Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of the 10-23 DNA enzyme using a combinatorial screen of paired oligonucleotides

Jacek Nowakowski,[a,b] Peter J. Shim,[b] Gerald F. Joyce[a,b] and C. David Stout[b]*

[a]Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA, and [b]Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA

Correspondence e-mail: dave@scripps.edu

Received 3 June 1999
Accepted 2 August 1999

One of the most difficult steps in the X-ray crystallography of nucleic acids is obtaining crystals that diffract to high resolution. The choice of the nucleotide sequence has proven to be more important in producing high-quality crystals than the composition of the crystallization solution. This manuscript describes a systematic procedure for identifying the optimal sizes of a multi-stranded nucleic acid complex which provide high-quality crystals. This approach was used to crystallize the *in vitro* evolved 10-23 DNA enzyme complexed with its RNA substrate. In less than two months, 81 different enzyme–substrate complexes were generated by combinatorial mixing and annealing of complementary oligonucleotides which differed in length, resulting in duplexes of varying length, with or without nucleotide overhangs. Each of these complexes was screened against a standard set of 48 crystallization conditions and evaluated for crystal formation. The screen resulted in over 40 crystal forms, the best of which diffracted to 2.8 Å resolution when exposed to a synchrotron X-ray source.

1. Introduction

X-ray structure determination of complex nucleic acids remains severely hampered by the difficulty of obtaining single crystals of high diffraction quality. The formation of an ordered three-dimensional crystal lattice requires specific intermolecular packing interactions between neighboring molecules. The negatively charged surfaces of nucleic acids are structurally monotonous and offer many sites for non-specific packing contacts, which can introduce defects into the crystal lattice. In addition, many nucleic acids are conformationally heterogeneous in solution, further contributing to crystalline disorder. The success of a nucleic acid crystallization experiment is often determined by the choice of the primary sequence of the molecule, which may favor a particular conformation and set of packing interactions which result in well ordered crystals (Anderson *et al.*, 1996).

Traditionally, one of the most successful methods for improving the quality of nucleic acid crystals has been the 'helix-engineering' approach, in which the length and sequence of double helices present in the molecule are varied in search of favorable crystallization properties (Jordan *et al.*, 1985). A double helix is an obvious target for sequence modification because the ends of helices are commonly involved in crystal packing contacts mediated by base-stacking interactions. The sequence of the terminal nucleotides determines the strength of stacking and the length of the helix dictates the distance between neighboring molecules in the crystal lattice. Inclusion of overhangs at the ends of helices can result in intermolecular base pairing to the lattice neighbor or

may stabilize the end of the helix through hydrogen-bonding interactions within either the major or minor groove (Cruse *et al.*, 1994). The helix-engineering method has been applied successfully to the crystallization of protein–DNA complexes (Aggarwal *et al.*, 1988; Jordan *et al.*, 1985; Schultz *et al.*, 1990), small structured oligonucleotides (Anderson *et al.*, 1996) and ribozymes (Scott *et al.*, 1995).

The identification of a favorable nucleic acid sequence requires the screening of a large number of nucleic acid constructs which vary with respect to helix length and the composition of any overhanging nucleotides. An elegant and efficient strategy for generating a pool of such molecules was devised for the crystallization of the CAP protein complexed with its DNA target (Schultz *et al.*, 1990). In this approach, five complementary DNA strands were annealed in a combinatorial fashion, resulting in 25 DNA duplexes which differed in length and the presence or absence of overhanging nucleotides at either end of the helix. Each of these DNA duplexes was complexed with the protein and subjected to a variety of crystallization conditions. Although this approach only allowed for the generation of DNAs with symmetrical ends, it proved successful in generating high-quality crystals of the CAP–DNA complex.

The present study involves application of a combinatorial screen of paired oligonucleotides to an all-nucleic acid system consisting of the *in vitro* evolved 10-23 DNA enzyme (Santoro & Joyce, 1997) complexed with an RNA substrate. Unlike the 'modular DNA' approach, the method described here generates all possible combinations of arm lengths and overhanging nucleotides, allowing a larger number of molecules to be tested. In this study, 81 different DNA–RNA complexes were synthesized and tested against 48 different crystallization conditions. In addition to identifying sequences with the best crystallization properties, the large number of crystallization trials provided estimates of the number of optimal sequences in the starting pool of molecules and allowed assessment of the productivity of various crystallization conditions.

The 10-23 DNA enzyme is composed of a single DNA strand that catalyzes the sequence-specific cleavage of a target RNA (Fig. 1). The cleavage reaction is dependent on a divalent metal cation and occurs between an unpaired purine and paired pyrimidine residue within the bound substrate (Fig. 1). The 10-23 enzyme is highly specific for the sequence of the substrate RNA and exhibits a catalytic efficiency exceeding that of all other known nucleic acid catalysts (Santoro & Joyce, 1998). The substrate-recognition domains can be made complementary to almost any target RNA. Thus, this enzyme is a general-purpose endoribonuclease. Its small size, ease of synthesis and high stability make it a useful tool for the manipulation of RNA, both *in vitro* (Pyle *et al.*, 1999; Unrau & Bartel, 1998) and *in vivo* (Warashina *et al.*, 1999).

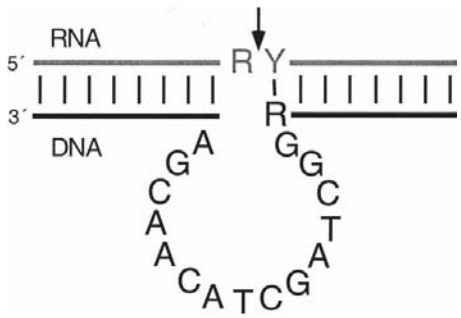

Figure 1
Secondary structure of the 10-23 DNA enzyme. The DNA strand is shown in black and the RNA strand in gray. R = G or A; Y = C or U. The arrow indicates the cleavage site. The 15 bulged nucleotides within the DNA strand comprise the catalytic domain of the enzyme.

Earlier attempts to crystallize the 10-23 DNA enzyme demonstrated that the length of the two substrate-recognition domains was extremely important for obtaining high-quality crystals. The enzyme–substrate complex containing six base pairs in each of the two recognition domains and a single guanosine overhang at the 5′ end of the DNA strand yielded large hexagonal crystals which diffracted to 3.0 Å. These crystals allowed determination of the structure of the complex by multiple isomorphous replacement employing chemically synthesized heavy-atom derivatives (Nowakowski *et al.*, 1999). Surprisingly, the 1:1 DNA–RNA complex present in solution rearranged in the crystal to form a non-catalytic structure composed of two strands of DNA and two strands of RNA (Nowakowski *et al.*, 1999). Formation of this alternative structure was made possible by a six-nucleotide palindromic sequence within the catalytic domain of the enzyme which facilitated dimerization of the DNA strand. Although the detailed mechanism of this rearrangement is not known, it is likely to involve opening of one of the substrate-recognition domains followed by cross-pairing between two enzyme–substrate complexes.

The sequence of the palindrome cannot be changed without reducing the activity of the enzyme. However, it should be possible to inhibit dimer formation by increasing the thermodynamic stability of the substrate-recognition domains. Accordingly, the length of each domain was increased from six to at least eight Watson–Crick pairs. In order to identify complexes that were readily crystallizable, the combinatorial screening method was designed to explore the preferred lengths of the two substrate-recognition domains. In less than two months, the screen resulted in over 40 crystal forms, the best of which diffracted to 2.8 Å resolution when exposed to a synchrotron X-ray source.

2. Materials and methods

RNA oligonucleotides were synthesized on a Pharmacia LKB Gene Assembler Special from commercially available phosphoramidites (Pharmacia). Each oligonucleotide was prepared on a 1.0 µmol scale using 5-ethylthio-1-*H*-tetrazole (Glen Research) as a coupling reagent, with typical yields of at least 95%. Molecules were cleaved from the solid support at 338 K using a 1:1(*v*/*v*) mixture of NH_3(aq)/$MeNH_2$, followed by deprotection of the 2′ hydroxyl groups and *n*-butanol precipitation (Wincott *et al.*, 1995). All DNA oligonucleotides were obtained from Operon Technologies (Alameda, CA). The following RNA oligonucleotides were used in the screen:

10-10, 5′-r(AA*x*AG)-3′; 10-9, 5′-r(AA*x*A)-3′; 10-8, 5′-r(AA*x*)-3′; 9-10, 5′-r(A*x*AG)-3′; 8-10, 5′-r(*x*AG)-3′; 9-9, 5′-r(A*x*A)-3′; 9-8, 5′-r(A*x*)-3′; 8-9, 5′-r(*x*A)-3′; 8-8, 5′-*x*-3′, where *x* is 5′-r(GGAGAGAG)d(A)r(UGGGUGCG)-3′. The following DNA oligonucleotides were used: 10-10, 5′-d(CT*y*TT)-3′; 10-9, 5′-d(T*y*TT)-3′; 10-8, 5′-d(*y*TT)-3′; 9-10, 5′-d(CT*y*T)-3′; 8-10, 5′-d(CT*y*)-3; 9-9, 5′-d(T*y*T)-3′; 9-8, 5′-d(*y*T)-3′; 8-9, 5′-d(T*y*)-3′; 8-8, 5′-*y*-3′, where *y* is 5′-d(CGCACCCAGGCTAGCTACAACGACTCTCTCC)-3′.

Crude oligonucleotides were purified in a denaturing 20% polyacrylamide/8 *M* urea gel, electroeluted and ethanol precipitated. The pellets were dissolved in 500 µl of deionized water and dialyzed for 48–72 h against pure water at 277 K. Dialyzed material was filtered through a 0.2 µm syringe filter, dried and stored frozen in water. RNA and DNA strands were mixed in a 1:1 molar ratio, lyophilized and dissolved in annealing buffer containing 2 m*M* $MgCl_2$ and 5 m*M* sodium cacodylate (pH 7.0). The solution was incubated at 343 K for 10 min and gradually cooled to room temperature over 12 h. The concentration of the annealed complex was typically 0.4 m*M*. 20 nmol of material was the minimum amount required to set up a screen with 48 crystallization drops.

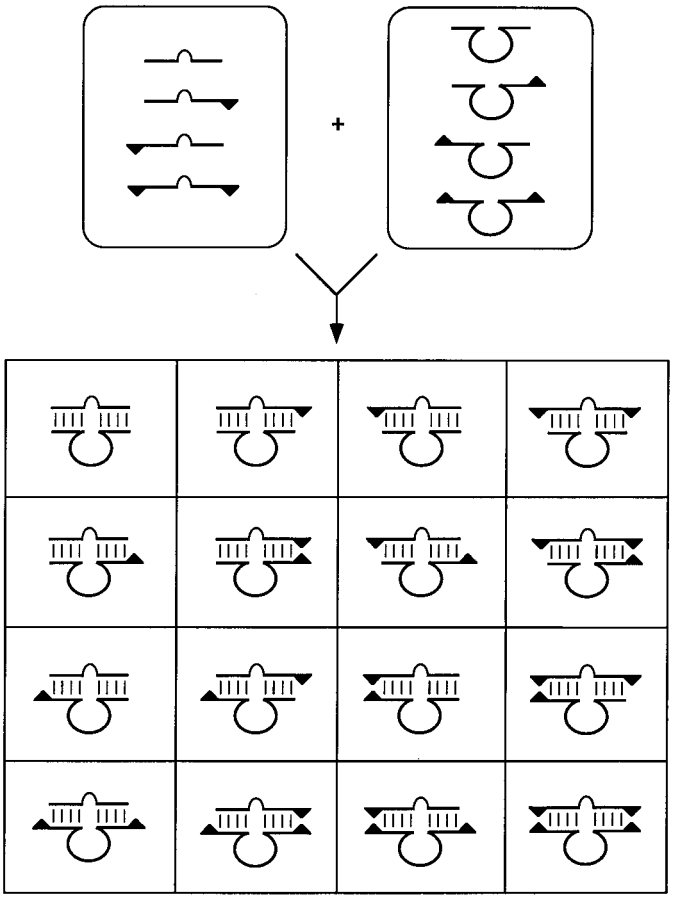

Figure 2
Scheme for combinatorial screening of paired oligonucleotides. Single-stranded oligonucleotides of different lengths are mixed individually, resulting in a large number of paired complexes. Triangles represent nucleotides in the variable regions which may be either paired or overhanging.

Crystallization was carried out by vapor diffusion using the sitting-drop method. The optimal volume of the drops was found to be 2 µl. Smaller drops resulted in faster nucleation rates and produced smaller crystals, but would still be useful for screening if only small amounts of material were available. 1 µl of nucleic acid solution was mixed with 1 µl of reservoir solution and placed on a plastic pedestal in a 24-drop Cryschem plate (Hampton Research). The typical volume of the reservoir liquor was 1 ml. In order to prevent evaporation of low molecular-weight alcohols present in many of the crystallization solutions, the plate was sealed with Crystal Clear plastic tape (Hampton Research). The plates were maintained at 297 K for at least two weeks before evaluation of crystal growth. The crystals were examined with a binocular microscope (40×) equipped with a polarizable filter.

The diffraction properties of crystals were evaluated at room temperature using Cu *Kα* radiation from a rotating-anode X-ray generator equipped with a graphite monochromator and a Siemens multiwire area detector. Crystals that diffracted well at room temperature were evaluated at 100 K using synchrotron radiation. Depending on the composition of the crystallization mixture, the crystals were either frozen directly in a stream of cold nitrogen or transferred briefly to the mother liquor containing increasing amounts of cryoprotectant and then flash-frozen in liquid nitrogen. The cryoprotectant was either MPD or glycerol, added in 10% increments to the mother liquor. Crystals were evaluated on beamline 5.0.2 at the Advanced Light Source (ALS) at Lawrence Berkeley Laboratory.

3. Results

3.1. Design of the combinatorial screen

The concept of the combinatorial screen of paired strands is illustrated in Fig. 2. Each strand is synthesized in several different lengths and purified individually. Each molecule of one strand is then mixed with each molecule of the other strand so that all combinations of arm lengths and overhangs are generated. Each of these complexes is then subjected to a standard set of crystallization conditions and evaluated for crystal formation. Because the crystallization solutions are identical for the various complexes, this method searches for the best combination of paired strands rather than the best crystallization conditions. Once the best complexes or sequences have been identified, the crystallization conditions can be optimized to increase the size and quality of the crystals.

Mixing oligonucleotides of different lengths results in duplexes which contain either blunt ends or nucleotide overhangs at either end of either strand. Depending on the range of strand lengths that are employed, overhangs of one, two or more nucleotides can be generated. In its simplest form, the screen samples only different arm lengths and not the primary sequence of the paired oligonucleotides. It is possible to vary the sequence as well, but the number of resulting complexes would increase enormously, making it difficult to complete the

screen. Previous experience suggests that arm-length variation alone generates sufficient diversity to provide molecules which result in high-quality crystals. The primary advantage of the combinatorial mixing strategy is the greatly reduced time needed to synthesize individual strands compared with the time which would be required to synthesize individual complexes. Mixing n forms of one strand with m forms of the other generates a total of $n \times m$ complexes, but requires only $n + m$ amount of time for synthesis and purification. Thus, the approach becomes more advantageous as the number of oligonucleotides involved in the screen increases.

The particular construct of the 10-23 DNA enzyme used in the screen is shown in Fig. 3. The substrate sequence corresponds to the start codon region of HIV-1 *gag-pol* mRNA. The choice of this sequence was arbitrary and was dictated primarily by the large amount of biochemical data that is available for this sequence (Santoro & Joyce, 1998). The length of each arm of the enzyme was varied from eight to ten nucleotides, requiring nine different DNA molecules. Similarly, the length of each arm of the substrate was varied from eight to ten nucleotides, requiring nine different RNA molecules. Each oligonucleotide was designated by two numbers corresponding to the length of its two arms. By convention, each RNA–DNA complex was referred to as x_1-x_2/y_2-y_1, where x_1 and x_2 are the lengths of the 5′ and 3′ arms of the substrate and y_1 and y_2 are the lengths of the 5′ and 3′ arms of the enzyme, respectively. For example, the 8-9/10-8 complex contains eight base pairs in each substrate-recognition domain, a two-nucleotide overhang at the 3′ end of the DNA strand and a one-nucleotide overhang at the 3′ end of the RNA strand. The sequence depicted in Fig. 3 is that of the 10-10/10-10 complex.

Each of the RNA oligonucleotides contained a single deoxyadenylate residue at the cleavage site in order to prevent RNA cleavage during crystallization. All molecules were synthesized chemically and combined in a 1:1 molar ratio, giving 81 unique complexes. The complete set of pairwise combinations included nine complexes with blunt ends (diagonal elements in Fig. 4), 24 complexes with a single one-nucleotide overhang, 16 complexes with two one-nucleotide overhangs, 12 complexes with a single two-nucleotide overhang, four complexes with two two-nucleotide overhangs and 16 complexes with a one-nucleotide overhang at one end of the duplex and a two-nucleotide overhang at the other. A single synthesis of each oligonucleotide on the 1.0 µmol scale provided a sufficient amount of material to carry out the screen.

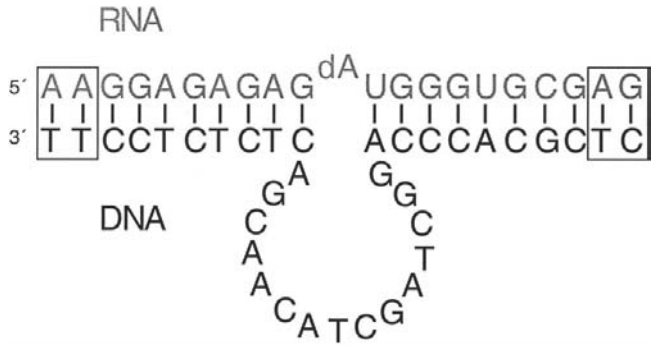

Figure 3
Composition of the enzyme–substrate complex used in the combinatorial screen. Boxed nucleotides represent the region of variable length.

3.2. Crystallization solutions

Each enzyme–substrate complex was evaluated for its ability to crystallize using a common set of 48 conditions which were designed specifically for this system (Table 1). The choice of crystallization conditions was based on prior experience with the DNA enzyme and on other nucleic acid crystallization studies that have been reported in the literature (Baeyens *et al.*, 1994; Berger *et al.*, 1996; Doudna *et al.*, 1993; Scott *et al.*, 1995). The set of solutions included the use of several different buffers in the pH range 5.5–8.0. The most commonly used precipitants were ethanol, 2-propanol, polyethylene glycol (PEG) and various inorganic salts. The approximate concentration of precipitant needed to induce the optimal degree of supersaturation was an important variable in designing the screen. Based on previous experience, it was estimated that either 10%(*v*/*v*) 2-propanol, 15%(*v*/*v*) ethanol or 2.0 *M* inorganic salts was appropriate for solutions which contained 0.4 m*M* enzyme–substrate complex. Because the activity of the 10-23 enzyme is dependent on divalent metal cations, magnesium or calcium salts were

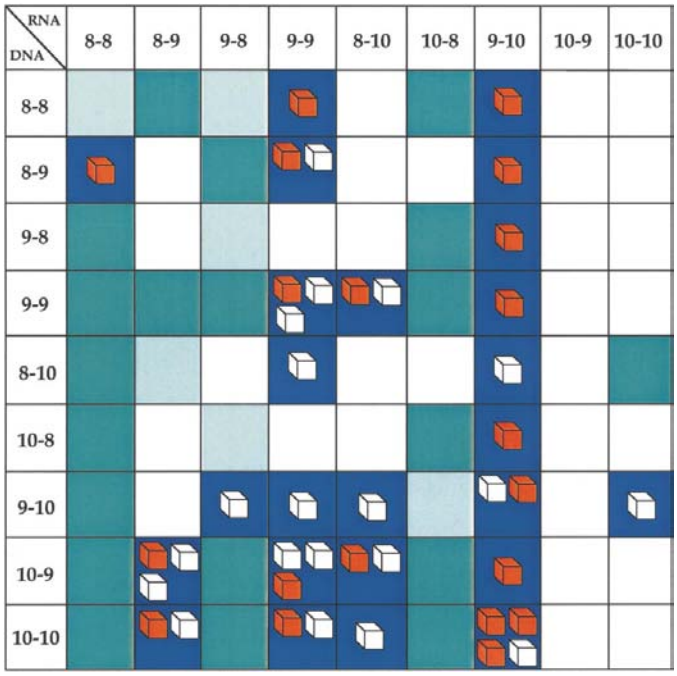

Figure 4
Results of the crystallization screen. Each element in the matrix corresponds to a particular complex of the DNA enzyme (rows) and RNA substrate (columns), summarizing the results for 48 different crystallization conditions. White square, no crystals; light blue square, some crystalline precipitates; medium blue square, twinned crystals; dark blue square, well formed single crystals. White cube, non-birefringent crystal form; red cube, birefringent crystal form.

Table 1
Sparse matrix of 48 crystallization conditions.

Buffers: Cac, cacodylate; Hep, HEPES; Suc, succinate – all were present at 50 m*M* concentration. CoHex, cobalt (III) hexamine; PEG, polyethylene glycol; MPD, 2-methyl-2,4-pentanediol; HD, 1,5-hexanediol; AS, ammonium sulfate; *i*-PrOH, 2-propanol; tart, tartrate; NaCit, sodium citrate.

No.	Buffer, pH	[Mg^{2+}] (m*M*)	[Spermine] (m*M*)	[CoHex] (m*M*)	Additive	Precipitant
1	Hep, 7.5	80	2.5	—	—	—
2	Cac, 6.0	18	2.25	—	1 m*M* $CuSO_4$	9% *i*-PrOH
3	Cac, 6.5	18	0.9	1.8	—	9% *i*-PrOH
4	Cac, 6.5	18	2.25	—	—	9% *i*-PrOH
5	Cac, 7.0	18	2.25	0.9	—	4.5% MPD
6	Cac, 6.5	36	2.25	—	—	5% PEG 400
7	Suc, 5.5	10	—	2.0	—	10% *i*-PrOH
8	Cac, 6.0	20	1.0	—	—	15% EtOH
9	Cac, 7.0	20	1.0	1.0	—	15% EtOH
10	Cac, 7.0	5	1.0†	—	—	10% *t*-BuOH
11	Cac, 7.0	30	2.5	—	—	5% PEG 400
12	Cac, 6.5	100	—	2.0	—	5% *i*-PrOH
13	Tris, 8.0	10	—	1.0	—	20% EtOH
14	Hep, 7.5	20	1.0	—	—	5% PEG 8000
15	Cac, 6.0	20	2.5	—	—	5% PEG 4000
16	Cac, 6.0	10	2.5	—	5 m*M* $CaCl_2$	10% *i*-PrOH
17	Cac, 7.0	9	2.25	1.8	0.9 m*M* spermidine	5% PEG 400
18	Cac, 6.5	10	2.5	—	1 m*M* $CuSO_4$	10% *i*-PrOH
19	Cac, 6.0	20	1.0	—	2 m*M* $CaCl_2$	10% HD
20	Hep, 7.5	15	1.0†	—	—	10% dioxane
21	Cac, 6.0	15	3.0	—	—	10% PEG 400
22	Cac, 6.5	—	2.5	—	18 m*M* $CaCl_2$	9% *i*-PrOH
23	Cac, 6.5	—	2.0	1.0	80 m*M* $CaCl_2$	—
24	Cac, 6.5	5	—	2.5	—	—
25	Cac, 6.5	30	1.0	—	—	1.3 *M* Li_2SO_4
26	Cac, 6.0	—	—	—	200 m*M* $Ca(OAc)_2$	5% *i*-PrOH
27	Cac, 6.5	100	—	1.0	—	10% EtOH
28	Cac, 6.0	10	2.5†	—	—	2.5 *M* NaCl
29	Cac, 6.5	10	—	—	200 m*M* NaCit	5% *i*-PrOH
30	Cac, 6.5	15	10.0	—	—	2.0 *M* Li_2SO_4
31	Cac, 6.5	20	1.0	—	—	2.0 *M* AS
32	Cac, 6.5	10	1.5	—	—	3.0 *M* AS
33	Hep, 7.5	15	1.0	—	—	1.0 *M* AS
34	Cac, 6.0	—	—	—	200 m*M* $Ca(OAc)_2$	2.5 *M* NaCl
35	Cac, 6.0	—	—	1.0	200 m*M* $Ca(OAc)_2$	2.0 *M* LiCl
36	Cac, 6.5	15	5.0†	1.0	—	2.0 *M* NaCl
37	Cac, 6.5	200	—	—	100 m*M* NaCl	20% PEG 1000
38	Tris, 7.5	50	—	—	—	1.0 *M* Na tart
39	Tris, 7.5	200	—	—	—	2.5 *M* NaCl
40	Cac, 6.0	200	—	—	—	2.5 *M* KCl
41	Tris, 8.0	200	—	—	—	15% EtOH
42	Cac, 6.0	15	5.0†	—	—	2.0 *M* Li_2SO_4
43	Cac, 6.0	20‡	0.5	—	100 m*M* NaCl	25% MPD
44	Suc, 5.5	20	0.5	—	—	3.0 *M* AS
45	Cac, 6.5	—	—	5.0	—	2.5 *M* KCl
46	Cac, 6.5	50	—	2.0	—	1.5 *M* Li_2SO_4
47	Cac, 6.5	—	1.0	2.0	30 m*M* $CaCl_2$	2.0 *M* LiCl
48	Cac, 6.5	10	50.0	—	—	—

† Spermidine instead of spermine. ‡ Magnesium acetate instead of magnesium chloride.

present in all solutions. Each RNA–DNA complex was annealed and maintained in a solution of 2 m*M* $MgCl_2$ prior to crystallization. Other cationic additives present in some of the crystallization solutions were spermine, spermidine, cobalt (III) hexamine and various monovalent metal ions. The 48 solutions can be divided into two groups: solutions 1–24 were of relatively low ionic strength, while solutions 25–48 were of higher ionic strength. All crystallization drops were vapor-equilibrated at 297 K against 1 ml of crystallization solution. The drops contained 1 µl of reservoir liquor and 1 µl of a 0.4 m*M* solution of the enzyme–substrate complex.

3.3. Crystallization results

A total of 162 crystallization plates (3888 drops) were set up by two persons over approximately six weeks. Evaluation of the results required an additional two weeks. The crystallization plates were incubated for at least 14 d before assessment of crystal growth. 1–6 d were typically required for the crystals to grow to their final size. Some drops produced amorphous precipitates which slowly turned into crystals over a few weeks. In two unusual instances, involving the 9-9/9-9 and 9-9/8-9 complexes, crystals which were present initially subsequently dissolved and then reappeared with an entirely different morphology.

The results of the screen are summarized graphically in Fig. 4. Based on visual inspection, 30 of the 81 complexes (37%) did not crystallize at all and six (7%) produced only small poorly developed crystals or crystalline precipitates. 20 of the constructs (25%) did not produce well developed crystals but showed good crystallization potential, as indicated by large but twinned crystals which grew under several different conditions. Finally, 25 constructs (31%) produced well formed single crystals that were adequate for further analysis by X-ray diffraction (dark blue squares in Fig. 4). The largest of these crystals grew to 0.3 mm in each dimension, making them suitable for X-ray diffraction experiments without further optimization (Fig. 5). The screen resulted in a total of 40 different crystal forms (Fig. 4). Representative crystal morphologies of six of these forms are shown in Fig. 5.

As anticipated, complexes that had the most favorable stem lengths crystallized under several different crystallization conditions. For example, the 9-10/10-10 complex produced four different crystal forms from 12 different solutions, the 9-9/8-9 complex produced two different forms from 14 different solutions and the 9-10/9-9 complex produced a single form from 16 different solutions (Table 2). The large number of complexes which were screened against the 48 different crystallization solutions allowed a statistical analysis of the efficiency of each solution in promoting crystal growth. The histogram in Fig. 6 shows the number of complexes that crystallized from each solution. Five different crystallization mixtures (solutions 3, 16, 17, 27 and 41 in Table 1) produced crystals of at least 14 different complexes. The most efficacious solution contained 100 m*M* $MgCl_2$, 50 m*M* Tris–HCl (pH 8.0)

and 15% ethanol (solution #41 in Table 1) and yielded crystals of 18 different complexes (Fig. 6). In general, more crystals were obtained under low ionic strength conditions (solutions 1–24) than at higher ionic strength. The volatile precipitants (ethanol and 2-propanol) performed better than other precipitants and the choices of buffer and pH were not crucial. For molecules that crystallized under more than one condition, it was straightforward to find a single preferred crystallization condition by combining ingredients from the various productive mixtures.

There are some interesting trends among the constructs that produced the best crystals. The length of the RNA strand appears to be more important than the length of the DNA strand. The most dramatic example of this is the observation that all the complexes containing the 9-10 RNA substrate yielded crystals (Fig. 4, column 7) regardless of the choice of the DNA strand. These crystals had similar but not identical morphologies and diffracted to varying resolution (Table 2). On the other hand, complexes containing the 10-9 RNA substrate did not produce any crystals (Fig. 4, column 8). Summarizing over all the complexes that were tested, the preferred lengths were nine nucleotides for the 5′ portion of the RNA strand (present in 27 crystal forms), either nine or ten nucleotides for the 3′ portion of the RNA strand (18 and 20 crystal forms, respectively), either nine or ten nucleotides for the 5′ portion of the DNA strand (19 and 17 crystal forms, respectively) and ten nucleotides for the 3′ portion of the DNA strand (19 crystal forms). There was no obvious common feature among the 25 complexes that produced the best crystals. Of these, two had blunt ends, nine had a single one-nucleotide overhang, six had two one-nucleotide overhangs, four had a single two-nucleotide overhang and four had a one-nucleotide overhang at one end of the duplex and a two-nucleotide overhang at the other.

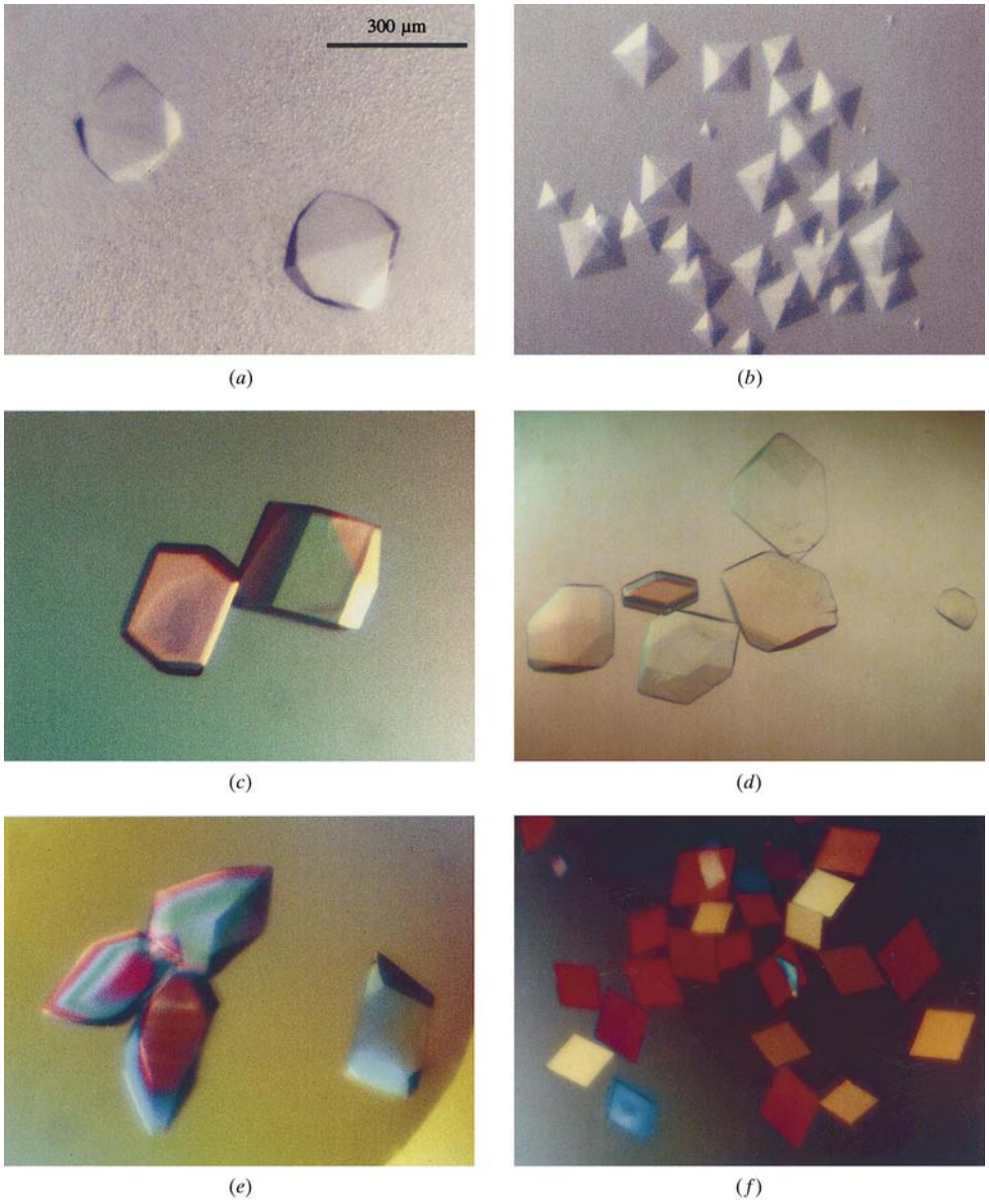

Figure 5
Representative crystal forms obtained from the combinatorial screen. (*a*) 9-9/9-9 complex; (*b*) 8-10/10-9 complex; (*c*) 9-10/10-8 complex; (*d*) 9-10/8-9 complex; (*e*) 9-10/10-10 complex; (*f*) 9-9/8-9 complex. Crystals shown in (*a*) and (*b*) were not birefringent and diffracted to much lower resolution compared with the birefringent forms shown in panels (*c*)–(*f*).

3.4. Characterization of crystals

Most of the crystals that were obtained in the screen were evaluated using a synchrotron X-ray source. The gain in resolution was typically 2–3 Å compared with results obtained using a conventional Cu $K\alpha$ rotating-anode X-ray source. Although the crystals were large and well formed, most diffracted to low resolution. 20 crystal forms diffracted to 15 Å or worse, 17 diffracted to 10–3.5 Å and three diffracted to 3.5 Å or better. Some of the crystals may have suffered internal damage owing to non-optimal freezing conditions and thus might show improved behavior upon optimizing the crystallization and freezing conditions. Table 2 summarizes the diffraction properties of the best crystal forms. Crystals of

Table 2
Summary of diffraction properties of the best crystal forms.

n.d., not determined

Complex†	Nt	Space group	Resolution‡ (Å)	Crystallization conditions
9-9/8-9	51	$I222$ or $I2_12_12_1$	2.8	3, 4, 6, 8, 9, 12, 15, 16, 21, 22, 27, 41, 43
9-10/9-10	54	$I222$ or $I2_12_12_1$	2.8	5, 8, 9, 13, 14, 22
9-10/10-10	55	Hexagonal	3.5	39, 40
9-10/8-9	52	Tetragonal	3.6	3, 5, 9, 17, 21
9-10/10-10	53	n.d.	3.6	4, 6, 9, 11, 13, 15, 16, 22
8-9/10-10	53	$P6_222$ or $P6_422$	3.9	3, 4, 27, 41, 43
9-10/9-9	53	n.d.	~4	4, 5, 6, 8, 9, 11, 16, 17, 20, 22, 27, 30, 37, 39, 41, 42
9-10/10-9	54	n.d.	~4	4, 16, 17, 27, 37, 41, 43
9-10/9-8	52	Tetragonal	~4	3, 9, 17
9-10/10-8	53	n.d.	~5	3, 17
8-9/10-9	52	n.d.	~6	3, 21, 41, 46
9-9/10-10	54	n.d.	~10	22, 24, 31, 42

† Each entry represents a different crystal form. ‡ Resolution limits were estimated by the highest angle observable reflections.

the 9-9/8-9 and 9-10/9-10 complexes were the most highly ordered, with diffraction patterns that extended to 2.8 Å, as judged by the highest angle observed reflections. These crystals belong to a body-centered orthorhombic space group $I222$ or $I2_12_12_1$, which is different from the hexagonal $P6_122$ space group of the 10-23 DNA enzyme molecule, the structure of which had been previously solved. Crystals of these two constructs can be grown under several different crystallization conditions (Table 2).

A strong correlation was observed between crystals that diffracted to high resolution and crystal birefringence. The ten crystal forms that diffracted to the highest resolution (Table 2) were all strongly birefringent (Figs. 5*c*–5*f*). The lack of birefringence might indicate cubic symmetry (or an uniaxial space group when the crystals are viewed along the optical axis), but may also reflect a disordered crystal lattice or an unusually high water content. The lack of any birefringence was a good indicator of poor crystal quality. About half of the crystal forms that were obtained in the screen were not birefringent (Fig. 4) and none of those diffracted to better than 15 Å resolution.

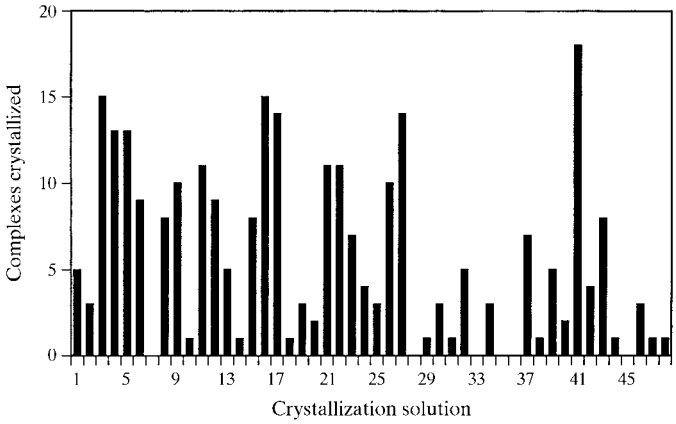

Figure 6
Productivity of the 48 crystallization solutions used in the screen (Table 1). The number of crystal forms obtained from each solution includes both crystals that were small and/or poorly formed and those that were large and well formed.

4. Discussion

The results of the screen demonstrate that the length of the duplex and the presence of overhanging nucleotides have a dramatic effect on crystal formation of the complex of the 10-23 DNA enzyme and its RNA substrate. Among the 81 complexes that were tested in the screen, 25 produced crystals, but only two produced crystals which diffracted to better than 3.0 Å resolution. This low rate of success emphasizes the importance of screening a large number of nucleic acid constructs in order to find the small fraction with desirable crystallization properties. The fraction of complexes that give high-quality crystals is likely to be different for other nucleic acid molecules and in some cases a larger number of combinations will need to be tested.

There are several possible ways to extend the scope of the screening method described in this study. Combinatorial mixing of complementary oligonucleotides in which both the primary sequence and length are varied will generate a much larger ensemble of complexes. Such an approach would require substantially more time to complete the screen. For example, if the identity of the overhanging nucleotides in this screen were allowed to vary, the number of complexes to be tested would increase from 81 to 2601. As an alternative that would be more manageable, the screen could be carried out in two steps. Firstly, the optimal length of the duplex and the preferred location of any nucleotide overhangs could be determined through a combinatorial screen of paired molecules of varying length. Then, if the quality of crystals proved unsatisfactory, the identity of the overhanging nucleotide(s) or the base pairs at the ends of the duplex could be optimized.

If the number of variants to be screened is very large, the time required to set up the crystallization trays can become unreasonable. It typically took 20 min to set up the plate of 24 drops for the screen described in this study. The task of distributing materials to the crystallization wells could be accomplished with a robot (Soriano & Fontecilla-Camps, 1993; Chayen *et al.*, 1994). The use of automated pipeting would significantly shorten the time required to set up the screen and thereby allow more variants to be surveyed.

Nucleic acid molecules that are composed of only a single strand can often be synthesized as two strands held together by strong base-pairing interactions. Nucleic acids that contain more than one double-stranded element and which can be divided into three or more strands provide an opportunity for optimizing multiple duplex elements simultaneously. Another useful feature of the combinatorial screening method is that it generates a large number of crystal forms of structurally and

functionally similar molecules. This might allow the complex to be crystallized in several slightly different conformations owing to differences in the crystallization conditions or the crystal packing interactions. These small structural variations may provide insight into the biochemical properties of the nucleic acid complex.

The authors wish to thank the staff at the Advanced Light Source for their generous assistance. This research was supported by the Skaggs Institute for Chemical Biology at The Scripps Research Institute.

References

Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M. & Harrison, S. C. (1988). *Science*, **242**, 899–907.
Anderson, A. C., Earp, B. E. & Frederick, C. A. (1996). *J. Mol. Biol.* **259**, 696–703.
Baeyens, K. J., Jancarik, J. & Holbrook, S. R. (1994). *Acta Cryst.* D**50**, 764–767.
Berger, I., Kang, C., Sinha, N., Wolters, M. & Rich, A. (1996). *Acta Cryst.* D**52**, 465–468.
Chayen, N. E., Shaw Stewart, P. D. & Baldock, P. (1994). *Acta Cryst.* D**50**, 456–458.
Cruse, W. B. T., Saludjian, P., Biala, E., Strazewski, P., Prange, T. & Kennard, O. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 4160–4164.
Doudna, J. A., Grosshans, C., Gooding, A. & Kundrot, C. E. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 7829–7833.
Jordan, S. R., Whitcombe, T. V., Berg, J. M. & Pabo, C. O. (1985). *Science*, **230**, 1383–1385.
Nowakowski, J., Shim, P. J., Prasad, G. S., Stout, C. D. & Joyce, G. F. (1999). *Nature Struct. Biol.* **6**, 151–156.
Pyle, A. M., Chu, V. T., Jankowski, E. & Boudvillain, M. (1999). In the press.
Santoro, S. W. & Joyce, G. F. (1997). *Proc. Natl Acad. Sci. USA*, **94**(9), 4262–4266.
Santoro, S. W. & Joyce, G. F. (1998). *Biochemistry*, **37**, 13330–13342.
Schultz, S. C., Shields, G. C. & Steitz, T. A. (1990). *J. Mol. Biol.* **213**, 159–166.
Scott, W. G., Finch, J. T., Grenfell, R., Fogg, J., Smith, T., Gait, M. J. & Klug, A. (1995). *J. Mol. Biol.* **250**, 327–332.
Soriano, T. M. B. & Fontecilla-Camps, J. C. (1993). *J. Appl. Cryst.* **26**, 558–562.
Unrau, P. J. & Bartel, D. P. (1998). *Nature (London)*, **395**, 260–263.
Warashina, M., Kuwabara, T., Nakamatsu, Y. & Taira, T. (1999). *Chem. Biol.* **6**, 237–250.
Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzales, C., Scaringe, S. & Usman, N. (1995). *Nucleic Acids Res.* **23**, 2677–2684.