Chapter 8

Generating Crystallographic Models of DNA Dodecamers from Structures of RNase H:DNA Complexes

Martin Egli and Pradeep S. Pallan

Abstract

The DNA dodecamer 5′-d(CGCGAATTCGCG)-3′ is arguably the best studied oligonucleotide and crystal structures of duplexes with this sequence account for a considerable portion of the total number of oligo-2′-deoxynucleotide structures determined over the last 30 years. The dodecamer has commonly served as a template to analyze the effects of sequence on DNA conformation, the conformational properties of chemically modified nucleotides, DNA–ligand interactions as well as water structure and DNA–cation binding. Although molecular replacement is the phasing method of choice given the large number of available models of the dodecamer, this strategy often fails as a result of conformational changes caused by chemical modification, mismatch pairs, or differing packing modes. Here, we describe an alternative approach to determine crystal structures of the dodecamer in cases where molecular replacement does not produce a solution or when crystals of the DNA alone cannot be grown. It is based on the discovery that many dodecamers of the above sequence can be readily co-crystallized with *Bacillus halodurans* RNase H, whereby the enzyme is unable to cleave the DNA. Determination of the structure of the complex using the protein portion as the search model yields a structural model of the DNA. Provided crystals of the DNA alone are also available, the DNA model from the complex then enables phasing their structures by molecular replacement.

Key words DNA, Molecular replacement, Phasing, Protein–DNA interactions, Ribonuclease H, RNase H

1 Introduction

The crystal structure of the DNA oligonucleotide 5′-d(CGCGAAT TCGCG)-3′, the so-called Dickerson-Drew Dodecamer (DDD), provided the first detailed view of a B-form duplex [1]. In the years since then some 160 structures based on the DDD have been deposited in the Nucleic Acid Database (NDB; http://ndbserver.rutgers.edu) [2], amounting to ca. 15 % of the total number of 1,040 structures of DNA duplexes in the NDB. The DDD exhibited unusual features that allowed a glimpse at the sequence-dependence of DNA structure. Among them were the extremely narrow minor groove in the central AATT section, subsequently

Eric Ennifar (ed.), *Nucleic Acid Crystallography: Methods and Protocols*, Methods in Molecular Biology, vol. 1320,
DOI 10.1007/978-1-4939-2763-0_8,

analyzed in depth using variations of the DDD sequence in the context of A-tract geometry and DNA bending [3, 4]. The striking water pattern in the minor groove, termed "spine of hydration" became the subject of numerous studies, and ultimately turned out to be two fused spines dissecting the groove [5]. Improvements in the synthetic preparation of oligonucleotides and purification procedures in combination with intense photon beams at third-generation synchrotrons yielded crystal structures of the DDD at atomic resolution [6]. These afforded intricate details of the water structure in the grooves and around phosphates [7] and the locations of mono- and divalent metal ions [5, 8, 9] as well as insights into the relative importance of crystal packing, base sequence and bound cations on DNA conformation [10]. The DDD also proved to be a fertile testing ground for computational simulations of DNA structure, for example by molecular dynamics simulations [11]. Dozens of structures of the DDD in complex with minor groove binding agents paved the road toward the design of highly specific probes and informed the discovery of drugs [12, 13]. The DDD also constitutes a very useful template sequence for analyzing the conformational properties of chemically modified nucleic acids. It contains all four building blocks, thus allowing incorporation of T, A, C or G analogs, crystals can be grown relatively easily, the duplex can be accommodated in several different space groups, including the most common orthorhombic $P2_12_12_1$ type, and crystals typically diffract to medium or high resolution [14–17]. The fact that the native duplex was studied in considerable detail is a further advantage, as the structures of modified DDDs can be compared to reference structures in order to establish the conformational consequences of a particular chemical modification. The DDD sequence also served as the template for the first crystal structures of all-modified DNA (N3′→P5′ phosphoramidate DNA) [18] and RNA (2′-*O*-(2-methoxyethyl)-RNA) [19].

In our efforts directed at the conformational analysis of chemically modified nucleic acids, we encountered quite a few cases where incorporation of a modified nucleotide precluded crystallization of the DDD alone [20–22]. Alternatively, crystals of the modified fragment could be grown, but the structure resisted phasing by molecular replacement using the canonical DDD duplex as the search model [23]. Both limitations can potentially be overcome by co-crystallizing such modified DDDs with *Bacillus halodurans* RNase H (*Bh*RNase H) [24]. RNase H endonucleolytically cleaves the RNA portion of DNA:RNA hybrids, but the enzyme also binds double-stranded RNA (dsRNA) and DNA (dsDNA; although it is unable to cleave either) and single-stranded nucleic acids, albeit with lower affinity [25]. We pursued the structure determination of non-specific RNase H:dsRNA and RNase H:dsDNA complexes. Although attempts to obtain a high-resolution structure of the former complex have thus far failed [26], we found that *Bh*RNase H can be readily crystallized

with the native [27] and chemically modified DDDs [20–23]. The structures of (MR) these complexes can be determined with molecular replacement using the RNase H portion as the search model. In cases where crystals of the modified DDD alone are at hand, the availability of the refined structure of the DDD from the complex then provides a handle to phase the crystal of the duplex alone. If the duplex cannot be crystallized without RNase H, the crystal structure of the complex furnishes a model of the modified DDD. The present protocol describes the overall approach of using RNase H complexes as a means to determine the crystal structures of DDDs alone that cannot be readily phased, most likely due to conformational changes as a result of the chemical modification that preclude the use of the native DDD as a viable search model, or to gain access to a model of a modified DDD for which crystals in the absence of protein cannot be grown.

2 Materials

2.1 Oligonucleotide Synthesis and Purification

1. Chemically synthesized native DDD 5′-d(CGCGAATTCGCG) DNA, 1 μmol scale.
2. The modified DDD 5′-d(CGCGAATFCGCG) (dF = 2′-deoxyribo-2,4-difluorotoluylnucleotide, a dT isostere [21]) was synthesized on a 1 μmol scale on an ABI 381A DNA synthesizer with the dF phosphoramidite (gift from Glen Research, Sterling, VA), by using a slightly prolonged wait time (10 min) for the phosphoramidite coupling.
3. The modified DDD 5′-d(CGCGAtcATTCGCG) (tcdA = [(5′R,6′R)-2′-deoxy-3′,5′-ethano-5′,6′-methano-β-D-ribofuranosyl]adenine) that contains an adenosine analog with a tricyclic sugar was obtained from Prof. Christian Leumann (University of Berne, Switzerland). For the preparation of tcDNA phosphoramidite building blocks and oligonucleotide synthesis, please *see* ref. 28.
4. The modified DDDs 5′-d(CGCGAAUsTCGCG) and 5′-d(CGCGAAUsUsCGCG) (Us = 2′-*S*Me-U) were synthesized following published procedures ([29] and cited refs.).
5. The modified DDDs containing 5-chloro-U (ClU) in place of T or C, i.e. 5′-d(CGCGAAClUTCGCG), 5′-d(CGCGAATClUCGCG), 5′-d(CGCGAAClUClUCGCG) and 5′-d(CGCGAATTClUGCG), were synthesized as previously reported [30]. For a general description of equipment and reagent setup for oligonucleotide synthesis please *see* ref. 31.
6. Ion exchange or reverse phase column for nucleic acids.
7. High performance liquid chromatography (HPLC) system.
8. 0.45 μM syringe filters.

2.2 Protein Expression and Purification

1. *Bacillus halodurans* (*B. halodurans*) genomic DNA (American Type Culture Collection, ATCC, Manassas, VA, USA).
2. The C-terminal fragment of *B. halodurans* RNase H (*Bh*-RNase H; Met58 to Lys196) with the Asp132 → Asn mutation was cloned into the PET15b vector with an N-terminal His tag and a thrombin cleavage site and expressed in *E. coli* BL21 cells and purified following published procedures [24, 27]. The protein solution was concentrated to ~20 mg/mL.

2.3 Crystallization

1. Nucleic Acid Mini Screen (Hampton Research, Aliso Viejo, CA).
2. Natrix and Crystal Screen I crystallization kits (Hampton Research, Aliso Viejo, CA).
3. Crystallization plates suitable for hanging drops.
4. Siliconized cover-slips.
5. 2-Methyl-2,4-pentanediol (MPD) 35 % (v/v).

2.4 Data Collection and Structure Solution

1. Nylon loop for crystals.
2. Liquid nitrogen and tools for crystal mounting.
3. Data processing program (HKL2000 or similar).
4. CCP4 software suite.
5. PHENIX software suite.
6. The Coot program for model building and visualization.

3 Methods

3.1 DNA Crystallization

1. Purify oligonucleotides either by ion exchange or reverse phase HPLC.
2. After collecting the HPLC fractions, filter solutions through a 0.45 μM syringe filter prior to setting up crystallization droplets.
3. Initial crystallization experiments should be undertaken with the DDD duplex alone. The concentration of the purified oligonucleotide is adjusted to ca. 2–4 mM in RNase/DNase free water. Alternatively, one can constitute the oligonucleotides in 5 mM Tris (pH 7.0), 50 mM NaCl buffer solution. The crystallization trials will typically be done with concentrations of the oligonucleotide that lie between 1 and 2 mM.
4. Strands are then annealed by heating the stock solution to 70–75 °C for 1–2 min., followed by slow cooling to room temperature.
5. Our first choice in terms of initial crystallization conditions to be tested with oligonucleotides is the hanging-drop vapor

diffusion technique in combination with the commercially available Nucleic Acid Mini Screen that consists of 24 different solutions [32]. Equal volumes of 1 or 2 μL of the oligonucleotide and crystallization solutions are mixed on siliconized cover-slips and the droplets equilibrated against a reservoir of 500 μL of 35 % MPD (v/v) in a 24-well plate.

6. The plates are then incubated at 18 °C and checked at regular intervals over the course of several weeks. In many cases, crystals obtained from this screen are suitable for diffraction experiments and data collection. However, it may be necessary to refine the conditions that resulted in initial growth of crystals in order to optimize their size and/or resolution limit. This is best achieved by varying one parameter at A time, e.g. pH, metal ion concentration, volume and concentration of the MPD reservoir solution as well as temperature. If the above sparse matrix screen does not yield diffraction-quality crystals, other conditions should be screened. In our own laboratory we often rely on the Natrix [33] and Crystal Screen I kits [34] from Hampton Research that are comprised of 48 and 50 individual conditions, respectively.

3.2 Phasing DDD Crystal Structures by Molecular Replacement

1. Once crystals are obtained, they are picked up from solution with a nylon loop, frozen in liquid nitrogen and checked for X-ray diffraction, either on in-house instrumentation or at a synchrotron source.
2. Data are processed and scaled with HKL2000 [35]. If the crystals are isomorphous with native DDD crystals, i.e. space group orthorhombic $P2_12_12_1$ with unit cell constants $a \approx 25$ Å, $b \approx 40$ Å and $c \approx 65$ Å, the structure can in all likelihood be determined by the Molecular Replacement (MR) technique, using a native DDD duplex as the search model.
3. MR is performed with commonly used software such as Molrep [36] in the CCP4 suite [37] (or any MR program of choice) in combination with a suitable model.
4. The obtained MR solution is checked for unfavorable packing contacts, and provided packing and the initial values for R-factor and R-free (i.e. values in the mid or low 30s) are indicative of a correct solution, the next steps consist of restrained refinement of atomic positions and temperature factors with, for example, the program PHENIX [38].
5. Visualization of the duplex and electron density as well as manual model building and water placement is performed, for example, using the program Coot [39].

In our investigations directed at the conformational consequences for DNA as a result of chemical modification, we frequently came across modified DDDs that resisted crystallization

despite an extensive number of trials. In other cases crystals could be grown but their structures subsequently not determined by MR. Alternative approaches such as multiple isomorphous replacement (MIR), single- or multi-wavelength anomalous dispersion (SAD or MAD, respectively), or combinations thereof (SIRAS or MIRAS, respectively) with suitable derivatives may then be pursued to phase the structure. However, this can result in a potentially protracted search for heavy atom derivatives or require the synthesis of chemically modified strands for SAD/MAD. By comparison, our approach described below involving co-crystallization of such DDDs with *Bh*RNase H and then solving the structure with MR using the protein as the search model offers a quicker solution to overcome the phasing problem.

3.3 DDDs That Resist Crystallization or DDD Crystal Structures That Resist Phasing by MR

1. Co-crystallization with *Bh*RNase H of chemically modified DDDs offers a rapid route to 3D-structural models of the duplexes in cases of DDDs for which crystals cannot be grown or crystals of DDDs that are non-isomorphous with the native DDD crystal form and resist phasing by MR. Crystals of the native DDD in complex with *Bh*RNase H diffract to high resolution (*see* Table 1, Fig. 1) and the structure can be readily determined by MR using the protein as the search model [27].
2. Unlike in the crystal of the native DDD alone, individual strands of duplexes in the complex crystal are related by two-fold rotational symmetry (*see* **Note 1**). Thus, DDD duplexes bound to RNase H display kinks into the major groove at both ends, whereas native duplexes in DDD crystals without protein are located in a general position and feature an asymmetric kink (Fig. 2).

3.3.1 DDDs That Resist Crystallization

1. If crystallization experiments with DDDs that contain modified nucleotides do not yield crystals or in cases where DDD crystals diffract only poorly, co-crystallization with *Bh*RNase H should be attempted (*see* **Note 2**). For crystallization experiments, the annealed DDD is mixed with the protein in a 2:1 M ratio in the presence of 5 mM $MgCl_2$ such that the final protein concentration is ~10 mg/mL.
2. Screening of crystallization conditions is best done with commercially available kits (*see* Subheading 3.1) and once crystals are obtained phasing with MR using RNase H as the search model will readily deliver a structural model of the modified DDD. One should be aware that not all crystals are of a complex, but that in quite a few cases crystals will contain only *Bh*RNase H (*see* Subheading 3.4). Examples of modified DDDs that were crystallized in complex with *Bh*RNase H but for which crystals of the duplex alone could not be grown are listed in Table 1. The complex between *Bh*RNase H and a DDD duplex containing two difluorotoluene residues in place of T is shown in Fig. 3.

Table 1
Selected crystal data and structure refinement statistics for complexes between *Bh*RNase H and native DDD, F-DDD, ClU-DDD, and tcdA-DDD

Complex	*Bh*RNase H/native d(CGCGAATTCGCG)	*Bh*RNase H/F = DFT d(CGCGAATFCGCG)	*Bh*RNase H/U* = 5-ClU d(CGCGAATU*CGCG)[a]	*Bh*RNase H/A* = tcdA d(CGCGAA*TTCGCG)
Crystal data				
Space group	Monoclinic, *C*2	Monoclinic, *C*2	Orthorhombic, $P2_12_12_1$	Tricilinic, *P*1
Unit cell constants *a*, *b*, *c* [Å] α, β, γ [°]	98.40, 66.66, 76.93 90, 122.3, 90	96.18, 66.71, 77.57 90, 121.0, 90	64.08, 64.76, 116.47 90, 90, 90	42.39, 47.50, 55.22 100.9, 101.8, 89.8
Wavelength [Å]	0.9785	1.0000	0.9787	1.0000
Resolution (outer shell) [Å]	1.80 (1.86–1.80)	1.61 (1.64–1.61)	1.49 (1.52–1.49)	1.54 (1.65–1.54)
No. of unique refls. (outer shell)	37,768 (3,643)	52,821 (2,202)	77,514 (3,752)	56,542 (9,820)
Completeness (outer shell) [%]	97.3 (94.7)	98.1 (82.0)	96.9 (95.1)	92.5 (86.6)
R-merge (outer shell)	0.061 (0.248)	0.053 (0.560)	0.071 (0.824)	0.047 (0.480)
I/σ (I) (outer shell)	42.7 (4.6)	39.4 (1.7)	27.3 (2.5)	19.5 (4.4)
Refinement parameters				
No. of protein molecules/ DNA strands per asymmetric unit	2 RNase H/2 single strands	2 RNase H/2 single strands	2 RNase H/2 duplexes	2 RNase H/2 single strands
R-work/R-free [%]	0.215/0.241	0.197/0.238	0.189/0.218	0.163/0.217
No. of protein/ DNA atoms	2,169/486	2,154/488	2,244/900	2,154/554
No. of waters/ ions/ligands	167/1 Na^+/–	232/–/4 glycerol	476/2 Mg^{2+}/4 glycer., 1 EGOH	200/–/2
R.m.s.d. bond lengths [Å]/ ang. [°]	0.016/1.7	0.030/2.5	0.009/1.4	0.021/2.0
Avg. B-factor, protein/DNA [Å^2]	30.9/29.4	16.9/13.0	22.1/48.5	32.3/34.7

(continued)

Table 1
(continued)

Complex	*Bh*RNase H/native d(CGCGAATTCGCG)	*Bh*RNase H/F = DFT d(CGCGAATFCGCG)	*Bh*RNase H/U* = 5-ClU d(CGCGAATU*CGCG)[a]	*Bh*RNase H/A* = tcdA d(CGCGAA*TTCGCG)
Avg. B-factor, H_2O/ions/ small molecules [Å^2]	37.3/53.2/–	24.4/–/42.5	36.5/13.7/29.0	40.7/–/47.2
PDB entry code	3D0P	3I8D	4HTU	4OPJ

[a]We determined crystal structures of *Bh*RNase H in complex with four different DDDs with incorporated ClU residues. Only one of them is listed here

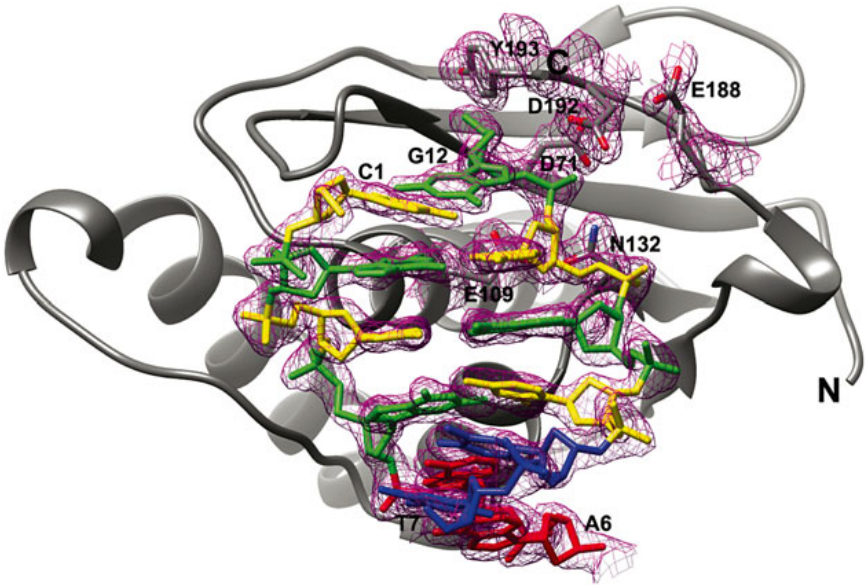

Fig. 1 Overall structure of the native DDD in complex with *Bh*RNase H and quality of the electron density. The DNA duplex is located on crystallographic dyad and only 6 bp are shown. The protein is depicted in a ribbon cartoon colored in gray and the DDD is viewed into the major groove with G, C, A and T nucleotides colored in *green*, *yellow*, *red*, and *blue*, respectively. Protein chain termini and terminal nucleotides are labeled. The Fourier 2Fo-Fc sum electron density around the 6 bp and amino acids at the active site (E109 and N132) or interacting with the terminal base pair is contoured at the 1σ level. Figures were generated with UCSF Chimera [43]

3.3.2 DDD Crystal Structures That Resist Phasing by MR

1. Many DDDs with incorporated chemically modified residues or bound to small molecules crystallize in the same orthorhombic crystal system as the native DDD and with cell constants that deviate only minimally from those of native crystals. However, there are numerous cases where chemical modification causes conformational changes that affect crystal packing, thus resulting in a variety of new crystals forms. Often MR with the native DDD as the search model will then fail

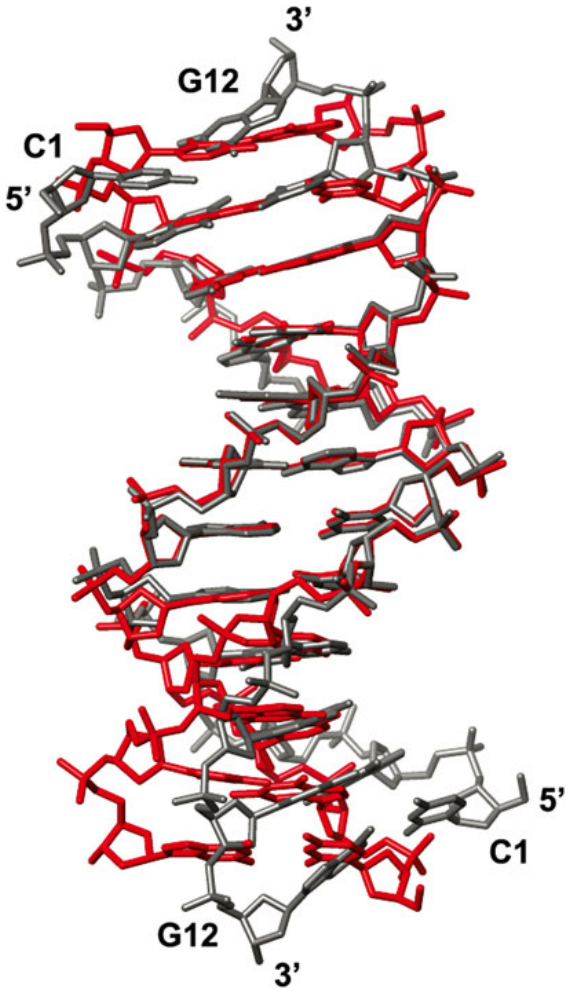

Fig. 2 Superimposition of the DDDs in the crystal structure of the complex with *Bh*RNase H (*gray*; Table 1) and crystallized alone (*red*; PDB ID code 436D [6]). Only phosphorus atoms in the central A-tract region were used for generating the overlay

because altered lattice interactions and/or chemical modification cause the DDD to adopt a conformation that is significantly different from that of the native duplex.

2. Instead of embarking on a potentially time-consuming search for alternative phasing approaches involving heavy atom soaks (the options with DNA are rather limited compared to protein crystals in this respect) or de novo synthesis of oligonucleotides with brominated pyrimidines (*see* **Note 2**), one should first attempt to co-crystallize such DDDs with *Bh*RNase H. MR with RNase H as the search model will then produce a model of the DDD duplex. For example, the 2′-*S*Me-U modification (Us) inserted into the DDD in place of T causes the narrow minor groove in the central section to open up significantly.

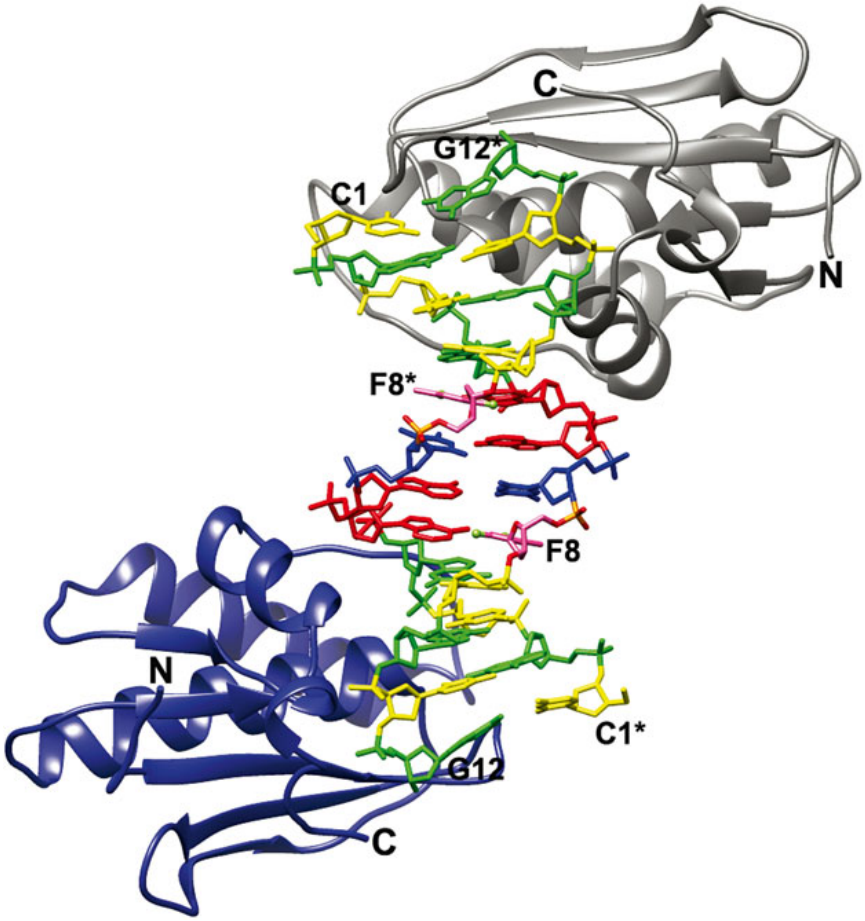

Fig. 3 Structure of a DDD containing two difluorotoluene residues (dF) in complex with *Bh*RNase H. The color code for G, C, A, and T nucleotides is the same as in Fig. 1, dF residues are *pink* with fluorine atoms highlighted as *green balls*, and *asterisks* indicate that the two strands of the DDD are symmetry-related in the crystal structure of the complex

Not surprisingly, the structures of DDDs with either one T or both Ts per strand replaced by Us could not be determined by MR [23]. By comparison, the crystal structures of their complexes with *Bh*RNase H yielded readily to MR with the RNase H search model (*see* Table 2). The modified DDD duplexes from the refined structures of the complexes were subsequently used for phasing by MR of the crystals of the duplexes alone.

3. The success of this approach is of course dependent on rather similar conformations of the duplexes in complex with *Bh*RNase H and in the free form (Fig. 4). Although the number of examples of modifications tested and complexes of DDDs with *Bh*RNase H determined thus far is still somewhat limited, the crystal data summarized in Tables 1 and 2 indicate a considerable variety of packing arrangements between complexes. We take this as evidence in support of the versatility of the co-crystallization approach for phasing of DDD structures (*see* **Notes 3–5**).

Table 2
Selected crystal data and structure refinement statistics for 2′-*S*Me-DDD:*Bh*RNase H complexes

Complex	*Bh*RNase H/U* = 2′-SMe-U d(CGCGAAU*U*CGCG)	*Bh*RNase H/U* = 2′-SMe-U d(CGCGAAU*TCGCG)
Space group	Monoclinic, *C*2	Monoclinic, *C*2
Unit cell constants *a*, *b*, *c* [Å] *β* [°]	81.93, 66.62, 38.77 103.3	98.14, 66.61, 77.90 122.1
Wavelength [Å]	0.9785	1.0000
Resolution (outer shell) [Å]	1.60 (1.66–1.60)	1.54 (1.57–1.54)
No. of unique reflections	26,499 (–)	59,362 (2,885)
Completeness (outer shell) [%]	99.4 (96.9)	94.2 (91.8)
R-merge (outer shell)	0.077 (–)	0.047 (0.675)
I/σ (I) (outer shell)	39.4 (1.7)	28.3 (1.3)
Refinement parameters		
No. of protein molecules/DNA strands per asymmetric unit	1 RNase H/1 single strand	2 RNase H/2 single strands
R-work/R-free [%]	0.193/0.261	0.192/0.259
No. of protein/DNA atoms	1,114/245	2,178/498
No. of waters/small molecules	102/3 glycerol	167/3 glycerol
R.m.s.d. bond lengths [Å]/angles [°]	0.030/2.5	0.023/2.3
Avg. B-factor, protein/DNA atoms [Å^2]	59.2/13.0	42.1/32.9
Avg. B-factor, H_2O/small mols [Å^2]	45.2/59.7	41.0/41.2
PDB entry code	3EY1	4OPK

3.4 Crystals of apo BhRNase H

1. Crystals obtained from mixtures of *Bh*RNase H and DDDs in some cases only contain the protein. In addition to the reported crystal form of *Bh*RNase H [24], we have so far identified two additional crystal forms (*see* Table 3).
2. If the unit cell dimensions and space groups of a "complex" crystal match one of these, it is more likely that the crystal obtained contains only RNase H (*see* **Note 6**). However, in order to make sure that the crystal is definitely not of a complex, it is advisable to collect a full data set, conduct a MR search, and inspect the generated Fourier 2Fo-Fc sum and Fo-Fc difference density maps.

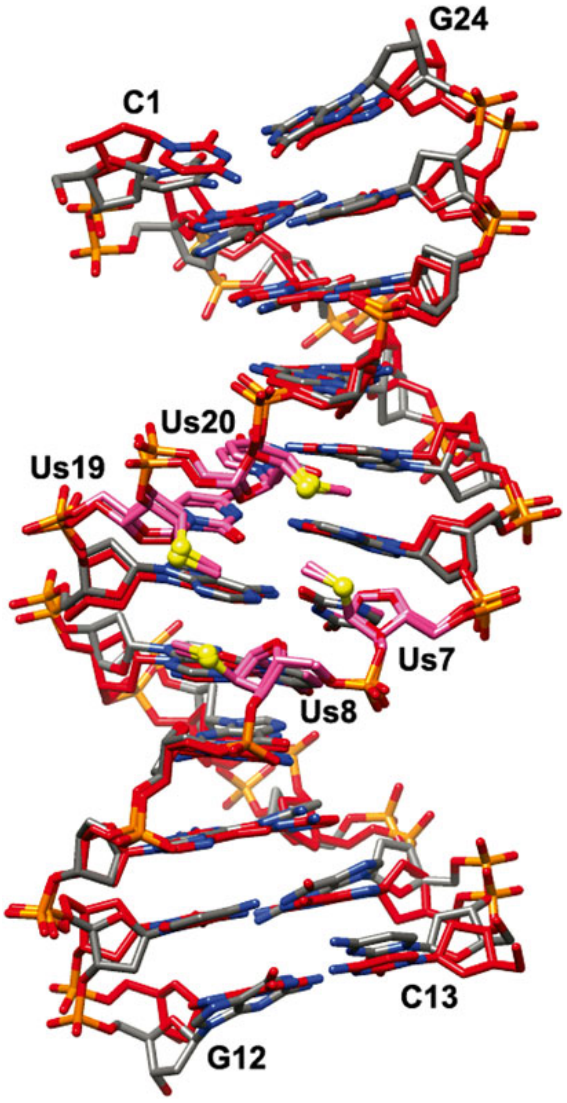

Fig. 4 Superimposition of the structures of the 2′-*S*MeU (Us)-modified DDD bound to *Bh*RNase H (*gray* carbon atoms; Table 2) and the Us-modified DDD crystallized alone (*red* carbon atoms). The central, modified section of the duplex is viewed into the minor groove and sulfur atoms in Us residues (*pink* carbon atoms) are highlighted as *yellow balls*

Table 3
Crystal forms of apo *Bh*RNase H

Space group	Trigonal, $P3_121$ [24]	Monoclinic, $P2_1$	Orthorhombic, $P2_12_12_1$
Unit cell constants a, b, c [Å] α, β, γ [°]	66.8, 66.8, 58.7 90, 90, 120	34.8, 63.3, 49.7 90, 94.0, 90	35.2, 63.8, 80.6 90, 90, 90

4 Notes

1. With most protein–DNA complexes, successful growth of diffraction-quality crystals is critically dependent on both oligonucleotide length and sequence. In many cases, the duplex mediates packing contacts, for example by end-to-end stacking between base pairs from adjacent duplexes. Crystals of *Bh*RNase H:DDD complexes exhibit various symmetries and associated lattice interactions that are dominated by protein–protein contacts (*see*, for example, refs. 22, 27). The central A-tract section of the minor groove in the DDD is very narrow and RNase H therefore contacts the minor groove in the outer G:C sections, where it is expanded and more similar in width to that in RNA:DNA hybrids. The AATT tetramer therefore serves as a spacer and one can imagine that its length could be varied somewhat without fundamentally altering the interactions between RNase H and the outer G:C sections. Because the proteins would remain bound to the terminal regions of duplexes, the former may mediate packing interactions.
2. Co-crystallizing a DDD with RNase H is a convenient approach to determine the structure of the nucleic acid duplex in cases where it cannot be crystallized alone or when the structure of the duplex alone cannot be determined by molecular replacement. A good case in point is the DDD CGCGAAUsUsCGCG with two 2′-*S*Me-U residues (Us) in place of T. Crystals of this modified DDD can be grown readily from multiple conditions. However, molecular replacement searches with the native DDD as the search model consistently failed. Halogenated pyrimidines such as Br^5U or Br^5C in combination with single wavelength anomalous dispersion (SAD) commonly offer a tested approach to phase a crystal structure that resists molecular replacement (although the use of Br^5U here would have required the synthesis of a special building block because of the 2′-*S*Me-U modification in the above DDD). However, this DDD also appeared to be an almost ideal example for replacement of either one or the other or both of the Us residues by 2′-*Se*Me-U for "SeMe" SAD phasing [40, 41]. We produced all three selenoated DDDs but despite exhaustive attempts none of them gave diffraction-quality crystals. The structure of the complex between CGCGAAUsUsCGCG and *Bh*RNase H then yielded a model of the duplex that was used to determine the structure of the Us-modified DDD alone [23] (*see* Table 2, Fig. 4).
3. Although the co-crystallization approach with *Bh*RNase H is especially useful for structure determination of chemically modified DDDs that resist crystallization alone, there are limits as to the extent (number of modified nucleotides) and

nature of the modification. For example, we were unsuccessful in our attempts to crystallize *Bh*RNase H with the all-peptide nucleic acid (PNA) DDD CGCGAATTCGCG-K that carries a C-terminal lysine.

4. We determined a crystal structure of *Bh*RNase H in complex with a DDD featuring 5-chlorouridine:G (ClU:G) mismatch pairs (*see* Table 1) [22]. This indicates that crystallization of complexes with DDDs exhibiting altered geometry in the outer regions is possible. In the case of the ClU nucleobase analog, the chlorine substituent is directed into the major groove and does not interfere with the protein binding from the minor groove side. However, as with the T:G mismatch, guanine paired to ClU is shifted into the minor groove, but the more exposed 2-amino moiety appears not to interfere with RNase H binding.
5. RNase H binds double-stranded RNAs (dsRNAs) [25] but is unable to cleave RNA opposite RNA (unlike RNA paired opposite DNA). We initially conducted co-crystallization experiments with RNAs in complex with the enzyme from *E. coli*, but were not successful in identifying a crystal with an ordered RNA duplex bound to RNase H [26]. Subsequently, we screened RNA duplexes with lengths of between 8 and 16 nucleotides in co-crystallization experiments with *Bh*RNase H. Although crystals were obtained in many cases, they all contained the enzyme alone and in only one case, the RNA duplex alone. Thus, a crystal structure of an RNase H:dsRNA complex has not been determined to date.
6. Hits obtained from crystallization experiments with complexes between *Bh*RNase H and DDDs may be the desired complex, or alternatively, may contain only the enzyme (*see* Table 3; we have not observed a case where a DDD alone crystallized from the complex mixture). To distinguish between crystals of complexes on the one hand and those of the apo form of RNase H on the other, it may be helpful to use color-labeled oligonucleotides [42]. Many of these are commercially available, e.g. the 5′-Cy5- or 5′-Cy3-labeled oligo-2′-deoxynucleotides from Integrated DNA Technologies, Coralville, Iowa, USA (http://www.idtdna.com).

Acknowledgments

This work was supported by the US National Institutes of Health grant R01 GM055237.

References

1. Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE (1980) Crystal structure analysis of a complete turn of B-DNA. Nature 287:755–758
2. Berman HM, Olson WK, Beveridge DL, Westbrook J, Gelbin A, Demeny T, Hsieh SH, Srinivasan AR, Schneider B (1992) The nucleic acid database: A comprehensive relational database of three-dimensional structures of nucleic acids. Biophys J 63:751–759
3. DiGabriele AD, Steitz TA (1993) A DNA dodecamer containing an adenine tract crystallizes in a unique lattice and exhibits a new bend. J Mol Biol 231:1024–1039
4. Allemann RK, Egli M (1997) DNA bending and recognition. Chem Biol 4:643–650
5. Minasov G, Tereshko V, Egli M (1999) Atomic-resolution crystal structures of B-DNA reveal specific influences of divalent metal ions on conformation and packing. J Mol Biol 291:83–99
6. Tereshko V, Minasov G, Egli M (1999) The Dickerson-Drew B-DNA dodecamer revisited – at atomic resolution. J Am Chem Soc 121:470–471
7. Egli M, Tereshko V, Teplova M, Minasov G, Joachimiak A, Sanishvili R, Weeks CM, Miller R, Maier MA, An H, Cook PD, Manoharan M (2000) X-ray crystallographic analysis of the hydration of A- and B-form DNA at atomic resolution. Biopolymers 48:234–252
8. Tereshko V, Minasov G, Egli M (1999) A "hydrat-ion" spine in a B-DNA minor groove. J Am Chem Soc 121:3590–3595
9. Egli M (2002) DNA-cation interactions: quo vadis? Chem Biol 9:277–286
10. Egli M, Tereshko V (2004) Lattice- and sequence-dependent binding of Mg^{2+} in the crystal structure of a B-DNA dodecamer. In: Stellwagen N, Mohanty U (eds) Curvature and deformation of nucleic acids: recent advances, new paradigms, ACS symposium, vol 884. Oxford University Press, New York, pp 87–109
11. Young MA, Jayaram B, Beveridge DL (1997) Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: fractional occupancy of electronegative pockets. J Am Chem Soc 119:59–69
12. Neidle S (2001) DNA minor-groove recognition by small molecules. Nat Prod Rep 18:291–309
13. Nanjunda R, Wilson WD (2012) Binding to the DNA minor groove by heterocyclic dications: from AT-specific monomers to GC recognition with dimers. Curr Protoc Nucleic Acid Chem 51:8.8.1–8.8.20
14. Egli M (1996) Structural aspects of nucleic acid analogs and antisense oligonucleotides. Angew Chem Int Ed Engl 35:1894–1909
15. Egli M (1998) Towards the structure-based design of nucleic acid therapeutics. In: Weber G (ed) Advances in enzyme regulation, vol 38. Elsevier, Oxford, pp 181–203
16. Egli M, Pallan PS (2007) Insights from crystallographic studies into the structural and pairing properties of nucleic acid analogs and chemically modified DNA and RNA oligonucleotides. Annu Rev Biophys Biomol Struct 36:281–305
17. Egli M, Pallan PS (2010) Crystallographic studies of chemically modified nucleic acids: a backward glance. Chem Biodivers 7:60–89
18. Tereshko V, Gryaznov S, Egli M (1998) Consequences of replacing the DNA 3′-oxygen by an amino group: high-resolution crystal structure of a fully modified N3′ → P5′ phosphoramidate DNA dodecamer duplex. J Am Chem Soc 120:269–283
19. Teplova M, Minasov G, Tereshko V, Inamati G, Cook PD, Manoharan M, Egli M (1999) Crystal structure and improved antisense properties of 2′-*O*-(2-methoxyethyl)-RNA. Nat Struct Biol 6:535–539
20. Pallan PS, Ittig D, Heroux A, Wawrzak Z, Leumann CJ, Egli M (2008) Crystal structure of tricyclo-DNA: an unusual compensatory change of two adjacent backbone torsion angles. Chem Commun 21:883–885
21. Pallan PS, Egli M (2009) The pairing geometry of the hydrophobic thymine analog 2,4-difluorotoluene in duplex DNA as analyzed by X-ray crystallography. J Am Chem Soc 131:12548–12549
22. Patra A, Harp J, Pallan PS, Zhao L, Abramov M, Herdewijn P, Egli M (2013) Structure, stability and function of 5-chlorouracil modified A:U and G:U base pairs. Nucleic Acids Res 41:2689–2697
23. Pallan PS, Prakash TP, Li F, Eoff RL, Manoharan M, Egli M (2009) A conformational transition in the structure of a 2′-thiomethyl-modified DNA visualized at high resolution. Chem Commun 21:2017–2019
24. Nowotny M, Gaidamakov SA, Crouch RJ, Yang W (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. Cell 121:1005–1016
25. Lima WF, Crooke ST (1997) Binding affinity and specificity of Escherichia coli RNaseH1:

impact on the kinetics of catalysis of antisense oligonucleotide-RNA hybrids. Biochemistry 36:390–398

26. Loukachevitch LV, Egli M (2007) Crystallization and preliminary X-ray analysis of Escherichia coli RNase HI-dsRNA complexes. Acta Crystallogr F 63:84–88
27. Pallan PS, Egli M (2008) Insights into RNA/DNA hybrid recognition and processing by RNase H from the crystal structure of a non-specific enzyme-dsDNA complex. Cell Cycle 7:2562–2569
28. Steffens R, Leumann C (1997) Preparation of [(5′R,6′R)-2′-Deoxy-3′,6′-ethano-5′,6′-methano-β-D-ribofuranosyl]thymine and -adenine, and the corresponding phosphoramidites for oligonucleotide synthesis. Helv Chim Acta 80:2426–2439
29. Lima WF, Nichols JG, Wu H, Prakash TP, Migawa MT, Wyrzykiewicz TK, Bhat B, Crooke S (2004) Structural requirements at the active site of the heteroduplex substrate for human RNase H1 catalysis. J Biol Chem 279: 36317–36326
30. Theruvathu JA, Kim CH, Rogstad DK, Neidigh JW, Sowers LC (2009) Base-pairing configuration and stability of an oligonucleotide duplex containing a 5-chlorouracil-adenine base pair. Biochemistry 48:7539–7546
31. Pallan PS, Egli M (2007) Selenium modification of nucleic acids. Preparation of phosphoroselenoate derivatives for crystallographic phasing of nucleic acid structures. Nat Protoc 2:640–646
32. Berger I, Kang C-H, Sinha N, Wolters M, Rich A (1996) A highly effective 24 condition matrix for the crystallization of nucleic acid fragments. Acta Crystallogr D 52:465–468
33. Scott WG, Finch JT, Grenfell R, Fogg J, Smith T, Gait MJ, Klug A (1995) Rapid crystallization of chemically synthesized hammerhead RNAs using a double screening procedure. J Mol Biol 250:327–332
34. Jancarik J, Kim SH (1991) Sparse matrix sampling: a screening method for crystallization of proteins. J Appl Crystallogr 24:409–411
35. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276:307–326
36. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. Acta Crystallogr D 66:22–25
37. CCP4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D 50: 760–763
38. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive python-based system for macromolecular structure solution. Acta Crystallogr D 66:213–221
39. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D 60:2126–2132
40. Teplova M, Wilds CJ, Wawrzak Z, Tereshko V, Du Q, Carrasco N, Huang Z, Egli M (2002) Covalent incorporation of selenium into oligonucleotides for X-ray crystal structure determination via MAD: proof of principle. Biochimie 84:849–858
41. Pallan PS, Egli M (2007) Selenium modification of nucleic acids. Preparation of oligonucleotides with incorporated 2′-SeMe-uridine for crystallographic phasing of nucleic acid structures. Nat Protoc 2:647–651
42. Jiang X, Egli M (2011) Use of chromophoric ligands to visually screen co-crystals of putative protein-nucleic acid complexes. Curr Protoc Nucleic Acid Chem 46:7.15.1–7.15.8
43. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera – a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612