp-9 is the smallest peptide containing enough amino acid residues to evince satisfactory footprints in the XPRK series. Quantitative footprinting by this and other peptides (Figs. 1–3, 4A–C) on the complementary strands provides valuable information about possible interstrand bidentate interactions[15,22] as well as intrastrand H-bond interactions[22] involving the side chains and amino protons of the peptides bonding to bases in the complementary DNA strands. These specific H-bond interactions between the peptides and DNA bases are well-recognized mediators of sequence recognition by ligands including proteins as well as small molecules[15,22] Interstrand bidentate interactions can be assigned from the differential cleavage plots at nucleotide positions where significant coincident DNase I blockages are observed (Fig. 4A–C).

On the upper strand, peptide PyHyp-9 binds strongly to two sites (Figs. 1 and 4A): position U107–113, corresponding to the sequence 5′-GAAAATA-3′ which shows marked positive cooperativity (n_H = 2.2); and position U84–89 comprising the sequence 5′-AA-TACC-3′, also associated with positive cooperativity (n_H = 2.2). On the complementary lower strand, PyHyp-9 footprints at two sites:

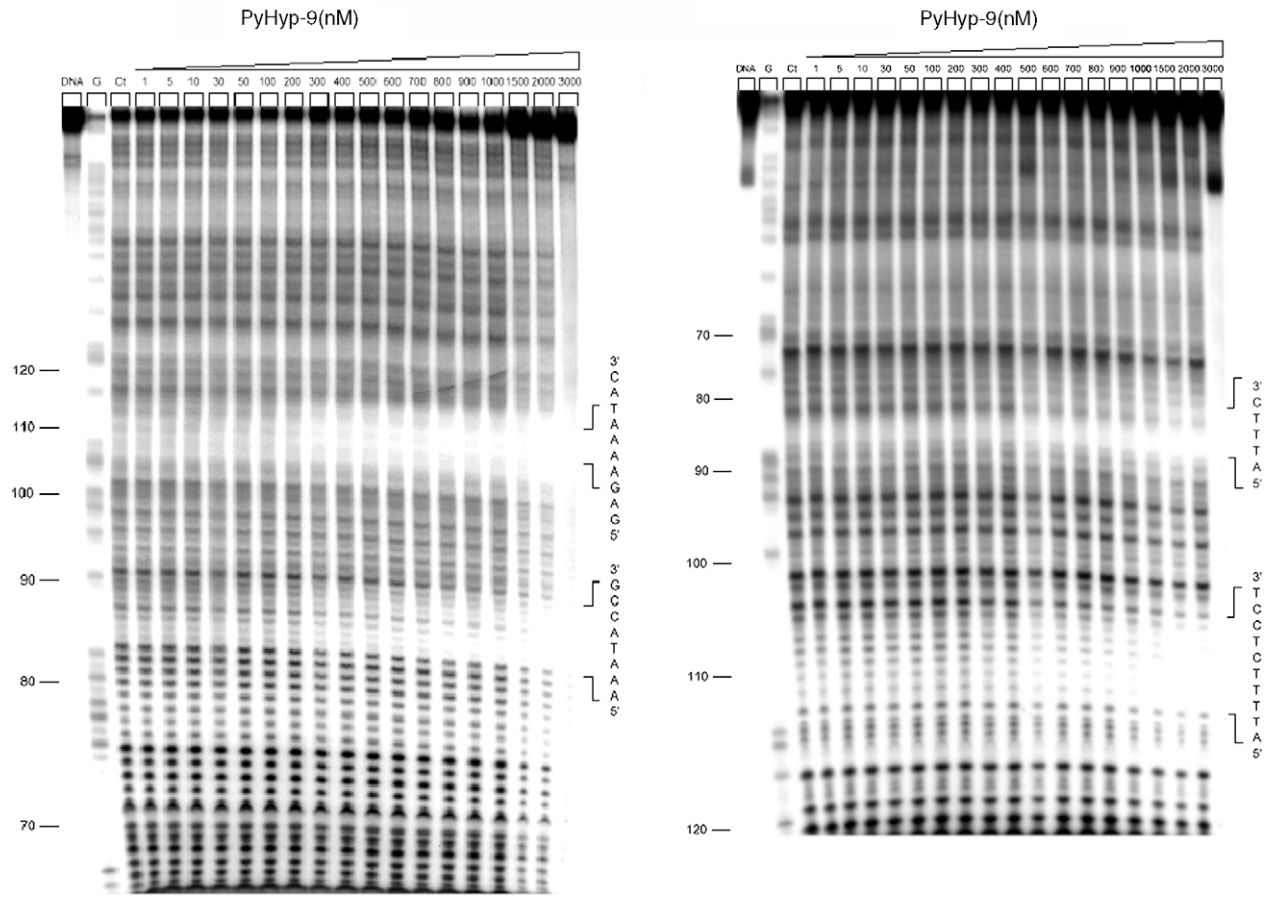


Figure 1. Autoradiograph showing DNase I footprinting of peptide PyHyp-9 on DNA duplexes labeled with [γ-^{32}P]ATP on the 5′ end: 5′-^{32}P-labeled 158-mer upper strand, left panel; and 5′-^{32}P-labeled 135-mer lower strand, right panel. PyHyp-9 was equilibrated with the DNA in 5 mM sodium cacodylate buffer, pH 6.5 at 37 °C for 60 min before nucleolytic cleavage. G represents a Maxam–Gilbert guanine sequencing track and Ct shows a DNase I digestion control lane.

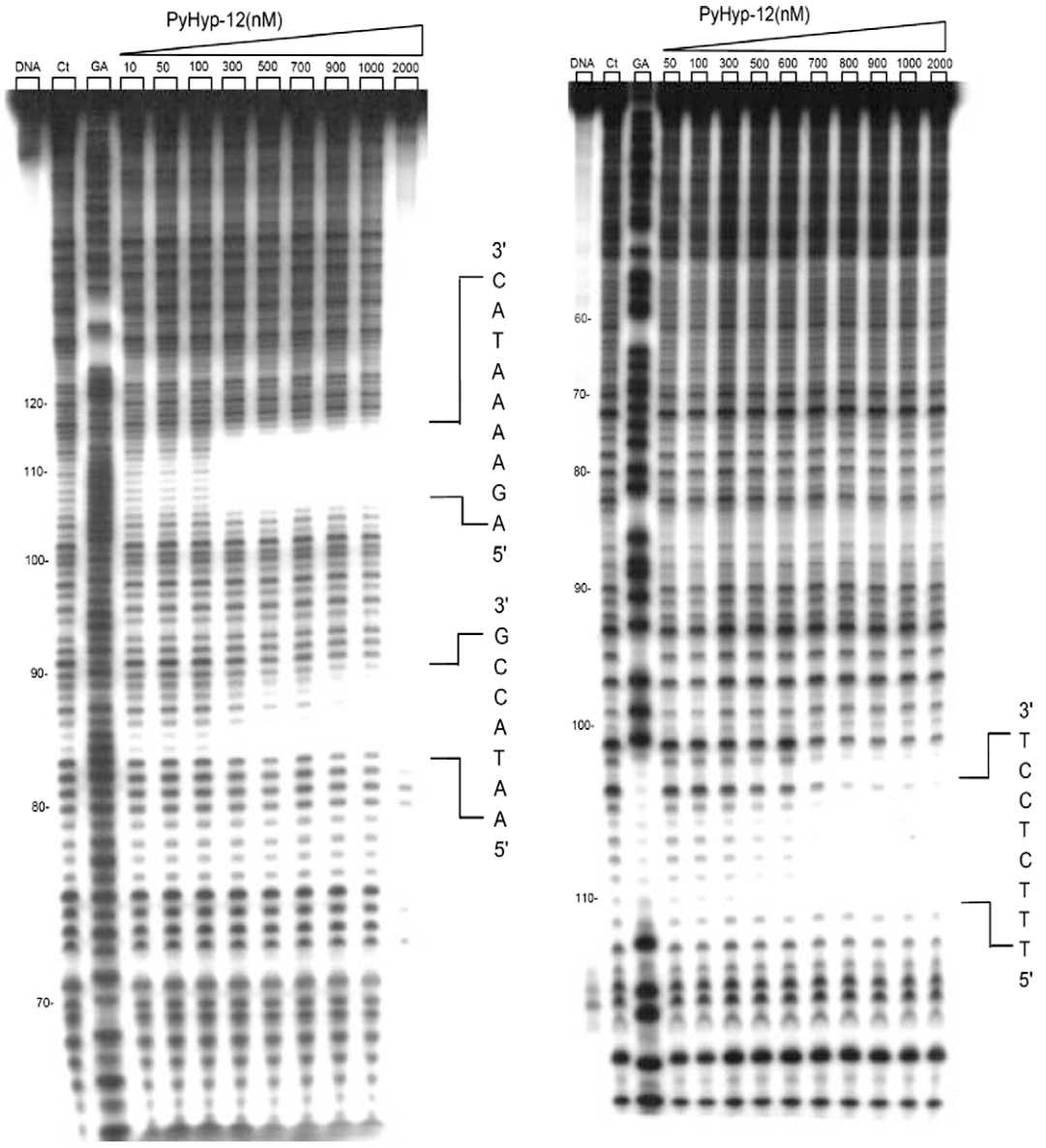


Figure 2. DNase I footprinting of peptide PyHyp-12 on DNA duplexes labeled with [γ-^{32}P]ATP on the 5′ end: 5′-^{32}P-labeled 158-mer upper strand, left panel; and 5′-^{32}P-labeled 135-mer lower strand, right panel. Reaction conditions as in the legend to Figure 1.

position L111–103, corresponding to the sequence 5′-TTTTCTCCT-3′ (n_H = 2.0), and position L84–83, comprising the sequence 5′-TT-3′ (n_H also 2.0). Examination of the differential plot thus allows us to assign interstrand bidentate interactions around positions 84 and 107–111. It is noteworthy that the nonapeptide displays significant positive cooperativity at all four binding sites on both strands.

The DNase I footprinting profile of peptide PyHyp-12 (Figs. 2 and 4B) appears to be quite different from that of the parent peptide PyPro-12. On the upper strand, peptide PyHyp-12 produces a broad DNase I blockage site around position U106–114 corresponding to the sequence 5′-AGAAAATAC-3′, displaying weak positive cooperativity (n_H = 1.4). Another blockage site occurs at position U84–88, comprising the sequence 5-AATAC-3′ displaying similarly weak positive cooperativity (n_H = 1.5). On the lower strand, PyHyp-12 binds preferentially to a broad site around position L110–103, corresponding to the sequence 5′-TTTCTCCT-3′ showing the same weak positive cooperativity (n_H = 1.4), but the binding site around position L84–83 that appears with peptide PyHy-9 is completely missing. In this case the differential cleavage plots indicate that for PyHyp-12 interstrand bidentate interactions can only be assigned around position 106–110. It is envisaged that these interstrand hydrogen bond interactions are involved in relaying conformational changes between local DNA sites, thereby facilitating binding of further peptide molecules across the complementary strands.

We repeated footprinting experiments on the parent peptide PyPro-12[15] using a wider concentration range (Figs. 3 and 4C). On the upper strand PyPro-12 binds quite well around position U84–90, comprising the sequence 5′-AATACCG-3′, showing some negative cooperativity (n_H = 0.7). Another broad blockage site occurs around position U106–116, comprising the sequence 5′-AGAAAATACCG-3′, showing weak positive cooperativity (n_H = 1.2). On the lower strand, corresponding blockage sites also appear around positions L85–83 (n_H = 1.1) and L110–105 (n_H = 1.5) that indicate binding of peptide molecules to sites on the opposite complementary strand, also manifesting modest positive cooperativity (Fig. 4C). So for this peptide interstrand bidentate interactions[15,22] can be assigned around position 84–85 as well as 106–110. The

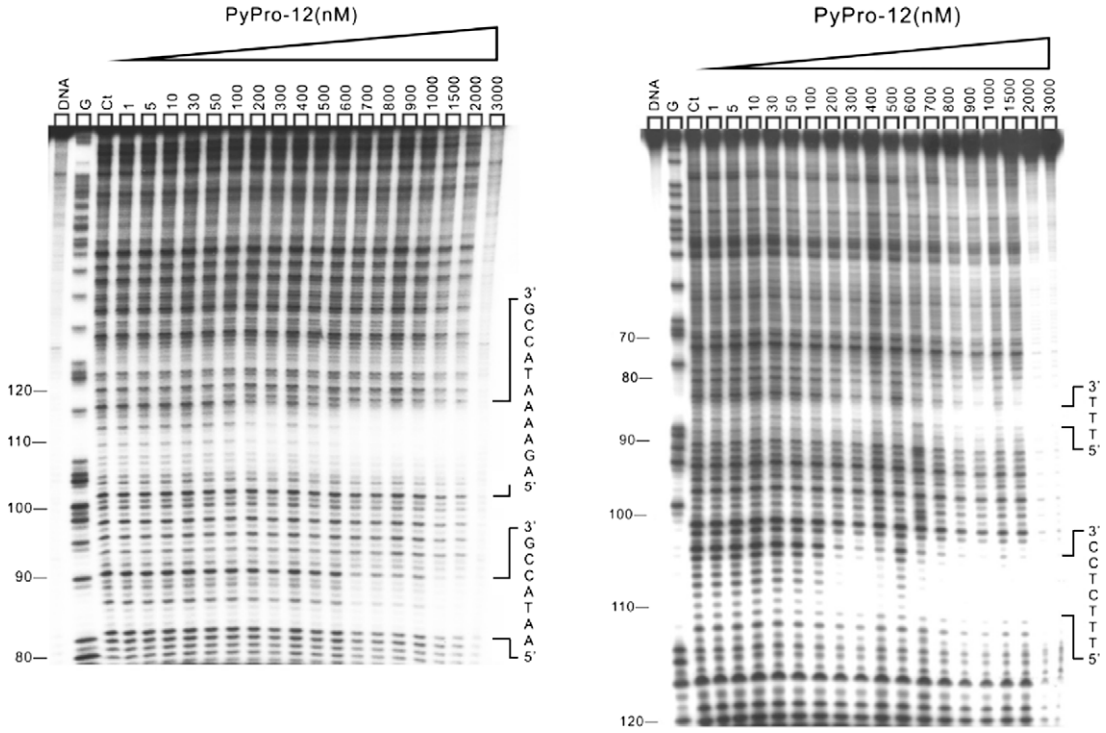


Figure 3. DNase I footprinting of peptide PyPro-12 on DNA duplexes labeled with [γ-^{32}P]ATP on the 5′ end: 5′-^{32}P-labeled 158-mer upper strand, left panel; and 5′-^{32}P-labeled 135-mer lower strand, right panel. Reaction conditions as in Figure 1.

apparent binding constant K_a of the three peptides on the four major binding sites of the pBR322 fragments lie rather even in the range of 2.2–3.8 $\times$ 10^6 M^{-1}, with the exception that peptide PyPro-12 has a rather high K_a (6.7 $\times$ 10^7 M^{-1} on position U80–84) (Table 1). Similarly, the apparent binding constant of peptides PyHyp-12 and PyHyp-9 on binding to the S-81 DNA also lie rather steady in the range of 1.4–2.5 $\times$ 10^6 M^{-1} (Table 2).

In view of the broad DNase I blockage sites seen with all three peptides, we envisage that two peptide molecules may be capable of binding in a dimeric head-to-head or head-to-tail fashion to some or all of these binding sites (Fig. 5). Several modes of binding of peptide molecules to complementary or adjacent binding sites are possible that might lead to cooperative binding. Interaction of a peptide molecule in a sequential manner with one sub-site (inside a binding site) on one DNA strand could facilitate the binding of a molecule to the other sub-site on the same strand. On the other hand, binding of a peptide molecule to one site might induce local DNA conformational change(s) mediated by interstrand bidentate interactions and facilitate sequential or concerted binding of further peptide molecules to an opposite site on the complementary strand. Of course, binding of peptide molecules to one site could invoke conformational change(s) and facilitate or hinder peptide binding to a neighboring site on the same strand. Thus, cooperative (or anti-cooperative) binding distributed between an array of binding sites may establish a network of cooperativity interconnecting several allosteric sites (Fig. 5).

To assess the nature and magnitude of allosteric communication between the d(AAAA)–d(TTTT) and d(AAA)–d(TTT) binding loci in the pBR322 fragments we designed a new 81-mer duplex (S-81) containing a single d(AAAA)–d(TTTT) sequence. The apparent size (number of nucleotides) of each binding locus was found to vary for individual peptides (Table 2). Quantitative footprinting experiments were carried out for peptides PyHyp-12 and PyHyp-9 binding to the 5′-^{32}P-labeled upper and lower strand duplexes (Supplementary data). Strikingly, the n_H values for peptides PyHyp-12 and PyHyp-9 binding to the 81-mer duplexes containing a single binding locus (Table 2) are significantly higher than those of the 158-mer and 135-mer duplexes containing both binding loci (Table 1). This suggests the existence of some negative cooperative effect associated with binding to the d(AAA)–d(TTT) and d(AAAA)–d(TTTT) positions, lowering the positive cooperativity between adjacent and complementary sub-sites, and resulting in smaller observed values of n_H at both sites, mediated over an intervening sequence of 12–14 base pairs.

Based on the quantitative footprinting behavior of the three peptides, summarized in Table 1, we have proposed network-based models in an effort to interpret the complex communication between DNA–peptide allosteric binding sites (Fig. 5). The 158-mer and 135-mer fragments used here contain two peptide binding loci: around position 83–89, comprising the d(AAA)–d(TTT) sequence, and position 103–116 comprising the d(AAAA)–d(TTTT) sequence. Each locus seems to afford two complementary sites that consist of two to four sub-sites, depending on individual peptides. In the model, conformational changes induced at one position affect binding to a neighboring site, and interstrand bidentate interactions are envisaged as playing an important role in relaying conformational changes between sites on the complementary strands, facilitating the binding of further peptide molecules. The different models for PyHyp-9, PyHyp-12 and PyPro-12 primarily reflect differences in the relay of conformational change between opposite sites on complementary strands via interstrand bidentate interactions. It is apparent that peptides PyHyp-9 and PyPro-12 form a circuit type of a communication: the allosteric communication between binding sites is complete. On the other hand, the communication in the DNA binding of peptide PyHyp-12 appears to be discontinuous and is referred to as an incomplete-circuit type. Footprinting experiments on the 81-mer duplex (S-81) containing a single d(AAAA)–d(TTTT) sequence (Table 2) suggest that the allosteric relay of peptide binding to adjacent sites between the d(AAA)–d(TTT) and d(AAAA)–d(TTTT) positions spanning over

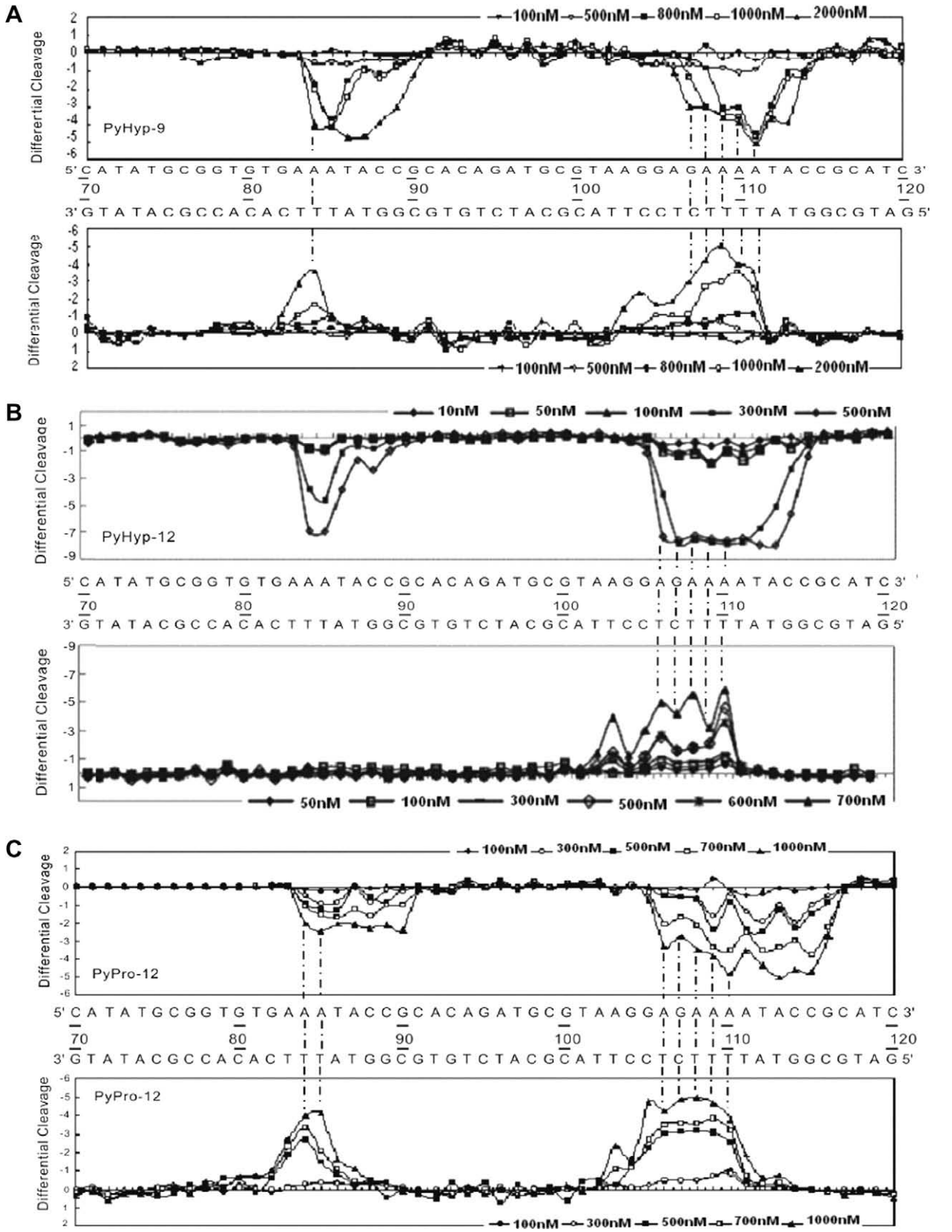


Figure 4. Differential cleavage plots comparing the susceptibility of DNA fragments to DNase I cleavage after incubation with each peptide in cacodylate buffer at room temperature for 60 min. In each panel A, B, C the upper traces represent the differential cleavage plot for a given peptide bound to the 5′-^{32}P-labeled upper strand (158-mer) DNA fragment; the lower traces represent the corresponding plots for each peptide bound to the 5′-^{32}P-labeled lower strand (135-mer) DNA fragment. The vertical dotted lines between DNA bases represent assignment of interstrand bidentate interactions where significant coincident H-bonding interactions occur between complementary bases in both strands.

Table 1
Physicochemical parameters for binding of peptides to recognition sites in pBR322 fragments: complementary 5′-^{32}P-labeled upper strand (158-mer) and lower strand (135-mer) DNA duplexes at 37 °C

Peptide	Binding site position	Recognition sequence	K_a (M^{-1})	n_H	Interstrand bidentate interactions (position)
PyHyp-9	U84–89	5′-AATACC-3′	2.5×10^6	2.2	84
	U107–113	5′-GAAAATA-3′	3.8×10^6	2.2	107–111
	L84–83	5′-TT-3′	2.7×10^6	2.0	
	L111–103	5′-TTTTCTCCT-3′	2.2×10^6	2.0	
PyHyp-12	U84–88	5′-AATAC-3′	2.5×10^6	1.5	106–110
	U106–114	5′-AGAAAATAC-3′	3.4×10^6	1.4	
	L110–103	5′-TTTCTCCT-3′	3.4×10^6	1.4	
PyPro-12	U80–94	5′-AATACCG-3′	6.7×10^7	0.7	84–85
	U106–116	5′-AGAAAATACCG-3′	3.2×10^6	1.2	106–110
	L85–83	5′-TTT-3′	3.7×10^6	1.1	
	L110–105	5′-TTTCTCC-3′	3.8×10^6	1.5	

K_a and n_H are the apparent association constant and Hill coefficient determined from concentration-dependent DNase I footprinting studies, respectively. The binding site positions on the upper and lower strands are abbreviated as U and L, respectively. Interstrand bidentate interactions are assigned where there are coincident effects on complementary bases in both strands.

Table 2
Physicochemical parameters for binding of peptides to the S-81 DNA: complementary 5′-^{32}P-labeled upper strand (81-mer) and lower strand (81-mer) duplexes at 37 °C

Peptide	Binding site position	Recognition sequence	K_a	n_H	Interstrand bidentate interactions (position)
PyHyp-9	U44–48	5′-AAAAT-3′	2.1×10^6	3.9	44–47
	L47–41	5′-TTTTCTC-3′	1.9×10^6	3.2	
PyHyp-12	U44–48	5′-AAAAT-3′	1.4×10^6	2.3	44–46
	L46–42	5′-TTTCT-3′	2.5×10^6	2.6	

Details as in the legend to Table 1.

an intervening sequence of 12–14 base pairs is affected by some negative cooperative effect (Fig. 5).

Seeking further evidence for conformational changes accompanying the molecular recognition process we resorted to circular dichroism spectroscopy (CD) as a means of probing changes in the helical structure associated with peptide–DNA binding. Two 11-mer deoxyribonucleotide duplexes were synthesized, d(GGAGA AAATAC)–d(GTATTTTCTCC) (4A–4T) and d(GTGAAATACCG)–d(CG GTATTTCAC) (3A–3T), corresponding to the recognition sites of the pBR322 fragments at positions 104–114 and 80–90 separately. With both duplexes a negative CD band around 253 nm and a positive band around 274 nm were observed (Fig. 6A). On increasing the concentration of peptide PyHyp-9 added to duplex 4A–4T, the 274 nm band is red-shifted to about 278 nm and a near-isoelliptic point appears suggesting a predominantly two-component binding process (Fig. 6B). This positive CD band extends as a shallow broad dose-dependent positive enhancement to 350 nm. In the difference spectrum (Fig. 6C) it can be seen that binding of peptide PyHyp-9 to the duplex results in dose-dependent CD intensity enhancement of two neighboring negative bands centering around 243 nm and 268 nm. Another curious feature of Figure 6C is a blue-shifted positive CD band around 273 nm induced by PyHyp-9 at low concentration (0.2 μM).

Addition of PyHyp-9 to the DNA duplex 3A–3T also induces a red-shift of the positive band at 274 nm to 281 nm with quite a good isoelliptic point at 283 nm (Fig. 6D). Again this positive band extends as a shallow broad band to about 350 nm. In the difference spectra two negative bands around 248 nm and 267 nm are observed (Fig. 6E), and at low concentrations of PyHyp-9 (0.2 μM and 1.0 μM), the induced positive CD band appears blue-shifted to about 282 nm. Odd spectral changes observed at low peptide concentrations are not entirely unexpected given the anomalous situation that might prevail when small numbers of ligand molecules are bound, probably largely to isolated sites, before propagated effects mediated via early conformational changes become predominant.

By contrast, titration of peptide PyHyp-12 versus the 4A–4T duplex produces a dose-dependent negative CD enhancement band around 250 nm as well as two strong positive CD enhancement bands around 283 nm and 325 nm (Fig. 7A). In the difference spectra two negative bands are seen around 243 nm and 268 nm as well as positive bands around 288 nm and 322 nm (Fig. 7B). A fairly consistent isoelliptic point also occurs near 280 nm, suggesting mainly a two-component binding process. Titration of PyHyp-12 versus the 3A–3T duplex (Fig. 7C) induces similar CD spectral changes to those observed with the 4A–4T duplex. In the difference spectra, there are strong and minor negative bands around 266 and 249 nm, while two strong positive CD bands appear around 282 and 325 nm (Fig. 7D).

For peptide PyHyp-9, the distinct separation of two unusual neighboring negative CD bands around 243 nm and 268 nm as well as the strong red-shifted positive bands indicate that significant local conformational changes are indeed induced in the duplex by the binding of the peptide, most likely resulting in the positively cooperative binding observed. Moreover, in the difference spectra (Figs. 6C and E), it is apparent that the enhancement of the CD bands is correlated with peptide–DNA binding stoichiometry. Accordingly we propose a three-step binding process (Fig. 8). At [peptide]/[duplex] ratios of 0.5–2, one molecule of PyHyp-9 or PyHyp-12 binds to the d(AAAA)–d(TTTT) or d(AAA)–d(TTT) locus, most probably in the minor groove. However, at [peptide]/[duplex] ratios above 1 and up to about 4, there is often a steeper variation in ellipticity consistent with strong positive cooperativity arising from binding of the first molecule facilitating the binding of a second molecule to form a 2:1 complex. At [peptide]/[duplex] ratios above 4.0, the progressive but smaller increase in $\Delta\theta$ suggests that more peptide molecules begin to bind non-selectively, perhaps occupying the wide major groove which can accommodate more peptide molecules than the minor groove. This idea is consistent with the footprinting studies which clearly indicate that peptide concentrations above 2 μM can result in non-sequence selective binding to DNA (Figs. 1 and 2).

There are signs that the peptides may bind in dimeric head-to-head or head-to-tail fashion to the d(AAAA)–d(TTTT)-containing 11-mer duplex with positive cooperativity. For peptide PyHyp-12, in addition to a positive CD band around 280 nm the appearance of a stronger positive band centered around 322 nm (Fig. 7) is noteworthy and the spectral position is similar to that seen with distamycin binding to an AAGTT-containing decameric duplex[26] as well as those of polyamide–DNA interactions (27). In the CD spectrum induced by

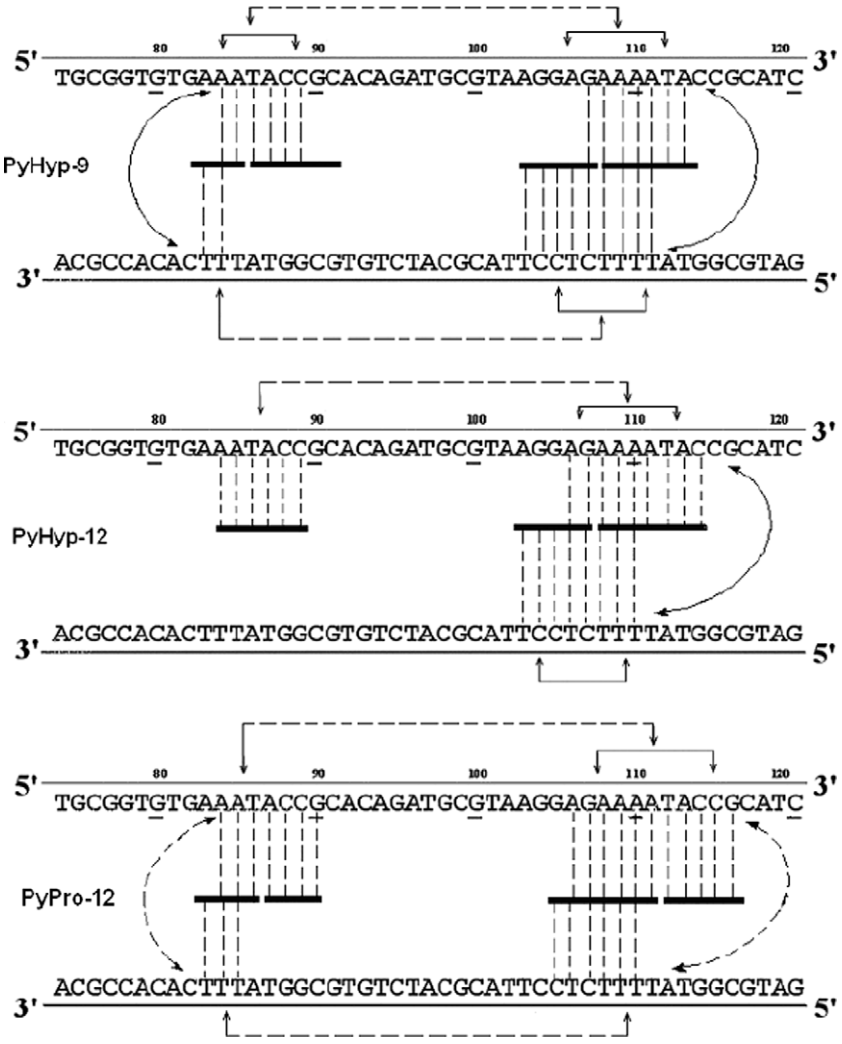


Figure 5. Proposed allosteric models for cooperative binding of peptides PyHyp-9, PyHyp-12 and PyPro-12 to 158-mer and 135-mer pBR322 fragments based on quantitative footprinting studies. The portion of the ligand binding to each DNA site/sub-site is represented by a thick horizontal line. Monodentate interactions and interstrand bidentate interactions are represented by vertical broken lines. The arrow lines represent communication of allosteric interaction between DNA binding sites. Arrow lines between the complementary strands represent possible influence of interstrand bidentate interactions mediated via DNA conformational changes. Broken arrow lines between neighboring binding loci in the top two panels are intended to represent negative cooperative communications.

DNA binding of a related peptide HR-12 (His-Pro-Arg-Lys)$_3$-NH_2, the positive CD band around 322 nm is absent (spectrum not shown). It has been demonstrated that molecules that bind to the minor groove typically elicit the appearance of a positive CD band.[27] Thus we reason that positive CD enhancement bands like those we detect around 322 nm are probably induced only by minor groove binding of ligands containing 4-amino-1-methylpyrrole-2-carboxylic acid (Py) residues. Compared to peptide PyHyp-12 with its four Py residues, peptide PyHyp-9 has only three, resulting in fewer H-bond interactions in the minor groove. Accordingly the observed change in ellipticity induced by PyHyp-9 is less than that of PyHyp-12 around 322 nm. Thus the CD spectral characterization of the binding of these peptides is consistent with the footprinting results and supports the allosteric model (Fig. 5) in which the recognition process involves cooperative binding of two peptide molecules to the d(AAAA)–d(TTTT)- or d(AAA)–d(TTT)-containing loci at appropriate [peptide]/[duplex] ratios. This accords with the finding that peptide molecules are hydrogen bonded to bases of complementary strands at these sites via interstrand bidentate interactions,[15,22] and is fully consistent with the interpretation that such interactions underlie the induction of local conformational changes in DNA as revealed by the CD experiments.

3. Conclusion

Quantitative DNase I footprinting studies have shown that network-based allostery featuring positive cooperativity appears to exist in the binding of some small peptides to adjacent sites or opposite complementary sites in DNA. However, this study also shows that allosteric communication between neighboring binding loci (separated from each other by 12–14 base pairs) can be negatively cooperative (Fig. 5, upper and middle panels). A special feature of the binding of peptides PyHyp-9 and PyPro-12 to duplex DNA molecules possessing multiple sites is the formation of a circuit type of allosteric effects within a network-based system. Unlike peptides PyHyp-9 and PyPro-12, for peptide PyHyp-12 the networks are referred to as an incomplete-circuit type. On the other hand, since peptides $(HPRK)_3NH_2$ [HR-12], and $(SPRK)_3NH_2$ [SP-12] bind to DNA fragments with GC-rich motifs lacking interstrand bidentate interactions, they form a non-circuit type of allosteric communication.[17] Thus, the present study and the previous one[17] enable us to hypothesize three different types of network-based allosteric communication in peptide–DNA molecular recognition.

The distinct positive CD bands induced by PyHyp-9 and PyHyp-12 around 288–323 nm indicate that significant local conforma-

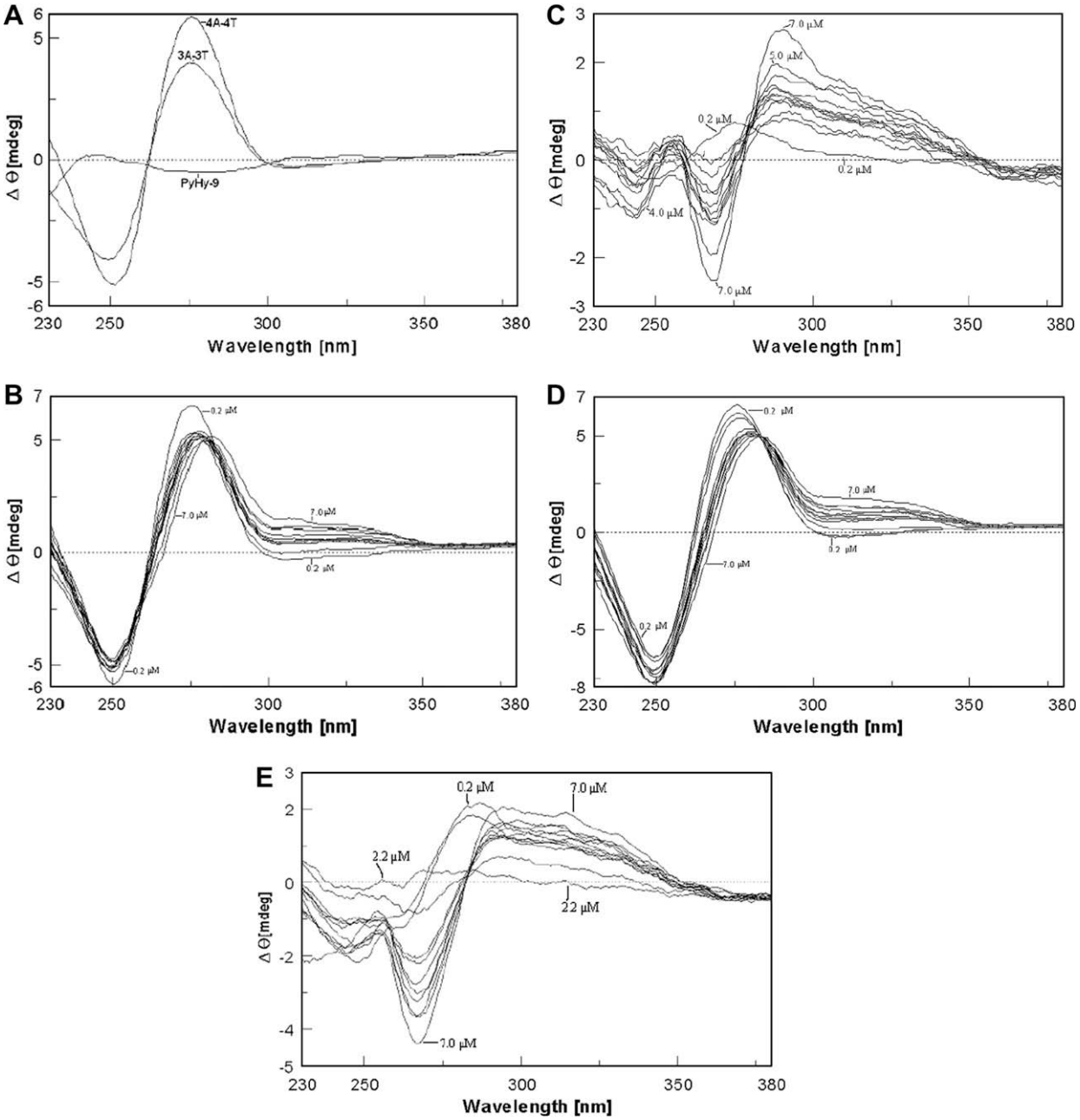


Figure 6. Panel A: CD spectra of DNA duplexes 4A–4T, 3A–3T alone and peptide PyHyp-9 alone. Panel B: titration of duplex 4A–4T versus PyHyp-9 at peptide concentrations of 0.2, 1.0. 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, 5.0, 7.0 μM. Panel C: corresponding CD difference spectra with the contribution of duplex 4A–4T and peptide subtracted. Panel D: titration of duplex 3A–3T versus PyHyp-9 over the same peptide concentration range. Panel E: corresponding CD difference spectra with the contributions of duplex 3A–3T and peptide subtracted.

tional changes are induced in the minor groove of the helix, resulting in positive cooperative binding of peptide molecules. Enhancement of the CD bands is correlated with peptide–DNA binding stoichiometry, suggesting a three-step stoichiometric binding process proceeding from 1:1 through 2:1 to non-specific *n*:1. This study therefore helps to provide insights into the network-based nature of allostery as it affects complex DNA–small molecule interactions.

4. Experimental

Protected amino acid derivatives were purchased from Bachem California (Torrance, CA) and AnaSpec, Inc. (San Jose, CA) or synthesized in our own laboratory according to published procedures. All other analytical reagents were purchased from Acros, Tedia or Sigma. The radiolabelled nucleoside triphosphate [γ-^{32}P]ATP was obtained from NEN Life Science Products at a specific activity of 6000 Ci/mmol. *Taq* polymerase, T4 polynucleotide kinase, and DNase I were purchased from Promega. DNA undecamers were purchased from Mdbio Inc., Taipei, Taiwan. Equal equivalents of complementary strands of DNA oligomers in 5 mM sodium cacodylate buffer, pH 6.5 were annealed by heating to 94 °C for 5 min followed by slow cooling to room temperature. Other chemicals were analytical grade reagents, and all solutions were prepared using deionized, Millipore filtered water. Melting points were determined on a Mel-Temp apparatus (Cambridge, Mass) and are uncorrected. Optical rotations were determined on a Rudolph Autopol II instrument. Semi-preparative and analytical HPLC (Vy-

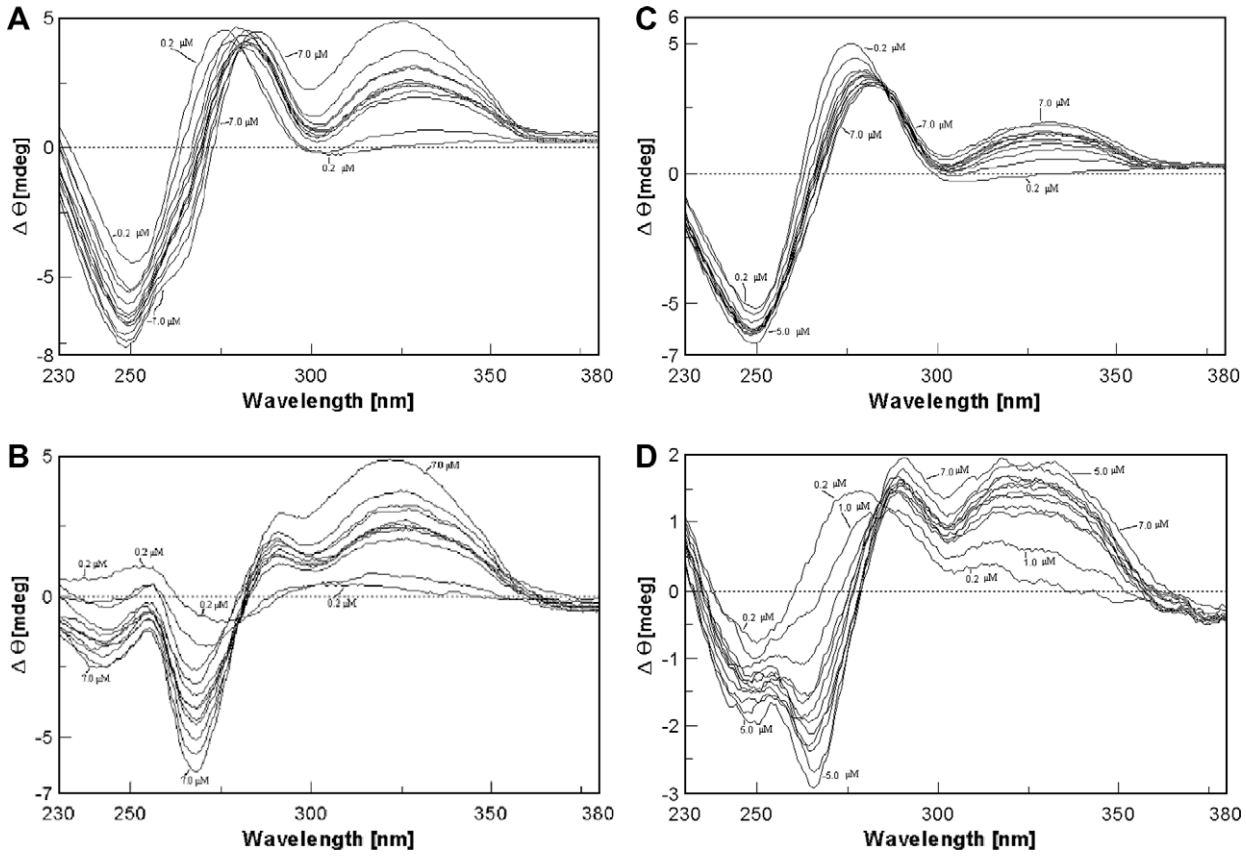


Figure 7. Panel A: titration of duplex 4A–4T versus PyHyp-12 at the same concentrations as in Figure 6B. Panel B: corresponding difference spectra with the contributions of duplex 4A–4T and peptide subtracted. Panel C: titration of duplex 3A–3T versus PyHyp-12 over the same peptide concentration range. Panel D: difference spectra with the contributions of duplex 3A–3T and peptide subtracted.

dac reversed-phase columns, TP201; column 1, 1 × 25 cm; column 2, 0.4 × 25 cm) were performed using a Hitachi L-7100 pump equipped with a gradient elution device and a Soma S-3702 variable wavelength UV detector which is connected to a PC computer installed with Hitachi HPLC analytical software. Mass spectra were determined with a Finnigan/Thermo Quest MAT 95XL instrument operating in the electrospray ionization (ESI) mode in National Chung-Hsing University.

4.1. Solid-phase peptide synthesis

4.1.1. His-Hyp-Arg-Lys-(Py)$_4$-His-Hyp-Arg-Lys-NH$_2$ (PyHyp-12)

This peptide was synthesized using solid phase methodology by manual operation of a Protein Technology PS3 peptide synthesizer. The first Fmoc-protected amino acid was coupled to the Nova Rink amide AM resin using PyBOP/NMM in DMF. All of the N_α-Fmoc-protected amino acids (in 4 equiv ratio excess to the resin) were coupled in a stepwise fashion using PyBOP/NMM in DMF after deprotection of the N_α-Fmoc group by piperidine. The side chains of Arg, Lys, Tyr and His were protected by the Pmc, Boc and Trt groups, respectively. After coupling the last *N*-terminal Fmoc-amino acid residue, the resin was treated with the cleavage reagent (0.75 g phenol, 10 mL TFA, 0.5 mL thioanisole, 0.25 mL EDT) for 1.5 h, and then lyophilized. The resin was washed first with dry ether (2 × 30 mL), filtered, and then with 5% acetic acid (200 mL). The combined filtrate was lyophilized and the product purified by semi-preparative reversed-phase HPLC (column 1) using gradient elution. Eluent A: 5% MeCN, 95% H_2O, 0.1% TFA; eluent B, 95% MeCN, 5% H_2O, 0.1% TFA. A linear gradient was achieved by increasing the MeCN content from eluent A to eluent B in 30 min. R_t (column 2), 11.59 min. Mp 154–157 °C, $[\alpha]_D^{24}$ −51.54 (*c* 0.0388, MeOH/H_2O, 1:1); ESIMS requires: 1574.75, found: 1576.0.

4.1.2. His-Hyp-Arg-Lys-(Py)$_3$-His-Hyp-Lys-Arg-NH$_2$ (PyHyp-9)

This peptide was synthesized by a procedure similar to that employed for peptide PyHyp-12. R_t (column 2), 12.43 min. Mp 150–153 °C, $[\alpha]_D^{27}$ −12.67 (*c* 0.071, H_2O); ESIMS requires: 1202.37, found: 1202.0.

4.2. Polymerase chain reaction (PCR) and end-labeling of PCR products

The pBR322 fragments: 158-mer DNA duplex (upper strand 5′-^{32}P-labeled) and 135-mer DNA duplex (lower strand 5′-^{32}P-labeled) were prepared by PCR amplification in a thermal cycler (ABI model 9700) as reported previously.[15] The sequence of the 81-mer DNA duplexes (S-81, 5′-^{32}P-labeled upper and lower strands) is shown below; these were also prepared by PCR amplification. The concentrations of DNA used as determined by UV spectroscopy are around 10^{-7} M.

```
1          10         20         30         40         50
5'-ACGTAGCGATAGCGGACTGATACTCATACTCATATCTATGGAGAAAATAC
3'-TGCATCGCTATCGCCTGACTATGAGTATGAGTATAGATACCTCTTTTATG

                   60         70         80
CGATGTCAATGTCATCAGGCGCTCTTCCGCT-3'
GCTACAGTTACAGTAGTCCGCGAGAAGGCGA-5'
```

4.3. DNase I footprinting

Reactions were conducted in a total volume of 10 μL. Briefly, radiolabelled DNA (2 μL) was mixed with varying concentrations of peptide (2 μL) dissolved in 5 mM sodium cacodylate buffer, pH 6.5) and equilibrated at room temperature for 60 min. DNase I (2 μL) was added and the reaction allowed to proceed at 37 °C for 10 min. The DNase I solution (in 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$) was adjusted to yield a final concentration of 0.009 unit/mL so as to limit the digestion to less than 30% of the starting material in order to minimize the incidence of multiple cleavages in any strand. The digestion was stopped by adding a solution (4 μL) containing 80% formamide, 10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol. The various samples were heated at 90 °C for 4 min and chilled on ice for 4 min and loaded on to the gel in a total volume of 10 μL/well. After electrophoresis (about 1.45 h at 70 W, 1800 V in TBE buffer, BRL sequencer model S2), gels were soaked in 10% acetic acid/10% methanol for 15 min, transferred to Whatman 3MM paper, dried under vacuum at 80 °C for 45 min, and placed against Kodak BioMax MR scientific imaging films with intensifying screens for about five days. The electrophoretic band areas were analyzed by a PC computer installed with Viber Lourmat BIO-ID software (Marne La Valle, Cedex 1, France). The apparent DNA binding site saturation is determined by the following equation:[11]

$$Y' = 1 - (I_{tot}/I_{ref})/(I^{o}_{tot}/I^{o}_{ref})$$

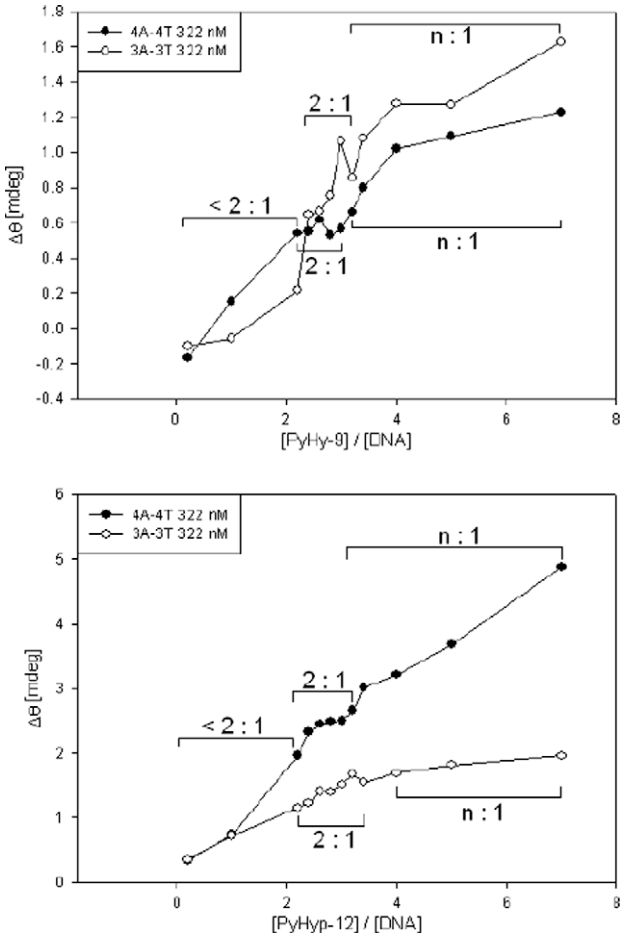


Figure 8. CD intensity at 322 nm as a function of [PyHyp-9]/[duplex] (upper panel) and [PyHyp-12]/[duplex] (lower panel). The proposed stoichiometric binding ratios are as indicated, with binding below 2:1 (<2:1) considered to be predominantly 1:1.

where Y' is the fractional saturation, I_{ref} and I^{o}_{ref} are integrated band volumes of 5 bases of a running lane (non-binding site, with peptide) and control lane (without peptide), respectively. I_{tot} and I^{o}_{tot} are integrated band volumes of the binding site locus and corresponding control lane locus, respectively.

The value of Y is optimized by the following equation:

$$Y = (Y' - Y'_{min})/(Y'_{max} - Y'_{min})$$

where Y'_{max} and Y'_{min} are the maximum and minimum site saturation Y' values, respectively.

The Hill coefficients were determined by the Hill equation:

$$\text{Log}\ [Y/1-Y] = \log K_a + n_H \log[L]$$

where $[L]$ is the peptide concentration, and n_H the Hill coefficient. At least six data points within or near the linear portion of binding isotherm (Y vs Log $[L]$) were carefully chosen for the Hill plot (Log $Y/(1-Y)$ vs Log $[L]$) using a linear least-squares fitting procedure by Sigma Plot software (version 8.0). The Hill coefficient was determined from the slope of the corresponding Hill plot. The apparent binding constant K_a is determined empirically as the peptide concentration at 50% fractional saturation from the binding isotherm.

4.4. Circular dichroism (CD) studies

CD spectra were measured at 37 °C with a Jasco J-815 instrument in the Institute of Chemistry, Academia Sinica. The duplex DNA was adjusted to 1.0 μM in 5 mM sodium cacodylate buffer (pH 6.5) and peptides, dissolved in the same buffer, were added to maintain final concentrations of 0.2, 1.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, 5.0 and 7.0 μM. CD spectra were recorded after 60 min incubation at 37 °C.

Acknowledgments

We thank Miss L. M. Hsu, National Chung-Hsing University for ESI mass analyses. This work was supported by Grant NSC93-2113-M029-003 from the National Science Council, ROC and Grant TCVGH-T-947807 to L.S.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2009.10.048.

References and notes

1. Fox, K. R.; Waring, M. J. *Methods Enzymol.* **2001**, *340*, 412.
2. Moravek, Z.; Neidle, S.; Schneider, B. *Nucleic Acids Res.* **2002**, *30*, 1182.
3. Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513. and references cited therein.
4. Waring, M. J. Echinomycin, Triostin and Related Antibiotics. In *Antibiotics*; Hahn, F. E., Ed.; Springer: Berlin, 1979; Vol. 5, pp 173–194.
5. Chaires, J. B.; Fox, K. R.; Herrera, J. E.; Britt, M.; Waring, M. J. *Biochemistry* **1987**, *26*, 8227.
6. Hudson, J. S.; Brooks, S. C.; Graves, D. E. *Biochemistry* **2009**, *48*, 4440.
7. Bailly, C.; Suh, D.; Waring, M. J.; Chaires, J. B. *Biochemistry* **1998**, *37*, 1033.
8. Chen, Y. H.; Lown, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 6995.
9. Walker, W. L.; Landaw, E. M.; Dickerson, R. E.; Goodsell, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5634.
10. Blasko, A.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10018.
11. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.
12. Fechter, E. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8476. and references cited therein.
13. Weyermann, P.; Dervan, P. B. *J. Am. Chem. Soc.* **2002**, *124*, 6872.
14. Yang, C. H.; Chou, P. J.; Luo, Z. L.; Chang, J. C.; Cheng, C. C.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *Bioorg. Med. Chem.* **2003**, *11*, 3279.
15. Chang, J. C.; Yang, C. H.; Chou, P. J.; Yang, W. H.; Chou, I. C.; Lu, C. T.; Lin, P. H.; Hou, R. C. W.; Jeng, K. C. G.; Cheng, C. C.; Sheh, L. *Bioorg. Med. Chem.* **2004**, *12*, 53.
16. Yang, C. H.; Chen, W. F.; Jong, M. C.; Jong, B. J.; Chang, J. C.; Waring, M. J.; Ma, L.; Sheh, L. *J. Am. Chem. Soc.* **2004**, *126*, 8104.
17. Yang, C. H.; Jeng, K. C. G.; Yang, W. H.; Chen, Y. L.; Hung, C. C.; Lin, J. W.; Chen, S. T.; Richardson, S.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *ChemBioChem.* **2006**, *7*, 1187.

18. Churchill, M. E. A.; Suzuki, M. *EMBO J.* **1989**, *8*, 4189.
19. Suzuki, M. *Nature* **1990**, *344*, 562.
20. Calvano, W. E. et al *Nature* **2005**, *437*, 1032.
21. Koshland, D. E., Jr.; Hamadani, K. *J. Biol. Chem.* **2002**, *277*, 46481. and references cited therein.
22. Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860.
23. Ciubotaru, M.; Ptaszek, L. M.; Baker, G. A.; Baker, S. N.; Bright, F. V.; Schatz, D. G. *J. Biol. Chem.* **2003**, *278*, 5584.
24. Ouameur, A. A.; Tajmir-Riahi, H.-A. *J. Biol. Chem.* **2004**, *279*, 42041.
25. Tato, I.; Zunzunegui, S.; de la Cruz, F.; Cabezon, E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8156.
26. Chen, F. M.; Sha, F. *Biochemistry* **1998**, *37*, 11143.
27. Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Le Minh, N.; Cox, K. K.; Nguyen, B.; Pacheco, K. A. O.; Wilson, W. D.; Moses, Lee *J. Am. Chem. Soc.* **2005**, *127*, 742.

Determination of Allosteric Effects and Interstrand Bidentate Interactions in DNA-peptide Molecular Recognition

Yen-Chung Chen[a] (陳彥中), Jonathan T. B. Huang[a] (黃則棓), Kee-Ching G. Jeng[b] (鄭啓清), Robin C. K. Yang[a] (楊智凱), Mo-Kai Liao[a] (廖穆凱), Chee-Shan Chen[c] (陳齊聖), Wei-Jyun Chien[c] (錢偉鈞), Ming-Tsair Wey[d] (魏明財), Lou-Sing Kan[d,e] (甘魯生) and Leung Sheh[a,*] (佘　亮)

[a]*Department of Chemistry and Life Science Research Center, Tunghai Christian University, Taichung, Taiwan 407, R.O.C.*

[b]*Department of Medical Research, Taichung Veterans General Hospital, Taichung, Taiwan 405, R.O.C.*

[c]*Department of Applied Chemistry, Chaoyang University of Technology, Taichung, Taiwan 413, R.O.C.*

[d]*Institute of Chemistry, Academia Sinica, Taipei, Taiwan 11529, R.O.C.*

[e]*Department of Bioengineering, Tatung University, Taipei, Taiwan 104, R.O.C.*

To study DNA allostery, quantitative DNase I footprinting studies were carried out on a newly designed peptide His-Hyp-Lys-Lys-$(Py)_4$-Lys-Lys-NH_2 (HypKK-10) containing the XHypKK (Hyp = hydroxyproline) and polyamide motifs. The interconnection of DNA footprints of peptides HypKK-10 and the parent peptide PyPro-12 supports the proposal that interaction network cooperativity is preferred in DNA-peptide interactions between multiple recognition sites. A simple method of determining interstrand bidentate interactions between the peptide moieties and DNA bases is introduced. It is envisaged that interstrand bidentate interactions also participate in the relay of conformational changes to recognition sites on the complementary strands. Circular dichroism studies of the titration of peptide HypKK-10 with an oligonucleotide duplex indicate that this peptide binds in a dimeric fashion to DNA in the minor groove. This work may prompt the design of new DNA binding ligands for the study of DNA-peptide allosteric interactions and DNA interaction network.

Keywords: DNA-peptide molecular recognition; Positive-cooperative binding; Footprinting; Circular dichroism; Interstrand bidentate interactions; Interaction network.

INTRODUCTION

In recent years, the growing importance of future gene therapy and gene regulation has prompted extensive DNA sequence-specific recognition studies of synthetic peptides and drugs which bind specifically to the minor groove.[1-14]

It is expected that drugs or peptides with high sequence-specificity are able to bind to target promoter regions of target genes and may switch off the expression of functional proteins, thus aid in the future control of malignant diseases. Among DNA sequence-specific ligands, synthetic conjugates containing the 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) are mimics of the naturally occurring antiviral antibiotics netropsin and distamycin and have been extensively studied.[6-10] The studies of sequence-specific or sequence-selective DNA binding agents are also important in the progress of chemical biology research.

In a continuous search for new DNA sequence-specific agents we have been focusing in the design and synthesis of peptides incorporating both the XPRK (X-Pro-Arg-Lys)[11-13] and polyamide[14] motifs. The design of the new XPRK motif is prompted by the observation that a SPXX (Ser-Pro-X-X) motif[15,16] is often found in repeating sequences in histones, steroidal hormone receptors and various oncogene products. Peptides containing the SPXX motif were considered to assume a β turn stabilized by two hydrogen bonds, and the side chains of the two basic residues engaged in ionic interactions with the DNA phosphate groups. We found that peptides containing various

* Corresponding author. Tel: +886-4-2359-0248; E-mail: Lsheh@thu.edu.tw

XPRK motifs can bind to DNA in sequence-specific or sequence-selective manners in submicromolar concentration range.[11-14] In addition, we recently demonstrated that peptides incorporating the HPRK (His-Pro-Arg-Lys) and polyamide motifs[14] are involved in interstrand bidentate interactions[14,17] with bases of DNA (two or more hydrogen bonds formed between a hydrogen donor of an amino acid residue and a base pair in the minor groove). We thus concluded that the sequence-specific recognition of minor groove binding peptides is significantly dependent on interstrand bidentate and monodentate interactions of the hydrogen donor or hydrogen acceptor of the peptide moieties (side chain functions) with DNA bases.[14,17] Recent studies show that the conjugation of cytotoxic drug with peptides containing the XPRK motif produces conjugates with higher cytotoxic activity than the parent cytotoxic drug, entailing detail studies of XPRK peptides.

DNA allostery[13,18-20] has emerged as an important chemical and biochemical research topic recently. Using quantitative DNase I footprinting we find that a number of designed peptides exhibit positive cooperativity in binding to discrete sites on duplex fragments of pBR322 DNA composed of a 5′-^{32}P-labeled 158-mer(upper strand) and a complementary 135-mer (lower strand). These include the parent peptide His-Pro-Arg-Lys-$(Py)_4$-His-Pro-Arg-Lys-NH_2 (PyPro-12, formerly named as PyH-12)[14] as well as a decapeptide His-Hyp-Lys-Lys-$(Py)_4$-Lys-Lys-NH_2 (HypKK-10) which is designed as a truncated derivative of the parent peptide PyPro-12. The proline residue in the parent peptide PyPro-12 is replaced by a hydroxyproline residue that furnishes an extra hydroxyl group for additional hydrogen bonding. Quantitative footprinting is employed to investigate the cooperativity of ligand binding to multiple DNA recognition sites. To gain further insights in the molecular aspects of DNA-peptide interactions we also seek spectral evidences from circular dichroism studies.

The objective of this study is to explore the DNA binding sequence-specificity of a short designed peptide and to compare the sequence-specificity with the parent peptide PyPro-12. Secondly, to study the cooperativity between DNA binding sites induced by the binding of peptide.

In a recent paper[13] we first reported an interaction network of cooperativity interconnecting multiple DNA allosteric sites in a latent membrane protein-1 (LMP-1) gene fragment comprising of CG-rich sequence on interactions with two peptide ligands incorporating the designed XPRK motif. In this study we wish to further investigate the pattern of allosteric network in common DNA such as the pBR322 fragments consisting of AT-rich base sequences.

RESULTS AND DISCUSSION

Monodentate interactions and interstrand bidentate interactions[14,17] between the peptide moieties with DNA bases have been recognized as an essential basis of DNA sequence-specific recognition of ligands including proteins and other smaller molecules.[17] In this study, we show that the position of interstrand bidentate interactions can be assigned simply by connecting bases (with dash lines) that yield significant simultaneous DNase I blockages on complementary strands in the differential plots (Fig. 1).

Quantitative footprinting is used to study the sequence-specific binding of peptide HypKK-10 that has a truncated C-terminal dipeptide fragment with respect to PyPro-12 (formerly named as PyH-12).[14] On the 5′-^{32}P-labeled 158-mer upper strand, peptide HypKK-10 shows a strong and broad DNase I blockage around position U107-116, corresponding to the sequence 5′-AGAAAATACC-3′ and displaying significant positive cooperativity (n_H = 2.0) (Figs. 1 & 2, Table 1). At a peptide concentration of 1000 nM, a weak and broad DNase I blockage is observed around position U83-90, corresponding to the sequence 5′-AAATACCG-3′. On the 135-mer lower strand, peptide HypKK-10 shows a major DNase I blockage site around position L110-104, corresponding to the sequence 5′-TTTCTCC-3′ and displaying significant positive cooperativity (n_H = 2.3). Another weak DNase I blockage site occurs around position L85-81. Four interstrand bidentate interactions are assigned around position 107-110. The wide DNA binding loci around position U107-116 and L110-104 that encompass 12 bases suggest that two peptide molecules must bind in a dimeric fashion to these sites (Fig. 1). The binding isotherms at individual binding sites exhibit sigmoidal curvature, typical of positive cooperativity (Fig. 3).

We repeated DNase I footprinting of the parent peptide PyPro-12 on the 5′-^{32}P-labeled 158-mer and 135-mer pBR322 fragments using a wider peptide concentration range (data not shown).[19] Peptide PyPro-12 exhibits a dif-

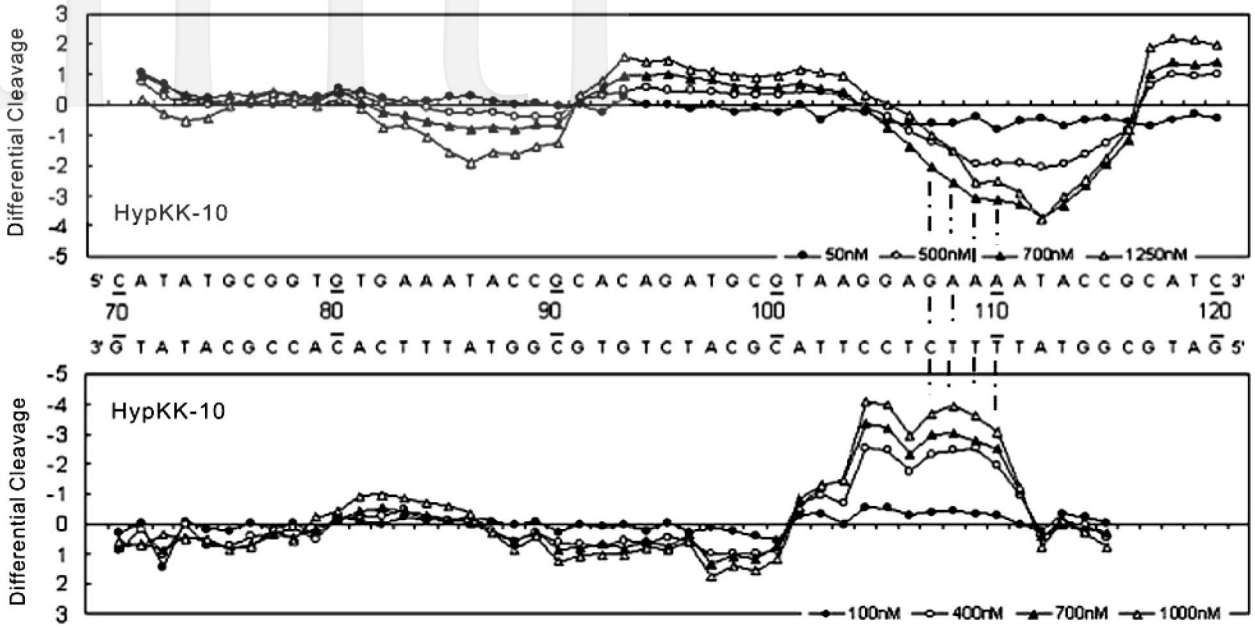


Fig. 1. Differential cleavage plots comparing the susceptibility of DNA fragments to DNase I cleavage after incubation with peptide HypKK-10 in cacodylate buffer at room temperature for 60 min. The upper traces represent the differential cleavage plot for peptide HypKK-10 bound to the 5′-^{32}P-labeled upper strand (158-mer) DNA fragment; the lower traces represent the corresponding plots for the peptide bound to the 5′-^{32}P-labeled lower strand (135-mer) DNA fragment. The vertical dotted lines between DNA bases represent assignment of interstrand bidentate interactions where significant coincident H-bonding interactions occur between complementary bases in both strands.

ferent DNA binding pattern in comparison to peptide HypKK-10. On the 158-mer upper strand, peptide PyPro-12 has two major DNase I blockage sites: a wide blockage site at position U106-116, corresponding to the sequence 5′-AGAAAATACCG-3′ and position U84-90, comprising the sequence 5′-AATACCG-3′ (Table 1). On the 135-mer lower strand, there are two major DNase I blockage sites: L85-83, corresponding to the sequence 5′-TTT-3′, and L110-105, comprising the sequence 5′-TTTCTCC-3′. Two and five interstrand bidentate interactions are assigned around positions 84-85 and 106-110, respectively (Table 1).

Table 1. Binding specificities and physicochemical parameters of sequence-specific binding of peptides to various recognition sites of complementary 5′-^{32}P-labeled upper (158-mer) and 5′-^{32}P-labeled lower (135-mer) DNA strands at 37 °C determined by quantitative DNase I footprinting

Ligand	Binding site position	Recognition sequence	K_a	n_H	Position of Interstrand bidentate interactions
HypKK-10	U83-90	5′-AAATACCG-3′	2.0×10^6	2.7	
	U107-116	5′-AGAAAATACC-3′	2.4×10^6	2.0	107-110
	L110-104	5′-TTTCTCC-3′	3.3×10^6	2.3	
	L85-81	5′-TTTCA-3′	1.6×10^6	-	
PyPro-12	U84-90	5′-AATACCG-3′	6.7×10^7	0.7	84-85
	U106-116	5′-AGAAAATACCG-3′	3.2×10^6	1.2	106-110
	L85-83	5′-TTT-3′	3.7×10^6	1.1	
	L110-105	5′-TTTCTCC-3′	3.8×10^6	1.5	

K_a nd n_H are the apparent association constant and Hill coefficient determined from concentration-dependent DNase I footprinting studies, respectively. The binding site positions on the upper and lower strands are abbreviated as U and L, respectively.

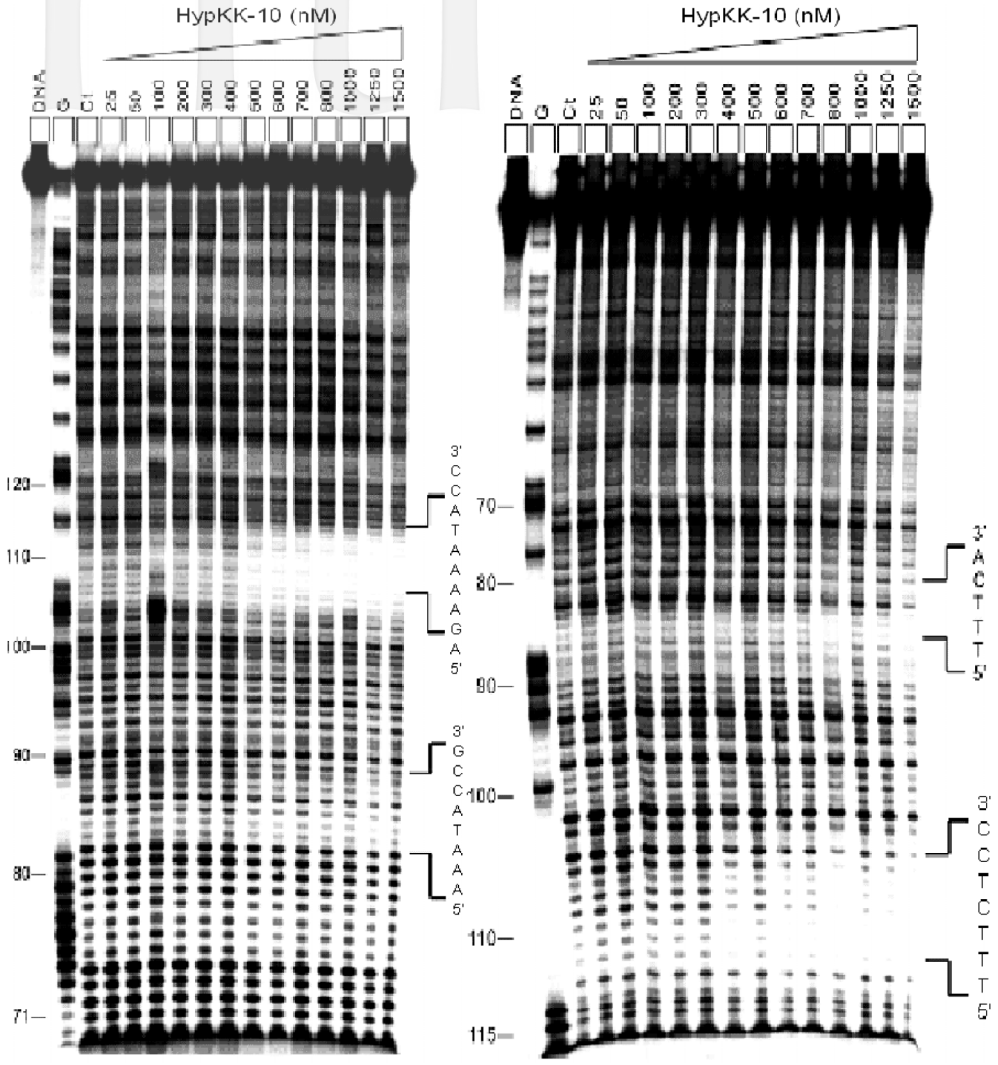


Fig. 2. Autoradiograph showing DNase I footprinting of peptide HypKK-10 on DNA duplexes labeled with [γ-^{32}P]ATP on the 5′ end: 5′-^{32}P-labeled 158-mer fragment (upper strand), left panel; and 5′-^{32}P-labeled 135-mer fragment (lower strand), right panel. Peptide HypKK-10 was equilibrated with the DNA in 5 mM sodium cacodylate buffer, pH 6.5 at 37 °C for 60 min before DNase 1 cleavage. G represents a Maxam-Gilbert guanine sequencing track and Ct shows a DNase I digestion control lane.

It appears that both peptides HypKK-10 and PyPro-12 share strong sequence preference on DNA loci to four sites on the upper and lower strands where the complementary sequences d(AAAA)-d(TTTT) and d(AAA)-d(TTT) occur around positions 108-111 and 83-85 respectively. However, it is notable that on the upper and lower strands, the parent peptide PyPro-12 possesses four major binding sites. On the other hand, peptide HypKK-10 has three major binding sites.

In this study we found that both peptide HypKK-10 and the parent peptide PyPro-12 display significant interstrand bidentate interactions[17] and intrastrand H-bond interactions of the side chain functions and amino protons of the peptides with bases in the complementary DNA strands. It is noteworthy that the antibiotic netropsin comprising of Py residues and without amino acid residues does not display interstrand interactions on complementary strands (autograph not shown). Both peptides HypKK-10 and PyPro-12 display interstrand bidentate interactions, suggesting that the amino acid side chains of these peptides are

responsible for interstrand interactions between AT base pairs (Fig. 2). It has been shown that the side chains of Arg, Lys, His and Tyr may involve in single strand H-bond interactions as well as interstrand bidentate interactions in proteins.[17] Complex interactions in which the side chain function of amino acid residues bind more than one base step simultaneously may also occur in DNA-peptide interactions.[17] In the contrast, footprinting studies of peptides with gene fragments of the latent membrane protein (LMP-1) showed that there are no interstrand H-bond interactions between the peptide moieties and DNA bases.[13]

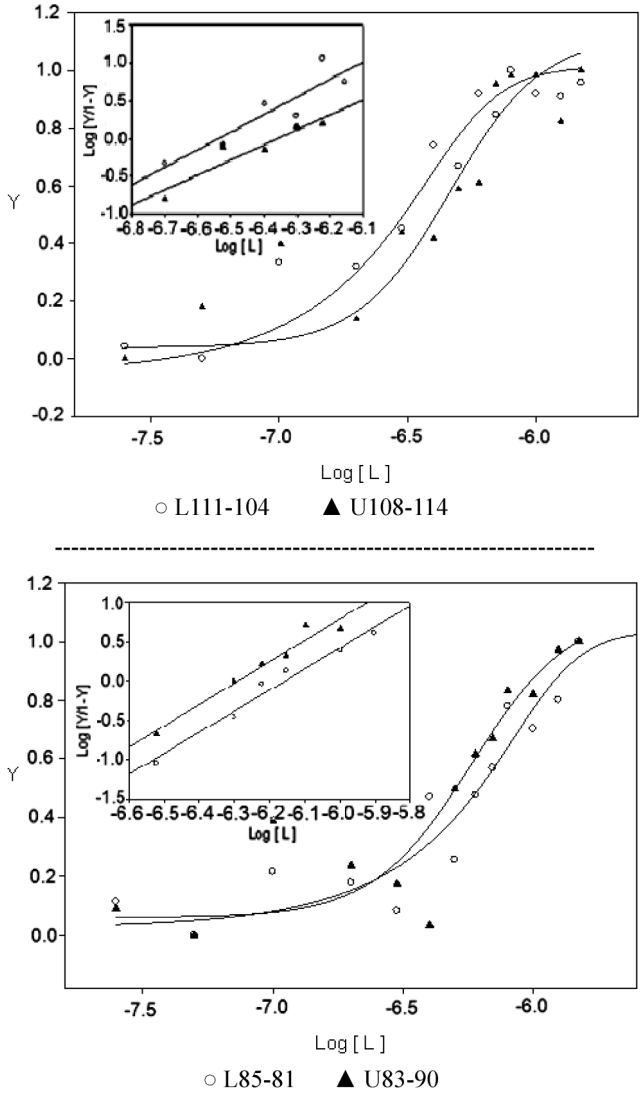


Fig. 3. Binding isotherms from DNase I footprinting titrations of peptides (log molar concentration *versus* Y, the fractional saturation). Insets are the respective Hill plots of log[Y/1-Y] *versus* log[L]. The concentration of ligand (peptide) is in nM.

Based on the footprinting results, interaction network models can be constructed to interpret the complex communication of DNA allosteric binding sites for the peptides (Fig. 4). It is clear that the models reveal network-based allosteric connections between recognition sites. Conformational changes of binding loci are transmitted to adjacent binding sites and also to sites on the complementary strand. It is likely that interstrand bidentate interactions also involve in the relay of conformational changes to binding sites on the complementary strands, facilitating the binding of further peptide molecules. To gain further understanding in the DNA-peptide cooperative binding and molecular recognition process we studied the DNA binding process of peptide HypKK-10 with circular dichroism spectroscopy (CD) that may provide spectral evidences. A 13-mer deoxyribonucleotide duplex d(TAGGAGAAAATAC)-d(GTATTTTCTCCTA) (U4A-L4T) (Fig. 5a) that represents the complementary recognition site of the 158-mer fragment, at positions 104-114 was used for CD positive band around 295 nm were observed. On increasing peptide concentration, the induced negative CD band is blue-shifted to 248 nm and a positive band is red-shifted to 283 nm (Fig. 5b). This positive CD band extends as a strong broad dose-dependent positive CD enhancement band to about 334 nm and centers around 325 nm. The 325 nm positive band is free of interference from free DNA and peptide, and is used for plotting binding isotherms (Fig. 6). In the difference spectrum (Fig. 5c), binding of peptide HypKK-10 to the duplex results in dose-dependent CD intensity enhancement of a negative band centering around 266 nm and a wide negative band extending under 220 nm. Two peptide concentration-dependent positive CD bands occur around 288 nm and 330 nm, respectively. A near-isoelliptic point occurs around 278 nm, suggesting predominantly a two-component binding process. The significant CD enhancement band around 325 nm of these peptides suggested that the major interaction is in the DNA minor groove.[19-21]

It is apparent that the enhancement of the CD bands is correlated with peptide-DNA binding stoichiometry. Fig. 6 shows the [Peptide]/[DNA duplex] plot of peptides HypKK-10 and PyPro-12. It appears that peptide HypKK-10 induces greater ellipticity change than that of PyPro-12, sug-

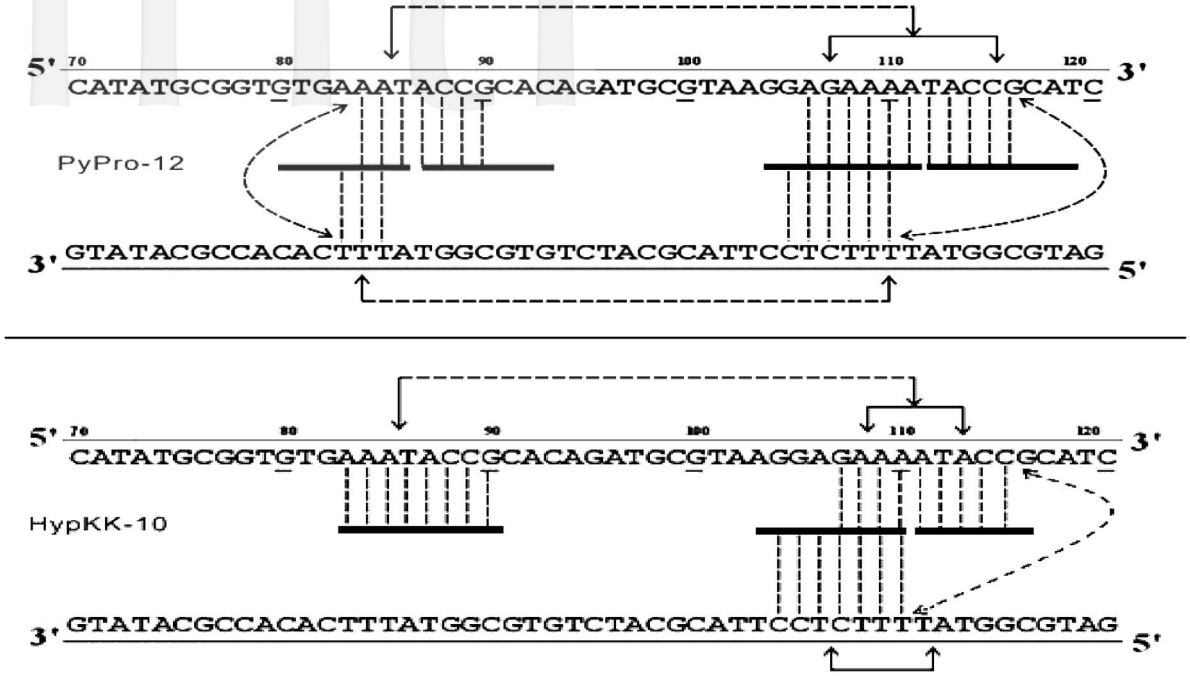


Fig. 4. Proposed allosteric models for cooperative binding of peptides PyPro-12 and HypKK-10 to 158-mer and 135-mer pBR322 fragments based on quantitative footprinting studies. The portion of the ligand binding to each DNA site/sub-site is represented by a thick horizontal line. Monodentate interactions and interstrand bidentate interactions are represented by vertical broken lines. Horizontal arrow lines represent communication of allosteric interaction between DNA binding sites. Arrow lines between the complementary strands represent possible allosteric influence by interstrand bidentate interactions which relay conformational changes to binding sites.

gesting that more drastic DNA conformational changes are accompanied with the binding of peptide HypKK-10. At [peptide]/[DNA duplex] ratio of 0.5 to 2.0, one molecule of peptide binds to the d(AAAA)-d(TTTT) locus of the minor groove. At [Peptide]/[DNA duplex] ratios 2.0-3.0, two peptide molecules bind in a dimeric fashion to the binding site in the minor groove. At ratios above 4.0, progressive increase in Δθ suggests more peptide molecules begin non-sequence selective binding, possibly to the major groove.

The CD results agree well with DNase I footprinting results that dimeric peptide binding to the d(AAAA)-d(TTTT) locus is favored. However, some peptides may induce non-sequence selective binding to DNA when their concentrations exceed 2 μM.

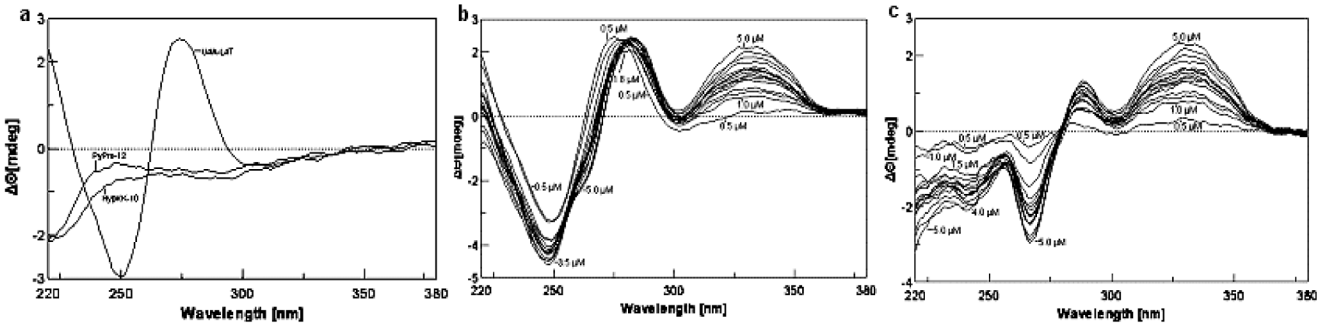


Fig. 5. Panel **a**: CD spectra of DNA duplex U4A-L4T alone and peptides HypKK-10 and PyPro-12 alone. Panel **b**: Titration of DNA duplex U4A-L4T versus HypKK-10 at peptide concentrations of 0.2, 1.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, 5.0 μM. Panel **c**: Corresponding CD difference spectra with the contribution of duplex and peptide subtracted.

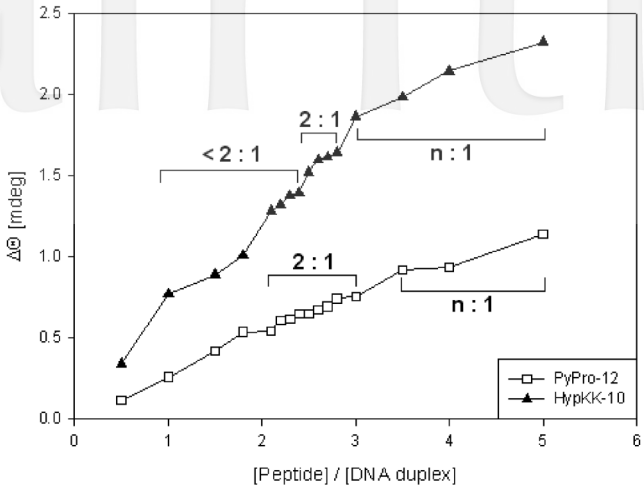


Fig. 6. CD intensity at 322 nm as a function of [Peptide]/[DNA duplex]. The proposed stoichiometric binding ratios are as indicated, with binding below 2:1 (< 2:1) considered to be predominantly 1:1.

CONCLUSION

This study shows that the interconnection of DNA footprints of peptides HypKK-10 and the parent peptide PyPro-12 supports the proposal that interaction network cooperativity is preferred in many DNA-peptide interactions between multiple recognition sites.[13,19,20] It is envisaged that interstrand bidentate interactions also participate in the relay of conformational changes to recognition sites on the complementary strands. This work may prompt the design of new DNA binding ligands for the study of DNA-small molecule allosteric interactions and DNA interaction network.

EXPERIMENTAL SECTION

Chemicals and Biochemicals

All of the protected amino acid derivatives were purchased from Bachem California (Torrance, CA) and AnaSpec, Inc. (San Jose, CA), and all peptides were synthesized in our laboratory. Analytical reagents were purchased from Acros, Tedia or Sigma. The radiolabeled nucleoside triphosphate [γ-^{32}P]ATP was obtained from NEN Life Science Products at a specific activity of 6000 Ci/mmol. Taq polymerase, T4 polynucleotide kinase, and DNase I were purchased from Promega. All other chemicals were analytical grade reagents, and all solutions were prepared using deionized, Millipore-filtered water.

Chemical methods

Melting points were determined on a Mel-Temp apparatus (Cambridge, Mass) and are uncorrected. Optical rotations were determined on a Rudolph Autopol II instrument. CD spectra were measured at 37 °C with a Jasco J-815 spectrometer. Semi-preparative and analytical HPLC (Vydac reversed-phase columns, TP201; column 1, 1 × 25 cm; column 2, 0.4 × 25 cm) were performed using a Hitachi L-7100 pump. Mass spectra were determined with a Finnigan/Thermo Quest MAT 95XL instrument operating in the electrospray ionization (ESI) mode in Chung-Hsing University.

His-Hyp-Lys-Lys-(Py)$_4$-Lys-Lys-NH$_2$(HypKK-10)

This peptide was synthesized using solid phase methodology by manual operation of a Protein Technology PS3 peptide synthesizer. The first Fmoc-protected amino acid was coupled to the Nova Rink amide AM resin using PyBOP/NMM in DMF. All of the Nα-Fmoc-protected amino acids (in 4 equivalent ratio excess to the resin) were coupled in a stepwise fashion using PyBOP/NMM in DMF after deprotection of the Nα-Fmoc group by piperidine. The side chains of Lys and His are protected by the Boc and Trt groups, respectively. After coupling the last *N*-terminal Fmoc-amino acid residue, the resin was treated with the cleavage reagent (0.75 g phenol, 10 mL TFA, 0.5 mL thioanisole, 0.25 mL EDT) for 1.5 h, and then lyophilized. The resin was washed with dry ether (2 × 30 mL), filtered, and then washed with 5% acetic acid (200 mL). The combined filtrate was lyophilized and the product was purified by semi-preparative reversed-phase HPLC (column 1) using gradient elution. Eluent A: 5% MeCN, 95% H_2O, 0.1% TFA; Eluent B, 95% MeCN, 5% H_2O, 0.1% TFA. A linear gradient was achieved by increasing the MeCN content from eluent A to eluent B in 30 minutes. R_t (column 2), 14.38 min., mp 155-158 °C, $[\alpha]_D^{27}$ -14.55 (*c* 0.183, MeOH/H_2O, 1:1); ESIMS requires: 1268.47, found: 1268.53.

Polymerase chain reaction (PCR) and end-labeling of PCR products

The 158-mer DNA duplex (upper strand 5′-^{32}P-labeled) and 135-mer DNA duplex (lower strand 5′-^{32}P-labeled) were prepared by PCR amplification in a thermal cycler (ABI model 9700) as reported previously.[13,14] The DNA concentration was determined by UV spectroscopy to lie in the range of 600-800 nM.

DNase I footprinting

Reactions were conducted in a total volume of 10 μL. Radiolabeled DNA (2 μL) was mixed with varying concentrations of peptide (2 μL) dissolved in 5 mM sodium cacodylate buffer, pH 6.5 and equilibrated at room temperature for 120 min. DNase I (2 μL) was added and the reaction allowed to proceed at 37 °C for 10 min. The DNase I solution (in 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$) was adjusted to yield a final concentration of 0.009 unit/mL so as to limit the digestion to less than 30% of the starting material in order to minimize the incidence of multiple cleavages in any one strand. The digestion was stopped by adding stop solution (4 μL) containing 80% formamide, 10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol. Samples were heated at 90 °C for 4 min and chilled on ice for 4 min prior to electrophoresis. The products of DNase I cleavage were resolved by polyacrylamide gel electrophoresis under denaturing conditions (0.3 mm thick, 8% acrylamide containing 8 M urea). After electrophoresis (about 1.45 h at 70 Watts, 1800 V in TBE buffer, BRL sequencer model S2), gels were soaked in 10% acetic acid/ 10% methanol for 15 min, transferred to Whatman 3MM paper, dried under vacuum at 80 °C for 45 min and placed against an imaging plate of a Fuji phosphoImager FLA-5100 overnight. The electrophoretic band areas were analyzed by a PC computer installed with Viber Lourmat BIO-ID software (Marne La Valle, Cedex 1, France). Differential cleavage plots were computed The Hill coefficients were determined by the Hill equation:

$$\text{Log}[Y/1\text{-}Y] = \log K_a + n_H \log[L]$$

where Y is the fractional saturation, [L] the peptide concentration, and n_H the Hill coefficient. The apparent DNA target site saturation Y was calculated according to the method described previously.[13,19] At least six data points within or near the linear portion of binding isotherm (Y versus Log[L]) were carefully chosen for the Hill plot (LogY/ (1-Y) versus Log[L]) using a linear least-squares fitting procedure by Sigma Plot software (version 8.0). The Hill coefficient was determined from the slope of the corresponding Hill plot. The apparent binding constant K_a is determined empirically as the peptide concentration at 50% fractional saturation from the binding isotherm.

Circular dichroism (CD) studies

CD spectra were measured at 37 °C with a Jasco J-815 spectrometer in the Institute of Chemistry, Academia Sinica. The duplex DNA was adjusted to 1.0 μM in 5 mM sodium cacodylate buffer (pH 6.5) and peptide, dissolved in the same buffer, were added to maintain final concentrations of 0.2, 1.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 4.0, and 5.0 μM. CD spectra were recorded after 60 min incubation at 37 °C.

ACKNOWLEDGEMENT

We thank Professors C. C. Cheng and M. J. Waring for helpful discussions, and Miss L. M. Hsu, National Chung-Hsing University for ESI mass analyses. This work was supported by grant NSC 97-2113-M029-005 from the National Science Council, R.O.C.

Received November 24, 2009.

REFERENCES

1. Waring, M. J. In *Antibiotics*; Hahn, F. E., Eds.; Springer-Verlag: Heidelberg, 1979.
2. Moravek, Z.; Neidle, S.; Schneider, B. *Nucleic Acids Res.* **2002**, *30*, 1182-1191.
3. Fox, K. R.; Waring, M. J. *Meth. Enzymol.* **2001**, *340*, 412-430.
4. Hudson, J. S.; Brooks, S. C.; Graves, D. E. *Biochemistry* **2009**, *48*, 4440-4447.
5. Bailly, C.; Suh, D.; Waring, M. J.; Chaires, J. B. *Biochemistry* **1998**, *37*, 1033-1045.
6. Chen, Y. H.; Lown, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 6995-7005.
7. Walker, W. L.; Landaw, E. M.; Dickerson, R. E.; Goodsell, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5634-5639.
8. Chen, F. M.; Sha, F.; Chin, K. H.; Chou, S. H. *Nucleic Acids Res.* **2004**, *32*, 271-277.
9. de Clairac, R. P. L.; Geierstanger, B. N.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7909-7916.
10. Fechter, E. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8476-8485, and references cited therein.
11. Yang, C. H.; Chou, P. J.; Luo, Z. L.; Chang, J. C.; Cheng, C. C.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *Bioorg. Med. Chem*. **2003**, *11*, 3279-3288.
12. Yang, C. H.; Chen, W. F.; Jong, M. C.; Jong, B. J.; Chang, J. C.; Waring, M. J.; Ma, L.; Sheh, L. *J. Am. Chem. Soc.* **2004**, *126*, 8104-8105.
13. Yang, C. H.; Jeng, K. C. G.; Yang, W. H.; Chen, Y. L.; Hung, C. C.; Lin, J. W.; Chen, S. T.; Richardson, S.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *Chem Bio Chem* **2006**, *7*, 1187-

1196.
14. Chang, J. C.; Yang, C. H.; Chou, P. J.; Yang, W. H.; Chou, I. C.; Lu, C. T.; Lin, P. H.; Hou, R. C. W.; Jeng, K. C. G.; Cheng, C. C.; Sheh, L. *Bioorg. Med. Chem.* **2004**, *12*, 53-61.
15. Churchill, M. E. A.; Suzuki, M. *EMBO J.* **1989**, *8*, 4189-4195.
16. Suzuki, M. *Nature* **1990**, *344*, 562-565.
17. Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860-2869.
18. Chaires, J. B. *A.C.S. Chem. Biol.* **2008**, *3*, 207-209.
19. Kao, K. L.; Jonathan, C. T.; Huang, J. C. T.; Yang, C. K.; Jeng, K. C. G.; Chang, J. C. C.; Yao, W. C.; Hsien, S. C.; Waring, M. J.; Chen, M. H.; Ma, L.; Sheh, L. *Bioorg. Med. Chem.* **2010**, *18*, 366-376.
20. Huang, J. T. B.; Chen, Y. C.; Chang, J. C.; Jeng, K. C. G.; Kao, K. L.; Yang, R. C. K.; Kan, L. S.; Wey, M. T.; Waring, M. J.; Chen, C. S.; Chien, W. J.; Sheh, L. *Bioorg. Med. Chem*. **2010**, *18*, 2575-2585.
21. Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Minh, Le, N.; Cox, K. K.; Nguyen, B.; Pacheco, K. A. O.; Wilson, W. D.; Lee, M. *J. Am. Chem. Soc*. **2005**, *127*, 742-750.