

Crystallization of Isoelectrically Homogeneous Cholera Toxin<sup>†</sup>Brenda D. Spangler<sup>\*,‡</sup> and Edwin M. Westbrook<sup>‡,§</sup>

Biological, Environmental, and Medical Research Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, and Department of Biochemistry and Molecular Biology, The University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

Received July 7, 1988; Revised Manuscript Received September 12, 1988

**ABSTRACT:** Past difficulty in growing good crystals of cholera toxin has prevented the study of the crystal structure of this important protein. We have determined that failure of cholera toxin to crystallize well has been due to its heterogeneity. We have now succeeded in overcoming the problem by isolating a single isoelectric variant of this oligomeric protein (one A subunit and five B subunits). Cholera toxin purified by our procedure readily forms large single crystals. The crystal form (space group  $P2_1$ ,  $a = 73.0$  Å,  $b = 92.2$  Å,  $c = 60.6$  Å,  $\beta = 106.4^\circ$ , one molecule in the asymmetric unit) has been described previously [Sigler et al. (1977) *Science (Washington, D.C.)* 197, 1277-1278]. We have recorded data from native crystals of cholera toxin to 3.0-Å resolution with our electronic area detectors. With these data, we have found the orientation of a 5-fold symmetry axis within these crystals, perpendicular to the screw dyad of the crystal. We are now determining the crystal structure of cholera toxin by a combination of multiple heavy-atom isomorphous replacement and density modification techniques, making use of rotational 5-fold averaging of the B subunits.

**C**holera toxin is one of a group of bacterial protein toxins that includes diphtheria, pertussis, shigella, tetanus, botulinum, and anthrax toxins and *Pseudomonas* exotoxin A [for a review, see Middlebrook and Dorland (1984)]. Each toxin consists of a subunit or peptide fragment responsible for enzymatic function (typically named fragment A or subunit A) coupled with one or more components (named fragment B or subunit B) associated with cell surface receptor binding and subsequent internalization (transmembrane transport) of the enzymatic component. Several bacterial toxins are ADP-ribosyltransferases, with protein substrates, and recent observations suggest that the enzymatic components of ADP-ribosylating toxins share some primary structure similarities (Locht & Keith, 1986; Capiou et al., 1986; Burns et al., 1987). Many of the substrates ADP-ribosylated by these bacterial protein toxins are "G-proteins" (Gilman, 1987), which are involved in signal transduction (the passage of information across membranes), and ADP-ribosylation has been recognized as one of the more significant posttranslational modifications of proteins (Collier, 1982; Ueda & Hayaishi, 1985).

Cholera toxin, the enterotoxin of *Vibrio cholerae*, is an oligomer of a single A subunit ( $M_r$  27 234) and five B subunits ( $M_r$  11 677 each), which are arranged as a pentameric ring with apparent 5-fold symmetry (Finkelstein & LoSpalluto, 1970; Gill, 1976, 1977; van Heyningen, 1976; Mekalanos et al., 1978; Dwyer & Bloomfield, 1982; Ribi et al., 1988). Subunit A, synthesized as a single polypeptide, is proteolytically nicked during secretion from the bacterium to give rise to two disulfide-linked polypeptides:  $A_1$  ( $M_r$  21 826) and  $A_2$  ( $M_r$  5407) (Gill, 1977; Lockman & Kaper, 1983; Mekalanos et al., 1979). It is the  $A_1$  fragment of subunit A, released by disulfide reduction, that acts enzymatically within the target cell as an ADP-ribosyltransferase (Gill & Coburn, 1987; Gill

& Meren, 1983; Kahn & Gilman, 1986). The entire oligomer, of  $M_r$  85 620, is required for toxic behavior. Choleragenoid, the intact pentamer of B subunits, interacts with a ganglioside  $GM_1$  membrane receptor (Van Heyningen, 1974; Dwyer & Bloomfield, 1985) but cannot activate adenylyl cyclase, whereas the  $A_1$  subunit alone does not enter the cell (Wisniewski & Bramhall, 1981).

The study of this system at the molecular level requires a crystal structure of the three-dimensional conformation of cholera toxin. Although cholera toxin crystals were reported 11 years ago (Sigler et al., 1977), that earlier study was halted by an inability to reproducibly obtain suitable crystals (P. B. Sigler, personal communication). Every sample of cholera toxin we have obtained and examined has been microheterogeneous. We believe that this microheterogeneity is the cause of poor crystallization behavior of cholera toxin in the past. We report here that we have isolated a single isoelectric variant of the toxin oligomer that crystallizes readily and reproducibly. Isoelectric homogeneity resulting in improved crystallization behavior has been observed before this for other proteins (Bott et al., 1982; McPherson, 1988; Geige et al., 1986).

## EXPERIMENTAL PROCEDURES

**Toxin.** Lyophilized, intact cholera toxin and choleragenoid were obtained either from List Biological Laboratories (Campbell, CA) or from Schwarz/Mann, Biotech. Freshly prepared unlyophilized toxin was a gift from Dr. Linda Shoer of List Biological Laboratories. Toxin from the El Tor strain 3083 was a gift from Dr. Richard Finkelstein. Toxin, lot 0972, prepared by Dr. Richard Finkelstein and dating from approximately 1972, was obtained from the National Institute of Allergy and Infectious Diseases, Bethesda, MD, through Dr. Robert Edelman.

We dissociated toxin or choleragenoid into monomeric subunits by adding 0.1 M HCl to solutions of protein to reach a pH of between 3 and 3.5 and then incubated these solutions at 37 °C for 30–60 min (Gill, 1976).

**Gel Electrophoresis.** Nondenaturing continuous polyacrylamide gel electrophoresis (PAGE)<sup>1</sup> was performed with

<sup>†</sup>Supported by the U.S. Department of Energy, Office of Health and Environmental Research, under Contract W-31-109-ENG-38.

<sup>\*</sup>To whom correspondence should be addressed.

<sup>‡</sup>Argonne National Laboratory.

<sup>§</sup>The University of Chicago.

8% polyacrylamide gels in barbital buffer (Paragon B-2 barbital buffer from Beckman). Buffer was made up to a final concentration of 5 mM diethylbarbituric acid and 25 mM sodium diethylbarbiturate, adjusted to pH 7.6 with HCl. These gels were stained with Xylene Cyanine Brilliant G, using trichloroacetic acid (TCA) as a precipitant to minimize protein diffusion in the gel during destaining (Blakesley & Boezi, 1977; Andrews, 1986). Nondenaturing discontinuous PAGE was performed with Tris buffer, in the Pharmacia PHAST system with precast 8–25% or 10–15% polyacrylamide gradient gels, following the manufacturer's instructions.

Sodium dodecyl sulfate (SDS)–PAGE was performed either on the Pharmacia PHAST system, with precast 8–25%, or 10–15%, polyacrylamide gradient PHAST gels and a Tris–SDS buffer according to the manufacturer's instructions, or by the method of Weber and Osborn (1969), in the presence or absence of 2-mercaptoethanol, as noted. Samples to be run were mixed with SDS immediately prior to being loaded onto the gel. For electrophoresis of completely denatured protein, these samples were heated for 30 s in a microwave oven, in the presence of SDS, before they were loaded onto the gel. This process separated the cholera toxin oligomer into its monomeric constituents. Cholera toxin samples in SDS that are not heated prior to loading separate into monomeric A subunit and pentamers of B subunits (Gill, 1976). All SDS–PAGE gels were stained with Coomassie Blue in a solution of 2-propanol/acetic acid/water (10:4:40 v/v).

Analytical isoelectric focusing (IEF) was performed on the Pharmacia PHAST system or on an LKB Ultraphor with precast gels, pH 3–9, according to manufacturer's instructions. Preparative IEF was performed with pH 5–8 ampholytes and Ultradex gel resin from LKB. Bands were eluted with 50 mM Tris–HCl, pH 7.5, and 1 mM EDTA. Proteins were concentrated and ampholytes removed by repeated washing and re-concentration of the samples with Tris buffer in a Centricon 10 microconcentrator (Amicon). Washes were repeated until no ampholytes were detected in the eluent by virtue of the absence of absorbance at 280 nm when read against buffer. Microdensitometry of analytical IEF gels was performed with a Shimadzu dual-wavelength TLC scanner, equipped with a recording densitometer.

**Column Chromatography.** Gel filtration was performed in a 1.25 × 50 cm column, with Sephadex G-100 resin. Protein was eluted with 50 mM Tris–HCl, pH 7.5, containing 1 mM EDTA. We evaluated chromatographic results by SDS–PAGE of unheated aliquots of each fraction. Protein bands were pooled and concentrated in a Centricon 30 (Amicon) to a final concentration of 5–10 mg/mL.

**Ion-Exchange Chromatography.** Buffer A consisted of 50 mM Tris–HCl, pH 7.70, and buffer B was 1.0 M NaCl/50 mM Tris–HCl, pH 7.70. Samples were dialyzed against buffer A at 4 °C overnight and then applied to a Pharmacia HR5/5 Mono Q column at 4 °C equilibrated with buffer A on a Pharmacia FPLC system. After an isocratic wash, the protein was eluted by a linear gradient from 0–10% of buffer B (0–100 mM NaCl), followed by a 300 mM salt wash and an extended wash with buffer A. Fractions were analyzed by IEF as described above. Peaks obtained in successive runs were combined, concentrated to 5–8 mg/mL in a Centricon 10

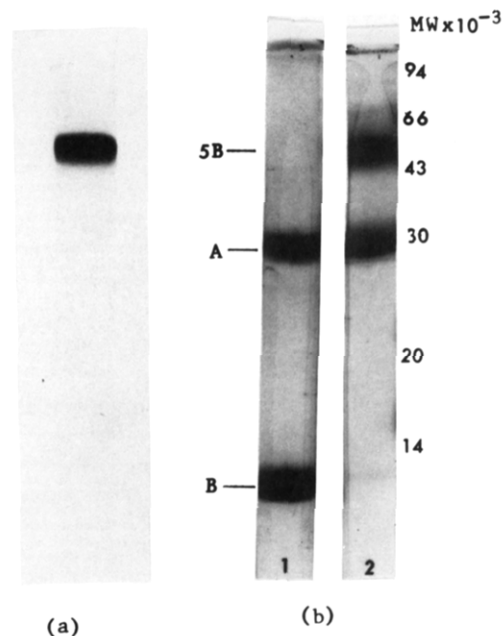


FIGURE 1: Analysis of cholera toxin. (a) Nondenaturing PAGE in barbital buffer, pH 7.6, of cholera toxin. All samples examined appear identical. (b) SDS–PAGE of cholera toxin. Lane 1, cholera toxin sample was heated in the presence of SDS. Subunit A, 28 000 Da, and monomeric subunit B, 11 600 Da. Lane 2, sample, in the presence of SDS, not heated prior to loading onto the gel. Subunit A, 28 000 Da, and pentameric subunit B, 58 000 Da. A small amount of monomeric B is visible in lane 2.

(Amicon), and stored at 4 °C.

**Crystallization.** Crystals were obtained by the vapor diffusion method (Hampel & Bock, 1970). Ten microliters of protein solution (5–10 mg/mL in 50 mM Tris–HCl, pH 7.7, and 100 mM NaCl) was mixed with an equal amount of equilibration buffer (14% PEG 4000 in 50 mM MOPS, pH 6.7). In many cases, crystallization occurred spontaneously within 48 h. Crystal seeding of drops (Thaller et al., 1985) was also used to initiate crystallization.

**Crystal Diffraction Data Collection.** Complete three-dimensional X-ray diffraction data were collected to 3.0-Å resolution with a Nicolet/Xentronics electronic area detector (Durbin et al., 1986) at a crystal to detector distance of 11.5 cm. A total of 900 data frames, each of 0.2° rotation around the vertical axis and consisting of an X-ray exposure of 320 s, were recorded from three different orientations of the crystal and processed by the XENGEN package of A. Howard (Howard et al., 1987).

**Rotation Function.** The rotation function (Rossmann & Blow, 1962) was calculated by the "fast" algorithm of Crowther (1972) and expressed in spherical polar notation (Tanaka, 1977). The program used for this calculation was written by one of us (E.M.W.). We calculated the self-rotation function, producing a three-dimensional map of rotational correlation within the Patterson function of the native cholera toxin crystal.

## RESULTS

**Gel Electrophoresis and Isoelectric Focusing.** Nondenaturing gels of fresh cholera toxin, run at pH 7.6 in barbital buffer (Figure 1a), produced a single band even when the gels were extremely overloaded (up to 40 µg of protein/tube gel). In the presence of SDS (Figure 1b), unheated preparations of fresh toxin contained subunit A and pentamers of subunit B, as well as varying amounts of monomeric B subunits. Heated samples contained monomeric A and B subunits. On

<sup>1</sup> Abbreviations: SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; EDTA, ethylenediaminetetraacetic acid; TCA, trichloroacetic acid; Tris, tris(hydroxymethyl)aminomethane; MOPS, 3-(*N*-morpholino)propanesulfonic acid; IEF, isoelectric focusing; PEG 4000, poly(ethylene glycol) of average molecular weight 3350;  $R$ -sym =  $\frac{h_i^{(h)}(h) - I_i^{(h)}(h)}{h_i^{(h)}(h)}$  where  $\langle I(h) \rangle$  is the mean of all measurements of reflection  $h$  and  $I_i^{(h)}(h)$  is the  $i$ th measurement of reflection  $h$ .

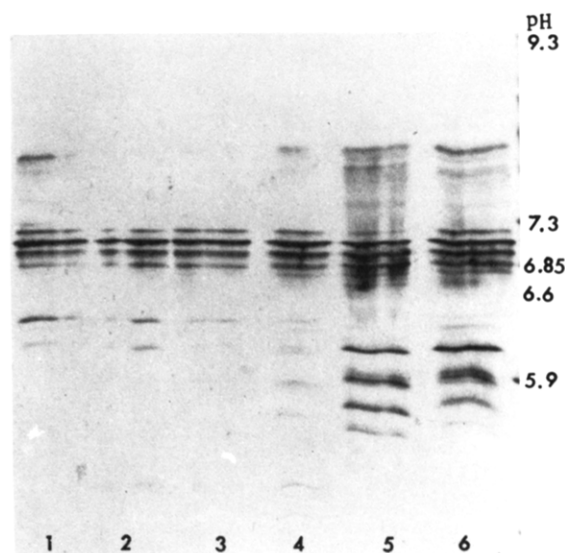


FIGURE 2: Analytical IEF (pH 3–9) of cholera toxin samples. Lane 1, 1 month old, unlyophilized, from List Labs. Lane 2, 1 year old, unlyophilized, from List Labs. Lane 3, 2 years old, lyophilized and freshly resuspended, from List Labs. Lane 4, unknown age, lyophilized and freshly resuspended, from Schwarz/Mann. Lane 5, El Tor strain 3083, 4 years old, stored in solution at 4 °C, from R. Finkelstein. Lane 6, lot 0972, 15 years old, lyophilized, stored at 4 °C and then freshly resuspended, from NIH.

Tris-buffered SDS-PAGE gels of either cholera toxin or cholera-genoid, we observed small amounts of high molecular weight aggregates with molecular weights of 110 000 and above, which we could demonstrate, by dissociation in acid, consisted only of B subunits (data not shown). On the basis of our observations, from PAGE in barbital buffer at pH 7.6 and SDS-PAGE, we concluded that there was no detectable contaminating protein in our fresh toxin preparations. However, nondenaturing electrophoresis run with Tris buffers at higher pH (8–9) on 8% gels or PHAST gradient gels produced one additional band just beneath the primary band. This observation suggested that a charge variant might exist in the toxin that can only be observed at higher pH or that exposure of the cholera toxin to these conditions may induce a charge variant, perhaps through deamidation.

Analytical IEF of all our cholera toxin samples produced at least six bands between pH 6.7 and 7.1 (Figure 2) with the major band at pH 7.0. There is a difference of approximately 0.1 pH unit between each of our bands. This banding pattern was observed in all the samples we tested, from several sources and ranging in age from 10 days to 15 years (Figure 2). In addition to these bands containing the bulk of our protein sample, minor bands, possibly subunit A, were seen at pH 6.1 and at pH 7.8, the reported *pI* of cholera-genoid (Finkelstein & LoSpalluto, 1970).

We observed the same pattern when we performed preparative IEF. Each of three bands (pH 6.7, 6.9, and 7.0) isolated by preparative IEF was found to contain only the 27 000-Da subunit A band and a 58 000-Da pentameric B band, when run as unheated samples on SDS-polyacrylamide gels, and bands of 27 000 and 12 000 Da, when run as heated samples in the presence of SDS, as would be expected from intact toxin.

Each of the bands isolated by preparative IEF refocused as a single band on analytical IEF. In each such refocusing, we also observed a minor band, probably due to imperfect isolation of the major band. We therefore concluded that these isoelectric variants were not due to variable ampholyte binding and that the isoelectric point of each variant was a stable property of the protein.

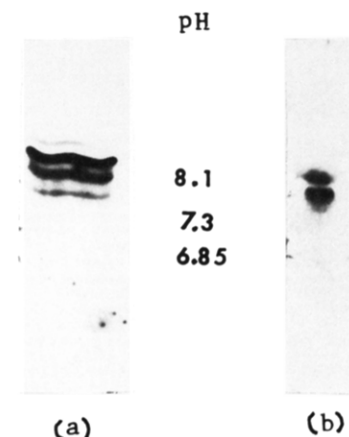


FIGURE 3: (a) Analytical IEF of pentameric B subunits (cholera-genoid). (b) Analytical IEF of monomeric B subunits, prepared by acid dissociation of cholera-genoid.

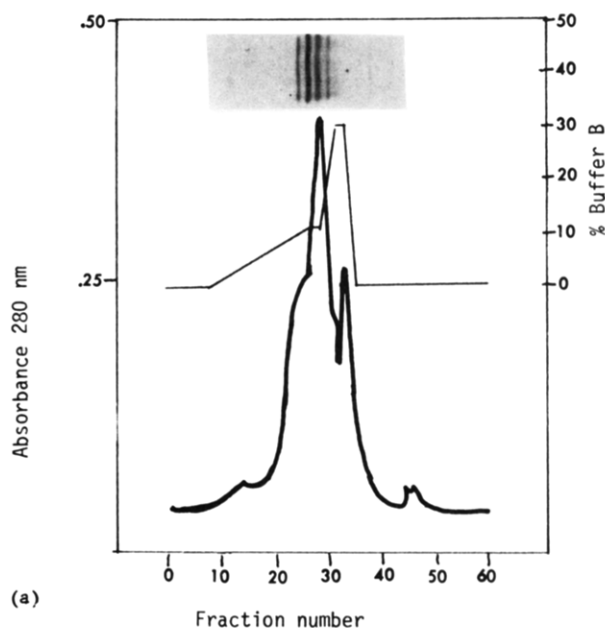
Analytical IEF of monomeric B subunits, obtained through dissociation of cholera-genoid by treatment with acid, produced two separate bands, approximately 0.3 pH unit apart (Figure 3). This banding pattern was also found in IEF gels of intact toxin, dissociated in the same way. We therefore conclude that there are two isoelectric variants of the B subunit in our samples of cholera toxin. The quantitative distribution between these two bands appears to be roughly 3:1 in fresh toxin samples, although this distribution is variable from batch to batch.

Since low pH is known to dissociate the toxin into monomeric subunits (Gill, 1976), we tested whether starting toxin samples at different locations on our analytical IEF gels might modify our observed banding patterns. However, all our samples focused identically. One additional observation we made during analytical IEF was that cholera toxin precipitated in the gel when it focused at its *pI*. The same would be true of material purified either by chromatofocusing or by preparative IEF. Since we were never able to crystallize cholera toxin that had been obtained from preparative IEF, we presume that the toxin became partially denatured or otherwise altered by this effect.

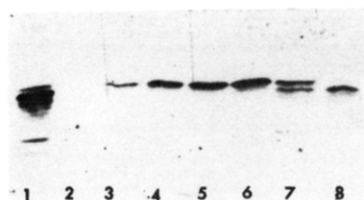
**Gel Filtration.** Complete toxin ran as a single peak on our Sephadex G-100 column. Analysis of fractions from this peak by SDS-PAGE showed that the protein in this peak contained a single A subunit and a pentamer of B subunits. No monomeric B subunits were seen in this analysis. Analytical IEF of this peak produced multiple bands, identical with those seen before column chromatography.

**Crystallization of Heterogeneous Material.** Freshly prepared cholera toxin, with analytical IEF banding patterns like those in Figure 2, lane 2, could be crystallized spontaneously within a few days. We grew both multiple and single crystals, up to 0.3 mm in length. When redissolved and subjected to analytical IEF, these crystals displayed the same multiple banding pattern seen in uncrystallized material. Diffraction patterns from these crystals all demonstrated a high degree of disorder within the crystalline lattice: mosaicity, twinning, and crystalline aggregation. None of these crystals could be used for quantitative diffraction data collection.

**Isolation of Isoelectrically Pure Protein.** Ion-exchange chromatography on the FPLC Mono Q resin, with careful analysis of each column fraction, was used to purify a single isoelectric variant of cholera toxin protein. Results from this procedure are illustrated in Figure 4, which shows a typical chromatogram, an IEF gel of the starting protein material, and analytical IEF of several sequential column fractions.



(a)



(b)

FIGURE 4: (a) Chromatogram of whole cholera toxin, eluted from Mono Q column. The column was eluted at a flow rate of 0.75 mL/min. Elution began with a 4-mL isocratic wash with buffer A, followed by a 10-mL linear gradient from 0% to 10% buffer B. This was followed by a wash with 10% buffer B for 1 mL, 30% buffer B for 1 mL, and then 0% buffer B for 12 mL. 0.5-mL fractions were collected. Buffer A consisted of 50 mM Tris-HCl, pH 7.70, and buffer B of 50 mM Tris-HCl/1 M NaCl, pH 7.70. Inset shows IEF of sample before chromatography. (b) Analytical IEF (pH 3–9) of cholera toxin samples. Lane 1, toxin before Mono Q column. Lanes 2–8, aliquots of column fractions 16–22. The primary peak was pooled from fractions 18 and 19. The separation of bands is clearly seen.

Each column fraction was checked by IEF as shown, before peak fractions were pooled. The chromatographic pattern corresponded to the analytical IEF pattern, and each chromatographic peak represents a separable isoelectric variant.

Analysis, by PAGE and SDS-PAGE, of the primary band purified by the above procedure demonstrated that this material was intact whole cholera toxin. Analytical IEF of the concentrated pooled primary chromatographic peak showed a single major band, although when greatly overloaded, the chromatogram also showed a small secondary band (Figure 5).

**Crystallization, Crystal Characterization, and Diffraction Data Collection.** Pooled material from the primary band crystallized spontaneously when we used the crystallization protocol described above. These crystals were identical in shape with those previously reported (Sigler et al., 1977). The crystals were of good quality and were neither twinned nor mosaic. They diffracted substantially beyond 2.8 Å, as shown in Figure 6. We characterized the crystal form with precession X-ray diffraction photographs. The three principal zones—(h0l), (hk0), and (0kl)—were recorded with 12° precession angles (minimum  $d$  spacing 3.7 Å). The crystal belongs to Laue group 2/ $m$ , and the (0k0) line manifests systematic absences of all  $k$ -odd Bragg diffraction peaks. Thus,

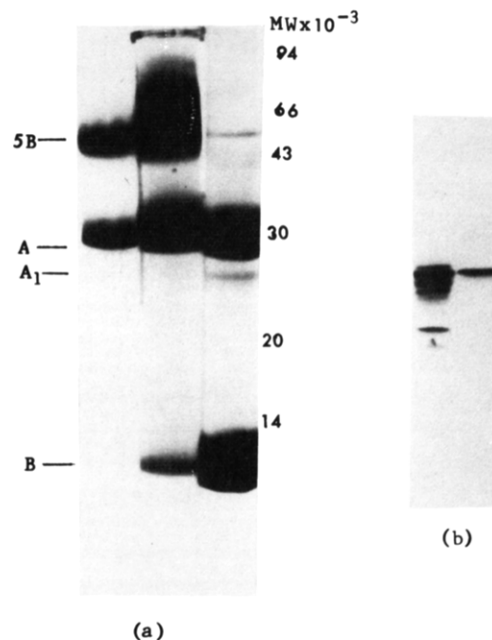


FIGURE 5: Electrophoretic analysis of concentrated, pooled primary peak fractions from Mono Q column. (a) SDS-PAGE. Lane 1, toxin sample prior to Mono Q column, not heated. Lane 2, overloaded gel of concentrated pool, not heated. Lane 3, same as lane 2, but heated in the presence of SDS. (b) Analytical IEF. Lane 1, toxin sample prior to Mono Q column. Lane 2, pooled sample of major peak from Mono Q.

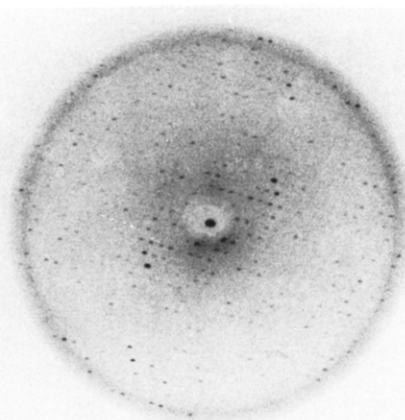


FIGURE 6: 12° precession photograph of the (h0l) zone (minimum  $d$  spacing 3.7 Å). Photograph was taken for 18 h with graphite monochromated CuK $\alpha$  X-rays from a rotating anode generator operating at 1.6 kW.

the crystal belongs to space group  $P2_1$ . The unit cell parameters are  $a = 73.0$  (3) Å,  $b = 92.2$  (3) Å,  $c = 60.6$  (3) Å,  $\beta = 106.4$  (5)°. These dimensions are essentially identical with those for the crystal form reported previously (Sigler et al., 1977). Although we have not, in this current study, measured the buoyant density of these crystals (Westbrook, 1985), that measurement was made 11 years ago by one of us (E.M.W.) and has been reported as 1.180 g/cm<sup>3</sup> in Ficoll/water (Sigler et al., 1977). This density value implies a protein volume fraction of 0.512, with one cholera toxin oligomer of  $M_r$  85 620 in the asymmetric unit. The Matthews number (Matthews, 1968) is  $V_M = 2.28$  Å<sup>3</sup>/Da.

Table I summarizes our data collection statistics.  $R$ -3, reported by the XENGEN data reduction program package, is equivalent to  $R$ -sym in intensity; our overall value for this parameter for these data was 10.1% and is quantum limited.

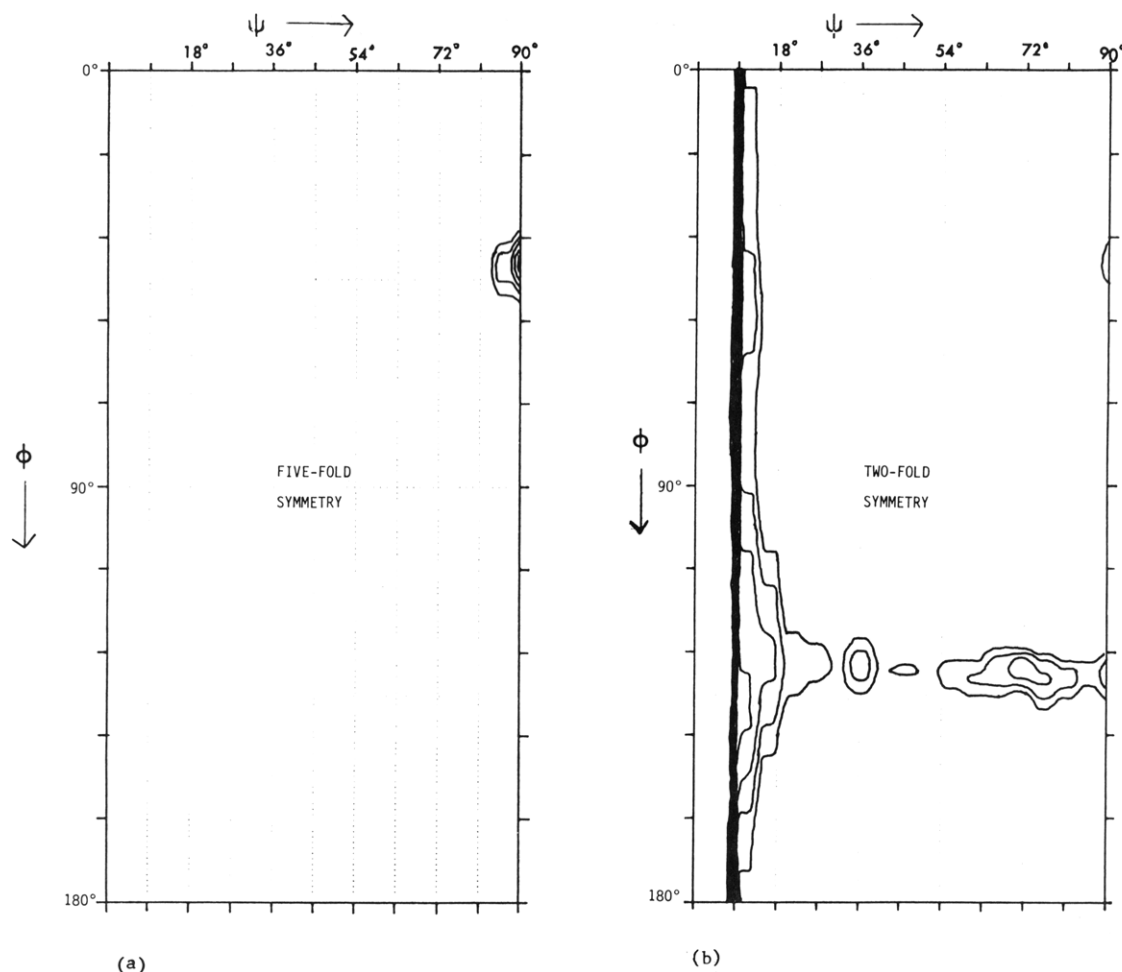


FIGURE 7: Cartesian projections of the self-rotation function of monoclinic cholera toxin crystal. All data between 35- and 4-Å resolution were used. Radius of integration 25 Å. Crystal a, axis is at  $\psi = 90^\circ$ ,  $\phi = -74^\circ$ ; crystal b, axis is at  $\psi = 0^\circ$ , independent of  $\phi$ ; crystal c, axis is at  $\psi = 90^\circ$ ,  $\phi = 180^\circ$ . (a) 5-fold symmetry ( $\kappa = 72^\circ$ ). (b) 2-fold symmetry ( $\kappa = 180^\circ$ ).

Table I: Summary of Native Diffraction Data<sup>a</sup>

parameters	min <i>d</i> spacing (Å) of shell				
	5.0	4.0	3.5	3.2	3.0
no. of unique reflections	3254	6361	9501	12421	15088
no. of obsd reflections	2890	5462	7704	9621	11248
% of unique reflections obsd	88	86	81	77	75
mean intensity (counts)	275	159	147	72	48
mean <i>I</i> / $\sigma$	17:1	12:1	8:1	6:1	3:1
<i>R</i> -sym ( <i>R</i> -3)	6.9	9.5	9.5	12.5	24.1

<sup>a</sup> Crystal to detector distance: 11.5 cm. Data were collected at the Midwest Area Detector Facility of Argonne National Laboratory with a Nicolet/Xenotronics multiwire detector.

That is, this value would be lower if we counted more X-rays in each Bragg reflection. Least-squares refined values of the unit cell parameters, calculated by XENGEN, corresponded within experimental error to those values obtained from film measurements.

**Rotation Function.** The rotation function, calculated with our current set of data, is similar to that reported by Zelano et al (1979), except that the previous result was reported with a left-handed coordinate system. Using data within resolution limits of 35-4 Å, and an integration sphere radius of 25 Å in the Patterson function, we clearly saw a single fivefold symmetry axis, normal to the unique b-axis of the crystal,  $140^\circ$  from the c-axis, and  $114^\circ$  from the a-axis. In addition, we find 10 noncrystallographic twofold dyads, spaced  $36^\circ$  apart, all perpendicular to the noncrystallographic fivefold, much as







wagon-wheel spokes are arranged about its axle. These results are illustrated in Figure 7.

## DISCUSSION

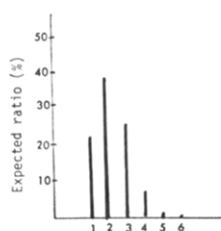
Despite apparent purity, a careful appraisal of protein homogeneity may be necessary to effect satisfactory crystallization and to obtain crystals suitable for X-ray diffraction. Although isoelectric separation methods have been used successfully in many instances to this end, they present some disadvantages, such as irreversible ampholyte binding (which can be observed spectroscopically) and denaturation due to precipitation at the isoelectric pH of the protein. For example, we observed this latter problem with cholera toxin and found it prevented successful crystallization. In addition, isoelectric focusing methods require low ionic strength (less than 100 mM), a condition that often is detrimental to protein stability. Finally, it is difficult to deal with proteins that have extreme isoelectric points. In the present study, we have applied ion-exchange chromatography, under carefully controlled conditions, to isolate isoelectric variants having very small differences in *pI*. This purification led to the successful crystallization of our protein, cholera toxin.

Cholera toxin has been previously reported to have only a single band that focuses at pH 6.6 (Finkelstein & LoSpalluto, 1970) or pH 6.65 (Mekalanos et al., 1978), in contrast to our observations. In all samples we examined by IEF, we found a characteristic pattern of multiple bands, spaced approximately 0.1 pH unit apart, primarily at pH 6.9–7.0. We cannot account for this discrepancy.

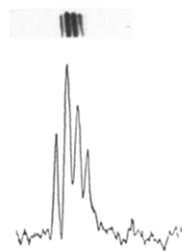


Expected Ratio	23.7%	39.6%	26.4%	8.8%	1.5%	0.1%
Aggregate Species	5x	4x1y	3x2y	2x3y	1x4y	5y
Model						
	1	2	3	4	5	6

(a)



(b)



(c)

FIGURE 8: (a) Binomial distribution model of combination of two species, taken in groups of five. (b) Expected ratios of combinations for two species,  $x = 75\%$ ,  $y = 25\%$ , arranged in groups of five. (c) Densitometry tracing of analytical IEF gel of a cholera toxin sample. The IEF is shown above the tracing. Although the precise fractional ratios of forms  $x$  and  $y$  are not easily measured in this way, the qualitative agreement between model and observation is clear.

We considered the possibility that the toxin might dissociate into A and B subunits during IEF. If so, the multiple bands that we see in whole toxin preparations might represent various oligomers of A and B, such as A<sub>1</sub>B, A<sub>2</sub>B, A<sub>3</sub>B, etc. However, our SDS-PAGE results do not support this conjecture, since we saw only one molecular weight species of the whole toxin.

We clearly observed two isoelectric forms of the monomeric B subunit, with ratios that vary among the samples we examined (Figure 3b). Intact cholera toxin is an oligomer containing five copies of subunit B. If we assume that each B subunit form is equally likely to be incorporated into whole toxin, then the probability of each of six possible combinations of B subunits follows a binomial distribution. Let  $x$  represent the fraction of all B subunits that are of one isoelectric variant, and let  $y$  represent the fraction of the other variant. Then

$$x + y = 1 \quad (1)$$

$$(x + y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5 = 1 \quad (2)$$

In this notation, each term in eq 2 represents the fraction of each of the isoelectric variants containing, respectively, five copies of  $x$ , four copies of  $x$ , and one copy of  $y$ , etc. The relative isoelectric point of each of these six combinations will differ by approximately a constant value, producing a "ladder" of IEF bands such as those seen in Figure 2 and Figure 3a.

In Figure 8 we compare a quantitative model of such a pattern, for which  $x = 75\%$ ,  $y = 25\%$ , with an observed pattern, analyzed and recorded with a microdensitometer, obtained from an actual IEF gel of a fresh cholera toxin sample. As seen in Figure 8, that observed IEF pattern fits well with our binomial distribution model. Thus, we suggest that the ladder of isoelectric points we observe in analytical IEF gels of cholera toxin can be explained by this model.

There are a number of possible reasons for the existence of two B-subunit forms in our samples. Toxin prepared commercially is derived from cultures of strain 569B, which carries a nontandem repeat of the cholera toxin gene (Mekalanos, 1983; Pearson & Mekalanos, 1982). We therefore considered the possibility that these two gene copies were different, producing nonidentical A or B subunits. Although the A<sub>1</sub> portion of each of these gene copies had been sequenced and

found to be identical (Mekalanos et al., 1983), the B-subunit portion of the second copy had not been sequenced completely (Mekalanos, personal communication).

Cholera toxin from E1 Tor strain 3083 contains a single copy of the cholera toxin gene. The sequence of this gene is identical with that published by Mekalanos for strain 2125 (T. J. Brickman, M. A. McIntosh, and R. A. Finkelstein, unpublished observations), so in this case, heterogeneity cannot be attributed to genetic differences between two copies of the gene. Figure 2, lane 5, shows that protein obtained from this strain has the same heterogeneous pattern as all other samples of toxin. The additional bands seen in that sample may have been generated by subunit dissociation and other processes that are accelerated by storage in solution at 4 °C (that sample had been stored 4 years). Therefore, we presume that subunit B undergoes a variable degree of some posttranscriptional modification or that our observation is some artifact of the isolation of our protein. We used protein from more than one source, isolated by two different procedures, and all our samples behaved identically. Thus, we do not believe that the formation of two separate forms of B subunits is an artifact of the isolation procedures.

Deamidation—the conversion of a glutamine residue to glutamate or an asparagine to aspartate—modifies the surface charge of a protein and reduces its isoelectric point (Kossiakoff, 1988). A change of 0.1 pH unit has been reported to result from a single deamidation (Hoffman, 1977; Williamson, 1973; Anderson et al., 1971). Heterogeneity due to deamidation could certainly affect the ability of a protein to form well-ordered crystals. Bott et al. (1982) provide examples of subtle structure variations that affect the crystallization of proteins, and they discuss mechanisms for these effects.

Deamidation may be the cause of the observed heterogeneity of subunit B in our samples. Deamidation is a posttranslational modification that can occur in the organism producing the protein, during isolation and purification of the protein, or during storage. The pH of the medium in which *V. cholerae* is grown must be carefully controlled, since cholera toxin produced in media with a pH above 8 becomes considerably deamidated, according to analysis on nondenaturing gels (L. Shoer, personal communication). Roda et al. (1977) reported that freshly isolated cholera toxin could be deamidated by incubating the protein in phosphate buffer, pH 8.2 at 37 °C, for 72 h. We treated several samples of toxin, which were less than 2 years old, in this way. Tris-PAGE, SDS-PAGE, and IEF patterns of these deamidated samples were practically identical with those of 15-year-old toxin that had not been so treated. We therefore conclude that deamidation of cholera toxin does occur, although slowly. Since deamidation should occur continuously, a toxin sample might reasonably be expected to become completely deamidated eventually, giving rise to a single homogeneous species, if no other processes are occurring. However, we do not see that, so we suspect that more than one chemical process occurs with aging, such as dissociation of B subunits or proteolysis.

Since we can now reproducibly grow single, nonmosaic crystals that diffract well, we are proceeding with the crystallographic structure determination of cholera toxin. Inability to grow acceptable crystals of this protein has frustrated the detailed study of its functional mechanisms. Low-resolution electron microscopic examinations of cholera toxin, in two-dimensional arrays, have been carried out (Ludwig et al., 1986; Ribí et al., 1988). These studies demonstrate the apparent 5-fold rotational symmetry of the B subunits, with the A subunit sitting on the 5-fold axis, displaced slightly above the

plane of the B subunits. Those results are entirely consistent with the much higher resolution results that we have obtained with the rotation function.

We are now carrying out the structure determination of cholera toxin by the classical multiple isomorphous replacement (MIR) method (Green et al., 1954; Blundell & Johnson, 1976). We intend to exploit the existence of 5-fold non-crystallographic symmetry in this crystal to improve the phase determination of its structure factors. Such averaging of electron density maps has been valuable in many crystal structure studies (Rossmann et al., 1985; Hogle et al., 1985; Wilson et al., 1981), although we must take care to perform our averaging only within the B subunits.

On the basis of our current knowledge, we cannot determine whether isoelectric heterogeneity has a bearing on the function of cholera toxin. We plan to investigate possible functional correlations with structural variation, in parallel with our crystallographic structure determination.

#### ACKNOWLEDGMENTS

We thank Mary Lyn Deacon for skilled technical assistance with data collection and analysis at the Midwest Area Detector facility. We acknowledge helpful discussions with David McKay, Paul Sigler, Richard Finkelstein, John Mekalanos, Linda Shoer, and Fred Stevens in the course of this work.

#### REFERENCES

- Anderson, B., Weigel, N., Kundig, W., & Roseman, S. (1971) *J. Biol. Chem.* **246**, 7023–7033.
- Andrews, A. T. (1986) in *Electrophoresis Theory, Techniques and Biochemical and Clinical Applications*, pp 29–30, Clarendon, Oxford, England.
- Blakesley, R. N., & Boeze, J. H. (1977) *Anal. Biochem.* **82**, 580–582.
- Blundell, T. L., & Johnson, L. N. (1976) *Protein Crystallography*, Academic, New York.
- Bott, R. R., Navia, M. A., & Smith, J. L. (1982) *J. Biol. Chem.* **257**, 9883–9886.
- Burns, D. L., Hausman, S. Z., Lindner, W., Robey, F. A., & Manclark, C. R. (1987) *J. Biol. Chem.* **262**, 17677–17682.
- Capiou, C., Petre, J., Van Damme, J., Puype, M., & Vandekerckhove, J. (1986) *FEBS Lett.* **204**, 336–340.
- Collier, R. J. (1982) in *ADP-ribosylation Reactions* (Hayashi, O., & Ueda, K., Eds.) pp 575–592, Academic, New York.
- Crowther, R. A. (1972) in *The Molecular Replacement Method* (Rossmann, M. G., Ed.) pp 172–178, Gordon and Breach, New York.
- Durbin, R. M., Burns, R., Moulai, J., Metcalf, P., Freymann, D., Blum, M., Anderson, J. E., Harrison, S. C., & Wiley, D. C. (1986) *Science (Washington, D.C.)* **232**, 1127–1132.
- Dwyer, J. D., & Bloomfield, V. A. (1982) *Biochemistry* **21**, 3227–3231.
- Finkelstein, R. A., & LoSpalluto, J. L. (1970) *J. Infect. Dis.* **121** (Suppl.) S63–S72.
- Geige, R., Dock, A. C., Kern, D., Lorber, B., Thierry, J. C., & Moras, D. (1986) *J. Crystal Growth* **76**, 554–561.
- Gill, D. M. (1976) *Biochemistry* **15**, 1242–1248.
- Gill, D. M. (1977) *Adv. Cyclic Nucleotide Res.* **8**, 85–118.
- Gill, D. M., & Meren, R. (1983) *J. Biol. Chem.* **258**, 11908–11914.
- Gill, D. M., & Coburn, J. (1987) *Biochemistry* **26**, 6364–6371.
- Gilman, A. G. (1987) *Annu. Rev. Biochem.* **56**, 615–649.
- Green, D. W., Ingram, V. M., & Perutz, M. F. (1954) *Proc. R. Soc. London, A* **225**, 287–307.
- Hampel, A., & Bock, R. M. (1970) *Biochemistry* **9**, 1873–1880.
- Hoffman, D. R. (1977) *Biological and Biomedical Applications of Isoelectric Focusing* (Catsimpoalas, N., & Drysdale, J., Eds.) p 132, Plenum, New York.
- Hogle, J., Chow, M., & Filman, D. J. (1985) *Science (Washington, D.C.)* **229**, 1358–1365.
- Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr.* **20**, 383–387.
- Kahn, R. A., & Gilman, A. G. (1986) *J. Biol. Chem.* **261**, 7906–7911.
- Kossiakoff, A. A. (1988) *Science (Washington, D.C.)* **240**, 191–194.
- Locht, C., & Keith, J. M. (1986) *Science (Washington D.C.)* **232**, 1258–1264.
- Lockman, H., & Kaper, J. B. (1983) *J. Biol. Chem.* **258**, 13722–13726.
- Ludwig, D. S., Ribí, H. O., Schoolnik, G. K., & Kornberg, R. D. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8585–8588.
- Matthews, B. W. (1968) *J. Mol. Biol.* **33**, 491–497.
- McPherson, A. (1989) *Useful Principles for the Crystallization of Proteins* (Michel, H., Ed.) CRC Press, Boca Raton, FL (in press).
- Mekalanos, J. J., Collier, R. J., & Romig, W. R. (1978) *Infect. Immun.* **20**, 552–558.
- Mekalanos, J. J., Collier, R. J., & Romig, W. R. (1979) *J. Biol. Chem.* **254**, 5849–5854.
- Mekalanos, J. J., Swartz, D. J., Pearson, G. D. N., Harford, N., Groyne, F., & de Wilde, M. (1983) *Nature (London)* **306**, 551–557.
- Middlebrook, J. L., & Dorland, R. B. (1984) *Microbiol. Rev.* **48**, 199–221.
- Pearson, G. D. N., & Mekalanos, J. J. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2976–2980.
- Ribi, H. O., Ludwig, D. S., Mercer, K. L., Schoolnik, G. K., & Kornberg, R. D. (1988) *Science (Washington D.C.)* **239**, 1272–1276.
- Roda, L. G., Tomasi, M., Battistini, A., Luzzi, I., Mastrantonio, P., Zampieri, A., & D'Agnolo, G. (1977) *Biochim. Biophys. Acta* **492**, 303–315.
- Rossmann, M. G., & Blow, D. M. (1962) *Acta Crystallogr.* **15**, 24–31.
- Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B., & Vriend, G. (1985) *Nature (London)* **317**, 145–157.
- Sigler, P. B., Druyan, M. E., Kiefer, H. C., & Finkelstein, R. A. (1977) *Science (Washington, D.C.)* **197**, 1277–1278.
- Tanaka, N. (1977) *Acta Crystallogr. Sect. A: Cryst. Phys., Diff., Theor. Gen. Crystallogr.* **A33**, 191–193.
- Thaller, C., Eichele, G., Weaver, L. H., Wilson, E., Karlsson, R., & Jansonius, J. N. (1985) *Methods Enzymol.* **114**, 132–135.
- Ueda, K., & Hayashi, O. (1985) *Annu. Rev. Biochem.* **54**, 73–100.
- van Heyningen, S. (1974) *Science (Washington, D.C.)* **183**, 656–657.
- van Heyningen, S. (1976) *J. Infect. Dis.* **133** (Suppl.), S5–S13.
- Weber, K., & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406–4412.
- Westbrook, E. M. (1985) *Methods Enzymol.* **114**, 187–196.
- Williamson, A. R. (1973) *Handbook of Experimental Immunology* (Weir, D. M., Ed.) pp 8.1–8.23, Blackwell, Oxford, England.

Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981) *Nature (London)* 289, 366-373.  
 Wisneiski, B. J., & Bramhall, J. S. (1981) *Nature (London)* 289, 319-321.

Zelano, J. A., Westbrook, E. M., Yonath, A., Druyan, M. E., & Sigler, P. B. (1979) *Molecular Mechanisms of Biological Recognition* (Balaban, M., Ed.) pp 157-163, Elsevier/North-Holland, Amsterdam.

## Stepwise Mechanism of HIV Reverse Transcriptase: Primer Function of Phosphorothioate Oligodeoxynucleotide<sup>†</sup>

Chirabrata Majumdar,<sup>‡§</sup> Cy A. Stein,<sup>||</sup> Jack S. Cohen,<sup>||</sup> Samuel Broder,<sup>||</sup> and Samuel H. Wilson<sup>\*†</sup>

Laboratory of Biochemistry and Clinical Oncology Program, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Received July 28, 1988; Revised Manuscript Received September 14, 1988

**ABSTRACT:** Primer recognition by purified HIV reverse transcriptase has been investigated. Earlier we found that the reaction pathway for DNA synthesis is ordered, with template-primer and free enzyme combining to form the first complex in the reaction sequence (Majumdar et al., 1988). We now find that d(C)<sub>28</sub> is a linear competitive inhibitor of DNA synthesis against poly[r(A)]-oligo[d(T)] as template-primer, indicating that d(C)<sub>28</sub> and the template-primer combine with the same form of the enzyme in the reaction scheme, i.e., the free enzyme. The phosphorothioate oligodeoxynucleotide Sd(C)<sub>28</sub> also is a linear competitive inhibitor against template-primer. However, the *K<sub>i</sub>* for inhibition (~2.8 nM) is ~200-fold lower than the *K<sub>i</sub>* for inhibition by d(C)<sub>28</sub>. Since the inhibition is linear competitive, the dissociation constant is equal to the *K<sub>i</sub>* for inhibition. Filter binding assays confirmed high-affinity binding between Sd(C)<sub>28</sub> and the enzyme and yielded a *K<sub>D</sub>* similar to the *K<sub>i</sub>* for inhibition. Substrate kinetic studies of DNA synthesis using Sd(C)<sub>28</sub> as primer, and poly[r(I)] as template, revealed that the *K<sub>m</sub>* for Sd(C)<sub>28</sub> is 24 nM. The *K<sub>m</sub>* for this primer is, therefore, 8-fold higher than the *K<sub>D</sub>* for enzyme-primer binding (2.8 nM). These results enable calculation of real time rate values for the enzyme-primer association (*k<sub>on</sub>* = 5.7 × 10<sup>8</sup> M<sup>-1</sup> s<sup>-1</sup>) and dissociation (*k<sub>off</sub>* = 1.6 s<sup>-1</sup>).

The HIV reverse transcriptase is responsible for replication of the HIV viral genome. Shortly after retrovirus infection of a cell, the enzyme catalyzes both RNA-directed DNA synthesis and DNA-directed DNA synthesis in the cytoplasm. In the process, the enzyme is capable of recognizing its natural primer, Lys-tRNA subspesies 3, and of the template switching steps necessary for replication of the complete viral genome (Gilboa et al., 1979). The enzymatic mechanism of these events and of the DNA synthesis reaction itself is the subject of the current investigation.

As a step toward understanding the mechanism of HIV reverse transcriptase, we recently applied steady-state kinetic and processivity analysis to obtain a kinetic scheme for the overall DNA synthetic reaction (Majumdar et al., 1988). Our data suggested that the free enzyme interacts with the primer substrate in the initial phase of the DNA synthesis reaction pathway. Here, we extended these studies by developing a primer analogue with exceptionally high affinity for the free enzyme. We then made use of this primer to determine both a thermodynamic dissociation constant for the enzyme-primer complex and a kinetic constant (*K<sub>m</sub>*) for DNA synthesis. A mathematical treatment of these constants then enabled calculation of on and off rates for the interaction between free enzyme and primer. The results are discussed in the context

of a kinetic model for reverse transcriptase activity and of the use of the high-affinity primer, a phosphorothioate oligodeoxynucleotide, as a potential antiviral agent (Matsukura et al., 1987).

### EXPERIMENTAL PROCEDURES

#### Materials

[<sup>3</sup>H]dTTP and [<sup>3</sup>H]dCTP were from New England Nuclear. [γ-<sup>32</sup>P]ATP was from Amersham. The chain length of each polynucleotide was calculated from the sedimentation coefficient or from gel electrophoretic analysis. Normal and sulfur-substituted oligodeoxynucleotides were synthesized chemically and were characterized by NMR spectroscopy, as reported earlier (Stec et al., 1984; Stein et al., 1988). Synthesis was conducted with a mixture of phosphorothioate diastereoisomers. Hence, each phosphorothioate oligomer is a mixture of stereoisomers. r(A)<sub>810</sub>,<sup>1</sup> where 810 is the average chain length, and r(I)<sub>127</sub>, where 127 is the average chain length, were

<sup>1</sup> Abbreviations: TCA, trichloroacetic acid; Tris, tris(hydroxymethyl)aminomethane; Pol I, *Escherichia coli* DNA polymerase I; poly[r(A)], polyriboadenylate; poly[r(C)], polyribocytidylate; poly[r(I)], polyriboinosinate; oligo[d(T)], oligomer of deoxythymidylate; nd(C)<sub>28</sub> or d(C)<sub>28</sub>, 28-residue-long normal oligodeoxycytidylate; Sd(C)<sub>28</sub>, 28-residue oligodeoxycytidylate with sulfur substituted at a nonbridge oxygen of each phosphate atom; 3'-S<sub>9</sub>-O<sub>18</sub>-5'-d(C)<sub>28</sub>, 3'-O<sub>9</sub>-S<sub>9</sub>-O<sub>9</sub>-5'-d(C)<sub>28</sub>, or 3'-O<sub>18</sub>-S<sub>9</sub>-5'-d(C)<sub>28</sub>, partially sulfur substituted oligodeoxynucleotides where sulfur substitutions are tandemly connected (blocks) and the subscript represents the number of substitutions relative to the 3' or 5' ends. Polynucleotide subscripts indicate precise chain length unless otherwise noted. The kinetic nomenclature and constants are according to Cleland (1963a,b) and Fersht (1985).

<sup>†</sup> This work was supported in part by U.S. Army Medical Research and Development Command Grant DAMD17-PP-7801.

<sup>‡</sup> Laboratory of Biochemistry.

<sup>§</sup> Present address: Laboratory of Comparative Carcinogenesis, Frederick Cancer Research Facility, Frederick, MD 21701.

<sup>||</sup> Clinical Oncology Program.