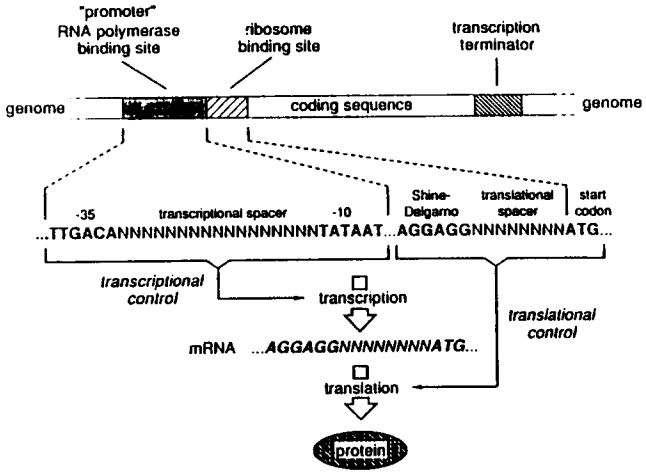
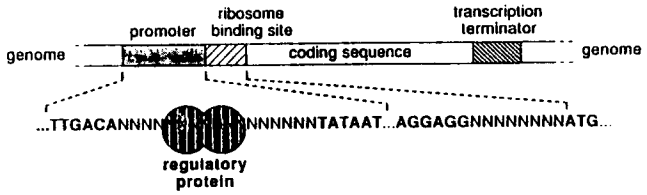
Architecture of a highly expressed bacterial gene:

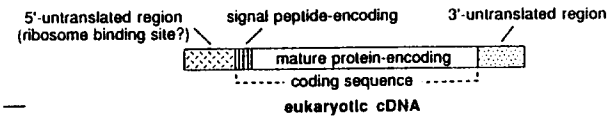


Most genes that one would want to express are nonideal:

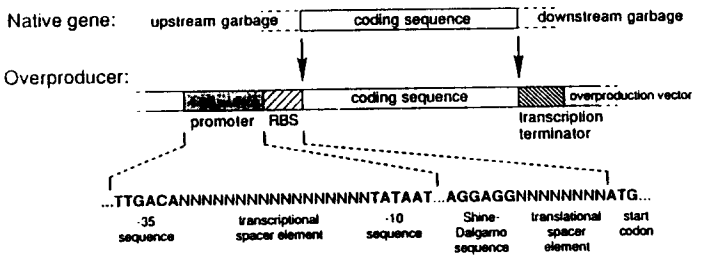
Bacterial genes: expression tightly controlled by sequences, binding of regulatory proteins.



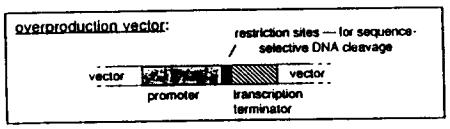
Eukaryotic genes: coding sequence interrupted by introns — bacteria can't splice mRNA precursor; cDNAs lack introns but have foreign flanking and signal sequences:



Engineering protein-overproducing DNA molecules — what operations are required?:



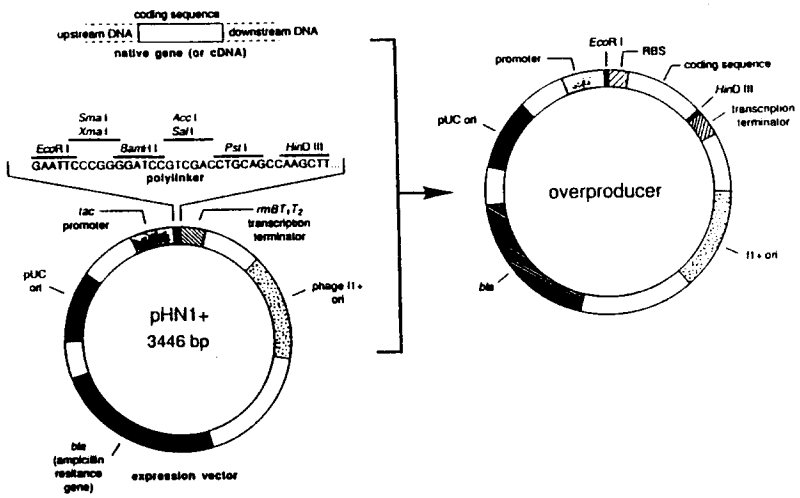
Some sequence elements are supplied by overproduction vector:



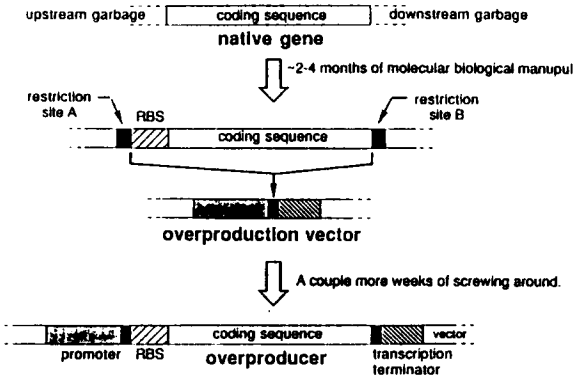
pHN1: H. M. Nash and G. L. Verdine, unpublished results.

Starting materials:

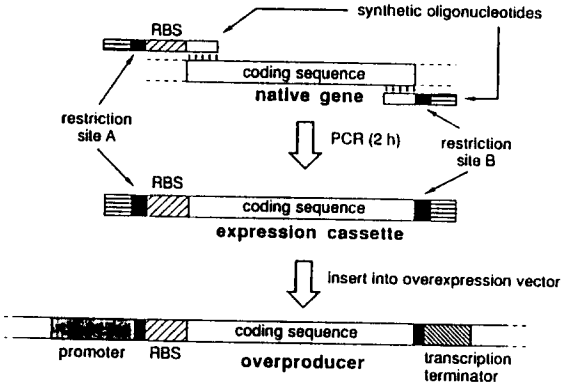
Desired product:



Engineering of protein-overproducing DNA molecules — before the Expression-Cassette Polymerase Chain Reaction:



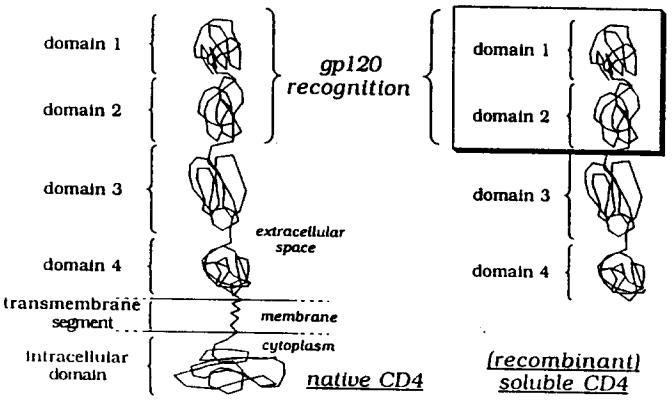
The Expression-Cassette Polymerase Chain Reaction:



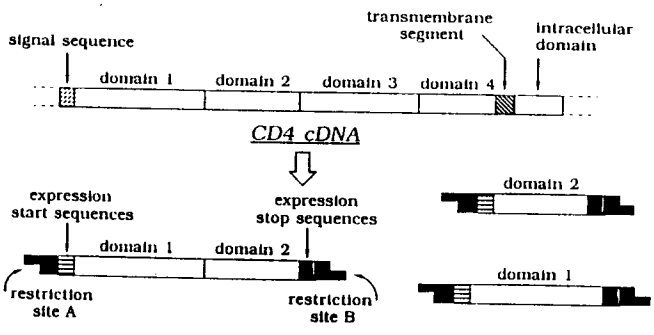
Only need sequence of target DNA — can amplify expression-cassette directly from library.

K. D. MacFerrin, M. P. Terranova, S. L. Schreiber, & G. L. Verdine *PNAS* 87, 1937-1941 (1990).

Redesigning the human CD4 protein:



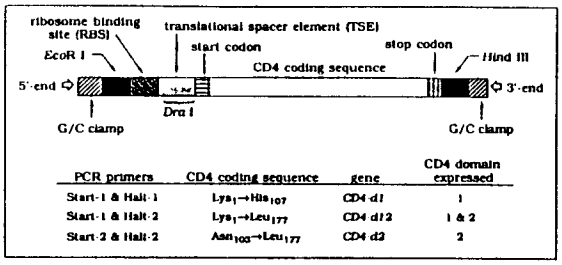
... by redesigning the human CD4 gene:



"Overproduction and Dissection of Proteins by the Expression-Cassette Polymerase Chain Reaction"

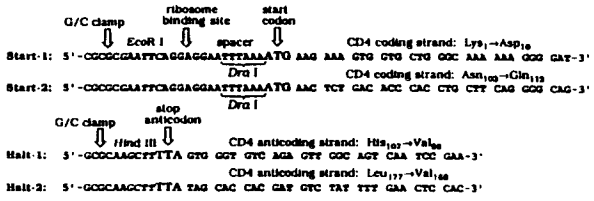
K. D. MacFerrin, M. P. Terranova, S. L. Schreiber & G. L. Verdine *PNAS* 87, 1937-1941 (1990).

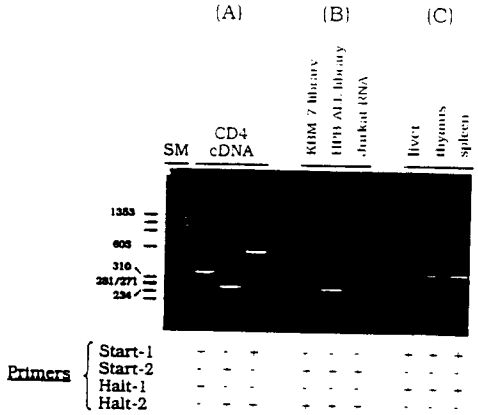
(A) CD4 gene cassettes general structure (Expression Cassette PCR; EC-PCR):



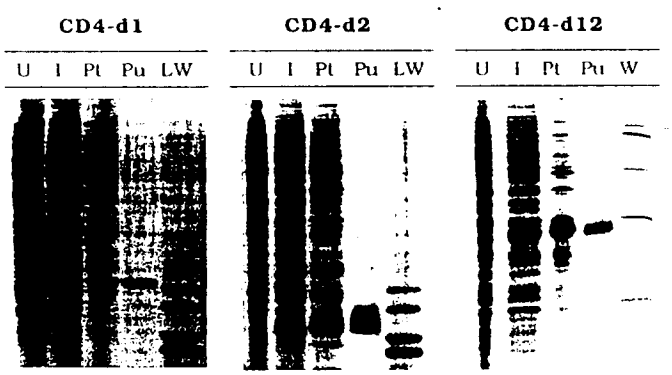
PCR primers	CD4 coding sequence	gene	CD4 domain expressed
Start-1 & Halt-1	$Lys_1 \rightarrow His_{107}$	CD4 d1	1
Start-1 & Halt-2	$Lys_1 \rightarrow Leu_{177}$	CD4 d12	1 & 2
Start-2 & Halt-2	$Asn_{103} \rightarrow Leu_{177}$	CD4 d2	2

(B) Gene redesign primers:



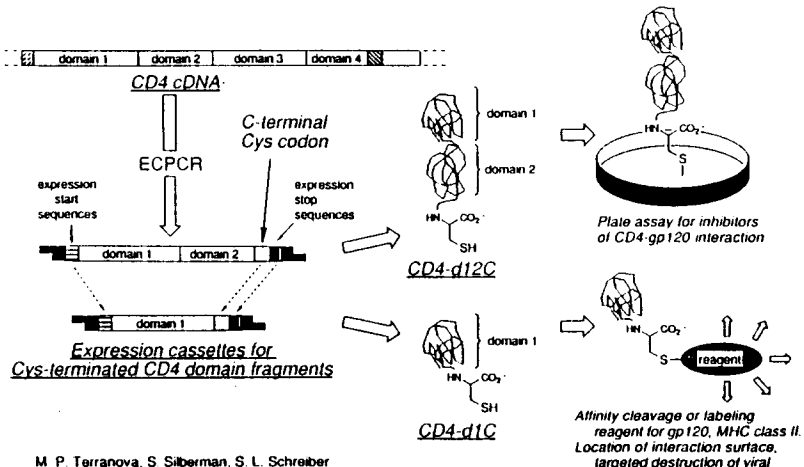


Expression of CD4 Domains in E. coli



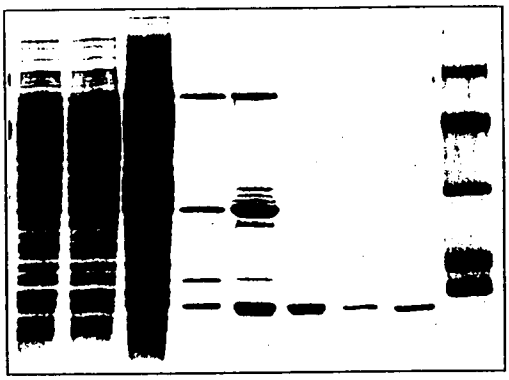
Oligonucleotide and peptide sequencing supports the proposed polypeptide sequence.

Semisynthetic CD4 Derivatives by ECPCR:



M. P. Terranova, S. Silberman, S. L. Schreiber
S. Burakoff, G. Verdine, unpublished results.

Overproduction of a Human FK506-Binding Protein, FKBP



R. F. Standaert, A. Galat, G. L. Verdine & S. L. Schreiber *Nature* 1990, in press.

The Expression-Cassette Polymerase Chain Reaction:

Reduces time required to engineer overproducer from months to days.

Eliminates need for training in molecular biology.

Can be used to truncate proteins or add useful chemical handles.

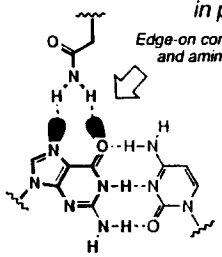
Can be used to direct periplasmic export — to avoid proteolysis, toxicity, folding problems.

Future directions:

Selectable ECPCR

Leaplrogging ECPCR

What is the structural basis for molecular recognition in protein-DNA complexes?

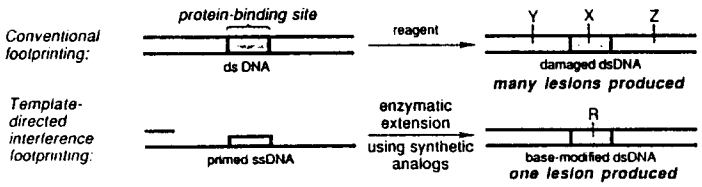


Edge-on contacts between specific DNA bases and amino acids — H-bonding and H-phobic.

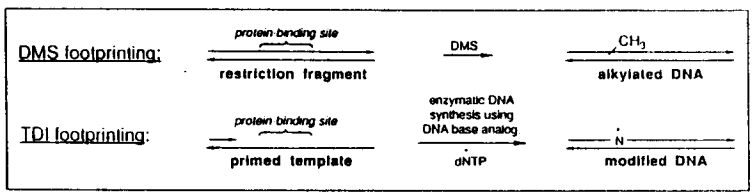
Crystallographic structures:	NMR structures:
*Eco*R I endonuclease-dodecamer	none
434 repressor-operator	
trp repressor-operator	Structural models:
λ repressor-operator	numerous

Over 1000 sequences reported to date.

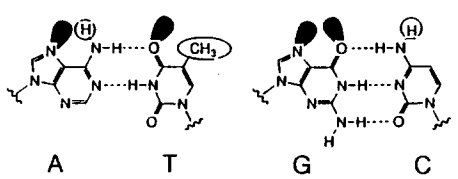
Chemical (footprinting) methods: several good reagents to probe backbone contacts; methods to analyze DNA base-contacts are marred by lack of base-selectivity.

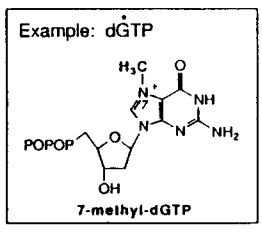


Template-Directed Interference Footprinting:



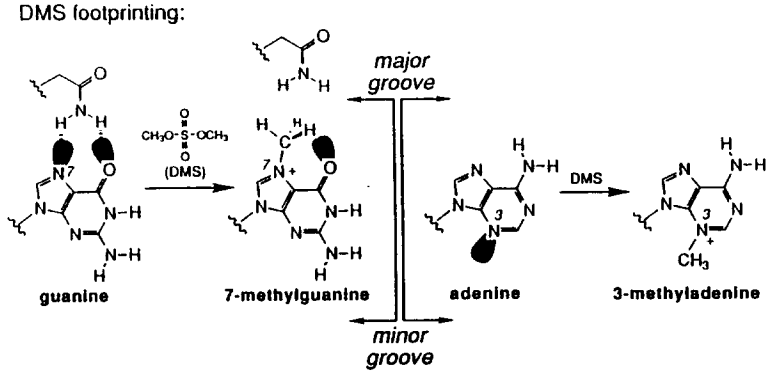
Incorporate DNA base analogs during enzymatic DNA synthesis — each analog has functionality that interferes with recognition elements of its natural counterpart:





Chemistry Underlying Base-Specific Footprinting:

DMS footprinting:



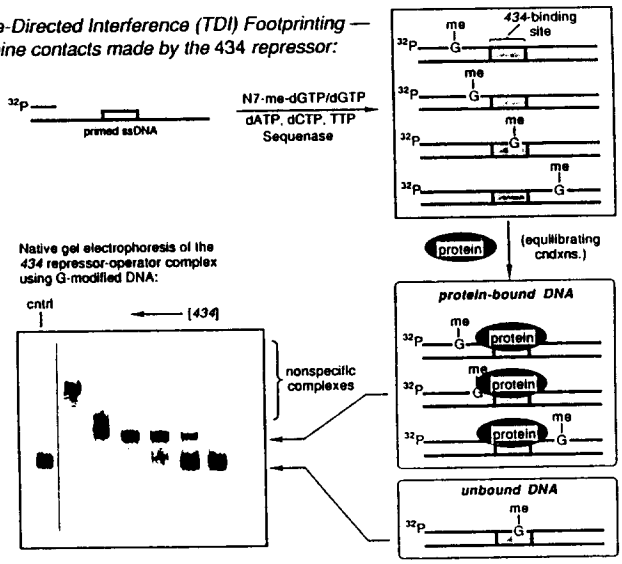
Problems:

DNA is extensively modified — many other adducts produced.

Requires handling of potent carcinogen (DMS).

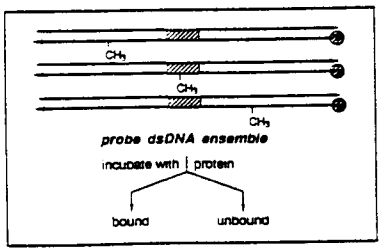
Only G's and A's may be assayed; only G useful. Lack of suitable chemoselective reagents for base modification.

Template-Directed Interference (TDI) Footprinting — guanine contacts made by the 434 repressor:

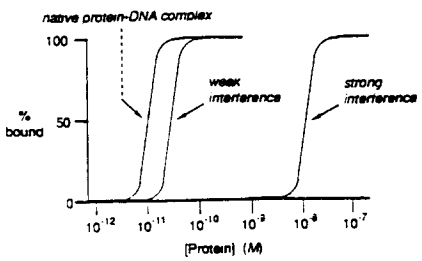


Template-directed interference footprinting of a 434 repressor-operator complex:

Thermodynamic considerations:



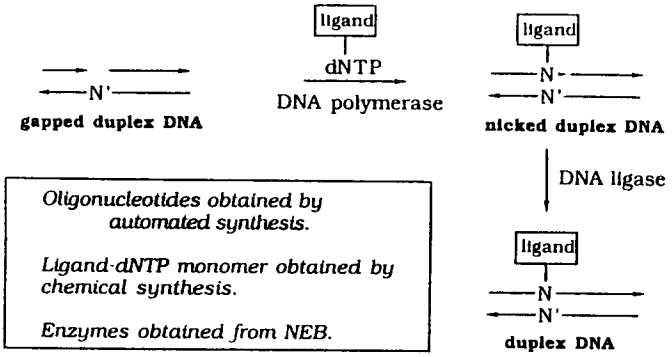
Hypothetical binding profiles:

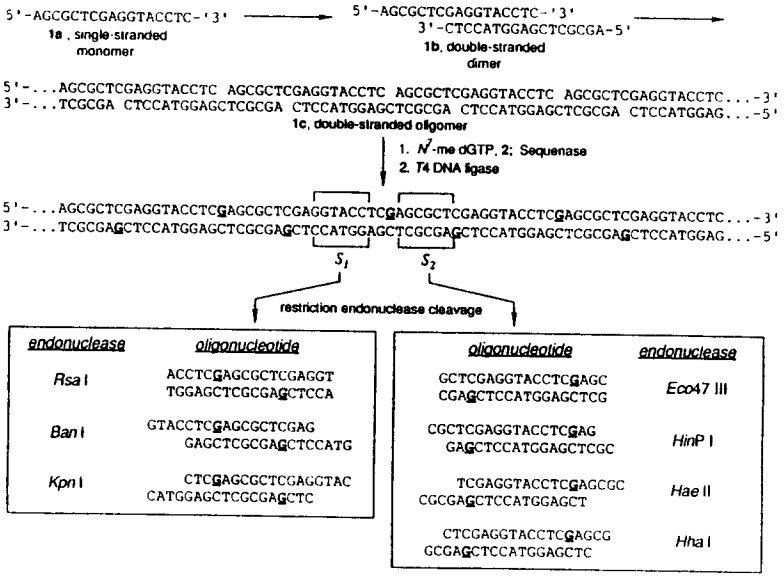


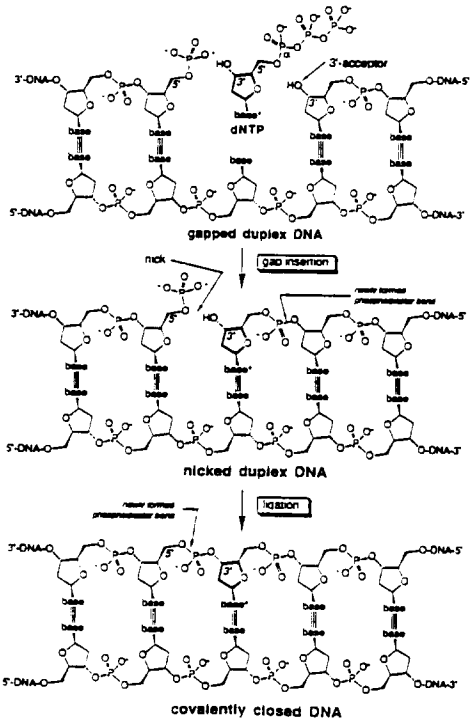
The practice of running interference experiment at a single concentration point should be avoided.
Experiment is best run under K_d sensitive conditions.

Semisynthetic approach to ligand-DNA complexes:

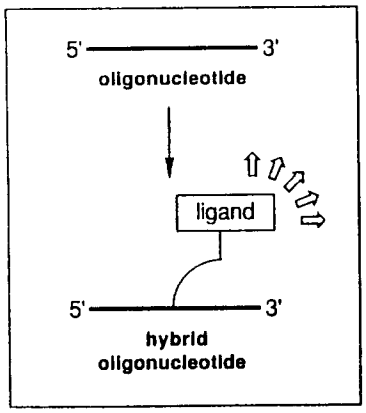
Gap Insertion/Ligation (GIL) Method:







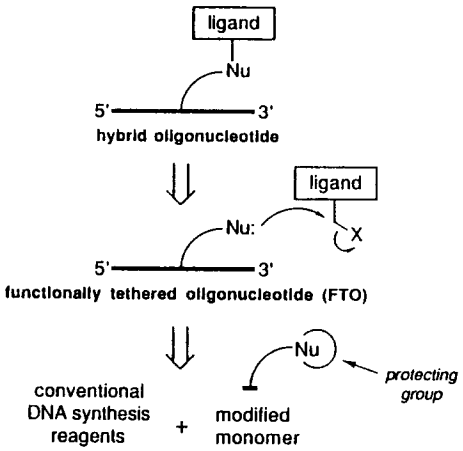
Equipping Oligonucleotides with a Ligand Effector Element:



Ligands (*inter alia*):

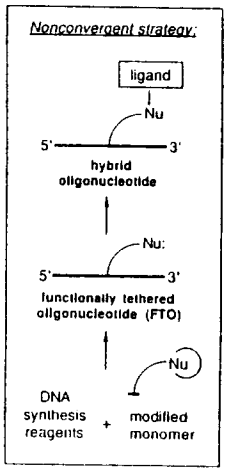
Hydroxyl radical generators
Fluorophores
High-affinity protein-binding elements
Intercalators
DNA/RNA cleaving enzymes
Signal amplification enzymes

Present synthetic strategy for generating hybrid oligonucleotides:



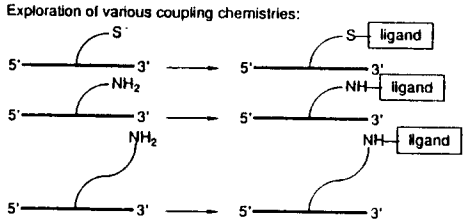
System is committed to specific tether structure at monomer level (nonconvergent).

Nonconvergent FTO syntheses — disadvantages:

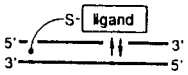


Difficult to permute tether structure.

Exploration of various coupling chemistries:



Use of chelate effect to drive ligand-DNA association — optimization of spacing, geometry.

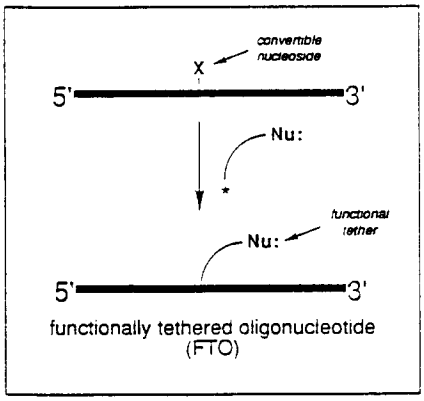


Ligand = *e.g.* peptides, carbohydrates

Highly functionalized tethers may necessitate complex protection/deprotection schemes (e.g., peptides, thioethers).

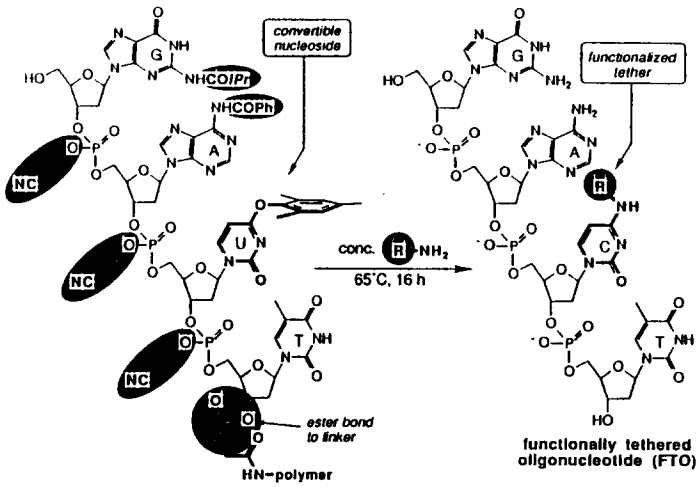
An alternative strategy for the synthesis of functionally tethered oligonucleotides:

The Convertible Nucleoside Approach

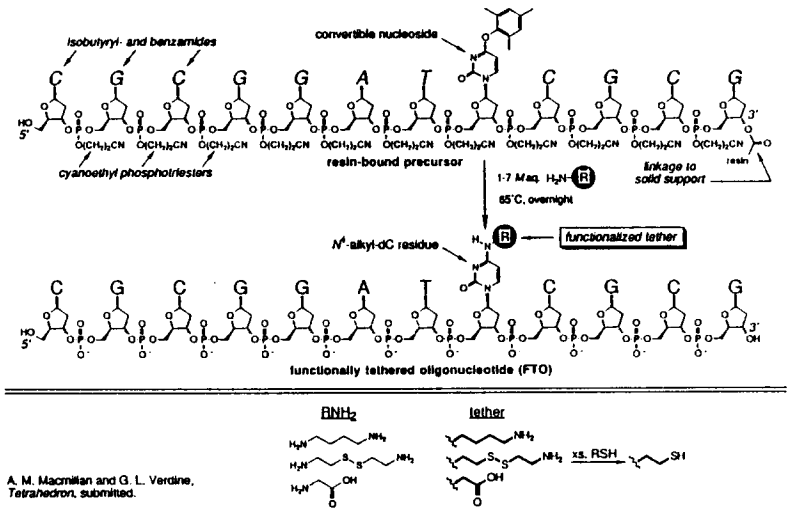


Install tether at the end of the synthesis.

A Convergent Synthesis of Functionally Tethered Oligonucleotides: Deprotection/Aminolysis:

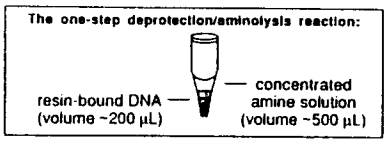


Synthesis of FTOs by One-Step Deprotection/Aminolysis:



A. M. Macmillan and G. L. Verdine, *Tetrahedron*, submitted.

Difficulties of the One-Step Deprotection/Aminolysis Reaction:

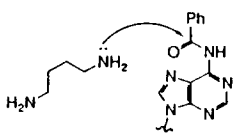


Difficulties:

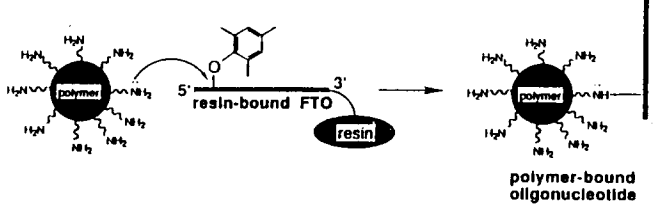
Removal of large excess of amine (centrifugal dialysis).

Prohibitive cost of expensive amines, *e.g.* peptides.

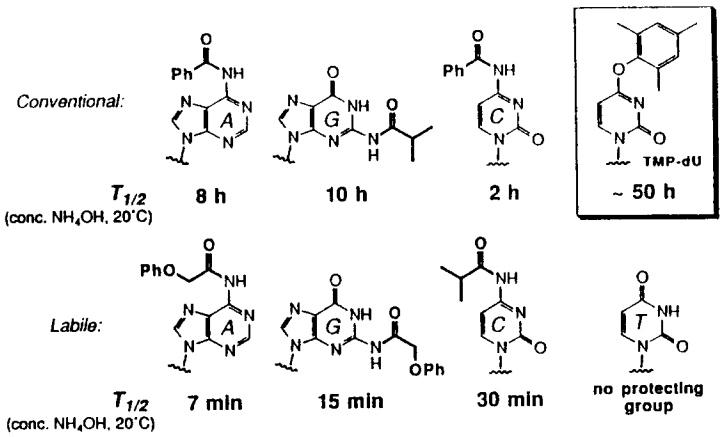
Transfer of base-protecting groups to amine:



Potential problems of two-polymer system:

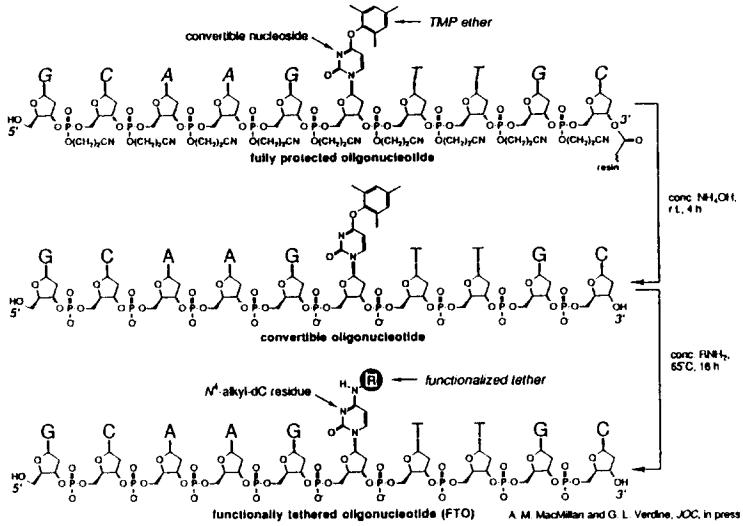


Labile Base-Protecting Groups: Comparison with Conventional and TMP-dU:

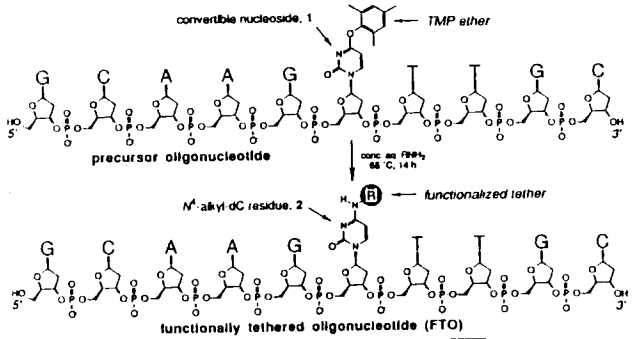


J. C. Schulhof, D. Molko, and R. Teoule, *Nucl. Acids Res.* 15, 397-416 (1987).
Commercially available: Pharmacia

The Two-Step Procedure: Deprotection, then Aminolysis:



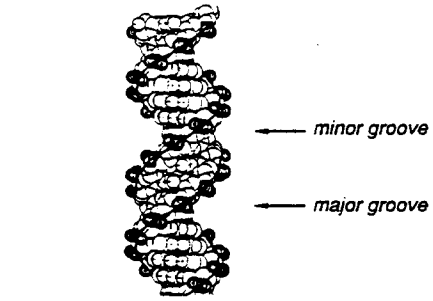
A. M. MacMillan and G. L. Verdine, *JOC*, in press

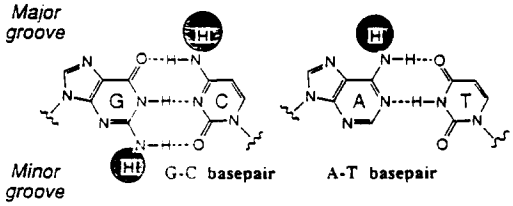


amine	tether	% conversion
H_2NH	H (none)	95
H_2N-CH_3	CH_3	100
H_2N-NH_2	NH_2	100
H_2N-NH_2	NH_2	100
H_2N-OH	OH	100
H_2N-CO_2H	CO_2H	89
$(H_2N$-S-$)_2$	S-S-NH_2	100

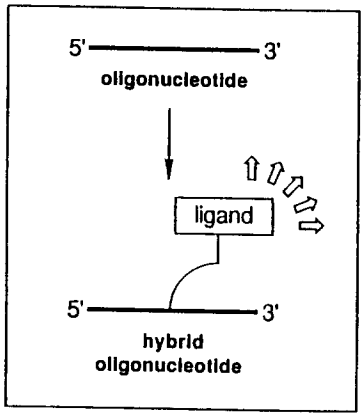
The Convertible Nucleoside Approach to Functionally Tethered Oligonucleotides:

1. Access to wide array of modified oligonucleotides.
2. Minimal modification of conventional synthesis procedure.
3. Maintains 5'- and 3'-ends for enzymatic manipulation.
4. Negligible disruption of DNA structure.
5. Can be extended to allow attachment to 3 of the 4 DNA bases.





Equipping Oligonucleotides with a Ligand Effector Element:



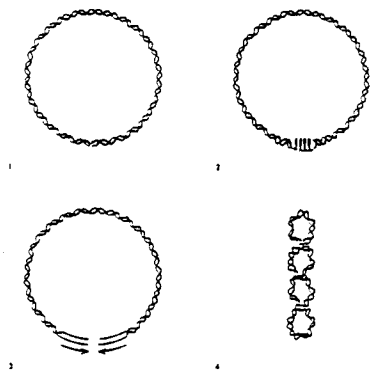
Ligands (*inter alia*):

Hydroxyl radical generators
Fluorophores
High-affinity protein-binding elements
Intercalators
DNA/RNA cleaving enzymes
Signal amplification enzymes

Can this synthetic chemistry be used to illuminate fundamental aspects of DNA structure and function?

Applications of FTO Technology:

Engineering Non-Ground-State DNA Structures



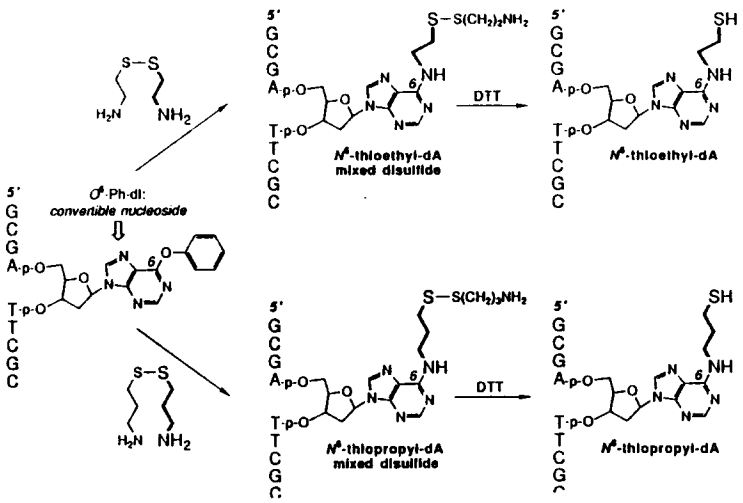
Non-ground-state (NGS) DNA: supercoiled DNA, Z-DNA, cruciform structures, bent & kinked DNA, H-DNA, etc.

Challenge: to generate non-ground-state (NGS) structures in oligonucleotides — will provide substrates for studying structure/function in proteins that specifically recognize NGS DNA.

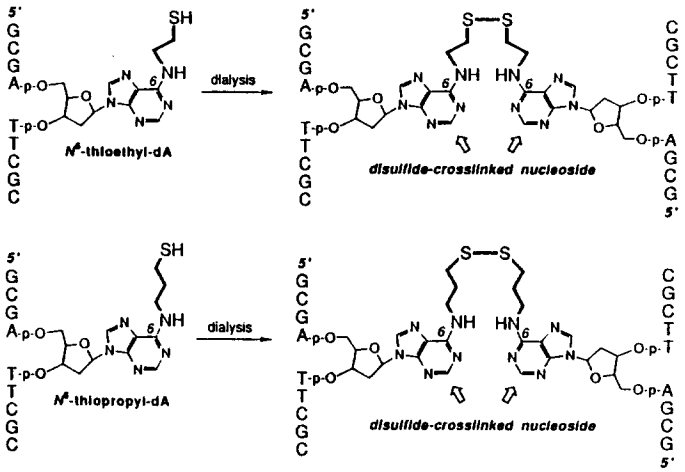
Strategy: drive conformational transitions in DNA by forming disulfide bonds.

Other attractive features: readily formed and cleaved under conditions that do not *per se* affect DNA.

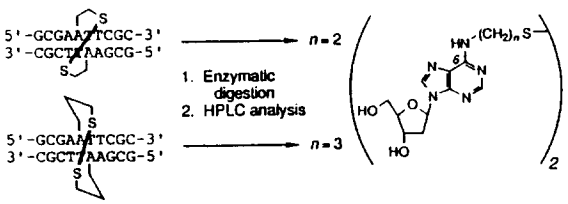
Synthesis of disulfide-crosslinked oligonucleotides by the FTO approach:



Synthesis of disulfide-crosslinked oligonucleotides by the FTO approach:



Characterization of disulfide-crosslinked oligonucleotides:



Native gel electrophoresis: dimer

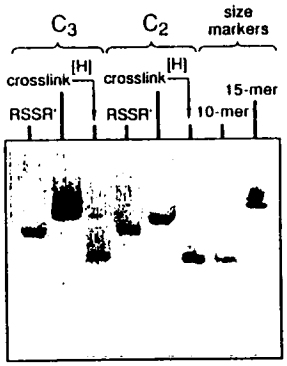
CD: B-DNA

Melting curves: single cooperative transition

oligo	T_m (°C)	Δ
5'-GCGAATTCGC-3' 3'-CGCTTAAGCG-5'	54.25	
		21.0 18.05
C_2 crosslink	75.25	
		2.95
C_3 crosslink	72.3	

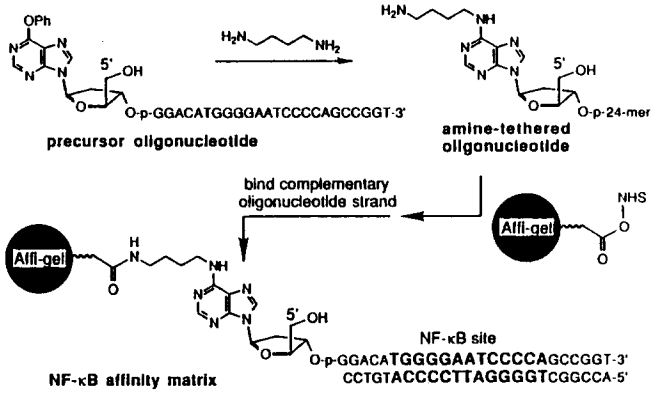
A. E. Ferentz & G. L. Verdine, submitted (*Science*).

Denaturing PAGE Analysis of Disulfide-Crosslinking Reaction:



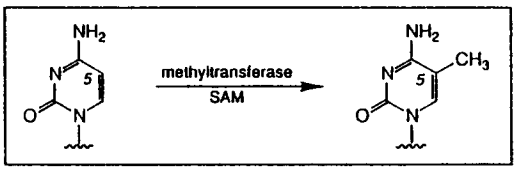
Isolation of NF-κB by Oligonucleotide Affinity Chromatography:

NF-κB: regulates transcription of immunoglobulin genes, activation of HIV enhancer; ultimate effector in signal transduction processes.



Column capacity: ~ 4 μmol (for 30 kDa protein, 120 mg); 10^3 higher than Kadonaga and Tjian.

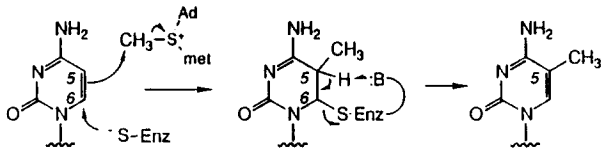
Active Site and Substrate Activation in Cytosine-C5 Methyltransferases:



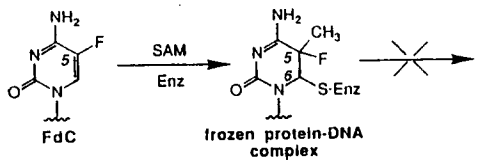
Biological significance:

eukaryotes: cell-type-specific and cell-cycle-specific regulation of gene expression.
prokaryotes: differentiation of self- from non-self-DNA; DNA repair.

Probable mechanism:



Subversion of Catalysis in a Protein-DNA Complex:



Strategy:

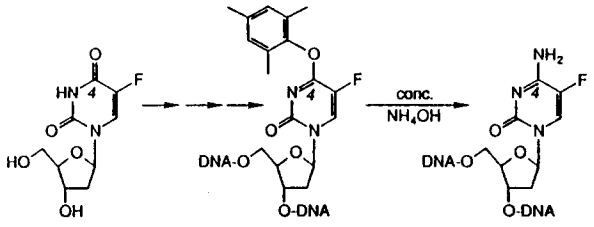
Synthesis of oligonucleotide containing site-specifically implanted FdC residue.

Application of chemical, spectroscopic, and crystallographic probes fro DNA structure.

Determination of active-site nucleophile.

First glimpse into substrate activation in a catalytic DNA-binding protein.

Application of the Convertible Nucleoside Approach to the Synthesis of an FdC-Containing Oligonucleotide:

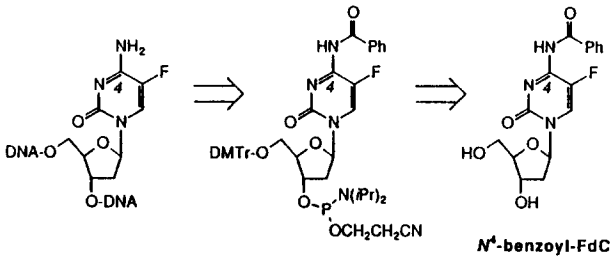


TMP-FdU is completely unaffected upon treatment under DNA conditions.

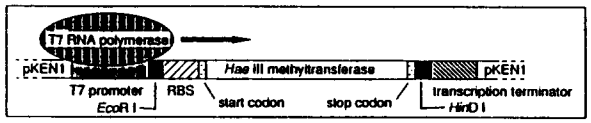
TMP ether acts as a protecting group for the FdC exocyclic amine function.

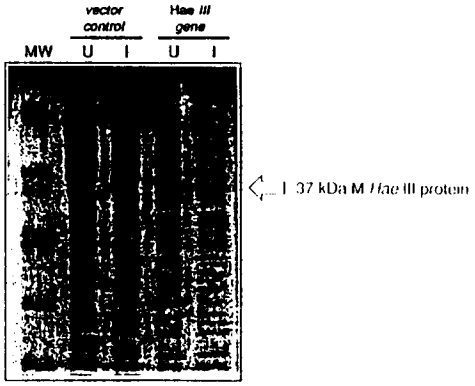
Quantitative yield, conversion.

Site-specific incorporation of FdC into DNA — problems with stability of protected monomer:



Overproduction of Hae III Methyltransferase — the Bacteriophage T7 Promoter:





PKEN1: K. Ezaz-Nikpay, K. Uchino, G. L. Verdine, unpublished; overproduction: L. Chen, G. L. Verdine, unpublished.

Co-workers:

- Lin Chen
  Dr. Peter Eckes
- Khosro Ezaz-Nikpay
- Dr. Dean Farmer
- Ann Ferentz
- Kathleen Hayashibara
  Chris Larson
- Andrew MacMillan
  Dina Romano
  Prof. Yang-Heon Song
- Mike Terranova

Collaborators:

- Stuart Schreiber
  Bob Standaert
  Andrew Galat
  Kurtis MacFerrin
- Stephen Burakoff
  Dr. Sandra Silberman

Phillip Sharp
Dr. David Potter

Ron Hoess
Gerald Crabtree

Camille and Henry Dreyfus Foundation
Searle Scholars Program
Bristol-Myers Squibb
Du Pont
Hoffmann-La Roche
ICI Pharmaceuticals
National Institutes of Health