CHAPTER 6

SUMMARY OF RESULTS
6.1 FINDINGS OF THE STUDY

6.2 TRANSLATION-INITIATION PROMOTING (TP) SITE ON mRNAs OF HIGHLY EXPRESSED GENES FROM PROKARYOTES AND LOWER EUKARYOTES

6.2.1 TP Site on Prokaryotic mRNAs

6.2.2 Extended TP Site on Lower Eukaryotic mRNAs

6.2.3 Functional Conservation of TP-like Site in Various Kingdoms

6.2.4 TP Base Pairing is only Transient and Released during the Subunit Association in Translation-initiation

6.3 CONTACT SITES ON rRNAs FOR SUBUNIT ASSOCIATION IN THE TRANSLATION-INITIATION PATHWAY

6.3.1 Involvement of Contact Sites in Building the 'Translation Domain'

6.3.2 Criteria for Phylogenetic Preservation of Inter-RNA Base Pairs

6.3.3 Conformational Switches Involving the Contact Sites

6.4 CONCLUDING REMARKS
6.1 FINDINGS OF THE STUDY

The work reported in this thesis demonstrates that sequence data analysis through functional approach, combined with inputs from biological features, can serve as a valuable tool to identify the characteristic sequence structures responsible for specific biological functions. The work has revealed the following functional feature sites on the RNA components of the translation process.

(a) A cis-acting element, needed for efficient translation-initiation, on mRNAs of highly expressed genes from prokaryotes and lower eukaryotes, and other related sites on rRNAs that help in carrying out the mechanism of modulating the efficiency of the translation process.

(b) A set of phylogenetically preserved inter-RNA contact sites between the small and the large subunit rRNAs that help not only in the association of subunits but also in building the 'translation domain' on the assembled ribosome.

The results of the study are summarised in the following sections.

6.2 TRANSLATION-INITIATION PROMOTING (TP) SITE ON mRNAs OF HIGHLY EXPRESSED GENES FROM PROKARYOTES AND LOWER EUKARYOTES

6.2.1 TP Site on Prokaryotic mRNAs

Firstly, it was shown that the 5' noncoding region of mRNAs of highly expressed genes from Escherichia coli contain a
sequence 5'-UGAUCC-3' (TP site) located upstream of Shine-Dalgarno sequence. Such a site on mRNAs has the potential to form base pairs with an anti-TP site (5'-GGAUCA-3') located 5' in continuation of the anti-(Shine-Dalgarno) sequence at the 3' terminal end of 16S rRNA. The analysis pointed out that both the TP and the anti-TP sites are free from any intramolecular RNA structure; and the anti-TP site is evolutionarily conserved. Such a TP site is proposed to function in the following ways in E. coli.

(a) The TP-(anti-TP) base pairing interaction contributes an additional free energy $\Delta G$ of -5.1 Kcal/mol towards the ribosome binding strength. As a result, it enhances the stability of the 30S initiation complex (involving highly expressed mRNAs) thereby leading to a higher rate of formation of 70S initiation complex.

(b) Along with the Shine-Dalgarno site, the TP site ensures a higher proportion of ribosomes sequestered on the 5' end of mRNA. As a result, the initiation occurs relatively more number of times on mRNAs containing the TP site than on those mRNAs having only Shine-Dalgarno sequence.

The TP site thus acts in concert with the SD site in controlling the differential rate of initiation. The observation, that the TP sites occur only on highly expressed genes, was found to be true in other prokaryotes also. Such an observation indicates that the mechanism of modulation of the level of gene expression through TP site base pairing is not restricted only to E. coli but is common to all prokaryotes. The analysis further pointed out that the region on mRNA that determines the efficien-
cy of translation-initiation is not confined only to the ribosome binding site (which encompasses the region -21 to +14 on mRNAs) but spans a larger region that extends at least up to -45 nucleotides 5' to the start codon.

6.2.2 Extended TP Site on Lower Eukaryotic mRNAs

Secondly, it was shown that in yeast, the 5' noncoding regions on highly expressed genes contain an extended TP site of the form 5'-UGAUCCACC-3' which has the potential to form base pairs with an extended anti-TP site, namely 5'-GGUGGAUCA-3', found on the 3' end of 18S rRNA. Both the extended TP site and the extended anti-TP site are formed by noncontiguous primary sequences brought together by weak intramolecular hairpin stems. The spacing between the extended TP site and the start codon on highly expressed mRNAs is found to be <6 nucleotides. The average free energy contributed by the extended TP site towards the association of mRNA to the ribosome is found to be -7 Kcal/mol. As a result, the extended TP site base pairing can facilitate the formation of a comparatively more stable 40S initiation complex, leading to a higher rate of formation of 80S initiation complex; this, in effect, could enhance the rate of translation initiation. It was noted that, in higher eukaryotes, the optimal 5' AUG-context sequence (5'-CC(A/G)CC-3') is a part of the extended TP site. Recently, it has been reported that there exists a prokaryotic-like cis element (5'-UUUCC-3') on picornaviral mRNAs (wherein the initiation is cap-independent); this element serves as the internal binding site for the ribosome on mRNAs. The region on 18S rRNA, with which the above cis-acting element is
proposed to base pair, overlaps with the extended anti-TP site.

6.2.3 Functional Conservation of TP-like Site in Various Kingdoms

It is interesting to note that the concept of the modulation of the rate of translation through the TP site appears to be evolutionarily conserved as indicated below.

Thus, the TP-like site brings about an overall similarity in the translation-initiation process among the kingdoms. Such a functional preservation of TP-like site might argue in favour of the validity of the proposed TP site.

6.2.4 TP Base Pairing is only Transient and Released during the Subunit Association in Translation-initiation

Based on the observations that (a) the RBS region on mRNA that is protected by the 70S(80S) initiation complex is smaller than the region protected by the 30S(40S) initiation complex, and
(b) the TP-like site occurs 5' to the RBS, it was proposed that the TP site base pairing is only transient and disrupted during the association of the 50S(60S) subunit with the 30S(40S) initiation complex to form the 70S(80S) initiation complex (brackets indicate the eukaryotic source). It is proposed that the release of the TP site base pairing may be brought about by a mechanism in which the anti-TP site on the 16S(18S) rRNA disrupts its base pairing with the TP site on mRNA and forms base pairing with a complementary site, called anti-(anti-TP) site, on the 23S(28S) rRNA during the association of the 50S(60S) subunit with the 30S(40S) initiation complex. The analysis pointed out that both the anti-TP and anti-(anti-TP) sites on 16S(18S) and 23S(28S) rRNAs are single-stranded in their respective subunits, and are well conserved in sequences from other species, supporting their involvement in the above mechanism. Such a mechanism enables the next available free 30S(40S) subunit to bind to mRNA, through the released TP site, even before the active ribosome finishes translating the initiation region. As a consequence, the size of the polysome would increase.

### 6.3 CONTACT SITES ON rRNAs FOR SUBUNIT ASSOCIATION IN THE TRANSLATION-INITIATION PATHWAY

The approach of comparative sequence data analysis was used to work out the possible sites of inter-rRNA interactions that help to bring about the subunit association. The analysis pointed out the possibility of the existence of eight sites each on 16S and 23S rRNAs. The base pairs of the inter-rRNA interactions were found to be well preserved in the phylogenetic spec-
trum of 39 species with a mismatch base pair change value of only 5.8%. The available experimental data on the functional status of the bases of rRNAs, existing in the literature, substantiated the choice of such contact sites and also revealed their functional significance.

6.3.1 Involvement of Contact Sites in Building the 'Translation Domain'

Examination of the functional status of the constituent bases of the contact sites indicated the possibility of the involvement of such inter-rRNA contact sites in building the translation domain. The regions on 16S rRNA, from which these 8 contact sites arise, were found to be part of the decoding site. The regions on 23S rRNA, from which the contact sites arise, were found to be part of the peptidyl transferase center. It is known that the anti-codon arm of tRNA is associated with the decoding site on 30S subunit and the acceptor arm of the tRNA is associated with the peptidyl transferase center on the 50S subunit. Thus the contact sites appear to be responsible for bringing the decoding site (on the 30S subunit) and the peptidyl transferase center (on the 50S subunit) into a topological proximity to build the translation domain.

6.3.2 Criteria for phylogenetic preservation of inter-RNA base pairs

The analysis brought forth a rethinking on the criteria for the phylogenetic preservation of inter-RNA base pairs. Some of
the differences between intra-RNA and inter-RNA interactions are worth noting, as follows.

(a) It has been sufficiently argued by earlier workers that the 'RNA world' might have entirely depended on RNA-RNA complexes for carrying out biological functions at the molecular level. If this was so, the inter-RNA base pairing interactions might have evolved earlier than the intra-RNA stems. Further evolution might have been centered around the proteins and their co-evolution with the intra-RNA stems to which they are known to bind leading to stabilised structures.

(b) Inter-rRNA interactions play a role in fulfilling the functional requirements, while intra-rRNA interactions are largely for structural requirements.

(c) As is the case with other functional sites, the sites of inter-rRNA interactions arise from single-stranded regions. The single-stranded regions on rRNAs are more conserved than the double-stranded regions. The average percentage of conservation of bases in single-stranded regions involved in inter-rRNA interactions was found to be 73% whereas that of bases in intra-rRNA interactions (stem regions) was found to be 50%. As a result, inter-rRNA base pairs would be expected to exhibit more conservation of bases than compensating base pair changes. In contrast, the intra-rRNA base pairs would exhibit more compensatory base pair changes.

As a consequence of the above mentioned differences, the criteria that one can think of for the phylogenetic preservation of inter-rRNA base pair should be based on the percentage value
of mismatch base pair occurrence; and this mismatch percentage should be as low as possible. The contact sites proposed in this work exhibit an average mismatch percentage of 5.8%, which is significantly lower than the average mismatch percentage (10%) exhibited by intra-rRNA stems (23S rRNA in this case). In addition, one can not ignore the fact that 5 out of 8 contact sites exhibited atleast one compensating base pair change.

6.3.3 Conformational Switches Involving the Contact Sites

Out of the proposed 8 contact sites, two sites share a common region on 16S rRNA (790-792 and 790-793) but occupy different regions on 23S rRNA (2754-2756 and 2478-2481). The region on 16S rRNA for the sites involves bases in the P-site on 30S subunit. One of the corresponding two sites on 23S rRNA is located at the peptidyl transferase center and the other slightly farther from it. Therefore, it is proposed that these two sites together may form a conformational switch in a way that the site on 16S rRNA may alternate its base pairing with the two sites on 23S rRNA. Based on Noller's proposal that tRNA translocation may involve some sort of relative movement between 30S and 50S subunits, it is quite possible that the above mentioned alternate sites, along with other unknown contact sites may be involved in tRNA translocation.

6.4 CONCLUDING REMARKS

The study has illustrated that RNA-RNA interaction plays a central role in executing biological functions, such as, translation; the theoretical approach of sequence data analysis is a
valuable tool to study the biological functions at molecular level. It further illustrates that through such an approach, one can make hypothesis that is consistent with known experimental observations and fundamental principle; and one can make predictions based on the hypothesis that are testable experimentally. It is heartening to mention here that the proposed TP site, since it was published in 1989, is being widely referred in the literature by the stalwarts in ribosomology (see Appendix); and experiments are being tried by other laboratories to test validity of the proposal. This is in tune with what Huszagh and Infante recently wrote in their article titled, 'The hypothetical way of progress',

"A theorist’s strength, and enjoyment, is in proposing testable ideas, where as that of an experimentalist is in devising test of such ideas"