Exploring the Recognition of Quadruplex DNA by an Engineered Cys_2-His_2 Zinc Finger Protein†

Sylvain Ladame,‡ James A. Schouten,‡ Jose Roldan,§ James E. Redman,‡ Stephen Neidle,§ and Shankar Balasubramanian*,‡

University Chemical Laboratories, University of Cambridge, Lensfield Road, Cambridge CB21EW, U.K., and Cancer Research U.K. Biomolecular Structure Group, The School of Pharmacy, University of London, 29-39 Brunswick Square, London WC1N 1AX, U.K.

Received February 8, 2005; Revised Manuscript Received November 22, 2005

ABSTRACT: We have recently described an engineered zinc finger protein (Gq1) that binds with high specificity to the intramolecular G-quadruplex formed by the human telomeric sequence 5′-$(GGTTAG)_5$-3′, and that inhibits the activity of the enzyme telomerase *in vitro*. Here we report site-directed mutagenesis, biophysical, and molecular modeling studies that provide new insights into quadruplex recognition by the zinc finger scaffold. We show that any one finger of Gq1 can be replaced with the corresponding finger of Zif268, without significant loss of quadruplex affinity or quadruplex versus duplex discrimination. Replacement of two fingers, with one being finger 2, of Gq1 by Zif268 results in significant impairment of quadruplex recognition and loss of discrimination. Molecular modeling suggests that the zinc fingers of Gq1 can bind to the human parallel-stranded quadruplex structure in a stable arrangement, whereas Zif268−quadruplex models show significantly weaker binding energy. Modeling also suggests that an important role of the key protein finger residues in the Gq1−quadruplex complex is to maintain Gq1 in an optimum conformation for quadruplex recognition.

The Cys_2-His_2 zinc finger proteins represent one of the largest classes of proteins encoded in the genomes of all eukaryotes that have been sequenced to date. These proteins typically have multidomain architectures in which individual zinc binding domains are connected by relatively short linker sequences to form arrays of two or more. The biological roles of specific zinc finger proteins have been extensively investigated, and in most cases these proteins function by binding to double-stranded nucleic acids in a sequence specific manner (*1*−*3*). The crystal structure of a complex of the murine transcription factor Zif268 with its double-stranded DNA[1] target sequence (*4*) has provided a structural framework for understanding a wide set of zinc finger protein−DNA interactions. This three zinc finger protein was shown to specifically recognize nine consecutive base pairs, with each zinc finger domain interacting with a triplet of base pairs.

Even prior to the determination of the Zif268 cocrystal structure, mutagenesis studies had demonstrated that it is possible to change the DNA sequence specificity of zinc finger proteins by substituting different amino acids within what is now known as the recognition helix. Various approaches have been used to design zinc finger proteins that recognize one particular DNA sequence and to demonstrate the possible existence of a recognition code. These include extensive mutagenesis experiments (*5*, *6*) and combinatorial selection using phage display (*7*−*9*). Although many points of contact between such engineered proteins and their DNA binding sites are similar to those made by Zif268, some additional interactions were also observed that involve residues at position 1 and 2 (*10*). Moreover, while Zif268 proved to interact preferentially with the primary DNA strand, some recently engineered proteins were found to form approximately the same number of contacts with the primary and secondary strands (Figure 1). These structural results demonstrate that the binding surfaces possible within the zinc finger framework are considerably more versatile than imagined from the first Zif268 structure (*11*).

During the past decade, there has been growing interest in the structure, recognition, and function of four stranded DNA G-quadruplexes. The best-studied example is the human telomeric DNA quadruplex that leads to inhibition of telomere extension by the enzyme telomerase (*12*), whose activity is up-regulated in cancer cells. Other quadruplexes with putative roles in biological mechanisms have also been identified (*13*, *14*). There is considerable interest in understanding the molecular recognition of quadruplexes and the design of quadruplex ligands. We have recently shown that zinc finger proteins could be engineered to recognize G-quadruplex DNA with high selectivity (*15*). The approach

† This work was supported by grants from the BBSRC and Cancer Research U.K.

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

‡ University of Cambridge.

§ University of London.

[1] Abbreviations: DNA, deoxyribonucleic acid; T, thymine base or thymidine nucleoside; A, adenine base or adenosine nucleoside; C, cytosine base or cytidine nucleoside; G, guanine base or guanosine nucleoside; Tris, tris(hydroxymethyl)aminomethane; BSA, bovine serum albumin; DTT, dithiothreitol; K_d, equilibrium dissociation constant; ELISA, enzyme-labeled immunosorbent assay; SPR, surface plasmon resonance.

10.1021/bi050229x CCC: $33.50
Published on Web 01/13/2006

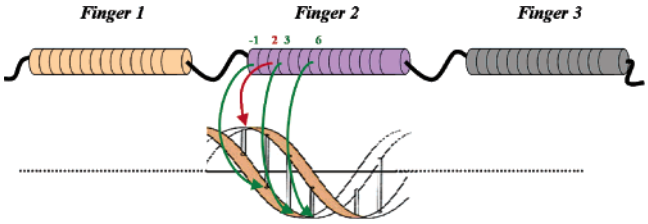

FIGURE 1: Schematic representation of a three zinc finger protein bound to its duplex DNA target. The numbering shown in finger 2 relates to DNA contact residues.

employed a phage display library of zinc finger variants engineered from Zif268 followed by affinity selection using the folded human telomeric G-quadruplex as a ligand. A small family of four zinc finger proteins was selected, all of which showed a significant degree of homology. One quadruplex binding protein, Gq1, was subjected to detailed investigation. Gq1 binds the intramolecular human telomeric quadruplex with a K_d value of 34 ± 10 nM and with a high selectivity for quadruplex as compared to duplex DNA. Gq1 was found to inhibit the action of telomerase *in vitro* with an IC_{50} of 74 ± 11 nM, consistent with the mechanistic hypothesis linking quadruplexes with telomerase (*16*). This clearly expands the scope of DNA sequences and DNA structures for which specific zinc finger-based proteins can be engineered and, furthermore, introduces the possibility of engineering a new class of functional zinc fingers.

We report here a series of studies employing mutagenesis and biophysical analysis in order to gain new insights into the recognition of a quadruplex by the zinc finger protein Gq1. These studies were complemented by molecular modeling studies on the two sets of possible complexes formed between the parallel human telomeric quadruplex and (a) Gq1 and (b) Zif268.

MATERIALS AND METHODS

Site-Directed Mutagenesis of Gq1 and Related Proteins. DNA fragments encoding for Gq1 and Zif268 were cloned into pGEX-3X plasmid (Amersham Pharmacia) as previously reported. Specific mutations were introduced into both Gq1- and Zif268-pGEX plasmids using the Quikchange site-directed mutagenesis kit (Stratagene) following the manufacturer's protocol. Briefly, appropriate complementary synthetic oligonucleotides introducing the desired mutations were used as primers for pseudo-polymerase chain reactions performed in a DNA thermal cycler using Pfu turbo DNA polymerase (Stratagene). The resultant double-stranded DNA was subsequently treated with *Dpn*I, purified by ethanol precipitation, and transformed into supercompetent *E.coli* XL1-Blue cells (Stratagene). Single colonies of transformants were grown overnight in LB and plasmid DNA extracted using a Qiagen plasmid purification kit. The nucleoside sequences of the newly mutated plasmids were confirmed by sequencing (DNA sequencing facility, Biochemistry Dept, Cambridge University). The two zinc finger construct Gq1F1F2 was amplified from Gq1-pGEX using standard polymerase chain reaction (PCR) techniques. The PCR product was digested with *Eco*RI and *Bam*HI to create cohesive ends and was inserted into pGEX-2T (Amersham Pharmacia).

Expression and Purification of GST-Fusion Zinc Finger Proteins. All DNA fragments were cloned into pGEX-3X, an *E.coli* expression vector that produces GST-fusion proteins. The plasmids were introduced into *E.coli* BL21 DE3 cells (Stratagene). Clones were cultured at 37 °C to an OD_{600} between 0.6 and 1.0 before fusion proteins were induced with 0.1 mM IPTG at 30 °C for 4 h. Bacteria were lysed by sonication in 50 mM Tris-HCl buffer (pH 8) containing 1 mM DTT, 50 μM zinc acetate, and 300 μL of protease inhibitor cocktail (Sigma). Fusion proteins were affinity purified using a glutathione sepharose 4B resin (Pharmacia Biotech). Protein concentrations were measured by the Bradford method (Bio-Rad) with BSA as a standard. The size and purity of the proteins were measured by electrophoresis in 10−15% SDS polyacrylamide gels. The purified fusion proteins were stored in aliquots at −80 °C.

DNA Oligonucleotide Preparation. Oligonucleotides for site-directed mutagenesis were purchased from Invitrogen. All concentrations were expressed in strand molarity with a nearest-neighbor approximation for the absorption concentrations of the unfolded species.

Three DNA 5′-biotinylated oligonucleotides (Invitrogen, HPLC purified) were used for SPR and ELISA binding studies: the *Htelo* human telomeric quadruplex d(biotin-[$(GGTTAG)_5$]) as used for the initial phage selection experiments (*15*); the *Atelo* single-stranded DNA control d(biotin-[$(AGTTAG)_5$]) derived from the telomeric quadruplex sequence but carrying a G→A mutation in the repeat unit to prevent quadruplex formation; the Zif_{DNA} Zif268 double-stranded DNA substrate d(biotin-[ggcatagt-GCGTGGGCGttagc]) hybridized with its complementary sequence and which contains the nine base pair Zif268 recognition site (uppercase).

Htelo at a concentration of 20 μM was annealed in buffer containing 50 mM Tris at pH 7.4 and 150 mM potassium chloride by heating to 95 °C for 5 min. After slow cooling for 3 h to room temperature, the oligonucleotide was stored at −20 °C. CD spectroscopy of the oligonucleotide shows a maximum at 295 nm and a shoulder at about 270 nm, consistent with observations made by ourselves and others (*15*, *17*) with the quadruplex formed by the sequence $(TTAGGG)_4$ in potassium. *Atelo*, in comparison, has no signature confirming its lack of structure.

To form the biotinylated Zif_{DNA} double-stranded DNA, complementary strands (one of them containing a 5′ biotin label) were annealed at 20 μM concentration as described above for *Htelo*.

Surface Plasmon Resonance. All experiments were carried out on a Biacore 2000 Biosensor using freshly filtered and degassed buffers. Biotinylated oligonucleotides were loaded onto a SA chip (Biacore) using the MANUAL INJECT command (Biacore 2000 control software) at a flow rate of 10 μL/min in loading buffer (20 mM Tris pH 7.4, 150 mM KCl). GST-fusion proteins were diluted in running buffer (20 mM Tris pH 7.4, 150 mM KCl, 1 mM DTT, 50 μM zinc acetate, 1 mM $MgCl_2$, 15 μg/mL calf thymus DNA (Pharmacia)) to concentrations of 1.25, 2.5, 5, 10, 20, and 40 nM. The KINJECT command was used to inject samples for a period of 4 min followed by a dissociation phase of 180 s. Between successive protein injections, the chip was regenerated by injecting running buffer containing 1 M NaCl for 1 min.

To calculate the binding constants for Gq1 to the quadruplex, sensorgrams measured for *Atelo* were subtracted from

Htelo flow cell data. A similar subtraction was performed with Zif268 data for the target Zif_{DNA}. Corrected sensorgrams were analyzed using a Langmuirian global kinetics fit model and the mass transfer model (Biaevaluation 3.1). The algorithm fits the data from each sensorgram globally to the closest fitting set of theoretical sensorgrams describing a binding event according to the equilibrium below where ligand (L) and analyte (A) interact (*18*).

$$L + A \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} LA$$

Enzyme-Labeled Immunosorbent Assay (ELISA). The wells of streptavidin coated microplates (96 well StreptaWell HighBind, Roche Diagnostics) were incubated with 300 μL aliquots of biotinylated oligonucleotide at a concentration of 25 nM in potassium phosphate buffer (50 mM KP_i, 100 mM KCl, pH 7.4) for 6 h at room temperature. The plate was washed four times with the same buffer and then incubated with 50 μL aliquots of serial dilutions of zinc finger-GST fusion protein in PBST-Zn buffer (50 mM KP_i, 100 mM KCl, pH 7.4, 0.2% Tween 20, 50 μM $Zn(OAc)_2$) containing 20 μg/mL salmon sperm DNA (Amersham) at 4 °C for 1 h. The plate was washed four times with PBST-Zn buffer at 4 °C, and then each well was incubated with 50 μL of AntiGST-HRP conjugate (Amersham, 5000× dilution) in PBST-Zn buffer for 1 h at 4 °C. After washing four times (PBST-Zn buffer, 4 °C), 100 μL TMB substrate (Amersham) was added to each well at room temperature, and following development of a blue coloration, the reaction was quenched with 100 μL of aqueous H_2SO_4 (0.19 M). The absorbance at 450 nM was read immediately using a Biokinetics plate reader.

After subtraction of the background (i.e. protein and AntiGST-HRP omitted), the absorbance, *A*, was fitted to the Langmuir adsorption isotherm

$$A = \frac{K_a A_{max} c}{1 + K_a c}$$

where the association constant, K_a, and A_{max} are adjustable parameters, and *c* is the protein concentration. The apparent dissociation constant for the zinc finger−DNA interaction is given by $K_d = 1/K_a$.

The evaluation of protein binding to *Htelo* and Zif_{DNA} permitted higher titers of protein (800 nM) without detectable background while nonspecific interactions were observed with the dextran matrix of the SPR chip when using protein concentrations higher than 320 nM.

Molecular Modeling. Docking Procedure. The crystal structure (*19*) of the G-quadruplex formed from the 22-mer human telomeric DNA sequence $d(AG_3[TTAG_3]_3)$ (PDB code 1KF1) was taken as the starting point for simulating quadruplex binding of the Gq1 protein. The crystal structure (*10*) of the Zif268 protein bound to an 11-mer double-stranded G-rich duplex DNA (PDB code 1AAY) was used to obtain the starting model for Gq1. The wild-type Zif268 protein structure was mutated to Gq1, maintaining the original backbone torsional angles, using the Biopolymer module in the InsightII package (Accelrys Inc.), followed by the addition of missing hydrogen atoms to the structure. This starting structure was minimized with the Sander module in the AMBER7 suite of programs (*20*−*22*), followed by an initial molecular dynamics trajectory (2 ns) in order to check that the model was stable.

The structure was used in a docking procedure with the Affinity Docking module in InsightII and an implementation of the AMBER force field, to find low-energy structures of complexes when bound to the quadruplex DNA structure. The previously calculated molecular dynamics trajectory, as well as other reported simulations of zinc finger proteins bound to double-stranded DNA (*23*), shows that these proteins comprise rigid α-helices held to β-sheet domains maintained by the coordination of a zinc cation with cysteines and/or histidines. Each zinc cation stabilizes one finger domain, but the amino acid sequence(s) linking the fingers allows total flexibility in the way that the protein interacts with a nucleic acid structure, as shown in the crystal structure of a zinc finger−RNA complex (*24*) where the zinc finger protein adopts a conformation very distinct from that in a duplex DNA complex. For this reason ϕ and ψ dihedral angles of nonlinking residues (i.e. residues 103−129, 134−155, and 162−186) were constrained to their initial values during the docking procedure by means of quadratic torsional restraints of 30 kcal mol^{-1} using the Discover3 module of InsightII, thereby providing appropriate flexibility in the linker regions.

A multipart docking protocol was used. First, 200 protein orientations were randomly generated. van der Waals radii were set to 10% of the full value, charges were not considered, and the nonbonded cutoffs were set to 5 Å. Each accepted structure was further minimized for 200 steps using the conjugate gradient method. The maximum allowable change for succeeding structures (energy tolerance) was set to 1×10^6 kcal mol^{-1}, and the energy range was set to 200 kcal mol^{-1}. The 75 lowest energy structures were used for the second phase of the protocol.

Simulated annealing was then used in order to further refine the initial placement for these 75 structures. During this phase the van der Waals radii were adjusted back to their full values, charges and a distance-dependent dielectric constant ($4r_{ij}$) were included, and the nonbonded cutoff was changed to 15 Å. Each structure was minimized for 100 steps of conjugate gradients prior to a simulated annealing calculation in which the system was cooled from 500 to 300 K over 10 ps. The resulting structures were finally minimized for 800 steps of conjugate gradients. The docking protocol took almost 200 h of computational time on a 2-processor SGI Octane. The minimization, equilibration, and dynamics runs with Sander were performed using the AMBER force field with the Parm99 parameters (*20*−*22*). As the Parm99 set has only parameters for nonbonded zinc atoms, this approximation was initially used to describe the zinc atom in each finger, since this procedure has been shown to be adequate for similar systems (*25*). NMR-type restraints were subsequently used (*26*), to constrain the distances between the zinc atom, the two cysteines, and the two histidines involved in the coordination scheme in each zinc finger. It is important to emphasize that no torsional restraints for nonlinking residues were used in these calculations. In that way, we could corroborate the stability of each domain and flexibility due to linker sequences. The stability of the final complex structures was also verified by running molecular dynamics simulations with a modified parameter set incor-

porating polarization effects (Roldan *et al.*, to be published). The complete procedure was repeated with the Zif268 zinc finger structure as a starting point. In both cases the 75 docked structures with the lowest total energy were examined.

Binding energies were calculated for the top 20 solutions in each case, which allows direct comparison to be made between molecules with different numbers of atoms (i.e. Zif268 complexes vs Gq1 complexes). Each docked structure was minimized using the steepest descent method with the same force field, parameter set, and program as used for the docking. As in the docking procedure, a distance-dependent dielectric constant ($4r_{ij}$) and a 15 Å cutoff were used. The binding energy E_B was calculated as $[E_{total} - (E_{zinc\ fingers} + E_{quadruplex})]$.

RESULTS AND DISCUSSION

Evaluating DNA Binding by SPR and ELISA. The binding affinities of native and mutant proteins for the intramolecular human telomeric quadruplex and Zif268 double-stranded DNA substrate were measured by two independent methods, surface plasmon resonance (SPR) and an enzyme-labeled immunosorbent assay (ELISA). The binding affinities of both Gq1 and Zif268, for quadruplex and duplex DNA targets, were reconfirmed by independent measurements using SPR and ELISA. The affinity of Zif268 for its double-stranded DNA substrate (K_d = 5.3 nM by SPR, 12 nM by ELISA) was comparable to that previously reported by Yang *et al.* (K_d = 6.5 nM) (*27*). The dissociation constant obtained for Gq1 binding to human telomeric quadruplex (24 nM by SPR, 59 nM by ELISA) was also comparable to that previously reported by us using either a whole phage ELISA or an electromobility shift assay (26 and 34 nM, respectively) (*16*). This would appear to validate SPR and ELISA as reasonable methods for measuring the binding interaction of the mutant zinc finger proteins with DNA targets. The general trend for the binding properties of all proteins evaluated was the same for both techniques. The K_d values obtained by ELISA were sometimes found to be slightly higher than those measured by SPR, which could be explained by reduced accessibility of the DNA target when immobilized on a Streptavidin coated ELISA plate, as compared to immobilization on a dextran matrix (SPR). However, ELISA enabled all zinc fingers to be tested against *Htelo* and *Zif*$_{DNA}$ using up to 800 nM protein without detectable background whereas nonspecific interactions were observed with the dextran matrix of the SPR chip at protein concentrations higher than 320 nM, which was advantageous for the evaluation of weak binding proteins.

Design and Binding Properties of Gq1 Mutants. The parent quadruplex binding protein, Gq1, was derived from the mouse transcription factor Zif268 based on its ability to bind the intramolecular human telomeric quadruplex (*15*). Gq1 differs from Zif268 by twelve amino acids: six located on finger 2, and three on each of fingers 1 and 3. To explore the contribution of specific amino acids and individual finger domains to both quadruplex affinity and specificity, we carried out a range of mutagenesis substitution experiments. A number of single amino acid replacements were made (see Figure 2), by alanine substitutions for a selected number of residues found to be highly conserved in the quadruplex binding family of zinc finger proteins originally selected (*15*).

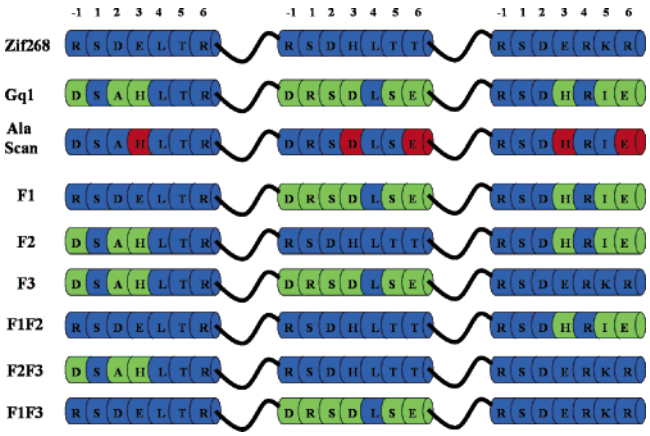

FIGURE 2: Schematic of Zif268, Gq1, and the mutant proteins. The zinc finger domains are shown as cylinders, and the linkers are shown as thick lines. For each finger, only the amino acids in positions −1, 1, 2, 3, 4, 5, and 6 are represented. Residues conserved with Zif268 are colored blue, and Gq1-specific residues are colored green. Residues that were individually mutated by alanine in the alanine scan are highlighted in red.

The alanine substituted mutants (Figure 2) were all found to exhibit affinities for quadruplex *and* duplex DNA that were indistinguishable from those of Gq1, within experimental error (Supporting Information), although there were slight differences in the on and off rates. These data show that the interaction between any one of these residues and the quadruplex target is expendable for molecular recognition. This contrasts with previous studies on zinc finger·DNA duplex recognition proteins where the replacement of a single key amino acid has often led to a significant alteration in either DNA binding properties (*28*), homo- or hetero-oligomerization (*29*), or even its ability to act as a nuclear localization signal (*30*). As a result of this initial study, we opted to explore the effects of more drastic alterations of Gq1 by the replacement of entire fingers, rather than pursue more point mutations. Figure 2 shows the finger swap mutants of Gq1, that involve replacement of entire zinc fingers with corresponding fingers from Zif268. Surprisingly, all single finger swapped proteins (F1, F2, and F3) were found to bind the intramolecular human telomeric quadruplex with an affinity, and specificity, comparable to that of Gq1 (Table 1). Furthermore, F1, F2, and F3 showed no binding to the duplex DNA at a maximum protein concentration of 800 nM. These results suggest that any combination of two fingers originating from Gq1, where the third is from Zif268, is sufficient for quadruplex recognition. These results prompted us to explore the properties of double finger replacements (Table 1). When combinations of two fingers of Gq1 were simultaneously replaced by the corresponding fingers of Zif268, little change in binding was observed for the case where F2 of Gq1 was retained (i.e. F1F3). However, double finger swaps lead to significantly reduced quadruplex binding affinity (10-fold or higher K_d) when one of the two fingers was F2 (i.e. F1F2 and F2F3). Moreover, F1F2 and F2F3 each also showed significant binding to Zif268 double-stranded DNA substrate with a K_d of 25 and 18 nM, respectively, thus indicating significantly depleted quadruplex specificity.

These data suggest that finger 2 of Gq1 plays an important role in quadruplex affinity and quadruplex versus duplex specificity, since replacement of both fingers 1 and 3 of Gq1

Table 1: Dissociation Constants (K_d) Determined by SPR[a] and ELISA[b]

protein	K_d(*Htelo*) (nM)		K_d(*Zif*$_{DNA}$) (nM)	
	SPR	ELISA	SPR	ELISA
Zif268	*c*	*c*	5.3 ± 1.5	12 ± 5
Gq1	24 ± 5	59 ± 15	*c*	*d*
F1	55 ± 12	67 ± 17	*c*	*d*
F2	67 ± 14	62 ± 15	*c*	*d*
F3	32 ± 7	72 ± 17	*c*	*d*
F1F3	57 ± 13	71 ± 17	*c*	*d*
F2F3	280 ± 80	280 ± 95	18 ± 5	10 ± 4
F1F2	*c*	*d*	25 ± 5	34 ± 8
Gq1F1F2	300 ± 95	780 ± 150	*e*	*d*

[a] Experiments were carried out in 20 mM Tris, pH = 7.4 containing 150 mM KCl, 1 mM DTT, 50 μM zinc acetate, 1 mM $MgCl_2$, and 15 μg/mL calf thymus DNA (Pharmacia). All experiments were carried out in duplicate. Fitting errors are less than ±5% for K_d values. [b] Experiments were carried out in PBST-Zn buffer (50 mM KP_i, 100 mM KCl, pH = 7.4, 0.2% Tween 20, 50 μM zinc acetate) containing 20 μg/mL salmon sperm DNA (Amersham). Experiments were carried out in duplicate. [c] No detectable binding at 320 nM protein concentration. [d] No detectable binding at 800 nM protein concentration. [e] No detectable binding at 1 μM protein concentration.

has no effect on protein−DNA binding properties, while replacement of finger 2 *and* either finger 1 or finger 3 abolishes quadruplex binding affinity and specificity. The importance of finger 2 is also supported by the fact that although all three fingers of Zif268 were each randomized to the same degree in the initial phage display library (*15*), Gq1 differs from the parent protein, Zif268, by twelve amino acid mutations, of which six are contained in finger 2 with only three mutations in each of fingers 1 and 3. However, the binding properties of mutant F2 indicated that Gq1 may also be able to adopt a binding mode where finger 2 is not absolutely essential.

Molecular Modeling Studies. To complement the experimental work on zinc finger−quadruplex binding, we carried out molecular modeling studies. In these studies, we simulated a large number of potential binding modes for both Gq1 and Zif268 zinc finger proteins binding to the human intramolecular quadruplex, representing the wide range of flexibility available to these two zinc finger proteins. The X-ray derived structure of the parallel-stranded G-quadruplex formed from the 22-mer human telomeric DNA sequence d[AGGG(TTAGGG)$_3$] (PDB code 1KF1) was used as the binding host. We have assumed, in the absence of any experimental data to the contrary, that the quadruplex itself retains the same fold as that in the native crystal structure; there was no indication during the molecular dynamics simulations that other folds were preferred since at no point was the quadruplex structure significantly perturbed. Equivalent docking calculations were carried out on Gq1 and Zif268 quadruplex DNA complexes. For each case the 10 structures with lowest total energy are listed in Table 2. The set of Zif268 complexes has three structures with binding energies greater than −400 kcal·mol^{-1}, with one, number 24, of outstandingly greater binding energy than any other. The set of Gq1 complexes (Table 2) has just two structures with binding energies greater than −400 kcal mol^{-1}. However, both are of greater binding energy than the most stable Zif268 one, and structure 37 is some 37 kcal mol^{-1} lower in energy than the best Zif268 complex (Figure 3a).

Table 2: Lowest Total Energy Docked Zif268− and Gq1−Quadruplex Structures, in kcal mol $^{-1}$[a]

structure number	total energy	interaction energy	binding energy	no. of inter H-bonds	no. of intra H-bonds
		Zif268			
24	**−414.51**	**−280.956**	**−476.44**	**20**	**77**
48	−406.729	−154.242	−326.35	12	95
61	−404.481	−254.612	−381.58	22	86
1	−387.216	−100.617	−100.47	7	89
71	−385.683	−204.461	−410.71	21	77
16	−385.613	−214.405	−178.26	13	80
57	−385.397	−248.283	−362.54	19	82
51	−384.282	−204.351	−307.33	16	84
47	−382.42	−225.872	−382.22	12	84
31	−375.785	−224.461	−435.8	18	79
		Gq1			
8	−441.406	−239.077	−337.4	17	78
14	−435.008	−160.452	−222.84	14	84
22	−430.256	−166.742	−225.89	14	81
65	−428.955	−184.772	−482.92	7	80
64	−425.065	−199.273	−371.04	17	80
11	−423.249	−61.3581	−101.22	6	102
52	−411.08	−175.214	−332.66	14	77
26	−407.639	−137.559	−171.7	5	83
69	−407.216	−186.536	−323.18	11	72
37	**−405.702**	**−201.83**	**−513.76**	**16**	**75**

[a] Values in bold are the lowest energy structures within each set.

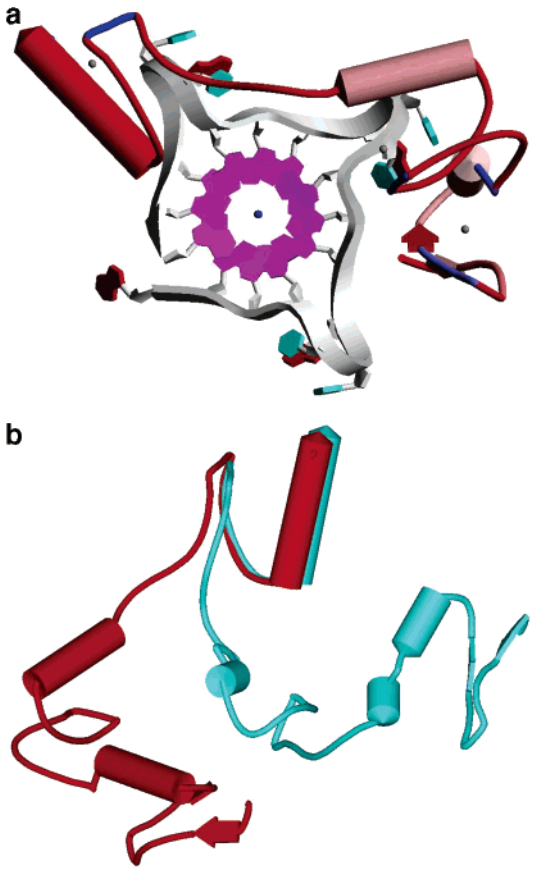

FIGURE 3: (a) The highest binding energy conformation of Gq1 (number 37 from Table 2) docked on the parallel intramolecular telomeric quadruplex. (b) Superposition of the highest binding energy structures of the Gq1 protein (number 37), in red, and the Zif268 protein (number 24), in blue.

Comparison of the Gq1 and Zif268 modeled structures suggests that they have very distinct conformations, as shown in Figure 3b. It is notable that all the low-energy Zif268

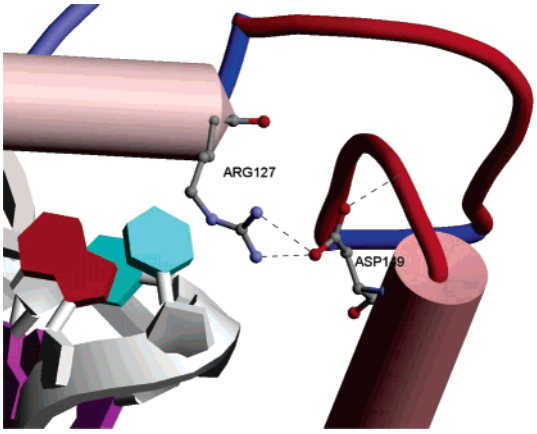

FIGURE 4: Specific residues from the helices in fingers 1 and 2 that are involved in intramolecular hydrogen bonds, as found in the highest binding energy structure (model number 37) for the Gq1 complex.

structures are only able to maximize their contacts with the quadruplex by perturbing the zinc finger arrangements, mostly by interruptions in the continuity of one or other of the helices, which is likely to be unfavorable (see Supporting Information). We also find that the human quadruplex structure is only able to make effective hydrogen-bonding contacts with two out of three zinc fingers from the Zif268 structures. The structure with the highest binding energy (number 24 in Table 2) only has one hydrogen-bonded contact from finger three with the DNA, involving lys183 and an external loop thymine. In striking contrast, all three fingers in the Gq1 structure with the lowest binding energy (number 37 in Table 2) make extensive hydrogen-bond contacts with the DNA, together with a large number of attractive van der Waals interactions.

No specific protein−DNA hydrogen-bond interactions involving the mutated residues were observed in any of the Gq1 models. However, intramolecular hydrogen-bond interactions were consistently found in both the Gq1 and Zif268 models, which stabilize particular (though very different) backbone conformations, especially sharp turns. Several of the Gq1 residues not present in Zif268 are involved in separate hydrogen-bond interactions that maintain the Gq1 structures in a quite distinct set of conformations. In particular there are a number of interactions that constrain finger 1 relative to finger 2. For example, an apparent key interaction conserved across the set of best Gq1 structures involves arg127 of finger 1 (also present in zif268) hydrogen-bonding to asp149 of finger 2 (unique to Gq1) (Figure 4).

A Two Finger Miniprotein. The domain swapping experiments, taken with the lowest energy Gq1-quadruplex model, suggested to us that a two finger miniprotein derived from Gq1 might retain significant Gq1-like binding properties. We tested this hypothesis by deleting finger 3 of Gq1 to generate a truncated protein, Gq1F1F2, composed only of fingers 1 and 2 of Gq1. Gq1F1F2 showed no binding to duplex DNA at 800 nM (ELISA) or 1 μM (SPR), whereas the K_d value for *Htelo* quadruplex by ELISA and SPR was 780 and 300 nM, respectively. Thus, Gq1F1F2 retains some of the specificity observed for Gq1, but with approximately 10-fold reduced affinity for quadruplex. This suggests there is potential for the design of two finger quadruplex binding proteins, with further optimization of the amino acid contacts. That significant specificity and affinity are retained by Gq1F1F2 is consistent with the multiple interactions between fingers 1 and 2 predicted in structure 37, whereas no significant interaction was predicted between fingers 2 and 3.

CONCLUSIONS

The mutational and binding studies show that any one finger of Gq1 could be replaced with the corresponding finger of Zif268, without significant loss of quadruplex affinity or duplex discrimination. The simultaneous replacement of two fingers of Gq1 by Zif268 only resulted in significant impairment of quadruplex recognition and loss of duplex discrimination, in cases that involved finger 2 replacement (i.e. F1F2 and F2F3). However, quadruplex recognition and duplex discrimination were unaffected when finger 2 was retained (i.e. F1F3). The double finger swap studies suggest that finger 2 is most critical for both quadruplex affinity and specificity. These results also suggest that all three fingers of Gq1 do not need to cooperate to bind quadruplex.

The molecular modeling studies are consistent with the major conclusions drawn from these experiments. All zinc finger proteins in the absence of a nucleic acid can adopt a wide range of conformations. Modeling shows that the energetically most preferred conformations of Gq1 when bound to the human parallel-stranded quadruplex structure are very distinct and also are of lower binding energy when compared to those obtained from the Zif268 protein. Thus, structure 37 for the Gq1−quadruplex complex showed a significantly greater binding energy to the quadruplex than that for the best Zif268−quadruplex complex. Structure 37 has a set of intramolecular interactions between fingers 1 and 2 that maintain their relative orientation in an optimal conformation for quadruplex binding, suggesting that fingers 1 plus 2 alone would still bind but with reduced affinity. Experimental evaluation of the truncated protein Gq1F1F2 shows it does indeed recognize quadruplex with some selectivity and reduced affinity.

SUPPORTING INFORMATION AVAILABLE

Experiments that describe the effect of specific point mutations on the ability of Gq1 to bind to the human telomeric quadruplex *Htelo*. Typical SPR sensorgrams obtained for mutants of Gq1 binding to *Htelo* are also presented. Coordinates for Zif268 model 24 and Gq1 model 37 are also available in pdb format. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000) DNA recognition by Cys2His2 zinc finger proteins, *Annu. Rev. Biophys. Biomol. Struct. 29*, 183−212.
2. Segal, D. J., and Barbas, C. F., 3rd. (2000) Design of novel sequence-specific DNA-binding proteins, *Curr. Opin. Chem. Biol. 4*, 34−39.
3. Segal, D. J., and Barbas, C. F., 3rd. (2001) Custom DNA-binding proteins come of age: polydactyl zinc-finger proteins, *Curr. Opin. Biotechnol. 12*, 632−637.
4. Pavletich, N. P., and Pabo, C. O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A, *Science 252*, 809−817.
5. Desjarlais, J. R., and Berg, J. M. (1992) Toward rules relating zinc finger protein sequences and DNA binding site preferences, *Proc. Natl. Acad. Sci. U.S.A. 89*, 7345−7349.

6. Desjarlais, J. R., Berg J. M. (1994) Length-encoded multiplex binding site determination: application to zinc finger proteins, *Proc. Natl. Acad. Sci. U.S.A. 91*, 11099−11103.
7. Choo, Y., and Klug, A. (1994) Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage, *Proc. Natl. Acad. Sci. U.S.A. 91*, 11163−11167.
8. Greisman, H. A., and Pabo, C. O. (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target site, *Science 275*, 657−661.
9. Segal, D. J., Dreier, B., Beerli, R. R., and Barbas, C. F., III (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences, *Proc. Natl. Acad. Sci. U.S.A. 96*, 2758−2763.
10. Elrod-Erickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions, *Structure 4*, 1171−1180.
11. For a recent review, see: Jantz, D., Amann, B. T., Gatto, G. J., Jr., and Berg, J. M. (2004) The design of functional DNA-binding proteins based on zinc finger domains, *Chem. Rev. 104*, 789−799.
12. Neidle, S., and Parkinson, G. (2002) Telomere maintenance as a target for anticancer drug discovery, *Nat. Rev. Drug Discovery 1*, 383−393.
13. Rangan, A., Fedoroff, O. Y., and Hurley, L. H. (2001) Induction of duplex to G-quadruplex transition in the c-myc promoter region by a small molecule, *J. Biol. Chem. 276*, 4640−4646.
14. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription, *Proc. Natl. Acad. Sci. U.S.A. 99*, 11593−11598.
15. Isalan, M., Patel, S. D., Balasubramanian, S., and Choo, Y. (2001) Selection of zinc fingers that bind single-stranded telomeric DNA in the G-quadruplex conformation, *Biochemistry 40*, 830−836.
16. Patel, S. D., Isalan, M., Gavory, G., Ladame, S., Choo, Y., and Balasubramanian, S. (2004) Inhibition of human telomerase activity by an engineered zinc finger protein that binds G-quadruplexes, *Biochemistry 43*, 13452−13458.
17. Balagurumoorthy, P., and Brahmachari, S. K. (1994) Structure and stability of human telomeric sequence, *J. Biol. Chem. 269*, 21858−21869.
18. Karlsson, R., and Falt, A. (1997) Experimental design for kinetic analysis of protein−protein interactions with surface plasmon resonance biosensors, *J. Immunol. Methods 200*, 121−133.
19. Parkinson, G. N., Lee, M. P., and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA, *Nature 417*, 876−880.
20. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) An all atom force-field for simulations of proteins and nucleic acids, *J. Comput. Chem. 7*, 230−253.
21. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
22. Ponder, J. W., and Case, D. A. (2003) Force fields for protein simulations, *Adv. Protein Chem. 66*, 27−85.
23. Marco, E., Garcia-Nieto, R., and Gajo, F. (2003) Assessment by molecular dynamics simulations of the structural determinants of DNA-binding specificity for transcription factor Sp1, *J. Mol. Biol. 328*, 9−32.
24. Lu, D., Searles, A., and Klug, A. (2003) Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition, *Nature 426*, 96−100.
25. Banci, L. (2003) Molecular dynamics simulations of metalloproteins, *Curr. Opin. Struct. Biol. 7*, 143−149.
26. Stote, R. H., and Karplus, M. (1995) Zinc binding in proteins and solution: a simple but accurate nonbonded representation, *Proteins 23*, 12−31.
27. Yang, W. P., Wu, H., and Barbas, C. F., 3rd. (1995) Surface plasmon resonance based kinetic studies of zinc finger-DNA interactions, *J. Immunol. Methods 183*, 175−182.
28. Schreiber, J., Enderich, J., and Wegner, M. (1998) Structural requirements for DNA binding of GCM proteins, *Nucleic Acids Res. 26*, 2337−2343.
29. Westman, B. J., Perdomo, J., Matthews, J. M., Crossley, M., and Mackay, J. P. (2004) Structural studies on a protein-binding zinc-finger domain of Eos reveal both similarities and differences to classical zinc fingers, *Biochemistry 43*, 13318−13327.
30. Pandya, K., and Townes, T. M. (2002) Basic residues within the Kruppel zinc finger DNA binding domains are the critical nuclear localization determinants of EKLF/KLF-1, *J. Biol. Chem. 277*, 16304−16312.

BI050229X