Automated Synthesis of Gene Fragments

Gabriel Alvarado-Urbina, Ganesh M. Sathe, Wing-Cheong Liu
Michael F. Gillen, Peter D. Duck, Robert Bender
Kelvin K. Ogilvie

Summary. The DNA/RNA Synthesizer provides a complete and automated procedure for the synthesis of DNA sequences. Each base unit is added in a 30-minute cycle, permitting a tetradecamer to be constructed in 6½ hours. The complete procedure is described, including a practical procedure for isolation and purification of the desired DNA sequence.

Genes from diverse sources can be placed into cells of specific living systems where they can reproduce and direct the synthesis of proteins (*1*). Bacterial cells have been widely used as host organisms, but yeasts and other eukaryotic cells may offer wider versatility in the future. Genes produced by total chemical synthesis can function in bacterial cells to direct the production of protein (*1*). In addition to the possibility of producing proteins of known importance, such as insulin and growth hormone, chemical synthesis of genes offers the capability of experimentally creating new polypeptides and enzymes with potential medical and industrial applications. Realization of this potential depends on the ability to produce DNA sequences rapidly and economically.

Genes are specific sequences of deoxyribonucleic acid (DNA). A short segment of a DNA molecule is shown in Fig. 1. Four different units can be used to construct a DNA chain. These units are called mononucleotides, and the order in which they are linked together determines the information carried by a specific DNA sequence, that is, by a specific gene. A gene may contain hundreds or thousands of units. In order to produce a synthetic gene, chemical techniques must be available to convert a nucleoside unit (Fig. 1), which is the normal starting point, into a nucleotide derivative. The nucleotide derivatives must then be successfully joined together to produce the desired DNA sequence. Sequences containing 10 to 14 units are those most generally produced by chemical synthesis.

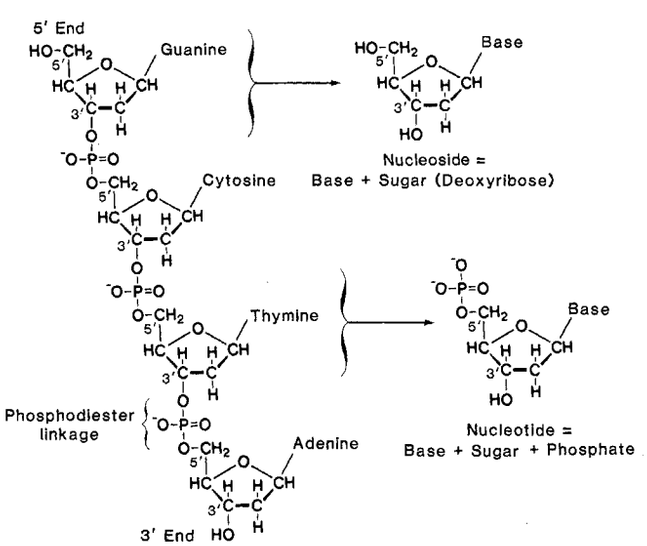


Fig. 1. The chemical structure of d-GCTA, a short DNA sequence.

Pieces of DNA 10 to 14 units in size have numerous practical applications. Probably their most dramatic use is joining them together by means of specific enzymatic techniques to form whole genes. In addition, pieces 10 to 14 units long can serve as probes to aid in isolating a desired natural gene from a living organism, or they can serve as primers to enable copying the genetic information carried in natural genes. They can also serve as linkers to aid in splicing together pieces of a desired gene or in inserting a gene into a plasmid or other vector. There are also numerous uses for pieces of this size in pure research such as the physical-chemical investigations of DNA.

Although the 10- to 14-unit sequence of DNA is a widely required and versatile size of DNA, such sequences had not been readily available to the majority of scientific groups needing them. Only a highly trained and skilled chemist could produce a single 12-unit sequence in less than 3 months. Moreover, because of the numerous procedures required to produce a DNA sequence, a new laboratory, starting with nucleosides, would probably require 6 to 12 months to make a single 12-unit sequence. In addition to the time required to join the units together, time is needed to prepare the reagents for the condensation and to purify the products after each condensation step, as well as to isolate and purify the final, desired nucleotide sequence.

We set out to develop an automated procedure that would be complete from synthesis to isolation and that could be used in any laboratory by technical staff.

History of DNA Synthesis

The first successful synthesis of a small piece of DNA containing the natural linkage (phosphate ester bridge from the 3'-oxygen of one unit to the 5'-oxygen of the next) (Fig. 1) was achieved by Michelson and Todd (*2*) in 1955. The procedure they used was reported to have limitations and was not widely used. The major developmental thrust in DNA synthesis came through the efforts of the Khorana group during the early

K. K. Ogilvie is professor of chemistry, McGill University, Montreal, Canada H3A 2K6; G. Alvarado-Urbina, G. M. Sathe, W.-C. Liu, M. F. Gillen, P. D. Duck, and R. Bender are members of ens Bio Logicals, Inc., 20 Victoria Street, Toronto, Canada M5C 2N8.

0036-8075/81/1016-0270$01.00/0

1960's into the mid-1970's (*3*). The Khorana team established the general nature of the problem of DNA synthesis and developed procedures that enabled others to investigate it.

The Khorana method (Fig. 2) was called a diester method. A suitably protected mononucleotide was condensed with a suitably protected nucleoside to form a dimer. By removal of the protecting group from either end of the chain, the process could be repeated and the chain extended. The protecting groups they developed are still in use in DNA synthesis. A major disadvantage of the method is that it is very slow. Times ranging from several hours to several days are required for condensations. The product at each step is a diester of phosphoric acid, which means that the third functional group on phosphate exists as an ionized acid in the form of a salt. The isolation and purification of salts require special techniques that are limited to small quantities and often require up to 2 weeks for a single purification. The next major breakthrough came in the mid-1960's with the introduction of the successful triester method (*4*). By this procedure (Fig. 2), the direct product of condensation is a triester of phosphoric acid. The product is a neutral organic molecule and can be isolated in large quantities by techniques widely used in organic chemistry, such as chromatography on silica gel. Condensation times were greatly speeded up.

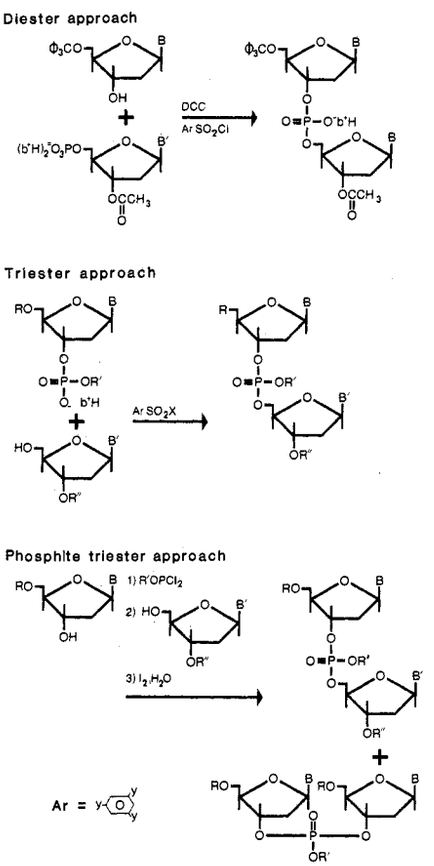


Fig. 2. General methods for the chemical synthesis of DNA. ϕ represents the phenyl groups; b is any base; B and B′ are bases in protected form; R, R′, and R″ are protecting groups; DCC is $C_6H_{11}N{=}C{=}NC_6H_{11}$; y is CH_3- or $(CH_3)_2CH-$; and X is chlorine, triazole, tetrazole, or nitroimidazole.

The basic triester method was expanded and improved by numerous research groups (*5–7*). Even though the practical improvements in the triester method during the 1970's were often described as new, novel, or modified (and sometimes all three), the synthesis of an oligodeoxynucleotide still required a highly trained and skilled chemist, and very little improvement was made in the rapid purification of the final sequence.

In 1975, Letsinger *et al.* (*8*) provided the basis for the next stage in the evolution of a chemical method for rapid nucleotide synthesis. With the phosphite triester procedure (Fig. 2), a dinucleotide could be successfully prepared in less than 1 hour at −78°C, starting from the protected nucleosides (*8*). No unusual or hidden manipulations were involved.

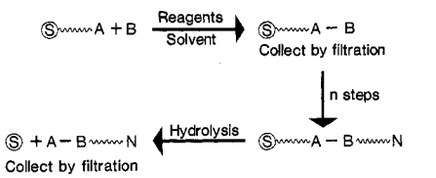
One of us (K.K.O.) began to use the phosphite method in our work on the synthesis of ribonucleotides (*9*). We were impressed by the speed and cleanliness of the reactions of this procedure, as well as by the ease with which products could be isolated. The procedure performed well even for the synthesis of long chains (*10*). After testing every practical nucleotide synthetic procedure previously proposed in the literature, we realized that the phosphite method offered the greatest promise for development of a procedure for the rapid synthesis of DNA. The phosphite method also offered a practical possibility for a solid support synthesis, which proved to be the key to automation of DNA synthesis.

Solid Support Method

In the stepwise assembly of a linear polymeric material such as a protein or nucleic acid (DNA), a solid support method offers the most advantageous route to automation because it eliminates separation and purification difficulties (*11*) (Fig. 3). The first unit, A, of a growing chain is attached to an insoluble macroscopic solid material such as polystyrene or silica beads. The second unit, B, and all necessary coupling reagents are dissolved in a suitable solvent and mixed with the solid support.

Reaction occurs at the A units on the surface of the solid, and the chain is extended to AB. At the end of the reaction period, all unused reagents are washed off with the solvent. This procedure allows a large excess of solution reagent to be used. Thus it it possible to drive the yield of the reaction toward quantitative levels. At the end of a predetermined number of chain extension steps, the chain AB . . . N is removed from the solid and collected in solution.

It is important in such a method that the condensation yields are as near 100 percent as possible. A 90 percent yield, usually considered excellent in any chemical synthesis, leads to an undesirable mixture of products on the support, and after ten such steps the desired product would constitute only 35 percent of the material attached to the support.



Condensation yield (%)	Products after two steps (S)-ABC	(S)-AB	(S)-AC	(S)-A	Yield of product after ten steps
90	81	9	9	1	35
95	90.2	4.8	4.8	0.2	60
99	98	1	1	0.01	90

Fig. 3. The general concept of a solid support synthesis of a polymeric molecule such as DNA.

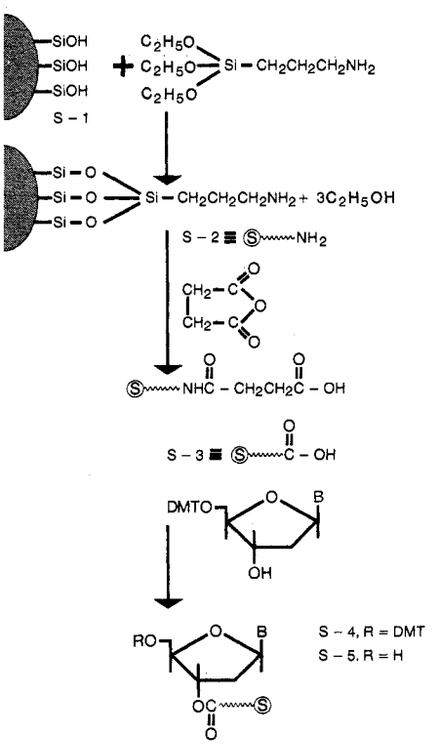
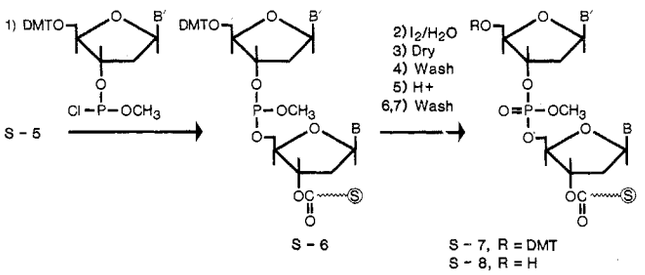


Fig. 4. Preparation of the solid support S-5 used in the automated synthesis of DNA. B represents thymine, *N*-benzoylcytosine, *N*-benzoyladenine, or *N*-isobutyrylguanine.



Automated chain extension cycle

Step	Process	Time (minutes)
1	Condensation	7
2	Oxidation	3
3	Drying	4
4	Washing ($CHCl_3$)	3
5	Detritylation	3
6	Washing ($CGCl_3$)	3
7	Washing (pyridine)	7
	Total	30 minutes

Time to prepare dodecamer, 5.5 hours

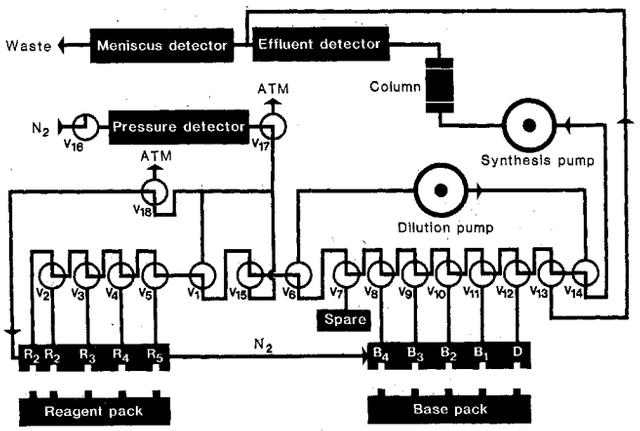


Fig. 5 (left). The steps involved in each cycle of the automated synthesis of DNA. Fig. 6 (right). Schematic diagram of the DNA/RNA Synthesizer used in the automated synthesis of gene fragments. *ATM*, vent to atmosphere; B_1 to B_4, solutions of the four protected nucleoside phosphites; *D*, detritylating solution; R_2, pyridine; R_3, chloroform; R_4, drying reagent; R_5, oxidizing reagent; V_1 to V_{18}, valves.

This often presents a problem in isolating the desired product at the end of the synthesis. These problems can be so great as to negate the other advantages of the method. However, if the yield is 99 percent at each step, the desired product constitutes 90 percent of the material at the end of ten cycles (Fig. 3). Thus a satisfactory solid support procedure depends on a very clean high-yielding condensation step.

None of the methods developed before the Letsinger phosphite procedure seemed adequate to meet the necessary requirements for an effective solid support synthesis. In fact the only limitation to the phosphite method in solution, the presence of the undesired 3′–3′ linked isomer (Fig. 2), is not a factor in a solid support system.

The Solid Phase

Numerous types of solid supports have been investigated for both proteins and nucleic acids. Most of these solids are of the polystyrene type which change volume (swell) in organic solvents; this causes severe operational difficulties. However, in the early 1970's silica was introduced as a solid support for synthesis (*12, 13*). The surface of silica is easy to derivatize, and particles are dimensionally stable in organic solvents. In 1977, Tundo (*14*) expanded the possibilities of surface derivatization of silica.

We recognized that the combination of the phosphite procedure and a silica solid phase could provide a practical method for the solid phase synthesis of oligonucleotides (*15*). This possibility was also recognized by Matteucci and Caruthers (*16*), who first reported the method. Our groups have been developing the chemistry of this method independently.

In this article, we present the chemistry and the totally automated procedure that we have developed. Our method includes a simple, economical, and practical method for isolating the desired final product. The total package is aimed at the synthesis, isolation, and purification of DNA segments up to 14 units in length and therefore represents another tool in the research arsenal of scientific investigators.

The Chemistry

The solid support is silica, and the preparation of the reactive solid phase is very similar to that previously described by Tundo (*14*) (Fig. 4). Silica gel (all gels tested give similar results) is surface activated by refluxing it with HCl (*15, 17*). This provides the maximum number of active SiOH groups on the surface. This activated gel (S-1 in Fig. 4) is derivatized by refluxing with γ-aminopropyltriethoxysilane in toluene (*14*) to provide the functionalized resin S-2. The amino group can now be derivatized in a number of ways (*14–16, 18*), including a standard procedure with succinic anhydride to give the activated resin S-3. The final step in preparing the resin for use in the automated procedure requires condensing a protected nucleoside to S-3, with dicyclohexylcarbodiimide (DCC) used as the condensing agent in pyridine. When diethylaminopyridine is used as catalyst (*15*), the resulting resin S-4 contains approximately 0.1 millimole of nucleoside per gram of resin. On removal of the dimethoxytrityl group (DMT) from the nucleoside with mild acid, the resin S-5 that is obtained is ready for chain extension. The resin S-5 (200 milligrams) is packed into a column (12 by 3 millimeters, internal volume), which snaps into the synthesizer.

Once the resin S-5 has been attached to the instrument, the chain is constructed as described in detail below. The actual chemical conversions that occur in the instrument are outlined in Fig. 5. A solution containing the active phosphorylating reagent 1 is passed through the column to produce the coupled phosphite product S-6. The phosphite intermediate is oxidized to a phosphate (S-7) with aqueous iodine. Water is removed from the resin with phenylisocyanate, a water scavenger. The resin is then washed with chloroform to remove excess reagents. The DMT group is removed with mild acid to give S-8, which, after being washed with chloroform and pyridine, is ready for the next cycle of chain extension. The total time for each cycle is 30 minutes. Thus it requires 5½ hours to produce a 12-unit nucleotide chain.

The DNA/RNA Synthesizer

The instrument (Fig. 6) consists of a series of pneumatic valves that are connected simultaneously to a series of reagent bottles and a series of solenoids. The solenoids respond to a microprocessor signal to open or close a particular valve. The solenoids are in turn controlled by a programmer that contains the instructions for the synthesis. When a valve is open to a particular reagent bottle, that reagent is pumped through the column containing the resin; reagents are pumped through the column in a continuous manner at a fixed rate of 5 milliliters per minute. The programmer has a fixed time program for each rea-

gent used. The time and order for each reagent in a cycle are fixed. The only variable is the nucleoside to be added in the first step of each cycle. The instruction for the nucleoside, in the form of the sequence to be constructed, is typed into the program before the synthesis begins. The programmer automatically records the sequence and opens the appropriate valve to the nucleoside derivatives at the correct point in the sequence. Once the reagent bottles have been filled and the sequence is typed into the program, the instruction to start is given and the synthesis continues uninterrupted until it is finished. After the last cycle is completed, the resin is removed from the column and the nucleotide chain is recovered as described below.

Chemistry in Detail

Preparation of the protected nucleoside derivative. The protected deoxynucleoside (5 mmole) (*7*) is dissolved in 20 ml of dioxane. This solution is added dropwise to a previously prepared solution of methyldichlorophosphite (0.45 ml) and collidine (3.5 ml) in dioxane (20 ml). All preparations are carried out at room temperature. The final solution is diluted with 500 ml of pyridine. The resulting solution acts as the stock solution of reagents.

Iodine oxidation. A stock solution containing 0.75 g of iodine in 100 ml of a mixture of tetrahydrofuran and water (3:1) is prepared and used as such.

Phenylisocyanate drying. A stock solution containing 1 ml of phenylisocyanate per 100 ml of pyridine is prepared and used as such.

Removal of the dimethoxytrityl group. A solution containing 3 percent trichloroacetic acid in chloroform is prepared and used as such.

Removal of the synthesized nucleotide from the support. At the end of the synthesis, the instrument cycles through a down sequence. The resin is removed and placed in a sealed tube containing concentrated ammonium hydroxide and pyridine (4:1). The system is heated at 50°C for 12 hours, after which the solution is collected by filtration or centrifugation. The resin is washed twice with 0.1 ml of a mixture of water and ethanol (3:1). The combined filtrate and washings are concentrated to dryness. The residue is then dissolved in 0.2 ml of water and ethanol (3:1). This solution is applied to four thin-layer chromatography (TLC) analytical sheets (Brinkmann Polygram SilG/UV 254), which are developed for 12 hours in a mixture of ammonium hydroxide, *n*-propanol, and water (35:55:10). The product band is by far the major ultraviolet-absorbing band and is easily detected. A tetradecamer has a relative mobility (R_F) value typically between 0.1 and 0.2. The band containing the product is scraped off of the plate. The product is washed from the silica with a mixture of water and ethanol (3:1) (two washings of 0.5 ml each). The washings are concentrated, and the residue is purified further if desired.

The product at this point is generally quite pure (~ 90 percent). The product can be completely purified by rechromatography on TLC plates or by high-performance liquid chromatography (HPLC). Sequences containing 10 to 15 units can be rapidly purified on anion exchange HPLC columns [Brownlee Anion Exchanger AX-10A, Lichrosorb AN (25 centimeters by 4.6 mm), or Whatman Anion Exchanger, Partisil PXS10/25 SAX (25 cm by 4.6 mm)] with solvent A, 0.2*M* triethylammonium acetate, *p*H 2.9, and solvent B, 0.5*M* triethylammonium acetate, *p*H 3.5. Separations are obtained with a 60-minute linear gradient of 0 to 100 percent solvent B. For a flow rate of 4 ml/min, retention times for units 10 to 14 bases long vary between 15 and 25 minutes. Figure 7 shows the HPLC characteristics of the dodecamer d-CCCGAATTCGGG after one TLC purification and after HPLC purification. This product was prepared with the standard machine cycle described above. After HPLC purification, the dodecamer was characterized by gel sizing (Fig. 8). The HPLC clearly establishes purity, and gel electrophoresis determines the size of the molecule. The precise sequence was determined with a modification of the Maxam-Gilbert technique (*19*).

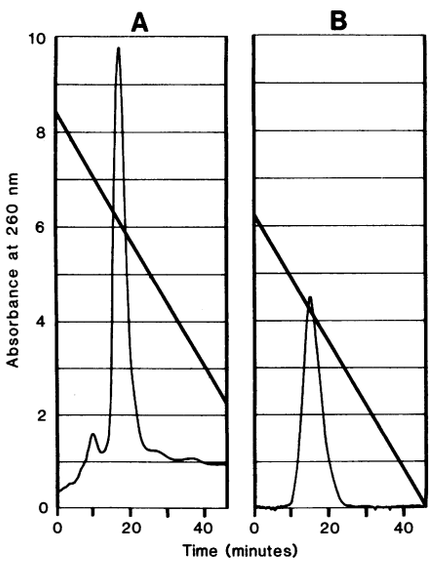


Fig. 7. (A) An HPLC tracing of the major band from the TLC separation of the reaction products obtained in the preparation of the dodecamer d-CCCGAATTCGGG (for TLC and HPLC solvent systems, see text). (B) An HPLC tracing of the pure dodecamer obtained from HPLC for yield determination.

Table 1. Yields of purified molecules determined by ultraviolet absorbance at 260 nm.

Nucleotide sequence	Length (number of nucleotides)	Overall yield (%)	Average yield at each step (%)
d-CTGTCTTTGCT	11	93	99.3
d-CCCGAATTCGGG	12	76	97.5
d-ACAACAAAAAGCAA	14	70	97.3

Yields of Product

The yields of purified (by HPLC) molecules prepared as described above are shown in Table 1. Three molecules are described, an undecamer, the above-mentioned dodecamer, and a tetradecamer. The yields are determined by the ultraviolet absorbance of the pure compound (from HPLC) at 260 nanometers and are based on the amount of the first nucleoside bound to the polymer; these

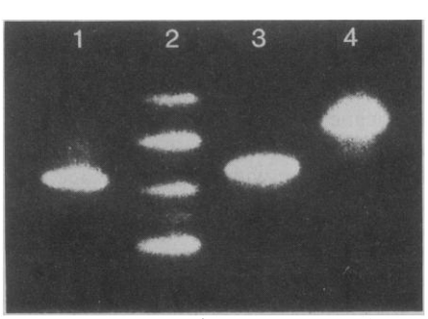


Fig. 8. Sizing gel results on the purified undecamer d-CTGTCTTTGCT (track 1), the dodecamer d-CCCGAATTCGGG (track 3), and the tetradecamer d-ACAACAAAAAGCAA (track 4). The results were obtained from a 24 percent polyacrylamide gel at 800 volts for 1½ to 2 hours, 40° to 50°C, under standard conditions. Track 2 is a standard mixture containing deoxynucleotides of 8, 10, 12, and 14 units.

yields represent overall conversion yields. The tetradecamer contains 11 deoxyadenosine units. Deoxyadenosine is the nucleoside most susceptible to glycoside bond hydrolysis during acid cleavage of the DMT group during each cycle. Figure 8 shows the sizing gel results of all three molecules purified by HPLC as previously described.

Conclusion

The synthesis of oligonucleotides up to 14 units in length is now a routine procedure. The DNA/RNA Synthesizer and reagent kits for the preparation of units up to 14 in length are commercially available. The synthesis can be carried out in any laboratory with the use of the synthesizer and the reagent kits, and no previous experience in nucleotide synthesis is necessary. The reagent kit is attached to the synthesizer; the sequence is typed into the memory; the start signal is given; and the synthesis continues uninterrupted until it is complete.

This basic procedure also applies to the synthesis of RNA sequences (*20*).

References and Notes

1. For a general review of these and related areas see the recombinant DNA issue of *Science*, 19 September 1980.
2. A. M. Michelson and A. R. Todd, *J. Chem. Soc.* **1955**, 2632 (1955).
3. H. G. Khorana, *Some Recent Developments in the Chemistry of Phosphate Esters of Biological Interest* (Wiley, New York, 1961), pp. 93–140; *Bioorg. Chem.* **1**, 351 (1978).
4. R. L. Letsinger and K. K. Ogilvie, *J. Am. Chem. Soc.* **89**, 4801 (1967); R. L. Letsinger and V. Mahadevan, *ibid.* **88**, 5319 (1966).
5. F. Eckstein and I. Rizk, *Angew. Chem.* **79**, 684 (1967).
6. C. B. Reese and R. Saffhill, *Chem. Commun.* (1968), p. 767.
7. S. A. Narang, R. Brousseau, H. M. Hsiung, J. J. Michniewicz, *Methods Enzymol.* **65**, 610 (1980).
8. R. L. Letsinger, J. L. Finnan, G. A. Heavner, W. B. Lunsford, *J. Am. Chem. Soc.* **97**, 3278 (1975).
9. K. K. Ogilvie, N. Theriault, K. L. Sadana, *ibid.* **99**, 7741 (1977).
10. K. K. Ogilvie and N. Y. Theriault, *Can. J. Chem.* **57**, 3140 (1979); K. K. Ogilvie and M. J. Nemer, *ibid.* **58**, 1389 (1980).
11. E. C. Blossey and D. C. Neckers, Eds., *Solid Phase Synthesis* (Academic Press, New York, 1975).
12. W. Parr and K. Grohmann, *Tetrahedron Lett.* (1971), p. 2633.
13. H. Koster, *ibid.* (1972), p. 1527.
14. P. Tundo, *Chem. Commun.* (1977), p. 641.
15. K. K. Ogilvie and M. J. Nemer, *Tetrahedron Lett.* (1980), p. 4159.
16. M. D. Matteucci and M. H. Caruthers, *ibid.*, p. 719.
17. J. F. Fritz and J. N. King, *Anal. Chem.* **48**, 570 (1976).
18. P. Tundo and P. Venturello, *J. Am. Chem. Soc.* **101**, 6066 (1979).
19. A. H. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
20. K. K. Ogilvie, in preparation.

21. We thank K. Deugau for the sequence analysis, and R. May and L. Rothwell for technical assistance.

Airborne Microwave Remote-Sensing Measurements of Hurricane Allen

W. Linwood Jones, Peter G. Black
Victor E. Delnore, Calvin T. Swift

In August 1980 the Langley Research Center's active and passive microwave sensors were flown into Hurricane Allen on a NOAA aircraft. This article presents the surface wind speed, wind direction, and rain rate inferred from the microwave sensors and compares these findings with data obtained by a P-3 aircraft that flew near the top of the hurricane planetary boundary layer. The results demonstrate the potential for safe and reliable hurricane measurements by aircraft flying at medium to high altitudes.

Hurricanes have been called the greatest storms on earth (*1*). Over the years considerable effort has been devoted to studying their dynamics and developing models for forecasting their intensity and path of travel. Much attention has been focused on the hurricane planetary boundary layer because it is there that convective processes which drive the hurricane's circulation transport heat and moisture from the ocean. Furthermore, it is the winds at the surface that cause the damage at landfall, either directly, or indirectly by the storm surge. The dynamics of these surface winds are of paramount interest to researchers and forecasters.

Present methods for estimating hurricane-force surface winds often have limited accuracy. For example, estimates derived from aircraft reconnaissance are frequently based on Beaufort state of the sea relations, but the validity of using these relations for hurricane-force winds has not been established. The sustained flight-level wind at 3300 meters and minimum surface pressures (*2*) are also used to infer maximum surface winds. To date, the only reliable method of obtaining the surface wind speed has been low-level penetration of the hurricane by aircraft equipped with inertial navigation systems.

To improve the reliability of surface wind estimates by applying microwave remote-sensing technology, a cooperative program was begun by the National Oceanic and Atmospheric Administration (NOAA) and the National Aeronautics and Space Administration (NASA).

Background

The use of passive and active microwave techniques for remote sensing of ocean surface winds is well documented (*3*). The brightness temperature of the ocean is strongly correlated with wind speed since wind creates roughness, foam, and breaking waves (whitecaps) (*4–6*). Radar backscatter from the sea can be used to measure wind speed (through the proportionality of the Bragg off-nadir radar return to the wind-dependent spectrum of water waves a few centimeters long) and wind direction (through the anisotropic scattering characteristic of the sea surface) (*7–9*).

The first airborne passive and active observations of a hurricane were conducted during Hurricane Ava (*10*). However, in this case, the L-band radar (1.3 gigahertz) observations were used only to generate wave images. This type of observation was later quantified in 1976 (*11*), demonstrating that L-band radar

When this article was written, W. Linwood Jones and Calvin T. Swift were aerospace technologists at the NASA Langley Research Center, Hampton, Virginia 23665. Jones is now project manager, Communications Systems, General Electric Valley Forge Space Center, King of Prussia, Pennsylvania 19406; Swift is now professor of electrical and computer engineering, University of Massachusetts, Amherst 01003. Peter G. Black is a meteorologist at the Hurricane Research Laboratory, National Oceanic and Atmospheric Administration, Coral Gables, Florida 33145. Victor E. Delnore is a senior environmental engineer at Kentron International, Hampton, Virginia 23665.

0036-8075/81/1016-0274$01.00/0