NOTE TO USERS

This reproduction is the best copy available.

UMI®

DISSERTATION

SYNTHESIS AND STUDY OF C-NUCLEOTIDE ANALOGUES OF FORMAMIDOPYRIMIDINE LESIONS IN DNA

Submitted by
Michael O. Delaney
Department of Chemistry

In partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

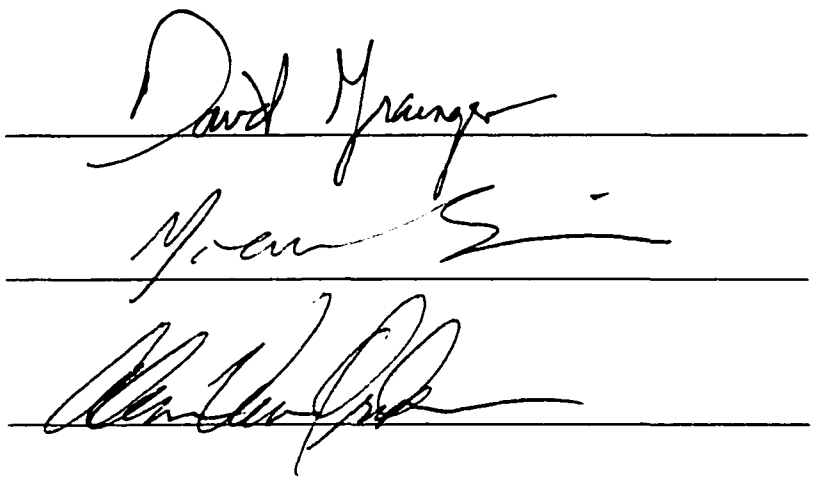
Colorado State University
Fort Collins, Colorado
Fall 2002

COLORADO STATE UNIVERSITY

June 14, 2002

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY MICHAEL O. DELANEY ENTITLED SYNTHESIS AND STUDY OF C-NUCLEOTIDE ANALOGUES OF FORMAMIDOPYRIMIDINE LESIONS IN DNA BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSPOPHY.

Committee on Graduate Work



Advisor

Department Chair

Abstract of Dissertation

SYNTHESIS AND STUDY OF C-NUCLEOTIDE ANALOGUES OF FORMAMIDOPYRIMIDINE LESIONS IN DNA

Formamidopyrimidine (Fapy) lesions are produced in DNA through the reaction of purines under oxidative stress conditions. Fapy•dA and Fapy•dG may exist as a mixture of anomers in DNA, and the equilibration between α- and β-anomers has been determined to be a rapid process. Fapy lesions may play an important role in mutagenic events and the onset of genetic diseases such as cancer. Another factor that suggests the importance of Fapy lesions is that they are substrates for the DNA repair protein Formamidopyrimidine DNA glycosylase (Fpg). In order to better understand the potential effects of the individual α- or β-anomers of Fapy lesions on DNA polymerases and repair proteins, non-epimerizable C-nucleotide analogues α-**18**, β-**18** and β-**19** were studied in DNA.

The use of α-**18**, β-**18** and β-**19** as structural probes of Fapy lesions was supported by molecular modeling results. Only minor variations exist between the analogues and the Fapy lesions. UV-melting studies of duplexes containing α-**18**, β-**18** and β-**19** do not show the same similarities with Fapy•dA and Fapy•dG and reveal that there is no preference for stabilization opposite any nucleotide.

Klenow (exo$^-$) prefers to incorporate the correct dNTP opposite β-**18** and β-**19**. However, misincorporation of dATP is also observed. Comparison of the relative

misinsertion frequencies of β-**18** and β-**19** to Fapy•dA and Fapy•dG respectively reveal similar preferences. This suggests that the β-anomer is the biologically relevant diastereomer of both Fapy lesions. The ability to accommodate the insertion of dATP opposite β-**18** or β-**19** can be rationalized through the adoption of the *syn*-conformation. The formamide can present a "T-like" base-pairing pattern that can induce Klenow (exo$^-$) to insert dATP.

The interactions of Fpg protein with duplex DNA containing β-**18** and β-**19** opposite the four native nucleotides are very similar to investigations involving Fapy•dA and Fapy•dG. The Fpg protein binds the duplexes containing the β-**18** opposite any of the four nucleotides with a K_d <27 nM. The Fpg protein discriminates between duplexes containing β-**19**, and only binds to the β-**19**:dC duplex (K_d = 11.8 nM).

Michael O. Delaney
Department of Chemistry
Colorado State University
Fort Collins, CO, 80523
Fall 2002

Table of Contents

List of Tables

List of Supplementary Tables

List of Figures

List of Supplementary Figures

Abbreviations:

AcN	Acetonitrile
AcOH	Acetic acid
AFA	Acetic formic anhydride
AP	Apurinic-apyrimidinic
BINAP	2,2'-Bis(diphenylphosphino)-1,1'-binapthyl
BER	Base excision repair
BSA	Bovine serum albumin
BSTFA	*N*-bis(trimethylsiliyl) trifluoroacetamide
CD	Circular dichroism
DCI	Dicyanoimidazole
DCM	Dichloromethane
DEA	Diethylaniline
DIBAL-H	Diisobutylaluminum hydride
DMAP	4,4-Dimetylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DMT	4,4'-Dimethoxytrityl
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide triphosphate

dsDNA	Double stranded DNA
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDCI	1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide
Endo III	Endonuclease III
ESI-MS	Electrospray ionization mass spectrometry
EtOAc	Ethylacetate
Et_2O	Diethylether
EtOH	Ethanol
FAB	Fast atom bombardment
Fpg	Formamidopyrimidine DNA glycosylase
FT-IR	Fourier transform-infared
GC-MS	Gas chromatography associated with mass spectrometry
Glu	Glutamate
HEPES	[4-(2-hydroxyethyl)-1-piperazineethansulfonic acid]
HPLC-EC	High performance liquid chromatography coupled to electrochemical detection
HPLC-MS/MS	High performance liquid chromatography associated with tandem mass spectrometry detection
HRMS	High resolution mass spectrometry
*i*PrPAC	4-isopropylphenoxyacetyl
KHMDS	Potassium bis(trimethylsilyl)amide
LDA	Lithium diisopropylamide
LRMS	Low resolution mass spectrometry

Lys	Lysine
MeOH	Methanol
MTBDSTFA	*N*-tertbutyldimethylsilyl-*N*-methyl-trifluoroacetamide
NaOAc	Sodium Acetate
NaOMe	Sodium methoxide
NaOEt	Sodium ethoxide
NER	Nucleotide excision repair
NHE	Normal hydrogen electrode
NMI	*N*-metylimidazole
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
PAC	Phenoxyacetyl
PAGE	Polyacrylamide gel electrophoresis
PDC	Pyridinium dichromate
PDG	Pyrimidine dimer glycosylase
PGSA	Polyacrylamide gel-shift assay
PIPES	Piperazinebis(ethansulfonic acid)
PyBOP	(benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
Pro	Proline
Pyr	Pyridine
ROS	Reactive oxygen species
RNA	Ribonucleic acid
Rxn	Reaction

ROSEY	Rotating frame Nuclear Overhauser effect spectroscopy
SAR	Structure activity relationship
ssDNA	Single stranded DNA
ST	Saline-tris buffer
TBDMS	Tertbutyldimethylsilyl
TBDPS	Tertbutydiphenylsilyl
TBE	Tris-borate-EDTA buffer
TFA	Trifluoroacidic acid
TfOH	Trifluoromethanesullfonic acid
TMS	Trimethylsilyl
TIPDS	1,3-(1,1,3,3-tetraisopropyldisiloxanylidene)
THF	Tetrahydrofuran
Tris	[Tris(hydroxymethyl)aminomethane]
UV	Ultraviolet
W-C	Watson-Crick

1 Introduction:

Deoxyribonucleic acid (DNA) is an important molecule in the life cycle of living organisms. It serves as the molecular blueprint for cells and encodes specific cellular function. DNA provides a mechanism by which organisms can express biological characteristics and also furnishes a method for genetic continuity from one generation to the next. The structure of DNA is remarkably stable, which facilitates its function as a hereditary agent. However, it is susceptible to alteration by various factors that are found within its environment such as by products of aerobic respiration, UV radiation and chemical agents.[1] DNA damage by environmental agents can alter the structure of DNA and depending upon the nature of the damage, DNA function can be disrupted. This can lead to mutations in the genome if not properly repaired.[2] Mutations result in loss of the genetic integrity of DNA and can also have severe implications on the synthesis of molecules coded for by the affected gene.

DNA damage occurs through a variety of conditions that are endogenous and exogenous to the cell and have been attributed to many different disease states within living organisms. Endogenous factors include free radicals released either by the incomplete reduction of molecular oxygen during aerobic respiration or redox active metal chelates.[3,4] Exogenous agents that damage DNA directly and indirectly include ionizing and ultraviolet radiation, as well as chemical agents. Indirect damage of DNA results from the formation of the hydroxyl radical and other reactive oxygen species (ROS) that are formed from exogenous agents.[2,5] Accumulation of genetic damage

can result in a loss of genetic integrity and the development of disease. Diseases associated with genetic damage are extensive and include cancer, heart disease, diabetes, chronic inflammation, and hereditary syndromes.[6-11] Genetic damage has also been implicated with diseases associated with aging.[12-15] By understanding the role that DNA damage plays in the development of disease states, potential treatments can be developed.

Nucleic acids are continuously subjected to environmental stress which modifies DNA. DNA repair is necessary to prevent loss of genetic information and the onset of disease. Cells have elaborate defense mechanisms in place that protect DNA from damage. Enzymes degrade ROS that damage DNA. For example, superoxide dismutase and catalase work in tandem to decompose superoxide to water and oxygen.[2] Antioxidants (e.g. Vitamin E) also reduce the amount of free radicals present. Furthermore, cells have intricate DNA repair processes that are responsible for the recognition and repair of specific types of DNA damage when they occur. Depending upon the nature of the DNA lesion, different mechanisms exist for the repair of DNA. The damaged nucleobase can be excised through base excision repair (BER), or the entire nucleotide can be excised through nucleotide excision repair (NER).[16]

Oxidative damage to DNA occurs through the action of ROS, namely the hydroxyl radical ($HO^{\bullet}$), which has been implicated in the formation of more than 100 different DNA lesions.[17] Damage occurs at both the nucleobases and sugar moieties in DNA. The reaction with the sugar moiety occurs through hydrogen atom abstraction that results in direct strand cleavage and alkali labile lesion formation.[5,18] The reaction of the nucleobases occurs through the addition of $HO^{\bullet}$ to double bonds at or near diffusion controlled rates.[5,19] Two classes of lesions that result from ROS reacting with purines

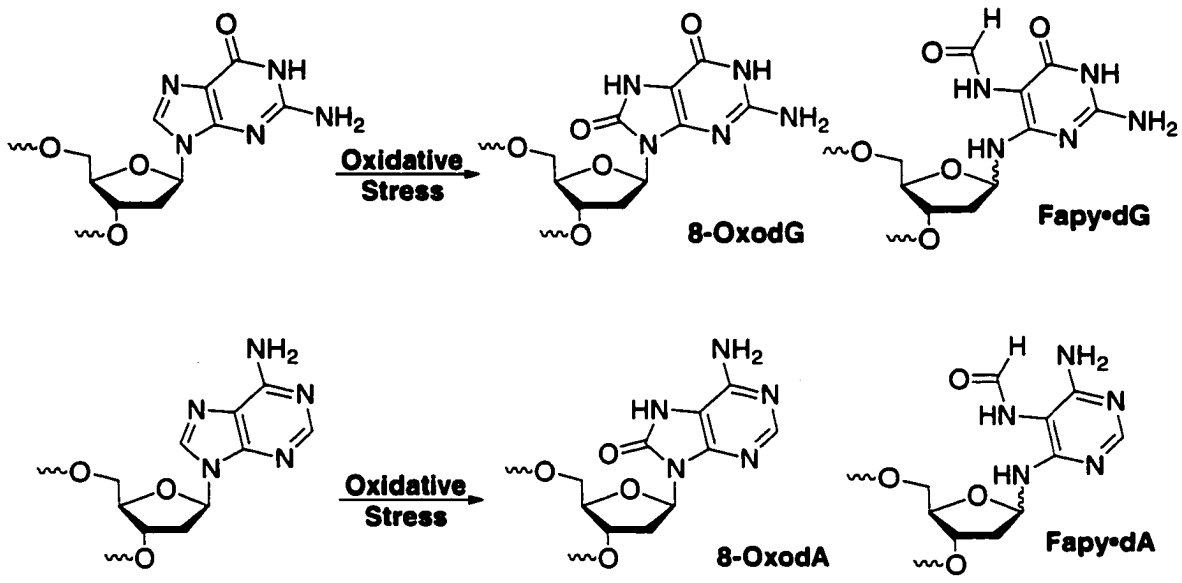


Figure 1. The products of oxidative damage to purines.

are oxopurines (8-OxodG and 8-OxodA) and formamidopyrimidines (Fapy•dG and Fapy•dA) (Figure 1). The biological significance of purine lesions in DNA is indicated by their relative 2-fold abundance over pyrimidine lesions.[20]

The potential mutagenicity of DNA lesions depends upon their interaction with DNA polymerases and repair proteins. If the DNA repair system recognizes the lesion, it can replace the damaged lesion through a BER pathway (if the damaged site occurs on the nucleobase).[16] If the DNA repair system fails to recognize the lesion the damaged site will remain to potentially affect biological processes. DNA polymerases are responsible for the replication of DNA. Some polymerases possess DNA repair features to enhance their proofreading ability and the overall fidelity of DNA synthesis.[21] When a polymerase comes in contact with a DNA lesion it can do one of two things. The polymerase can terminate DNA synthesis and dissociate from the DNA template. In this case, the lesion serves as block and unless repaired by a DNA repair process will prevent

the DNA from being replicated. Translesional DNA synthesis can result and depending upon what hydrogen-bonding pattern is presented by the lesion, the polymerase can incorporate the wrong nucleotide triphosphate.[22] Enzymes exist for repairing mismatches incorporated by polymerases. A specific mismatch repair process for the misinsertion of dA opposite dG and for other misinsertions opposite DNA lesions uses a BER mechanism to repair the mistake.[16] Other DNA repair processes recognize mismatches incorporated by polymerases, but generally require the removal of long stretches of DNA followed by resynthesis of the DNA.[21]

Active site requirements of DNA repair proteins have been probed using non-hydrolyzable analogues of DNA lesions to gain further insight into molecular recognition requirements of the binding event.[16] Advantages of such investigations are that the DNA repair protein can be competitively inhibited by the analogues, and binding affinities for their recognition can be determined. Additional modifications to substrate analogues can bring out important features involved in the molecular recognition of specific DNA lesions by removing potential hydrogen-bonding sites.[16,23] Most DNA repair proteins have an ability to recognize several different classes of DNA lesions.[24] Information about the generality of the binding event can also be acquired through the studies with non-hydrolyzable analogues of DNA lesions.

The interaction of DNA damages with DNA repair enzymes and DNA polymerases is an important area of study and can have profound effects on determining the etiology of disease. DNA lesions impart new structure and function on the oligonucleotides that they occupy. By studying how the structure and function of DNA is

affected by DNA lesions, information about the potential mutagenicity and recognition by DNA repair enzymes can be obtained.

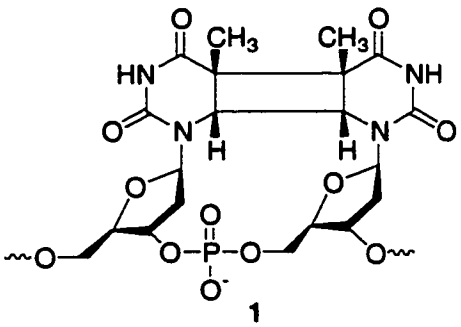
2 Background:

2.1 Formation of Formamidopyrimidine (Fapy) Lesions:

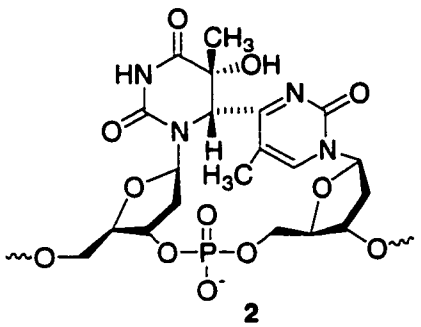
Fapy lesions are generated under a variety of conditions found within the cell. Endogenous and exogenous factors are involved in the formation of ROS that have been linked to the generation of Fapy lesions.[25,26] Generation of Fapy lesions can occur through electron transfer processes, the action of singlet oxygen, or through reaction with hydroxyl radical.[25-29] Fapy lesions have been detected in a variety of different disease states. They have been shown to be broadly present in cancerous tissues and have also been detected in Alzheimer's patients.[12,30,31]

2.1.1 Generation of Reactive Intermediates and Other Processes that Damage DNA:

Exogenous factors involved in the formation of Fapy lesions are sources of radiation, both ionizing and UV. UV and ionizing radiation can act in one of two ways to produce DNA damage.[5] They can react directly with the DNA molecule leading to electron transfer processes and subsequent damage. UV radiation (254 nm) can also excite nucleobases to excited states that stimulates the formation of photochemical cycloaddition products between adjacent pyrimidines.[2] The major products of this photocyclization are photocyclobutanedimers between two thymidines (**1**). Another product that occurs is the [2 + 2] photocyclization with the C-4 carbonyl of 3' pyrimidine and the C-5 / C-6 π-bond of the 5' pyrimidine. This forms an oxetane intermediate that decomposes to **2**.[32] By going to higher energy UV-light, direct strand breaks in DNA are



1



2

observed.[19] In an indirect manner, UV radiation can promote electron transfer events in DNA through the presence of photosensitizers. Photosensitizers reduce the intensity of radiation required for DNA damage. They promote energy transfer by absorption of the radiation and promotion of the sensitizer to an excited state (Equation 2.1). This excited state molecule can initiate a chemical reaction with DNA through the transfer of the absorbed energy (Equation 2.2). Biological photosensitizers include riboflavins and

$$\mathbf{PS} \xrightarrow{h\nu} \mathbf{PS^*} \qquad \textbf{Equation 2.1}$$

$$\mathbf{PS^*} + \mathbf{DNA} \longrightarrow \mathbf{DNA^*} + \mathbf{PS} \qquad \textbf{Equation 2.2}$$

porphyrins and biologically active drugs such as tetracycline react in a similar fashion.[33] The result of an electron transfer event is the formation of an intermediate radical cation that is responsible for the generation of DNA damage. Electron transfer processes predominately occur at deoxyguanosine due to its lower oxidation potential than deoxyadenosine (1.29 and 1.42 V vs NHE, respectively).[34]

Singlet oxygen ($^1O_2$) is another damaging agent that can arise through photoexcitation of molecular oxygen by UV light.[35] Biological photosensitizers are believed to play a role in the generation of $^1O_2$. DNA damage induced by $^1O_2$ is associated with a large degree of nucleobase modification with very little strand scission

occurring. Singlet oxygen can react with DNA nucleobases via cycloadditions with π-bonds in a 1,2, 1,3, and 1,4 manner.[29] With the exception of pyrimidine dimers that are formed from direct UV-light damage, the generation of $^1O_2$ is believed to be the largest source of DNA modifications resulting from UV-light.

Ionizing radiation can also damage DNA indirectly by reacting with water. The predominance of water in living systems suggests that it will interact with ionizing radiation more often than DNA.[5] Electronic excitation of water can occur (Equation 2.3), or direct ionization to form water cation radical (Equation 2.4) which forms ROS.

$$H_2O \xrightarrow{h\nu} H_2O^* \longrightarrow H^\bullet + HO^\bullet \qquad \textbf{Equation 2.3}$$

$$H_2O \xrightarrow{h\nu} H_2O^{\bullet+} + e^- \qquad \textbf{Equation 2.4}$$

$$H_2O^{\bullet+} + H_2O \longrightarrow H_3O^+ + HO^\bullet \qquad \textbf{Equation 2.5}$$

Excitation of water leads to an unstable intermediate that homolytically cleaves to form hydrogen atom and $HO^\bullet$ (Equation 2.3). The radical cation of water formed from direct ionization readily protonates the surrounding water molecules to form hydroxyl radical (Equation 2.5). Both of these processes lead to the formation of the highly reactive hydroxyl radical. The electron that is released in Equation 2.4 can also participate in DNA damage. However, solvated electrons are not as reactive as $HO^\bullet$.[5] In the presence of N_2O, the free electrons can be converted to hydroxyl radicals (Equation 2.6). By saturating a solution with N_2O, a system that consists of 90% hydroxyl radicals can be formed upon γ-radiolysis.[5]

$$H_2O + e^-_{aq} + N_2O \longrightarrow HO^\bullet + N_2 + HO^- \qquad \textbf{Equation 2.6}$$

Hydroxyl radicals are also generated through transition metal mediated processes. The transition metal generally involved in the reduction of hydrogen peroxide to hydroxyl radical is Fe^{+2}. This transformation is known as the Fenton reaction (Equation 2.7), wherein Fe^{+2} donates an electron to hydrogen peroxide to form a radical anion. Following homolytic cleavage of the radical anion, hydroxide and $HO^{\bullet}$ are formed. Superoxide ($O_2^{\bullet -}$) is also converted into the hydroxyl radical via the Haber-Weiss reaction

$$Fe^{+2} + H_2O_2 \longrightarrow Fe^{+3} + H_2O_2^{\bullet -} \longrightarrow HO^{-} + HO^{\bullet} \qquad \textbf{Equation 2.7}$$

$$O_2^{\bullet -} + H_2O_2 \longrightarrow HO^{\bullet} + HO^{-} + O_2 \qquad \textbf{Equation 2.8}$$

to form hydrogen peroxide (Equation 2.8).[5] This is then cycled through the Fenton reaction to generate hydroxyl radical.

The reaction of $HO^{\bullet}$ with DNA is known to produce many different types of damage. Damage includes base modifications, single strand breaks, double strand breaks, abasic sites, and protein-DNA crosslinks. Hydroxyl radical reaction with nucleobases of DNA is well studied.[5,19,36] The addition of $HO^{\bullet}$ to purines and pyrimidines is a very facile process occurring at greater than 10^9 $M^{-1} \bullet s^{-1}$. In the case of purines, addition occurs at the C4, C5 and C8 positions. Several products arising from

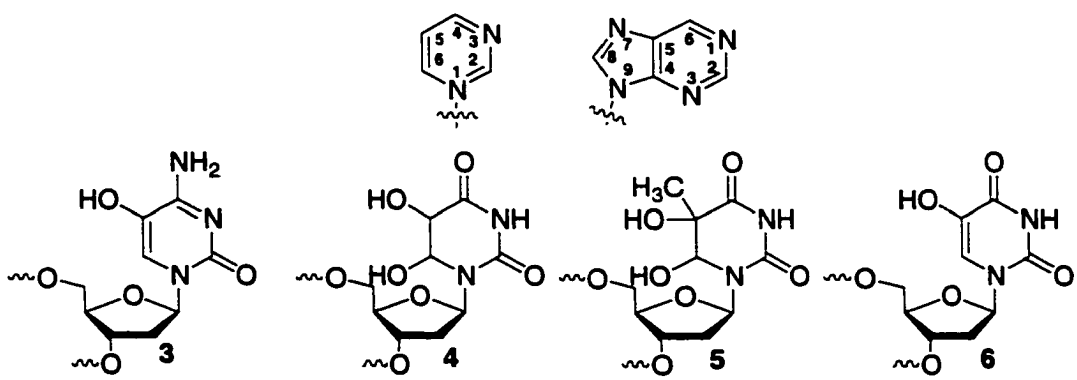


Figure 2. Pyrimidine damages from the action of hydroxyl radical.

the reaction of $HO^{\bullet}$ with purines and pyrimidines are known (Figure 1, 2).

The highly reactive nature of the $HO^{\bullet}$ prevents it from diffusing far in solution. The diffusion distance of $HO^{\bullet}$ has been approximated to be 4-6 nm and reaction typically occurs with the first molecule that it comes in contact with.[37] Under physiological conditions, superoxide and hydrogen peroxide do not damage DNA.[38,39] However, hydrogen peroxide and superoxide are diffusible species that generate hydroxyl radical upon contact with metal species (vide supra). Metal ions present in the chromatin can promote the formation of $HO^{\bullet}$ in close proximity to DNA. This predisposes the DNA for damage whenever hydrogen peroxide or superoxide approach DNA.

2.1.2 Methods of Detecting DNA and Quantifying DNA Damage:

Detection and quantification of DNA damage is an essential prerequisite for understanding the role of DNA lesions in the development of disease. The effects of reaction conditions on product distribution resulting from damaging agents can also shed light on the mechanisms involved in the formation of these lesions. Reactivity patterns have been revealed through treatment of DNA with specific damaging agents and can also be used in forming DNA damage profiles. The measurement of DNA lesions requires very sensitive techniques since DNA lesions are typically observed at a ratio of less than 1 lesion per 10^5-10^6 normal bases.[17,26,40]

Two main methods for quantifying DNA damage are gas chromatography mass spectrometry (GC-MS) and high-performance liquid chromatography with electrochemical detection (HPLC-EC). Recent experiments have coupled the HPLC separation to electrospray ionization mass spectrometry (ESI-MS) to detect the masses of

OxodG between different labs.[44,46,47] The use of HPLC-EC has diminished the problem of over oxidation and consistent measurements can be obtained.[47]

The second method, HPLC-EC requires enzymatic digestion of DNA into mononucleosides.[41,42,47,48] Digestion is typically performed with nuclease P1 or other phosphodiesterases that are tolerant to base damages to cleave the 3' phosphodiester linkage. Subsequent treatment with alkaline phosphatase removes the 5' phosphate to give the free nucleoside. Complete digestion of DNA is necessary for accurate quantification of DNA lesion.[48] Digestion conditions are optimized for complete deglycosylation of lesions that may inhibit cleavage by nucleases. The resulting nucleosides are separated using HPLC and quantified through an electrochemical detector that observes changes in current.

Another technique used to detect DNA damage is the use of single cell electrophoresis or the Comet assay.[49,50] This method detects the formation of strand breaks in DNA and can be used in concert with glycosylases that are specific for certain types of nucleobase damage. Therefore, strand breaks will only be generated at sites where the glycosylase recognizes a damage site. This technique cannot quantify the number of DNA damages as precisely as the techniques mentioned above.[51] The Comet assay offers a quick and efficient way to determine the relative amount of DNA damage occurring. An advantage of this technique is that only low doses of γ-radiation are required for detection of strand breaks by fluorescence microscopy. It can be used to determine if a particular damaging agent is involved in the generation of one or more lesions.

2.1.3 Generation of Formamidopyrimidine and Oxopurine Lesions:

Oxidative stress mechanisms are involved in purine damage through the processes and species described above. Oxopurines and formamidopyrimidines are formed in concert with one another and are believed to arise from a common intermediate. Varying ratios of products are formed depending upon the conditions (i.e oxidizing vs. reducing—Section 2.1.4). An intermediate radical species is generated through addition of the $HO^{\bullet}$ to the C8 position of the purine (Figure 3). This same intermediate can arise through an

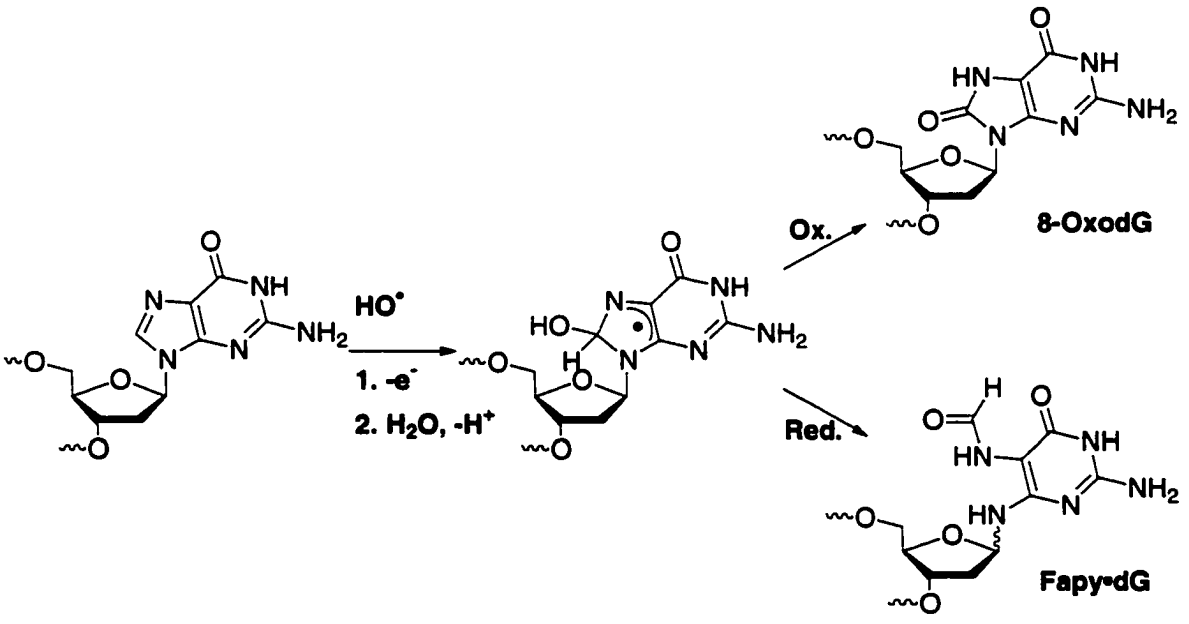


Figure 3. Formation of Fapy•dG and 8-OxodG through a common intermediate.

electron transfer process described above. Following the loss of an electron, water traps the intermediate radical cation to arrive at the same intermediate described for $HO^{\bullet}$ addition. The fate of this species depends greatly upon the reaction conditions. If an oxidizing environment is present, the radical can be trapped by oxygen (Figure 4).

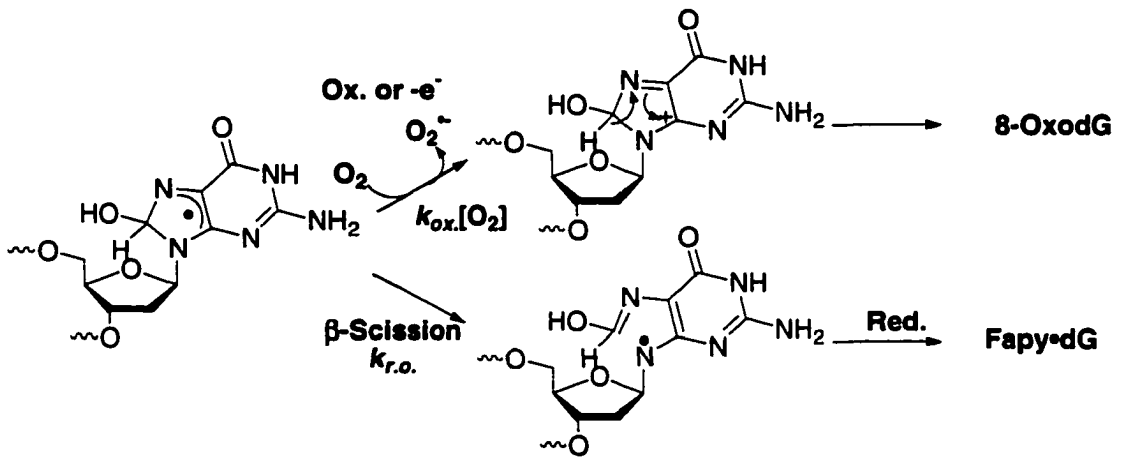


Figure 4. Conversion of Radical intermediate into Fapy•dG or 8-OxodG.

Loss of superoxide is formally a one electron oxidation and generates a cation. This same cation can be generated through the transfer of an electron to an adjacent metal species from the radical. Loss of a proton from the C8 position of the resulting cation forms the requisite oxopurine. In a reducing environment, the radical can undergo β-scission. Subsequent reduction of the radical forms the formamidopyrimidine. Steenken has rationalized that in the presence of oxygen the ring opening process to form Fapy lesions should be competitive with the oxidation step.[36,52] Measurement of the rates of oxidation and ring-opening ($k_{ox.}$ and $k_{r.o.}$ respectively—Figure 4) support this point. Since trapping of the radical with oxygen is diffusion controlled ($k_{ox.} = 4 \times 10^9\ M^{-1}s^{-1}$), and the concentration of O_2 in solution is 0.2 mM, the overall rate of the process is $8 \times 10^5\ s^{-1}$. This is very close to the rate of ring-opening which was measured in the absence of O_2 and reported as $k_{r.o.} = 2 \times 10^5\ s^{-1}$.[52] Thus, under normal aerobic conditions, it is predicted that both Fapy lesions and oxopurines should be present upon DNA damage with hydroxyl radical. Formation of Fapy•dA and 8-OxodA occur in a similar fashion.

However, due to the increased oxidation potential of dA versus dG (vide supra), electron transfer processes are less efficient at generating the intermediate radical cation.

Formation of hydroxyl radical adducts at the C4 and C5 positions of purines are less significant.[36,53] These adducts do not give rise to damage in the same fashion as addition to the C8 position. Instead, they dehydrate to form one electron oxidized purines (Figure 5).[36,53,54] The fate of the resulting species depends upon the conditions present.

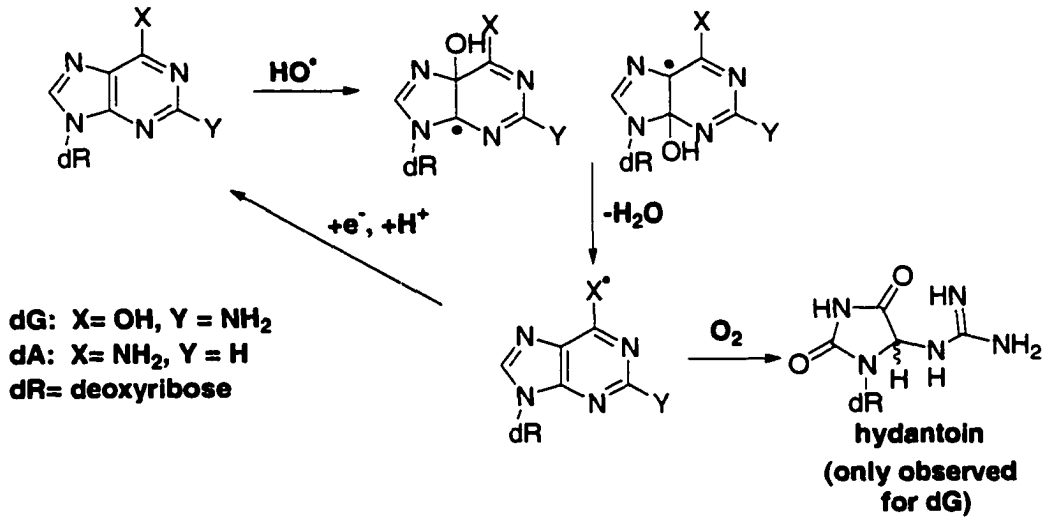


Figure 5. Hydroxyl radical adducts to the C4-C5 bond of purines.

Reduction of these species regenerates undamaged purines. Alternatively, the intermediate derived from deoxyguanosine can react with O_2 to produce a variety of different oxidation products.[54,55,56] Hydantoin is one of these products that has been studied by Burrows.[56]

2.1.3.1 Product Distribution of Fapy Lesions and Oxopurines in Nucleosides and Oligonucleotide Studies:

Product distribution studies for the reaction of purine nucleosides with ROS supports the proposed formation of Fapy lesions and oxopurines in Figure 4. In studies performed by Cadet on deoxyguanosine, under reducing conditions, the levels of 8-OxodG are decreased while the amounts of Fapy•dG are elevated. The reducing agent used was 1 mM cysteine to give a ratio of ~49:1 Fapy•dG: 8-OxodG.[27] These observations were corroborated in DNA.

In studies by Dizdaroglu, the importance of the radical environments for product distribution was investigated on mammalian chromatin *in vitro*.[57] The DNA solution was saturated with argon, air, N_2O, or N_2O/O_2 and exposed to a ^{60}Co γ-source. The DNA was analyzed by GC-MS to quantify the amounts of damage (Table 1). In the presence of O_2 the relative amounts of the oxopurines increases, while in the absence of O_2, the

Table 1.[57] Formation of DNA damage in chromatin by γ-radiolysis. Yields are reported in units of nmol•J^{-1}.

Product	Argon	Air	N_2O	N_2O/O_2
Fapy•dA	0.82 ± 0.04	0.60 ± 0.04	0.96 ± 0.04	1.02 ± 0.08
8-OxodA	0.45 ± 0.02	1.10 ± 0.08	0.80 ± 0.07	3.50 ± 0.30
Fapy•dG	0.96 ± 0.03	1.11 ± 0.05	1.81 ± 0.18	1.81 ± 0.18
8-OxodG	0.55 ± 0.03	3.85 ± 0.20	1.35 ± 0.12	8.05 ± 0.65

formamidopyrimidines predominated. With the addition of N_2O, the relative abundance of all the purine lesions increased, presumably due to the increased formation of HO$^{\bullet}$. This further substantiates that DNA damage arises through the action of the HO$^{\bullet}$ with purines. Other product distribution studies using different radical sources have been

performed and efforts were made to establish that the $HO^{\bullet}$ is responsible for generating the lesions quantified.[28,58]

These studies used GC-MS to quantify the distribution of damage products. Inconsistencies observed from the hydrolysis/derivatization and artifacts associated with the GC-MS quantification technique necessitated determination of product distributions under other conditions (Section 2.1.2). Using HPLC-MS/MS analyses the ratio of Fapy•dA to 8-OxodA stayed the same. However, they were 7-9 times less abundant than Fapy•dG and 8-OxodG (Table 2).[42] Furthermore, the amounts of Fapy•dG and 8-OxodG were found to be approximately the same using the HPLC-MS/MS technique. This is different than the GC-MS result where irradiation in an aerated solution gave a distribution of approximately 4:1 8-OxodG to Fapy•dG (Table 1) Presumably, the difference in observed product distributions is based upon the different conditions used to produce the lesions. The large difference in the distribution between damage to dA and dG may also have been due to an artifact in the quantification by GC-MS, or it could be

Table 2.[42] Distribution of Fapy lesions and oxopurines in calf thymus DNA upon γ-irradiation (yields are per 10^6 normal bases).

Product	γ-rays (yield/Gy)
Fapy•dA	5.0 ± 0.1
8-OxodA	3.9 ± 0.5
Fapy•dG	33 ± 9.4
8-OxodG	35.1 ± 0.9

Table 3.[25] Distribution of Fapy lesions and oxopurines in human leukemia cell line upon γ-irradiation (yields are per 10^6 normal bases).

Product	γ-rays (yield/Gy)
Fapy•dA	Not Detected
8-OxodA	Not Detected
Fapy•dG	0.027 ± 0.001
8-OxodG	0.011 ± 0.001

associated with the difference in DNA sources. The GC-MS studies were performed on chromatin DNA while the HPLC-MS/MS experiment was performed on calf thymus DNA. In chromatin, DNA is bound to histones that can interfere in the damaging events and potentially change the distributions of lesions.[57]

In vivo studies on product distributions of Fapy lesions and oxopurines also support the proposed partitioning of products (Figure 3).[25,42] Product distributions were analyzed in a human leukemia cell line that was exposed to γ-radiolysis. Because cancer cells are often hypoxic the amount of Fapy•dG relative to 8-OxodG was expected to be elevated (Table 3). This was the case as Fapy•dG was formed in a yield that was around 3 times greater than 8-OxodG. Another interesting observation in this experiment is that 8- OxodA or Fapy•dA were not detected. Cadet rationalizes that since the formation of Fapy•dA is generally 1 order of magnitude lower than Fapy•dG, Fapy•dA will not be detected (Table 2). However, it is also possible that 8-OxodA and Fapy•dA are efficiently repaired within the cell thereby making detection difficult. In a similar experiment using the same cell line, the occurrence of 8-OxodA and Fapy•dA appear to be within the experimental error of the basal cell level of these damages.[42]

2.1.4 Formation of Fapy Lesions Through Alkylation and Alkaline Treatment:

Formation of Fapy lesions can also occur through alkylation of the *N*7-position of deoxyguanosine, followed by alkaline treatment to open the imidazole ring. Due to the enhanced nucleophilicity of the *N*7 position of deoxyguanosine, alkylation is readily achievable by electrophilic species. Several biologically active alkylating agents have been shown to form Fapy adducts upon base treatment. Adducts with mitomycin,

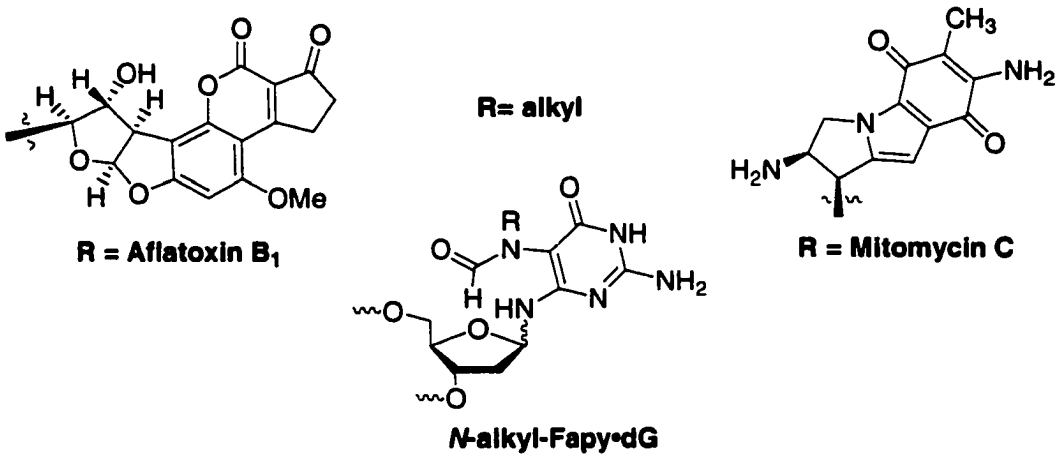


Figure 6. Adducts of Fapy•dG.

aflatoxin, *N*-methylnitrosourea, sulfur mustard, and dibromoethane have been observed.[59-64] The biological significance of the Fapy lesions derived from these adducts is questionable since a pH >9 is required to fracture the imidazole ring. Furthermore, the necessary alkylation events may occur in duplex DNA, but they are most likely minor products. Most biologically active DNA alkylating agents prefer to bind in the minor grove of DNA where the *N*7-position is inaccessible.

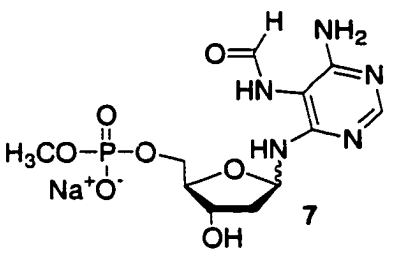
2.2 Structural Features and Stability of Fapy Lesions:

The stability and structural characteristics of Fapy lesions are very important for the investigation of their effects on the structure and function of DNA. Depending upon their longevity in DNA, implications about their biological significance can be made in regards to their effects on polymerase activity. If lesions are unstable, it is likely that they will deglycosylate readily, and thus their mutagenicity and repair will be similar to that of an abasic site. Furthermore, unique structural features that distinguish Fapy

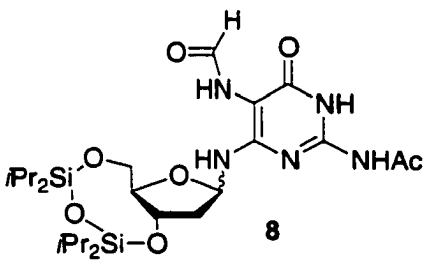
lesions from other lesions may play a role in their recognition by polymerases and DNA repair proteins.

2.2.1 Deglycosylation of Fapy Lesions:

The stability of the glycosidic linkage of the Fapy lesion was largely unknown until a method for the site-specific incorporation of Fapy lesions into DNA was developed.[65-67] Measurement of the monomer stability for Fapy•dA was accomplished



7



8

by the use of a nucleotide derivative of Fapy•dA (**7**).[68] The spontaneous formation of 4,6-diamino-5-formamidopyrimidine from the deglycosylation of **7** was measured by HPLC. The half-life at 37 °C was measured to be ~4 days, 20.5 h at 55 °C and < 6 h at 90°C. Studies on a substituted Fapy•dG monomer, **8**, demonstrated that the glycosidic bond is resistant to deglycosylation with slow decomposition observed after 24 h at 100 °C.[69] The half-life at 50 °C was measured to be 37.8 h for β-**8** and 65.2 h for α-**8**.[69] These observations are consistent with deglycosylation rates in oligonucleotides.[68] The half-life of Fapy•dA was measured to be 21.9 h at 55 °C. The stability of Fapy•dG was observed to be considerably more stable, with a $t_{½}$ > 3 weeks at 55 °C. The remarkable increase in stability of Fapy•dG can be rationalized by resonance stabilization (Figure 7).

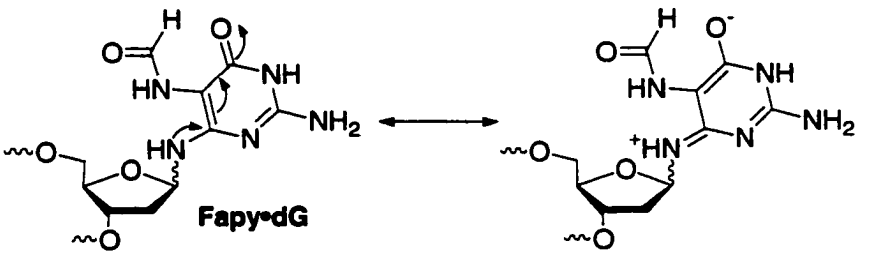


Figure 7. Resonance stabilization of Fapy•dG.

Fapy•dA cannot participate in resonance stabilization thereby increasing its rate of deglycosylation relative to Fapy•dG. These studies suggest that Fapy lesions are long-lived in DNA and should be biologically relevant.

2.2.2 Fapy Lesions Exist as a Mixture of Anomers:

Investigations of the formation of Fapy•dA from γ-radiolysis of deoxyadenosine reveal an interesting feature of formamidopyrimidines.[70] Upon γ-irradiation of deoxyadenosine in oxygen free conditions a mixture of α and β pyranosyl Fapy•dA was observed (Figure 8). Rearrangement of the furanose deoxyadenosine was explained through a ring-opened intermediate. The ring-opened imine recyclizes to the thermodynamically more stable pyranose configuration through attack by the C5' hydroxyl. The observation of rearrangement to pyranose derivatives was also detected in alkylation studies on deoxyguanosine.[59] The electron-rich nature of the pyrimidine ring is believed to play a role in this rearrangement. This hypothesis is supported by the observation that replacement of the formamide by a nitro group led to no rearrangement.[66] This was rationalized by the delocalization of the electron density of the glycosidic nitrogen into the pyrimidine ring thereby preventing the ring opening from

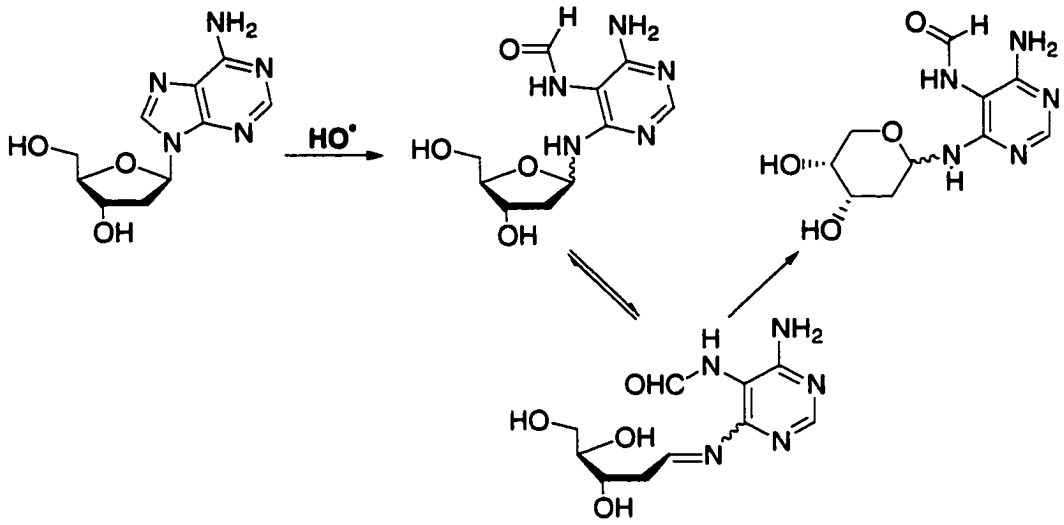


Figure 8. Rearrangement of Fapy•dA through ring-opened intermediate.

occurring. When the formamide is present the electron lone pair on *N*6 cannot delocalize into the electron rich pyrimidine and ring opening is observed.[66]

The rearrangement to the pyranose configuration is not relevant to the consequences of Fapy lesions in DNA, as the C5' hydroxyl is protected as a phosphodiester linkage to the next nucleotide. However, the ring-opening process is biologically relevant. The epimerization of Fapy•dA has been measured by NMR studies on **7** and demonstrate that the ring-opening and subsequent epimerization is a facile process (Figure 9).[68] The epimerization of **7** reaches equilibrium within 7 h to produce a

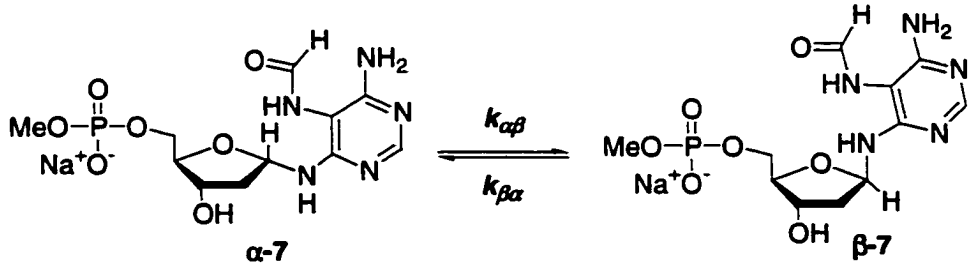


Figure 9. Equilibration of **α-7** to a mixture of anomers.

mixture of anomers. The equilibrium constant for the epimerization process is 1.33 ($k_{\alpha\beta}$/ $k_{\beta\alpha}$). Analogous studies have been performed on **8**, and demonstrate that Fapy•dG readily epimerizes as well within 6.5 h .[69]

Studies at the monomeric level provide limited insight into the real behavior of these lesions in duplex DNA. Other important interactions such as hydrogen bonding abilities are certain to play a role in the rate of epimerization and equilibrium mixtures of anomers. Only through NMR studies on duplex DNA can information pertaining to the behavior of Fapy lesions in DNA be obtained.

2.2.3 Fapy Lesions Exist as a Mixture of Formamide Rotamers:

Formamidopyrimidine lesions exist as a dynamic mixture of rotamers in solution. This can directly influence what type of hydrogen bonding pattern is presented to polymerases, and repair enzymes and affects the potential mutagenicity of Fapy lesions. Theoretical studies have been performed on Fapy•dA and Fapy•dG to determine the various rotamers that could exist.[71,72] In ^{1}H NMR studies on pyranosyl Fapy•dA two major rotamers are observed (Figure 10).[70] The assignments of *cis* and *trans* formamides is based upon the coupling constants observed for the splitting of the NH-5 by the formyl C-H. The relative ratio of 5:1 for the *cis* to *trans* rotamer was observed. The difference in energy between the two conformations was calculated to be 1.0 kcal/mol at room temperature favoring the *cis* rotamer. Similar findings for Fapy•dG were also observed using **8** (Figure 10).[69] A 2:1 mixture of *cis* versus *trans* rotamers is observed for **8** in

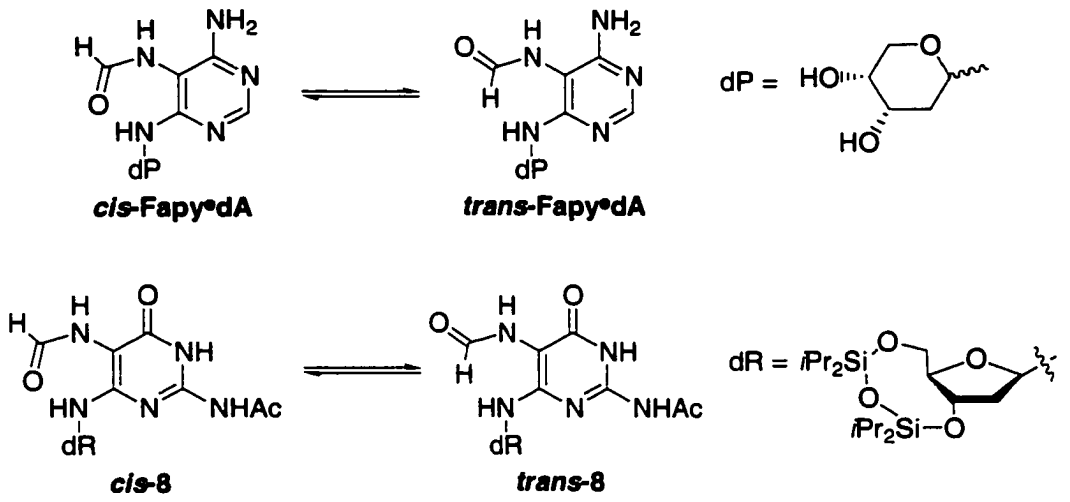


Figure 10. Major rotamers observed for Fapy•dA and Fapy•dG.

DMSO-*d*6.[69] Carell rationalizes that the abundance of the *cis*-amide is derived from a stabilizing hydrogen bond between the carbonyl of the formamide and the glycosidic nitrogen. The formation of a 7-membered hydrogen bonded ring shifts the equilibrium to the *cis* rotamer.[69] By changing NMR solvents to $CDCl_3$ the observation was more pronounced as the *trans* amide could not be detected. Carell predicts that the *cis*-amide is the biologically important conformation as the hydrophobic nature of biomolecules should be similar to the NMR analysis in apolar solvents.[69] However, due to the restrictive nature of the disiloxane-protecting group on **8**, the sugar could be locked into a conformation that is not biologically relevant and this may affect hydrogen bonding in the formamidopyrimidines.

2.3 Determining the Mutagenicity of DNA Lesions:

2.3.1 Polymerase Fidelity :

The ability of DNA lesions to cause DNA polymerases to miscode is an important area of research and has direct bearing on the genotoxicity of DNA lesions on cells. The

extensive size of the genome necessitates that DNA replication processes be highly efficient and accurate. The frequency of mistakes made by DNA polymerases and accessory proteins is remarkably low. Errors from DNA replication have been measured to be less than 1 mistake for every 10^9 to 10^{10} nucleotides added.[73] The high fidelity associated with DNA replication is a combination of many factors. The polymerase enzyme contributes significantly to the high fidelity of DNA replication making an error once every 10^5 to 10^6 nucleotides added.[2] Post-replicative mismatch corrections and other proteins associated with DNA replication account for the remaining accuracy in replication. Even though polymerases rarely make a mistake when replicating undamaged DNA, modifications can have a profound effect on this process.

The ability of DNA polymerases to incorporate nucleotides with high fidelity is believed to be due to two factors. The first is the ability of polymerases to specifically incorporate nucleotides opposite the proper base. Secondly, the presence of proofreading abilities in polymerases allows for recognition of any mistakes made during replication.[2] The mechanisms by which polymerases insert nucleotides are governed by the ability of the incoming nucleotide to bind to the active site. The active site imposes several restrictions on the selection of the nucleotide to be inserted.[22] Factors involved can include Watson-Crick (W-C) hydrogen bonding, base-stacking interactions, electrostatic interactions, as well as steric considerations. W-C hydrogen bonding has long been held as the distinguishing factor that leads to fidelity. However, recent observations indicate that the W-C hydrogen bonding is not as important in polymerase fidelity as are the W-C geometric requirements in the active site of the polymerase-ternary complex.[74,75]

The second feature that leads to high fidelity of polymerases is their ability to recognize mistakes.[2] It is hypothesized that polymerases check the geometry of the formed base pair through hydrogen bonding interactions, and through steric interactions in the minor groove of DNA leading to the repair of the lesion or misinsertion event.[21,76] The interaction of the polymerase in the minor groove of DNA is believed to serve as a proofreading mechanism for prior misinsertions that were allowed. The molecular recognition in the minor groove by polymerases is at the *N*3 of purines and the *O*2 of pyrimidines.[76]

DNA synthesis depends on the ability of the polymerase to bind to the primer/template and incoming nucleotide triphosphate (ternary complex) in a favorable geometry. If the lesion has a poor ability to interact with the incoming nucleotide, or presents significant structural distortions to the DNA molecule it can serve as a block to synthesis.[22] However, if the lesion has only moderate structural effects on the ternary structure, translesional synthesis can occur. The structure of the lesion presented to the polymerase is the primary determinant for which nucleotide will be inserted opposite a lesion. The distortion of the DNA structure that is introduced by a lesion can potentially be offset by the flexibility that is present in the catalytic center of the polymerase. Translesional synthesis not only affects the incorporation of nucleotides opposite a lesion in DNA, but it can also affect the insertion of subsequent nucleotides (mismatch extension).[22] The nature of the mismatch between the lesion and the inserted nucleotide dictates if the insertion of subsequent bases will be affected. General studies on full-length primer extension with DNA modifications typically show a preference for polymerase pausing before extending opposite the lesion, and in some cases 5' adjacent

extensions are also retarded for a few bases past the lesion.[19,77] Once the polymerase is beyond the influence of the lesion the chain is fully extended.

2.3.2 Measurement of Polymerase Fidelity:

The measurement of polymerase fidelity can be accomplished in assays that have been developed for both *in vitro* and *in vivo* analysis. The polyacrylamide gel fidelity assay developed by Goodman is used to measure relative rates of primer extension *in vitro*.[78,79] This can be used to determine the relative rates of insertion of individual nucleotide triphosphates (dNTPs) opposite a specific lesion. By employing a template strand that contains the lesion of interest and a radiolabeled primer, the rates of extension past the lesion can be determined by separating the products by denaturing polyacrylamide gel electrophoresis (PAGE) and subsequent quantification (Figure 11).

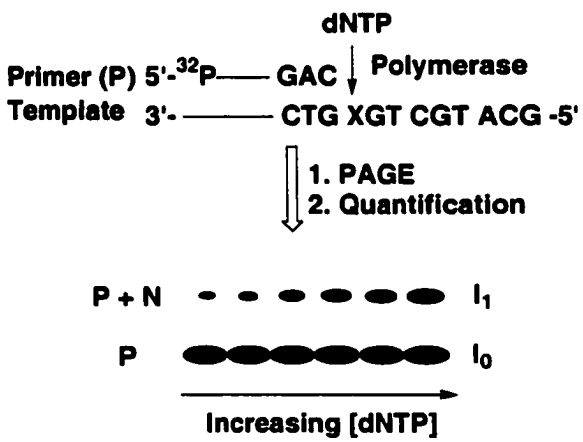


Figure 11. Schematic representation of standing start kinetic experiment.

The standing start method is commonly used to measure the kinetics of insertion opposite a lesion of interest and is based upon the kinetic assumption that single-hit conditions exist. In order for single hit conditions to predominate, less than 20% of the

primers are extended, and the amount of primer/template present is in large excess over the polymerase. This maximizes the probability that a DNA molecule interacts with only one enzyme molecule. The standing start method requires that the primer alignment is up to the lesion so that the first insertion event is directly across from the lesion. The selection of dNTP concentrations is chosen so as to preserve single hit conditions. Time and the concentration of polymerase are additional variables that can be changed to achieve the <20% extension requirement.

Quantification of the PAGE separated bands permits the calculation of the reaction velocity (v). Typically, two bands are observed for the extensions that are unreacted primer (I_0) and the extended primer (I_1). In the event that extension beyond the lesion is present, the additional bands are added into the intensity of I_1. The calculation of the velocities for the standing start reaction is obtained using Equation 2.9.[80] The

$$v = \frac{100\,(I_1)}{(I_0 + 0.5I_1)\,t} \qquad \textbf{Equation 2.9}$$

velocity is calculated in terms of percent per unit time where I_1 is measured as a function of time in minutes. The production of I_1 arose at the expense of I_0, thus the approximate intensity of the unextended primer band is the average for the amount of primer present initially and at time t (Equation 2.10).[80] The velocity for the extension is a function of

$$[(I_0 + I_1) + I_0] / 2 = (I_0 + 0.5I_1) \qquad \textbf{Equation 2.10}$$

the dNTP concentration used. Graphical analysis of velocity versus the dNTP concentrations through a Hanes-Woolf plot gives a linear relationship from which the kinetic parameters K_m and V_{max} are obtained. The slope of the line is inversely proportional to V_{max} and the y-intercept represents the K_m/V_{max} for the insertion event.

The fidelity of the enzyme is then obtained through a comparison of the ratio of the V_{max}/K_m for the insertion of the wrong nucleotide opposite the lesion, X, to the correct nucleotide (Equation 2.11). This ratio is termed the misinsertion frequency (f_{ins}). The

$$f_{ins} = \frac{(V_{max}/K_m)_{wrong}}{(V_{max}/K_m)_{correct}} \qquad \textbf{Equation 2.11}$$

value for the misinsertion frequency can be used as a point of reference in the comparison of the relative misincorporation of nucleotides opposite native nucleotides. Caution must be used when comparing misinsertion frequencies obtained from different polymerase enzymes, because the enzymes form different ternary complexes with the dNTP and primer-template duplex, resulting in different preferences for misinsertion. Furthermore, the rates measured are a combination of several steps that occur during the extension process. The observed rate represents the slowest of the steps that include the association of the polymerase with the primer/template duplex, binding of the dNTP, insertion of the dNTP, and the dissociation of the polymerase from the complex. Different polymerases and/or lesions are likely to have different microscopic rates for these steps and direct comparisons of insertion efficiencies between enzymes and lesions may not be accurate.[22]

In vivo analysis of the fidelity of DNA polymerases on DNA modifications can be used to obtain the mutagenic and cytotoxic properties of lesions directly in cells. This assay uses recombinant DNA methods to incorporate synthetic oligonucleotides containing lesions at defined sites into a single stranded vector.[81] The sequence near the lesion is engineered to possess restriction sites that facilitate analysis of the insertion event opposite the lesion. The vectors are transfected into cell lines and allowed to replicate. The replicated plasmids are isolated and a fragment containing the lesion is

excised. Different techniques are then employed to quantify the mutational frequency. Typically, the transformants are isolated and sequenced to determine the mutagenicity.[2] More recent techniques involve the use of molecular probes that are specific for the insertion events opposite the lesion. The hybridization efficiency of a molecular probe that is specific for the proper insertion or mismatch is correlated with the mutagenicity of the lesion.[82]

2.3.3 Thermodynamic Denaturation Studies:

Another technique used to support the mutagenicity findings above is thermodynamic denaturation studies. This technique allows one to measure the thermodynamic parameters associated with the ability of lesions to base pair and to distinguish between the relative stabilities of mismatches opposite lesions. Thermodynamic denaturation studies are based upon measuring the change in hyperchromicity of the DNA molecule as it is denatured. Double-stranded DNA has a lower extinction coefficient at 260 nm than does single-stranded DNA. The increase in absorbance at 260 nm between the transition of dsDNA to ssDNA is typically 20-30%.[83] The denaturation of the duplex is entropically favored and by heating the duplex, the increase in absorbance can be monitored to determine the melting temperature (T_m) for the duplex. The T_m for the duplex is the temperature at which 50% of the dsDNA is melted, and is obtained through analysis of a plot of absorbance versus temperature. The T_m for a duplex is dependent upon the concentration of DNA.[84] Thermodynamic parameters can be extracted from this melting process when performed at various concentrations of duplex. Through a van't Hoff plot of the inverse of the observed T_m

versus the natural log of the total concentration of DNA, a linear relationship is obtained. The slope of the line is inversely proportional to the enthalpy of melting (ΔH), and the entropy is extracted from the y-intercept (Figure 12).[84,85] The apparent free energy

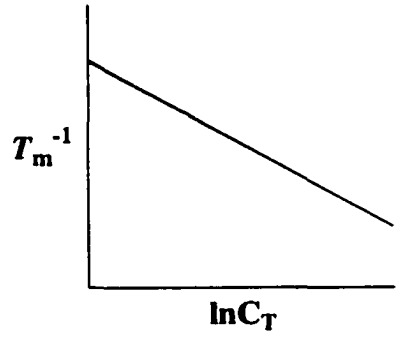


Figure 12. Determination of thermodynamic paramaters associated with thermal denaturation of dsDNA from a van't Hoff plot.

($\Delta G°$) for this process is calculated using the Gibbs free energy equation at a specified temperature (Equation 2.12). Thermodynamic values obtained for the mismatch sequences are compared to the properly paired sequence to obtain insight into the relative destabilization or stabilization of mismatches opposite a lesion.

$$\Delta G° = \Delta H - T\Delta S \qquad \textbf{Equation 2.12}$$

Upon analysis of a particular lesion in dsDNA general information about the ability to form stabilizing hydrogen bonds with other nucleotides can be obtained. This serves as a guide in interpreting results of gel fidelity assays for mutagenicity. However, thermodynamic parameters obtained from thermal denaturation studies should be used with caution when comparing to insertion events opposite a lesion. The nature of the two measured events are fundamentally different. In the insertion event, the polymerase serves as the stabilizing influence to allow the incoming nucleotide to align. In the duplex, thermodynamic melting studies are typically performed on duplexes wherein the

lesion is centrally located. Direct comparisons between the thermodynamic stability of the insertion process by thermodynamic denaturation studies and through gel fidelity assays have been made.[80] These studies demonstrate that the techniques do not always give comparable results and generally differ by an order of magnitude.

2.4 Mutagenic Properties of Purine Lesions:

Using the techniques mentioned in Section 2.3, the mutagenicity of many DNA lesions has been determined.[22] A significant finding of this work is that 8-OxodG is a highly mutagenic lesion whereas 8-OxodA is significantly less mutagenic.[17,22] The mutagenic properties of Fapy lesions are largely unknown.

2.4.1 Mutagenic Properties of 8-Oxopurines:

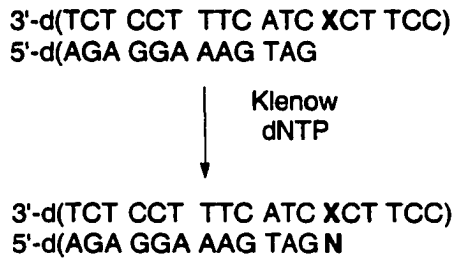
The mutagenicity of 8-oxopurines has been extensively studied both *in vitro* and *in vivo*. The site-specific incorporation of 8-OxodA and 8-OxodG into defined sequences of oligonucleotides by automated DNA synthesis has facilitated the investigations into the structural and functional effects that these lesions have on DNA.

2.4.1.1 The Effects of 8-OxodG on DNA Replication:

The mutagenic properties of 8-OxodG were first determined using the polyacrylamide gel fidelity assay. These experiments were the first to demonstrate that 8-OxodG produces dG→T transversions.[86] This result arises from the ability of polymerases to insert dATP opposite 8-OxodG. Upon replication of this newly synthesized strand of DNA a polymerase will insert a T opposite the misinserted dA

forming the mutagenic dG→T transversion. Transversion mutations involve substitution of a purine for a pyrimidine or pyrimidine for purine.[2] The misinsertion frequency of this event is 0.05 using the Klenow fragment polymerase (Table 4). This suggests that

Table 4.[86] Kinetic parameters of insertion events opposite 8-OxodG by Klenow fragement.



X	dNTP	K_m (μM)	V_{max} (% min^{-1})	F_{ins}[a]
dG	dC	2.04	14.5	1.0
8-OxodG	dC	6.62	16.3	0.35
8-OxodG	dA	28.3	9.5	0.05

[a] F_{ins} = (V_{max}/K_m for X= dG or 8-OxodG, N = dC or dA)/ (V_{max}/K_m for X = dG, N =dC).

whenever a polymerase comes to an 8-OxodG in the template strand dATP will be misincorporated 14% of the time. This may seem like an infrequent event, but when considered in context with the size of the genome that is replicated, the probability of the

Table 5.[87] Kinetic parameters of nucleotide insertion opposite 8-OxodG by Klenow exo$^-$.

X	dNTP	K_m (μM)	V_{max} (% min^{-1})	F_{ins}[a]	F_{rel}[b]
dG	C	1.1	0.4	1.0	--
8-OxodG	C	28.0	0.3	3.5×10^{-2}	1.0
8-OxodG	A	1.7	5×10^{-3}	8.7×10^{-3}	0.25

[a] F_{ins}= (V_{max}/K_m, X=G, or 8-OxodG; dNTP = C or A) /(V_{max}/K_m, X=G; dNTP = C).
[b] F_{rel}= (V_{max}/K_m, X= 8-OxodG; dNTP = C or A) /(V_{max}/K_m, X= 8-OxodG; dNTP = C).

misinsertion occurring is probable. Comparable values were obtained for insertion of dATP opposite 8-OxodG using *E. coli* Klenow (exo$^-$) (Table 5).[87]

Mutagenesis studies *in vivo* support the above findings.[81,82,88,89] Transformation of shuttle vectors containing 8-OxodG were incorporated into simian kidney cells (COS-7) and allowed to replicate.[82] The mutational frequency measured for this experiment was 5.2-6.0 % for dG→T transversions depending on the flanking sequence. A minor amount (0-0.8%) of dG→dA transitions were also detected.

The ability of the polymerase to insert dA opposite 8-OxodG is rationalized by the ability of 8-OxodG to adopt the *syn* conformation. This presents a hydrogen bonding pattern that is "T like" allowing dA to be misincorporated (Figure 13). The ability of 8-

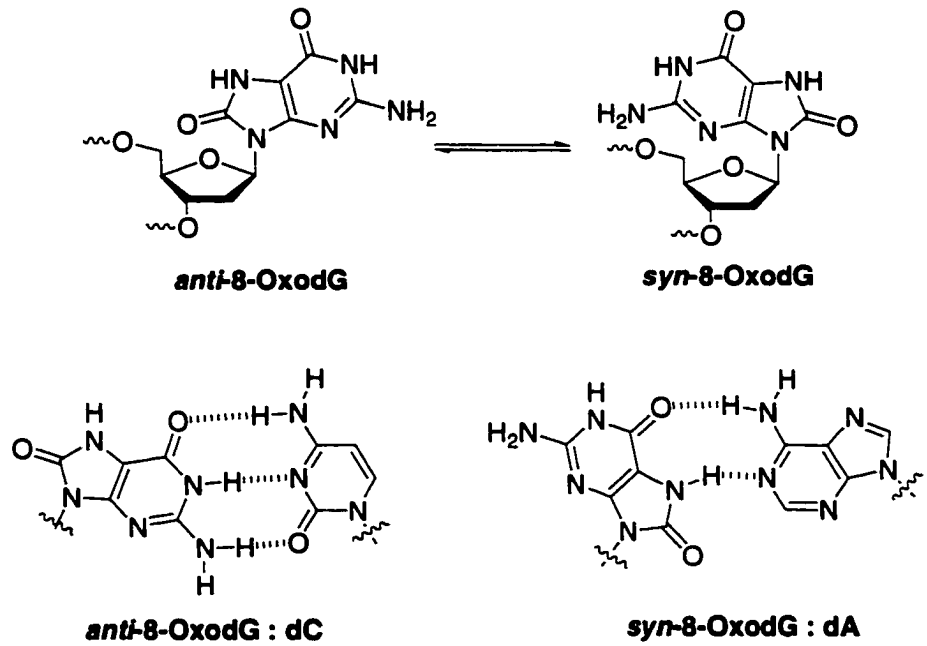


Figure 13. *Syn* vs. *anti* equilibrium of 8-OxodG and potential hydrogen-bonding patterns.

OxodG to form hydrogen bonds with deoxyadenosine has been demonstrated by X-ray crystallography and thermodynamic melting studies.[90,91] Thermodynamic melting studies show that the 8-OxodG:dA base pair is stabilized by 3.4 kcal/mol relative to the native dG: dA base pair (Table 6).[91] In comparison, the 8-OxodG: dC base pair is

Table 6.[91] Thermodynamic stability of 8-OxodG containing duplexes.

5'-d(GCG TAC **X**CA TGC G) X= G, 8-OxodG
3'-d(CGC ATG **Y**GT ACG C) Y= C, A

Duplex (X:Y)	T_m[a] (°C)	ΔG (kcal/mol)[a]	ΔΔG(kcal/mol)[a,b]
G:C	76.0 ± 0.2	19.3 ± 0.3	--
8-OxodG:C	73.5 ± 0.2	17.3 ± 0.4	-2.0 ± 0.7
G:A	64.8. ± 0.2	12.5 ± 0.4	--
8-OxodG:A	70.7 ± 0.2	15.9 ± 0.3	3.4 ± 0.7

[a] Calculated at 25 °C
[b] ΔΔG = ΔG (8-OxodG:Y)– ΔG (G:Y)

destabilized by 2.0 kcal/mol relative to the dG:dC base pair. Overall, the 8-OxodG:dC base pair is more stable than the 8-OxodG:A base pair by 1.4 kcal/mol. This is a small energy difference in dsDNA, suggesting that both base pairs can be accommodated in DNA.

2.4.1.2 The Effects of 8-OxodA on DNA Replication:

The mutagenicity of 8-OxodA is quite different from that observed for 8-OxodG. Analysis by the gel-fidelity assay shows that 8-OxodA is far less mutagenic than 8-OxodG.[92] Insertion events opposite 8-OxodA produce a low level of A→T transversions

as well as A→C transversions (Table 7). The misinsertion frequencies for these events are 5 x 10^{-4} and 6 x 10^{-4} respectively. Therefore, a mutagenic event will happen approximately once every 2000 times that the polymerase encounters 8-OxodA. This is 2 orders of magnitude lower than 8-OxodG. 8-OxodA is not considered a mutagenic lesion.

Table 7.[92] Kinetic parameters of insertion events opposite 8-OxodA by Klenow fragment.

3'-d(TCT CCT TTC ATC **X**CT TCC)
5'-d(AGA GGA AAG TAG

↓ Klenow, dNTP

3'-d(TCT CCT TTC ATC **X**CT TCC)
5'-d(AGA GGA AAG TAG **N**

X	dNTP	K_m (μM)	V_{max} (% min^{-1})	F_{ins}[a]
A	T	3.66	18.9	1.0
A	A	51.7	0.84	3.14 x 10^{-3}
A	G	15.3	0.07	8.76 x 10^{-4}
8-OxodA	T	7.08	10.4	0.28
8-OxodA	A	60.0	0.15	5.09 x10^{-4}
8-OxodA	G	26.1	0.06	5.50 x10^{-4}

[a] F_{ins} = (V_{max}/K_m for X= A or 8-OxodA, N = T, A, G, or C) /(V_{max}/K_m for X = A, N =T).

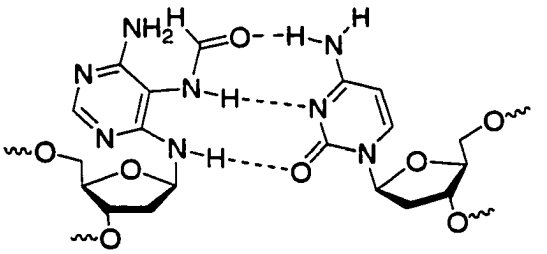
In vivo analysis of 8-OxodA was performed along side the experiments for 8-OxodG in COS-7 cells.[82] The mutational frequency for 8-OxodA was found to be 4-28 times less than that of 8-OxodG. The mutations that were observed were A→C

The mutagenic properties of Fapy lesions are largely unknown. This is due to the previous inability to selectively generate these lesions in DNA. Only recently have phosphoramidites of Fapy•dA and Fapy•dG been chemically synthesized and used to incorporate the lesions at defined sites in oligonucleotides.[65-67] This is greatly facilitating the discovery of the mutagenic properties of these compounds in DNA.

Previously, information regarding the mutagenicity of Fapy lesions was obtained through random generation of lesions in bulk DNA through oxidative stress mechanisms (γ-radiolysis and Fe-EDTA), or through alkylation of the *N*7 position followed by alkaline treatment.[95-99] The latter method generates the alkylated Fapy lesions discussed in Section 2.1.4. Recently, MeFapy•dG has been incorporated at defined sites in an oligonucleotide by the use of *N*7-methyl deoxyguanosine triphosphate, followed by alkaline treatment.[100-102] MeFapy•dG generated randomly blocks DNA replication *in vitro* and *in vivo*.[95-97] The use of oligonucleotides with MeFapy•dG incorporated site specifically showed that replication was inhibited, but not entirely. dG→T and dG→dC transversions were inferred. However, the frequency of mutations were small compared to the large number of MeFapy•dG lesions that were produced in DNA.[99] Further *in vitro* analysis with MeFapy•dG at defined sites demonstrates that it is not a mutagenic substrate.[102]

Inhibition of DNA replication *in vivo* was not observed when Fapy•dA and MeFapy•dA were randomly generated in DNA. Experiments suggest that Fapy•dA and MeFapy•dA are mutagenic lesions that can produce dA→dG transitions through the misinsertion of dCTP. The rationale proposed for this observation is that Fapy•dA can adopt the *syn* conformation presenting a dG-like hydrogen-bonding pattern (Figure

15).[98,99] However, the proposed hydrogen bonding pattern cannot account for the results with MeFapy•dA base pair.



syn-Fapy•dA : dC

Figure 15. Possible hydrogen bonding pattern for the Fapy•dA:dC base pair.

Another factor that has not been considered in indirect studies on the mutagenicity of Fapy lesions is that they exist as a mixture of anomers in DNA (See Section 2.2.2). The mutagenesis studies discussed above assumed that the β-anomer of Fapy lesions was involved. The presence of the α-anomer could also have profound effects on the activity of the DNA polymerase. If Fapy lesions are mutagenic or even genotoxic, it may be due to the contribution of one or both anomers.

2.4.2.1 Fapy•dG Induces Klenow (exo⁻) to Misinsert dATP:

Recent results in our lab have demonstrated that Fapy•dG at defined sites within an oligonucleotide induces Klenow (exo⁻) to produce predominantly G→T and to a lesser extent G→C transversions (Table 8).[103] The kinetics of insertion opposite Fapy•dG reveal that the misinsertion of dATP occurs 5% of the time that Klenow (exo⁻) comes in contact with Fapy•dG. This is in contrast to prior indirect experiments that suggest that Fapy•dG is a block of DNA synthesis.[96-97] Comparisons of the misinsertion preferences of

Table 8.[103] Kinetic parameters of nucleotide insertion opposite Fapy•dG.

3'-d(TGG TAC CCT GCA CGA CAX TGA CGT GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT GCT GT

↓ Klenow exo⁻
dNTP

3'-d(TGG TAC CCT GCA CGA CAX TGA CGT GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT GCT GTN

X	dNTP	K_m (μM)	V_{max} (% min^{-1})	F_{ins}[a]	F_{rel}[b]
dG	C	5.2×10^{-3}	5.1	1.0	--
Fapy•dG	C	0.4	7.9	1.9×10^{-2}	1.0
Fapy•dG	A	10.4	9.9	9.7×10^{-4}	4.8×10^{-2}
Fapy•dG	G	26.2	1.0	4.1×10^{-5}	1.9×10^{-3}
Fapy•dG	T	103.0	0.3	3.0×10^{-6}	1.5×10^{-4}

[a] F_{ins}= (V_{max}/K_m, X=G, or Fapy•dG; dNTP = T, A, G, or C) /(V_{max}/K_m, X=G; dNTP = C).
[b] F_{rel}= (V_{max}/K_m, X= Fapy•dG; dNTP = T, A, G, or C) /(V_{max}/K_m, X= Fapy•dG; dNTP = C).

Fapy•dG and 8-OxodG by Klenow (exo⁻) (Table 5) demonstrate that both of these lesions are potentially mutagenic. However, 8-OxodG induces Klenow (exo⁻) to misinsert dATP approximately ten-times more frequently than Fapy•dG. Comparison of the F_{rel} demonstrates that the Klenow (exo⁻) is five times as likely to misinsert dATP opposite 8-OxodG than Fapy•dG (25% of the time versus 5%, Table 5 and Table 8).

2.4.2.2 Fapy•dA Induces Klenow (exo⁻) to Misinsert dATP:

Recent reports on the mutagenicity of Fapy•dA at defined sites within an oligonucleotide suggest that Fapy•dA induces Klenow (exo⁻) to misincorporate dATP (Table 9).[104] Fapy•dA also causes Klenow (exo⁻) to misinsert dGTP, but is 2 times less likely to occur than the misinsertion of dATP. The observation of the misinsertion of A is contrary to previous indirect experiments on Fapy•dA that suggest that C is the

Table 9.[104] Kinetic parameters of nucleotide insertion opposite Fapy•dA.

3'-d(TAC CCT GCA CGA CAX TCA CGT GCA ACT TGC)
5'-d(ATG GGA CGT GCT GT

Klenow exo⁻
dNTP

3'-d(TAC CCT GCA CGA CAX TCA CGT GCA ACT TGC)
5'-d(ATG GGA CGT GCT GTN

X	dNTP	K_m (μM)	V_{max} (% min^{-1})	F_{ins}[a]	F_{rel}[b]
A	T	0.13	4.9	1.0	--
A	A	147.7	18.7	2.9×10^{-3}	--
A	G	231.1	9.9	9.9×10^{-4}	--
A	C	76.3	8.1	2.4×10^{-3}	--
Fapy•dA	T	0.34	3.8	0.28	1.0
Fapy•dA	A	86.7	14.7	3.9×10^{-3}	1.4×10^{-2}
Fapy•dA	G	16.3	1.4	1.5×10^{-3}	5.4×10^{-3}
Fapy•dA	C	75.2	1.0	4.3×10^{-4}	1.5×10^{-3}

[a] F_{ins}= (V_{max}/K_m, X=A, or Fapy•dA; dNTP = T, A, G, or C) /(V_{max}/K_m, X=A; dNTP = T).
[b] F_{rel}= (V_{max}/K_m, X=Fapy•dA; dNTP = T, A, G, or C) /(V_{max}/K_m, X=Fapy•dA; dNTP = T).

prevalent misincorporation event.[98] The F_{ins} for Fapy•dA suggests that A→T transversions occur ~0.4% of the time. This is a relatively low mutational frequency in comparison to 8-OxodG that induces misincorporation of dATP five percent of the time. However, the misinsertion of dATP is approximately 8 times more likely to occur than the similar insertion opposite 8-OxodA. This suggests that Fapy•dA is more likely to be mutagenic than 8-OxodA.

2.5 DNA Repair Processes Associated with Damaged Purines:

DNA repair processes contribute to the high fidelity associated with DNA replication. Post-replication processing can improve the overall fidelity of DNA

replication by three orders of magnitude.[2] This serves as a safeguard for correcting any mistakes that the DNA polymerases may make. There are several different repair processes present in biological systems. The nature of the DNA damage that is present typically dictates which process will be activated. In the case of DNA modifications that block polymerase synthesis, an SOS response is induced to remove a large portion of the DNA strand that contains the lesion.[2,22,105] This is followed by resynthesis by a DNA polymerase. The mechanism of this response has been the focus of much speculation and controversy. It is known that the induction of the SOS response involves a complicated cascade of events that is triggered by the association of the RecA protein to single-stranded DNA.[22]

For DNA modifications that promote translesional synthesis, other alternatives for repair exist that do not require extensive reconstruction of DNA. Nucleotide-excision repair (NER) generally deals with recognizing DNA modifications that distort the helix and interfere with base-pairing.[9] Base-excision repair (BER) is primarily targeted toward small chemical alterations to nucleobases. The types of lesions that BER enzymes recognize generally promote translesional synthesis and frequently lead to miscoding events.[9] The repair of these lesions are of utmost importance in preventing mutagenesis.

2.5.1 Enzymes Associated with BER Processes:

The BER process involves two key steps in the removal of a DNA modification. The first event involves a glycosylase that is specific for the removal of the particular DNA damage (Figure 16).[16] The product of this reaction is an apurinic-apyrimidic (AP) site. Following this event, an AP lyase removes the abasic site thereby producing a

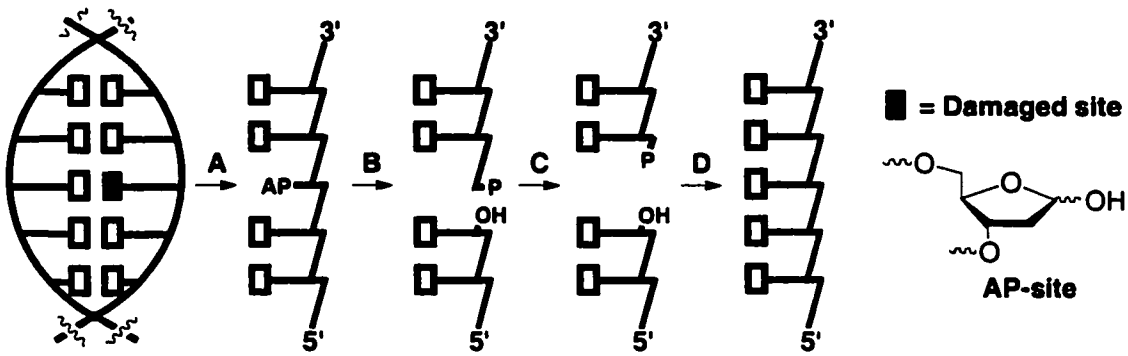


Figure 16. Schematic representation of BER. A. DNA glycosylase B. AP endonuclease. C. Excision of Sugar D. DNA polymerase/ligase.

strand break at the site of the lesion. The strand scission is then repaired by the action of a DNA polymerase and sealing of the backbone by ligase.

The specificity of glycosylases toward DNA modifications are broad and diverse.[16,106,107] Several different glycosylases are known to recognize more than 5 different substrates. In some instances several different glycosylases can recognize the same lesion. The presence of multiple systems for the repair of specific lesions increases the probability that repair will occur. It also suggests that the lesion has mutagenic properties as a variety of systems have evolved to protect the genome against it. Glycosylases are classified into two categories depending upon the presence of AP lyase activity. Monofunctional glycosylases do not possess AP lyase functionality and require the assistance of an AP lyase to facilitate the removal of the AP site. Bifunctional glycosylases possess both glycosylase and AP lyase activity (Figure 17).[16]

2.5.2 The Formamidopyrimidine DNA Glycosylase Protein (Fpg):

A DNA glycosylase that is responsible for the removal of Fapy lesions from DNA is the formamidopyrimidine DNA glycosylase protein (Fpg, MutM). The Fpg protein is a

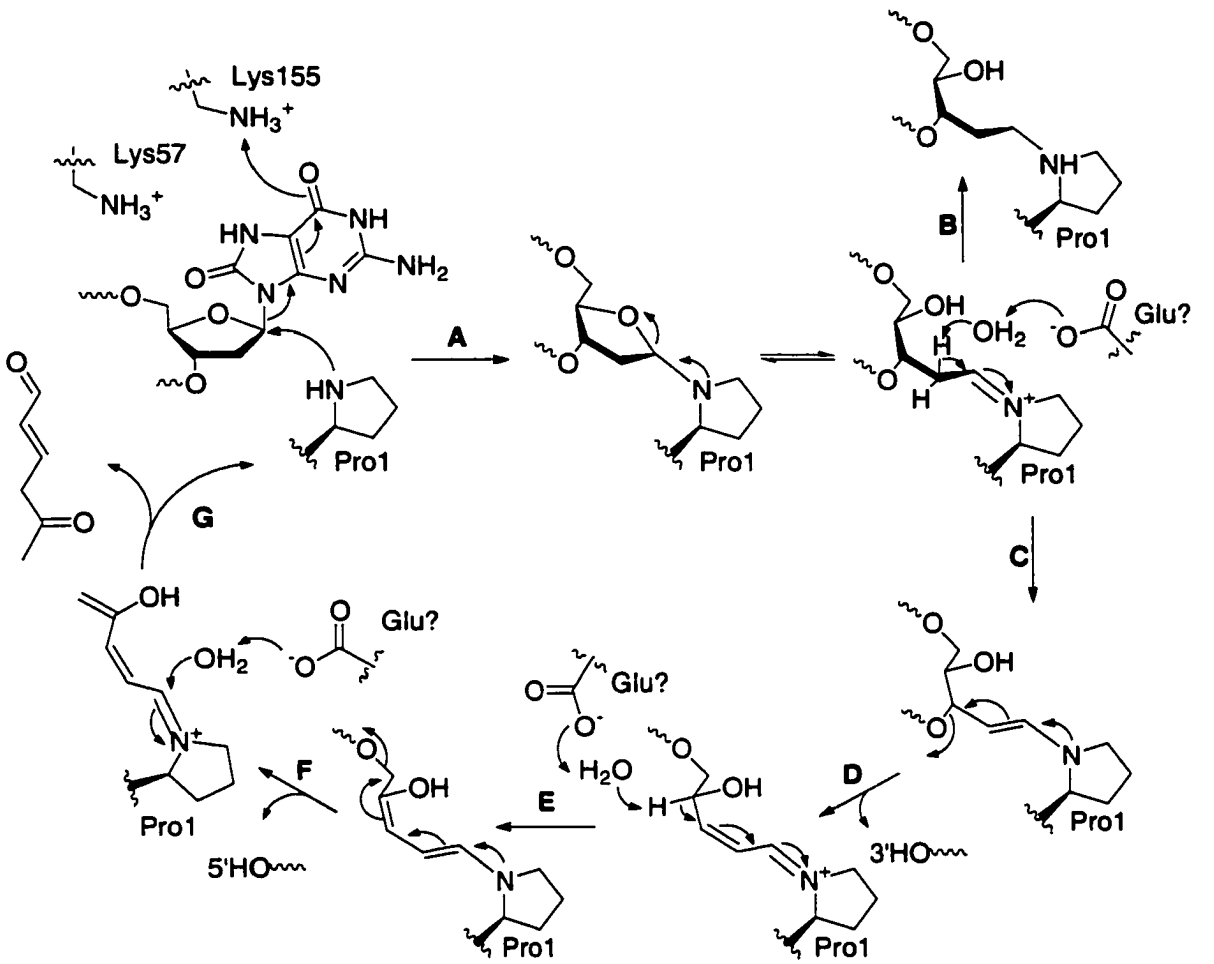


Figure 17. Schematic representation of the bifunctional glycosylase/ AP lyase activity of the Fpg protein. A) Deglycosylation; B) $NaBH_4$ trapping of iminium intermediate; C) Formation of eneamine; D) β-Elimination; E) Regeneration of eneamine; F) δ-Elimination; G) Regeneration of Fpg and release of 4-Oxo-2-pentenal.

bifunctional glycosylase that removes several types of DNA lesions. Overexpression of the Fpg protein in *E. coli* cells has facilitated mechanistic investigations into its function.[108] The general characteristics of the Fpg protein, the active site requirements, and the substrate specificity have been determined.

2.5.2.1 Characteristic Features of Fpg Protein:

The Fpg protein is highly conserved over several different species of bacteria. *E. coli* Fpg is composed of 269 amino acid residues with a molecular weight of 30.2 kDa.[108] A common feature amongst these proteins is that they contain the same *N*-terminal sequence of Pro-Glu-Leu-Pro-Glu-Val. They also include a Zn-finger binding motif at the C-terminus.[109] The Fpg protein has three known catalytic functions, it removes oxidatively damaged bases through glycosylase activity, participates in AP lyase activity by cleaving the 5' and 3' phosphodiester bonds through β- and δ-eliminations, and excises 5'-terminal phosphate groups (dRPase).[110-112] The activity of each of these functions is highly dependent upon the binding of the DNA by the Zn-finger motif.[113] The Fpg protein is essential in bacterial cells for DNA repair. Cell lines that are deficient in Fpg, are subject to higher levels of mutation.[114] This observation is amplified by inactivating the MutY gene (Section 2.5.3.1).[115]

2.5.2.2 The Active Site Requirements of the Fpg Protein:

The active site requirements of the Fpg protein have been extensively studied. The binding of the Fpg protein to dsDNA has been mapped by high resolution hydroxyl radical foot-printing. This technique revealed that the binding site of the Fpg protein spans 6 nucleotides along the strand containing the damaged site, and just one base on the opposite strand.[116]

The active site residues for the glycosylase and AP lyase activities of Fpg protein have been mapped through site directed mutagenesis experiments that were directed at conserved amino acid residues in the Fpg protein [117-120] These amino acid residues play a role in the nucleophilic catalysis that is required for the glycosylase and AP lyase

activities of the Fpg protein. Reductive trapping of the iminium ion followed by proteolysis has determined that the *N*-terminal proline is the amino acid responsible for the glycosylase activity.[121] Substituting for lysine-57 by glycine (K57G) reduced the glycosylase activity 50 fold.[118] Substitution of lysine 155 for alanine (K155A) also diminished the glycosylase activity by approximately 50 fold.[119] It is proposed that these two residues play a synergistic role in the deglycosylation event by interacting with C8-oxo group that is common to many of the substrates that Fpg recognize.[118,119] Presumably, these same residues interact with the formamido-carbonyl of Fapy lesions. Site-directed mutagenesis experiments were also directed at the acidic amino acid residues that are involved in the AP lyase activity. However, no loss in lyase activity is detected when the conserved glutamate and aspartate residues in Fpg are mutated.[120] Thus, the acidic amino acid residues involved in the removal of the AP site remain elusive.

X-ray crystal analysis of the Fpg protein isolated from the bacteria *T. thermophilus* shows that the overall structure consists of two major domains that are attached through a cleft.[109] This structural feature allows for the two domains to wrap around the dsDNA substrate. The amino acid residues responsible for the catalytic activity of the Fpg protein are found in the cleft region. The *N*-terminal proline residue occupies the lower domain. It is proposed that the glycosylase activity of the Fpg protein is analogous to other glycosylases that flip out the modified base from the duplex for further recognition of the cognate base.[122,123] This provides an additional feature that recognizes the estranged base and prevents excision of the damaged base if it is not properly base-paired. This serves as a safeguard from removing the damaged base if a

premutagenic event has taken place such as formation of an 8-OxodG:dA base pair. Modeling studies using the unbound crystal structure for Fpg, and dsDNA with a DNA lesion flipped out of the helix were performed. They demonstrate that amino acid residues in the proximity of the flipped out base can participate in this deglycosylation mechanism.[109]

2.5.2.3 The Substrate Specificity of the Fpg Protein:

The substrate specificity of the Fpg protein has been determined through an assortment of techniques. Using the GC-MS technique for quantifying oxidative damage of DNA (Section 2.1.2) coupled with deglycosylation by Fpg protein, products of excision were determined.[124] γ-Radiolysis treatment of calf thymus DNA randomly generates lesions in DNA that are substrates for the Fpg protein. Analysis of the pelleted DNA for modifications through acid hydrolysis/derivatization and subsequent GC-MS detection reveals the abundance of lesions that are generated in the DNA (Table 10).

Table 10.[124] Determination of the substrate specificity of the Fpg protein through GC-MS analysis of randomly generated lesions in calf thymus DNA by γ-radiolysis.

Product (nmol/mg of DNA)	Treatment[a]				
	1	2	3	4	5
Fapy•dA	0.09 ± 0.01	1.22 ± 0.02	1.13 ± 0.17	0.68 ± 0.10	0.63 ± 0.24
8-OxodA	0.20 ± 0.02	1.12 ± 0.06	1.11 ± 0.08	0.98 ± 0.06	0.17 ± 0.01
Fapy•dG	0.14 ± 0.01	1.76 ± 0.33	1.63 ± 0.24	0.76 ± 0.06	1.39 ± 0.51
8-OxodG	0.35 ± 0.06	2.67 ± 0.08	2.86 ± 0.16	1.17 ± 0.22	2.33 ± 0.47

[a] Treatment: 1) unirradiated DNA (pellet); 2) γ-irradiated DNA (pellet); 3) pellet from γ-irradiated DNA after treatment with inactivated Fpg; 4) pellet from γ-irradiated DNA after incubation with native Fpg; 5) Supernatant fraction from 4.

Treatment of the damaged DNA with the Fpg protein releases the lesions that it recognizes into solution. Analysis of the supernatant reveals the amount of lesions that are excised by the Fpg protein. Comparison of the lesions remaining in the DNA with the lesions that were released by the Fpg protein account for the total amount of lesions produced. This study demonstrates that the purine damage products Fapy•dA, Fapy•dG, and 8-OxodG are substrates for the Fpg protein. Of particular interest is that 8-OxodA is a poor substrate. Further analysis of this observation reveals that 8-OxodA is removed by the Fpg protein, but large amounts of protein are required for excision.[125]

The kinetic parameters for the release of Fapy•dA, Fapy•dG and 8-OxodG were measured.[126] The values for k_{cat} varied significantly for the three substrates, as did the K_m for excision (Table 11). However, the specificity for the Fpg protein (k_{cat}/K_m) was

Table 11.[126] Kinetic parameters for the excision of randomly generated Fapy•dA, Fapy•dG and 8-OxodG through H_2O_2 / Fe-EDTA treatment.

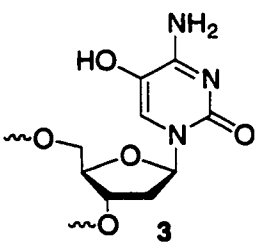
DNA Substrate	k_{cat} (min^{-1})	K_m (nM)	k_{cat}/K_m $(min \cdot nM)^{-1} x10^3$
Fapy•dA	0.054 ± 0.015	412 ± 79	0.132 ± 0.035
Fapy•dG	0.446 ± 0.050	2761 ± 330	0.162 ± 0.018
8-OxodG	0.127 ± 0.014	1004 ± 303	0.127 ± 0.014

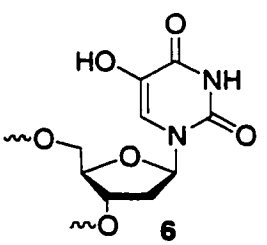
similar for all three substrates. Randomly distributing damage sites throughout DNA may affect the kinetics of excision by a DNA repair protein.[57] The Fpg protein must compete with all damaged sites present and potentially interferes with the recognition of Fapy•dA, Fapy•dG and 8-OxodG. The presence of other lesions also requires that large

amounts of Fpg be used. This is also reflected by the high K_m's that were observed for the excision.

Substrate requirements of the Fpg protein were probed by removing functionality from 8-oxopurines.[125] From these studies, it was noted that the C-6 carbonyl is not required for Fpg binding, but is necessary for deglycosylation. The poor ability of the Fpg protein to remove 8-OxodA from DNA was rationalized based upon this proposal. However, a C-6 carbonyl is not present in Fapy•dA. Therefore, its excision is not explained by the proposed structural requirements. It is speculated that the conformational flexibility of Fapy•dA is important in its recognition by the Fpg protein.[126]

Other substrates for the Fpg protein have been detected. Alkylated Fapy lesions





are substrates, including bulky aflatoxin adducts (Figure 6).[16] Oxidized pyrimidines have also been detected as substrates. Included in the oxidized pyrimidine substrates are 5-hydroxycytosine (**3**) and 5-hydroxyuracil (**6**).[127] These different substrates complicate the understanding of the general recognition events that the Fpg protein uses for excision of DNA modifications.

2.5.3 Other Glycosylases Involved in Recognition of Fapy Lesions:

2.5.3.1 The MutY/MutM (Fpg) System of DNA Repair:

A common defense mechanism that repairs 8-OxodG lesions and its premutagenic product (8-OxodG:dA) is known as the "GO system".[16] High mutational rates are observed when MutY and MutM (Fpg) activities are removed.[115] The frequency of mutations are much lower when only the activity of MutY or MutM are removed. This suggests that MutY and MutM work in tandem in preventing mutation. In the case of 8-OxodG, the 8-OxodG:dC base pair is recognized and repaired by the Fpg (MutM) protein and the 8-OxodG:dA misinsertion product is a poor substrate. However, the latter base pair is efficiently recognized by MutY, which catalyzes the excision of dA. The involvement of the Fpg protein in this system indicates that MutY could possibly have activity in the repair of the potentially deleterious products of Fapy lesions, although no reports that address this speculation have appeared. A third enzyme involved in the "GO system" is MutT.[16] The function of this enzyme is to protect against incorporation of 8-OxodGTP into DNA by polymerases. Therefore, MutT, or 8-OxodGTPase is responsible for removing 8-OxodGTP from the pool of dNTP's for DNA replication.

2.5.3.2 Fpg Protein Homologues:

Other glycosylases that have similar substrate specificities to the Fpg protein are known. These mainly consist of homologues of the Fpg protein found in other species. The predominant classification for these glycosylases is based upon their propensity to excise 8-OxodG from DNA. They are designated as 8-oxoguanine DNA glycosylases (Ogg). These proteins have very little sequence homology to the Fpg protein, but are remarkably similar in their specificities for the 8-OxodG:dC base pair.[16]

The Fpg homologue in the yeast *Saccharomyces cerevisiae* is yOgg1. The repair of Fapy•dG and MeFapy•dG has been reported for yOgg1.[128,129] However, Fapy•dA was not recognized by yOgg1. The inability to excise Fapy•dA was rationalized by the preference of yOgg1 to excise damages that are opposite a cytosine.[128] Since Fapy•dA prefers to base-pair with thymidine, it is not recognized by yOgg1. The isolation of the OGG homologues in mammalian systems has also been accomplished in humans and mice (hOgg1, mOgg1 respectively). These systems exhibit similar specificity for Fapy•dG and MeFapy•dG excision, but Fapy•dA is still not recognized.[101,130]

2.5.3.3 Endonuclease III Family of Glycosylases:

The Endonuclease III (Endo III) family is another class of glycosylases that recognize a wide variety of DNA lesions. The primary substrates for Endo III are damaged pyrimidine nucleobases.[16] Characterization of the substrate specificity of Endo III in *E. coli* (Nth), and yeast (Ntg) has been accomplished.[131,132] The activity of the Nth protein for DNA modifications includes Fapy•dA, but excision of Fapy•dG has not been observed.[131] In contrast to this, Ntg1 and Ntg2 of yeast recognize Fapy•dA, and Fapy•dG.[132] This difference in selectivity is unexpected as the Ntg and Nth proteins have closely related amino acid sequences.[132] Due to the purine structure possessed by 8-OxodA or 8-OxodG, Endo III does not recognize them as substrates.[131,132]

2.5.3.4 Pyrimidine Dimer Glycosylases:

Another class of glycosylases that recognize Fapy lesions is the pyrimidine dimer glycosylases (PDG). The primary responsibility of this glycosylase is to repair

pyrimidine dimers that block DNA synthesis.[16] However, Fapy lesions have been discovered to be substrates for PDG proteins.[133,134] The T4-PDG from *E. coli* has a low catalytic efficiency for Fapy•dA compared to other forms of UV-induced DNA damage.[133] Fapy•dG is not a substrate for T4-PDG.[133] The CV-PDG from *Paramecium Bursaria Chorella* virus has higher catalytic activity towards Fapy lesions and can excise both Fapy•dA and Fapy•dG. [134]

2.5.4 Substrate Analogues of DNA Lesions:

Another common method used to investigate the specificity of DNA repair proteins is through the use of substrate analogues of DNA lesions. Substrate analogues that form stable complexes with enzymes are useful for obtaining structural information about the protein-DNA-substrate recognition. Furthermore, they provide insights into the mechanism of action and the transition states that are involved in the protein-DNA complex. By constructing substrate analogues, a systematic analysis of the functional groups that are necessary for the recognition of a particular lesion can be accomplished. Structure activity relationships (SAR) between the functionalities that are responsible for substrate recognition can be built based upon their interaction with the protein. The development of automated DNA synthesis has made the incorporation of structural analogues of DNA lesions at defined sites in DNA possible.[135,136]

2.5.4.1 Modeling Transition State Requirements for the Design of Substrate Analogues:

The design of substrate analogues that destabilize the transition states in protein substrate complexes are often strong inhibitors for the catalytic activity of the protein. This approach has been used to design several substrate analogues that bind tightly to DNA repair proteins.[16] The transition state for deglycosylation involves the buildup of positive charge between the C_1' and O_4' atoms on the deoxyribose ring (Figure 18). This

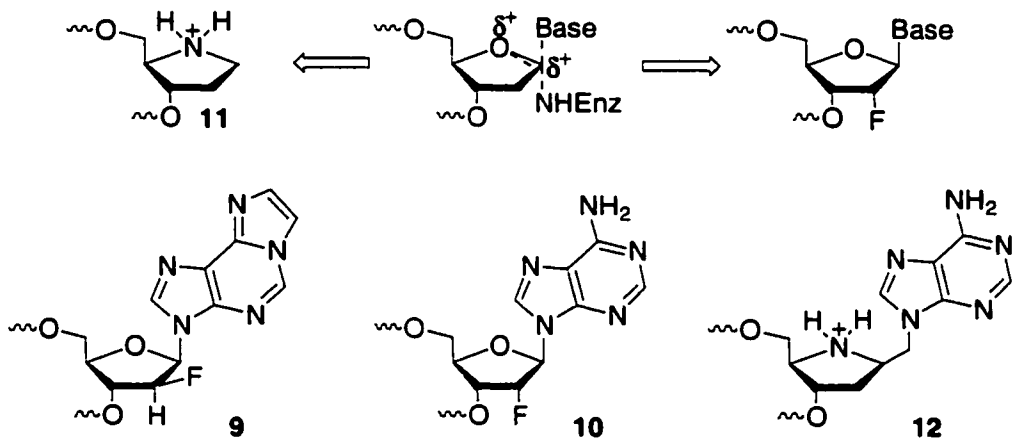


Figure 18. Transition state analogues.

intermediate transition state can be destabilized by incorporation of electron withdrawing groups at the C_2' position. This has been accomplished by introducing 2'-F into analogues. Usually, the addition of one 2'-F is sufficient to remove all glycosylase activity. The substrate analogue of an ethenoadenine (**9**) inhibits the glycosylase activity of alkyl-*N*-purine DNA glycosylase (AlkA).[137] Incorporation of (**10**) opposite a G in a duplex inhibits the repair of the A:G base pair analogue by MutY.[138]

Another modification that has been developed to mimic the transition state is the use of pyrrolidine substrates (**11**).[139-142] These substrate analogues mimic the build up of

positive charge at O4' in the transition state (Figure 18). Substrate analogue **11** binds AlkA very tightly, with a K_d of 16 pM as measured by a gel-shift assay (Section 2.5.4.4).[141] Analysis of **12** shows that MutY binds with a pM K_d.[142] Pyrrolidine-based inhibitors have poor affinity for other DNA glycosylases.[16]

2.5.4.2 Non-hydrolyzable Analogues :

Non-hydrolyzable substrate analogues are designed by altering the anomeric center directly. Stability to deglycosylation can be incorporated by substituting O_4' with CH_2 producing a carbacyclic analogue. The Carba-dG analogue of 8-OxodG (**13**) has a K_d of 34.7 nM for mOgg1, which is comparable to the binding affinity of 8-OxodG (51.5 nM, Figure 19).[143] The Carba-dA analogue (**14**) when opposite 8-OxodG is bound by MutY very tightly with a K_d of 0.3 nM.[144]

Removal of the glycosidic nitrogen can also be used to make a non-hydrolyzable C-glycosidic bond. Some C-glycosides are naturally occurring compounds that have

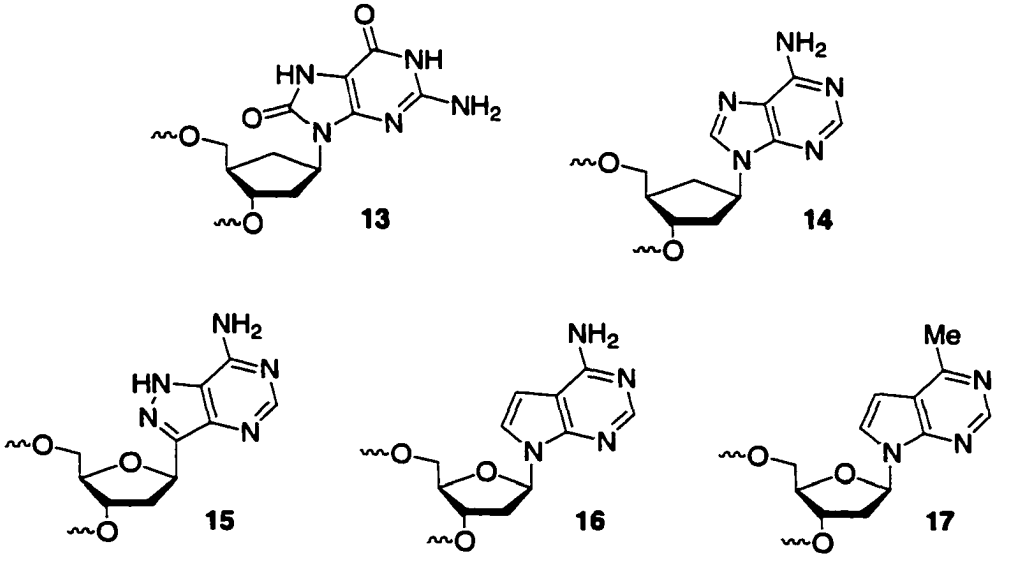


Figure 19. Examples of non-hydrolyzable substrate analogues.

unique biological properties.[146] Two biologically active substrate analogues that bind with high affinity to MutY are formycin A (**15**) and tubercidin (**16**).[138,145] Tubercidin (**16**) is not a C-nucleotide, but substitution of the *N*-7 nitrogen for a carbon prevents its cleavage by MutY. When placed opposite G in duplex DNA these substrate analogues are bound tightly by the MutY protein (K_d= 0.8 and 5.4 nM, respectively).[145] Further modification of **16** by substituting the *N*6 amino group for Me (**17**) decreases the binding affinity 8 fold to 40 nM.[23]

2.5.4.3 Gel Shift Assay for Determining Binding Affinity of Substrate Analogues:

Polyacrylamide gel-shift assay (PGSA) is a common technique for measuring the affinity of a DNA repair protein for a particular substrate.[147,148] The K_d's mentioned above for the substrate analogues were obtained through PGSA.

The K_d can be measured through an equilibrium expression for the dissociation event (Equation 2.13). The equilibrium expression for the association and dissociation can be derived from Equation 2.13 (Equations 2.14 and 2.15). The amount of bound substrate over a broad range of enzyme concentrations is measured and is analyzed

$$\text{DNA} + \text{Enzyme} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \text{[DNA•Enzyme]} \qquad \textbf{Equation 2.13}$$

$$K_a = \frac{\text{[DNA•Enzyme]}}{\text{[DNA] [Enzyme]}} \qquad \textbf{Equation 2.14}$$

$$K_d = \frac{\text{[DNA] [Enzyme]}}{\text{[DNA•Enzyme]}} \qquad \textbf{Equation 2.15}$$

through a plot of the log of enzyme concentration versus the amount of the DNA that is bound.[147] The point at which the DNA concentration is equal to the DNA•Enzyme

concentration is assigned as the K_d. This occurs when the DNA substrate is 50% bound by the enzyme and Equation 2.15 reduces simply to the concentration of the enzyme (Equation 2.16).

The observed values for the dissociation constants can vary between proteins. DNA repair proteins have different binding affinities for dsDNA, so comparisons of K_d

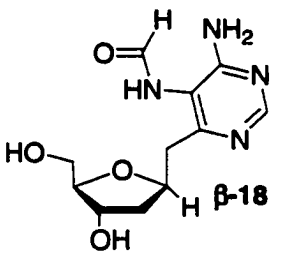
$$K_d = [\text{Enzyme}] \quad \text{when [DNA] =[DNA•Enzyme]} \qquad \textbf{Equation 2.16}$$

values between two different proteins are not meaningful. Furthermore, the purity of DNA repair proteins can also influence the value of the K_d.[147] If the specific activity of the DNA repair protein is low, it will not bind as effectively to the DNA substrate as a protein that is of higher purity.

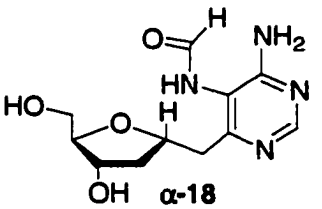
2.6 Research Objectives:

The primary goal of the project discussed is to investigate the effects that Fapy•dA and Fapy•dG have upon the structure and function on DNA. These effects will be addressed by the synthesis of C-nucleoside analogues of Fapy•dA and Fapy•dG and their incorporation into DNA at defined sites. Solid-phase oligonucleotide synthesis will be employed to construct oligonucleotides containing **α-18**, **β-18**, and **β-19**.

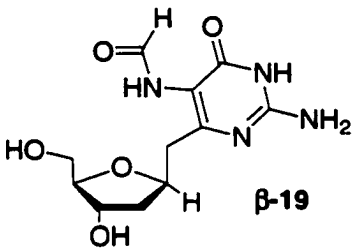
The hypothesis for this research is that C-nucleoside analogues will serve as configurationally stable models for the investigation of the structural and functional effects that Fapy lesions have on DNA. Oligonucleotides containing **α-18**, **β-18**, and **β-19** will serve as probes to determine the role of the individual anomers of Fapy lesions in recognition by polymerases and DNA repair proteins. They will also serve as non-



β-18



α-18



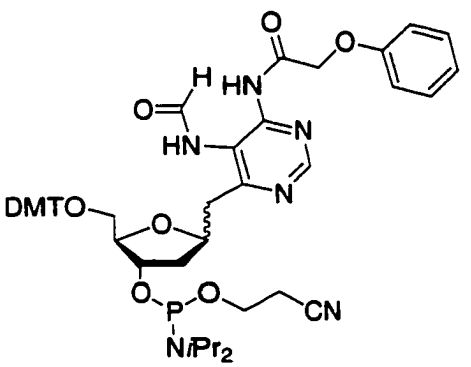
β-19

hydrolyzable analogues of Fapy lesions to probe the recognition requirements for Fapy lesions by DNA repair proteins and also serve as potential competitive inhibitors of DNA repair proteins.

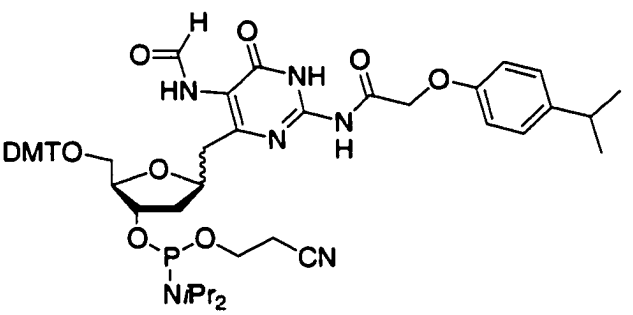
3. Results and Discussion:

3.1 Synthesis of C-Nucleoside Analogues of Fapy Lesions:

In order to probe the hypothesis that **18** and **19** can be used as configurationally stable models of α- and β-Fapy•dA and Fapy•dG, a straightforward synthesis of the respective phosphoramidites was developed (**20** and **21**). Suitable protecting groups that



20



21

are amenable to solid-phase DNA synthesis are also employed. The use of phenoxyacetamide protecting groups on the N2 and N4 positions of **20** and **21** respectively should allow for the selective deprotection the exocyclic amine protecting groups on the synthesized DNA without removal of the labile N5 formamide.[149]

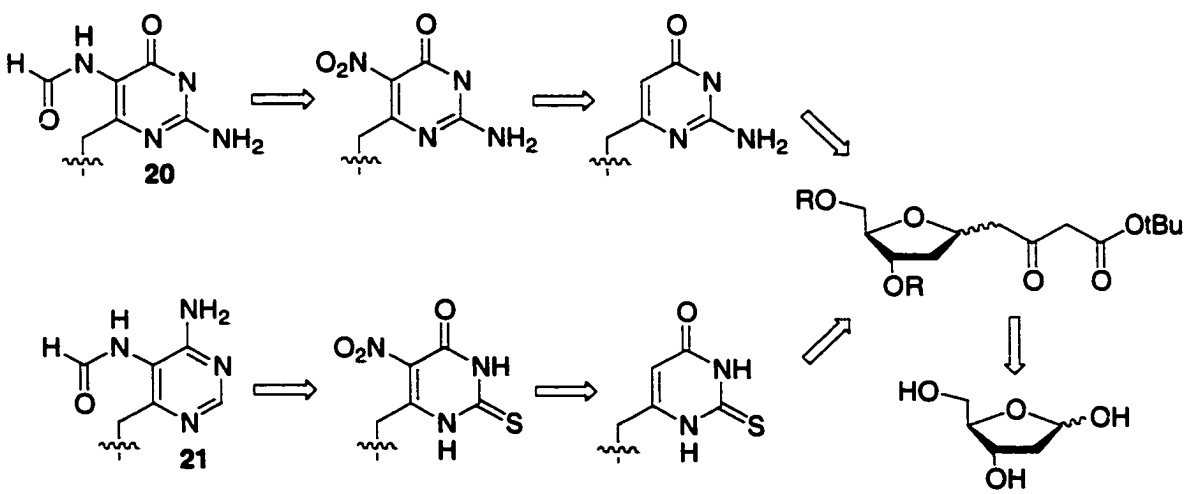
The isolation of C-glycosides from living systems has focused much attention on the construction of C-glycosides that are biologically active.[146] Results of this work include several different methods to generate C-glycosides.[146,150,151] Initially, we anticipated that both diastereomers of **20** and **21** would be needed as structural probes,

but based upon the results of α-**18** in DNA we viewed the use of α-**19** as less important. To accommodate the synthesis of both anomers, formation of the C-glycosidic linkage does not need to be sterospecific. Preferably, the ratios of the α- and β- products from this process are 1:1. It was anticipated that the mixture of diastereomers from this reaction would be separated later in the synthesis to optimize the efficiency of steps required. Based upon these criteria, we selected the Wittig approach for the installation of the C-glycosidic linkage of **20** and **21**.[150]

3.1.1 First Generation Approaches to the Synthesis of 20 and 21:

A general synthesis of C-nucleosides has been developed utilizing the cyclization of guanidine or thiourea with a β-ketoester C-glycoside intermediate.[152,153] Using this approach as a way to synthesize C-nucleosides, we proposed a general route to prepare **20** and **21** (Scheme 1). The β-ketoester intermediate is the point of divergence in the scheme and can be produced from deoxyribose in 3-5 steps depending upon the

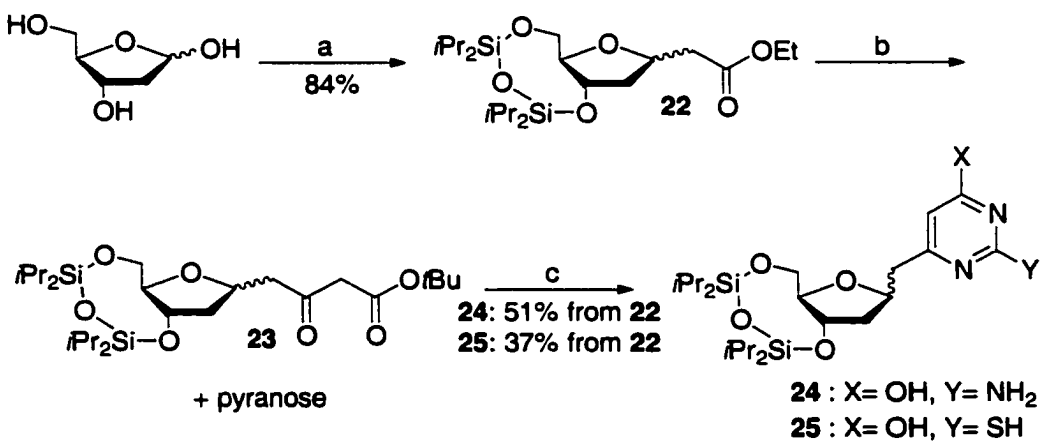
Scheme 1.



method.[154] Cyclization with thiourea or guanidine forms the two distinct C-nucleosides that can be further functionalized into **18** and **19**. The key to this route is the functionalization of the C-5 position of the C-nucleosides with nitrogen. In principle, nitrogen can be introduced through electrophilic aromatic substitution with a variety of electrophilic nitrogen sources.

The synthesis of **22** from deoxyribose is very efficient. Reaction of deoxyribose with PPh_3CHCO_2Et and subsequent protection with $TIPDSCl_2$ affords the C-glycoside (**21**) in high yields (Scheme 2). A disadvantage of this process is that an impurity was

Scheme 2.[a]

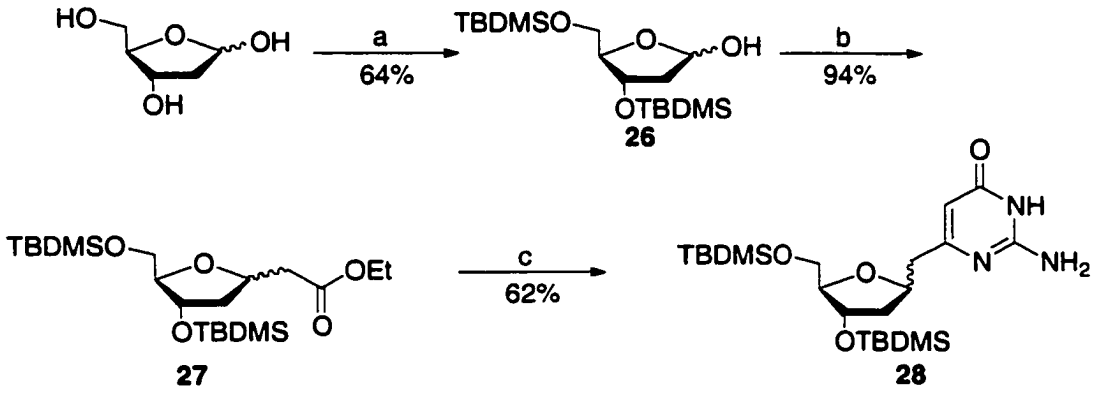


[a]Key: a) i. PPh_3CHCO_2Et, THF; ii. Catalytic NaOEt, EtOH; iii. Pyr, $TIPDSCl_2$, imidazole, DMAP; b) LDA, *t*BuOAc c) for **24**, guanidine•HCl, Na_2CO_3, EtOH; for **25**, thiourea, NaOEt, EtOH.

detected in as high as 20% yield. Analysis by ^1H NMR indicated that this impurity is the pyranose product. Treatment of the C-glycoside with lithium enolate of *t*-butyl acetate under Claisen conditions produces the intermediate β-ketoester (**23**). This material is difficult to purify from poly-alkylated material and the mixture was cyclized directly with guanidine or thiourea afford **24** and **25** respectively.

Efforts were made to improve the synthesis of the core pyrimidine C-nucleoside by pursuing a route that eliminated contamination by the pyranose impurities. The Wittig reaction was carried out on the protected lactol (**26**, Scheme 3). The anomeric position of

Scheme 3.[a]



[a]Key: a) i. Br_2, $CaCO_3$, H_2O; ii. TBDMSCl, DMF, imidazole; iii. DIBAL-H, toluene; b) i. PPh_3CHCO_2Et, THF; ii. MeOH, catalytic NaOMe; c) i. LDA, *t*BuOAc, ii. guanidine•HCl, Na_2CO_3, EtOH.

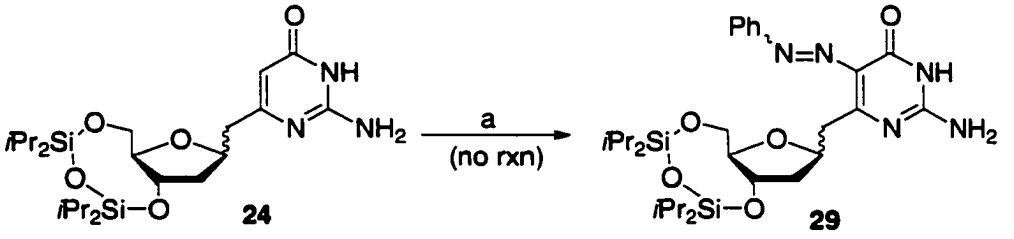
deoxyribose was oxidized with Br_2 and the resulting diol was protected with TBDMSCl to afford an intermediate lactone.[155,156] Reduction of the lactone with DIBAL-H[155] provided **26** and subsequent reaction with PPh_3CHCO_2Et affords **27** free from the pyranose impurity. This compound was treated as above to form **28.** The overall yield of **28** from deoxyribose is comparable to the synthesis of **25** (37% and 43% respectively) and is amenable to relatively large reaction scales on 5-10 g of material.

3.1.1.1 Attempts to Functionalize the C-5 Position by Electrophilic Aromatic Substitution:

Efforts to functionalize the C-5 position of **24** followed standard procedures for the functionalization of analogous pyrimidines at the C-5 position.[157] Diazo-couplings

are a common technique for the incorporation of nitrogen electrophiles at the C-5 position. The product of this reaction is an arylazopyrimidine that can be reduced to the corresponding amine.[158] However, only unreacted **23** was isolated upon treatment with *in situ* generated $PhN_2^+Cl^-$ (Scheme 4).[158]

Scheme 4.[a]

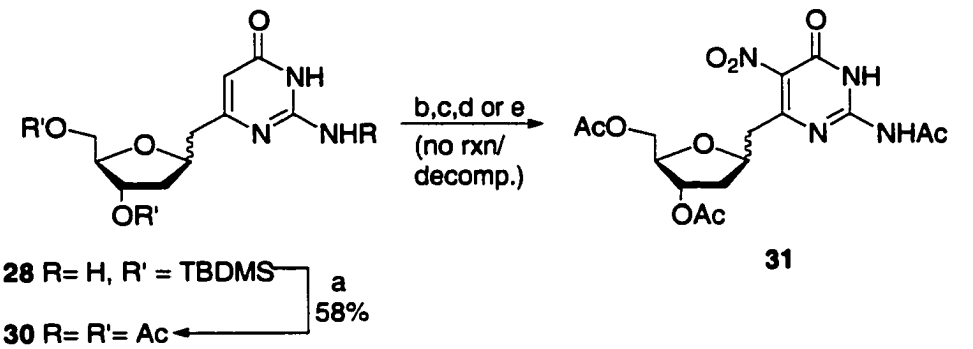


[a]Key: a) $PhN_2^+Cl^-$, Na_2CO_3.

Another method for the substitution of the C-5 positions of pyrimidines is nitration.[157] Nitration of pyrimidines is usually accomplished under extreme acidic conditions not amenable to nucleosides. Several mild nitration conditions have been developed for the nitration of nucleosides.[159-161] Most of these conditions use NO_2^+ surrogates that do not require the use of harsh acidic conditions for its formation.[159,160]

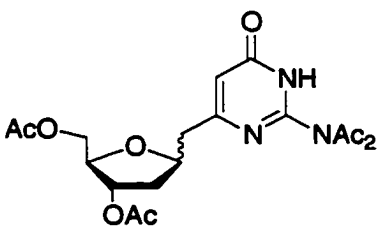
Initial attempts to nitrate **24** led to desilylation under mildly acidic conditions (TfOH /nitropyrazole)[159] or Lewis acidic conditions ($NO_2^+BF_4^-$).[161] We then switched to the more robust acetate groups. The use of acetate protecting groups can be problematic later in the synthesis due to the requirement that they be selectively removed in the presence of the C-5 formamide. Thus, additional steps would be required to convert the acetates back to silyl protecting groups prior to incorporating the formamide. The conversion to the triacetate (**30**) was accomplished by deprotection of **28** with $Et_3N{\bullet}3HF$

Scheme 5.[a]



[a]Key: a) i. $Et_3N{\bullet}3HF$, ii. Ac_2O, pyr, catalytic DMAP; b) $Cu(NO_3)_2$, Ac_2O; c)TfOH, nitropyrazole, AcN; d) $NO_2^+BF_4^-$, AcN; d) H_2SO_4, HNO_3 (dec.).

and treatment with acetic anhydride (Scheme 5). The triacetate **30** is stable to the reaction conditions of TfOH/nitropyrazole or $NO_2^+BF_4^-$, but no reaction is observed after using large excesses of these reagents. The use of $Cu(NO_3)_2/Ac_2O$ led to a small amount of the bis-amide of **30**.[160] Presuming that the C-glycoside would be stable to concentrated acidic conditions, we attempted nitration under classic conditions (H_2SO_4/HNO_3).[162] This only lead to decomposition of **30.** Presumably, the electron deficiency of the pyrimidine ring is contributing to the lack of reactivity of **30** towards electrophilic aromatic substitution. Removal of the *N*-acetamide could potentially improve the reactivity by increasing the electron density of the pyrimidine ring. However, efforts to nitrate the diacetate of **28** also led to no reaction.

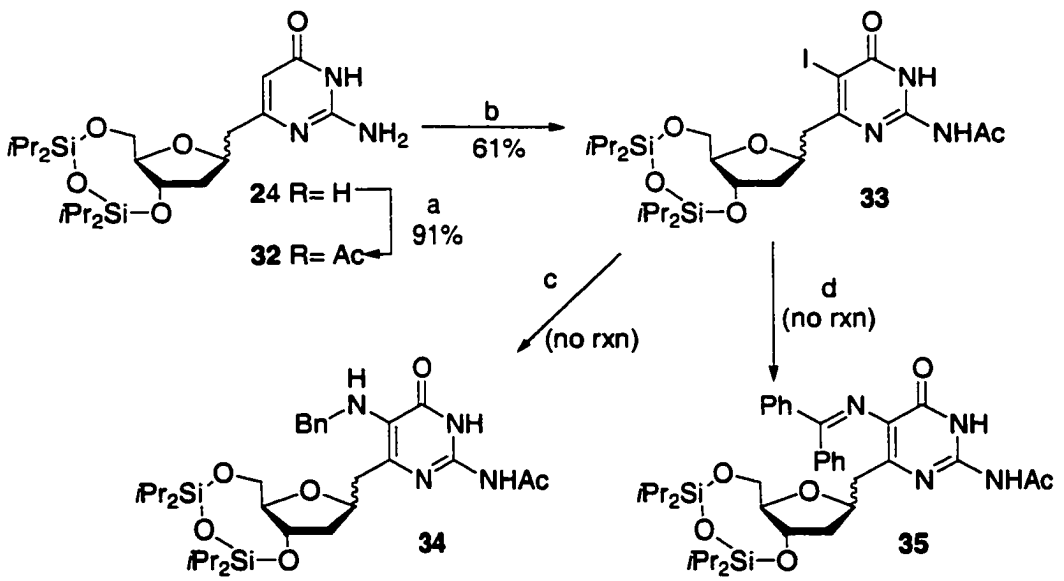


bis-amide of **30**

3.1.1.2 Attempts to Functionalize C-5 through Cross-Coupling Reactions:

Recent advances in transition metal cross-coupling reactions of aryl halides with amines have allowed for the efficient synthesis of aryl amines.[163,164] Several examples exist for application of these reactions to purine systems, but no examples for pyrimidine systems have been reported.[165,166] Since transition metal insertions have been commonly used with C-5 iodopyrimidines for Heck couplings, we rationalized that we could form a C-5 aminopyrimidine.[167] Acetylation and halogenation of **24** with ICl produced **33** in modest yields (Scheme 6).[168] Using similar reaction conditions employed for the purine

Scheme 6.[a]



[a]Key: a) Ac_2O, pyr, DCM; b) ICl, AcN; c) $Pd(dba)_3$ (5 mol%), BINAP (15 mol%), toluene, benzophenone imine; d) $Pd(dba)_3$ (5 mol%), BINAP (15 mol%), toluene, benzyl amine.

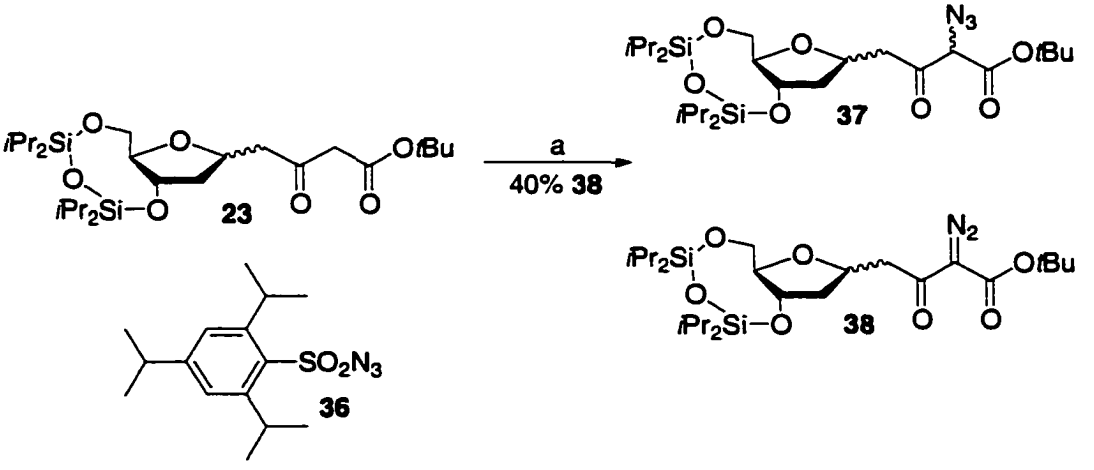
systems, the cross-couplings with benzophenone imine and benzyl amine were investigated.[165-166] After extended reaction times only a limited reaction is observed. Unreacted **33** is recovered along with 5% of the reduced product **32**. The latter resulted from Pd insertion into the aryl-I bond and subsequent rearrangement to the π-allyl system

through the γ-H at the methylene group. Following reductive elimination, **32** is formed along with the Pd-I species. This side reaction renders the catalyst useless. Due to the proximity of the methylenic H's to the insertion event, this side reaction cannot be avoided and this method cannot be used for the functionalization of C-5 in the molecule of interest.

3.1.1.3 Attempted Functionalization of the β-Ketoester:

Introduction of nitrogen functionality at the C-5 position can also be accomplished through functionalization of the α-position of the β-ketoester. Subsequent cyclization produces the desired C-5 nitrogen substituted pyrimidine.[157,169] The trapping of enolates by electrophilic nitrogen sources is a common way of introducing N-functionality adjacent to a carbonyl.[170] Diazonium cations, diazodicarboxylates, alkyl nitrites, and arylsulfonyl azides have been used as nitrogen electrophiles.[153,170-173] Initial attempts involved azidification.[171,172] Formation of the enolate of **23** was accomplished

Scheme 7.[a]

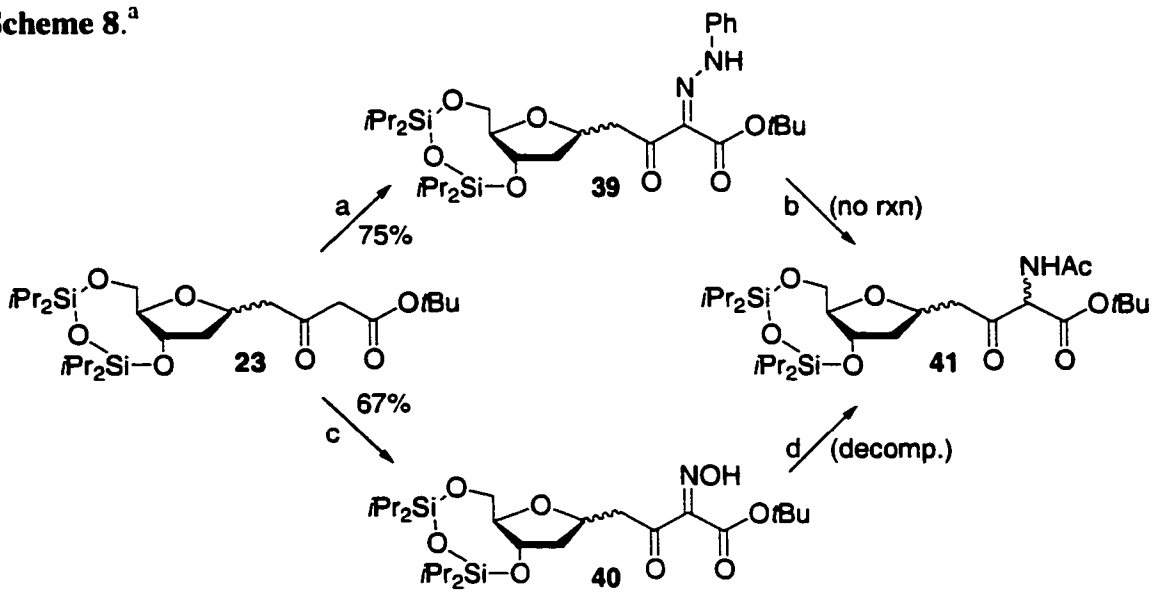


[a]Key: a) i. KHMDS, THF; ii. **36**, -78 °C; iii, AcOH, -78 to 25 °C.

with KHMDS (Scheme 7). Trapping of the enolate with trisyl azide (**36**) and quenching the reaction with acetic acid led to formation of the undesired diazo-transfer product (**38**). Structural assignment is supported by the isolation of trisyl amine from the reaction upon workup. Azide transfer is the major product for most cases where the potassium enolate is made from an ester. However, when the sodium and lithium enolates are trapped with **36**, mixtures of the diazo-transfer and azide transfer are observed.[171] It is rationalized that this observation is due to the decreased reactivity of the sodium and lithium enolates. Furthermore, stabilized enolates prefer to give diazo-transfer.[172] The observation of diazo-transfer in **23** is consistent with this since the enolate is delocalized between the two carbonyls of the β-ketoester,

Investigations into the reaction of the enolate of **23** with other electrophilic nitrogen species were examined. Trapping of the enolate with $PhN_2^+Cl^-$ produced the

Scheme 8.[a]

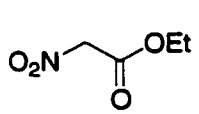


[a]Key: a) $PhN_2^+Cl^-$, NaOAc, EtOH; b) Raney Ni, EtOH, H_2 c) isoamyl nitrite, NaOEt, EtOH; d) i. 10% Pd/C, EtOH, H_2 (50 psi) ii. Ac_2O, pyr.

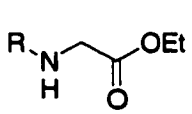
hydrazone intermediate (**39**) (Scheme 8).[153] However, reduction with Raney Nickel yielded unreacted **39**. Results obtained following formation of the oxime (**40**) were similar, although reduction with Pd/C led to a complex mixture of products that did not include **41**. The supposition is that once the amine is generated, it immediately reacts with the β-ketone of another molecule in solution giving rise to a complex mixture of dimeric products.

3.1.1.4 Attempted Functionalization Through Glycine Enolates:

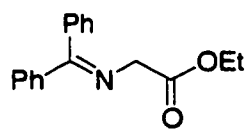
The construction of α-amino-β-ketoesters can be accomplished through other means. A common way to produce α-amino-β-ketoesters is through the condensation of glycine enolates with esters and acid halides.[174-176] Similar approaches are used with aldehydes, except an oxidation step is required to arrive at the α-amino-β-ketoester.[177-180]



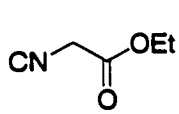
42



43: R=$COCF_3$
44: R=Bz



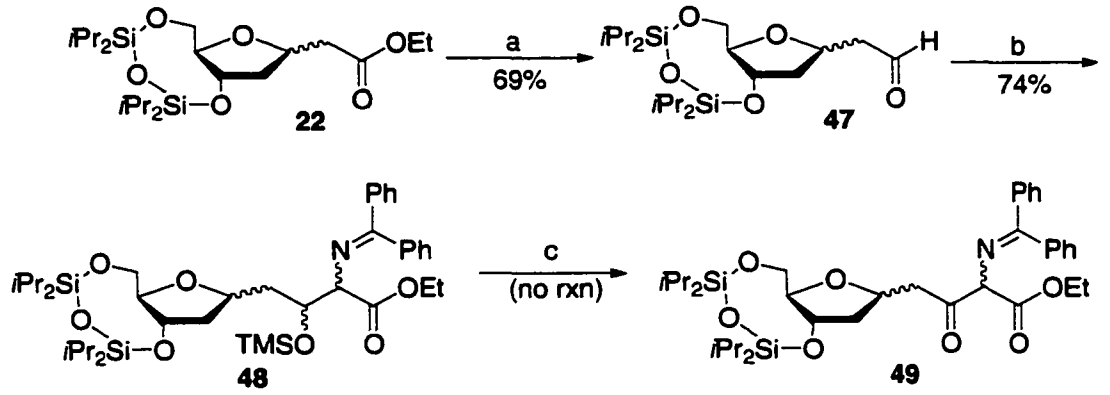
45



46

Initial attempts at this approach were performed with the enolate of **42**. However, model studies revealed that this enolate is unreactive due to the stabilizing effect of the nitro group. Similar studies showed that the dianion enolates of **43** and **44** were not suitable substrates for the reaction.[174,175] The reaction of **45** could not be performed on an acid halide of the C-glycoside of **22** due to an interesting rearrangement of 45 that was observed with acid halides that forms a substituted aziridine.[181] Therefore, the reaction with **47** was investigated (Scheme 9). Treatment of **22** with DIBAL-H affords **47**.

Scheme 9.[a]

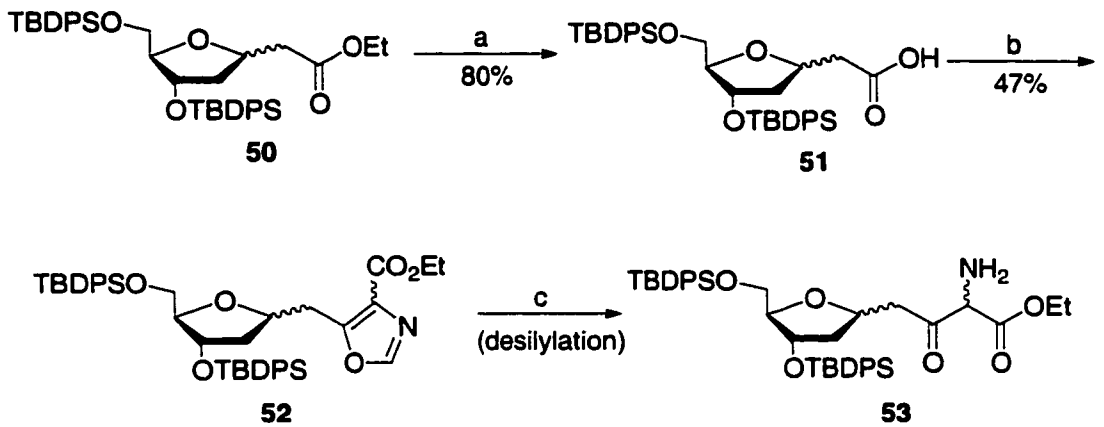


[a]Key: a) DIBAL-H, toluene; b) i. LDA, TMSCl, **45**; ii. $ZnCl_2$; c) Oxidation.

Reaction with the silyl-enol ether of **45** with **47** in the company of $ZnCl_2$ produces **48**.[179,180] Attempts to oxidize **48** directly to **49** were unsuccessful using PDC, Dess-Martin periodinane, Swern, and Moffat oxidation conditions.[182] These conditions are known to oxidize TMS ethers to carbonyls. Efforts to remove the TMS group under acidic conditions or using fluoride were unsuccessful. Although desilylation occurred with treatment with fluoride, selective desilylation of the TMS ether could not be obtained.

The reaction of glycine isocyanates with acid halides can also be used to produce α-amino-β-ketoesters.[176] An intermediate oxazole-4-carboxylate (**52**) was produced that rearranges to an α-amino-β-ketoester upon treatment with acid. In order to utilize this method, TBDPS protecting groups were used to prevent desilylation under the acidic conditions. The saponification of **50**[183] and conversion to the acid halide was accomplished (Scheme 10). Treatment of the acid halide with the enolate of **46** afforded **52** in modest yields. Attempted rearrangement to **53** was investigated under mild acidic

Scheme 10.[a]



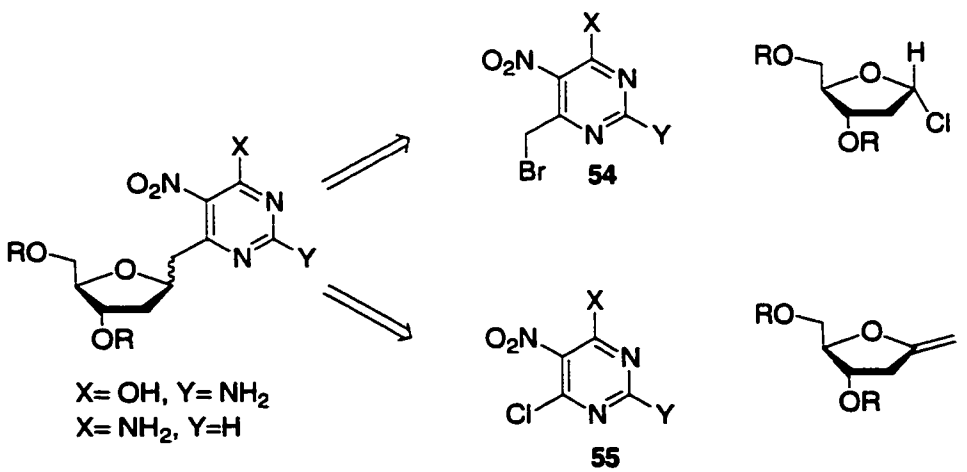
[a]Key: a) LiOH, MeOH b) i. $SOCl_2$, DMF, DCM; ii, **46**, *n*-BuLi, THF; c) 1:1 TFA:EtOH.

conditions (5% TFA-EtOH) at room temperature. Under these conditions no reaction was observed. Upon addition of 50% TFA hydrolysis of the silyl groups was observed.

3.1.2 Pyrimidine Ylide Approach to the Synthesis of 20 and 21:

Unsuccessful functionalization of the β-ketoester and the C-5 position of C-nucleosides with nitrogen functionality prevented us from preparing **20** and **21** by these

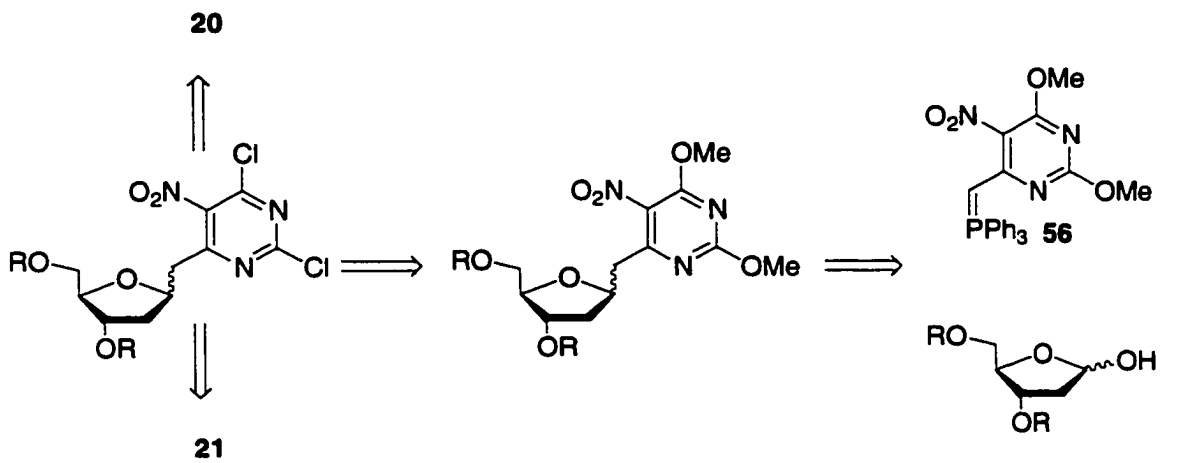
Scheme 11.



routes. Methods to incorporate the pyrimidine with a C-5 nitro group to form a C-nucleoside directly are possible. This avoids the limitations of the synthesis discussed above. Construction of the C-glycoside through traditional carbanion chemistry requires the formation of a benzylic anion (Scheme 11). It is unlikely that this approach will be successful because the electron withdrawing nature of the ortho nitro group suppresses the reactivity of the carbanion of **54**. Alternatively, C-glycosides can be constructed through tandem hydroboration/Suzuki couplings (Scheme 11).[184,185] Hydroboration of 1-exomethylene carbohydrates and Pd-cross coupling with an aryl halide produce C-glycosides in modest yields. However, the scope of substrates accommodated by this reaction is limited to 4-substituted phenyl groups. Cross-coupling with non-polar pyrimidine substrate results in poor yields.[184] In preliminary studies with Pd-cross-coupling on functionalized pyrimidines (**55**, X-OH, Y= NH_2), no reaction occurred due to the poor solubility of the pyrimidine.

With the improbable success of the above ideas, we turned our attention to back to another variation of the Wittig approach. The Wittig approach discussed above has been used successfully to make C-nucleosides through the use of pyrimidine- ylides.[186-189] This approach was considered for the synthesis of **20** and **21**, but the extensive modification of the pyrimidine core and synthesis of the pyrimidine-ylide initially appeared unattractive (Scheme 12). On the other hand, the clear advantage of the pyrimidine-ylide approach is that the C-5 nitrogen functionality is introduced prior to the formation of the C-nucleoside. The proposed syntheses of **20** and **21** involves an intermediate dichloride that can be selectively hydrolyzed or aminated at the C-4 chloride to produce the divergence to **20** or **21**.[157,190,191]

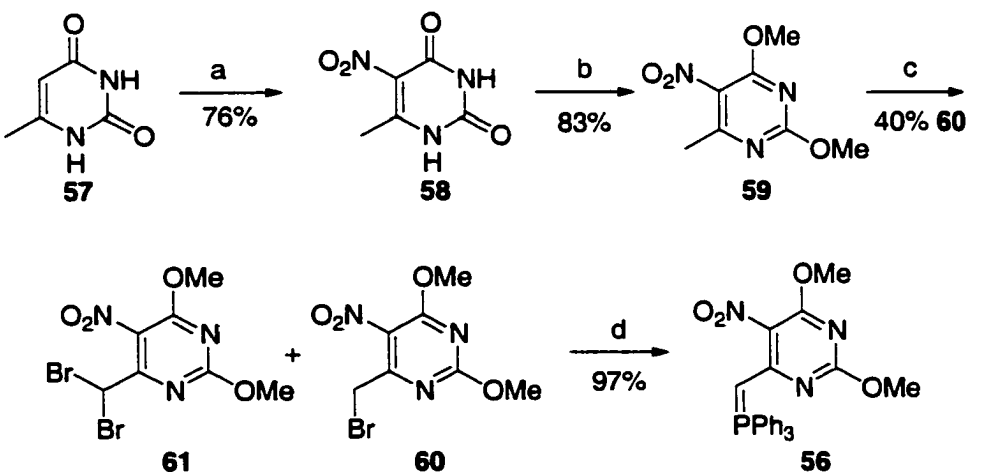
Scheme 12.



3.1.2.1 Synthesis of the Pyrimidine Ylide (56):

The synthesis of pyrimidine-ylide **56** is straight forward and was improved over the literature precedent.[189,192] Nitration of 6-methyl uracil (**57**) proceeds smoothly with H_2SO_4 and HNO_3 to produce **58** (Scheme 13). Chlorination of **58** is accomplished with

Scheme 13.[a]



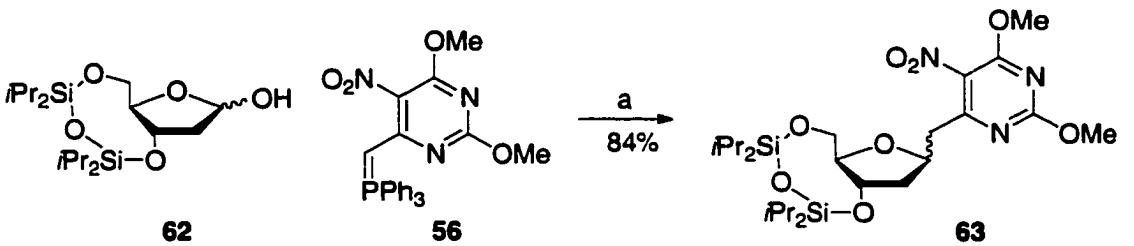
[a]Key: a) H_2SO_4, HNO_3; b) i. $POCl_3$, DEA; ii. MeOH, NaOMe c) Br_2, NaOAc, AcOH; d) i) PPh_3 ii) NaOH.

$POCl_3$ and the crude dichloride is treated with NaOMe to form **59**. Bromination of **59** creates a 1:1 mixture of **60** and **61**, along with a slight amount of unreacted starting material.[192] An attempt to drive the reaction to completion by the longer reaction times, or the addition of more Br_2, significantly increases the amount of **61**. Treatment of **60** with PPh_3 produces a phosphonium bromide salt that is converted to **56** by a biphasic treatment of aqueous NaOH in chloroform.[189]

3.1.2.2 Synthesis of 20:

The treatment of **62** (formed in a similar fashion as **26**)[193] with **56** in refluxing toluene for 65 h affords acyclic olefin, which cyclized directly to **63** upon treatment with catalytic NaOMe in MeOH (Scheme14).[192] The extensive reaction times required are a

Scheme 14.[a]



[a]Key: a) i. toluene; ii. MeOH, catalytic NaOMe.

result of the C-5 nitro group stabilizing the anion at the benzylic position. Separation of a small portion of the α- and β-diasteromers of **63** facilitated their sterochemical assignment at by NOE experiments involving the C_1' proton (Figure 20). The β-diastereomer was confirmed through observation of the diagnostic enhancement at C_4' upon irradiation at C_1'. The α-diastereomer was assigned through the resulting enhancement at C_3' upon irradiation at C_1'. The C-nucleosides also demonstrate a

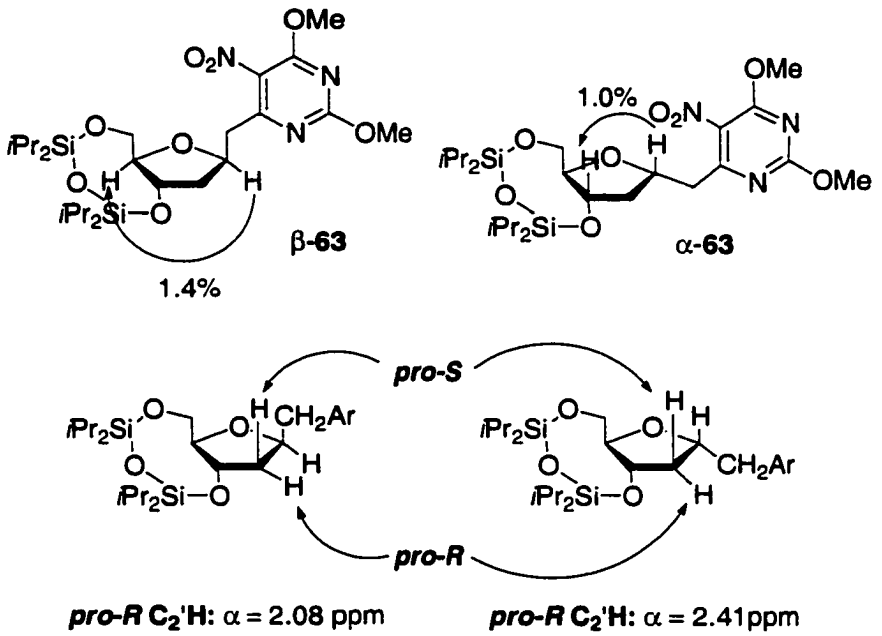
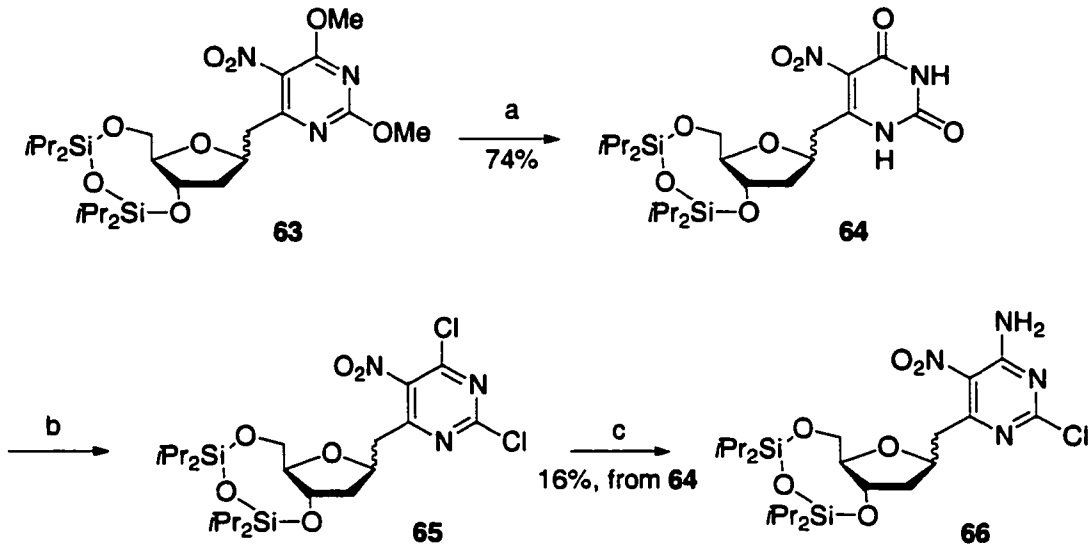


Figure 20. Determination of stereochemistry in **63** by NOE experiments.

characteristic chemical shift for the *pro-R* C_2'-H for the α- and β- anomers. The *pro-R* C_2'-H for the α-anomer is shifted downfield by 0.33 ppm relative to the β-anomer. The separation of the diastereomers of **63** was impractical on larger amounts of material, so the mixture was taken on to the next step. Removal of the methoxy groups of **63** was accomplished through treatment with *in situ* generated TMSI to produce **64** (Scheme 15).[194]

Chlorination of **64** under a variety of conditions produced the intermediate dichloride (**65**) in low yields. Aminolysis proceeded in 6:1 selectivity for the C-4 position in a 2-step yield of 16%. The assignment of the proper regioisomer was made at a latter stage in the synthesis confirming that the amine displacement occurred at the C-4 position. The low yield of the 2-step conversion was attributed to the instability of the

Scheme 15.[a]



[a]Key: a) TMSCl, NaI, $CHCl_3$; b) $POCl_3$, diethylaniline, toluene; c) THF, NH_3.

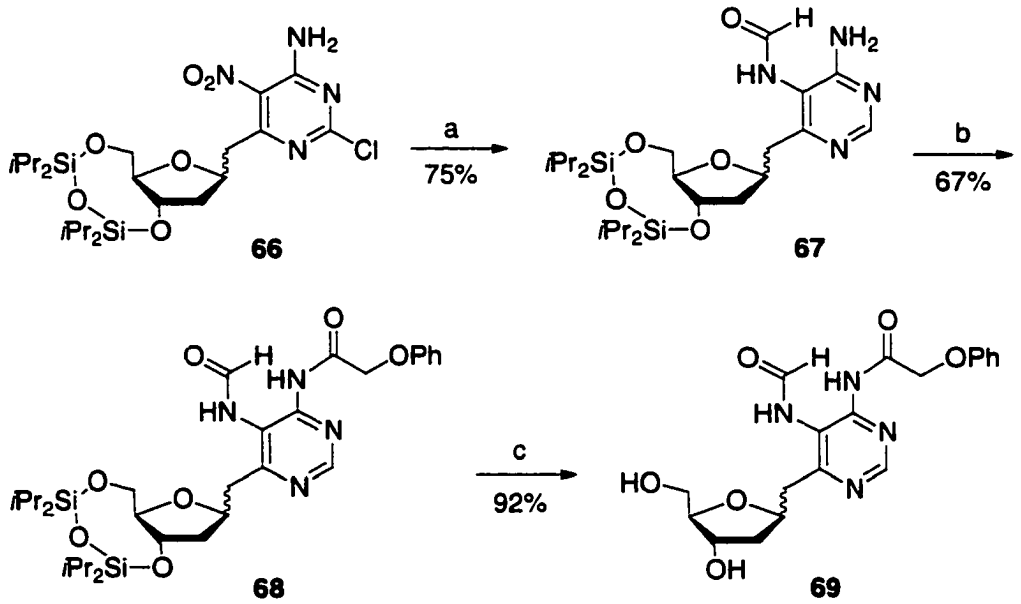
disiloxane protecting groups to acidic conditions and not to the nucleophilic substitution step. Changing the protecting groups to TBDPS did not improve the yield. Attempts were made to limit the acidity of the reaction by adding large excesses of diethylaniline to the reaction. However, large amounts of diethylaniline complicated product purification, and the isolated yields of **66** were only marginally improved.

The reason for the low yield is unknown. Presumably, the C-5 nitro-group deactivates the pyrimidine ring towards the chlorination by hindering the formation of the phosphorodichloridate.[195] Using alternate halogenation conditions of $SOCl_2$ or $PPh_3 \bullet Br_2$ led to no reaction.[196,197] Efforts to increase the electron density of the pyrimidine ring through reduction of the nitro-group did not improve the yield. Having been unable to improve the yields significantly for the chlorination reaction, the decision was made to proceed with the synthesis of **20**.

With limited amounts of **65** available, the selective hydrolysis with NaOAc was unable to be thoroughly investigated. Initial attempts led to no reaction and an alternative synthetic route that avoids the **65** would be needed for the synthesis of **21** (Section 3.1.2.3).[190]

Reduction of **66** with 10% Pd/C in the company of Na_2CO_3 removes the C-2 chloride and generates an intermediate diamino C-nucleoside (Scheme 16). The crude diamino C-nucleoside was formylated with acetic formic anhydride. Formylation

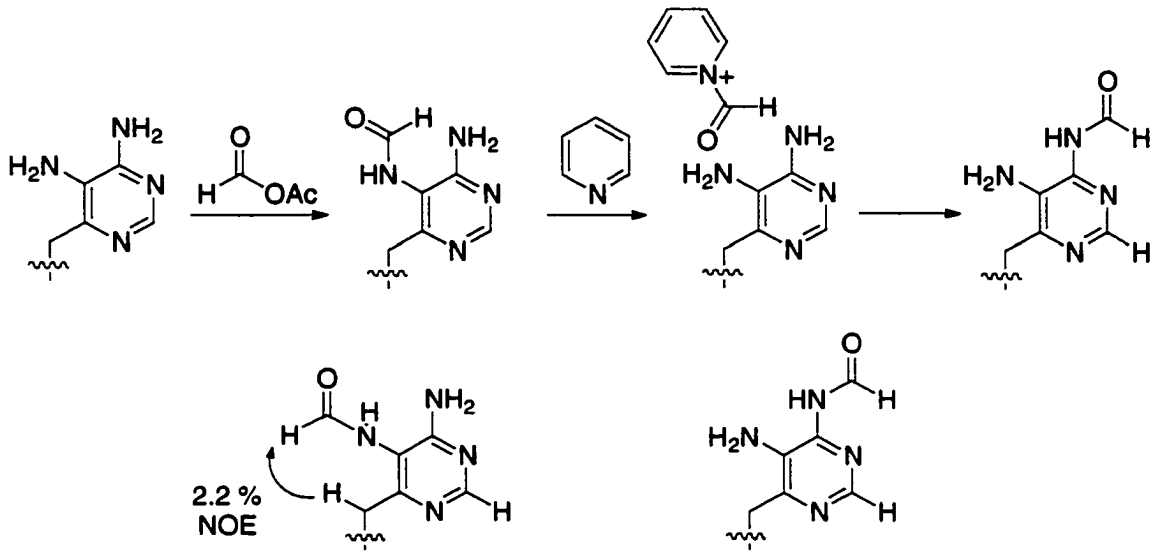
Scheme 16.[a]



[a]Key: a) i. 10% Pd/C, Na_2CO_3, EtOH, H_2; ii. AFA, pyridine, DCM; b) phenoxyacetic anyhdride, DCM; c) $Et_3N{\bullet}3HF$, THF.

occurred at the *N*-5 position selectively, but migration to the *N*-4 amine was observed in the presence of large amounts of pyridine under prolonged reaction times (Scheme 17). Separation of the regioisomeric formamides allowed for the assignment of the *N*-5 and *N*-4 formamides by ^1H NMR. The *N*-5 formamide exhibits a strong NOE between the

Scheme 17.[a]

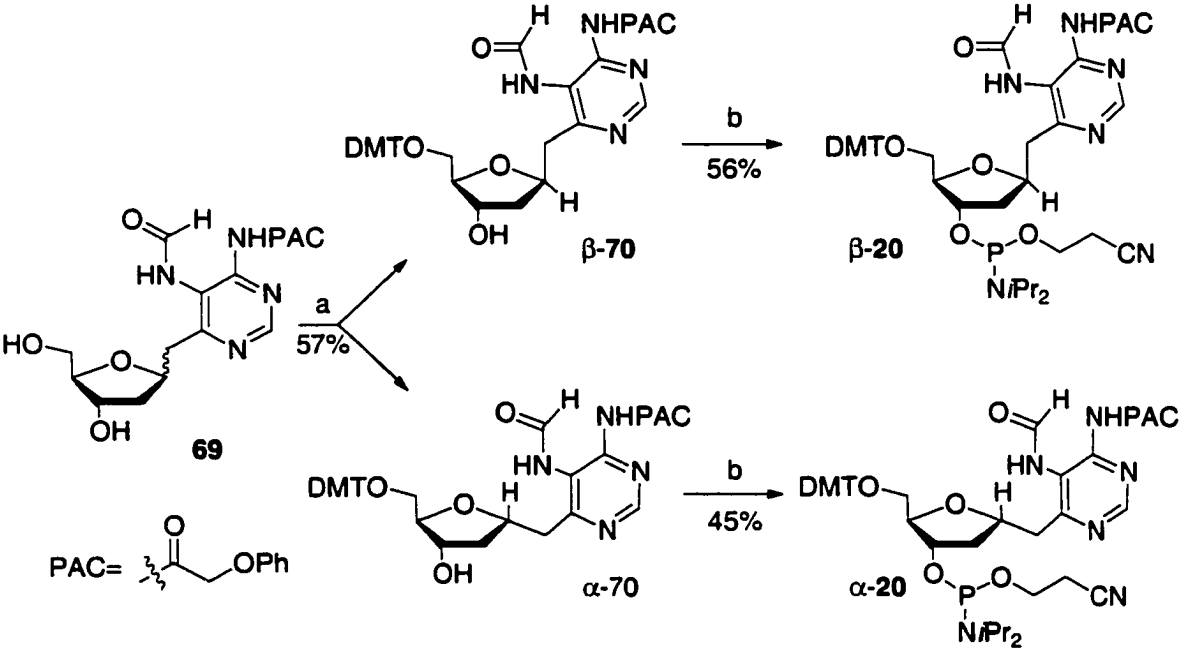


***N*5-Formamide C2-H** =8.28 ppm ***N*4-Formamide C2-H** = 8.56 ppm

formyl proton and the ortho methylene group. Furthermore, the *N*-5 formamide has a 0.2-0.3 ppm upfield shift of the C2 vinyl proton relative to the *N*-4 formamide. The migration event is ascribed to the effect of pyridine catalyzing the migration of the formyl group. Nucleophilic catalysis of the *N*-5 formamide by pyridine and subsequent intermolecular attack by the *N*-4 amine results in the migration. Reducing the amount of pyridine, and limiting the reaction times prevented the migration event forming **66** in 75% yield from **65**. The ability of the *N*-5 formamide to migrate illustrates the need to protect the *N*-4 amine.

Protection of the *N*-4 amino group in the absence of nucleophilic base was accomplished with *in situ* formation of phenoxyacetic anhdride[199] to produce **68**. Subsequent desilylation with $Et_3N{\bullet}3HF$ afforded **69**. Reaction of **69** with dimethoxytrityl chloride afforded **70** and the epimeric mixture of products was separated by column chromatography (Scheme 18). Elucidation of the stereochemistry at C_1' was

Scheme 18.[a]



[a]Key: a) DMTCl, pyr, catalytic DMAP (7:3 β:α); b) 2-Cyanoethyl-*N,N,N',N'*-tetraisopropyl phosphane, diisopropylamino tetrazolide, DCM.

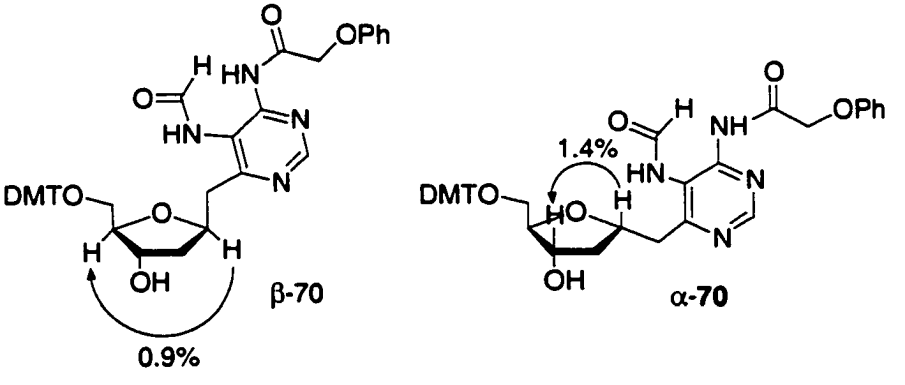


Figure 21. Determination of stereochemistry in **70** by NOE experiments.

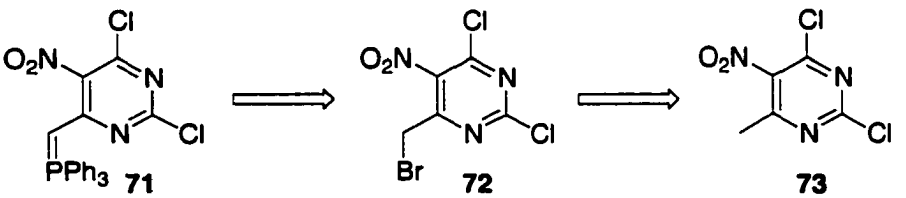
accomplished by NOE experiments (Figure 21). Irradiation at C_1' in the β-anomer produces an enhancement at C_4' in **β-70**. In contrast, an enhancement of C3' was observed for the α-anomer (**α-70**). The conversion of **70** to the phosphoramidite was

accomplished by the use of *N,N,N',N'*-tetraisopropylphosphane and the *in situ* generated tetrazolide to form **20**.[198]

3.1.2.3 Synthesis of 21:

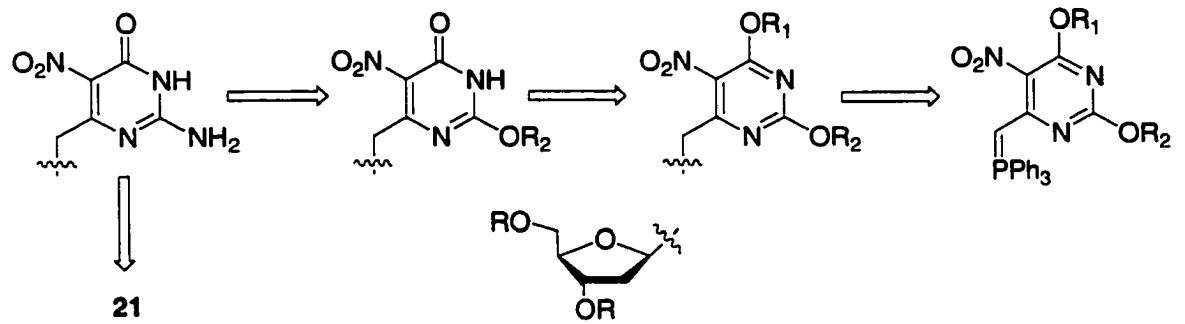
The limitations of the dichloride intermediate in the synthesis of **20** prevented the use of the same synthetic route to produce **21**. Another synthetic approach to the dichloride intermediate (**65**) is to introduce it through the pyrimidine-ylide (**71**). This would enhance the synthesis of **20** and **21** by reducing the number of transformations required on the C-nucleoside. However, attempts at brominating the methyl group of **73** were unsuccessful (Scheme 19). Bromination of the 6-methyl group is only successful

Scheme 19.



when two alkyl ethers are present at C-2 and C-4. One of the undesirable aspects of using the dichloride intermediate is that it is prepared from an O-4 functionalized pyrimidine. If we could retain this functionality, the need for the dichloride intermediate could be eliminated during the synthesis of **21**. This would also provide a shorter synthetic route to **21**. By designing a pyrimidine-ylide that contains different alkyl ethers at the C-2 and C-4 groups and selectively removing the C-4 alkyl ether, the O-4 functional group required for **21** can be introduced (Scheme 20). Aminolysis of the remaining C-2 alkyl ether installs the requisite functionality for **21**. Furthermore, by

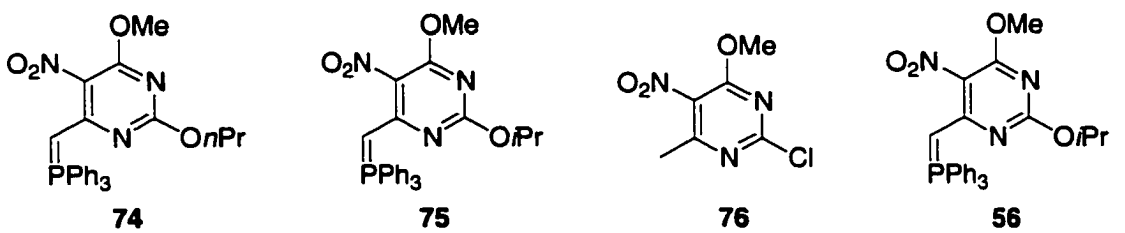
Scheme 20.



introducing two different alkyl ether groups onto the pyrimidine-ylide, the bromination of the 6-methyl group should be unaffected.

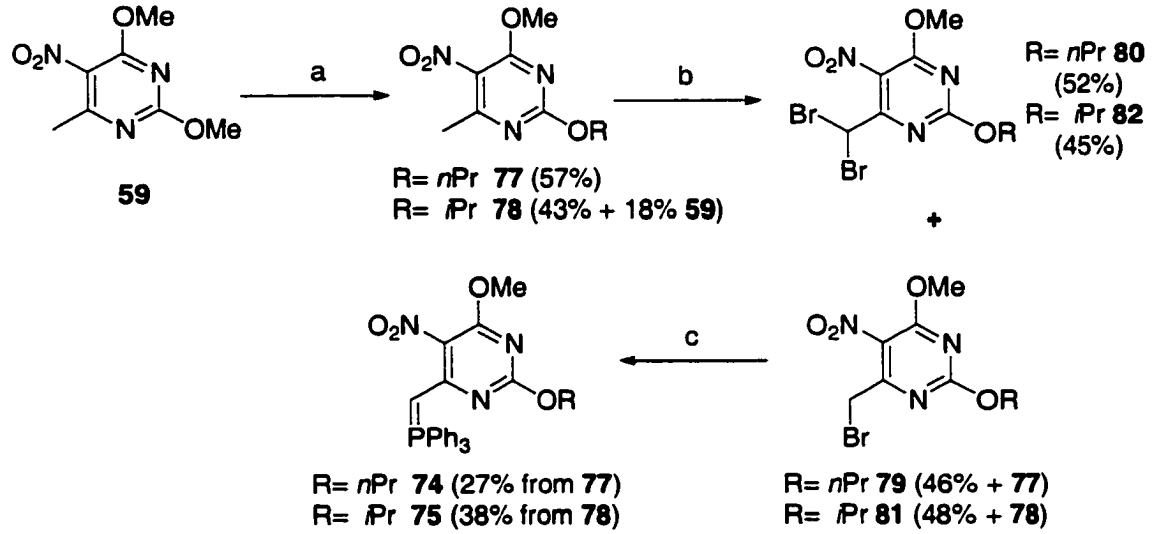
3.1.2.3.1 Synthesis of Pyrimidine Ylide 75:

The ability to demethylate alkyl ethers using TMSI was illustrated above during the preparation of **64** (Scheme 15). TMSI can demethylate primary and secondary alkyl ethers, but at a much slower rate than methyl ethers.[194] Based upon the preference for reactivity with methyl ethers, we decided to install a methyl ether at R_1 (Scheme 20) and to incorporate a *n*-propyl (**74**) or an isopropyl alkyl ether (**75**) at R_2. The syntheses of **74**



and **75**, are analogous to the synthesis of **56**. Initial attempts at the syntheses of **74** and **75** involved the displacement of the C-2-Cl of **76** with the respective Na-alkoxides. However, the inductive effects of the nitro-group activates the C-4 OMe for substitution.[157] As a result, competitive substitution at the C-2 and C-4 position occurred

Scheme 21.[a]



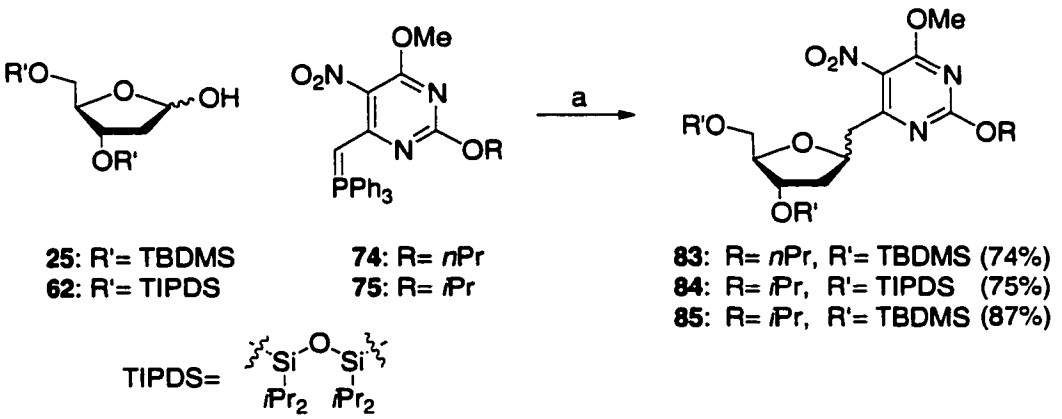
[a]Key: a) ROH, Ag_2O, reflux; b) Br_2, NaOAc, AcOH; c) i. PPh_3, Benzene; ii. $CHCl_3$, H_2O, NaOH.

producing a complex mixture of disubstituted, 4-substitued, 2-substituted, and unreacted products. The use of silver alkoxides in the transetherification of **59** overcame this limitation (Scheme 21).[200] The transetherification of **77** proceeded smoothly with small amounts of the disubstituted product being formed.[200] Through careful chromatography, **77** was purified and brominated to produce a 1:1 mixture of **79** and **80** with a small amount of starting material remaining. Chromatographic separation of **79** from **80** was accomplished, but separation of **79** from the remaining **77** was ineffective. Attempts to drive the reaction to completion increased the amounts of **80** at the expense of **79.** Therefore, the mixture of **79** and **77** was treated with PPh_3 to form the phosphonium bromide of **74**. The phosphonium bromide was separated from **77** by filtration. Treatment of the phosphonium bromide in the same fashion as **60** produced **74** in 27% yield from **77**. Similar transformations were used in the formation of **75**.

3.1.2.3.2 Selective Dealkylation Leads to the Formation of 21:

The Wittig reaction of **74** and **75** with lactols **25** and **62** produced a mixture of the intermediate *E*- and *Z*-acyclic olefins upon prolonged refluxing in toluene. Cyclization with catalytic NaOMe in MeOH generated C-nucleosides **83**, **84**, and **85** (Scheme 22).

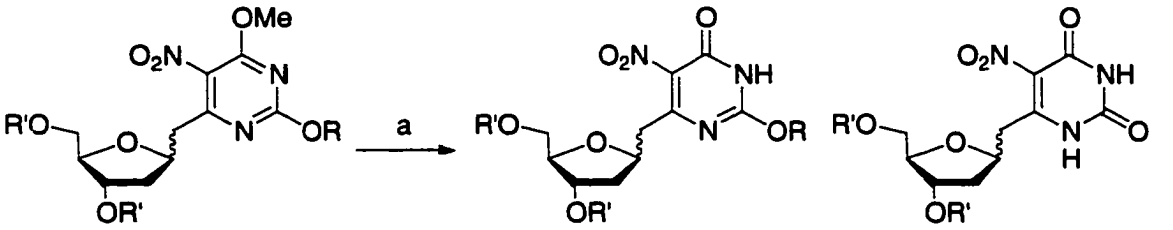
Scheme 22.[a]



[a]Key: a) i. toluene-reflux; ii. MeOH, catalytic NaOMe.

Investigations into the selective dealkylation of **83-85** initially focused on the use of TMSI for dealkylation. This method was successful in the synthesis of **64** (Scheme 15), however, since selective demethylation was not an issue large amounts TMSI was used to reduce reaction times. The use of less TMSI and lower reaction temperatures will be required to obtain selective dealkylation of **83-85**. Treatment of **83** with 1.2 equivalents of TMSI at 0 °C gave no selectivity for dealkylation upon 70% completion of the reaction (Table 12, entry 1). Furthermore, desilylation of **83** was also observed. For the synthesis of **64** (Scheme 15) we did not observe any desilylation using the TIPDS group. Furthermore, the inductive effect of the *p*-nitro group to C-2 facilitates the

Table 12. Optimization of dealkylation conditions.



83: R= *n*Pr, R'= TBDMS	**86**	**87**
84: R= *i*Pr, R'= TIPDS	**88**	**64**
85: R= *i*Pr, R'= TBDMS	**89**	**87**

a) Optimized conditons are for **85**: CH_3CN, TBDMSCl, NaI (67 % **89** + 27% **85**).

Entry	Silyl Chloride (eqv)[a]	R	R'	Conditions	Conv.	Selectivity[b]
1	TMSCl (1.2)	*n*Pr	TBDMS	0 °C, 0.1 M[c], 5 h	70%	1:1 + desilyl. **86**:**87**
2	TMSCl (1.2)	*i*Pr	TIPDS	0 °C, 0.1 M[c], 26 h	70%	2:1 + desilyl. **88**:**64**
3	$MeSiCl_3$ (1.2)	*i*Pr	TIPDS	-20 °C, 0.1 M[d], 22 h	10%	5:1 + desilyl. **88**:**64**
4	TBDMSCl (1.2)	*n*Pr	TBDMS	R.T., 0.1 M[d], 48 h	30%	1:0 + no desilyl. **86**
5	TBDMSCl (1.2)	*n*Pr	TBDMS	50 °C, 0.1 M[d], 1 h	100%	1:2 + no desilyl. **86**:**87**
6	TBDMSCl (1.2)	*n*Pr	TBDMS	R.T., 0.5 M[d], 18 h	25%	1:0 + no desilyl. **86**
7	TBDMSCl (1.5)	*i*Pr	TBDMS	R.T., 0.2 M[d], 44 h	70%	1:0 + no desilyl. **89**

[a]All reactions contain an equal amount of NaI.
[b]Selectivity refers to the ratio of product to the double deprotected material.
[c]Reactions are performed in $CHCl_3$.
[d]Reactions are performed in can

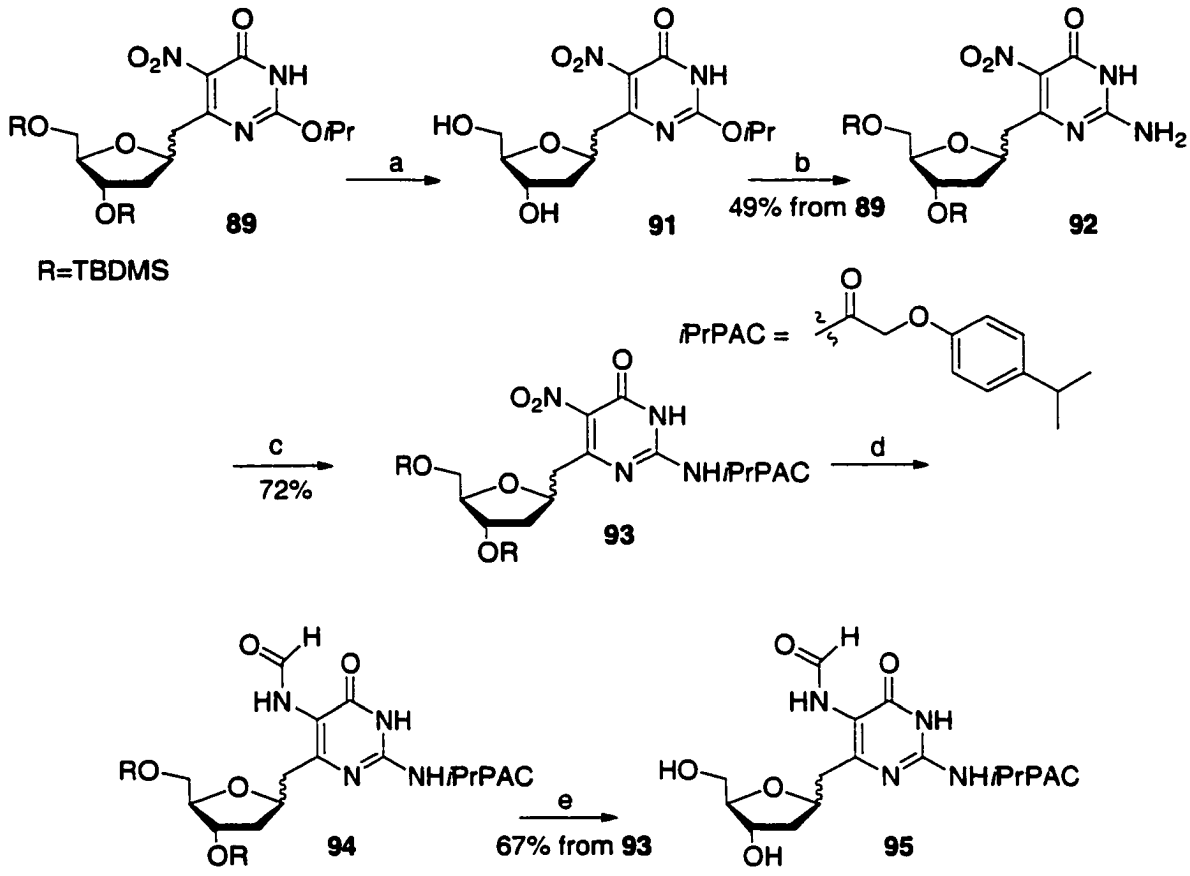
formation of the doubly deprotected material, reducing the selectivity for dealkylation. In order to minimize the desilylation during demethylation, TIPDS protected **84** was investigated. The C-2 alkyl group was also changed to isopropyl to improve the

selectivity of the dealkylation. Incorporation of the C-2-isopropyl ether improves the selectivity of **88**:**64** marginally to 2:1, but desilylation was still observed (Table 12, entry 2). The competing desilylation is presumed to occur from the long reaction times required to obtain the desired selectivity for dealkylation. In order to prevent the desilylation from occurring we used the milder Lewis acid CH_3SiCl_3. CH_3SiCl_3 works analogously to TMSCl, and imparts greater selectivity for the methyl ether versus more hindered alkyl-ethers.[201] Improved selectivity was observed but adventitious desilylation still occurred at lower reaction temperatures (Table 12, entry 3). Reducing the reaction temperature also decreases the rate of the reaction substantially. The undesirable desilylation forced us to consider other demethylation conditions. We rationalized that we could circumvent the problem of desilylation and improve the selectivity of the dealkylation of **83** and **85** with TBDMSI (Table 12, entries 4-7). TBDMSI eliminates the desilylation problem. The desired selectivity was also obtained but the reaction proceeds very slowly at room temperature (Table 12, entry 4). Attempts to improve the rate of the reaction focused on the concentration of reagents and the temperature of the reaction (Table 12, entries 5,6). At elevated temperatures the rate is increased, but the selectivity is diminished. The increased concentration of reagents improves the rate, but the conversion is still low after 18 h. The best conditions that were found involved reaction of **85** with 1.5 equivalents of TBDMSI for 44 h at room temperature. Only the desired product (**89**) was obtained. The reaction is still slow with 70% conversion, but unreacted **85** can be recovered and resubjected to the reaction conditions. Longer reaction times led to no improvement in conversion.

The selective demethylation of **85** allowed for the investigation of the aminolysis of **89**. Typical conditions required for aminolysis of alkyl ethers on 5-nitropyrimidine require the use of methanolic ammonia in a sealed tube at 100 °C.[157] When the leaving group is Cl (e.g. **90**), the reaction is complete in 1h. However, the bulky O*i*Pr leaving group is predicted to react much more slowly.[157,202] Initial attempts at the aminolysis of **89** resulted in desilylation in less than 2 h with no displacement of isopropoxide by ammonia. No reaction was observed using nucleophilic ammonia equivalents ($NaNH_2$, NH_2NH_2). Instead of investing large amounts of time trying to prevent desilylation, we decided to see if the nucleophilic displacement could be effected in the absence of silyl groups. Deprotection of **89** with NH_4F and subsequent treatment with methanolic ammonia at 90 °C for 43 h gave the desired product (**92**) along with unreacted diol (**91**, Scheme 23). The crude material was resilylated with TBDMSCl to produce a separable 3:1 mixture of **92**:**89**. Amidation of **92** with isopropyl phenoxyacetic acid and PyBOP produced **93** in good yields. The selection of the *i*PrPAC group instead of the PAC group was based upon the highly polar nature of **92**. Reduction of **93** with 10% Pd/C and subsequent formylation with acetic formic anhydride proceeded in high yields to produce **94.** The reduction was carried out in the presence of large amounts of catalyst (~50% w/w) to minimize the potential migration of the *i*PrPAC group to the C-5 amine.[67] Desilylation with Et_3N•3HF produced **95**. However, the polar nature of **95** presented difficulties in the purification from the Et_3N salts. Careful chromatography allowed for the majority of **95** to be separated from the Et_3N salts, but complete separation required several chromatographic steps.

90

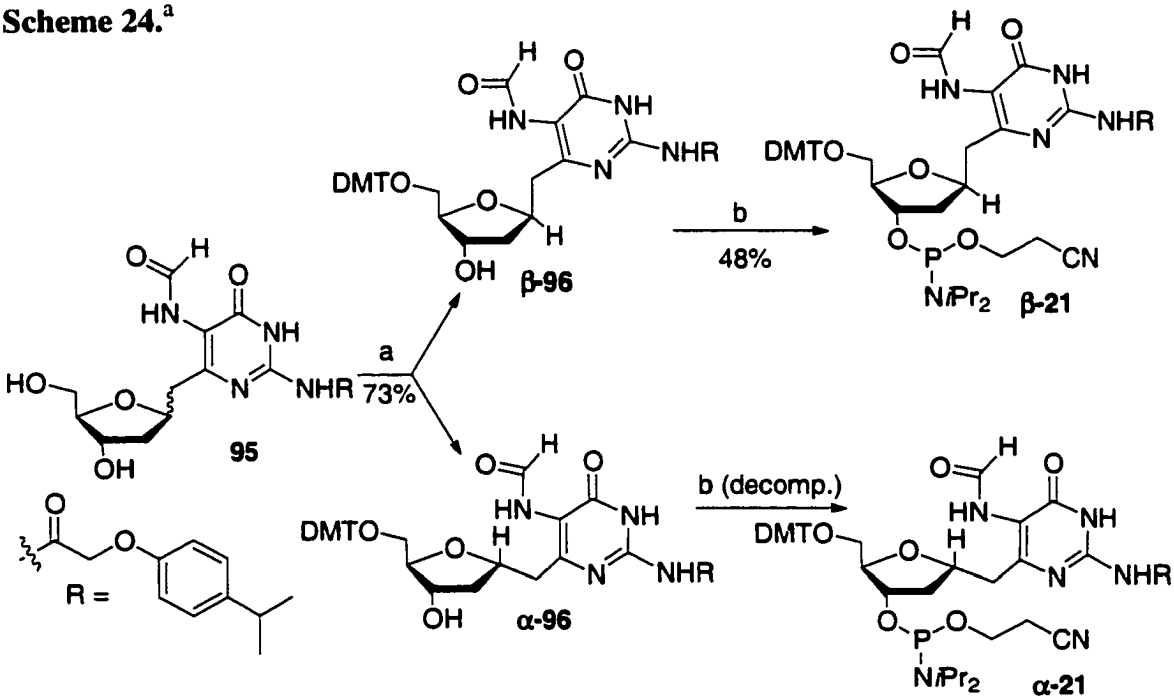
Scheme 23.[a]



[a]Key: a) NH_4F, MeOH; b) i. NH_3, MeOH, 90 °C, 43 h; ii. TBDMSCl, imidazole, DMF; c) PyBOP, *i*PrPAC-OH, iPr_2NEt; d) i. 10% Pd/C, 3:1 MeOH: THF, H_2; ii. AFA, pyridine, THF; e) $Et_3N{\bullet}3HF$, THF.

Reaction of **95** with dimethoxytrityl chloride in the prescence of catalytic DMAP produced the epimeric mixture of **96**, which were separated by column chromatography (Scheme 24). The stereochemistry at C_1' in **96** was assigned by ^1H NMR experiments. Irradiation of C_3' produced a strong NOE at C_1' in α-**96** (Figure 22). The respective irradiation in β-**96** did not produce an enhancement and even the defining C_4' enhancement was not observed upon saturating C_1'. The detection of the relaxation time for this process was not optimal using NOE experiments. Therefore, we performed a

Scheme 24.[a]



[a]Key: a) DMTCl, pyr, catalytic DMAP (3:1 β:α); b) 2-Cyanoethyl-*N,N,N',N'*-tetraisopropyl phosphane, diisopropylamino tetrazolide, DCM.

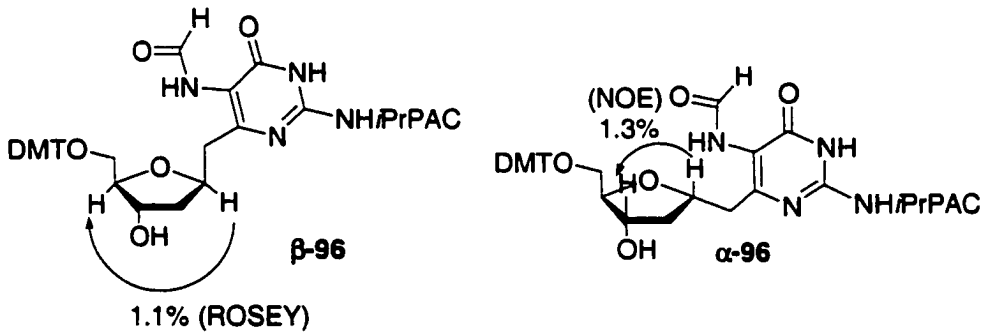


Figure 22. Determination of stereochemistry of **96** by NOE or ROSEY experiments.

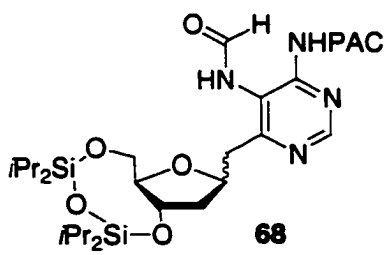
similar ^{1}H NMR experiment that was more sensitive to the relaxation event. ROSEY experiments were performed on β-**96** and clearly distinguished it as the β-anomer. Irradiation at C_1' produced a strong enhancement at the C_4' proton. Similar irradiation at C_4' also showed the enhancement at C_1'. Phosphitylation of β-**96** was accomplished in a similar fashion as **69**. Treatment of β-**96** with *N,N,N',N'*-tetraisopropylphosphane and

the *in situ* generated tetrazolide forms β-**21** in modest yields. Attempts to phosphitylate α-**96** were unsuccessful. Presumably, the C_3'-OH is significantly hindered by the α-configuration preventing phosphitylation. The steric congestion around the C_3'-OH in α-**96** is further compounded by the bulky N-2 isopropylphenoxyacetamide that prevents the approach of the phosphitylating agent. Extending the reaction times and the use of excess phosphitylating reagent leads to the decomposition of α-**96**.

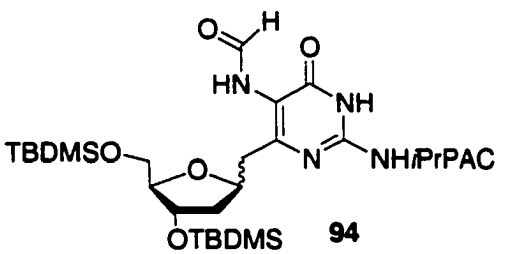
3.2 Oligonucleotide Synthesis and Characterization:

3.2.1 Stability to DNA Synthesis Conditions :

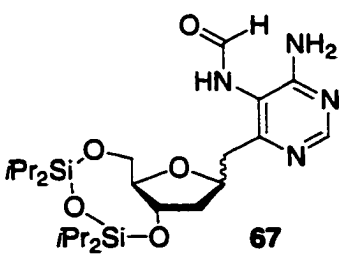
Prior to DNA synthesis of oligonucleotides with α,β-**20** and β-**21**, silylated monomers **68** and **94** were used as substrates to test the stability of the molecules to standard solid-phase oligodeoxynucleotide synthesis reagents and deprotection conditions. Monomers were stable to the oxidation and activator conditions used in



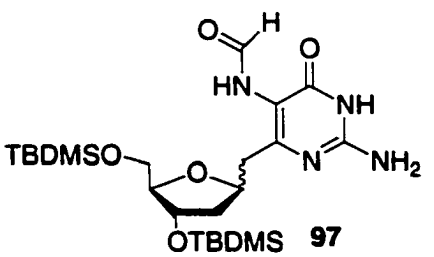
68



94



67

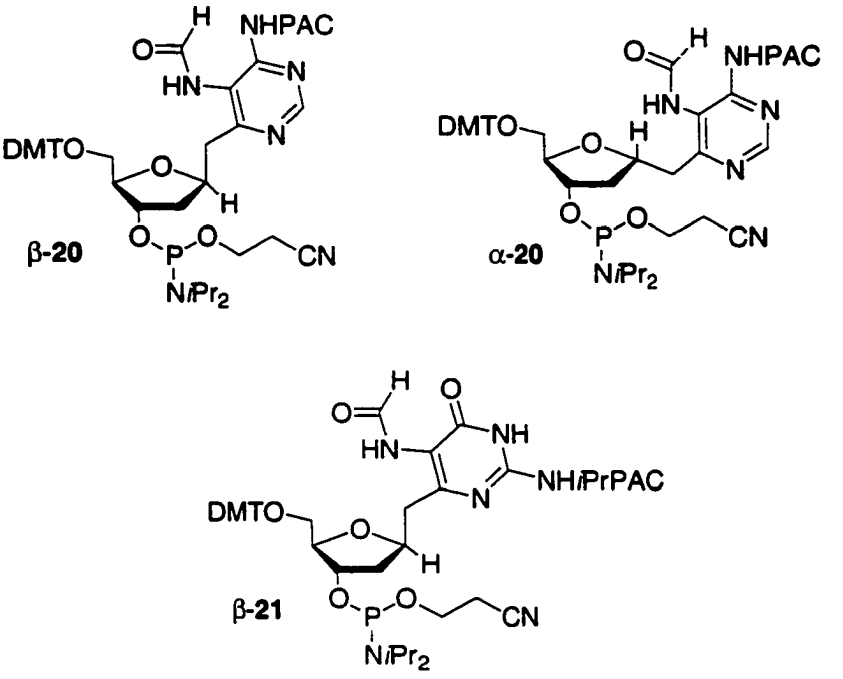


97

standard DNA synthesis. The monomers were not checked for stability to the deblocking conditions, as the silyl groups were not stable to acidic conditions. Due to the reactive nature of the *N*5-formamide, conventional capping conditions (Ac_2O, lutidine, *N*-methylimidazole (NMI)) were not suitable for the monomers. Transamidation products were observed in the presence of the NMI activator. Removal of NMI from the capping solution reduced the transamidation significantly, but trace amounts were still observed. Substitution of isobutyryl or pivaloyl anhydride for Ac_2O removed all traces of transamidation. Removal of the phenoxyacetamide groups of **68** and **94** was effected with K_2CO_3 (0.05 M in anhydrous MeOH) and occurred readily within 30 min at 25 °C to produce **67** and **97** without loss of the formamide. The formamide is stable for prolonged treatments of 5-6 h. Furthermore, no acyl migration was observed during the deprotection of **68** or **94**.

3.2.2 Modification of DNA Synthesis Cycles/Conditions:

The instability of **68** and **94** to conventional capping conditions was addressed by switching of the capping reagents to isobutyryl anhydride/lutidine and removal of NMI activator during and after coupling of the modified amidite. Introduction of the modified capping reagents later in the synthesis was done to preserve the overall capping efficiency as much as possible to facilitate purification. It was anticipated that longer coupling times would be required for the coupling α-**20**, β-**20**, and β-**21**. Initially, experiments extended the normal 25 s coupling time to 5 min. After coupling for 5 min, a second batch of activated phosphoramidite was allowed to react for 5 min ("double-couple") before continuing on with the synthesis cycle. A common problem with longer



coupling times is that the slightly acidic activator can cause slight detritylation during the coupling.[203] This event leads to branching from coupling at the deprotected site. To minimize this side-reaction, we employed the use of a less acidic activator than tetrazole, dicyanoimidazole (DCI) for coupling reactions.[204]

3.2.3 Synthesis of Oligonucleotides Containing α-18 and β-18:

Dodecameric oligonucleotides containing α-**18**, and β-**18** were prepared as described above. The limited amounts of α-**20** and β-**20** prevented optimization of the coupling step. In the synthesis of **98** and **99**, the yields for coupling α-**20** and β-**20** were

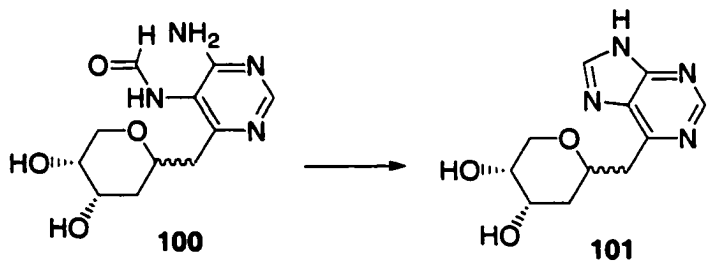
5'-d(GCT CTG **X**GT CGT)
98 a,b

5'-d(GCT CTC **X**CT CGT)
99 a,b

a, X = α-**18**; **b**, X = β-**18**

43 and 85%, respectively. The rationalization of the diminished yields of α-**20** was that the phosphorous center in the α-diastereomer was sterically encumbered. Admittedly the purity of the α-diastereomer may also have played a role in the diminished yield, as it was slightly contaminated with H-phosphonate. Isolation of **98** and **99** following deprotection (0.05 M K_2CO_3/MeOH) and gel purification afforded yields consistent with the trityl-responses. Precipitation of **98** and **99** and analysis by electrospray ionization mass-spectrometry (ESI-MS) revealed a minor impurity corresponding to dehydration product but no transacylated or deformylated product was observed in the dodecamers. Presumably, the dehydration occurs through the cyclization of the *N*4-amine and the formyl group. Examination of the ESI-MS capillary conditions revealed that the dehydration was an artifact of the spectrometer. Monomeric studies on **100** demonstrated that varying the cone voltage of the probe affects the relative distribution of the dehydrated (**101**) and uncyclized material (Scheme 25). By increasing the cone voltage

Scheme 25.



from 22 V to 42 V the ratio of **101** to **100** increases from ~1:7 to 1:1. Detection of ions below 22 V is not recommended due to an increase in noise.

Longer oligonucleotides **102** and **103** were also constructed with α,β-**20**. Oligonucleotides containing β-**18** were formed in modest yields. However, the

inefficient capping conditions, combined with modest phosphoramidite coupling yields, led to a mixture of product and oligonucleotides one nucleotide shorter that were difficult to separate by PAGE. In addition, ESI-MS analysis determined that transamidation was occurring during the synthesis of **102** and **103**. The decreased coupling efficiency of α-**20** prevented isolable amounts of **102b** and **103b** from being obtained.

3.2.3.1 Attempted Anion-Exchange HPLC Purification:

Purification of the crude mixture of **103b** by anion-exchange HPLC resulted in a broad peak that could not be cleanly resolved under a variety of different solvent gradients (100 mM Tris, pH 8: 0.5 M NH_4Cl). Fractionation of the peak into three portions and analysis by ESI-MS revealed that **103b** was present along with a mixture of impurities arising from transamidation and deformylation. The occurrence of transamidation was compounded by the instability of the isopropyl-PAC dG amidite to NMI/Ac_2O capping conditions used in the beginning stages of the oligonucleotide synthesis. After the synthesis of these oligonucleotides, our group demonstrated that the isopropyl-PAC dG amidite undergoes transamidation in the presence of the NMI activator.[205] To eliminate transamidation from occurring, pivaloyl anhydride was used in place of Ac_2O and the capping times were extended to 20 s. The use of these modified capping conditions can potentially reduce some of the transamidation that occurs, however, purification of the products would still be problematic.

3.2.3.2 Enzymatic Ligation Circumvents Purification Problems:

The instability of **18** in DNA to capping conditions limits the incorporation of

α-**20** and β-**20** into longer oligonucleotides due to difficulties in purification from n-1 sequences and transamidation products. The chemically synthesized dodecamers containing α-**18** and β-**18** provided a way of forming longer oligonucleotides containing **18** through enzymatic ligation. However, the use of enzymatic ligation was limited by the amount of material that could be obtained. Typically ligations are run on 0.1 nmol scales, but 1-5 nmol scale reactions have been reported with highly variable yields.[206-208] The feasibility of this approach depends upon the ability of the ligase enzyme to tolerate nucleotide modifications in the oligonucleotide. Enzymatic ligation can tolerate large modifications such as **1** and **2** when they are centered in the oligonucleotide.[207]

5'-d(AGG CGT TCA ACG)
104

5'-d(ACG TCC CAT GGT)
105

5'-d(TCA GAC ACC ATG GGA CGT ACG ACT CAG AGC CGT TGA ACG CCT)
106

5'-d(TCA GAC ACC ATG GGA CGT ACG AGT GAG AGC CGT TGA ACG CCT)
107

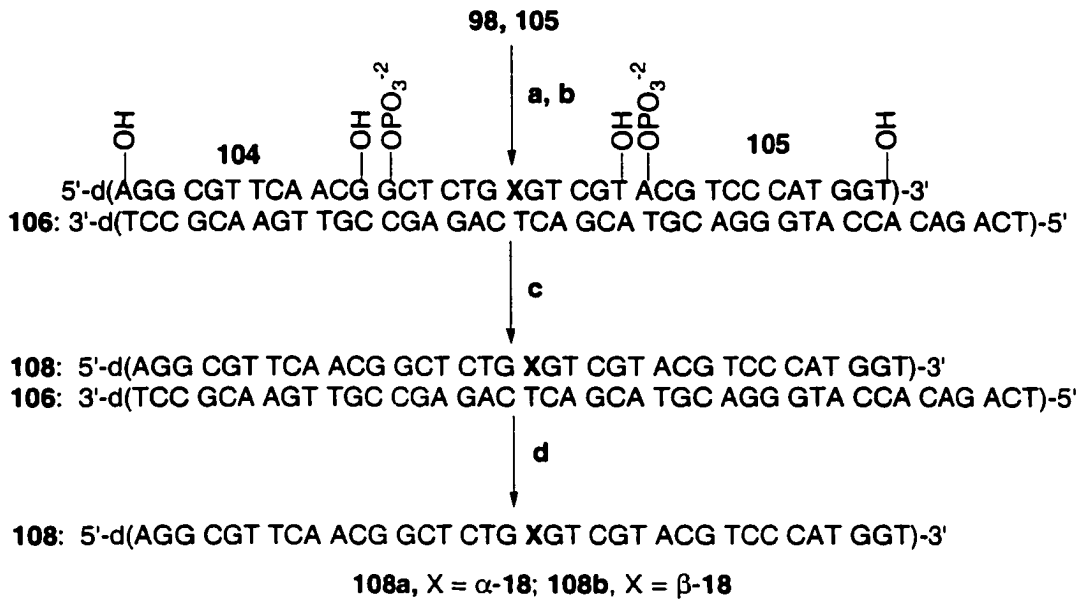
5'-d(AGG CGT TCA ACG GCT CTG **X**GT CGT ACG TCC CAT GGT)
108a,b

5'-d(AGG CGT TCA ACG GCT CTG **X**GT CGT ACG TCC CAT GGT)
109a,b

5'-d(AGG CGT TCA ACG GCT CTG **X**GT CGT)
110b

5'P-d(GCT CTG **X**GT CGT ACG TCC CAT GGT)
111b
a, X = α-**18**; **b**, X = β-**18**

Scheme 26.[a]



[a]Key: a) T4 polynucleotide kinase, ATP; b) hybridization; c) T4 DNA ligase; d) PAGE (50-60%).

Oligonucleotides **98** and **99** meet this criteria and were investigated for their potential in enzymatic ligations (Scheme 26). Initial experiments were performed on analytical scales using radiolabeled substrates to work out reaction conditions. Phosphorylation of **98b** and the 3'-terminal dodecamer (**105**) was accomplished with T4-kinase. The phosphorylated products were annealed to the template strand (**106**-2 equivalents) in the presence of the 5-terminal dodecamers (**104**) in a 1:1:1 ratio of dodecamers. The use of a longer template strand was implemented to facilitate PAGE purification of the ligated product from the template strand on preparative scales. The ligation was performed by adding T4-DNA ligase in 4-400 unit aliquots at 3, 6, 9, and 18 h. The experiment was performed along side a control (**98**; X= A). The progression of the ligation was monitored at 3, 6, 9, 18, and 24 h and the products were separated by PAGE (Figure 23). After 3 h, only monoligated (24 mer) and fully ligated material (36 mer) are present.

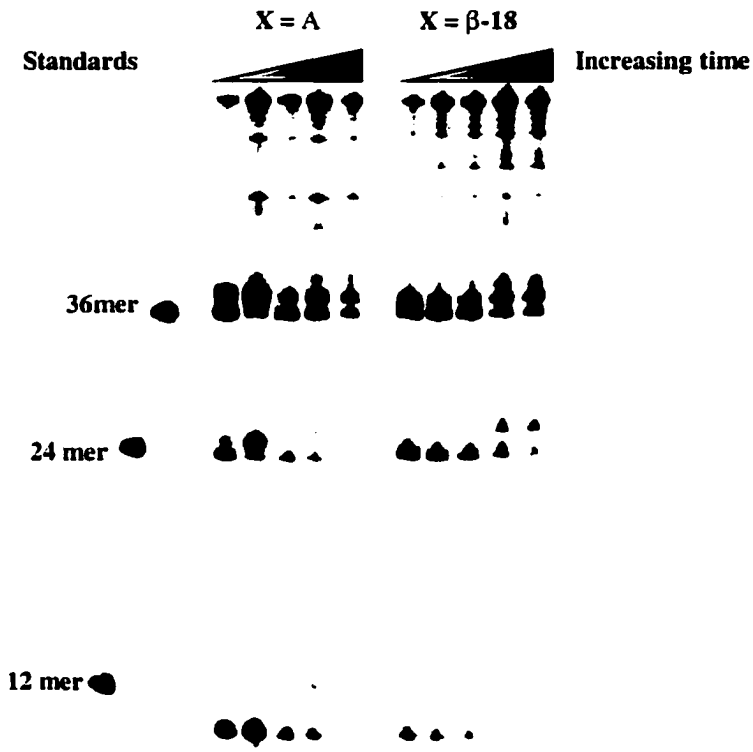


Figure 23. Enzymatic ligation of **98b** with T4 DNA ligase.

Complete conversion of the monoligated material to the product occurs after 24 h. Smearing of the product bands was attributed to the use of large amounts of enzyme relative to the DNA in this experiment and the difficulty in dehybridizing the product from the template. Preparative scale ligations were performed in a similar fashion, where 5 nmol of each of the dodecamers was employed for the ligation. Purification of the reaction by 15% denaturing PAGE provides **108b** in a 30% isolated yield that was still annealed with the template. The product appears to be renaturing prior to separation on the PAGE. Monoligated products **110b** and **111b** were also isolated as a mixture in 30% from the ligation. Incomplete denaturation was minimized by running the gels at elevated temperatures (between 50 and 60°C) to allow for the separation of product from the template prior to renaturation. The 1:1 mixture of product to monoligated material was overcome by using a 2-fold excess of the terminal dodecamers relative to **98b**.

These modifications provide **104b** in as high as 98% yield, but yields routinely are between 50-60%. Application of this methodology to **98a, 99a,b** allows for the formation of **108a**, **109a**, and **109b** in 53-66% yields.

3.2.4 Synthesis of Oligonucleotides Containing β-19:

The synthesis of oligonucleotides containing α-**18** and β-**18** uncovered several improvements in the DNA synthesis conditions that could be made to enhance the yield and purification of oligonucleotides containing β-**19**. Capping conditions prior to coupling of β-**21** were changed to pivalic anhydride, lutidine, and NMI to prevent transamidation at isopropyl PAC dG.[205] NMI was removed from the capping mixture during the coupling of β-**21** and all subsequent synthesis cycles. Coupling of β-**21** was performed with one batch of activated amidite for 15 min ("single-coupled") with tetrazole as the activator instead of DCI. Synthesis of 36 mers (**112** and **113**) containing

5'-d(AGG CGT TCA ACG TGC AGT **X**AC AGC ACG TCC CAT GGT)
112

5'-d(AGG CGT TCA ACG TGC AGT **X**TC AGC ACG TCC CAT GGT)
113

5'-d(TGC AGT **X**AC AGC)
114
X= β-**19**

19 was successful with an overall yield of 47%. The overall yields for the synthesis of the 36 mer was improved using tetrazole as the activator over prior results with α-**20** and β-**20** using DCI as the activator. The individual coupling yields of β-**21** was 80% and was comparable to the coupling yield of β-**20** (85%), despite using only a single

coupling. Deprotection of 36 mers for 5 h with K_2CO_3 (0.05 M in anhydrous MeOH) and purification by PAGE provided the desired materials which were characterized by ESI-MS. Synthesis of dodecamer **114** was also successful using tetrazole as the activator.

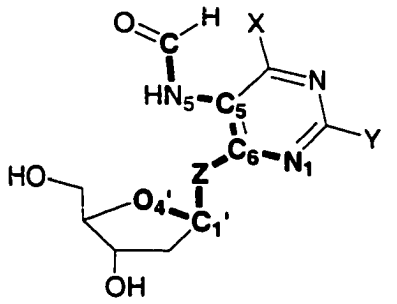
3.3 Molecular Modeling of Fapy Lesions and C-Nucleoside Analogues:

Molecular modeling studies were employed to determine the structural similarity of the C-nucleoside analogues **18** and **19** with Fapy•dA and Fapy•dG, as well as the similarities with dA and dG. The opening of the imidazole ring in the native purine allows for more rotational degrees of freedom in Fapy•dA and Fapy•dG and changes the hybridization of the glycosidic nitrogen from sp^2 to sp^3. Substitution of the glycosidic NH by a methylene group in **18** and **19** not only removes a potential hydrogen bond donor/acceptor from the analogue, but introduces two CH bonds that can affect the rotameric preferences of the formamide. This change may also bias the *anti–syn* conformational equilibrium and thereby directly relate to the potential hydrogen bonding patterns presented. Molecular modeling can offer insight into these conformational issues of the analogue/lesion and suggest valuable information as to the potential mutagenicity of Fapy lesions.

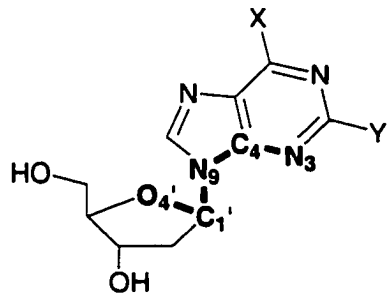
3.3.1 Molecular Modeling of β-Fapy•dA and β-18:

Molecular modeling was performed on monomeric structures of dA, α-Fapy•dA, β-Fapy•dA, α-**18**, and β-**18**. These structures were minimized using the Cerius2 software package and the glycosidic torsional angle, nucleobase tilt, and formamide torsional angles were calculated (Scheme 27). The glycosidic torsional angle was defined as the

Scheme 27.



Fapy•dA: **X** = NH_2, **Y**= H, **Z**= NH
α or β-**18**: **X** = NH_2, **Y**= H, **Z**= CH_2
Fapy•dG: **X** = OH, **Y**= NH_2, **Z**= NH
β-**19**: **X** = OH, **Y**= NH_2, **Z**= CH_2



dA: **X** = NH_2, Y= H
dG: **X** = OH, Y = NH_2

torsional angle between the bonds of O_4'-C_1' and Z-C6 for Fapy•dA/**18** (bold bonds in Scheme 27). The comparable torsional angle was calculated for dA between the bonds of O_4'-C_1' and N9-C4. The nucleobase tilt was measured from the torsional angle created between the C_1'-Z and C6-N1 bonds for Fapy•dA/**18** and dA was measured from the C_1'-N9 and C4-N3 bonds. Determination of the formamide torsional angle in Fapy•dA and **18** was accomplished through measurement of the angle between the C6-C5 and N5-C(O)H bonds.

Comparisons of the predicted structures of dA, Fapy•dA, and β-**18**, reveal similarities between the nucleosides, as well as the key structural differences between these molecules (Figure 24). The glycosidic torsional angle varies to a small extent between these molecules. The angles calculated for β-Fapy•dA and β-**18** are within 4° of that in dA. Ring-opening offers more rotational freedom for β-Fapy•dA and β-**18**, and allows for tilting of the nucleobase relative to dA. Measurement of the tilting as described above shows that Fapy•dA is tilted by 18° relative to dA. Substitution of the methylene group in β-**18** allows increases the tilt angle to 64° relative to dA.

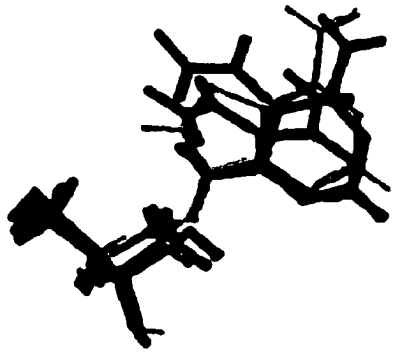


Figure 24. Overlay of predicted structures of dA, Fapy•dA, and β-**18** by molecular modeling. Yellow: dA; Blue: Fapy•dA; Red: β-**18**

β-Fapy•dA and β-**18** prefer to adopt the anti-conformation with the formamide in the *trans*-orientation (Figure 25). The formamide is rotated ~60° out of the plane of the pyrimidine for β-Fapy•dA. The adjacent C-H's in β-**18** force this angle to increase to 114°. Rotation of the formamide from the *trans*- to the *cis*-orientation and minimization

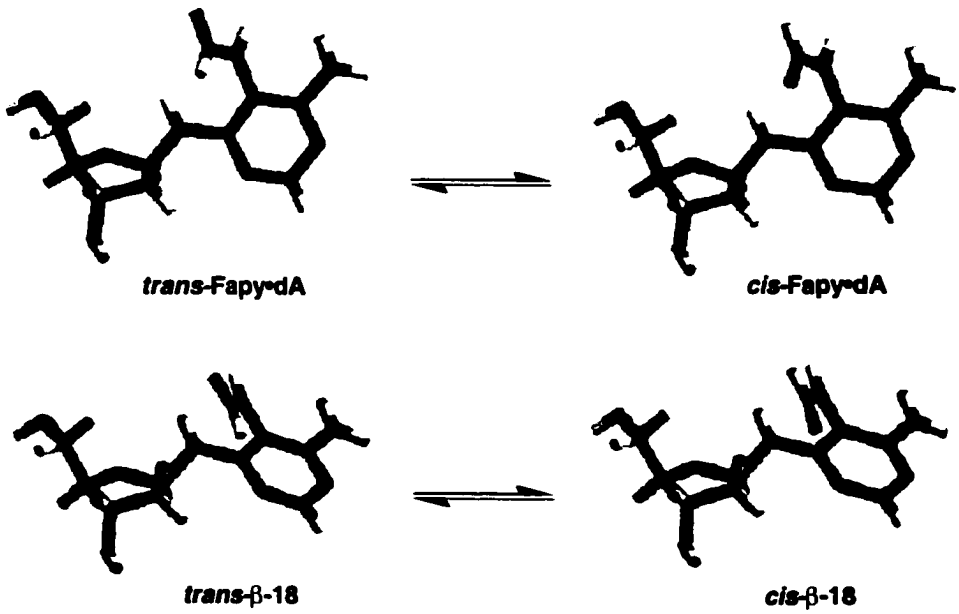


Figure 25 Molecular modeling of Fapy•dA and β-**18**

reveals that the *cis*-conformation is disfavored by 1.1 kcal/mol for β-Fapy•dA and **β-18**. In contrast to this, NMR studies on pyranosyl Fapy•dA by Cadet show that the *cis*-amide is the preferred conformation by 1.0 kcal/mol relative to the *trans*-amide.[70] Our observed 1.1 kcal/mol energy difference between the *trans*- and the *cis*-conformations represents a small energetic difference between these conformations as compared with common interactions that occur in DNA. It is believed that the structural differences between **β-18** and β-Fapy•dA in the monomer can be overcome in DNA. Stabilizing features such as base stacking and hydrogen binding that occur in the duplex could compensate for these small differences. Preliminary modeling studies of the molecules in a duplex were investigated and suggest that **β-18** is structurally similar to β-Fapy•dA in DNA.

The energetic differences between the *syn*- and *anti*-orientations were also determined for β-Fapy•dA and **β-18**. Minimization in the *syn*-conformation produces two minima that differ in rotation of the *trans*-formamide (Figure 26). In β-Fapy•dA these conformations are 0.7 and 0.9 kcal/mol higher in energy than the *anti*-conformation. The differences in energy between comparable conformational isomers are magnified in **β-18** (1.7 and 2.1 kcal/mol). The additional CH bond on the glycosidic linkage restricts the adoption of the *syn*-conformation relative to β-Fapy•dA. The presence of an additional C-H bond increases the energy required to adopt the syn conformation since more gauche and eclipsing interactions are present for **β-18** relative to β-Fapy•dA. However, the differences in energies of the transitions states for β-Fapy•dA and **β-18** would have to be very large to expect any significant discrimination between the two conformations. Examination of the *cis*-amide *syn*-conformations reveal that these

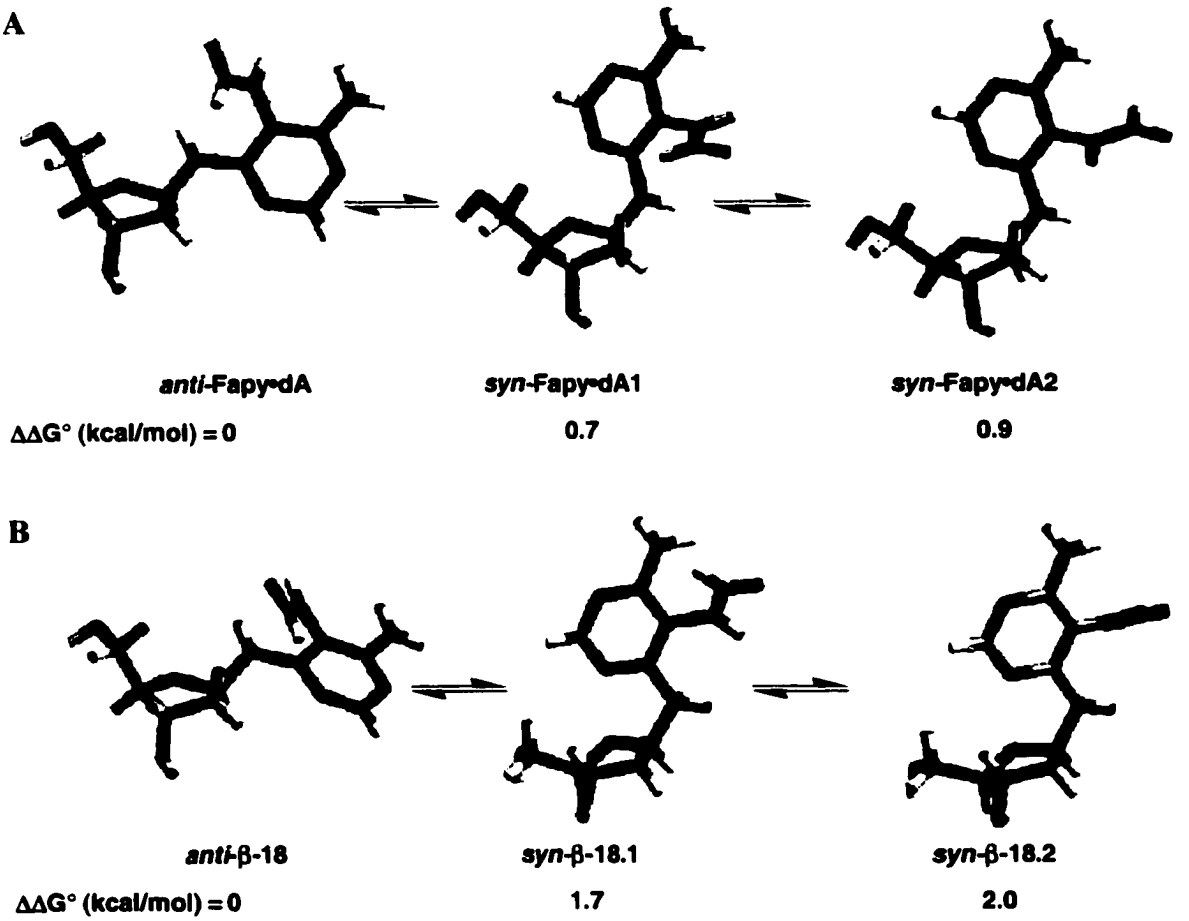


Figure 26. Molecular modeling of *anti* and *syn* conformations of Fapy•dA (A) and β-**18** (B).

isomers are destabilized significantly more than the *trans*-amides in β-Fapy•dA and β-**18** (1.9 and 3.3 kcal/mol, respectively).

3.3.2 Molecular Modeling of α-Fapy•dA and α-18:

Investigations into the structures of α-dA, α-Fapy•dA, and α-**18** reveal similar features that were observed in the β-anomer series. A large perturbation of the tilt angle was observed for α-Fapy•dA and α-**18** (17.6° and 62.5° relative to dA, Figure 27). An

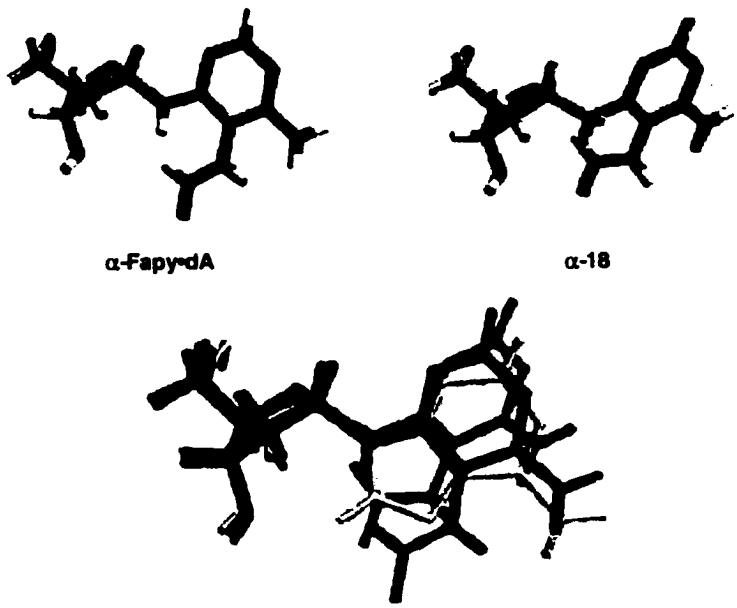


Figure 27. Overlay of predicted structures of α-dA, α-Fapy•dA, and β-**18** by molecular modeling. Yellow: dA; Blue: Fapy•dA; Red: β-**18**

interesting consequence of the α-configuration is that the glycosidic torsional angle in these molecules are approximately 70° greater than the respective β-compounds. This could have consequences on the ability of these compounds to hydrogen bond. Further investigations into the *syn*-conformations of the series of α-compounds were not performed, as biochemical results on α-**18** did not prove to be interesting.

3.3.3 Molecular Modeling of β-Fapy•dG and β-19:

Investigations on the structural similarities between β-Fapy•dG and β-**19** were performed in a similar fashion as β-Fapy•dA and β-**18**. Analysis of the *anti*-conformation of dG, β-Fapy•dG, and β-**19** reveal that these compounds exhibit similar features as β-Fapy•dA and β-**18**. The glycosidic torsional angles of β-Fapy•dG and β-**19** are within 6° of dG (Figure 28). The tilting of the nucleobase is very similar for β-

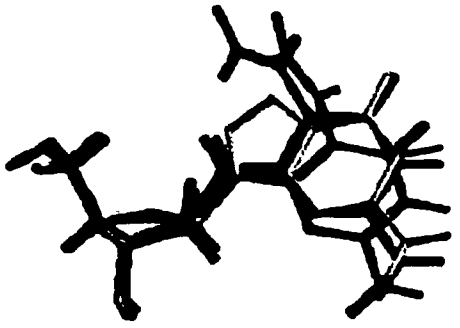


Figure 28. Overlay of predicted structures of dG, Fapy•dG, and β-**19** by molecular modeling. Yellow: dG; Blue: Fapy•dA; Red: β-**19**

Fapy•dG (19.4°) and β-Fapy•dA (18.2°), but the analogous measurement for β-**19** is very different than β-**18** (-1.1° versus 64.2°). This angle in β-**19** is very similar to the one predicted in dG (-0.2°, Figure 28). These studies demonstrate that β-**19** is structurally similar to Fapy•dG and should serve as a suitable model in DNA.

As in the Fapy•dA modeling, the *trans*-amide is favored thermodynamically in β-**19** and β-Fapy•dG. Rotation of the C-N bond into the *cis*-formamide conformation results in modest destabilizations of 1.3-2.3 kcal/mol for both the *syn*- and *anti*-conformations of β-**19** and β-Fapy•dG. The *cis*-amide is preferred for Fapy•dG in modeling studies carried out by Carrel (Figure 29).[69] The basis for favoring the *cis*-conformation is that a presumed hydrogen bond between the N6-NH and the formamide

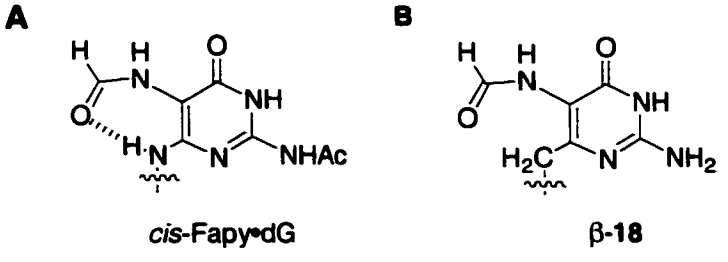


Figure 29.[69] Proposed intramolecular hydrogen bonding in *cis*-Fapy•dG. A) Hydrogen bonding of the *cis*-formamide with the N6-H; B) Analogus hydrogen bonding pattern is not possible in β-**18**.

carbonyl stabilizes this isomer. However, our modeling studies suggest that adoption of the *cis*- amide is energetically disfavored and minimized structures prefer the *trans*-amide for both Fapy•dA and Fapy•dG. In addition, if the formation of a hydrogen bond between the N6-H and the formamide was important in our modeling studies for Fapy•dA and Fapy•dG, a larger divergence in the molecular modeling results of the C-nucleotide analogues would have been predicted. Substitution of the N6-H for CH_2 prohibits the formation of a hydrogen bond with the formamide preventing stabilization of the *cis*-formamide (Figure 29).

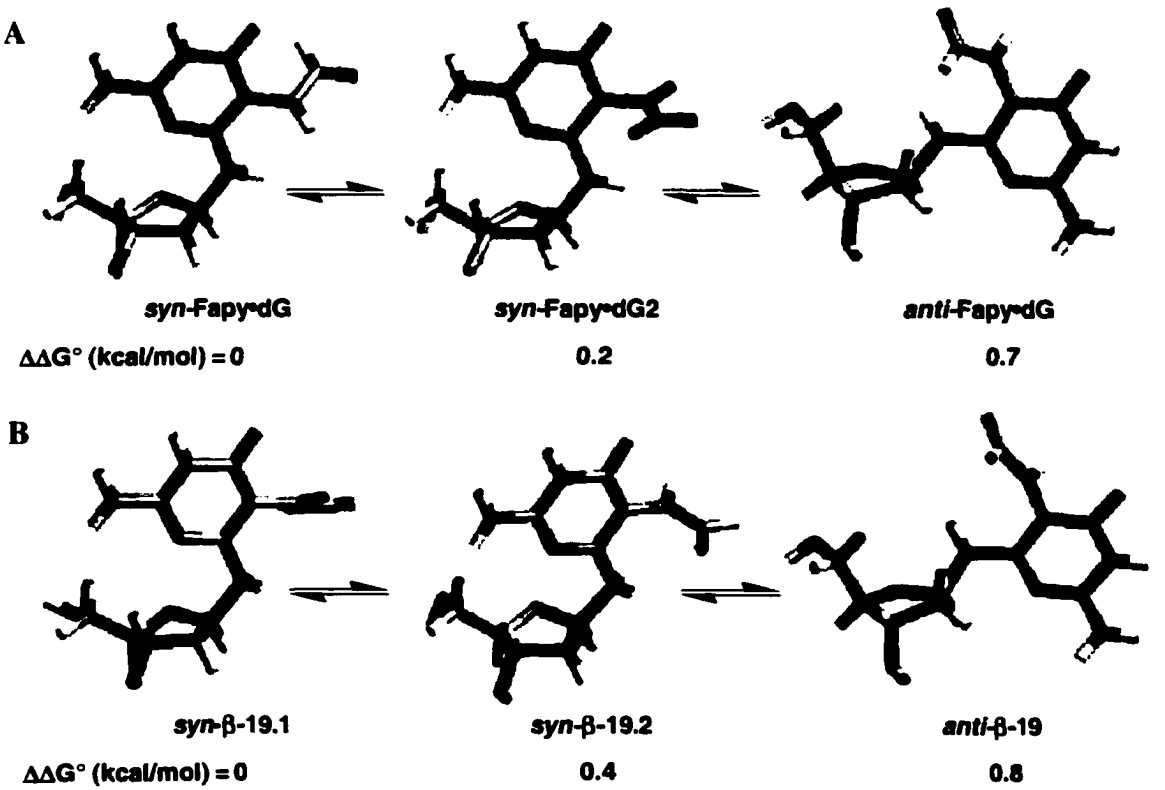


Figure 30. Molecular modeling of *anti* and *syn* conformations of β-Fapy•dG (A) and β-19(B).

Energy minimized structures for Fapy•dG and β-**19** reveal a slight preference for these compounds to adopt the *syn*-conformation (Figure 30). The minimum energy structure of the *anti*–configurational isomer is within 0.7 kcal/mol for β-Fapy•dG and 0.8 kcal/mol for β-**19**. Furthermore, the calculated energy differences between different rotameric conformations of *syn*-β-Fapy•dG (0.2 kcal/mol) and *syn*-β-**19** (0.4 kcal/mol) are also relatively small. These results suggest that many more potential hydrogen bonding patterns are energetically possible provided that the barriers for rotation are also small.

The overall conclusions that can be drawn from the molecular modeling calculations on Fapy•dA / **18**, and Fapy•dG/ **19** are that the predicted structures for the C-nucleoside analogues are very similar to the Fapy lesions that they were designed to mimic. However, the C-nucleoside analogue of FapydG (**19**) is predicted to have a smaller tilt of the pyrimidine ring relative to that of the lesion or dG. Comparisons of the minimized energies of the preferred *syn*- and *anti*-conformations of Fapy•dA / **18**, and Fapy•dG/ **19** are all negligible when compared to the energy contribution due to temperature. A close approximation of this contribution is approximately 0.6 kcal/mol (Boltzmann contribution—k_BT).[209] The predicted energy differences between the different *syn*- and *anti*-conformations of Fapy•dA / **18** were at most 1-2 kcal/mol and the differences for Fapy•dG /**19** being much lower in energy. With the Boltzmann contribution being very near to these differences in energy, the probability that the higher energy conformations will be populated is likely. As a result, the *syn*- and *anti*-conformations that are higher in energy cannot be excluded from hydrogen-bonding and multiple conformations are accessible for hydrogen bonding.

3.4 Biochemical and Physiochemical Characterization of Oligonucleotides Containing α-18, β-18, and β-19:

3.4.1 Thermodynamic Melting Studies of Oligonucleotides Containing α-18, β-18, and β-19:

The stability of duplexes containing C-nucleotide analogues can be tested for their usefulness as structural probes of Fapy lesions through UV-melting studies. Preferences for α-**18**, β-**18**, and β-**19** to base pair to the four native nucleotides may offer important information about Fapy lesions. If a particular anomer of **18** shows comparable stabilizing interactions to Fapy•dA, the potential biological relevance of the β- or α-anomers of Fapy•dA can be assigned.

3.4.1.1 Thermodynamic Melting (*T*m) Studies of Oligonucleotides Containing α-18 and β-18:

The effects of α-**18** and β-**18** on duplex stability were determined from van't Hoff plots of UV-melting using dodecamers **98** and **99**. Melting temperatures of duplexes **115a,b** and **116a,b** were compared to duplexes of native nucleotides (**115c** and **116c**). The observed T_m for duplexes containing α-**18**:T and β-**18**:T are reduced significantly relative to the native duplex (~10°C, Table 13). Mismatches opposite α-**18** or β-**18** also destabilize the duplex, but only slightly more than the α-**18**:T or β-**18**:T base-pair. This suggests that there is no preference for base pairing. The destabilization of mismatches opposite dA was much larger than that observed for the C-nucleotides. Thermodynamic melting studies on Fapy•dA do not produce a similar trend as duplexes containing **18** (Table 14).[104] The Fapy•dA:dA and Fapy•dA:dG base pairs are stabilized more than the

Table 13. UV-Melting thermodynamics of duplexes containing α-**18**, β-**18**, and dA.[a]

5'-d(GCT CTG **X**GT CGT)
3'-d(CGA GAC **Y**CA GCA)
115a,b,c
a(**X**=α-**18**),**b**(**X**=β-**18**), **c**(**X**=dA)

5'-d(GCT CTC **X**CT CGT)
3'-d(CGA GAG **Y**GA GCA)
116a,b,c
a(**X**=α-**18**),**b**(**X**=β-**18**), **c**(**X**=dA)

X:Y	T_m[b] (°C)	$\Delta G°_{298}$ (kcal/mol)	X:Y	T_m[b] (°C)	$\Delta G°_{298}$ (kcal/mol)
α-**18**:T	46.8 ± 0.1	13.7	α-**18**:T	44.6 ± 0.5	12.9
α-**18**:dA	46.4 ± 0.5	12.6	α-**18**:dA	40.8 ± 0.1	11.3
α-**18**:dG	42.9 ± 0.9	12.1	α-**18**:dG	40.9 ± 0.1	11.3
α-**18**:dC	45.8 ± 0.8	13.1	α-**18**:dC	36.5 ± 0.5	10.5
β-**18**:T	46.8 ± 0.1	13.5	β-**18**:T	43.4 ± 0.6	12.9
β-**18**:dA	45.9 ± 0.1	13.5	β-**18**:dA	40.4 ± 0.5	12.0
β-**18**:dG	41.8 ± 0.1	12.4	β-**18**:dG	41.9 ± 0.1	12.8
β-**18**:dC	46.6 ± 0.5	13.2	β-**18**:dC	33.5 ± 0.5	10.0
dA:T	56.3 ± 1.0	17.0	dA:T	56.1 ± 1.3	17.2
dA:dA	46.2 ± 0.5	13.2	dA:dA	44.9 ± 0.1	13.1
dA:dG	47.8 ± 0.1	13.6	dA:dG	52.0 ± 0.1	15.6
dA:dC	44.6 ± 0.6	13.6	dA:dC	39.9 ± 0.1	11.5

[a]Conditions: 10 mM PIPES (pH 7.0); 10 mM $MgCl_2$; 100 mM NaCl.
[b]Total oligonucleotide concentration = 4.4μM.

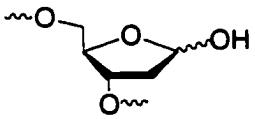
Table 14. [104] UV-Melting thermodynamics of duplexes containing Fapy•dA and dA.[a]

5'-d(TGC ACT **X**AC AGC)
3'-d(ACG TGA **Y**GT TCG)
117a,b
a(**X**=Fapy•dA), **b**(**X**=dA)

X:Y	T_m[b] (°C)	$\Delta G°_{298}$ (kcal/mol)	X:Y	T_m[b] (°C)	$\Delta G°_{298}$ (kcal/mol)
Fapy•dA:T	47.1± 0.4	12.2	dA:T	54.2 ± 0.4	13.9
Fapy•dA:dA	40.4 ± 0.7	10.5	dA:dA	41.6 ± 0.6	9.7
Fapy•dA:dG	42.0 ± 0.6	12.2	dA:dG	45.4 ± 0.5	11.8
Fapy•dA:dC	38.3 ± 0.6	8.9	dA:dC	40.9 ± 0.6	9.6

[a]Conditions: 10 mM PIPES (pH 7.0); 10 mM $MgCl_2$; 100 mM NaCl.
[b]Total oligonucleotide concentration = 4.4μM.

analogous dA:dA and dA:dG base pairs in duplex **117**. The differences in the thermodynamic stabilities of **115** /**116** and **117** suggest that α-**18** and β-**18** are behaving in a different fashion than Fapy•dA in the duplex. A potential explanation for the small discrimination between base pairing partners of α-**18** and β-**18** is that the C-nucleotides are extrahelical and do not form specific base pairs. This could also explain the remarkable similarities between the structurally different α-**18** and β-**18**. The apparent lack of base-pairing specificity by α-**18** and β-**18** is very similar to observations made using an abasic site model.[210]



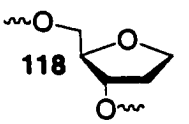
Abasic site

Solution structures of duplexes containing abasic sites have been obtained through NMR analysis.[211] In these studies, the stereochemistry of the C_1'-OH was important in determining the conformation that the abasic site adopts in the duplex. It is hypothesized that the β-configuration of the abasic site participates in stabilizing interactions with the opposite base through hydrogen bonding with a water molecule and an intrahelical conformation is favored. This same interaction is not available for the α-configuration and the abasic site prefers an extrahelical orientation. The stabilizing interactions of hydrogen bonding are also not available for the tetrahydrofuran abasic site model (**118**). The orientation of **118** as intra- or extrahelical depends largely upon sequence context.[212,213]

In thermodynamic denaturation studies on oligonucleotides containing **118**, no base pair is significantly stabilized opposite **118**.[210] The ΔΔG° are within 1.3 kcal/mol of each other (Table 15). Comparison of the duplexes of α-**18** and β-**18** (**115**) reported above (Table 13) reveals ΔΔG° values that are within 1.6 and 1.1 kcal/mol, respectively

Table 15. Thermodynamic comparisons between α-**18**, β-**18**, and **115** in duplex DNA.

5'-d(GCT CTG XGT CGT)
3'-d(CGA GAC YCA GCA)
115a,b,c
a(X=α-**18**),**b**(X=β-**18**), **c**(X=dA)



X:Y	$\Delta\Delta G°_{298}$[a]	X:Y	$\Delta\Delta G°_{298}$[a]	X:Y[b]	$\Delta\Delta G°_{298}$[c]
α-**18**:T	--	β-**18**:dT	--	**118**:T	-0.9
α-**18**:dA	-1.1	β-**18**:dA	0.0	**118**:dA	--
α-**18**:dG	-1.6	β-**18**:dG	-1.1	**118**:dG	-1.3
α-**18**:dC	-0.6	β-**18**:dC	-0.2	**118**:dC	-0.6

[a]$\Delta\Delta G°$ = $\Delta G°$ (**18**:T) - $\Delta G°$ (**18**:A,G, or C).
[b]Duplexes containing **118** are from a different sequence context where **118** is flanked by a 3'- and 5'- G.
[c]$\Delta\Delta G°$ = $\Delta G°$ (**118**:A) - $\Delta G°$ (**118**:A,G, or C).

(Table 15). The similarity in the lack of base pairing preferences between **118**, α-**18**, and β-**18** suggests that the C-nucleotides are extrahelical. In order for **18** to be extrahelical, a large distortion in the helix would be necessary. To test this hypothesis, CD spectroscopy on the duplexes **115a-c** were performed (Figure 31). Under normal salt conditions (100 mM NaCl) the CD spectra do not substantiate the observations from the UV melting studies. There is very little difference in the helicity of the native A:T duplex compared to the α-**18**:T and β-**18**:T duplex. Changing to low salt conditions (10 mM) further destabilizes the duplexes and reveals differences between the helicity of the A:T duplexes and those containing **18**. The characteristic relative maxima at 280 nm is diminished significantly for the α-**18**:T, and β-**18**:T duplexes. However, comparison of low-salt conditions with the thermodynamic melt data is not valid due to the large differences in salt conditions with the melting experiments.

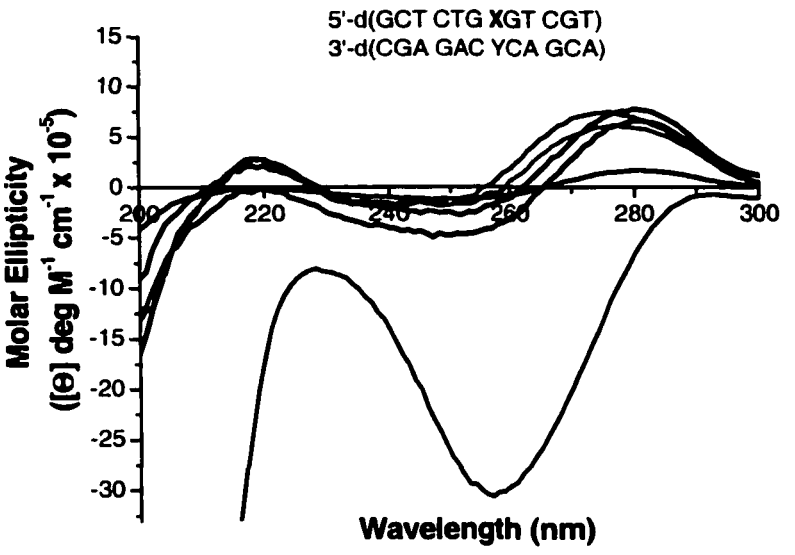


Figure 31. CD spectrometry on duplexes containing α-**18**:T and β-**18**:T under different NaCl concentrations. 100 mM NaCl, 10 mM Phosphate (pH 7.5): Blue (X=A), Aqua (X=α-**18**), Pink (X=β-**18**). 10 mM NaCl, 10 mM Phosphate (pH 7.5): Black (X=A), Red (X=α-**18**), Green (X=β-**18**).

3.4.1.2 Thermodynamic Melting (*Tm*) studies of Oligonucleotides containing β-19:

Behavior of β-**19** in oligonucleotides is very similar to α-**18** and β-**18**. The thermal stability of the β-**19**:dC base pair is reduced by ~11 °C relative to the native dG:dC base pair (Table 16). Furthermore, no base pairing preferences were observed other than β-**19**:dC. The energetic differences between the β-**19**:dC base pair and mismatches are more pronounced than in α-**18** or β-**18** ($\Delta\Delta G°$ = 2.2-3.3 kcal/mol). This suggests that the β-**19**:C base pair is instructive in base pairing relative to the mismatches. However, thermodynamic melting studies on Fapy•dG reveal a different preference towards mismatches.[103] Mismatches of dA and dG opposite Fapy•dG are stabilized in the duplex relative to the same mismatches in the native nucleotide. This is not the case for β-**19** (Table 16).

Table 16.[103] UV-Melting Thermodynamics of duplexes containing β-**19** and Fapy•dG compared to dG.[a]

5'-d(TGC ACT **X**AC AGC)
3'-d(ACG TGA **Y**GT TCG)
119a,b,c
a(**X**=β-**19**), **b**(**X**=Fapy•dG), **c**(**X**=dG)

X:Y	T_m[a] (°C)	$\Delta G°_{298}$ (kcal/mol)
β-**19**:dC	46.1 ± 0.3	13.0
β-**19**:dA	40.0 ± 0.4	10.6
β-**19**:dG	37.7 ± 0.6	10.5
β-**19**:T	39.8 ± 0.6	9.7
Fapy•dG:dC	54.1 ± 0.1	15.1
Fapy•dG:dA	51.7 ± 0.6	14.8
Fapy•dG:dG	45.1 ± 0.5	13.2
Fapy•dG:T	44.5 ± 0.3	12.3
dG:dC	57.4 ± 0.4	15.3
dG:dA	44.6 ± 0.5	12.9
dG:dG	46.3 ± 0.3	11.3
dG:dT	45.5 ± 0.3	13.4

[a]Conditions: 10 mM PIPES (pH 7.0); 10 mM $MgCl_2$; 100 mM NaCl.
[b]Total oligonucleotide concentration = 4.4μM.

3.4.2 Effects of α-18, β-18, and β-19 on DNA Polymerase Activity:

A major impetus for constructing **18** and **19** is that they are configurationally stable analogues of the α- and β-anomers of Fapy lesions and could be used as probes to determine the biological activity of the individual anomers. Molecular modeling predicts that the structure of the analogues are relatively similar to the native lesions. However, UV melting studies are only partially consistent with this. The incorporation of dNTP's opposite **18** and **19** by Klenow (exo$^-$) was investigated and compared with similar events for Fapy•dA and Fapy•dG to ascertain the usefulness of **18** and **19** as molecular mimics, and to assign the potential mutagenic properties of the α- and β-anomers of Fapy lesions.

The selection of Klenow (exo$^-$) for the polymerase in these experiments was chosen for its lack of exonuclease activity that the holoenzyme Klenow Fragment possesses.[214] Repair functionalities associated with polymerases can complicate the measurement of insertion kinetics opposite modified sites in a DNA template-primer complex.

3.4.2.1 Kinetics of Insertion Opposite α-18, β-18, and Fapy•dA:

3.4.2.1.1 Kinetics of Insertion Opposite α-18 and β-18:

The general hypothesis generated from the thermodynamic melting studies for α-**18** and β-**18** was that the analogues were extrahelical in nature and did not participate in hydrogen bonding. In order to further examine this hypothesis, comparisons were made to an abasic site model.[215] The general preference for incorporation of dNTP's opposite an abasic site is that dATP is generally incorporated more efficiently than any other dNTP.[215] The rationale for this occurrence is known as the "A-rule". It is based upon the ability of dA to base stack more efficiently to the adjacent nucleotide. Therefore, if α-**18** and β-**18** are extrahelical the kinetics of insertion should show a similar preference for the insertion of dNTP's opposite the analogue as the abasic site model. The standing start polyacrylamide gel fidelity assay was used to measure the insertion efficiencies of dNTP's opposite α-**18** and β-**18** in a template strand.[79] Klenow (exo$^-$) was used as the polymerase enzyme under single-hit conditions for the insertion event.[79] Enzyme kinetic paramaters were obtained through a Hanes-Woolf plot for the insertion of dNTP's opposite α-**18**, β-**18**, and dA in the primer-template **120** (Figure 32). The efficiency of insertion was reduced by ~40–fold for the insertion of TTP opposite α-**18** and β-**18**

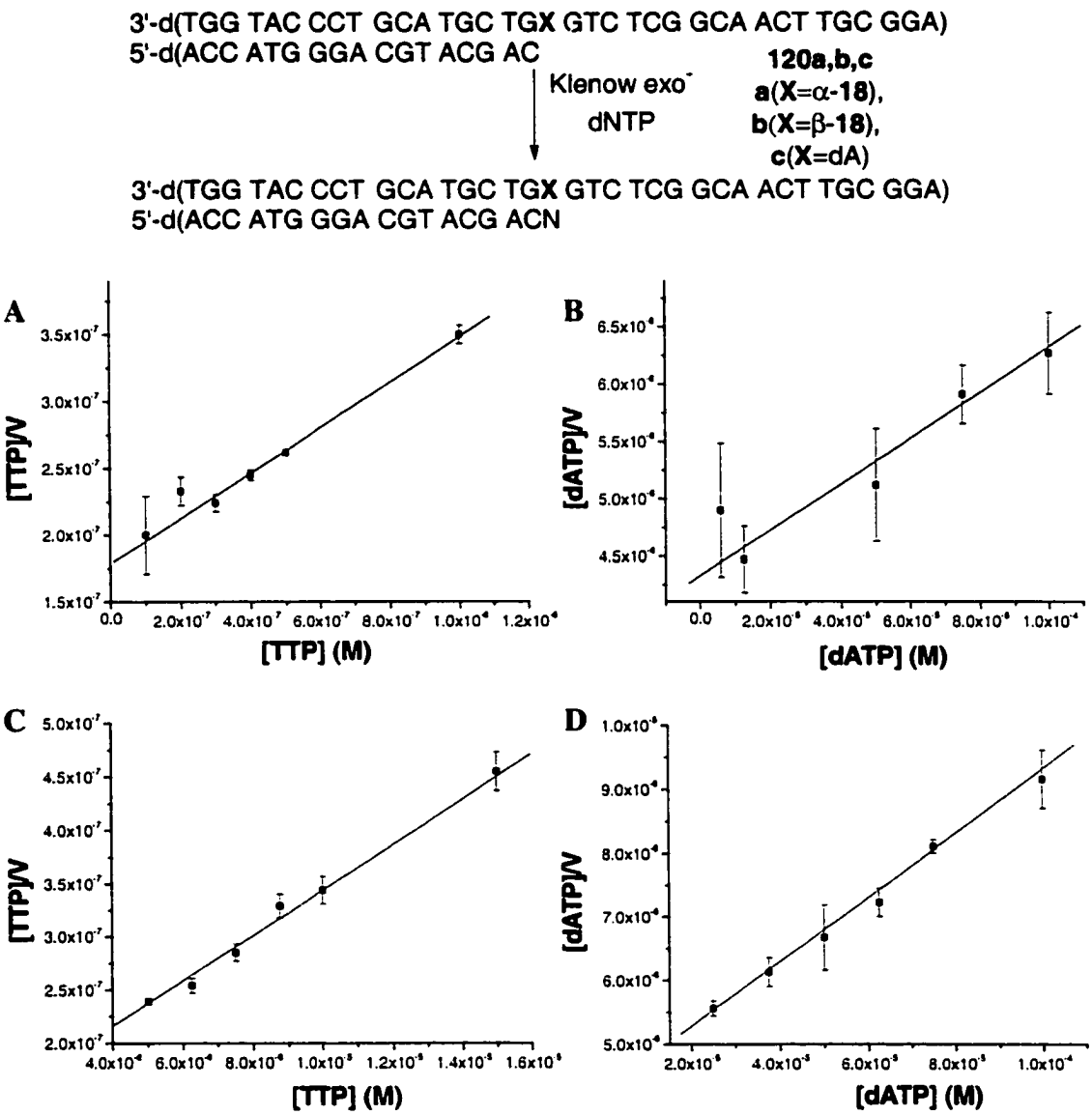
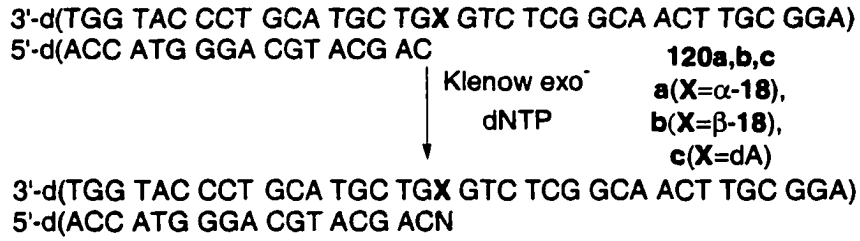


Figure 32. Sample Hanes-Woolf plots used to determine kinetic parameters for dNTP insertion opposite β-**18** and α-**18**. A. TTP→ β-**18**; B. dATP→β-**18**; C. TTP→ α-**18**; D. dATP→ α-**18**.

relative to dA (Table 17). Incorporation of mismatches (dA, dG, or dC) opposite the C-nucleotide analogues were similar in efficiency to the insertion events opposite dA. Comparable trends were not observed between of the efficiency of dNTP incorporation (V_{max}/K_m) for α-**18** and β-**18** and the abasic site model. The absolute magnitudes for the

Table 17. Comparison of the insertion efficiencies of dNTP's opposite α-**18**, β-**18**, dA, and **118**.



X	V_{max}/K_m (% $min^{-1} \cdot M^{-1}$)			
	TTP	**dATP**	**dCTP**	**dGTP**
α-**18**[a]	4.7×10^6	2.1×10^5	7.5×10^5	3.4×10^4
β-**18**[a]	6.2×10^6	1.8×10^5	1.8×10^4	3.4×10^4
dA[a]	2.3×10^8	4.1×10^5	5.4×10^4	2.3×10^5
118[b,c]	0.6×10^2	12×10^2	0.5×10^2	2.1×10^2

[a]Performed with Klenow exo⁻.
[b]A different sequence was used for the insertion across from **118**.
[c]Performed with *Drosophilia* Polymerase α.

efficiencies are different due to the use of another polymerase enzyme for the insertion opposite **118**. Although the A-rule is in effect for the insertion events opposite **118**, the insertion of TTP opposite α-**18** and β-**18** is more efficient than dATP. This suggests that in the context of a primer-template duplex, α-**18** and β-**18** are instructional in base pairing.

The lower efficiencies of insertion relative to dA suggest that α-**18** and β-**18** inhibit DNA synthesis, and they may prevent the polymerase from extending past the modified site. To investigate this event further, a qualitative full-length polymerization experiment with Klenow exo⁻ was performed on primer-template duplexes containing α-**18**, β-**18**, and dA (Figure 33). Polymerization was begun 3 bases prior to α-**18** and β-**18** in the presence of excess dNTP's and monitored at 5, 10, 15, 30, 60, and 120 min. Visual inspection of the gel shows that polymerization past dA occurs without any pause in

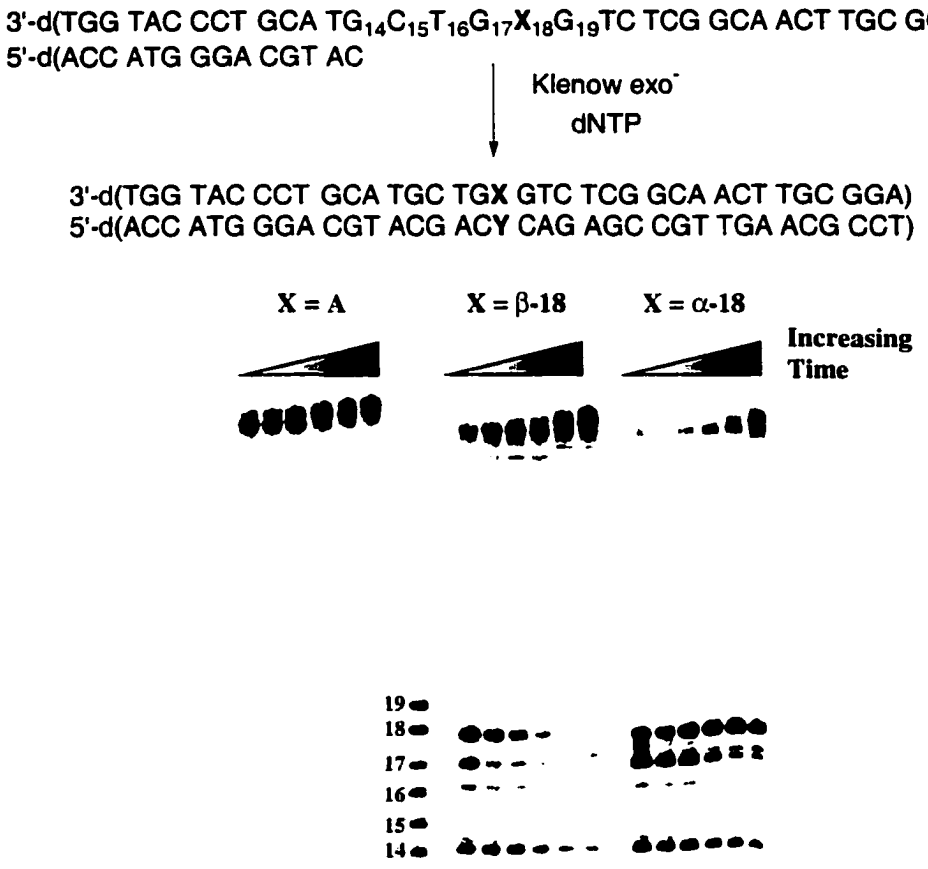


Figure 33. Full-length polymerization with Klenow exo$^-$ past dA, β-**18**, and α-**18**. Conditions: 50 nM duplex; 10 nM Klenow (exo$^-$); 20 μM dNTP's; 10 mM Tris-HCl (pH 7.5); 5 mM $MgCl_2$; 7.5 mM DTT; 0.1μg/μL BSA. Aliquots were removed at 5, 10, 15, 30, 60, and 120 min.

polymerase insertions. However, both α-**18** and β-**18** cause the build up of intermediate extension products before fully extending. A major pause product is 17 nucleotides long and corresponds to the insertion opposite α-**18** and β-**18**. Slight pausing is observed for the insertion at the 3' adjacent nucleotide of the lesion (16 mer). The predominant pause product is 18 nucleotides in length and represents extension past α-**18** or β-**18**. Therefore, the C-nucleosides are exerting a strong influence on the neighboring 5'

nucleotide. Moreover, qualitative inspection of the gel reveals that α-**18** is a stronger inhibitor of polymerase insertion relative to β-**18**.

The kinetic parameters for the insertion events opposite α-**18**, β-**18**, and dA were used to determine the misinsertion frequency (F_{ins}) (Table 18).[104] This reveals that misinsertion events opposite α-**18** and β-**18** are less efficient relative to dA. The relative misinsertion frequency (F_{rel}) was also determined for insertion events opposite α-**18** and β-**18** (Table 18). The calculation of F_{rel} compensates for the decrease in insertion

Table 18. Kinetic parameters of nucleotide insertion opposite α-**18** and β-**18**.

3'-d(TGG TAC CCT GCA TGC TGX GTC TCG GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT ACG AC

Klenow exo⁻
dNTP

120a,b,c
a(X=α-**18**),
b(X=β-**18**),
c(X=dA)

3'-d(TGG TAC CCT GCA TGC TGX GTC TCG GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT ACG ACN

X	dNTP	K_m (μM)	V_{max} (% min^{-1})	V_{max}/K_m	F_{ins}[a]	F_{rel}[b]
dA	T	0.04 ± 0.01	8.2 ± 1.1	2.3 ± 0.3 x 10^8	1.0	--
dA	A	126 ± 32	49.9 ± 3.5	4.1 ± 1.0 x 10^5	1.8 x 10^{-3}	--
dA	G	58.8 ± 33.0	11.6 ± 1.5	2.3 ± 1.0 x 10^5	1.0 x 10^{-3}	--
dA	C	56.3 ± 30.4	2.6 ± 0.1	5.4 ± 2.9 x 10^4	2.7 x 10^{-4}	--
β-**18**	T	0.9 ± 0.5	5.5 ± 2.7	6.2 ± 0.8 x 10^6	2.7 x 10^{-2}	1.0
β-**18**	A	322 ± 94	55.0 ± 13.3	1.8 ± 0.9 x 10^5	7.8 x 10^{-4}	2.9 x 10^{-2}
β-**18**	G	75.0 ± 7.3	2.6 ± 0.1	3.4 ± 0.1 x 10^4	1.5 x 10^{-4}	5.6 x 10^{-3}
β-**18**	C	96.0 ± 9.2	1.7 ± 0.2	1.8 ± 0.1 x 10^4	7.8 x 10^{-5}	2.9 x 10^{-3}
α-**18**	T	6.3 ± 0.1	29.3 ± 15.6	4.7 ± 2.6 x 10^6	2.0 x 10^{-2}	1.0
α-**18**	A	83.5 ± 10.0	17.1 ± 4.4	2.1 ± 0.7 x 10^5	9.1 x 10^{-4}	4.6 x 10^{-2}
α-**18**	G	39.9 ± 21.0	2.7 ± 0.9	8.5 ± 6.6 x 10^4	3.7 x 10^{-4}	1.9 x 10^{-2}
α-**18**	C	29.6 ± 3.2	21.9 ± 0.7	7.5 ± 1.0 x 10^5	3.3 x 10^{-3}	0.17

[a] F_{ins}= (V_{max}/K_m, X=A, α-**18**, or β-**18**; dNTP=T, A, G, or C) /(V_{max}/K_m, X=A; dNTP=T).
[b] F_{rel}= (V_{max}/K_m, X=α-**18**, or β-**18**; dNTP=T, A, G, or C) /(V_{max}/K_m, X=α-**18**, or β-**18**; dNTP=T)

efficiency of TTP opposite α-**18** and β-**18**. This also shows that there is a low preference for misinsertion events translesionally in β-**18**. In α-**18**, insertion of dATP is ~ 1.5 times greater and dGTP is ~3 times greater than the similar insertion across from β-**18**. Insertion of dCTP is much more prevalent across from α-**18** occurring 17% of the time relative to the insertion of TTP.

3.4.2.1.2 Comparison of the Kinetics of Insertion of α-18 and β-18 with Fapy•dA:

Investigations into the ability of Fapy•dA to promote misincorporation of dNTPs by Klenow (exo$^-$) were also performed in our group and reveal that Fapy•dA does not inhibit Klenow (exo$^-$) nearly as much as α-**18** or β-**18** (Table 9).[104] Insertion of TTP opposite Fapy•dA is decreased ~4-fold relative to the control (compared to ~40-fold for **18**, Table 18). Calculation of the F_{rel} for misinsertion events opposite Fapy•dA, α-**18**, and β-**18** normalizes the data for comparison (Table 19). The relative ability of β-**18** to induce Klenow (exo$^-$) to misincorporate dATP, dCTP, and dGTP are very similar to the values obtained for the same insertions opposite Fapy•dA. Similar comparisons with α-**18** do not demonstrate the same correlation. This suggests that β-Fapy•dA is the biologically relevant diastereomer.

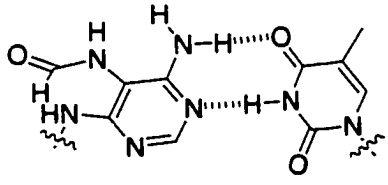
Table 19. Comparison of the F_{rel} for insertion events across from α-**18**, β-**18**, and Fapy•dA by Klenow (exo$^-$).

dNTP	Fapy•dA	β-18	α-18
A	1.4×10^{-2}	2.9×10^{-2}	4.6×10^{-2}
G	5.4×10^{-3}	5.6×10^{-3}	1.9×10^{-2}
C	1.5×10^{-3}	2.9×10^{-3}	0.17

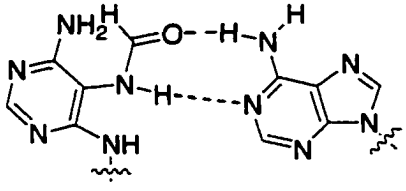
The ability of Fapy•dA and β-**18** to induce Klenow (exo$^-$) to misincorporate dATP can be explained by their ability to adopt the *syn*-conformation. Molecular modeling studies on the monomer of Fapy•dA demonstrate that adoption of the *syn*-conformation, as well as different rotamers of the formamide are energetically possible (Section 3.3.1).

The higher difference in energy between similar conformations in β-**18** is a possible factor in the reduced efficiency of insertion. In the *syn*-conformation Fapy•dA can present a "T-like" hydrogen bonding pairing pattern to permit the insertion of dATP (Scheme 28). Rotation about the formamide bond presents a hydrogen bonding structure that enables it to base pair with dG using bifurcated hydrogen bonds in a manner similar to that in 8-OxodA:dG base pairs.[94]

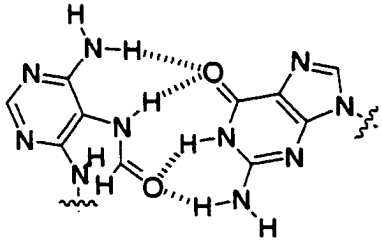
Scheme 28.



***anti*-Fapy•dA : T**



***syn*-Fapy•dA : dA**



***syn*-Fapy•dA : dG**

3.4.2.2 Kinetics of Insertion Opposite β-19 and Fapy•dG:

3.4.2.2.1 Kinetics of Insertion Opposite β-19:

The kinetic parameters for the insertion of dNTP's opposite β-**19** by Klenow (exo$^-$) were measured using the same protocol as for Fapy•dA and **18** (Figure 34, Table 20). Insertion of dCTP opposite β-**19** is reduced by ~10 fold relative to insertion opposite

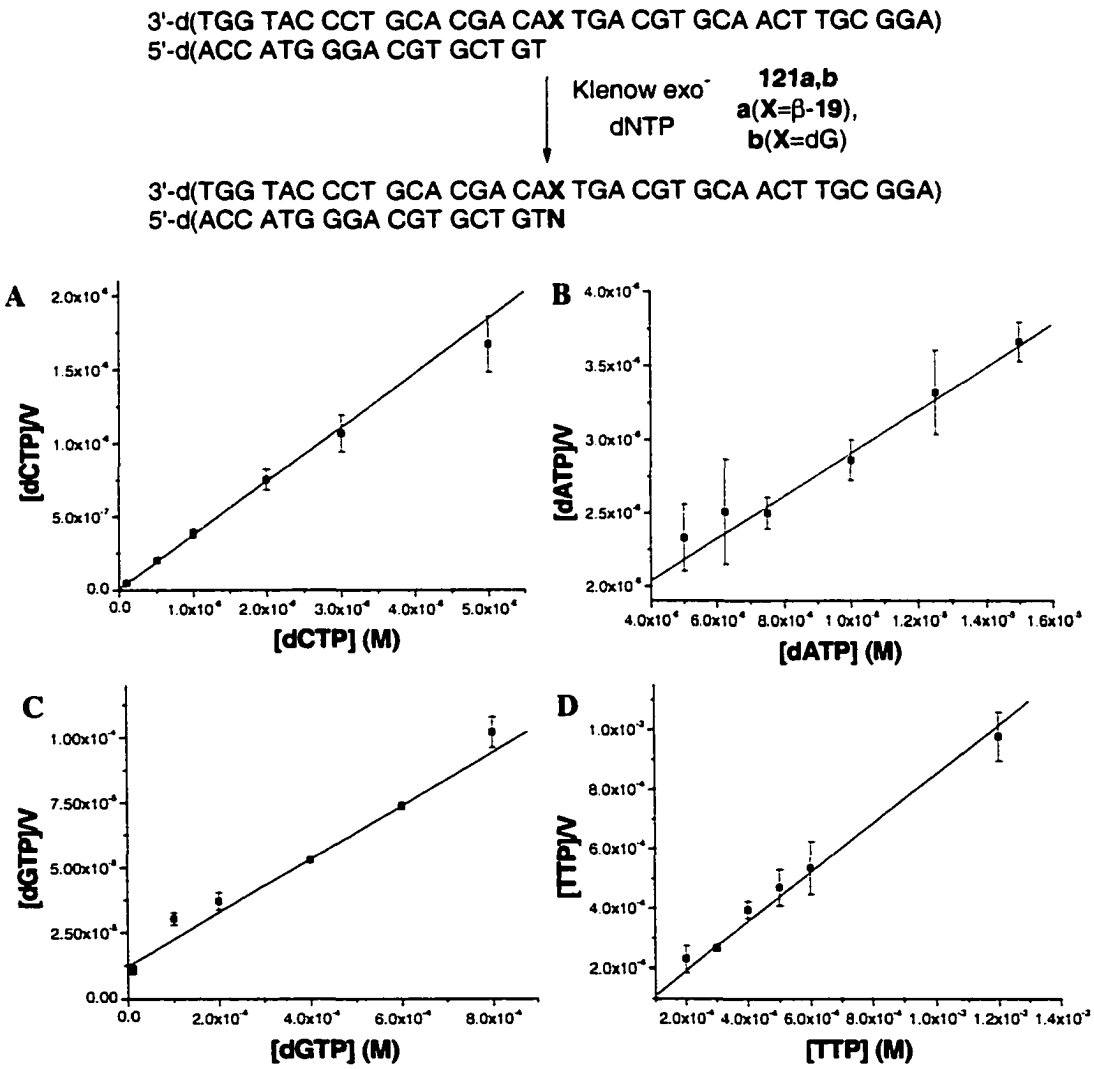


Figure 34. Sample Hanes-Woolf plots used to determine kinetic parameters for dNTP insertion opposite β-**19**. A. dCTP→ β-**19**; B. dATP→β-**19**; C. dGTP→ β-**19**; D. TTP→ β-**19**.

Table 20. Kinetic parameters of nucleotide insertion opposite β-**19**.

3'-d(TGG TAC CCT GCA CGA CAX TGA CGT GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT GCT GT

Klenow exo⁻ dNTP ↓ **121a,b a(X=β-19), b(X=dG)**

3'-d(TGG TAC CCT GCA CGA CAX TGA CGT GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT GCT GTN

X	dNTP	K_m (μM)	V_{max} (% min^{-1})	V_{max}/K_m	F_{ins}[a]	F_{rel}[b]
dG	C	$5.2 \pm 2.7 \times 10^{-3}$	5.1 ± 1.6	$9.8 \pm 3.6 \times 10^{8}$	1.0	--
dG	A	257 ± 135	14.0 ± 7.9	$5.4 \pm 3.2 \times 10^{4}$	5.6×10^{-5}	--
dG	G	74.2 ± 7.5	6.1 ± 2.1	$8.2 \pm 1.1 \times 10^{4}$	8.4×10^{-5}	--
dG	T	248 ± 60	4.7 ± 2.5	$1.9 \pm 0.7 \times 10^{4}$	1.9×10^{-5}	--
β-**19**	C	0.03 ± 0.001	2.8 ± 0.1	$9.7 \pm 0.1 \times 10^{7}$	9.9×10^{-2}	1.0
β-**19**	A	16.0 ± 5.4	8.7 ± 1.6	$5.7 \pm 1.1 \times 10^{5}$	5.8×10^{-4}	5.9×10^{-3}
β-**19**	G	50.7 ± 7.3	7.9 ± 2.8	$1.5 \pm 0.4 \times 10^{5}$	1.6×10^{-4}	1.6×10^{-3}
β-**19**	T	222 ± 59	2.4 ± 0.7	$1.1 \pm 0.1 \times 10^{4}$	1.1×10^{-5}	1.1×10^{-4}

[a] F_{ins}= (V_{max}/K_m, X=G, or β-**19**; dNTP = T, A, G, or C) /(V_{max}/K_m, X=G; dNTP = C).
[b] F_{rel}= (V_{max}/K_m, X=β-**19**; dNTP = T, A, G, or C) /(V_{max}/K_m, X= β-**19**; dNTP = C).

dG. This is four times more efficient than insertion of TTP opposite α-**18** or β-**18**. Interestingly, β-**19** induces Klenow (exo⁻) to misinsert dATP ~7 times more readily than dG. This was not the case for the Fapy•dA analogues. The relative misinsertion frequencies show that the incorporation of dATP is the predominant misinsertion event. dGTP is inserted ~4-times less frequently. A possible explanation for the difference in behavior of β-**19** than that of β-**18** is that molecular modeling predicted only a small difference of the nucleobase tilt of β-**19** relative to dG and a large difference was predicted for β-**18** relative to dA (Section 3.3).

3.4.2.2.2 Comparison of the Kinetics of Insertion Opposite Fapy•dG with β-19:

Recent results in our lab demonstrated that Fapy•dG induces Klenow (exo$^-$) to produce predominantly G→T and to a lesser extent G→C transversions (Table 8, Section 2.4.2.1).[103] Comparison of the relative frequency of misinsertion opposite Fapy•dG and β-**19** by Klenow (exo$^-$) show remarkable similarities for the insertion of dGTP and dCTP opposite these compounds (Table 21). The insertion of dATP opposite β-**19** is ~9

Table 21. Comparison of the F_{rel} for insertion events across from β-**19** and Fapy•dG by Klenow (exo$^-$).

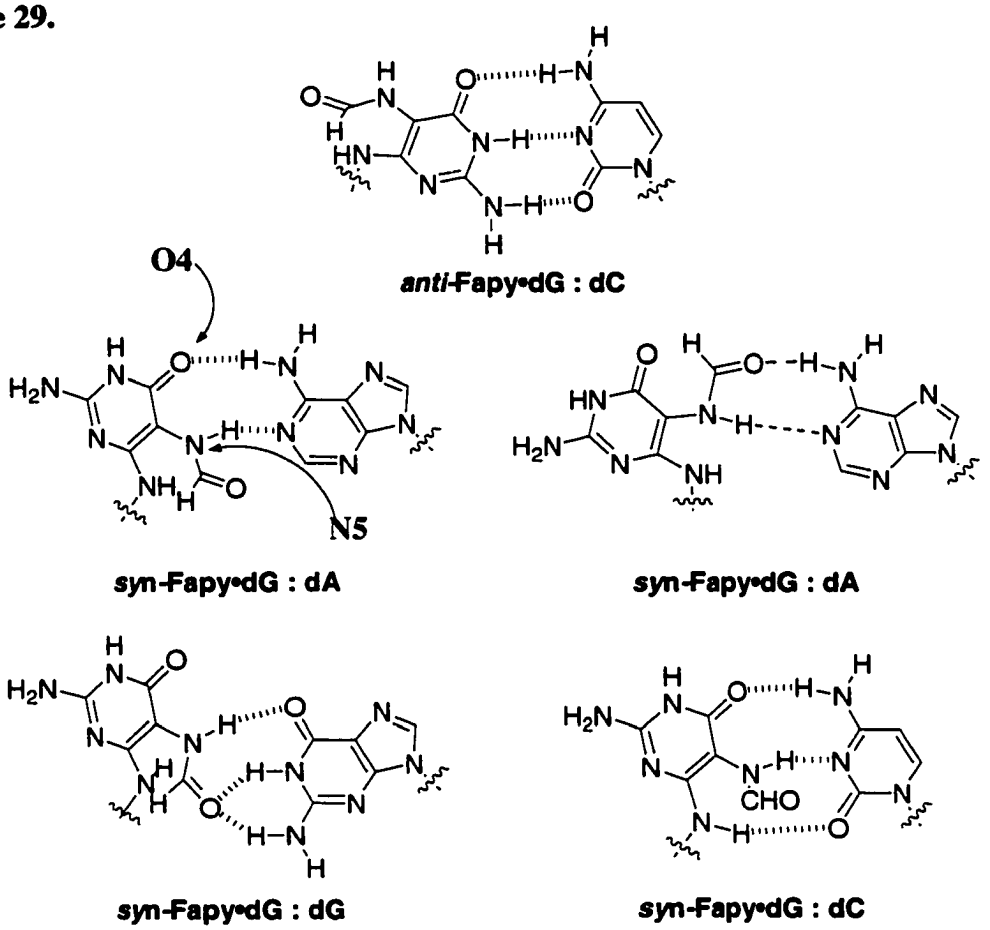
dNTP	Fapy•dG	β-19
C	1	1
A	0.05	5.9×10^{-3}
G	2.2×10^{-3}	1.6×10^{-3}
T	1.6×10^{-4}	1.1×10^{-4}

times less frequent than opposite Fapy•dG. These results suggest that β-Fapy•dG is responsible for the potential mutagenicity that is observed for Fapy•dG and that β-**19** is a suitable model for studying the structure and function of Fapy•dG in DNA.

If the β-anomer is the relevant isomer of Fapy•dA and Fapy•dG, the duplex may be stabilizing these isomers. In studies on the monomers of Fapy•dA and Fapy•dG, epimerizaition is a facile process (Section 2.). If the α-isomer of Fapy lesions was present it should have a greater influence on Klenow (exo$^-$) than what was observed.

UV melting studies and molecular modeling studies on the monomer of β-Fapy•dG are in agreement with the above observations. Fapy•dG:dA and Fapy•dG:dG base pairs are stabilized significantly over the native mismatches (Table 16). A number of different hydrogen bonding patterns that accommodate the insertion of dATP and

Scheme 29.



dGTP opposite Fapy•dG may be envisioned (Scheme 29). Molecular modeling studies on the monomer demonstrate that the *syn*-conformation is slightly favored. Very little energetic discrimination is observed between different rotamers of the formamide. Two potential conformations of Fapy•dG can explain the preference of Klenow (exo$^-$) to insert dATP. Hydrogen bonding of the O-4, *N*-5-H is analogous to the interactions between 8-OxodG and dA (Figure 13).[90] Rotation about the formamide provides a similar hydrogen-bonding pattern through the formamide carbonyl and the *N*-5-H. Both of these conformations are predicted to be minima for monomeric Fapy•dG by molecular modeling calculations, and are within 0.2 kcal/mol of one another. Base pairing of Fapy•dG with dG can be accomplished through formation of bifurcated hydrogen bonds

with the formamide carbonyl and hydrogen bonding through the *N*-5-H. Another interesting feature of the *syn*-conformation of Fapy•dG is that hydrogen bonding with dC can also be accommodated through involvement of the glycosidic NH. Recall that modeling studies predict that the formamide group is rotated out of the plane of the pyrimidine ring. This facilitates interaction between Fapy•dG's glycosidic-NH bond and the dC carbonyl oxygen. This conformation is potentially anti-mutagenic preserving the integrity of the G:C base pair. However, the efficiency of insertion of dCTP opposite β-**19** is similar to Fapy•dG. If the *syn*-conformation promotes the insertion of dC by polymerase opposite Fapy•dG, a larger preference to insert dCTP opposite Fapy•dG would have been observed relative to β-**19** due to the absence of the glycosidic NH in the latter.

An interesting correlation can be made between the molecular modeling predictions between the C-nucleoside analogues and the native Fapy lesions and the inhibition of polymerase activity that was observed for insertion of dNTP's opposite **18** and **19**. Molecular modeling calculations predicted a large variation in the nucleobase tilt for β-**18** relative to dA (Section 3.3.1). This is consistent with the ~10 fold inhibition of polymerase activity that was observed for the insertion of TTP across from β-**18** relative to Fapy•dA (Table 22). Calculation of the nucleobase tilt on β-**19** revealed a remarkably similar tilt as dG (Section 3.3.3). In contrast to observations involving the Fapy•dA/**18** molecule, the insertion of dCTP opposite β-**19** is more facile than insertion opposite Fapy•dG by Klenow (exo$^-$). The relationship between the nucleobase tilt and the inhibition of polymerase activity suggests that Klenow (exo$^-$) prefers to insert dNTP's

opposite modifications that possess a similar structure to the native nucleotide in the template.

Table 22. Correlation between the nucleobase tilt and the inhibition of polymerase activity.

X	V_{max}/K_m	Nucleobase Tilt (°)[a]	X	V_{max}/K_m	Nucleobase Tilt(°)[b]
dA	1.0[c]	1.1	dG	1.0[d]	-0.2
Fapy•dA	0.28[e]	18.2	Fapy•dG	0.019[f]	19.4
β-**18**	0.027[c]	64.2	β-**19**	0.099[d]	-1.1

[a]Supplementary Table 7
[b]Supplementary Table 9
[c]Table 18 (insertion of TTP opposite X)
[d]Table 20 (insertion of dCTP opposite X)
[e]Table 9 (insertion of TTP opposite Fapy•dA)
[f]Table 8 (insertion of dCTP opposite Fapy•dG)

3.4.3 Interactions of Oligonucleotides Containing α-18, β-18 and β-19 with DNA Repair Proteins:

The implications of Fapy lesion involvement in mutagenic events points to the importance of their recognition by DNA repair proteins. Incorporation of dNTP's opposite β-**18** and β-**19** revealed similar misinsertion preferences to the native lesions. Oligonucleotides containing the C-nucleotide analogues are suitable models to investigate the binding ability of DNA repair proteins and can potentially function as competitive inhibitors of DNA repair protein activity. Investigations into the effects of mismatches on the recognition of DNA repair proteins are also important experiments that reveal any preference DNA repair proteins have for mismatch recognition. Mismatches that warrant investigation are the products of misinsertion by polymerases.

3.4.3.1 Recognition of Oligonucleotides Containing α-18, β-18 and β-19 by Formamidopyrimidine DNA Glycosylase (Fpg):

3.4.3.1.1 Recognition of Fapy•dA Analogues by Fpg:

The binding affinities of the Fpg protein were measured for duplexes containing α-**18** or β-**18** (**122**) opposite the 4 native nucleotides (Figure 35, Table 23). Analysis of the binding affinities was accomplished through polyacrylamide gel-shift assays (PGSA).

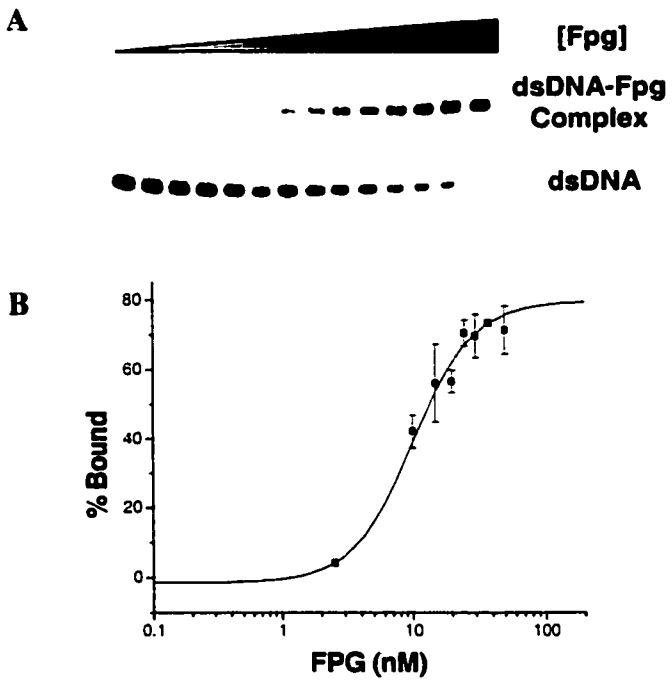


Figure 35. Polyacrylamide gel-shift assay of Fpg binding to β-**18**:T. A) Representative polyacrylamide gel-shift assay of 0-100 nM Fpg binding to β-**18**:T. B) Logarithmic plot of Fpg binding affinity.

Binding reactions were run in the presence of small amounts of duplex to minimize the amount of enzyme that is used in the experiment and non-denaturing PAGE was performed at low temperatures to prevent dissociation of the complex. Fpg strongly binds to duplexes containing β-**18** regardless of the opposing nucleotide. Investigation

Table 23. Dissociation constants of the Fpg protein for oligonucleotides containing α-**18** or β-**18**.

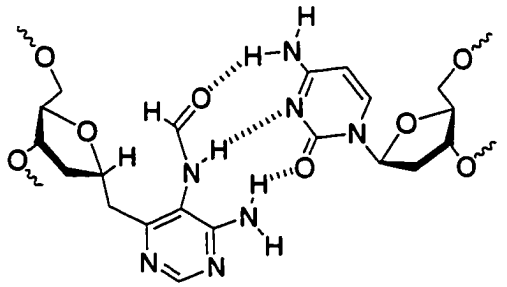
5'-d(AGG CGT TCA ACG GCT CTG **X**GT CGT ACG TCC CAT GGT)
3'-d(TCC GCA AGT TGC CGA GAC **Y**CA GCA TGC AGG GTA CCT)
122a,b,c
a(X=α-18), b(X=β-18), c(X=dA)

X	Y	K_d[a] (nM)
β-**18**	T	10.5 ± 1.3
β-**18**	A	12.8 ± 2.9
β-**18**	G	26.5 ± 2.9
β-**18**	C	10.9 ± 1.3
α-**18**	T	Non-specific binding
α-**18**	A	Non-specific binding
α-**18**	G	Non-specific binding
α-**18**	C	7.1 ± 1.5
dA	T	170 ± 37.5

[a]Classification of non-specific binding is based upon the inability to saturate the duplex at high concentrations of Fpg (>500 nM).

into the preference of Fpg to bind to α-**18** revealed different results. Non-specific binding was observed for α-**18**:T, dA, or dG base pairs. However, strong binding of the α-**18**:dC base pair by Fpg was observed. This correlates with Klenow (exo$^-$) experiments (Table 18). Oligonucleotides containing α-**18** induce Klenow (exo$^-$) to insert dCTP efficiently across from itself (Table 18). Thermodynamic melting studies do not reveal any stabilizing features of the α-**18**:dC base pair (Table 13) and molecular modeling does not predict an energy minimum that can form obvious hydrogen bonds with dC. Base pairing can be rationalized by α–**18** adopting a "dG-like" hydrogen bonding pattern (Scheme 30). However, the adoption of this base pairing conformation requires distortion of the DNA backbone and will disrupt the helicity of the duplex.

Scheme 30.



α-18: dC

The strong binding to mismatches opposite β-**18** suggests that Fpg can excise Fapy•dA from such duplexes. Kinetic analysis of Fapy•dA excision by Fpg supports the above predictions.[216] This finding has significant implications on the potential mutagenicity of Fapy•dA. Excision of Fapy•dA from mismatch sequences will result in a permanent alteration in the sequence of the DNA (Figure 36). The misinserted base will be the only information remaining to instruct another polymerase when the void is filled.

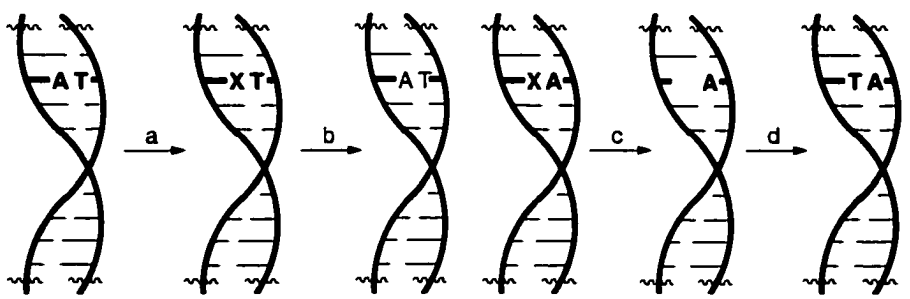


Figure 36. Potential mutagenicity of Fpg through the excision Fapy•dA from a Fapy•dA dA base pair. A) Oxidative damage generates Fapy•dA. B) Replication and misinsertion. C) Fpg excises Fapy•dA from mismatch. D) Repair of intermediate strand break. **X**= Fapy•dA

A potential explanation for the general recognition of mismatched base pairs of Fapy•dA is that the abundance of the misinsertion events are relatively minor and are not significant enough to merit specific repair. Thus, the Fpg protein has not evolved to discriminate between Fapy•dA:T, Fapy•dA:A, and Fapy•dA:G base-pairs. Another possibility is that a different repair protein is responsible for maintaining the mismatches opposite of Fapy•dA *in vivo*.

Recently, Carissa Wiederholt investigated the potential inhibitory properties of β-**18** on Fpg. Using the β-**18**:T duplex (**122b**), the inhibition excision of Fapy•dA by Fpg

5'-d(AGG CGT TCA ACG GCT CTG **X**GT CGT ACG TCC CAT GGT)
3'-d(T CC GCA AGT TGC CGA GAC TCA GCA TGC AGG GTA CCT)
122b (X=β-18)

substrates was measured. Measurement of the apparent K_m and V_{max} for the excision at fixed amounts of the inhibitor provides a K_i of 3.5 ± 0.3 nM.[216] The potential uses of β-**18** as a therapeutic compound are potentially limited by enzymatic incorporation of its triphosphate into DNA. Furthermore, investigations into the inhibition of DNA repair processes have been avoided for fear of damaging the integrity of DNA in living systems. The important role that repair processes play in maintaining the integrity of the genome have dissuaded scientists from exploring this avenue of research.

3.4.3.1.2 Recognition of β-Fapy•dG Analogue by Fpg:

The affinity of the Fpg protein for duplexes containing β-**19** opposite the four native nucleotides was quantified with the same protocol as duplexes containing **18** (Table 24). In contrast to the general preference of Fpg to recognize mismatches

Table 24. Dissociation constants of the Fpg protein for oligonucleotides containing β-**19**.

5'-d(AGG CGT TCA ACG TGC AGT **X**AC AGC ACG TCC CAT GGT)
3'-d(TCC GCA AGT TGA ACG TCA **Y**TG TCG TGC AGG GTA CCA)
123 (X= β-19)

X	Y	K_d[a] (nM)
β-**19**	C	11.8 ± 0.7
β-**19**	A	Non-specific binding
β-**19**	G	Non-specific binding
β-**19**	T	Non-specific binding

[a]Classification of non-specific binding is based upon the inability to saturate the duplex at high concentrations of Fpg (>500 nM).

opposite Fapy•dA, the β-**19**:dC base pair is the only duplex recognized. The observed K_d for this process is very similar to that measured for the binding of Fpg to β-**18**. The ability of Fpg to exclusively recognize the β-**19**:dC base pair is consistent with independently measured kinetic parameters for the excision of the Fapy•dG opposite dC and dA by Fpg.[217] This selectivity implies that genetic integrity will not be compromised by the excision of Fapy•dG opposite mismatches and provides further support that β-**19** is and excellent mimic of Fapy•dG, .

3.4.3.2 Recognition of Oligonucleotides Containing α-18, β-18 and β-19 by MutY:

Other investigations in our group have explored the ability of MutY to recognize Fapy lesions.[214,215] The importance of MutY in the repair of oxidative damage to purines is that it specifically recognizes mismatches and excises the base opposite the lesion.[16] Investigations into the ability of MutY to recognize Fapy•dA demonstrate that the Fapy•dA:dA base pair is not a substrate for excision of dA.[216] Interestingly, 8-OxodA:dA base pairs are substrates for the MutY protein.[218] This is significant because

MutY does not protect against removal of Fapy•dA from the mismatched duplex by Fpg. The consequences of this are that Fapy•dA mismatches will be deleterious if acted upon by Fpg. Similar arguments can be made about the relevance of Fapy•dA as a mutagenic lesion due to its low probability to induce misincorporation of dNTP's. However, the observation that MutY does not protect Fapy•dA:dA base pairs from Fpg enhances the mutagenic potential of Fapy•dA.

In contrast to duplexes containing base pairs Fapy•dA:dA, Fapy•dG:dA is a substrate for MutY. Furthermore, oligonucleotides containing β-**19**:dA are efficiently recognized by MutY with a K_d =5.4 nM.[217] These observations parallel the ability of MutY to recognize and excise dA from 8-OxodG:dA base pairs.[218] The potentially deleterious properties of Fapy•dG and 8-OxodG necessitates efficient repair of misinsertion events, as well as removal of the lesion. Thus, the Fpg and MutY proteins work in tandem to repair the consequences of Fapy•dG and 8-OxodG and the nucleotide misinsertions that they induce.

4 Conclusion:

The study of C-nucleoside analogues of Fapy lesions (**18** and **19**) with defined stereochemistry at C_1' in oligonucleotides has enabled us to probe the structure and function of Fapy lesions in DNA in more detail than what was otherwise possible due to the epimerizable nature of Fapy lesions. The synthetic goal of this project was the construction of phosphoramidites α-**20**, β-**20**, and β-**21**. Installation of the C-glycosidic linkage was efficiently completed through the use of the Wiitig olefination reaction and subsequent cyclization. The use of pyrimidine ylides **56** and **75** provided a way of introducing the C-5 nitrogen functionality prior to construction of the C-glycosidic linkage. Limitations in the synthesis of **20** led to the design and implementation of an improved synthetic approach for the construction of **21**. The key transformation that leads to the synthesis of **21** is the selective demethylation of **85**.

Syntheses of oligonucleotides containing α-**18**, β-**18** and β-**19** were accomplished using modified synthesis cycles. Reactivity of the N-5 formamide resulted in transamidation with the use of conventional capping agents. The use of isobutyryl or pivaloyl anhydride without NMI prevents transamidation at the expense of capping efficiency. Synthesis of dodecameric sequences containing α-**18** and β-**18** were successful using DCI as the activator and double coupling the respective phosphoramidite for 5 min. However, the poor capping efficiency prevented the purification of longer oligonucleotides. Enzymatic ligation of the dodecamers provided an efficient way of incorporating **98a**, **98b**, **99a**, and **99b** into 36 mers (**108** and **109**) on preparative scales.

Coupling of β-**21** with tetrazole activator for 15 min and the use of 20 s capping times with pivaloyl anhydride (no NMI) led to the successful synthesis of 36 mers containing β-**19** without the need for enzymatic ligation.

Comparison of molecular modeling predictions for the minimized structures of β-Fapy•dA and β-**18** reveal similar structural features. Large energy differences are present in the C-nucleoside that prevents it from adopting the *syn*-conformation as readily as β-Fapy•dA. Molecular modeling of β-Fapy•dG and β-**19** indicates that these molecules are also similar to one another and suggest that the preferred conformation is *syn*. Moreover, low energy differences were predicted between *syn*- and *anti*-conformations for β-Fapy•dG and β-**19**. However, the energy differences predicted for different conformations of β-**18**/Fapy•dA, and β-**19**/Fapy•dG are negligible when compared to the contribution of thermal energy at room temperature and suggest that multiple conformations are accessible for hydrogen bonding in DNA at biologically relevant temperatures.

Thermodynamic melting studies reveal that C-nucleoside analogues have no preference for favorable base pairing interactions with mismatches. This is inconsistent with the Fapy•dG:dA base pair that is significantly stabilized relative to a dG:dA base pair. Initial explanations of the uninformative interactions of C-nucleosides with mismatches focused on the abasic site model (**118**). Direct comparisons of oligonucleotides containing α-**18** and β-**18** demonstrate remarkable similarities with this model. However, polymerase extensions studies do not correlate with the efficiencies of insertion for the abasic site model suggesting that the C-nucleosides are instructional in the primer-template-polymerase complex.

Insertion of dNTP's opposite α-**18** and β-**18** by Klenow (exo$^-$) are retarded approximately 10-fold relative to the insertion opposite Fapy•dA. The decrease in insertion efficiencies can be potentially explained due to the differences observed in the nucleobase tilting revealed by molecular modeling. Comparison of the relative frequencies of insertion for Fapy•dA and β-**18** reveal a similar preference for the insertion of dNTP's by Klenow (exo$^-$) and β-Fapy•dA was assigned as the biologically relevant anomer. The insertion trends opposite β-**19** are directly comparable to Fapy•dG and also support assignment of β-Fapy•dG as the biologically relevant anomer. Molecular modeling predictions confirmed that β-Fapy•dG and β-**19** are structurally very similar and that there is very little energetic discrimination between potentially mutagenic conformations.

C-nucleotide analogues of Fapy lesions were also used to investigate their recognition by DNA repair proteins. The Fpg protein efficiently binds to oligonucleotides containing β-**18** opposite any of the four native nucleotides (K_d <27 nM). This observation suggests that the interaction of the Fpg protein with a Fapy•dA:dA, dG, or dT duplex may cause loss of genetic information. Investigations into the ability of Fpg to bind to duplexes containing β-**19** opposite mismatches reveal a different preference. Only the β-**19**:dC base pair is recognized by Fpg. This suggests that repair of Fapy•dG by Fpg will not promote mutagenic events when misinsertion of mismatches across from Fapy•dG have occurred.

5.0 Experimental Procedures;

General Methods: ^{1}H NMR spectra were recorded at 300 or 400 MHz. The NMR spectra are referenced according to the solvent ($CDCl_3$, 7.27; CD_3OD, 3.31; acetone-*d6*, 2.05; and DMSO-*d6*, 2.50) ^{13}C NMR were recorded at 75 or 100 MHz and referenced according to the solvent ($CDCl_3$, 77.23; CD_3OD, 49.15). ^{31}P NMR were obtained at 109 MHz and referenced according to H_3PO_4 as an external standard (H_3PO_4 = 0). IR spectra were obtained using a Perkin-Elmer Series 1600 FT-IR or an Avatar 320-FT-IR spectrophotometer. The Central Instrument Facility at Colorado State University performed LRMS and HRMS FAB. ESI-MS was performed on a VG Fisons Quattro or Finnigan LCQ-DUO.

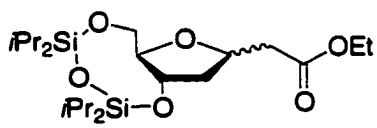
All reactions were carried out in oven-dried glassware, under an atmosphere of nitrogen, unless otherwise stated. The following chemicals were obtained commercially, dried, and purified according to literature methods prior to use.[219] THF was distilled from benzophenone and sodium. Pyridine, DCM, TMSCl, $CHCl_3$, toluene, DMF, diethyl aniline, iPr$_2$NEt, aniline, lutidine, AcN, and MeOH were distilled from CaH_2. t-Butyl acetate was distilled from $MgSO_4$. Diisopropyl amine was distilled from NaOH. Isobutyric anhydride and pivalic anhydride were distilled from P_2O_5. $POCl_3$ was distilled from itself. Phenoxyacetic acid and PPh_3 were recrystallized from EtOH and dried under reduced pressure. DMAP was recrystallized from toluene and dried under reduced pressure. NaI was recrystallized from acetone and dried under reduced pressure. All other chemicals obtained were used without further purification.

Oligonucleotide synthesis was carried out on an Applied Biosystems 394 DNA/RNA synthesizer using standard protocols. Isobutyryl-protected deoxycytidine-β-cyanoethyl phosphoramidite was purchased from Pharmacia Biotech. Phenoxyacetyl-protected deoxyadenosine, isopropylphenoxyacetyl-protected deoxyguanosine, and other standard phosphoramidites were purchased from Glen Research. All other oligonucleotide synthesis reagents were also obtained from Glen Research. Standard oligonucleotides were deprotected with 28% NH_4OH at 55 °C overnight and concentrated to dryness. Gel electrophoresis on 20 % polyacrylamide denaturing gels (PAGE) afforded the purified oligonucleotide. Oligonucleotides were precipitated from NH_4OAc/ EtOH prior to analysis by ESI-MS.[220]

Oligonucleotide melting studies were performed in 1 cm path length quartz cells on a Beckman DU 640 UV-VIS spectrophotometer equipped with a thermoprogrammer. CD spectra were performed in a 1 cm path length quartz cell on an AVIV Circular Dichroism Spectrophotometer Model 202. Anion-exchange HPLC Purification of 30 mers was carried out on a Vydac 301 VHP575 column. Gradient A: A, 10 mM Tris, pH 8.0; B, 10 mM Tris, 0.5M NH_4Cl, pH 8.0; 20-40% B linearly over 20 min; 40% B for 15 min. Gradient B: A, 10 mM Tris, pH 8.0, 10% AcN; B, 10 mM Tris, pH 8.0, 0.5M NH_4Cl, 10% AcN; 20-50% B linearly over 12 min; 50% B for 8 min; 50-60% B linearly over 2 min; 60% B for 8 min.

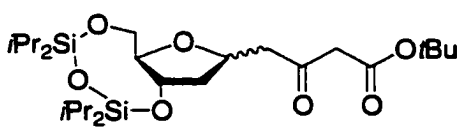
Enzymes for biochemical assays were purchased from New England Biolabs with the exception of *E. coli* formamidopyrimidine DNA glycosylase (Fpg) which was obtained from Jamie Milligan (UCSD) and from Trevigen. Radioactive [γ–^{32}P] ATP was purchased from Amersham Pharmacia Biotech. Water (ddH_2O) for biochemical studies

was obtained from a Barnstead Nanopure still and was autoclaved prior to use in experiments. Biochemical reactions were carried out in sterilized eppendorf tubes and sterilized pipette tips were employed for the transfer of solutions. Radioactive samples were counted with Cerenkov counting, using a Packard Tri-Carb 1500 scintillation counter. Gels were visualized by exposing the gel to a Molecular Dynamics Phosphor Screen (2-4 h for analytical gels, 24-36 h for dried gel-shift retardation assays) and scanning with a Molecular Dynamics Storm 820 Phosphorimager. Gel-shift assay gels were dried with a BioRad gel dryer for 1.5 h at 90 °C prior to exposure. The images were quantitated using Molecular Dynamics ImageQuant version 5.01 software. All data was plotted using Origin 6.1 (Microcal Software, Inc.) graphing software.



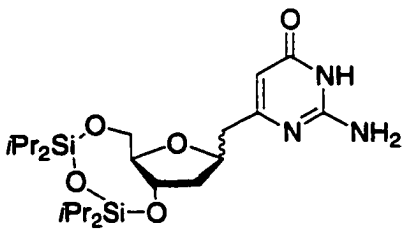
Preparation of 22:[154] A mixture of 2-deoxyribose (5.16 g, 38.5 mmol) and ethyl (triphenylphosphoranylidene)acetate (13.40 g, 38.5 mmol) was dissolved in THF (190 mL) and heated to reflux for 15 h. The solution was concentrated in vacuo to a colorless oil and chromatographed with 120 g of silica gel (elution with 2% methanol in DCM to 5% MeOH) to leave a colorless oil. The acyclic material was dissolved in absolute EtOH (180 mL) and reacted with sodium ethoxide (10 mL, 4.3 mmol) for 45 min. The reaction was quenched with 6 M HCl (0.8 mL, 4.8 mmol) and filtered over a fine centered glass funnel. The solution was concentrated to a colorless oil (7.43 g, 95%) and the crude material was carried on. ^{1}H NMR ($CDCl_3$) δ 4.58-4.47 (m, 1 H), 4.41-4.33 (m, 1 H), 4.18 (q, *J* =7.5 Hz, 2 H), 3.98-3.87 (m, 1 H), 3.78-3.60 (m, 2 H), 2.79-2.62 (m, 2 H), 2.48-2.40 and 2.08-2.02 (each as m, 1 H), 2.28 (b, 2 H), 1.98-1.79 (m, 1 H), 1.26 (t, *J* =7.5 Hz, 3 H); IR (film) 3422,

2935, 1731, 1198, 1094, 1029 cm^{-1}. A portion of the mixture of furanose and pyranose analogs from above (3.17 g, 15.51 mmol), imidazole (2.1 g, 31.02 mmol), and DMAP (0.19 g, 1.55 mmol) were coevaporated with pyridine (3 x 10 mL). The resulting oil was dissolved in pyridine (103 mL) and $TIPDSCl_2$ (6.04 g, 17.06 mmol) was added dropwise over a period of 30 min. The solution was allowed to stir for 14 h and concentrated under reduced pressure. The oil was then taken up in EtOAc (20 mL) and washed successively with 10% HCl (10 mL), water (10 mL), saturated $NaHCO_3$ (10 mL), and brine (10 mL). The clear solution was dried over $MgSO_4$ and concentrated to leave a colorless oil. The crude material was chromatographed on 120 g of silica gel (elution with 100% hexanes to 5% EtOAc) to give a mixture of α- and β-anomers of the furanose derivative (1.65 g, 24%), and a mixture of the furanose and pyranose analogues (4.44 g, 63%). ^1H NMR ($CDCl_3$) δ 4.48-4.40 (m, 2 H), 4.16 (q, J =7.2 Hz, 2 H), 4.04-3.97 (m, 1 H), 3.78-3.69 (m, 2 H), 2.74 and 2.65 (each as dd, J = 6.9, 15.5 Hz, and J = 6.6 and 8.7, 1 H), 2.45-2.40 and 2.16-2.12 (each as m, 1 H), 1.92-1.82 (m, 1H), 1.27 (t, J =7.2 Hz, 3 H), 1.12-0.92 (m, 28 H); IR (film) 2945, 2868, 1740, 1465, 1248, 1163, 1092, 1035, 886, 693 cm^{-1}.



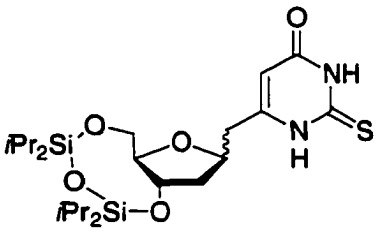
Preparation of 23: To a -78 °C solution of THF (10 mL) and sec-butyl lithium, 1.35 M in hexanes (22.4 mL, 30.24 mmol), was added diisopropylamine (2.53 g, 24.81 mmol) in THF (5 mL). The reaction was stirred for 10 min and allowed to warm to room temperature. The LDA solution was recooled to -78 °C and *t*-butyl acetate (3.51 g, 30.24 mmol) was added in THF (5 mL) dropwise. The enolate was allowed to warm to 0 °C after 30 min and **22** (3.35 g, 7.56 mmol) was added in THF (5 mL). The reaction was

stirred for 18 h and quenched with water (1 mL). Concentration of the reaction resulted in an oil that was taken up in EtOAc (100 mL). The product was washed successively with water (30 mL) and brine (30 mL). The organic layer was dried over $MgSO_4$ and concentrated to leave a brown oil. The crude product was chromatographed on 100 g of silica gel (elution with 3% EtOAc in hexanes to 10% EtOAc) to give a light yellow oil as a mixture of α- and β- anomers and keto-enol forms of **23**, as well as another impurity which was carried on (3.85 g, 99%). ^{1}H NMR ($CDCl_3$) δ 4.48-4.36, (m ,2 H), 4.02-3.92 (m, 1 H), 3.81-3.72 (m, 2 H), 3.40 (s, 2H), 2.98 (dd, 1 H, J = 4.8. 8.4 Hz), 2.78 (dd, 1 H, J = 3.0, 8.4 Hz) 2.54-2.40 (m, 1 H), 1.84-1.72 (m, 1 H), 1.48 (s, 9 H), 1.12-0.88 (m, 28 H); IR (film) 2945, 2868, 1719, 1647, 1465, 1250, 1148, 1075, 1036, 886, 695 cm^{-1}.

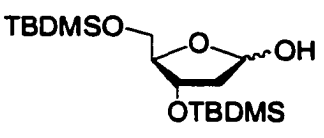


Preparation of 24: To a solution of **23** (0.57 g, 1.10 mmol) in absolute EtOH (37 mL) was added guanidine hydrochloride (0.13 g, 1.32 mmol) and sodium carbonate (0.14 g, 1.32 mmol). The solution was heated at reflux for 32 h and concentrated in vacuo. The resulting brown oil was taken up in EtOAc (10 mL) and washed with water (10 mL). The aqueous phase was reextracted with EtOAc (3 x 5 mL) and the combined organic layers were washed with brine (10 mL). The organic layer was dried over $MgSO_4$ and concentrated. The resulting brown foam was chromatographed on 50 g of silica gel (elution with 2% MeOH in EtOAc to 10% MeOH) to give a white foam as a 1:1 mixture of α- and β-anomers (0.28, 52%). mp= 135 °C (decomposed); ^{1}H NMR (actone-d_6) δ 5.59 and 5.56 (each as s, 1 H), 4.54-

4.28 (m, 1 H), 4.16-3.88 (m, 2 H), 3.82-3.74 (m, 1 H), 3.58-3.54 (m, 1 H), 2.76-2.35 (m, 2 H), 2.06-1.98 (m, 1 H), 1.92-1.78 (m, 1 H), 1.14-0.89 (m, 28 H); IR (film) 3566, 3146, 3023, 2945, 1659, 1468, 1388, 1143, 991, 885, 699 cm^{-1}; LRMS FAB (M+H) for $C_{22}H_{42}N_3O_5Si_2$: calcd. 484.3, found 484.2.



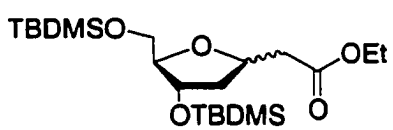
Preparation of 25: To a solution of **23** (2.00 g, 3.88 mmol) and thiourea (0.45 g, 5.82 mmol) in absolute EtOH (30 mL) was added NaOEt (10 mL, 8.54 mmol, 0.85 M in EtOH). The solution was brought to reflux and stirred for 18 h. The solution was concentrated and taken up in EtOAc (30 mL). The organic phase was washed with water (15 mL), and the aqueous phase was reextracted with EtOAc (3 x 15 mL). The combined organic phases were washed with brine (20 mL), dried over $MgSO_4$, and concentrated. The resulting oil was chromatographed on 40 g of silica gel (elution with 10 % EtOAc in hexanes to 50% EtOAc) to give an off-white foam as a mixture of anomers (0.72 g, 37%). ^1H NMR ($CDCl_3$) δ 10.16 (bd, 1 H), 9.80 (bd, 1 H), 5.72 (s, 1 H), 4.47-4.44 (m, 1 H), 4.09-4.05 (m, 1 H), 4.01-3.96 (m, 1 H), 3.82 (dd, 1 H, *J* = 4.8, 10.8 Hz), 3.71-3.64 (m, 1 H), 2.60-2.45 (m, 2 H), 1.87-1.81 (m, 1 H), 1.68 (m, 1H), 1.08-0.87 (m, 28 H); IR (KBr) 3133, 2945, 2867, 1665, 1549, 1187, 1154, 997 cm^{-1}.



Preparation of 26:[155,156] To a solution of 2-deoxyribose (4.00 g, 29.8 mmol) and $CaCO_3$ (3.0 g, 30 mmol) in water (12 mL) was added Br_2 (12.4 g, 77.5 mmol) dropwise at 0 °C. The exothermic reaction was allowed to stir for 5 min and sealed. After 48 h, the reaction was purged with N_2, diluted

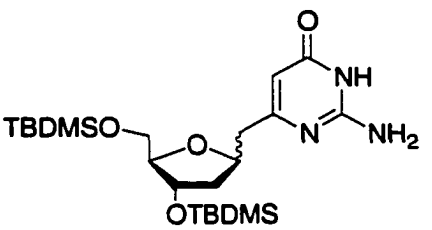
with water (50 mL), and treated with 4 g of oxalic acid. The solution was then neutralized with Ag_2CO_3 to give a green solution. The solution was filtered twice to remove the silver salts and concentrated. The resulting oil was coevaporated with pyridine (3 x 10 mL). Following coevaporation, imidazole (8.2 g, 120.5 mmol) and DMAP (0.4 g, 3.3 mmol) were added and the mixture was coevaporated with pyridine (2 x 10 mL). The oil was taken up in DMF (100 mL) and TBDMSCl (11.3 g, 75.0 mmol) was added. The reaction was stirred for 20 h and diluted with water (120 mL). The product was partitioned between the aqueous layer and Et_2O (100 mL). The aqueous layer was reextracted with Et_2O (2 x 100 mL). The combined organic layer was washed with water (50 mL), saturated $NaHCO_3$ (50 mL), and brine (50 mL). The organic phase was dried over $MgSO_4$ and concentrated to give a brown oil. The oil was purified by column chromatography with 75 g of silica gel (elution with 2% EtOAc in hexanes to 15% EtOAc) to give silylated lactone as a colorless oil (7.8 g, 72%.) ^1H NMR ($CDCl_3$) δ 4.53-4.50 (m, 1 H), 4.49-4.33 (m, 1 H), 3.84-3.74 (m, 2 H), 2.82 (dd, J = 3.3, 17.7 Hz, 1 H), 2.49 (dd, J = 2.7, 18.0 Hz, 1 H), 0.89 (s, 9 H), 0.87 (s, 9 H), 0.10-0.02 (m, 12 H). To a -78 °C solution of lactone (5.83 g, 16.2 mmol) in toluene (81 mL) was added DIBAL-H (19.4 mL, 1.0 M in toluene, 19.4 mmol) over 1 h. The reaction was stirred for 1 h and quenched with MeOH (10 mL). The solution was warmed to room temperature and diluted with Et_2O (100 mL). The solution was washed with saturated Rochelle's salt-potassium sodium tartrate (100 mL). The cloudy emulsion was reextracted with Et_2O (3 x 50 mL) and the combined organic was dried over $MgSO_4$. The solution was concentrated to dryness and purified by column chromatography on 70 g of silica gel (elution with 2% EtOAc in hexanes to 10% EtOAc) to give the product (4.98 g, 85%) as

a colorless oil in a 3:1 ratio of diastereomers. ^{1}H NMR ($CDCl_3$) δ 5.53-5.44 and 5.41 (first as m, second as dd, J = 4.8, 12.0 Hz, 1 H), 4.52 and 4.22 (each as dd, J = 6.0, 10.1 Hz, and J = 4.2, 7.4 Hz, 1 H), 4.40 (d, 3.9 Hz, 1 H), 4.01 (d, J = 12.0 Hz, 1 H), 3.78-3.57 (m, 1 H), 3.64 and 3.32 (each as dd, J = 3.9, 10.7 Hz, and J = 4.2, 7.4 Hz, 1 H), 2.16-1.93 (m, 2 H), 0.91 (s, 18 H), 0.13 (s, 6 H), 0.06 (s, 6 H); ^{13}C NMR ($CDCl_3$) δ 99.96, 99.22, 87.50, 87.35, 74.22, 72.36, 63.88, 63.60, 45.18, 41.50, 26.17, 26.03, 18.62, 18.16, -4.31, -4.44, -4.63, -5.11, -5.22; IR (film) 3430, 2960, 2927, 2851, 1258, 1115, 1086, 839 cm^{-1}.



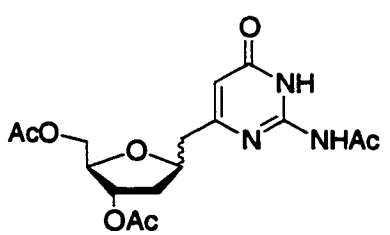
Preparation of 27: A mixture of **26** (4.80 g, 13.2 mmol) and (carbethoxymethylene)triphenylphosphorane (5.06 g, 14.5 mmol) in toluene (35 mL) was heated to 90 °C for 10 h. The solution was cooled to room temperature and concentrated. The crude material was purified by column chromatography on 60 g of silica gel (elution with 2% EtOAc in hexanes to 5% EtOAc) to give a mixture of *E* and *Z* isomers of the acyclic product (5.62 g, 98%). The acyclic material was dissolved in absolute EtOH (50 mL) and treated with 0.17 M NaOEt (2.5 mmol) in EtOH (15 mL). After stirring for 2 h, the NaOEt was quenched with 0.15 g NH_4Cl and concentrated. The resulting oil was dissolved in Et_2O (50 mL), washed with water (20 mL), and dried over $MgSO_4$. The solution was filtered and concentrated to give **27** as a 3:1 mixture of diastereomers (5.4 g, 94%). ^{1}H NMR ($CDCl_3$) δ 4.56-4.47 (m, 1 H), 4.38-4.34 and 4.32-4.27 (each as m, 1 H), 4.18-4.10 (m, 2 H), 3.91-3.86 and 3.82-3.78 (each as m, 1 H), 3.65-3.59 (m, 1 H), 3.53-3.41 (m, 1 H), 2.77 and 2.45 (each as dd, J = 6.9, 15.3 Hz, and J = 6.0, 15.5 Hz, 1 H), 2.64-2.55 (m, 1 H), 2.29-2.25 and 1.94 (first as m, second as ddd, J = 1.8, 6.4, 12.9 Hz, 1 H), 1.73-1.63 (m, 1 H), 1.26 and

1.25 (each as t, J = 6.9 Hz, and J = 7.2 Hz, 3 H), 0.89 (s, 18 H), 0.06 (s, 6 H), 0.05 (s, 6 H); ^{13}C NMR ($CDCl_3$) δ 171.57, 171.21, 87.93, 87.01, 75.61, 74.76, 74.12, 73.75, 63.93, 63.75, 60.64, 60.49, 41.74, 41.20, 40.96, 40.40, 26.23, 26.08, 18.66, 18.29, 14.49, -4.35, -4.39, -4.46, -4.98, -5.09; IR (film) 2952, 2851, 1740, 1740, 1262, 1099, 839 cm^{-1}; HRMS FAB (M+H) for $C_{21}H_{45}O_5Si_2$: calcd. 433.2806, found 433.2801.



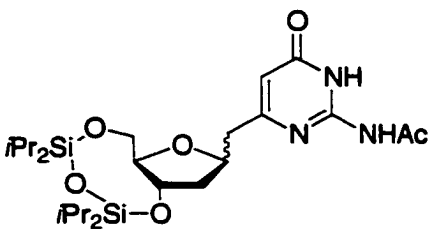
Preparation of 28: To a solution of diisopropylamine (1.71 g, 16.7 mmol) in THF (8 mL) at 0 °C was added *n*-butyl lithium (2.0 M in hexanes, 8.1 mL, 16.2 mmol). The LDA solution was cooled to -78 °C and *t*-butyl acetate (1.88 g, 16.2 mmol) was added in THF (8 mL). The solution was stirred for 30 min and **27** (2.34 g, 5.4 mmol) was added in THF (7 mL). The solution was allowed to slowly warm to 0 °C and stirred for 5 h. The reaction was quenched with water (1 mL), diluted with Et_2O (50 mL), and washed with saturated NH_4Cl (30 mL). The aqueous phase was reextracted with Et_2O (2 x 20 mL) and the combined organic phases were washed with brine (30 mL). The solution was dried over $MgSO_4$ and concentrated. The resulting oil was purified by column chromatography on 25 g of silica gel (elution with 5% EtOAc in hexanes to 10% EtOAc) to give a mixture of the product that is contaminated with some polyalkylated material (2.58 g, 95%). ^1H NMR ($CDCl_3$) δ 4.54-4.45 (m, 1 H), 4.31-4.27 (m, 1 H), 3.81-3.78 (m, 1 H), 3.65-3.59 (m, 1 H), 2.42 (s. 2 H), 3.48-3.36 (m, 1 H), 2.85-2.77 (m, 1 H), 2.73-2.66 (m, 1 H), 1.96 (ddd, J = 1.8, 4.8, 12.3 Hz, 1 H), 1.65-1.60 (m, 1 H), 1.47 (s, 9 H), 0.90 (s, 18 H), 0.06 (s, 12 H); IR (film) 2960, 2935, 2856, 1735, 1714, 1254, 1145, 831 cm^{-1}. To a suspension of Na_2CO_3 (0.69 g, 6.5

mmol) in absolute EtOH (30 mL) was added guanidine•HCl (0.62 g, 6.5 mmol) and the above β-ketoester (1.64 g, 3.3 mmol contaminated with some polyalkylated material). The suspension was heated to 75 °C for 60 h and cooled to room temperature. The reaction was concentrated to dryness, resuspended in EtOAc (50 mL), and washed with water (50 mL). The aqueous layer was reextracted with EtOAc (3 x 20 mL) and the combined organic phase was washed with brine (30 mL). The solution was dried over $MgSO_4$ and concentrated to leave a brown foam. The foam was purified by column chromatography on 20 g of silica gel (elution with 2% MeOH in DCM to 10 % MeOH) to give an off-white foam as a 3:1 mixture of diastereomers (1.00 g, 65%). ^{1}H NMR ($CDCl_3$) δ 6.91 (bs, 2 H), 5.63 (s, 1 H), 4.59-4.50 and 4.46-4.39 (each as m, 1 H), 4.36-4.31 and 4.29-4.26 (each as m, 1 H), 3.93-3.86 and 3.84-3.78 (each as m, 1 H), 3.62-3.57 (m, 1 H), 3.53-3.43 (m, 1 H), 2.85-2.77 and 2.64-2.56 (each as m, 1 H), 2.62-2.52 (m, 1 H), 2.28-2.19 and 1.93-1.87 (each as m, 1 H), 1.76-1.65 (m, 1 H), 0.87 (s, 18 H), 0.05-0.03 (m, 12 H); ^{13}C NMR ($CDCl_3$) δ 156.39, 88.04, 86.88, 73.94, 73.44, 63.95, 63.54, 41.04, 40.27, 26.08, 25.97, 18.16, 18.47, -4.44, -4.50, -4.59, -5.13, -5.24; IR (film) 3320, 3137, 2928, 2860, 1656, 1613, 1519, 1468, 1387, 1105, 836 cm^{-1}; HRMS FAB (M+H) for $C_{22}H_{44}N_3O_4Si_2$: calcd. 470.2870, found 470.2865.



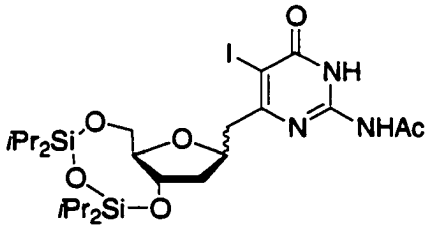
Preparation of 30: To a solution of **28** (0.99 g, 2.1 mmol) in THF (42 mL) was added Et_3N•3HF (3.38 g, 21.0 mmol). The reaction was allowed to stir for 12 h at room temperature and concentrated to dryness. The resulting oil

was coevaporated with pyridine (10 mL). The brown oil was resuspended in pyridine (11 mL) followed by the addition of DMAP (0.05g, 0.4 mmol) and acetic anhydride (2.18 g, 21.0 mmol). The mixture was reacted for 5 h at room temperature and concentrated. The resulting oil was dissolved in EtOAc (20 mL) and washed with 10 % HCl (10 mL), saturated $NaHCO_3$ (10 mL), and brine (10 mL). The organic phase was dried over $MgSO_4$ and concentrated. The product was purified by column chromatography on 20 g of silica gel (elution with 100% DCM to 1% MeOH) to give a 2.33:1 mixture of diastereomers (0.44 g, 57%). ^{1}H NMR ($CDCl_3$) δ 11.66 (b, 1 H), 6.03 and 6.01 (each as s, 1 H), (5.14-5.09 (m, 1 H), 4.53-4.48 and 4.44-4.35 (each as m, 1 H), 4.26-4.23 and 4.20-4.09 (each as m, 1 H), 2.85-2.65 (m, 2 H), 2.52-2.46 and 2.13-1.99 (each as m, 1 H), 2.31 (s, 3 H), 2.06 (s, 3 H), 2.05 (s, 3 H), 1.92-1.81 (m, 1 H); ^{13}C NMR ($CDCl_3$) δ 176.78, 172.61, 172.58, 170.41, 170.41, 150.72, 108.65, 82.68, 81.62, 77.14, 76.93, 76.67, 76.06, 64.34, 64.08, 38.13, 37.17, 25.84, 24.71, 21.17, 21.04, 21.00; IR (film) 3184, 2950, 1735, 1663, 1612, 1565, 1374, 1238, 1047 cm^{-1}; HRMS FAB (M + H) for $C_{16}H_{22}N_3O_7$: calcd. 368.1458, found 368.1460.

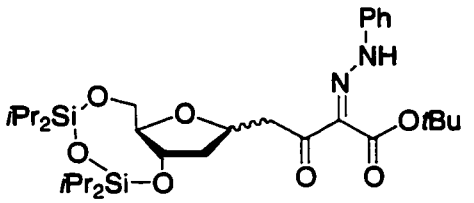


Preparation of 32: To a solution of **24** (1.07 g, 2.21 mmol) in DCM (19 mL) was added pyridine (0.70 g, 8.84 mmol) and acetic anhydride (2.29 g, 22.10 mmol). The solution was stirred for 18 h at room temperature and concentrated to an oil. The oil was taken up in EtOAc (30 mL) and washed with saturated $NaHCO_3$ (15 mL), brine (15 mL), and dried over $MgSO_4$. The solution was concentrated to an off-white foam and chromatographed on 40 g of silica gel

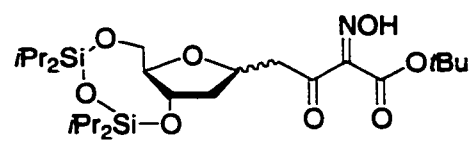
(elution with 50% EtOAc in DCM to 100% EtOAc) to give a white foam as a mixture of α- and β- anomers (1.08 g, 93%). mp= 98-102 °C; ^{1}H NMR (DMSO-d_6) δ 11.70 (b, 2 H), 5.92 and 5.88 (each as, 1 H), 4.42-4.26 (m, 1 H), 4.06-3.88 (m, 2 H), 3.74-3.63 (m, 1 H), 3.61-3.48 (m, 1 H), 2.78-2.48 (m, 2 H), 2.38-2.30 and 1.82-1.76 (each as m, 1 H), 2.14 (s, 3 H), 1.75-1.62 (m, 1 H), 1.08-0.91 (m, 28 H); IR (film) 3061, 2944, 2867, 1712, 1680, 1574, 1464, 1252, 990, 885, 696 cm^{-1}; LRMS FAB (M+H) for $C_{24}H_{44}N_3O_6Si_2$: calcd. 526.3, found 526.2.



Preparation of 33: To a solution of **32** (0.18 g, 0.35 mmol) in AcN (3 mL) was added ICl (0.53 mL, 1.0 M, 0.53 mmol) in DCM. The reaction was refluxed for 5 h and ICl (0.25 mL, 1.0 M, 0.25 mmol) was added. The reaction was refluxed for 2 h and diluted with DCM (10 mL). The solution was washed with 5% $NaHSO_3$ (5 mL) and the aqueous layer was reextracted with DCM (3 x 5 mL). The combined organic was successively washed with water (5 mL) and brine (5 mL). The organic layer was dried over $MgSO_4$ and concentrated. The resulting oil was chromatographed on 10 g of silica gel (elution with 50% EtOAc in hexanes to 100% EtOAc) to give a colorless oil as a mixture of anomers (0.14 g, 61%). ^{1}H NMR ($CDCl_3$) δ 12.14 (b, 1 H), 9.10 (b, 1 H), 4.76-4.39 (m, 2 H), 4.08-3.96 (m, 1 H), 3.90-3.76 (m, 2 H), 3.38-3.24 and 3.19-3.08 (each as m, 1 H), 2.96-2.84 (m, 1 H), 2.38 and 2.32 (each as s, 3 H), 2.21-2.16 and 2.10-2.04 (each as m, 1 H), 2.02-1.90 (m, 1 H), 1.08-0.86 (m, 28 H); IR (film) 3163, 2944, 2857, 1645, 1611, 1563, 1462, 1141, 1071, 1035 cm^{-1}; HRMS FAB (M+H) for $C_{24}H_{42}N_3O_6Si_2I$: calcd. 652.1735, found 652.1728.

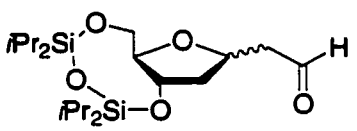


Preparation of 39: To a solution of 6 N HCl (0.5 mL) and $NaNO_2$ (0.04 g, 0.6 mmol) was added aniline (0.06 g, 0.6 mmol) at 0 °C. The solution was stirred for 15 min and added to a solution of **23** (0.31 g, 0.6 mmol), in 95% EtOH (19 mL) with NaOAc (0.25 g, 2.99 mmol). The solution immediately turned yellow and was stirred for 30 min and poured into DCM (50 mL). The solution was washed with 10% HCl (2 x 10 mL), saturated $NaHCO_3$ (10 mL), and brine (10 mL). The organic phase was dried over $MgSO_4$ and concentrated to leave an orange oil. The oil was chromatographed on 10 g of silica gel (elution with 5% EtOAc in hexanes to 20% EtOAc) to give an orange oil as a mixture of anomers (0.32 g, 75%). ^{1}H NMR ($CDCl_3$) δ 14.68 and12.79 (each as bs, 1 H), 7.42-7.35 (m, 4 H), 7.22-7.18 (m, 1 H), 4.62-4.58 (m, 1 H), 4.52-4.42 (m, 1 H), 4.02-3.96 (m, 1 H), 3.80-3.75 (m, 2 H), 3.49-3.39 (m, 1 H), 3.26-3.18 and 3.08-2.99 (each as m, 1 H), 2.57-2.43 and 2.23-2.12 (each as m, 1 H), 1.98-1.82 (m, 1 H), 1.58 (s, 9 H), 1.09-0.87 (m, 28 H); IR (film) 2944, 2867, 1702, 1524, 1246, 1152, 1035, 885, 692 cm^{-1}; LRMS FAB (M+H) for $C_{31}H_{53}N_2O_7Si_2$: calcd. 621.3, found 621.3.

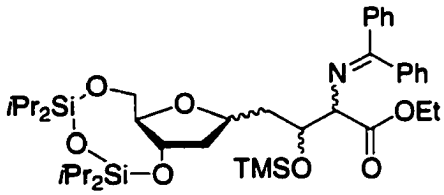


Preparation of 40: To a solution of **23** (0.42 g, 0.81 mmol) in absolute EtOH (2 mL) was added isopentyl nitrite (0.29 g, 2.43 mmol) and 0.67 M NaOEt (2 mL, 1.33 mmol) at 0 °C. The solution was warmed to room temperature and stirred for 18 h. The reaction was concentrated to an oil and taken up EtOAc (10 mL). The solution was washed with 10% HCl (10 mL),

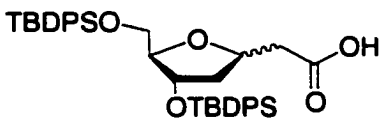
water (10 mL), brine (10 mL), and dried over $MgSO_4$. The solution was then filtered and concentrated in vacuo. The crude product was chromatographed on 30 g of silica gel (elution with 10% EtOAc in hexanes to 30% EtOAc) to give a complex mixture of transesterified products (0.30 g, 68.0%), which were carried on without further purification. ^{1}H NMR ($CDCl_3$) δ 9.37 (b, 1 H), 4.47-4.43 (m, 1 H), 4.39-4.32 (m, 1 H), 3.97-3.92 (m, 1 H), 3.78-3.68 (m, 1 H), 3.38-3.33 and 3.01-2.92 (each as m, 1 H), 3.19 (dd, J = 3.6, 9.0 Hz, 1 H), 3.78 (dd, J = 2.4, 9.0 Hz, 1 H), 1.96-1.82 (m, 2 H), 1.57 (s, 9 H), 1.12-0.92 (m, 28 H); IR (film) 3283, 2945, 2868, 1736, 1696, 1464, 1394, 1370, 1157, 919, 885, 693 cm^{-1}.



Preparation of 47: To a solution of **22** (1.03 g, 2.3 mmol) in 9 mL of toluene was added DIBAL-H (2.8 mL, 1.0 M in toluene, 2.8 mmol) at -78 °C. The solution was stirred for 2 h, quenched with water (1 mL), and warmed to room temperature. The solution was diluted with Et_2O (20 mL) and washed with 10% HCl (10 mL). The aqueous phase was reextracted with Et_2O (3 x 10 mL). The combined organic phase was washed with brine (20 mL), dried over $MgSO_4$ and concentrated. The resulting oil was chromatographed on 50 g of silica gel (elution with 10% EtOAc in hexanes to 25% EtOAc) to give a mixture of α- and β-anomers (0.64 g, 69%) as a colorless oil. ^{1}H NMR ($CDCl_3$) δ 9.81 (s, 1 H), 4.58-4.40 (m, 2 H), 4.03-3.98 (m, 1 H), 3.80-3.76 (m, 1 H), 2.84 and 2.68 (each as dd, J = 3.6, 8.4 Hz, and J = 3.0, 7.2 Hz, 1 H), 3.68-3.62 and 2.47-2.41 (each as m, 1 H), 2.21-2.18 and 1.90-1.86 (each as m, 1 H), 1.84-1.74 (m, 1 H), 1.09-0.94 (m, 28 H), IR (film) 2867, 2750, 1727, 1464, 1035, 885, 692 cm^{-1}.

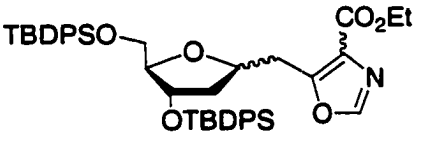


Preparation of 48: To a solution of diisopropylamine (0.076 g, 0.75 mmol) in THF (2 mL) at 0 °C was added *sec*-butyl lithium (0.55 mL, 1.26 M in hexanes, 0.69 mmol). The solution was then cooled to -78 °C and *N*-(diphenylmethylene)glycine ethyl ester (0.17 g, 0.69 mmol) was added in THF (2 mL). The solution was stirred for 10 min and TMSCl (0.23 g, 2.1 mmol) was added. The solution was warmed to 0 °C, and **47** (0.22 g, 0.53 mmol) was added in THF (2 mL) with $ZnCl_2$ (4 mg, 0.03 mmol). The solution was allowed to warm to room temperature and was stirred an additional 3 h. The solution was diluted with hexanes (25 mL), filtered over a pad of celite, and concentrated. The resulting oil was chromatographed on 10 g of silica gel (prewashed with 0.2% Et_3N and elution with 3% EtOAc in hexanes to 20% EtOAc) to give a mixture of α- and β- anomers (0.29 g, 74%) as a colorless oil. ^1H NMR ($CDCl_3$) δ 7.68-7.63 (m, 2 H), 7.42-7.29 (m, 6 H), 7.20-7.16 (m, 2 H), 4.41-4.26 (m, 2 H), 4.18-4.09 (m, 4 H), 4.08-3.92 (m, 2 H), 3.78-3.64 (m, 2 H), 2.38-2.24 and 2.04-1.98 (each as m, 1 H), 1.96-1.66 (m, 2 H), 1.30-1.21 (m, 3 H), 1.09-0.91 (m, 28 H), 0.03 and 0.06 (each as s, 9 H); IR (film) 2945, 2867, 1740, 1624, 1033, 885, 695 cm^{-1}, LRMS FAB (M+H) for $C_{39}H_{64}NO_7Si_3$: calcd. 742.4, found 742.7.



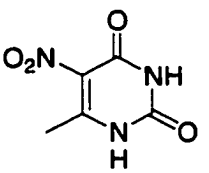
Preparation of 51: A solution of **50**[183] (1.0 g, 1.5 mmol) was taken up in a 0.18 M LiOH (10 mL) solution in MeOH. The solution was refluxed for 10 h and concentrated. The residue was taken up in EtOAc (10 mL) and washed with water (10 mL). The aqueous phase was reextracted

with EtOAc (4 x 10 mL). The combined organic phases were washed with brine (20 mL), dried over $MgSO_4$, and concentrated. The resulting oil was concentrated and chromatographed on 30 g of silica gel (elution with 20% ethyl acetate in hexanes to 50% EtOAc) to give a mixture of α- and β- anomers (0.74 g, 80%). ^1H NMR ($CDCl_3$) δ 7.78-7.29 (m, 20 H), 4.73-4.67 and 4.56-4.51 (each as m, 1 H), 4.53-4.48 (m, 1 H), 4.13-4.10 and 4.06-4.03 (each as m, 1H), 3.48-3.40 (m, 1 H), 3.26 (dt, J = 3.6, 11.1 Hz, 1 H), 2.97 and 2.73 (each as dd, J = 8.1, 15.6 Hz, and J = 5.4, 10.1 Hz, 1 H) (2.73, 2.64-2.62 (m, 1 H), 2.27-2.22 and 1.85-1.79 (each as m, 1 H), 2.08-2.02 and 1.67-1.57 (each as m, 1 H), 1.16 (s, 9 H), 0.94 (s, 9 H); IR (film) 3069, 2933, 2856, 1711, 1467, 1427, 1109, 703 cm^{-1}.

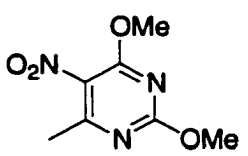


Preparation of 52: To a solution of **51** (0.15 g, 0.25 mmol) in DCM (1 mL) was added $SOCl_2$ (0.60 g, 5 mmol). The solution was refluxed for 3 hours, concentrated, and coevaporated with THF (3 x 1 mL). In a separate flask, *n*-butyl lithium (0.15 mL, 1.90 M in hexanes, 0.29 mmol) was diluted with THF (1 mL) at -78 °C. To this solution was added **46** (0.03 g, 0.28 mmol) dropwise. The solution was stirred for 5 min and the acid halide formed above was added to the enolate solution in THF (1.5 mL). The solution was stirred at -78 °C for 2 hours and allowed to warm to 0 °C. The reaction was stirred overnight and concentrated to an oil. The oil was taken up in of EtOAc (10 mL) and washed with water (5 mL). The aqueous phase was reextracted with EtOAc (3 x 5 mL), and the combined organic phases were washed with brine (10 mL), dried over $MgSO_4$, and concentrated. The resulting oil was chromatographed on 10 g of silica gel

(elution with 5% EtOAc in hexanes to 20% EtOAc) to give an anomeric mixture of products as an oil (0.09 g, 47%). ^{1}H NMR ($CDCl_3$) δ 7.76 and 7.74 (each as s, 1 H), 7.66-7.26 (m, 20 H), 4.64-4.58 (m, 1 H), 4.54-4.49 and 4.47-4.43 (each as m, 1 H), 4.35 (q, J = 3.6 Hz, 2 H), 4.12 (m, 0.5 H), 4.09-4.06 and 3.99-3.97 (each as m, 1 H), 3.59-3.56 and 3.48-3.44 (each as m, 1 H), 3.36-3.22 (m, 3 H), 1.98-1.92 (m, 1 H), 1.68-1.62 (m, 1 H), 1.36 (t, J = 3.6 Hz, 3 H), 1.09 and 1.08 (each as s, 9 H), 0.99 and 0.96 (each as s, 9 H); LRMS FAB (M+H) for $C_{44}H_{53}NO_6Si_2$: calcd. 748.3, found 748.5.

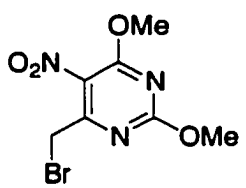


Preparation of 58:[191] A three neck flask equipped with a mechanical stirrer was charged with conc. H_2SO_4 (840 mL) and conc. HNO_3 (120 mL) and cooled to 0 °C. Addition of **57** (120.5 g, 956 mmol) to the solution was done in small portions over 1 h. After 3 h the reaction was allowed to warm to room temperature. The solution was poured unto ice water (2.5 L) and filtered to give an off-white solid (123.8 g, 76%) ^{1}H NMR (DMSO-*d6*) δ 11.79 (b, 1 H), 11.76 (b, 1 H), 2.28 (s, 3 H); ^{13}C NMR (DMSO-*d6*) δ 156.38, 154.18, 149.11, 127.21, 16.69; mp >214 °C decomp.



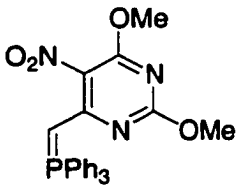
Preparation of 59:[191,192] To a solution of **58** (40.7 g, 238 mmol) in $POCl_3$ (438 g, 2.9 mol) was added diethylaniline (53.3 g, 356 mmol). The solution was equipped with a drying tube and heated at reflux for 24 h. The solution was cooled to room temperature, concentrated, and poured onto ice water (500 mL). The aqueous solution was extracted with Et_2O (5 x 250 mL). The combined organic layers were washed with water (300 mL), saturated $NaHCO_3$ (300

mL), and brine (200 mL). The solution was dried over $MgSO_4$ and concentrated to a brown solid (43.9 g, 89%), which was further reacted. ^1H NMR ($CDCl_3$) δ 2.63 (s, 3 H). To a solution of NaOMe (8.6 g, 160 mmol) in MeOH (50 mL) was added the dichloropyrimidine from above (13.3 g, 63.9 mmol) in MeOH (80 mL). The solution was warmed to reflux and stirred for 2 hours. The solution was cooled, filtered, and concentrated to leave an orange solid. The solid was taken up in $CHCl_3$ (100 mL) and washed with water (100 mL). The aqueous layer was reextracted with $CHCl_3$ (4 x 30 mL). The combined organic layers were washed with brine (40 mL), dried over $MgSO_4$ and concentrated. The resulting brown solid (11.8 g, 93%) was taken on without further purification. ^1H NMR ($CDCl_3$) δ 4.09 (s, 3 H), 4.05 (s, 3 H), 2.52 (s, 3 H). IR (film) 1573, 1518, 1496, 1377, 1347, 1095, 792 cm^{-1}; mp=81-82 °C.

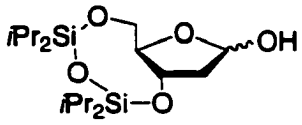


Preparation of 60:[192] To a solution of **59** (8.07 g, 40.5 mmol) and NaOAc (6.70 g, 81.1 mmol) in glacial AcOH (200 mL) was added Br_2 (4.03 g, 25.0 mmol). The reaction was refluxed for 45 min, cooled to room temperature and Br_2 (3.10 g, 19. 4 mmol) was added and returned to reflux for 1 h and cooled. Concentration in vacuo resulted in an orange paste, which was dissolved in EtOAc (100 mL). The solution was filtered and concentrated. The resulting solid was chromatographed on a 100 g plug of silica gel (elution with DCM). The mixture of monobromide, dibromide, and starting material was then separated using a 4 mm chromatotron plate (elution with 80 % CH_2Cl_2 in hexanes). The first fraction to elute contained **61** (6.24 g, 40 %) and the second fraction possessed **60** (4.37 g, 38 %) ^1H

NMR ($CDCl_3$) δ 4.49 (s, 2 H), 4.13 (s, 3 H), 4.09 (s, 3 H); IR (film) 1566, 1524, 1493, 1338, 1206 cm^{-1}.



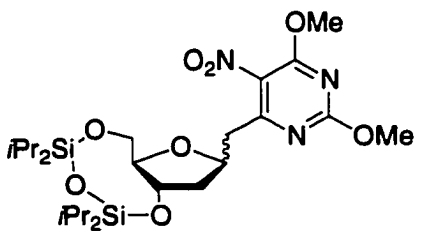
Preparation of 56:[187] To a solution of **60** (6.77 g, 24.3 mmol) in benzene (60 mL) was added PPh_3 (9.0 g, 34.1 mmol) in benzene (40 mL) dropwise over 5 min. The solution was stirred for 4.5 h and filtered. The phosphonium salt was washed 2 times with benzene (50 mL). The phosphonium salt was dissolved in a biphasic mixture of of $CHCl_3$ (160 mL) and ice water (80 mL). The mixture was stirred vigorously and 8.2 mL of 3 N NaOH (24.6 mmol) was added. The reaction was stirred for 1.5 h and the two phases were separated. The aqueous phase was extracted with $CHCl_3$ (3 x 100 mL). The combined organic phases were washed with brine (50 mL), dried over $MgSO_4$, filtered, and concentrated. The resulting orange solid was recrystallized from EtOH to give an orange powder (10.8 g, 96.7%). ^1H NMR ($CDCl_3$) δ 2.82 (s, 3 H), 3.99 (s, 3 H), 5.28 (d, J = 29 Hz, 1 H), 7.50 (m, 6 H), 7.64 (m, 9 H); ^{31}P NMR ($CDCl_3$) δ 17.06; IR (film) 3055, 2949, 1567, 1523, 1360, 1253, 1131, 1102 cm^{-1}.



Preparation of 62: To a solution of 2-deoxyribose (5.00 g, 37.3 mmol) and $CaCO_3$ (3.73 g, 37.3 mmol) in water (15 mL) was added Br_2 (15.5 g, 97.0 mmol) dropwise at 0 °C. The exothermic reaction was allowed to stir for 5 min and sealed. After 48 h, the reaction was purged with N_2, diluted with water (100 mL), and treated with 5 g of oxalic acid. The solution was then neutralized with Ag_2CO_3 (20 g) to give a green solution. The solution was filtered twice to remove the

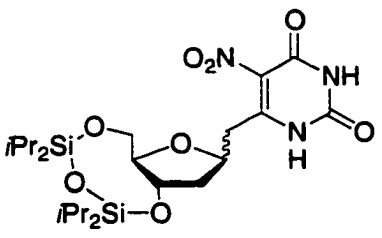
silver salts and concentrated. The resulting oil was coevaporated with pyridine (3 x 10 mL). Following coevaporation, imidazole (10.2 g, 149.2 mmol) was added and the mixture was coevaporated with pyridine (2 x 10 mL). The oil was taken up in DMF (190 mL) and TIPDSCl$_2$ (17.2 g, 44.8 mmol) was added. The reaction was stirred for 24 h and diluted with water (200 mL). The product was partitioned between the aqueous layer and Et_2O (200 mL). The aqueous layer was reextracted with Et_2O (2 x 200 mL). The combined organic layer was washed with water (100 mL), saturated $NaHCO_3$ (100 mL), and brine (100 mL). The organic phase was dried over $MgSO_4$ and concentrated to give a brown oil. The oil was purified by column chromatography with 75 g of silica gel (elution with 5 % EtOAc in hexanes to 20% EtOAc) to give silylated lactone as a colorless oil (11.3 g, 82%).[193] ^{1}H NMR ($CDCl_3$) δ 4.67-4.59 (m, 1 H), 4.21 (m, 1 H), 4.14 (dd, 1 H, *J*= 12.3, 3.6 Hz), 3.92 (dd, 1 H, *J*= 12.3, 6.7 Hz), 2.85 (dd, 1 H, *J*= 17.3, 8.0 Hz), 2.70 (dd, 1 H, *J*= 17.3, 9.2 Hz), 1.13-0.98 (m, 28 H). To a solution of the protected lactone (5.32 g, 14.2 mmol) in toluene (71 mL) at -78°C was added DIBAL-H (18.5 mL, 1.0 M in toluene) over 1.5 h. After 1.5 h the reaction was quenched with MeOH (8 mL), warmed to room temperature, and diluted with Et_2O (300 mL). The solution was washed with Rochelle's salt (100 mL) and the aqueous phase was extracted with Et_2O (3 x 50 mL). The combined organic phase was washed with brine (50 mL), dried over $MgSO_4$, and concentrated. The resulting oil was purified by column chromatography (5% EtOAc in hexanes to 20% EtOAc) to give **62** (4.59 g, 86%). ^{1}H NMR ($CDCl_3$) δ 5.42-5.48 (m, 1 H), 4.72-4.65 and 4.45-4.39 (each as m, 1 H), 4.12-4.06 (m, 1 H), 4.02-3.97 and 3.88-3.80 (each as m, 1 H), 3.68 and 3.29 (first as dd, *J* = 12.0, 7.5 Hz, second as bd s, 1 H), 3.17 (bd s, 1 H), 2.39-2.26 (m, 1 H), 2.14-1.97 (m, 1 H),

1.12-0.86 (m, 28 H); ^{13}C NMR ($CDCl_3$) δ 98.06, 97.53, 84.88, 84.04, 73.41, 72.84, 65.92, 63.23, 42.52, 41.77, 17.83, 17.70, 17.67, 17.61, 17.51, 17.38, 17.33, 17.26, 17.19, 13.70, 13.66, 13.57, 13.29, 13.08, 12.83, 12.77; IR (film) 3421, 2945, 2868, 1465 cm^{-1}.

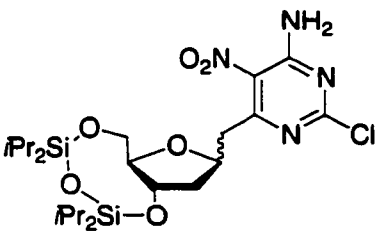


Preparation of 63: A solution of **62** (4.51 g, 12.0 mmol) and **56** (5.50 g, 12.0 mmol) in toluene (40 mL) was heated to reflux for 65 h. After removing the toluene in vacuo, the residue was stirred in NaOMe (0.13 g, 2.4 mmol) and MeOH (60 mL) for 1 h. The reaction was filtered, neutralized with 6 M HCl (0.4 mL), concentrated, and chromatographed (100% hexanes to 5% EtOAc) to give a mixture of epimers of **63** (5.59 g, 84%). A small portion of the mixture was separated by chromatatron (2 mm plate, elution with 2% EtOAc in hexanes) to give the α- and β- anomers of **63**. First fraction: β-**63**: ^{1}H NMR ($CDCl_3$) δ 4.53-4.46 (m, 1 H), 4.37-4.32 (m, 1 H), 4.04 (s, 3 H), 4.0 (s, 3 H), 3.94 (dd, *J* = 8.1, 2.1 Hz, 1 H), 3.77-3.59 (m, 2 H), 3.03 (dd, 1 H, *J* = 13.8, 7.2 Hz), 2.86 (dd, *J* = 13.8, 3.0 Hz, 1 H), 2.12-2.04 (m, 1 H), 1.95-1.86 (m, 1 H), 1.05-0.84 (m, 28 H); ^{13}C NMR ($CDCl_3$) δ 163.39, 163.23, 163.00, 130.75, 86.23, 75.68, 74.01, 64.14, 55.83, 55.45, 40.42, 39.48, 17.72, 17.56, 17.47, 17.20, 17.13, 13.63, 13.57, 13.13, 12.76; IR (film) 2945, 1574, 1464 cm^{-1}; HRMS FAB (M+H) calcd. for $C_{24}H_{44}N_3O_8Si_2$: 558.2667, found 558.2659. Second Fraction: α-**63**: ^{1}H NMR ($CDCl_3$) δ 4.53-4.48 (m, 1 H), 4.42-4.37 (m, 1 H), 4.06 (s, 3 H), 4.03 (s, 3 H), 3.93 (dd, *J* = 11.4, 2.4 Hz,1 H), 3.75 (d, *J* = 6.0 Hz, 1 H), 3.68-3.73 (m, 1H), 3.31 (dd, *J* = 13.6, 8.8 Hz, 1H), 2.87 (dd, *J* = 13.6, 4.8 Hz, 1 H), 2.41 (dt, *J* = 13.6, 4.8 Hz, 1 H), 1.85 (dt, *J* = 12.4, 7.2 Hz, 1 H), 1.07-0.91 (m, 28 H); ^{13}C NMR ($CDCl_3$) δ 163.73,

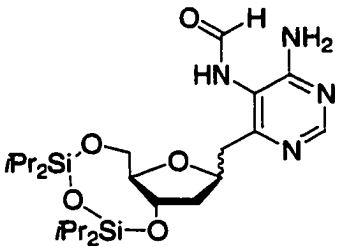
163.52, 163.32, 131.05, 83.30, 75.93, 72.62, 62.62, 55.88, 55.53, 40.01, 17.76, 17.61, 17.56, 17.38, 17.29, 17.20, 13.72, 13.50, 13.13, 12.81; IR (film) 2945, 2867, 1586, 1531, 1464 cm^{-1}; HRMS FAB (M+H) cacld. for $C_{24}H_{44}N_3O_8Si_2$: 558.2667, found 558.2675.



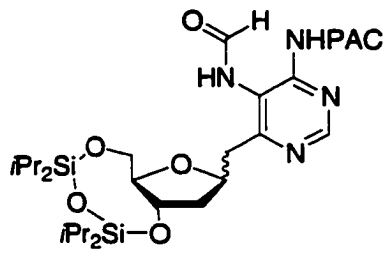
Preparation of 64: TMSCl (3.17 g, 28.9 mmol) was added to a solution of a mixture of anomers of **63** (3.22 g, 5.77 mmol) and NaI (8.6 g, 57.7 mmol) in $CHCl_3$ (58 mL). The suspension was stirred for 9 h, diluted with EtOAc (50 mL), and filtered. The organic phase was washed with 5% HCl. The aqueous phase was extracted with EtOAc (2 x 20 mL). The combined organic phase was washed sequentially with saturated $NaHSO_3$ (20 mL), saturated $NaHCO_3$ (20 mL), brine (20 mL), dried over $MgSO_4$, and concentrated. The resulting oil was purified by column chromatography (20% EtOAc in hexanes to 50% EtOAc) to give **64** as a mixture of anomers (2.25 g, 74%). ^1H NMR ($CDCl_3$) δ 10.28 (b, 2 H), 4.49-4.38 (m, 2 H), 4.03-3.93 (m, 1 H), 3.87-3.74 (m, 2 H), 3.09 and 2.72 (each as dd, J = 15.6, 2.4 Hz, and J = 15.0, 9.6 Hz, 1 H), 2.94 (d, J = 5.7 Hz, 1 H,), 2.55-2.47 and 2.28-2.20 (each as m, 1 H), 1.99-1.82 (m, 1 H), 1.07-0.95 (m, 28 H); ^{13}C NMR ($CDCl_3$) δ 156.71, 156.57, 153.42, 152.90, 149.54, 148.97, 128.58, 128.49, 89.87, 83.32, 74.97, 74.83, 72.30, 71.31, 62.93, 61.56, 40.20, 39.90, 36.29, 35.74, 17.73, 17.68, 17.56, 17.47, 17.34, 17.28, 17.25, 17.16, 13.85, 13.68, 13.50, 13.37, 13.18, 13.02, 12.75; IR (film) 3190, 3077, 2945, 2868, 1730, 1690, 1529 cm^{-1}; HRMS FAB (M+H) calcd. for $C_{22}H_{40}N_3O_8Si_2$: 530.2354, found 530.2347.



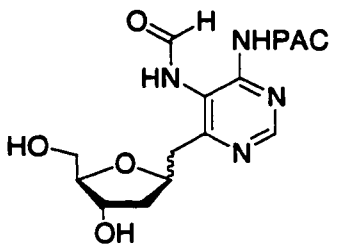
Preparation of 66: To a solution of **64** (1.72 g, 3.30 mmol) in toluene (33 mL) was added diethylaniline (5.42 g, 36.3 mmol) and $POCl_3$ (5.01 g, 33.3 mmol). The solution was heated to reflux for 5 h, cooled to room temperature, diluted with EtOAc (50 mL), and washed with 5% HCl. The aqueous phase was extracted with EtOAc (3 x 10 mL). The combined organic phase was washed with saturated $NaHCO_3$ (20 mL), brine (20 mL), dried over $MgSO_4$, and concentrated in vacuo. The resulting black oil was separated from baseline material by column chromatography (100% hexanes) to give **65** with diethylaniline as an impurity. The crude material was combined with another batch of the same reaction and taken up in THF (20 mL). The solution was cooled to 0 °C, saturated with NH_3, and stirred for 1 h. The solution was concentrated and purified by column chromatography (5% EtOAc in hexanes to 10% EtOAc) to give 0.44 g of **66** and 0.18 g of a 1:1 mixture of regioisomers (overall 16%). ^{1}H NMR ($CDCl_3$) δ 7.01 (b, 2 H), 4.56-4.35 (m, 2 H), 3.96-3.89 (m, 1 H), 3.80-3.70 (m, 1 H), 3.68-3.63 and 3.43-3.45 (each as m, 1 H), 3.18 and 3.04 (each as dd, J = 13.8, 5.1 Hz, and J = 13.8, 4.2 Hz, 1 H), 2.63- 2.46 and 2.19-2.11 (each as m, 1 H), 2.00-1.81 (m, 1 H), 0.91-1.08 (m, 28 H); ^{13}C NMR ($CDCl_3$) δ 167.34, 167.03, 160.60, 160.53, 158.02, 129.14, 129.82, 86.04, 82.91, 75.83, 75.61, 73.65, 71.88, 63.82, 61.91, 41.97, 41.60, 40.37, 39.91, 17.72, 17.67, 17.58, 17.51, 17.33, 17.22, 17.15, 14.41, 13.84, 13.65, 13.52, 13.40, 13.13, 12.99, 12.74; IR (film) 3465, 3300, 3182, 2946, 2867, 1623, 1580, 1521 cm^{-1}; HRMS FAB (M+H) calcd. for $C_{24}H_{40}N_4O_6Si_2Cl$: 547.2175, found 547.2172.



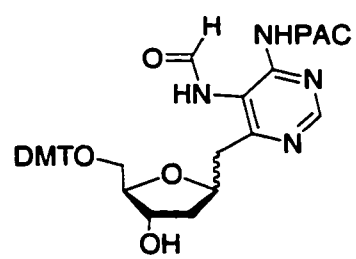
Preparation of 67: A mixture of **66** (0.28 g, 0.51 mmol), Na_2CO_3 (0.10 g, 1.02 mmol), and 70 mg of 10% Pd/C was suspended in absolute EtOH (5 mL), purged with N_2 and placed under H_2 atmosphere (latex balloon). After 12 h additional 10% Pd/C (80 mg) and Na_2CO_3 (120 mg) was added. After 24 h the reaction was filtered over a pad of celite and concentrated to leave a foam. The crude material was dried by coevaporation with pyridine (3 x 1 mL). The oil was then taken up in THF (5 mL) and pyridine (0.2 g, 2.6 mmol) was added. The solution was cooled to 0 °C and formic acetic anhydride (0.9 g, 1.02 mmol) was added. The reaction was stirred for 2 h and concentrated. The resulting oil was purified by column chromatography (50% EtOAc in hexanes to 90% EtOAc) to give **67** as a mixture of anomers (0.195 g, 75%). ^{1}H NMR ($CDCl_3$) δ 9.11 (s, 1 H), 8.44 and 8.42 (each as s, 1 H), 8.32 and 8.29 (each as s, 0.1 H), 5.41 (b, 2 H), 4.53-4.36 (m, 2 H), 4.00-3.92 (m, 1 H), 3.88-3.84 and 3.78-3.82 (first as m, second as dd, J = 12.5, 5.4 Hz, 1 H), 3.68-3.58 (m, 1H), 3.11 and 2.98-2.94 (first as dd, J = 10.2, 4.2 Hz, second as m, 1 H), 2.91-2.85 (m, 1 H), 2.41-2.31 and 2.25-2.21 (each as m, 1 H), 2.14-2.06 and 2.03-1.93 (each as m, 1 H), 1.13-0.91 (m, 28 H); ^{13}C NMR ($CDCl_3$) δ 159.76, 159.41, 158.73, 158.62, 155.85, 155.78, 115.52, 115.37, 84.84, 83.64, 78.40, 78.17, 73.25, 72.23, 71.45, 69.19, 68.58, 66.54, 62.70, 60.54, 41.61, 40.14, 39.73, 39.21, 38.41, 37.84, 17.83, 17.61, 17.51, 17.35, 17.20, 17.08, 14.64, 13.72, 13.59, 13.33, 13.15, 12.83, 12.69; IR (film), 3316, 3200, 2940, 2866, 1691, 1637, 1528 cm^{-1}; HRMS FAB(M+H) calcd. for $C_{23}H_{43}N_4O_5Si_2$: 511.2772, found 511.2763.



Preparation of 68: To a solution of phenoxyacetic acid (0.47g, 3.06 mmol) in DCM (5 mL) at 0°C was added EDCI (0.29 g, 1.53 mmol). The reaction was stirred for 30 min and then added to **67** (0.26 g, 0.51 mmol) in DCM (5 mL) at 0°C. The reaction was allowed to gradually warm to room temperature over 3 h and stirred for 60 h. The reaction was diluted with DCM (10 mL) and washed with water (10 mL). The aqueous phase was extracted with DCM (3 x 5 mL) and the combined organic phase was washed with brine (10 mL), dried over $MgSO_4$, and concentrated in vacuo. The resulting oil was purified by column chromatography (20% EtOAc in hexanes to 50% EtOAc) to give **68** (0.22 g, 67%) as a white foam. ^{1}H NMR ($CDCl_3$) δ 9.85 and 9.74 (each as s, 1 H), 9.32 and 9.27 (each as s, 1 H), 8.95 (s, 1 H), 8.30 and 8.21 (each as s, 1 H), 7.36-7.27 (m, 2 H), 7.07-6.92 (m, 3 H), 4.71-4.63 (m, 2 H), 4.54-4.42 (m, 2 H), 4.15-4.08 and 3.81-3.75 (each as m, 1 H), 3.96-3.89 (m, 1 H), 3.66-3.60 (m, 1 H), 3.24 and 3.00 (each as dd, J = 7.6, 10.8 Hz, and J = 4.3, 10.8 Hz, 1 H), 3.13-3.06 (m, 1 H), 2.44-2.36 and 2.32-2.25 (each as m, 1 H), 2.12-2.03 and 2.00-1.89 (each as m, 1 H), 1.11-0.94 (m, 28 H); ^{13}C NMR ($CDCl_3$) δ 171.24, 167.54, 167.31, 163.68, 160.37, 159.89, 157.08, 156.97, 155.85, 151.86, 129.84, 129.77, 129.68, 122.40, 122.26, 121.85, 121.48, 121.32, 114.88, 114.73, 85.02, 83.76, 78.08, 77.78, 72.41, 69.19, 67.68, 65.19, 62.77, 60.60, 42.20, 40.07, 39.32, 38.13, 17.54, 17.47, 17.40, 17.26, 17.12, 13.79, 13.63, 12.99, 12.76; IR (film) 3249, 2938, 2860, 1688, 1582, 1491 cm^{-1}; HRMS FAB (M+H) calcd. for $C_{31}H_{49}N_4O_8Si_2$: 645.3140, found 645.3142.

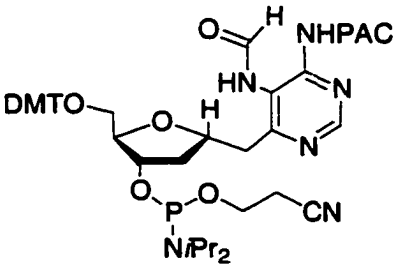


Preparation of 69: To a solution of **68** (0.22 g, 0.32 mmol) in THF (6.3 mL) was added $Et_3N•3HF$ (0.51 g, 3.2 mmol). The reaction was stirred at room temperature for 6 h and concentrated. The resulting oil was purified by column chromatography (elution with 100% EtOAc to 6% MeOH) to give the desired diol as a mixture of anomers (0.119 g, 92.4%). ^1H NMR (CD_3OD) δ 8.79 (s, 1 H), 8.32 and 8.29 (each as s, 1 H), 7.31-7.22 (m, 2 H), 7.01-6.88 (m, 3 H), 4.78-4.77 (m, 2 H), 4.56-4.50 (m, 1 H), 4.24-4.17 (m, 1 H), 3.86-3.75 (m, 1 H), 3.63-3.53 (m, 1 H), 3.48-3.44 and 3.22 (first as m, second as dd, J = 13.8, 7.8 Hz, 1 H), 3.11-3.05 (m, 1 H), 3.00-2.96 and 2.87-2.74 (each as m, 1 H), 2.40-2.32 and 2.11-2.03 (each as m, 1 H), 1.96-1.86 and 1.80-1.71 (each as m, 1 H); ^{13}C NMR (CD_3OD) δ 169.81, 169.73, 166.76, 166.41, 163.04, 158.86, 156.72, 156.63, 154.16, 130.78, 130.63, 123.09, 122.66, 121.99, 121.85, 121.85, 115.95, 115.69, 88.60, 87.42, 79.25, 78.96, 78.12, 73.37, 73.26, 72.24, 68.68, 67.91, 67.12, 66.96, 66.01, 63.37, 63.29, 52.72, 41.54, 41.18; IR (film) 3380, 3242, 2930, 1683, 1585, 1487 cm^{-1}; HRMS FAB (M+H) calcd. for $C_{19}H_{23}N_4O_6$: 403.1618, found 403.1599.



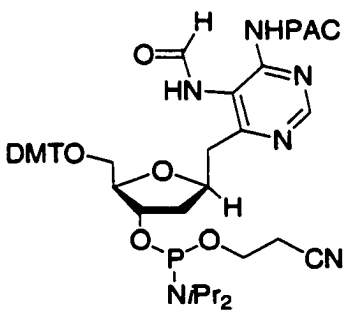
Preparation of 70: The above diol (0.12 g, 0.3 mmol) was coevaporated with pyridine (3 x 1 mL) and then dissolved in pyridine (1.5 mL). DMAP (4 mg, 0.03 mmol) and dimethoxytrityl chloride (0.12 g, 0.36 mmol) were added at 0°C in pyridine (1.5 mL). The reaction was stirred for 5 h, concentrated, and purified by column chromatography (elution with 50% EtOAc in hexanes to 100% EtOAc) to give **70** (130 mg, 61%) as a 7:3 (β:α) mixture of anomers. The anomers were separated using

a chromatotron (1mm plate, 2.5 % MeOH in DCM) to give 40 mg of α-**69** and 83 mg of β-**69**. α-**69**: ^{1}H NMR ($CDCl_3$) δ 9.77 (b, 1 H), 9.40 (b, 1 H), 8.94 (s, 1 H), 8.31 (s, 1 H), 7.39-7.17 (m, 11 H), 7.05-7.01 (m, 3 H), 6.82-6.80 (m, 4 H), 4.68 (s, 2 H), 4.54-4.50 (m, 1 H), 4.34-4.32 (m, 1 H), 4.02-3.98 (m, 1 H), 3.77 (s, 6 H), 3.75-3.73 (m, 1 H), 3.21-3.12 (m, 4 H), 2.50-2.43 (m, 1 H), 1.85-1.80 (m, 1 H); ^{13}C NMR ($CDCl_3$) δ 167.55, 163.99, 160.19, 158.74, 157.17, 155.97, 152.01, 144.76, 135.84, 130.20, 129.95, 128.26, 128.08, 127.15, 122.47, 121.55, 115.00, 113.39, 113.20, 86.69, 85.06, 79.08, 77.55, 73.90, 67.68, 64.29, 55.42, 40.94, 39.63; IR (film) 3260, 3057, 2931, 1600, 1583, 1509 cm^{-1}; HRMS FAB (M+H) calcd. for $C_{40}H_{41}N_4O_8$: 705.2924, found 705.2902. β-**69**: ^{1}H NMR ($CDCl_3$) δ 9.71 (b, 1 H), 9.38 (b, 1 H), 8.93 (s, 1 H), 7.90 (s, 1 H), 7.39-7.25 (m, 10 H), 7.18-7.14 (m, 1 H), 7.04-6.98 (m, 3 H), 6.82 (d, J = 8.8 Hz, 4 H), 4.67 (s, 2 H), 4.59-4.55 (m, 1 H), 4.21-4.18 (m, 1 H), 4.04-4.00 (m, 1 H), 3.75 (s, 3 H), 3.74 (s, 3 H), 3.28-3.19 (m, 2 H), 3.07-3.01 (m, 2 H), 2.09-2.05 (m, 1 H), 1.90-1.84 (m, 1 H); ^{13}C NMR ($CDCl_3$) δ 167.31, 163.67, 160.33, 158.87, 158.84, 157.16, 155.94, 152.12, 144.63, 135.55, 130.16, 129.88, 128.10, 127.29, 122.38, 121.14, 114.99, 113.46, 86.77, 86.52, 79.17, 73.64, 67.66, 64.46, 55.38, 40.66, 40.27; IR (thin film) 3254, 3058, 2932, 2837, 1684, 1582, 1509 cm^{-1}; HRMS FAB (M+H) calcd. for $C_{40}H_{41}N_4O_8$: 705.2924, found 705.2908.



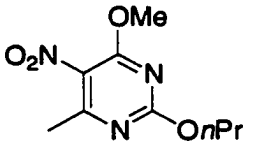
Preparation of α-20: A solution of α-**70** (40 mg, 0.06 mmol) was coevaporated with pyridine (2 x 1 mL). The oil was taken up in DCM (1 mL) and *N,N*-diisopropylaminotetrazolide (10 mg, 0.06 mmol) was added, followed by the addition of 2-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphane (23

mg, 0.074 mmol). The reaction was stirred at room temperature for 5 h, quenched with isopropyl alcohol (50 μL), and stirred for 20 min. The mixture was diluted with EtOAc (10 mL), and washed with brine (5 mL). The organic phase was dried over Na_2SO_4, concentrated, and chromatographed on oven-dried EM-silica gel (20% EtOAc, 79% Et_2O, 1% Et_3N to 50% EtOAc) to give α-**20** (23.3 mg, 45%). ^1H NMR ($CDCl_3$) δ 9.84 (b, 1 H), 9.44 (b, 1 H), 8.96 (s, 1 H), 8.35 and 8.33 (each as s, 1 H), 7.39-7.14 (m, 11 H), 7.04-6.99 (m, 3 H), 6.83-6.77 (m, 4 H), 4.75- 4.64 (m, 2 H), 4.64-4.58 (m, 1 H), 4.47-4.42 (m, 1 H), 4.18-4.12 (m, 1 H), 3.77 (s, 6 H), 3.70-3.45 (m, 4 H), 3.31-3.07 (m, 4 H), 2.56-2.43 (m, 1.0 H), 1.87 (m, 1.0 H), 1.30-1.05 (m, 12 H); ^{31}P NMR ($CDCl_3$) δ 148.59, 149.17. HRMS FAB (M + H) calcd. for $C_{49}H_{58}N_6O_9P$: 905.4003, found 905.3998.

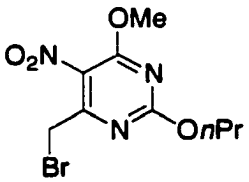


Preparation of β-20: A solution of β-**70** (46 mg, 0.065 mmol) was reacted as described above for the α-isomer with *N,N*-diisopropylaminotetrazolide (11 mg, 0.065 mmol) and 2-cyanoethyl-N,N,N',N'-tetraisopropylphosphane (26 mg, 0.085 mmol). The reaction was stirred at room temperature for 20 h and then worked up and purified as described above to give β-**20** (33.2 mg, 56%). ^1H NMR ($CDCl_3$) δ 9.72 (b, 1 H), 9.43 and 9.41 (each as s, 1 H), 8.95 (s, 1 H), 7.85 (s, 1 H), 7.40-7.14 (m, 11 H), 7.04-6.96 (m, 3 H), 6.80 (m, 4 H), 4.67 (s, 2 H), 4.57-4.48 (m, 1 H), 4.30-4.13 (m, 3 H), 3.74 (s, 6 H), 3.67-3.46 (m, 3 H), 3.32-3.20 (m, 2 H), 3.10-3.03 (m, 2 H), 2.76 (t, J =6.3 Hz, 1 H), 2.60 (t, J = 6.3 Hz, 1 H), 2.43 (t, J =6.3 Hz, 1 H), 2.30-2.15 (m, 1 H), 1.96-1.80 (m, 1 H), 1.29-1.08 (m, 12 H); ^{31}P NMR ($CDCl_3$) δ

148.74, 149.04; HRMS FAB (M + H) calcd. for $C_{49}H_{58}N_6O_9P$: 905.4003, found 905.4008.

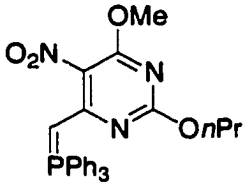


Preparation of 77:[200] To a solution of **59** (3.96 g, 19.9 mmol) in propanol (66 mL) was added Ag_2O (8.76 g, 37.8 mmol). The mixture was brought to reflux, stirred for 5 h, cooled to room temperature, and filtered through a pad of celite. The solution was concentrated in vacuo and the resulting oil was dissolved in EtOAc (100 mL). The organics were washed with saturated $NaHCO_3$ (30 mL), dried with $MgSO_4$, and concentrated in vacuo. The crude material was purified by column chromatography (elution with 100% hexane to 5% EtOAc) to produce a white solid (2.59 g, 57%). ^1H NMR ($CDCl_3$) δ 4.37 (t, *J*= 6.6 Hz, 2 H), 4.09 (s, 3 H), 2.51 (s, 3 H), 1.89-1.80 (m, 2 H), 1.05 (t, *J*= 7.5 Hz, 3 H).



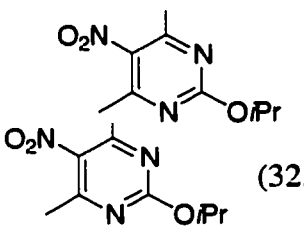
Preparation of 79: To a solution of **77** (3.48 g, 15.3 mmol) and NaOAc (2.50 g, 30.6 mmol) in AcOH (75 mL) was added a 9:1(v/v) solution of AcOH and Br_2 (2.7 mL, 5.24 mmol). The solution was heated to reflux until the red color dissipated. The reaction was cooled to room temperature, and additional 9:1(v/v) Br_2 solution in AcOH was added (2.6 mL, 5.04 mmol), and the reaction was returned to reflux. After 20 min another portion of Br_2 solution was added (2.6 mL, 5.04 mmol) in a similar manner and the reaction was refluxed for 30 min and concentrated in vacuo. The resulting oil was purified by column chromatography (100% hexanes to 2% EtOAc) to give **79** that was contaminated with residual starting material (2.13 g, 46%) and **80** (3.08 g, 52%). **79**: ^1H NMR ($CDCl_3$) δ

4.48 (s, 1 H), 4.40 (t, 2 H, *J*= 6.6 Hz), 4.12 (s, 3 H), 1.93-1.81 (m, 2 H), 1.06 (t, 3 H, *J*= 7.2 Hz); ^{13}C NMR ($CDCl_3$) δ 164.08, 163.53, 161.29, 128.55, 71.06, 55.92, 27.00, 22.11, 10.50; IR (film) 2971, 1586, 1526, 1493, 1431, 1333 cm^{-1}; HRMS (M+H) for $C_9H_{13}N_3O_4Br$: calcd. 306.0089, obsvd. 306.0083. **80**: ^{1}H NMR ($CDCl_3$) δ 6.83 (s, 1 H), 4.49 (t, *J*= 6.9 Hz, 2 H), 4.14 (s, 3 H), 1.94-1.87 (m, 2 H), 1.08 (t, *J*= 7.8 Hz, 3 H); ^{13}C NMR ($CDCl_3$) δ 163.53, 159.40, 123.97, 71.02, 56.14, 33.61, 21.88, 10.33; mp 59-60 °C; IR (film) 2967, 1568, 1529, 1490, 1433, 1330, 1066 cm^{-1}; HRMS (M+H) for $C_9H_{12}N_3O_4Br_2$: calcd. 385.9174, found 385.9169.



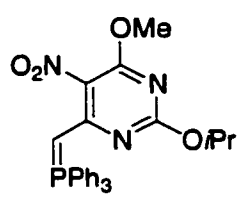
Preparation of 74: The mixture of **79** and **80** (2.74 g, 8.94 mmol) was dissolved in benzene (15 mL) and PPh_3 (3.30 g, 12.52 mmol) in benzene (20 mL) was added dropwise. The reaction was stirred for 5 h and filtered. The resulting white salt was washed with benzene (2 x 20 mL). The phosphonium bromide was dissolved in a biphasic mixture of $CHCl_3$ (55 mL) and ice water (35 mL). To this mixture was added 3 M NaOH (3.0 mL, 9.03 mmol) and the reaction was stirred vigorously. After 1.5 h the aqueous layer was extracted with $CHCl_3$ 2 x 20 mL), and the combined organic layers were washed with brine. The solution was concentrated to leave an orange powder **74** (3.17 g, 73 % or 27 % from **77**). ^{1}H NMR ($CDCl_3$) δ 7.65-7.45 (m, 15 H), 5.23 (d, *J*= 26.7 Hz, 1 H), 3.49 (s, 3 H), 2.95 (t, *J*= 6.9 Hz, 2 H), 1.23 (m, 2 H), 0.54 (t, *J*= 7.8 Hz, 3 H); ^{13}C NMR ($CDCl_3$) δ 166.60, 161.41, 161.37, 159.79, 132.92, 132.79, 132.36, 128.99, 128.83, 126.74, 125.53, 116.87, 68.36, 57.91, 56.38, 54.56, 21.77, 10.29; ^{31}P NMR ($CDCl_3$) δ 16.82; mp 245-246 °C; IR (film)

1564, 1524, 1436, 1409, 1246, 1100 cm^{-1}; HRMS (M+H) for $C_{27}H_{27}N_3O_4P$: calcd. 488.1739, found 488.1746.



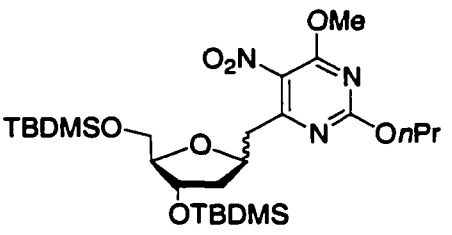
(32.60 g, 139.3 mmol) in 2-propanol (240 mL) was heated to reflux.

(32.60 g, 139.3 mmol) in 2-propanol (240 mL) was heated to reflux. After 24 h the solution was filtered through a pad of celite and concentrated. The resulting oil was dissolved in EtOAc (100 mL) and washed with saturated $NaHCO_3$ (50 mL). The aqueous layer was extracted with EtOAc (3 x 50 mL) and the combined organics were concentrated in vacuo. The crude material was purified by column chromatography (100% hexanes to 2% EtOAc) to give the product as a white solid (7.06 g, 42 %) and **59** (2.63 g, 18 %). ^1H NMR ($CDCl_3$) δ 5.27 (h, 1 H, *J*= 6.0 Hz), 4.01 (s, 3 H), 2.42 (s, 3 H), 1.34 (d, *J*= 6.4 Hz, 6 H); ^{13}C NMR ($CDCl_3$,) δ 163.79, 163.42, 162.48, 129.57, 72.11, 55.30, 21.76, 21.19; mp 82-84 °C; IR (film) 2982, 1581, 1519, 1428, 1378, 1316, 1113, 794 cm^{-1}; HRMS (M+H) for $C_9H_{14}N_3O_3$: calcd. 228.0984, found 228.0985.

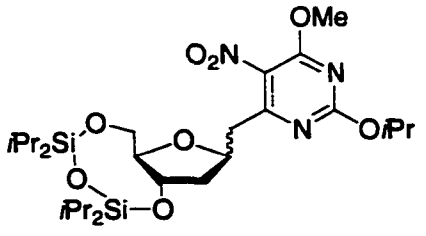


Preparation of 75: To a solution of **78** (5.13 g, 22.6 mmol) and NaOAc (3.70 g, 45.2 mmol) in AcOH (100 mL) was added a 9:1(v/v) solution of AcOH and Br_2 (4.0 mL, 7.76 mmol). The solution was heated to reflux until the red color dissipated. The reaction was cooled to room temperature, and additional 9:1 (v/v) of Br_2 in AcOH was added (4.0 mL, 7.76 mmol), and the reaction was returned to reflux. After 30 min another portion of Br_2 solution was added (5.0 mL, 9.70 mmol) added in a similar manner, and the reaction was refluxed for

30 min and concentrated in vacuo. The resulting oil was purified by column chromatography (100% hexanes to 2% EtOAc) to give the monobromide product (**81**) that was contaminated with residual starting material (3.32 g, 48%) and **82** (3.91 g, 45%). **81:** ^{1}H NMR ($CDCl_3$) δ 6.79 (s, 1 H), 5.40 (h, 1 H, *J*= 6.4 Hz), 4.10 (s, 3 H), 1.45 (d, 6 H, *J*= 6.4 Hz); ^{13}C NMR ($CDCl_3$) δ 164.04, 163.50, 159.91, 124.18, 73.66, 56.31, 33.67, 21.78; mp 83-84 °C; IR (film) 2982, 1577, 1523, 1485, 1427, 1386, 1324, 1105 cm^{-1}. The mixture of **81** and **78** (3.32 g, 10.85 mmol) was dissolved in benzene (20 mL) and PPh_3 (3.98 g, 15.2 mmol) in benzene (16 mL) was added dropwise. The reaction was stirred for 4.5 h and filtered. The resulting white salt was washed with benzene (2 x 30 mL). The phosphonium bromide was dissolved in a biphasic mixture of $CHCl_3$ (70 mL) and ice water (40 mL). To this mixture is added 3 M NaOH (3.65 mL, 11.0 mmol) and the reaction was stirred vigorously. After 1.5 h the aqueous layer was extracted with $CHCl_3$ (2 x 30 mL), and the combined organic layers were washed with brine. The solution was concentrated to leave an orange foam, **75** (4.19 g, 79% or 38% from **78**). ^{1}H NMR ($CDCl_3$) δ 7.68-7.45 (m, 15 H), 5.21 (d, 1 H, *J*= 26.4 Hz), 3.99 (s, 3 H), 3.81(h, 1 H, *J*= 6.0 Hz), 0.66 (d, 6 H, *J*= 6.3 Hz); ^{13}C NMR ($CDCl_3$) δ 166.64, 161.64, 161.58, 159.40, 132.92, 132.79, 132.38, 132.01, 131.88, 129.05, 128.90, 128.49, 128.32, 126.76, 125.55, 116.85, 68.80, 57.16, 55.64, 54.61, 21.62; ^{31}P NMR ($CDCl_3$) δ 16.66; mp 221-223 °C; IR (film) 3059, 2976, 1560, 1519, 1435, 1394, 1310, 1256, 1114, 715 cm^{-1}; HRMS (M+H) for $C_{27}H_{27}N_3O_4P$: calcd. 488.1739, found 488.1736.

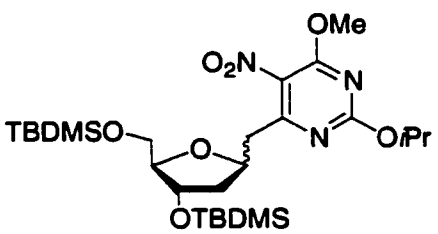


Preparation of 83: A solution of **26** (2.54 g, 6.99 mmol) and **74** (3.41 g, 6.99 mmol) in toluene (23 mL) was heated to reflux for 4.5 days. The solution was concentrated in vacuo and suspended in MeOH (35 mL). The solution was stirred for 45 min with NaOMe (0.08 g, 1.4 mmol). The reaction was filtered and quenched by the addition of excess NH_4Cl. The solution was concentrated in vacuo and chromatographed (elution with 2% EtOAc in hexanes to 10% EtOAc) to give a mixture of anomers as an oil (2.96 g, 74%). ^1H NMR ($CDCl_3$) δ 4.61-4.54 (m, 1 H), 4.35-4.28 (m, 3 H), 4.03 and 4.02 (each as s, 3 H), 3.86-3.83 and 3.77-3.74 (each as m, 1 H), 3.59-3.54 (m, 1 H), 3.47-3.42 (m, 1 H), 3.41-3.39 and 3.35-3.30 (each as m, 1 H), 3.03 (dd, *J*= 10.5, 5.7 Hz, 1 H), 2.88 and 2.71 (each as dd, *J*= 13.8, 4.8 Hz and *J*= 13.8, 4.4 Hz, 1 H), 2.27-2.20 and 1.91-1.87 (each as m, 1 H), 1.83-1.77 (m, 2 H), 1.75-1.66 (m, 1 H), 1.01 (t, *J*= 7.6 Hz, 3 H), 0.88, 0.86, and 0.84 (each as s, 9 H), 0.06, 0.05, 0.03, 0.02, 0.00, and -0.01 (each as s, 6 H); ^{13}C NMR ($CDCl_3$) δ 164.10, 163.59, 163.45, 163.39, 163.14, 163.08, 130.98, 130.70, 88.07, 87.14, 76.90, 76.79, 74.32, 73.86, 70.51, 70.40, 65.96, 63.83, 63.68, 55.36, 55.30, 41.02, 40.40, 39.62, 26.07, 25.91, 22.17, 18.46, 18.09, 10.55, -4.60, -4.67, -4.71, -5.09, -5.29, -5.36; IR (film) 2954, 2930, 1585, 1531, 1430, 1346, 1331, 1097, 837 cm^{-1}; HRMS (M+H) for $C_{26}H_{50}N_3O_7Si_2$: calcd. 572.3187, found 572.3193.



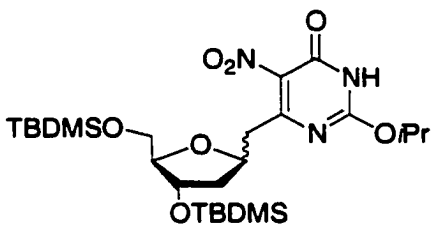
Preparation of 84: A solution of **62** (2.35 g, 6.23 mmol) and **75** (3.04 g, 6.23 mmol) in toluene (21 mL) was heated to reflux for 2 days. The solution was

concentrated in vacuo and suspended in MeOH (30 mL). The solution was stirred for 1.5 h with NaOMe (0.07 g, 1.2 mmol). The reaction was filtered and quenched by the addition of excess NH_4Cl. The solution was concentrated in vacuo and chromatographed (elution with 2% EtOAc in hexanes to 10% EtOAc) to give a mixture of anomers as an oil (2.74 g, 75%). ^1H NMR ($CDCl_3$) δ 5.31-5.26 (m, 1 H), 4.51-4.44 (m, 1 H), 4.37-4.33 (m, 1 H), 4.02 and 4.01 (each as s, 3 H), 3.96-3.87 (m, 1 H), 3.71-3.60 (m, 1 H), 3.70-3.60 and 3.55-3.42 (each as m, 1 H), 3.29 and 3.02 (each as dd, *J*= 13.6, 8.8 Hz, and *J*=14.0, 7.2 Hz, 1 H), 2.87-2.79 (m, 1H), 2.40-2.33 and 2.10-2.04 (each as m, 1 H), 1.93-1.79 (m, 1 H), 1.38-1.36 (m, 6 H), 1.02-0.88 (m, 28 H), ^{13}C NMR ($CDCl_3$) δ 163.73, 163.50, 163.02, 162.57, 130.66, 130.43, 86.26, 83.28, 75.97, 75.72, 74.04, 72.64, 72.26, 72.11, 71.84, 69.64, 69.09, 66.36, 64.15, 62.63, 55.36, 55.30, 40.43, 39.95, 39.44, 39.15, 38.57, 34.76, 31.70, 22.76, 21.83, 21.76, 20.78, 17.70, 17.64, 17.59, 17.48, 17.39, 17.28, 17.22,17.11, 17.06, 14.35, 14.20, 13.86, 13.58, 13.49, 13.36, 13.13, 13.07, 12.98, 12.93, 12.67; IR (film) 2941, 2862, 1579, 1525, 1426, 1385, 1323, 1113, 1034 cm^{-1}.



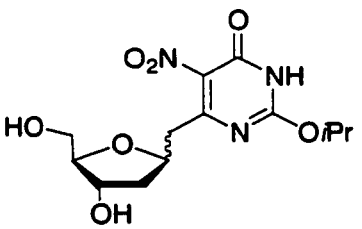
Preparation of 85: A solution of **26** (4.48 g, 12.35 mmol) and **75** (6.02 g, 12.35 mmol) in toluene (41 mL) was heated to reflux for 4.5 days. The solution was concentrated in vacuo and suspended in MeOH (62 mL). The solution was stirred for 1 h with NaOMe (0.2 g, 2.17 mmol). The reaction was filtered and quenched by the addition of excess NH_4Cl. The solution was concentrated in vacuo and chromatographed (elution with 2% EtOAc in hexanes to 10% EtOAc) to give a mixture of anomers as an oil (6.11 g, 87%). ^1H NMR ($CDCl_3$) δ 5.36-5.28 (m, 1 H),

4.64-4.57 (m, 1 H), 4.36-4.31 (m, 1 H), 4.06 and 4.05 (each as s, 3 H), 3.89-3.86 and 3.81-3.77 (each as m, 1H), 3.63-3.57 (m, 1 H), 3.52-3.42 (m, 1 H), 3.38-3.33 (m, 1 H), 3.11-3.04 (m, 1 H), 2.91-2.76 (m, 1 H), 2.29-2.22 and 1.95-1.88 (each as m, 1 H), 1.78-1.67 (m, 1 H), 1.42 (d, *J*= 6.3 Hz, 6 H), 0.91, 0.90, and 0.87 (each as s, 9H), 0.09, 0.08, 0.07, 0.06, 0.03, and 0.02 (each as s, 6 H); ^{13}C NMR ($CDCl_3$) δ 164.22, 164.00, 163.58, 163.38, 163.16, 162.78, 162.67, 130.56, 130.47, 88.20, 77.93, 74.39, 73.94, 72.48, 72.30, 72.15, 72.10, 72.01, 69.84, 64.70, 63.92, 63.76, 62.66, 55.45, 55.38, 44.10, 40.52, 39.70, 39.24, 26.17, 25.98, 25.87, 21.96, 21.89, 18.55, 18.17, 18.12, 14.36, -3.36, -4.52, -4.58, -4.61, -5.22, -5.27, -5.32, -5.44; IR (film) 2952, 2931, 1584, 1529, 1428, 1381, 1255, 1117, 838 cm^{-1}; HRMS (M+H) for $C_{26}H_{50}N_3O_7Si_2$: calcd. 572.3187, found 572.3183.

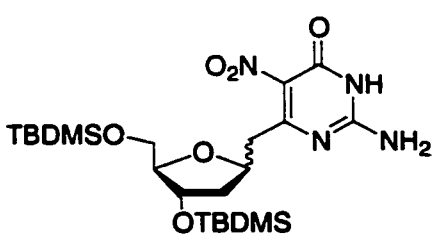


Preparation of 89: To a solution of **85** (5.27 g, 9.22) in AcN (46 mL), was added NaI (2.07 g, 13.83 mmol) and TBDMSCl (2.08 g, 13.83 mmol). The reaction was stirred at ambient temperature for 2 days and diluted with EtOAc (70 mL). The product was washed with saturated $NaHSO_3$ (30 mL), and the aqueous phase was extracted with EtOAc (2 x 30 mL). The combined organic phase was washed with brine and concentrated. The crude product was purified by column chromatography (elution with 100% DCM to 2% MeOH) to give the product as a white foam (3.43 g, 67%) and recovered starting material (1.42 g, 27 %). ^{1}H NMR ($CDCl_3$) δ 5.41-5.33 (m, 1 H), 4.63-4.56 (m, 1 H), 4.37-4.31 (m, 1 H), 3.92-3.88 and 3.82-3.76 (each as m, 1 H), 3.63-3.57 (m, 1 H), 3.50-3.32 (m, 1 H), 3.04-2.97 (m, 1 H), 2.79-2.67 (m, 1 H), 2.29-2.24 and 1.94-1.88 (each as m, 1 H), 1.77-1.68 (m, 1 H), 1.39

(m, 6 H), 0.91, 0.90, 0.89, and 0.87 (each as s, 9 H), 0.10, 0.08, 0.07, 0.06, 0.04, and 0.03 (each as s, 6 H); ^{13}C NMR ($CDCl_3$) δ 163.98, 163.47, 163.38, 162.48, 130.68, 130,41, 87.96, 87.03, 77.54,76.90, 76.72, 74.23, 73.77, 72.06, 71.86, 63.72, 63.59, 55.21, 55.16, 40.93, 40.33, 39.49, 25.98, 25.82, 21.79, 18.35, 17.98, -4.69, -4.76, -4.80, -5.38, -5.46; IR (film) 2951, 2926, 2851, 1579, 1425, 1324, 1111, 839 cm^{-1}; HRMS (M+H) for $C_{25}H_{48}N_3O_7Si_2$: calcd. 558.3031, found 558.3015.

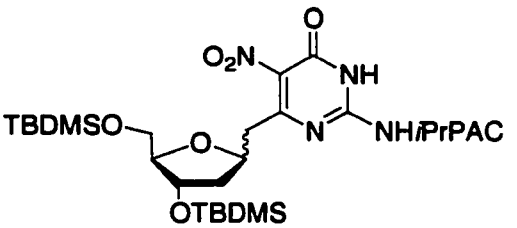


Preparation of 91: A solution of **89** (3.43 g, 6.16 mmol) and NH_4F (2.30 g, 61.6 mmol) in MeOH (41 mL) was heated at reflux for 18 h. The solution was concentrated and the product was purified by column chromatography (elution with 2% MeOH in DCM to 8% MeOH) to give the desilylated product as a white foam (1.62 g, 80%). ^1H NMR (CD_3OD) δ 5.27-5.19 (m, 1 H), 4.56-4.44 (m, 1 H), 4.24-4.18 (m, 1 H), 3.81-3.72 (m, 1 H), 3.61-3.46 (m, 2 H), 3.14-3.08 and 2.91-2.86 (each as m, 1 H), 2.86-2.71 (m, 1 H), 2.38-2.29 and 2.00-1.93 (each as m, 1 H), 1.87-1.71 (m, 1 H), 1.33 (d, *J*= 6.0 Hz, 6 H); ^{13}C NMR (CD_3OD) δ 167.37, 166.97, 163.45, 161.78, 161.70, 135.00, 134.60, 88.67, 86.91, 78.35, 78.14, 73.90, 73.91, 71.73, 71.59, 63.92, 63.28, 41.70, 41.21, 40.75, 40.41, 22.38; IR (film) 3361, 3174, 3071, 2922, 2849, 1687, 1568, 1308, 1100, 836 cm^{-1}; HRMS (M+H) for $C_{13}H_{20}N_3O_7$: calcd. 330.1301, found 330.1300.

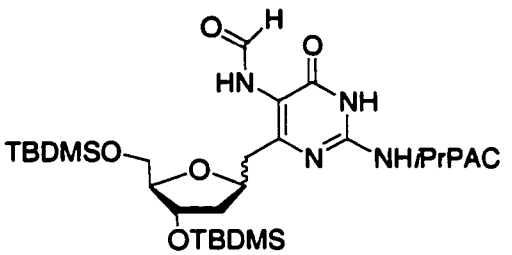


Preparation of 92: The diol **91**(1.50 g, 4.56 mmol) in 18mL of MeOH was saturated with NH_3 for 20 min at

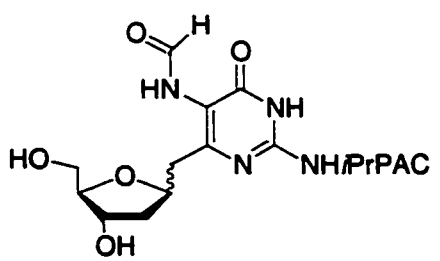
0 °C. The solution was then sealed and heated to 90 °C for 43 h. Concentration of the solution afforded the crude aminated diol (1.42 g) as a mixture with the unreacted diol. The crude material was azeotropically dried with pyridine (3 x 5 mL) and silylated with TBDMSCl (1.72 g, 11.4 mmol) and imidazole (1.37 g, 4.4 mmol) in DMF (23 mL) for 24 h. The reaction was concentrated in vacuo and the crude oil was partitioned between EtOAc (40 mL) and water (20 mL). The aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organics were washed successively with water (15 mL), saturated $NaHCO_3$ (15 mL), and brine (15 mL). The organics were concentrated and purified by column chromatography (elution with 100% DCM to 2% MeOH) to give the product (**92**) as a foam (1.42 g, 61%) and **89** (0.72 g, 28%). ^{1}H NMR ($CDCl_3$) δ 4.80-4.72 and 4.78-4.62 (each as m, 1 H), 4.41-4.37 and 4.31-4.30 (each as m, 1 H), 3.96-3.94 and 3.85-3.82 (each as m, 1 H), 3.70-3.64 (m, 1 H), 3.48-3.44 (m, 1 H), 3.17-3.07 (m, 1 H), 2.94-2.90 (m, 1 H), 2.39-2.34 and 2.01-1.96 (each as m, 1 H), 1.82-1.72 (m, 1 H), 0.91, 0.90, 0.88, and 0.86 (each as s, 9H), 0.09, 0.08, 0.05, 0.04, 0.02, and 0.00 (each as s, 6 H); ^{13}C NMR ($CDCl_3$) δ 170.66, 169.57, 157.84, 154.80, 154.65, 128.48, 87.92, 86.34, 77.38, 76.27, 73.93, 72.68, 63.90, 62.57, 41.50, 40.91, 26.20, 26.13, 26.09, 26.05, 26.00, 18.49, 18.17, 18.09, -4.14, -4.22, -4.42, 4.51, -4.56, -4.67, -5.27, -5.31; IR (thin film) 3327, 3214, 3152, 2929, 2855, 1683, 1591, 1486, 1254, 1106, 834 cm^{-1}; HRMS (M+H) for $C_{22}H_{43}N_4O_6Si_2$: calcd. 515.2721, found 515.2715.



Preparation of 93: To a solution of **92** (0.48 g, 0.93 mmol), PyBOP (0.730 g, 1.40 mmol), and isopropylphenoxyacetic acid (0.27 g, 1.40 mmol) in DCM (9.3 mL) was added iPr$_2$NEt (0.35 g, 2.78 mmol). The reaction was stirred at ambient temperature for 6 h and partitioned between DCM (50 mL) and saturated $NaHCO_3$ (20 mL). The aqueous layer was extracted with DCM (2 x 25 mL) and washed with brine (25 mL). The solution was concentrated and purified by column chromatography (elution with 100% DCM to 2% MeOH) to give a white foam (0.46 g, 72 %). ^{1}H NMR ($CDCl_3$) δ 9.35 (br s, 1 H), 7.22 (d, 2 H, J = 8.0 Hz), 6.92 (d, J = 8.0 Hz, 2 H), 4.70 (s, 2 H), 4.58-4.52 (m, 1H), 4.38-4.32 (m, 1 H), 3.92-3.89 and 3.82-3.77 (each as m, 1 H), 3.61-3.56 (m, 2 H), 3.47-3.44 and 3.40-3.35 (each as m, 1 H), 2.93-2.85 (m, 1 H), 2.81-2.77 and 2.68-2.65 (each as m, 1 H), 2.29-2.24 and 1.95-1.90 (each as m, 1 H), 1.77-1.70 (m, 1 H), 2.21 (d, J = 2.4 Hz, 6 H), 1.25 (s, 9 H), 1.23 (s, 9 H), 0.09, 0.08, 0.07, and 0.06 (each as s, 6 H), 0.05, 0.04, and 0.03 (each as s, 6 H); ^{13}C NMR ($CDCl_3$) δ 170.93, 161.39, 161.04, 154.52, 153.38, 148.59, 148.44, 144.05, 136.29, 128.06, 114.93, 88.30, 87.41, 76.10, 74.32, 73.86, 67.32, 63.88, 63.99, 41.15, 40.66, 40.55, 40.00, 33.52, 26.11, 25.98, 24.27, 18.51, 18.19, 18.11, -4.27, -4.51, -4.55, -4.64, -5.22, -5.31; IR (film) 3182, 2953, 1852, 1693, 1563, 1513, 1250, 1092, 837 cm^{-1}; HRMS (M+H) for $C_{33}H_{55}N_4O_8Si_2$: calcd. 691.3559, found 691.3566.

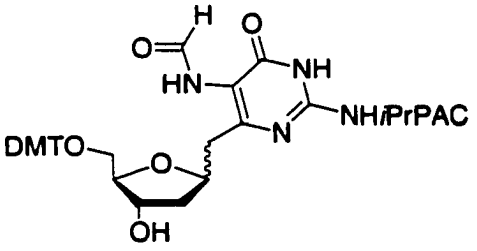


Preparation of 94: To a solution of **93** (0.46 g, 1.07 mmol) in 3:1 MeOH:THF (13.4 mL) was added 10% Pd/C (230 mg). The suspension was charged with H_2 (45 psi) and allowed to stir for 1 h. The reaction was filtered through a pad of celite and concentrated in vacuo. The crude amine was dissolved in 13.4 mL of THF and treated with AFA (0.18 g, 2.01 mmol) and pyridine (0.21 g, 2.70 mmol) at 0 °C. After 1.25 h the solvent was removed and the material purified by column chromatography (elution with 0.1 % MeOH in DCM to 5% MeOH) to afford a white foam (0.37 g, 80%). ^{1}H NMR (CD_3OD) δ 8.24 (s, 1 H), 7.16 (d, *J* = 8.8 Hz, 2 H), 6.93 (d, *J* = 8.8 Hz, 2 H), 4.74 (s, 2 H), 4.53-4.49 (m, 1 H), 4.45-4.40 and 4.38-4.35 (each as m, 1 H), 3.96-3.92 and 3.84-3.80 (each as m, 1 H), 3.68-3.64 and 3.54-3.50 (each as m, 1 H), 4.62-3.59 (m, 1 H), 3.02-2.99 and 2.90-2.84 (each as m, 1 H), 2.88-2.82 (m, 1 H), 2.72-2.68 (m, 1 H), 2.38-2.34 and 1.94-1.90 (each as m, 1 H), 1.83-1.71 (m, 1 H), 1.21 (d, *J*= 6.8 Hz, 6 H), 0.90 (s, 9 H), 0.89 (s, 9 H), 0.09 and 0.06 (each as s, 6 H), 0.05 and 0.03 (each as s, 6 H); ^{13} C NMR (CD_3OD) δ 173.24, 163.18, 163.09, 157.21, 143.86, 128.59, 115.94, 88.33, 78.47, 77.85, 75.49, 74.83, 68.46, 64.90, 64.39, 42.28, 41.71, 34.71, 26.73, 26.62, 26.59, 26,51, 24.78, 19.32, 18.99, -4.32, -4.41, -4.96, -5.03, -5.11; IR (film) 3189, 2953, 2853, 1673, 1594, 1507, 1249, 1096, 838 cm^{-1}; HRMS (M+H) for $C_{34}H_{57}N_4O_7Si_2$: calcd. 689.3766, found 689.3778.



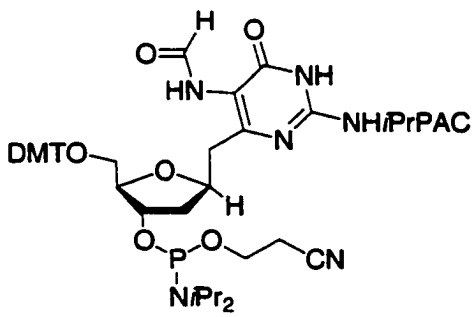
Preparation of 95: To a solution of the above formylated material, **94**, (0.37 g, 0.54 mmol) in THF (10.6 mL) was added Et_3N•3HF (0.85 g, 5.40 mmol).

The reaction was stirred for 24 h and concentrated to dryness. The crude material was purified by column chromatography (elution with 1% MeOH in EtOAc to 8% MeOH) to produce a white foam (0.21 g, 84%). ^{1}H NMR (CD_3OD) δ 8.27 and 8.25 (each as s, 1 H), 7.15-7.12 (m, 2 H), 6.93-6.69 (m, 2 H), 4.74 (s, 2 H), 4.51-4.46 (m, 1 H), 4.24-4.21 (m, 1 H), 3.93-3.90 and 3.82-3.78 (each as m, 1 H), 3.63-3.54 (m, 2 H), 2.84-2.76 (m, 2 H), 2.39-2.35 and 1.87-1.80 (each as m, 1 H), 2.04-1.98 (m, 1 H), 1.76-1.70 and 1.62-1.55 (each as m, 1 H), 1.19 (d, J = 6.8 Hz, 6 H); ^{13}C NMR (CD_3OD) d 173.32, 173.27, 167.97, 163.29, 163.20, 157.06, 150.03, 143.78, 128.54, 128.38, 117.45, 115.88, 115.57, 89.05, 87.40, 77.80, 77.70, 73.61, 73.34, 68.36, 66.22, 65.49, 63.74, 63.59, 63.41, 59.90, 52.70, 41.87, 41.76, 41.26, 34.59, 31.83, 24.73, 20.21; IR (film) 3476, 3068, 1694, 1477, 1277, 1167, 1043 cm^{-1}; HRMS (M+H) for $C_{22}H_{29}N_4O_7$: calcd. 461.2036, found 461.2033.



Preparation of 96: The above diol (0.10 g, 0.23 mmol) was coevaporated with pyridine (2 x 1 mL). The material was dissolved in pyridine (1.5 mL) and cooled to 0 °C. DMAP (3 mg, 0.03 mmol) and dimethoxytrityl chloride (0.11 g, 0.34 mmol) were added at 0°C in pyridine (1 mL). The reaction was stirred for 2 h at 0 °C and allowed to warm to room temperature. The reaction was stirred an additional 1 h, concentrated, and purified by column chromatography (elution with 70% EtOAc in hexanes to 2% MeOH in EtOAc) to give **96** (0.13 g, 73%) as a 3:1 (β:α) mixture of anomers. Resolution of anomers by TLC was accomplished with 10: 0.25: 4 EtOAc: *n*-propanol : water (organic phase). The separation of anomers was facilitated by chromatotron (1 mm plate, 100% EtOAc) to elute 93 mg of

β-**96** first, followed by 32 mg of α-**96**. α-**96**: ^{1}H NMR (CD_3OD) δ 8.32 and 8.26 (each as s, 1 H), 7.44-7.40 (m, 2 H), 7.31-7.09 (m, 9 H), 6.84-6.70 (m, 6 H), 4.68-4.63 (m, 2 H), 4.60-4.52 (m, 1 H), 4.35-4.31 (m, 1 H), 4.11-4.04 (m, 1 H), 3.77-3.72 (m, 6 H), 3.10-3.07 (m, 1 H), 2.86-2.79 (m, 2 H), 2.45-2.38 (m, 1 H), 1.76-1.70 (m, 1 H), 1.65-1.57 (m, 1 H), 1.19 (d, J =7.2 Hz, 6 H); ^{13}C NMR (CD_3OD) δ 173.25, 163.21, 160.21, 157.15, 146.72, 143.93, 137.52, 131.43, 129.52, 128.87, 128.60, 127.92, 115.97, 114.19, 87.52, 86.85, 78.24, 74.25, 74.12, 68.49, 65.61, 65.58, 55.83, 41.68, 34.74, 31.94, 24.75, 20.30; IR (film) 3175, 2957, 1660, 1603, 1512, 1250, 1172, 1028, 832 cm^{-1}; HRMS (M+H) for $C_{43}H_{47}N_4O_9$: calcd. 763.3343, found 763.3341. β-**96**: ^{1}H NMR (CD_3OD) d 8.25 and 8.14 (each as s, 1 H), 7.43-7.37 (m, 2 H), 7.30-7.13 (m, 9 H), 6.91-6.80 (m, 6 H), 4.69-4.64 (m, 2 H), 4.56-4.51 (m, 1 H), 4.26-4.23 (m, 1 H), 3.98-3.96 (m, 1 H), 3.72 (s, 6 H), 3.17-3.14 (m, 1 H), 3.09-2.97 (m, 1 H), 2.88-2.77 (m, 2 H), 1.97-1.93 (m, 1 H), 1.85-1.74 (m, 1 H), 1.66-1.56 (m, 1 H), 1.20 (d, J = 9.6 Hz, 6 H); ^{13}C NMR (1:1 $CDCl_3$:CD_3OD) δ 172.10, 166.48, 162.01, 159.32, 156.08, 145.70, 143.69, 136.79, 130.87, 129.36, 128.90, 128.47, 127.48, 115.44, 113.79, 87.45, 87.01, 77.18, 67.96, 65.07, 64.76, 59.66, 55.67, 41.46, 34.09, 24.56, 14.04, 8.48; IR (film) 3201, 2957, 1655, 1611, 1577, 1511, 1250, 1037, 832 cm^{-1}. HRMS (M+H) for $C_{43}H_{47}N_4O_9$: calcd. 763.3343, found 763.3349.



Preparation of β-21: A solution of β-**96** (47 mg, 0.06 mmol) was coevaporated with AcN (4 x 1 mL). The oil was taken up in DCM (1.5 mL) with diisopropylamine (6.5 mg, 0.06 mmol), and tetrazole (4.3 mg, 0.06 mmol) was added. This was

followed by addition of 2-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphane (24 mg, 0.08 mmol) to the reaction. The reaction was stirred at room temperature for 5 h, and diisopropylamine (1.4 mg, 0.014 mmol), tetrazole (0.9 mg, 0.013 mmol), and 2-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphane (3.7 mg, 0.012 mmol) were added. The reaction was allowed to stir for 12 h and quenched with MeOH (1 mL). The reaction was concentrated and chromatographed on oven-dried EM-silica gel (50% EtOAc, in hexanes with 0.1% Et_3N to 0.5% MeOH in EtOAc with 0.1% Et_3N) to give **β-97** (28.2 mg, 48%). ^1H NMR (CD_3OD) δ 8.27, 8.17, and 8.26 (each as s, 1 H), 7.44-7.38 (m, 2 H), 7.32-7.12 (m, 9H), 6.93-6.81 (m, 6 H), 4.71-4.66 (m, 2 H), 4.56-4.52 (m, 1 H), 4.46-4.42 (m, 1 H), 4.11-4.04 (m, 1 H), 3.75-3.59 (m, 9 H), 3.22-3.13 (m, 3 H), 2.90-2.80 (m, 3 H), 2.67 (t, *J* = 5.7 Hz, 1 H), 2.54 (t, *J* = 5.7 Hz, 1 H), 1.94-1.87 (m, 1 H), 1.64-1.58 (m, 1 H), 1.23-1.09 (m, 18 H); ^{31}P NMR ($CDCl_3$) δ 148.41, 148.34; HRMS (M+H) for $C_{52}H_{64}N_6O_{10}P$: calcd. 963.4422, found 963.4414.

Automated Synthesis of Oligonucleotides Containing α-18, β-18, and β-19: Standard 1 μmol scale cycles were modified for the synthesis of **98**, **99**, **102**, **103**, **112**, **113**, and **114**. Double-coupling of α-**20** or β-**20** with DCI as the activator was performed for 5 min each coupling, and capping solutions were substituted for 8:1:1 THF:lutidine:isobutyric anhydride during the coupling of α-**70** or β-**70**. Synthesis of oligonucleotides containing β-**19** was carried out differently. Modified capping conditions were used from the start of the synthesis (NMI, 8:1:1 THF:pivalic anhydride:lutidine—capping for 20 sec and were changed during the coupling of β-**21** by

substituting NMI for THF. Couplings were activated with tetrazole, and β-**21** was single-coupled for 15 min.

Oligonucleotides **98**, **99**, **102**, **103**, and **114** were deprotected by treatment with 0.05 M K_2CO_3 in MeOH (2 mL) for 3 h. Treatment of **112** and **113** with the same conditions required 5 h for complete deprotection. The deprotections were quenched with 2 equivalents of AcOH and concentrated to dryness. Oligonucleotides were purified by PAGE and desalted following standard protocols and characterized by ESI-MS.[220,221] **98a**: ESI-MS calcd. 3669.5, found 3669.6. **98b**: ESI-MS calcd. 3669.5, found 3670.0. **99a**: ESI-MS calcd. 3589.4, found 3590.0. **99b**: ESI-MS calcd. 3589.4, found 3589.8. **112:** ESI-MS calcd. 11103.3, found 11102.1. **113:** ESI-MS calcd. 11094.2, found 11093.4. **114:** ESI-MS calcd. 3687.5, found 3687.1.

Approximation of Extinction Coefficients (ε) for 18 and 19: The extinction coefficients (ε) for **18** and **19** were measured by analyzing the absorbance of **67** and **97** at known concentrations. Graphical analysis of the absorbance versus concentration provides ε as the slope. The ε for **67** is 2700 $L \cdot mol^{-1} \cdot cm^{-1}$, and 2430 $L \cdot mol^{-1} \cdot cm^{-1}$ for **97**. The measured values of ε for **67** and **97** were used as approximate values for nucleotides **18** and **19**. The DNA_{long} method for calculating ε was used for oligonucleotides containing **18** and **19**.[222]

Ligation Experiments (Analytical Scale): 5'-End-labeling of the 3' terminal dodecamer (**105**) was accomplished with [γ-^{32}P] ATP T4 and polynucleotide kinase in kinase buffer (70 mM Tris-HCl , pH 7.6, 10 mM $MgCl_2$, 5 mM DTT) according to the standard

protocol.[221] The kinase was heat deactivated at 95 °C for 5 min and purified by a G-25 sephadex spin column. Using the same procedure **98b** was phosphorylated with cold ATP. The above phosphorylated 12 mers and 5' terminal dodecamers (**104**, 2.5 pmol each) were annealed to template **106** (5 pmol) in kinase buffer such that the duplex concentration was 125 nM. The duplex was annealed at 90 °C for 5 min and cooled to 4 °C over 2 h. After annealing the mixture was incubated at 4 °C with 2 μL of 25 mM ATP and 1 μL of T4 DNA ligase (400 Units). Aliquots were removed at 3, 6, 9, 18, and 24 h and an additional 1 μL of T4 DNA Ligase (400 Units) was added at 3, 6, and 9 h. Aliquots were quenched with 95% formamide, denatured at 90 °C for 1 min, and immersed in a 0 °C bath. The aliquots were then electrophoresed on a 20% PAGE gel and visualized.

Preparation of 108 and 109 by Enzymatic Ligation. The 3'-terminal dodecamers, **105** (10 nmol) was treated with 670 nmol of ATP and T4 polynucleotide kinase (136 U) in 80 μL of kinase buffer (70 mM Tris-HCl , pH 7.6, 10 mM $MgCl_2$, 5 mM DTT). In a separate reaction the dodecamer (5 nmol) containing the lesion of interest (**98/99 a** or **b**) was phosphorylated with ATP (335 nmol) and T4 polynucleotide kinase (68 U) in kinase buffer (40 μL). The phosphorylation reactions were incubated at 37 °C for 4 h and then heat inactivated at 95 °C for 5 min. The phosphorylated dodecamers were combined and added to the 5'-terminal dodecamers, **104** (10 nmol) and the appropriate 42 nucleotide long template (**106** or **107**, 10 nmol) in kinase buffer (54 μL). The oligonucleotides were annealed at 90 °C for 5 min and cooled to 4 °C over 3 h. After annealing, ATP (315 nmol) and T4 DNA ligase (2000 U) was added to the duplex and incubated at 4 °C.

Additional T4 DNA ligase (2000 U) was added after 10 h and 20 h. After incubating for 34 h, the mixture was lyophilized, resuspended in 95% formamide loading buffer (100 μL), and denatured at 90 °C for 5 min. The denatured mixture was loaded without cooling unto a warm 15 % PAGE (50-60 °C) and electrophoresed at 40 mA. The gel was visualized by UV and the bands were crushed and eluted with elution buffer (200 mM NaCl, 1 mM EDTA) for 15 h. The oligonucleotide was desalted using a C18 Sep Pak cartridge and quantified. **108b**: 98%, ESI-MS calcd. 11085.2, found 11,086.0. **108a**: 59%, ESI-MS calcd. 11085.2, found 11085.0. **109b**: 56%, ESI-MS calcd. 11,005.2, found 11006.0. **109a**: 53%, ESI-MS calcd. 11,005.2, found. 11,007.0.

Molecular Modeling of α-Fapy•dA, α-18, α-dA, β-Fapy•dA, β-18, β-dA, β-Fapy•dG, β-19, and β-dG: Molecular modeling predictions were made using Molecular Simulations, Inc. (MSI) Cerius2 software installed on SGI workstations. The structures of β-dA, α-dA, and β-dG were minimized using two repetitions of 500 iterations of the "Smart Minimizer". Calculations assumed that molecules were in the gas phase and did not account for solvent interactions. The minimized structures of the native nucleosides were used as the point of divergence for the calculations on α-Fapy•dA, β-Fapy•dA, and β-Fapy•dG. The C8-N9 bond of the imidazole of the native nucleosides were removed and requisite N5-formamide was inserted to create the unminimized structures for α-Fapy•dA, β-Fapy•dA, and β-Fapy•dG. These structures were minimized as described above to give the total energy for the structures. Substitution of the glycosidic NH of the minimized structures of α-Fapy•dA, β-Fapy•dA, and β-Fapy•dG for CH_2 created

templates for α-**18**, β-**18**, and β-**19** that were further minimized to give the total energy for the C-nucleoside analogues.

Rotating the minimized structures of β-Fapy•dA/β-Fapy•dG, or β-**18**/β-**19** into the desired conformation followed by minimization gave the relative energy differences between conformations. Geometric measurements of the glycosidic torsional angle, nucleobase tilt, and the formamide torsional angles as described in Section 3.3 were calculated by measuring the respective torsional angles of the minimized structures.

DNA Melting Experiments: The samples for the oligonucleotide melting studies were composed of a total concentration of 1.0-14.0 μM of a 1:1 molar ratio of complementary oligonucleotides for duplexes containing α-**18** and β-**18**. Oligonucleotides containing β-**19** were composed of a total concentration range of 1.0-7.5 μM. The samples were prepared by the addition of appropriate amounts of complementary oligonucleotide stock solutions to 200 μL of PIPES buffer (20 mM PIPES, pH 7.0, 20 mM $MgCl_2$, 200 mM NaCl) followed by dilution to 400 μL with water. The complementary oligonucleotides were hybridized at 90 °C for 5 min and allowed to cool to room temperature overnight. The absorbance of the samples was then monitored at 260 nm while the temperature was increased at a rate of 1.0 °C /min over a range of 60 °C (25-85 °C) for oligonucleotides containing α-**18** or β-**18**. The temperature increase for duplexes containing β-**19** was 0.5 °C /min. Melting temperatures were calculated by computer fit of the first derivative of absorbance with respect to T^{-1}.[84,85] Thermodynamic parameters were obtained through van't Hoff analysis of the data.[85]

CD experiments: The samples for CD studies were composed of 10 μM of each complementary oligonucleotide, 100 mM NaCl, and 100 mM phosphate buffer (pH 7.5). Hybridization of a 2x solution (250 μL) from above at 90 °C for 5 min and cooling to room temperature overnight provided the duplex. Prior to running experiment the duplex solution was diluted to 500 μL with H_2O. Duplexes were analyzed by scanning from 300-200 nm while recording the absorbance. Each experiment is an average of 5 separate scans.

Qualitative Primer Extensions: The appropriate primer was labeled with [^{32}P] ATP and T4 polynucleotide kinase and was hybridized to the template (1.5 eqv.) at 90 °C for 5 min and slowly cooled to room temperature. Purification of the primer-template duplex from excess ATP was accomplished by passing it through a G-25 sephadex spin column. The duplex concentration for full-length polymerase extension was 200 nM in Klenow buffer (10 mM Tris, pH 7.5, 5 mM $MgCl_2$, 7.5 mM DTT). Klenow Fragment (exo$^-$) (5 U, 368 nM or 0.5 U, 36.8 nM) was added and the extension was initiated by the addition of 1 μL of a 10x solution of dNTP's (200 μM each NTP) and incubated at 37 °C. Aliquots were taken from the extension at 5, 10, 15, 30, 60, and 120 min and quenched with 95% formamide. The aliquots were denatured at 90°C for 2 min and immersed in a 0 °C bath. The reaction products were separated from unreacted primer by electrophoresis on a 20% PAGE.

Quantitative Primer Insertions Across from α-18 or β-18: The appropriate primer (75 pmol) was labeled for the standing start extension using standard protocols[221] and split

into three equal portions. The labeled primer was annealed to the appropriate template (375 pmol) and unlabled primer (225 pmol) in ST buffer (20 mM Tris, pH 7.5; 100 mM NaCl) for 5 min at 90° C and slowly cooled to room temperature. The duplexes were purified from excess [γ $-^{32}$P] ATP by passing through a G-25 sephadex spin column equilibrated with ST buffer. The final concentration of the duplex was 3.13 μM. Reactions were carried out by the addition of a 2x duplex-enzyme cocktail (5 μL) to a 2x dNTP (5 μL) solution. The reactions were run for a set period of time and quenched with 20 μL of 95% formamide loading buffer containing 20 mM EDTA. The samples were denatured (90° C for 3 min and cooled immediately to 0 ° C) and the unreacted primer was separated from the extended product by 20 % denaturing PAGE. The 2x duplex-enzyme cocktail consisted of 100 nM duplex, 20 nM Klenow fragment (exo$^-$), and 0.2 μg/μL BSA in 2x Klenow buffer (20 mM Tris, pH 7.5, 10 mM $MgCl_2$, 15 mM DTT). Klenow (exo$^-$) was pretreated with inorganic pyrophosphatase (0.2 U per unit of enzyme) at room temperature for 10 min before addition to the duplex. The 2x duplex-enzyme cocktail was incubated for 5 min at room temperature. The 2x dNTP solutions were made by diluting 100 mM stock solutions with the appropriate volume of water and stored at 4 °C. **Reaction Conditions:** The kinetics of insertion was measured in reactions incubated at room temperature. For the control [dNTP/time (min)]: T/3, A/4.5, C/4.5, G/4.5. Triphosphate concentrations for the control were 0.05-0.5 μM for TTP; 2-50 μM for dATP; and 50-500 μM for dCTP and dGTP. For β-**18** [dNTP/time (min)]: T/4.5, A/4.5, C/4.5, G/4.5. Triphosphate concentrations for the β-**18** were 0.1-1.0 μM for TTP; 6-100 μM for dATP; and 50-500 μM for dCTP and dGTP. For α-**18** [dNTP/time

(min)]: T/4.5, A/4.5, C/4.5, G/4.5. Triphosphate concentrations for the α-**18** were 5-15 μM for TTP; 25-100 μM for dATP; and 50-500 μM for dCTP and dGTP.

Quantitative Primer Insertions Across from β-19: Insertion experiments were performed in the same manner as described for the templates containing **18**. However, the amount of Klenow exo$^-$ used for insertions across from β-**19** was 10 nM in the 2x DNA-enzyme cocktail. **Reaction Conditions:** For the control [dNTP/time (min)]: C/3, A/4.5, G/3, T/4.5. Triphosphate concentration ranges for the control were 1-50 nM for dCTP, 25-250 μM dATP, 10-100 μM dGTP, and 50-600 μM TTP. For β-**19** [dNTP/time (min)]: C/4.5, A/4.5, G/4.5, T/15. Triphosphate concentration ranges for the control were 0.05-5.0 μM for dCTP, 1-25 μM dATP, 10-800 μM dGTP, and 0.2-1.2 mM TTP.

Gel-Shift Assays: Gel-shift assays were performed on duplexes containing α-**18**, β-**18**, and β-**19** with the modified oligonucleotide being 5'-^{32}P-end labeled. Binding solutions consisted of 50 pM duplex, 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM DTT, 1mM EDTA, 10 % glycerol, 0.1 μg/μL BSA, and varying amounts of the FPG protein in a volume of 20 or 100 μL. The FPG concentrations for duplexes β-**18**: T, A, or C and α-**18**: C were 0.1-200 nM. The FPG concentrations for duplexes β-**18**:G and β-**19**:C were 0.1-300 nM. The FPG protein was diluted with dilution buffer (10 mM HEPES (pH 7.4), 100 mM KCl, 10 mM EDTA, 1 μg/μL BSA) and stored at 0 ° C. The samples were incubated at room temperature for 20 min. Loading dye (30 % glycerol, 1x TBE, 0.25% bromophenol blue, 0.25% xylene cyanol) was added and the samples were electrophoresed on a 6% non-denaturing polyacrylamide gel with 0.5x TBE at 4 °C for 2

to 3 h at 500 V. The gel was dried for 1.5 h at 80 °C and exposed to a phosphorimager screen for 24 to 36 h and quantified. The data was plotted percent bound versus log [FPG concentration] and fitted to a sigmoidal curve to determine the K_d for the DNA-enzyme complex.

[13] Harman, D. "Role of Free Radicals in Ageing and Disease" *Ann. NY Acad. Sci.* **1992**, *673*, 126-141.

[14] Pacifici, R. E.; Davies, K. J. A. "Protein, Lipid and DNA Repair Systems in Oxidative Stress: The Free Radical Theory of Aging Revisited" *Gerontology* **1991**, *37*, 166-180.

[15] DePinho, R. A. "The Age of Cancer" *Nature* **2000**, *408*, 248-254.

[16] David, S. S.; David, S. D. "Chemistry of Glycosylases and Endonucleases Involved in Base-Excision Repair" *Chem. Rev.* **1998**, *98*, 1221-1261.

[17] Cadet, J.; Berger, M.; Douki, T.; Ravanat, J. L. "Oxidative Damage to DNA: Formation, Measurement, and Biological Significance" *Rev. Physiol. Biochem. Pharmacol.* **1997**, *131*, 1-87.

[18] Pogozelski, W. K.; Tullius, T. D. "Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety" *Chem. Rev.* **1998**, *98*, 1089-1107.

[19] Burrows, C. J.; Muller, J. G. "Oxidative Nucleobase Modifications Leading to Strand Scission" *Chem. Rev.* **1998**, *98*, 1109-1151.

[20] Boiteux, S.; Laval, J. "Repair of Oxidized Purines in DNA" In *Base Excision Repair of DNA Damage;* Hickson, I. D., Ed.; Landes Bioscience, Austin, Texas, 1997; pp31-44.

[21] Mullen, G. P.; Wilson, S. H. "Repair Activity in DNA Polymerases: a Structurally Conserved Helix-Hairpin-Helix Motif in Base Excision Repair Enzymes and in DNA Polymerase β" In *Base Excision Repair of DNA Damage;* Hickson, I. D., Ed.; Landes Bioscience, Austin, Texas, 1997; pp121-135.

[22] Hatahet, Z.; Wallace, S. S. "Translesional DNA Synthesis" In *DNA Damage and Repair: Volume I DNA Repair in Prokaryotes and Lower Eukaryotes*; Nickoloff, J. A.; Hoekstra, M. F., Eds. Humana Press, Totowa, New Jersey, 1998; pp229-262.

[23] Chepanoske, C. L.; Langlier, C. R.; Chmiel, N. H.; David, S. S. "Recognition of the Nonpolar Base 4-Metylindole in DNA by the DNA Repair Adenine Glycosylase MutY" *Org. Lett.* **2000**, *2*, 1341-1344.

[24] Laval, J.; Jurado, J.; Saparbaev, M.; Sidorkina, O. "Antimutagenic Role of Base-Excision Repair Enzymes Upon Free Radical-Induced DNA Damage" *Mutation Res.* **1998**, *402*, 93-102.

[25] Pouget, J.-P.; Douki, T.; Richard, M.-J.; Cadet, J. "DNA Damage Induced in Cells by γ and UVA Radiation As Measured by HPLC/GC-MS and HPLC-EC and Comet Assay" *Chem. Res. Toxicol.* **2000**, *13*, 541-549.

[26] Douki, T.; Martini, R.; Ravant, J.-L.; Turesky, R. J.; Cadet, J. "Measuremnet of 2,6-Diamino-4-Hydroxy-5-Formamidopyrimidine and 8-Oxo-7,8-Dihydroguanine in Isolated DNA Exposed to Gamma Radiation in Aqueous Solution" *Carcinogenesis* **1997**, *18*, 2385-2391.

[27] Douki, T.; Spinelli, S.; Ravanat, J.-L.; Cadet, J. "Hydroxyl Radical-Induced Degradation of 2'-Deoxyguanosine Under Reducing Conditions" *J. Chem. Soc., Perkin Trans. 2* **1999**, 1875-1880.

[28] Dizdaraglu, M. Aruoma, O. I.; Halliwell, B. "Modification of Bases in DNA by Copper Ion-1,10-Phenanthroline Complexes" *Biochemistry* **1990**, *29*, 8447-8451.

[29] Sies, H.; Menck, C. F. M. "Singlet Oxygen Induced DNA Damage" *Mutation Res.* **1992**, *275*, 367-375.

[30] Malins, D. C.; Haimanot, R. "Major Alterations in the Nucleotide Structure of DNA in Cancer of the Female Breast" *Cancer Res.* **1991**, *51*, 5430-5432.

[31] Malins, D. C.; Ostrander, G. K.; Haimanot, R.; Williams, P. "A Novel DNA Lesion in Neoplastic Livers of Feral Fish: 2,6-Diamino-4-hydroxy-5-formamidopyrimidine" *Carcinogenesis* **1990**, *11*, 1045-1047.

[32] Cadet, J.; Vigny, P. *Bioorganic Photochemistry*; Morrison, H. Ed.; Wiley, New York, 1990, vol. 1.

[33] Epe, B.; Pflaum, M.; Boiteux, S. "DNA Damage Induced by Photosensitizers in Cellular and Cell-Free Systems" *Mutation Res.* **1993**, *299*, 135-145.

[34] Steenken, S.; Jovanovic, S. "How Easily Oxidizable is DNA? One-Electron Reduction Potentials of Adenosine and Guanosine in Aqueous Solution" *J. Am. Chem. Soc.* **1997**, *119*, 617-618.

[35] Piette, J. "Biological Consequences Associated with DNA Oxidation Mediated by Singlet Oxygen" *J. Photochem. Photobiol. B* **1991**, *11*, 241-260.

[36] Redox Chemistry and Transformation Reactions of Their Radical Cations and e^- and OH Adducts" *Chem. Rev.* **1989**, *89*, 503-520.

[37] Fulford, J.; Nikjoo, H.; Goodhead, D. T.; O'Neill, P. O. "Yields of SSB and DSB Induced in DNA by Alk Ultrasoft X- Rays and α-particles: Comparison of Experimental and Simulated Yields" *Int. J. Radiat. Biol.* **2001**, *77*, 1053-1066.

[38] Aruoma, O. I.; Halliwell, B.; Dizdaroglu, M. "Iron Ion-Dependent Modification of Bases in DNA by the Superoxide Radical-Generating System Hypoxanthine / Xanthine Oxidase" *J. Biol. Chem.* **1989**, *264*, 13204-13028.

[39] Blakely, W .F.; Fuciarelli, A. F.; Wegher, B. J.; Dizdaraglu, M. "Hydrogen Peroxide-Induced Base Damage in Deoxyribonucleic Acid" *Radiat. Res.* **1990**, *121*, 338-343.

[40] Cadet, J.; Weinfeld, M. "Detecting DNA Damage" *Anal. Chem.* **1993**, *65*, 675a-682a.

[41] Podmore, I. D.; Cooper, D.; Evans, M. D.; Wood, M.; Lunec, J. "Simultaneous Measurement of 8-Oxo-2'-deoxyguanosine and 8-Oxo-2'-deoxyadenosine by HPLC-MS/MS" *Biochem. Biophys. Res. Commum.* **2000**, *277*, 764-770.

[42] Frelon, S.; Douki, T.; Ravanat, J.-L.; Pouget, J.-P.; Tornabene, C.; Cadet, J. "High-Performance Liquid Chromatography-Tandem Mass Spectrometry Measurement of Radiation-Induced Base Damage to Isolate and Cellular DNA" *Chem. Res. Toxicol.* **2000**, *13*, 1002-1010.

[43] Dizdaroglu, M.; "Chemical Determination of Oxidative DNA Damage by Gas Chromatography-Mass Spectroscopy" *Methods Enzymol.* **1994**, *234*, 3-16.

[44] England, T. G.; Jenner, A.; Aruoma, O. I.; Halliwell, B. "Determination of Oxidative DNA Base Damage by Gas Chromatography-Mass Spectrometry. Effect of Derivatization Conditions on Artifactual Formation of Certain Base Oxidation Products" *Free Rad. Res.* **1998**, *29*, 321-330.

[45] Mamer, O. A. "Stable Isotope Dilution Techniques" *J. Chromatogr. B* **1996**, *682*, 182-183.

[46] Cadet, J.; Douki, T.; Ravanat, J.-L. "Artifacts Associated with the Measurements of Oxidized DNA Bases" *Environmental Health Persp.* **1997**, *105*, 1034-1037.

[47] Evans, M. D.; Cooke, M. S.; Podmore, I. D.; Zheng, Q.; Herbert, K. E.; Lunec, J. "Discrepancies in the Measurement of UVC-Induced 8-Oxo-2'-deoxyguanosine: Implications for the Analysis of Oxidative DNA Damage" *Biochem. Biophys. Res. Commum.* **1999**, *259*, 374-378.

[48] Huang, Xi, Powell, J.; Mooney, L. A.; Chenlu, L.; Frenkel, K. "Importance of Complete DNA Digestion in Minimizing Variability of 8-OxodG Analyses" *Free Radic. Biol. Med.* **2001**, *31*, 1341-1351.

[49] Angelis, K. J.; Dusinska, M.; Collins, A. R. "Single Cell Electrophoresis: Detection of DNA Damage at Different Levels of Sensitivity" *Electrophoresis* **1999**, *20*, 2133-2138.

[50] Collins, A. R.; Dobson, V. L.; Dusinska, M.; Kennedy, G.; Stetina, R. "The Comet Assay: What Can it Really Tell Us?" *Mutation Res.* **1997**, *375*, 183-193.

[51] Pouget, J.-P.; Ravanat, J.-L.; Douki, T.; Richard, M.-J.; Cadet, J. "Measurement of DNA Base Damage in Cells Exposed to Low Doses of γ-Radiation: Comparison between the HPLC-EC and Comet Assays" *Int. J. Radiat. Biol.* **1999**, *75*, 51-58.

[52] Candeias, L. P.; Steenken, S. "Reaction of $HO^{\bullet}$ with Guanine Derivatives in Aqueous Solution: Formation of Two Different Redox-Active OH-Adduct Radicals and Their Unimolecular Transformation Reactions. Properties of $G(-H)^{\bullet}$" *Chem. Eur. J.* **2000**, *6*, 475-484.

[53] Viera, A. J.S.C.; Steenken, S. "Pattern of OH Radical Reaction with Adenine and Its Nucleosides and Nucleotides. Characterization of Two Types of Isomeric OH Adduct and Their Unimolecular Transformation Reactions" *J. Am. Chem. Soc.* **1990**, *112*, 6986-6994.

[54] Cadet, J.; Berger, M.; Buchko, G. W.; Joshi, P. C.; Raoul, S.; Ravanat, J.-L. "2,2-Diamino-4-[(3,5-di-*O*-acetyl-2-deoxy-β-D-erythro-pentofuranosyl)amino]-5-(2*H*)-oxazolone: A Novel and Predominant Radical Oxidation Product of 3',5'-Di-*O*-scetyl-2'-deoxyguanosine" *J. Am. Chem. Soc.* **1994**, *116*, 7403-7404.

[55] Breen, A. P.; Murphy, J. A. "Reactions of Oxyl Radicals with DNA" *Free Rad. Biol. Med.* **1995**, *18*, 1033-1077.

[56] Luo, W.; Muller, J. G.; Rachlin, E. M.; Burrows, C. J. "Characterization of Hydantoin Products from One-Electron Oxidation of 8-Oxo-7,8-dihydroxyguanosine in a Nucleoside Model" *Chem. Res. Toxicol.* **2001**, *14*, 927-938.

[57] Gajewski, E.; Rao, G.; Nackerdien, Z.; Dizdaroglu, M. "Modification of DNA Bases in Mammalian Chromatin by Radiation-Generated Free Radicals" *Biochemisrtry* **1990**, *29*, 7876-7882.

[58] Aruoma, O. I.; Halliwell, B.; Gajewski, E.; Dizdaroglu, M. "Damage to the Bases in DNA Induced by Hydrogen Peroxide and Ferric Ion Chelates" *J. Biol. Chem.* **1989**, *264*, 20509-20512.

[59] Tomasz, M.; Lipman, R.; Lee, M. S.; Verdine, G. L.; Nakanishi, K. "Reaction of Acid Activated Mitomycin C with Calf Thymus DNA and Model Guanines:

Elucidation of the Base-Catalyzed Degradation of *N*7-Alkylguanine Nucleosides" *Biochemistry* **1987**, *26*, 2010-2027.

60 Mao, H.; Deng, Z.; Wang, F.; Harris, T. M.; Stone, M. P. "An Intercalated and Thermally Stable FAPY Adduct of Aflatoxin B_1 in a DNA Duplex: Structural Refinement from ^{1}H NMR" *Biochemistry* **1998**, *37*, 4374-4387.

61 Kadlubar, F. F.; Beranek, D. T.; Weis, C. C.; Evans, F. E.; Cox, R.; Irving, C. C "Characterization of the Purine Ring-Opened 7-Methylguanine and its Persistence in Rat Bladder Epithelia DNA after Treatment with the Carcinogen *N*-Methylnitrosourea" *Carcinogenesis* **1984**, *5*, 587-592.

62 Li, Q. Laval, J.; Lindlum, D. B. "Fpg Protein Releases a Ring Opened *N*-7 Guanine Adduct from DNA That Has Been Modified by Sulfur Mustard" *Carcinogenesis* **1997**, *18*, 1035-1038.

63 Kim, D.-H.; Humphreys, W. G.; Guengrich, F. P. "Characterization of *S*-[2-(*N*^{1}-Adenyl)ethyl]glutathione as an Adduct Formed in RNA and DNA from 1,2-Dibromoethane" *Chem. Res. Toxicol.* **1990**, *3*, 587-594.

64 Humphreys, W. G.; Guengerich, F. P. "Structure of Formamidopyrimidine Adducts as Determined by NMR Using specifically ^{15}N-labled Guanosine" *Chem. Res. Toxicol.* **1991**, *4*, 632-636.

65 Haraguchi, K.; Greenberg, M. M. "Synthesis of Oligonucleotides Containing Fapy•dG (*N*-6-(2-Deoxy-α,β-D-*erythro*-pentofuranosyl)-2,6-diamino-4-hydroxy-5-formamidopyrimidine)" *J. Am. Chem. Soc.* **2001**, *123*, 8636-8637.

66 Haraguci, K.; Delaney, M. O.; Wiederholt, C. J.; Sambandam, A.; Hantosi, Z.; Greenberg, M. M. "Synthesis and Characterization of Oligodeoxynucleotides Containing Formamidopyrimidine Lesions and Nonhydrolyzable Analogues" *J. Am. Chem. Soc.* **2002**, *124*, 3263-3269.

67 Haraguci, K.; Delaney, M. O.; Wiederholt, C. J.; Sambandam, A.; Hantosi, Z.; Greenberg, M. M. "Synthesis and Characterization of Oligonucleotides Containing Formamidopyrimidine Lesions (Fapy•dA, Fapy•dG) at Defined Sites" *Nucleic Acids Res. Suppl. 1*, **2001**, 129-130.

68 Greenberg, M. M.; Hantosi, Z.; Wiederholt, C. J.; Rithner, C. D. "Studies on *N*4-(2-Deoxy-D-pentofuranosyl)-4,6-diamino-5-formamidopyrimidine (Fapy•dA) and *N*6-(2-Deoxy-D-pentofurnaosyl)-6-diamino-5-formamido-4-hydroxypyrimidine (Fapy•dG)" *Biochemistry* **2001**, *40*, 15856-15861.

69 Burgdorf, L. T.; Carell, T. "Synthesis, Stability, and Conformation of the Formamidopyrimidine G DNA Lesion" *Chem. Eur. J.* **2002**, *8*, 293-301.

[70] Raoul, S.; Bardet, M.; Cadet, J. "γ-Irradiation of 2'-Deoxyadenosine in Oxygen-Free Aqueous Solutions: Identification and Conformational Features of Formamidopyrimidine Nucleoside Derivatives" *Chem. Res. Toxicol.* **1995**, *8*, 924-933.

[71] Cysewski, P. 'Theoretical Studies on the Tautomeric Properties of Diamino-5-formamidopyrimidines" *Z. Naturforsch.* **1998**, *53c*, 1027-1036.

[72] Cysewski, P.; Olinski, R. "Theoretical Description of the Coding Potential of Diamino-5-formamidopyrimidines" *Z. Naturforsch.* **1999**, *54c*, 239-245.

[73] Echols, H.; Goodman, M. F. "Fidelity Mechanisms in DNA Replication" *Ann. Rev. Biochem.* **1991**, *60*, 477-511.

[74] Kool, E. T. "Replication of Non-hydrogen Bonded Bases by DNA Polymerases: A Mechanism for Steric Matching" *Biopolymers* **1998**, *48*, 3-17.

[75] Dzantiev, L.; Alekseyev, Y. O.; Morales, J. C.; Kool, E. T.; Romano, L. J. "Significance of Nucleobase Shape Complementarity and Hydrogen Bonding in the Formation and Stability of the Closed Polymerase—DNA Complex" *Biochemistry* **2001**, *40*, 3215-3221.

[76] Polymerase I (Klenow Fragment) and the Minor Groove of DNA by Amino Acid Substitution of the Polymerase and Atomic Substitution of the DNA" *Biochemistry* **2001**, *40*, 2647-2652.

[77] Paz-Elizur, T. Barak, Y.; Livneh, Z. "Anti-Mutagenic Activity of DNA Damage-Binding Proteins Mediated by Direct Inhibition of Translesion Replication" *J. Biol. Chem.* **1997**, *272*, 28906-28911.

[78] Goodman, M. F.; Creighton, S.; Bloom, L. B.; Petruska, J. "Biochemical Basis of DNA Replication Fidelity" *Crit. Rev. Biochem. Molec. Bio.* **1993**, *28*, 83-126.

[79] Creighton, S.; Bloom, Linda, B.; Goodman, M. F. "Gel Fidelity Assay Mesuring Nucleotide Misinsertion, Exonucleolytic Proofreading, and Lesion Bypass Efficiencies" *Methods Enzymol.* **1995**, *262*, 232-257.

[80] Petriska, J.; Goodman, M. F.; Boosalis, M. S.; Sowers, L. C.; Cheong, C.; Tinoco, I. "Comparison Between DNA Melting Thermodynamics and DNA Polymerase Fidelity" *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6252-6256.

[81] Moriya, M. "Single-Stranded Shuttle Phagemid for Mutagenesis Studies in Mammalian Cells: 8-Oxoguanine in DNA Induces G:C→T:A Transversions in Simian Kidney Cells" *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1122-1126.

[82] Tan, X.; Grollman, A. P.; Shibutani, S. "Comparison of the Mutagenic Properties of the 8-Oxo-7,8-dihydro-2'-deoxyadenosine, and 8-Oxo-7,8-dihydro-2'-deoxyguanosine DNA Lesions in Mammalian Cells" *Carcinogenesis* **1999**, *20*, 2287-2292.

[83] Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; Oxford University Press, Oxford, 1996.

[84] Breslauer, K. J. "Extracting Thermodynamic Data from Equilibrium Melting Curves for Oligonucleotide Order-Disorder Transitions" *Methods Enzymol.* **1995**, *259*, 221-242.

[85] Breslauer, K. J. *Methods in Molecular Biology, Volume 26: Protocols for Oligonucleotide Conjugates*; Agrawal, S., Ed.; Humana Press: New Jersey, 1994.

[86] Shibutani, S.; Takeshita, M.; Grollman, A. P. "Insertion of Specific Bases During DNA Synthesis Past the Oxidation-Damaged Base 8-OxodG" *Nature* **1991**, *349*, 431-434.

[87] Lowe, L. G.; Guengerich, F. P. "Steady-State and Pre-Steady-State Kinetic Analysis of dNTP Insertion Opposite 8-Oxo-7,8-dihydroguanine by *Escherichia coli* Polymerase I exo$^-$ and II exo$^-$" *Biochemistry* **1996**, *35*, 9840-9849.

[88] Moriya, M., Ou, C.; Bodepudi, V.; Johnson, F.; Takeshita, M.; Grollman, A. P. "Site-Specific Mutagenesis Using a Gapped Duplex Vector—A Study of Translesional Synthesis Past 8-Oxodeoxyguanosine in *Esherichia coli*" *Mutation Research* **1991**, *254*, 281-288.

[89] Wood, M. L.; Dizdaroglu, M.; Gajewski, E.; Essigmann, J. M. "Mechanistic Studies of Ionizing-Radiation And Oxidative Mutagenesis—Genetic-Effects of a Single 8-HydroxyGuanine (7-Hydro-8-oxoguanine) Residue Inserted at a Unique Site in a Viral Genome" *Biochemistry*, **1990**, *29*, 7024-7032.

[90] McAuley-Hecht, K. E.; Leonard, G. A.; Gibson, N. J.; Thomson, J. B.; Watson, W. P.; Hunter, W. N.; Brown, T. "Crystal Structure of a DNA Duplex Containing 8-Hydroxydeoxyguanine-Adenine Base Pairs" *Biochemistry* **1994**, *33*, 10266-10270.

[91] Plum, E. P.; Grollman, A. P.; Johnson, F.; Breslauer, K. J. "Influence of the Oxidatively Damaged Adduct 8-Oxodeoxyguanosine on the Conformation, Energetics, and Thermodynamic Stability of a DNA Duplex" *Biochemistry* **1995**, *34*, 16148-16160.

[92] Shibutani, S.; Bodepudi, V.; Johnson, F.; Grollman, A. P. "Translesional Synthesis on DNA Templates Containing 8-Oxo-7,8-dihydrodeoxyadenosine" *Biochemistry* **1993**, *32*, 4615-4621.

[93] Kamiya, H.; Miura, H.; Murata-Kamiya, N.; Ishikawa, H.; Sakaguchi, T.; Inoue, H.; Sasaki, T.; Masutani, H.; Hanaoka, F.; Nisimura, S.; Ohtsuka, E. "8-Hydroxyadenine(7,8-Dihydro-8-oxoadenine) Induces Misincorporation in *In Vitro* DNA Synthesis and Mutations in NIH 3T3 Cells" *Nucleic Acids Res.* **1995**, *23*, 2893-2899.

[94] Leonard, G. A.; Guy, A.; Brown, T.; Teoule, R.; Hunter, W. N. "Conformation of Guanine•8-Oxoadenine Base Pairs in the Crystal Structure of d(CGCGAATT(O8A)GCG)" *Biochemistry* **1992**. *31*, 8415-8420.

[95] Tudek, B.; Boitex, S.; Laval, J. "Biological Properties of Imidazole Ring-Opened *N*7-Methylguanine in M13mp18 phage DNA" *Nucleic Acids Res.* **1992**, *20*, 3079-3084.

[96] Boiteux, S.; Laval, J. "Imidazole Open Ring 7-Methylguanine: An Inhibitor of DNA Synthesis" *Biochem. Biophys. Res. Commun.* **1983**, *110*, 552-558.

[97] Ring-Opened 7-Methylguanine Residues in DNA are a Block to *In Vitro* DNA Synthesis" *Nucleic Acids Res.* **1988**, *16*, 5879-5895.

[98] Graziewicz, M.-A.; Zastawany, T. H.; Olinski, R.; Tudek, B. "SOS-Dependent A→G Transitions Induced by Hydroxyl Radical Generating System Hypoxanthine/Xanthine Oxidase/Fe^{+3}/EDTA are Accompanied by the Increase of Fapy-Adenine Content in M13mp18 Phage DNA" *Mutation Res.* **1999**, *434*, 41-52.

[99] Tudek, Graziewicz, Kazanova, O.; Zastawany, T. H.; Obtulowicz, T.; Laval, J. "Mutagenic Specificity of Imidazole Ring-Opened 7-Methylpurines in M13mp18 Phage DNA" *Acta Biochem. Pol.* **1999**, *46*, 785-799.

[100] Asagoshi, K.; Yamada, T.; Okada, Y.; Terato, H.; Ohyama, Y.; Seki, S.; Ide, H. "Recognition of Formamidopyrimidine by *Escherichia coli* and Mammalian Thymine Glycol Glycosylases" *J. Biol. Chem.* **2000**, *275*, 24781-24786.

[101] Asagoshi, Y.; Terato, H., Ohyama, Y.; Monden, Y.; Arai, T.; Nishimura, S.; Aburatani, H.; Lindahl, T.; Ide, H. "Distinct Repair Activities of 7,8-Dihydro-8-oxoguanine DNA Glycosylase and Formamidopyrimidine DNA Glycosylase for Formamidopyrimidine and 7,8-Dihydro-8-oxoguanine" *J. Biol. Chem.* **2000**, *275*, 4956-4964.

[102] Asagoshi, Y.; Terato, H.; Ohyama, Y.; Ide, H. "Influences of a Guanine-Derived Formamidopyrimidine Lesion on DNA Replication. Translesional DNA

Synthesis, Nucleotide Insertion and Extension Kinetics" *J. Biol. Chem.* **2002**, *277*, 14589-14597.

[103] Wiederholt, C. J.; Greenberg, M. M. "Fapy•dG Instructs Klenow Exo⁻ to Misincorporate Deoxyadenosine" *J. Am. Chem. Soc.* **2002**, (accepted for publication).

[104] Delaney, M. O.; Wiederholt, C. J.; Greenberg, M. M "Fapy•dA Induces Nucleotide Misincorporation Translesionally by a DNA Polymerase" *Angew. Chem. Int. Ed.* **2002**, *41*, 771-773.

[105] Echols, H.; Goodman, M. F. "Mutation induced by DNA Damage: A Many Protein Affair" *Mutation Res.* **1990**, *236*, 301-311.

[106] Singer, B.; Hang, B. "What Structural Features Determine Repair Enzyme Specificity and Mechanism in Chemically Modified DNA?" *Chem. Res. Toxicol.* **1997**, *10*, 713-732.

[107] Krokan, H, E,; Standal, R.; Slupphaug, G. "DNA Glycosylases in the Base Excision Repair of DNA" *Biochem. J.* **325**, 1-16.

[108] Boiteuz, S.; O'Connor, T. O.; Laval, J. "Formamido-DNA Glycosylase of *Escherichia coli*: Cloning and Sequencing of the *fpg* Structural Gene and Overproduction of the Protein" *EMBO J.* **1987**, *6*, 3177-3183.

[109] Sugahara, M.; Mikawa, T.; Kumasaka, Yamamoto, M.; Kato, R.; Fukuyama, K.; Inoue, Y.; Kuramitsu, S. "Crystal Structure of a Repair Enzyme of Oxidatively Damaged DNA, MutM (Fpg), From an Extreme Thermophile, *Thermus thermophilus* HB8" *EMBO J.* **2000**, *19*, 3857-3869.

[110] O'Connor, T. J.; Laval, J. "Physical Association of the 2,6-Diamino-4-hydroxy-5-*N*-formamidopyrimidine-DNA Glycosylase of *Escherichia Coli* and an Activity Nicking DNA at Apurinic/ Apyrimidic Sites" *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5222-5226.

[111] Bhagwhat, M.; Gerlt, J. A. "3'- and 5'-Strand Cleavage Reactions Catalyzed by the Fpg Protein from *Escherichia coli* Occur Via Successive β- and δ-Elimination Mechanisms" *Biochemistry* **1996**. *35*, 659-665.

[112] Graves, R.J.; Felzenszwabb, J.; Laval, J.; O'Connor, T. "Excision of 5'-Terminal Deoxyribose Phosphate From Damaged DNA is Catalyzed by the Fpg Protein of *Escherichia Coli*" *J. Biol. Chem.* **1992**, *267*, 14429-14435.

[113] O'Connor, T. R.; Graves, R. J.; de Murcia, G.; Castaing, B.; Laval, J. "Fpg Protein of *Escherichia Coli* is a Zinc Finger Protein Whose Cysteine Residues have a Structural and/ or Functional Role" *J. Biol. Chem.* **1993**, *268*, 9063-9070.

[114] Cabera, M.; Nghiem, Y.; Miller, J. H. "*mutM*, a Second Mutator Locus in *Escherichia Coli* that generates G•C.→T•A Transversions" *J. Bacteriol.* **1988**, *170*, 5405-5407.

[115] Michaels, M. L.; Cruz, C.; Grollman, A. P.; Miller, H. H. "Evidence that MutY and Mut M Combine to Prevent Mutations by an Oxidatively Damaged Form of Guanine in DNA" *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7022-7025.

[116] Castaing, B.; Fourrey, J.-L.; Hervouet, N.; Thomas, M.; Boiteux, S.; Zelwer, C. "AP Site Structural Determinants for Fpg Specific Recognition" *Nucleic Acids Res.* **1999**, *27*, 608-615.

[117] Kuznetsov, S. V.; Sidorkina, O. M.; Jurado, J.; Bazin, M.; Tauc, P,; Brochon, J.-C.; Laval, J. "Effect of Single Mutations on the Structural Dynamics of a DNA Repair Enzyme, the *Escherichia coli* Formamidopyrimidine-DNA Glycosylase. A Fluorescence Study Using Tryptophan Residues as Reporter Groups" *Eur. J. Biochem.* **1998**, *253*, 413-420.

[118] Sidorkina, O. M.; Laval, J. "Role of Lysine-57 in the Catalytic Activities of *Escherichia coli* Formamidopyrimidine-DNA Glycosylase (Fpg Protein)" *Nucleic Acids Res.* **1998**, *26*, 5351-5357.

[119] Rabow, L. E.; Kow, Y. W. "Mechanism of Action of Base Release by *Escherichia coli* Fpg Protein: Role of Lysine 155 in Catalysis" *Biochemistry* **1997**, *36*, 5084-5096.

[120] Lavrukhin, O. V.; Llyod, R. S. "Involvement of Phylogenetically Conserved Acidic Amino Acid Residues in Catalysis by an Oxidative DNA Damage Enzyme Formamidopyrimidine Glycosylase" *Biochemistry* **2000**, *39*, 15266-15271.

[121] Zharkov, D. O.; Rieger, R. A.; Iden, C. R.; Grollman, A. P. "NH_2-Terminal Proline Acts as a Nucleophile in the Glycosylase/ AP-Lyase Reaction Catalyzed by *Escherichia coli* Formamidopyrimidine-DNA Glycosylase (Fpg) Protein" *J. Biol. Chem.* **1997**, *272*, 5335-5341.

[122] Parikh, S. S.; Mol, C. D.; Slupphaug, G.; Bharati, S.; Krokan, H. E.; Tainer, J. A. "Base Excision Repair Initiation Revealed by Crystal Structures and Binding Kinetics of Human Uracil-DNA Glycosylase with DNA" *EMBO J.* **1998**, *17*, 5214-5226.

[123] Bruner, S. D. Norman., D. P. G.; Verdine, G. L. "Structural Basis for Recognition and Repair of the Endogenous Mutagen 8-Oxoguanine in DNA" *Nature* **2000**, *403*, 859-866.

[133] Glycosylase Activity of *E. coli* T4 Endonuclease V that Excises 4,6-Diamino-5-formamidopyrimidine form DNA, a UV-radiation- and Hydroxyl Radical-Induced Product of Adenine" *Mutation Res.* **1996**, *362*, 1-8.

[134] Jaruga, P.; Jabil, R.; McCullough, A. K.; Rodriguez, H.; Dizdaroglu, M.; Lloyd, R. S. "Chlorella Virus Pyrimidine Dimer Glycosylase Excises Ultraviolet Radiation-and Hydroxyl Radical-Induced Products 4,6-Diamino-5-formamidopyrimidine and 2,6-Diamino-4-hydroxy-5-formamidopyrimidine from DNA" *Photochem. Photobiol.* **2002**, *75*, 85-91.

[135] Beaucage, S. L.; Caruthers, M. H.; "Chemical Synthesis of DNA/ RNA" In *Bioorganic Chemistry: Nucleic Acids*; Hecht, S. M., Ed.; Oxford University Press, New York, 1996; pp36-75.

[136] Beaucage, S. L. In *Comprehensive Natural Products Chemistry, Vol. 7, DNA and Aspects of Molecular Biology*; Kool, E. T., Ed.; Pergamon: Oxford, 1999; pp105-152.

[137] Scharer, O. D.; Verdine, G. L. "A Designed Inhibitor of Base-Excision Repair" *J. Am. Chem. Soc.* **1995**, *117*, 10781-10782.

[138] Chepanoske, C. L.; Porello, S. L.; Fujiwara, T.; Sugiyama, H.; David, S. S. "Substrate Recognition by *Escherichia coli* MutY Using Substrate Analogs" *Nucleic Acids Res.* **1999**, *27*, 3197-3204.

[139] Horenstein, B. A.; Zabinski, R. F.; Schram, V. L. "A New Class of C-Nucleoside Analogues. 1-(S)-Aryl-1,4-dideoxy-1,4-imino-D-ribitols, Transition State Analogue Inhibitors of Nucleoside Hydrolase" *Tetrahedron Lett.* **1993**, *34*, 7213-7216.

[140] Horenstein, B. A.; Schram, V. L. "Correlation of the Molecular Electrostatic Potential Energy Surface of an Enzymatic Transition State with Novel Transition-State Inhibitors" *Biochemistry* **1993**, *32*, 9917-9925.

[141] Scharer, O. D.; Ortholand, J.-Y.; Ganesan, A.; Ezaz-Nikpay, K.; Verdine, G. L. "Specific Binding of the DNA Repair Enzyme AlkA to a Pyrrolidine-Based Inhibitor" *J. Am. Chem. Soc.* **1995**, *117*, 6023-6024.

[142] Deng, Li., Scharer, O. D.; Verdine, G. L. "Unusually Strong Binding of a Designed Transition-State Analog to a Base-Excision DNA Repair Protein" *J. Am. Chem. Soc.* **1997**, *119*, 7865-7866.

[143] Zharkov, D. O.; Rosenquist, T. A.;Gerchman, S. E.; Grollman, A. P. "Substrate Specificity and Reaction Mechanism of Murine 8-Oxoguanine-DNA Glycosylase" *J. Biol. Chem.* **2000**, *275*, 28607-28617.

[144] Zharkov, D. O.; Gilboa, R.; Yagil, I.; Kycia, J. H.; Gerchman, S. E.; Shoham, G.; Grollman, A. P. "Role for Lysine 142 in the Excision of Adenine from A:G Mispairs by MutY DNA Glycosylase of *Escherichia coli*" *Biochemistry* **2000**, 14768-14778.

[145] Porello, S. L.; Williams, S. D.; Kuhn, H.; Michaels, M. L.; David, S. S. "Specific Recognition of Substrate Analogs of the DNA Mismatch Repair Enzyme MutY" *J. Am. Chem. Soc.* **1996**, *118*, 10684-10692.

[146] Watanabe, K. A. "The Chemistry of C-Nucleosides" In *Chemistry of Nucleosides and Nucleotides, Volume 3*; Townsend, L. B., Eds.; Plenum Press, New York, 1994; 421-535.

[147] Carey, J. "Gel Retardation" *Methods Enzymol.* **1991**, *208*, 103-117.

[148] Carey, J. "Gel Retardation at Low pH Resolves *trp* Repressor-DNA Complexes for Quantitative Study" *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 975-979.

[149] Schulhof, J. C.; Molko, D.; Teoule, R. "The Final Deprotection Step in Oligonucleotide Synthesis is Reduced to a Mild and Rapid Ammonia Treatment by Using Labile Base-Protecting Groups" *Nucleic Acids Res.* **1988**, *15*, 397-416.

[150] Postema, M. H. D. *C-Glycoside Synthesis*; CRC Press, Boca Raton, 1995.

[151] Somsak, L. "Carbanionic Reactivity of the Anomeric Center in Carbohydrates" *Chem. Rev.* **2001**, *101*, 81-135.

[152] Secrist III, J. A. "Homo-*C*-Nucleosides. The Synthesis of Certain 6-Substituted 4-Pyrimidinone" *J. Org. Chem.* **1978**, *43*, 2925-2927.

[153] Cupps, T. L.; Wise, D. L.; Townsend, L. B. "Use of Allyltrimethylsilane in the Formation of Potential *C*-Nucleoside Precursors" *J. Org. Chem.* **1982**, *47*, 5115-5120.

[154] Guindon, Y.; Delorme, D.; Lau, C. K.; Zamboni, R. "Total Synthesis of LTB_4 and Analogues" *J. Org. Chem.* **1988**, *53*, 267-275.

[155] Walker, J. A.; Chen, J. J.; Wise, D. S.; Townsend, L. B. "A Facile, Multi-gram Synthesis of Ribofuranoid Glycals" *J. Org. Chem.* **1996**, *61*, 2219-2221.

[156] Wichai, U.; Woski, S. A. "An Improved Route to 1,2-Dideoxy-β-1-Phenyl-D-Ribofuransoe" *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3465-3568.

[157] Brown, D. J. *The Pyrimidines*, John Wiley & Sons, New York, 1994.

[158] DeGraw, J. I; Brown, V. H. "Synthesis of 2-Amino-4-hydroxy-1,3,5-triazanapthalenes" *J. Heterocyclic Chem.* **1976**, *13*, 439-441.

[159] Giziewicz, J.; Wnuk, S. F.; Robins, M. J. "Nucleic Acid Related Compounds. 107. Efficent Nitration of Uracil Base and Nucleoside Derivatives" *J. Org. Chem.* **1999**, *64*, 2149-2151.

[160] Pirrung, M. C.; Zhao, X.; Harris, S. V.; "A Universal, Photocleavable DNA Base: Nitropiperonyl 2'-Deoxyriboside" *J. Org. Chem.* **2001**, *66*, 2067-2071.

[161] Huang, G.-F.; Torrence, P. F. "Nitration of Pyrimidine Bases and Nucleosides Nitronium Tetrafluoroborate. Synthesis of 5-Nitro-2'-deoxyuridine" *J. Org. Chem.* **1977**, *42*, 3831-3824.

[162] Taylor, E. C.; Young, W. B.; "Pyrrolo[3,2,*d*]pyrimidine Folate Analogues: "Inverted" Analogues of the Cytotoxic Agent LY231514" *J. Org. Chem.* **1995**, *60*, 7947-7952.

[163] Yang, B. H.; Buchwald, S. L. "Palladium-Catalyzed Amination of Aryl Halides and Sulfonates" *J. Organomet. Chem.* **1999**, *576*, 125-146.

[164] Hartwig, J. H. "Transition Metal Catalyzed Synthesis of Arylamines and Aryl Ethers from Aryl Halides and Triflates: Scope and Mechanism" *Angew. Chem. Int. Ed.* **1998**, *37*, 2046-2067.

[165] Laksham, M. K.; Keeler, J. C.; Hilmer, J. H.; Martin, J. Q. "Palladium-Catalyzed C-N Bond Formation: Facile and General Synthesis of N^6-Aryl 2'-Deoxyadenosine Analogues" *J. Am. Chem. Soc.* **1999**, *121*, 6090-6091.

[166] Harwood, E. A.; Sigurdsson, S. T.; Edfeldt, N. B. F.; Reid, B. R.; Hopkins, P. B. "Chemical Synthesis and Preliminary Structural Characterization of a Nitrous Acid Interstrand Cross-Linked DNA" *J. Am. Chem. Soc.* **1999**, *121*, 5081-5082.

[167] Kahl, J. D.; Greenberg, M. M. "Introducing Structural Diversity in Oligonucleotides via Photolabile, Convertible C5-Substituted Nucleotides" *J. Am. Chem. Soc.* **1999**, *121*, 597-604 and references therein.

[168] Robins, M. J.; Barr, P. J.; Giziewicz, J. "Nucleic Acid Related Compounds. 38. Smooth Yield Iodination and Chlorination at C-5 of Uracil Bases and *p*-Toluyl-Protected Nucleosides" *Can. J. Chem.* **1982**, *60*, 554-557.

[169] Harnden, M. R.; Hurst, D. T. "The Chemistry of Pyrimidinethiols. III The Synthesis of Some Substituted Pyrimidinethiols and Some Thiazolo[5,4-*d*]pyrimidines" *Aust. J. Chem.* **1990**,*43*, 55-62.

[170] Greck, C.; Genet, J. P. "Electrophilic Amination: New Synthetic Applications" *Synlett* **1997**, 741-748.

[171] Evans, D. A.; Britton, T. C.; Ellman, J. A.; Dorow, R. L. "The Asymmetric Synthesis of α-Amino Acids. Electrophilic Azidation of Chiral Imide Enolates, a practical Approach to the Synthesis of (*R*)- and (*S*)-α-Azido Carboxylic Acids" *J. Am. Chem. Soc.* **1990**, *112*, 4011-4030.

[172] Evans, D. A.; Britton, T. C.; "Electrophilic Azide Transfer to Chiral Enolates. A General Approach to the Asymmetric Synthesis of α-Amino Acids" *J. Am. Chem. Soc.* **1987**, *109*, 6881-6883.

[173] Hauser, F. M.; Ellenberger, W. P. "Synthesis of *ribo* and *arabino* Deoxy- and Deoxyaminoepoxybenzoxocin Sugar Analogues" *J. Org. Chem.* **1988**, *53*, 118-1121.

[174] Harding, K. E.; Moreno, L. N.; Nace, V. M. "Hydroxylation and Acylation Reactions of Methyl Hippurate" *J. Org. Chem.* **1981**, *46*, 2809-2812.

[175] Lachicotte, R. J.; Hines, M. A.; Mazat, C. L.; Taylor, R. A.; Beam, C. F. "C-Acylation of Hippurate and Related Esters Utilizing Multiple Anion Intermediates and Ethyl-4-aminobenzoate" *Synthetic Commun.* **1990**, *20*, 63-69.

[176] Suzuki, M.; Iwasaki, T.; Miyoshi, M.; Okumura, K.; Matsumoto, K. "New Convenient Syntheses of α-*C*-Acylamino Acids and α-Amino Ketones" *J. Org. Chem.* **1973**, *38*, 3571-3575.

[177] Bachi, M. D.; Melman, A. "Stereocontrolled 5-*Exo-Trig* Cyclization of Imidoyl Radicals in the Synthesis of Substituted (Alkylthio)pyrroline, Pyroglutamates, and Thiopyroglutamates" *J. Org. Chem.* **1995**, *60*, 6242-6244.

[178] Fornicola, R. S.; Oblinger, E.; Montogomery, J. "A New Synthesis of α-Amino Acid Derivatives Employing Methyl Nitroacetate as a Versatile Glycine Template" *J. Org. Chem.* **1998**, *63*, 3528-3529.

[179] van der Werf, A. W.; Kellogg, R. M.; van Bolhuis, F. "Zinc Chloride Induced Stereoselection in Syntheses of α-Amino-β-hydroxy Acid Derivatives" *J. Chem. Soc. Chem. Commun.* **1991**, 682-683.

[180] Kanemasa, S.; Mori, T.; Wada, E.; Tatsukawa, A. "*anti*-Selective Aldol Reactions of Titanium Enolates of *N*-Alkylideneglycinates. Stereoselective Synthesis of *anti*-Isomers of β-Hydroxy-α-amino Esters" *Tetrahedron Lett.* **1993**, *34*, 677-680.

[181] Rao, M. N.; Holkar, A. G.; Ayyangar, N. R. "Unusual N-Acylation of Glycine Schiff Bases. A Simple approach to 3,3-Diphenylaziridine-2-carboxylates" *Tetrahedron Lett.* **1989**, *30*, 4717-4720.

[182] Muzart, J. "Silyl Ethers as Protective Groups for Alcohols: Oxidative Deprotection and Stability under Alcohol Oxidation Conditions" *Synthesis* **1993**, 11-27 and references therein.

[183] Bernstein, M. A.; Morton, H. E.; Guindon, Y. "A General Method for Determining the Anomeric Configuration of C-Furanoside Derivatives: A Proton Nuclear Magnetic Resonance Nuclear Overhauser Effect Study" *J. Chem. Soc., Perkin Trans. 2* **1986**. *8*, 1155-1163.

[184] Johnson, C. J.; Johns, B. A. "Suzuki Cross-Coupling if Carbohydrates: Synthesis of β-Arylmethyl-C-glycosides and Aryl-Scaffolded Trisaccharide Mimics" *Synlett* **1997**, 1406-1408.

[185] Link, J. T.; Sorensen, B. K. "A Method for Preparing *C*-Glycosides Related to Phlorizin" *Tetrahedron Lett.* **2000**, *41*, 9213-9217.

[186] Herrera, F. J. L.; Gonzalez, M. S. P.; Aguas, R. P.; "Structure and Reactivity of D-Ribofuranosylpyrimidine Homo-*C*-Nucleosides Halogenated on the Bridge C-Atom" *J. Chem. Soc. Perkin Trans. I* **1989**, 2401-2406.

[187] Katagiri, N.; Takashima, K.; Kato, Sato, S.; Tamura, C. "Sythesis of Homo-*C*-Nucleosides Using Pyrimidinylmethylenephosphoranes" *J. Chem. Soc. Perkin Trans. I* **1983**, 201-209.

[188] Renz, H.; Schlimme, E. "Sythesen von 2,4-Substituierten Homo-C-pyrimidin-nucleosiden" *Liebigs Ann. Chem.* **1986**, 957-966.

[189] Cupps, T. L.; Wise, Jr., D. S.; Townsend, L. B. "A Novel Three-Step Synthesis of a Pyttolo[3,2-*d*]pyrimidine C-Nucleoside" *J. Org. Chem.* **1986**, *51*, 1058-1064.

[190] Clark, J.; Ramsden, T. "Heterocyclic Studies. Part XVII. Some 2,3-Dihydroimidazo-[1,2,*a*]- and[1,2,*c*]-pyrimidines" *J. Chem. Soc. (C)* **1971**, 679-683.

[191] Albert, A.; Brown, D. J.; Wood, H. C. S. "Pteridine Studies. Part V. The monosubstituted Pteridines" *J. Chem. Soc.* **1954**, 3832-3839.

[192] Cupps, T. L.; Wise, D. S.; Townsend, L. B. "Synthetic Strategies Toward the Syntehsis of 2,4-Dimethoxypyrrolo[3,2-*d*] pyrimidine" *J. Org. Chem.* **1983**, *48*, 1060-1064 and references therin.

[193] Wichai, U.; Woski, S. A.; "Disiloxane-Protected 2-Deoxyribonolactone as an Efficient Precursor to 1,2-Dideoxy-1-β-aryl-D-ribofuranoses" *Org. Lett.* **1999**, *1*, 1173-1175.

[194] Jung, M. E.; Lyster, M. A.; Quantitative Dealkylation of AlkylEthers Via Treatment with Trimethylsilyl Iodide. A New Method for Ether Hydrolysis" *J. Org. Chem.* **1977**, *42*, 3761-3764.

[195] Robins, M. J.; Uznanski, B. "Nucleic Acid Related Compounds. 33. Conversion of Adenosine and Guanosine to 2,6-Dichloro, 2-Amino-6-chloro, and Derived Purine Nucleosides" *Can. J. Chem.* **1981**, *59*, 2601-2607.

[196] Zemlicka, J.; Sorm, F. "Nucleic Acids Components and Their Analogues. LXI. The Reaction of Dimethylchloromethyleneammonium Chloride with the 2',3',5'-Tri-O-Acyl Derivatives of Uridine and 6- Azauridine; a New Synthesis of 6-Azacytidine" *Collection Czechoslov. Chem. Commun.* **1965**, *30*, 2052-2067.

[197] Bridges, A. J. "Triaryl(oxy)phosphine Dibromide. A Convenient Reagent for the Preparation of *S*-Arylthioinosines and N^6, 5'-Disubstituted Adenosine Derivatives from Inosine" *Nucelosides Nucleotides* **1988**, *7*, 375-383.

[198] Hamamoto, S.; Takaku, H. "New Approach to the Synthesis of Deoxyribonucleoside Phosphoramidite Derivatives" *Chemistry Letters* **1986**, 1401-1404.

[199] Chen, F. M. F.; Benoiton, N. L. "A General Method for Formylating Sensitive Amino Acid Esters" *Synthesis* **1979**, 709-710.

[200] Brown, D. J.; Sugimoto, T. "Aza-analogues of Pteridine. Part II. The Novel Use of Silver Oxide in Transetherification of Alkoxy-1,2,4,6,8-penta-azanapthalenes, Alkoxy-nitropyrimidines, and Related Systems" *J. Chem. Soc. C* **1970**, 2661-2666.

[201] Olah, G. A; Husain, A.; Gupta, B. G. B.; Narang, S. C. "Trichloro(methyl)silane/Sodium Iodide, A New Regioselective Reagent for the Cleavage of Ethers" *Angew. Chem. Int. Ed. Engl.* **1981**, *20*, 690-691.

[202] Brown, D. J. "Pyrimidine Reactions. Part II. 2-Ethoxy-4-, and 4-Ethoxy-2-hydroxy-5-nitropyrimidine" *J. Chem. Soc.* **1959**, 3647-3648.

[203] Krotz, A. H.; Klopchin, P. G.; Walker, K. L.; Srivatsa, G. S.; Cole, D. L.; Ravikumar, V. T. "On the Formation of Longmers in Phosphorthioate Oligodeoxyribonucleotide Synthesis" *Tetrahedron Lett.* **1997**, *38*, 3875-3878.

[204] Vargeese, C.; Carter, J.; Yegge, J.; Krivjansky, S.; Settle, A.; Kropp, E.; Peterson, K.; Pieken, W. "Efficient Activation of Nucleoside Phosphoramidites with 4,5-

Dicyanoimidazole During Oligonucleotide Synthesis" *Nucleic Acids Res.* **1998**, *26*, 1046-1050.

[205] Zhu, Q.; Delaney, M. O.; Greenberg, M. M.; "Observation and Elimination of *N*-Acetylation of Oligonucleotides Prepared Using Fast-Deprotecting Phosphoramidites and Ultra-Mild Deprotection" *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1105-1107.

[206] Ordoukhanian, P.; Taylor, J.-S. "Solid Phase Supported Thymine Dimers fort the Construction of Dimer-Containing DNA by Combined Chemical and Enzymatic Synthesis: A Potentially General Method for the Efficient Incorporation of Modified Nucleotides into DNA" *Nucleic Acids Res.* **1997**, *25*, 3785-3786.

[207] Smith, C. A.; Taylor, J.-S. "Preparation and Characterization of a Set of Deoxyoligonucleotide 49-mers Containing Site-Specific Cis-syn, Trans-syn-I, (6,4), and Dewar Photoproducts of Thymidylyl (3'→5')-thymidine" *J. Biol. Chem.* **1993**, *268*, 11143-11151.

[208] Alekseyev, Y. O.; Romano, L. J. "In Vitro Replication of Primer-Templates Containing Benzo[a]pyrene Adducts by Exonuclease-Deficient *Escherichia coli* DNA Polymerase I (Klenow Fragment): Effect of Sequence Context on Lesion Bypass" *Biochemistry* **2000**, *39*, 10431-10438.

[209] Lowry, T. H.; Richardson, K. S. *Mechanism and Theory in Organic Chemistry, 3rd Ed.* Harpers Collins Publishers, New York, 1987.

[210] Gelfand, C. A.; Plum, G. E.; Grollmann, A. P.; Johnson, F.; Breslauer, K. J "Thermodynamic Consequences of an Abasic Lesion in Duplex DNA Are Strongly Dependent on Base Sequence" *Biochemistry* **1998**, *37*, 7321-7327.

[211] Goljier, I.; Kumar, S.; Bolton, P. H.; "Refined Solution Structure of a DNA Heteroduplex Containing an Aldehydic Abaisic Site" *J. Biol. Chem.* **1995**, *270*, 22980-22987.

[212] Lin, Z.; Hung, K.-N.; Grollmann, A. P.; de los Santos, C. "Solution Structure of Duplex DNA Containing an Extrahelical Abasic Site Analog Determined by NMR Spectroscopy and Molecular Dynamics" *Nucleic Acids Res.* **1998**, *26*, 2385-2391.

[213] Coppel, Y.; Berthet, N.; Coulombeau, C.; Colombeau, C.; Garcia, J.; Lhomme, J. "Solution Conformation of an Abasic DNA Undecamer Duplex d(CGCACXCACGC) d(GCGTGTGTGCG): The Unpaired Thymine Stacks Inside the Helix" *Biochemistry* **1997**, *36*, 4817-4830.

[214] Derbyshire, V.; Freemont, P. S.; Sanderson, M. R.; Beese, L.; Friedman, J. M.; Joyce, C. M.; Steitz, T. A. "Genetic and Crystallographic Studies of the 3', 5'-Exonucleolytic Site of DNA Polymerase I" *Science* **1988**, *240*, 199-201.

[215] Randall, S. K.; Eritja, Kaplan, B. E.; Petruska, J.; Goodman, M. F.; "Nucleotide Insertion Kinetics Opposite Abasic Lesions in DNA" *J. Biol. Chem.* **1987**, *262*, 6864-6870.

[216] Wiederholt, C. J.; Delaney, M. O.; Greenberg, M. M. "Interaction of DNA Containing Fapy•dA or its C-Nucleoside Analogues with Base Excision Repair Enzymes. Implications for Mutagenesis and Enzyme Inhibition." *Biochemistry* (submitted).

[217] Weiderholt, C. J. **2002**, (unpublished results).

[218] Bulychev, N. V.; Varaprasad, C. V.; Dorman, G.; Miller, J. H.; Eisenberg, M.; Grollman, A. P.; Johnson, F. "Substrate Specificity of *Escherichia coli* MutY Protein" *Biochemistry* **1996**, *35*, 13147-13156.

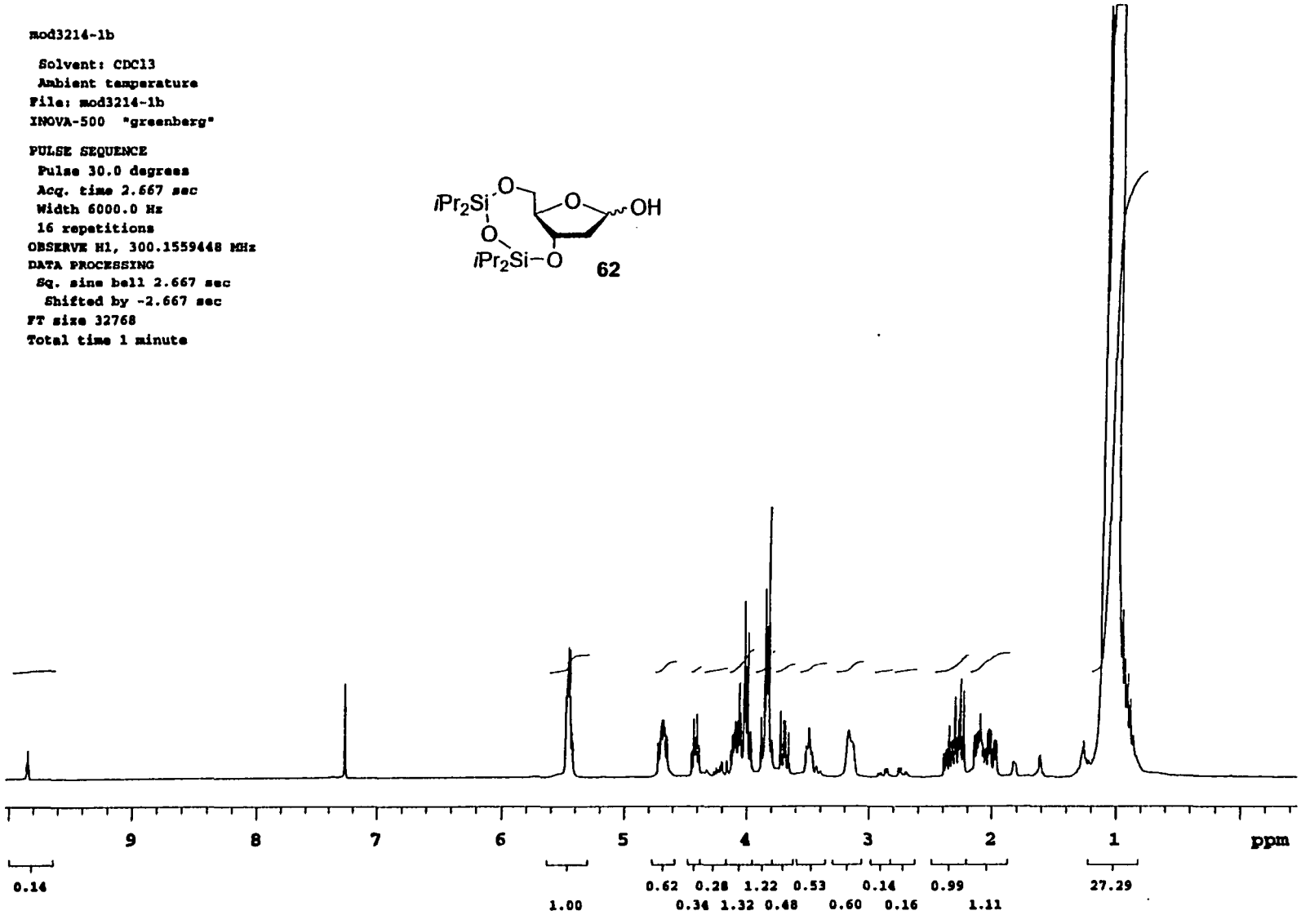
[219] Perrin, D. D.; Armarego, W. L. F. *Purification of Laboratory Chemicals.* Pergamon Press, New York, **1988**.

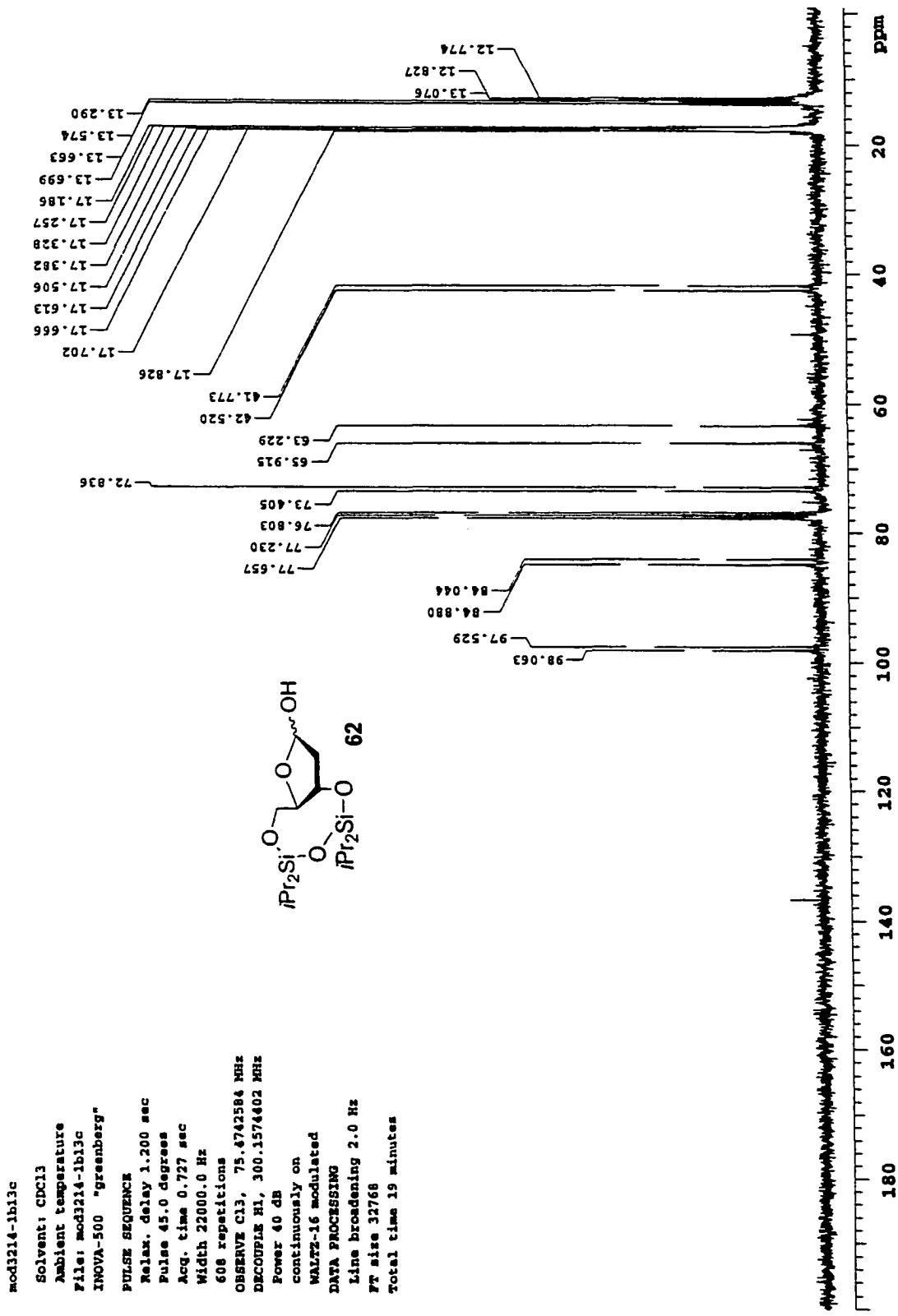
[220] Stults, J. T.; Marsters, J. C. "Improved Electrospray Ionization of Synthetic Oligodeoxynucleotides" *Rapid Commun. Mass Spectrom.* **1991**, *5*, 359-363.

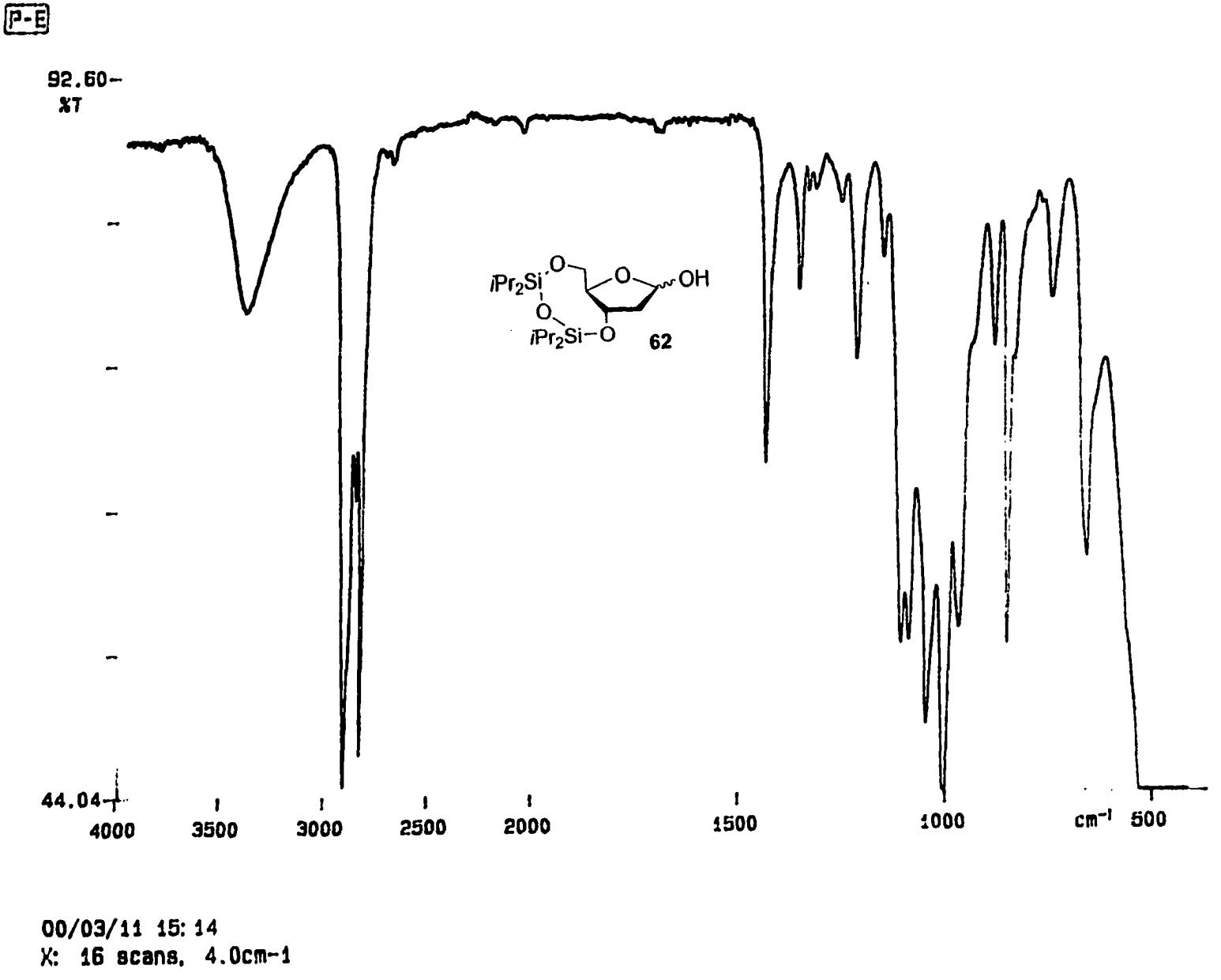
[221] Maniatis, T.; Fritsch, E. F.; Sambrook, J.; **1982** *Molecular Cloning : A Laboratory Manual;* Cold Springs Harbor Laboratory. Cold Springs Harbor, New York.

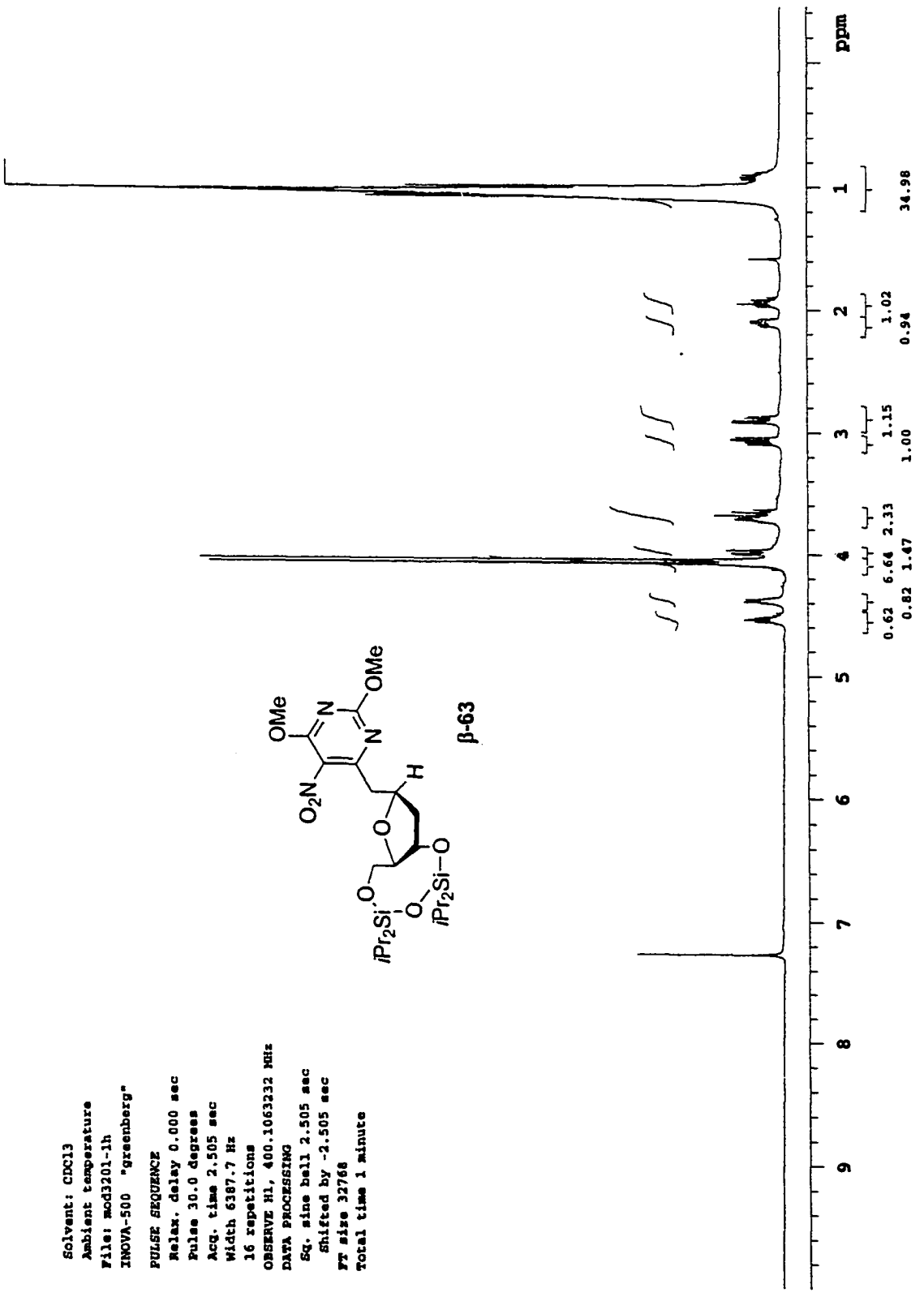
[222] Burer, P. N. In *Handbook of Biochemistry and Molecular Biology, 3rd ed.*; Fasman, G. D., Ed.; CRC Press: Cleveland, OH; Vol. 2, p. 589.

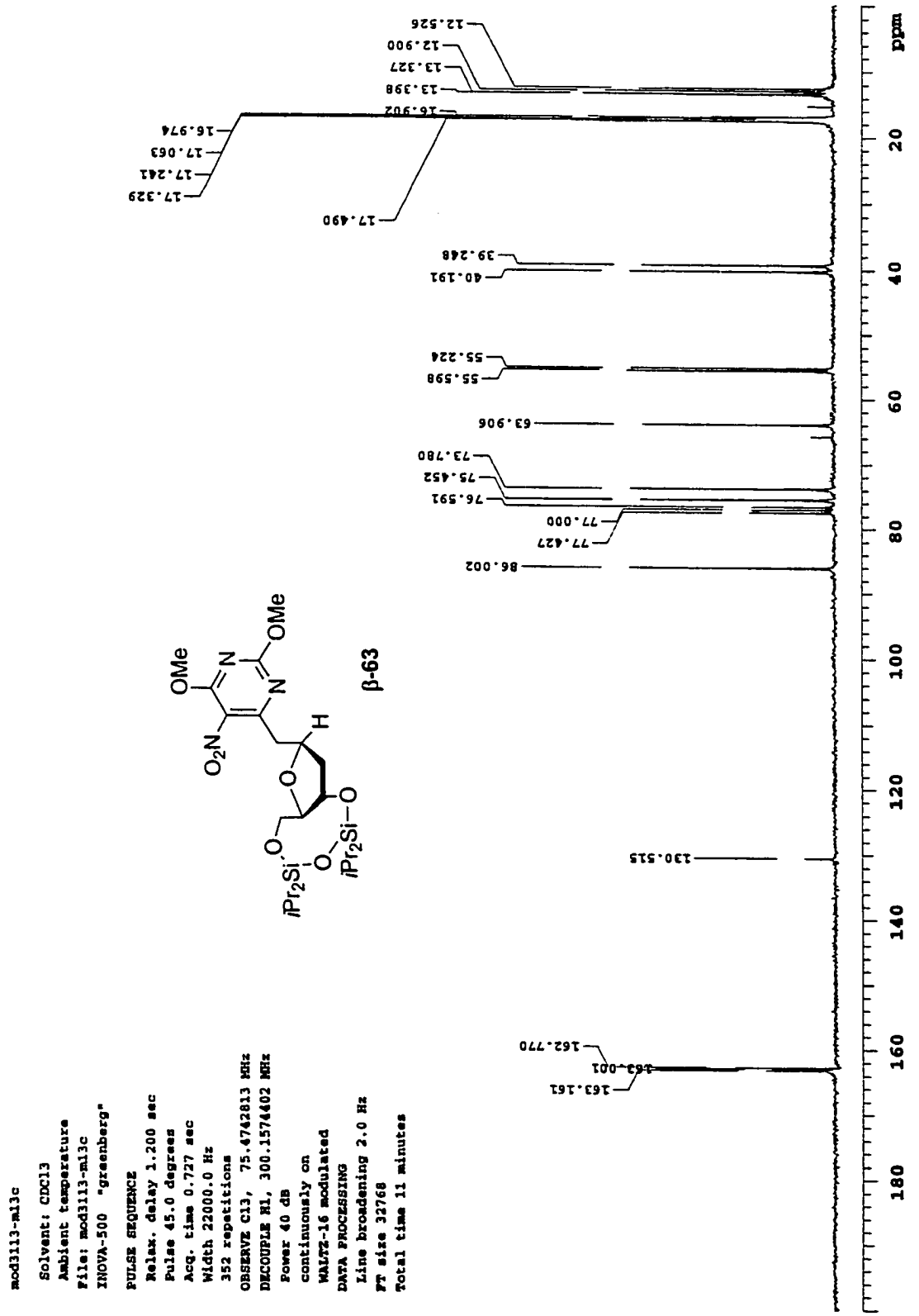
7 Appendix A: NMR, IR, and MS Spectra of Selected Compounds

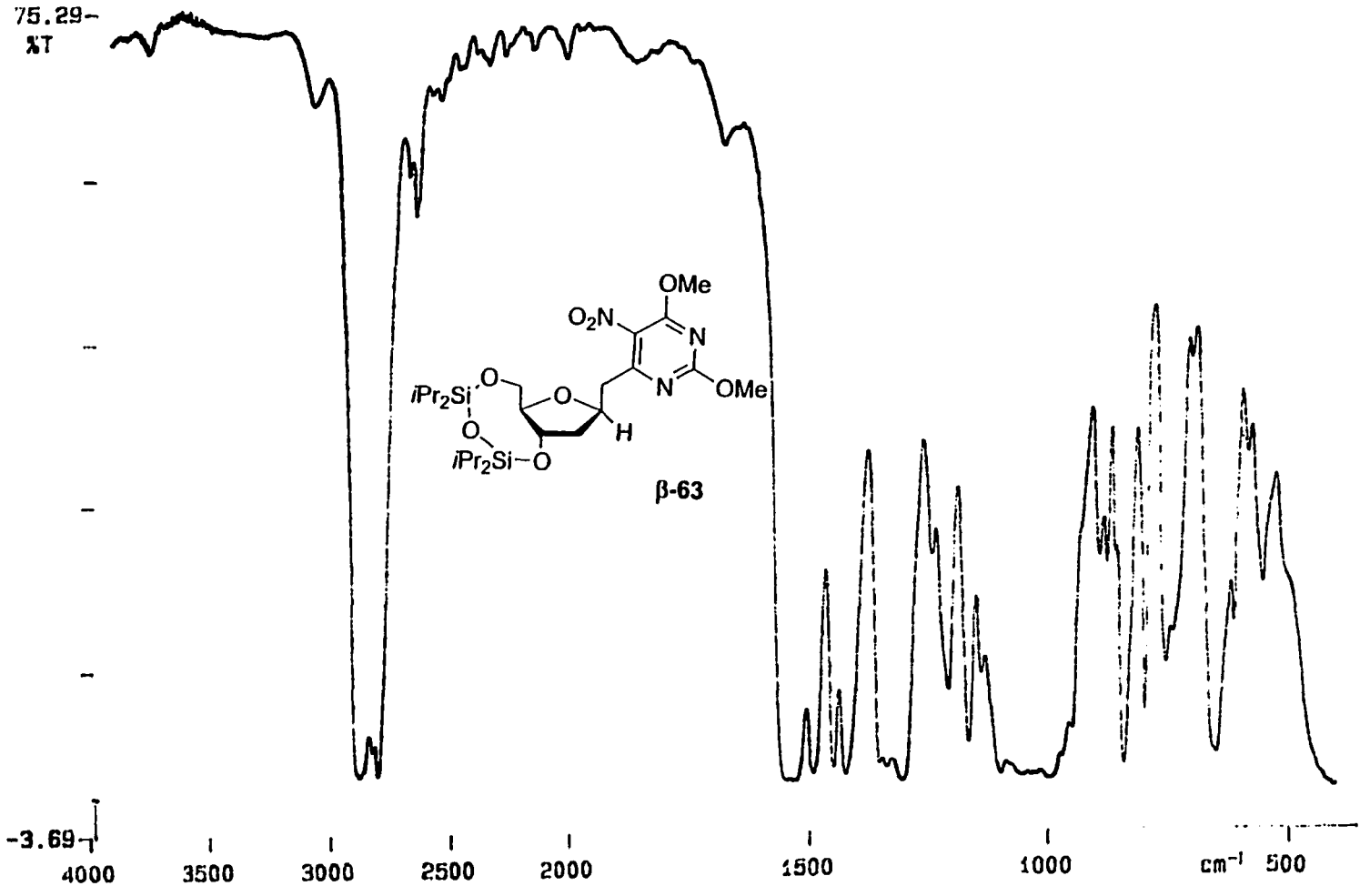




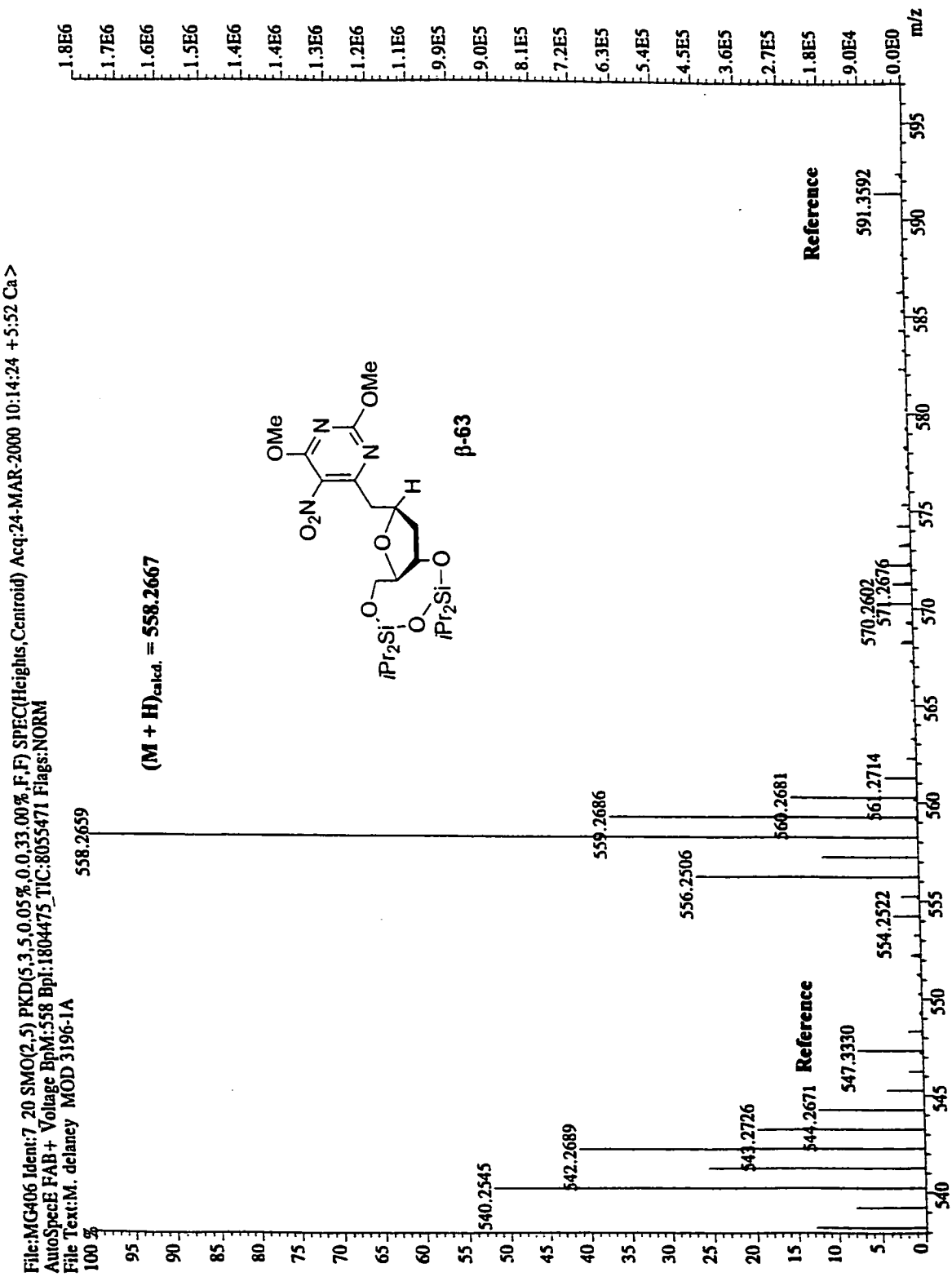


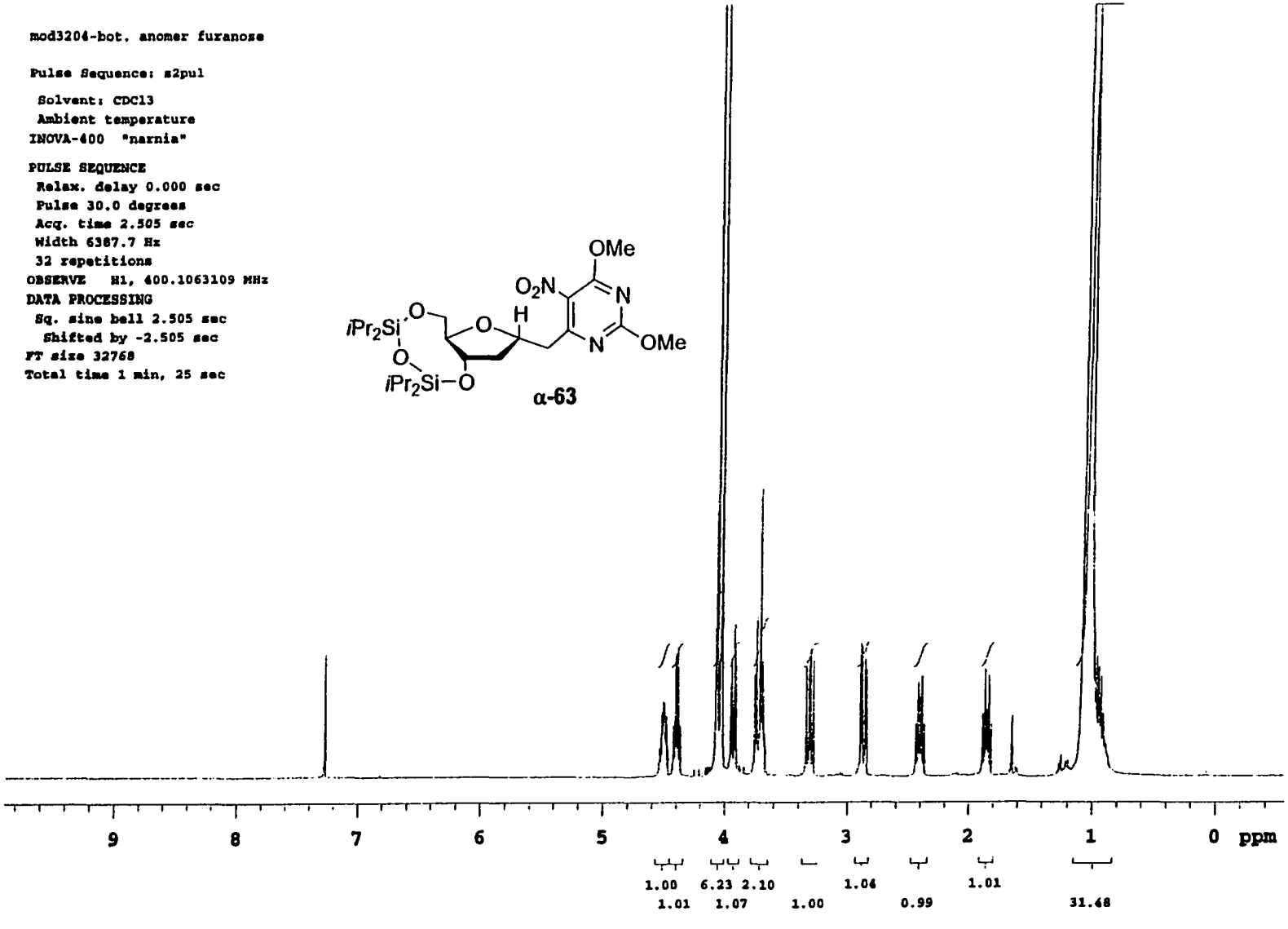


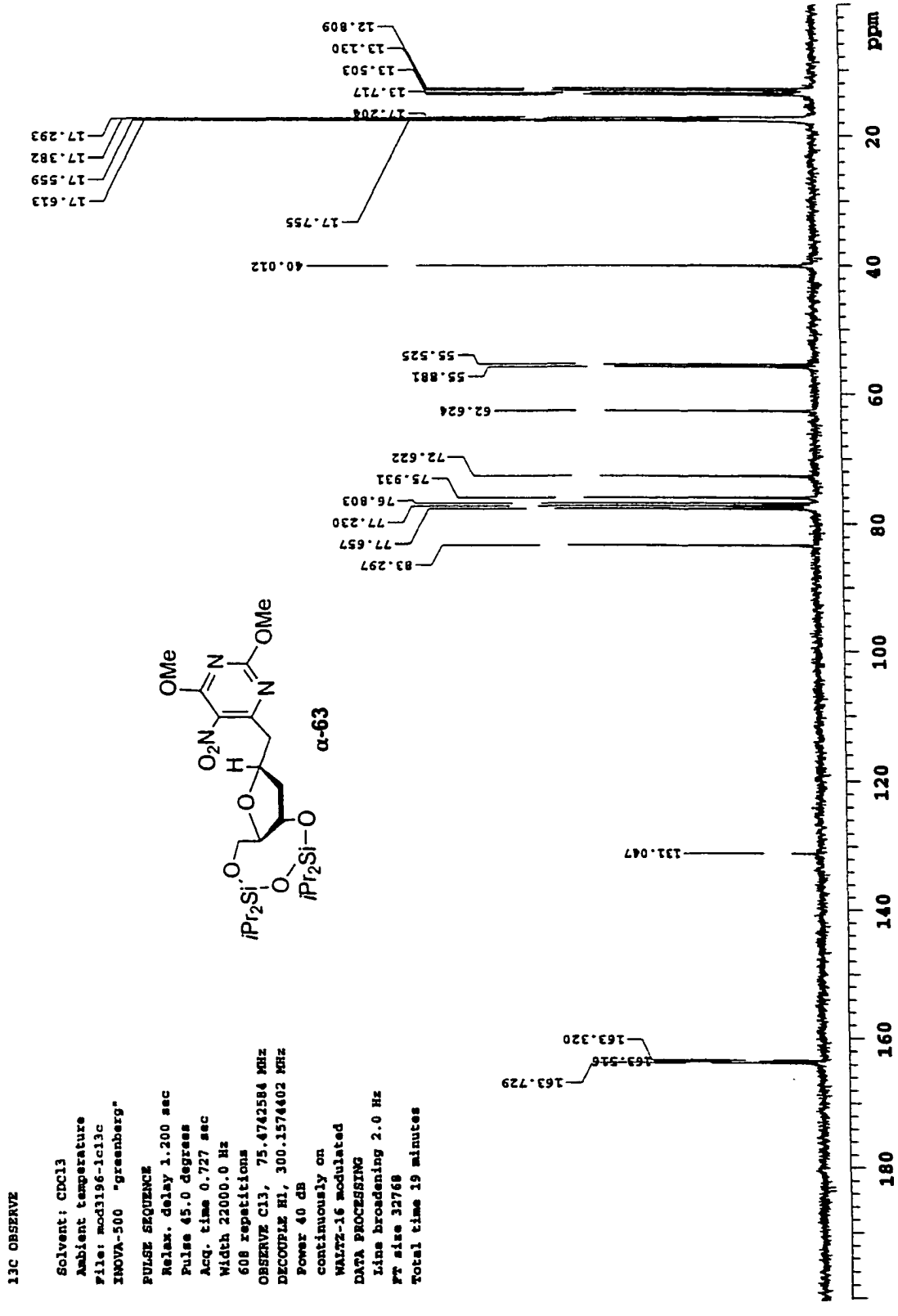


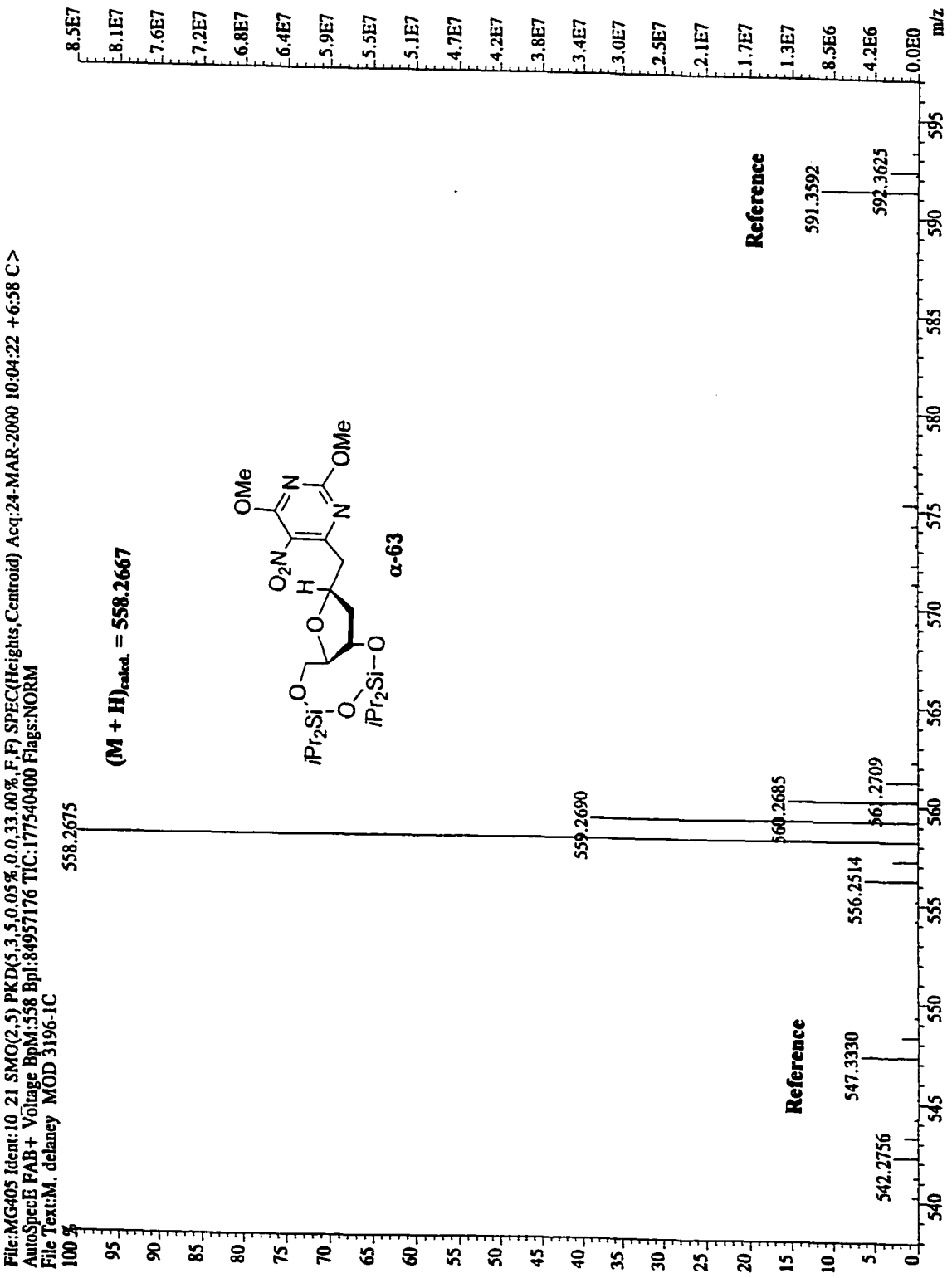


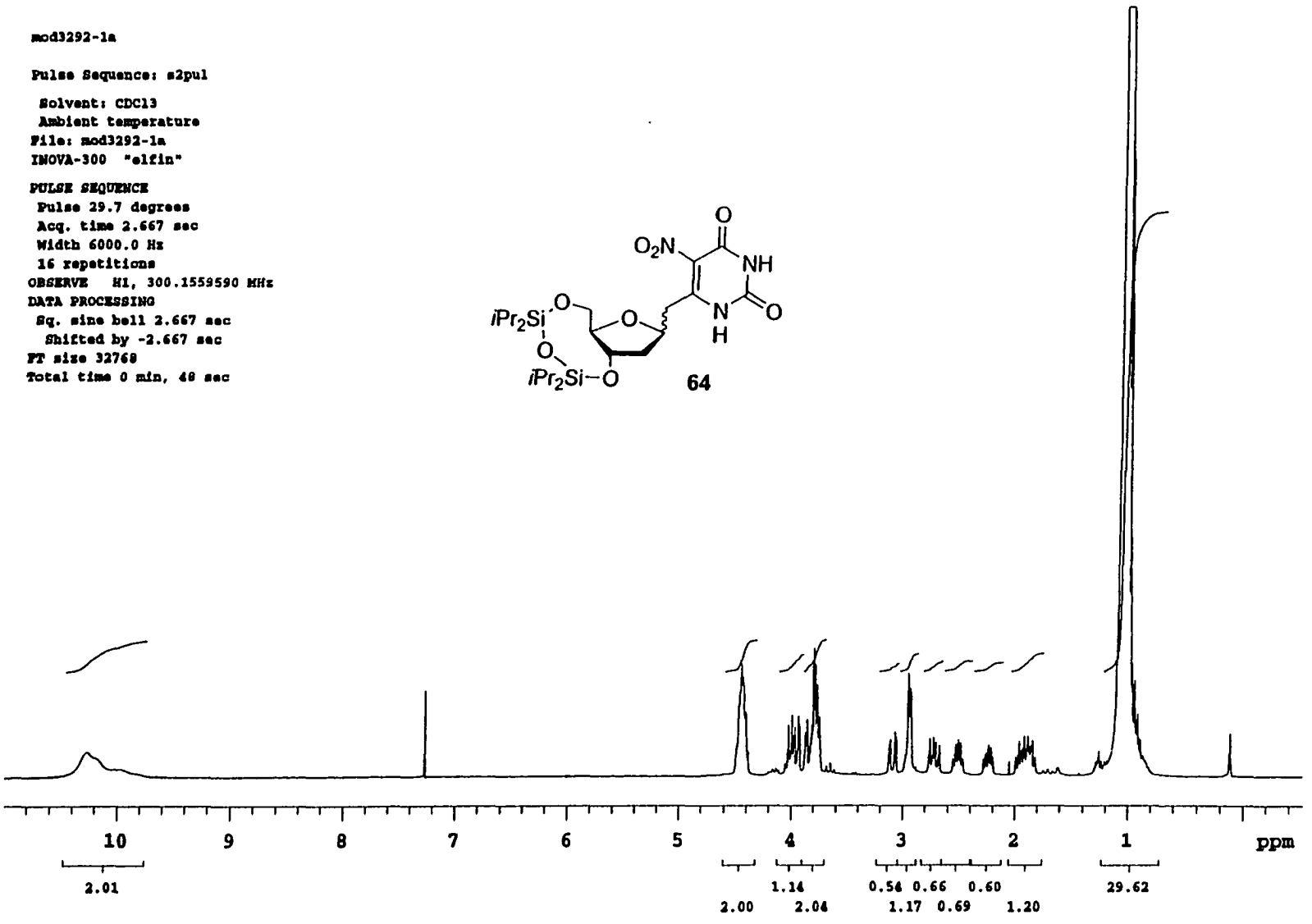
00/03/24 11:06
SCAN: 16 scans, 4.0cm-1

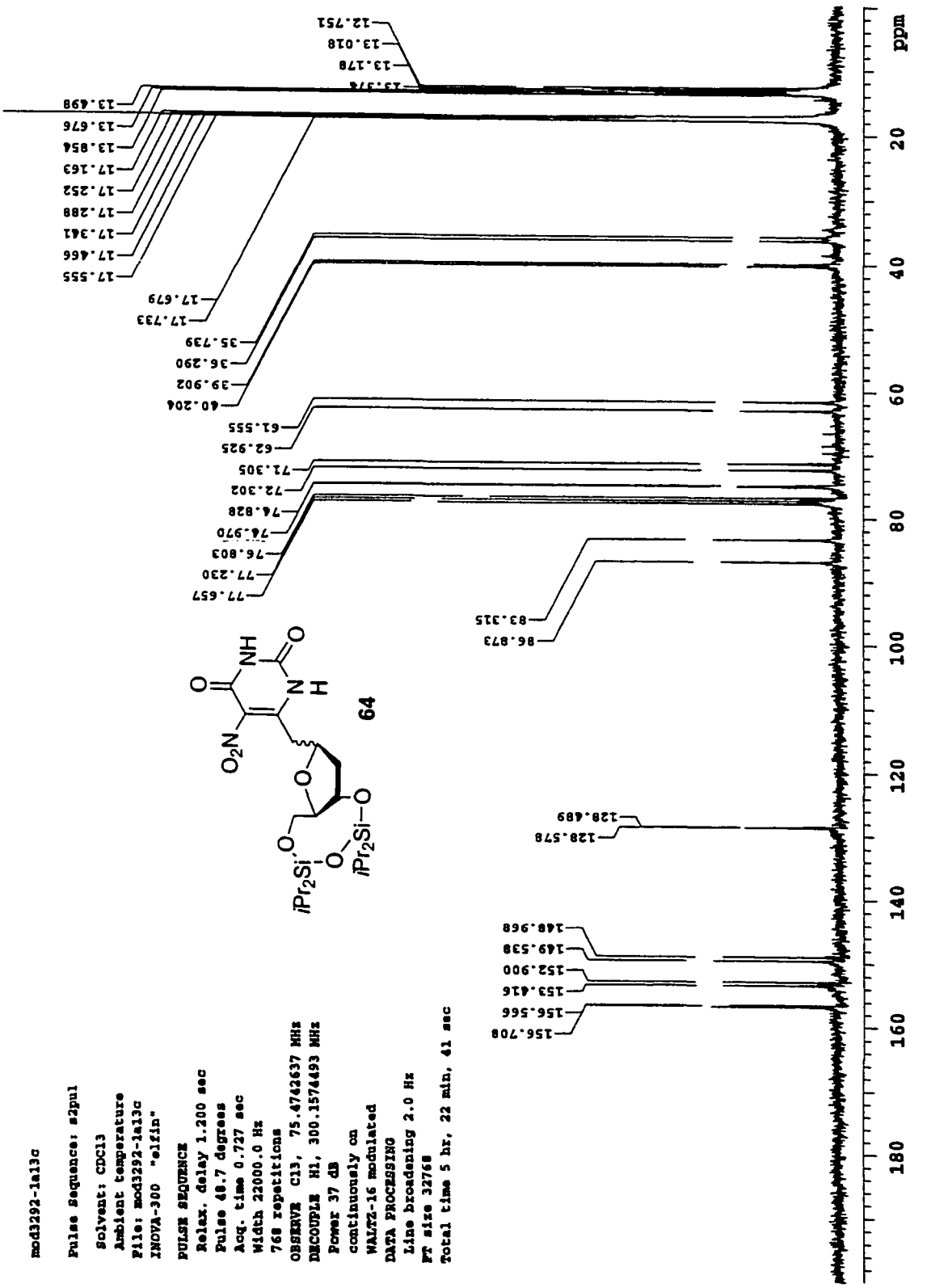


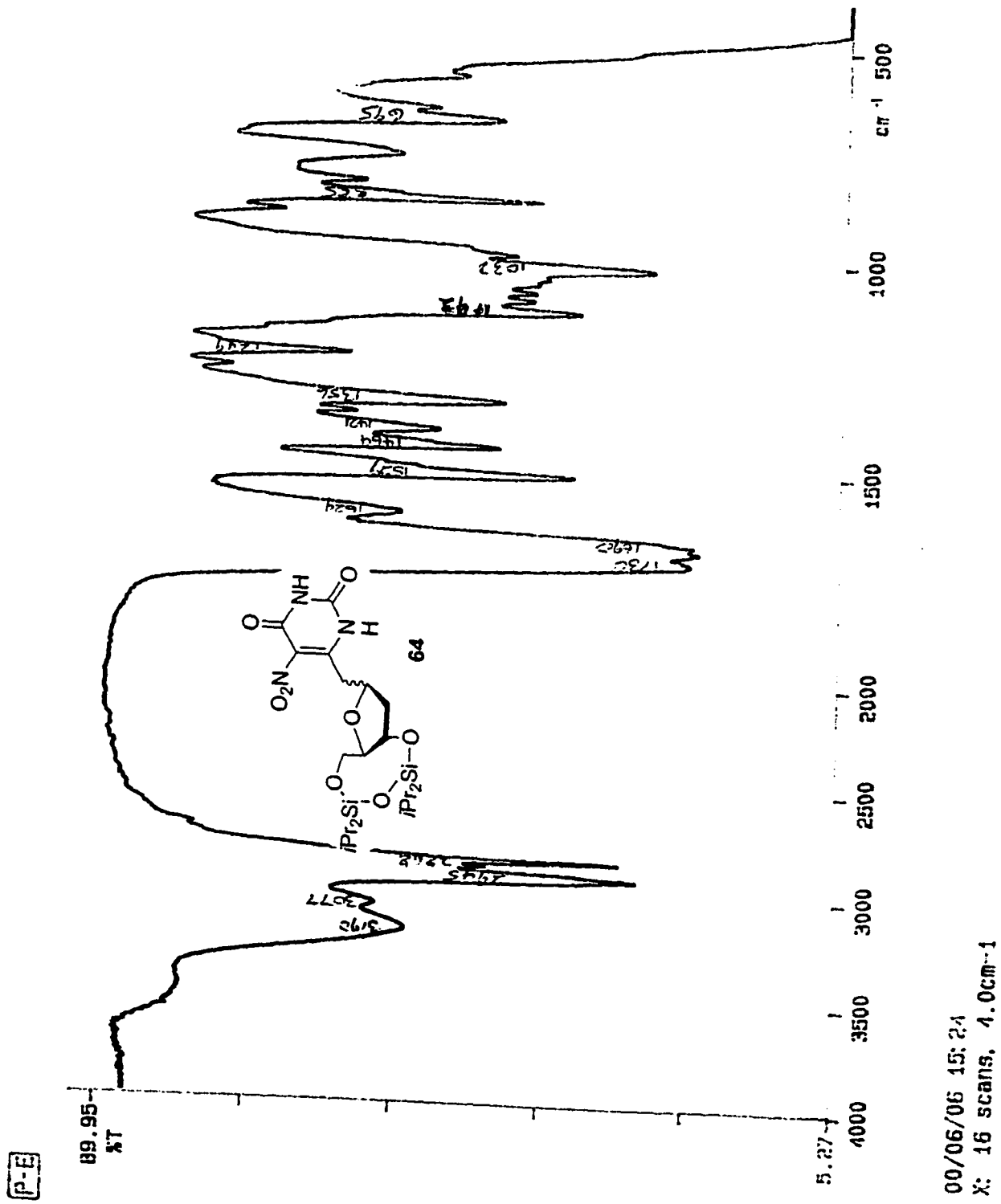


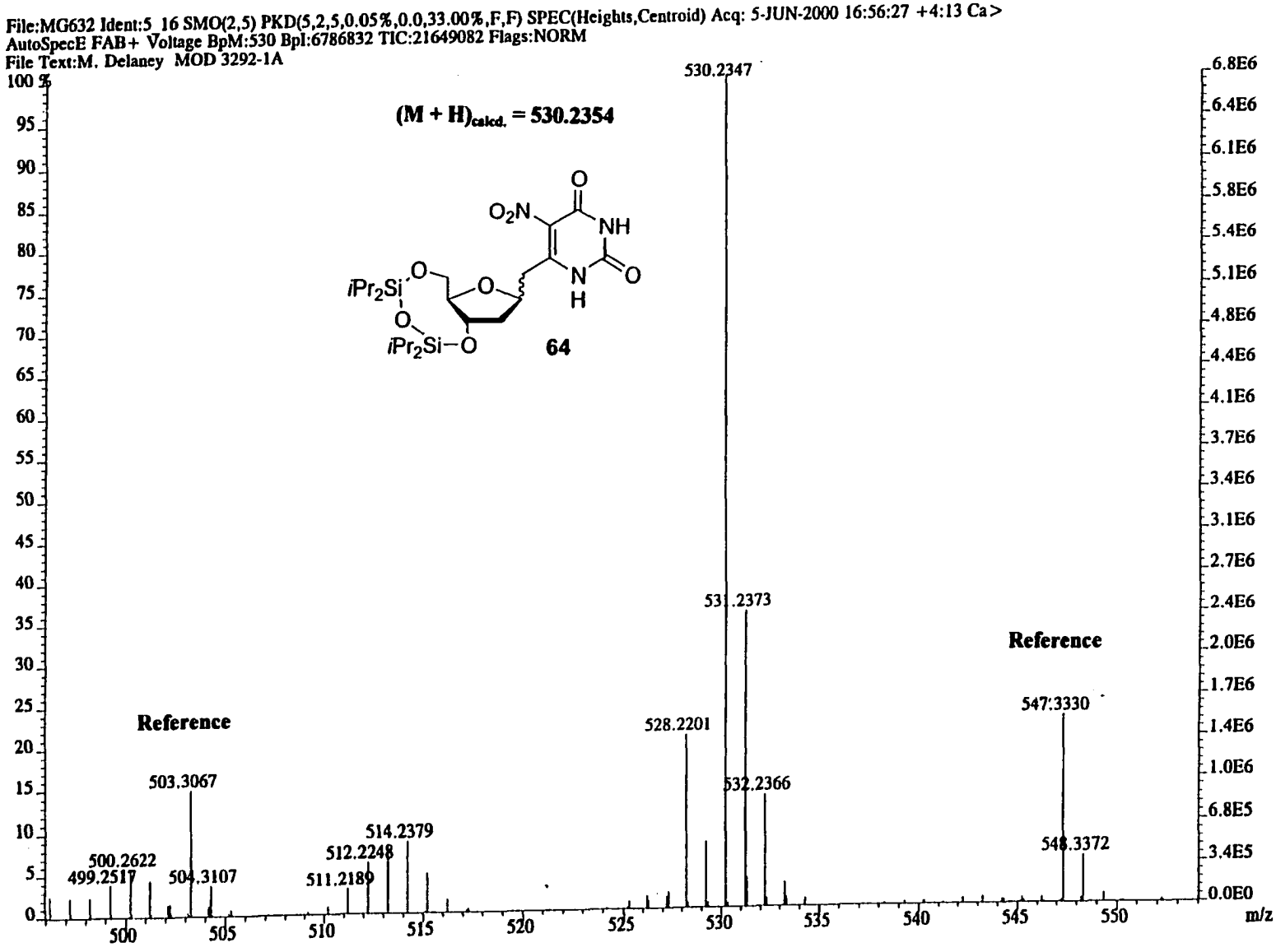


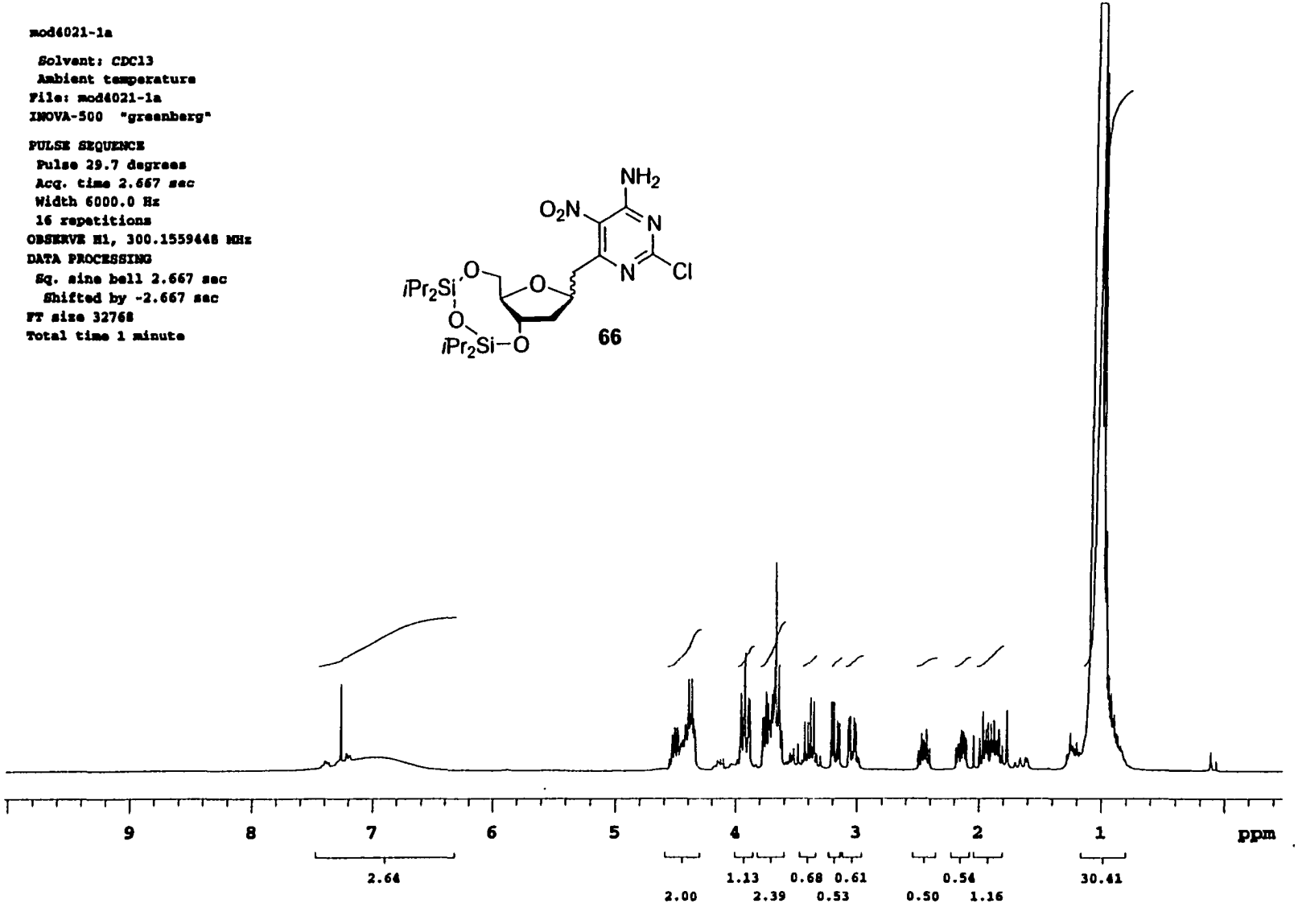


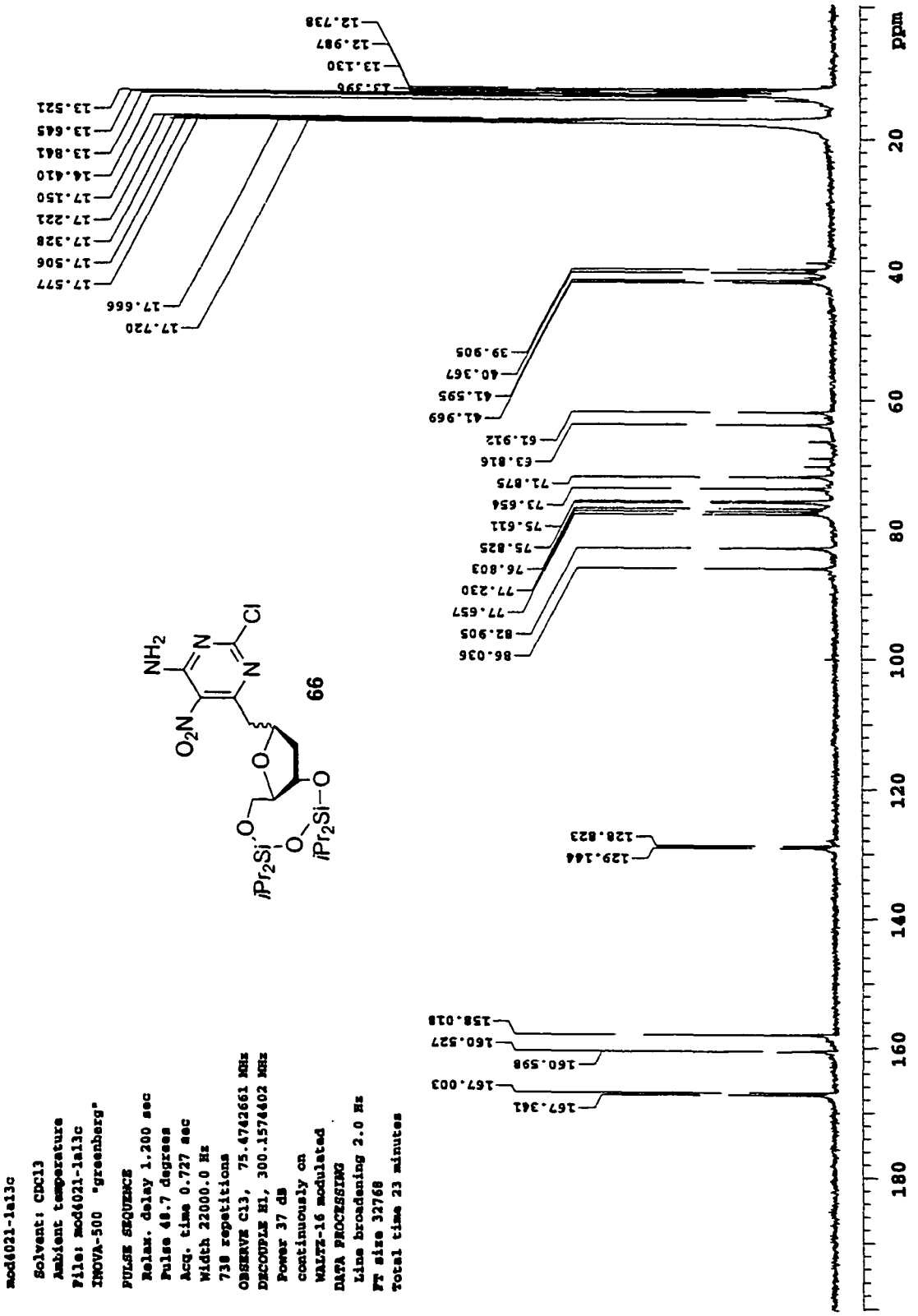


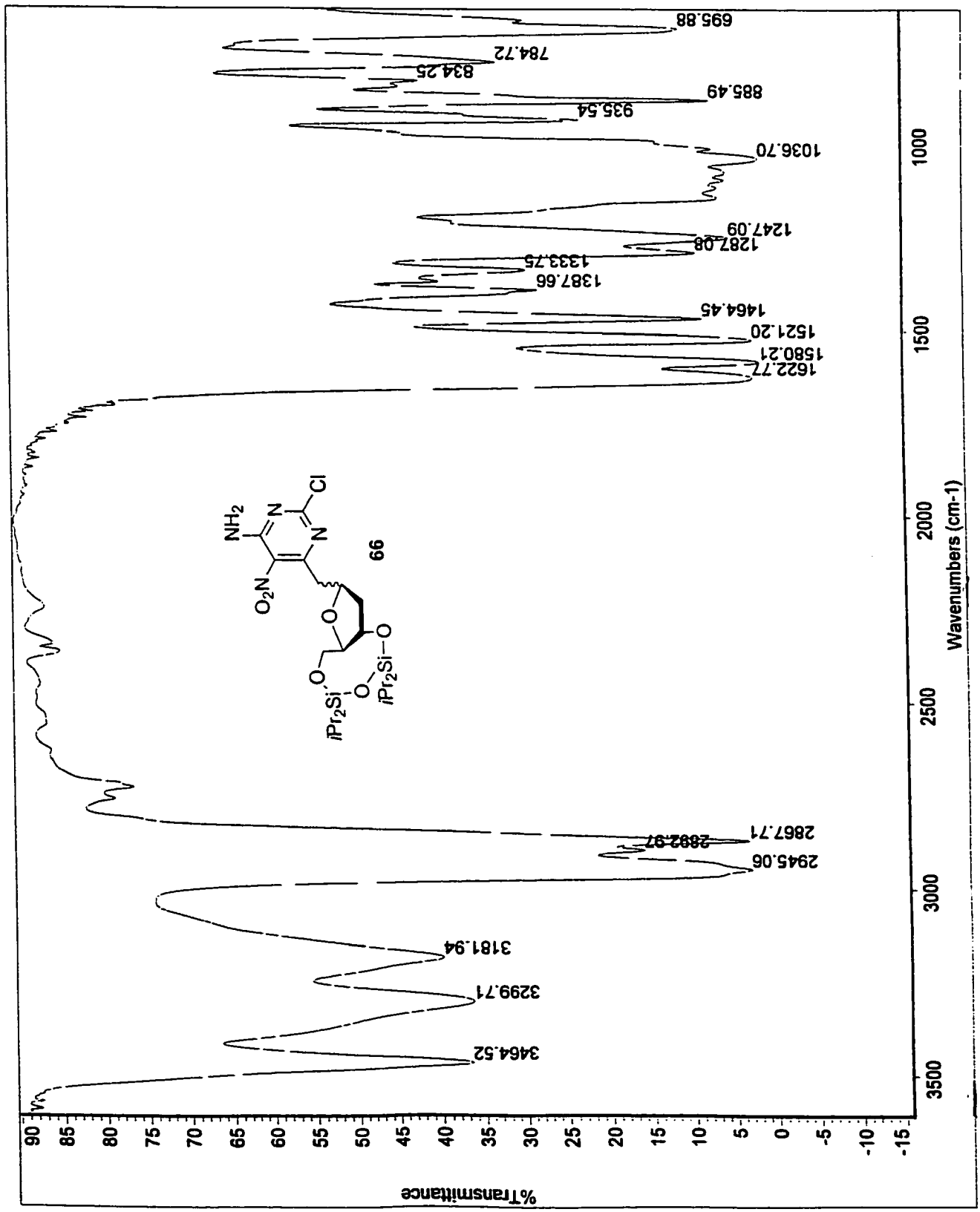


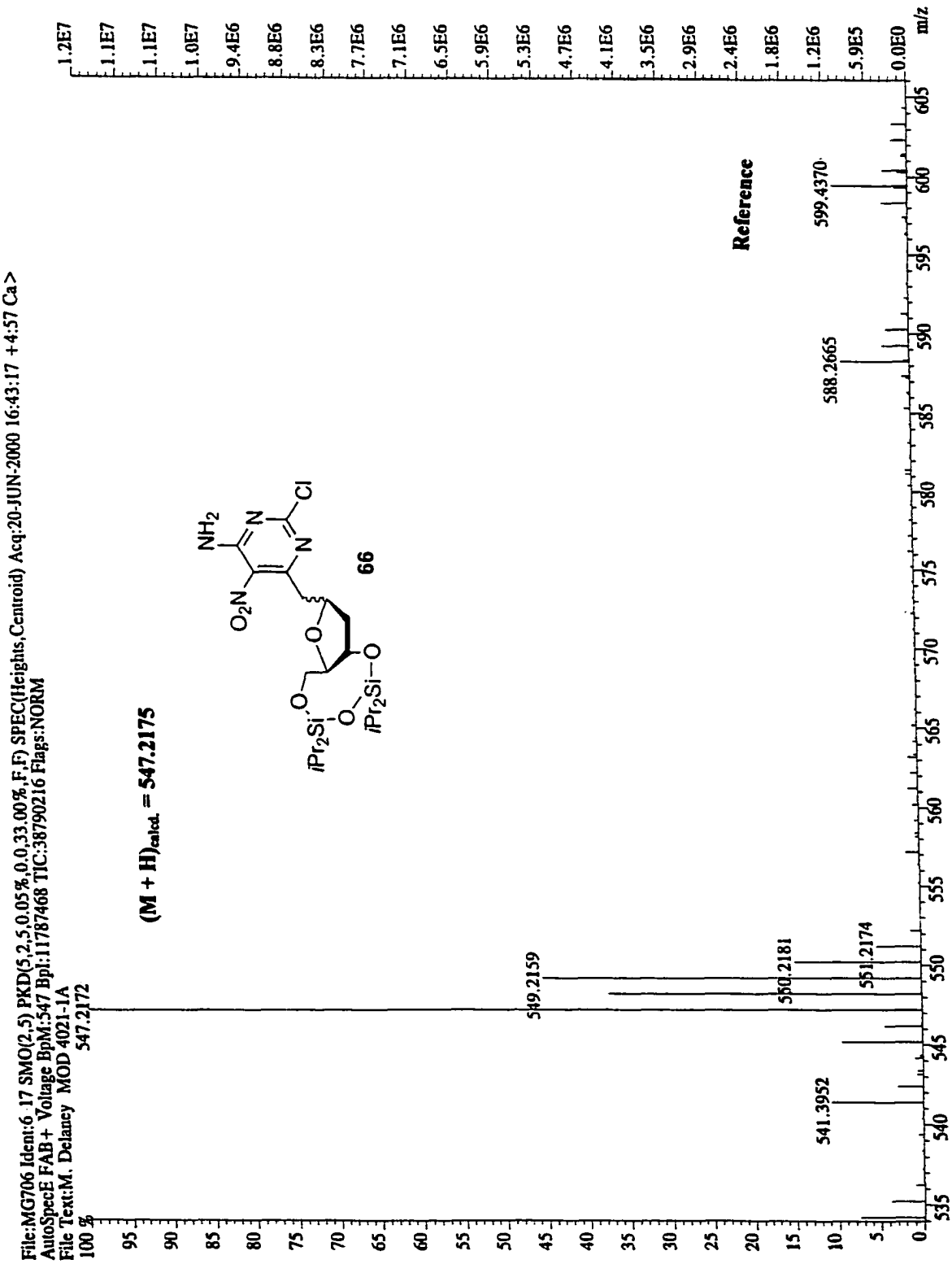


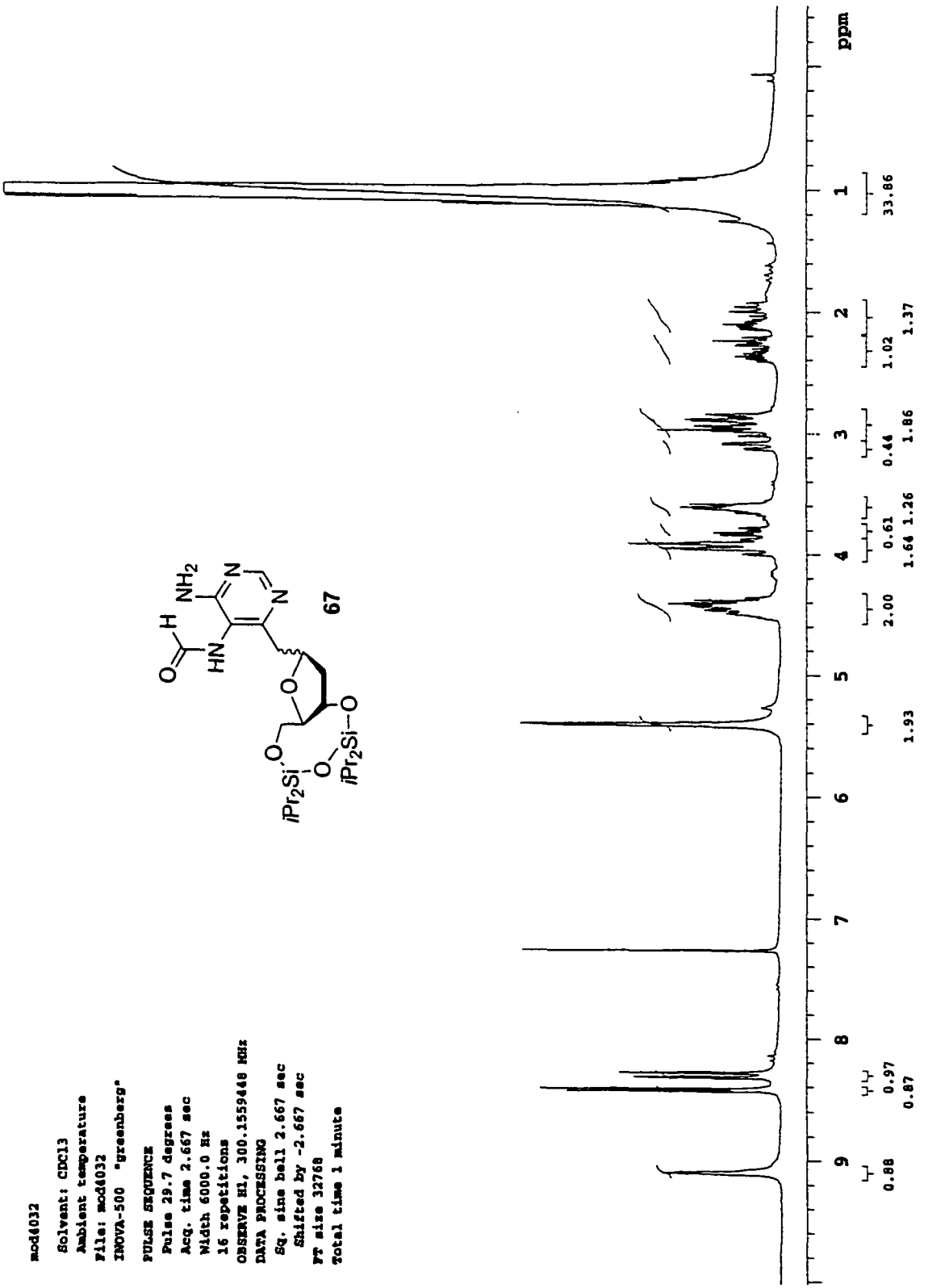


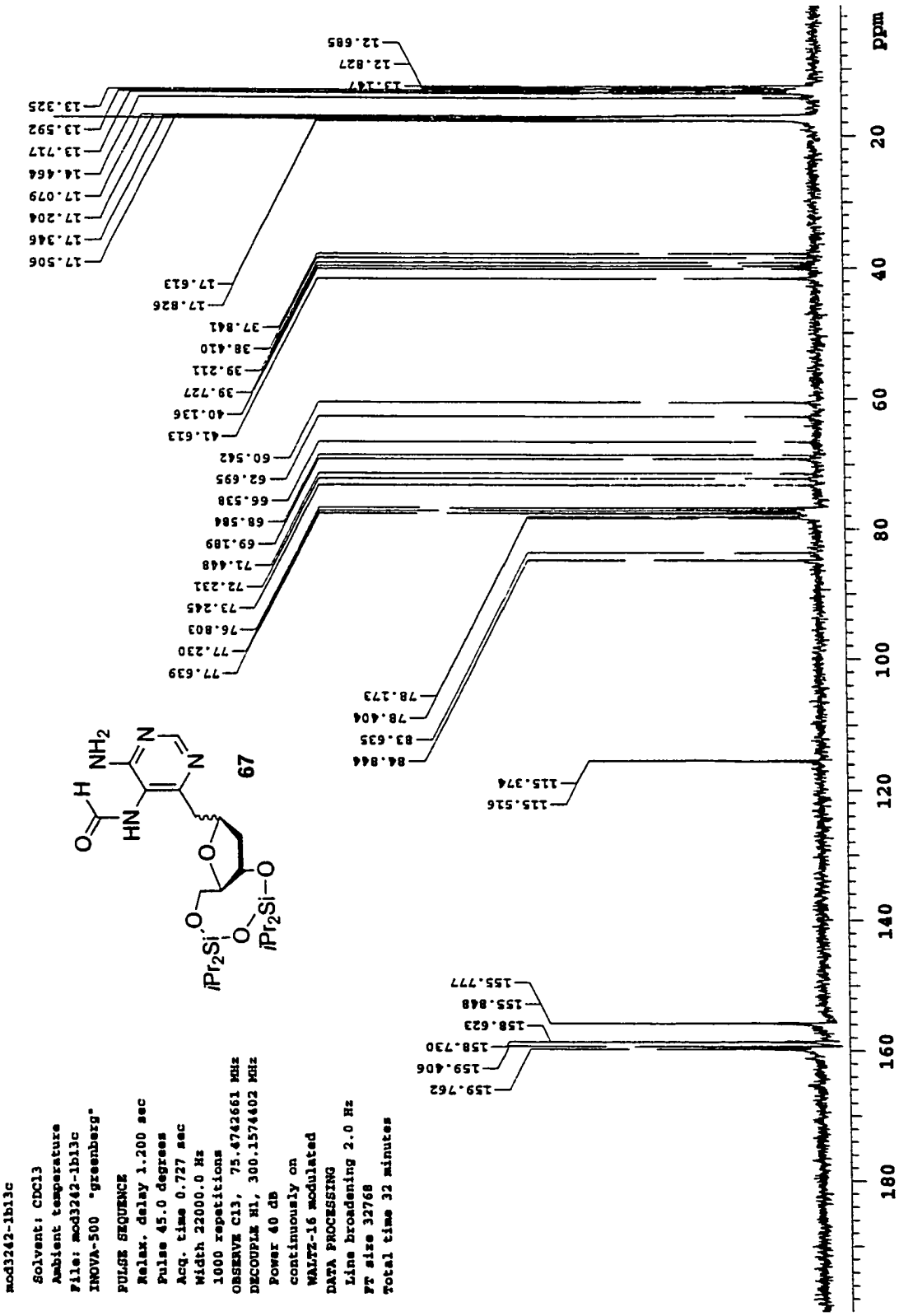


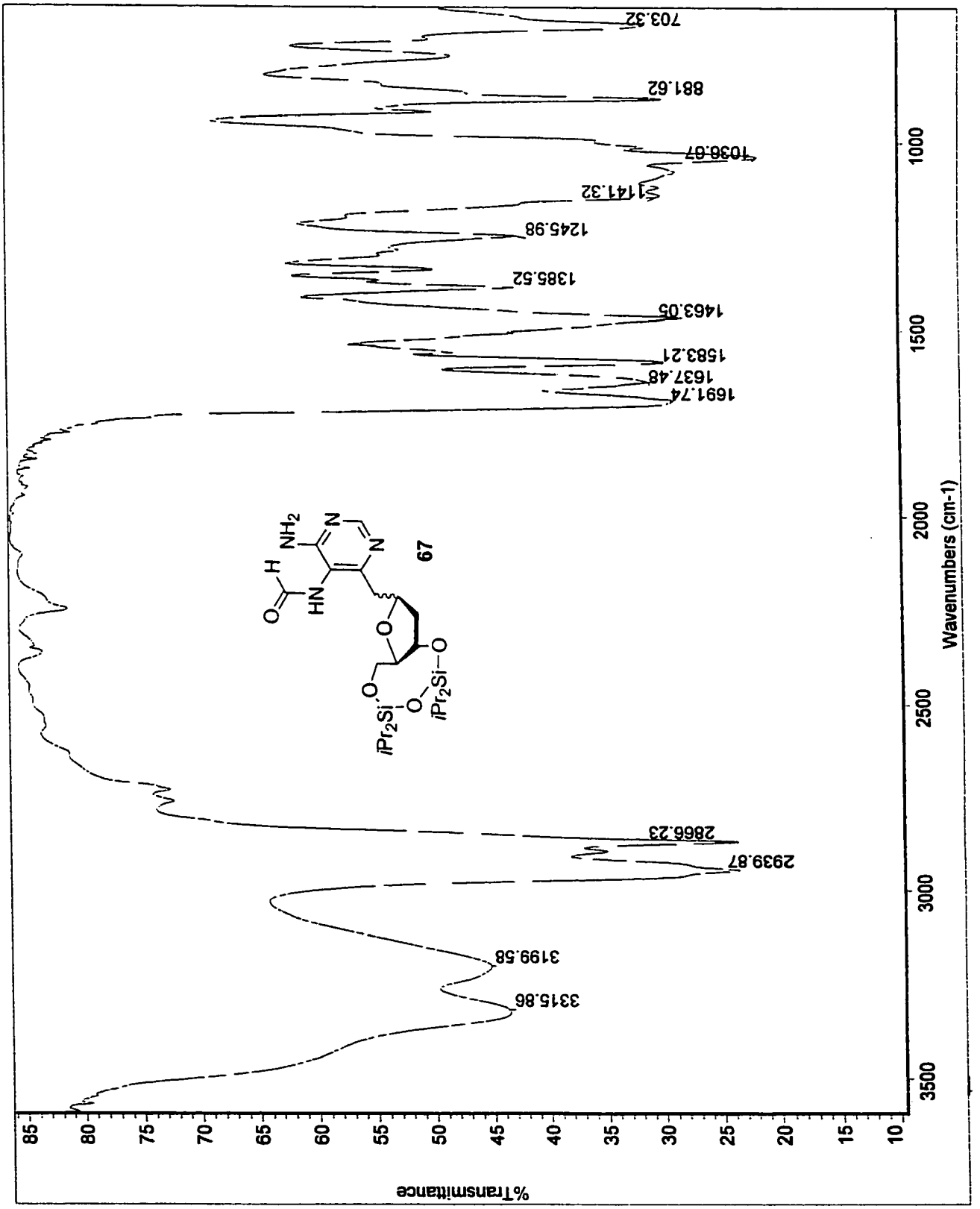


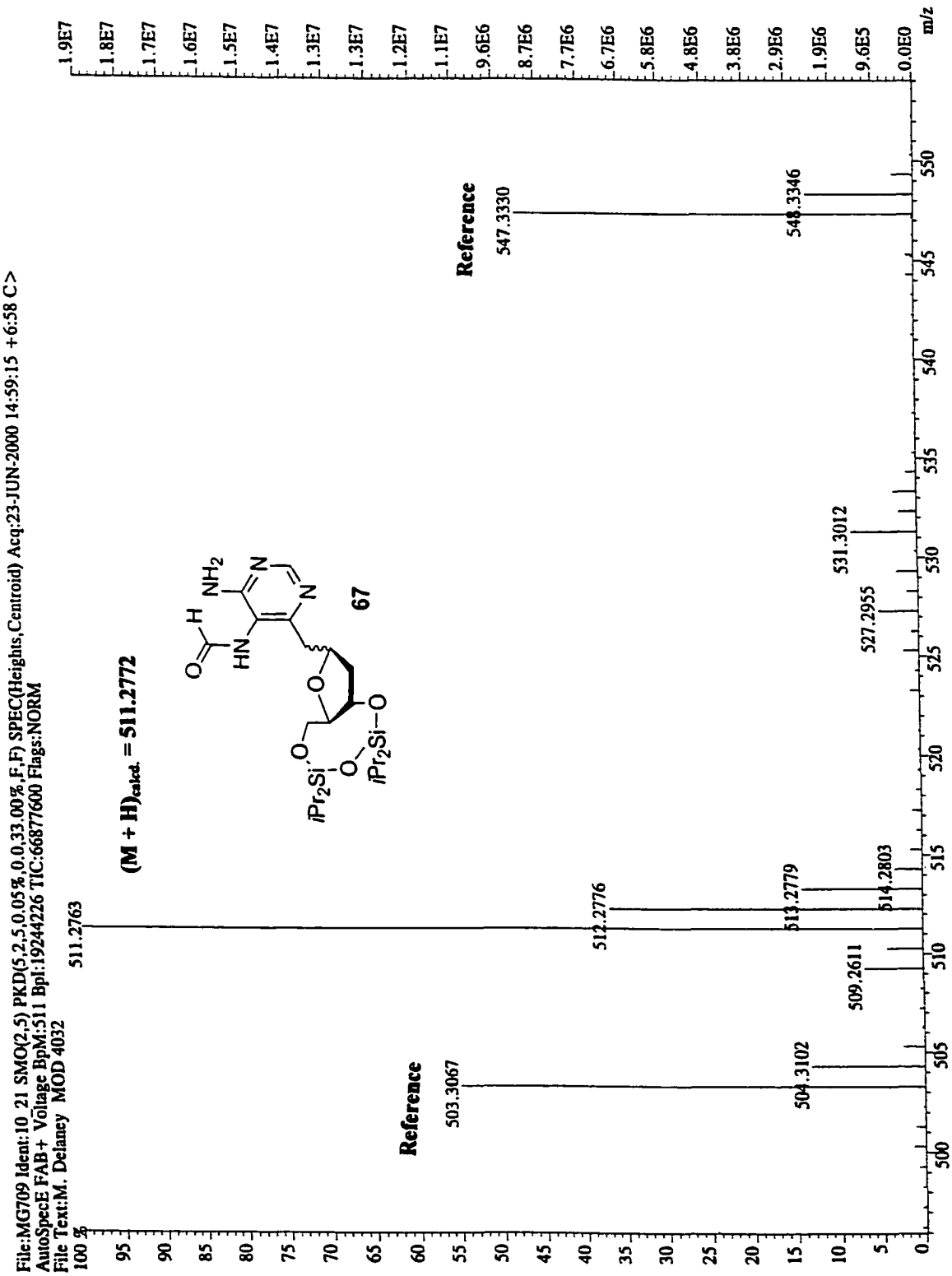


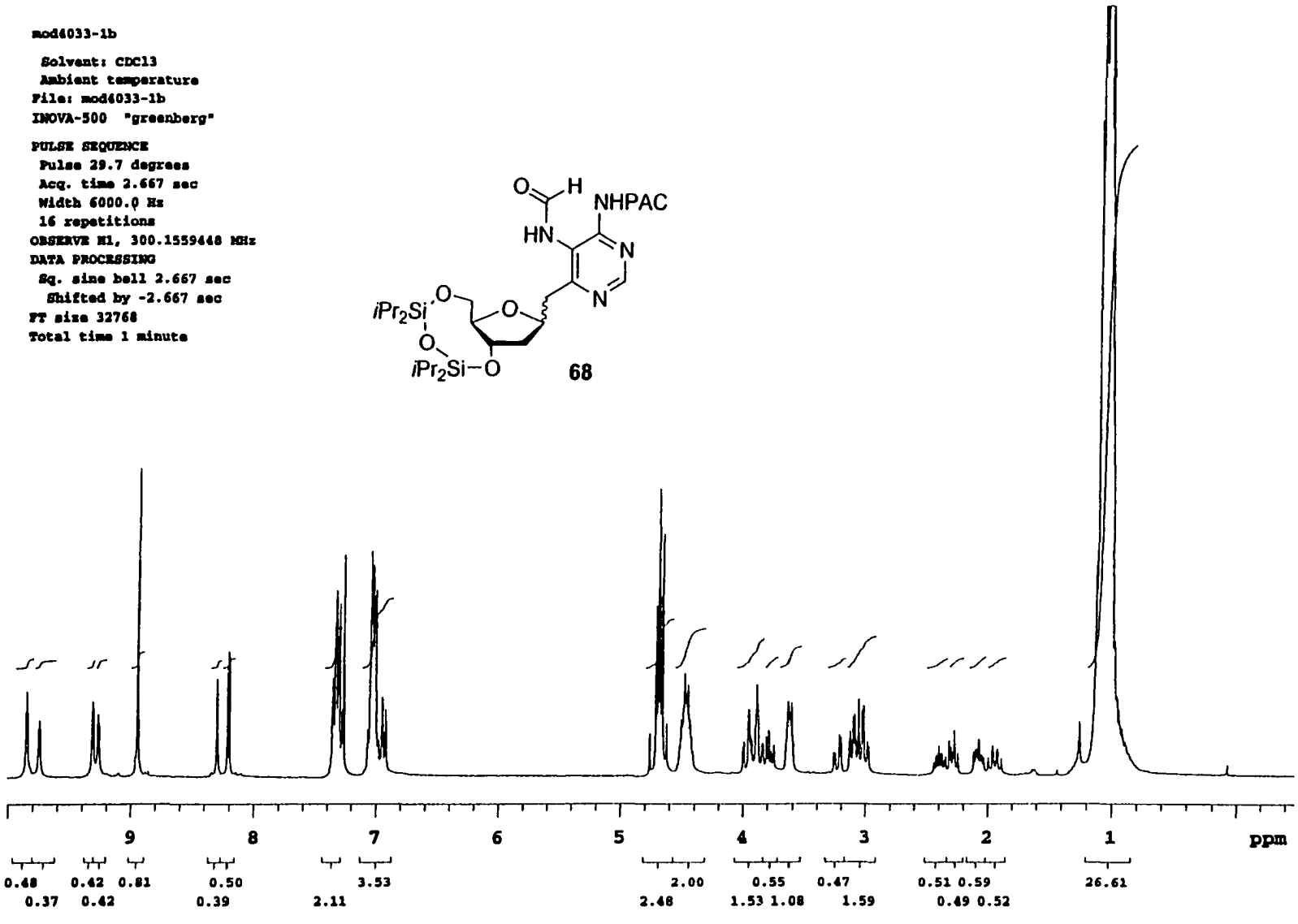


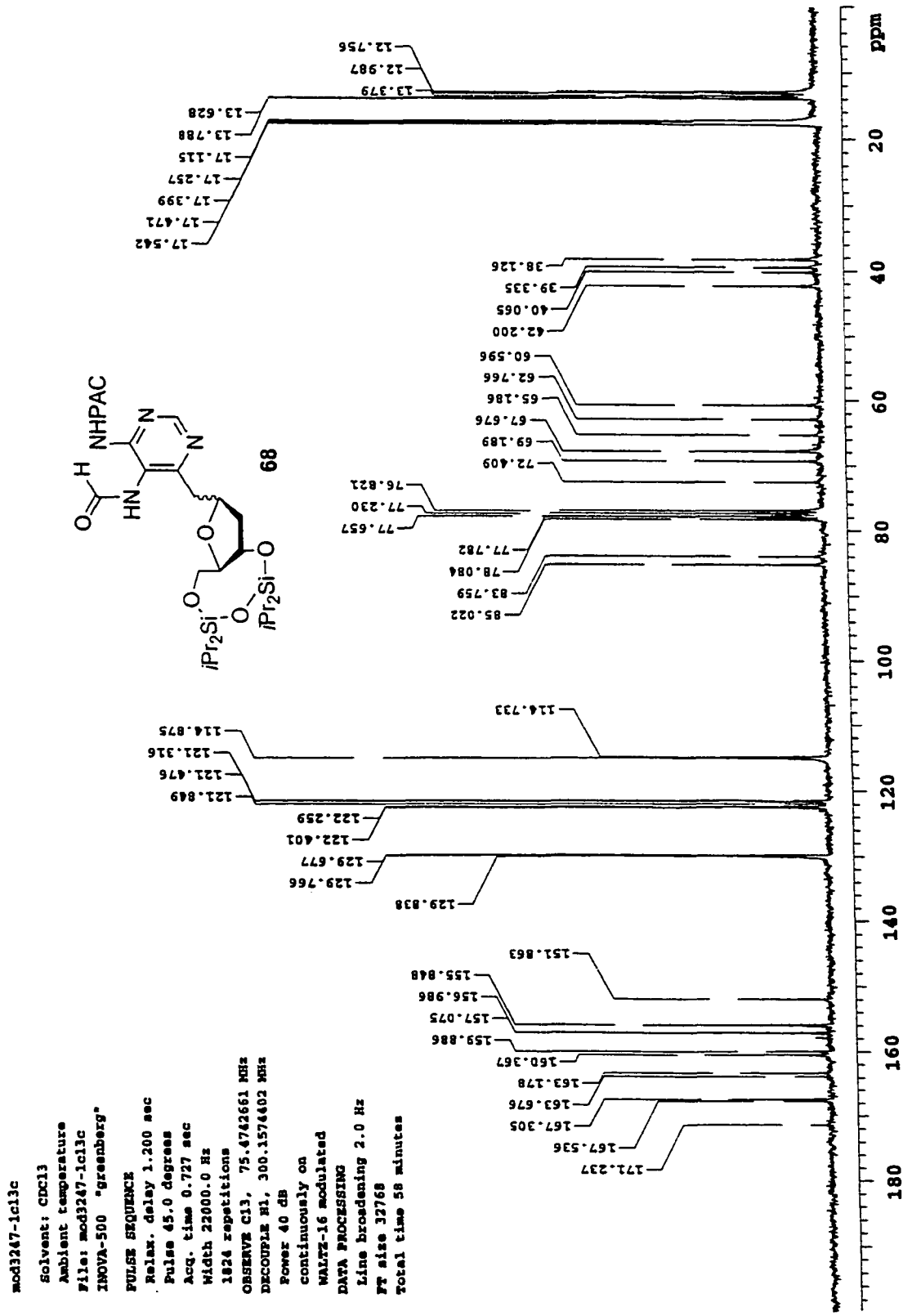


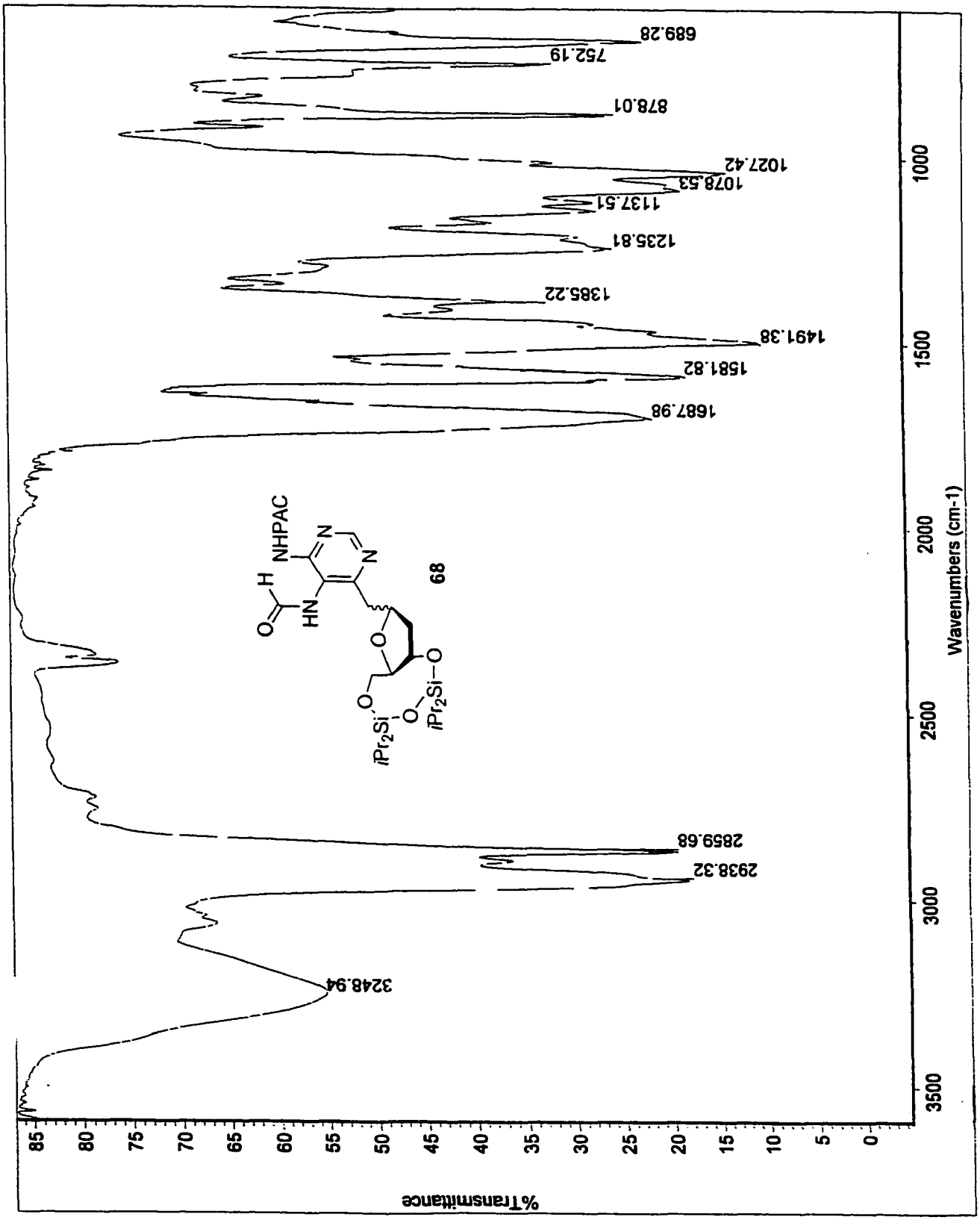


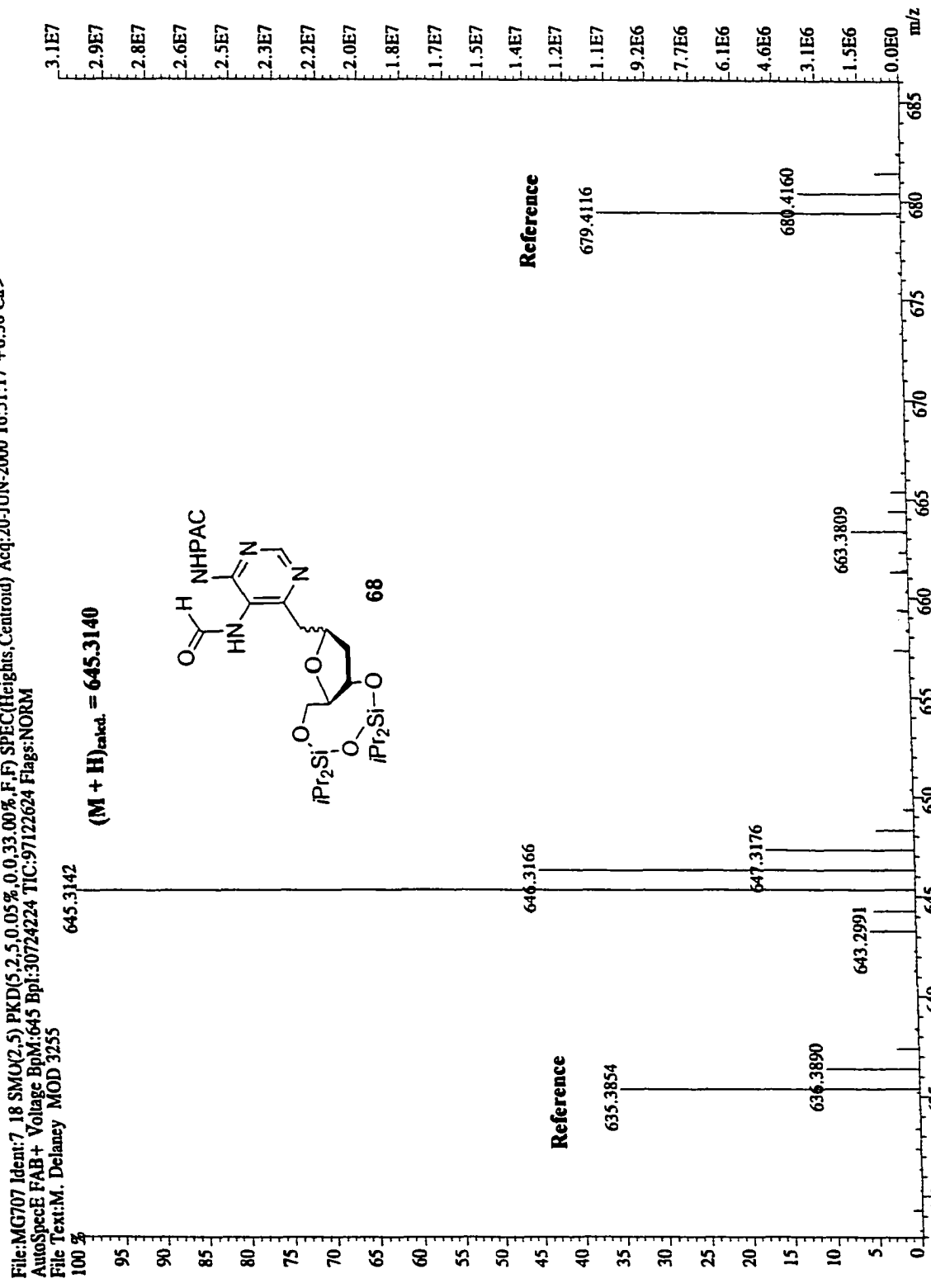


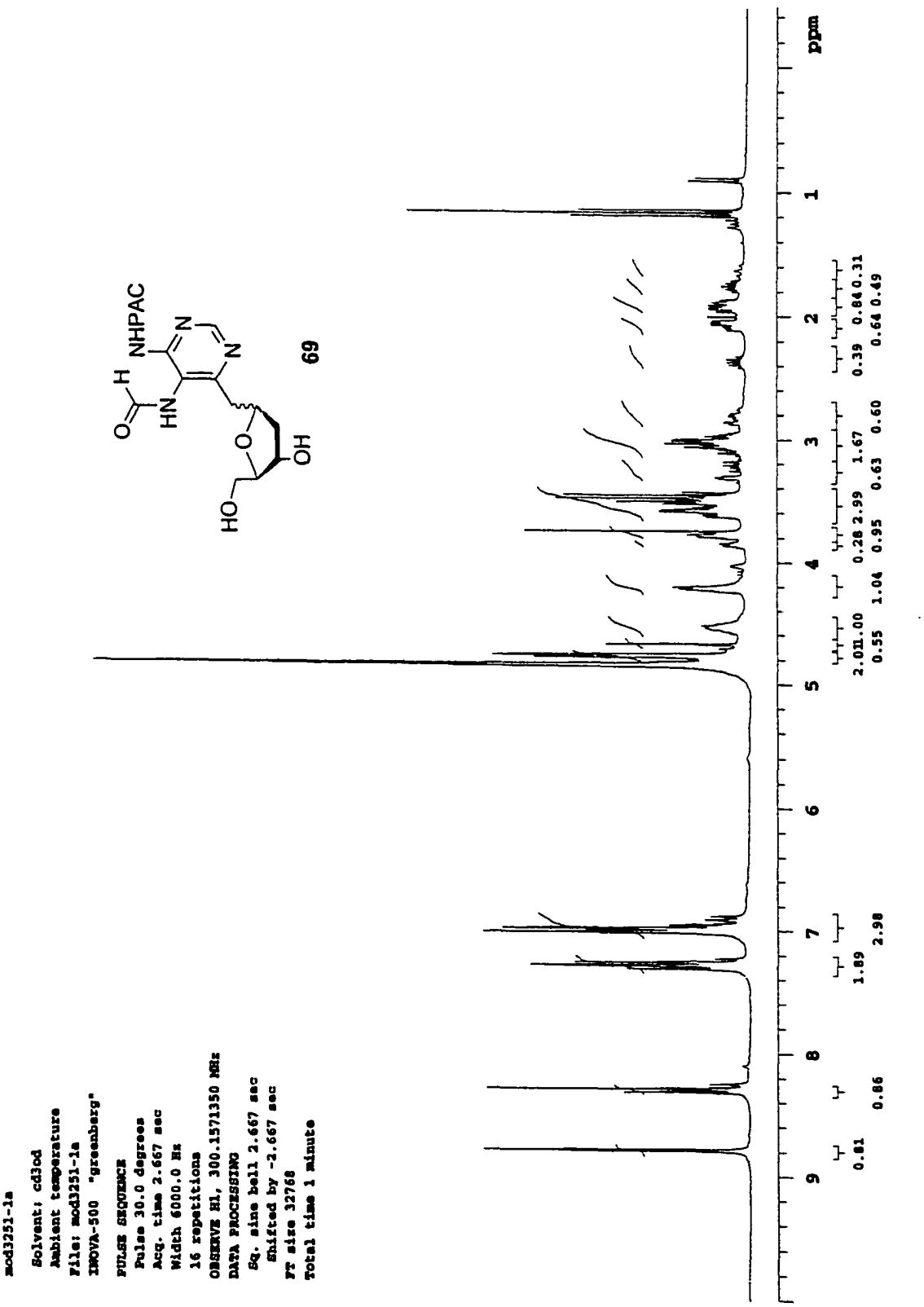


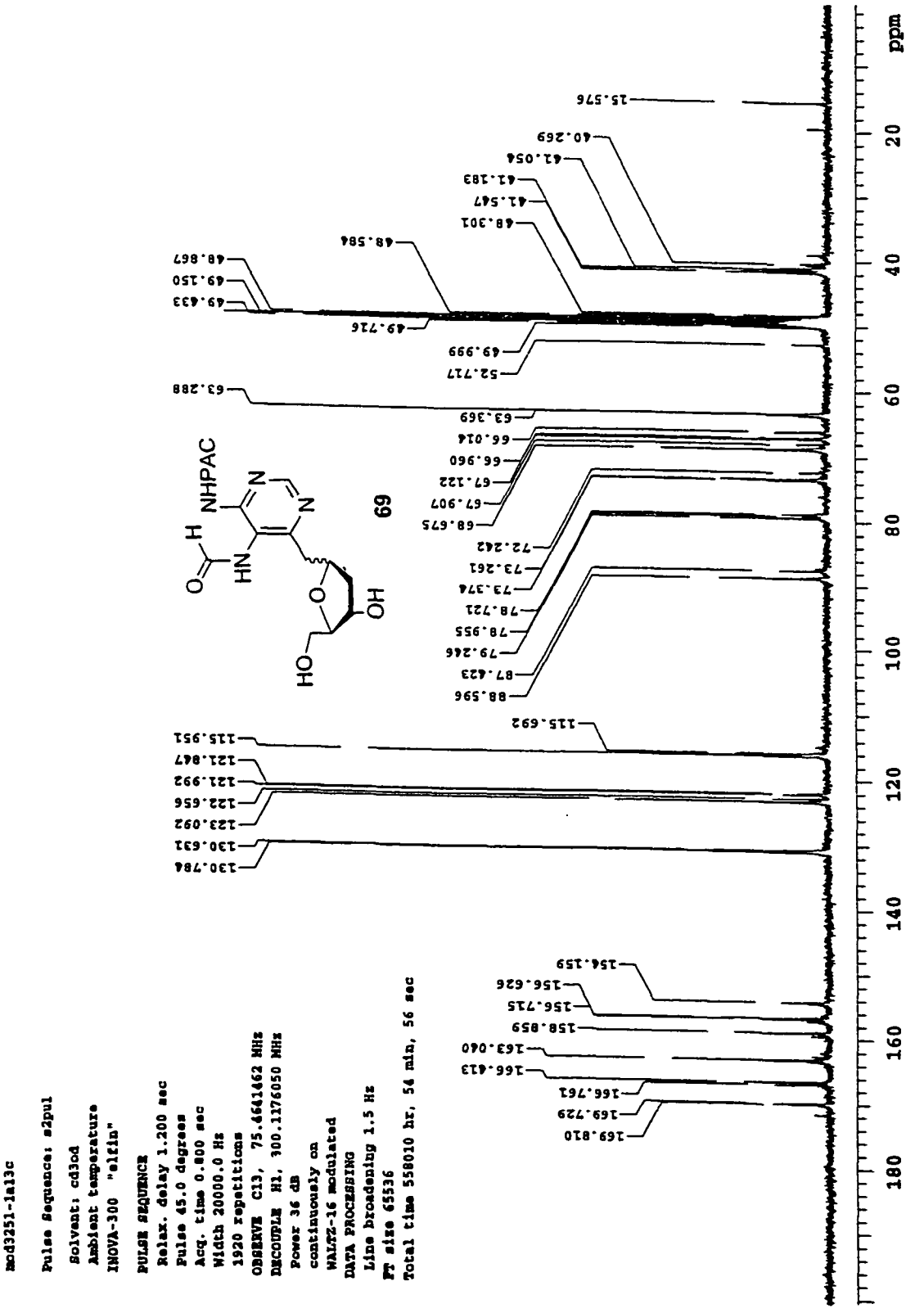


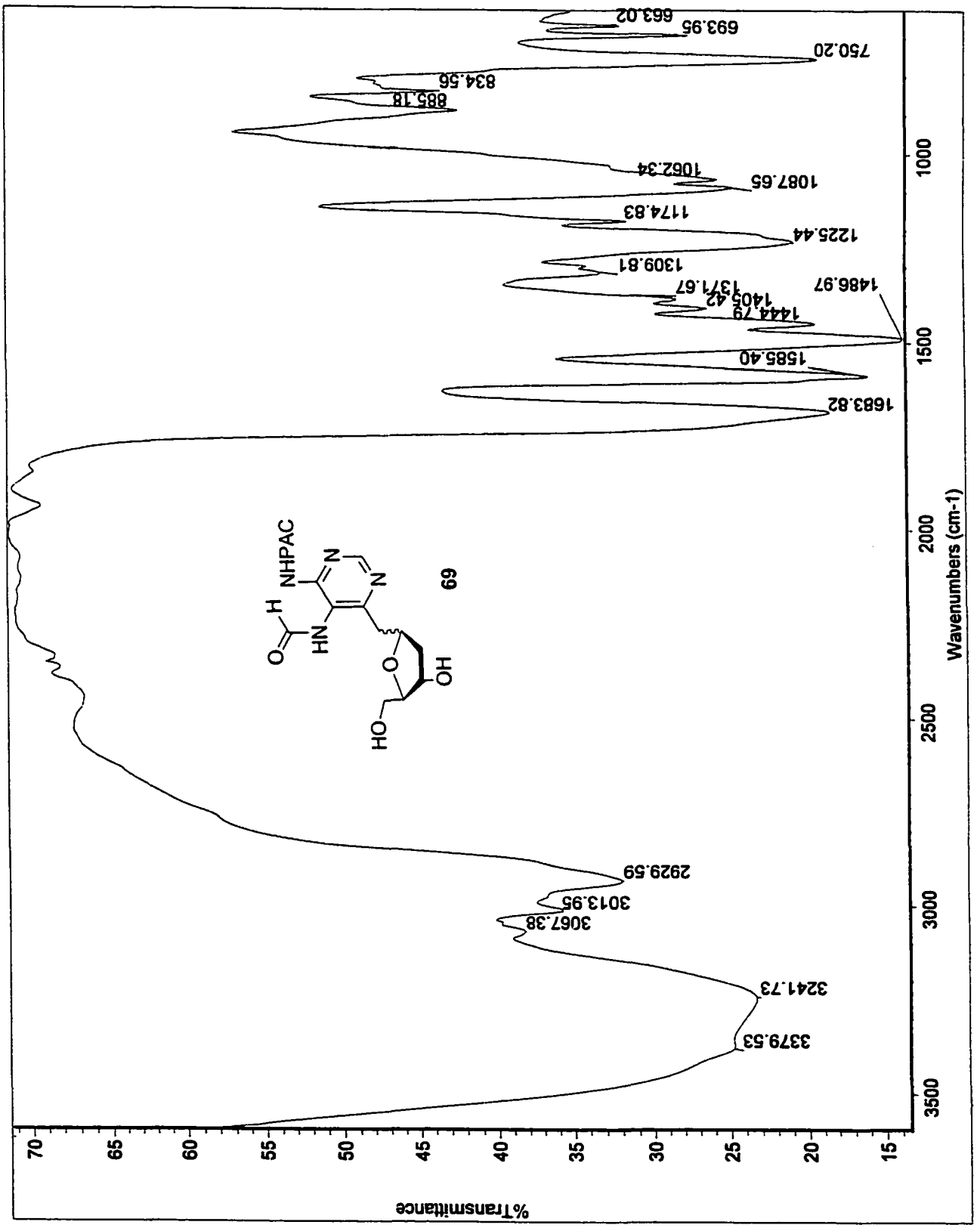


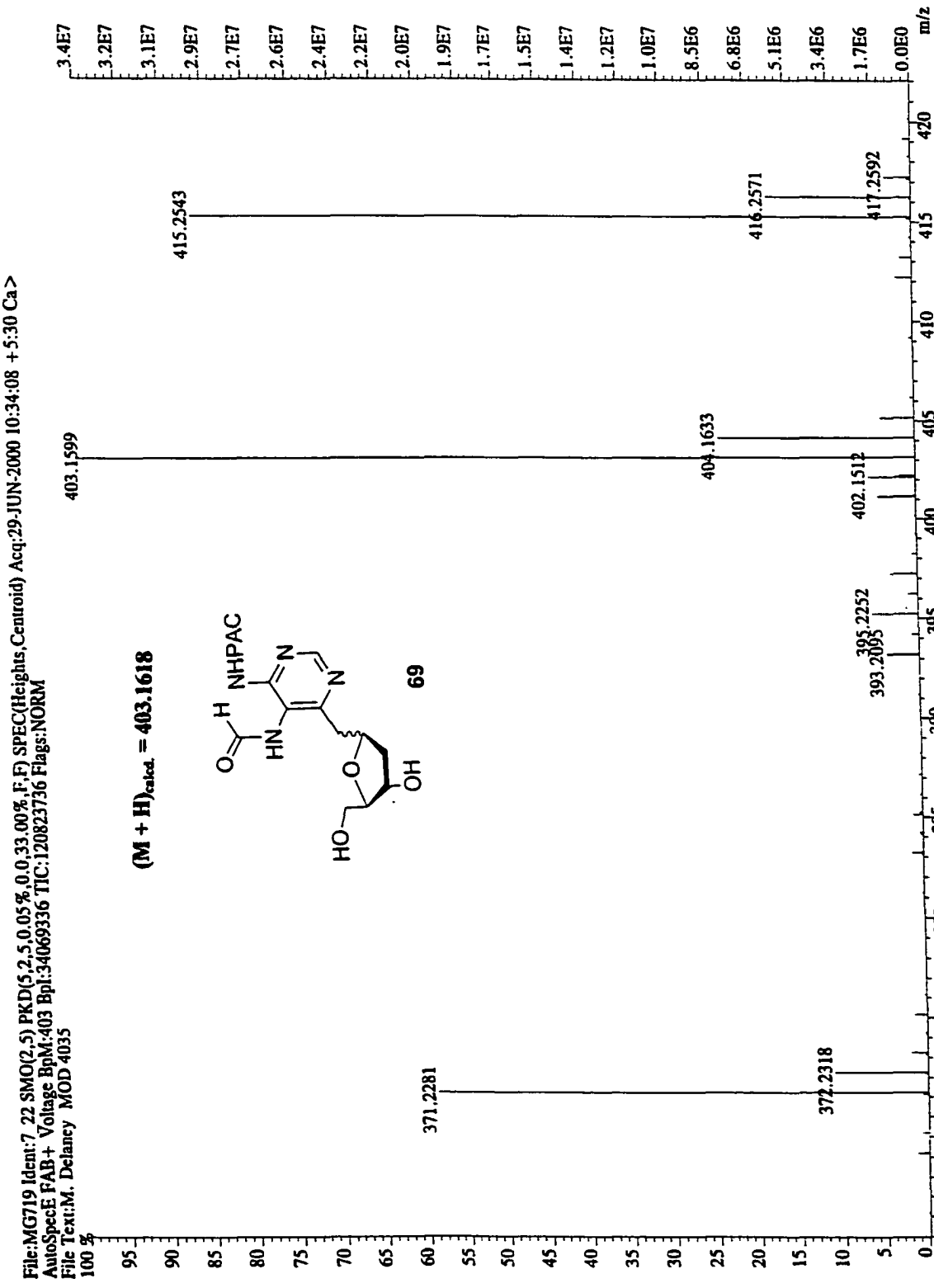


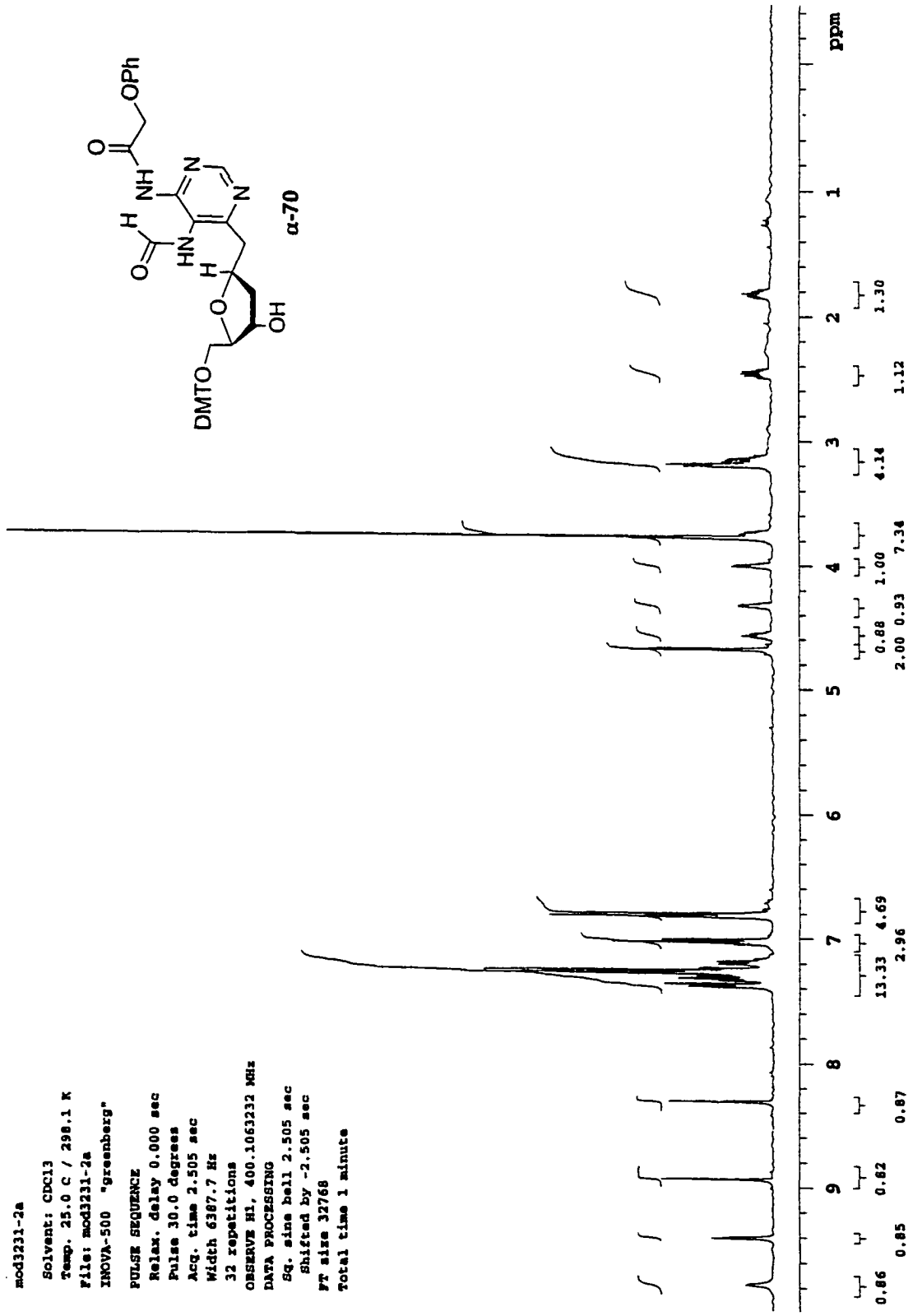


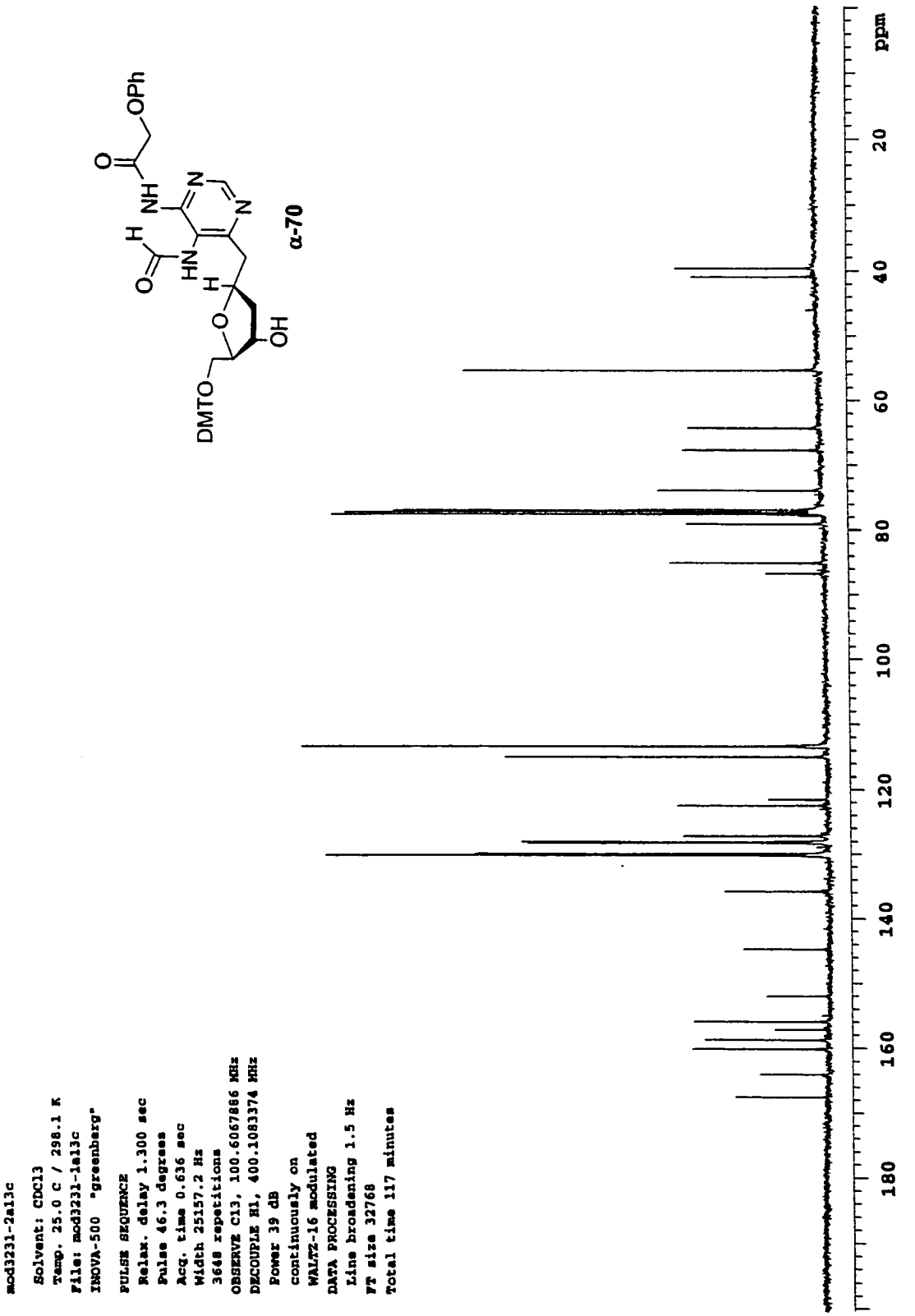


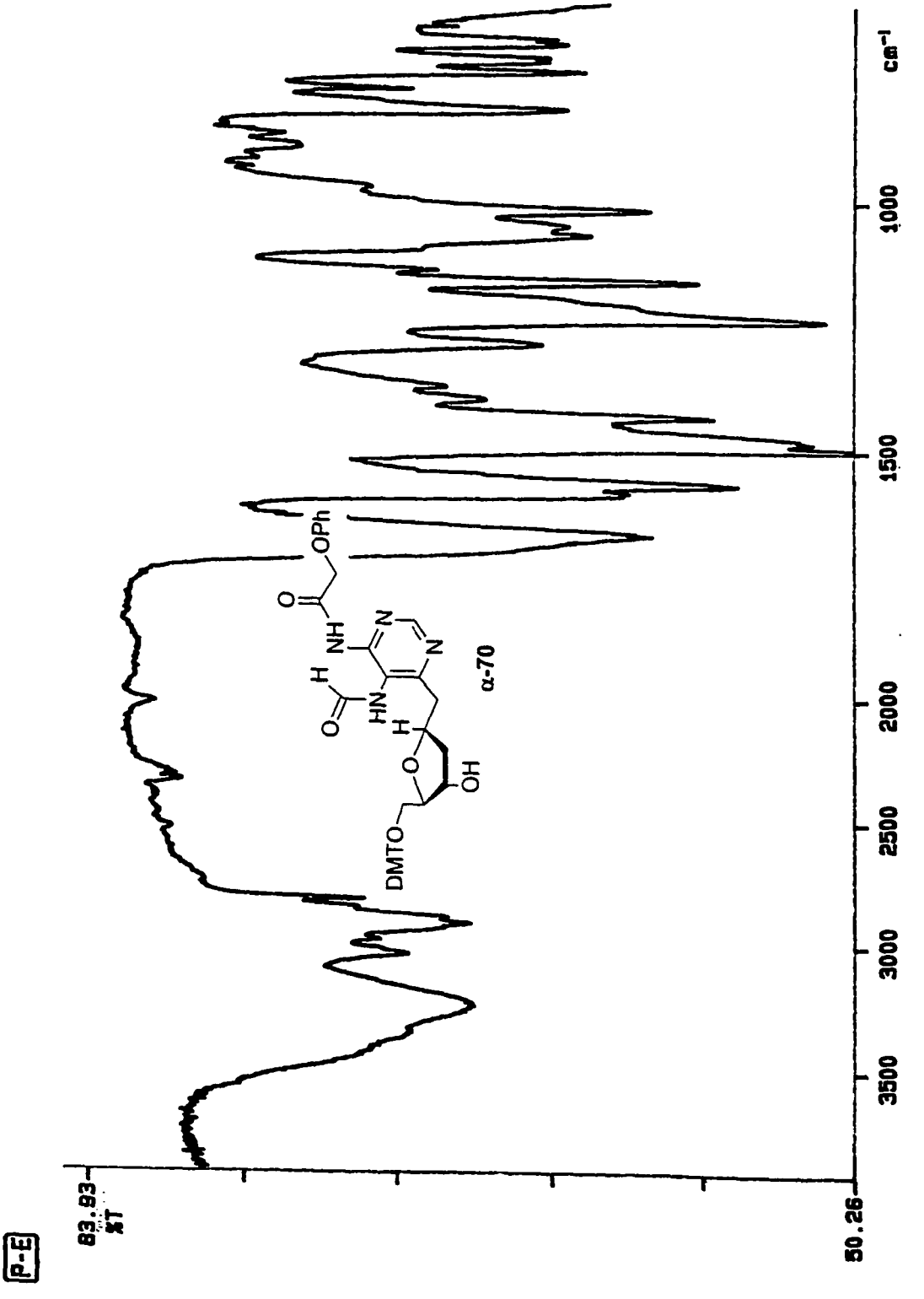


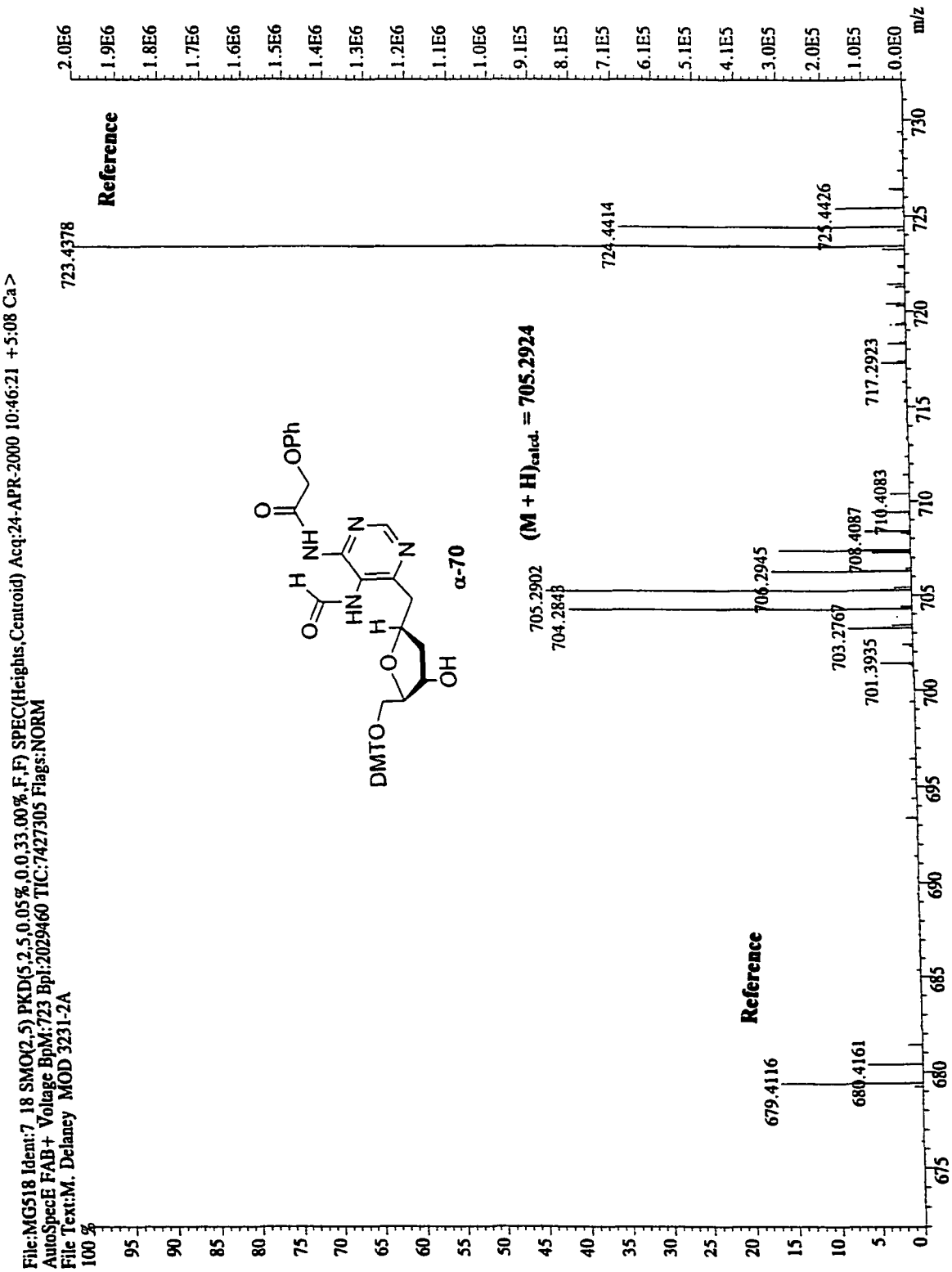


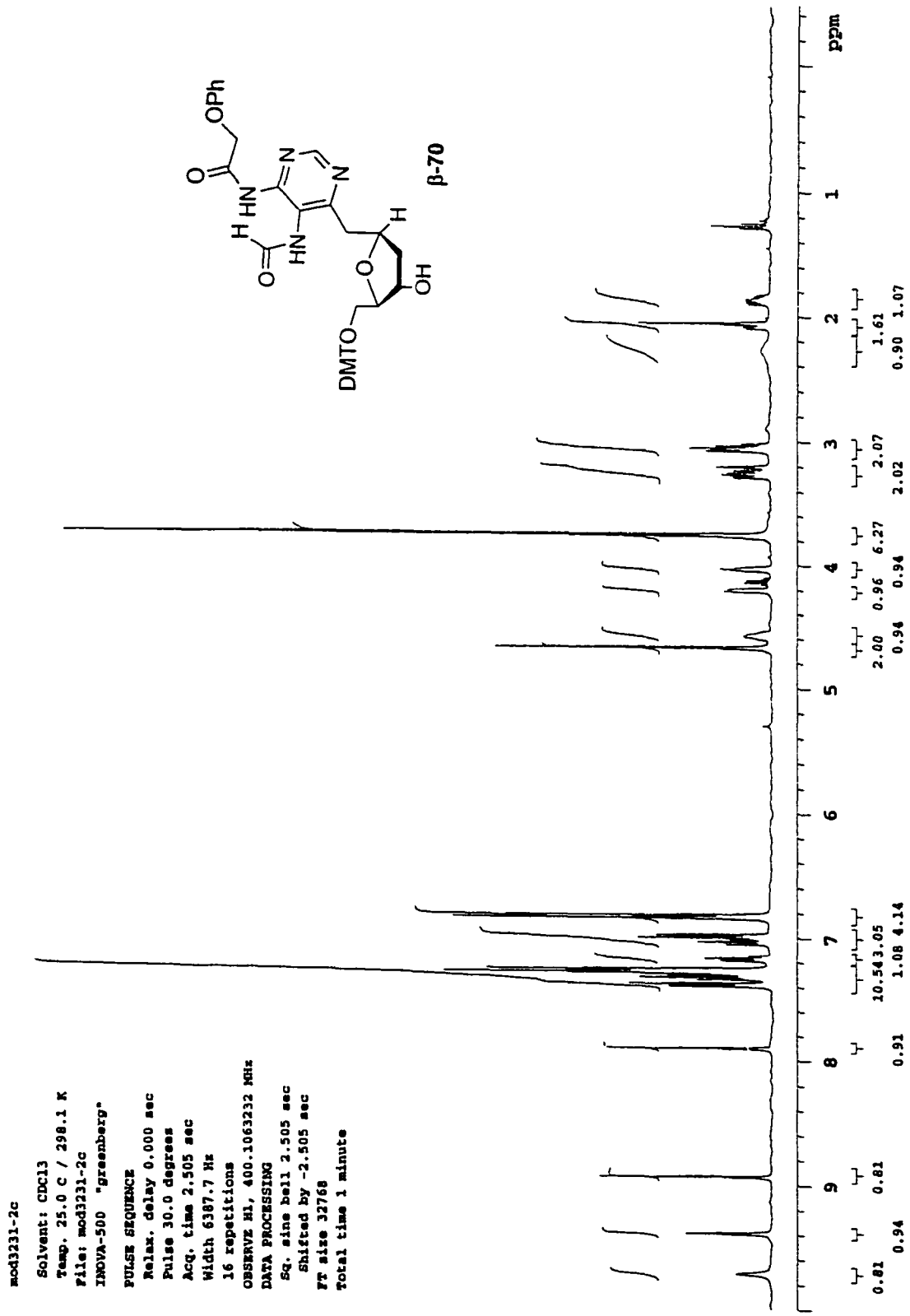


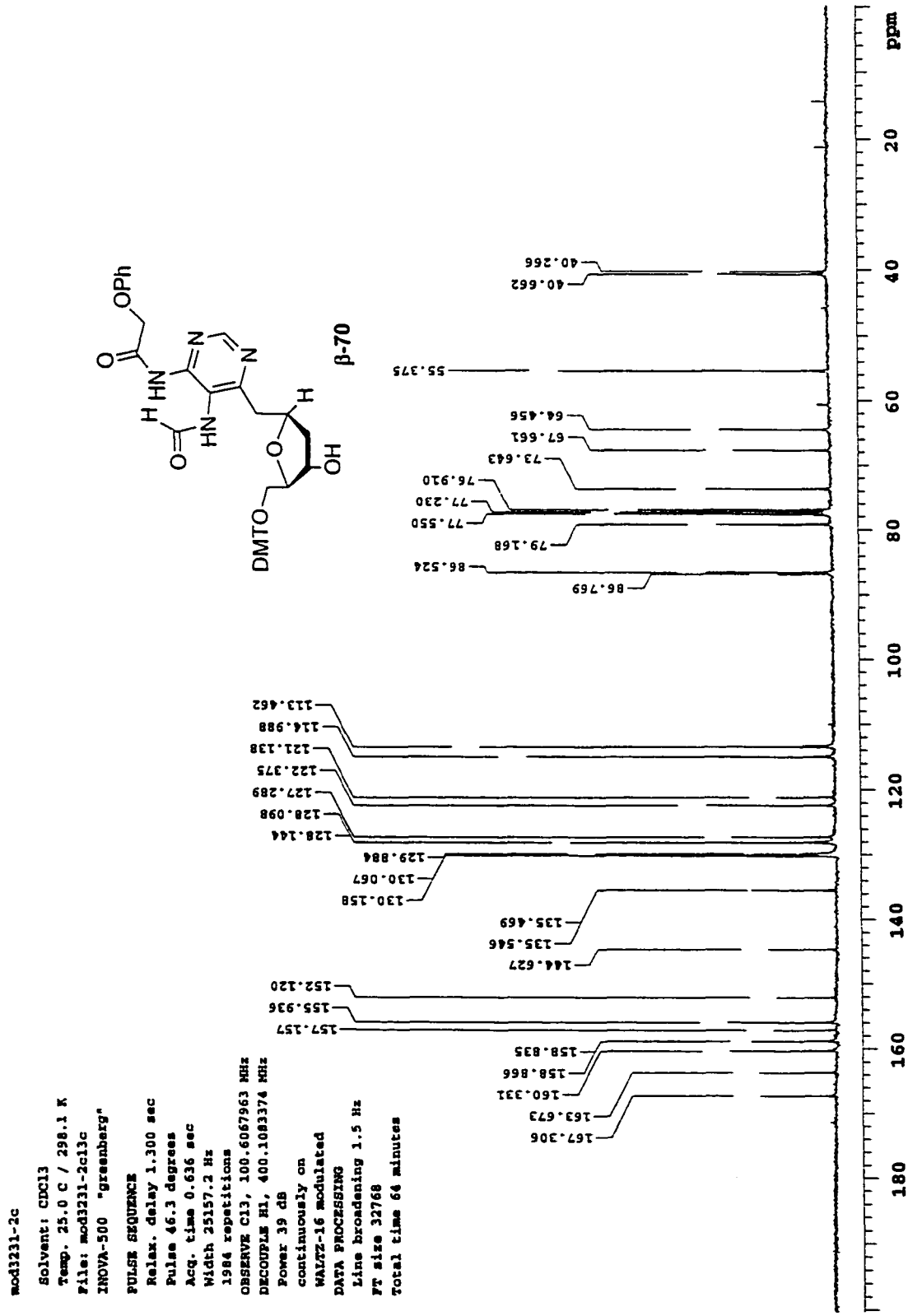


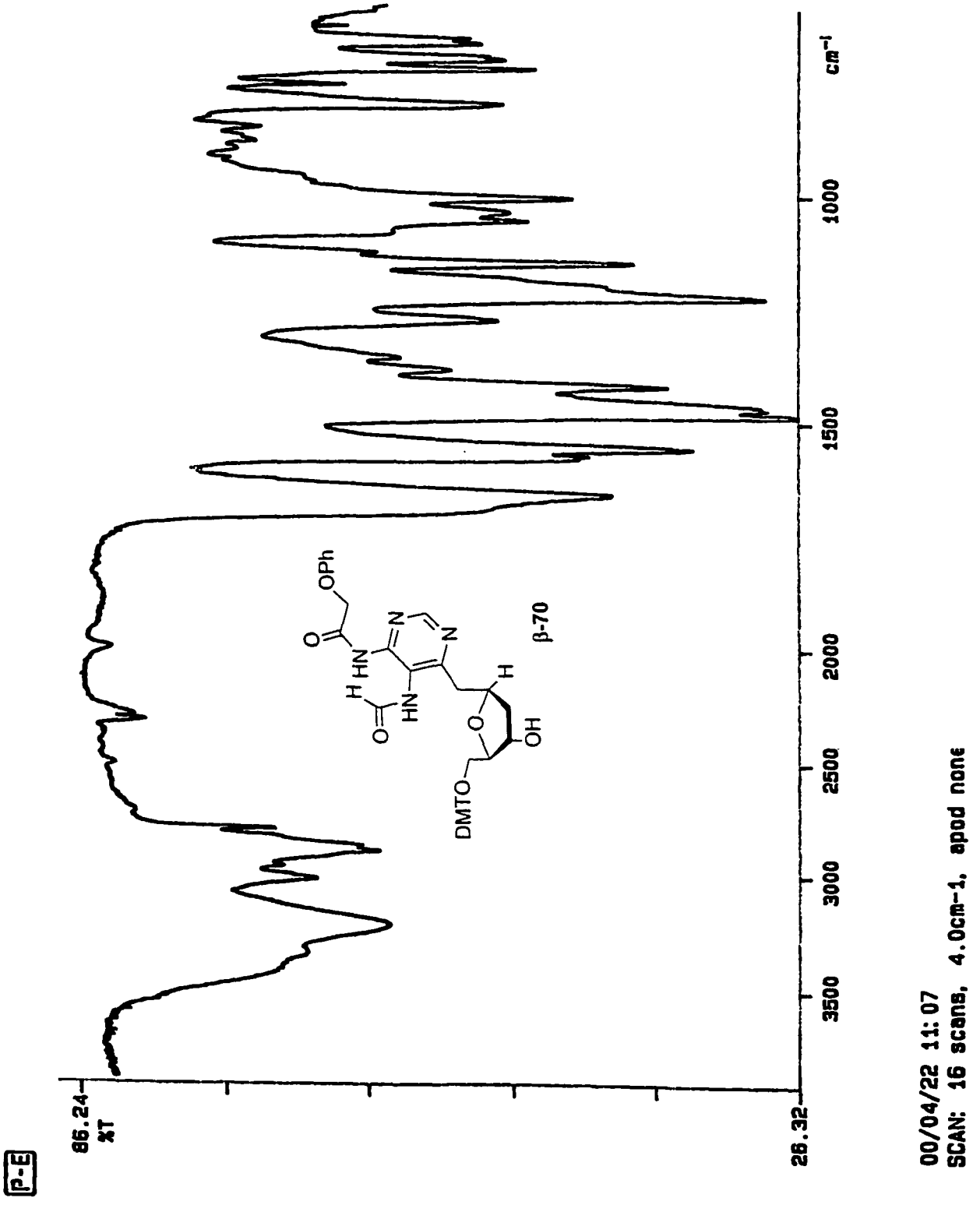


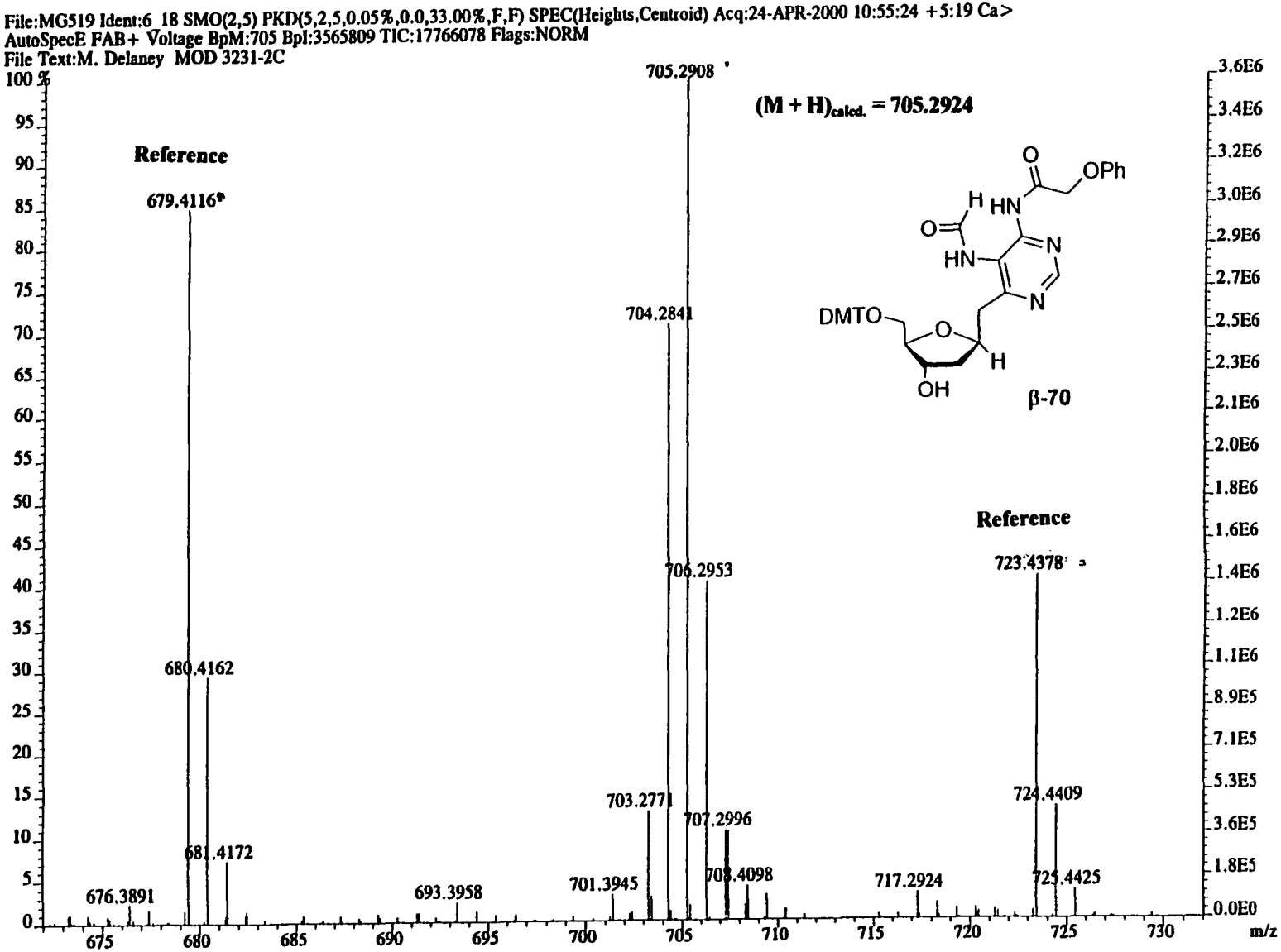


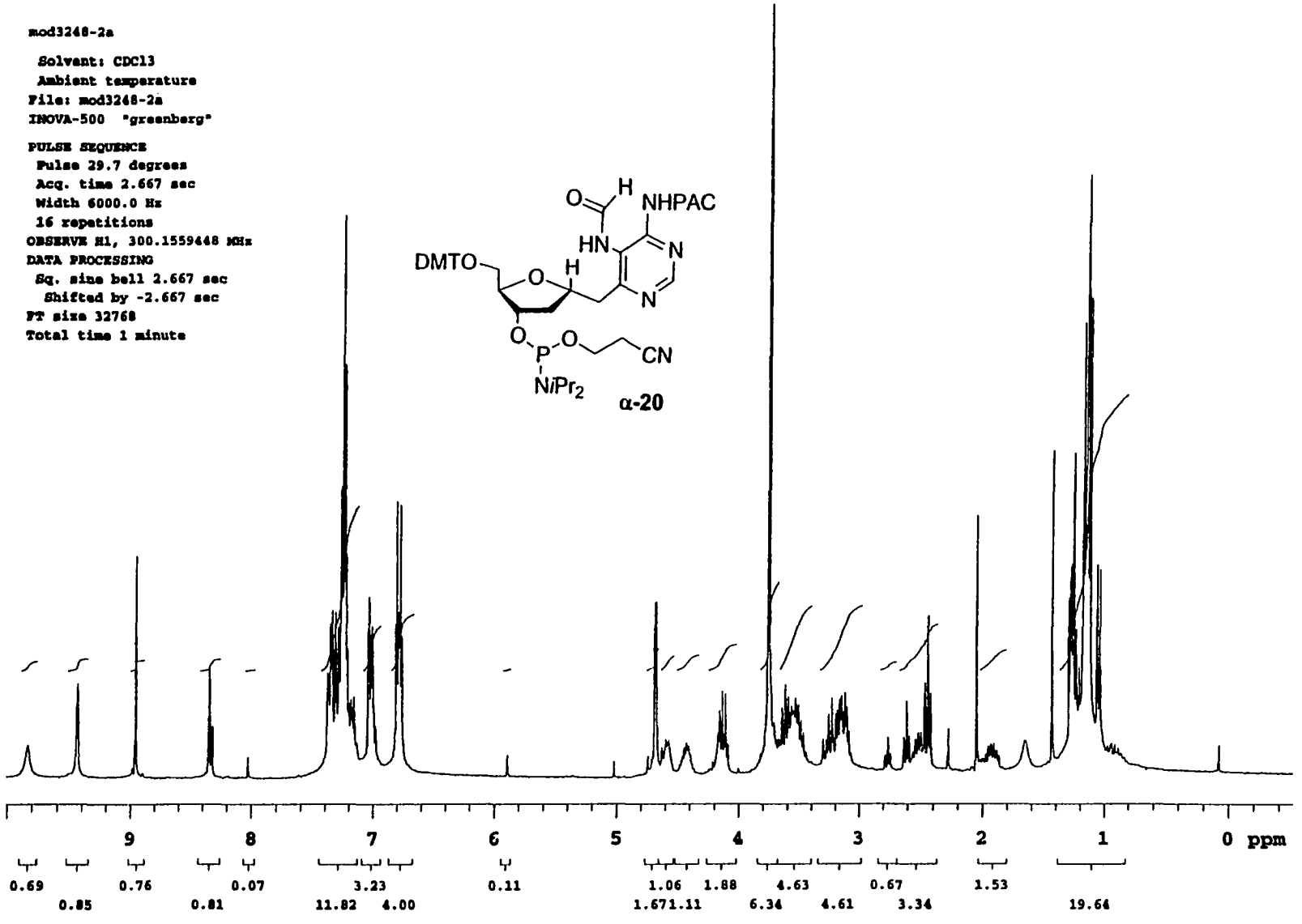


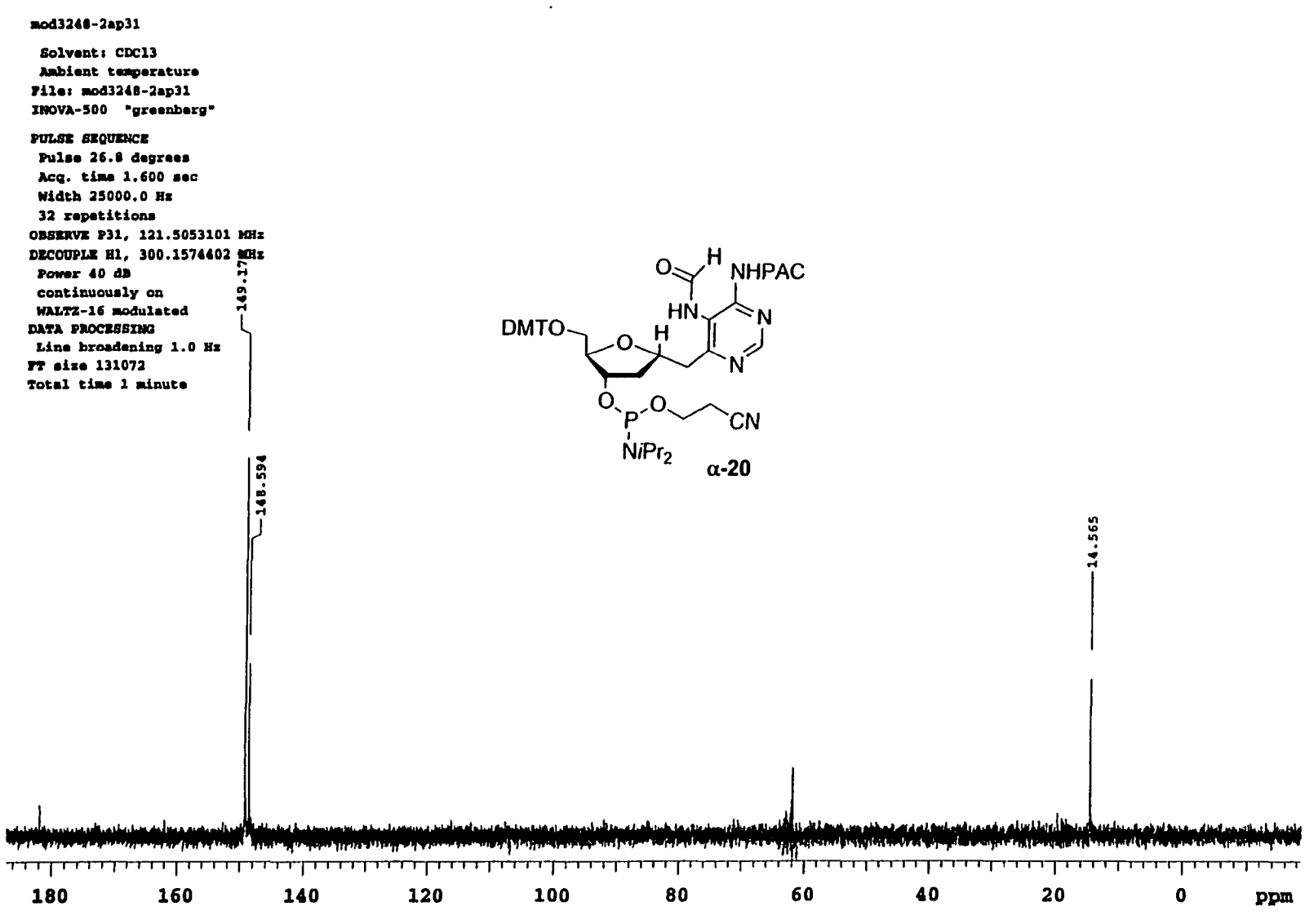


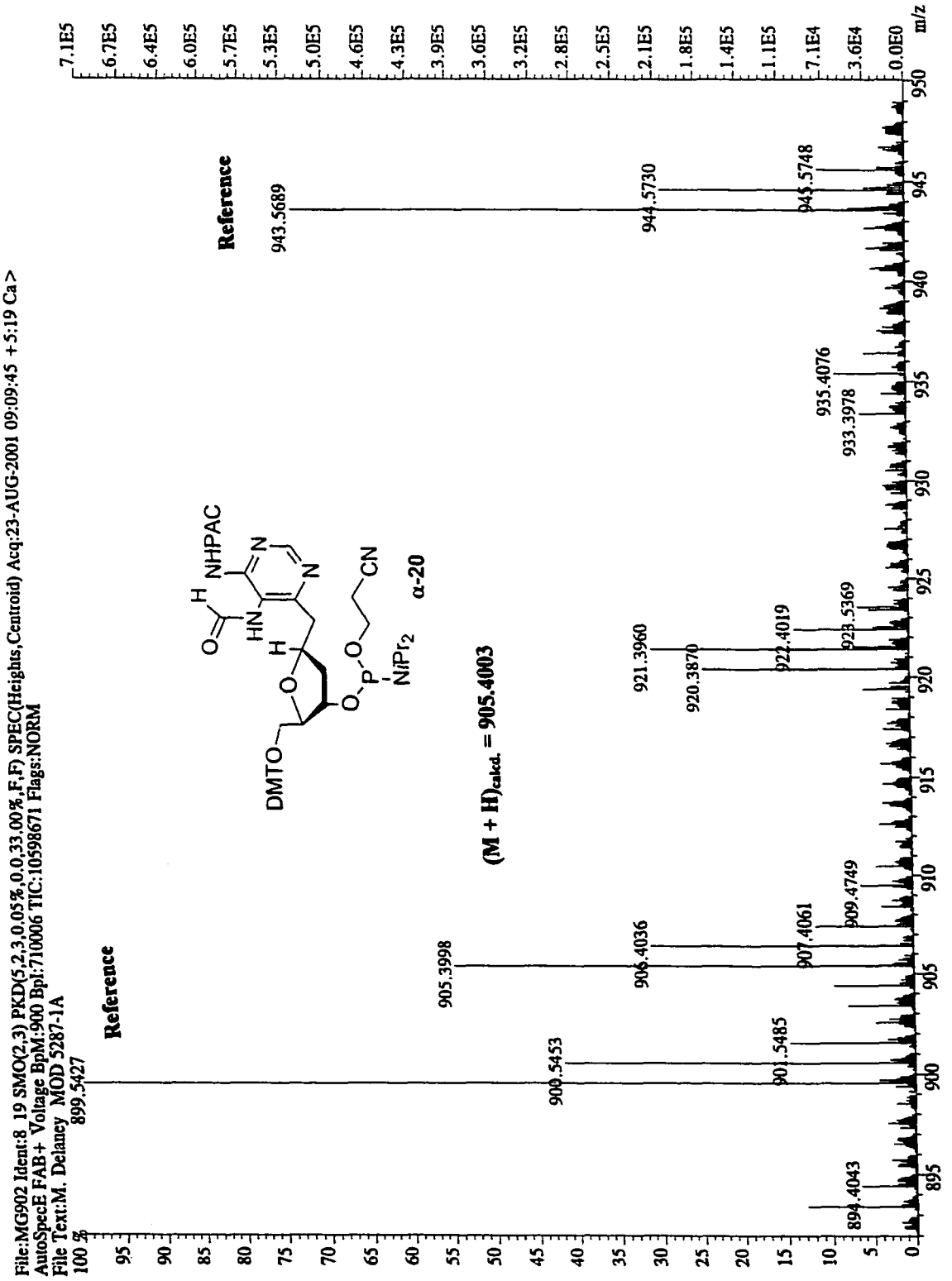


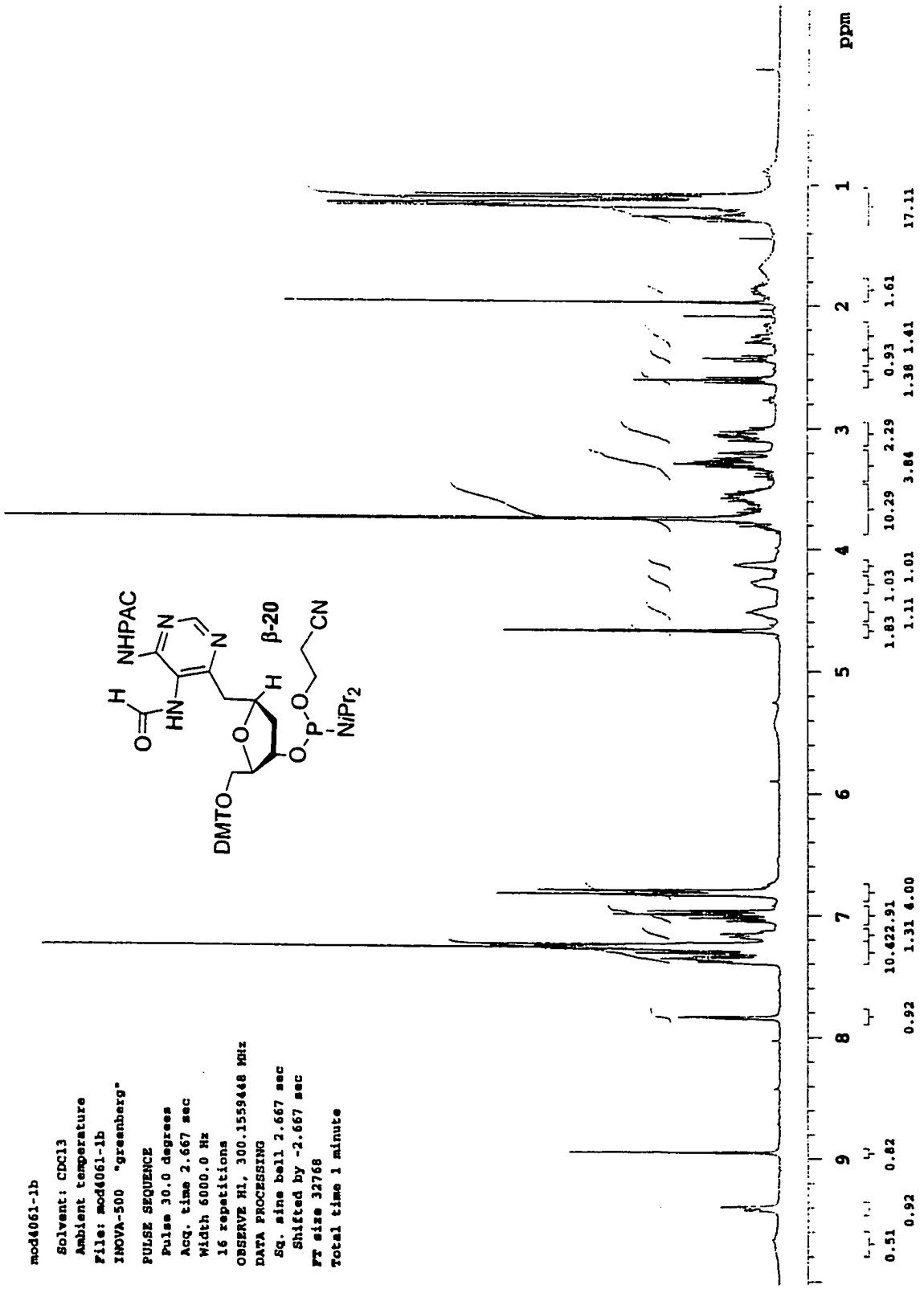








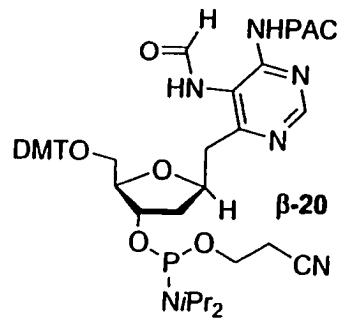


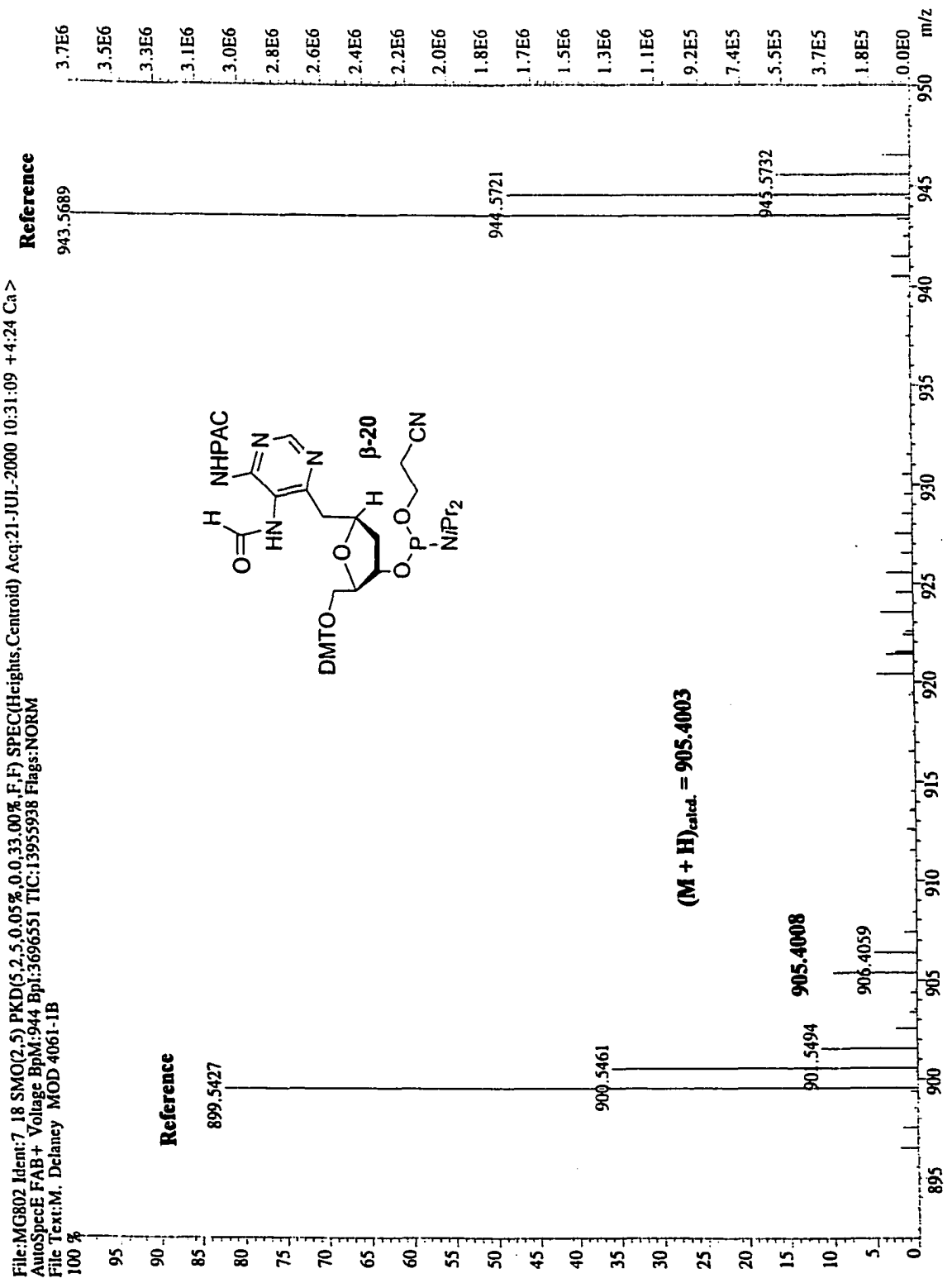


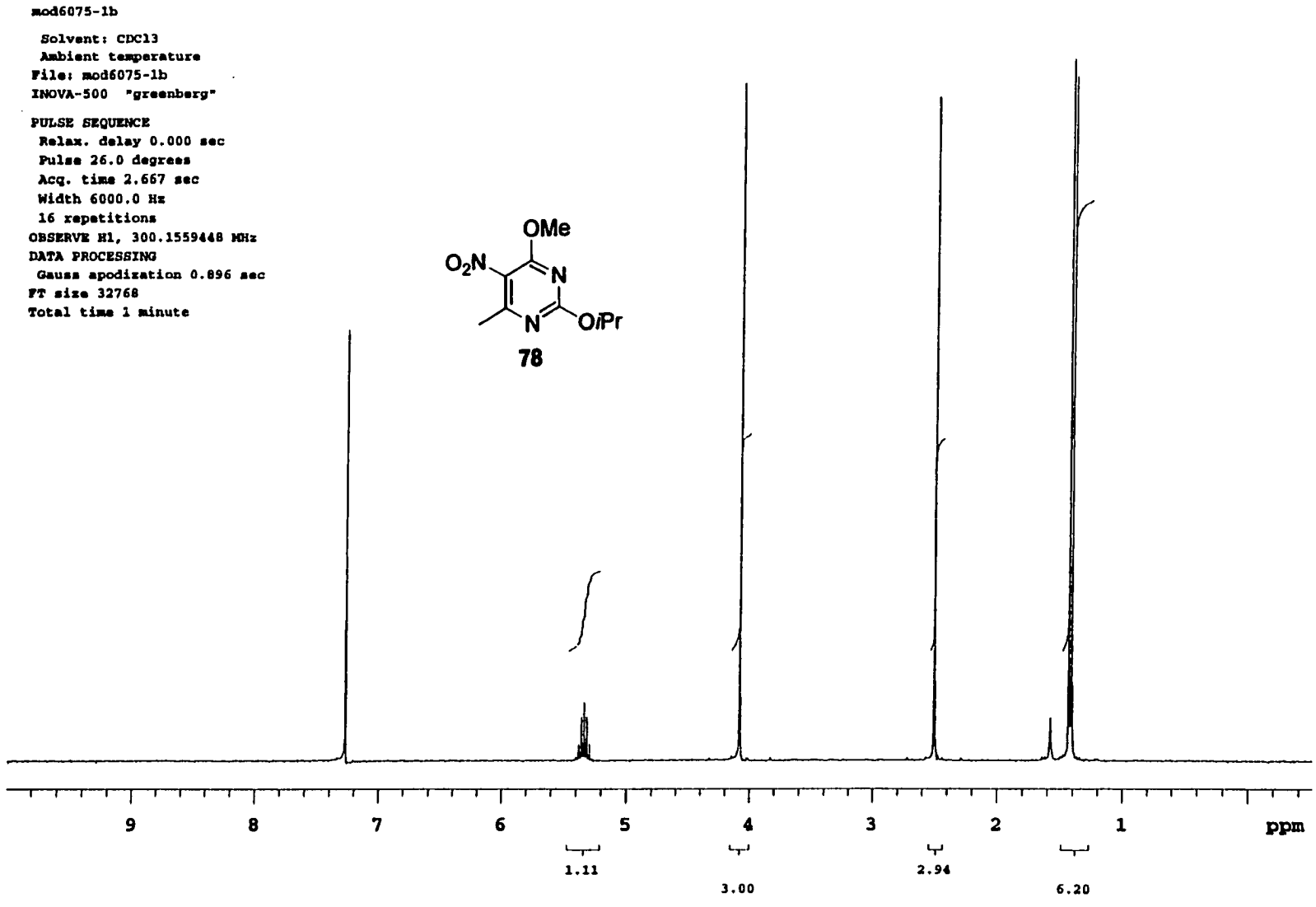
mod4061-1bp

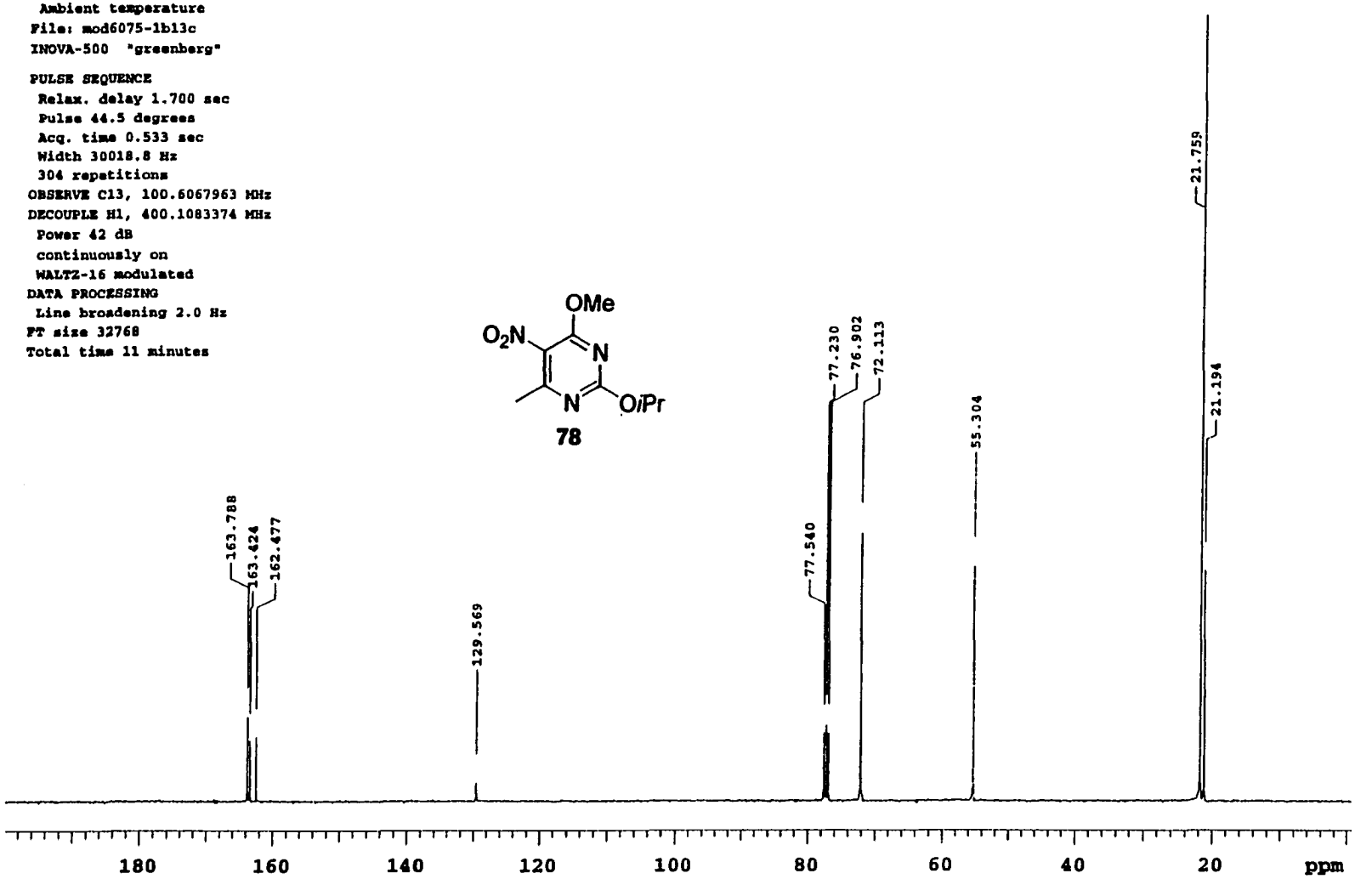
Solvent: CDCl3
Ambient temperature
File: mod4061-1bp
INOVA-500 "greenberg"

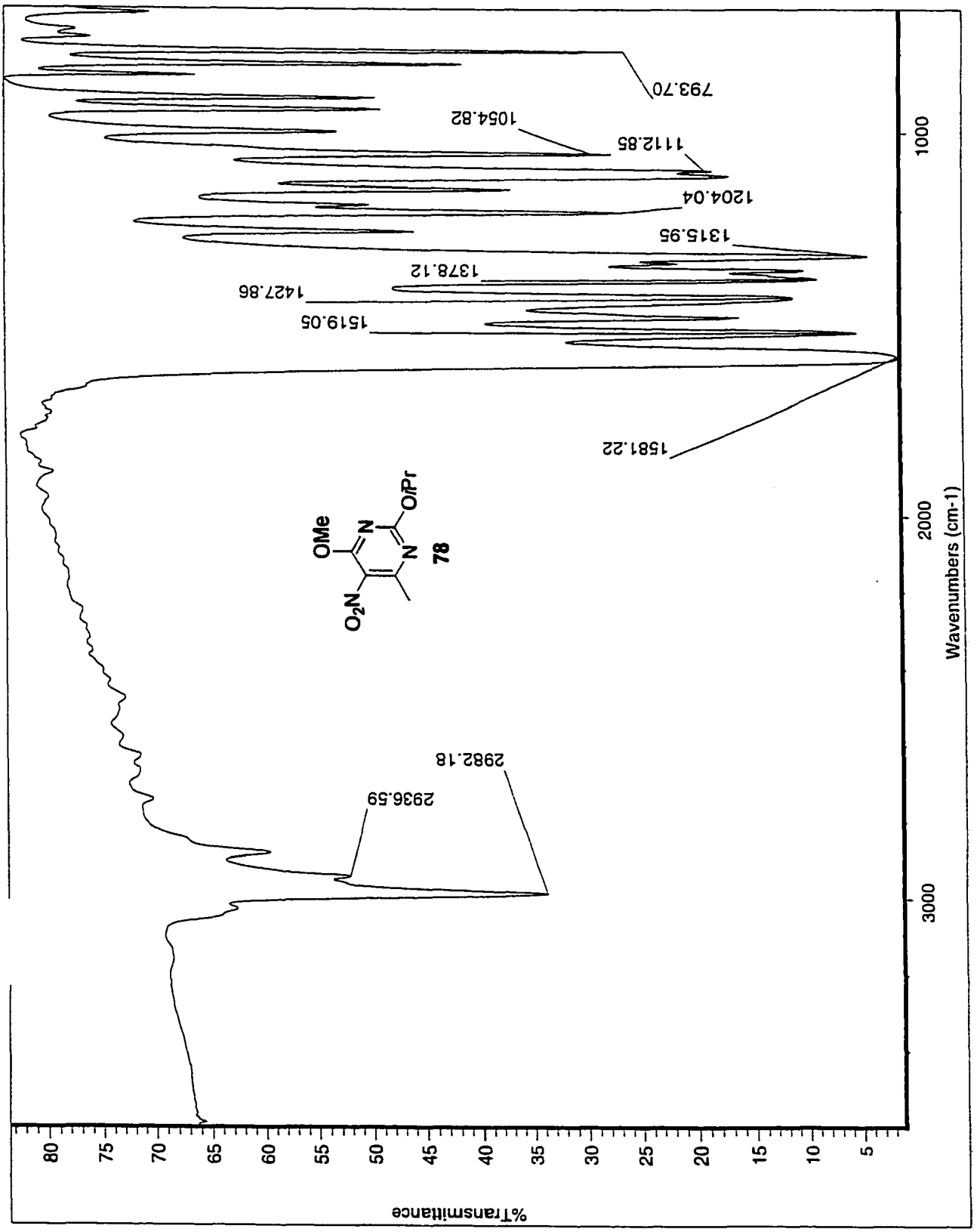
PULSE SEQUENCE
Pulse 26.8 degrees
Acq. time 1.600 sec
Width 25000.0 Hz
32 repetitions
OBSERVE P31, 121.5053101 MHz
DECOUPLE H1, 300.1574402 MHz
Power 40 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 131072
Total time 1 minute

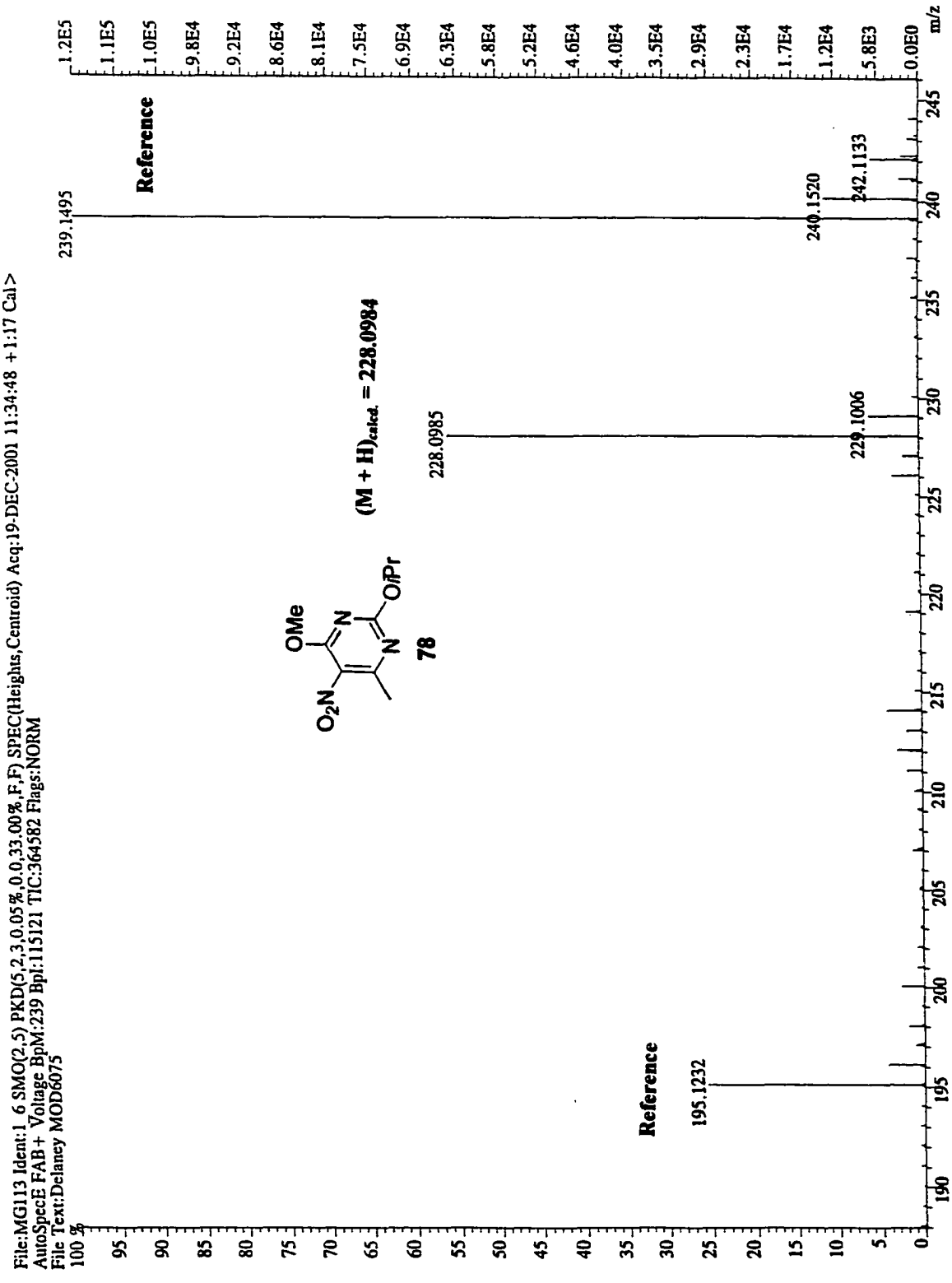


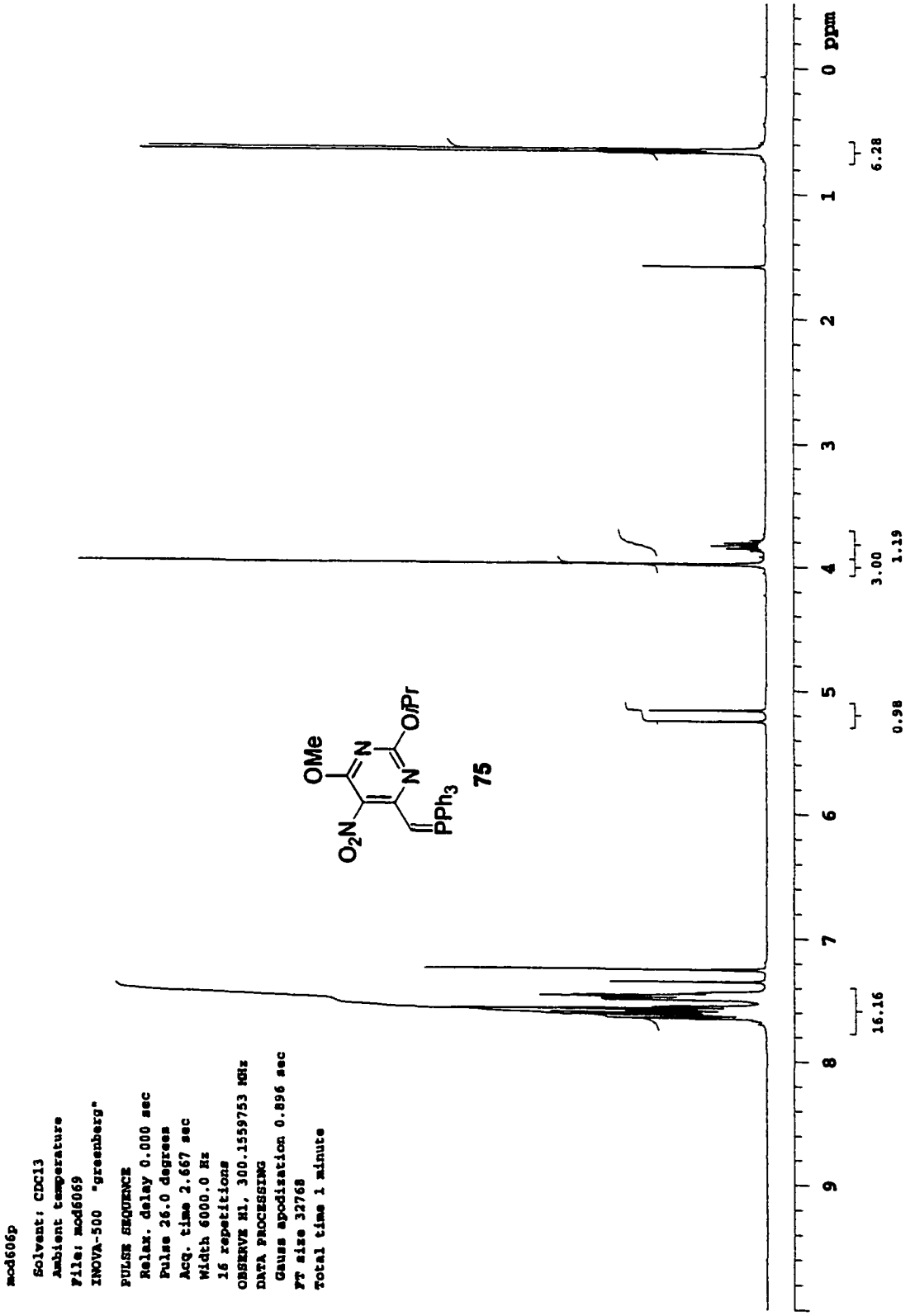








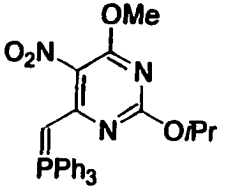




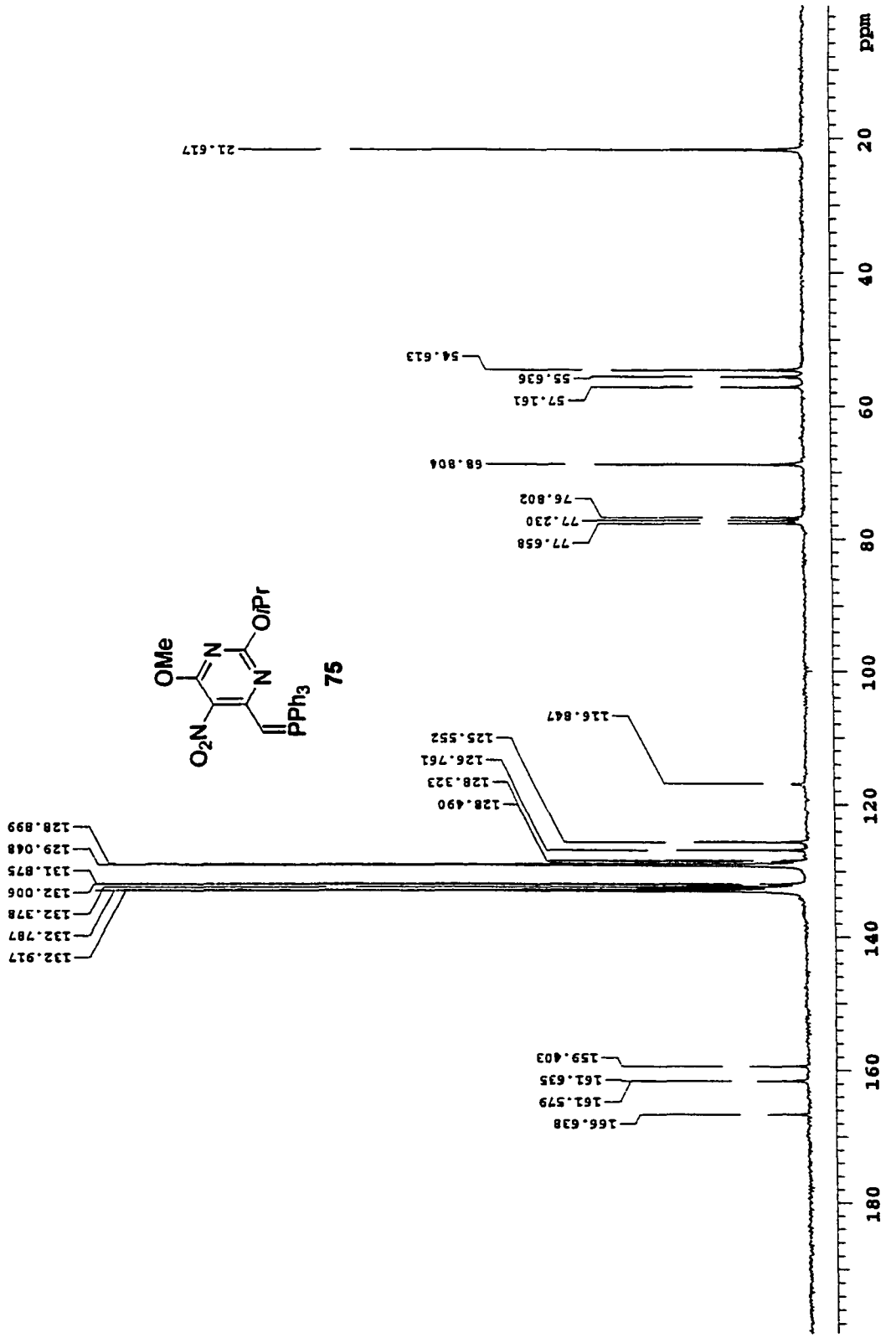
P-31 STANDARD PARAMETERS
PHOSPHATE REGION

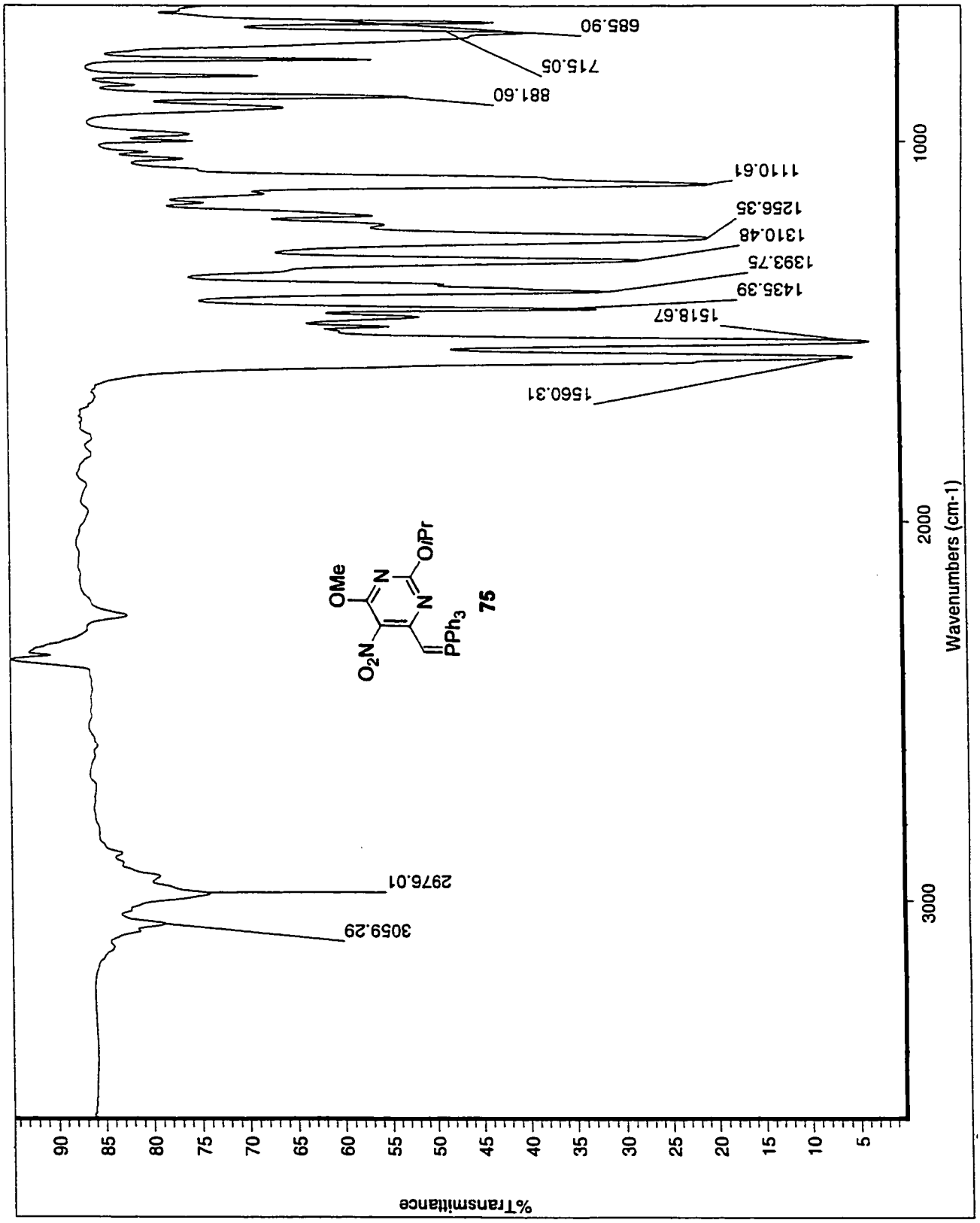
Solvent: CDCl3
Ambient temperature
File: mod6069-p31-2
INOVA-500 "greenberg"

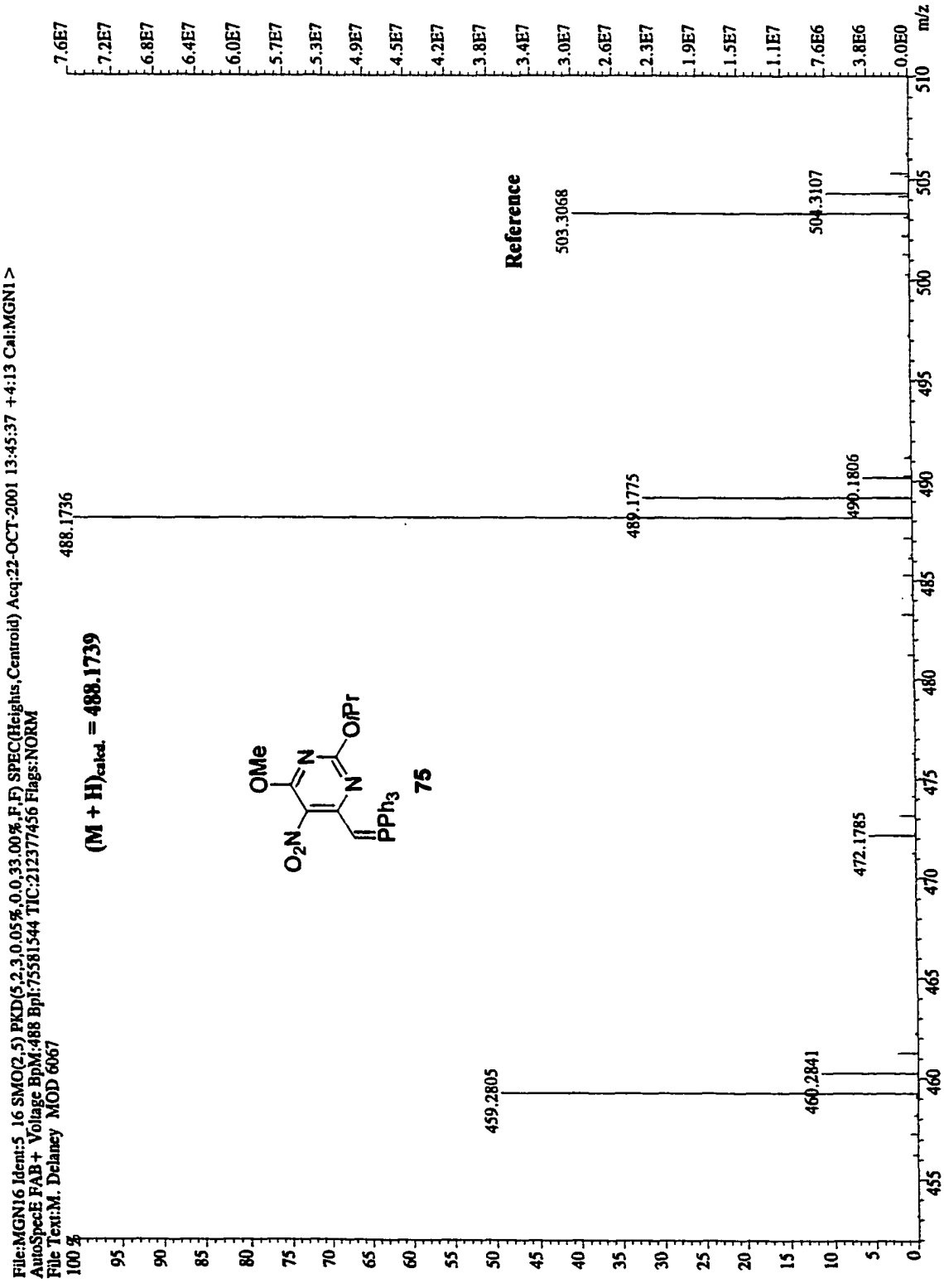
PULSE SEQUENCE
Relax. delay 0.000 sec
Pulse 45.0 degrees
Acq. time 1.067 sec
Width 30000.0 Hz
64 repetitions
OBSERVE P31, 121.5053253 MHz
DECOUPLE H1, 300.1574402 MHz
Power 40 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 1 minute

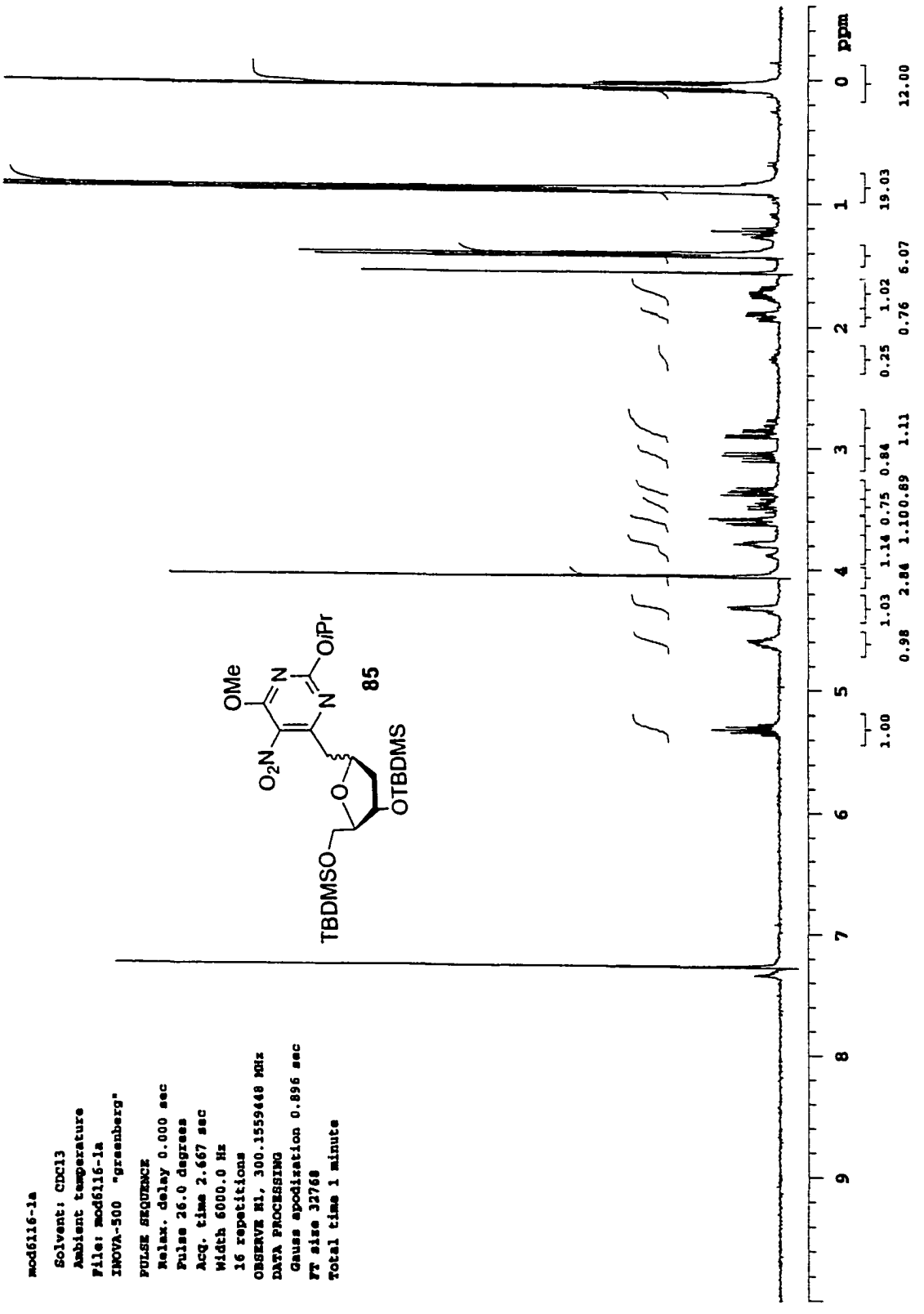


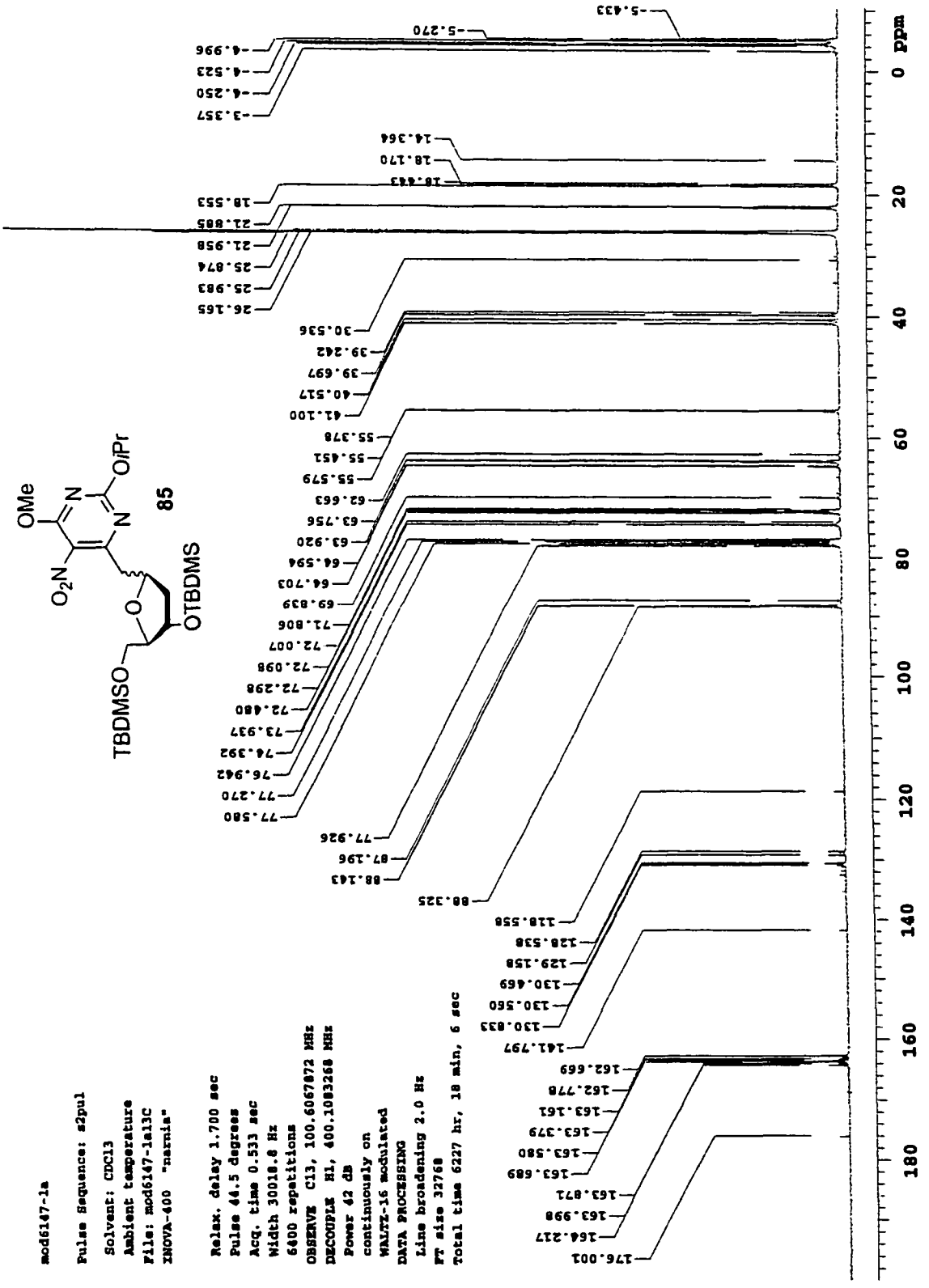
75

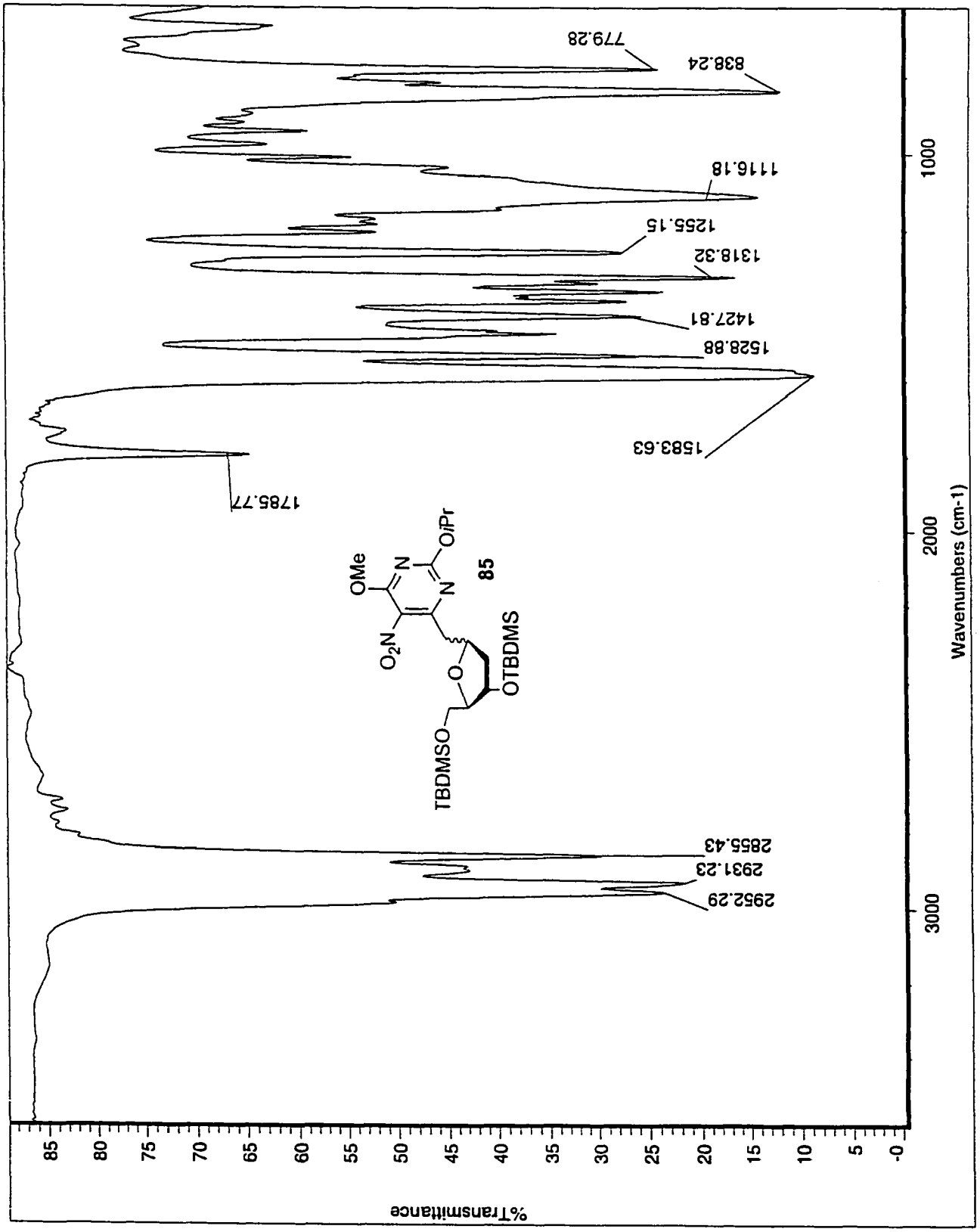


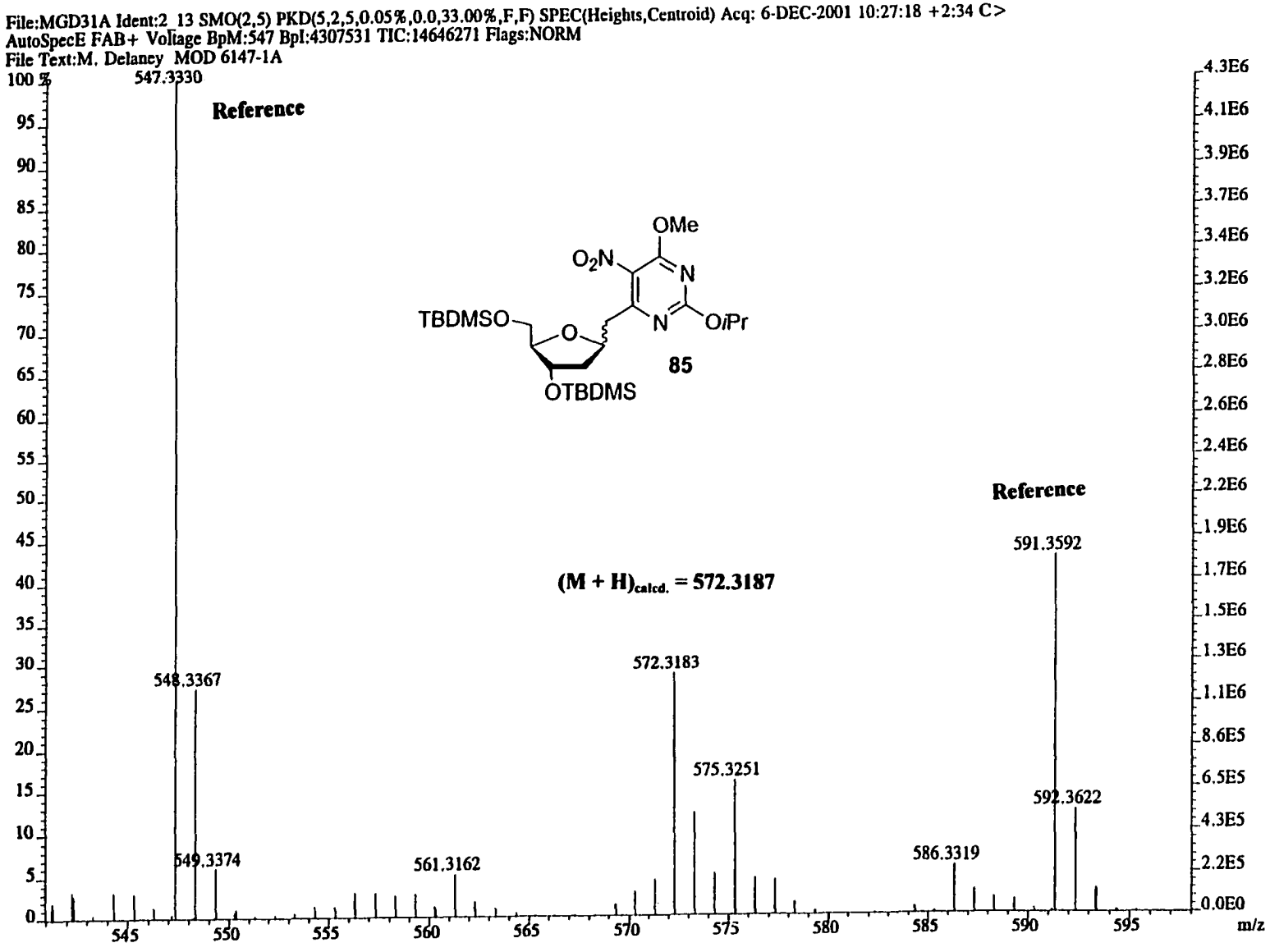


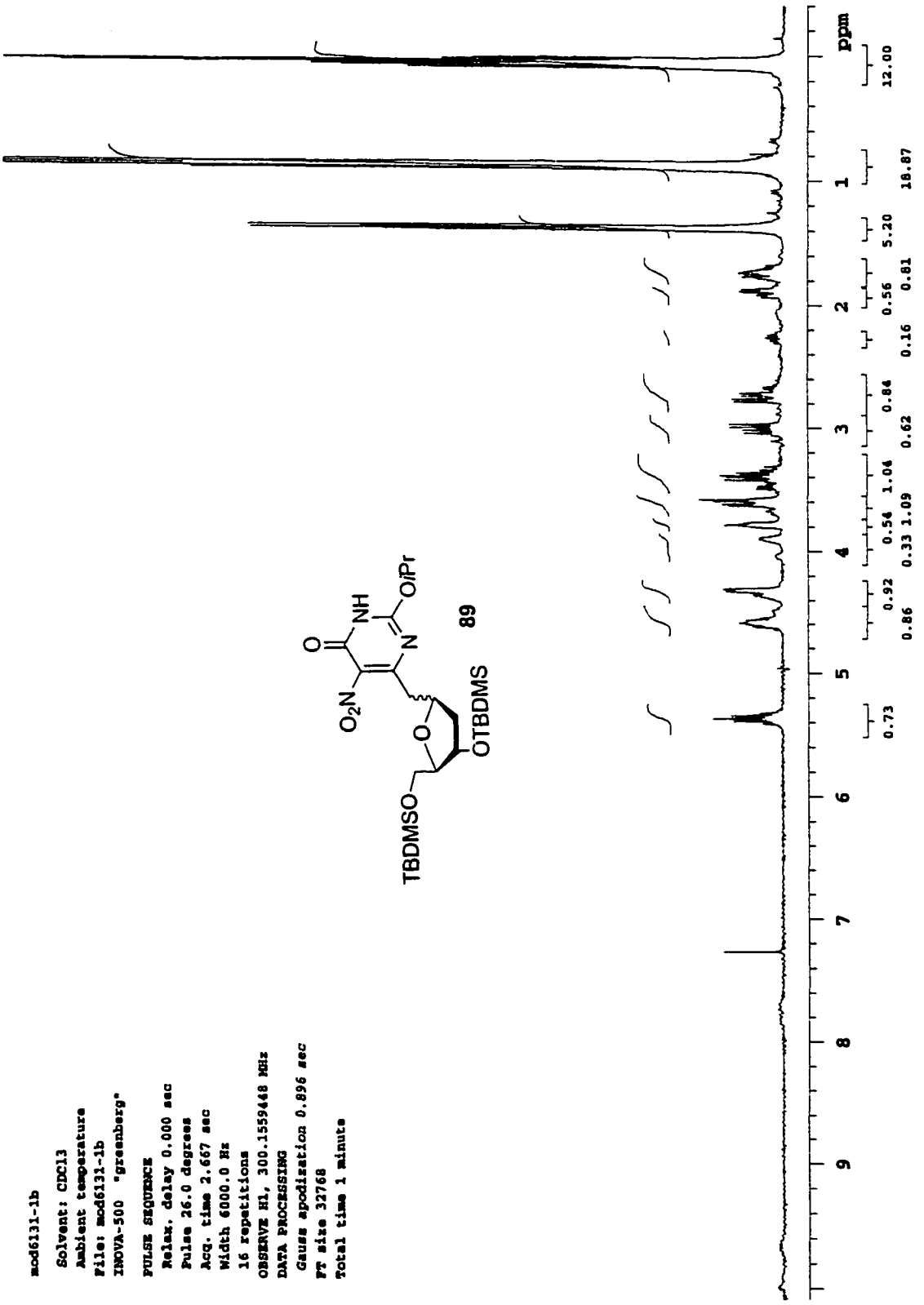


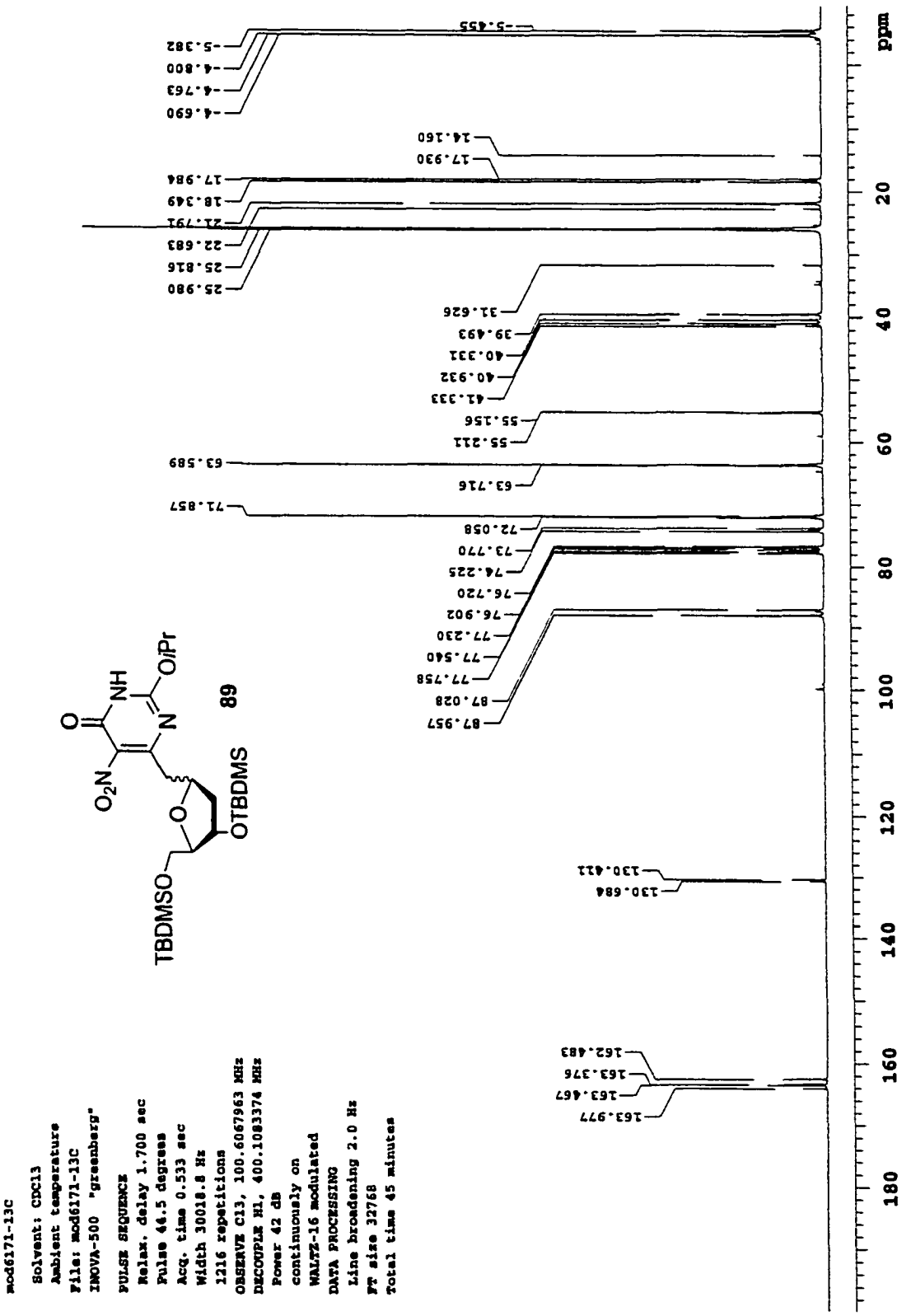


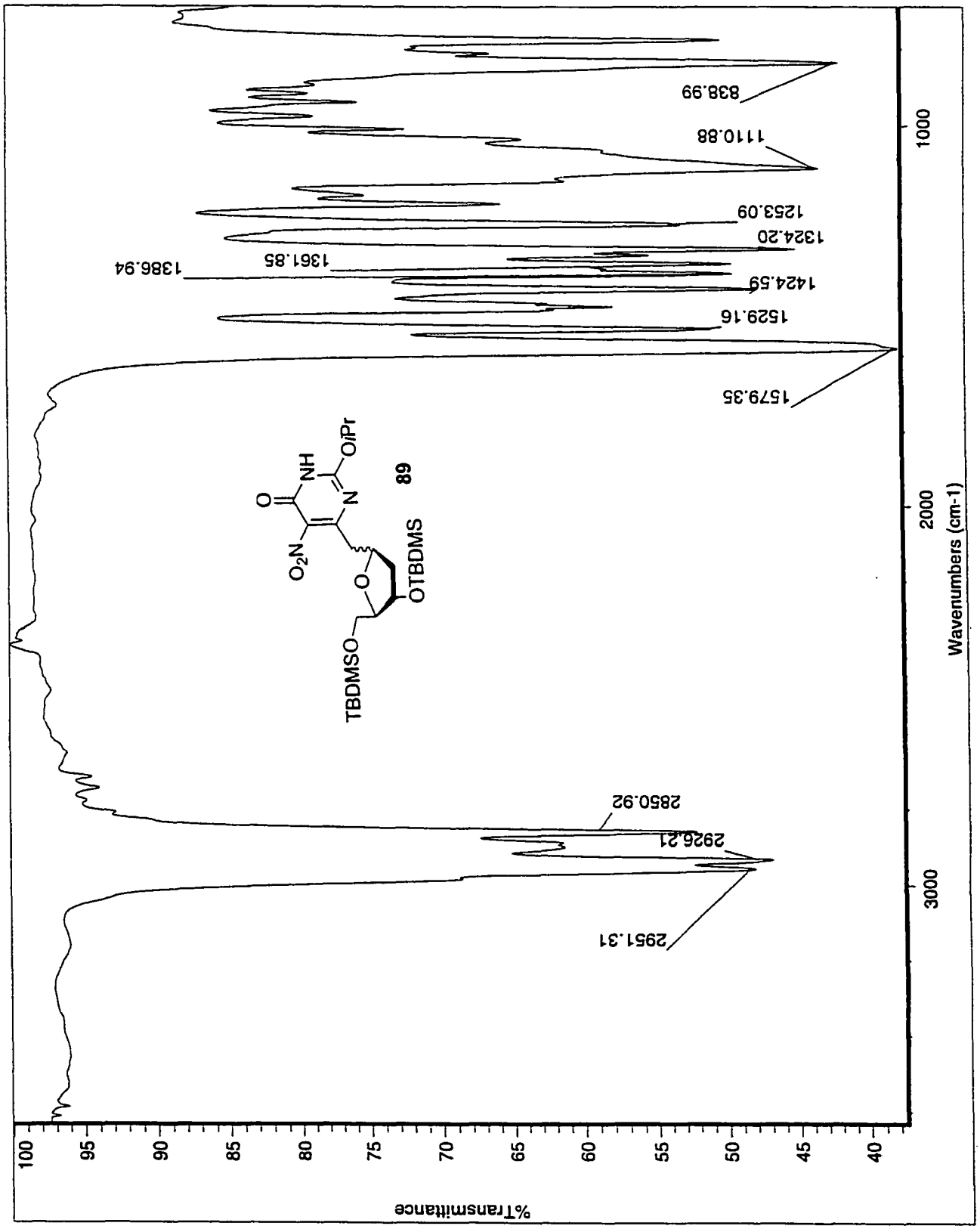


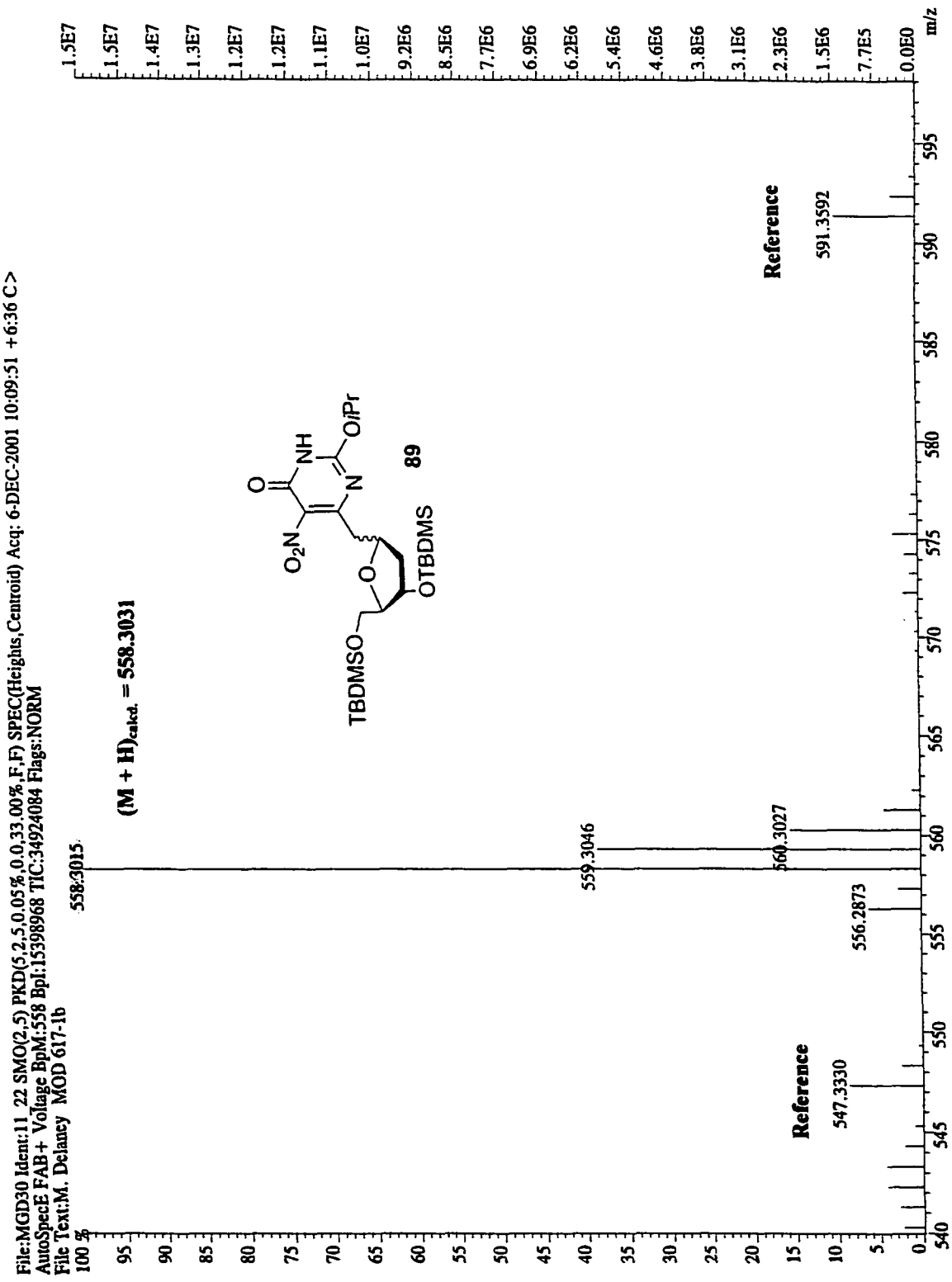


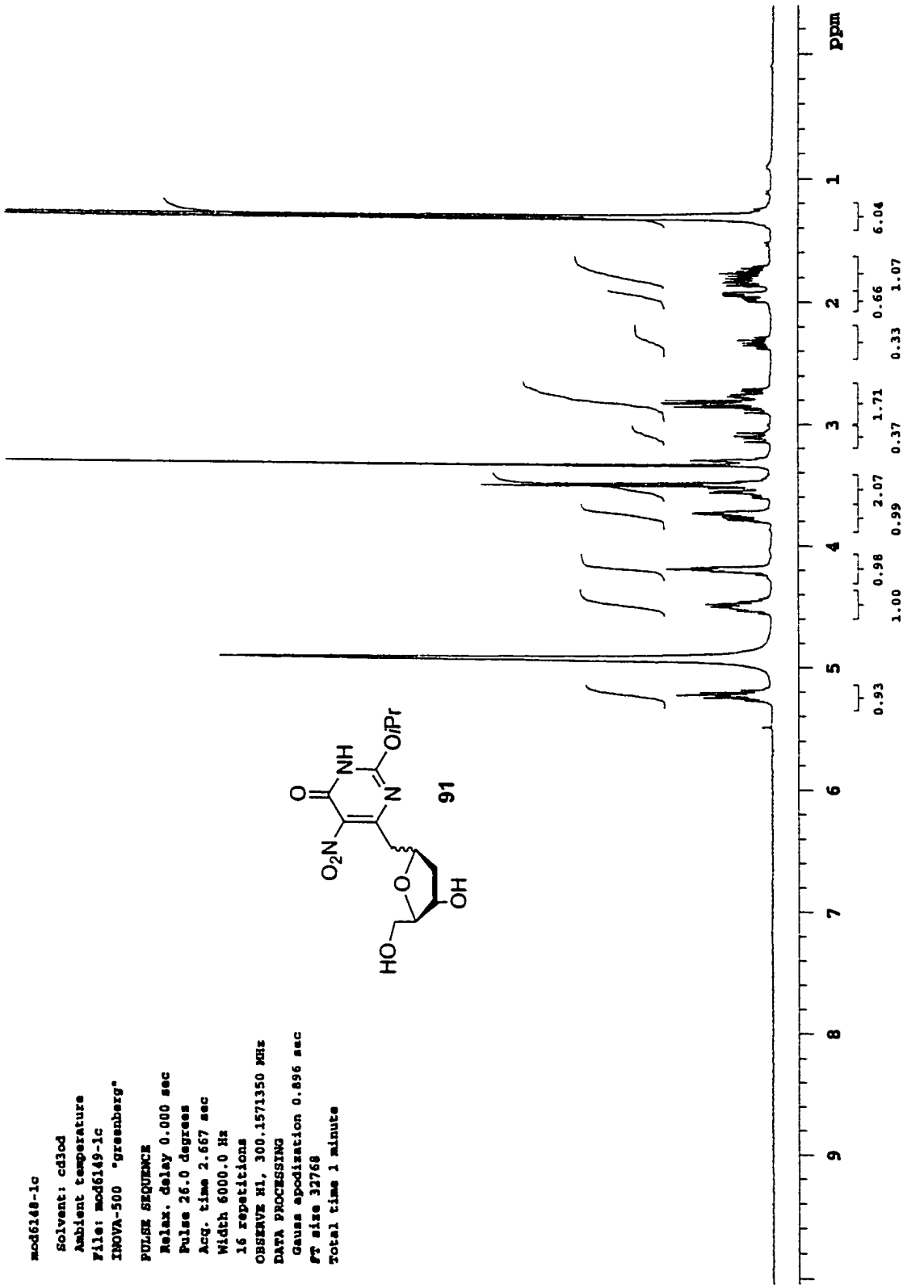


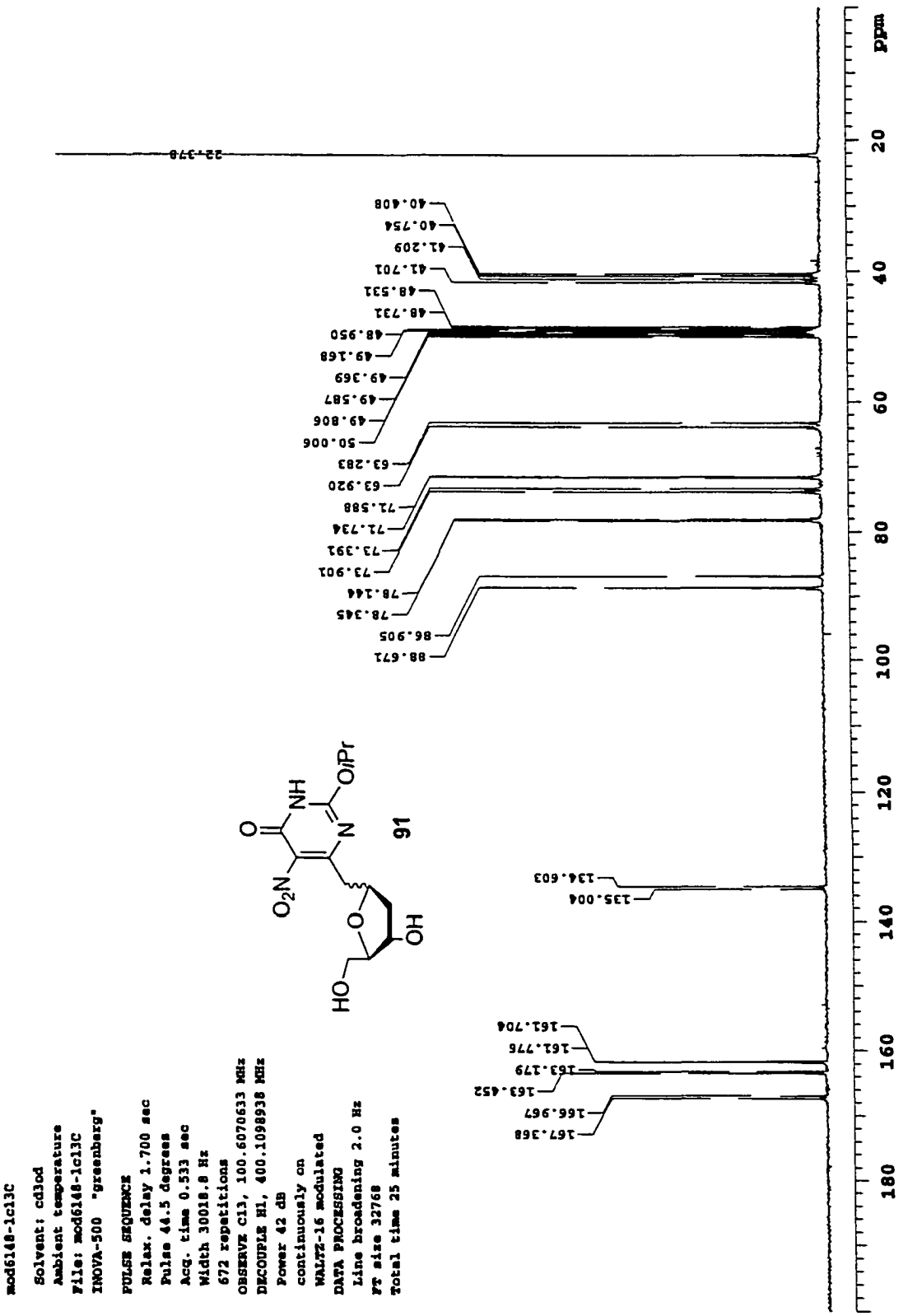


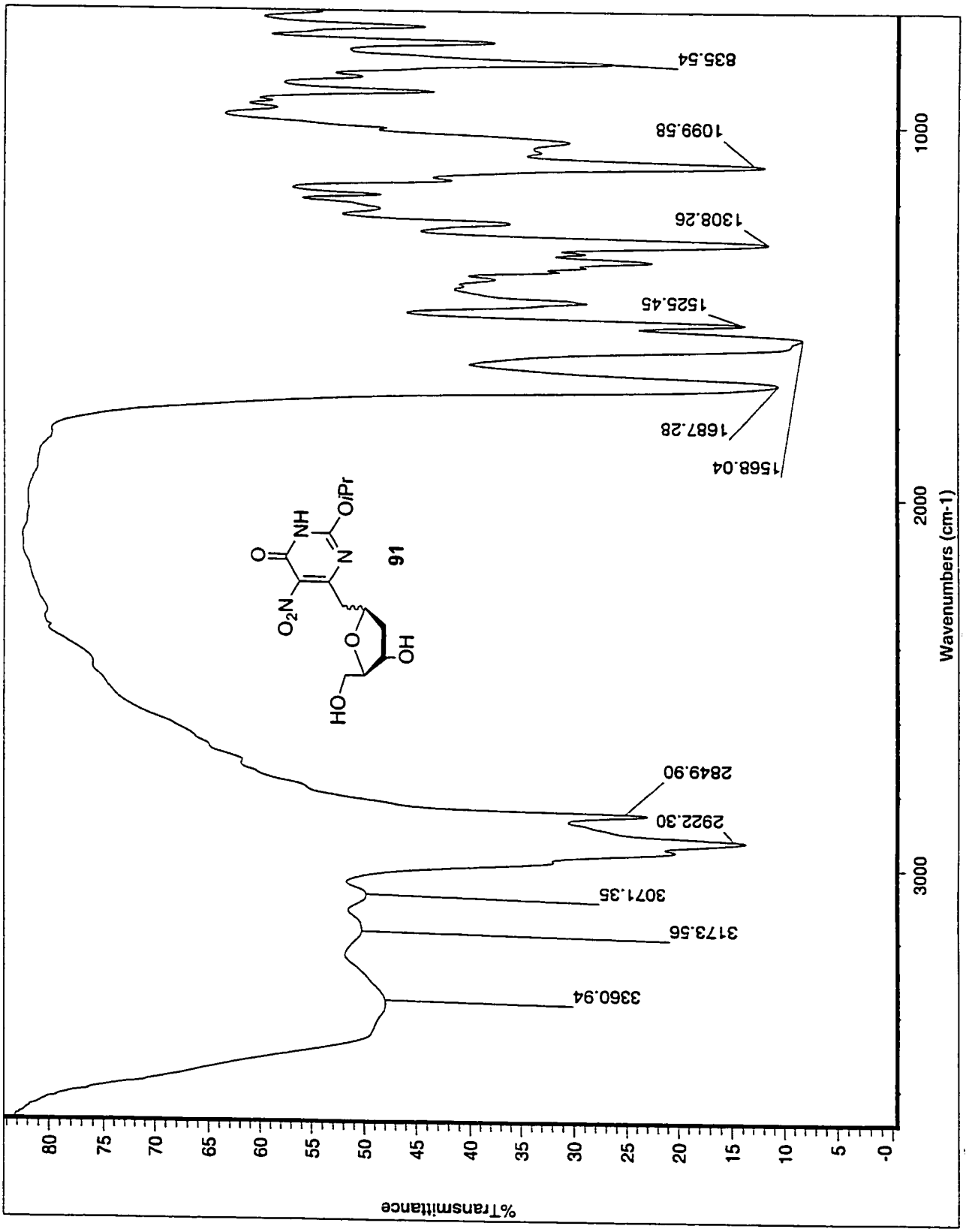


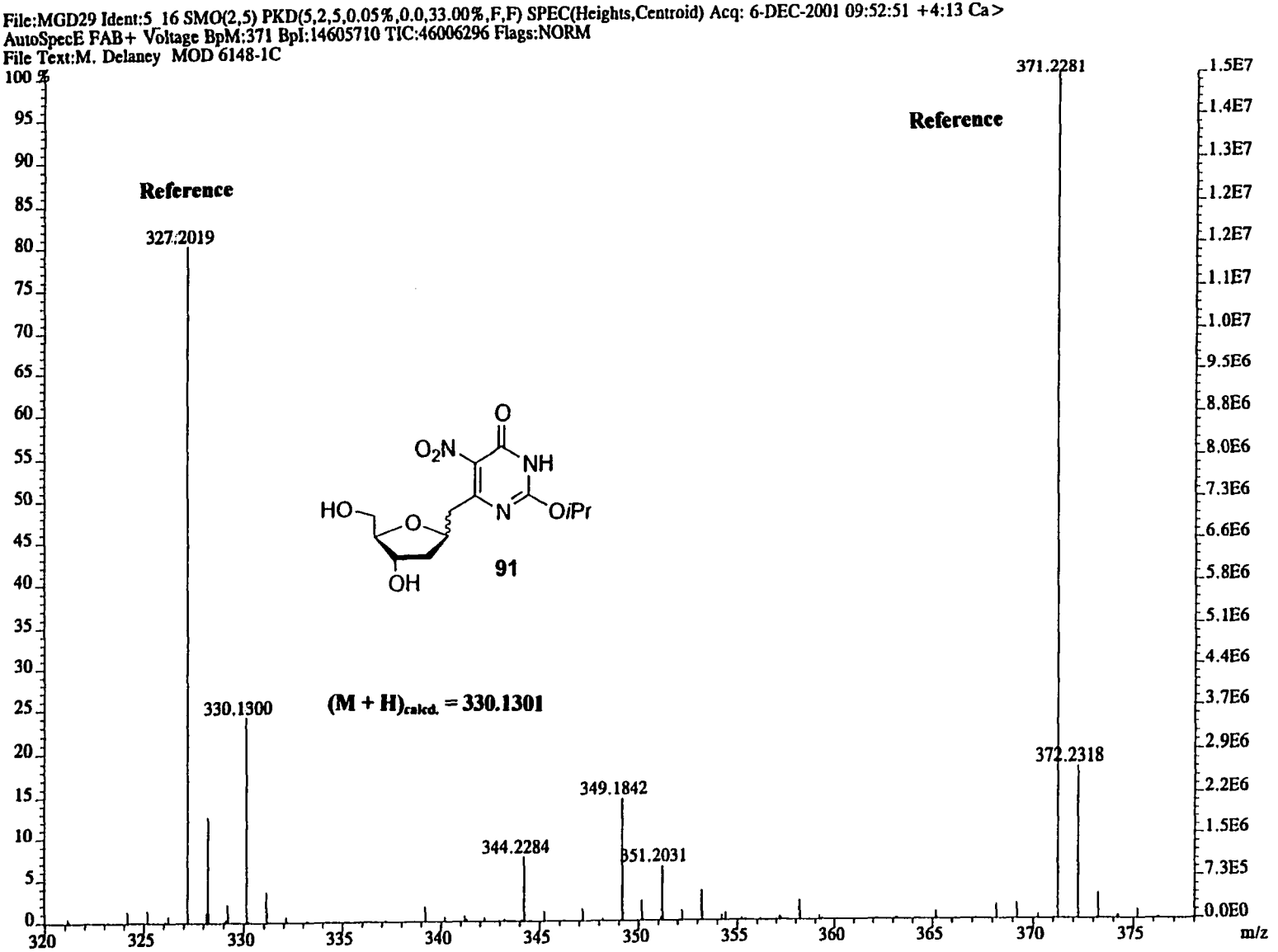


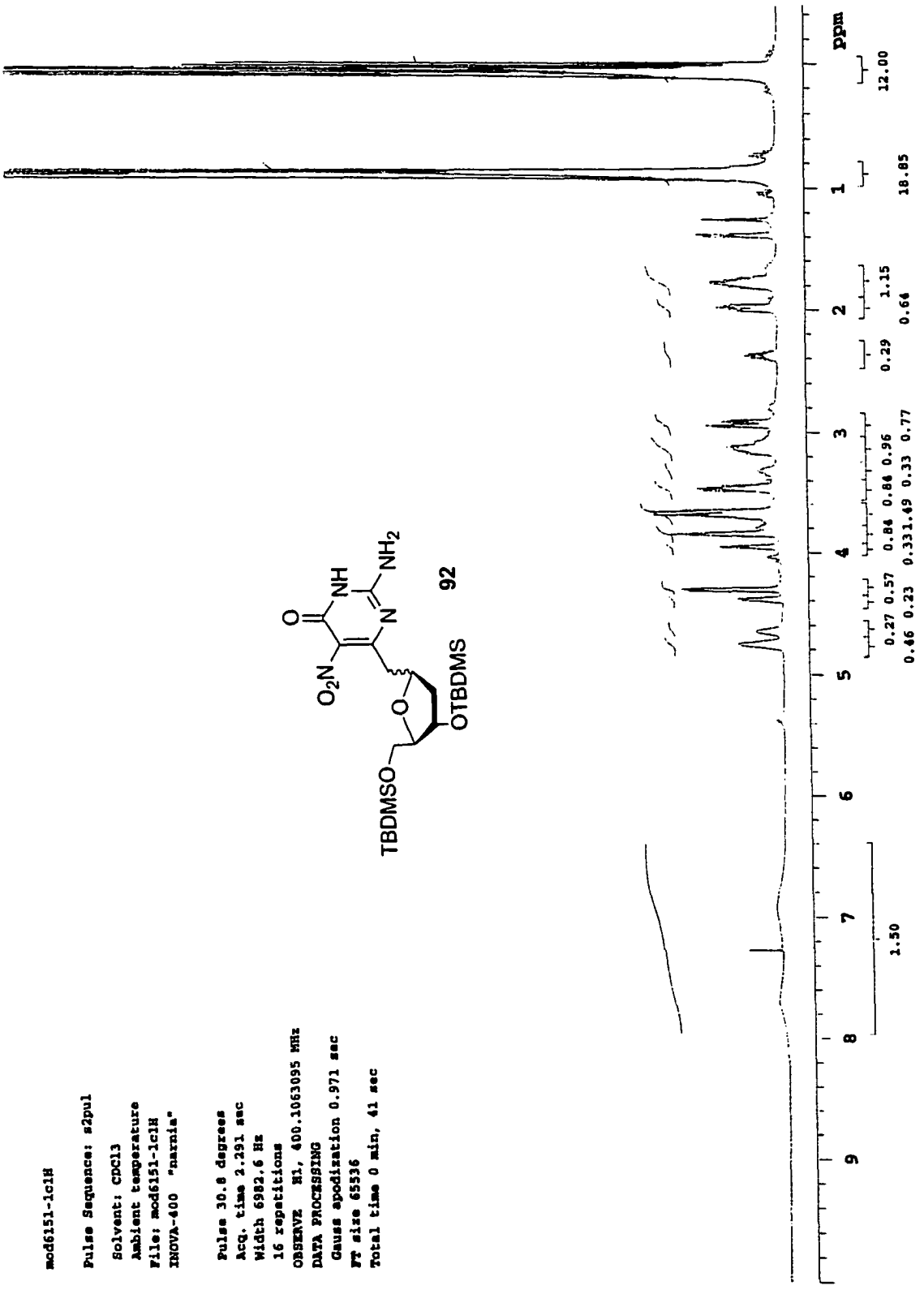


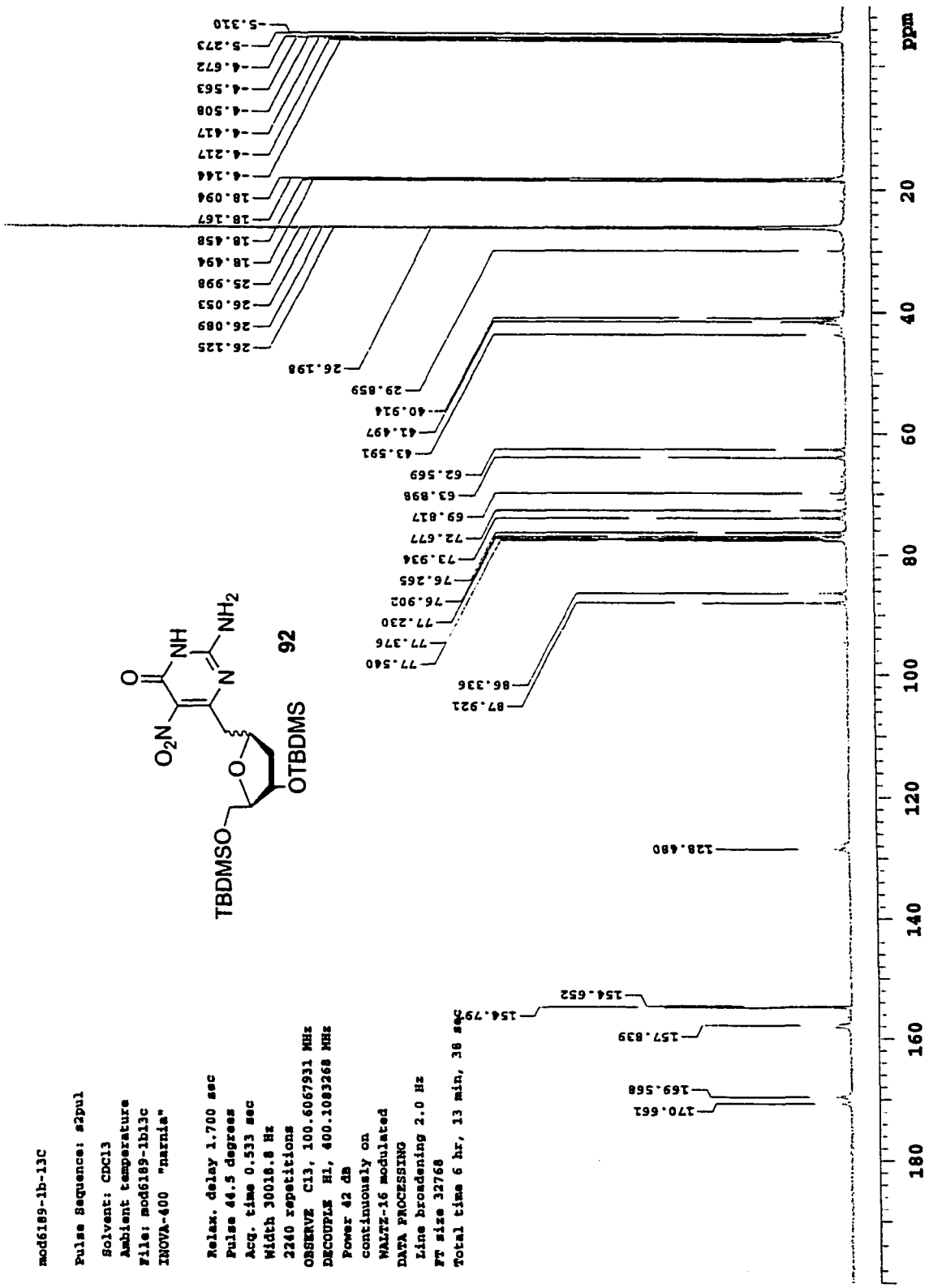


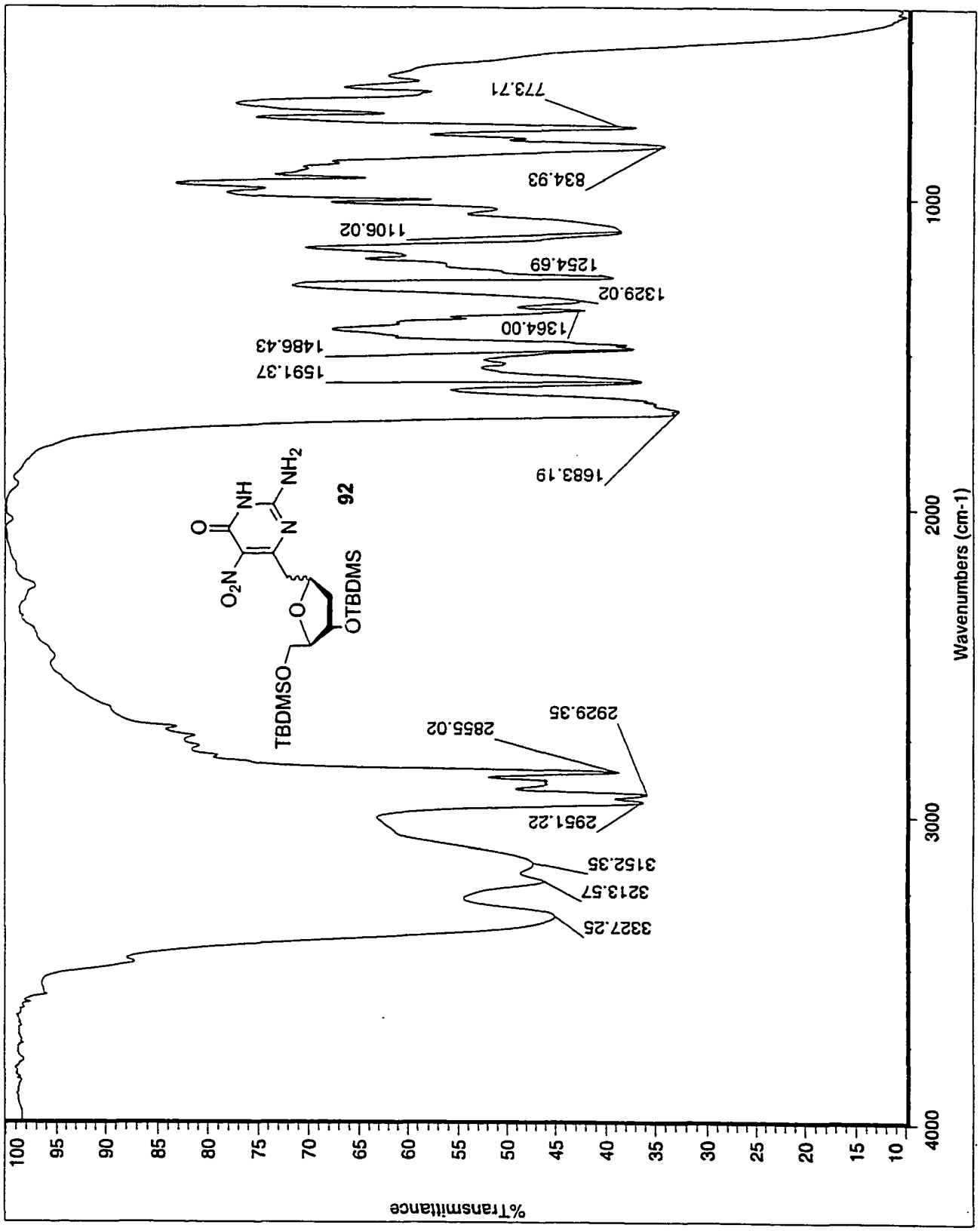


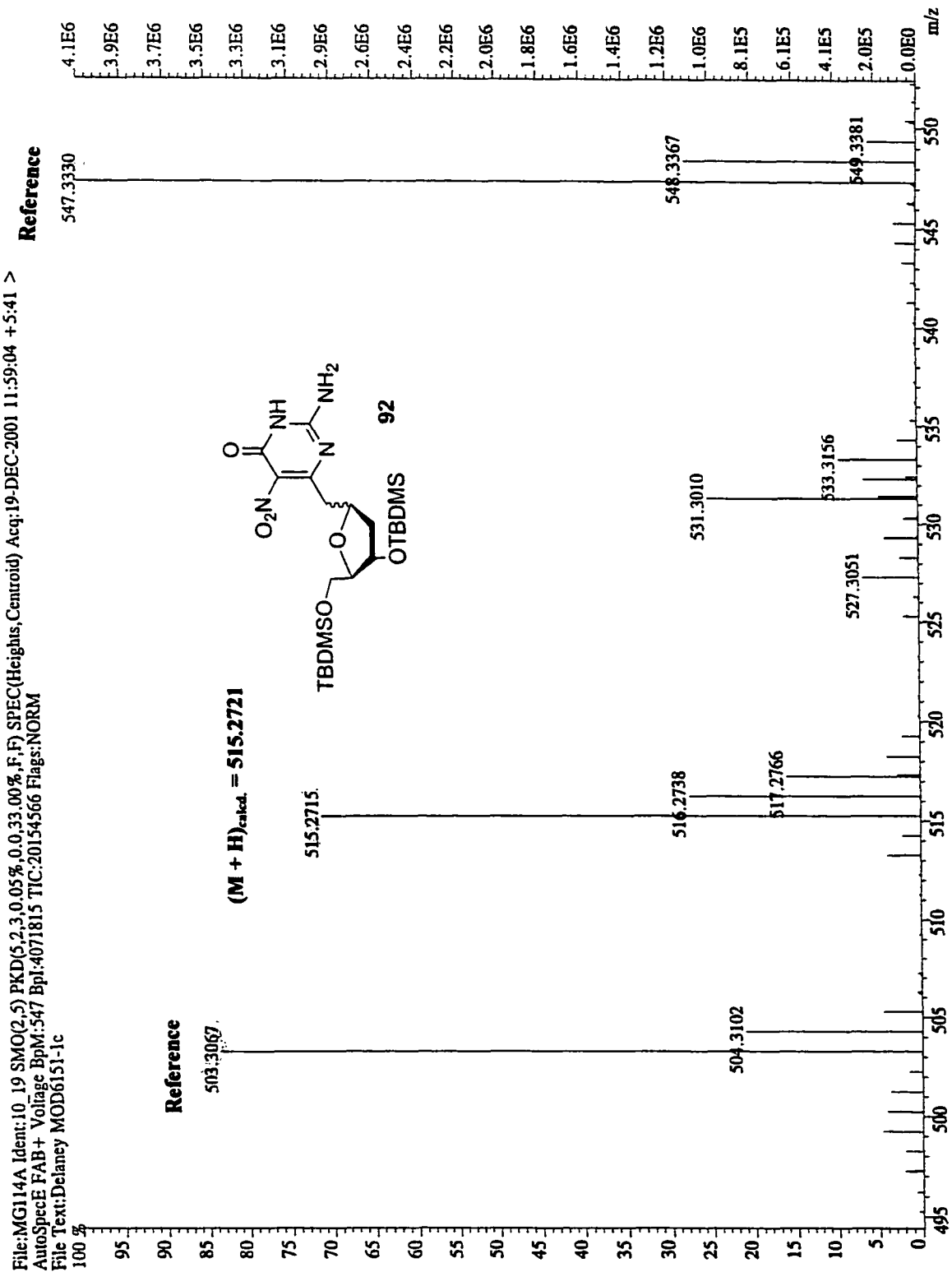


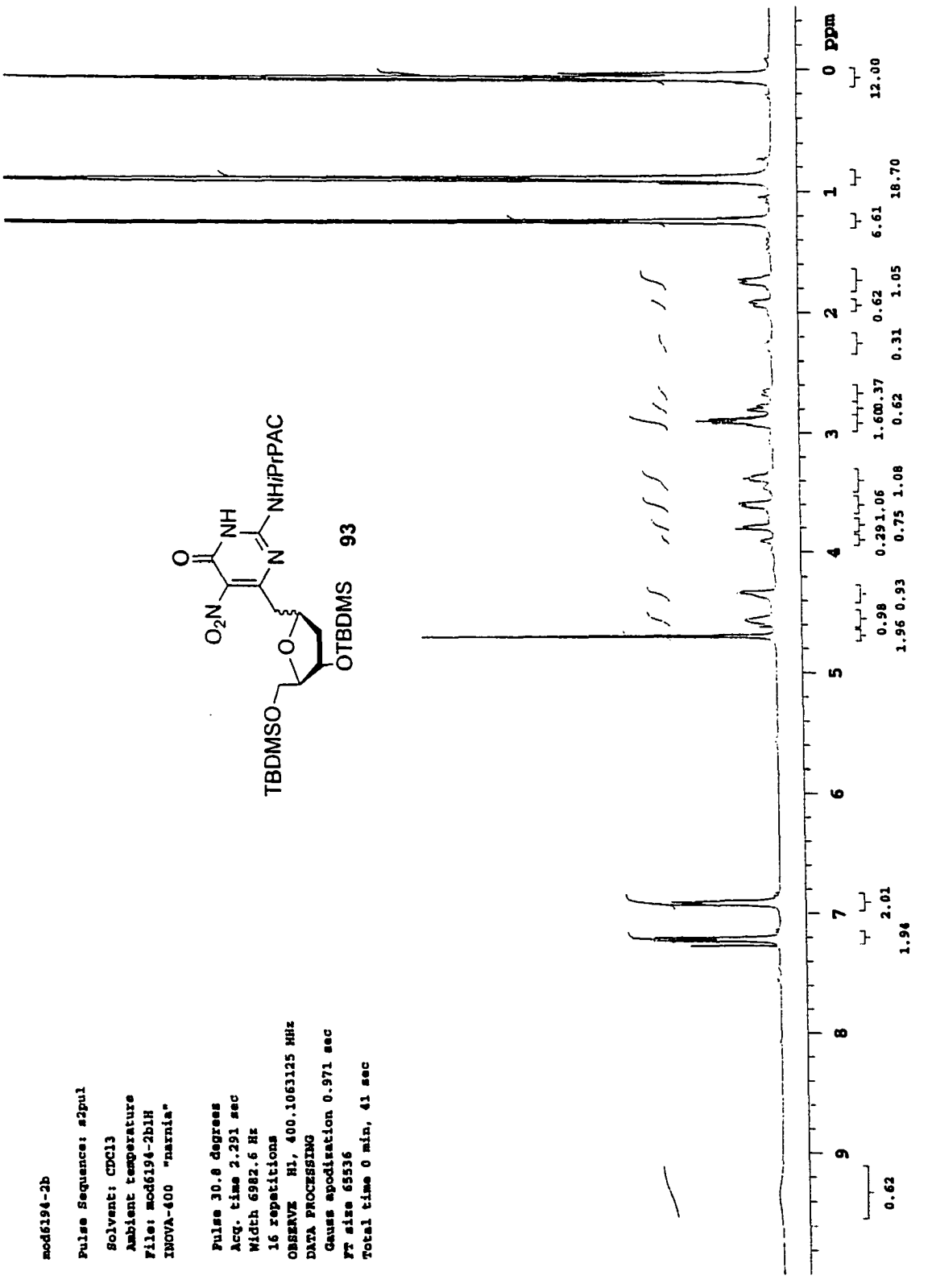


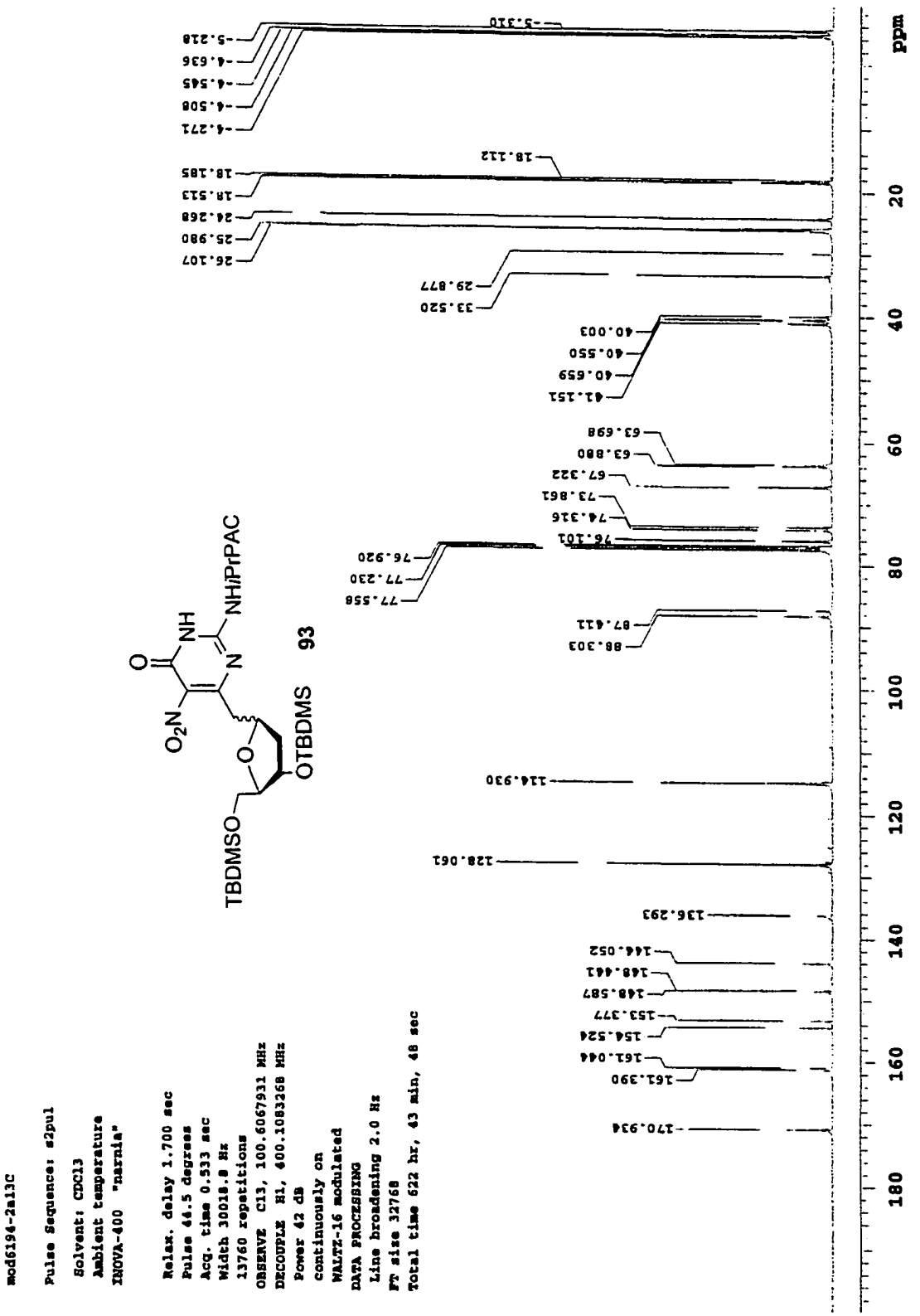


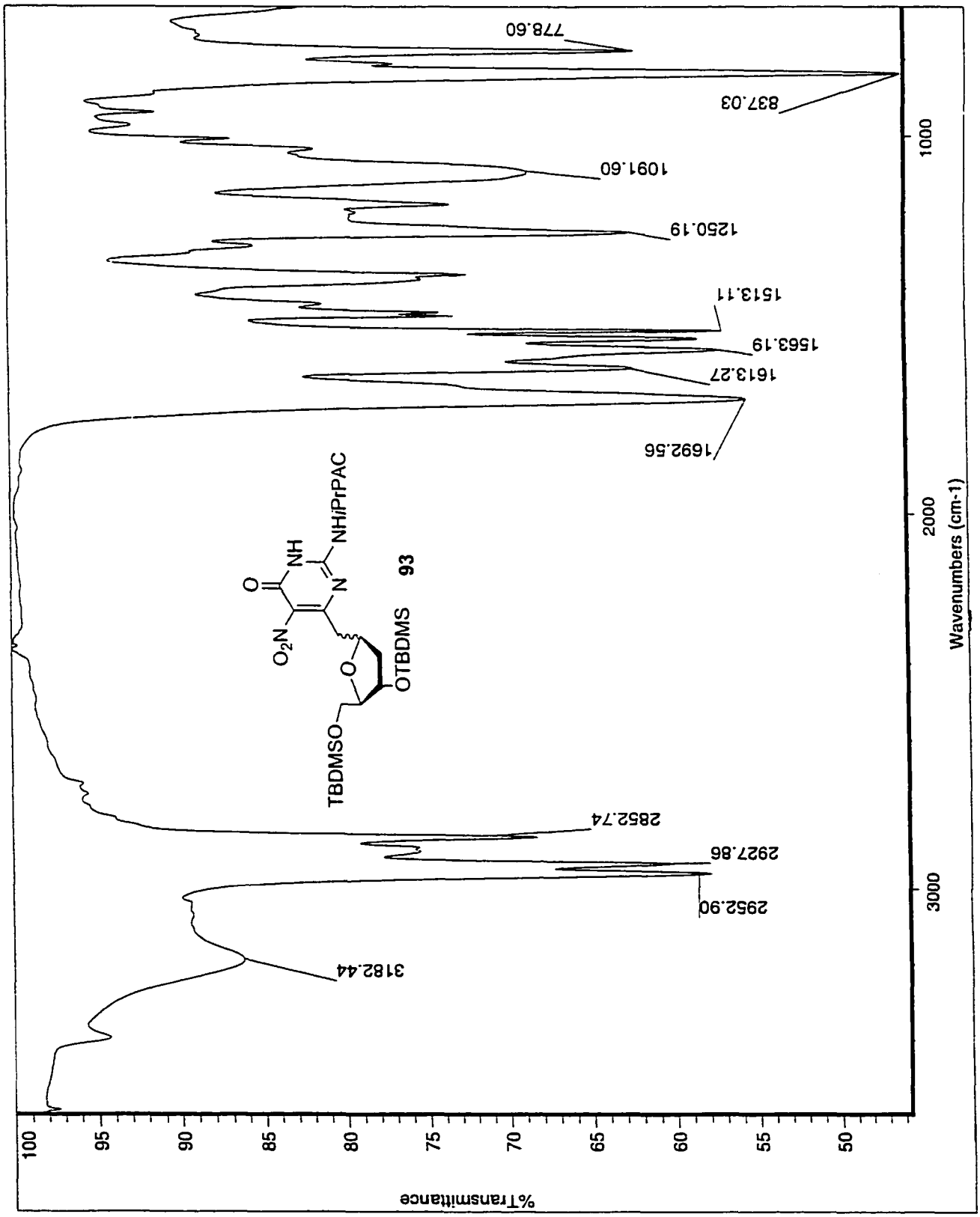


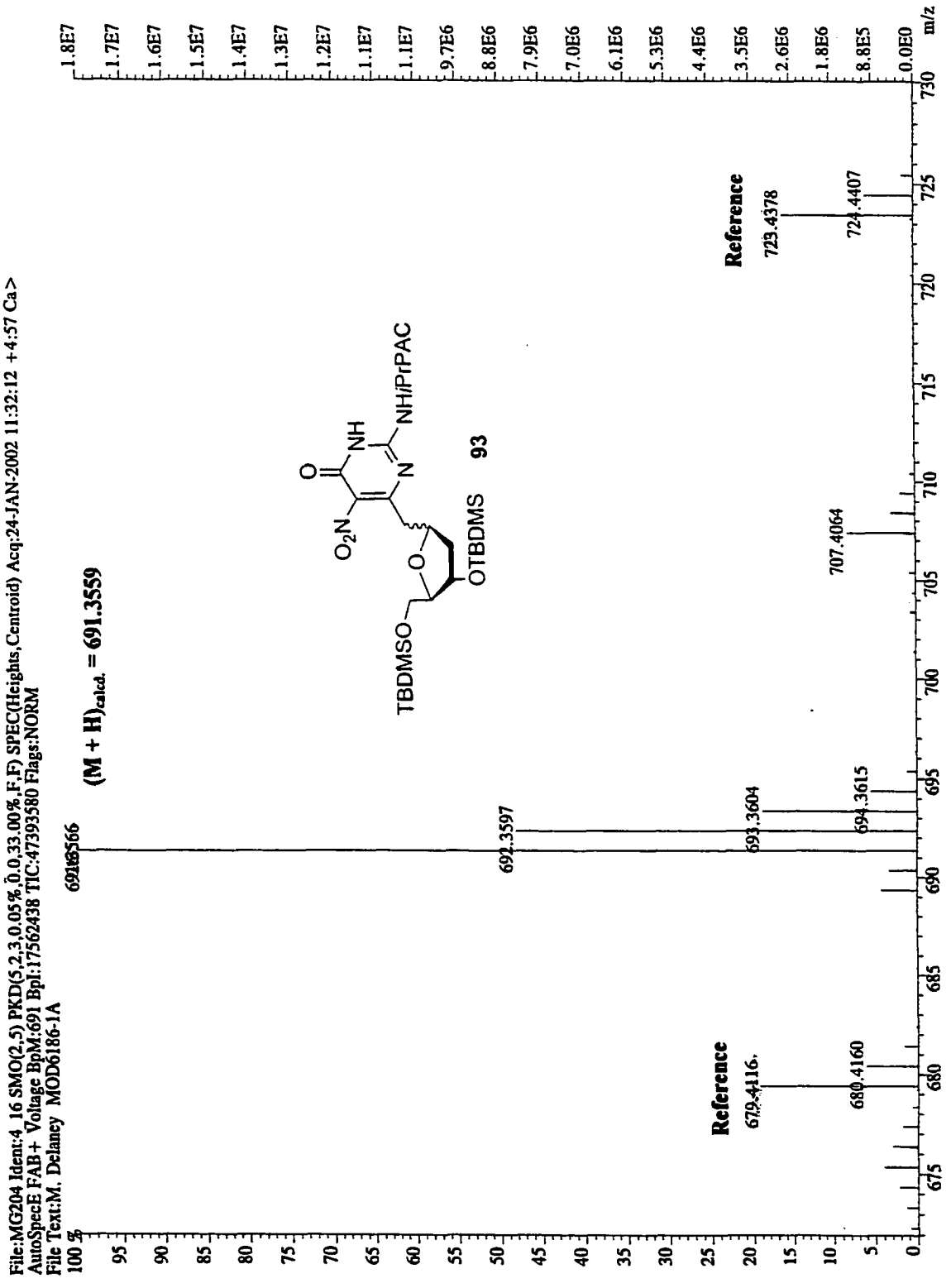


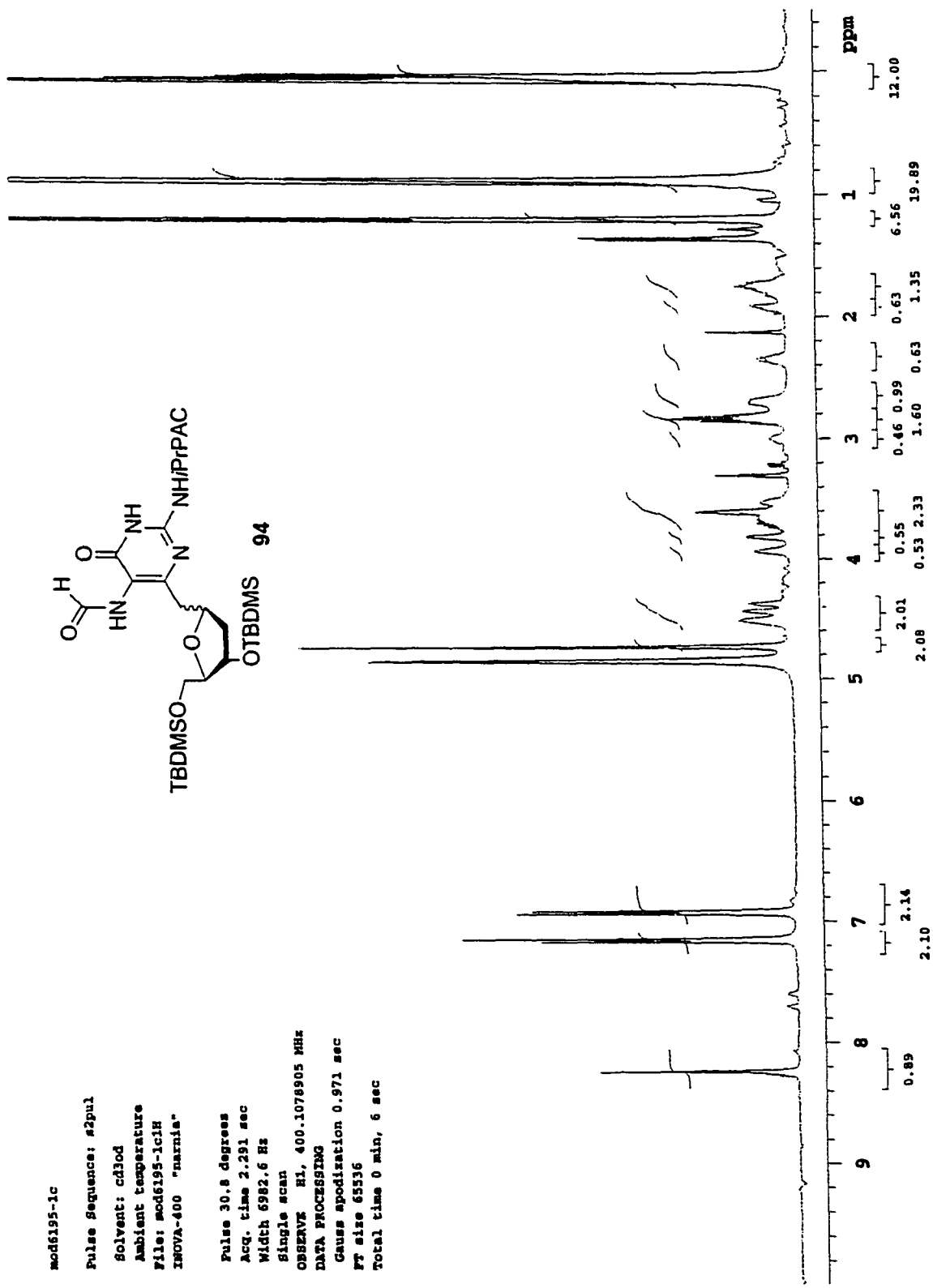


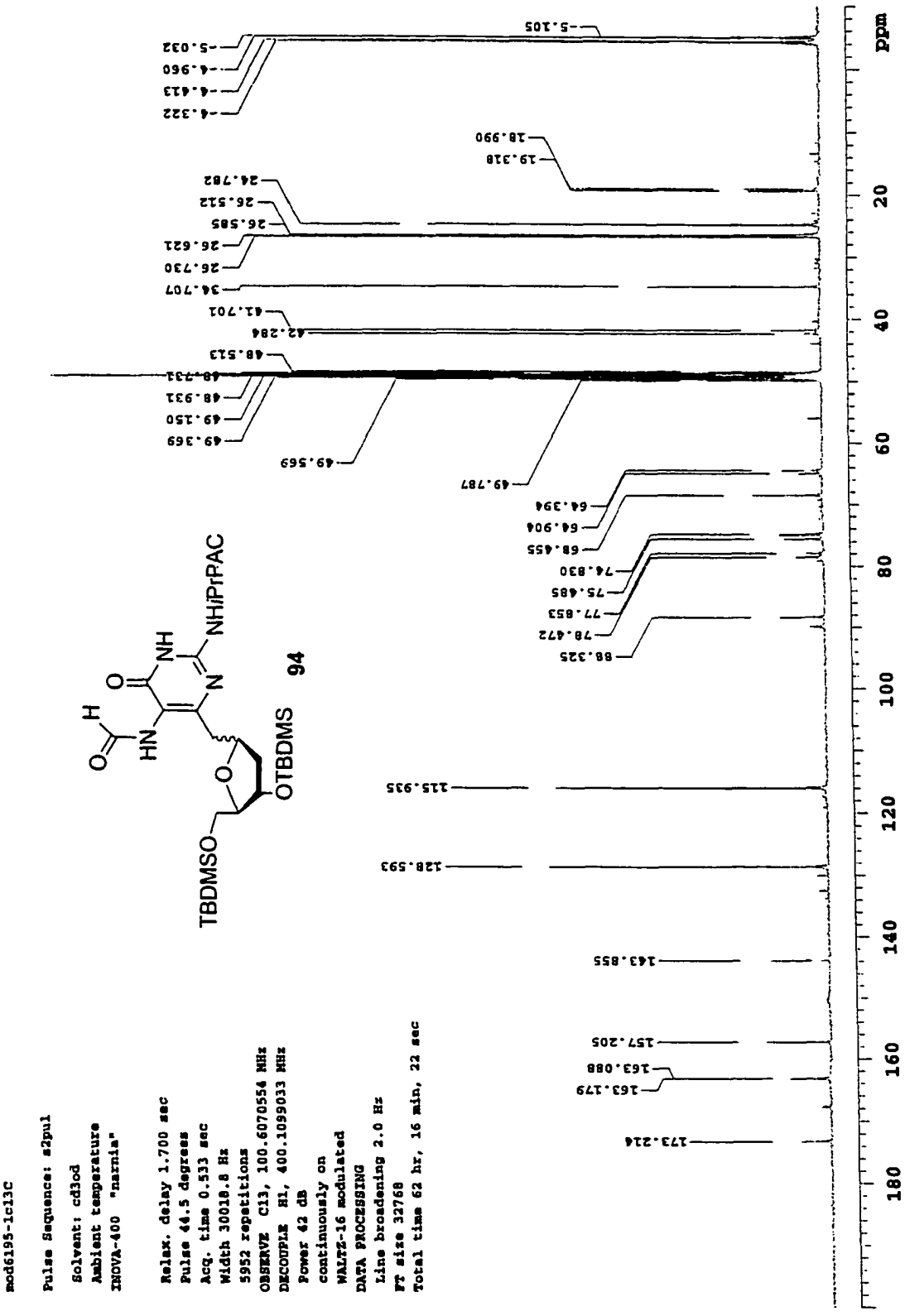


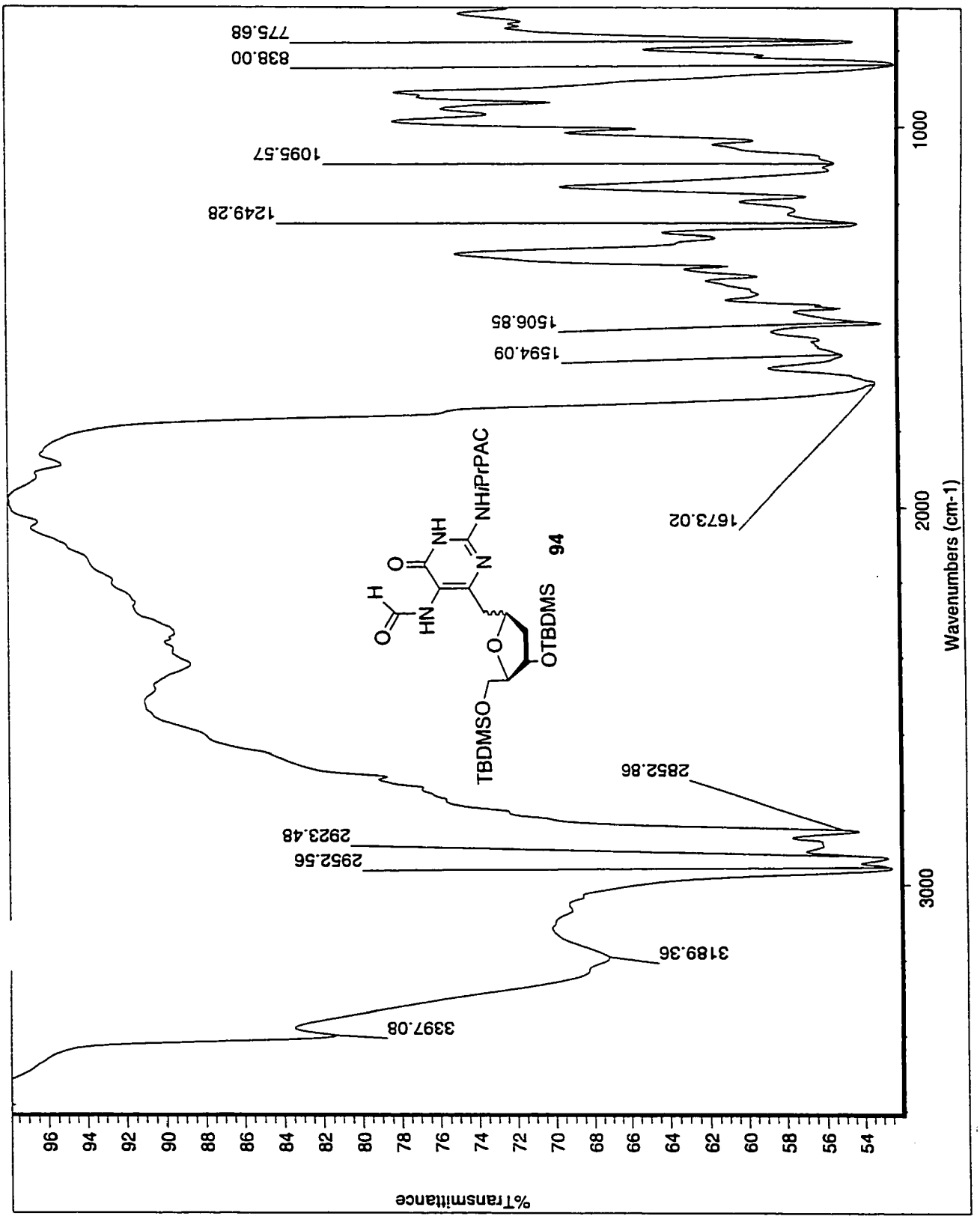


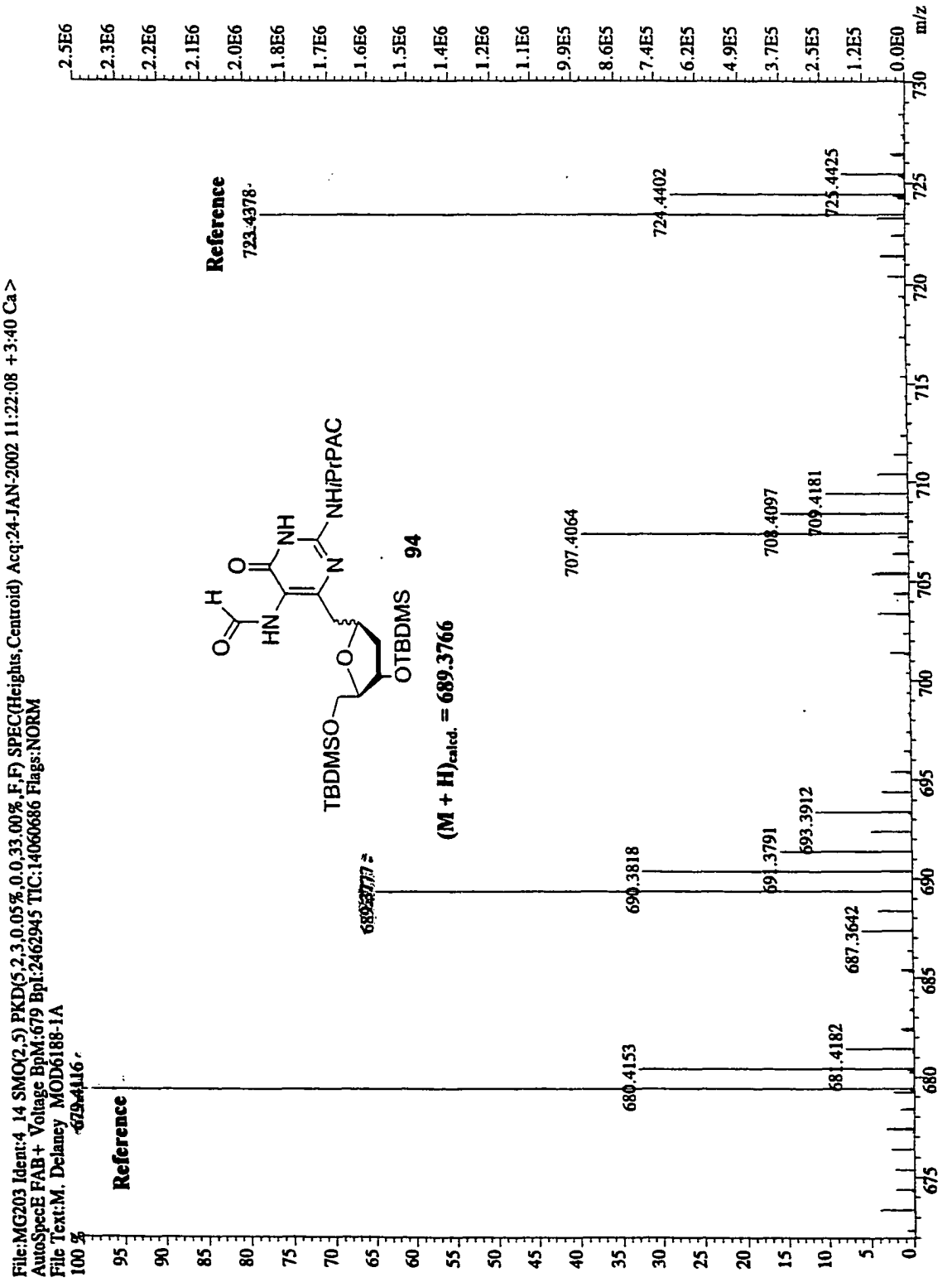








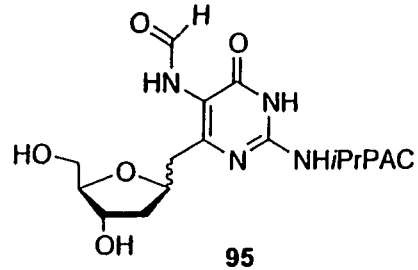




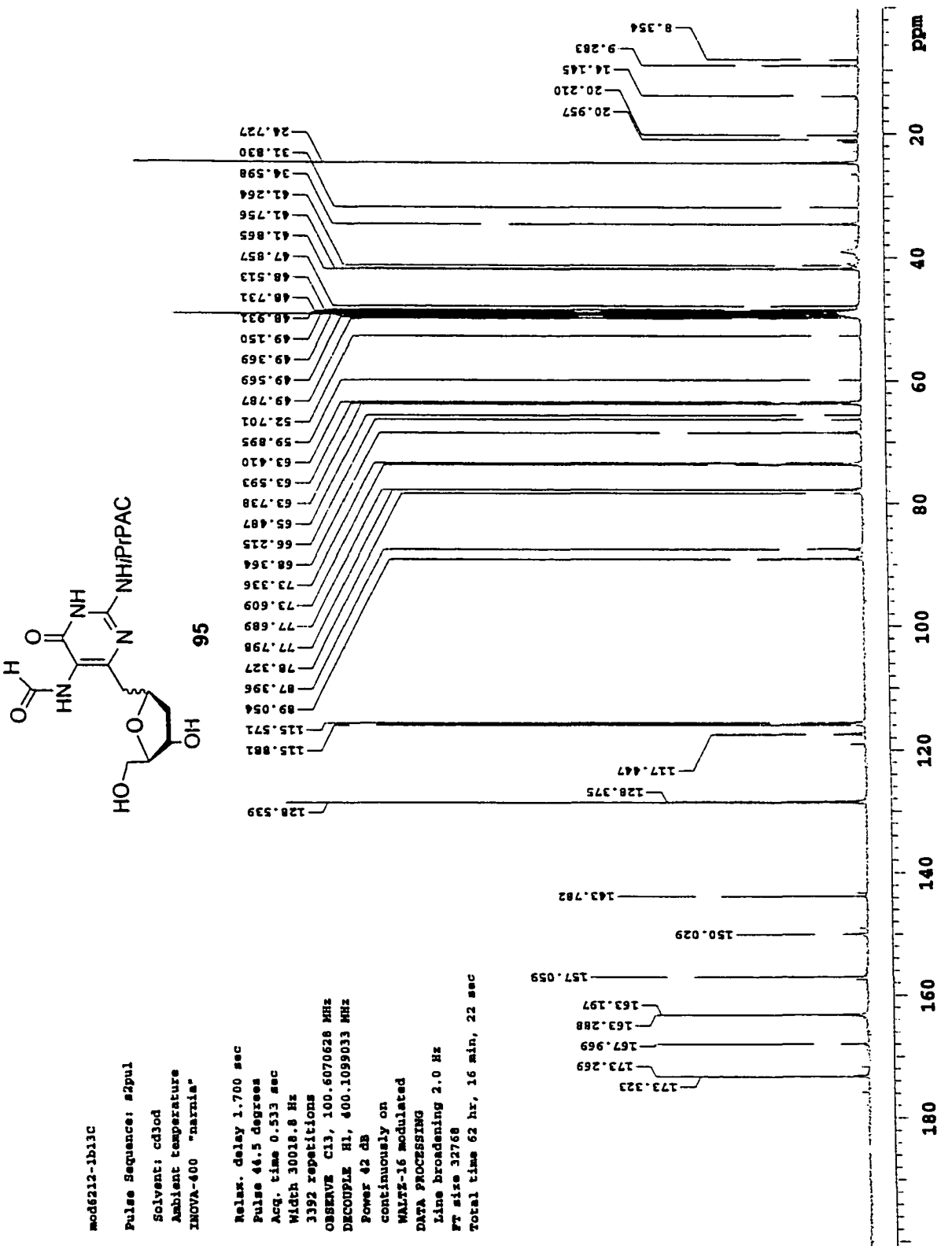
STANDARD 1H OBSERVE

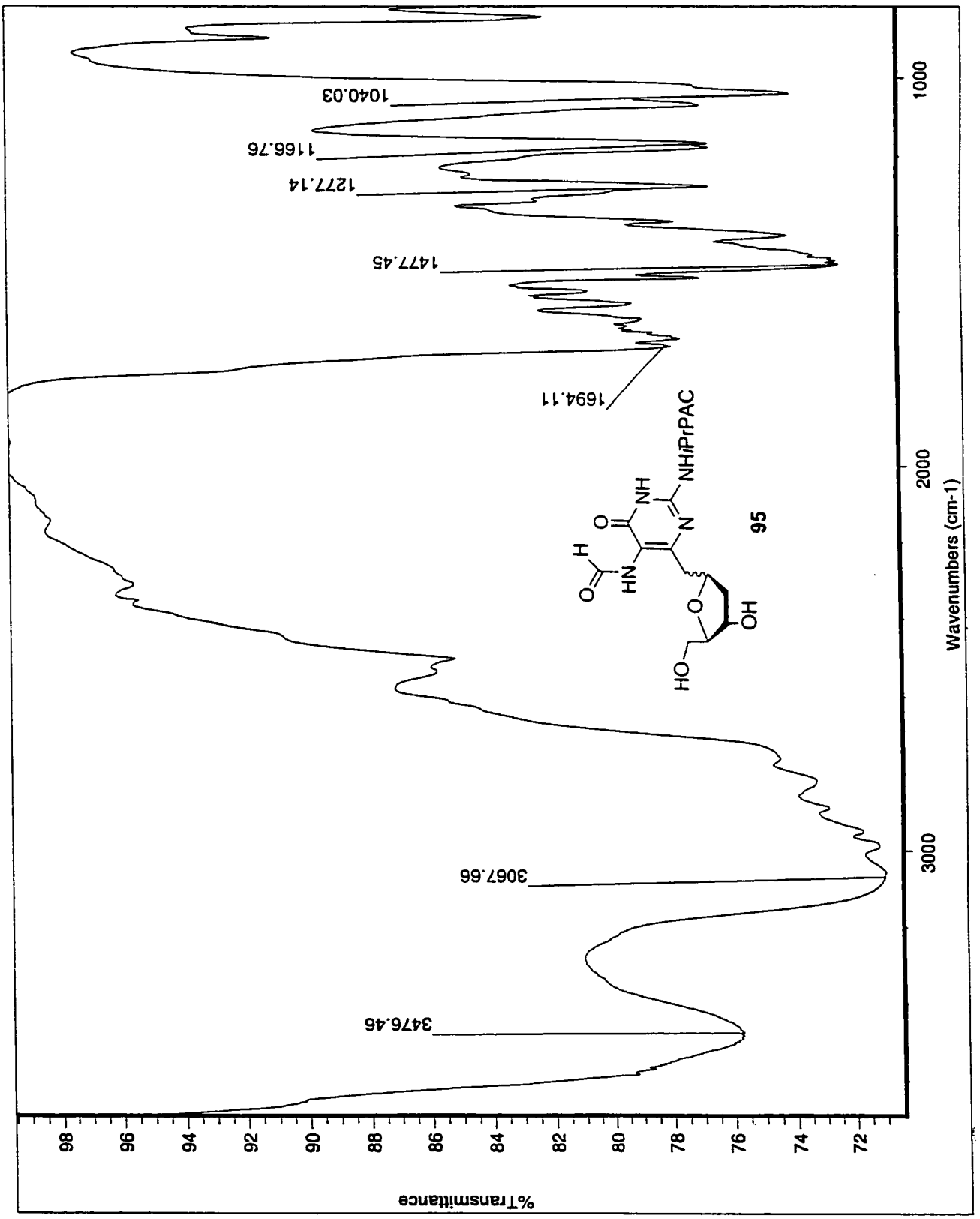
Pulse Sequence: s2pul
Solvent: cd3od
Ambient temperature
File: mod6212-1b1H
INOVA-400 "narnia"

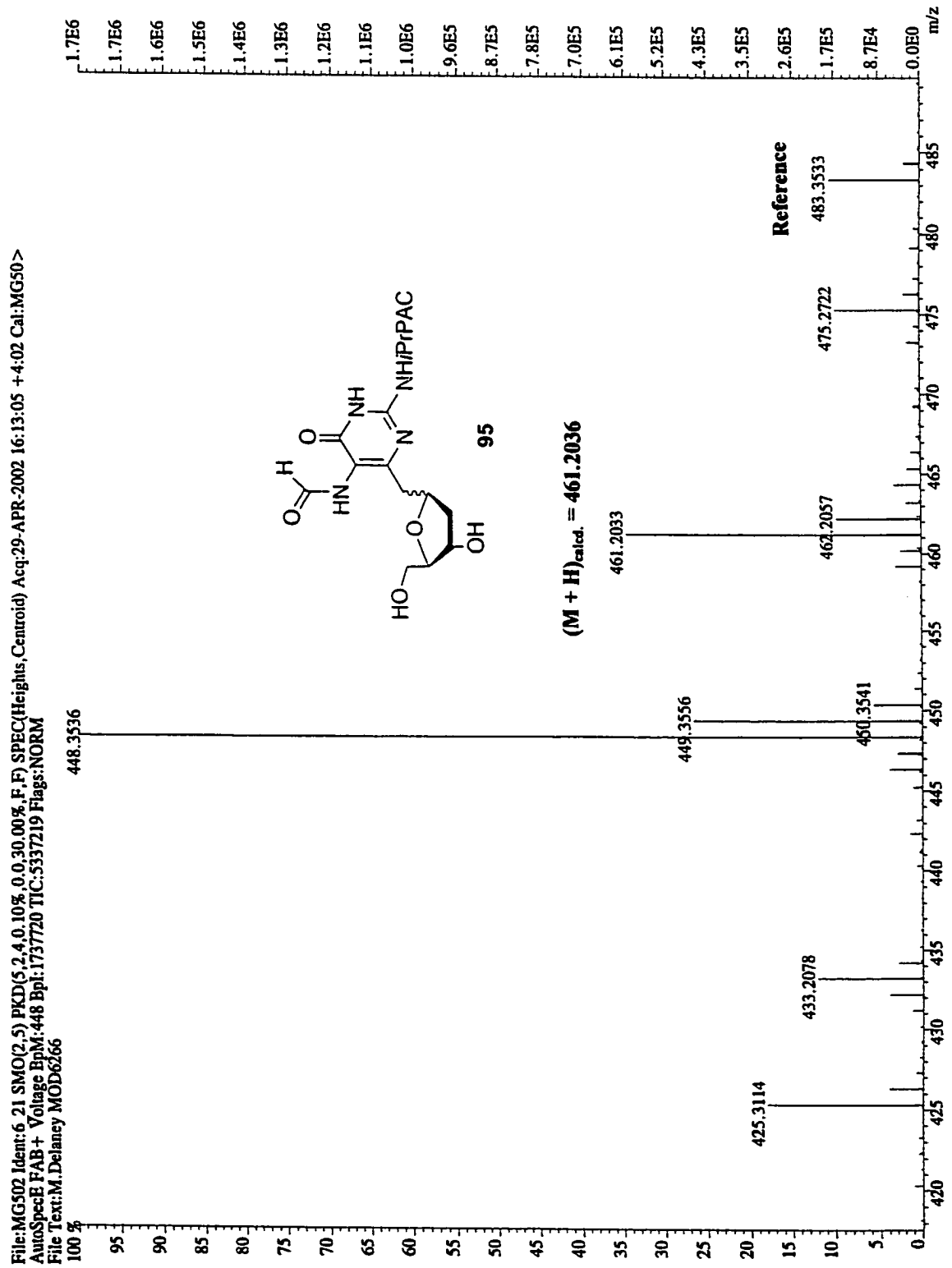
Pulse 30.8 degrees
Acq. time 2.291 sec
Width 6982.6 Hz
Single scan
OBSERVE H1, 400.1078910 MHz
DATA PROCESSING
Gauss apodization 0.971 sec
FT size 65536
Total time 0 min, 6 sec

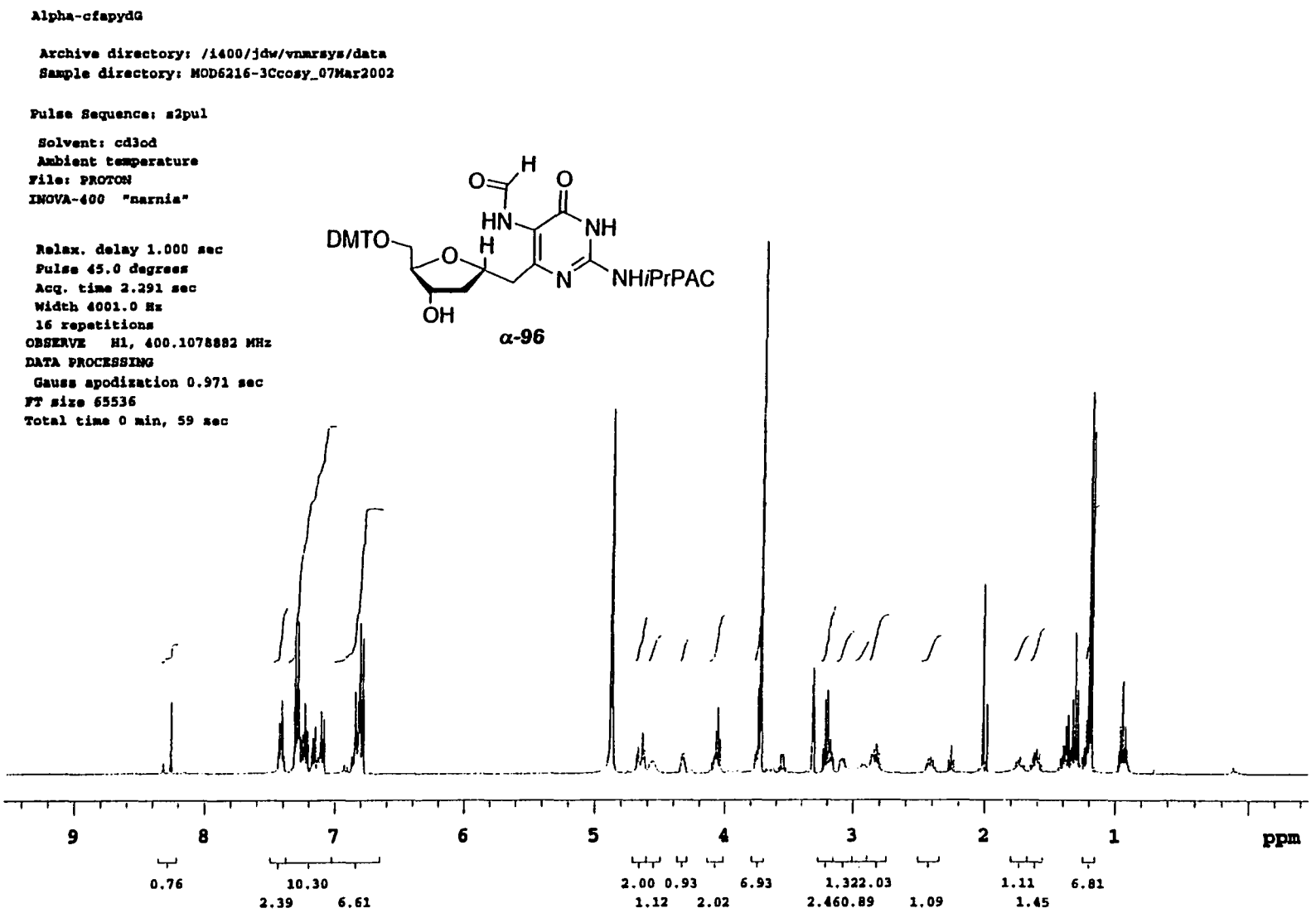


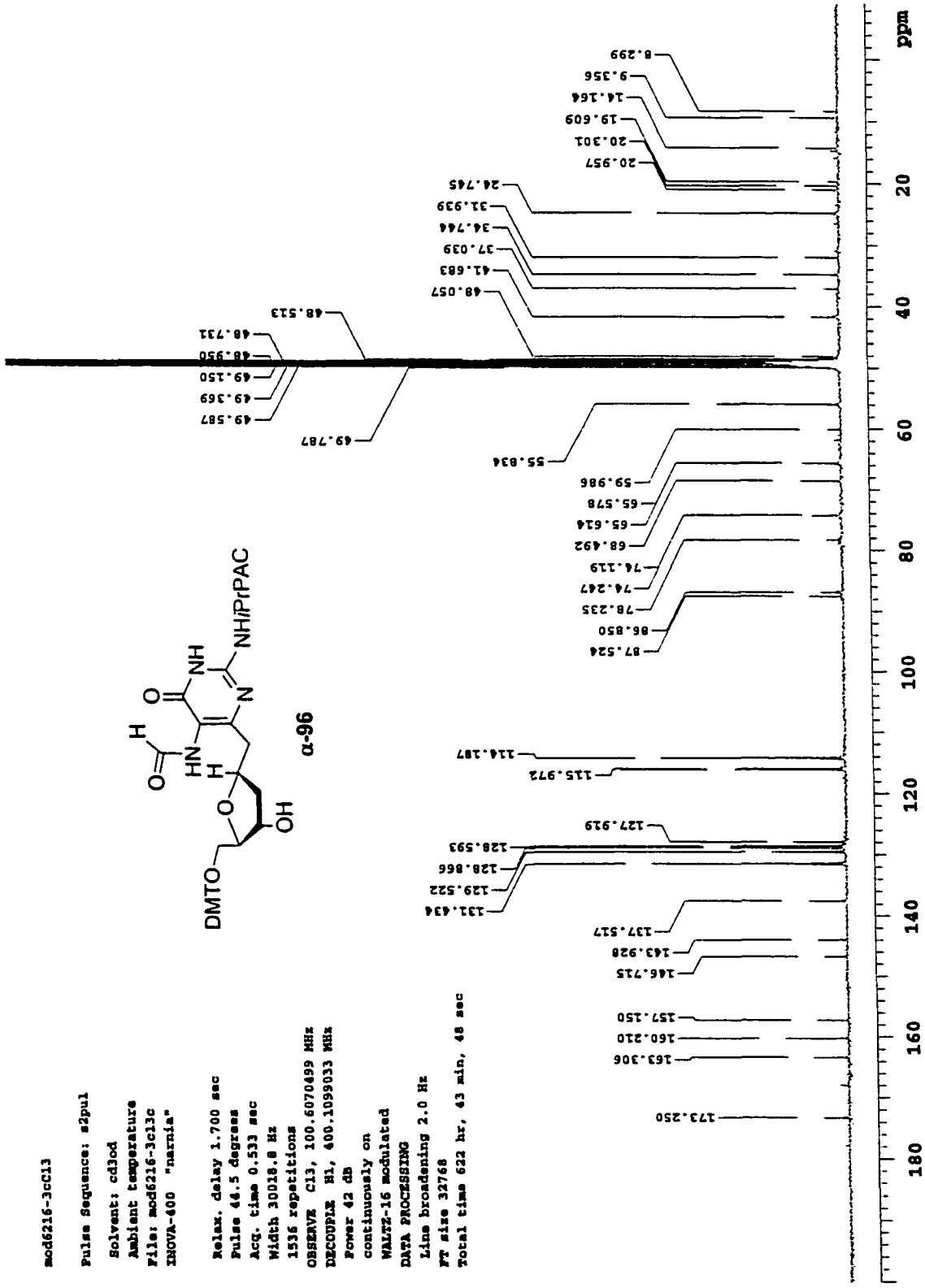
95

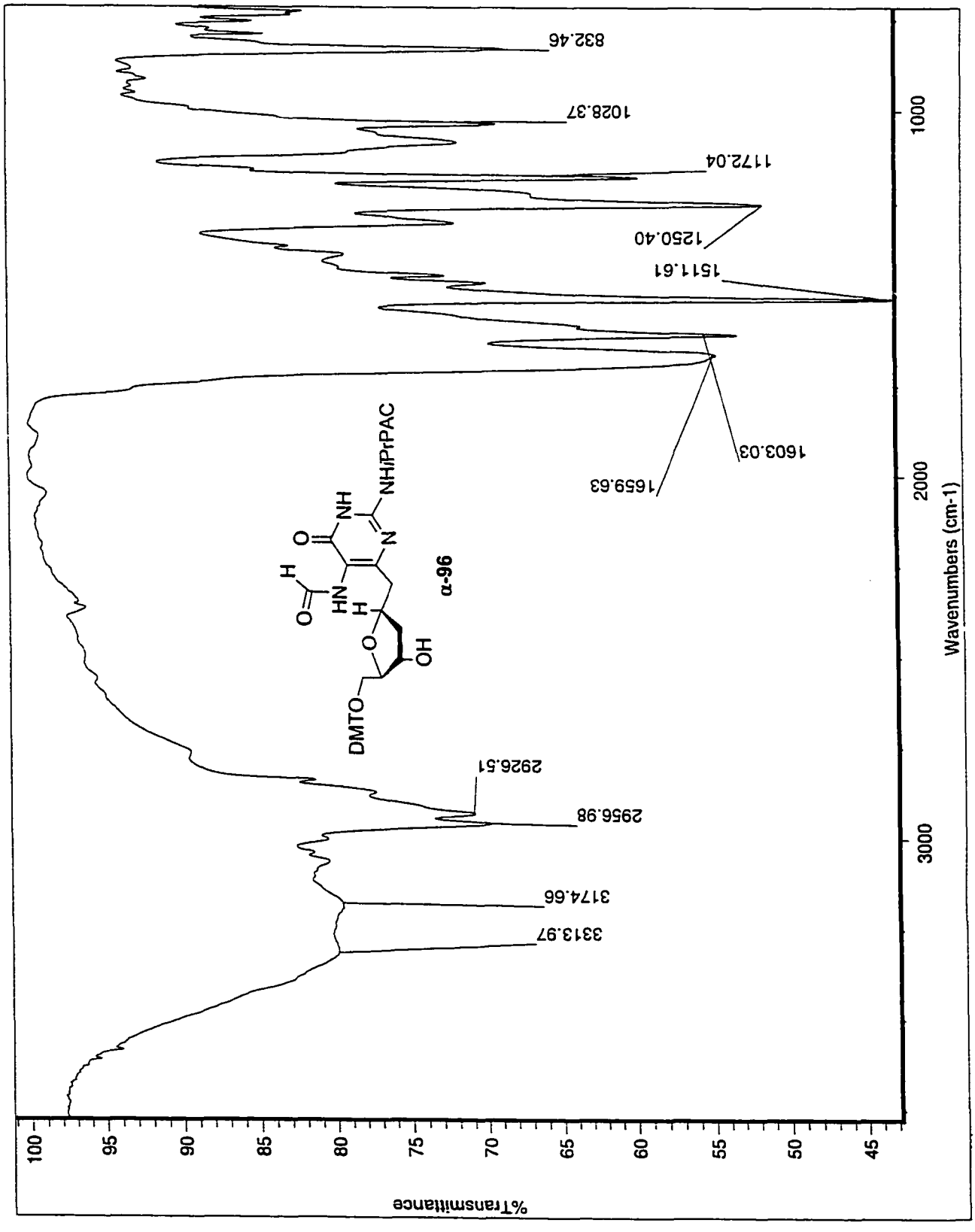


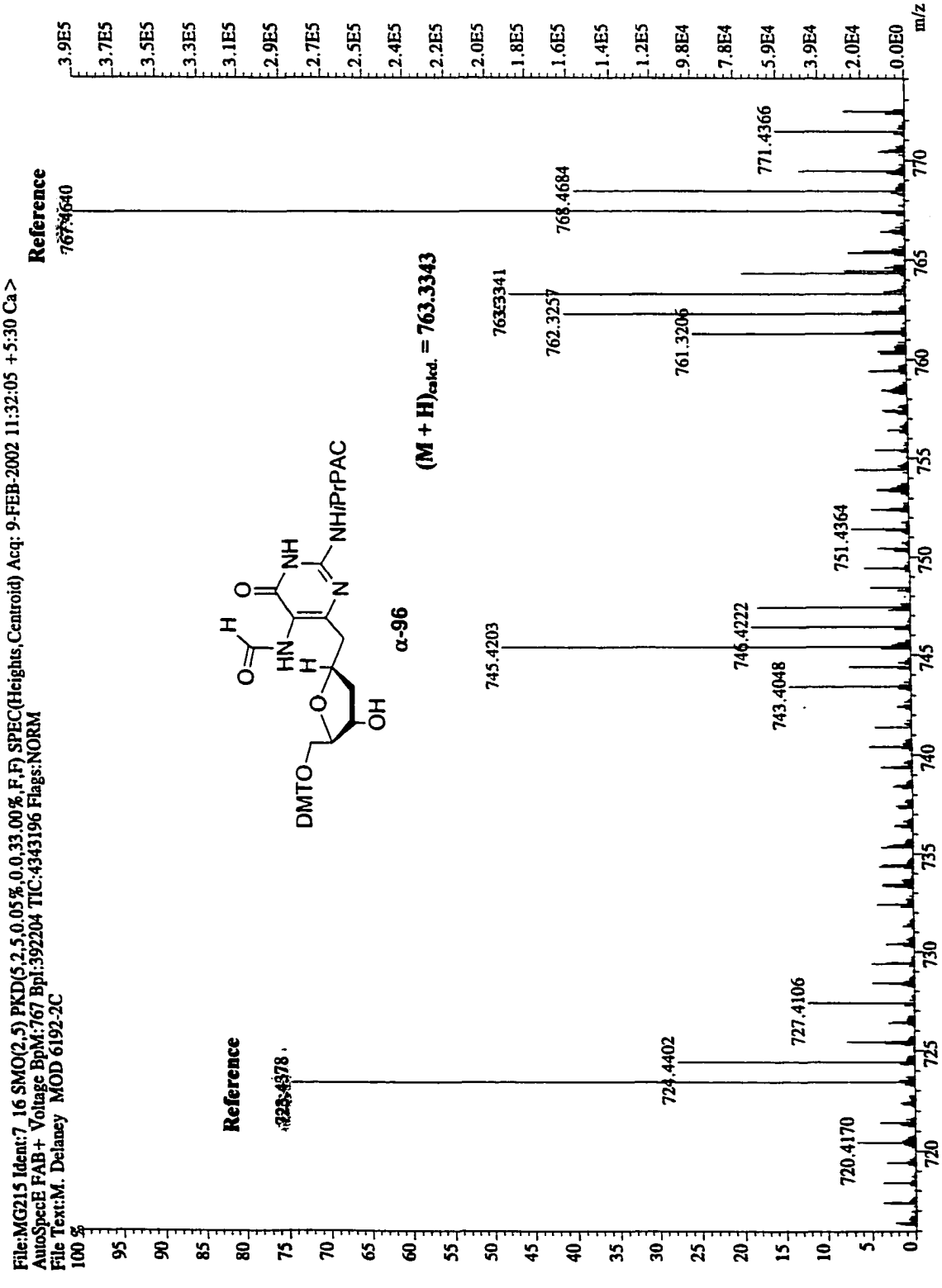


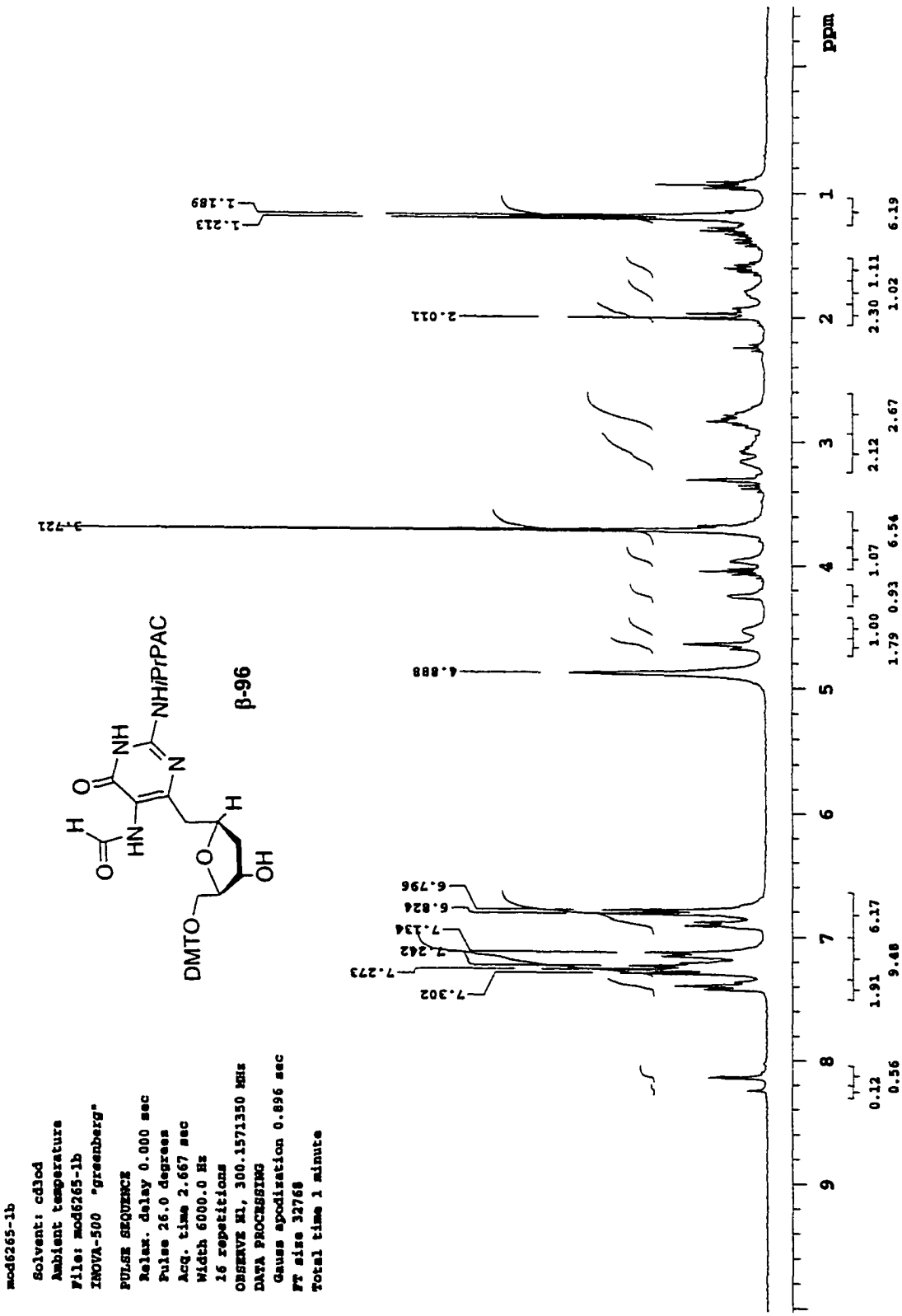


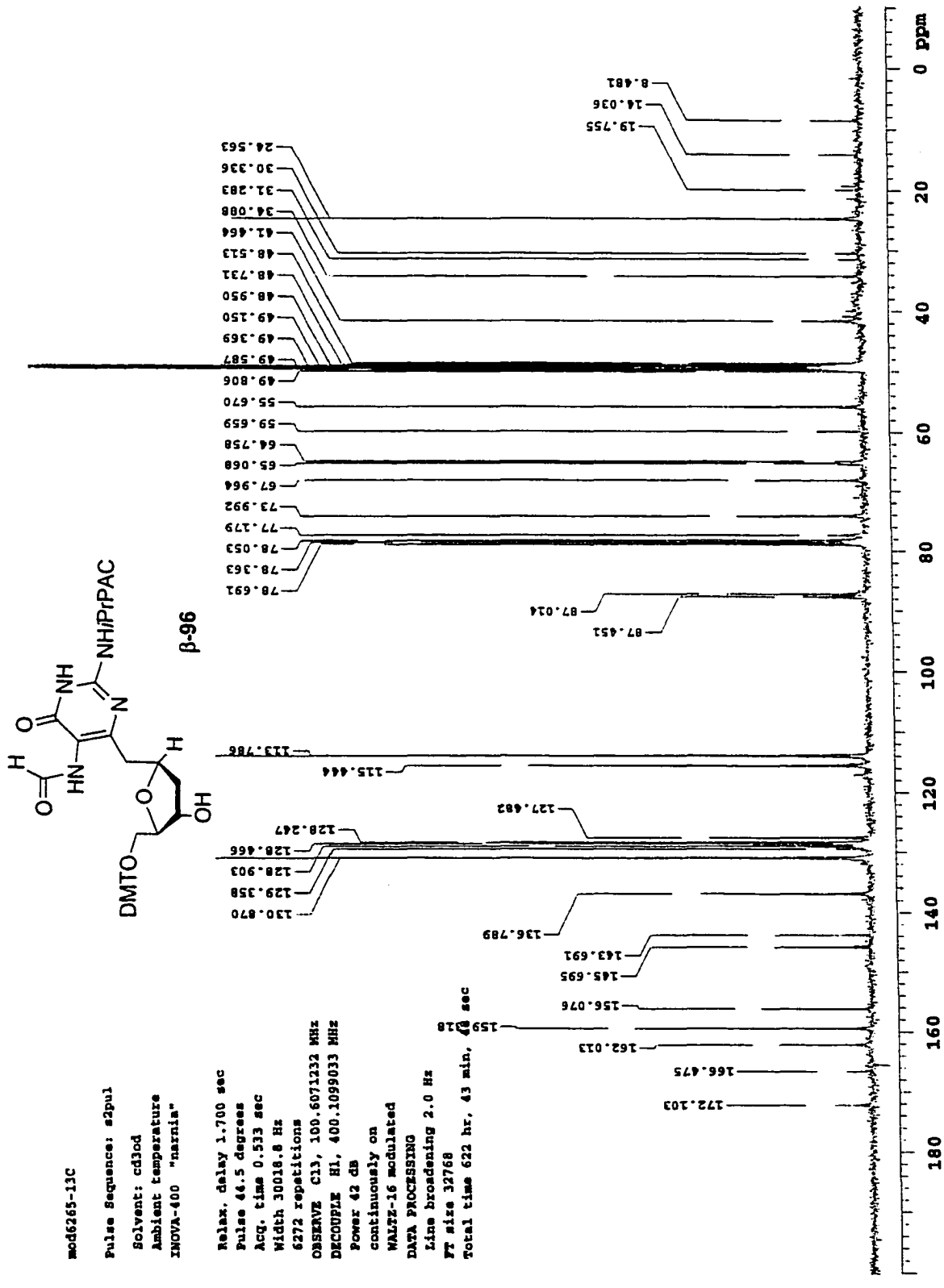


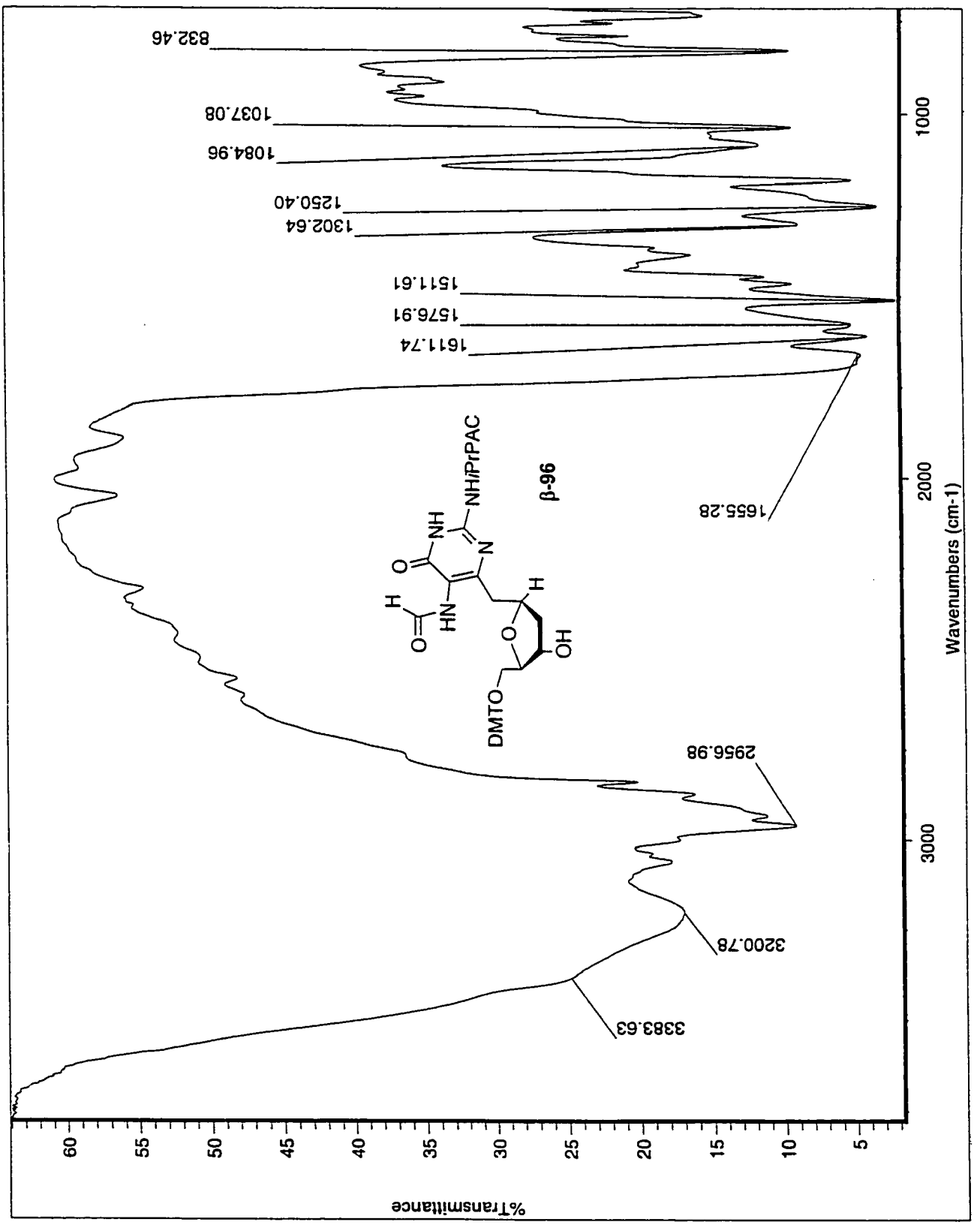


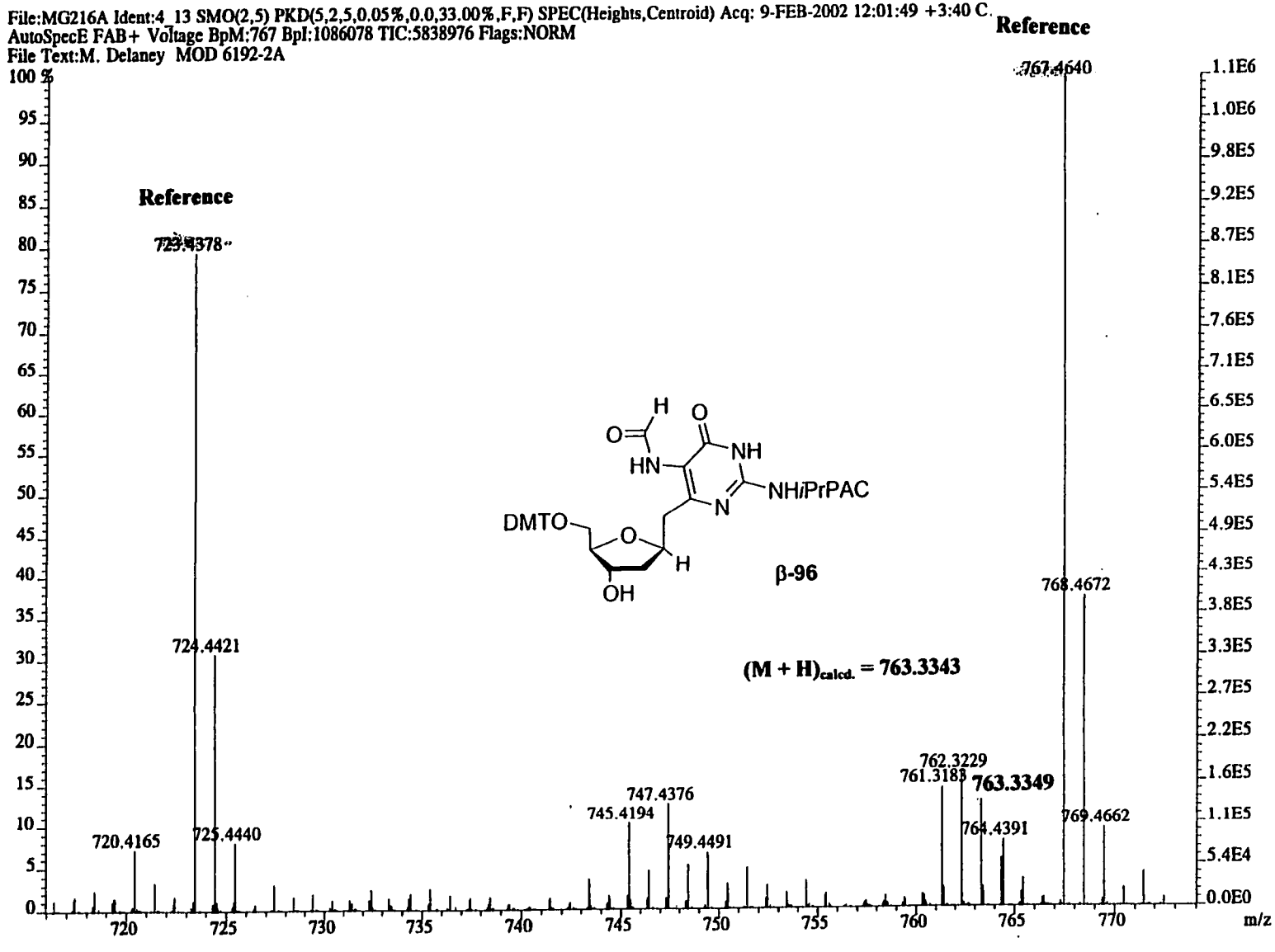


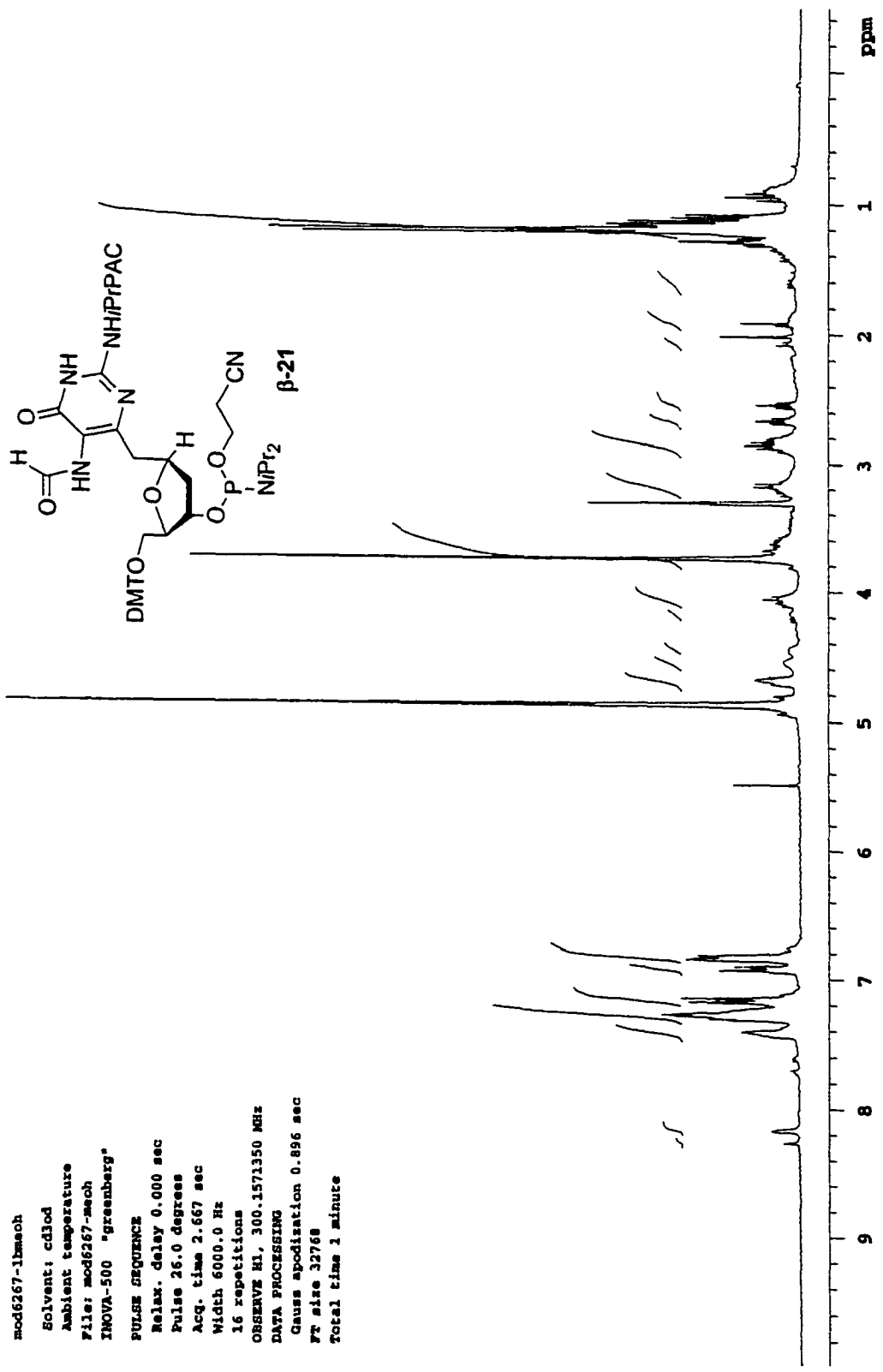








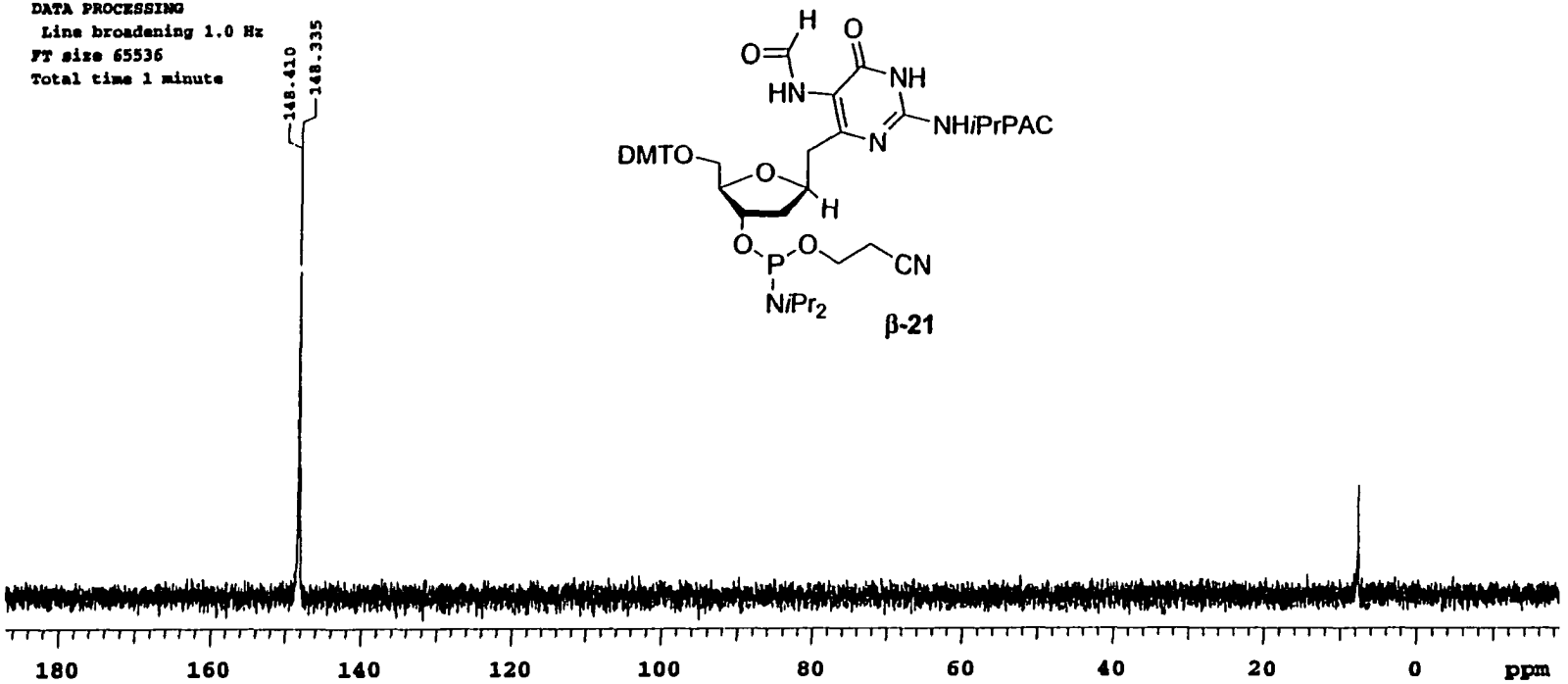




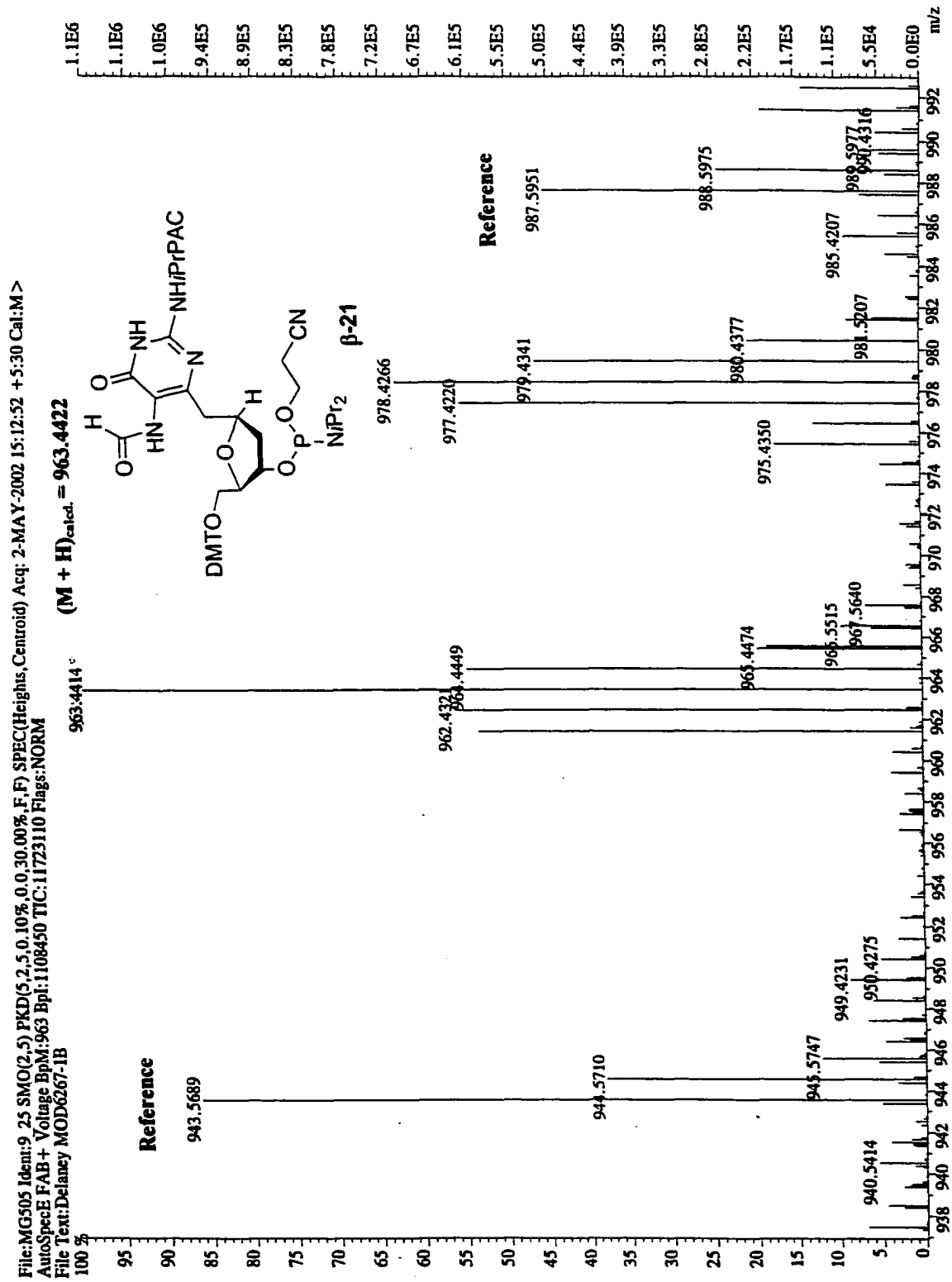
mod6215-1ap31

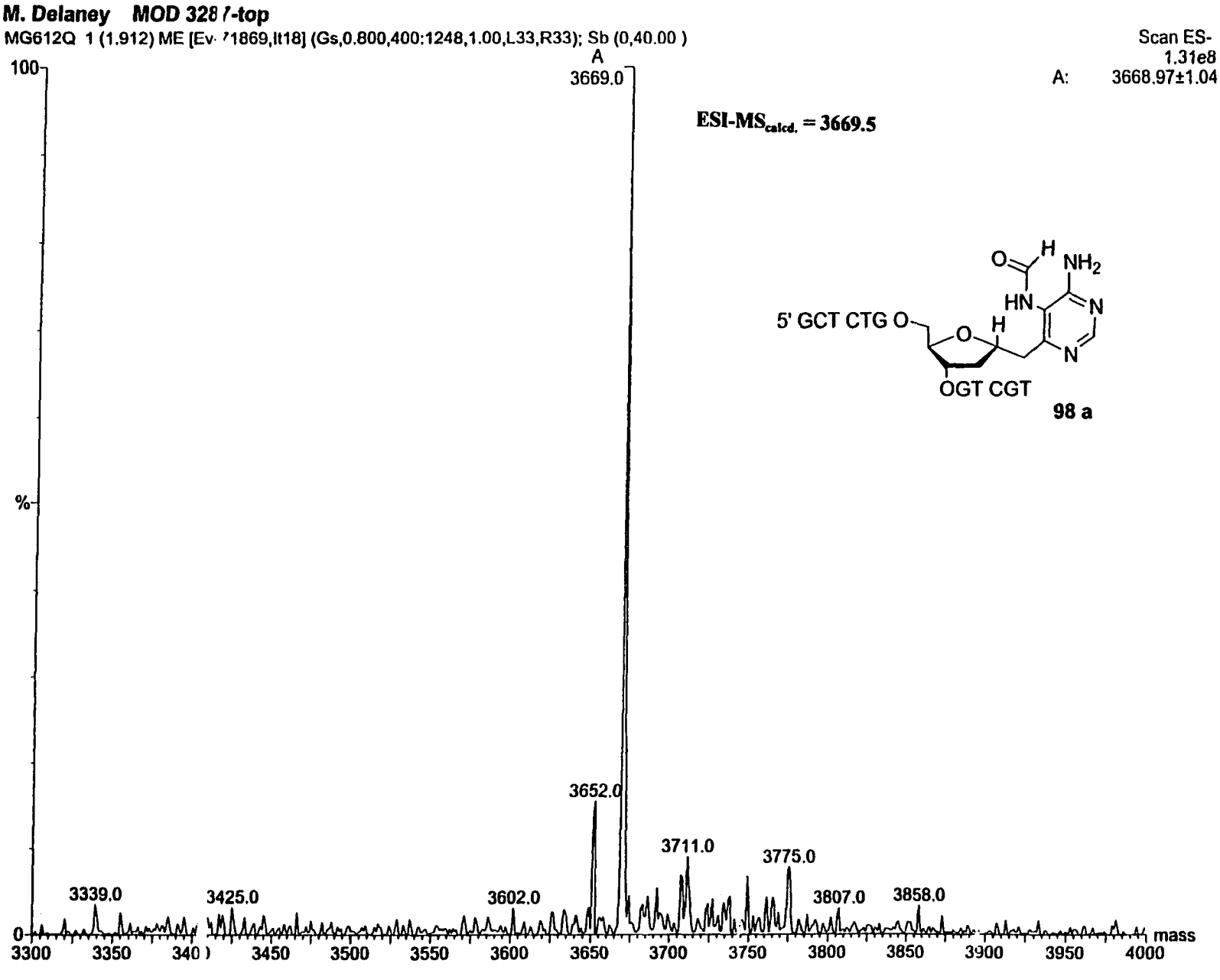
Solvent: CDCl3
Ambient temperature
File: mod6215-1ap31
INOVA-500 "greenberg"

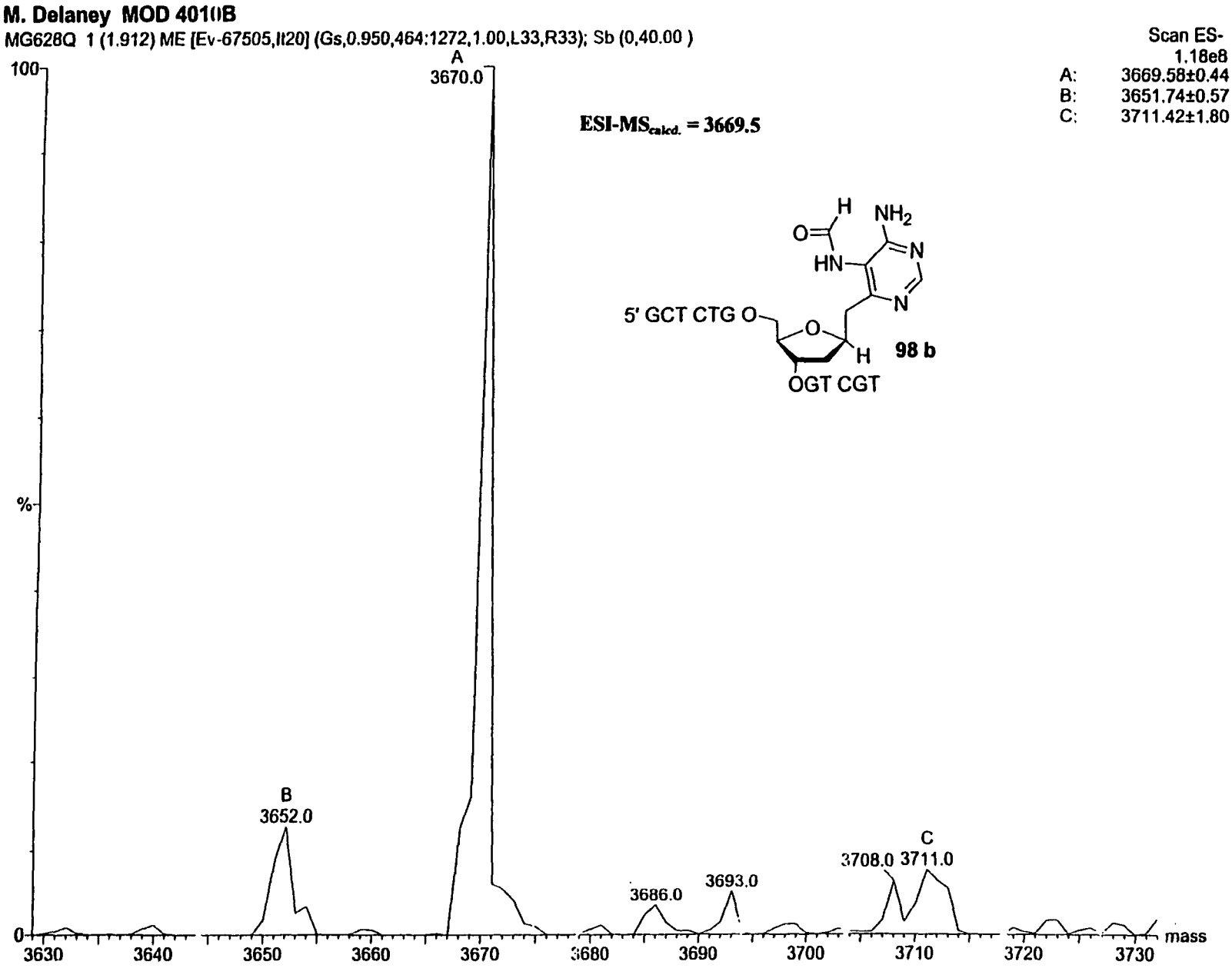
PULSE SEQUENCE
Relax. delay 0.000 sec
Pulse 45.0 degrees
Acq. time 1.066 sec
Width 25000.0 Hz
64 repetitions
OBSERVE P31, 121.5053253 MHz
DECOUPLE H1, 300.1574402 MHz
Power 40 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 1 minute

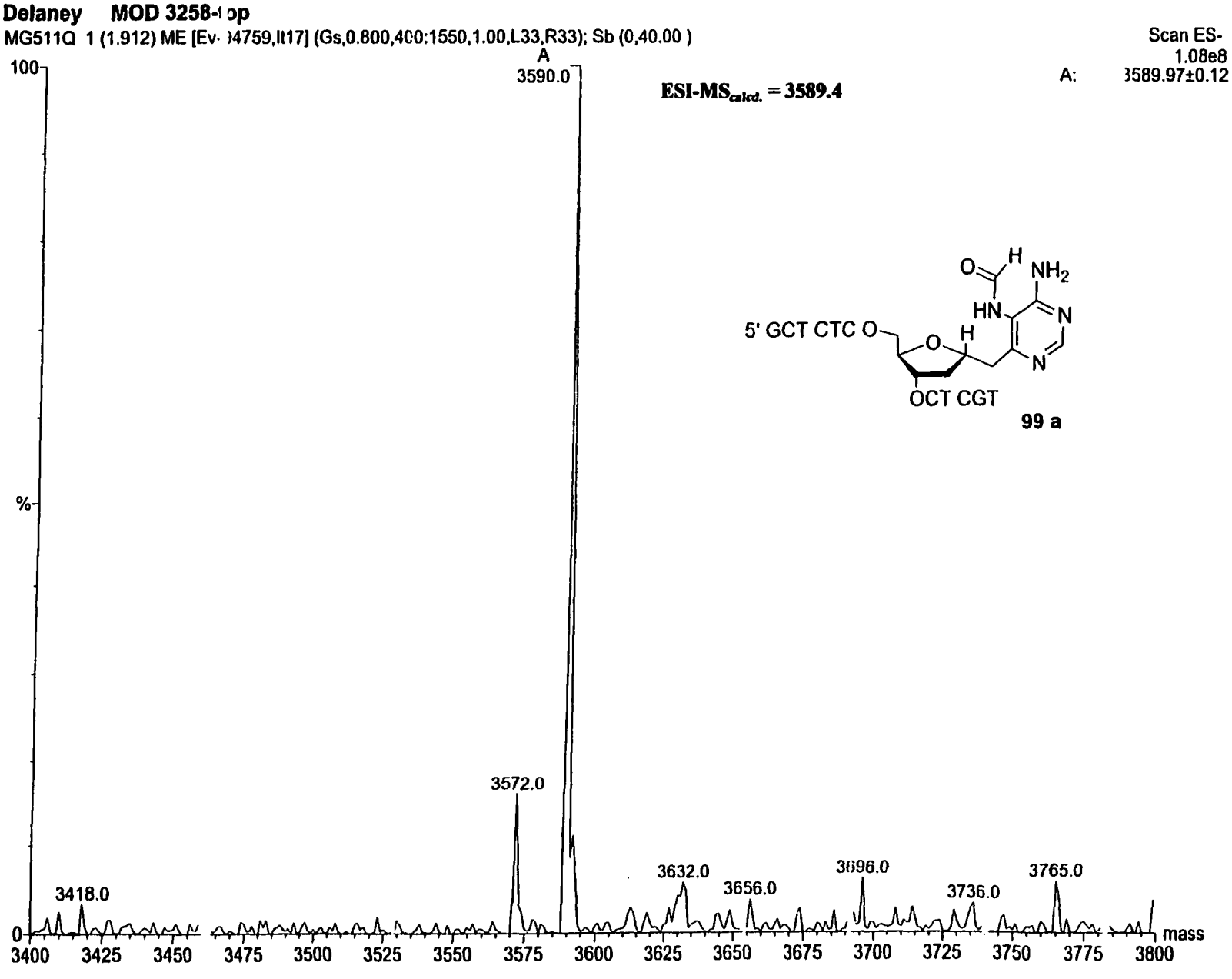


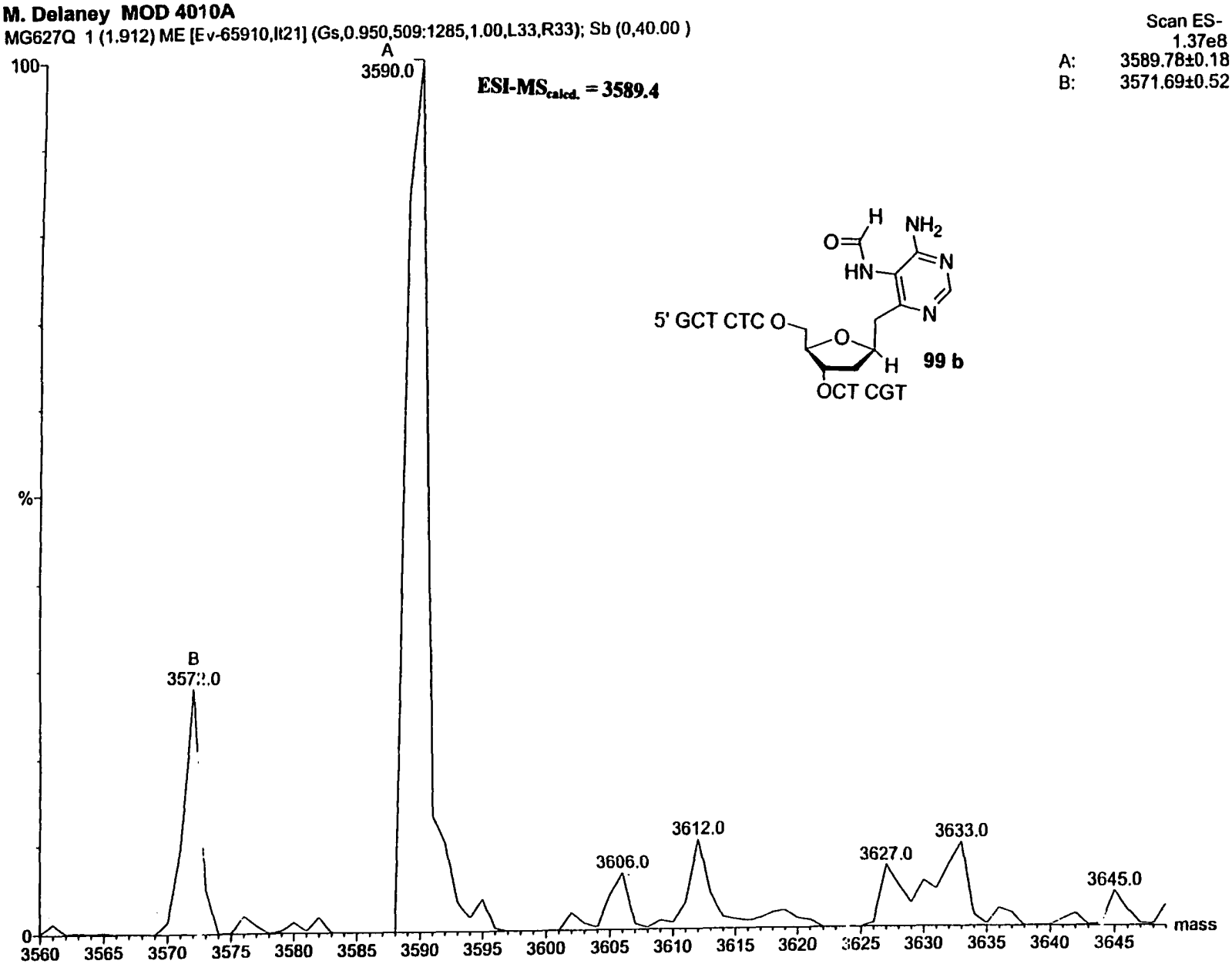
β-21

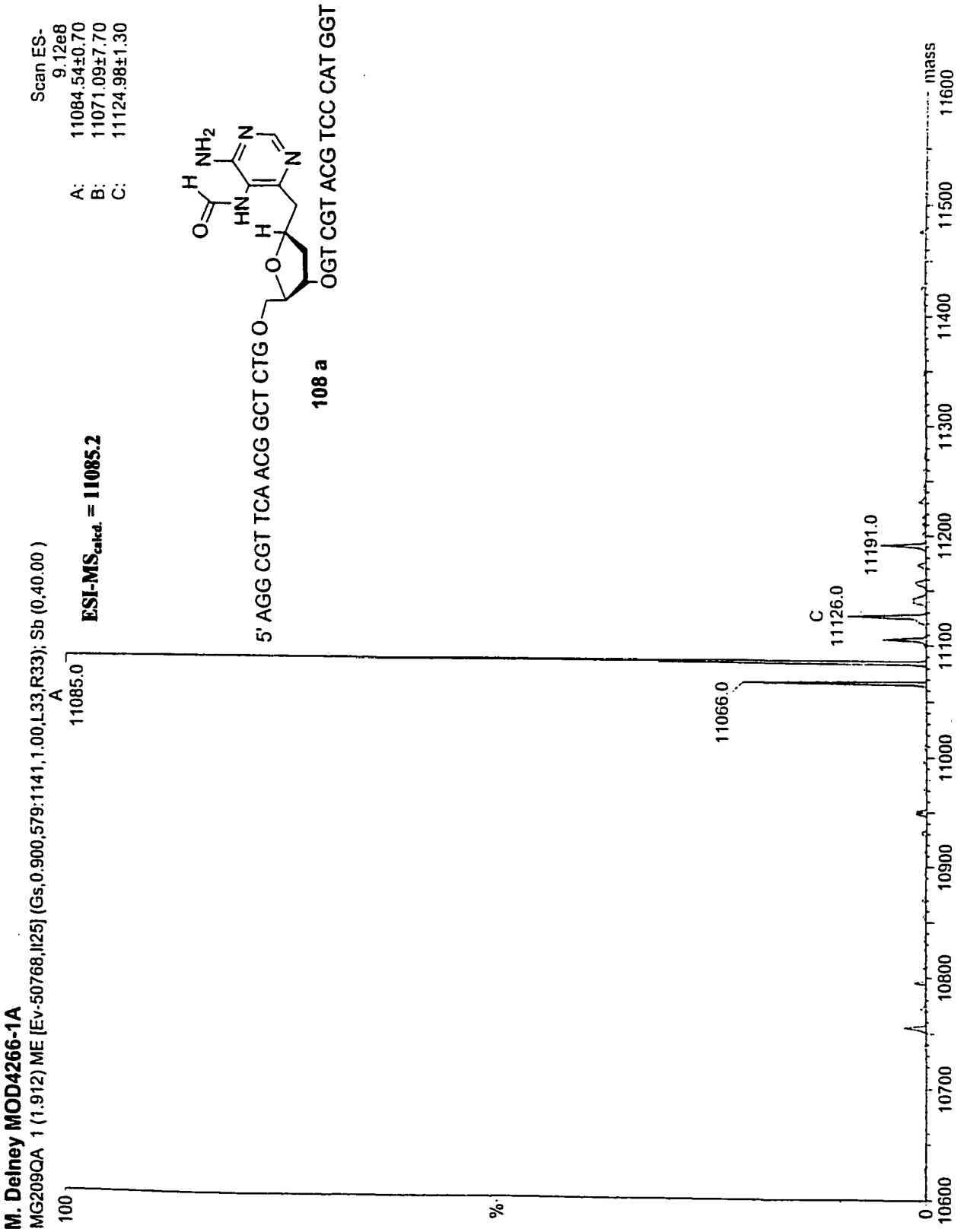


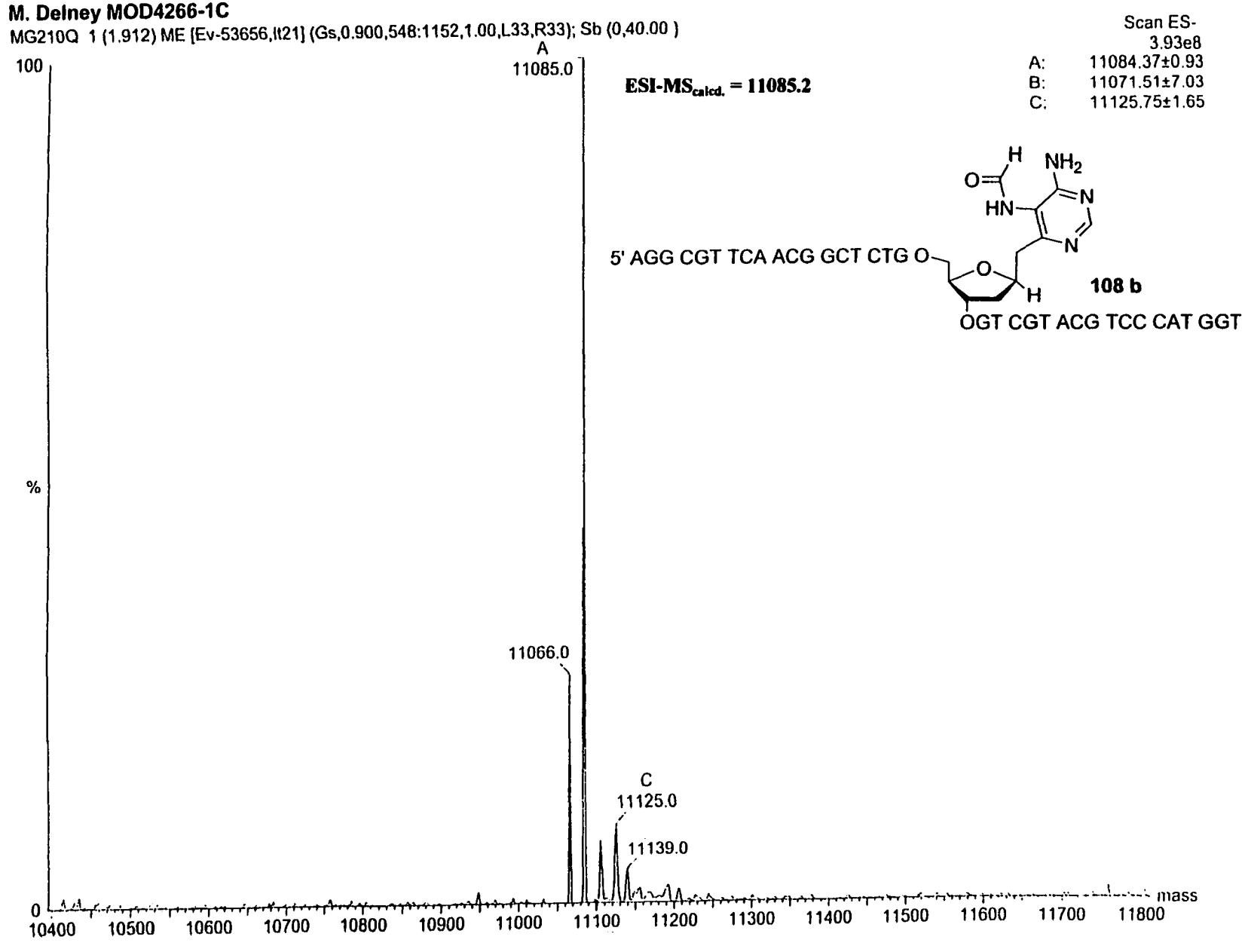


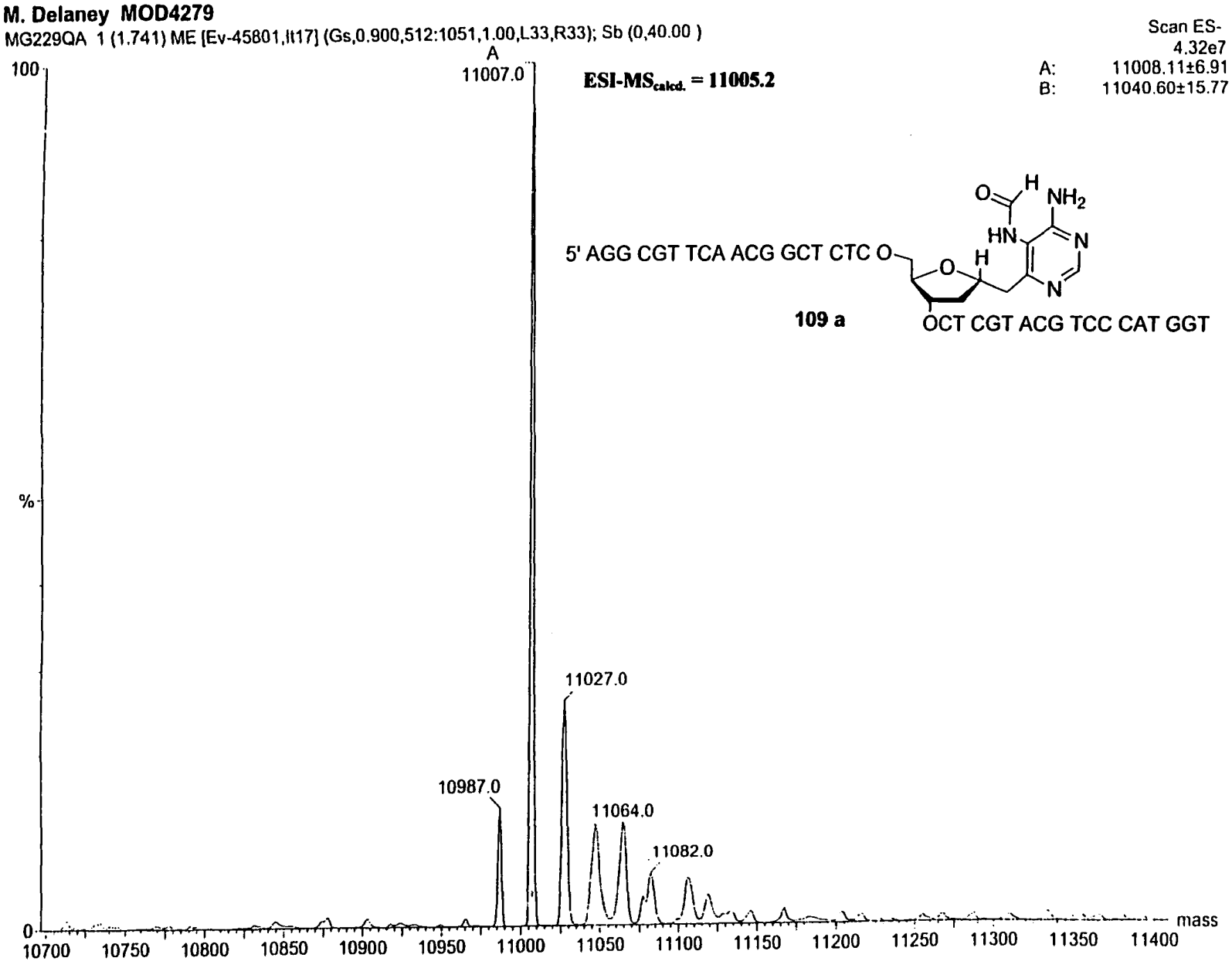


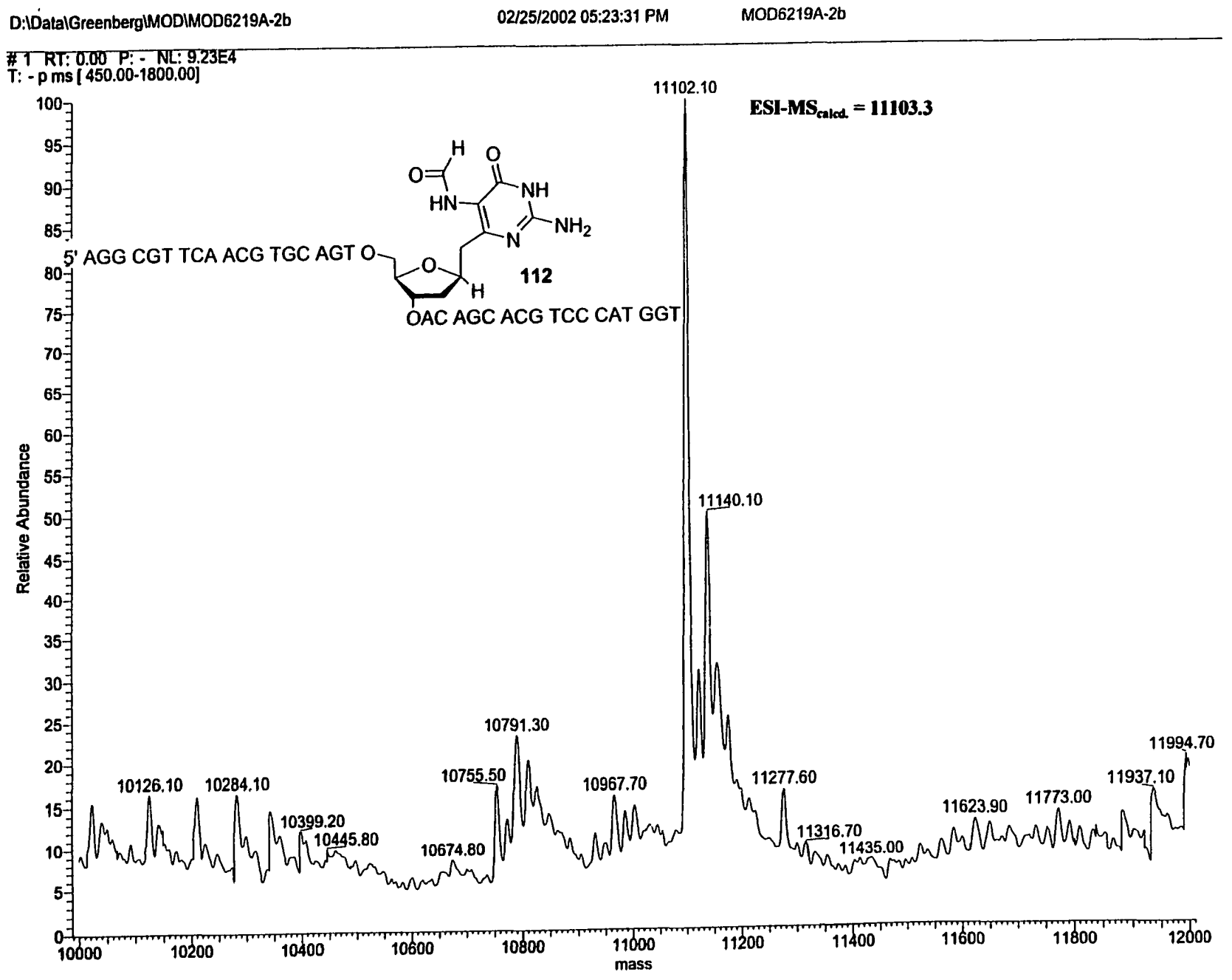


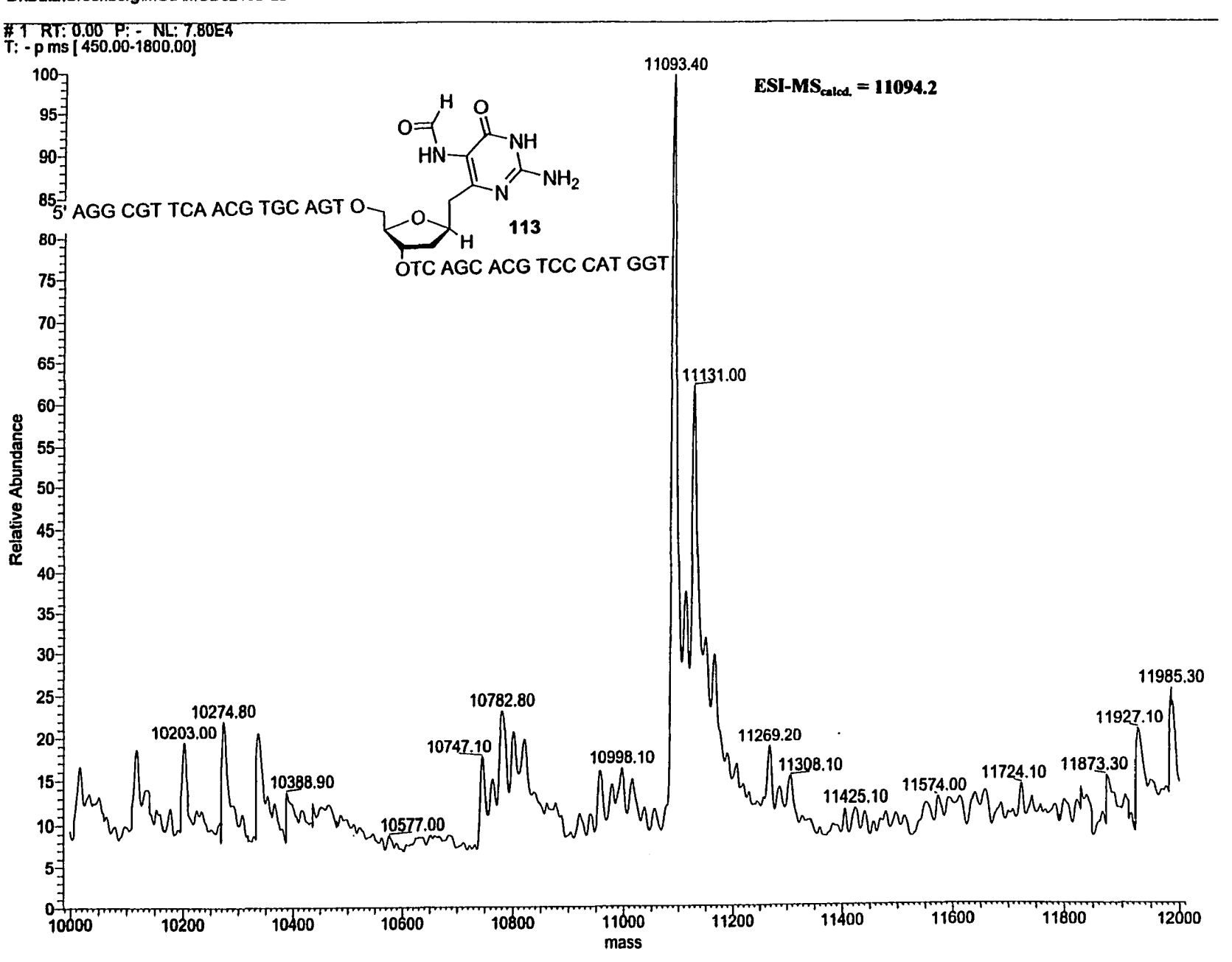


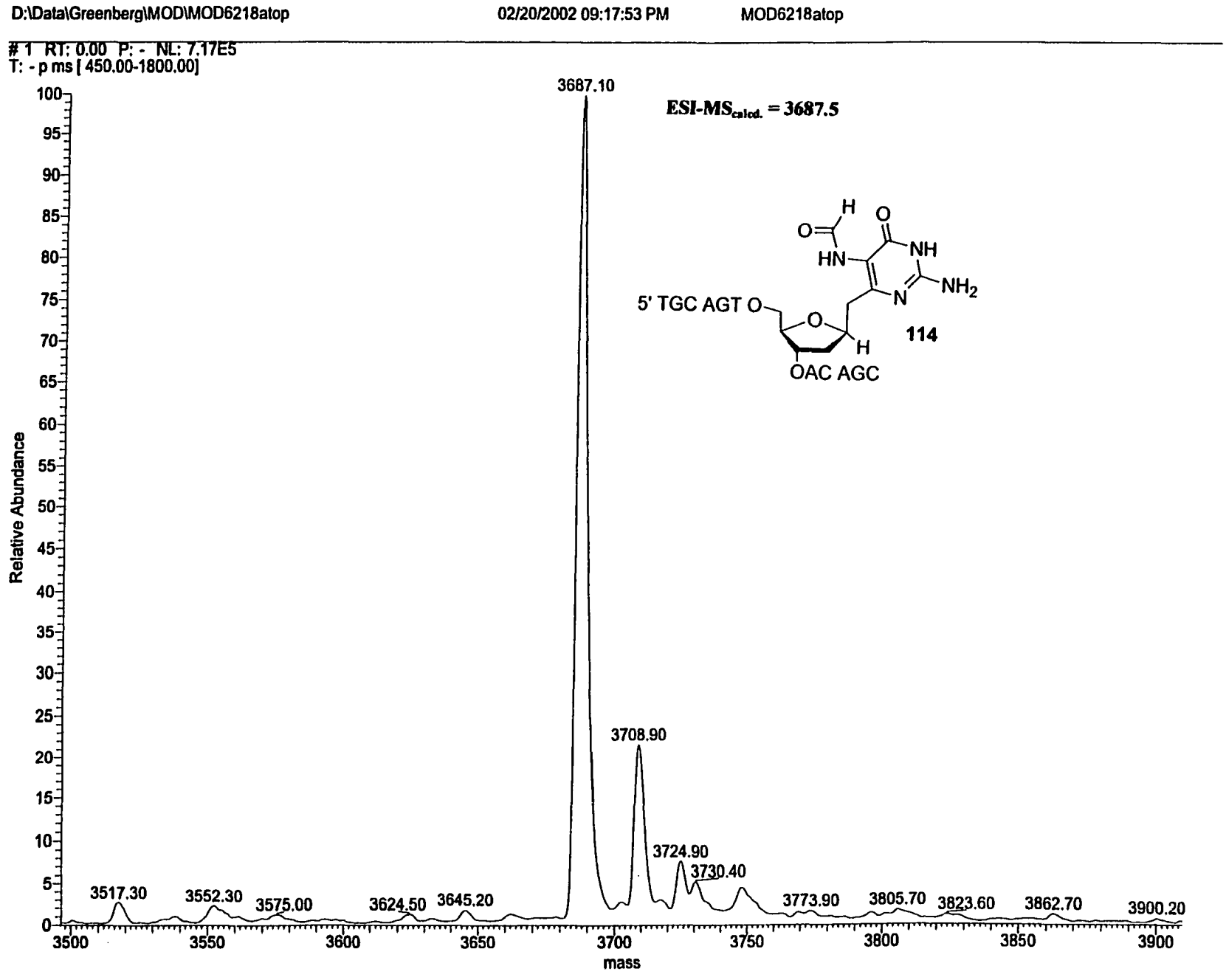












8 Appendix B: Assignment of Stereochemistry through NOE or ROESY NMR Experiments

Supplementary Table 1. NOE enhancements for β-**63**.

^{1}H irradiated	% Enhancement								
	C_1'	*pro-R* C_2'	*pro-S* C_2'	C_3'	C_4'	C_5'(A)[a]	C_5'(B)[a]	CH_2(A)[a]	CH_2(B)[a]
C_1'	---	2.15	0.8		1.39			1.53	1.14
pro-R C_2'	4.12	---	7.92	1.28	0.80				
pro-S C_2'	0.96	9.62	---	5.05				1.67	1.64
C_3'				---	2.77			4.45	
C_4'	1.51			1.16	---		5.07		
C_5'(A)[a]				2.22		---	14.23		
C_5'(B)[a]						22.78	---		
CH_2(A)[a]	2.45		1.36					---	13.44
CH_2(B)[a]	1.72		1.20				0.18	11.92	---

[a]A refers to the peak that is furthest upfield and B refers to the peak that is furthest downfield.

Supplementary Table 2. NOE enhancements for α-**63**.

^{1}H irradiated	% Enhancement							
	C_1'	*pro-R* C_2'	*pro-S* C_2'	C_3'	C_4'-C_5'(A)[a,b]	C_5'(B)[a]	CH_2(A)[a]	CH_2(B)[a]
C_1'	---	0.45	3.30	0.27	0.47		2.27	0.63
pro-R C_2'	0.79	---	18.80	0.90	1.67		1.99	2.25
pro-S C_2'	5.10	20.49	---	5.02				
C_3'	0.98	0.64		---	2.12	0.28	3.33	
C_4'-C_5'(A)[a,b]		0.87		0.29	---	11.06		0.95
C_5'(B)[a]				0.70	29.07	---		
CH_2(A)[a]	2.99	2.22					---	18.69
CH_2(B)[a]	1.46	2.22			1.72		19.50	---

[a]A refers to the peak that is furthest upfield and B refers to the peak that is furthest downfield.
[b]C_4' and C_5'(A) are unresolved.

Supplementary Table 3. NOE enhancements for β-70.

^{1}H irradiated	% Enhancement					
	C_1'	*pro-R* C_2'	*pro-S* C_2'	C_3'	C_4'	C_5'-CH_2[a]
C_1'	---		2.91	0.51		3.84
pro-R C_2'	0.82	---	16.03	1.06	1.06	2.61
pro-S C_2'	4.18	7.89	---	4.28	0.37	1.21
C_3'	0.89	1.11	4.00	---	0.58	1.81
C_4'		0.89	0.73	1.12	---	3.41
C_5'-CH_2[a]	4.14	2.22	0.11	-0.62	-1.35	---

[a] C_5' and the benzylic methylene peaks are unresolved.

Supplementary Table 4. NOE enhancements for α-70.

^{1}H irradiated	% Enhancement						
	C_1'	*pro-R* C_2'	*pro-S* C_2'	C_3'	C_4'	C_5'(A)- CH_2(A)[a,b]	C_5'(B)- CH_2(B)[a,c]
C_1'	---		2.67		1.36	1.37	1.00
pro-R C_2'	4.45	---	11.89	1.31	0.35		
pro-S C_2'	1.03	12.41	---	4.46		2.52	1.67
C_3'		1.03	2.88	---	0.78	1.20	1.20
C_4'	0.67			1.16	---	1.07	2.04
C_5'(A)- CH_2(A)[a,b]	1.35			1.70	1.73	---	9.52
C_5'(B)- CH_2(B)[a,c]	2.34			1.37	2.32	11.27	---

[a] A refers to the peak that is furthest upfield and B refers to the peak that is furthest downfield.
[b] C_5' (A) and CH_2(A) are unresolved.
[c] C_5' (B) and CH_2(B) are unresolved.

Supplementary Table 5. NOE enhancements for α-**96**.

^{1}H irradiated	% Enhancement		
	C_1'	C_3'	C_4'
C_1'	---		
C_3'	1.26	---	0.02
C_4'		0.50	---

Suplementary Table 6. ROSEY enhancements for β-**96**.

^{1}H irradiated	% Enhancement						
	C_1'	*pro-R* C_2'	*pro-S* C_2'	C_3'	C_4'	C_5'	CH_2
C_1'	---	3.02	0.55		1.06		2.62
pro-R C_2'	4.41	---	13.48	1.63	0.19		0.51
pro-S C_2'	1.19	19.53	---	6.30	0.38	1.24	4.35
C_3'		1.37	3.25	---	0.79	2.17	
C_4'	1.05			0.73	---	4.09	
C_5'			0.39	1.57	2.84	---	
CH_2	3.53	0.54	2.98				---

9 Appendix C: Molecular Modeling Parameters

Supplementary Table 7. Molecular modeling results for β-Fapy•dA and **β-18**.

model[a]	Glycosidic Torsional Angle(°)[b]	Nucleobase Tilt(°)[c]	Formamide Angle(°)[d]	Minimized Energy (ΔG, kcal/mol)
dA	-171.5[e]	1.1[f]	N/A	N/A
anti-Fapy•dA (*trans*)	-167.3	18.2	60.7	38.12
anti-Fapy•dA (*cis*)	-166.4	14.9	66.2	39.23
syn-Fapy•dA.1 (*trans*)	42.3	9.7	60.0	38.79
syn-Fapy•dA.2 (*trans*)	45.7	-179.3	-118.0	38.97
syn-Fapy•dA.3 (*cis*)	43.4	-174.8	65.4	39.96
anti-β-**18** (*trans*)	-175.7	64.2	113.7	37.30
anti-β-**18** (*cis*)	-175.4	63.8	111.2	38.36
syn-β-**18.1** (*trans*)	55.4	0.3	119.2	38.98
syn-β-**18.2** (*trans*)	54.5	3.1	-70.9	39.41
syn-β-**18.3** (*cis*)	55.2	0.4	79.4	40.50

[a]Formamide conformation is referenced in parentheses.
[b]Measured from the torsional angle created by O_4'-C_1' to N6(CH_2)-C6.
[c]Measured from the torsional angle created by C_1'-N6(CH_2) to C6-N1.
[d]Measured from the torsional angle created by C6-C5-N5-C(O)H.
[e]Measured from the torsional angle created by O_4'-C_1' to N9-C4.
[f]Measured from the torsional angle created by C_1'-N9 to C4-N3.

Supplementary Table 8. Molecular modeling results for α-Fapy•dA, α-**18**, and dA.

Model[a]	Glycosidic Torsional Angle(°)[b]	Nucleobase Tilt(°)[c]	Formamide Angle(°)[d]
dA	65.8[e]	1.0[f]	N/A
anti-Fapy•dA (*trans*)	61.5	18.6	58.8
anti-α-**18** (*trans*)	57.0	63.5	71.1

[a]Formamide conformation is referenced in parentheses.
[b]Measured from the torsional angle created by O_4'-C_1' to N6(CH_2)-C6.
[c]Measured from the torsional angle created by C_1'-N6(CH_2) to C6-N1.
[d]Measured from the torsional angle created by C6-C5-N5-C(O)H.
[e]Measured from the torsional angle created by O_4'-C_1' to N9-C4.
[f]Measured from the torsional angle created by C_1'-N9 to C4-N3

Supplementary Table 9. Molecular modeling results for β-Fapy•dG, β-**19** and dG.

model[a]	Glycosidic Torsional Angle(°)[b]	Nucleobase Tilt(°)[c]	Formamide Angle(°)[d]	Minimized Energy (ΔG, kcal/mol)
dG	-173.0[e]	-0.2[f]	N/A	N/A
anti-Fapy•dG (*trans*)	-169.3	19.4	54.6	37.66
anti-Fapy•dG (*cis*)	-167.2	15.3	63.6	39.96
syn-Fapy•dG.1 (*trans*)	47.0	12.6	53.8	37.18
syn-Fapy•dG.2 (*cis*)	48.2	10.2	60.9	38.56
syn-Fapy•dG.3 (*trans*)	55.9	-4.3	130.8	37.00
anti-β-**19** (*trans*)	-170.2	-1.1	68.1	39.21
anti-β-**19** (*cis*)	-171.1	2.2	77.3	40.46
syn-β-**19.1** (*trans*)	57.0	-6.4	67.2	38.42
syn-β-**19.2** (*trans*)	54.1	6.8	-66.1	38.81
syn-β-**19.3** (*cis*)	55.4	3.5	-76.5	40.12

[a]Formamide conformation is referenced in parentheses.
[b]Measured from the torsional angle created by O_4'-C_1' to N6(CH_2)-C6.
[c]Measured from the torsional angle created by C_1'-N6(CH_2) to C6-N1.
[d]Measured from the torsional angle created by C6-C5-N5-C(O)H.
[e]Measured from the torsional angle created by O_4'-C_1' to N9-C4.
[f]Measured from the torsional angle created by C_1'-N9 to C4-N3.

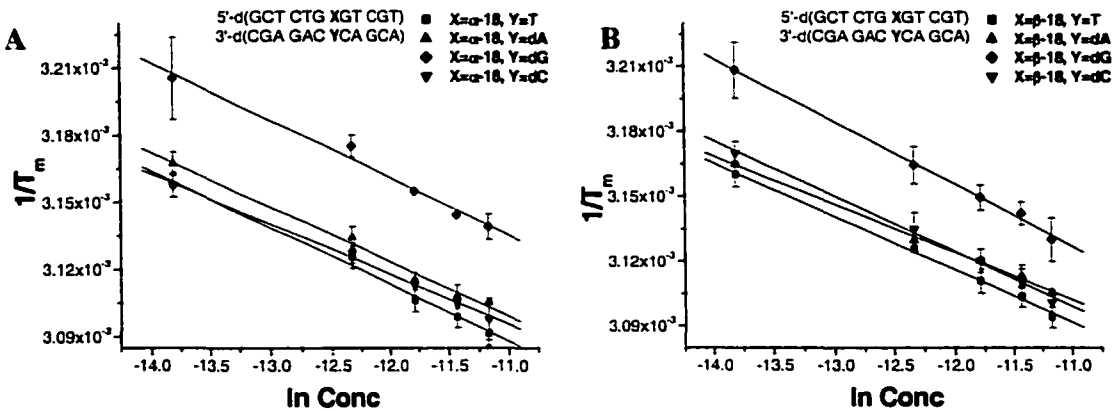
10 Appendix D: Thermodynamic Melting Parameters of Oligonucleotides Containing α-18, β-18, and β-19.

Supplementary Table 10. UV-melting thermodynamics of duplexes containing α-**18**, β-**18**, and dA.[a]

5'-d(GCT CTG **X**GT CGT)
3'-d(CGA GAC **Y**CA GCA)
115a,b,c
a(X=α-18),b(X=β-18), c(X=dA)

X:Y	T_m[b]	R^2	ΔH (kcal/mol)	ΔS (e.u.)	ΔG (kcal/mol)[c]
dA:T	56.3 ± 1.0	0.992	92.2 ± 8.7	252.0 ± 24.0	17.0 ± 8.7
dA:dA	46.2 ± 0.5	0.994	75.1 ± 9.9	207.6 ± 2.8	13.2 ± 1.0
dA:dG	47.8 ± 0.1	0.994	73.9 ± 2.7	202.1 ± 0.8	13.6 ± 0.4
dA:dC	44.6 ± 0.6	0.999	87.2 ± 4.7	246.7 ± 13.3	13.6 ± 4.7
β-**18**:T	46.8 ± 0.1	0.998	77.8 ± 2.3	215.6 ± 6.5	13.5 ± 2.3
β-**18**:dA	46.4 ± 0.5	0.998	81.9 ± 2.6	229.2 ± 7.7	13.5 ± 2.6
β-**18**:dG	42.9 ± 0.9	0.998	76.0 ± 2.5	213.5 ± 6.9	12.4 ± 2.5
β-**18**:dC	45.8 ± 0.8	0.997	77.5 ± 3.0	215.8 ± 8.5	13.2 ± 3.0
α-**18**:T	46.8 ± 0.1	0.994	80.0 ± 8.2	222.4 ± 22.8	13.7 ± 5.0
α-**18**:dA	45.9 ± 0.1	0.991	66.9 ± 3.7	182.2 ± 10.0	12.6 ± 0.5
α-**18**:dG	41.8 ± 0.1	0.991	73.1 ± 5.0	204.4 ± 14.1	12.1 ± 8.7
α-**18**:dC	46.6 ± 0.5	0.998	73.9 ± 5.3	203.6 ± 1.6	13.1 ± 1.0

[a]Conditions: 10 mM PIPES (pH 7.0); 10 mM $MgCl_2$; 100 mM NaCl.
[b]Total oligonucleotide concentration = 4.4μM.
[c]Calculated at 298 K.



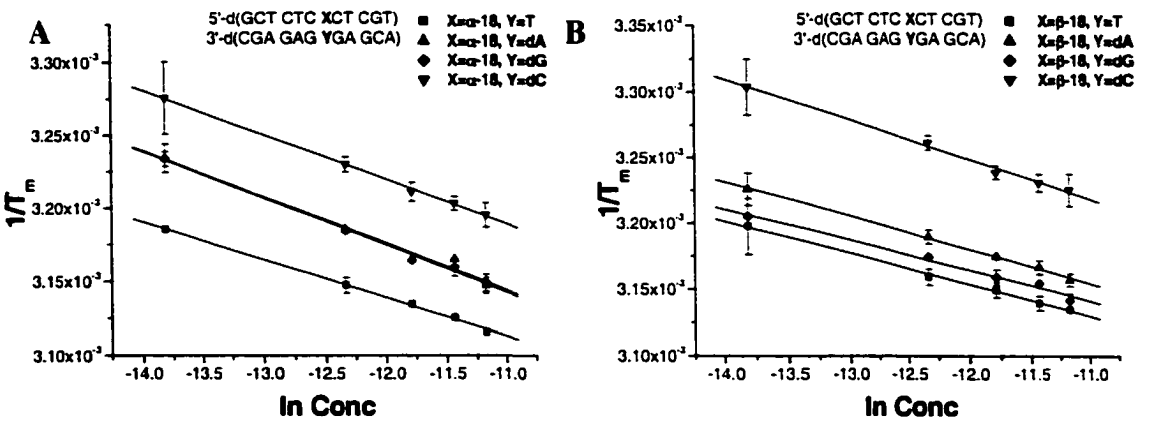
Supplementary Figure 1: van't Hoff plots of thermodynamic melting experiments on duplexes containing α-**18**(A) and β-**18**(B).

Supplementary Table 11. UV-melting thermodynamics of duplexes containing α-**18**, β-**18**, and dA.[a]

5'-d(GCT CTC **X**CT CGT)
3'-d(CGA GAG **Y**GA GCA)
116a,b,c
a(**X**=α-**18**),**b**(**X**=β-**18**), **c**(**X**=dA)

X:Y	T_m[b]	R^2	ΔH (kcal/mol)	ΔS (e.u.)	ΔG (kcal/mol)[c]
dA:T	56.1 ± 1.3	0.989	96.3 ± 1.0	265.2 ± 2.8	17.2 ± 1.2
dA:dA	44.9 ± 0.1	0.997	79.5 ± 1.2	222.7 ± 3.4	13.1 ± 1.2
dA:dG	52.0 ± 0.1	0.994	90.0 ± 2.2	249.5 ± 6.0	15.6 ± 2.2
dA:dC	39.9 ± 0.1	0.994	69.2 ± 1.2	193.4 ± 3.3	11.5 ± 1.6
β-**18**:T	43.4 ± 0.6	0.994	83.3 ± 2.7	236.3 ± 7.8	12.9 ± 2.7
β-**18**:dA	40.4 ± 0.5	0.994	77.9 ± 2.1	221.2 ± 5.9	12.0 ± 2.1
β-**18**:dG	41.9 ± 0.1	0.992	86.7 ± 5.6	247.9 ± 16.0	12.8 ± 5.6
β-**18**:dC	33.5 ± 0.5	0.993	65.7 ± 2.5	187.0 ± 7.3	10.0 ± 2.5
α-**18**:T	44.6 ± 0.5	0.999	78.0 ± 2.2	218.2 ± 6.2	12.9 ± 2.2
α-**18**:dA	40.8 ± 0.1	0.984	63.1 ± 4.7	173.8 ± 13.2	11.3 ± 4.7
α-**18**:dG	40.9 ± 0.1	0.999	62.5 ± 2.5	171.7 ± 7.0	11.3 ± 2.5
α-**18**:dC	36.5 ± 0.5	0.994	64.8 ± 1.4	182.0 ± 3.8	13.1 ± 1.4

[a]Conditions: 10 mM PIPES (pH 7.0); 10 mM $MgCl_2$; 100 mM NaCl.
[b]Total oligonucleotide concentration = 4.4μM.
[c]Calculated at 298 K.



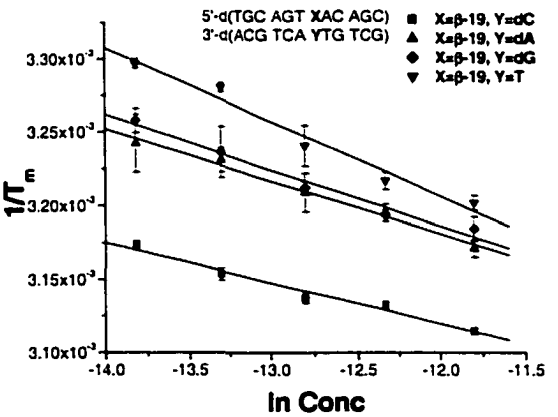
Supplementary Figure 2: van't Hoff plots of thermodynamic melting experiments on duplexes containing α-**18**(A) and β-**18**(B).

Supplementary Table 12. UV-melting thermodynamics of duplexes containing β-**19** and dG.[a]

5'-d(TGC ACT **X**AC AGC)
3'-d(ACG TGA **Y**GT TCG)
119a,b
a(X=β-19), b(X=dG)

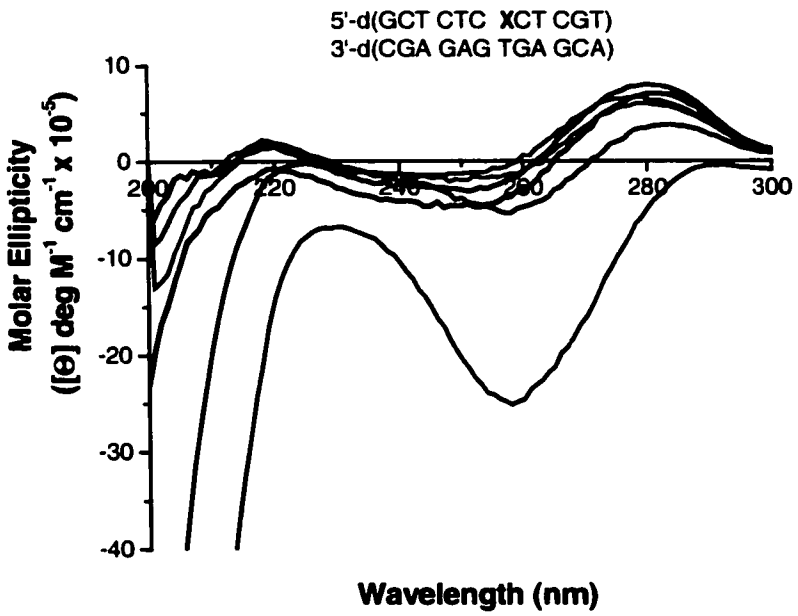
X:Y	T_m[b]	R^2	ΔH (kcal/mol)	ΔS (e.u.)	ΔG (kcal/mol)[c]
dG:dC	57.4 ± 0.4	0.994	71.9 ± 8.3	189.9 ± 21.8	15.3 ± 8.2
dG:dG	44.6 ± 0.5	0.982	58.3 ± 6.5	157.8 ± 11.2	11.3 ± 6.5
dG:T	46.3 ± 0.3	0.997	83.2 ± 8.7	234.1 ± 24.4	13.4 ± 8.7
dG:dA	45.5 ± 0.3	0.991	83.0 ± 8.4	235.1 ± 23.8	12.9 ± 8.4
β-**19**:dC	46.1 ± 0.3	0.987	70.9 ± 6.6	194.4 ± 11.7	13.0 ± 6.5
β-**19**:dG	40.0 ± 0.4	0.994	51.3 ± 3.3	136.8 ± 9.0	10.5 ± 3.3
β-**19**:T	37.7 ± 0.6	0.989	39.9 ± 3.6	101.5 ± 9.4	9.7 ± 3.6
β-**19**:dA	39.8 ± 0.6	0.986	51.5 ± 5.0	137.3 ± 13.8	10.6 ± 5.0

[a]Conditions: 10 mM PIPES (pH 7.0); 10 mM $MgCl_2$; 100 mM NaCl.
[b]Total oligonucleotide concentration = 4.4μM.
[c]Calculated at 298 K.



Supplementary Figure 3: van't Hoff plots of thermodynamic melting experiments on duplexes containing β-**19**.

11 Appendix E: CD Spectra of Oligonucleotides Containing α-18 and β-18 Under High and Low Salt Conditions



Supplementary Figure 4. CD spectrometry on duplexes containing α-**18**:T and β-**18**:T under different NaCl concentrations. 100 mM NaCl, 10 mM Phosphate (pH 7.5): Blue (X=A), Aqua (X=α-**18**), Pink (X=β-**18**). 10 mM NaCl, 10 mM Phosphate (pH 7.5): Black (X=A), Red (X=α-**18**), Green (X=β-**18**).

12 Appendix F: Statistical Analysis of Individual Standing Start Insertion Reactions

Supplementary Table 13. Statistical analysis of TTP insertion opposite α-**18**, β-**18**, and dA by Klenow (exo$^-$).

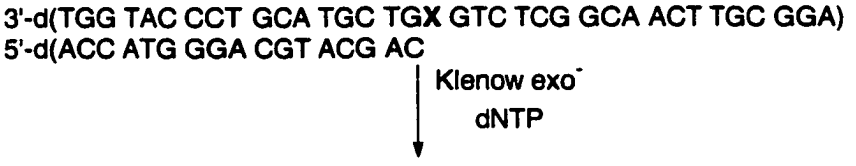
3'-d(TGG TAC CCT GCA TGC TGX GTC TCG GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT ACG AC

Klenow exo$^-$
dNTP

3'-d(TGG TAC CCT GCA TGC TGX GTC TCG GCA ACT TGC GGA)
5'-d(ACC ATG GGA CGT ACG ACN

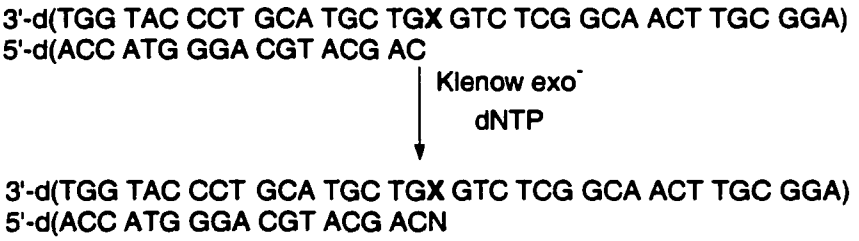
Experiment	X	R^2	K_m (μM)	V_{max} (% min^{-1})	V_{max}/K_m
MOD5263C	dA	0.99	0.028 ± 0.002	7.2 ± 0.2	2.6 ± 0.1 x 10^8
MOD5265C	dA	0.99	0.032 ± 0.005	8.0 ± 0.2	2.5 ± 0.4 x 10^8
MOD5267C	dA	0.99	0.048 ± 0.005	9.3 ± 0.2	1.9 ± 0.2 x 10^8
Average	**dA**		**0.036 ± 0.01**	**8.2 ± 1.1**	**2.3 ± 0.3 x 10^8**
MOD5267A	α-**18**	0.97	6.3 ± 1.9	23.7 ± 2.3	3.8 ± 0.4 x 10^6
MOD5270A	α-**18**	0.99	6.2 ± 0.5	47.0 ± 3.0	7.6 ± 0.5 x 10^6
MOD5272A	α-**18**	0.95	6.3 ± 0.7	17.2 ± 1.8	2.7 ± 0.1 x 10^6
Average	**α-18**		**6.3 ± 0.1**	**29.3 ± 15.6**	**4.7 ± 2.6 x 10^6**
MOD5265B	β-**18**	0.96	1.1 ± 0.1	5.9 ± 0.4	5.6 ± 0.2 x 10^6
MOD5270B	β-**18**	0.90	1.4 ± 0.4	7.0 ± 2.0	5.8 ± 0.5 x 10^6
MOD5272B	β-**18**	0.98	0.4 ± 0.1	2.6 ± 0.2	7.12 ± 0.2 x 10^6
Average	**β-18**		**0.9 ± 0.5**	**5.5 ± 2.7**	**6.2 ± 0.8 x 10^6**

Supplementary Table 14. Statistical analysis of dATP insertion opposite α-18, β-18, and dA by Klenow (exo$^-$).



Experiment	X	R^2	K_m (μM)	V_{max} (% min^{-1})	V_{max}/K_m
MOD6023C	dA	0.99	125.6 ± 14.7	53.9 ± 6.1	4.29 ± 0.1 x 10^5
MOD6026C	dA	0.91	158.0 ± 31.1	48.2 ± 9.4	3.1 ± 0.1 x 10^5
MOD6030C	dA	0.90	93.9 ± 19.6	47.6 ± 9.6	5.1 ± 0.1 x 10^5
Average	**dA**		**126 ± 32**	**49.9 ± 3.5**	**4.1 ± 1.0 x 10^5**
MOD5271A	α-18	0.99	84.8 ± 5.8	19.8 ± 1.1	2.3 ± 0.1 x 10^5
MOD5273A	α-18	0.98	73.0 ± 10.0	19.4 ± 2.6	2.7 ± 0.1 x 10^5
MOD5268A	α-18	0.97	92.8 ± 19.6	12.0 ± 2.1	1.3 ± 0.2 x 10^5
Average	**α-18**		**83.5 ± 10.0**	**17.1 ± 4.4**	**2.1 ± 0.7 x 10^5**
MOD6026B2	β-18	0.03	308 ± 53	44.9 ± 7.7	1.5 ± 0.1 x 10^5
MOD6030B1	β-18	0.97	441 ± 53	70.0 ± 8.4	1.6 ± 0.1 x 10^5
MOD6030B2	β-18	0.97	217 ± 53	50.1 ± 11.1	2.3 ± 0.1 x 10^5
Average	**β-18**		**322 ± 94**	**55.0 ± 13.3**	**1.8 ± 0.9 x 10^5**

Supplementary Table 15. Statistical analysis of dGTP insertion opposite α-**18**, β-**18**, and dA by Klenow (exo$^-$).



Experiment	X	R^2	K_m (μM)	V_{max} (% min^{-1})	V_{max}/K_m
MOD5277C	dA	0.96	82.3 ± 19.9	12.7 ± 1.1	1.5 ± 0.3 x 10^5
MOD5278C	dA	0.99	35.2 ± 9.0	10.6 ± 0.8	3.0 ± 0.7 x 10^5
Average	**dA**		**58.8 ± 33.0**	**11.6 ± 1.5**	**2.3 ± 1.0 x 10^5**
MOD5277A	α-**18**	0.99	25.2 ± 9.3	3.31 ± 0.2	1.3 ± 0.5 x 10^5
MOD5278A	α-**18**	0.99	54.5 ± 6.2	2.08 ± 0.1	3.8 ± 0.4 x 10^4
Average	**α-18**		**39.9 ± 21.0**	**2.7 ± 0.9**	**8.5 ± 6.62 x 10^4**
MOD5277B	β-**18**	0.99	80.1 ± 11.7	2.7 ± 0.1	3.3 ± 0.5 x 10^4
MOD5278B	β-**18**	0.99	69.8 ± 8.7	2.5 ± 0.1	3.5 ± 0.4 x 10^4
Average	**β-18**		**75.0 ± 7.3**	**2.6 ± 0.1**	**3.4 ± 0.1 x 10^4**

Supplementary Table 17. Statistical analysis of insertion of dNTP's opposite **β-19** by Klenow (exo$^-$).

3'-d(TGG TAC CCT GCA CGA CA**X** TGA CGT GCA ACT TGC AGG)
5'-d(ACC ATG GGA CGT GCT GT

Klenow exo$^-$
dNTP

3'-d(TGG TAC CCT GCA CGA CA**X** TGA CGT GCA ACT TGC AGG)
5'-d(ACC ATG GGA CGT GCT GT**N**

Experiment	X	R^2	dNTP	K_m (μM)	V_{max} (% min^{-1})	V_{max}/K_m
MOD6253C	β-19	0.99	C	0.030 ± 0.003	2.9 ± 0.1	9.7 ± 1.0 x 10^7
MOD6256C	β-19	0.99	C	0.028 ± 0.007	2.7 ± 0.1	9.7 ± 2.3 x 10^7
MOD6258C	β-19	0.99	C	0.028 ± 0.007	2.7 ± 0.1	9.7 ± 2.3 x 10^7
Average	**β-19**		**C**	**0.029 ± 0.001**	**2.8 ± 0.1**	**9. 7± 0.1 x 10^7**
MOD6258A	β-19	0.99	A	20.6 ± 1.7	9.8 ± 0.7	4.8 ± 0.3 x 10^5
MOD6262A	β-19	0.99	A	17.4 ± 2.0	9.4 ± 0.9	5.4 ± 1.0 x 10^5
MOD6263A	β-19	0.99	A	10.0 ± 1.9	6.9 ± 0.8	6.9 ± 0.2 x 10^5
Average	**β-19**		**A**	**16.0 ± 5.4**	**8.7 ± 1.6**	**5.7 ± 1.1 x 10^5**
MOD6256G	β-19	0.98	G	42.8 ± 1.9	4.8 ± 0.3	1.1 ± 0.1x 10^5
MOD6257G	β-19	0.99	G	52.0 ± 2.6	10.1 ± 0.3	1.9 ± 0.1x 10^5
MOD6262G	β-19	0.99	G	57.3 ± 2.4	8.9 ± 0.3	1.6 ± 0.1 x 10^5
Average	**β-19**		**G**	**50.7 ± 7.3**	**7.9 ± 2.8**	**1.5 ± 0.4 x 10^5**
MOD6253T	β-19	0.99	T	253 ± 36	2.7 ± 0.3	1.1 ± 0.1 x 10^4
MOD6257T	β-19	0.99	T	157 ± 23	1.8 ± 0.1	1.1 ± 0.2 x 10^4
MOD6261T	β-19	0.97	T	191 ± 44	2.0 ± 0.2	1.0 ± 0.2 x 10^4
MOD6263T	β-19	0.99	T	286 ± 63	3.2 ± 0.3	1.1 ± 0.2 x 10^4
Average	**β-19**		**T**	**222 ± 59**	**2.4 ± 0.7**	**1.1 ± 0.1 x 10^4**

13 Appendix G: Statistical Analysis of Individual PGSA Experiments

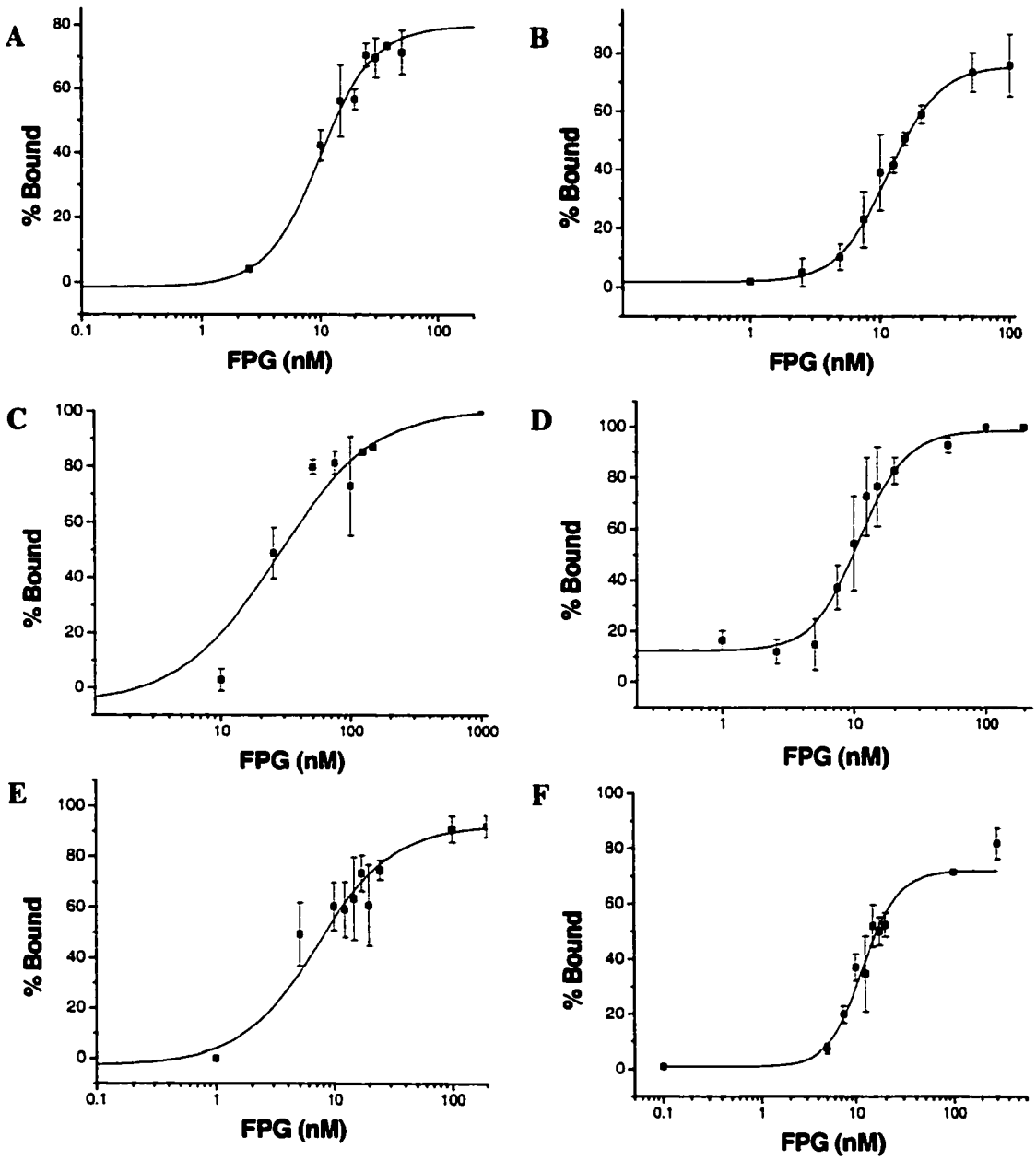
Supplementary Table 18. Dissociation constants of the Fpg protein for individual experiments on oligonucleotides containing α-**18**, β-**18**, β-**19**.

5'-d(AGG CGT TCA ACG GCT CTG **X**GT CGT ACG TCC CAT GGT)
3'-d(T CC GCA AGT TGC CGA GAC YCA GCA TGC AGG GTA CCT)
122a,b,c
a(X=α-**18**), **b**(X=β-**18**), **c**(X=dA)

5'-d(AGG CGT TCA ACG TGC AGT **X**AC AGC ACG TCC CAT GGT)
3'-d(TCC GCA AGT TGA ACG TCA YTG TCG TGC AGG GTA CCA)
123 (X=β-**19**)

Experiment	X	Y	K_d (nM)[a]
MOD5185	β-**18**	T	10.5 ± 1.3
MOD5198	β-**18**	T	16.4 ± 2.8
MOD5206	β-**18**	T	14.1 ± 2.4
MOD5180	β-**18**	dA	12.8 ± 3.0
MOD5182	β-**18**	dA	13.6 ± 2.9
MOD5205	β-**18**	dA	11.5 ± 1.1
MOD5203	β-**18**	dG	26.6 ± 3.5
MOD5183	β-**18**	dC	10.8 ± 3.2
MOD5186	β-**18**	dC	11.7 ± 2.0
MOD5199	β-**18**	dC	10.9 ± 1.3
MOD6019	α-**18**	dC	7.1 ± 1.5
MOD5243	dA	T	171.0 ± 37.5
MOD6240	β-**19**	dC	11.8 ± 0.7
MOD6242	β-**19**	dC	8.7 ± 0.4
MOD6243	β-**19**	dC	12.5 ± 2.9

[a]Each experiment reflects the average of three independent trials.



Supplementary Figure 5. Representative plots of gel-shift assays. A) β-**18**:T; B) β-**18**:dA; C) β-**18**:d; D) β-**18**:dC; E) α-**18**:dC; F) β-**19**:dC

14: Acknowledgements:

I am grateful for the many associates in the Greenberg Group that have contributed in my project during my graduate studies at Colorado State University. Specifically, Professor Marc M. Greenberg for advisement and financial funding during my thesis project. Carissa J. Wiederholt for her technical assistance in the biochemical evaluation of C-nucleotide analogues of Fapy lesions. I also need to acknowledge Carissa J. Wiederholt for performing the thermodynamic melting studies, insertion and extension kinetics, and for experiments involving Fpg and MutY on the native Fapy•dA and Fapy•dG lesions. I am appreciative for Professor Kazuhiro Haraguci, Dr. Aruna Sambandam, and Dr. Zsolt Hantosi for their contributions in the synthesis of the native Fapy lesions.

I would also like to acknowledge the contributions from my family. Misty A. Delaney for her continual support throughout the many years of my education in both undergraduate and graduate studies. For her constant willingness to offer encouragement and to sacrifice her time and other desires so that I could pursue my degree. In addition, I am grateful for the personal as well as financial support from George R. Delaney and Marlene K. Delaney during my graduate studies.