Europe PMC Funders Group
Author Manuscript
Angew Chem Int Ed Engl. Author manuscript; available in PMC 2008 June 02.

Published in final edited form as:
Angew Chem Int Ed Engl. 2008 ; 47(14): 2677–2680. doi:10.1002/anie.200705589.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Exploring the Differential Recognition of DNA G-Quadruplex Targets by Small Molecules Using Dynamic Combinatorial Chemistry**

Dr. Anthony Bugaut, **Dr. Katja Jantos**, **Dr. Jean-Luc Wietor**, **Dr. Raphaël Rodriguez**, **Prof. Jeremy K. M. Sanders**, and **Prof. Shankar Balasubramanian***
Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW (UK)

Keywords

carbohydrates; combinatorial chemistry; DNA; molecular recognition; quadruplexes

The search for small-molecule ligands of biological targets remains a challenge with major implications for both fundamental studies and drug discovery.[1] We are interested in the discovery of small molecules that specifically interact with regulatory nucleic acid elements. Such molecules have the potential to alter the expression of particular genes and thus influence cellular functions.

Certain guanine-rich (G-rich) regions in genomic DNA can form four-stranded structures, called G quadruplexes, which have emerged as biologically important elements.[2] G-quadruplex formation has been linked to cancer-related biology, most notably by remodeling of the telomere structure or by the regulation of oncogenic expression.[3] The two key challenges in the design of small-molecule[4] ligands for quadruplex DNA are: 1) to attain specificity for G-quadruplex-forming sequences over duplex DNA and 2) to achieve specificity for a given G-quadruplex structure and/or G-quadruplex-forming sequence. The latter criterion has become more important in the light of the recently revealed prevalence of G-quadruplex-forming sequences in the human genome,[5a,b] and particularly in promoter regions.[5c]

Although G quadruplexes all contain G quartets, there is considerable scope for structural variations within the loop and groove regions,[6] suggesting that specificity in the molecular recognition of a quadruplex is attainable. However, the rational design of quadruplex-binding molecules requires a good understanding of the interactions between the ligand and its host. Owing to the paucity of structural data and the dynamic nature of G quadruplexes, combinatorial searches are appealing.

Herein, we report on a study that employs a dynamic combinatorial approach to explore the differential recognition of G-quadruplex targets by closely related small molecules. Dynamic combinatorial chemistry (DCC) is a powerful approach for the rapid identification of binders for small molecules and biological targets.[7] Owing to its adaptive nature, small changes in the composition of a dynamic combinatorial library (DCL) upon introduction of a

**This study was supported by the Cancer Research UK, the EU, and EPSRC. We thank the EPSRC Mass Spectrometry Service for mass analysis.

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

* Fax: (+44) 1223-336-913, E-mail: sb10031@cam.ac.uk.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

target can be used as indicators for attractive interactions between the target and the DCL members. Since its conception, there have been few examples involving nucleic acids as targets.[8] We previously showed that the assembly of molecules that bind to a DNA quadruplex can be templated from a DCL and that DCC can be applied to the selection of ligands that specifically bind a duplex over a quadruplex.[8d,g]

We recently reported on a promising class of quadruplex-binding ligands based on an oxazole-peptide macrocycle.[9] We found that both the number and the length of simple alkylamine side chains appended to this platform could slightly influence quadruplex affinity. To investigate by DCC the potential of different chemical motifs for discrimination in quadruplex binding, we synthesized a thiol analogue of this platform (**1**) and two libraries of side-chain building blocks based on *para*-benzylic thiols (Scheme 1).[10]

Library **L1** includes various chemical motifs that are positively charged at physiological pH and exhibit different potential for hydrogen bonding and electrostatic interactions (**A**-**E**, Scheme 1). Library **L1** (100 μM in each of the building blocks) was combined with the quadruplex-binding platform **1** (100 μM) to generate **DCL1**.[10] The exchange buffer (50 mM Tris/HCl pH 7.4, 150 mM KCl) contained an excess of both reduced (1.28 mM) and oxidized glutathione (0.32 mM) to act as an exchange mediator.[8d,g]

DCL1 was prepared either in the absence or in the presence of various DNA targets (100 μM). The nucleic acid targets used were two intramolecular quadruplex-forming sequences (c-Kit21, c-Myc22) and a 22-mer duplex DNA (dsDNA; sequences are given in Table S1 in the Supporting Information). Oligonucleotides c-Kit21 and c-Myc22 are derived from G-rich sequences found in the promoters of c-KIT and c-MYC proto-oncogenes, respectively.[11,12] In potassium aqueous buffer at neutral pH, c-Myc22 folds into a single parallel-stranded G-quadruplex structure.[11] Using ^{1}H NMR and CD spectroscopy, we showed that c-Kit21 predominantly folds into a quadruplex with a parallel topology under nearly physiological conditions.[12] The main differences between c-Kit21 and c-Myc22 quadruplexes are likely to be in the sequence and size of the loops.

In a typical experiment, **DCL1** was left to equilibrate for three days at room temperature, under air, without stirring.[10] Then the exchange process was stopped by lowering the pH value to 2.[10] For mixtures that included DNA targets, the biotinylated targets were removed by using streptavidin-coated magnetic beads, were heat denatured, and washed several times to release any bound ligands. DCL compositions were then analyzed by UV-HPLC-MS.[10] In the absence of any target, **DCL1** predominantly contained the glutathione adducts of the macrocycle scaffold and of the side-chain building blocks. Only small traces of other homo- and heterodisulfides were detected (Figure 1 a). In the presence of c-Kit21, marked changes in the composition of **DCL1** occurred. Notably, there was a clear decrease in the glutathione-containing heterodimers. Indeed, glutathione was not expected to interact favorably with DNA because of its overall negative charge at pH 7.4. We observed a net amplification of the peaks corresponding to the macrocycle-side-chain conjugates (Figure 1 a,c).[13]

The level of amplification for the macrocyclic species in the presence of c-Kit21 was **1**-**E** > **1**-**A** ≫ **1**-**D** > **1**-**B** ≫ **1**-**C** (Figure 1 c).[14] Although all side chains are positively charged at pH 7.4, this discrimination suggests binding events that are not purely due to nondirectional electrostatic interactions between the positively charged molecules and the polyanionic DNA target, but rather due to the geometry and/or the hydrogen-bonding potential of the side chains. Molecules **1**-**E** and **1**-**A** were then resynthesized on a larger scale, and their binding affinities for c-Kit21 were evaluated by surface plasmon resonance (SPR).[10] Molecule **1**-**E** was found to bind the quadruplex with a dissociation constant (K_d) of 6.6 ±

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

0.1 μM. This value is an approximately 10-fold improvement in affinity as compared with the macrocycle platform **1**, which exhibits a K_d value of 67.5 ± 16.8 μM for c-Kit21. In contrast, the affinity of **1-A** for c-Kit21 (K_d = 10.9 ± 1.9 μM) is about sixfold better than **1**. These results are consistent with the relative amplifications obtained for **1-E** and **1-A**: +2200% and +1900%, respectively (Figure 1 c).

The thiol macrocycle platform **1** binds c-Myc22 with a K_d value of 82.5 ± 9.9 μM, which is slightly higher than the K_d value determined for c-Kit21. In the presence of c-Myc22, the two most amplified adducts from **DCL1** were the guanidinium derivatives **1-A** and **1-E** (Figure 2), as with c-Kit21. The other macrocycle-side-chain products (**1-B**, **1-C**, **1-D**) were only moderately amplified (Figure 2). However, by contrast with the experiment performed in the presence of c-Kit21, **1-A** was this time more strongly amplified (+2200%) than **1-E** (+1700%). Indeed, SPR experiments revealed that **1-A** (K_d = 6.8 ± 1.4 μM) binds approximately 12-fold better to c-Myc22 than **1**, while the acyl-guanidinium side chain **E** increases the affinity by approximately eightfold (K_d = 9.8 ± 0.2 μM).

Furthermore, in contrast to quadruplex targets, the presence of dsDNA did not induce any significant changes in the compositions of **DCL1**, thus suggesting no interaction between the components of the DCL and the DNA duplex, as was also confirmed by SPR. Taken together, these results reveal that subtle chemical variations of positively charged side chains can lead to differences in quadruplex-binding potential. We then investigated whether the stereochemistry of neutral carbohydrate molecules could also affect quadruplex binding.

Many carbohydrates, both neutral and positively charged, are known to be generally good binders for nucleic acids, mainly because of their hydrogen-bonding ability and large hydrophobic patches.[15] As opposed to the more widely studied RNA and duplex DNA, the interactions between carbohydrates and quadruplexes remain largely unexplored. To the best of our knowledge, only a single study dealing with the interaction between a carbohydrate, the aminoglycoside neomycin, and a DNA quadruplex has been published.[16] We thus prepared a series of carbohydrate-based thiols (**L2**, Scheme 1) containing α- and β-substituted derivatives of lyxose (**F** and **G**) and xylose (**H** and **I**). Library **L2** was combined with **1** under the same conditions as those described for **L1** to generate **DCL2**, either in the absence or in the presence of the c-Kit21, c-Myc22, and dsDNA targets (100 μM). As observed previously for **DCL1**, the composition of **DCL2** did not exhibit any detectable changes in the presence of the duplex DNA target. In contrast, significant changes occurred when quadruplex targets were introduced.

In the presence of c-Kit21, macrocycle conjugate **1-F** was the most strongly amplified product (+485%) of **DCL2**. It was amplified about twofold more than its epimer, **1-G** (+245%, Figure 1 d). By contrast, the xylose derivatives **1-H** (+400%) and **1-I** (+350%) exhibited more comparable levels of amplification (Figure 1 d). These results indicate that subtle changes in the carbohydrate geometry, such as the absolute configuration of a single stereogenic center, can give rise to significant differences in amplification, and thus in the affinity for a quadruplex target. To confirm this, binding of all four macrocycle-carbohydrate products were assessed by SPR. Molecule **1-F** was found to bind to c-Kit21 with a K_d value of 9.1 ± 1.1 μM, while **1-G** binds with a K_d value of 23.6 ± 5.1 μM. As expected from the amplification obtained, the discrimination between the macrocycle-xylose adducts was small. The K_d values of **1-H** and **1-I** for c-Kit21 are 16.2 ± 1.8 μM and 17.6 ± 2.6 μM, respectively.

When equilibrated in the presence of c-Myc22, slight variations in the composition of **DCL2** were observed as compared to its composition in the presence of c-Kit21. The most notable change was in the relative amplification of **1-F** and **1-H**. By contrast with amplifications

observed in the presence of c-Kit21, **1-F** and **1-H** were amplified at essentially identical levels in the presence of c-Myc22: +400 and +415%, respectively (Figure 3). Indeed, **1-F** and **1-H** were found to bind to c-Myc22 with similar K_d values of 24.4 ± 4.8 μM and 21.1 ± 3.7 μM, respectively. This result indicates that lyxose side chain **F** is not a generic solution for good quadruplex binding, even when two closely related quadruplex structures are considered. The least amplified species was **1-G** (+210%). It binds to c-Myc22 with a K_d value of 37.2 ± 1.7 μM, which is almost twofold weaker than its epimer **1-H**.

In conclusion, we have shown that DCC is a powerful approach to explore the effect of chemical modifications on the quadruplex-binding properties of a generic ligand (the oxazole-based peptide macrocycle) without the requirement for structural data. A key outcome of our study is the demonstration that subtle chemical and/or stereochemical changes can tune the affinity of the ligand for a particular quadruplex target, and can thus lead to differential recognition of DNA G quadruplexes. Furthermore, this work introduces carbohydrate molecules as promising motifs for selective quadruplex recognition.

Supporting Information

Refer to Web version on PubMed Central for supplementary material.

References

[1]. Schreiber SL. Science. 2000; 287:1964–1969. [PubMed: 10720315]

[2]. For a general review on quadruplexes, see: Davis JT. Angew. Chem. 2004; 116:684–716.; Angew. Chem. Int. Ed. 2004; 43:668–698. [PubMed: 14755695]

[3]. a) Zahler AM, Williamson JR, Cech TR, Prescott DM. Nature. 1991; 350:718–720. [PubMed: 2023635] b) Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. Proc. Natl. Acad. Sci. USA. 2002; 99:11593–11598. [PubMed: 12195017] c) Bejugam M, Sewitz S, Shirude PS, Rodriguez R, Shahid R, Balasubramanian S. J. Am. Chem. Soc. 2007; 129:12926–12927. [PubMed: 17918848]

[4]. For a recent report on G-quadruplex ligands, see: Searle, MS.; Balkwill, GD. Quadruplex Nucleic Acids. Neidle, S.; Balasubramanian, S., editors. Cambridge: RSC Biomolecular Sciences; 2006. p. 131-153.

[5]. a) Huppert JL, Balasubramanian S. Nucleic Acids Res. 2005; 33:2908–2916. [PubMed: 15914667] b) Todd AK, Johnston M, Neidle S. Nucleic Acids Res. 2005; 33:2901–2907. [PubMed: 15914666] c) Huppert JL, Balasubramanian S. Nucleic Acids Res. 2006; 35:406–413. [PubMed: 17169996]

[6]. a) Simonsson T. Biol. Chem. 2001; 382:621–628. [PubMed: 11405224] b) Phan AT, Kuryavyi V, Patel DJ. Curr. Opin. Struct. Biol. 2006; 16:288–298. [PubMed: 16714104]

[7]. Selected reviews: a) Lehn J-M. Chem. Eur. J. 1999; 5:2455–2463.b) Rowan SJ, Cantrill SJ, Cousins GRL, Sanders JKM, Stoddart JF. Angew. Chem. 2002; 114:938–993.; Angew. Chem. Int. Ed. 2002; 41:898–952. [PubMed: 12491278] c) Corbett PT, Leclaire J, Vial L, West KR, Wietor J-L, Sanders JKM, Otto S. Chem. Rev. 2006; 106:3652–3711. [PubMed: 16967917]

[8]. a) Klekota B, Hammond MH, Miller BL. Tetrahedron Lett. 1997; 38:8639–8642.b) Klekota B, Miller BL. Tetrahedron. 1999; 55:11687–11697.c) Karan C, Miller BL. J. Am. Chem. Soc. 2001; 123:7455–7456. [PubMed: 11472190] d) Whitney AM, Ladame S, Balasubramanian S. Angew. Chem. 2004; 116:1163–1166.; Angew. Chem. Int. Ed. 2004; 43:1143–1146. [PubMed: 14983458] e) Bugaut A, Toulmé J-J, Rayner B. Angew. Chem. 2004; 116:3206–3209.; Angew. Chem. Int. Ed. 2004; 43:3144–3147. [PubMed: 15199562] f) Bugaut A, Bathany K, Schmitter J-M, Rayner B. Tetrahedron Lett. 2005; 46:687–690.g) Ladame S, Withney AM, Balasubramanian S. Angew. Chem. 2005; 117:5882–5885.; Angew. Chem. Int. Ed. 2005; 44:5736–5739. [PubMed: 16078284] h) Bugaut A, Toulmé JJ, Rayner B. Org. Biomol. Chem. 2006; 4:4082–4088. [PubMed: 17312962]

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

[9]. Jantos K, Rodriguez R, Ladame S, Shirude PS, Balasubramanian S. J. Am. Chem. Soc. 2006; 128:13662–13663. [PubMed: 17044674]

[10]. Experimental details are given in the Supporting Information

[11]. Fernando H, Reszka AP, Huppert J, Ladame S, Rankin S, Venkitaraman AR, Neidle S, Balasubramanian S. Biochemistry. 2006; 45:7854–7860. [PubMed: 16784237]

[12]. Ambrus A, Chen D, Dai J, Jones RA, Yang D. Biochemistry. 2005; 44:2048–2058. [PubMed: 15697230]

[13]. It is noteworthy that a large number of homo- and heterodimers of side chains were amplified as well, which led to other emerging peaks apart from peaks containing the macrocycle conjugates. As we expected such products to bind less tightly and with lower specificity than the macrocycle conjugates, we did not attempt to perform a detailed analysis for these products. A selected homodimer, **A**-**A**, which was amplified by about 250% was found to bind to c-Kit21 with a K_d value of 54.8 ± 11.5 μM and a high stoichiometry of about 3:1 **A**-**A**/c-Kit21

[14]. Proportion changes are based on peak areas normalized with respect to an internal standard. Owing to overlapping peaks in **DCL1**, percentage changes were calculated on the basis of the integration of relevant ion count peaks on the MS spectra. Hotchkiss T, Kramer HB, Doores KJ, Gamblin DP, Oldham NJ, Davis BG. Chem. Commun. 2005:4264–4268. [PubMed: 16113716]

[15]. a) Kahne D. Chem. Biol. 1995; 2:7–12. [PubMed: 9383398] b) Hermann T, Westhof E. Biopolymers. 1998; 48:155–165. [PubMed: 10333743]

[16]. Kaiser M, Cian AD, Sainlos M, Renner C, Mergny J-L, Teulade-Fichou MP. Org. Biomol. Chem. 2006; 4:1049–1057. [PubMed: 16525549]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

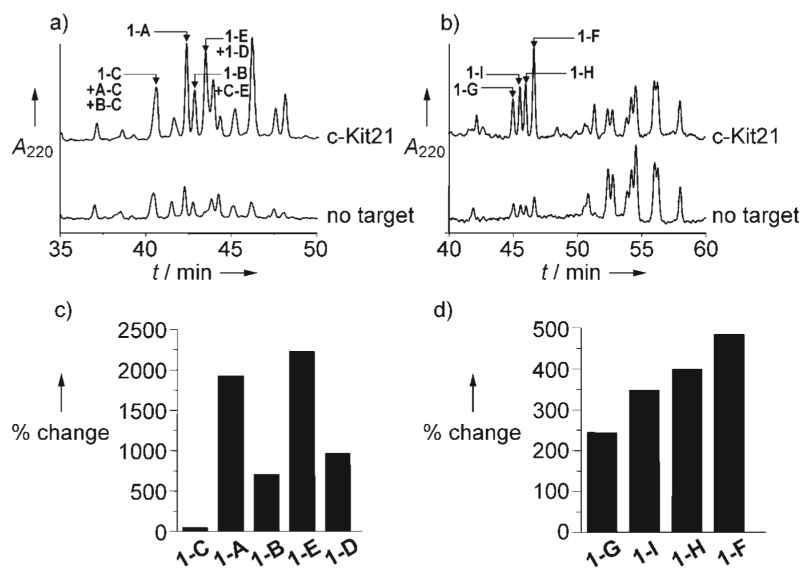


Figure 1.

Expansions of the HPLC traces of templated and untemplated a) **DCL1** and b) **DCL2**, and proportion changes of macrocycle-side-chain conjugates in c) **DCL1** and d) **DCL2** upon introduction of c-Kit21.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

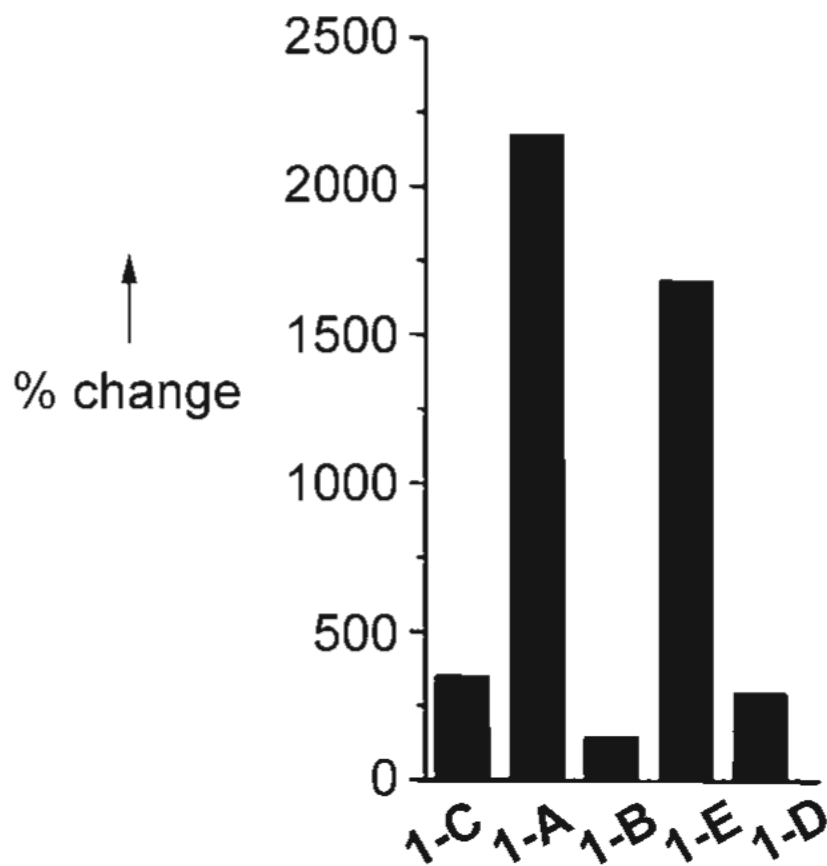


Figure 2.
Proportion changes of the macrocycle-side-chain conjugates in **DCL1** upon introduction of c-Myc22.

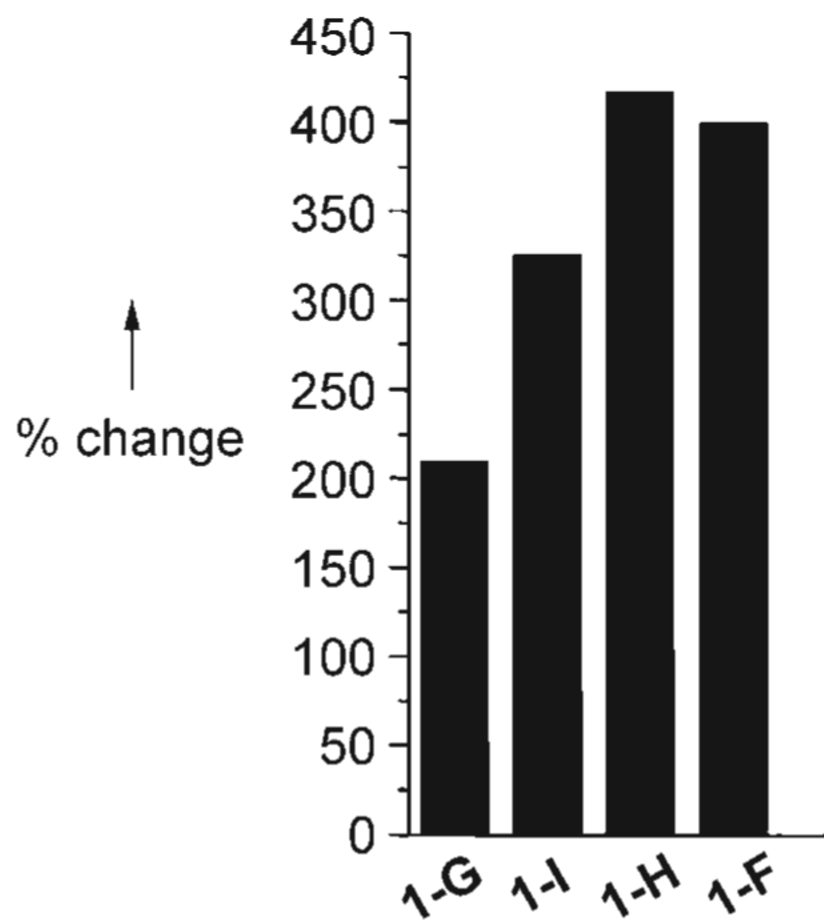
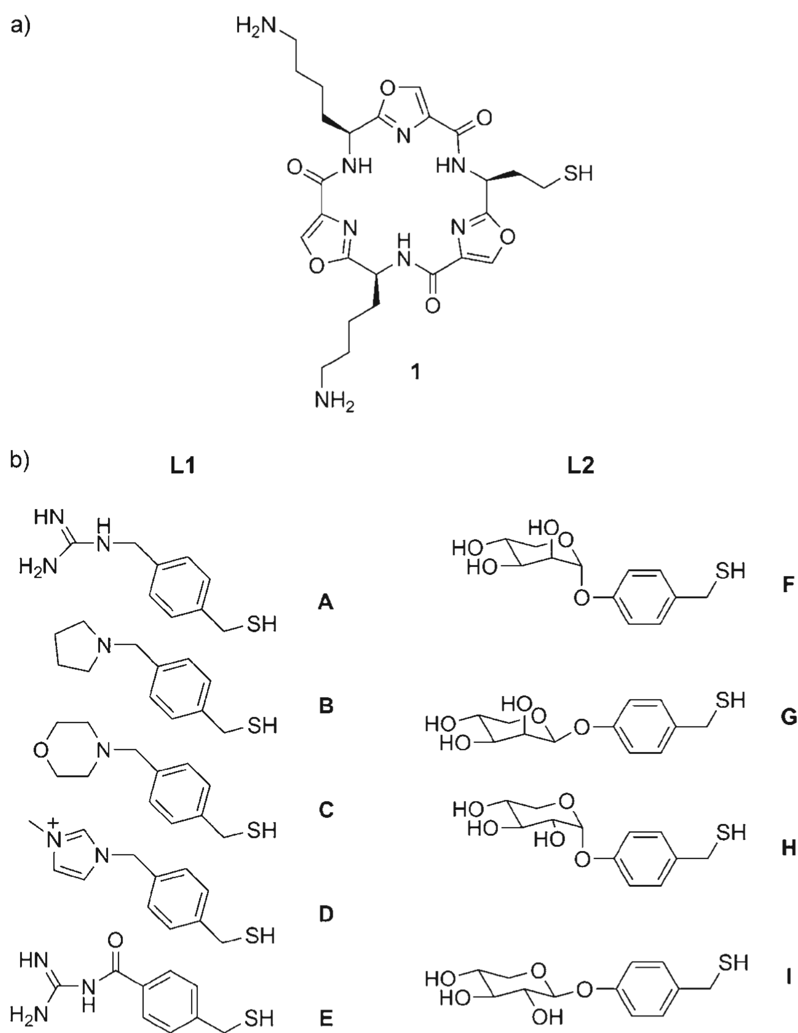


Figure 3.
Proportion changes of the macrocycle-side-chain conjugates in **DCL2** upon introduction of c-Myc22.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



Scheme 1.

a) Structure of the oxazole-based peptide macrocycle **1**. b) Structures of side chains that are cationic (**A**-**E**, library **L1**) and neutral (carbohydrate derivatives **F**-**I**, library **L2**) at physiological pH value.