Prof.Devons:

A

Chumban Telk

Good evening, ladies and gentlemen and welcome once again for a brief interval to this meeting of the Symposium on the Relationship of Biological and Physical Sciences. This Symposium is, as you know, supported by the NYState Foundation for Science & Technology. This meeting we have this evening is a delayed pleasure - it's been our ninth meeting Unfortunately Dr. Nirenberg was ill on that occasion and he had to disappoint to you. I am glad to say, as you can see for yourself, he is here and well this evening and will talk to us tonight, and I am going to ask Professor I. Rabi, University Professor of Columbia University to introduce our speaker this evening.

Prof Rabi:

Ladies and Genglemen , Dr. Nirenberg deserves a very much better introducer than he is getting.tonight. Because I am no kind of biologist - I call myself a physicist - my only connection with biology in a practical way was to plant a few trees some 50 years ago , in Brooklyn of course, (laughter) but I am very glad to introduce Dr. Nirenberg . In the first place his subject doesn't sound very biological to me on the translation of the genetic code - either physics or linguistics - now Dr. N. was born in New York and when he was a helpless little boy was taken down to Florida, where he grew up and proved that we get things from Florida, good things from π Florida apart from grapefruit and suntans. He was graduated from the University of Florida in 1948 and then got an M.A. there and finally worked his way north to get a doctor's degree from the University of Michigan. I will not recount $\operatorname{mu} h$ further in his career or the various things which he did because *xxx actually I couldn't put the proper feeling

into the topics for which he became justly famous. He's a young man, by my standards, and the year 1965 was an extraordinary year for him - probably the opening guns for a still more extraordinary career following - but I note in this biography given me that he became an Honorary Doctor of Science in the University of Michigan , Yale University, and the University of Chicago - all in one year and received a National Medal of Science from the President of the United States in that very same year. In addition to becoming a fellow of the American Academy of Arts & Sciences. What 1966 will bring for him I don't know - it has probably brought a great deal already but it's not on the record which I have. I happened to be present when the discussions were taking place for the nomination for this distinguished National Medal of Science and all agreed that man had very great talent indeed and I join with you in looking forward to his talk which would be as I said on the Translation of the Genetic Code. Dr. Nirenberg:

Thank you Dr. Rabi. Having been born in Brooklyn I also want to thank you for having planted some trees in Brooklyn. I am glad really to have this opportunity to tell you about the genetic language for perhaps 2 to billion years continuous dialogue between cells and their descendants has taken place.

Each generation passes to the next a library of information which specifies in detail the way to make many kinds of protein catalysts that the cells will need for their development. It has seems clear that all, or almost all forms

the Epole

of life on this planet at the use virtally the same language although recently several - a number of dialects have been found that I ll describe later.

May we have the first slide please - This -Very good - The last slide was a good introduction anyway to introduce some of the collaboratores who have worked on this problem in the last - oh 3 years I guess - 4 years This in a very real sense has been a collaborative project and I would like to call attention to these - to the collaborators. However, this is the last slide and I think if we could have the first slide (laughter) we will show it again XXXXX at the end. Good - this is an illustration from Watson's recent very fine book which diagramatically illustrates a double strand of DNA and an enzyme RNA preliminaries (?) which canalyzes the synthesis of message RNA as shown here. The point that I wish to make here is that only one strand of the DNA is copied by RNA prelimnates - is copies with a given polarity not known really what signals or what words specify the beginning and the end of the message RNA synthesis But there must be a code word of some sort which says Start and another one that says Stop. These remain to be defined. Next slide please

This shows also diagramatically a protein synthesis DNA shown here and the different cross hatches represent various segments of DNA corresponding to - each corresponding to a protein possibly or several proteins - messenger RNA ribozomes diagrammatically attached to the messenger RNA and reading begins and as soon as one moves down another one attaches until the message is virtually covered with ribozomes.

The reading takes place by means of SRNA, transfer RNA which carries specific amino acids and recognize particular code words on the ribozomes. Thus, the code word is recognized not by the amino acid per se, but by an adaptor molecule , adapter SRNA. Now, the rationale really for our previous work on the code was to use synthetic messages - compared with polynucleic type sporelegs? they were randomly ordered sequences messages, and in this way characteristics of the code could be determined. And also base compositions of code words. But not base sequences. The situation up to one - well, two - years ago was very much like that of an anagram. We knew the letters of the code , but not the order, of the letters within each word. Now, an approach towards determining the sequences of the code words, are shown on the next slide. This illustrates diagrammatically again the code on recognition process. This is a ribozome of E coli small sub-unit and large sub-unit . There are two binding sites for SRNA at least per ribozime, shown diagrammatically here, and this shows an amino SRNA presumably three bases in the transfer RNA molecule recognized and three bases in the message RNA which binds to the small sub-unit and amino acid as shown Now one of the binding sites presumably is for the peptftal SRNA which is shown here. The other for the incoming amino S-less RNA. Three enzymes plus GTP are required for the transfer, then, of the growing peptide chain to the next amino acid. A shift in some way occurs so that the next three bases would be read. Thus protein synthesis starts at a given place and proceeds, reads three bases sequentially and reads

with a given polarity. Next slide please ..

Well, actually in several laboratories, Schweets laboratory, Lippman's laboratory, Rockefeller, KG's , it is shown that if one added a synthetic message RNA, such as polyuradylic acid , to ribozomes, that phenalalamino SRNA bound to the ribozomes prior to peptide bonds synthesis, my colleague Dr. Philip Lieder, and I wondered how small a message would direct the binding of the SRNA to the ribozome and it turned out that only three bases were needed. That is, the triplet itself would direct the binding of the appropriate aminoSRNA to ribozomes . This provided a rather simple route towards determing the sequence of RNA code words. Now our major problem was to devise some techniques for synthesising triplets. At the time we started the work with the triplets a - oh perhaps 20 to 25 of the 64 possible permutations of bases were reported, had been either prepared by enzymatic digestion of RNA or by chemical synthesis, by some of the very elegant techniques devised by Karana and his associates . Two techniques were developed in our laboratory the first by Philip Lieder, Maxine Singer and Richard Rimakome employed polynucleic type phosphoralays which requires a doublet primer illustrated here, and adds nucleic psydiphosphates to this to make triplets, tetramers, pentamers and soforth. second method developed by Merton Bernfield in our laboratory which employed an observation that had been made by Heppel some years before. The pancreatic RNA-ase which is normally thought to be degradative enzyme will catalyze a transister ification from a two-prime, three-prime cyclic nuclear dite

to a five-prime hydroxol to form triplets and higher homolons. Thus, this technique added units to the left end of an accepter and the polynucleic type phosphoral aids added units to the right end of an accepter. With these two techniques it became rather simple to compare various triplets needed.

The next slide please ..

Now this just shows diagrammatically an oligonucleitide and I wish to illustrate only that ... a given sequence, say, triade, may exist in three chemically distinct forms, depending on its geography in a molecule, that is, a five-prime terminal code-on, internal code-on, and three-prime terminal code-on. All of our evidence todate indicates then that the biologic characteristics of the code-on recognition may in some cases .. many cases .. be influenced by the particular position of the code-on. Thus, each of the 64 possible triplets may exist in three different structural forms. Next slide please

This shows the influence of various substitutions on the hydroxyl groups - terminal hydroxyls of the sugar. This is a concentration of oligonucleitide plotted against the binding **xxx**xx*** aphenylalaminecern into ribosols.

A simple triplet of tri-U has an activity such as this fiveif one adds a phosphate to the phiprime hydroxylake the
sugar the template activity is greatly enhanced, Shown here.

A phosphate attached NEXE to the three-prime terminus lowers Rotman template activity., Recently Fritz Roten in our laboratory prepared some amalogs of tri-U also MEKRY with a metal group attached to the five-prime phosphate, shown here, and

this is, although this should be CH₃ it's a metal group also attached to the three-prime phosphate. Three-prime terminal phosphate. And this greatly reduces template activity. A triplet with a two-prime three-prime cyclic phosphate has very little template activity. Next slide please ...

We think it possible that terminal hydroxyls of sugars may be modified and that in some cases, possibly these may regulate the template activity of the code-on. certainly a substitution at the five-prime terminus may be important because this may serve as a signal which specifies the attachment and/or detachment of the ribozome to the message. Recently, Matra and Herwits , also Stent, have shown in vitro, at least, that messenger RNA contains a triphosphate attached to the five-prime terminal hydroxl. And, although it's not clear what the function physiological funcation of the triphosphate is, it's highly possible that this may in some way specify the phasing, the initiation of reading of the message . It would also select the first word to be read, phase the reading and perhaps affect the susceptibility to EXMUNERAREMEN exonucleasis. Internal codeons may be also important , may be modified in various ways two-prime hydroxyl, or the base could be modified ... I think that code-on neighbors may in some cases play a role also in ... may affect the code-on recognition process.

I would also like to call attention to the possible difference between internal initiation and termination vs terminal initiation and termination. The mes . . . often messages ... messenger RNA seems to contain the information from more than one purtsing ? if one starts reading

at/left end for the terminal initiation one reads in and then comes to a word that says Stop and an unknown mechansim for starting an internal position. And .. that so that internal initiation and termination may, it's not known, but it's possible that it may be different from a terminal initiation or termination mechanism. Next slide please ...

On this slide is plotted oliganucleitide concentration and binding of aminoacelesito ribozomes. I wimply want to illustrate several points that became obvious very early in thiswork. First, that synonym code-ons corresponded to various aminoaceles RNAs , for example, phenylalemine SRNA responded to both to UUU and UUC. And UUC was slightly more active than UUU. With light CSRNA responded to AAA and AAG there was a quite marked difference in template activity of each synonym. The degeneracy that was observed was a very logical degeneracy because the C=U in the third position and in this case, A=G in the third position. Next slide please...

On this slide is summarized virtually all of our work to date. We have synthesized 63 of the 64 possible triplet sequences and the sequences of RNA code-ons are shown. The underlining represents previous base composition assignements which were done with the protein synthesizing system and randomly ordered pinucleitides. And we've observed a very close correspondence between the earlier work with randomly ordered polynucleitides and base sequence studies. I should point out the types of synonymi that are observed. For example, glutannic acid corresponds to code-on ... GAA and GAG. Aspartic acid .. GAU GAC UC AG type of degeneracy. Another type of

degeneracy is illustrated in this column where the third position may be occupied by either UC A or G, completely variable. Another type of degeneracy is illustrated by methianate, that is a G may be present in the third position, may be recognized but not in A. I think that the logical degeneracy that has been found has its many consequences, one of the more obvious is that one may have a great deal of silent mutation, that is in one of the , one of the code words or groups of synonym code .. code-ons one may have mutation and ... which may convert any base in the third position to any other base without resulting in amino acid replacement in per Another obvious conclusion is that amino acids which are very similar chemically such as aspartic acid, and glutanic acid , both dicarboxcylic acids, have chemically similar, quite related code-on. It may reflect the evolution of the code , it's not clear, but certainly one consequence would be that when error does occur, usually two bases are read correctly, one incorrectly. And very often a chemicallyrelated amino acid xx substituted in protein when an error does occur. Thus the general picture of the code is quite a conservative one .. one that minimizes error most often.

There's one underlying question though which has not yet been answered, and that is, why ... why does one see this type of code ... why shouldn't phenylalimine, for example, be corresponding to GCU and GCC instead of alimine. This question has not been answered and I think that there may

well be a chemical meaning to this. It is not obvious yet

Alternatively, it ... simple chance alone may account for this.

I tend to think, however, this is ... is a personal preference

... that there will be an underlying meaning will be found.

Next slide please....

Could we have that last slide ... just a second ... I wanted to point out the ... previous slide ... I wanted to point out that some code words may ... appear to serve special functions. For example, the recent work of Brenner, Garen, Zinder and others indicate that UAA and UAG may specify the end of the message, although the precise mechanism for punctuation is not known. and CUG AUG and some cases GUG IIIG may correspond to the initiation .. may specify the start Our recent studies, and those also of CLark and Marker in England , have indicated that these code-ons in terminal positions at least, are recognized by enformulesionate and this may serve as an initiator of protein synthesis. I'll say more of special function words and .. in a few minutes. Let me have the next slide please ...

This simply summarizes the various patterns of synonym code-ons that are observed. You would see in the third position , and so forth. The enformulathiamine SRNA the initiator Next slide please ...

Some possible special function code-ons are listed on this slide. Sanger in England first observed the necoli that one of the two species, or at least a part of methianine SRNA could accept a formula moiety.

The amino group of the methianine after the methianine was linked to the SRNA could be formulated. And the work of Kapetchy and his colleagues in Watson's laboratories and others have suggested, Zinder's Lab also, have suggested that this may specify initiation. And as I indicated before, U UG AUG CUG and to some extent also GUG are recognized by the informulinine SRNA. UAA UAG may serve as a terminator, we think it likely also that AGU AGC AGA also AGG may serve as special function words. The functions, however, have not thus far been found. This is, I think, the situation now is extremely interesting because the tools are to hand to try to decipher the mechanism of special function words and also the various functions that they may play in protein synthesis.

Next slide please ... Dr. Hatfield in our laboratory has recently prepared some radioactive triplets and have looked at the binding of the triplet to the ribozome. In this box is plotted the binding of the triplet to the ribozome; in this the binding phenylalimine SRNA to the ribozome. As you see in the presence of the appropriate triplet phenylalimine SRNA binds to ribozomes times plotted against binding. With the radioactive triplet Dr. Hatfield asked the question "Is SRNA necessary for the binding of the code word to the ribozome?" And he finds that it is . Because very little triplet binds in the absence of SRNA, in the presence the triplet binds at approximately the same rate as phenylalimine SRNA and approximately the same amount of triplet binds as phenylalimine SRNA. Thus, the complex may well be a one-to-one complex on the ribozome.

I should also point out that major lucine SRNA binds to ribozome very weakly in response to triplets, and possibly this type of recognition corresponds to only recognition of two out of the three phases. Next slide please ...

This shows an experiment with some of Holly's SRNA plotting concentration of alimine SRNA against binding to ribozomes. And the dotted line represents hundred per cent binding, that is, all of the available SRNA in the reaction. This fraction of SRNA, which Holly sent to us, was estimated by their laboratory to be greater than 95% pure. And yet this SRNA fraction recognized quite well three, at least three of the alimine code-ons ... GCU GCC and GCA. It did not respond, or responded to a relatively slight extent to GCG. E-coli SRNA , unfractionated SRNA, responded quite well to GCG , as a matter of fact, this was the best code-on ... alimine code-on found as compared to the coli, rather the yeast, very poor pattern . The GCA GCU and GCC very little response to GCC ... GCC was an excellent code-on for alimine with yeast SRNA. Now since this SRNA was of high purity the results strongly suggest that one molecule of SRNA can recognize alternately at least three of the four alimine synonyms. Further, since Holly's laboratory has recently reported the sequence of alimine SRNA the data afford some insight into the mechanism of code-on recognition. Next slide please ...

This is the sequence of alimine SRNA as reported by Holly's laboratory. The amino acid is linked to the terminal denizine and this is shown in only one of the possible conformations that has been suggested by Holly's laboratory. I should point

This provides very simple and quite sensitive technique for detecting code-on recognition by SRNA which is not isolated with amino acids. If Some special function words are not recognized by activating enzymes, not asolated, this would provide a relatively simple route towards detecting such recognition.

Next slide please ... On this slide is summarized the work that has gone on in our laboratory with purified SRNA fractions. This work was done in collaboration with B. P. Docter at Walter Reed and also Donald Kellogg in our laboratory. Tyrazine SRNA both peaks of SRNA recognize both UAC and UAU and this shows the CU pattern in the third position. A minor peak availing recognizes GUC and GUU . The major vailing peak, however, recognizes the GU A GUG and to a smaller extent GUG. All of this work, with one exception, was done with E-coli SRNA fractions. The E-coli fractions, licine 1 and 2 recognize both licine code-ons. Recently carbon, however, has is reported, just at the Federation meetings a few weeks ago, that in mammalian liver one species of licine SRNA preferentially recognizes AAG. The other preferentially recognizes AAA. Lucine peak lA C recognizes CUG , another peak UUG. The small methianine peak SRNA peak recognizes AUG preferentially, whereas, as I said before, the major peak of methianine SRNA will accept a formile group and can recognize U UGC UG and Triptathane SRNA peak 2 recognizes UGG CGG and to a smaller MEMKERX extent AGG. Thus, one sees this pattern in the first position and Yeast alimine SRNA recognizes GCU GCC and GCA. The pattern is repeated in the third position. Now this experiment was done with ... this SRNA was obtained from Dr. Robert Holly with a highly purified fractions of yeast SRNA and I'll say more about this later.

that there are several possible single-stranded regions of interest. This sequence is GTpseudoUC has been found in virtually every SRNA that has been examined. Another interesting sequence is the CGGs surrounded by two dihydrauriditic acids. And the third is the IGC region which is right in the middle of a molecule. Next slide please ...

The two sequences of interest are those shown here.

The CGG between dihydroU's and the IGC . Now if phase-pairing ... if these really are the SRNA anti-code-ons, that recognize the alimine code-ons that are shown here, recognition would have to be by parallel pairing between C and G , and the G would then have to recognize U C and A. If, however, phase pairing were Watts and Krick hydrogen-bonding anti-parallel phase pairing C would phase pair with the G and G with the C and the inner scene in this position would base pair with alternately with U C or A but not G, which is the pattern which is observed. Krick has recently proposed a detailed mechanism for which would permit hydrogen bonding between intercine and U C or A. Next slide please ...

And this mechanism simply employs a "wobble" """ A MOVEment of either the SRNA or the messenger RNA on the ribozome
in the end position. And I think all of the experimental
results are in accord with this type of recognition mechanism
as shown here. The basis in the SRNA anti-code-on is shown
in this column, the basis in the messenger code-on is shown
in this column. The patterns that are found are listed here.
Thus in innercine in SRNA in an end position could recognize

by alternate base pairing U C or A a G, an end position of SRNA could recognize alternately C or U and A could recognize U C G and a U could recognize alternate pairing A or G . We would also predict that a ribothymadilic acid SRNA would pair also with A or G , perhaps the interaction with A would be stronger than a urydilic acid residue , that a pseudoU in SRNA might recognize alternately A G or U and we have noticed this pattern with SRNA rather often. And possibly/that a dihydrouridylic acid would not base pair so that the interaction with the triplet .. messenger triplet would be weak interaction but it's entirely possible that a U or a C in a terminal position would not greatly inhibit the interaction so this may be permissible. Also, that a metal group on a two-prime hydroxyl deribose might result in a weaker interaction and by permitting greater degree of motion on the ribozome might permit a greater ambiguity. ... a lower specificity.

Next slide please

Now the results with the innercine strongly .. and also the other trace spaces ... rather strongly suggests that .. actually a great deal of work suggests that SRNA is modified enzymatically after it is released from DNA templates. And it seems probable .. that there ... in the cell there exists a whole spectrum of intermediates. One gets a successive modification after modification because the level of trace spaces in SRNA is quite high. Now the consequences of this are rather simple to visualize. If an A, an admine, for example in SRNA is de-aminated and converted to an innercine the A would normally recognize the uradylic acid residue in the message and the conversion would result in a recognition of

a U C or A. Similarly, if a G were converted to an I when we get this type of inter-conversion, or a C were de-aminated we would get this type of conversion. The ... it;s too early really to say ... I think it likely that this type of inter-conversion occurs and that it has very possibly great biologic meaning. Next slide please ...

There really a great deal of work has been done recently which suggests that ... really a great deal of work has been done ... which shows that in many cases that it is possible to modify the specificity of code-on recognition. I Think that this is biologically a very very important point and will have truly profound biologic consequences. Gorini and Gilbert and Davies at Harvard have shown that streptomycin will bind to a protein on the 30 s ribozome the small sub-unit and all the available evidence $sugge\mathbf{x}ts$ that the binding of streptomycin to ribozome may in some way distort the geography of the code-on recognition site, to that a greater ambiguity, a lower specificity may occur and this may be at least one reason - it may not be the only reason - one reason which may account for the action of streptomycin on bacteria. Simply the greater degree of error occurs in protein synthesis. There may be other mechanisms, other effects of streptomycin on the cells as well.

Id like to actually list just a few ..pf the recent ... relatively recent examples of modification of specificity of code-on recognition that are being studied in various laboratories. Streptomycin effect is certainly one, Next slide please ...

Well, before I come to this slide ... I should mention also the ... some work that ... some very recent work that we've been doing in collaboration with the Siwilkets at Princeton . They observed that upon infection of bacterial cells of C-coli with a virus T2-phage that within one minute after infection a protein was synthesized by the bacteria which modified kke a pre-existing lucine SRNA component . This modification resulted in a shifted movement of the SRNA on mack columns so it could be purified. We have tested the code=on recognition of this modified ... this SRNA modification which is due to viral infection. And find that it recognizes only poly UG but it doesn't recognize any triplet. And we have tested all of the UG triplets . Now together with the modification of the SRNA one finds the cessation of post-protein synthesis and we don't understand the mechanism of this ... mechanism of turning off the host protein synthesis but we think it likely that .. that the enzyme modifies the lucine SRNA component and this lucine SRNA may in some way interfere with the initiation of host protein synthesis, but not a phage synthesis. There's a very subtle way of subverting the metabolism of the cell so that viral proteins may be synthesized in large amount. And currently this problem is being worked on .

And here also is some recent study that has been done in our laboratory by Richard Marshall and Tom Kaske which we have compared the specificity of code-on recognition in amphibian .. with amphibian SRNA .. zenith eslavor, liver and guinea pig liver SRNA with that of E-coli SRNA . E-coli

SRNA does not recognize AGG and recognizes CGG with biargine SRNA only to a very slight extent. Whereas in both amphibian and mammalian liver we see a very strong response of arginine SRNA to AGG also to CGG. The alimine SRNA I mentioned previously with the yeast .. contrast the yeast and coli patterns shown here. In both amphibian liver and guinea pig liver GCC is a very active code-on whereas in the amphibian liver GCG is a very ... has no activity for alimine SRNA as contrasted to coli-SRNA. AUA/recognized by isolucine SRNA in higher forms although we have not detected it in coli. all species tested AAA is recognized, whereas AAG has only slight activity in coli but is a very active code-on in higher forms. Seriun SRNA UCG ix recognition is variable, AGU and AGC recognition is variable also as shown here. recognition ACG is a variable as shown there. Now we have found no differences in the code-on recognition of SRNAs corresponding to aspartic acid, cystine, glutannic acid, histamine, phenylalimine, proline, tyrazine and also valine which should be on this slide.

Thus it seems clear that the same sequences correspond to quibalominic acids throughout .. from bacteria to mammalian liver. However, very subtle differences in relative responses of synonym code-ons are observed. I think that such differences may, in some cases, play biologic roles. We wonder whether such differences may be involved in some types of differentiation ... may affect the rate selectively of protein synthesis in some cases. These are problems I think for the future that will be worked on within the next year or so and .. we are very interested in determining the outcome of questions such

as these. Thank you very much.

Applause

Prof.Devons:

Well as Dr. Rabi said "it's not befitting for a physicist fo make comments of appreciation of these things." All I can say is that Dr. Nirenberg has indicated some of the rather breathtaking speed with which this subject is developing. I am going to ask Professor Zubay, professor zoology here to say a few more professional words of appreciation than I can offer. Prof. Zubay:

Prof.Zubay -

I was hoping Dr. Nirenberg would talk for another 5 minutes because I was working on some concluding remarks - However, let me say that I found that this lecture was somewhat more technical perhaps than most of us have been hearing on these Monday nights . I am sure that many of you were snowed by it - as a molecular biologist I assure you everything he said was correct (laughter) or at least 90% correct. I found it very exciting and I am sure you noticed how totally immersed and how exciting he finds his own work. This great crusade of his began in .. around 1961 when he made the very exciting announcement in Mostow - I was there at the International Biochemistry Congress - as to how polyurydelic acids stimulated the incorporation of phenylalimine into a polypeptide chain. Francis Krick at the time predicted the entire code would be solved in one year - it has taken a few years longer than that and the pace, in spite of Krick's prediction , has been miraculously fast. I think it fair to say that Marshall Nirenberg has carried the ball all the way. As a member of the newly formed Department of Biological Sciences and as a fellow molecular biologist - I might add, fortunately not a competitor of Dr. Nirenberg's - let me say that I am deeply honored to express our collective appreciation for his coming tonight and for his fine presentation. (Applause)

Prof.DevonsJust a couple of brief announcements. For those who would like to hear more about this subject, as you will see on the sheets that have been distributed, Dr. Nirenberg is giving a talk tomorrow at 10 o'clock in the Watson Laboratories. I also draw your attention that the date of our next meeting is Monday, May the Ninth, when Dr. Kendrew from Cambridge University, will be here to talk about physics, molecular structure and biological function. That's on Monday, May the ninth at the usual time. Thank you.