Peptide scalpels for site-specific dissection of the DNA-protein interface

Dieter Verzele[§], Lieselot L. Carrette[§], Annemieke Madder[*]

Laboratory for Organic & Biomimetic Chemistry, Department of Organic Chemistry, Ghent University, Krijgslaan 281 (S4), B-9000 Ghent, Belgium

Current review highlights emerging strategies for design of DNA-binding peptides and cross-linking techniques capable of interference at the DNA-protein interface with emphasis on bottom-up organic synthesis. Despite the obvious fundamental character, these inquiries could ultimately result in interesting biomedical applications as designed genome interfering agents and as potential (chemo)therapeutics or diagnostics and recent awareness of the potential of these larger peptidomimetic constructs has risen within the medicinal society.

Introduction

Aberrant modulation of gene expression at the transcriptional level is at the origin of numerous diseases. By disregulating cell growth and triggering cell proliferation, various oncoproteins carry out their biological functions as transcription factors, critically relying on their DNA-binding capacities. Artificial transcription inhibition or modification is possible via various approaches through different interfaces (Fig. 1). Selective recognition between the reaction partners on the interface is essential for success. Truncated mimics that lack, for example, the activation domain while retaining the DNA-binding requirements might interfere with DNA-complexation of the malignant natural counterpart, disrupt transcriptional induction and ultimately suppress the oncogenic activity [1-2]. This short review will treat

[§] Equal contribution of both authors
[*] *Corresponding author*: A. Madder (annemieke.madder@ugent.be)
URL: http://www.orgchem.ugent.be/biomimetic.html

such peptidic transcription factor mimics, acting on the protein-DNA interface by binding the duplex, ultimately towards irreversible blocking via interface cross-linking. Besides the simple desire to mimic Nature, these miniature probes hold potential for disentanglement of the molecular basis of DNA-targeting by natural proteins.

Figure 1. Different biological interfaces are involved in the processes of transcription and translation. **(a)** Typically several proteins have to associate to obtain an active transcription factor and/or to form the transcription initiation complex on the DNA. Disruption of these interactions can be modulated by small molecules, which block the interaction [3-5]. **(b)** Promising strategies exploiting the nucleic acid interface aim at inhibition of translation. This can be achieved by blocking mRNA with a short complementary RNA fragment (microRNA). Alternatively a double stranded RNA fragment (small interfering RNA) together with the RNA-induced silencing complex (RISC) can degrade mRNA [6, 7]. **(c)** The DNA-protein interface is an emerging site for new generation therapeutics [8]. Interference can be achieved either at the protein level with decoys (small synthetic DNA fragments that bind to the transcription factor by mimicking the specific DNA binding site) [9] or at the DNA level by blocking the target site. The latter can be realized by a small intercalating molecule, which slides in between DNA base pairs; a triplex forming oligonucleotide, which binds in the major groove of DNA through Hoogsteen interactions; a small molecule binding in the major groove or a large molecule binding in the major groove. In the latter approach, initial strategies focus on DNA-binding, and at most inhibition of gene expression. By linking this DNA-recognizing domain to an activating effector domain, selective modulation of the transcription process could ultimately be envisioned. While size of the compounds will probably hamper both synthesis and permeability, fascinating examples have been disclosed for the (synthetic) minor-groove binding Dervan polyamides or (mostly, if not all biotechnologically-engineered) major-groove targeting zinc fingers. Successes are most certainly related to the exquisite control of recognition specificity, the modularity of systems and the synthetic accessibility, as detailed hereafter.

Artificial transcription factor mimics: gradual expansion of the chemical space

Fifty years after the milestone discovery of the Watson and Crick double-helical structure of DNA, the official completion of the Human Genome Project (HGP) in 2003 was another landmark event in the history of natural sciences [10, 11]. Multidisciplinary research at the interface of various disciplines became prerequisite to distill the flood of new information and reap maximum reward from the genome data. The nascent discipline of chemical biology approaches biological problems from the chemist's point of view and employs its ability to proficiently design, prepare and characterize tailor-made compounds, which can be subsequently applied to probe biological systems in highly-controlled, well-defined experiments (reviewed in [12]). Organic synthesis occupies an indispensable position in the

discovery process, enabling the craftful preparation of tailor-made conjugates, not limited by the genetic code or the strictures of the ribosomal machinery. Through sophisticated strategies, the rigorous control of molecular structure and function provides access to the vast expanse of chemical space [13]. The targets of chemistry research in these days are no longer restricted to small drug-like molecules. Much of the intellectual adventure and challenge has gradually converged with biology and the size of aimed compounds has steadily increased. The activity, selectivity and stability of biopolymers become controllable, and they can be tailor-made like small organic compounds. Furthermore, chemical protein synthesis has enabled the systematic development of proteins with enhanced potency and specificity as candidate therapeutic agents [14]. Consequently eliciting a recent awareness within the pharmaceutical society for the unexplored opportunities of the generally ignored chemical space (GICS), medicinal chemists start thinking outside their Lipinski box [15].

Despite the promising relevance of probing simplified peptide versions derived from natural protein counterparts, progress has been remarkably slow, failing to keep pace with the plethora of resolved complexes. The fragile balance between structural minimization and biophysical outcome impedes rational design of down-sized mimics, demanding devoted trial-and-error tuning of empirical constraints.

Figure 2. Non-exhaustive, pictural overview of synthetic peptidic groove binders, based on the natural DNA binding moieties distamycin (a), zinc fingers (b) or basic region leucine zippers (c). Based on distamycin, Dervan developed hairpin minor groove binding polyamides of pyrrole (Py), imidazole (Im) and hydroxypyrrole (Hp) **(d)**. Small monomeric peptides can bind the major groove as α-helix, if the associated entropic penalties are compensated via preorganization **(e)**. Hybrid conjugates of a tripyrrole sequence and a peptide monomer represent another strategy to cope with thermodynamics. The former module assists in the major groove recognition by the latter module which mimics either the atypical zinc finger, GAGA **(f)** or the basic region leucine zipper GCN4 **(g)**. Furnishing genuine mimics of natural proteins, a final approach consists in the combination of several major groove binding peptides, such as homodimeric GCN4 **(h)** or heterodimeric cMyc-Max **(i)** sequences.

Minor Groove Recognition: synthetic polyamides

Dervans polyamides, based on the minor groove binder Distamycin A (Fig 2a), can be regarded as peptides with each α-carbon replaced by a five membered ring (Fig 2d). Natural Distamycine is a crescent shaped, small polyamide with three pyrrole (Py) rings that binds AT rich sequences in the minor groove selectively. Dervans group demonstrated that substituting pyrrole by imidazole (Im) and introducing hydroxypyrrole (Hp), it became possible to recognise all bases by a specific combination of two rings [16]. The genetic code was cracked. As the Dervan peptides bind the minor groove in a 2:1 fashion, affinity could be further increased by linking two peptides.

Additional modifications introduce specific properties to the DNA binders. Fusing a polyarginine to the tripyrrole increases binding affinity and promotes nuclear uptake even though these small molecules are already relatively cell permeable [17]. Linkage of a Dervan peptide, binding in the regulating domain of a gene, with the activator domain of a transcription factor allows for selective transcriptional activation of that specific gene [18]. Conjugation with a natural alkylating moiety, like CC-1065 or duocarmycin, enhances the efficiency and selectivity of the DNA alkylation and can be used for gene silencing, as alkylated coding regions cannot be read by RNA polymerase II [19, 20]. Except for the TATA box binding protein, most natural proteins bind in the major groove. Except for a few cases where major groove protein binding is inhibited by minor groove binding Dervan peptides [21], the latter are unsuitable for blocking the protein-DNA interface, as it was shown that both can bind at the same time.

Ultimately, to uniquely distinguish a site in the three gigabase human genome, a recognition sequence of 16 base pairs is required. Although examples of polyamides that can realise this exist, mismatch sequences will probably not disrupt binding, due to the great binding affinity of such a single molecule-DNA binder [22].

Major Groove Recognition: stabilised monomeric helices

The majority of well-characterized families of native DNA-binding proteins rely on recognition α-helices to make base contacts in the major groove. The overall shape and dimensions of an α-helix allow it to fit into the major groove in a number of related, but significantly different ways. This variability explains the preference of natural proteins for targeting this wider groove over the narrower minor one.

The design of peptides mimicking the DNA binding domain of such proteins is not straightforward, as most α-helices lose the capability to fold and thus bind DNA with high affinity when removed from their natural context.

Various possible strategies have been described [23 and references cited therein] and consist in the use of peptide fragments that stabilise an α-helix (Fig 2e). Successfull substitution of non-crucial amino acids with the helix promoting alanine can be achieved. Helical structures can also be stapled by the formation of a lactam bridge between lysine and aspartic acid, four positions further in the chain. Alternatively one can make use of stable α-helices, as in the small, well-folded avian pancreatic polypeptide, where an α-helix is stabilised by hydrophobic interactions on one side of the helix. The other side of this helix can then be used to graft residues required for DNA recognition and binding.

Exploitation of minor groove conjugation for major groove recognition

It was shown that the covalent attachment of peptides to the DNA backbone through a linker can compensate for the entropic cost associated with secondary structure formation and bind DNA as an α-helix. Next to intramolecular binding to increase the affinity for DNA, cooperative binding is another option [23 and references cited therein]. This can be achieved by fusing intercalators or Dervan peptides to the major groove binding peptide (Fig 2f and

2g). Combining different smaller DNA binding units, overall binding selectivity can be significantly improved.

Engineered Zinc Fingers: gripping the major groove

In the area of zinc finger transcription factors, an advanced level of insight and pairing control has been achieved [24]. Whereas other DNA-binding proteins generally make use of the twofold symmetry of the double helix, zinc fingers are linearly linked in tandem to recognize DNA stretches of varying lengths, with high discrimination fidelity. The ability to bind specifically to virtually any given DNA sequence, combined with the potential of fusing polydactyl zinc finger peptides with effector domains, has led to the technology of engineering of chimeric DNA modifying enzymes (equipped with e.g. methylase, nuclease or recombinase modules) and transcription factors (simple blockage without effector, repression vs. activation domains, and chromatin remodeling factors as examples) [25]. Opening the possibility of using the engineered zinc finger-based factors as novel human therapeutics, ongoing successes stimulated the creation of the first biotech companies to exploit this new platform (Sangamo BioSciences, Inc., in Richmond, California; and later Gendaq Ltd., as a spin-off of the MRC Laboratory of Molecular Biology, Cambridge, UK). Interestingly, customized zinc finger constructs for given DNA sequences can now be acquired commercially form Sigma-Aldrich, branded as Compo-Zr technology. Table 1 includes an example of a Zinc finger construct for therapeutic application. Although a vital sequel in the gradual expansion of the chemical space covering artificial transcription factors, most (if not all) zinc finger designs are biotechnologically-engineered and as such beyond the scope of present overview, emphasizing on miniature constructs synthesized via bottom-up organic synthesis. Despite numerous further studies and applications, ranging from genome

annotation tools to (pre-)clinical implementation, the interested reader is therefore referred to excellent reviews recently disclosed in literature [26,27].

Artificial zipper miniatures seizing the major groove

The α-helical basic zipper (bZIP) and basic helix-loop-helix-zipper (b-HLH-ZIP) motifs are among the simplest protein structures able to bind the DNA major groove in a sequence-specific fashion. The simplicity and tractability of the bipartite zipper-type proteins has made them attractive frameworks for the design of artificial counterparts. The variety of designs exhibiting native protein-like targeting of DNA-sequences by employing short basic region peptides, dissected from naturally occurring bZIP transcription factors, demonstrates the potential of artificial miniatures as biomolecular recognition devices. However, considering above discussion, the inherently-low DNA-binding affinity of short, isolated, natural basic region peptide monomers does not surprise. As mentioned earlier, suppression of bZIP binding capacities by presenting so-called monomeric reading heads illustrates the delicate balance between entropy and enthalpy.

Research demonstrated that the basic region as such contains sufficient information for DNA recognition, provided helical structures are stabilized (*vide supra*) or artificial dimerization is applied. Predominantly serving as a structural scaffold to pre-organize the relatively small, N-terminal recognition α-helices, the extended C-terminal dimerization domain can be replaced by various connectors, both covalent and non-covalent [23 and references cited therein, 28]. Since the original disclosure of a cystine-bridged peptide dimer mimicking the homodimeric yeast GCN4 leucine zipper protein, development of subsequent generation bZIP miniatures has been intensely pursued. Early work of Schepartz et al. focused on the chemical synthesis of (homodimeric) GCN4 mimics upon artificial dimerization via formation of a non-covalent bis(terpyridyl)-iron(II) coordination complex. Related contributions were further disclosed by

Morii et al., typically employing either a non-covalent β-cyclodextrin:adamantyl host-guest inclusion complex or covalently bridged enantiomeric biphenyl derivatives as artificial dimerization interfaces in a variety of homo- and heterodimeric GCN4-based designs.

The apparent hibernation in this research area notwithstanding, Mascareñas' artificial GCN4 homodimer, represented in Fig. 2h, illustrates current state-of-the-art in the comprehension of design principles and control over binding properties for simultaneous DNA major groove targeting. Incorporation of a photo-isomerisable azobenzene moiety allowed for generating a light-modulated DNA-binding peptide.

Although sharing a similar mode of binding DNA target sites with the extensively scrutinized bZIP-equivalents, artificial mimics of the b-HLH-ZIP transcription factors are conspicuously few. Despite the exceptional total chemical synthesis (172 residues) yielding a covalently-stitched replica of the vertebrate cMyc-Max (proto-) oncoprotein achieved by Kent et al. in 1995 [29], translation of this precedent towards miniature derivatives has been long overdue. Related to the bZIP and b-HLH-ZIP type proteins, the singular MyoD-MyoD bHLH-type mimic, reported by the Morii group [30], before finally resorting to the GCN4 standard, confirms the interest for corresponding miniaturization efforts.

Directly involved in human tumorgenesis and cancer, the pivotal cMyc member of the Myc/Max/Mad network is a prominent example of how these proteins are promising targets towards novel, potent, transcription-targeting chemotherapeutica [31]. However, (1) the influence of loop-projecting DNA-contacts, (2) the ambiguous role of loop-associated flexibility, (3) the constraining/stabilizing interactions with essential accessory mediators of the transcriptional machinery in vivo and (4) the seeming regulatory subtleness of the helix-loop-helix transcription factor family conspire against the exploitation of their peptide counterparts in design efforts. In this respect, the heterodimeric Jun-Fos bZIP oncoprotein

bears special mentioning, combining a cMyc-like cancer relevance with the GCN4-like structural tractability.

While a decade passed since Mascareñas' triggerable GCN4-GCN4 bZIP contribution mentioned above, the contrasting void of helix-loop-helix counterparts stimulated initiatives from our own group. Recently, we achieved a first generation of down-sized b(-HLH-)ZIP models with emphasis on the therapeutically relevant cMyc-Max combination (Verzele, D. (**2009**) Development of miniature b(-HLH-)ZIP peptidosteroid conjugates as new transcription factor models. *PhD dissertation*, Ghent University, Madder, A.). Based on covalent restriction by a steroid core with defined geometrical properties (as depicted in Fig. 2i), the orthogonal reactivity of the attachment points is an additional feature of the artificial linkage, permitting straightforward synthesis of not only homo- but also heterodimer peptidosteroid conjugates.

Table 1 highlights some of the therapeutic developments directly derived from the here discussed synthetic DNA binding peptidic constructs.

Table 1. Therapeutic developments of DNA-binding peptides

Name	Structure	Function	Stage	Ref
Distamycin A	Unmodified tripyrrole polyamide	No anti-tumor activity	No applications	[44]
Tallimustine	Benzoyl nitrogen mustard derivative of distamycin	Antitumor activity.	Phase II (haltet due to severe myelotoxicity)	[44]
Dervan-peptides	- Unmodified polyamide (binding in regulatory sequence)	Gene silencing of several genes (5S RNA, HIV-1 RNA, vascular endothelial growth factor, human transforming growth factor, prostate specific antigen, topoisomerase IIα, …)	Cell cultures	[19]
	-Polyamide (binding in regulatory or coding sequence) conjugated with an alkylating moiety (duocarmicine (Du86), CPI, CBI, CC-1065, dystamicine)	Gene silencing of several genes	Cell cultures	[19]
	-Polyamide conjugated with chlorambucil (nitrogen mustard)	Blocking tumorgenicity of carcinomic cells	Animal model (nude mice)	[45]
	-Polyamide modified with an activation domain	Activation of gene expression	In vitro (yeast nuclear extract)	[18]
	-Polyamide modified with a positive patch (N-aminoalkylpyrrole, Arg-Pro-Arg or polyamine)	Inhibition of major groove binding proteins (ex. bZip)	In vitro (purified TF)	[21]
PNU-145156E	Tetrasodium salt of tetra-pyrrole naphthalenedisulfonic acid	Antitumor activity: Believed to be anti-angiogenic and synergistic with other cytotoxic drugs	Phase I (halted due to its unique toxicity profile)	[44]
Brostallicin (PNU-166196)	Bromo-acrylamido tetra-pyrrole derivative	Antitumor activity	Phase III (generated significant interest in the clinic)	[46]
MVZ+426	Designer zinc finger transcription activator of the (VEGF)-A gene	Treatment of human peripheral arterial obstructive disease by stimulating vessel growth	Phase II	[47]

Crosslinking methods in development

In order to block the protein-DNA interface, the above described mimics need to compete with the natural DNA-binding proteins. Capitalizing on previous successes at the protein-protein interface, crosslinking moieties can be introduced to overcome the competition problem.

Exogenous bifunctional crosslinkers

For crosslinking purposes, it is most easy to add small reagents, like platina complexes, nitrogen mustards, psoralen or reactive aldehydes. In this case no modified DNA or protein is required and their size allows easy penetration in cell and nucleus. Most current chemotherapeutics belong to these compound classes and result in DNA intra- or interstrand crosslinking. The impact of covalent attachment of proteins was recognised more recently. It was shown that cisplatin, the best known anticancer drug, crosslinks nuclear matrix-bound transcription factors and cofactors to DNA [32]. The problem with these external reagents is their lack of selectivity. Selectivity can be dramatically increased by equipping either DNA or protein with reactive groups. Alternatively, incorporation of more selective photoactivatable crosslinking moieties has gained popularity. Over the years, an extensive collection of crosslinking techniques has been developed, mainly to study transient complexes between proteins or between proteins and DNA. Since it is not the intention of this paper to give an exhaustive list, the reader is referred to more specialised literature [33-35]. Table 2 intends to summarize the different chemical strategies for crosslinking.

DNA → Protein crosslinking

Many examples exist in which nucleotides are equipped with a reactive functionality, in order to form a covalent bond to a non-covalently bound protein. In first instance, it is possible to

modify the phosphate backbone, for example with pyrophosphate internucleotide groups [36] or disulfide groups. In the last case a PSS-backbone (phosphorothiosulfide) is formed, which can break and further react selectively with protein cysteine groups. Cells treated with a PSS-oligonucleotide of the binding site for NF-κB transcriptional factor underwent apoptosis [37]. Secondly, nucleotide building blocks equipped with a reactive functionality (aldehyde, aziridine, alkylating moiety) can be used in automated DNA synthesis for the production of reactive oligonucleotides [38]. Additionally, various approaches for postsynthetic regioselective modification to introduce reactive functionalities into DNA have been described [39]. So called decoys, as depicted in Fig 1, have been developed for irreversible crosslinking and inhibition of transcription proteins [36].

Protein → DNA crosslinking

As for crosslinking to DNA, strategies involving TFO (see Fig 1) for covalent blocking of the genome have been considered. Inhibition of transcription in cells was observed after application of a psoralen modified oligonucleotide and irradiation [40].

Alternative approaches involving crosslinking peptides or proteins for achievement of covalent DNA blocking have received considerably less attention. The most straightforward way to introduce a crosslinking functionality in peptides or proteins involves cysteine side chain modification. This was described for determination of the orientation of a DNA binding motif within a complex [41] and later applied for determination of other DNA-protein complexes. Furthermore, unnatural reactive amino acids can be included during solution or solid phase peptide synthesis. Even in proteins unnatural amino acids can be introduced via the natural or mutated translation system [42, 43]. Recently a photoactivatable crosslinking protein, obtained by genetically encoding for a benzophenone containing amino acid was crosslinked to its DNA binding site [43].

Although application of specifically engineered proteins for DNA crosslinking has thus proven feasible, the clear preference for the above described reverse approach can be explained in terms of the unpredictable influence of protein modification on DNA binding affinity. The DNA binding peptides described above, being potent candidates for mimicking their parent proteins, offer new possibilities in this area. As mentioned before, introduction of sequence selectivity to DNA targeting agents does improve their efficiency as anticancer drug. Following this trend, crosslinking transcription factor mimics might play an important role in future cancer treatment.

Table 2. Classification of chemical strategies for crosslinking

Technology	Small bifunctional crosslinking entities	Reactive DNA	DNA with inducible reactivity	Sulfur modified DNA
Cross-linking moieties	cisplatin, nitrogen mustards, psoralen, aldehyde, epoxides, ...	- nucleotide modified with aldehyde, crotonaldehyde, trans-4-hydroxynonenal, cyclopropapyrroindole, substituted pyrophosphate, … - post-synthetic regioselective modification of sugars in oligonucleotides with vicinal diols	- photoactivatable functionalities: azide, diazirine, benzophenone - intrinsic photoreactivity of DNA	- PSS-backbone oligonucleotide (phosphorothiosulfide) - N^2 thiol modified deoxyguanosine
Pro	- no modification required of DNA/protein/peptide - bioavailability	- embedded crosslink functionality - better selectivity	- on demand activation within formed complexes - increased selectivity for target protein	- selective for cysteine - very small destabilization of the DNA duplex
Con	- lack of selectivity	- non-specific reactivity before complex formation	- activation required - possible destruction of compounds under activation conditions - low yield	- correct and accurate positioning of the modification required - formed disulfide link metabolically unstable
Selected examples	Crosslinking of transcription factors CREB and JUN or nuclear matrix-bound transcription factors to their recognition sequence	Crosslinking of DNA to restriction-modification enzymes, HNF1 transcription factor and T7 RNA polymerase	Crosslinking of DNA to DNA polymerase, study positioning of Lex A repressor in the major groove, photoaffinity labelling	Crosslinking of NF-κB decoy to the p50 subunit or crosslinking of HIV-1 reverse transcriptase to its template
References	[32, 35, 48]	[36, 39, 49]	[33, 34]	[37, 38]

Conclusions

Illumination of the molecular mechanism by which cells adapt their phenotype in response to external stimuli is a central objective in modern life sciences. The use of exogenous chemical approaches based on simplified probes derived from natural proteins provides a potent alternative to interrogate and decipher cellular behavior, supplementing biological methods (e.g. knock-out organisms or expression profiles). Chemical biology is eliciting a paradigm change in the way scientists approach therapy-directed biomedical questions and diagnostics development, improving innovation and productivity towards drug discovery of the future. Prominent practitioners of chemical biology research recently advocated the eminent role of this burgeoning field in drug discovery, towards innovative cures of the 21st century [12]. Harboring the potential to bridge fundamental academic aspirations with pharmaceutical industry's objectives, i.e. to genuinely alter the perception of the pharmacologically achievable, this integrative discipline offers invaluable incentive to elicit uncovered therapeutic targets. The ultimate goal remains the expansion of the diagnostic toolkit of biomarkers and the enrichment of the limited arsenal of contemporary pharmacopeia. Corroborating the relevance and opportunities presented in current review, a profound impact of tailor-made bioconjugates, down-sized probes and small molecule libraries is anticipated, and so-called 'undruggable' targets (and processes) such as (oncogenic) transcription factors might accordingly shift into the realm of the medicinally conceivable. Bayer's acknowledgment for the fundamental research on transcription regulation at a molecular level by Prof. Dr. Patrick Cramer (see links) illustrates this change of mentality, while the 2006 Nobel Prize of Chemistry awarded to Prof. Dr. Roger Kornberg (see links) confirms the value of continuing efforts.

Designed mimics of transcription factors are valuable to assess the implications of the exquisite recognition selectivity of natural transcription factors on the control of gene

expression, towards molecular dissection of the complex regulatory networks and functional annotation of genetic information. Primarily unveiling fundamental details of the dynamic interplay between biomolecular systems, biomedicinal diagnostics and anti-gene chemotherapeutics might eventually arise from elucidation of key interactions, interrogation of aberrant transcriptional behavior, and identification of new genetic targets. Recent progress in peptide delivery systems and the prospect of developing peptidomimetic counterparts of in vitro active derivatives seem to warrant the ultimate medicinal potential of this research area. Tackling human disease in an early stage at the fundamental DNA-protein interface, the significantly rewarding ends for cancer research demonstrate the relevance and justify the means of ongoing endeavors.

Acknowledgements

The authors would like to apologize to all researchers whose contributions could not be included in this review article due to space limitations. Dieter Verzele gratefully acknowledges financial support from Ghent University (BOF 011D16403) and the FWO Vlaanderen (KAN 1.5.186.03; G.0347.04N). Lieselot Carrette is indebted to the FWO-Vlaanderen for an aspirant position.

Links

- The PIN Database (Proteins Interacting in the Nucleus): http://pin.mskcc.org/home.html
- The Nobel Prize in Chemistry 2006 to Roger D. Kornberg (Stanford University, USA) for his studies of the molecular basis of eukaryotic transcription: http://nobelprize.org/nobel_prizes/chemistry/laureates/2006/
- Recently acknowledged at the Bayer website, Prof. Dr. Patrick Cramer (LMU Munich, Germany) occupies a prominent role at the forefront of transcription research, elucidating the chemical processes which turn the building blocks of DNA into a living, biological process. Podcast: http://www.podcast.bayer.com/module/podcast/tv-research-en.xml
  2009 Hansen Family Award: http://www.research.bayer.com/edition-21/science_and_education.aspx
- Sangamo BioSciences, Inc., biopharmaceutical company exploiting the zinc finger technology: http://www.sangamo.com/
- The Sigma-Aldrich CompoZr^{TM} zinc finger technology offer: http://www.compozrzfn.com/

- The Zinc Finger Database (ZiFDB): http://bindr.gdcb.iastate.edu/ZiFDB
  (Fu, F. et al. (2009) Zinc Finger Database (ZiFDB): a repository for information on C2H2 zinc fingers and engineered zinc-finger arrays. *Nucleic Acids Res.* 37 (database), D279-D283)
- Zinc Finger Targeter (ZiFiT), a web-based tool to facilitate design of zinc finger proteins: http://bindr.gdcb.iastate.edu/ZiFiT
  (Sander, J. et al. (2007) Zinc Finger Targeter (ZiFiT): An engineered zinc finger/target site design tool. *Nucleic Acids Res.* 35 (web server), W599-W605)

Related articles

Hurley, L.H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer* 2 (3), 188-200
Uil, T. et al. (2003) Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities. *Nucleic Acids Res.* 31 (21), 6064-6078
Redell, M.S. and Redell. (2006) Targeting transcription factors in cancer: Challenges and evolving strategies. *Drug Discov. Today. Technologies* 3 (3), 261
Schmitz, K. and Schepers, U. (2004) Polyamides as artificial transcription factors: Novel tools for molecular medicine? *Angew. Chem. Int. Edit.* 43 (19), 2472-2475
Vázquez, M.E. et al. (2003) From transcription factors to designed sequence-specific DNA-binding peptides. *Chem. Soc. Rev.* 32 (6), 338-349

Outstanding issues

Although specific DNA recognition is claimed for most artificial transcription factor mimics, critical interpretation of assay results is needed since single target binding in a genome-wide context is yet to be verified.

In spite of the recent progress in structural and biochemical studies, a universal code for the recognition between proteins and nucleic acids via the major groove has yet to be generalized.

Despite the contemporary interest in compounds beyond the Lipinski rule of five, the medicinal applicability will benefit from ongoing miniaturization efforts.

References

1. Darnell, J.E. (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* 2 (10), 740-749
2. Brennan, P. et al. (2008) Targeting transcription factors for therapeutic benefit. *Mol. Biosyst.* 4 (9), 909-919
3. Strausberg, R.L. and Schreiber, S.L. (2003) From knowing to controlling: A path from genomics to drugs using small molecule probes. *Science* 300 (5617), 294-295
4. Reayi, A. and Arya, P. (2005) Natural product-like chemical space: search for chemical dissectors of macromolecular interactions. *Curr. Opin. Chem. Biol.* 9 (3), 240-247
5. Arndt, H.D. (2006) Small molecule modulators of transcription. *Angew. Chem. Int. Edit.* 45 (28), 4552-4560
6. Gewirtz, A.M. (2007) On future's doorstep: RNA interference and the pharmacopeia of tomorrow. *J. Clin. Invest.* 117 (12), 3612-3614
7. Karkare, S. and Bhatnagar, D. (2006) Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino. *Appl. Microbiol. and Biot.* 71 (5), 575-586
8. Boger, D.L. et al. (2003) Solution-phase combinatorial libraries: Modulating cellular signaling by targeting protein-protein or protein-DNA interactions. *Angew. Chem. Int. Edit.* 42 (35), 4138-4176

9. Sinha, N. and Michaud, D. (2007) Recent developments in the chemistry, analysis and control for the manufacture of therapeutic-grade synthetic oligonucleotides. *Curr. Opin. Drug Di. De.* 10 (6), 807-818
10. International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature*, 431, 931-945
11. Reiss, T. (2001) Drug discovery of the future: the implications of the human genome project. *Trends in Biotechnol.* 19 (12), 496-499
12. Altmann, K.H. et al. (2009) The state of the art of chemical biology. *Chembiochem* 10 (1), 16-29
13. Dobson, C.M. (2004) Chemical space and biology. *Nature* 432 (7019), 824-828
14. Rademann, J. (2004) Organic protein chemistry: Drug discovery through the chemical modification of proteins. *Angew. Chem. Int. Edit.* 43 (35), 4554-4556
15. Drews, J. (2006) What's in a number? *Nat. Rev. Drug Discov.* 5 (12), 975-975
16. White, S. et al. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* 391 (6666), 468-471
17. Vazquez, O. et al. (2008) Efficient DNA binding and nuclear uptake by Distamycin derivatives conjugated to octa-arginine Sequences. *Chembiochem* 9 (17), 2822-2829
18. Mapp, A.K. et al. (2000) Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. USA* 97 (8), 3930-3935
19. Bando, T. and Sugiyama, H. (2006) Synthesis and biological properties of sequence-specific DNA-alkylating pyrrole-imidazole polyamides. *Accounts Chem. Res.* 39 (12), 935-944
20. Gniazdowski, M. et al. (2005) Effects of anticancer drugs on transcription factor-DNA interactions. *Expert Opin. Ther. Tar.* 9 (3), 471-489
21. Bremer, R.E. et al. (1998) Inhibition of major-groove-binding proteins by pyrrole-imidazole polyamides with an Arg-Pro-Arg positive patch. *Chem. Biol.* 5 (3), 119-133
22. Moretti, R. and Ansari, A.Z. (2008) Expanding the specificity of DNA targeting by harnessing cooperative assembly. *Biochimie* 90 (7), 1015-1025
23. Vázquez, M.E. et al. (2003) From transcription factors to designed sequence-specific DNA-binding peptides. *Chem. Soc. Rev.* 32 (6), 338-349
24. Sera, T. and Uranga, C. (2002) Rational design of artificial zinc-finger proteins using a nondegenerate recognition code table. *Biochemistry* 41 (22), 7074-7081
25. Gommans, W.M. et al. (2005) Engineering zinc finger protein transcription factors: The therapeutic relevance of switching endogenous gene expression on or off at command. *J. Mol. Biol.* 354 (3), 507-519
26. Klug, A. (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem.* 79 (1), doi: 10.1146/annurev-biochem-010909-095056
27. Jamieson, A.C. et al. (2003) Drug discovery with engineered zinc-finger proteins. *Nat. Rev. Drug Discov.* 2 (5), 361-368
28. Sato, S. et al. (2002) Chemical approaches untangling sequence-specific DNA binding by proteins. *Chem-Eur. J.* 8 (22), 5066-5071
29. Canne, L.E. et al. (1995) Total chemical synthesis of a unique transcription factor-related protein - cMyc-Max. *J. Am. Chem. Soc.* 117 (11), 2998-3007
30. Morii, T. et al. (1993) Sequence-specific DNA-binding by a geometrically constrained peptide dimer. *J. Am. Chem. Soc.* 115 (3), 1150-1151
31. Ponzielli, R. et al. (2005) Cancer therapeutics: Targeting the dark side of Myc. *Eur. J. Cancer* 41 (16), 2485-2501
32. Samuel, S. et al. (1998) In situ cross-linking by cisplatin of nuclear matrix-bound transcription factors to nuclear DNA of human breast cancer cells. *Cancer Res.* 58 (14), 3004-3008
33. Vodovozova, E.L. (2007) Photoaffinity labeling and its application in structural biology. *Biochemistry-Moscow* 72 (1), 1-20
34. Meisenheimer, K.M. and Koch, T.H. (1997) Photocross-linking of nucleic acids to associated proteins. *Crit. Rev. Biochem. Mol.* 32 (2), 101-140
35. Rajski, S.R. and Williams, R.M. (1998) DNA cross-linking agents as antitumor drugs. *Chem. Rev.* 98 (8), 2723-2795

36. Kuznetsova, S.A. et al. (1996) Crosslinking of double-stranded oligonucleotides containing O-methyl-substituted pyrophosphate groups to the HNF1 transcription factor in nuclear cell extract. *Nucleic Acids Res.* 24 (23), 4783-4790
37. Lesage, D. et al. (2003) Specific covalent binding of a NF-kappa B decoy hairpin oligonucleotide targeted to the p50 subunit and induction of apoptosis. *Febs Lett.* 547 (1-3), 115-118
38. Peletskaya, E.N. et al. (2001) Cross-linking of the fingers subdomain of human immunodeficiency virus type 1 reverse transcriptase to template-primer. *J. Virol.* 75 (19), 9435-9445
39. Gritsenko, O.M. et al. (2002) Affinity modification of EcoRII DNA methyltransferase by the dialdehyde-substituted DNA duplexes: Mapping the enzyme region that interacts with DNA. *Nucleos. Nucleot. Nuc.* 21 (11-12), 753-764
40. Guieysse, A.L. et al. (2000) Psoralen adducts induced by triplex-forming oligonucleotides are refractory to repair in HeLa cells. *J. Mol. Biol.* 296 (2), 373-383
41. Pendergrast, P.S. et al. (1992) Determination of the orientation of a DNA-binding motif in a protein DNA complex by photo-cross-linking. *Proc. Natl. Acad. Sci. USA* 89 (21), 10287-10291
42. Suchanek, M. et al. (2005) Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat. Methods* 2 (4), 261-267
43. Lee, H.S. et al. (2009) Protein-DNA photo-crosslinking with a genetically encoded benzophenone-containing amino acid. *Bioorg. Med. Chem. Lett.* 19 (17), 5222-5224
44. Cai, X. et al. (2009) DNA minor groove binders: Back in the groove. *Cancer Treat. Rev.* 35 (5), 437-450
45. Dickinson, L. et al. (2004) Arresting cancer proliferation by small-molecule gene regulation. *Chem. Biol.* 11 (11), 1583-1594
46. Lorusso, D. et al. (2009) Brostallicin (PNU-166196), a new minor groove DNA binder: preclinical and clinical activity. *Expert Opin. Inv. Drug* 18 (12), 1939-1946
47. Rebar, E. et al. (2002) Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat. Med.* 8 (12), 1427-1432
48. Chu, B.C.F. and Orgel, L.E. (1992) Crosslinking transcription factors to their recognition sequences with PtII complexes. *Nucleic Acids Res.* 20 (10), 2497-2502
49. Lukhtanov, E.A. et al. (1997) Sequence and structure dependence of the hybridization-triggered reaction of oligonucleotides bearing conjugated cyclopropapyrroloindole. *J. Am. Chem. Soc.* 119 (27), 6214-6225

Figure 1
Click here to download high resolution image

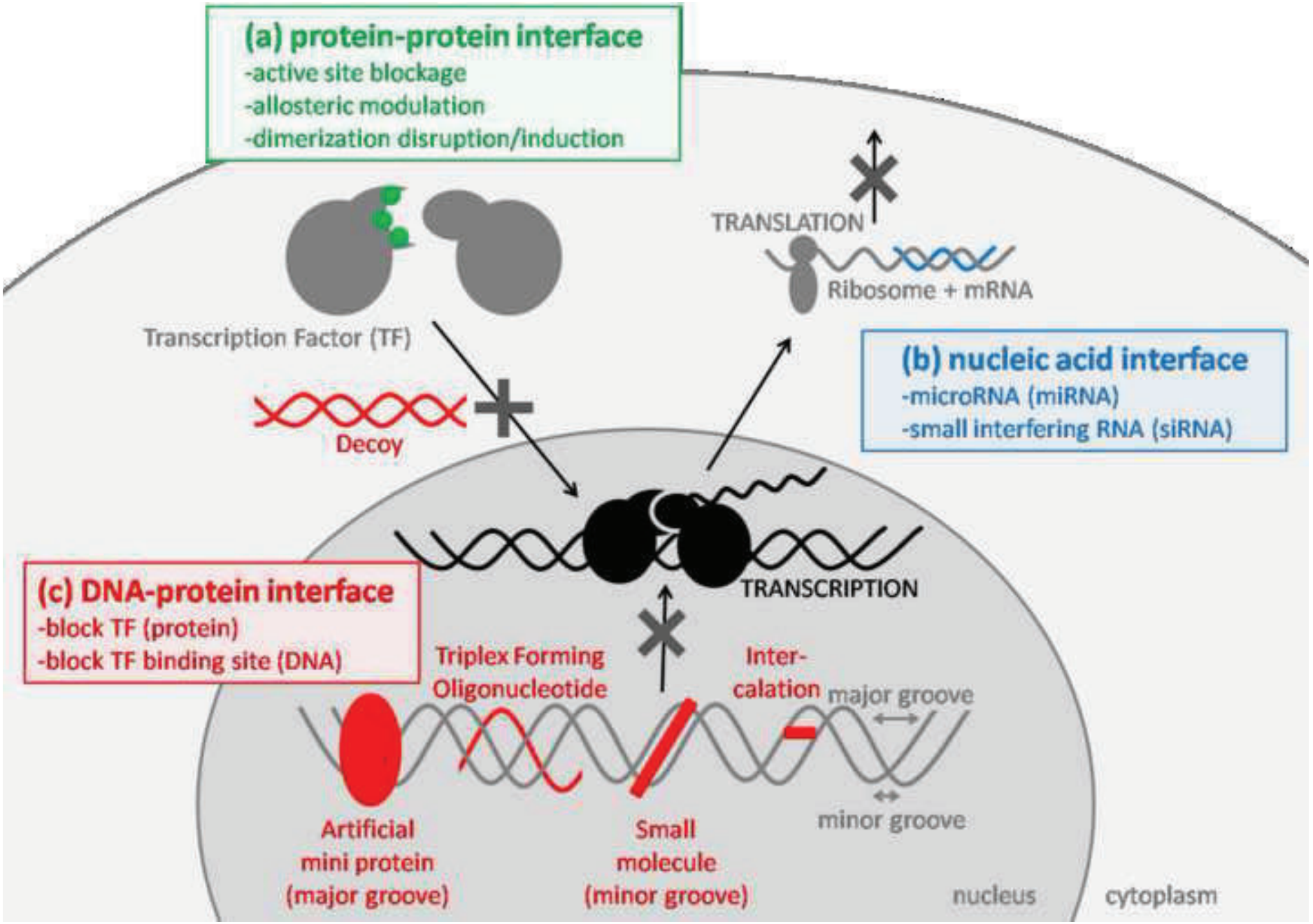

Figure 2

Click here to download high resolution image