Selection of Zinc Fingers that Bind Single-Stranded Telomeric DNA in the G-Quadruplex Conformation†

Mark Isalan,‡,§,|| Sachin D Patel,||,⊥ Shankar Balasubramanian,*,||,⊥ and Yen Choo‡,§,||

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, U.K., and MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

Received July 25, 2000; Revised Manuscript Received October 23, 2000

ABSTRACT: There is considerable interest in molecules that bind to telomeric DNA sequences and G-quadruplexes with specificity. Such molecules would be useful to test hypotheses for telomere length regulation, and may have therapeutic potential. The versatility and modular nature of the zinc finger motif makes it an ideal candidate for engineering G-quadruplex-binding proteins. Phage display technology has previously been widely used to screen libraries of zinc fingers for binding to novel duplex DNA sequences. In this study, a three-finger library has been screened for clones that bind to an oligonucleotide containing the human telomeric repeat sequence folded in the G-quadruplex conformation. The selected clones show a strong amino acid consensus, suggesting analogous modes of binding. Binding was found to be both sequence dependent and structure specific. This is the first example of an engineered protein that binds to G-quadruplex DNA, and represents a new type of binding interaction for a zinc finger protein.

The design of proteins that recognize DNA with high affinity and sequence specificity has been a major goal in the study of protein−nucleic acid interactions.[1] Significant progress has been made with the recognition of duplex DNA sequences using the zinc finger motif (reviewed in ref *1*). However, the potential for the zinc finger scaffold to recognize nonduplex nucleic acid structures has not been studied extensively, except for binding to RNA (*2*−*4*). The structure of the zinc finger allows a versatile binding surface to be created by varying a few amino acid positions on the α-helix. Large randomized libraries of zinc fingers have been screened, using phage display, to select fingers that bind to various duplex DNA sequences (*5*−*9*) and to RNA (*2*−*4*).

Telomeres consist of highly conserved DNA repeat sequences, associated with proteins, found at the ends of chromosomes in nearly all eukaryotes. They are widely studied because of their important roles in maintaining chromosome stability and in mediating normal chromosome segregation in mitosis and meiosis (*10*). Telomeric DNA sequences contain characteristic guanine-rich repeats of the form $d(T_{1-3}\text{-}(T/A)\text{-}G_{3-4})_n$ (reviewed in ref *11*). These sequences form four-stranded, guanine quadruplex secondary structures in vitro at physiological salt concentrations (K^+ or Na^+) and it has been proposed that such structures may be of biological significance in vivo. G-Quadruplexes have been extensively characterized in terms of structure, polymorphism, ion selectivity, stability and folding kinetics (reviewed in ref *12*). It has been suggested that intertelomeric G-quadruplexes may determine the correct association of homologous chromosomes in different stages of the cell cycle (*13*−*15*). More recently, it has been suggested that the G-quadruplex conformation of single-stranded telomeric DNA may be important to the mechanism and regulation of telomerase-mediated telomere extension (*16*, *17*). Furthermore, following the discovery of aberant telomere length regulation in various neoplasias, G-quadruplexes are being considered as a molecular target for cancer therapeutics (*18*, *19*).

Several naturally occurring G-quadruplex binding proteins have been reported, (*20*) although none have so far proved ideal candidates for use as diagnostic or therapeutic probes. Most of these, for example a recently reported DNA-binding autoantibody (*21*), have only moderate binding affinities and discriminate weakly between duplex and quadruplex DNA. Similarly, naturally occurring high-affinity telomere-binding proteins are also unable to discriminate these structures. For example, *Saccharomyces cerevisiae* RAP1 (*22*) has distinct but inseparable domains for binding quadruplexes and double stranded DNA.

† This work was supported by BBSRC Grant B08044.

* To whom correspondence should be addressed. Phone: +44 (0)1223 336347. Fax: +44 (0)1223 336913. E-mail: sb10031@cam.ac.uk.

‡ MRC Laboratory of Molecular Biology.

§ Present address: Gendaq Ltd, 1-3 Burtonhole Lane, Mill Hill, London NW7 1AD.

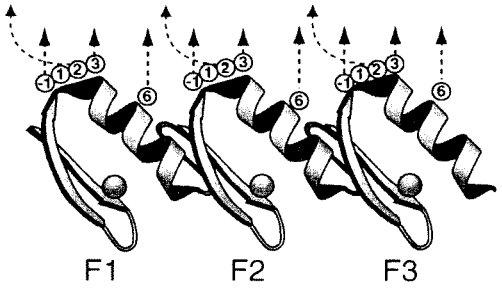
|| The MRC and Department of Chemistry groups contributed equally to this work.

⊥ University of Cambridge.

[1] Abbreviations: DNA, deoxyribonucleic acid; G-quadruplexes, guanine quadruplexes; RNA, ribonucleic acid; K^+, potassium ion; Na^+, sodium ion; Tris, tris(hydroxymethyl)aminomethane; DMS, dimethyl sulphate; Biotin-Htelo, biotin labeled human telomeric DNA $(GGTTAG)_5$; Htelo, human telomeric DNA $(GGTTAG)_5$; KCl, potassium chloride; NaCl, sodium chloride; dsHtelo, duplex human telomeric DNA $(GGTTAG)_5$; F1, F2, and F3, three fingers of Zif268; PCR, polymerase chain reaction; *E. coli*, *Escherichia coli*; $ZnCl_2$, zinc chloride; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; ELISA, enzyme-labeled immunosorbent assay; Gq1*, Gq1-GST fusion protein; GST, glutathione S-transferase; EDTA, ethylenediaminetetraacetic acid; BSA, bovine serum albumin; DTT, dithiothreitol; TB, Tris borate; K_d, dissociation constant; dsDNA, duplex DNA; Asp, aspartic acid; Glu, glutamic acid; K_d^E, ELISA dissociation constant; kDa, kilodalton.

10.1021/bi001728v CCC: $20.00
Published on Web 12/22/2000

a



b

	F1						F2							F3						
-1	1	2	3	4	5	6	-1	1	2	3	4	5	6	-1	1	2	3	4	5	6
A	S	A	A	L	I	A	A	S	A	A	L	S	A	A	K	A	A	R	I	A
D		D	D		T	E	D	R	D	D			E	D	N	D	D		K	E
H		H	H			K	H		H	H			K	H	R	H	H		T	K
N		K	N			N	N		K	N			N	N	S	K	N			N
Q		N	S			Q	Q		N	S			Q	Q		N	S			Q
R		Q	T			R	R		Q	T			R	R		Q	T			R
T		R	V			T	T		R	V			T	T		R	V			T
		S				V			S				V			S				V

FIGURE 1: (a) Schematic representation of the Zif268 DNA-binding domain, showing the three zinc finger helices (F1, F2, and F3). The circled numbers represent the key amino acid residues that interact with duplex DNA (relative to the first position of the α-helix, position +1). (b) Amino acids included in the phage display library used in this study. Amino acid residues in the helical regions of fingers 1−3 (F1−F3) are shown in single letter code, numbered relative to the first helical position (position +1). Note that library construction involved cloning a subset of the possible combinations shown above, although these clones were pre-enriched for DNA-binding potential (see Materials and Methods; Construction of phage display libraries).

We set out to engineer a zinc finger protein that would bind specifically to a telomeric G-quadruplex structure. A phage display library was constructed, based on the well-characterized three-finger DNA-binding domain of Zif268 (*23*, *24*). A zinc finger protein selected from this library binds to single-stranded human telomeric DNA with an affinity comparable to that of natural transcription factors. Data are presented supporting the hypothesis that the DNA in the bound complex is in the G-quadruplex conformation. There is also strong discrimination against the double-stranded form of the same sequence and against single-stranded variant sequences.

MATERIALS AND METHODS

DNA Oligonucleotide Preparation. Synthesized oligonucleotides (Oswel) were purified by fractionation in denaturing polyacrylamide-urea gels, recovered by elution and desalted further using Waters sep-Pack C-18 cartridges with final elution in 25 mM Tris, pH 7.5, as described by Giraldo et al. (*25*).

The sequence 5′-biotin-GGTTAG GGTTAG GGTTAG GGTTAG GGTTAG-3′ (Biotin-Htelo) was prepared for the phage selection experiments, and the unbiotinylated sequence (Htelo) was used for gel mobility shift and DMS protection experiments.

Annealing or Quadruplex Formation of Oligonucleotides for ELISA and Gel Assays. Oligonucleotides were diluted to 10 pmol/μL in 25 mM Tris (pH 7.5) or phosphate-buffered KCl or NaCl (pH 7.5) with cation concentrations ranging from 25 to 150 mM, as specified. Annealing or quadruplex formation was carried out by heating samples to 95 °C, on a thermal heating block, and cooling to 4 °C at a rate of 2 °C/min. The double-stranded DNA (ds Htelo**)** was made by primer extension with the *Klenow* fragment of DNA polymerase.

Construction of Phage Display Library. A zinc finger library was constructed comprising the amino acid framework of wild-type Zif268, but containing randomizations in amino acid positions over all three fingers (Figure 1). Due to the practical size constraints of library cloning ($\sim 10^6-10^7$ transformants), the final library had to be constructed from two complementary sublibraries. Sublibrary-1 contained randomisations in F1 (−1→6) and F2 (−1→3). Conversely, sublibrary-2 contained the randomizations in F2 (3→6) and F3 (−1→6). In both sublibraries, the nonrandomized regions retained the wild-type Zif268 framework.

The genes for each sublibrary were assembled from synthetic DNA oligonucleotides by directional end-to-end ligation using short complementary DNA linkers. The oligonucleotides contained selectively randomized codons, encoding a subset of the 20 amino acids, in the appropriate positions within the zinc fingers. Assembled constructs were amplified by PCR using primers containing *Not*I and *Sfi*I restriction sites, digested with the above endonucleases to produce cloning overhangs, and ligated into similarly prepared vector Fd-Tet-SN (*5*). Electrocompetent *Escherichia coli* TG1 cells were transformed with the recombinant vector and plated onto TYE medium [1.5% (w/v) agar, 1% (w/v) Bactotryptone, 0.5% (w/v) Bactoyeast extract, 0.8% (w/v) NaCl] containing 15 μg/mL tetracycline.

The sublibraries were enriched for DNA-binding members by selecting against random DNA sequences (see *Phage selections*). There are practical limits to the size of phage library that can be used; therefore, we used this limited library in the hope of maximizing the prospect of finding tight, specific binders for the target G-quadruplex, rather than a partially represented random library. The heterogeneous genes from the selected clones were recovered by PCR and recombined via a *Dde*I site, present in the sequence coding for positions +4 and +5 in F2 of both libraries (*26*). Recombinants were then recloned into phage vector, as described above. Ultimately, 3×10^6 selection-enriched library members were obtained, containing randomisations over all three zinc fingers.

Phage Selections. Tetracycline resistant colonies of *E. coli* TG1 cells were transferred from plates into 2xTY medium (16 g/L Bactotryptone, 10 g/L Bactoyeast extract, 5 g/L NaCl) containing 50 μM $ZnCl_2$ and 15 μg/mL tetracycline, and cultured overnight at 30 °C in a shaking incubator. Cleared culture supernatant containing phage particles was obtained by centrifuging at 300 g for 5 min.

For the first rounds of selection, specified quantities (see below) of biotinylated DNA target site were immobilized on streptavidin-coated tubes (Roche) in 50 μL of phosphate buffer (pH 7.4) containing 50 μM $ZnCl_2$ and 150 mM KCl for 30 min at room temperature. Bacterial culture supernatant containing phage was diluted 1:10 in selection buffer (phosphate buffer, pH 7.4 with 150 mM KCl) containing 50 μM $ZnCl_2$, 2% (w/v) fatfree dried milk (Marvel), 1% (v/v) Tween, 20 μg/mL sonicated salmon sperm DNA, and 1 mL was applied to each tube. After 1 h at 20 °C, the tubes were emptied and washed 20 times with selection buffer containing 50 μM $ZnCl_2$, 2% (w/v) fatfree dried milk (Marvel), and 1% (v/v) Tween. Retained phage were eluted in 0.1 M triethylamine and neutralized with an equal volume of 1 M Tris-HCl (pH 7.4). Logarithmic-phase *E. coli* TG1 were infected with eluted phage, and cultured overnight at 30 °C in 2×TY medium containing 50 μM $ZnCl_2$ and 15 μg/mL tetracycline, to amplify phage for subsequent rounds of selection.

For enrichment of the sublibraries 1 and 2, 50 pmol of biotinylated semi-random duplex DNA sequences(5′-tata-NNNNNNNggcgtgtcacagtcagcttcaacgtc-3′ and 5′-tatgtgcg-GNNNNNNNtcacagtcagtccacacgtc-3′) were used in selection round 1. These amounts were reduced to 20 and 10 pmol in rounds 2 and 3.

For selections against Biotin-Htelo, using the full recombined library, 100 pmol of the preannealed oligonucleotide was immobilized on streptavidin-coated tubes in the first round. In rounds 2 and 3, selection pressure was increased by reducing the amount of target site to 50 pmol and 1 pmol, respectively. In these rounds, 50 pmol of duplex and 50 pmol of single-stranded competitor oligonucleotides were also added of the sequence: 5′-TATANNNNNNNNNNNNNT-CACAGTCAGTCCACACGTC-3′. After three rounds of selection, *E. coli* TG1 infected with selected phage were plated. Individual colonies were picked and used to prepare phage for ELISA assays and DNA sequencing.

ELISA Assay for Binding Studies. The phage ELISA is adapted from previous assays (*5*). 5′-Biotinylated DNA sites were added to streptavidin-coated ELISA wells (Boehringer-Mannheim) in 50 mM potassium phosphate buffer (pH 7.5) containing 100 mM potassium chloride and 50 μM Zinc chloride (K/Zn buffer). Phage solution [overnight bacterial culture supernatant diluted 2:10 in K/Zn buffer containing 2% (w/v) fat-free dried milk (Marvel), 1% (v/v) Tween, and 20 μg/mL sonicated salmon sperm DNA] was applied to each well (50 μL/well). The phage were allowed to bind for 1 h at 20 °C. Unbound phage were removed by washing six times with K/Zn buffer containing 1% (v/v) Tween, and then three times with K/Zn buffer. Bound phage were detected by ELISA using horseradish peroxidase-conjugated anti-M13 IgG (Pharmacia Biotech), and the colorimetric signal was quantified using BIO KINETICS READER EL 340 (Bio-Tek Instruments).

Dimethyl Sulfate Protection Assay Htelo and Htelo-Gq1 Zinc Finger Complex. Htelo was 5′-labeled with ^{32}P and was denatured by heating at 95 °C for 10 min. Annealing or quadruplex forming reactions were carried out as described above, in 50 mM Tris-HCl buffer with or without 150 mM potassium. DMS protection was carried out as described by Maxam and Gilbert (*27*). One microliter of dimethyl sulfate (DMS) was added to 20 pmol of annealed *Htelo*, at 4 °C, in 200 μL of appropriate buffer. The mixture was incubated at 20 °C for 5 min. Reactions were stopped by adding $^1/_4$ volume of stop buffer containing 1 M β-mercaptoethanol and 1.5 M sodium acetate, pH 7.0. The reaction products were ethanol precipitated twice and treated with 100 μL of 1 M piperidine at 90 °C for 30 min. The cleaved products were resolved on a 20% denaturing urea−polyacrylamide gel.

For DMS footprinting of the Htelo-Gq1 zinc finger complex, the procedure described above was adapted: 2 μL of DMS were added to 0.2 pmol of annealed Htelo, in the absence or presence of 200 nM purified Gq1* (see below), in 200 μL of the appropriate buffer, containing 1 μg/mL calf thymus DNA. Reactions were carried out for 10 min at 20 °C, after which the procedure continued as described above.

Generation and Expression of Gq1: GST-Zinc Finger Fusion Proteins.* Zinc finger genes were amplified by PCR, using 1 μL of overnight bacterial culture supernatant (containing phage) as a template. The primers introduced *Bam*HI sites for ligation into vector pGEX-3X (Amersham-Pharmacia). The resulting construct (Gq1*), coding for GST fused in frame with C-terminal zinc fingers, was cloned in *E. coli* TG1 and verified by DNA sequencing. Expression of fusion protein was then carried out in *E. coli* BL21 DE3. Gq1* was purified from bacterial lysates by affinity chromatography using Glutathione Sepharose 4 Fast Flow (Pharmacia Biotech).

Gel Mobility Shift Assay. Binding reactions were performed in a final volume of 10 μL, using 10 fmol of labeled oligonucleotide and various amounts of purified Gq1* in binding buffer: 20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT, 6% glycerol, 100 μg/mL BSA, 1 μg/mL calf thymus DNA, and 50 μM $ZnCl_2$ and KCl to 150 mM. Binding reactions were carried out at room temperature for 1 h. The samples were loaded on a 8% polyacrylamide (acrylamide:bisacrylamide = 33:1) nondenaturing gel. The buffer in the gel and for electrophoresis was 0.5 × TB buffer (*28*). Electrophoresis was performed at 15 V/cm, for 2 h at 4 °C. The gels were exposed in a phosphorimager cassette and imaged (model 425E PhosphorImager; Molecular Dynamics, Inc). The bands were quantified using Imagequant

software. Fraction of DNA that is bound and free is determined after normalization by summing the total number of counts in each lane (*29*). To minimize any error due to perturbation of the equilibrium under electrophoretic conditions, the fraction of free DNA is measured at various protein concentrations rather than measuring the amount of complex formed (*29*−*32*). The data are plotted as Ø [1 − (fraction of free DNA)] vs protein concentration to determine the K_d, which is equal to the protein concentration at which half the free DNA is bound. Equilibrium dissociation constants (K_d) were extracted by nonlinear regression using the program Origin 4.1 and the following equation (*33*)

$$Ø = [P]/\{K_d + [P]\}$$

where Ø denotes the fractional saturation of DNA (i.e., fraction of DNA complexed with the protein). [P] represents the protein concentration in the experiment. Ø and [P] were inputs to the nonlinear regression; K_d was an unconstrained output.

DNA Sequence Analysis. The coding sequence of individual zinc finger clones was amplified by PCR from phage samples. PCR products were sequenced manually using Thermo Sequenase cycle sequencing (Amersham Life Science).

RESULTS AND DISCUSSION

Design of Phage Display Library. A zinc finger phage display library was used to select members that bind human telomeric DNA sequences, under conditions that promote G-quadruplex formation. The library comprised zinc fingers with selectively randomized residues, biased for dsDNA binding potential (*5*, *26*). To maximize the chance of finding proteins that bind the telomeric G-quadruplex, we used this three finger domain library rather than two or one finger. Similar libraries have been extensively characterized, both biochemically and structurally, but only in their capacity to bind duplex DNA sequences in the major groove (*1*, *34*). Owing to practical cloning restraints, the library could not represent all combinatorial variants of the natural amino acid residues as shown in Figure 1b. Consequently, two complete sublibraries were constructed and enriched for DNA-binding potential by selection against randomized dsDNA sequences (see Materials and Methods). The resulting clones were recombined in vitro to produce a single library containing randomizations over all three fingers.

Analysis of G-Quadruplex Structures Formed by Human Telomeric Sequences under Conditions Used for Phage Selection. It has been previously reported that the human telomeric sequence (5′-GTTAGG-3′)$_n$ forms G-quadruplex structures in vitro (*35*−*37*). The five repeat telomeric oligonucleotide sequence (5′-GTTAGG-3′)$_5$ (Htelo) was used as the ligand for affinity selection of phage. To confirm that Htelo was folded into a G-quadruplex, in the presence of sodium and potassium ions, a dimethyl sulfate (DMS) protection assay was carried out (*14*). G-Quadruplex formation requires Hoogsteen-type base pairing of guanines which protects N-7 of guanine against methylation on exposure to the potent methylating agent DMS. Subsequent cleavage of the DNA backbone at methylated guanines can be mediated by heating in aqueous piperidine (*27*). The gel pattern in Figure 2a clearly shows that potential quadruplex-forming

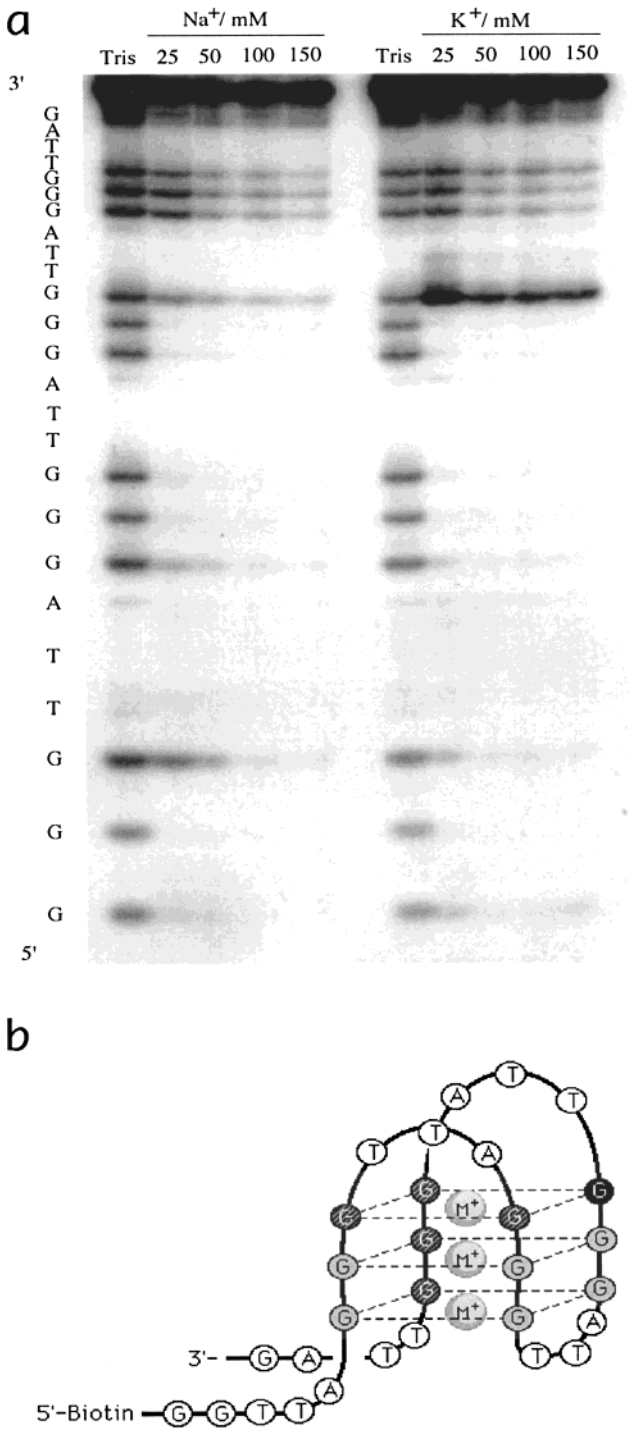


FIGURE 2: (a) DMS methylation protection analysis of Htelo. End-labeled [^{32}P]Htelo was annealed in KCl or NaCl at the indicated concentrations. Each sample was incubated with DMS for 5 min and then cleaved with piperidine. Methylation protection patterns, indicative of G-quadruplex formation, appear after resolution of the cleaved fragments on a 20% polyacrylamide gel. The Tris control lane indicates the reference methylation cleavage pattern of Htelo in the absence of Na^+ or K^+. (b) Schematic representation of an intramolecular antiparallel G-quadruplex formed by Htelo. Guanines in the G-quartet core are labeled in shaded circles with darker shading indicating a relatively higher amount of cleavage, as observed in the DMS methylation protection analysis. Note that the structure shown is only one possible isoform and that another "semi-parallel" conformation may be the most stable form of Htelo (W. Sundquist, personal communication). The alternative form comprises a pair of parallel "up" strands, facing a pair of parallel "down" strands created by "crossing-over" of the two top "TTA" sequences in the figure.

	F1							F2							F3						
PROTEIN	-1	1	2	3	4	5	6	-1	1	2	3	4	5	6	-1	1	2	3	4	5	6
Gq1	D	S	A	H	L	T	R	D	R	S	D	L	S	E	R	S	D	H	R	I	E
Gq2	R	S	D	H	L	I	N	D	R	A	D	L	S	E	T	S	S	H	R	T	N
Gq3	D	S	A	H	L	T	R	D	R	D	H	L	S	E	T	S	S	H	R	T	N
Gq4	T	S	H	H	L	I	Q	D	R	A	D	L	S	E	H	Q	H	Y	R	T	N
Zif268	R	S	D	E	L	T	R	R	S	D	H	L	T	T	R	S	D	E	R	K	R

FIGURE 3: Peptide sequences of the zinc finger helical domains of the four proteins Gq1−4, obtained after three rounds of selection. Amino acid residues in fingers 1−3 (F1−F3) are shown in single letter code, numbered relative to the first helical position (position +1). The zinc finger helices of the wild-type Zif268 DNA-binding domain are also shown for comparison.

guanines in Htelo are almost completely protected from cleavage, at K^+ or Na^+ concentrations above 100 mM, as compared to a Tris-HCl buffer control. Nondenaturing gels (data not shown) confirmed that these folded forms were of a single species and therefore likely to be antiparallel intramolecular G-quadruplexes similar to the structure illustrated in Figure 2b. Intermolecular G-quadruplexes were not observed under these conditions probably due to slow folding kinetics and the low concentrations of DNA used (*38*).

Library Screening for Zinc Finger Phage That Bind the Human Telomeric Sequence in the G-Quadruplex Conformation. The 3-finger phage library was screened with 5′-biotin-$(GGTTAG)_5$ (Biotin-Htelo), which had been annealed in a phosphate-buffered solution containing 150 mM potassium ions then immobilized on streptavidin-coated tubes. These salt conditions were maintained throughout the selection protocol to maintain the structural integrity of the G-quadruplex. After three rounds of selection, zinc finger clones were individually screened for binding to immobilized Biotin-Htelo by a phage ELISA assay (*5*). Four different zinc finger clones were isolated that bind to the target. Control binding assays confirmed that neither the vector phage nor Zif268-phage was able to bind to Biotin-Htelo (data not shown).

The four binding clones (Gq1−4) were sequenced, and the aligned sequences are shown in Figure 3. The clones have a significant degree of sequence similarity, which suggests analogous modes of binding for each clone. The sequence composition of the zinc finger helices from Zif268 is also shown for comparison in Figure 3. The symmetric charge distributions of the selected zinc fingers are very different to that of Zif268. It is interesting to note that finger 2 (F2) of Gq1−4 have each selected negatively charged acidic side chains (Asp or Glu) particularly in positions labeled −1, 3, and 6 (Figure 3). This pattern is unusual for DNA-binding zinc fingers as negatively charged residues are expected to repel the surface of the phosphodiester backbone. It is possible, however, that these acidic residues interact with guanine−NH groups which line all four grooves of an antiparallel G-quadruplex. There has also been some selection in fingers 1 and 3, most notably histidines at position 3 in each finger, and serine in position 1, finger 3, and arginine in position 6, finger 3 (Figure 3). It is possible that the aromatic side chains of histidine may undergo aromatic stacking interactions with the terminal tetrads of the G-quadruplex. The basic residues (Arg/His) on the selected

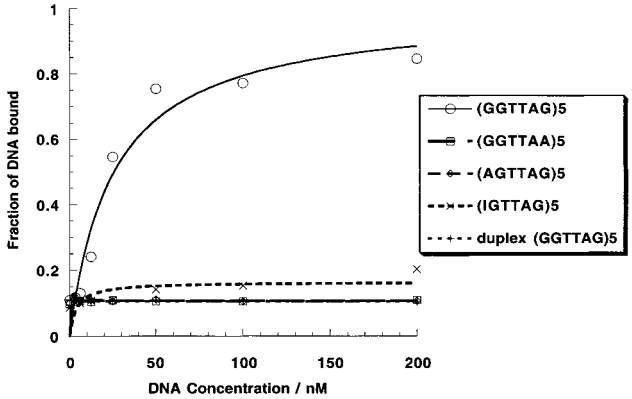


FIGURE 4: Apparent equilibrium binding curves for the protein Gq1 binding to single-stranded DNA sequences, and to the Htelo duplex sequence, as measured by phage ELISA. All ELISA procedures were carried out in the presence of 150 mM K^+, to stabilize G-quadruplexes.

portions of fingers 1 and 3 may also be contributing to binding through phosphate contacts with the quadruplex phosphodiester chains. The observed selection suggests that all three fingers contribute to binding which justifies the initial choice of a 3-finger library rather than 1 or 2.

Analysis of Binding Properties of Selected Zinc Fingers. Subsequent characterization of the binding properties of Gq1−4 showed that they behave very similarly, thus only one clone, Gq1, was used to study the binding specificity in more detail. Phage ELISA was performed using analogues of the Biotin-Htelo oligonucleotide, that contained adenine or inosine substitutions for critical guanine residues required in G-quadruplex formation (Figure 4). Although adenine and inosine are structurally related to guanine, both destabilize G-quadruplex formation (*15*). The adenine substitution leads to a hydrogen-bonding arrangement that is incompatible with G-quadruplex formation, while inosine lacks an N-2 exocyclic amino group required for fully stabilizing such structures. Under the ELISA binding assay conditions used (150 mM K^+), Gq1 has an apparent dissociation constant (K_d^E) of 26 nM for Biotin-Htelo (Figure 4). No significant binding of Gq1 was observed for any of the guanine-substituted analogues employed at 200 nM DNA concentration, suggesting Gq1 is highly sequence and/or structure specific. A double-stranded Htelo oligonucleotide ligand was made by DNA polymerase primer extension of the C-rich complementary sequence of Htelo. This complex was also analyzed for binding of G1q by ELISA and exhibited no significant binding. Thus, Gq1 is specific for the single-

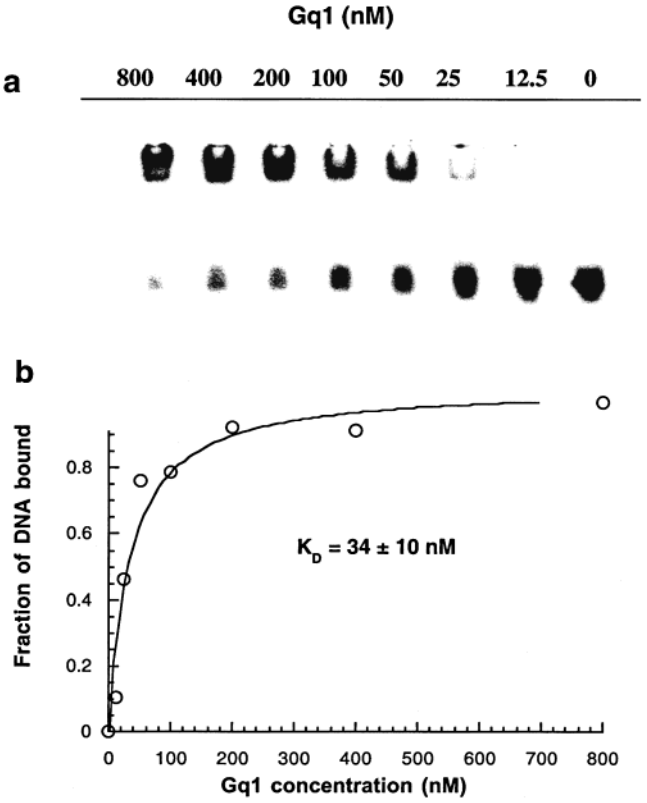


FIGURE 5: (a) Gel mobility shift assay of Gq1* binding to Htelo. The analysis was carried out in 8% nondenaturing polyacrylamide gel at 4 °C. The DNA concentration was fixed at 1 nM while the amount of protein added to the binding reaction is shown in each lane. (b) Apparent equilibrium binding curve obtained by calculating the fraction of Htelo bound at varying Gq1* concentrations (Imagequant software). The binding constant was determined by fitting to the equation $\emptyset = [P]/\{K_d + [P]\}$ as described in the Materials and Methods.

stranded Htelo sequence, under conditions that favor the G-quadruplex structure, but cannot bind the same sequence when present in a double helix.

Electromobility Shift Assays on G-Quadruplex DNAs and DMS Protection of the DNA−Protein Complex. To explore the nature of the Gq1−Htelo complex in more detail, the gene encoding Gq1 was cloned and overexpressed as a glutathione-*S*-transferase fusion protein (Gq1*) in *E. coli* (*39*, *40*). Gq1* was purified from bacterial lysates by affinity chromatography. The eluted protein appeared as a single band of >95% total protein on a protein gel, (data not shown) corresponding to the expected molecular weight of 37 kDa.

The complex between Gq1* and Htelo was studied by nondenaturing gel mobility shift analysis (*29*−*32*). Various concentrations of Gq1* were incubated with 5′ ^{32}P-labeled-Htelo, under conditions (150 mM K^+) that promote and stabilize the G-quadruplex conformation, and the resulting complex was run on an 8% nondenaturing polyacrylamide gel (Figure 5a). This analysis shows the transition of a low molecular weight band to a single, higher molecular weight species upon increasing Gq1* concentration. The gel mobility shift data was fitted to the equation $\emptyset = [P]/\{K_d + [P]\}$ (*33*) to obtain the equilibrium dissociation constant (K_d). The observed dissociation constant was $K_d = 34 \pm 10$ nM (Figure 5b), which was close to the apparent ELISA value (K_d^E) of 26 nM. No DNA-binding was observed for GST protein alone in the absence of Gq1 (data not shown).

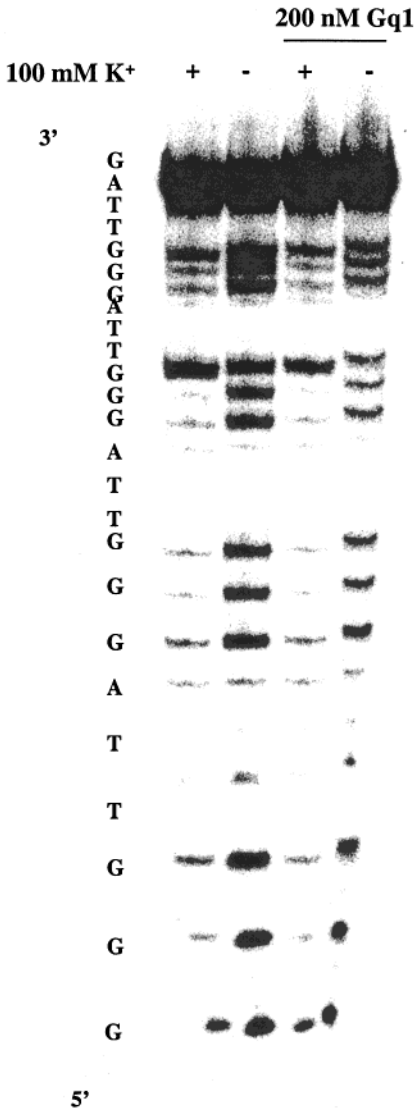


FIGURE 6: DMS methylation protection analysis of Htelo in the presence of Gq1* protein. Htelo was annealed in 100 mM K^+ or 50 mM Tris-HCl, and methylation protection patterns were obtained in the presence or absence of Gq1*, present at a concentration previously determined to give a near full shift (200 nM). The DNA concentration was fixed at 1 nM. Each sample was incubated with DMS for 5 min. Fragments formed by piperidine cleavage of methylated guanines, were resolved on a 20% polyacrylamide gel. Lane 1, methylation pattern of Htelo in the presence of 100 mM K^+; lane 2, reference methylation pattern of Htelo in the absence of K^+; lane 3, methylation pattern of Htelo in the presence of 100 mM K^+ and incubation with 200 nM Gq1*; lane 4, methylation pattern in the absence of K^+, incubated with 200 nM Gq1*.

To determine the conformation of the oligonucleotide in the Gq1*-Htelo complex, DMS protection experiments were carried out on the complex. Using 5′ ^{32}P-labeled-Htelo and buffer containing 100 mM K^+, the concentration of Gq1* was set at 200 nM which is ~6-fold higher than the K_d. These conditions correspond to a near total bandshift (Figure 5a), presumed to represent complete complexation of the DNA. In the absence of Gq1*, a cleavage protection pattern was generated that is both characteristic of G-quadruplex structure and that is dependent on the presence of 100 mM K^+ (Figure 6; lanes 1 and 3). In the presence of Gq1* and 100 mM K^+, there is still significant protection of the critical guanines (Figure 6; lane 3) indicative of G-quadruplex structure. In

the absence of potassium, the protein did not appear to alter the cleavage pattern of Htelo (Figure 6; lane 4). However, it is not known whether the protein was incapable of binding DNA under these conditions, or whether the cleavage pattern merely reflected a relative instability of the complexes formed at low salt concentrations.

In summary, these results strongly suggest that Gq1* is binding Htelo in the G-quadruplex conformation and that this protein has been selected specifically to recognize the structure of a folded G-quadruplex. There is still room for speculation regarding the precise nature of the protein−DNA interactions found in Gq1−Htelo complexes. Critically, an unanswered question is to what degree does Gq1 recognize "unfolded" forms of the Htelo sequence? Unfortunately, it has proved impossible to isolate true "single-stranded" species, even at the low metal cation concentrations used in control lanes in Figures 2 and 6. In such conditions, the DMS cleavage patterns merely indicate the relative destabilizing of G-quadruplex structures and not the complete absence of quadruplexes or "hairpins" from the mixture of species present. Therefore, we anticipate that elucidating the full recognition properties of Gq1 will ultimately require detailed structural studies of the protein−DNA complex.

CONCLUSION AND PERSPECTIVES

Zinc finger proteins have been selected from a phage display library to bind G-quadruplex DNA structures of single-stranded human telomeric sequences with high affinity and selectivity over a duplex DNA comprising the Htelo sequence and its complementary strand. Although further characterization is required to elucidate the molecular details of the mode of binding, the nature of the zinc finger−G-quadruplex interaction is likely to be quite distinct from known zinc finger−duplex DNA interactions. The widespread conservation of G-quadruplex-forming sequences at chromosome ends is intriguing, but it remains to be seen whether they actually play any biological role. The proteins generated in this study may be useful in vitro or in vivo molecular probes to explore the structure and maintenance of telomeres.

ACKNOWLEDGMENT

We thank Aaron Klug for suggesting useful experiments and for his comments on the manuscript.

REFERENCES

1. Choo, Y., and Klug, A. (1997) *Curr. Opin. Struct. Biol. 7*, 117−125.
2. Friesen, W. J., and Darby, M. K. (1997) *J. Biol. Chem. 272*, 10994−10997.
3. Friesen, W. J., and Darby, M. K. (1997) *Nat. Struct. Biol. 5*, 543−546.
4. Blancafort, P., Steinberg, S. V., Paquin, B., Klinck, R., Scott, J. K., and Cedergren, R. (1999) *Chem. Biol. 6*, 585−597.
5. Choo, Y., and Klug, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11163−11167.
6. Greisman, H. A., and Pabo, C. O. (1997) *Science 275*, 657−661.
7. Jamieson, A. C., Kim, S.-H., and Wells, J. A. (1994) *Biochemistry 33*, 5689−5695.
8. Wu, H., Yang, W.-P., and Barbas, C. F., III. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 344−348.
9. Isalan, M., Klug, A., and Choo, Y. (1998) *Biochemistry 37*, 12026−12033.
10. Rhodes, D., and Giraldo, R. (1995) *Curr. Opin. Struct. Biol. 5*, 311−322.
11. Blackburn, E. H., and Szostak, J. W. (1984) *Ann. Rev.Biochem. 53*, 163−194.
12. Williamson, J. R. (1994) *Ann. Rev. Biophys. Biomol. Struct. 23*, 703−730.
13. Sen, D., and Gilbert, W. (1988) *Nature 334*, 364−366.
14. Sundquist, W. I., and Klug, A. (1989) *Nature 342*, 825−829.
15. Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989) *Cell 59*, 871−880.
16. Salazar, M., Thompson, B. D., Kerwin, S. M., and Hurley, L. H. (1996) *Biochemistry 35*, 16110−16115.
17. Sun, D., LopezGuajardo, C. C., Quada, J., Hurley, L. H., and VonHoff, D. D. (1999) *Biochemistry 38*, 4037−4044.
18. Sun, D., B., T., Cathers, B. E., Salazar, M., Kerwin, S. M., Trent, J. O., Jenkins, T. C., Neidle, S., and Hurley, L. H. (1997) *J. Med. Chem. 40*, 2113−2116.
19. Perry, P. J., Reszka, A. P., Wood, A. A., Read, M. A., Gowan, S. M., Dosanjh, H. S., Trent, J. O., Jenkins, T. C., Kelland, L. R., and Neidle, S. (1998) *J. Med. Chem. 41*, 4873−4884.
20. Wellinger, R. J., and Sen, D. (1997) *Eur. J. Cancer 33*, 735−749.
21. Brown, B. A., Li, Y. Q., Brown, J. C., Hardin, C. C., Roberts, J. F., Pelsue, S. C., and Shultz, L. D. (1998) *Biochemistry 37*, 16325−16337.
22. Giraldo, R., Suzuki, M., Chapman, L., and Rhodes, D. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7658−7662.
23. ElrodErickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) *Structure 4*, 1171−1180.
24. Pavletich, N. P., and Pabo, C. O. (1991) *Science 252*, 809−817.
25. Giraldo, R., and Rhodes, D. (1994) *EMBO J. 13*, 2411−2420.
26. Isalan, M. D., Klug, A., and Choo, Y. International Patent Application no. WO98/5307.
27. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol. 65*, 499−560.
28. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular cloning. A laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
29. Senear, D. F., and Brenowitz, M. (1991) *J. Biol. Chem. 266*, 13661−13671.
30. Garner, M. M., and Revzin, A. (1981) *Nucleic Acids Res. 9*, 3047−3060.
31. Cann, J. R. (1989) *J. Biol Chem. 264*, 17032−17040.
32. Carey, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 975−979.
33. Gunasekera, A., Ebright, Y. W., and Ebright, R. H. (1992) *J. Biol. Chem. 267*, 14713−14720.
34. Choo, Y., and Isalan, M. D. (2000) *Curr. Opin. Struct. Biol. 10*.
35. Balagurumoorthy, P., Brahmachari, S. K., Mohanty, D., Bansal, M., and Sasisekharan, V. (1992) *Nucleic Acids Res. 20*, 4061−4067.
36. Balagurumoorthy, P., and Brahmachari, S. K. (1994) *J. Biol. Chem. 269*, 21858−21869.
37. Fletcher, T. M., Sun, D. K., Salazar, M., and Hurley, L. H. (1998) *Biochemistry 37*, 5536−5541.
38. Hardin, C. C., Henderson, E., Watson, T., and Prosser, J. K. (1991) *Biochemistry 30*, 4460−4472.
39. Chittenden, T., Livingston, D. M., and Kaelin, W. G. (1991) *Cell 65*, 1073−1082.
40. Smith, D. B., and Johnson, K. S. (1988) *Gene 67*, 31−40.

BI001728V