CiQUS
Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Designed metallopeptides as tools in Chemical Biology

Ph.D. 2017

Ghofrane Barka

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

D. MIGUEL VÁZQUEZ LÓPEZ, PROFESOR DEL DEPARTAMENTO DE QUÍMICA INORGÁNICA DE LA UNIVERSIDAD DE SANTIAGO DE COMPOSTELA Y **D. M. EUGENIO VÁZQUEZ SENTÍS**, PROFESOR DEL DEPARTAMENTO DE QUÍMICA ORGÁNICA DE LA UNIVERSIDAD DE SANTIAGO DE COMPOSTELA,

CERTIFICAN: que la memoria adjunta titulada ***"Designed Metallopeptides as Tools in Chemical Biology"*** que, para optar al grado de Doctora en Química, presenta Dra. Ghofrane Barka, ha sido realizada bajo nuestra dirección en los laboratorios del Centro Singular de Investigación en Química Biológica y Materiales Moleculares (CIQUS) de la Universidad de Santiago de Compostela.

Considerando que constituye trabajo de Tesis, autorizamos su presentación en la Universidad de Santiago de Compostela.

Y para que así conste, se expide el presente certificado en Santiago de Compostela a 11 de Septiembre de 2017.

Fdo. Prof. Miguel Vázquez López
Director y tutor

Fdo. Prof. M. Eugenio Vázquez Sentís
Codirector

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Acknowledgements

I would particularly like to thank my supervisors Eugenio and Miguel for their excellent cooperation and for all of the opportunities I was given to conduct my research and further my dissertation. I strongly acknowledge all their support and patience with my everyday work as a PhD student.

I am also very thankful with Eu-Metalic (II) Erasmus scholarship for giving me afterwards the opportunity to expand my research experience.

Special thanks to the international office Team for their help and advice.

I also want to express my deep appreciation for my colleagues, especially Iria, Gustavo, Diego, David, Jacobo, Sonia, Lidia, Renata, Cristina, Jessica, Sergio and Arcadio, for the help and support you have offered me on my recent project. I feel so grateful to you for taking time out of your busy schedule to answer so many questions. Without your help, I would not have been able to complete the project in such a proficient and timely manner.

Out of the lab, and for making my time in Santiago enjoyable, I would like to thank Irene, Marouan, Ahmad, Anas, Salam and Diana, for their friendship and good time.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study. Thank you.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Index

Abbreviations

δ	chemical shift	l	optical path lenght
ε	extinction coefficient	M	molar
λ	wavelength	m	multiplet
μL	microliter	mAU	absorbance miliunits
μM	micromolar	Me	methyl
μm	micrometer	MeCN	acetonitrile
Ac	acetyl	MeOH	methanol
AcOH	acetic acid	mg	miligram
a.u.	arbitrary units	MHz	megahertz
Boc	tert-butoxycarbonyl	min	minute
Bzl	bencyloxycarbonyl	mL	mililiter
cm	centimeter	mM	milimolar
c.p.s.	counts per second	mmol	milimol
d	doublet	mre	molar ellipticty per residue
dd	double doublet	MS	mass spectrometry
ds	double stranded	m/z	mass charge relation
DIEA	*N*,*N*-diisopropylethylamine	nM	nanomolar
DMF	dimethylformamide	nm	nanometer
dmol	decimol	PBS	phosphate-buffered saline
DMSO	dimethylsulfoxide	PhSiH3	phenylsilane
DNA	deoxyribonucleic acid	RMN	nuclear magnetic resonance
em	emission	rt	room temperature
equiv	equivalents	s	secon, singlet
ESI	electrospray ionization	SPPS	solid phase peptide synthesis
Et_3N	triethylamine	t-Bu	*tert*-butyl
Et_2O	diethyl ether	TCEP	tris(2-carboxyethyl)phosphine
EtOAc	ethyl acetate	TFA	trifluoroacetic acid
EtOH	ethanol	THF	tetrahydrofuran
exc	excitation	TIS	triisopropylsilane
Fmoc	9-fluorenylmethoxycarbonyl	TMS	tetramethylsilane
h	hour	TNBS	2,4,6-trinitrobenzene sulfonic acid
J	coupling constant		
K_D	dissociation constant	UV	ultraviolet

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Abstract

El mapeado del genoma humano ha proporcionado una gran cantidad de información y objetivos potenciales en el ADN, estimulado la búsqueda de nuevos agentes de unión al mismo. Entre los agentes clásicos de reconocimiento específico del ADN destacan las moléculas orgánicas pequeñas. Sin embargo, la mayoría de los compuestos estudiados hasta la fecha presentan muchas limitaciones, entre las que destacan los problemas de especificidad en la secuencia de unión de ADN (toxicidad), así como problemas de estabilidad química y para su internalización en las células, que dificultan su aplicación terapéutica. Los estudios sobre complejos metálicos como agentes de unión al ADN son menos numerosos en comparación con los realizados sobre moléculas orgánicas pequeñas, debido principalmente a su elevada labilidad en medio fisiológico, a la deficiente biocompatibilidad de sus ligandos precursores y sobre todo a los procesos de síntesis poco flexible que dificultan la obtención de análogos para la realización de estudios sistemáticos de estructura/actividad.

Los complejos de Ru(II) son particularmente atractivos como sondas y sensores de ADN aplicables en técnicas de imagen molecular, puesto que son altamente solubles en agua, poseen una gran estabilidad cinética y presentan longitudes de onda de excitación y emisión relativamente elevadas que permiten su utilización en tejidos biológicos, opacos a longitudes de onda UV y visible cercano. Además, sus bandas de emisión son sensibles a la interacción con el ADN, lo que los convierte en sensores de gran potencial. Debido a ello, los últimos años se han descrito multitud de complejos mononucleares de Ru(II), ejemplos de sistemas dinucleares e incluso precedentes de péptidos conjugados a complejos mononucleares polipiridínicos de Ru(II) con propiedades de unión al B-ADN, aunque en todos los casos los complejos metálicos son componentes estructuralmente aislados del fragmento peptídico.

Recientemente, se ha comprobado que los complejos polipiridínicos de Ru(II) derivados del ligando dipirido [3,2-a: 2 ', 3'-c] fenazina (dppz), son capaces de insertarse entre pares de bases no complementarias. Diversos estudios sugieren que el aumento de la emisión fluorescente que experimentan estos complejos al insertarse es mayor que cuando se intercalan, lo cual puede ser usado para detectar desajustes en el ADN. Es necesario indicar que el desarrollo de métodos de detección de pares de bases no complementarias en el ADN es un objetivo muy importante dentro del proceso de desarrollo de nuevos métodos

terapéuticos y de diagnosis para el cáncer, ya que, la existencia de deficiencias en el mecanismo de reparación de desajustes en el ADN de las células es directamente proporcional a la presencia de estos errores de emparejamiento, lo que a su vez está relacionado con diferentes tipos de cáncer.

Por otra parte, los péptidos son una clase de compuestos de enorme interés en el área de Química Biológica. En sus orígenes, el diseño de péptidos y proteínas se enfocó en la obtención de sistemas pequeños que permitieran entender su mecanismo de plegamiento. Posteriormente, comenzaron a construirse sistemas dotados de propiedades y funcionalidades específicas. Recientemente, los métodos de síntesis de péptidos y la expresión de proteínas, han permitido manipular las secuencias de aminoácidos para obtener nuevas proteínas "*de novo*" o modificar proteínas naturales de forma programada. La estructura secundaria y terciaria de las proteínas está estabilizada por interacciones no covalentes o por enlaces coordinados metal-ligando. Por otra parte, los péptidos, al ser más pequeños que las proteínas, suelen tener más dificultades para configurar un núcleo hidrofóbico que estabilice su estructura. Una alternativa para la obtener péptidos estructurados es explotar los enlaces metal-ligando a través del ensamblaje de secuencias peptídicas equipadas con cadenas laterales capaces de coordinar iones metálicos en las posiciones adecuadas. Para ello podemos hacer uso de los aminoácidos naturales o bien, de aminoácidos coordinantes artificiales. La gran ventaja de los aminoácidos coordinantes artificiales sobre los naturales es que éstos pueden orientar más de un átomo dador hacia el mismo ión metálico, actuando como ligandos quelato. Existen numerosos ejemplos de metalopéptidos artificiales en bibliografía que presentan diferentes funcionalidades, como agentes de unión al ADN y proteínas, agentes antitumorales, sondas para imagen molecular, entre otros. De forma general se puede decir que los metalopéptidos combinan las propiedades de unión al ADN de los complejos metálicos con las posibilidades de unión específica que ofrecen los péptidos a través de contactos específicos entre las cadenas laterales del esqueleto peptídico y los pares de bases del ADN. En este contexto, la molécula bidentada heteroaromática Bpy es uno de los ligandos más utilizados en Química de Coordinación y Supramolecular. Se trata de una base de Lewis frontera caracterizada por presentar una alta afinidad por una gran variedad de cationes metálicos del bloque *d* debido a su capacidad quelatante y de actuar como aceptor-π. Las grandes ventajas coordinativas que presenta la Bpy frente a los grupos presentes en las cadenas laterales de los aminoácidos naturales,

justifican su utilización en el diseño y síntesis de metalopéptidos artificiales bioactivos.

En este trabajo de tesis doctoral se ha sintetizado una serie de metalopéptidos de Ru (II) que contienen el ligando dipirido [3,2-a: 2 ', 3'-c] fenazina (dppz) mediante métodos de síntesis de péptidos en fase sólida (SPPS). Los estudios espectroscópicos indican que la funcionalización de estos complejos con una secuencia de oligoarginina provoca un cambio drástico en su modo de unión al ADN y un gran aumento de la afinidad de estos metalopéptidos por oligonucleótidos de ADN bien ensamblados frente a los que presentan bases desemparejadas en su secuencia. Además, los datos espectroscópicos sugieren que tanto la funcionalización de los complejos con oligoarginina como la naturaleza de los ligandos auxiliares no provocan un aumento de la selectividad de estos metalopéptidos por un tipo particular de error en oligonucleótidos de ADN con fallos. Por último, estudios de microscopía de fluorescencia indican que la funcionalización con oligoarginina causa la internalización eficiente de estos metalopéptidos en células Vero, causando la muerte celular apoptótica después de pocas horas.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

General Introduction

Chemical Biology: origins and relevance

The research at the interface between chemistry and biology has experienced an extraordinary development in the recent years. Chemical Biology is becoming an established discipline that involves the use of chemical tools to study the structure and function of biomolecules and the molecular mechanisms of biological systems.[1] It is defined as a field of science that uses chemistry to provide a molecular understanding of biology and takes advantage of the biological knowledge to advance research in chemistry;[2] its intrinsic interdisciplinarity complicates the precise definition of its scope, since it comprises diverse subjects such as biomolecular recognition, enzyme catalysis, cell signalling, cell internalization and transport, biosynthesis, molecular mechanisms of gene expression and regulation, bioconjugation methods or probe design.

Chemical Biology dissolves the boundaries between the classic disciplines of Chemistry and Biology. While Chemistry was centered in the study of the inert matter, maintaining a more analytical approach, Biology was basically phenomenological, mainly interested in processes that occurred in cells and living organisms.

A fortuitous discovery by Friedrich Wöhler in 1828, paved the way for the formulation of a chemical basis for life. In an attempt to synthesize ammonium cyanide, Wöhler heated a solution of silver cyanide and ammonium chloride, but instead of obtaining the expected silver chloride and ammonium cyanide mixture, the observed product was urea.[3] This experiment showed that inorganic materials could be used to synthesize substances previously associated only with living organisms. In other words, chemical synthesis does not require life or life force for obtaining biologically active compounds, thus the concept of vitalitsm collapsed. This finding was a chemical revolution, to the point that some chemists began to be interested in other fields like organic synthesis, or in the chemical reactions that occur in life organisms. This interest by combining the knowledge of chemistry and biology resulted in the apparition in the XIX

[1] B. Imperiali and S. Aldrige, *Chem. Commun.*, **2003**, *15*, 445–447.
[2] K. L. Morrison and G. a Weiss, *Nat. Chem. Biol.*, **2006**, *2*, 3–6
[3] F. Whöler, *Annalen der Physik und Chemie*, **1828**, *88*, 253-256.

century of a new discipline: Biochemistry. While this discipline is more focused to the study of biological processes and the behaviour of biomolecules and biopolymers from a biological perspective, Chemical Biology deals with the development of chemical tools to further understand the molecular bases of biology as well as to modify them in a selective manner.[4]

Figure 1. Wöhler's synthesis of urea, which represents a landmark achievement in organic synthesis and biological chemistry.

This new field of science has rapidly established in the last years as demonstrated by the launching of several specialized journals like *Chemistry & Biology*, *ChemBioChem*, *Current Opinion in Chemical Biology*, *ACS Chemical Biology*, *Organic & Biomolecular Chemistry* or *Nature Chemical Biology*.

Sequence-Specific DNA Recognition with artificial peptides

Introduction

Gene transcription in eukaryotic cells is largely controlled by the interaction of certain proteins called transcription factors (TFs) with specific DNA sequences,[5] which modify the basal levels of transcription, either enhancing or repressing their expression.[6] It is known that misregulation of TFs is at the origin of a number of diseases, including cancer,[7] and thus there is growing interest in understanding the molecular basis of specific DNA recognition, as well as in developing designed DNA binding agents. Furthermore, in addition to the potential applications of such synthetic binders, the development of artificial DNA binding agents represents an unmet challenge at the crossroads between supramolecular and biological chemistry and, beyond its biological relevance, it might find future applications in combination with DNA nanotechnology.[8] Most

[4] G. von Kiedrowski, Chembiochem, 2001, 2, 597–598.
[5] A. H. Brivanlou, J. E. Darnell Jr, *Science* **2002**, *295*, 813–818.
[6] a) A.-L. Todeschini, A. Georges, R. A. Veitia, *Trends Genet*. **2014**, 30, 211–219; b) M. Ptashne, *Nature* **1986**, *322*, 697–701.
[7] P. P. Pandolfi, *Oncogene* **2001**, *20*, 3116–3127.
[8] A. R. Chandrasekaran, *Nanoscale* **2016**, *8*, 4436–4446.

artificial DNA binding agents described to date are small molecules, including coordination compounds, as well as small organic molecules, which typically bind to the DNA through minor groove insertion or intercalation mechanisms, and have been the subject of a number of excellent reviews.[9] [10] In addition to those classic DNA-binding agents, researchers, inspired by natural TFs, have also explored the potential of designed peptides for the specific recognition of DNA, developing increasingly sophisticated systems that not only display excellent binding properties, but also are endowed with new properties not found in their natural counterparts, such as luminescence, photocontrol, nuclease activity, or stimuli-responsive binding, which we will review in the following pages.[11]

DNA recognition from a supramolecular perspective

The B-DNA, the most relevant DNA conformation under physiological conditions, is characterized by a rather uniform right-handed double helix formed by two antiparallel oligodeoxyribonucleotide chains, held together by the combination of stacking interactions between the base pairs (bps), and hydrogen bonds between complementary Watson-Crick bps.[12] The asymmetry of the deoxyribonucleotides generates two grooves in B-DNA double helix: the major groove, which is the wide side (~11.7 Å between phosphates across the groove) facing away the sugar-phosphate backbone, and the minor groove, which is the narrow side (~5.7 Å) of the bps facing towards the backbone (Figure 1). Importantly, the B-DNA conformation exposes the polar sugar/phosphate backbone while shielding the aromatic surface of the bases from the aqueous environment, thus giving rise to a monotonous

[9] a) S. Neidle, *Nat. Prod. Rep*. **2001**, 18, 291–309; b) A. Ali, S. Bhattacharya, *Bioorg. Med. Chem*. **2014**, *22*, 4506–4521; c) W. Han Ang, P. J. Dyson, *Eur. J. Inorg. Chem*. **2006**, *2006*, 4003–4018; d) C. Moucheron, *New J. Chem*. **2009**, *33*, 235–245; e) B. M. Zeglis, V. C. Pierre, J. K. Barton, *Chem. Commun*. **2007**, 4565–4579; f) C. Y. Majmudar, A. K. Mapp, *Curr. Opin. Chem. Biol.* **2005**, *9*, 467–474; g) P. B. Dervan, R. W. Bürli, *Curr. Opin. Chem. Biol*. **1999**, *3*, 688–693; h) P. B. Dervan, *Bioorg. Med. Chem*. **2001**, *9*, 2215–2235.

[10] a) J. B. Chaires, *Curr. Opin. Struct. Biol*. **1998**, *8*, 314–320; b) I. Romero-Canelón, P. J. Sadler, *Inorg. Chem.* **2013**, *52*, 12276–12291; c) A. C. Komor, J. K. Barton, *Chem. Commun*. **2013**, *49*, 3617–3630; d) G. S. Khan, A. Shah, Zia-ur-Rehman, D. Barker, *J. Photochem. Photobiol. B* **2012**, *115*, 105–118.

[11] a) E. Pazos, J. Mosquera, M. E. Vázquez, J. L. Mascareñas, *ChemBioChem* **2011**, *12*, 1958–1973; b) M. E. Vázquez, A. M. Caamaño, J. L. Mascareñas, *Chem. Soc. Rev*. **2003**, *32*, 338–349.

[12] V. A. Bloomfield, D. M. Crothers, I. Tinoco, *Nucleic Acids: Structures, Properties, and Functions*, Sterling Publishing Company, **2000**; b) C. R. Calladine, H. Drew, *Understanding DNA: The Molecule and How It Works*, Academic Press, **1997**.

physicochemical landscape, where particular base pairs can only be distinguished by relatively minor differences in their hydrogen bond donor/acceptor patterns at the bottom of the grooves (A·T/T·A, and G·C/C·G bps are indeed degenerate from the minor groove). Thus, B-DNA poses a great challenge for the development of specific binders, which given the geometric constraints, can directly "read" specific bp sequences mostly through the DNA major groove.

Despite the apparent regularity in the B-DNA, it is now recognized that the local properties of the B-DNA double helix are highly dependent on the base pair sequence, which can be considered not only a linear code, but also a structural code that influences both the conformation and molecular dynamics.[13] For example, A·T-rich tracts are known to induce bending of the DNA double helix, which is required for the activity of some regulatory sites.[14] A·T-rich sequences also display a narrower minor groove,[15] with a distinct hydration pattern along the floor of the groove,[16] and is also characterized by a much more negative electrostatic potential,[17] which can influence both reactivity,[18] and specific recognition of such sequences.[19]

Therefore, specific DNA recognition involves two types of interactions: direct readout through the formation of specific hydrogen bonds, polar interactions, and hydrophobic contacts between amino acids side chains and nucleotide functional groups—typically at the exposed edges of the bps in the DNA major groove,[20] and indirect readout, which is a more subtle mechanism related to the unique physicochemical properties of specific DNA sequences, such as readout of the stiffness of DNA double helix, its conformational flexibility, our particular electrostatic potential.[21] These effects can be even observed in the

[13] R. E. Dickerson, *J. Mol. Biol.* **1983**, *166*, 419–441.
[14] H. C. Nelson, J. T. Finch, B. F. Luisi, A. Klug, *Nature* **1987**, *330*, 221–226.
[15] D. G. Alexeev, A. A. Lipanov, Skuratovskii IYa, *Nature* **1987**, *325*, 821–823.
[16] a) S. K. Pal, L. Zhao, A. H. Zewail, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8113–8118; b) M. L. McDermott, H. Vanselous, S. A. Corcelli, P. B. Petersen, *ACS Cent. Sci.* **2017**, DOI 10.1021/acscentsci.7b00100.
[17] B. Jayaram, K. A. Sharp, B. Honig, *Biopolymers* **1989**, *28*, 975–993.
[18] E. P. Bishop, R. Rohs, S. C. J. Parker, S. M. West, P. Liu, R. S. Mann, B. Honig, T. D. Tullius, *ACS Chem. Biol.* **2011**, *6*, 1314–1320.
[19] V. Iyer, K. Struhl, *EMBO J.* **1995**, *14*, 2570–2579.
[20] S. A. Coulocheri, D. G. Pigis, K. A. Papavassiliou, A. G. Papavassiliou, *Biochimie* **2007**, *89*, 1291–1303.
[21] a) R. Rohs, S. M. West, P. Liu, B. Honig, *Curr. Opin. Struct. Biol*. **2009**, *19*, 171–177; b) R. Rohs, S. M. West, A. Sosinsky, P. Liu, R. S. Mann, B. Honig, *Nature* **2009**, *461*, 1248–1253; c) M. Y. Zhitnikova, A. V. Shestopalova, *J. Biomol. Struct. Dyn.* **2016**, 1–14.

DNA binding preferences of small molecules and, for example, it has been found that simple intercalators display some sequence selectivity due to the local deformability and conformational preferences of particular DNA sequences.[22] Evidently, natural DNA-binding proteins exploit both strategies to achieve high affinity and selectivity. Unfortunately, artificial systems are still much more rudimentary, and rational implementation of indirect readout in designed systems is still extremely complex.

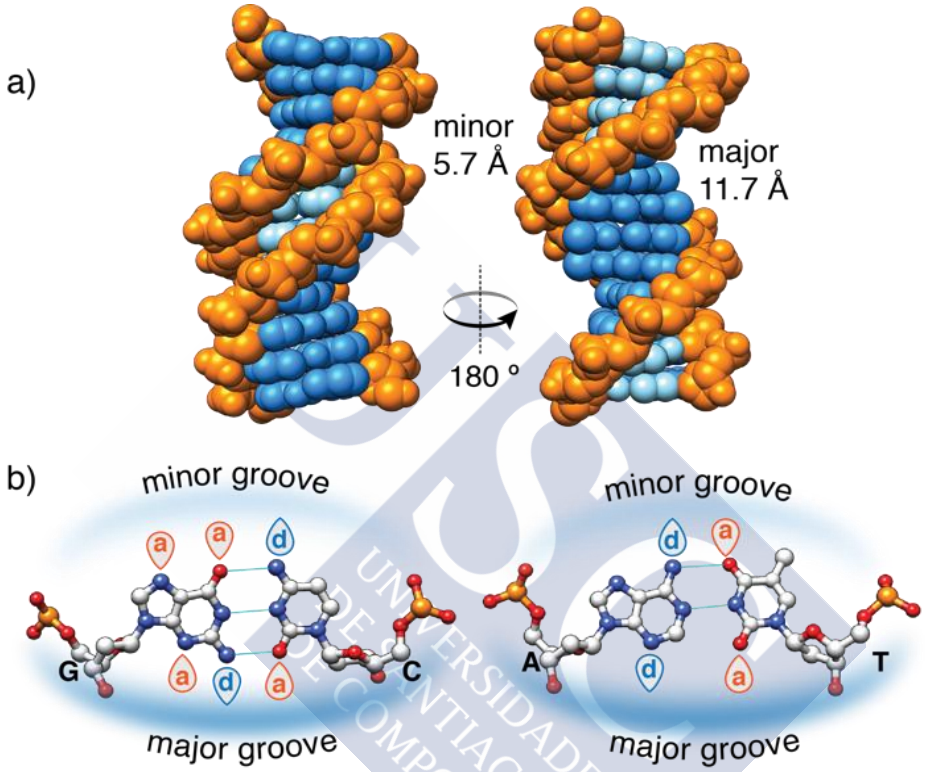


Figure 1. a) Classic Dickerson-Drew B-DNA dodecamer (PDB: 4C64) highlighting the major structural elements: sugar-phosphate backbone in orange, exposed bp edges in the major groove in dark blue, and exposed edges of the bps in the minor groove in light blue. b) structure of the G·C and A·T bps showing the pattern of hydrogen bond donors (d) and acceptors (a).[23]

Recognition of specific DNA sequences in eukaryotes not only relies on the readout mechanisms that have been just outlined that mediate the interactions between the DNA and each protein, but is also orchestrated through specific protein-protein interactions that lead to the formation of higher order assemblies of multiple TFs that cooperatively bind to the DNA. This allows the integration

[22] a) I. Haq, J. Ladbury, *J. Mol. Recognit*. **2000**, *13*, 188–197; b) J. B. Chaires, *Biopolymers* **1997**, *44*, 201–215.

[23] L. Lercher, M. A. McDonough, A. H. El-Sagheer, A. Thalhammer, S. Kriaucionis, T. Brown, C. J. Schofield, *Chem. Commun*. **2014**, *50*, 1794–1796.

of multiple signaling pathways into complex networks of TFs,[24] and also expands the number of unique DNA binding sites that can be addressed with a given number of TFs, which can combine in different ways to bind to composite sequences, thus, allowing the complex spatio-temporal control of the >30,000 human genes with a limited set of 2,000-3,000 TFs.[25] Despite the relevance of TF cooperativity, the complexities involved in the implementation of such effects have hampered the design of artificial systems that display some level of cooperativity in their DNA binding.

DNA-binding peptides

GCN4, the Jack of all trades, master of designed DNA-binding peptides

Transcription factors are grouped in families according to the fold of their DNA binding domains.[26] The GCN4 is an archetypical member of the Basic Leucine Zipper (bZIP) family of TFs. As all bZIP TFs, GCN4 binds to its target ATF/CREB (5′-ATGAC GTCAT-3′) or AP1/GCRE (5′-ATGA(c/g)TCAT-3′) sites as a leucine zipper-mediated dimer of uninterrupted α-helices. The specific contacts with the edges of the bases exposed in the major groove of the DNA take place through the N-terminal basic region (br), which folds into an α-helix upon DNA binding (Figure 2).[27] Importantly, it has been shown that monomeric bZIP TFs typically display low DNA binding affinity in the μM range,[28] and that dimerization is required for DNA binding by bZIP proteins, because of the high entropic cost involved in the folding of the basic region into the α-helical conformation.[29] Thus bZIP TFs behave as intrinsically disordered proteins that are largely unstructured under physiological conditions and display folding coupled DNA binding. This allows bZIP proteins to recognize their target DNA sites with exquisite selectivity while still forming relatively weak, and readily reversible complexes—in contrast with rigid molecules, in which stability of

[24] a) S. Neph, A. B. Stergachis, A. Reynolds, R. Sandstrom, E. Borenstein, J. A. Stamatoyannopoulos, *Cell* **2012**, *150*, 1274–1286; b) T. Ravasi, H. Suzuki, C. V. Cannistraci, S. Katayama, V. B. Bajic, K. Tan, A. Akalin, S. Schmeier, M. Kanamori-Katayama, N. Bertin, et al., *Cell* **2010**, *140*, 744–752.

[25] a) L. Chen, *Curr. Opin. Struct. Biol*. **1999**, *9*, 48–55; b) A. Reményi, H. R. Schöler, M. Wilmanns, *Nat. Struct. Mol. Biol*. **2004**, *11*, 812–815.

[26] C. W. Garvie, C. Wolberger, *Mol. Cell* **2001**, *8*, 937–946.

[27] a) T. E. Ellenberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **1992**, *71*, 1223–1237; b) W. Keller, P. König, T. J. Richmond, *J. Mol. Biol*. **1995**, *254*, 657–667.

[28] X. Wang, W. Cao, A. Cao, L. Lai, *Biophys. J.* **2003**, *84*, 1867–1875.

[29] a) C. Park, J. L. Campbell, W. A. Goddard, *J. Am. Chem. Soc*. **1996**, *118*, 4235–4239; b) M. Zhang, B. Wu, H. Zhao, J. W. Taylor, *J. Pept. Sci*. **2002**, *8*, 125–136.

their complexes usually correlates with their specificity—and allows increased rates of macromolecular associations within complex interaction networks.[30]

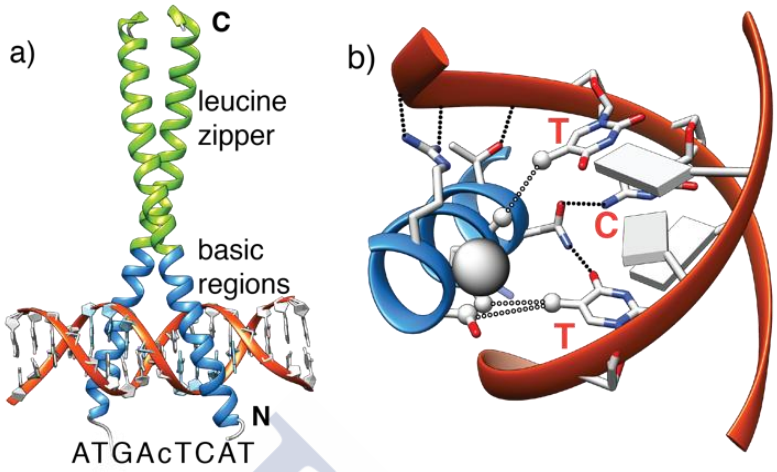


Figure 2. a) Structure of the DNA binding domain of the GCN4 TF bound to the AP1/GCRE site, highlighting the main structural elements and the key residues for specific DNA recognition. The leucine zipper in light green, and the br in blue. b) Details of the specific interactions between the GCN4 br with the AP1/GCRE half-site. Hydrophobic contacts (Ala238, Ala239, and Ser242) with white dot lines, polar/electrostatic contacts (Asn235, and Arg243) with black dot lines. Bases that do not participate in specific contacts are outlined as slabs for clarity.

Artificial GCN4 peptide dimers

The structural simplicity of bZIP TFs, added to the wealth of functional, biophysical and structural information about GCN4, made it into the preferred platform for the development of peptide-based DNA binders. A key development was reported in 1990 by the group of Peter S. Kim, who demonstrated in a seminal paper that the complete leucine zipper region of the GCN4 DNA binding domain could be replaced by a disulfide bond, and that the minimized 34-residue peptide corresponding largely to the GCN4 between residues 222 and 253 (plus a short Gly-Gly-Cys linker) could bind to the natural target site CREB with high (nM) affinity at 4 ºC in the form of a disulfide dimer (Figure 3a).[31] Additional studies reported two years later helped to identify the sequence of the minimal GCN4 basic region fragment that displays specific DNA binding in the form of disulfide dimer. Thus, a 23-residue peptide featuring a C-terminal Gly-Gly-Cys linker (Ac-DPAALKRARNTEAARRSRARKLQ-GGC) binds to both ATF/CREB and

[30] a) M. Miller, *Curr. Protein Pept. Sci*. **2009**, *10*, 244–269; b) J. Habchi, P. Tompa, S. Longhi, V. N. Uversky, *Chem. Rev.* **2014**, *114*, 6561–6588.
[31] R. V. Talanian, C. J. McKnight, P. S. Kim, *Science* **1990**, *249*, 769–771.

AP1/GCRE sites in a sequence-specific manner with high affinity in its oxidized form as a dimer.[32]

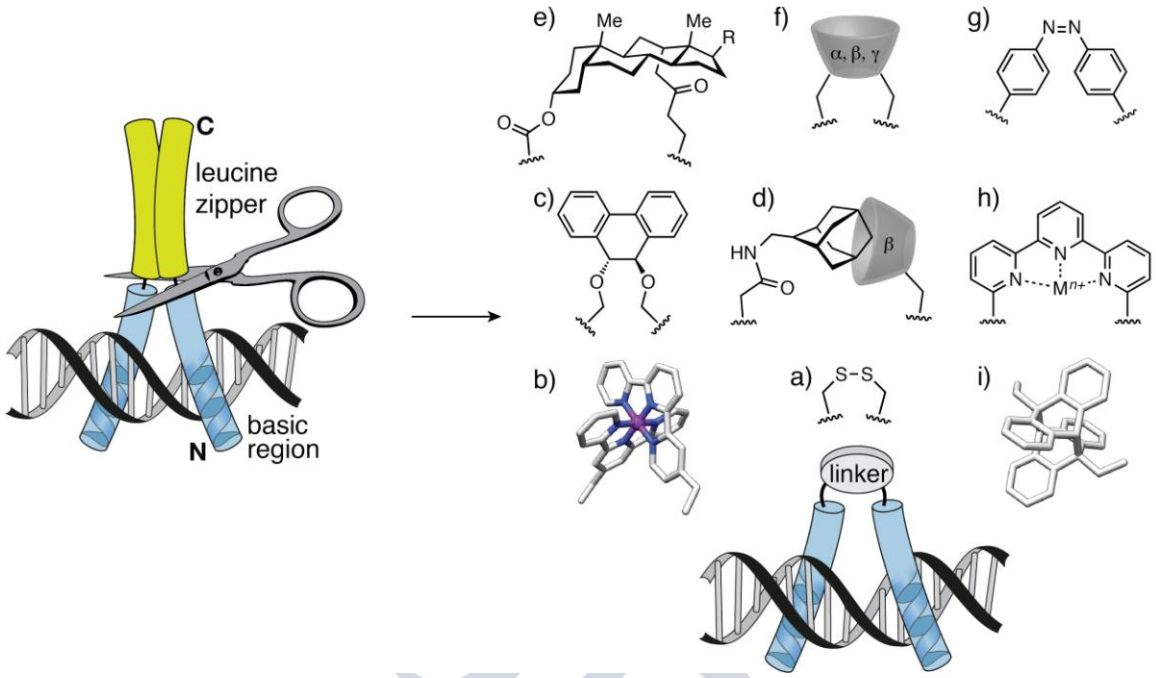


Figure 3. Schematic representation of the GCN4 dimer showing the helices as tubes, and the strategy to generate artificial GCN4 dimers by removing the leucine zipper and adding specific linkers between the br fragments. a) P. Kim's disulfide dimer,[27,28] b) A. Schepartz's transition metal coordination compound linker,[29] c) T. Morii's chiral linker derived from trans-9,10-dihydrophenanthrene-9,10-diol,[31] d) T. Morii's supramolecular linker based on a β-cyclodextrin/adamantane complex,[33,34] e) A. Madder's steroid linker,[36,37] f) A. Madder's cyclodextrin dimer,[35] g) J. L. Mascareña's cis/trans azobenzene photoswitch linker,[39] h) A. Peacock's metal switch linker,[30] i) A. Peacock's photocontrolled anthracene dimer.[53]

Following Peter Kim's work, a number of other groups have reported over the years the replacement of the leucine zipper with other functionalities. The group of Alanna Schepartz reported in 1993 the use of a set of transition metal complexes as linkers to systematically explore the effect of the geometry of the linker—and thus, of the relative geometry of the attached basic regions—on the DNA binding properties of the resulting dimers (Figure 3b). Unsurprisingly, they found that dimerization was not sufficient for DNA binding, and that the geometry of the metal complexes influenced not only the binding affinity of the dimers, but also the sequence selectivity.[33] The effect of the geometry in metal complexes was also exploited by the group of A. Peacock, who in 2014 reported

[32] R. V. Talanian, C. J. McKnight, R. Rutkowski, P. S. Kim, *Biochemistry* **1992**, *31*, 6871–6875.
[33] B. Cuenoud, A. Schepartz, *Science* **1993**, *259*, 510–513.

a dimer of the GCN4 basic region with a 2,2′:6′,2″-terpyridine linker. This linker would undergo a conformational change upon metal ion coordination, which in turn would affect the relative orientation of the basic regions, and influence the DNA binding of the dimer. CD and UV-vis spectroscopy demonstrated that coordination od this metal-chelating GCN4 dimer to Cu(II) or Zn(II) ions promoted DNA binding (Figure 3h).[34]

The influence of geometric effects on the DNA binding properties of GCN4 dimers were also studied by T. Morii and Y. Sugiura who, based on earlier studies using a 14-residue DNA-binding peptide derived from the basic region of the helix-loop-helix transcription factor MyoD,[35] synthesized various GCN4 dimers connected by C_2-symmetric, trans-9,10-dihydro phenanthrene-9,10-diol linkers with different chiralities (Figure 3c). In this case, the authors found that the stability of the different complexes was unaffected by the chirality of the linkers, which was consistent with the small effect of the chirality on the relative orientation of the attached basic regions. Interestingly, the authors also reported dimers of the GCN4 br through the N-terminus, which recognized altered palindromic DNA sequences (5′-TCATC GATGA-3′) in which the polarity of each half-site of the parent ATF/CREB) site was reversed.[36]

The group of T. Morii was the first to report the use of a non-covalent linker between GCN4 basic regions; for this, they synthesized two peptides by alkylation of a C-terminal Cys attached to the core GCN4 br peptide with N-(bromoacetyl)-1-adamantanemethylamine and mono-6-deoxy-6-iodo- β-cyclodextrin. Both peptides form a heterodimer, mediated by formation of an inclusion complex between β-cyclodextrin, and an adamantyl group, which specifically binds to the ATF/CREB (5′-ATGAC GTCAT-3′) site. The formation of the complex was inhibited by the addition of free β-cyclodextrin, once again demonstrating that dimerization of the GCN4 basic regions is required for specific DNA binding (Figure 3d).[37] Interestingly, their design relies on the high stability of the complex between adamantane derivatives and β-cyclodextrins ($K_D \approx 1.3$ μM), and subsequent studies showed that efficient and cooperative DNA binding was still possible if the original adamantane was replaced with norbornyl, K_D(Nb/β-CD) ≈ 3.1 μM, or noradamantyl groups,

[34] E. Oheix, A. F. A. Peacock, *Chem. Eur. J.* **2014**, *20*, 2829–2839.
[35] T. Morii, M. Simomura, S. Morimoto, I. Saito, *J. Am. Chem. Soc.* **1993**, *115*, 1150–1151.
[36] T. Morii, Y. Saimei, M. Okagami, K. Makino, Y. Sugiura, *J. Am. Chem. Soc.* **1997**, *119*, 3649–3655.
[37] M. Ueno, A. Murakami, K. Makino, T. Morii, *J. Am. Chem. Soc.* **1993**, *115*, 12575–12576.

K_D(NrA/β-CD) ≈ 2.7 μM), but no DNA recognition was observed with the cyclohexyl group, which displays a much lower binding affinity for β-cyclodextrin with a K_D(Ch/β-CD) ≈ 14 μM.[38]

A recent paper by A. Madder also makes use of (α, β, γ)-cyclodextrins for dimerization (Figure 3f). In their case, the cyclodextrins were not used as receptors, but as scaffolds for covalent attachment of the two GCN4 basic regions. The dimers were readily synthesized by a straightforward copper-catalyzed azide/alkyne cycloaddition (CuAAC) between the GCN4 br peptide bearing a C-terminal propargylglycine, and the corresponding diazido cyclodextrins. Curiously, while the α-cyclodextrin and β-cyclodextrin GCN4 br dimers display rather similar binding affinities for the ATF/CREB site with K_D(α-CD/DNA) = 50 ± 20 nM and K_D(β-CD/DNA) = 30 ± 20 nM, the γ-cyclodextrin analog binds with weaker affinity to the same DNA with a K_D(γ-CD/DNA) = 100 ± 60 nM, and appears to give rise to non-specific complexes in Electrophoretic Mobility Shift Assays (EMSA), suggesting that the distance between the two GCN4 basic regions in the γ-cyclodextrin dimer is not optimal for their simultaneous insertion as required in the specific complex.[39] Madder's group has also reported the application of steroid scaffolds for the homodimerization of GCN4 basic regions,[40] as well as for the heterodimerization of the basic regions of related c-Myc/Max basic Helix-Loop-Helix transcription factors, a class of TFs related to the bZIP family, containing an additional loop between the leucine zipper and the basic region.[41] The use of a steroidal cholic acid moiety as dimerizer element was justified because it provides a rigid, and synthetically-accessible scaffold to attach the peptides with reported benefits for improving the cell uptake of the conjugates (Figure 3e).

In addition to bZIP dimers that replace the leucine zipper dimerization with other functionalities, in 2012 J. L. Mascareñas reported a derivative of an heterodimeric cFos/cJun complex in which the basic region of the cFos bZIP TF

[38] Y. Aizawa, Y. Sugiura, M. Ueno, Y. Mori, K. Imoto, K. Makino, T. Morii, *Biochemistry* **1999**, *38*, 4008–4017.
[39] Y. Ruiz García, J. Zelenka, Y. V. Pabon, A. Iyer, M. Buděšínský, T. Kraus, C. I. E. Smith, A. Madder, *Org. Biomol. Chem*. **2015**, *13*, 5273–5278.
[40] L. L. G. Carrette, T. Morii, A. Madder, *Eur. J. Org. Chem*. **2014**, *2014*, 2883–2891.
[41] a) D. Verzele, A. Madder, *Eur. J. Org. Chem*. **2013**, 2013, 673–687; b) Y. R. García, Y. Vladimir Pabon-Martinez, C. I. Edvard Smith, A. Madder, *Chem. Commun*. **2017**, 53, 6653–6656.

was replaced by a small DNA-binding organic molecule.[42] Similarly to the role of cFos in the natural complex, the bisbenzamidine hybrid of the Fos leucine zipper (featuring residues from Arg→Trp^{157} to Leu^{179}) stabilized the cJun/DNA interaction but, in contrast with the natural cFos TF, which is bound to the DNA major groove, the cFos leucine zipper domain was delivered from the adjacent DNA minor groove. Furthermore, the dimer between cJun and the cFos-bisbenzamidine conjugate targeted a the composite DNA sequence different from the ATF/CREB binding site targeted by the natural cJun/cFos heterodimer (5′-ATGAC G AAATTT-3′), which contained the natural ATF/CREB half-site (ATGAC) next to the A·T-rich site preferred by the bisbenzamidine with an extra G as spacer between both recognition elements (Figure 4).

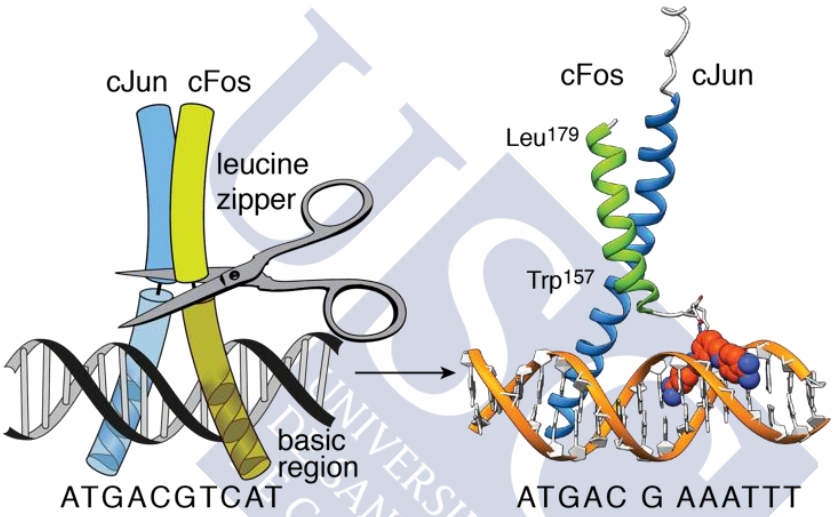


Figure 4. Left: Schematic representation of the cFos/cJun/DNA complex. Right: proposed interaction between the cFos-bisbenzamidine hybrid and cJun, binding as a heterodimer to a composite DNA site containing the consensus recognition sites for cJun and the bisbenzamidine.

Stimuli-responsive DNA-binding peptides

Light-responsive systems

A relevant subset of GCN4 dimers are those whose DNA binding can be modulated by light. In the year 2000 the group of J. L. Mascareñas reported the first artificial DNA binding peptide that displayed light-induced DNA binding.[43] Light responsiveness was achieved by incorporating an azobenzene unit, a well-

[42] M. I. Sánchez, O. Vázquez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.* **2012**, *3*, 2383.
[43] A. M. Caamaño, M. E. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed Engl.* **2000**, *112*, 3234–3237.

known photoisomerizable group,[44] in the linker between two GCN4 basic region peptides. The design takes advantage of the large geometrical change experienced by the azobenzene upon trans→cis isomerization, so that while the trans configuration forces the two GCN4 br away from each other, resulting in an unfavorable geometry for simultaneous interaction of both peptides in their target sites, the cis- conformation presents both basic regions in the appropriate orientation for both peptides to interact in their respective half-sites at the same time (Figure 3g and Figure 5). EMSA assays demonstrated that the cis isomer binds to the target ATF/CREB site with 60-70 times more affinity than the trans isomer. Curiously, the trans→cis isomerization can be carried out in the presence of the DNA, but in this case the isomerization is about eight times slower than in the absence of DNA, and leads to a slightly lower proportion of the cis isomer in the photostationary state.

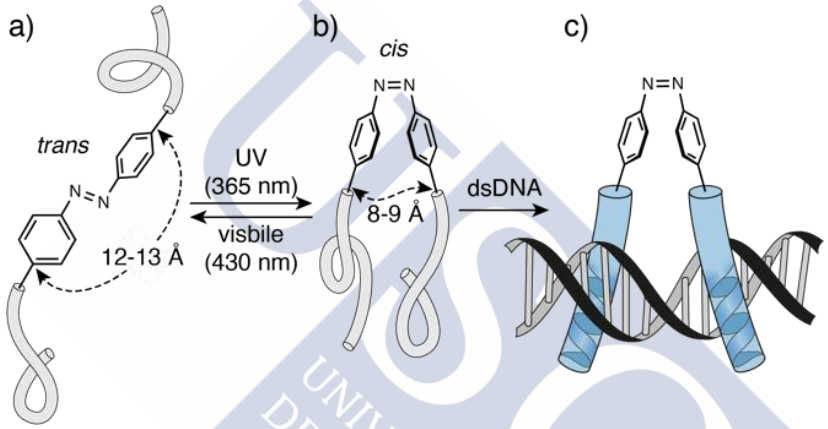


Figure 5. The DNA binding properties of GCN4 dimers featuring an azobenzene linker can be modulated by light. Thus, the opposite orientation, and large separation between the peptides in the trans azobenzene isomer does not support high-affinity DNA binding; a) irradiation with UV light induces the isomerization of the azobenzene unit, and b) the resulting cis GCN4 dimer presents the right geometry for high-affinity binding as shown in c). Unfolded peptides are represented with light-grey thin tubes.

Before presenting other examples of light-modulated systems we must introduce another of the preferred model systems in artificial DNA binding peptides, the Homedomain (HD) proteins. HDs are highly conserved TFs that play key roles in eukaryotic development. The HD DNA binding domain is folded into a helix-turn-helix motif, with an extended minor groove-binding N-terminal arm.[45] Most of the specific contacts with the DNA are made by helix 3 of the protein (h_3), which, in the complex with DNA, is inserted into the DNA major groove.

[44] a) R. J. Mart, R. K. Allemann, *Chem. Commun*. **2016**, *52*, 12262–12277; b) H. M. D. Bandara, S. C. Burdette, *Chem. Soc. Rev*. **2012**, *41*, 1809–1825.
[45] W. J. Gehring, Y. Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A. F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, K. Wüthrich, *Cell* **1994**, *78*, 211–223.

As in the case of the GCN4 and bZIP proteins, the short DNA-contacting helix h_3 isolated from the rest of the protein fails to fold into the α-helical conformation and shows negligible DNA binding affinity. The engrailed HD Q50K is an altered-specificity Gln^{50}→Lys variant of the engrailed HD that preferentially binds to the QRE site 5′-TAATCC-3′ (while the wild-type engrailed TF binds to the ERE sequence, 5′-TAATTA-3′). Furthermore, the engrailed Q50K displays higher affinity for the QRE site ($K_D \approx 8.8$ pM) than the wild-type engrailed for its ERE target sequence ($K_D \approx 79$ pM).[46]

After the initial work with azobenzene GCN4 dimers by J. L. Mascareñas, the group of G. A. Woolley reported in the year 2005 an alternative strategy for obtaining light-controlled DNA-binding peptides incorporating azobenzene units within DNA basic region (Figure 6).[47] Following their work using azobenzenes for photocontrolling the α-helical content in short peptides, their first design consisted on a 18-residue peptide fragment from the recognition helix h_3 of the Q50K engrailed homeodomain (Ac-EAQCKIAAKNARAKCKKA, mutated residues in bold typeface), in which residues Ile^{45} and Ile^{56} (at i, and i+11 positions) on the outer face of the h_3 α-helix, were mutated to Cys for introducing the 3,3′-bis(sulfo)-4,4′-bis(chloroacetamido)azobenzene cross-linker. Additionally, a number of residues that do not make direct contacts with the DNA were also replaced with Ala in order to increase the helical propensity of the peptide, and to avoid undesired steric clashes between bulky side chains with the azobenzene staple. The thermodynamically-stable trans azobenzene displayed significant α-helical content and formed stable complexes ($K_D \approx 7.5$ nM) with the QRE target sequence (5′-TAATCC-3′). Irradiation with UV light (360 nm) induced the trans→cis isomerization of the azobenzene, which significantly destabilized the α-helical conformation of h_3, and resulted in decreased affinity for the DNA ($K_D \approx 140$ nM).

[46] L. Tucker-Kellogg, M. A. Rould, K. A. Chambers, S. E. Ades, R. T. Sauer, C. O. Pabo, *Structure* **1997**, *5*, 1047–1054.

[47] a) J. R. Kumita, O. S. Smart, G. A. Woolley, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3803–3808; b) G. A. Woolley, *Acc. Chem. Res.* 2005, *38*, 486–493; c) J. R. Kumita, D. G. Flint, G. A. Woolley, O. S. Smart, *Faraday Discuss.* **2003**, *122*, 89–103.

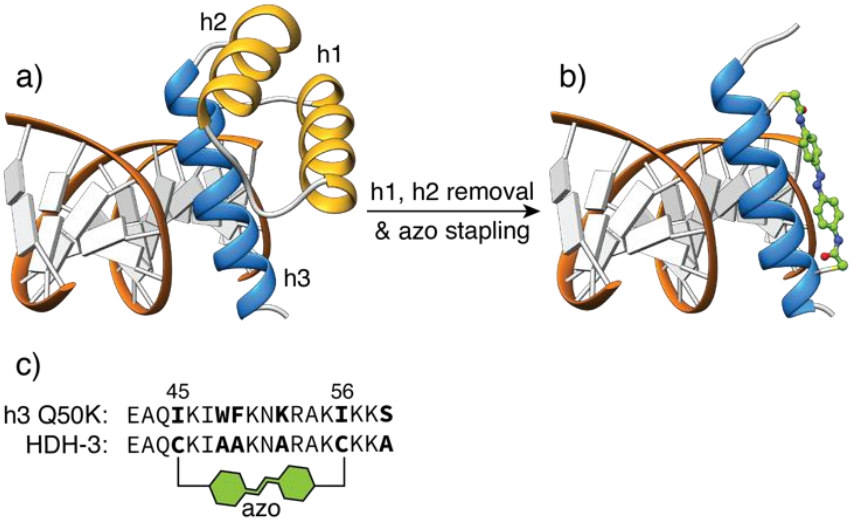


Figure 6. Miniaturized HD Q50K light-modulated DNA-binding peptide. a) Structure of the Q50K engrailed HD highlighting the DNA contacting helix (h3) in blue, and the accessory helices h1 and h2 in yellow (PDB: 2HDD). b) The light-responsive peptide contains of most of the HD Q50K sequence, and two Cys residues at positions 45 and 56 (i, i+11) for stapling with the azobenzene unit. The trans azobenzene stabilizes the DNA-binding α-helical conformation, while irradiation yields a low-affinity DNA configuration.

In the following years the group of G. A. Woolley refined their designs,[48] and reported modified azobenzene switches with improved optical properties for in vivo applications,[49] and 2011 they demonstrated for the first time the application of artificial peptides for the photocontrol of protein expression in living cells.[50] The authors focused their attention on the oncoproteins Fos and Jun, which are also members of the bZIP family of TFs that bind to the AP1/GCRE site as heterodimeric complexes. Their design was a modified version of a dominant-negative (DN) peptide reported by C. Vinson,[51] which contained an optimized Fos leucine zipper dimerization domain for increased affinity for the Jun protein,[52] and replaced the normal basic region critical for DNA binding with an acidic extension that also stabilized its complex with Jun, while at the same time prevented the resulting heterodimer from binding to the DNA. This DN peptide (AFosW) was modified with a 3,3′-bis(sulfonato)-4,4′-bis(chloroacetamido)

[48] G. A. Woolley, A. S. I. Jaikaran, M. Berezovski, J. P. Calarco, S. N. Krylov, O. S. Smart, J. R. Kumita, *Biochemistry* **2006**, *45*, 6075–6084.
[49] A. A. Beharry, O. Sadovski, G. A. Woolley, *J. Am. Chem. Soc.* **2011**, *133*, 19684–19687.
[50] F. Zhang, K. A. Timm, K. M. Arndt, G. A. Woolley, *Angew. Chem. Int. Ed Engl*. **2010**, *49*, 3943–3946.
[51] M. Olive, D. Krylov, D. R. Echlin, K. Gardner, E. Taparowsky, C. Vinson, *J. Biol. Chem*. **1997**, *272*, 18586–18594.
[52] J. M. Mason, M. A. Schmitz, K. M. Müller, K. M. Arndt, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 8989–8994.

azobenzene cross-link between two Cys residues replacing the Asp[174] and Thr[181] residues in the wild-type Fos leucine zipper sequence. The resulting photoisomerizable peptide (XAFosW) formed stable heterodimers with Jun when the azo group was in the trans configuration, and isomerization of the azobenzene crosslink to the cis configuration by irradiation with light at 365 nm reduced the stability of the XAFosW/Jun complex by about 10-fold. The in vivo effect of the XAFosW photoswitch was quantitatively studied by measuring the ratios of the activity between a luciferase reporter under the control of an AP1/GCRE promoter and a β-galactosidase reporter under the control of a constitutive rous sarcoma virus (RSV) promoter (pRSV-Gal). Thus, XAFosW showed concentration-dependent inhibition luciferase activity and, more importantly, a decrease of as much as 40% in AP-1 activity upon photo-irradiation. Subsequent work has focused on fine tuning these designs, for example by identifying the optimal position for introducing crosslinks,[53] as well as on developing encodable protein chimeras with photoactive yellow protein.[54]

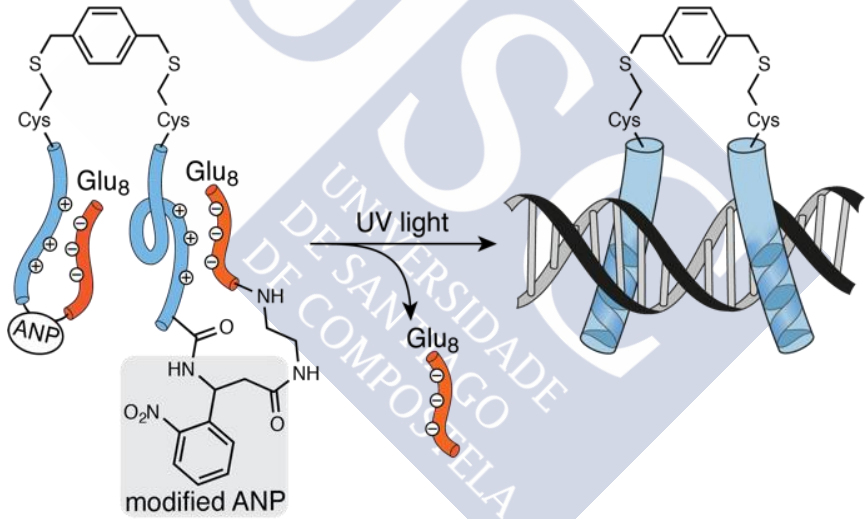


Figure 7. Strategy for electrostatic caging of the DNA binding domain of the GCN4 transcription factor. Oligoglutamic (Glu_8) domains were attached to the N-terminus of the GCN4 br through an orthogonal, photolabile ANP linker; after irradiation, the photoreleased br dimer binds to the DNA.

In 2012 our own group reported a different approach to obtaining light-activated DNA binding peptides.[55] Considering that the formation of thermodynamically

[53] A. M. Ali, M. W. Forbes, G. A. Woolley, *ChemBioChem* **2015**, *16*, 1757–1763.
[54] a) A. M. Ali, J. M. Reis, Y. Xia, A. J. Rashid, V. Mercaldo, B. J. Walters, K. E. Brechun, V. Borisenko, S. A. Josselyn, J. Karanicolas, et al., *Chem. Biol*. **2015**, *22*, 1531–1539.
[55] A. Jiménez-Balsa, E. Pazos, B. Martínez-Albardonedo, J. L. Mascareñas, M. E. Vázquez, *Angew. Chem. Int. Ed Engl*. **2012**, *51*, 8825–8829.

stable GCN4/DNA complexes relies on non-specific electrostatic interactions between positively-charged basic residues in the GCN4 basic region with the negatively-charged DNA phosphate backbone,[56] we envisaged that introducing repulsive interactions with short polyanionic tails would suppresses the DNA interaction. Furthermore, if the polyanionic tails were attached through a photolabile linker, then the interaction between the photoreleased GCN4 basic region and the target ATF/CREB DNA could be restored by irradiation with light. Thus, P. S. Kim's GCN4 basic region fragment was synthesized with an acidic oligo-glutamic domain (Glu_8) connected through a photolabile 3-amino-3-(2-nitrophenyl)-propionic acid (ANP) group. The resulting oligozwitterionic basic region containing a short linker with a C-terminal Cys residue (Glu_8-(ANP)- DPAALKRARNTEAARRSRAR KLQ-GGC) was dimerized by direct alkylation with 1,4-bis(bromomethyl)benzene, to yield the desired photoactivatable GCN4 derivative (Figure 7). As expected, the electrostatically-compromised dimer did not display significant DNA binding affinity for its target site, and irradiation with UV light irreversibly restored the DNA binding of the photoreleased GCN4 br dimer, as demonstrated by the appearance of a slow-migrating band in gel shift experiments. Circular dichroism assays also confirmed the folding of the photoreleased basic region in the presence of an oligonucleotide with the ATF/CREB target sequence.

An interesting interplay between the DNA binding of GCN4 peptides and anthracene photodimerization was described by the group of A. Peacock in 2016, when they studied monomeric GCN4 basic region peptides modified with anthracene units on their C-terminus.[57] While anthracene intercalation in the presence of non-target DNA sequence proved problematic, circular dichroism experiments showed a significant increase in the alpha helical content of the peptide upon irradiation in the presence of the target ATF/CREB (5′-ATGAC GTCAT-3′) site. Furthermore, they found that anthracene photodimerization took place only in the presence of such target site, while dimerization of the anthracene-basic region peptide was not observed when irradiation was made in the presence of half-site or random DNAs. As expected from the previous work with GCN4 dimers, the resulting photocrosslinked dimer displayed stronger DNA binding than the monomeric peptides.

[56] a) T. Härd, T. Lundbäck, Biophys. Chem. 1996, 62, 121–139; b) J. K. Strauss-Soukup, L. J. Maher 3rd, *Biochemistry* **1998**, *37*, 1060–1066.
[57] G. A. Bullen, J. H. R. Tucker, A. F. A. Peacock, *Chem. Commun.* **2015**, *51*, 8130–8133.

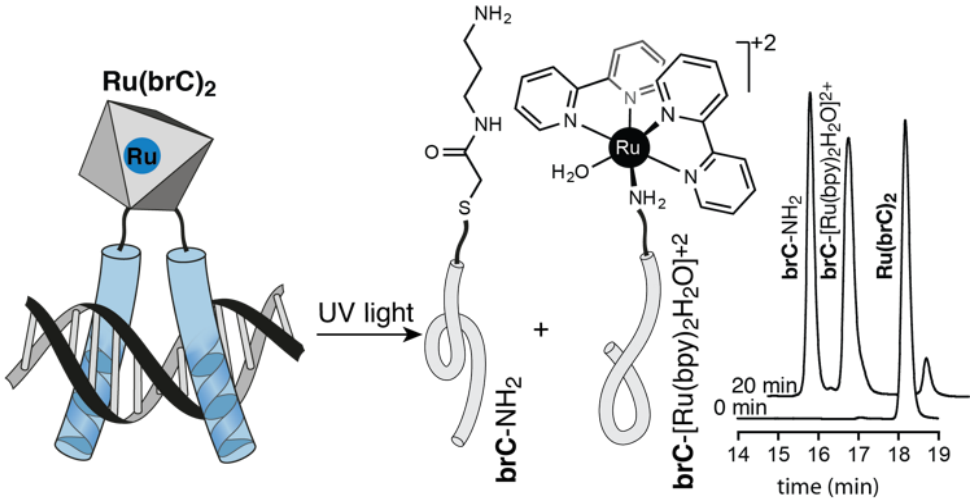


Figure 8. Photocleavable (2,2′-bipyridyne)Ru(II) linkers in bZIP basic region dimers (either GCN4 homodimers or C/EBP-GCN4 heterodimers). The high-affinity, specific DNA complex can be disassembled by irradiation with long-wavelength light. The different species resulting from the asymmetric photolysis can be identified by HPLC-MS (right).

One of the major problems associated with the application of photocontrolled peptides in vivo is the short wavelengths at which the photoswitches and caging groups respond, typically below 400 nm, which have poor tissue penetration, and are readily absorbed by other biomolecules. In this context, photolabile bisbipyridyl ruthenium(II) complexes have been explored as alternative photoactive components due to their advantageous spectroscopic properties, such as long wavelength of photolysis and high uncaging quantum yields. Thus, we reported in 2014 the use of ruthenium bipyridyl complexes as photocleavable dimerizers of cysteine-containing transcription factor fragments. More specifically, we reported the synthesis of a generally applicable cis-bromoalkyl-(2,2′-bipyridine)Ru(II) biselectrophilic reagent, which was derivatized to obtain homodimeric GCN4 br dimers, similar to those previously described in the 90s by A. Schepartz, as well as C/EBP-GCN4 heterodimers. In both cases the dimeric constructs displayed good DNA binding affinity and selectivity for their respective target site. Moreover, their DNA binding could be suppressed by irradiation at 455 nm with a LED source and photolysis of the ruthenium dimerizer (Figure 8).[58]

Ruthenium compounds have been used as photolabile protecting groups to control the DNA binding of small peptides.[59] In 2015 we reported the synthesis

[58] J. Mosquera, M. I. Sánchez, M. Eugenio Vázquez, J. L. Mascareñas, *Chem. Commun*. **2014**, *50*, 10975–10978.

[59] J. Mosquera, M. I. Sánchez, J. L. Mascareñas, M. Eugenio Vázquez, *Chem. Commun*. **2015**, *51*, 5501–5504.

of a light-sensitive histidine building block for Fmoc/tBu solid-phase peptide synthesis in which the imidazole side chain is protected with a ruthenium complex. This building block was used for the synthesis of a photoactivatable Arg-Gly-His tripeptide (RGH), which has been shown to be an efficient metal-chelating sequence ($K_D < 10^{-16}$ M), endowed with DNA binding and endonuclease properties in the presence of Ni(II) ions and oxidizing agents.[60] We demonstrated that the caged His analog was unable to coordinate the Ni(II) required form the catalytic metallopeptide. Moreover, we also showed that DNA binding and nuclease activity could be recovered upon irradiation at 455 nm, and uncaging of the His residue (Figure 9).

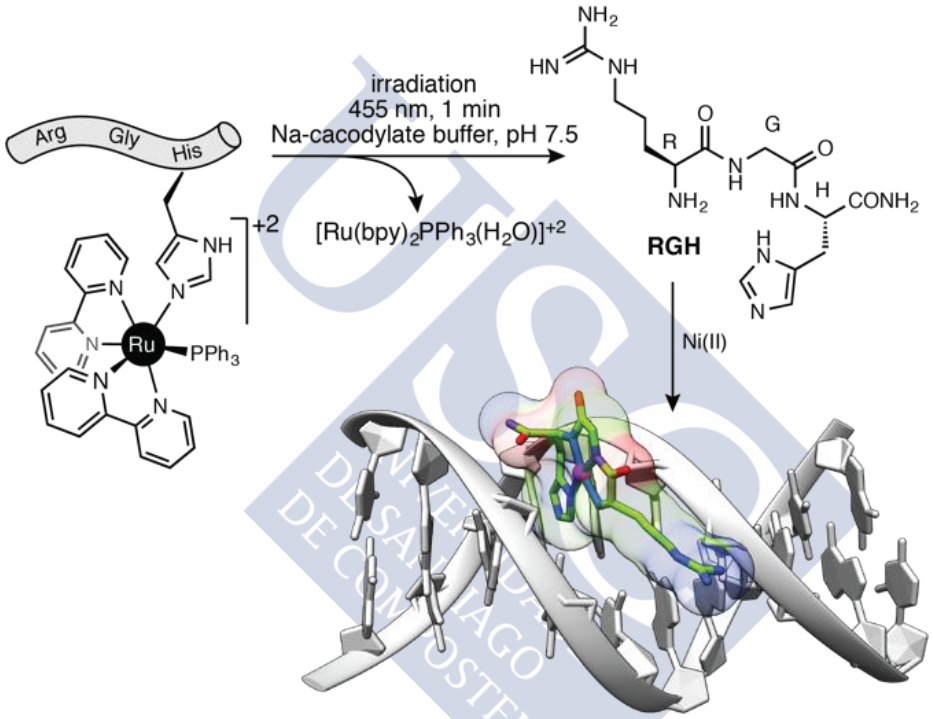


Figure 9. Caged RGH peptide does not bind to Ni(II). Irradiation with red ligt at 455 releases the metal-chelating peptide, which folds and binds into the DNA minor groove, where it can exert its nuclease activity.

Metal coordination and dynamic DNA-binding peptides

In contrast with the static designs which represent the majority of DNA-binding peptides reported until now, one of the first steps into the development of stimuli-responsive peptides derived from transcription factors was reported by Y. Sugiura in 2004.[61] In this work, the Fos leucine zipper is modified with a pair of metal-chelating iminodiacetic acid derivatives of lysine (Ida) at positions i,

[60] a) Q. Liang, P. D. Eason, E. C. Long, *J. Am. Chem. Soc*. **1995**, *117*, 9625–9631.
[61] S. Futaki, T. Kiwada, Y. Sugiura, *J. Am. Chem. Soc*. **2004**, *126*, 15762–15769.

and i+2. Incubation with Fe(III), and coordination of this metal ion, destabilizes the alpha helical conformation of the Fos(Ida)$_2$ peptide, thus preventing the formation of the heterodimer with the Jun leucine zipper. Unfortunately, the authors did not pursue this concept for modulating the DNA binding of the Fos-Jun network of homo- and heterodimers. In a related work, a successful DNA binding switch in response to a metal ion was reported by the group of Futaki in 2009, who expressed a recombinant GCN4 DNA binding domain (56-mer peptide) featuring two pairs of Cys residues at relative i and i+2 positions in the leucine zipper domain. These residues were selectively derivatized using N-(2-tosylthioethyl)iminodiacetic acid to turn them into metal-chelating residues. The resulting tetra-Ida-modified GCN4 DNA binding domain displayed significant helical structure, and high DNA binding affinity for the AP1/GCRE site (K_D = 22 ± 3.0 nM) similar to that of the natural GCN4 dimer (K_D = 15 ± 2.6 nM). Addition of excess of Co(II) to this metal-binding peptide induced a 33% decrease in its helical content, and also a drastic reduction in its DNA binding affinity (K_D > 1.0 μM). As expected, addition of EDTA to the mixture—and sequestering of the Co(II) ions—resulted in the recovery of the initial binding affinity in absence of the metal ion (K_D = 22 ± 1.1 nM). Curiously, the switching effect was not observed in control peptides that contained a single pair of Ida residues in their leucine zipper domains.[62]

The effect to metal ion coordination can be combined with other orthogonal signals, such as redox processes to yield more complex designs that respond to multiple stimuli. This was exemplified in 2013, when we described the synthesis of a GCN4 basic region fragment that contained a redox-sensitive Cys residue on its N-terminus, and a metal-chelating 5,5″-dimethyl-2,2′:6′,2″-terpyridine ligand orthogonally attached to a C-terminal Lys^{231} residue (Cys)br(tpy).[63] Incubation of this peptide with Ni(ClO_4)$_2$ in the presence of the target ATF/CREB site (5'-ATGA CG TCAT-3') gave rise to slow-migrating bands consistent with the formation of a specific peptide-DNA complex between the nickel-mediated dimer of the basic region and the DNA. Under the same conditions, the peptide did not display significant affinity for the inverted sequence (5'-TCAT CG ATGA-3'). In contrast, the N-terminal disulfide dimer formed by oxidation of the peptide with DTNB, 5,5'-dithiobis-(2-nitrobenzoic

[62] Y. Azuma, M. Imanishi, T. Yoshimura, T. Kawabata, S. Futaki, *Angew. Chem. Int. Ed.* **2009**, *48*, 6853–6856.
[63] J. Mosquera, A. Jiménez-Balsa, V. I. Dodero, M. E. Vázquez, J. L. Mascareñas, *Nat. Commun.* **2013**, *4*, 1874.

acid), only showed binding to the inverted site (Figure 10). As expected, reduction of the disulphide dimer by treatment TCEP (tris(2-carboxyethyl) phosphine) recovered the monomeric (Cys)br(tpy) peptide, and promoted the disassembly of the complex with the inverted sequence. Curiously, the disulfide dimer required the presence of Ni(II) ions to bind to its target inverted site.

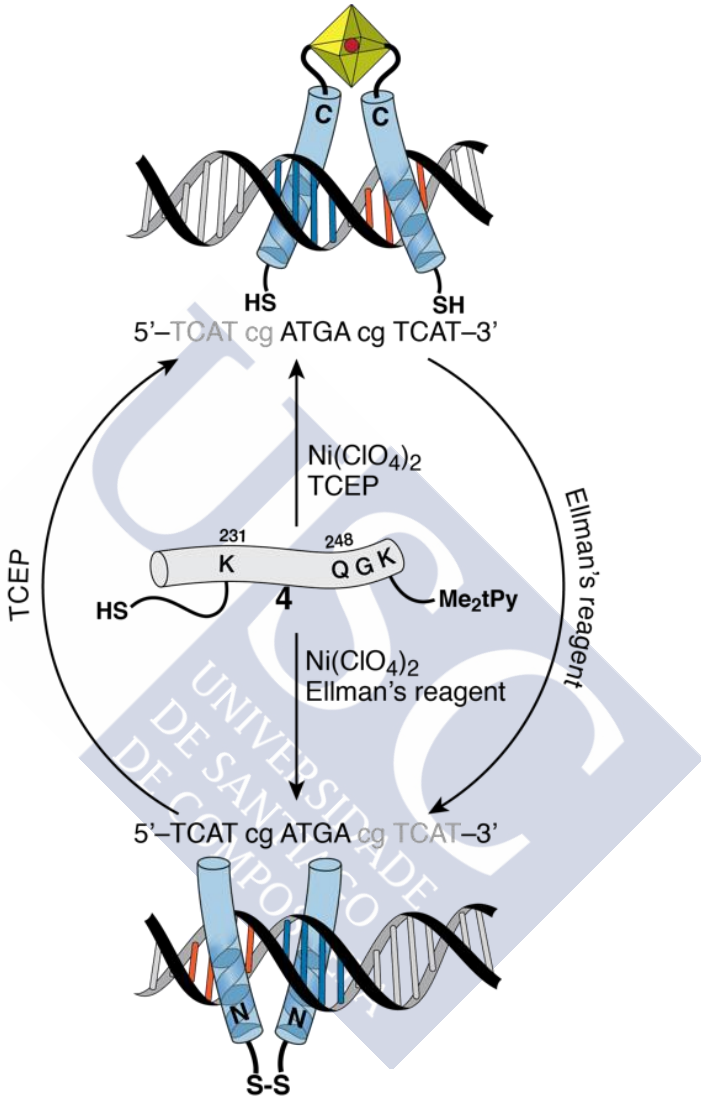


Figure 10. Dynamic DNA binding selection to multiple sites. The reduced monomeric (Cys)br(tpy) peptide in the presence of Ni(ClO4) binds to the consensus DNA (5'-ATGA CG TCAT-3') as a metal-mediated homodimeric complex ($K_D \approx 670$ nM at 4ºC). Oxidation of the N-terminal Cys side chain with DTNB yields the N-terminal disulfide dimer, which binds to the inverted polarity sequence 5'-TCAT CG ATGA-3' in the presemnce of Ni(II) salts ($K_D \approx 758$ nM). The original C-terminal complex can be recovered by reduction with TCEP (tris(2-carboxyethyl)phosphine). In both cases the binding of the peptides for their non-target DNAs were in the μM range).

Artificial GCN4 basic region monomers

Monomeric GCN4 basic regions do not bind to their target DNA half-site with high affinity due to the entropic cost associated to the folding of the peptide chain into the appropriate α-helical conformation, but residue grafting is a powerful strategy to obtain stabilized α-helices and thus obtain high-affinity DNA binding monomeric GCN4 peptides.[64] In this approach, selected residues that create the desired binding epitope are inserted (grafted) into a stable three-dimensional peptide scaffold. The group of A. Schepartz described in 1999 successful application of this strategy to obtaining high-affinity monomeric DNA-binding peptides.[65] Their design involved the introduction of the GCN4 residues mediating DNA binding in the basic region into the α-helix of the avian pancreatic polypeptide (aPP), a stable miniprotein consisting of a single α-helix stabilized by hydrophobic interactions with a type II polyproline helix. The resulting chimera was capable of recognizing the half site of the GCN4 DNA target sequence (5'-ATGA-3') with a K_D of 1.5 nM under physiological ionic strength. However, the mutations introduced in the aPP sequence for achieving DNA binding resulted in the disruption of the aPP hydrophobic core, and reduced structural stability of the construct. Further refinements of this initial binder using phage libraries resulted in a new peptide with extraordinary DNA affinity at 4 ºC (K_D ~ 23 pM), which even retained high affinity at 25 ºC (K_D ~ 1.6 nM).[66] Following the successful implementation of the grafting strategy to the synthesis of simplified versions of a bZIP transcription factor (GCN4), the Schepartz group demonstrated the versatility of this approach by obtaining miniature homeodomain proteins, again by dissecting the key DNA contacting residues and introducing them into the structure of the versatile aPP peptide scaffold.[67]

The group of G. L. Verdine demonstrated in 1995 that it was possible to drive specific DNA recognition by a monomeric GCN4 basic region if the peptide was delivered into the major groove intramolecularly.[68] Attaching the GCN4 br to the DNA significantly increased its effective concentration, and reduced the entropic loss associated with intermolecular binding, thereby allowing the formation of the desired (intramolecular) complex. The design required a Gly-

[64] a) J. A. Robinson, *ChemBioChem* **2009**, *10*, 971–973; b) J. Fernandez-Carneado, D. Grell, P. Durieux, J. Hauert, T. Kovacsovics, G. Tuchscherer, *Biopolymers* **2000**, *55*, 451–458.
[65] N. J. Zondlo, A. Schepartz, *J. Am. Chem. Soc*. **1999**, *121*, 6938–6939.
[66] J. W. Chin, A. Schepartz, *J. Am. Chem. Soc.* **2001**, *123*, 2929–2930.
[67] J. K. Montclare, A. Schepartz, *J. Am. Chem. Soc.* **2003**, 125, 3416–3417.
[68] D. Stanojevic, G. L. Verdine, *Nat. Struct. Biol.* **1995**, *2*, 450–457.

Gly-Cys connector to connect the C-terminus of the GCN4 br to an Adenine nucleotide immediately 5′ to the AP1/GCRE half-site (5′-A(c)TCAT-3′, modified base in bold) through a disulfide bond.

Taking the concept of intramolecular delivery for specific DNA binding one step further, the group of J. L. Mascareñas reported in 2001 the synthesis of conjugates between the GCN4 br and a small minor groove binding molecule that was capable of sequence-specific DNA recognition.[69] The role of the minor groove binding agent (a distamycin derivative) was analogous that of the disulfide linker in the original Verdine's intramolecular complex: to act as a supramolecular anchor that, upon binding to its target sequence in the DNA minor groove (5′-AAATT-3′), would deliver the GCN4 br into its binding site—the ATF/CREB half-site (5′-gTCAT-3′)—in the adjacent major groove. Based on the X-ray structures of the GCN4 dimer and distamycin bound to their respective DNA sites,[70] the authors built a hypothetical model of the simultaneous interaction of the GCN4 basic region and the small DNA binding agent bound to contiguous sites, which allowed the identification of Arg^{245} as the ideal position for installing short aminoalkyl linkers connecting with the distamycin bound to the adjacent minor groove. Unlike isolated GCN4 basic regions, the resulting conjugate displayed tight (low nM affinity at 4 ºC) and selective binding to a composite site containing the consensus sequences for both the GCN4 and the distamycin derivative (5′-gTCAT-AAATT-3′). As expected, control experiments with DNAs lacking the target sequence of the distamycin anchor did not bind to the DNA, thus confirming the key role of the accessory interactions in the minor groove for stabilizing the complex.

Following that initial report, the group of J. L. Mascareñas demonstrated that the conjugation strategy could be generally applied for the stabilization of TF fragments, describing a variety of peptide hybrids that recognized extended sites containing the sequence of both the peptides and the distamycin. Thus, for example, the DNA binding helix of the Skn-1 transcription factor could be conjugated through its Lys^{232} side chain to distamycin, and the resulting hybrid selectively bound to the site 5′-AAAA-TCAT-3′.[71] Importantly, this strategy could also be extended to other transcription factor families, so that conjugation

[69] M. E. Vázquez, A. M. Caamaño, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2001**, *40*, 4723–4725.
[70] M. Coll, C. A. Frederick, A. H. Wang, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8385–8389.
[71] J. B. Blanco, M. E. Vázquez, L. Castedo, J. L. Mascareñas, *ChemBioChem* **2005**, *6*, 2173–2176.

to distamycin restored the DNA binding properties of a fragment of the GAGA zinc finger transcription factor, which selectively bound to the composite sequence 5′-TTTT-GAGAG-3′ containing an A·T-rich site for distamycin binding (5′-TTTT-3′) next to the GAGA TF consensus binding sequence (5′-GAGAG-3′).[72] Similar conjugates with GCN4 and GAGA fragments, and even with homeodomain fragments,[73] were also described using bisbenzamidines as minor groove binding ligands,[74] further demonstrating the modular nature and general applicability of this strategy for obtaining sequence-selective DNA binding peptides derived from transcription factors (Figure 13).[75]

The thermodynamic stabilization of FT fragments to obtain high-affinity DNA binders does not require the use of artificial DNA binding agents such as distamycin or bisbenzamidines as described before, but can also be achieved with fully peptidic anchors. The AT-Hook is a short cationic peptide (RKPRGRPKK) found in eukaryotic HMG-I(Y) nuclear proteins.[76] Although AT-Hooks bind to their target DNA sites with poor affinity (in the millimolar range),[77] HMG-I(Y) proteins exploit the cooperative effect of three AT-Hook repeats to achieve high DNA binding affinity.[78] NMR and crystallography studies have provided a detailed structural picture of the interaction of the AT-Hook with the DNA, and have shown that its central RGR core is deeply inserted into the minor groove in an extended conformation, while the various lysines in the sequence introduce additional electrostatic contacts with the

[72] O. Vázquez, M. E. Vázquez, J. B. Blanco, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed Engl*. **2007**, *46*, 6886–6890.
[73] J. Mosquera, J. Rodríguez, M. E. Vázquez, J. L. Mascareñas, *ChemBioChem* **2014**, *15*, 1092–1095.
[74] a) J. B. Chaires, J. Ren, D. Hamelberg, A. Kumar, V. Pandya, D. W. Boykin, W. D. Wilson, *J. Med. Chem.* **2004**, *47*, 5729–5742; b) M. Munde, M. A. Ismail, R. Arafa, P. Peixoto, C. J. Collar, Y. Liu, L. Hu, M.-H. David-Cordonnier, A. Lansiaux, C. Bailly, et al., *J. Am. Chem. Soc.* **2007**, *129*, 13732–13743; c) O. Vázquez, M. I. Sánchez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Org. Lett.* **2010**, *12*, 216–219.
[75] a) J. B. Blanco, M. E. Vázquez, J. Martinez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Biol.* **2003**, *10*, 713–722; b) J. B. Blanco, O. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Eur. J.* **2005**, *11*, 4171–4178.
[76] a) C. Crane-Robinson, A. I. Dragan, P. L. Privalov, *Trends Biochem. Sci.* **2006**, *31*, 547–552; b) L. Aravind, D. Landsman, *Nucleic Acids Res.* **1998**, *26*, 4413–4421; c) R. Reeves, M. S. Nissen, *J. Biol. Chem*. **1990**, *265*, 8573–8582; d) T. Lund, K. H. Dahl, E. Mørk, J. Holtlund, S. G. Laland, *Biochem. Biophys. Res. Commun.* **1987**, *146*, 725–730.
[77] C.-C. Cheng, Y.-H. Jian, C.-J. Lo, J.-W. Cheng, *J. Chinese Chemical Soc* **1998**, *45*, 619–624.
[78] B. H. Geierstanger, B. F. Volkman, W. Kremer, D. E. Wemmer, *Biochemistry* **1994**, *33*, 5347–5355.

phosphates of the DNA backbone.[79] We recently reported the use of an AT-Hook as minor groove binding element in combination with the GCN4 br. The peptidic nature of the AT-Hook allows the straightforward synthesis of the GCN4 br/AT-Hook chimera following exclusively solid-phase peptide synthesis methods, and the lower DNA binding affinity of the AT-Hook compared to bisbenzamidines or distamycin results in better selectivity profile.[80]

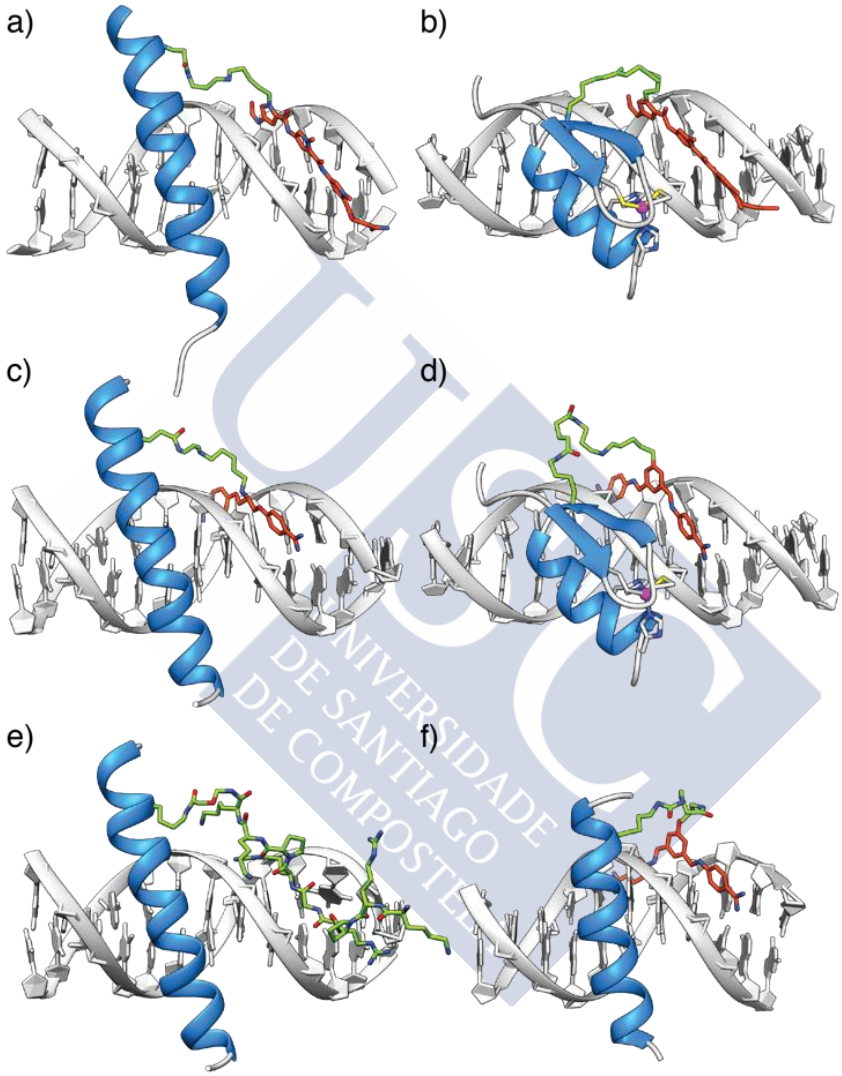


Figure 13. Examples of minimized TF stabilized by small minor groove binding agents. a) Original GCN4 basic region conjugate with distamycin,[73] b) GAGA/distamycin conjugate,[76] c) GCN4/bisbenzamidine hybrid, d) GAGA/bisbenzamidine hybrid; e) GCN4/AT-Hook chimera,[84] f) engrailed homeodomain/bisbenzamidine conjugate.

[79] a) J. R. Huth, C. A. Bewley, M. S. Nissen, J. N. Evans, R. Reeves, A. M. Gronenborn, G. M. Clore, *Nat. Struct. Biol.* **1997**, *4*, 657–665; b) E. Fonfría-Subirós, F. Acosta-Reyes, N. Saperas, J. Pous, J. A. Subirana, J. L. Campos, *PLoS One* **2012**, *7*, e37120.

[80] J. Rodríguez, J. Mosquera, J. R. Couceiro, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.* **2015**, *6*, 4767–4771.

A related example, recently published by the group of S. Roy in 2017, further demonstrates that the combination of short minor groove anchors and TF fragments is a robust approach to developing sequence-specific DNA binding peptides. In their work, S. Roy's group used a short minor groove-binding tail from the Serum Response Factor (SRF) in combination with the DNA-binding helix of Elk1 to yield a dominant negative of the two transcription factors, which upon DNA binding to the DNA upregulate the oncogene cFos.[81] The authors explored different linkers, such as GlyProGlyProGly, Ahx_3 (Ahx = aminohexanoic acid), Gly-Ahx_3, Gly_2-Ahx_3, and Gly_3-Ahx_3, between both DNA binding units, and studied as well conformationally-stabilized versions of the DNA contacting helix containing Aib residues, which are known to stabilize the α-helical conformation by restricting the available backbone conformations.[82] Importantly, in contrast with other conjugates that appear to get stuck inside endosomes upon internalization, the Gly_2-Ahx_3 chimera was readily internalized into the nuclei of A549 lung adenocarcinoma cells, and induced a specific response, down-regulating its targeted gene.

In an alternative approach to dominant negative DNA-binding peptides we described the synthesis of a dominant negative of the oncogenic c-Jun protein as an inactive complex by forcing its heterodimerization with a c-Fos-DNA conjugate.[83] This oligonucleotide-peptide receptor was based on the X-ray structure of the the c-Fos/c-Jun heterodimer bound to the AP1 site,[84] and included a fragment of the c-Fos leucine zipper (Arg^{158} to Lys^{192}) to induce c-Jun dimerization, attached to the 5′ end of a short oligonucleotide (5′-TGACTCATCCATTGCGCG-3′) containing the AP1 half-site (in bold) for c-Jun binding. The oligonucleotide provides increased affinity and selectivity for c-Jun, as demonstrated by its tight binding constant ($K_D \approx 200$ nM) that allowed it to efficiently compete with the natural c-Fos and disassemble the natural c-Fos/c-Jun heterodimer bound to the DNA.

[81] M. Chakraborty, S. Roy, *Chem. Commun*. **2017**, *53*, 376–379.
[82] De Filippis V., De Antoni F, M. Frigo, Polverino de Laureto P, A. Fontana, *Biochemistry* **1998**, *37*, 1686–1696.
[83] E. Pazos, C. Portela, C. Penas, M. E. Vázquez, J. L. Mascareñas, *Org. Biomol. Chem*. **2015**, *13*, 5385–5390.
[84] J. N. Glover, S. C. Harrison, *Nature* **1995**, *373*, 257–261.

Interaction of Rh(II) and Ru(II) complexes with DNA

De novo designed DNA-binding metallopeptides are relatively uncommon, and typically contain an octahedral metallo-intercalator complex that binds in the DNA major groove attached to a short peptide that modulates its DNA recognition or internalization properties. Prominent examples of this general design have been reported by J. Barton since the 90s, when they described conjugates between short 13-mer peptides and photoactive $[Rh(phi)_2(phen')]^{3+}$ (phi = 9,10-phenanthrenequinone diimine; phen′ = 5-(amidoglutaryl)-1,10-phenanthroline) complexes. The coordination compounds intercalate through the DNA major groove with high affinity, and upon photoactivation, induce DNA strand scission. The DNA site specificity was dependent on the sequence of the appended peptide, and the authors found that a single glutamate at position was essential in directing DNA site-recognition to the sequence 5′-CCA-3′.[85] Later examples demonstrated the general applicability of this strategy by inserting different peptides with alternative sequence preferences, such as the DNA recognition helix of the phage 434 repressor (preferentially targeting the sequence 5′-ACAA-3′).[86] Related designs featured α-helical metallopeptides with nuclease activity tethered to the $[Rh(phi)_2(phen')]^{3+}$ complex. Thus, the peptide sequence DPDELEHAAKHEAAAK, contains two His residues (in bold) in positions i, i+4 that create a zinc coordination site on one face of the α-helix. A number of residues were included in the sequence to increase the α-helical content, such as a C-terminal E-K salt bridge, and a number of Ala residues. The resulting metallointercalator-peptide nuclease chimera was found to convert a supercoiled pBR322 DNA plasmid to both the nicked and linear forms. Control experiments did not show any cleavage when the DNA was incubated with the unmodified rhodium intercalator, or the peptide nuclease without the intercalator, thus demonstrating the cooperative nature of the construct.[87] Similar effects were observed with rhodium conjugates with metallopeptide hairpin nucleases.[88]

A relatively recent example of these metallointercalator-peptide hybrids explored the effect of the peptide sequence in the internalization and

[85] a) N. Y. Sardesai, S. C. Lin, K. Zimmerman, J. K. Barton, Bioconjug. Chem. 1995, 6, 302–312; b) C. A. Hastings, J. K. Barton, *Biochemistry* **1999**, *38*, 10042–10051.
[86] N. Y. Sardesai, J. K. Barton, *J. Biol. Inorg. Chem.* **1997**, *2*, 762–771.
[87] M. P. Fitzsimons, J. K. Barton, *J. Am. Chem. Soc.* **1997**, *119*, 3379–3380.
[88] K. D. Copeland, M. P. Fitzsimons, R. P. Houser, J. K. Barton, *Biochemistry* **2002**, *41*, 343–356.

intracellular localization.[89] In this study the $[Ru(phen)(bpy')(dppz)_2]^{2+}$ complex (bpy′ = 4-(3-carboxypropyl)-4′-methyl-2,2′-bipyridine) was attached to a well-known (D)-octaarginine (arg_8) peptide transporter, and to that same peptide modified with a fluorescein tag (arg_8-Flu) (Figure 11). Interestingly, they found that while the simple Ru-[arg_8] conjugate was internalized by endocytosis, and remained in the cytoplasm in a punctuate pattern, the Ru-[arg_8]-Flu derivative was readily internalized into the cell nucleus under the same conditions for which the complex without fluorescein was excluded. This difference in intracellular localization is consistent with other studies that demonstrate the role of charge and hydrophobicity compensation in peptide transporters.[90] This result demonstrates the potential role of the peptide appendages not only in the DNA recognition process, but also for the modification of other properties, such as cell transport.

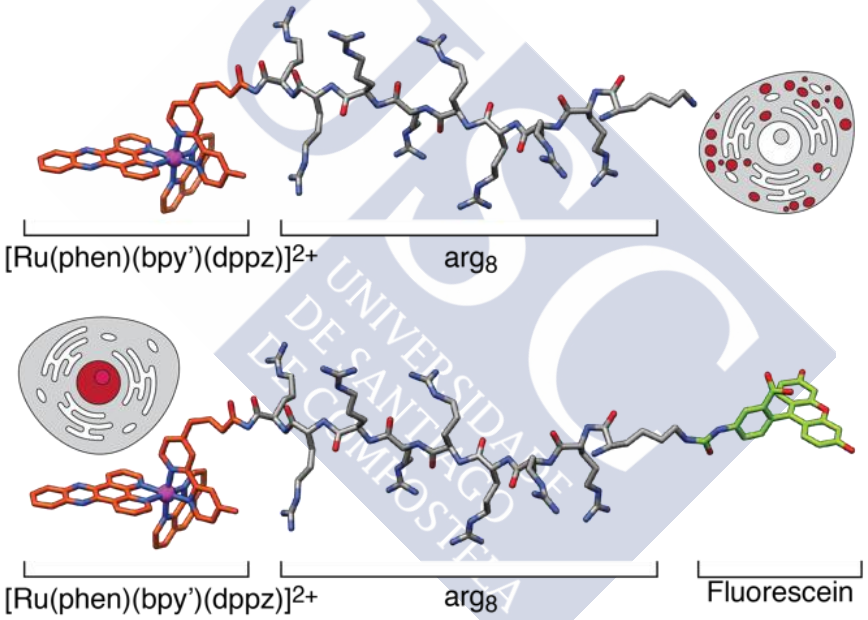


Figure 11. $[Ru(phen)(bpy')(dppz)_2]^{2+}$ arg_8 conjugates display different transport properties depending on their N-terminal modification with the hydrophobic fluorescein fluorophore. $[Ru(phen)(bpy')(dppz)_2]^{2+}$-$arg_8$ is internalized in endosomes, while the analog featuring an N-terminal fluorescein is efficiently transported into the cell nucleus.

In addition to the modification of discrete metallointercalators, peptides have also been used as integral structural elements in de novo DNA binding metallopeptides without any resemblance to natural transcription factors. In this context, in 2014 we described the application of solid-phase peptide synthesis

[89] C. A. Puckett, J. K. Barton, *Bioorg. Med. Chem*. **2010**, *18*, 3564–3569.
[90] a) C. N. Carrigan, B. Imperiali, *Anal. Biochem*. **2005**, *341*, 290–298; b) W. B. Kauffman, T. Fuselier, J. He, W. C. Wimley, *Trends Biochem. Sci*. **2015**, *40*, 749–764.

methodology to assemble peptide ligands as precursor of metal helicates, which are chiral metal complexes constituted by two (or more) metal ions. These metalocylinders have been shown to target unusual three-way DNA junctions,[91] and indeed, upon incubation with Fe(II), the peptide precursors enantioselectively assembled into ΛΛ- or ΔΔ-dimetallic helicates (Figure 12) that displayed high affinity for three-way DNA junctions ($K_D \approx 3.47$ μM for the ΛΛ- isomer, and 13.59 μM for the ΔΔ-helicate). In contrast, the binding of these species to double-stranded DNA was rather weak ($K_D \approx 40$ μM for A·T-rich oligonucleotides, and $K_D \approx 100$ μM for G·C-rich oligos). Following the same methodology, our group has recently described a new family of oligonuclear Ru(II) and Ir(III) polypyridyl organometallopeptides that exhibit exhibit high DNA binding affinity in the low nanomolar range, sequence selectivity, and high cytotoxicity against set cancer cell lines.[92]

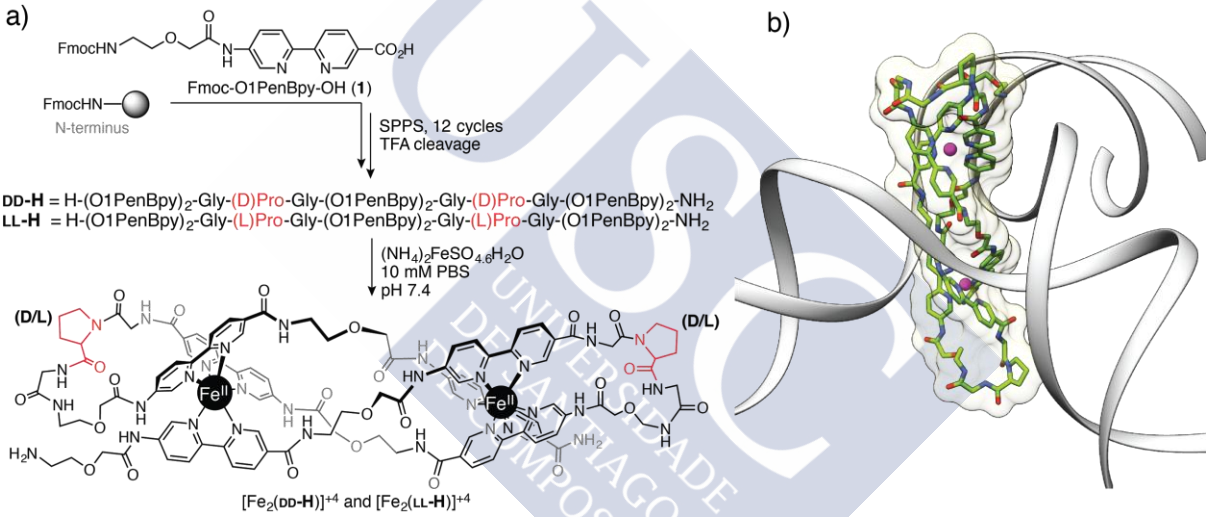


Figure 12. Solid phase peptide synthesis of the peptide helicates. Introduction of (L)-Pro or (D)-Pro residues in the loops selects a particular supramolecular chirality (ΛΛ- or ΔΔ-, respectively) in the final helicates. Bottom: Proposed structure of the peptide helicate bound to the center of a three-way DNA junction. DNA is only represented as ribbons for clarity.

[91] a) A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon, M. Coll, *Angew. Chem. Int. Ed.* **2006**, *45*, 1227–1231; b) S. E. Howson, A. Bolhuis, V. Brabec, G. J. Clarkson, J. Malina, A. Rodger, P. Scott, *Nat. Chem.* **2012**, *4*, 31–36; c) M. Albrecht, *Chem. Rev.* **2001**, *101*, 3457–3498; d) C. R. K. Glasson, G. V. Meehan, J. K. Clegg, L. F. Lindoy, J. A. Smith, F. R. Keene, C. Motti, *Chem. Eur. J.* **2008**, *14*, 10535–10538.

[92] a) I. Gamba, I. Salvadó, R. F. Brissos, P. Gamez, J. Brea, M. I. Loza, M. E. Vázquez, M. Vázquez López, *Chem. Commun.* **2016**, *52*, 1234–1237; b) I. Gamba, I. Salvadó, G. Rama, M. Bertazzon, M. I. Sánchez, V. M. Sánchez-Pedregal, J. Martínez-Costas, R. F. Brissos, P. Gamez, J. L. Mascareñas, et al., *Chem. Eur. J.* **2013**, *19*, 13369–13375; c) G. Rama, A. Ardá, J.-D. Maréchal, I. Gamba, H. Ishida, J. Jiménez-Barbero, M. E. Vázquez, M. Vázquez López, *Chem. Eur. J.* **2012**, *18*, 7030–7035.

In summary, since the early 90s a number of groups have developed efficient strategies for the development of miniaturized versions of transcription factors, these efforts have not only expanded the repertoire of DNA binding motifs, but also produced new systems with unique properties not found in natural systems, such as light-activation. The progress in the field, however, has been relatively slow partly due to our still rudimentary understanding of peptide and protein folding and limited computational methods for efficient in silico selection in the context of peptide-DNA interactions. On the other hand, while many systems have been described, most of them are isolated and static complexes, but complex and dynamic systems involving more than one DNA binding partner are still largely unexplored, and will probably be the focus of future development. Finally, limited by their inherent internalization problems, these peptide systems have not yet found real biological applications, but new developments in delivery platforms and strategies will open the way for their use in living systems.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Objetive

The main goal of this work is to investigate in detail the influence of a functionalization with a polyarginine sequence of a series Ru(II)-dppz complexes in both their B-DNA binding (with or without sequence failure) and cell-internalization properties.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Discusion

Oligoarginine Ru(II)-dppz metallopeptides that bind to matched and mismatched DNAs

Abstract

A series of Ru(II) metallopeptides featuring the dipyrido[3,2-a:2',3'-c]phenazine ligand (dppz) have been synthesized through solid-phase peptide synthesis (SPPS) methods. Spectroscopic studies indicate that functionalization with an oligoarginine basic tail causes a drastic change in the binding mode and a great increase of the affinity of these metallopeptides for well-matched B-DNA oligonucleotides over mismacthed DNAs. In addition, the spectroscopic data suggest that both the oligoarginine functionalization and the nature of the ancillary ligands do not cause an increase of the selectivity of these metallopeptides for a particular type of mismatched DNA oligonucleotide. Finally, fluorescence microscopy studies indicate that the oligoarginine functionalization causes the efficient internalization of the metallopeptides into Vero cells, causing apoptotic cell death after a few hours.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Introduction

Over the past few decades there has been an increased interest in the application of transition metal complexes for the development of DNA-based probes, reactive agents, and drugs.[93] However, severe side effects, such as nephrotoxicity, neurotoxicity, or drug-induced resistance, have hampered their clinical application, and fueled intense efforts to develop safer alternatives with reduced side effects and improve DNA binding.[94] In this context, polypyridyl ruthenium(II) complexes have been exhaustively studied as metal-based DNA binders,[95] These complexes usually contain one large heteroaromatic ligand that penetrates into the DNA through the major groove. Unfortunately, these metallointercalators suffer from the same lack of sequence specificity typical of organic intercalators.[96] Despite this limitation, metal complexes offer extraordinary opportunities for the modulation of the DNA binding properties, by taking advantage of ancillary ligands and stereochemical effects, not accessible to organic DNA binders. While these effects have been extensively explored in other areas, such as catalyst optimization,[97] their study in the context of DNA recognition by metal complexes must still be explored in depth. However, the functionalization of the intercalating/ancillary ligands in polypyridyl Ru(II) DNA binders implies the use demanding synthetic procedures, that complicate the access to multiple structural variants required for SAR studies, as well as the access to single systems endowed with specific biological properties. In this context, a promising and alternative approach to this end is to decorate this class of complexes with peptide sequences. Peptides are easy to synthesize[98] and they are highly versatile systems from a biological perspective, being also capable of endowing with versatile biological capabilities (*i.e.* transmembrane/cellular organelle delivery, protein/nucleic acid

[93] a) C. Moucheron, *New. J. Chem*. **2009**, *33*, 235; b) K. L. Haas K. J. Franz, *Chem. Rev*. **2009**, *109*, 4921.
[94] a) L. Kelland, *Nat. Rev. Cancer*. **2007**, *7*, 573; b) A. Cisnetti, A. Gautier, *Angew. Chem. Int. Ed*. **2013**, *52*, 11976; c) S. Komeda, A. Casini, *Curr. Top. Med. Chem*. **2012**, *12*, 219.
[95] M. R. Gill, J. Garcia-Lara, S. J. Foster, C. Smythe, G. Battaglia, J. A. Thomas, *Nat. Chem*. **2009**, *1*, 662.
[96] H.-K. Liu, P.J. Sadler, *Acc. Chem. Res*. **2011**, *44*, 349.
[97] D. J. Gorin, B.D. Sherry, F. D. Toste, *Chem. Rev*. **2008**, *108*, 3351.
[98] a) K. Heinze, K. Hempel, *Chem. Eur. J*. **2009**, *15*, 1346; b) B. Jung, A. G. Beck-Sickinger, *Angew. Chem. Int. Ed*. **1992**, *31*, 367; c) M. Nitz, K. J. Franz, R. L. Maglathlin, B. Imperiali, *ChemBioChem* **2003**, *4*, 272; d) G. Dirscherl, B. Koenig, *Eur. J. Org. Chem*. **2008**, 597.

targeting, etc…) to a variety of cargo molecules,[99] including Ru(II), Rh(III) or Ir(III) polypyridyl derivatives.[100,101]

On the other hand, a DNA mismatch is a defect that takes place when two non-complementary bases are faced one to each other in a duplex DNA.[102] Mismatches appear in the genome as a consequence of errors during DNA replication,[103] or due to DNA damage caused by ionizing radiation or genotoxic chemicals.[104] These mistakes are almost always corrected by the enzymatic mismatch repair machinery of the cell,[105] but if they are not fixed, they accumulate in the genome, often driving to genetic mutations which can lead to serious consequences, including cancer.[106] The selective recognition of single mismatches in duplex DNA by small molecules, and particularly if these systems are capable of emitting light as a result of such recognition, is a powerful strategy for the development of new theranostic drugs in cancer research. Although there are some examples in bibliography about the use of organic molecules for spectrofluorimetric mismatch detection,[107] the

[99] E. Pazos, J. Mosquera, M. E. Vázquez, J. L. Mascareñas, *ChemBioChem* **2011**, *12*, 1958.

[100] a) I. Gamba, I. Salvadó, R. F. Brissos, P. Gamez, J. Brea, M. I. Loza, M. E. Vázquez, M. Vázquez López, *Chem. Commun*., **2016**, *52*, 1234; b) I. Salvadó, I. Gamba, J. Montenegro, J. Martínez-Costas, J. M. Brea, M. I. Loza, M. Vázquez López, M. E. Vázquez, *Chem. Commun*., **2016**, *52*, 11008; c) I. Gamba, I. Salvadó, G. Rama, M. Bertazzon, M. I. Sánchez, V. M. Sánchez-Pedregal, J. Martínez-Costas, R. F. Brissos, P. Gamez, J. L. Mascareñas, M. Vázquez López, M. E. Vázquez, *Chem. Eur. J.* **2013**, *19*, 13369.

[101] a) U. Neugebauer, Y. Pellegrin, M. Devocelle, R. J. Forster,W: Signac, N. Moran, T. E. Keyes, *Chem. Commun*., **2008**, 5307; b) J. Kuil, P. Steunenberg, P. T. K. Chin, J. Oldenburg, K. Jalink, A. H. Velders,F. W. B. van Leeuwen, *ChemBioChem*, **2011**, *12*, 1897; c) N. Y. Sardesai, K. Zimmermann, J. K. Barton, *J. Am. Chem. Soc.*, **1994**, *116*, 7502; d) M. P. Fitzsimons, J. K. Barton, *J. Am. Chem. Soc.*, **1997**, *119*, 3379; e) K. D. Copeland, M. P. Fitzsimons, R. P. Houser, J. K. Barton, *Biochemistry*, **2002**, *41*, 343; f) C. Sissi, P. Rossi, F. Felluga, F. Formaggio, M. Palumbo, P. Tecilla, C. Toniolo, P. Scrimin, *J. Am. Chem. Soc.*, **2001**, *123*, 3169-3170; g) A. B. Tabor, *Tetrahedron*, **1996**, *52*, 2229.

[102] P. Modrich, *Annu. Rev. Biochem.*, **1987**, *56*, 435.

[103] M. F. Goodman, S. Creighton, L. B. Bloom, J. Petruska, *J. Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 83-126.

[104] C. Kunz, Y. Saito, P. Schär. *Cell. Mol. Life Sci.*, **2009**, *66*, 1021.

[105] P. Modrich, *J. Biol. Chem.* **2006**, *281*, 30305.

[106] a) I. I. Arizamanoglou, F. Gilbert, H. R. Barber, *Cancer* **1998**, *82*, 1808; b) L. A. Loeb, K. R. Loeb, J. P. Anderson, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 776.

[107] a) A. Granzhan, M.-P. Teulade-Fichou, *Chem. Eur. J.* **2009**, *15*, 1314; b) Y. Sato, A. Honjo, D. Ishikawa, S. Nishizawa, N. Teramae, *Chem. Commun*. **2011**, *47*, 5885.

development of DNA-mismatch fluorescent probes based on metal complexes is still poorly developed.[108,109]

In 2009, Barton *et al.* reported the mismatch binding properties of the classic metallointercalator $[Ru(bpy)_2dppz)]^{2+}$,[110] but this complex binds with almost the same strength both well-matched and mismatched DNA sites, even if they are in the same oligonucleotide sequence.[111,112] Later, the same research group showed that through structural modification of the inserting dppz ligand it is not possible to increase the specificity towards mismatched sequences,[109] but with the incorporation of methyl groups into the ancillary ligands of $[Ru(bpy)_2dppz)]^{2+}$ it is possible to get a 26-fold increase in the binding affinity for the mismatch compared to the well-matched duplex DNA, opening a window of opportunity to design improved mismatch-specific luminescent binding agents.[113]

In this context, we are currently embarking into a research program which aims to study the DNA-binding, cell-internalization properties and biological activity of artificial metallopeptides synthesized by SPPS methods and derived from the metal-chelating unit 2,2'-bipyridine.[100,114] In particular, we are especially interested in studying in which way the presence of an oligoarginine sequence

---

[108] Barton *et al*. developed a large family of octahedral Rh(III) polypyridyl complexes equipped with sterically expansive planar ligands which selectively recognise DNA mismatches and, upon photoactivation, cause the cleavage of the DNA skeleton in a highly efficient way, but this class of complexes are not emissive. a) K. M. Boyle, J. K. Barton, *Inorg. Chim. Acta*, **2016**, *452*, 3; b) A. C. Komor, J. K. Barton, *Chem. Commun*., **2013**, *49*, 3617.

[109] The Ru(II) analogues of the Barton's Rh(III) metalloinsertors[108] are not emissive at ambient temperature, despite being highly selective for DNA mismatches. A. J. McConnell, M. H. Lim, E. D. Olmon, H. Song, E. E. Dervan, J. K. Barton, *Inorg. Chem.* **2012**, *51*, 12511.

[110] Polypyridyl ruthenium(II) complexes have been deeply studied since more than 30 years ago due to their ability to bind the DNA in conjunction with their kinetic stability and convenient optical properties. From these, dipyrido[3,2-a:2',3'-c]phenazine (dppz) Ru(II) complexes are considered as "light-switch" models for DNA-binding, as they are practically non-emissive in water media but show a significant enhance in the emission when interacting with DNA: a) A. E. Friedman, J. C. Chambron, J. P. Sauvage, N. J. Turro, J. K.Barton, *J. Am. Chem. Soc*. **1990**, *112*, 4960; b) E. Amouyal, A. Homsi, J.-C. Chambron, J.-P. Sauvage, *J. Chem. Soc., Dalton Trans*., **1990**, 184; c) C. Hiort, P. Lincoln, B. Nordén, *J. Am. Chem. Soc*., **1993**, *115*, 3448.

[111] M. H. Lim, H. Song, E. D. Olmon, E. E. Dervan, J. K. Barton, *Inorg. Chem*. **2009**, *48*, 5392.

[112] H. Song, J. T. Kaiser, J. K. Barton, *Nat. Chem*. **2012**, *4*, 615-620.

[113] A. N. Boynton, L. Marcélis, J. K. Barton, *J. Am. Chem. Soc*. **2016**, *138*, 5020.

[114] I. Gamba, G. Rama, E. Ortega-Carrasco, J.-D. Maréchal, J. Martínez-Costas, M. E. Vázquez, M. Vázquez López, *Chem. Commun*., **2014**, *50*, 11097.

in the structure of the complexes modulate their DNA-binding and biological properties. We have yet reported the influence of this functionalization in the binding properties of B-DNA groove binders, but the same has never been studied for B-DNA metallointercalators[115] as, for example, dppz-derived Ru(II) complexes.[116,117]

Results and discussion

Metallopeptide synthesis

We first designed two peptidic ligands equipped with a single βAla-bpy coordinating unit (**1**,[100] scheme 1): **2-R_8** (scheme P1) is equipped with a C-terminal octarginine (R_8) sequence, and ligand **2**, which lacks the octargine tail. The ligands were synthesized following standard Fmoc/*t*Bu solid-phase protocols.[118] Once completed the ligand sequence, the peptides still attached to the solid support were reacted with the Ru(II) source ([$Ru(DMSO)_4Cl_2$]),[119] following by the coordination of the intercalating dppz moiety and the necessary second ancillary ligand.[120] To this end, the resins were reacted sequentially with 1,10-phenanthroline (phen) and later with dppz to achieve the desired Ru-dppz DNA "light-switches" [Ru(**2**)(phen)(dppz)]$^{2+}$ (**3**, scheme P1) and [Ru(**2-R_8**)(phen)(dppz)]$^{2+}$ (**4**, scheme P1). Acidic cleavage from the support, followed

[115] Barton *et al.* suggested in a very preliminary way that an oligoarginine functionalization may increase the nonspecific DNA binding affinity in chrysy-derived Rh(III) intercalators. However, this issue has never been studied in detail: J. Brunner, J. K. Barton, *Biochemistry* **2006**, *45*, 12295.

[116] Barton *et. al.* reported studies on the influence of an oligoarginine functionalization in the cell-internalization properties of a classic Ru(II)-dppz intercalator, but the DNA-binding properties of such metallopeptide conjugate have never been studied so far: a) C. A. Puckett, J. K. Barton, *J. Am. Chem. Soc.*, **2009**, *131*, 8738; b) C. A. Puckett, J. K. Barton, *Bioorg. Med. Chem.*, **2010**, *18*, 3564.

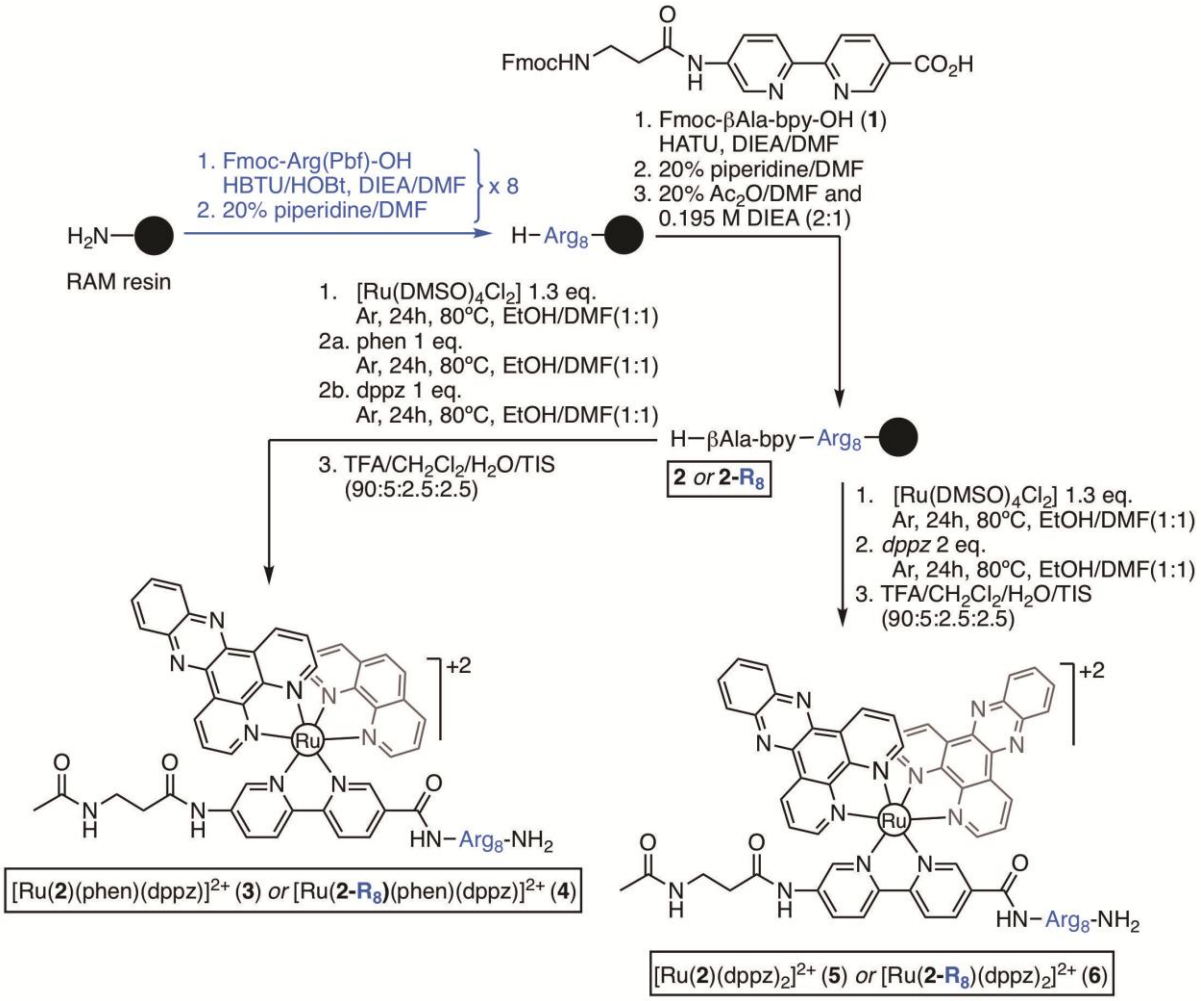
[117] Barton *et al.* reported the DNA binding properties of a series of "light-switch" Ru(II)-dppz peptide conjugates. In this study they suggested that the peptide functionalization slightly modulate the DNA-binding affinity of the Ru(II)-dppz intercalators and that the presence of positive charged residues in the peptide sequence has a quite discreet but positive effect on such affinity due the establishment of electrostatic interactions with the DNA polyphosphate backbone. However, the studies were made with call-thymus DNA and the peptide functionalizations studied were sequences of 5/6 residues (the maximum number of positively charged residues in the sequences is 3 and none of which was an octoarginine sequence: K. D. Copeland, A. M. K. Lueras, E. D. A. Stemp, J. K. Barton, *Biochemistry*, **2002**, *41*, 12785.

[118] I. Coin, M. Beyermann, M. Bienert, *Nat. Protoc.*, **2007**, *2*, 3247.

[119] P. Evans, A. Spencer, G. Wilkinson, *J. C. S. Dalton*, **1973**, 204.

[120] J. E. Dickenson, L. A. Summers, *Aust. J. Chem.*, **1970**, *23*, 1023.

by reverse-phase HPLC purification, afforded the Ru(II) metallopeptides as diasteromeric mixtures. It has to be noted that these class of metallopeptides have four different isomers (scheme S1): the arrangement of ligands around the metal center can be Δ or Λ, and the position of the peptide tail can be "axial" or "equatorial" to the dppz ligand.[117]



*Scheme P1. Solid-phase synthesis of the peptidic ligand **2** and its derived Ru(II) metallopeptides **3** and **5** as well as their oligoarginine derivatives **2-R$_8$**, **4** and **6**. The first step of the route (in blue) takes place only for the synthesis of the oligoarginine derivatives.*

DNA binding studies with synthetic oligonucleotides

Recognition of canonical dsDNA

Once we had the target Ru(II)-dppz derivatives at hand, the second step was to exploit their solvatochromic properties for studying their interaction with synthetic oligonucleotides, which would allow us to obtain quantitative information about their binding affinity towards particular well-matched dsDNA sequences: AAAATTT, GAATTC and GGCCC (see table S1 for the entire sequences). Thus, the addition of successive aliquots of a stock solution of the B-DNA hairpin oligonucleotides to 2 μM solutions of **3** in Tris-HCl buffer at pH 7.5, resulted in a progressive increase of the intensity of emission band of the Ru(II) complex, which could be fitted to a 1:1 binding model corresponding to the formation of specific **3**/B-DNA adducts (figure P1 and figures S1-S4). The resulting apparent dissociation constants obtained suggest that **3** interacts indistinctly and with almost the same strength with the three oligonucleotides. The DNA affinity constants calculated for **3** are in line with those previously reported for related Ru(II)-dppz complexes[121] and peptide conjugates (table S2).[117] After that, we repeated these studies with the oligoarginine derivative **4**. Interestingly, the titrations studies carried out indicated that this metallopeptide shows a more complex and interesting behavior than **3**. In particular, the fluorescence titration profiles at λ_{em}= 630 nm clearly suggest the existence of a two-step DNA-binding process. During the first additions of the DNA stock solution, the intensity of the emission band increases rapidly (and at a much higher rate than in the case of **3**). In fact, the emission intensity reaches its maximum when the [Ru]/[DNA] ratio is *c.a.* 8.0. After that, the emission intensity experiences first a slight decrease and then a very little pronounced but steady rise up until the end of the titration experiment (Figure P1 and Figures S5-S7). This multiphasic profile has been previously observed in reverse titrations of cationic proteins with DNA,[122] and also in titrations of bisbenzamidine-peptide hybrids with dsDNA,[123] and in all these cases it has been explained in terms of the competitive formation of multiple

[121] a) I. Haq, P. Lincoln, D. Suh, B. Nordén, B. Z. Chowdry, J. B. Chaires, J. B., *J. Am. Chem. Soc.*, **1995**, *117*, 4788; b) P. Lincoln, A. Broo, B. Nordén, *J. Am. Chem. Soc*., **1996**, *118*, 2644.

[122] a) J. J. Hollenbeck, M. G. Oakley, *Biochemistry* **2000**, *39*, 6380; b) C. Portela, F. Albericio, R. Eritja, L. Castedo, J. L. Mascareñas, *ChemBioChem*, **2007**, *8*, 1110; c) A. V. Fedorova, I.-S. Chan, J. A. Shin, *Biochim. Biophys. Acta* **2006**, *1764*, 1252.

[123] M I. Sánchez, O. Vázquez, M. E. Vázquez, J. L. Mascareñas, *Chem. Eur. J.*, **2013**, *19*, 9923.

nonspecific complexes with the DNA oligomers at high ligand/DNA ratios and the subsequent formation of specific 1:1 complexes at low ligand/DNA ratios. However, in the particular case of the metallopeptide **4** it should be noted that the increase in the emission intensity must be ascribed solely to the intercalation of the dppz moiety into the DNA double helix and that this intercalative binding mode cannot be considered as a nonspecific interaction. Thus, this unusual profile could be interpreted as the result of the existence of two different intercalative binding modes at different [Ru]/[DNA] ratios. In particular, we suggest the occurrence of an intercalative process of multiple metallopeptide units within a single DNA molecule in the first steps of the titration experiment, when the ratio of the DNA vs complex concentration in the media is very low. Later, in a second stage, when the concentration of DNA rises up, the metallopeptide units rearrange, and intercalate in other DNA molecules. Moreover, we were able to calculated a K_d of 1.09(0.26) μM for the second binding process of **4** with the GAATTC oligonucleotide (the fit was made only for the points represented by white circles). In the other two cases (AAATTT and GGCCC) the second part of the curve is too shallow to make the fit (Figure P1). This interpretation implies the existence of a greater affinity of **4** for the B-DNA with respect to that of **3** although the affinities of both metallopeptides cannot be quantitatively compared. The B-DNA binding behaviour of **4** is also quite different to those shown by related Ru-dppz peptide conjugates:[117] everything seems to indicate that the presence of such a high number of positive charged residues in the peptide tail of the Ru(II)-dppz complex is responsible for this unusual behavior.

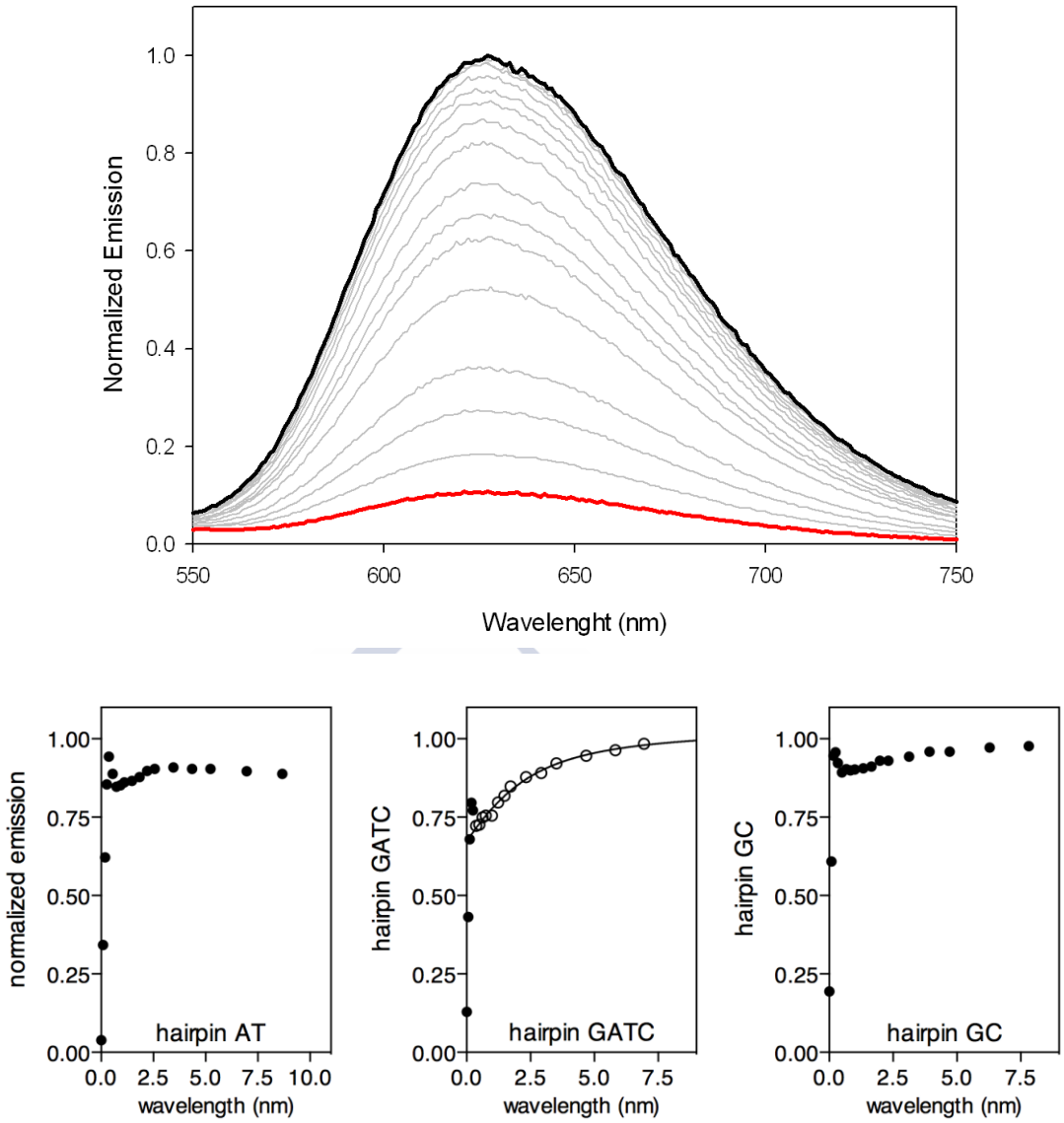


*Figure P1. Up, luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AAAATTT oligonucleotide solution until saturation. Back, profile of the titration experiments of **4** with AAAATTT (left), GAATTC (middle) and GGCCC (right) oligonucleotides at λ_{em} = 630 nm (emission intensity vs. concentration of DNA in the media). We were able to make the fit for the second part of the titration experiment only for the case of the GAATTC oligonucleotide [K_d= 1.09(0.26) μM].*

Absorption spectroscopy is a very useful technique to study the interaction of small molecules with the DNA.[124] In the particular case of polypyridyl DNA metallointercalators, the binding process usually involves both bathocromic and hypsochromic phenomenas of the MLCT band of the complex.[125] The

[124] J.-G. Liu, Q.-L. Zhang, X.-F. Shi, L.-N. Ji, *Inorg. Chem.*, **2001**, *40*, 5045.
[125] B. C. Poulsen, S. Estalayo-Adrián, S. Blasco, S. A. Bright, J. M. Kelly, D. Clive Williams, T. Gunnlaugsson, *Dalton. Trans.*, **2016**, *45*, 18208.

hypochromicity may be attributed to the interaction between the electronic states of the complex and those of the DNA bases,[126] while the bathochromic shift is related with the decrease in the energy gap between the HOMO/LUMO molecular orbitals after the binding process.[127] It has been suggested that the extent of hypochromism in the MLCT band of the complexes is related with the intercalative binding strength of the binder[128] and/or with different levels of penetration of the dppz ligand into the B-DNA base stack.[124] In this context, we studied the interaction of the metallopeptides **3** and **4** with the AAAATTT hairpin oligonucleotide by UV-vis spectroscopy. The intercalative binding mode of these metallopeptides proposed on the basis of the fluorescence studies are also obvious from the UV-data. In particular, the absorption spectra of these complexes are characterized by a MLCT band centered at 443 nm that suffers of both a bathochromic and a hypochromic phenomenon after the addition of the AT-rich B-DNA oligonucleotide (figures S51 and S52; table S4), indicating that **4** binds to B-DNA in an intercalative fashion.[124,129] Interestingly, the decrease in the absorbance at 443 nm is more pronounced in **3** (37.5%) than in 4 (19.0%), which in this case could reflect different levels of penetration of the dppz ligand into the B-DNA base stack, probably due to the steric hindrance caused by the oligoarginine tail of **4**.[124]

Recognition of mismatched DNAs

The next step in our research program was to investigate the affinity of the Ru(II)-dppz metallopeptides **3** and **4** for a set of B-DNA oligonucleotides with different number of mismatches in their sequences (A:A1': one mismatch GG; A:A2': two mismatches GG/AA; A:A3': three mismatches GG/GG/AA; see table S1 for the entire sequences). For comparison, parent well-matched (WM) B-DNA oligonucleotide was also studied (A:A'). Thus, the addition of successive aliquots of a stock solution of these mismatched (MM) B-DNA oligonucleotides to 2 μM solutions of **3** or **4** in Tris-HCl buffer at pH 7.5, resulted in a progressive increase of the intensity of emission band of the Ru(II) complex, which could be fitted to a 1:1 binding model corresponding to the formation of specific metallopeptide/MM B-DNA adducts (figures S15-S19 and

[126] B. D. Wang, Z. Y. Yang, P. Crewdson, D. Q. Wang, *J. Inorg. Biochem.*, **2007**, *101*, 1492.
[127] F. Arjmand, M. Aziz, *Eur. J. Med. Chem.*, **2009**, *44*, 834.
[128] J. K. Barton, J. J. Dennenberg, J. B. Chaires, *Biochemistry*, **1993**, *32*, 2573.
[129] a) A. K. F. Martensson, P. Lincoln, *Dalton. Trans.*, **2015**, *44*, 3604; b) T. Very, S. Despax, P. Hébraud, A. Monari, X, Assfeld, *Phys. Chem. Chem. Phys.*, **2012**, *14*, 12496.

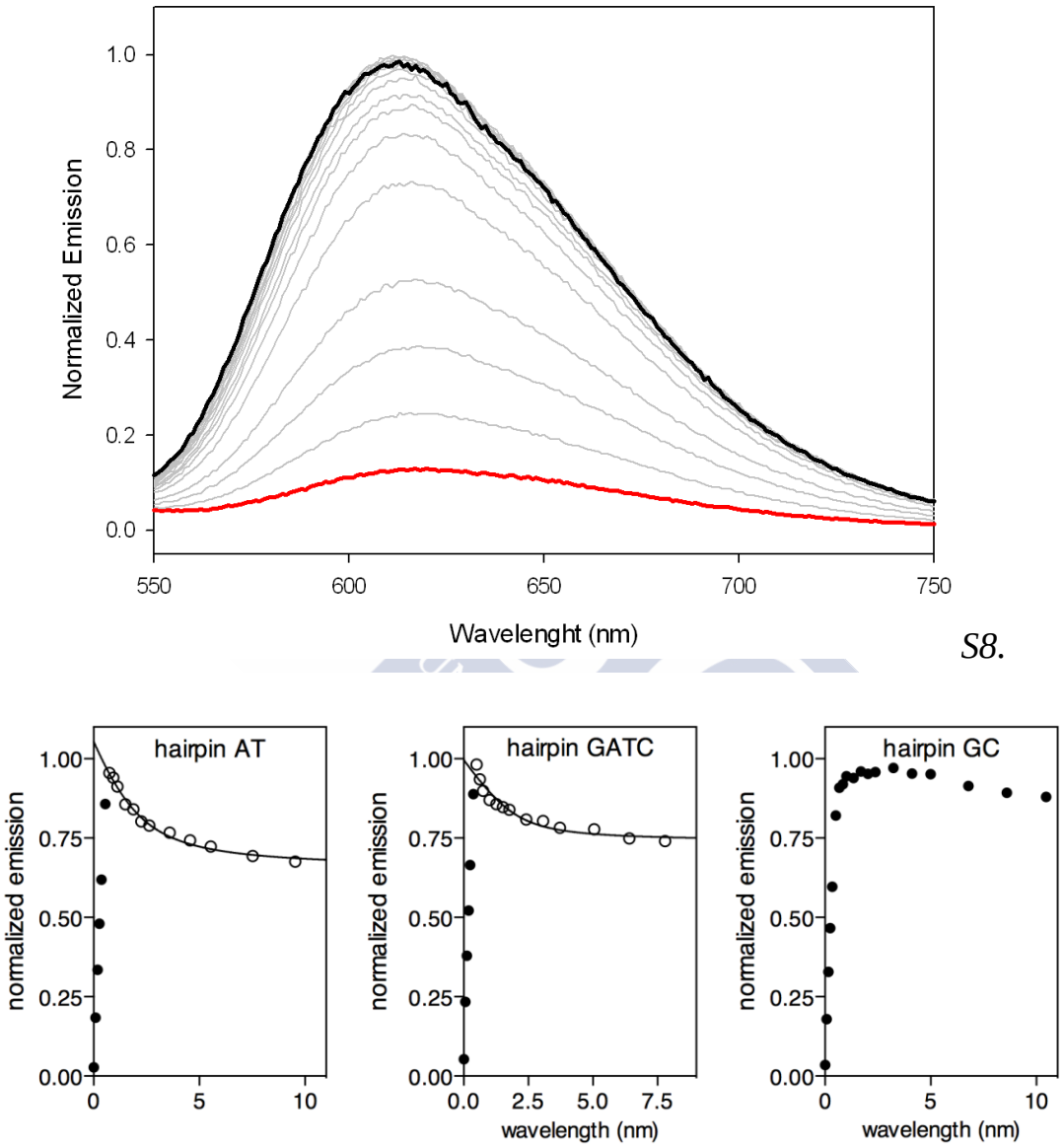
S24-S28). The K_d obtained are very similar (low micromolar) in all the cases and there does not seem to be any significant trend in them (table S3).

In order to check if **3** and **4** displays preference for a specific class of mismatched base pair, as it happens in Barton's Rh(III) metalloinsertors,[130] we also performed fluorescence titration experiments of these complexes with another set of mismatched oligonucleotides named B:B1' (which contains a CC mismatch), C:C1' (AA mismatch) and D:D1' (TT mismatch; see table S1 for the complete sequences). Again, the K_d obtained are very similar (low micromolar) in all the cases and there does not seem to be any significant trend in them (table S3, figures S20-S23 and S29-S32). Surprisingly, the K_d obtained for the interaction of **4** with the TT mismatched oligonucleotide is the one with the lowest value (*c.a.* 3-fold smaller than the K_d with CC, which is the most destabilized mismatch, and also with AA, and 5-fold smaller than the K_d with the more stable GG mismatch).[130]

Later, we decided to study the role of the non-peptidic ancillary ligand in the interaction of these metallopeptides with "well-matched" B-DNA. Therefore, we synthesized two new Ru(II) complexes, $[Ru(\mathbf{2})(dppz)_2]^{2+}$ (**5**) and $[Ru(\mathbf{2}\text{-}R_8)(dppz)_2]^{2+}$ (**6**) (scheme P1), following the same synthetic route described before for **3** and **4**. Once synthesized, we performed fluorescence titration studies with the entire set of hairpin B-DNA oligonucleotides (AAAATTT, GAATTC and GGCCC; see table S1 for the entire sequences; figures S8-S14). The calculated K_d values for **5** (1:1 binding model) shown that the substitution of the ancillary ligand causes a 5-fold increase of the affinity for the AT-rich sequence, whereas the affinity for the other two sequences remains more or less unchanged (figure P2 and table S2). By contrast, the pattern of the interaction of the oligoarginine derivative **6** against the hairpin B-DNA oligonucleotides is similar to that of its parent metallopeptide **4**, suggesting a very similar, and also unusual, binding behaviour (see table S2). However, there are some important differences between them. The first one has to do with the [Ru]/[DNA] ratio in which the emission maximum is reached which, in the case of **6**, is of *c.a.* 3.0

[130] Barton *et al.* demonstrated that there is a clear relationship between mismatch stability and the strength of metalloinsertor's binding: in general the more destabilized the mismatch, the tighter the binding. In this context, the binding affinity of a typical metalloinsertor is expected to decrease in the DNA mismatch series CC, TT, AA and GG: a) B. M. Zeglis, V. C. Pierre, J. K. Barton, *Chem. Commun.*, **2007**, 4565; b) B. A. Jackson, V. Y. Alekseyev, J. K. Barton, *Biochemistry*, **1999**, *38*, 4655; c) B. A. Jackson, J. K. Barton, *Biochemistry*, **2000**, *39*, 6176; d) H. Junicke, J. R. Hart, J. Kisko, O. Glebov, I. R. Kirsch, J. K. Barton, *Proc. Natl. Acad. Sci. U. S. A.*, **2003**, *100*, 3737.

(instead the value of *c.a.* 8.0 founded for **4**). The second one is the different behavior shown by **6** in the second part of the titration experiment with respect to those shown by **4**. If in the case of **4** the intensity increased slightly and constant, in the case of **6** the trend is the opposite: a slight and constant decrease in the emission intensity.



S8.

Figure P2. Up, luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AAAATTT oligonucleotide solution until saturation. Back, profile of the titration experiments of* ***6*** *with AAAATTT (left), GAATTC (middle) and GGCCC (right) oligonucleotides at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media). We were able to make the fit for the second part of the titration experiment only for the cases of the AAAATTT and GAATTC oligonucleotides [*K_d*= 0.72(0.16) and 0.33(0.21) μM, respectively].*

We also performed fluorescence titration experiments of **5** and **6** with the entire set of mismatched double-stranded B-DNAs. However, the data indicated that the substitution of the ancillary ligand does not have so much influence in the affinity and selectivity of the metallopeptides for these oligonucleotides (figures S33-S50). We also performed UV-vis studies with the entire set of mismatched B-DNA oligonucleotides under investigation (figures S67-S80). The values of the hypsochromic shifts (HS) undergone by the MLCT bands of the Ru(II) metallopeptides after the interaction with the oligonucleotides are collected in table S4. In agreement with fluorescence data, almost all the HS values are in a close range and there seems to be no clear trend emerging in them. However, it seems that the perturbative steric effect of the R_8 tail over the penetration level of the dppz ligand into the B-DNA base stack occurs only when the metallopeptides interact with the well-matched B-DNA oligonucleotides (AAAATTT and A:A'). In other words, the UV-vis data suggest that the R_8 tail hinders the penetration of the dppz ligand into double helix only when the nucleobases are perfectly matched. Interestingly, this behaviour is significantly more pronounced in the case of the phen-dppz derivatives (complexes **3** and **4**) than in the double-dppz metallopeptides (complexes **5** and **6**), which suggest a certain influence of the ancillary ligand in the intercalation process.

Circular dichroism studies

DNA is an optically active molecule as a result of the chiral sugar-phosphate backbone and its helicity. DNA can exist in three double helical forms, B, A or Z-DNA and indeed each of them has a different characteristic CD signal in the UV region. In particular, a solution of B-DNA exhibits a positive CD band (275 nm) due to base stacking and a negative band (245 nm) due to right-handed helicity of DNA. DNA CD signals are very sensitive to the mode in which this drugs interact with the DNA and, for this reason, CD spectroscopy is considered a quite useful technique in diagnosing the changes of DNA conformation during small molecule–DNA interactions.[131] In the other hand, nonchiral drug or racemic mixtures of chiral molecules have no CD signal. However, when bound to DNA, these systems often give rise to a CD spectrum due to an induced CD signal (ICD) resulting from the chiral environment around the molecule or due the enantiopreferential binding of one of the isomers of the racemic mixture to

[131] V. I. Ivanov, L. E. Minchenkova, A. K. Schyolkina, A. I. Poletayer, *Biopolymers*, **1973**, *12*, 89; b) B. Nordén, T. Kurucsev, *J. Mol. Recognit.*, **1994**, *7*, 141.

DNA.[132] Therefore, the CD technique has been also used to study the enantiopreferential DNA binding of rac-metal complexes.[133] Although there are lots of data about the study of the interaction of metal-complex DNA-binders with the B-DNA by CD spectroscopy,[134] the studies with mismatched B-DNA are very scarce.[135]

We performed CD studies on the four Ru(II) metallopeptides **3-6** with a selection of the set of well-matched and mismatched double-helix B-DNA oligonucleotides under study, in order to elucidate the role of the intercalation/insertion of the Dppz unit of the complexes as well as the influence of the oligoarginine functionalization and/or ancillary ligand substitution in the conformation of the B-DNA double helix (figures S81-S110). In the particular case of the metallopeptides **3** and **4** with the AAAATTT hairpin oligonucleotide (figures S81 and S82), no ICD signals are observed in the region of 200-320 nm, or also between 320-600 nm, where it is not expected that exist bands due to the DNA, indicating that there is no preferential binding of any of the possible 4 isomers of these complexes to the B-DNA double helix.[136] In addition, there is also not observed a displacement of the original CD bands of the AAAATTT oligonucleotide, suggesting that the conformation of the B-DNA is maintained after the intercalation process. Finally, contrary to what it was observed for other intercalators, there is observed a slight decrease of the intensity of the original B-DNA CD bands as a consequence of the recognition process, suggesting that there is a slight destabilization of the double helix after the intercalation of the metallopeptides. A very similar behaviour has been also

[132] a) B. N. A. Roger, *Circular and Linear Dichroism*, Oxford University Press, **1997**; b) M. Eriksson, B. Norden, *Drug-Nucleic Acid Interactions* **2001**, *340*, 68.

[133] a) S. Delaney, M. Pascaly, P. Bhattarcharya, K. Han, J. K. Barton, *Inorg. Chem.* **2002**, *41*, 1966; b) C. Hiort, B. Norden, A. Rodger, *J. Am. Chem. Soc*., **1990**, *112*, 1971; c) S. Delaney, M. Pascaly, P. Bhattarcharya, K. Han and J. K. Barton, *Inorg. Chem.*, **2002**, *41*, 1966.

[134] a) A. K. Mårtensson, P. Lincoln, *Dalton Trans*., **2015**, *44*, 3604; b) J. M. Kelly, A. B. Tossi, D. J. Mcconnell, C. Ohuigin, *Nucleic Acids Res.*, **1985**, *13*, 6017; c) R. B. P. Elmes, M. Erby, S. A. Bright, D. C. Williams, T. Gunnlaugsson, *Chem. Commun*., **2012**, *48*, 2588.

[135] A. C. Komor, J. K. Barton, *J. Am. Chem. Soc.* **2014**, *136*, 14160.

[136] The disappearance of the two original CD bands of the B-DNA and the subsequent formation of inverted signals with intensities much greater than the free DNA is a typical scene of exciton coupled ICD arising due to enantiopreferential binding of the D-enantiomer of the *rac*-complex and/or ligand-ligand interactions among DNA bound/unbound complexes: a) P. Uma Maheswari, V. Rajendiran, M. Palaniandavar, R. Parthasarathi,V. Subramanian, *J. Inorg Biochem*., **2006**, *100*, 3, b) P. Uma Maheswari, V. Rajendiran, H. Stoeckli-Evans, M. Palaniandavar, *Inorg. Chem.* **2006**, *45*, 37; c) V. Rajendiran, M. Murali, E. Suresh, M. Palaniandavar, V. Subbarayan Periasamy, M. Abdulkader Akbarsha, *Dalton Trans*., **2008**, 2157.

observed for the four metallopeptides (**3**-**6**) and the entire set of mismatched B-DNA oligonucleotides under study (figures S83-S110), suggesting analogous rearrangements of the double helices as a consequence of the binding processes.

Cell internalization studies by fluorescence microscopy

Intrigued by the unusual DNA binding properties of these Ru(II) metallopeptides we decided to study them by fluorescence microscopy. When Vero cell monolayers were incubated with the metallopeptide **4** at 5 μM for 24 hr, a punctate fluorescent pattern could be observed under the fluorescence microscope (figure P3A). The fluorescent spots on the cytoplasm of the Vero cells are consistent with the endocytic internalization of the compound. The irregular staining of some of the observed vesicles (see inlay) suggests that at least part of the compound is aggregating inside the endocytic vesicles. Such effect might be due to an elevated intra-vesicle concentration of the metallopeptide or to the particular composition or physico-chemical characteristics of the endosomes. It is noteworthy that all the cells in the culture are apparently live and healthy.

In a different experiment, we incubated Vero cells with an higher concentration of the compound **4** (25 μM) for 30 minutes. The cells were then washed three times with PBS before replacing the medium with fresh DMEM without serum. The fluorescence emission of the metallopeptide was continuously recorded together with brightfield monitoring of the culture at 37ºC for 4 hours after the addition of the compound (Figure P3B). At the high concentration, the metallopeptide forms non-uniform aggregates that settle on top of the cells. The presence of the aggregates cause cell breakage with leakage of the cytosol content, probably caused by destabilization of the cell membrane. Remarkably, all cells not associated with an aggregate die within the four hours of observation showing typical signs of apoptosis. Those results might suggest that the interaction of the compound with the extracellular face of the membrane triggers apoptosis. On the other hand, the fact that no cell death is observed at the lower concentration, but 24 hours continuous incubation seems to rule out this possibility. Direct cell-internalization at high concentration was previously observed for other octoarginine-containing compounds.[116] Thus, it is also possible that direct internalization of part of the metallopeptide is occurring at 25 μM that somehow causes apoptosis of the treated cell, while preventing its intracellular accumulation to a detectable level. Further investigation is needed

to clarify this issue. Finally, it has to be noted that metallopeptide 3 does not show either cytotoxic or internalization properties (Figure S111).

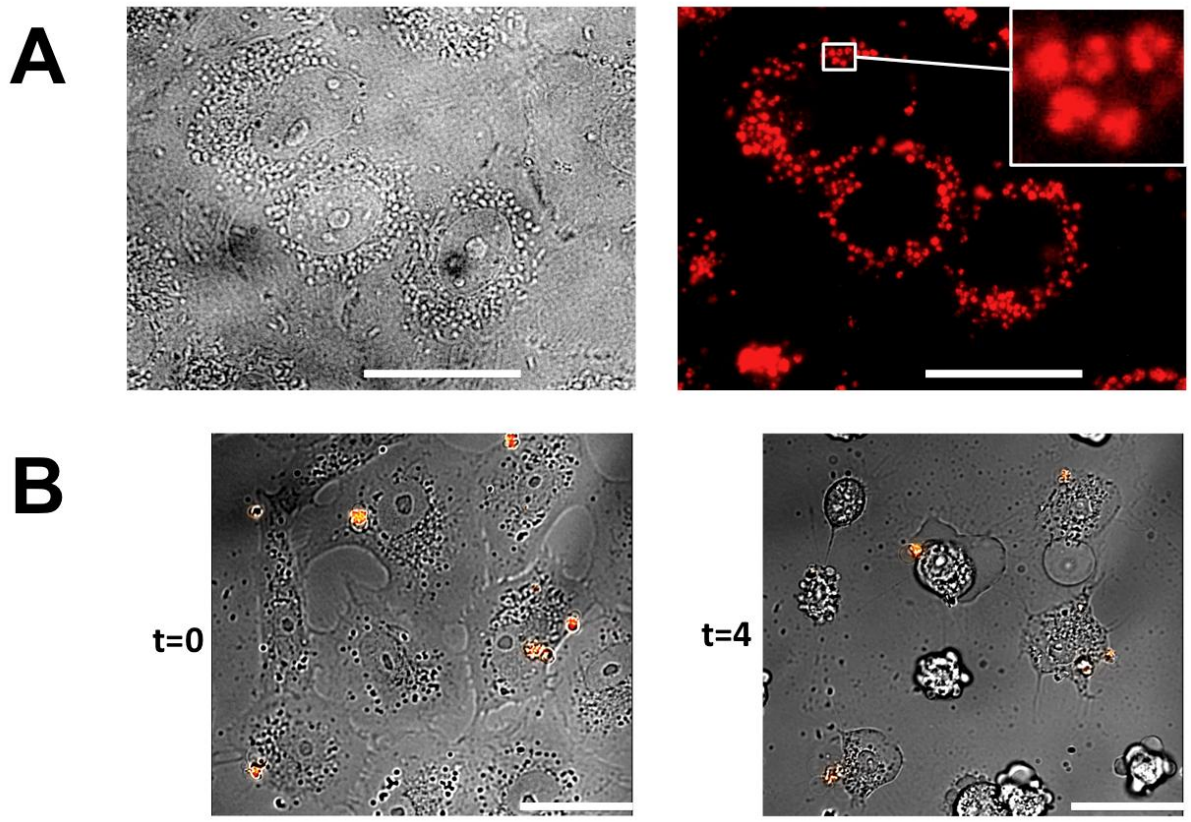


*Figure P3. A) Monolayers of Vero cells were incubated with 5 μM solution of **4** in medium DMEM without serum for 24 hours. Brightfield (left) and red-emission fluorescence pictures (right) were taken from unfixed living cells at 1000X magnification. The scale bar corresponds to 20 μm. B) Monolayers of Vero cells were incubated with 25 μM of **4** in DMEM without serum for 30 minutes. After that, the medium was removed, the cells washed three times with PBS, and further incubated at 37ºC in DMEM without serum. The figure shows captures at the starting incubation time (t=0) and after 4 hours of incubation (t=4) showing merged bright field and fluorescence (orange spots) images. As in A, the scale bar corresponds to 20 μm*

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Conclusions

We report herein a systematic study about the effect of the presence of an oligoarginine functionalization in the B-DNA and mismatched B-DNA binding properties of a Ru(II)-dppz metallointercalator. In particular, the fluorescence and UV-vis studies carried out suggested that the functionalization of a Ru(II)-dppz complex with a octaarginine tail causes an increase of its affinity for well-matched B-DNA. However, this peptide tail does not seem to have any influence on their binding properties for mismatched B-DNA. In the same way, the role of the non-peptidic ancillary ligand of these complexes over the B-DNA binding properties of these complexes to the entire set of studied oligonucleotides seems to be also negligible. Moreover, CD studies suggest that there is no preferential binding of any of the possible 4 isomers of this class of metallopeptides to the well-matched and mismatched B-DNA and also that their intercalation to the double helix cause a slight destabilization of the same. Finally, the R_8 functionalization endowed these Ru(II) metallopeptides with cell-internalization properties and, moreover, with appreciable cytotoxic capabilities.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Experimental Section

General

All reagents were acquired from commercial sources: Dimethyl sulfoxide (DMSO), Dimethylformamide (DMF) and Trifluoroacetic acid (TFA) were purchased from Scharlau, EtOH and CH_2Cl_2 from Panreac and CH_3CN from Merck. All peptide synthesis reagents, namely the coupling agents HBTU (O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluorophosphate) and HATU (2-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate methanaminium), and Fmoc amino acid derivatives were purchased from GL Biochem (Shanghai) Ltd. Fmoc-Rink amide AM resin was purchased from Iris Biotech. The oligonucleotides were purchased from Thermo Fisher Scientific GmbH. All other chemicals were purchased from Alfa-Aesar, Sigma-Aldrich or Fluka. All solvents were dry and of synthesis grade, unless specifically noted. $RuCl_3.3H_2O$ was purchased from Matthews Chemicals. Reactions were followed by analytical RP-HPLC with an Agilent 1100 series LC/MS using a Luna C18 (250 x 4.60 mm) analytical column from Phenomenex. Standard conditions for analytical RP-HPLC consisted on a linear gradient from 30% to 95% of solvent B for 30 min at a flow rate of 1 mL/min (A: water with 0.1% TFA, B: acetonitrile with 0.1% TFA). Compounds were detected by UV absorption at 222, 254 and 310 nm. High-Performance Liquid Chromatography (HPLC) was performed using an Agilent 1100 series Liquid Chromatograph Mass Spectrometer system. Analytical HPLC was run using a Luna C18 (250 x 4.60 mm) reverse phase analytical column; compounds were detected by UV absorption at 222, 254 and 310 nm. The purification of the peptides was performed on a Luna C18 (250 x 10 mm) semi-preparative reverse phase column from Phenomenex. The standard gradient used for analytical and semi-preparative HPLC was 70:30 to 5:95 over 30 min (water/acetonitrile, 0.1% TFA). Compounds were detected by UV absorption (222 nm) and by ESI^+−MS. The fractions containing the products were freeze-dried, and their identity was confirmed by ESI^+−MS and MALDI-TOF. Electrospray Ionization Mass Spectrometry (ESI/MS) was performed with an Agilent 1100 Series LC/MS model in positive scan mode using direct injection of the purified peptide solution into the MS. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI/MS) was performed with a Bruker Autoflex

MALDI/TOF model in positive scan mode by direct irradiation of the matrix-absorbed peptide. Luminescence experiments were made with a Jobin-Yvon Fluoromax-3 fluorescence spectrometers (DataMax 2.20), coupled to a Wavelength Electronics LFI−3751 temperature controller. All measurements were made with a Hellma semi-micro cuvette (114F-QS) at 20 ºC. CD experiments were made with a *Jasco J-715* coupled to a *Neslab RTE-111* thermostarted water bath at 20ºC. UV-vis absorption experiments were performed in a *Jasco V-630* spectrophotometer coupled to a *Jasco ETC-717* temperature controller at 20ºC.

Synthetic procedures

Synthesis of the unnatural coordinating residue Fmoc-βAla-bpy-OH (1)

The coordinating residue Fmoc-βAla-bpy-OH (1) was synthesized following a procedure previously reported by our research group.[100]

Synthesis of peptide ligands and metallopeptides

Synthesis of the peptide ligands 2 and 2-R_8

C-terminal amide peptides were synthesized by following standard SPPS protocols on a 0.1 mmol scale using a 0.45 mmol/g Fmoc-Rink-amide resin. Arginines were coupled, in 10-fold excess (vs. mmol of resin load), by using *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU) as an activating agent.

Fmoc-βAlaBpy-OH (**1**) was coupled in 5-fold excess using O-(7-Azabenzotriazol-1-yl)-*N,N,N',N'*- tetramethyluronium hexafluorophosphate (HATU) as activating agent.

Couplings were conducted for 1 h. Deprotection of the temporal Fmoc protecting-group was performed with 20% piperidine in DMF for 15 min.

Test cleavages were performed at a 1 mg scale for 2 h with CH_2Cl_2 (50 µL), H_2O (25 µL), TIS (triisopropylsilane, 25 µL), and TFA (900 µL) (~1 mL of cocktail for 20 mg of resin).

Synthesis of the Ru(II) metallopeptides 3-6

Common step

Once the peptide ligands **2** and **2-R_8** were synthesized, 222 mg of the resin with the corresponding peptide anchored was suspended in 3 mL of EtOH:DMF (1:1) in the dark and the resulting mixture were purged with Ar for 15 min. 54.6 mg (1.3 eq) of $[Ru(DMSO)_4Cl_2]$[102] was added and the resulting mixture was stirred under argon for 24 hours at 80 ºC. Then, the resin was washed with DMF (5 x 10 mL, 10 min) and dried under vacuum.

Phen-dppz Ru(II) metallopeptides (3 and 4)

The resin was then suspended in 3 mL of EtOH:DMF (1:1) in the dark and the resulting mixture were purged with Ar for 15 min. 18.0 mg (1.0 eq) of 1,10'-phenantroline was added and the resulting mixture was stirred under argon for

24 hours at 80 ºC. Then, the resin was washed with DMF (5 x 10 mL, 10 min) and dried under vacuum.

Finally, the resin was suspended in 3 mL of EtOH:DMF (1:1) in the dark and the resulting mixture were purged with Ar for 15 min. 28.2 mg (1.0 eq) of dipyrido[3,2-a:2',3'-c]phenazine (dppz)[17] was added and the resulting mixture was stirred under argon for 24 hours at 80 ºC. Then, the resin was washed with DMF (5 x 10 mL, 10 min) and dried under vacuum.

Double-dppz Ru(II) metallopeptides (5 and 6)

The resin was then suspended in 3 mL of EtOH:DMF (1:1) in the dark and the resulting mixture were purged with Ar for 15 min. 56.5 mg (2.0 eq) of dipyrido[3,2-a:2',3'-c]phenazine (dppz)[103] was added and the resulting mixture was stirred under argon for 24 hours at 80 ºC. Then, the resin was washed with DMF (5 x 10 mL, 10 min) and dried under vacuum.

General procedure for peptide cleavage-deprotection

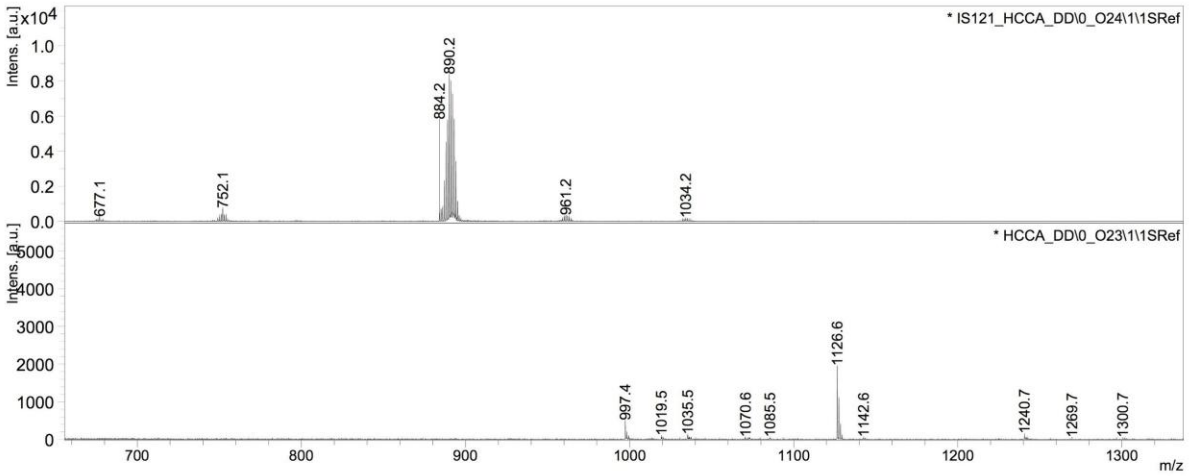
The resin was filtered, washed with DMF and CH_2Cl_2 and dried. The metallopeptide was cleaved with 5 mL of the standard TFA cocktail (TIS, H_2O, CH_2Cl_2 and TFA) over 2.5 hours. After that, the resin was filtered and washed with TFA (1 x 2 mL) and the filtrate was concentrated until 1 mL of volume with a N_2 stream. Then, 4 mL of H_2O and 5 mg of NH_4PF_6 were added to this solution. The red orange solid precipitated was separated by centrifugation, washed with H_2O (1 x 4 mL) and purified by semi-preparative HPLC to give the desired product.

[Ru(DMSO)Cl_2] and dipyrido[3,2-a:2',3'-c]phenazine (dppz) were synthesized following reported procedures.[102,103]

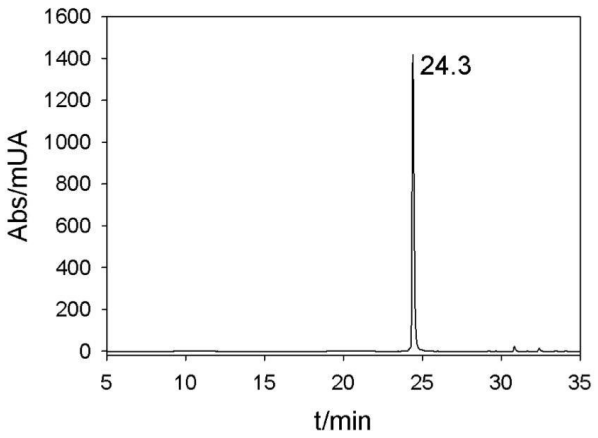
Mass spectra and HPLC chromatograms of the Ru(II) metallopeptides (3 to 6)

a) $[Ru(2)(phen)(dppz)]^{2+}$ (3)

MALDI-TOF: m/z calculated for $C_{43}H_{35}N_{11}O_3Ru$: 891.2, found: 890.2

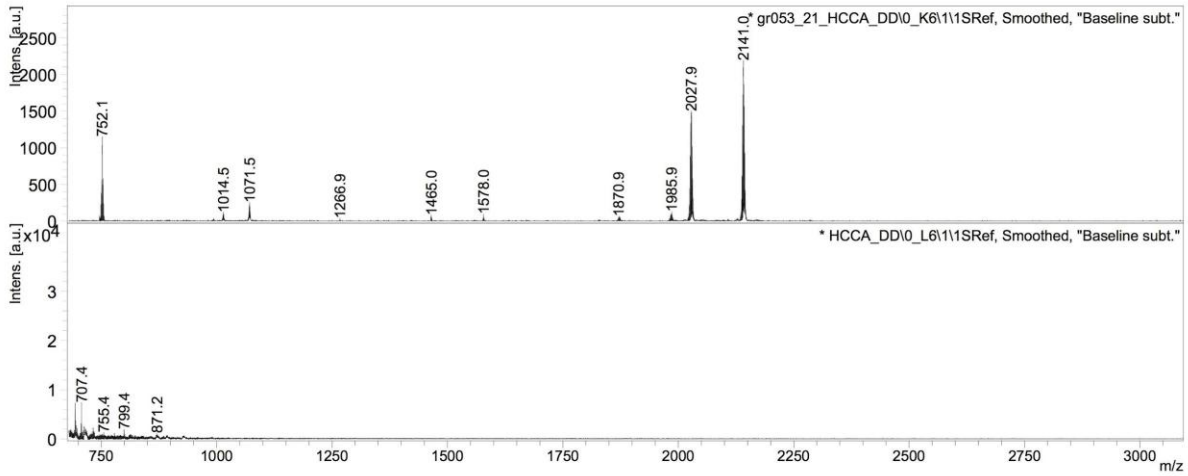


HPLC chromatogram: 1-75 %B, t_R=24.3´

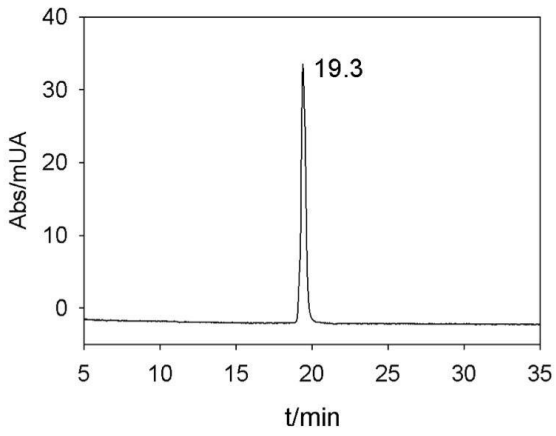


b) [Ru(2-R_8)(phen)(dppz)]$^{2+}$ (4)

MALDI-TOF: m/z calculated for $C_{94}H_{131}N_{43}O_{11}Ru$: 2140.0, found: 2141.1

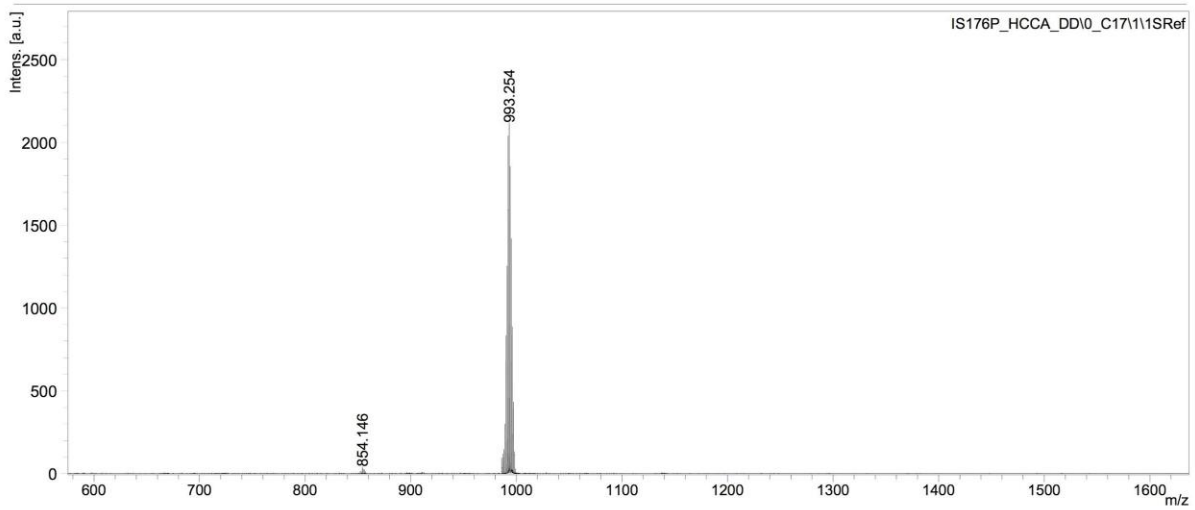


HPLC chromatogram: 1-75 %B, t_R=19.3'

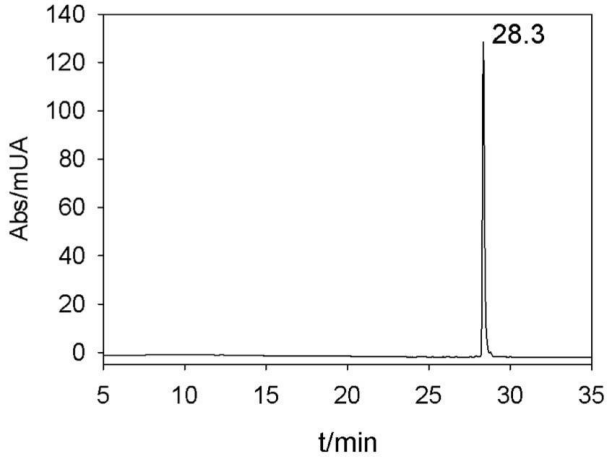


c) $[Ru(2)(dppz)_2]^{2+}$ (5)

MALDI-TOF: m/z calculated for $C_{52}H_{37}N_{13}O_3Ru$: 993.2, found: 993.2

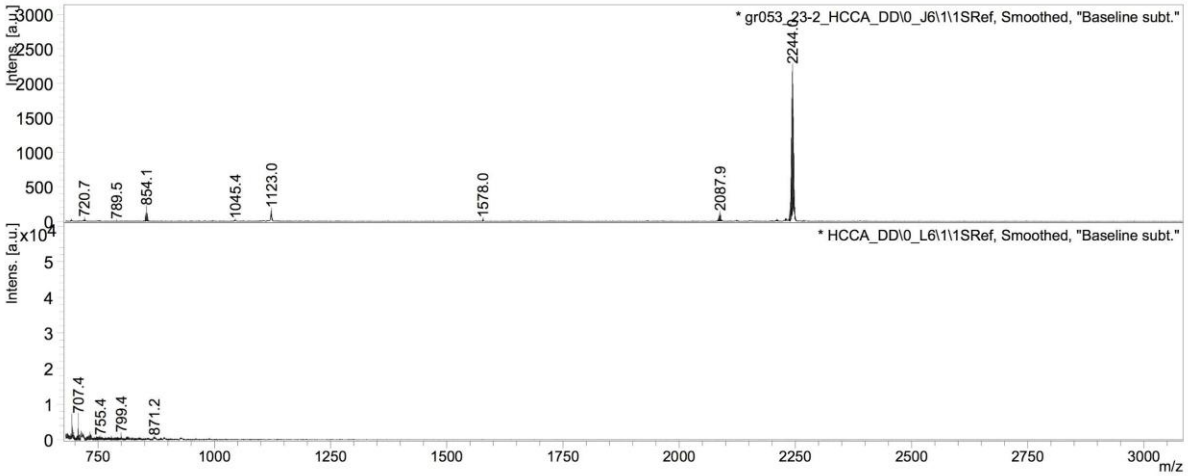


HPLC chromatogram: 1-75 %B, t_R=28.3´

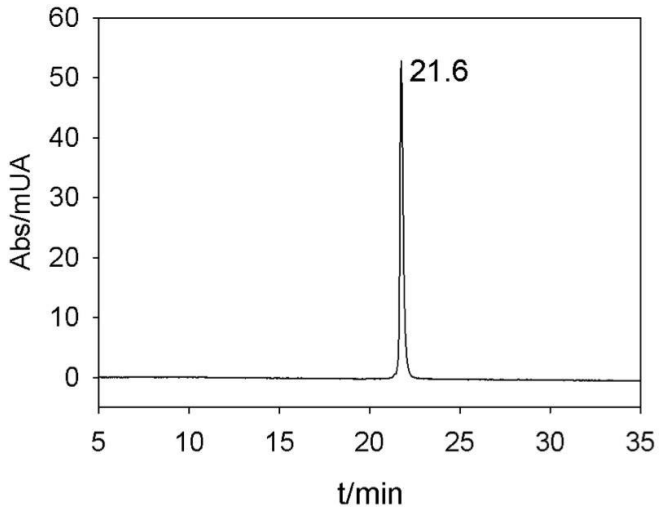


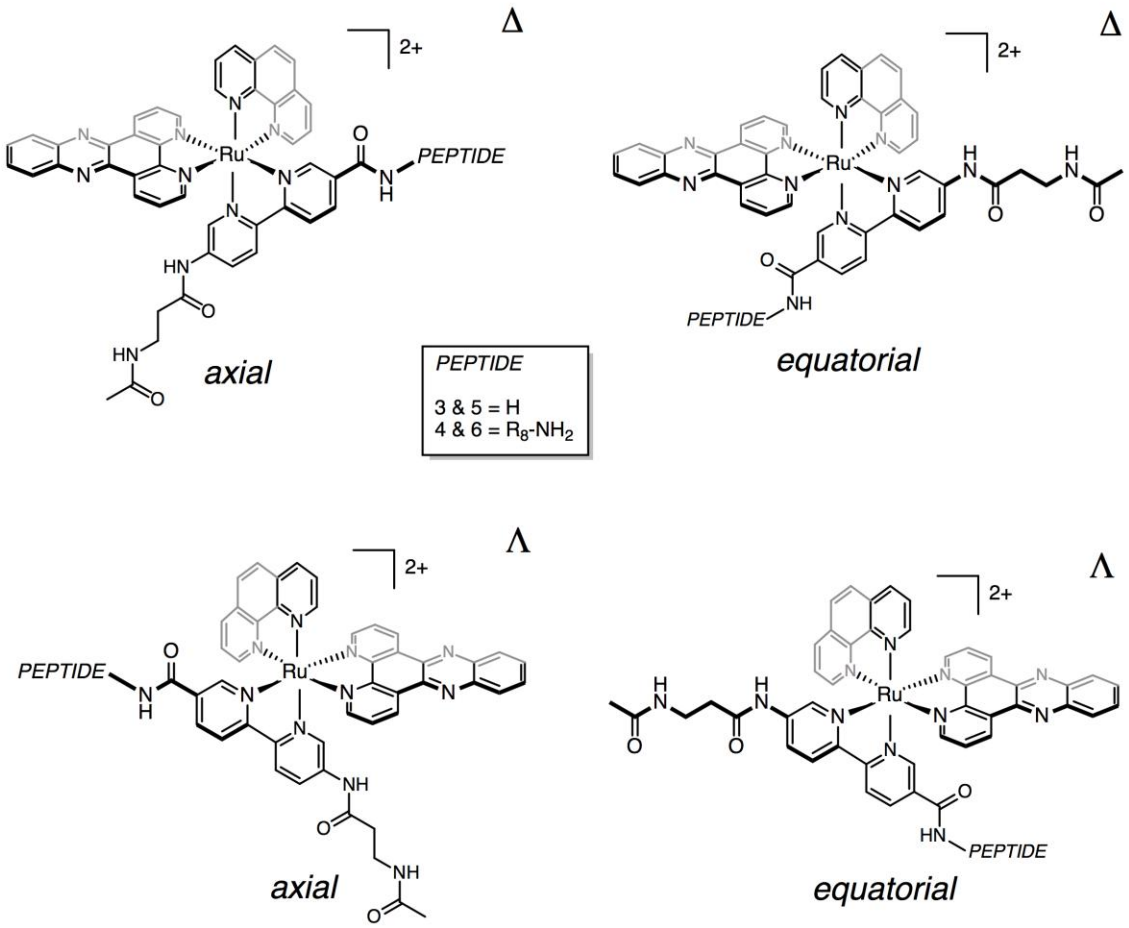
d) $[Ru(2\text{-}R_8)(dppz)_2]^{2+}$ (6)

MALDI-TOF: m/z calculated for $C_{100}H_{133}N_{45}O_{11}Ru$: 2242.0, found: 2244.0



HPLC chromatogram: 1-75 %B, t_R=21.6





Scheme S1. The isomers of the ruthenium-peptide conjugates. The ruthenium(II) metallopeptides have four isomers; the ruthenium center can be Δ or Λ, and the peptide linker can be axial or equatorial to the dppz ligand.

Fluorescence studies

a) Hairpin "well-matched" B-DNA binding studies

To a 2.0 μM solution of the selected Ru(II) metallopeptide (**3**-**6**) in Tris-HCl buffer (20 mM), pH 7.5 and NaCl (100 mM), aliquots of the selected hairpin B-DNA stock solution (in water) were added and the fluorescence spectrum was recorded after each addition. The additions were carried out until no further changes in the emission spectra were detected. The studied B-DNA oligonucleotides are listed in table S1.

In the case of metallopeptides **3** and **5** we were able to calculate the corresponding K_d (see table S2), The calculated values are very similar to those reported in literature by similar complexes.[4] However, the profiles of the titration experiments of **4** and **6** at $\lambda_{em} = 630\ nm$ suggest the existence of complicated two-step processes. As a consequence, it has not been possible to calculate the dissociation constants in these cases. Further experiments will be needed in order to study conveniently these unusual behaviours.

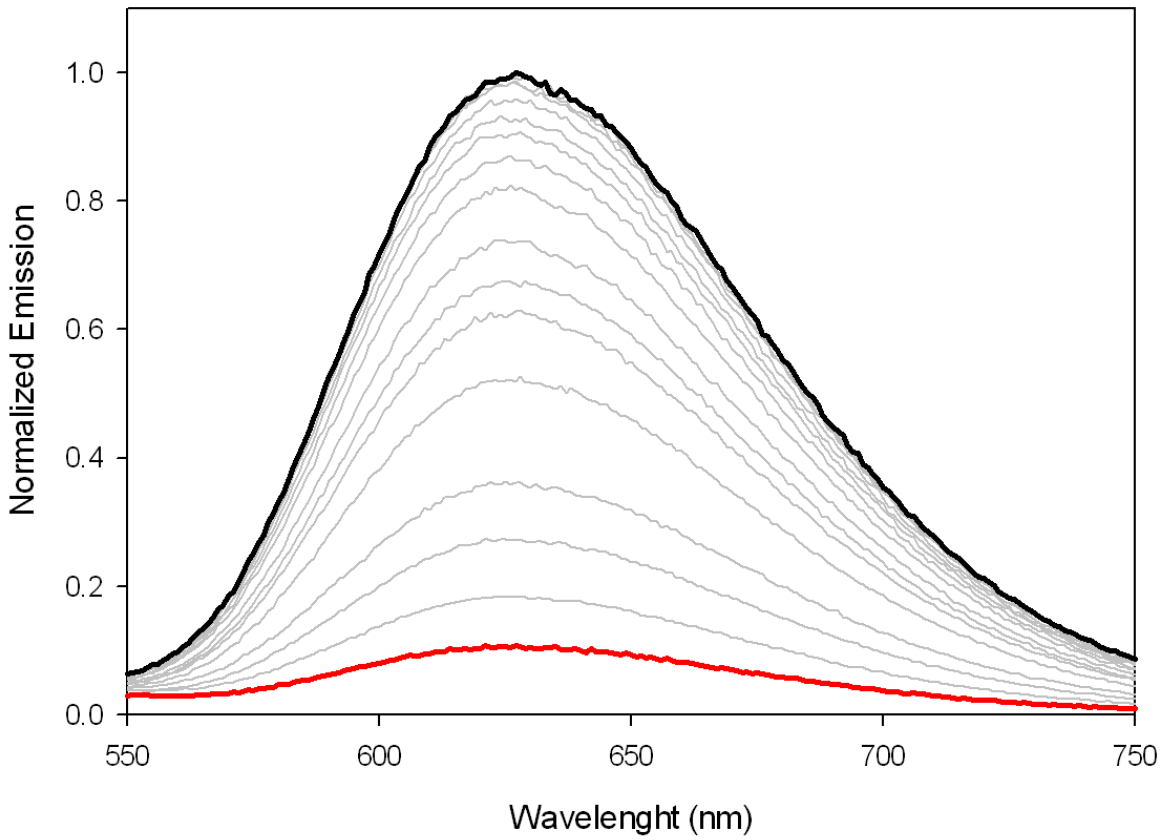


*Figure S1. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AAAATTT oligonucleotide solution until saturation.*

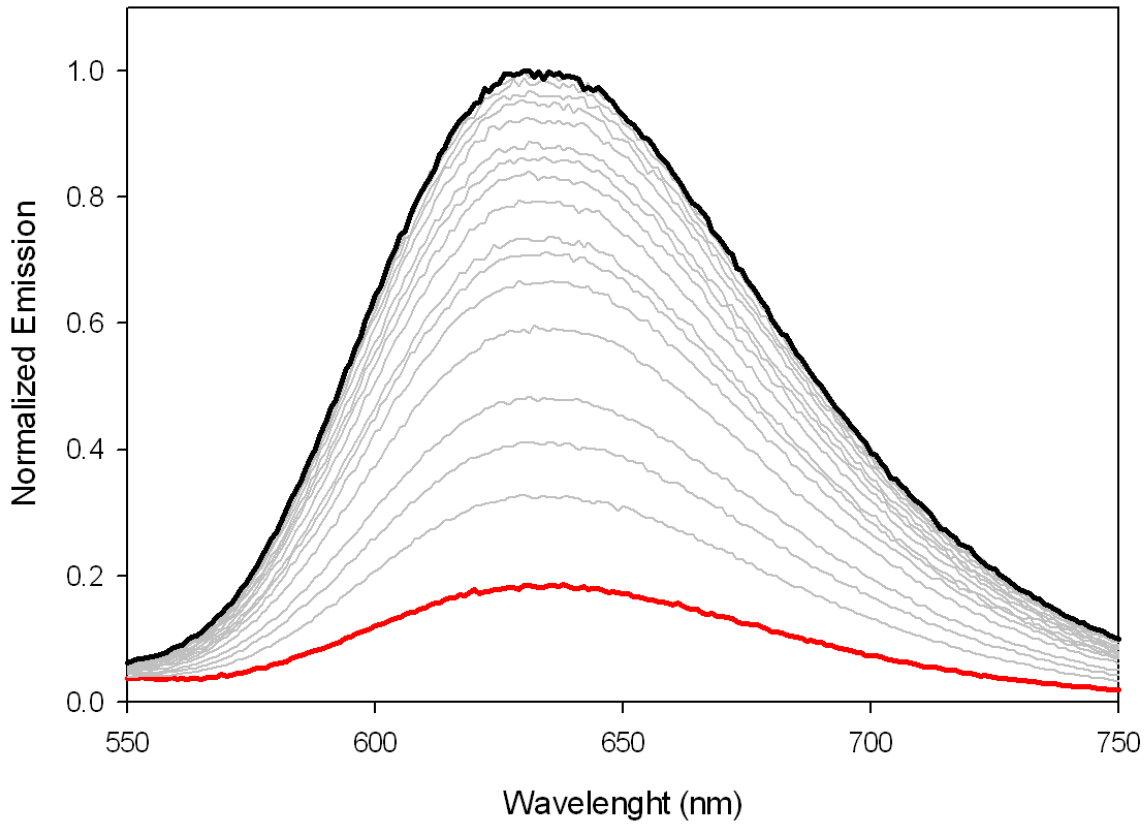


*Figure S2. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GAATTC oligonucleotide solution until saturation.*

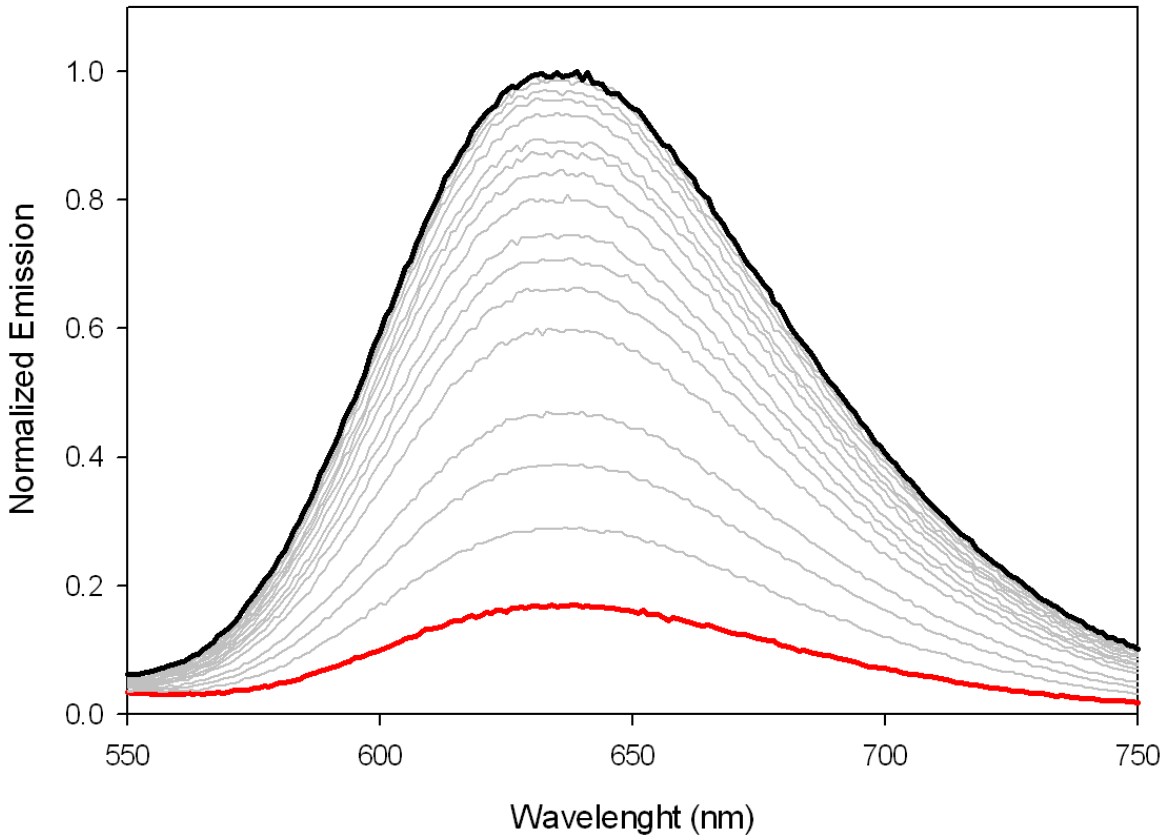


*Figure S3. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GGCCC oligonucleotide solution until saturation.*

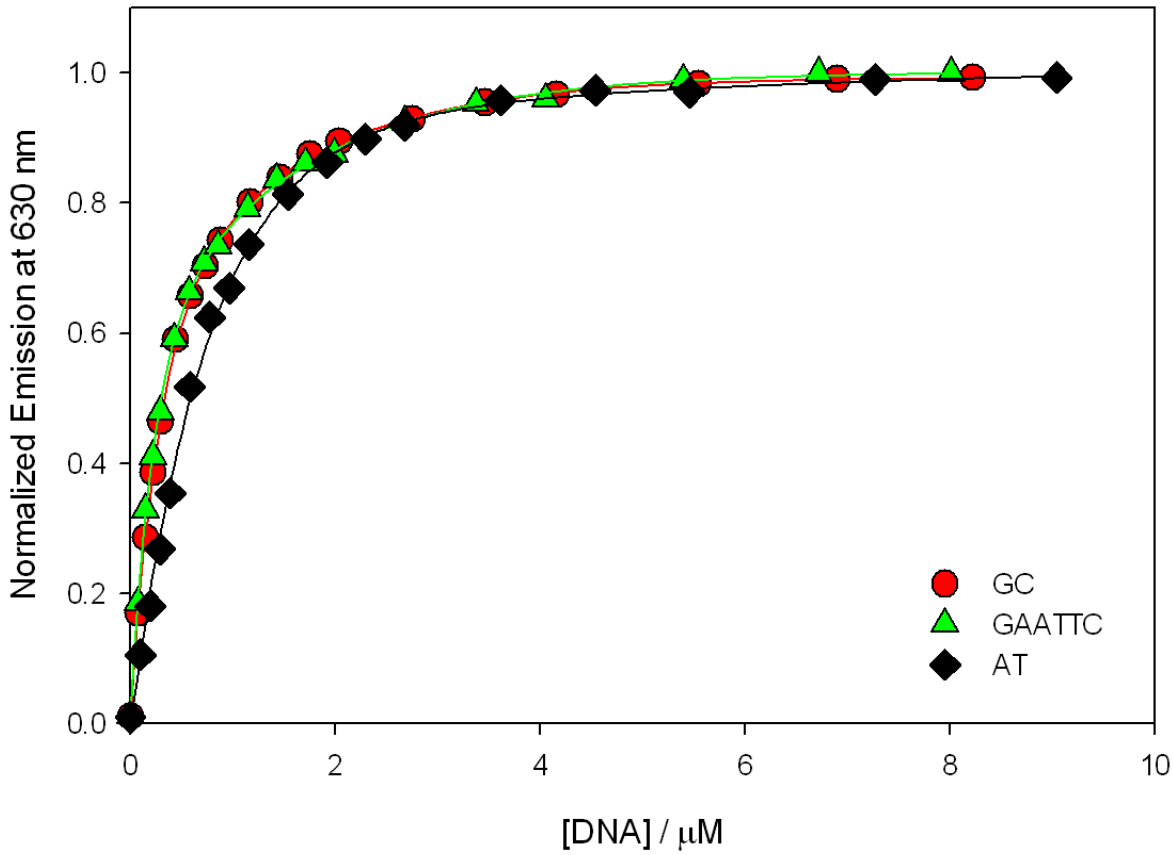


Figure S4. Profile of the titration experiments of ***3*** *with AAAATTT (black diamonds), GAATTC (green triangles) and GGCCC (red circles) oligonucleotides at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (black, green and red lines, respectively).*

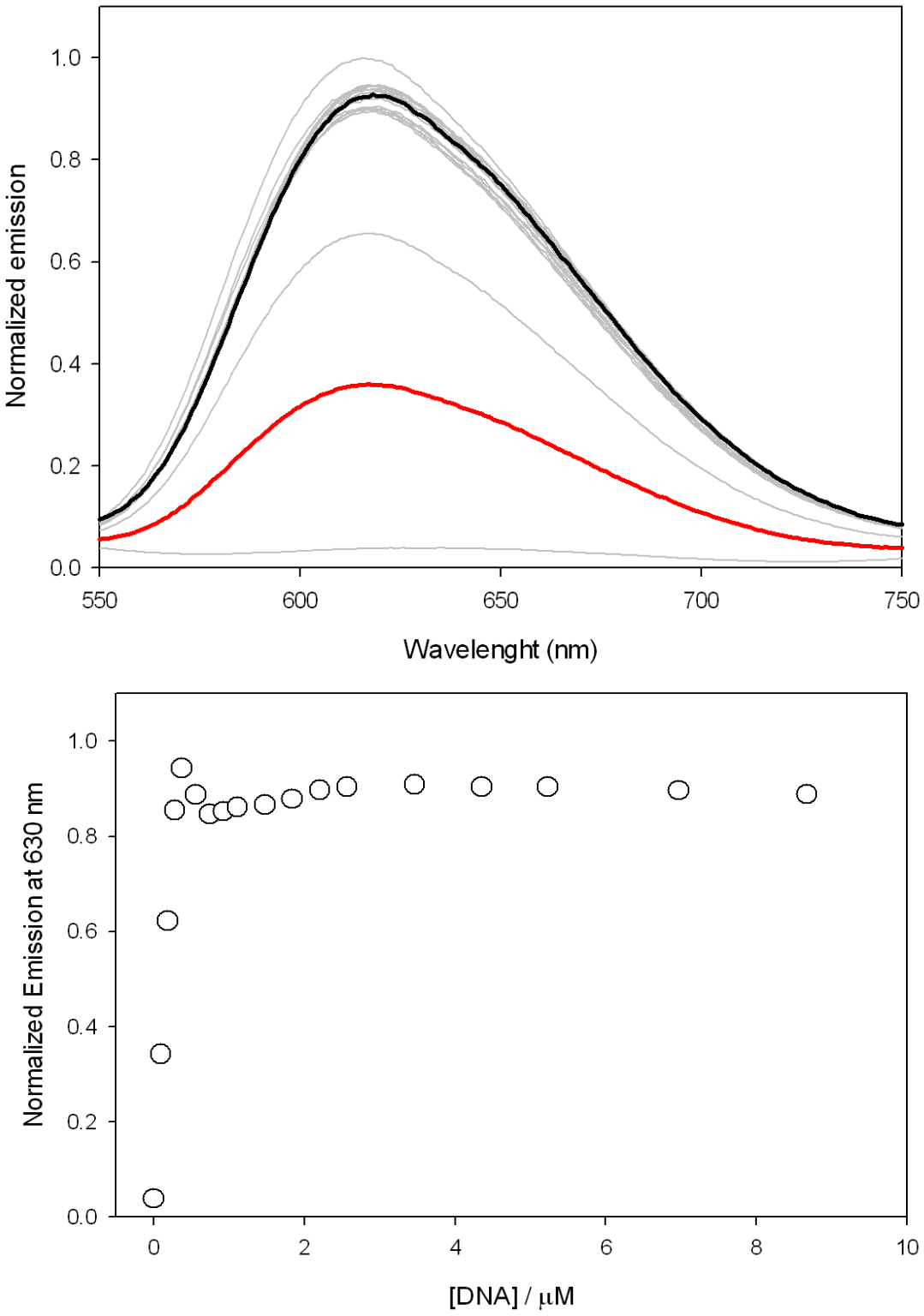


Figure S5. Top, luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AAAATTT oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of 4 with AAAATTT oligonucleotide at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media).*

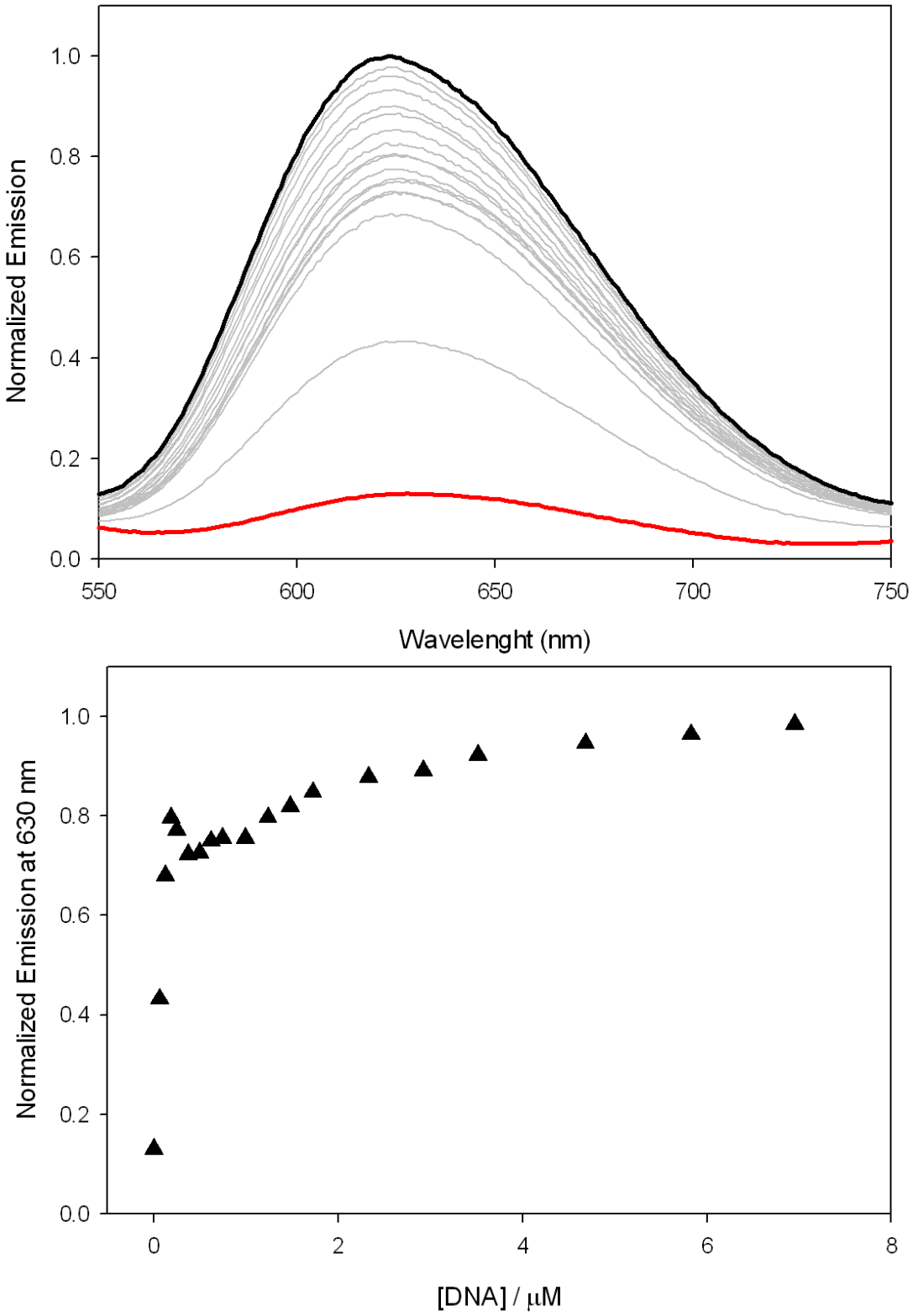


Figure S6. Top, luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GAATTC oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of 4 with GAATTC oligonucleotide at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media).*

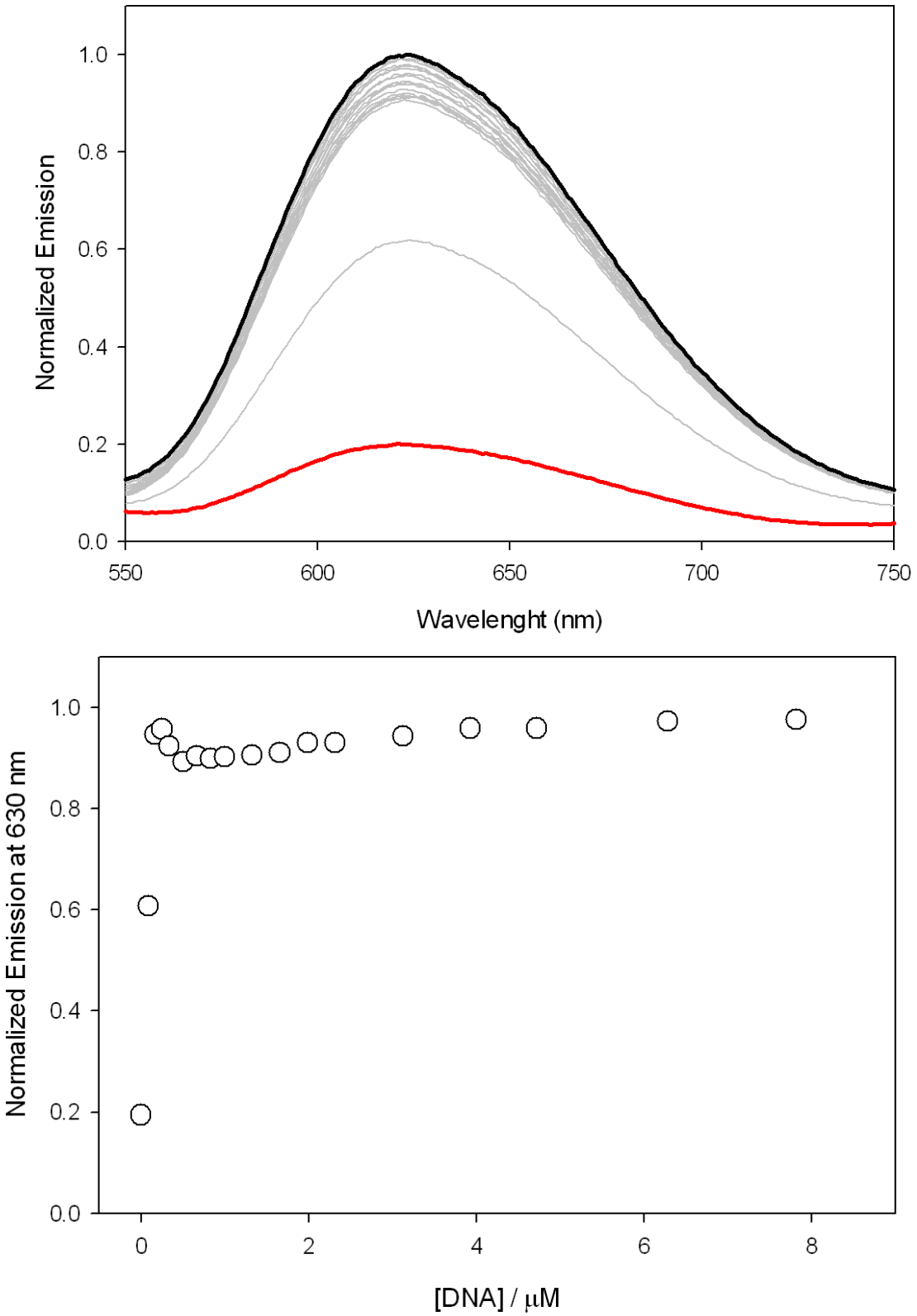


*Figure S7. Top, luminescence spectra of a 2.0 μM solution of **4** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GGCCC oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of 4 with GGCCC oligonucleotide at λ_{em} = 630 nm (emission intensity vs. concentration of DNA in the media).*

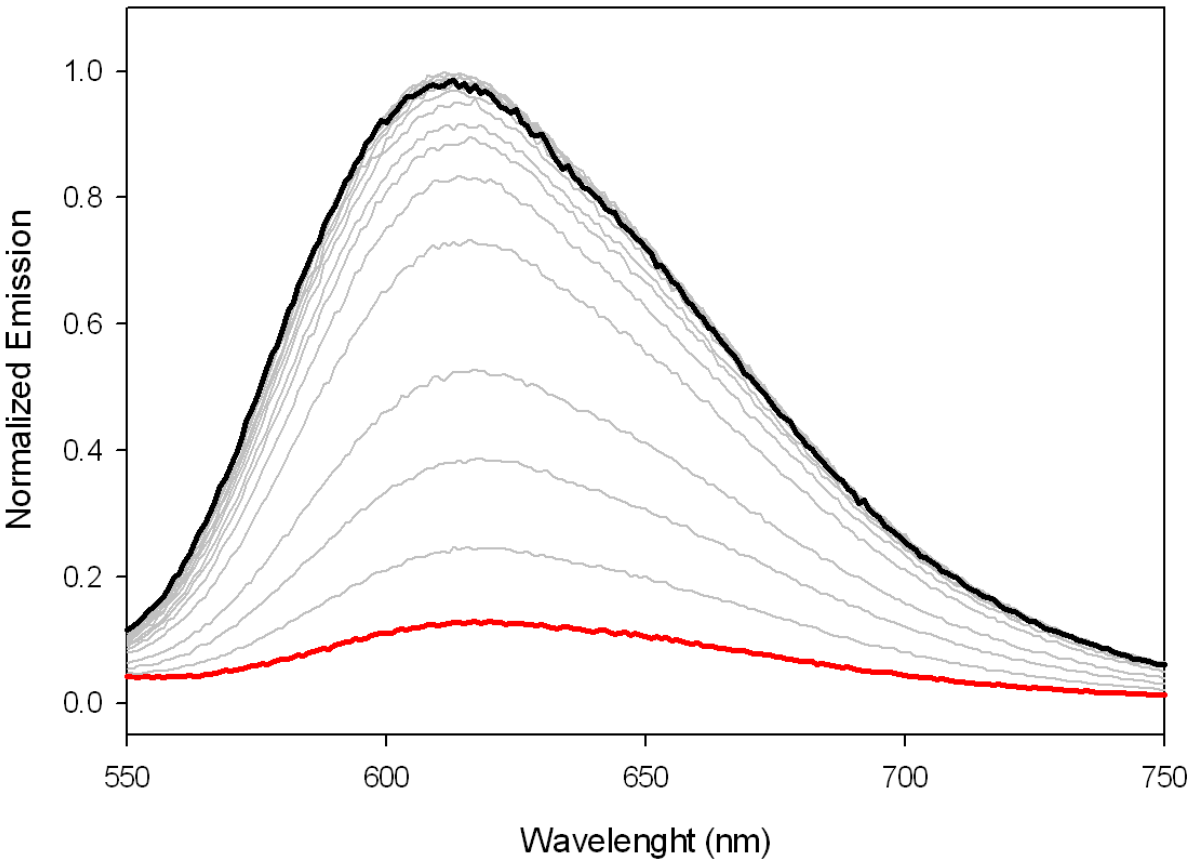


Figure S8. Luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AAAATTT oligonucleotide solution until saturation.*

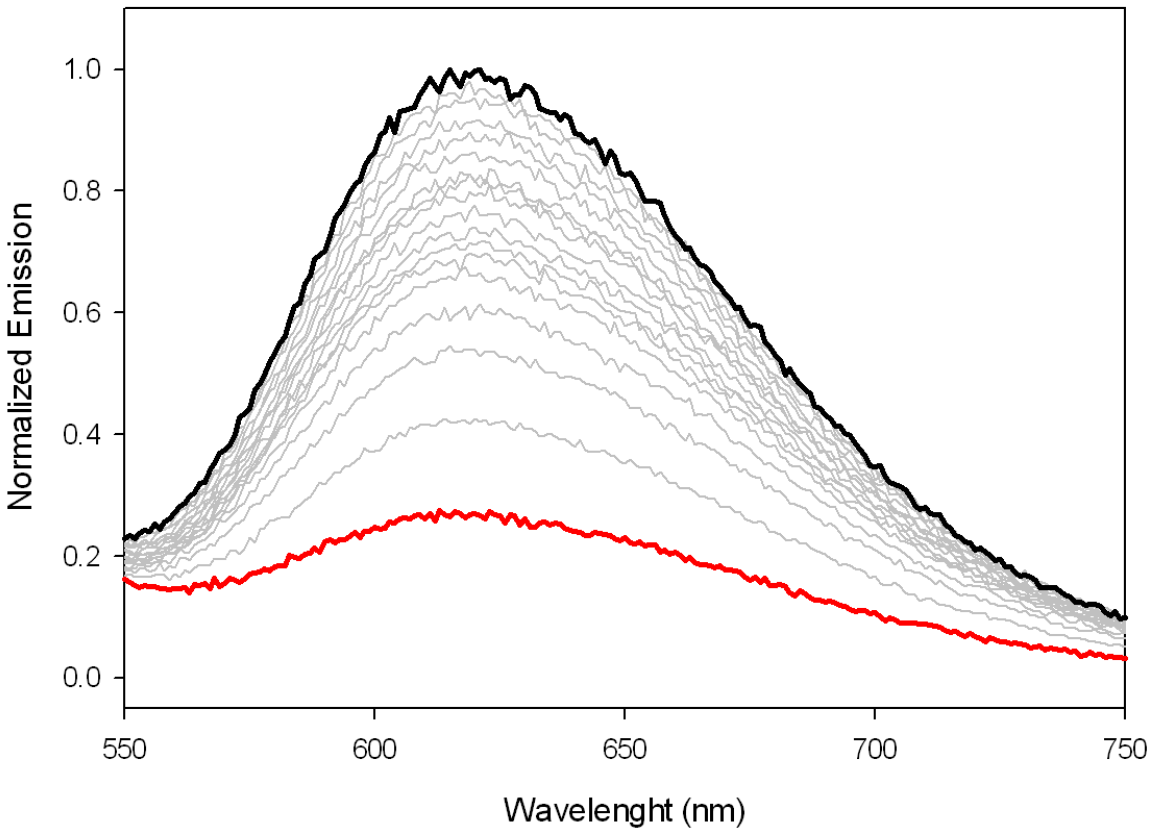


*Figure S9. Luminescence spectra of a 2.0 μM solution of **5** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GAATTC oligonucleotide solution until saturation.*

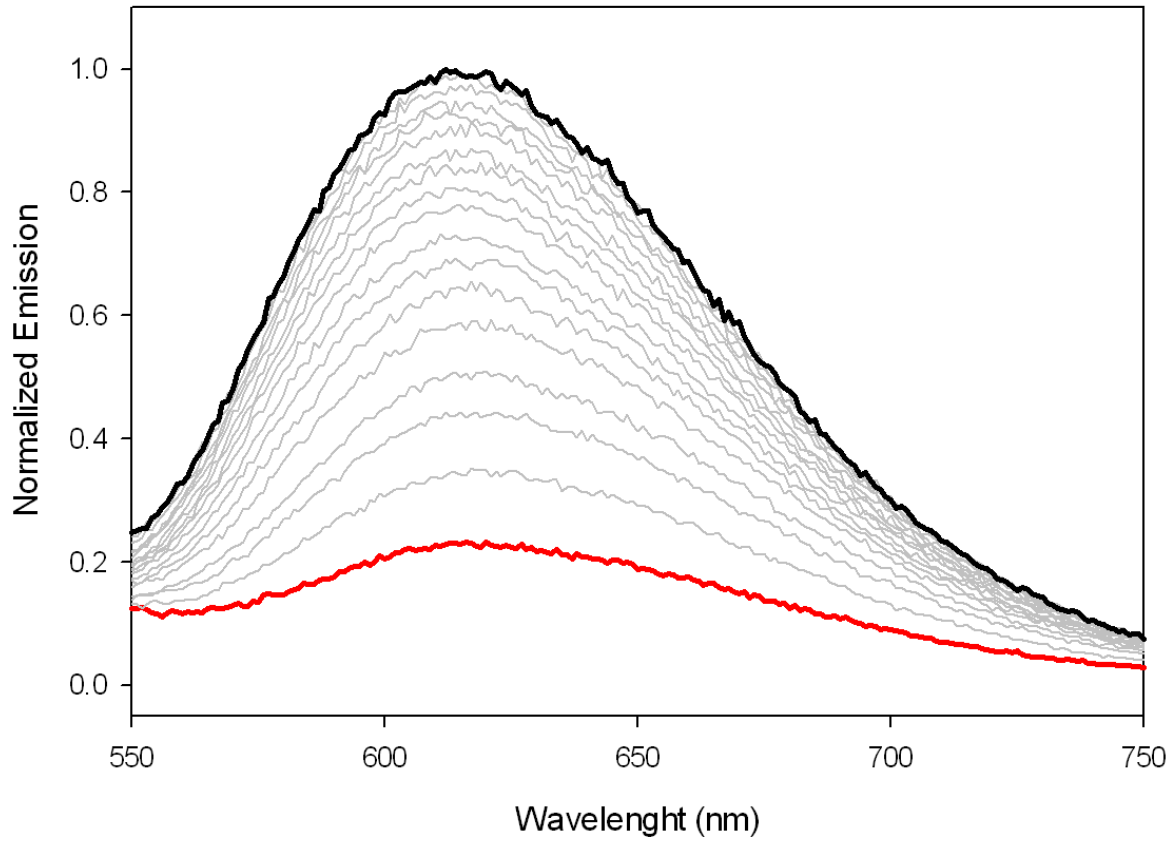


Figure S10. Luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GGCCC oligonucleotide solution until saturation.*

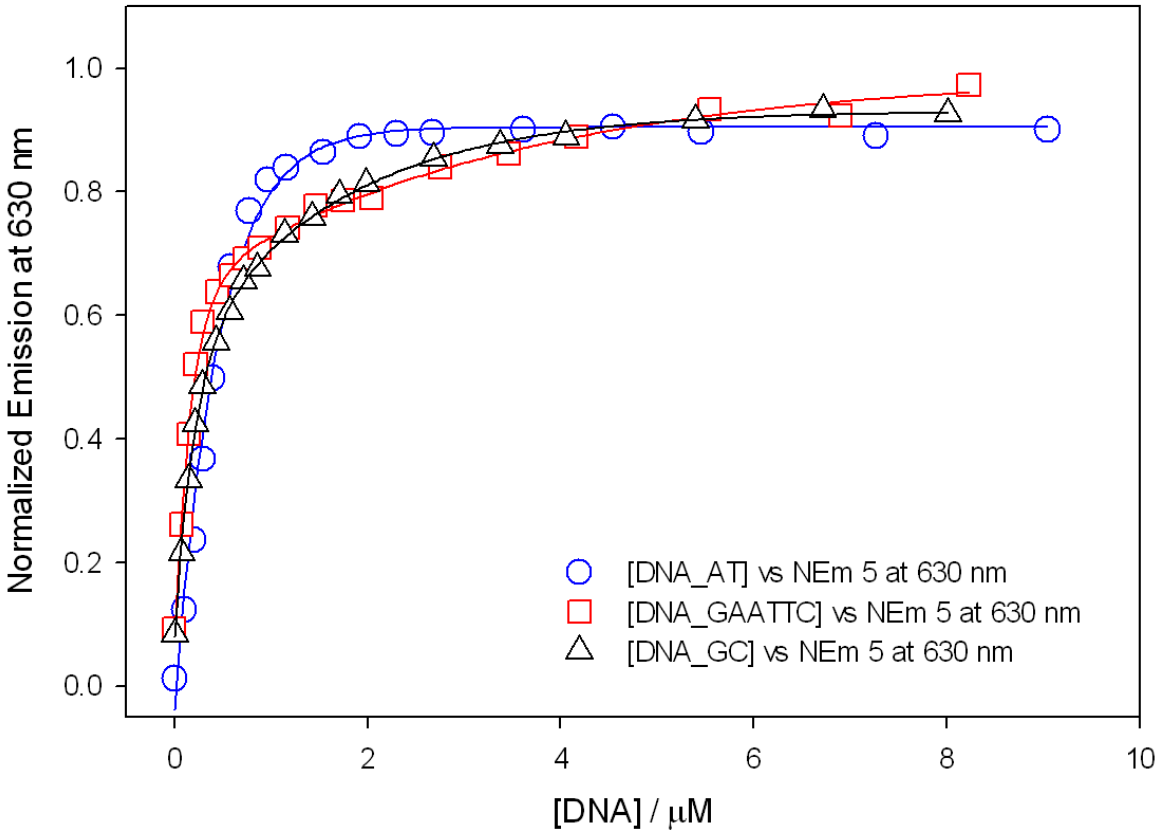


*Figure S11. Profile of the titration experiments of **5** with AAAATTT (blue circles), GAATTC (red squares) and GGCCC (black triangles) oligonucleotides at λ_{em} = 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (blue, red and black lines, respectively).*

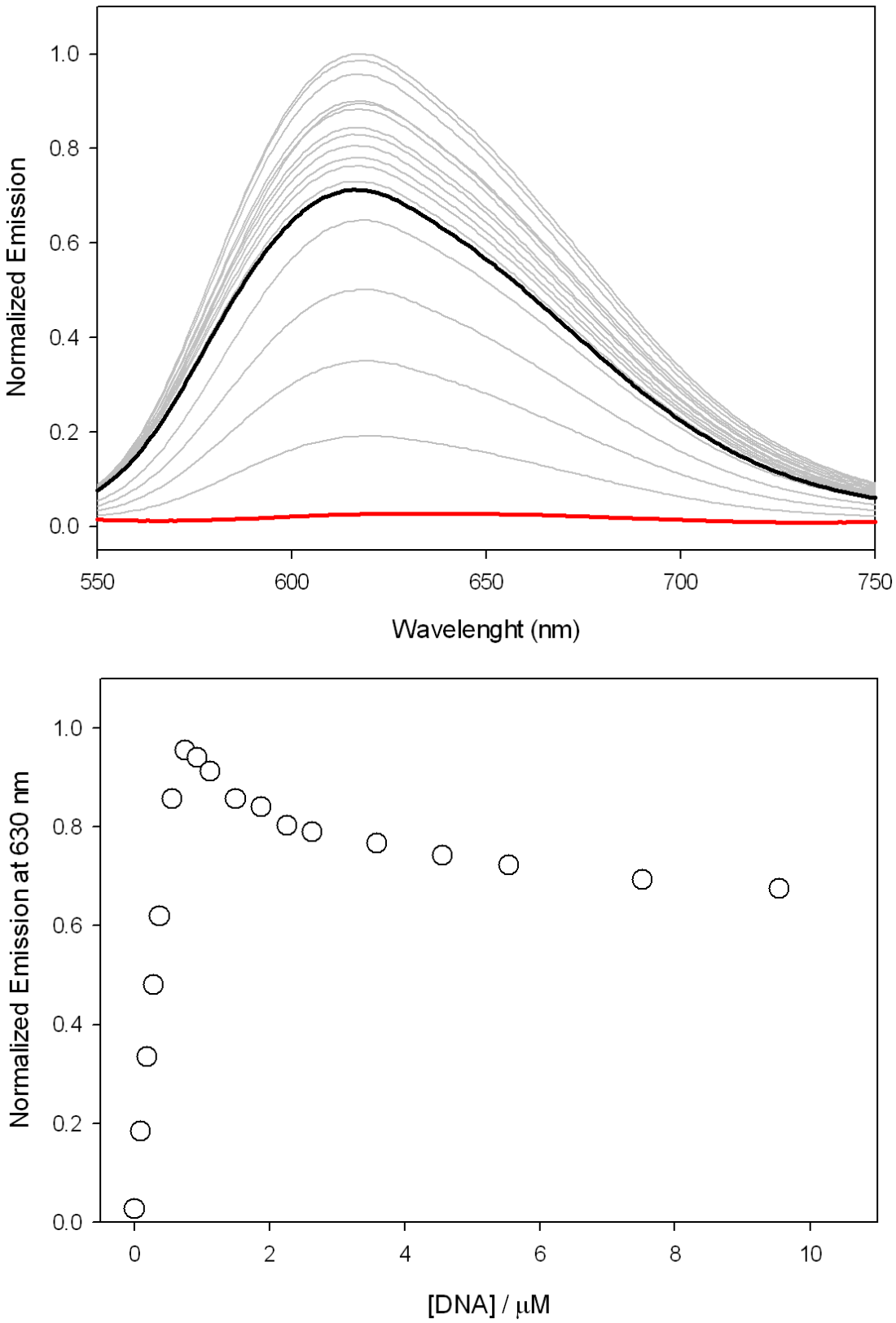


Figure S12. Top, luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AAAATTT oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of* ***6*** *with AAAATTT oligonucleotide at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media).*

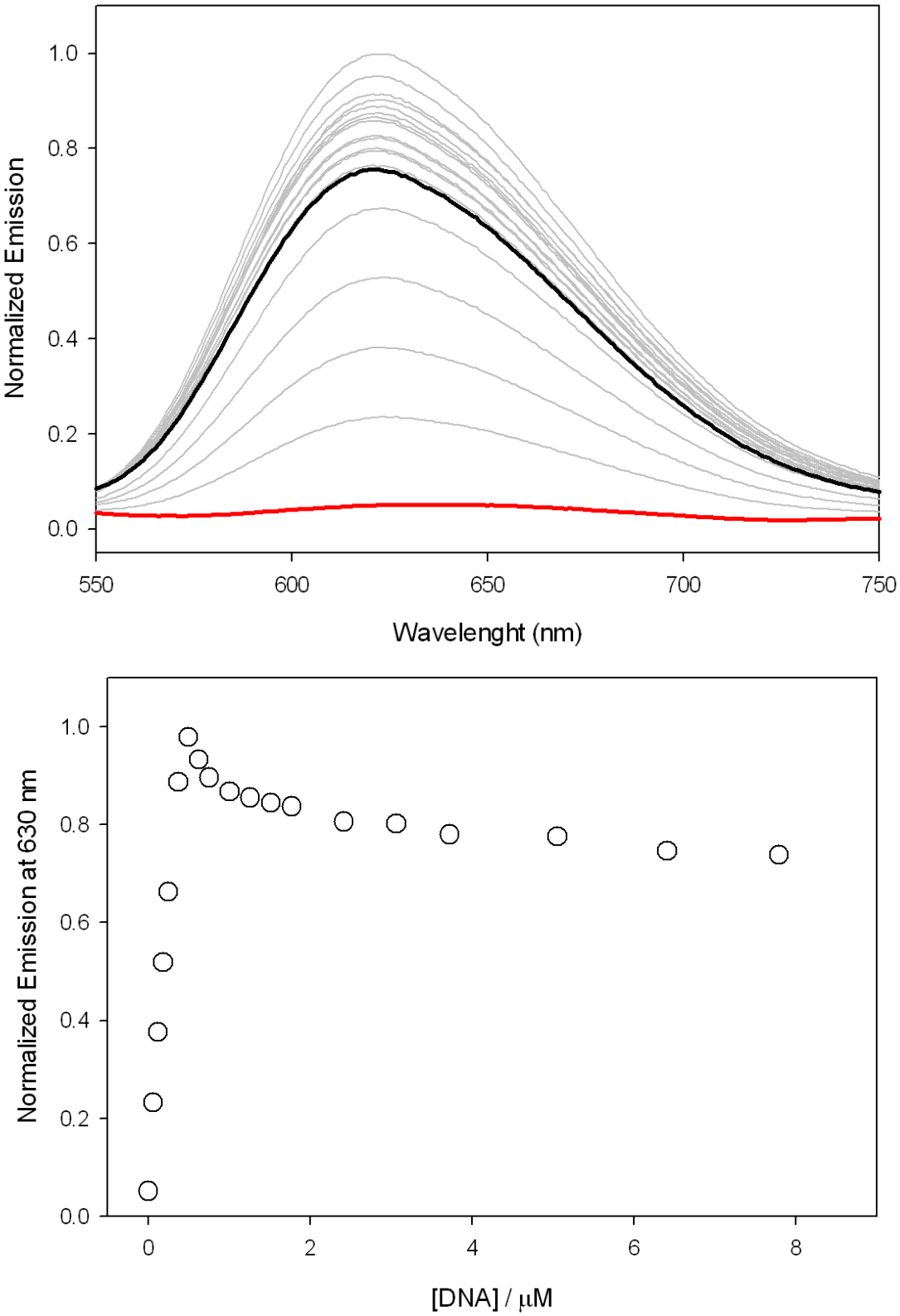


Figure S13. Top, luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GAATTC oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of* ***6*** *with GAATTC oligonucleotide at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media).*

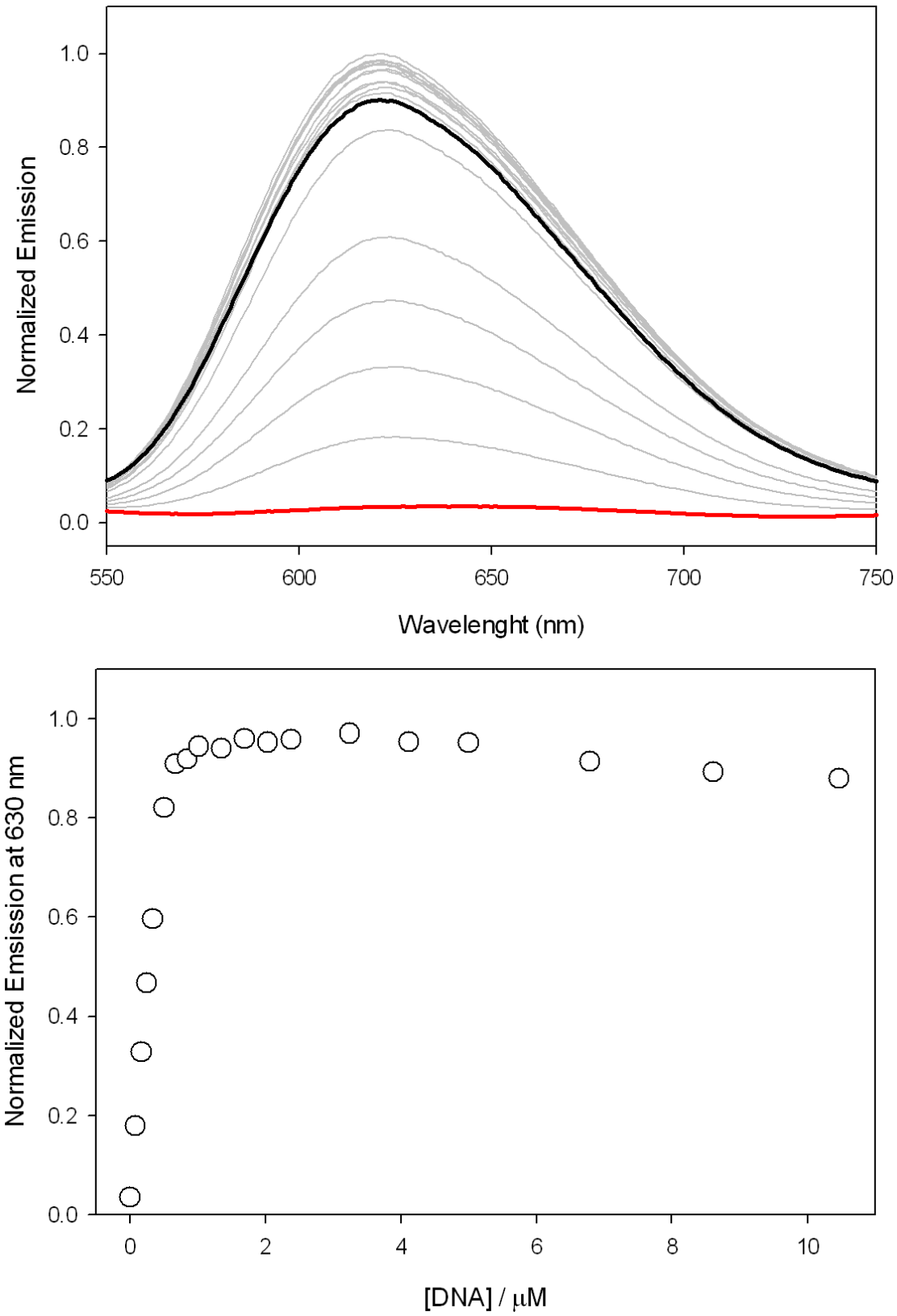


Figure S14. Top, luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GGCCC oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of* ***6*** *with GGCCC oligonucleotide at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media).*

b) Double-stranded "well-matched" and "mismatched" B-DNA binding studies

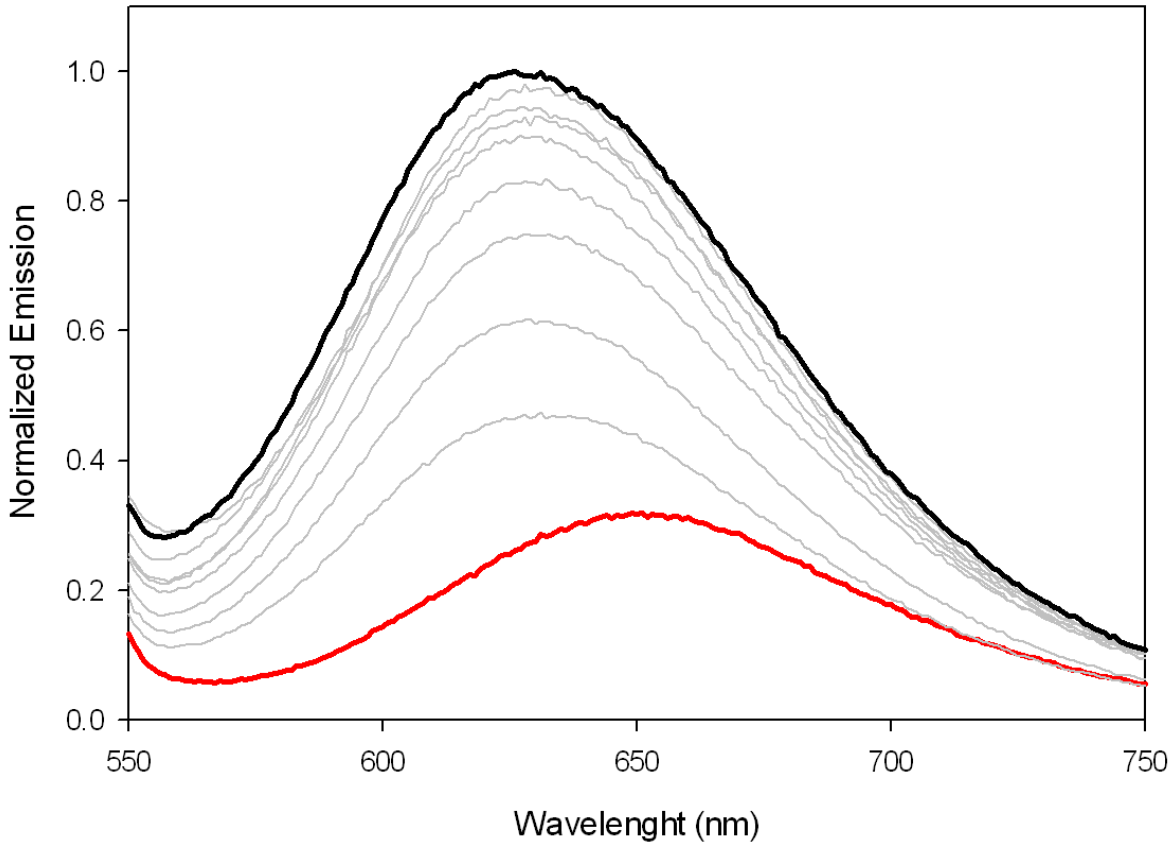


*Figure S15. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of "well-matched" A:A' oligonucleotide solution until saturation.*

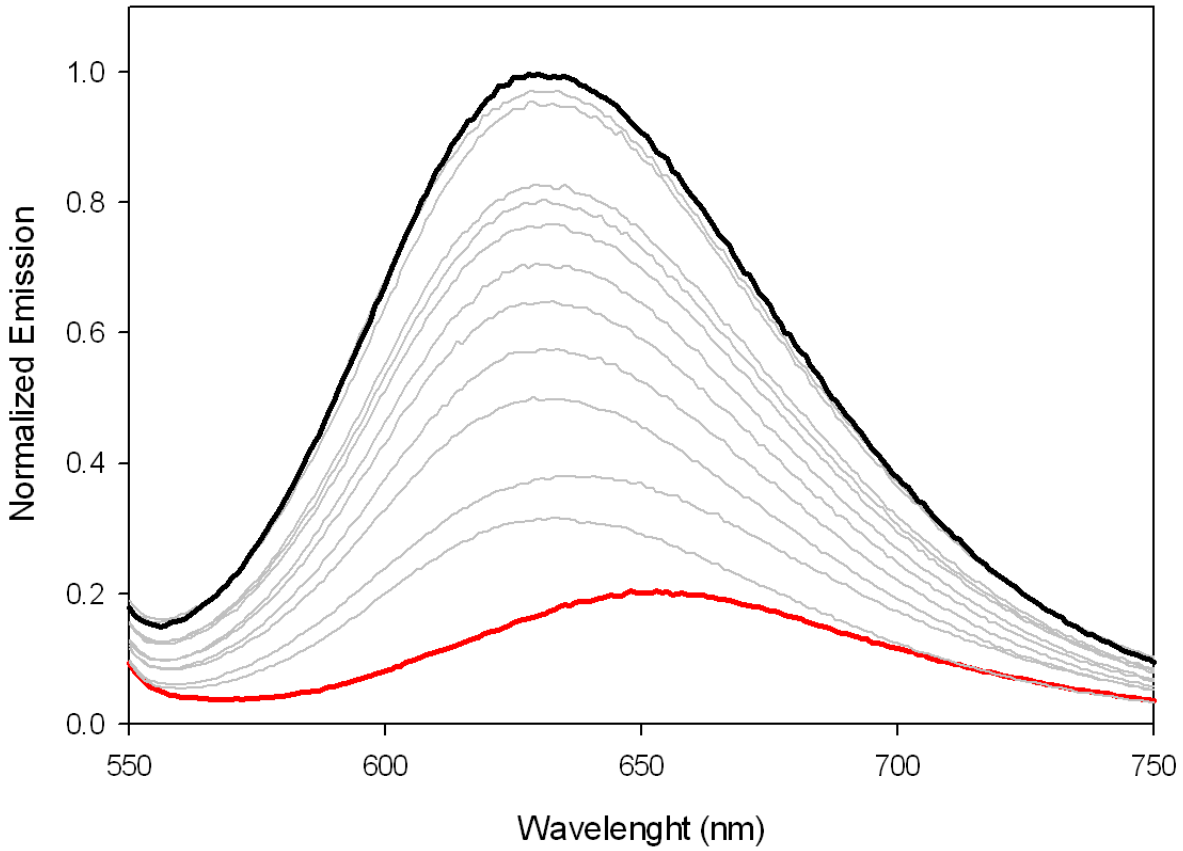


*Figure S16. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG mismatched A:A1' oligonucleotide solution until saturation.*

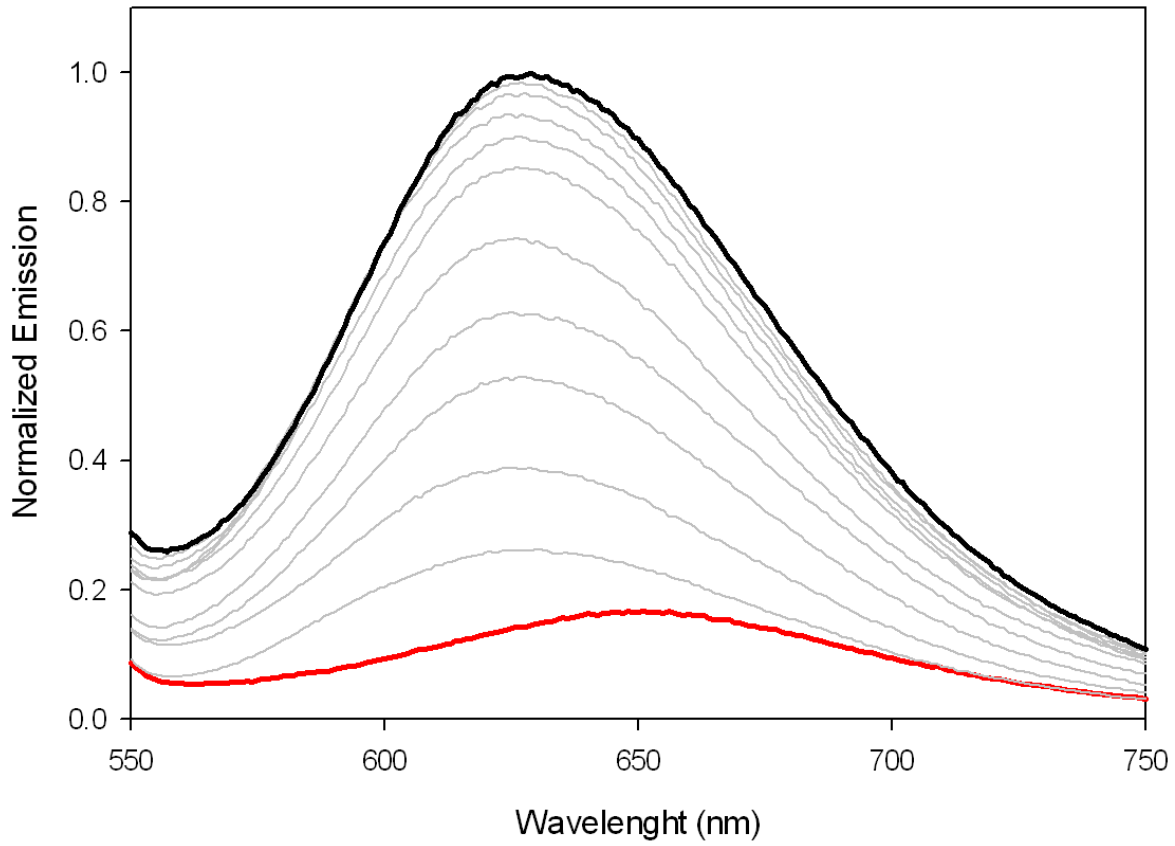


Figure S17. Luminescence spectra of a 2.0 μM solution of ***3*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/AA mismatched A:A2' oligonucleotide solution until saturation.*

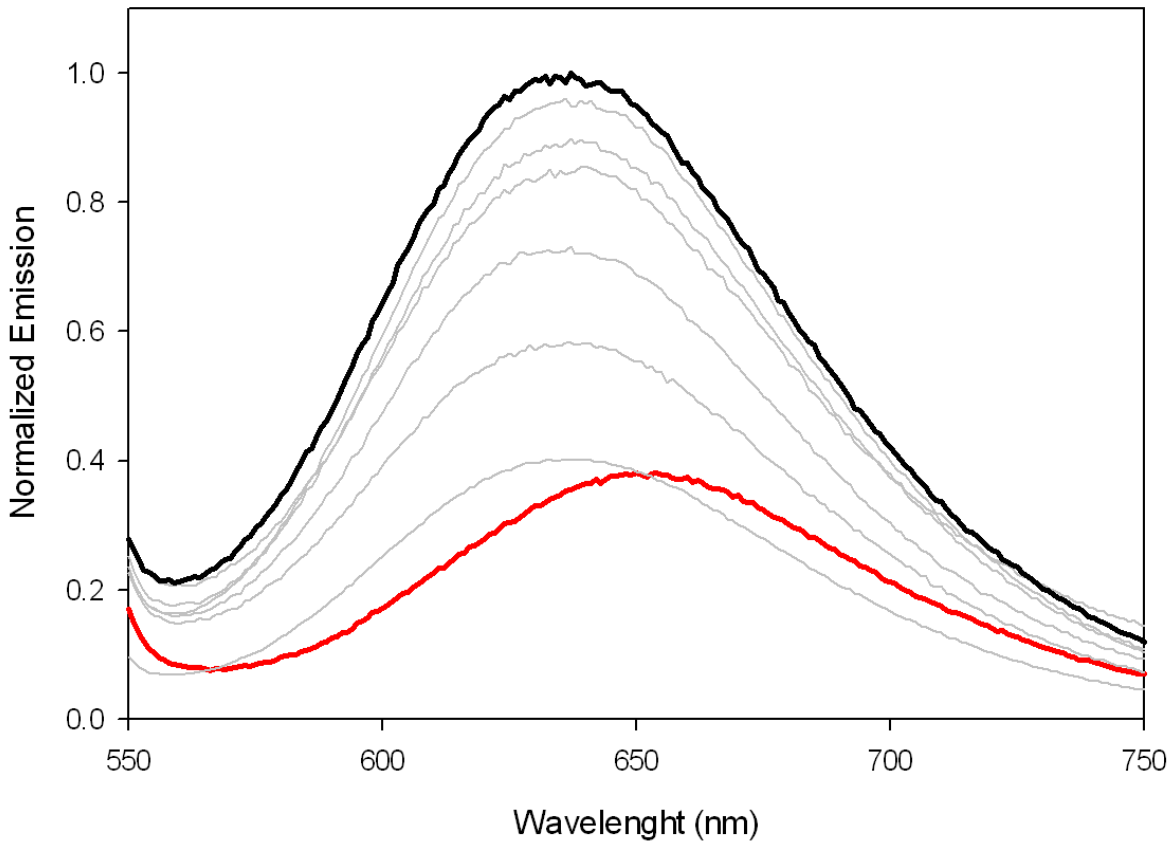


*Figure S18. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/GG/AA mismatched A:A3' oligonucleotide solution until saturation.*

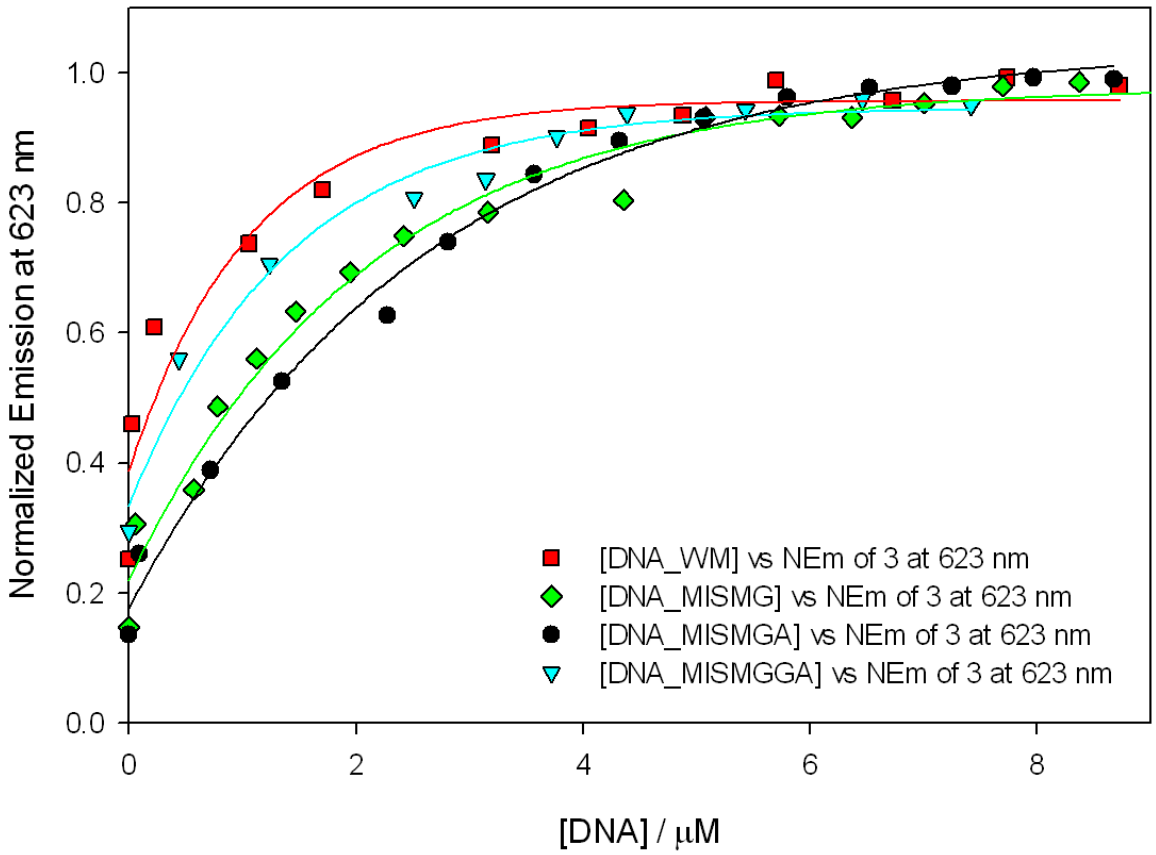


*Figure S19. Profile of the titration experiments of **3** with "well matched" (A:A'; red squares), GG mismatched (A.A1'; green diamonds), GG/AA mismatched (A:A2'; black circles) and GG/GG/AA mismatched (A:A3'; cyan triangles) oligonucleotides at λ_{em} = 623 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (red, green, black and cyan lines, respectively).*

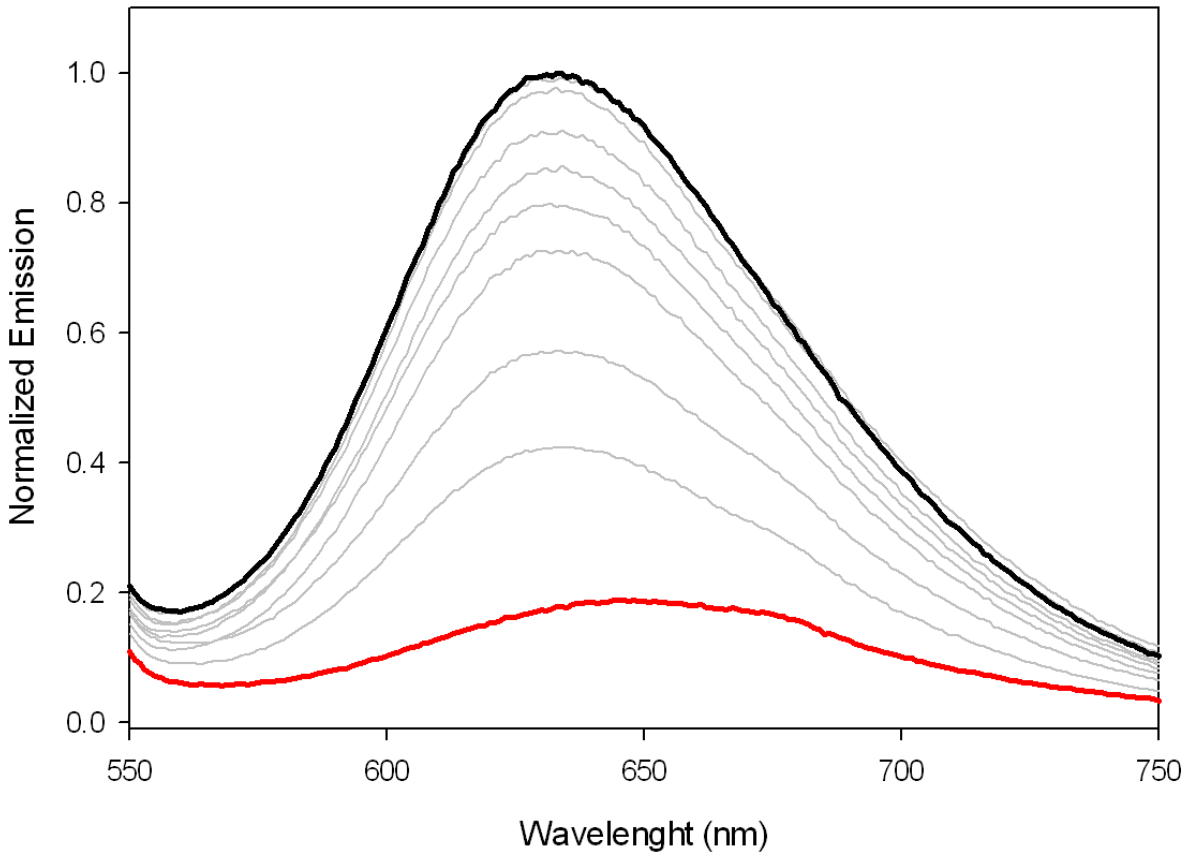


*Figure S20. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AA mismatched B:B1' oligonucleotide solution until saturation.*

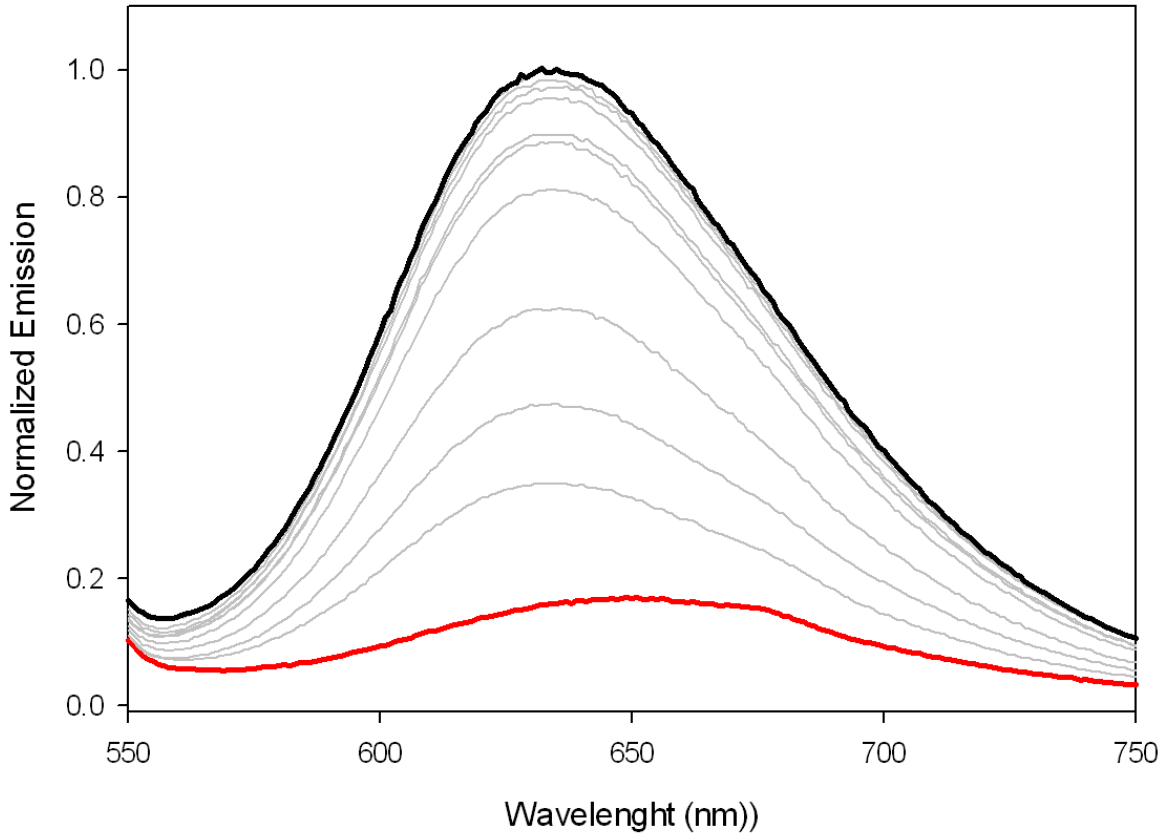


*Figure S21. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of CC mismatched C:C1' oligonucleotide solution until saturation.*

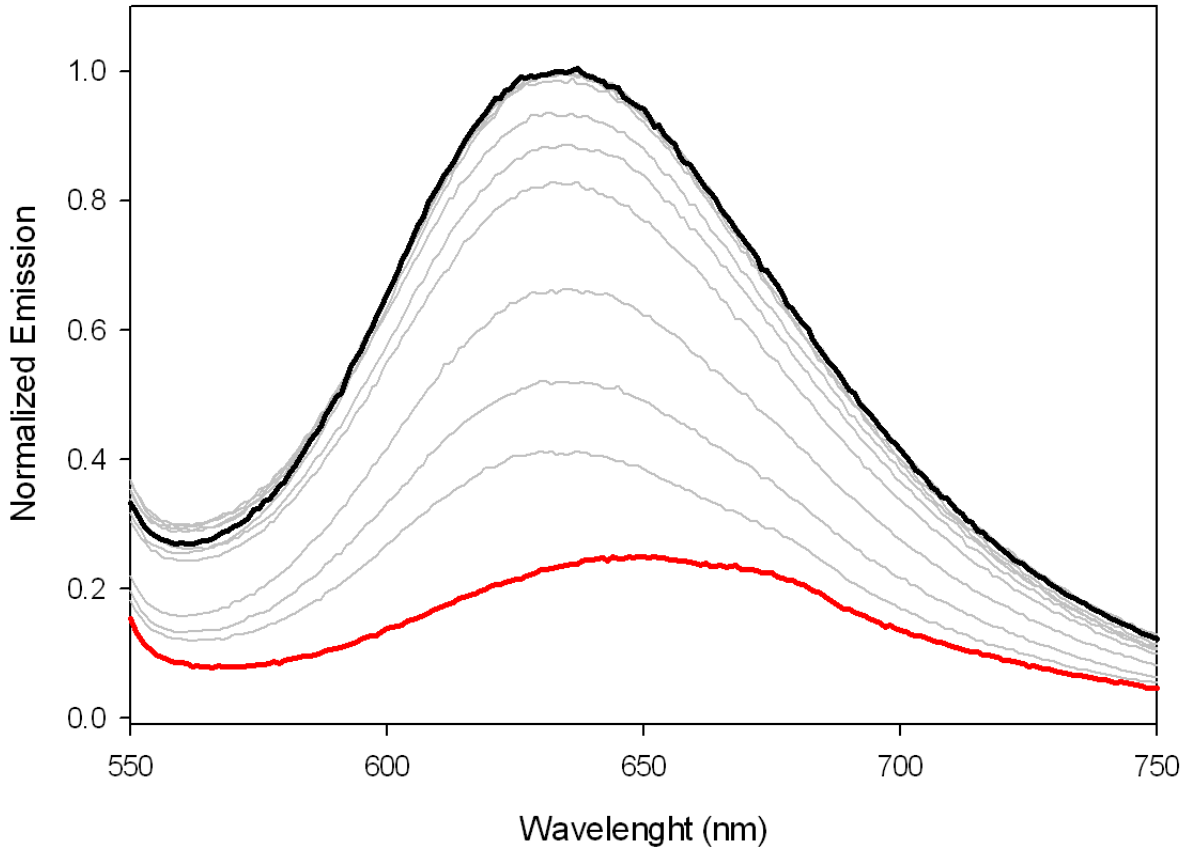


*Figure S22. Luminescence spectra of a 2.0 μM solution of **3** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of TT mismatched D:D1' oligonucleotide solution until saturation.*

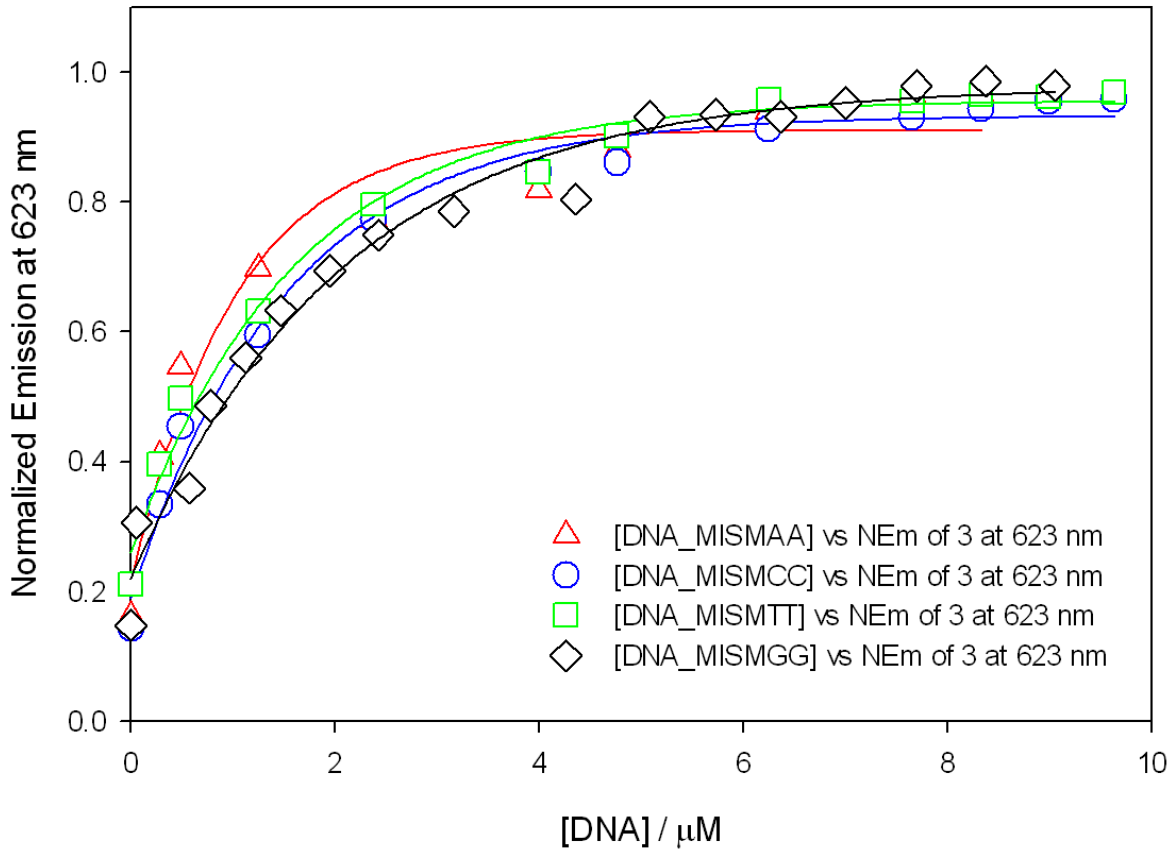


*Figure S23. Profile of the titration experiments of **3** with GG mismatched (A.A1'; black diamonds), AA mismatched (B:B1'; red triangles), CC mismatched (C:C1'; blue circles) and TT mismatched (D:D1'; green squares) oligonucleotides at λ_{em} = 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (black, red, blue and green lines, respectively).*

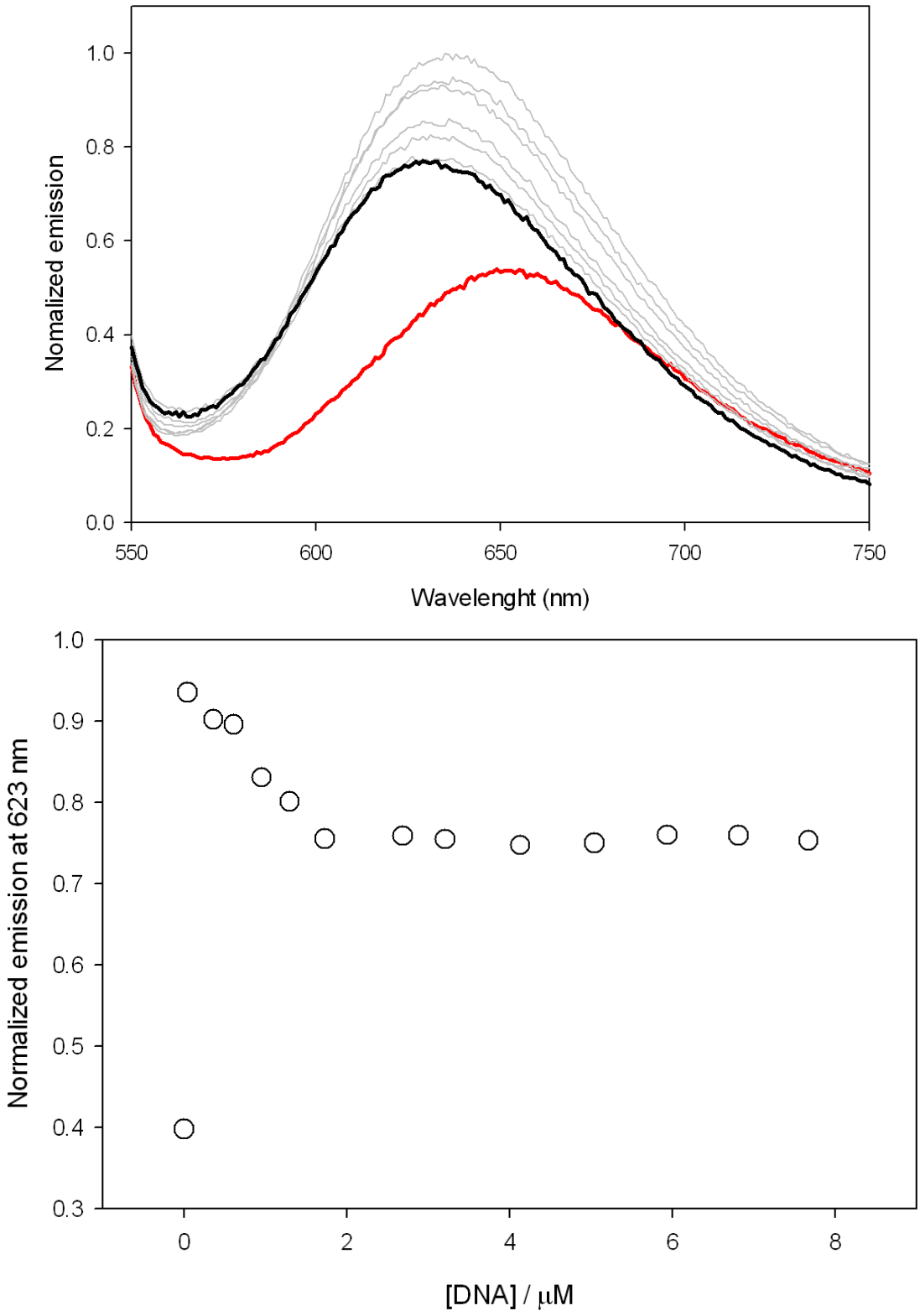


Figure S24. Top, luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of "well-matched" A:A' oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of* ***4*** *with "well-matched" A:A' oligonucleotide at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media).*

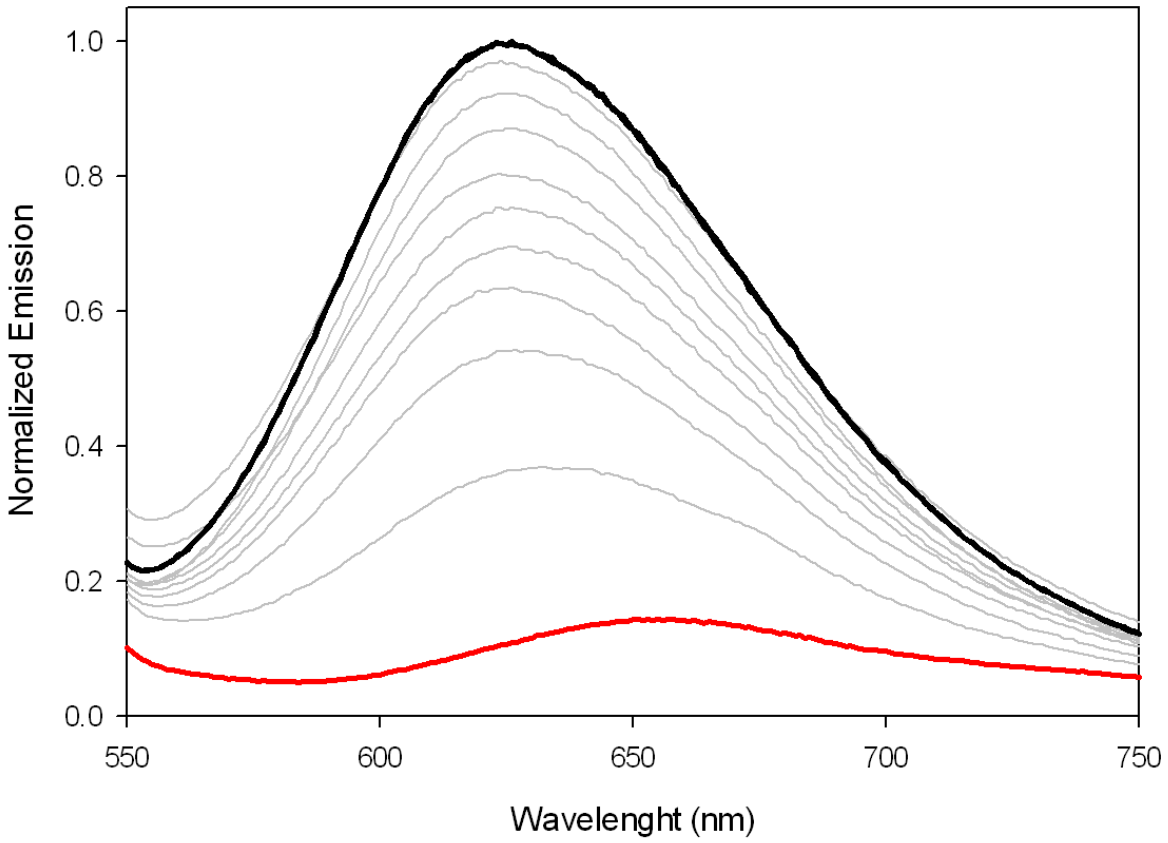


Figure S25. Luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG mismatched A:A1' oligonucleotide solution until saturation.*

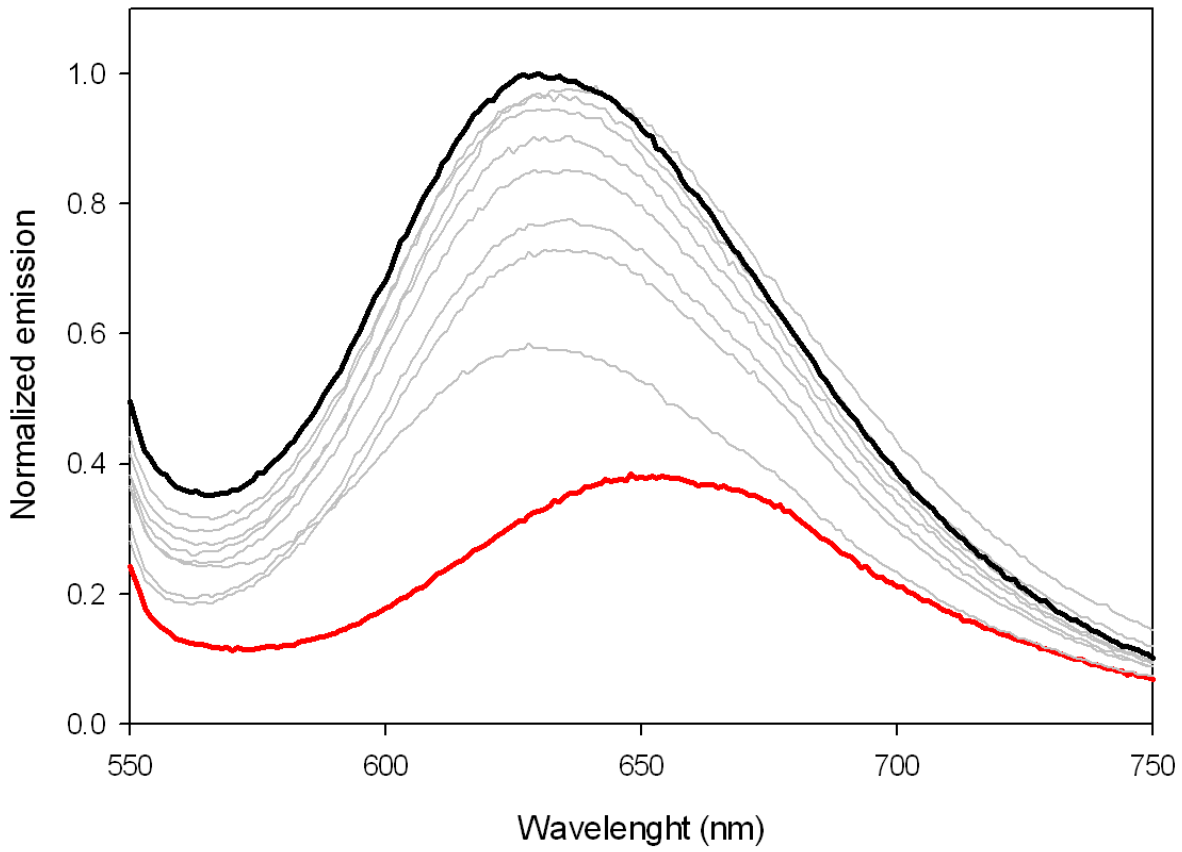


Figure S26. Luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/AA mismatched A:A2' oligonucleotide solution until saturation.*

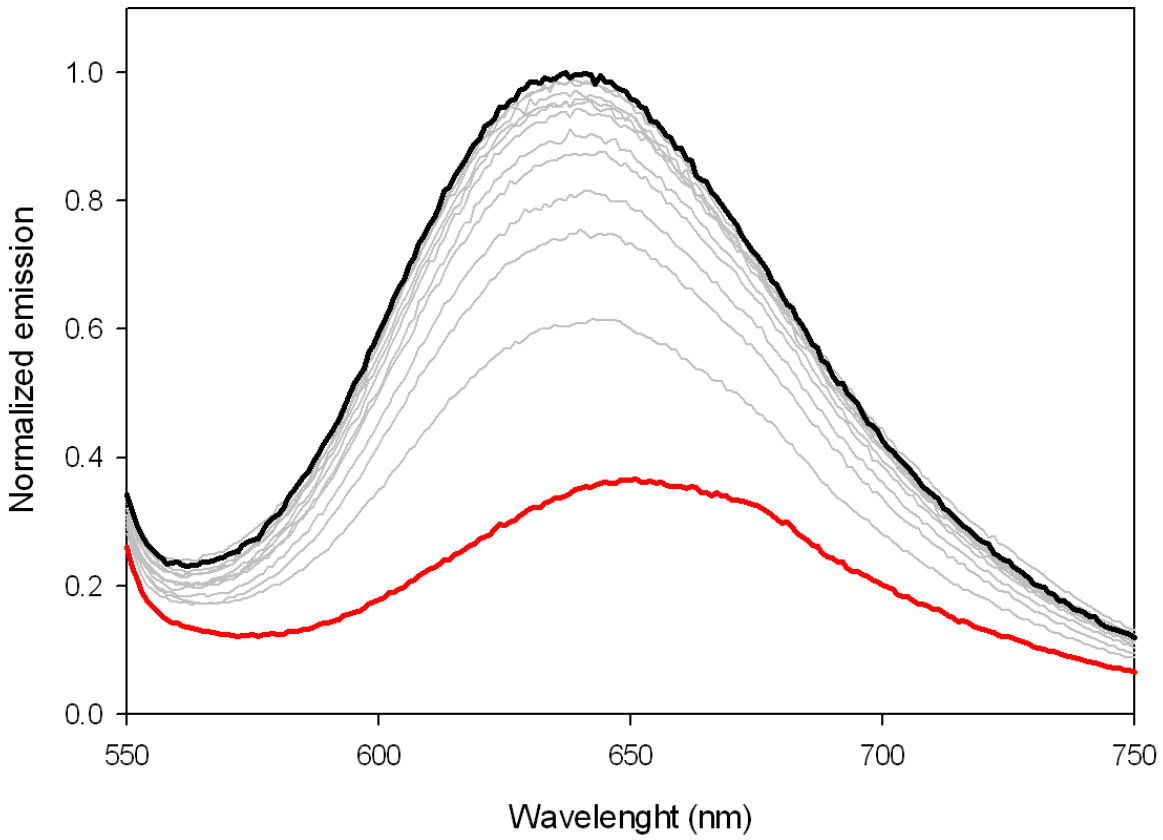


*Figure S27. Luminescence spectra of a 2.0 μM solution of **4** in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/GG/AA mismatched A:A3' oligonucleotide solution until saturation.*

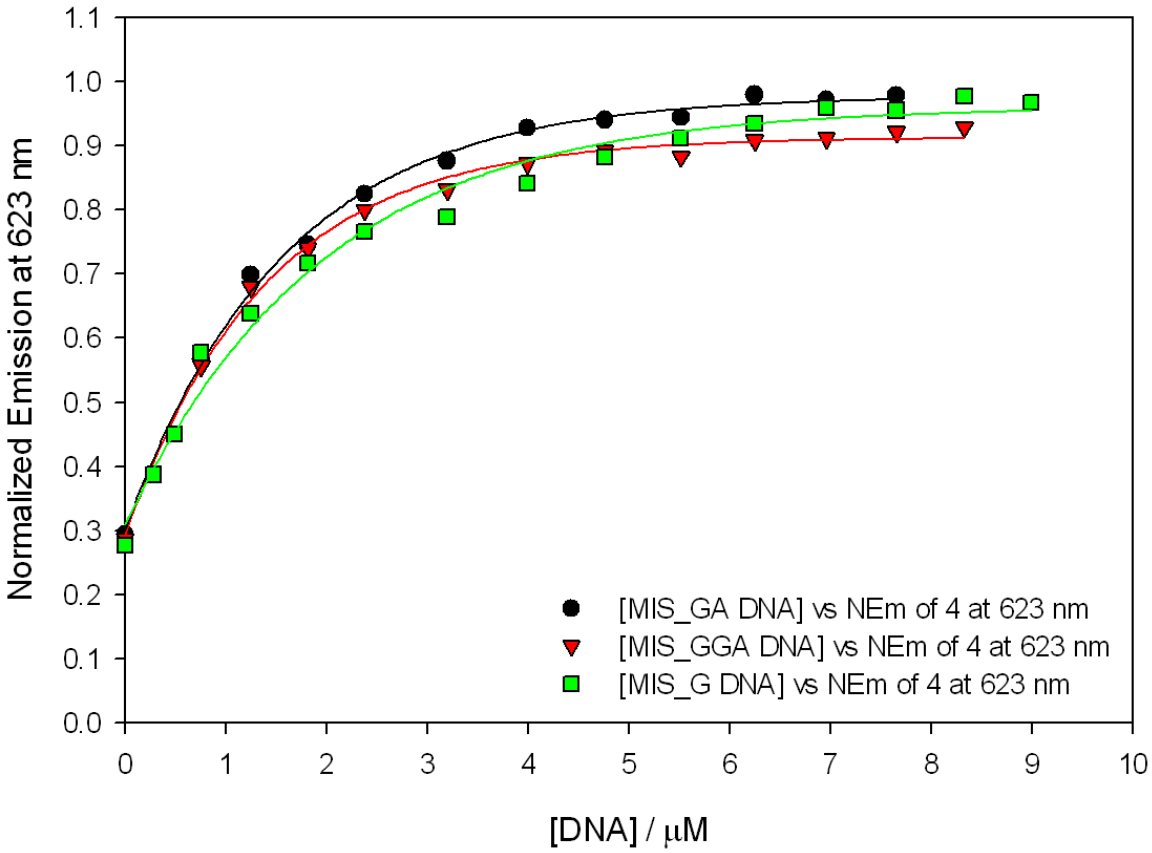


Figure S28. Profile of the titration experiments of ***4*** *with GG mismatched (A.A1'; green squares), GG/AA mismatched (A:A2'; black circles) and GG/GG/AA mismatched (A:A3'; red triangles) oligonucleotides at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (green, black and red lines, respectively).*

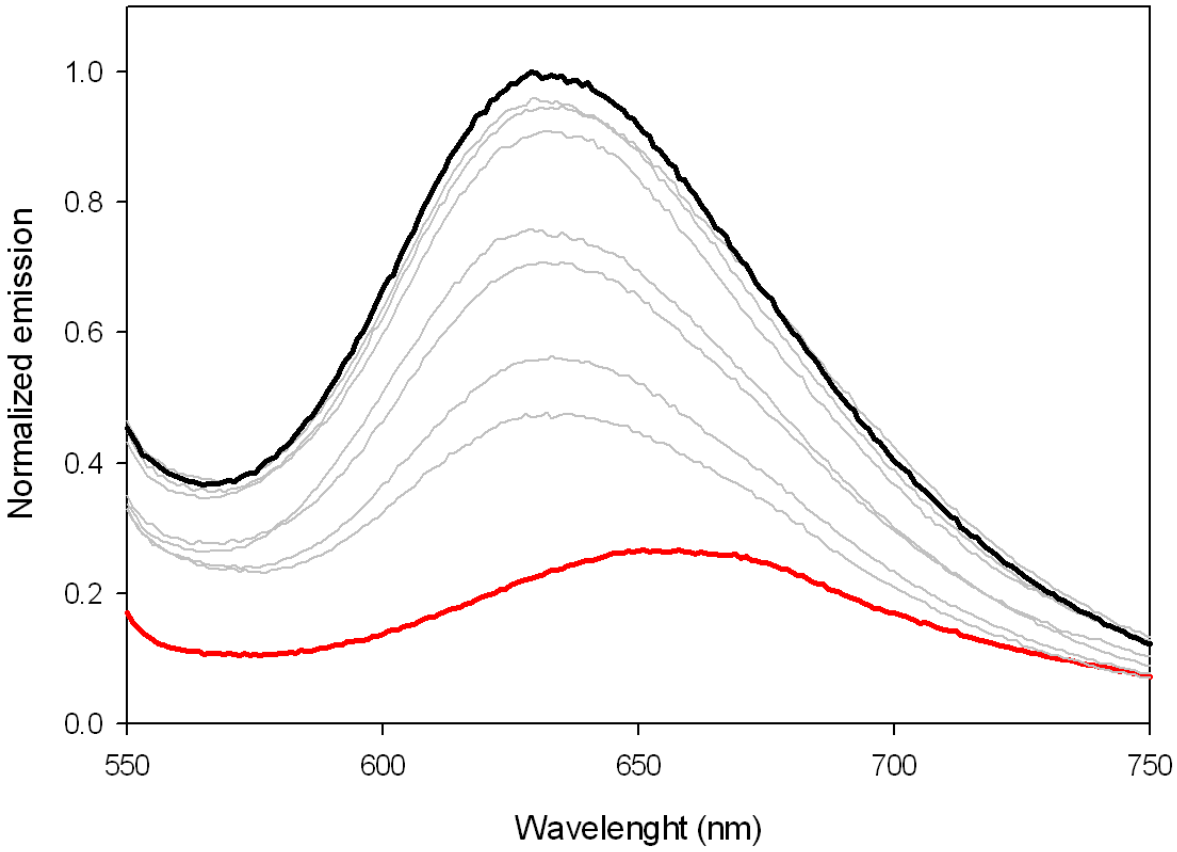


Figure S29. Luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AA mismatched B:B1' oligonucleotide solution until saturation.*

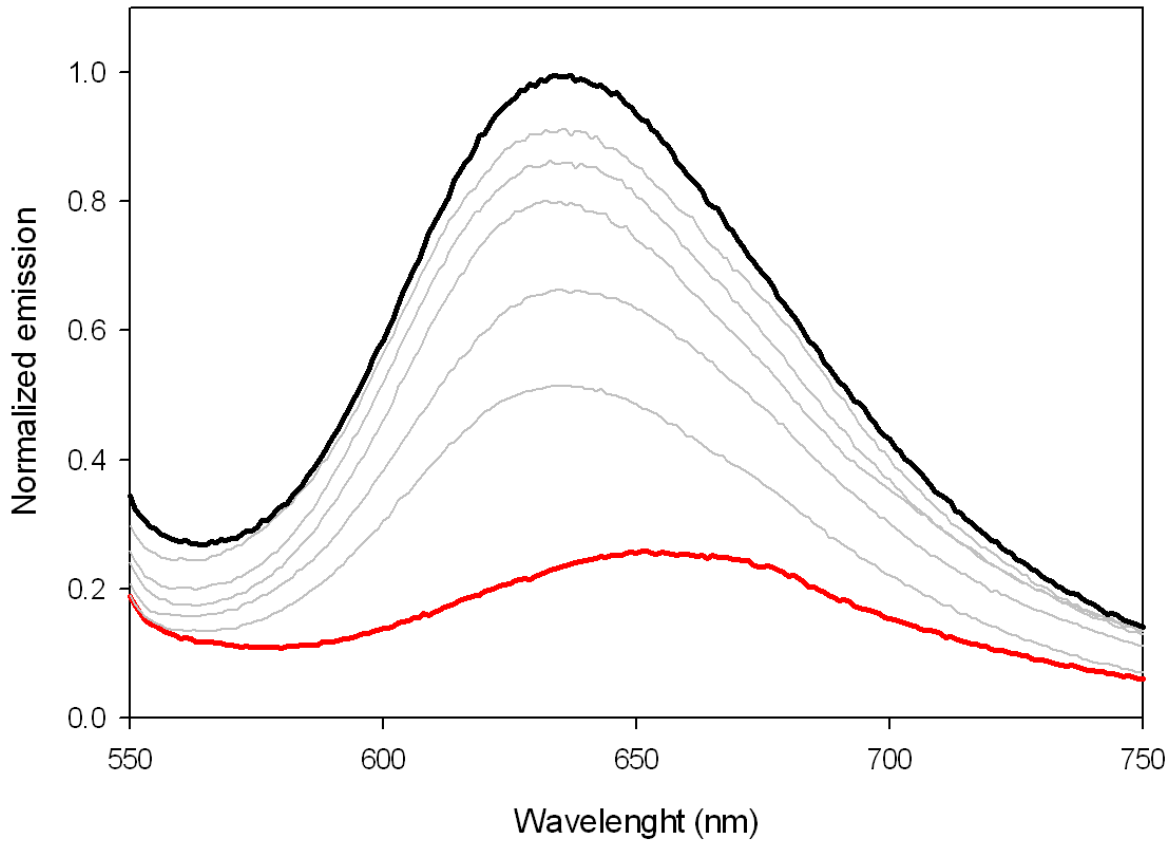


Figure S30. Luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of CC mismatched C:C1' oligonucleotide solution until saturation.*

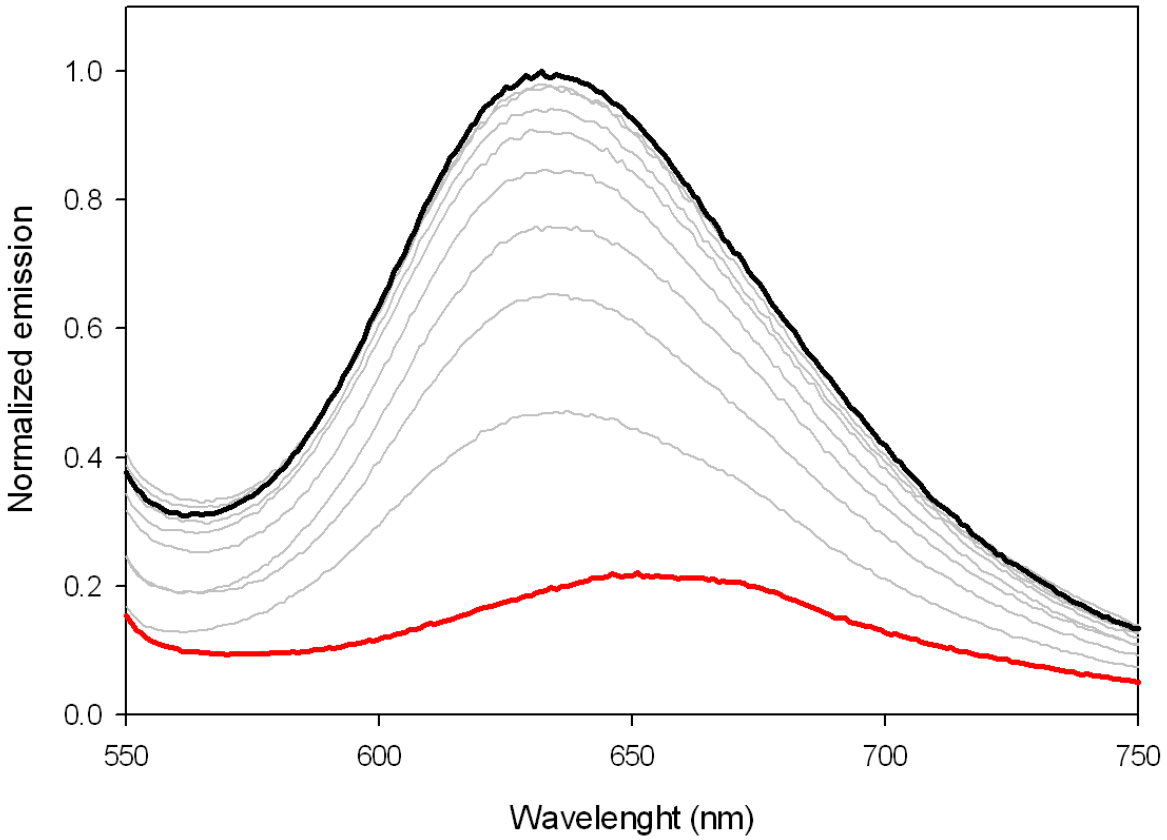


Figure S31. Luminescence spectra of a 2.0 μM solution of ***4*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of TT mismatched D:D1' oligonucleotide solution until saturation.*

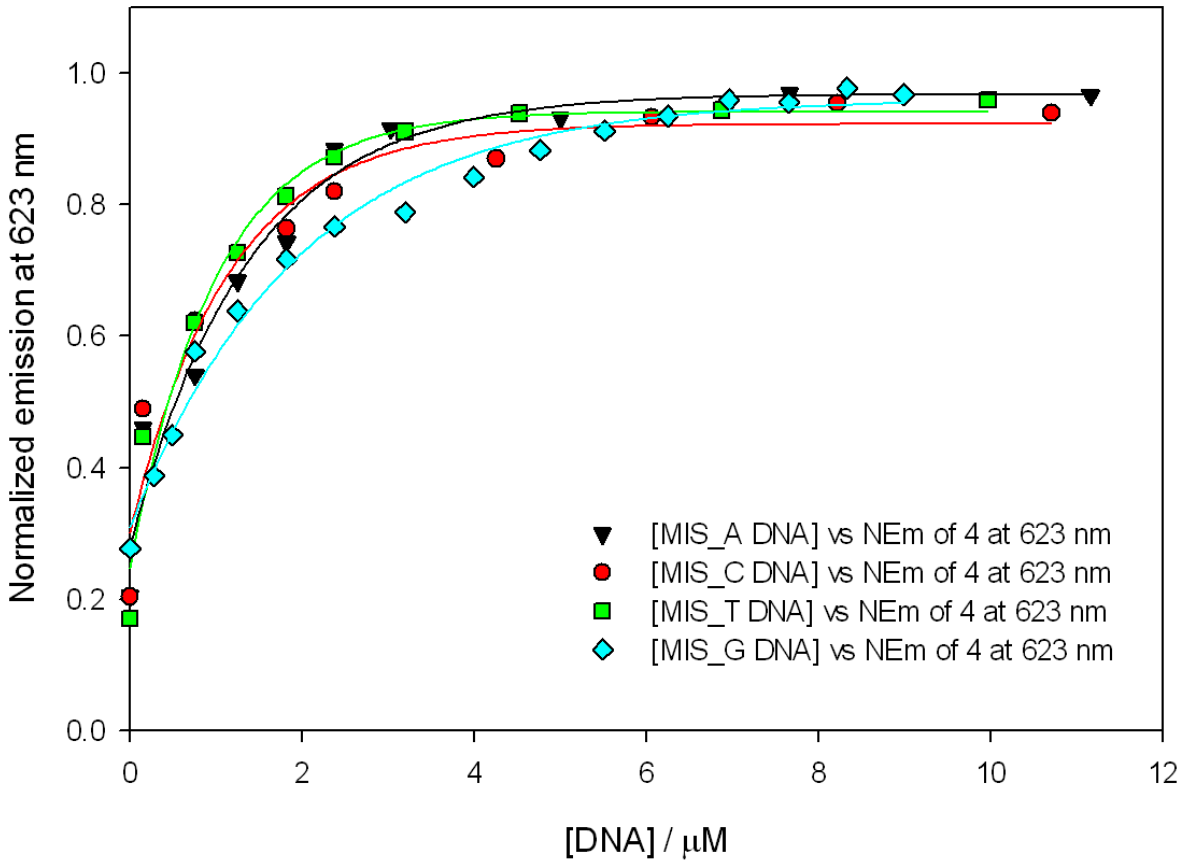


Figure S32. Profile of the titration experiments of ***4*** *with GG mismatched (A.A1'; cyan diamonds), AA mismatched (B:B1'; black triangles), CC mismatched (C:C1'; red circles) and TT mismatched (D:D1'; green squares) oligonucleotides at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (cyan, black, red and green lines, respectively).*

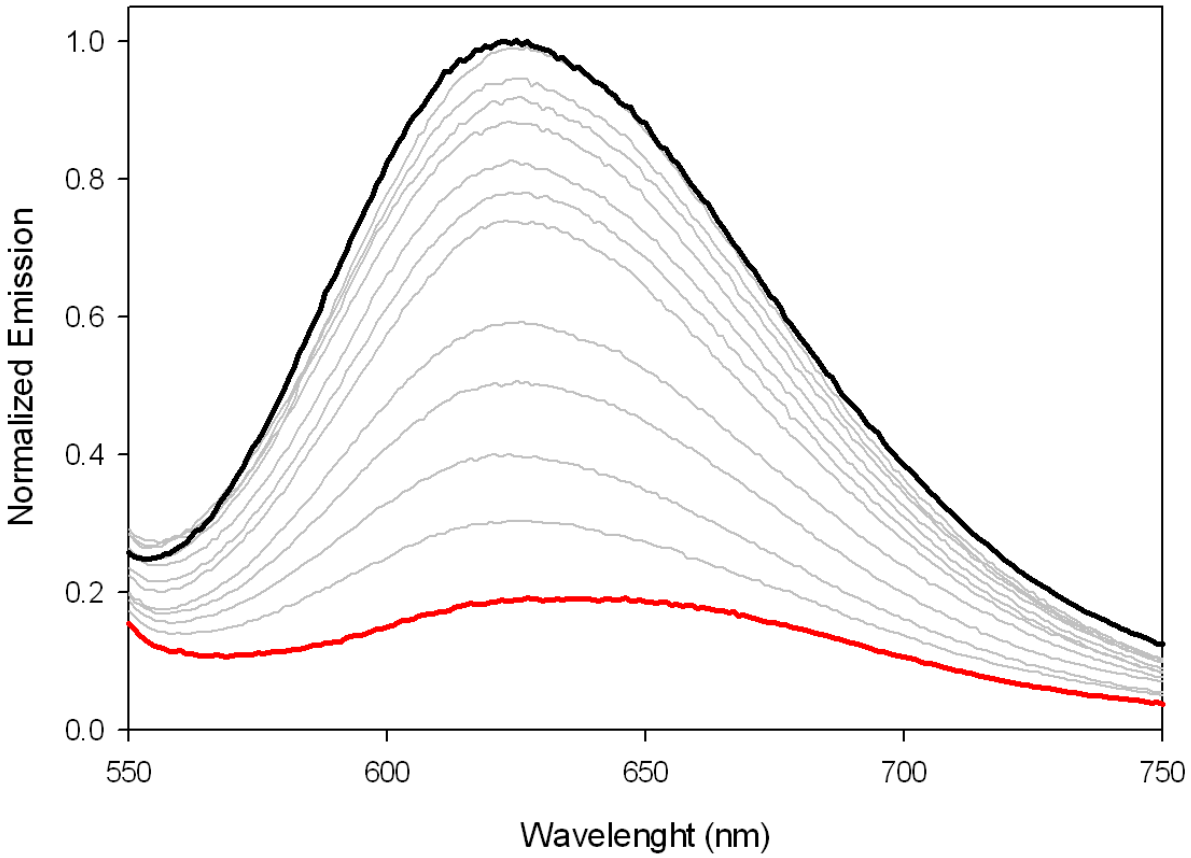


Figure S33. Luminescence spectra of a 2.0 μM solution of **5** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of "well-matched" A:A' oligonucleotide solution until saturation.*

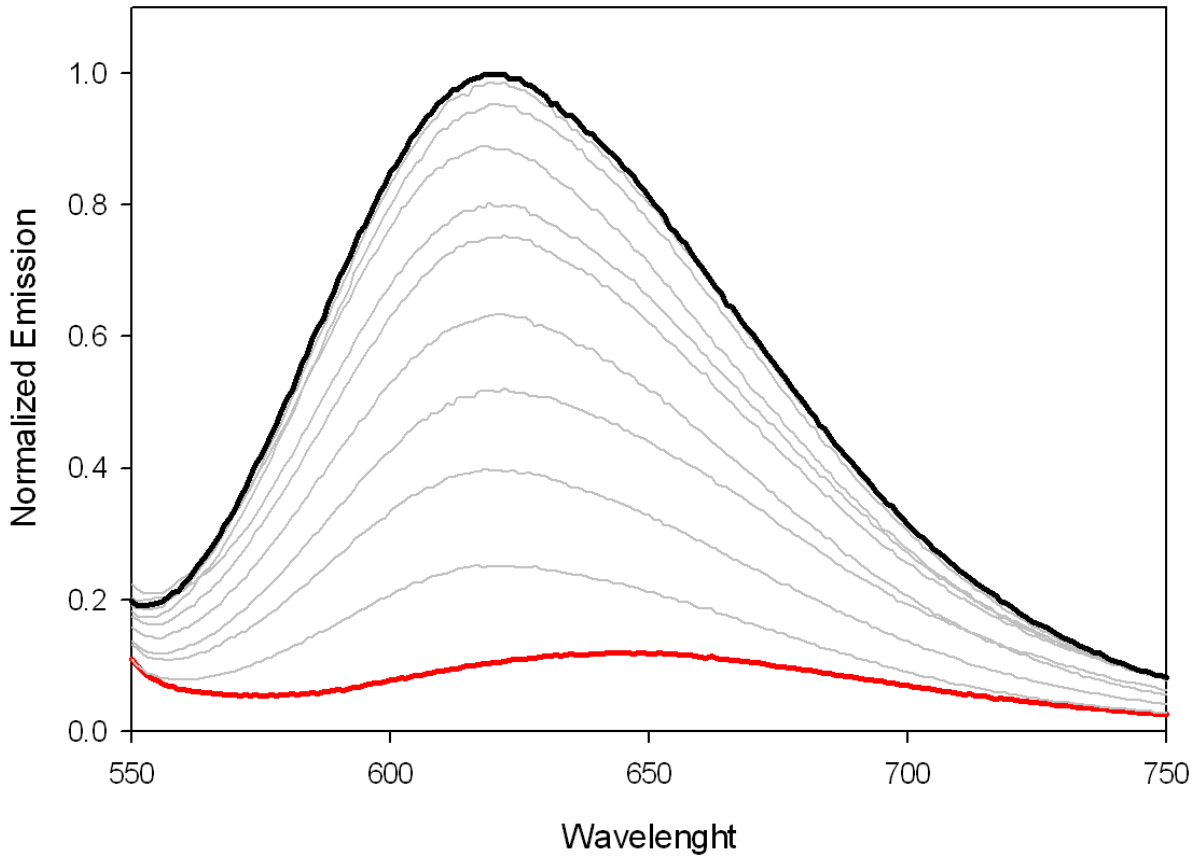


Figure S34. Luminescence spectra of a 2.0 μM solution of **5** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG mismatched A:A1' oligonucleotide solution until saturation.*

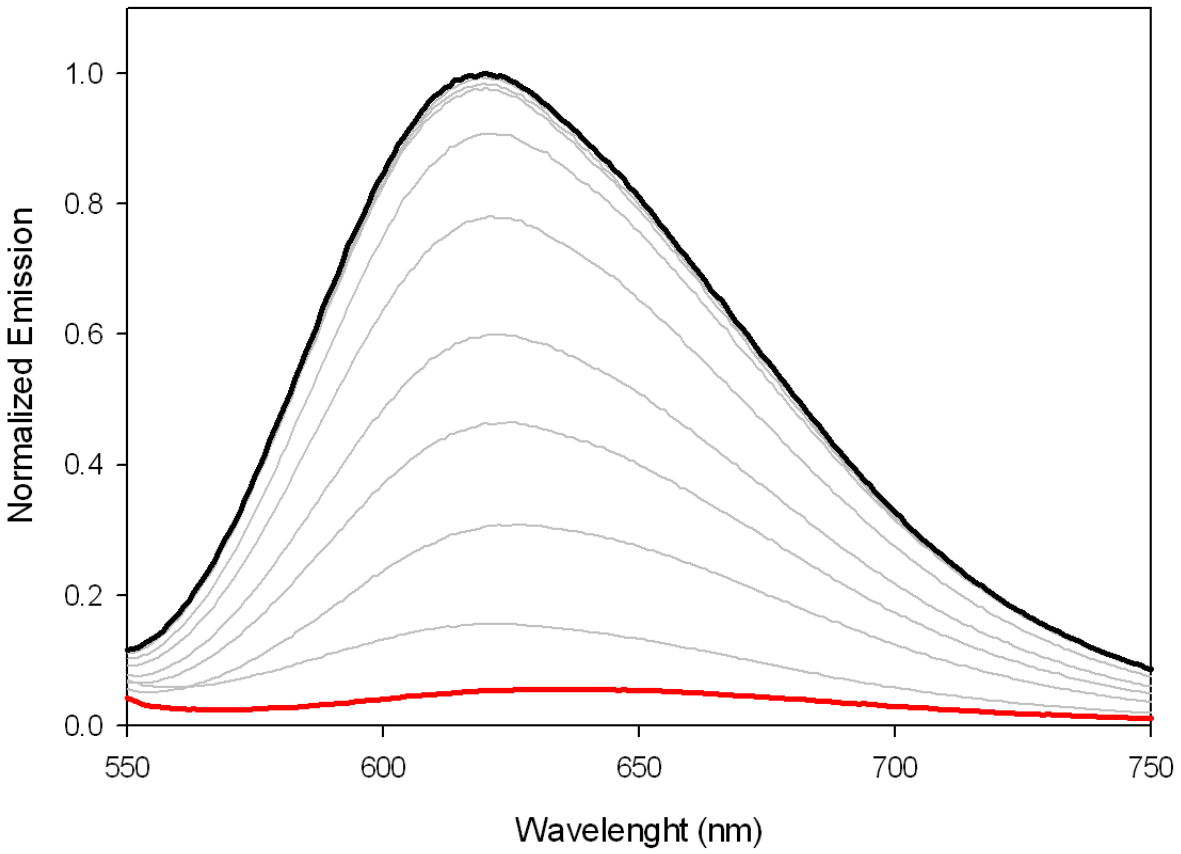


Figure S35. Luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/AA mismatched A:A2' oligonucleotide solution until saturation.*

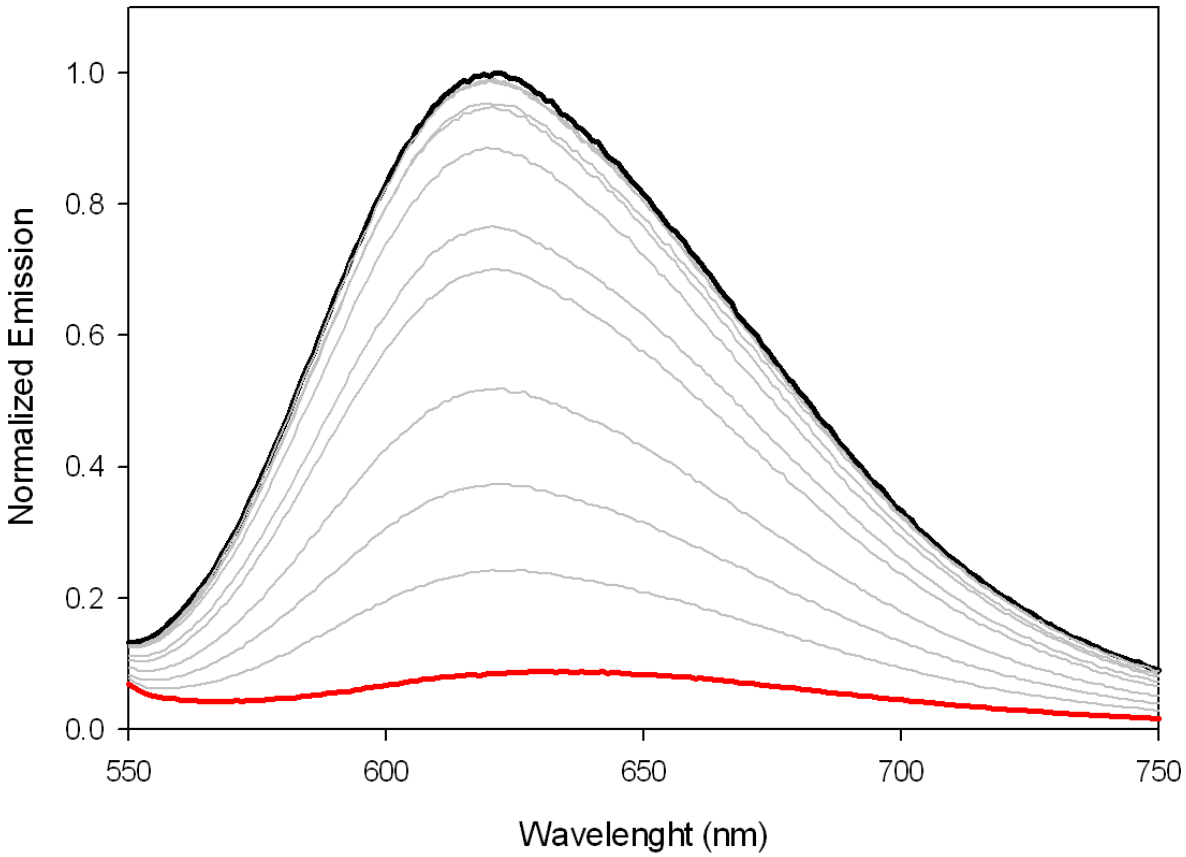


Figure S36. Luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/GG/AA mismatched A:A3' oligonucleotide solution until saturation.*

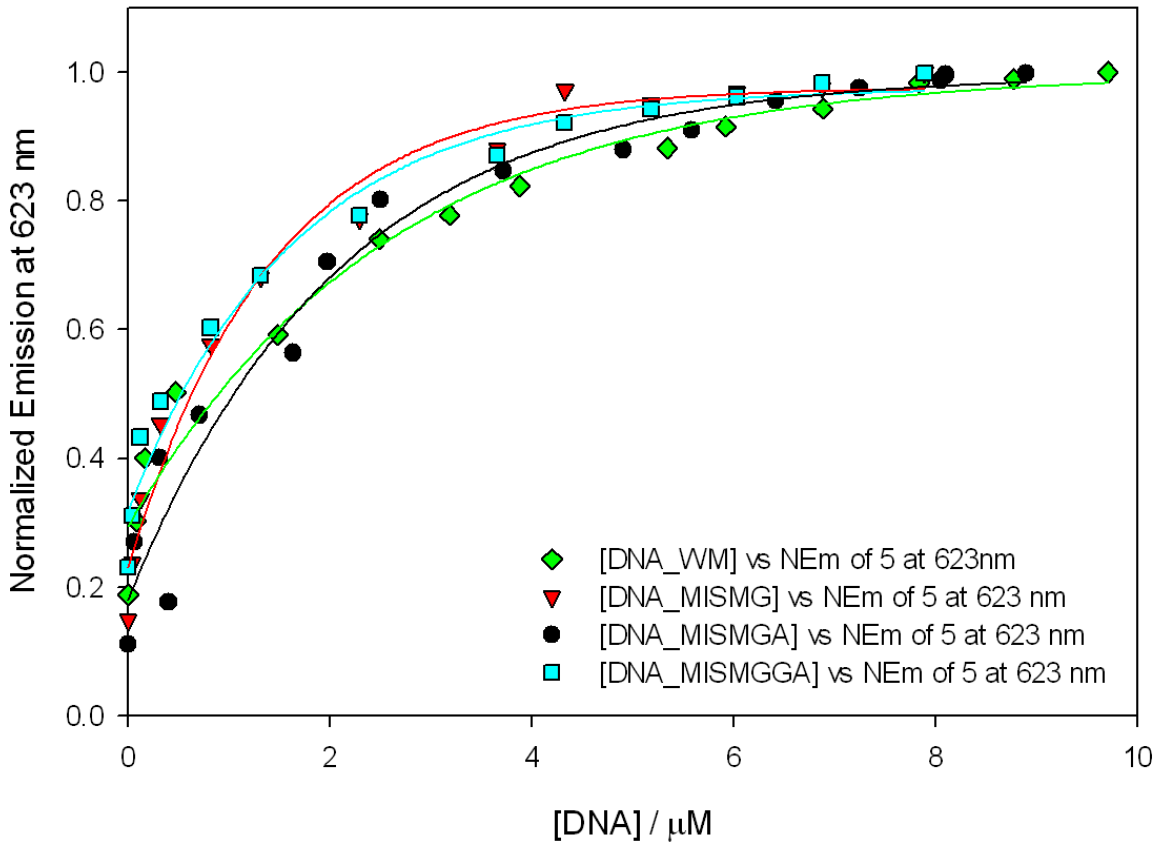


Figure S37. Profile of the titration experiments of ***5*** *with "well matched" (A:A'; green diamonds), GG mismatched (A.A1'; red triangles), GG/AA mismatched (A:A2'; black circles) and GG/GG/AA mismatched (A:A3'; cyan squares) oligonucleotides at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (green, red, black and cyan lines, respectively).*

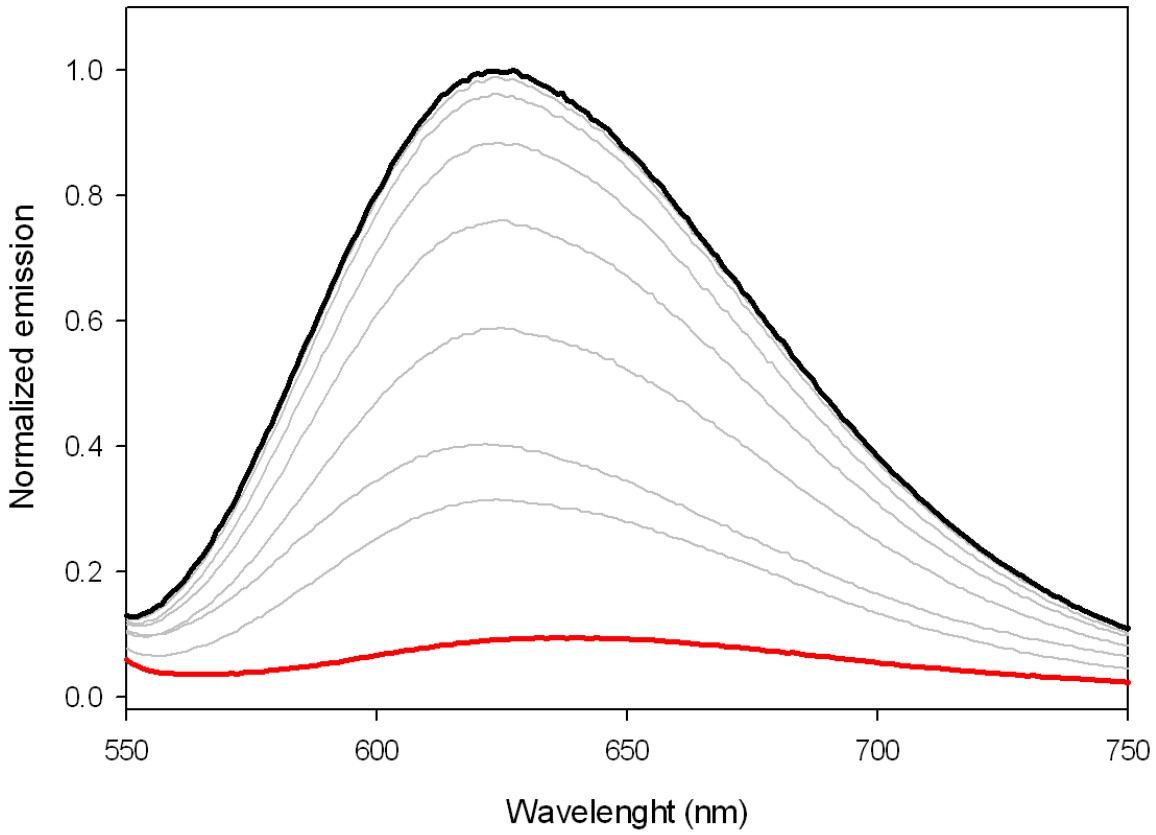


Figure S38. Luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AA mismatched B:B1' oligonucleotide solution until saturation.*

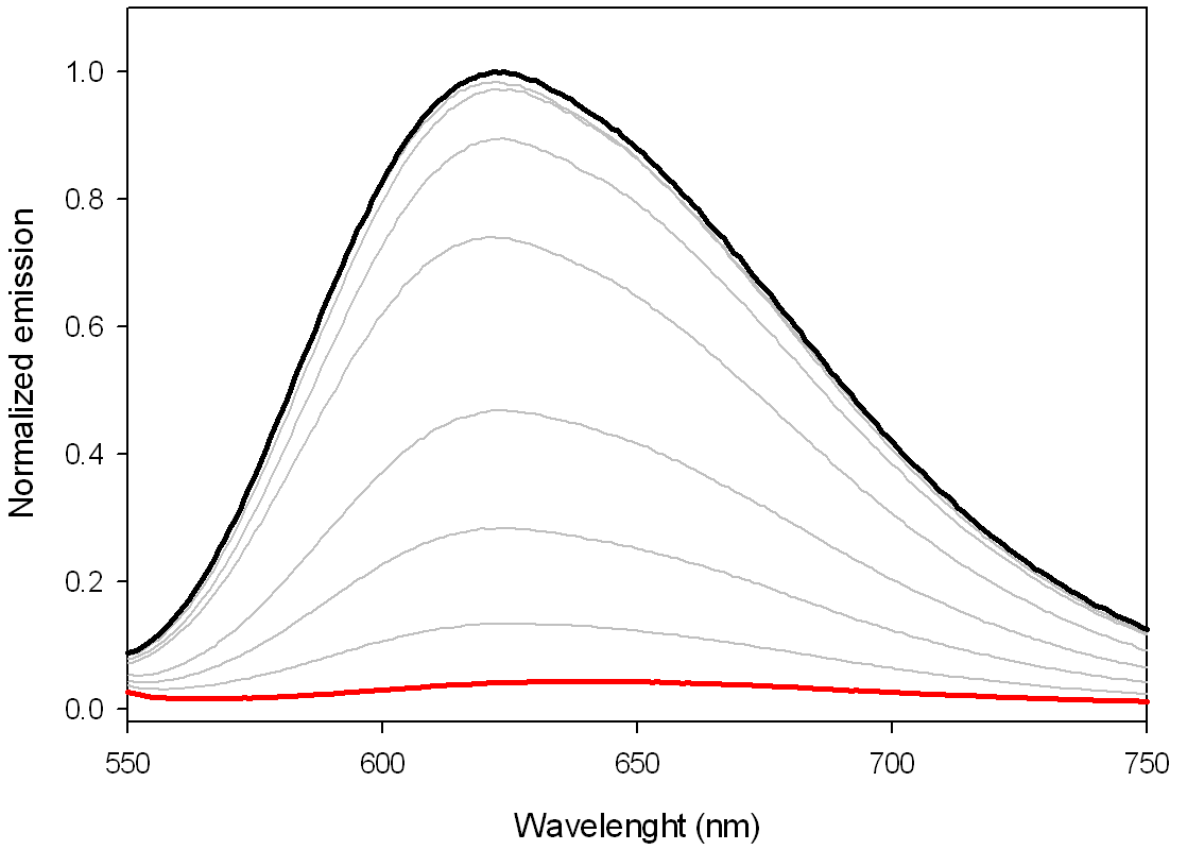


Figure S39. Luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of CC mismatched C:C1' oligonucleotide solution until saturation.*

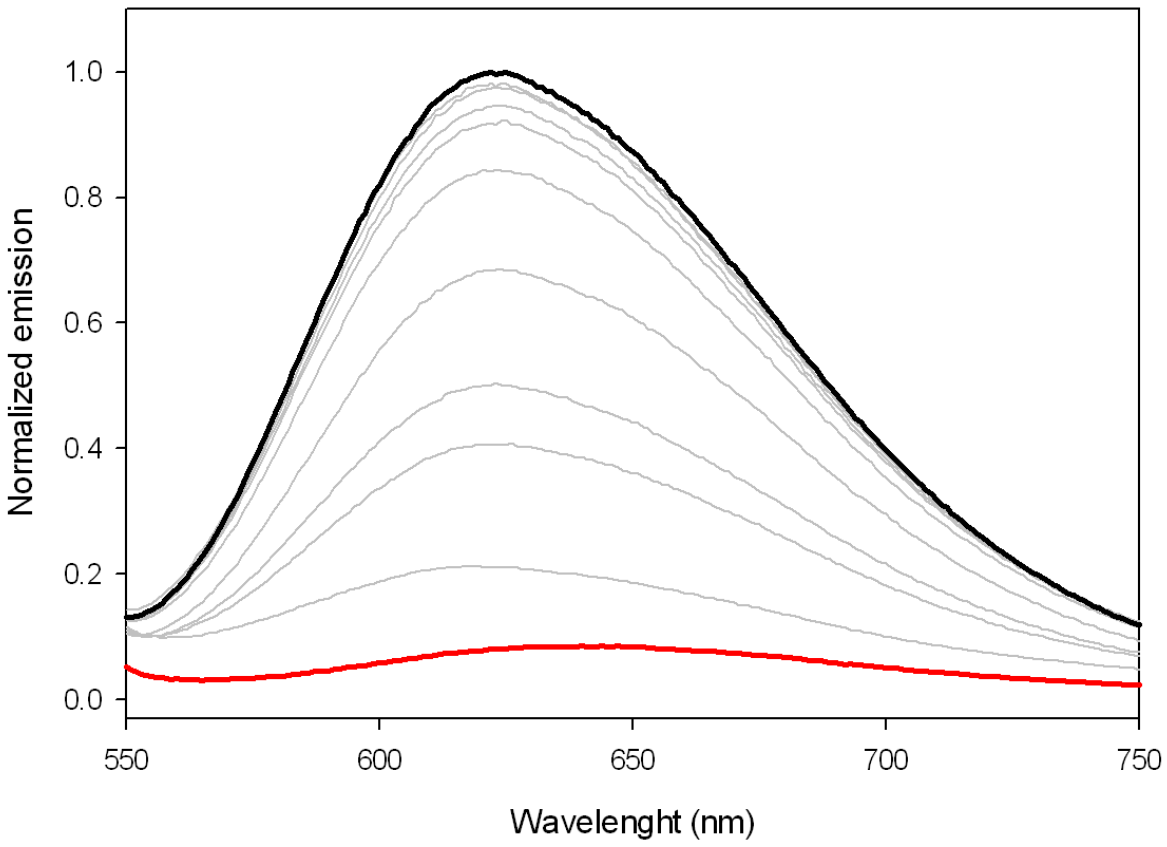


Figure S40. Luminescence spectra of a 2.0 μM solution of ***5*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of TT mismatched D:D1' oligonucleotide solution until saturation.*

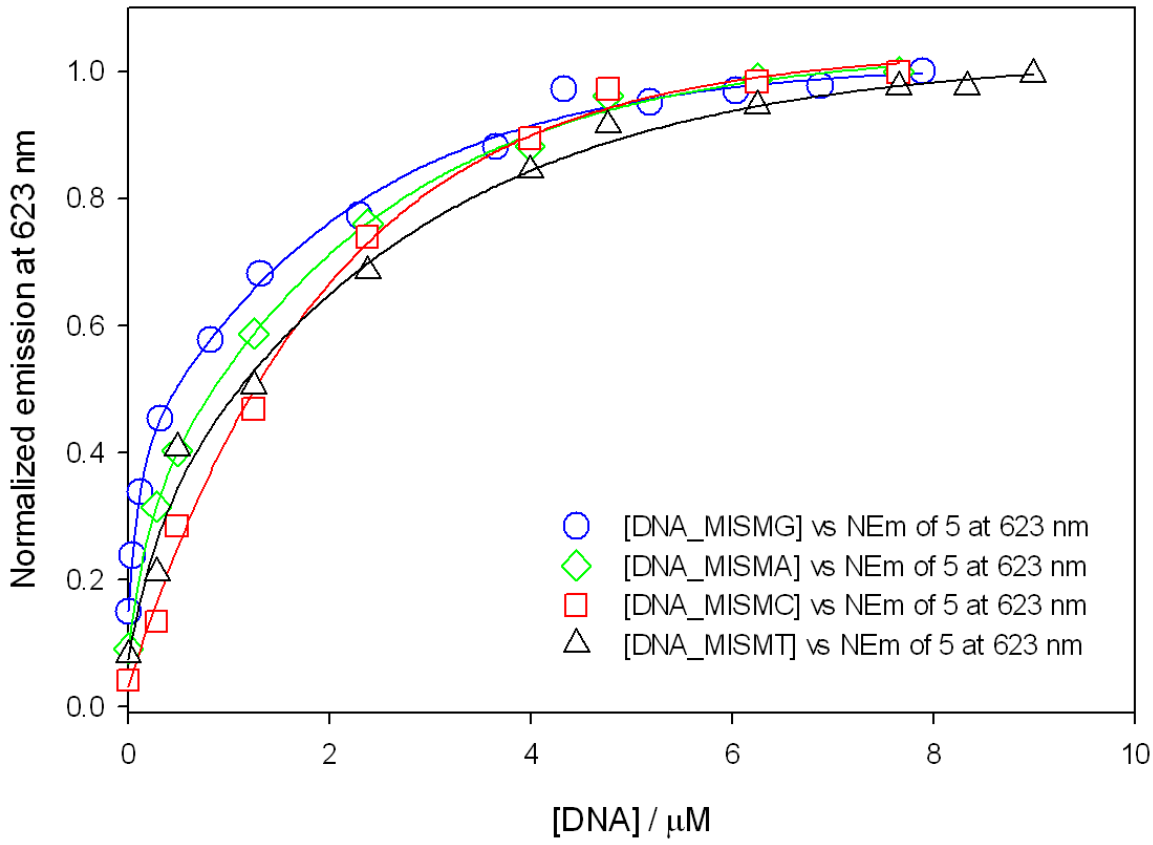


Figure S41. Profile of the titration experiments of ***5*** *with GG mismatched (A.A1'; blue circles), AA mismatched (B:B1'; green diamonds), CC mismatched (C:C1'; red squares) and TT mismatched (D:D1'; black triangles) oligonucleotides at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (blue, green, red and black lines, respectively).*

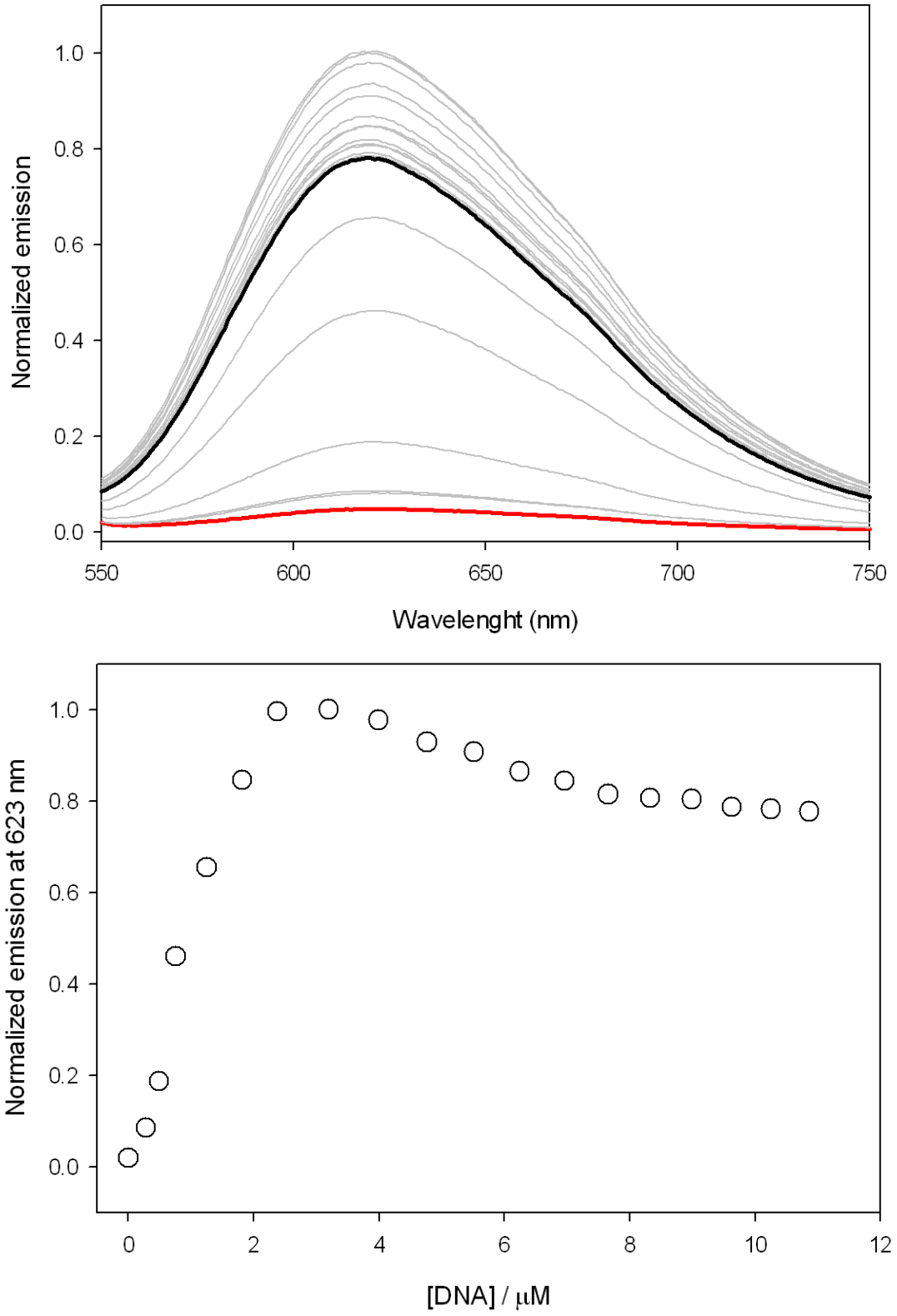


Figure S42. Top, luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of "well-matched" A:A' oligonucleotide solution until saturation; bottom, profile (black triangles) of the fluorescence titration experiment of* ***6*** *with "well-matched" A:A' oligonucleotide at* λ_{em} *= 630 nm (emission intensity vs. concentration of DNA in the media).*

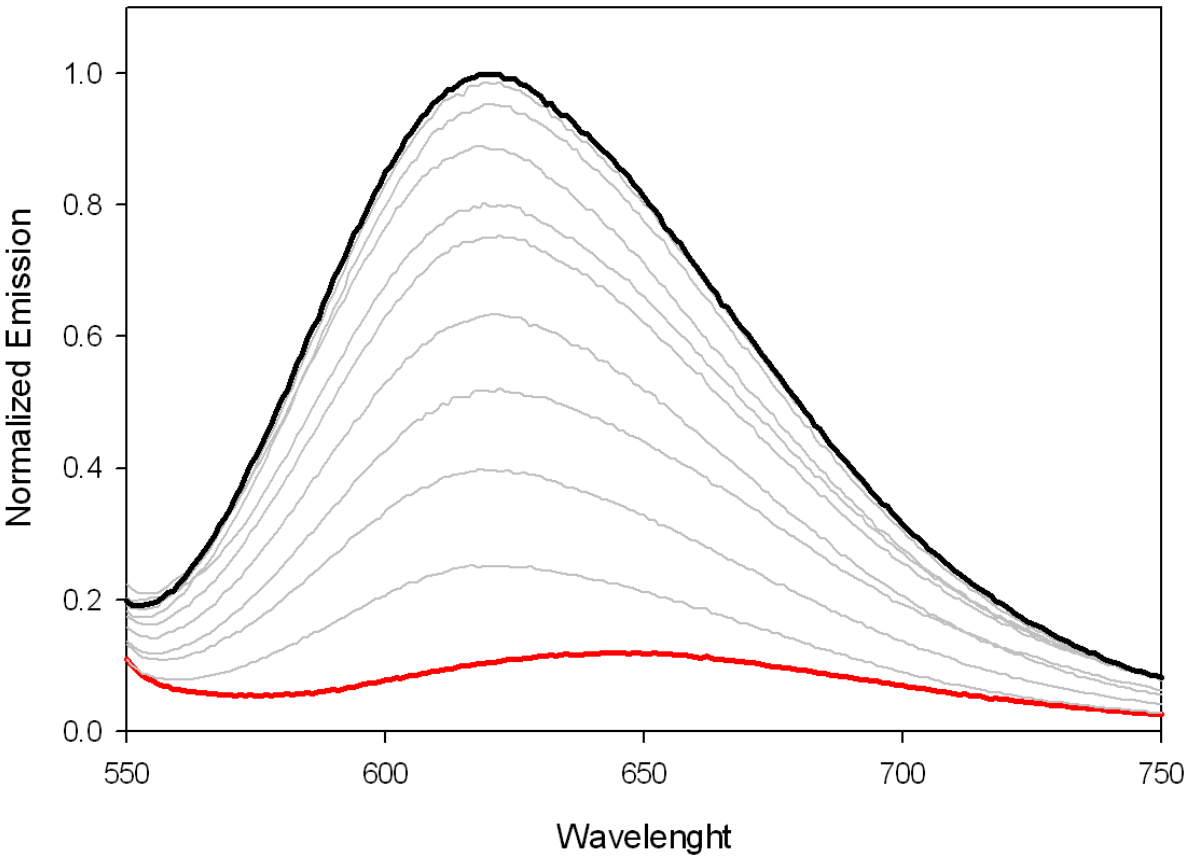


Figure S43. Luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG mismatched A:A1' oligonucleotide solution until saturation.*

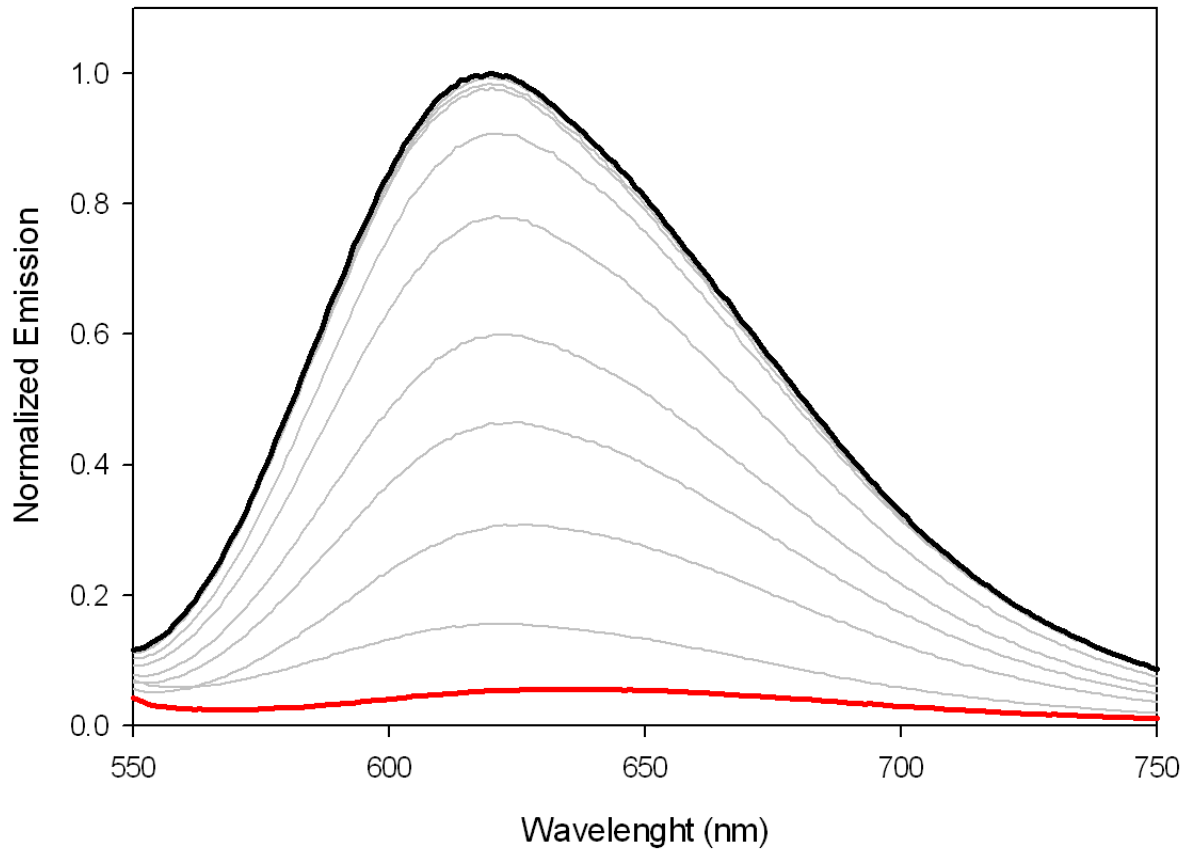


Figure S44. Luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/AA mismatched A:A2' oligonucleotide solution until saturation.*

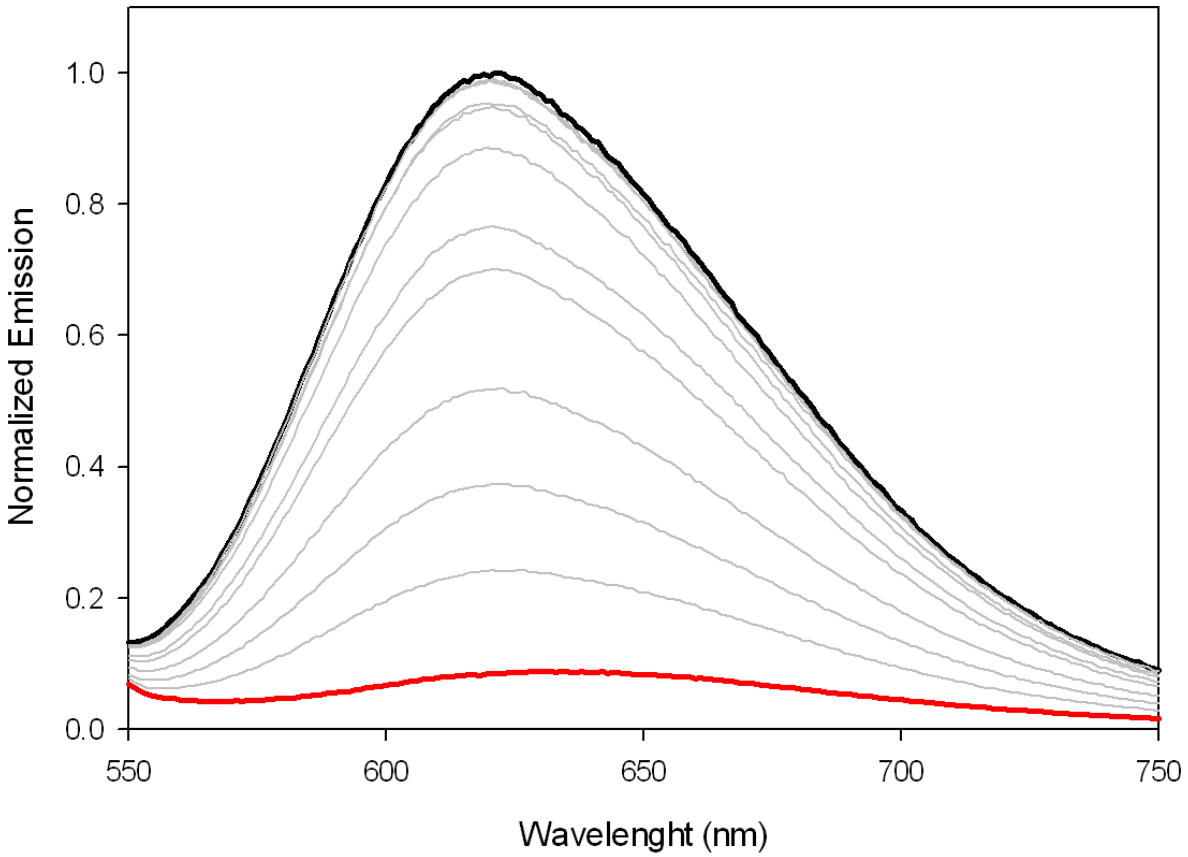


Figure S45. Luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of GG/GG/AA mismatched A:A3' oligonucleotide solution until saturation.*

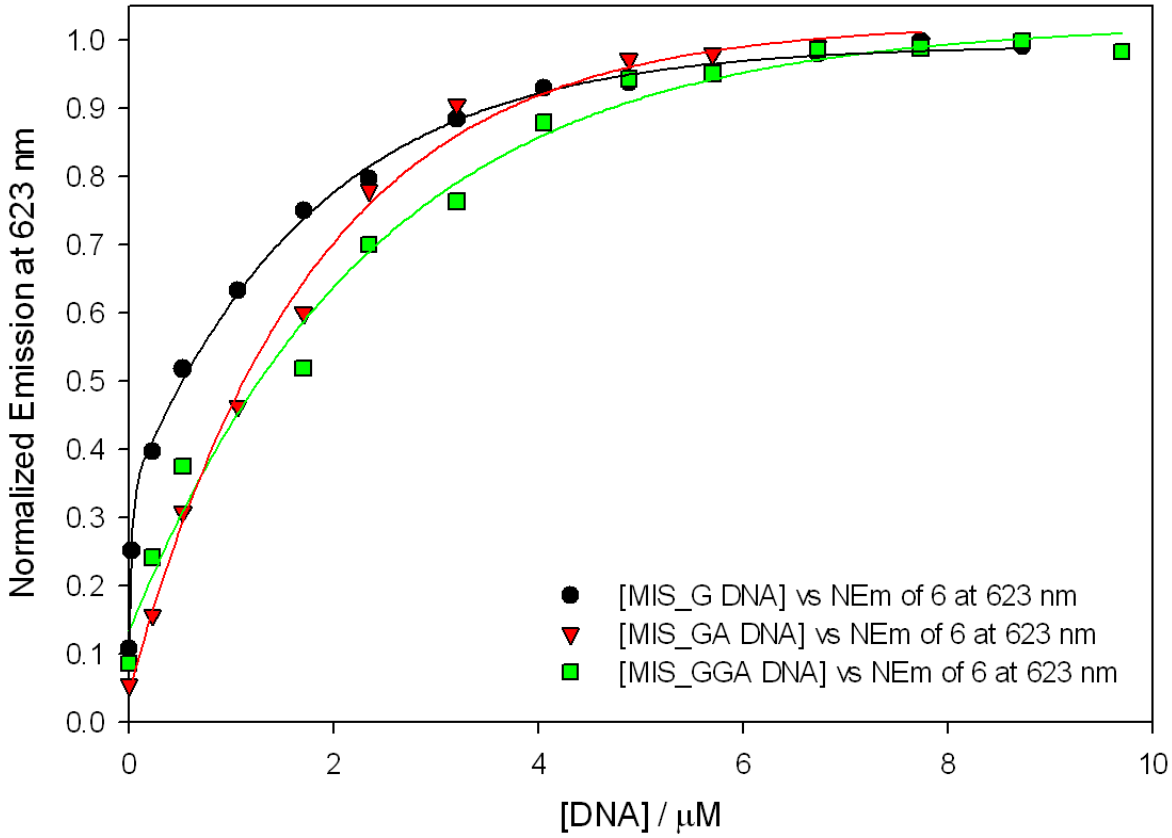


*Figure S46. Profile of the titration experiments of **6** with GG mismatched (A.A1'; black circles), GG/AA mismatched (A:A2'; red triangles) and GG/GG/AA mismatched (A:A3'; green squares) oligonucleotides at λ_{em} = 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (black, red and green lines, respectively).*

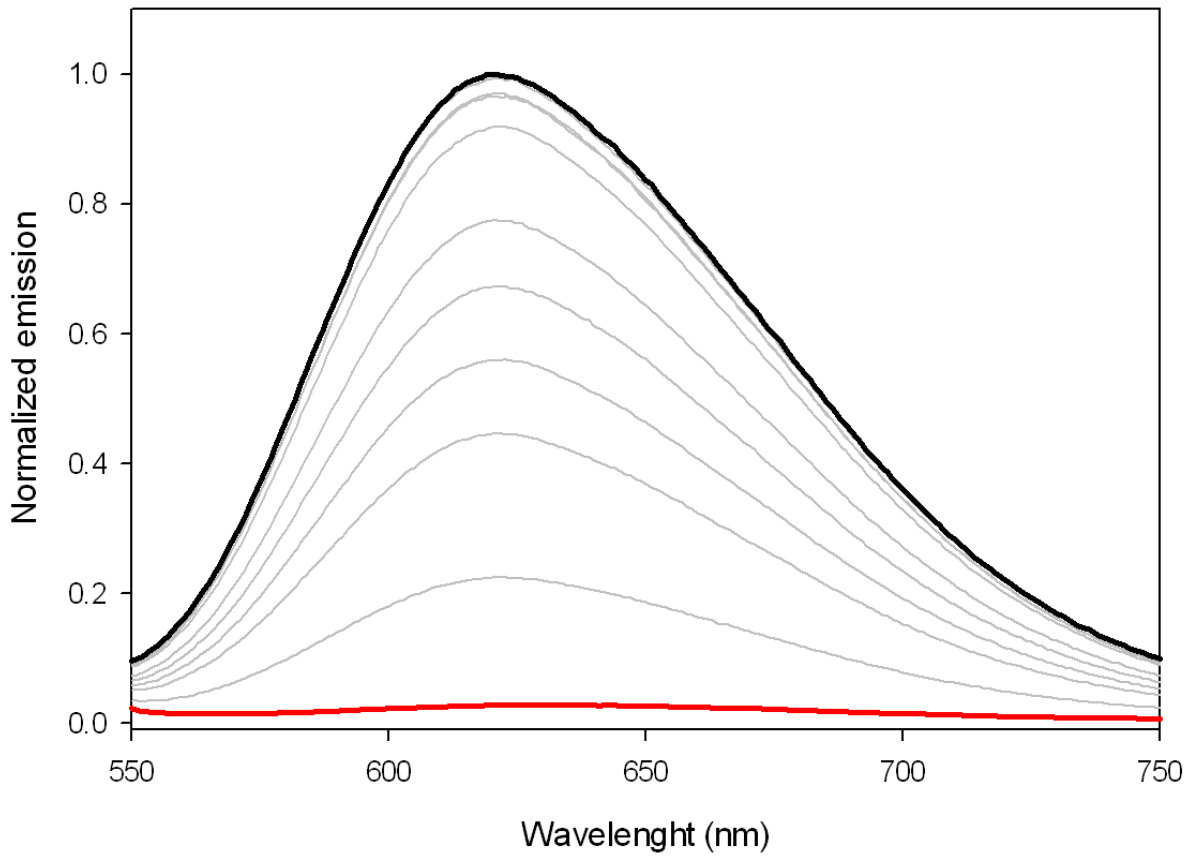


Figure S47. Luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of AA mismatched B:B1' oligonucleotide solution until saturation.*

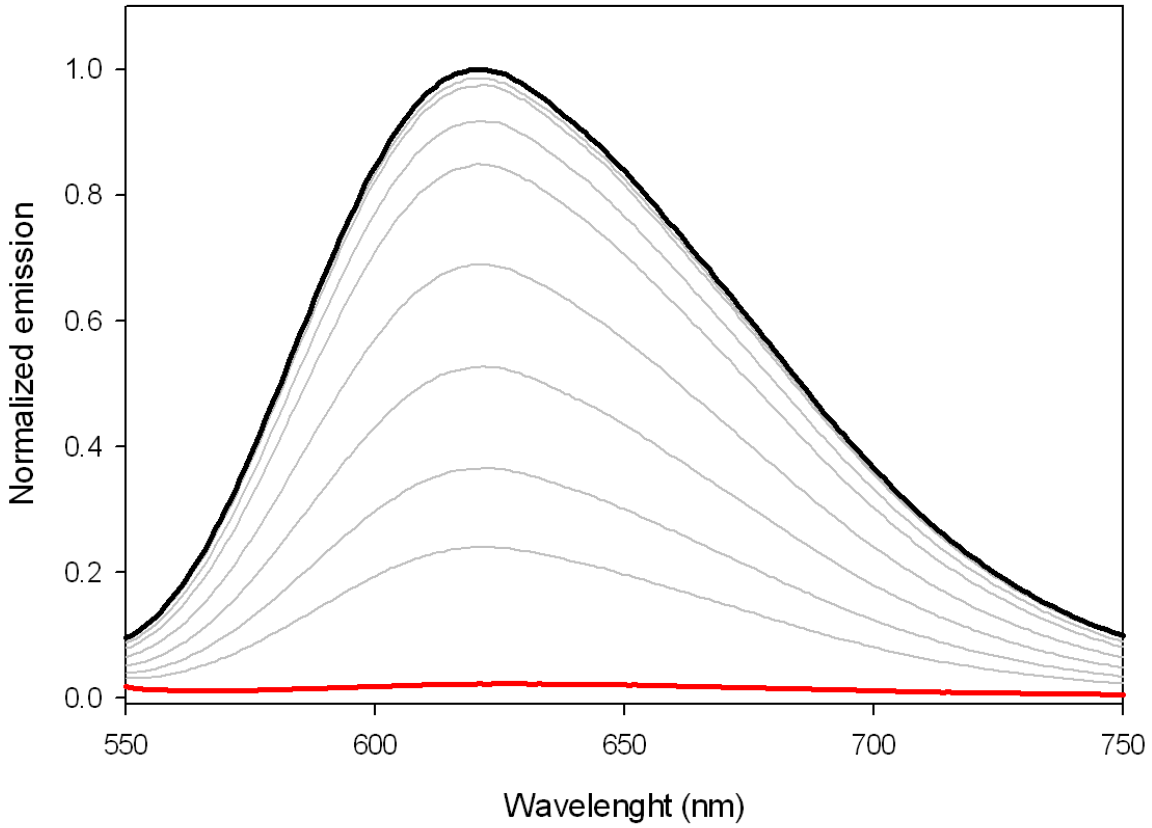


Figure S48. Luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of CC mismatched C:C1' oligonucleotide solution until saturation.*

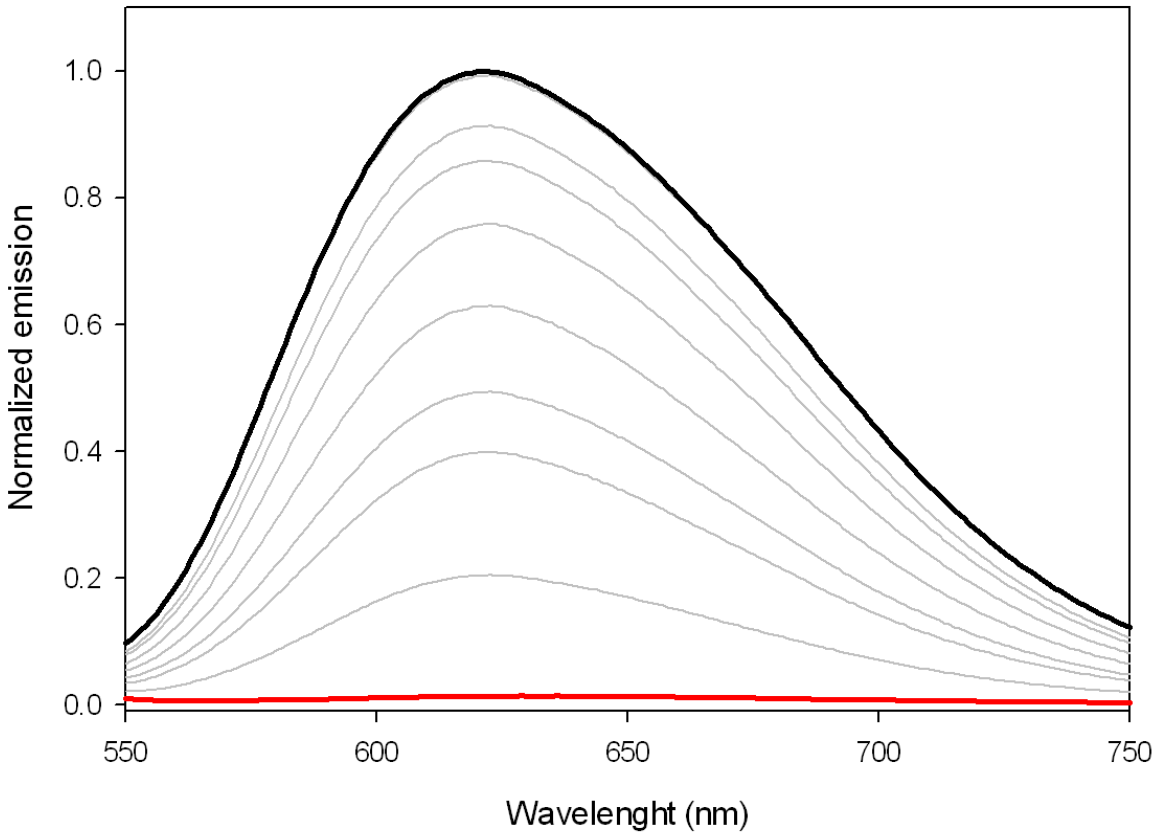


Figure S49. Luminescence spectra of a 2.0 μM solution of ***6*** *in Tris-HCl buffer (20 mM), NaCl (100 mM), pH 7.5 and evolution upon addition of aliquots of TT mismatched D:D1' oligonucleotide solution until saturation.*

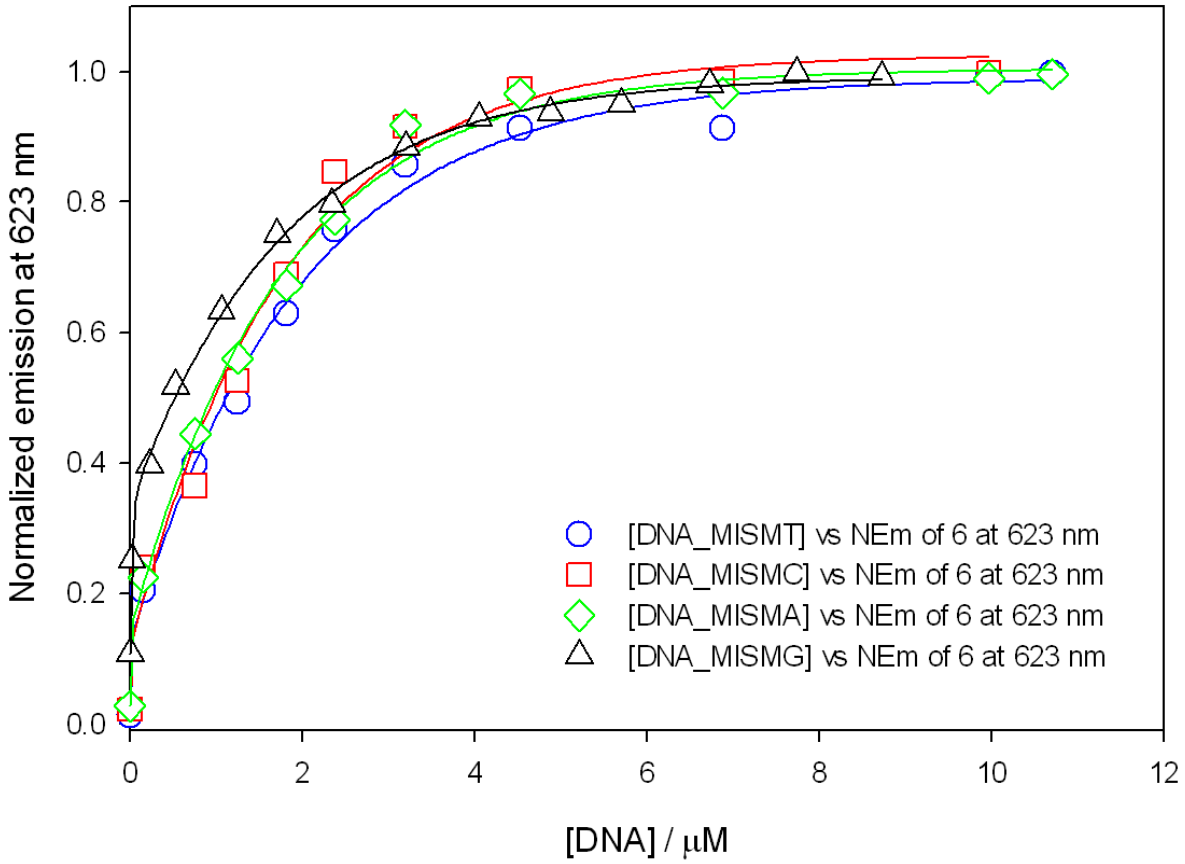


*Figure S50. Profile of the titration experiments of **6** with GG mismatched (A.A1'; black triangles), AA mismatched (B:B1'; green diamonds), CC mismatched (C:C1'; red squares) and TT mismatched (D:D1'; blue circles) oligonucleotides at λ_{em} = 630 nm (emission intensity vs. concentration of DNA in the media) with the corresponding best fits (black, green, red and blue lines, respectively).*

Table S1. Hairpin and double-stranded B-DNA sequences studied

	Code	***Complete Sequence (5'-3')***
Hairpin B-DNA oligonucleotides	***AAAATTT***	GGC**AAAATTT**CGTTTTTCGAAATTTTGCC
	GAATTC	GGC**GAATTC**AGCTTTTTGCTGAATTCGCC
	GGCCC	GGCA**GGCCC**AGCTTTTTGCTGGGCCT GCC
Double-Stranded B-DNA oligonucleotides	***A:A'*** *well matched*	**A:** AACACATGC***AGG***ACGGCGCTT **A':** AAGCGCCGT***CCT***GCATGTGTT
	A:A1' *GG mismatched*	**A1':** AAGCGCCGT***CGT***GCATGTGTT
	A:A2' *GG/AA mismatched*	**A2':** AAGCGCCGT***CGA***GCATGTGTT
	A:A3' *GG/GG/AA mismatched*	**A3':** AAGCGCCGT***GGA***GCATGTGTT
	B:B1' *AA mismatched*	**B:** AACACATGC***AAG***ACGCGCCTT **B1':** AAGCGCCGT***CAT***GCATGTGTT
	C:C1' *CC mismatched*	**C:** AACACATGC***ACG***ACGCGCCTT **C1':** AAGCGCCGT***CCT***GCATGTGTT
	D:D1' *TT mismatched*	**D:** AACACATGC***ATG***ACGCGCCTT **D1':** AAGCGCCGT***CTT***GCATGTGTT

*Table S2. Dissociation constants of the Ru(II) metallopeptides (**3-6**) with selected oligonucleotides. NC = Not calculated.*

Complex	***B-DNA oligonucleotides***		
	AAAATTT	**GAATTC**	**GGCCC**
3	0.32 (0.01)	0.19 (0.01)	0.18 (0.02)
4	NC	NC	NC
5	0.07 (0.006)	0.17 (0.006)	0.11 (0.03)
6	NC	NC	NC

*Table S3. Dissociation constants of the Ru(II) metallopeptides (**3**-**6**) with selected oligonucleotides. NC = Not calculated.*

Complex	*B-DNA oligonucleotides*						
	A:A'	**A:A1'**	**A:A2'**	**A:A3'**	**B:B1'**	**C:C1'**	**D:D1'**
	Well matched	*GG mismatch*	*GG/AA mismatches*	*GG/GG/AA mismatches*	*AA mismatch*	*CC mismatch*	*TT mismatch*
3	0.14 (0.2)	0.66 (0.21)	1.17 (0.25)	0.37 (0.28)	0.30 (0.36)	0.33 (0.16)	0.37 (0.18)
4	NC	0.36 (0.16)	0.31 (0.09)	0.20 (0.07)	0.16 (0.13)	0.20 (0.25)	0.07 (0.11)
5	1.22 (0.50)	0.31 (0.22)	0.65 (0.32)	0.47 (0.31)	0.48 (0.25)	0.57 (0.14)	0.76 (0.30)
6	NC	0.24 (0.16)	0.39(0.01)	0.95(0.28)	0.28 (0.10)	0.25 (0.11)	0.47 (0.15)

UV-vis studies

To a solution of the selected Ru(II) metallopeptide (**3**-**6**) in Tris-HCl buffer (20 mM), pH 7.5 and NaCl (100 mM), an aliquot of the selected B-DNA stock solution (in water) was added in order to reach a [Ru]/[DNA] ratio of 0.5. The absorption spectra of the metallopeptides were recorded before and after the addition of the corresponding oligonucleotide. The sequences of the oligonucleotides studied are listed in table S1.

a) Hairpin "well-matched" B-DNA binding studies

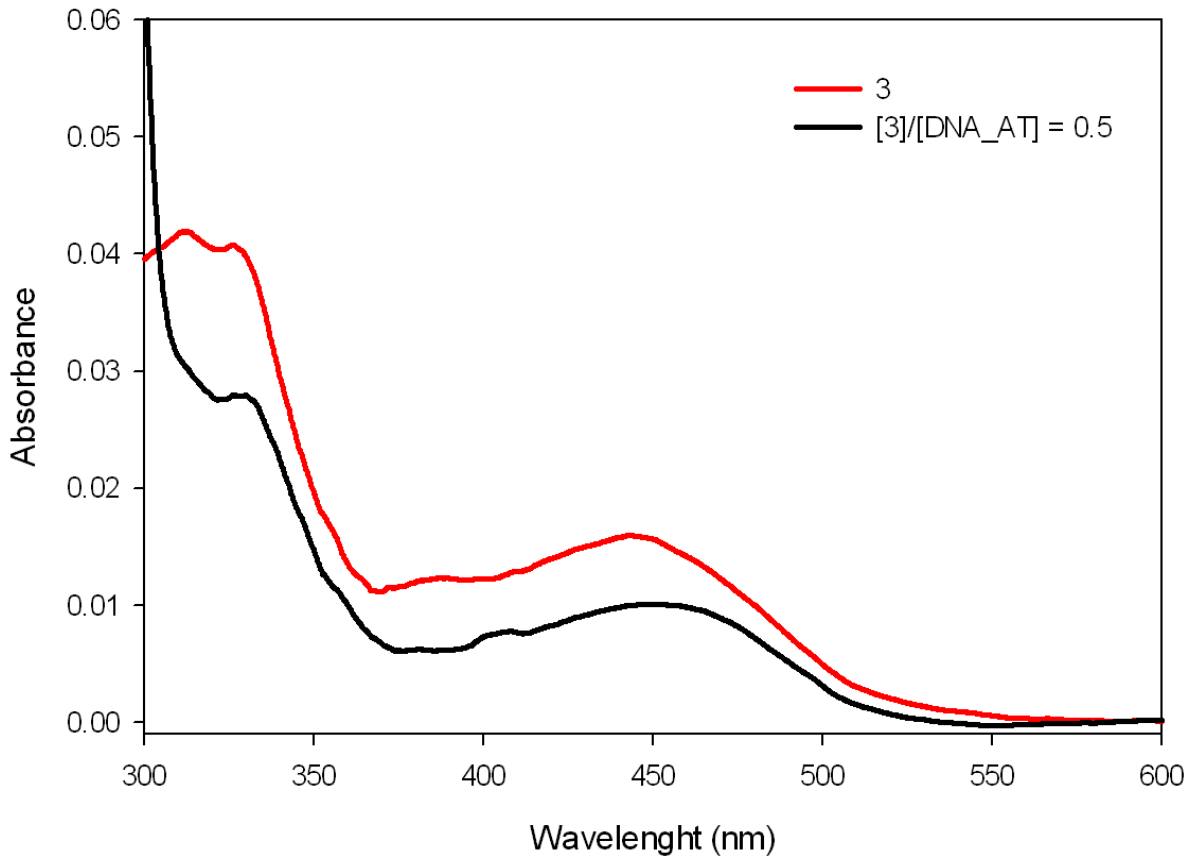


*Figure S51. UV-vis absorption spectra of metallopeptide **3** (2.9 μM) before (red line) and after (black line) the addition of a solution of AAAATTT oligonucleotide.*

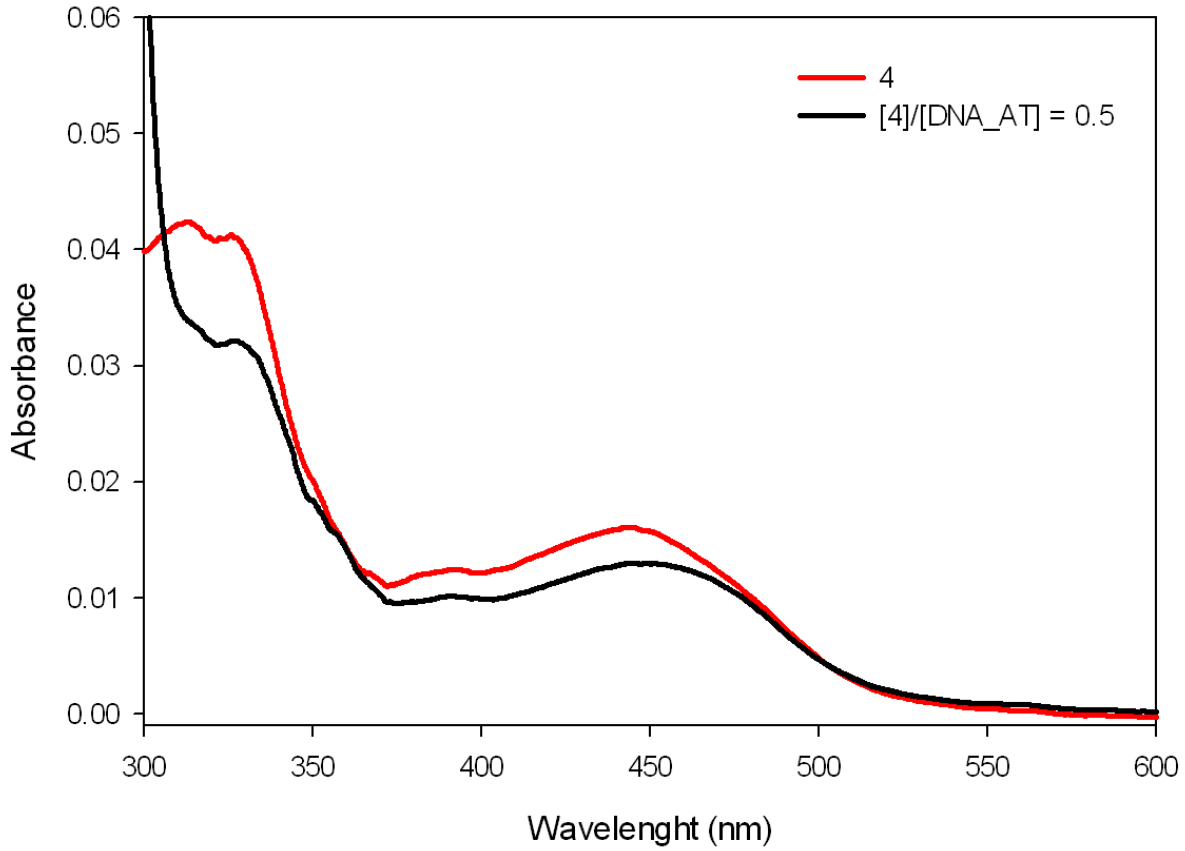


*Figure S52. UV-vis absorption spectra of metallopeptide **4** (2.7 μM) before (red line) and after (black line) the addition of a solution of AAAATTT oligonucleotide.*

b) Double-stranded "well-matched" and "mismatched" B-DNA binding studies

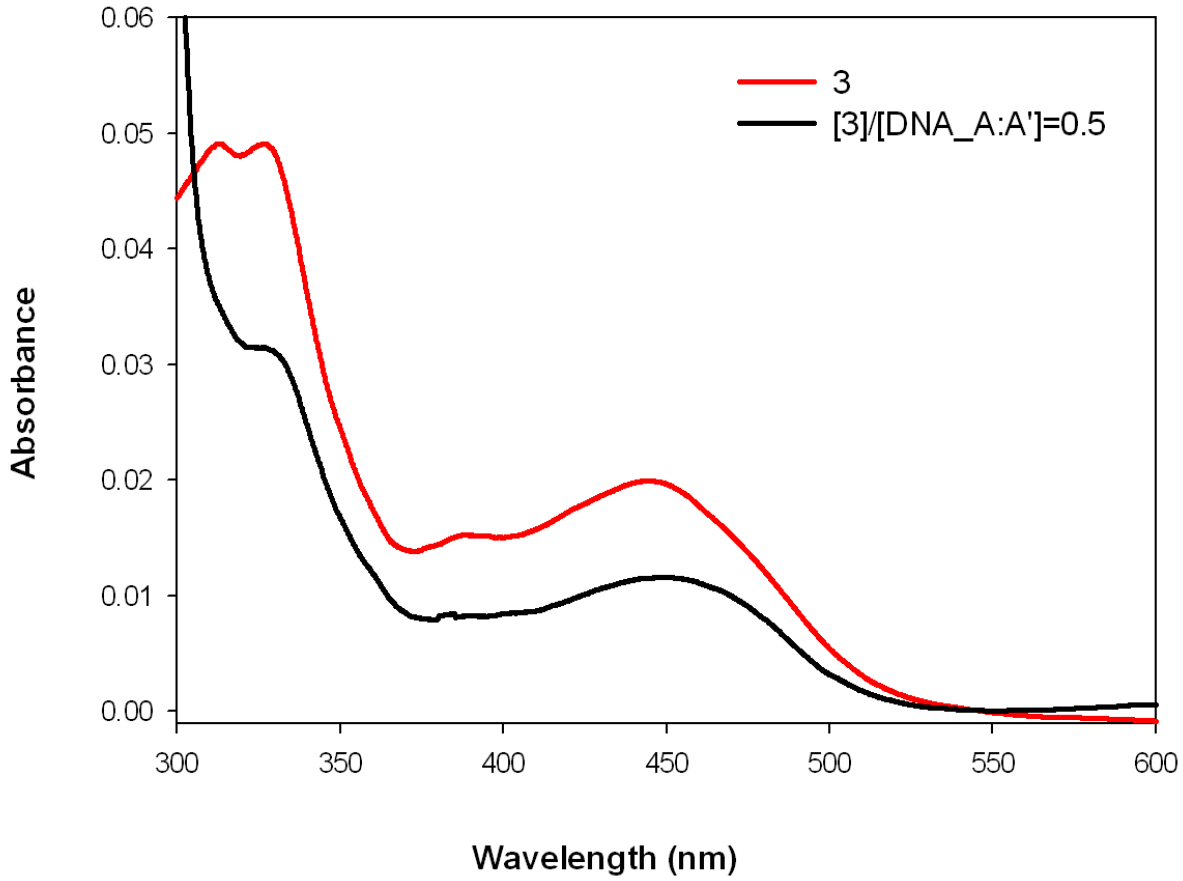


*Figure S53. UV-vis absorption spectra of metallopeptide **3** (3.4 μM) before (red line) and after (black line) the addition of a solution of "well matched" A:A' oligonucleotide.*

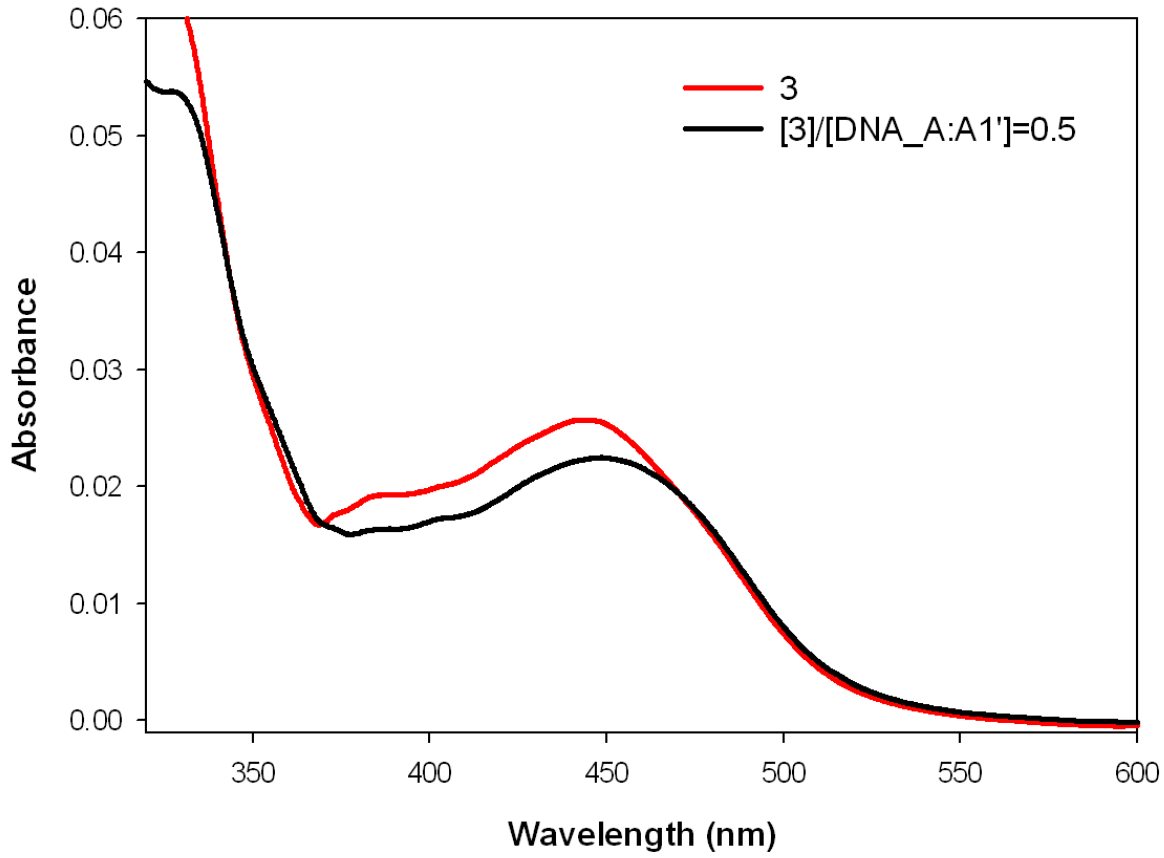


*Figure S54. UV-vis absorption spectra of metallopeptide **3** (4.4 μM) before (red line) and after (black line) the addition of a solution of GG mismatched A:A1' oligonucleotide.*

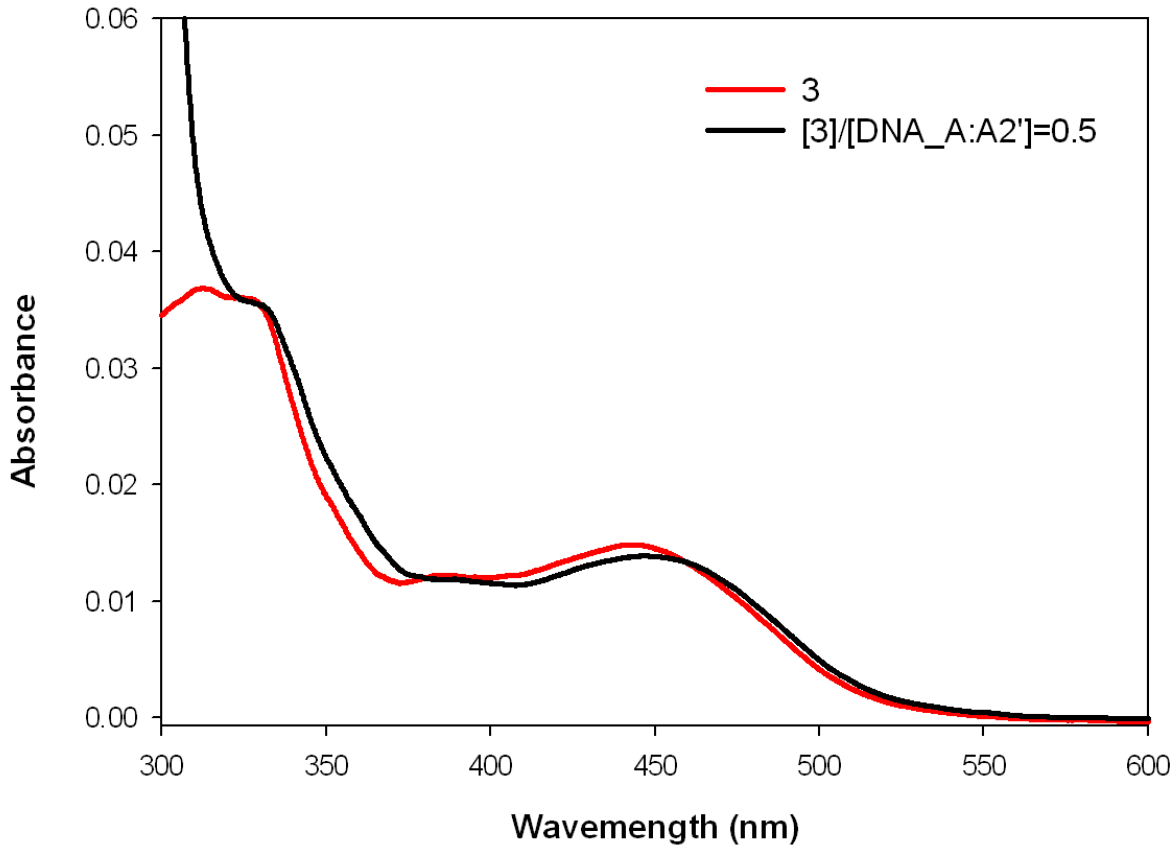


*Figure S55. UV-vis absorption spectra of metallopeptide **3** (2.5 μM) before (red line) and after (black line) the addition of a solution of GG/AA mismatched A:A2' oligonucleotide.*

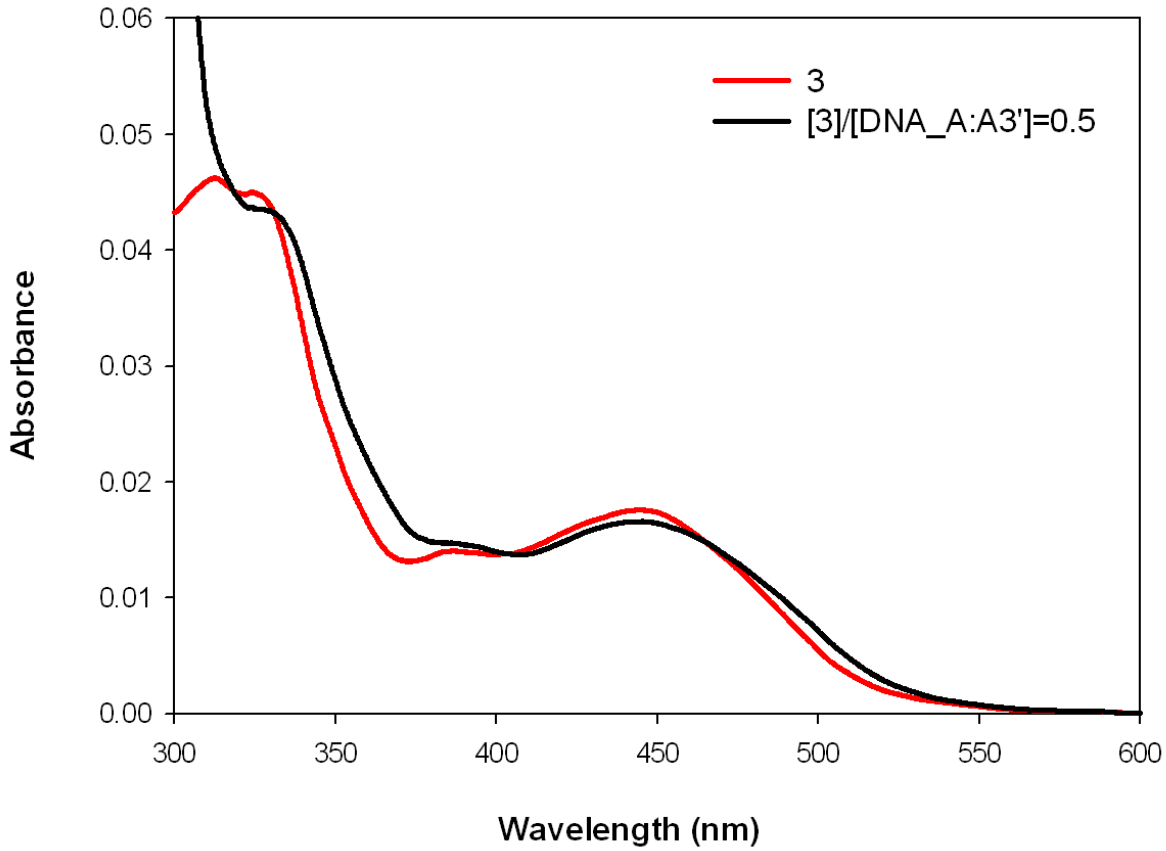


*Figure S56. UV-vis absorption spectra of metallopeptide **3** (3.0 μM) before (red line) and after (black line) the addition of a solution of GG/GG/AA mismatched A:A3' oligonucleotide.*

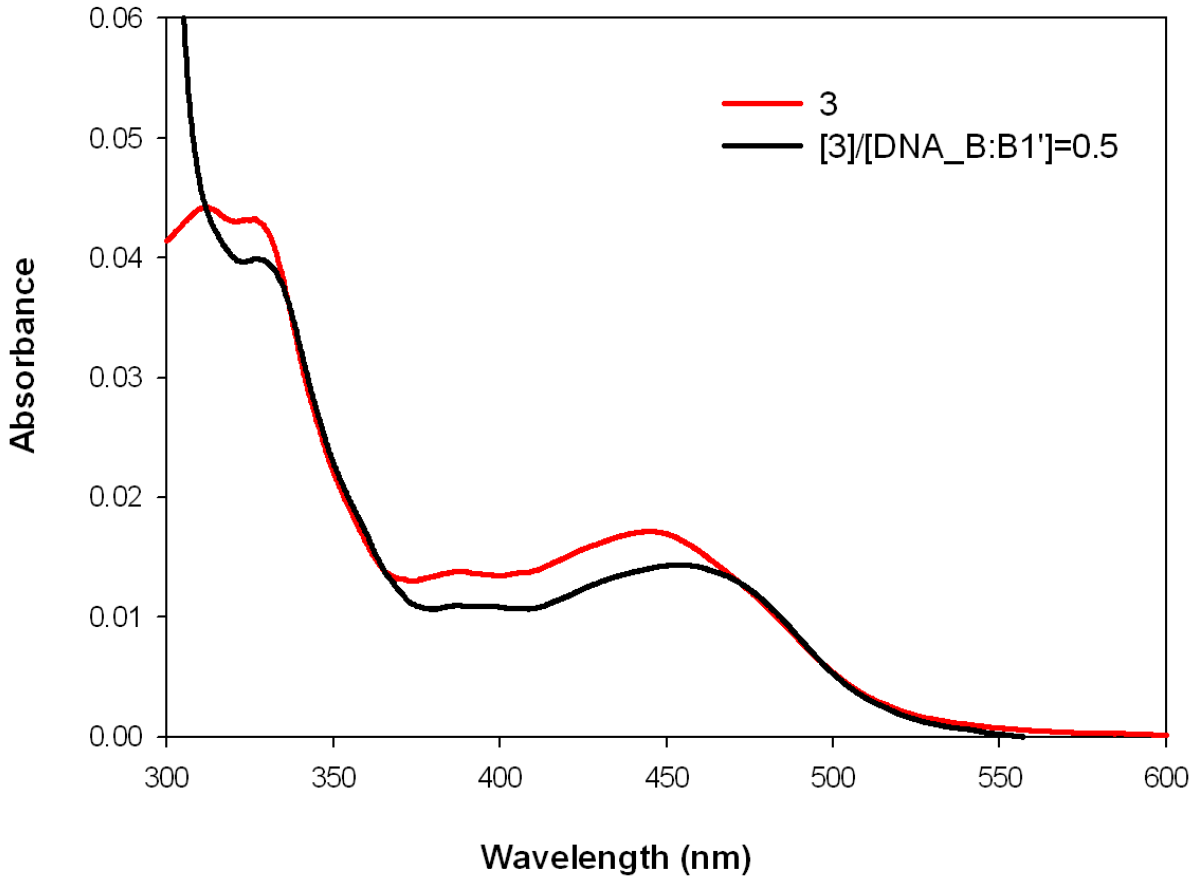


*Figure S57. UV-vis absorption spectra of metallopeptide **3** (2.9 μM) before (red line) and after (black line) the addition of a solution of AA mismatched B:B1' oligonucleotide.*

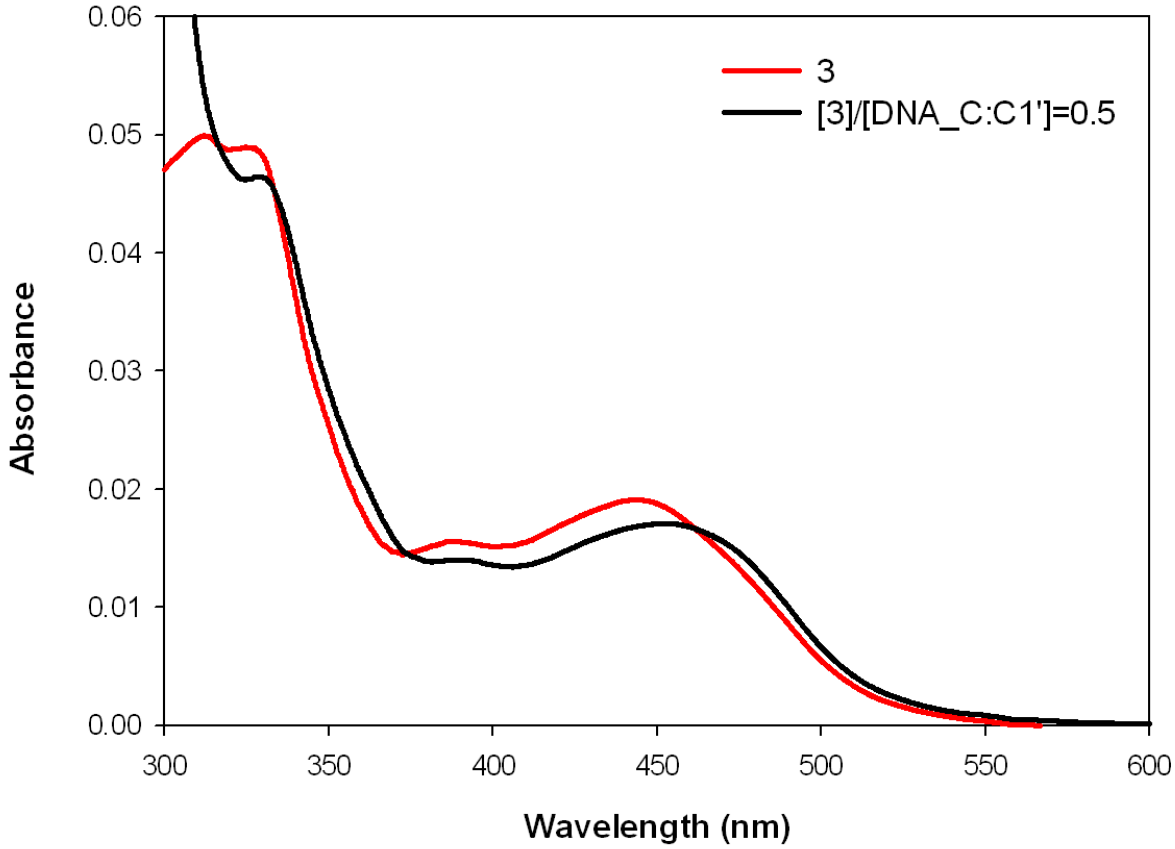


*Figure S58. UV-vis absorption spectra of metallopeptide **3** (3.2 μM) before (red line) and after (black line) the addition of a solution of CC mismatched C:C1' oligonucleotide.*

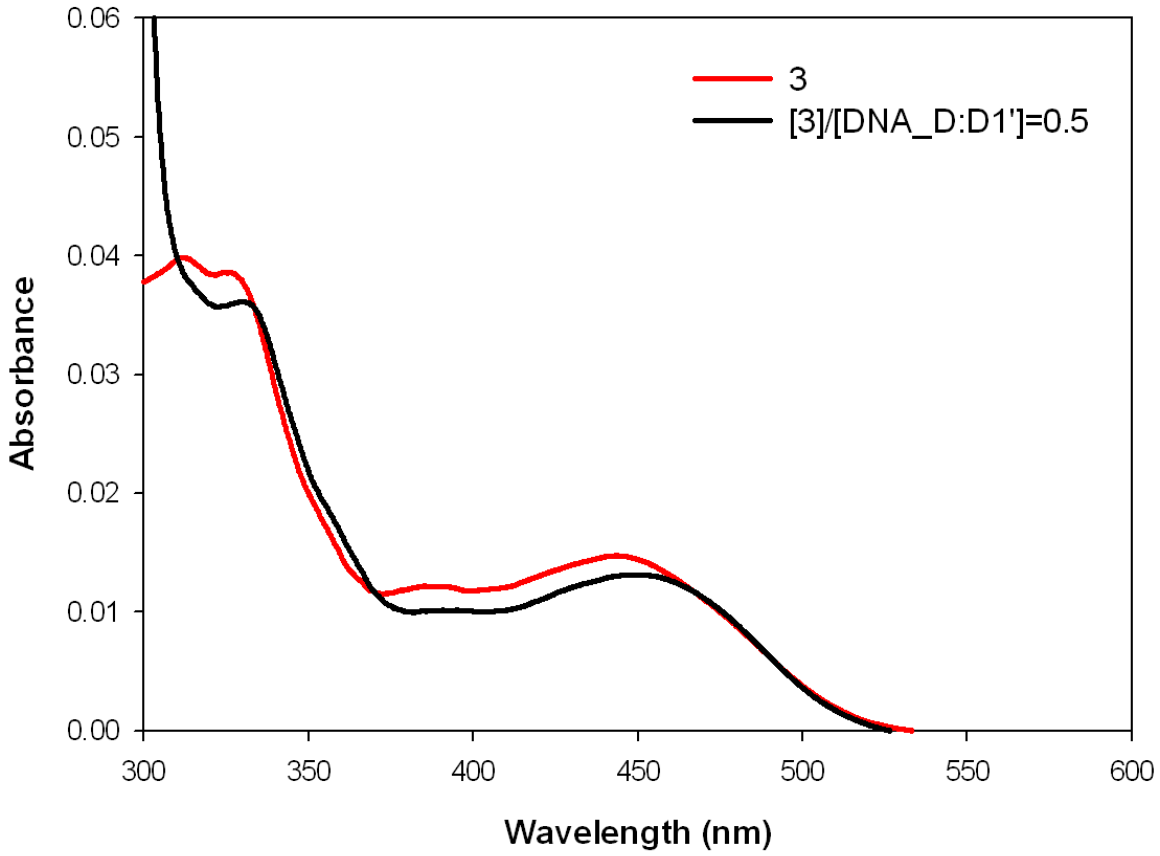


*Figure S59. UV-vis absorption spectra of metallopeptide **3** (2.7 μM) before (red line) and after (black line) the addition of a solution of TT mismatched D:D1' oligonucleotide.*

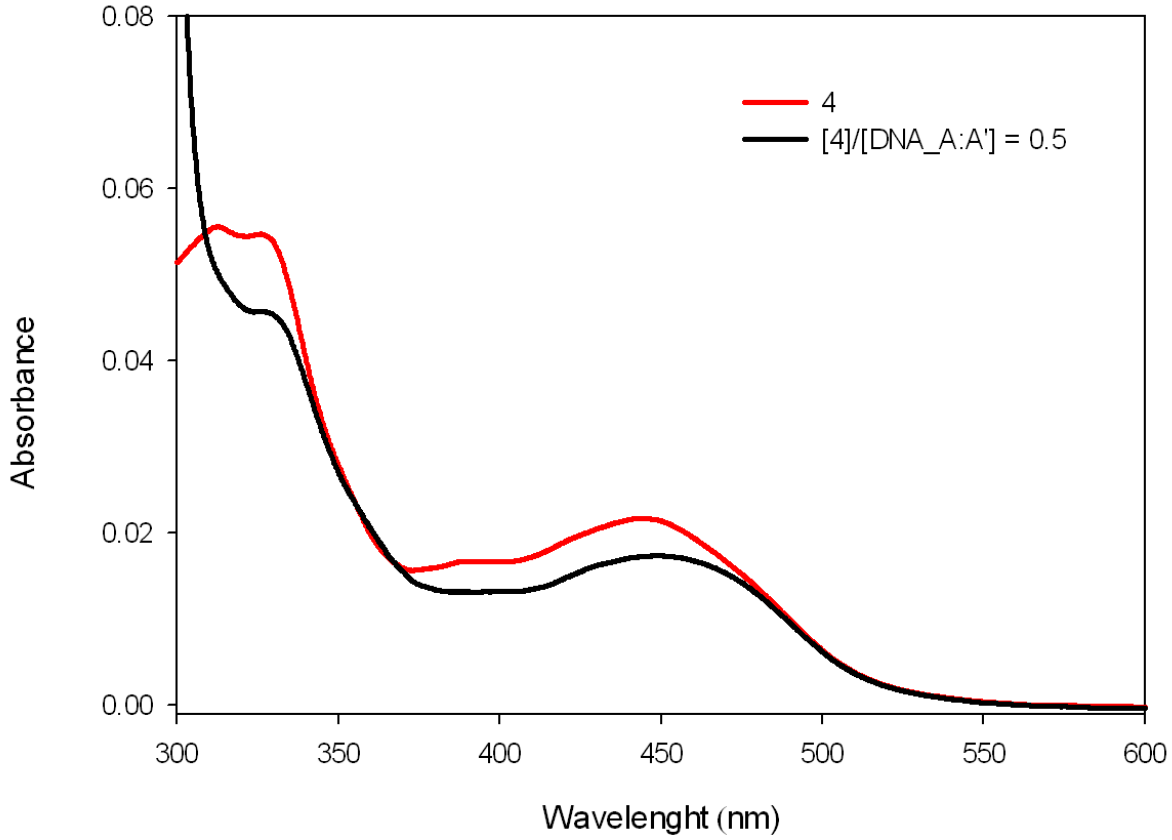


*Figure S60. UV-vis absorption spectra of metallopeptide **4** (3.7 μM) before (red line) and after (black line) the addition of a solution of "well matched" A:A' oligonucleotide.*

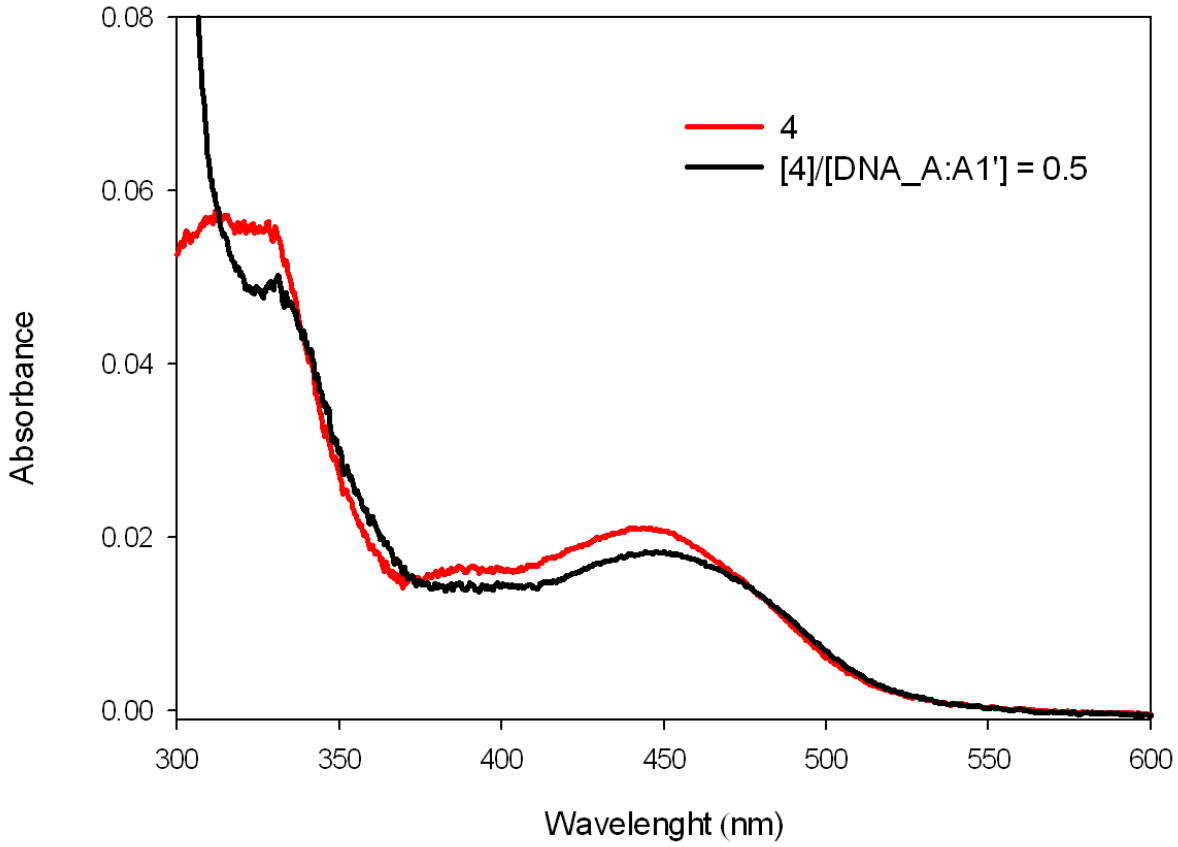


*Figure S61. UV-vis absorption spectra of metallopeptide **4** (3.7 μM) before (red line) and after (black line) the addition of a solution of GG mismatched A:A1' oligonucleotide.*

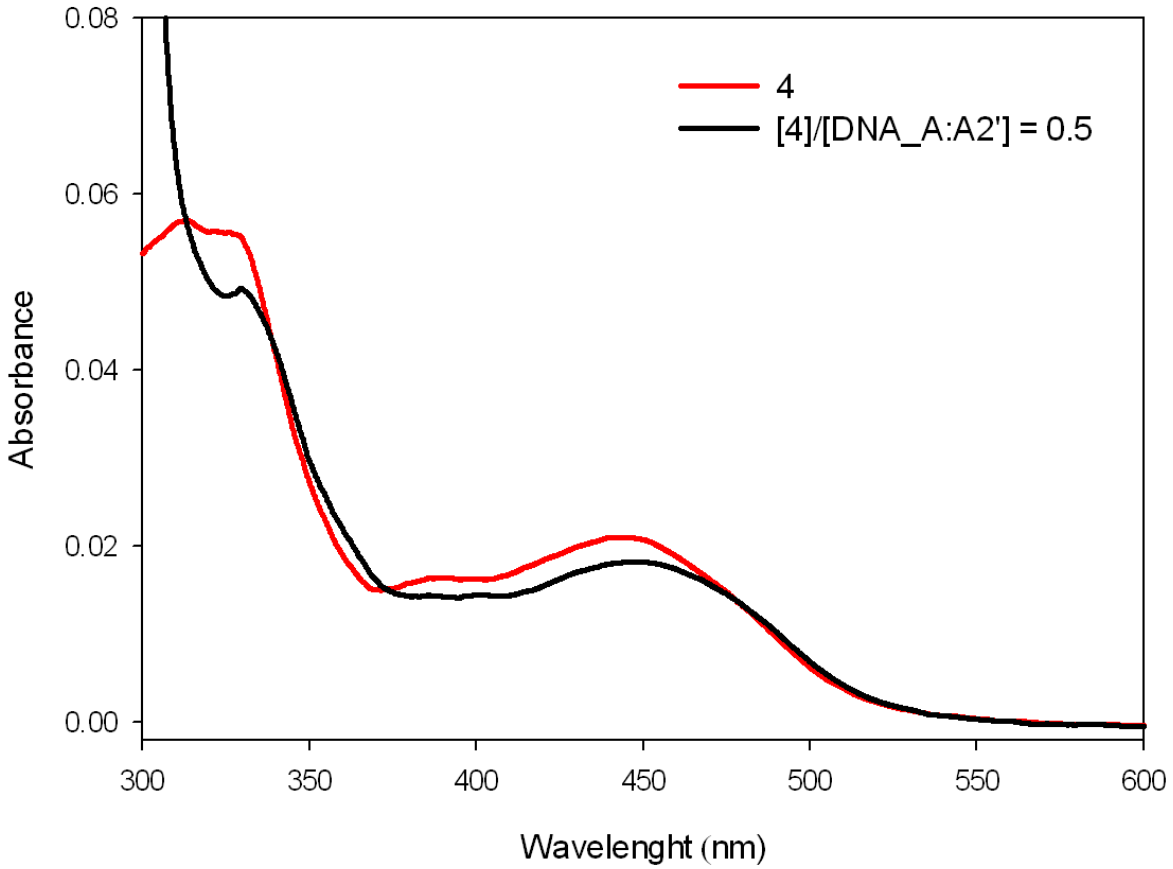


*Figure S62. UV-vis absorption spectra of metallopeptide **4** (3.5 μM) before (red line) and after (black line) the addition of a solution of GG/AA mismatched A:A2' oligonucleotide.*

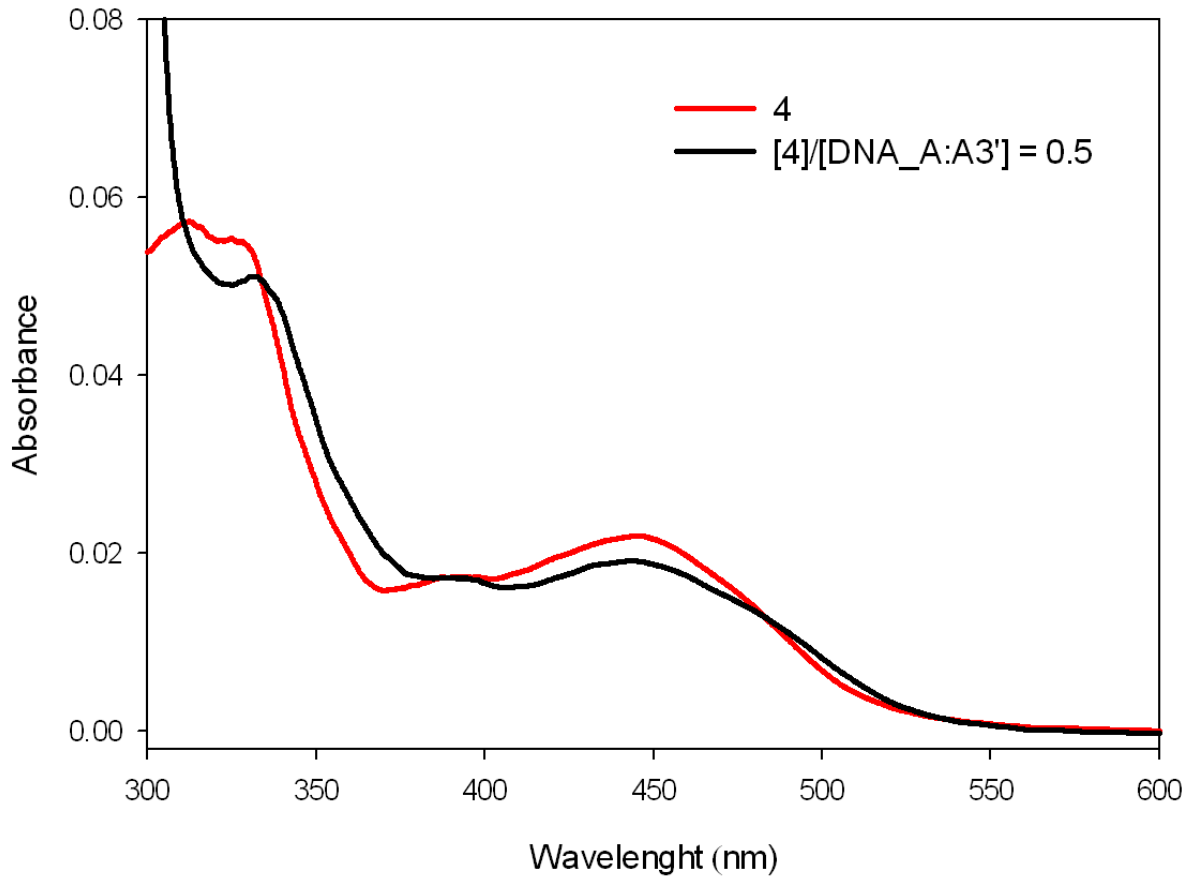


*Figure S63. UV-vis absorption spectra of metallopeptide **4** (3.5 μM) before (red line) and after (black line) the addition of a solution of GG/GG/AA mismatched A:A3' oligonucleotide.*

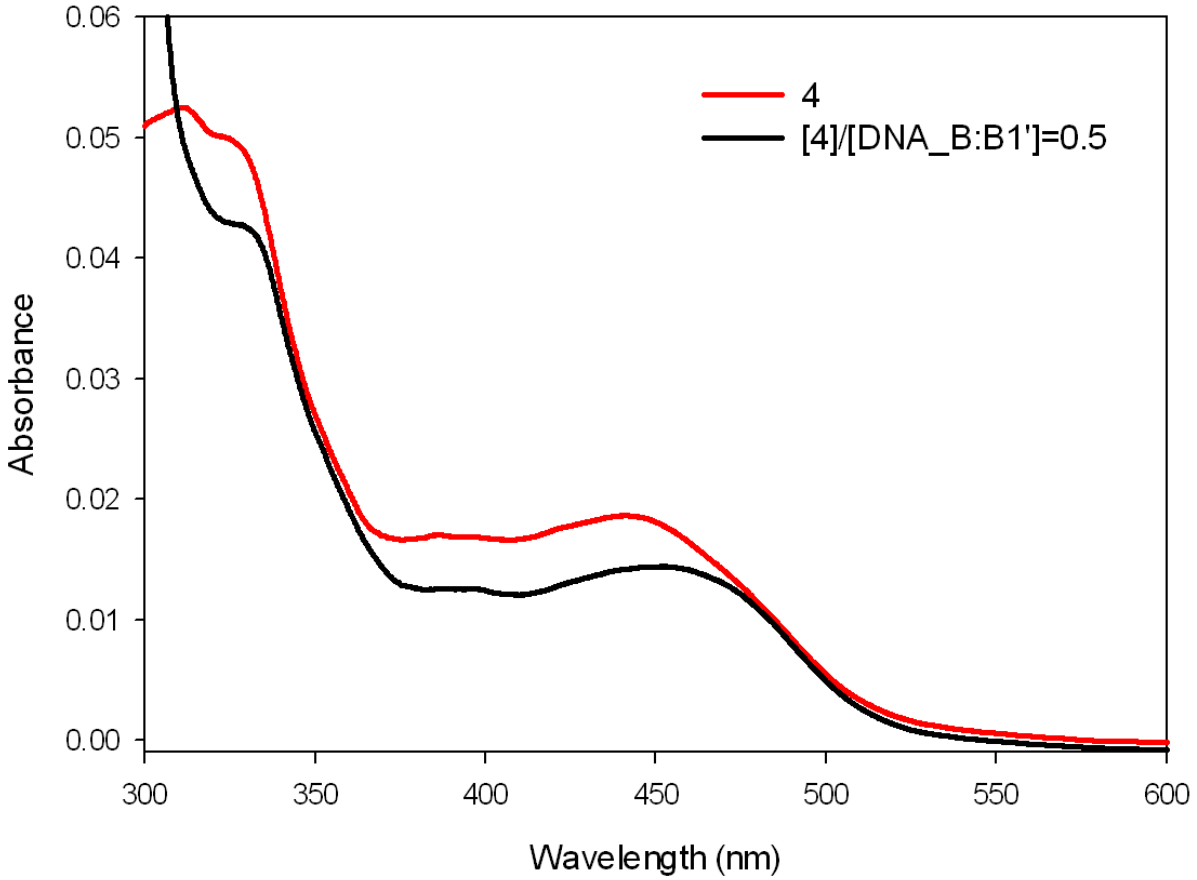


*Figure S64. UV-vis absorption spectra of metallopeptide **4** (3.5 μM) before (red line) and after (black line) the addition of a solution of AA mismatched B:B1' oligonucleotide.*

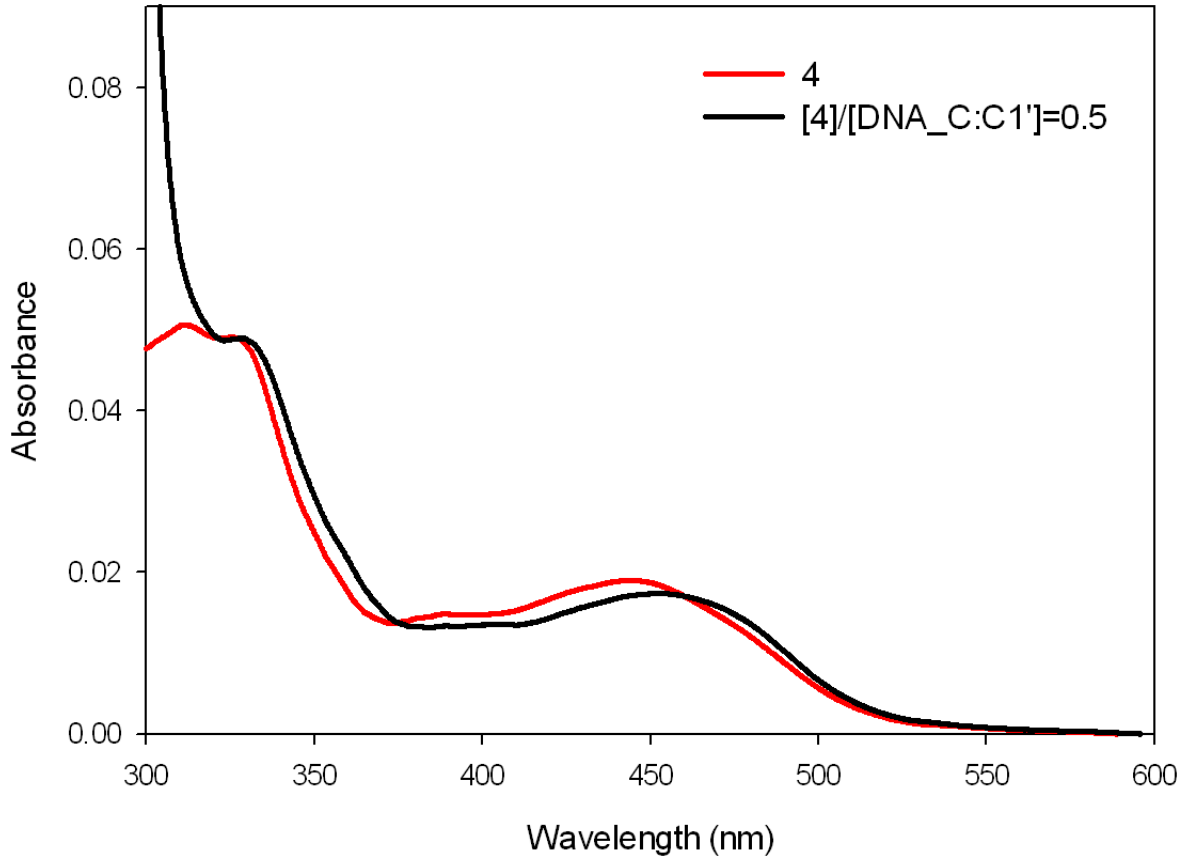


*Figure S65. UV-vis absorption spectra of metallopeptide **4** (3.6 μM) before (red line) and after (black line) the addition of a solution of CC mismatched C:C1' oligonucleotide.*

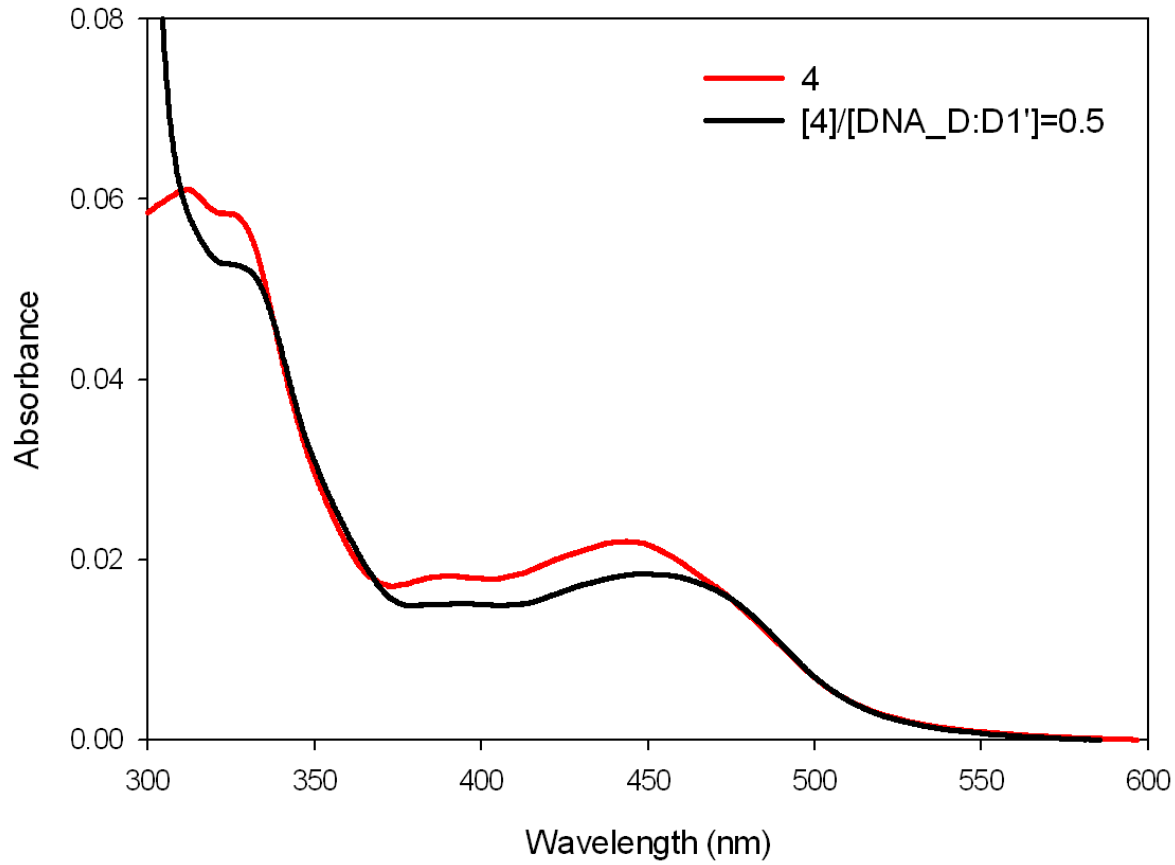


*Figure S66. UV-vis absorption spectra of metallopeptide **4** (3.7 μM) before (red line) and after (black line) the addition of a solution of TT mismatched D:D1' oligonucleotide.*

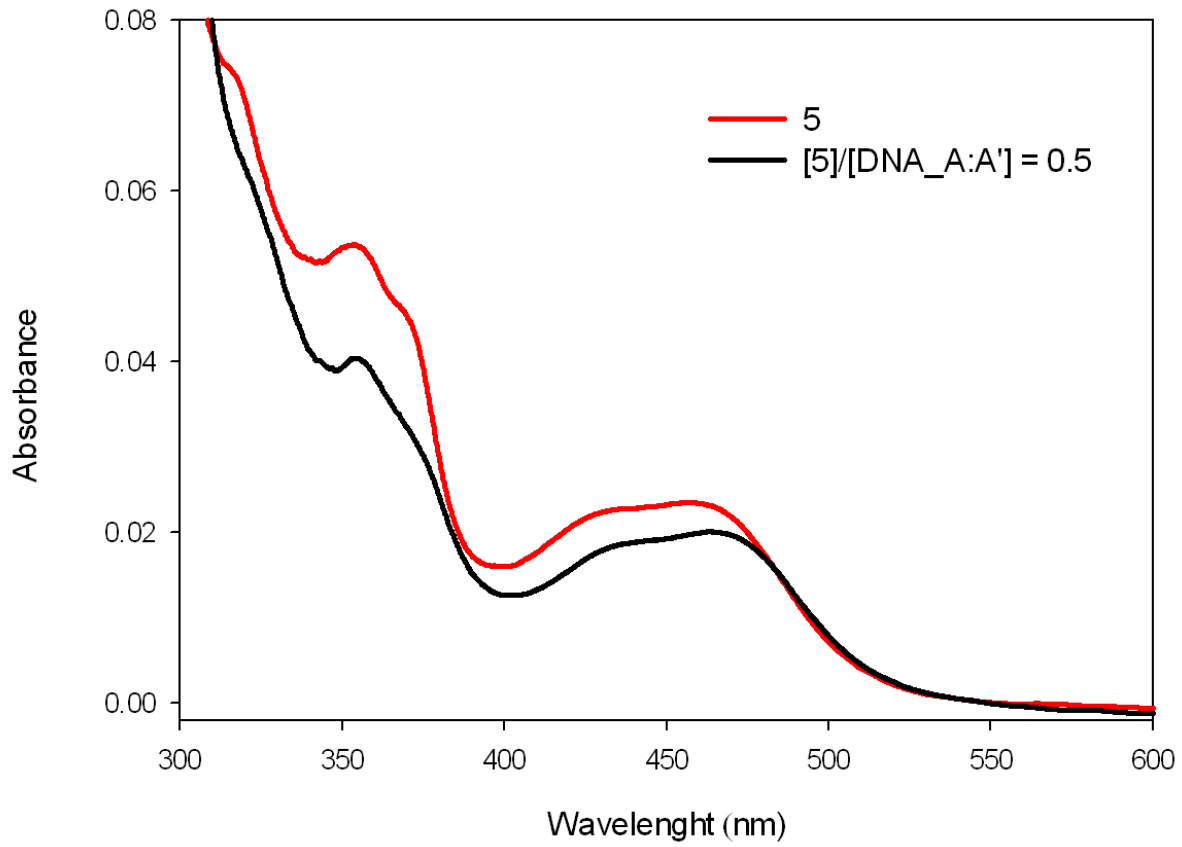


Figure S67. UV-vis absorption spectra of metallopeptide ***5*** *(3.7 μM) before (red line) and after (black line) the addition of a solution of "well matched" A:A' oligonucleotide.*

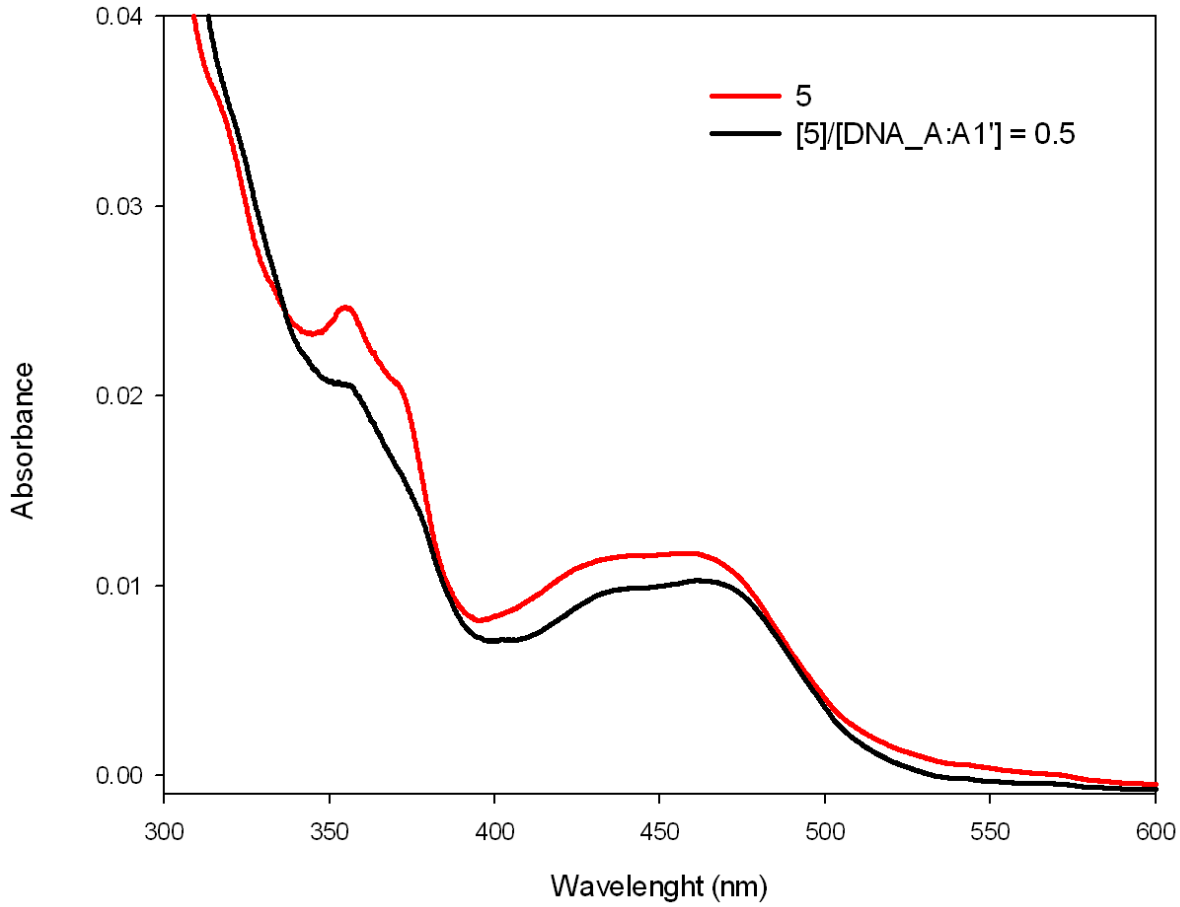


*Figure S68. UV-vis absorption spectra of metallopeptide **5** (2.0 μM) before (red line) and after (black line) the addition of a solution of GG mismatched A:A1' oligonucleotide.*

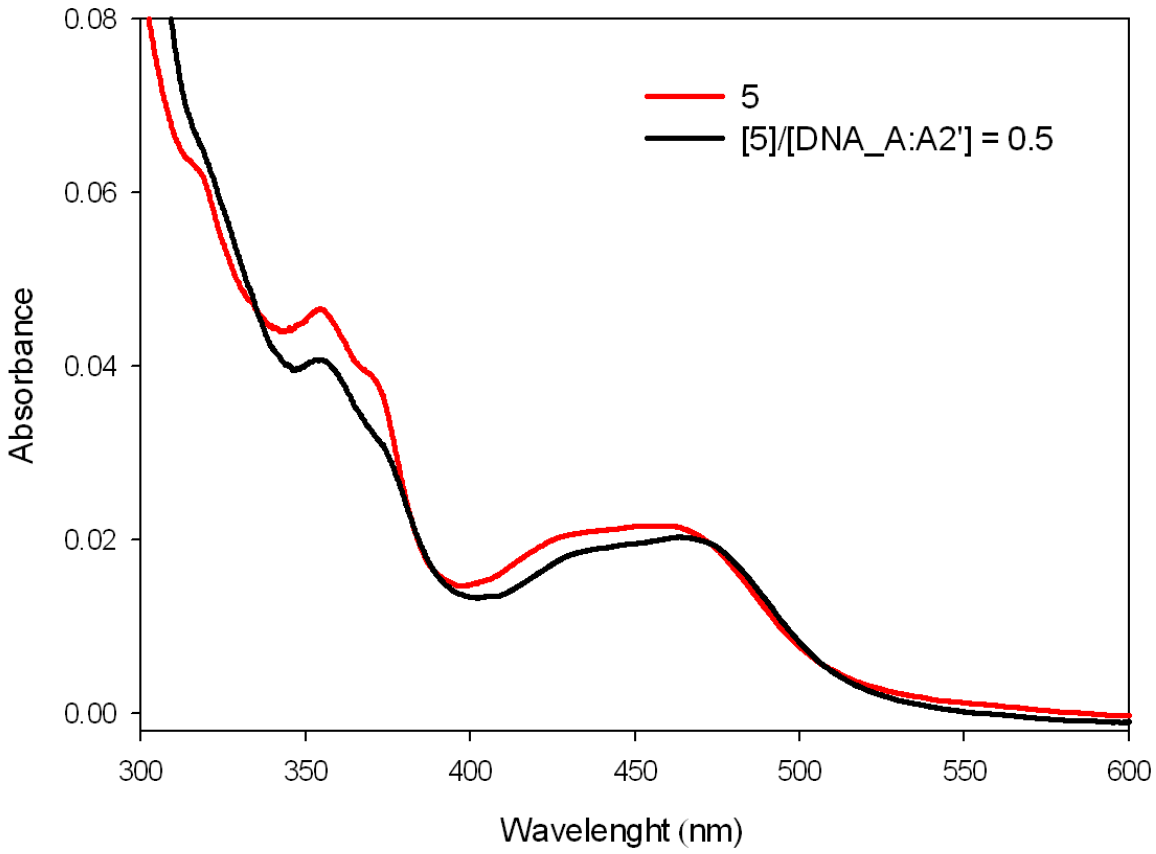


Figure S69. UV-vis absorption spectra of metallopeptide ***5*** *(3.6 μM) before (red line) and after (black line) the addition of a solution of GG/AA mismatched A:A2' oligonucleotide.*

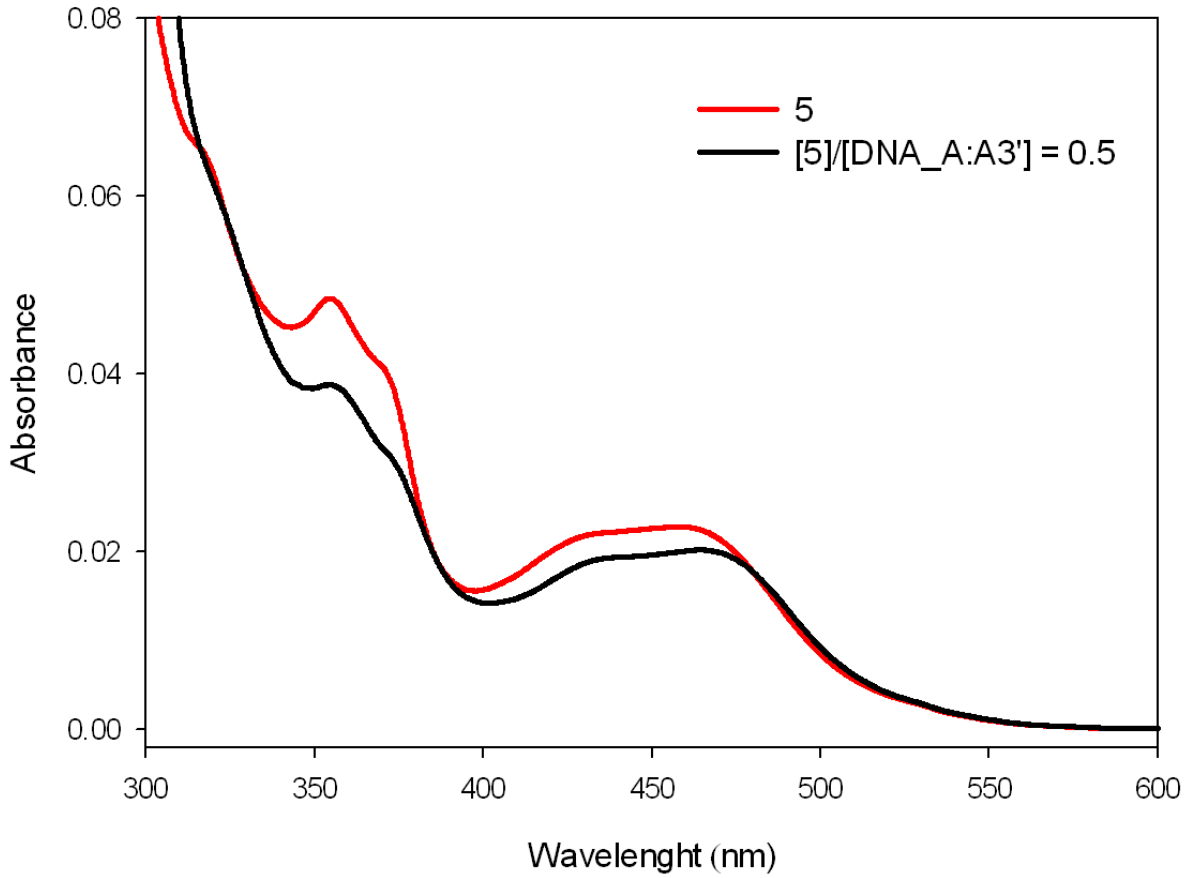


*Figure S70. UV-vis absorption spectra of metallopeptide **5** (3.7 μM) before (red line) and after (black line) the addition of a solution of GG/GG/AA mismatched A:A3' oligonucleotide.*

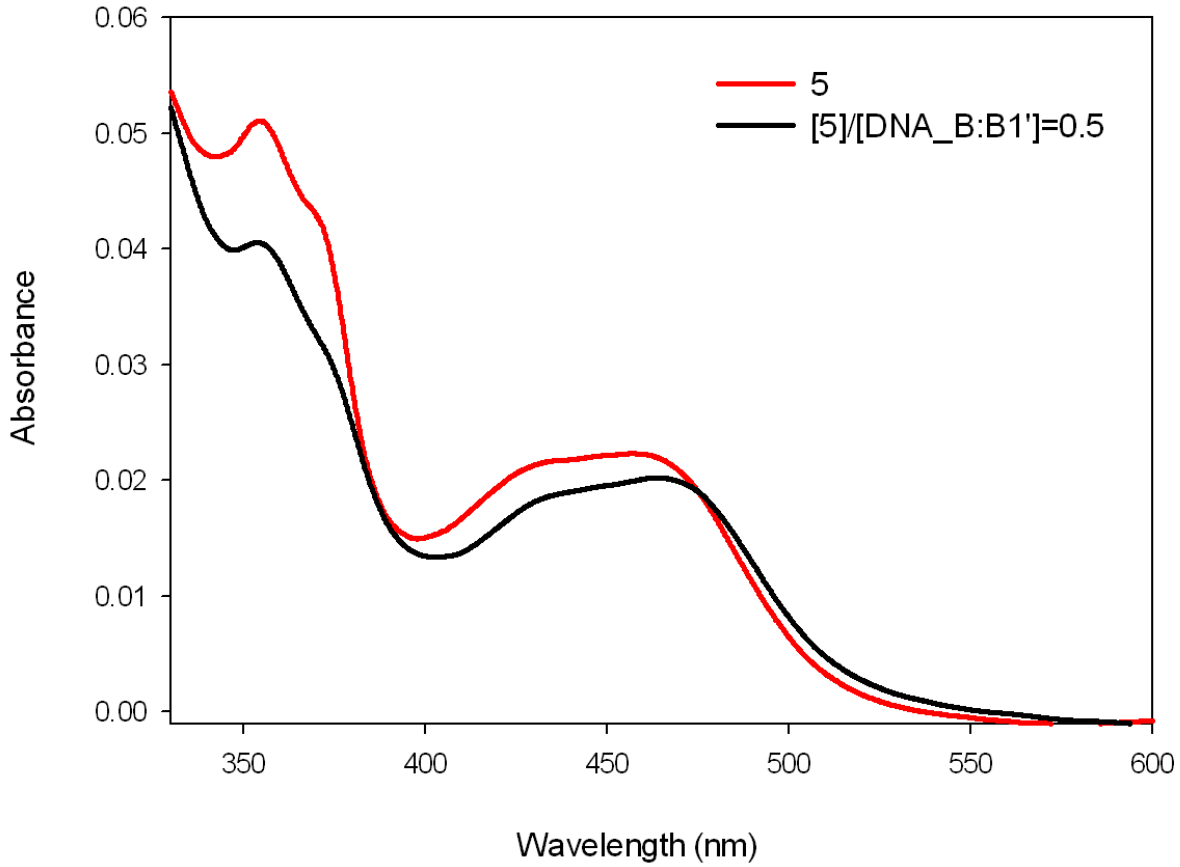


Figure S71. UV-vis absorption spectra of metallopeptide ***5*** *(3.5 μM) before (red line) and after (black line) the addition of a solution of AA mismatched B:B1' oligonucleotide.*

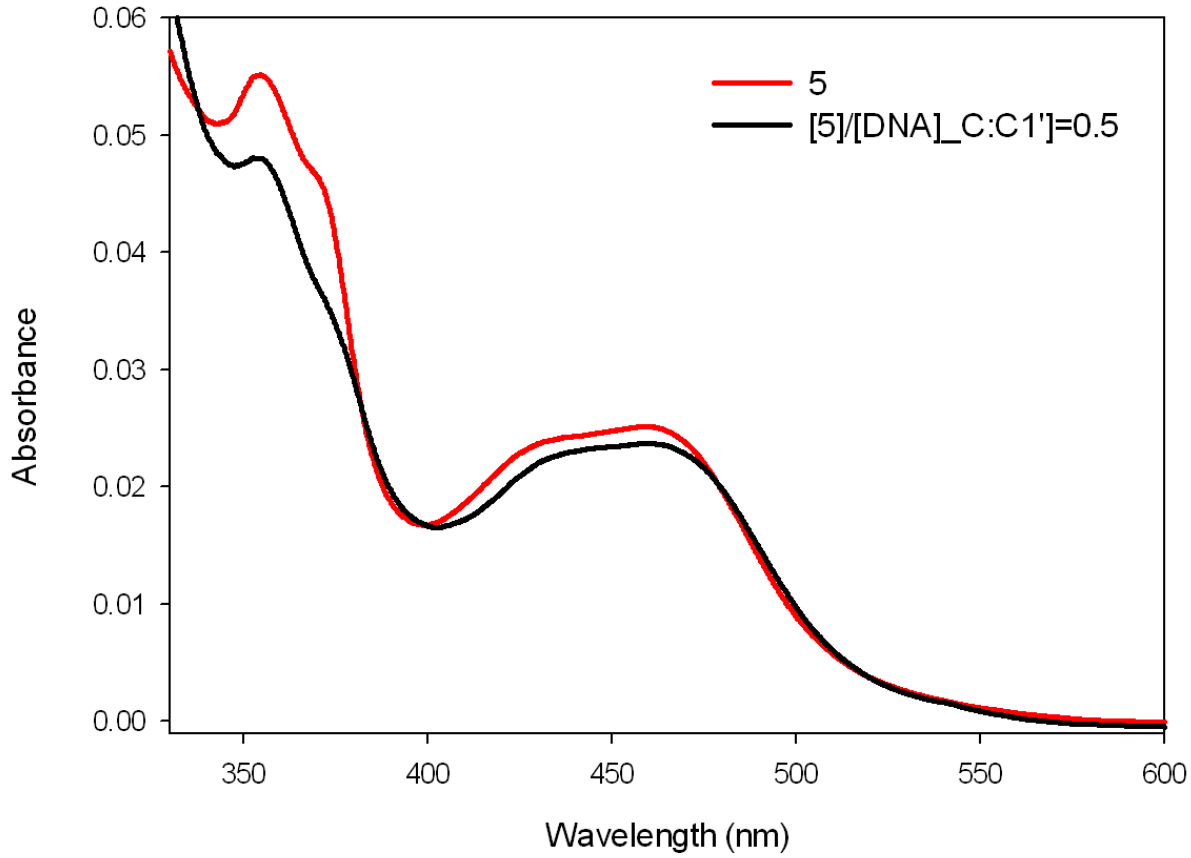


*Figure S72. UV-vis absorption spectra of metallopeptide **5** (4.0 μM) before (red line) and after (black line) the addition of a solution of CC mismatched C:C1' oligonucleotide.*

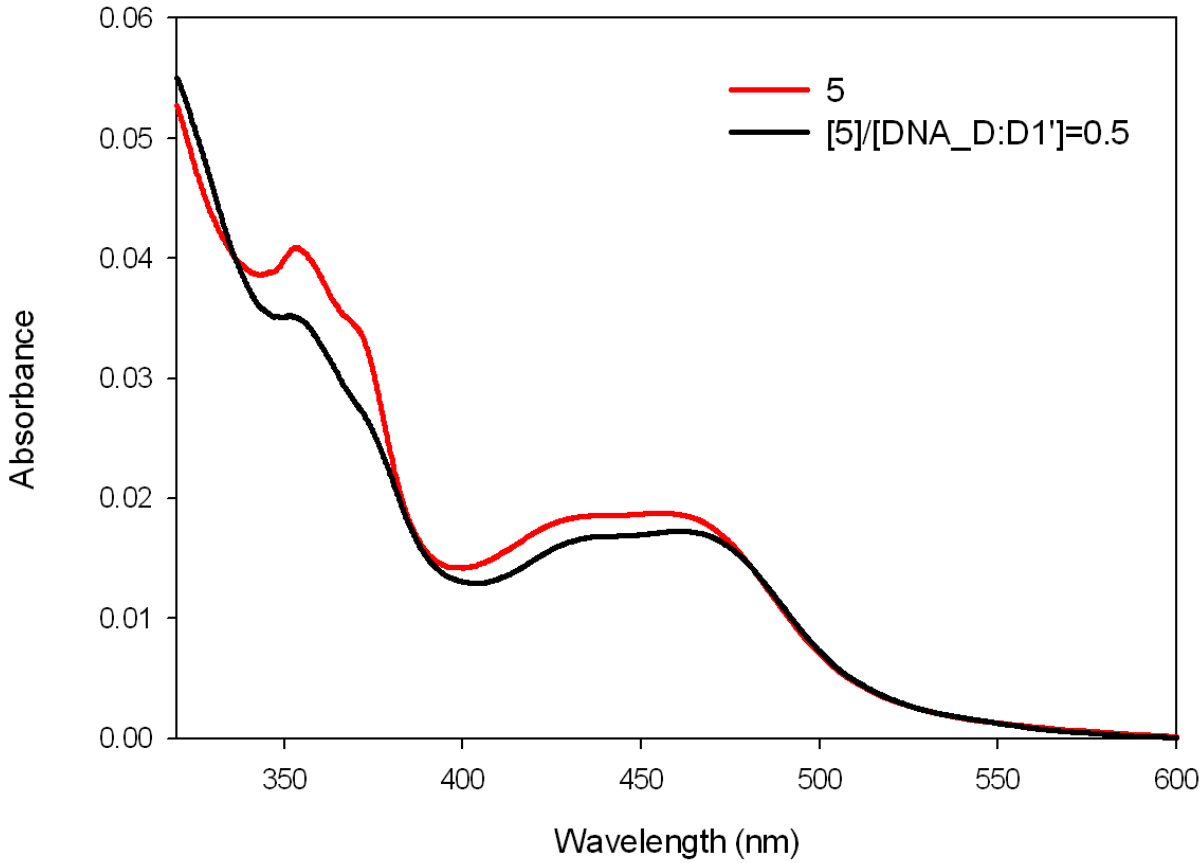


Figure S73. UV-vis absorption spectra of metallopeptide ***5*** *(3.1 μM) before (red line) and after (black line) the addition of a solution of TT mismatched D:D1' oligonucleotide.*

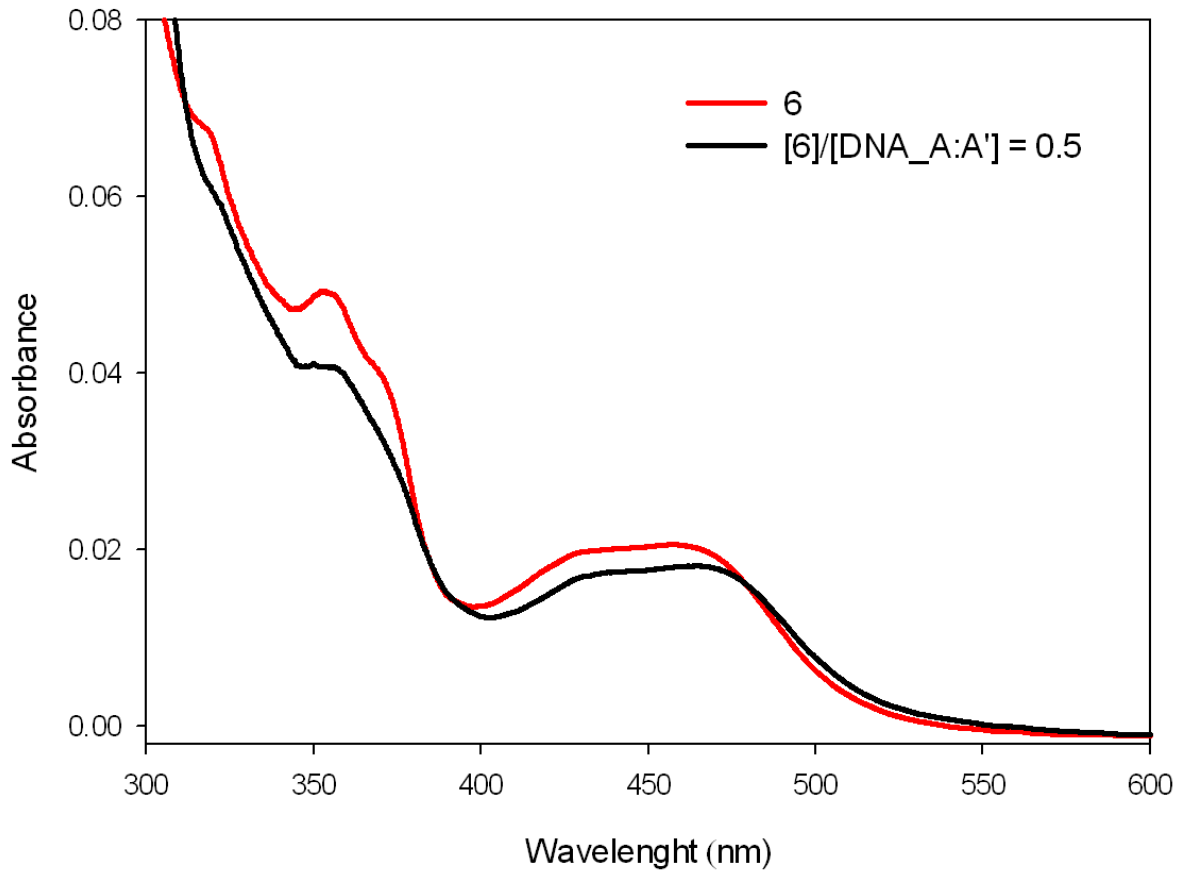


*Figure S74. UV-vis absorption spectra of metallopeptide **6** (3.6 μM) before (red line) and after (black line) the addition of a solution of "well matched" A:A' oligonucleotide.*

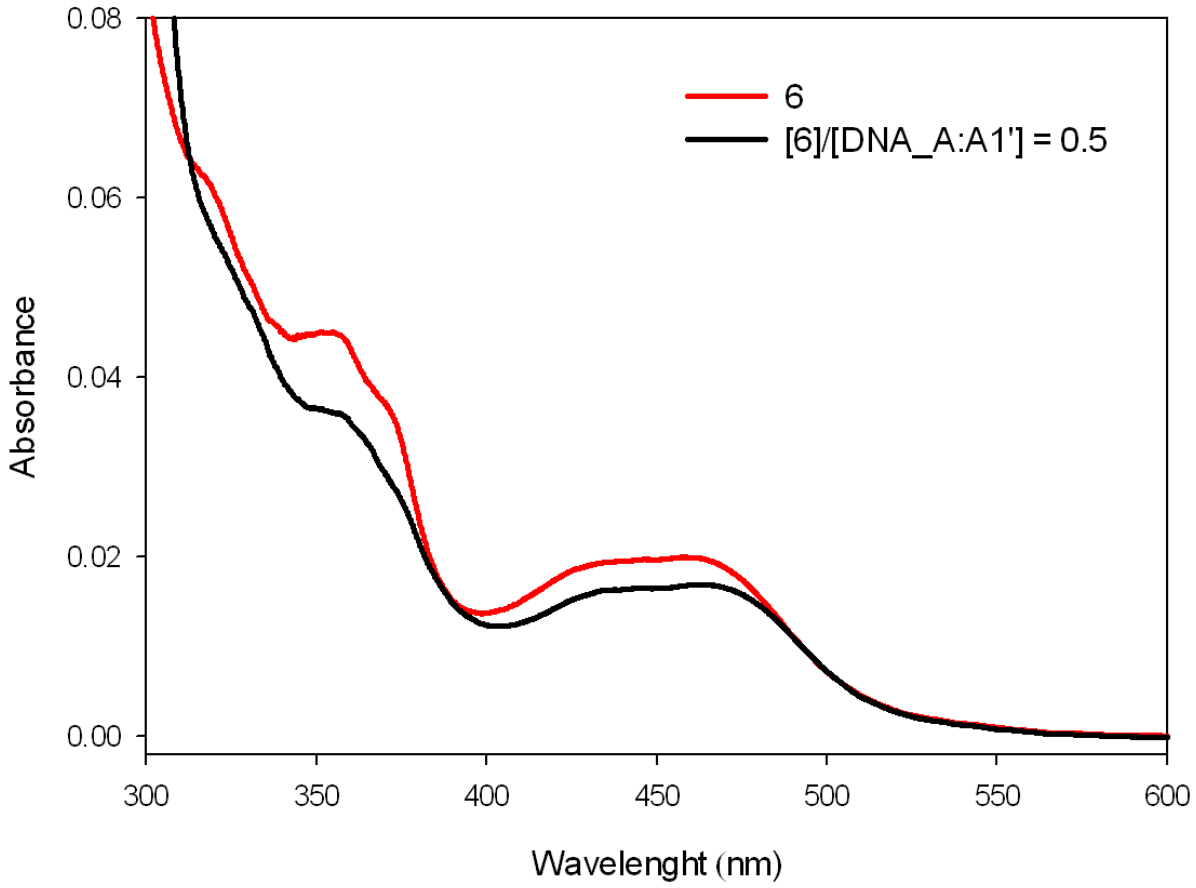


*Figure S75. UV-vis absorption spectra of metallopeptide **6** (3.5 μM) before (red line) and after (black line) the addition of a solution of GG mismatched A:A1' oligonucleotide.*

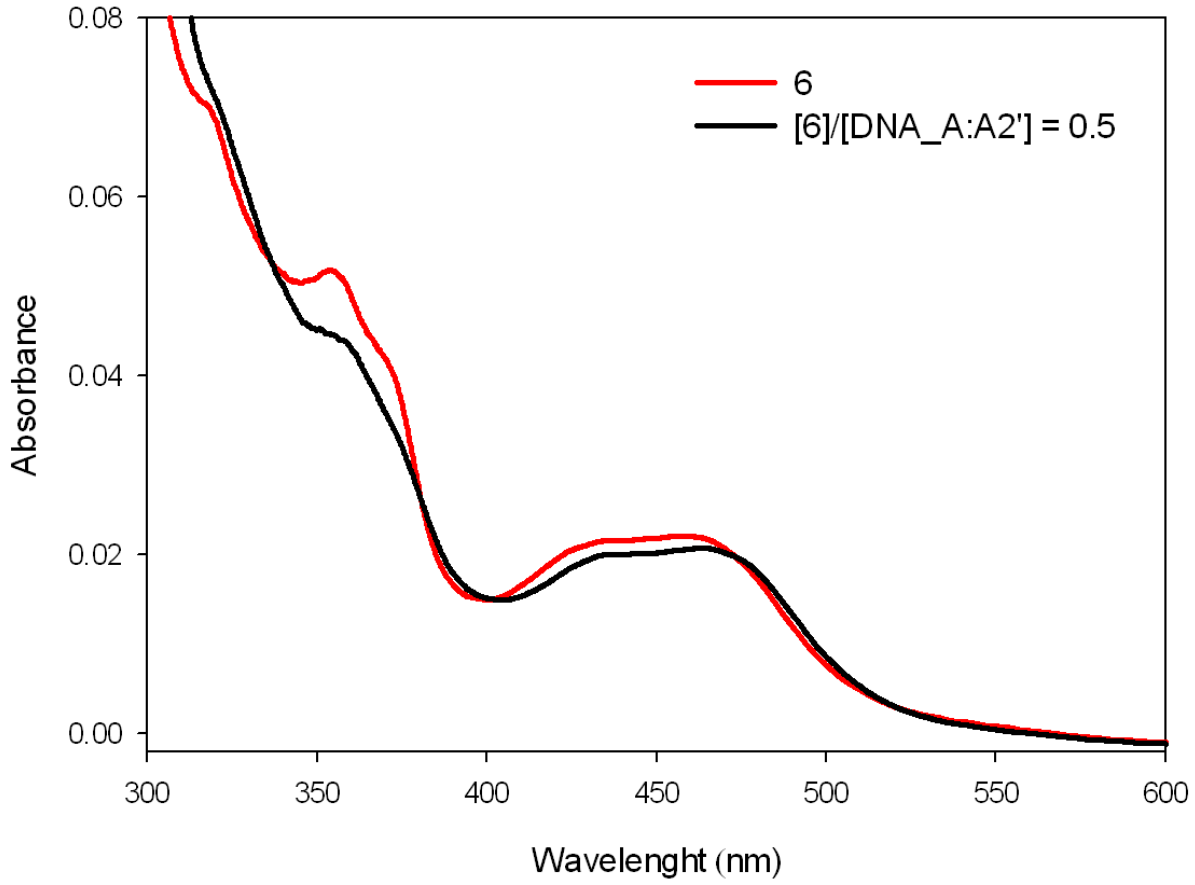


*Figure S76. UV-vis absorption spectra of metallopeptide **6** (3.7 μM) before (red line) and after (black line) the addition of a solution of GG/AA mismatched A:A2' oligonucleotide.*

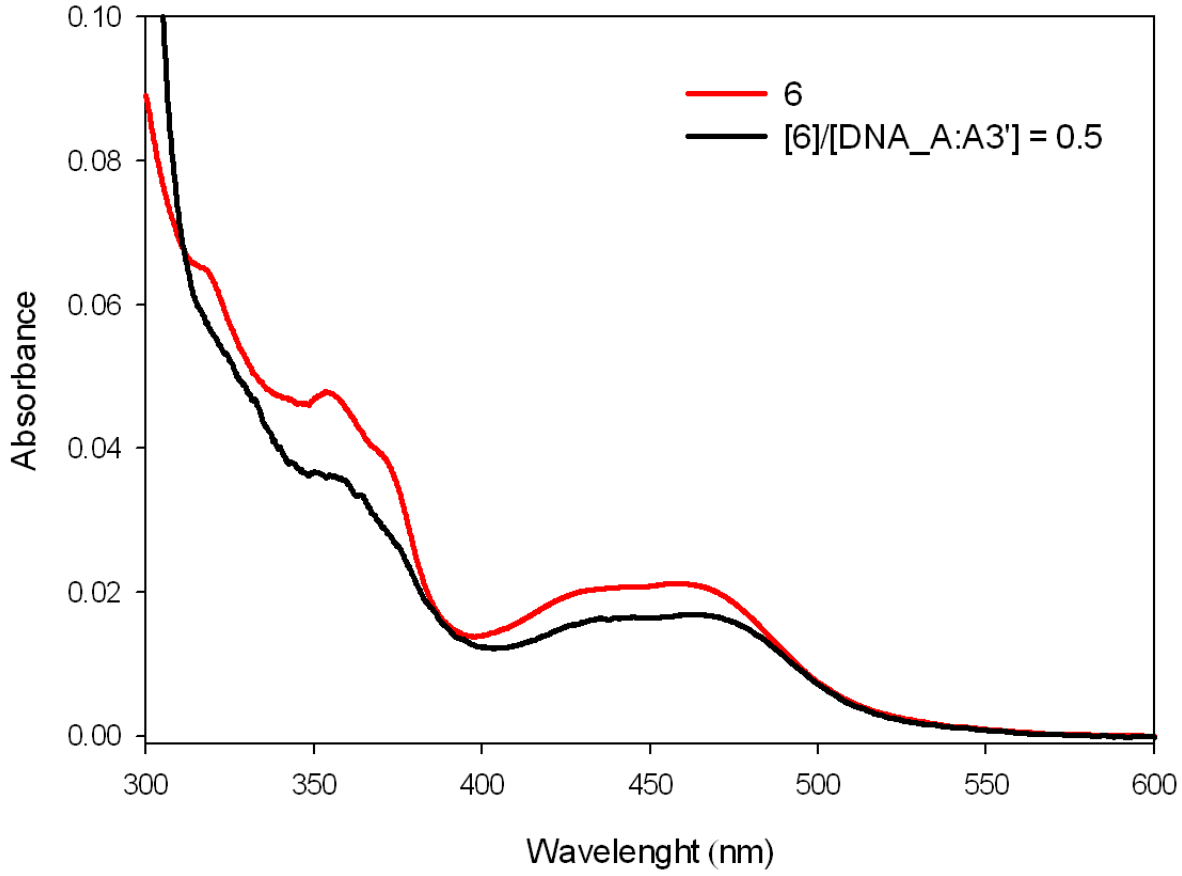


*Figure S77. UV-vis absorption spectra of metallopeptide **6** (3.7 μM) before (red line) and after (black line) the addition of a solution of GG/GG/AA mismatched A:A3' oligonucleotide.*

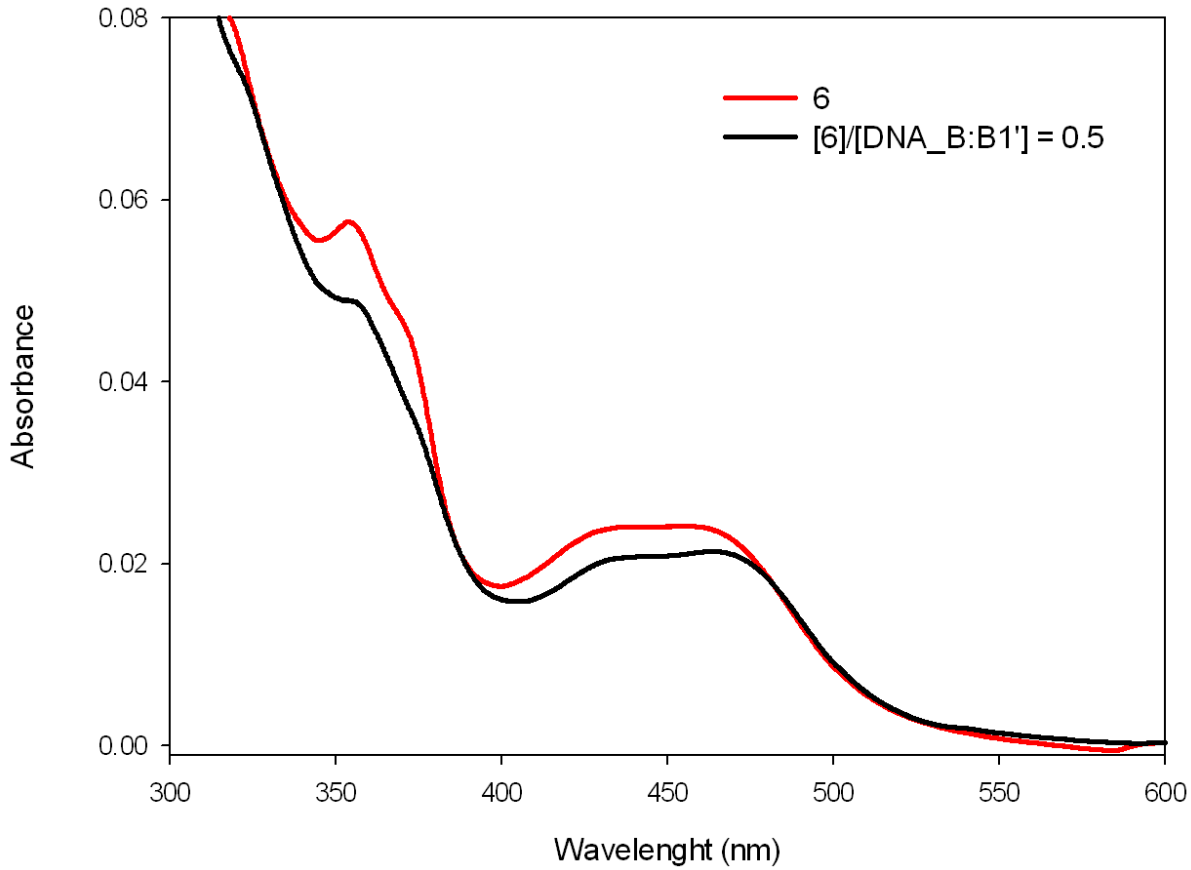


*Figure S78. UV-vis absorption spectra of metallopeptide **6** (4.0 μM) before (red line) and after (black line) the addition of a solution of AA mismatched B:B1' oligonucleotide.*

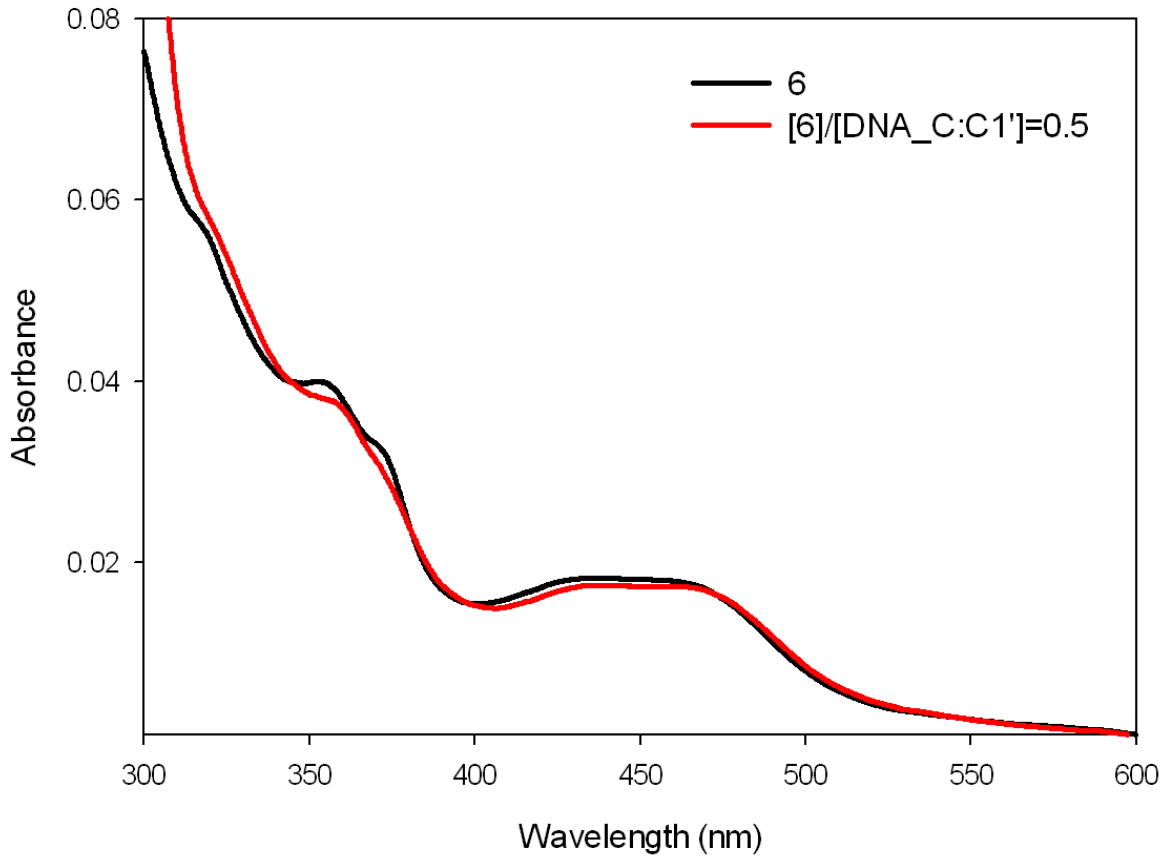


*Figure S79. UV-vis absorption spectra of metallopeptide **6** (3.0 μM) before (red line) and after (black line) the addition of a solution of CC mismatched C:C1' oligonucleotide.*

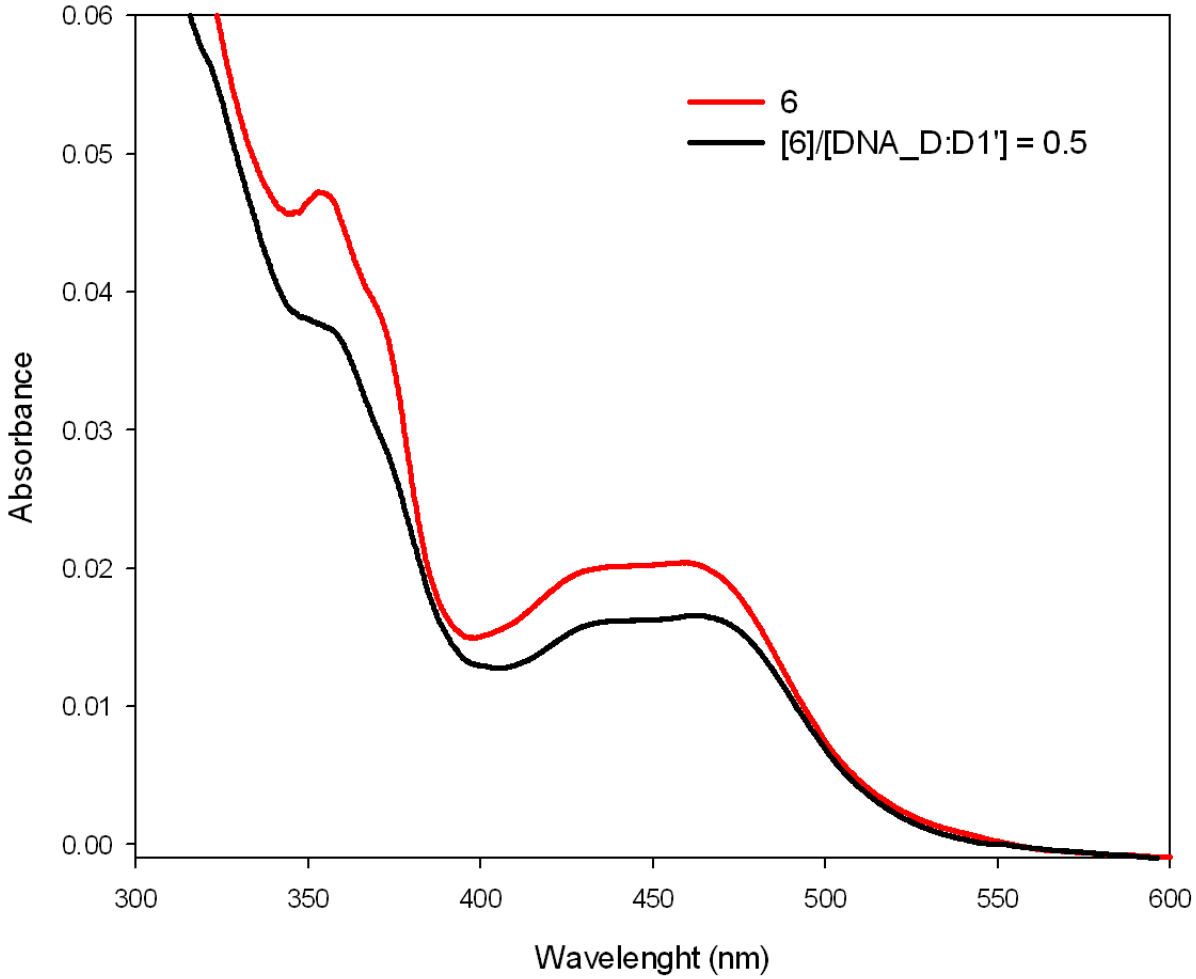


*Figure S80. UV-vis absorption spectra of metallopeptide **6** (3.4 μM) before (red line) and after (black line) the addition of a solution of TT mismatched D:D1' oligonucleotide.*

Table S4. Bathochromic (BS, in nm) and hypsochromic (HS, in %) shifts of the MLCT absorption band of the ruthenium metallopeptides 3-6 after the addition of B-DNA oligonucleotides in a ratio [Ru]/[DNA] = 0.5. ND = Not done

	Metallopeptides							
	3		***4***		***5***		***6***	
	BS *from 443 nm*	***HS*** *at 443 nm*	***BS*** *from 443 nm*	***HS*** *at 443 nm*	***BS*** *from 457 nm*	***HS*** *at 457 nm*	***BS*** *from 457 nm*	***HS*** *at 457 nm*
AAAAT TT	*8.2*	*37.52*	*2.0*	*19.03*	*ND*	*ND*	*ND*	*ND*
A:A'	*6.0*	*42.67*	*6.2*	*20.73*	*5.8*	*16.17*	*7.6*	*12.62*
A:A1'	*5.4*	*12.55*	*4.8*	*13.81*	*4.8*	*12.22*	*4.2*	*16.08*
A:A2'	*4.2*	*6.26*	*4.8*	*13.80*	*6.4*	*7.87*	*6.6*	*7.24*
A:A3'	*2.4*	*5.61*	*0.4*	*12.78*	*7.4*	*12.72*	*4.6*	*20.24*
B:B1'	*10.0*	*16.11*	*7.8*	*22.54*	*6.6*	*9.29*	*6.6*	*11.38*
C:C1'	*9.8*	*10.41*	*9.4*	*8.54*	*3.0*	*5.65*	*3.8*	*3.56*
D:D1'	*6.6*	*10.89*	*6.0*	*16.37*	*4.4*	*7.95*	*6.0*	*18.66*

CD studies

To a solution of the selected B-DNA oligonucleotide in Tris-HCl buffer (20 mM), pH 7.5 and NaCl (100 mM), an aliquot of a Ru(II) metallopeptide (3-6) in water was added in order to reach a [Ru]/[DNA] ratio of 0.5. The CD spectrum of the selected B-DNA oligonucleotide was recorded before and after the addition of the corresponding metallopeptide. The studied oligonucleotides are listed in Table S1.

a) Hairpin "well-matched" B-DNA binding studies

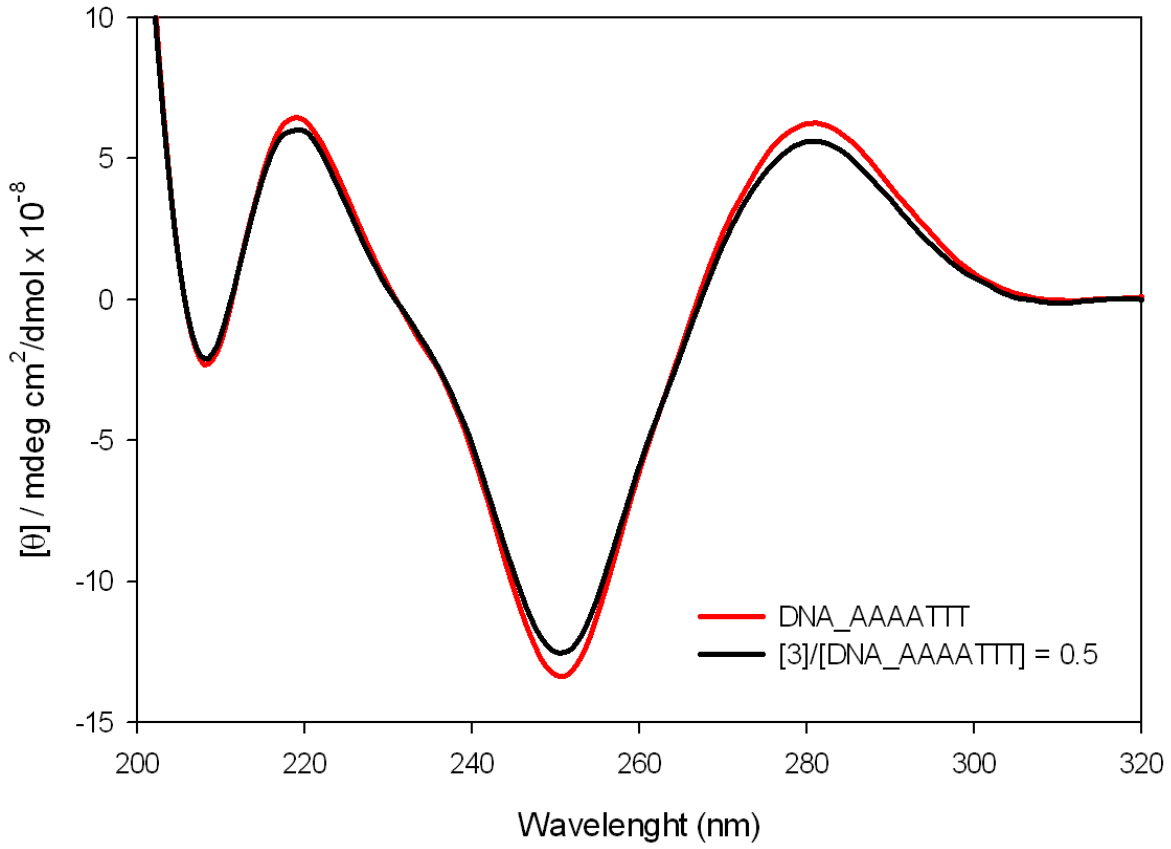


*Figure S81. CD spectra of AAAATTT oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide **3**.*

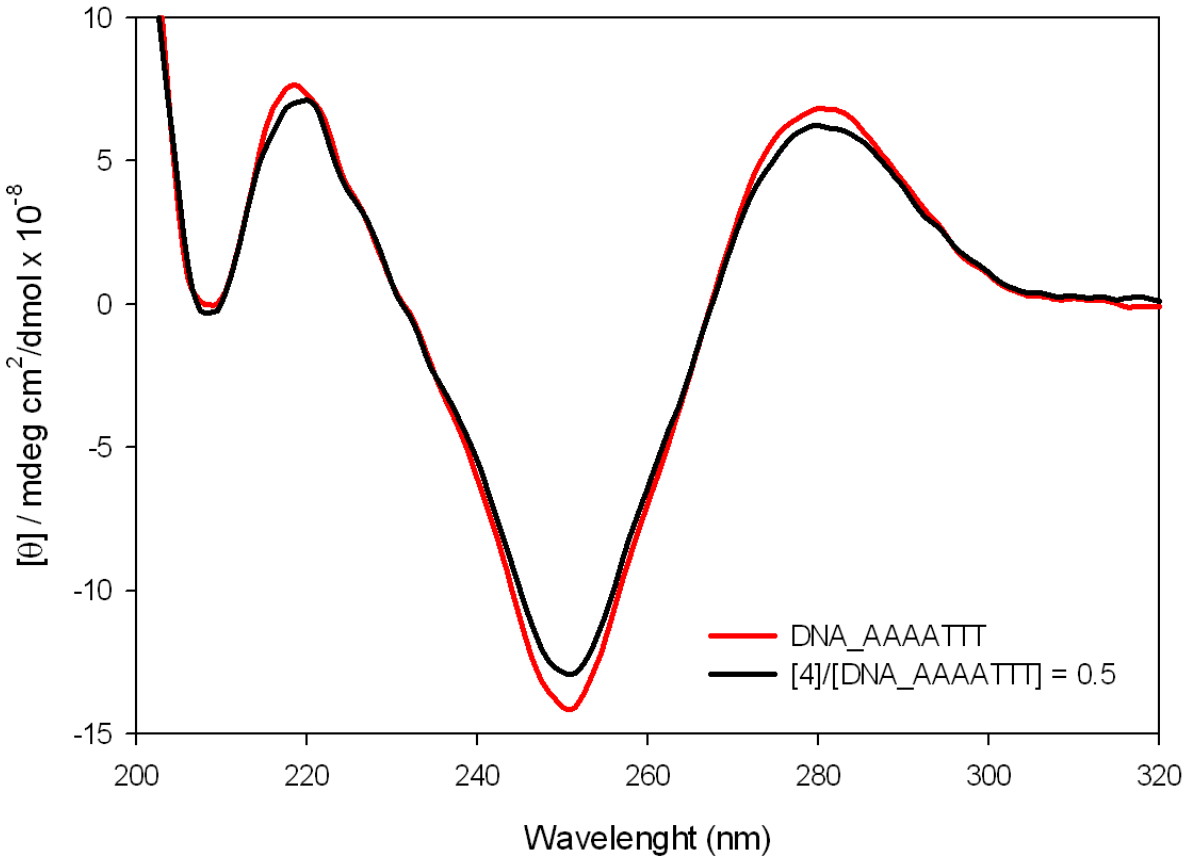


Figure S82. CD spectra of AAAATTT oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide ***4****.*

b) Double-stranded "well-matched" and "mismatched" B-DNA binding studies

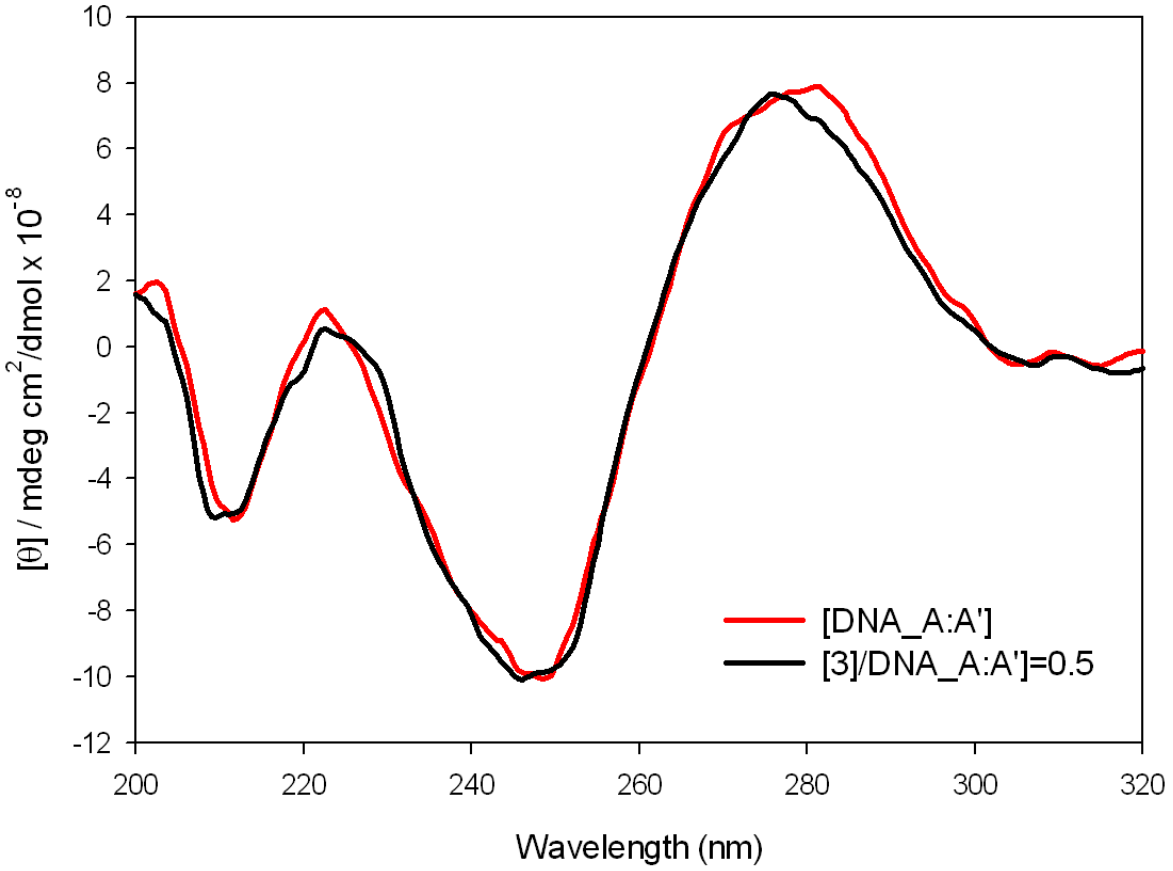


*Figure S83. CD spectra of "well matched" A:A' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide **3**.*

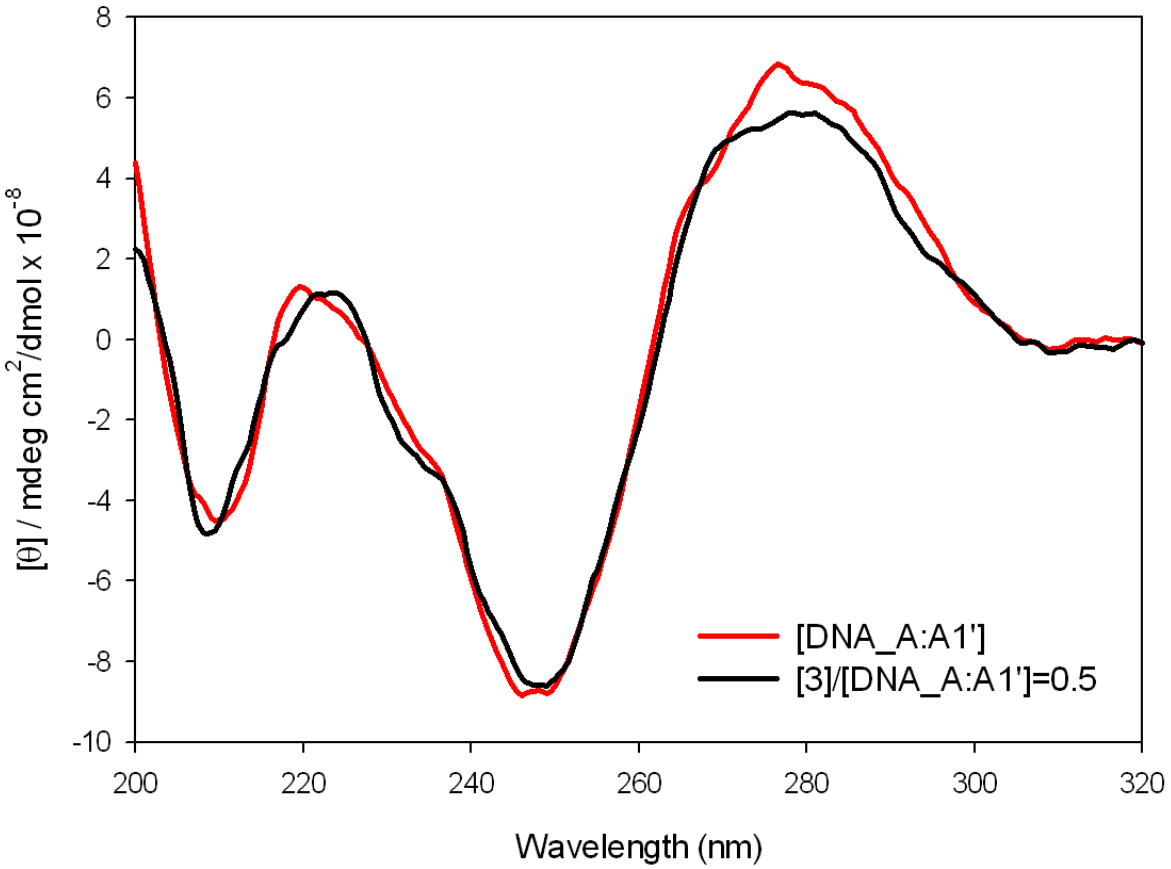


Figure S84. CD spectra of GG mismatched A:A1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide ***3****.*

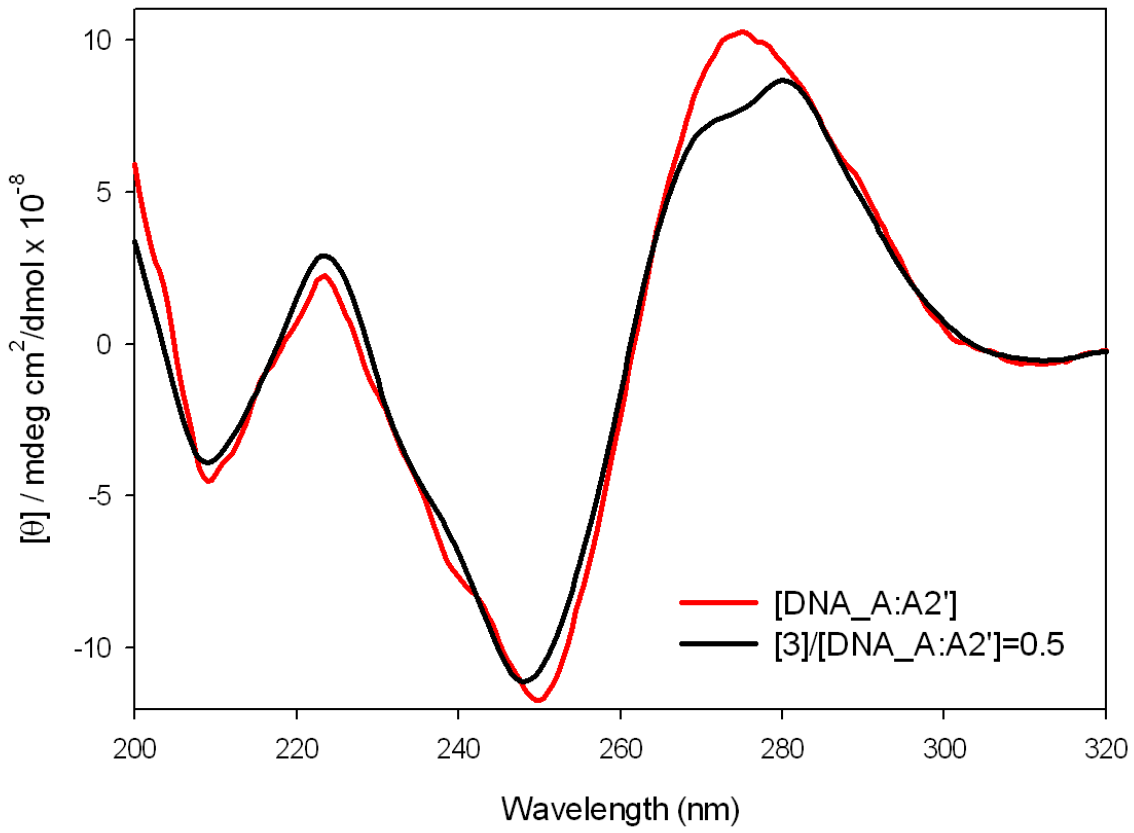


Figure S85. CD spectra of GG/AA mismatched A:A2' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide ***3****.*

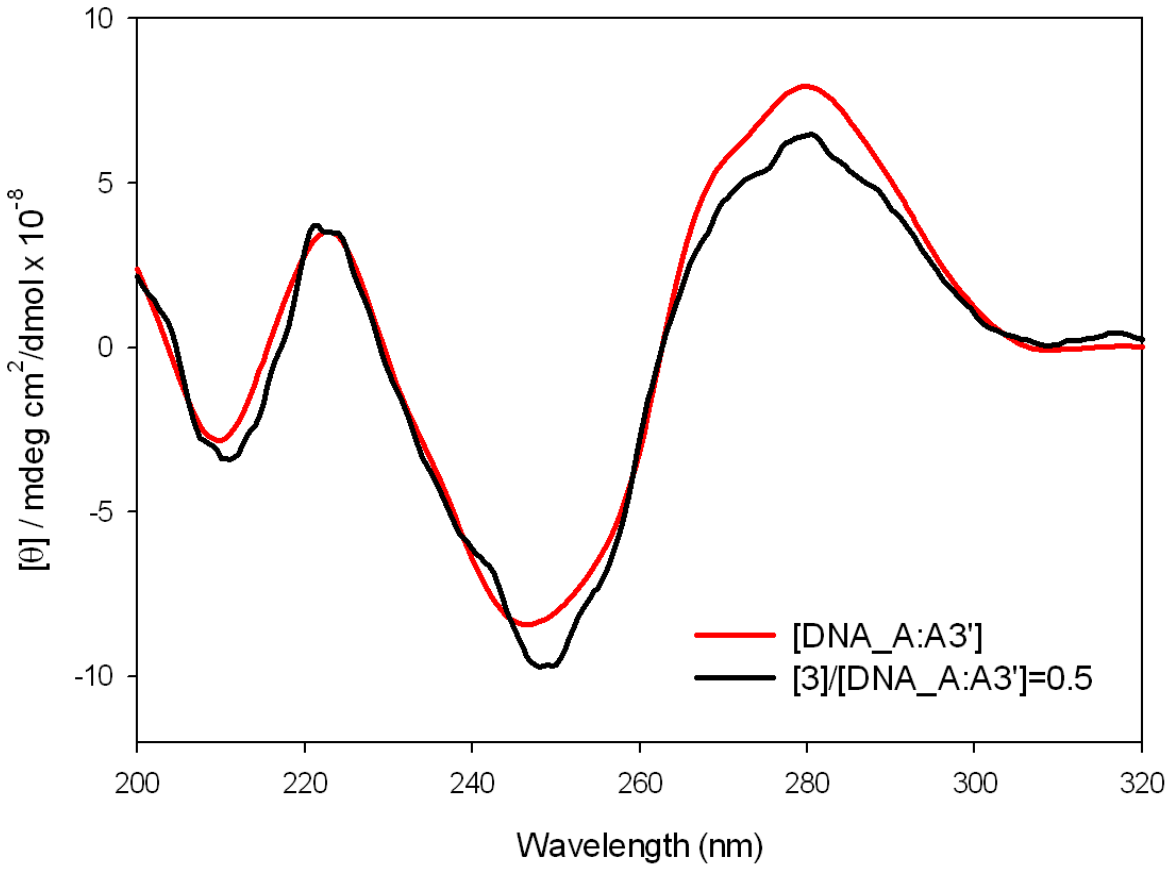


Figure S86. CD spectra of GG/GG/AA mismatched A:A3' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide ***3****.*

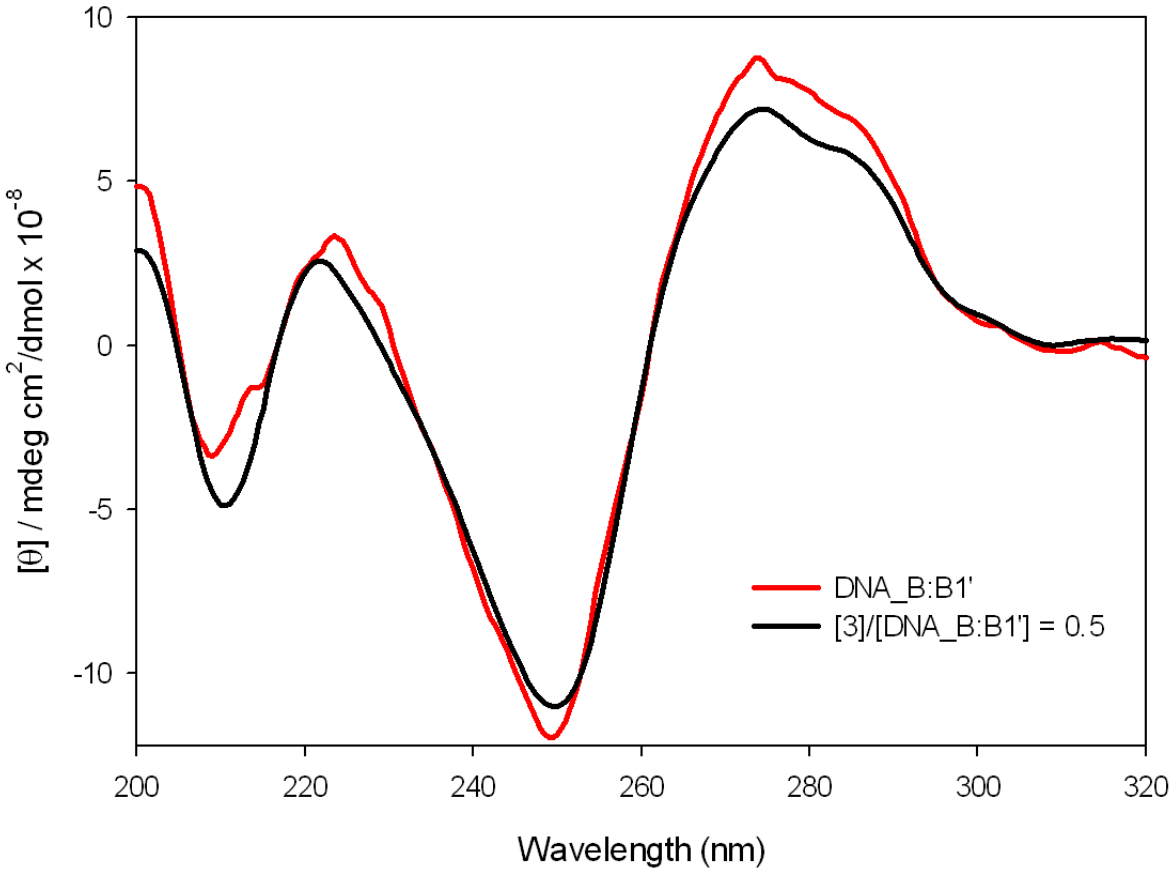


Figure S87. CD spectra of AA mismatched B:B1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide ***3****.*

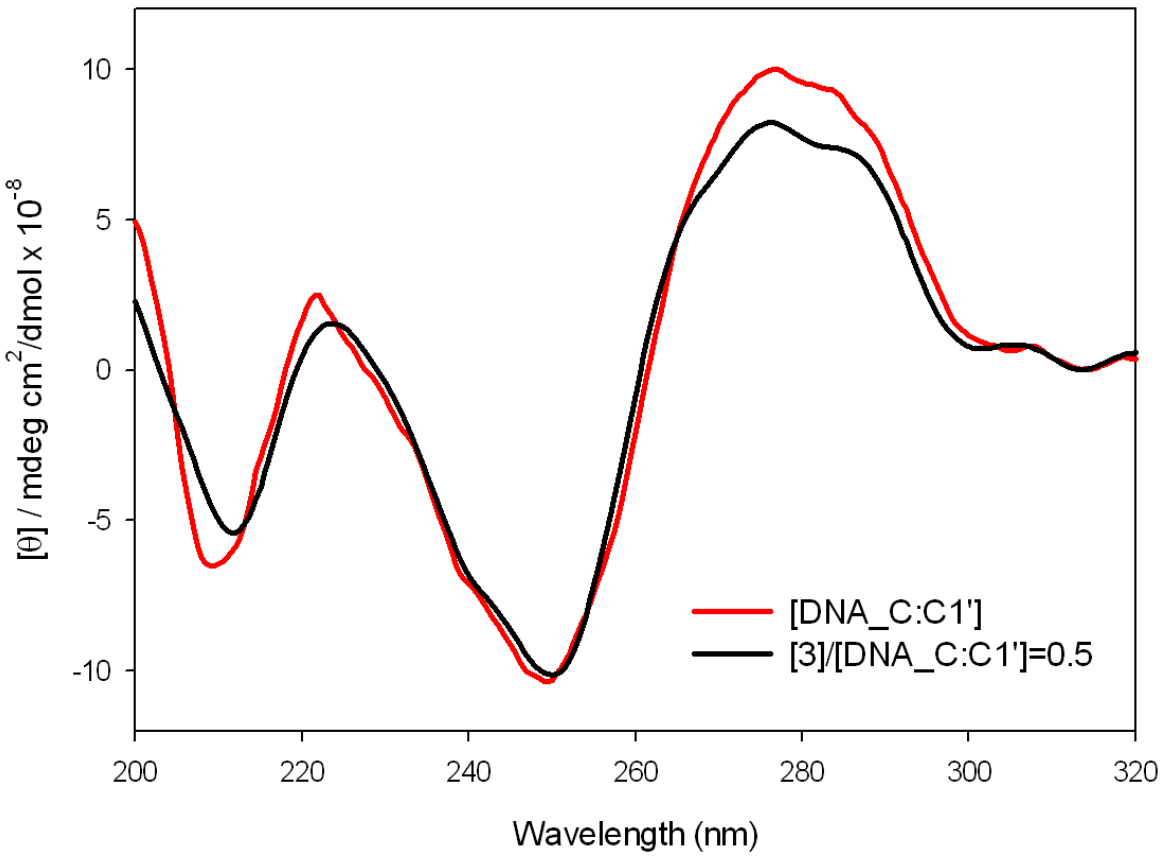


Figure S88. CD spectra of CC mismatched C:C1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide ***3****.*

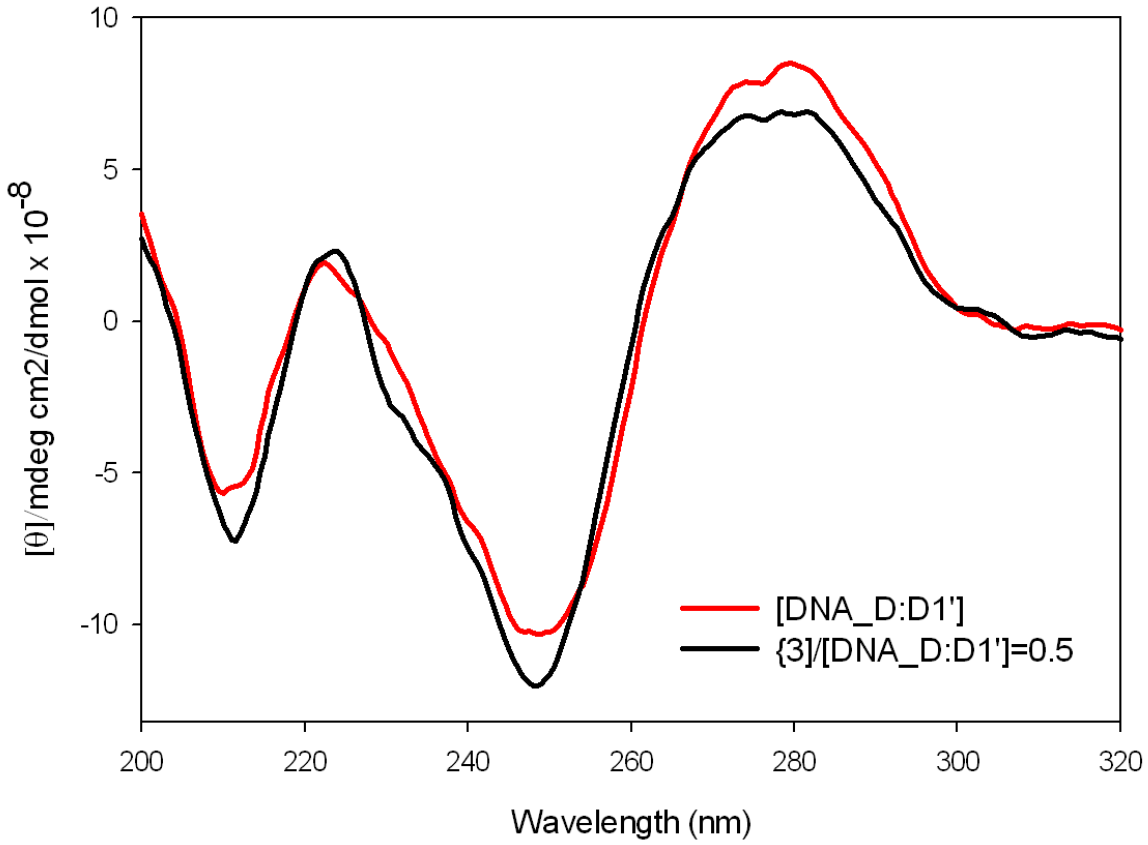


Figure S89. CD spectra of TT mismatched D:D1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide ***3****.*

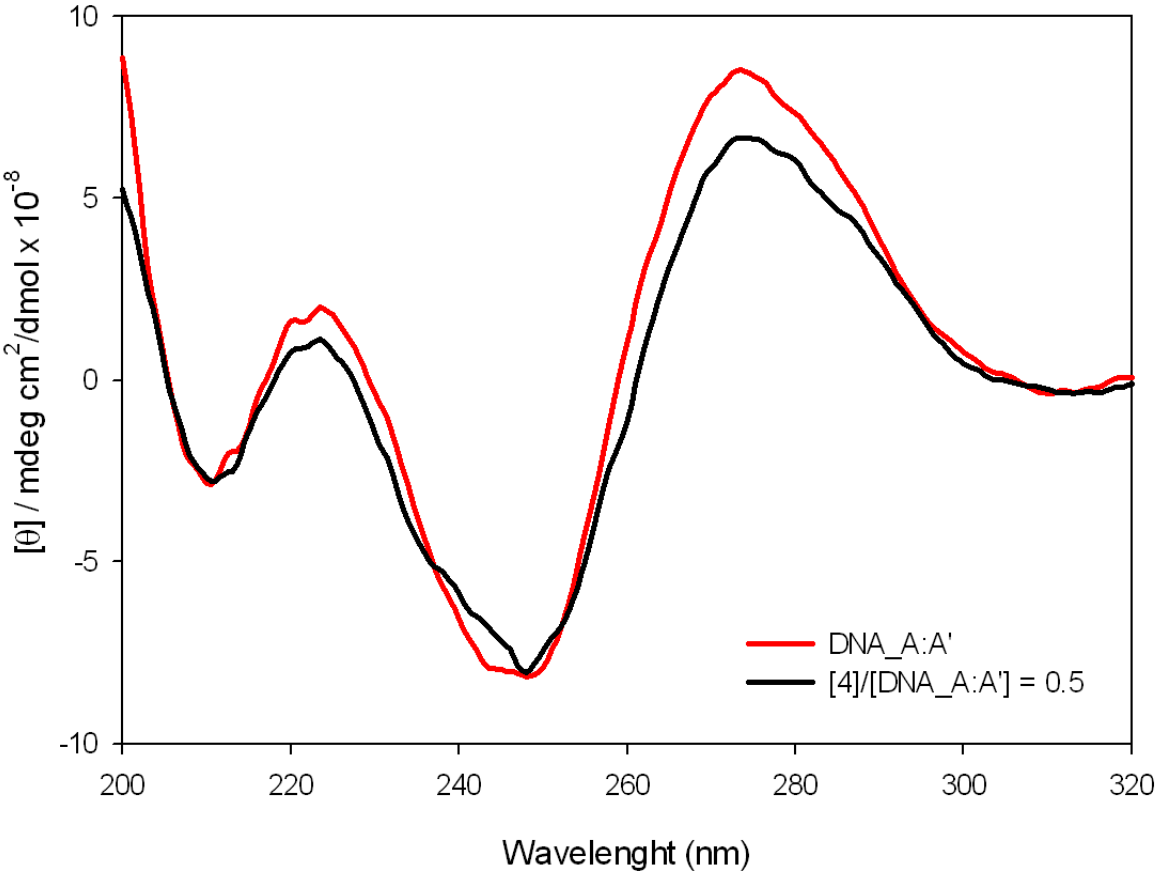


Figure S90. CD spectra of "well matched" A:A' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 4.

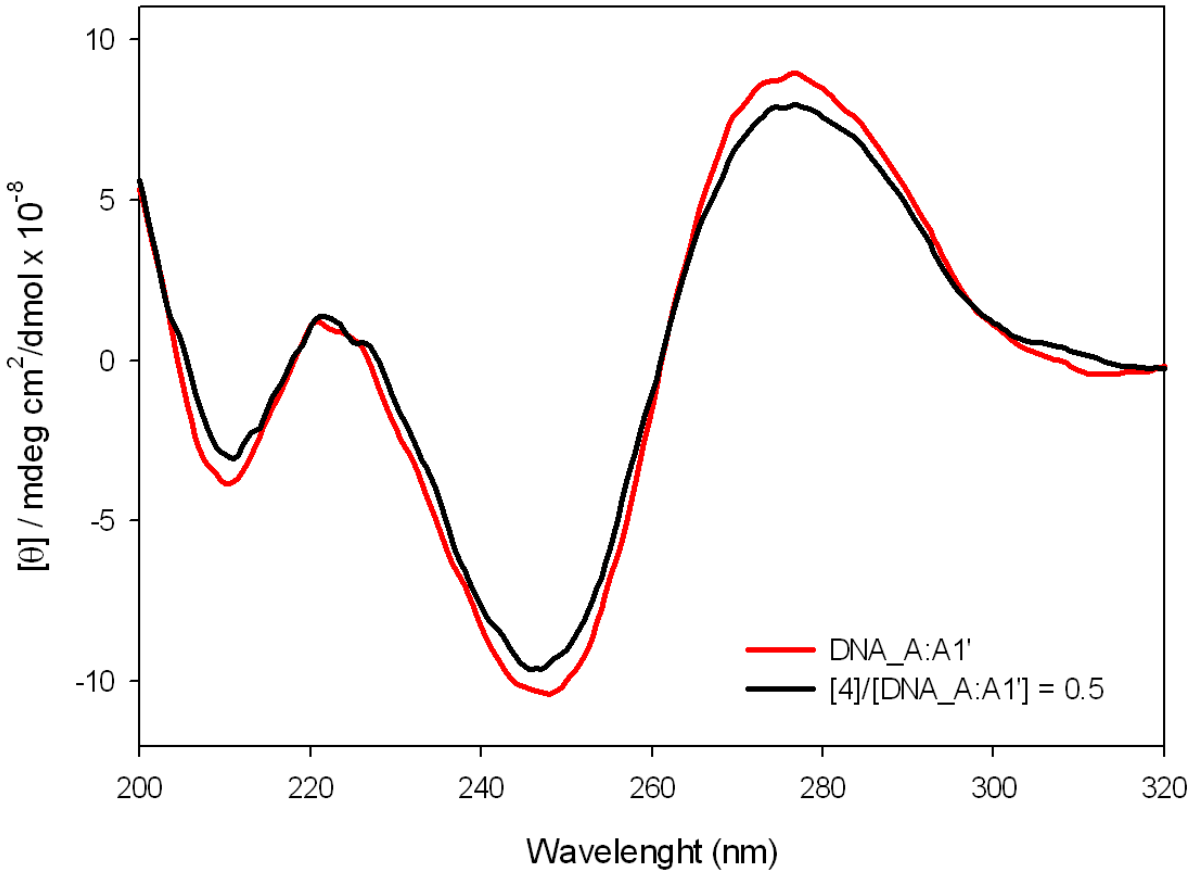


Figure S91. CD spectra of GG mismatched A:A1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 4.

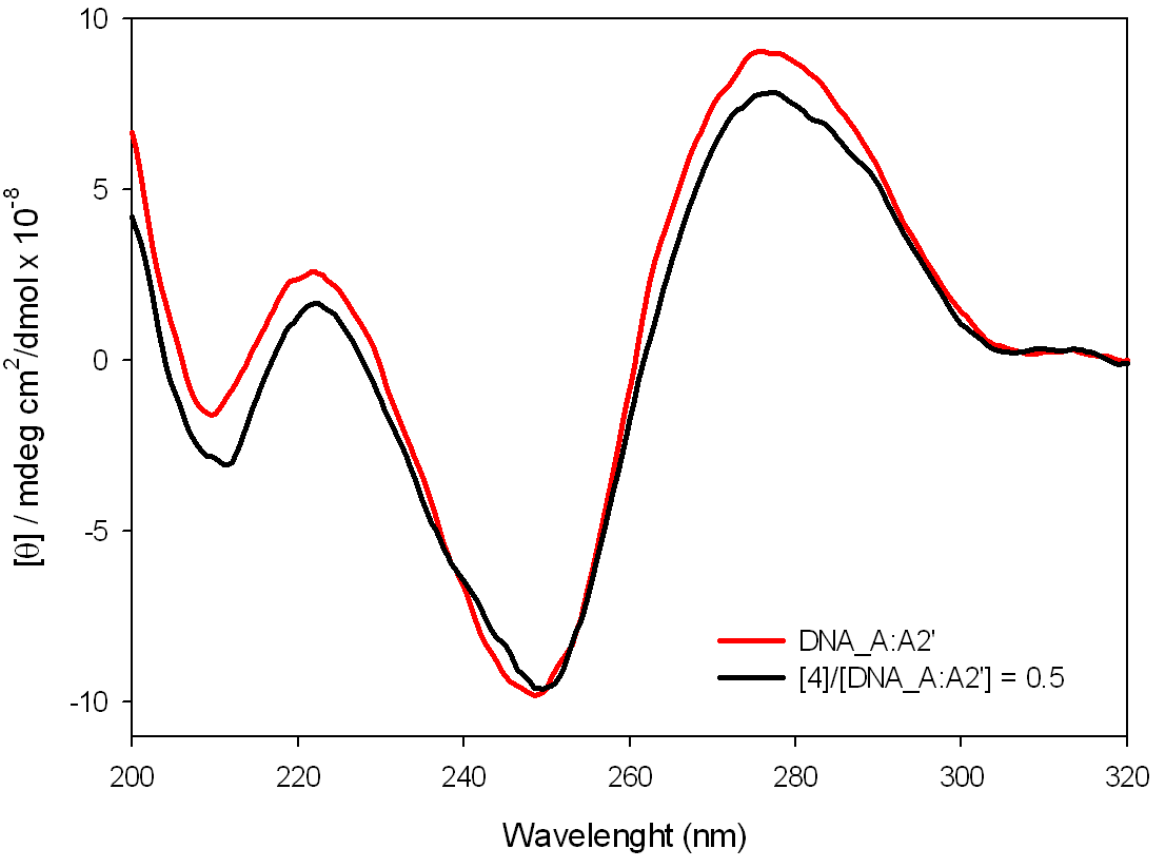


Figure S92. CD spectra of GG/AA mismatched A:A2' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 4.

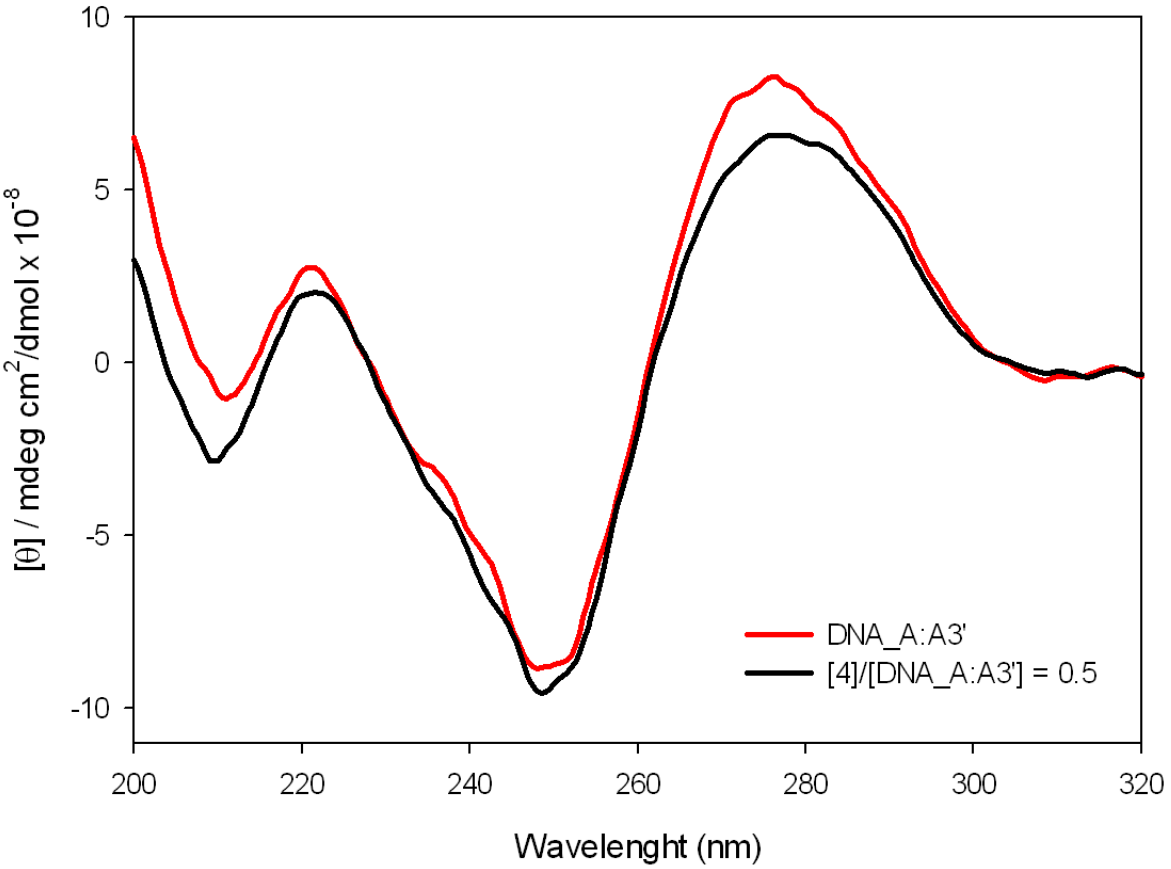


Figure S93. CD spectra of GG/GG/AA mismatched A:A3' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 4.

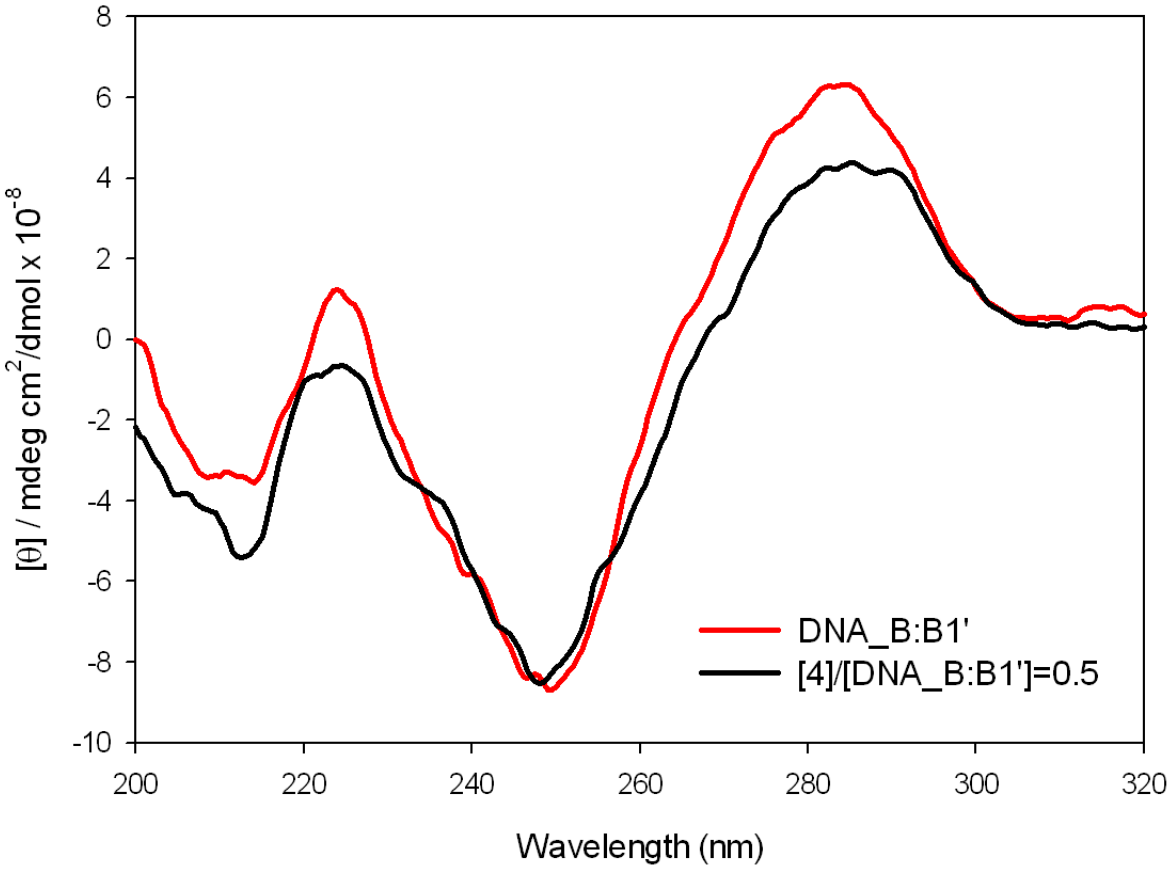


Figure S94. CD spectra of AA mismatched B:B1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 4.

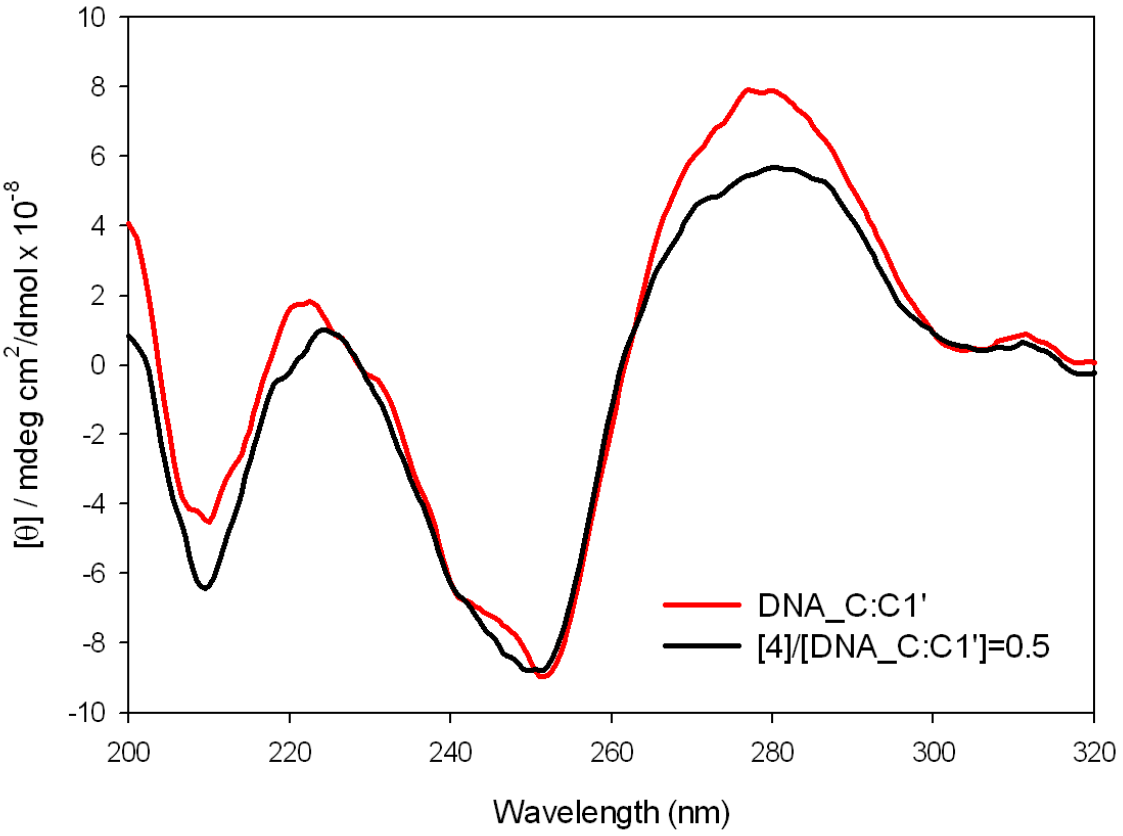


Figure S95. CD spectra of CC mismatched C:C1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 4.

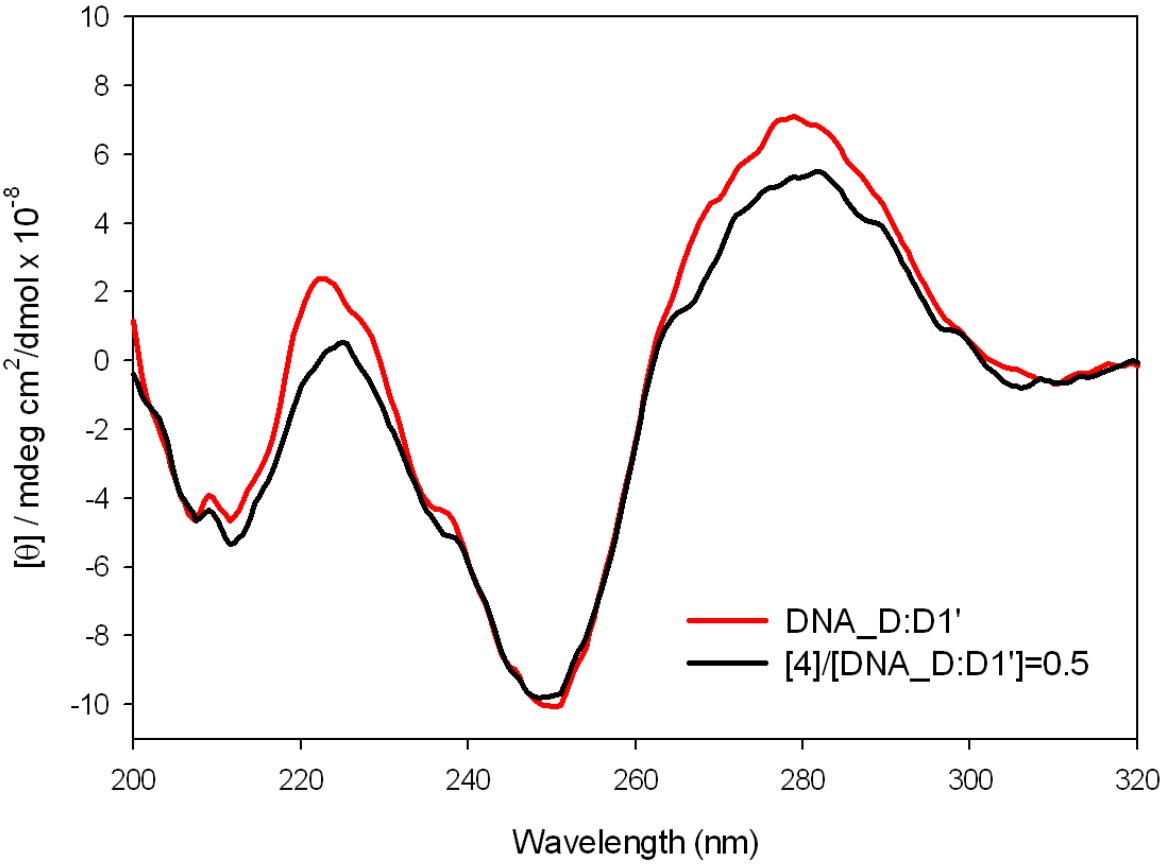


Figure S96. CD spectra of TT mismatched D:D1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 4.

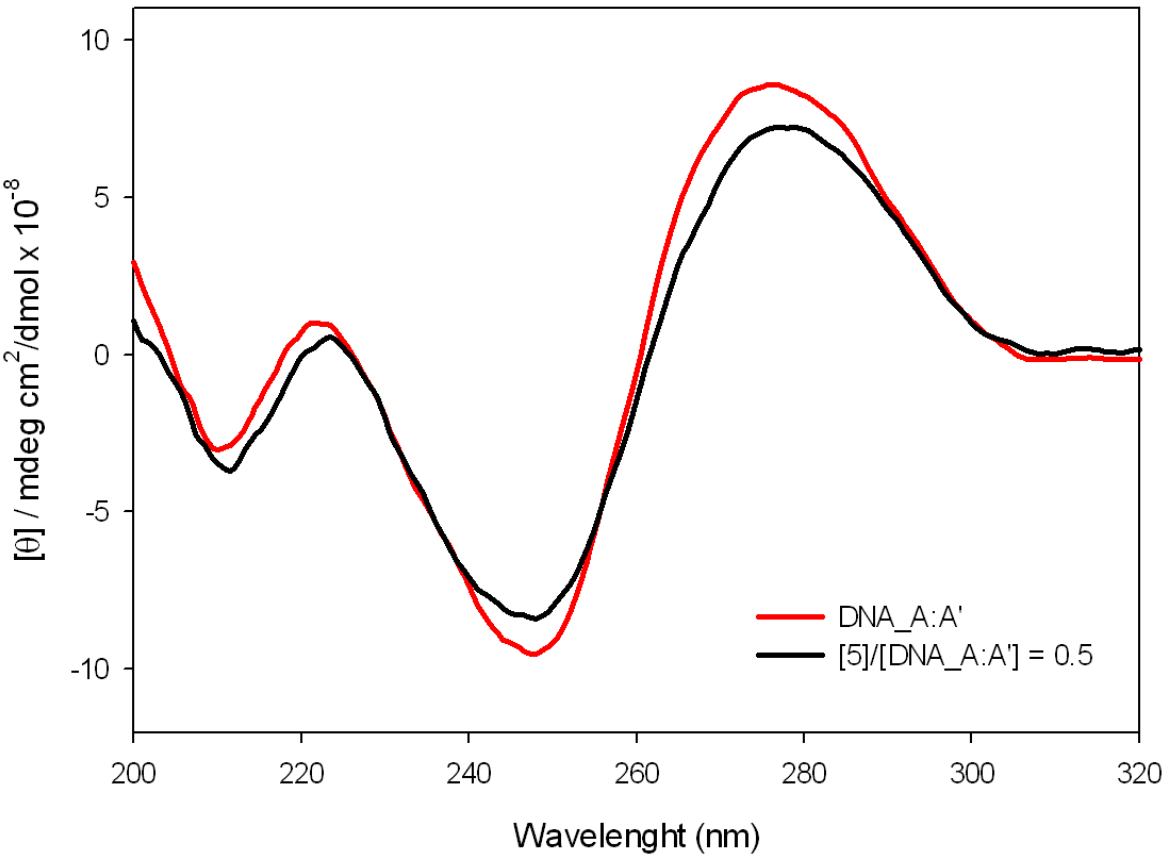


Figure S97. CD spectra of "well matched" A:A' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 5.

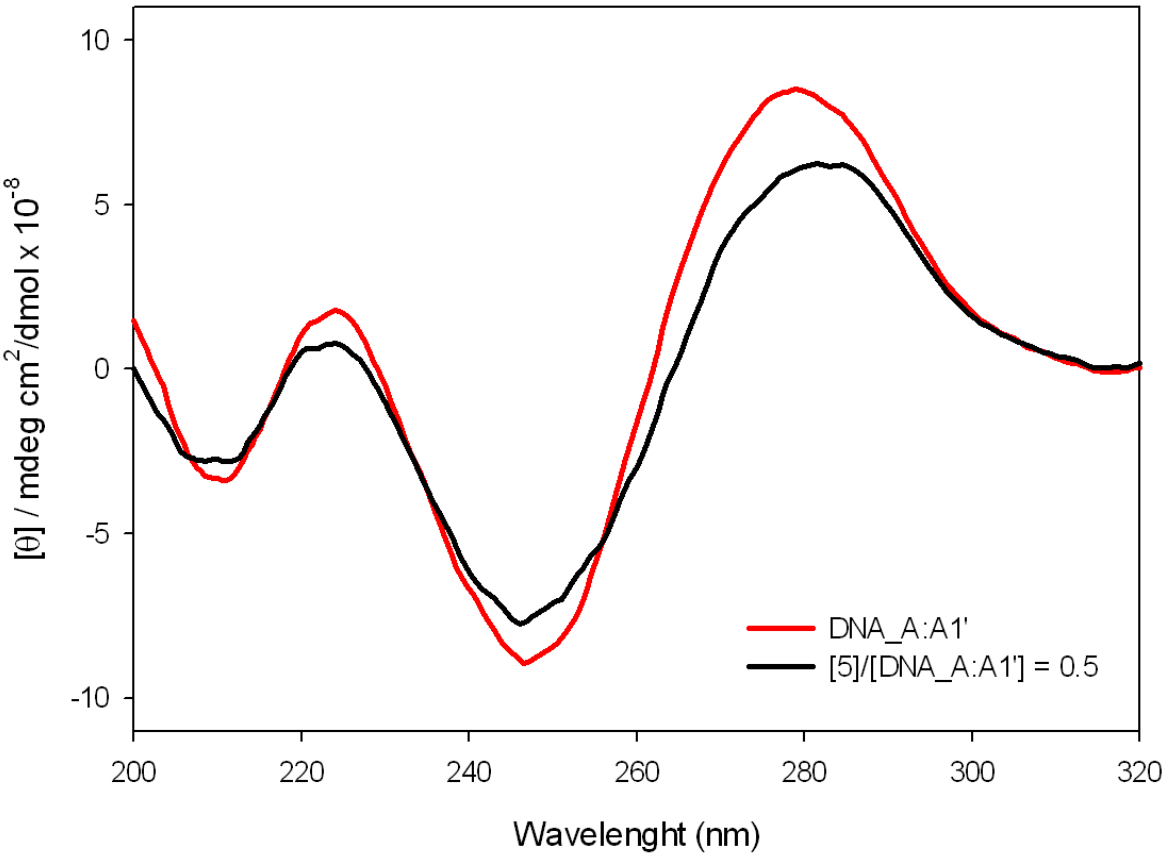


Figure S98. CD spectra of GG mismatched A:A1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 5.

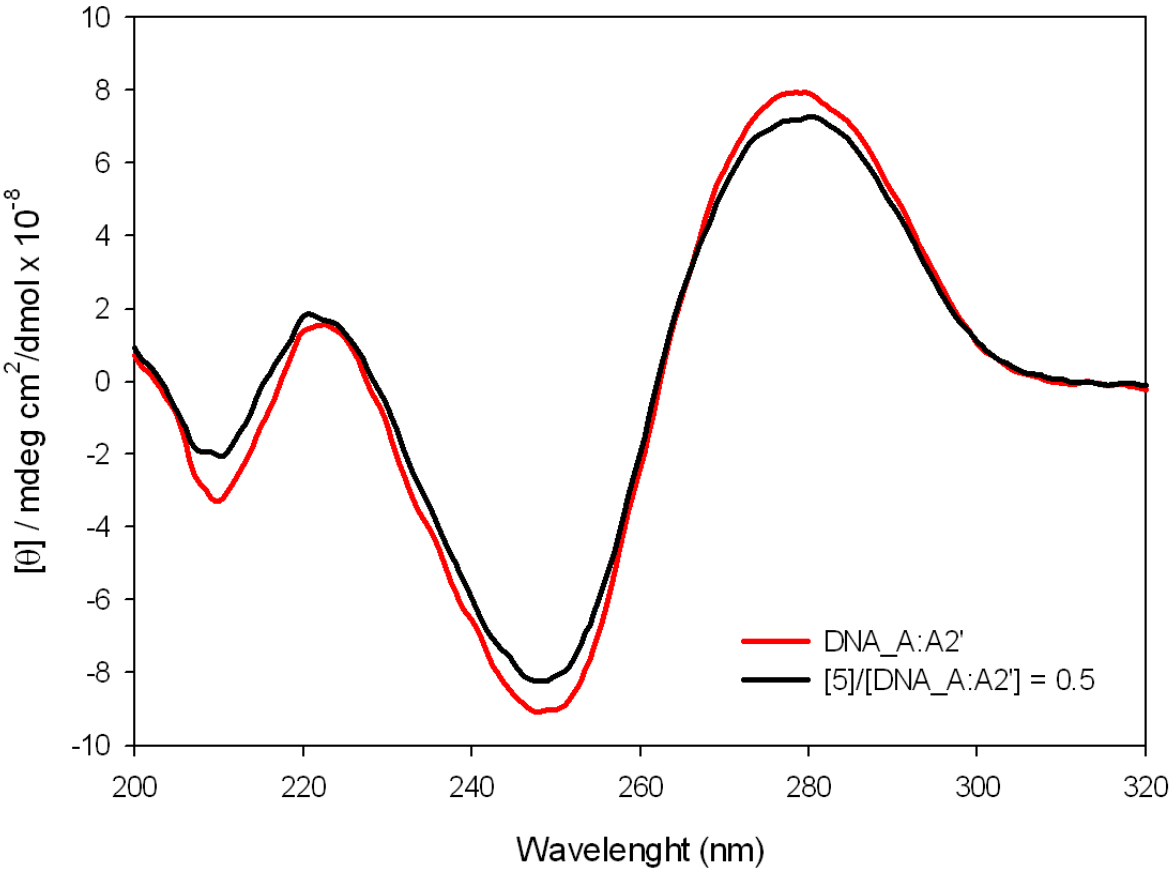


Figure S99. CD spectra of GG/AA mismatched A:A2' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 5.

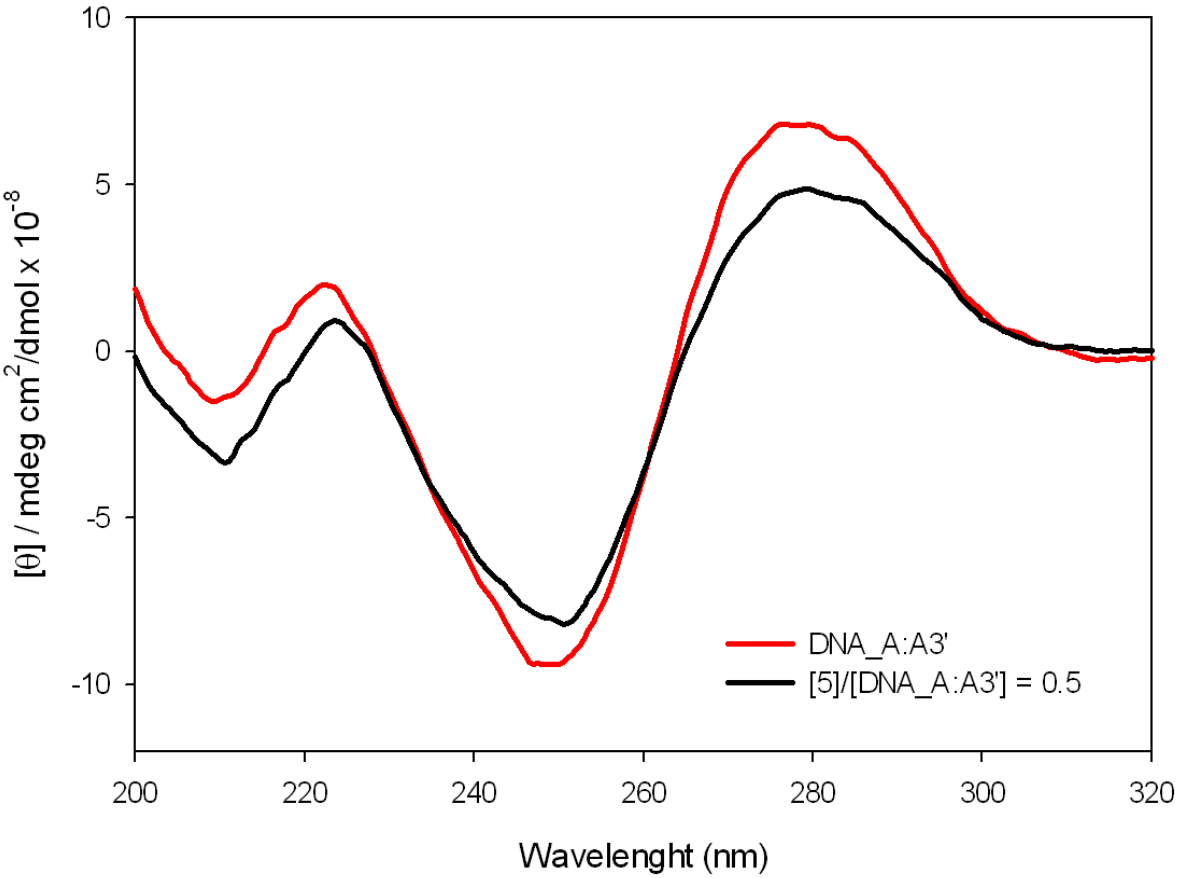


Figure S100. CD spectra of GG/GG/AA mismatched A:A3' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 5.

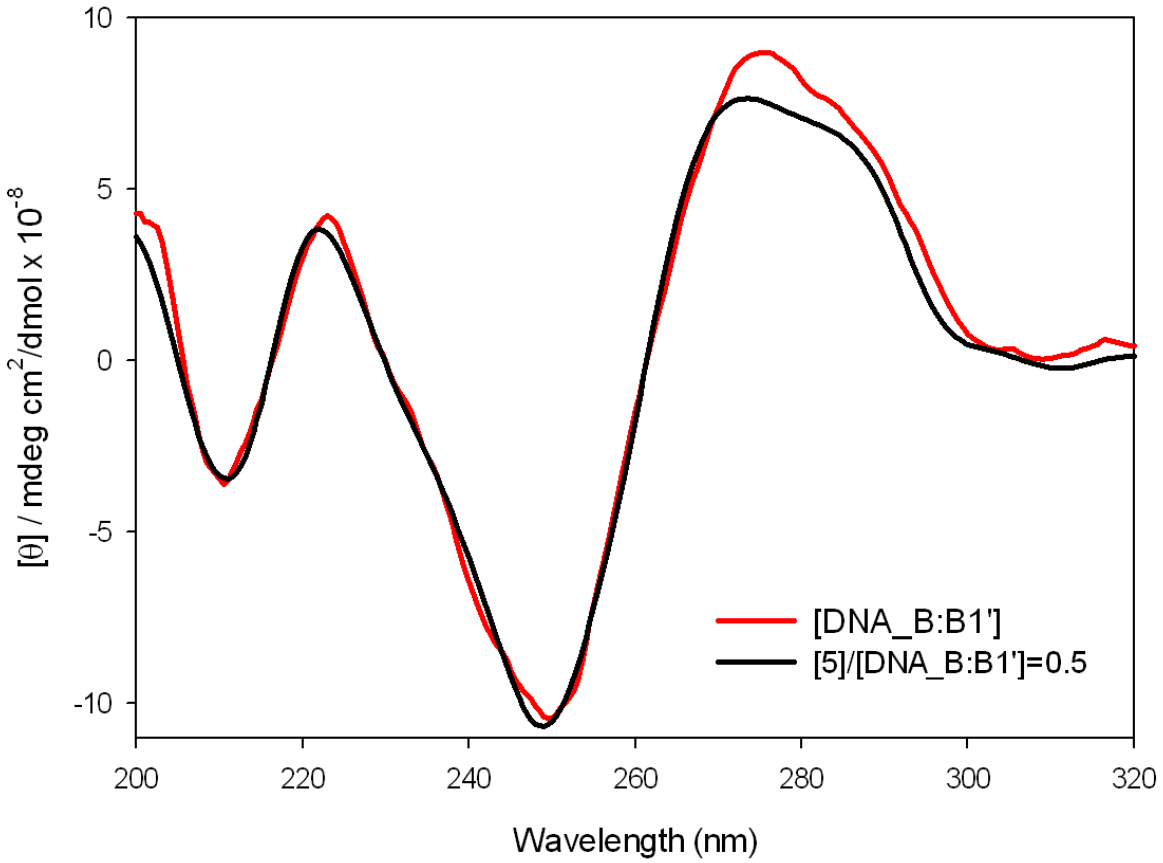


Figure S101. CD spectra of AA mismatched B:B1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 5.

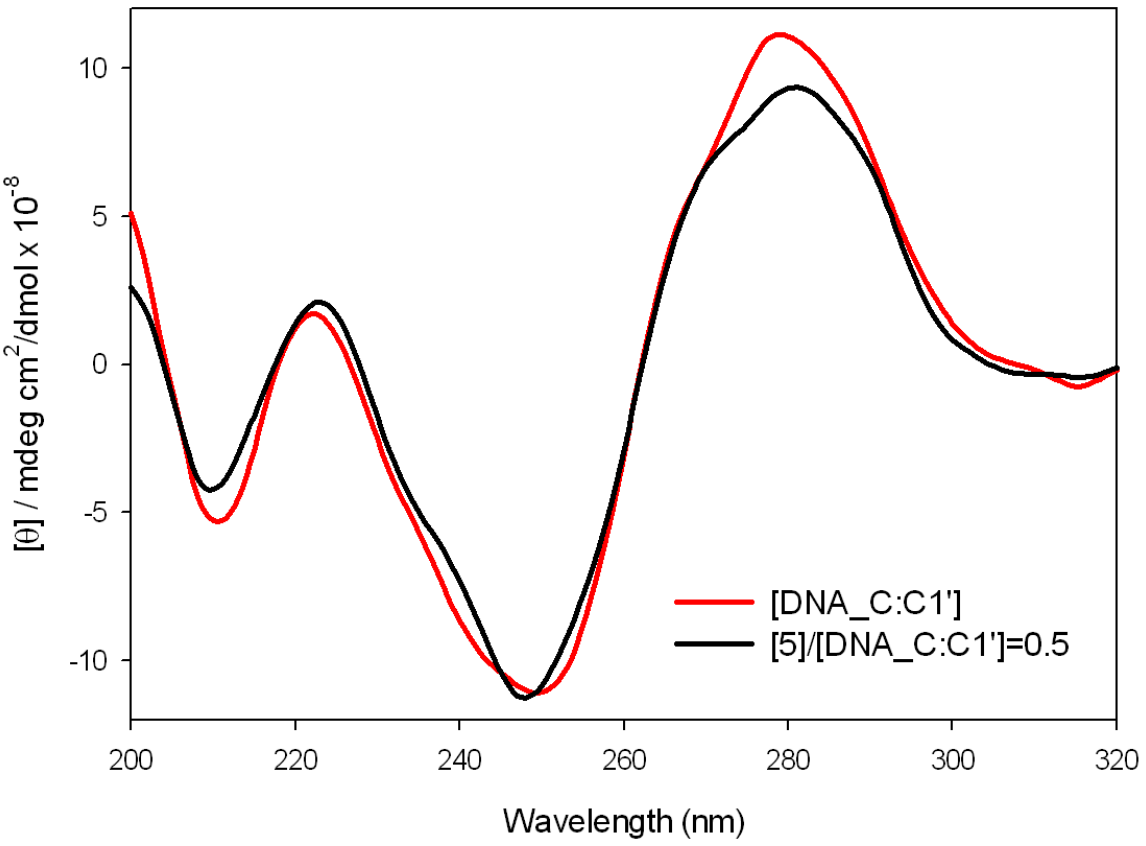


Figure S102. CD spectra of CC mismatched C:C1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 5.

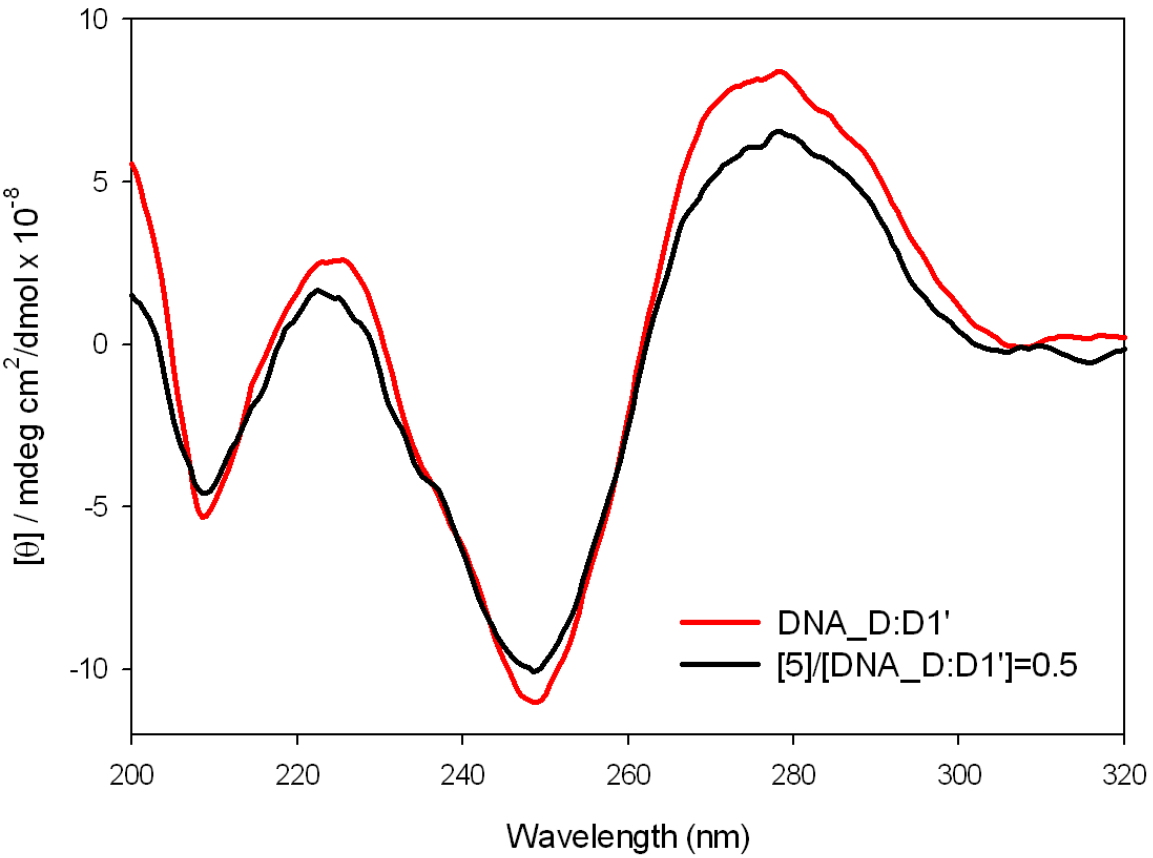


Figure S103. CD spectra of TT mismatched D:D1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 5.

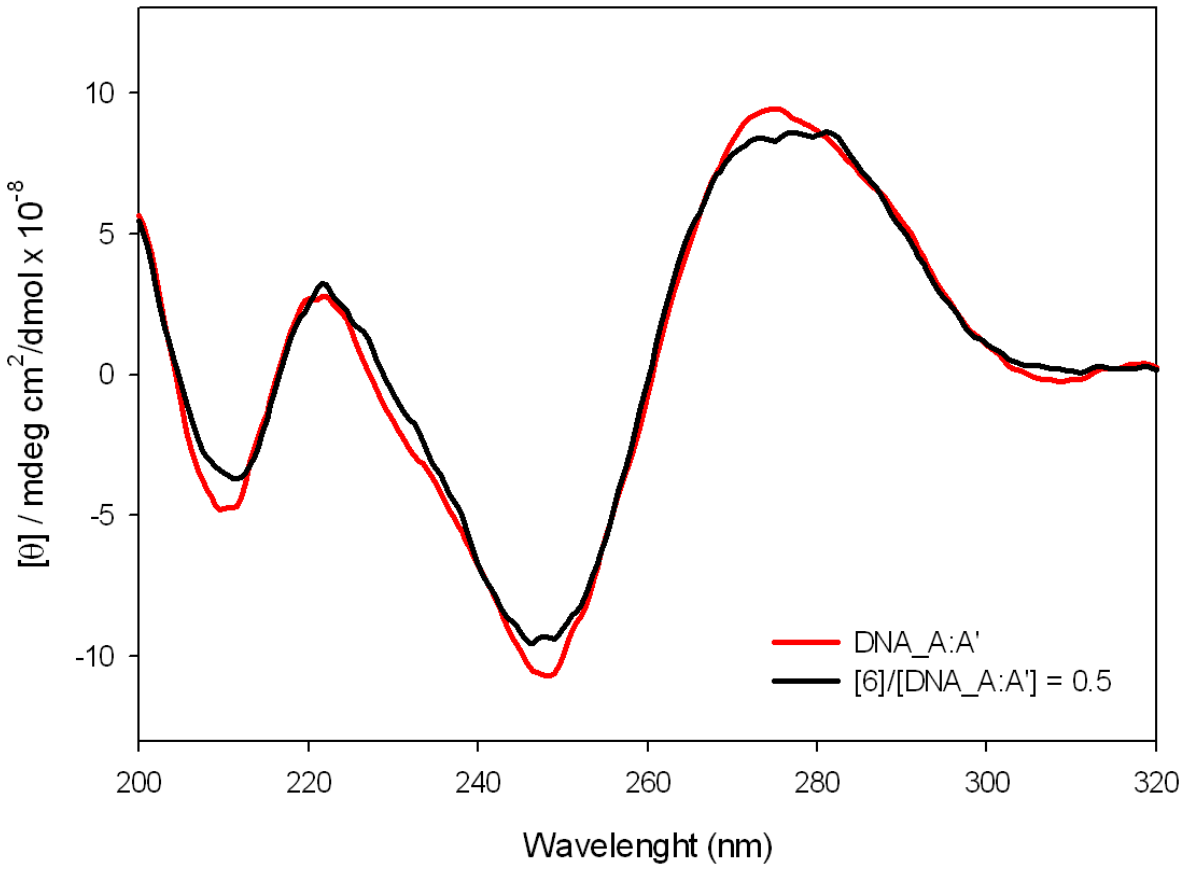


Figure S104. CD spectra of "well matched" A:A' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 6.

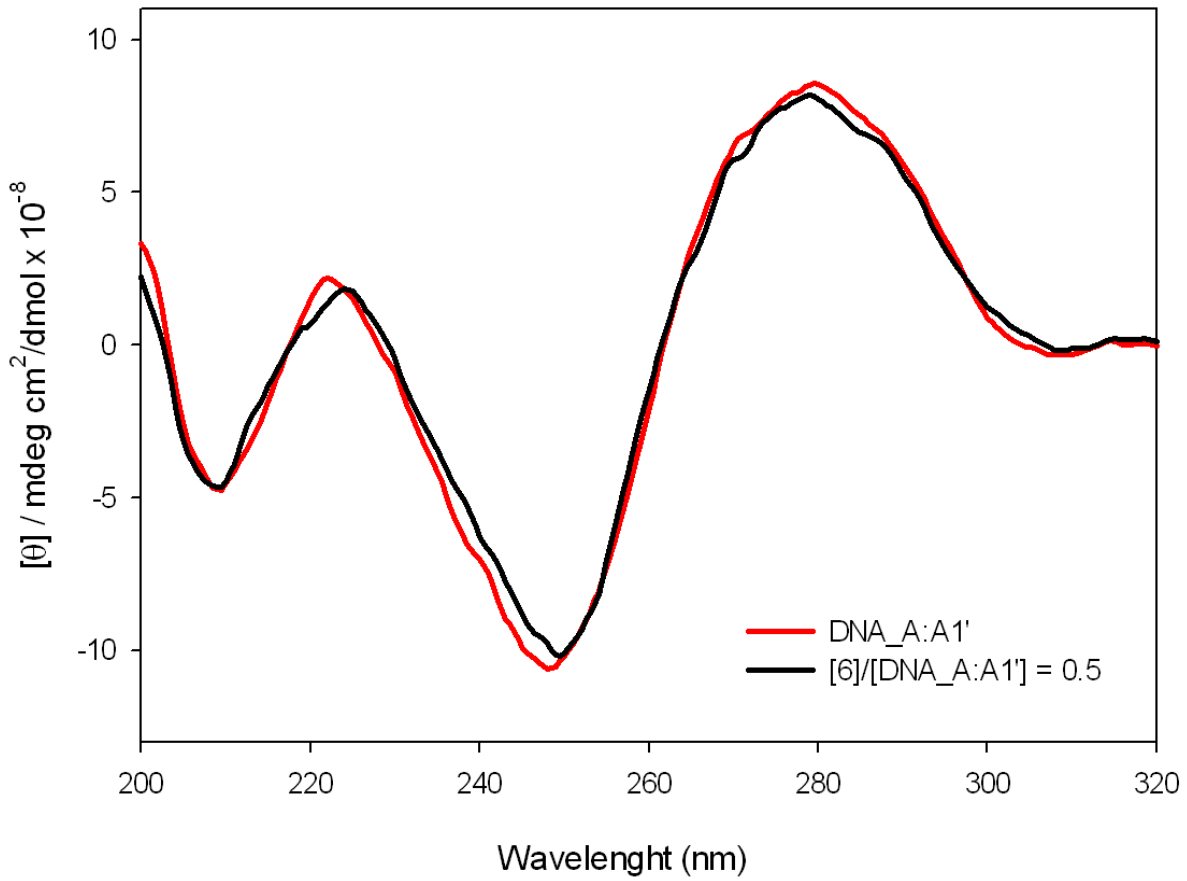


Figure S105. CD spectra of GG mismatched A:A1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 6.

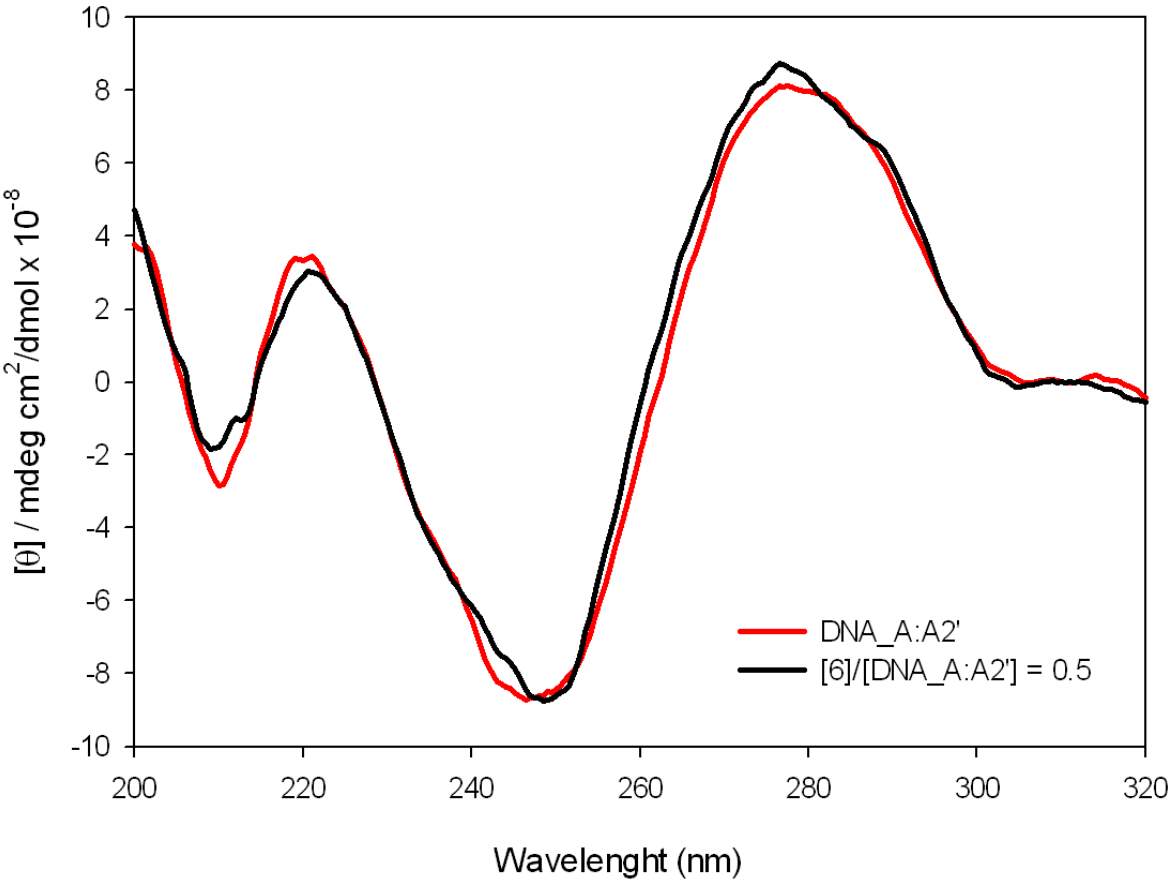


Figure S106. CD spectra of GG/AA mismatched A:A2' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 6.

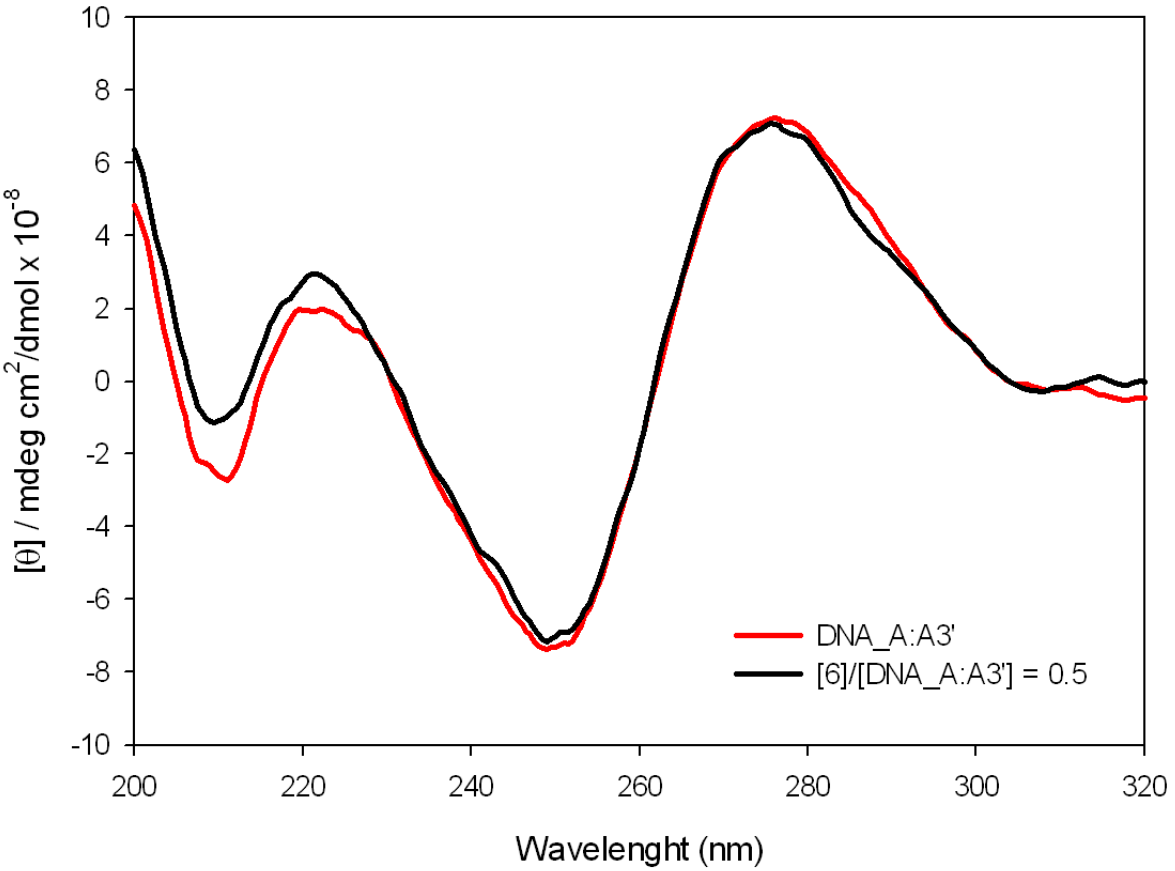


Figure S107. CD spectra of GG/GG/AA mismatched A:A3' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 6.

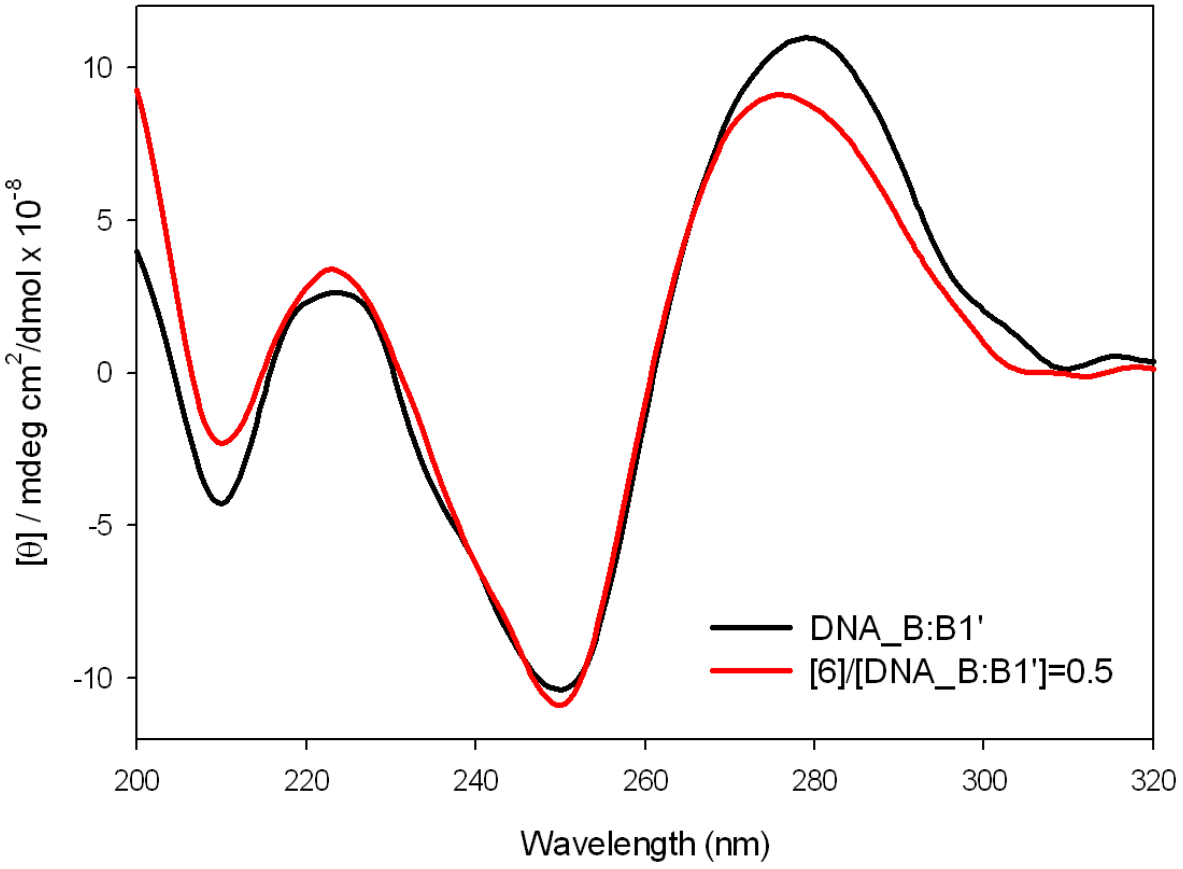


Figure S108. CD spectra of AA mismatched B:B1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 6.

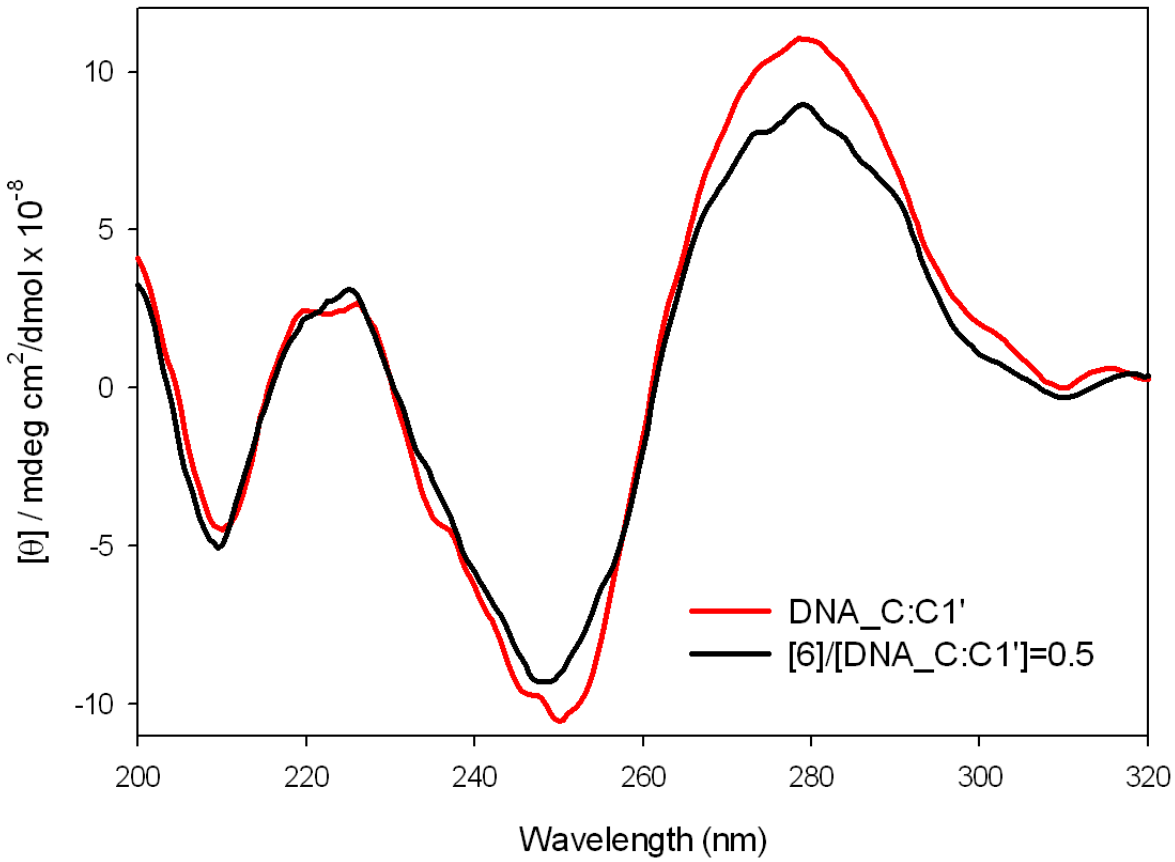


Figure S109. CD spectra of CC mismatched C:C1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 6.

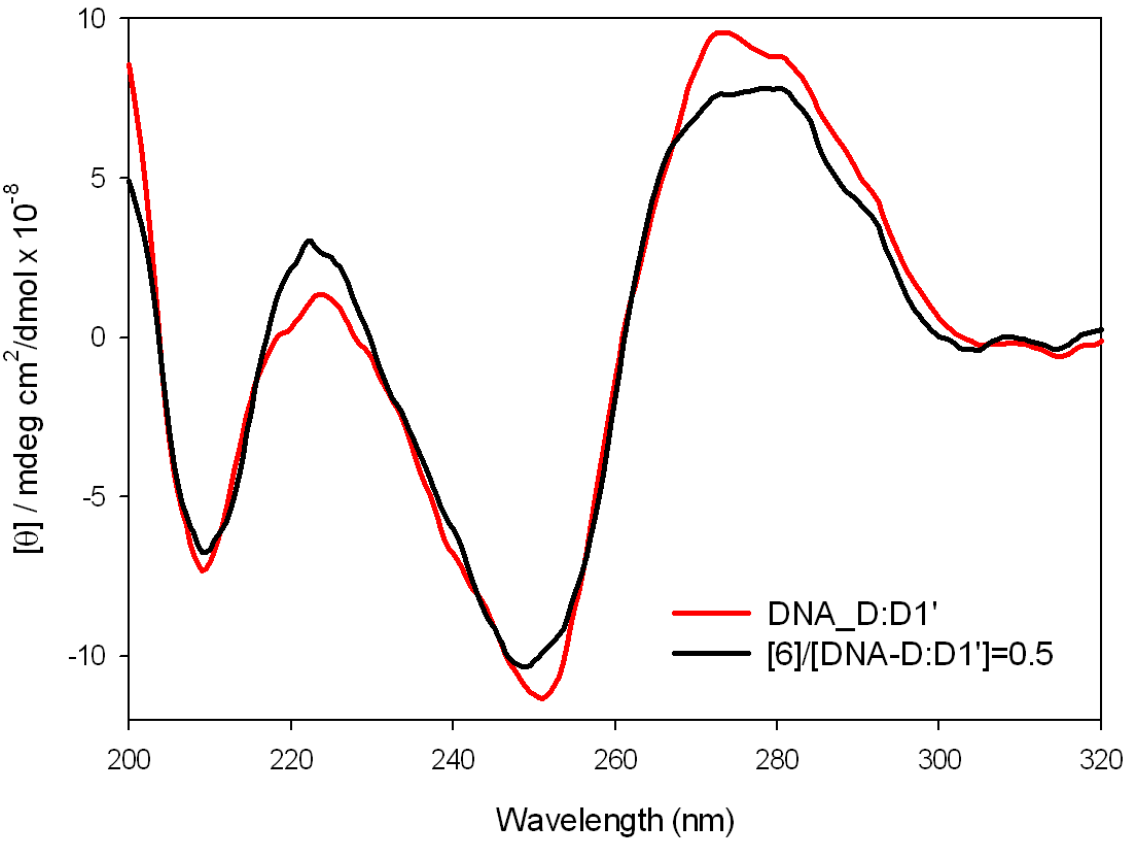


Figure S110. CD spectra of TT mismatched D:D1' oligonucleotide (10.0 μM) before (red line) and after (black line) the addition of a solution of metallopeptide 6.

Fluorescent microscopy studies

Semi-confluent monolayers of Vero cells were grown on glass bottom 35 mm dishes in Dulbeccos modified Eagle Medium (DMEM) containing 10% of FBS (fetal bovine serum). For the assays, the cells were washed three times with DMEM containing no FBS or antibiotics, and incubated with 5 μM of the compound 4 in DMEM without FBS during 24 hr, at 37 °C in a 5% CO_2 humidified atmosphere. Then, the cells were gently washed 5 times with phosphate buffered saline (PBS), and observed under the fluorescence microscope in DMEM without fixation. Alternatively, the cells were incubated for 30 minutes with 25 μM of the compounds 3 and 4 in DMEM without FBS during 30 minutes, at 37 °C in a 5% CO_2 humidified atmosphere. After that, the cells were washed 5 times with phosphate buffered saline (PBS), fresh DMEM was added and the cells were continuously monitored for four hours under the fluorescence microscope in DMEM without fixation. All images were obtained with a Zyla 4.2 camera (Andor) mounted on a NIKON Ti E inverted microscope equipped with incubation chamber. Images were manipulated with NIS software (Nikon) and Adobe Photoshop.

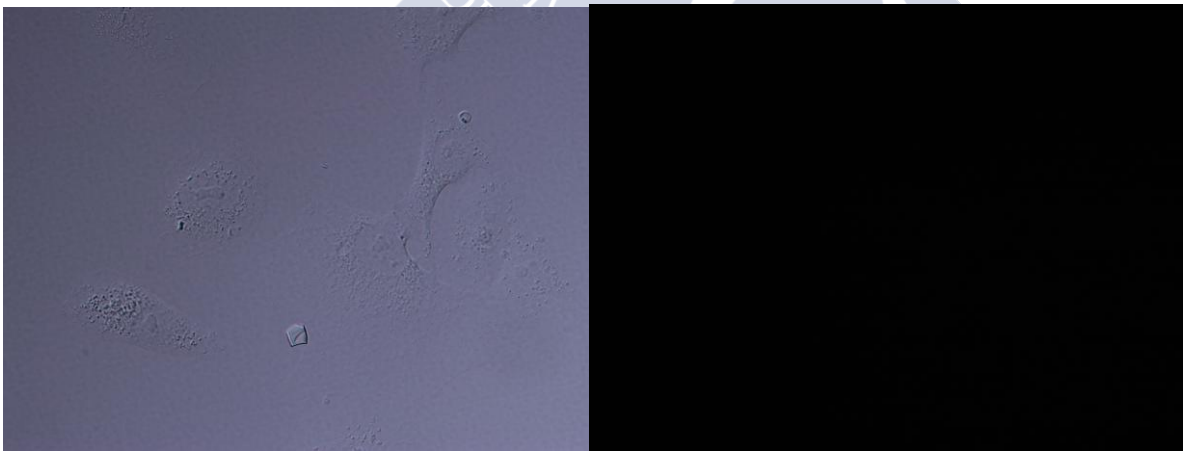


Figure S111. Monolayers of Vero cells were incubated with 25 μM of 3 in DMEM without serum for 30 minutes. After that, the medium was removed, the cells washed three times with PBS, and further incubated at 37ºC in DMEM without serum. Brightfield (left) and red-emission fluorescence pictures (right) were taken from unfixed living cells at 1000X magnification.