Tetrahedron 71 (2015) 7339–7343

Contents lists available at ScienceDirect

Tetrahedron

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

# Synthesis of 6-arylisocytosines and their potential for hydrogen bonding interactions

Alpa Patel [a], William Lewis [a], Mark S. Searle [a], Malcolm F.G. Stevens [b], Christopher J. Moody [a,*]

[a] School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, UK
[b] School of Pharmacy, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

ARTICLE INFO

*Article history:*
Received 3 February 2015
Received in revised form 10 April 2015
Accepted 23 April 2015
Available online 29 April 2015

Dedicated to the memory of Alan Katritzky, a towering figure in heterocyclic chemistry

*Keywords:*
Heterocyclic compounds
Pyrimidinones
Isocytosine
G-quadruplex

ABSTRACT

The synthesis of a number of 6-arylisocytosines, including linked bis-isocytosines, from the reaction of guanidine with β-ketoesters is described. The compounds were investigated for their ability to form hydrogen-bonded structural networks, and for their potential interactions with the telomeric quadruplex forming sequence AGGG(TTAGGG)$_3$.



## 1. Introduction

One of Alan Katritzky's pioneering physico-chemical studies was to exploit electronic absorption spectroscopy to determine tautomer preferences in amino- and hydroxy-heterocyclic compounds.[1–3] On many occasions at meetings of the Heterocyclic Group of The Royal Society of Chemistry, of which he was a co-founder, he objected—but always in a kindly manner—to an azin-one being described as a hydroxy-azine or an amino-azine being misrepresented as an imino-azine (Fig. 1). Such considerations of tautomeric structures were, of course, pivotal in the elucidation of the duplex hydrogen-bonded structure of DNA by Watson and Crick where guanine (G) residues form three hydrogen bonds to cytosine (C) bases and adenines (A) form two hydrogen bonds to thymines (T) on adjacent strands of DNA.[4] Arguably, guanine is the most interesting of the DNA bases. Its 5-membered ring is the well known imidazole, whilst the 6-membered component is the unusual isocytosine ring that is ideally disposed to form three hydrogen bonds to cytosine itself in the Watson-Crick C-G base pair (Fig. 1).

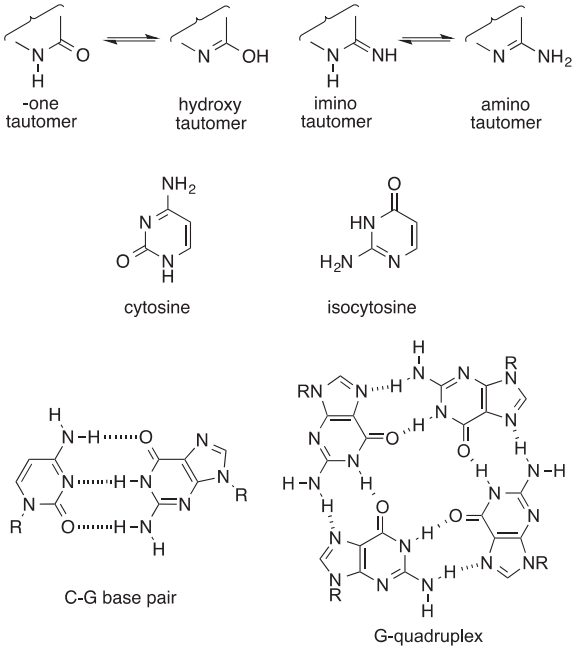


**Fig. 1.** Tautomerism in heterocyclic compounds and its role in hydrogen bonding.

* Corresponding author. Tel.: +44 115 846 8500; fax: +44 115 931 3564; e-mail address: c.j.moody@nottingham.ac.uk (C.J. Moody).

http://dx.doi.org/10.1016/j.tet.2015.04.084
0040-4020/

Guanine also features in the well-characterized series of repeats of the 5′-TTAGGG-3′ sequence, which occur in human DNA at the ends of chromosomes (telomeres) and are approximately 4–15 kilo base pairs in length.[5] These guanine-rich polynucleotides can form secondary structures known as G-quadruplexes by stacking square planes of G-tetrads, or G-quartets, that are assembled as a result of the unique hydrogen bonding properties of the isocytosine ring, in which the two nitrogen atoms of isocytosine are involved in hydrogen bonding along one face, while the oxygen of the carbonyl bond joins the imidazole N-3 atom in forming bonds along the perpendicular face in Hoogsteen pairing (Fig. 1). The whole structure is stabilized by the binding of a monovalent cation ($Na^+$ or $K^+$) at the centre of the tetrad (not shown in Fig. 1).

In 1991 it was shown that G-quadruplexes were the most stable form of telomeric DNA and that they were unable to serve as a primer for telomerase-catalysed telomere length maintenance, which requires a DNA single-stranded substrate.[6] There is mounting evidence that G-quadruplex structures may occur in many gene promoter sequences providing an opportunity to selectively target the oncogenes[7] that orchestrate development of the malignant 'hallmarks of cancer'.[8,9] A number of compounds that stabilize G-quadruplexes have been developed as possible cancer treatments,[10] and the structures of some prototypic ligands (**1**–**5**) are shown in Fig. 2. A common feature of compounds **1**–**4** is that the core of the molecules is essentially planar and charged, enabling them to undergo π-stacking interactions above or below the G-tetrad planes with the charged moiety acting as a surrogate to the monovalent metal ions.[11] On the other hand, the natural product telomestatin **5** is uncharged: however, a study utilizing mass spectrometry and molecular dynamics calculations indicates that the macrocyclic core can capture a potassium ion that can stabilize G-quadruplex structures.[12]

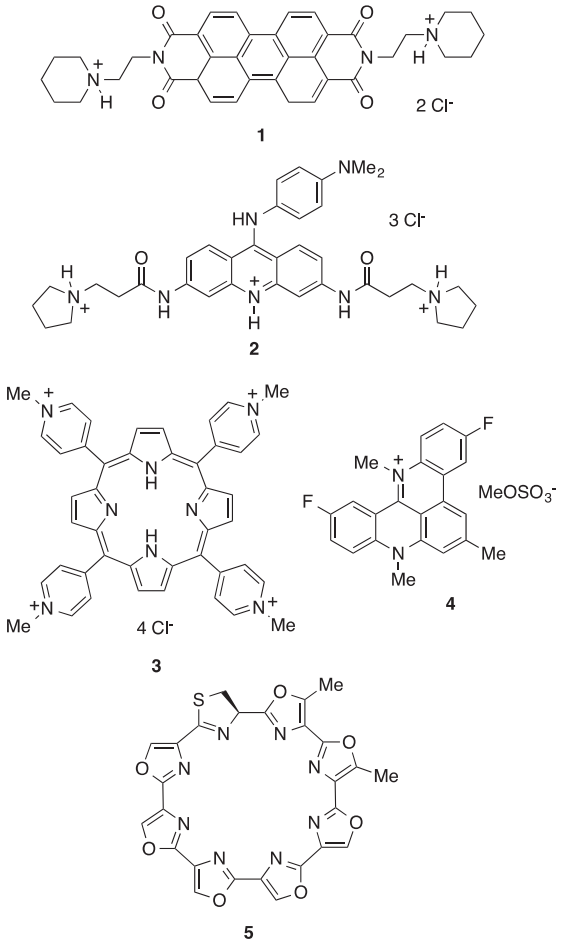


**Fig. 2.** Some DNA quadruplex ligands.

Isocytosine provides an interesting starting point for the design of neutral G-quadruplex ligands owing to its unusual hydrogen bonding properties. This versatility in hydrogen bonding is illustrated by 6-phenylisocytosine **6** and its derivatives. 6-Phenylisocytosine itself is known to exist in an extended hydrogen-bonded lattice of the 1*H*-tautomer in the solid form (Fig. 3A).[13] Our own interest in the 6-arylisocytosine class of neutral molecule was stimulated by the unusual hydrogen bonding patterns revealed in the X-ray structure of the *N*-methylformamide solvate of the isocytosine 2-amino-5-bromo-6-phenylpyrimidin-4(3*H*)-one (bropirimine) **7** that shows bonding between 1*H*- and 3*H*-tautomers (Fig. 3B).[14] Essentially the 3*H*-tautomer presents a 'guanine' face and the 1*H*-tautomer a 'cytosine' face for hydrogen bonding. We considered that such compounds, and linked analogues with the potential for extended hydrogen bonding networks (e.g., compound **8**, Fig. 3C), might be able to interact with G-quadruplex structures and that this encounter might be reflected in novel biological properties. In this paper we report the synthesis of several new 6-arylisocytosines, and an evaluation of their interaction with the 22 base oligonucleotide $AGGG(TTAGGG)_3$, using CD spectroscopy.

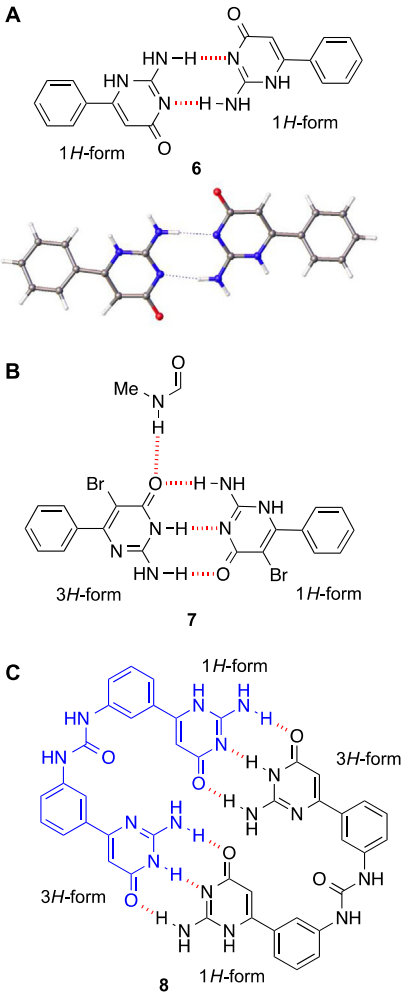


**Fig. 3.** Hydrogen bonding in 6-arylisocytosines. A, dimeric structure formed in solid state between two 1*H*-tautomers of 6-phenylisocytosine **6**;[13] B, dimeric structure formed in solid state between 3*H*- and 1*H*-tautomers of 5-bromo-6-phenylisocytosine **7**;[14] C, proposed extended hydrogen-bonded network for urea linked 6-arylisocytosine **8**.