In Vitro Selection of Hammerhead Ribozyme Sequence Variants

Fritz Eckstein,* Anilkumar R. Kore, and Kay L. Nakamaye^[a]

KEYWORDS:

gene technology \cdot molecular evolution \cdot nucleic acids \cdot ribozymes \cdot structure – activity relationships

1. Introduction

The hammerhead ribozyme, together with the hairpin and hepatitis delta virus ribozymes, belongs to the class of small ribozymes.^[1-3] It has been the subject of considerable efforts to establish its structure and its mechanism of action.^[4, 5] The structural motif of the hammerhead ribozyme occurs in nature as part of some circular plant pathogenic RNAs (viroids) (Figure 1). It is responsible for the intramolecular cleavage of the

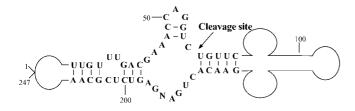


Figure 1. Section of the Avocado Sun Blotch viroid containing the hammerhead motif

concatameric transcripts, resulting from a rolling-circle replication mechanism, leading to the monomeric units.[6] It was established that the core region with the invariant nucleotides was responsible for cleavage and that the only requirement for nucleotides in the arms was their complementarity to form Watson-Crick base pairs. It is by no means clear that the sequence of the core was optimised in nature for most efficient cleavage as it might represent a compromise with other factors such as infectivity or regulation of transcription, which could be important in the life cycle of the viroid. An important step in obtaining more insight into the structure - function relationships of the ribozyme was its redesign for intermolecular cleavage, where substrate and ribozyme are on different RNA strands (Figure 2).[7, 8] This facilitated the determination of kinetic constants and thus a quantitative comparison of cleavage efficiencies of different sequences. X-ray structural analysis has established its three-dimensional (3D) structure at atomic resolution in the ground state as well as in a conformation approaching the transition state. [9-11] Fluorescence spectroscopic studies in particular have identified the conformational changes that the ribozyme undergoes as a function of Mg²⁺ concentration.^[12, 13] Despite these efforts, many questions as to the structure function relationships of this ribozyme have remained unresolved. Thus, for example, the basis for the substrate specificity

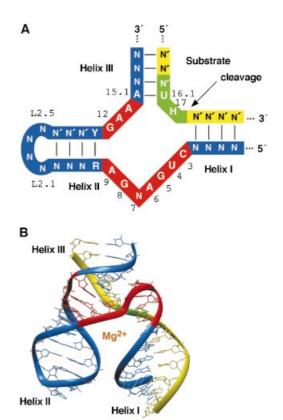


Figure 2. Generalised structure of the in-trans-cleaving hammerhead ribozyme. A: Two-dimensional structure representation. B: Three-dimensional structure representation. $^{[9]}$ Red = core region, green = NUH triplet, yellow = substrate beyond triplet, blue = helix-forming regions and loop II, N = any nucleotide, N' = nucleotide complementary to N, H = any nucleotide except G, Y = pyrimidine nucleotide, R = purine nucleotide.

for cleavage of triplets of the general formula NUH, where N is any nucleotide, U is uridine and H is any nucleotide except quanosine, is unclear.

- [a] Prof. Dr. F. Eckstein, Dr. A. R. Kore, Prof. Dr. K. L. Nakamaye⁽⁺⁾
 Max-Planck-Institut für experimentelle Medizin
 Hermann-Rein-Strasse 3, 37075 Göttingen (Germany)
 Fax: (+49)551-3899-388
 E-mail: eckstein@em.mpa.de
- [†] Further address: Gonzaga University Department of Chemistry Spokane, WA 99258 (USA)

CHEMBIOCHEM

It is generally accepted that the nucleotides in the core mainly determine cleavage rates and the specificity for cleavage at positions on the 3' side of NUH triplets. However, the structural basis for these functional relationships is not understood yet. We reasoned that a comparison of sequences with improved cleavage activities or altered specificities might identify nucleotides that are involved in determining these characteristics. In addition, new cleavage specificities will extend the applications of the hammerhead ribozyme for the inhibition of gene expression by the sequence-specific cleavage of mRNAs.

As single-nucleotide exchanges in the core lead to abolition of activity, a stepwise search for alternative sequences is impractical and in vitro selection is an attractive alternative. We have embarked on such studies to improve our understanding of the structure – function relationships of the hammerhead ribozyme.

2. General strategy

There are basically two possible approaches for in vitro selection of ribozymes with altered activity characteristics. One can create either a cis- or an in-trans-cleaving construct. We first chose the former for our studies. The latter was adopted from the method of Ishizaka et al.^[14] for the selection of more active hammerhead ribozymes and also used by Conaty et al.^[15] for the selection of ribozymes with lower Mg²⁺ requirements, as will be discussed later (see Section 8).

The in-cis strategy is based on the construction of doublestranded DNA containing restriction sites for cloning and sequencing, the T7 promoter for transcription and the sequence for the ribozyme with the nucleotides in the core region randomised. Typically, 10 or 22 nucleotides were randomised, resulting in a population of 10⁶ or 10¹³ sequences, respectively (Figure 3). Transcription with [α - 32 P]UTP in the presence of Mg $^{2+}$ gives full-length transcripts of which those containing active core sequences self-cleave. Products are separated by polyacrylamide gel electrophoresis (PAGE) and the longer of the cleavage products containing the active core ribozyme is isolated, reversetranscribed, a primer added to restore the cleaved segment of the ribozyme and amplified by PCR for further rounds of transcription. Selection pressure is increased by reducing the time of transcription per round from hours to less than a minute. Cloning and sequencing of active clones identifies ribozymes active for in-cis cleavage. Such ribozyme sequences are then redesigned for in-trans cleavage, which permits a quantitative kinetic comparison.

3. Search for alternative sequences cleaving at NUH triplets

Our search for hammerhead ribozymes with improved or altered cleavage properties for triplets began with the selection of ribozymes with an improved cleavage activity on the 3' side of the triplet AUA as a representative of the general formula NUH. The ten nucleotides in the core of the consensus ribozyme were randomised in these constructs. ^[16] Unfortunately, these efforts did not result in the selection of any active ribozyme although the same strategy using only three randomisations had been

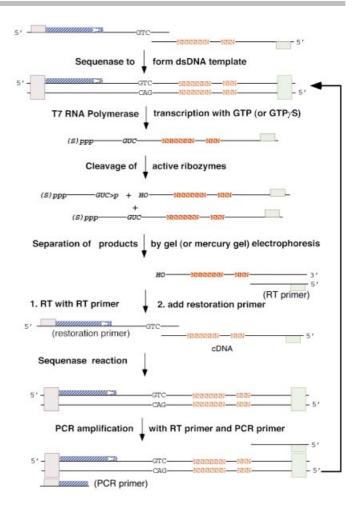


Figure 3. General scheme for in-cis selection of hammerhead ribozyme mutants. Grey- and green-shaded boxes = restriction sites, blue arrowed line = primer with T7 promoter, N = randomised positions, residues printed in bold = RNA. – ds = double-stranded, RT = reverse transcription.

successful in isolating active ribozymes of the satellite RNA of the Barley Yellow Dwarf virus. The difficulties apparently arose because transcription yielded transcripts of a size similar to that expected from the cleavage reaction either by premature termination or deletion during reverse-transcription polymerase chain reaction (RT-PCR). The desired cleavage products arise only from a small number of active sequences in the beginning of the selection, so that the pool for the RT-PCR amplification is dominated by the transcriptional byproducts preventing the accumulation of active ribozymes.

We therefore changed the selection protocol for a more active ribozyme directed against the triplet GUC by using GTP γ S instead of GTP for the initiation of the transcription. It was expected that separation of the transcription reaction by PAGE on a mercury-containing gel should retain all transcripts with an intact 5' end through the use of GTP γ S. Thus, not only inactive full-length transcripts should remain in the well but also prematurely terminated transcripts, thus eliminating at least one of the difficulties encountered before. Additionally, the double-stranded DNA (dsDNA) of the correct length resulting from the PCR amplification was purified from an agarose gel to further avoid contamination with PCR-derived deletion prod-

630 CHEMBIOCHEM **2001**, 2, 629–635

ucts. Indeed, this selection resulted in the isolation of active ribozymes. After five rounds, the most prominent change was either a uridine or a cytidine at position 7, which is known to be a variable position. Other sequences had an additional mutation at position 9 with a uridine instead of an adenosine. None of these ribozymes was more active than the consensus ribozyme, which interestingly was not found. Additional experiments indicated that this is likely due to the RT-PCR step, which shows the wild-type G-A tandem mismatch adjacent to stem II particularly prone to deletion. Thus, these results indicate that the consensus sequence derived from the viroid is probably the most active sequence. Support for this view comes from the work by Ishizaka et al.^[14] who searched for alternative sequences by the trans selection protocol and also ended up with the consensus sequence.

4. Search for more active sequences for GAC cleavage

The consensus hammerhead ribozyme cleaves on the 3' side of NAH sequences such as GAC with approximately 5% efficiency as compared to cleavage at GUC.^[18] Thus, there is a strong preference for the A15.1-U16.1 base pair over the inverted U15.1-A16.1 pair. We wondered whether this could be overcome by changes of certain nucleotides in the core. A selection for GAC cleavage was performed with ten randomisations for seven cycles as described above (see Section 3).^[19] Only one sequence active in cleaving the given triplet could be found (Figure 4). Clone 24 differed from the consensus sequence only at position 7 with a G instead of a U and had half the activity of the

consensus sequence (0.15 versus 0.3 min⁻¹). Unexpectedly, two ribozymes were identified in which the target triplet GAC had been changed to GUA although this region was not part of the randomised region and was part of the primers during the selection process. Thus, this mutation has probably arisen due to the poor fidelity of the Taq polymerase used for the PCR. Of these two ribozymes, clone 2, with A at position 7, was further investigated kinetically. It cleaved GAC very poorly, with a rate of 0.08 min⁻¹, but cleaved 3' to GUA quite efficiently with a rate of 1.0 min⁻¹. Thus, again this selection produced essentially the consensus sequence. This strongly suggests that this sequence was optimised by evolution. One restriction in our search for alternative sequences could of course be the limited sequence pool we generated by only randomising ten positions. However, this was dictated by our desire to be as close as possible to the consensus sequence to be able to make meaningful comparisons for the assignment of function to individual nucleotides.

Search for sequences that cleave at NUG triplets

The cleavage rate of the consensus ribozyme at the 3' side of a NUG triplet is barely measurable. Thus, a ribozyme with the capability of cleaving at this triplet is an obvious candidate for selection. The AUG triplet was chosen for the selection using the above described protocol (see Section 3) but with 22 randomisations in the core and in stem-loop II.^[20] After 13 cycles of transcription-amplification, several active sequences could be isolated (Figure 5). Among the most active was clone no. 48 (Figure 6). It cleaved at the 3' side of AUG with a rate of

5'-GGGAGACUGUAAACUUUUC GAC UCAGCCAGCUAUGAGGCUGA NNNNNNN		GUCCGCAAGGAC NNN UCGAAACUAGAAUU-3'		
	Consensus sequence	CUGAUGA	GAA	
	Clone No.			
GAC	24	CUGAGGA	GAZ	
GUA	2 (2x)	CUGAAGA	GAA	
GUA	8 (2x)	CUGAGGA	GAA	

Figure 4. Active sequences for GAC cleavage. Grey-shaded area = core region.

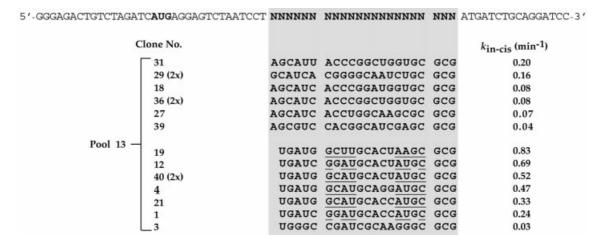


Figure 5. Active clones for AUG cleavage after 13 rounds of selection. The two-dimensional structure of ribozyme clone no. 40 is presented in Figure 6.

0.88 min⁻¹ in single-turnover experiments, which is similar to cleavage at the 3' side of GUC by the consensus ribozyme. Interestingly, this ribozyme was not specific for cleavage after G, but also cleaved at the 3' side of AUA and thus is a purine-specific cleaver. More surprising even was the observation that the central U of the triplet (position 16.1) can also be a C or a G with the appropriate base-pairing nucleotide at 15.1 being G or C, respectively.

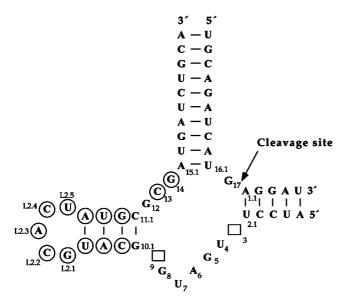


Figure 6. Two-dimensional structure representation of purine-specific hammer-head-like ribozyme clone no. 40. Circles = nucleotide changes from the consensus ribozyme, boxes = deletions.

There are a number of changes as compared to the consensus ribozyme which prevent an assignment of the importance of individual nucleotides for triplet specificity. An important functional difference is that the sequence in loop II cannot be changed without impairment of activity, whereas in the consensus ribozyme this can be replaced by a non-nucleotidic linker. Thus, even though the two-dimensional structure resembles that of the consensus ribozyme, this ribozyme is best described as hammerhead-like.

In search of a ribozyme which is more closely related to the hammerhead ribozyme, we used another method to arrive at a purine-specific ribozyme by subjecting the 22 nucleotides of the consensus sequence to evolution by mutagenic PCR with two nucleotide analogues^[21] (Figure 7). The two nucleotides were 6-(2-deoxy-β-D-ribofuranosyl)-3,4-dihydro-8*H*-pyrimido-[4,5-C]-[1,2]-oxazin-7-one (dP) and 8-oxo-2'-deoxyguanosine (8-oxodG). The nucleotide analogue dP pairs with A and G, leading to transitions, and oxodG pairs with A and C, which results in transversions. Surprisingly, this method led to a ribozyme identical to that obtained by the in vitro selection with 22 randomisations, with the exception of a C at position L2.5 instead of a U. The essentially identical sequence resulting from two quite different approaches strongly indicates that the sequence space for a purine-specific ribozyme of the hammerhead size is very limited.

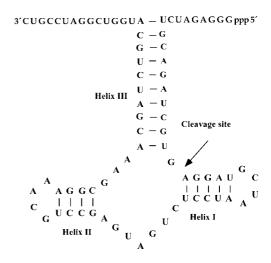


Figure 7. Two-dimensional structure representation of the starting ribozyme used for mutagenic evolution.

6. Attempts for the selection of sequences for cleavage of NGC triplets

The uridine at position 16.1 is required for optimal cleavage of triplets even though an A is tolerated, but with considerably lower activity. A cytidine at 16.1, however, is tolerated when 15.1 is occupied by an inosine.[22] Thus, a selection of a ribozyme which cleaves at the 3' side of triplets like UGC or AGC is most desirable. We attempted this first by using the protocol described above (see Section 3) with 22 randomisations. However, the selection was plagued by deletions and unfortunately did not result in active sequences. The method was therefore changed by introducing additional steps to improve the removal of deleted and other unspecified products (Figure 8). A 3'biotinylated oligonucleotide complementary to the 5' end of the transcripts was annealed to the material isolated by PAGE in the region of the expected cleavage product size. Purification on streptavidin beads was performed to retain inactive sequences and deleted products retaining the 5' end. The active sequences in the supernatant were incubated with a 5'-biotinylated oligonucleotide complementary to the 3' end of the cleavage product to separate them from non-specific RNA of the isolated length. Purification by streptavidin beads and subsequent RT and PCR directly on the beads produced the desired dsDNA. This was further purified by PAGE, which was far superior for length separation as compared to purification by agarose gel electrophoresis, and the DNA was reamplified by PCR to yield highpurity dsDNA in considerably better quality than obtained with the original method. Selection of two weak cis-cleaving ribozymes was successfully completed in eight cycles. Unfortunately, the in-trans constructs showed little activity. Thus, despite all the precautions to avoid contamination by non-active species in the amplification, no reasonably active ribozyme could be identified. The simple interpretation of this result is that no NGC cleavage activity can be accommodated in this hammerhead-size sequence space. This view is supported by the result of the additionally applied mutagenic PCR method, which also did not reveal any in-trans-active sequences.

632 CHEMBIOCHEM **2001**, 2, 629–635

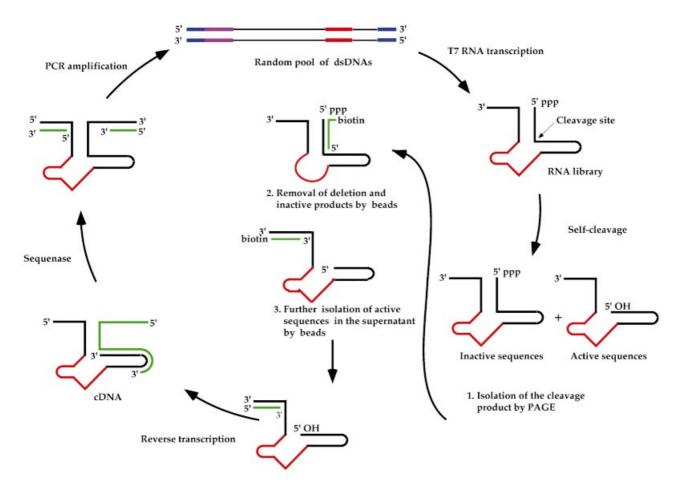


Figure 8. In-cis selection of ribozymes with an additional biotin - streptavidin affinity step.

7. Enlarged pool size

The failure of obtaining hammerhead ribozymes with significantly higher catalytic power for the cleavage of any triplet indicates the optimal natural evolution of this ribozyme within its size constraint. Different specificities apparently can only be obtained by ribozymes that do not entirely conform to the hammerhead structure, as demonstrated by the purine-specific hammerhead-like ribozyme described in Section 5. Thus, the size constraint seems to be considerable. However, ribozymes with different specificities should presumably be obtainable by increasing the sequence pool. [23] Indeed, when enlarging the randomised region to 40 nucleotides, twelve different classes of ribozymes with five different triplet cleavage activities could be isolated. Only one of these corresponded to the secondary structure of the consensus hammerhead ribozyme, except that it had an enlarged loop II (Figure 9).[24] Interestingly, this ribozyme also had an NUH triplet, UUA, as substrate, but with only a 30fold lower rate than the consensus ribozyme. Thus, this indicates again the close link between the consensus core sequence and the specificity for triplet cleavage. Other triplets not obeying the NUH rule such as AUG, UGA, CUG and UGG were also discovered as substrates, but their rates of cleavage were all well below that of the hammerhead ribozyme.

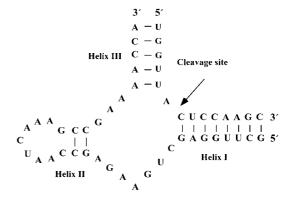


Figure 9. Two-dimensional structure representation of a hammerhead ribozyme selected from a large sequence pool.^[24]

8. Selection for ribozymes with lower Mg²⁺ requirement

The Mg²⁺ concentration for optimal catalytic activity of the hammerhead ribozyme is 10 mm. Independent of the precise role of the metal ion,^[5, 8, 25] ribozymes that function well at lower Mg²⁺ concentrations are highly desirable for their application for the sequence-specific inhibition of gene expression in vivo^[26] because the concentration of free Mg²⁺ in mammalian cells is

CHEMBIOCHEM **2001**, 2, 629–635 **633**

only approximately 500 μM, depending somewhat on the type of cell. [27, 28] Conaty et al. [15] have undertaken an in-trans selection based on a so-called minizyme, a version of the consensus hammerhead ribozyme with only one base pair in stem II, G10.1-C11.1. Selection, with increasing selection pressure by lowering the Mg²⁺ concentration, was performed with constructs in which all 18 positions of the core and stem – loop II were randomised. All selected active ribozymes had the central core of the consensus ribozyme with the G10.1.-C11.1 base pair (Figure 10).

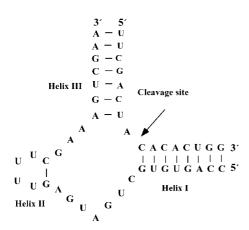


Figure 10. Two-dimensional structure representation of a minizyme with low Ma^{2+} concentration requirement. [15]

The most active species had the loop II tetranucleotide sequences UUUU, UUAA or UCCA. These ribozymes indeed had an activity of approximately $1 \, \text{min}^{-1}$ at $1 \, \text{mm} \, \text{Mg}^{2+}$, which is comparable to that of the consensus ribozyme at $10 \, \text{mm}$. As other tetraloop sequences had lower activity, this suggests that this part of the structure is involved in determining the requirement for the metal ion. As no metal ion has been observed in this region in any of the three-dimensional structures or been implicated by solution studies, the explanation for this effect remains to be elucidated. Such minizymes should certainly be beneficial for in vivo inhibition studies.

9. Potential applications

Although the main aim of these selections was to obtain insight into structure – function relationships within the hammerhead ribozyme, there is another aspect of this work which pertains to the application of ribozymes for the sequence-specific inhibition of gene expression. [26, 29–33] Although triplets that can be cleaved by the consensus ribozyme are abundant in any mRNA, its secondary structure renders their accessibility for ribozymes and oligonucleotides very limited. [34–36] An expansion of the triplet cleavage repertoire is expected to increase the probability of finding optimal regions for cleavage by ribozymes and thus their application for mRNA cleavage. The purine-specific ribozyme described above is a good example for such a case.

10. Conclusion

This article describes several selection strategies that might be used in other types of RNA selections. Although these selection experiments have not helped in gaining a better understanding of the structure-function relationships of the hammerhead ribozyme, they have nevertheless provided some interesting additional information. Thus, they strongly indicate that the core sequence of the consensus ribozyme seems to be optimal for the cleavage of triplets of the NUH rule within the sequence space of the hammerhead. Apparently, this sequence has a unique catalytically competent three-dimensional structure that has little tolerance for changes in sequence, including the sequence of the cleavage site. Several examples from the selections, such as the purine-specific ribozyme, [19, 20] those obtained from a large sequence pool $^{\![24]}$ and the minizyme operating at low Mg^{2+} concentration[15] suggest that the sequence of loop II is more important in maintaining optimal cleavage rates than had previously been assumed.[37, 38] These results will have to be accomodated in a future structure-function model of this ribozyme.

Besides these mechanistic aspects, the selection of a purine-specific ribozyme—even though it is not a bona fide hammer-head ribozyme—has some practical relevance for the application of the ribozyme for the inhibition of gene expression. ^[20, 21] This also applies to the discovery of the minizymes requiring low Mg²⁺ concentrations. ^[15] Thus, the selection experiments have yielded some interesting information on the hammerhead ribozyme as well as some improvements for practical applications.

- [1] S. T. Sigurdsson, J. B. Thomson, F. Eckstein in RNA Structure and Function (Ed.: R. W. Simons, M. Grunberg-Manago), Cold Spring Harbor Press, Cold Spring Harbor, 1998, pp. 339 – 376.
- [2] D. M. J. Lilley, Curr. Opin. Struct. Biol. 1999, 9, 330.
- [3] E. A. Doherty, J. A. Doudna, Annu. Rev. Biochem. 2000, 69, 597.
- [4] K. R. Birikh, P. A. Heaton, F. Eckstein, Eur. J. Biochem. 1997, 245, 1.
- [5] F. Eckstein, B. Bramlage, *Biopolymers* **1999**, *52*, 147.
- [6] R. H. Simons, Annu. Rev. Biochem. 1992, 61, 641.
- [7] O. C. Uhlenbeck, Nature 1987, 328, 596.
- [8] J. Haselof, W. L. Gerlach, Nature 1988, 334, 585.
- [9] H. W. Pley, K. M. Flaherty, D. B. McKay, Nature 1994, 372, 68.
- [10] J. E. Wedekind, D. B. McKay, Annu. Rev. Biophys. Biomol. Struct. 1998, 27, 475
- [11] J. B. Murray, D. P. Terwey, L. Maloney, A. Karpeisky, N. Usman, L. Beigelman, W. G. Scott. Cell 1998, 92, 665.
- [12] D. M. J. Lilley, T. J. Wilson, Curr. Opin. Chem. Biol. 2000, 4, 507.
- [13] M. Menger, F. Eckstein, D. Porschke, Nucleic Acids Res. 2000, 28, 4428.
- [14] M. Ishizaka, Y. Ohshima, T. Tani, Biochem. Biophys. Res. Commun. 1995, 214, 403.
- [15] J. Conaty, P. Hendry, T. Lockett, Nucleic Acids Res. 1999, 27, 2400.
- [16] K. L. Nakamaye, F. Eckstein, *Biochemistry* **1994**, 33, 1271.
- [17] N. K. Vaish, P. A. Heaton, F. Eckstein, *Biochemistry* **1997**, *36*, 6495.
- [18] A. R. Kore, N. K. Vaish, U. Kutzke, F. Eckstein, Nucleic Acids Res. 1998, 26, 4116.
- [19] A. R. Kore, C. Carola, F. Eckstein, Bioorg. Med. Chem. 2000, 8, 1767.
- [20] N. K. Vaish, P. A. Heaton, O. Fedorova, F. Eckstein, Proc. Natl. Acad. Sci. USA 1998, 95, 2158.
- [21] A. R. Kore, N. K. Vaish, J. A. Morris, F. Eckstein, J. Mol. Biol., **2000**, *301*, 1113.
- [22] J. Ludwig, M. Blaschke, B. S. Sproat, *Nucleic Acids Res.* 1998, 26, 2279.
- [23] P. C. Sabeti, P. J. Unrau, D. P. Bartel, *Chem. Biol.* **1997**, *4*, 767.
- [24] J. Tang, R. R. Breaker, Proc. Natl. Acad. Sci. USA 2000, 97, 5784.
- [25] R. Hanna, J. A. Doudna, Curr. Opin. Chem. Biol. 2000, 4, 166.

- [26] B. Bramlage, E. Luzi, F. Eckstein, Trends Biotechnol. 1998, 16, 434.
- [27] A. Romani, A. Scarpa, Arch. Biochem. Biophys. 1992, 298, 1.
- [28] R. E. London, Annu. Rev. Physiol. 1991, 53, 241.
- [29] B. Bramlage, E. Luzi, F. Eckstein, Nucleic Acids Res. 2000, 28, 4059.
- [30] A. R. Muotri, L. da Veiga Pereira, L. dos Reis Vasques, C. F. Menck, Gene 1999, 37, 303.
- [31] H. A. James, I. Gibson, *Blood* 1998, *91*, 371.
- [32] L. Q. Sun, M. J. Cairns, E. G. Saravolac, A. Baker, W. L. Gerlach, *Pharmacol. Rev.* 2000, 52, 325.
- [33] K.-Y. Jen, A. M. Gerwirtz, Stem Cells 2000, 18, 307.

- [34] M. Amarzguioui, G. Brede, E. Babaie, M. Grotli, B. Sproat, H. Prydz, Nucleic Acids Res. 2000, 28, 4113.
- [35] K. R. Birikh, B. A. Yuri, M. Soreq, F. Eckstein, RNA 1997, 3, 429.
- [36] M. Scherr, M. Reed, C.-F. Huang, A. D. Riggs, J. J. Rossi, Mol. Ther. 2000, 2, 26
- [37] T. Tuschl, F. Eckstein, Proc. Natl. Acad. Sci. USA 1993, 90, 6977.
- [38] D. J. Fu, F. Benseler, L. W. McLaughlin, J. Am. Chem. Soc. 1994, 116, 4591.

Received: January 17, 2001 [M 182]