Molecular Cloning of Mouse Immunoglobulin Heavy Chain Messenger Ribonucleic Acids Coding for μ, α, $\gamma 1$, $\gamma 2a$, and $\gamma 3$ Chains†

Jerry M. Adams,* Nicholas M. Gough, Elizabeth A. Webb, Brett M. Tyler, Jillian Jackson, and Suzanne Cory

ABSTRACT: To obtain immunoglobulin heavy chain nucleotide sequences in a pure form, we have constructed plasmid clones bearing complementary DNA (cDNA) copies of mouse plasmacytoma mRNAs coding for a μ chain, two α chains, two $\gamma 2a$ chains, a $\gamma 3$ chain, and a $\gamma 1$ chain. As substantially pure mRNAs were not available for identifying certain clones, we performed colony hybridization with cDNA made on mRNA from both homologous and heterologous plasmacytomas to eliminate clones bearing sequences common to plasmacytomas, a scheme having wider applicability. The results on the clones selected indicate that there is no detectable homology between μ, α, and γ nucleotide sequences. The identity of the sequences in representative clones was confirmed by showing that each could specifically arrest translation of the appropriate heavy chain mRNA in a cell-free system. Detailed restriction maps of cloned cDNAs were derived and correlated with known amino acid sequences to define the portion of the mRNA sequence represented. The largest HPC 76 μ cloned sequence represents nearly a complete copy of the mRNA, as does an S107 α and an HOPC 1 $\gamma 2a$ clone; these clones bear V_H as well as C_H sequences. The Y5606 $\gamma 3$ and MOPC 21 $\gamma 1$ cloned sequences are smaller and represent only C region sequences, while the largest McPC 603 α and MOPC 173 $\gamma 2a$ clones bear part of a V_H as well as the C_α sequence. The restriction maps indicate that, in addition to the five distinct C_H regions, three V_H regions and four joining (J_H) regions are represented. Restriction analysis of $\gamma 2a$ clones unexpectedly revealed a sequence difference in the C regions of HOPC 1 and MOPC 173 mRNAs, possibly reflecting heterogeneity of the C$\gamma 2a$ gene. A novel cloning aberration generating an inverted repeat sequence of several hundred base pairs was found in three clones.

An immunoglobulin contains two identical heavy and two identical light chains, each of which has an N-terminal variable (V)[1] and a C-terminal constant (C) region. Antigen binding specificity is determined by the V regions of both chains, while biological effector functions are determined by the C region of the different types of heavy chains, which in the mouse include μ, $\gamma 1$, $\gamma 2a$, $\gamma 2b$, $\gamma 3$, α, ϵ, and δ (for reviews, see Gally, 1973; Spiegelberg, 1974). The heavy chain gene system poses major biological questions regarding the origin of V_H diversity and the mechanisms by which separate genes encode a single polypeptide chain. In the germ line, V_H genes appear to be clustered at a locus distant from that bearing the C_H genes (Kemp et al., 1979; see Williamson (1976) for a review). During lymphocyte ontogeny there appears to be a DNA rearrangement to bring V_H and C_H sequences closer together (Early et al., 1979; Gough et al., 1980a; S. Cory, umpublished experiments), as is now well established for the κ and λ light chain genes (see, e.g., Brack et al., 1978; Seidman & Leder, 1978). An intriguing feature of the heavy chain system is that a single V_H region apparently can be expressed in association with more than one C_H region within the same cell or cell lineage (Gally & Edelman, 1972).

It seems clear that any detailed study of the extremely complex Ig heavy chain system will require a number of pure heavy chain nucleotide sequences, obtainable only by molecular cloning. Construction of well-characterized chimeric plasmids bearing Ig mRNA sequences allows those sequences to be analyzed in detail and provides pure probes for the corresponding Ig genes. In an accompanying paper (Gough et al., 1980b), we have described the successful cloning of seven κ mRNA sequences by insertion of duplex complementary DNA (cDNA) into the *Pst*I site of the plasmid vector pBR322 by dG·dC tailing, a procedure which permits subsequent excision of the cDNA insert (Bolivar et al., 1977). We have applied the same methodology here with a set of Ig heavy chain mRNAs isolated from mouse plasmacytomas (S. Cory, unpublished experiments). A problem faced in identifying certain clones was that the corresponding mRNA preparations were not pure enough to provide specific probes for the relevant clones. To overcome this problem, we performed colony hybridization with homologous and heterologous probes to eliminate all clones bearing sequences shared by the two probes. Reports have appeared during the latter stages of this work describing clones bearing a complete α chain sequence (Early et al., 1979), a $\gamma 2b$ C region (Schibler et al., 1978), and a $\gamma 1$ C region (Sakano et al., 1979; Rogers et al., 1979).

We have identified clones bearing sequences derived from heavy chain mRNAs encoding a μ, two α, a $\gamma 1$, a $\gamma 3$, and two $\gamma 2a$ chains. Restriction endonuclease mapping has enabled us to define the regions of the mRNAs present in the clones. Three V_H and five different C_H sequences are represented. The clones thus provide a set of well-characterized nucleotide sequences for molecular analysis of the heavy chain system.

Materials and Methods

Cloning of cDNA. Polyadenylated mRNA was isolated from plasmacytoma microsomes essentially as described previously (Cory et al., 1976). The tumors included Y5606 (Barstad et al., 1974), obtained from the Salk Institute collection, and HOPC 1 (Potter, 1972), in addition to those cited in the accompanying paper (Gough et al., 1980b). The 17–19S RNA fraction from each tumor was demonstrated to be

† From the Molecular Biology Laboratory, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria 3050, Australia. *Received September 24, 1979.* This work was supported by the U.S. National Cancer Institute (RO1 CA 12421), The American Heart Association, and the National Health and Medical Research Council (Canberra).

[1] Abbreviations used: Ig, immunoglobulin; C, constant region; V, variable region; cDNA, complementary DNA; poly(A), poly(adenylic acid); EDTA, ethylenediaminetetraacetic acid; EGTA, ethylene glycol bis(β-aminoethyl ether)-*N*,*N*′-tetraacetic acid; bp, base pair(s); kb, kilobase pair(s); Pipes, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); NaDodSO$_4$, sodium dodecyl sulfate.

0006-2960/80/0419-2711$01.00/0

substantially enriched for heavy chain mRNA by gel electrophoresis and characterization of in vitro translation products (S. Cory, unpublished experiments). The accompanying paper (Gough et al., 1980b) described the methods used for synthesis of duplex cDNA on Ig mRNA templates, enrichment of the largest cDNA molecules by gel filtration, addition of oligo(dC) tracts to the cDNA with terminal deoxynucleotidyltransferase, annealing to dG-tailed pBR322 DNA, transformation of χ1776, and colony hybridization. It also described screening of "mini-lysates" to identify the clones bearing the largest excisable inserts, large-scale preparation of plasmid DNA, and purification of the cDNA insert.

Hybrid Arrest of Translation (Paterson et al., 1977). A mixture of two mRNAs (0.2–1.0 μg each) was precipitated with ethanol together with DNA of a chimeric plasmid (or the cDNA insert from it) in twofold molar excess over the relevant mRNA. The nucleic acids were redissolved in 20 μL of freshly deionized formamide containing 20 mM sodium Pipes, pH 7, and 1 mM EDTA and incubated for 3 min at 65 °C to denature the DNA. Then 5 μL of 2 M NaCl, 0.1 M sodium Pipes, pH 7, and 5 mM EDTA was added, and hybridization was carried out at 42 °C for 1 to 3 h (1 h proved sufficient). The sample was then chilled and, after addition of 100 μL of cold water, divided into two portions, each of which was precipitated twice with ethanol and rinsed once with 70% ethanol. The two precipitates were resuspended in 10 μL of cold water, one sample was put into boiling water for 60 s to melt the hybrid, and both were stored frozen.

Translation System. A reticulocyte lysate, prepared as described by Hunt & Jackson (1974), was stored in 0.5-mL aliquots under liquid N_2. While a 0.5-mL aliquot was slowly thawed, various components were added to give the following concentrations: 1.5 mM GTP + 1.5 mM $MgCl_2$, 15 μM hemin, 0.4 mM dithiothreitol, 40 μg/mL creatine kinase, 10 mM creatine phosphate, 220 μM spermidine, 115 mM potassium acetate, 0.6 mM magnesium acetate, and 25–150 μM amino acids other than methionine. The supplemented lysate (660 μL) was made 1 mM in $CaCl_2$ and 7.5 μg/mL in micrococcal nuclease and incubated for 15 min at 20 °C; it was then chilled and made 2 mM in EGTA to inhibit the nuclease (Pelham & Jackson, 1976). Ascites tRNA was then sometimes added to 60 μg/mL, but it stimulated incorporation only slightly. A typical 50-μL reaction mixture containing 40 μL of treated lysate, 10 μL of mRNA or hybrid in water, and 50 μCi of [^{35}S]methionine (Amersham, 600 Ci/mmol) was incubated for 90 min at 30 °C.

Restriction Endonuclease Mapping by Partial Digestion of End-Labeled Fragments. To produce DNA fragments labeled at only one end, DNA was labeled at 3′ termini using DNA polymerase I, either at restriction sites within the cDNA sequence or at the oligo(dG·dC) tails, and recut with an appropriate enzyme, and the resulting fragments were separated on polyacrylamide gels. Partial digestion was performed at 30 °C in a 25-μL reaction containing 1 μg of pBR322 DNA and 0.2 unit of the appropriate restriction endonuclease (from New England Biolabs, using their suggested buffers). Samples (2.5 μL), removed at 5-min intervals, were pooled in 25 μL of 0.1 M EDTA and 20 μg/mL tRNA. The DNA was ethanol precipitated to remove excess salt, which causes smearing of bands, and redissolved in 7 μL of water. Then 2 μL of 50% glycerol, 20 mM EDTA, 0.15% bromophenol blue, and 0.15% xylene cyanol FF was added, and the samples were loaded onto a 0.35 mm thick 5% polyacrylamide gel (acrylamide:bis = 29:1) in 50 mM Tris-borate and 1 mM EDTA (Sanger & Coulson, 1978). Electrophoresis was at 2–5 mA, until the bromophenol blue dye marker had migrated about 28 cm (overnight). Nearly complete digests, containing 1 μg of pBR322 DNA and 0.5 unit of the appropriate enzymes, incubated for 60 min, were electrophoresed on neighboring tracks.

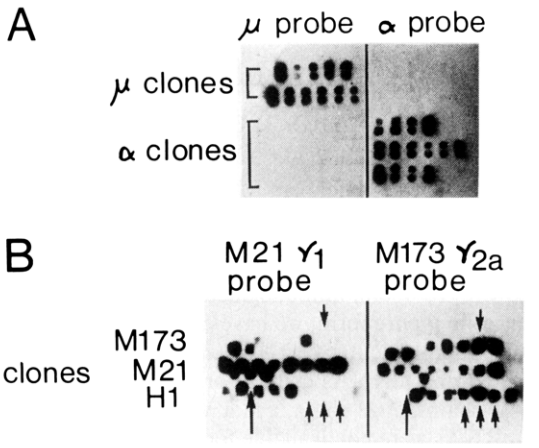

FIGURE 1: Cross-hybridization as a method for clone identification. (A) Arrays of previously selected μ and α candidate clones subjected to colony hybridization with impure μ and α ^{32}P-labeled cDNA probes, using modifications of the procedure of Grunstein & Hogness (1975), as described in the accompanying paper (Gough et al., 1980b). An autoradiogram of the filter is shown. Each clone was spotted twice (vertically) to assess the reproducibility of the signal. On the scale used in Tables I–III, the hybridization signals from the first three μ clones in the top row would be assigned values of +++, +, and ++, respectively. (B) An example of an actual screen in which replicas of clones generated from the indicated mRNAs were scored with cDNA made from impure MOPC 21 γ1 mRNA and MOPC 173 γ2a mRNA. The long arrow indicates a clone specifically revealed with the γ1 probe and the short arrows indicate clones specifically detected with the γ2a probe.

Results

A Comparative Hybridization Procedure for Clone Identification. Duplex cDNA was synthesized on Ig heavy chain mRNAs and cloned in the plasmid vector pBR322 by the procedures described in the accompanying paper (Gough et al., 1980b). To screen for Ig bearing clones, we first scored all transformants by colony hybridization (Grunstein & Hogness, 1975) using as probes [^{32}P]cDNA made on homologous mRNA preparations. This screen obviously might leave false positives because the probes were not pure. We reasoned, however, that abundant non-Ig mRNA species should be shared by different plasmacytomas, given the genetic and physiological near identity of Balb/c plasmacytomas (Potter, 1972). Hence, non-Ig clones should hybridize with heterologous probes, whereas Ig clones should not, if one makes the reasonable assumption that Ig sequences as distant as μ, α, and γ will not cross-hybridize, barring coincidental V region homology. Consequently, hybridizing impure probes corresponding to different types of Ig mRNAs to replicas of candidate clones should allow one to identify the clones bearing Ig sequences.

The validity of the cross-hybridization approach is documented in Figure 1A, which shows the extent of hybridization of μ and α cDNA probes to a set of previously selected μ and α candidate clones. For each colony the relevant probe gave clear hybridization, whereas the irrelevant probe gave none. Moreover, when 5% α probe was added to the μ probe, the mixed probe clearly hybridized to the α as well as μ clones (not shown). Hence, we conclude that irrelevant clones which bear sequences represented even at a level of a few percent in the probe can be recognized. Figure 1B shows an actual

Table I: HPC 76 μ Candidate Clones

clone	insert[a] size (kb)	insert[c] excisable	colony hybridization[g] H76μ[e]	3′ probe[f]	M173γ[e]
μ17	1.8[b]	+	++	+	–
.22	1.1[b]	p[d]	+	++	
.16	~1.6	+	++	++	–
.19	~1.5	+	++	+	
.7	0.9[b]	+	++	+	–
.20	~0.9	+	+	±	
.1	~0.8	–	++		–

[a] Measured by comparing the mobility on a 1% agarose gel of plasmid linearized with *Eco*RI to the mobility of linear plasmids with inserts of known size (Gough et al., 1980b). [b] More accurate determination from the size of *Pst*I fragments on a 3% agarose gel using a *Hae*III digest of ϕX174 DNA as markers. [c] Assessed by whether or not *Pst*I treatment reduced the chimeric plasmid to the size of pBR322 (Gough et al., 1980b). [d] *Pst*I digestion reduced the size of the chimeric plasmid but not to that of pBR322, showing that only one external *Pst*I site was cleavable. [e] Data representative of two independent determinations. [f] A cDNA synthesized with oligo(dT) priming on HPC 76 mRNA in the presence of dideoxy TTP as well as dTTP. Size determined by gel electrophoresis to be approximately 200 nucleotides. [g] The hybridization signal with cDNA probes was assigned values on the scale indicated in the legend to Figure 1.

screen in which impure γ1 and γ2a probes were used to score replicas of clones generated from MOPC 173 γ2a, MOPC 21 γ1, and HOPC 1 γ2a mRNAs. Clearly, most clones hybridized with both probes and were therefore rejected, but the large arrows indicate one clone which hybridized with the γ1 but not the γ2a probe, and the short arrows indicate four clones labeled only by the γ2a probe.

Identifying μ clones proved to be straightforward because the HPC 76 μ mRNA was our purest heavy chain mRNA preparation. Candidate μ clones, described in Table I, included plasmids with inserts ranging from 800 to 1800 bp long. All probably contain μ sequences, because colony hybridization indicated that none contained a sequence present at a detectable level in an S107 α probe (Figure 1) nor a very impure MPC 11 γ probe. Nor did any hybridize to a probe made from "peak A RNA", a common contaminant of heavy chain mRNAs (see Discussion), nor any of those tested with a MOPC 173 γ probe (Table I). Since the HPC 76 tumor also synthesizes κ chains, we verified that none hybridized to a κ probe. At least five clones appear to include 3′-terminal mRNA sequences, because they hybridized with an HPC 76 μ cDNA probe less than 200 nucleotides long (3′ probe). The inserts in five clones were fully excisable with *Pst*I, while DNA from one was merely linearized, and DNA from another (μ22) yielded only part of the insert, because an internal *Pst*I site was cleavable but not one external site.

Excellent candidates for α clones were obtained by cloning mRNA sequences from plasmacytomas S107 and McPC 603 (Table II). The clones were detected with a probe made from S107 α mRNA, the purer mRNA preparation. The inserts were from 800 to 1580 bp long and 8 out of 10 were fully excisable. At least six clones appear to include 3′-terminal mRNA sequences, because they hybridized with a short S107 α cDNA probe specific for that region (see below). All the clones probably bear α sequences, since none gave colony hybridization with probes made from HPC 76 μ mRNA (Figure 1), from cruder preparations of MOPC 173 γ and MPC 11 γ mRNA (Table II), from peak A RNA, or from MOPC 41A κ mRNA (not shown). In support of this conclusion, clones derived from S107 and McPC 603 mRNA appear to contain related sequences, since all those examined with *Pst*I, *Eco*RI and *Pvu*II endonucleases contained internal sites (Table II).

Comparative hybridization was particularly valuable in identifying γ clones, because our γ mRNA preparations were less pure. Since the mRNAs contained other major components, colony hybridization with γ probes yielded numerous false positives, examples of which are given with italicized clone numbers in Table III. Some non-Ig clones (e.g., pY5606.3) hybridized as strongly as the γ clones. The most common non-Ig clones were those giving a strong signal with a peak A probe (see Discussion). Most clones which hybridized with any one irrelevant probe also hybridized with all others tested (e.g., pM173.11 and pM21.12). Rarer clones (e.g., pH1.3 and pH1.5) hybridized well only with probes from certain plasmacytomas, presumably because the corresponding RNA sequences were more abundant in those lines.

The γ clones that passed the extensive cross-hybridization screen (Table III) included three strong γ2a candidates derived from MOPC 173 mRNA and four from HOPC 1 mRNA, all bearing excisable inserts from 800 to 1900 bp long, as well as two γ3 candidates from Y5606 mRNA, one with a 1050-bp excisable insert. Despite a low number of transformants obtained with MOPC 21 cDNA, we also isolated one strong γ1

Table II: S107 α and McPC 603 α Candidate Clones

clone	insert[a] size (kb)	*Pst*I[b] fragments	insert[c] excisable	restriction sites[e] RI	*Pst*I	*Pvu*II	colony hybridization with cDNA probes S107α	3′ probe[g]	M173γ	M11γ
pS107α.4	*1.58*	0.76, 0.82	+	1	1	2	++	*h*	–	–
.15	*1.35*	0.80, 0.55	+	1	1	2	++	++	–	–
.16	*1.21*	0.72, 0.49	+	1	1	1	+	±	–	–
.2	*1.11*	0.64, 0.47	+	1	1		++	±	–	–
.3	~0.8		+	1	1	1	+	+	–	–
.1	~0.8		–	1		1	+	±	–	–
pM603α.8	*1.32*	0.81, 0.51	+	1	1	2	++	+	–	–
.11	~1.4	0.50	p[d]		1		++	++	±	±
.1	~1.0	0.78	+	1	1	2	++	++	–	–
.4	~0.9	0.84	+	1	*f*		++	–	–	–

[a] Measured by comparing the mobility on a 1% agarose gel of plasmid linearized with *Bam*HI to the mobility of linear plasmids with inserts of known size, except for the more accurate numbers shown in italics, which are the sum of the sizes of the *Pst*I fragments. [b] Size (kilobase) determined on a 3% agarose gel with a *Hae*III digest of ϕX174 DNA as marker. [c] Assessed by whether or not *Pst*I treatment reduced the chimeric plasmid to pBR322 size. [d] *Pst*I digestion reduced the size of the chimeric plasmid but not to that of pBR322, showing that only one *Pst*I site was cleavable. [e] Restriction endonuclease sites examined within insert. A blank indicates that digestion was not tested. [f] Difficult to assess the presence of a *Pst*I site as one resulting fragment would have been too small to be detectable. [g] Probe synthesized using oligo(dT) priming on S107 α mRNA in the presence of dideoxy TTP as well as dTTP. Size determined by gel electrophoresis to be approximately 200 nucleotides. [h] Presence of 3′ sequence established in an orientation experiment (see text).

Table III: Candidate γ2a, γ3, and γ1 Clones

clone	insert[a] size (kb)	insert[c] excisable	colony hybridization with cDNA probes[d] M173γ2a	Y5606γ3	M21γ1	S107α	M603α	H1γ	peak A
pM173γ2a.15	1.90[b]	+	++	−	±	−	−	±	−
.20	~1.23[b]	+	++	−	±	−	−		−
.18	~0.8	+	++	−	±	−	−		−
.11	~1.5	+	+	++		+	+	++	++
pH1γ2a.8	1.72[b]	+	++	−	−	−	−	±	−
.17	~1.36[b]	+	+	−	−		−	±	−
.6	~1.4	+	++	−	−	−	−	±	−
.20	~1.2	+	+	−	±	−	−	±	−
.3	~1.7	+	+	++		±	+	±	++
pY5606γ3.15	1.05[b]	+	−	++	−	−	−	−	−
.18	~1.0	+	−	+	−		−	−	−
.3		+	++	++		+	++	±	++
pM21γ1.1	0.74[b]	−	−	−	++		−		−
.12	~0.2	+	+	++	++		++		++

[a] Measured by comparing the mobility on a 1% agarose gel of plasmid linearized with *Bam*HI and *Eco*RI to the mobility of linear plasmids with inserts of known size (kilobase). [b] Size determined by restriction mapping. [c] Assessed by whether or not *Pst*I treatment reduced the chimeric plasmid to the size of pBR322. [d] Blanks indicate that a probe was not tested.

candidate bearing a 750-bp insert, which was not, however, excisable with *Pst*I.

Hybrid Arrest of Translation. To establish the identity of representative clones more definitively, we tested whether the plasmid DNA could arrest translation of the appropriate mRNA in the hybrid-arrest assay of Paterson et al. (1977). In our hands, this procedure, which includes hybridization in formamide, often markedly reduced messenger activity, perhaps due to contaminants in the formamide. Nevertheless, we could guard against any effect due to nonspecific inactivation by including an irrelevant mRNA (usually κ mRNA) in each hybridization mixture and by verifying that melting the hybrid freed the heavy chain mRNA activity (Paterson et al., 1977). The polypeptides made in the messenger-dependent lysate system (Pelham & Jackson, 1976) were electrophoresed on $NaDodSO_4$–polyacrylamide gels.

Figure 2 shows hybrid-arrest experiments carried out with pS107α.4 DNA, with pH76μ.7 DNA, and with DNA from a clone bearing a non-Ig sequence ("control"). That the pS107α.4 DNA abolished α chain but not κ chain synthesis can be seen by comparing the products programmed by the unhybridized mixture of crude α and κ mRNAs (track 6), the hybrid (track 5), and the melted hybrid (track 4). Similarly, tracks 7–9 indicate that pH76μ.7 DNA arrested synthesis of the μ chain. The remote possibility that heavy chain synthesis had been selectively inhibited in some fashion other than by hybridization was excluded by showing that α chain synthesis was not arrested by the non-Ig clone (track 2), nor μ chain synthesis by a κ clone examined in the same experiment (not shown).

We have also demonstrated hybrid arrest with another μ clone bearing a larger insert (pH76μ.17), two γ2a clones derived from different mRNAs (pM173γ2a.15 and pH1γ2a.8), a γ3 clone (pY5606γ3.15), and the γ1 clone (pM21γ1.1), with results very similar to those in Figure 2. For each type of mRNA preparation, we had previously verified that the major translation product is an Ig heavy chain by immunoprecipitation with appropriate antisera (S. Cory, unpublished experiments). We also examined immunoprecipitates of the products from the hybrids and melted hybrids; the results confirmed those described above, although the melted hybrids yielded smaller polypeptides in addition to the heavy chain, presumably due to some mRNA degradation during the hybridization and/or melting steps.

Lengths of the Oligo(dG·dC) Tails. The lengths of the oligo(dG·dC) tails in several of the cloned cDNA sequences

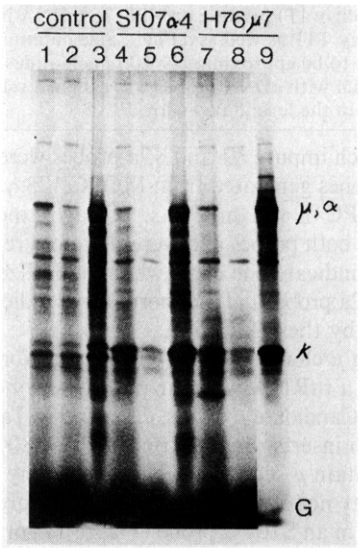

FIGURE 2: Identification of an α and a μ clone by hybird arrest of translation. The clones tested were plasmid DNA bearing a non-Ig sequence (tracks 1–3), pS107α.4 DNA (tracks 4–6), and pH76μ.7 DNA (tracks 7–9). Products from melted hybrids are shown in tracks 1, 4, and 7 and those from hybrids in tracks 2, 5, and 8. The products from the corresponding unhybridized mixtures of crude mRNAs are shown in tracks 3, 6, and 9: track 3, crude α and λ mRNA; track 6, crude α and κ mRNA; and track 9, crude μ and κ mRNA. The position of μ and α chains made in vitro, which have the same apparent molecular weight, is indicated, as is that of κ chains. Other than globin (G), no detectable polypeptides were produced by residual endogenous mRNA activity.

were determined by sizing the depurination products of insert DNA, labeled at 3′ termini with [α-^{32}P]dCTP, as described previously (Gough et al., 1980b). The tails measured were between 7 and 19 nucleotides long (Table IV), confirming that short dG·dC sequences are sufficient for an insert to anneal stably to the vector.

Orientation of the Inserts with the mRNA. We used different appraoches to establish which ends of cDNA inserts correspond to 5′ (V) and 3′ (C) portions of the mRNA. The first, most successful with the α clones, is diagrammed in Figure 3A. Since an *Eco*RI site occurs within the α inserts as well as the vector (Table II), an *Eco*RI digest of an α plasmid yields a large (L) and a small (S) fragment, each containing one end of the insert. A short cDNA probe should hybridize to only one of the two fragments, thereby establishing the orientation. Figure 3B shows that short α cDNA hy-

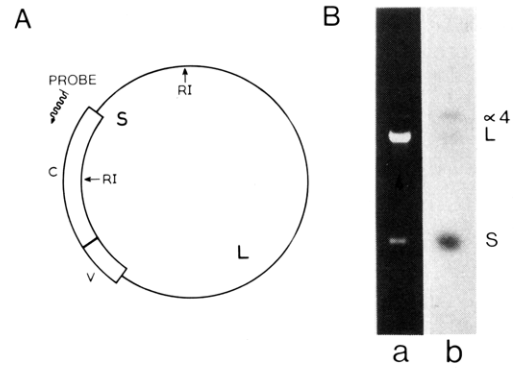

FIGURE 3: Orienting an S107 α insert within pBR322 DNA. (A) Diagram showing the fragments expected in an *Eco*RI digest of a recombinant plasmid having an *Eco*RI site within the insert as well as the vector. Short cDNA hybridizes to the plasmid DNA fragment bearing the 3′-terminal mRNA sequence. (B) *Eco*RI fragments obtained from pS107α.4 DNA. Fragments stained with ethidium bromide are shown on the left and those revealed by autoradiography after hybridization with short cDNA on the right. The probe was 50–200 nucleotides long, as judged by gel electrophoresis.

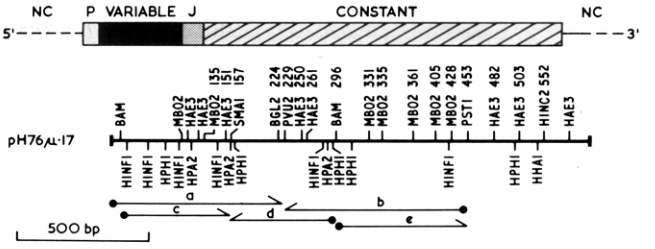

FIGURE 4: Restriction endonuclease cleavage map of a μ cDNA clone. The μ mRNA (top) consists of a constant region (hatched), a variable region (filled), approximately 60 nucleotides encoding a presumptive precursor peptide (P, open), a joining region (J, stippled) (Gough et al., 1980a), and 3′ and 5′ noncoding regions (NC). The restriction map of the cDNA insert in pH76μ.17 is shown below the mRNA. All of the restriction sites within the constant region can be related to the amino acid sequence of the MOPC 104E μ chain (Kehry et al., 1979), and corresponding amino acid residues are nominated above a number of the sites. Fragments a–e, labeled at the position of the filled circle, were used to derive the restriction map of the cDNA, by partial digestion.

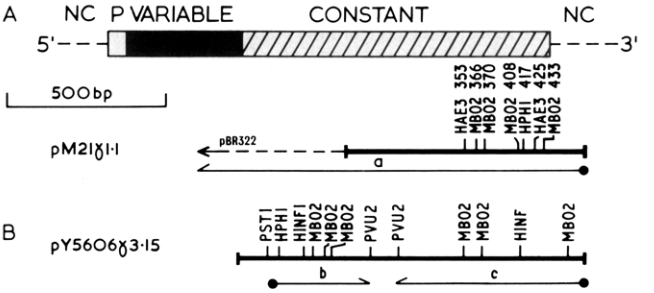

FIGURE 5: Restriction endonuclease cleavage maps of a γ1 and a γ3 cDNA clone. (A) The γ1 mRNA (top) is shown with the same convention used for Figure 4. In clone pM21γ1.1, only the *Pst*I site distal to the plasmid *Eco*RI site can be cleaved, but the cDNA sequence can be excised from the plasmid as a 1490 bp long *Eco*RI/*Pst*I fragment. The cDNA part of the fragment is shown as a thick line and the plasmid dashed. The restriction map was generated by partial digestion of fragment a. The restriction sites can be related to the MOPC 21 γ1 amino acid sequence (Adetugbo, 1978), and corresponding amino acid residues are given above the restriction sites. (B) The map of the γ3 cDNA clone was generated by partial digestion of fragments b and c and its orientation with respect to the γ3 mRNA determined as described under Results. As no amino acid sequence data are available for mouse γ3 chains, the map cannot be aligned with the mRNA, but it almost certainly corresponds to a portion of the constant region (see Results).

Following a method suggested by Dr. D. Kemp, we established the orientation of the pY5606γ3.15 sequence by determining which strand of the insert was complementary to the mRNA. To generate fragments with label in different strands, the pY5606γ3.15 insert was cleaved with *Pst*I (see Figure 5B) and 3′ terminally labeled with [^{32}P]dCTP (Gough et al., 1980b), and the two labeled DNA fragments were resolved by gel electrophoresis. Each was then hybridized with γ3 mRNA, and the hybrids were treated with the single-strand nuclease S1 (Berk & Sharp, 1978). Each of the two fragments is effectively labeled in only a single strand, at the "internal" *Pst*I site, because the labeled tails are not complementary to the mRNA and hence can be disregarded. Consequently, the mRNA will hybridize to and protect the label in one fragment from S1 digestion, but not the label in the other fragment, which is present in the opposite strand. The γ3 mRNA specifically protected the label in the large *Pst*I fragment from S1 digestion, establishing the orientation of the γ3 restriction map with the mRNA shown in Figure 5.

The orientation of the μ, γ2a, and γ1 sequences was apparent from the correlation of restriction data with amino acid sequences described below.

Restriction Endonuclease Mapping. The first approach used was complete digestion of unlabeled insert DNA with various restriction enzymes, both singly and in pairwise combinations. This was most useful for mapping sites of enzymes which made only one or two scissions. The second involved partial cleavage of a DNA fragment labeled at one end; this produces a nested set of radioactive fragments in which the fragment sizes reveal the order and position of cleavage sites (Smith & Birnstiel, 1976). Partial digestion was particularly useful for mapping multiple sites, but it proved necessary to check the complete digestion products.

The μ Map. *Pst*I digestion of pH76μ.17 yields two cDNA fragments, 400 and 1400 bp long, which were separable by sucrose gradient centrifugation. The restriction map of each was determined separately and later combined into the complete map (Figure 4). In the large *Pst*I fragment, the *Bam*HI, *Sma*I, *Bgl*II, and *Pvu*II sites were mapped by complete digestion. A large number of sites were then mapped by partial digestion of fragments a–e in Figure 4. Fragments a and b were generated by terminally labeling the *Pst*I fragment and recleaving with *Pvu*II. Fragments c–e were generated by labeling a *Bam*HI digest of the whole plasmid and recleaving at *Sma*I and *Pst*I sites. The data from these two overlapping sets of partial products were entirely consistent.

The map of the shorter fragment was determined only by the complete digestion approach, using the unique *Hinc*II and *Hph*I sites as reference points. To orient its map with respect to that of the larger *Pst*I fragment, we measured the distance between the *Sma*I site within the sequence of the large fragment and the *Hinc*II site within the sequence of the small fragment, by *Sma*I + *Hinc*II digestion of the whole plasmid. The restriction map of the C region of pH76μ.17 can be aligned unambiguously with the amino acid sequence of the MOPC 104E μ chain (Kehry et al., 1979), and amino acid residues corresponding to various restriction sites are nominated in Figure 4.

The γ1 and γ3 Maps. The maps of a γ1 and γ3 clone are given in Figure 5. In pM21γ1.1, only the *Pst*I site distal to

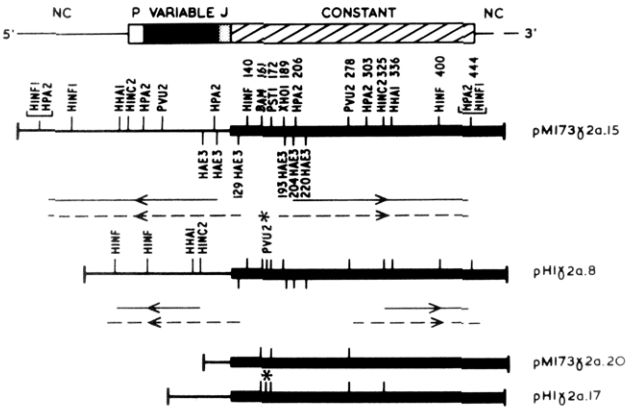

FIGURE 6: Restriction endonuclease cleavage maps of γ2a cDNA clones. The γ2a mRNA (top) is illustrated with the same convention as in Figure 4. Maps of the cDNA inserts in pM173γ2a.15, pH1γ2a.8, pM173γ2a.20, and pH1γ2a.17 are shown with their C region halves illustrated by thick lines. The lines below the first two clone maps indicate a region of inverted repeat sequence in those two clones but not in the other two (see text). The solid lines represent the minimal extent of the repeat inferred from the restriction maps, and the dashed lines represent the size and position of the repeat inferred from heteroduplexes; the arrowheads on the lines indicate 5′ to 3′ polarity with respect to the mRNA. A *Pvu*II site present in the C region of the HOPC 1 but not the MOPC 173 clones is indicated with an asterisk.

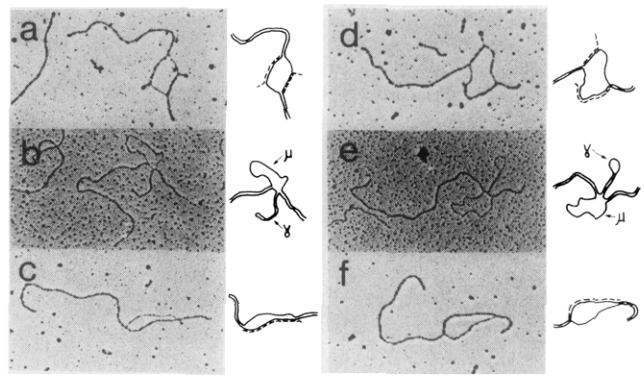

FIGURE 7: R loop and heteroduplex analysis of aberrant γ2a cDNA clones containing inverted repeats and of authentic γ2a cDNA clones. Plasmid DNAs were linearized with *Eco*RI. R loops were then formed as described (Gough et al., 1980a), and heteroduplexes were formed by denaturation of the DNAs (1 μg/mL each) at 80 °C for 3 min in R loop buffer (Gough et al., 1980a), followed by renaturation for 15 min at 30 °C. Both were mounted for electron microscopy essentially as described (Davis et al., 1971; Phillipsen et al., 1979). R loops: (a) aberrant clone pM173γ2a.15 with MOPC 173 γ2a mRNA; (c) authentic clone pM173γ2a.20 with MOPC 173 γ2a mRNA; (d) aberrant clone pH1γ2a.8 with HOPC 1 γ2a mRNA; (f) authentic clone pH1γ2a.17 with HOPC 1 γ2a mRNA. Heteroduplexes: (b) pM173γ2a.15 and pH76μ.17; (e) pH1γ2a.8 and pH76μ.17. The dashed lines represent mRNA. In the heteroduplexes, the pM173γ2a.15 or pH1γ2a.8 strand is indicated by γ and the pH76μ.17 strand by μ.

the plasmid *Eco*RI site can be cleaved, but an *Eco*RI + *Pst*I digest releases a 1490-bp fragment, 748 bp of which is pBR322 DNA (Sutcliffe, 1978). Hence, the length of the γ1 cDNA sequence is about 740 bp. As indicated in Figure 5A, the pM21γ1.1 sequence was mapped by partial digestion of plasmid DNA labeled at the cleavable *Pst*I site (fragment a). The map can be aligned accurately with the known amino acid sequence of the MOPC 21 γ1 heavy chain (Adetugbo, 1978), the restriction map of a cloned genomic Cγ1 sequence (Sakano et al., 1979), and the nucleotide sequence of an independent MOPC 21 γ1 cDNA clone (Rogers et al., 1979). pM21γ1.1 contains sequences corresponding to the CH2 and CH3 domains, as well as about 80 bp of the 3′ noncoding region (Figure 5A).

The restriction map of the γ3 cDNA insert was generated by partial digestion of the two fragments (b and c) indicated in Figure 5B. As no amino acid sequence data are available for mouse γ3 chains, the map cannot be aligned with the γ3 mRNA. However, this cloned sequence hybridizes only to genomic Cγ sequences (S. Cory, unpublished experiments) and thus probably does not contain any V region sequence.

Certain γ2a Clones Contain an Inverted Repeat. To facilitate mapping of the γ2a sequence in pM173γ2a.15, we subcloned the two *Pst*I fragments (Figure 6), which were then analyzed by the complete digestion approach. pH1γ2a.8 was also mapped by complete digestion, but without separation of the Pst fragments. All the sites within the C region of both clones can be aligned with the MOPC 173 γ2a amino acid sequence (Rocca-Serra et al., 1975), as indicated in Figure 6.

We were surprised to note a curious symmetry in the restriction map of pM173γ2a.15 and of pH1γ2a.8. Thus, the regions indicated by the lines below the maps in Figure 6 appear to be mirror images of one another. This raised the possibility that each clone contained an inverted repeat sequence. To test this, we examined R loops (Thomas et al., 1976) formed between each plasmid DNA and the mRNA from which it was derived. In both cases, we observed R loops in which mRNA molecules had hybridized in a nonoverlapping fashion to each strand of the duplex cDNA sequence (Figures 7a and d). The results suggested that the insert in both pM173γ2a.15 and pH1γ2a.8 represents a section of the mRNA sequence plus an inverted repeat of part of that section. To verify this, we formed heteroduplexes between DNA from each γ2a cDNA clone and DNA from an unrelated cDNA clone pH76μ17. Both the pM173γ2a.15 heteroduplexes and those of pH1γ2a.8 exhibited "snapback" structures characteristic of inverted repeat sequences (Figures 7b and e). The pM173γ2a.15 heteroduplexes exhibited a hairpin structure (Figure 7b), indicating that only a very short sequence (<150 bp) separated the repeats, whereas pH1γ2a.8 heteroduplexes exhibited a "lollipop" structure (Figure 7e), indicative of a much larger sequence (540 bp) between the repeats. Measurement of the heteroduplexes, together with the restriction analysis, established that the inverted repeats were of C region sequences and that in each clone the authentic copy of the mRNA terminated near the V–C junction, as indicated in Figure 6. To determine whether this aberration was common to γ2a clones, we screened two other MOPC 173 γ2a cDNA clones and four other HOPC 1 γ2a clones by R loop and heteroduplex analysis. One more HOPC 1 γ2a clone (pH1γ2a.6) contained an inverted repeat, but the remaining five clones contained faithful copies of the mRNAs, as is demonstrated by the R loop with pM173γ2a.20 shown in Figure 7c and that with pH1γ2a.17 in Figure 7f.

To determine whether any of the nonaberrant clones bore V_H sequences, we analyzed each by digestion with *Bam*H1 and *Pvu*II. From the results, we determined that the largest MOPC 173 γ2a clone (pM173γ2a.20) contains about 110 bp of the V_H region and the largest HOPC 1 γ2a clone (pH1γ2a.17), about 250 bp, as indicated in Figure 6.

A difference between the Cγ2a regions cloned in pM173γ2a.15 and pH1γ2a.8 was found on *Pvu*II digestion, which split the pH1γ2a.8 insert near the *Pst*1 site (within 30 bp) but not the pM173γ2a.15 insert (see Discussion).

The α Maps. Restriction sites in the pS107α.4 and pM603α.8 clones (Figure 8) were mapped by the complete digestion approach. The maps are compatible with one

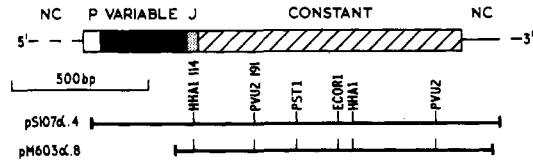

FIGURE 8: Restriction endonuclease cleavage maps of two α cDNA clones. The α mRNA (top) is shown with the same convention as in Figure 4. Maps of the cDNA inserts in pS107α.4 and pM603α.8 are shown, and the amino acid residues corresponding to the two restriction sites which occur within regions of known protein sequence are nominated.

Table IV: Summary of Information on Selected Heavy Chain cDNA Clones[a]

plasmid	chain	orientation[b]	length of tails (bp)[c]	length of insert (bp)	sequences represented
pH76μ.17	μ	V	13, 12	1850	V, J, Cμ
pS107α.4	α	C	19, 15	1520	V, J, Cα
pM603α.8	α	V	16, 11	1190	J, Cα
pM21γ1.1	γ1	V	*d*	740	Cγ1
pM173γ2a.15	γ2a	V	18, 17	1900[e]	J, Cγ2a
pM173γ2a.20	γ2a	V	*d*	1230	V, J, Cγ2a
pH1γ2a.8	γ2a	C	17, 7	1720[e]	J, Cγ2a
pH1γ2a.17	γ2a	V	*d*	1355	V, J, Cγ2a
pY5606γ3.15	γ3	*d*	*d*	1050	Cγ3

[a] The more extensively studied clones from each of the seven mRNAs are presented. [b] The region of the insert sequence proximal to the *Eco*RI site of the plasmid is given. The orientation of the inserts in pS107α.4 and pM603α.8 was determined as in Figure 3. The orientation of the other sequences was determined by restriction enzyme digestion of the chimeric plasmid. Note that pM21γ1.1 does not in fact contain any V region sequence. [c] The lengths of the oligo(dG·dC) tails were determined by depurination of terminally labeled insert DNA (Gough et al., 1980b). [d] Not determined. [e] These clones contain long inverted repeat sequences (see text).

presented recently by Early et al. (1979) for an independently cloned S107 α cDNA sequence. The amino acid residues corresponding to the two restriction sites which occur within regions of known protein sequence (Kabat et al., 1976; Robinson & Appella, 1977) are nominated in Figure 8. The map of pS107α.4 indicates that it contains the entire protein coding sequence and very short sequences from 3′ and 5′ noncoding regions. The pM603α.8 sequence extends a similar distance toward the 3′ end of the mRNA, but contains only ~100 bp of the V region.

Discussion

Clone Identification by Comparative Hybridization. Often only an impure mRNA or cDNA probe is available for screening plasmid clones, and this can hamper clone identification, particularly if contaminants in the probe are of comparable abundance to the relevant sequence. This problem may be largely overcome by comparative colony hybridization. In principle, the only requirement is one probe which contains the relevant sequence and a second probe having a similar complement of irrelevant sequences but which either lacks the specific sequence, or has a markedly reduced level of it. Irrelevant clones will bear sequences common to the two probes and hence can be readily recognized. Examples of applications would include preparation of one probe from a normal cell line and another from a mutant line lacking specific sequences in order to identify clones bearing those sequences. Other types of clones could be identified by using matched pairs of probes derived from cells in different physiological states, from closely related tissues, from the same tissue in related species, or from virally infected and uninfected cells. Schemes of this sort have proven very effective in identifying cloned sequences specific to different stages of development in *Drosophila melanogaster* (D. Kemp, personal communication). Our experience suggests that this approach works well if the desired sequence represents at least a few percent of the homologous probe. Rarer sequences may require a probe enrichment technique such as that devised by Alt et al. (1978).

A Common Contaminant of Chain mRNA Preparations. Microsomal polyadenylated RNA from certain plasmacytomas (notably MOPC 41A and MOPC 173) contains a prominent unidentified 24S peak which we have designated peak A (S. Cory, unpublished experiments). Partially degraded peak A molecules appear to contaminate the 17–19S heavy chain mRNA fraction, since clones registering a strong signal with a peak A probe have been obtained from every type of heavy chain mRNA that we have attempted to clone, as well as some light chain preparations. Intracisternal type A particles which contain a polyadenylated viral RNA are present in plasmacytomas and would be expected to contaminate microsome preparations (Lueders et al., 1977). We suggest that peak A RNA contains such sequences, because peak A sequences are highly reiterated in the mouse genome (S. Cory, unpublished experiments), as are sequences from type A particles (Lueders & Kuff, 1977).

Identity of the Heavy Chain Clones. That the clones studied in detail bear Ig heavy chain sequences is strongly supported by three general types of evidence. First, all gave reproducibly strong colony hybridization with a cDNA probe made on a homologous mRNA preparation, which was substantially enriched for a heavy chain sequence, while none hybridized with cDNA made on mRNA preparations from plasmacytomas synthesizing other types of Ig heavy chains (Tables I–III). Hence, each cloned sequence corresponds to an abundant mRNA species present only in the tumor of origin. Secondly, representative clones of each heavy chain type were shown to arrest translation of the corresponding mRNA in the hybrid-arrest assay of Paterson et al. (1977). Finally, restriction maps of the cloned sequences can be matched with known amino acid sequences (Figures 4–6), except for γ3, for which no amino acid sequence data are available. The correlation is particularly compelling for the μ and γ2a clones, for which detailed maps were generated (Figures 4 and 6).

Other evidence verifies the identity of particular clones. For μ clone pH76μ.17, several hundred nucleotides have now been sequenced (Bernard & Gough, 1980). Restriction analysis of pM173γ2a.15 and pHγ2a.8 (Figure 6) indicated that closely related sequences had been cloned from separate tumors, as in the case of pS107α.4 and pM603α.8 (Figure 8). Moreover, our restriction data on the α clones are compatible with those reported during the course of this work for another S107 α clone by Early et al. (1979), as are our data on pM21γ1.1 with that reported by Sakano et al. (1979) and Rogers et al. (1979) for two other MOPC 21 γ1 clones. Finally, the identity of all three types of γ clones is supported by evidence discussed below that their sequences are partially homologous.

Homology between Heavy Chain Sequences. As we anticipated from the marked differences in μ, α, and γ amino acid sequences, our colony hybridization data indicate that their nucleotide sequences exhibit no detectable homology, and this must extend to 3′ noncoding segments. Yamawaki-Kataoka et al. (1979) and Marcu et al. (1978) have reached the same conclusion from the hybridization of mRNA and cDNA in solution. Colony hybridization also gave no more than marginal cross-hybridization between the three types of γ sequences examined (Table III), but we are uncertain about the interpretation of this result, because other experiments in

this laboratory indicate that the cloned γ2a cDNA sequence does hybridize to some extent with genomic γ3 and γ1 sequences and that the cloned γ3 cDNA sequence hybridizes with a genomic γ2 sequence and to a very limited extent with a genomic γ1 sequence (our unpublished experiments). These conclusions are similar to those reached by Yamawaki-Kataoka et al. (1979).

Inverted Repeat Sequences Generated by a Cloning Aberration. Unexpectedly, we found that three of our γ2a clones contain inverted repeat sequences. Measurements on snapback structures (Figures 7b and e) and restriction mapping (Figure 6) indicated that the length of the repeat in pM173γ2a.15 is 720–770 bp and that in pH1γ2a.8 is 360–480 bp. R loop analysis established that these repeats were not present in the mRNAs, nor were such repeats present in five other γ2a clones derived from the same mRNAs. Hence, the repeats must be generated by the cloning procedure. We apparently inadvertently selected two clones having this aberration by choosing the largest γ2a clones for detailed analysis.

The mechanism by which the inverted repeats were formed is not known, but we favor the notion that it occurred during cDNA synthesis. Duplex cDNA synthesis required for cloning is thought to proceed through a hairpin stage. If the hairpin "breathes", the 3′ terminus might fold back onto the same strand and prime renewed synthesis by DNA polymerase I. The ensuing synthesis would convert the cDNA molecule to a larger hairpin bearing an inverted repeat. Why such structures were found only in the γ2a clones is not clear. One possibility is that the γ2a mRNA contains short inverted repeats which facilitate the aberrant priming hypothesized above.

A Sequence Difference in Two Cγ2a Regions. The restriction analysis revealed a sequence difference within the Cγ2a regions of MOPC 173 and HOPC 1, namely, a *Pvu*II site present near the *Pst*I site in pH1γ2a.8, but not in pM173γ2a.15. This site does not lie within the inverted repeat. We do not think the difference represents a cloning aberration, because two MOPC 173 clones (.15 and .20) were shown to lack the sites and four HOPC 1 clones (.8, .17, .20, and .21) to possess it. Hence, the difference is probably present in the MOPC 173 and HOPC 1 γ2a mRNAs. Possible explanations include a residual polymorphism in BALB/c mice, a mutation within one of the tumor lines, or a second Cγ2a gene.

Sequences Represented in the Clones. Table IV summarizes data on the more extensively studied clones from the seven heavy chain mRNAs. Three clones (pH76μ.17, pS107α.4, and pH1γ2a.17) contain large V_H as well as C_H sequences, while others contain a complete Cα sequence plus a short V_H sequence (pM603α.8), a complete Cγ2a sequence plus a short V_H sequence (pM173γ2a.17), most of a Cγ3 sequence (pY5606γ3.15), and the final two domains of a Cγ1 sequence (pM21γ1.1). Noncoding sequences are also represented in some clones.

Recent evidence suggests that the amino acid sequence immediately preceding the C region in heavy chains, as in light chains, is encoded not by a V gene but by an independent joining region (J_H) gene (Rao et al., 1979; Gough et al., 1980a). Our data indicate that a number of the clones contain J_H nucleotide sequences.

The cloned sequences thus represent five different C_H, three V_H, and four J_H sequences. Hence, they provide a set of well-characterized pure nucleotide sequences for analysis of the heavy chain gene system. Indeed, the value of these sequences has already been demonstrated in the cloning and characterization of a number of genomic V_H genes (Kemp et al., 1979) and a Cμ gene (Gough et al., 1980a).

Added in Proof

Nucleotide sequence analysis (O. Bernard and N. Gough, unpublished experiments) indicates that the restriction site difference between the HOPC 1 and the MOPC 173 Cγ2a region is due to a one nucleotide difference.

Acknowledgments

We thank Dr. David Kemp for useful advice and Jan Holton for making mRNA preparations.

References

Adetugbo, K. (1978) *J. Biol. Chem. 253*, 6068.
Alt, F. W., Kellems, R. E., Bertino, J. R., & Schimke, R. T. (1978) *J. Biol. Chem. 253*, 1357.
Barstad, P., Farnsworth, V., Weigert, M., Cohn, M., & Hood, L. (1974) *Proc. Natl. Acad Sci. U.S.A. 71*, 4096.
Berk, A. J., & Sharp, P. A. (1978) *Cell 14*, 695.
Bernard, O., & Gough, N. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77* (in press).
Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., & Boyer, H. W. (1977) *Gene 2*, 95.
Brack, C., Hirama, M., Lenhard-Schuller, R., & Tonegawa, S. (1978) *Cell 15*, 1.
Cory, S., Genin, C., & Adams, J. M. (1976) *Biochim. Biophys. Acta 454*, 248.
Davis, R. W., Simon, M., & Davidson, N. (1971) *Methods Enzymol. 21*, 413–428.
Early, P. W., Davis, M. M., Kaback, D. B., Davidson N., & Hood, L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 857.
Gally, J. A. (1973) in *The Antigens* (Sela, M., Ed.) Vol. 1, p 162, Academic Press, New York.
Gally, J. A., & Edelman, G. M. (1972) *Annu. Rev. Genet. 6*, 1.
Gough, N. M., Kemp, D. J., Tyler, B. M., Adams, J. M., & Cory, S. (1980a) *Proc. Natl. Acad. Sci. U.S.A. 77*, 554.
Gough, N. M., Webb, E. A., Cory, S., & Adams, J. M. (1980b) *Biochemistry* (preceding paper in this issue).
Grunstein, M., & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3962.
Hunt, T., & Jackson, R. J. (1974) in *Modern Trends in Human Leukaemia* (Neth, R., Gallo, R. C., Spigelman, S., & Stohlman, F., Eds.) pp 300–307, J. F. Lehmanns, Verlag, Munich.
Kabat, E. A., Wu, T. T., & Bilofsky, H. (1976) *Variable Regions of Immunoglobulin Chains*, Medical Computer Systems, Bolt, Beranek and Newman Inc., Cambridge, MA.
Kehry, M., Sibley, C., Fuhrman, J., Schilling. J., & Hood, L. E. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2932.
Kemp, D. J., Cory, S., & Adams, J. M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4627.
Leuders, K. K., & Kuff, E. L. (1977) *Cell 12*, 963.
Leuders, K. K., Segal, S., & Kuff, E. L. (1977) *Cell 11*, 83.
Marcu, K. B., Valbuena, O., & Perry, R. P. (1978) *Biochemistry 17*, 1723.
Paterson, B. M., Roberts, B. E., & Kuff, E. L. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4370.
Pelham, H. R. B., & Jackson, R. J. (1976) *Eur. J. Biochem. 67*, 247.
Phillipsen, P., Thomas, M., Kramer, R. A., & Davis, R. W. (1978) *J. Mol. Biol. 123*, 387.
Potter, M. (1972) *Physiol. Rev. 52*, 631.
Rao, D. N., Rudikoff, S., Krutzsch, H., & Potter, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2890.
Robinson, E. A., & Appella, E. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2465.
Rocca-Serra, J., Milili, M., & Fougereau, M. (1975) *Eur. J. Biochem. 59*, 511.

Rogers, J., Clarke, P., & Salser, W. (1979) *Nucleic Acids Res. 6*, 3305.
Sakano, H., Rogers, J. H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R., & Tonegawa, S. (1979) *Nature (London) 277*, 627.
Sanger, F., & Coulson, A. R. (1978) *FEBS Lett. 87*, 107.
Schibler, U., Marcu, K. B., & Perry, R. P. (1978) *Cell 15*, 1495.
Seidman, J. G., & Leder, P. (1978) *Nature (London) 276*, 790.
Smith, H. O., & Birnstiel, M. L. (1976) *Nucleic Acids Res. 3*, 2387.
Spiegelberg, H. L. (1974) *Adv. Immunol. 19*, 259.
Sutcliffe, J. G. (1978) *Nucleic Acids Res. 5*, 2721.
Thomas, M., White, R. L., & Davis, R. W. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2294.
Williamson, A. R. (1976) *Annu. Rev. Biochem. 45*, 467.
Yamawaki-Kataoka, Y., Sato, K., Shimizu, A., Kataoka, T., Mano, Y., Ono, M., Kawakami, M., & Honjo, T. (1979) *Biochemistry 18*, 490.

In Vitro Processing of Intervening Sequences in the Precursors of Messenger Ribonucleic Acid for Adenovirus 2 Deoxyribonucleic Acid Binding Protein†

Carlos J. Goldenberg and Heschel J. Raskas*

ABSTRACT: Early region 2 of the adenovirus 2 genome (map positions 61–75) specifies two polyadenylated nuclear RNAs (28 S and 23 S) that appear to be precursors of the 20S cytoplasmic mRNA [Goldenberg, C. J., & Raskas, H. J. (1979) *Cell 16*, 131–138]. Isolated nuclei were used to study the processing of region 2 RNA in vitro. Cultures infected with adenovirus 2 were pulse labeled with [^{3}H]uridine early in infection. Nuclei were purified from the labeled cultures and then incubated in vitro. Incubations were performed in the absence or presence of a cytosol extract isolated in hypotonic conditions from uninfected cells. During in vitro incubation, nuclear 28S and 23S region 2 RNAs were converted to 20S molecules. In the presence of the cytosol, the conversion to 20S RNA molecules was greatly accelerated and almost completed after 10 min of incubation. The conversion was quantitative, and the resulting 20S RNAs were stable for at least 30 additional min. The cytosol activity was concentration dependent and temperature sensitive. A cytosol prepared in isotonic conditions was inactive, suggesting that the processing activity was nuclear in origin. Early region 2 RNAs processed in vitro were analyzed for splicing patterns. RNA was purified by using a preparative membrane hybridization–selection procedure. The selected RNAs (20 S) were fractionated by size and then rehybridized to a series of DNA fragments. The data demonstrated (1) that the processed 20S RNAs had lost sequences as expected on the basis of known splicing patterns and (2) that these 20S RNAs contained leader sequences spliced to the body of the RNA molecule.

At least four regions of the adenovirus 2 genome are expressed at early times [see review by Flint (1977)]. Each of these early regions has at least one independent promoter for the initiation of transcription (Craig & Raskas, 1976; Berk & Sharp, 1977; Evans et al., 1977). We have been studying the processing of RNAs synthesized by early region 2. The early mRNA from this region is specified by DNA sequences within map positions 61–75 (Flint & Sharp, 1974; Pettersson et al., 1976); the mRNA is copied from the strand transcribed in the leftward direction (Sharp et al., 1974). The cytoplasmic mRNA consists of 70 nucleotides from positions 75 to 74.6 spliced to 170 nucleotides from 68.8 to 68.3 and 1700 nucleotides from 66.3 to 61.6 (Kitchingman et al., 1977; Berk & Sharp, 1978; see Figure 1). The polypeptide product of this mRNA is a 72 000-dalton DNA binding protein (Ginsberg et al., 1974; Grodziker et al., 1974; Van der Vliet et al., 1975; Lewis et al., 1976).

We have characterized the structures of three polyadenylated nuclear RNAs that appear to be precursors of the cytoplasmic mRNA for the DNA binding protein (Craig & Raskas, 1976; Goldenberg & Raskas, 1979) (Figure 1). All the intervening DNA sequences of the gene for the 72K DNA binding protein are contained in the largest (28 S) mRNA precursor. Apparently these sequences are removed sequentially to generate first a 23S intermediate and then a 20S nuclear RNA that is transported to the cytoplasm as mature mRNA. Our studies of the region 2 nuclear RNAs provide support for a processing model in which mRNA is generated by RNA cleavage followed by splicing (Klessig, 1977; Berget et al., 1977; Chow et al., 1977). The internal cleavage of mRNA precursors requires activities that recognize specific sequences; the resulting termini must then be ligated to form a covalent bond. Recently, splicing of yeast tRNA in vitro has been reported. The tRNA precursor contains a 15-nucleotide intervening sequence that is not present in the mature molecules (Valenzuela et al., 1978). A cell-free enzymatic activity which can remove the intervening sequences and religate the ends has been identified (Knapp et al., 1978, 1979; Peebles et al., 1979).

Several years ago, our laboratory described an in vitro nuclear system in which adenovirus high molecular weight RNA was converted to the size of mRNAs (Brunner & Raskas, 1972). These studies preceded the availability of specific DNA fragments, and therefore it was not possible to analyze pre-

† From the Department of Pathology, Washington University School of Medicine, St. Louis, Missouri 63110. *Received December 6, 1979.* This study was supported by NIH Grants CA09129 and CA15007 and American Cancer Society Grant VC-94E. Cell culture media were prepared in a cancer center facility funded by the National Cancer Institute. This study was also supported by the following companies: Brown & Williamson Tobacco Corp.; Larus and Brothers Co., Inc.; Liggett and Myers, Inc.; Lorillard, a Division of Loews Theatres, Inc.; Philip Morris, Inc.; R. J. Reynolds Tobacco Co.; United States Tobacco Co.; Tobacco Associates, Inc.

0006-2960/80/0419-2719$01.00/0