REITERATION FREQUENCY OF ANTIBODY GENES
Studies with κ-chain mRNA

S. TONEGAWA*, A. BERNARDINI*, B.J. WEIMANN+ and C. STEINBERG*

*Basel Institute for Immunology,
Grenzacherstrasse 487,
CH-4058 Basel, Switzerland

and

+Max-Planck Institut für Experimentelle Medizin,
Abteilung Molekulare Biologie,
Hermann-Rein-Strasse 3,
D-34 Göttingen, West Germany

Received 27 December 1973

1. Introduction

Two contrasting theories have been proposed in order to explain the diversity of antibody molecules: the germ line theory and the somatic mutation (or recombination) theory. Arguments for and against both theories have been exhaustively discussed in the literature. One of the few direct experimental approaches to the issue is to determine, by nucleic acid hybridization techniques, the number of genes coding for the variable region of immunoglobulin polypeptide chains. In this report we briefly describe our attempts along this line. The results strongly suggest that in the mouse there are a very limited number – possibly a single copy per subgroup – of genes coding for the variable region of κ-chains.

2. Methods

Mouse κ-chain mRNA was purified from plasmacytomas by a modification of the method described previously [1]; details will be reported elsewhere. The final step of the purification was preparative polyacrylamide gel electrophoresis in 99% formamide as described by Gould and Hamlyn [2].

RNA preparations were iodinated by the method of Commerford [3] to a specific activity of about 10^7 cpm/μg. Free and weakly bound iodine was removed by heating at 60°C for 30 min in the presence of 0.1 M Na_2SO_3 and passing through a hydroxyapatite column. Iodination does not detectably change the DNA-RNA hybridization reaction [3].

DNA was prepared from BALB/c mouse liver or MOPC 7OE tumor according to the method of Bishop [4]. Hybridization was carried out at 70°C in 2 X SSC, following the procedures described by Melli et al. [5] as closely as possible.

3. Results

3.1. *Preparation of κ-chain mRNA*

Fig. 1 shows the acrylamide gel electrophoresis pattern of MOPC 70E 13S RNA which had been purified from membrane-bound polysomes by poly dT-cellulose chromatography and repeated sucrose-density gradient centrifugation. The κ-chain mRNA activity was kindly assayed by M. Schreier and T. Staehelin in a cell-free translation system. This activity is

Correspondence should be addressed to: S. Tonegawa at the Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland.

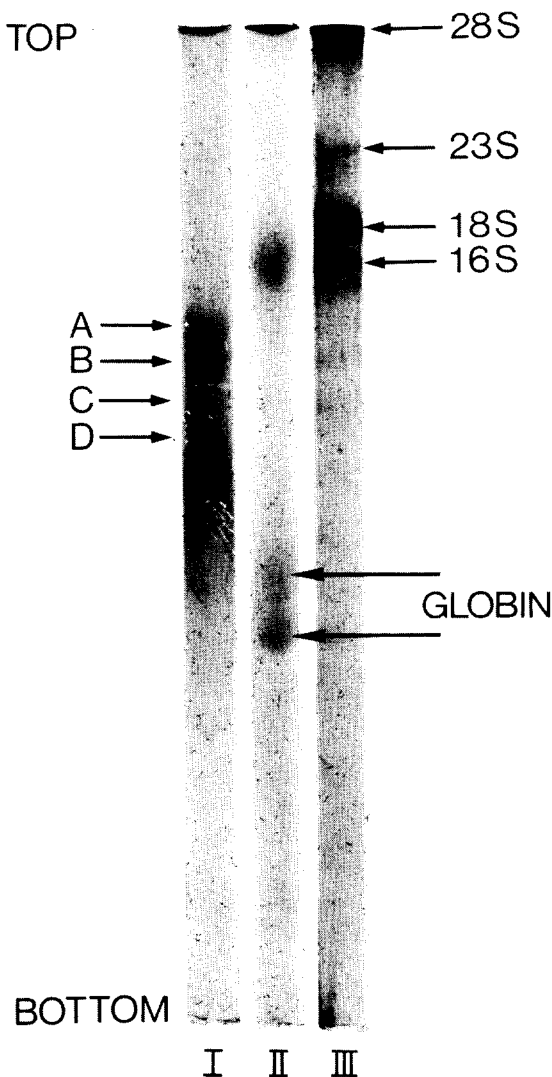

Fig. 1. Polyacrylamide gel electrophoresis of MOPC 70E 13 S RNA in 99% formamide. Gels were stained by Pyronin Y. Gel I: 20 μg 13 S MOPC 70E mRNA. Gel II: 10 μg 9 S rabbit globin mRNA. Gel III: Mixture of 16 μg 18 S and 28 S MOPC 70E rRNA and 17 μg 16 S and 23 S *E. coli* rRNA (Miles Laboratory).

associated with the major band A. The molecular weight of band A RNA was determined using MOPC 70E and *E. coli* ribosomal RNA, and rabbit globin mRNA as the migration standards. We obtained $4.0 \pm 0.2 \times 10^5$ daltons which corresponds to 1180 ± 60 nucleotides. Very little κ-chain mRNA activity was found in band B, which seems to code for another protein which migrates in SDS-polyacrylamide gel electrophoresis about 25% slower than κ-chain secreted by 70E. There are also two or three minor bands (Bands C,D) which migrate faster than the κ-chain mRNA. The ratio of RNA in band A to that in band B varies somewhat from one preparation of 13 S RNA to another within the range of 3:1 to 5:1. RNA was eluted from a 2 mm thick gel slice cut out of the center of band A. The distance between the centers of bands A and B was 7 mm. Eluted RNA was iodinated and used in the hybridization studies.

3.2. *Hybridization of κ-chain mRNA*

The hybridization of mouse liver DNA with MOPC 70E band A RNA and with 18 S ribosomal RNA is shown in fig. 2. The Cot curve for the mRNA is rather broader than that for rRNA, indicating heterogeneity. In fact, the mRNA curve can be decomposed into two components. The major component hybridizes with an apparent Cot½ of 1320. This component hybridizes as uniformly as rRNA. Taking *E. coli* cRNA as a standard, we obtained a maximal reiteration frequency of 4 for the major component. As discussed by Bishop et al. [6], this is an extreme upper limit. Henceforth, we will consider this component as consisting of unique sequences. Using the same standardization, we obtain a reiteration frequency of 250 for the 18 S ribosomal RNA gene. This is in good agreement with reported estimates (see ref. [5]). About 16% of the hybridizable RNA seems to react with a Cot½ of approximately 23, indicating a gene reiteration frequency of 230. While this fraction may not represent a homogeneous component, it is seen in all experiments where the points are closely spaced. Competition experiments show that it is not ribosomal RNA.

In other experiments, the range of Cot values corresponding to gene reiteration frequencies of 1–100 thousand was carefully explored, but no appreciable hybridization was found.

MOPC 70E band A RNA was also hybridized with DNA from the same tumor. The result was essentially the same as that shown in fig. 2, indicating that there is no specific amplification of κ-chain gene(s) in the myeloma cells.

3.3. *Competition experiments*

I^{125}-labeled MOPC 70E band A RNA was hy-

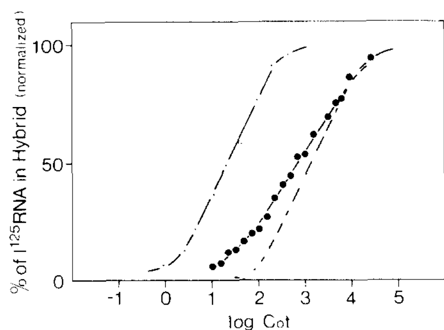

Fig. 2. Hybridization of MOPC 70E mRNA or 18 S rRNA with mouse liver DNA. (•—•—•) MOPC 70E band A RNA vs mouse liver DNA; DNA concentration, 22.4 mg/ml; DNA/RNA ratio, 1.1×10^7. After subtracting the intrinsic RNase resistant fraction (2.1%), the fraction of I^{125}-labelled RNA in the hybrid at each Cot value was normalized taking the maximum hybridization level (51%) as 100%. (○——○——○) The major component of the Cot curve was replotted after subtracting the contribution of the low Cot RNA. (•——•——•) 18 S rRNA vs mouse liver DNA; DNA concentration 2.2 mg/ml; DNA/RNA ratio, 5.5×10^5; intrinsic RNase resistant fraction, 4.3%; maximum hybridization level, 67.0%.

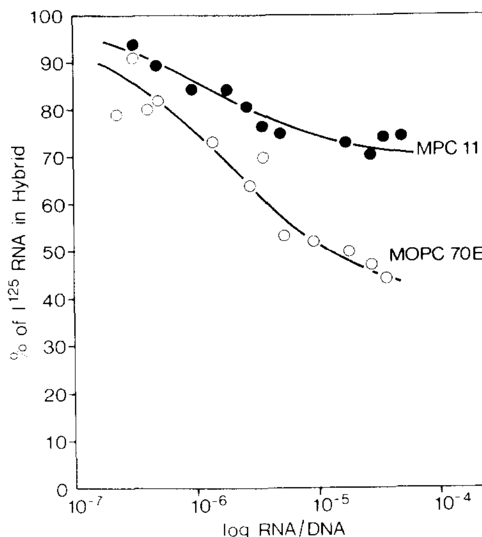

Fig. 3. Competition hybridization of I^{125}-labelled MOPC 70E band A RNA with mouse liver DNA in the presence of unlabeled MOPC 70E band A RNA (○——○——○) or unlabeled MPC 11 band A RNA (•——•——•). All samples were incubated to Cot = 6210; intrinsic RNase resistant fraction, 2.0%. The data are normalized against the hybridization level obtained without competitor (34.2%).

bridized with mouse liver DNA in the presence of varying amounts of unlabeled band A RNA prepared either from MOPC 70E cells or MPC 11 cells. Homologous competitor prevented 55% of the labeled RNA from hybridizing, while heterologous RNA competed out only 30%.

The amounts of competitor used were such that only unique or nearly unique sequences would be competed. The 16% of the RNA which hybridized at low Cot values (fig. 2) will not be competed. Since the hybridization mixture was incubated only to Cot 6210, where only 80% of the maximal hybridization level is reached (fig. 2), the estimate of low Cot RNA is increased from 16% to 20%. Furthermore, the κ-chain mRNA may not be 100% pure. Work by Brownless et al. [7] and preliminary fingerprinting studies of our mRNA (S. Dube, personal communication) suggest that the band A κ-chain mRNA used is contaminated by many different RNA species which may amount to 20–30%. The ratio of RNA to DNA for any individual species of contaminant is too low to allow effective competition with the amounts of competing RNA used.

The amount of competition (30%) by the heterologous RNA is a measure of sequences which are common to both mRNAs. The difference (25%) between the amount of competition by heterologous and by homologous mRNA is a measure of sequences which are specific to MOPC 70E. The ratio of RNA to DNA at which a half maximal competition was reached is a measure of gene reiteration frequency independent of the one obtained from the Cot curve. Apparent reiteration frequencies of 7.8 and 7.2 are obtained from the homologous and heterologous competition curves, respectively (fig. 3). Hence, the reiteration frequency seems to be the same for common and specific sequences.

Three correction factors must be applied to the apparent reiteration frequencies. These are due to: (a) impurities, (b) insufficient annealing, and (c) differential rate constants for DNA–DNA and DNA–RNA hybridization. The corrected reiteration frequencies are about 2–3. Thus, the reiteration frequency determined by this method agrees rather well with that determined from the Cot curve.

4. Discussion

The mouse κ-chain mRNA is composed of four sections. The poly A sequence of 200 nucleotides [7] and the 3′ end of the polynucleotide chain is not labeled by iodine and will not concern us further. Two sections, each of ~ 327 nucleotides (33.5% of the labeled portion of the molecule) code for the V and C regions of κ-chains. The function of the remaining section of ~ 322 nucleotides (33%) is unknown except that at least 30 to 60 nucleotides in this section should be involved in coding for the extra peptide [1,11–13] which is cleaved off before the κ-chain is secreted.

In the present work, in which κ-chain mRNA preparations with 70–80% purity was used, we have shown that about 84% of sequences hybridize with mouse liver or MOPC 70E DNA in a manner which is expected for non-reiterated genes (fig. 2). Of these 84%, about 32% are sequences common both in MOPC 70E and MPC 11 mRNA as demonstrated by the heterologous competition curve (data of fig. 3, corrected for insufficient annealing). Genetic studies indicate that the κ-C gene, at least in man and the rabbit, is not reiterated [14]. Thus, most of these common sequences are attributable to the constant region. The contaminating RNAs which could amount to 20–30% are very likely to be common in the two RNA preparations but they should not contribute much to this fraction for the reasons discussed above (part 3.3.).

Of the 84% of unique sequences, about 26% are uncommon between MOPC 70E and MPC 11 κ-chain preparations, as visualized by the difference in the final plateau levels of the homologous and heterologous competition curves (fig. 3). Using published amino acid sequences [8,9] and assuming that the minimum base sequence homology which hybridizes under our conditions is 15 consecutive bases [10], the variable region sequences of MPC 11 κ-chain mRNA are not expected to compete at all with those of MOPC 70E κ-chain mRNA. Therefore, the straightforward interpretation of this fraction is that these nucleotide sequences code for subgroup specific amino acid sequences in the variable region. The contaminating RNAs are most likely a population of mRNAs having a size very similar to that of κ-chain. It would be expected that most mRNA molecules would come from unreiterated genes. If this is so, most of the remaining unique sequences (about 26%) can be accounted for by the contaminating mRNAs.

The above considerations lead us to argue that the major part of the RNA in the first transition of the Cot curve is due to the sequences of κ-chain mRNA outside of the V and C regions. These sequences are reiterated roughly 200 times and comprise about 16% of the κ-chain mRNA preparation. A number of possibilities could be raised for sequences with such a reiteration frequency. These include sequences coding for the extra peptide, sequences involved in joining of the V and C genes, and sequences recognized by a particular component(s) of transcription and/or translation. It is, however, also probable that the limited sequences within the variable region which are common in all or most of the κ-chains are partly responsible for the reiterated RNA fraction. In any case, if these sequences are unique to and common in genes coding for κ-chains, then the reiteration frequency of roughly 200 may reflect the total number of mouse κ-chain V genes. M. Cohn, from an analysis of amino acid sequence of mouse light chains, proposed that there may be as many as 100 light chain subgroups in the mouse [15].

More extensive competition experiments now in progress should clarify some of the uncertain points discussed here.

Acknowledgements

We gratefully acknowledge the excellent technical assistance offered by Ms. Monica Schöld, Rita Schuler and Irene Vonzum. We thank Drs. Max Schreier and Theophil Staehelin for making their cell-free translation system available to us.

References

[1] Tonegawa, S. and Baldi, I. (1973) Biochem. Biophys. Res. Commun. 51, 81.
[2] Gould, H.J. and Hamlyn, P.H. (1973) FEBS Letters 30, 301.
[3] Commerford, S.L. (1971) Biochemistry 10, 1993.
[4] Bishop, J.O. (1972) Biochem. J. 126, 171.
[5] Melli, M., Whitfield, C., Rao, K.V., Richardson, M. and Bishop, J.O. (1971) Nature New Biol. 231, 8.

[6] Bishop, J.O., Pemberton, R. and Baglioni, C. (1972) Nature New Biol. 235, 231.
[7] Brownlee, G.G., Cartwright, E.M., Cowan, N.J., Jarvis, J.M. and Milstein, C. (1973) Nature New Biol. 244, 236.
[8] Gray, W.R., Dreyer, W.J. and Hood, L. (1967) Science 155, 465.
[9] Smith, G.P. (1973) Science 181, 941.
[10] Niyogi, S.K. and Thomas, C.A. (1967) Biochem. Biophys. Res. Commun. 26, 51.
[11] Brownless, G.G., Harrison, T.M., Mathews, M.B. and Milstein, C. (1972) FEBS Letters 23, 244.
[12] Swan, D., Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. U.S. 69, 1967.
[13] Mach, B., Faust, C. and Vassali, P. (1973) Proc. Natl. Acad. Sci. U.S. 70, 451.
[14] Gally, J.A. and Edelman, G.M. (1972) in: Annual Review of Genetics (Roman, H.L. ed.), pp. 1–46, Annual Reviews Inc., Palo Alto, California.
[15] Cohn, M. (1973) in: The Biochemistry of Gene Expression in Higher Organisms (Pollak, J.K. and Lee, J.W. eds.), pp. 574–592, D. Reidel, Dordrecht, Holland.