Published online 31 March 2010

Nucleic Acids Research, 2010, Vol. 38, No. 13 ***4325–4336***
doi:10.1093/nar/gkq175

Cell-type-specific long-range looping interactions identify distant regulatory elements of the *CFTR* gene

Nele Gheldof[1], Emily M. Smith[1], Tomoko M. Tabuchi[1], Christoph M. Koch[2], Ian Dunham[2], John A. Stamatoyannopoulos[3] and Job Dekker[1,*]

[1]Program in Gene Function and Expression and Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605-0103, USA, [2]European Bioinformatics Institute (EBI), The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK and [3]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

Received October 17, 2009; Revised March 1, 2010; Accepted March 3, 2010

ABSTRACT

Identification of regulatory elements and their target genes is complicated by the fact that regulatory elements can act over large genomic distances. Identification of long-range acting elements is particularly important in the case of disease genes as mutations in these elements can result in human disease. It is becoming increasingly clear that long-range control of gene expression is facilitated by chromatin looping interactions. These interactions can be detected by chromosome conformation capture (3C). Here, we employed 3C as a discovery tool for identification of long-range regulatory elements that control the cystic fibrosis transmembrane conductance regulator gene, *CFTR*. We identified four elements in a 460-kb region around the locus that loop specifically to the *CFTR* promoter exclusively in *CFTR* expressing cells. The elements are located 20 and 80 kb upstream; and 109 and 203 kb downstream of the *CFTR* promoter. These elements contain DNase I hypersensitive sites and histone modification patterns characteristic of enhancers. The elements also interact with each other and the latter two activate the *CFTR* promoter synergistically in reporter assays. Our results reveal novel long-range acting elements that control expression of *CFTR* and suggest that 3C-based approaches can be used for discovery of novel regulatory elements.

INTRODUCTION

Appropriate spatial and temporal control of gene expression depends on regulatory input from *cis*-acting elements such as promoters, enhancers and repressors. Identification of these elements is challenging because they can be located far from their target gene, sometimes up to several megabases (1–3). Detailed knowledge of such distant regulatory elements and their mechanism of action will greatly contribute to basic understanding of gene expression.

Abundant evidence suggests that human disease can be caused by mutations that affect distant regulatory elements, while leaving the disease gene itself intact. Examples include Aniridia, caused by loss of distant regulatory elements of the *PAX6* gene, and blepharophimosis syndrome (BPES) that can be caused by deletion of regulatory elements located >600 kb from the *FOXL2* gene (4,5).

Regulatory elements are often characterized by the presence of DNase I hypersensitive sites (DHS), which can mark the position where transcription factors are bound to DNA. Other chromatin features found at distant regulatory elements are increased levels of H3K4me1 and histone acetylation (6,7). In addition, these sequences are often conserved across species (8,9). All these features can be used to identify putative functional elements and these powerful strategies are currently widely applied (10,11). However, these analyses do not immediately reveal the target genes of these regulatory elements.

*To whom correspondence should be addressed. Tel: +1 508 856 4371; Fax: +1 508 856 4650; Email: job.dekker@umassmed.edu
Present addresses:
Nele Gheldof, Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland.
Christoph M. Koch, Abbott GmbH & Co. KG, Max-Planck-Ring 2, 65205 Wiesbaden, P.O. Box 21 03, Wiesbaden 65011, Germany.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2010. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

Regulatory elements can directly associate with target promoters through chromatin looping (1–3,12). These looping interactions can be detected using chromosome conformation capture, or 3C (13). The insight that regulatory elements physically associate with promoters provides a methodology to discover novel regulatory elements by performing systematic 3C analyses to search for genomic elements that are found to interact with a specific promoter. Here we tested the feasibility of such an approach by analysis of the cystic fibrosis transmembrane conductance regulator (*CFTR*) locus. Identification of extragenic regulatory elements for this locus (and other disease related loci) is especially important because (i) they could be screened for mutations in patients with no known mutations in the *CFTR* gene itself, and thus aid in proper diagnosis, and (ii) they could be included in gene therapy constructs (to recapitulate endogenous *CFTR* regulation). Finally, identification of *CFTR* regulatory elements will provide basic insights into the mechanisms that control expression of *CFTR*, which could also lead to new approaches to boost or manipulate *CFTR* expression in patients.

The *CFTR* gene is, when both alleles are mutated, responsible for cystic fibrosis. It contains a promoter that has many characteristics of a housekeeping gene including potential binding sites for SP1 (14). Also present is a critical CCAAT-like element, shown to bind C/EBP (15), implicating cAMP as a possible regulator. Supporting a role for cAMP in *CFTR* regulation is data showing that cAMP activation of protein kinase A can regulate basal *CFTR* expression (16) and the discovery that CREB and ATF-1 bind the *CFTR* promoter in a cAMP-responsive manner (17). A YY1 element has also been identified that, when mutated, significantly increases the expression of *CFTR* (18). However, despite the abundance of regulatory elements in the *CFTR* promoter, it is clear that additional elements are required for the complex spatio-temporal expression pattern of the *CFTR* gene (19). Indeed, work from the Harris laboratory has identified additional putative regulatory elements located within introns of the gene, as well up- and down-stream of the locus. For instance, HNF1α has been found to interact with a putative regulatory element in introns 1, 10, 17a and 20 and over-expression of this protein results in increased *CFTR* mRNA levels (20,21).

Here we applied a systematic 3C analysis to a 460 Kb chromosomal region surrounding the *CFTR* transcription start site (TSS). We examined different cell lines that either express or do not express the gene in order to identify regulatory elements that function specifically in *CFTR*-expressing cells. Our approach was validated by identification of previously discovered regulatory elements, e.g. an element located 203 Kb downstream of the promoter that coincides with a DHS [referred to as DHS4574+15.3 (22)] and that has been found to be able to act as an insulator (23). Importantly, we discovered an additional regulatory element: one located within intron 11 (109 kb downstream of the TSS) that interacts with the *CFTR* TSS exclusively in cells that express the gene. Additional 3C analyses allowed us to define the locations of long-range-acting elements at ~1 kb resolution. In *CFTR* expressing cells, these elements contain the characteristic features of known regulatory elements, such as the presence of a DHS and specific histone modification patterns. Interestingly, we find that regulatory elements also interact with each other, and in the case of the elements located 109 and 203 Kb downstream of the TSS synergistically activate the *CFTR* promoter in reporter assays. These studies identify novel *CFTR* regulatory elements and provide insights into combinatorial control of the gene. Finally, these results provide strong evidence that 3C-based approaches provide tools for *ab initio* discovery of regulatory elements and their target genes.

MATERIALS AND METHODS

Cell culture

We used six cell lines for the 3C analyses: Caco2, HT29, HeLa S3, GM06990, K562 and HepG2. The GM06990 cell line was obtained from Coriell Cell Repositories, and the Caco2, HT29, HeLa S3, K562 and HepG2 cell lines from the American Type Culture Collection (ATCC). All cell lines were grown at 37°C in 5% CO_2 in medium containing 1% penicillin–streptomycin. Caco2 cells were grown in MEM alpha medium supplemented with 20% fetal bovine serum (FBS), HT29 cells in DMEM medium with 10% FBS, HeLa S3 in F12K medium with 10% FBS, GM06990 and K562 cells in RPMI medium with 10% FBS and HepG2 in MEM alpha medium with 10% FBS. Suspension cells were harvested at log-phase and monolayer cells were grown to 95% confluency before harvesting for RT-PCR and 3C analysis.

RT-PCR

Total RNA from the six cell lines was isolated using the RNeasy Mini Kit (Qiagen). *CFTR* RNA transcript levels were analyzed using *Power* SYBR® Green RNA-to-C_T *1-Step* Kit (Applied Biosystems) on a *StepOnePlus*™ Real-Time PCR System (Applied Biosystems). *HPRT1* was used to normalize the data. Primer sequences are available in Supplementary Table S1.

3C analysis

The 3C analysis was performed using EcoRI as described previously (24–26). We generated a control library using BAC clones obtained from Invitrogen and the Children's Hospital Oakland Research Institute (CHORI). We used three minimally overlapping BAC clones spanning the investigated locus: RP11-35E12, RP11-450L14 and CTD-2034E23, as well as the BAC clone RP11-197K24 from a gene desert region (ENCODE region Enr313) as described in Dostie *et al*. (27). We normalized the data of each experiment using 20 interaction frequencies measured in this gene desert region to allow direct comparison of data obtained with different cell lines. 3C analysis of Caco2 versus GM06990 was performed in four independent experiments, HT29 versus K562 in 2, and HeLa S3 versus HepG2 in 1, and each experiment was quantified at least in triplicate. 3C fine mapping of elements III and

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

IV was preformed with BsrGI in the cell lines Caco2 and GM06990, and three independent experiments were performed. Primer sequences are available in Supplementary Table S1. 3C data is available in Supplementary Table S2.

Generation of reporter constructs

The pGL3 basic vector (Promega) was transformed into a Gateway-compatible destination vector by inserting the R4R2 cassette into the *Mlu*I restriction site according to the MultiSite Gateway Cloning manual (Invitrogen). Insertion of the cassette in the correct orientation to drive transcription of the luciferase gene was verified by sequencing. The R4R2 cassette was a generous gift by the Walhout Lab. A 1.7-kb fragment spanning the *CFTR* basal promoter (from −1691 to −35 bp from the ATG translation start site) was amplified with primers with B1B2 tails and cloned into the destination vector. Potential enhancer elements were amplified with primers containing B4B1R tails and each of these elements was cloned upstream of the promoter fragment. As the EcoRI DNA looping elements as identified by 3C were on average 4 kb in size, we split each element in smaller overlapping segments of 1.1–1.8 kb to allow positioning of the regulatory element more precisely: element III was split into two and element IV into three segments. The genomic coordinates (hg18, chromosome 7) of each segment are as follows: promoter fragment: 116905694–116907350; IIIa: 117 013 018–117 014 528; IIIb: 117 014 436–117 016 269; IVa: 117 105 596–117 107 448; IVb: 117 107 337–117 109 457; IVc: 117 109 425–117 110 818. As negative control, the pGL3-basic vector-promoter alone construct was made by cloning a short 51-bp fragment with B4B1R tails located upstream of the *CFTR* promoter insert (from −1825 to −1774 bp from the ATG translation start site; coordinates: 116 905 560–116 905 611). Inclusion of this 51 bp fragment in the promoter-only construct, but not in the element–promoter constructs did not affect the interpretation of the enhancing effects of the elements, as shown by analysis of a pair of promoter and promoter-element IV constructs that both lack this sequence yet gave identical levels of promoter activation by element IV (Supplementary Figure S3). As a positive control, we generated a construct containing a fragment of intron 1 of the *CFTR* gene. This fragment was previously shown to have significant enhancer activity (28). The intron 1 fragment was amplified using primers IA1R and TSR8 as described in ref. (28) and containing as well the B4B1R tails. We also generated a construct containing the segments with the highest enhancer activity from both elements III and IV as a fusion. To this end, we performed a fusion PCR (29) of elements IIIb and IVc, and cloned this fused fragment upstream of the *CFTR* promoter insert.

DNAse I mapping and analysis of histone modification patterns in the ENCODE regions

DHS data and histone modification data were generated as previously described (30,31). Data are available at http://genome.ucsc.edu/ENCODE/pilot.html, where detailed descriptions of the methods can also be found.

Transfection and luciferase assays

Transient transfections were performed on two monolayer cell lines, the *CFTR*-expressing Caco2 and non-*CFTR* expressing HepG2 cell line grown in 96-well plates to 70–80% confluency. In all transfection experiments the pGL3 constructs were cotransfected with 1/10 the amount of DNA of pRL-TK as a transfection control, using Effectene as transfection reagent, as indicated in the manufacturer's manual (Qiagen). Luciferase assays were carried out using the dual luciferase kit (Promega) on a 96-well plate reader. Each transfection experiment was carried out at least five times with individual constructs being assayed in duplicate in each experiment. Results are expressed as relative luciferase activity, with the pGL3-basic vector-promoter alone construct activity equal to 1. One-tailed *t*-tests were performed to test significance of the increased luciferase activity.

RESULTS

Identification of long-range interactions between the *CFTR* promoter and distant elements

Long-range looping interactions can be detected using 3C (13). 3C is a widely used method and the procedure has previously been described in detail (13,25,32,33). Briefly, 3C employs formaldehyde cross-linking to capture physically interacting chromatin segments. Cross-linked chromatin is then solubilized, digested and intra-molecularly ligated so that pairs of interacting genomic elements are converted into unique ligation products. Ligation products are then detected and quantified by semi-quantitative PCR (see 'Materials and Methods' section).

Here we used 3C as a discovery tool for *ab initio* identification of distant regulatory elements that interact specifically with the active *CFTR* promoter. We performed 3C with Caco2 cells, which express high levels of *CFTR*, and with GM06990 lymphoblastoid cells, which express very low levels of *CFTR* (Supplementary Figure S1) to identify genomic elements that physically associate with the *CFTR* promoter. We used a PCR primer located in the EcoRI fragment that contains the *CFTR* TSS and paired it with primers in restriction fragments throughout a 460-kb region surrounding the promoter (primer sequences available in Supplementary Table S1). This experimental setup will detect ligation (and thus long-range interaction) of the promoter fragment with any of the other restriction fragments. The results are shown in Figure 1A (3C data available in Supplementary Table S2). In GM06990 cells, we observe that the promoter fragment interacts most strongly with nearby restriction fragments and that interaction frequencies decrease precipitously for fragments located farther away. This inverse relationship between interaction frequency and genomic distance is expected for a flexible chromatin fiber in the absence of any specific long-range looping interactions (13,24,34–36). Specific long-range looping interactions would result in peaks of interaction frequency super-imposed upon this

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

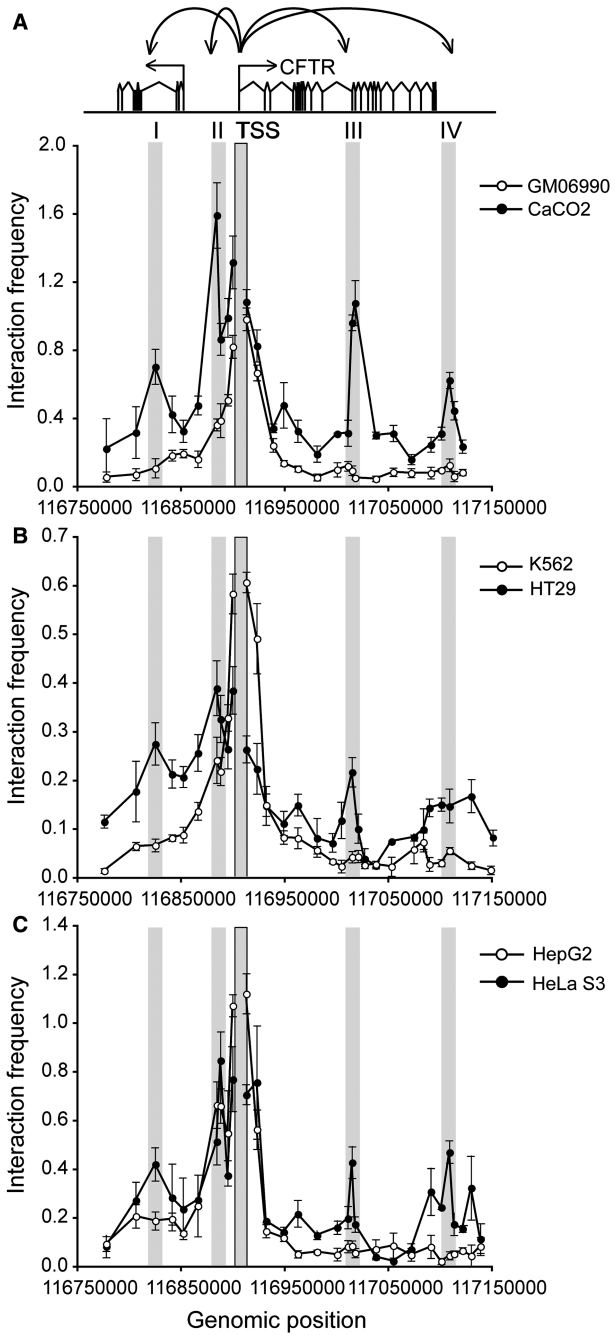


Figure 1. Chromosome conformation capture (3C) reveals four elements that interact with the *CFTR* transcription start site. All experiments used EcoRI to digest the chromatin. Interaction frequencies between the *CFTR* transcription start site and a 460 kb surrounding region are determined by 3C. Primer sequences can be found in Supplementary Table S1. 3C data can be found in Supplementary Table S2. (**A**) 3C profile of Caco2 cells (black circles), which express the *CFTR* gene, and GM06990 cells (white circles), which express the gene at low levels. (**B**) 3C profile of HT29 cells (black circles), which express the *CFTR* gene, and K562 cells (white circles) which do not express the gene. (**C**) 3C profile of HeLa S3 cells (black circles) which express the *CFTR* gene, and HepG2 cells (white circles), which do not express the gene.

background of interactions. We conclude that in GM06990 cells, the promoter does not engage in specific interactions with elements located in the surrounding 460 kb (Figure 1A).

In Caco2 cells, we also observe frequent interactions between the promoter fragment and nearby restriction fragments. In addition, we observe frequent interactions with several other restriction fragments, as evidenced by local peaks in interaction frequency above the background of non-specific associations. Two interacting restriction fragments are located 20 and 80 kb upstream of the TSS; one is located within the *CFTR* gene 109 kb downstream of the promoter in intron 11 and a fourth fragment is located 203 kb downstream of the TSS (15 kb downstream of the 3′-end of the gene).

To confirm these results we analyzed two additional pairs of cell lines, each composed of one cell line that expresses *CFTR* and one that does not. First, we chose HT29 cells, which express high levels of *CFTR* (37), and K562 cells, which do not express *CFTR* (as determined by RT-PCR, Supplementary Figure S1). We again performed 3C to analyze long-range interactions with the *CFTR* promoter fragment (Figure 1B). We observe that in both cell lines the promoter fragment frequently interacts with neighboring fragments, as expected, but these interactions are less frequent in HT29 cells than in K562 cells. We have previously observed that interaction frequencies around active promoters can be reduced, possibly as the result of local chromatin decondensation (24). Importantly, we found no evidence for specific long-range looping interactions in K562 cells, but we again observe local peaks in interaction frequency for multiple restriction fragments in *CFTR*-expressing HT29 cells. Importantly, these interacting restriction fragments are the same as those observed in Caco2 cells (Figure 1A). In addition, we observe a region encompassing several EcoRI restriction fragments around the 3′-end of the *CFTR* gene that all display increased interaction frequencies as compared to K562 cells, suggesting the presence of multiple interacting elements in that region.

Second, we analyzed HeLa S3 cells that express *CFTR* at a low level and HepG2 cells that do not express *CFTR* (as determined by RT-PCR, Supplementary Figure S1). In HeLa S3 cells we again detect specific long-range interactions between the *CFTR* promoter fragment and a set of distant fragments, whereas no specific long-range interactions were detected in HepG2 cells (Figure 1C). All interacting fragments correspond to the same set of fragments detected in Caco2 and HT29 cells, except one: in HeLa S3 cells the promoter fragment does not frequently interact with the fragment located at −20 kb. Instead, the promoter fragment associates more frequently with the restriction fragment immediately adjacent to this fragment, located at −17 kb. In HepG2 this −17 kb fragment also interacts with the promoter fragment somewhat more frequently than neighboring fragments. The two adjacent fragments located at −17 and −20 kb may contain two separate elements, or a single interacting region that displays cell type specific differences in the precise point of contact with the promoter. Finally, as in HT29 cells, we observe that in HeLa S3 cells the promoter fragment frequently interacts with several restriction fragments near and beyond the 3′ end of the gene.

Thus, multiple long-range interactions are observed between the *CFTR* promoter fragment and several

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

restriction fragments upstream and downstream of the gene as well as within an intron of the gene. Moreover, these interactions are observed only in *CFTR* expressing cells (Caco2, HT29 and HeLa S3) and not in non-expressing cells (K562 and HepG2). Interestingly, GM06990 cells do not display these long-range interactions, yet express *CFTR* at a very low, but detectable, level. Thus, the correlation between *CFTR* expression and looping between the promoter and these distal elements is not absolute (see 'Discussion' section). Analysis of digestion efficiency confirmed that these frequent interactions are not due to more efficient digestion of the corresponding restriction sites, consistent with many other studies that showed that digestion efficiency is not a major contributor to variation in 3C interaction frequencies [data not shown; (24,38–40)].

For further analyses, we focused on loci that are consistently found to interact with the *CFTR* promoter in all three *CFTR* expressing cell lines. These looping elements are located at −80, −20/−17, +109 and +203 kb from the TSS (highlighted by gray bars in Figure 1). From here on we refer to these looping elements with Roman numerals (I through IV, with element II containing the two adjacent restriction fragments located 20 and 17 kb upstream of the TSS). It is important to point out that other interacting elements may be present at other locations in and around the *CFTR* locus, as suggested by additional cell-line-specific peaks, e.g. around the 3′-end of the gene in HT29 cells. Consistently, one of these restriction fragments contains an element (located 6.8-kb downstream of the gene) that has recently been shown to bind CTCF and to also interact with the *CFTR* promoter in primary epididymis cells (23).

Three of the four identified elements (elements I, II and IV) had been previously identified, but only element II had been implicated in *CFTR* regulation. Element I may play a structural role but does not appear to be critical for *CFTR* regulation as constructs lacking this element faithfully reproduce the appropriate expression pattern of *CFTR* in mouse (41). Element II has previously been shown to play a role in *CFTR* regulation as deletion of this element reduces *CFTR* expression (41). We note that the ENCODE consortium recently identified sites that are bound by the CTCF protein in the restriction fragment directly adjacent to element II in Caco2 cells (J. Stamatoyannopoulos, unpublished results). The Harris lab also demonstrated CTCF binding to this element in Caco2 and Calu3 cells (42). This is interesting because CTCF has been directly implicated in mediating long-range looping interactions (43–45). In addition, DHSs, indicating the presence of putative regulatory elements, have been identified in or directly adjacent to both elements I and II (46). These data validate our 3C-based strategy for regulatory element discovery.

Long-range interactions between looping elements

Long-range regulatory elements not only interact with promoters, but in some cases have also been found to interact with other regulatory elements, e.g. in the case of the beta-globin locus (27,38,43). To determine whether any long-range interactions occur between the looping elements we identified above, we again employed 3C analysis.

First, we performed 3C analyses anchored on the elements I, II (anchored on the restriction fragment located 20-kb upstream of the TSS), III and IV in Caco2 and GM06990 cells (Figure 2). As expected, we find that these elements interact frequently with nearby restriction fragments in both cell lines. In addition, these elements interact frequently with the *CFTR* promoter fragment in Caco2 cells, but not in GM06990 cells. Elements I and II do appear to interact with elements III and IV, as minor peaks at the corresponding locations are visible (Figure 2A and C), although these interactions are considerably less frequent than their interaction with the promoter fragment. Interestingly, in Caco2 cells, elements III and IV are found to interact prominently with each other, as strong peaks of interaction frequencies are readily detected between the two restriction fragments (Figure 2E and G). In GM06990 cells we did not detect any obvious peaks in the 3C interaction profiles, suggesting that there are no specific looping interactions between these elements in that cell line. We note that interactions obtained with all four anchor elements are generally higher in Caco2 cells as compared to GM06990 cells throughout the region, which is as predicted for a locus that is more compact as a result of long-range looping interactions (34). We conclude that in Caco2 cells elements I–IV not only interact with the *CFTR* promoter fragment but also associate with each other, although somewhat less frequently.

We repeated the 3C analysis in HeLa S3 and HepG2 cells with the exception that for these cells the anchor for element II was on the restriction fragment located at −17 kb that we found to interact more frequently with the promoter fragment than the −20 kb restriction fragment (Figure 1C). We again observe that in HeLa S3 cells, but not in HepG2 cells, elements III and IV interact prominently with each other, as compared to background levels at other positions in the locus (Figure 2F and H). Elements I and II also display local peaks of interaction at the locations of the other looping elements (Figure 2B and D). The interaction between element I and IV is particularly prominent in HeLa S3 cells, whereas this interaction is much less pronounced in Caco2 cells.

We conclude that in *CFTR* expressing cells, elements I–IV not only interact with the promoter fragment but also with each other. In addition, the relative interaction frequency of these elements is somewhat different in the two cell lines.

Fine mapping of looping elements III and IV

Elements I and II are not further analyzed here, as they have been characterized before (41,42). Rather we focused on elements III and IV in Caco2 cells as they could represent novel regulatory elements that control the *CFTR* gene.

The above 3C analysis identified individual EcoRI restriction fragments that displayed the most prominent interactions with the promoter fragment. The size of

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

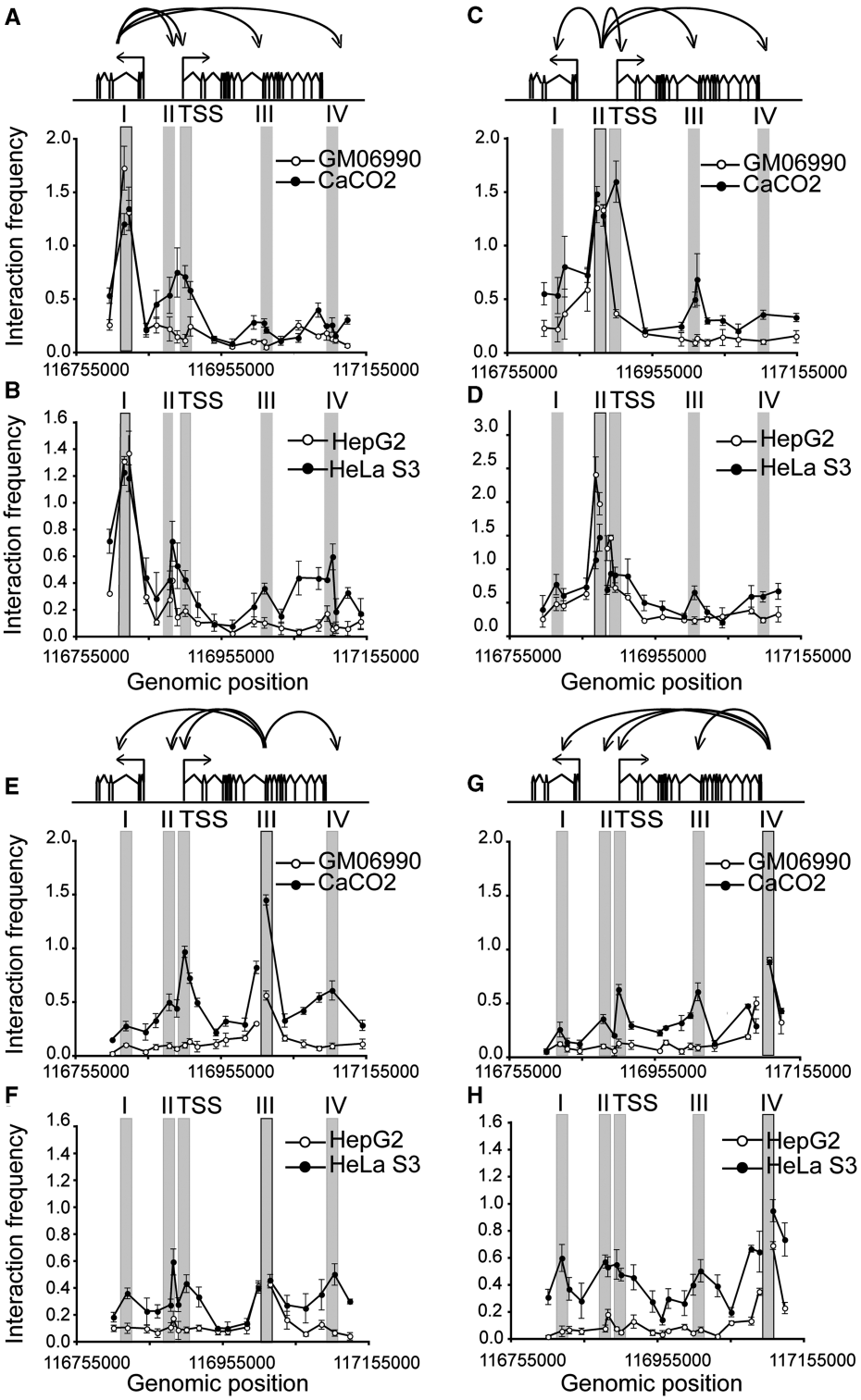


Figure 2. 3C interaction profiles for each of the four elements. All experiments used EcoRI to digest the chromatin. 3C data can be found in Supplementary Table S2. (**A**) The anchor point is set at element I in Caco2 and GM06990 cells, showing interactions between elements I and II, the transcription start site (TSS), and elements III and IV in Caco2 cells, with no notable interactions in GM06990 cells. (**B**) The anchor point is set at element I in HeLa S3 and HepG2 cells, showing interactions between element I and elements II, TSS, III and IV in HeLa S3 cells, with no notable interactions in HepG2 cells. (**C**) The anchor point is set at element II in Caco2 and GM06990 cells, showing interactions between elements II and I, TSS, elements III and IV in Caco2 cells, with no notable interactions in GM06990 cells. (**D**) The anchor point is set at element II in HeLa S3 and HepG2 cells, showing interactions between element II and elements I, TSS, III and IV in HeLa S3 cells, with no notable interactions in HepG2 cells. (**E**) The anchor point is set at element III in Caco2 and GM06990 cells, showing interactions between element III and elements I, II, TSS and element IV in Caco2 cells, with no notable interactions in GM cells. (**F**) The anchor point is set at element III in HeLa S3 and HepG2 cells, showing interactions between element III and elements I, II, TSS, and element IV in HeLa S3 cells, with no notable interactions in HepG2 cells. (**G**) The anchor point is set at element IV in Caco2 and GM06990 cells, showing interactions between element IV and elements I, II, TSS, and element III in Caco2 cells, with no notable interactions in GM06990 cells. (**H**) The anchor point is set at element IV in HeLa S3 and HepG2 cells, showing interactions between element IV and elements I, II, TSS, and element III in HeLa S3 cells, with no notable interactions in HepG2 cells.

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

these restriction fragments averages 4 kb, but the elements that mediate these interactions are most likely considerably smaller. In order to determine the position of the putative elements at higher resolution, we repeated the 3C analysis with another restriction enzyme, BsrGI. We chose this enzyme because it cuts the selected EcoRI fragments into smaller pieces. We then analyzed the interactions between the BsrGI fragment containing the TSS and several BsrGI fragments located around the positions of elements III and IV (Figure 3). We find a local peak of interaction between the promoter fragment and a 1066 bp fragment containing element III, located in the 3′-end of the original 3252 bp EcoRI fragment (Figures 3A and 4A). We also see a peak between the promoter fragment and a 1560 bp fragment, located near the 3′-end of the original 6069 EcoRI fragment containing element IV (Figures 3B and 4B). These interactions were not observed in GM06990 cells. These results confirm the presence of looping elements at these locations, and further define the positions of these elements. Based on the two independent 3C analyses performed with EcoRI and BsrGI, we conclude that element III is contained within an 806-bp BsrGI–EcoRI fragment and that element IV is contained within the 1560 bp BsrGI restriction fragment (Figure 4A).

Above we identified interactions between EcoRI fragments containing elements III and IV. We wanted to know if the interactions between elements III and IV are mediated by the same elements as the interactions between the promoter fragment and element III and IV, respectively, or whether these distinct interactions involved other elements contained within the EcoRI restriction fragments. Therefore, we tested whether the same BsrGI fragments that interact with the promoter fragment also interact with each other. 3C analysis confirmed that the BsrGI fragment containing element III interacts most prominently with the BsrGI fragment that contains element IV (Figure 3C). These results suggest that the same elements that interact with the promoter also interact with each other.

Functional analysis of looping elements III and IV

Next we used reporter assays to determine whether any of the looping elements display enhancer activity. For this experiment we tested subsections of the EcoRI fragments that we originally found to interact with the *CFTR* promoter fragment. This provided us with a complementary approach to the 3C analysis described above using BsrGI to narrow down the precise location of the functional elements. We generated a Gateway compatible pGL3-basic vector that contains the luciferase gene and an upstream Gateway recombination cassette (47). We used PCR to amplify the *CFTR* promoter and the looping elements identified above and added Gateway recombination tails (see 'Materials and Methods' section for genomic coordinates of these elements and the *CFTR* promoter fragment). Each of the looping elements was cloned upstream of the *CFTR* promoter, which, in a single Gateway recombination reaction, positioned the element and promoter immediately upstream of the luciferase coding region. We split each element fragment (defined as the EcoRI fragment that was found to interact with the *CFTR* promoter fragment) into smaller segments of 1.1–1.8 kb to allow locating the position of the regulatory element more precisely (Figure 3D): element III into two (IIIa and IIIb) and element IV into three segments (IVa, IVb and IVc). As a positive control we used the known enhancer located in intron 1 (28). Luciferase activity was measured in *CFTR*-expressing Caco2 cells and non-*CFTR*-expressing HepG2 cells and normalized to the level detected with the *CFTR* promoter alone (pGL3-basic vector-promoter alone construct; see 'Materials and Methods' section).

As shown in Figure 3E, we find that the known enhancer (intron 1) activates the *CFTR* promoter in both Caco2 cells and HepG2 cells, although to a significantly higher level in HepG2 cells. The strong activity of the intron 1 enhancer in HepG2 cells was unexpected, given that this enhancer has been reported to be specific to intestinal cells (28,48). Interestingly, we found that the DHS present at this enhancer in intestinal cells is also prominently present in HepG2 cells (J.A.S., unpublished results), confirming that this enhancer is active in these cells. Furthermore, only one fragment (IVc), encompassing the distal portion of the EcoRI fragment containing element IV, modestly activates the *CFTR* promoter to a similar extent in both cell types. Importantly, fragment IVc overlaps the BsrGI restriction fragment that most prominently interacted with the promoter fragment, allowing us to further narrow down the element's position to a 1002 bp locus (Figure 4B).

Elements III and IV act synergistically

The lack of observed enhancer activity for the other tested DNA fragments could arise because enhancer assays can present as false-negatives or because these elements may be involved in *CFTR* regulation in other ways. We considered an additional possibility. Our 3C analysis indicated that in Caco2 cells element III interacts not only with the *CFTR* promoter but also prominently with element IV. Therefore we hypothesized that element III acts in concert with element IV. To test this directly we generated constructs in which both elements were cloned upstream of the *CFTR* promoter–luciferase construct. Specifically, we tested the part of element III (IIIb) that overlaps the BsrGI–EcoRI restriction fragment that contains the looping element. Interestingly, we find that the presence of both elements IIIb and IVc results in a modest synergistic activation of the *CFTR* promoter, as measured by expression of the luciferase gene (Figure 3F). More importantly, this increase in luciferase expression occurs specifically in *CFTR* expressing Caco2 cells and not in HepG2 cells, indicating that element III has enhancer activity in the presence of the element IV. Thus, two elements separated by ~100 kb physically associate, and together synergistically activate the *CFTR* promoter specifically in Caco2 cells. We note that formally we cannot rule out an alternative explanation that in the context of the reporter construct element III on its own fails to activate the *CFTR* promoter due to its close proximity to the promoter.

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

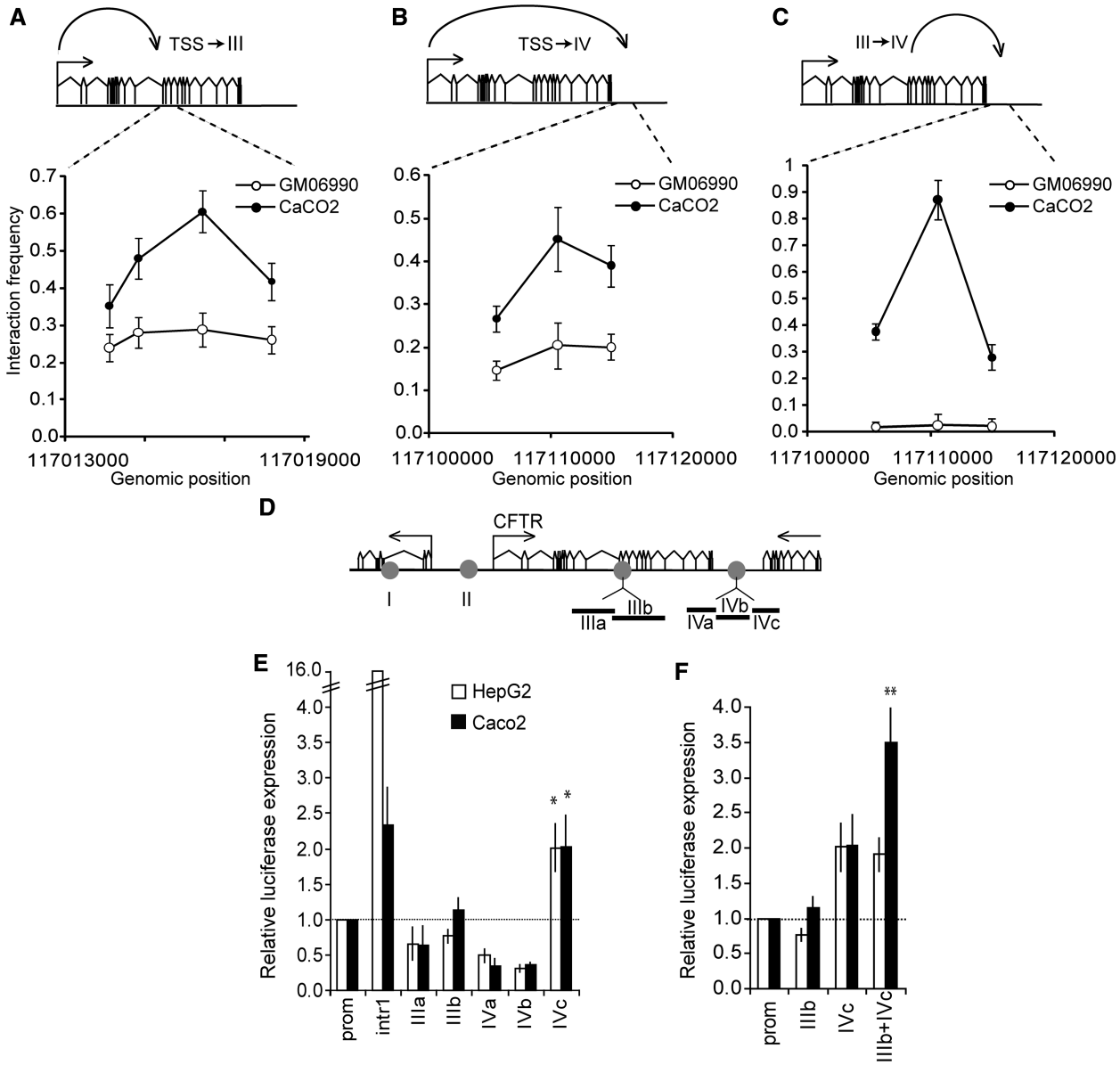


Figure 3. Fine mapping of the looping interactions between the *CFTR* promoter and elements III and IV, and luciferase assays showing that these two elements can synergistically activate the *CFTR* promoter. 3C primers and data can be found in Supplementary Tables S1 and S2. (**A–C**) 3C profiles using BsrGI to digest the chromatin focused on element III (A), anchored at the TSS. The interaction peak corresponds to a 1066 bp fragment. (B) 3C focused on element IV, anchored at the TSS. The interaction peak corresponds to a 1560 bp fragment. (C) 3C focused on element IV anchored at element III. The interaction peak corresponds to the same 1560 bp fragment that interacts with the *CFTR* TSS. (**D**) Outline of the fragments used in the luciferase assay. Element III was split into two fragments; element IV was split into three fragments. (**E**) Relative luciferase expression data, where 'Prom' refers to the fragment containing only the *CFTR* TSS and promoter region (pGL3-basic vector-promoter alone construct; see 'Materials and Methods' section), and was set as 1, and 'intr1' is a known enhancer element from intron 1 of the *CFTR* gene used as a positive control. Only element IVc showed enhanced luciferase expression compared to the promoter alone [$P<0.05$ (*t*-test) for both Caco2 and HepG2 cells]. (**F**) When elements IIIb and IVc were combined, a synergistic enhancer effect was observed only in the *CFTR*-expressing cell line Caco2 [$P<0.01$ (*t*-test)] when compared to element IVc alone in Caco2 cells. Asterisks indicate statistically significant differences from [Prom] (Panel E), or from [IIIb+IVc in HepG2] (Panel F). $*P<0.05$; $**P<0.01$.

Additional evidence for the presence of functional elements in the looping elements III and IV

The use of two enzymes in our 3C analysis, combined with the luciferase assays, allowed us to define the minimal interacting and functional regions containing elements III and IV. In Figure 4 we show, in modified UCSC genome browser shots, the EcoRI and BsrGI fragments that interacted most prominently with the promoter fragment. We also show the fragments that were tested in the luciferase assays. The minimal region of overlap between these fragments defines the positions of elements III and IV (indicated by vertical lines).

To obtain further evidence that the regions we identified here contain functional elements, we determined whether they display chromatin features that indicate the presence of regulatory elements. One hallmark of functional

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

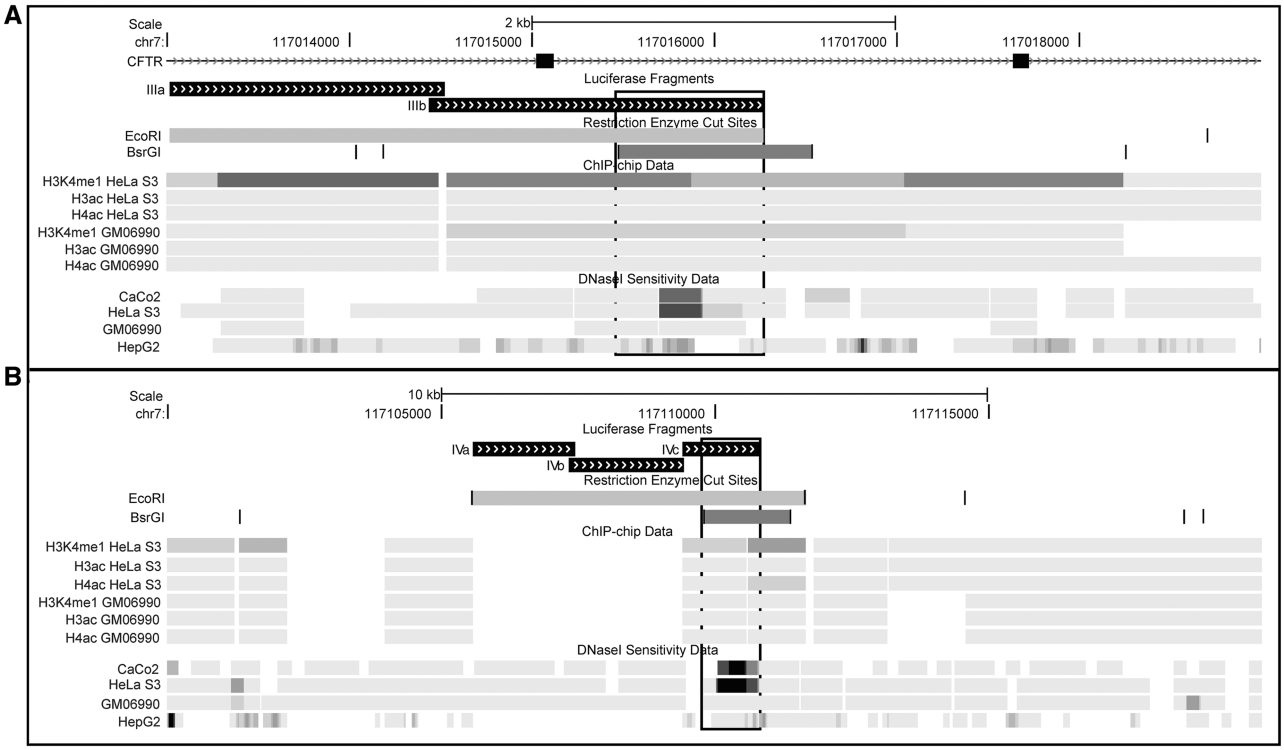


Figure 4. Data from the ENCODE consortium in the UCSC Genome Browser hg18 [(10); http://genome.ucsc.edu/ENCODE/]. Fragments used in the luciferase assay are indicated by black bars with white arrows. Restriction fragments that show 3C interaction peaks are outlined with grey boxes. A boxed outline indicates the narrowed-down functional region. Histone methylation, acetylation and DHS data are shown, with 'positive hits' shown as darker grey regions. (**A**) Region 117 013 000–117 019 000 on chromosome 7, which contains element III. (**B**) Region 117 100 000–117 120 000 on chromosome 7, which contains element IV.

elements is the presence of DHSs, corresponding to nucleosome free regions bound by regulatory protein complexes. In addition, regulatory elements, e.g. enhancers, have been found to display characteristic histone modifications including increased levels of acetylation and high levels of histone H3 monomethylation at the fourth lysine residue (H3K4Me1). The *CFTR* locus was selected for study by the ENCODE pilot project and as part of this consortium we generated extensive data related to chromatin structure and modification (publicly available at http://genome.ucsc.edu/ENCODE/). Figure 4 shows our ENCODE data representing DNase I hypersensitivity (in Caco2, GM06990 and HeLa S3 cells) and a variety of histone modifications (as detected in GM06990 and HeLa S3 cells) for the genomic regions containing elements III and IV. Strikingly, the minimal regions containing elements III and IV (indicated by vertical lines in Figure 4) coincide precisely with the presence of DHSs. These hypersensitive sites are found in *CFTR* expressing cells (Caco2 and HeLa S3), but not in GM06990 or HepG2 cells. Also, in HeLa S3 cells the regions around elements III and IV are enriched in H3K4Me1 and acetylated H4. Some H3K4Me1 is also found at element III in GM06990 cells.

These comparisons strongly suggest that elements III and IV contain functional elements, and indicate that our 3C-based element discovery approach identified *bona fide CFTR* regulatory elements.

DISCUSSION

We have employed 3C to identify long-range regulatory elements that may be involved in controlling the *CFTR* gene. We identified four elements that interact with the *CFTR* promoter specifically in *CFTR*-expressing cells. These looping elements appear to be gene regulatory elements based on the following observations. First, two of the elements had been previously identified as putative regulatory elements and one of these elements (element II) has been found to directly affect *CFTR* expression (28,41). Second, these elements contain DHSs and, in the case of elements III and IV patterns of histone modifications in *CFTR* expressing cells that have previously been found to be predictive of distal regulatory elements (6,7). Third, elements III and IV can synergistically activate the *CFTR* promoter in expressing cells indicating that these elements can function as transcriptional enhancers.

Our studies did not identify all previously discovered regulatory elements around the *CFTR* locus. There are several reasons for this. First, we did not analyze all restriction fragments throughout the 460 Kb surrounding the *CFTR* promoter. For instance, initially we did not analyze interactions between the *CFTR* promoter and a known enhancer that is located 10 kb downstream of the promoter in intron 1. The Harris lab has recently shown that this element interacts with and regulates the *CFTR* promoter (21). We have been able to confirm this looping interaction in subsequent 3C experiments using BsrGI

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

(Supplementary Figure S2). Second, we have focused on only those elements that consistently interact with the *CFTR* promoter in *CFTR* expressing cell lines (Caco2, HT29, HeLa S3). We did find that in some cell lines (e.g. HT29) there are additional elements that appear to frequently interact with the *CFTR* promoter, for instance a region just downstream of the gene. Consistently, the Harris lab has shown that the region contains additional elements that loop to the *CFTR* promoter in primary epididymis cells (23). Thus, we have been able to confirm previously detected regulatory elements (elements I and II, as well as the known enhancer in intron 1), thereby validating our approach, and we have discovered novel *CFTR* regulatory elements.

In this study we focused on elements III and IV. Further 3C experiments allowed us to narrow the regions of interaction of these elements down to ~1 kb. The results of the 3C mapping were fully consistent with reporter assays and both approaches identified the same putative regulatory elements. In addition, these elements coincide precisely with the presence of DHSs present in *CFTR* expressing cells and with regions enriched in histone modifications associated with enhancers (6). Combined, these independent lines of research provide strong evidence that *bona fide* *CFTR* regulatory elements were identified. In a very recent similar study the Harris lab identified the same elements III and IV as *CFTR* enhancers (49).

The long-range looping interactions described here are clearly correlated with expression of the *CFTR* gene. However, we note that the frequency of looping is not quantitatively related to the level of *CFTR* expression. The frequency of interaction between the *CFTR* promoter region and elements III and IV is quite comparable in Caco2 cells and HeLa S3 cells, despite the fact that Caco2 cells express the gene at a much higher level (Supplementary Figure S1). This indicates that these long-range interactions are not sufficient for high levels of expression. One possible explanation is that in Caco2 cells additional transcription factors and/or co-regulators bind these elements to further activate the gene. We also point out that the long-range looping interactions between the *CFTR* promoter and elements I–IV are not required for a low level of expression, as GM06990 cells express very low, but detectable, levels of *CFTR* while none of these long-range interactions are detected. It is possible that this low level is simply due to basal promoter activity or is the result of long-range interactions with other distal regulatory element. For instance, we find that in GM06990 cells the promoter is interacting with the enhancer element in intron 1 (Supplementary Figure S2). Finally, we cannot rule out the possibility that immortalized cell lines such as GM06990 display somewhat altered expression patterns as compared to the primary cells of origin.

The four elements we uncovered in our 3C analysis interact with the *CFTR* promoter in all three *CFTR* expressing cells. These elements also all interact with each other, suggesting they form a single cluster of interacting chromatin segments (indicated by the grey circular area in Figure 5). We do note that the frequency of interaction between pairs of elements varies (Figure 2), and is also somewhat different between cell lines. Thus, although the overall conformation of the locus is comparable in Caco2 and HeLa S3 cells, there may be slight differences in the frequencies with which long-range interaction occur.

Our reporter assays suggest that elements III and IV activate the *CFTR* promoter in a synergistic manner. This is interesting because elements III and IV directly associate with each other, which could provide a mechanism by which widely spaced elements can coordinately control gene expression. It is likely that specific protein complexes associate with these elements and mediate interactions between them and with the *CFTR* promoter. Previous work from the Harris Lab has started to identify some proteins that could potentially associate with the DHS that is present in element IV. They found that this element contains two C→T polymorphisms, and these mutations alter the DNA binding of ARP-1 and HNF-4 *in vitro* (50). They also found that C/EBP, CREB/ATF and AP-1 transcription factors can bind this region (22). In addition, the transcription factor Myc can bind to element III [(10), Iyer lab (University of Texas, Austin); http://genome.ucsc.edu/ENCODE/]. Future studies will be aimed at identifying these protein complexes in more detail so that the mechanisms of long-range associations and their effect on transcription can be further dissected.

Identification of extra-genic elements that affect expression of *CFTR* will not only provide basic insights into spatio-temporal regulation of this important gene, it may also be important for genetic diagnosis of cystic fibrosis. A significant number of patients with cystic fibrosis symptoms do not appear to carry mutations in the *CFTR* exons or promoter, suggesting that extra-genic or intronic mutations may be present, e.g. in long-range acting gene regulatory elements. Mutations in regulatory elements can also result in disease characteristics that are distinct form the full cystic fibrosis phenotype, such as

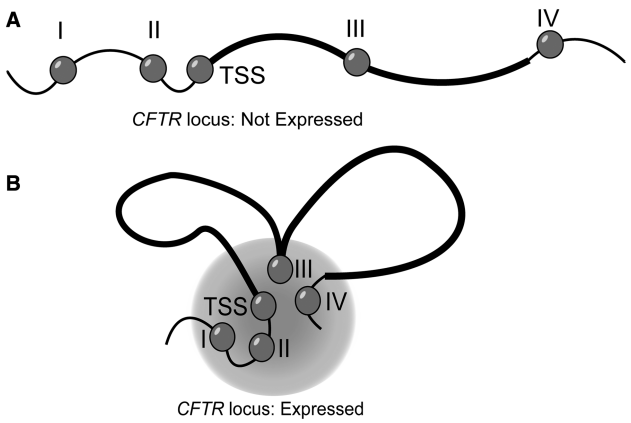


Figure 5. Different models of the conformation of the *CFTR* locus. (**A**) The *CFTR* locus in a cell line in which the gene is not expressed. The locus assumes a linear conformation without specific long-range interactions. Thick black line indicates the CFTR gene. (**B**) The *CFTR* locus in a cell line that does express the gene. All four elements interact with the promoter, and with each other. The grey circle symbolizes the spatial domain containing the clustered promoter and regulatory elements.

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

congenital bilateral absence of the vas deferens. In the absence of information of the positions of distant regulatory elements it is not feasible to screen for mutations in a very large genomic region. Thus, identification of *CFTR* regulatory elements, as we have described here, provides new targets for mutation screening. Furthermore, gene therapy approaches for cystic fibrosis could benefit from knowledge of gene regulatory elements by including such elements in *CFTR* gene targeting constructs.

We have shown that 3C technology can be used to discover novel regulatory elements throughout gene loci. A variety of 3C adaptations have recently been developed that allow large-scale detection of chromatin looping interactions (27,51–53). Using these technologies, it may become possible to map the regulatory elements that control genes throughout the genome.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dekker Lab members and Marian Walhout for fruitful discussions and for critical reading of the manuscript.

FUNDING

National Institutes of Health (grant HG003143 to J.D., grant HG004592 to J.S., grant U01 HG003168 to I.D.); Keck Foundation (to J.D.); Cystic Fibrosis Foundation (to J.D.); Wellcome Trust (to I.D.). Funding for open access charge: National Institutes of Health (HG003143 to J.D.).

Conflict of interest statement. None declared.

REFERENCES

1. Kleinjan,D.A. and van Heyningen,V. (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am. J. Hum. Genet.*, **76**, 8–32.
2. West,A.G. and Fraser,P. (2005) Remote control of gene transcription. *Hum. Mol. Genet.*, **14**, R101–111.
3. Dekker,J. (2008) Gene regulation in the third dimension. *Science*, **319**, 1793–1794.
4. Kleinjan,D.A., Seawright,A., Schedl,A., Quinlan,R.A., Danes,S. and van Heyningen,V. (2001) Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum. Mol. Genet.*, **10**, 2049–2059.
5. D'haene,B., Attanasio,C., Beysen,D., Dostie,J., Lemire,E., Bouchard,P., Field,M., Jones,K., Lorenz,B., Menten,B. *et al.* (2009) Disease-causing 7.4 kb cis-Regulatory deletion disrupting conserved non-coding sequences and their interaction with the FOXL2 promotor: implications for mutation screening. *PLoS Genet.*, **5**, e1000522.
6. Heintzman,N.D., Stuart,R.K., Hon,G., Fu,Y., Ching,C.W., Hawkins,R.D., Barrera,L.O., Van Calcar,S., Qu,C., Ching,K.A. *et al.* (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.*, **39**, 311–318.
7. Heintzman,N.D., Hon,G.C., Hawkins,R.D., Kheradpour,P., Stark,A., Harp,L.F., Ye,Z., Lee,L.K., Stuart,R.K., Ching,C.W. *et al.* (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, **459**, 108–112.
8. Maston,G.A., Evans,S.K. and Green,M.R. (2006) Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.*, **7**, 29–59.
9. Visel,A., Bristow,J. and Pennacchio,L.A. (2007) Enhancer identification through comparative genomics. *Semin. Cell Dev. Biol.*, **18**, 140–152.
10. ENCODE-consortium. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, **447**, 799–816.
11. Visel,A., Rubin,E.M. and Pennacchio,L.A. (2009) Genomic views of distant-acting enhancers. *Nature*, **461**, 199–205.
12. de Laat,W. and Grosveld,F. (2003) Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.*, **11**, 447–459.
13. Dekker,J., Rippe,K., Dekker,M. and Kleckner,N. (2002) Capturing chromosome conformation. *Science*, **295**, 1306–1311.
14. Yoshimura,K., Nakamura,H., Trapnell,B.C., Dalemans,W., Pavirani,A., Lecocq,J.P. and Crystal,R.G. (1991) The cystic fibrosis gene has a "housekeeping"-type promoter and is expressed at low levels in cells of epithelial origin. *J. Biol. Chem.*, **266**, 9140–9144.
15. Pittman,N., Shue,G., LeLeiko,N.S. and Walsh,M.J. (1995) Transcription of cystic fibrosis transmembrane conductance regulator requires a CCAAT-like element for both basal and cAMP-mediated regulation. *J. Biol. Chem.*, **270**, 28848–28857.
16. McDonald,R.A., Matthews,R.P., Idzerda,R.L. and McKnight,G.S. (1995) Basal expression of the cystic fibrosis transmembrane conductance regulator gene is dependent on protein kinase A activity. *Proc. Natl Acad. Sci. USA*, **92**, 7560–7564.
17. Matthews,R.P. and McKnight,G.S. (1996) Characterization of the cAMP response element of the cystic fibrosis transmembrane conductance regulator gene promoter. *J. Biol. Chem.*, **271**, 31869–31877.
18. Romey,M.C., Pallares-Ruiz,N., Mange,A., Mettling,C., Peytavi,R., Demaille,J. and Claustres,M. (2000) A naturally occurring sequence variation that creates a YY1 element is associated with increased cystic fibrosis transmembrane conductance regulator gene expression. *J. Biol. Chem.*, **275**, 3561–3567.
19. McCarthy,V.A. and Harris,A. (2005) The CFTR gene and regulation of its expression. *Pediatr. Pulmonol.*, **40**, 1–8.
20. Mouchel,N., Henstra,S.A., McCarthy,V.A., Williams,S.H., Phylactides,M. and Harris,A. (2004) HNF1alpha is involved in tissue-specific regulation of CFTR gene expression. *Biochem. J.*, **378**, 909–918.
21. Ott,C.J., Suszko,M., Blackledge,N.P., Wright,J.E., Crawford,G.E. and Harris,A. (2009) A complex intronic enhancer regulates expression of the CFTR gene by direct interaction with the promoter. *J. Cell Mol. Med.*, **13**, 680–692.
22. Nuthall,H.N., Moulin,D.S., Huxley,C. and Harris,A. (1999) Analysis of DNase-I-hypersensitive sites at the 3′ end of the cystic fibrosis transmembrane conductance regulator gene (CFTR). *Biochem. J.*, **341**, 601–611.
23. Blackledge,N.P., Ott,C.J., Gillen,A.E. and Harris,A. (2009) An insulator element 3′ to the CFTR gene binds CTCF and reveals an active chromatin hub in primary cells. *Nucleic Acids Res.*, **37**, 1086–1094.
24. Gheldof,N., Tabuchi,T.M. and Dekker,J. (2006) The active FMR1 promoter is associated with a large domain of altered chromatin conformation with embedded local histone modifications. *Proc. Natl Acad. Sci. USA*, **103**, 12463–12468.
25. Miele,A., Gheldof,N., Tabuchi,T.M., Dostie,J. and Dekker,J. (2006) Mapping chromatin interactions by chromosome conformation capture (3C). In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds), *Current Protocols in Molecular Biology*, Vol. Supplement 74. John Wiley & Sons, Hoboken, NJ, pp. 21.11.1–21.11.20.
26. Dostie,J. and Dekker,J. (2007) Mapping networks of physical interactions between genomic elements using 5C technology. *Nat. Protoc.*, **2**, 988–1002.

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011

27. Dostie,J., Richmond,T.A., Arnaout,R.A., Selzer,R.R., Lee,W.L., Honan,T.A., Rubio,E.H., Krumm,A., Lamb,J., Nusbaum,C. *et al.* (2006) Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.*, **16**, 1299–1309.
28. Smith,A.N., Barth,M.L., McDowell,T.L., Moulin,D.S., Nuthall,H.N., Hollingsworth,M.A. and Harris,A. (1996) A regulatory element in intron 1 of the cystic fibrosis transmembrane conductance regulator gene. *J. Biol. Chem.*, **271**, 9947–9954.
29. Heckman,K.L. and Pease,L.R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.*, **2**, 924–932.
30. Dorschner,M.O., Hawrylycz,M., Humbert,R., Wallace,J.C., Shafer,A., Kawamoto,J., Mack,J., Hall,R., Goldy,J., Sabo,P.J. *et al.* (2004) High-throughput localization of functional elements by quantitative chromatin profiling. *Nat. Methods*, **1**, 219–225.
31. Koch,C.M., Andrews,R.M., Flicek,P., Dillon,S.C., Karaöz,U., Clelland,G.K., Wilcox,S., Beare,D.M., Fowler,J.C., Couttet,P. *et al.* (2007) The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res.*, **17**, 691–707.
32. Splinter,E., Grosveld,F. and de Laat,W. (2004) 3C technology: analyzing the spatial organization of genomic loci in vivo. *Methods Enzymol.*, **375**, 493–507.
33. Simonis,M., Kooren,J. and de Laat,W. (2007) An evaluation of 3C-based methods to capture DNA interactions. *Nat. Methods*, **4**, 895–901.
34. Rippe,K. (2001) Making contacts on a nucleic acid polymer. *Trends Biochem. Sci.*, **26**, 733–740.
35. Dekker,J. (2006) The 3 C's of chromosome conformation capture: controls, controls, controls. *Nat. Methods*, **3**, 17–21.
36. Dekker,J. (2008) Mapping in vivo chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. *J. Biol. Chem.*, **283**, 34532–34540.
37. Sood,R., Bear,C., Auerbach,W., Reyes,E., Jensen,T., Kartner,N., Riordan,J.R. and Buchwald,M. (1992) Regulation of CFTR expression and function during differentiation of intestinal epithelial cells. *EMBO J.*, **11**, 2487–2494.
38. Tolhuis,B., Palstra,R.J., Splinter,E., Grosveld,F. and de Laat,W. (2002) Looping and Interaction between Hypersensitive Sites in the Active beta-globin Locus. *Mol. Cell*, **10**, 1453–1465.
39. Dekker,J. (2007) GC- and AT-rich chromatin domains differ in conformation and histone modification status and are differentially modulated by Rpd3p. *Genome Biol.*, **8**, R116.
40. Miele,A., Bystricky,K. and Dekker,J. (2009) Yeast silent mating type loci form heterochromatic clusters through silencer protein-dependent long-range interactions. *PLoS Genet.*, **5**, e1000478.
41. Nuthall,H.N., Vassaux,G., Huxley,C. and Harris,A. (1999) Analysis of a DNase I hypersensitive site located -20.9 kb upstream of the CFTR gene. *Eur. J. Biochem.*, **266**, 431–443.
42. Blackledge,N.P., Carter,E.J., Evans,J.R., Lawson,V., Rowntree,R.K. and Harris,A. (2008) CTCF mediates insulator function at the CFTR locus. *Biochem. J.*, **408**, 267–275.
43. Splinter,E., Heath,H., Kooren,J., Palstra,R.J., Klous,P., Grosveld,F., Galjart,N. and de Laat,W. (2004) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.*, **20**, 2349–2354.
44. Kurukuti,S., Tiwari,V.K., Tavoosidana,G., Pugacheva,E., Murrell,A., Zhao,Z., Lobanenkov,V., Reik,W. and Ohlsson,R. (2006) CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl Acad. Sci. USA*, **103**, 10684–10689.
45. Phillips,J.E. and Corces,V.G. (2009) CTCF: master weaver of the genome. *Cell*, **137**, 1194–1211.
46. Smith,A.N., Wardle,C.J. and Harris,A. (1995) Characterization of DNase I hypersensitive sites in the 120kb 5′ to the CFTR gene. *Biochem. Biophys. Res. Commun.*, **211**, 274–281.
47. Deplancke,B., Dupuy,D., Vidal,M. and Walhout,A.J. (2004) A gateway-compatible yeast one-hybrid system. *Genome Res.*, **14**, 2093–2101.
48. Rowntree,R.K., Vassaux,G., McDowell,T.L., Howe,S., McGuigan,A., Phylactides,M., Huxley,C. and Harris,A. (2001) An element in intron 1 of the CFTR gene augments intestinal expression in vivo. *Hum. Mol. Genet.*, **10**, 1455–1464.
49. Ott,C.J., Blackledge,N.P., Kerschner,J.L., Leir,S.H., Crawford,G.E., Cotton,C.U. and Harris,A. (2009) Intronic enhancers coordinate epithelial-specific looping of the active CFTR locus. *Proc. Natl Acad. Sci. USA*, **106**, 19934–19939.
50. Rowntree,R. and Harris,A. (2002) DNA polymorphisms in potential regulatory elements of the CFTR gene alter transcription factor binding. *Hum. Genet.*, **111**, 66–74.
51. Zhao,Z., Tavoosidana,G., Sjölinder,M., Göndör,A., Mariano,P., Wang,S., Kanduri,C., Lezcano,M., Sandhu,K.S., Singh,U. *et al.* (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.*, **38**, 1341–1347.
52. Simonis,M., Klous,P., Splinter,E., Moshkin,Y., Willemsen,R., de Wit,E., van Steensel,B. and de Laat,W. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.*, **38**, 1348–1354.
53. Lieberman-Aiden,E., van Berkum,N.L., Williams,L., Imakaev,M., Ragoczy,T., Telling,A., Amit,I., Lajoie,B.R., Sabo,P.J., Dorschner,M.O. *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, **326**, 289–293.

Downloaded from nar.oxfordjournals.org at Medical Center Library on April 26, 2011